Abstract

These notes are a compilation of pieces from the notes by Julien Berestycki, [1] Anton Bovier [2] and Ofer Zeitouni [9]. All mistakes are mine, all insights are theirs, even if occasionally misinterpreted.

1 Branching Brownian motion

In these notes, we will be concerned with the time evolution of a large number of particles. Usually, their number will grow in time as well. The simplest such example would be to take $N$ particles $x_1(t), x_2(t), \ldots, x_N(t)$, with each particle performing an independent copy of a fixed random process. For example, each $x_k(t)$ can be a standard Brownian motion in $\mathbb{R}^d$, $d \geq 1$, starting at the point $x_j(0) = 0$. However, if we fix the number $N$ of particles and consider what happens as $t \to +\infty$, the configuration would become really sparse in space. Hence, it may be more interesting to let the number of particles to grow in time, to keep their density from vanishing.

Such models involving an increasing number of particles appear very naturally in the context of biological invasions in ecology, as well as in SIR-type models of epidemics. A simple and common process of this type is the binary branching Brownian motion. It is described as follows. A single particle starts at a position $x \in \mathbb{R}^d$ at $t = 0$ and performs a standard Brownian motion. The particle carries an exponential clock that rings at a random time $\tau$, with

$$
P(\tau > t) = e^{-t}.
$$

At the time $\tau$ the particle splits into two particles that we will refer to as the children, and the original particle is sometimes called the parent. The original particle is removed at the branching event. The two children perform independent standard Brownian motions for $t > \tau$, both of them starting at the position of the branching event. Each of the children carries its own exponential clock, and when the corresponding clock rings, the particle splits into two, and the process continues. Thus, at each time $t > 0$ we have a collection of particles $x_1(t), \ldots, x_{N_t}(t)$. Here, $N_t - 1$ is the random number of times all the clocks rang until the time $t$. 

*Department of Mathematics, Stanford University, Stanford CA, 94305, USA; ryzhik@math.stanford.edu
A way to think of this process in terms of epidemics, is a SIR-type model where nobody is ever removed, but can infect. Each particle meets other particles at the times when the clock rings and infects one other particle at each ”infection event”.

The above model is usually referred to as a binary Brownian motion since the number of children is limited to two. A simple modification of the above model is a process where the particles may produce a random number of \( k \) children at each birth event, with

\[
\sum_{k=2}^{\infty} p_k = 1, \tag{1.2}
\]

and the average number of children

\[
\bar{N} = \sum_{k=2}^{\infty} kp_k. \tag{1.3}
\]

We will usually assume that \( p_k \) decay sufficiently fast as \( k \to +\infty \).

The total number \( N(t) \) of particles present at the time \( t > 0 \) can be thought of as a pure birth process – \( N(t) \) can go up but not down. As a warm-up, let us prove the following.

**Proposition 1.1** Let \( N(t) \) be the number of particles present in the binary BBM with the exponential clock as in (1.1), then

\[
\mathbb{E}(N(t)) = e^t. \tag{1.4}
\]

**Proof.** Let us consider the first branching time \( \tau_1 \). If \( \tau_1 > t \), then \( N(t) = 1 \). On the other hand, if \( \tau_1 \leq t \), then the total number of particles at the time \( t \) is the total number of particles coming from the first child born at the time \( \tau_1 \) plus the total number of particles coming from the second child born at the time \( \tau_1 \). The time elapsed between \( \tau_1 \) and \( t \) is \( t - \tau_1 \). This gives the following recursion relation:

\[
\mathbb{E}(N(t)) = \mathbb{P}(\tau_1 > t) + 2\mathbb{E}(N(t-\tau_1))\mathbb{P}(\tau_1 < t) = e^{-t} + 2\int_0^t \mathbb{E}(N(t-s))\mathbb{P}(\tau_1 \in ds) \tag{1.5}
\]

Let us set \( u(t) = \mathbb{E}(N(t)) \). Then we can write (1.5) as

\[
u(t) = e^{-t} + 2\int_0^t u(t-s)e^{-s}ds. \tag{1.6}\]

Differentiating in \( t \) gives

\[
\frac{du(t)}{dt} = -e^{-t} + 2u(0)e^{-t} + 2\int_0^t \frac{du(t-s)}{dt}e^{-s}ds = e^{-t} - 2\int_0^t e^{-s}d(u(t-s))ds \tag{1.7}
\]

\[
= e^{-t} - 2e^{-t}u(0) + 2u(t) - 2\int_0^t e^{-s}u(t-s)ds.
\]
We used above the fact that $u(0) = 1$ and integrated by parts the last integral in the first line. Now, we use (1.6) to replace the integral in the very right side of (1.7), and obtain
\[
\frac{du(t)}{dt} = -e^{-t} + 2u(t) - (u(t) - e^{-t}) = u(t). \tag{1.8}
\]
Since $u(0) = 1$, we deduce that $u(t) = e^t$. □

In the general case, with $k$ children born with probabilities $p_k$, the analog of (1.5) is
\[
E(N(t)) = \mathbb{P}(\tau > t) + \sum_{k=2}^{\infty} k p_k E(N(t - \tau_1)) \mathbb{P}(\tau_1 < t)
\]
\[
= e^{-t} + \sum_{k=2}^{\infty} k p_k \int_0^t E(N(t - s)) \mathbb{P}(\tau_1 \in ds) = e^{-t} + \bar{N} \int_0^t E(N(t - s)) e^{-s} ds. \tag{1.9}
\]
Hence, the function $u(t)$ satisfies the following analog of (1.6):
\[
u(t) = e^{-t} + \bar{N} \int_0^t u(t - s)e^{-s} ds. \tag{1.10}
\]

**Exercise 1.2** Show that in this case $E N(t) = \exp((\bar{N} - 1)t)$.

We can actually say more about the total number of particles. Very naively, since the number of particles is very large, one may think that, in the spirit of the law of large numbers, $N(t)$ behaves as $N_\infty e^t$, as $t \to +\infty$, where $N_\infty$ would be a fixed deterministic constant.

**Exercise 1.3** Show this is impossible because, for example, the time of the first branching event is random and the above guess would indicate that $N(t)$ should also behave as $2N_\infty e^{t - \tau_1}$, which would be a contradiction.

However, the guess is not very far from the correct picture, except that $N_\infty$ is random.

**Proposition 1.4** Let $N(t)$ be the number of particles in the binary BBM, then
\[
M(t) = e^{-t} N(t) \tag{1.11}
\]
is a martingale. Moreover, $M(t)$ converges, as $t \to +\infty$, almost surely and in $L^1$ to a random variable $M_\infty$.

**Proof.** First, note that $M(t)$ is a positive random variable with $E(M(t)) = 1$. Hence, it is integrable. Let us first check that $M(t)$ is a martingale. To see this, we first note that for $t > s$ we have
\[
\mathbb{E}(M(t)|\mathcal{F}_s) = \mathbb{E}(e^{-t} N(t)|\mathcal{F}_s) = e^{-t} \mathbb{E}(N(t)|\mathcal{F}_s). \tag{1.12}
\]
Recall that the exponential clocks have the "lack of memory" property:
\[
\mathbb{P}(\tau > s + t|\tau > t) = \mathbb{P}(\tau > s). \tag{1.13}
\]
This is because
\[
\mathbb{P}(\tau > s + t | \tau > t) = \frac{\mathbb{P}(\tau > s + t \text{ and } \tau > t)}{\mathbb{P}(\tau > t)} = \frac{\mathbb{P}(\tau > s + t)}{\mathbb{P}(\tau > t)} = \frac{e^{-(t+s)}}{e^{-t}} \quad (1.14)
\]
\[
= e^{-s} = \mathbb{P}(\tau > s).
\]
Note that the lack of memory property is very specific to the exponential clocks, and that is one reason why using an exponential clock is convenient.

The lack of memory property means that we can reset all the clocks at the time \(s\) to zero, without changing the law of the process for \(t > s\), and deduce from Proposition 1.1 that
\[
\mathbb{E}(N(t) | \mathcal{F}_s) = N(s)e^{(t-s)}. \quad (1.15)
\]
Using this in (1.12) gives
\[
\mathbb{E}(M(t) | \mathcal{F}_s) = e^{-t}\mathbb{E}(N(t) | \mathcal{F}_s) = e^{-t}N(s)e^{(t-s)} = e^{-s}N(s) = M(s). \quad (1.16)
\]
Thus, \(M(t)\) is a martingale. Doob’s theorem implies that, as \(M(t)\) is a martingale, it converges almost surely, as \(t \to +\infty\) to a random limit \(M_\infty\). To prove converges in \(L^1\), it suffices to show that \(M(t)\) is uniformly integrable. For that, we can look at
\[
\mathbb{E}(M^2(t)) = e^{-2t}\mathbb{E}(N^2(t)). \quad (1.17)
\]
We have the following recursion relation for the function \(\phi_2(t) = \mathbb{E}(N^2(t))\), analogous to (1.5):
\[
\phi_2(t) = \mathbb{P}(\tau_1 < t) + \int_0^t (2\phi_2(t-s) + 2e^{2(t-s)})\mathbb{P}(\tau_1 \in ds)
\]
\[
= e^{-t} + 2 \int_0^t (\phi_2(t-s) + e^{2(t-s)})e^{-s}ds. \quad (1.18)
\]
Hence, the function
\[
\psi_2(t) = \phi_2(t) + e^{2t} \quad (1.19)
\]
satisfies
\[
\psi_2(t) = e^{-t} + e^{2t} + 2 \int_0^t \psi_2(t-s)e^{-s}ds. \quad (1.20)
\]
Differentiating in \(t\) gives
\[
\frac{d\psi_2(t)}{dt} = -e^{-t} + 2e^{2t} + 2e^{-t}\psi_2(0) + 2 \int_0^t \frac{d\psi_2(t-s)}{dt}e^{-s}ds
\]
\[
= 3e^{-t} + 2e^{2t} - 2 \int_0^t e^{-s}\frac{d\psi_2(t-s)}{ds}ds = 3e^{-t} + 2e^{2t} - 2e^{-t}\psi_2(0) + 2\psi_2(t) \quad (1.21)
\]
\[
- 2 \int_0^t e^{-s}\psi_2(t-s).
\]
Inserting (1.20) into (1.21) and using the fact that \(\psi_2(0) = 2\), gives
\[
\frac{d\psi_2(t)}{dt} = 2e^{2t} - e^{-t} + 2\psi_2(t) - (\psi_2(t) - e^{-t} - e^{2t}) = 3e^{2t} + \psi_2(t). \quad (1.22)
\]
It follows that
\[ \psi_2(t) = 3e^{2t} - e^t, \]  
and
\[ \phi_2(t) = 2e^{2t} - e^t. \]  
This gives
\[ \gamma_2(t) := \mathbb{E}(M^2(t)) = e^{-2t}\phi_2(t) = 2 - e^{-t}. \]  
Thus, \( M(t) \) is uniformly integrable, and converges to \( M_\infty \) in \( L^1 \), as well. □

We can use the above recursive strategy to compute higher moments of \( N(t) \). Let us set
\[ \phi_k(t) = \mathbb{E}(N^k(t)), \quad \gamma_k(t) = \mathbb{E}(M^k(t)) = e^{-kt}\phi_k(t). \]
These functions satisfy the following analog of (1.18):
\[ \phi_k(t) = \mathbb{P}(\tau_1 < t) + \int_0^t \sum_{m=0}^k \binom{k}{m} \phi_m(t-s)\phi_{k-m}(t-s)\mathbb{P}(\tau_1 \in ds) 
= e^{-t} + \int_0^t \sum_{m=0}^k \binom{k}{m} \phi_m(t-s)\phi_{k-m}(t-s)e^{-s}ds, \]  
and
\[ \gamma_k(t)e^{kt} = \mathbb{P}(\tau_1 < t) + \int_0^t \sum_{m=0}^k \binom{k}{m} e^{m(t-s)}\gamma_m(t-s)\gamma_{k-m}(t-s)e^{(k-m)(t-s)}e^{-s}ds. \]  
The last equation becomes
\[ \gamma_k(t) = e^{-(k+1)t} + \int_0^t \sum_{m=0}^k \binom{k}{m} \gamma_m(t-s)\gamma_{k-m}(t-s)e^{(k-m)(t-s)}e^{-s}ds. \]  
Let us assume that the limits
\[ \bar{\gamma}_k = \lim_{t \to +\infty} \gamma_k(t) \]  
exist. Then, passing to the limit \( t \to +\infty \) in (1.28) leads to a recursive equation
\[ \bar{\gamma}_k = \frac{1}{k+1} \sum_{m=0}^k \binom{k}{m} \bar{\gamma}_m \bar{\gamma}_{k-m} = \frac{k!}{k+1} \sum_{m=0}^k \binom{k}{m} \bar{\gamma}_m \bar{\gamma}_{k-m}. \]  
We have already computed that \( \bar{\gamma}_1 = 1 \) and \( \bar{\gamma}_2 = 2 \). It is immediate to see from (1.30) that
\[ \bar{\gamma}_k = k!. \]  

**Exercise 1.5** Show that \( M_\infty \) is an exponential variable with parameter 1. Explain why \( M_\infty \) is determined by what happens early in the branching process.

**Exercise 1.6** Generalize the claims of Proposition 1.4 and Exercise 1.5 to a general non-binary branching with probabilities \( p_k \) to have \( k \) children at each branching event, as in (1.2).
2 Maximum of a family of independent particles

A consequence of Proposition 1.4 is that we can think of \(x_1(t), \ldots, x_{N(t)}(t)\) as a collection of approximately \(M_\infty e^t\) random particles, for \(t \gg 1\), even though \(M_\infty\) itself is random. An important point is that the random variables \(x_1(t), \ldots, x_{N(t)}(t)\) are correlated, and the nature of these correlations will be crucial to us. In particular, the correlations will be very important when we consider the extrema of such branching processes.

As a first step in this direction, in this section we consider a much simplified model where no correlations are present. Let us think of \(N \in \mathbb{N}\) as an analog of the time variable for the BBM and take a large number \(M\) of independent particles \(X_1, \ldots, X_M\), with \(M\) that depends on \(N\). In order to mimic the concept that \(N\) is the time of the BBM, we assume that \(X_k\) are mean zero Gaussian random variables with variance \(N\). Thus, we can think of each \(X_k\) as a snapshot of a Brownian motion at the time \(t = N\). To make sure that the number \(M\) of these variables also mimics BBM, we assume that it grows exponentially in \(N\), as is the case for BBM. We take \(M = 2^N\), to ensure that \(M\) is an integer. To complete the model we need to prescribe the correlations between \(X_k\). The key difference with the BBM is that here, for an infinitely greater simplicity, we assume that all \(X_k\) are independent. This assumption will allow us to calculate many quantities explicitly.

To summarize, one can think of this model as an "uncorrelated BBM", in the sense that at the time \(N\) we have \(2^N\) particles and each one has variance \(N\), that of the standard Brownian motion. We are interested in the large \(N\) behavior of this system.

2.1 The law of the maximum

The first object we would like to understand is the location of the maximal particle

\[
\bar{M}_N = \max(X_1, \ldots, X_M).
\] (2.1)

In this simple model, since all particles are identical, this can be done by writing

\[
\mathbb{P}(\bar{M}_N < y) = \left(\mathbb{P}(X_1 < y)\right)^2 = \left(1 - \mathbb{P}(X_1 > y)\right)^2 = e^{-2N}.
\] (2.2)

The first question is for which \(y\) is this probability of order one. In particular, we are interested in the median location \(m_N\) such that

\[
\mathbb{P}(\bar{M}_N > m_N) = 1/2.
\] (2.3)

The second is the width of the transition layer: if we fix \(\varepsilon_0 \in (0, 1)\) and consider \(Y_N(\varepsilon_0)\) defined by

\[
\mathbb{P}(\bar{M}_N > Y_N) = \varepsilon_0,
\] (2.4)

the question is if the difference

\[
|Y_N(\varepsilon_0) - m_N|
\] (2.5)

grows in \(N\) or stays of order 1 as \(N \to +\infty\), with \(\varepsilon_0 \in (0, 1)\) fixed.

We see from (2.2) that for \(\mathbb{P}(\bar{M}_N < y)\) to be of order one but not too close to 1, we need \(\mathbb{P}(X_1 > y)\) to be of the order \(2^{-N}\):

\[
2^N \mathbb{P}(X_1 > y) \sim O(1).
\] (2.6)
Note that for large \( y \) we have

\[
P(X_1 > y) = \int_y^\infty e^{-x^2/(2N)} \frac{dx}{\sqrt{2\pi N}} = \int_{y/\sqrt{2N}}^\infty e^{-x^2} \frac{dx}{\sqrt{\pi}} = (1 + o(1)) \frac{\sqrt{N}}{y\sqrt{2\pi}} \int_{y/\sqrt{2N}}^\infty 2x e^{-x^2} \, dx
\]

\[
= (1 + o(1)) \frac{\sqrt{N}}{y\sqrt{2\pi}} e^{-y^2/(2N)}, \quad \text{as } N \to \infty.
\]

Thus, we are looking for \( y \) such that

\[
2^N \frac{\sqrt{N}}{y\sqrt{2\pi}} e^{-y^2/(2N)} \sim O(1),
\]

which means that, to the leading order, we should be looking at \( y \) such that

\[
\frac{y^2}{2N} \approx N \log 2, \quad y \approx \sqrt{2 \log 2 N}.
\]

Note that \( y \) is growing linearly in the "time" \( N \), and not diffusively, which would be \( \sqrt{N} \):

\[
y \approx c_* N, \quad c_* = \sqrt{2 \log 2}.
\]

Going back to (2.8), we see that the next order correction requires that

\[
2^N \sqrt{N} e^{-y^2/(2N)} \sim y \approx \sqrt{2 \log 2 N}.
\]

This gives

\[
N \log 2 - \frac{y^2}{2N} \approx \frac{1}{2} \log N,
\]

meaning that

\[
y \approx \left(2N^2 \log 2 - N \log N\right)^{1/2} = N \sqrt{2 \log 2} \left(1 - \frac{\log N}{2N \log 2}\right)^{1/2}
\]

\[
\approx N \sqrt{2 \log 2} \left(1 - \frac{\log N}{4N \log 2}\right) = N \sqrt{2 \log 2} - \frac{1}{2 \sqrt{2 \log 2}} \log N.
\]

Let us set

\[
c_* = \sqrt{2 \log 2}, \quad \lambda_* = \sqrt{2 \log 2},
\]

then (2.13) can be written as

\[
y \approx c_* N - \frac{1}{2\lambda_*} \log N.
\]

The fact that \( \lambda_* = c_* \) in (2.14) is an unfortunate coincidence and should be ignored. They do not always coincide in other models and their roles are different, as we will see. This structure will appear repeatedly later on, and, for reasons to become clear later, the pre-factor 1/2 in (2.13) indicates that the particles are uncorrelated.
Exercise 2.1 Make the above approximations precise, controlling the errors and show that there exists $x_0$ so that the median $m_N$ defined in (2.3) satisfies

$$m_N = c_s N - \frac{1}{2\lambda_s} \log N + x_0 + o(1), \text{ as } N \to +\infty.$$  

(2.16)

(ii) More generally, show that for any $\varepsilon_0 > 0$ there exists $x_\varepsilon$ so that $Y_N(\varepsilon_0)$ defined in (2.4) satisfies

$$Y_N(\varepsilon_0) = c_s N - \frac{1}{2\lambda_s} \log N + x_\varepsilon + o(1), \text{ as } N \to +\infty.$$  

(2.17)

Now that we know how we should choose the centering for $y$, let us consider the function

$$u_N(y) = \mathbb{P}(\max_{1 \leq k \leq M} X_k > c_s N - \frac{1}{2\lambda_s} \log N + y).$$  

(2.18)

Using (2.7), we obtain

$$2^N \mathbb{P}(X_1 > c_s N - \frac{1}{2\lambda_s} \log N + y)$$

$$= \frac{\sqrt{N} 2^N}{(c_s N - (1/(2\lambda_s)) \log N + y) \sqrt{2\pi}} e^{-(c_s N - (1/(2\lambda_s)) \log N + y)^2/(2N)} + o(1)$$

$$= \frac{1}{c_s \sqrt{2\pi} N} \frac{\log N}{2\lambda_s \sqrt{N}} \frac{y}{\sqrt{N}} \exp \left( N \log 2 - \frac{c_s^2}{2} N + \frac{c_s}{2\lambda_s} \log N - c_s y \right) + o(1)$$

$$= \frac{1}{c_s \sqrt{2\pi}} e^{-\lambda_s y} + o(1), \text{ as } N \to +\infty.$$  

(2.19)

Going back to (2.2), we see that

$$1 - u_N(y) = \left(1 - \mathbb{P}(X_1 > c_s N - \frac{1}{2\lambda_s} \log N + y)\right)^{2^N} = \left(1 - 2^{-N} \frac{1}{c_s \sqrt{2\pi}} e^{-\lambda_s y}\right)^{2^N} + o(1)$$

$$= \exp \left(- \frac{1}{c_s \sqrt{2\pi}} e^{-\lambda_s y}\right) + o(1), \text{ as } N \to +\infty,$$  

(2.20)

so that

$$u_N(y) = 1 - \exp \left(- \frac{1}{c_s \sqrt{2\pi}} e^{-\lambda_s y}\right) + o(1), \text{ as } N \to +\infty.$$  

(2.21)

Note that the leading order term in the right side above is the Gumbel distribution.

Exercise 2.2 We have used both large $y$ asymptotics in (2.7), and also large $N$ asymptotics later on. Check the above approximations, to make sure that $y$ taken in (2.13) is sufficiently large, so that (2.21) does hold for all $y \in \mathbb{R}$.

2.2 An upper bound for the maximum: the first moment method

The above computation for the location of the maximum, and the asymptotics in (2.21) is a little too explicit for an analyst’s taste. Let us now, instead, get some bounds on the location of the maximum. They will not be as precise as (2.21) but will allow us to introduce two
methods to bound the distribution of the maximum from above and from below that will be useful in the analysis of the maximum of BBM.

For the upper bound, we will use what is known as the first moment method. The total number of particles located to the right of a given \( y \in \mathbb{R} \) can be written as follows (recall that \( M = 2^N \)):

\[
M_N(y) = \sum_{k=1}^{M} \mathbb{1}(X_k > y).
\] (2.22)

Note that we have, on one hand,

\[
\mathbb{P}( \max_{1 \leq k \leq M} X_k > y) = \mathbb{P}(M_N(y) \geq 1),
\] (2.23)

and on the other,

\[
\mathbb{E}(M_N(y)) = \sum_{k=1}^{M} k\mathbb{P}(M_N(y) = k) \geq \sum_{k=1}^{M} \mathbb{P}(M_N(y) = k) = \mathbb{P}(M_N(y) \geq 1).
\] (2.24)

This gives the following ballpark estimate which is the essence of the first moment method:

\[
\mathbb{P}( \max_{1 \leq k \leq M} X_k > y) \leq \mathbb{E}(M_N(y)) = 2^N \mathbb{P}(X_1 > y).
\] (2.25)

We used the fact that the particles are identically distributed in the last step.

This allows us to bound the function \( u_N(y) \) introduced in (2.18) using (2.19):

\[
u_N(y) = \mathbb{P}( \max_{1 \leq k \leq M} X_k > c_\ast N - \frac{1}{2\lambda_\ast} \log N + y) \\
\leq 2^N \mathbb{P}(X_1 > c_\ast N - \frac{1}{2\lambda_\ast} \log N + y) = \frac{1}{c_\ast \sqrt{2\pi}} e^{-\lambda_\ast y} + o(1), \quad \text{as } N \to +\infty.
\] (2.26)

Thus, we have an upper bound

\[
u_N(y) = \mathbb{P}( \max_{1 \leq k \leq M} X_k > c_\ast N - \frac{1}{2\lambda_\ast} \log N + y) \leq \frac{1}{c_\ast \sqrt{2\pi}} e^{-\lambda_\ast y} + o(1), \quad \text{as } N \to +\infty.
\] (2.27)

### 2.3 A lower bound for the maximum: the second moment method

We may also get a lower bound for \( u_N(x) \), using what is known as the second moment method. The starting point is, again, (2.23):

\[
\mathbb{P}( \max_{1 \leq k \leq M} X_k > y) = \mathbb{P}(M_N(y) \geq 1).
\] (2.28)

The second moment method estimates the probability in the right side of (2.28) using the Cauchy-Schwartz inequality

\[
(\mathbb{E}M_N(y))^2 = (\mathbb{E}[M_N(y) \mathbb{1}(M_N(y) \geq 1)])^2 \leq \mathbb{E}(M_N(y))^2 \mathbb{E}(\mathbb{1}(M_N(y) \geq 1)) = \mathbb{E}(M_N(y))^2 \mathbb{P}(M_N(y) \geq 1).
\] (2.29)
It follows that
\[ \mathbb{P}( \max_{1 \leq k \leq M} X_k > y) = \mathbb{P}(M_N(y) \geq 1) \geq \frac{(\mathbb{E}M_N(y))^2}{\mathbb{E}(M_N(y))^2}. \] (2.30)
This is convenient as it reduces the computation of a probability to a computation of moments.

The dominator in (2.30) can be written as
\[ \mathbb{E}(M_N(y)^2) = \mathbb{E}\left( \sum_{k=1}^{M} \mathbb{1}(X_k > y) \right)^2 = 2^N\mathbb{P}(X_1 > y) + 2^N(2^N - 1)[\mathbb{P}(X_1 > y)]^2 \] (2.31)
\[ = \mathbb{E}(M_N(y)) + (1 - 2^{-N})(\mathbb{E}(M_N(y)))^2. \]

Hence, if \( y \) in (2.31) is such that \( \mathbb{E}(M_N(y)) \sim O(1) \), then we have
\[ \mathbb{E}(M_N(y)) \sim O(1), \] (2.32)
then we have
\[ \mathbb{P}( \max_{1 \leq k \leq M} X_k > y) \geq \frac{(\mathbb{E}M_N(y))^2}{\mathbb{E}(M_N(y))^2} = \frac{(\mathbb{E}M_N(y))^2}{\mathbb{E}(M_N(y)) + (1 - 2^{-N})(\mathbb{E}(M_N(y)))^2} \] (2.33)
\[ = \frac{\mathbb{E}M_N(y)}{1 + \mathbb{E}(M_N(y))} + o(1), \text{ as } N \to +\infty. \]

In order to have (2.32), we take \( y \) as in (2.26):
\[ \mathbb{E}(M_N(c_\ast N - \frac{1}{2\lambda_\ast} \log N + y)) = \frac{1}{c_\ast \sqrt{2\pi}} e^{-\lambda_\ast y} + o(1), \text{ as } N \to +\infty. \] (2.34)

The following lower bound follows then from (2.33):
\[ \mathbb{P}( \max_{1 \leq k \leq M} X_k > c_\ast N - \frac{1}{2\lambda_\ast} \log N + y) \geq \frac{e^{-\lambda_\ast y}}{c_\ast \sqrt{2\pi} + e^{-\lambda_\ast y}} + o(1), \text{ as } N \to +\infty. \] (2.35)

Summarizing, we have now obtained upper and lower bounds for the function \( u_N(y) \):
\[ \frac{e^{-\lambda_\ast y}}{c_\ast \sqrt{2\pi} + e^{-\lambda_\ast y}} + o(1) \leq u_N(y) \leq \frac{1}{c_\ast \sqrt{2\pi}} e^{-\lambda_\ast y} + o(1), \text{ as } N \to +\infty. \] (2.36)

These bounds match as \( y \to +\infty \), so that, in particular, we have
\[ e^{\lambda_\ast y}u_N(y) = \frac{1}{c_\ast \sqrt{2\pi}} + o(1), \text{ as } N \to +\infty \text{ and } y \to +\infty. \] (2.37)

This tells us that the maximum of \( X_1, \ldots, X_M \) is located around
\[ m_N = c_\ast N - \frac{1}{2\lambda_\ast} \log N \] (2.38)
and gives us the limiting distribution when \( y \) is centered at that location. Of course, these estimates are less precise than the exact expression (2.21) but they give a very good approximation to it.
2.4 The extremal process

Let us now try to understand how the collection $X_1, X_2, \ldots, X_M$, $M = 2^N$, looks like as a cloud of points on $\mathbb{R}$. We recenter $X_k$ according to the typical location of the maximum in (2.38): set $Z_k$ by

$$X_k = m_N + Z_k.$$  

(2.39)

Here, $m_N$ is taken to be

$$m_N = c_* N - \frac{1}{2 \lambda_*} \log N;$$  

(2.40)

with

$$c_* = \lambda_* = \sqrt{2 \log 2}.$$  

(2.41)

Recall that the median $\bar{m}_N$ of the maximum $\bar{M}_N = \max_{1 \leq k \leq M} X_k$, defined as,

$$\mathbb{P}(\bar{M}_N > \bar{m}_N) = 1/2,$$  

(2.42)

satisfies the asymptotics (2.16)

$$\bar{m}_N = c_* N - \frac{1}{2 \lambda_*} \log N + x_0 + o(1), \text{ as } N \to +\infty.$$  

(2.43)

Hence, the difference between $\bar{m}_N$ and $m_N$ is just the constant $x_0$ that is convenient to eliminate, to shorten some computations.

**Exercise 2.3** Show that there exists $\varepsilon_0 \in (0, 1)$ so that

$$\mathbb{P}(\bar{M}_N > m_N) = \varepsilon_0 + o(1), \text{ as } N \to +\infty.$$  

(2.44)

Consider the measure, called a random point process

$$\mathcal{E}_N(x) = \sum_{k=1}^{M} \delta(x - Z_k).$$  

(2.45)

The process $\mathcal{E}_N$ is also known as the extremal process because it is re-centered near the maximum of $X_N$. One should note the difference between the process $\mathcal{E}_N$ coming from $Z_k$ that are centered at a deterministic location $m_N$ given by (2.40), and the process $\bar{\mathcal{E}}_N$ seen from the tip:

$$\bar{\mathcal{E}}_N(x) = \sum_{k=1}^{M} \delta(x - \bar{Z}_k),$$  

(2.46)

generated by

$$\bar{Z}_k = X_k - \bar{M}_N, \quad \bar{M}_N = \max_{1 \leq k \leq M} X_k.$$  

(2.47)

The points $\bar{Z}_k$ are centered at a random location $\bar{M}_N$. Here, we will focus on $\mathcal{E}_N$.

Note that the sum in (2.45) involves $M = 2^N$ terms, so the total mass of $\mathcal{E}_N$ is $2^N$. However, most of the points $Z_k$ are very large: each of $Z_k$ is a Gaussian with mean $(-m_N)$ and variance $\sqrt{N}$. Thus, relatively few of $Z_k$ will lie in any given compact set, hence it is perfectly plausible that the measure $\mathcal{E}_N(x)$ may have a weak limit as $N \to +\infty$, even though
its total mass blows up as $N \to +\infty$. We will see how this blowup of the mass is reflected in the nature of the limit.

We will study $\mathcal{E}_N$ via its Laplace transform. The Laplace transform of a random point process $\mathcal{E}_N$ is defined as follows: given a test function $\phi \in C_b^+ (\mathbb{R})$ (the set of bounded non-negative continuous functions), set

$$
\Psi_N (\phi) = \mathbb{E} \left( \exp \left[ - \int \phi (x) \mathcal{E}_N (x) dx \right] \right) = \mathbb{E} \left( \exp \left[ - \sum_{k=1}^{M} \phi (Z_k) \right] \right). \tag{2.48}
$$

A basic result in the theory of point processes is that the Laplace transform determines the law of a point process, and that convergence of the Laplace transforms implies convergence in law of the point processes to some limit – see Appendix to [2].

In the present case, the asymptotics of the Laplace transform of $Z_1, \ldots, Z_M$ can be computed essentially explicitly in the limit $N \to +\infty$. First, we write

$$
\Psi_N (\phi) = \mathbb{E} \left( \exp \left[ - \sum_{k=1}^{2N} \phi (Z_k) \right] \right) = \left[ \mathbb{E} (\exp (-\phi (Z_1))) \right]^{2N}, \tag{2.49}
$$

and decompose

$$
\mathbb{E} (\exp (-\phi (Z_1))) = 1 + \left[ \mathbb{E} (\exp (-\phi (Z_1))) - 1 \right]. \tag{2.50}
$$

Let us look at the second term in the parentheses: recall that $Z_1$ is a Gaussian with the mean $(-m_N)$ and variance $N$. Hence, we have

$$
\mathbb{E} (\exp (-\phi (Z_1))) - 1 = \frac{1}{\sqrt{2\pi N}} \int e^{-(x+m_N)^2/2N} \left[ e^{-\phi (x)} - 1 \right] dx
= \frac{1}{\sqrt{2\pi N}} \int \exp \left\{ - \frac{x^2}{2N} - \frac{m_N^2}{2N} - \frac{m_N x}{N} \right\} \left[ e^{-\phi (x)} - 1 \right] dx
= \frac{1}{\sqrt{2\pi N}} \int \exp \left\{ - \frac{c_*^2 N^2}{2} - \left( c_* / \lambda_* \right) N \log N + 1/(4\lambda_*^2) \log^2 N \right\} \left[ e^{-\phi (x)} - 1 \right] dx
\times \exp \left( - \frac{x^2}{2N} \right) \left[ e^{-\phi (x)} - 1 \right] dx.
$$

Recall that $c_* = \sqrt{2 \log 2}$ and $\lambda_* = c_*$. This gives

$$
\mathbb{E} (\exp (-\phi (Z_1))) - 1 = \frac{2^{-N}}{\sqrt{2\pi}} \int e^{-\lambda_* x} \left[ e^{-\phi (x)} - 1 \right] dx + o(2^{-N}), \quad \text{as } N \to +\infty. \tag{2.52}
$$

Thus, the second term in the parentheses in (2.50) is of the order $2^{-N}$. Hence we have, combining (2.49), (2.50) and (2.52):

$$
\Psi_N (\phi) = \left( 1 + \left[ \mathbb{E} (\exp (-\phi (Z_1))) - 1 \right] \right)^{2N}
= \exp \left( 2^N \left[ \mathbb{E} (\exp (-\phi (Z_1))) - 1 \right] \right) + o(1)
\quad \text{as } N \to +\infty. \tag{2.53}
$$
We deduce that the process $\mathcal{E}_N(x)$ converges as $N \to +\infty$ to a point process with the Laplace transform

$$\Psi(\phi) = \exp\left( \int e^{-\lambda_* x} \left[ e^{-\phi(x)} - 1 \right] \frac{dx}{\sqrt{2\pi}} \right). \quad (2.54)$$

In order to identify this limit, let us recall the definition of a Poisson point process with intensity $\mu$. Here, $\mu$ is a $\sigma$-finite non-negative measure on $\mathbb{R}^d$. The Poisson point process $N$ with intensity $\mu$ is characterized by the following properties: first, given any Borel set $B \subset \mathbb{R}^d$, we have, for any $k \geq 0$:

$$P[N(B) = k] = \frac{(\mu(B))^k}{k!} e^{-\mu(B)}, \quad \text{if } \mu(B) < +\infty,$$

$$P[N(B) = k] = 0, \quad \text{if } \mu(B) = +\infty. \quad (2.55)$$

Second, if $A$ and $B$ are two Borel sets such that $A \cap B = \emptyset$, then $N(A)$ and $N(B)$ are independent random variables. The Laplace transform of a Poisson point process is described by the following exercise.

**Exercise 2.4** Let $N$ be a Poisson point process with intensity $\mu$, then its Laplace transform is defined, by its action on non-negative continuous bounded functions as

$$\Psi(\phi) = \exp\left( \int \left[ e^{-\phi(x)} - 1 \right] \mu(dx) \right). \quad (2.56)$$

It is helpful to start with non-negative simple functions (even though they are not continuous). If in doubt, this is Proposition 4.8 of [2].

Comparing the limit (2.54) of the Laplace transforms $\Psi_N$ of the point processes $\mathcal{E}_N$ and (2.56), we see that we have proved the following.

**Theorem 2.5** The point processes

$$\mathcal{E}_N = \sum_{k=1}^{2^N} \delta(x - Z_k), \quad (2.57)$$

converges in law as $N \to +\infty$ to a Poisson point process with intensity

$$\mu(x) = \frac{1}{\sqrt{2\pi}} e^{-\lambda_* x}. \quad (2.58)$$

Thus, $\lambda_*$ is the exponential rate of growth of the density of the random point process associated to $X_1, \ldots, X_M$. Note that the limit process has an infinite mass, and that the density of the measure $\mu(x)$ grows exponentially as $x \to -\infty$. The former property reflects the fact that the original extremal point process $\mathcal{E}_N$, before the limit $N \to +\infty$, had the mass $2^N$. The latter comes from the fact that most of the particles $X_k$ are located far to the left of the maximum $\bar{M}_N$.

**Exercise 2.6** Perform a similar analysis of the extremal process seen from the tip, that is, $\bar{\mathcal{E}}_N$, defined in (2.46), and describe the limit of $\bar{\mathcal{E}}_N$.

**Exercise 2.7** Perform the analysis of this section for independent Gaussian random variables with mean zero and variance $\sigma^2 N$. Compute the resulting $c_*(\sigma)$ and $\lambda_*(\sigma)$. 

13
3 The connection of the BBM to partial differential equations

Before continuing with the random processes points of view, let us describe the connection between branching random processes and nonlinear partial differential equations. The starting point is the following fact. Let $B_t$ be the $d$-dimensional Brownian motion with variance $\sigma$, and $g(x)$ be a bounded function. Consider the function

$$u(t, x) = \mathbb{E}_x g(B(t)).$$ (3.1)

The notation in the right side means that $B(t)$ starts at the point $x$ at $t = 0$. A key property of the Brownian motion we will be using is that $u(t, x)$ satisfies the heat equation

$$\frac{\partial u}{\partial t} = \frac{\sigma^2}{2} \Delta u, \quad t > 0, \quad x \in \mathbb{R}^d,$$ (3.2)

with the initial condition $u(0, x) = g(x)$. This result can be found in numerous textbooks on stochastic analysis.

3.1 Discrete linear equations and random walks

Let us first explain informally how the linear partial differential equation (3.2) can be explained in the context of random walks, in a very simple way. The starting point is a discrete time Markov jump process $X_{n\tau}$, with a time step $\tau > 0$, defined on one-dimensional a lattice with mesh size $h$:

$$h\mathbb{Z} = \{0, \pm h, \pm 2h, \ldots \}.$$

The particle position evolves as follows: if the particle is located at a position $x \in h\mathbb{Z}$ at the time $t = n\tau$ then at the time $t = (n + 1)\tau$ it jumps to a random position $y \in h\mathbb{Z}$, with the transition probability

$$P(X_{(n+1)\tau} = y \mid X_{n\tau} = x) = k(x - y), \quad x, y \in h\mathbb{Z}.$$ (3.3)

Here, $k(x)$ is a prescribed non-negative kernel such that

$$\sum_{y \in h\mathbb{Z}} k(y) = 1.$$ (3.4)

The classical symmetric random walk with a spatial step $h$ and a time step $\tau$ corresponds to the choice $k(\pm h) = 1/2$, and $k(y) = 0$ otherwise – the particle may only jump to the nearest neighbor on the left and on the right, with equal probabilities.

In order to connect this process to an evolution equation, let us take a function $f : h\mathbb{Z} \to \mathbb{R}$, defined on the lattice, and introduce

$$u(t, x) = \mathbb{E}(f(X_t(x))).$$ (3.5)

Here, $X_t(x)$, $t \in \tau\mathbb{N}$, is the above Markov process starting at a position $X_0(x) = x \in h\mathbb{Z}$ at the time $t = 0$. If $f \geq 0$ then one may think of $u(t, x)$ as the expected value of a “prize” to
be collected at the time \( t \) at a (random) location of \( X_t(x) \) given that the process starts at the point \( x \) at the time \( t = 0 \). An important special case is when \( f \) is the characteristic function of a set \( A \). Then, \( u(t,x) \) is the probability that the jump process \( X_t(x) \) that starts at the position \( X_0 = x \) is inside the set \( A \) at the time \( t \).

As the process \( X_t(x) \) is Markov, the function \( u(t,x) \) satisfies the following relation

\[
\begin{align*}
\frac{\partial u}{\partial t}(t,x) &= a \frac{\partial^2 u}{\partial x^2}(t,x),
\end{align*}
\]

(3.8)

This is because after the initial step when the particle jumps at the time \( \tau \) from the starting position \( x \) to a random position \( y \), the process “starts anew”, and runs for time \( \tau \) between the times \( \tau \) and \( t + \tau \) – this is where the second equality in (3.6) comes from. The third equality simply uses the definition (3.3) of \( k(x) \). Equation (3.6) can be re-written, using (3.4) as

\[
\begin{align*}
\frac{\partial u}{\partial t}(t,x) &= a \frac{\partial^2 u}{\partial x^2}(t,x),
\end{align*}
\]

(3.7)

The key point is that the discrete equation (3.7) leads to various interesting continuous limits as \( h \downarrow 0 \) and \( \tau \downarrow 0 \), depending on the choice of the transition kernel \( k(y) \), and on the relative size of the spatial mesh size \( h \) and the time step \( \tau \). In other words, depending on the microscopic model – the particular properties of the random walk – we will end up with different macroscopic continuous models.

**The heat equation and random walks**

Let us show how this can be done to obtain the heat equation

\[
\begin{align*}
\frac{\partial u}{\partial t} &= a \frac{\partial^2 u}{\partial x^2},
\end{align*}
\]

(3.8)

with a constant diffusivity constant \( a > 0 \). This gives an informal explanation for the interpretation (3.1)-(3.2) of the heat equation in terms of Brownian motion. We will assume that the transition probability kernel has the form

\[
\begin{align*}
k(x) &= \phi \left( \frac{x}{h} \right), \quad x \in h \mathbb{Z},
\end{align*}
\]

(3.9)

with a non-negative function \( \phi(m) \geq 0 \) defined on \( \mathbb{Z} \), such that

\[
\begin{align*}
\sum_{m} \phi(m) &= 1.
\end{align*}
\]

(3.10)

This form of \( k(x) \) allows us to re-write (3.7) as

\[
\begin{align*}
\frac{\partial u}{\partial t}(t,x) &= a \frac{\partial^2 u}{\partial x^2}(t,x),
\end{align*}
\]

(3.11)

or, equivalently,

\[
\begin{align*}
\sum_{m} \phi(m)[u(t, x - mh) - u(t, x)] = \sum_{y \in h \mathbb{Z}} \phi \left( \frac{x - y}{h} \right) [u(t, y) - u(t, x)],
\end{align*}
\]

(3.12)
In order to arrive to the heat equation in the limit, we will make the assumption that jumps are symmetric on average:
\[
\sum_{m \in \mathbb{Z}} m \phi(m) = 0. \quad (3.13)
\]
Then, expanding the right side of (3.12) in \( h \) and the left side in \( \tau \), we obtain
\[
\tau \frac{\partial u(t, x)}{\partial t} = \frac{ah^2}{2} \frac{\partial^2 u}{\partial x^2}(t, x) + \text{lower order terms}, \quad (3.14)
\]
with
\[
a = \sum_{m} |m|^2 \phi(m). \quad (3.15)
\]
To balance the left and the right sides of (3.14), we need to take the time step \( \tau = h^2 \) – note that the scaling \( \tau = O(h^2) \) is essentially forced on us if we want to balance the two sides of this equation. Then, in the limit \( \tau = h^2 \downarrow 0 \), we obtain the heat equation
\[
\frac{\partial u(t, x)}{\partial t} = a \frac{\partial^2 u(t, x)}{2 \partial x^2}. \quad (3.16)
\]
The diffusion coefficient \( a \) given by (3.15) is the second moment of the jump size – in other words, it measures the “overall jumpiness” of the particles. This is a very simple example of how the microscopic information, the kernel \( \phi(m) \), translates into a macroscopic quantity – the overall diffusion coefficient \( a \) in the macroscopic equation (3.16).

**Exercise 3.1** Show that if (3.13) is violated and
\[
b = \sum_{m \in \mathbb{Z}} m \phi(m) \neq 0, \quad (3.17)
\]
then one may take \( \tau = h \), and the formal limit of (3.12) is the advection equation
\[
\frac{\partial u(t, x)}{\partial t} + b \frac{\partial u(t, x)}{\partial x} = 0, \quad (3.18)
\]
without any diffusion.

**Exercise 3.2** Relate the limit in (3.18) to the law of large numbers and explain the relation \( \tau = h \) in these terms. How can (3.16) and the relation \( \tau = h^2 \) between the temporal and spatial steps be explained in terms of the central limit theorem?

### 3.2 Linear parabolic equations and branching random walks

Let us now explain how we can obtain a probabilistic interpretation for a parabolic equation with a zero-order term:
\[
\frac{\partial u}{\partial t} = \frac{\sigma^2}{2} \frac{\partial^2 u}{\partial x^2} + mu, \quad (3.19)
\]
with some \( m \in \mathbb{R} \) fixed. Let us consider a branching Brownian motion of variance \( \sigma > 0 \), with an exponential clock such that
\[
P(\tau > t) = e^{-t}. \quad (3.20)
\]
We consider general branching so that a particle branches into \( k \) children with probabilities \( p_k \), so that

\[
\sum_{k=1}^{\infty} p_k = 1,
\]

and the average number of off-spring is

\[
\bar{N} = \sum_{k=1}^{\infty} k p_k. \tag{3.21}
\]

Let us assume that the BBM starts at \( t = 0 \) at the position \( x \) and denote the locations of the BBM particles at the time \( t > 0 \) by \( X_1(t), \ldots, X_{N_t}(t) \). Given a bounded function \( g(x) \), consider the function

\[
u(t, x) = E_x \sum_{k=1}^{N_t} g(X_k(t)). \tag{3.22}\]

This is the analog of (3.1) that led to the standard heat equation in the case of the standard Brownian motion, with no branching. In order to get an equation for \( u(t, x) \) let us write a recursive relation, very similar to what we have seen in the proofs of Propositions 1.1 and 1.4. Looking at the first branching event gives the renewal relation

\[
u(t, x) = E_x(g(B_t))P(\tau_1 > t) + \sum_{k=1}^{\infty} k p_k \int_0^t E_x(u(t-s, B_s))P(\tau_1 \in ds) \tag{3.23}\]

Note that (3.1)-(3.2) implies that the function

\[
u(t, x) = E_x(g(B_t))\]

is the solution to the heat equation

\[
\frac{\partial \nu}{\partial t} = \frac{\sigma^2}{2} \Delta \nu, \tag{3.24}\]

with the initial condition \( \nu(0, x) = g(x) \). Hence, it can be written as

\[
u(t, x) = [e^{\sigma^2 t \Delta / 2} g(\cdot)](x). \tag{3.25}\]

In addition, for \( 0 < s < t \) fixed, the function

\[
w(\tau, x) = E_x(u(t-s, B_\tau)) \tag{3.26}\]

is the solution to the heat equation

\[
\frac{\partial w}{\partial \tau} = \frac{\sigma^2}{2} \Delta w, \tag{3.27}\]

with the initial condition \( w(0, x) = u(t-s, x) \). That is, we have

\[
w(\tau, x) = [e^{\sigma^2 \tau \Delta / 2} u(\tau-s, \cdot)](x). \tag{3.28}\]
It follows that
\[ E_x(u(t - s, B_s)) = u(s, x) = [e^{\sigma^2 s \Delta/2} u(t - s, \cdot)](x). \] (3.29)

Hence, (3.23) has the form
\[ u(t, x) = e^{-t} [e^{\sigma^2 \Delta/2} g(\cdot)](x) + \bar{N} \int_0^t [e^{\sigma^2 s \Delta/2} u(t - s, \cdot)](x)e^{-s}ds. \] (3.30)

This is simply a way to write the Duhamel formula for the initial value problem
\[ \frac{\partial u}{\partial t} = \frac{\sigma^2}{2} \Delta u + (\bar{N} - 1)u, \]
\[ u(0, x) = g(x). \] (3.31)

This is exactly (3.19), with \( m = \bar{N} - 1 \).

**Exercise 3.3** Show why (3.30) implies (3.31). More generally, consider an evolution equation of the form
\[ \frac{du}{dt} = \mathcal{L}u + F(t), \] (3.32)
with an initial condition \( u(0) = u_0 \). Show that, at least, on the formal level, we have
\[ u(t) = e^{\mathcal{L}t}u_0 + \int_0^t e^{\mathcal{L}s}F(t - s)ds, \quad \text{for } t > 0. \] (3.33)

The assumption in Exercise 3.3 is that the exponential \( \exp(\mathcal{L}t) \) is well-defined for all \( t > 0 \).

**Exercise 3.4** Note that we have obtained (3.19) with \( m = \bar{N} - 1 > 0 \). Modify the branching process to cover also equations with \( m < 0 \). Which range of \( m \) can be obtained this way? What needs to be done to get \( m < -1 \)?

**Exercise 3.5** (i) Given an open set \( A \subset \mathbb{R}^n \) with a smooth boundary, let \( N_A(t) \) be the number of the binary BBM particles inside \( A \) at a time \( t > 0 \). Use (3.31) with an appropriate initial condition to find \( \mathbb{E}(N_A(t)) \).

(ii) Let \( N(t) \) be the total number of BBM particles present at a time \( t > 0 \). Use (3.31) to show that \( \mathbb{E}(N(t)) = \exp((\bar{N} - 1)t) \). This is the same as the result of Proposition 1.1.

(iii) Use (3.31) and the properties of the standard heat equation to show that \( M(t) = N(t)e^{-t} \) is a martingale.

(iv*) Find out if it is possible to use the properties of the standard heat equation to show that \( M_\infty = \lim_{t \to +\infty} M(t) \) is exponentially distributed.

## 3.3 The Fisher-KPP equation and the branching Brownian motion

### 3.3.1 Derivation of the Fisher-KPP equation

Let us now explain how the Fisher-KPP equation
\[ \frac{\partial u}{\partial t} = \frac{\sigma^2}{2} \Delta u + u - u^2 \] (3.34)
arises from very similar considerations, in the context of BBM. This was discovered by Henry McKean in [8].

Given a bounded function \( g(x) \), we now consider a functional of the branching Brownian motion not of the additive form (3.22) but multiplicative:

\[
v(t, x) = \mathbb{E}_x \left( \prod_{k=1}^{N_t} g(X_k(t)) \right). \tag{3.35}
\]

In order to get an equation for \( v(t, x) \) let us write a renewal relation, very similar to what we have seen in (3.23). Looking at the first branching event gives

\[
v(t, x) = \mathbb{E}_x(g(B_t))\mathbb{P}(\tau_1 > t) + \sum_{k=1}^{\infty} p_k \int_0^t \mathbb{E}_x \left( [v(t-s, x+B_s)]^k \right) \mathbb{P}(\tau_1 \in ds)
\]

\[
= \mathbb{E}_x(g(B_t))e^{-t} + \sum_{k=1}^{\infty} p_k \int_0^t \mathbb{E}_x \left( [v(t-s, x+B_s)]^k \right)e^{-s} ds. \tag{3.36}
\]

Recall that we consider a branching Brownian motion with a variance \( \sigma > 0 \). Hence, as in (3.29), we have

\[
\mathbb{E}_x \left[ v(t-s, x+B_s) \right]^k = [e^{\sigma^2 s \Delta/2}v^k(t-s, \cdot)](x). \tag{3.37}
\]

Recall also (3.25):

\[
\mathbb{E}_x(g(B_t)) = [e^{\sigma^2 t \Delta/2}g(\cdot)](x). \tag{3.38}
\]

Now, (3.36) becomes

\[
v(t, x) = [e^{\sigma^2 t \Delta/2}g(\cdot)](x)e^{-t} + \sum_{k=1}^{\infty} p_k \int_0^t \left[ e^{\sigma^2 s \Delta/2}v^k(t-s, \cdot) \right](x)e^{-s} ds
\]

\[
= [e^{\sigma^2 t \Delta/2}g(\cdot)](x)e^{-t} + \int_0^t \left[ e^{\sigma^2 s \Delta/2}F(v(t-s, \cdot)) \right](x)e^{-s} ds. \tag{3.39}
\]

This is the Duhamel representation for the initial value problem

\[
\frac{\partial v}{\partial t} = \frac{\sigma^2}{2} \Delta v - v + F(v),
\]

\[
v(0, x) = g(x). \tag{3.40}
\]

Here, the nonlinearity \( F(v) \) is given by the generating function for the branching process:

\[
F(v) = \sum_{k=1}^{\infty} p_k v^k. \tag{3.41}
\]

From the PDE point of view, it is often convenient to use instead the function

\[
u(t, x) = 1 - v(t, x) = 1 - \mathbb{E}_x \left( \prod_{k=1}^{N_t} g(X_k(t)) \right). \tag{3.42}
\]
It satisfies the initial value problem

\[
\begin{aligned}
\frac{\partial u}{\partial t} &= \frac{\sigma^2}{2} \Delta u + f(u), \\
0(0, x) &= 1 - g(x).
\end{aligned}
\] (3.43)

Here, we have defined

\[
f(u) = 1 - u - F(1 - u) = 1 - u - \sum_{k=1}^{\infty} p_k (1 - u)^k.
\] (3.44)

In the case of the purely binary branching, when \( p_2 = 0 \) and all other \( p_k = 0 \), the function \( f(u) \) takes the form

\[
f(u) = 1 - u - (1 - u)^2 = u(1 - u).
\]

Then, (3.43) becomes the classical Fisher-KPP equation

\[
\begin{aligned}
\frac{\partial u}{\partial t} &= \frac{\sigma^2}{2} \Delta u + u - u^2, \\
0(0, x) &= 1 - g(x).
\end{aligned}
\] (3.45)

As a small remark, applied mathematicians usually take \( \sigma = \sqrt{2} \) and write (3.45) as

\[
\begin{aligned}
\frac{\partial u}{\partial t} &= \Delta u + u - u^2, \\
0(0, x) &= 1 - g(x).
\end{aligned}
\] (3.46)

**Other symmetric functionals**

We have now seen how a partial differential equation can be derived for two functionals of the Brownian motion: the additive functional

\[
u(t, x) = \mathbb{E}_x \left( \sum_{k=1}^{N_t} g(X_k(t)) \right),
\] (3.47)

and the multiplicative one

\[
u(t, x) = \mathbb{E}_x \left( \prod_{k=1}^{N_t} g(X_k(t)) \right).
\] (3.48)

The function \( u(t, x) \) satisfies the linear equation

\[
\begin{aligned}
\frac{\partial u}{\partial t} &= \frac{\sigma^2}{2} \Delta u + (\bar{N} - 1)u, \\
0(0, x) &= 1 - g(x).
\end{aligned}
\] (3.49)

and \( v(t, x) \) satisfies the nonlinear equation (3.40)

\[
\begin{aligned}
\frac{\partial v}{\partial t} &= \frac{\sigma^2}{2} \Delta v - v + F(v).
\end{aligned}
\] (3.50)

Both \( u(t, x) \) and \( v(t, x) \) are expectations of symmetric functions of \( g(X_1), \ldots, g(X_{N_t}) \). The next exercise asks you to consider more general symmetric functionals of \( g(X_1), \ldots, g(X_{N_t}) \).
Exercise 3.6 (i*) Consider the function
\[ z(t, x) = \mathbb{E}_x \left( \sum_{1 \leq i < j \leq N_t} g(X_i(t))g(X_j(t)) \right). \] (3.51)

Write down a renewal relation for \( z(t, x) \) and try to describe it as a solution to some partial differential equation.

(ii*) Let \( F \) be a symmetric function of its arguments. Investigate if
\[ w(t, x) = \mathbb{E}_x \left[ F(g(X_1(t)), g(X_2(t)), \ldots, g(X_{N_t}(t))) \right] \] (3.52)
satisfies a tractable problem. One should probably first consider the case when \( F \) is a symmetric polynomial, which is the case for \( u(t, x) \) and \( v(t, x) \) defined in (3.47) and (3.48). Can one characterize a class of partial differential equations that have such interpretation in terms of a symmetric functional of the branching Brownian motion?

3.3.2 The law of the maximum of BBM

Let us now make a particular choice of the function \( g(x) \) as
\[ g(x) = 1(x \geq 0). \] (3.53)

Then, the function \( v(t, x) \) defined by (3.35)
\[ v(t, x) = \mathbb{E}_x \left( \prod_{k=1}^{N_t} g(X_k(t)) \right), \] (3.54)
has the meaning
\[ v(t, x) = \mathbb{P}_x \{ \text{all } X_k(t) \geq 0 \}, \] (3.55)
hence the meaning of \( u(t, x) \) is
\[ u(t, x) = \mathbb{P}_x \{ \text{some } X_k(t) \leq 0 \}. \] (3.56)

Therefore, the solution to the initial value problem
\[ u_t = \frac{\sigma^2}{2} u_{xx} + f(u), \]
\[ u(0, x) = 1(x < 0), \] (3.57)
is
\[ u(t, x) = \mathbb{P}_x \left\{ \min_{1 \leq k \leq N_t} X_k(t) \leq 0 \right\}. \]

The translational and reflection invariance of the process means that \( u(t, x) \) can be written also as
\[ u(t, x) = \mathbb{P}_0 \left\{ \max_{1 \leq k \leq N_t} X_k(t) \geq x \right\}, \]
that is, the solution \( u(t, x) \) to the Fisher-KPP equation (3.57) with the (reflected) Heaviside function as the initial condition is the probability distribution function of the maximum of
the branching Brownian motion starting at \( x = 0 \). In particular, the median point \( m(t) \) such that
\[
\mathbb{P}_0 \left\{ \max_{1 \leq k \leq N_t} X_k(t) \geq m(t) \right\} = \frac{1}{2}
\] (3.58)
can be characterized by
\[
u(t, m(t)) = \frac{1}{2}.
\] (3.59)
Later we will see that there exists \( x_0 \in \mathbb{R} \) so that \( m(t) \) has the asymptotics
\[
m(t) = c_* t - \frac{3}{2\lambda_*} \log t + x_0 + o(1), \quad \text{as } t \to +\infty.
\] (3.60)
Here, the speed \( c_* \) and the exponent \( \lambda_* \) are given by
\[
c_* = \sqrt{2(\bar{N} - 1)}, \quad \lambda_* = c_*.
\] (3.61)
This asymptotics should be contrasted with the asymptotics (2.16)
\[
m_N = c_* N - \frac{1}{2\lambda_*} \log N + x_0 + o(1), \quad \text{as } N \to +\infty,
\] (3.62)
with
\[
c_* = \sqrt{2 \log 2}, \quad \lambda_* = c_*.
\] (3.63)
for median of the distribution of the maximum of \( 2^N \) uncorrelated Gaussian random variables of variance \( N \). This number of particles corresponds to \( \bar{N} = 1 + \log 2 \), hence the speed \( c_* \) and the exponential rate of decay \( \lambda_* \) of the distribution match those for the corresponding BBM. What is different is the pre-factor \( 3/2 \) in front of the \( \log t \) term in (3.60) as opposed to \( 1/2 \) in expression (3.62) for the uncorrelated model.

3.3.3 The Laplace transform of the point process for the branching Brownian motion

We now explain how the Laplace transform of the point process of the branching Brownian motion can be connected to the Fisher-KPP equation. Let \( X_1(t), \ldots, X_{N_t}(t) \) be the locations of the BBM particles at a time \( t > 0 \), and set
\[
\mathcal{E}(t) = \sum_{k=1}^{N_t} \delta(x - X_k(t)).
\] (3.64)
For the moment, we assume that the BBM starts at the position \( x = 0 \) at \( t = 0 \). Note that at the moment we do not center the locations of the particles near the maximum, as we did for uncorrelated Gaussians in Section 2.4. This will be done later. The Laplace functional of \( \mathcal{E}(t) \) is
\[
\Psi(\phi)(t) = \mathbb{E}_0 \exp \left( - \int \phi(x) d\mathcal{E}(t) \right) = \mathbb{E}_0 \exp \left( - \sum_{k=1}^{N_t} \phi(X_k(t)) \right).
\] (3.65)
Here, $\phi(x)$ is a non-negative bounded test function. The subscript 0 in (3.65) refers to the starting point of the branching Brownian motion. A simple but important observation is that (3.65) can be written as

$$
\Psi(\phi)(t) = \mathbb{E}_0 \exp \left( - \int \phi(x) d\mathcal{E} \right) = \mathbb{E}_0 \exp \left( - \sum_{k=1}^{N_t} \phi(X_k(t)) \right) = \mathbb{E}_0 \left( \prod_{k=1}^{N_t} g(X_k(t)) \right), \quad (3.66)
$$

with the function $g(x)$ given by

$$
g(x) = e^{-\phi(x)}. \quad (3.67)
$$

Combining with we have done in Section 3.3.1, we conclude that if we let $u(t, x)$ be the solution to the initial value problem

$$
\frac{\partial u}{\partial t} = \frac{\sigma^2}{2} \Delta u + f(u),
\quad u(0, x) = 1 - e^{-\phi(x)}, \quad (3.68)
$$

then

$$
\Psi(\phi)(t) = 1 - u(t, 0). \quad (3.69)
$$

Let us see what happens if $\phi(x) = \lambda \mathbb{1}_{[a,b]}(x)$ is a step function. In that case, the initial condition in (3.68) is a multiple of $\phi(x)$:

$$
u(0, x) = (1 - e^{-\lambda}) \mathbb{1}_{[a,b]}(x). \quad (3.70)
$$

Exercise 3.7 Let $u(t, x)$ be the solution to (3.68) with the initial condition (3.70), corresponding to $\phi(x) = \lambda \mathbb{1}_{[a,b]}(x)$. Use the definition of the Laplace transform $\Psi(\phi)(t)$ and (3.69) to interpret what it means, in terms of the branching Brownian motion, that $u(t, 0)$ is small, or that $u(t, 0) \approx 1$.

Thus, the Laplace transform of the point process of the branching Brownian motion can be directly computed in terms of a solution of the Fisher-KPP equation with a suitable initial condition. This provides an occasionally powerful tool to understand the statistics of the branching Brownian motion in terms of solutions of PDEs. For example, one may be interested in finding a location $y$ that would make the limit of the re-centered point process, with

$$Z_j(t) = X_j(t) - y \quad (3.71)$$

be non-trivial. This is what we did in Section 2.4 by re-centering at the locations $y = m_N$. The re-centered point process is

$$\mathcal{E}(t, x, y) = \sum_{k=1}^{N_t} \delta(x + y - X_k(t)). \quad (3.72)$$
The Laplace functional of $\mathcal{E}(t, y)$ is

$$
\Psi(\phi)(t, y) = \mathbb{E}_0 \exp \left( - \int \phi(x) d\mathcal{E}(t, x, y) \right) = \mathbb{E}_0 \exp \left( - \sum_{k=1}^{N_t} \phi(X_k(t) - y) \right) = \mathbb{E}_0 \exp \left( - \sum_{k=1}^{N_t} \phi_y(X_k(t)) \right). \tag{3.73}
$$

Here, we have denoted

$$
\phi_y(x) = \phi(x - y).
$$

It follows that

$$
\Psi(\phi(t, y)) = 1 - \tilde{u}(t, 0; y), \tag{3.74}
$$

and $\tilde{u}(t, x; y)$ is the solution to the Fisher-KPP equation with a shifted initial condition

$$
\begin{align*}
\frac{\partial \tilde{u}}{\partial t} &= \frac{\sigma^2}{2} \Delta \tilde{u} + f(\tilde{u}), \\
\tilde{u}(0, x; y) &= 1 - e^{-\phi(x-y)},
\end{align*} \tag{3.75}
$$

We see that

$$
\tilde{u}(t, x, y) = u(t, x - y), \tag{3.76}
$$

with $u(t, x)$ that is the solution to (3.68). In particular, we have

$$
\tilde{u}(t, 0; y) = u(t, -y), \tag{3.77}
$$

and

$$
\Psi(\phi)(t, y) = u(t, -y). \tag{3.78}
$$

Thus, solution to (3.68) encodes the information about the shifted processes $Z_k(t)$ as well. In particular, to find the shift $y$ that would make the limit of $Z_k(t)$ non-trivial, we need to find the locations where $u(t, y)$ is neither close to 0 nor to 1. A natural question is if this range of points would depend on the initial condition to (3.68), and we will soon see that it does not.

References


