Branching Brownian Motion and Partial Differential Equations

Zachariah O'Dowd

Magdalen College

University of Oxford

A thesis submitted for the degree of

MMath Mathematics CCD Mathematics dissertation

Hilary 2019

Contents

List of Figures

1 Purpose

Our aim is to explore the intimate relationship between the semilinear heat equation

$$\frac{\partial u}{\partial t} = \frac{1}{2}\frac{\partial^2 u}{\partial x^2} + F(u), \qquad 0 \leq u(0, x) \leq 1$$

and the theory of stochastic processes. This can be regarded as a natural extension of the fundamental connection between Brownian Motion and the heat equation. Our key probabilistic object will be Branching Brownian Motion; we will demonstrate how the theory of stochastic processes can be used to characterize solutions to PDEs in terms of Branching Brownian Motion, and conversely how analytic PDE theory on the semilinear heat equation can uncover properties of Branching Brownian Motion.

We begin with a rigorous construction of Branching Brownian Motion, before examining two methods by which one can use it to fashion probabilistic representations of solutions to PDEs; the classical approach pioneered by Skorokhod ([Sko65]) and McKean ([McK75]) on the F-KPP equation, and the novel approach by Etheridge-Freeman-Penington ([EFP17]) which deploys the concept of voting schemes. Our primary interest is in describing the class of PDEs for which such a characterization exists, and our major contribution is to identify and characterize precisely which PDEs can be solved through voting scheme mechanisms, and to establish some duality results between voting schemes on Branching Brownian Motions with different branching structures. This is entirely original research, and our study is granular; we demonstrate how the notion of voting schemes can be gradually generalized in order to solve larger classes of PDEs. Lastly, we turn to studying the properties of the maximal process of Branching Brownian Motion through analytic PDE theory, with emphasis on the importance of branching structure. The existing literature is vast but often laconic; our antidote is to provide a self-contained account of this topic, discussing the contributions of Kolmogorov-Petrovskii-Piskunov ([KPP37]), McKean ([McK75]), Bramson ([Bra78], [Bra83]) and Lalley-Sellke ([LS87]), while adding considerable rigour to some of their insights.

Branching Brownian Motion has an incredibly rich mathematical structure; our study therefore draws from a real breadth of disciplines, including phase-plane analysis, stochastic analysis, functional analysis, and even basic combinatorics. Of course, analytic PDE theory predates the study of Branching Brownian Motion quite considerably, so the interplay between these two theories is often disorderly. We aim to give an exposition that draws fluidly from the two disciplines, and hopefully inspires a further interest in both.

2 Preliminaries

2.1 Measure Theory and Stochastic Processes

We shall assume familiarity with basic measure-theoretic and probabilistic concepts such as filtered probability spaces, random variables, distributions, measurability, and martingales. Concepts from Lebesgue integration theory, such as uniform integrability, conditional expectations, and the Monotone Convergence and Dominated Convergence Theorems, will also be assumed. For completeness' sake, we cement some standard notation.

(i) Given a collection $\mathcal{A}$ of random variables, $\sigma(\mathcal{A})$ denotes the smallest σ-algebra with respect to which each $X \in \mathcal{A}$ is measurable.

(ii) Given $A \subseteq \mathbb{R}$, $\mathcal{B}(A)$ is the associated Borel σ-algebra.

We will work on a filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ satisfying the usual conditions, and all stochastic processes will be assumed to be measurable, with state space $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Given a stochastic process $X = (X_t)_{t \geq 0}$, $x \in \mathbb{R}$, we shall write

(i) $\mathcal{F}_t^X = \sigma\{X_s : s \leq t\}$ for the natural filtration of X,

(ii) X^x for the process $x + X_t - X_0$,

(iii) $\mathbb{E}_x$ for the expectation under the law of X^x,

(iv) $\mathbb{P}_x$ for the probability measure under the law of X^x.

Definition 2.1.1. Let $\sigma > 0$. A stochastic process $B{=}(B_t)_{t \geq 0}$ is a σ^2-*Brownian Motion* if

(i) $B_0 = 0$ almost surely,

(ii) $B_{t+s} - B_s \sim \mathcal{N}(0, \sigma^2 t)$ for all $t > 0, s \geq 0$,

(iii) $\forall n \in \mathbb{N}$ and $0 \leq t_0 < t_1 < ... < t_n$, $\{B_{t_i} - B_{t_{i-1}}\}_{i=1}^{i=n}$ are independent,

(iv) The sample paths $t \mapsto B_t(\omega)$ are continuous for almost all $\omega \in \Omega$.

If $\sigma{=}1$ then B is a *standard Brownian Motion.*

In the context of stochastic calculus, Brownian Motion is in many senses the canonical example of a continuous stochastic process[1]. However, it should be noted that the existence of a Brownian Motion is far from trivial. It was first proven by Wiener ([Wie23]) in 1923, and later by Lévy ([Lév48]). We shall need the following results. Proofs can be found in [LG16] (*pp*.58, *pp*.61) and [Øks03] (*pp*.143), for example.

THEOREM 2.1.2 (Doob's Martingale Convergence Theorem). Let X_t be an L^1-bounded supermartingale with right-continuous sample paths. Then there exists $X_\infty \in L^1(\Omega)$ such that $X_t \to X_\infty$ almost surely.

THEOREM 2.1.3 (Optional Stopping Theorem). Let X_t be an $\mathcal{F}_t$-martingale with right-continuous sample paths. Let $\rho \leq \tau$ be bounded $\mathcal{F}_t$-stopping times. Then $X_\rho = \mathbb{E}[X_\tau | \mathcal{F}_\rho]$.

THEOREM 2.1.4 (Feynman-Kăc Formula). Let $c : [0, \infty) \times \mathbb{R} \to \mathbb{R}$ be bounded and continuous. Suppose that $u(t, x)$ satisfies

$$\frac{\partial u}{\partial t} = \frac{1}{2} \frac{\partial^2 u}{\partial x^2} - c(t, x)u,$$

that u is bounded on every $[0, t] \times \mathbb{R}$, and that $u(0, x)$ is continuous on $\mathbb{R} \backslash \{0\}$. Then

$$M_s = u(t - s, B_s) e^{-\int_0^s c(t-r, B_r) dr}$$

defines a bounded continuous martingale on $[0, t)$, where B is a standard Brownian Motion, and

$$u(t, x) = \mathbb{E}_x \left[u(0, B_t) e^{-\int_0^t c(t-r, B_r) dr} \right].$$

2.2 The Heat Equation

A fundamental object in PDE theory is the heat equation:

$$\frac{\partial u}{\partial t} = \kappa \frac{\partial^2 u}{\partial x^2}.$$

It is derived naturally from studying the distribution of heat in a given region over time, but appears in many other crevices of applied mathematics, including the study of image analysis, machine learning, and the Black-Scholes theory of options pricing.

[1] See, for example, the theorem of Dambis and Dubins-Schwarz (1965) in [RY99] (*pp*.181) or [LG16] (*pp*.121).

Owing to its omnipresence and simple appearance, the heat equation is our prototypical example of a parabolic PDE. It is intriguing, therefore, that there is a rather intimate relationship between Brownian Motion, which has a purely probabilistic construction, and the canonical parabolic PDE, which models physical phenomena. The analytic description of the following theorem was first demonstrated by Fourier ([Fou22]) in 1822, a century before Wiener's construction of Brownian Motion.

THEOREM 2.2.1. Let B_t be a standard Brownian Motion, $\phi : \mathbb{R} \to \mathbb{R}$ measurable. Then $u(t,x) = \mathbb{E}_x[\phi(B_t)]$ solves the Cauchy problem

$$\frac{\partial u}{\partial t}(t,x) = \frac{1}{2}\frac{\partial^2 u}{\partial x^2}(t,x) \qquad\qquad u(x,0) = \phi(x).$$

Proof.

$$\mathbb{E}_x[\phi(B_t)] = \int_{-\infty}^{\infty} \frac{e^{-\frac{(x-z)^2}{2t}}}{\sqrt{2\pi t}}\phi(z)dz.$$

From here we proceed analytically, noting that

$$\begin{aligned}
\frac{\partial}{\partial t}\left(\frac{e^{-\frac{(x-z)^2}{2t}}}{\sqrt{2\pi t}}\right) &= \frac{1}{2}\left(\frac{(x-z)^2}{\sqrt{2\pi t^5}} - \frac{1}{\sqrt{2\pi t^3}}\right)e^{-\frac{(x-z)^2}{2t}} \\
&= \frac{1}{2}\frac{1}{\sqrt{2\pi t^3}}\left(\frac{(x-z)^2}{t} - 1\right)e^{-\frac{(x-z)^2}{2t}} \\
&= \frac{1}{2}\frac{\partial^2}{\partial x^2}\left(\frac{e^{-\frac{(x-z)^2}{2t}}}{\sqrt{2\pi t}}\right).
\end{aligned}$$

The result follows. □

Remark 2.2.2. As Theorem 2.2.1 suggests, it is a convenience and a probabilistic convention to take $\kappa = \frac{1}{2}$ when working with the semilinear heat equation. While we will adhere to this convention, it should be noted that the PDE literature usually does not.

Definition 2.2.3. The *heat kernel* is the function $K(t,x,z) = \frac{1}{\sqrt{2\pi t}}e^{-\frac{(x-z)^2}{2t}}$.

Throughout this paper, we will make use of the fact that the heat kernel satisfies the heat equation. The following lemma will also be useful.

Lemma 2.2.4 (Sifting Property)**.** Let $\phi : \mathbb{R} \to \mathbb{R}$ be continuous and bounded. Let $x \in \mathbb{R}$. Then

$$\lim_{t\to 0}\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi t}}e^{-\frac{(x-z)^2}{2t}}\phi(z)dz = \phi(x).$$

Proof. We have

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi t}} e^{-\frac{(x-z)^2}{2t}} \phi(z)dz = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi t}} e^{-\frac{y^2}{2t}} \phi(x-y)dy$$
$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{v^2}{2}} \phi(x-\sqrt{t}v)dv.$$

By boundedness there exists $M > 0$ such that $|\phi| \leq M$ on $\mathbb{R}$, so

$$\left| \frac{1}{\sqrt{2\pi}} e^{-\frac{v^2}{2}} \phi(x-\sqrt{t}v) \right| \leq M \frac{1}{\sqrt{2\pi}} e^{-\frac{v^2}{2}}$$

for each $t > 0, v \in \mathbb{R}$. Now $\int_{-\infty}^{\infty} e^{-\frac{v^2}{2}} dv < \infty$, so the Dominated Convergence Theorem yields

$$\lim_{t\to 0} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{v^2}{2}} \phi(x-\sqrt{t}v)dv = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{v^2}{2}} \lim_{t\to 0} \phi(x-\sqrt{t}v)dv$$
$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{v^2}{2}} \phi(x)dv$$
$$= \phi(x).$$

□

At this point, a natural question to ask is can we represent solutions to more complicated parabolic PDEs of the form

$$\frac{\partial u}{\partial t} = \frac{1}{2}\frac{\partial^2 u}{\partial x^2} + F(u),$$

in terms of stochastic processes? It turns out that the requisite process is Branching Brownian Motion.

3 The Classical Approach

3.1 Branching Brownian Motion

Consider a particle B_t which follows a standard Brownian Motion. After an exponentially distributed time T, the particle dies, and simultaneously produces a random number of offspring. Each of these offspring then moves according to an independent standard Brownian Motion, started at B_T, for an independent lifetime with the same distribution as T, at the end of which it leaves behind a random number of offspring, and so on. Our Branching Brownian Motion X_t will describe the position of the particles alive at time $t > 0$.

In order to define Branching Brownian Motion rigorously, we first introduce the concept of Galton-Watson trees, which describe the dimensional properties of the process, and should be interpreted as a mathematical formulation of a family tree. Indeed, Galton-Watson introduced these objects while studying the extinction of family names, though their construction ([GW74]) is predated by Bienaymé's (largely overlooked) paper of 1845 ([Bie45]). Our construction follows [Ber15].

Definition 3.1.1. An *offspring distribution* is a probability measure on $\mathbb{N}_{\geq 0}$.

Let $\mathcal{U}:=\bigcup_{n=1}^{\infty} \mathbb{N}^n$. An element $u = (u_1, u_2, ..., u_n) \in \mathcal{U}$ can be thought of as a member of the nth generation. We write $|u| = n$ for the generation of u, and $p(u) = (u_1, u_2, ..., u_{n-1})$ for the parent of u. Note that $p^2(u)$ can be though of as the grandparent of u, and so on.

Definition 3.1.2. Let μ be an offspring distribution and $(\mathcal{C}_u : u \in \mathcal{U})$ a family of i.i.d random variables with law μ. The associated *Galton-Watson tree* is the (random) tree $\mathcal{T} \subset \mathcal{U}$ defined by

$$\mathcal{T} := \left\{u = (u_1, u_2, ..., u_n) \in \mathcal{U} : u_k \leq \mathcal{C}_{p^{n-k+1}(u)} \quad \forall k \leq n\right\} \cup \{\emptyset\}.$$

Remark 3.1.3. We can think of $\mathcal{C}_u$ as the number of children of u. Note that the following properties are both satisfied:

(i) If $\emptyset \neq u \in \mathcal{T}$, then $p(n) \in \mathcal{T}$.

(ii) For each $u = (u_1, u_2, ..., u_n) \in \mathcal{T}$, $(u_1, u_2, ...u_n, k) \in \mathcal{T}$ if and only if $k \leq \mathcal{C}_u$.

These correspond to the quite reasonable assumptions that any member of the tree has a parent and cannot produce infinitely many children.

For $u = (u_1, u_2, ..., u_n) \in \mathcal{T}$, we write $c(u) = \{(u_1, u_2, ..., u_n, k) \in \mathcal{T} : k \leq \mathcal{C}_u\}$ for the set of children of u. Note that $c \circ p(u)$ can be thought of as the set consisting of all 'siblings' of u and u itself.

Recall that we want to model a system where each particle produces offspring just as it dies. We enrich a Galton-Watson tree $\mathcal{T} \subset \mathcal{U}$ with this extra structure by endowing each $u \in \mathcal{T}$ with a lifespan ℓ_u, a birth-time $b_u := \sum_{k=1}^{|u|-1} \ell_{p^k(u)}$ ($b_\emptyset = 0$), and a death-time $d_u := b_u + \ell_u$. Then

$$d_{p(u)} = b_{p(u)} + \ell_{p(u)} = \sum_{k=1}^{|p(u)|-1} \ell_{p^{k+1}(u)} + \ell_{p(u)} = \sum_{k=1}^{|u|-2} \ell_{p^{k+1}(u)} + \ell_{p(u)} = \sum_{k=1}^{|u|-1} \ell_{p^k(u)} = b_u$$

as required. Note that, given $b_\emptyset = 0$, all other birth-times and death-times are uniquely specified by the lifespans. We now have all the machinery required to construct our Branching Brownian Motion.

Suppose that we are given an offspring distribution μ, and $\lambda > 0$. Let $\mathcal{T} \subset \mathcal{U}$ be a Galton-Watson tree with offspring distribution μ. Let $\boldsymbol{\ell} = (\ell_u)_{u \in \mathcal{T}}$ be i.i.d exponential(λ)-distributed lifespans. Let $\mathbf{W} = (W_u)_{u \in \mathcal{T}}$ be independent standard Brownian Motions. We define

$$\mathcal{N}_t := \{u \in \mathcal{T} : b_u \leq t < d_u\}$$

to be the set of particles alive at time t. We let $X_\emptyset(t) = W_\emptyset(t)$ for $0 \leq t < \ell_\emptyset$, and inductively define

$$X_u(t) := W_u(t - b_u) + X_{p(u)}(b_u-),$$

for each $u \in \mathcal{T} \backslash \{\emptyset\}$, to be the position of $u \in \mathcal{T}$ at time $t \in [b_u, d_u)$. It is convenient to let $X_u(t) := X_{p^k(u)}(t)$ for $t \in [b_{p^k(u)}, d_{p^k(u)})$, for each $k < |u|$, so that each X_u is defined on $[0, d_u)$.

Definition 3.1.4. The process $X_t := (X_u(t) : u \in \mathcal{N}_t)$ is a *Branching Brownian Motion with branching rate λ and offspring distribution μ.*

Remark 3.1.5. By taking $\mu = \delta_1$, where δ is the Dirac delta function, we see that a standard Brownian Motion is embedded in our definition of Branching Brownian Motion.

Remark 3.1.6. By the memoryless property of the exponential distribution and the Markov property of Brownian Motion, Branching Brownian Motion is a *branching process.* That is, for each $t > 0$ and $u \in \mathcal{N}_t$, the process

$$X_s^{(u)} = \big(X_v(t+s) - X_u(t) : v \in \mathcal{N}_{t+s}, u = p^k(v) \text{ for some } k \in \mathbb{N}\big)$$

is a Branching Brownian Motion with the same offspring distribution and branching rate as X_t. Furthermore, the distinct $X^{(u)}$ are independent. This property will be crucial in many of our ensuing proofs.

Definition 3.1.7. A Branching Brownian Motion is *n-adic* if its offspring distribution is $\mu(k) = \delta_{n,k}$. If $n = 2$, the process is *dyadic*; if $n = 3$, it is *triadic.*

Write $N(t) = |\mathcal{N}_t|$ for the number of particles alive at time t. We give two elementary properties of Branching Brownian Motion. The technique for proving this first lemma is implicit in many subsequent proofs.

Lemma 3.1.8. Let $X_t = (X_u(t) : u \in \mathcal{N}_t)$ be a Branching Brownian Motion with branching rate $\lambda > 0$ and offspring distribution μ, with $\gamma := \sum_{k=0}^{\infty} k\mu(k) < \infty$. Then $\mathbb{E}[N(t)] = e^{\lambda(\gamma-1)t}$.

Proof. Let $\ell = \ell_\emptyset$ be the lifespan of the original particle $\emptyset$. Let V be the number of offspring produced at time ℓ, so that V has law μ. We consider the partition

$$\Omega = \{\ell > t\} \cup \bigcup_{k=0}^{\infty} \{\ell \leq t, V = k\}.$$

On $\{\ell > t\}$, $\emptyset$ is still alive, so $\mathbb{E}[N(t)|\ell > t] = 1$. On $\{\ell \leq t, V = k\}$, we regard the process $(X_s)_{s \geq \ell}$ as k independent Branching Brownian Motions with branching rate λ and offspring distribution μ, started from time ℓ. Therefore, for each $0 \leq s \leq t$, we have $\mathbb{E}[N(t)|\ell = s, V = k] = k\mathbb{E}[N(t-s)]$. Since $\ell \sim \exp(\lambda)$, this gives

$$\begin{aligned}
\mathbb{E}[N(t)] &= \mathbb{E}[N(t)|\ell > t]\mathbb{P}(\ell > t) + \sum_{k=0}^{\infty} \mathbb{P}(V = k)\mathbb{E}[N(t)|\ell \leq t, V = k] \\
&= e^{-\lambda t} + \sum_{k=0}^{\infty} \mu(k) \int_0^t \lambda e^{-\lambda s} k\mathbb{E}[N(t-s)]ds \\
&= e^{-\lambda t} + \lambda\gamma e^{-\lambda t} \int_0^t e^{\lambda(t-s)}\mathbb{E}[N(t-s)]ds \\
&= e^{-\lambda t} + \lambda\gamma e^{-\lambda t} \int_0^t e^{\lambda u}\mathbb{E}[N(u)]du.
\end{aligned}$$

Differentiating,

$$\begin{aligned}
\frac{d}{dt}\mathbb{E}[N(t)] &= -\lambda e^{-\lambda t} + \lambda\gamma \left(-\lambda e^{-\lambda t} \int_0^t e^{\lambda u}\mathbb{E}[N(u)]du + \mathbb{E}[N(t)] \right) \\
&= \lambda\gamma\mathbb{E}[N(t)] - \lambda \left(e^{-\lambda t} + \lambda\gamma e^{-\lambda t} \int_0^t e^{\lambda u}\mathbb{E}[N(u)]du \right) \\
&= \lambda(\gamma - 1)\mathbb{E}[N(t)].
\end{aligned}$$

Therefore $\mathbb{E}[N(t)] = \mathbb{E}[N(0)]e^{\lambda(\gamma-1)t} = e^{\lambda(\gamma-1)t}$. □

Corollary 3.1.9 (Many-to-one Lemma). Fix $t > 0$ and let $F : C[0,t] \to \mathbb{R}$ be measurable. Then

$$\mathbb{E}\left[\sum_{u\in\mathcal{N}_t} F(X_u(s) : 0 \le s \le t)\right] = e^{\lambda(\gamma-1)t}\mathbb{E}[F(B_s : 0 \le s \le t)],$$

where B is a standard Brownian Motion.

Proof. Conditioning on $N(t)$, we have

$$\begin{aligned}\mathbb{E}\left[\sum_{u\in\mathcal{N}_t} F(X_u(s) : 0 \le s \le t)\right] &= \sum_{k=0}^{\infty}\mathbb{E}\left[\sum_{u\in\mathcal{N}_t} F(X_u(t) : 0 \le s \le t)|N(t) = k\right]\mathbb{P}(N(t) = k)\\ &= \sum_{k=0}^{\infty} k\mathbb{E}[F(B_s : 0 \le s \le t)]\mathbb{P}(N(t) = k)\\ &= \mathbb{E}[N(t)]\mathbb{E}[F(B_s : 0 \le s \le t)]\\ &= e^{\lambda(\gamma-1)t}\mathbb{E}[F(B_s : 0 \le s \le t)].\end{aligned}$$

□

The Many-to-one Lemma is useful because it allows us to describe the behaviour of the branches of a Branching Brownian Motion (of which there are a random number) in terms of one standard Brownian Motion.

3.2 The F-KPP Equation

The first class of semilinear PDEs we wish to consider is motivated by two seminal papers written in 1937, both concerned with biological problems. Kolmogorov-Petrovskii-Piskunov ([KPP37]) studied the semilinear heat equation

$$\frac{\partial u}{\partial t} = k\frac{\partial^2 u}{\partial x^2} + F(u)$$

for $k > 0$ and F satisfying $F(0) = F(1) = 0$, $F'(0) = \alpha > 0$, $F'(u) \le \alpha$ for all $u \in (0, 1]$. In the same year, the statistician and geneticist Fisher studied the diffusion of an advantageous gene through a linear habitat after a genetic mutation occurs. Letting $u(t, x)$ be the frequency of the mutant gene at time t, spatial position x, Fisher ([Fis37]) derived the semilinear heat equation with $F(u) = mu(1 - u)$. This latter class of PDEs is contained in the former, and consequently PDEs of this form (and more besides) are known as the F-KPP equation. After years of dormancy, these have recently garnered a great deal of interest in the PDE community. A survey paper, [Saa03], lists 453 papers devoted to its applications in physics, for example. Solutions to the F-KPP equation were studied in a purely analytic framework by Fife-McLeod ([FM75]) and Aronson-Weinberger ([AW75]), while probabilistic approaches were adopted by Skorokhod ([Sko65]), McKean ([McK75]), Bramson ([Bra78], [Bra83]), and countless others since. In this probabilistic setting, it is a convention to take $k = \frac{1}{2}$ and consider[1] the transformation $u \to 1 - u$. This yields the F-KPP equation:

$$\frac{\partial u}{\partial t} = \frac{1}{2}\frac{\partial^2 u}{\partial x^2} + mu(u-1).$$

The following theorem is somewhat groundbreaking, as it gives a connection between the F-KPP equation and Branching Brownian Motion. It is usually attributed to McKean ([McK75]), though it appears in the prior works of Skorokhod ([Sko65]) and Ikeda-Nagasawa-Watanabe ([INW65]).

THEOREM 3.2.1. Let $f : \mathbb{R} \to [0,1]$ be measurable. Let $X_t = (X_u(t) : u \in \mathcal{N}_t)$ be a dyadic Branching Brownian Motion with branching rate 1. Then

$$u(t,x) := \mathbb{E}\left[\prod_{u\in\mathcal{N}_t} f(x - X_u(t))\right]$$

solves the F-KPP equation

$$\frac{\partial u}{\partial t} = \frac{1}{2}\frac{\partial^2 u}{\partial x^2} + u(u-1) \qquad\qquad u(0,x) = f(x).$$

We actually prove a much more general theorem, which readily specializes to give McKean's result.

THEOREM 3.2.2. Let $f : \mathbb{R} \to [0,1]$ be measurable. Let $X_t = (X_u(t) : u \in \mathcal{N}_t)$ be a Branching Brownian Motion with branching rate λ and offspring distribution $\mu(k) = p_k$. Let $\Phi(x) = \sum_{k=0}^{\infty} p_k x^k$ be the generating function associated with μ. Then

$$u(t,x) := \mathbb{E}\left[\prod_{u\in\mathcal{N}_t} f(x - X_u(t))\right]$$

solves the semilinear PDE

$$\frac{\partial u}{\partial t} = \frac{1}{2}\frac{\partial^2 u}{\partial x^2} + \lambda(\Phi(u) - u) \qquad\qquad u(0,x) = f(x).$$

Proof. Our condition on f guarantees that the expectation exists. Let V be the number of offspring produced at time ℓ, where $\ell = \ell_\emptyset$ is the lifetime of the original particle $\emptyset$, so that V has law μ. We condition on $\{\ell > t\}$ and $\{\ell \le t, V = k\}$, for each

[1] Bramson is a notable and frustrating exception to this rule.

$k \in \mathbb{N}$. On $\{\ell > t\}$, $\emptyset$ is still alive at time t, so $(X_s)_{0\le s\le t}$ is just a standard Brownian Motion B. Therefore

$$\mathbb{E}\left[\prod_{u\in\mathcal{N}_t} f(x - X_u(t))|\ell > t\right] = \mathbb{E}[f(x - B_t)] = \int_{-\infty}^{\infty} \frac{e^{-\frac{z^2}{2t}}}{\sqrt{2\pi t}} f(x-z)dz.$$

On $\{\ell \le t, V = k\}$, we can think of the process $(X_t)_{t\ge\ell}$ as k independent Branching Brownian Motions with branching rate λ and offspring distribution μ, started from the point B_ℓ, where B is a standard Brownian Motion. Therefore, for each $0 \le s \le t$,

$$\begin{aligned}\mathbb{E}\left[\prod_{u\in\mathcal{N}_t} f(x - X_u(t))|\ell = s, V = k\right] &= \mathbb{E}\left[\mathbb{E}_{B_s}\left[\prod_{u\in\mathcal{N}_{t-s}} f(x - X_u(t-s))\right]^k\right] \\ &= \mathbb{E}[u(t-s, x - B_s)^k] \\ &= \int_{-\infty}^{\infty} \frac{e^{-\frac{z^2}{2s}}}{\sqrt{2\pi s}} u(t-s, x-z)^k dz.\end{aligned}$$

Since $\ell \sim \exp(\lambda)$, we can then express $u(t,x)$ as

$$\begin{aligned}u(t,x) &= \mathbb{P}(\ell > t)\mathbb{E}\left[\prod_{u\in\mathcal{N}_t} f(x - X_u(t))|\ell > t\right] + \sum_{k=0}^{\infty}\mathbb{P}(V = k)\mathbb{E}\left[\prod_{u\in\mathcal{N}_t} f(x - X_u(t))|\ell \le t, V = k\right] \\ &= e^{-\lambda t}\int_{-\infty}^{\infty} \frac{e^{-\frac{z^2}{2t}}}{\sqrt{2\pi t}} f(x-z)dz + \sum_{k=0}^{\infty} p_k \int_0^t \lambda e^{-\lambda s}\int_{-\infty}^{\infty} \frac{e^{-\frac{z^2}{2s}}}{\sqrt{2\pi s}} u(t-s, x-z)^k dzds.\end{aligned}$$

The proof is now entirely analytic. We differentiate term by term with respect to t. For ease of notation, let

$$h(t,x) = e^{-\lambda t}\int_{-\infty}^{\infty} \frac{e^{-\frac{z^2}{2t}}}{\sqrt{2\pi t}} f(x-z)dz, \quad g_k(t,x) = \int_0^t \lambda e^{-\lambda s}\int_{-\infty}^{\infty} \frac{e^{-\frac{z^2}{2s}}}{\sqrt{2\pi s}} u(t-s, x-z)^k dzds.$$

Now

$$\begin{aligned}\frac{\partial h}{\partial t}(t,x) &= -\lambda h(t,x) + e^{-\lambda t}\frac{\partial}{\partial t}\left(\int_{-\infty}^{\infty} \frac{e^{-\frac{z^2}{2t}}}{\sqrt{2\pi t}} f(x-z)dz\right) \\ &= -\lambda h(t,x) + \frac{1}{2}e^{-\lambda t}\frac{\partial^2}{\partial x^2}\left(\int_{-\infty}^{\infty} \frac{e^{-\frac{z^2}{2t}}}{\sqrt{2\pi t}} f(x-z)dz\right) \\ &= -\lambda h(t,x) + \frac{1}{2}\frac{\partial^2 h}{\partial x^2}(t,x)\end{aligned}$$

where we used the fact that the heat kernel satisfies the heat equation. The g_k terms are more troublesome. The substitution $v = t - s$ yields

$$g_k(t,x) = \int_0^t \lambda e^{-\lambda(t-v)}\int_{-\infty}^{\infty} \frac{e^{-\frac{z^2}{2(t-v)}}}{\sqrt{2\pi(t-v)}} u(v, x-z)^k dzdv.$$

Differentiating with respect to t and using Leibniz's Integral Rule[2] gives

$$\frac{\partial g_k}{\partial t}(t,x) = \int_0^t \lambda \frac{\partial}{\partial t}\left(e^{-\lambda(t-v)}\int_{-\infty}^{\infty}\frac{e^{-\frac{z^2}{2(t-v)}}}{\sqrt{2\pi(t-v)}}u(v,x-z)^k dz\right)dv + \lambda \lim_{v\to t}\int_{-\infty}^{\infty}\frac{e^{-\frac{z^2}{2(t-v)}}}{\sqrt{2\pi(t-v)}}u(t,x-z)^k dz$$

$$= -\lambda g_k(t,x) + \int_0^t \lambda e^{-\lambda(t-v)}\frac{\partial}{\partial t}\left(\int_{-\infty}^{\infty}\frac{e^{-\frac{z^2}{2(t-v)}}}{\sqrt{2\pi(t-v)}}u(v,x-z)^k dz\right) + \lambda u(t,x)^k$$

$$= -\lambda g_k(t,x) + \int_0^t \lambda e^{-\lambda(t-v)}\frac{1}{2}\frac{\partial^2}{\partial x^2}\left(\int_{-\infty}^{\infty}\frac{e^{-\frac{z^2}{2(t-v)}}}{\sqrt{2\pi(t-v)}}u(v,x-z)^k dz\right) + \lambda u(t,x)^k$$

$$= -\lambda g_k(t,x) + \frac{1}{2}\frac{\partial^2 g_k}{\partial x^2}(t,x) + \lambda u(t,x)^k,$$

where we used the sifting property of Lemma 2.2.4, and the fact that the heat kernel satisfies the heat equation. Piecing these calculations together,

$$\begin{aligned}
\frac{\partial u}{\partial t}(t,x) &= \frac{\partial h}{\partial t}(t,x) + \sum_{k=0}^{\infty} p_k \frac{\partial g_k}{\partial t}(t,x)\\
&= -\lambda h(t,x) + \frac{1}{2}\frac{\partial^2 h}{\partial x^2}(t,x) + \sum_{k=0}^{\infty} p_k\left(\frac{1}{2}\frac{\partial^2 g_k}{\partial x^2}(t,x) + \lambda u(t,x)^k - \lambda g_k(t,x)\right)\\
&= \frac{1}{2}\frac{\partial^2 u}{\partial x^2}(t,x) + \lambda\sum_{k=0}^{\infty} p_k u(t,x)^k - \lambda u(t,x)\\
&= \frac{1}{2}\frac{\partial^2 u}{\partial x^2}(t,x) + \lambda(\Phi(u(t,x)) - u(t,x)),
\end{aligned}$$

and clearly $u(0,x) = f(x)$. □

Remark 3.2.3. The assumption that $0 \le f \le 1$ can be relaxed. We really only need that f is measurable.

Remark 3.2.4. By taking X_t to be dyadic (so that $\Phi(x) = x^2$) with branching rate 1, we recover Theorem 3.2.1.

Theorem 3.2.2 gives us a simple probabilistic representation for solutions to a fairly large class of PDEs, some of which might have appeared rather fiendish at first glance. For example, by taking $\mu \sim$ Poisson(α), we can represent solutions to PDEs of the form

$$\frac{\partial u}{\partial t} = \frac{1}{2}\frac{\partial^2 u}{\partial x^2} + \lambda(e^{\alpha(u-1)} - u)$$

in terms of Branching Brownian Motion. Our class of PDEs can also be characterized as follows:

[2] See Appendix.

Corollary 3.2.5. Let $f : \mathbb{R} \to [0,1]$ be measurable, and let $F(u) = \sum_{k=0}^{\infty} a_k u^k$, with $a_1 \leq 0$, $a_k \geq 0$ for all $a \neq 1$, and $\sum_{k=0}^{\infty} a_k = 0$. Then there exists a probabilistic representation for the solution to the PDE

$$\frac{\partial u}{\partial t} = \frac{1}{2}\frac{\partial^2 u}{\partial x^2} + F(u) \qquad\qquad u(0,x) = f(x),$$

in terms of Branching Brownian Motion.

Proof. If $a_1 = 0$ then $a_k = 0$ for all $k \in \mathbb{N}$, so we just have the Cauchy problem, for which there certainly exists such a representation by Remark 3.1.5. Suppose then that $a_1 < 0$. Let $\lambda = -a_1$, $p_1 = 0$, and $p_k = \frac{a_k}{\lambda}$ for $k \neq 1$. Then $p_k \geq 0$ for all $k \in \mathbb{N}$, $\sum_{k=0}^{\infty} p_k = 1$, so that $\mu(k) := p_k$ defines an offspring distribution. Now

$$F(u) = \lambda \left(\sum_{k=0}^{\infty} \frac{a_k}{\lambda} + u - u \right) = \lambda \left(\sum_{k=0}^{\infty} p_k u^k - u \right).$$

Therefore we take X_t to be a Branching Brownian Motion with branching rate λ and offspring distribution μ, and apply Theorem 3.2.2. $\square$

4 Voting Schemes

4.1 Motivation

While the class of PDEs considered in the previous chapter is in some sense large, there are notable examples of semilinear PDEs which it does not accommodate, including one from [KPP37].

Consider an infinite population, where individuals carry two copies (alleles) of a gene that occurs as A or a. Let $u(t,x)$ be the concentration of the A-alleles at time t, position x. Assume that the A-alleles are uniformly distributed in space, and that after random mating, the expected genotype (allele pairs) concentrations follow the Hardy-Weinberg principle; that is, their concentrations are given[1] by the following table:

AA	Aa	aa
u^2	$2u(1-u)$	$(1-u)^2$

Suppose that the relative fitness of the genotypes is given by

AA	Aa	aa
$1+s$	$1+s$	1

where $s>0$ is small, so that cells with AA or Aa genotype produce $1+s$ times as many cells as those with genotype aa. Then, after random mating, the concentration of A-alleles is

$$\begin{aligned}\frac{(u^2+u(1-u))(1+s)}{(u^2+2u(1-u))(1+s)+(1-u)^2} &= \frac{u+su}{1-s(u^2-2u)}\\ &= (u+su)(1+s(u^2-2u)+\mathcal{O}(s^2))\\ &= u+s(u^3-2u^2+u)+\mathcal{O}(s^2)\\ &= u+su(u-1)^2+\mathcal{O}(s^2).\end{aligned}$$

Now write $s=\frac{\alpha}{N}$, where N describes the number of generations. If we measure time t in units of N generations, then taking $N\to\infty$ and adding diffusion gives

$$\frac{\partial u}{\partial t}=\frac{1}{2}\frac{\partial^2 u}{\partial x^2}+\alpha u(u-1)^2. \tag{4.1}$$

It is clear that $F(u)=\alpha u(1-u)^2$ does not lie in the class of functions considered in the previous chapter. In order to tackle this problem probabilistically, we use voting schemes. Etheridge-Freeman-Penington introduced majority voting schemes in [EFP17]. Here, we consider more general voting schemes.

[1] Biologically, there is a distinction between Aa and aA genotypes. Mathematically, this is not relevant and we write Aa to mean either.

4.2 Construction of Voting Schemes

For now, we let $X_t = (X_u(t) : u \in \mathcal{N}_t)$ be an n-adic Branching Brownian Motion with branching rate λ. We wish to construct a system where, at time t, each $u \in \mathcal{N}_t$ randomly votes 1 or 0 according to the value of some function at its position (this function will correspond to the initial condition of our PDE). The ancestors of $\mathcal{N}_t$ then vote 1 or 0 based on the votes of their children, so that the voting process descends the generations until $\emptyset$ has voted. This construction will require rigorous treatment, because the order in which the particles vote will depend on the branching structure of X_t; an individual can only vote once all its descendants have done so. Recall that $c \circ p(u)$ is the set consisting of u and its 'siblings'.

Definition 4.2.1. An *n-adic voting scheme* is a pair (θ, q), where $\theta : \{0,1\}^n \to \{0,1\}$ and $q : \mathbb{R} \to [0,1]$ is measurable.

Given an n-adic voting scheme (θ, q), the associated voting procedure $\mathbb{V}_q$ is constructed on X_t as follows. First each $u \in \mathcal{N}_t$ votes 1 with probability $q(X_u(t))$ and 0 otherwise, and we write $\mathbb{V}_q^u(t)$ for the vote of u. Let $k = \max\{|u| : u \in \mathcal{N}_t\}$ and define $\mathcal{N}^k := \{u \in \mathcal{N}_t : |u| = k\}$. Since each $u \in \mathcal{N}^k$ has already voted, and $c \circ p(\mathcal{N}^k) = \mathcal{N}^k$, the parents $p(\mathcal{N}^k)$ can all vote. By our Galton-Watson tree construction, for each $u \in p(\mathcal{N}^k)$ there is a natural ordering associated with its offspring, which we write as $c(u) = \{u_1, u_2, ..., u_n\}$. We define the vote of u to be

$$\mathbb{V}_q^u(t) = \theta\left(\mathbb{V}_q^{u_1}(t), \mathbb{V}_q^{u_2}(t), .., \mathbb{V}_q^{u_n}(t)\right).$$

Inductively, for $1 < r \leq k$, once the $p(\mathcal{N}^r)$ have all voted we define $\mathcal{N}^{r-1} = p(\mathcal{N}^r) \cup \{u \in \mathcal{N}_t : |u| = r-1\}$ and repeat the same process as above; note that the $\mathcal{N}^{r-1}$ have all voted, and $c \circ p(\mathcal{N}^{r-1}) = \mathcal{N}^{r-1}$, so the parents $p(\mathcal{N}^{r-1})$ can all vote, and do so by applying θ as above. For $u \in p(\mathcal{N}^{r-1})$, we write the vote of u as $\mathbb{V}_q^u(t)$. Since $p(\mathcal{N}^1) = \{\emptyset\}$, the process terminates with $\emptyset$ casting its vote, for which we just write $\mathbb{V}_q(t)$.

Remark 4.2.2. The function $\theta : \{0,1\}^n \to \{0,1\}$ should be thought of as a partition of $\{0,1\}^n$ into two, where $\theta^{-1}(1)$ is the set of all combinations of offspring votes which induce a 1 vote, and $\theta^{-1}(0)$ the set of all combinations of offspring votes which induce a 0 vote.

PROPOSITION 4.2.3. Let (θ, q) be an n-adic voting scheme and $\mathbb{V}_q$ the associated voting procedure on X_t. Then $u(t,x) := \mathbb{E}_x[\mathbb{V}_q(t)] = \mathbb{P}_x(\mathbb{V}_q(t) = 1)$ solves the PDE

$$\frac{\partial u}{\partial t} = \frac{1}{2}\frac{\partial^2 u}{\partial x^2} + \lambda\left(\mathbb{E}[\theta(V_1(u), ..., V_n(u))] - u\right) \qquad u(0,x) = q(x),$$

where the $V_k(u(t,x))$ are i.i.d Bernoulli($u(t,x)$) random variables.

Proof. Let $\ell = \ell_\emptyset$ be the lifespan of the original particle $\emptyset$. On $\{\ell > t\}$, $\mathbb{V}_q(t)$ is determined entirely by the position of $\emptyset$ at time t, so

$$\mathbb{E}_x[\mathbb{V}_q(t)|\ell > t] = \mathbb{E}_x[q(B_t)] = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi t}} e^{-\frac{z^2}{2t}} q(x-z)dz,$$

where B is a standard Brownian Motion. On $\{\ell \le t\}$, $\mathbb{V}_q(t) = \theta\left(\mathbb{V}_q^{(1)}(t), ..., \mathbb{V}_q^{(n)}(t)\right)$ and we can think of the process $(X_s)_{s\ge\ell}$ as n independent n-adic Branching Brownian Motions with branching rate m started from B_ℓ, which all inherit the voting scheme (θ, q) and voting procedure $\mathbb{V}_q$, so that the particles $(1), (2), ..., (n)$ independently vote 1 with probability $\mathbb{E}_{B_\ell}[\mathbb{V}_q(t-\ell)] = u(t-\ell, B_\ell)$, and 0 otherwise. Hence $\mathbb{V}_q^{(k)}(t) = V_k(u(t-\ell, B_\ell))$, where the $V_k(u)$ are i.i.d Bernoulli(u) random variables, for $k = 1, ..., n$. Therefore,

$$\begin{aligned}
\mathbb{E}_x[\mathbb{V}_q(t)|\ell = s] &= \mathbb{E}_x[\theta(\mathbb{V}_q^{(1)}(t), ..., \mathbb{V}_q^{(n)}(t))|\ell = s] \\
&= \mathbb{E}_x[\mathbb{E}[\theta(V_1(u(t-s, B_s)), ..., V_n(u(t-s, B_s)))]] \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi s}} e^{-\frac{z^2}{2s}} \mathbb{E}[\theta(V_1(u(t-s, x-z)), ..., V_n(u(t-s, x-z)))]dz
\end{aligned}$$

for each $0 \le s \le t$. Since $\ell \sim \exp(\lambda)$,

$$u(t,x) = e^{-\lambda t}\int_{-\infty}^{\infty} \frac{e^{-\frac{z^2}{2t}}}{\sqrt{2\pi t}} q(x-z)dz + \int_0^t \lambda e^{-\lambda s}\int_{-\infty}^{\infty} \frac{e^{-\frac{z^2}{2s}}}{\sqrt{2\pi s}} \mathbb{E}[\theta(V_1, ..., V_n)u(t-s, x-z)]dzds.$$

We now use precisely the same analytic argument as in Theorem 3.2.2. Letting

$$\begin{aligned}
h(t,x) &= e^{-\lambda t}\int_{-\infty}^{\infty} \frac{e^{-\frac{z^2}{2t}}}{\sqrt{2\pi t}} q(x-z)dz, \\
g(t,x) &= \int_0^t \lambda e^{-\lambda s}\int_{-\infty}^{\infty} \frac{e^{-\frac{z^2}{2s}}}{\sqrt{2\pi s}} \mathbb{E}[\theta(V_1(u(t-s, x-z)), ..., V_n(u(t-s, x-z)))]dzds,
\end{aligned}$$

we find, by the sifting property of the heat kernel and the fact that it satisfies the heat equation, that

$$\begin{aligned}
\frac{\partial h}{\partial t}(t,x) &= -\lambda h(t,x) + \frac{1}{2}\frac{\partial^2 h}{\partial x^2}(t,x), \\
\frac{\partial g}{\partial t}(t,x) &= -\lambda g(t,x) + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(t,x) + \lambda\mathbb{E}[\theta(V_1(u(t,x)), ..., V_n(u(t,x)))].
\end{aligned}$$

Hence

$$\begin{aligned}\frac{\partial u}{\partial t} &= -\lambda(h+g) + \frac{1}{2}\left(\frac{\partial^2 h}{\partial x^2} + \frac{\partial^2 g}{\partial x^2}\right) + \lambda\mathbb{E}[\theta(V_1(u), ..., V_n(u))] \\ &= \frac{1}{2}\frac{\partial^2 u}{\partial x^2} + \lambda\left(\mathbb{E}[\theta(V_1(u), ..., V_n(u))] - u\right),\end{aligned}$$

and clearly $u(0, x) = q(x)$. $\square$

Definition 4.2.4. We say that a function $F(u)$ has an *n-adic voting scheme representation* if there exists a function $\theta : \{0,1\}^n \to \{0,1\}$ and $\lambda > 0$ such that, for all measurable $q : \mathbb{R} \to [0,1]$, if $\mathbb{V}_q$ is the voting procedure associated with the voting scheme (θ, q) on an n-adic Branching Brownian Motion with branching rate λ, then $u(t,x) = \mathbb{E}_x[\mathbb{V}_q(t)]$ solves the PDE

$$\frac{\partial u}{\partial t} = \frac{1}{2}\frac{\partial^2 u}{\partial x^2} + F(u), \qquad u(0,x) = q(x).$$

Remark 4.2.5. Extensions of this notion, such as *n-adic voting measure representations* and *generalized voting scheme representations*, will be introduced as our granular study develops. It is cumbersome to define them explicitly and we refrain from doing so, but their meaning will be obvious from their context.

Writing PDEs in terms of independent Bernoulli random variables, as in Proposition 4.2.3, is a bit of an annoyance. Fortunately, there is a much nicer way of presenting this class of PDEs.

Lemma 4.2.6. There is a 1-1 correspondence between functions of the form $G(u) = \mathbb{E}[\theta(V_1(u), ..., V_n(u))]$, where $\theta : \{0,1\}^n \to \{0,1\}$, and the $V_k(u)$ are i.i.d Bernoulli(u) random variables, and polynomials of the form $G(u) = \sum_{k=0}^n a_k u^k (1-u)^{n-k}$, where $a_k \in \mathbb{N}$, $0 \le a_k \le \binom{n}{k}$.

Proof. Given $\theta : \{0,1\}^n \to \{0,1\}$, there exists $a_k \in \mathbb{N}$ with $0 \le a_k \le \binom{n}{k}$, and a_k distinct partitions $(\{r_{1,j}, r_{2,j}, ..., r_{k,j}\}, \{r_{k+1,j}, ..., r_{n,j}\})_{j=1}^{a_k}$ of $\{1, 2., , , n\}$ into two sets of size k and $n-k$, such that

$$\theta(x_1, ...x_n) = \sum_{k=0}^{n}\sum_{j=1}^{a_k}\prod_{i=1}^{k} x_{r_{i,j}} \prod_{i=k+1}^{n}(1 - x_{r_{i,j}}).$$

While messy, this representation gives

$$\begin{aligned}
\mathbb{E}[\theta(V_1(u),...,V_n(u))] &= \sum_{k=0}^{n}\sum_{j=1}^{a_k}\mathbb{E}\left[\prod_{i=1}^{k}V_{r_{i,j}}(u)\prod_{i=k+1}^{n}(1-V_{r_{i,j}}(u))\right] \\
&= \sum_{k=0}^{n}\sum_{j=1}^{a_k}\prod_{i=1}^{k}\mathbb{E}[V_{r_{i,j}}(u)]\prod_{i=k+1}^{n}(1-\mathbb{E}[V_{r_{i,j}}(u)]) \\
&= \sum_{k=0}^{n}\sum_{j=1}^{a_k}\prod_{i=1}^{k}u\prod_{i=k+1}^{n}(1-u) \\
&= \sum_{k=0}^{n}a_k u^k(1-u)^{n-k},
\end{aligned}$$

where the second line follows from independence of the $V_k(u)$. Conversely, suppose that $G(u) = \sum_{k=0}^{n} a_k u^k (1-u)^{n-k}$, with $a_k \in \mathbb{N}$ and $0 \leq a_k \leq \binom{n}{k}$. For each k we pick distinct partitions $(\{r_{1,j}, r_{2,j}, ..., r_{k,j}\}, \{r_{k+1,j}, ..., r_{n,j}\})_{j=1}^{a_k}$ of $\{1, 2, ..., n\}$ into two sets of size k and $n-k$, and define $\theta : \{0,1\}^n \to \{0,1\}$ by

$$\theta(x_1, ...x_n) = \sum_{k=0}^{n}\sum_{j=1}^{a_k}\prod_{i=1}^{k}x_{r_{i,j}}\prod_{i=k+1}^{n}(1-x_{r_{i,j}}).$$

Then $\mathbb{E}[\theta(V_1(u),...,V_n(u))] = G(u)$ by the same argument as above, where the $V_k(u)$ are i.i.d Bernoulli(u) random variables. □

Corollary 4.2.7. The class of functions which have n-adic voting scheme representations is precisely the polynomials of the form

$$F(u) = \lambda\left(\sum_{k=0}^{n}a_k u^k(1-u)^{n-k} - u\right),$$

where $\lambda > 0$, $n \in \mathbb{N}$, and $a_k \in \mathbb{N}$ with $0 \leq a_k \leq \binom{n}{k}$.

Armed with this representation, Kolmogorov-Petrovskii-Piskunov's PDE (4.1) is easily solved. We simply note that

$$\begin{aligned}
\alpha u(1-u)^2 &= \alpha(u(1-u)^2 + u - u) \\
&= \alpha(u(1-u)^2 + u(u+1-u)^2 - u) \\
&= \alpha(u^3 + 2u^2(1-u) + 2u(1-u)^2 - u),
\end{aligned}$$

so that $F(u) = \alpha u(1-u)^2$ has an n-adic voting scheme representation. We might, for example, consider a triadic Branching Brownian Motion with branching rate α, and a voting scheme defined by the function

$$\theta(x_1, x_2, x_3) = x_1x_2x_3 + x_1x_2(1-x_3) + x_1x_3(1-x_2) + x_1(1-x_2)(1-x_3) + x_2(1-x_1)(1-x_3).$$

Heuristically, it is clear that an n-adic voting scheme can be embedded in an m-adic voting scheme when $m > n$, for we can consider a voting procedure on an m-adic Branching Brownian Motion which is 'blind' with respect to $m-n$ of the m offspring. More formally, we show that we can embed an n-adic voting scheme inside an m-adic voting scheme.

Lemma 4.2.8. Suppose that $F(u)$ has an n-adic voting scheme representation and let $m > n$. Then $F(u)$ has an m-adic voting scheme representation.

Proof. By induction, it suffices to show that $F(u)$ has an $(n+1)$-adic voting scheme representation. By Corollary 4.2.7, F is of the form

$$F(u) = \lambda \left(\sum_{k=0}^{n} a_k u^k (1-u)^{n-k} - u \right),$$

where $\lambda > 0$ and $a_k \in \mathbb{N}$ with $0 \leq a_k \leq \binom{n}{k}$. Then

$$\begin{aligned}
\sum_{k=0}^{n} a_k u^k (1-u)^{n-k} &= \sum_{k=0}^{n} a_k u^k (1-u)^{n-k} (u + 1 - u) \\
&= \sum_{k=0}^{n} a_k u^{k+1} (1-u)^{n-k} + \sum_{k=0}^{n} a_k u^k (1-u)^{n+1-k} \\
&= a_0 (1-u)^{n+1} + \sum_{k=1}^{n} (a_{k-1} + a_k) u^k (1-u)^{n+1-k} + a_n u^{n+1}.
\end{aligned}$$

Note that for each $k = 1, ..., n$, we have

$$0 \leq a_{k-1} + a_k \leq \binom{n}{k-1} + \binom{n}{k} = \binom{n+1}{k},$$

so letting $b_0 = a_0$, $b_k = a_{k-1} + a_k$ for $k = 1, ..., n$, and $b_{n+1} = a_n$, we have

$$F(u) = \lambda \left(\sum_{k=0}^{n+1} b_k u^k (1-u)^{n+1-k} - u \right)$$

with $b_k \in \mathbb{N}$ and $0 \leq b_k \leq \binom{n+1}{k}$. Therefore, by Corollary 4.2.7, $F(u)$ has an $(n+1)$-adic voting scheme representation. $\square$

There are several directions in which we might wish to generalize our notion of voting schemes in order to solve a larger class of semilinear PDEs. We shall see that increased generality in our constructions allows us to consider more general functions $F(u)$.

4.3 Voting Measures

Our first generalization concerns voting procedures which choose between different voting schemes each time the Branching Brownian Motion splits. More formally, instead of associating a voting scheme with a fixed function $\theta : \{0,1\}^n \to \{0,1\}$, we wish to consider a probability measure on the space of all such functions. As before, let $X_t = (X_u(t) : u \in \mathcal{N}_t)$ be an n-adic Branching Brownian Motion with branching rate λ. We write $\mathcal{D}$ for the space of all functions $\theta : \{0,1\}^n \to \{0,1\}$.

Definition 4.3.1. An *n-adic voting measure* is a pair (μ, q), where $\mu : \mathcal{D} \to [0,1]$ is a probability measure and $q : \mathbb{R} \to [0,1]$ is measurable.

Given a voting measure (μ, q) on X_t, we define the associated voting procedure $\mathbb{V}_q$ as before, except this time each $u \in p(\mathcal{N}^r)$ independently chooses $\theta \in \mathcal{D}$ with probability $\mu(\theta)$ before casting its vote as

$$\mathbb{V}_q^u(t) = \theta\left(\mathbb{V}_q^{u_1}(t), \mathbb{V}_q^{u_2}(t), .., \mathbb{V}_q^{u_n}(t)\right),$$

where $c(u) = \{u_1, u_2, ..., u_n\}$ are the offspring of u.

Remark 4.3.2. Voting measures subsume voting schemes, for we can consider a probability measure μ that takes the value 1 on some fixed $\theta \in \mathcal{D}$, and 0 elsewhere. The associated voting procedures of θ and μ are then precisely the same.

PROPOSITION 4.3.3. Let (μ, q) be a voting measure on X_t. Let $\mathbb{V}_q$ be the associated voting procedure. Then $u(t,x) := \mathbb{E}_x[\mathbb{V}_q(t)] = \mathbb{P}_x(\mathbb{V}_q(t) = 1)$ solves the PDE

$$\frac{\partial u}{\partial t} = \frac{1}{2}\frac{\partial^2 u}{\partial x^2} + \lambda\left(\sum_{\theta \in \mathcal{D}} \mu(\theta)\mathbb{E}[\theta(V_1(u), ..., V_n(u))] - u\right) \qquad u(0,x) = q(x),$$

where the $V_k(u(t,x))$ are i.i.d Bernoulli($u(t,x)$) random variables.

Proof. Let $\ell = \ell_\emptyset$ be the lifespan of the original particle $\emptyset$ and let (Θ, q) be the n-adic voting scheme adopted by $\emptyset$. Then $\ell \sim \exp(\lambda)$ and Θ has law μ. Conditioning on the events $\{\ell > t\}$, and $\{\ell \le t, \Theta = \theta\}$, for each $\theta \in \mathcal{D}$, and using precisely the same argument as in Proposition 4.2.3, we have

$$\mathbb{E}_x[\mathbb{V}_q(t)|\ell > t] = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi t}} e^{-\frac{z^2}{2t}} q(x-z)dz,$$

$$\mathbb{E}_x[\mathbb{V}_q(t)|\ell = s, \Theta = \theta] = \int_{-\infty}^{\infty} \frac{e^{-\frac{z^2}{2s}}}{\sqrt{2\pi s}} \mathbb{E}[\theta(V_1(u(t-s, x-z)), ..., V_n(u(t-s, x-z)))]dz,$$

for each $0 \leq s \leq t$ and $\theta \in \mathcal{D}$. We have

$$u(t,x) = e^{-\lambda t}\int_{-\infty}^{\infty}\frac{e^{-\frac{z^2}{2t}}}{\sqrt{2\pi t}}q(x-z)dz + \sum_{\theta\in\mathcal{D}}\mu(\theta)\left(\int_0^t \lambda e^{-\lambda s}\int_{-\infty}^{\infty}\frac{e^{-\frac{z^2}{2s}}}{\sqrt{2\pi s}}\mathbb{E}[\theta(V_1,...,V_n)u(t-s,x-z)]dzds\right),$$

and precisely the same analytic argument as in Theorem 3.2.2 yields

$$\frac{\partial u}{\partial t} = \frac{1}{2}\frac{\partial^2 u}{\partial x^2} + \lambda\left(\sum_{\theta\in\mathcal{D}}\mu(\theta)\mathbb{E}[\theta(V_1(u),...,V_n(u))] - u\right).$$

Clearly, $u(0,x) = q(x)$. □

Analogously to Corollary 4.2.7, we seek a nicer presentation for this class of PDEs.

Lemma 4.3.4. There is a 1-1 correspondence between functions of the form $G(u) = \sum_{\theta\in\mathcal{D}}\mu(\theta)\mathbb{E}[\theta(V_1(u),...,V_n(u))]$ - where $\mu : \mathcal{D} \to [0,1]$ is a probability measure and the $V_k(u)$ are i.i.d Bernoulli(u) random variables - and polynomials of the form

$$G(u) = \sum_{r=1}^{l} p_r \sum_{k=0}^{n} a_{k,r}u^k(1-u)^{n-k},$$

where $m > 0$, $p_r \geq 0$ with $\sum_{r=1}^{l} p_r = 1$ and, for each r, $a_{k,r} \in \mathbb{N}$ with $0 \leq a_{k,r} \leq \binom{n}{k}$.

Proof. By Lemma 4.2.6, for each $\theta \in \mathcal{D}$, we can write

$$\mathbb{E}[\theta(V_1(u),...,V_n(u))] = \sum_{k=0}^{n} a_{k,\theta}u^k(1-u)^{n-k},$$

where $a_{k,\theta} \in \mathbb{N}$ and $0 \leq a_{k,\theta} \leq \binom{n}{k}$ for each k. Therefore, if $\mu : \mathcal{D} \to [0,1]$ is a probability measure, we have

$$\sum_{\theta\in\mathcal{D}}\mu(\theta)\mathbb{E}[\theta(V_1(u),...,V_n(u))] = \sum_{\theta\in\mathcal{D}}\mu(\theta)\sum_{k=0}^{n} a_{k,\theta}u^k(1-u)^{n-k}.$$

Conversely, suppose that G is of the form

$$G(u) = \sum_{r=1}^{l} p_r \sum_{k=0}^{n} a_{k,r}u^k(1-u)^{n-k},$$

where $p_r \geq 0$ with $\sum_{r=1}^{l} p_r = 1$ and, for each r, $a_{k,r} \in \mathbb{N}$ with $0 \leq a_{k,r} \leq \binom{n}{k}$. Then by Lemma 4.2.6, for each $r = 1,...,l$ there exists $\theta_r \in \mathcal{D}$ such that

$$\sum_{k=0}^{n} a_{k,r}u^k(1-u)^{n-k} = \mathbb{E}[\theta_r(V_1(u),...,V_n(u))],$$

where the $V_k(u)$ are i.i.d Bernoulli(u) random variables. Then

$$\mu(\theta) := \begin{cases} p_r & \text{if } \theta = \theta_r \text{ for some } 0 \leq r \leq l \\ 0 & \text{otherwise,} \end{cases}$$

defines a probability measure on $\mathcal{D}$, and now

$$G(u) = \sum_{\theta \in \mathcal{D}} \mu(\theta) \mathbb{E}[\theta(V_1(u), ..., V_n(u)].$$

$\square$

We have the following characterization of our new class of semilinear PDEs.

Corollary 4.3.5. The class of functions which have n-adic voting measure representations is precisely the polynomials of the form

$$F(u) = \lambda \left(\sum_{r=1}^{l} p_r \sum_{k=0}^{n} a_{k,r} u^k (1-u)^{n-k} - u \right),$$

where $\lambda > 0$, $p_r \geq 0$ with $\sum_{r=1}^{l} p_r = 1$ and, for each r, $a_{k,r} \in \mathbb{N}$ with $0 \leq a_{k,r} \leq \binom{n}{k}$.

Therefore, voting measures allow us to relax the conditions on the coefficients of the polynomials $F(u)$. The drawback is that these coefficients are elusive, and consequently the voting measures can be difficult to construct. However, there is a subclass which does have a nice representation.

Proposition 4.3.6. Let

$$F(u) = \sum_{k=0}^{n} a_k u^k (1-u)^{n-k},$$

where $a_k \geq 0$ and $a_n = 0$. Then $F(u)$ has a voting measure representation.

Proof. The proof hinges on the trivial observation that $1 = u + 1 - u$. Writing $a = \sum_{k=0}^{n} a_k$, we have

$$F(u) = a \sum_{k=0}^{n-1} \frac{a_k}{a} \left(u^k (1-u)^{n-k} + u - u \right),$$

and for $k = 0, ..., n-1$,

$$\begin{aligned} u^k(1-u)^{n-k} + u &= u^k(1-u)^{n-k} + u(u+1-u)^{n-1} \\ &= u^k(1-u)^{n-k} + u \sum_{i=0}^{n-1} \binom{n-1}{i} u^i (1-u)^{n-1-i} \\ &= u^k(1-u)^{n-k} + \sum_{i=1}^{n} \binom{n-1}{i-1} u^i (1-u)^{n-i}. \end{aligned}$$

Since $\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}$, we have $\binom{n-1}{k-1} + 1 \leq \binom{n}{k}$ for $k = 1, ..., n-1$. Letting $p_k = \frac{a_k}{a}$, so that $p_k \geq 0$ and $\sum_{k=0}^{n-1} p_k = 1$, we see that $F(u)$ lies in the class of functions that admit n-adic voting measure representations. □

The benefit of functions of this class is twofold. Firstly, given any polynomial, there is a clear algorithm which will determine whether of not it lies in this class. Secondly, Proposition 4.3.6 gives us a canonical method for determining the branching rate of the corresponding n-adic Branching Brownian Motion, and constructing an appropriate n-adic voting measure.

4.4 Generalized Voting Schemes

We continue to fix an n-adic Branching Brownian Motion $X_t = (X_u(t) : u \in \mathcal{N}_t)$ with branching rate λ. In our original construction of voting schemes, each particle's vote was deterministic once all of its offspring had voted. We now consider the case when each vote is probabilistic; that is, we assign each combination of offspring votes a probability, which corresponds to the probability that a particle will vote 1 given that it has observed that particular combination of offspring votes.

Definition 4.4.1. A *generalized n-adic voting scheme* is a pair (θ, q), where $\theta : \{0,1\}^n \to [0,1]$ and $q : \mathbb{R} \to [0,1]$ is measurable.

The associated voting procedure $\mathbb{V}_q$ is defined in line with previous constructions, but now each $u \in p(\mathcal{N}^r)$ votes 1 with probability $\theta\left(\mathbb{V}_q^{u_1}(t), \mathbb{V}_q^{u_2}(t), .., \mathbb{V}_q^{u_n}(t)\right)$ and 0 otherwise, where $c(u) = \{u_1, u_2, ..., u_n\}$ are the offspring of u.

Proposition 4.4.2. Let (θ, q) be a generalized n-adic voting scheme. Let $\mathbb{V}_q$ be the associated voting procedure on X_t. Then $u(t,x) := \mathbb{E}_x[\mathbb{V}_q(t)] = \mathbb{P}_x(\mathbb{V}_q(t) = 1)$ solves the PDE

$$\frac{\partial u}{\partial t} = \frac{1}{2}\frac{\partial^2 u}{\partial x^2} + \lambda\left(\mathbb{E}[\theta(V_1(u), ..., V_n(u))] - u\right) \qquad u(0,x) = q(x),$$

where the $V_k(u(t,x))$ are i.i.d Bernoulli$(u(t,x))$ random variables.

Proof. The proof is identical to that of Proposition 4.2.3, once we observe that

$$\mathbb{E}_x[\mathbb{V}_q(t)|\ell = s] = \mathbb{E}_x[\theta\left(\mathbb{V}_q^{(1)}(t), ..., \mathbb{V}_q^{(n)}(t)\right)|\ell = s]$$

still holds for all $0 \leq s \leq t$, where $\ell = \ell_\emptyset$ is the lifespan of $\emptyset$. □

Lemma 4.4.3. There is a 1-1 correspondence between functions of the form $G(u) = \mathbb{E}[\theta(V_1(u), ..., V_n(u))]$, where $\theta : \{0,1\}^n \to [0,1]$ and the $V_k(u)$ are i.i.d Bernoulli(u) random variables, and polynomials of the form $G(u) = \sum_{k=0}^n a_k u^k (1-u)^{n-k}$, where $0 \leq a_k \leq \binom{n}{k}$.

Proof. For each $0 \leq k \leq n$, there exist $\binom{n}{k}$ distinct partitions

$$(\{r_{1,j}, r_{2,j}, ..., r_{k,j}\}, \{r_{k+1,j}, ..., r_{n,j}\}) \qquad \left(j = 1, ..., \binom{n}{k}\right)$$

of $\{1, 2., , , n\}$ into two sets of size k and $n-k$. If $\theta : \{0,1\}^n \to [0,1]$ then there exists $a_{k,j} \in [0,1]$ such that

$$\theta(x_1, ..., x_n) = \sum_{k=0}^{n} \sum_{j=1}^{\binom{n}{k}} a_{k,j} \prod_{i=1}^{k} x_{r_{i,j}} \prod_{i=k+1}^{n} (1 - x_{r_{i,j}}).$$

By linearity of expectation and independence of the $V_k(u)$, we have

$$\begin{aligned}\mathbb{E}[\theta(V_1(u), ..., V_n(u))] &= \sum_{k=0}^{n} \sum_{j=1}^{\binom{n}{k}} a_{k,j} \prod_{i=1}^{k} \mathbb{E}[V_{r_{i,j}}(u)] \prod_{i=k+1}^{n} (1 - \mathbb{E}[V_{r_{i,j}}(u)]) \\ &= \sum_{k=0}^{n} \sum_{j=1}^{\binom{n}{k}} a_{k,j} u^k (1-u)^{n-k},\end{aligned}$$

with $0 \leq \sum_{j=1}^{\binom{n}{k}} a_{k,j} \leq \binom{n}{k}$ for $k = 0, ..., n$. Conversely if $G(u) = \sum_{k=0}^n a_k u^k (1-u)^{n-k}$ with $0 \leq a_k \leq \binom{n}{k}$, then take $a_{k,j} = \frac{a_k}{\binom{n}{k}}$ for each $j = 0, ..., \binom{n}{k}$ and define

$$\theta(x_1, ...x_n) = \sum_{k=0}^{n} \sum_{j=1}^{\binom{n}{k}} a_{k,j} \prod_{i=1}^{k} x_{r_{i,j}} \prod_{i=k+1}^{n} (1 - x_{r_{i,j}}).$$

By the same argument as above, $\mathbb{E}[\theta(V_1(u), ..., V_n(u))] = G(u)$, where the $V_k(u)$ are i.i.d Bernoulli(u) random variables. $\square$

Corollary 4.4.4. The class of functions which have generalized n-adic voting scheme representations is precisely the polynomials of the form

$$F(u) = \lambda \left(\sum_{k=0}^{n} a_k u^k (1-u)^{n-k} - u \right),$$

where $\lambda > 0$, $n \in \mathbb{N}$, and $0 \leq a_k \leq \binom{n}{k}$.

It is clear from Corollary 4.3.5 that this class of polynomials contains those with n-adic voting measure representations, and that the coefficients have fewer restrictions. We now demonstrate just how large this class of functions is.

THEOREM 4.4.5. Let $n \in \mathbb{N}$. For $k = 0, ..., n$, let $a_k \in \mathbb{R}$ with $a_0 \geq 0$ and $a_n \leq 0$. Then

$$F(u) = \sum_{k=1}^{n} a_k u^k (1-u)^{n-k}$$

has a generalized n-adic voting scheme representation.

Proof. Pick $a > 0$ large enough so that

$$0 \leq \frac{a_0}{a} \leq 1, \qquad -1 \leq \frac{a_n}{a} \leq 0, \qquad -\binom{n-1}{k-1} \leq \frac{a_k}{a} \leq \binom{n-1}{k} \text{ for } 1 \leq k \leq n-1$$

all hold. Then, using that $1 = u + 1 - u$,

$$\begin{aligned}
F(u) &= a\left(\sum_{k=0}^{n} \frac{a_k}{a} u^k (1-u)^{n-k} + u - u\right) \\
&= a\left(\sum_{k=0}^{n} \frac{a_k}{a} u^k (1-u)^{n-k} + u(u+1-u)^{n-1} - u\right) \\
&= a\left(\sum_{k=0}^{n} \frac{a_k}{a} u^k (1-u)^{n-k} + u\sum_{k=0}^{n-1} \binom{n-1}{k} u^k (1-u)^{n-1-k} - u\right) \\
&= a\left(\sum_{k=0}^{n} \frac{a_k}{a} u^k (1-u)^{n-k} + \sum_{k=1}^{n} \binom{n-1}{k-1} u^k (1-u)^{n-k} - u\right) \\
&= a\left(\frac{a_0}{a}(1-u)^n + \sum_{k=1}^{n} \left(\frac{a_k}{a} + \binom{n-1}{k-1}\right) u^k (1-u)^{n-k} - u\right).
\end{aligned}$$

Write $b_0 = \frac{a_0}{a}$, $b_k = \frac{a_k}{a} + \binom{n-1}{k-1}$ for $k = 1, ..., n$. Then

$$F(u) = a\left(\sum_{k=0}^{n} b_k u^k (1-u)^{n-k} - u\right).$$

Note that $0 \leq b_0 \leq 1$, $0 \leq b_n \leq 1$, and for $k = 1, ..., n-1$, we have

$$0 \leq b_k \leq \binom{n-1}{k} + \binom{n-1}{k-1} = \binom{n}{k}.$$

Therefore, $0 \leq b_k \leq \binom{n}{k}$ for $k = 0, ..., n$. Hence, by Corollary 4.4.4, $F(u)$ has a generalized voting scheme representation. □

We have the following equivalent characterization.

Corollary 4.4.6. Let $P(u)$ be any real polynomial with $P(0) \geq 0$ and $P(1) \leq 0$. Then $P(u)$ has a generalized n-adic voting scheme representation.

Proof. By induction on the degree of P, it is clear that certainly P can be expressed in the form

$$P(u) = \sum_{k=0}^{n} a_k u^k (1-u)^{n-k},$$

for some $a_k \in \mathbb{R}$. Then $a_0 = P(0) \geq 0$ and $a_n = P(1) \leq 0$, and we appeal to Theorem 4.4.5. ☐

Remark 4.4.7. By considering the transformation $u \to 1-u$, the same result holds for polynomials $P(u)$ with $P(0) \leq 0$ and $P(1) \geq 0$.

4.5 Voting with Arbitrary Offspring Distributions

Naturally, we can also consider voting schemes on general Branching Brownian Motions. Let $X_t = (X_u(t) : u \in \mathcal{N}_t)$ be a Branching Brownian Motion with branching rate λ and offspring distribution μ, where $\mu(0) = 0$.

Definition 4.5.1. A *voting scheme* is a pair (Θ, q), where $\Theta = (\theta_n)_{n\in\mathbb{N}}$ is a sequence of functions $\theta_n : \{0,1\}^n \to \{0,1\}$, and $q : \mathbb{R} \to [0,1]$ is measurable.

The associated voting procedure $\mathbb{V}_q$ on X_t is constructed as with n-adic voting schemes, but with each $u \in p(\mathcal{N}^r)$ casting its vote as

$$\mathbb{V}_q^u(t) = \theta_n \left(\mathbb{V}_q^{u_1}(t), \mathbb{V}_q^{u_2}(t), .., \mathbb{V}_q^{u_n}(t)\right)$$

where $c(u) = \{u_1, ..., u_n\}$ are the offspring of u (n will vary according to μ).

Proposition 4.5.2. Let (Θ, q) be a voting scheme. Let $\mathbb{V}_q$ be the associated voting procedure on X_t. Then $u(t,x) := \mathbb{E}_x[\mathbb{V}_q(t)] = \mathbb{P}_x(\mathbb{V}_q(t) = 1)$ solves the PDE

$$\frac{\partial u}{\partial t} = \frac{1}{2}\frac{\partial^2 u}{\partial x^2} + \lambda \left(\sum_{k=1}^{\infty} \mu(k)\mathbb{E}[\theta_k(V_1(u), ..., V_k(u))] - u\right) \qquad u(0,x) = q(x),$$

where $\Theta = (\theta_k)_{k\in\mathbb{N}}$ and the $V_i(u(t,x))$ are i.i.d Bernoulli$(u(t,x))$ random variables.

Proof. By precisely the same arguments as in Theorem 3.2.2 and Proposition 4.2.3,

$$u(t,x) = e^{-\lambda t}\int_{-\infty}^{\infty} \frac{e^{-\frac{z^2}{2t}}}{\sqrt{2\pi t}} q(x-z)dz + \sum_{k=1}^{\infty}\mu(k)\left(\int_0^t \lambda e^{-\lambda s}\int_{-\infty}^{\infty}\frac{e^{-\frac{z^2}{2s}}}{\sqrt{2\pi s}}\mathbb{E}[\theta_k(V_1, ..., V_k)u(t-s, x-z)]dzds\right)$$

and the standard analytic argument gives

$$\frac{\partial u}{\partial t} = \frac{1}{2}\frac{\partial^2 u}{\partial x^2} + \lambda\left(\sum_{k=1}^{\infty} \mu(k)\mathbb{E}[\theta_k(V_1(u), ..., V_k(u))] - u\right).$$

Clearly $u(0, x) = q(x)$, as required. □

At first glance, it might appear that Proposition 4.5.2 has provided us with a fresh batch of functions with voting scheme representations. Actually, our study of n-adic voting measures has already familiarized us with a large number of these functions.

Definition 4.5.3. An offspring distribution μ is *bounded* if there exists $n \in \mathbb{N}$ such that $\mu(k) = 0$ for all $k > n$.

We have the following duality result.

THEOREM 4.5.4. Let $F(u)$ be a function. F has a voting scheme representation on a Branching Brownian Motion with bounded offspring distribution if and only if it has an $n-$adic voting measure representation, for some $n \in \mathbb{N}$.

Proof. Suppose that F has a voting scheme representation on a Branching Brownian Motion X_t, with bounded offspring distribution μ and branching rate $\lambda > 0$. Let $n = \max\{k \in \mathbb{N} : \mu(k) \neq 0\}$. We will show that the voting scheme can be embedded into an n-adic voting measure. By Proposition 4.5.2, there exist functions $\theta_k : \{0, 1\}^k \to \{0, 1\}$ such that

$$F(u) = \lambda\left(\sum_{k=1}^{n} \mu(k)\mathbb{E}[\theta_k(V_1(u), ..., V_k(u))] - u\right),$$

where the $V_i(u)$ are i.i.d Bernoulli(u) random variables. By Lemma 4.2.6, for each k, there exist $a_{0,k}, ..., a_{k,k} \in \mathbb{N}$ with $0 \leq a_{i,k} \leq \binom{k}{i}$, such that

$$\mathbb{E}[\theta_k(V_1(u), ..., V_k(u))] = \sum_{i=0}^{k} a_{i,k} u^i (1 - u)^{k-i}.$$

Applying our familiar trick $1 = u + 1 - u$, we have

$$\begin{aligned}
\mathbb{E}[\theta_k(V_1(u), ..., V_k(u))] &= \sum_{i=0}^{k} a_{i,k} u^i (1-u)^{k-i} (u+1-u)^{n-k} \\
&= \sum_{i=0}^{k} a_{i,k} u^i (1-u)^{k-i} \sum_{j=0}^{n-k} \binom{n-k}{j} u^j (1-u)^{n-k-j} \\
&= \sum_{i=0}^{k} \sum_{j=0}^{n-k} a_{i,k} \binom{n-k}{j} u^{i+j} (1-u)^{n-i-j} \\
&= \sum_{i=0}^{n} \sum_{j=0}^{i} a_{j,k} \binom{n-k}{i-j} u^i (1-u)^{n-i}.
\end{aligned}$$

Then

$$F(u) = \lambda \left(\sum_{k=1}^{n} \mu(k) \sum_{i=0}^{n} \left(\sum_{j=0}^{i} a_{j,k} \binom{n-k}{i-j} u^i (1-u)^{n-i} \right) - u \right).$$

Now note that for $i = 0, ..., n$,

$$0 \leq \sum_{j=0}^{i} a_{j,k} \binom{n-k}{i-j} \leq \sum_{j=0}^{i} \binom{k}{j} \binom{n-k}{i-j} = \binom{n}{i}$$

by Vandermonde's identity[2]. Therefore, by Corollary 4.3.5, $F(u)$ has an n-adic voting measure representation. Conversely, if $F(u)$ has an n-adic voting measure representation, then by Corollary 4.3.5,

$$F(u) = \lambda \left(\sum_{r=0}^{l} p_r \sum_{k=0}^{n} a_{k,r} u^k (1-u)^{n-k} - u \right),$$

where $\lambda > 0$, $p_r \geq 0$ with $\sum_{r=0}^{l} p_r = 1$, and for each r, $a_{k,r} \in \mathbb{N}$ with $0 \leq a_{k,r} \leq \binom{n}{k}$. For each r, by Lemma 4.2.8, there exist $b_{0,r}, b_{1,r}, ..., b_{n+r,r} \in \mathbb{N}$ with $0 \leq b_{k,r} \leq \binom{n+r}{k}$, such that

$$\sum_{k=0}^{n} a_{k,r} u^k (1-u)^{n-k} = \sum_{k=0}^{n+r} b_{k,r} u^k (1-u)^{n+r-k},$$

and by Lemma 4.2.6 there exists a function $\theta_{n+r} : \{0,1\}^{n+r} \to \{0,1\}$ such that

$$\sum_{k=0}^{n+r} b_{k,r} u^k (1-u)^{n+r-k} = \mathbb{E}[\theta_{n+r}(V_1(u), ..., V_{n+r}(u))],$$

where the $V_k(u)$ are i.i.d Bernoulli(u) random variables. Define $\mu : \mathbb{N} \to [0,1]$ by

$$\mu(k) := \begin{cases} p_r & \text{if } k = n + r \text{ for some } 0 \leq r \leq l \\ 0 & \text{otherwise.} \end{cases}$$

[2]See Appendix.

Then μ is clearly a bounded offspring distribution, $\mu(0) = 0$, and

$$F(u) = \lambda \left(\sum_{k=1}^{\infty} \mu(k) \mathbb{E}[\theta_k(V_1(u), ..., V_k(u))] - u \right).$$

Therefore, F has a voting scheme representation on a Branching Brownian Motion with offspring distribution μ and branching rate λ. $\square$

It follows that introducing voting schemes on arbitrary Branching Brownian Motions is only additive to our study if the offspring distribution is unbounded. It is worth noting that voting schemes readily generalize to the n-dimensional case, in the natural way.

5 The Maximal Process

5.1 Motivation

Of course, Branching Brownian Motion is far more than just a tool for solving PDEs; it is an interesting mathematical object in its own right. We now demonstrate how PDE theory can help us to better understand some of its subtle properties.

Definition 5.1.1. The *maximal process* of a Branching Brownian Motion $X_t = (X_u(t) : u \in \mathcal{N}_t)$ is $M(t) := \max_{u\in\mathcal{N}_t} X_u(t)$.

Definition 5.1.2. The *Heaviside function* is the function $H : \mathbb{R} \to \mathbb{R}$ defined by $H(x) = \mathbb{1}_{[0,\infty)}(x)$.

As McKean ([McK75]) remarks, there is a close relationship between the maximal process and the F-KPP equation:

PROPOSITION 5.1.3. Let $X_t = (X_u(t) : u \in \mathcal{N}_t)$ be a Branching Brownian Motion with branching rate $\lambda > 0$ and offspring distribution $\mu(k) = p_k$. Let $\Phi(x) = \sum_{k=0}^{\infty} p_k x^k$. Then $u(t,x) = \mathbb{P}(M(t) \le x)$ solves

$$\frac{\partial u}{\partial t} = \frac{1}{2}\frac{\partial^2 u}{\partial x^2} + \lambda(\Phi(u) - u), \qquad u(0,x) = H(x).$$

Proof. We simply note that by Theorem 3.2.2, our solution is given by

$$\begin{aligned}
\mathbb{E}\left[\prod_{u\in\mathcal{N}_t} H(x - X_u(t))\right] &= \mathbb{E}\left[\prod_{u\in\mathcal{N}_t} \mathbb{1}_{\{x - X_u(t) \ge 0\}}\right] \\
&= \mathbb{E}\left[\mathbb{1}_{\bigcap_{u\in\mathcal{N}_t}\{X_u(t)\le x\}}\right] \\
&= \mathbb{E}\left[\mathbb{1}_{\{M(t)\le x\}}\right] \\
&= \mathbb{P}(M(t) \le x).
\end{aligned}$$

$\square$

This connection forms a basis for studying the maximal process through the lens of PDE theory.

5.2 Travelling Wave Solutions to F-KPP

One of the most interesting properties of the F-KPP equation is that it admits travelling wave solutions, which exhibit particularly nice behavior. Indeed, Kolmogorov-Petrovskii-Piskunov, Fife-McLeod and many more besides, were all chiefly concerned with the existence and behaviour of travelling wave solutions. It will emerge that this property is incredibly useful in studying the maximal process.

Definition 5.2.1. A solution $u(t,x)$ to the F-KPP equation is a *travelling wave solution of speed* $c > 0$ if $u(t,x) = U(x-ct)$, for some function $U : \mathbb{R} \to \mathbb{R}$.

The following theorem is implicit in [KPP37]. Our proof generalizes the strategy deployed in [McK75], which only considers the case $\Phi(x) = x^2$.

THEOREM 5.2.2. Let $p_k \geq 0$ with $\sum_{k=2}^{\infty} p_k = 1$ and $\gamma := \sum_{k=2}^{\infty} kp_k < \infty$ Write $\Phi(x) = \sum_{k=2}^{\infty} p_k x^k$. Then the semilinear PDE

$$\frac{\partial u}{\partial t} = \frac{1}{2}\frac{\partial^2 u}{\partial x^2} + \lambda(\Phi(u) - u)$$

admits a travelling wave solution $U : \mathbb{R} \to (0,1)$ of speed $c > 0$, with $\lim_{x\to\infty} U(x) = 1$ and $\lim_{x\to-\infty} U(x) = 0$, if and only if $c \geq \sqrt{2\lambda(\gamma-1)}$. Furthermore, any such solution is unique up to an additive constant, and if $c = \sqrt{2\lambda(\gamma-1)}$ then U is strictly increasing with bounded derivative.

Proof. Let $c > 0$. Note that $u(t,x) = U(x-ct)$ is a travelling wave solution if and only if

$$-cU'(x) = \frac{1}{2}U''(x) + \lambda(\Phi(U(x)) - U(x))$$

for all $x \in \mathbb{R}$. Letting $\xi = U$, $\eta = U'$, we have a plane autonomous system of ODEs:

$$\begin{aligned}\dot{\xi} &= \eta \\ \dot{\eta} &= -2c\eta - 2\lambda(\Phi(\xi) - \xi)\end{aligned}$$

Our proof is now a simple phase plane analysis. In the (ξ,η)-plane, we have critical points[1] at $(0,0)$ and $(1,0)$. We classify these as follows:

<u>At $(0,0)$</u>:

For small ξ,

$$\Phi(\xi) - \xi = \sum_{k=2}^{\infty} p_k\xi^k - \xi = -\xi + o(\xi),$$

[1]For a concise exposition on the classification of critical points, consult [Col06] (*pp.*327).

so linearizing gives $\begin{pmatrix}\dot{\xi}\\ \dot{\eta}\end{pmatrix} = \begin{pmatrix}0 & 1\\ 2\lambda & -2c\end{pmatrix}\begin{pmatrix}\xi\\ \eta\end{pmatrix}$. Then

$$\begin{vmatrix}-\kappa & 1\\ 2\lambda & -2c-\kappa\end{vmatrix} = \kappa(2c+\kappa) - 2\lambda = \kappa^2 + 2c\kappa - 2\lambda$$

has zeroes $\kappa = -c \pm \sqrt{c^2 + 2\lambda}$, which are of different sign, and hence $(0,0)$ is a saddle.
<u>At $(1,0)$</u>:
Let $\xi = 1 - \zeta$, so that

$$\begin{aligned}\dot{\zeta} &= -\eta\\ \dot{\eta} &= -2c\eta - 2\lambda(\Phi(1-\zeta) - (1-\zeta)).\end{aligned}$$

For small ζ, we have

$$\begin{aligned}\Phi(1-\zeta) - (1-\zeta) &= \sum_{k=2}^{\infty} p_k(1-\zeta)^k - (1-\zeta)\\ &= \sum_{k=2}^{\infty} p_k - \zeta\sum_{k=2}^{\infty} kp_k - (1-\zeta) + o(\zeta)\\ &= -(\gamma-1)\zeta + o(\zeta),\end{aligned}$$

so linearizing gives $\begin{pmatrix}\dot{\zeta}\\ \dot{\eta}\end{pmatrix} = \begin{pmatrix}0 & -1\\ 2\lambda(\gamma-1) & -2c\end{pmatrix}\begin{pmatrix}\zeta\\ \eta\end{pmatrix}$. Then

$$\begin{vmatrix}-\kappa & -1\\ 2\lambda(\gamma-1) & -2c-\kappa\end{vmatrix} = \kappa(2c+\kappa) + 2\lambda(\gamma-1) = \kappa^2 + 2c\kappa + 2\lambda(\gamma-1)$$

has zeroes $\kappa = -c \pm \sqrt{c^2 - 2\lambda(\gamma-1)}$.
If $c < \sqrt{2\lambda(\gamma-1)}$ then both zeroes have non-trivial complex parts and strictly negative real part, and the point $(1,0)$ (in the (ξ,η)-plane) is a stable spiral. But then for our corresponding solution U, we cannot have $0 < U < 1$.
If $c > \sqrt{2\lambda(\gamma-1)}$, then we have two strictly negative zeroes, and $(1,0)$ is a stable node. If $c = \sqrt{2\lambda(\gamma-1)}$, then we have only one (strictly negative) zero, and $(1,0)$ is a stable star. In both cases we have a solution U satisfying $\lim_{x\to\infty} U(x) = 1$ and $\lim_{x\to-\infty} U(x) = 0$. For uniqueness, we appeal to the Picard-Lindelöf Theorem[2].
Define

$$\begin{aligned}f_1(\xi,\eta) &:= \dot{\xi} = \eta\\ f_2(\xi,\eta) &:= \dot{\eta} = -2c\eta - 2\lambda(\Phi(\xi) - \xi),\end{aligned}$$

[2]See Appendix.

and write $\mathbf{f}=\begin{pmatrix} f_1 \\ f_2 \end{pmatrix}$. We claim that $\mathbf{f}$ is globally Lipschitz. Let $\underline{\mathbf{x}}, \underline{\mathbf{y}} \in [0,1] \times \mathbb{R}$. Then

$$\begin{aligned}
\left\|\mathbf{f}(\underline{\mathbf{x}}) - \mathbf{f}(\underline{\mathbf{y}})\right\|_1 &= |x_2 - y_2| + |2c(y_2 - x_2) + 2\lambda(\Phi(y_1) - \Phi(x_1) + x_1 - y_1)| \\
&\le (1+2c)|x_2 - y_2| + 2\lambda|x_1 - y_1| + 2\lambda|\sum_{k=0}^{\infty} p_k(y_1^k - x_1^k)| \\
&\le (1+2c)|x_2 - y_2| + 2\lambda|x_1 - y_1| + 2\lambda\sum_{k=0}^{\infty} p_k|x_1 - y_1| \\
&= (1+2c)|x_2 - y_2| + 4\lambda|x_1 - y_1| \\
&\le (1+2c+4\lambda)\left\|\underline{\mathbf{x}} - \underline{\mathbf{y}}\right\|_1
\end{aligned}$$

Hence $\mathbf{f}$ is globally Lipschitz, which gives uniqueness up to an additive constant. Finally, suppose that $c = \sqrt{2\lambda(\gamma-1)}$, and consider the trajectory in the $(\xi,\eta)-$plane corresponding to U. At (0,0), its motion is governed by the eigenvalue

$$\kappa = -c + \sqrt{c^2 + 2\lambda} = \sqrt{2\lambda\gamma} - \sqrt{2\lambda(\gamma-1)},$$

and its associated eigenvector $\begin{pmatrix} 1 \\ \sqrt{2\lambda\gamma} - \sqrt{2\lambda(\gamma-1)} \end{pmatrix}$. Therefore the trajectory begins in the first quadrant of the $(\xi,\eta)-$ plane, and from a sketch of the phase plane it is clear that it cannot then pass below the $\xi-$axis. It follows that U is strictly increasing with bounded derivative.

□

For the remainder of this chapter, we only consider the following F-KPP equation:

$$\frac{\partial u}{\partial t} = \frac{1}{2}\frac{\partial^2 u}{\partial x^2} + u(u-1), \qquad\qquad u(0,x) = H(x).$$

Accordingly, we write $M(t) = \max_{u\in\mathcal{N}_t} X_u(t)$ for the maximal process of a dyadic Branching Brownian Motion $X_t = (X_u(t) : u \in \mathcal{N}_t)$ with branching rate 1.

Since $\gamma = 2$, $m = 1$, there exist travelling wave solutions of speed c satisfying the conditions of Theorem 5.2.2 whenever $c \ge \sqrt{2}$. The limiting case $c = \sqrt{2}$ is of particular interest. The remainder of this section is devoted to proving the following theorem.

Theorem 5.2.3 ([KPP37]). Let m_t be the median of the maximal process $M(t)$, so that $u(t, m_t) = \frac{1}{2}$. Then $u(t, x + m_t)$ converges uniformly to a travelling wave solution $w(x)$ with speed $\sqrt{2}$.

We loosely follow [Bra82], with added rigour. Since $0 \leq u \leq 1$, for pointwise convergence it is enough to show that $u(t, x + m_t)$ is monotone in t for each fixed $x \in \mathbb{R}$. We will prove this using the Extended Maximum Principle, which applies to a far more general class of semilinear heat equations. It is essentially due to McKean ([McK75]).

Proposition 5.2.4 (Extended Maximum Principle). Let $F : \mathbb{R} \to \mathbb{R}$ be such that $F'(u)$ is bounded and continuous, and suppose that $u_1(t, x)$ and $u_2(t, x)$ solve

$$\frac{\partial u}{\partial t} = \frac{1}{2}\frac{\partial^2 u}{\partial x^2} + F(u).$$

Suppose further that whenever $x_1 < x_2$, we have that $u_1(0, x_1) < u_2(0, x_1)$ implies $u_1(0, x_2) \leq u_2(0, x_2)$. Then for all $t > 0$, and $x_1 < x_2$, we have that

1. $u_1(t, x_1) \leq u_2(t, x_1)$ implies $u_1(t, x_2) \leq u_2(t, x_2)$.
2. $u_1(t, x_1) < u_2(t, x_1)$ implies $u_1(t, x_2) < u_2(t, x_2)$.

Proof. Let $v(t, x) := u_2(t, x) - u_1(t, x)$. Then for all $x_1 < x_2$, we have that $v(0, x_1) > 0$ implies $v(0, x_2) \geq 0$. By considering the contrapositive of 1, and noting the symmetry in the ensuing proof, we shall see that it suffices to show that for all $t > 0$ and $x_1 < x_2$, $v(t, x_1) > 0$ implies $v(t, x_2) > 0$. Now

$$\frac{\partial v}{\partial t} = \frac{1}{2}\frac{\partial^2 v}{\partial x^2} + F(u_2) - F(u_1),$$

and by the Mean Value Theorem there exists a continuous function $\theta : \mathbb{R} \to (0, 1)$ such that

$$F(u_2) - F(u_1) = F'(u_1 + \theta(u_2 - u_1))(u_2 - u_1) = F'(u_1 + \theta(u_2 - u_1))v.$$

Then $k(t, x) := F'(u_1(t, x) + \theta(u_2(t, x) - u_1(t, x)))$ is bounded and continuous, and we have

$$\frac{\partial v}{\partial t} = \frac{1}{2}\frac{\partial^2 v}{\partial x^2} + k(t, x)v.$$

Now fix $t_0 > 0$ and $x_1 < x_2$. Suppose that $v(t_0, x_1) > 0$. We claim that $v(t_0, x_2) > 0$. By Feynman Kăc,

$$M_s^x := v(t_0 - s, B_s^x)e^{\int_0^s k(t_0 - r, B_r^x)dr}$$

defines a bounded continuous martingale on $[0, t_0)$ for all $x \in \mathbb{R}$, and furthermore

$$v(t_0, x) = \mathbb{E}_x\left[v(0, B_{t_0})e^{\int_0^{t_0} k(t_0 - r, B_r)dr}\right],$$

where B is a standard Brownian Motion. Let $0 \leq \tau \leq t_0$ be a stopping time. Then by the Optional Stopping Theorem, $\mathbb{E}[M_\tau^x] = \mathbb{E}[M_{t_0}^x]$ for all $x \in \mathbb{R}$, so

$$v(t_0, x) = \mathbb{E}_x \left[v(t_0 - \tau, B_\tau) e^{\int_0^\tau k(t_0 - r, B_r) dr} \right] \tag{5.1}$$

holds for any stopping time $0 \leq \tau \leq t_0$, $x \in \mathbb{R}$.
First let $x = x_1$, and

$$\tau = \inf_{0 \leq s \leq t_0} \{ s : M_s^{x_1} = 0 \} \wedge t_0 = \inf_{0 \leq s \leq t_0} \{ s : v(t_0 - s, B_s^{x_1}) = 0 \} \wedge t_0.$$

in Equation (5.1). Then

$$0 < v(t_0, x_1) = \mathbb{E}_{x_1} \left[v(t_0 - \tau, B_\tau) e^{\int_0^\tau k(t_0 - r, B_r) dr} \right],$$

so that

$$0 < \mathbb{P}_{x_1} \left[v(t_0 - \tau, B_\tau) e^{\int_0^\tau k(t_0 - r, B_r) dr} > 0 \right] = \mathbb{P}[v(t_0 - \tau, B_\tau) > 0] = \mathbb{P}[\tau = t_0].$$

Therefore, by almost sure continuity of Brownian Motion, there exists $\omega \in \Omega$ such that $\gamma(s) := B_s^{x_1}(\omega)$ defines a (continuous) curve $[0, t_0] \to \mathbb{R}$, with $\gamma(0) = x_1$ and $v(t_0 - s, \gamma(s)) > 0$ for all $0 \leq s \leq t_0$. Now let $x = x_2$ and

$$\tau = \inf_{0 \leq s \leq t_0} \{ s : B_s^{x_2} = \gamma(s) \} \wedge t_0.$$

in Equation (5.1). Define $\psi : [0, t_0] \to \mathbb{R}$ by $\psi(s) = B_s^{x_2} - \gamma(s)$, so that ψ is almost surely continuous with

$$\begin{aligned} \psi(0) &= x_2 - x_1 > 0, \\ \psi(t_0) &= B_{t_0}^{x_2} - \gamma(t_0). \end{aligned}$$

By almost sure continuity and an intermediate value argument, we have

$$\mathbb{P}[\tau < t_0] \geq \mathbb{P}[\psi(t_0) < 0] = \mathbb{P}_{x_2}[B_{t_0} < \gamma(t_0)] > 0.$$

Now

$$\begin{aligned} v(t_0, x_2) &= \mathbb{E}_{x_2} \left[v(t_0 - \tau, B_\tau) e^{\int_0^\tau k(t_0 - r, B_r) dr} \right] \\ &= \mathbb{E}_{x_2} \left[v(t_0 - \tau, \gamma(\tau)) e^{\int_0^\tau k(t_0 - r, B_r) dr} \mathbb{1}_{\{\tau < t_0\}} \right] + \mathbb{E}_{x_2} \left[v(0, B_{t_0}) e^{\int_0^{t_0} k(t_0 - r, B_r) dr} \mathbb{1}_{\{\tau = t_0\}} \right], \end{aligned}$$

and $v(t_0 - \tau, \gamma(\tau)) > 0$ on $\{\tau < t_0\}$, which has strictly positive measure, so the first term is strictly positive. For the second term, note that the inclusions

$$\{\tau = t_0\} \subseteq \{B_s^{x_2} > \gamma(s) \, \forall s \in [0, t_0]\} \subseteq \{B_{t_0}^{x_2} > \gamma(t_0)\}$$

hold almost surely by the almost sure continuity of Brownian Motion. But now $v(0, \gamma(t_0)) > 0$ by construction, so on $\{B_{t_0}^{x_2} > \gamma(t_0)\}$ we have, by our original assumption on v, that $v(0, B_{t_0}^{x_2}) \geq 0$. In particular, $v(0, B_{t_0}^{x_2}) \geq 0$ almost surely on $\{\tau = t_0\}$, so the second term is positive. Hence $v(t_0, x_2) > 0$, as required. □

Let $u(t,x) = \mathbb{P}(M(t) \le x)$, so that u solves

$$\frac{\partial u}{\partial t} = \frac{1}{2}\frac{\partial^2 u}{\partial x^2} + u(u-1) \qquad\qquad u(0,x) = H(x).$$

For each $\varepsilon \in (0,1)$, there exists a unique function $r_\varepsilon : (0,\infty) \to \mathbb{R}$ satisfying $u(t, r_\varepsilon(t)) = \varepsilon$.

COROLLARY 5.2.5. Fix $\varepsilon \in (0,1)$. Then

$$u(t, x + r_\varepsilon(t)) \begin{cases} \uparrow \text{ as } t \to \infty & \text{if } x < 0 \\ \downarrow \text{ as } t \to \infty & \text{if } x > 0 \end{cases}.$$

Furthermore, $\frac{\partial u}{\partial x}(t, r_\varepsilon(t))$ is decreasing in t.

Proof. Fix $t_0 > 0$, $a > 0$, and define

$$\begin{aligned} u_1(t,x) &= u(t+a, x + r_\varepsilon(t_0+a)), \\ u_2(t,x) &= u(t, x + r_\varepsilon(t_0)). \end{aligned}$$

Then $u_1(t,x)$ and $u_2(t,x)$ are both solve

$$\frac{\partial u}{\partial t} = \frac{1}{2}\frac{\partial^2 u}{\partial x^2} + u(u-1).$$

Clearly $F(u) = u(u-1)$ satisfies the conditions of the Extended Maximum Principle. We simply exploit the fact that $u(0,x) = H(x)$. Let $x_1 < x_2$ and suppose that $u_1(0,x_1) < u_2(0,x_1)$. Then

$$0 \le u(a, x_1 + r_\varepsilon(t_0+a)) < u(0, x_1 + r_\varepsilon(t_0)),$$

so $x_1 + r_\varepsilon(t_0) \ge 0$. Hence $x_2 + r_\varepsilon(t_0) > 0$, which gives $u_2(0,x_2) = u(0, x_2 + r_\varepsilon(t_0)) = 1$, so certainly $u_1(0,x_2) \le u_2(0,x_2)$. Therefore the Extended Maximum Principle tells us that whenever $x_1 < x_2$, we have

1. $u_1(t_0,x_1) \le u_2(t_0,x_1)$ implies $u_1(t_0,x_2) \le u_2(t_0,x_2)$,

2. $u_1(t_0,x_2) \ge u_2(t_0,x_2)$ implies $u_1(t_0,x_1) \ge u_2(t_0,x_1)$.

Note that $u_1(t_0,0) = \varepsilon = u_2(t_0,0)$. For $x > 0$, taking $x_1 = 0, x_2 = x$ in 1. gives $u_1(t_0,x) \le u_2(t_0,x)$. For $x < 0$, taking $x_1 = x, x_2 = 0$ in 2. gives $u_1(t_0,x) \ge u_2(t_0,x)$. Hence

$$u(t_0 + a, x + r_\varepsilon(t_0+a)) - u(t_0, x + r_\varepsilon(t_0)) = u_1(t_0,x) - u_2(t_0,x)$$

is negative for $x > 0$ and positive for $x < 0$. Since t_0 and a were arbitrary it follows that

$$u(t, x + r_\varepsilon(t)) \begin{cases} \uparrow \text{ as } t \to \infty & \text{if } x < 0 \\ \downarrow \text{ as } t \to \infty & \text{if } x > 0 \end{cases}.$$

Because

$$u(t_0 + a, r_\varepsilon(t_0 + a)) - u(t_0, r_\varepsilon(t_0)) = 0,$$

it also follows that

$$u(t_0 + a, x + r_\varepsilon(t_0 + a)) - u(t_0, x + r_\varepsilon(t_0)) \le u(t_0 + a, r_\varepsilon(t_0 + a)) - u(t_0, r_\varepsilon(t_0))$$

for all $x > 0$. Rearranging,

$$u(t_0 + a, x + r_\varepsilon(t_0 + a)) - u(t_0 + a, r_\varepsilon(t_0 + a)) \le u(t_0, x + r_\varepsilon(t_0)) - u(t_0, r_\varepsilon(t_0)),$$

and taking $x \downarrow 0$ gives

$$\frac{\partial u}{\partial x}(t_0 + a, r_\varepsilon(t_0 + a)) \le \frac{\partial u}{\partial x}(t_0, r_\varepsilon(t_0)).$$

Again this holds for all $t_0 > 0, a > 0$, so $\frac{\partial u}{\partial x}(t, r_\varepsilon(t))$ is decreasing in t. $\square$

By Theorem 5.2.2, for each $\varepsilon \in (0, 1)$, there exists a unique travelling wave solution $U_\varepsilon : \mathbb{R} \to (0, 1)$ of speed $\sqrt{2}$ satisfying $\lim_{x\to\infty} U_\varepsilon(x) = 1$, $\lim_{x\to-\infty} U_\varepsilon(x) = 0$ and $U_\varepsilon(0) = \varepsilon$. Furthermore, each U_ε is strictly increasing with bounded derivative.

Corollary 5.2.6. Let $\varepsilon \in (0, 1)$. Then

$$u(t, x + r_\varepsilon(t)) \begin{cases} \le U_\varepsilon(x) & \text{if } x < 0 \\ \ge U_\varepsilon(x) & \text{if } x > 0 \end{cases}.$$

for all $t > 0$. Furthermore,

$$\frac{\partial u}{\partial x}(t, r_\varepsilon(t)) \ge U'_\varepsilon(0).$$

Proof. We fix $t_0 > 0$ and let

$$\begin{aligned} u_1(t, x) &= U_\varepsilon(x), \\ u_2(t, x) &= u(t, r_\varepsilon(t_0) + x). \end{aligned}$$

Let $x_1 < x_2$ and suppose that $u_1(0, x_1) < u_2(0, x_1)$. Then

$$0 \le U_\varepsilon(x_1) < u(0, r_\varepsilon(t_0) + x_1),$$

so $u(0, r_\varepsilon(t_0) + x_1) = 1$ and hence $r_\varepsilon(t_0) + x_1 \ge 0$. But then $r_\varepsilon(t_0) + x_2 > 0$, so that $u_2(0, x_2) = u(0, r_\varepsilon(t_0) + x_2) = 1$ and hence $u_1(0, x_2) \le u_2(0, x_2)$. Therefore we may apply the Extended Maximum Principle. Note that $u_1(t_0, 0) = \varepsilon = u_2(t_0, 0)$, so for $x > 0$ we have

$$U_\varepsilon(x) = u_1(t_0, x) \leq u_2(t_0, x) = u(t_0, r_\varepsilon(t_0)),$$

and for $x < 0$ we have (using the contrapositive) that

$$U_\varepsilon(x) = u_1(t_0, x) \geq u_2(t_0, x) = u(t_0, r_\varepsilon(t_0)).$$

Since $u(t, r_\varepsilon(t)) = U_\varepsilon(0)$ it also follows immediately that

$$\frac{\partial u}{\partial x}(t, r_\varepsilon(t)) \geq U'_\varepsilon(0).$$

□

Now $m_t = r_{\frac{1}{2}}(t)$, so taking $\varepsilon = \frac{1}{2}$ in Corollary 5.2.5 gives

$$u(t, x + m_t) \begin{cases} \uparrow \text{ as } t \to \infty & \text{if } x < 0 \\ \downarrow \text{ as } t \to \infty & \text{if } x > 0 \end{cases}.$$

and $u(t, m_t) = \frac{1}{2}$ for all $t > 0$. Since $0 \leq u \leq 1$, monotone convergence gives that $w(x) := \lim_{t\to\infty} u(t, x + m_t)$ exists pointwise. The following observation is crucial.

Remark 5.2.7. By Corollary 5.2.5 and Corollary 5.2.6, for fixed $t_0 > 0$,

$$\begin{aligned} \tfrac{1}{2} = U_{\frac{1}{2}}(0) < U_{\frac{1}{2}}(x) \leq w(x) \leq u(t_0, x + m_{t_0}) < 1 \quad \text{for } x > 0, \\ \tfrac{1}{2} = U_{\frac{1}{2}}(0) > U_{\frac{1}{2}}(x) \geq w(x) \geq u(t_0, x + m_{t_0}) > 0 \quad \text{for } x < 0. \end{aligned}$$

It follows that $w : \mathbb{R} \to (0, 1)$, $\lim_{x\to-\infty} w(x) = 0$, $\lim_{x\to\infty} w(x) = 1$, and that $\int_0^x (w(\xi) - \frac{1}{2})d\xi \neq 0$ for all $x \neq 0$.

We now show that the limit $w(x)$ is in fact uniform. We present a detailed adaptation of the proof given in [KPP37].

Proposition 5.2.8. $u(t, x + m_t) \to w(x)$ uniformly in t.

Proof. Fix $t > 0$ and define $r_t : (0, 1) \to \mathbb{R}$ by $r_t(\varepsilon) = r_\varepsilon(t)$. Note that r_t is strictly increasing and hence invertible. Differentiating the expression $\varepsilon = u(t, r_t(\varepsilon))$ gives

$$1 = \frac{\partial u}{\partial x}(t, r_t(\varepsilon))r'_t(\varepsilon),$$

so that

$$\begin{aligned}\int_{\frac{1}{2}}^{u(t,x+m_t)} \left(\tfrac{\partial u}{\partial x}(t, r_t(\varepsilon))\right)^{-1} d\varepsilon &= r_t(u(t, x+m_t)) - r_t(\tfrac{1}{2}) \\ &= r_t(u(t, r_t \circ r_t^{-1}(x+m_t))) - m_t \\ &= r_t(r_t^{-1}(x+m_t)) - m_t \\ &= x.\end{aligned}$$

The integrand is increasing by Corollary 5.2.5, and bounded above by Corollary 5.2.6. Therefore it converges, say to $a(\varepsilon)$. For $x > 0$, we have

$$x = \int_{\frac{1}{2}}^{1} \left(\tfrac{\partial u}{\partial x}(t, r_t(\varepsilon))\right)^{-1} d\varepsilon - \int_{u(t,x+m_t)}^{1} \left(\tfrac{\partial u}{\partial x}(t, r_t(\varepsilon))\right)^{-1} d\varepsilon,$$

with $u(t, x+m_t)$ decreasing in t, and for $x < 0$ we have

$$x = \int_{0}^{u(t,x+m_t)} \left(\tfrac{\partial u}{\partial x}(t, r_t(\varepsilon))\right)^{-1} d\varepsilon - \int_{0}^{\frac{1}{2}} \left(\tfrac{\partial u}{\partial x}(t, r_t(\varepsilon))\right)^{-1} d\varepsilon$$

with $u(t, x+m_t)$ increasing in t. In both cases, the Monotone Convergence Theorem yields

$$x = \int_{\frac{1}{2}}^{w(x)} a(\varepsilon) d\varepsilon.$$

for all $x \in \mathbb{R}$. For $0 < \delta < \frac{1}{2}$, $u(t, x+m_t) \to w(x)$ uniformly on $[\delta, 1-\delta]$. Therefore, by boundedness of $(\frac{\partial u}{\partial x}(t, r_t(\varepsilon)))^{-1}$,

$$u(t, x+m_t) = \int_{\frac{1}{2}}^{u(t,u(t,x+m_t)+m_t)} \left(\tfrac{\partial u}{\partial x}(t, r_t(\varepsilon))\right)^{-1} d\varepsilon$$

converges uniformly on $w^{-1}([\delta, 1-\delta])$ to

$$\int_{\frac{1}{2}}^{w(w(x))} a(\varepsilon) d\varepsilon = w(x),$$

so in fact $u(t, x+m_t) \to w(x)$ uniformly on each $w^{-1}([\delta, 1-\delta])$. Fix $\varepsilon > 0$. Then there exists $N \in \mathbb{N}$ such that $\sup_{x>N} |w(x) - 1| < \frac{\varepsilon}{2}$ and $\sup_{x<-N} |w(x)| < \frac{\varepsilon}{2}$. Using Corollary 5.2.5 and the triangle inequality, $\sup_{|x|>N} |u(t, x+m_t) - w(x)| < \varepsilon$ for all $t > 0$. Now pick $\delta > 0$ such that $[-N, N] \subseteq w^{-1}([\delta, 1-\delta])$. Then there exists $M \in \mathbb{N}$ such that $\sup_{x \in w^{-1}([\delta,1-\delta])} |u(t, x+m_t) - w(x)| < \varepsilon$ for all $t > M$. It follows that $\sup_{x\in\mathbb{R}} |u(t, x+m_t) - w(x)| < \varepsilon$ for all $t > M$, so that $u(t, x+m_t) \to w(x)$ uniformly in t. $\square$

Our next step is to show that w is a travelling wave solution to the F-KPP equation. The following result is due to McKean, [McK75].

Proposition 5.2.9. There exists $c > 0$ such that $w(x)$ satisfies

$$cw' + \frac{1}{2}w'' + w(w-1) = 0.$$

Proof. Where convenient, we shall write $m(t)$ for m_t. Let $v(t,x) = u(t, x+m_t)$, then

$$\begin{aligned}\frac{\partial v}{\partial t}(t,x) &= \frac{\partial u}{\partial t}(t, x+m_t) + \frac{\partial u}{\partial x}(t, x+m_t)m'(t) \\ &= \frac{1}{2}\frac{\partial^2 u}{\partial x^2}(t, x+m_t) + u(t,x+m_t)(u(t,x+m_t)-1) + \frac{\partial u}{\partial x}(t, x+m_t)m'(t) \\ &= \frac{1}{2}\frac{\partial^2 v}{\partial x^2}(t,x) + v(t,x)(v(t,x)-1) + \frac{\partial v}{\partial x}(t,x)m'(t).\end{aligned}$$

Hence

$$\frac{\partial v}{\partial t} = \frac{1}{2}\frac{\partial^2 v}{\partial x^2} + m'\frac{\partial v}{\partial x} + v(v-1). \tag{5.2}$$

We fix $x \in \mathbb{R}$, $t > 0$, and integrate Equation (5.2) three times:

$$\int_t^{t+1}\int_0^x\int_0^\xi \frac{\partial v}{\partial s}(s,\eta)d\eta d\xi ds = \int_t^{t+1}\int_0^x\int_0^\xi \left(\frac{1}{2}\frac{\partial^2 v}{\partial \eta^2} + m'(s)\frac{\partial v}{\partial \eta} + v(v-1)\right)(s,\eta)d\eta d\xi ds.$$

Then take $t \to \infty$. On the LHS, we have

$$\int_t^{t+1}\int_0^x\int_0^\xi \frac{\partial v}{\partial s}(s,\eta)d\eta d\xi ds = \int_0^x\int_0^\xi (v(t+1,\eta) - v(t,\eta))d\eta d\xi,$$

so

$$\lim_{t\to\infty}\int_t^{t+1}\int_0^x\int_0^\xi \frac{\partial v}{\partial s}(s,\eta)d\eta d\xi ds = \int_0^x\int_0^\xi (w(\eta) - w(\eta))d\eta d\xi = 0.$$

On the RHS we appeal to the Mean Value Theorem[3]. The first term gives

$$\begin{aligned}\lim_{t\to\infty}\int_t^{t+1}\int_0^x\int_0^\xi \frac{1}{2}\frac{\partial^2 v}{\partial \eta^2}(s,\eta)d\eta d\xi ds &= \lim_{t\to\infty}\int_0^x\int_0^\xi \frac{1}{2}\frac{\partial^2 v}{\partial \eta^2}(t,\eta)d\eta d\xi \\ &= \lim_{t\to\infty}\int_0^x \frac{1}{2}\left(\frac{\partial v}{\partial \eta}(t,\xi) - \frac{\partial v}{\partial \eta}(t,0)\right)d\xi \\ &= \lim_{t\to\infty}\tfrac{1}{2}\left(v(t,x) - v(t,0) - x\frac{\partial v}{\partial x}(t,0)\right) \\ &= \tfrac{1}{2}\left(w(x) - \tfrac{1}{2} - x\lim_{t\to\infty}\frac{\partial v}{\partial x}(t,0)\right).\end{aligned}$$

[3]Specifically, we apply the Mean Value Theorem to functions of the form $F(t) = \int_0^t f(x)dx$, where $f : \mathbb{R} \to \mathbb{R}$ is continuous. If $\lim_{t\to\infty} f(t)$ exists, then we have $\lim_{t\to\infty}\int_t^{t+1} f(x)dx = \lim_{t\to\infty} f(t)$.

The third term gives

$$\lim_{t\to\infty}\int_t^{t+1}\int_0^x\int_0^\xi v(s,\eta)(v(s,\eta)-1)d\eta d\xi ds = \int_0^x\int_0^\xi w(\eta)(w(\eta)-1)d\eta d\xi.$$

The second term requires a little more effort. We have

$$\int_t^{t+1}\int_0^x\int_0^\xi m'(s)\frac{\partial v}{\partial \eta}(s,\eta)d\eta d\xi ds = \int_t^{t+1} m'(s)\int_0^x (v(s,\xi)-\tfrac{1}{2})d\xi ds.$$

Now $m:(0,\infty)\to(0,\infty)$ is strictly increasing and hence invertible, so we may define $F:(0,\infty)\to\mathbb{R}$ by

$$F(t) = \int_0^{m^{-1}(t)} m'(s)\int_0^x (v(s,\xi)-\tfrac{1}{2})d\xi ds.$$

We wish to apply the Mean Value Theorem to F on the interval $[m_t, m_{t+1}]$. Continuity of F is obvious, and differentiability will follow from showing that m' is bounded. Note

$$\begin{aligned}0 &= \frac{d}{ds}u(s,m_s)\\ &= \frac{\partial u}{\partial t}(s,m_s) + \frac{\partial u}{\partial x}(s,m_s)m'(s)\\ &= \frac{\partial u}{\partial t}(s,m_s) + \frac{\partial v}{\partial x}(s,0)m'(s),\end{aligned}$$

and by Corollary 5.2.5 and Corollary 5.2.6, we know that $\frac{\partial v}{\partial x}(s,0)$ is decreasing in s and bounded below by $U'_{\frac{1}{2}}(0)$. But we know from Theorem 5.2.2 that $U_{\frac{1}{2}}(s)$ is strictly increasing, so that $U'_{\frac{1}{2}}(0) > 0$. Therefore

$$|m'(s)| \le \left|\frac{\frac{\partial u}{\partial t}(s,m_s)}{U'_{\frac{1}{2}}(0)}\right| < \infty.$$

It follows that m' is bounded on each $[0,t]$, and hence F is differentiable with

$$F'(t) = \int_0^x (v(m^{-1}(t),\xi)-\tfrac{1}{2})d\xi ds.$$

Fix $t>0$. Then by the Mean Value Theorem there exists $\eta_t\in(m_t,m_{t+1})$ such that

$$F(m_{t+1}) - F(m_t) = F'(\eta_t)(m_{t+1}-m_t).$$

Then

$$\begin{aligned}\int_t^{t+1} m'(s)\int_0^x (v(s,\xi)-\tfrac{1}{2})d\xi ds &= F(m_{t+1}) - F(m_t)\\ &= F'(\eta_t)(m_{t+1}-m_t)\\ &= (m_{t+1}-m_t)\int_0^x (v(m^{-1}(\eta_t),\xi)-\tfrac{1}{2})d\xi\\ &= (m_{t+1}-m_t)\int_0^x (w(\xi)-\tfrac{1}{2})d\xi + o(1).\end{aligned}$$

Putting all this together gives

$$\tfrac{1}{2}(w(x)-\tfrac{1}{2}-ax)+(m_{t+1}-m_t)\int_0^x (w(\xi)-\tfrac{1}{2})d\xi+\int_0^x\int_0^\xi w(\eta)(w(\eta)-1)d\eta d\xi=o(1),$$

where $a=\lim_{t\to\infty}\frac{\partial v}{\partial x}(t,0)$. By Remark 5.2.7, $\int_0^x(w(\xi)-\frac{1}{2})d\xi\neq 0$ for all $x\neq 0$. Then it is clear that $c:=\lim_{t\to\infty}(m_{t+1}-m_t)$ exists, and satisfies

$$c\int_0^x (w(\xi)-\tfrac{1}{2})d\xi=-\tfrac{1}{2}(w(x)-\tfrac{1}{2}-ax)-\int_0^x\int_0^\xi w(\eta)(w(\eta)-1)d\eta d\xi$$

for all $x\in\mathbb{R}$. Differentiating gives

$$c(w(x)-\tfrac{1}{2})=-\tfrac{1}{2}(w'(x)-a)-\int_0^x w(\eta)(w(\eta)-1)d\eta, \tag{5.3}$$

and differentiating a second time gives

$$cw'(x)=-\tfrac{1}{2}w''(x)-w(x)(w(x)-1),$$

as required. □

By Remark 5.2.7 and Proposition 5.2.9, $w:\mathbb{R}\to(0,1)$ is a travelling wave solution with $\lim_{x\to-\infty}w(x)=0$ and $\lim_{x\to\infty}w(x)=1$. It follows immediately from Theorem 5.2.2 that $c\geq\sqrt{2}$. In order to prove that conversely, $c\leq\sqrt{2}$, we shall use the following lemma from [McK75].

Lemma 5.2.10. $c\leq\liminf_{t\to\infty}\dot{m}_t$.

Proof. Let $v(t,x)=u(t,x+m_t)$, so that

$$\frac{\partial v}{\partial t}=\frac{1}{2}\frac{\partial^2 v}{\partial x^2}+\dot{m}_t\frac{\partial v}{\partial x}+v(v-1).$$

By Corollary 5.2.5, $\frac{\partial v}{\partial t}(x,t)\geq 0$ for all $x<0$, so integrating over $(-\infty,0]$ gives

$$\begin{aligned}0&\leq\int_{-\infty}^0\left(\frac{1}{2}\frac{\partial^2 v}{\partial x^2}(t,x)+\dot{m}_t\frac{\partial v}{\partial x}(t,x)+v(t,x)(v(t,x)-1)\right)dx\\&=\left[\frac{1}{2}\frac{\partial v}{\partial x}(t,x)+\dot{m}_t v(t,x)\right]_{-\infty}^0+\int_{-\infty}^0 v(t,x)(v(t,x)-1)dx\\&\leq\frac{1}{2}\left(\frac{\partial v}{\partial x}(t,0)+\dot{m}_t\right)+\int_{-\infty}^0 v(t,x)(v(t,x)-1)dx,\end{aligned}$$

where the final inequality holds because $\frac{\partial v}{\partial x}(t,x)\geq 0$ and $\lim_{x\to-\infty}v(t,x)=0$. Taking $t\to\infty$ gives

$$\tfrac{1}{2}\liminf_{t\to\infty}\dot{m}_t\geq\int_{-\infty}^0 w(x)(1-w(x))dx-\tfrac{1}{2}\lim_{t\to\infty}\frac{\partial v}{\partial x}(t,0).$$

On the other hand, taking $x \to -\infty$ in Equation (5.3) gives

$$-\tfrac{1}{2}c = \tfrac{1}{2} \lim_{t\to\infty} \frac{\partial v}{\partial x}(t,0) - \int_{-\infty}^{0} w(x)(1-w(x))dx.$$

Therefore $\liminf_{t\to\infty} \dot{m}_t \geq c$. $\square$

In order to show that $c \leq \sqrt{2}$, it will suffice to show that $\limsup_{t\to\infty} \dot{m}_t \leq \sqrt{2}$. We will in fact prove a slightly stronger claim, the proof of which is due to McKean, [McK75]. First, though, we need the following technical lemma:

Lemma 5.2.11. Let B be a standard Brownian Motion, $x > 0$. Then

$$\frac{1}{\sqrt{2\pi}} \frac{x\sqrt{t}}{x^2+t} e^{-\frac{x^2}{2t}} \leq \mathbb{P}(B_t > x) \leq \frac{1}{\sqrt{2\pi}} \frac{\sqrt{t}}{x} e^{-\frac{x^2}{2t}}.$$

Proof. the upper bound is straightforward:

$$\begin{aligned}
\mathbb{P}(B_t > x) &= \int_x^\infty \frac{1}{\sqrt{2\pi t}} e^{-\frac{y^2}{2t}} dy \\
&= \int_{\frac{x}{\sqrt{t}}}^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy \\
&\leq \int_{\frac{x}{\sqrt{t}}}^\infty \left(\frac{y\sqrt{t}}{x}\right) \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy \\
&= \frac{1}{\sqrt{2\pi}} \frac{\sqrt{t}}{x} e^{-\frac{x^2}{2t}},
\end{aligned}$$

where the inequality follows from the observation that $\frac{y\sqrt{t}}{x} \geq 1$ on $\{y > \frac{x}{\sqrt{t}}\}$. For the lower bound, consider the function $F : \mathbb{R} \to \mathbb{R}$ defined by

$$F(x) = \mathbb{P}(B_t > x) - \frac{1}{\sqrt{2\pi}} \frac{x\sqrt{t}}{x^2+t} e^{-\frac{x^2}{2t}}.$$

We claim that F is always positive. Clearly $F(0) = \frac{1}{2} > 0$, and $\lim_{x\to\infty} F(x) = 0$. Therefore it suffices to show that F is decreasing:

$$\begin{aligned}
F'(x) &= -\frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}} - \frac{1}{\sqrt{2\pi}} \left(\frac{\sqrt{t}(t-x^2)}{(x^2+t)^2} - \frac{x^2}{\sqrt{t}(x^2+t)} \right) e^{-\frac{x^2}{2t}} \\
&= \frac{1}{\sqrt{2\pi}} \left(\frac{-(x^2+t)^2 - t(t-x^2) + x^2(x^2+t)}{\sqrt{t}(x^2+t)^2} \right) e^{-\frac{x^2}{2t}} \\
&= -\frac{1}{\sqrt{2\pi}} \frac{2t^2}{\sqrt{t}(x^2+t)^2} e^{-\frac{x^2}{2t}} \\
&< 0,
\end{aligned}$$

as required. $\square$

PROPOSITION 5.2.12 ([McK75]). For all t sufficiently large,

$$m_t \leq \sqrt{2}t - \frac{1}{2\sqrt{2}} \log t.$$

Proof. We need to show that for all t sufficiently large, $\mathbb{P}(M(t) \leq \sqrt{2}t - \frac{1}{2\sqrt{2}} \log t) > \frac{1}{2}$. Fix $x \in \mathbb{R}$. Then

$$\begin{aligned}
\mathbb{P}(M(t) \geq x) &= \mathbb{P}\left(\#\{u \in \mathcal{N}_t : X_u(t) \geq x\} > 0\right) \\
&\leq \sum_{k=0}^{\infty} \mathbb{P}\left(\#\{u \in \mathcal{N}_t : X_u(t) \geq x\} > k\right) \\
&= \mathbb{E}\left[\#\{u \in \mathcal{N}_t : X_u(t) \geq x\}\right] \\
&= \mathbb{E}\left[\sum_{u \in \mathcal{N}_t} \mathbb{1}_{\{X_u(t) \geq x\}}\right] \\
&= e^t \mathbb{E}\left[\mathbb{1}_{\{B_t \geq x\}}\right] \\
&= e^t \mathbb{P}(B_t \geq x).
\end{aligned}$$

where B is a standard Brownian Motion and the penultimate line follows from the Many-to-one Lemma. We now appeal to Lemma 5.2.11, with $x + \sqrt{2}t - \frac{1}{2\sqrt{2}} \log t$ in place of x. Note first that

$$e^{-\frac{1}{2t}\left(x+\sqrt{2}t-\frac{1}{2\sqrt{2}}\log t\right)^2} = e^{-\left(t-\frac{1}{2}\log t+\sqrt{2}x+o(1)\right)} = \sqrt{t}e^{-\left(t+\sqrt{2}x+o(1)\right)}.$$

Therefore

$$\begin{aligned}
\frac{1}{\sqrt{2\pi}} \frac{\left(x+\sqrt{2}t-\frac{1}{2\sqrt{2}}\log t\right)\sqrt{t}}{\left(x+\sqrt{2}t-\frac{1}{2\sqrt{2}}\log t\right)^2+t} e^{-\frac{1}{2t}\left(x+\sqrt{2}t-\frac{1}{2\sqrt{2}}\log t\right)^2} &= \frac{1}{\sqrt{2\pi}} \frac{\left(x+\sqrt{2}t-\frac{1}{2\sqrt{2}}\log t\right)t}{\left(x+\sqrt{2}t-\frac{1}{2\sqrt{2}}\log t\right)^2+t} e^{-\left(t+\sqrt{2}x+o(1)\right)} \\
&= \frac{1}{2\sqrt{\pi}} e^{-(t+\sqrt{2}x+o(1))}(1+o(1)),
\end{aligned}$$

and similarly

$$\begin{aligned}
\frac{1}{\sqrt{2\pi}} \frac{\sqrt{t}}{x+\sqrt{2}t-\frac{1}{2\sqrt{2}}\log t} e^{-\frac{1}{2t}\left(x+\sqrt{2}t-\frac{1}{2\sqrt{2}}\log t\right)^2} &= \frac{1}{\sqrt{2\pi}} \frac{t}{x+\sqrt{2}t-\frac{1}{2\sqrt{2}}\log t} e^{-\left(t+\sqrt{2}x+o(1)\right)} \\
&= \frac{1}{2\sqrt{\pi}} e^{-(t+\sqrt{2}x+o(1))}(1+o(1)).
\end{aligned}$$

It follows from Lemma 5.2.11 that

$$\mathbb{P}\left(B_t \geq x + \sqrt{2}t - \frac{1}{2\sqrt{2}} \log t\right) = \frac{1}{2\sqrt{\pi}} e^{-(t+\sqrt{2}x+o(1))}(1+o(1)). \tag{5.4}$$

Finally, we have

$$\begin{aligned}\lim_{t\to\infty} \mathbb{P}\left(M(t) \le x + \sqrt{2}t - \frac{1}{2\sqrt{2}}\log t\right) &\ge 1 - \lim_{t\to\infty} e^t \mathbb{P}\left(B_t \ge x + \sqrt{2}t - \frac{1}{2\sqrt{2}}\log t\right) \\ &= 1 - \lim_{t\to\infty} \frac{1}{2\sqrt{\pi}} e^{-(\sqrt{2}x + o(1))}(1 + o(1)) \\ &= 1 - \frac{1}{2\sqrt{\pi}} e^{-\sqrt{2}x}.\end{aligned}$$

Setting $x = 0$ in this expression yields

$$\lim_{t\to\infty} \mathbb{P}\left(M(t) \le \sqrt{2}t - \frac{1}{2\sqrt{2}}\log t\right) \ge 1 - \frac{1}{2\sqrt{\pi}} > \frac{1}{2},$$

as required. $\square$

Differentiating gives $\dot{m}_t \le \sqrt{2} - \frac{1}{2\sqrt{2}t}$ for all sufficiently large t, so $\limsup_{t\to\infty} \dot{m}_t \le \sqrt{2}$. Hence we have

$$\sqrt{2} \le c \le \liminf_{t\to\infty} \dot{m}_t \le \limsup_{t\to\infty} \dot{m}_t \le \sqrt{2},$$

so that $c = \sqrt{2}$. The proof is complete.

5.3 Martingale Representations

In the previous section, we saw that $m_t = \sqrt{2}t + o(t)$, and that $m_t \le \sqrt{2}t - \frac{1}{2\sqrt{2}}\log t$ for large t. Bramson significantly improved these estimates in [Bra78] and [Bra83].

THEOREM 5.3.1 ([Bra83]). Let m be any function of the form

$$m(t) = \sqrt{2}t - \frac{3}{2\sqrt{2}}\log t + c + o(1),$$

where c is a constant. Then $u(t, x + m(t)) \to w(x)$ uniformly as $t \to \infty$, where $w : \mathbb{R} \to (0, 1)$ is a travelling wave solution of speed $\sqrt{2}$ with $\lim_{x\to-\infty} w(x) = 0$, $\lim_{x\to\infty} w(x) = 1$.

Remark 5.3.2. By Theorem 5.2.2, a change in the constant c simply corresponds to a linear shift in the function w. It follows that there exists a constant C such that

$$m_t = \sqrt{2}t - \frac{3}{2\sqrt{2}}\log t + C + o(1).$$

The following estimate will also be needed later.

Lemma 5.3.3 ([Bra78])**.** Let w be as in Theorem 5.3.1. Then

$$1 - w(x) \sim Cxe^{-\sqrt{2}x}$$

as $x \to \infty$, for some $C > 0$.

Let

$$m(t) = \sqrt{2}t - \frac{3}{2\sqrt{2}} \log t + c + o(1).$$

Since $u(t, x + m(t)) = \mathbb{P}(M(t) - m(t) \le x)$, Theorem 5.3.1 tells us that $M(t) - m(t)$ converges in distribution to a random variable with distribution function $w(x)$; but how can we characterize this random variable? Lalley-Sellke resolved this question in [LS87] using martingales. We reproduce a streamlined version of their argument here.

Lemma 5.3.4 (Additive martingale)**.** Let

$$Y_t = \sum_{u \in \mathcal{N}_t} e^{\sqrt{2}X_u(t) - 2t}.$$

Then Y_t is an $\mathcal{F}_t^X$-martingale. Furthermore, Y_t converges almost surely to a finite non-negative random variable Y_∞.

Proof. By the Many-to-one Lemma,

$$\begin{aligned}
\mathbb{E}[Y_t] &= e^t \mathbb{E}\left[e^{\sqrt{2}B_t - 2t}\right] \\
&= e^t \int_{\infty}^{\infty} \frac{1}{\sqrt{2\pi t}} e^{-\frac{z^2}{2t}} e^{\sqrt{2}z - 2t} dz \\
&= e^t \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi t}} e^{-\frac{1}{2t}\left((z - \sqrt{2}t)^2 + 2t^2\right)} dz \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi t}} e^{-\frac{1}{2t}(z - \sqrt{2}t)^2} dz \\
&= 1.
\end{aligned}$$

for all $t > 0$. Now let $s < t$. Recall that by the Markov property of X_t, we have

$$X_t = \left(X_u(s) + X_v^{(u)}(t - s) : u \in \mathcal{N}_s, v \in \mathcal{N}_{t-s}^u\right),$$

where the $X_r^{(u)} = (X_v^{(u)}(r) : v \in \mathcal{N}_r^u)$ are dyadic Branching Brownian Motions with branching rate 1, independent of $\mathcal{F}_s^X$, so that

$$Y_t = \sum_{u \in \mathcal{N}_s} \sum_{v \in \mathcal{N}_{t-s}^u} e^{\sqrt{2}\left(X_u(s) + X_v^{(u)}(t-s)\right) - 2t} = \sum_{u \in \mathcal{N}_s} e^{\sqrt{2}X_u(s) - 2s} \sum_{v \in \mathcal{N}_{t-s}^u} e^{\sqrt{2}X_v^{(u)}(t-s) - 2(t-s)}.$$

Therefore

$$\begin{aligned}\mathbb{E}\left[Y_t|\mathcal{F}_s^X\right] &= \sum_{u\in\mathcal{N}_s} e^{\sqrt{2}X_u(s)-2s}\mathbb{E}\left[\sum_{v\in\mathcal{N}_{t-s}^u} e^{\sqrt{2}X_v^{(u)}(t-s)-2(t-s)}|\mathcal{F}_s^X\right]\\ &= Y_s\mathbb{E}\left[\sum_{v\in\mathcal{N}_{t-s}^u} e^{\sqrt{2}X_v^{(u)}(t-s)-2(t-s)}\right]\\ &= Y_s,\end{aligned}$$

where in the last line we used the Many-to-one Lemma just as above. Therefore Y_t is a bounded martingale, and Y_t is clearly positive and almost surely right-continuous. The conclusion then follows from the Martingale Convergence Theorem. □

Remark 5.3.5. Since Y_∞ is finite, and $N(t)\to\infty$ almost surely, we have

$$\min_{u\in\mathcal{N}_t}\left(2t-\sqrt{2}X_u(t)\right)\uparrow\infty$$

as $t\to\infty$.

Now consider the PDE

$$\frac{\partial u}{\partial t}=\frac{1}{2}\frac{\partial^2 u}{\partial x^2}+u(u-1) \qquad\qquad u(0,x)=w(x).$$

By Theorem 3.2.1, it has solution

$$u(t,x)=\mathbb{E}\left[\prod_{u\in\mathcal{N}_t} w(x-X_u(t))\right].$$

Let $v(t,x):=u(t,x+\sqrt{2}t)$, then

$$\frac{\partial v}{\partial t}=\frac{1}{2}\frac{\partial^2 v}{\partial x^2}+\sqrt{2}\frac{\partial v}{\partial x}+v(v-1)$$

with $v(0,x)=w(x)$. But $v(t,x)=w(x)$ clearly solves this PDE, so we have

$$w(x)=\mathbb{E}\left[\prod_{u\in\mathcal{N}_t} w\left(x+\sqrt{2}t-X_u(t)\right)\right]. \tag{5.5}$$

Lemma 5.3.6 (Multiplicative martingale)**.** For each $x\in\mathbb{R}$,

$$W_t(x)=\prod_{u\in\mathcal{N}_t} w\left(x+\sqrt{2}t-X_u(t)\right)$$

defines an $\mathcal{F}_t^X$-martingale. Furthermore, $W_t(x)$ converges almost surely and in L^1 to a random variable $W_\infty(x)$.

Proof. Since $0 \le w \le 1$, W is clearly bounded. Fix $s < t$. We have

$$X_t = \left(X_u(s) + X_v^{(u)}(t-s) : u \in \mathcal{N}_s, v \in \mathcal{N}_{t-s}^u\right),$$

with $X_r^{(u)} = (X_v^{(u)}(r) : v \in \mathcal{N}_r^u)$ as in Lemma 5.3.4, so

$$W_t(x) = \prod_{u \in \mathcal{N}_s} \prod_{v \in \mathcal{N}_{t-s}^u} w\left(x + \sqrt{2}s - X_u(s) + \sqrt{2}(t-s) - X_v^{(u)}(t-s)\right).$$

Then

$$\begin{aligned}
\mathbb{E}[W_t(x)|\mathcal{F}_s^X] &= \prod_{u \in \mathcal{N}_s} \mathbb{E}\left[\prod_{v \in \mathcal{N}_{t-s}^u} w\left(x + \sqrt{2}s - X_u(s) + \sqrt{2}(t-s) - X_v^{(u)}(t-s)\right) |\mathcal{F}_s^X\right] \\
&= \prod_{u \in \mathcal{N}_s} w\left(x + \sqrt{2}s - X_u(s)\right) \\
&= W_s(x),
\end{aligned}$$

where the second line follows from independence and Equation (5.5). Therefore $W(x)$ is a bounded martingale. Almost sure continuity is obvious, and hence the conclusion follows from the Martingale Convergence Theorem. $\square$

Now fix $x \in \mathbb{R}$ and notice that

$$W_t(x) = \prod_{u \in \mathcal{N}_t} e^{\log w\left(x + \sqrt{2}t - X_u(t)\right)}.$$

For each $u \in \mathcal{N}_t$, $\log w\left(x + \sqrt{2}t - X_u(t)\right) \uparrow \infty$ as $t \to \infty$ by Remark 5.3.5, so by Lemma 5.3.3 there exists $C > 0$ such that

$$\begin{aligned}
\log w\left(x + \sqrt{2}t - X_u(t)\right) &\sim \log\left(1 - C\left(x + \sqrt{2}t - X_u(t)\right)e^{-\sqrt{2}\left(x + \sqrt{2}t - X_u(t)\right)}\right) \\
&\sim -C\left(x + \sqrt{2}t - X_u(t)\right)e^{-\sqrt{2}\left(x + \sqrt{2}t - X_u(t)\right)}
\end{aligned}$$

as $t \to \infty$. Therefore

$$\begin{aligned}
W_t(x) &\sim \prod_{u \in \mathcal{N}_t} e^{-C\left(x + \sqrt{2}t - X_u(t)\right)e^{-\sqrt{2}(x + \sqrt{2}t - X_u(t))}} \\
&= e^{-C\sum_{u \in \mathcal{N}_t}\left(x + \sqrt{2}t - X_u(t)\right)e^{-\sqrt{2}(x + \sqrt{2}t - X_u(t))}} \\
&= e^{-C\left(xe^{-\sqrt{2}x}Y_t + e^{-\sqrt{2}x}\sum_{u \in \mathcal{N}_t}\left(\sqrt{2}t - X_u(t)\right)e^{-\sqrt{2}(\sqrt{2}t - X_u(t))}\right)},
\end{aligned}$$

as $t \to \infty$.

Definition 5.3.7. The *derivative martingale* is the process

$$Z_t = \sum_{u \in \mathcal{N}_t} \left(\sqrt{2}t - X_u(t)\right) e^{-\sqrt{2}\left(\sqrt{2}t - X_u(t)\right)}.$$

We then have

$$W_t(x) \sim e^{-C\left(xe^{-\sqrt{2}x}Y_\infty + e^{-\sqrt{2}x}Z_t\right)}$$

as $t \to \infty$. Our next aim is to establish almost sure convergence of Z_t. Unlike our previous two martingales, Z_t is not L^1-bounded, so we cannot appeal to Doob's Martingale Convergence Theorem - in fact the martingale property will not be used at all.

Lemma 5.3.8. Z_t converges almost surely to a positive finite random variable Z_∞.

Proof. Note that

$$\begin{aligned} Z_t &= \sum_{u \in \mathcal{N}_t} \left(\sqrt{2}t - X_u(t)\right) e^{\sqrt{2}X_u(t) - 2t} \\ &\geq \min_{u \in \mathcal{N}_t} \left(\sqrt{2}t - X_u(t)\right) Y_t \\ &\sim \min_{u \in \mathcal{N}_t} \left(\sqrt{2}t - X_u(t)\right) Y_\infty, \end{aligned}$$

as $t \to \infty$. Let $A = \{Y_\infty > 0\}$. By Remark 5.3.5, $Z_t \to \infty$ on A, so that for each $x \in \mathbb{R}$, $W_\infty(x) = 0$ on A. Suppose for contradiction that $\mathbb{P}(A) > 0$. For each $x \in \mathbb{R}$,

$$\mathbb{E}[W_\infty(x)] = \mathbb{E}[W_\infty(x) : \Omega \backslash A] \leq \mathbb{P}(\Omega \backslash A) < 1,$$

because $0 \leq W_\infty(x) \leq 1$, and since $W_t(x)$ is a martingale we have

$$\mathbb{E}[W_\infty(x)] = \mathbb{E}[W_0(x)] = w(x).$$

It follows that

$$w(x) \leq \mathbb{P}(\Omega \backslash A) < 1.$$

But we had $\lim_{x \to \infty} w(x) = 1$. Hence, by contradiction, $Y_\infty = 0$ almost surely. Therefore

$$W_t(x) \sim e^{-Ce^{-\sqrt{2}x}Z_t}$$

as $t \to \infty$. It follows from Lemma 5.3.6 that Z_t converges almost surely to a positive finite random variable Z_∞. ∎

We can now proof the main theorem of Lalley-Sellke ([LS87]).

THEOREM 5.3.9. Fix $x \in \mathbb{R}$. Then

$$\lim_{t\to\infty} \mathbb{P}(M(t) - m(t) \le x) = \mathbb{E}\left[e^{-Ce^{-\sqrt{2}x} Z_\infty}\right],$$

where $1 - w(x) \sim Cxe^{-\sqrt{2}x}$ as $x \to \infty$.

Proof. Fix $s > 0$. For each $t > 0$ we have

$$X_{t+s} = \left(X_u(s) + X_v^{(u)}(t) : u \in \mathcal{N}_s, v \in \mathcal{N}_t^u\right),$$

where $X_r^{(u)} = (X_v^{(u)}(r) : v \in \mathcal{N}_r^u)$ as in Lemmas 5.3.4 and 5.3.6. Letting

$$M^u(t) := \max_{v \in \mathcal{N}_t^u} X_v^{(u)}(t),$$

so that the M^u are independent, we have

$$M(t+s) = \max_{u \in \mathcal{N}_s} \left(X_u(s) + M^u(t)\right).$$

Then

$$\begin{aligned}\mathbb{P}\left(M(t+s) - m(t+s) \le x | \mathcal{F}_s^X\right) &= \mathbb{P}\left(\max_{u\in\mathcal{N}_s} \left(X_u(s) + M^u(t)\right) \le x + m(t+s) | \mathcal{F}_s^X\right) \\ &= \prod_{u\in\mathcal{N}_s} \mathbb{P}\left(M^u(t) \le x + m(t+s) - X_u(s) | \mathcal{F}_s^X\right),\end{aligned}$$

by independence of the M^u. Hence

$$\begin{aligned}\lim_{t\to\infty} \mathbb{P}(M(t) - m(t) \le x) &= \lim_{t\to\infty} \mathbb{P}(M(t+s) - m(t+s) \le x) \\ &= \lim_{t\to\infty} \mathbb{E}\left[\mathbb{P}\left(M(t+s) - m(t+s) \le x | \mathcal{F}_s^X\right)\right] \\ &= \lim_{t\to\infty} \mathbb{E}\left[\prod_{u\in\mathcal{N}_s} \mathbb{P}\left(M^u(t) \le x + m(t+s) - X_u(s) | \mathcal{F}_s^X\right)\right] \\ &= \mathbb{E}\left[\prod_{u\in\mathcal{N}_s} \lim_{t\to\infty} \mathbb{P}\left(M^u(t) \le x + m(t+s) - X_u(s) | \mathcal{F}_s^X\right)\right],\end{aligned}$$

where the last equality follows from the Dominated Convergence Theorem. Now

$$m(t+s) - m(t) - \sqrt{2}s = \frac{3}{2\sqrt{2}} \left(\log(t+s) - \log s\right) + o(1),$$

and $\log(t+s) - \log s \to 0$ as $t \to \infty$, so

$$\begin{aligned}\lim_{t\to\infty} \mathbb{P}\left(M^u(t) \le x + m(t+s) - X_u(s) | \mathcal{F}_s\right) &= \lim_{t\to\infty} \mathbb{P}\left(M^u(t) \le x + m(t) + \sqrt{2}s - X_u(s) | \mathcal{F}_s\right) \\ &= w\left(x + \sqrt{2}s - X_u(s)\right)\end{aligned}$$

for each $u \in \mathcal{N}_s$. We therefore have

$$\lim_{t\to\infty} \mathbb{P}(M(t) - m(t) \le x) = \mathbb{E}\left[\prod_{u\in\mathcal{N}_s} w\left(x + \sqrt{2}s - X_u(s)\right)\right] = \mathbb{E}[W_s(x)].$$

But we know that

$$\mathbb{E}[W_s(x)] \sim \mathbb{E}\left[e^{-Ce^{-\sqrt{2}x}Z_s}\right],$$

so

$$\lim_{t\to\infty} \mathbb{P}(M(t) - m(t) \le x) \sim \mathbb{E}\left[e^{-Ce^{-\sqrt{2}x}Z_s}\right]$$

for all $s \in \mathbb{R}$. Since the LHS does not depend on s, it follows from taking $s \to \infty$ that

$$\lim_{t\to\infty} \mathbb{P}(M(t) - m(t) \le x) = \mathbb{E}\left[e^{-Ce^{-\sqrt{2}x}Z_\infty}\right].$$

□

We therefore have a nice probabilistic interpretation of Bramson's result.

5.4 The Role of Branching Structure

Theorem 5.3.1 also gives us an insight into the branching structure of X_t. In this section, we consider the process obtained by replacing the particles of X_t with independent Brownian Motions, and how it differs from X_t. In this context, it will emerge that McKean's upper bound

$$m_t \le \sqrt{2}t - \frac{1}{2\sqrt{2}}\log t, \tag{5.6}$$

is quite meaningful, despite being a severe overestimate. Recall that $N(t) = |\mathcal{N}_t|$. The following lemma is new.

Lemma 5.4.1. $N(t)$ has geometric(e^{-t}) distribution.

Proof. We will show by induction that $\mathbb{P}(N(t) = k) = e^{-t}(1 - e^{-t})^{k-1}$ for each $k \in \mathbb{N}$. Let $\ell = \inf_{s>0}\{N(s) = 2\}$, so that $\ell \sim \exp(1)$. Clearly $\mathbb{P}(N(t) = 1) = \mathbb{P}(\ell > t) = e^{-t}$, so that the result holds for $k = 1$. Suppose that the result holds for $k = 1, 2, ..., n-1$. For $i = 1, 2$, we let

$$\mathcal{N}_s^i = \left\{v \in \mathcal{N}_{\ell+s} : p^k(v) = (i) \text{ for some } k \in \mathbb{N}\right\}$$

be the particles alive at time $\ell + s$ that are descended from particle (i). Let $N^i(s) = |\mathcal{N}_s^i|$. Then $N(t) = N^1(t-\ell) + N^2(t-\ell)$ on $\{\ell \le t\}$, and by the branching property

N^1 and N^2 are independent with the same law as N. Then

$$\begin{aligned}
\mathbb{P}(N(t)=n) &= \int_0^t e^{-s}\mathbb{P}(N(t)=n|\ell=s)ds \\
&= \int_0^t e^{-s}\mathbb{P}(N^1(t-s)+N^2(t-s)=n)ds \\
&= \int_0^t e^{-s}\sum_{k=1}^{n-1}\mathbb{P}(N^1(t-s)=k)\mathbb{P}(N^2(t-s)=n-k)ds \\
&= \int_0^t e^{-s}\sum_{k=1}^{n-1} e^{-(t-s)}\left(1-e^{-(t-s)}\right)^{k-1} e^{-(t-s)}\left(1-e^{-(t-s)}\right)^{n-k-1} ds \\
&= e^{-t}\int_0^t e^{-(t-s)}\sum_{k=1}^{n-1}\left(1-e^{-(t-s)}\right)^{n-2} ds \\
&= (n-1)e^{-t}\int_0^t e^{-(t-s)}\left(1-e^{-(t-s)}\right)^{n-2} ds \\
&= (n-1)e^{-t}\left[-\frac{1}{n-1}\left(1-e^{-(t-s)}\right)^{n-1}\right]_0^t \\
&= e^{-t}\left(1-e^{-t}\right)^{n-1}.
\end{aligned}$$

□

We have the following analogue of Lemma 5.3.8. Our proof is adapted from [Bov17], with considerable detail added.

Proposition 5.4.2. $M_t = e^{-t}N(t)$ is a right-continuous $\mathcal{F}_t^N$-martingale. Furthermore, M_t converges almost surely and in L^1 to a random variable $M_\infty \in L^1(\Omega)$

Proof. Fix $s > 0$. For $u \in \mathcal{N}_s$, let

$$\mathcal{N}_t^u = \left\{v \in \mathcal{N}_{t+s} : p^k(v) = u \text{ for some } k \in \mathbb{N}\right\}$$

be the particles alive at time $t+s$ that are descended from u and let $N^u(t) = |\mathcal{N}_t^u|$. By the branching property, the N^u are independent of $\mathcal{F}_s^N$ and have the same law as N. Let $t > s$. Then

$$N(t) = \sum_{u\in\mathcal{N}_s} N^u(t-s),$$

and therefore

$$\begin{aligned}
\mathbb{E}[M_t|\mathcal{F}_s^N] &= \mathbb{E}\left[e^{-t}\sum_{u\in\mathcal{N}_s} N^u(t-s)|\mathcal{F}_s^N\right] \\
&= e^{-t}\sum_{u\in\mathcal{N}_s}\mathbb{E}\left[N^u(t-s)|\mathcal{F}_s^N\right] \\
&= e^{-t}N(s)\mathbb{E}[N(t-s)] \\
&= M_s,
\end{aligned}$$

where the final equality follows from Lemma 3.1.8. For each $t \geq 0$, M_t is positive and $\mathbb{E}[M_t] = 1$ by Lemma 3.1.8, so that M is an L^1-bounded martingale, and right continuity follows from our construction of N. Then by the Martingale Convergence Theorem, M_t converges almost surely to a random variable $M_\infty \in Ł^1(\Omega)$. For L^1 convergence it suffices to show that M_t is uniformly integrable. This will follow from showing that M_t is L^2-bounded. Let $\phi(t) = \mathbb{E}[M_t^2]$, and $\ell = \inf_{s>0}\{N(s) = 2\}$. Then

$$\begin{aligned}
\phi(t) &= e^{-2t}\mathbb{E}[N(t)^2] \\
&= e^{-2t}\left(e^{-t}\mathbb{E}[N(t)^2|\ell > t] + \int_0^t e^{-s}\mathbb{E}[N(t)^2|\ell = s]ds\right) \\
&= e^{-3t} + e^{-2t}\int_0^t e^{-s}\mathbb{E}[N(t)^2|\ell = s]ds.
\end{aligned}$$

For $0 \leq s \leq t$, define N^1, N^2 as in Lemma 5.4.1, so that $N(t) = N^1(t-s) + N^2(t-s)$ on $\{\ell = s\}$. Then

$$\begin{aligned}
\mathbb{E}[N(t)^2|\ell = s] &= \mathbb{E}[(N^1(t-s) + N^2(t-s))^2] \\
&= 2\mathbb{E}[N(t-s)^2] + 2\mathbb{E}[N(t-s)]^2 \\
&= 2e^{2(t-s)}\phi(t-s) + 2e^{2(t-s)},
\end{aligned}$$

where the second line follows from the Markov property. We have

$$\begin{aligned}
\phi(t) &= e^{-3t} + 2\int_0^t e^{-3s}\phi(t-s)ds + 2\int_0^t e^{-3s}ds \\
&= \frac{1}{3}(e^{-3t} + 2) + 2\int_0^t e^{-3s}\phi(t-s)ds.
\end{aligned}$$

Differentiating gives

$$\begin{aligned}
\phi'(t) &= -e^{-3t} + 2e^{-3t}\phi(0) + 2\int_0^t e^{-3s}\phi'(t-s)ds \\
&= e^{-3t} - 2\int_0^t e^{-3s}\frac{\partial}{\partial s}\phi(t-s)ds \\
&= e^{-3t} - 2\left(\left[e^{-3s}\phi(t-s)\right]_0^t + \int_0^t 3e^{-3s}\phi(t-s)ds\right) \\
&= -e^{-3t} + 2\phi(t) - 6\int_0^t e^{-3s}\phi(t-s)ds \\
&= -\phi(t) + 3\left(\phi(t) - \frac{1}{3}e^{-3t} - 2\int_0^t e^{-3s}\phi(t-s)ds\right) \\
&= 2 - \phi(t).
\end{aligned}$$

Since $\phi(0) = 1$, this yields $\phi(t) = 2 - e^{-t}$, so that M is L^2-bounded, as required. $\square$

Unlike Z_∞ in the previous section, we can easily characterize the distribution of M_∞. The following proof is original.

Lemma 5.4.3. M_∞ has Exponential(1) distribution.

Proof. Fix $x > 0$. By Proposition 5.4.2, it suffices to show that for some sequence $t_n \uparrow \infty$, we have $\mathbb{P}(M_{t_n} \leq x) \to 1 - e^{-x}$ as $n \to \infty$. Indeed, let $t_n = \log(\frac{n}{x})$. Then

$$\begin{aligned}
\mathbb{P}(M_{t_n} \leq x) &= \mathbb{P}(N(t_n) \leq xe^{t_n}) \\
&= \mathbb{P}(N(t_n) \leq n) \\
&= \sum_{k=1}^{n} e^{-t_n}(1 - e^{-t_n})^{k-1} \\
&= \tfrac{x}{n}\sum_{k=0}^{n-1}(1 - \tfrac{x}{n})^k \\
&= 1 - (1 - \tfrac{x}{n})^n \\
&\to 1 - e^{-x}, \qquad \text{as } n \to \infty
\end{aligned}$$

$\square$

Now let $Y_t = (B_u(t) : u \in \mathcal{N}_t)$, where the B_u are independent standard Brownian Motions. Let $R(t) = \max_{u\in\mathcal{N}_t} B_u(t)$. We have the following analogue of Theorem 5.3.9 for the process $R(t)$. Our proof follows [Ber15].

Theorem 5.4.4. Let $r(t) = \sqrt{2}t - \frac{1}{2\sqrt{2}}\log t + C$, where C is a constant. Then

$$\lim_{t\to\infty} \mathbb{P}(R(t) - r(t) \leq x) = \mathbb{E}\left[e^{-\frac{1}{2\sqrt{\pi}}e^{-\sqrt{2}C}M_\infty e^{-\sqrt{2}x}}\right]. \tag{5.7}$$

Proof. Fix $x \in \mathbb{R}$. By independence of the B_u, we have

$$\begin{aligned}\mathbb{P}(R(t) - r(t) \le x) &= \mathbb{P}\left(\max_{u \in \mathcal{N}_t} B_u(t) \le x + r(t)\right) \\ &= \mathbb{E}\left[\mathbb{P}(B_t \le x + r(t))^{N(t)}\right] \\ &= \mathbb{E}\left[(1 - \mathbb{P}(B_t \ge x + r(t)))^{N(t)}\right].\end{aligned}$$

By Proposition 5.4.2, $N(t) = e^t(M_\infty + o(1))$ almost surely, and from Equation (5.4), we have

$$\mathbb{P}(B_t \ge x + r(t)) = \frac{1}{2\sqrt{\pi}} e^{-\left(t + \sqrt{2}(x+C) + o(1)\right)}(1 + o(1)).$$

Hence

$$\begin{aligned}\lim_{t\to\infty} (1 - \mathbb{P}(B_t \ge x + r(t)))^{N(t)} &= \lim_{t\to\infty} \left(1 - \frac{1}{2\sqrt{\pi}} e^{-\left(t+\sqrt{2}(x+C)\right)}\right)^{e^t M_\infty} \\ &= \lim_{t\to\infty} \left(1 - \frac{1}{2\sqrt{\pi}} \frac{M_\infty}{t} e^{-\sqrt{2}(x+C)}\right)^{t} \\ &= e^{-\frac{1}{2\sqrt{\pi}} e^{-\sqrt{2}C} M_\infty e^{-\sqrt{2}x}}\end{aligned}$$

almost surely. Finally, we note that

$$(1 - \mathbb{P}(B_t \ge x + r(t)))^{N(t)}$$

is always bounded by 1, so we may apply the Dominated Convergence Theorem, which gives

$$\lim_{t\to\infty} \mathbb{P}(R(t) - r(t) \le x) = \mathbb{E}\left[e^{-\frac{1}{2\sqrt{\pi}} e^{-\sqrt{2}C} M_\infty e^{-\sqrt{2}x}}\right].$$

□

By choosing C in Equation (5.7) such that the RHS is equal to $\frac{1}{2}$, we see that the limiting behaviour of the median of $R(t)$ is in fact described by McKean's upper bound (5.6).

The discrepancy between $M(t)$ and $R(t)$ is rooted in the branching structure of X_t; in particular, the $X_u(t)$ are not independent. Rather, their correlation depends on their genealogical history. Indeed, let $u, v \in \mathcal{N}_t$ and let $n = \min\{k \in \mathbb{N} : p^k(u) = p^k(v)\}$, so that $w = p^n(u)$ is the most recent common ancestor of u and v. Then

$$X_u(t) = B_{d_w} + B_u(t - d_w), \qquad X_v(t) = B_{d_w} + B_v(t - d_w),$$

where d_w is the death time of w, and B, B_u and B_v are independent standard Brownian Motions. We have

$$\begin{aligned}\mathbb{E}[X_u(t)X_v(t)|d_w = s] &= \mathbb{E}[B_s^2 + B_s(B_u(t-s) + B_v(t-s)) + B_u(t-s)B_v(t-s)] \\ &= s,\end{aligned}$$

by independence. Therefore, the branching structure of Branching Branching Brownian Motion is integral to understanding the behaviour of the maximal process.

In our interdisciplinary study, we have seen how Branching Brownian Motion and analytic PDE theory are complementary topics. On the one hand, Branching Brownian Motion can characterize solutions to the semilinear heat equation; on the other, PDE theory grants us fruitful insights into the properties of Branching Brownian Motion.

A Appendix

THEOREM A.0.1 (Leibniz's Integral Rule). Let $a, b \in C^1(\mathbb{R})$, let $f : \mathbb{R}\times\mathbb{R} \to \mathbb{R}$ be such that $f(t,x)$ and $\frac{\partial f}{\partial t}(t,x)$ are continuous on $\{(t,x) \in \mathbb{R}^2 : t \in \mathbb{R}, a(t) \leq x \leq b(t)\}$. Then

$$\frac{d}{dt}\int_{a(t)}^{b(t)} f(t,x)dx = \int_{a(t)}^{b(t)} \frac{\partial f}{\partial t}(t,x)dx + f(t,b(t))b'(t) - f(t,a(t))a'(t).$$

THEOREM A.0.2 (Picard-Lindelöf Theorem). Let $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^n$ be globally Lipschitz continuous (with respect to any norm). Then the initial value problem

$$y'(\underline{x}) = \mathbf{f}(y(\underline{x})) \qquad y(\underline{x_0}) = \underline{y_0},$$

has a unique solution $y : \mathbb{R}^n \to \mathbb{R}^n$.

THEOREM A.0.3 (Vandermonde's Identity). Let $r, m, n \in \mathbb{N}$ with $r \leq m + n$. Then

$$\binom{m+n}{r} = \sum_{k=0}^{r} \binom{m}{k}\binom{n}{r-k}.$$

Bibliography

[AW75] D.G. Aronson and H.F. Weinberger. *Nonlinear Diffusion in Population Genetics, Combustion, and Nerve Pulse Propagation*, volume 446, pages 5–49. Springer, Berlin/Heidelberg, 1975.

[Ber15] J. Berestycki. Topics on branching Brownian motion, 2015. Retrieved from http://www.stats.ox.ac.uk/ berestyc/articles.html.

[Bie45] I.J Bienaymé. De la loi de la multiplication et de la durée des familles. *Bulletin de la Société Philomathique de Paris*, 5:37–39, 1845.

[Bov17] A. Bovier. *Gaussian Processes on Trees : From Spin Glasses to Branching Brownian Motion.* Cambridge University Press, Cambridge, 2017.

[Bra78] M. Bramson. Maximal displacement of branching Brownian motion. *Communications on Pure and Applied Mathematics*, 31(5):531–581, 1978.

[Bra82] M. Bramson. The Kolmogorov nonlinear diffusion equation, 1982. Retrieved from the University of Minnesota Digital Conservancy, http://hdl.handle.net/11299/151582.

[Bra83] M. Bramson. *Convergence of Solutions of the Kolmogorov Equation to Travelling Waves.* American Mathematical Society, Providence, Rhode Island, 1983.

[Col06] P.J. Collins. *Differential and Integral Equations.* Oxford University Press, Oxford, 2006.

[EFP17] A. Etheridge, N.P. Freeman, and S. Penington. Branching Brownian motion, mean curvature flow and the motion of hybrid zones. *Electronic Journal of Probability*, 22(103):1–40, 2017.

[Fis37] R.A. Fisher. The wave of advance of advantageous genes. *Annals of Eugenics*, 7:355–369, 1937.

[FM75] P.C. Fife and J.B. McLeod. The approach of solutions of nonlinear diffusion equations to travelling wave solutions. *Bulletin of the American Mathematical Society*, 81(6):1076–1078, 1975.

[Fou22] J-B.J. Fourier. *The Analytical Theory of Heat*. Cambridge University Press, Cambridge, 1822. translated by A. Freeman.

[GW74] F. Galton and H. Watson. On the probability of the extinction of families. *Journal of the Anthropological Institute of Great Britain and Ireland*, 4(1):138–144, 1874.

[INW65] N. Ikeda, M. Nagasawa, and S. Watanabe. On branching Markov processes. *Proceedings of the Japan Academy*, 41(9):816–821, 1965.

[KPP37] A. Kolmogorov, I. Petrovskii, and N. Piskunov. Étude de l'équation de la diffusion avec croissance de la quantité de la matière et son application à un problème biologique. *Moscow University Mathematics Bulletin*, 1(6):1–26, 1937.

[Lév48] P. Lévy. *Processus Stochastiques et Mouvement Brownien*. Gauthier-Villars, Paris, 1948.

[LG16] J-F. Le Gall. *Brownian Motion, Martingales, and Stochastic Calculus*. Springer, Cham, 2016.

[LS87] S.P. Lalley and T. Sellke. A conditional limit theorem for the frontier of a branching Brownian motion. *The Annals of Probability*, 15(3):1052–1061, 1987.

[McK75] H.P. McKean. Application of Brownian motion to the equation of Kolmogorov-Petrovskii-Piskunov. *Communications on Pure and Applied Mathematics*, 28(3):323–331, 1975.

[Øks03] B. Øksendal. *Stochastic Differential Equations: an Introduction with Applications*. Springer, Berlin, 6th edition, 2003.

[RY99] D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion*. Springer, Berlin/London, 3rd edition, 1999.

[Saa03] W. van Saarloos. Front propagation into unstable states. *Physics Reports*, 386(2):29–222, 2003.

[Sko65] A.V. Skorokhod. *Studies in the Theory of Random Processes*. Addison-Wesley, Reading, Massachusetts, 1965.

[Wie23] N. Wiener. Differential-Space. *Journal of Mathematics and Physics*, 2:131–174, 1923.