Working w/ Diffusion limit

That's great for formalism... what can we actually do w/ SDEs/Langevin eqs./ Fokker-Planck eqs?

$\Rightarrow$ Gaussian random walk was trivially solvable... other choices for $\mu(x), \sigma^2(x)$ not so easy to integrate

* one classic example from physics (& will arise in many evolution problems as well)

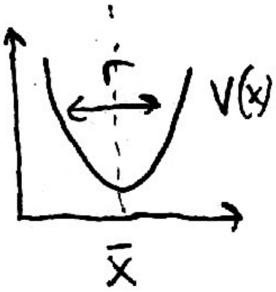
is $$\frac{dx}{dt} = \overbrace{-r(x-\bar{x})}^{\frac{\partial V}{\partial x}} + \sqrt{D}\,\eta(t)$$

r: restoring force

$\bar{x}$: equilibrium point

D: diffusion constant ($\propto \frac{1}{kT}$ in physics)

"Brownian particle in quadratic potential"



we'll use this e.g. to illustrate mechanics of SDE manipulation

e.g. say we're interested in dynamics of mean, $\langle x(t) \rangle$

From definition of SDE, $x(t+\delta t) = x(t) - r(x-\bar{x})\delta t + \sqrt{D\delta t}\, \underset{\to N(0,1)}{Z_t}$

taking averages, we have

$$\langle x(t+\delta t)\rangle = \langle x(t)\rangle - r\langle x(t)\rangle\delta t + r\bar{x}\,\delta t + \overset{\langle Z_t\rangle = 0}{0}$$

$$\Rightarrow \frac{\langle x(t+\delta t)\rangle - \langle x(t)\rangle}{\delta t} = -r\left[\langle x(t)\rangle - \bar{x}\right]$$

$$\Rightarrow \frac{d\langle x\rangle}{dt} = -r\left[\langle x\rangle - \bar{x}\right] \Rightarrow \langle x(t)\rangle - \bar{x} = \left(\langle x(0)\rangle - \bar{x}\right)e^{-rt}$$

$\Rightarrow \langle x(t)\rangle \to \bar{x}$ at rate r $\boxed{\text{Just like deterministic equation!}}$

* what about ~~spread~~ around this value? e.g. if $\bar{x}=0$, want $\langle x(t)^2\rangle$

Again, from definition: $\langle x(t+\delta t)^2\rangle = \left\langle\left[x(t) - r(x(t))\delta t + \sqrt{D\delta t}\,Z_t\right]^2\right\rangle$

expand to lowest order in δt:

$$\langle x(t+\delta t)^2\rangle = \langle x(t)^2\rangle - 2r\langle x(t)^2\rangle\delta t + \langle D\delta t\,\cancelto{1}{Z_t^2}\rangle + \cancelto{0}{2\langle x(t)\sqrt{D\delta t}\,Z_t\rangle}$$

$$\Rightarrow \frac{d\langle x^2\rangle}{dt} = \underbrace{-2r\langle x^2\rangle}_{\text{same as deterministic version}} + \underbrace{D}_{\text{new part from stochasticity}}$$

$\Rightarrow \langle x^2\rangle = \dfrac{D}{2r}$ $\boxed{\text{balance between noise \& deterministic restoring force.}}$

can actually get full distn @ long times:

E.g. for any SDE of form: $\frac{dx}{dt} = -\frac{dV(x)}{dx} + \sqrt{D}\,\eta(t)$

$\Rightarrow$ Fokker-Planck eq: $\frac{dp}{dt} = -\frac{\partial}{\partial x}\left[-\frac{dV(x)}{dx} p(x)\right] + \frac{\partial^2}{\partial x^2}\left[\frac{p(x) D}{2}\right]$

@ stationarity, $\partial_t p = 0 \Rightarrow \frac{dp}{dx} = -\frac{2}{D}\frac{dV(x)}{dx} p(x)$

$\Rightarrow p(x) \propto e^{-\frac{2V(x)}{D}}$ "Boltzmann" distribution $\Longrightarrow$ if $\frac{dV}{dx} = r(x-\bar{x})$ $p(x) \propto e^{-\frac{2r(x-\bar{x}_0)^2}{2D}}$ Gaussian dist'n

$\approx$ deterministic dynamics + a little fuzzyness from noise

this is standard physics case ... what about evolutionary model?

e.g. $\frac{df}{dt} = sf(1-f) + \sqrt{\frac{f(1-f)}{N}}\,\eta(t)$

($+\mu(1-f)$ for mutation)

($-\nu f$ for back-mutation)

2 key differences:

1. Diffusion const depends on mutant freq.
2. selection term is nonlinear

#2 becomes important if we want to calculate avgs, e.g. $\langle f(t)\rangle$.

using same approach as above $\left[f(t+\delta t) = f(t) + sf(1-f)\delta t + \sqrt{\frac{f(1-f)\delta t}{N}}\, Z_t \right]$

find: $\frac{d\langle f\rangle}{dt} = s\left[\langle f\rangle - \langle f^2\rangle\right] \Rightarrow$ need $\langle f^2(t)\rangle$ to find $\langle f(t)\rangle$

$\langle f^2\rangle$ NOT $= \langle f\rangle^2$

ok... do same thing for $\langle f(t+\delta t)^2 \rangle$:

$\Rightarrow$ find $$\frac{d\langle f^2\rangle}{dt} = 2s\langle f\cdot f(1-f)\rangle + \frac{f(1-f)}{N}$$

from deterministic part — from 2 stochastic terms

$\Rightarrow$ depends on $\langle f^3\rangle$ in addition to $\langle f^2\rangle$ and $\langle f\rangle$

$\Rightarrow$ and so on for higher moments. known as "moment hell"
(general consequence of nonlinearity)

since nonlinearity caused by selection, one sol'n is to only look @ evolution problems w/o selection [i.e., $s=0$ or "neutral theory"]

$\Rightarrow$ much of classical pop. gen focuses on this limit.
we will revisit later when we talk about multi site genomes

what about <u>stationary distribution</u>?

also trickier in evolution setting. e.g. one way mutation ($WT \xrightarrow[\mu]{} mut$)
$f=1$ is absorbing state, so
$f\to 1$ @ long times (boring)

If turn on back-mutations ($WT \underset{\nu}{\overset{\mu}{\rightleftarrows}} mut$) then no absorbing state.

In this case, can show that $p(f) \propto f^{-1}(1-f)^{-1} e^{-2N\Lambda(f)}$ is solution to Fokker-Planck equation (when $\partial_t p = 0$) if we choose $\Lambda(f)$ such that

$$\frac{\partial f}{\partial t} = f(1-f)\left[-\frac{\partial \Lambda(f)}{\partial f}\right] + \sqrt{\frac{f(1-f)}{N}}\,\eta(t)$$

$\Rightarrow$ to see this, just plug in and check:

$$\frac{1}{2N}\frac{\partial^2}{\partial f^2}\left[f(1-f)p(x)\right] = \frac{1}{2N}\frac{\partial}{\partial f}\left[\frac{\partial}{\partial f}\left[Ce^{-2N\Lambda(f)}\right]\right]$$

$$= -\frac{\partial}{\partial f}\left[\frac{\partial \Lambda}{\partial f}\, Ce^{-2N\Lambda(f)}\right]$$

$$= +\frac{\partial}{\partial f}\left[f(1-f)\left[-\frac{\partial \Lambda}{\partial f}\right]p(f)\right] \quad \checkmark$$

end of lecture 4

in this case, note that (deterministically)

$$\frac{d\Lambda}{dt} = \frac{\partial \Lambda}{\partial f}\frac{\partial f}{\partial t} = -f(1-f)\left(\frac{\partial f}{\partial t}\right)^2 \leq 0$$

so dynamics act to minimize $\Lambda(f)$ [just like "energy"] thus, since $p(f) \propto f^{-1}(1-f)^{-1} e^{-2N\Lambda(f)}$

$\Lambda(f)$ is analogy of "energy" for this SDE w/

N is analogy of "1/temp" non-constant diffusion eq.

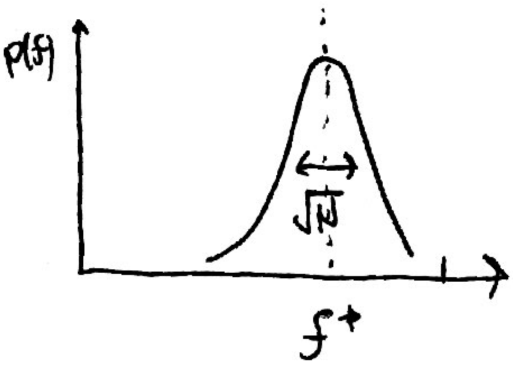
in this particular case, $-\frac{\partial \Lambda}{\partial f} = s + \frac{\mu}{f} - \frac{\nu}{1-f}$

so $\Lambda(f) = sf + \mu \log f + \nu \log(1-f)$

and $\boxed{p(f) \propto f^{N\mu-1}(1-f)^{N\nu-1} e^{2Nsf}}$ "mutation-selection-drift" balance (Wright 1930's)

What does dist'n look like? Strongly depends on $N\mu, N\nu$!

(a) if $N\mu, N\nu \gg 1$, then $p(f)$ is strongly peaked around some characteristic frequency $f^* \in (0,1)$ minimum of $\Lambda(f)$



$\Rightarrow \frac{\partial \Lambda}{\partial f} = 0 \Rightarrow s + \frac{\mu}{f} - \frac{\nu}{1-f} = 0$

note: same as deterministic solution to $\frac{df}{dt} = 0$.

$\Rightarrow$ "deterministic mutation-selection balance"

(or just mutation balance if $s=0$)

full distribution is expansion around f^*: $\to 0$ by def. of f^*

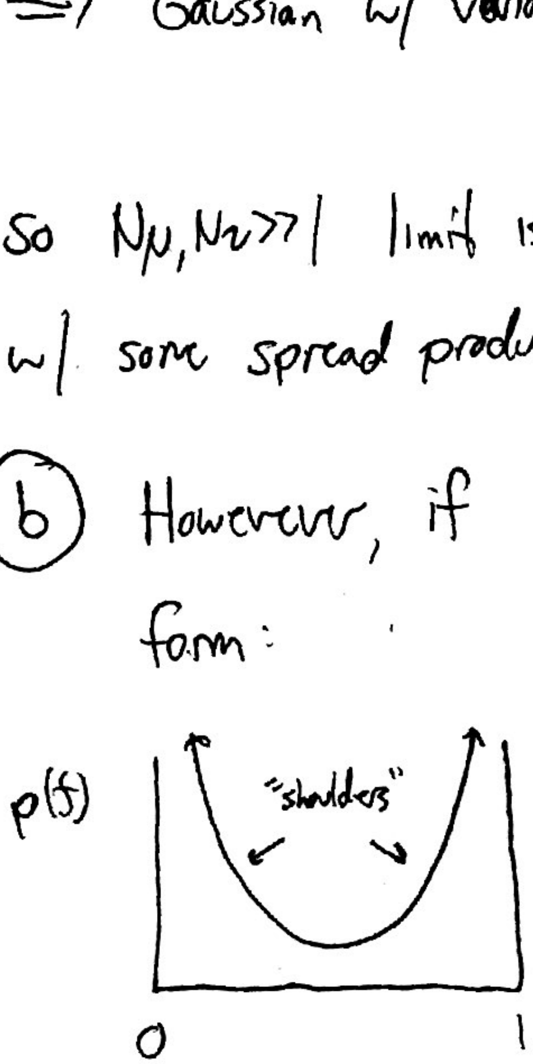
$$p(f) \propto \approx f^{*-1}(1-f^*)^{-1} \exp\Big[-2N\Lambda(f^*) + \cancel{2N \frac{d\Lambda(f^*)}{df}(f-f^*)} - \frac{2N}{2}\frac{d^2\Lambda(f^*)}{df^2}(f-f^*)^2 \Big]$$

$\Rightarrow$ Gaussian w/ variance $\propto \frac{1}{\sqrt{N}}$.

So $N\mu, N\nu \gg 1$ limit is standard situation of mostly deterministic, w/ some spread produced by noise.

b However, if $N\mu, N\nu < 1$, distn takes on "U-shaped" form:

where "height" of shoulders (roughly speaking) differ by factor of e^{2Nsf}

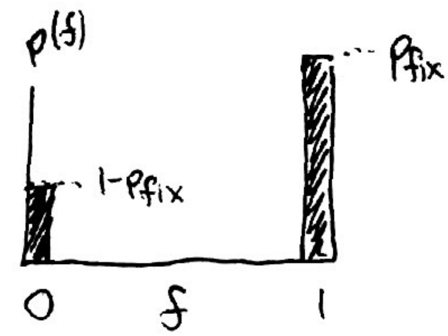


$\Rightarrow$ definitely not deterministic + (a little noise.) even if N itself is big!

what's going on here under the hood? what are "shoulders"?
how long does it take to reach this stationary state?
is it ever relevant in practice? (e.g. data?)

can gain a little more insight into these Qs by considering final stationary dist'n scenario:

<u>no mutation</u>: $\frac{df}{dt} = sf(1-f) + \sqrt{\frac{f(1-f)}{N}}\,\eta(t)$



$\Rightarrow$ in this case, 0 and 1 are *both* absorbing states, so $p(f)$ will be mixture: w/ weight $P_{fix} \equiv Pr(f=1)$ that depends on initial freq f_0 $\Rightarrow$ fundamentally out-of-equilibrium question (though posed in terms of eq. measurement)

$\Rightarrow$ recall when $s=0$ used trick that $\langle f(t)\rangle = \text{const}$ to show $P_{fix}(f_0) = f_0$

How does natural selection change this?

Unfortunately, Fokker-Planck eq doesn't work well for discrete dist'n (what does $\frac{dp}{df}$ mean?). But <u>generating function</u> is still useful:

$$H(z,t) \equiv \langle e^{-zf(t)} \rangle \equiv \int e^{-zf} p(f,t)\,df$$

Using same approach as we did for other moments, $\langle f(t)\rangle$, $\langle f(t)^2\rangle$, can work out equation of motion for $H(z,t)$:

$$H(z,t+\delta t) \equiv \left\langle e^{-zf(t+\delta t)} \right\rangle = \left\langle e^{-z\left[f(t) + sf(1-f)\delta t + \sqrt{\frac{f(1-f)\delta t}{N}}\, Z_t\right]} \right\rangle$$

$=$ talyor expand through $\mathcal{O}(\delta t)$ and avg over Z_t

$$= \underbrace{\left\langle e^{-zf(t)} \right\rangle}_{H(z,t)} + \left\langle e^{-zf} \left[-z\underbrace{sf(1-f)}_{\text{deterministic part}} + \frac{z^2}{2N} \underbrace{f(1-f)}_{\text{2 stochastic terms avg'd}} \right] \delta t \right\rangle$$

$$\Rightarrow \quad \frac{H(z,t+\delta t)-H(z,t)}{\delta t} = \left\langle -\left[sz - \frac{z^2}{2N} \right] \underbrace{f(1-f)e^{-zf}}_{\left[-\frac{\partial}{\partial z} + \frac{\partial^2}{\partial z^2}\right] e^{-zf}} \right\rangle$$

$$\Rightarrow \quad \frac{\partial H}{\partial t} = \left[sz - \frac{z^2}{2N} \right] \left[\frac{\partial H}{\partial z} - \frac{\partial^2 H}{\partial z^2} \right]$$

(can also get from Laplace transform of Fokker-Planck eq.)

still hard to solve... but for one particular value of z, very easy

E.g. let $z^* = 2Ns$. then

$$\frac{\partial H(z^*,t)}{\partial t} = \left[\cancelto{0}{s(2Ns) - \frac{(2Ns)^2}{2N}} \right] \left[\frac{\partial H}{\partial z} - \frac{\partial^2 H}{\partial z^2} \right] = 0$$

and hence $H(z^*,t) = \text{const} = H(z^*, t=0) = e^{-zf_0}$

(z^* is known as a characteristic curve - will see more later)

* this is really cool - allows us to connect initial condition w/ property of dist'n @ long times.

as in neutral case, @ $t \to \infty$: $f \begin{cases} \nearrow 0 & \text{w/ prob } 1-p_{fix} \\ \searrow 1 & \text{w/ prob } p_{fix} \end{cases}$

so $H(z^*, t=\infty) = \underbrace{e^{-z^* \cdot 0} \cdot (1-p_{fix}) + e^{-z^* \cdot 1} \cdot p_{fix}}_{\text{defintion of } H(z) \text{ @ } t=\infty} \overset{\uparrow}{=} \underbrace{e^{-zf^*}}_{\text{definition of } H(z) \text{ @ } t=0}$

characteristic curve.

$\Rightarrow$ solve for $p_{fix}(f_0)$ $\Rightarrow$ $\boxed{p_{fix}(f_0) = \dfrac{1-e^{-2Nsf_0}}{1-e^{-2Ns}}}$ "fixation probability" (Kimura 1950's)

Fixation prob is battle between selection & genetic drift.

(a) if $Ns \ll 1$ $\Rightarrow$ $p_{fix}(f_0) = f_0$ as before. (drift wins) ("weak selection") ("neutrality")

(b) if $Ns \gg 1$ ("strong selection")

$$p_{fix}(f_0) \approx \begin{cases} 1 & \text{if } f_0 \gg \frac{1}{2Ns}\ ;\ s>0 \longrightarrow \text{"selection wins"} \\ 2Nsf_0 & \text{if } f_0 \ll \frac{1}{2Ns}\ ;\ s>0 \longrightarrow \text{outcome uncertain...} \\ e^{-2N|s|(1-f_0)} & \text{if } s<0;\ \longrightarrow \approx 0 \text{ "selection wins"} \end{cases}$$

e.g if we extrapolate to new mutation ($f_0 = \frac{1}{N}$)

$\Rightarrow$ $\boxed{P_{fix} = 2s \quad \text{(independent of } N!)}$ "Haldane's formula" (Haldane, 1930's)

e.g. if $s = 0.01$ $\Rightarrow$ only 2% chance that mutation fixes
(i.e. pretty beneficial on lab timescales)
$\Rightarrow$ 98% of these mutations go extinct due to genetic drift.

$\Rightarrow$ but same mutant mixed @ 50-50 will rapidly, take over on lab timescales.
↳ & consistently

what's going on here?

@ least naively, as $N \to \infty$, we expect behavior to look like deterministic dynamics $\frac{df}{dt} = sf(1-f) \Rightarrow f(t) = \frac{f(0)e^{st}}{f(0)e^{st} + 1 - f(0)}$

+ small bit of fuzzyness due to noise.

How can we understand this?

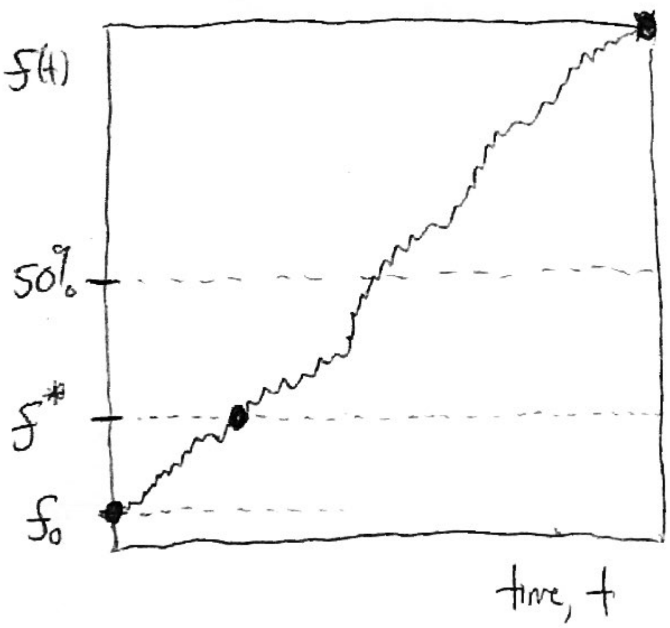
Our formula for fixation probability already contains clues:

$\Rightarrow$ e.g. we see that $p_{fix} \approx 1$ when $f_0 \gg \frac{1}{2Ns}$

even when f_0 itself is rare ($f_0 \ll 1$)

i.e. outcome only uncertain when f_0 gets as small as $\frac{1}{2Ns}$ ($\ll 1$ if $Ns \gg 1$)

$\Rightarrow$ can go one step further by breaking fixation probability into two components: before + after reaching some intermediate freq f^*



Since all paths must pass through f^* on way to fixation, we have

$$\Pr[f_0 \to 1] = \Pr[f_0 \to f^* \text{ before } f_0 \to 0] \times \Pr[f^* \to 1]$$

$\Rightarrow$ or $\Pr[f_0 \to f^* \text{ before } f_0 \to 0] = \dfrac{2Ns f_0}{\Pr[f^* \to 1]}$ $\left(\text{for } f_0 \ll \frac{1}{2Ns f_0}\right)$

if $f^* \gg \frac{1}{2Ns}$ then $\Pr[f^* \to 1]$, and all uncertainty in mutation's fate takes place between $0 \leq f \leq f^* \ll 1$

i.e. "selection wins" when $f(t) \gg \frac{1}{2Ns}$

$\Rightarrow$ on the other hand, if $f^* \ll \frac{1}{2Ns}$, then

$$Pr\left[f_0 \to f^* \text{ before } f_0 \to 0 \right] = \frac{2Nsf_0}{2Nsf^*} = \left(\frac{f_0}{f^*} \right)$$

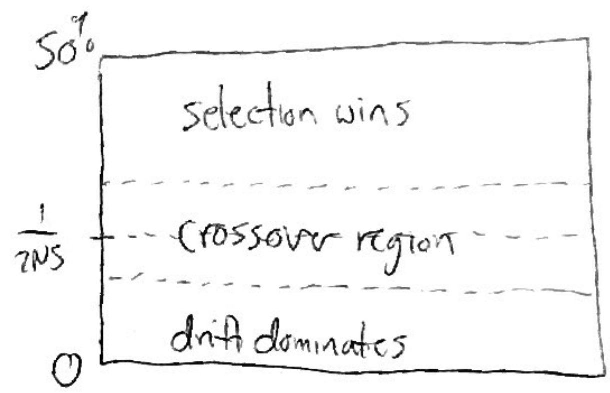
independent of selection strength!

this result can also be derived from a symmetry argument

$\Rightarrow$ in other words, looks like neutral mutation for $f(t) \ll \frac{1}{2Ns} \ll 1$

"genetic drift dominates"

$\Rightarrow$ suggests interesting partitioning of frequency space:



this shows why it's not possible to just drop noise term in $N \to \infty$. Although selection dominates in ever greater region of frequency space, always a narrow "boundary layer" below $\frac{1}{Ns}$ where noise is dominant factor!

$\Rightarrow$ important for evolution, since new mutations typically enter @ $\frac{1}{N} \ll \frac{1}{Ns}$

Fortunately, this analysis suggests that when $Ns \gg 1$, we can gain a complete picture of what's going on by focusing on $f \ll 1$ limit (since $\frac{1}{2Ns} \ll 1$)

$\Rightarrow$ then, once $f \gg \frac{1}{2Ns}$ (but still $\ll 1$), we can patch back on to the deterministic limit, $\frac{\partial f}{\partial t} = sf(1-f)$

$\Rightarrow$ this approach is known as "asymptotic matching"
it is a powerful method that works whenever you have 2 approx's that agree in overlap region
(in this case, $\frac{1}{2Ns} \ll f \ll 1$)

When $f \ll 1$, single-locus model reduces to

$$\frac{\partial f}{\partial t} = sf + \sqrt{\frac{f}{N}}\,\eta(t)$$

$\boxed{\text{known as "linear branching process"}}$

($+\mu$ for forward mutation, WT $\to$ mut)
($-\nu f$ for back-mutation, mut $\to$ WT)

(technically, continuous-time & continuous state B.P.)

$\Rightarrow$ turns out that this process is simple enough that can get complete picture of dynamics as well as stationary quantities like P_{fix}, $p(f)$

~~$\Rightarrow$ this gives lots of intuition for evolutionary patterns~~,

$\Rightarrow$ understanding these dynamics will give us lots of intuition for what's going on in evolutionary problems, and they will be a natural starting point when we start to consider more complicated scenarios later in the course.

(also increasingly relevant for analyizing any kind of longitudinal data, e.g. lab expts, ancient DNA, etc.)

$\Rightarrow$ we will take a deeper dive into these dynamics in next few lectures.