
Derivation of Results and Description of Simulation Procedure

A1: Conditions for an Evolutionarily Stable Strategy (ESS)

Here we derive conditions for a defense strategy to be singular (condition [4a]) and for a singular strategy to be evolutionarily stable (condition [4b]). The mutant fitness function is \( W_2(D, \bar{A}) = F(D) \int_0^{\frac{1}{t}} \exp \left\{ -F \left( 1 - S(D) \right) \right\} \Phi(dP) \), where attack rate \( \bar{A} \) is a function of the resident strategy \( \bar{D} \). We maintain the shorthand notation adopted in the text and write \( W_2(D, \bar{A}) = F(D)H(t) \), where \( t = - (1 - S(D)) \).

A strategy \( D' \) is singular if \( \frac{\partial}{\partial D} W_2(D, \bar{A}) |_{D = \bar{D} = D'} = 0 \). Writing out the derivative, we obtain

\[
\frac{F'(D')}{F(D')S'(D')} = - \frac{\bar{A}H(t)}{H(t)} \bigg|_{D = \bar{D} = D'},
\]

which is equivalent to condition (4a) in the text: \( \left[ \ln F(D') \right]'S'(D') = -\bar{A}\left[ \ln H(t) \right]' |_{D = \bar{D} = D'} \).

A singular strategy \( D' \) is also evolutionarily stable if \( \frac{\partial}{\partial D} W_2(D, \bar{A}) |_{D = \bar{D} = D'} < 0 \), which expands to

\[
\left[ H(t) \left( \frac{2F(D)S(D)'\bar{A}H(t)}{H(t)} + F(D)S(D)'\bar{A}H(t) \right) + F(D)'S(D)'H(t) \right] |_{D = \bar{D} = D'} < 0.
\]

Using that \( H(t) > 0 \) and that \( H'(t)/H(t) - (H'(t)/H(t))^2 = [\ln H(t)]'' \), this condition becomes

\[
\left[ F''(D) + 2F'(D)S(D)'\bar{A}H(t) / H(t) + F(D)S(D)'\bar{A}H(t) / H(t) + S(D)'F(D) \left( \frac{\bar{A}H(t)}{H(t)} \right)^2 + \bar{A}^2 [\ln H(t)]'' \right] |_{D = \bar{D} = D'} < 0.
\]

Using condition (A1) to simplify, the condition for evolutionary stability reduces to the form presented in condition (4b) in the text:

\[
F''(D') - \frac{2F'(D')^2}{F(D')} - \frac{F'(D')S''(D')}{S'(D')} + \frac{F''(D')^2}{F(D')} + S'(D')^2F(D')[\bar{A}^2[\ln H(t)]'' |_{D = \bar{D} = D'} < 0
\]

\[
\Leftrightarrow \frac{1}{S'(D')} \left( \frac{F'(D')}{F(D')} \right)^2 - \frac{F''(D')^2}{F(D')S'(D')} < -\bar{A}^2[\ln H(t)]'' |_{D = \bar{D} = D'}
\]

\[
\Leftrightarrow \frac{1}{S'(D')} \left[ \left( \ln F(D') \right)' \right]'' < -\bar{A}^2[\ln H(t)]'' |_{D = \bar{D} = D'}.
\]

A2: Convergence Stability of Monomorphic ESSs

Here we show that condition (5) in the text,

\[
[\ln H(t)]' + t[\ln H(t)]'' \geq 0, \text{ for all } t < 0,
\]

(A2)
is a sufficient condition for convergence stability of a monomorphic ESS. Notation is still such that \( H(t) = \int_0^t \exp[-t\bar{P}\Lambda(1 - S(D))]\Phi(P)dP, \ t = -(1 - S(D))\bar{A}, \bar{A} = A(\hat{D}), \) and \( W_2(D, \bar{A}) = F(D)H(t). \)

A sufficient condition for convergence stability of a singular strategy \( D' \) in the interior is (Eshel 1983; Christiansen 1991) \((\partial^3/\partial D\partial \hat{D})W_2(D, \bar{A})|_{D = \hat{D} < D'} < 0. \) The first term of this condition is negative at any ESS, and thus a sufficient condition for an ESS to be convergence stable is that

\[
(\partial^2/\partial D\partial \hat{D})W_2(D, \bar{A})|_{D = \hat{D} < D'} \leq 0.
\]

Writing out the cross-derivative and setting \( D = D' \), we obtain

\[
\left[ F(D) \left( \frac{\partial}{\partial D} F(D) H'(t) + \frac{\partial}{\partial D} F(D) \left( \frac{\partial}{\partial D} H(t) + \frac{\partial}{\partial D} H'(t) \right) \right) \right]_{D = \hat{D} < D'} \leq 0.
\]

This can be rearranged and simplified further using that \( A' \leq 0 \) and condition (A1):

\[
\frac{F(D')}{F(D') S'(D')} (S(D') - 1) (H(t))_{D = \hat{D} < D'} + (H(t))_{D = \hat{D} < D'} + (S(D') - 1) (\bar{A}H'(t))_{D = \hat{D} < D'} \geq 0.
\]

Because \( H(t) > 0 \) for all \( t < 0 \), a sufficient condition for \((\partial^2/\partial D\partial \hat{D})W_2(D, \bar{A})|_{D = \hat{D} < D'} \leq 0 \) is that \([\ln H(t)]' + t[\ln H(t)]'' \geq 0 \) for all \( t < 0 \), which is exactly condition (A2).

It is readily shown that condition (A2) holds under gamma and exponential distributions. \( H(t) \) has the form of the moment-generating function (mgf): \( H(t) = \int_0^\infty \exp[\theta t] \Phi(P)dP = E(e^{\theta P}). \)

The mgf of the gamma distribution is \( H(t) = (1 + \theta t)^{-k} \) for \( \theta, k > 0. \) Entering into condition (A2), we obtain \( \lambda \theta(1 - \theta) \geq 0, \) which holds for all \( t. \) Likewise, the mgf of the exponential distribution is \( H(t) = (1 - t/\lambda)^{-1} \) for \( \lambda > 0. \) Under condition (A2), we must then have \( \lambda/t(\lambda - 2) \geq 0, \) which also holds for all \( t. \) For the special case when predator density \( P \) is constant across patches, \( \ln H(t) \) reduces to \( tP, \) for which condition (A2) always holds.

Convergence stability of an ESS on the boundary follows directly from evolutionary stability.

### A3: Ecological Stability of Dimorphic Equilibria

Here we prove that condition (5) in the text (condition [A2] above) is sufficient for the stability of any dimorphic equilibrium and consequently that ecological equilibria must be unique and always comprise mutually invasible strategies.

Consider a pair of defense strategies \( D_1 \) and \( D_2 \) where \( D_1 < D_2. \) Let \( \rho \) denote the frequency of \( D_2 \) at the start of the season, when both strategies are present in the population. The resulting attack rate is \( \bar{A} = (1 - \rho)A(D_1) + \rho A(D_2), \) and, since \( A(D_2) < A(D_1), \) we have that \( (\partial/\partial \rho)\bar{A} < 0. \) At an ecological equilibrium \( \hat{D}, D_1 \) and \( D_2 \) have equal fitness:

\[
W_2(D_1, \bar{A}) = W_2(D_2, \bar{A}) \Rightarrow \frac{F(D_1)}{F(D_2)} = \frac{H(t_2)}{H(t_1)}, \tag{A3}
\]

where \( t_i = -(1 - S(D_i))\bar{A} \) and \( t_1 < t_2. \) Stability of the equilibrium under continuous allele frequency dynamics requires that a small perturbation in \( \rho \) from \( \hat{\rho} \) results in a lower fitness for the strategy that is present at a
marginally higher frequency than at equilibrium. The sufficient and necessary condition is $(\partial / \partial \rho)W_2(D_2, \overline{A})|_{\rho = \rho} > (\partial / \partial \rho)W_2(D_2, \overline{A})|_{\rho = \rho}$, which simplifies to

$$F(D_2)H(t_1)(1 - S(D_2)) > F(D_2)H(t_2)(1 - S(D_2)).$$

(A4)

Using condition (A3) and multiplying with $-\overline{A}$, we can express the necessary condition for a stable equilibrium as $t_1 H'(t_1) / H(t_1) < t_2 H'(t_2) / H(t_2)$, or equivalently $t_1 [\ln H(t_1)]' < t_2 [\ln H(t_2)]'$. This holds for any $t_i < t_2$ if $(\partial / \partial t)([\ln H(t)])' > 0$ for all $t < 0$, which is equivalent to condition (A2). Thus, if (A2) holds, all ecological equilibria involving any two strategies $D_1$ and $D_2$ will be stable. If all ecological equilibria are stable, it follows from an argument of continuity that any equilibrium must also be unique and moreover that ecological equilibria exist if (and only if) the two strategies are mutually invasible.

**A4: Absolute Convergence Stability of Dimorphic ESSs**

Here we show that a dimorphic ESS, comprising strategies $D_1$ and $D_2$, both in the interior, is always absolute convergence stable (sensu Leimar 2001). The following two conditions must be met (Kisdi 2006; Leimar 2001; Matessi and Pasquale 1996):

$$\left[ \frac{\partial^2}{\partial D^2} W_2(D, \overline{A}) + \frac{\partial^2}{\partial D\partial D_j} W_2(D, \overline{A}) \right]_{|\delta, \delta_j} < 0,$$

for $j = 1$ and $j = 2$. \hspace{1cm} (A5)

and

$$\left| \left[ \frac{\partial^2}{\partial D^2} W_2(D, \overline{A}) + \frac{\partial^2}{\partial D\partial D_j} W_2(D, \overline{A}) \right]_{|\delta, \delta_j} \right| > \left| \left[ \frac{\partial^2}{\partial D^2} W_2(D, \overline{A}) + \frac{\partial^2}{\partial D\partial D_j} W_2(D, \overline{A}) \right]_{|\delta, \delta_j} \right|.$$

(A6)

In the case where the cross-derivatives $(\partial^2 / \partial D\partial D_j)W_2(D, \overline{A})$ are 0 for both $j = 1$ and $j = 2$, condition (A5) will hold, since the first term within the brackets is negative at an ESS. In addition, condition (A6) will hold because we are left with 0 on the right-hand side and the product of two negative numbers on the left-hand side. A sufficient condition for absolute convergence stability of a dimorphic ESS is, thus, that all cross-derivatives evaluate to 0. We now show that this is, in fact, the case.

Writing out the cross-derivative $(\partial^2 / \partial D\partial D_j)W_2(D, \overline{A})$, keeping in mind that $\overline{A}$ is a function of $D_j$ and using our shorthand notation, we obtain

$$\frac{\partial}{\partial D_j} \left( F'(D)H(t) + F(D)H'(t) \right) = \frac{\partial}{\partial D_j} \left( F'(D)H(t) + F(D)S(D)\overline{A}H'(t) \right)$$

$$= F'(D)H'(t) \frac{\partial}{\partial D_j} t$$

$$+ F(D)S'(D) \left( H'(t) \frac{\partial}{\partial D_j} \overline{A} + \overline{A}H''(t) \frac{\partial}{\partial D_j} t \right)$$

$$= - (1 - S(D))F'(D)H'(t) \frac{\partial}{\partial D_j} \overline{A}$$

$$+ F(D)S'(D) \left( H'(t) \frac{\partial}{\partial D_j} \overline{A} - (1 - S(D)\overline{A}H''(t) \frac{\partial}{\partial D_j} \overline{A} \right)$$

$$= \frac{\partial}{\partial D_j} \overline{A} \left( F(D)S'(D)(H'(t) - (1 - S(D)\overline{A}H''(t)) \right.$$

$$- (1 - S(D))F'(D)H'(t).$$

3
We have that all cross-derivatives are proportional to \((\partial / \partial \tilde{D}) \tilde{A}\). The attack rate \(\tilde{A}\) in a stable dimorphic population is implicitly defined by \(G(\tilde{A}, \tilde{D}_1, \tilde{D}_2) = W_2(\tilde{D}_2, \tilde{A}) - W_1(\tilde{D}_1, \tilde{A}) = 0\).

Implicit differentiation of \(G\) with respect to \(\tilde{D}_1\) gives

\[
\frac{\partial}{\partial \tilde{A}} \tilde{A} = - \frac{\partial}{\partial \tilde{D}_1} G \left| \frac{\partial}{\partial \tilde{A}} G \right.
\]

\[
= - \frac{\partial}{\partial \tilde{D}_1} W_2(\tilde{D}_1, \tilde{A}) \left[ \frac{\partial}{\partial \tilde{A}} W_1(\tilde{D}_1, \tilde{A}) - \frac{\partial}{\partial \tilde{A}} W_1(\tilde{D}_1, \tilde{A}) \right]
\]

\[
= - \frac{F(\tilde{D}_1)H(t_1) + \tilde{A}S(\tilde{D}_1)H'(t_1)}{F(\tilde{D}_1)H'(t_1)\frac{\partial}{\partial \tilde{A}} \tilde{D}_1 - F(\tilde{D}_2)H'(t_2)\frac{\partial}{\partial \tilde{A}} \tilde{D}_2}
\]

\[
= - \frac{S'(\tilde{D}_1)F(\tilde{D}_1)H(t_1) + \tilde{A}H'(t_1)}{-F(\tilde{D}_1)H'(t_1)(1 - S(D_1)) + F(\tilde{D}_2)H'(t_2)(1 - S(D_2))}
\]

Using condition (A1), it is clear that the numerator of this expression is 0 when \(\tilde{D}_1 = D^*_1\) and from condition (A4) it is clear that the denominator is nonzero at ecological equilibrium. We can conclude that \((\partial / \partial \tilde{D}) \tilde{A} = 0\) for singular strategies at ecological equilibrium. The same holds for \(\tilde{D}_2\). Thus, both cross-derivatives are always 0 at a singular strategy; conditions (A5) and (A6) are satisfied for both component strategies of a dimorphic ESS, which will therefore always be absolute convergence stable.

If the dimorphic ESS includes a strategy on the lower boundary, say \(D_1 = 0\), the sufficient conditions for absolute convergence stability are that \((\partial / \partial \tilde{D}) W_j(\tilde{D}, \tilde{A})\}_{D_1=0} < 0\) and that condition (A5) holds for \(j = 2\). It follows from the above that this will be the case at a dimorphic ESS.

**A5: Invasion Boundaries**

Here we show that the invasion boundaries are linear in the case of constant predator density and are concave when predator density varies between patches.

An invasion boundary for a given resident strategy is defined by all combinations of fecundity \(F\) and survival \(S\) (taken as unconstrained variables) that yield the same expected reproductive output as the resident strategy when the attack environment is set by the resident. Let \(Q\) denote the fitness of the resident strategy of interest, and let \(F_S(S)\) be the decreasing function that satisfies \(F_S(S)H(t) = Q\) for all \(S\), where \(H(t) = \int_0^t \exp\{tP\} \Phi(t) dP\), \(t = -(1 - S)\tilde{A}\tilde{P}\), and \(\tilde{A}\) is set by the resident strategy. Since our trade-off curve is plotted in the \((S, \ln F\) plane, it is natural to consider \(\ln F_S(S)\) as the invasion boundary. Taking the natural logarithm on each side and rearranging, we obtain \(\ln F_S(S) = \ln Q - \ln H(t)\). In the case of constant population density \(P\) across all patches, \(H(t)\) reduces to \(e^P\), and the invasion boundary \(\ln F_S(S)\) becomes linear with slope \(-\tilde{A}\tilde{P}\). When predator density varies between patches, \(\ln H(t)\) is strictly convex in \(t\) (and thus in \(S\)), and it follows that \(\ln F_S(S)\) must be strictly concave.

**A6: Evolutionary Stability of the Randomizing Strategy**

Here we show that the randomizing strategy \(D^* = (D^*_1, D^*_2, \rho^*)\), whose component defense levels equal those of the genetically dimorphic ESS \(D_i\) and \(D^*_2\) (\(0 < D^*_1 < D^*_2\)), is also an ESS when \(\rho^*\) (the probability of inducing \(D^*_2\)) equals the ecological equilibrium frequency of \(D^*_1\) of the genetic dimorphism—that is, \(\rho^* = \tilde{\rho}\).

The fitness of a randomizing strategy \(D = (D_1, D_2, \rho)\) in a resident population that elicits attack rate \(\tilde{A}\) is

\[
U(D, \tilde{A}) = (1 - \rho)W(D_1, \tilde{A}) + \rho W(D_2, \tilde{A}), \quad (A7)
\]

where \(W\) signifies either \(W_1\) or \(W_2\) from the text.
Appendix from T. O. Svennungsen et al., Evolution of Inducible Defenses

The randomization strategy $D^*$ is an ESS if (1) the mutant fitness gradients are 0 at all components and (2) mutant fitness is also concave:

1. 
   a. $\frac{\partial}{\partial D_1} U((D_1, D_2, \rho^*), \bar{A}) \bigg|_{D_1=\rho^*} = 0$
   b. $\frac{\partial}{\partial D_2} U((D_1, D_2, \rho^*), \bar{A}) \bigg|_{D_2=\rho^*} = 0$
   c. $\frac{\partial}{\partial \rho} U((D_1', D_2', \rho), \bar{A}) \bigg|_{\rho=\rho^*} = 0$

2. 
   a. $\frac{\partial^2}{\partial D_1^2} U((D_1, D_2, \rho^*), \bar{A}) \bigg|_{D_1=\rho^*} < 0$
   b. $\frac{\partial^2}{\partial D_2^2} U((D_1, D_2, \rho^*), \bar{A}) \bigg|_{D_2=\rho^*} < 0$
   c. $\frac{\partial^2}{\partial \rho^2} U((D_1', D_2', \rho), \bar{A}) \bigg|_{\rho=\rho^*} < 0$

At a dimorphic ESS, we have that $(\partial/\partial D_i) W(D_i, \bar{A}) \big|_{D_i=\rho^*} = 0$ and $(\partial^2/\partial D_i^2) W(D_i, \bar{A}) \big|_{D_i=\rho^*} < 0$, for $i \in \{1, 2\}$, and that $W(D_i, \bar{A}) = W(D_i, \bar{A})$. Using this, it is readily verified that conditions (1a), (1b), and (1c) and conditions (2a) and (2b) are satisfied. Condition (2c), however, fails because $U(D, \bar{A})$ is linear in $\rho$:

$$(\partial^2/\partial \rho^2) U((D_1', D_2', \rho), \bar{A}) \big|_{\rho=\rho^*} = 0.$$ 

This means that the randomizing strategy $D^* = (D_1', D_2', \rho^*)$ with $\rho^* = \hat{\rho}$ will be neutral against the initial invasion of mutants that deviate slightly in $\rho$. In this case, for the randomizing strategy to be an ESS we must have that the fitness of $D^*$ is higher than the fitness of a mutant $D = (D_1', D_2', \rho^* + \epsilon)$ in a population that includes some small fraction of this mutant (Taylor and Jonker 1978):

$$U((D_1', D_2', \rho^*), \bar{A}_e) > U((D_1', D_2', \rho^* + \epsilon), \bar{A}_e),$$

where $\bar{A}_e$ is the attack rate in the mixed population. For positive $\epsilon$, the frequency of $D_1'$ in the population is increased, and the attack rate is decreased. We write out condition (A8) using condition (A7) and obtain $$(1 - \rho^*) W(D_1', \bar{A}_e) + \rho^* W(D_2', \bar{A}_e) > (1 - \rho^* - \epsilon) W(D_1', \bar{A}_e) + (\rho^* + \epsilon) W(D_2', \bar{A}_e),$$ which reduces to

$$W(D_1', \bar{A}_e) > W(D_2', \bar{A}_e)$$

(A9) when $\epsilon$ is positive. For negative $\epsilon$, the direction of the inequality is opposite:

$$W(D_1', \bar{A}_e) < W(D_2', \bar{A}_e).$$

(A10)

It is readily seen that conditions (A9) and (A10) are equivalent to the stability conditions of an ecological equilibrium (cf. section A3 of the appendix), and thus both conditions (A9) and (A10) hold when $\rho^* = \hat{\rho}$. We can thereby conclude that the randomizing strategy $D^* = (D_1', D_2', \rho^*)$ is an ESS.

A7: Simulation Procedure

To numerically simulate adaptive walks, we first discretize reaction norms by dividing the observational scale into $n$ intervals $C_1, C_2, ..., C_n$, with each interval corresponding to a separate point on the reaction norm. The resident reaction norm is then represented by a $n \times 3$ matrix, with elements in the $k$th row, $(D_{ik}, \hat{D}_{ik}, \hat{\rho}_k)$, pertaining to observations within interval $C_i$. The last interval, $C_n$, is open-ended, with lower boundary $C_n$ satisfying $\int_{C_n} \Phi(P) dP < 10^{-3}$. That is, the probability of a prey animal settling in a patch with the predator density corresponding to an observation from within $C_n$ is less than $10^{-3}$. The probability of observations falling within the $k$th interval is calculated as $\Pr(k|P) = \int_{C_k} \Phi(C|P) dC$, and attack rate (cf. eq. [6] in the text) is calculated as $\bar{A}(D, P) = \sum_{k=1}^{n} (1 - \hat{\rho}_k) A(D_{ik}) + \hat{\rho}_k A(D_{ik}) \Pr(k|P)$. For each observation interval, we assign a corresponding conditional distribution of predator densities, $\phi(P|C)$, whose parameters are in each case based on the midvalue...
of the observational interval except for $C_n$, where parameters are based on the value $C_n + 0.1$. Romberg’s method was used to numerically integrate over $P$, using the transform $u = (P + 1)^{-1}$ for $2^n + 1$ regularly spaced values of $u$ on $[0, 1]$.

In each round of a simulation, we first calculated the gradients (using eqq. [8] in the text) at all entries of the resident reaction norm. A new resident reaction norm was then created by adding small random mutations to all entries of the old resident in the direction of the gradient. Entries with a value on the boundary were left unchanged in cases where the gradient pointed out of the strategy space. The new resident strategy then replaced the old one for the next round. Simulations were run until the gradient was less than $10^{-8}$ at all entries with values from the interior. Simulations were initialized using interval widths of 0.1, and resolution was increased during simulations by further subdividing intervals in regions where the reaction norms were steep or seemed discontinuous (e.g., the interval width in the enlarged part of fig. 5A is 0.005). The default initial resident reaction norm used in the simulations was $D_{kl} = 0, D_{kl} = 0.5, \bar{p}_k = 0.5$ for all $k$. This initialization always converged on a “global” ESS—global in the sense that no strategy that for any $k$ deviates within the boundaries $(0 \leq D_{kl} \leq 1, 0 \leq D_{kl} \leq 1, 0 \leq \bar{p}_k \leq 1)$ can invade. Alternative initializations could lead to convergence on alternative local ESSs in cases where a threshold trait was supported. In cases where the reaction norm was continuous, even simulations initialized with a highly erratic reaction norm (each entry independently drawn from a uniform distribution on $[0, 1]$) converged on the global ESS.

**Literature Cited Only in the Appendix**


