The $\Lambda$-Fleming-Viot process and a connection with Wright-Fisher diffusion

Bob Griffiths
University of Oxford
A \( d \)-dimensional \( \Lambda \)-Fleming-Viot process \( \{X(t)\}_{t \geq 0} \) representing frequencies of \( d \) types of individuals in a population has a generator described by

\[
\mathcal{L} g(x) = \int_0^1 \sum_{i=1}^{d} x_i \left( g(x(1 - y) + ye_i) - g(x) \right) \frac{F(dy)}{y^2}.
\]

The population is partitioned at events of change by choosing type \( i \in [d] \) to reproduce with probability \( x_i \), then rescaling the population with additional offspring \( y \) of type \( i \) to be \( x(1 - y) + ye_i \) at rate \( y^{-2}F(dy) \).
Examples

Eldon and Wakeley (2006). A model where $F$ has a single point of increase in $(0, 1]$ with a possible atom at 0.

A natural class that arises from discrete models is when $F$ has a Beta$(\alpha, \beta)$ density, particularly a Beta$(2 - \alpha, \alpha)$ density coming from a discrete model where the offspring distribution tails are asymptotic to a power law of index $\alpha$. Birkner, Blath, Capaldo, Etheridge, Möhle, Schweinsberg, Wakolbinger (2005) give a connection to stable processes.

Birkner and Blath (2009) describe the $\Lambda$-Fleming-Viot process and discrete models whose limit gives rise to it.
If $F$ has a single atom at 0, then $\{X_t\}_{t \geq 0}$ is the $d$-dimensional Wright-Fisher diffusion process with generator

$$\mathcal{L} = \frac{1}{2} \sum_{i,j=1}^{d} x_i (\delta_{ij} - x_j) \frac{\partial^2}{\partial x_i \partial x_j}$$

$X_1(t)$ is a one-dimensional Wright-Fisher diffusion process with generator

$$\mathcal{L} = \frac{1}{2} x (1 - x) \frac{\partial^2}{\partial x^2}$$
The $\Lambda$-coalescent process is a random tree process back in time which has multiple merger rates for a specific $k$ lineages coalescing while $n$ edges in the tree of

$$\lambda_{nk} = \int_0^1 x^k (1 - x)^{n-k} \frac{F(dx)}{x^2}, \quad k \geq 2$$

After coalescence there are $n - k + 1$ edges in the tree.

This process was introduced by Pitman (1999), Sagitov (1999) and has been extensively studied. Berestycki (2009); recent results in the $\Lambda$-coalescent.
There is a connection between continuous state branching processes and the $\Lambda$-coalescent. The connection is through the Laplace exponent

$$\psi(q) = \int_0^1 \left( e^{-qy} - 1 + qy \right) y^{-2} F(dy)$$

Bertoin and Le Gall (2006) showed that the $\Lambda$-coalescent comes down from infinity under the same condition that the continuous state branching process becomes extinct in finite time, that is when

$$\int_1^\infty \frac{dq}{\psi(q)} < \infty$$
Some papers on the $\Lambda$-coalescent


A Wright-Fisher generator connection

**Theorem**
Let $\mathcal{L}$ be the $\Lambda$-Fleming-Viot generator, $V$ be a uniform random variable on $[0,1]$, $U$ a random variable on $[0,1]$ with density $2u$, $0 < u < 1$ and $W = YU$, where $Y$ has distribution $F$ and $V, U, Y$ are independent. Denote the second derivatives of a function $g(x)$ by $g_{ij}(x)$.

Then

$$
\mathcal{L}g(x) = \frac{1}{2} \sum_{i,j=1}^{d} x_i (\delta_{ij} - x_j) \mathbb{E} \left[ g_{ij}(x(1 - W) + WV e_i) \right]
$$

where expectation $\mathbb{E}$ is taken over $V, W$. 
Wright-Fisher generator

\[
\mathcal{L} g(\mathbf{x}) = \frac{1}{2} \sum_{i,j=1}^{d} x_i \delta_{ij} - x_j g_{ij}(\mathbf{x})
\]

Λ-Fleming-Viot generator

\[
\mathcal{L} g(\mathbf{x}) = \frac{1}{2} \sum_{i,j=1}^{d} x_i (\delta_{ij} - x_j) \mathbb{E}\left[ g_{ij}(\mathbf{x}(1 - W) + WV \mathbf{e}_i) \right]
\]
Method of proof

\[ \mathcal{L} g(\mathbf{x}) = \int_0^1 \sum_{i=1}^d x_i \left( g(\mathbf{x}(1 - y) + y\mathbf{e}_i) - g(\mathbf{x}) \right) \frac{F(dy)}{y^2} \]

\[ \mathcal{L} g(\mathbf{x}) = \frac{1}{2} \sum_{i,j=1}^d x_i (\delta_{ij} - x_j) \mathbb{E} \left[ g_{ij}(\mathbf{x}(1 - W) + W\mathbf{V} \mathbf{e}_i) \right] \]

Show that the generators have the same answer acting on

\[ g(\mathbf{x}) = \exp \left\{ \sum_{i=1}^d \eta_i x_i \right\}, \quad \eta \in \mathbb{R}^d \]
1-dimensional generator

Wright-Fisher diffusion generator

\[ \mathcal{L}g(x) = \frac{1}{2}x(1 - x)g''(x) \]

Λ-Fleming-Viot process generator

\[ \mathcal{L}g(x) = \frac{1}{2}x(1 - x)\mathbb{E}\left[g''(x(1 - W) + WV)\right] \]

or

\[ \mathcal{L}g(x) = \frac{1}{2}x(1 - x)\mathbb{E}\left[\frac{g'(x(1 - W) + W) - g'(x(1 - W))}{W}\right] \]
The Laplace transform of $W$ is related to the Laplace exponent by

$$
\mathbb{E}[e^{-\eta W}] = 2 \int_0^1 \frac{e^{-\eta y} - 1 + \eta y}{(y\eta)^2} F(dy)
$$

$W = UY$ is continuous in $(0, 1)$ with a possible atom at $0$.

$$P(W = 0) = P(Y = 0)$$
Adding mutation

The generator has an additional term added of

\[ \frac{\theta}{2} \sum_{i=1}^{d} \left( \sum_{j=1}^{d} p_{ji}x_j - x_i \right) \frac{\partial}{\partial x_i} \]

If mutation is parent independent \( \theta p_{ij} = \theta_j \), not depending on \( i \), and the additional term is

\[ \frac{1}{2} \sum_{i=1}^{d} \left( \sum_{j=1}^{d} \theta_j x_j - \theta x_i \right) \frac{\partial}{\partial x_i} \]
Eigenstructure of the $\Lambda$-Fleming-Viot process

Theorem
Let $\{\lambda_n\}, \{P_n(x)\}$ be the eigenvalues and eigenvectors of $\mathcal{L}$, the generator which includes mutation, satisfying

$$\mathcal{L}P_n(x) = -\lambda_n P_n(x)$$

Denote the $d-1$ eigenvalues of the mutation matrix $P$ which have modulus less than 1 by $\{\phi_k\}_{k=1}^{d-1}$.

The eigenvalues of $\mathcal{L}$ are

$$\lambda_n = \frac{1}{2} n(n-1) \mathbb{E} \left[ (1 - W)^{n-2} \right] + \frac{\theta}{2} \sum_{k=1}^{d-1} (1 - \phi_k) n_k$$
Polynomial Eigenvectors

Denote the $d-1$ eigenvalues of $P$ which have modulus less than 1 by $\{\phi_k\}_{k=1}^{d-1}$ corresponding to eigenvectors which are rows of a $d - 1 \times d$ matrix $R$ satisfying

$$
\sum_{i=1}^{d} r_{ki} p_{ji} = \phi_k r_{kj}, \ k = 1, \ldots, d - 1.
$$

Define a $d - 1$ dimensional vector $\xi = Rx$.

The polynomials $P_n(x)$ are polynomials in the $d - 1$ terms in $\xi = (\xi_1, \ldots, \xi_{d-1})$ whose only leading term of degree $n$ is

$$
\prod_{j=1}^{d-1} \xi_j^{n_j}
$$
In the parent independent model of mutation

\[ \lambda_n = \lambda_n = \frac{1}{2} n \left\{ (n - 1) \mathbb{E} \left[ (1 - W)^{n-2} \right] + \theta \right\} \]

repeated \( \binom{n+d-2}{n} \) times with non-unique polynomial eigenfunctions within the same degree \( n \).

The non-unit eigenvalues of the mutation matrix with identical rows are zero.

Wright-Fisher diffusion, general mutation structure.

\[ \lambda_n = \frac{1}{2} n(n - 1) + \frac{\theta}{2} \sum_{k=1}^{d-1} (1 - \phi_k) n_k \]
$\Lambda$-coalescent eigenvalues and rates

$$\frac{1}{2}n(n - 1)E[(1 - W)^{n-2}] = \sum_{k=2}^{\infty} \binom{n}{k} \lambda_{nk}$$

which is the total jump rate away from $n$ individuals.

These are the eigenvalues in the $\Lambda$-coalescent tree.
The individual rates can be expressed as

\[
\binom{n}{k} \lambda_{nk} = \binom{n}{k} \int_0^1 y^k (1 - y)^{n-k} \frac{F(dy)}{y^2}
\]

\[
= \frac{n}{2} \mathbb{E} \left[ \frac{P_k(n, W) - P_{k-1}(n, W)}{W^2} \right],
\]

where

\[
P_k(n, w) = \binom{n - 1}{k - 1} (1 - w)^{n-k} w^k
\]

is a negative binomial probability of a waiting time of \(n\) trials to obtain \(k\) successes, where the success probability is \(w\).
Two types

The generator is specified by

$$\mathcal{L}g(x) = \frac{1}{2}x(1 - x)\mathbb{E}\left[g''\left(x(1 - W) + WV\right)\right] + \frac{1}{2}(\theta_1 - \theta x)g'(x)$$

The eigenvalues are

$$\lambda_n = \frac{1}{2}n \left\{ (n - 1)\mathbb{E}\left[(1 - W)^{n-2}\right] + \theta \right\}$$

and the eigenvectors are polynomials satisfying

$$\mathcal{L}P_n(x) = -\lambda_n P_n(x), \ n \geq 1.$$
Polynomial eigenvectors

\[
\frac{1}{2}x(1-x)E \left[ \frac{P'_n(x(1-W)+W) - P'_n(x(1-W))}{W} \right] \\
+ \frac{1}{2}(\theta_1 - \theta x)P'_n(x) \\
= \frac{1}{2}n \left[ (n-1)E[(1-W)^{n-2}] + \theta \right] P_n(x)
\]

The monic polynomial \( P_n(x) \) is uniquely defined by recursion of its coefficients.
Stationary distribution $\psi(x)$

\[ \int_0^1 \mathcal{L} g(x) \psi(x) \, dx = 0 \]

\[ \sigma^2(x) = x(1-x), \; \mu(x) = \theta_1 - \theta x \]

\[ k(x) = \mathbb{E} \left[ (1 - W)^{-2} g(x(1-W) + VW) \right] \]

An equation for the stationary distribution

\[ 0 = \int_0^1 \left[ k(x) \frac{1}{2} \frac{d^2}{dx^2} \left[ \sigma^2(x) \psi(x) \right] - g(x) \frac{d}{dx} \left[ \mu(x) \psi(x) \right] \right] \, dx \]

\[ + k(x) \frac{d}{dx} \left[ \frac{1}{2} \sigma^2(x) \psi(x) \right] \bigg|_0^1 + g(x) \mu(x) \psi(x) \bigg|_0^1 \]
In a diffusion process \( k(x) = g(x) \) and the boundary terms vanish. Then there is a solution found by solving

\[
\frac{1}{2} \frac{d^2}{dx^2} \left[ \sigma^2(x) \psi(x) \right] - \frac{d}{dx} \left[ \mu(x) \psi(x) \right] = 0
\]

however \( k(x) \neq g(x) \) so we do not have an equation like this.
Green's function, $\gamma(x)$

Solve, for a given function $g(x)$

$$\mathcal{L}\gamma(x) = -g(x), \quad \gamma(0) = \gamma(1) = 0.$$ 

Then

$$\gamma(x) = \int_0^1 G(x, \xi) g(\xi) d\xi$$

A non-linear equation, equivalent to

$$\frac{1}{2}x(1 - x) \mathbb{E}\left[ \gamma''(x(1 - W) + VW) \right] = -g(x)$$
Green’s function solution

Define

\[ k(x) = \mathbb{E} \left[ (1 - W)^{-2\gamma} (x(1 - W) + VW) \right] \]

then

\[ k''(x) = -2 \frac{g(x)}{x(1 - x)} \]

with a solution

\[ k(x) = k(0)(1 - x) + k(1)x + (1 - x) \int_0^x \frac{2g(\eta)}{1 - \eta} d\eta + x \int_x^1 \frac{2g(\eta)}{\eta} d\eta \]
Mean time to absorption

If \( g(x) = 1, \ x \in (0, 1) \) then \( \gamma(x) \) is the mean time to absorption at 0 or 1 when \( X(0) = x \).

\[
\gamma(x) = k(0)(1 - x) + k(1)x
\]

\[
+ (1 - x) \int_0^x \frac{2}{(1 - \eta)} d\eta + x \int_x^1 \frac{2}{\eta} d\eta
\]

There is a non-linear equation to solve of

\[
k(x) = k(0)(1 - x) + k(1)x - 2(1 - x) \log(1 - x) - 2x \log x
\]

where

\[
k(x) = E\left[ (1 - W)^{-2} \gamma(x(1 - W) + VW) \right]
\]
Stationary distribution, $\Lambda$-Fleming process with mutation

Let $Z$ be a random variable with the size-biassed distribution of $X$, $Z_*$ a size-biassed $Z$ random variable and $Z^*$ a size-biassed random variable with respect to $1 - Z$.

Let $B$ be Bernoulli random variable, independent of the other random variables in the following equation such that

$$P(B = 1) = \frac{\theta - \theta_1}{\theta(\theta_1 + 1)}$$

An interesting distributional identity

$$VZ_* =^\mathcal{D} (1 - B)VZ + B(Z^*(1 - W) + WV)$$
The frequency spectrum in the infinitely-many-alleles model

Take a limit from a $d$-allele model with $\theta_i = \theta/d$, $i \in [d]$. The limit is a point process $\{X_i\}_{i=1}^{\infty}$. The 1-dimensional frequency spectrum $h(x)$ is a non-negative measure such that for suitable functions $k$ on [0,1] in the stationary distribution

$$\mathbb{E}\left[\sum_{i=1}^{\infty} k(X_i)\right] = \int_{0}^{1} k(x)h(x)dx.$$

Symmetry in the $d$-allele model shows that

$$\int_{0}^{1} k(x)h(x)dx = \lim_{d \to \infty} d\mathbb{E}\left[k(X_1)\right].$$

The classical Wright-Fisher diffusion gives rise to the Poisson-Dirichlet process with a frequency spectrum of

$$h(x) = \theta x^{-1}(1-x)^{\theta-1}, \ 0 < x < 1.$$
Let $Z$ have a density

$$f(z) = zh(z), \quad 0 < z < 1$$

Interesting identity

$$VZ_* \overset{\mathcal{D}}{=} Z^*(1 - W) + WV$$

where $Z_*$ is size-biassed with respect to $Z$, and $Z^*$ is size-biassing with respect to $1 - Z$.

The constant $\theta$ appears in the identity through scaling in the size-biassed distributions.

Limit distribution of excess life in a renewal process

$$P(VZ_* > \eta) = \int_{\eta}^{1} \frac{P(Z > z)}{\mathbb{E}[Z]} dz$$
Typed dual $\Lambda$-coalescent

The $\Lambda$-Fleming-Viot process is dual to the system of coalescing lineages $\{L(t)\}_{t \geq 0}$ which takes values in $\mathbb{Z}^d_+$ and for which the transition rates are, for $i, j \in [d]$ and $l \geq 2$,

$$q_\Lambda(\xi, \xi - e_i(l - 1)) = \int_{[0,1]} \binom{|\xi|}{l} y^l(1 - y)|\xi| - l F(dy) \frac{y^2}{y^2} \times \frac{\xi_i + 1 - l \mathcal{M}(\xi - e_i(l - 1))}{|\xi| + 1 - l \mathcal{M}(\xi)}$$

$$q_\Lambda(\xi, \xi + e_i - e_j) = \mu_{ij}(\xi_i + 1 - \delta_{ij}) \frac{\mathcal{M}(\xi + e_i - e_j)}{\mathcal{M}(\xi)}$$

The process is constructed as a moment dual from the generator.
A different dual process

Define a sequence of monic polynomials \( \{g_n(x)\} \) by the generator equation

\[
\frac{1}{2}x(1-x)Eg''_n(x(1-W)+VW) + \frac{1}{2}(\theta_1 - \theta x)g'_n(x)
\]

\[
= \binom{n}{2}E(1-W)^{n-2}[g_{n-1}(x) - g_n(x)] + n\frac{1}{2}[\theta_1 g_{n-1}(x) - \theta g_n(x)]
\]

The defining equation mimics the Wright-Fisher diffusion acting on test functions \( g_n(x) = x^n \)

\[
\frac{1}{2}x(1-x)\frac{d^2}{dx^2}x^n + \frac{1}{2}(\theta_1 - \theta x)\frac{d}{dx}x^n
\]

\[
= \binom{n}{2}(x^{n-1} - x^n) + \frac{1}{2}n(\theta_1 x^{n-1} - \theta x^n)
\]
Jacobi polynomial analogues
The eigenfunctions are polynomials \( \{P_n(x)\} \) satisfying

\[
\mathcal{L} P_n(x) = \lambda_n P_n(x)
\]

\[
P_n(x) = g_n(x) + \sum_{r=0}^{n-1} c_{nr} g_r(x)
\]

The coefficients are

\[
c_{nr} = \frac{\lambda_{r+1} \cdots \lambda_n}{(\lambda_r - \lambda_n) \cdots (\lambda_{n-1} - \lambda_n)}
\]

where

\[
\lambda_n = \frac{n}{2} \left[ (n - 1) \mathbb{E}(1 - W)^{n-2} + \theta \right]
\]

\[
\lambda_n^\circ = \frac{n}{2} \left[ (n - 1) \mathbb{E}(1 - W)^{n-2} + \theta_1 \right]
\]
In the stationary distribution

$$\mathbb{E}[g_n(X)] = \omega_n$$

where $\omega_n$ is a Beta moment analog

$$\omega_n = \frac{\prod_{j=1}^{n} \left( (j - 1)\mathbb{E}(1 - W)^{j-2} + \theta_1 \right)}{\prod_{j=1}^{n} \left( (j - 1)\mathbb{E}(1 - W)^{j-2} + \theta \right)}$$

Let

$$h_n = \frac{g_n}{\omega_n}$$

so

$$\mathbb{E}[h_n(X)] = 1$$
Dual Generator

\[ \mathcal{L}h_n = \lambda_n \left[ h_{n-1} - h_n \right] \]

Dual equation

\[ \mathbb{E}_{X(0)=x}\left[ h_n(X(t)) \right] = \mathbb{E}_{N(0)=n}\left[ h_N(t)(x) \right] \]

where \( \{N(t), t \geq 0\} \) is a death process with rates

\[ \lambda_n = \frac{1}{2} n \left( (n - 1) \mathbb{E}[ (1 - W)^{n-2} ] + \theta \right) \]

The process \( \{N(t), t \geq 0\} \) comes down from infinity if and only if

\[ \sum_{n=2}^{\infty} \lambda_n^{-1} < \infty \]

which implies the \( \Lambda \)-coalescent comes down from infinity.
The transition functions are then

\[
P(N(t) = j \mid N(0) = i) = \sum_{k=j}^{i} e^{-\lambda_k t} (-1)^{k-j} \frac{\prod\{l:j \leq l \leq k+1\} \lambda_l}{\prod\{l:j \leq l \leq k+1, l \neq k\} (\lambda_l - \lambda_k)}
\]

\[
P(N(t) = j \mid N(0) = \infty)
\]

is well defined if \(N(t)\) comes down from infinity.

In the Kingman coalescent the death rates are \(n(n - 1 + \theta)/2\) and \(N(t)\) describes the number of non-mutant lineages at time \(t\) back in the population.

Let $P(x)$ be the probability that the 1st type fixes, starting from an initial frequency of $x$. $P(x)$ is the solution of

$$L^\sigma P(x) = \frac{1}{2} x (1 - x) P''(x) - \sigma x (1 - x) P'(x) = 0$$

with $P(0) = 0$, $P(1) = 1$. The solution of this differential equation is

$$P(x) = \frac{e^{2\sigma x} - 1}{e^{2\sigma} - 1}$$
λ-Fleming Viot process. Fixation probability with selection.

Let \( P(x) \) be the probability that the 1st type fixes, starting from an initial frequency of \( x \). \( P(x) \) is the solution of

\[
\mathcal{L}^\sigma P(x) = \frac{1}{2}x(1-x)\mathbb{E}\left[P''((1-W)+(x-W)V)\right]-\sigma x(1-x)P'(x) = 0
\]

with \( P(0) = 0 \), \( P(1) = 1 \).

Der, Epstein and Plotkin (2011, 2012). For some measures \( F \) and \( \sigma \) it is possible that \( P(x) = 1 \) or \( 0 \) for all \( x \in (0, 1) \). If \( \sigma > 0 \), fixation is certain if

\[
\sigma > -\int_0^1 \frac{\log(1 - y)}{y^2} F(dy)
\]
A computational solution for $P(x)$ when fixation or loss is not certain from $x \in (0, 1)$.

Define a sequence of polynomials $\{h_n(x)\}_{n=0}^\infty$ for a choice of pre-specified constants $\{h_n(0)\}$ as solutions of

$$
\mathbb{E} \left[ \frac{h_n(x(1 - W) + W) - h_n(x(1 - W))}{W} \right] = nh_{n-1}(x)
$$

where the leading coefficient in $h_n(x)$ is

$$
\frac{1}{\prod_{j=1}^{n-1} \mathbb{E}[(1 - W)^j]}
$$
Polynomial solution for $P(x)$

$$P(x) = (e^{2\sigma} - 1)^{-1} \sum_{n=1}^{\infty} \frac{(2\sigma)^n}{n!} H_n(x),$$

where $\{H_n(x)\}$ are polynomials derived from

$$H_n(x) = \int_{0}^{x} nh_{n-1}(\xi)d\xi$$

and the constants $\{h_n(0)\}$ are chosen so that

$$\int_{0}^{1} nh_{n-1}(\xi)d\xi = 1$$

The coefficients of $H_n(x)$ are well defined by a recurrence relationship with the coefficients of $H_{n-1}(x)$. 