PHD

Some aspects of a large deviations theory

Git, Yoav

Award date:
1998

Awarding institution:
University of Bath

Some Aspects of Large Deviations Theory

submitted by

Yoav Git

for the degree of Ph.D

of the

University of Bath

1998

Signature of Author .... Yoav Git

Yoav Git

Contents

"I apologise for writing such a long exposition, I did not have the time to write a short one"

Friedrich Engels (1810-1895)

Prologue

Having an Israeli friend is a thankless task and a blemish on your reputation. Consequently, my friends who have been forced to read this manuscript at the expense of an after dinner conversation will be relieved to find their anonymity preserved.

I am grateful to Simon Harris, David Hobson, David Marles and Jon Warren who have exhibited extreme constraint while attempting to answer my many wild and unstructured questions.

My thanks are also extended to EPSRC who has kindly supported this research and to Professor Grimmett and the Statistical Laboratory in Cambridge for hosting me over the past year. While in Cambridge, my trips to Bath were made possible thanks to the warm hospitality of uncle Ryan.

I can not begin to thank my family (old and new) who enriched my life with so much happiness over the past three years, but I owe the most to the magical ingredient of this thesis, Professor David Williams, my supervisor.

Notations and Standard Abbreviations

Some standard notations (e.g. ":=" meaning "defined to be") are used throughout this work. We draw the attention of the reader to a small ambiguity in the definition of $\bar{X}_n$ and $\bar{X}_{\sqrt{n}}$.

General Notations

- LHS := Left hand side
- RHS := Right hand side
- =: := "RHS is defined to be LHS"
- w.r.t. := With respect to
- WLOG := Without loss of generality
- iff := If and only if

Spatial Notations

- $\mathcal{R}$:= The real numbers
- $\mathcal{R}^+$:= The non negative real numbers $\{x \in \mathcal{R} : x \geq 0\}$
- $\mathcal{Z}$:= The integers
- $\mathcal{Z}^+$:= The positive integers $\{n \in \mathcal{Z} \colon n > 0\}$
- $\mathcal{N}$:= The natural numbers $\{n \in \mathcal{Z} \colon n \geq 0\}$

- $[a, b] := \{x : a \leq x \leq b\}$, a closed interval
- $(a, b) := \{x : a < x < b\}$, an open interval
- 1_A := The indicator function of A

- $||x||$:= The norm of x

- $< x, y >$:= The inner product of x and y

- $C^n(X, Y)$:= Functions from X to Y with a continuous nth derivative

- $f^{(n)}$:= The nth derivative of f

- C_0:= $\{x \in C^0([0,1], \mathcal{R}) : x(0) = 0\}$, the real valued continuous paths on $[0,1]$ starting at 0.

- C_1:= $\{\dot{x} \in C^0([0,1], \mathcal{R}) : x(0) = 0\}$, the real valued differentiable with continuous derivatives paths on $[0,1]$ starting at 0.

- $a_n \sim b_n := \lim_{n\to\infty} \frac{a_n}{b_n} = 1$

Probabilistic Notations

- p.d.f. := Probability density function

- m.g.f. := Moment generating function

- IID := Independent, identically distributed

- C_k^n := n choose k, defined as $C_k^n = n!/k!(n-k)!$

- $P(A)$:= Probability of an event

- $E(X)$:= Expected value of a random variable

- $\mathrm{Var}(X)$:= The variance of a random variable $\mathrm{Var}(X) = E(X^2) - E(X)^2$

- $\bar{X}_n$:= The mean of n IID random variables $\bar{X}_n = \frac{1}{n}\sum_{i\le n} X_i$

- $\bar{X}_{\sqrt{n}} := (n\mathrm{Var}(X))^{-\frac{1}{2}} \sum_{i\le n}(X_i - E(X))$

- f_X := The probability density function of X

- $M(\theta)$:= The moment generating function, $M(\theta) = E(e^{\theta X})$

- $c(\theta)$:= The log moment (cumulant) generating function, $c(\theta) = \log M(\theta)$

- $X \sim N(\mu, \sigma^2)$:= The normal distribution. $E(X) = \mu$, $\mathrm{Var}(X) = \sigma^2$

$$f_X(x) = \tfrac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$
$$\Phi(x) = P(N(0,1) < x) = \int_{-\infty}^{x} dN$$

- $X \sim \mathcal{E}(\lambda)$:= The exponential distribution. $E(X) = \mathrm{Var}(X) = \lambda^{-1}$

$$f_X(x) = \lambda e^{-\lambda x} \quad \text{for all } x \in \mathcal{R}^+$$

- $X \sim \mathcal{B}(n, p)$:= The binomial distribution. $E(X) = np$, $\mathrm{Var}(X) = np(1-p)$

$$P(X = k) = C_k^n p^k (1-p)^{n-k}$$

- CLT := Central Limit Theorem
- BCL := Borel Cantelli Lemma
- BM := Brownian Motion
- BBM := Branching Brownian Motion
- BDP := Branching Diffusion Process

Abstract

This thesis is concerned with the study of branching diffusion processes. These systems have been studied extensively using the interplay between differential equations and probability theory building on ideas by McKean and Neveu. These unite, for example, the analytical study of the Fisher-Kolmogorov-Petrowski-Piscounov (FKPP) equation and the study of martingales associated with a branching Brownian motion. Although in chapter 3 we touch upon these ideas, this work takes a different path. We attempt to re-formalise our understandings of branching diffusion processes paths from a large-deviations theory perspective. This allows us to gain additional insight into the almost sure behaviour of such systems by using ideas from Lagrangian dynamics. Chapter 4 investigates the paths of a BBM, chapter 5 follows this with the large-deviations principle associated with the integrated BBM's path. Finally, in chapter 6 we partially explore the paths of branching diffusion processes with invariant measures (e.g. the Ornstein-Uhlenbeck process).

Independently, chapter 2 contains a refinement of Cramér's theorem on $\mathcal{R}$.

Chapter 1

Since this thesis is concerned mainly with large-deviations theory, we include a short introduction to the subject in chapter 1. The chapter does not contain new results but reading it is advisable as we call upon most of the results proved in it.

Chapter 2

Chapter 2 is almost independent from the rest of the thesis. In it we present a new proof of a refined Cramér's theorem. It is mainly concerned with expressions of the form

$$A_n := \int_L e^{nI(x)} dP_n(x).$$

Here, P_n denotes the law of $\bar{X}_n$ the mean on n IID random variables and I denotes the Cramér rate function. Varadhan's theorem tells us that if L is compact then the above equation is finite and has no exponential growth. We present a result concerning the power-law governing A_n.

Chapter 3

Chapter 3 contains no new theory. It does contain a succinct introduction aimed at familiarising the reader to the field of branching diffusion processes. A complete survey of this vast topic is beyond the scope of this work.

Chapter 4

The main result of this thesis is described in chapter 4 and is concerned with almost-sure path properties of a branching Brownian motion. It is well known that for a dyadic branching Brownian motion, the particles disperse at a rate proportional to the time elapsed. Work by [Uchiyama, 82] has allowed us to estimate accurately the growth rate, $1 - \frac{1}{2}\lambda^2$, of particles along a ray $x(t) = \lambda t$. Our result extends this further. We consider a general path x in $C^0([0,1], \mathcal{R})$. We produce two rate functions $J(x)$ and $K(x)$. These measure the rate of growth of particles along a path x both in expectation and almost surely. Interestingly, these functions are not identical. The almost-sure growth is obviously bounded by the rate of growth in expectations ($J(x) \leq K(x)$) though the rates agree if for example x is a convex function. By observing that $\sup_{x(1)=\lambda} J(x) = \sup_{x(1)=\lambda} K(x) = 1 - \frac{1}{2}\lambda^2$, we see that the difference does not manifest itself as different BBM wavefront speeds. In fact, we get back to the result by [Uchiyama, 82].

Chapter 5

Thinking of a Brownian motion as the speed of a particle, we integrate to get a differential random process representing its position. We demonstrate using current techniques that the almost-sure wavefront of this process is different from the expected wavefront even in the first order of magnitude. Using the large-deviations contraction principle and utilising $J(x)$ and $K(x)$ derived in the previous chapter, we are able to explain this behaviour. We end this chapter by studying the phase-plane of a branching Brownian motion.

Chapter 6

We conclude this thesis with the study of a BBM with a drift μ towards the origin.

$$dU = dB - \mu(U)dt.$$

Our first task is to demonstrate that the expectation wavefront and the almost-sure wavefront agree. Since the process is pulled increasingly towards the origin, its wavefront speed r_t satisfies $r_t/t \to 0$. The process is highly ergodic, which means that the past behaviour of the Brownian motion is not much correlated to the its current position U_t. We quantify this notion by proving a large-deviations principle for the phase-plane

$$p_i(t) = \left(\frac{1}{t}B_i(t), \frac{1}{r_t}U_i(t)\right).$$

We find that essentially, the two coordinates are independent not only in the sense that

$$\lim_{t\to\infty} \frac{1}{t} \log P\left(p_i(t) \in (dx, dy)\right) = \lim_{t\to\infty} \frac{1}{t} \log P(\frac{1}{t}B_i(t) \in dx) + \lim_{t\to\infty} \frac{1}{t} \log P(\frac{1}{r_t}U_i(t) \in dy)$$

but also that almost surely, the number of particles in (dx, dy) grows exponentially at a rate

$$1 - \lim_{t\to\infty} \frac{1}{t} \log P\left(p_i(t) \in (dx, dy)\right).$$

Finally, we investigate the integral processes $\int U_t dt$ and $\int \mu(U_t)dt$. It turns out that the integral process $\int \mu(U_t)dt$ can be though of as a smoothed Brownian motion. This will allow us not only to apply results proved in chapter 4 but even to strengthen them by refining the topology on which the rate function is defined.

Chapter 1

Introduction to Large-Deviations Theory

Large-deviations theory is a way of studying the extreme behaviour of probabilistic systems. This chapter contains a brief outline of the classical results and techniques. More thorough expositions can be found in [Ellis, 85], [Dembo *et al.*, 95], [Varadhan, 84] and [Strook, 84].

1.1 Example: Mean of IID random variables

Suppose I was in a fair casino, betting on the outcome of independent gambles. The bets being fair, one would expect me to neither win nor lose any money over a prolonged period of time. Large-deviations theory captures this idea by looking at the extreme behaviour of the systems (a gambler consistently over-performing is such an event) and measuring the exponential decay of the probability of such rare events occurring.

Let P_n be the measure on $\mathcal{R}$ which corresponds to $\bar{X}_n := \frac{1}{n}\sum_{i\leq n} X_i$, my average gains at time n. We assume X_i to be IID random variables satisfying $E(X_i) = 0$. We wish to estimate $P(A_y^n)$, the probability of the rare event of having an average earning per bet greater than $y > 0$ after n bets.

$$A_y^n := \left\{\omega \in \Omega : \bar{X}_n \geq y\right\}.$$

1.1.1 Upper Bound

We let $M(\theta) = e^{c(\theta)} = E(e^{\theta X})$ denote the moment generating function and the cumulant generating function. We ignore X_i which do not have a finite moment generating function and refer the reader to [Heyde, 68] and [Wentzell, 90] who analyse sub-exponential decays of random variables. It follows that $E(e^{\theta \bar{X}_n}) = M^n(\theta/n)$ and hence by Chebychev's inequality:

$$P(\bar{X}_n > y) \leq \inf_\theta M^n(\theta/n)e^{-\theta y} = e^{-n \sup_\theta (\theta y - c(\theta))},$$

and so

$$\limsup_{n\to\infty} \frac{1}{n} \log P_n([y,\infty)) \leq -I(y),$$

where we define the rate function I to be

$$I(y) = \sup_\theta \{\theta y - c(\theta)\}.$$

1.1.2 Lower Bound

Assume $\sup_\theta \{\theta y - c(\theta)\}$ is attained by a point α so that

$$I(y) = \alpha y - c(\alpha)$$

Since $\sup_\theta \{\theta y - c(\theta)\}$ attains its supremum at α, $\partial I/\partial \alpha = y - \dot{c}(\alpha) = 0$. Indeed, this is another way of characterising α. We now define a new random variable X^y with law P^y using the change of measure formula

$$P^y(A) := P(X^y \in A) := M(\alpha)^{-1} \int_A e^{\alpha x} dP(x). \qquad (1.1)$$

This new random variable has mean

$$E(X^y) = E(Xe^{\alpha X})/M(\alpha) = \dot{M}(\alpha)/M(\alpha) = \dot{c}(\alpha) = y.$$

Consider now $A(y, \epsilon)$,

$$A(y, \epsilon) = \{x \in \mathcal{R} : 0 < x - y < \epsilon\}.$$

If $\bar{X}_n \in A(y, \epsilon)$, then using the change of measure n times, we deduce that

$$\begin{aligned}
P_n(A(y,\epsilon)) &= \int_{\bar{X} \in A(y,\epsilon)} dP(x_1) \ldots dP(x_n), \\
&= \int_{\bar{X} \in A(y,\epsilon)} e^{-n\alpha \bar{X}_n} \times e^{\alpha x_1} dP(x_1) \ldots e^{\alpha x_n} dP(x_n), \\
&\geq M(\alpha)^n e^{-n\alpha y - n|\alpha|\epsilon} \int_{\bar{X}_n \in A(y,\epsilon)} dP^y(x_1) \ldots dP^y(x_n), \\
&= e^{-nI(y) - n|\alpha|\epsilon} P(\bar{X^y}_n \in A(y,\epsilon)).
\end{aligned}$$

Since $E(X^y) = y$, as $n \to \infty$, using the Central Limit Theorem, $P(\bar{X^y}_n \in A(y, \epsilon))$ converges to $\frac{1}{2}$. We conclude that

$$\liminf_{n \to \infty} \tfrac{1}{n} \log P_n(A(y, \epsilon)) \geq -I(y) - |\alpha|\epsilon.$$

Letting $\epsilon \downarrow 0$ we conclude that $\liminf_{n \to \infty} \frac{1}{n} \log P_n([y, \infty)) \geq -I(y)$.

1.1.3 The Large-Deviations Principle

We generalise the notion of $\bar{X}_n$ to consider a family of measures P_n.

- We will also consider also families of measures indexed by a continuous parameter $\{P_t\}_{t \geq 0}$ as well.
- Although X_n was defined on $\mathcal{R}$, in general we will consider P_n to be supported on any complete separable metric space S.

- We assumed $\bar{X}_n \rightarrow_{\text{law}} \delta_{E(X)}$. Maintaining this notion, we will assume that P_n converges weakly to δ_{s_0}, the unit mass at $s_0 \in S$.

Definition. *We say that $\{P_n\}$ obeys the large-deviations principle with a rate function $I(s)$, if there exists a function $I : S \rightarrow [0, \infty]$ satisfying:*

- *I is lower semi-continuous.*
- *For each $l < \infty$ the set $\{s : I(s) \leq l\}$ is compact in S.*
- **Upper Bound.** *For each closed set $D \subset S$*

$$\limsup_{n\to\infty} \tfrac{1}{n} \log P_n(D) \leq - \inf_{s\in D} I(s).$$

- **Lower Bound.** *For each open set $A \subset S$*

$$\liminf_{n\to\infty} \tfrac{1}{n} \log P_n(A) \geq - \inf_{s\in A} I(s).$$

1.1.4 Cramér's Theorem

Essentially, we have proved in sections 1.1.1 and 1.1.2 most of the following.

Theorem 1 (Cramér's Theorem). *Let P_n denote the law of the mean $\bar{X}_n$ of $\mathcal{R}$-valued IID random variables X_i. Assume that the moment generating function*

$$M(\theta) := E(e^{\theta X}) = \int_{\mathcal{R}} e^{\theta x} dP(x) =: e^{c(\theta)}$$

is defined over an interval containing 0. Then, the sequence $\{P_n\}$ satisfies the large-deviations principle with the rate function I defined as

$$I(x) = \sup_{\theta}\{\theta x - c(\theta)\}$$

Proof. There are four properties in the definition which must be satisfied. The first two describe I and follow from the convexity of $c(\theta)$ and the properties of the Legendre transform. See [Ellis, 85] or [Dembo *et al.*,95] for details. The upper bound and lower bound follow easily from previous discussion or from consulting [Varadhan, 84]. □

1.2 The Contraction Principle

In this section we state two general results which allow us to transfer the large-deviations principle from one space to another. The following two theorems are straight forward to prove and can be found in [Dembo *et al.*, 95] or [Ellis, 85].

Theorem 2 (The Contraction Principle). *Let P_n on X satisfy a large-deviations principle with rate function I. Also, let $\pi : X \to Y$ be a continuous map. Then, the family $Q_n := P_n \pi^{-1}$ satisfies a large-deviations principle with rate function J*

$$J(y) = \inf_{x \in \pi^{-1}(y)} I(x).$$

Akin to the contraction principle is the Inverse Contraction Theorem.

Theorem 3 (The Inverse Contraction Theorem). *Let $\pi : X \to Y$ be a bijection with π^{-1} continuous. Let Q_n be exponentially tight in Y and assume $P_n := Q_n \pi$ satisfy a LDP with rate function I. Then, Q_n satisfy a LDP with rate function $J = I\pi^{-1}$.*

Let us re-examine Cramér's theorem (assuming of course $E(X) = 0$). It certainly implies the Strong Law of Large Numbers. For

$$\lim_{n\to\infty} \sum_{m>n} P(\bar{X}_m > \epsilon) < \sum_{m>n} e^{-mI(\frac{1}{2}\epsilon)} < \infty,$$

and using BCL, almost surely $\limsup_{n\to\infty} \bar{X}_n < \epsilon$. Now, suppose we have a mapping $\pi : \mathcal{R} \to Y$ and π was only continuous near 0. Although we do not have a large-deviations principle on Y, it is clear that almost surely

$$\lim_{n\to\infty} \pi(\bar{X}_n) = \pi(0).$$

This observation is entirely obvious and yet extremely useful. When we discuss almost-sure large-deviations type results in future chapters, we will use this argument to transfer a large-deviations principle from X to Y using maps which are *almost-surely* continuous.

1.3 An Extension of Cramér's Theorem: The paths of a diffusion process

Let X_i satisfy the conditions of Cramér's theorem with rate function I. We assume I to be continuous, strictly convex, increasing on $\mathcal{R}^+$ and decreasing on $\mathcal{R}^-$. Consider the random walk defined as $Y_n = \sum_{i\leq n} X_i$. We know $\frac{1}{n}Y_n$ satisfy a large-deviations principle. We wish to think of the path $(0, \frac{1}{n}Y_1, \dots, \frac{1}{n}Y_n)$ that the Markov chain might take and try and deduce a large-deviations principle. Thus we define the (random) piece-wise linear function $y_n : [0,1] \to \mathcal{R}$ as

$$y_n(k/n) = \frac{1}{n}\sum_{i\leq k} X_i \quad 0 \leq k \leq n.$$

This induces a probability measure P_n on the space of continuous functions satisfying $y(0) = 0$.

Theorem 4. *$\{P_n\}_{n>0}$ satisfy a large-deviations principle on $C^0([0,1],\mathcal{R})$ endowed with the supremum norm and with a rate function $J(y)$,*

$$J(y) := \begin{cases} \int_0^1 I(\dot{y})dt & \text{if } y \in C^1([0,1],\mathcal{R}), \\ \infty & \text{otherwise.} \end{cases}$$

The theorem can be seen as an application of the Inverse Contraction Theorem. A more general result is in fact true, consult [Russel, 1996] for more details. We give a sketch proof which utilises some ideas which will be used later in this work.

Sketch of Proof. Let m be finite. Each coordinate in the expression $\{\frac{1}{nm}\sum_{(i-1)n<j<in} X_j\}_{i\leq m}$ is independent. Hence $\{\frac{1}{nm}\sum_{(i-1)n<j<in} X_j\}_{i\leq m}$ satisfies a large-deviations principle in $\mathcal{R}^m$ with a rate function

$$J\left(\frac{1}{m}x_1, \dots, \frac{1}{m}x_m\right) = \frac{1}{m}\sum_{i\leq m} I(x_i).$$

For every $z \in C([0,1],\mathcal{R})$ let $J_m(z)$ be defined as

$$J_m(z) = \frac{1}{m}\sum_{i\leq m} I\left(\frac{z(\frac{i}{m}) - z(\frac{i-1}{m})}{1/m}\right).$$

Now let $m_k = 2^k$. Because I is convex, J_{m_k} are increasing. If z is differentiable $\lim_{k\to\infty} J_{m_k}(z) = J(z)$. For a suitable z which is not differentiable on some interval, it can be shown that $\lim_{k\to\infty} J_{m_k}(z) = \infty$ (remember that $\{x : I(x) \leq l\}$ is a compact set). Consequently, for *suitable* closed $D \subset C([0,1],\mathcal{R})$,

$$\limsup_{n\to\infty} \frac{1}{n}\log P_n(D) \leq -\lim_{m\to\infty}\inf_{z\in D} J_m(z) \leq -\inf_{z\in D}\limsup_{m\to\infty} J_m(z) = -\inf_{z\in D} J(z).$$

Determining which closed sets are *suitable* and showing that they generate all closed sets is tedious and uninformative. Instead, let us look at the lower bound. For every y, define its open ϵ-neighbourhood $A(y,\epsilon)$,

$$A(y,\epsilon) := \{z : |z(t) - y(t)| < \epsilon \quad \text{for all } 0 \leq t \leq 1\}.$$

We aim to show that $\liminf_{n\to\infty} \frac{1}{n}\log P_n(A(y,\epsilon)) \geq -J(y)$. Consider first the linear function $y = xt$ (hence $I(x) = J(y)$) and let $E = A(0,x)^c \cap A(y,\epsilon)^c$. Since $P_n(A(0,x)^c) \geq P(\bar{X}_n > x)$, using Cramér's theorem and the assumed continuity of I we have

$$\liminf_{n\to\infty} \frac{1}{n}\log P_n(A(0,x)^c) \geq -\inf_{z>x} I(z) = -I(x).$$

On the other hand, in lemma 1 we will show that for some $\delta > 0$

$$\limsup_{n\to\infty} \frac{1}{n}\log P_n(E) \leq -J(y) - \delta.$$

Since $A(0,x)^c \subset E \cup A(y,\epsilon)$ it follows that

$$\liminf_{n\to\infty} \frac{1}{n}\log P_n(A(y,\epsilon)) \geq -J(y)$$

as required. Finally, we easily extend $\liminf_{n\to\infty} \frac{1}{n} \log P_n(A(y,\epsilon)) > -J(y)$ to piece-wise linear functions and appeal to the fact that piece-wise functions are dense in $C([0,1],\mathcal{R})$ to complete the proof.

□

Lemma 1. *For some $\delta > 0$*

$$\limsup_{n\to\infty} \frac{1}{n} \log P_n(E) \leq -J(y) - \delta.$$

Proof. Let $E = \cup_{t\leq 1}\{U_t \cup V_t\}$ where we define U_t and V_t to be the set of all functions that first "leave" $A(y,\epsilon)$ at time $t \in (0,1]$ from above or below respectively.

$$U_t := \left\{z \in E : \inf_s\{|z(s) - xs| \geq \epsilon\} = t, z(t) = xt + \epsilon\right\},$$
$$V_t := \left\{z \in E : \inf_s\{|z(s) - xs| \geq \epsilon\} = t, z(t) = xt - \epsilon\right\}.$$

For all $z \in U_t, V_t$, by considering $z(0) = 0, z(t)$ and $z(1) \geq x$ and using convexity of I we see that

$$\inf\{J(z) : z \in U_t\} \geq J(u_t),$$
$$\inf\{J(z) : z \in V_t\} \geq J(v_t).$$

Here u_t and v_t are the piece-wise linear function satisfying $u_t(0) = v_t(0) = 0$, $u_t(1) = v_t(1) = x$ and

$$u_t(t) = xt + \epsilon, \quad v_t(t) = xt - \epsilon.$$

It now follows by strict convexity of I and the lower semi-continuity of the function

$$t \to \min\{J(u_t), J(v_t)\}$$

that $\sup_{z\in E} J(z) > J(y) + \delta$ for some positive $\delta > 0$. □

By assuming $X_i \sim N(0,1)$ in theorem 4 we get the following well known result first proved by [Schilder, 66].

Theorem 5 (Schilder's Theorem.). *Let $y^T \in C^0([0,1],\mathcal{R})$ be the scaled path of a Brownian motion.*

$$y^T(t) = \frac{1}{T}B(tT). \tag{1.2}$$

Then, the family of laws $\{P_T\}_{T>0}$ induced on $C^0([0,1],\mathcal{R})$ satisfy a a large-deviations principle with rate function

$$J(y) = \int_0^1 \tfrac{1}{2}\dot{y}^2 dt$$

1.4 More Extensions to Cramér's Theorem

- Extending the space on which X is defined.
  By resorting to the minimax principle we can assume X to be an $\mathcal{R}^d$-valued random variable. Using sub-additivity we can let X be defined on a locally convex Hausdorff real vector spaces. We refer the reader to [Dembo *et al.*, 95] for further details.

- Ellis-Gartner Theorem.
  In the proof of Cramér's Theorem, we depended mostly on the properties of $M(\theta)$, the moment generating function. The Ellis-Gartner Theorem makes this dependence explicit. It dispenses with the notion of IID random variables, and replaces them with a family of measures P_n, whose moment generating functions M_n obey a convergence condition. We will state a slightly weaker version.

  Theorem 6 (Ellis-Gartner). *Suppose we have a family of laws P_n on $\mathcal{R}^d$, with the moment generating functions*

  $$M_n(\lambda) := e^{c_n(\lambda)} := E_n(e^{<\lambda,x>}),$$

having (around a neighbourhood of 0*), a well defined convex, differentiable limit*

$$c(\lambda) = \lim_{n\to\infty} \tfrac{1}{n} c_n(n\lambda). \tag{1.3}$$

Then, P_n *satisfy the large-deviations principle with rate function* $I(x)$*, the Legendre transform of* $c(\lambda)$*,*

$$I(x) := \sup_{\lambda\in\mathcal{R}^d} \{< x, \lambda > -c(\lambda)\}.$$

Condition 1.3 seems obscure until we observe that for a family of IID random variables $\frac{1}{n}c_n(n\lambda) = c(\lambda)$.

- Improved "resolution".
  Although we know that $P_n(A)$ decays exponentially, what is the polynomial correction term? We will discuss this problem in the next chapter.

Chapter 2

The Correction Term in Cramér's Theorem

2.1 Introduction

We recall briefly Cramér's Theorem. X_i is a sequence of IID random variables. We let P_n denote the law of the mean $\bar{X}_n := \frac{1}{n}\sum_{i\leq n} X_i$. This can be shown to satisfy the large-deviations principle, with a rate function I defined as:

$$I(x) := \sup_\theta \{\theta x - \log E(e^{\theta x})\}. \tag{2.1}$$

Let us consider two concrete examples.

- $X_i \sim N(0, \sigma^2)$
  The sum of normal random variables is normal. The probability distribution function of $\bar{X}_n$ is given by

$$f_n(x) = \left(\frac{n}{2\pi}\right)^{\frac{1}{2}} \sigma^{-1} e^{-nI(x)}, \tag{2.2}$$

  where $I(x) = \frac{1}{2}(x/\sigma)^2$.

- $X_i \sim \mathcal{E}(1)$
  The sum of exponentials is Gamma, $\bar{X}_n \sim \Gamma(n, n)$ and f_n is given by

$$f_n(x) = \frac{n^n}{(n-1)!} x^{n-1} e^{-nx}.$$

Substituting Stirling's formula, we get

$$f_n(x) \sim \left(\frac{n}{2\pi}\right)^{\frac{1}{2}} x^{-1} e^{-nI(x)}, \tag{2.3}$$

where $I(x) = x - 1 - \log x$. $I(x)$ can also be evaluated using formula 2.1.

The keen reader may have spotted the similarity between equations 2.2 and 2.3. Of course, both distributions satisfy the large-deviations principle. More interestingly though, they share a similar correction term. This was identified by [Daniels, 54] who derived the following.

Theorem 7 (Daniels). *Let X_i be a sequence of continuous IID random variables. Let f_n denote the probability distribution function of $\bar{X}_n$. Then,*

$$f_n(x) = \sqrt{\frac{n\ddot{I}(x)}{2\pi}} \exp\{-nI(x)\}(1 + O(n^{-1})), \tag{2.4}$$

where I is defined as in equation 2.1.

Sketch of Proof. Consider the Fourier transform of the probability distribution function f_n. We let $M(\theta) = e^{c(\theta)} = E(e^{\theta X})$ and assume that it is well defined on some interval $\mathcal{D}$ containing 0. This implies that $M(\theta)$ is an analytic function on the strip $\mathcal{S} := \{x+iy : x \in \mathcal{D}\}$ and that

$$f_n(x) = \frac{n}{2\pi} \int_{-\infty}^{\infty} M^n(i\theta) e^{-nix\theta} d\theta = \frac{n}{2\pi} \int_{\theta_x - i\infty}^{\theta_x + i\infty} e^{-n[\theta x - c(\theta)]} d\theta. \tag{2.5}$$

The constant $\theta_x \in \mathcal{D}$ is chosen so as to maximise $I(\theta) = c(\theta) - \theta x$ and is therefore the solution of the equation $c'(\theta_x) = x$. We then approximate the contour integral using the Taylor expansion of $c(\theta) - \theta x$ near θ_x. It is not clear that the correction terms of increasing powers of n actually converge, but by deforming the path slightly and integrating equation 2.6 along the path of steepest descent, [Daniels, 54] arrives at equation 2.4 which is uniformly good in x. □

The approximation clearly does not have any meaning when X is a discrete random variable, but if X is a lattice-valued random variable, [Daniels, 54] has extended the Fourier-transform method so as to estimate the size of each atom.

2.2 A New Approach

2.2.1 An Old Approach

Well, actually the new approach is probably not new at all. After reading some of the references kindly provided J. D. Biggins, it seems to me that this chapter will simply reinvent the wheel. In particular, the following scheme to arrive at an integral form of equation 2.4 has been suggested by [Stone, 67] shortly after he stated a higher-dimensions version of equation 2.4 (Stone's theorem 1). Nevertheless, there is some value in carrying out the scheme explicitly which is what we will do next.

2.2.2 Outline of Approach

We present a method of deriving the integral-form of equation 2.4 and show that for a general class of random variables (including strongly non lattice-valued)

$$\frac{1}{\sqrt{n}} \int_L e^{nI(x)} dP_n(x) = \frac{1}{\sqrt{2\pi}} \int_L \sqrt{\ddot{I}(x)} dx \times [1 + O(n^{-1})]. \tag{2.6}$$

Our method involves the following steps:

- Dissect L into smaller intervals L_i.
- Think of the function $e^{nI(x)}$ as a change of measure on each L_i.
- Rescale and use the Central Limit Theorem to evaluate the integral over L_i.

We will need to assume that

- X satisfies $\limsup_{t \to \infty} |E(e^{itX})| < \infty$ (strongly non-lattice).
- $I(x)$ is five-differentiable, $I(x) \in C^5(L, \mathcal{R})$.

Let us consider a short example, demonstrating the approach we will be taking.

2.2.3 Example

For simplicity, we assume that the X_i's satisfy $E(X) = 0$ and $\mathrm{Var}(X) = 1$. Because $P_n \to_{\text{law}} \delta_0$, the point mass at 0, for all smooth functions g,

$$\int g(x)dP_n(x) \to g(0).$$

Let us gauge how quickly this convergence occurs. Let $\bar{X}_{\sqrt{n}} := \frac{1}{\sqrt{n}}\sum_{i\leq n} X_i$ and let $P_{\sqrt{n}}$ denote the law of $\bar{X}_{\sqrt{n}}$. The forgiving reader will observe an ambiguity in the sense that $\bar{X}_{\sqrt{4}} \neq \bar{X}_2$, but will ignore it.

$$\int g(x)dP_n(x) = \int g\left(\tfrac{y}{\sqrt{n}}\right) dP_{\sqrt{n}}(y)$$

We apply Taylor's expansion to g. We also observe that by CLT, $P_{\sqrt{n}} \to_{\text{law}} N(0,1)$. We conclude that

$$\begin{aligned}
\int g(x)dP_n(x) &= \int g\left(\tfrac{y}{\sqrt{n}}\right) dP_{\sqrt{n}}(y) \\
&\approx \int \left(g(0) + \dot{g}(0)\tfrac{y}{\sqrt{n}} + \tfrac{1}{2}\ddot{g}(0)\tfrac{y^2}{n} + \ldots\right) dN(y) \\
&= g(0) + \tfrac{1}{2}g^{(2)}(0)n^{-1} + \ldots
\end{aligned}$$

This is not completely rigorous because we do not know how fast $P_{\sqrt{n}}$ converges to $N(0,1)$. We will review current results on this issue next.

2.2.4 Asymptotic Expansion of the Central Limit Theorem

We need results concerning the convergence of the CLT which are proven in [Petrov, 95]. Let X_i be independent identically distributed random variables with $M(it) := e^{c(it)} := E(e^{iX})$ denoting the Fourier transform. Let X be a random variable such that $\limsup_{t\to\infty} |M(it)| < 1$. The nth cumulant of X is equal to $c^{(n)}(0)$, the nth derivative of the log moment generating function at 0. We let $\gamma_n = c^{(n)}(0)$ and thus we have:

$$c(it) = \log M(it) = \sum_{j\geq 1} \frac{\gamma_j}{j!}(it)^j.$$

WLOG we may assume that $\gamma_1 = E(X) = 0$, $\gamma_2 = \text{Var}(X) = 1$. Let us also assume that $E|X|^m < \infty$ for some integer $m \geq 3$. Define the Hermite polynomials as

$$H_r(x) := (-1)^r e^{\frac{1}{2}x^2} \frac{d^r}{dx^r} e^{-\frac{1}{2}x^2}.$$

Define also the following sequence of polynomials

$$Q_r(x) := -(2\pi)^{-\frac{1}{2}} e^{-\frac{1}{2}x^2} \sum H_{r+2s-1}(x) \prod_{i=1}^{r} \frac{1}{j_i!} \left(\frac{\gamma_{i+2}}{(i+2)!} \right)^{j_i}.$$

The summation is extended over all non-negative integer solutions $(j_1, j_2, \ldots, j_r)$ of the equations $\sum_{i \leq r} i j_i = r$, and $s = \sum_{i \leq r} j_i$. Finally, we define the cumulative distribution functions $F_n(x) := P\left(n^{-\frac{1}{2}} \sum_{j \leq n} X_j < x\right)$ and $\Phi(x) = P(N(0,1) < x)$. We now quote results 1 and 2 from [Petrov, 95].

Result 1. *Let F_n, Q_r and Φ be defined as above. Then uniformly in x we have*

$$|F_n(x) - \Phi(x) - \sum_{i \leq r} Q_i(x) n^{-\frac{1}{2}i}| \leq BOUND$$

where

$$BOUND = c_r \left\{ 2n^{-\frac{1}{2}(r+1)} E\left(|X|^{r+3}\right) + \left(\sup_{|t| \geq \delta} |M(it)| + \frac{1}{2n} \right)^n n^{\frac{1}{2}(r+2)(r+3)} \right\}.$$

c_r is a universal constant independent of X and $\delta^{-1} = 12E(|X|^3)$.

Result 2. *In addition, if f is any $C^1(L, \mathcal{R})$ function, then by partial integration,*

$$\left| \int_L f \left\{ dF_n - dN - \sum_{i \leq r} dQ_i n^{-\frac{1}{2}i} \right\} \right| \leq \left\{ |f(a)| + |f(b)| + \int_L |\dot{f}| dx \right\} \times BOUND.$$

2.2.5 Statement of Main Theorem

We formally state the theorem which we will prove in the next section.

Theorem 8. *Let X_i be a sequence of random variables for which the moment generating function satisfies $\limsup_{t \to \infty} |E(e^{itX})| < \infty$. Let $\bar{X}_n$ have a large-deviations rate function I*

with bounded continuous 3rd and 5th derivatives on (a possibly unbounded) interval $L = [a, b]$. Then,

$$\frac{1}{\sqrt{n}} \int_L e^{nI} dP_n = \frac{1}{\sqrt{2\pi}} \int_L \sqrt{\ddot{I}(x)} dx \times [1 + O(n^{-1})].$$

2.3 Proof

We divide the interval $L = [a, b]$ into small intervals of length at most $2k_n$, where $k_n < n^{-1}$ is to be determined exactly later. We let K_n denote the indexing set of these intervals. For each $i \in K_n$, we denote the corresponding interval by L_i and let x_i be its midpoint. Clearly

$$\int_L e^{nI(x)} dP_n(x) = \sum_{i \in K_n} \int_{L_i} e^{nI(x)} dP_n(x)$$

We aim to evaluate the integral on each interval L_i on its own. As mentioned, we wish to think of the function $e^{nI(x)}$ as an approximate change of measure. This idea was exploited before in proving the lower bound of Cramér's Theorem. On each L_i we let $J(x)$ denote the first three terms in the Taylor expansion of I around x_i and let $R(x)$ denote the remainder.

$$\begin{aligned} J(x) &= I(x_i) + \dot{I}(x_i)(x - x_i) + \tfrac{1}{2}\ddot{I}(x_i)(x - x_i)^2 \\ R(x) &= I(x) - J(x) \end{aligned}$$

We employ the Mean Value Theorem twice and use the boundedness of $||I^{(3)}||_\infty$ to discover that for all L_i and for all $x \in L_i$ we have $|R(x)| \leq ||I^{(3)}|| k_n^3$. Therefore,

$$\sup_{i \in K_n} \sup_{x \in L_i} |e^{nR(x)} - 1| \leq Cnk_n^3.$$

Consequently,

$$
\begin{aligned}
\left|\int_{L_i} e^{nI(x)} - e^{nJ(x)} dP_n\right| &= \left|\int_{L_i} e^{nJ(x)}(e^{nR(x)} - 1) dP_n\right|, \\
&\leq \sup_{x \in L_i} |e^{nR(x)} - 1| \times \int_{L_i} e^{nJ(x)} dP_n, \\
&\leq Cnk_n^3 \times \int_{L_i} e^{nJ(x)} dP_n.
\end{aligned}
$$

Summing over all intervals L_i, we see that

$$
\int_L e^{nI(x)} dP_n = \sum_{i \in K_n} \int_{L_i} e^{nJ(x)} dP_n \times [1 + O(nk_n^3)]. \tag{2.7}
$$

2.3.1 Changing Measure Locally

Let θ_i be the unique constant in $\mathcal{D}$ satisfying $\dot{c}(\theta_i) = x_i$. We have $I(x_i) = x_i\theta_i - c(\theta_i)$ and also $\dot{I}(x_i) = \theta_i$. We define the following new random variable X^i with a law P^i as

$$
P(X^i \in A) := P^i(A) := \frac{\int_A e^{\theta_i x} dP(x)}{E(e^{\theta_i X})} = e^{-c(\theta_i)} \int_A e^{\theta_i x} dP(x). \tag{2.8}
$$

This transformation is identical to transformation 1.1, see [Varadhan, 84] for more details. The Laplace transforms of X and X^i differ by a translation of θ_i. Consequently, the nth cumulant of X^i is equal to $c^{(n)}(\theta_i)$, the nth derivative at θ_i. In particular, $E(X^i) = x_i$. Since I is the Legendre transform of c, it follows that $c^{(n)}(\theta_i)$ and $I^{(n)}(x_i)$ are intimately related. More specifically,

$$
\mathrm{Var}(X^i) = \ddot{c}(\theta_i) = 1/\ddot{I}(x_i).
$$

The law P_n^i of $\bar{X}_n^i := \frac{1}{n}\sum_{j \leq n} X_j^i$ is given by

$$
P(\bar{X}_n^i \in A) = e^{-nc(\theta_i)} \int_A e^{n\theta_i x} dP_n = \int_A e^{n[\theta_i x - c(\theta_i)]} dP_n. \tag{2.9}
$$

We now observe that

$$
\begin{aligned}
J(x) &= I(x_i) + \dot{I}(x_i) \times (x - x_i) + \tfrac{1}{2}\ddot{I}(x_i)(x - x_i)^2 \\
&= \theta_i x_i - c(\theta_i) + \theta_i \times (x - x_i) + \tfrac{1}{2}\ddot{I}(x_i)(x - x_i)^2 \\
&= \theta_i x - c(\theta_i) + \tfrac{1}{2}\ddot{I}(x_i)(x - x_i)^2
\end{aligned}
$$

and hence

$$\int_{L_i} e^{nJ(x)} dP_n = \int_{L_i} e^{\frac{n}{2}\ddot{I}(x_i)(x-x_i)^2} dP^i_n.$$

Finally, if we define $\tilde{X}^i$, a random variable of mean zero and unit variance

$$\tilde{X}^i := (X^i - x^i)\ddot{I}(x_i)^{\frac{1}{2}},$$

then

$$\int_{L_i} e^{nJ(x)} dP_n = \int_{\tilde{L}_i} e^{\frac{n}{2}x^2} d\tilde{P}^i_n.$$

where of course $\tilde{L}_i := (L_i - x_i)\ddot{I}(x_i)^{\frac{1}{2}}$.

2.3.2 Rescaling - Central Limit Theorem

We rescale the integral by $\sqrt{n}$ to arrive at the law $\tilde{P}^i_{\sqrt{n}}$ of $\frac{1}{\sqrt{n}}\sum_{j\leq n} \tilde{X}^i_j$. Since

$$\int_{L_i} e^{nJ(x)} dP_n(x) = \int_{\sqrt{n}\tilde{L}_i} e^{\frac{1}{2}x^2} d\tilde{P}^i_{\sqrt{n}}, \tag{2.10}$$

we would like to replace the law $d\tilde{P}^i_{\sqrt{n}}$ with dN, the law of $N(0,1)$, the normal distribution to which $\tilde{P}^i_{\sqrt{n}}$ converge as $n \to \infty$. We therefore recall results 1 and 2 concerning the asymptotic expansion of the Central Limit Theorem.

2.3.3 Applying CLT-Expansion to the Law of $\tilde{X}^i$.

We need to

- Verify that the conditions of results 1 and 2 hold.

We assumed that X has $\limsup_{t\to\infty}|E(e^{itX})| < \infty$. X^i which is an exponential tilted with respect to X (and hence $\tilde{X}^i$) must satisfy a similar condition. We conclude that we may apply the results.

- Estimate how good is the bound provided in results 1 and 2.

We recall that the bound provided in the these results satisfies:

$$\text{BOUND} = c_r\left\{2n^{-\frac{1}{2}(r+1)}E\left(|\tilde{X}^i|^{r+3}\right) + \left(\sup_{|t|\geq\delta}|M(it)| + \frac{1}{2n}\right)^n n^{\frac{1}{2}(r+2)(r+3)}\right\}.$$

The size of the second term. We assume that $I^{(3)}$ is bounded on L. Thus, $\delta = 12E(|\tilde{X}^i|^3)^{-1}$ is bounded below. It follows that $\sup_{t\geq\delta}|M(it)| < 1$ and the second term of the BOUND decays exponentially in n, uniformly in $\tilde{X}^i$.

The size of the first term. Since the Laplace transform of X^i is well defined on some open interval containing 0, all the moments must exist. Moreover, if $I^{(r+3)}$ is uniformly bounded on L, then $E(|X^i|^{r+3})$ is uniformly bounded on L.

We now let $F_n^i(x) = \tilde{P}^i_{\sqrt{n}}((-\infty, x])$ and deduce the following.

Corollary 1. *Let I be $r+3$-differentiable with bounded derivatives on L. For all $x_i \in L$ there exists N such that for all $n > N$ and for all x,*

$$|F_n^i(x) - \Phi(x) - \sum_{i\leq r} Q_i(x)n^{-\frac{1}{2}i}| \leq c_r n^{-\frac{1}{2}(r+1)}||I^{(r+3)}||_\infty.$$

Moreover, if f is any continuously differentiable function, then

$$
\begin{aligned}
&\left|\int_\alpha^\beta f\left\{dF_n^i - dN - \sum_{i\le r} n^{-\frac{1}{2}i} dQ_i\right\}\right| \\
&\qquad \le c_r n^{-\frac{1}{2}(r+1)} ||I^{(r+3)}||_\infty \left(|f(\alpha)| + |f(\beta)| + \int_\alpha^\beta |\dot{f}| dx\right).
\end{aligned}
$$

We substitute $f(x) := e^{\frac{1}{2}x^2}, -\alpha_i = \beta_i = k_n\sqrt{n}\ddot{I}(x_i)^{\frac{1}{2}}$ and compute the various terms in the above equation.

The RHS: We need to make sure that β_i is bounded, which implies the boundedness of $|f(\alpha)| + |f(\beta)| + \int_\alpha^\beta |\dot{f}| dx$. Since $\beta_i = k_n\sqrt{n}\ddot{I}(x_i)^{\frac{1}{2}}$, when we decide on the size of k_n, we will remember to impose $k_n\sqrt{n} \le 1$. Also, WLOG we may assume $\ddot{I}^{\frac{1}{2}}$ is bounded on L because...

- If L is a bounded interval, then the continuity of $I^{(3)}$ implies the uniform boundedness of $\ddot{I}^{\frac{1}{2}}$ on L.
- If L is unbounded then we dissect L into the intervals $\{L_j = L \cap [j, j+1] : j \in \mathcal{Z}\}$. On each one of L_j we simultaneously perform the analysis separately with different sizes of k_n^j. Because $I^{(3)}$ is bounded on the entire interval L, the convergence to the desired result is uniform across the entire of L. We will again get equation 2.12.

The LHS: The first term on the LHS is the one we want to estimate. The second term is straightforward:

$$
\int_{\sqrt{n}\tilde{L}_i} e^{\frac{1}{2}x^2} dN = \frac{1}{\sqrt{2\pi}} 2\beta_i.
$$

We now consider the case when $r = 1$. The function $Q_1(x)$ is given by

$$
Q_1(x) = -(2\pi)^{-\frac{1}{2}} e^{-\frac{1}{2}x^2} \times \frac{1}{3!}\gamma_3(x^2 - 1)
$$

Looking at $\int e^{\frac{1}{2}x^2} dQ_1$ we see it is an odd function integrated over a symmetric interval which therefore must integrate to 0. We deduce therefore that

$$\left| \int_{L_i} e^{nJ(x)} dP_n - \frac{1}{\sqrt{2\pi}} 2\beta_i \right| \leq c_1 n^{-1} ||I^{(4)}||_\infty. \tag{2.11}$$

This estimate is almost sufficient for our purposes but not quite. We thus consider the case where $r = 2$. We will not be interested in the correction term of order $n^{-\frac{3}{2}}$. We simply wish to evaluate the correction term of order n^{-1} precisely. Assuming $I^{(5)}$ is bounded we get

$$Q_2(x) = -(2\pi)^{-\frac{1}{2}} e^{-\frac{1}{2}x^2} \times \left[\frac{1}{4!}\gamma_4(x^2 - 1) + \tfrac{1}{2} \left(\frac{1}{3!}\gamma_3 \right)^2 (x^5 - 10x^3 + 15x) \right].$$

Splitting Q_2 into its odd and even components and integrating it against $f(x) = e^{\frac{1}{2}x^2}$ we notice that the component containing γ_4 vanishes. After some manipulations, we are left with

$$\left| \int_{L_i} e^{nJ(x)} dP_n - \frac{1}{\sqrt{2\pi}} 2\beta_i \right| \leq c_2 n^{-1} ||I^{(3)}||_\infty^2 2\beta_i.$$

We now sum over all the intervals L_i to deduce that

$$\left| \sum_i \int_{L_i} e^{nJ(x)} dP_n - \sum_i \frac{1}{\sqrt{2\pi}} 2\beta_i \right| \leq c_2 n^{-1} ||I^{(3)}||_\infty^2 \sum_i 2\beta_i. \tag{2.12}$$

2.3.4 Evaluating $\sum_i \beta_i$

We recall that $\beta_i = \sqrt{n} k_n \sqrt{\ddot{I}(x_i)}$. Hence as $n \to \infty$ we must have

$$\lim_{n\to\infty} \frac{1}{\sqrt{n}} \sum_i 2\beta_i = \int_L \ddot{I}(x)^{\frac{1}{2}} dx. \tag{2.13}$$

There is of course an error term but we may choose k_n such that the correction term in the above equation decays exponentially in n. Equation 2.12 now becomes

$$\left|\sum_i \int_{L_i} e^{nJ(x)} dP_n - \frac{\sqrt{n}}{\sqrt{2\pi}} \int_L \ddot{I}(x)^{\frac{1}{2}} dx\right| \leq c_3 n^{-\frac{1}{2}} \int_L \ddot{I}(x)^{\frac{1}{2}} dx.$$

We divide by $\sqrt{n}$ and get

$$\frac{1}{\sqrt{n}} \sum_i \int_{L_i} e^{nJ(x)} dP_n = \frac{1}{\sqrt{2\pi}} \int_L \ddot{I}(x)^{\frac{1}{2}} dx \times [1 + O(n^{-1})]. \tag{2.14}$$

We combine this with equation 2.10

$$\int_L e^{nI(x)} dP_n = \sum_i \int_{L_i} e^{nJ(x)} dP_n \times [1 + O(nk_n^3)].$$

We let k_n to be of size at most n^{-1} and as fine as required to allow an exponential convergence of the summation in equation 2.13. The proof of theorem 8 is now complete.

2.3.5 Two Related Observations

We first prove a direct extension of the above theorem and then look if we can "invert" the result.

Suppose G and H are any bounded continuous functions. [Varadhan, 84] showed that

$$\lim_{n\to\infty} \frac{1}{n} \log \int He^{nG} dP_n = \sup_{x\in\mathcal{R}} \{G(x) - I(x)\}.$$

Assume $\sup_{x\in\mathcal{R}}\{G(x) - I(x)\} = 0$. Let L be the compact set on which I and G agree.

$$L := \{x : G(x) = I(x)\}$$

We let ∂L denote the boundary L counting points which are not part of an interval twice. We assume that $\ddot{I}(x) - \ddot{G}(x)$ is strictly positive for all $x \in \partial L$. We further assume that

L is composed of the union of closed intervals. Repeating the method of theorem 8 at the boundary points, we conclude that

$$\int He^{nG}dP_n = \frac{\sqrt{n}}{\sqrt{2\pi}}\int_L H(x)\ddot{I}(x)^{\frac{1}{2}}dx + \tfrac{1}{2}\sum_{x\in\partial L} H(x)\left(\frac{\ddot{I}(x)}{\ddot{I}(x)-\ddot{G}(x)}\right)^{\frac{1}{2}} + O(n^{-\frac{1}{2}}).$$

We have shown that the Central Limit Theorem can be used to prove the large-deviations principle. The reverse is also true. Let X be a random variable satisfying the large-deviations principle with rate function $I(x) = \sup_\theta\{\theta x - c(\theta)\}$. WLOG let $E(X) = 0$ and $\mathrm{Var}(X) = 1$. Then $\bar{X}_{\sqrt{n}} = \frac{1}{\sqrt{n}}\sum_{i\leq n} X_i$ also satisfies a large-deviations principle, with a rate function I_n satisfying:

$$I_n(\sqrt{n}x) = \sup_\theta\left\{\sqrt{n}\theta x - nc(\tfrac{\theta}{\sqrt{n}})\right\} = nI(x).$$

Letting $n \to \infty$ we expand $I_n(x) = nI(\frac{x}{\sqrt{n}})$ near 0 using Taylor's expansion. Observing that $I(0) = \dot{I}(0) = 0$ and that $\ddot{I}(0) = \mathrm{Var}(X) = 1$ we deduce that

$$I_n(x) \to \tfrac{1}{2}x^2$$

uniformly on compact subsets of $\mathcal{R}$. This implies the Central Limit Theorem either directly (using total variation distance) or by observing that by the properties of the Legendre transform, $c_n(\theta)$ must converge to the moment generating function of a Normal random variable.

$$c_n(\theta) \to \tfrac{1}{2}\theta^2.$$

2.4 Final Comments

Our new approach as it stands, is of limited interest. After all, the result we have proved has been proven already almost half a century ago. Nevertheless it is of interest to see if we can extend it to "improve the resolution" of other results. There are a few natural candidates.

- Cramér's Theorem in higher dimensions.

This should be relatively straightforward, as the Central Limit Theorem clearly holds.

- The Ellis-Gartner Theorem.

  The Ellis-Gartner Theorem replaces independent random variables with convergent moment generating functions. We know that the rate of convergence of the moment generating function, controls the rate of convergence in law. Consequently, if $\frac{1}{n}c_n(n\lambda)$ converges fast enough, this should provide us with a weaker version of the CLT. Otherwise, our approach lends itself quite naturally.

Chapter 3

Introduction to Branching Diffusion Processes

In future chapters we study the almost-sure behaviour of branching diffusion processes using probabilistic methods. Indeed, we will almost exclusively ignore the two primary uses of branching diffusion processes:

- Modelling the spatial behaviour of a population.
  The foundations of this branch of Mathematical Biology were laid by [Fisher, 37] and [Kolmogorov *et al.*, 37]. We refer the reader to [Murray, 89] and [Stekel *et al.*, 95] for current progress.

- Studying differential Reaction-Diffusion equations.
  The FKPP equation,

  $$\phi_t = \tfrac{1}{2}\phi_{xx} + \phi(\phi - 1) \tag{3.1}$$

  is an example of such a Reaction-Diffusion equation. It was [McKean, 75] who first realised that the solutions to this equation can be expressed as integrals over P, a probability space associated with a branching diffusion process.

The aim of this chapter is thus merely to acquaint the reader to our notations and direct him to areas where our results fit contextually. [Ikeda *et al.*, 68], [McKean, 75], [Bramson, 83], [Biggins, 77,79], [Uchiyama, 82], [Neveu, 87], [Chauvin *et al.*, 88,90,91] and [Champney *et al.*, 95] do more justice to this beautiful subject than this short chapter.

3.1 Constructing a Branching Diffusion Process

A branching diffusion process is a conceptually simple model. It consists of a system of particles each breeding and diffusing. We assume each particle obeys the following "commandments"

- Thou shall inherit your position at birth from your ancestor.
- Thou shall behave independently from all other particles.
- Thou shall diffuse in space according to diffusion process $X(t)$.
- Thou shall die after an exponential time $\tau \sim \mathcal{E}(r)$.
- Thou shall give birth at death to $1 + C$ offspring.

Throughout, we restrict ourselves to death-free continuous-time processes (see [Biggins, 77] for discrete-parameter BDP's) on $\mathcal{R}$. We formally construct the probability space for such a branching diffusion process.

3.1.1 The Branching Process (Galton-Watson)

Consider a Markovian birth process $|N_t|$ representing the number of particles alive. Each particle lives for an exponential time of rate r and gives particle to $1 + C$ particles, where $E(C) = \mu$ and

$$c_n := P(C = n).$$

The probability generating function $H(\theta)$, and the infinitesimal generator function $a(x)$, of the process are

$$H(\theta) := E(\theta^C) = \sum_{n \in \mathcal{N}} c_n \theta^n \tag{3.2}$$

$$a(x) := rx\{H(x) - 1\}. \tag{3.3}$$

We note that $a(1) = a(0) = 0$ and $\lim_{x \uparrow 1} a(x) = -rc_\infty$. On $0 < x < 1$, $a(x)$ is convex and negative.

3.1.2 The Probability Space

For simplicity we make the following assumptions about the breeding process.

- **Assumption:** $E(C \log C) < \infty$.
  We want our process to be finite so that $c_\infty = 0$ and $a(x)$ is continuous at 1. If $|N_t|$ denotes the number of particles alive at time t, we would also like the martingale $e^{-r\mu t}|N_t|$ to converge to a positive limit. This happens if and only if $E(C \log C) < \infty$.

- **Assumption:** For simplicity $r = 1$.

We denote a typical path of the diffusion process by $X(t)$. At birth, the parent particle and its offspring share the same spatial position. Consequently, All particles share the same spatial law. From birth onwards, each offspring follows an independent diffusion path, but equality in law is maintained. We follow [Neveu, 87]'s construction. A finite sequence i of numbers will label each particle, starting with the first particle labelled $\emptyset$. Each particle i has $1 + C_i$ descendants $i0, i1 \ldots iC_i$. Let $I = \cup_{n \in \mathcal{N}} \mathcal{Z}^{+n}$ be the space of labels. Let τ_i be the lifetime of particle i ($i \in I$). Particle i will thus be born at time

$$T_i = \sum_{k=0}^{k=n-1} \tau_{j_1 \ldots j_k} \quad \text{if } i = j_1 \ldots j_n.$$

The τ_i are assumed to be strictly positive random variables satisfying the non-explosion condition: $\{i : T_i \leq t\}$ is finite for all t. The trajectories of particles are continuous maps X_i of the time intervals $[T_i, T_i + \tau_i]$ into $\mathcal{R}$ such that $X_{ic}(T_{ic}) = X_i(T_i + \tau_i)$ for every $i \in I$ and $c \in C_i$.

A point $\omega \in \Omega$ is a collection $\{\tau_i, X_i, C_i : i \in I\}$ satisfying the above conditions. Let

$$N_t(\omega) = \{i : T_i \leq t < T_i + \tau_i\}$$

be the set of particles alive at time t. The filtration $\{\mathcal{F}_t : t \in \mathcal{R}^+\}$ on Ω is generated by $\{N_t, (X_i(t), i \in N_t)\}_{t \in \mathcal{R}^+}$. There exists a unique probability measure P on $(\Omega, \mathcal{F}_\infty)$ such that $\{X_i : i \in I\}$ is an independent family of diffusion processes with each X_i started at $X_i(T_i)$, stopped after an exponential time τ_i of mean 1, and giving birth to $1 + C_i$ offspring at its time of death.

3.2 BDP's and Reaction-Diffusion Equations

One of the main difficulties in studying BDP's is that the behaviour in expectation and almost surely do not necessarily agree. The behaviour in expectation is well understood because the following many-to-one picture holds: If $A \subset \mathcal{R}$ then the expected number of particles in A satisfies

$$E\sum_{i\in N_t} 1_A(X_i(t)) = E\sum_{i\in N_t} P(X_i(t) \in A) = E(|N_t|)P(X(t) \in A). \tag{3.4}$$

Thus we can study the growth process and the diffusion process separately.

- **Example: A dyadic branching Brownian motion.**
  Our system consists of particles performing independent Brownian motion, each particle splitting into two particles at death. The expected number of particles is $E(|N_t|) = e^t$ and $B(t) \sim N(0,t)$. We define the *expectation* wavefront u_t, by imposing

$$E\left\{\sum_{i\in N_t} 1_{[u_t,\infty)}(B_i(t))\right\} = 1$$

  By substituting in equation 3.4 we see that

$$u_t = t\sqrt{2} - \frac{1}{2\sqrt{2}}\log t + O(1). \tag{3.5}$$

To study the almost sure behaviour of a BDP, we must study the associated Reaction-Diffusion equation using martingale theory. We give two examples of such an approach.

3.2.1 The Galton-Watson Process

We recall that $|N_t|$ is an integer valued Markovian birth process representing the number of particles alive, with a probability generating function $H(\theta) = E(\theta^C)$, and an infinitesimal generator function $a(x) = rx\{H(x) - 1\}$. Suppose that for $\phi : \mathcal{R} \to \mathcal{R}$, we wanted to solve the differential equation

$$\dot{\phi} = a(\phi)$$

under the boundary condition $\phi(0) = u \in (0,1)$. We would observe that M_s, defined on $[0,\infty]$ as

$$M_s = \prod_{i \in N_s} \phi(t-s) = \phi(t-s)^{|N_s|},$$

satisfies

$$E(dM_s) = \sum_{i \in N_s} ds\{a(\phi) - \dot{\phi}\} \prod_{i \neq j} \phi(t-s) = 0.$$

More precisely, M_s is a true martingale (see [Watanabe, 67] for more details). By the martingale property $E(M_t) = M_0$, ϕ is determined explicitly as

$$\phi(t) = E(\phi(0)^{|N_t|}) = E(u^{|N_t|}).$$

Letting $t = 0$ and running M_s on $(-\infty, 0]$ we deduce that

$$W := -\lim_{s \to \infty} |N_s| \log \phi(-s)$$

exists and defines a non-negative random variable. With more care (see [Neveu, 87]) we can determine its Laplace transform. We deduce that as $t \to \infty$, the function $\phi(t) \to 0$ exponentially, and that

$$e^{r\mu t} \log \phi(-t) \tag{3.6}$$

always has a finite limit. We now prove a small lemma concerning $G(\lambda, t) := E(\lambda^{|N_t|})$.

Lemma 2. *Let* $\epsilon > 0$. *Then,*

$$G(1-e^{-r\mu(1-\epsilon)t}, t) \to 0 \quad \textit{exponentially as } t \to \infty.$$

Proof. Fix $\lambda = 1 - e^{-(1-\epsilon)r\mu t}$ and observe how $G(\lambda, s)$ behaves over time $s \in [0, t]$. We condition on the first time of birth to obtain:

$$\begin{aligned} G(\lambda, s) &= \lambda e^{-rs} + \int_0^s re^{-ru} \sum_{n=0}^{\infty} c_n G(\lambda, s-u)^{n+1} du, \\ &= \lambda e^{-rs} + e^{-rs} \int_0^s re^{ru} G(\lambda, u) H(G(\lambda, u)) du. \end{aligned}$$

Differentiating with respect to time, we obtain

$$\dot{G} = a(G) = -rG(1 - H(G)).$$

We see that G_λ and ϕ are identical modulo translation, and we can evaluate $G(\lambda, t)$ using ϕ. Since $G(\lambda, 0) = \lambda = 1 - e^{-r\mu(1-\epsilon)t}$, we can see using equation 3.6 that

$$G_\lambda(t) = \phi(t - \tau)$$

for some $\tau < (1 - \frac{1}{2}\epsilon)t$. We conclude that $G(\lambda, t) < \phi(\frac{1}{2}\epsilon t)$ and both decay exponentially in time. □

The following corollary has been proved by (amongst others) [Dekking & Grimmmet, 1988] for discrete time, so here is the continuous version.

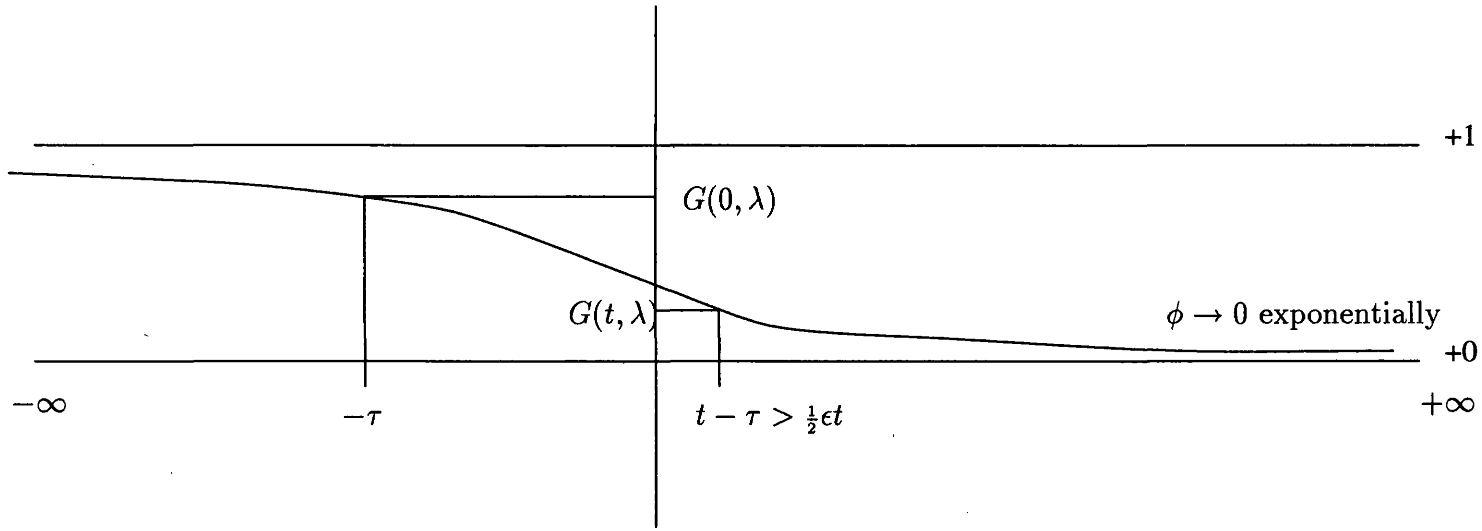
Corollary 2. *For every $\epsilon > 0$, $P(N_t < e^{r\mu(1-\epsilon)t})$ decays exponentially.*

Proof. Let $n_t = e^{r\mu(1-\epsilon)t}$. We have

$$G(1 - \frac{1}{n_t}, t) \geq P(N_t < n_t) \left(1 - \frac{1}{n_t}\right)^{n_t} \to e^{-1} P(N_t < n_t).$$

Since the LHS decays exponentially, so must the RHS. □

A fixed travelling wave $\dot{\phi} = a(\phi)$



3.2.2 Dyadic Branching Brownian Motion

The dyadic branching breeding process has an infinitesimal generating function

$$a(x) = x^2 - x,$$

while the generator of a Brownian motion is given by $\mathcal{G}f = \frac{1}{2}f_{xx}$. It was first observed by [McKean, 75] that if $\phi(x,t)$ is a solution to

$$\phi_t = \mathcal{G}\phi + a(\phi) = \tfrac{1}{2}\phi_{xx} + \phi(\phi - 1), \tag{3.7}$$

with bounded initial conditions $|\phi(x,0)| \leq 1$, then

$$M^x(s) = \prod_{i \in N_s} \phi(x + B_i(s), t - s), \tag{3.8}$$

is a bounded (hence true) martingale on $[0,t]$. To see this, condition on both the position and time of the first birth and then differentiate to see that $E(dM^x_s) = 0$. Using the martingale property $E(M^x(t)) = M^x(0)$, we get

$$\phi(x,t) = E \prod_{i \in N_t} \phi(x + B_i(t), 0). \tag{3.9}$$

Thus we have found the unique bounded solution.

From an analysis point of view, we may be interested in finding travelling waves. These correspond to solutions of the form $\phi(x,t) = w_\lambda(x - \lambda t) : \mathcal{R} \to [0,1]$ which solve Kolmogorov's equation

$$\tfrac{1}{2}\ddot{w} + \lambda\dot{w} + w(w-1) = 0.$$

Monotone solutions exist if $|\lambda| \geq \sqrt{2}$ (for studies on non-monotone solutions when $|\lambda| < \sqrt{2}$ see [Britton, 86] and [Murray, 89]). See [Warren, 95] for a probabilistic investigation of the convergence of initial conditions to monotone travelling waves. From a probabilistic point of view, the most interesting initial condition is the step function. This is because of the following interpretation: Let R_t denote the position right-most particle of the dyadic BBM,

$$R_t := \sup\{B_i(t) : i \in N_t\}.$$

By conditioning on the first dyadic splitting, we can show that

$$\phi(x,t) := P(R_t < x)$$

satisfies equation 3.7. By studying McKean's representation (equation 3.9) carefully, [Bramson, 83] was able to show that $\phi(x + v(t), t)$ converges to a travelling wave, where

$$v(t) = t\sqrt{2} - \frac{3}{2\sqrt{2}} \log t + O(1).$$

This result seemingly contrasts with the faster expectation wavefront speed u_t defined earlier in equation 3.5. We conjecture that almost surely

$$\limsup \frac{R_t - t\sqrt{2}}{\log t} = \frac{1}{2\sqrt{2}}$$

which agrees with the expectation wavefront. Our belief stems from the fact the $P(R_t > u_t) \sim t^{-1}$. If we could use BCL we would then conclude that $R_t > u_t$ infinitely often. Unfortunately, the experiments we perform are not independent and several attempts to prove the above conjecture have failed.

The FKPP can also be linearised near the equilibrium solutions $\phi \equiv 1$ and $\phi \equiv 0$. For example, when ϕ is near 0 the FKPP equation is approximated by

$$\phi_t = \tfrac{1}{2}\phi_{xx} - \phi.$$

This is mirrored by the existence of a family of linearised martingales

$$Z_\lambda(t) = e^{-t} \sum_{i \in N_t} e^{\lambda B_i(t) - \frac{1}{2}\lambda^2 t}.$$

For $|\lambda| < \sqrt{2}$, these converge to $Z_\lambda(\infty)$, strictly positive random variables which are analytical in λ [Biggins, 92]. In fact, using Fourier expansion techniques not dissimilar to the ideas in [Daniels, 54], [Uchiyama, 82] showed that for a function g with compact support, if $E(C^3) < \infty$ then

$$e^{-t} \sum_{i \in N_t} \sqrt{t} g(B_i(t) - \lambda t) e^{\lambda B_i(t) - \frac{1}{2}\lambda^2 t}$$

converge. More details on this result can be found in the next chapter. However, multiplicative martingales contain more information than additive martingales. For example, $Z_{\sqrt{2}}(t) \to 0$ almost surely near the wavefront, but the corresponding product martingale still converges to a non-trivial random variable. Similarly, when $E(C \log C) = \infty$ the limit of product martingales exists but not the limit of linear martingales. This has promoted the use of tree methods as ways of investigating branching diffusion processes. Tree methods involve the study of the particles who first among their ancestors cross a space-time boundary l_τ. As we deform the boundary continuously in τ, the set of particles crossing l_τ forms a branching process which we can study. To fully appreciate this technique see the pioneering paper by [Neveu, 87] and further work by [Chauvin, 91].

To summarise, we have seen how the study of BDP's is intimately linked to studying the solutions to Reaction-Diffusion equations of the form

$$\phi_t = \mathcal{G}\phi + a(\phi).$$

In expectation, we can separate the "diffusion term" $\mathcal{G}\phi$ and the "reaction term" $a(\phi)$ but not when studying the almost sure behaviour.

Chapter 4

Almost Sure Path Properties of a Branching Brownian Motion

4.1 Introduction

In this chapter we analyse the almost-sure behaviour of a dyadic branching Brownian motion using large-deviations techniques. We formulate a large-deviations principle for the *almost sure* rate of growth of particles along any (suitably scaled) path. The result follows directly from theorem 5 and work by [Biggins, 77], [Uchiyama, 82] and [Chauvin *et al.*, 88]. We combine Schilder's Theorem with the many-to-one picture (equation 3.4) to deduce the expected rate function for each BBM-path. This provides the upper bound for the *almost sure* rate function. We then pull together the results by [Biggins, 77] and [Chauvin *et al.*, 88] to prove the lower bound.

4.1.1 Scaling The Branching Brownian Motion

At a fixed time T, let us scale the BBM by a factor of T in both the space and time coordinates. We utilise the projection π_T defined in equation 1.2. We get a branching process on the time-parameter set $[0,1]$. Specifically, for every $i \in N_T$, let $x_i^T \in C_0$; the space of continuous functions from $[0,1]$ to $\mathcal{R}$ started at 0; be the T-scaled path of particle i, defined as

$$\pi_T x_i(t) := x_i^T(t) := \tfrac{1}{T} B_{a_{tT}(i)}(tT).$$

Here, $a_{tT}(i)$ denotes the unique ancestor of particle i at time tT. If $D \subset C_0$, then let $M_D(T)$ denote the set of particles at time T whose T-scaled path is in D. Also if $D|_\theta := \{x \in \mathcal{C}^0([0,\theta],\mathcal{R}) : \exists z \in D, \quad x(t) = z(t) \quad \forall t \in [0,\theta]\}$, let $M_D(T,\theta)$ denote the set of particles whose T-scaled path is in $D|_\theta$ up to time $\theta \leq 1$.

$$\begin{aligned} M_D(T) &:= \{i \in N_T : x_i^T \in D\}, \\ M_D(T,\theta) &:= \{i \in N_{\theta T} : x_i^T|_\theta \in D|_\theta\}. \end{aligned}$$

The function $x|_\theta \in \mathcal{C}^0([0,\theta],\mathcal{R})$ is x truncated at time θ.

4.2 Rate of Growth in Expectation

We denote the law of a standard Brownian Motion run until time 1 as P_1 and the law of an individual T-scaled path x_i^T path by P_T. Using Schilder's Theorem (theorem 5), we know that P_T satisfy a large-deviations principle with rate function I.

$$I(x) := \begin{cases} \frac{1}{2}\int_0^1 \dot{x}(t)^2 dt & \text{if } x \in C_1, \\ \infty & \text{otherwise} \end{cases}$$

By running the BBM until time θT where $\theta \in [0,1]$, we get a slightly modified rate function $I(x,\theta) = \frac{1}{2}\int_0^\theta \dot{x}^2 dt$.

Let D be a subset of C_0. By applying the many-to-one picture (equation 3.4) we get

$$E(|M_D(T,\theta)|) = E(|N_{\theta T}|)P(x^T|_\theta \in D|_\theta),$$

whence the following result is immediate:

Result 3. *Let $J(x,\theta) := \theta - I(x,\theta)$. If A and D are an open subset and a closed subset of $C_0|_\theta$ respectively, then*

$$\liminf_{T\to\infty} T^{-1} \log E(|M_A(T,\theta)|) \geq \sup_{x\in A} J(x,\theta),$$
$$\limsup_{T\to\infty} T^{-1} \log E(|M_D(T,\theta)|) \leq \sup_{x\in D} J(x,\theta).$$

As a matter of convenience, for all θ and for all sets B, we let

$$I(B,\theta) := \inf_{x\in B} I(x,\theta),$$
$$J(B,\theta) := \sup_{x\in B} J(x,\theta).$$

Also, the reader should always assume $\theta = 1$ unless otherwise specified. So $I(x) = I(x,1)$, $J(x) = J(x,1), \ldots$. We note that I is lower-semicontinuous while J is upper-semicontinuous in the sense that $\lim_{z\to x} I(z) \geq I(x)$ and $\lim_{z\to x} J(z) \leq J(x)$.

4.3 Rate of Growth Almost-Surely

We wish to transform the result in probability to an almost sure result, so that for some function $K(x)$ to be determined later (which we might hope looks like $J(x)$), we have almost surely;

$$\lim_{T\to\infty} T^{-1} \log |M_A(T)| \geq \sup_{x\in A} K(x),$$
$$\lim_{T\to\infty} T^{-1} \log |M_D(T)| \leq \sup_{x\in D} K(x).$$

We certainly expect $K(x) \leq J(x)$ for all $x \in C_1$. We can improve this upper bound by considering the following: Suppose that for some $\theta \in [0,1]$ we have $J(D,\theta) < 0$. Then, using result 3 and Chebychev's inequality, we deduce that as T tends to infinity,

$$P(|M_D(T,\theta)| > 0) \leq \exp\{TJ(D,\theta)\} \to 0.$$

Intuitively, this implies that $\lim_{T\to\infty} |M_D(T,\theta)| = 0$, and consequently also $\lim_{T\to\infty} |M_D(T)| = 0$ almost surely. This is a better indication as to how J "controls" K. It turns out that this upper bound is actually tight and distinguishes exactly between the different rates of growth. We now begin the rigorous study.

4.3.1 Upper Bound

Lemma 3. *Let D be a closed subset of C_0. Then for every $\theta \in [0,1]$, we have almost surely:*

$$\limsup_{T\to\infty} T^{-1}\log|M_D(T,\theta)| \leq J(D,\theta).$$

Proof. Suppose that the result is false. Then there exists a θ and an event W with $P(W) > 0$ such that, for every $\omega \in W$, $\limsup_{T\to\infty} T^{-1}\log|M_D(T,\theta)| > J(D,\theta)$. Hence if

$$W_n := \{\omega \in \Omega : \limsup_{T\to\infty} T^{-1}\log|M_D(T,\theta)| > J(D,\theta) + n^{-1}\},$$

then $P(W_n) > 0$ for some n. It is now clear that

$$\limsup T^{-1}\log E(|M_D(T,\theta)|) \geq J(D,\theta) + n^{-1}$$

contradicting result 3. □

In particular, we see that if for some $\theta \leq 1$ we have $J(D,\theta) < 0$, then, almost surely, $\lim_{T\to\infty}|M_D(T,\theta)| = 0$. Since $x_i^T \in D$ implies that $x_i^T|_\theta \in D|_\theta$ we must also have that $\lim_{T\to\infty}|M_D(T)| = 0$ almost surely. This leads us to the following definition and the upper bound result:

Definition (The Almost Sure Rate Function). *Let $\theta_0 \in [0,1] \cup \{\infty\}$ be the last time at which $J(x,\theta)$ is non-negative, $\theta_0 := \inf\{\theta \in [0,1] : J(x,\theta) < 0\}$. Define $K(x,\theta)$ as:*

$$K(x,\theta) := \begin{cases} J(x,\theta) & \text{if } \theta \leq \theta_0, \\ -\infty & \text{otherwise} \end{cases}$$

Result 4. *Let $\theta \in [0,1]$ and let $D \subset C_0$ be closed. Then,*

$$\limsup_{T\to\infty} T^{-1}\log|M_D(T)| \leq \sup_{x\in D} K(x).$$

4.3.2 Lower Bound

We shall prove the lower bound in stages. We first consider open sets around linear functions, then open sets around piecewise-linear functions, and finally arbitrary open sets. We use the following definition of an open ϵ-neighbourhood:

$$A(x,\epsilon) := \{z \in C_0 : ||z-x|| < \epsilon\} = \{z \in C_0 : \sup_t |x(t) - z(t)| < \epsilon\}.$$

Lemma 4. *Let $x(t) = \lambda t$ be a linear function with $0 \leq \lambda < \sqrt{2}$. For every $\epsilon > 0$, we have almost surely,*

$$\liminf_{T\to\infty} T^{-1} \log |M_{A(x,\epsilon)}(T)| \geq 1 - \tfrac{1}{2}\lambda^2.$$

Our proof relies on a result by [Biggins, 77] which we quote again

Result 5. *If $E(C \log C) < \infty$ then for a branching Brownian motion, if $|\lambda| < \sqrt{2}$, then*

$$Z_\lambda(t) = e^{-t} \sum_{i \in N_t} e^{\lambda B_i(t) - \frac{1}{2}\lambda^2 t}$$

converge to $Z_\lambda(\infty)$, strictly positive random variables.

Proof. For every ω we define a sequence of probability measures P_t on $C([0,1],\mathcal{R})$.

$$P_t(A) = |N_t(A)|/|N_t|$$

By the upper bound result, these actually satisfy a large-deviations principle, namely for all closed sets D

$$\lim_{t\to\infty} \frac{1}{t} \log P_t(D) \leq \sup_{y \in D} K(y) - 1.$$

Following a scheme by [Dembo *et al.*, 95] for every $x \in C([0,1],\mathcal{R})$ we define a large-deviations rate function $L(x)$ by taking $A(x, \frac{1}{n})$, a sequence of decreasing open intervals containing x and defining $L(x)$ by

$$L(x) := -\lim_{n\to\infty}\lim_{t\to\infty}\frac{1}{t}\log P_t(A_n).$$

We remind the reader that this function may depend on the particular ω we chose. From above discussion we have that $L(x) \geq 1 - K(x)$. Now, for $x = \lambda t$, we notice that

$$Z_\lambda(t) = e^{-t}|N_t|\int_{y\in C[0,1]} e^{t(<x,y>-\frac{1}{2}<x,x>)}dP_t(y).$$

By Varadhan's theorem, we therefore have that

$$\limsup_{t\to\infty}\frac{1}{t}\log Z_\lambda(t) \leq \sup_y\{<x,y> -\tfrac{1}{2} < x,x > -L(y)\}.$$

From result 7 we have that *for every* ω the LHS is actually 0. Because $L(y) \geq 1 - K(y) \geq I(y)$ we have that the RHS is negative unless $L(x) = \frac{1}{2} < x, x >$. We conclude that for every ω we have $L(x) = \frac{1}{2}\lambda^2$ and $K(x) = 1 - \frac{1}{2}\lambda^2$. □

We now wish to glue together several linear functions.

Definition. *Let x be a piecewise linear function. We say x satisfies the lower bound condition until time $\theta_1 > 0$, if for all $\epsilon > 0$, almost surely*

$$\liminf_{T\to\infty} T^{-1}\log|M_{A(x,\epsilon)}(T,\theta_1)| \geq J(x,\theta_1).$$

Suppose x satisfies the lower bound condition until θ_1. If from θ_1 until θ_2, x is a linear function satisfying $\dot{x} = \lambda$, then we wish to show that x satisfies the lower bound condition until θ_2. We first assume $|\lambda| < \sqrt{2}$. We will run the process until time $\theta_1 T$, arriving at $M_{A(x,\epsilon)}(T,\theta_1)$ particles. We will then run $M_{A(x,\epsilon)}(T,\theta_1)$ independent copies from time $\theta_1 T$ to time $\theta_2 T$, and add them all together. We require the following two definitions. The first simply introduces the change in the rate function over the interval $[\theta_1,\theta_2]$. The second defines a random variable, very much like $M_{A(x,\epsilon)}(T)$, for each $i \in N(\theta_1 T)$ which simply counts the offspring of i whose T-scaled paths follow x closely over the interval $[\theta_1,\theta_2]$. Formally,

$$J(x,\theta_1,\theta_2) := \theta_2 - \theta_1 - \tfrac{1}{2}\int_{\theta_1}^{\theta_2} \dot{x}^2(t)dt,$$
$$M^i_{A(x,\epsilon)}(T,\theta_1,\theta_2) :=$$
$$\left\{j \in N(\theta_2 T) : a(j) = i, \left|(x_j^T(t) - x_j^T(\theta_1)) - (x(t) - x(\theta_1))\right| < \epsilon \text{ for all } t \in [\theta_1,\theta_2]\right\}.$$

It is a simple matter to verify that since all particles are independent, M^i all share the same law and that

$$M_{A(x,\epsilon)}(T,\theta_2) \supseteq \sum_{i \in M_{A(x,\frac{1}{2}\epsilon)}(T,\theta_1)} M^i_{A(x,\frac{1}{2}\epsilon)}(T,\theta_1,\theta_2).$$

Also apparent is the additivity of the rate function:

$$J(x,\theta_2) = J(x,\theta_1) + J(x,\theta_1,\theta_2),$$

Lemma 5. *Let $x \in C_1$ be piecewise-linear satisfying the lower bound condition up until time θ_1. Let $\dot{x} = \lambda$ on $[\theta_1,\theta_2]$, with $|\lambda| < \sqrt{2}$. Then, for every $\epsilon > 0$,*

$$\liminf_{T\to\infty} T^{-1} \log |M_{A(x,\epsilon)}(T,\theta_2)| \geq J(x,\theta_2).$$

Proof. Let $\delta > 0$ be arbitrary. We use the Strong Law of Large Numbers. From previous discussion, we have

$$\begin{aligned}
e^{-(J(x,\theta_2)-2\delta)T}|M_{A(x,\epsilon)}(T,\theta_2)| &\geq e^{-(J(x,\theta_2)-2\delta)T} \sum_{i \in M_{A(x,\frac{1}{2}\epsilon)}(T,\theta_1)} |M^i_{A(x,\frac{1}{2}\epsilon)}(T,\theta_1,\theta_2)| \\
&\geq e^{-(J(x,\theta_1)-\delta)T}|M_{A(x,\frac{1}{2}\epsilon)}(T,\theta_1)| \ldots \\
&\quad \times \frac{1}{|M_{A(x,\frac{1}{2}\epsilon)}(T,\theta_1)|} \sum_{i \in M_{A(x,\frac{1}{2}\epsilon)}(T,\theta_1)} e^{-(J(x,\theta_1,\theta_2)-\delta)T}|M^i_{A(x,\frac{1}{2}\epsilon)}(T,\theta_1,\theta_2)|
\end{aligned}$$

The IID random variables $e^{-(J(x,\theta_1,\theta_2)-\delta)T}|M^i_{A(x,\frac{1}{2}\epsilon)}(T,\theta_1,\theta_2)|$ inside the summation tend a.s. to ∞ as $T \to \infty$ (lemma 4). We average over *an independent* random number $|M_{A(x,\frac{1}{2}\epsilon)}(T,\theta_1)|$ of particles which tends to ∞ a.s. as $T \to \infty$ so that the SLLN still holds. By the induction hypothesis, x satisfies the lower bound condition until time θ_1, and thus the random variable $e^{-(J(x,\theta_1)-\delta)T} M_{A(x,\frac{1}{2}\epsilon)}(T,\theta_1)$ also tends almost surely to infinity. We conclude that the RHS (and hence the LHS) tends to infinity almost surely as T tends to infinity, and hence for any sequence of increasing times, $T_n \uparrow \infty$, almost surely,

$$\liminf_{T_n \to \infty} T_n^{-1} \log |M_{A(x,\epsilon)}(T_n, \theta_2)| \geq J(x,\theta_2) - 2\delta.$$

To extend this result for all time T, we use the fact that for every open interval A, the map $T \to M_A(T)$ is continuous almost surely. We use lemma 6 (see later) to deduce that almost surely

$$\liminf_{T \to \infty} T^{-1} \log |M_{A(x,\epsilon)}(T, \theta_2)| \geq J(x,\theta_2) - 2\delta.$$

Letting $\delta \downarrow 0$ concludes the proof. □

We turn our attention to the case where x satisfies the lower bound condition until time θ_1, while between θ_1 and θ_2, the gradient $\dot{x} = \lambda > \sqrt{2}$. We of course insist on $J(x,\theta_2) > 0$.
Heuristics: The following proof is in principle the same as that of lemma 5 above. We run the process until time θ_1, arriving (using lemma 4) at an almost sure $M_{A(x,\epsilon)}(T,\theta_1)$ particles. We then run independent copies on $[\theta_1 T, \theta_2 T]$.

We replace the almost sure number of particles $M^i_{A(x,\epsilon)}(T,\theta_1,\theta_2)$ produced by an independent copy, with $P^i_{A(x,\epsilon)}(T,\theta_1,\theta_2)$, the probability of an independent copy with $x^T_i(\theta_1) = x(\theta_1)$, still remaining close to x by time $\theta_2 T$. Formally, we define

$$P^i_{A(x,\epsilon)}(T,\theta_1,\theta_2) := P\left(|M^i_{A(x,\epsilon)}(T,\theta_1,\theta_2)| > 0\right)$$

These probabilities are identical for all i, and are equal to the probability of finding a particle started at 0, at an ϵ-neighbourhood of $y = \lambda t$ at time $(\theta_2 - \theta_1)T$. [Chauvin *et al.*,88] showed that the probability of a particle starting at 0 ascending to level λT at time T decays at the rate $1 - \frac{1}{2}\lambda^2$. We will need to modify her result slightly to prove that $\liminf_{T \to \infty} P^i_{A(x,\epsilon)}(T,\theta_1,\theta_2) \geq J(x,\theta_1,\theta_2)$. This will be done by a method analogous to the one used in lemma 4.

We think of $e^{J(x,\theta_1)T}$ copies, each performing an independent trial, with probability of success $P^i \approx e^{J(x,\theta_1,\theta_2)T}$. We see that since $J(x,\theta_1) + J(x,\theta_1,\theta_2) = J(x,\theta_2) > 0$, the expected number of particles succeeding, increases exponentially. Using an estimate on the Binomial distribution, we show that the probability that the growth rate is less than $J(x,\theta_2) - \delta$, decays exponentially for all $\delta > 0$. Finally, this result is true only *in probability.* To get an almost sure result, we have to use some sort of Borel-Cantelli Lemma. Basically, we show that if we had a particle inside $A(x,r)$ at time t, for some $r < \epsilon$. Then the particle was inside $A(x,\epsilon)$ for some interval before t. This allows us to divide time into countably many intervals, and use BCL.

We state and prove the three supporting lemmas.

Lemma 6. *Let $x \in C_1$ be a piecewise linear function. We claim that for every $\epsilon > 0$, there exists $r > 0$, such that, for all sufficiently large T, if $x^T \in A(x,r)$, then $x^\tau \in A(x,\epsilon)$ for all $\tau \in [T-1,T]$.*

Proof. We define the look-back transformation for all $\tau < T$:

$$L_\tau^T z(t) := \tfrac{T}{\tau} z(t\tfrac{\tau}{T}).$$

$L_\tau^T : A(x,r) \to A(L_\tau^T x, \frac{T}{\tau} r)$ and $\lim_{T\to\infty} \sup_{T-1<\tau<T} ||L_\tau^T x - x|| = 0$. We let $r = \frac{1}{4}\epsilon$. Pick T sufficiently large such that $\sup_{T-1<\tau<T} ||L_\tau^T y - y|| < r$ and $\frac{T}{T-1} < 2$. We deduce that for such T sufficiently large,

$$L_\tau^T A(x,r) \subseteq A(x,3r) \subset A(x,\epsilon) \quad \text{for all } \tau \in [T-1,T].$$

□

Result 6 (Right-Most Particle At The Subcritical Region - Chauvin). *Let $\lambda > \sqrt{2}$ and let R_T be the position of the right-most particle of a dyadic branching Brownian motion at time T. Then*

$$\liminf_{T\to\infty} T^{-1} \log P(R_T > \lambda T) = 1 - \tfrac{1}{2}\lambda^2$$

Corollary 3. *Let $y(t) = \lambda t$ where $\lambda > \sqrt{2}$. Then for all $r > 0$*

$$\liminf_{T\to\infty} T^{-1} \log P\left(|M_{A(y,r)}(T)| > 0\right) \geq 1 - \tfrac{1}{2}\lambda^2.$$

It follows that if $\dot{x} = \lambda$ on $[\theta_1, \theta_2]$, then for all $r > 0$

$$\liminf_{T\to\infty} T^{-1} \log P^i_{A(x,r)}(T, \theta_1, \theta_2) \geq J(x, \theta_1, \theta_2).$$

Proof. Define the closed set $D := \{z \in C_0 \setminus A(y,r) : z(1) \geq \lambda_2\}$. It is easy to show that $J(D) < 1 - \frac{1}{2}\lambda^2$ and hence

$$\limsup_{T\to\infty} T^{-1} \log P(|M_D(T)| > 0) < 1 - \tfrac{1}{2}\lambda^2.$$

Since $P(R_T \geq \lambda T) \leq P(|M_D(T)| > 0) + P(|M_{A(y,r)}(T)| > 0)$, the result follows. □

Finally, an estimate on the binomial distribution $\mathcal{B}(n,p)$.

Lemma 7. *Let $\alpha < 1$ and let $p < \frac{1}{2}$. Then, $\frac{1}{n} \log P\left(\mathcal{B}(n,p) < pn\alpha\right) < -\frac{p}{2(1-p)}(1-\alpha)^2$.*

Proof. This is a consequence of the Cramér's theorem for the Binomial distribution $\mathcal{B}(1,p)$ with rate function $b(x)$ which is a consequence Chebychev's inequality.

$$\frac{1}{n} \log P\left(\frac{1}{n}\mathcal{B}(n,p) \in [0, \alpha p]\right) \leq -b(\alpha p).$$

We calculate $b(x)$. If $X \sim \mathcal{B}(1,p)$ then $E(e^{\theta X}) = pe^\theta + q$. Hence $b(x) = \sup_\theta\{\theta x - \log(pe^\theta + q)\}$ and hence

$$b(x) = x \log \frac{x}{p} + (1-x) \log \frac{1-x}{1-p}.$$

Differentiating $b(x)$ we get $b'(x) = \log \frac{x(1-p)}{p(1-x)}$. Differentiating once more we get $b''(x) = \frac{1}{x(1-x)}$ which when $x < p < \frac{1}{2}$ satisfies $b''(x) \geq \frac{1}{p(1-p)}$. Integrating the inequality twice on $[x,p]$ using the boundary conditions $b(p) = b'(p) = 0$ we get $b(x) \geq \frac{1}{2p(1-p)}(x-p)^2$ and putting $x = \alpha p$ completes the proof.

□

Let us now state and prove the main result.

Lemma 8. *Let $x \in C_1$ be piecewise linear satisfying the lower bound condition until time θ_1. Let $\dot{x} = \lambda$ on $[\theta_1, \theta_2]$, with $\lambda > \sqrt{2}$, but with $J(x, \theta_2) > 0$. Then, for every $\epsilon > 0$,*

$$\liminf_{T \to \infty} T^{-1} \log |M_{A(x,\epsilon)}(T, \theta_2)| \geq J(x, \theta_2).$$

Proof. Pick $r < \epsilon$ as in lemma 6. At integer times $T_m := m$ define the following events:

$$U_m := \{\omega \in \Omega : |M_{A(x,\frac{1}{2}r)}(T_m, \theta_1)| < e^{(J(x,\theta_1)-\delta)T_m}\},$$
$$V_m := \{\omega \in \Omega \setminus U_m : |M_{A(x,r)}(T_m, \theta_2)| < e^{(J(x,\theta_2)-3\delta)T_m}\}.$$

Since x satisfies the lower bound condition until time θ_1, we know that almost surely, there exists $M(\omega)$ such that for all $m \geq M$, U_m does not occur. To work out the probability of V_m we use lemma 7 with the values $n \geq e^{(J(x,\theta_1)-\delta)T_m}$, $p \geq e^{(J(x,\theta_1,\theta_2)-\delta)T_m}$ and $\alpha = e^{-\delta T_m}$. We take n to represent the number of particles which stayed within $A(x, \frac{1}{2}r)$ up to time $\theta_1 T_m$. Since we are not in U_m we know that n is large (i.e. $n > e^{(J(x,\theta_1)-\delta)T}$). We take p to represent the probability for each of these particles that we could find a descendent in $A(x,r)$ by time $\theta_2 T_m$. This probability is decaying (and so $p < \frac{1}{2}$) yet is greater than $P^i_{A(x,\frac{1}{2}r)}(T, \theta_1, \theta_2)$ which was evaluated in corollary 3.
We deduce that $P(V_m)$ decays exponentially, and using BCL, V_m does not occur almost surely. Thus, almost surely,

$$\liminf_{m \to \infty} T_m^{-1} \log |M_{A(x,r)}(T_m)| \geq J(x, \theta_2) - 3\delta.$$

We now use lemma 6 to deduce that for all m and for all $\tau \in [T_{m-1}, T_m]$

$$\liminf_{\tau \to \infty} \tau^{-1} \log |M_{A(x,r)}(\tau)| \geq J(x, \theta_2) - 3\delta.$$

□

Corollary 4. *Let $x \in C_1$ be a piecewise-linear function such that $K(x) > 0$. Then, for every $\epsilon > 0$,*

$$\liminf_{T \to \infty} T^{-1} \log |M_{A(x,\epsilon)}(T)| \geq K(x).$$

Proof. Clearly, $\dot{x}(0) \leq \sqrt{2}$. Since $A(x, \epsilon)$ is open, there is no problem of finding a piecewise-linear function in $A(x, \epsilon)$ with $\dot{z}(0) < \sqrt{2}$. Now, proceed to glue together each linear segment of z using the previous lemmas. Please note that we avoided the case where $K(x) = 0$. □

We now have the lower bound result for the almost sure rate function. We ignored the case where $K(x) = 0$ because if A is any open set, and $x \in A$ satisfies $K(x) = 0$, then for some $0 < \alpha < 1$ we have $\alpha x \in A$ and $K(\alpha x) > 0$.

Theorem 9. *Let A be an open subset in C_0. Then, almost surely,*

$$\liminf_{T \to \infty} T^{-1} \log |M_A(T)| \geq \sup_{x \in A} K(x).$$

Proof. Since the piecewise-linear functions are dense in C_0 and $I(x)$ is lower-semicontinuous in the supremum topology, the result follows directly from the above corollary. □

4.4 Natural Extensions

In this section we consider possible extensions to theorem 9.

- We aim to extend the rate of breeding $r(x, t)$, (currently assumed to be constant) to a time and space dependent function. Let us count the number of particles along path $x \in C_1$ at time T. For every $s \leq 1$, the breeding rate at time sT is given by $r(sT, x(s)T)$. We observe that if $r(x, t) = \phi(x/t)$, then the observed breeding is T-independent.

- What is the correction term in theorem 9? More precisely, how does

$$e^{-nK(A)} |N_A(t)|$$

behave? This question corresponds to the question we asked about Cramér's theorem in chapter 2.

4.4.1 Position & Time Dependent Breeding

We state and sketch-prove a more general result. Consider a BBM with a birth process C satisfying $E(C) = 1 + \mu$. We also assume that $P(C = 0) = 0$ and $E(C \log C) < \infty$. Let each particle die at an exponential rate $r\,(B_i(t)/1 + t)$. The breeding rate $r \geq 0$ is assumed to be a continuous function. For every $x \in C_1$, the adjusted expectation rate function is defined as:

$$\bar{J}(x, \theta) := \int_0^\theta \mu r\left(\frac{x}{t}\right) - \tfrac{1}{2}\dot{x}^2 dt.$$

As before, let $\theta_0 := \inf\{\theta : \bar{J}(x, \theta) < 0\}$. Also let the almost sure rate function $\bar{K}$ be defined as

$$\bar{K}(x, \theta) := \begin{cases} \bar{J}(x, \theta) & \text{if } \theta \leq \theta_0, \\ -\infty & \text{otherwise.} \end{cases}$$

Theorem 10. *Let A, D be open and closed sets in C_0. Then*

$$\limsup_{T\to\infty} T^{-1} \log E(|M_D(T)|) \leq \bar{J}(D),$$
$$\liminf_{T\to\infty} T^{-1} \log E(|M_A(T)|) \geq \bar{J}(A).$$

Also, almost surely,

$$\limsup_{T\to\infty} T^{-1} \log |M_D(T)| \leq \bar{K}(D),$$
$$\liminf_{T\to\infty} T^{-1} \log |M_A(T)| \geq \bar{K}(A).$$

Proof: Almost-sure lower bound. We prove the lower bound for an open neighbourhood of a piecewise linear function x. Take $\mathcal{D}$, a partition of $[0,1]$. $\mathcal{D} := \{0 = t_0 < t_1 \ldots < t_n = 1\}$. Over the interval $[t_i, t_{i+1}]$, along the path $\{x(t) : t_i < t < t_{i+1}\}$ the process "observes" breeding at a rate greater or equal to $\inf\{r(\frac{x(t)}{t+1/T}) : t_i \leq t \leq t_{i+1}\}$. Thus, using the lower bound lemmas 4, 5 and 7, almost surely,

$$\liminf_{T\to\infty} T^{-1}\log|M_A(T)| \ge \mu\sum_{i<n}(t_{i+1}-t_i)\inf_{t_i\le t\le t_{i+1}} r(x(t)/t) - \int_0^1 \tfrac{1}{2}\dot{x}^2 dt.$$

Since x is piecewise continuous, by taking the supremum over all partitions we get the result. We cheated slightly, as we are only allowed to consider partitions which satisfy for all $j < n$,

$$\mu\sum_{i\le j}(t_{i+1}-t_i)\inf_{t_i\le t\le t_{i+1}} r(x(t)/t) > \int_0^{t_{j+1}} \tfrac{1}{2}\dot{x}^2 dt.$$

Since $\bar{K}(x,\theta) > 0$ for all θ, it can be shown that this constraint does not matter. □

4.4.2 The Correction Term in Theorem 9

This section is more about future possible research than actual results. We assume from now on that $E(|C|^3) < \infty$. Under this assumption a finer result than lemma 4 has been achieved by [Uchiyama, 82].

Result 7. *Let $|\lambda| < \sqrt{2}$. Then, almost surely*

$$\lim_{T\to\infty} \sqrt{T}e^{-(1-\frac{1}{2}\lambda^2)T}|N_{[\lambda T-a,\lambda T+b]}(T)| \to (2\pi)^{-\frac{1}{2}}\left\{\int_a^b e^{-\lambda x}dx\right\} \times Z_\lambda(\infty),$$

where $Z_\lambda(\infty)$ is a strictly positive random variable defined as the almost sure limit:

$$Z_\lambda(\infty) = \lim_{T\to\infty} e^{-T}\sum_{i\in N_T}\exp\{\lambda B_i(t) - \tfrac{1}{2}\lambda^2 T\}.$$

Guided by it, we make the following definition.

Definition. *For every $x \in C_1$, let $N_{x,\epsilon}(T)$ be the number of particles at time T, which are in the $\epsilon\sqrt{T\log T}$ neighbourhood of x.*

$$N_{x,\epsilon}(T) := \left\{ i \in N_T : |B_i(t) - Tx(t/T)| \leq \epsilon\sqrt{T \log T} \quad \textit{for all } t \leq T \right\}.$$

Theorem 11. *Let x be a piecewise linear function with $K(x) > 0$. Then almost surely*

$$\sqrt{T} e^{-TK(x)} |N_{x,\epsilon}(T)| \to +\infty.$$

The proof is identical in structure to the proof of theorem 9, with $M_{A(x,\epsilon)}$ replaced by $N_{x,\epsilon}$ in all the supporting lemmas. The most difficult part is the new proof of lemma 4 which we outline below. We assume that $x(t) = \sqrt{\lambda} t$ where $|\lambda| < \sqrt{2}$.

Proof: Assuming $x(t) = \lambda t$. Fix $k > 0$. We know that almost surely

$$\sqrt{T} |N_{[\lambda T - k, \lambda T + k]}(T)| e^{-TK(x)} \to (2\pi)^{-\frac{1}{2}} \left\{ \int_{-k}^{k} e^{-\lambda x} dx \right\} \times Z_\lambda(\infty).$$

.Consider now the set S_T defined as

$$S_T := N_{[\lambda T - k, \lambda T + k]}(T) \setminus N_{x,\epsilon}(T).$$

This set S_T, contains particles who at time T are inside $[\lambda T - k, \lambda T + k]$ but for some $s < T$ have $|B_i(s) - Tx(s/T)| > \epsilon\sqrt{T \log T}$. Consider a single particle which is known to be inside $[\lambda T - k, \lambda T + k]$ at time T. Conditioned upon its position, it behaves like a Brownian motion with a drift of at most $\lambda + \frac{K}{T}$ and at least $\lambda - \frac{K}{T}$. Subtracting these drifts and using the reflection principle we can see that

$$P\left(\sup_{s \leq T} |B_i(s) - Tx(s/T)| > \epsilon\sqrt{T \log T} | B_T \in [\lambda T - k, \lambda T + k] \right) \leq c_\epsilon \frac{1}{T}.$$

We deduce that for some constant $d(\epsilon, k)$

$$\sqrt{T} e^{-TK(x)} E(|S_T|) \leq d(\epsilon, k) \frac{1}{T} E(Z_\lambda(\infty)).$$

Looking at integer times and then applying a look-back argument, we see that almost surely

$$\sqrt{T}|S_T|e^{-TK(x)} \to 0,$$

and thus almost surely

$$\sqrt{T}e^{-TK(x)}|N_{x,\epsilon}(T) \cap N_{[\lambda T-k,\lambda T+k]}(T)| \to (2\pi)^{-\frac{1}{2}}\{\int_{-k}^{k} e^{-\lambda x}dx\} \times Z_\lambda(\infty).$$

Finally, letting $k \uparrow \infty$ we get the result. □

Conjecture 1. *We conjecture that for all $x \in C_1$, if $K(x) > 0$, then almost surely*

$$e^{-TK(x)}|N_{x,\epsilon}(T)| \to \textit{constant} \times Z_{\dot{x}(0)}(\infty).$$

We are still a long way from proving this conjecture.

4.5 A Final Note

Recently, we discovered a book by [Revesz, 94]. He considered a split-at-integer-times branching Brownian motion and showed the space of paths to be the closure of $\{f : K(f) \geq 0\}$ without counting the actual growth rate along each path. His result is similar in nature although the methods he used are different.

Chapter 5

The Phase Plane of an Integrated Branching Brownian Motion.

5.1 An Overview

As in the previous chapter we will concentrate on a dyadic branching Brownian motion. We will present an application of theorem 9 and result 3. These two results are concerned with the rate of growth of BBM particles along (scaled) paths. Our application is based on the simple observation, that we can project the path of a Brownian Motion, to the path of another diffusion process. We utilise this observation to study the following point process on $\mathcal{R}^2$.

We take a BBM and assumes that for each particle, $B_i(t)$ represents the the particle's *velocity.* Its position can then be obtained by integrating $\{B_{a_i(s)}(s)\}_{s\leq t}$.

$$Y_i(t) = \int_0^t B_{a_i(s)}(s)ds,$$

where $a_i(s)$ denotes the ith particle's unique ancestor at time s. We arrive at a two dimensional point-process on the plane $(B_i(t), Y_i(t))$. The process $Y(t)$ is an integral with respect to a continuous path and is therefore a differentiable finite-variation process. It is also Gaussian and its variance is given by

$$2E\left(\int_0^t B_r dr \int_r^t B_s ds\right) = 2\int_0^t \int_r^t r ds dr = \frac{1}{3}t^3$$

We will analyse the phase plane picture using theorem 9 and result 3. Before that though, let us demonstrate using current methods that there *must* be a difference between the behaviour in expectation and almost surely of $(B_i(t), Y_i(t))$. We do so by considering the wavefront speeds.

5.2 The Wavefront Speeds of $Y_i(t)$

5.2.1 The Expectation Wavefront

If we let $N_{[x,\infty)}(t) = \{i \in N_t : Y_i(t) \geq x\}$ by using the many-to-one picture (equation 3.4) we get

$$E(|N_{[u_t,\infty)}(t)|) = E(|N_t|)P(Y_t > u_t) \approx e^t \exp(-\frac{3}{2t^3}u_t^2).$$

We deduce that the expectation wavefront travels at the speed $t^2\sqrt{2/3}$ in the sense that

$$\lim_{t\to\infty} \frac{u_t}{t^2} = \sqrt{2/3}. \tag{5.1}$$

5.2.2 The Almost-Sure Wavefront

[Neveu, 87] observed that almost surely R_t, the rightmost particle of a branching Brownian Motion satisfies $\limsup_{t\to\infty} R_t - t\sqrt{2} = -\infty$. Because $\sup_{i\in N_t} Y_i(t) \leq \int_0^t R_s ds$ by integrating the bound on R_t we get an instant upper bound on v_t, the almost-sure wavefront speed.

$$\limsup_{t\to\infty} \frac{v_t}{t^2} \leq 1/\sqrt{2}. \tag{5.2}$$

Before proving that equality holds in equation 5.2 we want to point out that the almost-sure wavefront speed is below the expectation wavefront speed already in the first order of magnitude! A more comprehensive explanation of this phenomenon will be offered when we study the phase plane.

Theorem 12. *Let v_t denote the rightmost particle's position of a branching integrated Brownian motion. Then, almost surely,*

$$\lim_{t\to\infty} \frac{v_t}{t^2} = 1/\sqrt{2}.$$

Proof (Lower Bound). We look at the branching Brownian motion. We follow [Neveu, 87] and define Z_s^λ to be the number of particles which first among their ancestors crossed the line $x = s - \lambda t$.

$$Z_s^\lambda = |\{i \in I : \exists t \in [T_i, T_i + \tau_i] \quad B_i(t) > s - \lambda t, \quad \forall t < T_i \quad B_{a(i)}(t) < s - \lambda t\}|.$$

Result 8 (Neveu, Proposition 3). *For each $\lambda \geq \sqrt{2}$ the integer valued process $(Z_s^\lambda, s > 0)$ is a Galton-Walton process without extinction whose infinitesimal generating function a is given by*

$$a = \psi' \circ \psi^{-1} \quad on\ (0,1)$$

where $\psi : \mathcal{R} \to (0,1)$ is the solution of Kolmogorov's equation

$$\tfrac{1}{2}\psi'' - \lambda\psi' = \psi - \psi^2. \tag{5.3}$$

The result is based on the Brownian path decomposition by [Williams, 74]. We now consider what happens if $0 < \lambda < \sqrt{2}$. Z_s^λ can still be defined as a birth-death process. Since a Brownian Motion almost surely hits the downward sloping line $x(t) = s - \lambda t$ we see that Z_s^λ is without extinction. Reproducing [Neveu, 87]'s proof we arrive at Kolmogorov's equation. From differential equations theory we know that Kolmogorov's equation 5.3 does not have a monotone solution on $(0,1)$. Looking at the definition of the infinitesimal generator function (equation 3.3), this implies that a possesses a discontinuity at 1 which means $a_\infty > 0$. Thus the process explodes almost surely. We let $T(\omega)$ denote the explosion time. Spatially this corresponds to there being, at all times, a particle below the line $x = T(\omega) - \lambda t$ all of whose ancestors have also been below that line. (If after some time τ, there is no such particle, then $Z_T^\lambda \le N_\tau < \infty$). Integrating the Brownian Path of this particle and its unique ancestors we deduce that almost surely $v_t(\omega) > \frac{1}{2}\lambda t^2 - T(\omega)t$ and hence, almost surely,

$$\liminf_{t\to\infty} \frac{v_t}{t^2} \ge \lambda/2.$$

We now let $\lambda \uparrow \sqrt{2}$ to complete the proof. □

5.3 The Phase Plane Picture

We use the projection from the BBM to the Branching point process, to project the space of paths of BBM to the phase plane. We deduce a large-deviations principle for the phase plane, both in expectation and almost surely. We find the two rate functions to be different, and the difference explains the different wavefront speeds we observed earlier.

5.3.1 Scaling The Process

As before, at a fixed time T, let us scale the branching Brownian Motion by a factor of T in both the space and time coordinates. We get a branching process on $[0,1]$. For every $i \in N_T$ let $x_i^T \in \mathcal{C}^0([0,1],\mathcal{R})$ and $y_i^T \in \mathcal{C}^0([0,1],\mathcal{R})$ be defined as

$$\begin{aligned} x_i^T(t) &:= \frac{1}{T} B_{a(i)}(tT), \\ y_i^T(t) &:= \int_0^t x_i^T(s)ds = \frac{1}{T^2} Y_{a(i)}(tT). \end{aligned}$$

We define the projection map $\Pi : \mathcal{C}_0 \to \mathcal{R}^2$ as

$$\Pi(z) := (z(1), \int_0^1 z(t)dt),$$

which is clearly continuous in the supremum norm. For every $D \subset \mathcal{C}_0$ we let $M_D(T)$ denote the particles at time T whose path is in D. For every $D \subset \mathcal{R}^2$ we define let $\Pi M_D(T)$ to be $M_{\Pi^{-1}D}(T)$.

$$M_D(T) := \{i \in N_T : z_i^T \in D\},$$
$$\Pi M_D(T) := \{i \in N_T : \Pi z_i^T \in D\}.$$

We must apologise to the reader for the slight change of notations which is about to occur. From now on, we will use $z \in C_0$ to denote a path of a BBM. x and y will now represent the coordinates in the phase plane.

5.3.2 The Expectation Picture

The Expectation large-deviations result tells us that if $D \subset \mathcal{R}^2$ is closed and $A \subset \mathcal{R}^2$ is open then

$$\limsup_{T\to\infty} \tfrac{1}{T} \log E|\Pi M_D(T)| \leq \sup_{z\in\Pi^{-1}D} J(z) = \sup_{(x,y)\in D} \quad \sup_{\Pi z=(x,y)} J(z),$$
$$\liminf_{T\to\infty} \tfrac{1}{T} \log E|\Pi M_A(T)| \geq \sup_{z\in\Pi^{-1}A} J(z) = \sup_{(x,y)\in A} \quad \sup_{\Pi z=(x,y)} J(z).$$

We recall that $J(z) = J(z,1)$ is the expectation rate function for a branching Brownian motion:

$$J(z,\theta) := \theta - \tfrac{1}{2}\int_0^\theta \dot{z}^2(t)dt.$$

For every $(x,y) \in \mathcal{R}^2$ we define $\Pi J(x,y) := \sup\{J(z) : \Pi z = (x,y)\}$ and use calculus of variation with Lagrange multiplier optimising procedure (see section 5.4.1) to find that there is a unique z maximising $\Pi J(x,y)$:

$$z(t) = 3(x-2y)t^2 + 2(3y-x)t.$$

Accordingly, $\Pi J(x,y) = J(z) = 1 - \frac{1}{2}x^2 - 6(y - \frac{1}{2}x)^2$. We immediately have the following result.

Result 9. *Let $D \subset \mathcal{R}^2$ be closed and let $A \subset \mathcal{R}^2$ be open. Then,*

$$\limsup_{T\to\infty} \tfrac{1}{T}\log E|\Pi M_D(T)| \leq \sup_{(x,y)\in D} \Pi J(x,y),$$
$$\liminf_{T\to\infty} \tfrac{1}{T}\log E|\Pi M_A(T)| \geq \sup_{(x,y)\in A} \Pi J(x,y).$$

Before carrying on please take time to consider how natural this formula is. For particles whose Brownian position at time T is xT we know that their rate function is given by $1-\frac{1}{2}x^2$. Conditioning on their final position we know that the particles have the law of a Brownian Motion with drift x so that most of these particles arrive in a straight line $x_i^T(t) = xt$ yielding $y_i^T(1) = \frac{1}{2}x$. Some of them will deviate from that path and are penalised by the amount $6(y-\frac{1}{2}x)^2$.

Corollary 5. *Let u_t be the expectation wavefront speed. Then*

$$\lim_{t\to\infty} \frac{u_t}{t^2} = \sqrt{2/3}.$$

Proof. The boundary of the region $\{(x,y) : \Pi J(x,y) \geq 0\}$ defines the expectation wavefront. In particular, maximising y subject to $\Pi J(x,y) \geq 0$, we find $x = \sqrt{3/2}$ and $y = \sqrt{2/3}$. Since $y = \frac{1}{T^2}Y_i(T)$ the result follows.

□

5.3.3 The Almost-Sure Picture

We know from the large-deviations contraction principle that almost surely

$$\limsup_{T\to\infty} \tfrac{1}{T}\log|\Pi M_D(T)| \leq \sup_{z\in\Pi^{-1}D} K(z) = \sup_{(x,y)\in D} \sup_{\Pi z=(x,y)} K(z),$$
$$\liminf_{T\to\infty} \tfrac{1}{T}\log|\Pi M_A(T)| \geq \sup_{z\in\Pi^{-1}A} K(z) = \sup_{(x,y)\in A} \sup_{\Pi z=(x,y)} K(z).$$

Here $K(z) = K(z,1)$, is the almost-sure rate function for a branching Brownian motion. If $\theta_z := \inf\{\theta \in [0,1] : J(z,\theta) < 0\}$ then K is defined as

$$K(z,\theta) := \begin{cases} J(z,\theta) & 0 < \theta \leq \theta_z \\ -\infty & \theta_z \leq \theta \leq 1 \end{cases}$$

Finding $\Pi K(x,y) := \sup\{K(z) : \Pi z = (x,y)\}$ is more involved as in addition to $\Pi z = (x,y)$, we also impose that $J(z,\theta) \geq 0$ for all $\theta \in [0,1]$, but see section 5.4.2 for a solution. We find that for the half-plane $\{y \geq \frac{1}{2}x\}$ the following holds:

Let $\alpha, \beta, \gamma \in \mathcal{R}^2$ be defined as

$$\alpha = (\sqrt{2}, 1/\sqrt{2}), \quad \beta = (-1/\sqrt{2}, 0), \quad \gamma = (-\sqrt{2}, -1/\sqrt{2}).$$

Let f, g, h be the functions defined as

$$f(x) = \tfrac{1}{3}x + \tfrac{1}{3\sqrt{2}}, \quad g(x) = \tfrac{1}{2}x + \tfrac{1}{6}\sqrt{6 - 3x^2}, \quad h(x) = \tfrac{1}{2}x.$$

Note that f links α to β, the function g links β to γ while clearly h links γ and α. So that they form a region D_1 (also see diagram)

$$D_1 = \left\{(x,y) : x \in [-\sqrt{2}, \sqrt{2}], y \in [h(x), f(x) \wedge g(x)]\right\}.$$

In addition let $l(x) = \frac{1}{\sqrt{2}} - \frac{\sqrt{2}}{9}(\sqrt{2} - x)^2$ be another function linking α and β and let D_2 be the region enclosed by $l(x)$ and $f(x)$.

$$D_2 = \left\{(x,y) : x \in [-1/\sqrt{2}, \sqrt{2}], y \in [f(x), l(x)]\right\}.$$

We find that in D_1 (and by symmetry in $-D_1$ too) the almost-sure behaviour and the behaviour in expectation agree so that the z which maximises the expectation satisfies $K(z) = J(z)$ and so $\Pi K = \Pi J$. Inside D_2 we find that the z maximising is given by

$$\dot{z}(t) = \begin{cases} \sqrt{2} & 0 \le t \le \theta \\ \sqrt{2} - \mu(t-\theta) & \theta \le t \le 1 \end{cases}$$

with

$$\theta = 1 - 3(\tfrac{1}{\sqrt{2}} - y)/(\sqrt{2} - x),$$
$$\mu = \frac{2}{9}(\sqrt{2} - x)^3/(\tfrac{1}{\sqrt{2}} - y)^2.$$

We conclude that for $(x, y) \in D_2$,

$$\Pi K(x,y) = \sqrt{2}(\sqrt{2} - x)\left(1 - \frac{2}{9}(\sqrt{2} - x)^2/(1 - \sqrt{2}y)\right).$$

Otherwise we find $\Pi K(x,y) = -\infty$. We now do the same analysis for the other half plane $\{y \le \frac{1}{2}x\}$ and get the almost-sure result.

Result 10. *Let $D \subset \mathcal{R}^2$ be closed and let $A \subset \mathcal{R}^2$ be open. Then, almost surely*

$$\limsup_{T\to\infty} \tfrac{1}{T} \log |\Pi M_D(T)| \le \sup_{(x,y)\in D} \Pi K(x,y),$$
$$\liminf_{T\to\infty} \tfrac{1}{T} \log |\Pi M_A(T)| \ge \sup_{(x,y)\in A} \Pi K(x,y).$$

Corollary 6. *Let v_t be the almost-sure wavefront speed, then*

$$\lim_{t\to\infty} \frac{v_t}{t^2} = 1/\sqrt{2}.$$

Proof. The boundary of the region $D_1 \cup D_2 = \{(x,y) : \Pi K(x,y) \ge 0\}$ defines the almost-sure wavefront. In particular, maximising y subject to $\Pi K(x,y) \ge 0$, we find that $\frac{dl}{dx}|_{x=\sqrt{2}} = 0$ and the supremum is attained at $x = \sqrt{2}$ and $y = 1/\sqrt{2}$. Since $y = \frac{1}{T^2} Y_i(T)$ the result follows.

□

5.4 Optimisation of The Rate Functions

In this section we explain briefly how the optimisations for $\Pi J(x,y)$ and $\Pi K(x,y)$ were derived.

5.4.1 The Expectation Rate Function

If $(x,y) \in \mathcal{R}^2$ we wish to maximise $\{J(z) : \Pi(z) = (x,y)\}$. Alternatively, we minimise $\frac{1}{2}\int_0^1 \dot{z}^2$ subject to the constraint $z(0) = 0, z(1) = x, \int_0^1 z(t)dt = y$. We get using Lagrange multiplier the unconstrained problem of minimising F.

$$F(z,\dot{z},t) = \int_0^1 \tfrac{1}{2}\dot{z}^2 - \lambda(y-z)dt.$$

From calculus of variations we have $F_z - \dot{F_{\dot{z}}} = 0$ from which we get that $\ddot{z} =$ constant and so $z(t) = (x-\alpha)t^2 + \alpha t$. Substituting $\int_0^1 z dt = y$ we arrive at the optimal path in expectations:

$$z = 3(x-2y)t^2 + 2(3y-x)t,$$

from which we deduce that

$$J(z) = 1 - \tfrac{1}{2}\int_0^1 \dot{z}^2 dt = 1 - \tfrac{1}{2}x^2 - 6(y - \tfrac{1}{2}x)^2.$$

5.4.2 The Almost-Sure Rate Function

Throughout, we assume that $y \geq \frac{1}{2}x$. When $y < \frac{1}{2}x$ we use the symmetry $\Pi K(-x,-y) = \Pi K(x,y)$. Clearly if the z which optimises $\{J(z) : \Pi(z) = (x,y)\}$ also has $K(z) = J(z)$ we are done. This amounts to ensuring $\dot{z}(0) \leq \sqrt{2}$ and we find that if $(x,y) \in D_1$, this is indeed the case.

Outside D_1, although we can not follow the same optimising procedure, some points are clear. Keeping x fixed, as y increases ΠJ and ΠK are decreasing. To maximise y while keeping J constant, we must have $\dot{z}$ as a non increasing function. Conditioning on the first time θ when $\dot{z} < \sqrt{2}$ we find that on $[\theta, 1]$, the in-expectation optimisation must also be valid almost surely so that z must be of the form

$$\dot{z}(t) = \begin{cases} \sqrt{2} & 0 \leq t \leq \theta \\ \sqrt{2} - \mu(t - \theta) & \theta \leq t \leq 1 \end{cases}$$

which we integrate to get

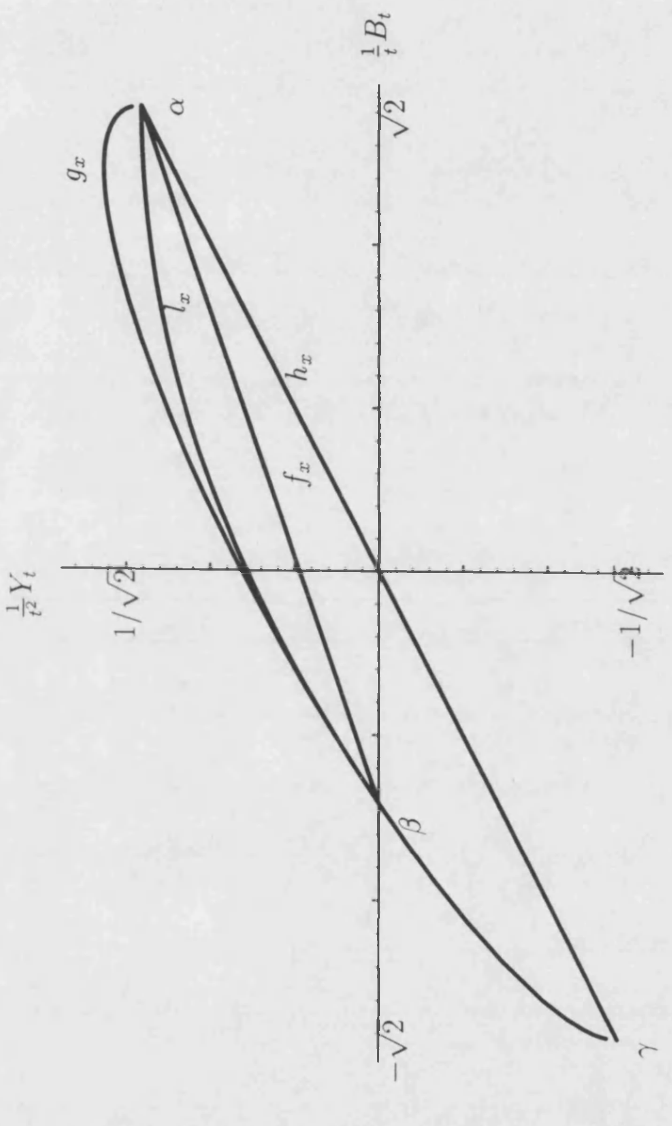
$$z(t) = \begin{cases} \sqrt{2}t & 0 \leq t \leq \theta \\ \sqrt{2}t - \frac{1}{2}\mu(t - \theta)^2 & \theta \leq t \leq 1 \end{cases}$$

and finally we deduce that

$$\int_0^1 z(t)dt = \tfrac{1}{\sqrt{2}} - \tfrac{1}{6}\mu(1 - \theta)^3.$$

We substitute boundary conditions $z(1) = x$ and $\int_0^1 z(t)dt = y$ to complete the analysis.

5.4.3 The Phase Plane Diagram



Chapter 6

Brownian Motion with Drift

6.1 Introduction

In this chapter we study a BBM with drift μ which attracts the particles towards the origin. This is a branching process whose diffusion process satisfies

$$dU_t = dB_t - \mu(U_t)dt.$$

We call this a μ-BBM which we set up as follows. We consider a BBM probability space (see section 3.1.2). The only assumption we make about the birth process is that $E(C) < \infty$ and that by rescaling time, WLOG we assume that $E(|N_t|) = e^t$. We associate with this probability space, a family of diffusion process $U_i(t)$ defined by

$$dU_i(t) = dB_i(t) - \mu(U_i(t))dt,$$

under the boundary conditions that the initial particle starts at the origin and subsequent particles inherit their ancestor's position at birth. Formally, for all $i \in I$ and for all $c \in \mathcal{N}$,

$$U_{ic}(T_{ic}) = U_i(T_i + \tau_i).$$

This chapter contains three distinct sections. In the first (rather tedious) section we study the diffusion process without any breeding. We derive probabilistic and path-wise estimates

on the rate at which the diffusion spreads in space. Under the assumptions that the particle is being"pulled" increasingly to the origin by the drift, we will be able to get some very reasonable estimates.
In the second section we prove that the almost-sure and expectation wavefronts first order behaviour are the same. This is quickly extended to an almost-sure large-deviations result about the number of particles in space. The result is very similar to the BBM result but is less informative in the sense that we count particles not along paths but simply according to their final position. Essentially, the diffusion process "forgets" its past so characterising the entire path is meaningless. We emphasise this point by proving that the diffusion processes and the Brownian motion driving it are essentially independent for large t. From a large-deviations perspective, the phase-space (U_t, B_t) satisfies a large-deviations principle which looks like a large-deviations principle associated with the product space of two independent processes.
In the third section we study the integral process $dV_t = \mu(U_t)dt$. This can be thought of as the "energy" (force applied over time) the particle expends. It will not be surprising therefore that the process $dV_t = \mu(U_t)dt$ and dB_t will turn out to share similar large deviations behaviour. More generally, we will study the BDP whose diffusion process satisfies

$$dV_t = f(U_t)dt$$

for some suitable f.

6.2 The Diffusion Process

6.2.1 An Example

The example we have in mind is the Ornstein-Uhlenbeck process, which for the purposes of this work has drift $\mu(x) = \frac{1}{2}x$ and stochastic equation:

$$dU_t = dB_t - \tfrac{1}{2}U_t dt.$$

The Ornstein-Uhlenbeck process has an invariant distribution $N(0,1)$ and an explicit solution:

$$U_t = e^{-\frac{1}{2}t}\left(U_0 + \int_0^t e^{\frac{1}{2}s} dB_s\right). \tag{6.1}$$

To make our analysis work, we will impose an increasing amount of assumptions on the process U_t. All these assumptions can be demonstrated to hold for the Ornstein-Uhlenbeck process. To start with, we restrict μ to be a differentiable, increasing odd function.

$$\mu'(x) = \mu'(-x) \geq 0. \tag{6.2}$$

Further, we impose that

$$\lim_{x\to\infty} \mu(x) = \infty. \tag{6.3}$$

The Ornstein-Uhlenbeck process with $\mu(x) = \frac{1}{2}x$ satisfies all these restrictions. More generally, $\mu_\alpha(x) = \frac{1}{2}\text{sign}(x)|x|^\alpha$ satisfies these restrictions when $\alpha \geq 0$.

Note

The reason why we assume that $\lim_{x\to\infty} \mu(x) = \infty$ is that if $\mu := \lim_{x\to\infty} \mu(x)$ is finite, then essentially we have a Brownian motion with constant drift. We can project the paths of an ordinary BBM to study the paths of this process. The optimal path for the expectation/almost-sure wavefronts to travel is given by $u(t) = 0$ on $[0, \theta]$ and $u(t) = \lambda(t - \theta)$ on $[\theta, 1]$ where we maximise $u(1) = \lambda(1 - \theta)$ subject to $\frac{1}{2}(\lambda + \mu)^2(1 - \theta) = 1$ and $0 \leq \theta \leq 1$. The optimal value for λ is $\max(\mu, \sqrt{2} - \mu)$, for large μ we have the expectation and almost sure wavefronts travel at speed $\frac{1}{2\mu}t$. The problem is thus solved.

6.2.2 The Invariant Distribution

By considering the stationary solution of the differential equation $P(U_t > x)$ satisfies, one can show that U_t tends to an invariant distribution U whose probability density function is given by

$$\frac{1}{2}\frac{d}{dx}\log\phi(x) = -\mu(x). \tag{6.4}$$

This can be solved with the constant of integration chosen to ensure $\int \phi(x)dx = 1$. For example, if $\mu(x) = \frac{1}{2}x$, we have the Ornstein-Uhlenbeck process whose invariant distribution is given by $N(0,1)$.

6.2.3 The Expectation Wavefront

Starting at the origin with $U_0 = 0$, we define the expectation wavefront speed $r(t)$ to be the solution to

$$P(U_t^0 \geq r_t) = e^{-t}. \tag{6.5}$$

We use the notation U_t^x to denote the process at time t started at $U_0 = x$.

Lemma 9. *The function r_t is increasing in t but satisfies $\lim_{t\to\infty} \frac{r_t}{t} = 0$.*

Proof. Consider the reflected process $V_t = |U_t|$. Since μ is antisymmetric, this is a Markov process. Moreover, by symmetry, $P(U_t^0 > r_t) = \frac{1}{2}P(V_t^0 > r_t)$. Conditioning on ϕ_t, the distribution of V_t^0, we have

$$\begin{aligned}
P(V_{t+s}^0 > r_s) &= \int_0^\infty \phi_t(x)P(V_s^x > r_s)dx \\
&\geq \int_0^\infty \phi_t(x)P(V_s^0 > r_s)dx \\
&= P(V_s^0 > r_s) = 2e^{-s} \\
&> 2e^{-s+t} \\
&= P(V_{t+s}^0 > r_{t+s}).
\end{aligned}$$

The first inequality follows from a coupling argument of two processes, one started at 0 and the other at x. Because V_{t+s}^0 is just a random variable have that $r_s < r_{t+s}$ and r_t is increasing.

Now, V_t^0 is in itself stochastically dominated by the process $c + W_t$. Here W_t is a reflected

Brownian motion with a constant drift $\mu(c)$ towards the origin. Recall that for a fixed drift, the problem is essentially solved (see note above). We have that for every $\epsilon > 0$ there exists a level c such that $\mu(c) > \frac{2}{\epsilon}$ is sufficiently large to ensure that

$$\lim_{t\to\infty} \frac{1}{t} \log P(V_t^0 > \epsilon t) < \lim_{t\to\infty} \frac{1}{t} \log P(c + W_t > \epsilon t) < -2.$$

We deduce that $\lim_{t\to} \frac{r_t}{t} < \epsilon$. Letting $\epsilon \downarrow 0$ completes the proof.

□

Since $U_t \to U$, the invariant distribution, we have that $\lim_{t\to\infty} r_t = \infty$. We now get a tighter upper bound on the rate at which r_t is increasing. We find the invariant distribution using equation 6.4. Let U_t^ϕ be the process started at the invariant distribution. By conditioning on U_0, we have

$$P(U_t^\phi > r_t) \geq \int_0^\infty \phi(x) P(U_t^x > r_t) dx \geq \int_0^\infty \phi(x) P(U_t^0 > r_t) dx = P(U > 0) P(U_t^0 > r_t) = \tfrac{1}{2} e^{-t}.$$

Hence

$$\lim_{t\to\infty} \frac{1}{t} \log P(U > r_t) \geq -1.$$

For example, let $\mu_\alpha(x) = \frac{1}{2}\mathrm{sign}(x)|x|^\alpha$. By equation 6.4 the invariant distribution ϕ_α satisfies

$$\log \phi_\alpha(x) \sim -\frac{1}{1+\alpha}|x|^{1+\alpha}.$$

After some calculations we get that for some constant c,

$$r_t \sim c t^{\frac{1}{1+\alpha}}.$$

On the other hand, the diffusion process is certainly positive recurrent and converges to the origin exponentially fast as the following results show. Recall that U_t^x denotes the process

U started at x. The following gives us a comparison of U_t^x with the deterministic path u_t^x which satisfies $\dot{u}^x = -\mu(u^x)$ and $u_0^x = x$.

Lemma 10. *Let U_t^x and u_t^x be as above. If B_t is the Brownian motion driving U_t, with minimum m_t and maximum M_t denote W_t^+ as the positive reflected Brownian motion $W_t^+ = B_t - m_t$. Similarly, let $W_t^- = B_t - M_t$ be the negative reflected Brownian motion. Finally, let $U_t^\pm$ be defined as:*

$$\begin{aligned} U^+(t) &= u_t^x + W_t^+, \\ U^-(t) &= u_t^x + W_t^-. \end{aligned}$$

Then, for all $t \leq T^x$ we have path-wise

$$U_t^- \leq U_t^x \leq U_t^+.$$

Proof. Stochasticly differentiating U_t^+ we get

$$dU^+(t) = dW_t^+ - \mu(u_t^x)dt.$$

Since $W_t^+ = B_t - m_t \geq 0$, we have $U_t^+ \geq u_t^x$. Therefore, using the monotonicity of μ, we have

$$dU_t^+ \geq dB_t - \mu(U_t^+)dt.$$

It follows that whenever $U_t^x = U_t^+$, we have $dU^x \leq dU^+$, and the upper bound follows. A similar argument yields the lower bound.

□

Suppose now we start the particle at $x = r_T$. While $u_t > \epsilon r_T$ the process is stochastically dominated by $r_T + B_t^\epsilon$. Here B^ϵ is a Brownian motion with drift $-\mu(\epsilon r_T)$. We get even after a short period δT

$$P\left(B_{\delta T}^{\epsilon} > -\tfrac{1}{2}\mu(\epsilon r_T)\delta T\right) \leq e^{-\frac{1}{8}\mu(\epsilon r_T)^2\delta T}.$$

Since the diffusion process satisfies $\frac{r_t}{t} \to 0$, the drift of B^{ϵ} dominates the initial position and we get the following lemma.

Lemma 11. *For every $\epsilon > 0$, define $\tau = \inf\{t : U_t^{r_T} \leq \epsilon r_T\}$ to be the first time the process U^{r_T} hits level ϵr_T. Then, for all $\delta > 0$ and for all $c > 0$,*

$$\lim_{T\to\infty} \frac{1}{T} \log P(\tau > \delta T) \leq -c.$$

Proof. If $\tau > \delta T$, then

$$P\left\{U_{\delta T} > r_T - \tfrac{1}{2}\mu(\epsilon r_T)\delta T\right\} \leq e^{-\frac{1}{8}\mu(\epsilon r_T)^2\delta T}.$$

Pick T large such that $\frac{1}{8}\mu(\epsilon r_T)^2\delta > c$ and $r_T - \frac{1}{2}\mu(\epsilon r_T)\delta T < 0$ (this can be done because $\lim_{T\to\infty} r_T = \infty$ and $r_T/T \to 0$). The proof follows because

$$\begin{aligned} P(\tau > \delta T) &= P\{\tau > \delta T, U_{\delta T} > 0\} \\ &\leq P\left\{U_{\delta T} > r_T - \tfrac{1}{2}\mu(\epsilon r_T)\delta T\right\} \\ &\leq e^{-\frac{1}{8}\mu(\epsilon r_T)^2\delta T}. \end{aligned}$$

□

Effectively the above lemma says that the diffusion process will *not* be below level ϵr_T at some point during the interval $[(1-\delta)T, T]$ with a very small probability indeed. Hence we get the following corollary.

Corollary 7. *For every $\delta > 0$,*

$$\lim_{T\to\infty} \frac{1}{T} \log P(U_{\delta T}^0 > r_T) = -1.$$

Similarily, supoose we start the process at the invariant distribution. Then, by conditioning on the initial position of the particle (either bigger than $r_{(1+\epsilon)T}$ or smalller) we get the following corollary.

Corollary 8. *If $f(\theta) = \lim_{t\to\infty} \frac{r_{\theta t}}{r_t}$ is a continuous function at* 1*. Then,*

$$\lim_{T\to\infty} \frac{1}{T} \log P(U_T^{\phi} > r_T) = 1.$$

6.3 Large Deviations Theorems

Theorem 13. *Consider a μ-BBM which satisfies*

- *The drift is differentiable with $\mu'(x) = \mu'(-x) > 0$ and $\mu(0) = 0$.*
- $\lim_{x\to\infty} \mu(x) = \infty$
- *The increasing function r_t satisfying $P(U_t^0 > r_t) = e^{-t}$ satisfies*

$$f(\theta) := \lim_{t\to\infty} \frac{r_{\theta t}}{r_t}$$

exists and is continuous in $[0, 1]$

Further, we assume that for every $\epsilon > 0$ there exists a sequence of increasing stopping times $T_n \uparrow \infty$ which satisfy

- *There exists a large N such that for all $n > N$*

$$T_{n+1} - T_n \le \epsilon \frac{r_{T_n}}{\mu(r_{T_n})}.$$

- *T_n grow sufficiently fast such that for all $\delta > 0$*

$$\sum_n e^{-\delta T_n} < \infty.$$

- *T_n grow not too fast in the sense that for all $\delta > 0$ for large n*

$$r_{T_{n+1}} - r_{T_n} < \delta r_{T_n}.$$

Then, almost surely,

$$\limsup_{t\to\infty} \frac{R_t}{r_t} \leq 1. \quad \textit{[Upper Bound]}$$
$$\liminf_{t\to\infty} \frac{R_t}{r_t} \geq 1. \quad \textit{[Lower Bound]}$$

Before we try to prove the theorem let us satisfy ourselves that the set of diffusions satisfying the above conditions is not empty. The drift $\mu(x) = \frac{1}{2}\mathrm{sign}(x)|x|^\alpha$ satisfies the first two conditions imposed on the drift. We already saw that $r_t \sim t^{\frac{1}{1+\alpha}}$. This implies that $f(\theta) = r_{\theta t}/r_t$ is continuous and tends to 1 as $\theta \to 1$. Further, consider the stopping times $T_n = n^\beta$ which are increasing as long as $\beta > 0$. Certainly for all $\delta > 0$ we have

$$\sum_n e^{-\delta n^\beta} < \infty.$$

We observe that

$$\epsilon \frac{r_{T_n}}{\mu(r_{T_n})} \sim n^{\frac{\beta}{1+\alpha} - \frac{\alpha\beta}{1+\alpha}}.$$

Since $T_{n+1} - T_n \sim n^{\beta-1}$, by picking β small enough to satisfy

$$\beta - 1 \leq \beta \frac{1-\alpha}{1+\alpha},$$

the fourth condition (in the second part of the theorem) is satisfied as well.

6.3.1 A Failed Approach

We owe the reader an explanation why we do not attempt to prove this theorem using Theorem 9. When $\lim_{x\to\infty}\mu(x)$ is finite, we can do this. But if $\mu(x)$ increases to ∞, for every T we get a map $C([0,1],\mathcal{R}) \to C([0,1],\mathcal{R})$ where $x \to u$ is defined by the differential equation

$$\dot{u} = \dot{x} - \mu(Tu) \quad \text{for all } t \leq 1. \tag{6.6}$$

Alas this map $x \to u$ is not continuous under the supremum topology on C_0 which prevents us from utilising this approach. Interestingly, suppose we allow space-dependent rate of breeding. We believe that the rate function $r(x) = \frac{1}{2}\mu(x)^2$ will yield a continuous projection of the Brownian paths to the diffusion paths. However, we leave further discussion on the matter until a later time. Finally, watch out for section 6.4.3 in which we do use the contraction principle.

6.3.2 Proof of the Upper Bound

Let $2\epsilon > 0$. because $f(\theta) \to 1$ as $\theta \to 1$ we have that for some $\delta > 0$

$$\lim_{t\to\infty} \frac{r_{(1+\delta)t}}{r_t} = 1 + \epsilon.$$

We get that for sufficiently large t

$$\begin{aligned} P(R_t > (1+\epsilon)r_t) &\leq \sum_{i\in N_t} P(N_t = n)nP(U_i(t) > r_{(1+\delta)t}) \\ &\leq E(N_t)e^{-(1+\delta)t} = e^{-\delta t}. \end{aligned}$$

Take a sequence T_n of increasing times and define the stopping time $\tau_n \in [T_n, \infty]$ when we first after time T_n have $R_\tau > (1+2\epsilon)r_\tau$. If $\tau_n \in [T_n, T_{n+1}]$, there exists at time τ_n a particle above level $(1+2\epsilon)r_{\tau_n}$. Using lemma 10 we can estimate the probability that by time T_{n+1} the particle will still be above level $(1+\epsilon)r_{T_n}$. The deterministic path $u^{(1+2\epsilon)r_n}(\bullet)$ is bounded above by

$$u^{(1+2\epsilon)\tau_n}(T_{n+1}) \geq (1+2\epsilon)r_{T_n} - (T_{n+1}-T_n)\mu((1+2\epsilon)r_{T_{n+1}}).$$

Therefore, as long as the RHS satisfies

$$(1+2\epsilon)r_{T_n} - (T_{n+1}-T_n)\mu((1+2\epsilon)r_{T_{n+1}}) > (1+\epsilon)r_{T_{n+1}}, \tag{6.7}$$

the probability of the particle being above level $(1+\epsilon)r_{T_{n+1}}$ at time T_{n+1} is greater than some constant c. We deduce that

$$P\left(R_{T_{n+1}} > (1+\epsilon)r_{T_{n+1}}\right) \geq cP(\tau_n \in [T_n, T_{n+1}]).$$

We pick the increasing sequence T_n which satisfies equation 6.7 and observe that

$$\sum_n P(\tau_n \in [T_n, T_{n+1}]) \leq \frac{1}{c}\sum_n e^{-\delta T_n} < \infty$$

Hence almost surely

$$\limsup_{t\to\infty} \frac{R_t}{r_t} \leq 1+2\epsilon.$$

6.3.3 Proof of the Lower Bound

Let $T = T_k$. Define the event A_n as

$$A_n = \{\omega \in \Omega : |N_{(1-2\epsilon)T}| = n\}$$

Define $B_n \subset A_n$ as the event where one of these particles happened to be below $-2r_T$ at time $(1-2\epsilon)T$. Assuming we are in $A_n \setminus B_n$, we say a particle 'succeeds' if

- **At some time $\tau \in [(1-2\epsilon)T, (1-\epsilon)T]$ we have $U_i > -\epsilon r_T$.**
  Using lemma 11 we know that the probability of failure decays exponentially so that

the probability of success is greater than $\frac{1}{2}$.

and

- **By time T we have $U_i(T) > (f(1-5\epsilon)-\epsilon)r_T$.**
  Because a particle started at time $\tau < (1-\epsilon)T$ at a level greater or equal to $-\epsilon r_T$, we bound (from below) the probability of a particle succeeding by $P(U^0_{\epsilon T} > f(1-5\epsilon)r_T)$. Because $r_{\theta t}/r_t \to f(\theta)$ we bound this below by

$$P(U^0_{\epsilon T} > r_{(1-4\epsilon)T}).$$

  Using corollary 7 we know this is eventually greater than $e^{-(1-3\epsilon)T}$.

and

- **During a short period of time $[T_k, T_{k+1}]$ the particle stayed above level $(f(1-5\epsilon)-2\epsilon)r_{T_k}$.**
  As in the proof of the upper bound, If $(T_{k+1}-T_k)\mu(r_{T_k})$ is small when compared to ϵr_{T_n}, the probability of success is greater than some constant c.

Combining the three independent experiments we see that

$$p_T := P(\text{a particle succeeds}) \geq \tfrac{1}{2}ce^{-(1-3\epsilon)T}.$$

Given A_n, we say the system failed if either B_n occurred or all the particles failed. The probability that all $n = |N_{(1-2\epsilon)T}|$ particles in $A_n \setminus B_n$ failed is thus smaller than

$$P(\text{system failed}|A_n \setminus B_n) \leq (1-p_T)^{|N_{(1-2\epsilon)T}|}$$

Summing B_n and over all $A_n \setminus B_n$ separately, we get that the probability of a system failure is bounded above by

$$\sum_{n} P(B_n) + E\left((1-p_T)^{|N_{(1-2\epsilon)T}|}\right) = P(R_{(1-2\epsilon)T} > 2r_T) + G(1-p_T,(1-2\epsilon)T).$$

Both terms on the RHS decay exponentially in T. The exponential decay for the first term follows from the upper bound result, the exponential decay of the moment generating function $G(\lambda, t) = E(\lambda^{|N_t|})$ follows from lemma 2.
Summing over the T_k we have that almost surely, eventually we will always succeed in finding a particle in the whole of the interval $[T_k, T_{k+1}]$ above level $(f(1-5\epsilon)-2\epsilon)T_k$. Because $T_{k+1} - T_k < \epsilon T_k$, we know this level is above $(f(1-5\epsilon)-3\epsilon)T_{k+1}$. Hence

$$\liminf_{t\to\infty} \frac{R_t}{r_t} \geq f(1-5\epsilon) - 3\epsilon.$$

Let $\epsilon \downarrow 0$ and use the continuity of f to complete the proof.

6.3.4 A Large-Deviations Principle

By replacing r_t with $r_{\theta t}$ in the above proof of theorem 13 we get the following large-deviations principle.

Theorem 14. *Consider the μ-BBM satisfying the same conditions as those in theorem 13. Let $N_\theta(t)$ denote the particles whose position at time t is above $r_{\theta t}$.*

$$N_\theta(t) := \{i \in N_t : U_i(t) > r_{\theta t}\}.$$

Then, almost surely,

$$\lim_{t\to\infty} \frac{1}{t} \log |N_\theta(t)| = 1 - \theta.$$

6.3.5 The Point Process in $\mathcal{R}^2$ of a μ-BBM

We conclude this section with another application of theorem 13. Suppose we associate with each particle a vector in $\mathcal{R}^2$, representing $U_i(t)$ and $B_i(t)$. Formally, we let

$$p_i(t) = \left(\frac{1}{r_t}U_i(t), \frac{1}{t}B_i(t)\right).$$

Both theorems 13 and 14 hold for an arbitrary birth process satisfying only $E(C) < \infty$. Alas, to apply the large-deviations principle for a Brownian motion requires us to assume from now that $E(C \log C) < \infty$. We let $|\lambda| < \sqrt{2}$. We now alter the steps in the proof of theorem 13 as follows.

- **Run the process until time $(1-2\epsilon)T$ where $T = T_n$ and ϵ is small.**
  We now consider only particles which satisfy

$$\left\{ i \in N_{(1-2\epsilon)T} : |B_i((1-2\epsilon)T) - \lambda(1-2\epsilon)T| < 1 \right\}.$$

  We should get approximately $e^{(1-2\epsilon)(1-\frac{1}{2}\lambda^2)T}$ particles.

- **Perform an individual experiment with each particle**
  We say a particle "succeeds" if it manages to
  (I) Reach an ϵr_T neighbourhood of the origin before time $(1-\epsilon)T$.
  (II) Satisfy at time T, $U_i(T) > r_{\theta(1-8\epsilon)T}$.
  (III) Stay above $(1-\epsilon)r_{\theta(1-8\epsilon)t}$ for all $t \in [T_n, T_{n+1}]$.
  (IV) Its Brownian position at time $t \in [T_{n+1}, T_{n+1}]$ satisfies

$$|B_i(t) - \lambda t| \leq 4\epsilon t.$$

We now duplicate the proof of theorem 13. The only significant change we notice is that instead of of using result 2 to count the number of particles on which we perform the test, we use lemma 4. We deduce the following theorem.

Theorem 15. *Consider a μ-BBM satisfying the same conditions as in theorem 13. For all $A \subset \mathcal{R}^2$, define $N_A(t)$ as the set of all particles satisfying $p_i(t) \in A$. Let us also define the rate function I*

$$I(f(\theta), \lambda) = 1 - \theta - \tfrac{1}{2}\lambda^2.$$

Since f is increasing, this is well defined.

We claim that for all open sets $A \subset \mathcal{R}^2$ and closed set $D \subset \mathcal{R}^2$, almost surely

$$\limsup_{t \to \infty} \frac{1}{t} \log |N_D(t)| \leq \sup_{z \in A} I(z),$$
$$\liminf_{t \to \infty} \frac{1}{t} \log |N_A(t)| \geq \sup_{z \in A} I(z).$$

6.4 The Integrated μ-BBM

The final section of this work is somewhat analogous to chapter 5. It contains a beautiful proof of the large-deviations principle associated with the integral process of a μ-BBM. It is our hope that the reader will not be put off by the abundance of notations.

6.4.1 Notations

We associate two random processes Y and Z (henceforth referred to as the μ-integral process and the integral process respectively) with each particle.

$$dY_i(t) = \mu(U_i(t))dt,$$
$$dZ_i(t) = U_i(t)dt.$$

We assume that both processes start at 0 and that offspring inherit their ancestral position. At time T, we scale both of these by T in both the space and time coordinates.

$$y_i^T(t) = \frac{1}{T} Y_i(tT),$$
$$z_i^T(t) = \frac{1}{T} Z_i(tT).$$

These two paths are related using the map

$$z(t) = \int_0^t \mu^{-1}(\dot{y}(s))ds. \tag{6.8}$$

For every $A \subset C^0([0,1], \mathcal{R})$ we define

$$\mu M_A(T) = \{i \in N_T : y_i^T \in A\},$$
$$M_A(T) = \{i \in N_T : z_i^T \in A\}.$$

Finally, we let $q_i(t)$ be the point process on $\mathcal{R}^2$

$$q_i(t) = \left(\frac{1}{r_t}U_i(t), \frac{1}{t}Y_i(t)\right),$$

and for every $A \subset \mathcal{R}^2$ we let $\mu N_A(t) := \{i \in N_t : q_i \in A\}$.

6.4.2 The μ-Integral Process

The following is a trivial yet a powerful observation. Integrating

$$dB_i(t) - dY_i(t) = dU_i(t) \tag{6.9}$$

and dividing by t, we deduce that

$$||p_i(t) - q_i(t)||_{\mathcal{R}^2} = \frac{1}{t}|U_i(t)|,$$

and so almost surely, for all particles,

$$\sup_{i \in N_t} ||p_i(t) - q_i(t)|| \leq \frac{1}{t}R_t \to 0. \tag{6.10}$$

Corollary 9. *Theorem 15 still holds when we replace $N_A(t)$ with $\mu N_A(t)$.*

Proof. The family of projections $\pi_t : p_i(t) \to q_i(t)$ is almost surely uniformly continuous. We simply apply the contraction principle with $\pi_t \to \text{Id}$. □

6.4.3 The Path Space

There is more to observation 6.10. We can project a particle's entire scaled Brownian path to y_i^T and for all i, this will be almost surely uniformly continuous in the supremum topology on $C_1 := \{y \in C^1([0,1], \mathcal{R}) : y(0) = 0\}$.

Corollary 10. *Results 3, 4 and theorem 9 all hold when we replace $M_A(t), M_D(t)$ and C_0 with $\mu M_A(t), \mu M_D(t)$ and C_1 respectively.*

6.4.4 The Integral Process

The space of functions on which corollary 10 observed a large-deviations principle is C_1 and not C_0. Since C_1 differentiable, we prove next that we can refine the topology on C_1.
For every $\alpha \in \mathcal{R}^+$ define $||y||_\alpha$ on C_1 as

$$||y||_\alpha := \left\{ \int_0^1 |\dot{y}(t)|^\alpha dt. \right\}^{1/\alpha} \tag{6.11}$$

The topologies become finer as α increases with $||\bullet||_1$ still finer than the supremum topology. We complete C_1 with respect to this topology and arrive at the Sobolev spaces $W^{1,\alpha}([0,1])$, the spaces of functions with α-integrable weak derivatives.

Theorem 16. *Corollary 10 holds when we replace $(C_1, ||\bullet||_\infty)$ with $(W^{1,2}, ||\bullet||_2)$.*

Proof. Let $K_n = B_{\sqrt{n}}$ be the compact ball

$$K_n := \{y \in C_1 : 2I(y) = \int_0^1 \dot{y}^2 dt = ||y||_2^2 \leq n\}.$$

Clearly, if $y \in K_n^c$ then $K(y) \leq J(y) \leq 1 - \frac{1}{2}n$ and so the family of probability measures is still exponentially tight in the $||\bullet||_2$ topology. The result follows from the Inverse Contraction Theorem applied to the identity map. □

Corollary 11. *For every $\alpha \geq \frac{1}{2}$, consider a μ_α-BBM where $\mu_\alpha(x)$ is defined as*

$$\mu_\alpha(x) := \begin{cases} \frac{1}{2}x^\alpha & \textit{if } x > 0, \\ -\frac{1}{2}|x|^\alpha & \textit{if } x < 0, \end{cases}$$

Then, under the supremum topology in C_1, for every open set A and closed set D almost surely,

$$\liminf_{T\to\infty} T^{-1} \log |M_A(T)| \geq \sup_{z\in A} \mu K(z),$$
$$\limsup_{T\to\infty} T^{-1} \log |M_D(T)| \leq \sup_{z\in D} \mu K(z),$$

and also

$$\liminf_{T\to\infty} T^{-1} \log E|M_A(T)| \geq \sup_{z\in A} \mu J(z),$$
$$\limsup_{T\to\infty} T^{-1} \log E|M_D(T)| \leq \sup_{z\in D} \mu J(z).$$

where

$$\mu K(z) := K(y),$$
$$\mu J(z) := J(y).$$

J and K were defined for a BBM in chapter 4 and $y \to z$ using the bijection map 6.8.

Proof. We endow C_1 with the $||\bullet||_2$ topology. Looking back at equation 6.8, $y \to z$ defined as

$$(W^{1,2}, ||\bullet||_2) \to (W^{1,2\alpha}, ||\bullet||_{2\alpha}),$$
$$y \to \int \mu_\alpha^{-1}(\dot{y}),$$

is continuous. As long as $\alpha \geq \frac{1}{2}$, the topology on the RHS can be replaced with the supremum topology and the map will still be continuous. The result follows from the contraction principle. $\square$

Corollary 12. *For the μ-BBM described in corollary 11, the almost-sure and the expectation wavefronts of the integral process Z agree in the first order. Almost surely,*

$$\lim_{T\to\infty} T^{-1} \sup_i Z_i(T) = 2^{3/2\alpha}.$$

Proof. Pick a given growth rate $1-c$. Optimising $\int_0^1 \mu_\alpha^{-1}(\dot{y})dt$ subject to $\frac{1}{2}\int \dot{y}^2 = c$ we get the optimal linear path $\dot{y} = \sqrt{2c}$. Consequently $\int_0^1 \mu_\alpha^{-1}(\dot{y})dt = 2c^{3/2\alpha}$. Since these path exhibit identical behaviour in expectations and almost surely the equality of wavefronts follow. Letting $c \uparrow 1$ concludes the proof. $\square$

Bibliography

[1] Biggins, J D. *Martingales convergence in the branching random walk.*
J. Appl. Probability 14 (1977) p25-37.

[2] Biggins, J D. *Growth rates in the branching random walk.*
Z. Wahrscheinlichkeitstheorie 48 (1979) p17-34.

[3] Biggins, J D. *Uniform convergence in the branching random walk.*
Annals of Probability 20 (1992) p137-151.

[4] Bramson, M D. *Convergence of solutions of the Kolmogorov non-linear diffusion equations to travelling waves.*
Mem. Amer. Math. Society 285 (1983) p1-190.

[5] Britton, N F. *Reaction-diffusion equations and their application to biology.*
Academica Press, New York (1986).

[6] Champneys, A; Harris, S C; Toland J F; Warren J; Williams, D. *Algebra, analysis and probability for the coupled system of reaction-diffusion equations.*
Philosophical Transactions of the Royal Society A 350 (1995) p69-112.

[7] Chauvin, B; Rouault, A. *KPP equation and branching Brownian Motion in the subcritical speed-area. Application to spatial trees.*
Prob. Th. and Rel. Fields 80 (1988) p299-314.

[8] Chauvin, B; Rouault, A. *Supercritical branching brownian motion and the KPP equation in the critical speed area.*
Math. Nachr. 149 (1990) p41-59.

[9] Chauvin, B. *Product martingales and stopping lines for branching brownian motion.*
Ann. Prob. 19.3 (1991) p1195-1205.

[10] Daniels, H. *Saddlepoint approximations in statistics.*
Ann. Math. Statitistics 25 (1954) p631-650.

[11] Dekking, F M; Grimmett, G R. *Superbranching processes and projections* Prob. The. Rel. Fields, 78 (1988) p335-355.

[12] Dembo A; Zeitouni O. *Large-deviations techniques and applications.* Jones and Bartlett Publishers Boston and London (1995).

[13] Ellis, R S. *Entropy, large-deviations and statistical mechanics* Springer-Verlag, Heidelberg and New York (1985).

[14] Fisher, R A. *The wave of advance of an advatageous gene* Annals of Eugenics, 7 (1937) p353-369

[15] Git, Y. *Phase-plane analysis of branching brownian motion with drift.* Seminaré de Probabilités 17. (In print)

[16] Heyde, C C. *On large-deviation probabilities in the case of attraction to the non-normal stable law.* Sankkhya A 30.3 (1968) p253-258.

[17] Ikeda, N; Nagasawa, M; Watanabe, S. *Branching markov processes.* J. Math. Kyoto University 8 (1968) p233-278 and 9 (1969) p95-160.

[18] Kolmogorov, A N; Petrowski, I; Piscounov, N. *Étude de l'équation de la diffusion avec croissance de la quantité de la matière et son application à un problème biologique* Moscow University Bulletin of Mathematics 1 (1937) p1-25

[19] Marles, D. *Large deviations* Part III essay, University of Cambridge (1991).

[20] McKean, H P. *Applications of brownian motion to the equation of KPP* Comm. Pure Appl. Math. 28 (1975) p323-331 and 29 (1976) p553-554

[21] Murray J D. *Mathematical biology* Biomathematics Text 19. Springer-Verlag Heidelberg and New York (1989).

[22] Neveu, J. *Multiplicative martingales for spatial branching process.* Sem. Stochastic Processes Princeton (1987) p223-242.

[23] Petrov *Limit theorems of probability theory: sequences of independent random variables* OUP, Oxford (1995).

[24] Revesz, P. *Random walks of infinitely many particles* World Scientific, Singapore (1994).

[25] Schilder, M. *Some Asymptotic formulae for wiener integrals.*
Trans. Amer. Math. Soc. 125 (1966) p63-85.

[26] Stekel, D J; Rashbass, J; Williams, E D. *A computer graphic simulation of squamous epithelium.*
J. Theor. Biology 175 (1995) p283-293.

[27] Stone, C. *On local and ratio limit theorems.*
Fifth Berkeley Symposium 2.2 (1967) p217-224

[28] Strassen, V. *An invariance principle for the law of the iterated logarithm.*
Z. Wahrsch. Verw. Gebiete 3 (1964) p227-246.

[29] Strook D W. *An introduction to the theory of large-deviations.*
Springer-Verlag, Heidelberg and New York (1984).

[30] Uchiyama, K. *Spatial growth of a branching process of particles living in* $\mathcal{R}^d$.
Annals of Probability 10 (1982) p896-918.

[31] Varadhan, S R S. *Large-deviations and applications.*
CBMS-NSF Regional Conference Series in Applied Mathematics (1984).

[32] Watanabe, S. *Limit theorems for a class of branching processes*
Markov processes and potential theory, Wiley, New York (1967) p 205-232.

[33] Wentzell, A D. *Limit theorems on large-deviations for markov stochastic processes.*
Kluwer Academic Publishers, Dordrecht (1990).

[34] Warren, J. *Some aspects of branching processes.*
Ph.D. dissertation, University of Bath (1995).

[35] Williams, D. *Probability with martingales.*
Cambridge Mathematical Textbooks (1994).

[36] Williams, D. *Path decompostion and continuity of local time for one dimensional diffusions.*
Proc. London Math. Soc. 3.28 (1974) p738-768.

[37] Williams, D; Rogers, L C G. *Diffusions, markov processes, and martingales.*
Wiley, Chichester, New York (1987)