

# Coevolution under constraint, and the special role of mutualism

Lee Worden

## Introduction

To be written.

Use of the four types (quadrants) of Lotka-Volterra interactions to classify ecological relationships, interest in the causes of mutualism and antagonism

Natural history examples: emergence of mutualism, and loss of mutualism

Evolution of constrained phenotypes; connection to Levins fitness sets

Coevolutionary dynamics, frequency-dependent selection

Related questions about cooperation, incentives, and externalities

## Sidebar: Coevolution of constrained characters

When there are two or more phenotypic characters, or ecological characteristics of a population such as efficiency of resource uptake, that are in fact a function of a smaller number of “underlying” characters, the population is constrained to a subset of the space of characters. See figure 1.

We consider the selection gradient describing motion in the constrained space and an abstract selection gradient giving the direction of motion that would obtain if all variation in the unconstrained space were available (see appendix for details).

The constrained selection gradient is at an acute angle to the abstract selection gradient. That is, it must always be located within the half-space of vectors less than  $90^\circ$  from the abstract gradient vector.

At any fixed point of adaptation, either the abstract gradient vanishes, or all motion in the constraint space is  $90^\circ$  or more from the abstract gradient.

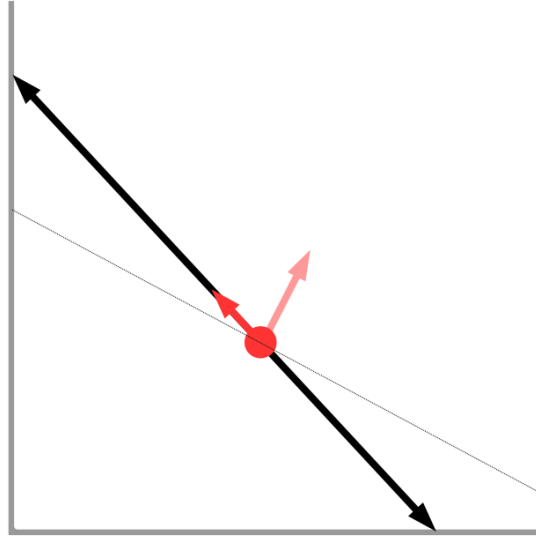


Figure 1: Evolution of two characters of a single population, constrained to a one-dimensional curve (the heavy black link). Direction of change in one dimension (constrained gradient of selection) is in the direction closer to the abstract gradient of selection in two dimensions. In the terms used in the Appendix, this figure shows the population's phenotype  $\mathbf{p}(\mathbf{u})$  (red dot), the unconstrained gradient of selection on that phenotype  $\mathbf{S}(\mathbf{p})$  (pink arrow), and the motion of  $\mathbf{p}$  when constrained to a one-dimensional set of possibilities (red arrow). The set of values available to  $\mathbf{p}(\mathbf{u})$  is the heavy black line, and the half-space of directions of motion that are selectively possible is the half-plane above the thin dotted line. The red arrow is the only possible direction of resulting motion of  $\mathbf{p}(\mathbf{u})$ .

## Coevolution of interaction terms

Lotka-Volterra interaction terms,  $a_{ij}$  for each pair of populations  $i$  and  $j$ , can be treated similarly to phenotypic characters as above, but parametrized by the phenotypes of two populations at a time, rather than one.

That is, each interaction term  $a_{ij}$  is a function of the characters of populations  $i$  and  $j$ :  $a_{ij} = a(u_i, u_j)$ . In this expression, population  $i$  is the **patient** – the entity affected by the interaction – and population  $j$  is the **agent** – the entity who affects the other.

Adaptive change in population  $i$  is predicted by gradient of selection on  $a_{ij}$  terms for which  $i$  is the patient. (See appendices for detailed treatment of all this.)

The “direct effect” of selection on  $a_{ij}$  is change due to selection on population  $i$ .

There is also the “indirect effect”: change in each interaction  $a_{ij}$  due to selection on the agent  $j$ .

Because of the nature of the Lotka-Volterra equations, the abstract selection gradient  $S$  is always positive because the  $j$ th entry of the  $S$  vector is the mean size  $\hat{X}_j$  of population  $j$ .

The direct effect vector  $D$  must be within  $90^\circ$  of that positive vector  $S$ . Therefore the entries of  $D$  are on average positive (though they need not all be positive).

The actual motion of each row  $(a_{i1}, \dots, a_{in})$  of the Lotka-Volterra matrix is given by the sum  $\mathbf{D}_i + \mathbf{I}_i$ . Assuming  $\mathbf{D}_i$  is biased toward positive and  $\mathbf{I}_i$  is unbiased, coevolution is biased toward motion in the positive direction, i.e. toward increase of mutual benefit and decrease of harmful interactions. This assumption may not hold however, see below.

Figure: Evolution of  $(a_{ij}, a_{ji})$  pair on the familiar plane of Lotka-Volterra interactions. The four quadrants of the plane are labeled competition, predation (twice), and mutualism.

Figure: The vector  $A_i = (a_{i1}, \dots, a_{in})$  moves in an  $n$ -dimensional space (in the illustration  $n = 2$ ). It's constrained in two ways when each population's variation is parametrized by one-dimensional character  $u_i$ .

Figure: Gradient of selection on  $A_i$  in one dimension is in direction closest to abstract gradient of selection  $S$ .

Figure: Each population  $j$  changes due to selection on  $a_{ji}$  terms. It causes change in  $a_{ij}$ , which is the source of the indirect effects on  $A_i$ .

Figure: The direction of the direct effect is always in the half-space within  $90^\circ$  of the abstract selection gradient  $S$ .  $S$  is always positive, which implies that the direct effect vector  $D$  is largely positive. The resultant direction of actual change in  $A_i$  is the sum  $\mathbf{D}_i + \mathbf{I}_i$ . The positive bias in  $\mathbf{D}_i$  may produce a positive bias in the motion of  $A_i$ .

Figure: direct and indirect effects on the pair  $(a_{ij}, a_{ji})$ .

## Claims and questions about evolution of interactions

The above argument establishes that if direct effects and indirect effects were IID random variables, interaction terms would drift in the direction of increase, tending to convert antagonistic interactions into mutualistic ones.

They may not be independent, though. Interaction terms are structured and correlated, and this structure may block the emergence of mutualism. I propose a three-part analysis of evolutionary change in interactions:

- The “default”, or *ceteris paribus*, outcome is positive drift, toward the mutualistic quadrant. When interactions do not move in that direction, there are two categories of constraints interfering with positive drift:
- Populations may not be able to increase the value of all their interactions due to tradeoffs in the different interaction terms affecting the population as patient. That is, the shape of the “fitness set” of the interactions affecting the population.
- Gains in the value of interactions may be lost due to negative indirect effects of agents’ adaptation on patients. This can imply a zero-sum situation in which one population’s gain is another’s loss, or a Hardin-like tragedy (“race to the bottom”, or social trap), in which a population’s direct gain is indirectly its own loss.

In an earlier paper (Worden and Levin 2007), we proposed that

emphasizing the difficulty of achieving cooperation in prisoner’s dilemma, social dilemma and tragedy of the commons situations without acknowledging the existence of other situations in which cooperation is much easier is both incomplete theoretically and dangerous socially. It can lead us to make overly pessimistic and destructive policy decisions, and can cause us to overlook opportunities to construct easy forms of cooperation, as well as making it more difficult to explain the ubiquity of cooperative human behavior.

I propose the above three-part structure as an alternative framework for researching the evolution of cooperation, focusing on **factors that interfere with the emergence of cooperation** in addition to studying means for mitigating tragic situations in which cooperation is hard.

This research may help us to explain both mutualistic and antagonistic situations in nature and society, and may also help illuminate opportunities to transform antagonistic situations or social traps to opportunities for mutual benefit.

## Example: Resource competition

Here we analyze the classic MacArthur-Levins formulation of  $n$  species competing for use of  $n$  resources [insert citation]. To construct this model, we begin by proposing a simple model for dynamics of  $n$  species and  $n$  resources:

$$\begin{aligned}\frac{dX_i}{dt} &= b_i X_i (\sum_{\ell} c_{i\ell} w_{\ell} R_{\ell} - m_i) \\ \frac{dR_{\ell}}{dt} &= r_{\ell} (K_{\ell} - R_{\ell}) - \sum_i c_{i\ell} X_i,\end{aligned}$$

where  $X_i$  is the density of population  $i$ ,  $R_{\ell}$  is the abundance of resource  $\ell$  and the parameters are  $b_i$ , an intrinsic population growth rate,  $m_i$ , mortality rate,  $c_{i\ell}$ , the rate at which population  $i$  captures resource  $\ell$ ,  $w_{\ell}$ , the amount a unit of resource  $\ell$  contributes to population growth,  $r_{\ell}$ , the resupply rate of resource  $\ell$ , and  $K_{\ell}$  its maximum possible abundance.

We then reduce that model to a Lotka-Volterra form by making the simplifying assumption that the resources come to equilibrium very quickly compared to the populations. Under this assumption, we can hold the population sizes  $X_i$  fixed and solve the second equation above for  $R_{\ell}$  when  $\frac{dR_{\ell}}{dt} = 0$ :

$$R_{\ell}^* = K_{\ell} - \frac{1}{r_{\ell}} \sum_i c_{i\ell} X_i.$$

Then we simply use that value of  $R_{\ell}$  in the first equation, and we have a system of population sizes only:

$$\frac{dX_i}{dt} = b_i X_i (\sum_{\ell} c_{i\ell} w_{\ell} (K_{\ell} - \frac{1}{r_{\ell}} \sum_j c_{j\ell} X_j) - m_i).$$

We can rearrange the terms of this to have the standard Lotka-Volterra form:

$$\frac{dX_i}{dt} = k_i X_i + \sum_j a_{ij} X_i X_j,$$

where

$$k_i = b_i (\sum_{\ell} c_{i\ell} w_{\ell} K_{\ell} - m_i)$$

(ordinarily I'd call this  $r_i$ , but the name is already in use), and

$$a_{ij} = -b_i \sum_{\ell} \frac{c_{i\ell} c_{j\ell} w_{\ell}}{r_{\ell}}.$$

In this model, ecological competition is inevitable by assumption, because the above forms are constructed so that  $a_{ij}$  and  $a_{ji}$  are always negative. This constraint has implications in both the direct and indirect effects emergent in the model.

## Evolution of ecological parameters

The parameters that can vary as population  $i$  adapts are those that are indexed by  $i$ :  $b_i$ ,  $m_i$ , and  $c_{i\ell}$  for each  $\ell$ . Let us give those functional forms,  $b_i = b(\mathbf{u}_i)$ ,  $m_i = m(\mathbf{u}_i)$ , and  $c_{i\ell} = c_{\ell}(\mathbf{u}_i)$ , assuming some underlying phenotype  $\mathbf{u}_i$ , and collect them into an ‘‘ecological phenotype’’ vector,

$$\mathbf{p}(\mathbf{u}_i) = \begin{pmatrix} b_i(\mathbf{u}_i) \\ m_i(\mathbf{u}_i) \\ c_{i1}(\mathbf{u}_i) \\ \vdots \\ c_{im}(\mathbf{u}_i) \end{pmatrix}.$$

This allows us to apply the analysis developed in the appendices. The selection gradient on  $\mathbf{p}_i$  is

$$\mathbf{S}(\mathbf{p}_i) = \begin{pmatrix} 0 \\ -b_i \\ b_i \hat{R}_0 \\ \vdots \\ b_i \hat{R}_n \end{pmatrix},$$

indicating that selection on these populations tends to reduce mortality and increase the conversion rate of each resource to births in proportion to the abundance of that resources. Interestingly, selection on the overall birth rate  $b_i$  is neutral, because variation in  $b_i$  alone does not give a variant an advantage over its resident parent type. (It is possible to imagine that the overall birth rate may be indirectly selected in some circumstances, however, since it increases the rate of adaptation in other characters, potentially allowing populations with higher birth rate to evade extinction better than others.)

That vector is therefore the predicted mean direction of adaptation of population  $i$ , if variation is available in all those quantities equally: mortality  $m_i$  will decline and resource uptake rates  $c_{i\ell}$  will increase in proportion to each resource's availability, while overall birth rate  $b_i$  remains constant or drifts neutrally.

The equilibrium resource availabilities  $\hat{R}_\ell$  decline as the populations' uptake rates increase; this is not immediately obvious from the expression  $\hat{R}_\ell = K_\ell - \frac{1}{r_\ell} \sum_i c_{i\ell} \hat{X}_i$  because equilibrium population sizes  $\hat{X}_i$  need to be found by a matrix inversion, whose result is not intuitively obvious. However, it is a well known result of resource competition theory (Tilman 1982) that the result of successive replacements in this community has the effect of reducing the availability of all the affected resources.

As a special case, let us consider variation restricted to the uptake rates  $c_{i\ell}$ , since those are the quantities most relevant to our concerns about competition.

In a community of one population consuming one resource, the uptake rate increases, resource availability decreases, and as can be seen from their forms, the intrinsic growth rate  $k_1$ , which is positive, increases, and the Lotka-Volterra interaction term  $a_{11}$ , which is negative, becomes larger in absolute value, signifying an increase in competition (figure 2).

In a community of two populations and two resources, when  $b_i$  and  $m_i$  are fixed but the four  $c_{i\ell}$  values are subject to evolutionary change, the selection gradient on  $\mathbf{p}_i = \begin{pmatrix} c_{i1} \\ c_{i2} \end{pmatrix}$  is  $b_i \begin{pmatrix} \hat{R}_1 \\ \hat{R}_2 \end{pmatrix}$ .

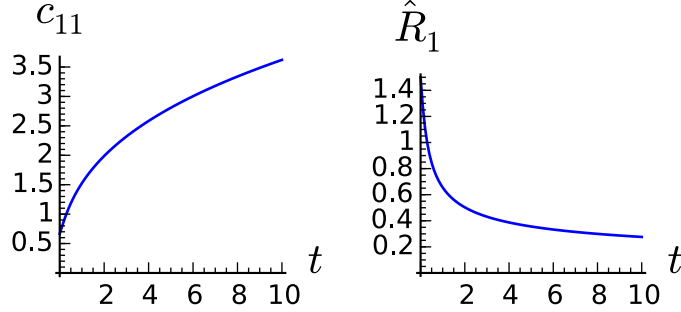


Figure 2: Evolution of ecological parameters in a one-population, one-resource model: change in  $c_{11}$  and  $\hat{R}_1$  when  $c_{11}$  is considered to be a phenotypic character subject to heritable variation.

When  $c$  values vary independently, they move according to the selection gradient, increasing and drawing  $\hat{R}_1$  and  $\hat{R}_2$  down: see figure 3.

When the pair  $(\frac{c_{i1}}{c_{i2}})$  is constrained by dependence on a single real value  $u_i$ , the direction of motion is described by the constrained selection gradient on  $\mathbf{p}$ , as discussed in the appendix. The selection gradient on  $u$  is

$$\begin{aligned} \mathbf{S}(u_i) &= \frac{\partial \mathbf{p}_i}{\partial u_i}^T \mathbf{S}(\mathbf{p}_i) \\ &= b_i \sum_{\ell} \hat{R}_{\ell} \frac{dc_{i\ell}}{du_i}, \end{aligned}$$

and  $u_i$  will either increase or decrease, according as  $\mathbf{S}(u_i)$  is positive or negative. By choosing a functional form  $c_{\ell}(u_i)$  that is either convex or concave, we can construct communities that select for a single population generalized to consume both resources, or for a pair of specialized populations. [To do: details of these models. I could also discuss geometrical connections between the selection gradient and ZNGI analysis of species replacement on the  $R_1 - R_2$  plane. For this draft I want to move on to the question about  $a_{ij}$ .]

## Evolution of Lotka-Volterra interactions

In the above models we have seen that contrary to this paper's analysis, the Lotka-Volterra competition terms  $a_{ij}$  do not decrease in strength, let alone increase from negative to positive values. We have explained their negative trajectory in terms of the selection on the ecological parameter  $c_{i\ell}$ . Now let us

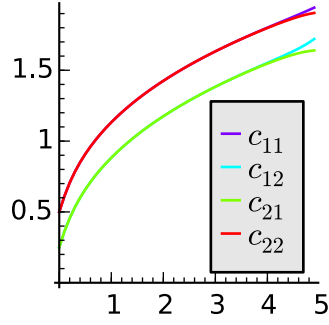


Figure 3: Evolution of ecological parameters in a two-population, two-resource model when each population's two  $c_{i\ell}$  values mutate independently and identically. [Plot of  $R$  to come. Also, redo with less symmetry so that all four  $c$  values and  $a$  values are visible.]

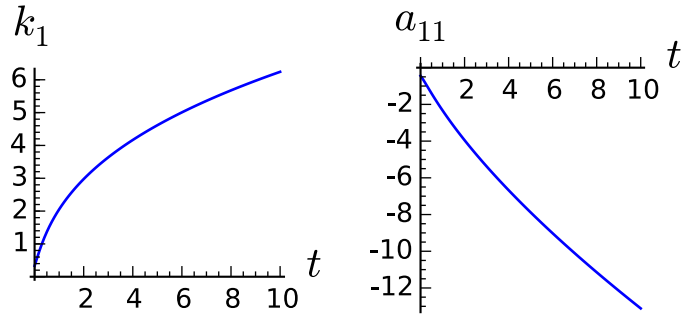


Figure 4: Change in  $k_1$  and  $a_{11}$  due to adaptation in the  $c_{11}$  parameter.



analyze the evolution of the Lotka-Volterra terms in terms of selection on those terms and constraints among them.

[to do somewhere along here: tragedy of commons of the one a term in monomorphic maclev model, in 1-d with both partials visible.]

The key step here is to understand the direction of selection on each of the the Lotka-Volterra terms and how those effects produce the resulting selection on the  $c_{i\ell}$  values.

The Lotka-Volterra form of the MacArthur-Levins model breaks the dynamics of the model into three parts, all relevant to understanding when and why competition increases, as it does in the cases presented above.

First and second are the two terms of the intrinsic growth term,

$$k_i = b_i \sum_{\ell} c_{i\ell} w_{\ell} K_{\ell} - b_i m_i,$$

and third is the competition term

$$a_{ij} = -b_i \sum_{\ell} c_{i\ell} w_{\ell} \frac{c_{j\ell}}{r_{\ell}}.$$

This division of terms breaks the resource abundances

$$R_{\ell}^* = K_{\ell} - \sum_j \frac{c_{j\ell}}{r_{\ell}} X_j$$

into parts and invites us to consider them separately.  $K_{\ell}$  is the abiotic abundance of  $R_{\ell}$ , the amount that would be present if all consumer populations were absent. Each coefficient  $\frac{c_{i\ell}}{r_{\ell}}$  is the rate of draw-down of resource  $\ell$  per unit of consumer population  $i$ , and the weighted sum of abiotic abundance and draw-down is the true abundance of  $R_{\ell}$ .

The intrinsic growth rate  $k_i$ , then, is the growth rate that population  $i$  would obtain if all resources were fixed at their abiotic abundances. Density dependence and competition are caused by draw-down of the resources, and the Lotka-Volterra competition term  $a_{ij}$  directly reflects the effect of all the draw-down caused by population  $j$  on the growth rate of population  $i$ .

The interaction selection gradient,  $\mathbf{S}(\mathbf{A}_i)$ , is always the same independent of the form of the Lotka-Volterra model:

$$\mathbf{S}(\mathbf{A}_i) = \begin{pmatrix} 1 \\ \hat{X}_1 \\ \vdots \\ \hat{X}_n \end{pmatrix},$$

indicating that selection on these populations favors increase in the intrinsic growth rate  $k_1$  at constant rate 1, and increase in the interaction terms  $a_{ij}$  at rate proportional to the long-term average population sizes  $\hat{X}_j$ .

Since  $k_i$  increases with  $c_{i\ell}$  and decreases with  $m_i$ , selection for increase in  $k_i$  is selection for increase in  $c_{i\ell}$  and decrease in  $m_i$ . More precisely, the direction of selection on  $b_i$ ,  $m_i$ , and  $c_{i\ell}$  due to selection on  $k_i$  are given by

$$\frac{\partial k_i}{\partial b_i} = \sum_{\ell} c_{i\ell} w_{\ell} K_{\ell} - m_i$$

$$\frac{\partial k_i}{\partial m_i} = -b_i$$

$$\frac{\partial k_i}{\partial c_{i\ell}} = b_i w_{\ell} K_{\ell}.$$

The direction of selection in  $c_{i\ell}$  is positive, and on  $m_i$  is negative. The direction of selection on  $b_i$  has the same sign as the population growth rate given abiotic resource abundances, which in any useful community model must be positive.

The competition term  $a_{ij}$ , on the other hand, decreases with both  $b_i$  and the  $c_{i\ell}$  terms, and is independent of  $m_i$ :

$$\frac{\partial a_{ij}}{\partial b_i} = -\sum_{\ell} c_{i\ell} w_{\ell} \sum_j \frac{c_{j\ell}}{r_{\ell}}$$

$$\frac{\partial a_{ij}}{\partial m_i} = 0$$

$$\frac{\partial a_{ij}}{\partial c_{i\ell}} = -b_i w_{\ell} \frac{c_{j\ell}}{r_{\ell}}.$$
<sup>1</sup>

Therefore selection on  $a_{ij}$  tends to decrease the overall growth rate  $b_i$  and the uptake rates  $c_{i\ell}$ .

The combination of these values adds up to the actual selection gradient on the ecological phenotype variables, which we already know from the previous section:

$$\mathbf{S}(\mathbf{b}_i) = \begin{pmatrix} \sum_{\ell} c_{i\ell} w_{\ell} (K_{\ell} - \sum_j \hat{X}_j \frac{c_{j\ell}}{r_{\ell}}) - m_i \\ -b_i \\ b_i w_1 (K_1 - \sum_j \hat{X}_j \frac{c_{j1}}{r_1}) \\ \vdots \\ b_i w_n (K_n - \sum_j \hat{X}_j \frac{c_{jn}}{r_n}) \end{pmatrix} = \begin{pmatrix} 0 \\ -b_i \\ \hat{R}_1 \\ \vdots \\ \hat{R}_n \end{pmatrix}.$$

(This is the sum  $\partial_1 \mathbf{A}_i^T \mathbf{S}(\mathbf{A}_i)$  discussed in the appendix.)

### Selection on $k$ causes competition to increase

However, considering them separately reveals additional information about the causes of selection: specifically, that there is a tension between selection on the  $k$  term and selection on the  $a$  terms with regard to the direction of change in the  $c$  values. Selection on the  $k$  terms favors increase in  $c$  values, because increasing them increases the population's benefit from all the resources. Selection on  $c$  terms via the  $a$  terms, though, isolates the effect on the patient population  $i$

<sup>1</sup>Note that this partial derivative is taken only with respect to the  $c_{i\ell}$  term corresponding to the *patient of the interaction*, that is, the first argument of each interaction term  $a(u_i, u_j)$  – see the appendix for details.

of resource draw-down due to each agent population  $j$ , which is harmful to the patient. Thus the two effects on each  $c_{i\ell}$  are at odds.

However, the amount of resource drawn down by the populations is never as much as the amount available in the absence of draw-down. Thus the net effect of these different factors is always positive, and thus the net selection on each  $c_{i\ell}$  is in the direction of increase, with the consequence that selection on the  $c_{i\ell}$  terms will always increase  $k_i$  and decrease the  $a_{ij}$  coefficients.

This is a **direct effect** on the  $a_{ij}$  terms, in which the dependence of both  $k_i$  and  $a_{ij}$  on  $c_{i\ell}$  produces a tradeoff such that an increase in one is a decrease in the other.

There is also an **indirect effect**, in which adaptation of each population  $j$  in the agent role has a side effect of decreasing the benefit to  $i$  of their interaction, i.e. of decreasing the value of  $a_{ij}$ . This is the effect that increasing each  $c_{j\ell}$  has on the value of  $a_{ij}$ : competition between populations intensifies as resources become less abundant. This includes the case  $j = i$ , in which population  $i$  acts on itself by drawing down the resources it needs to consume.

### Selection on $a$ terms alone decreases competition

The observation that increasing competition (negative change in  $a$  terms) is caused by selection on  $k$  terms, not on  $a$  terms, suggests an experiment: what would happen to  $a$  terms if  $k$  terms were held fixed?

We can do this by imposing a trade-off between the  $c_{i\ell}$  values, which are involved in both  $a$  and  $k$  terms, and the mortality  $m_i$ , which is involved only in  $k_i$ , constructed such that  $k_i = b_i \sum_{\ell} c_{i\ell} w_{\ell} K_{\ell} - b_i m_i$  is constant. This is readily done by setting  $m_i = m_0 + \sum_{\ell} c_{i\ell} w_{\ell} K_{\ell}$  for some constant  $m_0 < 0$ .

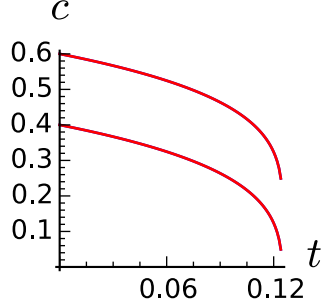
This gives  $k_i = -b_i m_0$ , while all  $a_{ij}$  terms are still as above, and the net direction of selection on each  $c_{i\ell}$  is now negative. As a result, the  $c$  values decrease, and the competition terms  $a_{ij}$  shrink by increasing toward zero, indicating a decrease in competition.

[To do: analysis of this case by defining an “underlying phenotype”  $\mathbf{u}_i$ , equal to the vector of  $c_{i\ell}$  values, with the ecological vector  $\mathbf{p}_i$  as a function of those. The dependence of both  $m_i$  and  $c_{i\ell}$  on  $\mathbf{u}_i$  imposes a constraint that translates selection on  $a_{ij}$  terms into negative selection on the terms of  $\mathbf{u}_i$ .]

[To do: move much of this example’s analysis to appendix, leave only discussion of it here.]

What is the general condition for competition to decrease? In a one-population system, when  $r$  and  $a$  are constrained by a scalar phenotype  $u$ ,

In general, given some functions  $b, m, c_1$  of a scalar  $u$ ,



[WorkingWiki encountered errors: Error:  
 <ww-make-failed: maclev-2-2-a-c-a-vs-t.pdf,  
 /Selection\_Gradients/maclev-2-2-a-c-a-vs-t.pdf.make.log> ]

Figure 5: Change in  $c$  and  $a$  values due to adaptation in the  $c$  values when  $k$  values are held fixed.

## Example: Predator-prey model

[In this model, interaction pairs are confined to the predation quadrants by a negative correlation between  $a_{ij}$  and  $a_{ji}$ .

This should be a matter of indirect effects. Presumably if left to indirect effects, predators would get fuller and fuller while eating fewer and fewer prey. When this doesn't happen, it's due to some kind of arms race which takes the form of a tug-of-war in the indirect effects.

Fuller analysis to come.]

Along with competition in conditions of scarcity, the most common scenario for non-mutualistic interactions is predation. To investigate the constraints maintaining predation interactions, let us consider a simple community model structured by predator-prey relationships.

In this model, a community is defined by a directed graph of predation relationships, with an arrow connecting each prey population to each of its predators. Each population has a vector  $\mathbf{u}_i$  of underlying phenotype values, and the Lotka-Volterra interaction term for each pair of populations is

$$a_{ij} = a(\mathbf{u}_i, \mathbf{u}_j) = \begin{cases} -f(\mathbf{u}_i, \mathbf{u}_j) & \text{if } i \rightarrow j \\ kf(\mathbf{u}_j, \mathbf{u}_i) & \text{if } j \rightarrow i \\ -1 & \text{if both } i \leftrightarrow j \\ 0 & \text{otherwise,} \end{cases}$$

where we write  $i \rightarrow j$  to mean that  $j$  eats  $i$ , and  $i \leftrightarrow j$  to mean that both  $i$  and  $j$  are autotrophs, meaning populations that have no prey.  $k$ , between 0 and 1, is a rate of conversion efficiency from prey killed to increase in predator population. To provide some flow into the system and keep populations bounded, we set

$$r_i = r(\mathbf{u}_i) = \begin{cases} 1 & \text{if } i \text{ is an autotroph} \\ -1 & \text{otherwise.} \end{cases}$$

Thus the population dynamics equations are

$$\frac{dX_i}{dt} = X_i(r_i + \sum_{j \rightarrow i} k f(\mathbf{u}_j, \mathbf{u}_i) X_j - \sum_{i \rightarrow j} f(\mathbf{u}_i, \mathbf{u}_j) X_j - \sum_{i \leftrightarrow j} X_j)$$

for each  $i$ .

Populations adapt by changing their phenotypes  $\mathbf{u}_i$ , while their placement in the food web stays fixed. By choosing a functional form  $f(\mathbf{u}_i, \mathbf{u}_j) > 0$ , we can investigate the interplay between coevolution and strength of interactions, while interactions are constrained to be either predator-prey interactions or the density-dependent competition among autotrophs that keeps the populations' sizes bounded.

For simplicity, let's say  $\mathbf{u}_i$  (in this case,  $u_i$ ) is a single real number.

## Selection pressures on predators and prey

The selection gradients follow directly from the analysis of the appendices. Because  $r$  is constant for each population, selection is only through the  $a$  terms, and that only through the terms involving the predation strength function  $f$ . Given the long-term average population sizes  $\hat{X}_i$ , the selection gradient on  $u_i$  is

$$\begin{aligned} \mathbf{S}(u_i) &= \sum_j \partial_1 a(u_i, u_j) \hat{X}_j \\ &= \sum_{j \rightarrow i} k \partial_2 f(u_j, u_i) \hat{X}_j - \sum_{i \rightarrow j} \partial_1 f(u_i, u_j) \hat{X}_j. \end{aligned}$$

Below I will use the functional form

$$f(u_i, u_j) = 2 + \alpha \cos(u_i - u_j),$$

with some  $\alpha \in (0, 1)$ .

When  $f(u_i, u_j)$  depends only on the difference  $u_i - u_j$ , with an antisymmetry relationship  $f'(-x) = -f'(x)$ , then we have  $\partial_2 f(u_j, u_i) = \frac{d}{du_i} f(u_j - u_i) = -f'(u_j - u_i) = f'(u_i - u_j) = \partial_1 f(u_i, u_j)$ , which gives

$$\mathbf{S}(u_i) = \sum_{j \rightarrow i} k \hat{X}_j f'(u_i - u_j) - \sum_{i \rightarrow j} \hat{X}_j f'(u_i - u_j).$$

Thus the direction of selection on  $u_i$  depends on a weighted sum of influences of its prey and/or predators, shaped by each other population's size, the difference in their phenotype values  $u$ , and the form of the function  $f$ . In the trigonometric case it is

$$\mathbf{S}(u_i) = \sum_{j \rightarrow i} k \hat{X}_j \sin(u_i - u_j) - \sum_{i \rightarrow j} \hat{X}_j \sin(u_i - u_j).$$

With this trigonometric functional form, predation is most intense when predator's and prey's  $u$  values match (give or take a multiple of  $2\pi$ ), and least intense when they differ by an odd multiple of  $\pi$ .

### Example: one prey, one predator

When there is only population 1 eating population 0, the Lotka-Volterra system is

$$\begin{aligned} \frac{dX_0}{dt} &= X_0(1 - X_0 - X_1(1 + \cos(u_0 - u_1))) \\ \frac{dX_1}{dt} &= X_1(-1 + kX_0(1 + \cos(u_0 - u_1))) \end{aligned}$$

These populations' adaptive dynamics follow their selection gradients:

$$\begin{aligned} \mathbf{S}(u_0) &= -\hat{X}_1 \sin(u_0 - u_1) \\ \mathbf{S}(u_1) &= k\hat{X}_0 \sin(u_0 - u_1). \end{aligned}$$

Each population's full adaptive dynamics is proportional to its own size as well:

$$\begin{aligned} \frac{du_0}{dt} &= -\gamma \hat{X}_0 \hat{X}_1 \sin(u_0 - u_1) \\ \frac{du_1}{dt} &= k\gamma \hat{X}_0 \hat{X}_1 \sin(u_0 - u_1). \end{aligned}$$

From this we can readily see, since  $0 < k < 1$ , that

$$\frac{d(u_0 - u_1)}{dt} = \gamma \hat{X}_0 \hat{X}_1 (k - 1) \sin(u_0 - u_1)$$

is opposite in sign to  $\sin(u_0 - u_1)$ , so that it is positive when  $u_0 - u_1$  is small and positive, and negative when  $u_0 - u_1$  is small and negative. Thus  $u_0$  can escape from  $u_1$  when the two are close.

This example perhaps follows the “life-dinner hypothesis”, in which selection acts more strongly on prey because they have more to lose from each interaction (Dawkins and Krebs 1979).

Equilibria of this system are configurations where one or more populations go extinct ( $\hat{X}_i = 0$ ) or where  $u_0 - u_1$  is a multiple of  $\pi$ . As we have seen,  $u_0 - u_1 = 0$

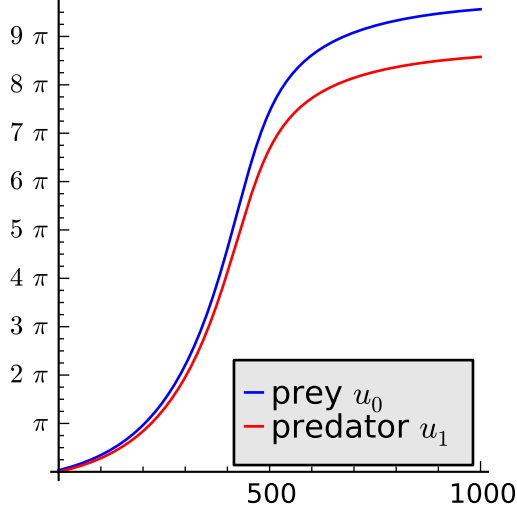


Figure 6: Coevolutionary chase of predator and prey's phenotypes

is unstable. By the analogous argument,  $u_0 - u_1 = (2n + 1)\pi$  is stable, which is the configuration that minimizes the strength of predation  $f$ , and maximizes the negative interaction term  $a_{01}$ .

**Another example here?**

### Direct and indirect effects in the 2-species example

The direct and indirect effects predicting the evolution of interactions can be written directly using the definitions in the appendix. The terms of the interaction phenotype vector for population  $i$  are the terms involved in the population dynamics equation for  $X_i$ :

$$\mathbf{A}_0 = \begin{pmatrix} r_0 \\ a_{00} \\ a_{01} \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ -f(u_0 - u_1) \end{pmatrix}$$

$$\mathbf{A}_1 = \begin{pmatrix} r_1 \\ a_{10} \\ a_{11} \end{pmatrix} = \begin{pmatrix} -1 \\ kf(u_0 - u_1) \\ 0 \end{pmatrix}.$$

The direct effect,  $\mathbf{D}_i = \gamma \hat{X}_i \partial_1 \mathbf{A}_i \partial_1 \mathbf{A}_i^T \mathbf{S}(\mathbf{A}_i)$ , is

$$\begin{aligned}
\mathbf{D}_0 &= \gamma \hat{X}_0 \begin{pmatrix} 0 \\ 0 \\ -f'(u_0 - u_1) \end{pmatrix} \begin{pmatrix} 0 & 0 & -f'(u_0 - u_1) \end{pmatrix} \begin{pmatrix} 1 \\ \hat{X}_0 \\ \hat{X}_1 \end{pmatrix} \\
&= \gamma \hat{X}_0 \hat{X}_1 \begin{pmatrix} 0 \\ 0 \\ f'(u_0 - u_1)^2 \end{pmatrix} \\
\mathbf{D}_1 &= \gamma \hat{X}_1 \begin{pmatrix} 0 \\ kf'(u_0 - u_1) \\ 0 \end{pmatrix} \begin{pmatrix} 0 & kf'(u_0 - u_1) & 0 \end{pmatrix} \begin{pmatrix} 1 \\ \hat{X}_0 \\ \hat{X}_1 \end{pmatrix} \\
&= \gamma \hat{X}_0 \hat{X}_1 \begin{pmatrix} 0 \\ k^2 f'(u_0 - u_1)^2 \\ 0 \end{pmatrix}.
\end{aligned}$$

Due to the sparsity of terms depending on phenotypes in the  $\mathbf{A}$  vectors, we find direct effects only affecting the predation terms,  $a_{01}$  and  $a_{10}$ . The prey species 0 adapts in a way that directly reduces the predation on it, and the predator, species 1, adapts so as to increase the predation feeding it.

The indirect effects are also concentrated on those two terms:

$$\begin{aligned}
\mathbf{I}_0 &= \gamma \begin{pmatrix} 0 \\ \hat{X}_0 \partial_2 a(u_0, u_0) \mathbf{S}(u_0) \\ \hat{X}_1 \partial_2 a(u_0, u_1) \mathbf{S}(u_1) \end{pmatrix} = \gamma \begin{pmatrix} 0 \\ 0 \\ -\hat{X}_0 \hat{X}_1 f'(u_0 - u_1) \mathbf{S}(u_1) \end{pmatrix} \\
&= \gamma \hat{X}_0 \hat{X}_1 \begin{pmatrix} 0 \\ 0 \\ -kf'(u_0 - u_1)^2 \end{pmatrix} \\
\mathbf{I}_1 &= \gamma \begin{pmatrix} 0 \\ \hat{X}_0 \partial_2 a(u_1, u_0) \mathbf{S}(u_0) \\ \hat{X}_1 \partial_2 a(u_1, u_1) \mathbf{S}(u_1) \end{pmatrix} = \gamma \begin{pmatrix} 0 \\ \hat{X}_0 \hat{X}_1 kf'(u_0 - u_1) \mathbf{S}(u_0) \\ 0 \end{pmatrix} \\
&= \gamma \hat{X}_0 \hat{X}_1 \begin{pmatrix} 0 \\ -kf'(u_0 - u_1)^2 \\ 0 \end{pmatrix}.
\end{aligned}$$

The indirect effects are the effects of each species's evolution on the other: change in the predator species increases predation on the prey, and change in the prey decreases the benefit to the predator.

Putting these together gives the resulting change in interactions:



$$\begin{aligned}\frac{d\mathbf{A}_0}{dt} &= \gamma \hat{X}_0 \hat{X}_1 \begin{pmatrix} 0 \\ 0 \\ (1-k)f'(u_0 - u_1)^2 \end{pmatrix} \\ \frac{d\mathbf{A}_1}{dt} &= \gamma \hat{X}_0 \hat{X}_1 \begin{pmatrix} 0 \\ (k^2 - k)f'(u_0 - u_1)^2 \\ 0 \end{pmatrix}.\end{aligned}$$

The form of the predator-prey interaction terms imposes the constraint  $a_{10} = -ka_{01}$ , so that their derivatives must also have a  $-1 : k$  ratio, and the above expression confirms that.

The representation in terms of Lotka-Volterra pairs is especially relevant here, since there is only one pair of interest:

$$\begin{aligned}\frac{da_{(0,1)}}{dt} &= \frac{d}{dt} \begin{pmatrix} a_{01} \\ a_{10} \end{pmatrix} \\ &= D_{(0,1)} + I_{(0,1)} \\ &= \gamma \hat{X}_0 \hat{X}_1 \left[ \begin{pmatrix} 1 \\ k^2 \end{pmatrix} + \begin{pmatrix} -k \\ -k \end{pmatrix} \right] f(u_0 - u_1)^2 \\ &= \gamma \hat{X}_0 \hat{X}_1 \begin{pmatrix} 1 \\ -k \end{pmatrix} (1-k)f(u_0 - u_1)^2\end{aligned}$$

## Direct and indirect effects in the general case

In an unspecified foodweb, with an unspecified functional form  $f(u_i, u_j)$ , we have as above

$$\mathbf{S}(u_i) = \sum_{j \rightarrow i} k \partial_2 f(u_j, u_i) \hat{X}_j - \sum_{i \rightarrow j} \partial_1 f(u_i, u_j) \hat{X}_j,$$

which describes a population  $u_i$  adapting by moving towards its prey and away from its predators, each weighted by its long-term population size, and with the selective pressure due to prey attenuated by a factor of  $k < 1$ .

Lotka-Volterra interactions fall naturally into predator-prey pairs  $a_{(i,j)}$ , and the motion of each pair (where we assume without loss of generality that  $i$  is the prey of  $j$ ) is

$$\begin{aligned}
\frac{da_{(i,j)}}{dt} &= D_{(i,j)} + I_{(i,j)} \\
&= \gamma \left[ \left( -\hat{X}_i \partial_1 f(u_i, u_j) \mathbf{S}(u_i) \right) + \left( -\hat{X}_j \partial_2 f(u_i, u_j) \mathbf{S}(u_j) \right) \right] \\
&= \gamma \left[ \binom{-1}{k} \hat{X}_i \partial_1 f(u_i, u_j) \mathbf{S}(u_i) + \binom{-1}{k} \hat{X}_j \partial_2 f(u_i, u_j) \mathbf{S}(u_j) \right] \\
&= \gamma \left[ \binom{-1}{k} \hat{X}_i \partial_1 f(u_i, u_j) \partial_1 \mathbf{A}_i + \binom{-1}{k} \hat{X}_j \partial_2 f(u_i, u_j) \partial_1 \mathbf{A}_j \right] \mathbf{S}(\mathbf{A}) \\
&= \gamma \binom{-1}{k} \left[ \hat{X}_i \partial_1 f(u_i, u_j) \partial_1 \mathbf{A}_i + \hat{X}_j \partial_2 f(u_i, u_j) \partial_1 \mathbf{A}_j \right] \mathbf{S}(\mathbf{A})
\end{aligned}$$

## Externalities and the threat of “cheating” in evolution of mutualisms

In the natural history of fig-fig wasp mutualisms, the coevolving tree and insect species have transitioned from parasitism to mutualism, and from mutualism to parasitism, a number of times.

[To come]

## Example: statistical mechanics of interactions.

Create a model of a large number of coexisting species with an interaction function ( $a(u_i, u_j)$ ) just complex enough to allow pairs of species to move among all four quadrants independently. Hypothesis: coevolution produces a statistical drift toward the upper right.

## Example: adaptive change in a game theory payoff matrix

[to do: this is just discussion, not a full example]

Adaptation of economic actors by learning, under just about any reasonable simple model assumptions (within the rational-choice framework), follows the same gradient as the above evolutionary models. Thus all the geometrical analysis holds, including the three-part research framework.

In this section I demonstrate how the terms of the game's payoff matrix change in just the same way as the Lotka-Volterra interaction terms.

Then make a little sketch of how a social interaction might transition among prisoner's dilemma, snowdrift, stag hunt, and byproduct cooperation structures, seen as a matter of coevolving constrained terms of the payoff matrix.

## Discussion

## Appendices

### Adaptive change in constrained characters

Let  $\mathbf{p}_i$  be a vector of numbers representing phenotypic characters undergoing adaptive change in the  $i$ th population under consideration. Under adaptive dynamics assumptions (Metz et al. 1996), the dynamics of the population's change in evolutionary time is (Ma, Worden, and Levin 2007)

$$\frac{d\mathbf{p}_i}{dt} = \gamma \hat{X}_i \left. \frac{\partial \mathcal{I}(\mathbf{q}|\mathbf{p}_1, \dots, \mathbf{p}_n)}{\partial \mathbf{q}} \right|_{\mathbf{q}=\mathbf{p}_i} = \gamma \hat{X}_i \partial_1 \mathcal{I}(\mathbf{p}_i)^T,$$

where  $\mathcal{I}(q|E)$  is the **invasion fitness** function: the long-term growth rate of a very small population of type  $q$  in environment  $E$ . In the above formulation we assume that the community defined by all coexisting populations' monomorphic phenotypes  $\mathbf{p}_j$  is sufficient to define the selective environment  $E$ .

In the above notation,  $\mathbf{p}_i$  is a column vector,  $\frac{\partial \mathcal{I}}{\partial \mathbf{p}_i}$  is a row vector,  $\hat{X}_i$  is the long-term average size of population  $i$ , the superscript  $T$  stands for the vector transpose, and  $\partial_1$  stands for the partial derivative with respect to the function's first argument. The variable  $\gamma$  is a constant encapsulating the overall rate of evolutionary change.

Both  $\gamma$  and  $\hat{X}_i$  in the above expression are positive scalar quantities, so the direction of adaptive change in  $\mathbf{p}_i$  is in the direction of greatest increase in the invasion fitness available to a rare variant  $\mathbf{v}$ . Let us refer to that direction as the **gradient of selection**, or **selection gradient**, of  $\mathbf{p}_i$ :

$$\mathbf{S}(\mathbf{p}_i) = \partial_1 \mathcal{I}(\mathbf{p}_i)^T,$$

[To do: the important thing here is that it's the partial derivative with respect to the rare type, as opposed to the partial derivative with respect to the second  $\mathbf{p}_i$  argument (in which it is the resident type), or the total derivative with respect to  $\mathbf{p}_i$ . We are really interested in patterns of change in the components of  $\mathcal{I}$  determined by the total derivative, when change is driven only by the first partial derivative. This pattern of change in the value, fitness, or utility driven by its first partial derivative is in some sense the signature of adaptive change, whether in adaptive dynamics models, population genetics, economics, or animal behavior

(and organizational change? acquisition of immunity? look into it). Make this generality explicit, to clarify that this paper’s conclusions apply to all these things.

The corresponding thing in population genetics is presumably the change in allele frequencies,  $\frac{dp_i}{dt}$ , in the direction of  $\frac{\partial w}{\partial p_i}$ , when fitness  $w$  is frequency-dependent, and has just the same considerations about the arguments  $p_i$  representing patients and agents (analogous to rare types and residents – terminology defined below), and taking the partial derivative with respect to the patients, while impacts of agents on various patients are left as externalities not considered by selection.]

When a population’s phenotype is confined to only a bounded set within the full space of possible values, the above dynamics is only applicable at points where all degrees of freedom are available. Often the population will evolve to the boundary of that set, in which case its motion comes to be confined to that boundary. This become the dynamics of a vector constrained to a lower-dimensional subset of the space of possible values. That constraint can be represented by considering the phenotype to be a function of a smaller number of variables. We represent that by introducing an “underlying” phenotype vector  $\mathbf{u}_i$ , with a smaller number of entries, and writing the phenotype  $\mathbf{p}_i$  as a function  $\mathbf{p}(\mathbf{u}_i)$ .

The dynamics of the underlying phenotype evolves in the standard way:

$$\frac{d\mathbf{u}_i}{dt} = \gamma \hat{X}_i \frac{\partial \mathcal{I}(\mathbf{v}|\mathbf{u}_1, \dots, \mathbf{u}_n)}{\partial \mathbf{v}} \Big|_{\mathbf{v}=\mathbf{u}_i} = \gamma \hat{X}_i \partial_1 \mathcal{I}(\mathbf{u}_i)^T,$$

and the constrained dynamics of the constrained phenotype  $\mathbf{p}_i$  follows from the chain rule:

$$\begin{aligned} \mathbf{S}(\mathbf{u}_i) &= \partial_1 \mathcal{I}(\mathbf{u}_i)^T \\ &= \left( \partial_1 \mathcal{I}(\mathbf{p}_i) \frac{\partial \mathbf{p}}{\partial \mathbf{u}_i} \right)^T \\ &= \frac{\partial \mathbf{p}}{\partial \mathbf{u}_i}^T \partial_1 \mathcal{I}(\mathbf{p}_i)^T \\ &= \frac{\partial \mathbf{p}}{\partial \mathbf{u}_i}^T \mathbf{S}(\mathbf{p}_i) \\ \frac{d\mathbf{p}(\mathbf{u}_i)}{dt} &= \frac{\partial \mathbf{p}}{\partial \mathbf{u}_i} \frac{d\mathbf{u}_i}{dt} \\ &= \gamma \hat{X}_i \frac{\partial \mathbf{p}}{\partial \mathbf{u}_i} \mathbf{S}(\mathbf{u}_i) \\ &= \gamma \hat{X}_i \frac{\partial \mathbf{p}}{\partial \mathbf{u}_i} \frac{\partial \mathbf{p}}{\partial \mathbf{u}_i}^T \mathbf{S}(\mathbf{p}_i) \end{aligned}$$

While in the standard, unconstrained case, adaptive change in the phenotype  $\mathbf{p}_i$  follows the direction of the gradient of the invasion fitness,  $\mathbf{S}(\mathbf{p}_i)$ ,

$$\frac{d\mathbf{p}_i}{dt} = \gamma \hat{X}_i \mathbf{S}(\mathbf{p}_i),$$

in the constrained case, it is the underlying phenotype  $\mathbf{u}$  that follows its selection gradient:

$$\frac{d\mathbf{u}_i}{dt} = \gamma \hat{X}_i \mathbf{S}(\mathbf{u}_i),$$

while the constrained phenotype is stopped from following its selection gradient by its constraints, which limit it to change in a smaller space of available directions. Thus its derivative is in a different direction:

$$\frac{d\mathbf{p}(\mathbf{u}_i)}{dt} = \gamma \hat{X}_i S_c(\mathbf{p}_i),$$

where  $S_c(\mathbf{p}_i) = \frac{\partial \mathbf{p}}{\partial \mathbf{u}} \frac{\partial \mathbf{p}}{\partial \mathbf{u}}^T \mathbf{S}(\mathbf{p})$  is the **constrained selection gradient** of  $\mathbf{p}_i$  determined by the available variation in  $\mathbf{u}_i$ .

This derivative is a vector derived from the gradient of unconstrained selection,  $\mathbf{S}(\mathbf{p}_i)$ , by multiplying it by a matrix  $\frac{\partial \mathbf{p}}{\partial \mathbf{u}} \frac{\partial \mathbf{p}}{\partial \mathbf{u}}^T$ . That matrix contains information about the directions available to change in  $\mathbf{p}$  as  $\mathbf{u}$  varies, and the product  $\frac{\partial \mathbf{p}}{\partial \mathbf{u}} \frac{\partial \mathbf{p}}{\partial \mathbf{u}}^T \mathbf{S}(\mathbf{p})$  is a vector pointing in a direction of change in  $\mathbf{p}$  that can be obtained by varying the underlying vector  $\mathbf{u}$ .

The following lemma establishes an important relationship between the direction of constrained change and the gradient of unconstrained selection: the angle between the two vectors is always  $90^\circ$  or less.

**Lemma.** The direction of motion of the above phenotype vector  $\mathbf{p}_i$  constrained by an underlying phenotype  $\mathbf{u}_i$  is always at an angle no greater than  $90^\circ$  to the unconstrained selection gradient  $\mathbf{S}(\mathbf{p}_i)$ .

*Proof:* Consider the dot product of the two vectors:

$$\begin{aligned} \left( \frac{\partial \mathbf{p}}{\partial \mathbf{u}_i} \frac{\partial \mathbf{p}}{\partial \mathbf{u}_i}^T \mathbf{S}(\mathbf{p}_i) \right) \cdot \mathbf{S}(\mathbf{p}_i) &= \left( \frac{\partial \mathbf{p}}{\partial \mathbf{u}_i} \frac{\partial \mathbf{p}}{\partial \mathbf{u}_i}^T \mathbf{S}(\mathbf{p}_i) \right)^T \mathbf{S}(\mathbf{p}_i) \\ &= \mathbf{S}(\mathbf{p}_i)^T \frac{\partial \mathbf{p}}{\partial \mathbf{u}_i} \frac{\partial \mathbf{p}}{\partial \mathbf{u}_i}^T \mathbf{S}(\mathbf{p}_i) \\ &= \left( \frac{\partial \mathbf{p}}{\partial \mathbf{u}_i}^T \mathbf{S}(\mathbf{p}_i) \right)^T \left( \frac{\partial \mathbf{p}}{\partial \mathbf{u}_i}^T \mathbf{S}(\mathbf{p}_i) \right) \\ &= \left( \frac{\partial \mathbf{p}}{\partial \mathbf{u}_i}^T \mathbf{S}(\mathbf{p}_i) \right) \cdot \left( \frac{\partial \mathbf{p}}{\partial \mathbf{u}_i}^T \mathbf{S}(\mathbf{p}_i) \right) \\ &\geq 0, \end{aligned}$$

because the dot product of a vector with itself is always nonnegative. The dot product of two vectors is the product of their magnitudes, a positive quantity,

multiplied by the cosine of the angle  $\theta$  between them. When the dot product is positive, it follows that  $\cos(\theta) \geq 0$ , which implies that  $\theta$  is  $90^\circ$  or less.  $\square$

Here is another way of expressing the same conclusion:

**Corollary.** The selection gradient  $\mathbf{S}(\mathbf{p}_i)$  defines a half-space bounded by the vectors perpendicular to  $\mathbf{S}(\mathbf{p}_i)$  and including  $\mathbf{S}(\mathbf{p}_i)$ , and the direction of motion of a constrained vector  $\mathbf{p}_i$  must always be within that half-space.

When  $\mathbf{p}_i$  is constrained to a curve, this half-space generally determines the direction of its motion, as only one direction on the curve falls within that half-space.

## Direct and indirect effect of adaptation on interaction terms

The Lotka-Volterra interaction term between populations  $i$  and  $j$ ,  $a_{ij}$ , is similar to a phenotype constrained by a vector of underlying characters, but is parametrized by two populations' characters rather than one:  $a_{ij} = a(\mathbf{u}_i, \mathbf{u}_j)$ . As a result, the motion of  $a_{ij}$  as the community of populations adapts is similar but more complex to analyze.

In a Lotka-Volterra model defined by the population dynamics equations

$$\frac{dX_i}{dt} = \left( r(\mathbf{u}_i) + \sum_j a(\mathbf{u}_i, \mathbf{u}_j) X_j \right) X_i,$$

the invasion fitness is the function

$$\mathcal{I}(\mathbf{v}|\mathbf{u}_1, \dots, \mathbf{u}_n) = r(\mathbf{v}) + \sum_j a(\mathbf{v}, \mathbf{u}_j) \hat{X}_j,$$

and the gradient of selection on population  $i$  is

$$\begin{aligned} \mathbf{S}(\mathbf{u}_i) &= \partial_1 \mathcal{I}(\mathbf{v}|\mathbf{u}_1, \dots, \mathbf{u}_n)^T \\ &= \frac{\partial}{\partial \mathbf{v}} \left( r(\mathbf{v}) + \sum_j a(\mathbf{v}, \mathbf{u}_j) \hat{X}_j \right)^T \bigg|_{\mathbf{v}=\mathbf{u}_i} \\ &= r'(\mathbf{u}_i)^T + \sum_j \partial_1 a(\mathbf{u}_i, \mathbf{u}_j)^T \hat{X}_j. \end{aligned}$$

[To do: flesh out the derivation of that invasion fitness, since it's a key step.]

We already know the motion of the underlying phenotypes:

$$\frac{d\mathbf{u}_i}{dt} = \gamma \hat{X}_i \mathbf{S}(\mathbf{u}_i),$$

and from that we can derive the motion of the Lotka-Volterra terms, both the intrinsic growth  $r$  and the interaction terms  $a_{ij}$ .

$$\begin{aligned}
\frac{dr_i}{dt} &= r'(\mathbf{u}_i) \frac{d\mathbf{u}_i}{dt} \\
&= r'(\mathbf{u}_i) \gamma \hat{X}_i \mathbf{S}(\mathbf{u}_i) \\
\frac{da_{ij}}{dt} &= \partial_1 a(\mathbf{u}_i, \mathbf{u}_j) \frac{d\mathbf{u}_i}{dt} + \partial_2 a(\mathbf{u}_i, \mathbf{u}_j) \frac{d\mathbf{u}_j}{dt} \\
&= \partial_1 a(\mathbf{u}_i, \mathbf{u}_j) \gamma \hat{X}_i \mathbf{S}(\mathbf{u}_i) + \partial_2 a(\mathbf{u}_i, \mathbf{u}_j) \gamma \hat{X}_j \mathbf{S}(\mathbf{u}_j).
\end{aligned}$$

We have seen, above, that the selection gradient on population  $i$ ,  $\mathbf{S}(\mathbf{u}_i)$ , is determined not by all the interaction terms involving population  $i$ , but only by the terms in which  $\mathbf{u}_i$  appears *in the first of the two arguments*, together with the  $r$  term for population  $i$ 's intrinsic growth.

I will refer to these as the terms in which population  $i$  is the **patient** – the entity acted on – distinct from the **agent** – the entity acting on the other. The adaptation of population  $i$  is driven by selection on the interactions that affect population  $i$ , not by the effects population  $i$  has on others (even on itself, as we will see).

We can draw a useful analogy with the analysis of constrained phenotypes, above, by defining a sort of “interaction phenotype” of population  $i$ , consisting of the quantities that drive selection on population  $i$ . That is a vector

$$\mathbf{A}(\mathbf{v}) = \begin{pmatrix} r(\mathbf{v}) \\ a(\mathbf{v}, \mathbf{u}_1) \\ \vdots \\ a(\mathbf{v}, \mathbf{u}_n) \end{pmatrix}.$$

Let

$$\mathbf{A}_i = \mathbf{A}(\mathbf{u}_i).$$

Then we can reformulate the motion of  $\mathbf{u}_i$  in terms of this vector. Since

$$\partial_1 \mathbf{A}_i = \begin{pmatrix} r'(\mathbf{u}_i) \\ \partial_1 a(\mathbf{u}_i, \mathbf{u}_1) \\ \vdots \\ \partial_1 a(\mathbf{u}_i, \mathbf{u}_n) \end{pmatrix}$$

and

$$\frac{\partial \mathcal{I}(\mathbf{v} | \mathbf{u}_1, \dots, \mathbf{u}_n)}{\partial \mathbf{A}_i} = (1, \hat{X}_1, \dots, \hat{X}_n),$$

we can define an **interaction selection gradient**

$$\mathbf{S}(\mathbf{A}_i) = \left( \frac{\partial \mathcal{I}(\mathbf{v} | \mathbf{u}_1, \dots, \mathbf{u}_n)}{\partial \mathbf{A}_i} \right)^T \Big|_{\mathbf{v}=\mathbf{u}_i} = \begin{pmatrix} 1 \\ \hat{X}_1 \\ \vdots \\ \hat{X}_n \end{pmatrix}$$

and rewrite the selection gradient of the underlying phenotype just as we did in the constrained- $\mathbf{p}$  case:

$$\mathbf{S}(\mathbf{u}_i) = \partial_1 \mathbf{A}_i^T \mathbf{S}(\mathbf{A}_i).$$

However, we can not simply write the motion of  $\mathbf{A}_i$  as  $\frac{d\mathbf{A}_i}{dt} = \partial_1 \mathbf{A}_i^T \frac{d\mathbf{u}_i}{dt}$  by analogy to the  $\mathbf{p}$  case, because the motion of the  $a_{ij}$  terms depends on their agent populations  $\mathbf{u}_j$  as well as the patient  $\mathbf{u}_i$ .

The full equation of motion for  $\mathbf{A}_i$ , and thus for the motion of all the  $r$  and  $a$  terms of the Lotka-Volterra model, requires two terms, which I call a **direct effect**,  $\mathbf{D}_i$ , of selection on population  $i$ , and an **indirect effect**,  $\mathbf{I}_i$ , of selection on all the populations in the community on the  $i$ th interaction phenotype:

$$\begin{aligned} \mathbf{D}_i &= \gamma \hat{X}_i \partial_1 \mathbf{A}_i \partial_1 \mathbf{A}_i^T \mathbf{S}(\mathbf{A}_i) = \gamma \hat{X}_i \begin{pmatrix} r'(\mathbf{u}_i) \\ \partial_1 a(\mathbf{u}_i, \mathbf{u}_1) \\ \vdots \\ \partial_1 a(\mathbf{u}_i, \mathbf{u}_n) \end{pmatrix} \mathbf{S}(\mathbf{u}_i) \\ \mathbf{I}_i &= \gamma \begin{pmatrix} 0 \\ \hat{X}_1 \partial_2 a(\mathbf{u}_i, \mathbf{u}_1) \mathbf{S}(\mathbf{u}_1) \\ \vdots \\ \hat{X}_n \partial_2 a(\mathbf{u}_i, \mathbf{u}_n) \mathbf{S}(\mathbf{u}_n) \end{pmatrix} \\ \frac{d\mathbf{A}_i}{dt} &= \mathbf{D}_i + \mathbf{I}_i. \end{aligned}$$

[TODO:  $\mathbf{S}$  is a column vector,  $\mathbf{A}_i$  is a column vector,  $\partial_1 \mathbf{A}_i$  is either a column or a rectangular matrix, etc.]

The direct effect is analogous to the constrained selection gradient of a constrained phenotype  $\mathbf{p}$ . It must be within  $90^\circ$  of the selection gradient  $\mathbf{S}(\mathbf{A}_i)$ , for the same reasons as in the constrained phenotype case.

Here is a summary of the different components of adaptive motion of Lotka-Volterra coefficients:

- Given  $n$  populations, there are  $n$  intrinsic growth terms  $r_i$  and  $n^2$  interaction terms  $a_{ij}$ .
- We arrange those values into  $n$  “interaction phenotypes”  $\mathbf{A}_i$ .
- For each  $\mathbf{A}_i$  there is a selection gradient  $\mathbf{S}(\mathbf{A}_i)$ , which is the direction of greatest increase in the invasion fitness of population  $i$ , when we vary population  $i$  as a rare invading population with all resident populations’ phenotypes held fixed (including population  $i$ ), that is, as the first argument to  $k(\mathbf{u}_i)$  and  $a(\mathbf{u}_i, \mathbf{u}_j)$ , the patient.
- Adaptive change in the  $i$ th population must be in a direction that would move  $\mathbf{A}_i$  within  $90^\circ$  of its selection gradient  $\mathbf{S}(\mathbf{A}_i)$ , considering  $i$  as an invader with all resident populations held fixed.



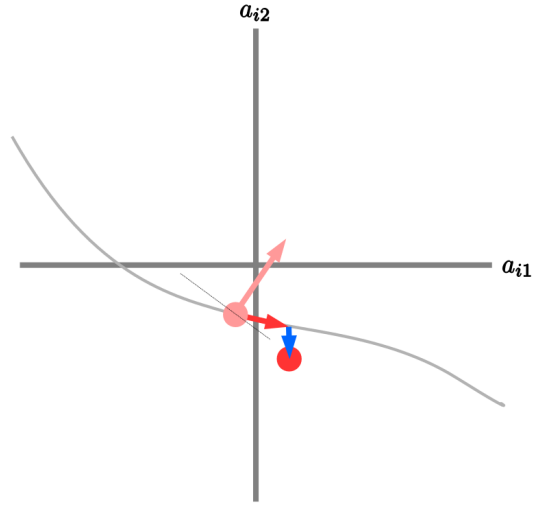


Figure 7: If we assume for the sake of illustration that  $r$  does not vary, and that there are only two populations, we can visualize the interaction phenotype as a two-dimensional vector  $\begin{pmatrix} a_{i1} \\ a_{i2} \end{pmatrix}$ . Sketched here are the components of motion of  $\mathbf{A}_i$ :  $\mathbf{A}_i$  (red dot),  $\mathbf{S}(\mathbf{A}_i)$  (pink arrow),  $\mathbf{D}_i$  (red arrow),  $\mathbf{I}_i$  (blue arrow), and the resultant motion of  $\mathbf{A}_i$  (red dot). The set of values available to  $\mathbf{A}_i$  as  $\mathbf{u}_i$  varies is the gray curve. The half-plane defined by  $\mathbf{S}(\mathbf{A}_i)$  is above the thin dotted line, and  $\mathbf{D}_i$  is in the direction of motion on the gray curve that is allowed within that half-plane. The indirect effect is not constrained by that curve, and due to the indirect effects of selection on both populations 1 and 2, the vector  $\mathbf{A}_i$  does not generally remain confined to that curve as adaptation proceeds.

- The direction of change in  $\mathbf{A}_i$  due to that change in the  $i$ th population is the “direct effect” of selection on  $\mathbf{A}_i$ ,  $\mathbf{D}_i$ .
- The direction of change in  $\mathbf{A}_i$  due to changes in all the resident populations, considered as agents, i.e. as the second argument to  $a(\mathbf{u}_i, \mathbf{u}_j)$ , is the “indirect effect” of selection on  $\mathbf{A}_i$ ,  $\mathbf{I}_i$ . Note that this includes an indirect effect of population  $i$  on itself.

## Upward bias in direct effects

In general, there is nothing apparent that we can assume about the form of the functions  $r$  and  $a$  and their derivatives.  $\partial_1 a$  and  $\partial_2 a$  may be positive, negative, or zero at any time during the evolution of a community.

However, the above results imply a bias toward a positive sign in the derivatives of  $r$  and  $a$  terms.

The selection gradient  $\mathbf{S}(\mathbf{A}_i)$ , which defines the directions of motion available to  $\mathbf{A}_i$  resulting from selection on  $\mathbf{u}_i$ , is a special vector that is the same for all  $i$ :

$$\mathbf{S}(\mathbf{A}_i) = \begin{pmatrix} 1 \\ \hat{X}_1 \\ \vdots \\ \hat{X}_n \end{pmatrix},$$

where  $\hat{X}_j$  is the long-term average population size of population  $j$ .

It immediately follows that the entries of this vector are all nonnegative. If we assume that extinct populations have been removed from the model, **the entries of the selection gradient  $\mathbf{S}(\mathbf{A}_i)$  are all positive.**

Since the direct effect vector  $\mathbf{D}_i$ , just like the constrained selection vector  $S_c(\mathbf{p}_i)$ , must be within  $90^\circ$  of the selection gradient, the direct effect is confined to directions that are within  $90^\circ$  of this positive vector. While this does not imply that all the entries of  $\mathbf{D}_i$  are positive, it does imply that at least some of them must be positive, and it suggests that in some sense, “all else being equal,” entries of this vector should be more often positive than not.

We can draw no conclusions about the sign of  $\mathbf{I}_i$  one way or another. If we were to assume naively that, sampling over many realistic population biology models, the first and second partial derivatives of the  $a$  function would be distributed around a mean of zero, for lack of a reason to imagine them biased one way or another, we would have to conclude that  $\mathbf{I}_i$  would be unbiased with mean zero and  $\mathbf{D}_i$  would be biased with positive mean entries, and consequently that the motion of  $\mathbf{A}_i$  and therefore the motion of the Lotka-Volterra terms  $r_i$  and  $a_{ij}$  would tend to increase.

Of course, in specific cases, those terms may not increase. These results suggest a framework for examining when they do and do not increase:

- The all-positive entries of the interaction selection gradient can be interpreted in terms of the meaning of a specific model’s parameters, describing what selection “would like to do” if all degrees of freedom were available.
- Constraints on actually available variation in  $r$  and  $a$  terms cause the direct effect to point in a different direction from the interaction selection gradient. This can be interpreted in terms of tradeoffs and constraints imposed by the functional form of  $r$  and  $a$ , as functions of the model’s parameters. Unpacking these constraints can shed light on how model assumptions constrain the evolution of interactions.
- The indirect effects of each population’s adaptation on other populations, and on itself, can change the resulting dynamics of interactions from what they would be due to direct effects alone. These effects can be interpreted in terms of interplays or tensions between evolutionary “self-interest” and impacts of ones change on others.

Understanding the three-part structure of the selection gradient, direct effect, and indirect effect can also suggest ways to investigate how a model would behave differently if certain assumptions were changed.

In the examples in this paper, I attempt to develop these themes by exploring them in relation to familiar model scenarios of resource competition, predation, evolution of mutualism, and the general development of communities defined by Lotka-Volterra matrices.

## Direct and indirect effects in the Lotka-Volterra plane

Lotka-Volterra interactions are classified into competition, predation, and mutualism by the location of the pair  $(a_{ij}, a_{ji})$  on the two-dimensional plane. Here we examine the action of selection on that pair in that plane, which must not be confused with the motion of one population’s “interaction phenotype”,  $\mathbf{A}_i$ , above.

Let  $a_{(i,j)}$  stand for the pair  $\begin{pmatrix} a_{ij} \\ a_{ji} \end{pmatrix}$ .

Let  $D_{ij} = \gamma \hat{X}_i \partial_1 a(\mathbf{u}_i, \mathbf{u}_j) \mathbf{S}(\mathbf{u}_i)$  and  $I_{ij} = \gamma \hat{X}_j \partial_2 a(\mathbf{u}_i, \mathbf{u}_j) \mathbf{S}(\mathbf{u}_j)$  be the components of the direct and indirect effect vectors acting on the term  $a_{ij}$ , so that

$$\frac{da_{ij}}{dt} = D_{ij} + I_{ij}.$$

Let  $D_{(i,j)} = \begin{pmatrix} D_{ij} \\ D_{ji} \end{pmatrix}$  and  $I_{(i,j)} = \begin{pmatrix} I_{ij} \\ I_{ji} \end{pmatrix}$ . Then, of course, [TODO: the dimensions of these matrices seem weirdly wrong]

$$\begin{aligned}
\frac{d}{dt}a_{(i,j)} &= D_{(i,j)} + I_{(i,j)} \\
&= \gamma \begin{pmatrix} \hat{X}_i \partial_1 a(\mathbf{u}_i, \mathbf{u}_j) \partial_1 \mathbf{A}_i^T \mathbf{S}(\mathbf{A}_i) + \hat{X}_j \partial_2 a(\mathbf{u}_i, \mathbf{u}_j) \partial_1 \mathbf{A}_j^T \mathbf{S}(\mathbf{A}_j) \\ \hat{X}_j \partial_1 a(\mathbf{u}_j, \mathbf{u}_i) \partial_1 \mathbf{A}_j^T \mathbf{S}(\mathbf{A}_j) + \hat{X}_i \partial_2 a(\mathbf{u}_j, \mathbf{u}_i) \partial_1 \mathbf{A}_i^T \mathbf{S}(\mathbf{A}_i) \end{pmatrix} \\
&= \gamma \begin{pmatrix} \hat{X}_i \partial_1 a(\mathbf{u}_i, \mathbf{u}_j) \partial_1 \mathbf{A}_i^T + \hat{X}_j \partial_2 a(\mathbf{u}_i, \mathbf{u}_j) \partial_1 \mathbf{A}_j^T \\ \hat{X}_j \partial_1 a(\mathbf{u}_j, \mathbf{u}_i) \partial_1 \mathbf{A}_j^T + \hat{X}_i \partial_2 a(\mathbf{u}_j, \mathbf{u}_i) \partial_1 \mathbf{A}_i^T \end{pmatrix} \begin{pmatrix} 1 \\ \hat{X}_1 \\ \vdots \\ \hat{X}_n \end{pmatrix},
\end{aligned}$$

where the last line makes use of the fact that the interaction selection gradient,  $\mathbf{S}(\mathbf{A}_i)$ , is the same for all  $i$ .

There is no geometrical argument making this vector necessarily positive, and in fact it can be negative. Note, however, that the terms of all the  $a_{(i,j)}$  pairs taken together are the terms of the  $\mathbf{A}_i$  vectors rearranged (without the  $r$  terms), so any positive bias occurring in those terms will also appear here. Another way to look at positive bias here is that the  $\partial_1 a$  values appearing in the above expressions are also entries in the  $\partial_1 \mathbf{A}$  vectors they are multiplying, so there will be some squared values there, skewing the overall value toward the positive.

Alternatively, rather than separating direct from indirect effects, we can look at the motion of  $a_{(i,j)}$  due to change in  $i$  and in  $j$ :

$$\begin{aligned}
\frac{d}{dt}a_{(i,j)} &= a_{(1)(i,j)} + a_{(2)(i,j)} \\
&= \gamma \begin{pmatrix} \hat{X}_i \partial_1 a(\mathbf{u}_i, \mathbf{u}_j) \partial_1 \mathbf{A}_i^T \mathbf{S}(\mathbf{A}_i) \\ \hat{X}_i \partial_2 a(\mathbf{u}_j, \mathbf{u}_i) \partial_1 \mathbf{A}_i^T \mathbf{S}(\mathbf{A}_i) \end{pmatrix} + \gamma \begin{pmatrix} \hat{X}_j \partial_2 a(\mathbf{u}_i, \mathbf{u}_j) \partial_1 \mathbf{A}_j^T \mathbf{S}(\mathbf{A}_j) \\ \hat{X}_j \partial_1 a(\mathbf{u}_j, \mathbf{u}_i) \partial_1 \mathbf{A}_j^T \mathbf{S}(\mathbf{A}_j) \end{pmatrix} \\
&= \gamma \left[ \hat{X}_i \begin{pmatrix} \partial_1 a(\mathbf{u}_i, \mathbf{u}_j) \\ \partial_2 a(\mathbf{u}_j, \mathbf{u}_i) \end{pmatrix} \partial_1 \mathbf{A}_i^T + \hat{X}_j \begin{pmatrix} \partial_2 a(\mathbf{u}_i, \mathbf{u}_j) \\ \partial_1 a(\mathbf{u}_j, \mathbf{u}_i) \end{pmatrix} \partial_1 \mathbf{A}_j^T \right] \mathbf{S}(\mathbf{A}),
\end{aligned}$$

where  $\mathbf{S}(\mathbf{A}) = \mathbf{S}(\mathbf{A}_i) = \mathbf{S}(\mathbf{A}_j)$  is the positive interaction selection gradient.

## Direct effects of constraints on interactions

In the terminology I'm using<sup>2</sup>, the “direct effect” of constraint on evolution is how the restriction to interactions  $\mathbf{A}$  that can be generated by the available variation in phenotypes  $\mathbf{u}$  changes the dynamics of interactions, in comparison to the unconstrained direction of selection on interactions  $\mathbf{S}(\mathbf{A})$ .

Let the “effect of constraint” be that difference:  $C_i = D_i - \gamma \hat{X}_i \mathbf{S}(\mathbf{A}_i)$ , so that the actual motion of the interaction vector can be written

---

<sup>2</sup>Which I may want to change: “direct effect” is used differently in Roughgarden's treatment of ecological evolution

$$\frac{d\mathbf{A}_i}{dt} = \gamma \hat{X}_i \mathbf{S}(\mathbf{A}_i) + C_i + I_i,$$

as the interaction selection gradient plus adjustments due to constraints and due to externalities.

The form of the direct effect is

$$\begin{aligned} D_i &= \gamma \hat{X}_i \partial_1 \mathbf{A}_i \partial_1 \mathbf{A}_i^T \mathbf{S}(\mathbf{A}_i) \\ &= \gamma \hat{X}_i \begin{pmatrix} r_i'^2 & r_i' \partial_1 a_{i1} & \cdots & r_i' \partial_1 a_{in} \\ r_i' \partial_1 a_{i1} & \partial_1 a_{i1}^2 & \cdots & \partial_1 a_{i1} \partial_1 a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ r_i' \partial_1 a_{in} & \partial_1 a_{i1} \partial_1 a_{in} & \cdots & \partial_1 a_{in}^2 \end{pmatrix} \begin{pmatrix} 1 \\ \hat{X}_1 \\ \vdots \\ \hat{X}_n \end{pmatrix} \\ &= \gamma \hat{X}_i \begin{pmatrix} r_i'(r_i' + \sum_j \partial_1 a_{ij} \hat{X}_j) \\ \partial_1 a_{i1}(r_i' + \sum_j \partial_1 a_{ij} \hat{X}_j) \\ \vdots \\ \partial_1 a_{in}(r_i' + \sum_j \partial_1 a_{ij} \hat{X}_j) \end{pmatrix} \end{aligned}$$

and the form of the effect of constraint, then, is

$$\begin{aligned} C_i &= \gamma \hat{X}_i (\partial_1 \mathbf{A}_i \partial_1 \mathbf{A}_i^T \mathbf{S}(\mathbf{A}_i) - \mathbf{S}(\mathbf{A}_i)) \\ &= \gamma \hat{X}_i (\partial_1 \mathbf{A}_i \partial_1 \mathbf{A}_i^T - \mathbf{I}) \mathbf{S}(\mathbf{A}_i) \\ &= \gamma \hat{X}_i \begin{pmatrix} r_i'(r_i' + \sum_j \partial_1 a_{ij} \hat{X}_j) - 1 \\ \partial_1 a_{i1}(r_i' + \sum_j \partial_1 a_{ij} \hat{X}_j) - \hat{X}_1 \\ \vdots \\ \partial_1 a_{in}(r_i' + \sum_j \partial_1 a_{ij} \hat{X}_j) - \hat{X}_n \end{pmatrix}. \end{aligned}$$

Evolution scenarios that have this effect of constraint but no indirect effect of adaptation are familiar in evolutionary theory as **density-dependent selection**, whereas cases with indirect effects are **frequency-dependent selection**.

## Indirect effects of externalities on interactions

### Structure of direct and indirect effects in monomorphic models

#### Direct effect

The direct effect predicts what adaptation in the population would do if there were no frequency-dependent selection, often referred to as density-dependent selection (cite Roughgarden or whatever). Interestingly, when there is only a single species, density-dependent selection can reduce the intrinsic growth  $r$  or

the interaction term  $a$ , but not both. [Todo: is this true if  $u$  is more than one dimensional?]

The  $\mathbf{A}$  vector is  $\begin{pmatrix} r(u) \\ a(u, u) \end{pmatrix}$ , the line of available variation is parallel to the vector  $\partial_1 \mathbf{A} = \begin{pmatrix} r'(u) \\ \partial_1 a(u, u) \end{pmatrix}$ , and the vector of adaptive change due to the direct effect of selection is  $\partial_1 \mathbf{A} \partial_1 \mathbf{A}^T \mathbf{S}(\mathbf{A})$ .

Selection on  $r$  and  $a$  are in *conflict* if increasing one requires decreasing the other. This is the case if the elements of  $\partial_1 A$  are opposite in sign: that is, if

$$\frac{\partial_1 a(u, u)}{r'(u)} < 0.$$

If  $r$  and  $a$  are not in conflict, then the direct effect of selection can never decrease both: since  $\hat{X}$  is positive,

$$\partial_1 \mathbf{A} \partial_1 \mathbf{A}^T \mathbf{S}(\mathbf{A}) = (r'(u) + \hat{X} \partial_1 a(u, u)) \begin{pmatrix} r'(u) \\ \partial_1 a(u, u) \end{pmatrix}$$

is a nonnegative vector.

If  $r$  and  $a$  are in conflict, then the direct effect will increase one at the expense of the other.

The condition for increase in  $r$  is for  $(r'(u) + \hat{X} \partial_1 a(u, u))$  and  $r'(u)$  to have the same sign, which is satisfied if and only if

$$\begin{aligned} \frac{(r'(u) + \hat{X} \partial_1 a(u, u))}{r'(u)} &> 0 \\ 1 + \hat{X} \frac{\partial_1 a(u, u)}{r'(u)} &> 0; \end{aligned}$$

since  $\hat{X} = -r(u)/a(u, u) > 0$ , this is equivalent to

$$\begin{aligned} 1 - \frac{r(u)}{a(u, u)} \frac{\partial_1 a(u, u)}{r'(u)} &> 0 \\ 1 &> \frac{r(u)}{a(u, u)} \frac{\partial_1 a(u, u)}{r'(u)} \\ \frac{a(u, u)}{r(u)} &< \frac{\partial_1 a(u, u)}{r'(u)} \end{aligned}$$

A similar process gives the condition for increase in  $a$ , which can be bypassed by observing that it must be the opposite of the above condition:

$$\frac{\partial_1 a(u, u)}{r'(u)} < \frac{a(u, u)}{r(u)}.$$

Thus, to summarize, there are three cases, determined by the ratio  $\frac{\partial_1 a(u, u)}{r'(u)}$ :

$$\begin{cases} \frac{\partial_1 a(u, u)}{r'(u)} < \frac{a(u, u)}{r(u)} < 0 & r \text{ decreases, } a \text{ increases} \\ \frac{a(u, u)}{r(u)} < \frac{\partial_1 a(u, u)}{r'(u)} < 0 & r \text{ increases, } a \text{ decreases} \\ 0 > \frac{\partial_1 a(u, u)}{r'(u)} & \text{both } r \text{ and } a \text{ increase.} \end{cases}$$

There is no case in which both  $r$  and  $a$  decrease due to direct effects of selection.

### Externalities

Externalities of selection on interactions, represented by the indirect effect vector  $\mathbf{I}$ , act on  $a$  but not on  $r$ . In a monomorphic model, the indirect effect is

$$\mathbf{I} = \gamma \begin{pmatrix} 0 \\ \hat{X} \partial_2 a(u, u) S(u) \end{pmatrix}.$$

The nontrivial term of that expands out to

$$\gamma \hat{X} \partial_2 a(u, u) (r'(u) + \hat{X} \partial_1 a(u, u)),$$

the sign of which depends on the sign of  $\partial_2 a(u, u)$  and the direction of change in  $u$ , which is  $S(u)$ .

Of at least as much interest is the actual change in  $a$  due to the combined direct and indirect effects. The full direct effect of selection on  $a(u, u)$  is

$$\gamma \hat{X} \partial_1 a(u, u) S(u) = \gamma \hat{X} \partial_1 a(u, u) (r'(u) + \hat{X} \partial_1 a(u, u)),$$

and the combined change in  $a$  is

$$\frac{da(u, u)}{dt} = \gamma \hat{X} (\partial_1 a(u, u) + \partial_2 a(u, u)) (r'(u) + \hat{X} \partial_1 a(u, u)).$$

The condition on  $\partial_2 a(u, u)$  for  $a$  to decrease is

$$\begin{aligned} \partial_2 a(u, u) \operatorname{sgn}(r'(u) + \hat{X} \partial_1 a(u, u)) &< -\frac{\partial_1 a(u, u) (r'(u) + \hat{X} \partial_1 a(u, u))}{|r'(u) + \hat{X} \partial_1 a(u, u)|} \\ &< -\partial_1 a(u, u) \operatorname{sgn}(r'(u) + \hat{X} \partial_1 a(u, u)). \end{aligned}$$

This requires the two partials of the  $a$  function to be opposite in sign.

Note, however, that the right hand side is biased toward negative values by the double occurrence of  $\partial_1 a$ , so that this condition typically requires  $\partial_2 a(u, u)$  to be not only opposite in sign to  $\partial_1 a(u, u)$  but additionally to exceed some distance from zero.

## Details of resource competition model

## Details of predator-prey model

## Details of mutualism model

## Details of “statistical mechanics” model

Note to self:  $a(u_i, u_j) = \frac{u_i - 2u_j}{10} + \sin(u_i - 2u_j)$  might work as an  $a$  function for this

Or to choose  $\theta, \rho$  randomly and construct  $a$  s.t. those are the directions of  $\partial_1 a, \partial_2 a$ ? That is, with  $\mathbf{u} \in \mathbb{R}^2$ ,  $\partial_1 a(\mathbf{u}_i, \mathbf{u}_j) = c(\cos \theta, \sin \theta)$ ,  $\partial_2 a(\mathbf{u}_i, \mathbf{u}_j) = d(\cos \rho, \sin \rho)$ .

So maybe  $a(\mathbf{u}_i, \mathbf{u}_j) = b + c(\cos \theta, \sin \theta) \mathbf{u}_i + d(\cos \rho, \sin \rho) \mathbf{u}_j$ ? Will need to draw a separate  $(b, c, d, \theta, \rho)$  for each  $i$ - $j$  pair, with angles uniformly from  $[0, 2\pi)$ , scalars maybe from  $N(0, 1)$ ?

I think I’ll also need to do a community-assembly process to get a bunch of them to coexist, then start the adaptive dynamics.

But no, not with  $\mathbf{u} \in \mathbb{R}^2$ , I think what I wanted was for  $\theta$  and  $\rho$  to be the directions of  $a_{(i,j)}$  as  $i$  and  $j$  shift. To do that I would want  $u \in \mathbb{R}$  and  $(\partial_1 a(u_i, u_j), \partial_2 a(u_j, u_i)) = c(\cos \theta, \sin \theta)$  and the other thing for  $\rho$ .

So with  $i$  and  $j$  sorted, for the  $(i, j)$  pair choose  $\theta, \rho$ , etc., and set

$$\begin{aligned} a(u_i, u_j) &= b_{ij} + c_{ij} \cos \theta_{ij} u_i + d_{ij} \cos \rho_{ij} u_j \\ a(u_j, u_i) &= b_{ji} + c_{ij} \sin \theta_{ij} u_i + d_{ij} \sin \rho_{ij} u_j. \end{aligned}$$

Then

$$\begin{aligned} \begin{pmatrix} \partial_1 a(u_i, u_j) \\ \partial_2 a(u_j, u_i) \end{pmatrix} &= c_{ij} \begin{pmatrix} \cos \theta_{ij} \\ \sin \theta_{ij} \end{pmatrix} \\ \begin{pmatrix} \partial_2 a(u_i, u_j) \\ \partial_1 a(u_j, u_i) \end{pmatrix} &= d_{ij} \begin{pmatrix} \cos \rho_{ij} \\ \sin \rho_{ij} \end{pmatrix}, \end{aligned}$$

as desired.

## Bibliography

Dawkins, R., and J. R. Krebs. 1979. “Arms Races Between and Within Species.” *Proceedings of the Royal Society of London. Series B, Biological Sciences* 205 (1161): 489–511.

Ma, J., L. Worden, and S. A. Levin. 2007. “Evolutionary Branching of Single Traits.” In *From Energetics to Ecosystems: the Dynamics and Structure of*



*Ecological Systems*, edited by N. Rooney, K. S. McCann, and D. L. G. Noakes, 191–212. Springer.

Metz, J. A. J., S. A. H. Geritz, G. Meszéna, F. J. A. Jacobs, and J. S. van Heerwaarden. 1996. “Adaptive Dynamics: a Geometrical Study of the Consequences of Nearly Faithful Reproduction.” In *Stochastic and Spatial Structures of Dynamical Systems*, edited by S. J. Van Strien and S. M. Verduyn Lunel, 183–231. Amsterdam: North Holland.

Tilman, D. 1982. *Resource Competition and Community Structure*. Monographs in Population Biology 17. Princeton University Press.

Worden, L., and S. A. Levin. 2007. “Evolutionary Escape from the Prisoner’s Dilemma.” *Journal of Theoretical Biology* 245: 411–422.