

---

## 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 | $u - 1$<br>$u - 1$ | $u + 2$<br>$u$     |
| Dove | $u$<br>$u + 2$     | $u + 1$<br>$u + 1$ |

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.