

# 4

## Evolution

“such things survived, being organized spontaneously in a fitting way; whereas those which grew otherwise perished and continue to perish. . . .”

Aristotle, *Physics*<sup>1</sup>

### Evolution

The ideas of natural selection and survival of the fittest existed already in Greek philosophy. Aristotle is not describing his own view—he believed in the fixity of the species—but rather a rival theory according to which unsuccessful species go to extinction. Aristotle is referring to Empedocles of Sicily.<sup>2</sup> Empedocles was a statesman and a physician as well as a mystic, philosopher, and poet. His theory was put forward in a long poem, *On Nature*. Empedocles’ account of the origin of species begins with a haphazard combination of parts into a great variety of organisms, only the fittest of which survived. Empedocles influenced Democritus, and both Empedocles and Democritus influenced Lucretius. As Lucretius puts it in his own poem, *On the Nature of Things*:

<sup>1</sup> Aristotle, *Physics* II 8, 198b29.

<sup>2</sup> There are secondary sources, such as Simplicius’ commentary on the foregoing passage in Aristotle’s *Physics*, and Lucretius’ poem. See Sedley 2003b and Campbell 2003 on the connections between Empedocles and Lucretius.

Perforce there perished many a stock, unable:  
By propagation to forge a progeny.

Empedocles even had a theory of how traits are transmitted from generation to generation. Small copies of organs form in the male and female, and in reproduction some from the father and some from the mother combine to form the new organism.

He thus has in hand a rudimentary theory of recombination. Empedocles influenced Hippocrates (probably both directly and through Democritus). Hippocrates' theory of inheritance is remarkably similar to that put forward by Darwin in *The Variation of Plants and Animals under Domestication* nine years after the publication of *The Origin of Species*.<sup>3</sup> Darwin did not know about Hippocrates at the time, but in a letter to William Ogle in 1868, Darwin writes:

I thank you most sincerely for your letter, which is very interesting to me. I wish I had known of these views of Hippocrates before I had published, for they seem almost identical with mine—merely a change of terms—and an application of them to classes of facts necessarily unknown to the old philosopher. The whole case is a good illustration of how rarely anything is new.

Darwin and Hippocrates were wrong about inheritance. But Darwin was right about the broad outlines of the theory of evolution. Traits are inherited by some unknown mechanism. There is some process that produces natural variation in these traits. The traits may affect the ability of the organism to reproduce, and thus the average number of individuals bearing the traits in the next generation. Therefore, those traits that enhance reproductive success increase in frequency in the population, and those that lead to reproductive success below the average

<sup>3</sup> I owe my knowledge of Darwin's theory to my colleague P. Kyle Stanford. See Stanford 2007.

decrease in frequency. The three essential factors in Darwin's account are (i) *natural variation*, (ii) *differential reproduction*, and (iii) *inheritance*.

## Evolutionarily stable strategies

Darwinian processes lead to adaptation to a fixed environment, at least where the genetic mechanism doesn't get in the way.<sup>4</sup> The story is more complicated when fitness depends on the frequencies of different types who interact with one another. Here the fitness landscape may be constantly changing, along with the population proportions. John Maynard Smith, following the lead of William Hamilton,<sup>5</sup> realized that this kind of interactive evolution is a biological version of von Neumann and Morgenstern's Theory of Games.<sup>6</sup>

In 1973, John Maynard Smith and George Price introduced a strengthening of the Nash equilibrium concept of game theory—the concept of an *evolutionarily stable strategy*. The context was the explanation of “limited war” in animal contests. Since hyper-aggressive types, Hawks, defeat peaceful types, Doves, to win resources, why don't they take over the population? The general answer is that selection here is frequency-dependent. If most of the population is occupied by Hawks, they usually interact with each other in fights that lead to serious injury or death. It is only good to be a Hawk if there are enough Doves around to exploit.

Hawk-Dove interactions are modeled as a game. Payoffs for a typical example are shown in the following table, with the numbers

<sup>4</sup> As it does in the case of heterozygote superiority.

<sup>5</sup> “In the way in which the success of a chosen sex ratio depends on choices made by the co-parasitizing females, this problem resembles certain problems discussed in the ‘theory of games.’ In the foregoing analysis a game-like element, of a kind, was present and made necessary the use of the word *unbeatable* to describe the ratio finally established. This word was applied in just the same sense in which it could be applied to the ‘minimax’ strategy of a zero-sum two-person game” (Hamilton 1967).

<sup>6</sup> von Neumann and Morgenstern 1944.

being payoffs (in Darwinian fitness) of row strategy against column strategy:

	Hawk	Dove
Hawk	0	3
Dove	1	2

(In our evolutionary context, payoffs only depend on strategies, not on who is row and who is column. The whole payoff table listing *row payoff*, *column payoff* in each cell looks like this:

	Hawk	Dove
Hawk	0, 0	3, 1
Dove	1, 3	2, 2

In what follows we will use the first, simpler form of giving evolutionary games.)

It is evident that where you are meeting Hawks, it is better to be a Dove (column 1) and where you are meeting Doves (column 2) it is better to be a Hawk. Consequently, a population of All Hawks cannot be evolutionarily stable in that in such a population a few mutant Doves would do better than the natives. Likewise a population of All Doves would be vulnerable to invasion by a few Hawks.

An evolutionarily stable strategy is one such that if the whole population played it, a few mutants would always do worse against the resulting population (including the mutants) than the natives would. Thus the mutants would fade away. If the population is large and individuals are randomly paired to have an interaction there is a simple test for evolutionary stability in terms of the payoffs to the game. A strategy, S, is evolutionarily stable if for any other strategy, M, either:

- (i) Fitness (S played against S) > Fitness (M played against S)  
 or:  
 (ii) Fitnesses are equal against S, but Fitness(S against M) > Fitness(M against M)

This is how evolutionary stability is defined by Maynard Smith and Price.<sup>7</sup>

For instance, in the Hawk–Dove game Hawk is not evolutionarily stable because Fitness (Hawk against Hawk) is less than Fitness (Dove against Hawk). Dove is not evolutionarily stable because Fitness (Dove against Dove) is less than Fitness (Hawk against Dove).

The Maynard Smith–Price test is easily applied to other familiar simple games. For instance, consider the Stag Hunt game. Players can either hunt *Stag* or hunt *Hare*. Hunting Stag is a cooperative enterprise. It fails if both players do not hunt Stag, but it pays off well if they do. Hare hunting is a solitary enterprise. Hare hunters do equally well if the other hunts Hare or Stag, but worse than successful Stag hunters. The Stag Hunt has this kind of payoff structure:

	Hare	Stag
Hare	3	3
Stag	0	4

Applying the test of Maynard Smith and Price, we see that both *Stag* and *Hare* are evolutionarily stable strategies. Stag against Hare does worse than Hare against Hare; Hare against Stag does worse than Stag against Stag. A population of each type is stable against invasion by a few mutants of the other type.

For an example where there is exactly one evolutionarily stable strategy, consider the most widely discussed game theory model in the social sciences, the Prisoner’s Dilemma:

<sup>7</sup> If the first condition is satisfied, mutants are driven out rapidly. If the second condition holds, mutants fade away more slowly.

	Cooperate	Defect
Cooperate	3	1
Defect	4	2

Defect is an evolutionarily stable strategy; cooperate is not.

But what about all the models that explain the evolution of altruism, which is usually taken as cooperation in the Prisoner's Dilemma? All these accounts, in one way or another, explain the evolution of cooperation by some correlation mechanism.<sup>8</sup> Cooperators tend to meet cooperators; defectors tend to meet defectors. Pairing is not random. If pairing is not random the Maynard Smith–Price test of evolutionary stability is *wrong*. This is transparent if correlation is perfect. Then a population of defectors could be invaded by a few mutant cooperators. The cooperators meet each other for a payoff of 3, while the native defectors have a payoff of 2. Correlation can change everything.

## Differential reproduction

Stability is really a dynamic concept. A rest state is *strongly stable* if all states near to it are carried to it by the dynamics. You could think of a marble at the bottom of a bowl. It is just *stable* if states near to it are not carried away by the dynamics. Think of the marble sitting on table top as being stable but not strongly stable. Otherwise it is unstable, like a marble balanced at the top of an inverted bowl. Maynard Smith and Price clearly have in mind something like dynamic stability. Where is the dynamics?

To build a dynamic foundation for the notion of an evolutionarily stable strategy, Taylor and Jonker introduced the *replicator dynamics*.<sup>9</sup> This is a model of differential reproduction in a large

<sup>8</sup> See Bergstrom 2002; Skyrms 1996, 2004.

<sup>9</sup> Taylor and Jonker 1978.

population, where types are inherited with complete fidelity. For simplicity, Mendelian genetics is left out of the picture. Reproduction proceeds as if by cloning.

Replicator dynamics is driven by *Darwinian fitness*—expected number of progeny. If the expected number of progeny of a type is for instance two, then some individuals might have four and some three and some one or zero. But in a large enough population these differences will almost surely average out, and the average number of progeny will equal the expectation. On average, you get what you expect. This gives us replicator dynamics as introduced by Taylor and Jonker to provide a dynamical foundation for evolutionary game theory.

Suppose that reproduction takes place in discrete time—for instance, every spring. What proportion  $x_{\text{new}}(S)$  of the new generation will play a given strategy,  $S$ ? It is just the number who play  $S$  in the new population divided by total number in the population. The number who play  $S$  in the new population is equal to the total number in the old population,  $N$ , multiplied by the proportion who had strategy  $S$ ,  $x_{\text{old}}(S)$ , multiplied by the average number for offspring of those who had strategy  $S$ ,  $\text{Fitness}(S)$ . We have to divide this by the total number of the new population which is just the number of the old population,  $N$ , multiplied by the average number of offspring throughout the old population,  $\text{Average Fitness}$ .

$$x_{\text{new}} = [N x_{\text{old}}(S)\text{Fitness}(S)]/[N \text{Average Fitness}]$$

$N$  drops out and we get  $x_{\text{new}}$  from  $x_{\text{old}}$  by multiplying by a Darwinian success factor:

$$x_{\text{new}} = x_{\text{old}}[\text{Fitness}(S)/\text{Average Fitness}]$$

This is discrete time replicator dynamics. There is an associated (idealized) continuous time replicator dynamics that gives the rate of change of population proportions,  $dx/dt$  at a point in time:

$$dx/dt = x[\text{Fitness}(S) - \text{Average Fitness}]$$

This is what Taylor and Jonker gave us as a simple model of differential reproduction.

What about *cultural evolution*? We want to discuss dynamics of signaling for cultural evolution as well as for biological evolution. There are cases of each, and mixed cases, that are all of interest. We would like a theory of cultural evolution to be more than just a story about how culture evolved. In all honesty, a full theory at this point is out of the question; the cognitive processes involved are too various, complex and poorly understood. The best we can do is to start with a simple basic model that we have some hope of understanding.

One basic process is *imitation*. Suppose that individuals look around them and see which behaviors or strategies are paying off for others, and imitate those strategies with probability proportional to their success. This process and a number of variations on it have been analyzed.<sup>10</sup> What we get, when the population is large and chance fluctuations average out, is just our simple model of differential reproduction—the *replicator dynamics*.<sup>11</sup>

But what is the currency here, in which payoffs are measured? It has to be whatever drives differential imitation. This has to be empirically determined for the context of application. The specific application of the theory derives its content from this determination. The relevant payoffs for cultural evolution may or may not correlate well with Darwinian fitness. In conditions of hardship, both may correlate with eating well and surviving attacks of predators; in conditions of affluence they may be decoupled. Even if the form of the dynamics is the same for biological and cultural evolution the substantive conclusions may be different. Care in interpretation is required.

The replicator dynamics may or may not lead to a dynamical equilibrium (a rest point of the dynamics). If individuals are paired at random and there are just two strategies, it must do so. We can

<sup>10</sup> Björnerstedt and Weibull 1995; Weibull 1995; Schlag 1998.

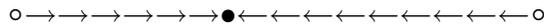
<sup>11</sup> Or some slight variant. This route to the replicator dynamics is even more straightforward, because there is no diploid genetics being suppressed.

visualize the situation by plotting the proportion of one of the strategies on an interval from 0 to 1. We could have:

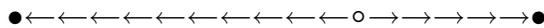
- (i) the dynamics carrying one strategy to fixation, no matter what the interior starting point:



- (ii) the dynamics carrying the population to a mixed state, no matter what the starting point:



- (iii) the dynamics carrying one or the other strategy to fixation, depending on the starting point:



- (iv) the dynamics not moving at all:



Case (i) is exemplified by the Prisoner’s Dilemma. *All cooperate* is a rest point of the dynamics because defectors are all extinct.<sup>12</sup> But it is dynamically unstable. We mark an unstable rest point with a hollow point. All other points are carried to *All Defect*, which is dynamically strongly stable. We mark a strongly stable rest point with a filled circle. With Hawk-Dove, we have case (ii). *All Hawk* and *All Dove* are dynamically unstable. The dynamically stable equilibrium is a mixed (or polymorphic) state of the population with some Hawks and some Doves. The Stag Hunt is case (iii). Here the polymorphic rest state is an unstable “knife-edge.” Any movement off it carries the population to one of the strongly stable equilibria—*All Stag* or *All Hare*.

<sup>12</sup> Differential reproduction by itself does not introduce new types.

For case (iv) consider the *game of inconsequential actions*. Here the available actions have no payoff consequences whatever. I believe that much of life has this structure.

	Do This	Do That
Do This	o	o
Do That	o	o

Replicator dynamics does not move any point. (We can't really draw it.) All points are *Stable* in that nearby points stay nearby, but no state is *strongly stable*.

## The Red Queen

When we have three strategies, however, replicator dynamics may not lead to equilibrium at all! Consider the familiar game of *rock-scissors-paper*. Rock breaks scissors, scissors cuts paper, paper covers rock, so we get the following sort of payoffs:

Rock-scissors-paper			
	R	S	P
R	1	2	0
S	0	1	2
P	2	0	1

This structure is also found outside children's games. Christof Hauert, Silvia de Monte, Josef Hofbauer, and Karl Sigmund find rock-scissors-paper structure in a social dilemma with the possibility of opting out.<sup>13</sup>

The pure social dilemma is a generalization of the Prisoner's Dilemma to many players. Individuals can either choose to contribute to the public good or to free ride. Contributions are

<sup>13</sup> Hauert et al. 2002.

multiplied by the synergy of the joint project, and the joint public good is divided among everyone. If everyone contributes everyone does well. But the multiplier is smaller than the group number, so your dollar contribution gets you personally less than a dollar in return although it can get the group much more. Thus, whatever others do, it is in an agent's own selfish interest to free-ride and share the benefits of others' contributions. If everyone free rides, the public good project fails. There is nothing to distribute, and all do very poorly. Thus we have the n-Person Prisoner's dilemma. To this basic setup is added the possibility of opting out and being a loner. Loners are less successful than those in cooperative groups, but more successful than those in failed public-goods projects. In a population of cooperators, free-riders do better than natives. In a population of free-riders, loners do better. In a population of loners, cooperators do better.

Barry Sinervo and Curtis Lively find rock-scissors-paper structure in mating strategies of side-blotched lizards in California.<sup>14</sup> There are three types of males, which exhibit different coloration. Orange-throated males are very aggressive and guard large territories. Blue-throated males guard smaller territories and are able to guard their mates. Yellow-throated males resemble females, and mate with females on the sly. In a population of mate-guarding blue throats, the ultra-dominant orange throats do better. But they can be invaded by yellow-throated sneakers. And these can be invaded in turn by the blue throats. Field studies confirm the presence of cycles.

Benjamin Kirkup and Margaret Riley find rock-paper-scissors being played by bacteria in the gut of a living mouse.<sup>15</sup> One strain of *E.coli* both produces a poison and maintains immunity to this poison. There are two metabolic costs, one for the poison and one for the immunity, which reduce reproductive potential. These poisoners beat normal *E.coli*, which are not immune, in the

<sup>14</sup> Sinervo and Lively 1996.

<sup>15</sup> Kirkup and Riley 2004.

spatial interactions in the gut. A third strain maintains immunity to the poison, but does not produce it. It free-rides, so to speak, on the spite of the poisoners.

These free-riders flourish in a population of poisoners, because of the lower metabolic load. But in a population of such free-riders, the normals will do best. Here there is no poison, and the cost of maintaining immunity is a drag on the free riders. This rock-scissors-paper type of interaction structure explains the maintenance of all three types in the wild. As the Red Queen said to Alice, “*Now, here, you see, it takes all the running you can do, to keep in the same place.*”

The replicator dynamics for rock-scissors-paper is shown in figure 4.1.

Rock-scissors-paper has four rest points (or equilibria) of the replicator dynamics. The three possible pure populations (*all rock*, *all scissors*, *all paper*) are all dynamically unstable. The other equilibrium is the mixed state where one-third of the population plays each strategy. This is stable, since points near it stay near it, but not strongly stable. The equilibria are not so important here. No initial population state that is not already an equilibrium converges to any of the equilibria.

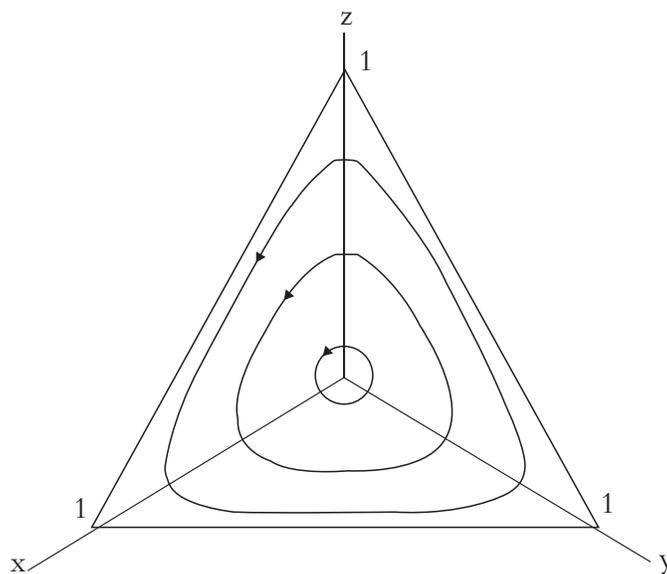


Figure 4.1: Cycles in rock-scissors-paper.

This changes if we modify the payoffs slightly:<sup>16</sup>

Rock-scissors-paper II

For small positive  $\varepsilon$ , trajectories of the replicator dynamics spiral inward to the point where the population proportions are equal.

	R	S	P
R	$1-\varepsilon$	2	0
S	0	$1-\varepsilon$	2
P	2	0	$1-\varepsilon$

This equilibrium has changed character. It is not only stable; it is *strongly stable*. Population states near it get carried to it. More impressively, it is *globally stable*. Every state in which none of the strategies is extinct converges to it.

The example illustrates another concept that will be important to us. The tiniest change in the dynamics changed the equilibrium structure radically. (A negative  $\varepsilon$  rather than a positive one would have caused the trajectories to spiral outward, changing the central equilibrium from stable to unstable.) Our original rock-scissors-paper game with replicator dynamics is said to be *structurally unstable*. In a structurally unstable situation, small local changes in the rates of change of population proportions can lead to a radically different global dynamic structure. Our game of inconsequential actions was also structurally unstable. The littlest consequence could change everything. Although structural instability in the replicator dynamics is rare in the space of games in general, in signaling games it happens *all the time!*

<sup>16</sup> Zeeman 1980; Hofbauer and Sigmund 1998.

## Natural variation

Let us now add Darwin's third principle—natural variation. In a Mendelian setting, sexual reproduction adds a lot of natural variation through genetic recombination of contributions from both parents—just as Empedocles shrewdly hypothesized. But in line with our minimalist approach so far, preserving compatibility with both biological and cultural evolution, we will just add *mutation*.

With high probability types or strategies are inherited, but with some small probability any type may mutate into any other. On the cultural side these mutations may be viewed as imperfect imitation, leavened by error, which keeps all strategies in play and prevents an absolutely monomorphic culture. In principle it might be easier for a given type to mutate into a second than into a third. However, we will concentrate on the case of uniform mutation. Every type has the same probability of mutating into any other type, so there is only one mutation rate. We again assume a large population, so chance fluctuations average out. This gives us *replicator-mutator* dynamics.<sup>17</sup>

Those population states that were dynamic equilibria only by virtue of all other types being extinct do not survive mutation. Consider one population playing Prisoner's Dilemma:

	Cooperate	Defect
Cooperate	3	1
Defect	4	2

With replicator dynamics there are two equilibria, *All Cooperate* and *All Defect*. The former is unstable, since introduction of any defectors would lead to them taking over the population. With replicator-mutator dynamics, defectors are automatically introduced by mutation and only one equilibrium survives. This is the *All Defect*

<sup>17</sup> Introduced by Haderler 1981 and analyzed by Hofbauer 1985.

equilibrium perturbed slightly by mutation. For a small mutation rate it is an *Almost-All Defect* equilibrium.

Let us return to our original rock-scissors-paper game. Instead of changing the payoffs a little, as we did earlier, we can keep the payoffs the same but introduce mutation. We change the dynamics to replicator-mutator with a small mutation rate. Since we are starting with a structurally unstable situation, we expect that this small change might have large consequences. Indeed, it is so. As before, all cycles vanish and the only surviving equilibrium is the population state where each of rock, scissors, and paper is played with probability  $1/3$ . This is a global attractor—all trajectories lead to it. Since mutants from more frequent strategies to less frequent ones are more numerous than those in the converse direction, mutation gives the dynamics a little nudge in the direction of equality. That is all it takes to destabilize the cycles and turn them into inward spirals.

Rock-scissors-paper has a lot to teach us about evolutionary games. The first big lesson is the importance of dynamical analysis. If we look for evolutionarily stable strategies—strategies that if established could repel any invaders—there aren't any. If we concentrate on equilibrium analysis, we miss the cycles. The second big lesson is the importance of attention to structural stability. If the model is structurally unstable, a small change in the model may make a big change in its dynamics.