Friedrich-Alexander-Universität Erlangen
Mathematisches Institut
Lehrstuhl für Mathematische Stochastik
Bismarckstraße $1\frac{1}{2}$
91054 Erlangen

Asymptotic Genealogy of a Catalytic Branching Process with a General Catalyst

---

Diplomarbeit
von Thomas Rippl

betreut durch
Prof. Dr. A. Greven

vorgelegt im Mai 2009.

Abstract

We consider two random populations and want to study the asymptotic behavior of their genealogies, when the number of initital individual increases. Both populations are continuous-time Galton-Watson branching processes. The first population, called catalyst, consists of individuals which do not branch independently. The second population, called reactant, consists of individuals which branch independently of the others but the branching rate depends on the number of catalyst individuals alive.

For the description of the genealogy we use total mass processes, $\mathbb{R}$-tree-valued processes, contour processes and point-processes of MRCAs (most recent common ancestors). We present weak limit results for any of these descriptions. Moreover we discuss some results which are due to the catalytic structure, especially there, where we are close to the extinction time of the catalyst.

The ideas and techniques used include $\mathbb{R}$-trees with Gromov-Hausdorff metric, excursion theory, Kingman coalescent, theory of one-dimensional diffusions and Stochastic averaging.

Zusammenfassung

Wir betrachten zwei zufällige Populationen und wollen, im Falle steigender Anzahl von Anfangsindividuen, das asymptotische Verhalten ihrer Genealogien beschreiben. Beide Populationen sind zeitstetige Galton-Watson Verzweigungsprozesse. Die erste Population, genannt Katalyst, besteht aus Individuen, die nicht unabhängig verzweigen. Die zweite Population, genannt Reaktant, besteht aus Individuen, die unabhängig verzweigen, jedoch hängt die Verzweigungsrate von der Anzahl der Katalyst-Individuen zu dieser Zeit ab.

Zur Beschreibung dieser Genealogien verwenden wir Totale Massen Prozesse, $\mathbb{R}$-Baum-wertige Prozesse, Konturprozesse und Punktprozesse von MRCAs (letzter gemeinsamer Vorfahr). Wir beweisen schwache Limes Aussagen für alle diese Beschreibungen. Zudem besprechen wir einige Ergebnisse, die durch die katalytische Struktur des Prozesses gegeben sind, besonders dort, wo der Katalyst fast ausgestorben ist.

Unter den verwendeten Techniken und Ideen sind $\mathbb{R}$-Bäume mit Gromov-Hausdorff-Metrik, Exkursionstheorie, Kingman Koaleszent, Theorie eindimensionaler Diffusionen und Stochastic Averaging.

Contents

1 Introduction

1.1 History of branching processes

In the end of the nineteenth century the English Reverend Henry William Watson raised the question of extinction of noble surnames within English society in an article in an English journal. He supposed that due to life conditions the fertility of well-situated families shrank and therefore the names disappeared. One year later he and Francis Galton published jointly an article in the Journal of Anthropology [WG75] and took a mathematical approach to the question: In each generation a father, who passes his surname to his sons, has probabilities $p_0, p_1, p_2, \ldots$ to have $0, 1, 2, \ldots$ sons. His sons, independently of each other, have the same probabilities to have sons themselves as well. Galton and Watson came to the wrong conclusion that almost all surnames would die out. Later it was corrected to the true assertion that the surname almost surely dies out, when the expected number of sons is less or equal to 1.

The model just described was given the name Galton-Watson process and includes the following characteristics, amongst others, to be noted here: It considers a population consisting of individuals that live a certain given time and after that have a random number of offspring. Additionally the process is

(i) time discrete,

(ii) Markovian,

(iii) counts the number of individuals not their relationships and

(iv) sons evolve independently of each other.

Clearly all of these four assumptions have some problems when applying to reality or even do not seem rich enough in description:

ad (i): Time is not measured in generations but in years, so to get an idea about what happens after 200 years, the model does not help.

ad (ii): If one father has already few sons due to genetic endowments, his sons might also suffer lesser fertility. Therefore generations might rely on each other, so the real evolution is probably not Markovian.

ad (iii): It is not possible to answer questions such as "When did the most recent common ancestor of two individuals die?"

ad (iv): Overpopulation decreases the amount of food and probably the number of children, so the independence of all other individuals at the same time is disputable. Also being the brother of ten other brothers makes it more likely that a son will be be sent to the monastery instead of having a family.

During the last century many adjustments were made to extend the original models. Some of the more recent models contain more information than the older model of Galton and Watson. All of these extensions of the basic model are somehow grouped together in the phrase *Branching Processes*.

In recent years some letters written by Bienayme were discovered, who already took up the problem in 1849 and was able to give a correct solution to the problem. More about the history of the (still-called) Galton-Watson processes can be found in [Ken66] and [Ken75] or in a talk by Peter Jagers [Jag09]. The classic books about branching processes were written by Harris [Har61] and Arthreya and Ney [AN72].

The applications of these processes exceed the problem of surname evolution and include phenomena in physics, e.g. cosmic rays, or in biology, e.g. phylogenetic trees. Here the applications will play no role, but for the sake of understanding of the mathematics it is sometimes helpful to keep the "picture" in mind. Notation and nomenclature will be influenced by possible applications as well.

Note that it is possible to consider higher information models and lower information ones. Some keep track of all of the history, others only of the recent state. We will say that the lower information model is a *functional* of the other one. Indeed we will in a moment introduce a rather general model and consider various functionals of it.

1.2 The Catalytic Branching Model with a modified catalyst

In this diploma thesis the focus is made on a special kind of a population evolution model. It will be a *general* model and differs from the aforementioned in several points, e.g. in the fact that it considers two populations instead of one. We will call them the *catalyst* η and the *reactant* ξ and both are stochastic processes evolving in time starting with one individual. Together they form the population model (η, ξ). To get an idea of the process a short description is given before an exact definition:

(i) The *catalyst* η evolves like an autonomous binary branching process. That means that each individual lives until the first jump of a time changed Poisson-process. At this instant of time it dies and after death has zero or two offspring. The sons evolve in the same manner.

(ii) For a given catalyst realization the *reactant* ξ evolves like a binary branching process:

- Each individual lives until the first jump of a time changed Poisson-process. At this instant of time it dies and after death has zero or two offspring. The sons evolve in the same manner.

- This time change depends on the number of catalyst individuals present at the current time. The more catalyst *individuals* are present, the faster the reactant individuals branch. That is clearly why the populations are called catalyst and reactant.

This description contains a lot of information and needs some clearer probabilistic foundation. Hence we will give a more formal definition of a probability space on which the previous description can be realized. The ingredients, "alarm clocks and coins", are already visible and they will become clearer in the exact definitions of the catalyst and reactant to come. For alarms clocks we could use the a collection of exponential waiting times but we will use jump times of Poisson processes instead, which makes the formulation easier.

The names starting with c or r are related to catalyst and reactant, respectively.

Definition 1.2.1 (The basic probability space)**:**
Let $(\Omega, \mathcal{F}, \mathbf{P}) = (\Omega_c \times \Omega_r, \mathcal{F}_c \otimes \mathcal{F}_r, \mathbf{P}_c \otimes \mathbf{P}_r)$ *be a probability space, which contains*

- *a sequence* $N_1^c, N_2^c, \ldots$ *of Poisson-1-processes,*
- *a sequence* $N_1^r, N_2^r, \ldots$ *of Poisson-1-processes,*
- *a sequence* $C_1, C_2, \ldots$ *of* $0, 2$ *coin-tossing random variables, i.e.*
$$\mathbf{P}(C_1 = 0) = \mathbf{P}(C_1 = 2) = 1/2 \text{ and}$$
- *a sequence* $R_1, R_2, \ldots$ *of* $0, 2$ *coin-tossing random variables, i.e.*
$$\mathbf{P}(R_1 = 0) = \mathbf{P}(R_1 = 2) = 1/2,$$

all of which are independent.

When looking at the formal definitions to come it can be helpful to look at Figure 1.1. There both populations are given in a planar embedding.

We now describe the catalyst $(\eta_t)_{t\geq 0}$, where the total number of catalyst individuals at a time t will be called η_t^{tot}. Let a function $g : [0, \infty) \to [0, \infty)$ be given, which shall be fixed throughout the paper. The catalyst starts with one individual ($\eta_0^{\text{tot}} = 1$) and evolves as a Markovian process. This first individual faces a branching event at the first jump of $N_1^c(\int_0^{\cdot} g(\eta_s^{\text{tot}})/\eta_s^{\text{tot}} ds) = N_1^c(\int_0^{\cdot} g(1) ds)$, say after time t_1. At this branching event t_1 this first individual has C_1 sons, i.e. either 0 or 2 sons. If $C_1 = 0$, the process is over since no more individuals are alive. In the case $C_1 = 2$ the first son is labelled *son00*, the second one *son01*. Then *son00* branches at

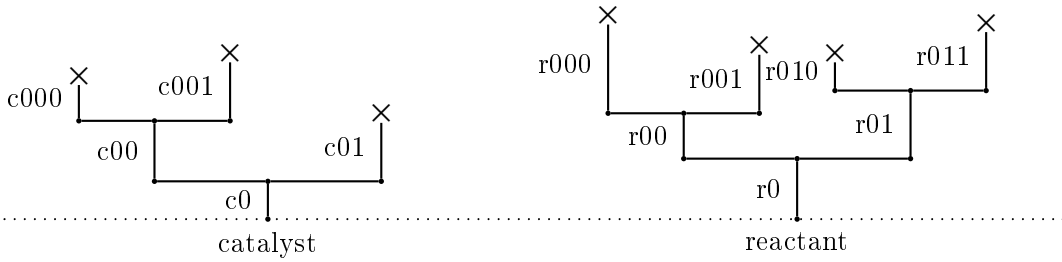
$$\text{the first jump of the process } N_2^c(\int_{t_1}^{\cdot} g(\eta_s^{\text{tot}})/\eta_s^{\text{tot}} ds), \tag{1.1}$$

with offspring C_2. The same for *son01*. Forthcoming sons are given labels by adding 0 or 1 to the label of their father, so the first son is called *son(label of father)0*, the second one *son(label of father)1*. The label does not play any role for their behavior. To be a bit more formal define the following function

$$\pi : \begin{cases} \bigcup_{n\in\mathbb{N}} \{0,1\}^n & \to \mathbb{N} \\ (\omega_1, \ldots, \omega_n) & \mapsto 2^n + \omega_1 2^{n-1} + \cdots + \omega_n 2^0 \end{cases}. \tag{1.2}$$

When, for the sake of clearness, catalyst individuals are called *csons* = catalyst sons, then the definition of the *general catalyst process* η is given by:

Figure 1.1: A planar embedding of the process



Definition 1.2.2 (The general catalyst process)**:**
The catalyst $(\eta_t)_{t\geq 0}$ *is a continuous time process starting with one individual, called cson0, where the suffix 0 is called the label. Each individual alive cson(label), born at time* $t_{\pi(label)}$*, lives until the first jump time of the process*

$$N^c_{\pi(label)}(\int_{t_{\pi(label)}}^{\cdot} g(\eta_s^{\text{tot}})/\eta_s^{\text{tot}}\, ds). \tag{1.3}$$

Then this individual branches, i.e. it has $C_{\pi(label)}$ *csons, either 0 or 2. In the first case individual son(label) is called dead. Otherwise its two csons are called cson(label)0 and cson(label)1.*

Similarly the reactant is given by the following definition. Here the individuals are called *rsons*, b is a branching parameter and each of them branches faster if there are more catalyst individuals "around":

Definition 1.2.3 (The general reactant process)**:**
For a given catalyst realization $(\eta_t)_{t\geq 0}$*, the* reactant $(\xi_t)_{t\geq 0}$ *is a continuous time process starting with one individual, called rson0, where the suffix 0 is called the label. Each individual alive rson(label), born at time* $t'_{\pi(label)}$*, lives until the first jump time of the process*

$$N^r_{\pi(label)}(\int_{t'_{\pi(label)}}^{\cdot} b\eta_s^{\text{tot}}\, ds). \tag{1.4}$$

Then this individual branches, i.e. it has $R_{\pi(label)}$ *rsons, either 0 or 2. In the first case individual rson(label) is called dead. Otherwise its two sons are called rson(label)0 and rson(label)1.*

Be aware that η_t represents the whole catalyst evolution up to time t. By definition the process η is Markovian, since any information needed for the further evolution after time t is available in η_t, namely the number of individuals at time t: η_t^{tot}. The same is true for the reactant ξ, if a catalyst realization is given, and for the pair (η, ξ) as well.

With the means of studying diffusion limits of the catalyst and the reactant, rescaled versions of the previous processes are defined. The number of starting individuals of catalyst and reactant is increased to n, where n is a natural number.

Each individual is given mass $1/n$, so that the initial total mass is $n \times 1/n = 1$ for catalyst and reactant. For the catalyst, time is sped up by a factor n as well. That means the Poisson process runs n times the original speed.

To define the rescaled process more explicitly let the following modified labelling function be given:

$$\pi^n : \begin{cases} \bigcup_{k\in\mathbb{N}}\{1,2,\ldots,n\}\times\{0,1\}^k & \to \{1,2,\ldots,n\}\times\mathbb{N} \\ (\omega_0,\omega_1,\ldots,\omega_k) & \mapsto (\omega_0, 2^k+\omega_1 2^{k-1}+\cdots+\omega_k 2^0) \end{cases}. \tag{1.5}$$

With this labelling at hand it is clear how the formal ingredients of the rescaled process need to look like.

Definition 1.2.4 (The rescaled catalyst process)**:**
The rescaled catalyst $\eta^n = (\eta^n_t)_{t\geq 0}$ *is a continuous time process on the probability space* $(\Omega^n, \mathcal{F}^n, \mathbf{P}^n)$*, which consists of* n *independent copies of the probability space defined in Definition 1.2.1. It starts with* n *individuals called cson1, cson2, ... csonn, where the suffix is called the label. Each individual alive cson(label), born at time* $t_{\pi^n(label)}$*, has mass* $1/n$ *and lives until the first jump time of the process*

$$N^c_{\pi^n(label)}(n\int_{t_{\pi(label)}}^{\cdot} g(\eta^{\mathrm{tot},n}_s)/\eta^{\mathrm{tot},n}_s\, ds). \tag{1.6}$$

Then this individual branches, i.e. it has $C_{\pi^n(label)}$ *csons, either 0 or 2. In the first case individual son(label) is called dead. Otherwise its two csons are called cson(label)0 and cson(label)1.*

For the reactant one needs to remember that the branching behavior depends on the *number* of catalyst individuals alive at a certain time $n\eta^{\mathrm{tot},n}_t$. Therefore the rescaled reactant is given by the following definition, where time needs not to be rescaled since the speeding up will be automatic by the number of catalyst individuals alive:

Definition 1.2.5 (The rescaled reactant process)**:**
The rescaled reactant $\xi^n = (\xi^n_t)_{t\geq 0}$ *is a continuous time process on the probability space* $(\Omega^n, \mathcal{F}^n, \mathbf{P}^n)$*, which consists of* n *independent copies of the probability space defined in Definition 1.2.1. It starts with* n *individuals, called rson0, rson1, ... rsonn, where the suffix is called the label. Each individual alive rson(label), born at time* $t'_{\pi^n(label)}$*, lives until the first jump time of the process*

$$N^r_{\pi(label)}(n\int_{t'_{\pi^n(label)}}^{\cdot} b\eta^{\mathrm{tot},n}_s\, ds). \tag{1.7}$$

Then this individual branches, i.e. it has $R_{\pi^n(label)}$ *rsons, either 0 or 2. In the first case individual rson(label) is called dead. Otherwise its two sons are called rson(label)0 and rson(label)1.*

Remark 1.2.6:
When speaking about rescaled process, then the probability space $(\Omega, \mathcal{F}, \mathbf{P})$ *does not suffice. In fact we introduced* $(\Omega^n, \mathcal{F}^n, \mathbf{P}^n)$*, consisting of* n *independent copies of*

the original probability space. In order to avoid confusion, this overload of notation in e.g. $\mathbf{P}$*, will be suppressed in most parts of the paper.*
For the understanding it is often helpful to have the planar embedding as in Figure 1.1 in mind. ◇

The definitions so far have given a description of the process, but yet there is no embedding into a good state space. We do this in the next section and want to present some functionals of the process.

1.3 Functionals of Branching Processes

By the definitions of the general process (η, ξ) a full genealogy of catalyst and reactant is given. But instead of looking at the process (η, ξ) in general certain functionals of it will be considered, which live in some reasonable state spaces. Four different functionals are used and considered following [GPW06], when looking at the non-rescaled processes:

The total mass, Chapter 2: It is close to the classical Galton-Watson idea, since the process at time t is the number of individuals of the respective types at time t. The state space is $\mathbb{N}^2$. The process is indicated with a superscript tot:

$$\eta^{\text{tot}} \text{ and } \xi^{\text{tot}}. \tag{1.8}$$

The random tree, Chapter 4: It encodes the genealogy evolution up to time t into a metric space, called an $\mathbb{R}$-tree, which evolves in time. It is in fact like a genealogical tree, where the life times are made distances. State space is the set of $\mathbb{R}$-trees. The notation will be

$$\eta^{\text{for}} \text{ and } \xi^{\text{for}}. \tag{1.9}$$

The contour process, Chapter 5: It is possible to code the full genealogical tree up to a fixed time t by a continuous real-valued function, when a linear order is put on the individuals, i.e. it is possible to distinguish the first and the second son. State space is $C_{[0,\infty)}[0,\infty)$. This process will be

$$B \text{ and } C. \tag{1.10}$$

The point process, Chapter 6: For a given time $t > 0$ consider the extant individuals at time t. Any pair of them has a most recent common ancestor (MRCA). Collect labels and death times of the MRCAs and make points in the product space *set of labels* $\times[0,t)$. These random points constitute the point process description. State space is the set of integer-valued measures on $labels \times [0,t)$. It will be denoted by

$$\Pi^t \text{ and } \Xi^t. \tag{1.11}$$

These functionals were put in an ascending order in concern of information. In fact the last two descriptions contain the same information as the general process (η, ξ). Hence one could also start from the description given by the contour process or the point process. Starting from there one could also find the way to the other

descriptions and we would end up with the same process. So the processes (η, ξ), (B, C), (Π, Ξ) are just different sides of the same coin. As intuition and the relationship to Galton and Watson's ideas would have been lost, this path of representation was not selected. The first two functionals contain less information than the general process (η, ξ).

1.4 Main goals, methods and tools, quenched vs. annealed and context

1.4.1 Main goals

First of all we are interested in obtaining asymptotic results for the random genealogy of the reactant, especially the random tree that was the 2nd functional mentioned before. For this purpose we require the help of other functionals given in the previous section. For some questions it is helpful to use the contour process, for others the point process has advantages for the calculations.

The asymptotic results are comparable to diffusion limit results. For example the functional limit theorem for Brownian motion, where a rescaling of space and time gives a limit object (for Brownian motion: space-time rescaled random walks converge to Brownian motion). For each of the functionals a scaling as given in Definitions 1.2.4 and 1.2.5 is done. The rescaled processes will be given a superscript "n".

We will be able to show that there exist diffusion limits

- for the total mass process $(\eta^{\text{tot}}, \xi^{\text{tot}})$,
- for the reactant tree-valued forest ξ^{for},
- for a truncated reactant contour process C^{δ} and
- for the reactant point process Ξ^{t}.

One might wonder why there are only results for the reactant processes and not for the catalytic ones. So far there are no results for the catalyst in the literature, since the proofs for the reactant rely heavily on the independence of the individuals and this is only given for the reactant.

During all of this work we will link the different functionals, e.g. we will compare reactant limit contour and reactant limit point process.

Finally we want to establish some comparison results between reactant trees and ordinary Galton-Watson trees in the diffusion limit.

1.4.2 Organization of the diploma thesis, methods and tools

The diploma thesis is separated into two parts: a first part where the results are presented and a second part where the proofs are given. Each chapter containing the results is related to a chapter containing the corresponding proofs. In order to obtain these results several techniques are used throughout this diploma thesis.

The ideas for Chapters 2 and 7, describing the total mass process, consist of standard probability theory, including Markov processes and martingale problems. One

can consult [RW79], [KS00] and [EK86].

The chapter about the tree-valued process (Chapters 4 and 8) relies on the definition of *real trees*, where a superb introduction is given in [EPW06] and [EW06]. The techniques used in these chapters make use of the Kingman coalescent [Kin82] and its connection with trees, e.g. in [DK96].

The contour-process chapter (Chapters 5 and 8) is strongly related to the tree-chapter and hence its proofs are joined in one chapter. For some introduction see e.g. [NP89] or [LG96] and some proofs require martingale problem methods, e.g. Stochastic averaging as in [Kur92].

The point process chapter (Chapters 6 and 9) uses some elementary understanding of Poisson point processes as given in [RY91, Chapter XII.1]. The techniques for the proofs are Poisson approximation, local times and excursion theory as given in [RW79, Chapter VI.57].

1.5 Catalytic branching

In this closing section of the introduction we speak about an important characteristic of the catalytic branching scheme and the context into which it can be put.

1.5.1 Quenched vs. annealed

An important feature for the analysis will be the following observation. There are two different possibilities to understand the catalytic branching processes and to establish the limit results for the reactant functionals mentioned in the previous subsection:

Since the catalyst evolves autonomously one could first let the catalyst run. Then the catalyst gets frozen = quenched and for each catalyst sample path a reactant evolution is started. It is like conditioning on the catalyst total mass process, where some justifications will be given later. This conditioning will be called *quenched analysis*.

On the other hand we can let the process (η^{tot}, ξ) run aside. This is possible since the evolution of ξ at time t depends only on the catalyst total mass at time t: η_t^{tot}. This is called *annealed analysis*.

1.5.2 Context

The formulation itself can be brought into a wider context of processes with two (or more) populations. The catalyst-reactant scheme is somewhere in the middle of the independent branching scheme and the mutually catalytic one. In the next table the different names and references are presented, where A and B are stochastic processes and $A \to B$ means "A influences B" and so on.

name	relation	i.b.r. A	i.b.r. B	Papers
independent	$A \quad B$	Δt	Δt	any branching
catalytic	$A \to B$	Δt	$A_t \Delta t$	[Pen03], [GPW06]
catalytic, modified catalyst	$A \to B$	$\frac{g(A_t)}{A_t} \Delta t$	$A_t \Delta t$	this paper
mutually catalytic	$A \leftrightarrow B$	$B_t \Delta t$	$A_t \Delta t$	[CDG04]

Here "i.b.r." denotes the branching rate per individual. In the top and bottom cases also a modified version might be of interest.

Part I

Definitions and Results

2 The total mass process $(\eta^{\text{tot}}, \xi^{\text{tot}})$

This section is devoted to considering the *total mass process* of the catalytic branching process defined in the introduction. Proceeding from that definition the first section presents a definition of the total mass process. That means that we are going to look at the process counting the number of now-living individuals.

Within the first section we will set conditions on g, the branching modification of the catalyst, which will be valid throughout the diploma thesis. In the second section the corresponding diffusion total mass process will be presented. As a final result we will establish a diffusion limit theorem between the discrete process and the diffusion. The whole chapter is independent of the introduction in one sense that we could also start with Definition 2.1.1.

All proofs are given in chapter 7.

2.1 The catalyst and the reactant total mass process

We recall the definition of (η, ξ) from Section 1.2 in the introduction. Each catalyst individual has attached an alarm clock ringing after an exponential time and then it branches. At time t there are η_t^{tot} catalyst individuals alive. The first branching event of one of them happens after the minimum of these η_t^{tot} exponential times, i.e. for the catalyst after the first jump of $N^c(\int_t^{\cdot} g(\eta_s^{\text{tot}})\, ds)$. Then one individual dies (the one where the clock rings) and gets replaced by 0 or 2, i.e. total mass increases by 1 or decreases by 1. For the reactant similar ideas hold true and thus we get the following description:

- The total mass functional $(\eta^{\text{tot}}, \xi^{\text{tot}})$ of the general process, is an $\mathbb{N}^2$-valued Markov-process. The first coordinate η^{tot} is the *catalyst total mass*, the second coordinate ξ^{tot} is the *reactant total mass*. Both processes will have the structure of a critical binary branching process, i.e. in a branching event catalyst mass increases by 1 or decreases by 1.

- The catalyst η^{tot} emerges autonomously like a classical continuous-time Galton-Watson process, but with a reproduction function $g : [0, \infty) \to [0, \infty)$, which shall only depend on the individuals at time t, but not on time explicitly. That means that if there are n individuals alive then the reproduction rate is not like in a normal Galton-Watson-process n, but $g(n)$, i.e. the probability of a branching event occurring between time t and $t + s$ is

$$\int_t^{t+s} g(\eta^{\text{tot}}(u)) e^{-g(\eta^{\text{tot}}(u))u}\, du$$

  for small time steps s.

- The reactant ξ evolves depending on the catalyst. The probability of a branching event taking place in the reactant population between time t and $t + s$

is

$$\int_t^{t+s} b\eta^{\text{tot}}(u)\xi^{\text{tot}}(u)e^{-b\eta^{\text{tot}}(u)\xi^{\text{tot}}(u)u}\,du.$$

Hence we write down the following definition of the branching process.

Definition 2.1.1 (The discrete total mass process)**:** *(i) The catalyst total mass process $\eta^{\text{tot}} = (\eta_t^{\text{tot}})_{t\geq 0}$ is a critical binary modified Galton-Watson-branching process started in 1 with branching modification $g \in C([0,\infty),[0,\infty))$:*

$$\eta_t^{\text{tot}} \equiv (\eta_t^{\text{tot}};1) \mapsto \begin{cases} \eta_t^{\text{tot}}+1 \\ \eta_t^{\text{tot}}-1 \end{cases} \quad \text{each at rate } \tfrac{1}{2}g(\eta_t^{\text{tot}}). \tag{2.1}$$

(ii) For a given catalyst η^{tot} the reactant total mass process $\xi^{\text{tot}} = (\xi_t^{\text{tot}})_{t\geq 0}$ is a critical binary time-inhomogeneous branching process started in 1 with branching rate $b\eta_t^{\text{tot}}$:

$$\xi_t^{\text{tot}} \equiv (\xi_t^{\text{tot}};1) \mapsto \begin{cases} \xi_t^{\text{tot}}+1 \\ \xi_t^{\text{tot}}-1 \end{cases} \quad \text{each at rate } \tfrac{1}{2}b\eta_t^{\text{tot}}\xi_t^{\text{tot}}. \tag{2.2}$$

Remark 2.1.2:

- *Even if throughout the paper we will speak of branching processes, when considering the catalyst, an important feature which "characterizes" branching behavior is lost by introducing the modification g: independence of the individuals.*
  The evolution of one catalyst individual relies on the number of other individuals alive by the function g. Only if g is linear we have independence of the catalyst individuals. The reactant, however, evolves as a branching process.

- *The reactant branches in an "environment" given by the catalyst. The question arises arises whether this branching is somehow connected to what is understood as "branching in random environments", i.e. the branching events arise as in continuous-time Galton-Watson branching, but the offspring distributions are random. We do not discuss that question deeper, but refer the reader to [AK72] for elementary results about that question.*

- *The definition describes behavior in small time steps. Indeed this will allow us to denote a pre-generator U_1. This pre-generator defines a uniquely determined process $(\eta^{\text{tot}},\xi^{\text{tot}})$, for a broad range of g. Hence the definition given is detailed enough. This is made more precise in Lemma 2.1.6.*

◇

In the definition of the catalyst the function g emerges. In the forthcoming we will specify this function but let us first mention that we will already consider g defined as a function for all $x \geq 0$. This might seem cumbersome since so far we are only dealing with an integer-valued process, but later we will scale the masses of the individuals from 1 to $\frac{1}{n}$. Then the state space becomes $(\frac{1}{n}\mathbb{N})^2$ and then it makes sense to already have a function defined for all non-negative real numbers rather than non-negative integers. Additionally we require some more "local" properties, which are useful in the diffusion limit approximation.

Condition 2.1.3:
Let $g : [0,\infty) \to [0,\infty)$ be a function, which satisfies the following criteria:

(G1) *g locally Lipschitz-continuous on $[0,\infty)$,*

(G2) *$g(x) = 0 \Leftrightarrow x = 0$ and $\exists\, g_0 > 0\ \exists\, x_0$ s.t. $g(x) \geq g_0\ \forall\, x \geq x_0$,*

(G3) *$\exists\, C > 0$ and $\alpha \in [0,2)$ s.t. $g(x) \leq C(1 + x^\alpha)$ and*

(G4) *$\exists\, 0 \leq \beta < 1,\ c' > 0$ s.t. $\lim_{x\to 0} \frac{x^{1+\beta}}{g(x)} = c'$.*

For some justification of these conditions, we refer the reader to Section 7.1.1. Only let us mention that **G2** and **G3** let the catalyst total mass live in a world between Brownian motion with absorbing boundary at 0 and Anderson diffusion.

We will sometimes speak of the *usual conditions* on g.

As already mentioned before we are interested in the behavior of rescaled processes. Here we give the definition of the rescaled total mass processes. We will do *time* and *space* rescaling as when considering random walk limits for Brownian motion. Here *time* is related to time and *space* is related to individual masses.

- For the *space* rescaling we will consider the mass of one individual as $m = 1/n$, but starting at time $t = 0$ with total mass 1, each for catalyst and reactant. Therefore n starting individuals are present at time $t = 0$. So if we call $\eta^{\text{tot},n}$ the *rescaled catalyst total mass* and $\xi^{\text{tot},n}$ the *rescaled reactant total mass*, we start with:
$$\eta_0^{\text{tot},n} = 1,\ \xi_0^{\text{tot},n} = 1. \tag{2.3}$$
- The *time* rescaling is a bit different. Bear in mind that the number of catalyst individuals is given by $n\eta^{\text{tot},n}$, and the number of reactant individuals is $n\xi^{\text{tot},n}$.
  - The *time* rescaling for the catalyst is given by speeding up by factor n. Note that there are $n\eta^{\text{tot},n}$ catalyst individuals. All of them have assembled a Poisson-processes running with speed $ng(\eta^{\text{tot},n})/\eta^{\text{tot},n}$. Hence the first jump, the first branching, arises after an exponential $n^2 g(\eta^{\text{tot},n})$-time.
  - There is no *time* rescaling for the reactant. There are $n\xi^{\text{tot},n}$ reactant individuals and each of them has a Poisson-process with $nb\eta^{\text{tot},n}$-speed assembled. Hence the first jump, i.e. the first branching emerges after an exponential $n^2 b\eta^{\text{tot},n}\xi^{\text{tot},n}$-time.

The oncoming definitions can be made on the probability space as given in Definitions 1.2.4 and 1.2.5.

Definition 2.1.4 (Rescaled total mass process)**:** *(i) The* rescaled catalyst total mass process $\eta^{\text{tot},n} = (\eta_t^{\text{tot},n})_{t\geq 0}$ *is a critical binary modified Galton-Watson-branching process started in* 1 *with branching modification* $g \in C([0,\infty),[0,\infty))$*:*

$$\eta_t^{\text{tot},n} \equiv (\eta_t^{\text{tot},n}; \tfrac{1}{n}) \mapsto \begin{cases} \eta_t^{\text{tot},n} + \frac{1}{n} \\ \eta_t^{\text{tot},n} - \frac{1}{n} \end{cases} \quad \textit{each at rate } \tfrac{n^2}{2} g(\eta_t^{\text{tot},n}). \tag{2.4}$$

(ii) For a given catalyst $\eta^{\mathrm{tot},n}$ *the* rescaled reactant total mass process $\xi^{\mathrm{tot},n} = (\xi_t^{\mathrm{tot},n})_{t\geq 0}$ *is a critical binary time-inhomogeneous branching process started in* 1 *with individual branching rate* $n\eta_t^{\mathrm{tot},n}$*:*

$$\xi_t^{\mathrm{tot},n} \equiv (\xi_t^{\mathrm{tot},n}; \tfrac{1}{n}) \mapsto \begin{cases} \xi_t^{\mathrm{tot},n} + \frac{1}{n} \\ \xi_t^{\mathrm{tot},n} - \frac{1}{n} \end{cases} \quad \text{each at rate } \tfrac{n^2}{2} b \eta_t^{\mathrm{tot},n} \xi_t^{\mathrm{tot},n}. \tag{2.5}$$

Remark 2.1.5:
It is worth having a look at the scaling idea of the catalyst with branching modification g*. Consider sample paths for different* n*, so a different individual mass. By the definition just made we think of the branching behavior not changing for constant* total mass *of the process. That means if we change* n*, then the branching behavior is only dependent on the total mass of the catalyst* $\eta^{\mathrm{tot},n}$*, since* g *is directly a function of the catalyst total mass. Hence the catalyst individuals evolve like in a medium given by the total mass of the catalyst.*

Another possibility of scaling would be to think of the branching behavior depending on the number *of catalyst individuals alive. That means for mass* $m = \frac{1}{n}$ *to think of* $ng(n\eta^{\mathrm{tot},n})$ *as the catalyst's branching rate. This might match more the idea of the reactant's behavior, whose branching events are determined by the number of catalyst and reactant individuals alive. But clearly for convergence (given the time rescaling) it is crucial that either the rescaling depends on the* ∞*-close behavior of* g *or we restrict ourselves to* $g(x) = \mathcal{O}(x)$*. Hence one either needs to adapt the time rescaling or restrict oneself in the choice of functions* g*.* ◇

After the definitions it is time to speak about existence and uniqueness of the total mass catalytic branching process.

Lemma 2.1.6: *There is a unique process* $(\eta^{\mathrm{tot},n}, \xi^{\mathrm{tot},n})$ *(up to indistinguishability) with sample paths in the Skorokhod space* $\mathcal{D}_{\mathbb{R}_+^2}[0,\infty)$ *satisfying the definitions* (2.4) *and* (2.5)*. Moreover* $(\eta^{\mathrm{tot},n}, \xi^{\mathrm{tot},n})$ *is a martingale and a Feller-process.*

This existence and uniqueness theorem also holds for the process defined in Definition 2.1.1, as then $n = 1$. The proof is done by denoting a generator corresponding to the *jump process* behavior given in the definitions. This generator can then be shown to create a strongly continuous contraction semigroup $S_t^{(n)}$. The proof can be found in Lemma 7.2.1 with the help of Section 7.1.2.

With the lemma at hand we will henceforth only talk about the modification with càdlàg paths. As the process is a Feller-process, there exists a Feller-semigroup, which we will denote by

$$S_t^{(n)} f(x,y) := \mathbf{E}[f(\eta_t^{\mathrm{tot},n}, \xi_t^{\mathrm{tot},n}) | \eta_0^{\mathrm{tot},n} = x, \xi_0^{\mathrm{tot},n} = y], \tag{2.6}$$

when $f \in C_0(\mathbb{R}_+^2, \mathbb{R})$ and $x, y \in \mathbb{R}_+^2$.

From the definition of the processes it is obvious that if the catalyst or the reactant reach 0, they never leave again. As we consider a branching process, the discrete (time) theory says that extinction depends on the expectation of the offspring distribution. For the catalyst we have the following lemma:

Lemma 2.1.7: *Every (discrete) catalyst process, defined as in (2.4) with g satisfying the usual conditions dies out almost sure, i.e.*

$$T^{n,0} := \inf\{t \geq 0 : \eta_t^{\text{tot},n} = 0\} < \infty \ a.s. \tag{2.7}$$

and 0 is an exit boundary.

The proof of this lemma is given in Lemma 7.3.1.
We face a special problem, when the catalyst reaches zero: Not only the catalyst branching rate is zero, but the reactant branching rate is zero as well. So there will not be any more branching events. From that point on, the process is frozen, i.e. the catalyst stays in zero and the reactant will remain constant. This point will be of great importance throughout the diploma thesis.

Remark 2.1.8:
With the previous lemma at hand we already know that almost surely only a finite number of branching events take place for catalyst and reactant up to time $T^{n,0}$, since $T^{n,0} < \infty$ almost sure. ◇

2.2 The diffusion process (X, Y)

Like Brownian motion can be introduced as a diffusion limit of simple random walks with appropriate time- and space-scaling, the total-mass-process $(\eta^{\text{tot}}, \xi^{\text{tot}})$ has a diffusive limit process. By a look at Definition 2.1.4 we get an idea of what the generator for the diffusion could look like. We will define the diffusion process now by means of a system of SDEs. Later we will also discuss the process as a solution to a martingale problem.

Let a probability space $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbf{P}})$ with two independent Brownian motions be given.

Definition 2.2.1 (The diffusion total mass process)**:**
A diffusion total-mass process (X, Y) is given as a solution of the following system of SDEs

$$dX_t = \sqrt{g(X_t)}\, dW_t^1, \tag{2.8}$$

$$dY_t = \sqrt{bX_tY_t}\, dW_t^2, \tag{2.9}$$

where W^1 and W^2 are two independent Brownian motions.

The first question is, if there exist solutions for this SDE-system at all. Here the catalyst-reactant-type of the problem helps obtaining a positive answer. We will show that there exists a strong unique solution X for equation (2.8), since g is locally Lipschitz-continuous. That means for any Brownian motion path $W^1(\omega)$ on a probability space $(\Omega, \mathcal{A}, \mathbf{P})$ there exists a path $X(\omega)$ which solves the catalyst's SDE for this specific $\omega \in \Omega$. Then we can solve (2.9) for this single path $X(\omega)$ and we get a strong unique solution for the second equation.

This means that given two independent Brownian motions W^1 and W^2 there exists a strong unique solution (X, Y) for the SDE system. So the Definition 2.2.1 gives a well-defined total-mass-diffusion process (X, Y) for two independent Brownian motions on a given probability space:

Lemma 2.2.2 (Existence, uniqueness and Feller-property of the diffusion total mass)**:** *If g satisfies Condition 2.1.3, the total-mass-diffusion-process (X, Y) is a continuous unique strong solution of (2.8) and (2.9). Moreover (X, Y) is a Feller process and we denote the Feller-semigroup as $(S_t)_{t\geq 0}$:*

$$\text{For } x, y \in \mathbb{R}_+,\ f \in C_0(\mathbb{R}_+^2, \mathbb{R}) : S_t f(x, y) = E[f(X_t, Y_t)|X_0 = x, Y_0 = y]. \quad (2.10)$$

The proof of this lemma is split in several parts and can be found in Chapter 7 in two different sections. The first (Section 7.4) is showing existence and uniqueness according to the program presented before stating the lemma. The second section (Section 7.6) deals with the Feller-property. It is proven in several steps via coupling arguments using the catalytic setting of the system.

As we now know that there exists a *unique* strong solution, we can also speak of *the* total-mass-diffusion process. The process (X, Y) could have been introduced as well as a solution of the martingale problem $(U, \delta_1 \times \delta_1)$, where U is given by

$$Uf(x, y) = \frac{1}{2} g(x) \frac{\partial^2 f}{\partial x^2}(x, y) + \frac{b}{2} xy \frac{\partial^2 f}{\partial y^2}(x, y), \quad (2.11)$$

for $f \in C_0^2(\mathbb{R}_+^2, \mathbb{R})$.

Proposition 2.2.3: *The total-mass-diffusion process (X, Y) is given as the unique solution of the martingale problem $(U, \delta_x \times \delta_y)$, i.e. (X, Y) is a process, s.t.*

$$f(X_t, Y_t) - \int_0^t Uf(X_s, Y_s) ds, \quad (2.12)$$

is a $\hat{\mathbf{P}}$ -martingale for any given $f \in C_0^2(\mathbb{R}_+^2, \mathbb{R})$, where $\hat{\mathbf{P}}$ is the law related to X and Y on $C_{\mathbb{R}_+^2}[0, \infty)$.

There is a result, similar to the one given for the discrete setting, describing the extinction behavior of the catalyst:

Lemma 2.2.4 (Extinction of the catalyst total mass)**:** *Every catalyst process X with g satisfying the usual conditions dies out almost sure in finite time, i.e.*

$$\tau^0 := \inf\{t \geq 0 : X_t = 0\} < \infty \ a.s. \ , \quad (2.13)$$

and 0 is an exit boundary for X.

After this extinction time, the reactant will remain constant, which is obvious by the Strong Markov property of the process. The proof of this lemma relies mainly on looking at the speed measure of the catalyst.

As in the lemma just before we will sometimes only consider one coordinate of the pair (X, Y). In this case we will not always be totally stringent in description and will just say "the reactant process" Y, but we will understand "reactant process Y for a given catalyst X".

2.3 Convergence of the total mass process to (X, Y)

The main result of this chapter states the convergence of the discrete total mass process $(\eta^{\text{tot},n}, \xi^{\text{tot},n})$ to the continuous total mass process (X, Y). It is the first step towards obtaining "asymptotic results" for the catalytic branching process. As both processes have càdlàg paths, they both induce a measure on the set of càdlàg paths $\mathcal{D}_{\mathbb{R}_+^2}[0,\infty)$. This space equipped with the Skorokhod-metric d^{Sk} is a complete and separable metric space. Then the set of probability measures on $(\mathcal{D}_{\mathbb{R}_+^2}[0,\infty), d^{\text{Sk}})$ can be equipped with the Prohorov metric (for more, see [EK86, Chapter 3]). The Prohorov metric generates a topology on the set of probability measures on $\mathcal{D}_{\mathbb{R}_+^2}[0,\infty)$, for which the following theorem needs to be understood.

Theorem 2.3.1 (Weak convergence of the total masses)**:**
When g satisfies Condition 2.1.3, then

$$\mathcal{L}[(\eta^{\text{tot},n}, \xi^{\text{tot},n})] \Longrightarrow \mathcal{L}[(X, Y)] \text{ as } n \to \infty, \tag{2.14}$$

where convergence is weak convergence in the path space $\mathcal{D}_{\mathbb{R}_+^2}[0,\infty)$.

The proof is done by applying some martingale problem results from the book of Ethier and Kurtz [EK86]. As an easy consequence we get the following corollary:

Corollary 2.3.2 (Convergence of catalyst killing times)**:** *The catalyst killing times converge weakly as random variables on the state space $\mathbb{R}_+$:*

$$\mathcal{L}[T^{n,0}] \Longrightarrow \mathcal{L}[\tau^0] \text{ as } n \to \infty, \tag{2.15}$$

Proof: This is true by convergence of the finite dimensional distributions of $\eta^{\text{tot},n}$ to X: Let $t > 0$:

$$\mathbf{P}(T^{n,0} \le t) = \mathbf{P}(\eta_t^{\text{tot},n} = 0) \to \hat{\mathbf{P}}(X_t = 0) = \hat{\mathbf{P}}(\tau^0 \le t) \tag{2.16}$$

□

As always when having weak convergence we can create a probability space where almost sure convergence holds. Formally this can be done by Theorem 3.1.8 from [EK86]:

There is a realization of $\eta^{\text{tot},n}$ and X on a probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbf{P}})$ such that for a given $T > 0$:

$$\lim_{n\to\infty} \sup_{t \le T} |\eta_t^{\text{tot},n} - X_t| = 0, \tag{2.17}$$

Note that the Skorokhod metric was not used here, but the metric induced by the ∞-norm on the compact set $[0, T]$. This is possible since the catalyst X is continuous.

By these considerations it is possible to attach a (general) reactant process to the catalyst. That means for each $\omega \in \tilde{\Omega}$ we find reactants $\xi^n(\omega)$ s.t. $\xi^n(\omega)$ is a reactant process for a catalyst with total mass process $\eta^{\text{tot},n}(\omega)$. That is the basis for the quenched analysis and we give a definition of this probability space, which is given a product structure.

Definition 2.3.3 (The quenched reactant process)**:**
For each $n \in \mathbb{N}$ let the probability space $(\tilde{\Omega} \times \Omega_r^n, \tilde{\mathcal{F}} \otimes \mathcal{F}_r^n, \tilde{\mathbf{P}} \otimes \mathbf{P}_r^n)$ be given. It supports catalyst total mass processes $\eta^{\mathrm{tot},k}$ for each $k \in \mathbb{N}$, a limit catalyst total mass X and the general reactant process ξ^n, which is defined as in Definition 1.2.5. Additionally it holds

$$\lim_{k\to\infty} \sup_{t\leq T} |\eta_t^{\mathrm{tot},k} - X_t| = 0. \tag{2.18}$$

3 The real tree

To describe the genealogy of branching processes some tree-like-structure is needed. With this aim in mind in a first section motivation and definition of real trees are given. Next some simple properties and operators on real trees are defined. Additional concepts such as linearly ordered trees are treated in a third section.

3.1 Definition of the real tree

3.1.1 The graph-theoretical tree and definition of the real tree

This section is devoted to show that the concept of real trees ($\mathbb{R}$-trees) is a good way to describe a genealogy. Think of a genealogical tree of an ordinary binary continuous-time Galton-Watson process starting with one individual. It consists of a set of vertices V representing the individuals and edges E representing the relations between fathers and sons. One vertex $root \in E$ is distinguished as the first individual. Additionally we know the lifetime of each individual, so somehow a function $L : V \to [0, \infty)$. This is all the information we want to have for the *tree-valued process* and it is coded into

- a graph-theoretical tree (E, V),
- a root $\rho \in E$ and
- a lifetime function $L : V \to [0, \infty)$.

This concept of representing a genealogy contains some difficulties. Who are the individuals being alive 100 years after the start? One needs to add up lifetimes of all ancestors to get the actual time. The clue which helps is to use the edges for more. One has to switch a bit ideas: Up to now the vertices were understood as the individuals, now the edges will be the individuals, the vertices are the birth-and-death (or branching) events. That means the root individual is related to an edge, which either ends (0 sons) or splits into two edges (two sons) and so on. As before we assemble the lifetime of each individual to its edge. A good way to do this is not just to introduce another function $\tilde{L} : E \to [0, \infty)$, but to use the fact that the one-dimensional edges can code the lifetime much better than the zero-dimensional vertices. So we introduce distances: length of the father edge =lifetime of father.

This concept is rather something analytical than graph-theoretical. We have an object, the genealogical tree with distances and it will be the right idea to consider *metric spaces* that have a tree-like shape. These metric spaces are introduced now and get the name $\mathbb{R}$-tree or real tree.

Definition 3.1.1 (The real tree)**:**
A complete metric space (T, d) is called an $\mathbb{R}$-tree if it satisfies the following two conditions:

- *For every $x, y \in T$ there exists a unique isometric embedding*

$$\phi_{x,y} : [0, d(x,y)] \to T, \tag{3.1}$$

  such that $\phi_{x,y}(0) = x$ and $\phi_{x,y}(d(x,y)) = y$.

- *For every injective continuous map $\psi : [0,1] \to T$, it is true that*

$$\psi([0,1]) = \phi_{\psi(0),\psi(1)}([0, d(\psi(0), \psi(1))]). \tag{3.2}$$

The first condition in fact says that the tree is one-dimensional and that all points are connected by a continuous path in the tree. The second condition guarantees that there are no loops included in the tree. So in fact this definition gives what one might suspect to be a tree.

There is an equivalent definition which says that a metric space (T, d) is an $\mathbb{R}$-tree if it is path-connected and satisfies the four-point-condition. The latter means that for all $x_1, \ldots, x_4 \in T$ it is true that

$$d(x_1, x_2) + d(x_3, x_4) \leq \max\{d(x_1, x_3) + d(x_2, x_4), d(x_1, x_4) + d(x_2, x_3)\} \tag{3.3}$$

If T contains a point ρ, called the root, then we will say that (T, d, ρ) is a *rooted $\mathbb{R}$-tree*. Yet this point ρ does not need to have any special properties.

It is now possible to think of the set of all rooted $\mathbb{R}$-trees. In fact a lot of them will be "the same" in the sense that for two rooted $\mathbb{R}$-trees (T_1, d_{T_1}, ρ_1) and (T_2, d_{T_2}, ρ_2) there exists an isometric isomorphism $\iota : T_1 \to T_2$, s.t. $\iota(\rho_1) = \rho_2$. Therefore we will denote by $\mathbb{T}^{\text{root}}$ the set of all equivalence classes of rooted $\mathbb{R}$-trees. Henceforth we will not distinguish between a member of an equivalence class and the equivalence class itself.

For more about real trees consult [Chi01], [EPW06] or [EW06].

3.1.2 Extended definitions of real rooted trees and genealogical trees

For a rooted $\mathbb{R}$-tree some more terminology is used. The motivations for the names comes from tree-ideas or a genealogical point of view, which should be kept in mind for further considerations.

Definition 3.1.2 (Tree terminology)**:**
Let $(T, d, \rho) \in \mathbb{T}^{\text{root}}$ be given.

- *Write $[x, y]$ for the set*

$$\phi_{x,y}([0, d(x,y)]), \tag{3.4}$$

  which is the unique path connecting x and y. Sometimes the term geodesic *is used for this path.*

- *The path $[\rho, x]$ connecting the root ρ and an arbitrary point $x \in T$ is called an* arc.

- *If $x \in T$ is such that no other points lie beyond x, i.e. if for any $y \in T$: $x \in [\rho, y] \Rightarrow x = y$, then x is called a* leaf.

- *For two points $x, y \in T$ there are the two paths going to the root and the point where they first meet is denoted by $x \wedge y$. This point is called the* most recent common ancestor*(MRCA). Clearly this point is defined well, since there are unique geodesics.*
- *The* degree *of the root is defined as the number of different edges leaving the root, if there are finitely many. Otherwise the degree is ∞.*

There is a construction which allows to "build" trees, which we describe in the following remark.

Remark 3.1.3 (Gluing together rooted trees)**:**
Let two rooted $\mathbb{R}$-tress (T_1, d_{T_1}, ρ_1) and (T_2, d_{T_2}, ρ_2) be given, where the tree sets T_1, T_2 are disjoint. The second tree T_2 shall be glued to a point $a \in T_1$, and this new rooted $\mathbb{R}$-tree is called (T, d, ρ), where clearly $T = T_1 \cup T_2$, $\rho = \rho_1$. Additionally the metric d is defined to be the metric of single trees if both points are contained in either T_1 or T_2 and to be $d(x_1, x_2) = d_1(x_1, a) + d_2(\rho_2, x_2)$ for $x_1 \in T_1, x_2 \in T_2$. It can easily be checked that this construction still is a rooted $\mathbb{R}$-tree. $\diamond$

Consider a family history with a starting individual and known lifetimes of each individual until it vanishes and has an offspring. Then it is possible to construct a rooted $\mathbb{R}$-tree T representing the family history by giving to each individual a rooted $\mathbb{R}$-tree, (in fact only a line with length equals lifetime) which is then glued to the ancestor's death point.

By this construction we get a rooted $\mathbb{R}$-tree, were the root ρ is the birth point of the starting individual. The distance within the tree is the so called *genealogical distance metric* d_{gen}, represented for two points $x, y \in T$ by:

$$d_{\text{gen}}(x, y) = \hat{d}(x, \rho) + \hat{d}(y, \rho) - 2\hat{d}(x \wedge y, \rho), \tag{3.5}$$

where $\hat{d}(x, \rho)$ is the "age" of x, measured from the birth of the first individual on. The very simple tree consisting only of a line with a certain length will in the forthcoming sometimes be called *line segment of length L*. By this is meant a rooted $\mathbb{R}$-tree (T, d, ρ) with a leaf point t, s.t.: $T = \phi_{\rho,t}([0, L])$.

3.2 Operators and properties for rooted real trees

On the set of rooted trees $\mathbb{T}^{\text{root}}$ some operations are required. Later some basic properties about this set of real trees will be given.

Definition 3.2.1:
For $(T, d, \rho) \in \mathbb{T}^{\text{root}}$ define:

(i) the height $h(T) = \sup\{d(\rho, x) : x \in T\}$,

(ii) the cut operator $Q_t : \mathbb{T}^{\text{root}} \to \mathbb{T}^{\text{root}}$*, which cuts the tree at height t:*

$$Q_t(T, d, \rho) = (\{x \in T : h(x) \leq t\}, d, \rho), \tag{3.6}$$

(iii) the border operator $\partial Q_t : \mathbb{T}^{\text{root}} \to \{\text{set of labels}\}$*, which takes all elements of the tree with height t:*

$$\partial Q_t(T, d, \rho) = \{x \in T : h(x) = t\}, \tag{3.7}$$

(iv) the ϵ-trimming $S_\epsilon : \mathbb{T}^{\text{root}} \to \mathbb{T}^{\text{root}}$ *of* T*:*

$$S_\epsilon(T) = \{\rho\} \cup \{x \in T : \exists y \in T s.t.\, x \in [\rho, y] \text{ and } d(x,y) > \epsilon\} \tag{3.8}$$

There are some possible ways to make the space $\mathbb{T}^{\text{root}}$ a metric space. One way is to start with the Hausdorff metric d_H for compact subsets A and B of a single metric space (X, r):

$$d_H(A,B) := \inf\{\epsilon > 0 : B \subset A^\epsilon,\ A \subset B^\epsilon\} \tag{3.9}$$

where $A^\epsilon = \{x \in X : r(x,A) < \epsilon\}$. Two rooted $\mathbb{R}$-trees are metric spaces themselves and do normally not "live" on a common metric space. So for comparing two trees we need to embed them in a common metric space (Z, d_Z). It is useful to remember that we were only going to think of root-invariant equivalence classes of trees. The rooted Gromov-Hausdorff-distance between two rooted trees (T_1, d_{T_1}, ρ_1) and (T_2, d_{T_2}, ρ_2) is defined by:

$$d_{GH^{\text{rooted}}}(T_1, T_2) := \inf\{d_H^{(Z,d_Z)}(T_1', T_2') \vee d_Z(\rho_1', \rho_2')\}, \tag{3.10}$$

where the infimum is taken over all rooted $\mathbb{R}$-trees T_1', T_2' that are root-invariant isomorphic to T_1, T_2 as metric spaces and that are subspaces of a common metric space (Z, d_Z).

With this metric at hand we will list some results which were proven in [EPW06] and can be found as Theorem 2, Lemma 2.5 and Lemma 2.6. They are put together in the following proposition

Proposition 3.2.2:

- *The metric space* $(\mathbb{T}^{\text{root}}, d_{GH^{\text{rooted}}})$ *is separable and complete.*
- *A subset* $\mathcal{T} \subset \mathbb{T}^{\text{root}}$ *is pre-compact if for every* $\epsilon > 0$ *there exists a positive integer* $N(\epsilon)$ *such that each* $T \in \mathcal{T}$ *has an* ϵ*-net with at most* $N(\epsilon)$ *points. An* ϵ*-net for* T *is a set of points in* T*, s.t. that every* $x \in T$ *is* ϵ*-close to at least one of these points.*
- *For any* $(T, \rho) \in \mathbb{T}^{\text{root}}$*:*

$$d_{GH^{\text{rooted}}}(T, S_\epsilon(T)) \le \epsilon. \tag{3.11}$$

Later on we are going to consider "rescaled populations" in the sense that we will start with n individuals. Thus we need to consider various trees but the following remark tells that this does not make a new definition necessary:

Remark 3.2.3 (The forest)**:**
A set of real trees $\{(T_i, d_i, \rho_i) : i \in I\}$ *can be put together into one real tree* (T, d, ρ) *by gluing together all the trees to the new root* ρ*. The distance between* ρ *and* ρ_i *is chosen to be a non-negative constant. Such a set of trees is called a* real forest*. But the techniques for handling it are the same as the ones for trees by the previous considerations.* ◇

More advanced ideas about rooted trees can be found in [EPW06] and [GPW07]. In the latter additionally a measure on a tree is given in the definition of metric measure spaces. We do in fact not need this more advanced description, but for somebody interested in a dynamic description of the oncoming processes we refer to that paper.

3.3 Linearly ordered rooted real trees

Real trees do so far not allow to distinguish "family names". That means, if two trees are mirrored copies of each other, then they are considered the same. If one assembles labels to the vertices of the tree then the before "same" trees might differ. Hence this is a possible way to create a higher information functional than an ordinary real tree. In this section we are not too deeply interested in an exact labelling, but we are satisfied to have an ordering on these labels and consider only this order relation. We will indeed give an ordering for all elements of the tree not only for single arcs.

Therefore we will think of ways to characterize an ordering on the tree. This will be necessary for the contour process and the point process given in later chapters. In fact we will see that we just extend the ordering given as suffixes of the sons (e.g. son011001) in Definition 1.2.2 and 1.2.3.

The first intuitive idea of an order is to define that $x \leq_{\text{partial}} y$ if $x \in [\rho, y]$, but of course this only compares points on the same arc from the root, so it is a partial order. But we will require that the total order respects this partial order. The second idea is that when thinking of traversing the tree along the ordering then an already started subtree should be "finished" before starting to traverse the next disjoint subtree.

Definition 3.3.1 (Rooted, linearly ordered trees)**:**
The metric space $(T, d, \rho, \leq_{\text{lin}})$ is called a rooted, linearly ordered $\mathbb{R}$-tree, if (T, d, ρ) is a rooted $\mathbb{R}$-tree and the linear order $\leq_{\text{lin}}$ on T respects for all $x, y \in T$:

(i) If $x \leq_{\text{partial}} y$, then $x \leq_{\text{lin}} y$.

(ii) If $x \leq_{\text{lin}} y$ and for any $x', y' \in T$, s.t. $x \wedge y <_{\text{partial}} x' \wedge x$ and $x \wedge y <_{\text{partial}} y' \wedge y$, then $x' \leq_{\text{lin}} y'$.

The set of equivalence-classes of such trees is called $\mathbb{T}^{\text{root,lin}}$. Clearly the isometry here is required to conserve the linear order.

Additionally we will say that a rooted, linearly ordered $\mathbb{R}$-tree is finite, if it has only finitely many branch points. The set of all these trees is then denoted by $\mathbb{T}^{\text{root,lin}}_{\text{fin}}$.

The operators given in the previous section are also defined here. It is often helpful to think of a planar embedding of the rooted linearly ordered tree where the order is "left-to-right".

One should note that the set of "unordered" trees can be embedded into the set of ordered trees by putting a random ordering on all trees. So ordered trees are more complicated and therefore can somehow "store" more information. This will come back to our mind, when we look at the contour process and the point process on one side and the random tree on the other side.

4 The tree valued process $(\eta_t^{\text{for}}, \xi_t^{\text{for}})$

In this chapter the tree-valued catalytic branching process will be introduced. For this the concept of $\mathbb{R}$-trees, developed in the previous chapter will extensively be used. In a first section the definitions for the tree-valued catalyst and reactant random-variable are made. Rescaled forests, i.e. collections of trees, are presented in a next section. Then we will talk about basic properties and a convergence result for the reactant tree process. A last section contains a result comparing the reactant tree with the classical Galton-Watson forest.

This chapter is the "core" of the diploma thesis since it states results about the asymptotic genealogy of the catalytic branching process. All proofs are given in Chapters 8 and 9.

4.1 The catalyst and the reactant forest

In this section we are interested in describing the catalyst and the reactant forest in the discrete setting.

It is tempting to write down a description of the tree-valued reactant process as in the definition of the total mass processes in Definition 2.1.1: extending leafs in small time steps a bit or in a branching event assembling two new arcs to the branching leaf. The problem of this idea, is that it does not suffice the strong Markov property any more. Stopping the first time it jumps to two individuals two arc stubs, but of length zero, are attached to the tree. But distinguishing the two stubs contradicts the properties of metric spaces.

There are several attempts to solve this problem of giving a dynamic description of the process. A good idea is to assemble weights to the individuals, which means extending the state space from metric spaces to metric measure spaces. Consult the paper of Greven, Pfaffelhuber and Winter ([GPW07]) for more information about that.

But here the dynamic approach to cover the problem will *not* be taken. For the results to be proven it is sufficient to consider the "already done" tree. That means that from the general process (η, ξ) and a fixed $t > 0$ the tree can already be squeezed out. It can be constructed via the gluing technique described in the previous chapter. So we make the following definition on the probability space mentioned in Definition 1.2.1:

Definition 4.1.1:
The tree-valued process $(\eta_t^{\text{for}}, \xi_t^{\text{for}})_{t\geq 0}$ *with state space* $\mathbb{T}^{\text{root}} \times \mathbb{T}^{\text{root}}$ *is defined as follows:*

(i) The catalyst tree η_t^{for} *for a fixed* $t > 0$ *is obtained by gluing together the individuals described in the Definition 1.2.2 of* η *which live until time* t *to a root* ρ^{cat}*. The metric is the genealogical distance metric.*

(ii) The reactant tree ξ_t^{for} *for a given catalyst realization* $(\eta_s)_{s\geq 0}$ *and time* $t > 0$ *is obtained by gluing together the individuals described in the Definition 1.2.3 of* ξ *which live until time* t *to a root* ρ^{reac}*. The metric is the genealogical distance metric.*

Remark 4.1.2:

- *Dynamically the trees look like Markov processes growing, but in fact a dynamic definition would face the problems mentioned before the definition.*
- *Up to a fixed time* t *only a finite number of branching events will arise almost surely. Therefore* η_t^{for} *and* ξ_t^{for} *are finite trees almost surely and hence the state space (compact metric spaces) was described correctly.*
- *It is possible to consider the tree pair and then not having a reactant relying on the general process* η*, since the functional* η^{tot} *can be obtained from* η^{for} *with the help of* $\#\partial Q_t$*.*

◇

Clearly since the total mass process is a functional of the tree-valued one, the results from Chapter 2 on page 16 will play a role. As the catalyst dies out after an almost sure finite time $T^{1,0}$, the tree does not grow any more then. So all interesting things happen up to that time. Therefore it makes sense to consider the catalyst tree at that time. Moreover the reactant individuals keep on living without any branching after that killing time, if there are still some alive (for linear g see for example (2.8) in [Pen03]). Thus the reactant tree will neither show any interesting behavior after that killing time of the catalyst. Hence special interest in that chapter will be laid on the tree-valued random variable:

$$(\eta_{T^{1,0}}^{\text{for}}, \xi_{T^{1,0}}^{\text{for}}). \tag{4.1}$$

We take up the rescaling presented in Definition 1.2.4 and 1.2.5 from the Introduction. We recapitulate briefly that the number of starting individuals is increased from 1 to n. Therefore we have n single trees and we will call such an object a forest, even if all of the n rescaled trees in the forest are glued to the same root, which has then degree n. The individual mass is changed to $\frac{1}{n}$ and the catalyst is sped up by a factor n.

The rescaled processes are again defined in a static way on the probability space mentioned in Definitions 1.2.4 and 1.2.5:

Definition 4.1.3:
For $n \in \mathbb{N}$ *the process* $(\eta_t^{\text{for},n}, \xi_t^{\text{for},n})_{t\geq 0}$ *with state space* $\mathbb{T}^{\text{root}} \times \mathbb{T}^{\text{root}}$ *is defined as follows:*

(i) The rescaled catalyst forest $\eta_t^{\text{for},n}$ *for a fixed* $t > 0$*, is obtained by gluing together the individuals described in the Definition 1.2.4 of* η^n *which live until time* t *to a root* ρ^{cat}*. All* n *starting individuals are glued to the root* ρ^{cat}*, with different arcs, so that the root has degree* n*. The metric is the genealogical distance metric.*

(ii) The rescaled reactant forest *for a given catalyst realization* $(\eta_s^n)_{s\geq 0}$ *for a fixed* $t > 0$*,* $\xi_t^{\text{for},n}$*, is obtained by gluing together the individuals described in the*

Definition 1.2.5 of ξ which live until time t to a root ρ^{reac}. All n starting individuals are glued to the root ρ^{cat}, so that the root has degree n. The metric is the genealogical distance metric.

Again we face a problem after the extinction of the catalyst as in (4.1).

4.2 Tightness and Convergence of the reactant forest

Now we have to separate the two processes $\eta^{\text{for},n}$ and $\xi^{\text{for},n}$. We will give a convergence result for the latter, the reactant forest, and will leave the first one. This is done, because catalyst individuals do not evolve independently and applying the proofs which work for the reactant is not possible. When we will talk later about the contour process, then a catalyst process would not even be Markovian and its limit clearly not a diffusion.

The result, however, will not be a result for the whole process $\xi^{\text{for},n}$. It will describe the behavior of the "finished" reactant forest, i.e. the forest cut, when the catalyst has died out as in (4.1). We will give first a quenched result and later an annealed one.

For the quenched result introduce the following notation for conditional probability:

$$\mathcal{L}(\xi^{\text{for},n};\eta) := \mathcal{L}\left[\xi^{\text{for},n}_{T^{n,0}} | \eta^{\text{tot},n} = \eta\right]. \tag{4.2}$$

Note that this is the law of a $\mathbb{T}^{\text{root}}$-valued random-variable.

We fix an $\omega \in \tilde{\Omega}$ as in Definition 2.3.3. That means we have a convergent sequence of catalyst total mass processes $(\eta^{\text{tot},n}(\omega))_{n\in\mathbb{N}}$:

$$\lim_{k\to\infty} \sup_{t\le T} |\eta^{\text{tot},k}_t(\omega) - X_t(\omega)| = 0. \tag{4.3}$$

We will leave out the ω in what follows and will just use the word "quenched".

Then the following proposition holds:

Proposition 4.2.1 (Tightness of the reactant forest)**:** *The sequence of the quenched rescaled catalytic forest $\{(\xi^{\text{for},n}_{T^{n,0}};\eta^{\text{tot},n})\}_{n\in\mathbb{N}}$ is tight in the topology of the space $(\mathbb{T}^{\text{root}}, d_{GH^{\text{root}}})$.*

Since the space $\mathbb{T}^{\text{root}}$ with the Gromov-Hausdorff metric is complete and separable it is true that $\{(\xi^{\text{for},n};\eta^{\text{tot},n})\}_{n\in\mathbb{N}}$ is sequentially compact and therefore has a convergent subseries. Clearly one would like to extend that result by describing a unique limit and this can be done in the next theorem. For that purpose the δ-hitting time of the catalyst

$$\tau^{\delta} = \inf\{t \ge 0 : X_t \le \delta\} \tag{4.4}$$

needs to be defined. Then the following quenched result holds:

Theorem 4.2.2 (The reactant limit forest exists)**:**
There exists a random variable $Y^{\text{for}} \in \mathbb{T}^{\text{root}}$ (depending on X), s.t.:

$$\mathcal{L}\left[\xi^{\text{for},n};\eta^{\text{tot},n}\right] \xrightarrow{n\to\infty} \mathcal{L}\left[Y^{\text{for}};X\right]. \tag{4.5}$$

The law of Y^{for} *is given by*

$$\mathcal{L}\left[Y^{\text{for}}; X\right] = \lim_{\delta \to 0} \mathcal{L}\left[\mathcal{T}((\zeta_u^\delta)_{0 \le u \le \alpha_{4/b}})\right], \tag{4.6}$$

where convergence is in the Prohorov-metric of probability measures, $\mathcal{T}$ *is the mapping described on page 31.*
The diffusion ζ^δ *is the unique solution of the* $(A^\delta, \mathbb{D}(A^\delta))$ *martingale problem, where* $\mathbb{D}(A^\delta)$ *is given by:*

$$\mathbb{D}(A^\delta) = \{h \in C^1([0, \tau^\delta], \mathbb{R}) : h'|_{\{0,\tau^\delta\}} = 0, \frac{h'(\cdot)}{X_\cdot} \in C^2_{[0,\tau^\delta]}([0,\infty))\} \tag{4.7}$$

and for each $f \in \mathbb{D}(A^\delta)$*:*

$$A^\delta f(c) = \left(\frac{f'}{bX_c}\right)'(c), \tag{4.8}$$

Furthermore $\alpha_{4/b}$ *is the local time inverse at level* 0 *of* ζ^δ *of* $\frac{4}{b}$*.*

Note that when we write $\mathcal{L}\left[Y^{\text{for}}; X\right]$ it is not yet clear that this can be understood in the spirit of (4.2). So far we could say that we use this notation to express the dependence of the law on X. After the next annealed theorem we will see that we also can understand that expression as a conditional probability.

The proof occupies the biggest part of Chapter 8. First we cut the rescaled reactant forest and relate it to a contour process. The cut contour process sequence converges to a diffusion. As we can recover the cut tree from the contour and we have tightness of the forest sequence we get the result.

It it is even possible to extend this result to an annealed point of view. As we cannot state a limit result for the catalyst tree we have to restrict ourselves to a convergence theorem of the joint law of catalyst total mass and reactant forest.

Theorem 4.2.3 (Convergence of the joint law for the reactant forest)**:**
The sequence of the pair of rescaled catalyst total mass and rescaled reactant forest converges:

$$\mathcal{L}(\eta^{\text{tot},n}, \xi^{\text{for},n}_{T^{n,0}}) \Rightarrow \mathcal{L}(X, Y^{\text{for}}) \text{ as } n \to \infty. \tag{4.9}$$

Here convergence is understood as weak convergence on the set of probability measures on $\mathcal{D}_{\mathbb{R}^1_+}[0,\infty) \times \mathbb{T}^{\text{root}}$ *with the product topology.*

The proof of this theorem can be found in Section 8.6 due to the techniques used, which are developed in the point process Chapter.

4.3 Comparison result between the classical forest and the catalytic forest

In this last section we want to compare the forest of an ordinary Galton-Watson branching forest, which will be called Z^{for}, and the catalytic forest Y^{for}. Results for the classical (=Galton-Watson = Continuum Random Tree = CRT) forest are given

in [LG96] and [Ald93]. For Z^{for} the contour process is easily given as the excursion process of a Brownian motion β until local time at level zero reaches 2 (this can be seen for the case of a fixed catalyst $X \equiv 1$ and $b = 2$ by the arguments in the proof of the limit contour).

For this section fix a catalytic background $(X_r)_{0\le r\le \tau^0}$. The way to compare the two trees is to take a fixed time t and to relate the extant individuals of Y at time t to the extant individuals of Z at a non-random time $s(t)$, depending on the fixed catalyst. If the metric is also changed, then in the end we see that Y^{for} looks like a stretched CRT Z^{for}.

We define the scaling function s_t for $t < \tau^0$:

$$s_t : \begin{cases} [0,\tau^0] & \to [0,\infty) \\ h & \mapsto \frac{b}{2}\int_{t-h}^{t} X_s\, ds \end{cases}. \tag{4.10}$$

By an easy argument one can show (see Lemma 9.2.2) that the contour processes of the CRT, which is equal to β and the reactant contour ζ^δ are related by s_t:

$$(\zeta^\delta_u)_{u\ge 0} \overset{d}{=} (s_t^{-1}(\beta_{\gamma^{-1}(u)}))_{u\ge 0}, \tag{4.11}$$

where γ is the time change depending on s_t (see (9.24) for the definition of γ).

For the trees things are similar and we get the following proposition:

Proposition 4.3.1 (Stretching tree metric)**:** *Let Z^{for} be a classical Galton-Watson forest and Y^{for} a catalytic branching forest with fixed catalyst $(X_s)_{0\le s\le \tau^0}$. Then for any $t < \tau^0$ let*

$$\tilde{Y}^{\text{for}}_t := \partial Q_{s_t(t)}(Z^{\text{for}}) \tag{4.12}$$

and for $u_1, u_2 \in \tilde{Y}^{\text{for}}_t$, i.e. $u_1, u_2 \in \partial Q_{s(t)}(Z^{\text{for}})$ define:

$$\mathrm{d}_{\tilde{Y}^{\text{for}}}(u_1,u_2) := 2s_t^{-1}\left(\tfrac{1}{2}d_{Z^{\text{for}}}(u_1,u_2)\right). \tag{4.13}$$

Then it holds that

$$(\tilde{Y}^{\text{for}}, \mathrm{d}_{\tilde{Y}^{\text{for}}}) \overset{d}{=} (\partial Q_t Y^{\text{for}}; X), \tag{4.14}$$

where equality in distribution is meant to be on the set of ultrametric spaces.

The proof is done via the point processes π^t and $\pi^{\beta,t}$, which describe distances between extant individuals.

Remark 4.3.2:

- *The idea for a statistician is to have a sample of extant individuals that are known to evolve according to an inhomogeneous branching mechanism. The question is whether it is possible to determine the unknown inhomogeneity, i.e. the catalyst. But the catalyst is encoded in the scale function s_t and one can compare the distances with the CRT-distances.*

- *The metric space describing the extant individuals is in fact an ultra-metric space, that means for any $u_1, u_2, u_3 \in \partial Q_t Y^{\text{for}}$ it holds that:*

$$\mathrm{d}_{Y^{\text{for}}}(u_1,u_3) \le \max\{\mathrm{d}_{Y^{\text{for}}}(u_1,u_2), \mathrm{d}_{Y^{\text{for}}}(u_2,u_3)\}. \tag{4.15}$$

◇

5 The contour process (B, C)

Closely related to the forest-valued processes is the description of the populations (η, ξ) as contour processes. By this a coding of the tree-structure into a positive continuous function is meant. The major difference to the preceding two functionals is that the contour process keeps track of all information available from the general process (η, ξ). In comparison to the tree-valued process it also remembers "family names". That means that the contour process allows to distinguish scenarios where in the first one son00 has a long history, son01 a short history and in the second scenario the other way round. The process description we will give is not a dynamic one.

5.1 Contour processes and branching populations

Imagine a (finite) genealogical tree with names (=labels) at the vertices be given. Additionally an ordering of the names should be available, e.g. like in an address book. One can traverse the tree starting from the root visiting the vertices according to the order, i.e. visiting the individuals as listed in the telephone book, along their lifetimes. This walk through the tree visits every point in the tree.

The idea now is to code this walk into a function $e : [0, \infty) \to [0, \infty)$ by traversing the genealogical tree. We go along the unique geodesics following the ordered labels and denote on the ordinate the distance to the root. By this procedure a continuous positive function is given. Shortly, the contour is the function where elapsed transversal time is mapped to recent height of the traversal, i.e. the distance from root to the recent point of the tree. Of course to walk continuously one also has to walk in the opposite direction of the total order at some times to reach the next point, with a higher order. After finishing the walk through the tree, the height stays zero. See Figure 8.1 to get an idea of the procedure.

It is true that the linearly ordered trees in $\mathbb{T}^{\mathrm{root,lin}}$ have a special ordering respecting Definition 3.3.1. Therefore we can understand the contour as the walk "around" the tree. This walk around the tree can be done with a given speed $\sigma > 0$ and different σ results in a different contour. We will denote this mapping from finite compact linearly ordered rooted $\mathbb{R}$-trees to continuous functions by $\mathcal{C}(\cdot : \sigma)$:

Definition 5.1.1 (Tree to contour mapping)**:**
For a $\sigma > 0$ the mapping $\mathcal{C}(\cdot : \sigma) : \mathbb{T}_{\mathrm{fin}}^{\mathrm{root,lin}} \to C_{[0,\infty)}^{0}[0, \infty)$ is defined as the mapping which maps a finite linearly-ordered rooted tree $(T, d, \rho, \leq)$ to the continuous function $\mathcal{C}(T : \sigma) : [0, \infty) \to [0, h(T)]$ in the following way:
Denote the number of branch points of T in ascending linear order: $x_0 = \rho \leq x_1 \leq x_2 \leq \cdots \leq x_{\hat{N}}$. Whenever $\rho \in [x_n, x_{n+1}]$, then add ρ to the sequence of branch

points $x_1, \ldots, x_n, \rho, x_{n+1}, \ldots x_{\hat{N}}$. *Then* $\mathcal{C}(T : \sigma)$ *is given by*

$$u \mapsto \begin{cases} d(\rho, \phi_{\rho,x_1}(\tilde{u})) & \text{for } 0 \le \tilde{u} := \sigma u \le d(\rho, x_1) \\ d(\rho, \phi_{x_1,x_2}(\tilde{u})) & \text{for } 0 \le \tilde{u} := \sigma u - d(\rho, x_1) \le d(\rho, x_2) \\ \vdots & \vdots \\ d(\rho, \phi_{x_N,\rho}(\tilde{u})) & \text{for } 0 \le \tilde{u} := \sigma u - d(\rho, x_1) - \ldots d(x_{N-1}, x_N) \le d(x_N, \rho) \\ 0 & \text{otherwise} \end{cases}. \tag{5.1}$$

For another construction of this mapping even for non-finite trees see remark 3.2 of [EW06].

Remark 5.1.2:
By the contour process we have indeed a planar embedding of a real tree. For more about that, see [NP89]. ◇

One also can go the other direction: Let a continuous function $e : [0,1] \to [0,\infty)$ be given with $e(0) = e(1) = 0$ and $e(x) \ge 0$ for all $x \in (0,1)$. Then define an equivalence relation $\sim_e$ on $[0,1]$ by

$$x \sim_e y \text{ if } e(x) = \min_{z \in [x \wedge y, x \vee y]} e(z) = e(y). \tag{5.2}$$

The metric space $([0,1]/\sim_e, d)$, where

$$d(x,y) = e(x) + e(y) - 2 \min_{z \in [x \wedge y, x \vee y]} e(z), \tag{5.3}$$

is then easily checked to be a rooted compact $\mathbb{R}$-tree (see Lemma 3.1 of [EW06]). To be more formal we introduce a mapping from the following set of functions

$$C^{0,*}_{[0,\infty)}[0,L] := \{f \in C([0,L],[0,\infty)) : \ f(0) = f(L) = 0,\ f(x) \ge 0 \ \forall x \in (0,L)\} \tag{5.4}$$

with the previously described properties, where L replaces 1 and make the following definition:

Definition 5.1.3 (Contour to tree mapping)**:**
The mapping $\mathcal{T} : C^{0,*}_{[0,\infty)}[0,L] \to \mathbb{T}^{\text{root,lin}}$ *is defined for a continuous* $e \in C^{0,*}_{[0,\infty)}[0,L]$ *by:*

$$\mathcal{T}(e) = ([0,L]/\sim_e, d, \le_{\text{lin}}), \tag{5.5}$$

where the equivalence relation $\sim_e$ *and the metric* d *are defined as in* (5.2) *and* (5.3). *The root* ρ *is the equivalence class corresponding to zero and* $\le_{\text{lin}}$ *is the linear order induced from the interval.*
Additionally define the mapping

$$\mathcal{T}_{\text{unord}} : C^{0,*}_{[0,\infty)}[0,L] \to \mathbb{T}^{\text{root}}, \tag{5.6}$$

which does the same as $\mathcal{T}$*, but gives up the ordering of the tree.*

Remark 5.1.4:
- *Both mappings are continuous by definition (quotient mapping!).*

- *For finite rooted linearly ordered $\mathbb{R}$-trees and arbitrary $\sigma > 0$ it is then true that $\mathcal{T} \circ \mathcal{C} = id$. The same is true the other way round if the traversal speed σ is adjusted correctly. But we will in fact use the first identity only; thus the traversal speed is* not *affecting the tree structure. For more information, see [EW06].*

$\diamond$

For our purposes the tree-valued process does not yet give an ordered tree, which could be put into the mapping $\mathcal{C}(\cdot : \sigma)$ and would give us a contour. This problem needs to be taken up now. So, when starting with the non-linearly ordered tree-valued process from the previous chapter we could think of $\mathcal{C}$ as a multi-valued mapping. This multi-valued mapping would transport equal weights to each possible ordering of a given finite tree η_t^{for} or ξ_t^{for}. But in fact it will be easier to extract the functional "contour" directly from the general process (η, ξ). Therefore define analogously to the previous chapter a catalyst and reactant process $(\tilde{\eta}, \tilde{\xi})$, but with values in linearly ordered trees $\mathbb{T}^{\text{root,lin}}$:

Definition 5.1.5:
For $n \in \mathbb{N}$ the random-variable $(\tilde{\eta}^n, \tilde{\xi}^n)$ with values in $\mathbb{T}^{\text{root,lin}} \times \mathbb{T}^{\text{root,lin}}$ is defined as follows:

(i) The first coordinate $\tilde{\eta}^n$, is obtained by gluing together the individuals described in Definition 1.2.4 of η^n, which live until time $T^{n,0}$, to a root ρ^{cat}. All n starting individuals are glued to the root ρ^{cat}, with different arcs, so that the root has degree n. The metric is the genealogical distance metric and the linear order is the one obtained from the labels of the general process η.

(ii) The second coordinate $\tilde{\xi}^n$, for a given catalyst realization $(\eta_s^n)_{s\geq 0}$, is obtained by gluing together the individuals described in Definition 1.2.5 of ξ, which live until time $T^{n,0}$, to a root ρ^{reac}. All n starting individuals are glued to the root ρ^{reac}, so that the root has degree n. The metric is the genealogical distance metric and the linear order is the one obtained from the labels of the general process ξ.

This definition gives us the full genealogical tree of catalyst and reactant, at least until the time $T^{n,0}$, after which nothing interesting happens any more.
We will leave out the superscripts n sometimes and will understand this as the case $n = 1$.

5.2 The catalyst and the reactant contour

The catalyst and the reactant contour are now defined. With the mapping $\mathcal{C}$ and the random variable $(\tilde{\eta}, \tilde{\xi})$ from the previous section we can give a precise definition of the contour process. Before starting remember that the killing time of the catalyst $T^{1,0}$ is $< \infty$ almost surely. We define the contour processes by traversing the catalyst and the reactant tree each with speed 1.

Definition 5.2.1 (The discrete contour process)**:**
The catalyst and the reactant contour process $(B_u, C_u)_{u\geq 0}$ is a $\mathbb{R}_+^2$ valued process defined by:

$$(B_u)_{u\geq 0} = \mathcal{C}(\tilde{\eta} : 1), \tag{5.7}$$

$$(C_u)_{u\geq 0} = \mathcal{C}(\tilde{\xi} : 1). \tag{5.8}$$

Note that in the definition the height of the contour is random, it is $T^{1,0}$. If we consider a catalyst realization to be given, then the reactant contours all have the same maximum height. The contour process consists of line segments with slope $+1$ and -1 which have random length.

We note that this contour process leaves zero (the root) once and returns to zero (once). After this returning time the contour stays at zero. So the reactant contour is zero after a time $L(C, 4/b)$ given by the following random-variable:

$$L(C, \tfrac{4}{b}) := \inf\{u \geq 0 : \lim_{\epsilon \to 0} \frac{1}{\epsilon} \int_0^u \mathbb{1}_{\{C_v \in [0,\epsilon)\}} \frac{2}{b\eta^{\mathrm{tot}}_{C_v}}\, dv = \frac{4}{b}\}. \tag{5.9}$$

Even if this seems to be a quite difficult description of the time until which the reactant contour runs, it is in fact the easiest way to obtain a local time characterization of the limit contour.

The same is true for the catalyst and a similar time $L(B, 4)$. Later we will take that idea up. Next a remark about the linear ordering on extant individuals is described:

Remark 5.2.2:
By the definition of the contour a linear ordering on the reactant individuals

$$y_1, y_2, \dots y_m \in \partial Q_t(\xi^{\mathrm{for}})$$

alive at time t can be given:

$$y_i \leq_{\mathrm{lin}} y_j \text{ is true}, \tag{5.10}$$

if the point y_i is traversed by the contour before y_j. We will use this idea for the point process. ◇

Now we give the rescaled contour processes.

The rescaled reactant contour process is defined analogously, only one has to consider what speed σ to choose for traversing the tree. The total population size is of order $\mathcal{O}(n^2)$ and the length of each line segment is of order $\mathcal{O}(\frac{1}{n})$. Hence the right choice to get a non-trivial limit-contour is traversal speed $\sigma = n$ for the n-th approximation step.

Definition 5.2.3 (The rescaled contour process)**:**
The catalyst and the reactant contour process $(B^n_u)_{u\geq 0}$ and $(C^n_u)_{u\geq 0}$ are $\mathbb{R}_+$-valued random variables defined by

$$(B^n_u)_{u\geq 0} = \mathcal{C}(\tilde{\eta}^n : n), \tag{5.11}$$

$$(C^n_u)_{u\geq 0} = \mathcal{C}(\tilde{\xi}^n : n). \tag{5.12}$$

The reactant contour is zero after a time $L(C^n, 4/b)$ given by the following random-variable:

$$L(C^n, \tfrac{4}{b}) := \inf\{u \geq 0 : \lim_{\epsilon \to 0} \frac{1}{\epsilon} \int_0^u \mathbb{1}_{\{C^n_v \in [0,\epsilon)\}} \frac{2}{b\eta^{\mathrm{tot},n}_{C^n_v}}\, dv = \frac{4}{b}\}. \tag{5.13}$$

Outgoing from these definitions we can establish some properties but most of them restrict to the reactant. The next result shows what happens for a fixed catalyst realization $(\eta_t^n)_{t\geq 0}$.

Before stating the lemma it is a good idea to define the slope sign V of the reactant contour as an process:

$$V_u^n := \text{sign}(\text{slope}(C_u^n)) \in \{-1, 1\}, \tag{5.14}$$

for all $0 \leq u \leq T^0$, where C_u does not change direction. If it is chosen to have càdlàg paths then the following lemma holds:

Lemma 5.2.4: *The process $(C_u^n, V_u^n)_{u\geq 0}$ is a $[0, T^{n,0}] \times \{-1, 1\}$-valued Markov-process, whose generator is given by the closure of the operator $(A^n, \mathbb{D}(A^n))$, where:*

$$A^n f(c, v) = nv\frac{\partial}{\partial c} f(c, v) + \frac{b}{2} n^2 \eta_c^{\text{tot},n}(f(c, -v) - f(c, v)), \tag{5.15}$$

for all $f \in \mathbb{D}(A^n)$, when

$$\mathbb{D}(A^n) = \{h \in C^{1,0}([0, T^{n,0}] \times \{-1, 1\}, \mathbb{R}) : \frac{\partial h}{\partial c}|_{\{0,T^{n,0}\}\times\{-1,1\}} \equiv 0\}. \tag{5.16}$$

until the time, when the n trees in the reactant forests have been traversed, i.e. until $L(C^n, \frac{4}{b})$ given as in (5.13). After that time the reactant contour stays zero.

For the proof, see Lemma 8.4.2.

We note that the catalyst η^{tot} plays a big role in the behavior of the contour. The higher η^{tot} is, the faster the contour changes its direction. That means that the line segment in the contour with constant slope get shorter. In the case when we get close to the extinction time of the catalyst then η^{tot} gets very low and the arcs in the tree get longer and grow until they are reflected at the extinction time. This fact will play a role when we want to extend the limit result we give in the next section.

For the catalyst contour we see that non-independence of the individuals disallows the contour to be Markovian. Even if g is a piecewise linear function one does not necessarily get a Markov process for the catalyst contour.

5.3 Convergence of the truncated reactant contour

Within this section a quenched convergence result for the reactant contour process is given. We assume to be in a situation as in (4.2), i.e. we have a fixed convergent sequence of catalyst realizations. Unfortunately it is not possible to state a convergence theorem for the whole contour. We need to cut the tree before the extinction of the catalyst. Therefore define the δ-hitting times of the rescaled catalyst total mass process $\eta^{\text{tot},n}$:

$$T^{n,\delta} = \inf\{t \geq 0 : \eta_t^{\text{tot},n} \leq \delta\}, \tag{5.17}$$

and of the limit catalyst X:

$$\tau^\delta = \inf\{t \geq 0 : X_t \leq \delta\}. \tag{5.18}$$

Remember that both of these stopping times are almost surely finite by Lemmas 2.1.7 and 2.2.4. Then one can define the δ-cut contour $C^{n,\delta}$ by:

$$C^{n,\delta} := \mathcal{C}(Q_{T^{n,\delta}}(\tilde{\xi}^{n,\delta}) : n). \tag{5.19}$$

With these definitions the following convergence result for the cut contour holds:

Theorem 5.3.1 (Reactant limit contour)**:**
Let a realization of the catalyst be given such that (2.18) holds. Consider the linear operator $(A^\delta, \mathbb{D}(A^\delta))$, where

$$A^\delta f(c) = \left(\frac{f'(\cdot)}{bX_\cdot}\right)'(c), \tag{5.20}$$

for all $f \in \mathbb{D}(A^\delta)$, when

$$\mathbb{D}(A^\delta) = \{h \in C^1([0,\tau^\delta], \mathbb{R}) : h'|_{\{0,\tau^\delta\}} = 0, \frac{h'}{X_\cdot} \in C^1_{[0,\tau^\delta]}[0,\infty)\}. \tag{5.21}$$

Then:

(i) The $(A^\delta, \mathbb{D}(A^\delta))$ martingale problem is well-posed and

(ii) if ζ^δ is the solution of the $(A^\delta, \mathbb{D}(A^\delta))$ martingale problem, then:

$$\mathcal{L}[(C^{n,\delta}_u)_{0\le u\le L(C^{n,\delta},4/b)}] \xrightarrow{n\to\infty} \mathcal{L}[(\zeta^\delta_u)_{0\le u\le \alpha_{4/b}}], \tag{5.22}$$

where convergence is weak convergence of continuous processes.

Here $\alpha_{4/b}$ is the shortcut for $(l^0_\cdot(\zeta^\delta))^{-1}(\frac{4}{b})$, the inverse of local time at level zero. The proof is given in Section 8.4. The convergence is shown via Stochastic Averaging techniques. Uniqueness is done by transforming the martingale problem to an easier one without drift.

Remark 5.3.2:
Looking back at the discrete process we see that higher catalyst total mass is speeding up the change of direction. In the limit case it seems to be the other way round: High catalyst total mass decreases the quadratic variation. In fact there is a mistake in the first sentence: When going to limits the part where catalyst total mass occurs runs with high speed (n^2) and it is better if it does not change direction too often in order not to disappear in the limit.

One could ask whether there is an annealed result as well. If in the previous chapter we had an annealed result for the convergence of the trees, then there also should be one in that case. The proof of the following corollary is in fact part of the annealed result for trees:

Corollary 5.3.3: *For fixed $\delta > 0$ the sequence*

$$((\eta^{\mathrm{tot},n}_t)_{0\le t\le T^{n,\delta}}, C^{n,\delta})_{n\in\mathbb{N}} \tag{5.23}$$

converges weakly to an $\mathbb{R}^2$-valued process. Topology is the product topology in the Skorokhod space of càdlàg functions.

5.4 On top of the limit reactant tree

Naturally after Theorem 5.3.1 one has the question if it is possible to extend the τ^δ-cut reactant limit contour to the case $\delta \to 0$. We will regard that as a quenched question and in fact there is not a single answer: all depends on the given catalyst and the height of each contour path. To speak about that we define the reactant extinction time

$$\rho^0 := \inf\{t > 0 : Y_t = 0\}. \tag{5.24}$$

We differentiate between two cases:

$$\{\rho^0 < \tau^0\} \text{ and } \{\tau^0 < \rho^0\}. \tag{5.25}$$

The first one means that the entire reactant contour ζ^δ lies below the given horizontal line τ^0. This case will be easy to deal with, since no extension of the contour is necessary. That means in this case we have, when for example $\rho^0 < \tau^\delta$:

$$(\zeta^0; \rho^0 < \tau^0) := \zeta^\delta. \tag{5.26}$$

We wonder if the case $\rho^0 < \tau^0$ arises and will give a partial answer in the oncoming proposition.

The second event described above is more difficult to deal with, since the reactant tree was cut at height τ^δ and even for $\delta \to 0$ there remain some branches above that line. Then something interesting happens: Approaching τ^0 the catalyst approaches 0 and hence the reactant loses branching likelihood, i.e. its branches get longer. Since there are "many" individuals alive the reactant contour gets looking like a hedgehog and this can be expressed by quadratic variation going to infinity.

Both aspects get reflected in the following proposition:

Proposition 5.4.1: *Let a fixed catalyst $(X_t)_{t\geq 0}$ and its killing time τ^0 be given. For $\delta > 0$ let ζ^δ be the reactant limit contour. Then*

$$\mathbf{P}[\lim_{\delta\to 0}\langle\zeta^\delta, \zeta^\delta\rangle_{\alpha_{4/b}} < \infty | X] = \mathbf{P}[\rho^0 < \tau^0] \tag{5.27}$$

and

$$\mathbf{P}[\lim_{\delta\to 0}\langle\zeta^\delta, \zeta^\delta\rangle_{\alpha_{4/b}} = \infty | X] = \mathbf{P}[\rho^0 > \tau^0]. \tag{5.28}$$

Since diffusions have a well-defined finite quadratic variation, we obtain the following corollary:

Corollary 5.4.2: *The reactant tree Y^{for} can be associated with a diffusion ζ^0 via $\mathcal{T}$ only on the event:*

$$\{\rho^0 < \tau^0\}. \tag{5.29}$$

For linear branching modification $g(x) = ax$ this event has probability $(4a/b+1)^{-\frac{1}{2}}$. Otherwise in the case $\{\rho^0 > \tau^0\}$ the reactant forest Y^{for} cannot be associated with a diffusion process.

6 The point process (Π^t, Ξ^t)

Another way to describe the genealogy of the catalytic branching model is the point process. To do that we first introduce a short general description of point processes in the branching setting. Then we will specify this for the catalyst and the reactant setting. Finally the chapter ends with a convergence result for the rescaled reactant point process.

6.1 Point processes and genealogy

Let a genealogy starting with one individual be given, for example as in Definition 1.2.2 or 1.2.3. Then a linearly ordered tree, say T, as $\tilde{\eta}^n$ in Definition 5.1.5 can be constructed.

Now fix a time $t > 0$. At this time t the population consists of $\#\partial Q_t(T)$ extant individuals. If one considers the minimal subtree spanned by these extant individuals, there are exactly $\#\partial Q_t(T) - 1$ most recent common ancestors (MRCAs) of them. We label the MRCAs in ascending linear order and then they are given by a subset of T consisting of the following points:

$$x_1 \leq_{\text{lin}} x_2 \leq_{\text{lin}} \cdots \leq_{\text{lin}} x_{\#\partial Q_t(T)-1}. \tag{6.1}$$

These points and a constant $\nu > 0$ are the ingredients of the point process $\mathcal{P}^t$. We want to make a point in $(k\nu, s) \in \mathbb{R}^2$, when x_k dies at time s. Hence the point process just denotes the extinction times of the MRCAs, one after another. Thus we make the following definition:

Definition 6.1.1 (Point process of a genealogy)**:**
For a fixed $t > 0$, a given genealogy T the point process $\mathcal{P}^t(T;\nu)$ is an integer-valued random measure on the set

$$\{\nu, 2\nu, \ldots, (\#\partial Q_t(T) - 1)\nu\} \times [0, t)$$

. Its distribution for $0 \leq a < b < t$, $m \leq \#\partial Q_t(T) - 1$ is given by:

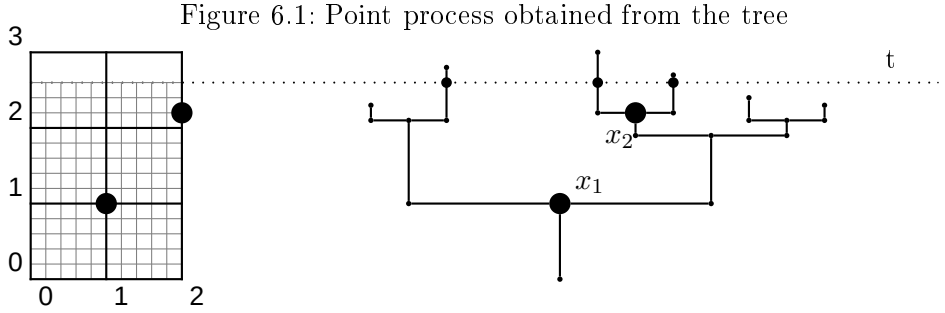
$$\mathcal{P}^t(T;\nu)(\{\nu, 2\nu, \ldots, m\nu\} \times [a, b]) = \#\{i \in \{1, 2, \ldots, m\} : a \leq x_i \leq b\}. \tag{6.2}$$

We would like to point out that this point process should *not* be understood as a process in the variable t. But in fact with the collection $(\mathcal{P}^t)_{t\geq 0}$ the whole genealogy can indeed be recovered.

Sometimes we will refer to the first coordinate of the random measure as the time coordinate and the second one as the level coordinate.

Indeed we will also face the fact, that we are dealing with a set of genealogies, as in the situation of a genealogical forest consisting of genealogical trees. Then the question is where the MRCA of two individuals alive at time t, which are not in the same tree of the forest, is located. The MRCA is then at the level of the root,

Figure 6.1: Point process obtained from the tree



where the trees are glued together (see Remark 3.2.3). Therefore in these cases the time level 0 can, in the rescaled setting, contain a positive mass, different to any other fixed level.

In the next section we will apply this definition to the catalytic branching setting.

6.2 The catalyst and reactant point process

For the whole section let a fixed $t > 0$ be given. To define the catalyst point process we remember that the catalyst genealogy with ordering of the individuals was described either in the contour process (see Remark 5.2.2) or in Definition 5.1.5 by the ordered tree $\tilde{\eta}$. We take the second description and define the catalyst point process by:

Definition 6.2.1 (Catalyst point process)**:**
The catalyst point process is an integer-valued random measure Π^t *on*

$$\{1, 2, \ldots, \eta_t^{\text{tot}} - 1\} \times [0, t), \tag{6.3}$$

given by

$$\Pi^t := \mathcal{P}^t(\tilde{\eta}^1; 1). \tag{6.4}$$

The same is done for the reactant process where we also make use of Definition 5.1.5, where $\tilde{\xi}^1$ is introduced. Hence we define the following:

Definition 6.2.2 (Reactant point process)**:**
The reactant point process is an integer-valued random measure Ξ^t *on*

$$\{1, 2, \ldots, \xi_t^{\text{tot}} - 1\} \times [0, t), \tag{6.5}$$

given by

$$\Xi^t := \mathcal{P}^t(\tilde{\xi}^1; 1). \tag{6.6}$$

As already mentioned we also could have extracted the genealogy from the contour process. The contour will be quite helpful for the proofs we are going to give. The essential connection between MRCAs and the contour is given by the following observation, which is put into a remark:

Remark 6.2.3:
For a fixed catalyst η^{tot} *and* $t < T^{1,0}$*, i.e. less than the catalyst extinction time, let the reactant contour process* $(C_u^1)_{u \geq 0}$ *be given as in Definition 5.2.1. The reactant*

individuals alive at time t are corresponding to the starting points and endpoints of a downward excursion of the contour from level t. The set of MRCAs of these extant individuals are corresponding to the infima of the downward excursions of C^1 from level t. See Figure 8.1 for more details. ◇

For the reactant point process we can describe the distribution of the most recent common ancestors by the following lemma:

Lemma 6.2.4: *For a fixed catalyst $(\eta_s^{\text{tot}})_{s\geq 0}$ realization and fixed $t < T^{1,0}$ the reactant point process Ξ^t has total mass*

$$\Xi^t(\mathbb{N} \times [0,t)) = \xi_t^{\text{tot}} - 1. \tag{6.7}$$

The point process Ξ^t is given by the random points $\{(i,\sigma_i) : 1 \leq i \leq \xi_t^{\text{tot}} - 1\}$, where the σ_i are independent and identically distributed $[0,t)$-valued random-variables. They have distribution given by

$$P(\sigma_1 \geq h) = \frac{\frac{2}{b} + \int_0^t \eta_s^{\text{tot}}\, ds}{\int_0^t \eta_s^{\text{tot}}\, ds} \frac{\int_h^t \eta_s^{\text{tot}}\, ds}{\frac{2}{b} + \int_h^t \eta_s^{\text{tot}}\, ds}, \tag{6.8}$$

for every $0 \leq h < t$.

The idea is to rescale the point processes from the beginning of this section to obtain the rescaled catalyst and reactant point process. This rescaling will be consistent with the previous rescaling procedures.

Think of the rescaled reactant forest $\xi^{\text{for},n}$ and its contour C^n. For large n the contour looks like a diffusion process. For a fixed time level t the number of individuals alive at time t is $n\xi_t^{\text{tot},n} - 1$ and hence the number of MRCAs diverges with n.

Keeping the distance ν in the point process equal to 1 would result in the information diverging to the right. Only low-linear-order information would stay available. We will do a spatial rescaling and set $\nu = \frac{1}{n}$. This will be the right choice to keep information available and to make the limit process a σ-finite point process.

The interesting fact that we might also have mass at time level 0 was already mentioned in the first section of this chapter. Therefore we give the following definition:

Definition 6.2.5 (Rescaled Catalyst point process)**:**
The catalyst point process $\Pi^{t,n}$ is an integer-valued random measure on

$$[0, \eta_t^{\text{tot},n} - \tfrac{1}{n}] \times [0,t), \tag{6.9}$$

given by

$$\Pi^{t,n} := \mathcal{P}^t(\tilde{\eta}^n; \frac{1}{n}). \tag{6.10}$$

And the rescaled reactant is given by the following definition:

Definition 6.2.6 (Rescaled Reactant point process)**:**
The catalyst point process $\Xi^{t,n}$ is an integer-valued random measure on

$$[0, \xi_t^{\text{tot}n} - \tfrac{1}{n}] \times [0,t), \tag{6.11}$$

given by

$$\Xi^{t,n} := \mathcal{P}^t(\tilde{\xi}^n; 1/n). \tag{6.12}$$

For further purposes we will define the killing time of the reactant:

$$R^{n,0} := \inf\{t \geq 0 : \xi_t^{\mathrm{tot},n} = 0\}. \tag{6.13}$$

This time will be helpful in the oncoming considerations, since after this time, there are no reactant individuals alive any more and the reactant point process degenerates to a point process without points.

We give now a quenched result describing the distribution of the reactant point process for a given fixed catalyst. The law for a random catalyst can then be obtained by mixing binomially distributed processes. Therefore let a $t > 0$ and a fixed catalyst path $\eta^{\mathrm{tot},n}$ (and with this a fixed catalyst extinction time) be given. Condition the reactant point process on the event $\{t < T^{n,0} \wedge R^{n,0}\}$ and on the reactant total mass at time t: $\xi_t^{\mathrm{tot},n}$. Then the following result holds:

Proposition 6.2.7: *We can specify the distribution of the reactant point process $\Xi^{t,n}$ at time t. For $k \in \{1, 2, \ldots, n\xi_t^{\mathrm{tot},n} - 1\}$*

(i) the number of points at level 0 is given by

$$\kappa_n := \Xi^{t,n}(\{\tfrac{1}{n}, \tfrac{2}{n}, \ldots, \tfrac{k_n}{n}\} \times \{0\}) \stackrel{d}{=} \mathrm{Bin}\ (k, \mathbb{P}(\sigma_n \geq t)) \ \text{and} \tag{6.14}$$

(ii) the number of points between 0 and $t - h$ is given by

$$\Xi^{t,n}(\{\tfrac{1}{n}, \tfrac{2}{n}, \ldots, \tfrac{k}{n}\} \times (0, t-h)) \stackrel{d}{=} \mathrm{Bin}\ (k - \kappa_n, \mathbb{P}(\sigma_n \geq h | \sigma_n < t)). \tag{6.15}$$

Here $\mathrm{Bin}(n, p)$ is the law of a binomially distributed random variable with parameters n, p and σ_n is the extinction time of a birth-and-death process with reproduction and death rate $(\frac{nb}{2}\eta_{t-s}^{\mathrm{tot},n})_{0 \leq s \leq t}$.

This proposition prepares the limit theorem of the point process. We will give a remark to explain the previous proposition a bit more:

Remark 6.2.8:
- *Since we are starting with n individuals there arise two possibilities: The first line represents branch points at level 0, that means separated trees. The second line are death points of MRCAs within one tree.*
- *The probabilities denoted by $\mathbb{P}$ can be calculated more explicitly by the help of Lemma 6.2.4 adjusting η^{tot} to $n\eta^{\mathrm{tot},n}$.*

◇

6.3 Convergence of the reactant point process

In this section the limit point process of the reactant is calculated. We will therefore fix $t > 0$ throughout this section, but we will also require some restrictions on t, explained later. The limit point process will be a point process on $[0, Y_t] \times [0, t)$, which consists of two different types of domains as in the proposition just before:

- the line at time level zero representing the division into different trees and
- the branch points within the trees.

We first give a quenched result with the following prerequisites. Let ρ^0 be the reactant extinction time:

$$\rho^0 = \inf\{t > 0 : Y_t = 0\}. \tag{6.16}$$

Let $t_n, t > 0$ and assume catalyst paths $\eta^{\text{tot},n}, X$, s.t.:

$$t_n \to t \text{ as } n \to \infty, \tag{6.17}$$

$$\sup_{0 \le s \le t \wedge t_n} |\eta_s^{\text{tot},n} - X_s| \to 0 \text{ as } n \to \infty, \tag{6.18}$$

be given (see Definition 2.3.3 and (4.3)). Then conditioned on the event $\{t_n < R^{n,0} \wedge T^{n,0}\}$ and $t < \rho^0 \wedge \tau^0$, the following holds:

Theorem 6.3.1:
The point process $\Xi^{t_n,n}$ converges to a point process π^t on $[0, Y_t] \times [0, t)$. The distribution of π^t is given by:

$$\pi^t([0, uY_t] \times \{0\}) = \text{Poisson}\left(\frac{2uY_t}{\int_0^t bX_s\, ds}\right), \tag{6.19}$$

$$\pi^t([0, uY_t] \times (0, h)) = \text{Poisson}\left(uY_t\left(\frac{2}{\int_h^t bX_s\, ds} - \frac{2}{\int_0^t bX_s\, ds}\right)\right), \tag{6.20}$$

for $u \in [0, 1]$ and $0 < h < t$.

Remark 6.3.2:

- *The theorem shows us that for a positive time there are only finitely many trees surviving in the forest. This is because the number of MRCAs at time level zero in* (6.19) is a Poisson number and therefore almost surely finite.
- This reactant limit point process is σ-discrete as defined in [RY91, Chapter XII.1]. Only finitely MRCAs are below time level $t - h'$ for fixed time h'. That means almost all of the branch points are very close to the top of the cut tree.
- The faster the reactant branches, i.e. the greater b is, the less points can be expected far from the top-level, i.e. far below level t.
- In the second line consider t close to the catalyst extinction time τ^0 and h close to t. Then the first summand gets very big and we have the situation that many points lie below h. This is true, since an almost killed catalyst makes the reactant rather lazy to branch and its branches get long. Hence the MRCAs are far below level t.

◇

One would expect that there is a strong link between this limit point process π and the limit contour process ζ^δ (see Remark 6.2.3). Before we start with describing that link we give some definitions from excursion theory (see [RY91, Chapter XII] or [RW79, Chapter VI]):

For a fixed catalyst X let ζ^δ be the limit contour process. It is given as the solution of the martingale problem A^δ as in Theorem 5.3.1. Fix a level $t < \tau^\delta$.

Let $(l_u^t(\zeta^\delta))_{u\geq 0}$ denote its local time process at level t. The inverse of the local time is defined as:

$$\alpha_u^t(\zeta^\delta) := (l_\cdot^t(\zeta^\delta))^{-1}(u) = \inf\{r > 0 : l_r^t(\zeta^\delta) > u\}, \tag{6.21}$$

where we sometimes omit the relation to ζ^δ, if no confusion arises. To describe an excursion we denote the set of downward excursions from level t by:

$$U_t^- := \{f : [0,\infty) \to (-\infty, t] : f(0) = t \text{ and } f(a) = t \Rightarrow f(a+s) = t \; \forall s > 0\} \cup \{\Delta\}, \tag{6.22}$$

where Δ denotes the *no excursion* state. The downward excursion ϵ_u^- from level t (at local time u) is given as an element of U_t^- by:

$$\epsilon_u^-(s) := \begin{cases} \zeta^\delta_{(\alpha^t_{u-}+s)\wedge\alpha^t_u} & \text{if } \alpha^t_{u-} \neq \alpha^t_u \\ \Delta & \text{if } \alpha^t_{u-} = \alpha^t_u \end{cases}. \tag{6.23}$$

Additionally we set

$$\alpha_{4/b} := \alpha^0_{4/b}(\zeta^\delta), \tag{6.24}$$

as the time, when local time at level 0 reaches $4/b$ for the first time, or speaking about populations, when the total initial mass of the population has reached 1. This rather awkward $4/b$ comes from the quadratic variation term of the limit contour which is needed to calculate local time (consult (5.9) and Section 9.2.

The point process of maximal depths of downward excursions is then defined as

$$\pi^{\zeta^\delta,t} := \{(u, \inf(\epsilon_u^-) :, \text{ when } \alpha^t_{u-} \neq \alpha^t_u \text{ and } u \leq \alpha_{4/b}\}. \tag{6.25}$$

For this construction the following relation between contour and point process holds:

Proposition 6.3.3 (Quenched contour and point process)**:** *Let a fixed catalyst X and $t < \tau^0$ be given. Choose $\delta > 0$ such that $t < \tau^\delta$ and let ζ^δ denote the solution of the $(A^\delta, \mathbb{D}(A^\delta))$ martingale problem. Then it holds that*

$$\pi^t \overset{d}{=} \pi^{\zeta^\delta,t}. \tag{6.26}$$

Be aware that the process π^t measures the time in Lebesgue measure, whereas the process on the right hand side is related to increasing local time of the contour.

After this quenched result the question arises if also an annealed result holds. One might think of achieving that for a special class of branching modifications g, where the proof can be comparable to the one in Section 8.6. But we are not intending a proof of this result.

Part II

Proofs

7 Proofs of the main results from Chapter 2

In this chapter we put together the main proofs from Chapter 2 and try to lead the reader through the structure of the problems. Some of the theorems were put in the appendix.

7.1 Remarks and techniques

7.1.1 Properties of g

We will try to justify the conditions set on g on page 14:

- Continuity makes sense since the branching behaviour should not change significantly if there is a slight change in the population size.
- Lipschitz-continuity is necessary to receive solutions for the catalyst SDE in Section 2.2.
- Vanishing of g in 0 expresses the fact that branching stops, when all individuals have died out, since there is nobody left to branch.
- The third condition G3 is required to be in a "sub-Anderson-model", since otherwise the branching behaviour cannot be controlled easily and the diffusion would not die out.
- The last condition makes it possible that 0 is an exit boundary and we can obtain Lemma 2.1.7 and so we do not get into the following situation:

  > The catalyst is close to zero but not dying out. Then the reactant faces almost no more branching events and therefore does not change.

  As later want to focus on contour processes and trees, we want to avoid that situation.

Let us also note that Lipschitz-continuity in 0 and (G4) imply that g must approach zero faster than "$x \mapsto x$", but slower than "$x \mapsto x^2$".

Remark 7.1.1:
In some proof we will not need to use condition G2 from page 14, but a weaker requirement on g:

$$g(x) \leq C(1 + x^2) \ \forall x \in [0, \infty) \tag{7.1}$$

which is certainly true, probably for a different constant C. ◇

For the further analysis it will be helpful to extend the definition of g to negative real numbers. There we will set $g = 0$. This assumption does neither violate the Lipschitz-continuity nor the estimates just presented. In fact this extension is only of theoretical use, since it will allow us to use some theorems which were designed for problems defined on $\mathbb{R}$ rather than on $\mathbb{R}_+$, e.g. in Chapter 2.2 dealing with diffusions.

7.1.2 The generator of the discrete total mass process

We refer to the definition of the discrete total mass process $(\eta^{\text{tot},n}, \xi^{\text{tot},n})$ in Definition 2.1.4 for fixed n and want to explain how to derive a generator U_n for this Markov jump process. After having done this we will see that this generator U_n is uniquely closable in $C_0(\mathbb{R}_+^2, \mathbb{R})$ and there is a unique process $(\eta^{\text{tot},n}, \xi^{\text{tot},n})$ corresponding to the generator for a given initial distribution. So we are invited to speak of *the* discrete catalytic branching process.

First we extract the infinitesimal generator U_n of the total mass functional from the Definitions 1.2.4 and 1.2.5 (for a fine introduction see the book of Breiman [Bre68, page 332]). From then on we will understand the generator U_n as the preferable description of the process.

If the process $(\eta_t^{\text{tot},n}, \xi_t^{\text{tot},n})_{t\geq 0}$ starts in $\eta_0^{\text{tot},n} = x, \xi_0^{\text{tot},n} = y$, where x and y are multiples of $1/n$ and greater than zero.

Each catalyst individual (and there are nx of them dies after an exponential time with rate $\frac{ng(x)}{x}$. Just after that zero or two individuals with mass $\frac{1}{n}$ each are born. So either the total decreases by $\frac{1}{n}$ or it increases by $\frac{1}{n}$, each with probability $\frac{1}{2}$. Hence after an infinitesimal time step Δt we expect the following situation:

$(\eta^{\text{tot},n}, \xi^{\text{tot},n})$	with probability
$(x + \frac{1}{n}, y)$	$\frac{n^2}{2} g(x)\Delta t + \mathcal{O}(\Delta t)$
$(x - \frac{1}{n}, y)$	$\frac{n^2}{2} g(x)\Delta t$
$(x, y + \frac{1}{n})$	$\frac{n^2}{2} bxy\Delta t$
$(x, y - \frac{1}{n})$	$\frac{n^2}{2} bxy\Delta t$
(x, y)	$1 - n^2 g(x)\Delta t - n^2 bxy\Delta t$

So the generator U_n acts on a given function $f \in C_0(\mathbb{R}_+^2, \mathbb{R})$ as follows:

$$
\begin{aligned}
U_n f(x,y) &= \lim_{\Delta t \to 0} \frac{1}{\Delta t} E[f(\eta_{\Delta t}^{\text{tot},n}, \xi_{\Delta t}^{\text{tot},n}) - f(x,y)] = \\
&= \lim_{\Delta t \to 0} \frac{1}{\Delta t} \Big[f(x + \frac{1}{n}, y) \frac{n^2}{2} g(x)\Delta t + f(x - \frac{1}{n}, y) \frac{n^2}{2} g(x)\Delta t \\
&\quad + f(x, y + \frac{1}{n}) \frac{n^2}{2} bxy\Delta t + f(x, y - \frac{1}{n}) \frac{n^2}{2} bxy\Delta t \\
&\quad + f(x,y)(1 - n^2 g(x)\Delta t - n^2 bxy\Delta t) - f(x,y) \Big] = \\
&= \frac{n^2}{2} \Big[g(x) \Big(f(x + \frac{1}{n}, y) - 2f(x,y) + f(x - \frac{1}{n}, y) \Big) \\
&\quad + bxy \Big(f(x, y + \frac{1}{n}) - 2f(x,y) + f(x, y - \frac{1}{n}) \Big) \Big].
\end{aligned}
\tag{7.2}
$$

We will need to be a bit more precise about the aforementioned to write down the exact generator. What happens if we already start with $y = 0$? Then we cannot go

to $\xi^{\text{tot},n}_{\Delta t} = -\frac{1}{n}$. If one considers all the problematic cases we get

$$\begin{aligned} U_n f(x,y) =& \frac{n^2}{2} g(x)[f(x+\frac{1}{n},y) - 2f(x,y) + f(x-\frac{1}{n},y)], \\ &+ \frac{n^2}{2} bxy[f(x,y+\frac{1}{n}) - 2f(x,y) + f(x,y-\frac{1}{n})], \text{ for } x,y \geq \frac{1}{n}, \\ U_n f(0,y) =& 0 \text{ for } y \geq 0, \\ U_n f(x,0) =& \frac{n^2}{2} g(x)[f(x+\frac{1}{n},0) - 2f(x,0) + f(x-\frac{1}{n},0)] \text{ for } x \geq \frac{1}{n}, \\ \mathbb{D}(U_n) =& C_0(\mathbb{R}^2_+, \mathbb{R}). \end{aligned} \tag{7.3}$$

Now with the generator U_n at hand it is possible to show that, for g sufficing the condition $g(x) \leq C(1+x^2)$, from (7.1), there is only one process satisfying Definition 2.1.4. The other conditions on g do clearly not play a role, since they are *local*, but here we are dealing with a process on a discrete grid $\frac{1}{n}\mathbb{N}^2$.

Within the next section we will tackle this uniqueness problem, but will use an arbitrary mass m, since this makes the arguments and the looking clearer.

7.2 Existence, Uniqueness and Feller-property of $(\eta^{\text{tot},n}, \xi^{\text{tot},n})$

We start with the proof of Lemma 2.1.6. The proof will be done without giving special attention to the parameter n, but it will be done with an arbitrary mass m of any individual. To get the factor n^2 in U_n just multiply the oncoming μ with n^2.

Lemma 7.2.1: *[Existence, Uniqueness and Feller-property of $(\eta^{\text{tot},n}, \xi^{\text{tot},n})$] For any $g : [0,\infty) \to [0,\infty)$ satisfying Condition 2.1.3 in Section 2.2 the closure of generator U_n from Section 2.2 is single-valued and generates a Feller-semigroup on $C_0(\mathbb{R}^2)$. Moreover $C_c(\mathbb{R}^2)$ is a core for this generator. So there exists $(\eta^{\text{tot}}, \xi^{\text{tot}})$, uniquely determined by U_n and this process is a Feller-process.*

PROOF: The proof relies on Theorem 8.3.1 in [EK86], which can be found in the appendix. As function λ and measure μ we use:

$$\begin{aligned} \lambda(x,y) &= g(x) + xy, \\ \mu((x,y),(\tilde{x},\tilde{y})) &= \begin{cases} \frac{g(x)}{2\lambda(x,y)} & for \quad |x-\tilde{x}| = m, y = \tilde{y} \\ \frac{bxy}{2\lambda(x,y)} & for \quad x = \tilde{x}, |y-\tilde{y}| = m \end{cases}, \end{aligned} \tag{7.4}$$

where $x, \tilde{x}, y, \tilde{y} \geq 0$. Then

$$U_n f(x,y) = n^2 \sum_{\tilde{x},\tilde{y}} f(\tilde{x},\tilde{y}) \lambda(x,y) \mu((x,y),(\tilde{x},\tilde{y})). \tag{7.5}$$

In the case of the catalyst being 0, the reactant being y, the measure $\mu((0,y),\cdot)$ is concentrated on this point.

The state space $E = m\mathbb{N}^2_+$ is a locally compact, noncompact, separable metric space. The one-point-compactification of E will be denoted by $E^\Delta = E \cup \Delta$, where Δ is the point at infinity. The real-valued function λ is nonnegative and continuous on $\mathbb{N}^2_+$ and μ is continuous in the first coordinate pair, since the preimage space

is discrete. We have to define some functions γ and η. There should not be any confusion with the η which is the catalyst population during the rest of this work. We set γ and η as functions which are positive, continuous and the inverse vanishing at infinity:

$$\gamma(x,y) = \eta(x,y) = x^2 + y^2 + 1. \tag{7.6}$$

Now we have to check the properties (3.2) to (3.5) from [EK86, Theorem 8.3.1]. We start with condition (3.2) of [EK86, Theorem 8.3.1]:

$$\begin{aligned}\sup_{x,y\geq 0} \frac{\lambda(x,y)}{\gamma(x,y)} &= \sup_{x,y\geq 0} \frac{g(x) + bxy}{x^2+y^2+1} \\ &\leq \sup_{x,y\geq K} \frac{C(1+x^2)}{x^2+y^2+1} + \sup_{x,y\geq 0} \frac{bxy}{x^2+y^2+1} \\ &\leq C + \sup_{x,y\geq K} \frac{Cx^2}{x^2} + \frac{b}{2} = 2C + \frac{b}{2} = C_1 < \infty.\end{aligned} \tag{7.7}$$

So the first condition is fulfilled. Next we consider condition (3.3) of [EK86, Theorem 8.3.1]:

$$\lim_{(x,y)\to\Delta} \lambda(x,y)\mu((x,y),K) = 0 \text{ for every compact } K \subset \mathbb{R}^2_+. \tag{7.8}$$

This is certainly true since $\mu((x,y),\cdot)$ only has positive measure at the neighbouring points having distance not more than m from (x,y) Then for diverging (x,y) there will be no more such neighbour point in any compact K. So the μ-factor vanishes.

Conditions (3.4) and (3.5) of [EK86, Theorem 8.3.1] are basically similar in their proof. For both cases we will only check values of $x, y \geq m$, since the other cases are even easier to verify and would make the proof only lengthier. Indeed it makes sense to check them separately. Let us start with (3.4), keeping in mind that we only wanted to allow $m \leq 1$ and $g(x) \leq C(1+x^2)$ (G3):

$$\begin{aligned}&\sup_{x,y\geq m} \lambda(x,y) \Big| \int \frac{\gamma(x,y) - \gamma(\bar{x},\bar{y})}{\gamma(\bar{x},\bar{y})} \mu((x,y), d(\bar{x},\bar{y})) \Big| = \\ &\quad = \sup_{x,y\geq m} \frac{g(x)}{2} \frac{x^2+y^2+1 - \big((x+m)^2+y^2+1\big)}{(x+m)^2+y^2+1} \\ &\qquad + \frac{g(x)}{2} \frac{x^2+y^2+1 - \big((x-m)^2+y^2+1\big)}{(x-m)^2+y^2+1} \\ &\qquad + \frac{bxy}{2} \frac{x^2+y^2+1 - \big(x^2+(y+m)^2+1\big)}{x^2+(y+m)^2+1} \\ &\qquad + \frac{bxy}{2} \frac{x^2+y^2+1 - \big(x^2+(y-m)^2+1\big)}{x^2+(y-m)^2+1} \\ &\quad \leq \sup_{x,y\geq m} \frac{g(x)}{2} \frac{m^2}{(x-m)^2+y^2+1} + \frac{bxy}{2} \frac{m^2}{x^2+(y-m)^2+1}\end{aligned} \tag{7.9}$$

$$\begin{aligned}
&\leq \sup_{x,y\geq m} \frac{g(x)}{(x-m)^2+y^2+1} + \frac{bxy}{x^2+(y-m)^2+1} \\
&\leq \sup_{x,y\geq m} C\frac{1+x^2}{(x-m)^2+y^2+1} + b\frac{x^2+y^2}{x^2+(y-m)^2+1} \\
&\leq \sup_{x\geq m} C\frac{1+x^2}{(x-m)^2+1} + \sup_{x\geq m}\frac{bx^2}{x^2+1} + \sup_{y\geq m}\frac{by^2}{(y-m)^2+1} \qquad (7.10)\\
&\leq C + \sup_{x\geq m}\frac{(C+b)x^2}{(x-m)^2+1} + \sup_{y\geq m}\frac{by^2}{(y-m)^2+1} \\
&\leq C + (C+2b)\sup_{x\geq m}\frac{x^2}{(x-m)^2+1} < \infty,
\end{aligned}$$

where the last inequality is easily believed since the denominator stays strictly away from 0. This is the desired result.

Now at last we have to show (3.5):

$$\sup_{x,y\geq 0} \lambda(x,y)\int \frac{\gamma(x,y)-\gamma(\bar{x},\bar{y})}{\gamma(x,y)}\mu((x,y),d(\bar{x},\bar{y})) < \infty. \qquad (7.11)$$

But this can be shown the same way as (3.4) with even less problems in the denominator.

□

7.3 Extinction and compact containment condition of $(\eta^{\text{tot},n}, \xi^{\text{tot},n})$

As a next result we show Lemma 2.1.7:

Lemma 7.3.1: *Every (discrete) catalyst process, defined as in (2.1) with g satisfying the usual conditions dies out almost sure, i.e.*

$$T^{n,0} = \inf\{t \geq 0 : \eta_t^{\text{tot},n} = 0\} < \infty \ a.s. \qquad (7.12)$$

Furthermore 0 is an exit boundary (= absorbing point), i.e. the catalyst does not leave 0 after having reached it.

PROOF: We will consider the situation for a fixed n and do a proof by contradiction.

Clearly the discrete catalyst $\eta^{\text{tot},n}$ is a non-negative martingale, so there exists a limit process $\eta_\infty^{\text{tot},n}$ s.t. $\eta^{\text{tot},n}$ converges almost sure to this process. That means that for a given $\epsilon > 0$ there exists a $T > 0$ with

$$\mathbf{P}[\,|\eta_t^{\text{tot},n} - \eta_\infty^{\text{tot},n}| \leq \epsilon \ \forall t \geq T] \geq 1-\epsilon. \qquad (7.13)$$

If we select $\epsilon < \frac{1}{n}$, then the event in consideration means that no more jumps occur after time T, except probably on a set of measure ϵ.

We assume that for given $0 < a < b$ the induced probability measure on $[0,\infty)$ of the limit process is given as:

$$\mathbf{P} \circ (\eta_\infty^{\text{tot},n})^{-1}([a,b]) = r, \text{ with } 0 \leq r \leq 1.$$

Let us now assume that

$$2\epsilon < r$$

and we will lead this to a contradiction for any ϵ.

Then by (7.13) we get that there must be a set A in Ω^n with measure at least $r-\epsilon$, s.t. for $\omega \in A$ the catalyst $\eta^{\text{tot},n}(\omega)$ faces no more branching events after time $T(\epsilon)$ and

$$\eta_t^{\text{tot},n}(\omega) \in [a,b] \ \forall t \geq T(\epsilon).$$

But the probability of no branching event from time $T(\epsilon)$ to $T(\epsilon)+t$ is bounded from above by $\max_{x\in[a,b]} e^{-g(x)t}$. If t tends to infinity, the no-branching probability $\mathbf{P}(A)$ goes to zero, as g stays away from zero in any interval $[a,b]$ not containing 0 (G1 and G2). This contradicts that it would be more than $r-\epsilon$. So the assumptionthat $\mathbf{P} \circ (\eta_\infty^{\text{tot},n})^{-1}([a,b]) = r$ cannot hold for any positive r and we have that $\mathbf{P} \circ (\eta_\infty^{\text{tot},n})^{-1}([a,b]) = 0$ for any positive a and b. So all the mass of $\eta_\infty^{\text{tot},n}$ is concentrated in zero.

Additionally we have shown that $T^{n,0}$ is almost surely finite, since by (7.13) $\mathbf{P}[T^{n,0} \leq t] = \mathbf{P}[\eta_t^{\text{tot},n} = 0] > 1-\epsilon$ if we choose $t \geq T(\epsilon)$ and $\epsilon < \frac{1}{n}$.

Clearly 0 is absorbing, since $g(0) = 0$. □

Remark 7.3.2:
It is well known that in discrete time the population gets extinct depending on the expectation of the offspring distribution. This also could have been a way to prove the previous lemma.

Another result will describe how far the discrete total mass process moves away from the origin when starting in $(\eta_0^{\text{tot},n}, \xi_0^{\text{tot},n})$. As an additional result we get the *compact containment condition* for the set of processes $(\eta^{\text{tot},n}, \xi^{\text{tot},n})_{n\in\mathbb{N}}$, which states this set moves within a compactum with a probability close to 1 up to a given time t. Therefore let $\|\cdot\|_2$ denote the Euclidean norm on $\mathbb{R}^2$.

Lemma 7.3.3 (Compact containment condition)**:** *Let $(\eta_t^{\text{tot},n}, \xi_t^{\text{tot},n})_{t\geq 0}$ be a process corresponding to the solution of the discrete martingale problem $(U_n, \delta_{\eta_0^{\text{tot},n}} \otimes \delta_{\xi_0^{\text{tot},n}})$. Then for $T > 0, k > 0$, it is true that:*

$$\mathbf{P}\left[\sup_{0\leq t\leq T} \|(\eta_t^{\text{tot},n}, \xi_t^{\text{tot},n})\|_2 > k\right] \leq \frac{\sqrt{2}}{k}(\eta_0^{\text{tot},n} + \xi_0^{\text{tot},n}). \tag{7.14}$$

Additionally the compact containment condition holds:
When $\eta_0^{\text{tot},n} = \xi_0^{\text{tot},n} = 1$ for all n, then for any $\lambda > 0$ and $t > 0$ there exists a $\Gamma_{\lambda,T} \subset \mathbb{R}_+^2$ such that:

$$\inf_{n\in\mathbb{N}} \mathbf{P}\left[(\eta_t^{\text{tot},n}, \xi_t^{\text{tot},n}) \in \Gamma_{\lambda,T} \ \forall\, 0 \leq t \leq T\right] \geq 1-\lambda. \tag{7.15}$$

Proof: Since $\eta^{\text{tot},n}$ and $\xi^{\text{tot},n}$ are martingales by Lemma 7.2.1, the maximum

inequality holds and expectation values are conserved.

$$\begin{aligned}
&\mathbf{P}\left[\sup_{0\le t\le T}\|(\eta_t^{\mathrm{tot},n},\xi_t^{\mathrm{tot},n})\|_2 > k\right] \\
&\le \mathbf{P}\left[\sup_{0\le t\le T}|\eta_t^{\mathrm{tot},n}| > \frac{k}{\sqrt{2}}\right] + \mathbf{P}\left[\sup_{0\le t\le T}|\xi_t^{\mathrm{tot},n}| > \frac{k}{\sqrt{2}}\right] \le \\
&\le \frac{\sqrt{2}}{k}\left(\mathbf{E}[\eta_T^{\mathrm{tot},n}] + \mathbf{E}[\xi_T^{\mathrm{tot},n}]\right) = \frac{\sqrt{2}}{k}(\eta_0^{\mathrm{tot},n} + \xi_0^{\mathrm{tot},n}).
\end{aligned} \tag{7.16}$$

The second statement is just a reformulation. □

7.4 Existence and Uniqueness of (X, Y)

7.4.1 The main result and the strategy of the proof

Within this section a proof of the main result of section 2.2 is given. This theorem stated that

Theorem 7.4.1:
If g satisfies Condition 2.1.3 the SDE system

$$dX_t = \sqrt{g(X_t)}\, dW_t^1, \tag{7.17}$$
$$dY_t = \sqrt{bX_tY_t}\, dW_t^2, \tag{7.18}$$

where W^1 and W^2 are two independent Brownian motions has a unique strong solution.

Here the catalyst-reactant-type of the problem helps obtaining a proof for this theorem. In a first step we will show that there exists a strong unique solution X for equation (7.17), since g is locally Lipschitz-continuous. That means for any Brownian motion path $W^1(\omega)$ on a probability space $(\Omega', \mathcal{A}', \mathbf{P}')$ there exists a path $X(\omega)$ which solves the catalyst's SDE for this specific $\omega \in \Omega'$. After that we show in the second step that the solutions of the catalyst die out after almost sure finite time. Then in the last step we take a fixed catalyst path $X(\omega)$ and consider (7.18) for this single path. This fixed reactant SDE has a strong unique solution and altogether we get that whenever starting with two independent Brownian motions we can construct a solution (X, Y) of the SDE system.
The proof of this theorem is split up in three parts. We will put the proof of this theorem in the next subsection and denote the steps here in a short overview:

Step 1: The catalyst SDE has a unique strong solution.

Step 2: The catalyst dies out after almost sure finite time.

Step 3: The reactant SDE for fixed catalyst has a unique strong solution.

7.4.2 The proof

Step 1: *The catalyst SDE has a unique strong solution*
Let us start with a lemma about the catalyst:

Lemma 7.4.2 (The catalyst SDE has a unique strong solution)**:** *The SDE* $dX_t = \sqrt{g(X_t)}\, dW_t^1$ *has a unique strong solution, for* g *satisfying Condition 2.1.3.*

Remark 7.4.3:
The theorems used for the proof in this first step are due to Engelbert and Schmidt and are well presented in [KS00]. Their approach to one-dimensional problems allows deeper results than the multi-dimensional theory via Lipschitz-continuity. The latter is presented in Theorem 5.2.1 in the SDE book [Øks05] of Øksendal. They cannot be used as there was not postulated a global Lipschitz-continuity for g. We did not postulate that, since we also wanted to treat cases where g grows faster than linear. ◇

PROOF: This proof relies on Corollary 5.10 of [KS00], which says that there exists a strong unique solution to $dX_t = \sigma(X_t)dW_t$, if the four conditions (E) and (i)-(iii) hold for functions $f : \mathbb{R} \to [0, \infty]$ and $h : [0, \infty] \to [0, \infty]$:

(E)

$$I(\sigma) \subseteq Z(\sigma), i.e. : \{x \in \mathbb{R} : \exists \epsilon > 0 \text{ s.t. } \int_{x-\epsilon}^{x+\epsilon} \frac{dy}{\sigma^2(y)} = \infty\} \subseteq \{x \in \mathbb{R} : \sigma(x) = 0\} \tag{7.19}$$

(i) at every $x \in I(\sigma)^c$, the quotient $(f/\sigma)^2$ is locally integrable; i.e., there exists $\epsilon > 0$ (depending on x) such that

$$\int_{x-\epsilon}^{x+\epsilon} \left(\frac{f(y)}{\sigma(y)}\right)^2 dy < \infty; \tag{7.20}$$

(ii) the function h is strictly increasing and satisfies $h(0) = 0$ and

$$\int_0^\epsilon h^{-2}(u)du = \infty; \ \forall \epsilon > 0 \tag{7.21}$$

(iii) there exists a constant $a > 0$ such that

$$| \sigma(x + y) - \sigma(x) | \leq f(x)h(| y |); \quad \forall x \in \mathbb{R}, y \in [-a, a]. \tag{7.22}$$

Assertion (E) is relatively easy to prove: Since g is continuous, $g(x)$ must be zero, if $x \in I(\sigma)$. The remaining three points will be proven in reverse order, since it seems more intuitive: As g is locally Lipschitz-continuous, for any compactum K there will be a constant L_K, s.t.:

$$| g(x) - g(y) | \leq L_K | x - y |; \quad \forall x, y \in K. \tag{7.23}$$

See the Appendix B.4 for a proof. Set $K(x) := [x - a, x + a]$. Then $L_{K(x)}$ is finite for any x and a since the interval is compact.
Next define $f(x) = L_{K(x)}^{1/2}$ and we have:

$$| \sqrt{g(x + y)} - \sqrt{g(x)} |^2 \leq | g(x + y) - g(x) | \leq L_{K(x)} | y | \ \ \forall y \in [-a, a]. \tag{7.24}$$

But this is the squared version of (iii) for $h(u) = u^{1/2}$. This function h also satisfies condition (ii) since $u \mapsto 1/u$ is not integrable in a neighbourhood of zero and the

other conditions being clearly satisfied.
It remains to show (i). But the integral will be finite if one chooses $x - \epsilon > 0$, since the denominator stays strictly away from zero and the nominator is bounded from above by L_K for $K = [x - \epsilon - a, x + \epsilon + a]$. □

STEP 2: *The catalyst dies out after almost sure finite time*

In the third step we want to prove that for a given sample path $X_t(\omega) = x(t)$ of the catalyst, there is a solution to the reactant SDE. Before we can do that we need to know a bit more about the catalyst. The following lemma shows that it almost surely dies out after finite time and this will prove Lemma 2.2.4 about the extinction time of the catalyst:

Lemma 7.4.4: *Every catalyst process X with g satisfying the usual conditions dies out almost sure in finite time, i.e.*

$$\tau^0 = \inf\{t \geq 0 : X_t = 0\} < \infty \ a.s. \ , \tag{7.25}$$

and 0 is an exit boundary for X.

PROOF: We will use the notations and theorems from [RW79, Sections V.44-51]. Another fine account for boundary behavior of diffusions is given in [KT81].

Since the catalyst SDE (7.17) has no drift, X is on the natural scale (i.e. the scaling function is the identity). From equation (47.30) in [RW79] we see that the speed measure m is given by:

$$m(dx) = \frac{1}{g(x)}\, dx. \tag{7.26}$$

First we need to check that $\tau^0 < \infty \ a.s.$ by Theorem 51.2(ii) and for that purpose calculate by taking into account (G4) from page 14:

$$\begin{aligned} \int_{0+} x m(dx) = \int_{0+} x \frac{dx}{g(x)} \leq \int_0^{\delta(\epsilon)} x \frac{c' + \epsilon}{x^{1+\beta}}\, dx = \\ = (c' + \epsilon) \left[\frac{x^{1-\beta}}{1-\beta} \right]_0^{\delta(\epsilon)} = \frac{c' + \epsilon}{1-\beta} \delta^{1-\beta} < \infty, \end{aligned} \tag{7.27}$$

where one should remember that $0 \leq \beta < 1$. Thus it follows that $P^x[\tau^0 < \infty] = 1 \ \forall x > 0$.

Now let us check that 0 is an exit boundary by Definition (51.3): Since $P^x[\tau^0 < \infty] = 1 > 0$, we have that 0 is accessible. However with

$$\int_{0+} m(dx) > (c' - \epsilon) \int_0^{\delta} x^{-1-\beta}\, dx = \infty, \tag{7.28}$$

it follows that 0 is an exit boundary. □

STEP 3: *The reactant SDE for fixed catalyst has a unique strong solution*

With the lemma from the previous step at hand we will now show existence and uniqueness of the reactant SDE for a fixed catalyst realization $x(t) = X_t(\omega)$. We already know that almost surely $x(t)$ will go to zero after finite time T. For such a function the following lemma holds:

Lemma 7.4.5 (The reactant SDE has a unique strong solution)**:** *For any continuous function $x : [0,\infty) \to [0,\infty)$, with $x([T,\infty)) = \{0\}$ for a $T < \infty$ the SDE*

$$dY_t = \sqrt{bx(t)Y_t}\, dW_t^2 \tag{7.29}$$

has a unique strong solution.

Proof: Since x is a continuous function on a compactum and identically zero outside, it is bounded from above by a constant k, i.e. $x(t) < k\ \forall t \in [0,\infty)$. Then due to the continuity of $\sigma(t,y) = \sqrt{bx(t)}\sqrt{y}$ in both coordinates there exists a weak solution to (7.29), see e.g. [EK86, Theorem 5.3.10]. Since σ is Hölder-1/2-continuous in the second coordinate, pathwise uniqueness follows by a theorem of Yamada and Watanabe for one-dimensional SDEs [EK86, Remark 5.3.9]. But pathwise uniqueness and weak existence imply existence and uniqueness of a strong solution (e.g. [Kle08, Theorem 26.18] or [KS00, Corollary 5.3.23]). □

Combining the three previous steps we have shown existence and uniqueness of a strong solution to the SDE system 7.17 and 7.18. □

7.5 Simple properties of the diffusion process

We will introduce some notation when we start the processes. If $X_0 = x > 0$ and $Y_0 = y > 0$, then we will denote the solution of the SDE (or the martingale problem) for these initial values by $(X_t^x, Y_t^{x,y})$. This makes sense, since the catalyst evolves autonomously of the reactant and so the initial value of the reactant does not influence the evolution of the catalyst. Sometimes we also might consider probabilities or expectations for given initial values. Then it sometimes is helpful not to fill the formula with two many superscripts and we will denote the probability as $\mathbf{P}_{(x,y)}$ and the expectation as $\mathbf{E}_{(x,y)}$. If not required for our considerations at all we try to avoid the sub- and superscripts for a better reading.

We begin with some expectation values of the catalyst and the reactant diffusions:

Proposition 7.5.1: *[Expectation values of the SDE system] If $(X^x, Y^{x,y})$ is a unique strong solution for the martingale problem $(A, \delta_x \times \delta_y)$, then we get the following expectation values, where γ is a constant only depending on g and α, C is as in (G2) from page 14:*

$\underline{0 < \alpha < 2}$*:*

$$\begin{aligned} \mathbf{E}\left[X_t^x\right] = x,\ \mathbf{E}\left[Y_t^{x,y}\right] = y, \\ \mathbf{E}\left[X_t^x Y_t^{x,y}\right] = xy, \end{aligned} \tag{7.30}$$

$0 \leq \alpha < 1$:

$$\mathbf{E}\left[(X_t^x)^\alpha\right] \leq 1 + x, \tag{7.31}$$
$$\mathbf{E}\left[g\left(X_t^x\right)\right] \leq C(2 + x), \tag{7.32}$$

$1 \leq \alpha < 2$:

$$\mathbf{E}\left[(X_t^x)^\alpha\right] \leq \left(x^\alpha + \frac{C+\gamma}{C}\right) e^{C\alpha(\alpha-1)t} - \frac{C+\gamma}{C}, \tag{7.33}$$
$$\mathbf{E}\left[g\left(X_t^x\right)\right] \leq \left(Cx^\alpha + C + \gamma\right) e^{C\alpha(\alpha-1)t} - \gamma. \tag{7.34}$$

PROOF: The first statements in (7.30) are clear, since X and Y are martingales. Additionally XY is also a martingale, since U does not contain any cross-derivatives of x and y.

More interesting are the other results. It is clear that (7.31) implies (7.32) and (7.33) implies (7.34), since from the condition (G2) on g (in page 14), we have:

$$\mathbf{E}\left[g\left(X_t^x\right)\right] \leq \mathbf{E}\left[C(1 + (X_t^x)^\alpha)\right] \leq C + C\,\mathbf{E}\left[(X_t^x)^\alpha\right]. \tag{7.35}$$

So now we will show (7.31)

$0 \leq \alpha < 1$: An easy calculation gives:

$$\begin{aligned}\mathbf{E}\left[(X_t^x)^\alpha\right] &= \mathbf{E}\left[(X_t^x)^\alpha \mathbb{1}_{\{X_t^x \leq 1\}} + (X_t^x)^\alpha \mathbb{1}_{\{X_t^x > 1\}}\right] \leq \\ &\leq 1 + \mathbf{E}\left[X_t^x \mathbb{1}_{\{X_t^x > 1\}}\right] \leq 1 + x.\end{aligned} \tag{7.36}$$

$1 \leq \alpha < 2$: For this case we will apply Itô's rule for the function $x \mapsto x^\alpha$ and take expectations:

$$\begin{aligned}\mathbf{E}\left[(X_t^x)^\alpha\right] =& x^\alpha + \mathbf{E}\left[\int_0^t \alpha (X_s^x)^{\alpha-1} \sqrt{g(X_s^x)}\, dW_s\right] \\ &+ \mathbf{E}\left[\int_0^t \alpha(\alpha-1)(X_s^x)^{\alpha-2} g(X_s^x)\, ds\right] = \\ =& x^\alpha + \alpha(\alpha-1)\int_0^t \mathbf{E}_x\left[\mathbb{1}_{\{X_s \leq 1\}}(X_s)^{\alpha-2} g(X_s)\right] ds \\ &+ \alpha(\alpha-1)\int_0^t \mathbf{E}_x\left[\mathbb{1}_{\{X_s > 1\}}(X_s)^{\alpha-2} g(X_s)\right] ds \leq \\ \leq& x^\alpha + \alpha(\alpha-1)\int_0^t \mathbf{E}_x\left[\mathbb{1}_{\{X_s \leq 1\}}(X_s)^{\alpha-2} \gamma(X_s)^1\right] ds \\ &+ \alpha(\alpha-1)\int_0^t \mathbf{E}_x\left[\mathbb{1}_{\{X_s > 1\}}(X_s)^{\alpha-2} C(1 + (X_s)^2)\right] ds.\end{aligned} \tag{7.37}$$

Here γ is a real number, s.t. $g(x) \leq \gamma x$ holds on $[0, 1]$. This exists by the conditions set on g on page 14:
If a is s.t. $[0, a)$ is a neighbourhood of zero, where the Lipschitz-constant L_0 in zero bounds $g(x)/x$, then we have

$$g(x) \leq L_0 x \text{ for } x \in [0, a) \tag{7.38}$$

on the other hand $g(x)/x$ is continuous on [a,1], so bounded by a constant c_0. If γ is chosen to be the maximum of L_0 and c_0/a, then we are done.

To continue with (7.37), we look at the first summand of the expectation and use the fact that the random variable was truncated to be smaller than 1. For the second summand use that $\alpha - 2 < 0$:

$$\mathbf{E}[(X_t^x)^\alpha] \leq x^\alpha + \alpha(\alpha-1)(\gamma+C)t + \alpha(\alpha-1)\int_0^t CE[(X_s^x)^\alpha]\,ds \tag{7.39}$$

This already makes you look for Gronwall's inequality, which we adapt now to get:

$$\begin{aligned}
&\mathbf{E}[(X_t^x)^\alpha] \leq x^\alpha + \alpha(\alpha-1)(\gamma+C)t \\
&\quad + C\alpha(\alpha-1)\int_0^t (x^\alpha + \alpha(\alpha-1)(\gamma+C)s)e^{C\alpha(\alpha-1)(t-s)}\,ds \\
&= x^\alpha + \alpha(\alpha-1)(\gamma+C)t + C\alpha(\alpha-1)x^\alpha e^{C\alpha(\alpha-1)t}\frac{[e^{-C\alpha(\alpha-1)s}]_0^t}{-C\alpha(\alpha-1)} \\
&\quad + C\alpha^2(\alpha-1)^2(\gamma+C)e^{C\alpha(\alpha-1)t}\int_0^t se^{-sC\alpha(\alpha-1)}\,ds = \\
&= x^\alpha + \alpha(\alpha-1)(\gamma+C)t + x^\alpha(e^{C\alpha(\alpha-1)t}-1) \\
&\quad + C\alpha^2(\alpha-1)^2(\gamma+C)e^{C\alpha(\alpha-1)t}\left(\left[s\frac{e^{-C\alpha(\alpha-1)s}}{-C\alpha(\alpha-1)}\right]_0^t + \right. \\
&\quad \left. + \frac{1}{C\alpha(\alpha-1)}\int_0^t e^{-C\alpha(\alpha-1)s}ds\right) = \\
&= \alpha(\alpha-1)(\gamma+C)t + x^\alpha(e^{C\alpha(\alpha-1)t}) - t\alpha(\alpha-1)(\gamma+C) \\
&\quad + \frac{\gamma+C}{C}e^{C\alpha(\alpha-1)t}\left[-e^{-C\alpha(\alpha-1)s}\right]_0^t = \\
&= (x^\alpha + \frac{\gamma+C}{C})e^{C\alpha(\alpha-1)t} - \frac{\gamma+C}{C}.
\end{aligned} \tag{7.40}$$

Then we are done and have shown the proposition. □

Another easy to verify consequence of the SDE-structure is the following proposition

Proposition 7.5.2 (Quadratic Variations)**:** *If $(X^x, Y^{x,y})$ is a strong unique solution for the martingale problem U with initial values $X_0 = x$ and $Y_0 = y$, then we have:*

$$\begin{aligned}
\langle X, X\rangle_t &= \int_0^t g(X_s)\,ds, \\
\langle Y, Y\rangle_t &= \int_0^t X_s Y_s\,ds, \\
\langle X, Y\rangle_t &= 0.
\end{aligned} \tag{7.41}$$

We will omit the proof of this proposition, which can be done as in [Kle08], page 555. The third statement can already be concluded by proposition 7.5.1.

7.6 The Feller-property of the diffusion process

The next a bit more tedious thing to prove is the existence of a Feller semigroup for the diffusion process (X, Y).

7.6.1 The main result and the strategy of the proof

With the oncoming lemma the proof of Lemma 2.2.2 will be complete:

Lemma 7.6.1: *If g satisfies Conditions 2.1.3, then the diffusion process (X, Y) is a Strong Markov process and moreover a Feller process.*

We will do the proof in several steps and separate them in some subsections.

Very fast we will see that the process is Markovian and this will not cost us a lot of time. It will be a bit more exhausting to show the Feller-property explicitly. From the Markov property we are given a Markov-semigroup S_t:

$$\begin{aligned} S_t :& C_0(\mathbb{R}^2, \mathbb{R}) \to \{h : \mathbb{R}^2 \to \mathbb{R}\}, \text{ where} \\ & S_t f(x, y) = \mathbf{E}[f(X_t, Y_t) | X_0 = x, Y_0 = y] \quad \forall f \in C_0(\mathbb{R}^2, \mathbb{R}). \end{aligned} \tag{7.42}$$

Unfortunately we cannot conclude the Feller-property of this semigroup easily, since we neither have boundedness nor ellipticity of the diffusion matrix. So we take a different approach and prove the Feller-property in a direct way. For this we will use the abbreviation C_0 for $C_0(\mathbb{R}^2, \mathbb{R})$. And we need to show the following two statements:

(i) $\forall f \in C_0, \forall x, y \geq 0 : \lim_{t \to 0}(S_t f)(x, y) = f(x, y)$.

(ii) $S_t(C_0) \subseteq C_0 \ \forall t > 0$.

They are a sufficient condition for S_t being a strongly continuous contraction semigroup on C_0. This can be found in [RY91, Proposition III.2.4]. Then we know that (X, Y) is a Feller-process.

Before these two statements are proven a compact containment condition is shown. The proof of the second property is split up in two parts as well: vanishing at infinity and continuity. The first requires one step the latter four steps. Altogether we get the following proofing scheme in eight steps:

Step 1: (X, Y) is a Markov process and has the Markov-semigroup S_t.

Step 2: (X, Y) satisfies a compact containment condition.

Step 3: Proof of (i). Convergence of the sequence $(S_t f)(x, y)$ for t tending to zero to $f(x, y)$.

Step 4: If f vanishes at infinity, then so does $S_t f$ for any $t \geq 0$.

Step 5: If f is continuous, then so is $S_t f$: Preparation.

Step 6: If f is continuous, then so is $S_t f$ in a point (x, y) with $x > 0, y > 0$.

Step 7: If f is continuous, then so is $S_t f$ in a point $(x, 0)$ with $x > 0$.

Step 8: If f is continuous, then so is $S_t f$ in a point $(0, y)$ with $y \geq 0$.

The proof is not put into a "proof environment", but held within these steps. Some lemmas and propositions will arise, but all will be belonging to the corresponding parts of the proof.

7.6.2 The proof

STEP 1: *The process (X,Y) is Markov process with Markov semigroup S_t*
Due to the existence of a unique strong solution for the SDE

$$\begin{aligned} dX_t &= \sqrt{g(X_t)}\, dW_t^1, \\ dY_t &= \sqrt{bX_tY_t}\, dW_t^2, \end{aligned}$$

the process (X,Y) is a Markov-process. Hence there is a Markov-semigroup S_t, which satisfies:

$$S_t f(x,y) = \mathbf{E}[f(X_t,Y_t)|X_0 = x, Y_0 = y] \quad \forall f \in C_0(\mathbb{R}^2,\mathbb{R}). \tag{7.43}$$

STEP 2: *Compact containment condition of (X,Y)*
In this step we will write an estimate describing how far the bivariate process (X_t,Y_t) can go away from a starting point (x,y) for a given finite time t. We will find it appropriate to use the notation $a \vee b$ for indicating that in the case $\alpha < 1$ we have a and for $\alpha \geq 1$ we have b. We start with an easy calculation about the modulus of X and Y:

$$\begin{aligned} \mathbf{E}[\|(X_t^x,Y_t^{x,y}) - (x,y)\|^2] &= \mathbf{E}[(X_t^x - x)^2 + (Y_t^{x,y} - y)^2] \\ &= \mathbf{E}[(\int_0^t \sqrt{g(X_s^x)}\, dW_s^1)^2 + (\int_0^t \sqrt{bX_t^xY_t^{x,y}}\, dW_s^2)^2] \\ = \mathbf{E}[\int_0^t g(X_s^x) ds + \int_0^t bX_s^xY_s^{x,y}\, ds] & \\ &= \int_0^t (\mathbf{E}[g(X_s^x)] + b\mathbf{E}[X_s^xY_s^{x,y}])\, ds. \end{aligned}$$

Now we will use what we calculated in Proposition 7.5.1 :

$$\begin{aligned} \mathbf{E}[\|(X_t^x,Y_t^{x,y}) - (x,y)\|^2] &\leq \int_0^t \Big[(1+x) \vee \Big((Cx^\alpha + C + \gamma)e^{C\alpha(\alpha-1)s} - \gamma\Big) - \\ &= bxyt + t(1+x) \vee \frac{C(1+x^\alpha)+\gamma}{C\alpha(\alpha-1)}(e^{C\alpha(\alpha-1)t} - 1) - \gamma t. \end{aligned} \tag{7.44}$$

Now we can use the Doob-inequality for $p = 2$, since $(X_t^x,Y_t^{x,y}) - (x,y)$ is a martingale with expectation 0 and we get the following estimate:

$$\begin{aligned} &\mathbf{P}[\sup_{0\leq s\leq t} \|(X_t^x,Y_t^{x,y}) - (x,y)\| > k_0] \leq \\ &\qquad \leq \frac{1}{k_0^2}\left(xyt + t(1+x) \vee \frac{C(1+x^\alpha)+\gamma}{C\alpha(\alpha-1)}(e^{C\alpha(\alpha-1)t} - 1)\right). \end{aligned} \tag{7.45}$$

Remark 7.6.2:
Note that the right hand side is strictly increasing in t. ◇

Lemma 7.6.3: *For any given $\epsilon > 0, t \geq 0, (x,y) \in \mathbb{R}^2_+$, there is a compactum $K = K(\epsilon, t, x, y)$ s.t. the process $(X^x_s, Y^{x,y}_s)$ stays within K with probability $\geq 1-\epsilon$, whenever $0 \leq s \leq t$.*

PROOF: The proof is clear by (7.45). □

STEP 3: *Continuity of S_t for $t \to 0$*
We start by showing the first condition (i) for the Feller-property. Let therefore $f \in C_0$ a function, $(x,y) \in \mathbb{R}^2_+$ a starting point and an $\epsilon > 0$ be given. Then all we need to show is that we can find a t_0, s.t.

$$|(S_t f)(x,y) - f(x,y)| < \epsilon \ \forall t < t_0. \tag{7.46}$$

By continuity of f in the point (x,y) it is true that for given ϵ we can find a δ s.t. for all $(\tilde{x}, \tilde{y})$ with $\|(x,y) - (\tilde{x}, \tilde{y})\| \leq \delta$:

$$|f(x,y) - f(\tilde{x}, \tilde{y})| < \frac{\epsilon}{2}. \tag{7.47}$$

Additionally choose t_0 in (7.45), s.t. for $k_0 = \delta$, the right hand side is smaller than $\epsilon/2\|f\|_\infty$. Then we get by the isotonity of (7.45) for $t \leq t_0$:

$$\begin{aligned} |(S_t f)(x,y) - f(x,y)| &= |\,\mathbf{E}_{(x,y)}[f(X_t, Y_t) - f(x,y)]\,| \leq \\ &\leq \mathbf{E}_{(x,y)}\left[\,|f(X_t, Y_t) - f(x,y)|(\mathbb{1}_{\{\|(X_t,Y_t)-(x,y)\| \leq \delta\}} + \mathbb{1}_{\{\|(X_t,Y_t)-(x,y)\| > \delta\}})\right] \leq \\ &\leq \frac{\epsilon}{2} + 2\|f\|_\infty \frac{\epsilon}{2\|f\|_\infty} = \epsilon. \end{aligned} \tag{7.48}$$

This is already what we needed to show for (i).

STEP 4: *If f vanishes at infinity, then so does $S_t f$*
Let us consider the behaviour of $S_t f$ close to infinity. Therefore let an $f \in C_0$, a fixed time t and an $\epsilon > 0$ be given. We need to check whether $S_t f$ vanishes close to infinity:

$$\lim_{(x,y)\to\infty} (S_t f)(x,y) = 0. \tag{7.49}$$

Or equivalently:

$$\forall \epsilon \ \exists n \text{ s.t. } (S_t f)(x,y) < \epsilon \ \forall (x,y) \text{ with either } x > n \text{ or } \ y > n. \tag{7.50}$$

It is clear that we need to consider two cases, one where x is diverging and another one, where y is diverging. We will only deal with the first one, since the methods for both cases would be the same and the second one is even easier to show.

So let $x_0 > 0$ be given. We will use the Doob-inequality estimate in (7.45) with $k_0 = x_0/2$ to get:

$$\begin{aligned}
&\mathbf{P}\left[\|(X_t^{x_0}, Y_t^{x_0,y}) - (x_0, y)\| > \frac{x_0}{2}\right] \leq \mathbf{P}\left[\sup_{0\leq s\leq t} \|(X_s^{x_0}, Y_s^{x_0,y}) - (x_0, y)\| > \frac{x_0}{2}\right] \\
\leq &\frac{4}{bx_0^2}\left(x_0 yt + t(1+x_0) \vee \frac{C(1+x_0^\alpha)+\gamma}{C\alpha(\alpha-1)}(e^{C\alpha(\alpha-1)t} - 1)\right) = \\
= &\frac{4byt}{x_0} + 4t\frac{1+x_0}{x_0^2} \vee 4\frac{C(1+x_0^\alpha)+\gamma}{x_0^2\ C\alpha(\alpha-1)}(e^{C\alpha(\alpha-1)t} - 1).
\end{aligned} \tag{7.51}$$

Now it is clear that for given t the right hand side can be bounded from above by any positive bound if x_0 is sufficiently large (remember that $\alpha < 2$). Intuitively we showed that if we start in x_0 far away from zero, (X_t, Y_t) can reach an area with $X_t < x_0/2$ within a short time with only small probability and we shall use that now to get (7.49). Therefore we choose x_0 so large that (7.51) is smaller than $\epsilon/2\|f\|_{\mathbb{R}_+^2}$ and s.t. $\|f\|_{[x_0/2,\infty)\times\mathbb{R}_+} < \epsilon/2$ (remember $f \in C_0$). Then we get

$$\begin{aligned}
(S_t f)(x_0, y) = &\mathbf{E}[f(X_t^{x_0}, Y_t^{x_0,y})] = \\
= &\mathbf{E}_{(x_0,y)}\left[f(X_t, Y_t)(\mathbb{1}_{\{\|(X_t,Y_t)-(x_0,y)\|\leq\frac{x_0}{2}\}} + \mathbb{1}_{\{\|(X_t,Y_t)-(x_0,y)\|>\frac{x_0}{2}\}})\right] \leq \\
\leq &\|f\|_{[x_0/2,\infty)\times\mathbb{R}_+} + \|f\|_{\mathbb{R}_+^2}\frac{\epsilon}{2\|f\|_{\mathbb{R}_+^2}} \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} \leq \\
\leq &\epsilon.
\end{aligned} \tag{7.52}$$

Clearly the argument still holds for any any other starting point $> x_0$. So for given $\epsilon > 0$ there is a x_0 s.t. (7.50) is true for all $x > x_0$. For the second coordinate a y_0 can be found the same way. Then $n = \max(x_0, y_0)$ makes (7.50).

Step 5: *If f is continuous, then so is $S_t f$: Preparation*
The most difficult thing to show is that for given $f \in C_0$, the image under S_t is a continuous function for any t. Hence for the oncoming steps keep $t > 0$ and $f \in C_0$ fixed.

Suppose the points x and y are fixed. Assume that $\tilde{x}$ and $\tilde{y}$ are close to x and y and we have solutions to the martingale problem (2.2.3) with initial values $(X_0^x, Y_0^{x,y}) = (x, y)$ and $(X_0^{\tilde{x}}, Y_0^{\tilde{x},\tilde{y}}) = (\tilde{x}, \tilde{y})$. What we need to show is the following:

$$\forall\epsilon > 0\ \exists\delta > 0\ s.t.\ \forall\tilde{x},\ \tilde{y}\text{ with }\|(\tilde{x},\tilde{y}) - (x,y)\| < \delta : |S_t f(x,y) - S_t f(\tilde{x},\tilde{y})| < \epsilon \tag{7.53}$$

We will do that by showing that the solutions "come together" as time goes on. To do that we require a coupling argument:

$$\begin{aligned}
|S_t f(x,y) - S_t f(\tilde{x},\tilde{y})| \leq &|\mathbf{E}\left[f(X_t^x, Y_t^{x,y}) - f(X_t^x, Y_t^{x,\tilde{y}})\right]| + \\
&+ |\mathbf{E}\left[f(X_t^x, Y_t^{x,\tilde{y}}) - f(X_t^{\tilde{x}}, Y_t^{\tilde{x},\tilde{y}})\right]| \qquad (7.54) \\
\leq &\|f\|\,\mathbf{E}\left[P(T_{y,\tilde{y}} > t | X^x)\right] + \|f\|\,\mathbf{P}(T_{x,\tilde{x}} > \tilde{t} \cup \hat{T}_{y,\tilde{y}} > t), \qquad (7.55)
\end{aligned}$$

where the following coupling times are used

$$\begin{aligned} T_{y,\tilde{y}} &:= \inf\{s > 0 : Y_s^{x,y} = Y_s^{x,\tilde{y}}\}, \\ T_{x,\tilde{x}} &:= \inf\{s > 0 : X_s^{x} = X_s^{\tilde{x}}\}, \\ \hat{T}_{y,\tilde{y}} &:= \inf\{s > T_{x,\tilde{x}} : Y_s^{x,\tilde{y}} = Y_s^{\tilde{x},\tilde{y}}\} \text{ and} \\ & \tilde{t} < t, \text{ where } \tilde{t} \text{ is fixed.} \end{aligned} \tag{7.56}$$

The idea is to couple the processes mentioned in the previous lines according to the technique presented in [Lin92, page 214]. To describe the idea shortly we let the processes in (7.54) run independently until its coupling times, i.e. the first time they meet. Then the processes run together and that is why we only need to worry about the question whether the coupling times can be bounded from above by any positive ϵ, if we choose the initial points $\tilde{x}$ and $\tilde{y}$ sufficiently close to x and y.

We will benefit from the catalytic setting of the problem and we note that for the second summand we first want the catalyst to couple at time $T_{x,\tilde{x}}$, which will be a very small time and then as the reactants have not moved apart to dramatically we can still be sure that the reactant couples sufficiently fast.

We treat several cases in the next steps and start with the most difficult one.

STEP 6: *If f is continuous, then so is $S_t f$ in a point (x, y) with $x > 0, y > 0$*
(A) We start with the first summand and give a preparatory lemma:

Lemma 7.6.4: *Let a Brownian motion W be given. In the case $x > 0, y > 0$ the following holds:*
For all $s > 0, \epsilon > 0$ there exist $\delta_1 > 0, \delta_2 > 0$ s.t.

$$\mathbf{P}(\langle Y^{x,y}, Y^{x,y}\rangle_s > \delta_1) > 1 - \epsilon, \tag{7.57}$$

$$\mathbf{P}(W(r) < \delta_2 \ \forall r \leq \delta_1) < \epsilon. \tag{7.58}$$

Proof. The proof is given by simple calculations using (7.45):

$$\langle Y^{x,y}, Y^{x,y}\rangle_s = \int_0^s X_r^x Y_r^{x,y}\, dr \geq \tag{7.59}$$

$$\geq \tilde{s}\frac{xy}{4}\mathbb{1}_{\{|X_r^x - x| < \frac{x}{2}, |Y_r^{x,y} - y| < \frac{y}{2} \ \forall r < \tilde{s}\}}. \tag{7.60}$$

And by (7.45) and a choice of $\tilde{s}$ sufficiently small (note that the quadratic variation is monotonous in s) this can be bounded from below with a constant δ_1 with probability $> 1 - \epsilon$.
Additionally by the distribution of level-hitting times for Brownian motion (compare [RY91, p.107]) it is clear that the second claim also holds by choosing δ_2 appropriately. □

Then we can start thinking about the coupling time $T_{y,\tilde{y}}$ for a given catalyst X^x started in a fixed point x. By the Martingale Representation Theorem for the martingale $Y_s^{x,y} - Y_s^{x,\tilde{y}}$ it holds for a Brownian Motion W that:

$$\{T_{y,\tilde{y}} > t\} = \{Y_s^{x,y} \neq Y_s^{x,\tilde{y}} \ \forall s \leq t\} = \tag{7.61}$$

$$= \{W(\langle Y^{x,y} - Y^{x,\tilde{y}}, Y^{x,y} - Y^{x,\tilde{y}}\rangle_s) \neq y - \tilde{y} \ \forall s \leq t\} \tag{7.62}$$

and since $\langle A-B, A-B\rangle_s \geq \langle A, A\rangle_s$ for independent processes, we have:

$$\{T_{y,\tilde{y}} > t\} \subseteq \{W(\langle Y^{x,y}, Y^{x,y}\rangle_s) \neq y - \tilde{y} \ \forall s \leq t\}. \tag{7.63}$$

By the preceding lemma we get by the first line for a good choice of δ_1 that

$$\mathbf{P}(T_{y,\tilde{y}} > t) \leq (1-\epsilon)\mathbf{P}(W(s) \neq y - \tilde{y} \ \forall s \leq \delta_1) + \epsilon, \tag{7.64}$$

and then by the second line for choosing $y - \tilde{y} < \delta_2$:

$$\mathbf{P}(T_{y,\tilde{y}} > t) \leq (1-\epsilon)\epsilon + \epsilon < 2\epsilon. \tag{7.65}$$

So the right choice of $\tilde{y}$ is already done to bound the first summand in (7.55).
(B) To control the second summand we need to follow some more complicated paths. First we give a lemma similar to the one before:

Lemma 7.6.5: *Let a Brownian motion W be given. In the case $x > 0$ the following holds:*
For all $\tilde{t} > 0, \epsilon > 0$ there exist $\delta_1 > 0, \delta_2 > 0$ s.t.

$$\mathbf{P}(\langle X^x, X^x\rangle_{\tilde{t}} > \delta_1) > 1 - \epsilon, \tag{7.66}$$

$$\mathbf{P}(W(r) < \delta_2 \ \forall r \leq \delta_1) < \epsilon. \tag{7.67}$$

Proof. The proof uses (7.45) and Condition 2.1.3 ((C2) and (C4)) and is similar to the one before. For $\tilde{\tilde{t}} < \tilde{t} < t$ and $\alpha < 1$:

$$\begin{aligned}
\langle X^x, X^x\rangle_{\tilde{t}} &= \int_0^{\tilde{t}} g(X_r^x)dr \\
&\geq \int_0^{\tilde{\tilde{t}}} g(X_r^x)\mathbb{1}_{\{|X_s^x - x| < \alpha x \ \forall s \leq \tilde{\tilde{t}}\}} + \mathbb{1}_{\{|X_s^x - x| \geq \alpha x \ \forall s \leq \tilde{\tilde{t}}\}} dr \\
&\geq \int_0^{\tilde{\tilde{t}}} \min_{z \in [x-\alpha x, \infty)} g(z)\mathbb{1}_{\{|X_s^x - x| < \alpha x \ \forall s \leq \tilde{\tilde{t}}\}} \, dr \\
&\geq \tilde{\tilde{t}}(c'(x-\alpha x)^{1+\beta} \wedge g_0)\mathbb{1}_{\{|X_s^x - x| < \alpha x \ \forall s \leq \tilde{\tilde{t}}\}}.
\end{aligned} \tag{7.68}$$

But by the Doob estimate in (7.45) the right hand side can be bounded from below by a positive constant with probability $> 1 - \epsilon$ as in the previous lemma.
The second line is proven the same way as above. □

The idea is not to let the reactant move too far apart its starting point until time $\tilde{t}$. Hence, for given $\epsilon > 0$, $\tilde{\delta} > 0$ fix $0 < \tilde{t} < t$ s.t. (by (7.45)):

$$\begin{aligned}
&\mathbf{P}(|Y_s^{x,\tilde{y}} - \tilde{y}| < \tilde{\delta} \ \forall s < \tilde{t}) > 1 - \epsilon \text{ and} \\
&\mathbf{P}(|Y_s^{\tilde{x},\tilde{y}} - \tilde{y}| < \tilde{\delta} \ \forall s < \tilde{t}) > 1 - \epsilon.
\end{aligned} \tag{7.69}$$

Then find, by Lemma 7.6.5 and the ideas at the end of part (A), $\tilde{x}$ in the neighborhood of x s.t.:

$$\mathbf{P}(T_{x,\tilde{x}} > \tilde{t}) < \epsilon. \tag{7.70}$$

(C) To bound the second expression in (7.55) note that:

$$P(T_{x,\tilde{x}} > \tilde{t} \cup \hat{T}_{y,\tilde{y}} > t) \leq P(T_{x,\tilde{x}} > \tilde{t}) + P(T_{x,\tilde{x}} > \tilde{t} \cap \hat{T}_{y,\tilde{y}} > t). \tag{7.71}$$

The first summand can be bounded by part **(B)**. For the second summand observe the following:

At time $\tilde{t}$ the processes $Y^{x,\tilde{y}}$ and $Y^{\tilde{x},\tilde{y}}$ differ at most by $2\tilde{\delta}$ by (7.69) and the catalysts have already "merged". Hence at time $\tilde{t}$ we are in the situation of case (A) where we need to wait for the merger of two reactants with fixed catalytic background. But this can be done with a probably different choice of $\tilde{x}$.

STEP 6: *If f is continuous, then so is $S_t f$ in a point $(x,0)$ with $x > 0$*
We will argue similar as in the previous step, but we cannot apply Lemma 7.6.4: we cannot bound $Y^{x,\tilde{y}}$ from below.

We go back to (7.53) and write for $\tilde{x} > 0, \tilde{y} > 0$:

$$\begin{aligned} |S_t f(x,0) - S_t f(\tilde{x},\tilde{y})| \leq & |\mathbf{E}\left[f(X_t^x, Y_t^{x,0}) - f(X_t^{\tilde{x}}, Y_t^{\tilde{x},0})\right]| + \\ & + |\mathbf{E}\left[f(X_t^{\tilde{x}}, Y_t^{\tilde{x},0}) - f(X_t^{\tilde{x}}, Y_t^{\tilde{x},\tilde{y}})\right]| \qquad (7.72) \\ \leq & \|f\|\,\mathbf{P}(T_{x,\tilde{x}} > t) + \mathbf{E}[f(X_t^{\tilde{x}},0) - f(X_t^{\tilde{x}}, Y_t^{\tilde{x},\tilde{y}})]. \qquad (7.73) \end{aligned}$$

The bound on the first summand is clear by part (B) of the previous step. The second summand is not too difficult to bound since it is just a one-dimensional problem. Hence set

$$T_0 := \inf\{s > 0 : Y^{\tilde{x},\tilde{y}} = 0\} \text{ and} \qquad (7.74)$$

$$T_{\sqrt{\tilde{y}}} := \inf\{s > 0 : Y_s^{\tilde{x},\tilde{y}} = \sqrt{\tilde{y}}\}. \qquad (7.75)$$

Now we divide the probability space into the following three (not necessarily disjoint) sets:

$$\{T_0 < T_{\sqrt{\tilde{y}}}, T_0 < t\} \cup \{T_{\sqrt{\tilde{y}}} < T_0, T_{\sqrt{\tilde{y}}} < t\} \cup \{T_0 > t, T_{\sqrt{\tilde{y}}} > t\}. \qquad (7.76)$$

Clearly we want to bound the second summand from above with the help of these sets. Note that

$$\mathbf{P}[T_{\sqrt{\tilde{y}}} < T_0 | X_s^{\tilde{x}}] = \sqrt{\tilde{y}}. \qquad (7.77)$$

Hence,

$$\begin{aligned} & \mathbf{E}[f(X_t^{\tilde{x}},0) - f(X_t^{\tilde{x}}, Y_t^{\tilde{x},\tilde{y}})] \qquad (7.78) \\ & \leq 0 + 2\|f\|\mathbf{P}[T_{\sqrt{\tilde{y}}} < T_0 | (X_s^{\tilde{x}}] + \mathbf{E}[\sup_{z \in (0,\sqrt{\tilde{y}})} |f(X_t^{\tilde{x}},0) - f(X_t^{\tilde{x}}, z)|]. \qquad (7.79) \end{aligned}$$

And since f is continuous in the second coordinate 0, we can choose $\tilde{y}$ so close to zero, s.t. the right hand side is bounded by ϵ. If additionally for the first summand we had chosen $\tilde{x}$ so close to x that

$$\|f\| P(T_{x,\tilde{x}} > t) < \epsilon, \qquad (7.80)$$

then we obtain:

$$|S_t f(x,0) - S_t f(\tilde{x},\tilde{y})| \leq 2\epsilon. \qquad (7.81)$$

Hence this step is done.

STEP 7: *If f is continuous, then so is $S_t f$ in a point $(0,y)$ with $y \geq 0$*
This step involves similar problems as the previous one: non-applicability of Lemma

7.6.5. The advantage, however, is that we only have to deal with some sort of one-dimensional problem. We have:

$$|S_t f(0,y) - S_t f(\tilde{x},\tilde{y})| \leq |f(0,y) - f(0,\tilde{y})| + |\mathbf{E}\left[f(0,\tilde{y}) - f(X_t^{\tilde{x}}, Y_t^{\tilde{x},\tilde{y}})\right]|. \quad (7.82)$$

The first summand can be bounded by the right choice of $\tilde{y}$ and the second one needs a similar argument as just before: Let us first couple the catalyst $X^{\tilde{x}}$ to zero and after that "short" time the reactant has not left $\tilde{y}$ too much:

$$T_0 := \inf\{s > 0 : X_s^{\tilde{x}} = 0\}, \quad (7.83)$$

$$T_{\sqrt{\tilde{x}}} := \inf\{s > 0 : X_s^{\tilde{x}} = \sqrt{\tilde{x}}\}. \quad (7.84)$$

Then we divide the probability space as in the step just before in (7.76) and get:

$$\begin{aligned} |\mathbf{E}\left[f(0,\tilde{y}) - f(X_t^{\tilde{x}}, Y_t^{\tilde{x},\tilde{y}})\right]| \leq & \mathbf{E}[|f(0,\tilde{y}) - f(0,Y_t^{\bullet,\tilde{y}})|] + 2\|f\|\mathbf{P}(T_0 > T_{\sqrt{\tilde{x}}}) \\ & + \mathbf{E}[\sup_{z\in[0,\sqrt{\tilde{y}}]} |f(0,\tilde{y}) - f(z,Y_t^{\bullet,\tilde{y}})|]. \end{aligned} \quad (7.85)$$

Here we introduced $Y_t^{\bullet,\tilde{y}}$ as a stochastic process given as a reactant, whose catalyst total mass is in between 0 and $\sqrt{\tilde{x}}$. We need to show that $Y_t^{\bullet,\tilde{y}}$ is not too far from $\tilde{y}$. By the Doob-inequality:

$$\begin{aligned} \mathbf{P}[\sup_{0\leq s\leq t} |Y_t^{\bullet,\tilde{y}} - \tilde{y}| > \delta_1] &\leq \frac{1}{\delta_1^2}\mathbf{E}[\int_0^t bX_r Y_r^{\bullet,\tilde{y}}\, dr] \\ &\leq \frac{bt}{\delta_1^2}\sqrt{\tilde{x}}\tilde{y}. \end{aligned} \quad (7.86)$$

And if we choose $\tilde{x}$ so close to zero s.t.

$$\sqrt{\tilde{x}} < \frac{\delta_1^2\epsilon}{3bt\tilde{y}\|f\|} \wedge \frac{\epsilon}{2\|f\|} \quad (7.87)$$

and $\tilde{y}$ s.t.

$$|f(0,y) - f(0,\tilde{y})| < \epsilon/3, \quad (7.88)$$

we get that by (7.85):

$$|S_t f(0,y) - S_t f(\tilde{x},\tilde{y})| \leq \epsilon. \quad (7.89)$$

And that was all we needed to show.

Step 8: *If f is continuous, then so is $S_t f$ in the point $(0,0)$*
The shortest and easiest step is hopefully a good time to recover for the reader:

$$|S_t f(0,0) - S_t f(\tilde{x},\tilde{y})| \leq |f(0,0) - f(0,\tilde{y})| + |f(0,\tilde{y}) - S_t f(\tilde{x},\tilde{y})|. \quad (7.90)$$

The first summand can be bounded by continuity of f and the second summand can be done as the previous Step 7.

Now all the steps we needed to prove the Feller-property of (X,Y) are taken. Therefore Lemma 7.6.1 is proved. □□□

7.7 Convergence of the total mass process to (X, Y)

Now we will prove the convergence of the total-mass-processes, Theorem 2.3.1:

Theorem 7.7.1 (Weak convergence of the total masses)**:**
When g satisfies Condition 2.1.3, then

$$\mathcal{L}[(\eta^{\text{tot},n}, \xi^{\text{tot},n})] \Longrightarrow \mathcal{L}[(X, Y)] \text{ as } n \to \infty, \tag{7.91}$$

where convergence is weak convergence in the path space $\mathcal{D}_{\mathbb{R}_+}[0, \infty)$.

The proof of this theorem can be put into a wider context. For diffusion limits there is a wide number of theorems available, some of which can be found in [EK86]. We will give a rather general proposition, which does not use the Feller-property shown in the previous section. We will show the following proposition:

Proposition 7.7.2: *Let $(Z^n)_{n\in\mathbb{N}}$ be a sequence of martingales in $\mathbb{R}^d$ corresponding to closable pregenerators Ω^n on $C_c(\mathbb{R}^d, \mathbb{R})$. Z^n has càdlàg paths and satisfies the compact containment condition:*

> *For every $\lambda > 0$ and $T > 0$, there exists a compact set $\Gamma_{\lambda,T} \subset \mathbb{R}^d$, independent of n such that*
>
> $$\inf_n P\left[Z^n(t) \in \Gamma_{\lambda,T} \text{ for } 0 \le t \le T\right] \ge 1 - \lambda. \tag{7.92}$$

Let additionally Z be the $\mathbb{R}^d$-valued unique solution of the $(\Omega, \mathbb{D}(\Omega))$ martingale problem, where $\mathbb{D}(\Omega) \subset C_0(\mathbb{R}^d, \mathbb{R})$. Assume that C_c^∞ is a core for Ω and the following two properties hold:

(i) $\mathcal{L}[Z_0^n] \to \mathcal{L}[Z_0]$,

(ii) For every $f \in C_c^\infty$ it is true that:

$$\lim_{n\to\infty} \|\Omega^n f - \Omega f\|_{\mathbb{R}^d} = 0. \tag{7.93}$$

Then $Z^n \Longrightarrow Z$, where convergence is in distribution in $\mathcal{D}_{\mathbb{R}^d}[0, \infty)$.

PROOF: The proof will go along some theorems from [EK86], explicitly Theorem 3.9.1, 3.9.4 and 4.8.10, which are all quoted in the appendix.

The proof is split into three parts. In a first part we will show tightness of the sequence $(Z^n)_{n\in\mathbb{N}}$ in $D_{\mathbb{R}^d}[0, \infty)$ by Theorem 3.9.1 and 3.9.4. In a second part we will show convergence of the finite-dimensional distributions of Z^n to those of Z and the third part puts together these results and shows weak convergence in the path space $D_E[0, \infty)$.

(A) We use Theorem 3.9.1 in [EK86] and notice that the compact containment condition is already given. So it suffices to verify relative compactness of $(f \circ Z^n)_{n\in\mathbb{N}}$ for any $f \in H = C_c^\infty(\mathbb{R}^d, \mathbb{R})$. Thus we can look at Theorem 3.9.4 and we only need to show that for any $f \in H$, $T > 0$ condition (3.9.18) holds for $p = 2$:

$$\sup_n E\left[\int_0^T |(\Omega^n f)(Z_t^n)|^2 dt\right] < \infty\,. \tag{7.94}$$

Let us start the calculation:

$$|(\Omega^n f)(Z_t^n)|^2 \leq 2(|(\Omega f)(Z_t^n)|^2 + |(\Omega^n f)(Z_t^n) - (\Omega f)(Z_t^n)|^2). \tag{7.95}$$

The first summand is constant in n and bounded from above by a bound on derivatives on f (remember that f has compact support). The second one is bounded since we have (7.93) on compacta. With this easy argument, the first part is already done.

(B) In the second part we use Theorem 4.8.10 and show (c) to get the convergence of f.d.d.'s in (a) and we note that Z was the unique solution of the $(\Omega, \mathbb{D}(\Omega))$-martingale problem. By the first part $(Z^n)_{n\in\mathbb{N}}$ is relatively compact and we take $M = \{\mathbb{1}_{(-\infty,a)}(\cdot_s) : a, s \in \mathbb{Q}, s \geq 0\}$.

Let $f \in C_c^\infty(\mathbb{R}^d, \mathbb{R})$, $T > 0$, $\xi_n(t) = f(Z_t^n)$, $\phi_n(t) = (\Omega^n f)(Z_t^n)$. Now we have to check (8.51),(8.52), (8.53), (8.54) in chapter 4 and the martingale property (9.16) in chapter 3. The martingale property is satisfied, since Z^n is a solution of the discrete martingale problem. The first condition (8.51) is verified, since f is bounded. We show (8.52):

$$E[\,|\phi_n(t)|\,] = E[|\Omega^n f(Z_t^n)|] \leq E[\,|(\Omega f)(Z_t^n)|\,] + E[\,|(\Omega_n f)(Z_t^n) - (\Omega f)(Z_t^n)|\,]. \tag{7.96}$$

With the same arguments as in the first part this expectation is uniformly bounded in n and in a time $0 \leq t \leq T$. So (8.52) is shown. Clearly (8.53) holds by definition, so it only remains to show (8.54):

$$E[\,|\phi_n(t) - \Omega f(Z_t^n)|\,] = E[\,|\Omega^n f(Z_t^n) - \Omega f(Z_t^n)|\,] \to 0 \text{ as } n \to \infty, \tag{7.97}$$

since f has compact support and by (7.93). So the second part is also done.

(C) In this part we use Theorem 3.7.8 of [EK86]. The requirements of (b): relative compactness and f.d.d. convergence are fulfilled and were shown in the first and second part. So convergence in path space follows: $\mathcal{L}[Z^n] \to \mathcal{L}[Z]$ in $D_{\mathbb{R}^d}[0, \infty)$. □

Now it is an easy exercise to adopt this proposition for the proof of Theorem 7.7.1:

PROOF: *(of Theorem 7.7.1)*
We set $Z^n = (\eta^{\text{tot},n}, \xi^{\text{tot},n})$ and $Z = (X, Y)$. Then all we have to do is to show the requirements of Proposition 7.7.2, where the index in the bracket indicates where they were already proven:

- $(\eta^{\text{tot},n}, \xi^{\text{tot},n})_{n\in\mathbb{N}}$ satisfies the compact containment condition (Lemma 7.3.3),
- $(\eta^{\text{tot},n}, \xi^{\text{tot},n})$ has càdlàg paths for each n (Lemma 7.2.1),
- The $(U, \mathcal{D}(U))$-martingale problem has a unique solution (Theorem 7.4.1) .
- For $f \in C_c^\infty$, we have $\|U_n f - Uf\|_{\mathbb{R}_+^2} \to 0$ as $n \to \infty$.

We only need to show the last argument and let $f \in C_c^\infty$ be given. Let K be the compactum in $\mathbb{R}_+^2$, s.t. $f|_{K^c} \equiv 0$. Then we have

$$\begin{aligned}\|U_n f - Uf\|_{\mathbb{R}_+^2} &= \|U_n f - Uf\|_K \\ &= \|\frac{n^2}{2} g(x)[f(x + \frac{1}{n}, y) - 2f(x, y) + f(x - \frac{1}{n}, y)] + \frac{n^2}{2} bxy[f(x, y + \frac{1}{n}) \\ &\quad - 2f(x, y) + f(x, y - \frac{1}{n})] - \frac{1}{2} g(x) \frac{\partial^2 f}{\partial x^2}(x, y) - \frac{1}{2} bxy \frac{\partial^2 f}{\partial y^2}(x, y)\|_K.\end{aligned} \tag{7.98}$$

With Taylor's expansion up to second order of the non-$f(x, y)$-summands around (x, y) we get:

$$\begin{aligned} \|A_n f - Af\|_{\mathbb{R}^2_+} &= 2\|\frac{1}{2n} g(x) f_{xxx}(\hat{x}, y) + \frac{1}{2n} xy f_{yyy}(x, \hat{y})\|_K \leq \\ \leq \frac{1}{n} &\|g(x) f_{xxx}(\hat{x}, y) + bxy f_{yyy}(x, \hat{y})\|_{(x,y),(\hat{x},\hat{y})\in K} \to 0 \text{ as } n \to \infty. \end{aligned} \tag{7.99}$$

This is all we needed to show. So we can use Proposition 7.7.2 to show Theorem 2.3.1. □

8 Proofs of the main results from Chapters 4 and 5

Within this chapter the proofs of the main results from Chapter 4 and 5 are presented. They are put together in one chapter, since the proofs used for the tree-valued-process and the contour-process depend on each other by the mappings $\mathcal{C}$ and $\mathcal{T}$, which map trees to contours and vice versa.

8.1 Preliminary considerations for the quenched analysis

8.1.1 Regular conditional probabilities and quenched analysis

The quenched analysis will present some results about the reactant process conditioned on the catalyst total mass process. In order to describe results about the quenched point of view it is sufficient to consider single catalyst sample paths only. This is true since $\mathbb{T}^{\text{root}}$ and $\mathcal{D}_{\mathbb{R}_+^2}[0,\infty)$ are complete and separable metric spaces. Therefore by Theorem 5.3.19 in [KS00] regular conditional probabilities, i.e. kernels, exist (see Theorem A.2.1 in the Appendix):

$$K(\eta, A) = \mathbf{P}\left[((\eta_t^{\text{tot}})_{t\geq 0}, \xi_r^{\text{for}}) \in A | (\eta_t^{\text{tot}})_{t\geq 0} = \eta\right]. \tag{8.1}$$

Clearly the same is true for the rescaled processes and the diffusion, where for each

we get a transition kernel: All of these kernels can be brought together as different kernels on the common probability space as in Definition 2.3.3. Indeed we think of all rescaled catalysts realized on the common probability space as in Chapter 2, (2.18):

$$K^n(\eta, A) := \text{ prob. of } n\text{-rescaled process in A under } \eta^{\text{tot},n} = \eta\ , \tag{8.2}$$

$$K(x, A) := \text{ prob. of limit process in A under } X = x\ . \tag{8.3}$$

Note that yet we cannot really speak about the second line, since we do not know if a limit object exists for trees or contours. As later we will be able to show that, we do not hesitate to introduce the notation already.

8.1.2 Specification of the quenched catalysts

This is the point where we do not take the catalyst processes to be generated by a catalyst branching scheme, but by a wider class of processes, and as we already conditioned, by a wider class of functions.

Take a fixed sequence of càdlàg functions x^n and a continuous function x s.t.

$$\begin{aligned} x^n &: [0, \hat{t}^n] \to \mathbb{R}_+ \text{ and} \\ x &: [0, \hat{t}] \to \mathbb{R}_+. \end{aligned} \tag{8.4}$$

The first function x shall play the role of the total mass diffusion X as part of the SDE (2.9). The functions x^n shall play the role of the discrete total mass processes $\eta^{\text{tot},n}$ as part of the individual branching rate of the reactant. Therefore the functions are additionally required to have the following properties:

Condition 8.1.1: *(i)* $x(0) = x^n(0) = 1$,

(ii) $x(\hat{t}) = x^n(\hat{t}^n) = 0$,

(iii) before the absorption times x and x^n are strictly positive: $x^n(c) > 0$, when $c \in (0, \hat{t}^n)$,

(iv) $\hat{t}^n \to \hat{t}$ as n tends to infinity and

(v) for any $T > 0$ it holds:

$$\lim_{n\to\infty} \sup_{t\leq T} |x^n(t) - x(t)| = 0. \tag{8.5}$$

Indeed the catalyst total mass process and its rescaled versions definitely satisfy all these conditions, see Corollary 2.3.2 and (2.18).

But note that there are also some properties, which η^{tot} has, but which are not necessary for the proofs to come: it is not required that the jumps of x^n are of magnitude $1/n$ only. The range of possible catalysts is therefore indeed wider than the ones given by the catalyst total mass process in Chapter 2. Any rescaled sequence of functions satisfying the Condition can be a catalyst and the proofs can be done.

Remark 8.1.2:
From now on take the functions x^n and x to be fixed for the whole chapter. All probabilities even if not explicitely indicated are understood to be quenched, i.e. conditioned on $\eta^{\text{tot},n} = x^n, X = x$.
We will abbreviate notation and simply write $\xi^{for,n}$ for the reactant forest with catalyst x^n and which is cut at the height of the non-random extinction time $\hat{t}^n$ of the catalyst. ◇

8.2 The reactant limit forest exists: The proof strategy

The main theorem from Chapter 4 claims that the sequence of rescaled trees converges:

Theorem 8.2.1 (The reactant limit forest exists)**:**
There exists a random variable $Y^{\text{for}} \in \mathbb{T}^{\text{root}}$, s.t.:

$$\mathcal{L}\left[\xi^{\text{for},n}; \eta^{\text{tot},n}\right] \xrightarrow{n\to\infty} \mathcal{L}\left[Y^{\text{for}}; X\right]. \tag{8.6}$$

The law of Y^{for} is given by

$$\mathcal{L}\left[Y^{\text{for}}; X\right] = \lim_{\delta\to 0} \mathcal{L}\left[\mathcal{T}((\zeta_u^\delta)_{0\leq u\leq \alpha_{4/b}})\right], \tag{8.7}$$

where convergence is in the Prohorov-metric of probability measures, $\mathcal{T}$ is the mapping described on page 31.

The diffusion ζ^δ is the unique solution of the $(A^\delta, \mathbb{D}(A^\delta))$ martingale problem, where $\mathbb{D}(A^\delta)$ is given by:

$$\mathbb{D}(A^\delta) = \{f \in C^1([0,\tau^\delta],[0,\infty)) : f'|_{\{0,\tau^\delta\}} = 0, \frac{f'(\cdot)}{X_\cdot} \in C^2_{[0,\tau^\delta]}([0,\infty))\} \tag{8.8}$$

and for each $f \in \mathbb{D}(A^\delta)$:

$$A^\delta f(c) = \left(\frac{2f'}{X_c}\right)'(c). \tag{8.9}$$

Furthermore $\alpha_{4/b}$ is the local time inverse at level 0 of ζ^δ of $\frac{4}{b}$.

The proof of this theorem will occupy the rest of this chapter, since it is rather involved. We will split the proof in several sections containing various lemmas, but also some theorems, which are worth to be mentioned as well. In a short description the proof goes like this:

- First we show that the sequence $(\xi_t^{\text{for},n}; \eta^{\text{tot},n})_{n\in\mathbb{N}}$ is tight in $\mathbb{T}^{\text{root}}$. This is done with an argument that shows that the structure of the reactant tree is rather regular with respect to the catalyst: one can find an ϵ-net in each tree, where the overall number of net-points is bounded uniformly in n for a fixed catalyst. For this see Section 8.3.

- In a second step we translate the problem from the tree-setting to the contour setting. In the contour setting the cut reactant contours of trees which are cut at a certain height form a tight series and all possible limit contours solve a martingale problem. This martingale problem is then shown to be uniquely solvable. Therefore the cut contours converge weakly to a cut limit contour. This can be found in Section 8.4.

- In the end we show that the unique cut tree corresponding to the unique limit contour form a Cauchy sequence in the contour cutting parameter. But any limit point of the tight tree sequence $\left(\mathcal{L}[\xi^{\text{for},n}; \eta^{\text{tot},n}]\right)_{n\in\mathbb{N}}$ must have the same law on the cut trees. Therefore there can only be one limit law on $\mathbb{T}^{\text{root}}$, since the cut-tree law is uniquely determined by the unique contour law. This part is done in Section 8.5.

Each of these points is shown in an own section.

8.3 Tightness of the reactant tree-valued process

8.3.1 The main result and the strategy of the proof

In this section we do the first step to prove Theorem 8.2.1 and prove Proposition 4.2.1. This proposition claims:

Proposition 8.3.1: *Under the hypotheses of Condition 8.1.1 the family $(\xi^{\text{for},n})_{n\in\mathbb{N}}$ is tight in $\mathbb{T}^{\text{root}}$.*

Since the proof is a bit longer first the strategy of the proof is given and then the formal proof follows.

In fact by Lemma 2.5 of [EPW06] what needs to be shown is the following: For each $\gamma > 0$ there is a compact subset Γ_γ of $\mathbb{T}^{\text{root}}$ such that for all $n \in \mathbb{N}$:

$$\mathbf{P}[\xi^{\text{for},n} \in \Gamma_\gamma] \geq 1 - \gamma. \tag{8.10}$$

To get this for fixed n assume each tree $\xi^{\text{for},n}$ to be cut in ϵ-thick slices. Each slice contains a certain number of different arc pieces. Out of each arc piece choose a single point.

If with high probability an upper bound $n(\epsilon)$ on the number of these chosen points can be given for each ϵ, then by Lemma 2.5 in [EW06] we have a precompact set. Moreover this upper bound can be established with high probability for the random set of forests $\xi^{\text{for},n}$. So the main task will become estimating numbers of ancestors and hence it is good to look at the tree from a backward perspective (look-down). Indeed this bound can be given via an argument from [DK96], which relates looking at branching in inverse time on the one hand and to the Kingman coalescent on the other hand.

Therefore, the proof is split into the following steps:

Step 1: The Kingman coalescent and its behaviour close to zero.

Step 2: Description of a set $\Gamma_{\gamma,L_m} \subset \mathbb{T}^{\text{root}}$ of trees which is precompact in the Gromov-Hausdorff-metric using [EPW06].

Step 3: Branching looked at in inverse time direction is like a time-changed Kingman coalescent, as in [DK96].

Step 4: Proof of the tightness property, via showing that the rescaled reactant trees are in Γ_{γ,L_m} with probability $1 - \gamma$.

8.3.2 The proof

STEP 1: *The Kingman coalescent*

In 1982 J.F.C Kingman presented in his paper [Kin82] the so-called "Kingman"-coalescent. It is a stochastic process describing the evolution of a partition of an arbitrary countable set. Each pair of partition elements merges after exponential time independent of all other possible partition elements partners.

To be more precise let a $n \in \mathbb{N}$ and a set A containing at least n elements be given. The *n-coalescent* $(R_t^{(n)})_{u\geq 0}$ is a process starting with the partition of the set A into n different partition elements $A_1, A_2, \ldots, A_n$. There are $\binom{n}{2}$ possible *pairs* of partition elements, which merge at constant rate 1 each. After the first merging of two partition elements there are $n-1$ partition elements left and the process continues as before by merging at constant rate of the now possible $\binom{n-1}{2}$ pairs. The process continues until it reaches the exit state, where there is only one partition element left, the whole set A. Let the times between two mergers from k to $k-1$ partition elements be called T_k. Then this time is exponentially distributed with mean $2/k(k-1)$.

It is possible to extend this definition to starting with a given countably infinite partition of the set A. When reaching a partition with n partition elements this

process continues as the merging for the n-coalescent. This extension is possible because of the convergence of the sum of the expectations of T_k and the process shall be called the *coalescent* $(R_t)_{t\geq 0}$. More about this is reflected in the Appendix B.1.

Quite many things are known about the Kingman coalescent, some of which can be found in [Ald99], page 27. Here we will state an easy lemma we use later on:

Lemma 8.3.2: *For the Kingman coalescent $(R_t)_{u\geq 0}$ it is true that*

$$E[\#R_t] \leq 16/t + 3. \tag{8.11}$$

PROOF: The proof consists of four small parts and essentially only uses the Markov inequality:
(A) By definition $\#R_t$ is the number partition elements at time t. It is a death process starting at $infinity$ with rate $\binom{k}{2}$ in state k. Then each of the waiting times between k and $k-1$ partition elements is exponentially-$\binom{k}{2}$-distributed. Let $T_1, T_2, \ldots$ be this sequence of independent exponentially distributed random variables with $E[T_k] = 1/\binom{k}{2}$. Then it holds for

$$S_n = \sum_{k=n+1}^{\infty} T_k, \tag{8.12}$$

that

$$\{S_n \geq t\} = \{\#R_t \geq n\}. \tag{8.13}$$

(B) Additionally calculate the Legendre transform of S_n for $\lambda < \binom{n}{2}$:

$$\begin{aligned} E[e^{\lambda S_n}] &= \prod_{k=n+1}^{\infty} E[e^{\lambda T_k}] = \prod_{k=n+1}^{\infty} \left(1 - \frac{\lambda}{\binom{k}{2}}\right)^{-1} \\ &= \exp\left(-\sum_{k=n+1}^{\infty} \log(1 - \frac{\lambda}{\binom{k}{2}})\right) \leq \exp\left(-\sum_{k=n+1}^{\infty}\sum_{m=1}^{\infty} (\frac{\lambda}{\binom{k}{2}})^m\right) \\ &= \exp\left(\sum_{k=n+1}^{\infty} \infty \frac{2\lambda}{k(k-1) - 2\lambda}\right). \end{aligned} \tag{8.14}$$

By choosing $\lambda_n = \lambda = \frac{n(n-1)}{4}$ we get:

$$E[e^{\lambda_n S_n}] \leq \exp\left(4n(n-1) \sum_{k=n+1}^{\infty} \infty (k(k-1))^{-1}\right) = \exp(4(n-1)). \tag{8.15}$$

Then we get by the Markov inequality:

$$P[\#R_t \geq n] = P[S_n \geq t] \leq e^{-\lambda_n t} E[e^{\lambda_n S_n}] \leq \exp(-\frac{n(n-1)}{4} + 4(n-1)). \tag{8.16}$$

(C) With the help of the preceding we can bound the estimation of $\#R_t$:

$$\begin{aligned} E[\#R_t] &= \sum_{n=1}^{\infty} P[\#R_t \geq n] \\ &\leq \sum_{n=1}^{\lfloor 16/t \rfloor} P[\#R_t \geq n] + \sum_{n=\lfloor 16/t+1 \rfloor+1}^{\infty} P[\#R_t \geq n] \\ &\leq \frac{16}{t} + \sum_{n=\lfloor 16/t \rfloor+1}^{\infty} e^{-(n-1)}. \end{aligned} \tag{8.17}$$

An estimate of the last sum yields:

$$E[\#R_t] \leq \frac{16}{t} + 3 \tag{8.18}$$

□

STEP 2: *A precompact set Γ_γ in $\mathbb{T}^{\text{root}}$*

The construction of Γ_γ is the next point to do now. Lemma 2.5 from [EW06] will be used, which states that a set in $\mathbb{T}^{\text{root}}$ is precompact, if it has an ϵ-net, i.e. for each tree in this set, there are at most $n(\epsilon)$ points in each tree, s.t. every point in the tree is covered by balls with radius ϵ around these $n(\epsilon)$ points.

Let a set of trees and $\epsilon > 0$ be given. First cut the trees horizontally into slices with height ϵ. Each slice contains a certain number of different arc pieces. Out of each arc piece choose a single point. To get precompactness of this set of trees by Lemma 2.5 from [EPW06] it is sufficient that the number of these chosen points is bounded uniformly by a constant in the set of trees. This is in fact the way to construct the compact subset Γ_γ.

First some notation is introduced. Therefore remember the definitions of the tree-height h, the cut-operator Q_t, the leaf-operator ∂Q_t and the trimming operator S_ϵ in Chapter 3 (Definition 3.2.1). Then define $A^t_{t-\epsilon}(T,\rho)$ to be the *set of ancestors* at time $t-\epsilon$ of the individuals alive at time t, i.e.:

$$A^t_{t-\epsilon}(T,\rho) = S_\epsilon(Q_t(T,\rho)) \cap \partial Q_{t-\epsilon}(T,\rho). \tag{8.19}$$

Then we construct precompact sets as follows:

Lemma 8.3.3: *For a positive non-decreasing sequence of integers $(L^m)_{m\in\mathbb{N}}$, the set*

$$\Gamma := \{(T,\rho) \in \mathbb{T}^{\text{root}} : \sum_{k=0}^{\lfloor 2^{(m+1)} h(T,\rho)-1 \rfloor} \#A^{k2^{-(m+1)}}_{(k-1)2^{-(m+1)}}(T,\rho) \leq L^m \forall m \in \mathbb{N}\} \tag{8.20}$$

is precompact in $\mathbb{T}^{\text{root}}$.

PROOF: To prove this lemma we use the already mentioned Lemma 2.5 in [EPW06]. We use the following observation: If we choose $m_0 = \lceil -\log_2 \epsilon \rceil \vee M$, then $2^{-(m_0+1)} \leq \epsilon/2$. That means the slices corresponding to m_0 are thinner than $\epsilon/2$.

Now we will prove that for a given $(T, d, \rho) \in \Gamma$ the following set R is an ϵ-net of T:

$$R(\epsilon, T) = \bigcup_{k=0}^{\lfloor 2^{(m_0+1)} h(T,\rho) \rfloor - 1} A^{k2^{-(m_0+1)}}_{(k-1)2^{-(m_0+1)}}(T, \rho). \tag{8.21}$$

Given a point $x \in T$ and its height $h(x) = d(\rho, x)$, the point x must have an ancestor a_1 at the time $2^{-(m_0+1)} \lfloor 2^{(m_0+1)} h(x) \rfloor$ just before (on the floor of the x-slice). This ancestor clearly has another ancestor a_2 in the slice before at time $2^{-(m_0+1)} \lfloor 2^{(m_0+1)} h(x) - 1 \rfloor$. But this ancestor a_2 has children in the next slice, namely a_1. So for individual a_2 it holds that:

$$a_2 \in A^{2^{-(m_0+1)}(\lfloor 2^{(m_0+1)} h(x) \rfloor)}_{2^{-(m_0+1)} \lfloor 2^{(m_0+1)} h(x) - 1 \rfloor}(T, \rho), \tag{8.22}$$

and since slices are at most $\epsilon/2$:

$$x \in B_\epsilon(A^{2^{-(m_0+1)} \lfloor 2^{(m_0+1)} h(x) \rfloor}_{2^{-(m_0+1)} \lfloor 2^{(m_0+1)} h(x) - 1 \rfloor}(T, \rho)) \tag{8.23}$$

Thus all points lie in balls around the ϵ-net (x is in fact in the annulus between radius $\epsilon/2$ and ϵ). When thinking of the slices-idea then we left out the last slice, but this will be helpful further on, since the last slice can contain some problematic branching behaviour.

What remains to show now for the application of the cited lemma is that the number of points $n(\epsilon, T)$ in the ϵ-net $R(\epsilon, T)$ for each $T \in \Gamma$ is bounded uniformly by $n(\epsilon) < \infty$. But this is clear by definition of Γ, since $n(\epsilon, T) \leq L^{m_0}$. So for any given $\epsilon > 0$ choose $n(\epsilon) = L^{m_0}$. Hence the lemma is proved. □

STEP 3: *The look-down: From the tree to the coalescent*

The interesting thing to do now is to give a uniform estimate in k, m and n for the number $A^{k2^{-(m+1)}}_{(k-1)2^{-(m+1)}}(T)$, when $T = \xi^{\text{for},n}$. We get that estimate by the look-down from the top of the tree. Branching in forward time means coalescing in inverse time.

To calculate this estimate a concept developed by Donnelly and Kurtz (see [DK96, Section 3] using the Kingman coalescent introduced in Step 1 (see [Kin82]) is used. In their paper they consider a population of particles evolving in continuous time with branchings after exponential-one-times. Each of the particles is assigned a different level. They proof in Theorem 3.2 that the number of parental levels at time $t-u$ of the children levels at time t has the same distribution as the Kingman coalescent at time u. What we need here is just the same: we do not speak of levels, but of different individuals.

Here, however, the situation is a bit different. A branching event after time t, at time $t+u$ of a single particle occurs not at the first jump of $N(u)$ as in [DK96], but at the first jump of $N(n \int_t^{t+u} bx^n(s)\, ds)$.

Consider now the non-random time change $\nu^{t,n}(u)$ given by

$$\int_{t-\nu^{t,n}(u)}^{t} nbx^n(s)\,ds = u. \tag{8.24}$$

This change is strictly monotone in u since the integrand is positive as long as $t < \hat{t}^0$. Then for a Poisson process N it is trivially true that:

$$N(\int_{t-\nu^{t,n}(u)}^{t} nbx^n(s)\,ds) = N(\int_{t-u}^{t} ds). \tag{8.25}$$

Now fix $u, t > 0$, where $0 \vee (t - \int_0^t nbx^n(s)\,ds) < t-u < t$, which is always possible, for example if $0 < u < t - \delta$. Then

$$\begin{aligned} \#A^t_{t-\nu^{t,n}(u)}(\xi^{\text{for},n,x^n}) &= \text{number of levels at time } t-\nu^{t,n}(u), \\ &\quad \text{which have descendants at time } t, \text{ where branching} \\ &\quad \text{occurs with rate } nbx^n(s) \\ &= \text{number of levels at time } t-u, \\ &\quad \text{which have descendants at time } t, \text{ where branching} \\ &\quad \text{occurs with rate one} \\ &= \#R(u), \end{aligned} \tag{8.26}$$

where R is the Kingman coalescent and this last result was shown in Theorem 3.1 in [DK96]. Furthermore

$$(\nu^{t,n}(\cdot))^{-1}(\nu^{t,n}(u)) = u = \int_{t-\nu^{t,n}(u)}^{t} nbx^n(s)\,ds, \tag{8.27}$$

which implies

$$(\nu^{t}(\cdot))^{-1}(\tilde{u}) = \int_{t-\tilde{u}}^{t} nbx^n(s)\,ds. \tag{8.28}$$

Thus we get that

$$\#A^t_{t-\tilde{u}}(\xi^{\text{for},n,x^n}) = \#R(\int_{t-\tilde{u}}^{t} nbx^n(s)\,ds). \tag{8.29}$$

Now recall what we have done in Step 1. In Lemma 8.3.2 an estimate for $\mathbf{E}[\#R(t)]$ was given. By applying this result, we get

$$\mathbf{E}[\#A^t_{t-\tilde{u}}(\xi^{\text{for},n})] \le 16 \left(\int_{t-\tilde{u}}^{t} nbx^n(s)\,ds\right)^{-1} + 3. \tag{8.30}$$

Step 4: *Proof of Tightness via the previous steps*

Now to show the proposition we apply the previous results to show that (8.10) holds.

First we note that the height of the reactant tree is bounded by $\hat{t}^n$, the hitting time of zero of the catalyst sample path x^n. Since these hitting times converge

to the hitting time $\hat{t}$ of the diffusion catalyst, all of the reactant tree heights are bounded from above. We manipulate the expression on the left hand side by simple rewriting and the Markov inequality:

$$\begin{aligned}
\mathbf{P}[\xi^{\mathrm{for},n} \in \Gamma] == \mathbf{P}[& \sum_{k=1}^{\lfloor 2^{(m+1)} h(\xi^{\mathrm{for},n},\rho)-1 \rfloor} \#A^{2^{-(m+1)}k}_{2^{-(m+1)}(k-1)}(\xi^{\mathrm{for},n}) \ \leq L^m \forall m \in \mathbb{N}] \\
\geq & 1 - \sum_{m\geq 1} \mathbf{P}[\sum_{k=1}^{\lfloor 2^{(m+1)} \hat{t}^n-1 \rfloor} \#A^{2^{-(m+1)}k}_{2^{-(m+1)}(k-1)}(\xi^{\mathrm{for},n}) \geq L^m] \\
\geq & 1 - \sum_{m\geq 1} \frac{1}{L^m} \mathbf{E}[\sum_{k=0}^{\lfloor 2^{(m+1)} \hat{t}^n-1 \rfloor} \#A^{2^{-(m+1)}k}_{2^{-(m+1)}(k-1)}(\xi^{\mathrm{for},n})] \\
\geq & 1 - \sum_{m\geq 1} \frac{1}{L^m} \sum_{k=0}^{\lfloor 2^{(m+1)} \hat{t}^n-1 \rfloor} \mathbf{E}[\#A^{2^{-(m+1)}k}_{2^{-(m+1)}(k-1)}(\xi^{\mathrm{for},n})].
\end{aligned} \tag{8.31}$$

At this point we use on the right hand side above the result proven in the previous step:

$$\mathbf{P}[\xi^{\mathrm{for},n} \in \Gamma] \geq 1 - \sum_{m\geq 1} \frac{2^{m+1}\hat{t}^n}{L^m} \left(16 + 16 \left(\int_{2^{-(m+1)}(k-1)}^{2^{-(m+1)}k} nbx^n(s)\, ds \right)^{-1} \right). \tag{8.32}$$

To control this expression in n we have to use the convergence of the catalyst to the diffusive limit x. Therefore consider a fixed m and take any $\epsilon_m > 0$, such that

$$\epsilon_m < 1/2 \min_{0\leq s\leq \hat{t}^0-2^{-m}} x(s). \tag{8.33}$$

Then one can select a natural number $N_{\epsilon,m}$ such that $\forall r \leq \hat{t}^0 - 2^{-m}$ $\forall n \geq N_{\epsilon,m}$:

$$1/2 \min_{0\leq s\leq \hat{t}^0-2^{-m}} x(s) \leq x(r) - \epsilon_m < x^n(r). \tag{8.34}$$

Now choose L^m so big that it satisfies the following two equations:

$$L^m \geq 4^{(m+1)}\hat{t} \left(\frac{1+\gamma}{\gamma} \right)^m 16 \left((\min_{0\leq s\leq \hat{t}-2^{(m+1)}} bX_s)^{-1} + 1 \right) \tag{8.35}$$

and

$$L^m \geq 4^{(m+1)}\hat{t}^n \left(\frac{1+\gamma}{\gamma} \right)^m 16 \sup_{n\leq N_{\epsilon_m,m}} \left((\min_{0\leq s\leq \hat{t}-2^{(m+1)}} bx^n(s))^{-1} + 1 \right). \tag{8.36}$$

Then obviously for this fixed m it holds that in the case $n \geq N_{m,\epsilon_m}$:

$$\begin{aligned}
& \frac{4^{m+1}\hat{t}^n}{L^m} 16 \left(\left(\min_{0\leq s\leq \hat{t}^{0,n}} nbx^n(s)\, ds \right)^{-1} + 1 \right) \\
& \qquad \leq \frac{4^{m+1}\hat{t}^n}{L^m} 16 \left(\left(1/2 \min_{0\leq s\leq \hat{t}^0-2^{-m}} bX_s \right)^{-1} + 1 \right) \leq \\
& \qquad \leq \left(\frac{\gamma}{1+\gamma} \right)^m
\end{aligned} \tag{8.37}$$

and in the case $n < N_{m,\epsilon_m}$:

$$\frac{4^{m+1}\hat{t}^n}{L^m} 16 \left(\left(\min_{0 \le s \le \hat{t}^{0,n}} nbx^n(s)\, ds \right)^{-1} + 1 \right) \le \left(\frac{\gamma}{1+\gamma} \right)^m. \tag{8.38}$$

So we are done since then in both cases we obtain for any $n \in \mathbb{N}$:

$$\begin{aligned} \mathbf{P}[\xi^{\text{for},n} \in \Gamma] &\ge 1 - \sum_{m \ge 1} \frac{2^{2m+2}\hat{t}^n}{L^m} \left(\min_{0 \le s \le \hat{t}^0 - 2^{-m}} nbx^n(s) \right)^{-1} \\ &\ge 1 - \sum_{m \ge 1} \left(\frac{\gamma}{1+\gamma} \right)^m = 1 - \gamma. \end{aligned} \tag{8.39}$$

So we have shown that with probability $> 1 - \gamma$ the reactant forest $\zeta^{\text{for},n}$ stays within a precompact set. Hence the sequence $(\zeta^{\text{for},n})_{n \in \mathbb{N}}$ is tight.

After finishing all four steps we are done with the proof of Proposition 8.3.1. □□□

Next we would like to show that there exists a unique limit. This task will take several steps and will start with results about the contour process.

8.4 Convergence of the truncated reactant contour

This section is devoted to show Theorem 5.3.1 about the convergence of the reactant contour process. Since contour and tree are related via the mappings $\mathcal{C}$ and $\mathcal{T}$ this is the next step to come closer to an asymptotic result about trees.

8.4.1 The main result and the strategy of the proof

In order to obtain a result about a diffusion-limit of the contour-process we introduce some further notation. The key point is to cut off the tree once the catalyst is close to zero, since there the reactant branching behavior slows down significantly. Hence we define a δ-hitting time of the catalyst. For $\delta > 0$ set:

$$\begin{aligned} \hat{t}^{n,\delta} &= \inf\{t \ge 0 : x^n(t) \le \delta\} \le \hat{t}^n < \infty, \\ \hat{t}^{\delta} &= \inf\{t \ge 0 : x(t) \le \delta\} \le \hat{t} < \infty. \end{aligned} \tag{8.40}$$

By the cut-operator Q_t the cut reactant tree, when the catalyst falls below δ, is given by:

$$\tilde{\xi}^{n,\delta} = Q_{\hat{t}^{n,\delta}}(\tilde{\xi}^n). \tag{8.41}$$

The theorem we are going to show is in a slightly more general setting than the original Theorem 5.3.1. The reactant tree is traversed with a general positive speed kn in the n-th approximation step. In the notation we omit further reference to the dependence of C on k. In the original setting of the theorem it was $k = 1$ and this will come out to be a good choice for some applications. So the cut reactant contour with traversal speed kn is given by:

$$C^{n,\delta} := \mathcal{C}(\tilde{\xi}^{n,\delta}; kn). \tag{8.42}$$

Then it is possible to establish the following convergence result of the reactant contour, when the catalyst is given as x^n and x respectively:

Theorem 8.4.1 (Reactant limit contour)**:**
Consider the linear operator $(A^\delta, \mathbb{D}(A^\delta))$*, where*

$$A^\delta f(c) = \frac{1}{2}\left(\frac{f'}{\frac{b}{2k}x_c}\right)'(c), \tag{8.43}$$

defined for $f \in \mathbb{D}(A^\delta)$ *with*

$$\mathbb{D}(A^\delta) = \{f \in C^1([0,\tau^\delta],[0,\infty)) : h'|_{\{0,T^\delta\}} = 0, \frac{h'}{x} \in C^1_{[0,\infty)}[0,T^\delta]\}. \tag{8.44}$$

Then under the hypotheses (8.1.1)*, the following holds:*

(i) The $(A^\delta, \mathbb{D}(A^\delta))$ *martingale problem is well-posed and*

(ii) if ζ^δ *is the solution of the* $(A^\delta, \mathbb{D}(A^\delta))$ *martingale problem, then:*

$$\mathcal{L}[(C^{n,\delta}_u)_{0\le u\le L(C^{n,\delta},\frac{4}{b})}] \xrightarrow{n\to\infty} \mathcal{L}[(\zeta^\delta_u)_{0\le u\le (l^0(\zeta^\delta))^{-1}(\frac{4}{b}))}], \tag{8.45}$$

where convergence is weak convergence of continuous processes.

To show this theorem, first it is necessary to get a clearer description of the discrete contour process. Therefore it is augmented to a two-dimensional process, the second coordinate being the sign of the slope of the contour. This $\mathbb{R}^2$-valued process is Markovian and a generator can be identified up to the time the tree is traversed. The same is done for rescaled and cut contour processes and we forget about the traversal time $L(C, 4/b)$ for a while. So we think of the Markov process as given by its generator. With the stochastic averaging technique one can show tightness of the rescaled reactant contours and identify one limit process. This limit process is shown to be unique. In the next step the traversal times of the rescaled contours are analysed and compared with the local time of the limit. A last step puts together the ideas.

The formal proof of this theorem is quite long and therefore needs to be split in several steps according to the following program:

Step 1: Discrete (contour, slope of contour)-process (C^1, V^1) is shown to be Markovian in $[0, \hat{t}^1] \times \{-1, 1\}$ and its generator is identified.

Step 2: The δ-hitting time cut (contour, slope of contour)-process $(C^{1,\delta}, V^{1,\delta})$ is shown to be Markovian in $[0, \hat{t}^{1,\delta}] \times \{-1, 1\}$ and its generator is identified.

Step 3: The rescaled, δ-hitting time cut (contour, sign of slope of contour)-process $(C^{n,\delta}, V^{n,\delta})$ is shown to be Markovian in $[0, \hat{t}^{n,\delta}] \times \{-1, 1\}$ and its generator is identified.

Step 4: Using Stochastic averaging ([Kur92]) the sequence $(C^{n,\delta})_{n\in\mathbb{N}}$ is tight and limit points solve the $(A^\delta, \mathbb{D}(A^\delta))$- martingale problem.

Step 5: Uniqueness of the $(A^\delta, \mathbb{D}(A^\delta))$- martingale problem is established.

Step 6: Traversal-time $L(C^{n,\delta}, \frac{4}{b})$ and inverse of local time of ζ^δ are compared.

Step 7: Here the pieces are put together to observe the claim.

The proof is not put into a "proof environment", but held within these steps. Some lemmas and propositions will arise, but all will be belonging to the corresponding parts of the proof.

For the understanding of the catalytic setting and the quenched analysis we refer the reader to the notation at the beginning of this chapter, see page 68.

8.4.2 The proof

STEP 1: *Identification of the (C_u, V_u)-generator*
The contour process $(C_u)_{u\geq 0}$ is obtained by traversing the reactant forest ξ^{for} with constant speed k and denoting the height, i.e. the distance of the root and the traversed point against the yet passed traversing time (for a definition see Section 5.1). This yet passed traversing time will be denoted with small Latin letters u or v in contrast to the time of the branching process, which we so far denoted with s or t.

We define the slope (sign) $(V_u)_{u\geq 0}$ of the contour process:

$$V_u := \text{sign}\ (\ \text{slope}(C_u)) \in \{-1, 1\}. \tag{8.46}$$

So far the "sign" could have been substituted by $\frac{1}{k}$, since slopes are $+k$ and $-k$, but later it will be helpful to already introduce it like that. There are some points, where the slope is not well defined, exactly, when the contour changes its sign from 1 to -1 or the other way round. Then V is chosen to be right continuous at these points, so that it has càdlàg paths. The pairing $(C_u, V_u)_{u\geq 0}$ is then a process in the state space

$$E^1_{\text{cont}} \times E_{\text{slope}} = [0, \hat{t}^1] \times \{-1, 1\} \tag{8.47}$$

and the following lemma holds, which is just the extended version of Lemma 5.2.4:

Lemma 8.4.2: *The process $(C_u, V_u)_{u\geq 0}$ is a $E^1_{\text{cont}} \times E_{\text{slope}}$-valued Markov-process stopped at a random time. Its generator is given by the closure of the operator $(A^1, \mathbb{D}(A^1))$, where:*

$$\mathbb{D}(A^1) = \{h \in C^{1,0}(E_{\text{cont}} \times E_{\text{slope}}, \mathbb{R}) : \frac{\partial h}{\partial c}|_{\partial E_{\text{cont}} \times E_{\text{slope}}} \equiv 0\} \tag{8.48}$$

and for $f \in \mathbb{D}(A^1)$:

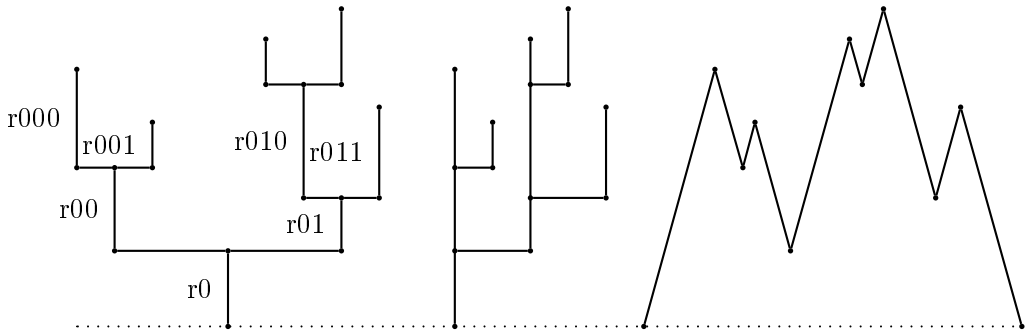
$$A^1 f(c, v) = kv\frac{\partial}{\partial c} f(c, v) + \frac{b}{2} k x^1(c)(f(c, -v) - f(c, v)). \tag{8.49}$$

The random time is reached when the reactant tree has been traversed, i.e. until $L(C, \frac{4}{b})$ given by:

$$L(C, \tfrac{4}{b}) = \inf\{u \geq 0 : \lim_{\epsilon \to 0} \frac{1}{\epsilon} \int_0^u \mathbb{1}_{\{C_v \in [0,\epsilon)\}} \frac{2k}{b x^1(v)}\, dv = \tfrac{4}{b}\}. \tag{8.50}$$

After that time the reactant contour stays zero.

Figure 8.1: Branching tree, birth-and-death tree, contour process



PROOF: The proof of this lemma consists itself of several parts. First it is shown that the lengths of line-segments of the contour are independent. After this the Markov property and the pre-generator are obtained.
A: *Indepence of the line-segments.* At a first glance the independence of contour line segments seems surprising, when looking at the branching tree on the left and the contour on the right in Figure 8.1. A contour line segment is the sum of the lifetimes of various reactant individuals. The trick is to look at and to understand the tree related to the contour (via the mappings $\mathcal{C}$ and $\mathcal{T}$) and its planar embedding in several different ways:

- We were dealing so far with a branching process whose individuals have a branching rate of $x^1(t) = b\eta_t^{\text{tot}}$ at time t with 0 or 2 offspring (left of Figure 8.1).
- The other way is to look at the process as a birth- death -process whose individuals die at rate $\frac{b}{2}x^1(t) = \frac{b}{2}\eta_t^{\text{tot}}$ and during all their lifetime have birth event (with offspring 1)with rate $\frac{b}{2}x^1(t) = \frac{b}{2}\eta_t^{\text{tot}}$, everything independent of other individuals (middle of Figure 8.1).

In the Appendix B.3 we show that the contour process related to these two processes are indeed the same.

But the planar embeddings can be made different: The right (birth and death) process will be called the one where children are attached to the right of the tree. Now we will show that when traversing the genealogical tree, the length of the newly attached line-segments are independent of the history. First we will not worry about the case a contour touches the critical upper and lower levels $T^{1,0}$ and 0:

Each (birth and death) individual I lives until the first jump of a Poisson process $M_d^I(\frac{b}{2}\int_{\tilde{t}}^{\cdot} x^1(s)\,ds)$ and gives birth at the jumps of an independent process $M_b^I(\frac{b}{2}\int_{\tilde{t}}^{\cdot} x^1(s)\,ds)$ (the subscript b means birth and has nothing to do with the branching rate b). Here $\tilde{t}$ denotes the birth time of individual I.

We will start with the first individual starting at the root. Let us then call:

$$\begin{aligned}
\theta &:= \inf\{t \geq 0 : M_d^I(\tfrac{b}{2}\int_0^t x^1(s)\,ds) = 1\}\text{the death time,}\\
\sigma_0 &:= 0,\\
\sigma_n &:= \inf\{t \geq \sigma_{n-1} : M_b^I(\tfrac{b}{2}\int_{\sigma_{n-1}}^t x^1(s)\,ds) = 1\}\text{ the n-th birth event and}\\
\Lambda &:= M_b^I(\tfrac{b}{2}\int_0^\theta x^1(s)\,ds)\text{ the number of children.}
\end{aligned} \tag{8.51}$$

The contour process starts with a slope $+1$ line segment of length θ. Then the length of the next line segment is independent since it is given by the first jump of an independent (reverse) Poisson process $M_b^I(\frac{b}{2}\int_{\theta-t}^\theta x^1(s)\,ds)$ in s. This distribution is independent of θ, since it does not matter for its jump time distribution, where we start a Poisson process (here in θ) and the direction it runs (here time-inverse). So in the contour the length of the first -1 line segment is independent of the past.

After this birth event of the last child, let us call it Λ, has assembled two Poisson processes, one for birth events and one for the death event. Both are independent among themselves and of anything previous, since it is a new individual. The slope $+1$ line segment is indepedent of anything before just by construction.

Now comes the key observation: The birth times of this individual Λ are given by the jump times of the M_b^Λ-process. We again let it run backward from the killing time, which does not affect the distribution. If the first jump time lies above the level of the birth of individual Λ then we can proceed in our argumentation as before and get independence of the -1 line segment. But if the first jump lies below the level of the birth of individual Λ then it has no children. That is clear. But for the contour process it means that it goes down until the next child of the initial individual, down to $\Lambda - 1$.

Then if individual Λ was killed at time θ^Λ, then the length of the -1 line segment is given by the first jump of the following process in s:

$$M_b^\Lambda(\tfrac{b}{2}\int_{\theta_\Lambda - s\vee\sigma_\Lambda}^{\theta^\Lambda} x^1(s)\,ds) + M_b^I(\tfrac{b}{2}\int_{\theta^\Lambda - s}^{\sigma_\Lambda\vee(\theta_\Lambda - s)} x^1(s)\,ds). \tag{8.52}$$

But the segment of the Poisson process M_b^I we plugged in, was unused before, so it is independent of all previously constructed elements. That is why this -1 line segment is independent of the past. One can continue like this until the end of the tree.

Now we need to be a bit more careful about what happens if jumps of the inverse birth process happen below level (= time) zero. We stop the previously constructed bunch of line segments when its building together ends up below zero. Then the contour process is just decreasing to zero with slope -1 and stops there. Then the tree is traversed.

Conversely if an individual reaches the upper level $\hat{t}^1$, then the death time of the individual is just $\hat{t}^1$. The next attached line-segment is an independent decreasing line-segment. Its time-inverse Poisson process starts from that level $\hat{t}^1$ on and is independent of all previous line-segments.

So we have shown that the length of new line segments are independent of the ones before. Then it is clearthat $(C_u, V_u)_{u\geq 0}$ is Markovian, since the whole history of $(C_u, V_u)_{0\leq u\leq v}$ up to time v has the same information for further development as just knowing the value (C_v, V_v). This is true, since the evolution of C in a short time interval after time v is just depending if we are actually running up or down in the contour, i.e. in the previous setting if we are on a birth (M_b) or on a death (M_d) Poisson process. Additionally by the value of C_v we know where to start the process, which gives the starting time in the time change integral of the Poisson process.

B: *Markov property and pregenerator.* Now in a second argument we want to identify the generator of this Markov process (C, V). How do the components evolve? C_u in fact simply grows linearly with slope kV_u. So its part of the generator is simple. The slope V_u is a jump-process with values 1 and -1. The jump rate is the individual death rate at this point at this point of the tree (where the contour currently traverses it). But this is $\frac{b}{2}x^1(C_u)$ and multiplying it with the speed of traversion k we get $\frac{b}{2}kx^1(C_u)$, since C_u is the distance root to the currently traversed point. So the generator A^1 for $f \in C^{1,0}(E^1_{\text{cont}} \times E_{\text{slope}})$ is given by:

$$A^1 f(c, v) = kv\frac{\partial}{\partial c}f(c, v) + \tfrac{b}{2}kx^1(c)(f(c, -v) - f(c, v)). \tag{8.53}$$

The domain of A^1 will need some restrictions, since when the contour reaches 0 or the possible maximal tree top $\hat{t}^1$, then it should be reflected not to leave the domain E_{cont}. The right way to do this is to restrict the domain of A^1 to:

$$\mathbb{D}(A^1) = \{h \in C^{1,0}_{E^1_{\text{cont}}\times E_{\text{slope}}}[0, \infty) : \frac{\partial h}{\partial c}|_{\partial E^1_{\text{cont}}\times E_{\text{slope}}} \equiv 0\}. \tag{8.54}$$

Since $\mathbb{D}(A^1)$ is dense in $C_{E^1_{\text{cont}}\times E_{\text{slope}}}[0, \infty)$ and A^1 is the sum of a closable operator (first summand) and a perturbation (second summand), also A^1 is closable (see e.g. Theorem 1.7.1. in [EK86]).

Clearly the time $L(C, \frac{4}{b})$ gives the time, when the tree is traversed once. □

Remark 8.4.3:
Sometimes we will suppress the superscript 1, *or later* n *for the contour state space* E_{cont} *for the sake of readibility.*

Secondly we will not worry about the time $L(C, \dot{})$ *in the forthcoming until Step 6 and will treat the contour process as a process traversing not only one tree (forest in the next steps) and staying zero, but traversing an infinite set of random trees (or forests).* ◇

STEP 2: *Identification of the* $(C^{1,\delta}, V^{1,\delta})$ *generator*
The same what was shown for the reactant tree or contour cut at height $\hat{t}^1$ is also true for tree and reactant cut at height $\hat{t}^{1,\delta}$. This process is called $C^{1,\delta}$ and to it a slope process $V^{1,\delta}$ is attached, the same way as in step 1:

$$\begin{aligned} C^{1,\delta} &:= \mathcal{C}(Q_{T^{1,\delta}}(\tilde{\xi}^1) : k), \\ V^{1,\delta}_u &:= \text{sign }(\text{ slope}(C^{1,\delta}_u)) \in \{-1, 1\}. \end{aligned} \tag{8.55}$$

The only difference is the state space of the contour changes from

$$E_{\text{cont}} = [0, \hat{t}^1] \text{ to } E^{1,\delta}_{\text{cont}} = [0, \hat{t}^{1,\delta}] \tag{8.56}$$

and we have:

Lemma 8.4.4: *The process* $(C^{1,\delta}_u, V^{1,\delta}_u)_{u\geq 0}$ *is a* $E^{1,\delta}_{\text{cont}} \times E_{\text{slope}}$*-valued Markov-process stopped at a random time. Its generator is given by the closure of the operator* $(A^{1,\delta}, \mathbb{D}(A^{1,\delta}))$*, where:*

$$\mathbb{D}(A^{1,\delta}) = \{h \in C^{1,0}(E_{\text{cont}} \times E_{\text{slope}}, \mathbb{R}) : \frac{\partial h}{\partial c}|_{\partial E^{1,\delta}_{\text{cont}} \times E_{\text{slope}}} \equiv 0\} \tag{8.57}$$

and for $f \in \mathbb{D}(A^1)$*:*

$$A^{1,\delta} f(c, v) = kv \frac{\partial}{\partial c} f(c, v) + \frac{b}{2} k x^1(c)(f(c, -v) - f(c, v)). \tag{8.58}$$

The random time is reached when the reactant tree has been traversed, i.e. until $L(C, \frac{4}{b})$ *given by:*

$$L(C, \tfrac{4}{b}) = \inf\{u \geq 0 : \lim_{\epsilon \to 0} \frac{1}{\epsilon} \int_0^u \mathbb{1}_{\{C^{1,\delta}_v \in [0,\epsilon)\}} \frac{2k}{b x^1(v)} \, dv = \tfrac{4}{b}\}. \tag{8.59}$$

After that time the reactant contour stays zero.

This result does require a proof, but the proof would just be the same. Or one can state that the cutting time $\hat{t}^1$ was chosen arbitrarily in Lemma 8.4.2 and did not play any role yet. So it can also be replaced by $\hat{t}^{1,\delta}$.

Step 3: *Identification of the generator of the* $(C^{n,\delta}, V^{n,\delta})$
In this step a rescaled version of the lemma proven in Step 2 will be given. For that purpose we recall the definition of the δ-hitting times $\hat{t}^{n,\delta}$ of x^n and the rescaled cut contour process $(C^{n,\delta}_u)_{u\geq 0}$ in (8.42). The slope of this process is kn and as before a sign of slope process $(V^{n,\delta}_u)_{u\geq 0}$ is attached the same way as it was with the non-rescaled process:

$$V^{n,\delta}_u := \text{sign}(\text{slope}(C^{n,\delta}_u)) \in \{-1, 1\} \tag{8.60}$$

The state space of the contour E_{cont} is given by

$$E^{n,\delta}_{\text{cont}} = [0, \hat{t}^{n,\delta}] \tag{8.61}$$

and the slope state space E_{slope} stays the same:

$$E_{\text{slope}} = \{-1, 1\} \tag{8.62}$$

Then the following lemma holds:

Lemma 8.4.5: *The process* $(C^{n,\delta}_u, V^{n,\delta}_u)_{u\geq 0}$ *is a* $E^{n,\delta}_{\text{cont}} \times E_{\text{slope}}$*-valued a Markov-process stopped at a random time. Its generator is given by the closure of the operator* $(A^{n,\delta}, \mathbb{D}(A^{n,\delta}))$*, where*

$$\mathbb{D}(A^{n,\delta}) = \{h \in C^{1,0}(E^{n,\delta}_{\text{cont}} \times E_{\text{slope}}, \mathbb{R}) : \frac{\partial h}{\partial c}|_{\partial E^{n,\delta}_{\text{cont}} \times E_{\text{slope}}} \equiv 0\} \tag{8.63}$$

and for $f \in \mathbb{D}(A^{n,\delta})$:

$$A^{n,\delta} f(c,v) = knv \frac{\partial}{\partial c} f(c,v) + \tfrac{b}{2} kn^2 x^n(c)(f(c,-v) - f(c,v)). \tag{8.64}$$

The random time is reached when the reactant forest has been traversed, i.e. until $L(C^n, \frac{4}{b})$ given by:

$$L(C^n, \tfrac{4}{b}) = \inf\{u \geq 0 : \lim_{\epsilon \to 0} \frac{1}{\epsilon} \int_0^u \mathbb{1}_{\{C_v^{n,\delta} \in [0,\epsilon)\}} \frac{2k}{bx^1(v)}\, dv = \tfrac{4}{b}\}. \tag{8.65}$$

After that time the reactant contour stays zero.

For the sake of not repeating the same proof as of Lemma 8.4.2 another time, argumentation is kept sparse here. Just two things are mentioned:

First we explain where the factors n and n^2 come from. The factor n dates from the changed traversal speed through the reactant tree as given in definition 5.2.3. The factor n^2 splits up in one factor "n", which multiplied with $x^n(c)$, so $nx^n(c)$, represents the number of catalyst individuals speeding up branching. The other factor n belongs to the increased traversal speed through the tree. Additionally we see that after traversing a tree of the forest the next tree is independent of everything previous.

Secondly the contour is finished when the contour has already traversed the n trees in the reactant forest. But since the slope of the line segment is kn the random time was chosen correctly.

The pairing $(C_u^{n,\delta}, V_u^{n,\delta})_{u \geq 0}$ can in some sense also be seen as a random evolution with driving process $V^{n,\delta}$ and driven process $C^{n,\delta}$. More about that can be found in Chapter 9 of [EK86].

Step 4: *Tightness of the contour and description of limit contours*
Now clearly the next task is to show convergence and to obtain a limit of the process $(C^{n,\delta}, V^{n,\delta})$ for $n \to \infty$. This is not straightforward, since standard results as for example about generator convergence cannot be used, e.g. Proposition 7.7.2 here or Lemma 4.5.1 in [EK86]. But what will happen for large n is that the slope "averages" out and will be positive or negative with equal probability. These techniques were introduced in [Kur92] and are called *stochastic averaging*. This step is subdivided in two parts, where the first one corresponds to Theorem 2.1 and the second one to Example 2.3, both in Kurtz's paper.

***A:** the limit contour solves a martingale problem* Let us define the following occupation times measure on $\mathcal{B}(\mathbb{R}_+) \times \mathcal{P}(\{-1,1\})$ for the slope process, where $y \in E_{\text{slope}} = \{-1,1\}$ by:

$$\Gamma^{n,\delta}([0,u] \times y) = \int_0^u \mathbb{1}_{\{y\}}(V_v^{n,\delta})\, dv. \tag{8.66}$$

Additionally $l_m(E_{\text{slope}})$ is defined to be the set of measures on $\mathcal{B}(\mathbb{R}_+) \times \mathcal{P}(\{-1,1\})$, s.t. for every $\mu \in l_m(E_{\text{slope}})$ and every $u \geq 0$:

$$\mu([0,u] \times E_{\text{slope}}) = u. \tag{8.67}$$

Then the following lemma holds:

Lemma 8.4.6: *The sequence* $((C_u^{n,\delta}, \Gamma^{n,\delta})_{u\geq 0})_{n\in\mathbb{N}}$ *is relatively compact in the space* $\mathcal{D}_{E_{\text{cont}}}[0,\infty)\times l_m(E_{\text{slope}})$*, and for any limit point* $(\zeta^\delta, \Gamma^\delta)$ *there exists a filtration* $\{\mathcal{G}_t\}$ *such that*

$$f(\zeta_u^\delta) - \int_0^u \int_{E_{\text{slope}}} A^\delta f(\zeta_s^\delta, y)\Gamma^\delta(ds \times dy), \tag{8.68}$$

is a $\{\mathcal{G}_u\}$*-martingale for each* $f \in \mathbb{D}(A^\delta)$ *where* A^δ *is as in Theorem 8.4.1.*

Proof: The result is just taken from Theorem 2.1 in [Kur92], where one has to check some prerequisites before applying this theorem. We will put them in the order as given in the quoted theorem to avoid confusion:

Clearly $E_{\text{cont}}^{n,\delta} = [0, \hat{t}^{n,\delta}]$ and E_{slope} are complete separable metric spaces. The sample paths of the process $(C^{n,\delta}, V^{n,\delta})$ are càdlàg by definition. Additionally $(C^{n,\delta})_{n\in\mathbb{N}}$ satisfies the compact containment condition since $\hat{t}^{n,\delta} \to \hat{t}^\delta < \infty$ and therefore the union $\bigcup_{n\in\mathbb{N}}[0, \hat{t}^{n,\delta}]$ is contained in a compactum. The sequence $(V^{n,\delta})_{n\in\mathbb{N}}$ is relatively compact, since its state space E_{slope} is already compact. After these preliminary thoughts the harder things to prove are the following four assertions:

(i) For each $f \in \mathbb{D}(A^\delta)$ there is a process $(\epsilon_u^{f,n,\delta})_{u\geq 0}$ for which the following expression is a $\{\mathcal{F}_u^n\}$-martingale:

$$f(C_u^{n,\delta}) - \int_0^u A^\delta f(C_s^{n,\delta}, V_s^{n,\delta})\, ds + \epsilon_t^{f,n,\delta}, \tag{8.69}$$

where $\{\mathcal{F}_u^n\}$ is the filtration generated by $(C^{n,\delta}, V^{n,\delta})$.

(ii) The domain of A^δ, $\mathbb{D}(A^\delta)$, is dense in $C_b(E_{\text{cont}}^\delta, \mathbb{R})$.

(iii) For each $f \in \mathbb{D}(A^\delta)$ and each $T > 0$, there exists $p > 1$ such that

$$\sup_n \mathbf{E}\left[\int_0^T |A^\delta f(C_u^{n,\delta}, V_u^{n,\delta})|^p\, du\right] < \infty. \tag{8.70}$$

(iv) The correction process $\epsilon^{f,n,\delta}$ vanishes in the following sense:

$$\lim_{n\to\infty} \mathbf{E}\left[\sup_{u\leq T} |\epsilon_u^{f,n,\delta}|\right] = 0. \tag{8.71}$$

The third expression is clear, since the state spaces are contained in a compactum ($\hat{t}^{n,\delta} \to \hat{t}^\delta$!) and $A^\delta f$ is continuous. For the first, second and fourth expression more effort is necessary. The first and the fourth expression are somehow related so we begin with them:

We want to construct $\epsilon^{f,n,\delta}$ and therefore note that all $f \in \mathbb{D}(A^\delta)$ can be written with a continuously differentiable function $h \in C^1_{[0,\infty)}[0, \hat{t}^\delta]$ with $h(0) = h(\hat{t}^\delta) = 0$:

$$f(c) = f(0) + \int_0^c x(s)h(s)\, ds, \tag{8.72}$$

for all $0 \leq c \leq \hat{t}^\delta$. The idea now is to define a function which is "close" to f, but lies in the domain of $A^{n,\delta}$ instead of A^δ. Therefore set for $f \in \mathbb{D}(A^\delta)$

$$\tilde{f}^n(c) = f(0) + \int_0^c x^n(s)h(\frac{\hat{t}^{n,\delta}}{\hat{t}^\delta}s)\, ds. \tag{8.73}$$

Now the boundary conditions on $\tilde{f}^n$ already coincide with the ones for functions in $\mathbb{D}(A^{n,\delta})$ and it remains to augment $\tilde{f}^n$ with a second coordinate. Hence define another function

$$f^n(c,v) = \tilde{f}^n(c) + \frac{v}{bnx^n(c)}(\tilde{f}^n)'(c), \tag{8.74}$$

which is in the domain of the operator $A^{n,\delta}$ and by applying this operator to that function one obtains:

$$\begin{aligned} A^{n,\delta}f^n(c,v) =& knv(\tilde{f}^n)'(c) + \frac{knv^2}{bn}\left(\frac{(\tilde{f}^n)'}{x^n(c)}\right)'(c) - \\ &- \frac{bk}{2}n^2x^n(c)\frac{2v}{bnx^n(c)}(\tilde{f}^n)'(c) \\ =& \left(\frac{(\tilde{f}^n)'}{\frac{k}{b}x^n(c)}\right)'(c). \end{aligned} \tag{8.75}$$

Then the correction process $\epsilon^{f,n,\delta}$ is defined as:

$$\epsilon^{f,n,\delta}_u = \int_0^u (A^\delta f - A^{n,\delta}f^n)(C^{n,\delta}_s, V^{n,\delta}_s)\,ds + \tag{8.76}$$

Clearly by Lemma 8.4.5 the expression in (8.69) is a martingale and assertion (i) is shown.

Now we prove the fourth statement and call in the definition of $\epsilon^{f,n,\delta}$ the first summand with the integral *part (I)* and the other three summands *part (II)*. We will show that both of them vanish in the sense of (8.71):

$$\begin{aligned} \mathbf{E}\left[\sup_{u\le T}|\epsilon^{f,n,\delta}_u|\right] \le& T\|A^\delta f - A^{n,\delta}f^n\|_{c\in[0,\hat{t}^\delta]} + \\ &+ \|\tilde{f}^n(c) - f(c) + \frac{v}{bnx^n(c)}(\tilde{f}^n)'(c)\|_{(c,v)\in[0,\hat{t}^\delta]\times\{-1,1\}}. \end{aligned} \tag{8.77}$$

For part (I) we get by using the triangle-inequality:

$$\begin{aligned} (I) =& \|A^\delta f - A^{n,\delta}f^n\|_{[0,\hat{t}^\delta]} \le T\|h'(c) - \left(h(\frac{\hat{t}^{n,\delta}}{\hat{t}^\delta}\cdot)\right)'(c)\| \le \\ \le& T\|h'(c) - \frac{\hat{t}^{n,\delta}}{\hat{t}^\delta}h'(c)\| + T\frac{\hat{t}^{n,\delta}}{\hat{t}^\delta}\|h'(c) - h'(\frac{\hat{t}^{n,\delta}}{\hat{t}^\delta}c)\|. \end{aligned}$$

And this expression vanishes for $n\to\infty$ by Corollary 2.3.2 and the uniform continuity of h' as it vanishes outside a compactum by Lemma 2.2.4. Then it remains to treat part (II). It is

$$\begin{aligned} (II) \le& \int_0^{\hat{t}^\delta} |x^n(s)h(\frac{\hat{t}^{n,\delta}}{\hat{t}^\delta}s) - x(s)h(s)|\,ds + \frac{1}{bn}\|h\| \\ \le& \int_0^{\hat{t}^\delta} |x^n(s) - x(s)|\,|h(\frac{\hat{t}^{n,\delta}}{\hat{t}^\delta}s)|\,ds + \\ &+ \int_0^{\hat{t}^\delta} |x(s)|\,|h(\frac{\hat{t}^{n,\delta}}{\hat{t}^\delta}s) - h(s)|\,ds + \frac{1}{bn}\|h\|. \end{aligned}$$

Since $\hat{t}^{n,\delta} \to \hat{t}^{\delta}$, (8.1.1) and the continuity of h all terms vanish for $n \to \infty$.

For condition (ii) it is sufficient to approximate all piecewise linearly continuous functions on $[0, \hat{t}^{\delta}]$, since they are already dense in $C_b(E^{\delta}_{\text{cont}}, \mathbb{R})$. But any piecewise linear function can be approximated arbitrarily close, since x is bounded below by δ.

□

B: *the limit occupation measure of V is identified.* After this proof the question is whether the rather complicated expression in the previous lemma with the measure Γ can be simplified. The following lemma gives the answer:

Lemma 8.4.7: *The process ζ^{δ} of the previous lemma is a solution of the $(A^{\delta}, \mathbb{D}(A^{\delta}))$-martingale problem.*

PROOF: Note that Lemma 1.4 from [Kur92] shows us that the measure $\Gamma^{\delta}(ds \times dy)$ can be decomposed into the measure $\gamma^{\delta}_s(dy)\, \lambda(ds)$, where λ is the Lebesgue measure. Then what we need to do is to determine $\gamma^{\delta}_s(dy)$.
The idea of the proof is taken from Example 2.3 in the same work. Therefore define an operator $B : C(\{-1,1\}, \mathbb{R}) \to C([0, \hat{t}^{\delta}] \times \{-1,1\}, \mathbb{R})$, where

$$h(v) \mapsto Bf(c,v) = \frac{bk}{2} x(c)(h(-v) - h(v)). \tag{8.78}$$

As before it is necessary to define a correction term $\alpha^{h,n,\delta}_t$ in this way:

$$\alpha^{h,n,\delta}_t = \frac{bk}{2} n^2 \int_0^t \left(x(C^{n,\delta}_s) - x^n(C^{n,\delta}_s) \right) h(V^{n,\delta}_s)\, ds. \tag{8.79}$$

Then by Lemma 8.4.5, for each $h \in C(\{-1,1\}, \mathbb{R})$ the following process is a martingale

$$h(V^{n,\delta}_t) - \int_0^t n^2 Bh(C^{n,\delta}_s, V^{n,\delta}_s)\, ds + \alpha^h_{n,\delta}(t), \tag{8.80}$$

and for each $T > 0$:

$$\begin{aligned}
&\lim_{n\to\infty} \mathbf{E}\left[\sup_{t \le T} n^{-2} |\alpha^{h,n,\delta}_t| \right] = \\
&= \lim_{n\to\infty} \mathbf{E}\left[\frac{bk}{2} \sup_{t\le T} n^{-2} n^2 \Big| \int_0^t x(C^{n,\delta}_s) - x^n(C^{n,\delta}_s) h(V^{n,\delta}_s)\, ds \Big| \right] \le \\
&\le \lim_{n\to\infty} \|h\|_{\infty}\, T\, \frac{bk}{2} \mathbf{E}\left[\sup_{s>0} |x(C^{n,\delta}_s) - x^n(C^{n,\delta}_s)| \right] \le \\
&\le \lim_{n\to\infty} \|h\|\, T\, \frac{bk}{2} \|x(c) - x^n(c)\|_{\infty, c \in [0, \hat{t}^{\delta} \vee \hat{t}^{n,\delta}]} = 0.
\end{aligned} \tag{8.81}$$

The second modulus expression tends to zero by (8.1.1).

Note that for any $n \in \mathbb{N}$ the following expression is still a martingale:

$$n^{-2} h(V^{n,\delta}_t) - \int_0^t Bh(C^{n,\delta}_s, V^{n,\delta}_s)\, ds + n^{-2} \alpha^h_{n,\delta}(t). \tag{8.82}$$

Then by taking weak limits $n \to \infty$ (and changing the sign) we get with the help of the previous Lemma 8.4.6 the following expression:

$$\begin{aligned} &\int_0^t \int_{\{-1,1\}} Bh(\zeta^\delta, y)\, \Gamma^\delta(ds \times dy) = \\ &= (h(1) - h(-1))\frac{bk}{2} \int_0^t (1 - 2\gamma_s^\delta(1))x(\zeta^\delta)ds. \end{aligned} \tag{8.83}$$

Since we are in a compact metric state space and we have càdlàg paths, this weak limit is still a martingale, now w.r.t. the filtration of ζ^δ. Additionally this process is continuous, since it is an integral expression. Furthermore, the total variation of this continuous martingale is finite, because it is an integral expression. Therefore it must already be constant and because of the initial value, this constant must be zero. Hence, for all $t \geq 0$ it is valid that:

$$(h(1) - h(-1)) \int_0^t (1 - 2\gamma_s^\delta(1))x(\zeta^\delta)ds = 0. \tag{8.84}$$

Now remember that the values of x were always required to be $\geq \delta$, because the catalyst was to be stopped before reaching δ. So the only way to get the left hand side expression equal to zero for any function h is to have:

$$\gamma_s^\delta(1) = \frac{1}{2} \; \forall s > 0. \tag{8.85}$$

Then the previous Lemma 8.4.6 can be reformulated by using the decomposition of Γ^δ and this result. Then the lemma is proven. □

Step 5: *Uniqueness of the Martingale problem*
In the previous step existence for the $(A^\delta, \mathbb{D}(A^\delta))$ problem was shown. In this step the uniqueness is shown. To do this firstly another martingale problem (transformation of drift) is constructed and then uniqueness of the latter is proved.
A: *transformation of drift* Define the *scale function* s, given by

$$s : \begin{cases} [0, \hat{t}^\delta] & \to [0, \int_0^{\hat{t}^\delta} x(u)du] \\ t & \mapsto \frac{b}{2k} \int_0^t x(u)du \end{cases}, \tag{8.86}$$

and the process $B_u = s(\zeta_u^\delta)$ are defined. Then for a sufficiently regular f, when c denotes the replacement character:

$$\begin{aligned} A^\delta(f \circ s)(\zeta_u^\delta) &= \frac{k}{b}\left(\frac{(f \circ s)'}{x(c)}\right)'(\zeta_u^\delta) \\ &= \frac{k}{b}\left(\frac{(f' \circ s(c))s'}{x(c)}\right)'(\zeta_u^\delta) \\ &= \left(\frac{(f' \circ s(c))x(c)}{2x(c)}\right)'(\zeta_u^\delta) \\ &= \frac{1}{2}((f'' \circ s)(c)s'(c))(\zeta_u^\delta) = \frac{b}{4k}f''(B_u)x(\zeta_u^\delta) \\ &= \frac{b}{4k}f''(B_u)x(s^{-1}(B_u)). \end{aligned} \tag{8.87}$$

Thus $(B_u)_{u \geq 0}$ is a martingale under the law of ζ^δ. Even the following lemma holds:

Lemma 8.4.8: *If ζ^δ solves the $(A^\delta, \mathbb{D}(A^\delta))$ martingale problem, then the process $(B_u = s(\zeta_u^\delta))_{u\geq 0}$ solves the $(\mathcal{B}, \mathbb{D}(\mathcal{B}))$ problem, where*

$$\mathcal{B}f = \frac{b}{4k} x(s^{-1}(\cdot)) f''(\cdot),$$

defined for $f \in \mathbb{D}(\mathcal{B})$, where

$$\mathbb{D}(\mathcal{B}) = \{H \in C^2([0, s(\hat{t}^\delta)], \mathbb{R}) : H'|_{\{0, s(\hat{t}^\delta)\}} \equiv 0\}.$$

Proof: Let $H \in \mathbb{D}(\mathcal{B})$, then

- $H \circ s : [0, \hat{t}^\delta] \to [0, \infty]$ and this mapping is C^1 since it is composed of two C^1-mappings.
- $(H \circ s)'|_{\{0,\hat{t}^\delta\}} \equiv 0$.
- $H' \circ s \in C^1$ since both mappings are C^1 mappings.

So the composition $H \circ s$ is in the domain of the operator A^δ and we can do a similar calculation as done before the lemma to obtain:

$$A^\delta (H \circ s)(u) = \frac{b}{4k} H''(s(u)) x(u). \tag{8.88}$$

Since ζ^δ solves the A^δ martingale problem, we have that the following process is a martingale:

$$\begin{aligned} &\left(H \circ s(\zeta_t^\delta) - \int_0^t A^\delta H \circ s(\zeta_u^\delta) du \right)_{t\geq 0} = \\ &= \left(H(B_t) - \int_0^t \frac{b}{4k} x(s^{-1}(B_u)) H''(B_u) du \right)_{t\geq 0}. \end{aligned} \tag{8.89}$$

So $(B_t)_{t\geq 0}$ solves the $(\mathcal{B}, \mathbb{D}(\mathcal{B}))$ problem. □

***B:** uniqueness of the martingale problem* The last lemma is not yet the desired result for this step, but it directly leads to:

Lemma 8.4.9: *The $(A^\delta, \mathbb{D}(A^\delta))$ martingale problem has a unique solution ζ^δ.*

Proof: The existence was already given in Step 4. Assume the contrary about the uniqueness, i.e. assume there are two solutions $\zeta^{\delta,1}$ and $\zeta^{\delta,2}$. Then by the previous lemma $s(\zeta^{\delta,1})$ and $s(\zeta^{\delta,2})$ are two solutions of the $(\mathcal{B}, \mathbb{D}(\mathcal{B}))$-martingale problem. But by Corollary IX 1.14 in [RY91] the $(\mathcal{B}, \mathbb{D}(\mathcal{B}))$-martingale problem is well-posed, so uniqueness holds. Therefore it is valid that:

$$\int_{[0,\zeta^{\delta,1}]} x(u)\, du = \int_{[0,\zeta^{\delta,2}]} x(u)\, du. \tag{8.90}$$

So the one-dimensional distributions of $\zeta^{\delta,1}$ and $\zeta^{\delta,2}$ are the same, since x is a positive function. Then by Theorem 4.4.2 in [EK86] the two processes $\zeta^{\delta,1}$ and $\zeta^{\delta,2}$ have the same distribution in $C_{[0,\hat{t}^\delta]}[0, \infty)$. Furthermore this theorem also states that the unique ζ^δ is a Markov process. □

STEP 6: *Analysis of the contour traversal time* $L(C^{n,\delta}, \frac{2}{k})$
So now we need to go back to considering contour processes as processes traversing trees. Since for the rescaled contours C^n the slope increased to n it is still truethat the traversal time of the reactant tree is given by:

$$L(C^{n,\delta}, \tfrac{4}{b}) = \inf\{u \geq 0 : \lim_{\epsilon \to 0} \frac{1}{\epsilon} \int_0^u \mathbb{1}_{\{C_v^{n,\delta} \in [0,\epsilon)\}} (x^n(v))^{-1} \, dv = \frac{4}{b}\}. \tag{8.91}$$

For the limit contour ζ^δ we wonder if a limit of this functional $L(\cdot, \frac{4}{b})$ exists and what it is. In fact if it turned out to be zero, then it would be the case that the reactant limit contour is defined on the degenerate interval $[0, 0]$ and it would contain no useful information.

But we see that the functional $L(\cdot, \cdot)$ on continuous functions $C([0, \infty), [0, \infty))$ is measurable with respect to an infinity-norm with decreasing weights ($d(f, g) = \sum_{n=1}^{\infty} 2^{-n} (\sup_{t \in [0,n]} |f(t) - g(t)| \wedge 1)$). Then we can use a version of Theorem 13.29 in [Bre68].

Hence we see that $L(C^{n,\delta}, \frac{4}{b})$ converges weakly to $L(\zeta^\delta, \frac{4}{b})$ and the latter is nothing else than the inverse of ζ^δ's local time at level 0 of local time $\frac{4}{b}$ (for more about local times of the reactant contour see 9.2.3):

$$\mathcal{L}\left[L(C^{n,\delta}, \frac{4}{b})\right] \xrightarrow{n \to \infty} \mathcal{L}\left[(l^0(\zeta^\delta))^{-1}(\frac{4}{b})\right]. \tag{8.92}$$

Hence the "duration" of the contour converges weakly.

STEP 7: *Putting the steps of the proof together*
In Step 3 the generators of the discrete contour processes $C^{n,\delta}$ were identified. In Step 4 we got tightness of the discrete contour processes and saw that the limit points were described by a pregenerator A^δ. The corresponding $(A^\delta, \mathbb{D}(A^\delta))$ martingale problem is well posed, i.e. a unique solution ζ^δ exists (Step 5). In the next step we showed convergence of the tree traversal time to a local time inverse (Step 6). So the tight sequence $(C^{n,\delta})_n$ must already converge to the diffusion ζ^δ, which is considered until its local time hits at zero hits a certain level. Hence we are done with the proof of Theorem 8.4.1.

With having completed all the steps we have shown convergence of the reactant contour. □□□

8.5 Convergence of the reactant limit forest

With all the preceding steps done it is not difficult any more to show Theorem 4.2.2, which is the main result about trees:

Proposition 8.5.1: *For any $\delta > 0$ there exists a subsequence of the sequence of cut reactant trees $(Q_{\hat{t}^{n,\delta}}(\xi^{\text{for},n_m}); \eta^{\text{tot},n_m})_{m \in \mathbb{N}}$ that converges in distribution to $(Y^{\text{for},\delta}; X) = \mathcal{T}_{\text{unord}}(\zeta^\delta)$.*

PROOF: By Proposition 8.3.1 it is clear that one can take a subsequence of forests $(\xi^{\text{for},n_m}; \eta^{\text{tot},n_m})_{m \in \mathbb{N}}$ which converges in distribution to a random forest $(Y^{\text{for}}; X)$.

The truncation mappings $Q_\cdot(\cdot) : \mathbb{R}^+ \times \mathbb{T}^{\text{root}} \to \mathbb{T}^{\text{root}}$ do not violate this convergence, since they are jointly continuous and the killing times $t^{n,\delta}$ converge:

$$\mathcal{L}[Q_{\hat{t}^{n,\delta}}(\xi^{\text{for},n_m}); x^{n_m}] \xrightarrow{m\to\infty} \mathcal{L}[Q_{\hat{t}^\delta}(Y^{\text{for}}); x]. \tag{8.93}$$

This convergence is to be understood as weak convergence in the probability measures on rooted compact $\mathbb{R}$-trees.

But in Theorem 8.4.1 it was shown that there is also weak convergence in $\mathcal{D}_{\mathbb{R}}([0, \hat{t}^\delta])$ for the reactant contour:

$$\mathcal{L}[\mathcal{C}(Q_{\hat{t}^{n_m,\delta}}(\tilde{\xi}^{n_k}) : kn_m^2)] \xrightarrow{m\to\infty} \mathcal{L}[\zeta^\delta]. \tag{8.94}$$

The mapping $\mathcal{T}_{\text{unord}}$ from Definition 5.1.3 is continuous and bounded and therefore it holds that

$$\mathcal{L}[\mathcal{T}_{\text{unord}}(\mathcal{C}(Q_{\hat{t}^{n_m,\delta}}(\tilde{\xi}^{n_m}) : kn_m^2))] \xrightarrow{k\to\infty} \mathcal{L}[\mathcal{T}_{\text{unord}}(\zeta^\delta)]. \tag{8.95}$$

But it is true by definition of the reactant forest that

$$\mathcal{T}_{\text{unord}}(\mathcal{C}(Q_{\hat{t}^{n_m,\delta}}(\tilde{\xi}^{n_m}) : kn_m^2)) = \xi^{\text{for},n_m}. \tag{8.96}$$

Hence the result is shown. □

Proof: (Proof of Theorem 4.2.2: The reactant limit tree)
The question roughly speaking is whether there exists a limit object of $\mathcal{T}_{\text{unord}}(\zeta^\delta)$ for $\delta \to 0$. From Proposition 8.3.1 we know that the sequence $(\xi^{\text{for},n})_{n\in\mathbb{N}}$ is tight. But from the previous proposition we know that any limit forest Y^{for} needs to fulfil $\mathcal{L}[Q_{\hat{t}^\delta}(Y^{\text{for}})] = \mathcal{L}[\mathcal{T}_{\text{unord}}(\zeta^\delta)]$, so this law is independent of the limit forest chosen.

Clearly we need to show that there is a unique limit forest Y^{for}. To do so we show that the sequence $(\mathcal{L}[\mathcal{T}_{\text{unord}}(\zeta^\delta])_{\delta>0}$ is a Cauchy sequence as $\delta \to 0$ for the Prohorov metric on the probability measures $\mathcal{M}_1(\mathbb{T}^{\text{root}})$. Thus since $\mathcal{M}_1(\mathbb{T}^{\text{root}})$ is complete , there is a limit forest distribution $\mathcal{L}[Z]$. Moreover for *any* limit forest Y^{for}, the sequence $(\mathcal{L}[Q_{\hat{t}^\delta}(Y^{\text{for}})])_{\delta>0}$ is Cauchy with the *same* limit distribution $\mathbb{P} = \mathcal{L}[Z]$. So the distribution of Y^{for} is uniquely determined and given by the distribution of $\lim_{\delta\to 0} \mathcal{T}_{\text{unord}}(\zeta^\delta)$.

To verify that $(\mathcal{L}[\mathcal{T}_{\text{unord}}(\zeta^\delta)])_{\delta>0}$ is Cauchy , fix $0 < \delta' < \delta$ and recall a characterization of the Prohorov metric (Theorem 3.1.2 in [EK86]). Then:

$$\begin{aligned} &d_{\text{Pr}}(\mathcal{L}(\mathcal{T}_{\text{unord}}(\zeta^\delta), \mathcal{L}(\mathcal{T}_{\text{unord}}(\zeta^{\delta'}))) = \\ &= \inf_\mu \inf\{\epsilon > 0 : d_\mu(((T,\rho),(T',\rho')) : d_{GH^{root}}((T,\rho),(T',\rho')) > \epsilon) \le \epsilon\}, \end{aligned} \tag{8.97}$$

where the infimum for μ is taken over all probability measures $\mu \in \mathcal{M}_1(\mathbb{T}^{\text{root}} \times \mathbb{T}^{\text{root}})$, with marginal distributions

$$\mu(\cdot \times \mathbb{T}^{\text{root}}) = \mathcal{L}(\mathcal{T}_{\text{unord}}(\zeta^\delta)) \tag{8.98}$$

and

$$\mu(\mathbb{T}^{\text{root}} \times \cdot) = \mathcal{L}(\mathcal{T}_{\text{unord}}(\zeta^\delta)). \tag{8.99}$$

As our goal is to give an upper bound for this distance, it will be sufficient to describe a specific probability measure μ fulfilling the properties and showing that the right hand side expression for this specific μ vanishes as $\delta, \delta' \to 0$. We choose μ to be the probability measure on $\mathbb{T}^{\text{root}} \times \mathbb{T}^{\text{root}}$, s.t. positive measure is associated to those tree pairs $((T,\rho),(T',\rho'))$, which have the same root and the same truncated excursion $\mathcal{C}(Q_{\hat{t}^\delta}(\cdot))$. Additionally the marginals need to obey the postulated properties. By this a measure μ on $\mathbb{T}^{\text{root}} \times \mathbb{T}^{\text{root}}$ is well-defined.

Note that in the Gromov-Hausdorff-distance between a pair of trees from the support of μ is clearly less than $\hat{t}^{\delta'} - \hat{t}^{\delta}$. Thus

$$d_{\text{Pr}}\left(\mathcal{L}(\mathcal{T}(\zeta^\delta), \mathcal{L}(\mathcal{T}(\zeta^{\delta'}))\right) \leq \hat{t}^{\delta'} - \hat{t}^{\delta}. \tag{8.100}$$

Furhermore, since $\hat{t}^\delta \to \hat{t}$ as $\delta \to 0$, the sequence is Cauchy. □

8.6 Convergence of the joint law for the reactant forest

Within this section we provide a proof of Theorem 4.2.3. This is the first non-quenched result in this chapter. Be aware that we leave the world of fixed catalyst sequences x and x^n here.

8.6.1 The main result and the strategy of the proof

We restate Theorem 4.2.2 which gives the annealed tree convergence:

Theorem 8.6.1:
The sequence of the pair of rescaled catalyst total mass and rescaled reactant forest converges:

$$\mathcal{L}(\eta^{\text{tot},n}, \xi^{\text{for},n}) \Rightarrow \mathcal{L}(X, Y^{\text{for}}). \tag{8.101}$$

Here convergence is understood as weak convergence on the set of probability measures on $\mathcal{D}_{\mathbb{R}^1_+}[0,\infty) \times \mathbb{T}^{\text{root}}$ with the product topology.

We are going to split up the proof in four steps, since it is rather long. A part of one step is left in the appendix, since it is a general result and its proof can be done easier when using some ideas developed in the chapter about point processes.

For the proof we assume that the limit law is the product measure of the limit catalyst total mass times the limit quenched probability kernels. We will start with considering the first coordinate, the catalyst total mass process, and restrict its paths and the time, where it is positive to a compactum. In a second step we take a Lipschitz-function

$$H : \mathcal{D}_{\mathbb{R}^2_+}[0,\infty) \times \mathbb{T}^{\text{root}} \to \mathbb{R} \tag{8.102}$$

and need to show the ordinairy weak convergence argument. Since we cannot do that directly we apply the quenched reactant tree convergence and the catalyst total mass convergence to reduce the problem to a question of continuity of a mapping depending on the catalyst. In the third step we relate trees and contours via $\mathcal{T}_{\text{unord}}$. Therefore, the question translates to continuous dependence of the contour on its catalyst, which is nothing else than the derivative of the scale function. In the last step we show that diffusions depend continuously on their scale function under certain restrictions and this will end the proof.

Step 1: Restrictions of catalyst total mass $\eta^{\text{tot},n}$, X to compacta.

Step 2: Rewriting the claim to a question of continuity.

Step 3: Translation of the continuity question to a contour question of continuity.

Step 4: Continuity of the contour diffusion on its scale function.

All $\epsilon, \tilde{\epsilon}$ mentioned later are positive and fixed. When writing ϵ_n, then we want to indicate that n needs to be chosen appropriately to get ϵ. So n depends on ϵ in that case.

8.6.2 The proof

Step 1:
We want to speak about the law of the catalyst total mass processes. Therefore we define for a set of functions A in $\mathcal{D}_{\mathbb{R}_+}[0,\infty)$

$$\|A\| := \sup_{f\in A} \|f\| \tag{8.103}$$

and

$$T(A) := \sup_{f\in A}\{t \geq 0 : f(t) = 0\}. \tag{8.104}$$

Then we can bound the measure $\mathbf{P}(\eta^{\text{tot},n} \in \cdot)$ in the following way:

$$\begin{aligned}\mathbf{P}(\eta^{\text{tot},n} \in A) \leq \quad & \mathbf{P}(\eta^{\text{tot},n} \in A, \|A\| \leq M, T(A) \leq M) \\ & + \mathbf{P}(\eta^{\text{tot},n} \in A, \|A\| > M, T(A) \leq M) \\ & + \mathbf{P}(\eta^{\text{tot},n} \in A, T(A) > M).\end{aligned} \tag{8.105}$$

Observe that by Lemmas 7.3.3 and 7.3.1 the last two summands can be bounded from above by any positive constant for an appropriate choice of M.

The same can be done for the measure $\mathbf{P}(X \in \cdot)$ on $C_{\mathbb{R}_+}[0,\infty)$ by Lemmas 7.6.3 and 7.4.4.

Hence let us define

$$\nu_n(dx) := \mathbf{P}[\eta^{\text{tot},n} \in dx, \|x\| \leq M, T(x) \leq M] \tag{8.106}$$

and

$$\nu(dx) := \mathbf{P}[X \in dx, \|x\| \leq M, T(x) \leq M]. \tag{8.107}$$

Then we can write for given $\epsilon > 0$ and an appropriate $M = M(\epsilon)$ the following

$$\begin{aligned}\mathbf{P}(\eta^{\text{tot},n} \in dx) &\leq \nu_n(dx) + \epsilon_M, \\ \mathbf{P}(X \in dx) &\leq \nu(dx) + \epsilon_M.\end{aligned} \tag{8.108}$$

Step 2:
For a continuous Lipschitz continuous function

$$H : \mathcal{D}_{\mathbb{R}_+^2}[0,\infty) \times \mathbb{T}^{\text{root}} \to \mathbb{R} \tag{8.109}$$

we need to show the following:

$$\int d\mathbf{P}[\eta^{\mathrm{tot},n} = x, \xi^{\mathrm{for},n} = y]H(x,y) \xrightarrow{n\to\infty} d\mathbf{P}[X = x, Y^{\mathrm{for}} = y]H(x,y). \quad (8.110)$$

We rewrite this using the quenched probability kernels $K^n(\cdot,\cdot)$ and $K(\cdot,\cdot)$ defined in (8.2) and (8.3):

$$\left| \int d\mathbf{P}(X = x) \int K(x,dy)H(x,y) - \int d\mathbf{P}(\eta^{\mathrm{tot},n} = x) \int K^n(x,dy)H(x,y) \right| \quad (8.111)$$

This can be bounded by a decomposition as follows:

$$\begin{aligned} &\leq \left| \int d\mathbf{P}(X = x) \int K(x,dy)H(x,y) - \int d\mathbf{P}(\eta^{\mathrm{tot},n} = x) \int K(x,dy)H(x,y) \right| \\ &+ \left| \int d\mathbf{P}(\eta^{\mathrm{tot},n} = x) \int K(x,dy)H(x,y) - \int d\mathbf{P}(\eta^{\mathrm{tot},n} = x) \int K^n(x,dy)H(x,y) \right|. \end{aligned}$$

The second summand tends to zero by the quenched result about tree convergence in Theorem 4.2.2 (for a fixed catalyst sequence). When choosing n sufficiently high it can be achieved to bound it by a positive ϵ_n.

With the help of part (A) the first summand can be bounded from above by

$$2\|H\|\,\epsilon_M + \left| \int \nu(dx) \int K(x,dy)H(x,y) - \int \nu_n(dx) \int K(x,dy)H(x,y) \right| \quad (8.112)$$

and we rewrite that to

$$2\|H\|\,\epsilon_M + \left| \int \nu(dx)\hat{H}(x) - \int \nu_n(dx)\hat{H}(x) \right|, \quad (8.113)$$

where

$$\hat{H}(x) = \int K(x,dy)H(x,y). \quad (8.114)$$

Thus, by the total mass convergence ($\nu_n \Rightarrow \nu$ in Theorem 2.3.1) we are done if we can show that $\hat{H}$ is bounded and continuous (in Skorokhod topology). The first is clear, since H is bounded and we are dealing with probability kernels. To show continuity in x is more subtle and we use two parts to do that. First we relate the problem to contour processes.

Step 3:
For two càdlàg catalyst paths x and $\tilde{x}$ observe that:

$$\begin{aligned} |\hat{H}(x) - \hat{H}(\tilde{x})| \leq &| \int K(x,dy)H(x,y) - \int K(x,dy)H(x,Q_{\tau^\delta}y)| \\ &+ | \int K(x,dy)H(x,Q_{\tau^\delta}y) - \int K(\tilde{x},dy)H(\tilde{x},Q_{\tau^\delta}y)| \\ &+ | \int K(\tilde{x},dy)H(\tilde{x},Q_{\tau^\delta}y) - \int K(\tilde{x},dy)H(\tilde{x},y)|. \end{aligned} \quad (8.115)$$

Since H is Lipschitz-continuous, note that the first and the third line can be bounded from above by any possible constant $\tilde{\epsilon}$ if we choose δ sufficiently small. Now we remember (Proposition 8.5.1) that we were able to express the cut trees with limit the contour process via the Lipschitz-continuous mapping $\mathcal{T}_{\text{unord}}$:

$$Q_{\tau^\delta}(Y^{\text{for}}) \stackrel{d}{=} \mathcal{T}_{\text{unord}}(\zeta^\delta). \tag{8.116}$$

Hence

$$|\hat{H}(x)-\hat{H}(\tilde{x})| \leq 2\delta\|H\|+|\mathbf{E}[H(x,\mathcal{T}_{\text{unord}}(\zeta^\delta))|X=x]-\mathbf{E}[H(\tilde{x},\mathcal{T}_{\text{unord}}(\zeta^\delta))|X=\tilde{x}]|.$$

The continuity in the first coordinate is evident and giving up the first coordinate we can define $G := H \circ \mathcal{T}_{\text{unord}}$. Clearly this mapping is Lipschitz-continuous since both factors are. We will show continuity not only for mappings G of that special form but for any kind of Lipschitz mappings G. Therefore we can also deal with the following mapping

$$G : (C_{[0,M]}[0,\infty), \|\cdot\|_{\sup}) \to \mathbb{R} \tag{8.117}$$

and we need to take care of

$$\hat{G} : \begin{cases} \mathcal{D}_{[0,M]}[0,M] & \to \mathbb{R} \\ x & \mapsto \mathbf{E}[G(\zeta^\delta)|X=x] \end{cases}. \tag{8.118}$$

We note that we are done when showing continuity of that mapping by (8.117).

Step 4:
Let us call $\zeta^{\delta,x}$ and $\zeta^{\delta,\tilde{x}}$ the contours corresponding to catalysts x and $\tilde{x}$. By Theorem 5.3.1 the limit contour was the solution of a martingale problem. Hence we can define scale-function and a random time-change to relate the contour to a Brownian motion β (this becomes clearer in Section 9.2, especially Lemma 9.2.2). Hence, define

$$s_x(t) = \frac{b}{2}\int_0^t x(r)\,dr s_{\tilde{x}}(t) = \frac{b}{2}\int_0^t \tilde{x}(r)\,dr \tag{8.119}$$

Now we use a result proven in the appendix (Chapter B.2) about the difference of two diffusions stemming from two different scale functions. The time until we need to consider the two diffusions is almost surely finite and hence we can use the result presented in the appendix. Therefore, we can choose the two functions x and $\tilde{x}$ so close that the following expression can be bounded by:

$$\begin{aligned} |\hat{G}(x)-\hat{G}(\tilde{x})| \leq & \mathbf{E}\left[|G(\zeta^{\delta,x})-G(\zeta^{\delta,\tilde{x}})|\right] \\ \leq & \|G\|\,\mathbf{E}\left[\|\zeta^{\delta,x}-\zeta^{\delta,\tilde{x}}\|\right] \\ \leq & \|G\|\,\delta^{-3}\epsilon. \end{aligned} \tag{8.120}$$

This ends the proof since we were able to establish continuity of the mapping $\hat{G}$ in Skorokhod topology by (8.118).

After this result the proof is done. □□□

Remark 8.6.2:
Naïvely one could guess that $\nu_n \Rightarrow \nu$ *and* $K^n \Rightarrow K$ *implies*

$$\nu_n \times K^n \Rightarrow \nu \times K. \tag{8.121}$$

But we have seen in the proof that this is a question of continuous dependence of some elements.

9 Proofs of the main results from Chapter 6

In this chapter we give the proofs concerning the genealogical point processes.

The proofs for the first rectant point process results and its limit are put in a first section. They are mostly done by classical Poisson approximation ideas. A second section inhabits a proposition treating the relationship between limit point process and limit contour. The proof of this proposition needs quite some effort and uses excursion theory. The third section of this chapter contains a proof of a comparison result between classical and catalytic forests. And in the last section we give a proof of results about joint-convergence of the trees and limit-contour extension above level τ^δ. These last proofs are put here at the end of this chapter, since either they need notation from here or they require the genealogical point-process as a full description of the process.

Before starting we give a remark.

Remark 9.0.3:
Within this chapter there might arise confusion about the word "time" or "time index". We will talk about

- *time level or simply level if we mean the height in the tree (e.g. t in η_t^{tot}) and*
- *the term "time", if we mean the time index in the diffusion process to the genealogical tree (e.g. u in ζ_u^δ).* ◇

9.1 Law and convergence of the reactant point process

First we provide the proof of a lemma about the reactant point process.

Lemma 9.1.1: *For a fixed catalyst $(\eta_s^{\mathrm{tot}})_{s\geq 0}$ realization and fixed $t < T^{1,0}$ the reactant point process Ξ^t has total mass*

$$\Xi^t(\mathbb{N}\times[0,t]) = \xi_t^{\mathrm{tot}} - 1. \tag{9.1}$$

The point process Ξ^t is given by the random points $\{(i,\sigma_i) : 1 \leq i \leq \xi_t^{\mathrm{tot}} - 1\}$, where the σ_i are independent and identically distributed $[0,t]$-valued random-variables. They have distribution given by

$$P(\sigma_1 \geq h) = \frac{\frac{2}{b} + \int_0^t \eta_s^{\mathrm{tot}}\,ds}{\int_0^t \eta_s^{\mathrm{tot}}\,ds}\,\frac{\int_h^t \eta_s^{\mathrm{tot}}\,ds}{\frac{2}{b} + \int_h^t \eta_s^{\mathrm{tot}}\,ds}, \tag{9.2}$$

for every $0 < h < t$.

Proof: In the case of $\xi_t^{\mathrm{tot}} \leq 1$ nothing needs to be shown, since there are no MRCAs. So let us assumethat $\xi_t^{\mathrm{tot}} - 1 \geq 2$. Then the σ_i are the depths of a

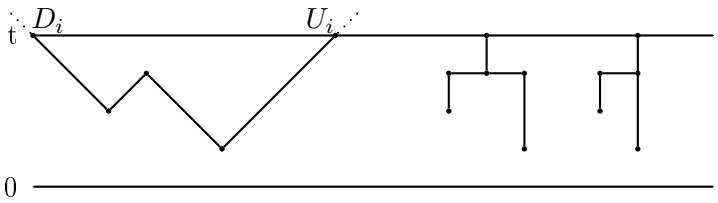


Figure 9.1: The downward excursion relates to an upward-down birth-and-death process.

downward excursion from level t of the contour process (C, V), as given in Lemma 5.2.4 cut at height $T^{1,0}$. This process was strong Markovian, so these downward-excursions are independent and identically distributed.
Let us define the times, when the contour C crosses the level t:

$$\begin{aligned} D_0 := 0; U_i := \inf\{s > D_i : C_s = t\} \text{ for } i \geq 0; \\ D_i := \inf\{s > U_{i-1} : C_s = t\} \text{ for } i \geq 1 \end{aligned} \tag{9.3}$$

are the upcrossing and downcrossing times of the contour. These times are the "contour times" of each individual alive at time t.

Now we have to think of the depth σ_i of this downward excursion between D_i and U_i. But we can look at this downward excursion as the contour of a downward tree as in Figure 9.1. The downward tree behaves like a birth-and-death process with rates $\frac{b}{2}\eta_s^{\text{tot}}$ for birth and death each (the rates were identified in the first part of the proof of Lemma 8.4.2). But the depth of a downward excursion is nothing else than the extinction time of this birth-and-death process.

An exercise in Feller's book ([Fel68, Problem XVII.10.11]) shows that a birth-and-death process $(N_s)_{s\geq 0}$, $N_0 = 1$ with rates $(\lambda_s)_{s\geq 0}$ for each is extinct at time s with the following probability:

$$P(N_s = 0) = \frac{\int_0^s \lambda_r \, dr}{1 + \int_0^s \lambda_r \, dr}. \tag{9.4}$$

In our situation it is $\lambda_s = \frac{b}{2}\eta_{t-s}^{\text{tot}}$ and we obtain:

$$\begin{aligned} P(\sigma_1 \geq h) &= P(N_{t-h} = 0 | N_t = 0) = \\ &= \frac{P(N_{t-h} = 0)}{P(N_t = 0)} = \\ &= \frac{1 + \frac{b}{2}\int_0^t \eta_s^{\text{tot}} \, ds}{\frac{b}{2}\int_0^t \eta_s^{\text{tot}} \, ds} \, \frac{\frac{b}{2}\int_h^t \eta_s^{\text{tot}} \, ds}{1 + \frac{b}{2}\int_h^t \eta_s^{\text{tot}} \, ds}, \end{aligned} \tag{9.5}$$

which had to be proved. □

Now we want to prove a result for the rescaled reactant point process. The situation is more complex, since two extant individuals can be members of different trees. Hence their MRCA lies at time level 0. This is reflected in Proposition 6.2.7 which we state here again and give a proof of it:

Proposition 9.1.2: *We can specify the distribution of the reactant point process* $\Xi^{t_n,n}$ *at time* t_n. *For* $k_n \in \{1, 2, \ldots, n\xi^{\mathrm{tot},n}_{t_n} - 1\}$

(i) the number of points at level 0 *is given by*

$$\kappa_n := \Xi^{t_n,n}(\{\tfrac{1}{n}, \tfrac{2}{n}, \ldots, \tfrac{k_n}{n}\} \times \{0\}) \stackrel{d}{=} \text{Bin } (k_n, \mathbb{P}(\sigma_n \geq t_n)) \text{ and} \tag{9.6}$$

(ii) the number of points between 0 *and* $t_n - h_n$ *is given by*

$$\Xi^{t_n,n}(\{\tfrac{1}{n}, \tfrac{2}{n}, \ldots, \tfrac{k_n}{n}\} \times (0, t_n - h_n)) \stackrel{d}{=} \text{Bin } (k_n - \kappa_n, \mathbb{P}(\sigma_n \geq h_n | \sigma_n < t_n)) . \tag{9.7}$$

Here $\mathrm{Bin}(n, p)$ *is the law of a binomially distributed random variable with parameters* n, p *and* σ_n *is the extinction time of a birth-and-death process with reproduction and death rate* $(\frac{nb}{2}\eta^{\mathrm{tot},n}_{t_n-s})_{0\leq s\leq t_n}$.

Proof: Since we have conditioned on the total mass of the reactant at time t_n, we know that there are $n\xi^{\mathrm{tot},n}_{t_n}$ reactant individuals alive at that time. Among the $n\xi^{\mathrm{tot},n}_{t_n} - 1$ most recent common ancestors of these individuals there are some which lie at time level zero and some which lie above. The first ones contribute to κ_n the others to the second line in the proposition. The idea, similar to the previous proof, is to look at the contour process and relate MRCAs to minimal points of the downward excursions.

First we note that there are $n\xi^{\mathrm{tot},n}_{t_n} - 1$ downward excursions from level t_n to consider. We already showed that the contour process reaches its minimum in a downward excursion at the extinction time of a birth-and-death process with rate $(\frac{b}{2}n\eta^{\mathrm{tot},n}_{t_n-s})_{s\geq 0}$ each. In the case $n = 1$ we knew that extinction would occur before the excursion reaches level zero. For the general n it is in fact the casethat if this extinction does not happen before reaching time level zero, then this means the contour goes to zero and then starts traversing the next tree of the forest. This gives a point at level zero.

As all of the excursions are independent the number of excursions "dropping" below level zero among the first k_n excursions is therefore given by:

$$\kappa_n := \Xi^{t_n,n}(\{\tfrac{1}{n}, \tfrac{2}{n}, \ldots, \tfrac{k_n}{n}\} \times \{0\}) \stackrel{d}{=} \text{Bin}\,(k_n, \mathbb{P}(\sigma_n \geq t_n)) . \tag{9.8}$$

Thus the first line is already proven. It remains to show the second line. There are now $k_n - \kappa_n$ most recent common ancestors that lie above time level zero. Each of these contours independent of the others has probability $\mathbb{P}(\sigma_n > h_n | \sigma_n < t_n)$ to "drop" at least below level $t_n - h_n$:

$$\Xi^{t_n,n}(\{\tfrac{1}{n}, \tfrac{2}{n}, \ldots, \tfrac{k_n}{n}\} \times (0, t_n - h_n)) \stackrel{d}{=} \text{Bin}\,(k_n - \kappa_n, \mathbb{P}(\sigma_n > h_n | \sigma_n < t_n)) . \tag{9.9}$$

And that is all we needed to show for the proposition. □

We continue clasically and want to prove a convergence result for $\Xi^{t_n,n}$ when $n \to \infty$ as Theorem 6.3.1. Its proof is given by the following theorem, when additionally mixing the (here) non-random reactant total masses at time t_n.

It will be a quenched result , i.e. we fix a catalytic background as in Definition 2.3.3. Additionally let a fixed time $t < \tau^0$ be given and a sequence of t_n approaching t as $n \to \infty$. Furthermore fix a sequence of reactant total masses s.t.:

$$\lim_{n\to\infty} |Y_t - \xi_{t_n}^{\text{tot},n}| = 0. \tag{9.10}$$

The next theorem uses these ingredients and shows convergence of the rescaled point process. Therefore condition the rescaled processes on a catalyst total mass process $\eta^{\text{tot},n}$ and reactant total mass $\xi_{t_n}^{\text{tot},n}$ at time t_n.

Then the following theorem holds.

Theorem 9.1.3:
The point process $\Xi^{t_n,n}$ converges to a point process π^t on $[0, Y_t] \times [0, \tau^0]$. For $u \in [0,1]$ and $0 < h < t$ the limit point process is given by:

- *a Poisson process at level zero, i.e. on $[0, Y_t] \times \{0\}$ specified by*

$$\pi^t([0, uY_t] \times \{0\}) = \text{Poisson}\left(\frac{2uY_t}{\int_0^t bX_s\, ds}\right) \tag{9.11}$$

- *and another Poisson process on the set $[0, Y_t] \times (0, \tau^0]$ specified by*

$$\pi^t([0, uY_t] \times (0, h)) = \text{Poisson}\left(uY_t\left(\frac{2}{\int_h^t bX_s\, ds} - \frac{2}{\int_0^t bX_s\, ds}\right)\right) \tag{9.12}$$

PROOF: By Theorem 4.2 in [Kal83] it is sufficient to show the convergence of $\Xi^{t_n,n}$ to π^t by considering intensity measures on the sets given in the theorem. Both lines will be proven via the usual Poisson approximation and we will only do the first line, since the proof of the second line uses the same ideas.
Therefore let $u \in [0,1]$ be given:

$$\begin{aligned}
&\Xi^{t_n,n}([0, uY_t] \times \{0\}) = \\
&\quad = \Xi^{t_n,n}([0, u\xi_{t_n}^{\text{tot},n}] \times \{0\}) + \\
&\qquad \pm \Xi^{t_n,n}([uY_t \wedge u\xi_{t_n}^{\text{tot},n}, uY_t \vee u\xi_{t_n}^{\text{tot},n}] \times \{0\}) \\
&\quad = \Xi^{t_n,n}\left(\{\tfrac{1}{n}, \tfrac{2}{n}, \ldots, \tfrac{\lfloor nu\xi_{t_n}^{\text{tot},n}\rfloor}{n}\} \times \{0\}\right) + \\
&\qquad \pm \Xi^{t_n,n}\left(\{\tfrac{1}{n}, \tfrac{2}{n}, \ldots, \tfrac{\lfloor nu|Y_t - \eta_{t_n}^{\text{tot},n}| - 1\rfloor}{n}\} \times \{0\}\right) \\
&\quad \overset{d}{=} \text{Bin}\left(\lfloor nu\xi_{t_n}^{\text{tot},n}\rfloor, \mathbb{P}(\sigma_n \geq t_n)\right) + \\
&\qquad \pm \text{Bin}\left(\lfloor nu|Y_t - \xi_{t_n}^{\text{tot},n}|\rfloor, \mathbb{P}(\sigma_n \geq t_n)\right).
\end{aligned} \tag{9.13}$$

Now it is helpful to remember that the probability $\mathbb{P}(\sigma_n \geq t_n)$ can be calculated similar as in Lemma 9.1.1 (especially (9.4)) and we get

$$\begin{aligned}
\mathbb{P}(\sigma_n \geq t_n) &= \mathbb{P}(\text{Poisson process with rate } (n\eta_{t_n - s}^{\text{tot},n})_{s\geq 0} \text{ has} \\
&\qquad \text{not jumped before time } t_n) \\
&= \frac{1}{1 + \int_0^{t_n} \frac{b}{2} n\eta_s^{\text{tot},n} ds}.
\end{aligned} \tag{9.14}$$

And in a limit for $n \to \infty$ we obtain that:

$$n\mathbb{P}(\sigma_n \geq t_n) \xrightarrow{n\to\infty} \frac{2}{b}\left(\int_0^t X_s ds\right)^{-1}. \tag{9.15}$$

And as $\xi_{t_n}^{\text{tot},n}$ goes to Y_t (remember that by conditioning they are not random) the first summand gives what we want by the Poisson approximation for $n \to \infty$. The second summand by Poisson approximation tends to unit mass at zero. Then as $n \to \infty$:

$$\Xi^{t_n,n}([0, uY_t] \times \{0\}) \to \text{Poisson}\left(\frac{2uY_t}{\int_0^t bX_s ds}\right) \tag{9.16}$$

□

9.2 The relationship between limit point process and limit contour

After the limit reactant one wonders, if there is a connection to the other functionals of the catalytic branching setting. We can affirmate such a question and prove Proposition 6.3.3, which relates the limit point process with the minima of downward excursions of the contour process. The proof is quite long and is subdivided into several steps. For the understanding of the proof we will use the slightly more general contour description as in Section 8.4 with tree traversal speed $= k$. The contour runs until local time hits $\frac{4}{b}$.

9.2.1 Main result and strategy of the proof

The proposition to prove is

Proposition 9.2.1: *Let a fixed catalyst X and $t < \tau^0$ be given. If $\delta > 0$ is such that $t < \tau^\delta$, then let ζ^δ denote the solution of the $(A^\delta, \mathbb{D}(A^\delta))$ martingale problem. Then it holds that*

$$\pi^t \overset{d}{=} \pi^{\zeta^\delta,t}. \tag{9.17}$$

Here the process on the right hand side is given as the point process of minima of downward excursions from level t:

$$\pi^{\zeta^\delta,t} := \{(u, \inf(\epsilon_u^-) :, \text{ when } \alpha_{u-}^t \neq \alpha_u^t \text{ and } u \leq \alpha_{4/b}\}, \tag{9.18}$$

where ϵ_u^- is a downward excursion of ζ^δ from level t and α_1 is the first time, when the local time at level zero reaches $\frac{4}{b}$ and α_u^t is the inverse of local time at level t of u (for the exact definitions see page 42).

The proof will be done in several steps. The first four preparatory steps are in a very general context that can be applied to any excursion question. The last step focuses on the contour process ζ^δ and its downward excursion point process and then relates it to the limit point process.

Step 1: Scale function and speed measure of a diffusion ζ and relation to Brownian motion.

Step 2: Local times of a diffusion ζ and relationship to Brownian local time.

Step 3: Excursion depths of a diffusion ζ: Depths coordinate.

Step 4: Excursion depths of a diffusion ζ: Time coordinate.

Step 5: Application of the previous result to the point process $\pi^{\zeta^\delta,t}$ and comparison with π^t.

9.2.2 The proof

STEP 1: *Scale function and speed measure of a diffusion ζ and relation to Brownian motion*
In this step we follow the description in [RY91, Chapter VII.3] and references refer to that book. Let $(\zeta_u)_{u\geq 0}$ be a regular diffusion on a compactum $[l,r]$ with generator $(A,\mathbb{D}(A))$. Then there exists a function s and a measure m:

$$\begin{aligned} s :&[l,r] \to \mathbb{R}, \\ m :&\mathcal{B}([l,r]) \to [0,\infty), \end{aligned} \tag{9.19}$$

called the scale function s and the speed measure m, s.t.

- s is continuous, strictly increasing and unique up to affine transformations (Proposition VII.3.1).
- $s(\zeta)$ is a local martingale.
- if m has density m' w.r.t. Lebesgue measure. Then it is

$$Af(x) = \frac{1}{2}\frac{d}{m'(x)dx}\frac{d}{ds}f(x), \tag{9.20}$$

  for bounded $f \in \mathbb{D}(A)$ and almost any $x \in (l,r)$ (Theorem VII.3.12).

If s is linear then ζ is said to be on the natural scalethat means it contains no drift. If the speed measure density m' is constantthat means the process is running at the same "speed" as Brownian motion. In the case $s = id$, $m = \lambda$, we have $\zeta = \beta$, where β is a Brownian motion.
To use the third property let us henceforth assume that m has density with respect to the Lebesgue measure λ, i.e. that we can write:

$$m(dx) = m'(x)\lambda(dx). \tag{9.21}$$

Now we want to start some calculations to express the distribution of a diffusion ζ, started in t, as a "transformed" Brownian motion β, started in 0. First we note that $B := (s(\zeta_u) - s(t))_{u\geq 0}$ is a martingale started in 0. As shown in Lemma 8.4.8 for special m and s, it is true that B is a solution of the martingale problem with generator

$$\tilde{A}f(x) = \frac{1}{2}\frac{s'(s^{-1}(x+s(t)))}{m'(s^{-1}(x+s(t)))}f''(x). \tag{9.22}$$

So for a Brownian motion β we can write

$$dB_u = \sqrt{\frac{s'(s^{-1}(B_u+s(t)))}{m'(s^{-1}(B_u+s(t)))}}\,d\beta_u. \tag{9.23}$$

To use the Martingale Representation Theorem set the random time change

$$\gamma(t) := \int_0^t \frac{m'(s^{-1}(\beta_v + s(t)))}{s'(s^{-1}(\beta_v + s(t)))}\, dv, \tag{9.24}$$

and its well-defined inverse

$$\gamma^{-1}(u) := \inf\{t \geq 0 : \gamma(t) \geq u\}. \tag{9.25}$$

Then an easy calculation as in [KS00, Proof of Theorem 5.5.4] gives that

$$(B_u)_{u\geq 0} \stackrel{d}{=} (\beta_{\gamma^{-1}(u)})_{u\geq 0}, \tag{9.26}$$

and even the following lemma holds:

Lemma 9.2.2: *It is true that*

$$(\zeta_u)_{u\geq 0} \stackrel{d}{=} (s^{-1}(\beta_{\gamma^{-1}(u)} + s(t)))_{u\geq 0}. \tag{9.27}$$

STEP 2: *Local times of a diffusion ζ and relationship to Brownian local time*
The setting and the notation is the same as in the previous step. If we call $l_u^h(\zeta)$ the local time of the diffusion ζ at level h at time u, then we get the following lemma:

Lemma 9.2.3: *In the setting as above ($\zeta_0 = t$, (9.27)) it holds that*

$$l_u^h(\zeta) \stackrel{d}{=} (s'(h))^{-1} l^{s(h)-s(t)}_{\gamma^{-1}(u)}(\beta),. \tag{9.28}$$

PROOF: Before we start with calculating local times we do an easy but helpful calculation of the quadratic variation, which also will be of help in a later proof:

$$\begin{aligned} d\langle \zeta^\delta, \zeta^\delta\rangle_u &= \left(s'(s^{-1}(B_u + s(t)))\right)^{-2} d\langle B, B\rangle_u \\ &= \left(s'(s^{-1}(\beta_{\gamma^{-1}(u)} + s(t)))\right)^{-2} d\gamma^{-1}(u) \\ &= \left(m'(s^{-1}(\beta_{\gamma^{-1}(u)} + s(t)))\right)^{-1} \left(s'(s^{-1}(\beta_{\gamma^{-1}(u)} + s(t)))\right)^{-1} du. \end{aligned} \tag{9.29}$$

Hence by the definition of local time we get the relationship between local time of ζ and Brownian motion β:

$$\begin{aligned} l_u^h(\zeta) &= \lim_{\epsilon\to 0} \frac{1}{\epsilon} \int_0^u \mathbb{1}_{\{\zeta_v \in [h, h+\epsilon)\}}\, d\langle \zeta^\delta, \zeta^\delta\rangle_v \\ &= \lim_{\epsilon\to 0} \frac{1}{\epsilon} \int_0^u \mathbb{1}_{\{s(\zeta_v)-s(t)\in [s(h)-s(t), s(h)-s(t)+s'(h)\epsilon+o(\epsilon))\}}\, d\langle \zeta^\delta, \zeta^\delta\rangle_v \qquad (9.30) \\ &= \lim_{\epsilon\to 0} \frac{1}{\epsilon} \int_0^u \mathbb{1}_{\{\beta_{\gamma^{-1}(v)}\in [s(h)-s(t), s(h)-s(t)+s'(h)\epsilon+o(\epsilon))\}} \qquad (9.31) \\ &\quad \left(m'(s^{-1}(\beta_{\gamma^{-1}(v)} + s(t)))\right)^{-1} \left(s'(s^{-1}(\beta_{\gamma^{-1}(v)} + s(t)))\right)^{-1} dv. \end{aligned}$$

Now a change of variables $z := \gamma^{-1}(v)$ gives:

$$= \lim_{\epsilon\to 0} \frac{1}{\epsilon} \int_0^{\gamma^{-1}(u)} \mathbb{1}_{\{\beta_z \in [s(h)-s(t), s(h)-s(t)+s'(h)\epsilon+o(\epsilon))\}} \left(s'(s^{-1}(\beta_z + s(t)))\right)^{-2} dz = \tag{9.32}$$

and Occupation Times Formula leads to:

$$\begin{aligned}&= \lim_{\epsilon \to 0} \frac{1}{\epsilon} \int \mathbb{1}_{\{x \in [s(h)-s(t), s(h)-s(t)+s'(h)\epsilon + o(\epsilon))\}} \left(s'(s^{-1}(x+s(t)))\right)^{-2} l^x_{\gamma^{-1}(u)}(\beta)\, dx \\ &= (s'(h))^{-1} l^{s(h)-s(t)}_{\gamma^{-1}(u)}(\beta), \end{aligned} \tag{9.33}$$

where in the last step we used that Brownian local time is continuous in x and s' is right-continuous. □

In fact it looks if we had made our life a bit complicated with some factors and summands cancelling out in the end of the proof. But we will see the merits of this puzzling calculation in the next steps, where we can use the setting just presented.

STEP 3: *Excursion depths of a diffusion ζ: Depths coordinate*
In this step we describe the "depths" part of the excursion point process. In Step 4 the "time" part (or coordinate) is dealt with.

Let ζ be a regular diffusion as above and let t be in the interior of the state space (l, r) of ζ. Remember the definitions of $\alpha^t_u(\zeta)$, ϵ^-_u, U^-_t and $\pi^{\zeta,t}$ from page 42.
The point process $\pi^{\zeta,t}$ measures the depth of the downward excursions from level t of ζ. Then it holds the following lemma, describing the depths coordinate:

Lemma 9.2.4: *The process $\pi^{\zeta,t}$ is a Poisson point process and the second coordinate $n^{\zeta,t}$ of the intensity measure for $0 < h < t$ is given by*

$$n^{\zeta,t}(dh) = s'(t) \frac{s'(h)}{2(s(t)-s(h))^2} dh. \tag{9.34}$$

And the intensity measure of excursions reaching zero is given by

$$n^{\zeta,t}(\{0\}) = s'(t) \frac{1}{2s(t)}. \tag{9.35}$$

In the last line of the lemma there should not be $\{0\}$, but $(-\infty, 0]$). If we restrict to diffusions ζ reflected at zero, as we will do later, then it does not matter.

PROOF: By Theorem VI.57.6 from [RW79] we know that the measure on downward excursions is a Poisson point process

$$N(dl \otimes de) = dl \otimes n(de), \tag{9.36}$$

where the first coordinate is measured in local time.

If we set for $H = t - h$

$$G_h := \{f \in U^-_t : \inf(f) < h = t - H\}, \tag{9.37}$$

the set of excursions going below h we are interested in the law of $n(G_h)$. To obtain that law we adapt the proof of Theorem VI.57.12 in the same book to our situation. Therefore set

$$T := \inf\{l > 0 : N((0, l] \times G_h) > 0\}, \tag{9.38}$$

that is the first time an excursion with depth $t-h$ arises. Since N is Poisson with rate $n(G_h)$ we know that the waiting time T is exponentially distributed and therefore

$$E[T] = (n(G_h))^{-1}. \tag{9.39}$$

We additionally define the time when the process ζ reaches the level $t-h$, when it is started in t:

$$\tau := \inf\{r : \zeta_r = t - H, \zeta_0 = t\}. \tag{9.40}$$

By Lemma 9.2.2 we have that for a Brownian motion started in 0:

$$\tau \stackrel{d}{=} \inf\{r : \beta_{\gamma^{-1}(r)} = s(t-H) - s(t), \beta_0 = 0\}. \tag{9.41}$$

For Brownian motion reflected at zero the Tanaka formula gives that the following expression is a martingale:

$$\left((\beta_u)^- - \frac{1}{2} l_u^{s(t)}(\beta)\right)_{u \geq 0}. \tag{9.42}$$

Now the Optional Sampling Theorem for the optional time $\gamma^{-1}(\tau)$ yields:

$$E[(\beta_{\gamma^{-1}(\tau)})^-] = \frac{1}{2} E[l^0_{\gamma^{-1}(\tau)}(\beta)]. \tag{9.43}$$

But the left hand side is equal to $(s(t-H) - s(t))^-$ and for the right hand side we calculate:

$$\begin{aligned} \frac{1}{2} E[l^0_{\gamma^{-1}(\tau)}(\beta)] &= \frac{1}{2} s'(t) E[l^t_\tau(\zeta)] = \\ &= \frac{1}{2} s'(t) E[T] \\ &= \frac{1}{2} s'(t) \left(n(G_h)\right)^{-1}. \end{aligned} \tag{9.44}$$

And putting left and right hand side of (9.43) together we get:

$$n(G_h) = s'(t) \left(2s(t) - s(h)\right)^{-1}. \tag{9.45}$$

That is just the first line of the lemma. The second line does not need to be proven additionally, since it is just the special case of $h = 0$. □

Step 4: *Excursion depths of a diffusion ζ: Time coordinate*
In this step we will talk about the first coordinate of the point process $\pi^{\zeta,t}$. Excursions arise, when local time grows. So the natural first coordinate for the intensity measure $\aleph^{\zeta,t}$ of $\pi^{\zeta,t}$ is given by:

$$dl^t_u(\zeta) := (l^t_\cdot(\zeta))^{-1}(du). \tag{9.46}$$

This is the uniform distribution on the set $\{u : \zeta_u = t\}$.

As we are going to deal with tree excursions later, we put the following question: What happens if we want to consider the diffusion ζ only until its local time at zero reaches $\frac{4}{b}$? Until which u can the uniform measure for the local time be interesting?

We already notice that things here get a bit complicated. So to make things clear: Here we are going to deal with a diffusion ζ starting at zero and we will apply the results from Step 2, but in the case $t = 0$. We will write $\tilde{\zeta}$ instead of ζ in this step and only in this step, where we deal with a diffusion ζ starting in $t = 0$.

To be more precise we give some definitions and a lemma, which are put in a wider framework:
For a Brownian motion β and $r > 0$ set:

$$r* := \inf\{u \geq 0 : l_u^0(\beta) = rs'(0)\}. \tag{9.47}$$

By Lemma 9.2.3 it is then:

$$\gamma(r*) \stackrel{d}{=} \alpha_r := \inf\{u \geq 0 : l_u^0(\tilde{\zeta}) = r\}, \tag{9.48}$$

and the following lemma holds:

Lemma 9.2.5: *If $(Z_h)_{h\geq 0}$ is a solution of the following SDE*

$$dZ_h = \sqrt{4Z_h s'(h)}d\beta_h\,,\ Z_0 = rs'(0), \tag{9.49}$$

then

$$(l^h_{\alpha_r}(\tilde{\zeta}))_{h\geq 0} \stackrel{d}{=} (\frac{Z_h}{s'(h)})_{h\geq 0}. \tag{9.50}$$

Remark 9.2.6:
This result describes the total mass of the local time on the set $\{u : \zeta_u = h\}$. Hence it can be calculated via an inhomogeneous Feller diffusion with catalyst rate $4s'$. $\diamond$

Proof: By the second Ray-Knight theorem, we know that $(l^a_{r*}(\beta))_{a\geq 0}$ is a squared zero-dimensional Bessel process started in $rs'(0)$ (see Chapter VI.52 in [RW79]). Using Lemma 9.2.3 we have that the following expression is also a squared zero-dimensional Bessel process started in $rs'(0)$:

$$\left(s'(s^{-1}(a))\ l^{s^{-1}(a)}_{\alpha_r}(\tilde{\zeta})\right)_{a\geq 0}. \tag{9.51}$$

Then we calculate:

$$\begin{aligned}\left(s'(s^{-1}(a))\ l^{s^{-1}(a)}_{\alpha_r}(\tilde{\zeta})\right) =& s'(0)l^0_{\alpha_r}(\tilde{\zeta}) + \int_0^a \sqrt{4s'(s^{-1}(u))\ l^{s^{-1}(u)}_{\alpha_r}(\tilde{\zeta})}d\beta_u \\ =& s'(0)r + \int_0^{s^{-1}(a)} \sqrt{4s'(z)\ l^z_{\alpha_r}(\tilde{\zeta})\ s'(z)}d\beta_z,\end{aligned} \tag{9.52}$$

and get by setting

$$Z_z := s'(z)l^z_{\alpha_r}(\tilde{\zeta}), \tag{9.53}$$

the following SDE:

$$dZ_z = \sqrt{4Z_z s'(z)}d\beta_z \text{ and } Z_0 = s'(0)l^0_{\alpha_r} = s'(0)r. \tag{9.54}$$

Hence the lemma is shown by rearranging the definition of Z. $\square$

STEP 5: *Application of the previous result to the point process $\pi^{\zeta^\delta,t}$ and comparison with π^t.*
The previous general results can now be taken up to for the case of the diffusion ζ^δ, i.e. where:

$$\begin{aligned} m(dx) &= \lambda(dx), \\ s(h) &= \frac{b}{2k}\int_0^h X_s\, ds, \\ \zeta_0^\delta &= 0. \end{aligned} \tag{9.55}$$

Additionally we have to stop the contour, when its local time at level 0 reaches $\frac{4}{b}$, i.e. at time $\alpha_{4/b}$. Note that it is sufficient to compare intensity measures of π^t and $\pi^{\zeta^\delta,t}$ to obtain equality in distribution. The point process $\pi^{\zeta^\delta,t}$ of depths of downward excursions from level t has the intensity measure $\aleph^{\zeta^\delta,t}$. We will calculate this intensity measure on a small rectangle of size $du \times dh$.

Its first (=time) coordinate is uniformly distributed with a total mass given by Lemma 9.2.5 for $r = \frac{4}{b}$:

$$l^t_{\alpha_{4/b}}(\zeta^\delta) = \frac{Z_t}{s'(t)} = \frac{2kZ_t}{bX_t} \overset{d}{=} \frac{4Y_t}{bX_t}, \tag{9.56}$$

since in this case $\frac{k}{2}Z$ behaves like a reactant Y with a fixed catalyst medium $(X_s)_{s\geq 0}$ (consider (9.49)).

Secondly the depth measure $n^{\zeta^\delta,t}$ is given by Lemma 9.2.4:

$$n^{\zeta,t}(dh) = X_t \frac{X_h}{2(\int_{t-h}^t X_s\, ds)^2} dh. \tag{9.57}$$

Hence, in the product measure evaluated on a rectangle the factors X_t cancel out and we get:

$$\aleph^{\zeta^\delta,t}(du \otimes dh) = \mathbb{1}_{[0,Y_t]}(u)du \otimes \frac{2X_h}{(\int_{t-h}^t bX_s\, ds)^2} dh. \tag{9.58}$$

And this is almost the proof of Proposition 6.3.3, since this is the same intensity measure than that of π^t. One can argue similarly for the points at level 0, but we omit the reproduction of the ideas.

Now after having done all the necessary steps we are done with the proof of Proposition 9.2.1. □□□

9.3 Comparison result between classical forest and the catalytic forest

In this section we give the proof of Proposition 4.3.1. Therefore we recall the following definition of an integral over the catalyst total mass . This integral is a bit different, than in the previous section, since here we are only going to deal with

contour downward excursions from level t. Therefore the integral starts from level t on and accumulates the catalyst downwards:

$$s_t : \begin{cases} [0,\tau^0] & \to [0,\infty) \\ h & \mapsto \frac{b}{2k}\int_{t-h}^{t} X_s\, ds \end{cases} \tag{9.59}$$

In fact it is $s_t(h) = s(t) - s(t-h)$ for the case $k = 1$ and this shortcut is more handy than always writing the difference.

Then we want to prove the following proposition:

Proposition 9.3.1 (Stretching tree metric)**:** *Let Z^{for} be a classical Galton-Watson forest, i.e. branching rate equal to 2 with contour process β run until local time at level 0 reaches 2. Let Y^{for} a catalytic branching forest with fixed catalyst contour $(X_s)_{0\le s\le \tau^0}$. Then for any $t < \tau^0$ let*

$$\tilde{Y}_t^{\text{for}} := \partial Q_{s_t(t)}(Z^{\text{for}}) \tag{9.60}$$

and for $u_1, u_2 \in \tilde{Y}_t^{\text{for}}$, i.e. $u_1, u_2 \in \partial Q_{s(t)}(Z^{\text{for}})$ define:

$$\mathrm{d}_{\tilde{Y}^{\text{for}}}(u_1,u_2) := 2s_t^{-1}\left(\tfrac{1}{2}\mathrm{d}_{Z^{\text{for}}}(u_1,u_2)\right). \tag{9.61}$$

Then it holds that

$$(\tilde{Y}^{\text{for}}, \mathrm{d}_{\tilde{Y}^{\text{for}}}) \overset{d}{=} (\partial Q_t Y^{\text{for}}; X). \tag{9.62}$$

Proof: First we claim that the contour of the ordinary Galton-Watson process is indeed equal to reflected Brownian motion β if we choose the traversal speed $k_Z = 1$ and $b = 2$ by looking at Proposition 5.3.1:

$$\mathcal{C}(Z^{\text{for}}, 1) = \beta. \tag{9.63}$$

Note that comparing the distribution of extant individuals $\partial Q_t Y^{\text{for}}$ and $(\tilde{Y}^{\text{for}}, d_{\tilde{Y}^{\text{for}}})$ as metric spaces means comparing them in the Gromov-Hausdorff metric.

But any ultrametric space is characterized by its minimal spanning tree [GPW08, Remark 2.2]. Therefore it is sufficient to compare distributions of the minimal spanning trees. These spanning trees themselves are characterized by the MRCA points, which denote the most recent common ancestors of two individuals alive at time t. These MRCA points are just the points of the point process π^t (in the case of $\partial Q_t Y^{\text{for}}$) and $\pi^{\beta,t}$ (in the case of $\partial Q_{s_t(t)} Z^{\text{for}}$). Hence we only need to show that the law of π^t is the same as the s_t-transformed law of $\pi^{\beta,t}$ and we do this by showing the equivalence of their intensity measures $\aleph^t$ and $\aleph^{\beta,t}$ similar as in the preceding proof and hence choose $\delta > 0$ s.t. $t < \tau^\delta$.

We will use notation from the previous section, in the sense that the contour traversal speed is set to be $k > 0$ and we will specify this k. Especially the reader should be familiar with Steps 4 and 5 in that section.

As the intensity measures consist of two independent coordinates, we separate the proof into two parts and start by comparing the time coordinate (or local time coordinate, to be more precise) and treat the excursion coordinate, i.e. the excursion measures n^t and $n^{\beta,t}$ later.

The first coordinate is Lebesgue measure in both cases, which is supported until local time at level 0 hits 2 for Z and $4/b$ for Y. Then the only question is the total

mass of the Lebesgue measure in each case, i.e. how much local time at level t (for Y^{for}) and at level $s_t(t)$ (for Z^{for}) until the end of the excursion.

Then by Lemma 9.2.3:

$$s_t'(t) l^t_{\alpha_{4/b}}(\zeta^\delta) = l^{s_t(t)}_{\gamma^{-1}(\alpha_{4/b})}(\beta). \tag{9.64}$$

To specify the lower index on the right hand side observe that:

$$\begin{aligned} r = \gamma^{-1}(\alpha_{4/b} &\Leftrightarrow \gamma(r) = \alpha_{4/b} \\ &\Leftrightarrow l^0_{\gamma_r}(\zeta^\delta) = \frac{4}{b} \\ &\Leftrightarrow l^0_r(\beta) = s'(0)\frac{4}{b} = \frac{b}{2k}\frac{4}{b} = \frac{2}{k}. \end{aligned} \tag{9.65}$$

We want to call α^β the inverse local time of β at zero. Then by choosing $k = 1$ we get for (9.64):

$$s_t'(t) l^t_{\alpha_{4/b}}(\zeta^\delta) = l^{s_t(t)}_{\gamma^{-1}(\alpha_2^\beta)}(\beta). \tag{9.66}$$

So the time parts of the point processes are already related.

Secondly we need to show that the "excursion-parts" n^t and $n^{\beta,t}$ of the intensity measures correspond. Therefore let $0 < h_1 < h_2 < t$ be given and calculate on the one hand for the measure n^t by Theorem 9.1.3:

$$n^t([h_1, h_2]) = \frac{s_t'(t)}{2}\left(\frac{1}{s_t(h_2)} - \frac{1}{s_t(h_1)}\right). \tag{9.67}$$

On the other hand we observe for the intensity measure $n^{\beta,t}$ of Z^{for} that the metric transformation for two extant individuals $u_1, u_2 \in \partial Q_{s_t(t)}(Z^{\text{for}})$ works like the following:

$$\frac{1}{2}\mathrm{d}_{\tilde{Y}^{\text{for}}}(u_1, u_2) \in [h_1, h_2] \Leftrightarrow \frac{1}{2}\mathrm{d}_{Z^{\text{for}}}(u_1, u_2) \in [s_t(h_1), s_t(h_2)]. \tag{9.68}$$

Hence it is:

$$\begin{aligned} n^{t,\tilde{Y}^{\text{for}}}([h_1, h_2]) &= n^{\beta,t}([s_t(h_1), s_t(h_2)]) \\ &= \frac{1}{2}\left(\frac{1}{s_t(h_2)} - \frac{1}{s_t(h_1)}\right). \end{aligned} \tag{9.69}$$

By multiplying the first and the second coordinate of the intensity measures the factor $s_t'(t)$ cancels out and we get similar as in Step 5 of the previous sectionthat $\tilde{Y}_t^{\text{for}}$ and $\partial Q_t Y^{\text{for}}$ have the same distribution. □

The next proposition was given in the contour process chapter. Its proof requires results shown in this chapter about the local time process of the contour. First we restate it:

Proposition 9.3.2: *Let a fixed catalyst $(X_t)_{t\geq 0}$ and its killing time τ^0 be given. For $\delta > 0$ let ζ^δ be the reactant limit contour. Then*

$$\mathbf{P}[\lim_{\delta\to 0}\langle\zeta^\delta, \zeta^\delta\rangle_{\alpha_{4/b}} < \infty | X] = \mathbf{P}[\rho^0 < \tau^0] \tag{9.70}$$

and

$$\mathbf{P}[\lim_{\delta\to 0}\langle\zeta^\delta, \zeta^\delta\rangle_{\alpha_{4/b}} = \infty | X] = \mathbf{P}[\rho^0 > \tau^0]. \tag{9.71}$$

PROOF: In the first case there is nothing to show, since the quadratic variation of a bounded diffusion until an almost sure finite time is finite.

In the second case we condition on the event $\rho^0 > \tau^0$ and we can write by (9.64):

$$\begin{aligned} \{\rho^0 > \tau^0\} = \{Y_t > 0 \ \forall t < \tau^0\} = \\ \{X_t l^t_{\alpha_{4/b}}(\zeta^\delta) > 0 \ \forall t < \tau^0\} = \\ \{l^t_{\alpha_{4/b}}(\zeta^\delta) > 0 \ \forall t < \tau^0\}. \end{aligned} \tag{9.72}$$

So by Occupation Times Formula we get:

$$\begin{aligned} \mathbb{1}_{\{\rho^0>\tau^0\}}\langle\zeta^\delta,\zeta^\delta\rangle_{\alpha_{4/b}} &= \int_0^{\alpha_{4/b}} \mathbb{1}_{\{\rho^0>\tau^0\}} d\langle\zeta^\delta,\zeta^\delta\rangle_u \\ &= \int_0^{\tau^\delta} l^v_{\alpha_{4/b}}(\zeta^\delta)\mathbb{1}_{\{l^t_{\alpha_{4/b}}(\zeta^\delta)>0 \ \forall t<\tau^0\}} dv \\ &= \int_0^{\tau^\delta} \frac{2l^{s(v)}_{2*}(\beta)}{bX_v}\mathbb{1}_{\{l^{s(t)}_{2*}(\beta)>0 \ \forall t<\tau^0\}} dv. \end{aligned} \tag{9.73}$$

Hence, the indicator just transformed to an indicator for non-vanishing of the β-local time. Therefore the denominator can be bounded from below by an almost sure positive random variable ϕ. Hence

$$\langle\zeta^\delta,\zeta^\delta\rangle_{\alpha_{4/b}} \geq \phi \int_0^{\tau^\delta} \frac{1}{X_v} dv. \tag{9.74}$$

The integral expression is then a random variable which for $\delta \to 0$ tends to infinity by Lemma 3.1 in [AW05] in the case that $g(x) = x^{1+\beta}$ for $\beta \in [0,1)$. In fact we wanted to have a slightly wider range of branching modifications g according to G4 of Condition 2.1.3:

$$\exists \beta \in [0,1) \ \lim_{x\to 0} \frac{g(x)}{x^{1+\beta}} = c'. \tag{9.75}$$

The integral above has its difficulties only in the region, where the catalyst approaches zero.

Hence, we are only going to consider the integral for a catalyst started in $X_0 = \epsilon'$ and where g lies already ϵ-close to its approximation in $[0, 2\epsilon']$. We set the iterated δ-hitting times τ^ϵ_k, $k \geq 1$ of X (jumps of local time). Then it is clear that

$$\int_0^{\tau^0} \frac{1}{X_v}\, dv \geq \sup_{k\geq 1} \int_{\tau^\epsilon_k}^{\tau^0} \frac{1}{X_v}\, dv. \tag{9.76}$$

Remember that $dX_t = \sqrt{g(X_t)}\, dB_t$ and we define another well-defined diffusion $\tilde{X}$, also started in ϵ' by:

$$d\tilde{X}_t = \sqrt{\tilde{X}_t^{1+\beta}}\, dB_t. \tag{9.77}$$

If we set

$$\gamma_t = \int_0^t \frac{ds}{g(B_s+\delta)}, \ \alpha_t = \int_0^t \frac{ds}{(B_s+\delta)^{1+\beta}}, \tag{9.78}$$

then

$$X_t = B_{\gamma^{-1}(t)} + \epsilon', \ \tilde{X}_t = B_{\alpha^{-1}(t)} + \epsilon' \tag{9.79}$$

Thus, we can consider the integral, where the catalyst stays in a region close to the origin and we get by a change of variables:

$$\int_{\tau_k^\epsilon}^{\tau^0} \frac{1}{X_v}\, dv = \int_0^{\tilde{\tau}^0} \frac{dr}{\tilde{X}_r} \frac{(B_{\alpha_r^{-1}+\delta} + \epsilon')^{1+\delta}}{g(B_{\alpha_r^{-1}+\delta} + \epsilon')} \geq \int_0^{\tilde{\tau}^0} \frac{dr}{\tilde{X}_r}(c' - \epsilon). \tag{9.80}$$

And the last integral was already shown to have the value infinity in the previously mentioned Lemma 3.1 of [AW05]. □

Part III

Appendix

A Important theorems required in the proofs

Here will be denoted some of the theorems quoted from books. The notation will be convenient with the usual notation of this diploma thesis, so some of the theorems will look different than the ones in the mentioned books. They were mostly put in the order of appearance in the book rather than appearance in this paper.

Clearly the theorems do not substitute a thourough study of the topics with the help of the books mentioned.

A.1 Theorems from the book of Ethier and Kurtz: Markov Processes

The book of Stewart Ethier and Tom Kurtz is one of the widest collection of what is known about Markov processes in general. It covers semigroup theory, the basics of stochastic analysis martingale problem theory and not to forget many applications. Special emphasis is laid on convergence of Markov processeses.

A.1.1 Semigroup theory

A first result about semigroups and generators is Theorem 1.7.1 from [EK86, p.37], which states:

Theorem A.1.1 (Theorem 1.7.1 from [EK86]):
Let A be a linear operator on L sucht that $\bar{A}$ is single-valued and generates a strongly continuous contraction semigroup on $L = C_b(E^1_{\text{cont}} \times E_{\text{slope}}, \mathbb{R})$. Let B be a dissipative linear operator on L such that $\mathbb{D}(B) \supset \mathbb{D}(A)$. If

$$\|Bf\| \leq \alpha\|Af\| + \beta\|f\|, \;\; f \in \mathbb{D}(A), \tag{A.1}$$

where $0 \leq \alpha < 1$ and $\beta > 0$, then $\overline{A+B}$ is single-valued and generates a strongly continuous contraction semigroup on L.

Another result about Markov jump processes states the following:
Theorem 8.3.1 from [EK86, p.376]

Theorem A.1.2 (Theorem 8.3.1 from [EK86]):
Let

$$Af(x) = \lambda(x) \int (f(y) - f(x))\mu(x, dy).$$

Let E be a locally compact, noncompact, separable metric space and let $E^\Delta = E \cup \Delta$ be its one-point-compactification. Let $\lambda \in C(E)$ be nonnegative and let $\mu(x, \Gamma)$ be a transition function on $E \times \mathcal{B}(E)$ such that the mapping $x \mapsto \mu(x, \cdot)$ of E into $\mathcal{P}(E)$

is continuous. Let γ and η be positive functions in $C(E)$ such that $1/\gamma$ and $1/\eta$ belong to $C_0(E)$ and

$$\sup_{x\in E}\frac{\lambda(x)}{\gamma(x)} \equiv C_1 < \infty,$$
$$\lim_{x\to\Delta}\lambda(x)\mu(x,K) = 0 \text{ for every compact } K \subset E,$$
$$\sup_{x\in E}\lambda(x)\int\frac{\gamma(x)-\gamma(y)}{\gamma(y)}\mu(x,dy) \equiv C_2 < \infty,$$
$$\sup_{x\in E}\lambda(x)\int\frac{\eta(x)-\eta(y)}{\eta(y)}\mu(x,dy) \equiv C_3 < \infty.$$

Then the closure of $\{(f,Af) : f \in C_0(E), \gamma f \in C_b(E), Af \in C_0(E)\}$ is single-valued and generates a Feller semigroup on $C_0(E)$. Moreover, $C_c(E)$ is a core for this generator.

A.1.2 Convergence theorems for Markov processes

The next theorems are being useful to show some convergence results for Markov processes. In fact we first state a theorem, which shows where the idea of proofs will go along, but we will not use explicitely within this thesis. It holds for any kind of stochastic processes with càdlàg paths:

Theorem A.1.3 (Theorem 3.7.8 from [EK86])**:**
Let E be separable and let $Z_n, n = 1, 2, \ldots,$ and Z be processes with sample paths in $\mathcal{D}_E[0,\infty)$.
If $\{Z_n\}$ is relatively compact and there exists a dense set $D \subset [0,\infty)$ such that

$$(Z_n(t_1),\ldots,Z_n(t_k)) \Rightarrow (Z(t_1),\ldots,Z(t_k)) \tag{A.2}$$

holds for every finite set $\{t_1,\ldots,t_k\} \subset D$, then $Z_n \Rightarrow Z$, where convergence is convergence in the distribution on the path-space $\mathcal{D}_E[0,\infty)$.

To establish the two criteria relative compactness and f.d.d.-convergence there are several tools available:

An important indicator for the first is the compact containment condition, which states that one can find a compact subset of the state space in which the process stays with high probability up to a given time.

Then one way to show relative compactness is to transfer the problem of finding a compact set in the metric space E to finding a compact set in $\mathbb{R}$ for a big class of mappings f from the state space E to $\mathbb{R}$:

Theorem A.1.4 (Theorem 3.9.1 from [EK86])**:**
Let (E,r) be a complete and separable metric space, and let $\{Z_n\}$ be a family of processes with sample paths in $\mathcal{D}_E[0,\infty)$. Suppose that the compact containment condition holds. That is, for every $\lambda > 0$ and $T > 0$ there exists a compact set $\Gamma_{\lambda,T} \subset E$ for which

$$\inf_n P[Z_n(t) \in \Gamma_{\lambda,T} \text{ for } 0 \le t \le T] \ge 1-\lambda. \tag{A.3}$$

Let H be a dense subset of $\mathcal{C}_b(E)$ in the topology of uniform convergence on compact sets. Then $\{Z_n\}$ is relatively compact if and only if $\{f \circ Z_n\}$ is relatively compact (as a family of processes with sample paths in $\mathcal{D}_{\mathbb{R}}[0,\infty)$) for each $f \in H$.

By the previous theorem the relative compactness question of the sequence (Z_n) can be shifted to the real-valued sequence $(f \circ Z_n)$ for sufficiently many f. With the next theorem relative compactness can be shown for the latter:

Theorem A.1.5 (Theorem 3.9.4 from [EK86]):
Let (E,r) be an arbitrary metric space, and let $\{Z_n\}$ be a family of processes with sample paths in $\mathcal{D}_E[0,\infty)$. Let C_a denote a subalgebra of $\mathcal{C}_b(E)$ (e.g. the space of bounded, uniformly continuous functions with bounded support), and let D be the collection of $f \in \mathcal{C}_b(E)$ such that for every $\epsilon > 0$ and $T > 0$ there exist real-valued $\mathcal{F}^n$-progressive-measurable (Y_n, Z_n) with uniformly bounded expectation in t with:

$$Y_n(t) - \int_0^t Z_n(s)\,ds \text{ is a martingale,} \tag{A.4}$$

$$\sup_n E\left[\sup_{t\in[0,T]\cap\mathbb{Q}} |Y_n(t) - f(X_n(t))|\right] < \epsilon, \tag{A.5}$$

$$\sup_n E[\int_0^T |Z_n(t)|^p] < \infty \text{ for some } p \in (1,\infty]. \tag{A.6}$$

If C_a is contained in the closure of D (in the sup norm), then $\{f \circ X_n\}$ is relatively compact for each $f \in C_a$.

The following theorem puts together most of the previously mentioned ideas and states the convergence of the Markov process Z_n to a diffusion Z. This convergence holds if something closely related to the convergence of the corresponding generators holds and relative compactness is true.

Theorem A.1.6 (Theorem 4.8.10 from [EK86]):
Let (E,r) be complete and separable. Let $A \subset \mathcal{C}_b(E) \times \mathcal{C}_b(E)$ and $\nu \in \mathcal{P}(E)$, and suppose that the $\mathcal{D}_E[0,\infty)$ martingale problem for (A,ν) has at most one solution. Suppose X_n, $n = 1,2,\ldots$, is a $\{\mathcal{G}_t^n\}$-adapted process with sample paths in $\mathcal{D}_E[0,\infty)$, $\{X_n\}_{n\in\mathbb{N}}$ is relatively compact, $PX_n(0)^{-1} \Rightarrow \nu$, and $M \subset \mathcal{C}_b(E)$ is separating. Then the following are equivalent:

(a) *There exists a solution X of the $\mathcal{D}_E[0,\infty)$ martingale problem for (A,ν), and $X_n \Rightarrow X$.*

(c) *There exists a countable set $\Gamma \subset [0,\infty)$ such that for each $(f,g) \in A$ and $T > 0$, there exist integrable (ξ_n, ϕ_n), such that:*

$$(\xi_n(t) - \int_0^t \phi_n(s)\,ds)_{t\geq 0} \text{ is an } \{\mathcal{G}_t^n\}\text{- martingale,} \tag{A.7}$$

$$\sup_n \sup_{s\leq T} E[|\xi_n(s)|] < \infty, \tag{A.8}$$

$$\sup_n \sup_{s\leq T} E[|\phi_n(s)|] < \infty, \tag{A.9}$$

$$\lim_{n\to\infty} E\left[(\xi_n(t) - f(X_n(t)))\prod_{i=1}^k h_i(X_n(t_i))\right] = 0, \tag{A.10}$$

$$\lim_{n\to\infty} E\left[(\phi_n(t) - g(X_n(t)))\prod_{i=1}^k h_i(X_n(t_i))\right] = 0, \tag{A.11}$$

for all $k \geq 0$, $0 \leq t_1 < t_2 < \cdots < t_k \leq t \leq T$ with $t_i, t \notin \Gamma$, and $h_1,\ldots,h_k \in M$.

A.1.3 Martingale problems

Another group of important theorems are put under the headline of Martingale problems. First a uniqueness result for one-dimensional SDEs says:

Remark A.1.7 (Remark 5.3.9 from [EK86]):
Let $\sigma : [0,\infty) \times \mathbb{R} \to \mathbb{R}_+$ and $b : [0,\infty) \times \mathbb{R} \to \mathbb{R}$ be locally bounded, measurable and satisfy

$$|\sigma(t,x)^2 - \sigma(t,y)^2| + |b(t,x) - b(t,y)| \leq K|x-y|, \tag{A.12}$$

$$t \geq 0,\ x \in \mathbb{R}^d, \tag{A.13}$$

for some constant K. Given two solutions

$$(\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}, W, X) \text{ and } (\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}, W, Y) \tag{A.14}$$

of the SDE corresponding to (σ, b) it is true that

$$P[X(0) = Y(0)] = 1 \tag{A.15}$$

implies

$$P[X(t) = Y(t)\ \forall t \geq 0] = 1. \tag{A.16}$$

Next a multi-dimensional analogue:

Theorem A.1.8 (Theorem 5.3.10 from[EK86]):
Let $\sigma : [0,\infty) \times \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$ and $b : [0,\infty) \times \mathbb{R}^d \to \mathbb{R}^d$ be continuous and satisfy

$$|\sigma(t,x)|^2 \leq K(1+|x|^2),\ x \cdot b(t,x) \leq K(1+|x|^2), \tag{A.17}$$

$$t \geq 0,\ x \in \mathbb{R}^d, \tag{A.18}$$

for some constant K, and let $\mu \in \mathcal{M}_1(\mathbb{R}^d)$. Then there exists a solution of the stochastic differential equation corresponding to (σ, b, μ).

The following theorem allows to derive uniqueness of a martingale problem by knowing the uniqueness of the one-dimensional distributions. In a special case it also allows to prove the strong Markov property of a solution of the martingale problem.

Theorem A.1.9 (Theorem 4.4.2 from [EK86]):
Let E be separable, and let $A \subset B(E) \times B(E)$. Suppose that for each $\mu \in \mathcal{P}(E)$ any two solutions X, Y of the martingale problem for (A, μ) have the same one-dimensional distribution that is, for each $t > 0$,

$$P[X(t) \in \Gamma] = P[Y(t) \in \Gamma],\ \Gamma \in \mathcal{B}(E). \tag{A.19}$$

Then the following hold.

(a) Any solution of the martingale problem for A with respect to a filtration $(\mathcal{G}_t)$ is a Markov process with respect to $(\mathcal{G}_t)$, and any two solutions of the martingale problem for (A, μ) have the same finite-dimensional distributions.

(c) Let X^x be the unique solution of the $(A, \mu = \delta_x)$-martingale problem with càdlàg paths and P_x is the law of X^x on $\mathcal{D}_E[0,\infty)$. If such a solution exists for any $x \in E$ and the mapping $x \to P_x(B)$ is measurable for any $B \in \mathcal{B}(\mathcal{D}_E)$, then X is a strong Markov process.

A.2 Theorems from the book of Karatzas and Shreve: Brownian Motion and Stochastic Calculus

From the book Brownian Motion and Stochastic Calculus [KS00] by Ioannis Karatzas and Steven Shreve we take several results which are losely connected. First we quote a result about regular conditional probabilities. This theorem states the existence of probability kernels describing the conditional probabilities as long as we are on a complete and separable metric space. This is always the case if the state space of the processes we consider itself is complete and separable and the processes are either continuous or have càdlàg paths. In this thesis the state spaces are either subsets of $\mathbb{R}^d$, the set of compact rooted real trees $\mathbb{T}^{\text{root}}$.

Theorem A.2.1 (Theorem 5.3.9 from [KS00])**:**
Let Ω be a complete, separable metric space with Borel σ-field $\mathcal{F} = \mathcal{B}(\Omega)$ and a probability measure P. Furthermore let X a measurable mapping from this space into a measurable space $(S, \mathcal{S})$, on which it induces the distribution PX^{-1}. There exists then a function $Q(x; A) : S \times \mathcal{F} \to [0, 1]$, called a regular conditional probability for $\mathcal{F}$ given X such that

(i) for each $x \in S$,$Q(x; \cdot)$ is a probability measure on $(\Omega, \mathcal{F})$,

(ii) for each $A \in \mathcal{F}$, the mapping $x \mapsto Q(x, A)$ is $\mathcal{S}$-measurable, and

(iii) for each $A \in \mathcal{F}$, $Q(x; A) = P[A|X = x]$, PX^{-1} -a.e. $x \in S$.

This probability is unique in the sense that for PX^{-1} -a.e. $x \in S$:

$$Q(x; \{\omega : X(\omega) = x\}) = 1 \tag{A.20}$$

To apply it to the quenched situation consider in the case of the tree-valued process, the probability space $(\Omega, \mathcal{F}, \mathbf{P})$ from the introduction. Then the functional $((\eta_t^{\text{tot}})_{t\geq 0}, \xi_r^{\text{for}})$, where r is fixed induces a probability measure $P = \mathbf{P} \circ ((\eta_t^{\text{tot}})_{t\geq 0}, \xi^{\text{for}})^{-1}$ on the measurable space $(\mathcal{D}_{\mathbb{R}^1_+}[0, \infty) \times \mathbb{T}^{\text{root}}, \mathcal{B})$, where $\mathcal{B}$ is the Borel-σ-algebra on the product space. Then the projection on the first coordinate in this space is a measurable mapping by definition of the product-σ-algebra. With the help of the theorem a regular conditional probability exists, i.e. there is a probability kernel $Q(\cdot, \cdot) : \mathcal{D}_{\mathbb{R}^1_+}[0, \infty) \times \mathcal{B} \to [0, 1]$ s.t. for $A \in \mathcal{B}$ and $\mathbf{P}$- almost surely any $\eta \in \mathcal{D}_{\mathbb{R}^1_+}[0, \infty)$ one has the existence of the kernels.

A next result tells us about the existence and the strong uniqueness of a one-dimensional SDE. It is pretty close to the results of Engelbert and Schmidt and tells that existence and strong uniqueness hold under quite general conditions:

Theorem A.2.2 (Corollary 5.5.10 from [KS00])**:**
Let $\sigma : \mathbb{R} \to \mathbb{R}$ be given. The equation $dX_t = \sigma(X_t)dW_t$ possesses a unique strong solution for every initial distribution μ, if the four conditions (E) and (i)-(iii) hold for functions $f : \mathbb{R} \to [0, \infty]$ and $h : [0, \infty] \to [0, \infty]$:

(E)

$$I(\sigma) \subseteq Z(\sigma), i.e. : \{x \in \mathbb{R} : \int_{x-\epsilon}^{x+\epsilon} \frac{dy}{\sigma^2(y)} = \infty\} \subseteq \{x \in \mathbb{R} : \sigma(x) = 0\} \tag{A.21}$$

(i) at every $x \in I(\sigma)^c$, the quotient $(f/\sigma)^2$ is locally integrable; i.e., there exists $\epsilon > 0$ (depending on x) such that

$$\int_{x-\epsilon}^{x+\epsilon} \left(\frac{f(y)}{\sigma(y)}\right)^2 dy < \infty; \tag{A.22}$$

(ii) the function h is strictly increasing and satisfies $h(0) = 0$ and

$$\int_0^{\epsilon} h^{-2}(u)du = \infty; \ \forall \epsilon > 0 \tag{A.23}$$

(iii) there exists a constant $a > 0$ such that

$$| \sigma(x+y) - \sigma(x) | \leq f(x)h(| y |); \quad \forall x \in \mathbb{R}, y \in [-a, a]. \tag{A.24}$$

Solutions of SDEs and solutions of martingale problems have a strong link and the next result tells in a quite general version about the Strong Markov property of such solutions:

Theorem A.2.3 (Theorem 5.4.20 from [KS00])**:**
Let a linear operator A on a subset of $C_0(\mathbb{R}^d, \mathbb{R})$ be given by:

$$Af(x) = \sum_{1 \leq i,j \leq d} \sigma_{ij}(x) \frac{\partial^2 f}{\partial x_i \partial x_j}(x) + \sum_{1 \leq i \leq d} b_i(x) \frac{\partial f}{\partial x_i}(x) \tag{A.25}$$

Suppose that the coefficients b, σ are bounded on compact subsets of $\mathbb{R}^d$, and that the (A, δ_x)-martingale problem is well-posed for any $x \in \mathbb{R}^d$ with solution $P_x \in \mathcal{M}_1(C_{\mathbb{R}^d}[0, \infty))$. Then the family P_x satisfies the strong Markov property.

A.3 Theorems from the book of Rogers and Williams: Diffusions, Markov Processes and Martingales

From this very good book [RW79] of Chris Rogers and David Williams we only take a theorem about the boundary behaviour of a diffusion X on the state space $[0, \infty)$, with measure P_x on $C([0, \infty), [0, \infty))$, when $X(0) = x$. This theorem states that the hitting time of he boundary 0 can already be calculated by checking an integrability criterion for the speed measure of the diffusion:

Theorem A.3.1 (Theorem V.51.(ii))**:**
Let X be a diffusion with natural scale (e.g. no drift) and speed measure m. For $H_0 := \inf\{t > 0 : X_t = 0\}$ it is true that:

$$P^x[H_0 < \infty] = 1 \text{ for all } x > 0 \text{ if and only if } \int_{0^+} xm(dx) < \infty. \tag{A.26}$$

B Additional concepts and proofs

B.1 The Kingman coalescent

In 1982 J.F.C Kingman presented in his paper [Kin82] the so-called "Kingman"-coalescent. It is a stochastic process describing the merging of sets, where merging of two sets arises after exponential time independent of all other possible two-set-combinations.

To be more precise we use the notation of [GLW07] and let a countable set S be given. We say that $\{\pi_\lambda\}$ is a partition of S, iff

- $\cup\pi_\lambda = S$,
- $\pi_\lambda \cap \pi_{\lambda'} = \emptyset$, when $\lambda \neq \lambda'$ and
- $\pi_\lambda \neq \emptyset$ for all λ.

The sets π_λ are called partition elements. The set of all partitions of S will be called Π^S.

An equivalence relation $\sim_P$ on S is induced by a partition $P \in \Pi^S$:

$$s_1 \sim_P s_2 :\Leftrightarrow \exists \pi \in P \text{ s.t. } s_1, s_2 \in \pi. \tag{B.1}$$

For a subset $S' \subset S$ we define a mapping $\rho_{S'} : \Pi^S \to \Pi^{S'}$ by the corresponding equivalence relations:

$$s_1 \sim_{\rho_{S'}P} s_2 :\Leftrightarrow s_1, s_2 \in S' \text{ and } \exists \pi \in P \text{ s.t. } s_1, s_2 \in \pi. \tag{B.2}$$

This mapping restricts a partition of S to a partition of one of its subsets S'. Additionally define a partial ordering $\prec$ on Π^S:

$$P \prec P' :\Leftrightarrow \forall \pi' \in P' \exists \pi \in P : \pi \subseteq \pi', \tag{B.3}$$

and to count the number of partition elements of P, we write $\#P$.
If we restrict our attention to the case $S = \mathbb{N}$ and subsets $S_n = \{1, 2, \ldots, n\}$, what we will do henceforth, a metric d can be introduced on the space Π^S:

$$d(P, P') := 2^{-N(P,P')}, \tag{B.4}$$

$$\text{where } N(P, P') = \sup\{n \in \mathbb{N} : \rho_{\{1,2,\ldots n\}}P = \rho_{\{1,2,\ldots n\}}P'\}. \tag{B.5}$$

With this metric the mapping $\rho_{S'}$ is continuous for any finite subset $S' \subset \mathbb{N}$. One can even then check that (Π^S, d) is a compact, complete, separable metric space (the first can be shown via sequential compactness).

Before we start with the construction of the coalescent we set the following special partitions of S_n and S:

$$\begin{aligned}\Delta &:= \{\{1\},\{2\},\dots,\{n\}\} \text{ in the case of } S_n\\ \Delta &:= \{\{1\},\{2\},\dots,\} \text{ in the case of} S\\ \Theta &:= \{1,2,\dots,n\} \text{ in the case of } S_n\\ \Theta &:= \{1,2,\dots\} \text{ in the case of } S\end{aligned} \tag{B.6}$$

Now to start with the coalescent we give the following definition:

Definition B.1.1 (The n-coalescent)**:**
The n-coalescent $(R_t^n)_{t\geq 0}$ *is a Markov-process with state space* Π^{S_n}*, starting at* $R_0^n = \Delta$ *with the generator* Ω^n*:*

$$\Omega^n f(\alpha) = \sum_{\beta\prec\alpha,|\beta|=|\alpha|-1} (f(\beta) - f(\alpha))\,. \tag{B.7}$$

First some properties of this process can be shown, which can be found in [Kin82]:

- the process $(D_t^n)_{t\geq 0} := (\#R_t^n)_{t\geq 0}$ is a death process starting in n with rate $\binom{k}{2}$ in state k.
- the process $(\mathfrak{R}_k^n)_{1\leq k\leq n}$ which gives the sequence of states in Π^{S_n} for the n-coalescent, is ordered in the following sense:

  $$\Delta = \mathfrak{R}_n^n \prec \mathfrak{R}_{n-1}^n \prec \cdots \prec \mathfrak{R}_2^n \prec \mathfrak{R}_1^n = \Theta and \tag{B.8}$$

  is a Markov process.
- These two processes D and $\mathfrak{R}$ are independent.

The idea now is to extend this definition to a process defined on Π^S and starting in the diffuse partition Θ. This "limit" process should have the property, that when restricting it to a finite subset of $\mathbb{N}$, then we got the n-coalescent. This was first done by Kingman in [Kin82] and we will follow his ideas in a very short description, called the paintbox-construction:

Given a probability vector $\mathbf{x} = (x_1, x_2, \dots)$ define an i.i.d. sequence of random variables $Z_1, Z_2, \dots$, s.t.:

$$\mathbf{P}(Z_1 = r) = x_r \ \forall r \in \mathbb{N}_0. \tag{B.9}$$

This gives us an exchangeable probability measure $P^{\mathbf{x}}$ on Π^S, where this measure is induced by the following equivalence relation:

$$\mathbf{R} = \{(i,j) : Z_i = Z_j \geq 1\}. \tag{B.10}$$

We can do the same even if the vector $\mathbf{x}$ is random by defining the law on the partitions by:

$$\mathbf{P} = \int \mathbf{P}^x \,\mu(dx). \tag{B.11}$$

As a special case we take it to be uniformly distributed on the $k-1$-dimensional simplex Δ_k, that means:

$$x_0 = x_{k+1} = x_{k+2} = \cdots = 0, \tag{B.12}$$

$$dx_1\, dx_2 \cdots dx_{k-1} = d\lambda^{k-1}|_{\Delta_k}(x_1, \ldots, x_{k-1}), \tag{B.13}$$

$$x_k = 1 - (x_1 + \cdots + x_{k-1}), \tag{B.14}$$

where $d\lambda^{k-1}$ is the $k-1$-dimensional Lebesgue measure. Then after a uniform random choice of two of the coordinates $x_1, \ldots, x_k$, say x_i and x_j and adding up these two components we obtain the vector

$$(x_1, \ldots, x_i + x_j, \ldots, \hat{x_j}, \ldots, x_k). \tag{B.15}$$

It is easy to see that this vector is uniformly distributed over the $k-1$-dimensional simplex. By this consistent procedure we can by a projective limit argument define a Markov process $(\mathfrak{R}_k)_{k\in\mathbb{N}}$ with state space $\Pi^S = \Pi^{\mathbb{N}}$:

$$\mathbf{P}[\mathfrak{R}_{k-1} = \eta | \mathfrak{R}_k = \xi] = \begin{cases} 2/k(k-1) & \text{if } \eta \prec \xi \text{ and } \#\eta = \#\xi - 1, \\ 0 & \text{otherwise,} \end{cases} \tag{B.16}$$

where $\xi \in \Pi^S, \#\xi = k$.
Let us additionally we define an independent death process $(D_t)_{t\geq 0}$ starting from ∞ with death rates $\frac{1}{2}k(k-1)$ in state k ([Don91]). Then the process

$$R_0 := \Delta,\ R_t := \mathfrak{R}_{D_t} \text{ for } t > 0, \tag{B.17}$$

is a Π^S-valued Markov process, called the *coalescent*. The restriction $\rho_{S_n}(R)$ of R to the first n coordinates is the ordinary n-coalescent.

In fact in the topology of Π^S we get that for $(i,j) \in \mathbb{N}^2$, with $i \neq j$, we have with the aid of exchangeability of the law that for $k \in \mathbb{N}$:

$$\mathbf{P}[(i,j) \in \mathfrak{R}_k] = \mathbf{P}[(1,2) \in \mathfrak{R}_k] = \frac{2}{k+1}. \tag{B.18}$$

Therefore it is true that

$$\mathbf{P}[(i,j) \in R_t] = \mathbf{E}[\frac{2}{D_t + 1}] \tag{B.19}$$

and this expression tends to 0 as $t \to 0$. Hence the initial distribution of R is $R_0 = \Delta$.

More ideas and newer ideas, for example the "look-down" idea are presented in the work of Donnelly and Kurtz [DK96], or in Alison Etheridge's book about superprocesses [Eth00] the entrance law at ∞ is discussed in a paper of Donnelly [Don91].

B.2 Diffusions and scale functions

Let two diffusions ζ^1 and ζ^2 be given on a compact time interval $[0,T]$. Their state space is a compactum $[L,H] \subset \mathbb{R}$ and they are reflected on the boundary. Call s_1 and s_2 the scale functions of the diffusion and assume that the speed measure is Lebesgue measure. Our aim is to show that, when the derivatives s_1' and s_2' differ slightly in Skorokhod-metric, then the diffusions stay close in the sense that

$$E[\sup_{0\le u\le T} |\zeta_u^1 - \zeta_u^2|]$$

can be bounded by the Skorokhod distance. Hence we want to prove some sort of continuity from scale functions to diffusions.

We state the following proposition:

Proposition B.2.1: *Assume that the derivatives of the scale functions are càdlàg and lie in a compact interval:*

$$s_1', s_2' \in [a,b] \subset (0,\infty). \tag{B.20}$$

Then, for all $\epsilon > 0$ there exists a $\delta > 0$, s.t.:

$$d^{\mathrm{Sk}}(s_1', s_2') < \delta \;\Rightarrow\; E[\sup_{0\le u\le T} |\zeta_u^1 - \zeta_u^2|] < a^{-3}\epsilon. \tag{B.21}$$

Proof: The proof is rather straightforward. First we note that similar to Section 9.2 in Step 1 we can set a random time-change for a Brownian motion β:

$$\begin{aligned} \gamma_1(t) &= \int_0^t \frac{1}{s_1'(\beta_v)}\, dv, \\ \gamma_2(t) &= \int_0^t \frac{1}{s_2'(\beta_v)}\, dv. \end{aligned} \tag{B.22}$$

With that time change one can rewrite the diffusions as time-changed Brownian motions as in Lemma 9.2.2. Now we will use the $\|\cdot\|$ to indicate the supremum norm.

$$\begin{aligned} E[\sup_{0\le u\le T} |\zeta_u^1 - \zeta_u^2|] &= E[\|s_1^{-1}(\beta_{\gamma_1^{-1}(u)}) - s_2^{-1}(\beta_{\gamma_2^{-1}(u)})\|] \\ &\le E[\|s_1^{-1}(\beta_{\gamma_1^{-1}(u)}) - s_2^{-1}(\beta_{\gamma_1^{-1}(u)})\|] + \\ &\quad + E[\|s_2^{-1}(\beta_{\gamma_1^{-1}(u)}) - s_2^{-1}(\beta_{\gamma_2^{-1}(u)})\|] \\ &\le \|s_1^{-1}\|\, E[\|\beta_v - s_1(s_2^{-1}(\beta_v))\|] \\ &\quad + \|s_2^{-1}\| E[\|\beta_{\gamma_1^{-1}(u)} - \beta_{\gamma_2^{-1}(u)}\|]. \end{aligned} \tag{B.23}$$

(A) For the first line we observe that for $\alpha := s_2^{-1}(u)$:

$$u = s_2(\alpha) \ge \alpha a \|u - s_1(s_2^{-1}(u))\| \quad = \|s_2(\alpha) - s_1(\alpha)\| \le \|\alpha\|\, d^{\mathrm{Sk}}(s_1', s_2')(1+b)$$

(B.24)

$$\le a^{-1}\|u\|\, d^{\mathrm{Sk}}(s_1', s_2')(1+b). \tag{B.25}$$

Hence we get

$$E[\|\beta_v - s_1(s_2^{-1}(\beta_v))\|] \leq a^{-1} E[\|\beta_v\|]\, d^{\mathrm{Sk}}(s_1', s_2')(1+b). \tag{B.26}$$

(B) For the second line we need to show more and this will require some time. We need to find a way to say that the time-changes do not differ too much. Hence for a constant $r > 0$ we split up the problem in two lines

$$\begin{aligned} E[\|\beta_{\gamma_1^{-1}(u)} - \beta_{\gamma_2^{-1}(u)}\|] = \\ E[\|\beta_{\gamma_1^{-1}(u)} - \beta_{\gamma_2^{-1}(u)}\| \mathbb{1}_{\|\gamma_1^{-1}(u) - \gamma_2^{-1}(u)\| \leq r\sqrt{d^{\mathrm{Sk}}(s_1', s_2')}}] + \\ E[\|\beta_{\gamma_1^{-1}(u)} - \beta_{\gamma_2^{-1}(u)}\| \mathbb{1}_{\|\gamma_1^{-1}(u) - \gamma_2^{-1}(u)\| > r\sqrt{d^{\mathrm{Sk}}(s_1', s_2')}}]. \end{aligned}$$

Triangle inequality and Cauchy-Schwarz inequality yield:

$$\begin{aligned} \cdots \leq & E[\sup_{0 \leq \tilde{r} \leq r\sqrt{d^{\mathrm{Sk}}(s_1', s_2')}} |\beta_{\tilde{r}}|] \\ & + 2E[\sup_{0 \leq s \leq T} |\beta_s^2|] P[\|\gamma_1^{-1}(u) - \gamma_2^{-1}(u)\|^2 > r^2 d^{\mathrm{Sk}}(s_1', s_2')]. \end{aligned}$$

With Brownian scaling and Doob's inequality we obtain

$$\cdots \leq \sqrt[4]{d^{\mathrm{Sk}}(s_1', s_2') r^2}\, E[\sup_{0 \leq \tilde{r} \leq 1} |\beta_{\tilde{r}}|] + 8E[|\beta_T^2|] P[\cdots]. \tag{B.27}$$

Hence, what remains to be done is to bound the probability expression on the right-hand-side. We do that by using a strictly increasing smooth function λ which is the "time-change" to relate s_1' and s_2' as Skorokhod functions (see Section 3.5 in [EK86], especially (3.5.5)):

$$\|\gamma_1^{-1}(u) - \gamma_2^{-1}(u)\| \leq \|(\gamma_1^{-1})'\| \quad \|u - \gamma_1(\gamma_2^{-1}(u))\|,$$

by setting $t := \gamma_1^{-1}(u)$ we get

$$\begin{aligned} \leq & H \,\| \int_0^t \frac{1}{s_1'(\beta_v)}\, dv - \int_0^t \frac{1}{s_2'(\beta_v)}\, dv \| \\ \leq & H \int_0^t \frac{|s_2'(\beta_v) - s_1'(\lambda(\beta_v))| + |s_1'(\lambda(\beta_v)) - s_1'(\beta_v)|}{s_1'(\beta_v)\, s_2'(\beta_v)}\, dv \\ \leq & H a^{-2} \|\gamma_1^{-1}(u)\| d^{\mathrm{Sk}}(s_1', s_2') + H a^{-2} \| \int_0^t s_1'(\lambda(\beta_v)) - s_1'(\beta_v)\, dv \|. \end{aligned} \tag{B.28}$$

The first summand looks well for a good choice of r but the second summand still needs further treatment. Unfortunately we will not go into details now. First one needs to go back to considering the probability of the event in (B.27) and apply Markov inequality. Looking at the line just written we can split up s_1' into finitely many continuity components which cover most of the interval. Then use a uniform continuity argument to finish the idea.

(C) Hence we get the following:

$$\begin{aligned} E[\|\beta_{\gamma_1^{-1}(u)} - \beta_{\gamma_2^{-1}(u)}\|] \leq \\ \leq & a^{-2} E[\|\beta_v\|]\, d^{\mathrm{Sk}}(s_1', s_2')(1+b) \\ & + a^{-2}\sqrt{d^{\mathrm{Sk}}(s_1', s_2')r}\, E[\sup_{0\leq\tilde{r}\leq 1} |\beta_{\tilde{r}}|] \\ & + 8E[|\beta_T^2|]\frac{Ha^{-3}}{r}\sqrt{d^{\mathrm{Sk}}(s_1', s_2')} \\ & + 8E[|\beta_T^2|]\frac{Hba^{-3}}{r}\sqrt{d^{\mathrm{Sk}}(s_1', s_2')}. \end{aligned} \tag{B.29}$$

And the expression on the right hand side can be bounded from above by $a^{-3}\epsilon$ when choosing s_1' and s_2' sufficiently close in Skorokhod metric. □

B.3 Link between Birth-and-Death processes and Branching processes

Let two models be given for a rate $\lambda > 0$:

- a continuous time branching process Br with binary offspring. Branching of each individual happens independently of the others after an $\mathrm{Exp}(\int_t^{t+\cdot} 2\lambda_s\, ds)$ time. The process starts with one individual: $Br_0 = 1$.
- a continuous time birth-and-death process BD. Each individual independent of the others dies after an exponential time with rate λ and gives birth after an exponential time with rate λ. New-born individuals are "attached to the right", i.e. in linear order they are put after the father individual.

By simply looking at the generators, the two total mass processes related to the two models are the same. It even holds the following lemma:

Lemma B.3.1: *The contour processes of Br and BD are equal in distribution:*

$$\mathcal{C}(Br : \sigma) \stackrel{d}{=} \mathcal{C}(BD : \sigma). \tag{B.30}$$

Remark B.3.2:
It is crucial for this lemma to have binary offspring for the branching process. To relate other offspring distributions or even random offspring one gets other results.

Proof: We compare the two contours from the start on and show that the lengths of the line segments are the same.

Let X_1 be the length of the first constant slope line segment in the BD contour. The first individual after an exponential time. Hence for $x > 0$

$$P(X_1 > x) = \mathrm{e}^{-\lambda x}. \tag{B.31}$$

On the other hand let Y_1 be the length of the first constant slope line segment of the Br contour. The length Y_1 is the sum of several line segments in the genealogical tree. It has probability $\frac{1}{2}$ to consist of one line segment, probability $\frac{1}{4}$ to consist of two line segments, probability $\frac{1}{8}$ to consist of three line segments and so on.

Each of these line segments is equal to a lifetime of a branching individual and this is an Exp($\int 2\lambda\, dr$)-time. If we call $E_1, E_2, \ldots$ a sequence of such exponential random variables, each starting, where the other one ends, then we can write:

$$\begin{aligned} P(Y_1 > x) &= \sum_{k=1}^{\infty} P(E_1 + \cdots + E_k > x, \#\text{line segments} = k) \\ &= \sum_{k=1}^{\infty} P(E_1 + \cdots + E_k > x) P(\#\text{line segments} = k) \\ &= \sum_{k=1}^{\infty} \Gamma(k, 2\lambda)([x, \infty)) 2^{-k}, \end{aligned} \tag{B.32}$$

where $\Gamma(k, \lambda)$ is Gamma-distribution with parameters k and λ. After a short calculation one gets:

$$P(Y_1 > x) = \int_x^{\infty} \lambda \mathrm{e}^{-\lambda s}\, ds = \mathrm{e}^{-\lambda x}. \tag{B.33}$$

But this is the same distribution as X_1, so the first length of the contours coincide. One can continue this idea for the oncoming line segments as well. □

B.4 Additional proofs

Lemma B.4.1 (Local Lipschitz-continuity implies Lipschitz-continuity on compacta)**:** *Let (E, d), $(F, \tilde{d})$ be metric spaces and $f : E \to F$ a locally Lipschitz-continuous function. Then f is globally Lipschitz-continuous on any compactum $K \subset E$*

Proof: Assume the contrary:
Let (x_n), (y_n) sequences in K, s.t.: $\tilde{d}(f(x_n), f(y_n)) > n\ d(x_n, y_n)$ Since K is compact, there exists a subsequence (x_{n_k}) of (x_n) and a subsequence $(y_{n_{k_l}})$ of (y_{n_k}) which converge in K. Not to get confused with notation, we will without loss of generality assume that $x_n \to x$ and $x_n \to y$ as $n \to \infty$. If $x \not= y$ then the right hand side in (reference) is positive for $n \to \infty$, so the left hand side would need to grow to infinity. But this cannot be since f is continuous and K compact. In the case $x = y$, assume that V is the neighbourhood of x, in which Lipschitz-continuity holds for the Lipschitz-constant $L_x < \infty$. As x_n, y_n converge to x, there will be an $N \in \mathbb{N}$ s.t. $x_n, y_n \in V\ \forall n \geq N$. So we have:

$$n\ d(x_n, y_n) < \tilde{d}(f(x_n), f(y_n)) < L_x d(x_n, y_n)\ \forall n \geq N$$

But this cannot be true. □

Notation

$\|\cdot\|_E$	supremum norm on the set of functions from E to $\mathbb{R}$	
α_r^t	$= \inf\{u \geq 0 : l_u^t(\zeta) = r\}$	p. 42
α_r	$= \inf\{u \geq 0 : l_u^0(\zeta) = r\}$	
β	one-dimensional Brownian motion started in 0	
C_0	space of continuous functions vanishing at infinity	
$C(E,F)$	space of continuous functions from E to F	
$C_E[0,\infty)$	space of continuous functions from $[0,\infty)$ to E	
$C^{0,*}_{[0,\infty)}[0,L]$	$= \{f \in C([0,L],[0,\infty)) : \ f(0) = f(L) = 0,\ f(x) > 0\ \forall x \in (0,L)\}$	p. 30
$\mathcal{C}(\cdot : \sigma)$	ordered tree to contour mapping	p. 30
$\mathcal{D}_{\mathbb{R}^2_+}[0,\infty)$	Banach space of càdlàg functions $f : [0,\infty) \to \mathbb{R}^2_+$	
$\mathbb{D}(A)$	domain of the linear operator A	
$d_{GH}(\cdot,\cdot)$	Gromov-Hausdorff-metric on $\mathbb{T}^{\text{root}}$	p. 23
$l_u^t(\zeta)$	local time of ζ at time u at level t	
$\mathcal{M}_1(\Omega)$	the set of probability measures on Ω	
$\mathbb{N}$	the natural numbers, $\{1,2,\dots\}$	
$\mathfrak{P}(\Omega)$	power set of Ω	
R	Kingman coalescent	p. 70
$\mathbb{R}^2_+$	$= \{(x,y) \in \mathbb{R}^2 : x \geq 0, y \geq 0\}$	
$\mathbb{R}_+$	$= [0,\infty)$	
$\mathcal{T}$	contour to ordered tree mapping	p. 31
$\mathcal{T}_{\text{unord}}$	contour to tree mapping	p. 31
$\mathbb{T}^{\text{root}}$	space of rooted $\mathbb{R}$-trees	p. 21
$\mathbb{T}^{\text{root,lin}}$	space of rooted, linearly ordered tree	p. 24
$\mathbb{T}^{\text{root,lin}}_{\text{fin}}$	space of rooted, linearly ordered tree, with finitely many branch points	p. 24
W^1, W^2	independent Brownian motions	

Index

In this index only few "n-rescaled objects" are listed. Indeed the definitions of them can be found shortly after the other definitions of the non-rescaled objects throughout the thesis. Mostly they are analagous.

Bibliography

[AK72] Krishna Arthreya and Samuel Karlin. Branching processes with random environments. *The Annals of Mathematical Statistics*, 42:1499 – 1520, 1972.

[Ald93] David Aldous. The continuum random tree. III. *Annals of Probability*, 21(1):248–289, 1993.

[Ald99] David J. Aldous. Deterministic and stochastic models for coalescence: a review of the mean-field theory for probabilists. *Bernoulli*, 5:3–48, 1999.

[AN72] Krishna Arthreya and Peter Ney. *Branching Processes.* Springer, New York, second edition, 1972.

[AW05] Siva Athreya and Anita Winter. Spatial coupling of neutral measure-valued population models. *Stochastic Processes and their Applications*, 115(6):891–906, Jun 2005.

[Bre68] Leo Breiman. *Probability.* Addison Wesley Publishing, Reading, Massachusetts, first edition, 1968.

[CDG04] Theodore Cox, Donald Dawson, and Andreas Greven. Mutually catalytic super branching random walks: large finite systems and renormalization analysis. *Memoirs of the American Mathematical Society*, 171:97pp, 2004.

[Chi01] Ian Chiswell. *Introduction to* Λ *tree.* World Scientific, Singapore, 2001.

[DK96] Peter Donnelly and Thomas Kurtz. A countable representation of the fleming viot measure-valued process. *Annals of Probability*, 24:698 – 742, 1996.

[Don91] Peter Donnelly. Weak convergence to a markov chain with an entrance boundary: anestral processes in population genetics. *Annals of Probability*, 19:1102 – 1117, 1991.

[EK86] Stewart N. Ethier and Thomas G. Kurtz. *Markov processes – characterization and convergence.* Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, 1986.

[EPW06] Steven N. Evans, Jim Pitman, and Anita Winter. Rayleigh processes, real trees, and root growth with re-grafting. *Probab. Theory Related Fields*, 134(1):81–126, 2006.

[Eth00] Alison M. Etheridge. *An introduction to superprocesses*, volume 20 of *University Lecture Series.* American Mathematical Society, Providence, RI, 2000.

[EW06] Steven N. Evans and Anita Winter. Subtree prune and regraft: A reversible real tree-valued markov process. *Annals of Probability*, 34:918 – 961, 2006.

[Fel68] William Feller. *An Introduction to Probability and its Applications 1.* John Wiley and Sons, New York, third edition, 1968.

[GLW07] Andreas Greven, Vlada Limic, and Anita Winter. Coalescent processes arising in a study of diffusive clustering. *Preprint*, 2007. arXiv:math/0703875v1 [math.PR].

[GPW06] A. Greven, L. Popovic, and A. Winter. Genealogy of catalytic branching processes. *Preprint*, 2006. arXiv:math/0606313v1 [math.PR].

[GPW07] Andreas Greven, Peter Pfaffelhuber, and Anita Winter. Convergence in distribution of random metric measure spaces. *Preprint*, 2007. arXiv:0806.2224v1 [math.PR].

[GPW08] Andreas Greven, Peter Pfaffelhuber, and Anita Winter. Tree-valued resampling dynamics. *Preprint*, 2008. arXiv:0806.2224v1 [math.PR].

[Har61] Robert Harris. *The Theory of Branching Processes.* Springer, New York, second edition, 1961.

[Jag09] Peter Jagers. Some notes on the history of branching processes, from my perspective. Talk in Oberwolfach, http://www.math.chalmers.se/ jagers/Branching%20History.pdf, January 2009.

[Kal83] Olav Kallenberg. *Random Measures.* Academic Press, New York, third edition, 1983.

[Ken66] D.G. Kendall. Branching processes since 1873. *Journal of the London Mathematical Society*, 41:385 – 406, 1966.

[Ken75] D.G. Kendall. Branching processes since (and after) 1873. *Bulletin of the London Mathematics Society*, 7:225 – 253, 1975.

[Kin82] J.F.C. Kingman. The coalescent. *Stochastic Processes and their Applications*, 13:235 – 248, 1982.

[Kle08] Achim Klenke. *Wahrscheinlichkeitstheorie.* Springer, Berlin, second edition, 2008.

[KS00] I. Karatzas and S. Shreve. *Brownian Motion and Stochastic Calculus.* Springer, New York, second edition, 2000.

[KT81] Samuel Karlin and Howard M. Taylor. *A Second Course in Stochastic Processes.* Academic Press, New York, second edition, 1981.

[Kur92] Thomas Kurtz. *Averaging for martingale problems and stochastic approximation*, volume 177 of *Lecture Notes in Control and Information Sciences.* Springer, Berlin, 1992.

[LG96] Jean Francois Le-Gall. Random trees and applications. *Probability Surveys*, 2:245 – 311, 1996.

[Lin92] Torgny Lindvall. *Lectures on the coupling method.* John Wiley and Sons, New York, 1992.

[NP89] J. Neveu and J. Pitman. The branching process in a Brownian excursion. In *Séminaire de Probabilités XXIII*, volume 1372 of *Lecture Notes in Math.*, pages 248–257. Springer, 1989.

[Øks05] Bernt Øksendal. *Stochastic Differential Equations.* Springer, Berlin, third edition, 2005.

[Pen03] Christian Penssel. *Interacting Catalytic Feller Diffusions.* Logos Verlag, Berlin, second edition, 2003.

[RW79] L.C.G. Rogers and David Williams. *Diffusions, Markov Processes and Martingales.* John Wiley and Sons, New York, first edition, 1979.

[RY91] Daniel Revuz and Marc Yor. *Continuous Martingales and Brownian Motion.* Springer, New York, first edition, 1991.

[WG75] Henry William Watson and Francis Galton. On the probabilities of extinction of families. *Journal of the Anthropological Institute of Great Britain*, 4:138 – 144, 1875.

Danksagung

Hiermit möchte ich Herrn Prof. Dr. Andreas Greven für die Betreuung meiner Diplomarbeit danken. In vielen Gesprächen stand er mir immer mit vielen wertvollen Tipps beiseite.

Zudem gilt mein besonderer Dank Dr. Anita Winter und Lea Popovic für hilfreiche Erläuterungen und unterstützende Worte. Auch Maxim Drabkin, Patric Glöde, Sven Piotrowiak, Peter Seidel und Frank Schirmeier gebührt viel Dank.

Ganz besonders danken möchte ich meinen Eltern, meinem Bruder Michael und meiner Schwester Christine für fortwährende Unterstützung im Laufe des gesamten Studiums. Vor allem meiner Freundin Lucía möchte ich für ihre Geduld und Unterstützung danken.

Zuletzt gilt mein Dank noch allen Freunden, die mir während der Diplomarbeit und während des gesamten Studiums mit Rat beiseite standen.

Declaration

I hereby declare that this document has been composed by me and is based on my own work, unless otherwise acknowledged in the text.

Erlangen, May 4, 2009

Selbstständigkeitserklärung

Hiermit erkläre ich, dass ich diese Arbeit selbstständig und nur mit den im Literaturverzeichnis angegebenen Quellen angefertigt habe.

Erlangen, 4. Mai 2009