The genealogy of branching Brownian motion with drift and absorption

Julien Berestycki[1]

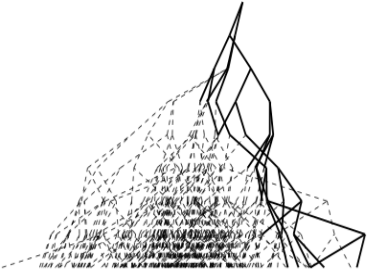
[1] Université Pierre et Marie Curie, CMAP Polytechnique

18/12/2009

Joint work with N. Berestycki and J. Schweinsberg

Outline

Population of fixed size $N, X_1(t), X_2(t), \ldots, X_N(t)$ on the real line. Position $=$ *fitness*.



- Discrete generations
- Reproduction $=$ Each ptc produces k offspings (iid centered displacements $\sim \mu$)
- Selection $=$ Keep N rightmost of the kN produced.

Durrett-Mayberry (09) study a related model in context of predator-prey system. See also Bérard-Gouéré (09) and Durrett-Remenik (09).

Conjecture

$L_t = \max\{X_i(t), i = 1, \ldots, N\}$ then a.s. $L_t/t \to v_N$. Furthermore, limit $v_\infty = \lim_{N\to\infty} v_N$ exists, and $\exists C$ such that

$$v_\infty - v_N \sim \frac{C}{(\log N)^2}. \tag{1}$$

Recently rigorously proved by Bérard and Gouéré using ideas from Gantert, Hu and Shi (when $k = 2$ and suitable regularity cond. on μ).

Brunet-Derrida-Mueller-Munier(06,07) study the solution $u(x,t)$ to the noisy FKPP equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + u - u^2 + \sqrt{u(1-u)/N}\, W(x,t), \qquad (2)$$

where $W(x,t) =$ space-time white noise.
Without noise this is the FKPP equation (1937) Admits traveling wave solutions $u(t,x) = w(x - vt)$, where $w =$function
$v =$velocity of wave.
Brunet-Derrida conjecture :
$v_N - v_\infty + \pi^2/(\log N)^2 = o(1/(\log N)^2)$ (cutoff ansatz) In fact even more precise conjecture
$v_N - v_\infty + \pi^2/(\log N)^2 \simeq 3\pi^2 \log\log N/(\log N)^3)$ Recent progresses by Mytnik Mueler and Quastel.

Conjecture

Sample 2 ptc at some time. Then age of most recent common ancester (MRCA) is of order $(\log N)^3$.

Conjecture

Sample n ptc from population at random at some generation. Their ancestral lines are traced backwards in time. The tree they form = Bolthausen-Sznitman coalescent.

This suggests that a good model for the genalogy of a population under selection is the Bolthausen-Sznitman coalescent !

Let Λ be a finite measure on $[0,1]$. The Λ-coalescent is a Markov process in $\mathcal{P} = \{\text{partitions of } \mathbb{N}\}$ such that when b blocks each k-tuple ($2 \le k \le b$) merge at rate

$$\lambda_{b,k} = \int_0^1 p^k (1-p)^{b-k} p^{-2} \Lambda(dp)$$

If $\Lambda(\{0\}) = 0$ Poissonian construction. Start with PPP on $[0,\infty) \times [0,1]$ with intensity $dt \times p^{-2}\Lambda(dp)$. If (t,p) is an atom a p merger occurs at time t. That is we flip a coin for each lineage with probability p of heads, and merge all lineages whose coin is head.
Bolthausen-Sznitman coalescent : $\Lambda(dx) = dx$. Connection with Derrida's GREM (Bovier Kurkova 07), random recursive trees (Goldschmidt Martin 05).

Introduced by Laperti (67) A continuous state branching process (CSBP) $=$ $[0,\infty)$-valued Markov process $(X(t), t \geq 0)$ whose transition function

$$p_t(a+b,\cdot) = p_t(a,\cdot) * p_t(b,\cdot).$$

CSBPs $=$ scaling limits of GW processes.
Let $(Y(s), s \geq 0) =$ a Lévy process with no < 0 jumps and $Y(0) > 0$ stopped when it hits 0. Let

$$S(t) := \inf\{u : \int_0^u Y(s)^{-1} ds > t\}$$

Then $X(t) = Y(S(t))$ is a CSBP. All CSBP can b obtained this way.
If $Y(0) = a$ then $E[e^{-uY(t)}] = e^{au+\psi(u)}$ where the branching mechanism ψ is

$$\psi(u) = \alpha u + \beta u^2 + \int_0^\infty (e^{-ux} - 1 + ux 1_{x \leq 1} \nu(dx)$$

Neveu (92) considers

$$\psi(u) = au + bu\log u = cu + \int_0^\infty (e^{-ux} - 1 + ux1_{x\leq 1}bx^{-2}dx$$

Bertoin-Le Gall (00) : the genealogy of Neveu's CSBP is the Bolthausen-Sznitman coalescent.
Let A be the current pop. size. After a jump of size x a fraction $p = x/(A+x)$ was born of a single indiv. at this time. Tracing ancestral lineages, we have a p-merger.
Since $x/(A+x) \geq r$ iff $x \geq Ar/(1-r)$ p-mergers with $p \geq r$ occur at rate

$$A\int_{Ar/(1-r)}^\infty bx^{-2}dx = b\frac{1-r}{r}$$

In the Bolthausen-Sznitman coalescent rate of $p-$mergers with $p \geq r$ is $\int_r^1 p^{-2}\Lambda(dp) = \frac{1-r}{r}$.

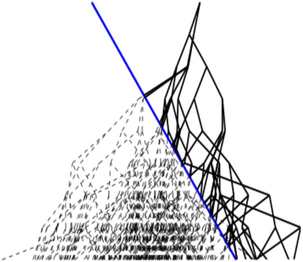
Begin with some configuration of ptcles at time 0. Each ptcl independently performs standard 1-d BM. Each ptcl splits into two at rate 1.
Early work on $M(t)$ position of right-most ptcl (started from 1 ptcl at 0):

- McKean (75) $u(t,x) = P(M(t) \leq x)$ solves the FKPP equation.
- Bramson (78) is $m(t) =$ median of $M(t)$ then

$$m(t) = \sqrt{2}t - (3/2\sqrt{2}) \log t + O(1)$$

- Sharper results by Bramson (83), Lalley-Sellke (87).

Selection = kill ptcl too far to the left.
$M_N(t)$ (random) particles $X_1(t), X_2(t), \dots, X_{M_N(t)}(t)$ on the real line.



- BBM where ptcl. have drift $-\mu$
- kill ptcl. when hit 0

Theorem (Kesten, 1978)

Start with 1 ptc at $x > 0$. This process dies out a.s. if $\mu \geq \sqrt{2}$. If $\mu < \sqrt{2}$ the $\#$ of ptc grows exponentially with positive proba.

We take $\mu = \mu_N = \sqrt{2 - \frac{2\pi^2}{(\log N + 3 \log\log N)^2}}$ (the $O(1/(\log N)^2)$ correction is related to conjecture 1).

$M_N(t) = \#$ of ptc. alive at time t

$(X_{1,N}, \ldots, X_{M_N(t),N})$ = positions of ptc alive at time t

$L = \frac{1}{\sqrt{2}}(\log N + 3 \log\log N)$ natural space scale.

$Y_N(t) = \sum_1^{M_N(t)} e^{\mu X_{i,N}(t)}$

$Z_N(t) = \sum_1^{M_N(t)} e^{\mu X_{i,N}(t)} \sin(\frac{\pi X_{i,N}(t)}{L}) 1_{X_{i,N}(t) \leq L}$ measure the size of the pop. to left of L.

Theorem (B.-Berestycki-Schweinsberg 09)

Suppose $Z_N(0)/[N(\log N)^2] \Rightarrow \nu$ *and* $Y_N(0)/[N(\log N)^3] \Rightarrow 0$. *Then* $\exists a \in \mathbb{R}$ *s.t. the fdd of*

$$\left(\frac{1}{2\pi N} M_N((\log N)^3 t), t \geq 0 \right)$$

converge to those of the CSBP with initial pop. $\sim \nu$ *and branch. mech.* $\psi(t) = au + 2\pi^2 u \log U$.

Note : The initial condition will be satisfied if we start with N ptc in a "stable" config. with no ptc. too far to the right.
Neveu (1992) studied a CSBP with $\psi(u) = u \log u$. Bertoin and Le Gall (2000) have shown that the genealogy of Neveu CSBP is the Bolthausen Sznitman coalescent.

Theorem

Same conditions. Fix $t > 0$ and chose n ptc at random from the $M_N((\log N)^3 t)$ alive at time $(\log N)^3 t$. Let $\Pi_N(s)$ be the ancestral partition at time $(\log N)^3(t-s)$. Then fdd of $(\Pi_N(s), 0 \leq s \leq 2\pi t)$ converge to those of Bolthausen - Sznitman coalescent run at speed 2π for time $2\pi t$.

Ancestral partition : $\Pi_N(s)$ partition of $\{1, \ldots, n\}$, i and j in the same block iff i and j have the same ancestor at time $(\log N)^3(t - s/2\pi)$.

From Brunet- Derrida-Mueller-Munier (06,07). All stochastic effects come from the tip of the front.
Occasionally 1 ptc gets very far to the right. This ptc. does not feel the killing and grows like a BBM for a while producing a large # of offsprings who avoid the wall.
This lead to jumps in the # of ptcl and multiple mergers in the genealogy.
Our proof strategy follows this heuristic:

- Find level L that ptc must reach to produce a jump in # of ptc.
- Behavior of BBM with ptc killed at 0 and L is roughly determinist ("LLN" or "fluid limit" type result by 1st and 2nd moments).
- Separately consider the ptc that reach L.

Some questions I want to answer :

1. Why is the drift which keeps the population stable

$$\mu = \mu_N = \sqrt{2 - \frac{2\pi^2}{(\log N + 3\log\log N)^2}}$$

2. Why must a ptcl gets near

$$L = \frac{1}{\sqrt{2}}(\log N + 3\log\log N)$$

   to produce a jump in the pop. size?
3. Why does $Z_N(t)$ measure the size of the pop.?
4. Why is the characteristic time scale $(\log N)^3$?
5. Why does the Bolthausen-Sznitman coalescent arise?
   1. Why is jump rate prop. to the # of ptcls?
   2. Why is the rate of jumps of size greater than x prop. to $\int_x^\infty y^{-2}dy = x^{-1}$?

BM killed at 0 and L. Start with 1 ptc at x. "Density" at time t is

$$q_t(x,y) = \frac{2}{L}\sum_{n=1}^{\infty} e^{-\pi^2 n^2 t/2L^2} \sin\left(\frac{n\pi x}{L}\right) \sin\left(\frac{n\pi y}{L}\right).$$

Add branching an drift $-\mu$. "Density" becomes

$$p_t(x,y) = q_t(x,y) e^t e^{\mu(x-y)-\mu^2 t/2}.$$

When $t >> L^2$

$$p_t(x,y) \equiv \frac{2}{L} e^{(1-\mu^2/2-\pi^2/2L^2)t} \cdot e^{\mu x} \sin\left(\frac{\pi x}{L}\right) \cdot e^{-\mu y} \sin\left(\frac{\pi y}{L}\right)$$

$$p_t(x,y) \equiv \frac{2}{L} e^{(1-\mu^2/2-\pi^2/2L^2)t} \cdot e^{\mu x} \sin\left(\frac{\pi x}{L}\right) \cdot e^{-\mu y} \sin\left(\frac{\pi y}{L}\right)$$

- If $\mu = \mu_L = \sqrt{2 - \frac{\pi^2}{L^2}}$ then no time dependence. # of ptcl stays relatively stable. $Z_N(t)$ is a martingale.
- Formula prop. to $e^{\mu x} \sin(\pi x/L)$. so summing over ptcl., at time 0 this becomes $Z_N(0)$. Thus $Z_N(0$ determines the # of ptcls in a set at future times.
- # of ptc at y is prop to $e^{-\mu y} \sin(\pi y/L)$. When $t >> L^2$ ptc. have density

$$g(x) = \frac{\mu^2}{\pi} L e^{-\mu y} \sin(\pi y/L)$$

Choice of L

If we start with N ptc. picked iid $\sim g(x) = \frac{\mu^2}{\pi} L e^{-\mu y} \sin(\pi y/L)$ a quick calculation shows

$$Z_N(0) = \sum_{i=1}^{N} e^{\mu X_i(0)} \sin\left(\frac{\pi X_i(0)}{L}\right) \simeq \frac{NL^2}{\pi}$$

By martingale property $Z_N(t) = O(NL^2)$ for larger t.
If we start with 1ptc at L, usually the right-most descendent will only get to $L + \alpha$ where α is a constant, so

$$Z_N(t) \approx e^{\mu L} \sin\left(\frac{L}{L+\alpha}\right) = O(L^{-1} e^{\mu L})$$

and note that NL^2 and $L^{-1} e^{\mu L}$ are of the same order when

$$L = \frac{1}{\sqrt{2}}(\log N + 3 \log\log N) + O(1).$$

time scale $(\log N)^3$

How often does a ptcl hit L ? Fix $\beta > 0$ a ptcl in $[L-\beta, L]$ at time t has a > 0 proba of hitting L before $t+1$.
Configuration at time t is like N ptcls with density $g(y) = CLe^{-\mu y} \sin\left(\frac{\Pi y}{L}\right)$.
Expected # of ptcls in $[L-\beta, L]$ is thus

$$N \int_{L-\beta}^{L} CLe^{-\mu y} \sin\left(\frac{\Pi y}{L}\right) dy = O((\log N)^{-3}).$$

A continuous time branching process : effect of ptc that reach L

Consider BBM with drift $\mu = -\sqrt{2}$ started with 1 ptc at L.
Let $M(y) = \#$ of ptc that reach $L - y$ when ptc are killed at $L - y$.
Cond. on $M(x) = n, M(x+y) =_d$ the sum of n iid variables $\sim M(y)$. Therefore $(M(z), z \geq 0)$ is a continuous time branching process. Offspring dist. has finite mean but not in the $L \log L$ class.

BBM with drift $\mu = -\sqrt{2}$ started with 1 ptc at L.

Theorem (Neveu 1987)

$\exists$ *r.v.* W *such that a.s.*

$$\lim_{y\to\infty} y e^{-\sqrt{2}y} M(y) = W$$

Furthermore, for all $u \in \mathbb{R}$

$$E[e^{-e^{-\sqrt{2}u}W}] = \psi(u)$$

where ψ *solves*

$$\frac{1}{2}\psi'' - \sqrt{2}\psi' = \psi(1-\psi).$$

Tail behavior of W

Lemma

$P(W > x) \sim \frac{1}{\sqrt{2x}}$ *as* $x \to \infty$.

By Tauberian theorem this reduces to study of asymptotic of $E(e^{-\lambda W})$ when $\lambda \to 0$.
Since $E[e^{-e^{-\sqrt{2u}}W}] = \psi(u)$, this reduces to asymptotic of $\psi(u)$ when $u \to \infty$.
Following an idea of Harris (1999), this asymptotic is obtained through connection with 3d- Bessel process.

- Waiting time for a ptc to hit L is typically $O((\log N)^3)$. Rate at which ptc reach L is proportional to $Z_N(t)$.
- Once a ptc hit L its contribution is prop. to $\#$ of descendants that hit $L - y$ for large y. This is roughly $We^{\sqrt{2}y}/y$ Take $y = O(L)$ to find $WN(\log N)^2$
- Hence, the proba. that $Z = Z_N(t)/(N(\log N)^2)$ jumps by at least r is roughly $P(W > r) \sim Cr^{-1}$.

After a jump of size r, a fraction $p = r/(Z + r)$ of the population is descended from the ptc that hit L. Tracing lineages backward in time a p merger occur at this time.
Since $r/(Z + r) \geq x$ iff $r \geq Zx/(1 - x)$ we get p-mergers with $p \geq x$ at rate

$$CZ.\frac{1-x}{Zx} = C\frac{1-x}{x}.$$

For Bolthausen Sznitman coalescent, rates of p merger with $p \geq x$ is

$$\int_x^1 p^{-2} dp = \frac{1-x}{x}.$$