Ancestral lines under selection and recombination

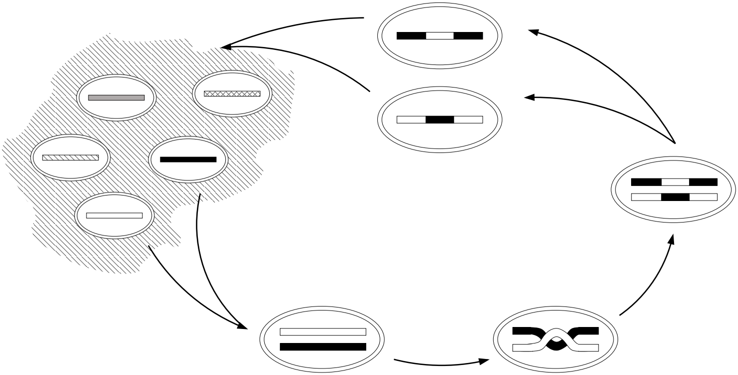
Ellen Baake

Bielefeld University

joint work with Frederic Alberti

1. The selection-recombination equation (forward in time)
2. Its solution via genealogical thinking

Recombination



Sequences and selection

individual: sequence of n sites, $S = \{1, \dots, n\}$

types: $x := (x_1, \dots, x_n) \in \bigtimes_{i \in S} X_i =: X, \; X_i = \{0, 1\}$

marginal types: $x_U := (x_i)_{i \in U}, \; U \subseteq S$

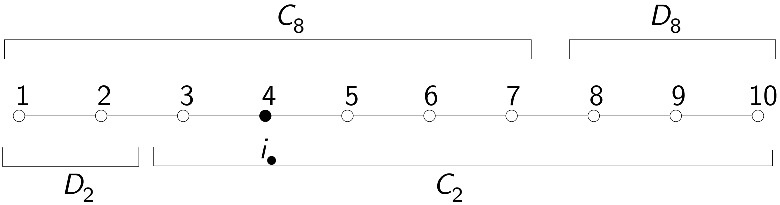
reproduction and selection:

- single selected site: $i_\bullet \in S$ (all others ‘neutral’)
- types x with $x_{i_\bullet} = 1$: reproduce at rate 1 (‘bad, unfit’)
- types x with $x_{i_\bullet} = 0$: reproduce at rate $1 + s$ (‘good, fit’)
- offspring replaces randomly chosen individual

Ordering the sites

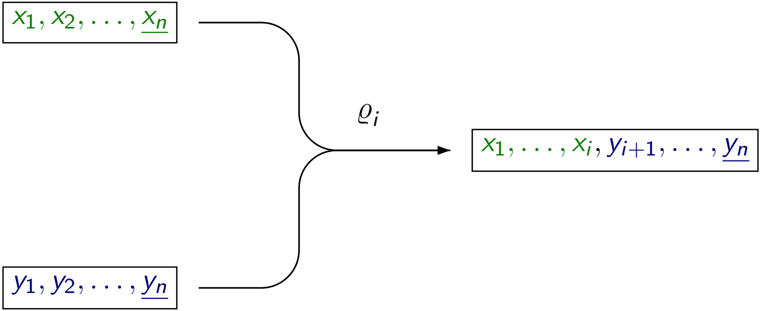
partial order on S:

- $i \preccurlyeq j$ if either $i_\bullet \leqslant i \leqslant j$ or $i_\bullet \geqslant i \geqslant j$
- $i \prec j$ if $i \preccurlyeq j$ and $i \neq j$
- $D_i := \{j \in S \mid i \preccurlyeq j\}$ (i-tail), $C_i := S \setminus D_i$ (i-head), $i \in S \setminus i_\bullet$



- $\rightsquigarrow i_\bullet \in C_i$
- (single-crossover) recombination between C_i and D_i (partitions S into C_i and D_i) at rate ϱ_i

Sequences and recombination

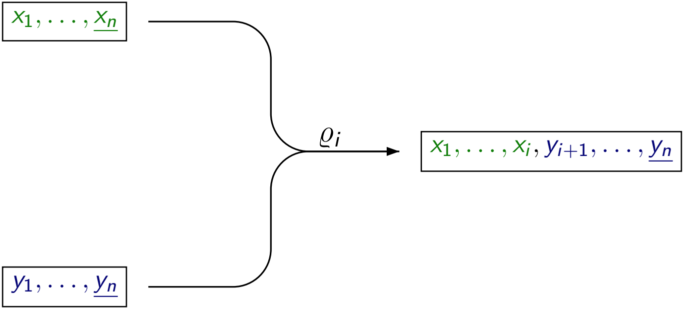


Selection-recombination equation

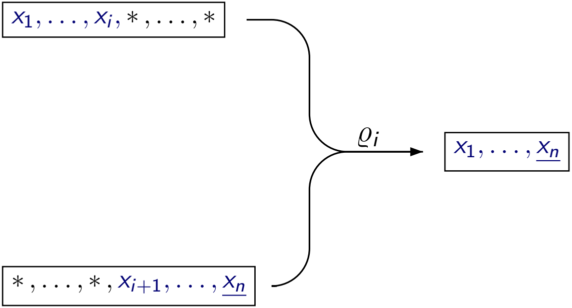
- $\omega_t = \big(\omega_t(x)\big)_{x \in X}$ probability measure on X
  ($\omega_t(x)$ proportion of individuals of type x at time t, $x \in X$)
- solves selection-recombination equation (SRE)

$$\dot{\omega}_t(x) = s\big[\delta_{x_{i_\bullet},0} - \omega_t(*,0,*)\big]\,\omega_t(x) + \sum_{i \in S \setminus i_\bullet} \varrho_i \big[\omega_t(x_{C_i},*)\,\omega_t(*,x_{D_i}) - \omega_t(x)\big], \quad x \in X$$

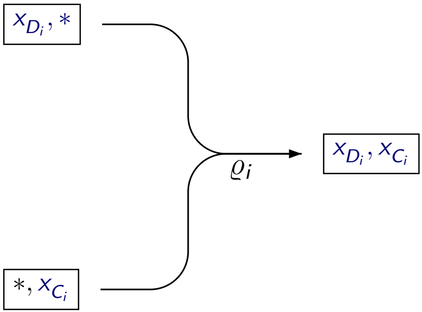
Recombination



Recombination



Recombination



Selection-recombination equation

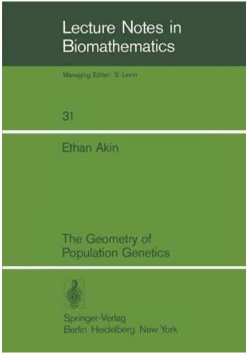
- $\omega_t = \big(\omega_t(x)\big)_{x \in X}$ probability measure on X
  ($\omega_t(x)$ proportion of individuals of type x at time t, $x \in X$)
- solves selection-recombination equation (SRE)

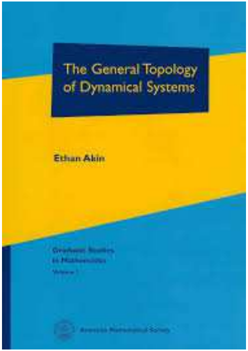
$$\dot{\omega}_t(x) = s\big[\delta_{x_{i_\bullet},0} - \omega_t(*,0,*)\big]\,\omega_t(x) + \sum_{i \in S \setminus i_\bullet} \varrho_i \big[\omega_t(x_{C_i},*)\,\omega_t(*,x_{D_i}) - \omega_t(x)\big], \quad x \in X$$

Research context

- hitchhiking: neutral x_i ($i \neq i_\bullet$) selectively favoured if linked with $x_{i_\bullet} = 0$
- Akin 1979: "The differential equations which model the action of selection and recombination are nonlinear equations which are impossible to solve explicitly."

Ethan Akin


Lecture Notes in
Biomathematics
Managing Editor: S. Levin
31
Ethan Akin
The Geometry of
Population Genetics
Springer-Verlag
Berlin Heidelberg New York


The General Topology
of Dynamical Systems
Ethan Akin

Research context

- hitchhiking: neutral x_i ($i \neq i_\bullet$) selectively favoured if linked with $x_{i_\bullet} = 0$
- Akin 1979: the "selection-recombination equation is a large, nonlinear system of differential equations that are impossible to solve explicitly"
- Stephan, Song, Langley 2006: approximate solution ($n = 3$, wild approximations, technical result, numerical evaluation only, not generalisable)

Recombinators

- canonical projection: for $\varnothing \neq U \subseteq S$,

$$\pi_U : X \to \bigtimes_{i \in U} X_i = X_U, \quad \pi_U(x) = (x_i)_{i \in U} = x_U$$

- marginal measure wrt sites in U: for $\nu \in \mathcal{P}(X)$,

$$\pi_U.\nu = \nu \circ \pi_U^{-1} =: \nu_U$$

  type distribution of sites in U
  for $x_U \in X_U$: $\quad \nu_U(x_U) = \nu(x_U, *)$
- recombinator: for $i \in S \setminus i_\bullet$,

$$\mathcal{P}(X) \longrightarrow \mathcal{P}(X)$$

$$R_i(\nu) := \nu_{C_i} \otimes \nu_{D_i}$$

$$\big(R_i(\nu)\big)(x) = \nu(x_{C_i}, *) \cdot \nu(*, x_{D_i})$$

  distribution of sequences pieced together from random i-heads and i-tails

Selection-recombination equation

$$\begin{aligned}\dot{\omega}_t(x) &= s\big[\delta_{x_{i_\bullet},0} - \omega_t(*,0,*)\big]\,\omega_t(x)\\ &\quad + \sum_{i\in S\setminus i_\bullet} \varrho_i\big[\omega_t(x_{C_i},*)\,\omega_t(*,x_{D_i}) - \omega_t(x)\big], \quad x\in X\end{aligned}$$

or

$$\dot{\omega}_t = s\big(F - f(\omega_t)\big)\omega_t + \sum_{i\in S\setminus i_\bullet} \varrho_i(R_i - \mathbb{1})\omega_t = \Psi_{\text{sel}}(\omega_t) + \Psi_{\text{rec}}(\omega_t)$$

with

$$F\omega(x) = \delta_{x_{i_\bullet},0}\,\omega(x)$$

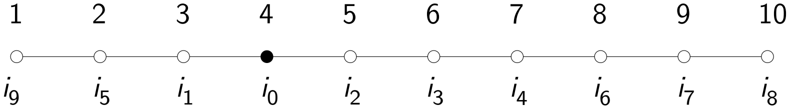
and

$$f(\omega) := \omega_{\{i_\bullet\}}(0)$$

(proportion of fit individuals = proportion of individuals with 0 at selected site)

Recursive solution

- $(i_k)_{0 \leqslant k < n}$ non-decreasing permutation of S (sensu $\preccurlyeq$)
- $i_0 = i_\bullet$
- $C^{(k)} = C_{i_k}, D^{(k)} = D_{i_k}, \varrho^{(k)} := \varrho_{i_k} \quad (k > 0)$



hierarchy of solutions:

- $\big(\omega_t^{(0)}\big)_{t \geqslant 0}$ solution of SRE with $\varrho^{(\ell)} = 0$ for all ℓ
  (pure selection equation)
- $\big(\omega_t^{(k)}\big)_{t \geqslant 0}$ solution of SRE truncated at k
  (with $\varrho^{(\ell)} = 0$ for all $\ell > k$), $\quad 1 \leqslant k < n$

Recursive solution

Theorem

The solution of the pure selection equation is given by

$$\omega_t^{(0)} = \frac{\mathrm{e}^{stF}\omega_0}{\mathrm{e}^{st} f(\omega_0) + 1 - f(\omega_0)} =: \varphi_t(\omega_0),$$

and the solutions $\omega_t^{(k)}$ *can be computed via the recursion*

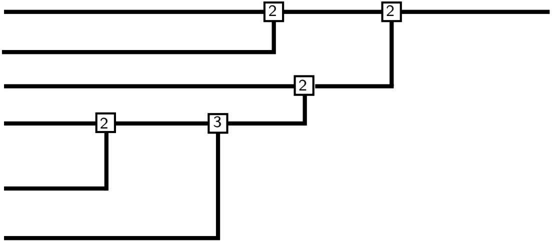
$$\omega_t^{(k)} = \mathrm{e}^{-\varrho^{(k)} t} \omega_t^{(k-1)} + \omega_{C^{(k)},t}^{(k-1)} \otimes \int_0^t \varrho^{(k)} \mathrm{e}^{-\varrho^{(k)}\tau} \omega_{D^{(k)},\tau}^{(k-1)} \, \mathrm{d}\tau.$$

analytical proof ✓

(genealogical) meaning?

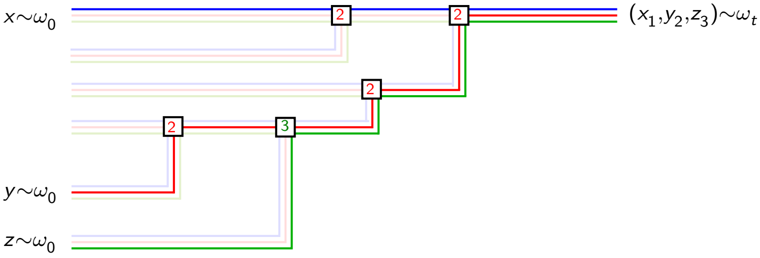
Ancestral graph

i-splitting events on every line at rate ϱ_i, $i \in S \setminus i_\bullet$ (backward!)
$S = \{1, 2, 3\}, \; i_\bullet = 1:$



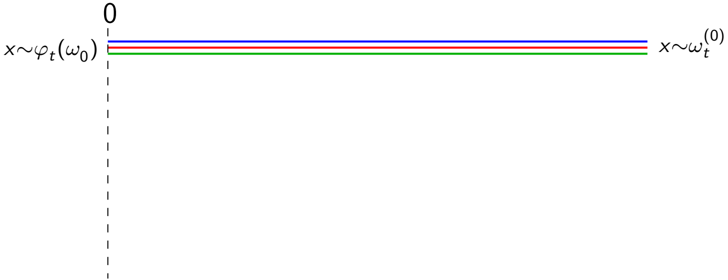
Ancestral graph

i-splitting events on every line at rate ϱ_i, $i \in S \setminus i_\bullet$ (backward!)
$S = \{1, 2, 3\}$, $i_\bullet = 1$:



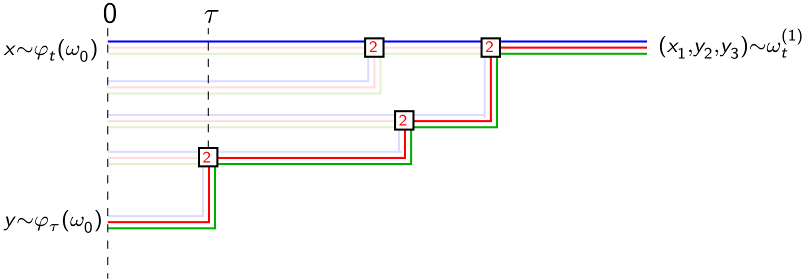
$s = 0$: partitioning of sequence across ancestors
sampling from ω_0

Ancestral graph, $s > 0$



$$\omega_t^{(0)} = \varphi_t(\omega_0)$$

Ancestral graph, $s > 0$

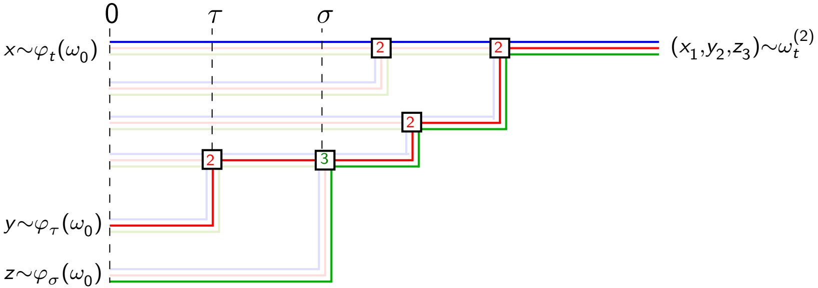


decomposition according to leftmost 2-event on ancestral line of tail (since $\Psi_{\text{sel}}(\nu_{\{i_\bullet\}} \otimes \nu_{\{S \setminus i_\bullet\}}) = \Psi_{\text{sel}}(\nu_{\{i_\bullet\}}) \otimes \nu_{\{S \setminus i_\bullet\}})$

$$\rightsquigarrow \omega_t^{(1)} = \mathrm{e}^{-\varrho_2 t} \varphi_t(\omega_0) + \varphi_{\{1\},t}(\omega_0) \otimes \int_0^t \varrho_2 \, \mathrm{e}^{-\varrho_2 \tau} \varphi_{\{2,3\},\tau}(\omega_0) \, \mathrm{d}\tau$$

$$= \mathrm{e}^{-\varrho^{(1)} t} \omega_t^{(0)} + \omega_{C^{(1)},t}^{(0)} \otimes \int_0^t \varrho^{(1)} \, \mathrm{e}^{-\varrho^{(1)} \tau} \omega_{D^{(1)},\tau}^{(0)} \, \mathrm{d}\tau$$

Ancestral graph, $s > 0$



decomposition according to leftmost k-event on ancestral line of $D^{(k)}$

$$\rightsquigarrow \omega_t^{(k)} = \mathrm{e}^{-\varrho^{(k)} t} \omega_t^{(k-1)} + \omega_{C^{(k-1)},t}^{(k-1)} \otimes \int_0^t \varrho^{(k)} \, \mathrm{e}^{-\varrho^{(k)}\sigma} \omega_{D^{(k-1)},\sigma}^{(k-1)} \mathrm{d}\sigma$$