Gradient Dynamics and Speciation

Daniel Friedman
Economics Department
University of California, Santa Cruz

Theoretical and Experimental Evolutionary Biology Seminar
Oxford 5/6/16

How do new species begin?

Consequently, I cannot doubt that in the course of many thousands of generations, the most distinct varieties of any one species [...] would always have the best chance of succeeding and of increasing in numbers, and thus of supplanting the less distinct varieties; and varieties, when rendered very distinct from each other, take the rank of species. (Darwin 1859, p. 155)

- Darwin famously argued that all species, from aardvarks to zynnias, evolved from a common ancestor...
- But how, exactly? Darwin had only partial answers, and we still lack adequate models...
- I'll quickly review a 20 year-old approach, and suggest a new direction.

Charles Darwin, Sewall Wright, and sympatry

- Allopatry: geographically subdivided populations build up genetic differences (e.g., Ernst Mayr, 1985; featured in Pitt-Rivers).
- Darwin (1859) hinted at a more endogenous process of speciation, now called sympatry.
- Wright's shifting balance theory (1932, 1982): escaping local fitness maxima to reach unoccupied niches.
- True sympatry requires both
  - (a) disruptive (frequency-dependent?) selection to generate distinct phenotypes, and
  - (b) some force to swamp hybridization, e.g., assortative mating mechanisms.

  Today's discussion focuses on (a)

Sticklebacks

Figure: Limnetic and benthic morphs

Long run evolution and continuous trait space

- SR evolution: replicator dynamics for a specified finite list of alternative traits. E.g., FS16 (OUP 3/16, for applied researchers in all fields).
- In LR, new varieties enter via mutations, and even small mutations accumulate.
- Examples: foot speed in terrestrial animals, and beak size or migration dates of birds.
- My goal: to model when and how subpopulations might diverge, in terms of a single continuous trait.

Adaptive Dynamics: Background

In the 1990s, Hans Metz, Ulf Diekmann, et al. developed calculus-based phenotypic models of frequency-dependent evolution, esp. speciation (a). My list of their key assumptions:

0. 1-D trait space $A \subset (-\infty, \infty)$, LR dynamics via fitness gradient.

1. For a.e. $t \in [0, \infty)$ there is a normally distributed trait distribution with constant variance & variable mean $r(t) \in A$ ("resident" trait).

2. A twice continuously differentiable function $f(m, r)$ gives the fitness of any nearby "mutant" trait $m \in A$.

2a. Fitness is LR expected growth rate, so it is natural to normalize so that $f(r, r) = 0$ for all $r \in A$.

3. If $f(m, r) > 0$, then that mutant (initially with an infinitesimal population share) takes over (via unmodelled SR dynamics).

Adaptive Dynamics analysis

By Taylor's theorem, for m near r,

$$f(m, r) \approx f(r, r) + (m - r)\frac{\partial f(m, r)}{\partial m}|_{m=r} = (m - r)g(r), \qquad (1)$$

where $g(r) \equiv f_1(r, r)$ is called the selection gradient.

- If $g(r) > 0$ (resp. < 0) then $f(m, r) \approx (m - r)g(r) > 0 = f(r, r)$ for $m > r$ (resp. $m < r$).
- Thus LR interior steady states, where nearby mutants can't invade, occur only at critical points, i.e., at roots r^* of g.

"Canonical" Equation for Adaptive Dynamics

LR dynamics governed by selection gradient:

$$\dot{r} = kg(r). \tag{2}$$

- Thus the resident population evolves according to the selection gradient at a rate $k > 0$.
- Standard biological theory says that k is proportional to the heritability of the trait, to $1/$age to maturity, and to the (constant) variance around the mean trait value $r(t)$.

Evolutionary Stability and Convergence Stability

- Def: The critical point r^* is ES if $0 > f_{11}(r^*, r^*) \equiv \frac{\partial^2 f(m,r)}{\partial m^2}|_{m=r=r^*}$.

- Then r^* is a local fitness maximum, hence is dynamically stable in SR because immune to invasion by nearby mutant traits.

- Def: The critical point r^* is CS if
  $0 > g'(r^*) = \frac{df_1}{dr}|_{m=r=r^*} \equiv f_{11}(r^*, r^*) + f_{12}(r^*, r^*)$.

- That is, the *total* derivative of the selection gradient, not a first partial derivative, is negative. LR dynamic stability condition.

- CS $\implies$ given r near r^*, mutation m can invade only if it is closer than r to r^*.

- An equivalent condition for CS is $f_{11}(r^*, r^*) < f_{22}(r^*, r^*)$. Hint: twice differentiate the normalization identity...

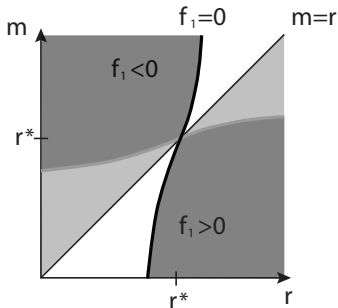# Pairwise Invasion Plots for two CS critical points r^*
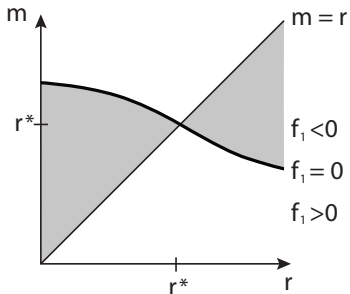
Figure: Fitness is positive in shaded regions. **Panel A**: CS since mutations $m > r^*$ have positive fitness only when below the diagonal (and mutations $m < r^*$ have positive fitness only when above the diagonal) so mutations succeed only when they lie between r and r^*. Also ES, since vertical line thru r^* doesn't intersect shaded area. **Panel B**: Also CS since a mutation is successful only when it is closer than r to r^* (it may lie on the opposite side of r^*). But not ES since r^* is a local min, not max, on its vertical line.

CSS and Branch Points

Classify fitness functions in a neighborhood of a critical point by signs and $|\cdot|$ ranks of pure second partials $f_{11}(r^*, r^*)$, $f_{22}(r^*, r^*)$. Of 8 subquadrants,

- 3 satisfy $f_{11} < 0, f_{22} \implies$ both ES and CS. Called CSS (Eshel, 1983).
- Our interest is in the subquadrant $0 < f_{11} < f_{22} \implies$ CS but not ES, as in Panel B. Called branch points.
- A m in Panel B shaded region satisfies $f(m, r) > 0$ so it can invade resident level r,
- In darker shaded region (obtained by reflecting the $f_1 = 0$ locus across the diagonal) it is also true that $f(r, m) > 0$ so trait level r can invade when the resident level is m.
- Interpretation (some hand-waving): Since r^* is not ES, successive mutations split the resident level r^* into two different protospecies, which should move apart.

A Thought Experiment

A remote island initially has a single species of seed-eating birds and several species of seed bearing plants. The birds' beak size is well-adapted to eating mid-range seeds but not especially efficient for the largest and toughest seeds or the smallest.

- Suppose climate change reduces availability of mid-range seeds, and
- this pushes intermediate size beak from a CSS to a BP.
- Selective pressure encourages separate subpopulations of large beaked and of small beaked birds.

Two Finch Morphs

Polynomial example

- Let $f(m, r) = 2m^2 + 3r^2 - 5mr$.
  - Normalization $f(r, r) = 0$ already imposed.
  - Here $g(r) = f_1(r, r) = 4r - 5r = -r$.
- So $r^* = 0$ is a critical point. Easy to check that it is a branch point:
- we have $0 < f_{11} = 4 < 6 = f_{22}$.
- Beef up the example later by adding higher order terms $r^4 - m^4$.

Gaps in Adaptive Dynamics analysis

- Normality clearly fails at branch points.
- Justification for "canonical" equation then breaks down.
- Indeed, meaning of fitness function not clear in neighborhood of branch point.
- Not clear how to best to proceed.
  - Metz (e.g., 2012) writes $f(m, r_1, r_2)$, but that only deals with the last point.
  - Dieckmann and Doebeli (1999) run highly parametrized stochastic agent-based sims.
  - Cressman and Hofbauer (2005) try replicator.
- Is there some way to retain gradient dynamics?

Fitness landscapes

- Biologists from Sewall Wright to Stuart Kauffman (1993) invite us to visualize the graph of a fitness function as a landscape.
- With a continuous trait, evolution pushes the population uphill.
- My contribution: the landscape is not static.
    - Fitness of any trait value is "frequency dependent,"
    - affected by the current distribution of trait values.
    - As the population moves uphill in the landscape, the distribution changes, so
    - the landscape shape-shifts, producing non-linear dynamics.
- A landscape approach allows us to do gradient dynamics properly, and drop restrictive assumptions 1,2,3 of AD
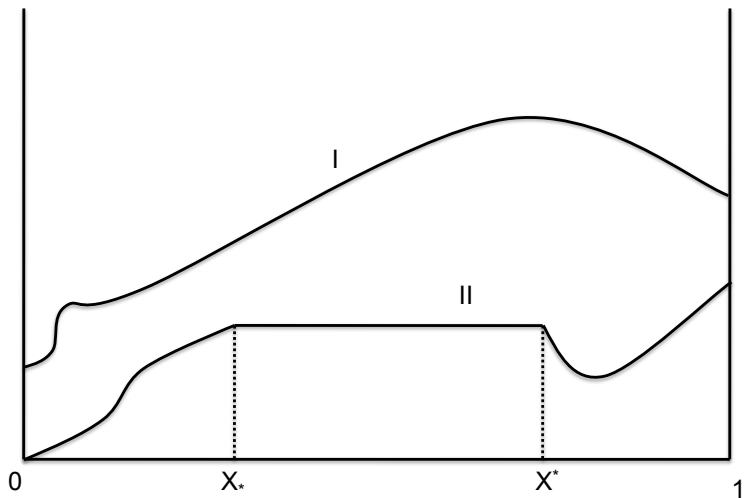- (but keep 0, 2a).

Two Landscapes over $A = [0, 1]$

Figure: Landscape I might eventually morph into Landscape II.

Trait distributions and fitness functions

Drop the artificial distinction between m, r and allow for traits x in some natural interval $A \subset (-\infty, \infty) \equiv \mathbb{R}$.

- The current state of a population is described by a cdf $F : A \to [0, 1]$.
- F need not be Normal or even have connected support, but of course it is surjective, monotone and right-continuous: in $\mathcal{F}$.
- Fitness function is $\phi : [0, \infty) \times A \times \mathcal{F} \to \mathbb{R}$.
- Landscape is graph of $x \mapsto \phi(t, x, F)$, for fixed time t and current distribution F.
- Evolving state is $F(t, \cdot) \in \mathcal{F}$; the landscape co-evolves.
- Now denote densities (when they exist) by $f = F_x$.

Gradient dynamics

- Evolution pushes population towards higher fitness, with speed proportional to steepness:

$$\dot{x} = V(x, t) = \phi_x(x, F(t, \cdot)). \tag{3}$$

- With a small random component, we have $dx = V dt + \sigma dB$.

- Then the Fokker-Planck-Kolmogorov equation (conservation of population mass) yields

$$f_t = -(Vf)_x + \frac{1}{2}\sigma^2 f_{xx}. \tag{4}$$

Gradient adjustment equation

- We focus on the $\sigma \searrow 0$ limit,

$$f_t = -(Vf)_x, \tag{5}$$

  or, after integrating from $-\infty$ to x,

$$F_t = -VF_x. \tag{6}$$

- Conservation of population mass with no jumps; aka "continuity equation" in fluid dynamics.
- Here it simply captures the standard notion of selection gradient acting on a *general* population distribution.

Solving the Continuity Equation

- Given an initial trait distribution $F(0, \cdot) = F^0 \in \mathcal{F}$,
- we seek a trajectory $[F(t, \cdot), t \geq 0] \subset \mathcal{F}$ that satisfies the continuity equation (6).
- Of particular interest are fitness functions ϕ such that
  - a unimodal F^0 with connected support
  - converges to a bimodal limit with disconnected support.
 - This would capture speciation, part (a).

Connection to Adaptive Dynamics

- Piecewise smooth fn $h : A \times A \to \mathbb{R}$ is fitness of trait level x when interacting with individuals of trait level y. (denoted f previously; apologies)

- Population state at time $t \in [0, \infty)$ is cdf $F(t, y)$,

- so overall fitness of trait level x is

$$\phi(x, F(t, \cdot)) = E_F h = \int_A h(x, y) f(t, y) dy. \qquad (7)$$

(Use Stieltjes integral if the density $f = F_x$ doesn't exist.)

- Reduces to basic AD fitness function when F is $\delta_{y=r(t)}$.

- So (7) says that fitness is a weighted avg of AD fitnesses across y's.

Graph of polynomial example
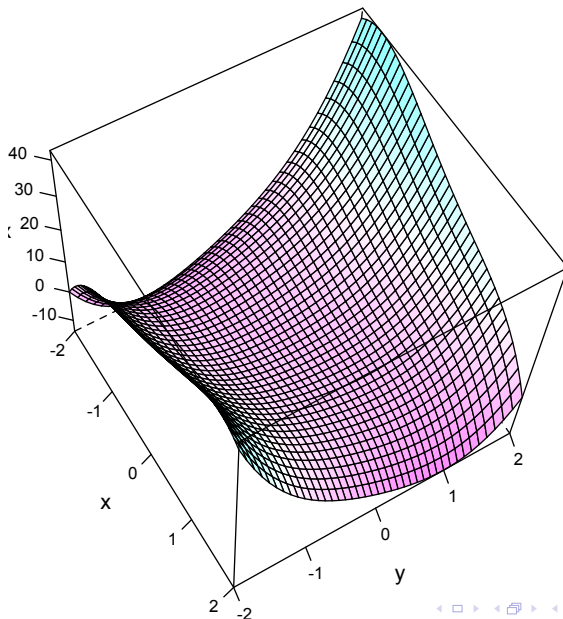
Polynomial example, revisited

- Recall $h(x, y) = 2x^2 + 3y^2 - 5xy + y^4 - x^4$; suppose $[-1, 1] \subset A$.
- Then $V = \phi_x = \int_A h_x(x, y) f(t, y) dy = 4x(1 - x^2) - 5\mu_F(t)$, where $\mu_F(t)$ is the mean trait at time t.
- Let the initial distribution F^o be symmetric around 0 and continuous at 0. Evidently $\mu_F(t) = 0$, and the PDE (6) becomes

$$0 = F_t + 4x(1 - x^2)F_x. \tag{8}$$

Solving the continuity equation

Theorem 1 of FO13 ensures that, given a nice initial density (bounded first derivative), the solution to (8) exists and has nice density for all finite time. To find explicit solution, use method of characteristics.

$$\frac{d\xi(t)}{dt} = 4x(1-x^2), \qquad\qquad \xi(0) = x_o \qquad (9)$$

$$\frac{dF(t,\xi(t))}{dt} = F_t + F_x\frac{d\xi(t)}{dt}, \qquad F(0,x_o) = F^o(x_o). \quad (10)$$

Solution to (9) is obtained via separation of variables, partial fractions, etc. It turns out to be

$$\xi(t) = \pm\sqrt{\frac{x_o^2}{x_o^2 - e^{-8t}(x_o^2-1)}}.$$

$F(t,x)$ is constant along a ch curve, so rearranging and using (10),

$$F(t,x) = F^o(\frac{x}{+\sqrt{x^2 + (1-x^2)e^{8t}}}). \qquad (11)$$

Characteristic curves for (8)

no characteristics!
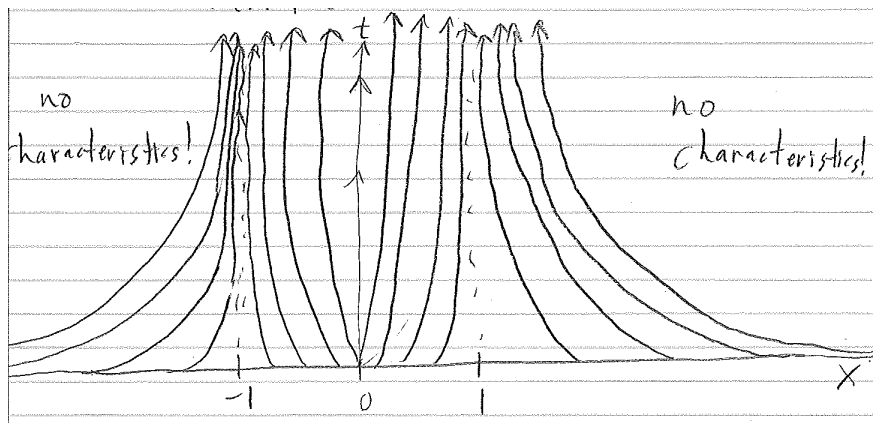
no characteristics!

t

-1 0 1

x

Figure: Curves cluster at $x = \pm 1$ as $t \to +\infty$.

Polynomial example, solved.

- Thus, for any initial condition F^o symmetric around 0, as $t \to \infty$, we see that $F(x, t) \to 0, 0.5$, or 1 for $x < -1, x \in [1, 1]$, or $x > 1$.
- Thus the distribution converges under gradient dynamics (8) to equal point masses at $x = -1$ and $x = +1$.
- With a little noise, these would be tight Normal distributions, as envisioned in AD.
- True speciation if gene pools don't hybridize.
- For an arbitrary initial distribution, it seems that qualitatively similar behavior occurs; the two cluster points might drift a bit while the mean settles down.

Finch species
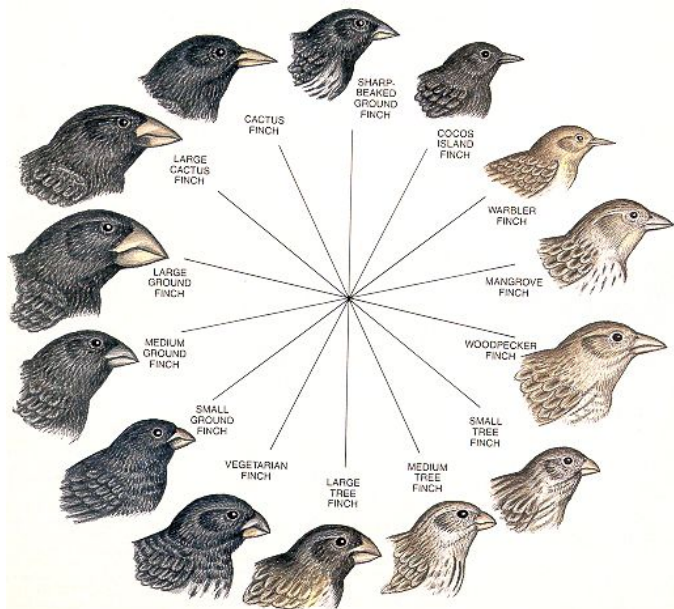
More general examples

- The same sort of analysis could be applied to any other $h(x, y)$.

- Say trait is preferred altitude, and fitness depends mainly on direct competition, so $h(x, y) = 0$ except near the diagonal $x = y$.

- Indeed, the fitness function need not be a population average of pairwise fitness $h(x, y)$, but could be defined directly.

- For example, $\phi(x, F)$ might be local — depending only on F of $F_x = f$ in a neighborhood of x.

- E.g., benefit of x is position in dominance hierarchy, but it has a convex cost, so $\phi(x, F) = F(x) - c(x)$.

- Or ϕ might depend on F via a moment such as μ_F or σ_F, or an order statistic, or the mode.

Solving more general models

In all these cases, one proceeds in the same fashion:

- Find the gradient $V = \phi_x$
- Write out the PDE

$$
\begin{aligned}
f_t &= -(Vf)_x + \frac{1}{2}\sigma^2 f_{xx} \text{ , or} & (12)\\
f_t &= -(Vf)_x \text{ , or} & (13)\\
F_t &= -VF_x & (14)
\end{aligned}
$$

- Use techniques from fluid mechanics to solve, e.g., method of characteristics.

Caveats and Comments

- Ten years ago I thought that ODEs on space of (prob) measures would work just as well.
  - ▶ OK in interior (densities with full support), but interesting solutions hit edges or corners, often of infinite codimension.
  - ▶ Fortunately, fluid dynamics research developed very helpful PDE tricks over the last century, e.g., Rankine-Hugoniot conditions.
- Five years ago I thought that weak-star topology was appropriate for studying convergence.
  - ▶ Especially in light of weak solutions, needed where F isn't differentiable.

Weak Solutions

- In interesting cases, F might have kinks or discontinuities, due to disconnected support or point masses. How to interpret PDE ?
- Return to the FPK equation. For viscosity (or volatility) $\sigma > 0$, it is known to have a unique solution F^σ.
- Define ("vanishing viscosity" or "entropy") solution as $\lim_{\sigma \to 0} F^\sigma$.
- It is a "weak" solution of the PDE: for every smooth test function $\psi(t, x)$ with compact support,

$$\int_0^\infty \int_A f\psi_t + Vf\psi_x dx dt = 0 \tag{15}$$

(integration by parts shifts derivatives from f or F to ψ.)

When characteristics collide

- Characteristics nicely foliate the (t, x)-plane in the classical case.
- But not necessarily so when flux $H = VF$ is non-linear.
- E.g., $H(t, x, p) = p(1 - p)$ arises naturally from homophily. Here slope of characteristics is $1 - 2p = 1 - 2F(t, \xi(t))$, and F is constant along each ξ. Thus they are straight lines.
- Let $F(0, \cdot)$ have support $[a, b]$. Then the line emanating from $(t, x) = (0, a)$ has slope $1 - 2 \cdot 0 = 1$, while the line emanating from $(t, x) = (0, b)$ has slope $1 - 2 \cdot 1 = -1$. They must collide at time $t = (b - a)/2$, trapping all mass in between.
- Collisions imply discontinuities: the values of F on the two characteristics generally differ.
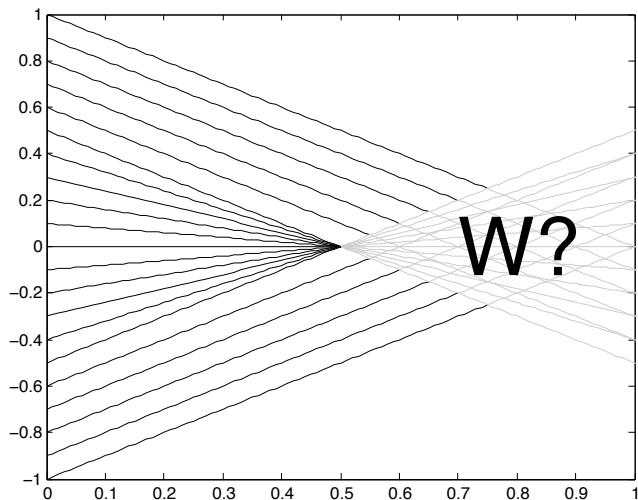
A collision

Figure: Characteristics for gradient adjustment equation for homophilic ϕ

R-H to the rescue

- When characteristics for the weak solution collide, they do so along a "shock curve" $s(t)$.

- It satisfies the Rankine-Hugoniot jump condition

$$\frac{ds}{dt} = \lim_{\varepsilon \to 0^+} \frac{H(t, s(t), f(t, s(t) + \varepsilon)) - H(t, s(t), f(t, s(t) - \varepsilon))}{f(t, s(t) + \varepsilon) - f(t, s(t) - \varepsilon)}. \tag{16}$$

- A second condition, called "entropy," guarantees uniqueness: characteristics can terminate on $[t, s(t)]$ (destroying information) but never emanate from it.

- In homophilic example (with uniform initial distribution on $[-\frac{1}{2}, \frac{1}{2}]$), the full solution is uniform on $[-\frac{1}{2} + t, \frac{1}{2} - t]$ for $t < \frac{1}{2}$ and is Θ_0 thereafter.

Characteristic and shock curves for previous example
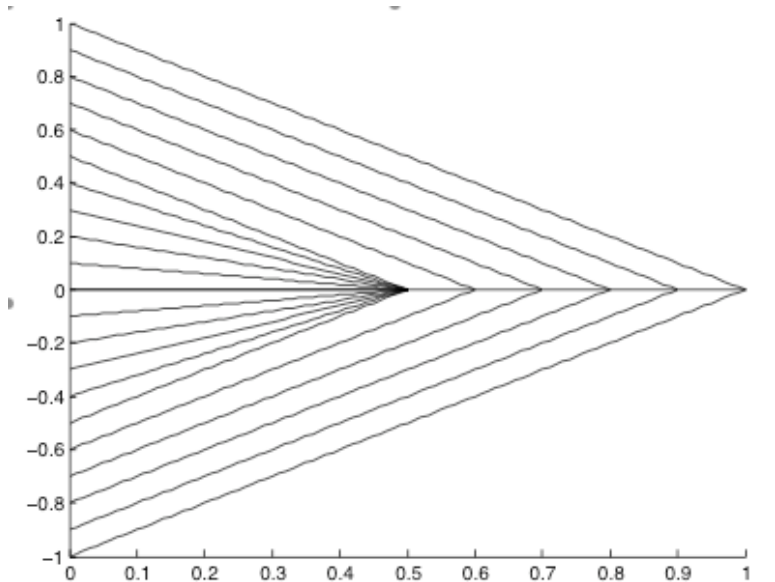
Summary

- Sympatric speciation occurs when frequency-dependent speciation splits the trait distribution into two (or more) components, which eventually become separate gene pools.
- Darwin, and especially S. Wright, discussed processes and examples, but wrote no formal models.
- Metz, Dieckmann et al wrote static calculus models, and invoked gradient dynamics, but key pieces are missing.
- Friedman and Ostrov (2009, 2013) develop gradient or landscape dynamics exploiting fluid dynamics PDEs.
- Those papers focussed on econ applications, but the ideas seem to work at least as well for biological evolution.
- E.g., "dynamic insufficiency" is no problem for landscape dynamics.

Ideas to take home

- Fitness landscapes coevolve with trait distributions.
- (in FO papers...) Equilibrium landscapes F^* are like the Southwest: plateaus and peaks above $supp(F^*)$.
- Some fitness functions ϕ naturally lead to disconnected limit distributions, i.e., to speciation.
- Fluid dynamics is well developed for multi-dimensional traits and ideas also could be extended to deal with coevolving populations.
- I hope to work with biologists to develop a killer app.
- Thank you for your time and thoughts!