

---

## Evolutionary Dynamics

Through the animal and vegetable kingdoms, nature has scattered the seeds of life abroad with the most profuse and liberal hand; but has been comparatively sparing in the room and nourishment necessary to rear them.

T. R. Malthus

Fifteen months after I had begun my systematic enquiry, I happened to read for amusement “Malthus on Population” . . . it at once struck me that . . . favorable variations would tend to be preserved, and unfavorable ones to be destroyed. Here, then, I had at last got a theory by which to work.

Charles Darwin

Our study of evolutionary dynamics is built around the replicator equations. We begin by defining the replicator dynamic, deriving it in several distinct ways, and exploring its major characteristics (§12.1–§12.6). The next several sections make good on our promise to justify Nash equilibrium in terms of dynamical systems, as we exhibit the relationship between dynamic stability of evolutionary models, on the one hand, and dominated strategies (§12.7), Nash equilibria (§12.8), evolutionarily stable strategies (§12.9), and connected sets of Nash equilibria, on the other. Many of the results we obtain remain valid in more general settings (e.g., when the dynamic has an aggregate tendency toward favoring more fit strategies, but not necessarily as strongly as the replicator dynamic).

We next turn to asymmetric evolutionary games (§12.17), which have the surprising property, a property that is extremely important from the point of view of understanding real-world evolutionary dynamics, that strictly mixed Nash equilibria are never asymptotically stable but are often neutrally stable, leading to generically stable orbits (the Lotka-Volterra model has orbits, but it is not generic, as small changes in the coefficients lead to the equilibrium being either a stable or an unstable focus). In asymmetric games, the limit points of dynamic processes are generally Nash equilibria.

ria, but virtually nothing stronger than this can be asserted, including the elimination of weakly dominated strategies (Samuelson and Zhang 1992).

### 12.1 The Origins of Evolutionary Dynamics

The central actor in an evolutionary system is the *replicator*, which is an entity having some means of making approximately accurate copies of itself. The replicator can be a gene, an organism (defining “accurate copy” appropriately in the case of sexual reproduction), a strategy in a game, a belief, a technique, a convention, or a more general institutional or cultural form. A *replicator system* is a set of replicators in a particular environmental setting with a structured pattern of interaction among agents. An *evolutionary dynamic* of a replicator system is a process of change over time in the frequency distribution of the replicators (and in the nature of the environment and the structure of interaction, though we will not discuss these here), in which strategies with higher payoffs reproduce faster in some appropriate sense.

In addition to having an evolutionary dynamic, evolutionary systems may generate novelty if random errors (“mutations” or “perturbations”) occur in the replication process, allowing new replicators to emerge and diffuse into the population if they are relatively well adapted to the replicator system.

The stunning variety of life forms that surround us, as well as the beliefs, practices, techniques, and behavioral forms that constitute human culture, are the product of evolutionary dynamics.

Evolutionary dynamics can be applied to a variety of systems, but we consider here only two-player *evolutionary games*, which consist of a *stage game* played by pairs of agents in a large population, each “wired” to play some pure strategy. We assume the game is symmetric (§10.1), so the players cannot condition their actions on whether they are player 1 or player 2. In each time period, agents are paired, they play the stage game, and the results determine their rate of replication. We generally assume there is random pairing, in which case the payoff to an agent of type  $i$  playing against the population is given by equation (10.1); note that this assumes that the population is very large, so we treat the probability of an agent meeting his own type as equal to fraction  $p_i$  of the population that uses the  $i$ th pure strategy. More generally, we could assume spatial, kinship, or other patterns of assortment, in which case the probability of type  $i$  meeting

type  $j$  depends on factors other than the relative frequency  $p_j$  of type  $j$  in the population.

There are various plausible ways to specify an evolutionary dynamic. See Friedman (1991) and Hofbauer and Sigmund (1998) for details. Here we discuss only *replicator dynamics*, which are quite representative of evolutionary dynamics in general. Our first task is to present a few of the ways a replicator dynamic can arise.

## 12.2 Strategies as Replicators

Consider an evolutionary game where each player follows one of  $n$  pure strategies  $s_i$  for  $i = 1, \dots, n$ . The play is repeated in periods  $t = 1, 2, \dots$ . Let  $p_i^t$  be the fraction of players playing  $s_i$  in period  $t$ , and suppose the payoff to  $s_i$  is  $\pi_i^t = \pi_i(p^t)$ , where  $p^t = (p_1^t, \dots, p_n^t)$ . We look at a given time  $t$ , and number the strategies so that  $\pi_1^t \leq \pi_2^t \leq \dots \leq \pi_n^t$ .

Suppose in every time period  $dt$ , each agent with probability  $\alpha dt > 0$  learns the payoff to another randomly chosen other agent and changes to the other's strategy if he perceives that the other's payoff is higher. However, information concerning the difference in the expected payoffs of the two strategies is imperfect, so the larger the difference in the payoffs, the more likely the agent is to perceive it, and change. Specifically, we assume the probability  $p_{ij}^t$  that an agent using  $s_i$  will shift to  $s_j$  is given by

$$p_{ij}^t = \begin{cases} \beta(\pi_j^t - \pi_i^t) & \text{for } \pi_j^t > \pi_i^t \\ 0 & \text{for } \pi_j^t \leq \pi_i^t \end{cases}$$

where  $\beta$  is sufficiently small that  $p_{ij} \leq 1$  holds for all  $i, j$ . The expected fraction  $\mathbf{E}p_i^{t+dt}$  of the population using  $s_i$  in period  $t + dt$  is then given by

$$\begin{aligned} \mathbf{E}p_i^{t+dt} &= p_i^t - \alpha dt p_i^t \sum_{j=i+1}^n p_j^t \beta(\pi_j^t - \pi_i^t) + \sum_{j=1}^i \alpha dt p_j^t p_i^t \beta(\pi_i^t - \pi_j^t) \\ &= p_i^t + \alpha dt p_i^t \sum_{j=1}^n p_j^t \beta(\pi_i^t - \pi_j^t) \\ &= p_i^t + \alpha dt p_i^t \beta(\pi_i^t - \bar{\pi}^t), \end{aligned}$$

where  $\bar{\pi}^t = \pi_1^t p_1^t + \dots + \pi_n^t p_n^t$  is the average return for the whole population. If the population is large, we can replace  $\mathbf{E}p_i^{t+dt}$  by  $p_i^{t+dt}$ . Subtracting  $p_i^t$  from both sides, dividing by  $dt$ , and taking the limit as  $dt \rightarrow 0$ , we

get

$$\dot{p}_i^t = \alpha\beta p_i^t(\pi_i^t - \bar{\pi}^t), \quad \text{for } i = 1, \dots, n, \quad (12.1)$$

which is called the *replicator dynamic*. Because the constant factor  $\alpha\beta$  merely changes the rate of adjustment to stationarity but leaves the stability properties and trajectories of the dynamical system unchanged, we often simply assume  $\alpha\beta = 1$  (§12.5).

Several points are worth making concerning the replicator dynamic. First, *under the replicator dynamic, the frequency of a strategy increases exactly when it has above-average payoff*. In particular, this means that the replicator dynamic is not a best-reply dynamic; that is, agents do not adopt a best reply to the overall frequency distribution of strategies in the previous period. Rather, the agents in a replicator system have limited and localized knowledge concerning the system as a whole. Some game theorists call such agents “boundedly rational,” but this term is very misleading, because the real issue is the distribution of information, not the degree of rationality.

Second, if we add up all the equations, we get  $\sum_i \dot{p}_i^t = 0$ , so if  $\sum_i p_i^t = 1$  at one point in time, this remains true forever. Moreover, although a particular replicator can become extinct at  $t \rightarrow \infty$ , a replicator that is not represented in the population at one point in time will never be represented in the population at any later point in time. So, replicator dynamics deal poorly with innovation. A more general system adds a term to the replicator equation expressing the spontaneous emergence of replicators through mutation.

Third, our derivation assumes that there are no “mistakes;” that is, players never switch from a better to a worse strategy. We might suspect that small probabilities of small errors would make little difference, but I do not know under what conditions this intuition is valid.

Note that taking expected values allows us to *average* over the possible behaviors of an agent, so that even if there is a positive probability that a player will switch from better to worse, on average the player will not. The replicator dynamic compels a dynamical system *always* to increase the frequency of a strategy with above average payoff. If we do *not* take expected values, this property fails. For instance, if there is a probability  $p > 0$ , no matter how small, that a player will go from better to worse, and if there are  $n$  players, then there is a probability  $p^n > 0$  that *all* players will switch from better to worse. We would have a “stochastic dynamic” in which movement over time probably, but not necessarily, increases the

frequency of successful strategies. If there is a single stable equilibrium, this might not cause much of a problem, but if there are several, such rare accumulations of error will inevitably displace the dynamical system from the basin of attraction of one equilibrium to that of another (see chapter 13).

It follows that the replicator dynamic, by abstracting from stochastic influences on the change in frequency of strategies, is an idealized version of how systems of strategic interaction develop over time, and is accurate only if the number of players is very large in some appropriate sense, compared to the time interval of interest. To model the stochastic dynamic, we use stochastic processes, which are Markov chains and their continuous limits, diffusion processes. We provide an introduction to such dynamics in chapter 13.

It is satisfying that as the rate of error becomes small, the deviation of the stochastic dynamic from the replicator dynamic becomes arbitrarily small with arbitrarily high probability (Freidlin and Wentzell 1984). But the devil is in the details. For instance, as long as the probability of error is positive, under quite plausible conditions a stochastic system with a replicator dynamic will make regular transitions from one asymptotically stable equilibrium to another, and superior mutant strategies may be driven to extinction with high probability; see chapter 13, as well as Foster and Young (1990) and Samuelson (1997) for examples and references.

### 12.3 A Dynamic Hawk-Dove Game

There is a desert that can support  $n$  raptors. Raptors are born in the evening and are mature by morning. There are always at least  $n$  raptors alive each morning. They hunt all day for food, and at the end of the day, the  $n$  raptors that remain search for nesting sites (all raptors are female and reproduce by cloning). There are two types of nesting sites: good and bad. On a bad nesting site, a raptor produces an average of  $u$  offspring per night, and on a good nesting site, she produces an average of  $u + 2$  offspring per night. However, there are only  $n/2$  good nesting sites, so the raptors pair off and vie for the good sites.

There are two variants of raptor: *hawk raptors* and *dove raptors*. When a dove raptor meets another dove raptor, they do a little dance and with equal probability one of them gets the good site. When a dove raptor meets a hawk raptor, the hawk raptor takes the site without a fight. But when two hawk raptors meet, they fight to the point that the expected number of offspring for each is one less than if they had settled for a bad nesting site. Thus the payoff to the two “strategies” hawk and dove are as shown in the diagram.

	Hawk	Dove
Hawk	$u - 1$ $u - 1$	$u + 2$ $u$
Dove	$u$ $u + 2$	$u + 1$ $u + 1$

Let  $p$  be the fraction of hawk raptors in the population of  $n$  raptors. We assume  $n$  is sufficiently large that we can consider  $p$  to be a continuous variable, and we also assume that the number of days in the year is sufficiently large that we can treat a single day as an infinitesimal  $dt$  of time. We can then show that *there is a unique equilibrium frequency  $p^*$  of hawk raptors and the system is governed by a replicator dynamic.*

In time period  $dt$ , a single dove raptor expects to give birth to

$$f_d(p)dt = (u + 1 - p)dt$$

little dove raptors overnight, and there are  $n(1 - p)$  dove raptors nesting in the evening, so the number of dove raptors in the morning is

$$n(1 - p)(1 + (u + 1 - p)dt) = n(1 - p)(1 + f_d(p)dt).$$

Similarly, the number of hawk raptors in the evening is  $np$  and a single hawk raptor expects to give birth to

$$f_h(p)dt = (u + 2(1 - p) - p)dt$$

little hawk raptors overnight, so there are

$$np(1 + (u + 2(1 - p) - p)dt) = np(1 + f_h(p)dt)$$

hawk raptors in the morning. Let

$$f(p) = (1 - p)f_d(p) + pf_h(p),$$

so  $f(p)dt$  is the total number of raptors born overnight and  $n(1 + f(p)dt)$  is the total raptor population in the morning. We then have

$$p(t + dt) = \frac{np(t)(1 + f_h(p)dt)}{n(1 + f(p)dt)} = p(t) \frac{1 + f_h(p)dt}{1 + f(p)dt},$$

which implies

$$\frac{p(t + dt) - p(t)}{dt} = p(t) \left\{ \frac{f_h(p) - f(p)}{1 + f(p)dt} \right\}.$$

If we now let  $dt \rightarrow 0$ , we get

$$\frac{dp}{dt} = p(t)[f_h(p) - f(p)]. \quad (12.2)$$

This is of course a replicator dynamic, this time derived by assuming that agents reproduce genetically but are selected by their success in playing a game. Note that  $p(t)$  is constant (that is, the population is in equilibrium) when  $f_h(p) = f(p)$ , which means  $f_h(p) = f_d(p) = f(p)$ .

If we substitute values in equation (12.2), we get

$$\frac{dp}{dt} = p(1 - p)(1 - 2p). \quad (12.3)$$

This equation has three fixed points:  $p = 0, 1, 1/2$ . From our discussion of one-dimensional dynamics (§11.8), we know that a fixed point is asymptotically stable if the derivative of the right-hand side is negative, and is unstable if the derivative of the right-hand side is positive. It is easy to check that the derivative of  $p(1 - p)(1 - 2p)$  is positive for  $p = 0, 1$  and negative for  $p = 1/2$ . Thus, a population of all dove raptors or all hawk raptors is stationary, but the introduction of even one raptor of the other type will drive the population toward the heterogeneous asymptotically stable equilibrium.

## 12.4 Sexual Reproduction and the Replicator Dynamic

Suppose the *fitness* (that is, the expected number of offspring) of members of a certain population depends on a single *genetic locus*, at which there are two genes (such creatures, which includes most of the “higher” plants and animals, are called *diploid*). Suppose there are  $n$  alternative types of genes (called *alleles*) at this genetic locus, which we label  $g_1, \dots, g_n$ . An individual whose gene pair is  $(g_i, g_j)$ , whom we term an “ $ij$ -type,” then has fitness  $w_{ij}$ , which we interpret as being the expected number of offspring surviving to sexual maturity. We assume  $w_{ij} = w_{ji}$  for all  $i, j$ .

Suppose sexually mature individuals are randomly paired off once in each time period, and for each pair  $(g_i, g_j)$  of genes,  $g_i$  taken from the first and  $g_j$  taken from the second member of the pair, a number of offspring of type  $ij$  are born, of which  $w_{ij}$  reach sexual maturity. The parents then die.

**THEOREM 12.1** For each  $i = 1, \dots, n$  let  $p_i(t)$  be the frequency of  $g_i$  in the population. Then, fitness of a  $g_i$  allele is given by  $w_i(t) = \sum_{j=1}^n w_{ij} p_j(t)$ , the average fitness in the population is given by  $w(t) = \sum_{i=1}^n p_i w_i(t)$ , and the following replicator equations hold:

$$\dot{p}_i = p_i [w_i(t) - w(t)] \quad \text{for } i = 1, \dots, n. \quad (12.4)$$

**PROOF:** For any  $i = 1, \dots, n$ , let  $y_i$  be the number of alleles of type  $g_i$ , and let  $y$  be the total number of alleles, so  $y = \sum_{j=1}^n y_j$  and  $p_i = y_i/y$ . Because  $p_j$  is the probability that a  $g_i$  allele will meet a  $g_j$  allele, the expected number of  $g_i$  genes in the offspring of a  $g_i$  gene is just  $\sum_j w_{ij} p_j$ , and so the total number of  $g_i$  alleles in the next generation is  $y_i \sum_j w_{ij} p_j$ . This gives the differential equation

$$\dot{y}_i = y_i \sum_{j=1}^n w_{ij} p_j.$$

Differentiating the identity  $\ln p_i = \ln y_i - \ln y$  with respect to time  $t$ , we get

$$\frac{\dot{p}_i}{p_i} = \frac{\dot{y}_i}{y_i} - \sum_{j=1}^n \frac{\dot{y}_j}{y} = \sum_{j=1}^n w_{ij} p_j - \sum_{j=1}^n \frac{\dot{y}_j}{y_j} p_j = \sum_{j=1}^n w_{ij} p_j - \sum_{j,k=1}^n w_{jk} p_j p_k,$$

which is the replicator dynamic.

The following important theorem was discovered by the famous biologist R. A. Fisher.

**THEOREM 12.2** Fundamental Theorem of Natural Selection. *The average fitness  $w(t)$  of a population increases along any trajectory of the replicator dynamic (12.4), and satisfies the equation*

$$\dot{w} = 2 \sum_{i=1}^n p_i (w_i - w)^2.$$

Note that the right-hand side of this equation is twice the fitness variance.

**PROOF:** Let  $W$  be the  $n \times n$  matrix  $(w_{ij})$  and let  $p(t) = (p_1(t), \dots, p_n(t))$  be the column vector of allele frequencies. The fitness of allele  $i$  is then

$$w_i = \sum_{j=1}^n w_{ij} p_j,$$

and the average fitness is

$$w = \sum_{i=1}^n p_i w_i = \sum_{i,j=1}^n p_i w_{ij} p_j.$$

Then,

$$\begin{aligned} \dot{w} &= 2 \sum_{i,j=1}^n p_j w_{ji} \dot{p}_i = 2 \sum_{i,j=1}^n p_j w_{ji} p_i (w_i - w) \\ &= 2 \sum_{i=1}^n p_i (w_i - w) w_i = 2 \sum_{i=1}^n p_i (w_i - w)^2, \end{aligned}$$

where the last equation follows from  $\sum_{i=1}^n p_i (w_i - w) w = 0$ . ■

The above model can be extended in a straightforward manner to a situation in which the parents live more than one generation, and the fundamental theorem can be extended to include many genetic loci, provided they do not interact. However, it is a bad mistake to think that the fundamental theorem actually holds in the real world (this is often referred to as the *Panglossian fallacy*, named after Voltaire's Dr. Pangloss, who in *Candide* declared that "all is for the best in this, the best of all possible worlds"). Genes *do* interact, so that the fitness of an allele depends not just on the allele, but on the other alleles in the individual's genetic endowment. Such genes, called *epistatic genes*, are actually quite common. Moreover, the fitness of populations may be *interdependent* in ways that reduce fitness over time (see, for instance, section 11.4, which describes the Lotka-Volterra predator-prey model). Finally, stochastic effects ignored in replicator dynamics can lead to the elimination of very fit genes and even populations.

## 12.5 Properties of the Replicator System

Given the replicator equation (12.1), show the following:

- a. For  $1 \leq i < j \leq n$ , show that

$$\frac{d}{dt} \left( \frac{p_i}{p_j} \right) = \left( \frac{p_i}{p_j} \right) (\pi_i - \pi_j).$$

- b. Suppose that there is an  $n \times n$  matrix  $A = (a_{ij})$  such that for each  $i = 1, \dots, n$ ,  $\pi_i = \sum_j a_{ij} p_j$ ; that is,  $a_{ij}$  is the payoff to player  $i$  when paired with player  $j$  in the stage game. Show that adding a constant to a column of  $A$  does not change the replicator equation and hence does not change the dynamic properties of the system. Note that this allows us to set the diagonal of  $A$  to consist of zeros, or set the last row of  $A$  to consist of zeros, in analyzing the dynamics of the system.
- c. How does the column operation described in the previous question affect the Nash equilibria of the stage game? How does it affect the payoffs?

A more general form of (12.1) is

$$\dot{p}_i^t = a(p, t) p_i^t (\pi_i^t - \bar{\pi}^t) \quad \text{for } i = 1, \dots, n, \quad (12.5)$$

where  $p = (p_1, \dots, p_n)$ ,  $\pi_i^t$  and  $\bar{\pi}^t$  are defined as in (12.1) and  $a(p, t) > 0$  for all  $p, t$ . We will show that for any trajectory  $p(t)$  of (12.5) there is an increasing function  $b(t) > 0$  such that  $q(t) = p(b(t))$  is a trajectory of the original replicator equation (12.1). Thus, multiplying the replicator equations by a positive function preserves trajectories and the direction of time, altering only the time scale.

## 12.6 The Replicator Dynamic in Two Dimensions

Suppose there are two types of agents. When an agent of type  $i$  meets an agent of type  $j$ , his payoff is  $\alpha_{ij}$ ,  $i, j = 1, 2$ . Let  $p$  be the fraction of type 1 agents in the system.

- a. Use section 12.5 to show that we can assume  $\alpha_{21} = \alpha_{22} = 0$ , and then explain why the replicator dynamic for the system can be written

$$\dot{p} = p(1 - p)(a + bp), \quad (12.6)$$

where  $a = \alpha_{12}$  and  $b = \alpha_{11} - \alpha_{12}$ .

- b. Show that in addition to the fixed point  $p = 0$  and  $p = 1$ , there is an interior fixed point  $p^*$  of this dynamical system (that is, a  $p^*$  such that  $0 < p^* < 1$ ) if and only if  $0 < -a < b$  or  $0 < a < -b$ .
- c. Suppose  $p^*$  is an interior fixed point of (12.6). Find the Jacobian of the system and show that  $p^*$  is an asymptotically stable equilibrium if and only if  $b < 0$ , so  $0 < a < -b$ . Show in this case that both of the other fixed points of (12.6) are unstable.

- d. If  $p^*$  is an unstable interior fixed point of (12.6), show that the fixed points  $p = 0$  and  $p = 1$  are both asymptotically stable equilibria.
- e. Show that if  $z = p/(1 - p)$ , then the replicator equation becomes

$$\dot{z} = (1 - p)z(\alpha_{11}z + \alpha_{12}), \quad (12.7)$$

and this has an interior asymptotically stable equilibrium  $z^* = -\alpha_{12}/\alpha_{11}$  if and only if  $\alpha_{11} < 0 < \alpha_{12}$ .

- f. Now use section 12.5 to show that this has the same trajectories as the simpler differential equation

$$\dot{z} = z(\alpha_{11}z + \alpha_{12}). \quad (12.8)$$

Show that the general solution to (12.8) is given by  $z(t) = \alpha_{12}/(ce^{-\alpha_{11}t} - \alpha_{11})$ , where  $c = \alpha_{12}/z(0) + \alpha_{11}$ . In this case we can verify directly that there is an interior asymptotically stable equilibrium if and only if  $\alpha_{11} < 0 < \alpha_{12}$ .

## 12.7 Dominated Strategies and the Replicator Dynamic

All Nash equilibria of a game survive the iterated elimination of strongly dominated strategies, but not of weakly dominated strategies (see chapter 3). Not surprisingly, strongly dominated strategies do not survive in a replicator dynamic. Suppose there are  $n$  pure strategies in the stage game of an evolutionary game in which  $p_i(t)$  is the fraction of the population playing strategy  $i$  at time  $t$ . Recall that a strategy is *completely mixed* if  $p_i(t) > 0$  for all  $i$ . We have the following theorem.

**THEOREM 12.3** *Let  $p(t) = (p_1(t), \dots, p_n(t))$  be a completely mixed trajectory of the replicator dynamic (12.1) and suppose strategy  $i$  is recursively strongly dominated (§4.1). Then, strategy  $i$  does not survive the replicator dynamic; that is,  $\lim_{t \rightarrow \infty} p_i(t) = 0$ .*

To see this, first suppose  $i$  is strongly dominated by  $p^0$ . We write  $\pi(p, q)$  for the payoff to strategy  $p$  against strategy  $q$ . Then,  $\pi(p^0, p) > \pi_i(p)$  for all mixed strategies  $p$ . Because the set of mixed strategies is closed and bounded,  $\epsilon = \min_p(\pi(p^0, p) - \pi_i(p))$  is strictly positive. Let

$$f(p) = \ln(p_i) - \sum_{j=1}^n p_j^0 \ln(p_j).$$

It is easy to check that  $df(p(t))/dt \leq -\epsilon$ , so  $p_i(t) \rightarrow 0$ .

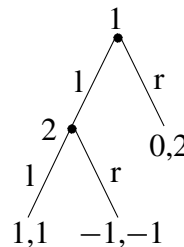
It seems likely that this proof can be extended to the case of iterated domination. For instance, suppose strategy  $j$  is not strongly dominated until strongly dominated strategy  $i$  has been eliminated. By the preceding argument, when  $t$  is sufficiently large,  $i$  is used with vanishingly small probability, so now  $j$  is strongly dominated, and hence we can apply the preceding argument to  $j$ . And so on. The theorem is proved for the case of strategies that are iteratively strongly dominated by pure strategies in Samuelson and Zhang (1992). The case of strategies strongly dominated by mixed strategies is also treated in Samuelson and Zhang (1992), but a stronger condition on the dynamic, which they call *aggregate monotonic*, is needed to ensure elimination.

What about weakly dominated strategies? If the pure strategies against which a weakly dominated strategy does poorly are *themselves* driven out of existence by a replicator dynamic, then the weakly dominated strategy may persist in the long run. However, we do have the following two theorems.

**THEOREM 12.4** *Let  $p(t) = (p_1(t), \dots, p_n(t))$  be a completely mixed trajectory of the replicator dynamic (12.1), and suppose  $p(t)$  converges to a limit  $p^*$  as  $t \rightarrow \infty$ . Then  $p^*$  cannot assign unitary probability to a weakly dominated strategy.*

**THEOREM 12.5** *Let  $p(t) = (p_1(t), \dots, p_n(t))$  be a completely mixed trajectory of the replicator dynamic (12.1), and let  $\alpha_{ij}$  be the payoff of pure strategy  $s_i$  against  $s_j$ . Suppose pure strategy  $s_i$  is weakly dominated by  $s_k$ , so  $\pi(s_k, s_j) > \pi(s_i, s_j)$  for some pure strategy  $s_j$ . Suppose  $\lim_{t \rightarrow \infty} p_j(t) > 0$ . Then,  $\lim_{t \rightarrow \infty} p_i(t) = 0$ .*

It is worthwhile thinking about the implications of this theorem for the persistence of a non-subgame-perfect Nash equilibrium under a replicator dynamic. Consider the little game in the diagram. Clearly, there are two Nash equilibria. The first is (1,1), where each player gets 1. But if player 2 is greedy, he can threaten to play r, the best response to which on the part of player 1 is r. Thus, (r,r) is a second Nash equilibrium. This equilibrium is, however, not subgame perfect, because player 2's threat of playing r is not credible.



- a. Construct the normal form game and show that strategy r for player 2 is weakly dominated by strategy l.
- b. Write the replicator equations for this system and find the fixed points of the replicator dynamic. Show that the replicator equations are

$$\dot{\alpha} = \alpha(1 - \alpha)(2\beta - 1) \tag{12.9}$$

$$\dot{\beta} = -2\beta(1 - \beta)(1 - \alpha). \tag{12.10}$$

Note that the underlying game is not symmetric in this case.

- c. Find the Jacobian of the dynamical system at each of the fixed points, draw a phase diagram for the system, and show that any trajectory that does not start at a fixed point tends to the subgame perfect equilibrium as  $t \rightarrow \infty$ . Compare your results with the phase diagram in figure 12.1.
- d. If you are inquisitive, study some other non-subgame-perfect Nash equilibria of various games and try to generalize as to (1) the relationship between non-subgame perfection and weakly dominated strategies, and (2) the conditions under which a non-subgame perfect Nash equilibrium can persist in a replicator dynamic.

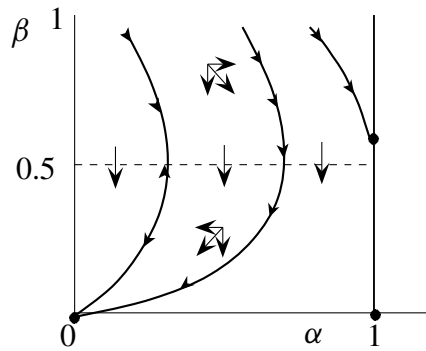


Figure 12.1. Phase diagram for dynamics of equations (12.9) and (12.10)

### 12.8 Equilibrium and Stability with a Replicator Dynamic

Consider an evolutionary game with  $n$  pure strategies and stage game payoff  $\pi_{ij}$  to an  $i$ -player who meets a  $j$ -player. If  $p = (p_1, \dots, p_n)$  is the frequency of each type in the population, the expected payoff to an  $i$ -player is then  $\pi_i(p) = \sum_{j=1}^n p_j \pi_{ij}$ , and the average payoff in the game

is  $\bar{\pi}(p) = \sum_{i=1}^n p_i \pi_i(p)$ . The replicator dynamic for this game is then given by

$$\dot{p}_i = p_i(\pi_i(p) - \bar{\pi}(p)). \quad (12.11)$$

We are now at the point of motivating the importance of the Nash equilibrium as the fundamental equilibrium concept of game theory. We have

**THEOREM 12.6** *The following hold, provided an evolutionary game satisfies the replicator dynamic (12.11).*

- a. *If  $p^*$  is a Nash equilibrium of the stage game,  $p^*$  is a fixed point of the replicator dynamic.*
- b. *If  $p^*$  is not a Nash equilibrium of the stage game, then  $p^*$  is not an evolutionary equilibrium.*
- c. *If  $p^*$  is an asymptotically stable equilibrium of the replicator dynamic, then it is an isolated Nash equilibrium of the stage game (that is, it is a strictly positive distance from any other Nash equilibrium).*

The first of these assertions follows directly from the fundamental theorem of mixed-strategy Nash equilibrium (§3.6). To prove the second assertion, assume  $p^*$  is not isolated. Then, there is an  $i$  and an  $\epsilon > 0$  such that  $\pi_i(p^*) - \bar{\pi}(p^*) > \epsilon$  in a ball around  $p^*$ . But then the replicator dynamic implies  $p_i$  grows exponentially along a trajectory starting at any point in this ball, which means  $p^*$  is not asymptotically stable. The third part is left to the reader.

In general, the converse of these assertions is false. Clearly, there are fixed points of the replicator dynamic that are not Nash equilibria of the evolutionary game, because if an  $i$ -player does not exist in the population at one point in time, it can never appear in the future under a replicator dynamic. Therefore, for any  $i$ , the state  $p_i = 1, p_j = 0$  for  $j \neq i$  is a fixed point under the replicator dynamic.

Also, a Nash equilibrium need not be an asymptotically stable equilibrium of the replicator dynamic. Consider, for instance, the two-player pure coordination game that pays each player one if they both choose  $L$  or  $R$ , but zero otherwise. There is a Nash equilibrium in which each chooses  $L$  with probability  $1/2$ . If  $p$  is the fraction of  $L$ -choosers in the population, then the payoff to an  $L$ -player is  $\pi_L(p) = p$  and the payoff to an  $R$ -player is  $\pi_R(p) = 1 - p$ . The average payoff is then  $\bar{\pi}(p) = p^2 + (1 - p)^2$ , so  $\pi_L(p) - \bar{\pi}(p) = p - p^2 - (1 - p)^2$ . The Jacobian is then  $3 - 4p$ , which is positive at  $p^* = 1/2$ , so the fixed point is unstable. This is of course

intuitively clear, because if there is a slight preponderance of one type of player, then all players gain from shifting to that type.

**12.9 Evolutionary Stability and Asymptotically Stability**

Consider the replicator dynamic (12.11) for the evolutionary game described in section 12.8. We have the following theorem.

**THEOREM 12.7** *If  $p^*$  is an evolutionarily stable strategy of the stage game, then  $p^*$  is an asymptotically stable equilibrium of the replicator dynamic (12.11). Moreover, if  $p^*$  uses all strategies with positive probability, then  $p^*$  is a globally stable fixed point of the replicator dynamic.*

This theorem, which is due to Taylor and Jonker (1978), is proved nicely in Hofbauer and Sigmund (1998:70–71).

The fact that a point is an asymptotically stable equilibrium in a symmetric game does *not* imply that the point is an ESS, however. The diagram represents the stage game of an evolutionary game that has a locally stable fixed point that is not an evolutionarily stable strategy. Show the following:

	$s_1$	$s_2$	$s_3$
$s_1$	2,2	1,5	5,1
$s_2$	5,1	1,1	0,4
$s_3$	1,5	4,0	3,3

- a. The game has a unique Nash equilibrium, in which the three strategies are used in proportions (15/35, 11/35, 9/35).
- b. This Nash equilibrium is not evolutionarily stable, because it can be invaded by the third strategy.
- c. The eigenvalues of the Jacobian of the replicator dynamic equations are  $3(-3 \pm 2i\sqrt{39})/35$ , so the fixed point is a stable focus.

**12.10 Trust in Networks III**

In trust in networks (§6.23), we found a completely mixed Nash equilibrium, which in section 10.8 we found to be evolutionarily unstable, because it could be invaded by trusters. We now show that this equilibrium is in fact globally stable under the replicator dynamic. We illustrate this dynamic in figure 12.2. Note that south of the equilibrium, the fraction of trusters increases, but eventually the path turns back on itself and the fraction of trusters again increases. This is another example of an evolutionary equilibrium that is not an evolutionarily stable strategy: near the equilibrium, the

dynamic path moves away from the equilibrium before veering back toward it.

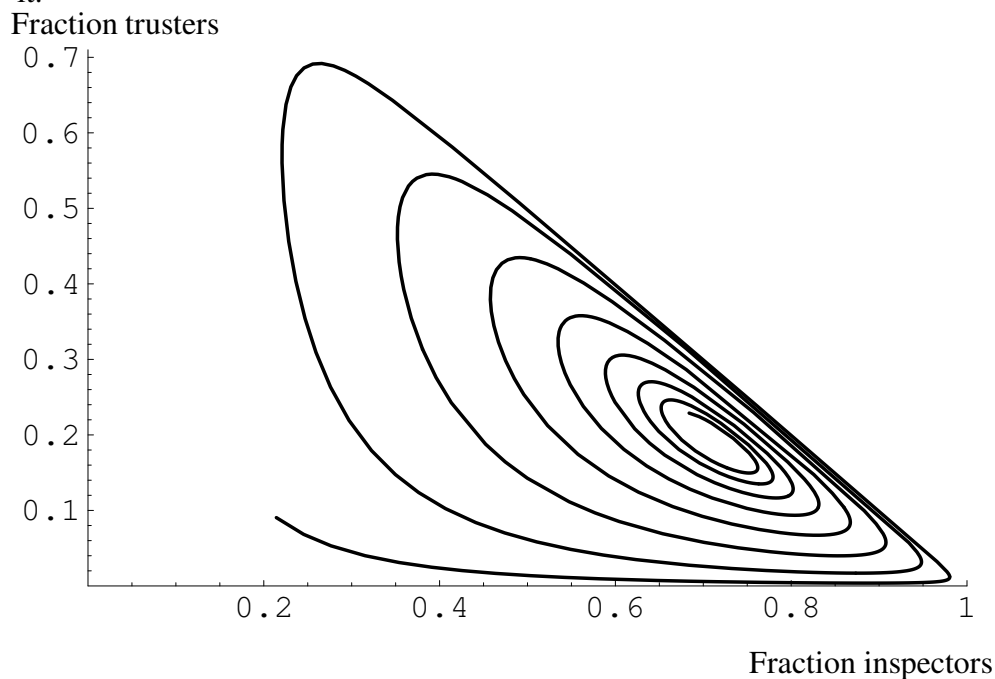


Figure 12.2. A typical path of the trust in networks dynamical system

### 12.11 Characterizing $2 \times 2$ Normal Form Games II

Suppose a normal form game is generic in the sense that no two payoffs for the same player are equal. Suppose  $A = (a_{ij})$  and  $B = (b_{ij})$  are the payoff matrices for Alice and Bob, so the payoff to Alice's strategy  $s_i$  against Bob's strategy  $t_j$  is  $a_{ij}$  for Alice and  $b_{ij}$  for Bob. We say two generic  $2 \times 2$  games with payoff matrices  $(A, B)$  and  $(C, D)$  are *equivalent* if, for all  $i, j = 1, 2$

$$a_{ij} > a_{kl} \equiv c_{ij} > c_{kl}$$

and

$$b_{ij} > b_{kl} \equiv d_{ij} > d_{kl}.$$

In particular, if a constant is added to the payoffs to all the pure strategies of one player when played against a given pure strategy of the other player, the resulting game is equivalent to the original.

Show that equivalent  $2 \times 2$  generic games have the same number of pure Nash equilibria and the same number of strictly mixed Nash equilibria. Show also that every generic  $2 \times 2$  game is equivalent to either the prisoner's dilemma (§3.11), the battle of the sexes (§3.9), or the hawk-dove (§3.10). Note that this list does not include throwing fingers (§3.8), which is not generic. Hint: Refer to the figure in section 3.7. First order the strategies so the highest payoff for player 1 is  $a_1$ . Second, add constants so that  $c_1 = d_1 = b_2 = d_2 = 0$ . Because the game is generic,  $a_1 > 0$ , and either  $a_2 > 0$  (case I) or  $a_2 < 0$  (case II). Third, explain why only the signs of  $c_2$  and  $b_1$ , rather than their magnitudes, remain to be analyzed. If either is positive in case I, the game has a unique equilibrium found by the iterated elimination of dominated strategies, and is equivalent to the prisoner's dilemma. The same is true in case II if either  $b_1 > 0$  or  $c_2 < 0$ . The only remaining case I situation is  $b_1, c_1 < 0$ , which is equivalent to the battle of the sexes, with two pure- and one mixed-strategy equilibrium. The only remaining case II is  $b_1 < 0, c_2 > 0$ , which is equivalent to hawk-dove, and there is a unique mixed-strategy equilibrium.

Now show that if two  $2 \times 2$  normal form games are equivalent, then their corresponding Nash equilibria have the same stability properties. It follows that there are really only three types of generic,  $2 \times 2$  two-player games: (a) a single, stable, pure-strategy equilibrium; (b) a coordination game, with two stable pure-strategy Nash equilibrium, separated by an unstable mixed-strategy equilibrium; and (c) a hawk-dove game, which has a unique, stable Nash equilibrium in mixed strategies. Show that a  $2 \times 2$  two-player ESS is asymptotically stable in the replicator dynamic.

## 12.12 Invasion of the Pure-Strategy Nash Mutants II

In section 10.13 we exhibited a Nash equilibrium that cannot be invaded by any pure strategy mutant but can be invaded by an appropriate mixed-strategy mutant. We can show that this Nash equilibrium is unstable under the replicator dynamic. This is why we insisted that the ESS concept be defined in terms of mixed- rather than pure-strategy mutants; an ESS is an asymptotically stable equilibrium only if the concept is so defined.

To show instability, let  $\gamma$ ,  $\alpha$ , and  $\beta$  be the fraction of agents using strategies 1, 2, and 3, respectively. It is straightforward to check that the replica-

tor equation governing strategies 2 and 3 are given by

$$\begin{aligned}\dot{\alpha} &= \alpha(-\alpha + \beta(a - 1) + (\alpha + \beta)^2 - 2\alpha\beta a) \\ \dot{\beta} &= \beta(-\beta + \alpha(a - 1) + (\alpha + \beta)^2 - 2\alpha\beta a).\end{aligned}$$

The Jacobian of the equations is the zero matrix at the fixed point  $\gamma = 1$ , that is, where  $\alpha = \beta = 0$ . Thus, linearization does not help us. However, we can easily calculate that when  $\alpha = \beta$ ,  $\gamma > 0$ ,  $\dot{\alpha} + \dot{\beta}$  is strictly positive when  $\alpha = \beta$  and  $a > 2$ , so  $\dot{\gamma} < 0$  arbitrarily close to the equilibrium  $\gamma = 1$ . This proves that the Nash equilibrium using strategy 1 is unstable. Note that the Nash equilibrium using strategies 2 and 3 with probability 1/2 is an ESS, and is evolutionarily stable under the replicator dynamic.

Use the same method to check that the two evolutionarily stable equilibria in section 10.14 are asymptotically stable equilibria, and the completely mixed Nash equilibrium, which you showed was resistant to invasion by pure but not mixed strategies, is a saddle point under the replicator dynamic.

### 12.13 A Generalization of Rock, Paper, and Scissors

The game in the diagram, where we assume  $-1 < \alpha < 1$  and  $\alpha \neq 0$ , is a generalization of Rock, Paper, and Scissors in which agents receive a nonzero payoff  $\alpha$  if they meet their own type. We can show that the game has a unique Nash equilibrium in which each player chooses each strategy with probability 1/3, but this equilibrium is not evolutionarily stable for  $\alpha > 0$ . Moreover, we can show that the equilibrium is a hyperbolic fixed point under the replicator dynamic and is a stable focus for  $\alpha < 0$  and an unstable focus for  $\alpha > 0$ .

	<i>R</i>	<i>S</i>	<i>P</i>
<i>R</i>	$\alpha, \alpha$	1, -1	-1, 1
<i>S</i>	-1, 1	$\alpha, \alpha$	1, -1
<i>P</i>	1, -1	-1, 1	$\alpha, \alpha$

### 12.14 *Uta stansburiana* in Motion

Determine the dynamic behavior of the male lizard population in section 6.25 under a replicator dynamic.

**12.15 The Dynamics of Rock, Paper Scissors**

Consider the rock-paper-scissors type game in the diagram, where  $r$  and  $s$  are nonzero. Suppose  $\alpha$ ,  $\beta$ , and  $\gamma = 1 - \alpha - \beta$  are the fraction of the population playing the three strategies, and suppose in each period members are randomly paired and they play the game. What is the replicator dynamic for the game? How does the behavior of the system depend on  $r$  and  $s$ ? Prove the following:

	$\alpha$	$\beta$	$\gamma$
$\alpha$	0,0	$r,s$	$s,r$
$\beta$	$s,r$	0,0	$r,s$
$\gamma$	$r,s$	$s,r$	0,0

Show that when  $r, s < 0$ , this system has three stable pure strategy equilibria, as well as three unstable Nash equilibria using two pure strategies. Then show that rock, paper, and scissors has the mixed-strategy Nash equilibrium  $(\alpha, \beta) = (1/3, 1/3)$  with the following dynamic properties:

- a. The fixed point cannot be a saddle point.
- b. The fixed point is an asymptotically stable equilibrium if  $r + s > 0$  and unstable if  $r + s < 0$ .
- c. The fixed point is a focus if  $r \neq s$ , and a node if  $r = s$ .
- d. If  $r + s = 0$ , as in the traditional Rock, Paper and Scissors game, the fixed point of the linearized system is a center, so the system is not hyperbolic. Thus, we cannot determine the dynamic for this case from the Hartman-Grobman theorem. However, we can show that the fixed point is a center, so all trajectories of the system are periodic orbits.

**12.16 The Lotka-Volterra Model and Biodiversity**

Suppose two species interact in a fixed environment. If  $u$  and  $v$  represent the number of individuals of species  $A$  and  $B$  respectively, the system follows the differential equations

$$\begin{aligned} \dot{u} &= u \left[ a \frac{u}{u+v} + b \frac{v}{u+v} - k(u+v) \right] \\ \dot{v} &= v \left[ c \frac{u}{u+v} + d \frac{v}{u+v} - k(u+v) \right], \end{aligned}$$

where  $k > 0$  and  $(d - b)(a - c) > 0$ . We interpret these equations as follows: the growth rate of each species is frequency dependent, but all share

an ecological niche and hence are subject to overcrowding, the intensity of which is measured by  $k$ . For instance, suppose individuals meet at random. Then, an  $A$  meets another  $A$  with probability  $u/(u + v)$ , and they may reproduce at rate  $a$ , although an  $A$  meets a  $B$  with probability  $v/(u + v)$ , in which case the  $A$  eats the  $B$  ( $b > 0$ ) or vice versa ( $b < 0$ ). Show the following.

- a. Let  $w = u + v$ , the size of the total population, and  $p = u/w$ , the fraction of species  $A$  in the population. The stationary fraction  $p^*$  of species  $A$  is given by

$$p^* = \frac{d - b}{a - c + d - b},$$

which is strictly positive and independent of the crowding factor  $k$ .

- b. If we think of  $w$  as the whole population, then the payoff  $\pi_A$  to species  $A$ , the payoff  $\pi_B$  to species  $B$ , and the mean payoff  $\bar{\pi}$  to a member of the population, *abstracting from the overcrowding factor  $k$* , are given by

$$\begin{aligned}\pi_A &= ap + b(1 - p), \\ \pi_B &= cp + d(1 - p), \\ \bar{\pi} &= p\pi_A + (1 - p)\pi_B.\end{aligned}$$

Show that  $p$  satisfies the replicator dynamic

$$\dot{p} = p(\pi_A - \bar{\pi}),$$

even taking into account the overcrowding factor. This equation indicates that the frequency of species  $A$  in the population is independent of the crowding factor in the dynamic interaction between the two species. Moreover, the stability conditions for  $p$  are also independent of  $k$ , so we conclude: *If neither species can become extinct when the crowding factor  $k$  is low, the same is true no matter how large the crowding factor  $k$ .*

- c. We can generalize this result as follows. Suppose there are  $n$  species, and let  $u_i$  be the number of individuals in species  $i$ , for  $i = 1, \dots, n$ . Define  $w = \sum_j u_j$  and for  $i = 1, \dots, n$  let  $p_i = u_i/w$ , the relative frequency of species  $i$ . Suppose the system satisfies the equations

$$\dot{u}_i = u_i [a_{i1}p_1 + \dots + a_{in}p_n - ku]$$

for  $k > 0$ . We assume the  $\{a_{ij}\}$  are such that there is a positive stationary frequency for each species. Show that the system satisfies the differential equations

$$\frac{\dot{p}_i}{p_i} = \sum_{j=1}^n a_{ij} p_j - \sum_{j,k=1}^n a_{jk} p_j p_k$$

for  $i = 1, \dots, n$ . Show that this represents a replicator dynamic if the payoffs to the various species abstract from the crowding factor  $k$ . Once again we find that the frequency of each species is independent of the crowding factor, and if the ecology is sustainable with low crowding factor (that is, no species goes extinct), then it remains so with high crowding factor.

This result is surprising, perhaps. How do we account for it? It is easy to see that the *absolute* number of individuals in each species in equilibrium is proportional to  $1/k$ . Thus, when  $k$  is large, the justification for using a replicator dynamic is no longer valid: with considerable probability, the stochastic elements abstracted from in the replicator dynamic act to reduce some  $p_i$  to zero, after which it cannot ever recover unless the ecological system is repopulated from the outside. For an example of dynamics of this type, see Durrett and Levin (1994).

### 12.17 Asymmetric Evolutionary Games

Consider two populations of interacting agents. In each time period, agents from one population (row players) are randomly paired with agents from the other population (column players). The paired agents then play a game in which row players have pure strategies  $S = \{s_1, \dots, s_n\}$  and column players have pure strategies  $T = \{t_1, \dots, t_m\}$ . Agents are “wired” to play one of the pure strategies available to them, and the payoffs to an  $i$ -type (that is, a row player wired to play  $s_i$ ) playing a  $j$ -type (that is, a column player wired to play  $t_j$ ) are  $\alpha_{ij}$  for the  $i$ -type and  $\beta_{ij}$  for the  $j$ -type. We call the resulting game an *asymmetric evolutionary game*.

Suppose the frequency composition of strategies among column players is  $q = (q_1, \dots, q_m)$ , where  $q_j$  is the fraction of  $j$ -types among column

players. Then, the payoff to an  $i$ -type row player is

$$\alpha_i(q) = \sum_{j=1}^m q_j \alpha_{ij}.$$

Similarly if the frequency composition of strategies among row players is  $p = (p_1, \dots, p_n)$ , where  $p_i$  is the fraction of  $i$ -types among row players, then the payoff to a  $j$ -type column player is

$$\beta_j(p) = \sum_{i=1}^n p_i \beta_{ij}.$$

We say  $s_i \in S$  is a *best response* to  $q \in Q$  if  $\alpha_i(q) \geq \alpha_k(q)$  for all  $s_k \in S$ , and we say  $t_j \in T$  is a *best response* to  $p \in P$  if  $\beta_j(p) \geq \beta_k(p)$  for all  $t_k \in T$ . A *Nash equilibrium* in an asymmetric evolutionary game is a frequency composition  $p^* \in P$  of row players and  $q^* \in Q$  of column players such that for all  $s_i \in S$ , if  $p_i^* > 0$ , then  $s_i$  is a best response to  $q^*$ , and for all  $s_j \in T$ , if  $q_j^* > 0$ , then  $t_j$  is a best response to  $p^*$ .

Note that there is a natural correspondence between the mixed-strategy Nash equilibria of a two-player normal form game as defined in section 3.4 and the Nash equilibria of an asymmetric evolutionary game. Thus, if we take an arbitrary two-player game in which row players and column players are distinguished and place the game in an evolutionary setting, we get an asymmetric evolutionary game. Hence, the dynamics of asymmetric evolutionary games more or less represent the dynamics of two-player games in general.<sup>1</sup>

A replicator dynamic for an asymmetric evolutionary game is given by the  $n + m - 2$  equations

$$\begin{aligned} \dot{p}_i &= p_i(\alpha_i(q) - \alpha(p, q)) \\ \dot{q}_j &= q_j(\beta_j(p) - \beta(p, q)), \end{aligned} \tag{12.12}$$

<sup>1</sup>I say more or *less* because in fact the assumption of random pairings of agents is not at all characteristic of how agents are paired in most strategic interaction settings. More common is some form of *assortative interaction*, in which agents with particular characteristics have a higher than chance probability of interacting. Assortative interactions, for instance, are a more favorable setting for the emergence of altruism than panmictic interactions.

where  $\alpha(p, q) = \sum_i p_i \alpha_i(q)$ ,  $\beta(p, q) = \sum_j q_j \beta_j(p_i)$ ,  $i = 1, \dots, n - 1$ , and  $j = 1, \dots, m - 1$ . Note that although the *static* game pits the row player against the column player, the *evolutionary* dynamic pits row players against themselves and column players against themselves. This aspect of an evolutionary dynamic is often misunderstood. We see the conflict between a predator and its prey, or between a pathogen and its host, and we interpret the “survival of the fittest” as the winner in this game. But, in fact, in an evolutionary sense predators fight among themselves for the privilege of having their offspring occupy the predator niche in the next period and improve their chances by catching more prey. Meanwhile the prey are vying among themselves for the privilege of having their offspring occupy the prey niche, and they improve their chances by evading predators for an above-average period of time.

What nice properties does this dynamic have? Theorem 12.3 continues to hold: only strategies that are not recursively strongly dominated survive the replicator dynamic. A version of theorem 12.6 also holds in this case: a Nash equilibrium of the evolutionary game is a fixed point under the replicator dynamic, a limit point of a trajectory under the replicator dynamic is a Nash equilibrium, and an asymptotically stable equilibrium of the replicator dynamic is a Nash equilibrium. Even theorem 12.7 continues to hold: an evolutionarily stable strategy is an asymptotically stable equilibrium under the replicator dynamic. However, as we have seen in section 10.16, *an evolutionarily stable strategy of an asymmetric evolutionary game must be a strict Nash equilibrium*; that is, both row and column players must be monomorphic in equilibrium, there being only one type of player on each side. So, in all but trivial cases, evolutionary stability does *not* obtain in asymmetric evolutionary games. Because evolutionary stability is closely related to being an asymptotically stable equilibrium under the replicator dynamic the following theorem (Hofbauer and Sigmund 1998) is not surprising.

**THEOREM 12.8** *A strictly mixed-strategy Nash equilibrium of asymmetric evolutionary games is not an asymptotically stable equilibrium under the replicator dynamic.*

Actually, this situation applies to a much larger class of evolutionary dynamics than the replicator dynamic. See Samuelson and Zhang (1992) for details.

For a simple example of theorem 12.8, consider the case where  $n = m = 2$ ; that is, row and column players each have two pure strategies. We have the following theorem.

**THEOREM 12.9** *In the asymmetric evolutionary game in which each player has two pure strategies, a mixed-strategy Nash equilibrium  $(p^*, q^*)$  is either unstable or an evolutionary focal point. In the latter case, all trajectories are closed orbits around the fixed point, and the time average of the frequencies  $(p(t), q(t))$  around an orbit is  $(p^*, q^*)$ :*

$$\begin{aligned} \frac{1}{T} \int_0^T q(t) dt &= \frac{\alpha}{\gamma} = q^* \\ \frac{1}{T} \int_0^T p(t) dt &= \frac{\beta}{\delta} = p^*. \end{aligned} \tag{12.13}$$

When the time average of a dynamical system equals its equilibrium position, we say the system is *ergodic*.

**PROOF:** Check out the following fact. If a constant is added to each entry in a column of the matrix  $A = \{\alpha_{ij}\}$ , or to each row of the matrix  $B = \{\beta_{ij}\}$ , the replicator equations (12.12) remain unchanged. We can therefore assume  $\alpha_{11} = \alpha_{22} = \beta_{11} = \beta_{22} = 0$ . Writing  $p = p_1$  and  $q = q_1$ , the replicator equations then become

$$\begin{aligned} \dot{p} &= p(1-p)(\alpha - \gamma q) \\ \dot{q} &= q(1-q)(\beta - \delta p), \end{aligned}$$

where  $\alpha = \alpha_{12}$ ,  $\beta = \beta_{12}$ ,  $\gamma = \alpha_{12} + \alpha_{21}$ , and  $\delta = \beta_{12} + \beta_{21}$ . A mixed-strategy equilibrium then occurs when  $0 < \alpha/\gamma, \beta/\delta < 1$ , and is given by  $p^* = \beta/\delta$ ,  $q^* = \alpha/\gamma$ . The Jacobian at the fixed point is

$$J(p^*, q^*) = \begin{bmatrix} 0 & -\gamma p^*(1-p^*) \\ -\delta q^*(1-q^*) & 0 \end{bmatrix}.$$

Note that if  $\alpha$  and  $\beta$  (or equivalently  $\gamma$  and  $\delta$ ) have the same sign, this is a saddle point (theorem 11.4) and hence unstable. You can check that in this case at least one of the monomorphic fixed points is asymptotically stable. Because the mixed-strategy equilibrium is hyperbolic, the fixed point is also a saddle under the replicator dynamic, by the Hartman-Grobman theorem (theorem 11.7). In case this argument whizzed by you, make a phase diagram to get a feel for this very common situation.

If  $\alpha$  and  $\beta$  have opposite signs, the linearized system is neutrally stable, so the mixed-strategy equilibrium is not hyperbolic. Although we cannot apply the Hartman-Grobman theorem, a sketch of the phase diagram shows that trajectories spiral around the fixed point. We can then determine that trajectories are closed orbits by exhibiting a function that is constant on trajectories. To see this, we divide the second replicator equation by the first, getting

$$\frac{dq}{dp} = \frac{(\beta - \delta p)q(1 - q)}{(\alpha - \gamma q)p(1 - p)}.$$

Separating variables, we get

$$\frac{\alpha - \gamma q}{q(1 - q)} dq = \frac{\beta - \delta p}{p(1 - p)} dp.$$

Integrating both sides and simplifying, we get

$$\alpha \ln(q) - (\alpha - \gamma) \ln(1 - q) - \beta \ln(p) + (\beta - \delta) \ln(1 - p) = C,$$

for some constant  $C$ . Suppose  $\alpha > \gamma$ . Then, this function is monotonic in the  $q$  direction, so the spirals must in fact be closed orbits. If  $\alpha \leq \gamma$ , then we must have  $\beta > \delta$ , so the function is monotonic in the  $p$  direction, so again the spirals are closed orbits.

To check the ergodic property of the system in the case of neutral stability, consider a trajectory  $(p(t), q(t))$  starting at a point  $(p(0), q(0)) = (p_0, q_0)$ . We integrate both sides of the equation  $\dot{p}/p(1 - p) = \alpha - \gamma q$  with respect to time, getting

$$\ln(p(t)) + \ln(1 - p(t)) = A + \alpha t - \gamma \int_0^t q(\tau) d\tau,$$

where the constant of integration  $A$  satisfies  $A = \ln(p_0) + \ln(1 - p_0)$ . If the period of the trajectory is  $T$ , so  $p(T) = p_0$  and  $q(T) = q_0$ , then letting  $t = T$  in the previous expression gives

$$\frac{1}{T} \int_0^T q(t) dt = \frac{\alpha}{\gamma} = q^*.$$

A similar argument justifies the second equation in (12.13) as well. This proves the theorem. ■

### 12.18 Asymmetric Evolutionary Games II

To gain some feeling for the argument in section 12.17, check out the dynamic properties of the asymmetric evolutionary game versions of the following games. Hint: In most cases the results follow easily from performing the row and column manipulations that leave zeros on the diagonals of the two payoff matrices.

- Section 12.17 (draw the phase diagram).
- The mixed-strategy equilibrium of the Big John and Little John game (§3.1).
- The mixed-strategy equilibrium of the battle of the sexes (§3.9).

### 12.19 The Evolution of Trust and Honesty

Consider an asymmetric evolutionary game with buyers and sellers. Each seller can be either honest ( $H$ ) or dishonest ( $D$ ), and each buyer can either inspect ( $I$ ) or trust ( $T$ ). Let  $p$  be the fraction of buyers who inspect and let  $q$  be the fraction of sellers who are honest. Suppose the payoff matrix for an encounter between a buyer and a seller is given as in the figure in the diagram. The payoff to inspect is then  $3q + 2(1 - q) = q + 2$ , the payoff to trust is  $4q + (1 - q) = 3q + 1$ , the payoff to be honest is  $2p + 3(1 - p) = -p + 3$ , and the payoff to be dishonest is  $p + 4(1 - p) = -3p + 4$ .

	$H$	$D$
$I$	3,2	2,1
$T$	4,3	1,4

Suppose we have a replicator dynamic, such that the fraction of inspectors grows at a rate equal to its fitness minus the average fitness of buyers. Buyer average fitness is  $p(q + 2) + (1 - p)(3q + 1) = 3q + 1 - p(2q - 1)$ , so the inspector growth rate is  $q + 2 - [3q + 1 - p(2q - 1)] = (1 - p)(1 - 2q)$ , and we have the replicator equation

$$\dot{p} = p(1 - p)(1 - 2q). \quad (12.14)$$

Similarly, the fraction of honest sellers grows at a rate equal to its fitness minus the average fitness among sellers, giving

$$\dot{q} = q(1 - q)(2p - 1). \quad (12.15)$$

- Show that these two coupled differential equations have five fixed points,  $(0,0)$ ,  $(0,1)$ ,  $(1,0)$ ,  $(1,1)$ , and  $(1/2,1/2)$ .

- b. Show that the first four fixed points are unstable.
- c. Show that the equilibrium at  $(1/2, 1/2)$  is not hyperbolic: its linearization is a center. It follows that we cannot use the Hartman-Grobman theorem to ascertain the type of fixed point.
- d. Draw a phase diagram and show that the trajectories are spirals moving counterclockwise around the fixed point.

How might we prove that the fixed point is a center? Suppose we could find a function  $f(p, q)$  that is constant on trajectories of the system. If we could then show that  $f$  is strictly increasing along an appropriate ray from the fixed point to the northeast, we would be done, because only closed orbits are then possible. This is precisely what we did in section 11.4 to show that the trajectories of the Lotka-Volterra equations are orbits around the fixed point. See also sections 12.14 and 12.17.

Eliminating  $t$  from (12.14) and (12.15), we get

$$\frac{dq}{dp} = \frac{(q - q^2)(2p - 1)}{(p - p^2)(1 - 2q)}.$$

Separating the variables, this becomes

$$\frac{1 - 2p}{p - p^2} dp = -\frac{1 - 2q}{q - q^2} dq.$$

Integrating both sides and combining terms, we get  $\ln(p - p^2)(q - q^2) = C$  for some constant  $C$ . We simplify by taking the antilogarithm of both sides, getting  $(p - p^2)(q - q^2) = e^C$ . Thus,  $f(p, q) = p(1 - p)q(1 - q)$  is constant on trajectories of the dynamical system. Consider a ray from the origin through the fixed point. We may parametrize this by  $p = q = s$ , which hits the fixed point when  $s = 1/2$ . Then,  $f(p(s), q(s)) = s^2(1 - s)^2$ , so  $df(p(s), q(s))/ds = 2s(1 - s)(1 - 2s)$ . This is strictly positive for  $1/2 < s < 1$ . If the trajectory were not a center, it would hit this ray more than once, and  $f(p(s), q(s))$  would have a larger value the second time than the first, which is impossible. This proves that  $(1/2, 1/2)$  is a center.