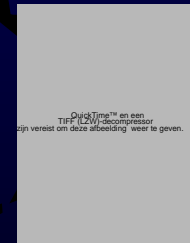




# Effective population sizes

and the canonical equation  
of adaptive dynamics



Hans Metz

Vincent Jansen

Ulf Dieckmann

ECP

IIASA

# the ecological theatre and the evolutionary play

---

Long term adaptive evolution proceeds through the continual filtering of new mutations by selection.

The supply of new phenotypes by mutation depends on the genotypic architecture and the genotype to phenotype map.

Selection is a population dynamical process.

Convenient idealisation:  
the time scales of the production of novel variation  
and of gene substitutions are separated.

## the adaptive play

---

Given a population dynamics one can graft onto it an adaptive dynamics:

Just assume that individuals are characterised by traits that may change through mutation, and that affect their demographic parameters.

The speed of adaptive evolution is proportional to the population size  $n$ .

## a second, silent, play: random genetic drift

---

Given a population dynamical model one can also graft onto it a “random genetic drift”.

Just imagine that each individual harbours two alleles which may be, say, either black or white, and which in the reproductive process are reassorted according to Mendel’s laws.

The speed at which variation is lost over time is inversely proportional to  $n$ .

# a formal connection between the two plays

---

## Result:

The constants that appear in front of  $n$  in the formulas for the speeds of adaptive evolution and random genetic drift are the same.

# Part I

## The ecological theatre

# i-states

---

Sensible population models are individual-based!

Let the physiological state of an individual (i-state) be denoted as  $\underline{S}$ , where  $\underline{S}$  takes values in  $S$ .

By the definition of the state concept,  $\underline{S}(t)$  and the intervening stretch of environmental history,  $E[t, s)$ ,  $s > t$ , together fully determine the future state  $\underline{S}(s)$ ,

i.e., there exists a family of sub-Markov transition operators  $u_E$  such that

$$P_E \left( \underline{S}(s) \in A \mid \underline{S}(t) = S \right) = u_E \left( A ; s, t, S \right)$$

# Example: deterministic i-state transitions

Individuals are born equal in state  $S_b$ , move deterministically

through  $S$ , according to  $\frac{dS}{dt}(t) = G(S(t), E(t))$

and

die at a rate  $d(S(t), E(t))$ .

Then  $u_E$  satisfies the backward equation  $\frac{d}{dt} \bar{f} = L_i \bar{f}$  with

$$\bar{f}(s, t)(S) := \mathbf{E}[f(\underline{S}(s) \mid \underline{S}(t) = S)] \quad f \in BC^1(S)$$

the boundedly differentiable functions on  $S$ , and

$$L_i \bar{f}(s, t)(S) = -G(S, E(t)) \frac{\partial}{\partial S} \bar{f}(s, t)(S) + \bar{f}(s, t)(S) d(S, E(t))$$

# the “general” case

---

In other models in addition to the deterministic movement through  $S$  also diffusive and jump movements may occur,

the former as result of many very small local disturbances,

the latter, for example, as the result of swallowing prey in models that explicitly take satiation into account (in addition to energy reserves and body size).

In addition individuals may be born in more than one birth state, with a distribution depending on  $S$  and  $E$ .

As an example you may think of cells that grow and distribute their mass  $S$  roughly equally over two daughter cells.

# births may present a conceptual problem

---

The most common assumption in deterministic population models is that individuals give birth at a rate  $\beta(S(t), E(t))$ .

At the individual level this corresponds to the assumption that births occur in a Poisson process, or in clutches released according to a Poisson process.

This assumption goes against mass balance principles: to produce a (clutch of) young individuals have to accumulate the necessary energy and mass.

Moreover, usually the next brood can be started only after the womb / brood pouch / egg chamber has been emptied.

⇒ On a microscopic scale the occurrence of reproductive events is (often far) more regular than Poisson.

Accommodating this regularity requires extra state variables.

In the case of deterministic i-dynamics the resulting models tend to become inherently degenerate. *How should this be handled?*

# the environment (well mixed case)

Individuals impinge on the environment as  $\Omega^{-1} \gamma(\underline{S}, \underline{E})$ ,  
 $\Omega$  the so-called system size.

The effects of different individuals add up.

Hence the total impingement on the environment can be written as

$$\underline{I}(t) = \Omega^{-1} \sum_{i=1}^{n(t)} \gamma(\underline{S}_i(t), \underline{E}(t))$$

## Some possible special assumptions

1. The environment moves like

$$\frac{d\underline{E}}{dt}(t) = H(\underline{E}(t), \underline{I}(t))$$

(to be added to the example).

2. The environment is some instantaneous function of  $I$ .

# the p-state (well mixed case)

---

If the environment is everywhere the same the p(population)-state can be given in the form of a list  $\underline{S} = (\underline{S}_1, \dots, \underline{S}_n)$ , taking values in

$$\bigcup_{n=0}^{\infty} \mathcal{S}^n \quad \text{with} \quad \mathcal{S}^0 = 0, \quad \mathcal{S}^1 = \mathcal{S}, \quad \mathcal{S}^n = \mathcal{S} \times \mathcal{S}^{n-1}$$

As the ordering in a list is arbitrary, this state representation is too large, and p-states can in principle be reduced by taking equivalence classes under permutation of the indices.

(It is not possible to use sets instead of lists, as there may exist times at which there is more than one individual with exactly the same i-state.)

# simulations

Example, cont'd:

Simulations can be done by integrating the sum of all birth and death rates

$$\frac{d\underline{E}}{dt}(t) = \underline{\xi}(t) \quad \underline{\xi}(t) = \sum_{i=1}^{n(t)} \left[ \beta(\underline{S}_i(t), \underline{E}(t)) + d(\underline{S}_i(t), \underline{E}(t)) \right]$$

together with the differential equations for the  $\underline{S}_i$  and  $\underline{E}$ .

When  $\underline{E}(t)$  encounters a tick of a rate 1 Poisson clock on the  $\underline{E}$ -axis, which happens at times, say,  $\underline{\tau}_k$ , an event happens. The probability that this is the death of the  $i^{\text{th}}$  individual is

and that it is a birth from the  $i^{\text{th}}$  mother is

# a different p-state representation

---

Let  $M_F(\mathcal{S})$  be the space of finite measures on  $\mathcal{S}$  and let  $\delta_S$  denote a unit point mass at  $S$ .

Then we can also represent the population state as

$$\mu = \sum_{i=1}^n \delta_{S_i} \in M_F(\mathcal{S})$$

Under special assumption 1 the state of the community at any particular time can be written as

$$\left( \underline{\mu}(t), \underline{E}(t) \right) = \left( \sum_{i=1}^{\underline{n}(t)} \delta_{\underline{S}_i(t)}, \underline{E}(t) \right)$$

# backward equations

---

Just as in the case of the individual level transition operator  $u_E$  we can then write a differential generator by telling what happens to the expectations of functions on  $M_F(S) \times E$ . The set of considered functions should be sufficiently rich that by combining them in appropriate manners one can generate all bounded measurable functions on  $M_F(S) \times E$ .

One such set are the functions

$$F_f(\mu, E) := F(\langle \mu, f \rangle, E), \quad F \in BC^1(\mathbb{R}^2) \quad f \in BC^1(S)$$

with

$$\langle \mu, f \rangle := \int_S f(S) \mu\{dS\}$$

# backward equations, cont'd

Example, cont'd:

Let again  $\bar{F}_f(t)(\mu, E) = \mathbf{E} \left[ F_f(\underline{\mu}(t), \underline{E}(t)) \mid (\underline{\mu}(0), \underline{E}(0)) = (\mu, E) \right]$  then

$\frac{d}{dt} \bar{F}_f = L_p \bar{F}_f$  with

$$L_p \bar{F}_f(\mu, E) =$$

$$\left[ \partial_1 F \right]_f(\mu, E) \left\langle \mu, \left\langle \partial f, G(\cdot, E) \right\rangle \right\rangle \quad \text{i-movement}$$

$$+ \left\langle \mu, d(\cdot, E) \left( F(\langle \mu, f \rangle - f(\cdot), E) - F_f(\mu, E) \right) \right\rangle \quad \text{deaths}$$

$$+ \left\langle \mu, \beta(\cdot, E) \left( F(\langle \mu, f \rangle + f(X), E) - F_f(\mu, E) \right) \right\rangle \quad \text{births}$$

$$- \left\langle \left[ \partial_2 F \right]_f(\mu, E), H \left( \Omega^{-1} \langle \mu, \gamma(\cdot, E) \rangle, E \right) \right\rangle \quad \text{environment}$$

A backward operator also defines the population process.

# large systems size limits

---

## Conjecture:

When  $\Omega \rightarrow \infty$ ,  $\mu/\Omega$  converges weakly to a deterministic flow in  $M_F(S) \times E$ .

Sylvie Méléard and various coworkers have already proved this for various special cases.



Nicolas Champagnat



Sylvie Méléard



Viet Chi Tran

# large systems size limits, cont'd

In general the non-linear infinite dimensional semigroup associated with the flow does not have a classical differential generator.

It is therefore necessary to construct the semi-group directly from the individual level ingredients  $u_E, \gamma, H$ , plus  $\Lambda(a)$ , the mean number of offspring produced up to age  $a$ .



Odo Diekmann



Mats Gyllenberg



Huang Haiyang



Marcus Kirkilionis

me



Horst Thieme

# direct interactions

---

The interactions considered so far were indirect, through the intervening variables  $I$  and  $E$  (think of competition for non-particulate renewable resource, or available space).

Other interactions, like predation or mating, directly involve more than one individual, leading to a dependence between the those individuals that has to be accounted for in a different manner.

## Conjecture:

In the limit for infinite system size also such dependencies can be captured by global variables (predation pressure, mate availability), which then can be absorbed in  $I$  and  $E$ .

## direct interactions, cont'd

---

For the microscopic description of direct interactions problems similar to those encountered for birth rates raise their head:

In a microscopic model the capturing and eating of a prey corresponds to a state jump of the predator.

In the final deterministic model usually all that fine structure is replaced by a growth rate dependent on prey availability.

A microscopic mating model may be formulated in terms of meta-individuals comprising ordinary individuals and married pairs.

In the final deterministic model marriage status may become an additional state variable of individuals.

*A general theory for dealing with such issues is still missing.*

# extensions: 1. space

---

The assumption of well-mixing can be removed by characterising individuals with their h(eterogeneity)-state, consisting of their i-state and their spatial location.

$E$  then also becomes a function of space, with individuals reacting only to its local value.

Similarly,  $I$  becomes a measure over space calculated as

$$\underline{I}(t)(A) = \Omega^{-1} \int_{S \times A} \gamma((S, Z), \underline{E}(t, Z)) \underline{\mu}(\{dS\} \times \{dZ\})$$

## extensions: 2. traits

---

To capture evolutionary change individuals should be characterised by hereditary traits  $X \in X \subset R^k$ , parametrising their i-behaviour, and changing only over a birth event, and then only rarely, and usually only in small steps.

This trait vector can also be added as a component to the h-state.

### Aside about Mendelian inheritance:

Mendelian inheritance can be handled by considering alleles as individuals and ordinary biological individuals as meta-individuals, uneasy aggregates of alleles at all loci.

Phenotypes are generated by a map from genotypical compositions to the trait vectors of meta-individuals.

# with traits different limits become possible

classical large  
number limit



$$\Omega \rightarrow \infty$$

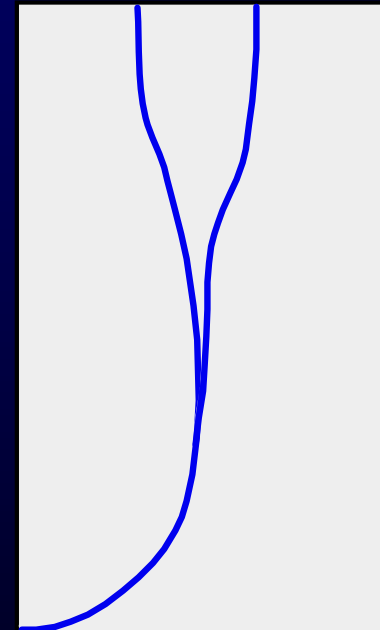
rescale numbers to densities

individual-based  
stochastic process



$$x \rightarrow$$

adaptive dynamics  
limit



$$\Omega \rightarrow \infty, \Omega\varepsilon \rightarrow 0$$

rescale time, only consider traits

$\Omega$  = system size,  $\varepsilon$  = mutations / birth

# with traits different limits become possible, cont'd

The classical limit is similar to all large system size limits.

The adaptive dynamics limit was first considered for ODE models by Diekmann & Law (1996) and for structured models by Metz, Geritz et al. (1996)



Ulf Diekmann



Richard Law

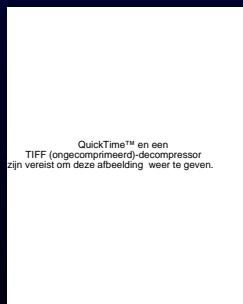


Stefan Geritz

The first hard proofs [for various special cases] are in papers by Champagnat, Méléard, Tran and Ferrière.



Regis Ferrière



Nicolas Champagnat

Sylvie  
Méléard



Viet Chi Tran

# nature of the adaptive dynamics limit

---

Due to (1) the ensuing time scale separation between the community and evolutionary dynamics  
(2) the large population sizes  
the community resides almost all the time at a deterministic attractor for the current trait value.

New mutants have a positive probability to invade if and only if their invasion fitness is positive.

If mutational steps are not too large,  
away from evolutionary singularities (like branching points and Evolutionarily Stationary Strategies)  
mutants oust former residents,  
while inheriting their attractor branch.

(OK for stable equilibria and limit cycles and some chaotic attractors. Unsure for general chaotic attractors.)

# an alternative representation

---

Since

1. most deterministic population processes do not have a differential generator,
2. usually the number of h-states at birth (b-states) is far far smaller than the number of h-states in general,

it is often much easier to work in terms of an

integral equation for the birth rate vector  $B$

(c.f. Lotka's equation from demography, or Feller's renewal equation).

The corresponding community state space is

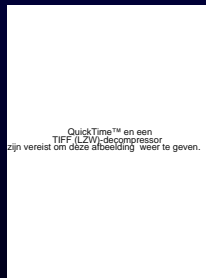
$\{\text{potential histories of } B\} \times \{\text{potential histories of } E\}$

# an alternative representation, cont'd

---

An additional advantage is that in this space it becomes possible to prove a principle of linearised stability, a Hopf bifurcation theorem etc..

This is in general not possible in  $M_{\mathbb{F}}(S) \times E$  since in this space in a sense the functions  $F_f$  encountered earlier are the only Fréchet differentiable ones.



Odo Diekmann



Philipp Getto



Mats Gyllenberg

# complicated ecological theatres

---

Doing the bookkeeping from births to births  
is also by far the easiest way of  
accommodating all sort of life history detail  
in a single overarching formalism

By allowing multiple birth states it is possible to accommodate i.a.

- parentally mediated differences in offspring size, social status, etc.
- genetic sex determination, including haplo-diploid genetics, etc
- genetic polymorphisms
- spatially distributed populations (use location as birth state component)
- cyclic environments (use phase of cycle as birth state component)

Assumption (to avoid functional analytical complications):

the number of b-states is finite.

# community equilibria

---

Equilibria of deterministic structured populations generally satisfy:

$$B = L(X|E_X)B, \quad E_X = F(I), \quad I = G(X|E_X)B$$

Diekmann, Gyllenberg & Metz (2003) TPB 63: 309-338

trait vector  
(affects the two operators that  
describe individual behaviour)

# community equilibria

Equilibria of deterministic structured populations generally satisfy:

$$\boxed{B} = L(X|E_X)B, \quad \boxed{E_X} = F(I), \quad I = G(X|E_X)B$$

equilibrium  
Birth rate vector  
per unit of area

environmental input,  
i.e., the Environment as  
perceived by the individuals

per capita lifetime  
impingement on  
the environment

next generation operator  
(i.e.,  $L_{ij}$  is the Lifetime number of births  
in state  $i$  expected from a newborn in state  $j$ )

population output,  
i.e., Impingement  
on the environment

# community equilibria

---

Equilibria of deterministic structured populations generally satisfy:

$$\boxed{B} = \underbrace{\boxed{L(X|E_X)}}_{\text{next generation operator}} B, \quad \boxed{E_X} = F(I), \quad I = G(X|E_X)B$$

equilibrium  
Birth rate vector  
per unit of area

environmental input,  
i.e., the Environment as  
perceived by the individuals

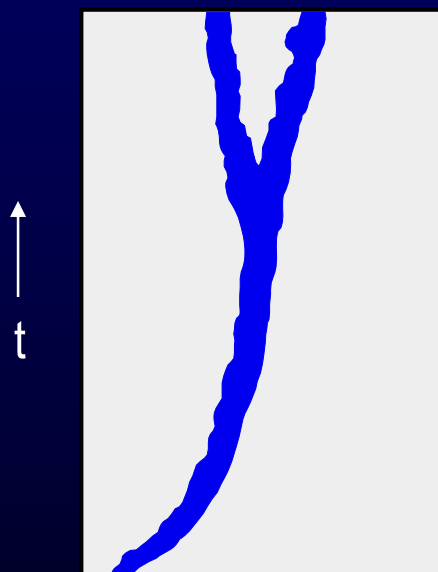
next generation operator  
(i.e.,  $L_{ij}$  is the Lifetime number of births  
in state  $i$  expected from a newborn in state  $j$ )

# Part II

## The adaptive play

# two subsequent limits

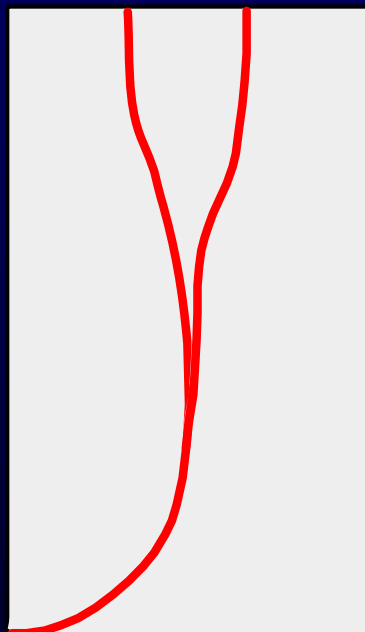
individual-based  
stochastic process



x →  
trait value

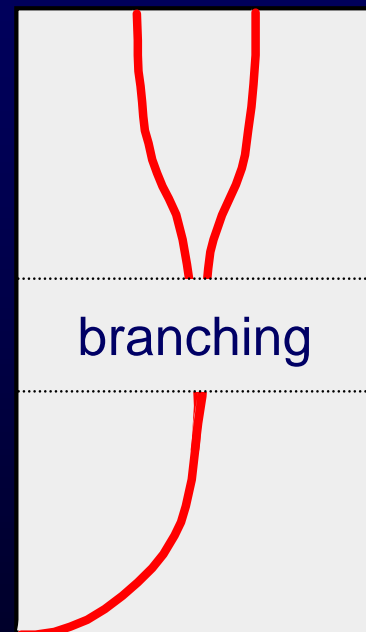
limit type:

adaptive dynamics  
limit



system size  $\rightarrow \infty$   
successful mutations/time  $\rightarrow 0$

canonical equation  
limit



mutational step size  
 $\rightarrow 0$

# the canonical equation of adaptive dynamics

---

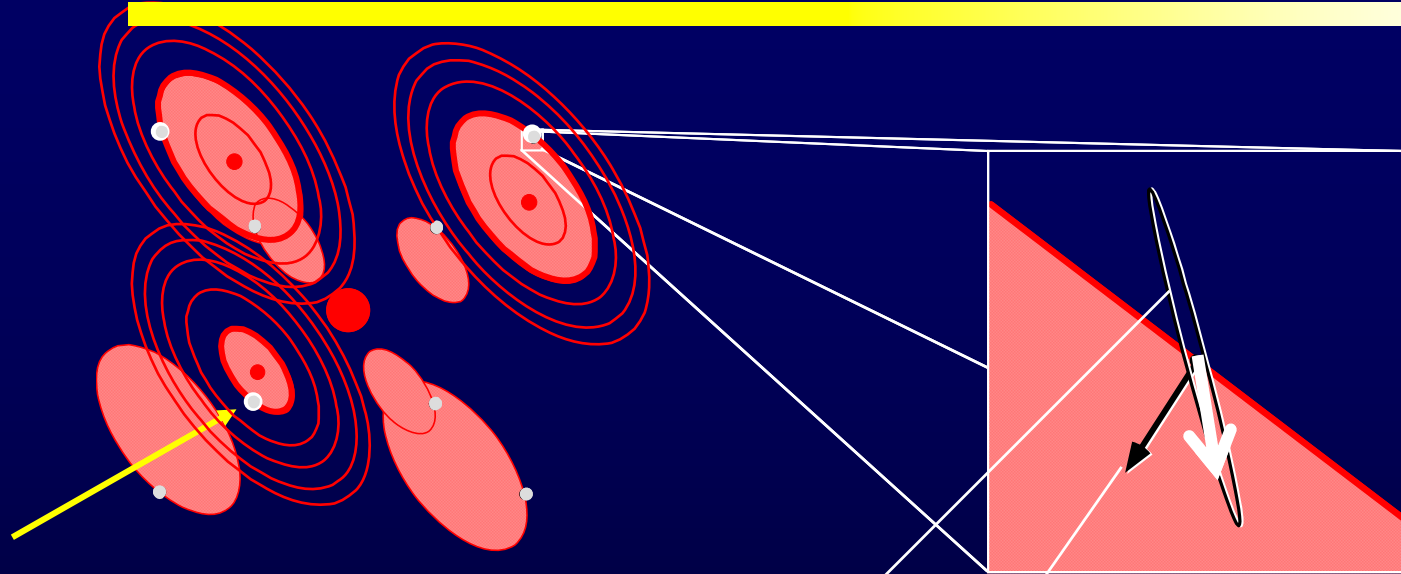
The “canonical equation of adaptive dynamics” was derived to describe the evolution of quantitative traits in realistic ecologies.

Assumptions: large population sizes, mutation limitation, small mutational steps\*.

$$\frac{dX}{dt} = \frac{1}{2} n \varepsilon C \left[ \frac{\partial s(Y|X)}{\partial Y} \Big|_{Y=X} \right]^T$$

\* and, initially, simple ODE population models

# the canonical equation of adaptive dynamics



$$\frac{dX}{dt} = \frac{1}{2} n \varepsilon C \left[ \frac{\partial s(Y|X)}{\partial Y} \Big|_{Y=X} \right]^T$$

$X$ : value of trait vector predominant in the population

$n$ : population size,  $\varepsilon$ : mutation probability per birth event

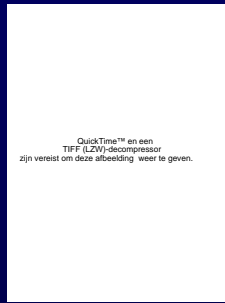
$C$ : mutational covariance matrix,  $s$ : invasion fitness, i.e., initial relative growth rate of a potential  $Y$  mutant population.

# history

basic ideas and first derivation (1996)



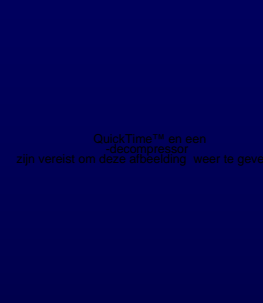
Ulf Dieckmann & Richard Law



Nicolas Champagnat & Sylvie Méléard



hard proofs (2003)



extensions (2008)



Michel Durinx & me

Mendelian  
diploids

$$\frac{dX}{dt} = 2^{n_{e,A}} C \left[ \frac{\partial s(Y|X)}{\partial Y} \Big|_{Y=X} \right]^T$$

general  
life histories

$$n_{e,A} = n_{e,D}$$

---

The intriguing result is that

$$n_{e,A} = n_{e,D}$$

the effective population size for random genetic drift.

# derivation of the canonical equation

From this starting point the canonical equation appears first in another form:

$$\frac{dX}{dt} = b \varepsilon \left( \frac{2}{\sigma_e^2} C \left[ \frac{\partial R_0(Y|E_X)}{\partial Y} \Big|_{Y=X} \right]^T \right)$$

births per  
unit of time  
 $B = bU, \mathbf{1}^T U = 1$

probability of mutation  
per birth event

mean of  
[mutational step  $\times$  approximation for the  
probability that a  $Y$ -mutant invades]

evolution is mutation limited

mutational steps are small

# demographic ingredients 1

$$\frac{2}{\sigma_e^2} \left[ \frac{\partial R_0(Y|E_X)}{\partial Y} \Big|_{Y=X} \right]$$

\*  $R_0(Y|E_X) \equiv$  average life-time offspring number of a mutant allele producing  $Y$  when singly substituted in the resident genotype

(calculated as dominant eigenvalue of a next generation operator  $L(Y|E_X)$ ).

For the resident:  $R_0(X|E_X) = 1.$

$U$   $\equiv$  stable birth state distribution of resident( allele)s ,  
normalised dominating right eigenvector of

$V$   $\equiv$  (birth rate based) reproductive values of newborn resident( allele)s,  
co-normalised dominating left eigenvector of  $L(X|E_X)$ ,  $V^T U =$

1.

# demographic ingredients 2

$$\frac{2}{\sigma_e^2} \left[ \frac{\partial R_0(Y|E_X)}{\partial Y} \Big|_{Y=X} \right]$$

\* 
$$\sigma_e^2 := \sum_j \text{Var}_j \left( \sum_i v_i \underline{m}_{ij} \right) u_j$$

with  $\underline{m}_{ij}$  the lifetime number of offspring of type  $i$  begotten by a resident allele born in state  $j$ ,

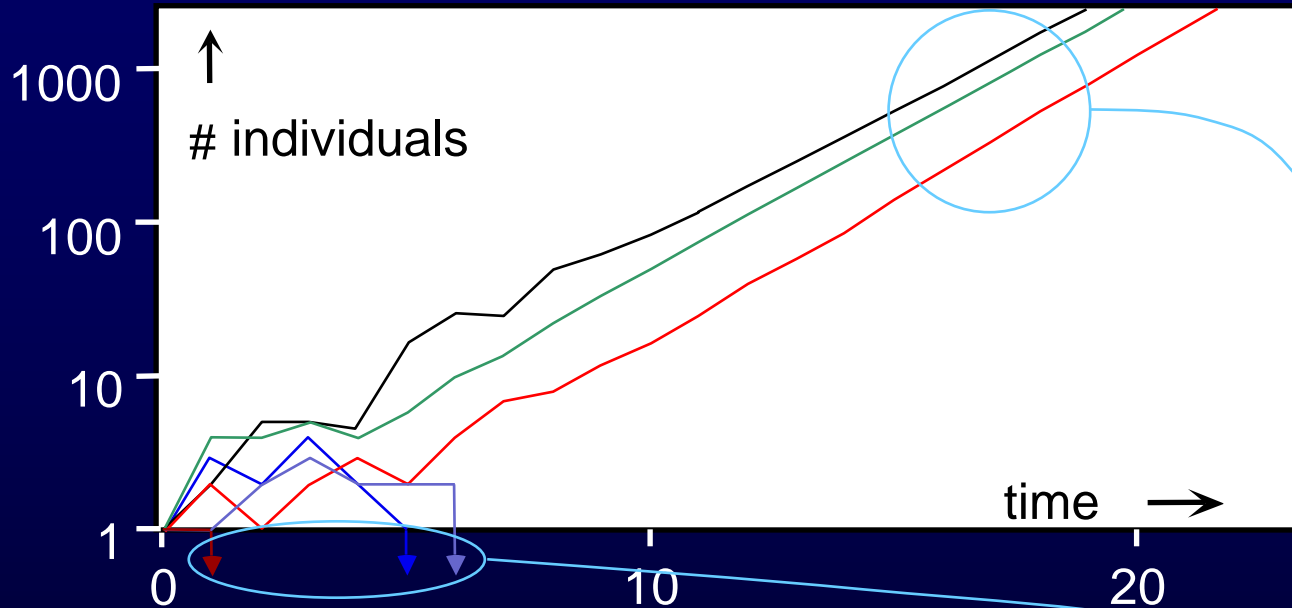
For the resident:  $R_0(X|E_X) = 1.$

$U \equiv$  stable birth state distribution of resident( allele)s ,  
normalised dominating right eigenvector of

$L(X|E_X)$ ,  $1^T U = 1$ ,  
 $V \equiv$  (birth rate based) reproductive values of newborn resident( allele)s,  
co-normalised dominating left eigenvector of  $L(X|E_X)$ ,  $V^T U =$

1.

# from the theory of branching processes



In an ergodic environment  $E_X$   
 a population starting from a single individual:  
 either goes extinct, with probability  $Q(Y|E_X)$ ,  
 or "grows exponentially" at a relative rate  $s(Y|X)$ .

For constant  $E_X$  and small  $s$ :

$$P = 1 - Q \approx \left[ \frac{2}{\sigma_e^2} \ln(R_0) \right]_{+}$$

QuickTime™ en een  
 TIFF (LZW)-decompressor  
 zijn vereist om deze afbeelding weer te geven.

QuickTime™ en een  
 TIFF (LZW)-decompressor  
 zijn vereist om deze afbeelding weer te geven.

# the average effective step

$$\frac{dX}{dt} = b \varepsilon \frac{2}{\sigma_e^2} \mathbf{C} \left[ \frac{\partial R_0(Y|E_X)}{\partial Y} \Big|_{Y=X} \right]^T$$

$$Z = Y - X$$

approximation for the probability that a  $Y$ -mutant invades

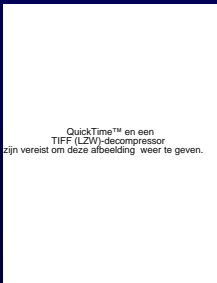
probability density of mutational steps

$$\int_{-\infty}^{\infty} Z$$

$$Z^T$$



$$g(Z) dZ$$



Brook Taylor

||

Smooth genotype to phenotype maps lead to locally additive genetics.  
 $\Rightarrow$  In diploids the mutational effect doubles over a full substitution

$$\frac{2}{\sigma_e^2} \mathbf{C} \left[ \frac{\partial R_0(Y|E_X)}{\partial Y} \Big|_{Y=X} \right]^T$$

mutational covariance matrix



## aside: mutational covariance may still vary

---

In the Mendelian case the derivation from the basic ingredients, genotype to phenotype map and genotypic mutation structure, shows that the canonical equation represents but the lowest level in a moment expansion.

The next level is a differential equation for change of the mutational covariance matrix, depending on mutational 3rd moments, etc.

# recovering population size and fitness

$$\frac{dX}{dt} = b \varepsilon \frac{2}{\sigma_e^2} C \left[ \frac{\partial R_0(Y|E_X)}{\partial Y} \Big|_{Y=X} \right]^T = 2 \frac{n_{e,A}}{\sigma_e^2} n \varepsilon C \left[ \frac{\partial s(Y|X)}{\partial Y} \Big|_{Y=X} \right]^T$$

$T_r$ : average age at giving birth,

$T_s$ : average survival time

} of the residents.

$$n = b T_s \Rightarrow b = \frac{n}{T_s} \quad s(Y|X) \approx \frac{\ln[R_0(Y|E_X)]}{T_r}$$

# demographic ingredients 3

$$\frac{dX}{dt} = 2 \frac{T_r}{T_s \sigma_e^2} n \varepsilon C \left[ \frac{\partial s(Y|X)}{\partial Y} \Big|_{Y=X} \right]^T$$

$$* \quad T_s = \int_0^{\infty} a F^T(a) U \, da$$

with  $f_i(a)$  the probability density of the age at death of an individual born in state  $i$ .

$$* \quad T_r = \int_0^{\infty} a V^T d\Lambda(a) U$$

with  $\Lambda(a)$  composed of the average pro capita births up to age  $a$ .

$$\left( \int_0^{\infty} V^T d\Lambda(a) U = V^T L(X | E_X) U = 1. \right)$$

# aside: robustness of the CE

---

QuickTime™ en een  
TIFF ILBM-decompressor  
zijn vereist om deze afbeelding weer te geven.

Géza Meszéna

For small mutational steps  
the influence of the mutants on the environment at higher densities  
comes in only as a term with a higher order in the mutational step size  
than the terms accounted for by the canonical equation.



Whether or not some previous mutants  
have not yet gone to fixation  
has little influence on the invasion of new mutants.



The applicability of the canonical equation  
extends well beyond the case of strict mutation limitation.

# random genetic Drift

---

## Definition:

The effective population size for random genetic Drift,  $n_{e,D}$ , is most easily defined as the parameter occurring in the diffusion approximation for the temporal development of the probability density  $\phi$  of the frequency  $\tilde{x}$  of a neutral gene.

$$\frac{\partial \phi}{\partial t} = \frac{1}{2} \frac{1}{n_{e,D}} \frac{\partial^2 \tilde{x}(1-\tilde{x})\phi}{\partial \tilde{x}^2}$$

# relating the effective population sizes

---

## Proposition:

The effective population sizes  $n_{e,A}$  for Adaptive evolution and  $n_{e,D}$  for random genetic Drift are equal, whatever the life history or ecological embedding.

# idea of the "proof"

Connect the approximation formula for the invasion probability derived from branching process theory  $P_B(s)$  (Eshel's formula) with the probability for take-over for a diffusion approximation including selection,  $P_D(\tilde{x}_0, \tilde{s})$ :

$$2 \frac{T_r}{\sigma_e^2} = \lim_{s \rightarrow 0} \frac{P_B(s)}{s} = \lim_{\substack{n \rightarrow \infty \\ s \rightarrow 0}} \frac{P(s, n)}{s} = \lim_{\substack{ns \rightarrow \infty \\ n_{e,D}/n \text{ constant}}} \lim_{s \rightarrow 0} \frac{P_D\left(\frac{1}{2n} \frac{T_s}{T_r}, T_r s\right)}{s} = 2T_s \frac{n_{e,D}}{n}$$

initial frequency of mutant allele      selection coefficient

- Why this particular combination of limits ?
- Why this formula for the initial mutant allele frequency  $\tilde{x}_0$ ?
- Why this formula for the selection coefficient  $\tilde{s}$ ?

# idea of the “proof”

---

$$2 \frac{T_r}{\sigma_e^2} = \lim_{s \rightarrow 0} \frac{P_B(s)}{s} = \lim_{\substack{ns \rightarrow \infty, \\ n_{e,D}/n \text{ constant}}} \lim_{s \rightarrow 0} \frac{P_D\left(\frac{1}{2n} \frac{T_s}{T_r}, T_r s\right)}{s} = 2T_s \frac{n_{e,D}}{n}$$

Why this formula for the selection coefficient  $\tilde{s}$  ?

---

In the diffusion approximation time is customarily measured in generations. Hence  $\tilde{s} = T_r s$ .

# idea of the “proof”

---

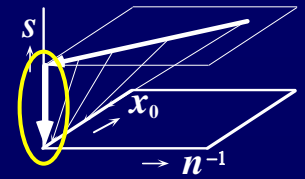
$$2 \frac{T_r}{\sigma_e^2} = \lim_{s \rightarrow 0} \frac{P_B(s)}{s} = \lim_{\substack{ns \rightarrow \infty \\ n_{e,D}/n \text{ constant}}} \lim_{s \rightarrow 0} \frac{P_D\left(\frac{1}{2n} \frac{T_s}{T_r}, T_r s\right)}{s} = 2T_s \frac{n_{e,D}}{n}$$

Why this particular combination of limits ?

---

- $n_{e,D}/n$  is a dimensionless ratio of life history parameters.
- We want to determine the first order term for small positive  $s$  in the Taylor expansion of the invasion probability.
- We want to remove the effect of the finiteness of  $n$ .

# idea of the “proof”



$$2 \frac{T_r}{\sigma_e^2} = \lim_{s \rightarrow 0} \frac{P_B(s)}{s} = \lim_{\substack{ns \rightarrow \infty \\ n_{e,D}/n \text{ constant}}} \lim_{s \rightarrow 0} \frac{P_D\left(\frac{1}{2n} \frac{T_s}{T_r}, T_r s\right)}{s} = 2T_s \frac{n_{e,D}}{n}$$

Why this particular combination of limits ?

- Left and right we take subsequent limits in different ways:
    - (1) from full population dynamics to branching process, followed by an approximation for the invasion probability,
    - (2) from full population dynamics to diffusion process, followed by a two-step approximation for the invasion probability.
- To get equality we want the corresponding paths in parameter space to approach the eventual limit point from the same direction.

# idea of the “proof”

---

$$2 \frac{T_r}{\sigma_e^2} = \lim_{s \rightarrow 0} \frac{P_B(s)}{s} = \lim_{\substack{ns \rightarrow \infty, \\ n_{e,D}/n \text{ constant}}} \lim_{s \rightarrow 0} \frac{P_D\left(\frac{1}{2n} \frac{T_s}{T_r}, T_r s\right)}{s} = 2T_s \frac{n_{e,D}}{n}$$

Why this formula for the initial mutant allele frequency  $\tilde{x}_0$ ?

---

# belief

Any relevant life history process can be uniformly approximated by finite state processes.

In continuous time at population dynamical equilibrium ( $E = E_X$ ):

**B**: individual level state transition generator

**A**: average birth rate operator

Order the states such that the birth states come first. Then the next generation operator  $L(X, E_X)$  can be expressed as

$$L(X, E_X) = K^T A B^{-1} K \quad \text{with} \quad K^T = \begin{pmatrix} 1 & 0 & L & 0 & 0 & L & L & 0 \\ 0 & 0 & 0 & M & M & & & M \\ M & 0 & 0 & 0 & M & & & M \\ 0 & L & 0 & 1 & 0 & L & L & 0 \end{pmatrix}$$

# idea of the “proof”

$$2 \frac{T_r}{\sigma_e^2} = \lim_{s \rightarrow 0} \frac{P_B(s)}{s} = \lim_{\substack{ns \rightarrow \infty, \\ n_{e,D}/n \text{ constant}}} \lim_{s \rightarrow 0} \frac{P_D \left( \frac{1}{2n} \frac{T_s}{T_r}, T_r s \right)}{s} = 2T_s \frac{n_{e,D}}{n}$$

Why this formula for the initial mutant allele frequency  $\tilde{x}_0$ ?

For  $|N''_0| = |N'_0| = |\bar{N}| = n$  the fast process results in

$$\tilde{x}_0 = \frac{1}{2n} \tilde{\mathbf{V}}^T \mathbf{N}'_0$$

$\tilde{\mathbf{V}}^T$ : co-normalised left eigenvector of  $\mathbf{A} + \mathbf{B}$

# idea of the "proof"

$$2 \frac{T_r}{\sigma_e^2} = \lim_{s \rightarrow 0} \frac{P_B(s)}{s} = \lim_{\substack{ns \rightarrow \infty, \\ n_{e,D}/n \text{ constant}}} \lim_{s \rightarrow 0} \frac{P_D \left( \frac{1}{2n} \frac{T_s}{T_r}, T_r s \right)}{s} = 2T_s \frac{n_{e,D}}{n}$$

Why this formula for the initial mutant allele frequency  $\tilde{x}_0$ ?

For  $|N''_0| = |N'_0| = |\bar{N}| = n$  the fast process results in

$$\tilde{x}_0 = \frac{1}{2n} \tilde{V}^T N'_0$$

$\tilde{V}^T$ : co-normalised left eigenvector of  $A+B$

For a just appeared mutant:  $N'_0 = KU$

$$\tilde{x}_0 = \frac{1}{2n} \frac{T_s}{T_r}$$

Lemma:

$$\tilde{V}^T K = \frac{T_s}{T_r} V^T$$

# corollary

$$\left(2 \frac{T_r}{\sigma_e^2}\right) = \lim_{s \rightarrow 0} \frac{P_B(s)}{s} = \lim_{\substack{ns \rightarrow \infty, \\ n_{e,D}/n \text{ constant}}} \lim_{s \rightarrow 0} \frac{P_D\left(\frac{1}{2n} \frac{T_s}{T_r}, T_r s\right)}{s} = \left(2 T_s \frac{n_{e,D}}{n}\right)$$



$$n_{e,D} = \frac{T_r}{T_s} \frac{n}{\sigma_e^2},$$

a result already reached by different means by  
William G Hill (1972)

for the simple age-dependent case,  
and by

Edward Pollak (1979, ...)

for the age dependent case with multiple birth states.



William G Hill



Edward Pollak

# a tricky point

---

The calculation linking the initial condition for the structured mutant population to the initial condition of the diffusion of the gene frequency assumes that there is no need to account for extinctions before the reaching of the slow manifold, in blatant disagreement with the fact that mutants arrive singly.

The justification lies in a relation linking the invasion probabilities of single invaders to those of mutants arriving in groups that are

- so small that their influence on the environment can be neglected, yet
- so large that the probability of their extinction before their  $i$ -state distribution has stabilised can be neglected.

# a tricky point, cont'd

Consider a mutant population starting from an initial cohort of newborns  $\underline{\xi} = (\xi_1, \dots, \xi_k)$ ,  $\xi_i$  the number in birth state  $i$ , drawn from some law  $\mu$ .

Let  $P(\mu, s)$  be the corresponding probability of invasion.

\_\_\_\_\_ Ansatz:  $P(\mu, s) = \alpha(\mu)s + o(s)$ ,  $\alpha(\delta_i) =: \eta_i$ .

Then (under some conditions on the “tail” of  $\mu$ )

$$P(\mu, s) = \mathbf{E}_\mu \left[ 1 - \prod_{i=1}^k (1 - P(\delta_i, s))^{\xi_i} \right] + o(\mathbf{E}_\mu \|\underline{\xi}\|) + o(s)$$

Hence

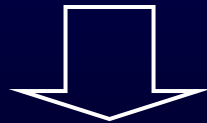
$$\alpha(\mu) = \lim_{s \rightarrow 0} \mathbf{E}_\mu \frac{\left[ 1 - \prod_{i=1}^k (1 - P(\delta_i, s))^{\xi_i} \right]}{s} = \mathbf{E}_\mu \left[ \sum_{i=1}^k \eta_i \xi_i \right] = H^T \mathbf{E}_\mu \underline{\xi}$$

# a consequence for biology

---

At equal genetic and developmental architectures  
and population dynamical regimes

the speeds of adaptive evolution and random genetic drift  
are inversely proportional.



Already for moderately large effective population sizes  
adaptive processes dominate, and  
neutral substitutions will be largely caused by genetic draft.



Thank you for your attention

serendipitous pay off:

quicker and more insightful ways  
for calculating  $n_{e,D}$

# example: micro- and macrogametes

Assumption: everybody can be considered to be born (stochastically) equal.

$$* \quad R_0(Y|E_X) = \frac{1}{2} [f(Y|E_X) + m(Y|E_X)]$$

with

$f(Y|E_X)$  } the average lifetime { macrogametes } of the mutant  
 $m(Y|E_X)$  } number of successful { microgametes } heterozygote

For the resident:  $f(X|E_X) = m(X|E_X) = 1.$

# example: micro- and macrogametes

contribution of Mendelian sampling



$$* \quad \sigma_e^2 = \frac{1}{4} \left[ \sigma_f^2 + 2(c_{f,m} + 1) + \sigma_m^2 \right]$$

with

$\left. \begin{array}{l} \sigma_f^2 \\ \sigma_m^2 \end{array} \right\}$  the variance of the lifetime number of successful  $\left\{ \begin{array}{l} \text{macrogametes} \\ \text{microgametes} \end{array} \right.$

of the resident, and

$c_{f,m}$  the covariance of these two lifetime numbers.

# example: micro- and macrogametes

Denote the average pro capita rate of parenting at age  $a$  as  $\lambda(a)$ .

$$\lambda(a) = \lambda_f(a) + \lambda_m(a)$$

$$\int_0^{\infty} \lambda_f(a) da = f(X|E_X) = 1$$

$$\int_0^{\infty} \lambda_m(a) da = m(X|E_X) = 1$$

\* 
$$T_r = \int_0^{\infty} a \frac{1}{2} \lambda(a) da = \frac{1}{2} (T_{r,f} + T_{r,m})$$

# example: micro- and macrogametes

---

\* 
$$T_s = \int_0^{\infty} a f(a) da$$

with  $f(a)$  the probability density of the age at death of an individual.

## special case: separate sexes

Call the relative frequencies at which the two sexes are born into the population  $q_f$  and  $q_m$  and their average life-lengths  $T_{s,f}$  and  $T_{s,m}$ , then

$$T_s = q_f T_{s,f} + q_m T_{s,m}.$$

# example: micro- and macrogametes

Let a subscript  $+$  indicate that the quantity in question refers to the number of offspring of a female or male, then

$$f(Y|E_X) = q_f(Y|E_X)f_+(Y|E_X), \quad m(Y|E_X) = q_m(Y|E_X)m_+(Y|E_X).$$

Hence

$$R_0(Y|E_X) = \frac{1}{2} \left[ q_f(Y|E_X)f_+(Y|E_X) + q_m(Y|E_X)m_+(Y|E_X) \right].$$

For the resident  $f=1$  and  $m=1$ . Hence

$$f_+(X|E_X) = \frac{1}{q_f}, \quad m_+(Y|E_X) = \frac{1}{q_m}.$$

Moreover, the production of macro- and microgametes are mutually exclusive. Hence

$$c_{fm} = -1.$$

# on calculating variances

Consider a discrete nonnegative random variable  $\underline{h}$  such that

with probability  $p$ :  $\underline{h} = 0$

with probability  $q = (1-p)$ :  $\underline{h} = \underline{k}$

with  $\underline{k}$  another nonnegative random variable,

then  $\mathbf{E}\underline{h} = q\mathbf{E}\underline{k}$  and

$$\mathbf{Var}(\underline{h}) = \mathbf{E}\underline{h}^2 - (\mathbf{E}\underline{h})^2 = q \left[ \mathbf{Var}(\underline{k}) + (\mathbf{E}\underline{k})^2 \right] - (q\mathbf{E}\underline{k})^2.$$

When moreover  $\mathbf{E}\underline{h} = 1$  so that  $\mathbf{E}\underline{k} = q^{-1}$ :  $\theta_k^2$

$$\mathbf{Var}(\underline{h}) = \mathbf{E}\underline{h}^2 - 1 = q \left[ \mathbf{Var}(\underline{k}) + q^{-2} \right] - 1 = \frac{\mathbf{Var}(\underline{k}) / (\mathbf{E}\underline{k})^2 + 1 - q}{q}$$

# example: separate sexes

$$\sigma_e^2 = \frac{1}{4} \left[ \sigma_f^2 + \cancel{2(c_{f,m} + 1)} + \sigma_m^2 \right] = \frac{\theta_{f+}^2 + 1 - q_f}{4q_f} + \frac{\theta_{m+}^2 + 1 - q_m}{4q_m}$$

$\theta$  the coefficient of variation (standard deviation/mean).

$$n_e = \frac{T_r}{T_s} \frac{n}{\sigma_e^2} = \frac{T_r}{T_s} \frac{1}{\frac{\theta_{f+}^2 + 1 - q_f}{4(T_s/T_{s,f})n_f} + \frac{\theta_{m+}^2 + 1 - q_m}{4(T_s/T_{s,m})n_m}}$$

With the definitions

$$n_{e,f} = \frac{n_f}{\theta_{f+}^2 + 1 - q_f}, \quad n_{e,m} = \frac{n_m}{\theta_{m+}^2 + 1 - q_m}, \quad \alpha_f = \frac{T_{r,f}}{T_r}, \quad \alpha_m = \frac{T_{r,m}}{T_r},$$

this gives

$$n_e = \frac{4n_{e,f}n_{e,m}}{\alpha_f n_{e,f} + \alpha_m n_{e,m}}.$$

Generalises the textbook formula for the case  $T_{r,f} = T_{r,m} = 1 (\Rightarrow \alpha_f = \alpha_m = 1)$ .



The end

# slowing down by within population variation

---

Away from full mutation limitation, evolution is slowed down a bit by the additional variability in the offspring number of newly introduced alleles coming from the variability in the partners with whom they team up to make a body:

$$\sigma_e^2 \leftarrow \sigma_e^2 + \frac{1}{2} \left( \frac{dR_0}{dX} \right)^T C_p \left( \frac{dR_0}{dX} \right)$$

$C_p$  the covariance matrix of the variation of  $X$  in the population.

If  $\sigma_e^2$  would be calculated based on empirically determined components the additional term is automatically incorporated.

However, the above expression only gives a very partial picture since  $C_p$  is bound to fluctuate in time.

# dissecting the limit

$$\lim_{\substack{ns \rightarrow \infty, \quad s \rightarrow 0 \\ n_{e,D}/n \text{ constant}}} P_D \left( \frac{1}{2n} \frac{T_s}{T_r}, T_r s \right)$$

# why this particular limit?

---

First hint:

For Karlin models the chosen limit procedure naturally recovers the result from their branching process limits.

# Karlin models (without any population structure)

In Karlin's conditioned direct product branching process models

$n/n_{e,D} = \sigma^2 = \text{variance of offspring number}$

$$P_D(n^{-1}, s) = \frac{1 - \exp[-2s/\sigma^2]}{1 - \exp[-2ns/\sigma^2]}$$



$$\lim_{ns \rightarrow \infty, s \rightarrow 0} \frac{P_D(n^{-1}, s)}{2s} = \frac{1}{\sigma^2},$$

which equals the limit result from branching process theory.

# why this particular limit?

$$\frac{1 - \exp\left[-2n_{e,D} \frac{s}{\rho n}\right]}{1 - \exp\left[-4n_{e,D} \frac{s}{\rho}\right]}$$

First hint:

For Karlin models the chosen limit procedure naturally recovers the result from their branching process limits.

General heuristics:

- The effect of finite population size should disappear. (This effect occurs primarily in the integrals making up the denominator of the diffusion formula.)  $\rightarrow$  Let  $n_{e,D} \tilde{s} \rightarrow \infty$ .
- The goal is to obtain the coefficient of the linear term in an expansion in  $s$ .  $\rightarrow$  Divide by  $\tilde{s}$  and let  $\tilde{s} \rightarrow \infty$ .

Eventual “justification”:

A study of the paths in parameter space corresponding to the two ways of taking subsequent limit shows that in both paths the limit point is approached in a similar manner.

# diffusion approximation versus limit

Population geneticists usually express the diffusion approximation in the variables relative gene frequency,  $x$ , and generations,  $t$ .

Mathematicians express the diffusion limit in rescaled time  $\tau = t/n$ , leading to (for the equation with selection)

$$\frac{\partial \phi}{\partial \tau} = \frac{1}{2} \frac{1}{2 \left[ n_{e,D}/n \right]} \frac{\partial^2 x(1-x)\phi}{\partial x^2} + \left[ n\tilde{s} \right] \frac{\partial x(1-x)\phi}{\partial x}$$

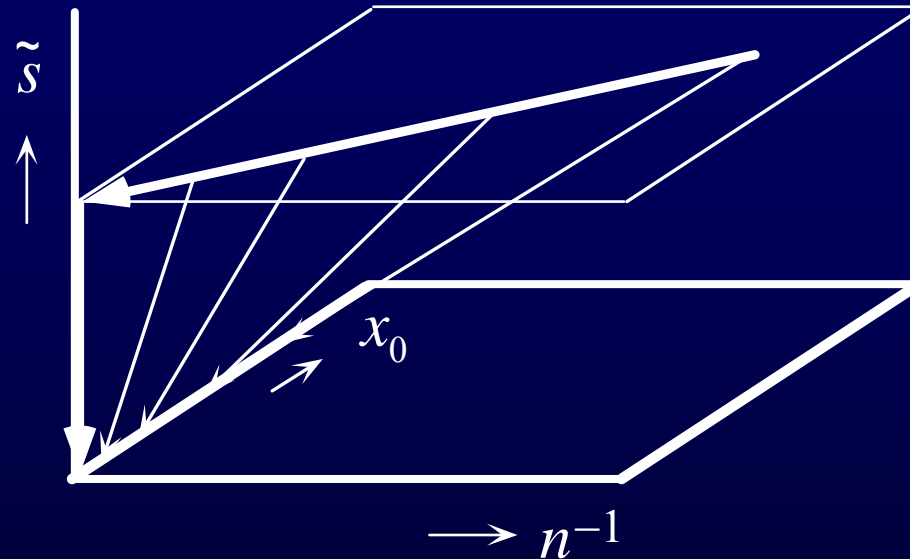
where the expressions

$$\left[ n_{e,D}/n \right] \quad \left[ n\tilde{s} \right]$$

should be interpreted as single symbols instead of as formulas:

$$\left[ n_{e,D}/n \right] := \lim_{n \rightarrow \infty} n_{e,D}/n \quad \left[ n\tilde{s} \right] := \lim_{n \rightarrow \infty} n\tilde{s} \quad \Rightarrow \quad \tilde{s} \approx \left[ n\tilde{s} \right] n^{-1}$$

# the paths in parameter space



**Bold arrows:** branching process limit followed by  $\tilde{s} \rightarrow 0$ .

**Thin arrows:** each arrow corresponds to a separate diffusion limit; their angles correspond to the values of  $[n\tilde{s}]$ , their endpoints to the values of  $x_0$ .

The sequence of arrows corresponds to the limits  $n\tilde{s} \rightarrow \infty$ ,  $\tilde{s} \rightarrow 0$ .

proof of the lemma

$$\tilde{\mathbf{V}}^T \mathbf{K} = \frac{T_s}{T_r} \mathbf{V}^T$$

# proof of the lemma

$$\tilde{\mathbf{V}}^T \mathbf{K} = \frac{T_s}{T_r} \mathbf{V}^T$$

$S$  :  $s_{ij}$  = expected residence time in state  $i$  of an individual born in state  $j$

$\tilde{\mathbf{U}}$  : normed stationary distribution of h-states (right eigenvector of  $\mathbf{A} + \mathbf{B}$ )

$\tilde{\mathbf{V}}$  : reproductive values of h-states (left eigenvector of  $\mathbf{A} + \mathbf{B}$ ,  $\tilde{\mathbf{V}}^T \tilde{\mathbf{U}} = 1$ )

The upper part of  $\tilde{\mathbf{V}}$  is proportional to  $\mathbf{V}$

$\tilde{\mathbf{L}}$  :  $\tilde{l}_{ij}$  = expected future # kids in state  $i$  from parent who is now in state  $j$

The leftmost square block of  $\tilde{\mathbf{L}}$  equals  $\mathbf{L}$ .

$$T_s = \mathbf{1}^T \mathbf{S} \mathbf{U}$$

$$\tilde{\mathbf{U}} = T_s^{-1} \mathbf{S} \mathbf{U}$$

$$\tilde{\mathbf{V}}^T = c \mathbf{V}^T \tilde{\mathbf{L}}$$

since  $v_i$  is the expected long term birth rate if the mutant population is started with a single individual in state  $i$  and  $\tilde{v}_i$  is the expected long term population size when the mutant population is started with a single individual in state  $i$ .

$$\tilde{\mathbf{V}}^T \tilde{\mathbf{U}} = c \mathbf{V}^T \tilde{\mathbf{L}} T_s^{-1} \mathbf{S} \mathbf{U} = 1 \quad \Rightarrow \quad c = \frac{T_s}{\mathbf{V}^T \tilde{\mathbf{L}} \mathbf{S} \mathbf{U}} = T_r ?$$

# proof of the lemma

$$c = \frac{T_s}{\mathbf{V}^T \mathbf{L} \mathbf{S} \mathbf{U}} = T_r ?$$

$T_r$  is defined as

$$T_r = \int_0^{\infty} a \mathbf{V}^T \Lambda(a) \mathbf{U} da = \mathbf{V}^T \int_0^{\infty} \int_a^{\infty} \Lambda(\alpha) d\alpha da \mathbf{U}$$

with  $\Lambda(a)$  containing the average pro capita birth rates at age  $a$ .

For a finite state process

$$\Lambda(a) = \mathbf{K}^T \mathbf{A} e^{Ba} \mathbf{K} \quad \text{so that} \quad T_r = \mathbf{V}^T \mathbf{K}^T \mathbf{A} \mathbf{B}^{-2} \mathbf{K} \mathbf{U}.$$

Similarly

$$\mathbf{S} = \int_0^{\infty} e^{-Ba} da = -\mathbf{B}^{-1} \mathbf{K} \quad \text{and} \quad \tilde{\mathbf{L}} = \int_0^{\infty} \mathbf{K}^T \mathbf{A} e^{-Ba} da = -\mathbf{K}^T \mathbf{A} \mathbf{B}^{-1}.$$

Therefore indeed

$$\mathbf{V}^T \tilde{\mathbf{L}} \mathbf{S} \mathbf{U} = T_r.$$

# proof of the lemma, discrete time case

$T_r$  is defined as

$$T_r = \sum_{a=1}^{\infty} a V^T \Lambda(a) U = V^T \sum_{a=1}^{\infty} \sum_{\alpha=a}^{\infty} \Lambda(\alpha) U$$

with  $\Lambda(a)$  containing the average pro capita births at age  $a$ .

For a finite state process

$$\Lambda(a) = K^T A B^a K \quad \text{so that} \quad T_r = V^T K^T A B (I - B)^{-2} K U.$$

Similarly

$$S = \sum_{a=1}^{\infty} B^a K = B (I - B)^{-1} K \quad \text{and} \quad \tilde{L} = \sum_{a=0}^{\infty} K^T A B^a = K^T A (I - B)^{-1}.$$

Therefore indeed  $V^T \tilde{L} S U = T_r$ .

# example 2: simplest cyclic environment

$$n_j = p_j b_j \qquad b_j = \lambda_{j-1} n_{j-1}$$

$p_i$  probability that a newborn in phase  $i-1$  survives till phase  $i$

$b_i$  number of births at phase  $i$

$\lambda_i$  pro capita number of births at phase  $i$

$$L(X | E_X) = \begin{pmatrix} 0 & \cdot & \cdot & 0 & \lambda_k p_k \\ \lambda_1 p_1 & 0 & \cdot & \cdot & 0 \\ 0 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & 0 & \lambda_{k-1} p_{k-1} & 0 \end{pmatrix} = \begin{pmatrix} 0 & L & L & 0 & b_1/b_k \\ b_2/b_1 & 0 & L & L & 0 \\ 0 & O & O & & M \\ M & O & O & O & M \\ 0 & L & 0 & b_k/b_{k-1} & 0 \end{pmatrix}$$

$$V = \frac{\sum_i b_i}{k} \begin{pmatrix} 1/b_1 & L & 1/b_k \end{pmatrix} \qquad U = \frac{1}{\sum_i b_i} \begin{pmatrix} b_1 \\ M \\ b_k \end{pmatrix}$$

## example 2: simplest cyclic environment

$$\begin{aligned} \sigma_e^2 &= \sum_j \text{Var}_j \left( \sum_i v_i m_{ij} \right) u_j = \sum_j \left[ \left( \frac{\sum_i b_i}{k b_{j+1}} \right)^2 \sigma_j^2 \right] \frac{b_j}{\sum_i b_i} = \\ &= \frac{\sum_i b_i}{k} \frac{1}{k} \sum_j \frac{1}{b_{j+1}^2} \sigma_j^2 b_j = \frac{\sum_i b_i}{k} \frac{1}{k} \sum_j \frac{\sigma_j^2 / \bar{m}_j}{b_{j+1}} \end{aligned}$$

Next use  $\bar{m}_j = b_{j+1}/b_j = \lambda_j p_j$  and write  $n_{e,j} := \frac{b_{j+1}}{\sigma_j^2 / \bar{m}_j} = \frac{\lambda_j n_j}{\sigma_j^2 / \bar{m}_j} = \frac{n_j}{p_j \theta_j^2}$  to rewrite this as

$$\sigma_e^2 = \frac{1}{k} \sum_j b_i \frac{1}{k} \sum_j \frac{1}{n_{e,j}}.$$

The end result is the textbook harmonic mean expression:

$$n_e = \frac{T_r}{T_s} \frac{n}{\sigma_e^2} = \frac{1}{\langle pb \rangle / \langle b \rangle} \frac{\langle pb \rangle}{\sigma_e^2} = \frac{\langle b \rangle}{\sigma_e^2} = \langle n_{e,j}^{-1} \rangle^{-1}.$$