SOME MATHEMATICAL MODELS FROM POPULATION GENETICS

Alison Etheridge
University of Oxford

with thanks to numerous collaborators, especially Nick Barton, IST Austria

CMAP5, May 2023

The spatial Λ-Fleming-Viot process

Barton - E - Véber and friends

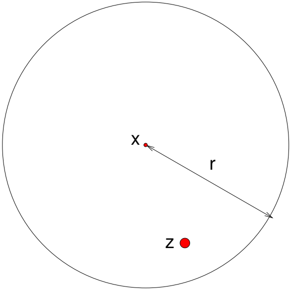
State $\{\rho(t,x,\cdot)\in\mathcal{M}_1(K), x\in\mathbb{R}^2, t\geq 0\}$. Π Poisson point process rate $dt\otimes dx\otimes\xi(dr,du)$ on $[0,\infty)\times\mathbb{R}^2\times[0,\infty)\times[0,1]$.

Dynamics: for each $(t,x,r,u)\in\Pi$,

- $z\sim U(B_r(x))$
- $k\sim\rho(t-,z,\cdot)$.

For all $y\in B_r(x)$,

$$\rho(t,y,\cdot)=(1-u)\rho(t-,y,\cdot)+u\delta_k.$$



Backwards in time

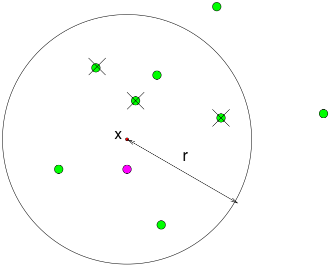
- A *single* ancestral lineage evolves in series of jumps with intensity

$$dt \otimes \int_{(|x|/2,\infty)} \int_{[0,1]} \frac{L_r(x)}{\pi r^2}\, u\, \xi(dr, du)dx$$

on $\mathbb{R}_+ \times \mathbb{R}^2$ where $L_r(x) = |B_r(0) \cap B_r(x)|$.

- Lineages can coalesce when hit by same 'event'.

Note: If $\xi(dr, du) = \mu(dr) \otimes \delta_u$, rate of jumps $\propto u$.



Introducing selection to the SLFV

$K = \{a, A\}$, $w(t,x) = \rho(t,x,a)$ proportion of type a

- (i) Two types, a, A. Weight type a by $(1-s)$. If a reproduction event affects a region $B(x,r)$ in which current proportion of a-alleles is $\overline{w}$, then probability offspring are type a is
$$\frac{(1-s)\overline{w}}{1-s\overline{w}} = \overline{w}(1-s) + s\overline{w}^2 + \mathcal{O}(s^2).$$
- (ii) Neutral events rate $\propto (1-s)$, selective events rate $\propto s$. At selective reproduction events, sample two potential parents. If types aa, then an a reproduces, otherwise an A does.

c.f. what we did for Moran model

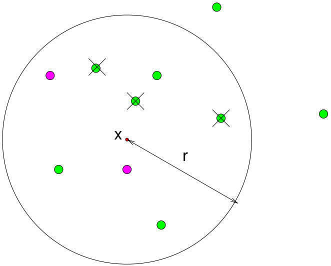
(Spatial) Ancestral selection graph

Evolution of ancestry due to neutral events as before:

- lineages evolve in a series of jumps;
- they can coalesce when covered by same event.

At *selective* events

- *Two 'potential' parents must be traced;*
- Lineages can coalesce when hit by same 'event'.



A sampled individual is type a *iff* all lineages in the corresponding ASG are type a at any previous time.

Zooming out: recovering classical models

The spread of a *favoured* allele is classically modelled through the (stochastic) Fisher-KPP equation:

$$du = \big(\frac{1}{2}\Delta u + su(1-u)\big)dt + \mathbf{1}_{d=1}\epsilon\sqrt{u(1-u)}W(dt, dx).$$

Over sufficiently large spatial and temporal scales, does the proportion of favoured alleles in the SLFV with selection look like a solution to the (stochastic) Fisher-KPP equation?

Key tool: ancestral selection graph.

Stochastic Fisher-KPP is dual to branching and coalescing Brownian motion

Reminder: parameters in SLFV with selection

- Events driven by Poisson Point Process Π that specifies
  - centre and *radius* event
  - *impact* event
- *selection coefficient* determines proportion of selective events

Reminder: parameters in SLFV with selection

- Events driven by Poisson Point Process Π that specifies
  - centre and *radius* event
  - *impact* event
- *selection coefficient* determines proportion of selective events

For simplicity suppose r, u, s fixed

Reminder: parameters in SLFV with selection

- Events driven by Poisson Point Process Π that specifies
  - centre and *radius* event
  - *impact* event
- *selection coefficient* determines proportion of selective events

For simplicity suppose r, u, s fixed

Dual lineages make jumps of length $O(r)$ at rate proportional to ur^d, and branch at rate proportional to sur^d

lineages can only coalesce when at separation less than $4r$

Scaling limits I: High neighbourhood size

Fixed impact u and event radius r, selection coefficient s

- Set $u_n = u/n^{1/3}$, $s_n = s/n^{2/3}$, $w^{(n)}(t,x) = w(nt, n^{1/3}x)$,

$$dw = \frac{1}{2}\Delta w dt + sw(1-w)dt + \mathbf{1}_{d=1}\epsilon\sqrt{w(1-w)}W(dt,dx)$$

E. Véber, Yu.

Scaling limits I: High neighbourhood size

Fixed impact u and event radius r, selection coefficient s

- Set $u_n = u/n^{1/3}$, $s_n = s/n^{2/3}$, $w^{(n)}(t,x) = w(nt, n^{1/3}x)$,

$$dw = \frac{1}{2}\Delta w dt + sw(1-w)dt + \mathbf{1}_{d=1}\epsilon\sqrt{w(1-w)}W(dt, dx)$$

E. Véber, Yu.

Here, $u_n \to 0$, corresponding to high *neighbourhood size*.

Scaling limits I: High neighbourhood size

Fixed impact u and event radius r, selection coefficient s

- Set $u_n = u/n^{1/3}$, $s_n = s/n^{2/3}$, $w^{(n)}(t,x) = w(nt, n^{1/3}x)$,

$$dw = \frac{1}{2}\Delta w dt + sw(1-w)dt + \mathbf{1}_{d=1}\epsilon\sqrt{w(1-w)}W(dt,dx)$$

E. Véber, Yu.

Here, $u_n \to 0$, corresponding to high *neighbourhood size*.

$\gamma > \beta$ (even bigger neighbourhood size) $\rightsquigarrow$ deterministic equation in all dimensions

Establishment probability: high neighbourhood size

$$dw = \frac{1}{2}\Delta w dt + sw(1-w)dt + \epsilon\sqrt{w(1-w)}W(dt,dx).$$

While rare,

$$dw \approx \frac{1}{2}\Delta w dt + swdt + \epsilon\sqrt{w}W(dt,dx).$$

Establishment probability: high neighbourhood size

$$dw = \frac{1}{2}\Delta w dt + sw(1-w)dt + \epsilon\sqrt{w(1-w)}W(dt, dx).$$

While rare,

$$dw \approx \frac{1}{2}\Delta w dt + swdt + \epsilon\sqrt{w}W(dt, dx).$$

Writing X for total mass rare allele,

$$dX \approx sXdt + \epsilon\sqrt{X}dB_t,$$

a continuous state branching process.

Establishment probability: high neighbourhood size

$$dw = \frac{1}{2}\Delta w dt + sw(1-w)dt + \epsilon\sqrt{w(1-w)}W(dt,dx).$$

While rare,

$$dw \approx \frac{1}{2}\Delta w dt + swdt + \epsilon\sqrt{w}W(dt,dx).$$

Writing X for total mass rare allele,

$$dX \approx sXdt + \epsilon\sqrt{X}dB_t,$$

a continuous state branching process. Establishment probability is independent of spatial structure.

Establishment probability: high neighbourhood size

$$dw = \frac{1}{2}\Delta w dt + sw(1-w)dt + \epsilon\sqrt{w(1-w)}W(dt, dx).$$

While rare,

$$dw \approx \frac{1}{2}\Delta w dt + sw dt + \epsilon\sqrt{w}W(dt, dx).$$

Writing X for total mass rare allele,

$$dX \approx sX dt + \epsilon\sqrt{X}dB_t,$$

a continuous state branching process. Establishment probability is independent of spatial structure.

When neighbourhood size is high, spatial structure hardly perturbs establishment probability.

Establishment probability: high neighbourhood size

$$dw = \frac{1}{2}\Delta w dt + sw(1-w)dt + \epsilon\sqrt{w(1-w)}W(dt,dx).$$

While rare,

$$dw \approx \frac{1}{2}\Delta w dt + swdt + \epsilon\sqrt{w}W(dt,dx).$$

Writing X for total mass rare allele,

$$dX \approx sXdt + \epsilon\sqrt{X}dB_t,$$

a continuous state branching process. Establishment probability is independent of spatial structure.

When neighbourhood size is high, spatial structure hardly perturbs establishment probability. ... but in a spatial continuum, neighbourhood size can be small.

Small neighbourhood size: Why rescale?

Neutral mutation rate, μ, sets timescale

- Mutation rates are low;

Natural question:
When will we see a signature of a favourable allele in data?

Scaling limits II: Small neighbourhood size:

Fix $u \in (0, 1)$.

Fix radius events.

Scaling limits II: Small neighbourhood size:

Fix $u \in (0, 1)$. Fix radius events.

Set $n = 1/\mu$ and rescale: $w(nt, \sqrt{n}x)$.

Scaling limits II: Small neighbourhood size:

Fix $u \in (0,1)$.

Fix radius events.

Set $n = 1/\mu$ and rescale: $w(nt, \sqrt{n}x)$.

Heuristics:

- At a 'branching' event in ASG, two lineages born at separation $\mathcal{O}(1/\sqrt{n})$.
- Probability they separate to $\mathcal{O}(1)$ before coalescing is
    - $d = 1$: $\mathcal{O}(1/\sqrt{n})$,
    - $d = 2$: $\mathcal{O}(1/\log n)$,
    - $d \geq 3$: $\mathcal{O}(1)$.
- Selection will only be visible if expect to see at least one pair 'separate' by time 1.
- Order one coalescence probability when meet, so in low dimensions need lots of branches.

Scaling limits II: Small neighbourhood size:

Fix $u \in (0,1)$. Fix radius events.

Set $n = 1/\mu$ and rescale: $w(nt, \sqrt{n}x)$.

Ability to detect selection depends on dimension:

- $d = 1$, selection only visible if $s = \mathcal{O}(1/\sqrt{n})$, *if $u = 1$ limiting ASG embedded in Brownian net*;
- $d = 2$, selection only visible if $s = \mathcal{O}(\log n/n)$, *limiting ASG 'Branching BM'*;
- $d \geq 3$, selection only visible if $s = \mathcal{O}(1/n)$, *limiting ASG Branching BM*.

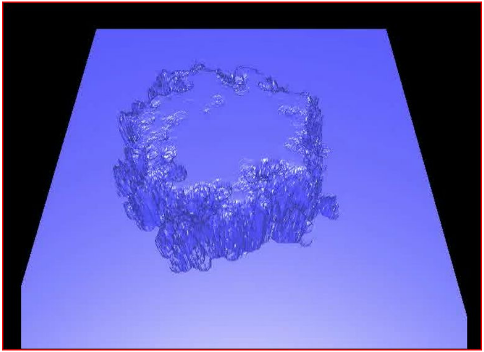
Technical challenges because $ns_n \to \infty$.

Straulino (2015); E., Freeman, Straulino (2017); E., Freeman, Penington, Straulino (2017).

SPACE MATTERS!

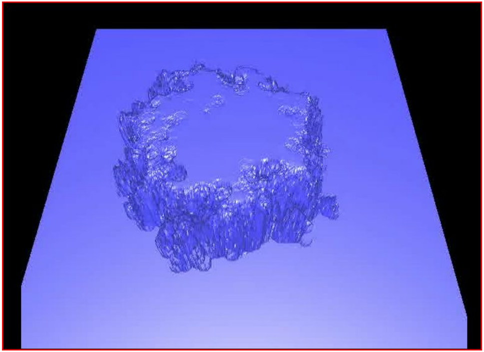
Spread of a favoured allele

Two types, a, A, relative fitnesses $1 : 1 + s$. If a reproduction event affects a region $B(x, r)$ in which current proportion of a-alleles is w, then probability offspring are type a is $\frac{w}{1+s(1-w)}$.



Spread of a favoured allele

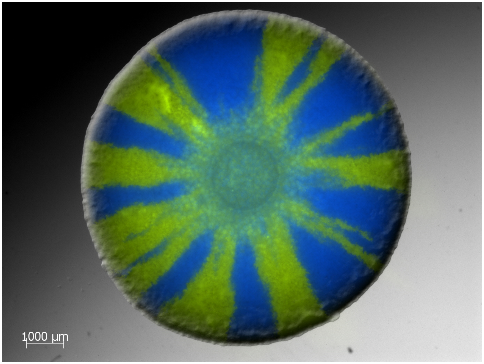
Two types, a, A, relative fitnesses $1 : 1+s$. If a reproduction event affects a region $B(x,r)$ in which current proportion of a-alleles is w, then probability offspring are type a is $\frac{w}{1+s(1-w)}$.



Alternative interpretation: strong selection $\sim$ range expansion

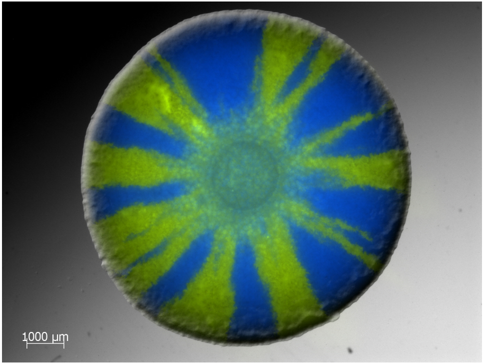
Range expansion

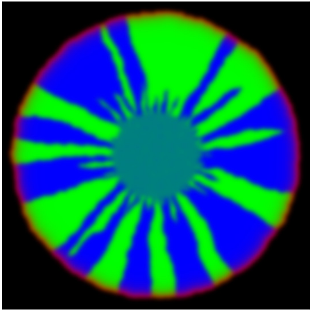
Pseudomanas aeruginosa (Kevin Foster)



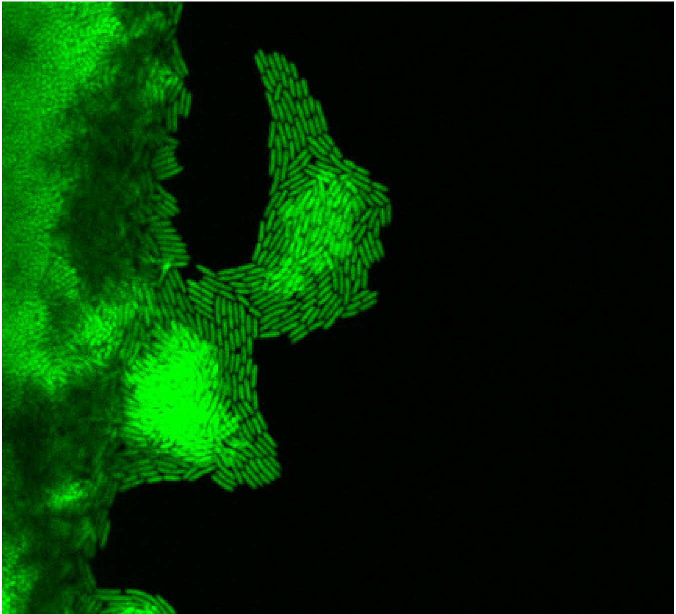
Range expansion

Pseudomanas aeruginosa (Kevin Foster)





What's really happening?



Recap: The Wright-Fisher model with selection

Two types a, A, relative fitnesses

a	A
$1-s$	1

During reproduction, each individual produces large number of juveniles from which next generation sampled.

If proportion a-alleles in parental population is p, proportion in pool of juveniles is

$$p^* = \frac{(1-s)p}{1-sp} \approx p - sp(1-p).$$

Population size N (fixed).

Recap: The Wright-Fisher model with selection

Two types a, A, relative fitnesses

a	A
$1-s$	1

During reproduction, each individual produces large number of juveniles from which next generation sampled.

If proportion a-alleles in parental population is p, proportion in pool of juveniles is

$$p^* = \frac{(1-s)p}{1-sp} \approx p - sp(1-p).$$

Population size N (fixed).
But selection can take many forms

Hybrid Zones

A hybrid zone is a narrow geographic region where two genetically distinct populations are found close together and hybridise to produce offspring of mixed ancestry.

They are maintained by a balance between selection and dispersal.

With thanks to Nick Barton and his group

A mathematical model

We focus on selection against heterozygosity

Individuals carry two copies of a gene that occurs as a or A.

Hardy-Weinberg proportions: $\overline{w} =$ proportion of a-alleles,

aa	aA	AA
$\overline{w}^2$	$2\overline{w}(1-\overline{w})$	$(1-\overline{w})^2$

Relative fitnesses:

aa	aA	AA
1	$1-s$	1

Reproduction

- Each heterozygote (aA) produces $(1-s)$ times as many germ cells (cells of same genotype) as a homozygote (aa or AA);
- Germ cells split into effectively infinite pool of gametes (containing just one copy of gene),

Reproduction

- Each heterozygote (aA) produces $(1-s)$ times as many germ cells (cells of same genotype) as a homozygote (aa or AA);
- Germ cells split into effectively infinite pool of gametes (containing just one copy of gene), with proportion of type a

$$\overline{w}^* \quad = \quad \frac{\left(\overline{w}^2 + \overline{w}(1-\overline{w})(1-s)\right)}{\left(\overline{w}^2 + 2\overline{w}(1-\overline{w})(1-s) + (1-\overline{w})^2\right)}$$

Reproduction

- Each heterozygote (aA) produces $(1-s)$ times as many germ cells (cells of same genotype) as a homozygote (aa or AA);
- Germ cells split into effectively infinite pool of gametes (containing just one copy of gene), with proportion of type a

$$\begin{aligned}
\overline{w}^* &= \frac{\left(\overline{w}^2 + \overline{w}(1-\overline{w})(1-s)\right)}{\left(\overline{w}^2 + 2\overline{w}(1-\overline{w})(1-s) + (1-\overline{w})^2\right)} \\
&= \frac{\overline{w}^2 + \overline{w}(1-\overline{w})(1-s)}{1-2s\overline{w}(1-\overline{w})}
\end{aligned}$$

Reproduction

- Each heterozygote (aA) produces $(1-s)$ times as many germ cells (cells of same genotype) as a homozygote (aa or AA);
- Germ cells split into effectively infinite pool of gametes (containing just one copy of gene), with proportion of type a

$$\begin{aligned}
\overline{w}^* &= \frac{\left(\overline{w}^2 + \overline{w}(1-\overline{w})(1-s)\right)}{\left(\overline{w}^2 + 2\overline{w}(1-\overline{w})(1-s) + (1-\overline{w})^2\right)} \\
&= \frac{\overline{w}^2 + \overline{w}(1-\overline{w})(1-s)}{1 - 2s\overline{w}(1-\overline{w})} \qquad (s \text{ small}) \\
&= (1-s)\overline{w} + s(3\overline{w}^2 - 2\overline{w}^3) + \mathcal{O}(s^2) \\
&= \overline{w} + s\overline{w}(1-\overline{w})(2\overline{w}-1) + \mathcal{O}(s^2).
\end{aligned}$$

Reproduction

- Each heterozygote (aA) produces $(1-s)$ times as many germ cells (cells of same genotype) as a homozygote (aa or AA);
- Germ cells split into effectively infinite pool of gametes (containing just one copy of gene), with proportion of type a

$$\begin{aligned} \overline{w}^* &= \frac{\Big(\overline{w}^2 + \overline{w}(1-\overline{w})(1-s)\Big)}{\Big(\overline{w}^2 + 2\overline{w}(1-\overline{w})(1-s) + (1-\overline{w})^2\Big)} \\ &= \frac{\overline{w}^2 + \overline{w}(1-\overline{w})(1-s)}{1-2s\overline{w}(1-\overline{w})} \qquad (s \text{ small}) \\ &= (1-s)\overline{w} + s(3\overline{w}^2 - 2\overline{w}^3) + \mathcal{O}(s^2) \\ &= \overline{w} + s\overline{w}(1-\overline{w})(2\overline{w}-1) + \mathcal{O}(s^2). \end{aligned}$$

$$\overline{w}^* - \overline{w} = s\overline{w}(1-\overline{w})(2\overline{w}-1) + \mathcal{O}(s^2).$$

A deterministic approximation

$$\overline{w}^* - \overline{w} = s\overline{w}(1 - \overline{w})(2\overline{w} - 1) + \mathcal{O}(s^2).$$

A deterministic approximation

$$\overline{w}^* - \overline{w} = s\overline{w}(1-\overline{w})(2\overline{w}-1) + \mathcal{O}(s^2).$$

In an infinite population, if $s = \frac{\alpha}{M}$ (where M is large), measuring time in units of M generations,

$$\frac{\delta\overline{w}}{\delta t} = \alpha\overline{w}(1-\overline{w})(2\overline{w}-1) + \mathcal{O}(s^2).$$

A deterministic approximation

$$\overline{w}^* - \overline{w} = s\overline{w}(1-\overline{w})(2\overline{w}-1) + \mathcal{O}(s^2).$$

In an infinite population, if $s = \frac{\alpha}{M}$ (where M is large), measuring time in units of M generations,

$$\frac{\delta\overline{w}}{\delta t} = \alpha\overline{w}(1-\overline{w})(2\overline{w}-1) + \mathcal{O}(s^2).$$

$$\frac{d\overline{w}}{dt} = \alpha\overline{w}(1-\overline{w})(2\overline{w}-1).$$

A deterministic approximation

$$\overline{w}^* - \overline{w} = s\overline{w}(1-\overline{w})(2\overline{w}-1) + \mathcal{O}(s^2).$$

In an infinite population, if $s = \frac{\alpha}{M}$ (where M is large), measuring time in units of M generations,

$$\frac{\delta \overline{w}}{\delta t} = \alpha\overline{w}(1-\overline{w})(2\overline{w}-1) + \mathcal{O}(s^2).$$

$$\frac{d\overline{w}}{dt} = \alpha\overline{w}(1-\overline{w})(2\overline{w}-1).$$

Add dispersal:

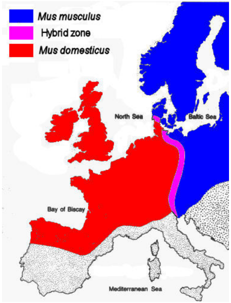
$$\frac{\partial w}{\partial t} = \frac{m}{2}\Delta w + \alpha w(1-w)(2w-1).$$

Examples of hybrid zones

Maintained by selection?

$$\frac{\partial w}{\partial t} = \frac{m}{2}\Delta w + \alpha w(1-w)(2w-1).$$

plus noise



©2019, Steven M Carr

Width of zone

$$\approx \sqrt{\frac{2m}{\alpha}}$$

Examples of hybrid zones

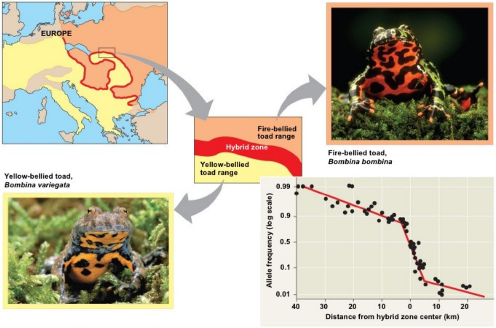
Maintained by selection?

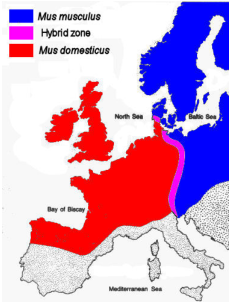
$$\frac{\partial w}{\partial t} = \frac{m}{2}\Delta w + \alpha w(1-w)(2w-1).$$

plus noise

or, eg changes in environment?

Fig. 24-13





Width of zone

$$\approx \sqrt{\frac{2m}{\alpha}}$$

Zooming out

Applying a diffusive rescaling $t \mapsto \frac{t}{\varepsilon^2}$, $x \mapsto \frac{x}{\varepsilon}$, the Allen-Cahn equation becomes

$$\frac{\partial w}{\partial t} = \frac{m}{2}\Delta w + \frac{\alpha}{\varepsilon^2} w(1-w)(2w-1).$$

For convenience, set $m = 2$, $\alpha = 1$.

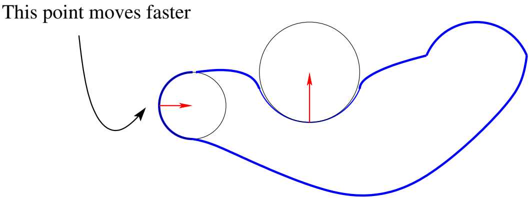
For sufficiently regular initial conditions, as $\varepsilon \to 0$, the solution converges to the indicator function of a region whose boundary evolves according to *curvature flow*.

(Mean) Curvature flow

- $\mathbf{\Gamma}_t : S^1 \to \mathbb{R}^2$ smooth embeddings;
- $\mathbf{n}_t(u)$ unit (inward) normal vector to $\mathbf{\Gamma}_t$ at u;
- $\kappa = \kappa_t(u)$ curvature of $\mathbf{\Gamma}_t$ at u.

$$\frac{\partial \mathbf{\Gamma}_t(u)}{\partial t} = \kappa_t(u)\mathbf{n}_t(u).$$

Defined up to fixed time T



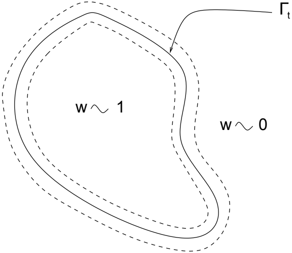
The Allen-Cahn equation and curvature flow

$d(x,t) =$ signed distance x to $\mathbf{\Gamma}_t$

$\mathbf{\Gamma}_0 = \{x \in \mathbb{R}^2 : w_0(x) = \frac{1}{2}\}$

$w_0 > \frac{1}{2}$ inside $\mathbf{\Gamma}$, $< \frac{1}{2}$ outside

$$\frac{\partial w}{\partial t} = \Delta w + \frac{1}{\varepsilon^2} w(1-w)(2w-1).$$



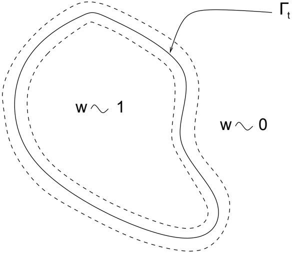
The Allen-Cahn equation and curvature flow

$d(x,t) =$ signed distance x to $\mathbf{\Gamma}_t$

$\mathbf{\Gamma}_0 = \{x \in \mathbb{R}^2 : w_0(x) = \frac{1}{2}\}$

$w_0 > \frac{1}{2}$ inside $\mathbf{\Gamma}$, $< \frac{1}{2}$ outside

$$\frac{\partial w}{\partial t} = \Delta w + \frac{1}{\varepsilon^2} w(1-w)(2w-1).$$



Theorem (Chen 1992)
Fix $T^* \in (0,T)$. Let $k \in \mathbb{N}$. There exists $\varepsilon(k) > 0$, and $a(k), c(k) \in (0,\infty)$ such that for all $\varepsilon \in (0,\varepsilon(k))$ and t satisfying $a\varepsilon^2|\log\varepsilon| \le t \le T^*$,

1. for x such that $d(x,t) \ge c\varepsilon|\log\varepsilon|$, we have $w(t,x) \ge 1-\varepsilon^k$;
2. for x such that $d(x,t) \le -c\varepsilon|\log\varepsilon|$, we have $w(t,x) \le \varepsilon^k$.

A probabilistic proof (E. Freeman, Penington, 2017)

Ternary branching Brownian motion

- Individual lifetime $\texttt{Exp}(1/\varepsilon^2)$;
- During lifetime follows Brownian motion;
- Replaced by three offspring.

Majority voting in (Historical) BBM

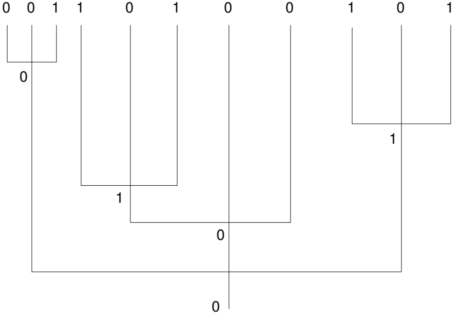
Adaptation of idea of de Masi, Ferrari & Lebowitz (1986)
$\boldsymbol{W}(t) =$ historical ternary BBM.

For a fixed function $w_0 : \mathbb{R}^2 \to [0,1]$, define a voting procedure on $\boldsymbol{W}(t)$ as follows.

1. Each leaf, independently, votes 1 with probability $w_0(W_i(t))$ and otherwise votes 0.
2. At each branch point the vote of the parent particle is the majority vote of the votes of its three children.

This defines an iterative voting procedure, which runs inwards from the leaves of $\boldsymbol{W}(t)$ to the root.
Define $\mathbb{V}_{w_0}(\boldsymbol{W}(t))$ to be the vote associated to the root.

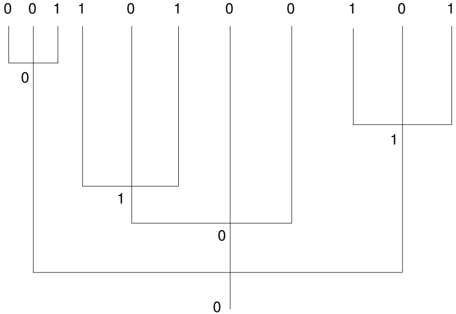
Majority voting and the Allen-Cahn equation



$\boldsymbol{W}(t) =$ historical BBM, branching rate $\frac{1}{\varepsilon^2}$; $w_0 : \mathbb{R}^2 \to [0,1]$.

$$w(t,x) = \mathbb{P}_x^{\varepsilon}\left[\mathbb{V}_{w_0}(\boldsymbol{W}(t)) = 1\right]$$

Majority voting and the Allen-Cahn equation

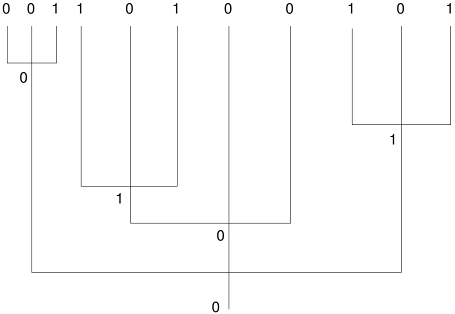


$\boldsymbol{W}(t) =$ historical BBM, branching rate $\frac{1}{\varepsilon^2}$; $w_0 : \mathbb{R}^2 \to [0,1]$.

$$w(t,x) = \mathbb{P}_x^{\varepsilon}\left[\mathbb{V}_{w_0}(\boldsymbol{W}(t)) = 1\right]$$

Note that if probability of voting 1 is w, the probability that the majority of 3 independent votes is 1 is $w^3 + 3w^2(1-w) = w(1-w)(2w-1) + w$.

Majority voting and the Allen-Cahn equation



$\boldsymbol{W}(t) =$ historical BBM, branching rate $\frac{1}{\varepsilon^2}$; $w_0 : \mathbb{R}^2 \to [0,1]$.

$$w(t,x) = \mathbb{P}_x^{\varepsilon}\left[\mathbb{V}_{w_0}(\boldsymbol{W}(t)) = 1\right]$$

solves

$$\frac{\partial w}{\partial t} = \Delta w + \frac{1}{\varepsilon^2} w(1-w)(2w-1), \quad w(0,x) = w_0(x).$$

Probabilistic proof of Chen's result

Representation reduces result to

1. for x with $d(x,t) \geq c\varepsilon|\log \varepsilon|$, $\mathbb{P}_x^\varepsilon\left[\mathbb{V}_{w_0}(\boldsymbol{W}(t)) = 1\right] \geq 1 - \varepsilon^k$;
2. for x with $d(x,t) \leq -c\varepsilon|\log \varepsilon|$, $\mathbb{P}_x^\varepsilon\left[\mathbb{V}_{w_0}(\boldsymbol{W}(t)) = 1\right] \leq \varepsilon^k$.

Probabilistic proof of Chen's result

Representation reduces result to

1. for x with $d(x,t) \geq c\varepsilon|\log\varepsilon|$, $\mathbb{P}^{\varepsilon}_{x}\left[\mathbb{V}_{w_0}(\boldsymbol{W}(t)) = 1\right] \geq 1 - \varepsilon^{k}$;
2. for x with $d(x,t) \leq -c\varepsilon|\log\varepsilon|$, $\mathbb{P}^{\varepsilon}_{x}\left[\mathbb{V}_{w_0}(\boldsymbol{W}(t)) = 1\right] \leq \varepsilon^{k}$.

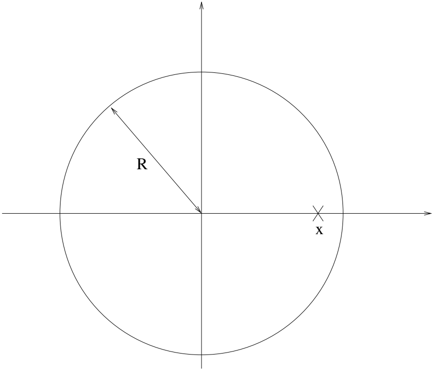
Two mechanisms:

- Majority voting amplifies voting bias;
  $(p > \frac{1}{2} \implies p^3 + 3p^2(1-p) > p$;
  $p < \frac{1}{2} \implies p^3 + 3p^2(1-p) < p)$
- for two-dimensional BM W and one-dimensional BM B, couple so that $d(W_s, t-s) \approx B_s$ when W_s is close to $\mathbf{\Gamma}_{t-s}$ (uses regularity assumptions on initial condition)

Some heuristics

Small $\varepsilon \implies$ many rounds of majority voting $\rightsquigarrow$ generation of an interface.

Suppose there is already a sharp (circular) interface.



For the point x,

$\mathbb{P}_x[W_{\delta t}$ outside ball$] = 1/2$

$\mathbb{P}_x[B_{\delta t} + \frac{1}{R}\delta t > R] = 1/2$

$x = R - \frac{1}{R}\delta t$.

What if homozygotes not equally fit?

Relative fitnesses:

aa	aA	AA
$1+\gamma s$	$1-s$	1

Equation becomes

$$\frac{\partial w}{\partial t} = \Delta w + sw(1-w)((2+\gamma)w-1).$$

Take $\gamma = \mathcal{O}(\varepsilon)$ and rescale:

$$\frac{\partial w}{\partial t} = \Delta w + \frac{1}{\varepsilon^2} w(1-w)(2w-(1-\nu\varepsilon)).$$

Modified voting scheme

$$\frac{\partial w}{\partial t} = \Delta w + \frac{1}{\varepsilon^2} w(1-w)(2w-(1-\nu\varepsilon)).$$

- Ternary branching Brownian motion, branching rate $(1+\varepsilon\nu)/\varepsilon^2$;
- Each leaf independently votes 1 with probability $w_0(W_i(t))$;
- At branch point, parental vote is majority vote of children *unless* precisely one offspring vote is 1 in which case the parent votes 1 with probability $2\varepsilon\nu/(3+3\varepsilon\nu)$.

Modified voting scheme

$$\frac{\partial w}{\partial t} = \Delta w + \frac{1}{\varepsilon^2} w(1-w)(2w-(1-\nu\varepsilon)).$$

- Ternary branching Brownian motion, branching rate $(1+\varepsilon\nu)/\varepsilon^2$;
- Each leaf independently votes 1 with probability $w_0(W_i(t))$;
- At branch point, parental vote is majority vote of children *unless* precisely one offspring vote is 1 in which case the parent votes 1 with probability $2\varepsilon\nu/(3+3\varepsilon\nu)$.

The probabilistic representation is far from unique

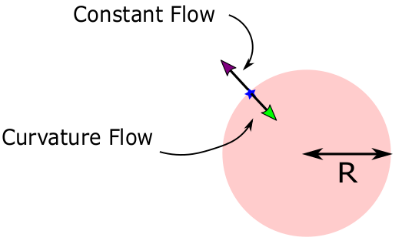
Sensitivity to asymmetry (Gooding, 2018)

$$\frac{\partial w}{\partial t} = \Delta w + \frac{1}{\varepsilon^2} w(1-w)(2w-(1-\nu\varepsilon)).$$

Limit a mixture of curvature flow and 'constant flow':

$$\frac{\partial \mathbf{\Gamma}_t(u)}{\partial t} = \big(-\nu + \kappa_t(u)\big)\mathbf{n}_t(u).$$

Defined up to fixed time T



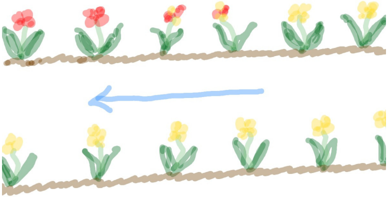
Invasions

$$\frac{\partial w}{\partial t} = \Delta w + \frac{1}{\varepsilon^2} w(1-w)(2w-(1-\nu\varepsilon)).$$

In $d=1$, travelling wave solution (pushed wave)

$$w(x,t) = \left(1+\exp\left(-\frac{x+\nu t}{\varepsilon}\right)\right)^{-1}$$

wave speed $-\nu$, connects 0 at $-\infty$ to 1 at ∞



Blocking (E., Gooding, Letter, 2022)

Consider a domain $\Omega \subseteq \mathbb{R}^2$ (and containing the x-axis, say)

When do we have invasion?

$$\frac{\partial w}{\partial t} = \Delta w + \frac{1}{\varepsilon^2} w(1-w)(2w-(1-\nu\varepsilon)), \quad w(0,x) = \mathbf{1}_{x_1 \geq 0}.$$

Theorem (Berestycki et al., 2016) (paraphrased)
Depending on the geometry of the domain:

1. complete invasion;
2. partial propagation;
3. total blocking.

Blocking (E., Gooding, Letter, 2022)

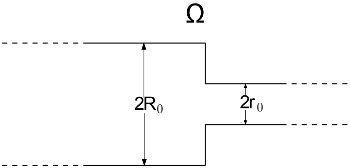
Consider a domain $\Omega \subseteq \mathbb{R}^2$ (and containing the x-axis, say)

When do we have invasion?

$$\frac{\partial w}{\partial t} = \Delta w + \frac{1}{\varepsilon^2} w(1-w)(2w-(1-\nu\varepsilon)), \quad w(0,x) = \mathbf{1}_{x_1 \geq 0}.$$

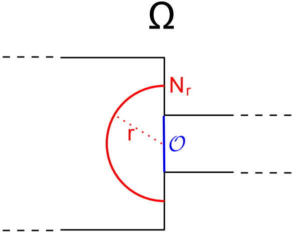
Theorem (Berestycki et al., 2016) (paraphrased)
Depending on the geometry of the domain:

1. complete invasion;
2. partial propagation;
3. total blocking.



A more precise statement

$$\begin{cases} \frac{\partial w}{\partial t} = \Delta w + \frac{1}{\varepsilon^2} w(1-w)(2w-(1-\nu\varepsilon)); \\ \\ \frac{\partial w}{\partial n} = 0, \qquad w(x,0) = 1_{x_1 \geq 0}; \end{cases}$$

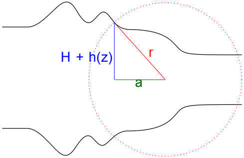


Theorem
Suppose $r_0 < r < \frac{\mathbb{d}-1}{\nu} \wedge R_0$. Let $k \in \mathbb{N}$. Then for $\varepsilon \in (0, \hat{\varepsilon}(k))$

$$x \in \{x = (x_1, \ldots, x_d) : x_1 < -r - M(k)\varepsilon|\log(\varepsilon)|\} \implies w(x,t) \leq \varepsilon^k.$$

Other domains

$$\Omega = \left\{ (x_1, x'), x_1 \in \mathbb{R}, x' \in \mathbb{R}^{\mathbb{d}-1}, \|x'\| \leq H + h(-x_1) \right\}$$



Theorem

Suppose that,

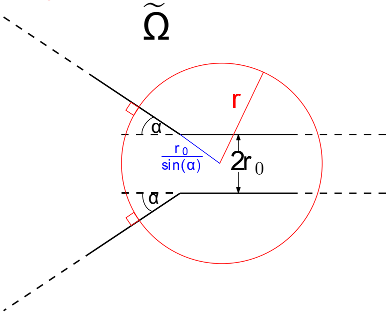
$$\inf_{z>0} \left\{ H + h(z) - \left(\frac{\mathbb{d}-1}{\nu} \right) \frac{h'(z)}{\sqrt{1+h'(z)^2}} \right\} < 0.$$

Fix $k \in \mathbb{N}$. There exist $x_0 < 0$, $\widehat{\varepsilon}(k) > 0$ and $M(k) > 0$ such that for all $\varepsilon \in (0, \widehat{\varepsilon})$, and $t \geq 0$,

$$\text{if } x_1 \leq x_0 - M(k)\varepsilon|\log(\varepsilon)| \text{ then } u^\varepsilon(x,t) \leq \varepsilon^k.$$

Key argument

Key is coupling around a portion of a spherical shell



If $r_0\nu < (\mathbb{d}-1)\sin\alpha$ wave blocked for small ε.

Effect of noise

Two dimensions:

- If genetic drift is weak (population density high), the spread of the favoured type is blocked;
- If genetic drift is strong (population density low), the interface is broken down.

Selection is more effective in bigger populations

Proof uses duality with branching/coalescing system

Somewhat more precisely

- Set up SLFV as for directional selection, except sample three potential parents at selective events.
- Dual process a *ternary* branching and coalescing system.
- Duality via the modified voting scheme on historical process of branching and coalescing lineages.

Somewhat more precisely

- Set up SLFV as for directional selection, except sample three potential parents at selective events.
- Dual process a *ternary* branching and coalescing system.
- Duality via the modified voting scheme on historical process of branching and coalescing lineages.

As for directional selection, we can identify different regimes depending on the relative strength of noise and selection.

Two regimes

Weak noise/selection ratio

$$\gamma_n = \nu\varepsilon_n, \quad u_n = \frac{u}{n^{1-2\beta}}, \quad s_n = \frac{1}{\varepsilon_n^2 n^{2\beta}}$$

$\beta \in (0, 1/4)$, $w(nt, n^\beta x)$ close to solution to deterministic equation: blocking occurs. Note $n s_n u_n = u n^{2\beta} s_n$, diffusive scaling $(n^{2\beta}t, n^\beta x)$ for lineage motion

Two regimes

Weak noise/selection ratio

$$\gamma_n = \nu\varepsilon_n, \quad u_n = \frac{u}{n^{1-2\beta}}, \quad s_n = \frac{1}{\varepsilon_n^2 n^{2\beta}}$$

$\beta \in (0, 1/4)$, $w(nt, n^\beta x)$ close to solution to deterministic equation: blocking occurs. Note $ns_n u_n = un^{2\beta}s_n$, diffusive scaling $(n^{2\beta}t, n^\beta x)$ for lineage motion

Strong noise/selection ratio

$$\frac{s_n n^{2\beta}}{u_n \log n} \to 0 \quad \liminf_{n\to\infty} u_n \log n = \infty \text{ and } d = 2.$$

$\beta \in (0, 1/2)$, u_n any sequence of impact parameters, $\widehat{u}_n = u_n n^{1-2\beta}$. $\mathbb{E}\big[w(nt/\widehat{u}_n, n^\beta x)\big]$ behaves like solution to the heat equation: the interface is broken down. Note $(n/\widehat{u}_n)u_n s_n = n^{2\beta}s_n$, diffusive scaling $(n^{2\beta}t, n^\beta x)$ for lineage motion