Mutation and Selection in Age-structured Populations

"No, I don't want to live forever, but I damn sure don't want to be dead forever, either."

David Steinsaltz

Dept. of Statistics

Oxford University

Mutation and Selection in Age-structured Populations

- Introduction to the evolutionary problem
- Description of the model
- Feynman-Kac Solution
- Adding recombination
- Implications of the Solution

Joint work with Ken Wachter and Steve Evans

What are the questions to which an evolutionary theory of ageing could be (part of) the answer?

1. Why are organisms allowed to fall apart after being painstakingly built up?

2. Given that substantially longer life is possible in many organisms, sometimes from simple mutations, why are these not prevalent?

"It is indeed remarkable that after a seemingly miraculous feat of morphogenesis a complex metazoan should be unable to perform the much simpler task of merely maintaining what is already formed." - George Williams

The answers might not be the same, or even compatible.

Question 1: Why does optimality fail?

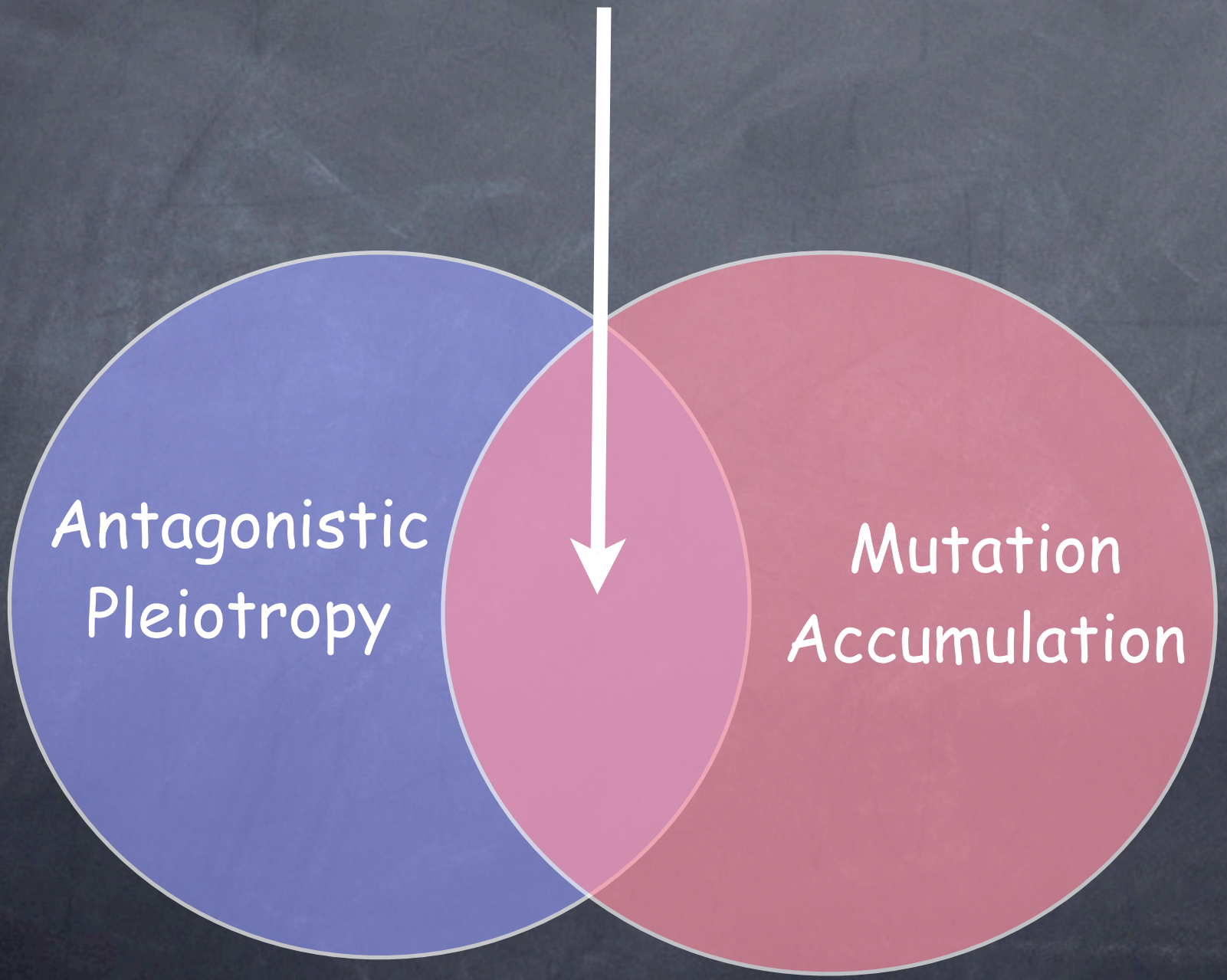
Question 2: Why is longer life (or potential immortality) not optimal?

Evolution of Aging

General idea goes back to A. Weismann (late 19th C.), P. Medawar and G. Williams (1950s):

Late-acting deleterious mutations are subject to less stringent selection control

The nature of age-dependent mutations



Mutation-Selection Equilibrium

Intuitive single-locus model: Mutant allele arises at rate ν. Selective cost s.

Evolution equation: Let p_t be the frequency of the mutant allele at time t.

$$\frac{dp_t}{dt} \approx \nu - sp_t \text{ when } p_t \text{ small.}$$

Equilibrium when frequency of mutant is about ν/s (when ν/s is small).

Mutation-Selection Equilibrium

Intuitive single-locus model: Mutant allele arises at rate ν. Selective cost s. Equilibrium when frequency of mutant is ν/s.

B. Charlesworth (2001):

- constant reproduction rate λ
- high "background mortality" μ
- mutation increases mortality by m at age x
- constant mutation rate ν

cost = $\lambda m e^{-\mu x}$ of total reproduction

Expect equilibrium frequency $\frac{\nu}{m\lambda} e^{\mu x}$.

Haldane's principle

"the loss of fitness to the species depends entirely on the mutation rate and not at all on the effect of the gene upon fitness of the individual carrying it ..." -- Haldane (1936)

In our model, overall loss of fitness in population is (ν /s)*s, so does depends only on mutation rate ν .

How do we extend this to multiple sites?

Kimura-Murayama model:

Individual with k mutations has fitness $(1-s)^k$.

Each newborn gets extra Pois(ν) mutations.

Evolution equation: Population defined at generation t as distribution on number of mutations. This is always Poisson with mean p_t, satisfying $p_{t+1}=p_t(1-s) + \nu$.

Equilibrium when frequency of mutant is ν/s.

Hamilton (1966): Study evolution of ageing by considering "mutations" that raise mortality at one age.

What is the "cost" of mortality?

Simple model: Cost=lost future reproduction. Decrease in Net Reproduction Ratio (NRR)

$$NRR(g) = \int_0^\infty f_x(g)\ell_x(g)e^{-rx}dx,$$

where $f_x(g)$=fertility at age x, $\ell_x(g)$=survivorship to age x, r=population growth rate

Wait a minute! Isn't selective cost given by Lotka's r?

Lotka's r= unique solution to $1 = \int_0^\infty \ell_x f_x e^{-rx} dx.$

Answer: No.

Why not?

Answer: The populations in individual mutation classes are never in demographic equilibrium.

Reference: Wachter, Steinsaltz, Evans "Vital rates from the action of mutation accumulation". Big source for this material B. Charlesworth, Evolution in age-structured populations

Problems:

Mathematical framework for single locus, applied to infinite-locus setting.

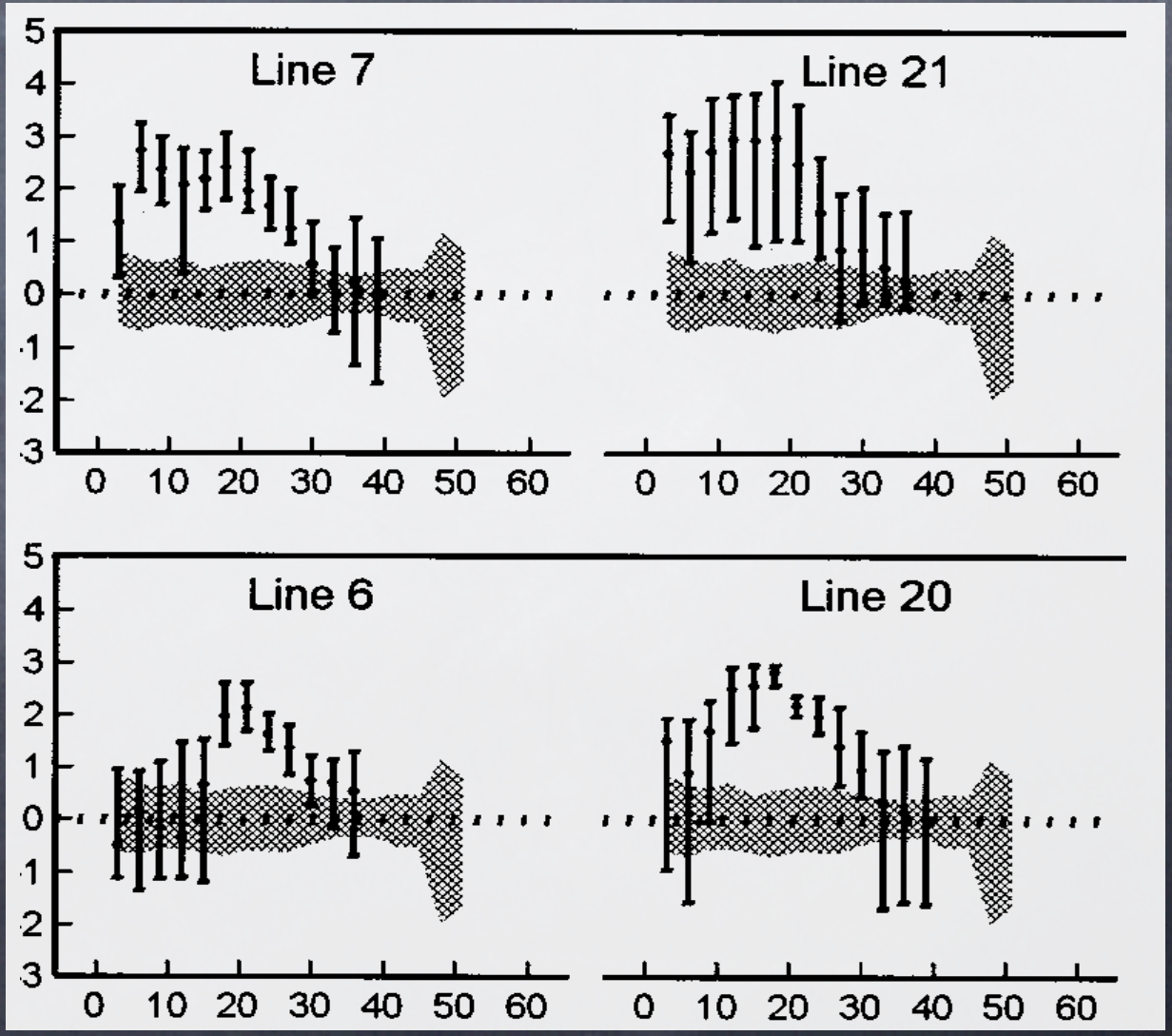
Selective cost of multiple mutations non-additive.

Problems:

Mathematical framework for single locus, applied to infinite-locus setting.

Selective cost of multiple mutations non-additive.

Mutations which act only at one age are extremely unrealistic.



Linear version of Hamilton's formula:

$$h(a) \approx \frac{\nu(a)}{w(a)} = \frac{\nu(a)}{\int_a^\infty \exp\left(-\int_0^x \lambda(y)dy\right)\ f_x\, dx}$$

Nonlinear version of Hamilton's formula:

$$h(a) = \frac{\nu(a)}{\int_a^\infty \exp\left(-\int_0^x \left(\lambda(y) + h(y)\right) dy\right)\ f_x\, dx}$$

Improved model (S. Evans, K. Wachter, and DS):

Mutation space M

Mutation rate ν = σ-finite measure on M

Genotype space G = {integer measures on M}

State of system P = probability on G

Selection cost $S : G \to \mathbb{R}^+$

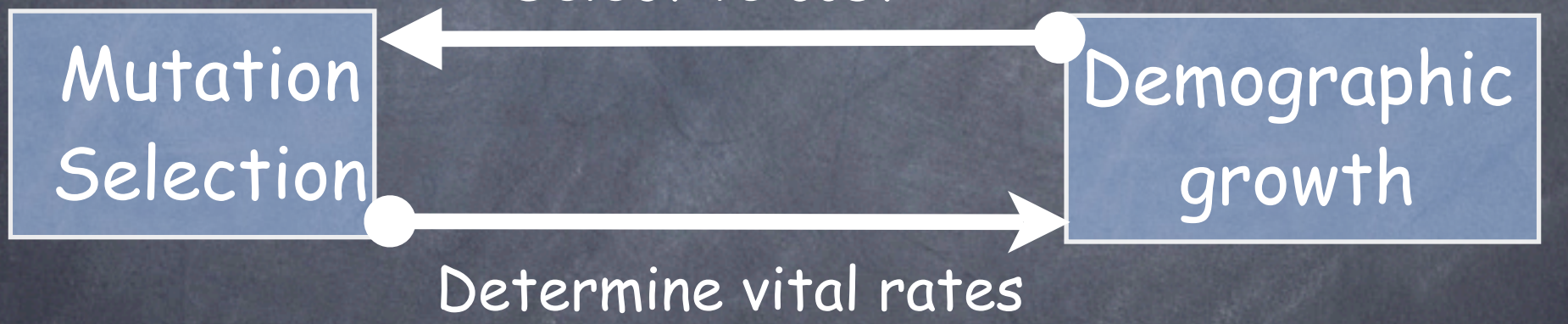
State of the population at time t is described by a probability distribution P_t on genotypes.

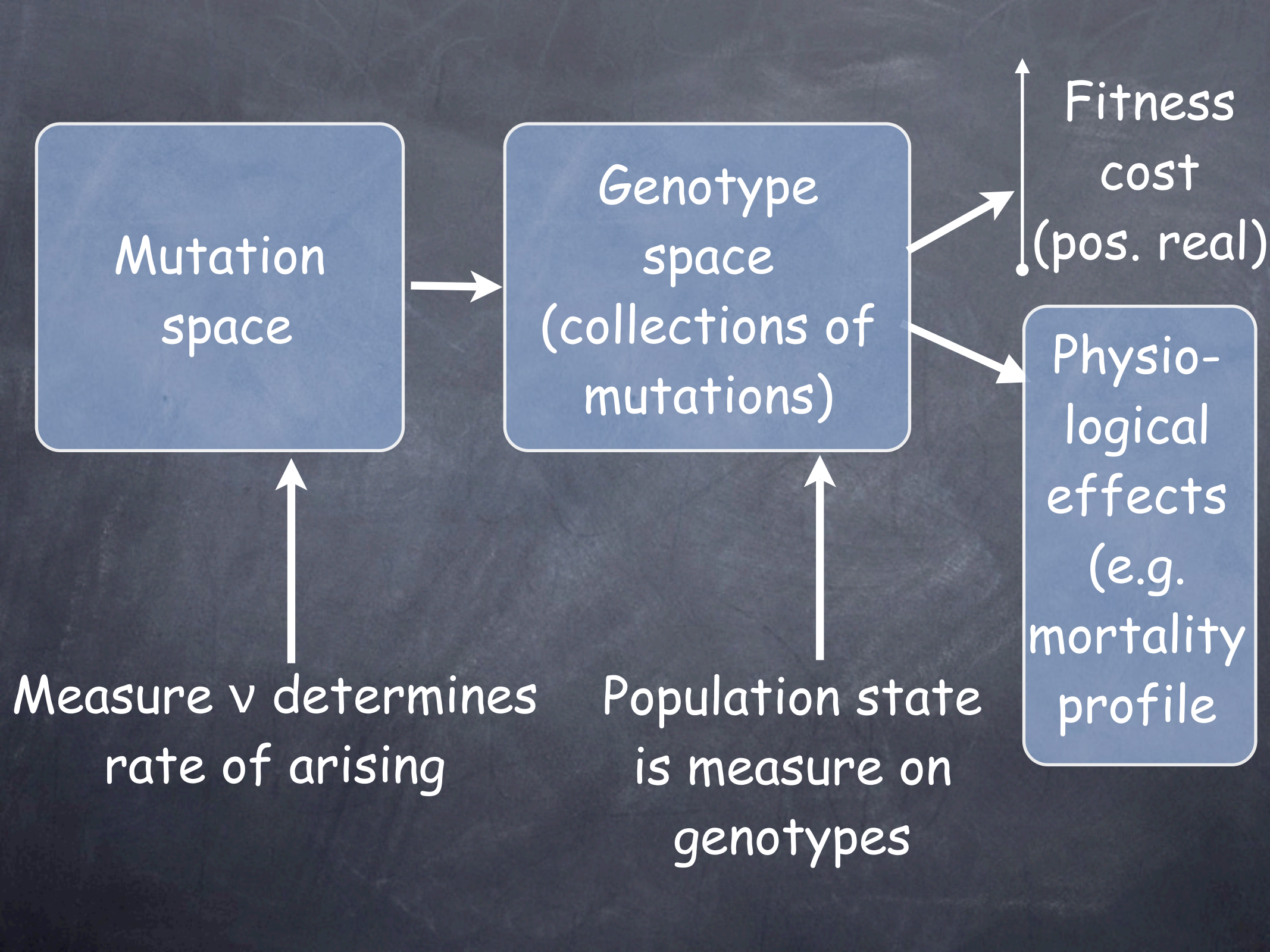
$$\frac{d}{dt} P_t F = P_t \left(\int \left[F(\cdot + \delta_m) - F(\cdot) \right] d\nu(m) \right)$$

$$- P_t(FS) + (P_t F)(P_t S)$$

Coupling Time scales

Currently linear (fitness purely function of individual genotype) but could substitute a nonlinear process here.




Mutation space
Genotype space (collections of mutations)
Fitness cost (pos. real)
Physio-logical effects (e.g. mortality profile
Measure ν determines rate of arising
Population state is measure on genotypes

Quantitative genetics: Compare and contrast

- "Genotype" in QG is mortality function
- Heritable "mutations" are point changes
- Different ages connected only by covariance of transmission (only two-point link)
- Is this the only difference? We're working to explore the connections

Solution: Define an operator A by

$$AF = \int [F(\cdot + \delta_m) - F(\cdot)]\, d\nu(m) - S(\cdot)F(\cdot).$$

A is the generator of a sub-Markovian semigroup Γ_t. By Feynman-Kac,

$$\Gamma_t F(g) = \mathbb{E}\left[\exp\left(-\int_0^t S(g + X_u - X_0)\, du\right) F(g + X_t - X_0)\right]$$

What is the solution? Let X_t be a Poisson point process with intensity ν. Then

$$P_t F = \frac{\mathbb{E}\left[\exp\left(-\int_0^t S(X_u)\,du\right) F(X_t)\right]}{\mathbb{E}\left[\exp\left(-\int_0^t S(X_u)\,du\right)\right]}.$$

When S is linear ("non-epistatic"), the solution reduces to a Poisson random measure with intensity

$$\frac{1-e^{-S(m)t}}{S(m)} d\nu(m).$$

In this case, the solution is unique.

What does this tell us?

1. Series expansion for P_t and limiting distribution: Let $Y_1, Y_2, \ldots, Y_n$ be an increasing random choice of n mutations (from distribution ν.) Then

$$\lim_{t\to\infty} P_t F = \frac{\sum_{n=0}^{\infty} \nu(\mathcal{M})^n \mathbb{E}\left[\left(S(Y_1)\ldots S(Y_n)\right)^{-1} F(Y_n)\right]}{\sum_{n=0}^{\infty} \nu(\mathcal{M})^n \mathbb{E}\left[\left(S(Y_1)\ldots S(Y_n)\right)^{-1}\right]}.$$

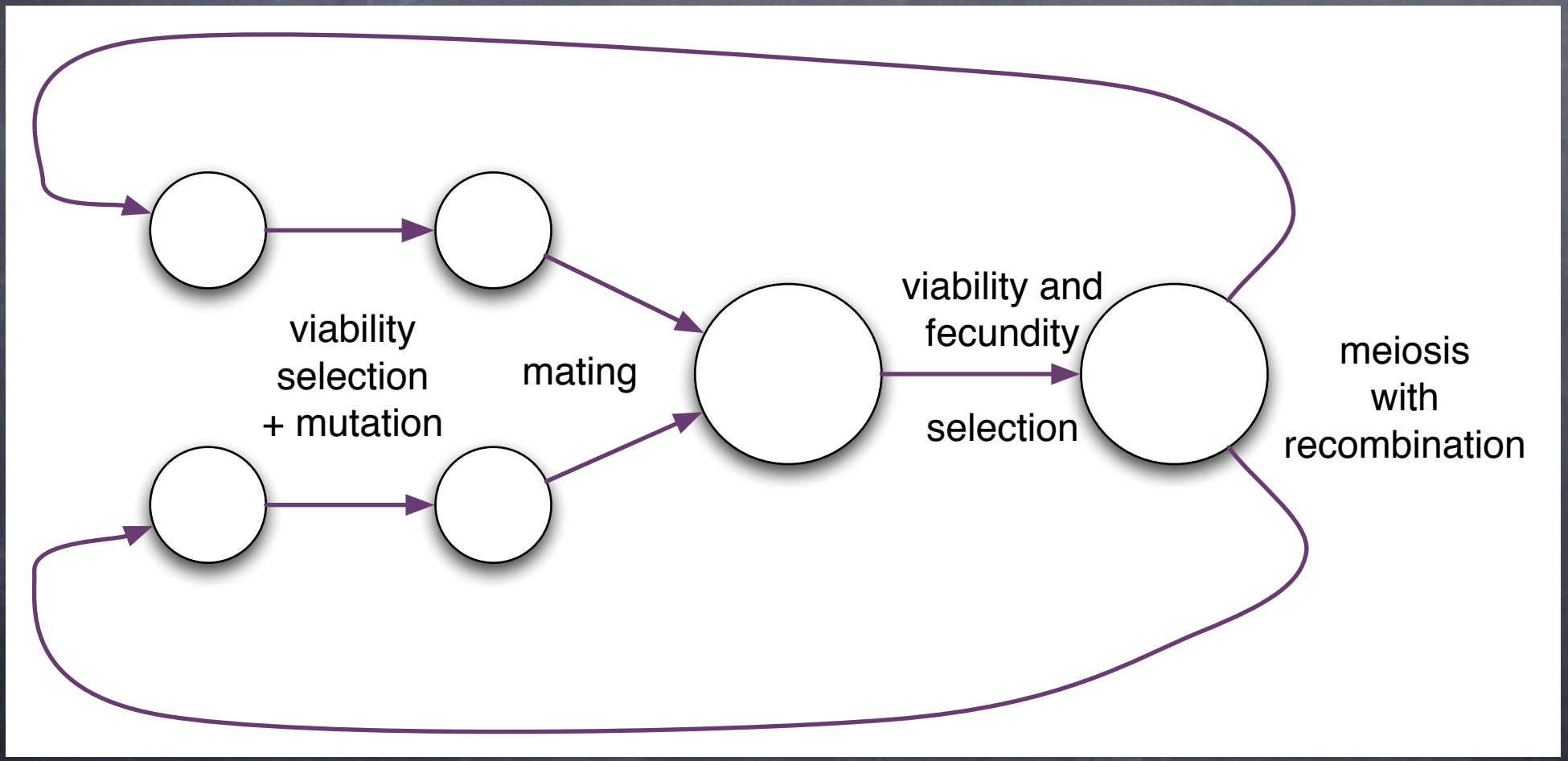
There is a corresponding finite-time formula.

What does this tell us?

2. Explosion: If B is a set s.t. $S(g+b)-S(g)<v(B)$ when $b \in B$, then the number of mutations in B goes to infinity.

Implies "wall of death" rather than Gompertz.

Recombination



Barton-Turelli model

Recombination: Pick a random subset of mutations A from a distribution r. New genotype gets A mutations from one parent, and A^c mutations from the other.

If we iterate this process, the genotypes get completely reshuffled.

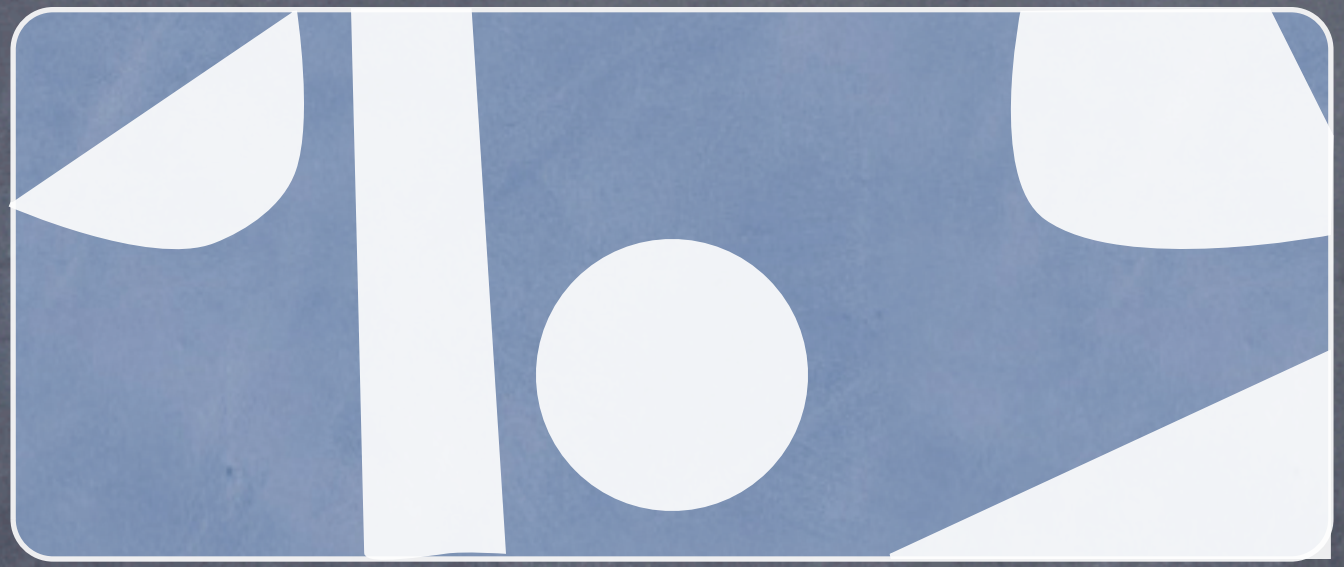
End up with a Poisson random measure, with the same marginal intensities as the genotype distribution we start with.

Mutation space

Parent 1

Parent 2

Recombination set



Parent 1

Parent 2

Offspring genotype

Definition: The recombination measure R is the distribution on subsets of M, defining which sites come from the same parent. R is shattering if there is a positive constant α such that $\mathrm{E}\big[\nu(A \cap R)^2\big] \leq \frac{1}{2}\nu(A)^2 - \alpha\nu(A)^3$.

Intuitively, shattering means that points get separated. For example, if mutation space is a line segment, and R is chosen by splitting at a single point it's shattering if

$$P\big\{N([a,b]) = 1\big\} \geq C\nu([a,b]).$$

Definition: The recombination measure R is the distribution on subsets of M, defining which sites come from the same parent. R is shattering if there is a positive constant α such that $\mathrm{E}\left[\nu(A \cap R)^2\right] \leq \frac{1}{2}\nu(A)^2 - \alpha\nu(A)^3$.

Definition: A distribution P on genotypes is dispersive if there is a constant β such that for any Borel set A,

$$\int g(A)\mathbf{1}_{\{g(A)\geq 2\}} dP(g) \leq \beta\mu P(A)^2.$$

Easy part: Repeated recombination without mutation or selection (or linear selection) converges to Poisson distribution.

Theorem: If P is dispersive and R is shattering, then

$$\|\mathfrak{R}^k P - \mathfrak{P} P\|_{Was} \leq (3\beta + 2)\left(|\nu|^2 \vee 2\alpha|\nu|\right)(k+1)^{-1}.$$

This is a process version of Le Cam's Theorem on convergence to Poisson distribution.

This justifies defining a dynamical system concentrated only on Poisson random measures.

Since mutation and selection are much slower, this converges to a process that is always Poisson.

Let $s_P(m)$ be the average cost of mutations m, averaged over the genotype dist. P.

If P is Poisson, then

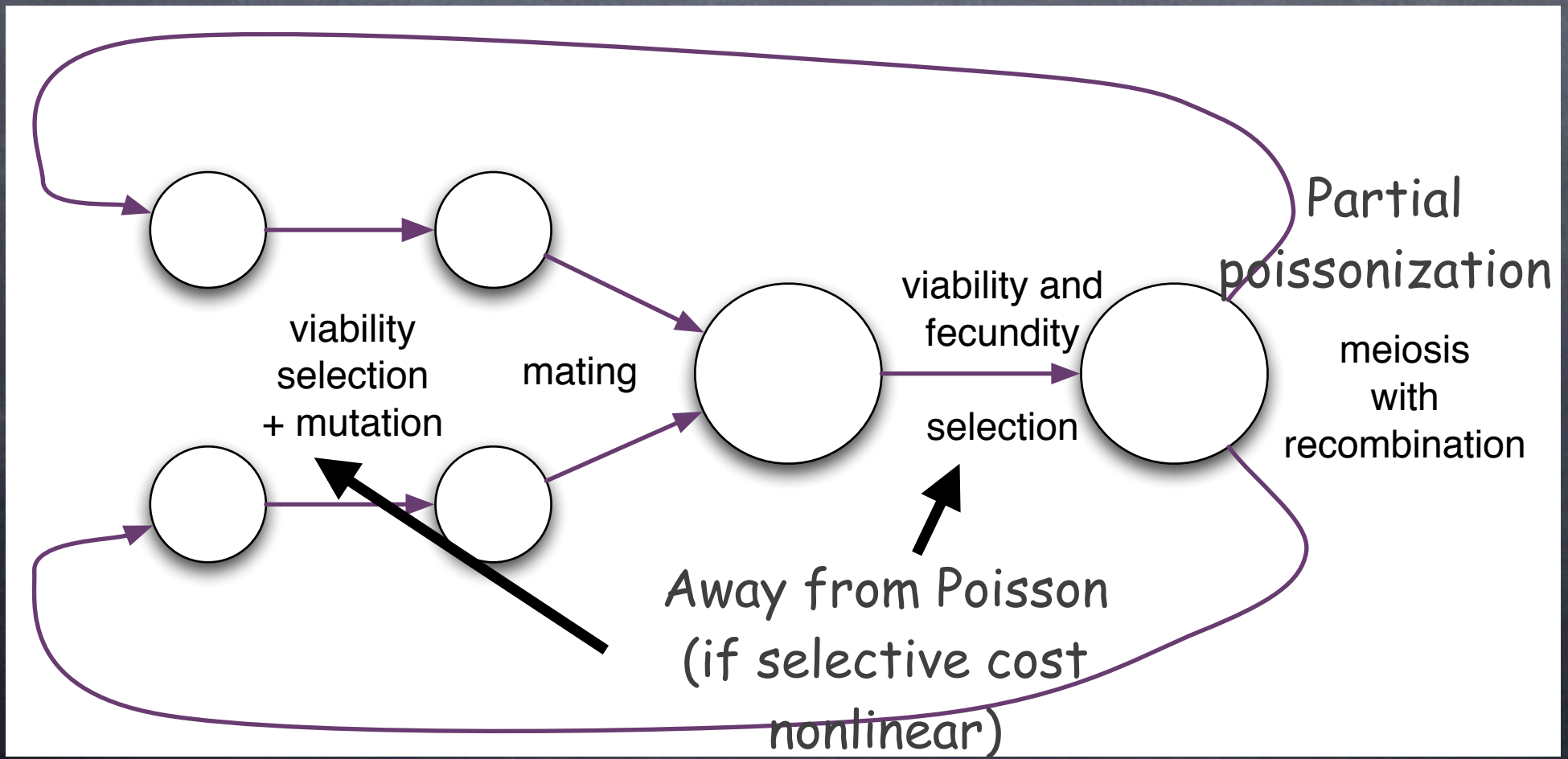
$$s_P(m) := \int \left[S(g + \delta_m) - S(g)\right] dP(g).$$

Let ρ_t be the Poisson intensity at time t.

$$\frac{d\rho_t}{dt} = \nu - s_{P_t}(m)\rho_t.$$

Theorem (Evans, DS, Wachter): If ν is finite, this equation has a unique solution, which remains finite for all t.

Does this poissonization really work?



Main result

Let Q_k be the distribution after k rounds of mutation, selection, and recombination.

If the initial P_0 is Poisson and R is shattering, then for any positive T,

$$\lim_{n\to\infty} \sup_{t\le T} \left\| \Pi_{\rho_t} - Q_{\lfloor tn \rfloor} \right\|_{Was} = 0.$$

Π_ρ is the Poisson measure with intensity ρ.

Example: Gamma profiles

$$\kappa(m,x) = \int_{\alpha}^{x} \frac{1}{\Gamma(m)} \phi^{m} (y-\alpha)^{m-1} e^{-\phi(y-\alpha)} dy.$$

mean age of action $\alpha + m/\phi$

standard deviation $\sqrt{m}/\phi$.

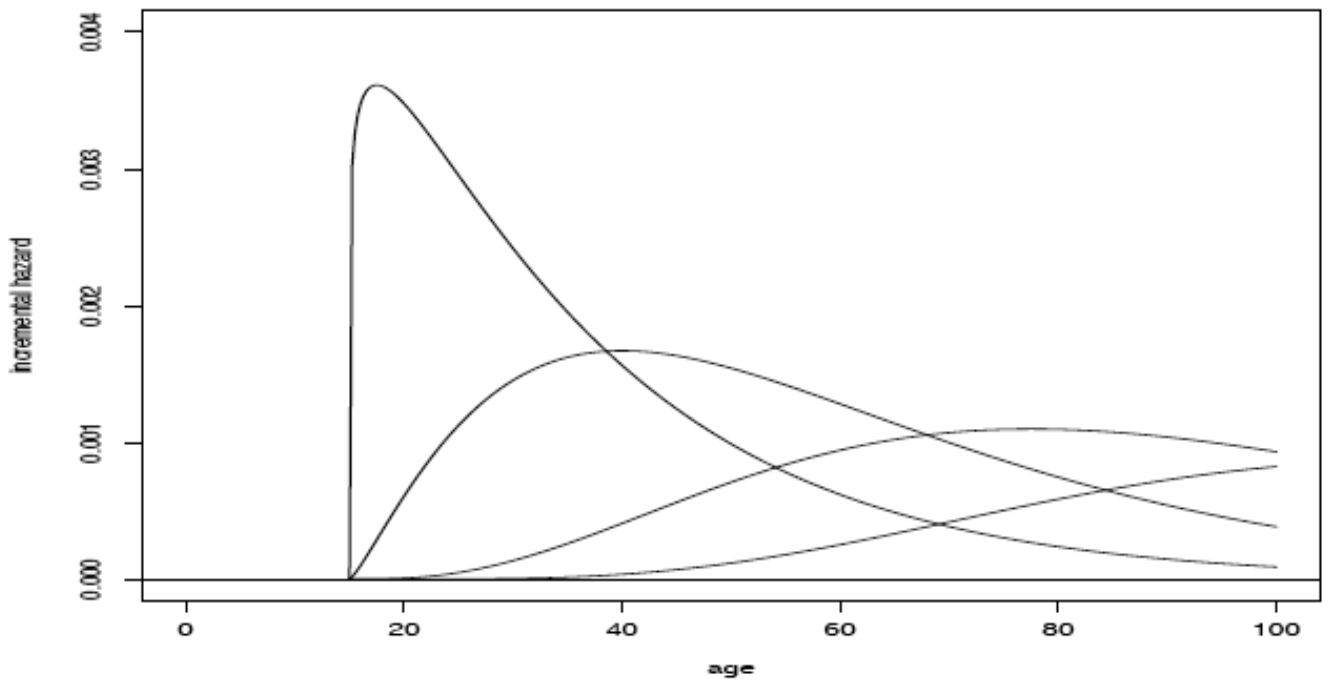


Figure 1: Gamma profiles for increments to the hazard function for four selected values of the mutation index m.

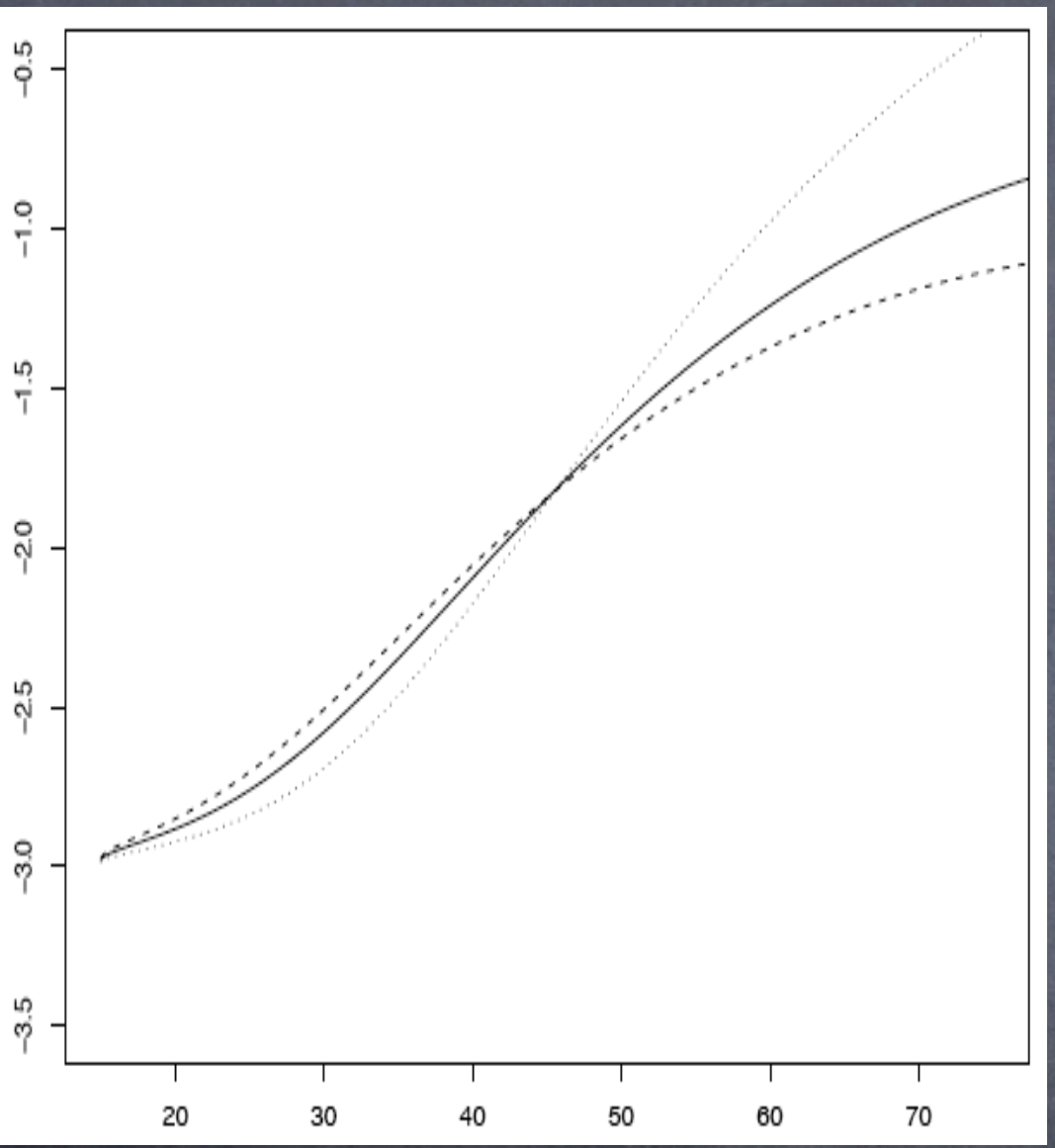
Take background mortality constant $\lambda = .05$

Gamma parameter $\phi = .05$

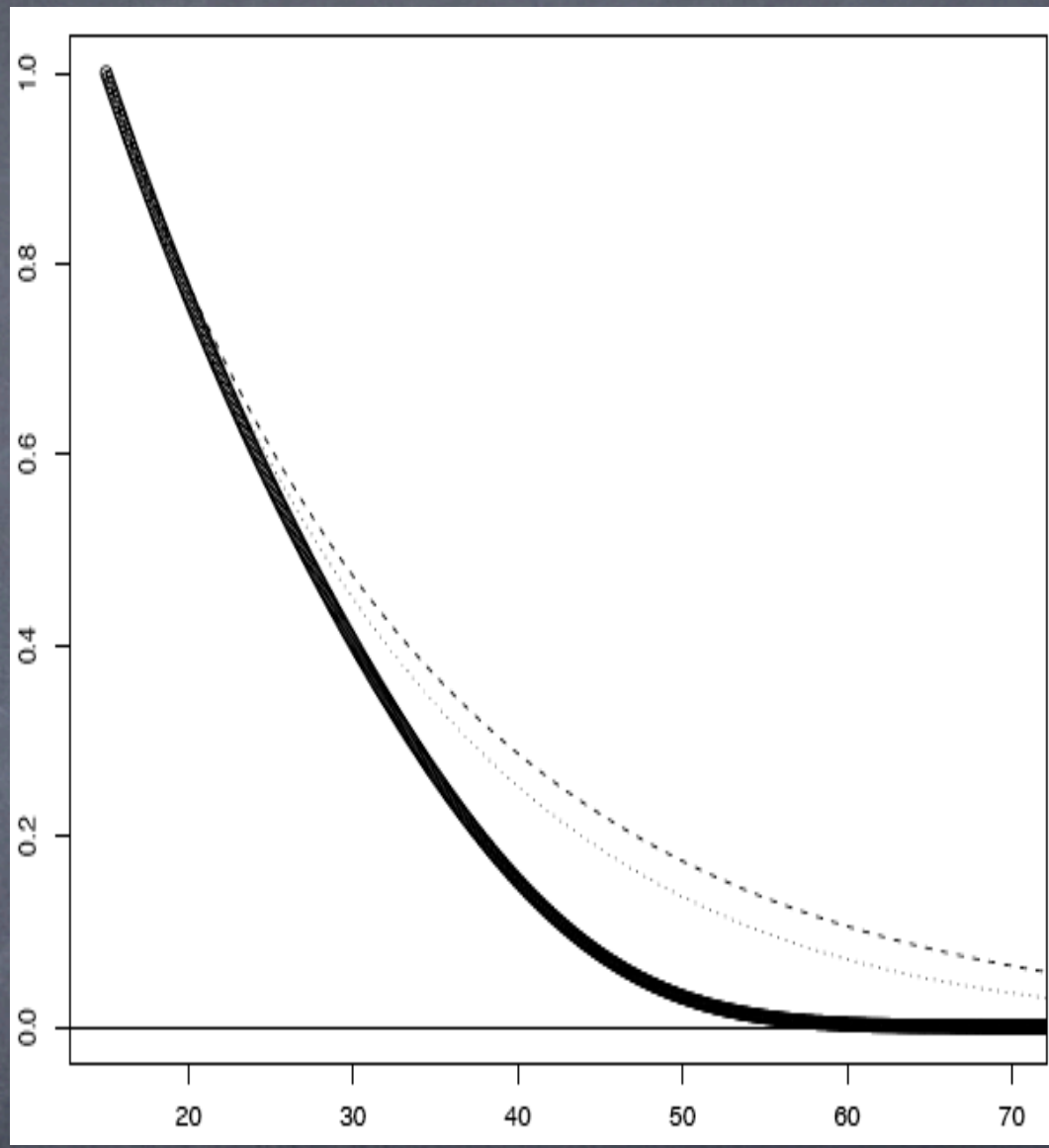
total mutation rate $\nu = .12, .15, .17$

Mutations uniform on $[0, \xi\,]$ where $\xi = 7, 6, 5.5$

Age at first reproduction $\alpha = 15$



equilibrium
log hazard



equilibrium
survival rate

Example: Polynomial Fitness cost

"Mutation-selection balance with recombination: Convergence to equilibrium for polynomial selection costs." A Clayton, S Evans. SIAM J Applied Math.

$$S(g) = \sum_I a_I g^I$$

Sum over ordered subsets of mutations.

Theorem: Unique equilibrium if minimum cost of a mutation bounded away from 0. This equilibrium is globally stable.

General Implication 1: Unraveling

- Hamilton/Charlesworth setting: Point-mass mutation effects on mortality, mutation rate constant across all ages, constant fertility
- Increasing mortality erases selective pressure against mortality at ever younger ages.
- No equilibrium. In the limit, survivorship at all ages goes to 0.
- Different from no-recombination case.

Implication 2: Equilibrium condition

$$\nu(a) = \rho(a) \int_{\alpha}^{\infty} (1 - e^{-\eta(a)\kappa(a,x)}) \, \mathbb{E}_{\rho} [\, f_x l_x(G) \,] \, dx$$

Mutations labelled by "age of effect" a, producing increment to mortality η (a) κ (a,x) at age x.

Classical: κ =step function at age a. Define

"remaining life expectancy"

$$T(a) := \int_{a}^{\infty} \exp\left(- \int_{\alpha}^{x} [\lambda + h(y)] dy \right) \, dx$$

Then at equilibrium, $\nu(a) = h(a) \, f \, T(a)$

Implication 3: Haldane's Principle generalised

Note that the equilibrium formula $\nu(a) = h(a)\, f\, T(a)$ doesn't depend on η.

In general, at equilibrium

$$\nu(da) = \rho(da) \int_{\alpha}^{\infty} (1 - e^{-\eta(a)\kappa(a,x)})\, \mathbb{E}_{\rho}\left[\, f_x l_x(G)\,\right] dx$$

Integrate both sides. If F is any linear function of genotypes $F(g) = \phi \cdot g$, we have the Fourier transform

$$\log \mathbb{E}_{\rho} e^{-F(G)} = \int_{M} \left[1 - e^{-\phi(a)}\right] d\rho(a).$$

Implication 3: Haldane's Principle generalised

$$\nu(M) = \int_0^\infty \left(-\log \mathbb{E}_\rho[\ell_x(G)/\ell_x(0)]\right) f_x \mathbb{E}_\rho\left[\ell_x(G)\right] dx$$

This is <u>not</u> the total loss of fitness in the population. Changing the size of mutations leaves the "aggregate" survivorship unchanged, not the expected survivorship.

Total loss of fitness is not invariant under changes in the age-pattern of effects.

Conclusions

- Evolutionary theory of ageing requires recognition of nonlinear fitness effects
- Recombination produces qualitative changes in behaviour

References

(all Evans, DS, Wachter unless otherwise noted)

- ``A generalized model of mutation-selection balance with applications to aging'' Advances in Applied Mathematics, 2004.
- "A mutation-selection model for general genotypes with recombination" (The basic results for the recom model. A version is posted on the ArXiv, but a substantially improved version is nearing completion.)
- "Vital rates from the action of mutation accumulation" (Demographic implications, mainly, with some computations and plots for the gamma case.)
- "The age-specific force of natural selection and walls of death"
- "Mutation-selection balance with recombination: Convergence to equilibrium for polynomial selection costs." A Clayton, S Evans.