Population genetics

- understand patterns of genetic diversity
- diffusion theory (→ Luca Peliti)

$$\frac{\partial P}{\partial t} = \frac{\partial^2}{\partial x^2} \frac{x(1-x)}{2N} P(x,t) \stackrel{!}{=} 0$$ steady state

$$\Rightarrow \quad \frac{\partial}{\partial x} \frac{x(1-x)}{2N} P(x) = C$$

- new mutations arise at rate $N\mu$ at freq $\frac{1}{N}$
- $P(x) = \frac{2N\mu}{x} \Rightarrow$ flux μ
- pairwise diversity $\int_0^1 dx\, 2x(1-x) P(x) = 2N\mu \int_0^1 dx\, 2x = 2N\mu$

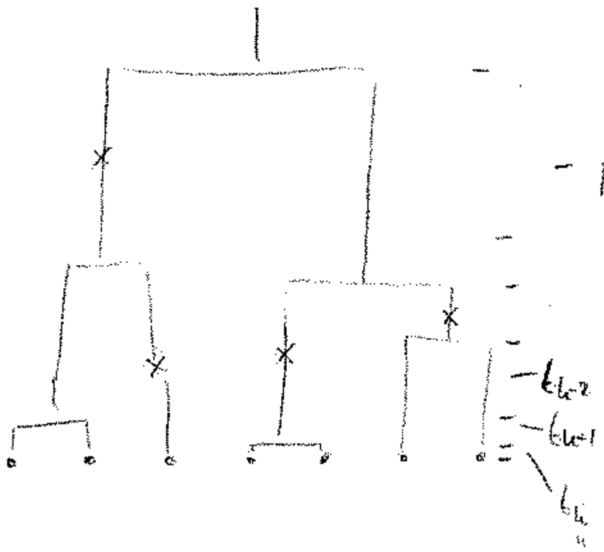
Kingman Coalescent

- every individual picks a parent at random
  → eq. to poisson offspring distribution
- Prob that a pair has a common parent

$$p_2 = \frac{1}{N}$$

- prob that any pair has a common parent

$$\frac{k(k-1)}{2N}$$



$$P(t_i) = \frac{k(k-1)}{2N} e^{-\frac{k(k-1)}{2N} t_i}$$

$$\langle T_{MRCA} \rangle = \sum_{i=2}^{k} \frac{2N}{i(i-1)} = 2N\left(1 - \frac{1}{k}\right) \longrightarrow 2N$$

$$\langle T_2 \rangle = N$$

– mutations

- if neutral → sprinkled on tree
- pairwise distance : $2\langle T_2\rangle\mu = 2N\mu$ (comp diffusion model)

2a

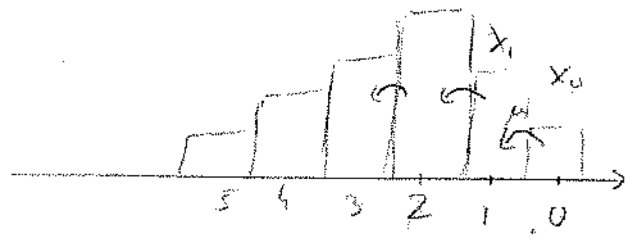
Selection – deleterious

- single locus

$$\frac{dx}{dt} = \mu x - \mu(1-x) + s\,x(1-x) \;\rightarrow\; \mu x + sx$$

if $s<0$ & $\mu \ll |s|$

deleterious mutation-selection balance



$\dot{x}_k = \mu x_{k-1} - (ks+\mu)x_k$

$\Rightarrow x_k = \frac{\lambda^k}{k!}e^{-\lambda} \quad \lambda = \frac{\mu}{s}$

$\langle ks\rangle = -\mu$

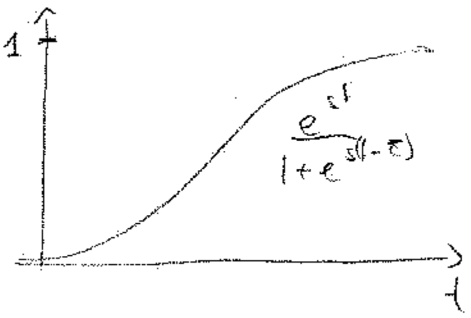
categories with $k=0,1,2\ldots$ mutations
each mutation has fitness effect $s<0$

- mutation free class $n_0 = Ne^{-\mu/s}$
- source of all future individuals
  → reduced $T_{MRCA} = 2Ne^{-\mu/s}$
- Muller's ratchet · n_0 fluctuates, if it hits zero, no way back

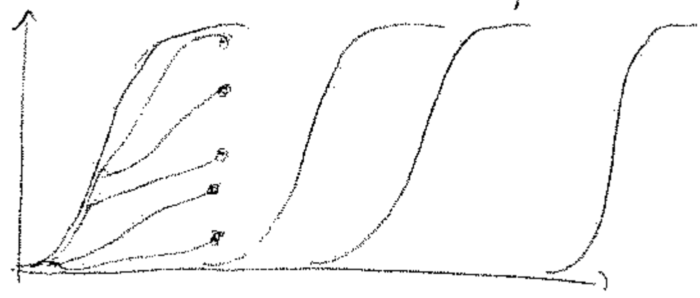
rate $\quad e^{-\alpha n_0 s} = e^{-\alpha N s e^{-\mu/s}}$

Beneficial mutations

$\dot{x} = sx(1-x)$



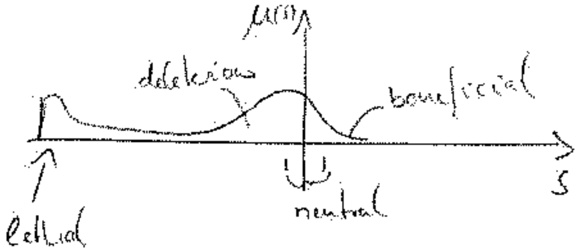
$p_{fix} = s$, $\quad N u_b s$ = rate at which beneficial mutation arise

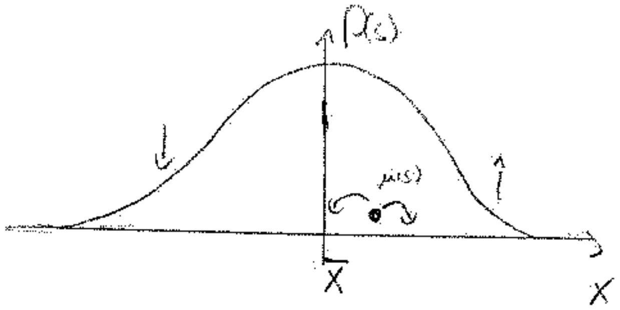


- always only a single locus polymorphic
- each sweep moves the population forward by s

$\Rightarrow \quad v = N u_b s^2$

general mutation distribution $\mu(s)$





$$\frac{\partial P}{\partial t} = \underbrace{(x - \bar{x}_{\cdot})}_{\text{selection / population size constraint}} P(x,t) + \overbrace{\int ds\, \mu(s)\left(P(x-s) - P(x)\right)}^{\text{mutation}}$$

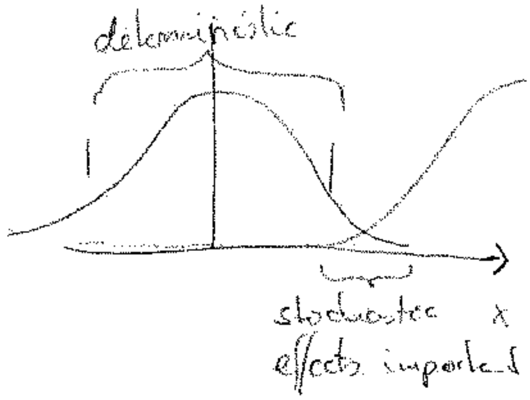
$$\bar{x} = \int dx\, P(x,t)$$

$$\dot{\bar{x}} = \int dx\, x\, \dot{P}(x,t) = \int dx\, x(x-\bar{x})P(x,t) - \int ds\, \mu(s) \int dx\, x\left(P(x-s) - P(x)\right)$$

$$= \sigma^2 - \int ds\, s\mu(s)$$

$$= \sigma^2 - \mu\langle s\rangle .$$

- moment hierarchy: σ^2 depends on skewness, etc. not closed
- diversity will decrease in absence of new mutations
- how is diversity maintained?
- as with u., only very fit individuals have any chance to leave offspring
- deterministic description valid in bulk, but tip is most important



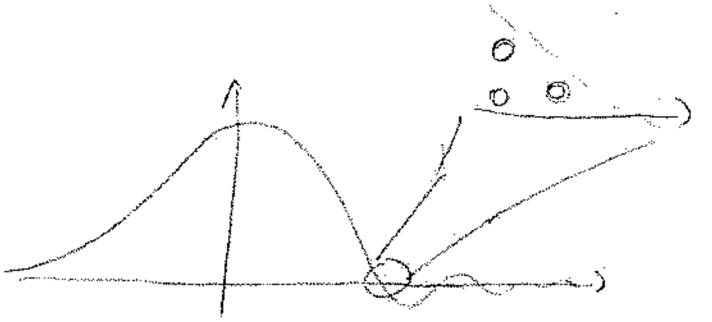
Diffusive limit

$\mu \gg s$ $$\frac{\partial P}{\partial t} = (x-\bar{x})P + \int ds\, \mu(s) \left[-s\frac{\partial P}{\partial x} + \frac{s^2}{2}\frac{\partial^2 P}{\partial x^2} \right]$$

$$= (x-\bar{x})P + \frac{\langle \mu s^2 \rangle}{2}\frac{\partial^2 P}{\partial x^2} - \langle \mu s \rangle \frac{\partial P}{\partial x}$$

- rapid mutation $\Rightarrow$ diffusion
  $\Rightarrow$ similar to FKPP equation
- comoving frame $\chi = x - (v + \langle \mu s \rangle)t$
  (assumption of steady state)

$$(v+\langle \mu s \rangle)\frac{\partial P}{\partial \chi} = \chi P + D\frac{\partial^2 P}{\partial \chi^2}$$ $\Rightarrow$ Airy function solution
$\rightarrow v$ is not yet determined



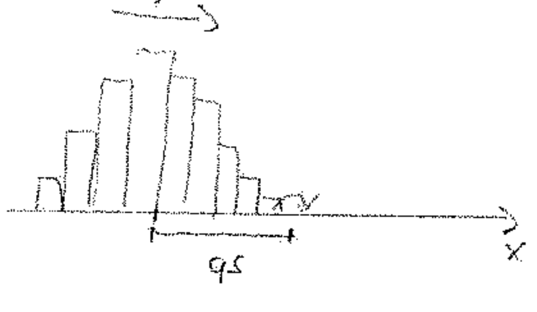
- turns negative at some large χ
- discreteness of the individuals
- selection requires ≥ 1 individual
- introduce cutoff $P(\chi)\langle \mu s^2 \rangle^{1/3} \simeq 1$

$\Rightarrow$ fixes v

$$v \approx (\langle \mu s^2 \rangle)^{1/3} \log^{1/3} N$$

Discrete limit

$\mu \ll s$, assume one discrete s for simplicity



$n_0 e^{qst} = \mu/qs \Rightarrow t \approx \frac{1}{qs}\log\mu$

$\Rightarrow v_{max} = s^2 q / \log \mu$

$t_s = qs/v$

- need to relate v_{nose} to v_{bulk}
- time for the bulk to arrive at nose

$$t_c = \frac{sq}{v}$$

- in that time, class has to grow to $\mathcal{O}(N)$
- average growth rate $\frac{sq}{2}$

$$N \simeq e^{\frac{qs}{2} t_c} = e^{\frac{q^2 s^2}{2v}} \Rightarrow v \simeq \frac{q^2 s^2}{2 \log N}$$

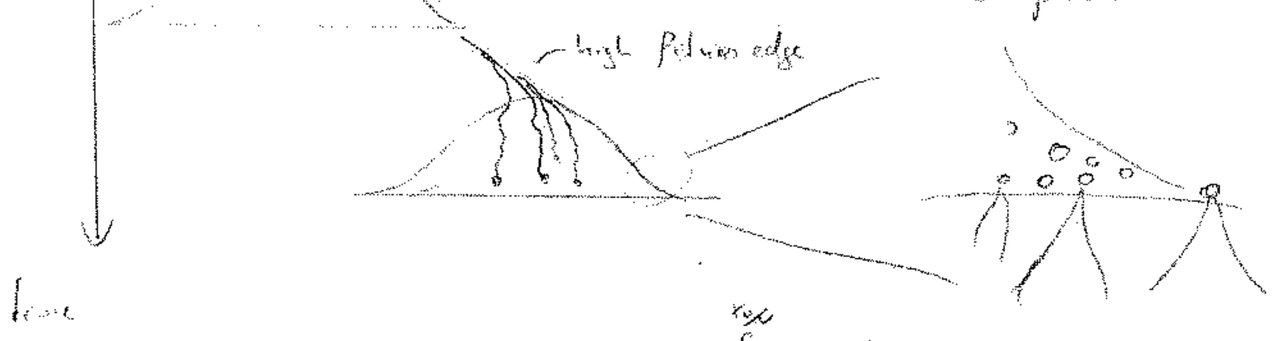
$$\Rightarrow \frac{s^2 q}{\log \mu} = \frac{q^2 s^2}{2 \log N} \Rightarrow q = \frac{2 \log N}{\log \mu}$$

$$\Rightarrow v = 2 s^2 \frac{\log N s}{\log^2 \mu}$$

- v depends logarithmic on N & μ

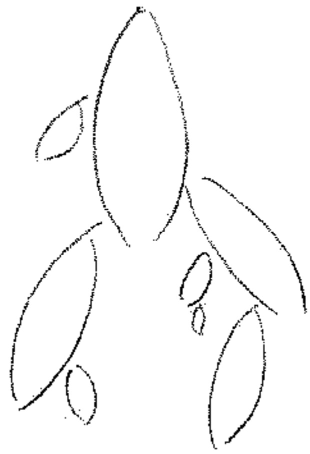
Genetic diversity in traveling waves

- lineages concentrate in nose
- competition in the nose



"bubble size": $n(x_0) = e^{\int^{x_0} dx\, x_0 - vt} = e^{\frac{x_0^2}{2v}}$

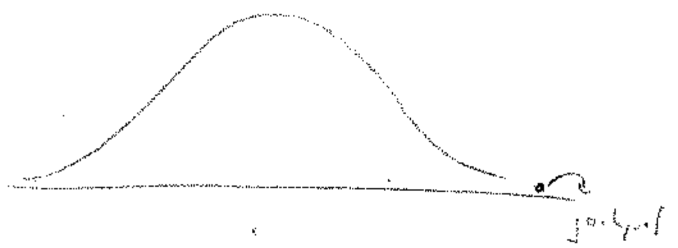
- integration over x $\rightarrow$ $p(n) \sim \frac{1}{n^2}$
- effective coalescence of bubbles



Offspring distribution $\sim n^{-2}$

$\Rightarrow$ Bolthausen - Sznitman coalescent

$\Rightarrow$ coalescence is dominated by jackpot events



• Site frequency spectra