Heuristics

In this lecture, we will present a powerful heuristic approach for deriving many of the exact results we have discussed so far

$\Rightarrow$ may seem sloppy or arbitrary @ first, but w/ practice, can be done in way that keeps track of approximations in controlled manner, while highlighting key physical intuition.

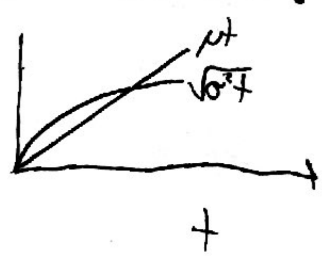
$\Rightarrow$ enables progress in more complicated settings where exact results are not available.

Start by returning to Gaussian random walk:

$$\partial_t x = \mu + \sqrt{\sigma^2}\,\eta(t) \implies x(t) = \mu t + \sqrt{\sigma^2 t}\, Z$$

$\Rightarrow$ when are stochastic vs. deterministic effects dominant?

* since deterministic contribution $\propto t$
stochastic contribution $\propto \sqrt{t}$



$\Rightarrow$ stochastic term always dominant at short t.
deterministic term always dominant @ long t.

$\Rightarrow$ crossover @ $\mu t^* = \sqrt{\sigma^2 t^*} \implies t^* = \dfrac{\sigma^2}{\mu^2}$

- $t \gg t^*$ deterministic ($x \approx \mu t$)
- $t \ll t^*$ stochastic ($x \sim \sqrt{\sigma^2 t}\, Z$)

Now we return to our evolution problem:

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

$\Rightarrow$ can't apply same approach because <u>det</u> and <u>stoch</u> terms both depend on $f(t)$, which influenced by <u>det</u> and <u>stoch</u> terms, etc, etc.

$\Rightarrow$ need to integrate SDE (moment eqs, gen funcs, etc.) $\Rightarrow$ Hard!

Heuristics $\approx$ way to do this approximately $\approx$ "poor man's integration"

or

"Euler's method for analytical sol'ns"

<u>Idea</u>: if interested in logarithmic precision [i.e., $\log(x(t)) \pm \Theta(1)$]

short time approx $f(\Delta t) = f(0) + sf(0)\Delta t + \sqrt{\frac{f(0)\Delta t}{N}}\, Z$

works pretty well until $\log f(\Delta t) \simeq \log(f(0)) \pm \Theta(1)$, since this is when Δf_{sel} & Δf_{drift} start to deviate by $\Theta(1)$ factors.

$\Rightarrow$ call this time Δt_{reset}. occurs when $\log(\Delta x) \approx \log x \pm \Theta(1)$

[" $\Delta x \sim x$ "]

At this point, set $f(0) = f(\Delta t_{reset})$ and repeat entire process, ...

$\Rightarrow$ iterative method for building up $f(t)$ for $t \gg \Delta t_{reset}$.

<u>Question then becomes</u>: Are deterministic forces (selection) or stochastic forces (drift) dominant on timescales $\sim \Delta t_{reset}$?

- <u>Approach</u>: guess & check (self consistency)

1. if deterministic forces dominant ($\Delta f_{sel} \gg \Delta f_{drift}$),

   must have $f \sim |\Delta f_{sel}| \sim |s| f \Delta t_{reset} \Rightarrow \Delta t_{reset} \sim T_{sel} = \frac{1}{|S|}$

   (really, $\Delta t_{reset} \approx \frac{c_1}{|S|}$ for $\Theta(1)$ const c_1)

   on this timescale, contribution from drift is

   $|\Delta f_{drift}| \sim \sqrt{\frac{f \Delta t_{reset}}{N}} \sim \sqrt{\frac{f}{N|S|}} \Rightarrow |\Delta f_{drift}| \ll |\Delta f_{sel}| \sim f$ when $\boxed{f \gg 1/N|S|}$

   selection dominant

   After k resets, have $\Theta(1)$ const. c

   $\log f(t) \approx \log f(0) + c_2 k \simeq \log f(0) + c \cdot s \cdot t$

   $\Rightarrow f(t)$ grows exponentially @ rate $\Theta(s)$.

2. If stochastic forces dominant ($\Delta f_{drift} \gg \Delta f_{sel}$) then

   $f \sim |\Delta f_{drift}| \sim \sqrt{\frac{f \Delta t_{reset}}{N}} \Rightarrow \Delta t_{reset} \sim T_{drift} = Nf$.

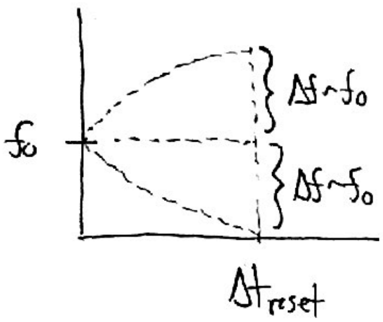
   contribution from selection on same timescale is $\Delta f_{sel} \sim NSf^2$

so $|\Delta f_{sel}| \ll |\Delta f_{drift}|$ when $\boxed{f \ll 1/N|s|}$ (drift dominates)

$\Rightarrow$ in this case, behavior is not as simple as unbiased random walk since diffusion coefficient depends on $f(t)$.

$\Rightarrow$ but can still understand behavior by gluing together several iterated random walks.

start w/ freq f_0



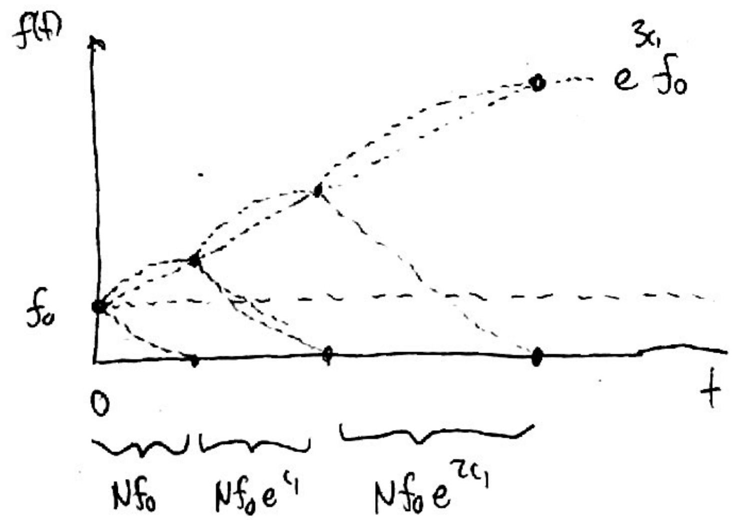
After Δt_{reset} gens, $f(t) \approx f_0 \pm f_0$

$\Rightarrow$ decent chance of going extinct!

w/ prob $\approx e^{-c_1}$ [$c_1 \to O(1)$ factor] [e.g. $1/2$]

mutation is not extinct and must have size $f \approx f_0/e^{-c_1} = e^{c_1} f_0$

then process repeats itself starting from $f(0) = e^{c_1} f_0$:



can see that after k iterations:

* probability of survival is $p_{survival} \simeq e^{-c_1 k}$
* size is $f(t) \approx f_0 e^{c_1 k}$
* total time elapsed is

$$t \approx Nf_0 + Nf_0 e^{c_1} + \ldots + Nf_0 e^{c_1 k} \approx \frac{e^{c_1 k} - 1}{e^{c_1} - 1}$$

$$\sim Nf_0 e^{c_1 k} \quad (k \gg 1)$$

Rewriting in terms of t: * survival probability is $\approx Nf_0/t \cdot c$

* size is $f(t) \approx \frac{c \cdot t}{N}$

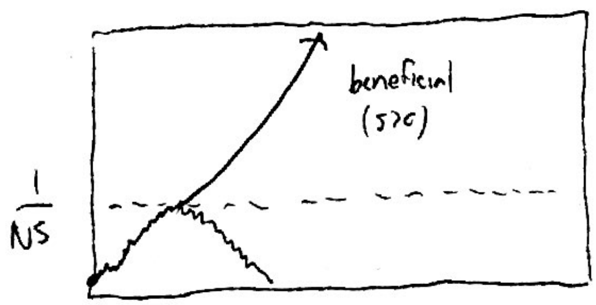
$\Rightarrow$ i.e. w/ probability $\sim \frac{Nf_0}{t}$, survives for t gens & reaches size $\sim t/N$

$\Rightarrow$ alternatively, in terms of final size $f(t) = f$:

w/ probability f_0/f, drifts to size $\geq f$ on timescale $t \sim Nf$ gens.

* Heuristic approach pretends that division between drift dominated & sel dominated is infinitely sharp, and can <u>patch</u> 2 regimes together (note: $\neq$ asymptotic matching from before)

$\Rightarrow$ incurs $\Theta(1)$ errors in $\log f(t)$ & t, but that's w/in our tolerance anyway.



selection dominated [grows exponentially w/ rate $\Theta(s)$]

drift dominated [w/ prob $\approx \frac{Nf_0}{t}$, survives for t gens & reaches size $\sim t/N$]

① For beneficial mut ($s>0$), drifts to size $\sim 1/Ns$ w/ prob $\sim \frac{1/N}{1/Ns} \sim s$, takes $1/s$ gens to do so.

$\Rightarrow$ then grows exponentially @ rate $\sim s$.

② deleterious mut ($s<0$), drifts to size $\sim 1/N|s|$ w/ prob $\sim |s|$, but can't grow any higher

$\Rightarrow$ prob of surviving another $1/|s|$ gens is $\sim e^{-c}$ $\Rightarrow$ $p_{survive}(t) \sim |s| e^{-c \cdot s \cdot t} \to 0$.

③ Neutral mutations look like triangles w/ height $\frac{t}{N}$, width t, w/ prob $p(t) \propto 1/t^2$