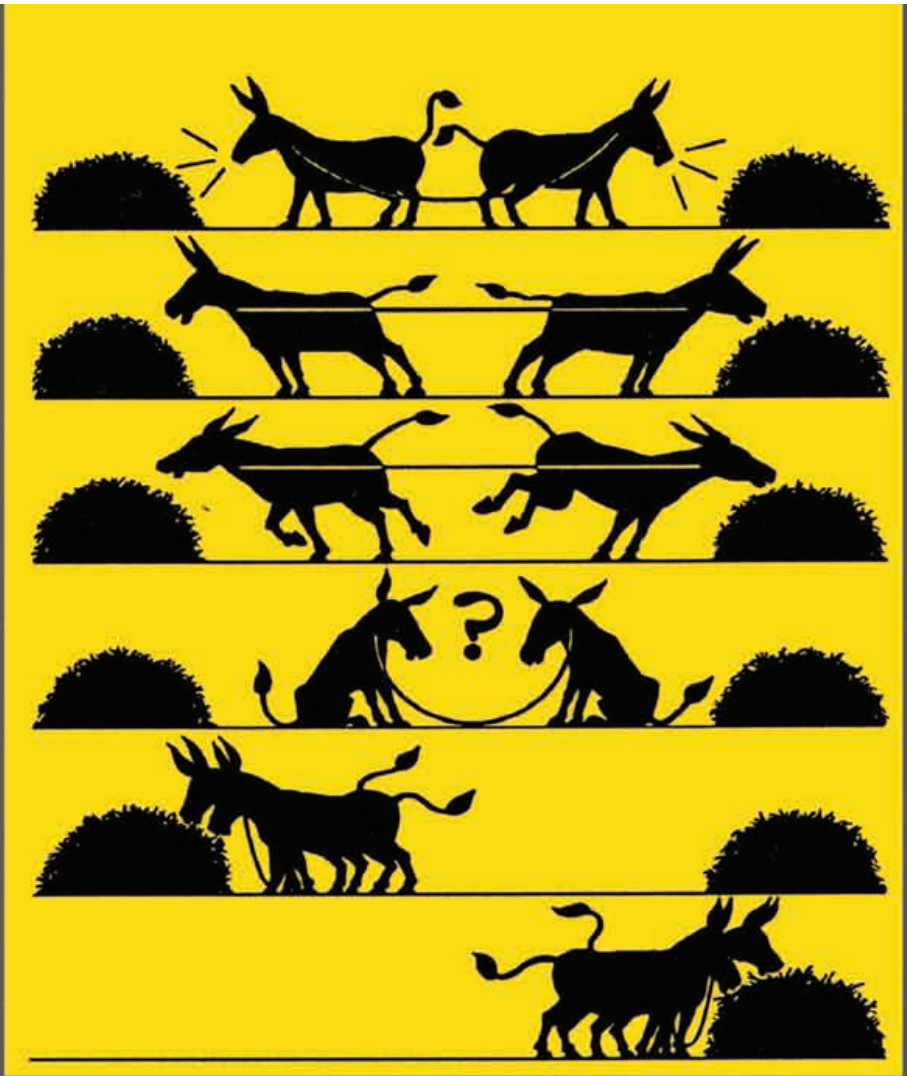

Princeton Series in Theoretical and Computational Biology

# The Calculus of Selfishness

Karl Sigmund

# Chapter Three

## Direct Reciprocity: The Role of Repetition

### 3.1 HELP

As Darwin wrote, "The small strength and speed of man, his want of natural weapons, etc., are more than counterbalanced . . . by his social qualities which lead him *to give and receive* aid from his fellow-men" (italics added). In its simplest form, to help means to confer a benefit $b$ to another individual, at a cost $c$ to oneself. This can be viewed as an atom of social interaction.

In the *Donation game*, two players have to decide simultaneously (more precisely, in ignorance of the co-player's decision) whether to give help to their co-player or not. The two strategies $\mathbf{e}_1$ and $\mathbf{e}_2$ will be denoted by C (for *cooperate*) and D (for *defect*), respectively. This yields the following payoff matrix:

$$\begin{pmatrix} b-c & -c \\ b & 0 \end{pmatrix}. \tag{3.1}$$

If not otherwise stated, we will assume $b > c > 0$. The second strategy D dominates the first. This is an example of a Prisoner's Dilemma game, as described in section 1.3, i.e., a symmetric $2 \times 2$ game whose payoff matrix

$$\begin{pmatrix} R & S \\ T & P \end{pmatrix} \tag{3.2}$$

satisfies

$$T > R > P > S. \tag{3.3}$$

The Prisoner's Dilemma game encapsulates the tug-of-war between the common interest ($R$ is larger than $P$) and the selfish interest (D dominates C). Selfishness ought to win in this conflict. Indeed, the game has a unique Nash equilibrium, namely defection; and imitation of successful individuals leads inexorably to the demise of cooperation, see section 2.10.

It can be interesting to compare this Donation game with the Snowdrift game (see section 1.4). Both players can receive a benefit $b$ each, if they come up with a fee $c < b$. They have to decide simultaneously whether to pay the fee or not, knowing that if both decide to pay, they will share the cost. The payoff matrix is

$$\begin{pmatrix} b-\frac{c}{2} & b-c \\ b & 0 \end{pmatrix}. \tag{3.4}$$

Obviously, it is best to do the opposite of what the other player does. If your co-player is willing to pay the fee, you yourself can safely skip it. But if your co-player is unwilling to pay the fee, you should better pay. Clearly, a player would prefer to be the one who does not pay the cost. The game has a unique symmetric Nash equilibrium. It consists in paying the fee with a probability of $\frac{2(b-c)}{2b-c} = 1 - \frac{c}{2b-c}$. We note that if the Donation game is played twice, then the two players would do twice as well to both play C each time than to take turns in playing C. In two turns of the Snowdrift game, they would do as well to both play C each time than to take turns in playing C. The Snowdrift game is an example of the so-called Chicken game, a symmetric $2 \times 2$ game whose payoff matrix (3.2) satisfies

$$T > R > S > P. \tag{3.5}$$

The small difference in rank order ($S$ and $P$ are permuted) has a considerable effect.

## 3.2 ITERATED GAMES

Let us now consider several rounds of the simultaneous Donation game. If the number of rounds is known to both players, then backward induction predicts, as seen in section 1.5, that selfish players ought to play D in each round.

Let us suppose instead that the two players do *not* know how many rounds their game will last. Usually, one assumes that after every round, a further round can occur with a constant probability $w < 1$. (One could alternatively assume that the number of rounds is given by a Poisson distribution, for instance.) We number the initial round by 0, and by $n$ the round obtained at the $n$-th iteration. The probability that the game is iterated at least $n$ times is given by $w^n$. The probability that the game has *exactly* $n + 1$ rounds (the initial round followed by $n$ iterations) is $w^n(1 - w)$. The number of rounds is a random variable with a geometric distribution, and its expected value is

$$1(1-w) + 2w(1-w) + \cdots + nw^{n-1}(1-w) + \cdots = \frac{1}{1-w}. \tag{3.6}$$

Let us denote by $A(n)$ the payoff in the $n$-th round. The expected value of the total payoff is given by

$$\sum_{n=0}^{+\infty} w^n(1-w)[A(0) + \cdots + A(n)], \tag{3.7}$$

which by *Abel's summation formula* is the power series $A(0) + wA(1) + \cdots$. Since $A(n) \in \{R, S, T, P\}$, all $A(n)$ are uniformly bounded, and hence expression (3.7) always converges to some value $A(w)$, for $0 \le w < 1$. The average payoff *per round* is given by

$$(1-w)A(w) = (1-w)\sum_{n=0}^{+\infty} w^n A(n). \tag{3.8}$$

It is often instructive to analyze the limiting case $w = 1$. In this case, the game consists of infinitely many rounds, and the total payoff $A(0) + A(1) + \cdots$ may diverge. It is convenient, in that case, to consider the average (over time) of the payoff *per round*, namely

$$\lim_{n \to +\infty} \frac{A(0) + \cdots + A(n)}{n+1}, \tag{3.9}$$

provided this limit exists. The *theorem of Frobenius* implies that in this case, expression (3.9) is given by the limit of equation (3.8), i.e., by $\lim_{w \to 1}(1-w)A(w)$.

## 3.3 THE GOOD, THE BAD, AND THE RECIPROCATOR

Let us first consider the interaction of three strategies only. The cooperator always decides to help; the defector always refuses to help; and the reciprocator refuses to help if and only if the co-player refused to help in the previous round. (By default, thus, the reciprocator donates in the initial round.) These are the strategies $\mathbf{e}_1 = AllC$, $\mathbf{e}_2 = AllD$ and $\mathbf{e}_3 = TFT$, respectively.

We consider a large, well-mixed population. The frequencies of the three strategies are given by $x$, $y$, and $z$, respectively (with $x + y + z = 1$). With $P_x$, $P_y$, and $P_z$ we denote the expected values for the total payoff obtained by players using these strategies (rather than by $(A\mathbf{x})_1$, etc., as in the previous chapter). The average payoff in the population is $\bar{P} = xP_x + yP_y + zP_z$. We shall assume that more successful strategies are more likely to be imitated, as in section 2.7. Hence the evolution of the frequencies of the three strategies in the population is given by the replicator equation

$$\begin{aligned} \dot{x} &= x(P_x - \bar{P}) \\ \dot{y} &= y(P_y - \bar{P}) \\ \dot{z} &= z(P_z - \bar{P}). \end{aligned} \tag{3.10}$$

We will frequently use the fact that the replicator equation remains unchanged (on the simplex $S_3$) if the same function is added to each payoff term (see section 2.8), and by abuse of notation still design the corresponding terms with $P_x$, $P_y$, $P_z$, and $\bar{P}$. In particular, we can normalize the payoff matrix by adding an appropriate constant to each column.

*AllD* against *AllD* obtains payoff $A(n) = 0$ in every round, so that $A(w) = 0$. *TFT* against *AllD* earns $A(0) = -c$ in the initial round, and henceforth $A(n) = 0$ for $n \geq 1$, so that $A(w) = -c$, etc.

The payoff matrix for the three strategies *AllC*, *AllD*, and *TFT* is given by

$$\begin{pmatrix} b-c & -c & b-c \\ b & 0 & b(1-w) \\ b-c & -c(1-w) & b-c \end{pmatrix}, \tag{3.11}$$

where we omitted the factor $(1-w)^{-1}$, (i.e., considered the average payoff *per round*). Setting $w = 1$ yields the infinitely repeated case.

In the general Prisoner's Dilemma game, the payoff matrix corresponding to matrix (3.11) is

$$\begin{pmatrix} R & S & R \\ T & P & (1-w)T + wP \\ R & (1-w)S + wP & R \end{pmatrix}. \tag{3.12}$$

## 3.4 PYRRHIC VICTORIES

Let us stick with the Donation game and normalize the corresponding replicator equation such that $P_y$, the payoff for defectors, is 0. Then we obtain

$$P_x = -c + wbz \qquad P_z = P_x + wcy. \tag{3.13}$$

We note that $P_z - \bar{P} = yg$, with

$$g = w(b-c)z - c(1-w). \tag{3.14}$$

On the edge of the state space simplex $S_3$ with $z = 0$ (no reciprocators), *AllD* clearly dominates. On the edge with $x = 0$, i.e., in a population consisting of defectors and TFT players, we have bi-stable dynamics. The unstable equilibrium is $F_{yz} = (0, 1 - \hat{z}, \hat{z})$, with

$$\hat{z} = \frac{(1-w)c}{w(b-c)}, \tag{3.15}$$

provided $\hat{z} < 1$, i.e., $w > c/b$. In particular, *TFT* is risk-dominant (see section 2.10) if

$$w > \frac{2c}{b+c}, \tag{3.16}$$

and selectively advantageous (see section 2.17) if

$$w > \frac{3c}{b+2c}. \tag{3.17}$$

Since $\hat{z}$ is small if $w$ is close to 1, a small *TFT* population is able to invade a population of defectors if $w$, i.e., the "shadow of the future" is sufficiently large.

The edge $y = 0$ consists of fixed points only. Clearly, a population of *AllC* and *TFT* players will always cooperate, and none of the two strategies is favored. On the edge $y = 0$, those points with $z \geq c/wb$ are Nash equilibria, and the others are not. To see this, we have only to look at the sign of $P_y - \bar{P}$, i.e., of $P_x = -c + wbz$, and recall from section 2.10 that the Nash equilibria are exactly those fixed points that are saturated (i.e., if $y = 0$, then $P_y \leq \bar{P}$).

The other Nash equilibria of the game are the vertex $y = 1$ (defectors only) and the point $F_{yz}$. In the interior of the simplex $S_3$, there is no fixed point, since $P_z > P_x$ whenever $y > 0$. It is easy to see that the function

$$V = x^{\frac{1-w}{w}} z^{-\frac{1}{w}} g \tag{3.18}$$

is an invariant of motion, i.e., satisfies $\dot{V} = 0$.

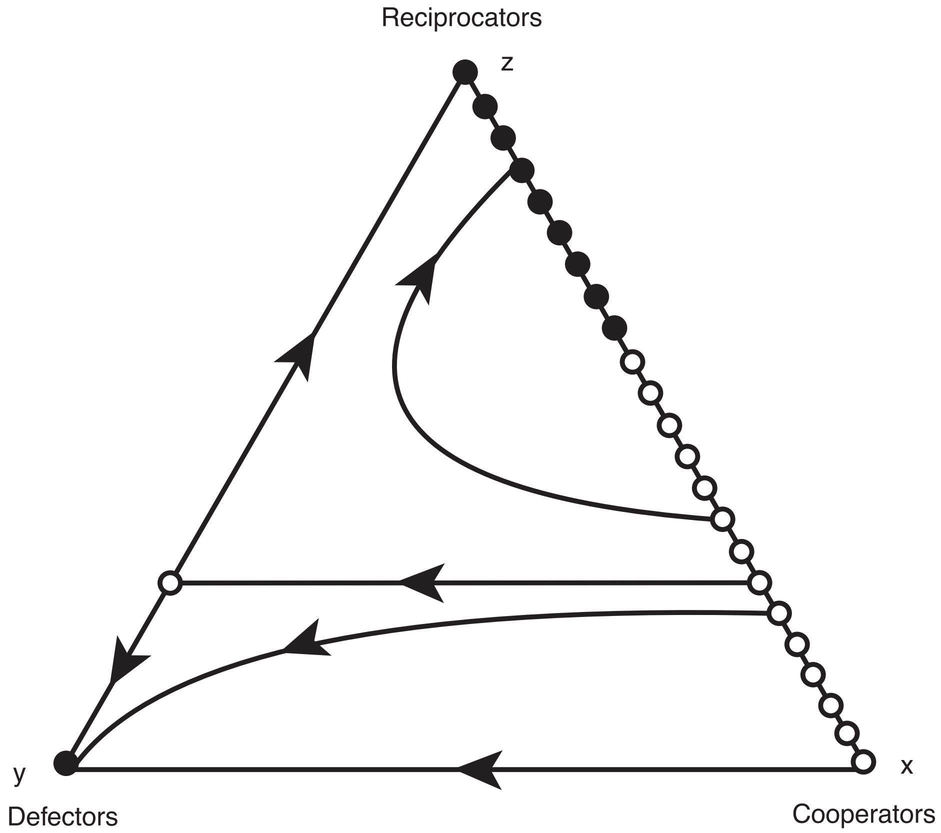


Figure 3.1 The good, the bad, and the reciprocator, in the absence of errors. A horizontal line $z = c/wb$ divides the state space. Below the line, defectors win; above the line, defectors are eliminated. Here and in all other figures, filled circles correspond to stable rest points, and empty circles to unstable rest points.

In the case $c < wb$, the dynamics shows an interesting behavior, see figure 3.1. The segment with $g = 0$ consists of a single orbit parallel to the edge $z = 0$, which converges to the saddle point $F_{yz}$ and separates the simplex into two parts. Below this line, $z$ decreases, and $y$ converges to 1, i.e., defectors win. Above the line $g = 0$, $z$ increases, and $y$ converges to 0, i.e., defectors lose.

In the absence of defectors, any mixture of *TFT* players and *AllC* players corresponds to a rest point. Such a mixture can be viewed as a mixture of discriminating and indiscriminating altruists. If we assume that occasionally, small random shocks perturb the system, then these will send the system up and down the defectors-free edge $y = 0$. If a random shock introduces a small amount of defectors while $z > c/wb$, the defectors will forthwith be eliminated. If the defectors are introduced while $z < (1 - w)c/w(b - c)$, they will take over. But if the defectors are introduced in the "middle zone" where

$$c/wb > z > (1 - w)c/w(b - c), \tag{3.19}$$

the amount of defectors will first increase, and then vanish. During the phase of their invasion, the *AllD* players will exploit and eventually deplete the *AllC* players. This is a kind of Pyrrhic victory: the defectors end up meeting mostly *TFT* players, and this is their undoing.

Looking at it from the point of view of defectors, any invasion attempt while $z > \hat{z}$ is doomed to failure and will result in a state with $y = 0$ and $z > c/wb$. Figuratively speaking, the only hope for the defectors is to wait with their invasion attempt until drift, i.e., a succession of small random shocks, has moved the population state along the edge $y = 0$, to the region where $z < \hat{z}$. This drift needs time. If the invasion attempts occur too often, the drift will never have enough time to lead into the zone that favors the defectors. Thus defectors should not try to invade too often. In other terms, cooperators are safe only if invasion attempts by defectors are sufficiently frequent. If the invasion attempts are too rare, a cooperative society can lose its immunity—random fluctuations can lead to a population state with too few reciprocators to repel an invading minority of defectors.

## 3.5 REACTIVE STRATEGIES

So far, we have assumed that the players execute their intentions faultlessly. If we assume that they occasionally commit errors, we obtain very different results. This leads to the investigation of stochastic strategies, described by the probabilities, in each round, to cooperate or not.

To begin with, let us consider strategies given by triplets $(f, p, q)$, where $f$ is the probability to cooperate in round 0, and $p$ resp. $q$ are the probabilities to cooperate after a cooperation resp. defection by the co-player in the previous round. For such *reactive strategies*, the propensity to cooperate depends uniquely on what the co-player did in the previous round. The pair $(p, q)$ defines the *reaction norm* of the strategy. It is a point in the unit square $[0, 1]^2$, and it is said to be *deterministic* if it corresponds to one of the corners. For instance, *TFT* corresponds to $(1, 1, 0)$ and *AllD* to $(0, 0, 0)$; both have deterministic reaction norms. A (small) probability $\epsilon$ to implement the unintended move would change this to $(1-\epsilon, 1-\epsilon, \epsilon)$ resp. $(\epsilon, \epsilon, \epsilon)$. We shall use the notation $\rho := p - q$. Clearly $|\ \rho\ | < 1$ except for some strategies with deterministic reaction norm, such as *TFT*.

Let us consider an $(f, p, q)$ player encountering an $(f', p', q')$ player. In each round, there are four possible outcomes, namely $(C, C)$, $(C, D)$, $(D, C)$, and $(D, D)$, depending on the moves of the first and the second player. This outcome can also be described by the payoff obtained by the first player, namely $R$, $S$, $T$, or $P$, which we enumerate by 1, 2, 3, 4. (Note that an $S$ for the first player corresponds to a $T$ for the second player.)

In the initial round, the probabilities $x_i(0)$ for outcome $i \in \{1, 2, 3, 4\}$ are given by the quadruple

$$\mathbf{x}(0) = (ff', f(1-f'), (1-f)f', (1-f)(1-f')). \tag{3.20}$$

In the following rounds, these probabilities change according to the reaction norms of the two players. We denote by $p_{ij}$ the probability that from one round to the next, the state changes from $i$ to $j$ (with $i, j \in \{1, 2, 3, 4\}$). Thus $\mathbf{x}(n)$ turns into $\mathbf{x}(n+1)$ according to the transition rule

$$\mathbf{x}(n+1) = \mathbf{x}(n)\mathbf{P}, \tag{3.21}$$

where $\mathbf{P} = (p_{ij})$ is the stochastic matrix

$$\mathbf{P} = \begin{pmatrix} pp' & p(1-p') & (1-p)p' & (1-p)(1-p') \\ qp' & q(1-p') & (1-q)p' & (1-q)(1-p') \\ pq' & p(1-q') & (1-p)q' & (1-p)(1-q') \\ qq' & q(1-q') & (1-q)q' & (1-q)(1-q') \end{pmatrix}. \tag{3.22}$$

This yields a Markov chain.

## 3.6 LINKAGE

If the probabilities $x_i$ to be in state $i$ satisfy the condition

$$x_1 x_4 = x_2 x_3, \tag{3.23}$$

then the moves of the two players are independent. Indeed, $x_1$ is the probability that both player I *and* player II play $C$. The probability that $I$ plays $C$ is $x_1 + x_2$, and the probability that II plays $C$ is $x_1 + x_3$. Independence means that $x_1 = (x_1 + x_2)(x_1 + x_3)$, which for $\mathbf{x} \in S_4$ is equivalent with $x_1 x_4 = x_2 x_3$. In this sense the *linkage* $D = x_1 x_4 - x_2 x_3$ measures the interdependence of the two players: $D = 0$ means that their moves are independent.

A straightforward computation shows that

$$D(n+1) = \rho\rho' D(n), \tag{3.24}$$

where $\rho = p - q$ and $\rho' = p' - q'$ as before. Indeed, we have only to replace $x_j(n+1)$ in $D(n+1)$ by $\sum_i x_i(n) p_{ij}$, using equation (3.21), and then compare the coefficients of the product terms $x_i(n) x_j(n)$. Most of the coefficients cancel obligingly, since $p_{k1} p_{k4} = p_{k2} p_{k3}$ and $p_{1k} p_{4k} = p_{2k} p_{3k}$ for $k = 1, 2, 3, 4$.

It follows that the *linkage disequilibrium* $D(n)$, which is 0 in the initial round, remains 0. (If it were initially distinct from 0, it would converge to 0 exponentially if at least one of the reaction norms is non-deterministic.) This confirms that the moves of the two players are independent in every round, as expected.

## 3.7 COOPERATION LEVELS

Players using reactive strategies play a kind of ping-pong with each other: if player II cooperates with a probability $y$, then player I cooperates with probability

$$\alpha(y) = py + q(1-y) = q + \rho y \tag{3.25}$$

in the following round. Thus if player I's *cooperation level* in round $n$ is denoted by $c_n = x_1(n) + x_2(n)$, then

$$c_{n+2} = q + \rho(q' + \rho' c_n) = A + u c_n, \tag{3.26}$$

where $u := \rho\rho'$ and $A = q + \rho q'$ (which is $\alpha\alpha'(0)$). Equation (3.26) defines an affine-linear mapping from the unit interval (the set of all cooperation levels) into itself. The mapping can be iterated, starting from the initial round:

$$\begin{aligned} c_0 = f \mapsto c_2 &= A + uf \mapsto c_4 = A + u(A + uf) \\ &= A(1+u) + u^2 f \mapsto \cdots . \end{aligned} \tag{3.27}$$

Since

$$c_{2n} = A(1 + u + \cdots + u^{n-1}) + u^n f = \frac{A}{1-u} + u^n \left( f - \frac{A}{1-u} \right), \tag{3.28}$$

we obtain

$$c_{2n} = v + u^n (f - v), \tag{3.29}$$

where

$$v := \frac{A}{1-u} = \frac{q + \rho q'}{1 - \rho\rho'} \tag{3.30}$$

is just the fixed point of $y \mapsto A + uy$. A similar equation holds for $c_{2n+1}$ (with $f$ replaced by $c_1 = q + \rho f'$). The cooperation level $c_n$ thus converges to $v$. The same holds for the other player, whose cooperation level converges to $v'$. Clearly, one has $\alpha(v') = v$, etc. It is only if both strategies have deterministic reaction norms that the cooperation levels may periodically oscillate forever, for instance if a *TFT* player encounters a "suspicious *TFT* player" using $(0, 1, 0)$.

In addition to the stationary cooperation levels $v$ and $v'$ of the two players against each other, we can also consider the hypothetical cooperation levels $s$ and $s'$ which the players would obtain, in the limit, against a co-player using their own strategy. An $(f, p, q)$ player reaches a cooperation level

$$s := \frac{q}{1-\rho} \tag{3.31}$$

against another $(f, p, q)$ player. Interestingly, $v - v'$ has the same sign as $s - s'$ (and as $v - s'$, as well as $s - v'$). In particular, if two of the limits $v$, $v'$, $s$, and $s'$ of cooperation levels are equal, so are all four. It is useful to note that

$$v - s = \rho(v' - s). \tag{3.32}$$

This leads to a simple interpretation linking cooperation levels to reaction norms. All reaction norms $(p', q')$ lying on the line from $(p, q)$ to $(1, 0)$ (the *TFT* norm) have the same asymptotic cooperation level against themselves, and consequently against each other. If a reaction norm $(p', q')$ lies above the line from $(p, q)$ to $(1, 0)$, it has a higher asymptotic cooperation level (against itself, and against $(p, q)$), and vice versa.

## 3.8 PAYOFF VALUES

We shall not consider the case $u^2 = 1$, (which occurs only if both strategies have a deterministic reaction norm).

Since the decisions of the two players are independent, the player using $(f, p, q)$ obtains in round $n$ against a player using $(f', p', q')$ the payoff

$$A(n) = Rc_nc'_n + Sc_n(1 - c'_n) + T(1 - c_n)c'_n + P(1 - c_n)(1 - c'_n). \tag{3.33}$$

In the special case of the Donation game, this reduces to

$$A(n) = bc'_n - cc_n. \tag{3.34}$$

For the infinitely iterated case $w = 1$ this means that the average payoff per round is

$$(R - S - T + P)vv' + (S - P)v + (T - P)v' + P, \tag{3.35}$$

which reduces to

$$bv' - cv \tag{3.36}$$

for the Donation game. These expressions do not depend on the initial propensities to cooperate, namely $f$ and $f'$.

In order to obtain the total payoff for the Donation game with $w < 1$, we have to compute $\sum w^n c_n$. By equation (3.29),

$$\sum w^n c_n = \sum w^{2n}[v + u^n(f - v)] + \sum w^{2n+1}[v + u^n(c_1 - v)], \tag{3.37}$$

which is, up to the factor $[(1 - w)(1 - uw^2)]^{-1}$,

$$\begin{aligned} & v(1 - uw^2) - v(1 - w) - vw(1 - w) + f(1 - w) + c_1w(1 - w) && (3.38)\\ &= vw^2(1 - u) + (1 - w)(f + qw + w\rho f') \\ &= (q + \rho q')w^2 + (1 - w)(f + qw + w\rho f'). && (3.39) \end{aligned}$$

Collecting the terms in $f$, $f'$, $q$, and $q'$, and setting $e := (1 - w)f + wq$, $e' := (1 - w)f' + wq'$, we obtain

$$\sum w^n c_n = \frac{e + w\rho e'}{(1 - w)(1 - uw^2)}. \tag{3.40}$$

Thus the average payoff per round is given by

$$\frac{-c(e + w\rho e') + b(e' + w\rho' e)}{1 - uw^2}. \tag{3.41}$$

## 3.9 THE GOOD, THE BAD, AND THE RECIPROCATOR WITH ERRORS

We will assume that an intended donation can fail with probability $\epsilon$, and an intended refusal with probability $k\epsilon$, for some $k \geq 0$. It makes sense to distinguish between these two errors in implementation, and in particular to keep the case $k = 0$ in mind. For instance, players who want to donate, but are out of funds, are failing to implement their intention. But it is unlikely that players who do not want to give anything away are absentminded enough to donate. Thus the three strategies *AllC*, *AllD* and *TFT* are now given by $\mathbf{e}_1 = (1-\epsilon, 1-\epsilon, 1-\epsilon)$, $\mathbf{e}_2 = (k\epsilon, k\epsilon, k\epsilon)$ and $\mathbf{e}_3 = (1-\epsilon, 1-\epsilon, k\epsilon)$, respectively.

Applying expression (3.41) to these three strategies, we obtain a $3 \times 3$ payoff matrix $M$ which, at first glance, looks somewhat daunting. But it can be simplified considerably, especially as the $\rho$ values of the two unconditional strategies are 0 (i.e., $p = q$). Once more we use the fact that the replicator dynamics in $S_3$ is unchanged if we subtract, in each column of $M$, the diagonal from all elements. Up to the multiplicative factor $c(1-(k+1)\epsilon)$, the normalized matrix of payoff values per round is of the form

$$M = \begin{pmatrix} 0 & -1 & \delta\sigma \\ 1 & 0 & -\kappa\sigma \\ \delta & -\kappa & 0 \end{pmatrix} \tag{3.42}$$

where we used

$$\delta := w\epsilon, \quad \kappa := 1 - w + wk\epsilon, \quad \sigma := \frac{b\theta - c}{c - c\theta}, \text{ and } \theta = w(1-(k+1)\epsilon). \tag{3.43}$$

We note that $\bar{P} = z(1+\sigma)P_z$. Using

$$P_z - \bar{P} = P_z[1-(1+\sigma)z], \tag{3.44}$$

we see that in the interior of $S_3$, $\dot{z} = 0$ holds whenever $g := 1-(1+\sigma)z$ vanishes. It is easy to see that $g = 0$ corresponds to an orbit connecting the fixed points $F_{yz} := (0, 1-\hat{z}, \hat{z})$ and $F_{xz} := (1-\hat{z}, 0, \hat{z})$, where $\hat{z} := (1+\sigma)^{-1}$. On the edge $x = 0$ defectors and reciprocators are engaged in a bi-stable competition, their basins of attraction separated by $F_{yz}$. On the edge $y = 0$, reciprocators and *AllC* players are stably coexisting at the point $F_{xz}$. On the edge $z = 0$ of unconditional players, the defectors dominate the cooperators.

In the interior of $S_3$ we obtain an invariant of motion

$$V := x^A y^B z^C [1-(1+\sigma)z] \tag{3.45}$$

with $A = \kappa/\theta$, $B = \delta/\theta$, and $C = -1/\theta$ (note that $A + B + C + 1 = 0$). The interior rest point is

$$F = \frac{1}{1+\sigma(\kappa+\delta)}(\kappa\sigma, \delta\sigma, 1). \tag{3.46}$$

The dynamics is shown in figure 3.2. There is a horizontal orbit on the line $z = \hat{z}$, connecting the fixed points $F_{xz}$ and $F_{yz}$ (the latter is a Nash equilibrium). Below this line, all orbits converge to $y = 1$, the defectors win. The part above the line is filled with periodic orbits surrounding the unique fixed point: they correspond to

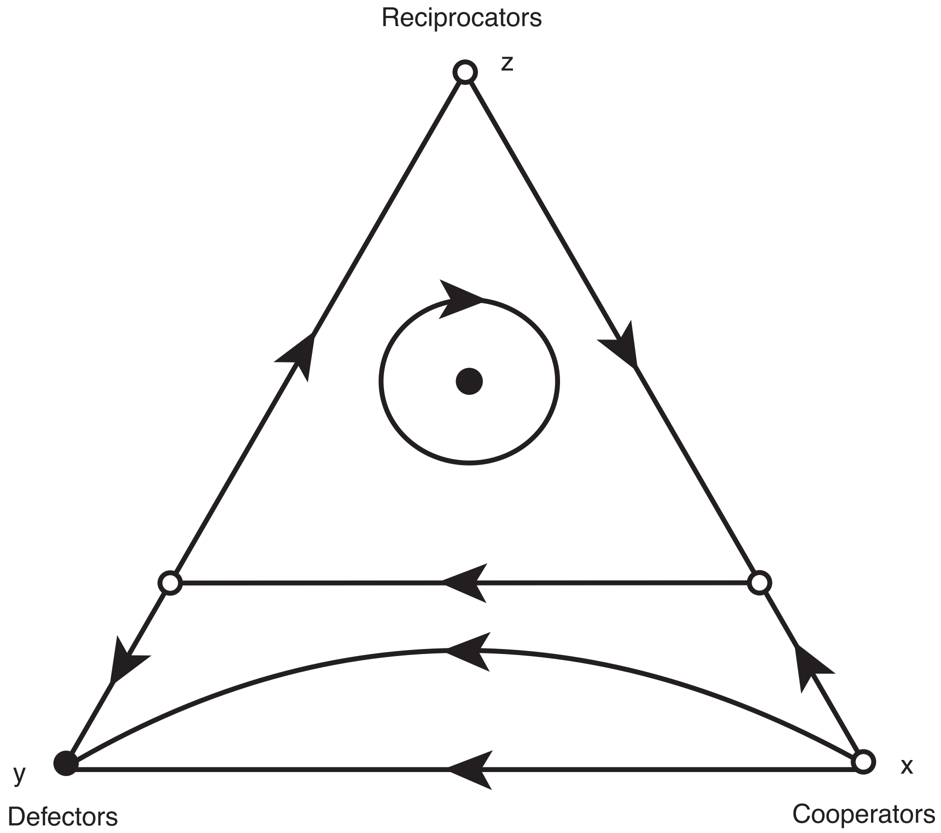


Figure 3.2 The good, the bad, and the reciprocator, with errors. If $z$ is below a threshold, defectors win; if $z$ is above the threshold, all three strategies co-exist, their frequencies oscillating periodically.

the constant level curves of the invariant of motion $V$ given by expression (3.45). The time averages correspond to the values at the rest point $F$. This rest point is stable, but not asymptotically stable. We note that the amount of defectors at $F$ can be made arbitrarily small if the error rate $\epsilon$ is sufficiently reduced. On the other hand, the basin of attraction of the *AllD* state ($y = 1$) can be arbitrarily small if $w$ is sufficiently close to 1.

## 3.10 LIMITING CASES

For $w = 1$ we obtain, up to the multiplicative factor $c(1 - (k + 1)\epsilon)$, the payoff matrix

$$M = \begin{pmatrix} 0 & -1 & \beta \\ 1 & 0 & -k\beta \\ \epsilon & -k\epsilon & 0 \end{pmatrix} \tag{3.47}$$

where

$$\beta := \frac{1}{c}\left(\frac{b-c}{1+k} - \epsilon b\right). \tag{3.48}$$

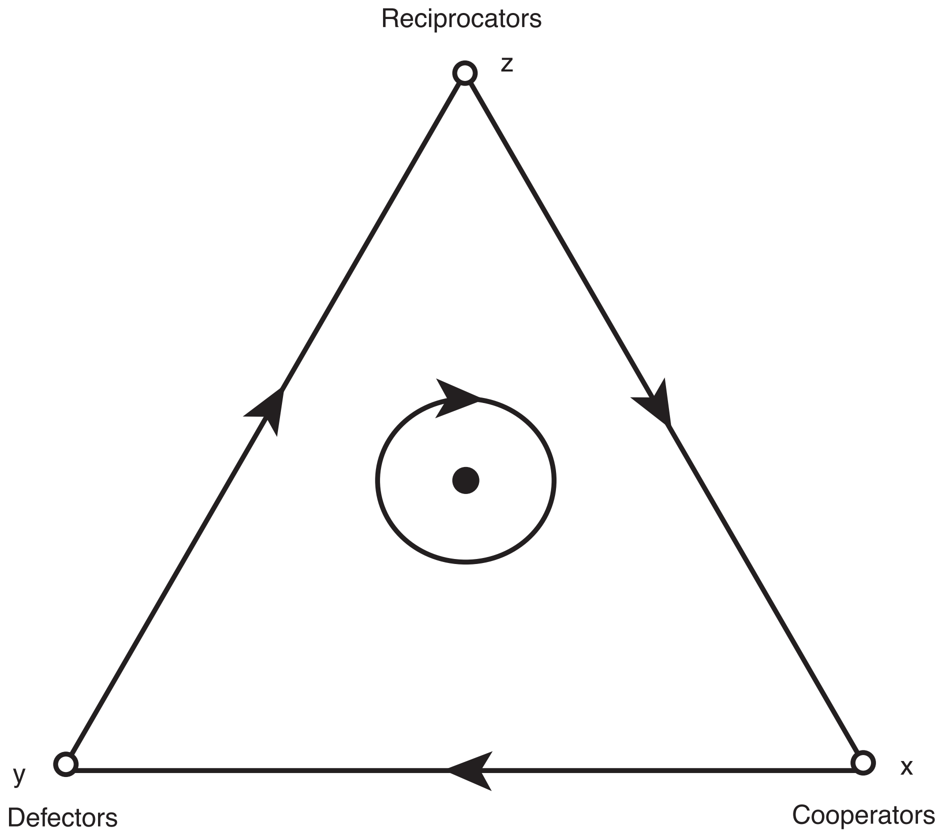


Figure 3.3 The infinitely iterated Donation game ($w = 1$), if there is a positive probability that intended moves (donation or refusal) are mis-implemented.

If $k > 0$ (i.e., if there is a positive probability for a donation, even if a refusal is intended), the dynamics is the same as in figure 3.2, the $z$ coordinate of the separatrix is given by

$$\hat{z} := \frac{c}{(b-c)} \frac{(k+1)\epsilon}{(1-(k+1)\epsilon)}. \tag{3.49}$$

If $\epsilon \to 0$ the separatrix merges with the edge $z = 0$ and we obtain a system whose payoff matrix is

$$M = \begin{pmatrix} 0 & -c & (b-c)/(1+k) \\ c & 0 & -k(b-c)/(1+k) \\ 0 & 0 & 0 \end{pmatrix}. \tag{3.50}$$

This is a Rock-Scissors-Paper game: *AllD* is out-competed by *TFT*, which is out-competed by *AllC*, which is out-competed by *AllD* in turn. The unique rest point in the interior of $S_3$ is $F = \frac{1}{b}(k(b-c)/(k+1), (b-c)/(k+1), c)$. The replicator dynamics is as in figure 3.3.

If, on the other hand, we first consider the limiting case $\epsilon = 0$ (with $w < 1$), we obtain the dynamics shown in figure 3.1 If we then let $w$ converge to 1, we obtain fig. 3.4. We note that the limits $w \to 1$ and $\epsilon \to 0$, therefore, do not commute.

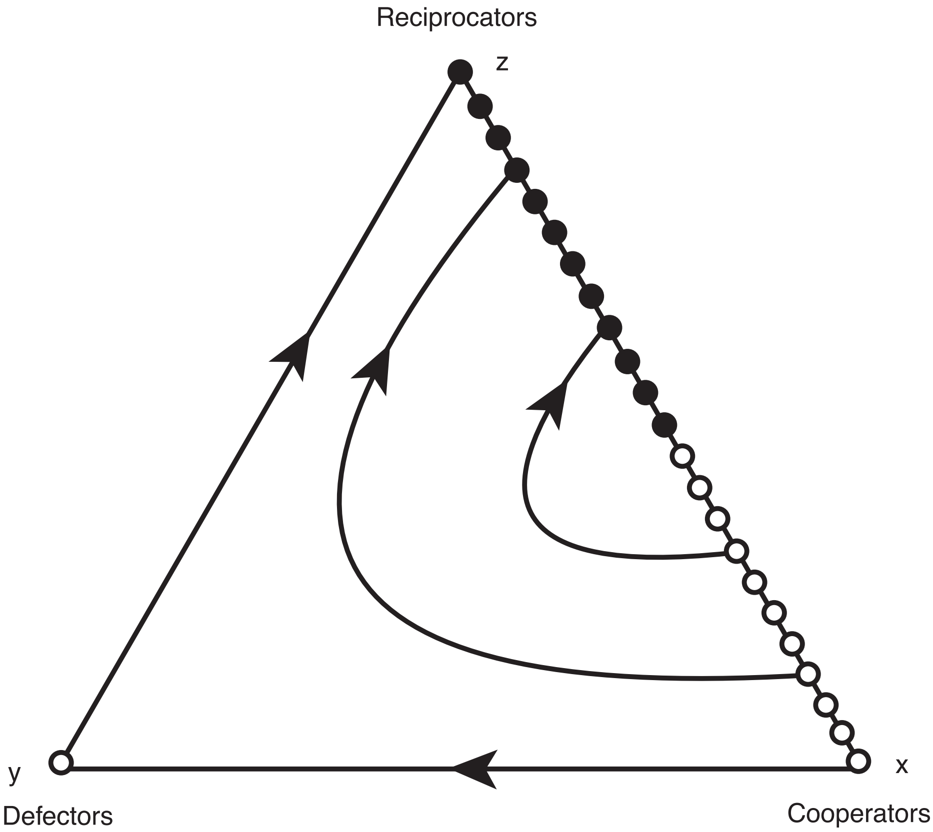


Figure 3.4 The infinitely iterated Donation game ($w = 1$), in the absence of errors (i.e., $\epsilon = 0$).

Suppose now that $k = 0$, i.e., that an intended refusal never fails. This is not without plausibility. In the limiting case $w = 1$, the payoff matrix is given, up to the factor $c(1 - \epsilon)$, by

$$M = \begin{pmatrix} 0 & -1 & \beta \\ 1 & 0 & 0 \\ \epsilon & 0 & 0 \end{pmatrix} \tag{3.51}$$

with $\beta = [(1-\epsilon)b - c]/c$. This yields a completely different picture. The edge $x = 0$ consists of fixed points. Intuitively, this is clear: errors between two *TFT* players will eventually lead to mutual defection in each round, and this can never be redressed by another error. Thus the *TFT* players' average payoff per round will be 0. The rest points with $z \le \bar{z}$ are Nash equilibria, with

$$\bar{z} = c/[b(1 - \epsilon)]. \tag{3.52}$$

The dynamics looks as in figure 3.5, which is an intriguing mirror-image of figure 3.1.

## 3.11 ADAPTIVE DYNAMICS

The reactive strategies $(f, p, q)$ form a continuum. A heterogeneous population consisting of three or four such strategies can have a complex dynamics displaying

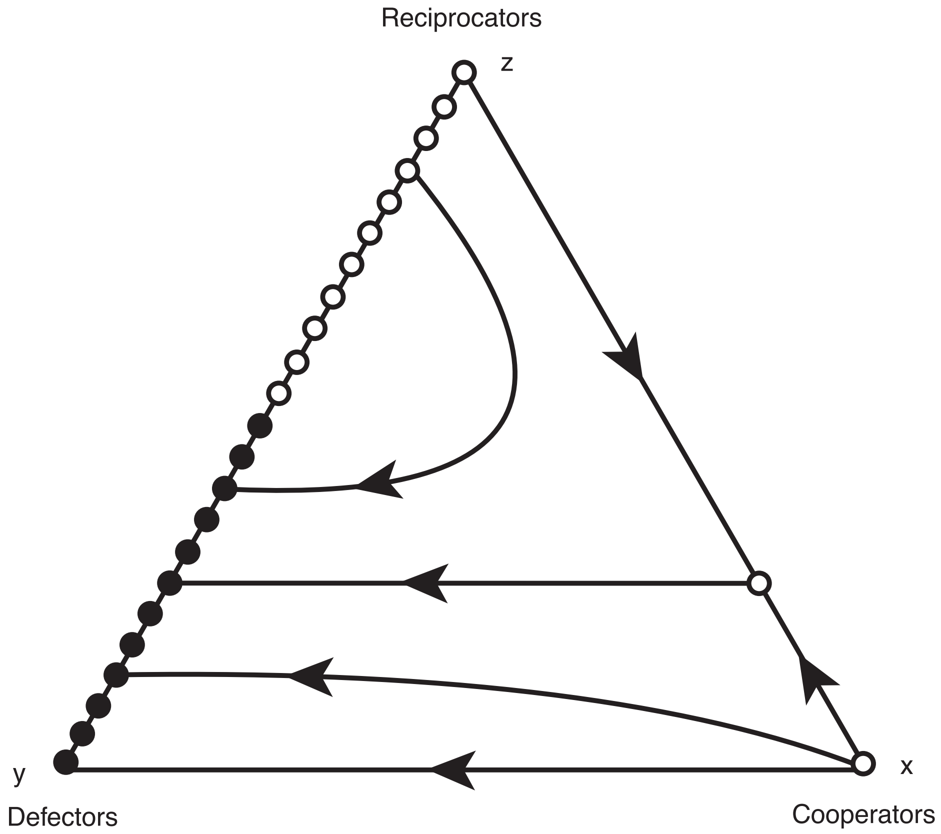


Figure 3.5 The replicator dynamics of the infinitely iterated Donation game, if only donations can be mis-implemented, but refusals are not. Cooperation vanishes in the long run.

limit cycles, heteroclinic cycles, or chaotic oscillations. Rather than pursue this point, let us ask how a *homogeneous* population evolves.

More precisely, we shall assume that the resident population is homogeneous, and that from time to time, a small minority of another type enters. These dissidents can do better or less well than the residents. Imitation will occur, and usually lead either to the elimination or to the fixation of this new type. After this, another minority can try its luck, etc. Such a limiting situation (with very rare innovations and strong imitation, or in a biological framework with very rare mutations and strong selection) can be described by a sequence of homogeneous populations. We shall describe an *adaptive dynamics* pointing towards the most favorable direction of evolution.

Let us first consider the limiting case $w = 1$. If we denote with $\mathbf{n} := (p, q)$ the reaction norm of the resident type and with $\mathbf{n}' := (p', q')$ that of the rare invading minority, we have to check whether invaders or the resident population are doing better. Individuals of both types are essentially interacting with the resident (since the dissidents are rare). Let $A(\mathbf{n}', \mathbf{n})$ be the average payoff of a player using the strategy $\mathbf{n}'$ against a player using $\mathbf{n}$. Hence the type $\mathbf{n}'$ can invade if and only if the payoff difference $A(\mathbf{n}', \mathbf{n}) - A(\mathbf{n}, \mathbf{n})$ is positive.

Let us denote, as in section 3.7, the asymptotic cooperation level of a $(p, q)$ player against another $(p, q)$ player by $s$, and the asymptotic levels of cooperation

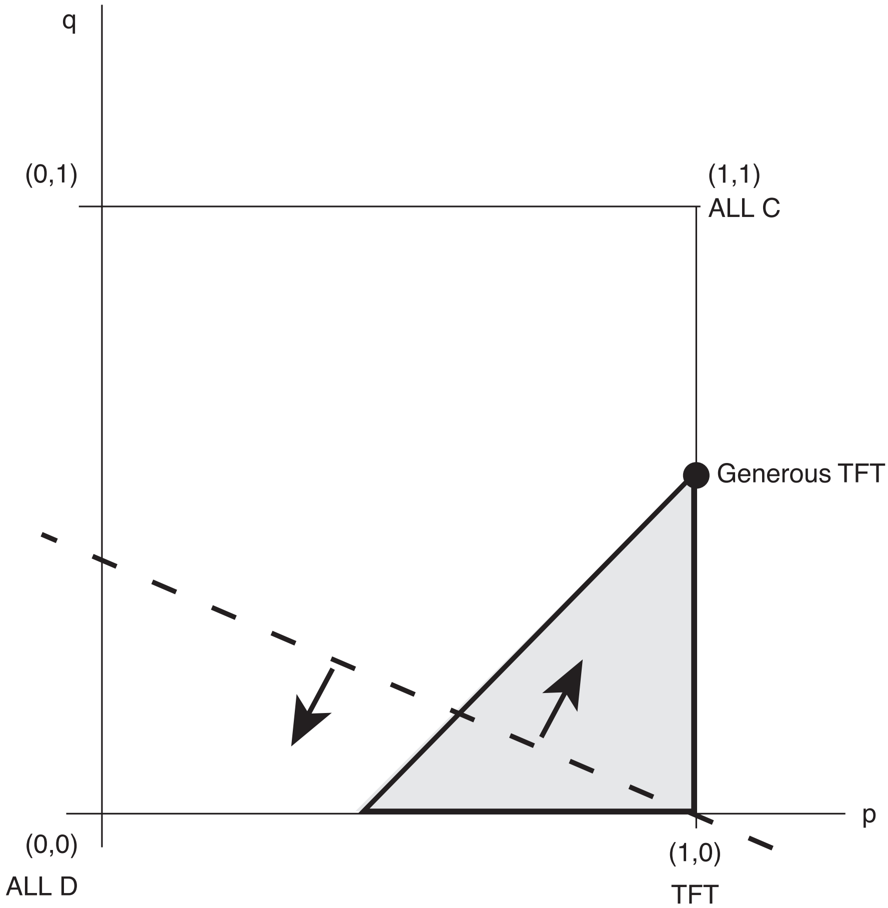


Figure 3.6 The cooperation-rewarding zone (shaded in grey) is a subset of the space of reaction norms $(p, q)$ for the infinitely iterated Donation game. The arrows point in the direction of the most favorable adaptation. This direction is always orthogonal to the line connecting the norm with $(1, 0)$.

between a $(p, q)$ player and a $(p', q')$ player by $v$ resp. $v'$. From equation (3.34) we see that

$$A(\mathbf{n}, \mathbf{n}) = bs - cs \tag{3.53}$$

and

$$A(\mathbf{n}', \mathbf{n}) = bv - cv'. \tag{3.54}$$

Hence, using $\rho = p - q$, we obtain

$$A(\mathbf{n}', \mathbf{n}) - A(\mathbf{n}, \mathbf{n}) = (v' - s)(b\rho - c). \tag{3.55}$$

The line $\rho = c/b$, i.e., $q = p - (c/b)$, divides the square $[0, 1]^2$ of reaction norms $(p, q)$ into two regions (see figure 3.6), namely the southeast corner, (which includes the *TFT* strategy $(1, 0)$) and the rest. As mentioned in section 3.7, the sign of $v' - s$ is positive resp. negative depending on whether $\mathbf{n}' = (p', q')$ lies above or below the line from $\mathbf{n} = (p, q)$ to $(1, 0)$. It follows that if $\mathbf{n}$ lies in the southeast corner, then precisely those strategies $\mathbf{n}'$ that are more cooperative can invade: indeed, if $s' > s$, then $v' > s$ and the invader's payoff is larger than that of the resident. We denote this region as the *cooperation-rewarding* zone. Conversely, if the homogeneous population adopts a strategy $\mathbf{n} = (p, q)$ that does not lie in this cooperation-rewarding zone, then every less cooperative strategy can invade. If $\mathbf{n}$ lies on the boundary of the cooperation-rewarding zone, i.e., satisfies $\rho = c/b$, then all strategies do exactly as well, against $\mathbf{n}$, as $\mathbf{n}$ does against itself.

If the invader's strategy $\mathbf{n}'$ is close to the resident's strategy $\mathbf{n}$, we can approximate the invader's payoff difference $A(\mathbf{n}', \mathbf{n}) - A(\mathbf{n}, \mathbf{n})$ by its first-order Taylor expansion, i.e., by

$$(p' - p)\frac{\partial A}{\partial p'}(\mathbf{n}', \mathbf{n}) + (q' - q)\frac{\partial A}{\partial q'}(\mathbf{n}', \mathbf{n}), \tag{3.56}$$

where the partial derivatives of the function $\mathbf{n}' \mapsto A(\mathbf{n}', \mathbf{n})$ are evaluated at $\mathbf{n}' = \mathbf{n}$. We accordingly define the adaptive dynamics in the space $[0, 1]^2$ of reaction norms $(p, q)$ as

$$\dot{p} = \frac{\partial A}{\partial p'}(\mathbf{n}', \mathbf{n}) \qquad \dot{q} = \frac{\partial A}{\partial q'}(\mathbf{n}', \mathbf{n}), \tag{3.57}$$

where the derivatives are evaluated at $\mathbf{n}' = \mathbf{n}$. This yields a vector field pointing, for every homogeneous state $\mathbf{n}$, into the direction that is most advantageous for the invader. A straightforward computation yields the derivatives of $A(\mathbf{n}', \mathbf{n}) = bv - cv'$. One obtains

$$\dot{p} = q\frac{b\rho - c}{(1-\rho)(1-\rho^2)}, \tag{3.58}$$

$$\dot{q} = (1-p)\frac{b\rho - c}{(1-\rho)(1-\rho^2)}. \tag{3.59}$$

Thus the vector $(\dot{p}, \dot{q})$ at the point $\mathbf{n} = (p, q)$ is orthogonal to the line from $\mathbf{n}$ to the *TFT* corner $(1, 0)$. In the cooperation-rewarding zone, and only there, this vector points upwards: if it pays to increase $p$ (the gratitude), it pays to increase $q$ (forgiveness), and vice versa.

The same holds for the general Prisoner's Dilemma case (if $w = 1$), except that the cooperation-rewarding zone is of a different shape: in equations (3.58) and (3.59), the term $b\rho - c$ is replaced by

$$(R - S - T + P)q\left(\frac{1+\rho}{1-\rho}\right) + (T - P)\rho + S - P. \tag{3.60}$$

There is no evolutionary tendency towards *TFT*: this strategy is a pivot, rather than a target, of adaptation.

## 3.12 GENEROUS TIT FOR TAT

Any strategy $\mathbf{n}$ at the boundary of the cooperation-rewarding zone, where $q = p - (c/b)$, has the property that every strategy $\mathbf{n}'$ yields the same payoff against $\mathbf{n}$, namely

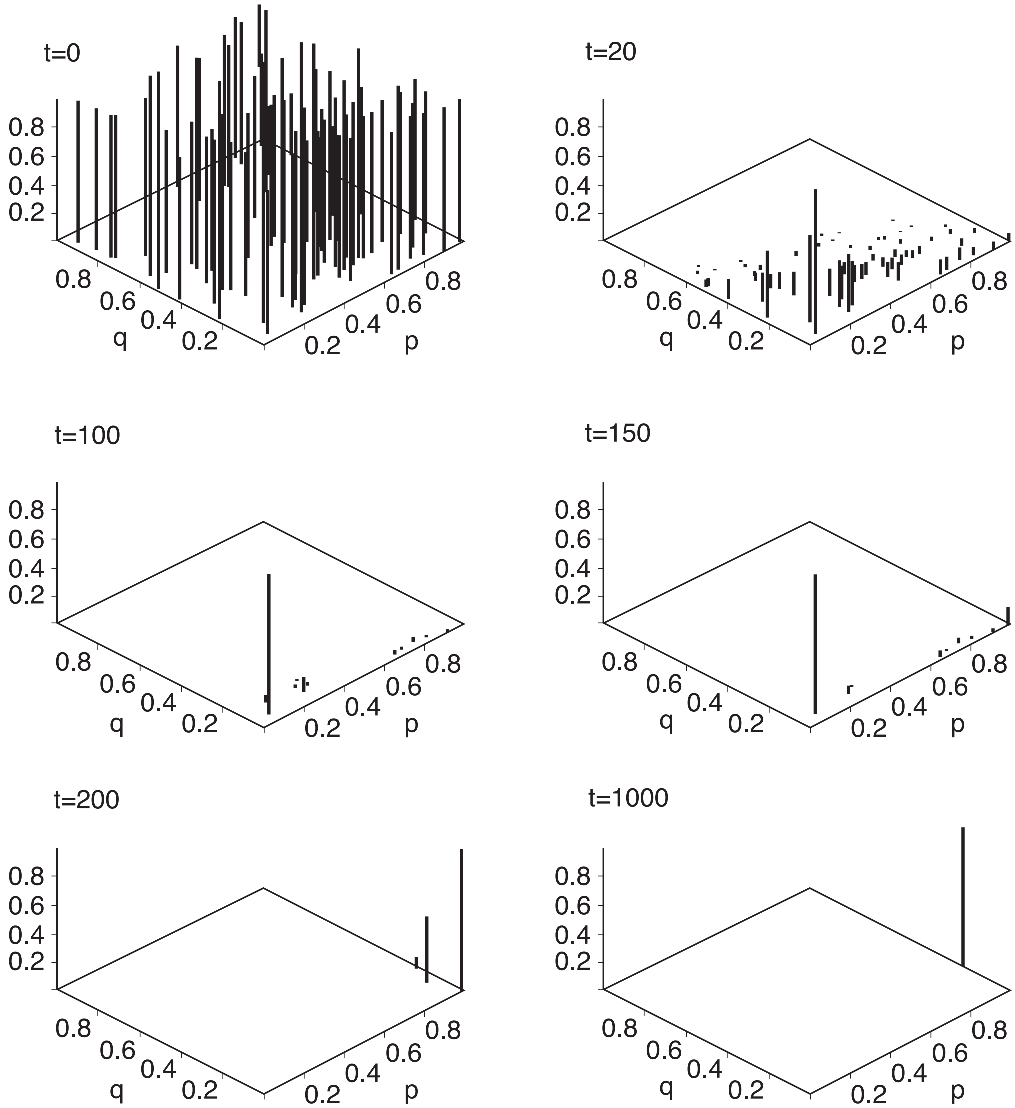


Figure 3.7 The evolution of a finite population with randomly chosen $(p, q)$ strategies. First, "*AllD*" seems to win, then "*TFT*." But in the end, "*Generous TFT*" carries the day. (After Nowak and Sigmund (1992).)

$bq$. The largest such value, namely $b - c$, is obtained for $(1, 1 - (c/b))$. This strategy is called *Generous TFT* (*GTFT*). A *Generous TFT* player always cooperates after a co-player's C, but does not always defect after a co-player's D. Rather, such a player forgives with a well-specified probability, namely $(b - c)/b$.

*Generous TFT* shows up in individual based computer simulations, see figure 3.7. Let us consider a large fictitious population of players who are assigned strategies chosen at random in the $(p, q)$ square. Thus the initial population is not at all homogeneous. It can consist of hundreds of different types. Let us assume that players meet randomly and play a repeated Donation game against each other, with a large number of rounds. Let us furthermore assume that they update their strategy from time to time, by imitating more successful players. Quickly, most of the strategies will be eliminated from the population. In general, only three out of the initial set of strategies will play a role: those closest to *AllD*, *TFT*, and *GTFT*. We shall denote these approximations by "*AllD*", "*TFT*", and "*GTFT*", respectively. What one observes at first is a strong tendency towards "*AllD*." The other strategies seem hopelessly outclassed. But then, it frequently happens (for instance if "*TFT*" is below the line from "*AllD*" to *TFT*) that "*TFT*" experiences an upsurge, and displaces "*AllD*." But this is not the end of the story. The population has reached the cooperation-rewarding zone, and strategies that have higher $p$ and $q$ values can return. In particular, the more tolerant "*GTFT*" supersedes the stern "*TFT*," and becomes fixed in the population. The striking point is that "*GTFT*" on its own can never beat "*AllD*." It needs the catalytic action of "*TFT*." It seems almost like the succession of three social phases: first the "dog-eat-dog" world of *AllD*, then the "law of the talion" represented by *TFT* and finally the age of the tolerant, but not too tolerant *GTFT*.

Similar results hold for the adaptive dynamics of the Donation game if $w < 1$. In this case, the probability $f$ to cooperate in the initial round is an additional trait. The adaptive dynamics at $\mathbf{n} = (f, p, q)$ is given by

$$\dot{f} = \frac{bw\rho - c}{1 - w^2\rho^2}. \tag{3.61}$$

$$\dot{p} = \dot{f}\left(\frac{w}{1-w}\right)\left(\frac{e}{1 - w\rho}\right) \tag{3.62}$$

$$\dot{q} = \dot{f}\left(\frac{w}{1-w}\right)\left(1 - \frac{e}{1 - w\rho}\right). \tag{3.63}$$

Here, $e = (1 - w)f + wq$ as in section 3.8. Once again, all the components have the same sign (because $0 < e < 1 - w\rho$), so that we may again speak of a cooperation-rewarding zone. A direct computation shows that we can display the adaptive dynamics in a suggestive way:

$$\dot{f} = \frac{1-w}{1 - w^2\rho^2}[A(AllC, \mathbf{n}) - A(AllD, \mathbf{n})] \tag{3.64}$$

$$\dot{p} = \frac{w}{1 - w^2\rho^2}[A(\mathbf{n}, \mathbf{n}) - A(AllD, \mathbf{n})] \tag{3.65}$$

$$\dot{q} = \frac{w}{1 - w^2\rho^2}[A(AllC, \mathbf{n}) - A(\mathbf{n}, \mathbf{n})]. \tag{3.66}$$

## 3.13 MEMORY-ONE STRATEGIES

So far, we have considered *reactive strategies* that depend only on the co-player's previous move. But it seems reasonable to assume that players also take their own move into account. It is probably easier to forgive a co-player's defection if it was matched by one's own defection, rather than if it exploited one's own cooperativeness. Hence we shall consider stochastic strategies $(f, q_R, q_S, q_T, q_P) \in [0, 1]^5$

where $f$, as before, is the propensity to play C in the initial round, and $q_R, q_S, \dots$ are the propensities to play C after having experienced an $R, S, \dots$ in the previous round.

Let us assume that player $I$ using $(f, q_R, q_S, q_T, q_P)$ encounters a co-player $II$ using $(f', q'_R, q'_S, q'_T, q'_P)$. Again, we are dealing with a Markov chain; in every round, the state is specified by the payoff obtained by player $I$. The transition probabilities are given by the matrix

$$Q = \begin{pmatrix} q_R q'_R & q_R(1-q'_R) & (1-q_R)q'_R & (1-q_R)(1-q'_R) \\ q_S q'_T & q_S(1-q'_T) & (1-q_S)q'_T & (1-q_S)(1-q'_T) \\ q_T q'_S & q_T(1-q'_S) & (1-q_T)q'_S & (1-q_T)(1-q'_S) \\ q_P q'_P & q_P(1-q'_P) & (1-q_P)q'_P & (1-q_P)(1-q'_P) \end{pmatrix}, \tag{3.67}$$

(again one player's $S$ is the other player's $T$).

The initial probabilities for the four states are given by the vector

$$\mathbf{x}(0) = (ff', f(1-f'), (1-f)f', (1-f)(1-f')), \tag{3.68}$$

which we denote by $\mathbf{f}$. In the next round, the probabilities are given by $\mathbf{f}Q$, and in round $n$ by $\mathbf{f}Q^n$. For $n \geq 1$, the probabilities need no longer be in linkage equilibrium (the matrix $Q$ satisfies $q_{k1}q_{k4} = q_{k2}q_{k3}$ but not $q_{1k}q_{4k} = q_{2k}q_{3k}$). If we denote by $\mathbf{g}$ the vector $(R, S, T, P)$, then the payoff for player $I$ in round $n$ is given by

$$A(n) = \mathbf{g} \cdot \mathbf{f}Q^n. \tag{3.69}$$

For $w < 1$ the average payoff per round, as shown in equation (3.8), is $(1-w) \sum w^n A(n)$, i.e.,

$$(1-w)\mathbf{g} \cdot \mathbf{f}(Id - wQ)^{-1}, \tag{3.70}$$

where $Id$ is the $4 \times 4$ identity matrix. For $w = 1$ we must proceed differently. If the matrix $Q$ is mixing, i.e., if there exists an $m$ such that all entries of $Q^m$ are strictly positive, then there exists a unique vector $\pi \in S_4$ that is a left eigenvector of $Q$ for the eigenvalue 1, i.e., $\pi = \pi Q$. The components $\pi_R, \pi_S, \pi_T$, and $\pi_P$ denote the stationary probabilities of the four states, and we have

$$\mathbf{f}Q^n \to \pi \tag{3.71}$$

for every initial state $\mathbf{f}$. The average payoff per round, in this case, is $\mathbf{g} \cdot \pi$, which for the Donation game reduces to

$$b(\pi_R + \pi_T) - c(\pi_R + \pi_S). \tag{3.72}$$

## 3.14 THE SPACE OF REACTION NORMS

For $w = 1$ we can neglect the initial probability to cooperate and concentrate on the space of reaction norms $(q_R, q_S, q_T, q_P)$. This unit cube is spanned by its sixteen corners, i.e., by the quadruples $(u_R, \dots, u_P)$ where $u_i$ is 1 or 0 depending on whether the strategy prescribes to use C or D after outcome $i \in \{R, S, T, P\}$. We can number these strategies as $S_j$, where $j$ ranges from 0 to 15 and is given, in binary notation, by $u_R u_S u_T u_P$. Hence $AllD = (0, 0, 0, 0)$ is $S_0$, $AllC = (1, 1, 1, 1)$ is $S_{15}$, $TFT = (1, 0, 1, 0)$ is $S_{10}$, etc. If we compute the transition matrix $\mathbf{P}$ for an $S_i$ player meeting an $S_j$ co-player, we find that in general it is not irreducible, let alone mixing: there are too many zeros, only one entry in each row does not vanish. Hence the stationary distributions are not uniquely determined.

This is different if we assume that every strategy is subject to errors in implementation: with a probability $\epsilon$, the move is the opposite of what the strategy prescribes. Then each 1 turns into $1 - \epsilon$ and each 0 into $\epsilon$. Strategy $S_j$ turns into $S_j(\epsilon)$. For instance, *TFT*, i.e., $S_{10} = (1, 0, 1, 0)$, turns into $S_{10}(\epsilon) = (1 - \epsilon, \epsilon, 1 - \epsilon, \epsilon)$, etc.

It is straightforward to compute the payoff for strategy $S_i(\epsilon)$ against $S_j(\epsilon)$. The corresponding transition matrix is $Q(\epsilon)$, its elements are quadratic polynomials in $\epsilon$. We can develop

$$Q(\epsilon) = Q + \epsilon Q_1 + \epsilon^2 Q_2, \tag{3.73}$$

where $Q$ is a stochastic matrix with exactly one 1 in each row and $Q_1$ and $Q_2$ have row sums 0. We may view $Q(\epsilon)$ as a perturbation of the matrix $Q$ and treat the problem of finding the left eigenvector $\mathbf{s}(\epsilon)$ of $Q(\epsilon)$ as a perturbation problem. Thus we set

$$\mathbf{s}(\epsilon) = \pi + \epsilon\mathbf{x} + \epsilon^2\mathbf{y} + \cdots, \tag{3.74}$$

where the stochastic vector $\pi$ is a solution of the unperturbed eigenvalue problem $\pi Q = \pi$, whereas the components of the vectors $\mathbf{x}$ and $\mathbf{y}$ must sum up to 0. By expanding $\mathbf{s}(\epsilon)Q(\epsilon) = \mathbf{s}(\epsilon)$ and comparing powers of $\epsilon$, this yields not only the limiting value $\pi$ for the payoff (if $\epsilon \to 0$), but also the first order term $\mathbf{x}$.

Let us consider, for example, $S_8 = (1, 0, 0, 0)$ against $S_{11} = (1, 0, 1, 1)$. $S_8$ is also called *Grim*, because it is a grim variant of *TFT*, prescribing to defect except after a round of mutual cooperation; whereas $S_{11}$, also know as *Firm But Fair* (*FBF*) is a tolerant brother of *TFT*, prescribing to play C if both players defected in the previous round. In that case,

$$Q = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \tag{3.75}$$

which is a reducible matrix, and

$$Q_1 = \begin{pmatrix} -2 & 1 & 1 & 0 \\ 1 & 0 & -2 & 1 \\ 0 & 1 & 1 & -2 \\ 1 & 0 & -2 & 1 \end{pmatrix}. \tag{3.76}$$

The equation $\pi Q = \pi$ yields $\pi_2 = 0$ and $\pi_3 = \pi_4$, i.e., $\pi = (1 - 2a, 0, a, a)$ for unknown $a$. The equation $\pi Q_1 + \mathbf{x}Q = \mathbf{x}$ yields $a = 2/5$, so that $\pi = (1/5, 0, 2/5, 2/5)$. We note that in this case, we did not need the $\epsilon^2$ term, but sometimes we do.

In table 3.1 we display, for the Donation game, the resulting $16 \times 16$ matrix $A$, with $a_{ij}$ denoting the payoff for an $S_i(\epsilon)$ player against a $S_j(\epsilon)$ player (or more

Table 3.1 The Simultaneous Donation Game with Errors in Implementation.

| $*$ | $S_0$ | $S_1$ | $S_2$ | $S_3$ | $S_4$ | $S_5$ | $S_6$ | $S_7$ | $S_8$ | $S_9$ | $S_{10}$ | $S_{11}$ | $S_{12}$ | $S_{13}$ | $S_{14}$ | $S_{15}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $S_0$ | $\mathbf{0}$ | $\frac{b}{2}$ | $0$ | $\frac{b}{2}$ | $\frac{b}{3}$ | $b$ | $\frac{b}{2}$ | $b$ | $0$ | $\frac{b}{2}$ | $0$ | $\frac{b}{2}$ | $\frac{b}{2}$ | $b$ | $\frac{2b}{3}$ | $b$ |
| $S_1$ | $-\frac{c}{2}$ | $\mathbf{\frac{b-c}{2}}$ | $\frac{b-c}{3}$ | $\frac{b-c}{2}$ | $\frac{b-2c}{5}$ | $\frac{2b-c}{3}$ | $b$ | $\frac{3b-c}{4}$ | $-\frac{c}{2}$ | $\frac{2b-c}{3}$ | $\frac{b-c}{3}$ | $\frac{2b-c}{3}$ | $\frac{2b-c}{4}$ | $b$ | $b$ | $b$ |
| $S_2$ | $0$ | $\frac{b-c}{3}$ | $\mathbf{\frac{b-c}{4}}$ | $\frac{b-c}{2}$ | $0$ | $\frac{2b-c}{3}$ | $0$ | $\frac{2b-c}{3}$ | $0$ | $\frac{b-c}{3}$ | $\frac{b-c}{3}$ | $\frac{b-c}{2}$ | $\frac{2b-c}{4}$ | $\frac{2b-c}{2}$ | $\frac{4b-2c}{5}$ | $\frac{2b-c}{2}$ |
| $S_3$ | $-\frac{c}{2}$ | $\frac{b-c}{2}$ | $\frac{b-c}{2}$ | $\mathbf{\frac{b-c}{2}}$ | $-\frac{c}{2}$ | $\frac{b-c}{2}$ | $\frac{b-c}{2}$ | $\frac{b-c}{2}$ | $-\frac{c}{2}$ | $\frac{b-c}{2}$ | $\frac{b-c}{2}$ | $\frac{b-c}{2}$ | $\frac{b-c}{2}$ | $\frac{2b-c}{2}$ | $\frac{2b-c}{2}$ | $\frac{2b-c}{2}$ |
| $S_4$ | $-\frac{c}{3}$ | $\frac{2b-c}{5}$ | $0$ | $\frac{b}{2}$ | $\mathbf{\frac{b-c}{4}}$ | $b$ | $\frac{b}{3}$ | $b$ | $-\frac{c}{3}$ | $\frac{2b-c}{5}$ | $0$ | $\frac{b}{2}$ | $\frac{3b-c}{6}$ | $b$ | $\frac{2b}{3}$ | $b$ |
| $S_5$ | $-c$ | $\frac{b-2c}{3}$ | $\frac{b-2c}{3}$ | $\frac{b-c}{2}$ | $-c$ | $\mathbf{\frac{b-c}{2}}$ | $\frac{b-c}{2}$ | $\frac{2b-c}{3}$ | $-c$ | $\frac{b-c}{2}$ | $\frac{b-c}{2}$ | $\frac{2b-c}{3}$ | $\frac{b-c}{2}$ | $b$ | $b$ | $b$ |
| $S_6$ | $-\frac{c}{2}$ | $-c$ | $0$ | $\frac{b-c}{2}$ | $-\frac{c}{3}$ | $\frac{b-c}{2}$ | $\mathbf{0}$ | $\frac{2b-c}{3}$ | $-\frac{2c}{3}$ | $-c$ | $\frac{b-c}{2}$ | $\frac{2b-2c}{3}$ | $\frac{b-c}{2}$ | $\frac{4b-3c}{5}$ | $\frac{4b-2c}{5}$ | $\frac{2b-c}{2}$ |
| $S_7$ | $-c$ | $\frac{b-3c}{4}$ | $\frac{b-2c}{3}$ | $\frac{b-c}{2}$ | $-c$ | $\frac{b-2c}{3}$ | $\frac{b-2c}{3}$ | $\mathbf{\frac{b-c}{2}}$ | $-c$ | $-c$ | $\frac{2b-2c}{3}$ | $\frac{2b-2c}{3}$ | $\frac{2b-3c}{4}$ | $\frac{4b-3c}{5}$ | $\frac{2b-c}{2}$ | $\frac{2b-c}{2}$ |
| $S_8$ | $0$ | $\frac{b}{2}$ | $0$ | $\frac{b}{2}$ | $\frac{b}{3}$ | $b$ | $\frac{2b}{3}$ | $b$ | $\mathbf{0}$ | $\frac{3b-c}{5}$ | $0$ | $\frac{3b-c}{5}$ | $\frac{3b-c}{6}$ | $\frac{3b-c}{3}$ | $\frac{3b-c}{4}$ | $\frac{3b-c}{3}$ |
| $S_9$ | $-\frac{c}{2}$ | $\frac{b-2c}{3}$ | $\frac{b-c}{3}$ | $\frac{b-c}{2}$ | $\frac{b-2c}{5}$ | $\frac{b-c}{2}$ | $b$ | $b$ | $\frac{b-3c}{5}$ | $\mathbf{b-c}$ | $\frac{b-c}{2}$ | $b-c$ | $\frac{b-c}{2}$ | $\frac{3b-2c}{3}$ | $\frac{3b-c}{3}$ | $\frac{2b-c}{2}$ |
| $S_{10}$ | $0$ | $\frac{b-c}{3}$ | $\frac{b-c}{3}$ | $\frac{b-c}{2}$ | $0$ | $\frac{b-c}{2}$ | $\frac{b-c}{2}$ | $\frac{2b-2c}{3}$ | $0$ | $\frac{b-c}{2}$ | $\mathbf{\frac{b-c}{2}}$ | $\frac{2b-2c}{3}$ | $\frac{b-c}{2}$ | $b-c$ | $b-c$ | $b-c$ |
| $S_{11}$ | $-\frac{c}{2}$ | $\frac{b-2c}{3}$ | $\frac{b-c}{2}$ | $\frac{b-c}{2}$ | $-\frac{c}{2}$ | $\frac{b-2c}{3}$ | $\frac{2b-2c}{3}$ | $\frac{2b-2c}{3}$ | $\frac{b-3c}{5}$ | $b-c$ | $\frac{2b-2c}{3}$ | $\mathbf{\frac{3b-3c}{4}}$ | $\frac{2b-3c}{4}$ | $b-c$ | $b-c$ | $b-c$ |
| $S_{12}$ | $-\frac{c}{2}$ | $\frac{b-2c}{4}$ | $\frac{b-2c}{4}$ | $\frac{b-c}{2}$ | $\frac{b-3c}{6}$ | $\frac{b-c}{2}$ | $\frac{b-c}{2}$ | $\frac{3b-2c}{4}$ | $\frac{b-3c}{6}$ | $\frac{b-c}{2}$ | $\frac{b-c}{2}$ | $\frac{3b-2c}{4}$ | $\mathbf{\frac{b-c}{2}}$ | $\frac{5b-3c}{6}$ | $\frac{5b-3c}{6}$ | $\frac{2b-c}{2}$ |
| $S_{13}$ | $-c$ | $-c$ | $\frac{b-2c}{2}$ | $\frac{b-2c}{2}$ | $-c$ | $-c$ | $\frac{3b-4c}{5}$ | $\frac{3b-4c}{5}$ | $\frac{b-3c}{3}$ | $\frac{2b-3c}{3}$ | $b-c$ | $b-c$ | $\frac{3b-5c}{6}$ | $\mathbf{\frac{3b-3c}{4}}$ | $\frac{3b-2c}{3}$ | $\frac{3b-2c}{3}$ |
| $S_{14}$ | $\frac{-2c}{3}$ | $-c$ | $\frac{2b-4c}{5}$ | $\frac{b-2c}{2}$ | $\frac{-2c}{3}$ | $-c$ | $\frac{2b-4c}{5}$ | $\frac{b-2c}{2}$ | $\frac{b-3c}{4}$ | $\frac{b-3c}{3}$ | $b-c$ | $b-c$ | $\frac{3b-5c}{6}$ | $\frac{2b-3c}{3}$ | $\mathbf{b-c}$ | $b-c$ |
| $S_{15}$ | $-c$ | $-c$ | $\frac{b-2c}{2}$ | $\frac{b-2c}{2}$ | $-c$ | $-c$ | $\frac{b-2c}{2}$ | $\frac{b-2c}{2}$ | $\frac{b-3c}{3}$ | $\frac{b-2c}{2}$ | $b-c$ | $b-c$ | $\frac{b-2c}{2}$ | $\frac{2b-3c}{3}$ | $b-c$ | $\mathbf{b-c}$ |

precisely, its limit for $\epsilon \to 0$). We note an obvious symmetry: if $a_{ij} = xb - yc$, then $a_{ji} = yb - xc$.

If the resident population is playing $S_0$, i.e., *AllD*, then no strategy can invade except $S_2 = (0, 0, 1, 0)$, the "*Grim*" strategy $S_8 = (1, 0, 0, 0)$, and the *TFT* strategy $S_{10} = (1, 0, 1, 0)$. Since $S_2$ is dominated by $S_{10}$ (in the absence of other strategies), this means that *TFT* can overcome *AllD*. But *TFT* can be superseded by more tolerant strategies, such as $S_{15}$, i.e., *AllC*, and these can in turn be displaced by *AllD*. However, this tendency to cycle can be broken up by $S_9$. This strategy dominates $S_{10}$, $S_2$, and (if $b > 3c$) also $S_8$, and it cannot be invaded by *AllD* as long as $b > 2c$, i.e., the cost-to-benefit ratio is less than $1/2$.

We note that $S_9$ is the only strategy that cannot be invaded by any other $S_i$ (for $b > 3c$). Moreover, $S_9$ is very good against itself: a population of $S_9$ players earns $b - c$, which is the best a homogeneous population can achieve. Only $S_{14}$ and $S_{15}$ do as well, but these are easy prey to $S_1$ or $S_0$.

## 3.15 WIN-STAY, LOSE-SHIFT

The strategy $S_9 = (1, 0, 0, 1)$, for reasons difficult to fathom, is called *Pavlov*. It has the remarkable property of being *error-correcting*. If two players using *Pavlov* play against each other, they will cooperate most of the time. If player II, say, defects by mistake, then in the next round both players will play D, and thereafter resume mutual cooperation, like an old couple after a row (see fig. 3.8). Moreover, if a *Pavlov* player plays against *AllC*, it will shamelessly exploit the co-player. After the first accidental D, it will continue playing D until a further error occurs. This is an important property for safeguarding the population against eventual invasions by defectors. A *TFT* population, for instance, will quickly be subverted by *AllC* players, and these will be open to exploitation by *AllD*.

| | | |
|---|---|---|
| (a) | Pavlov-player I | C C C ... C C D C C ... |
| | Pavlov-player II | C C C ... C D D C C ... |
| (b) | TFT-player I | C C C ... C C D C D ... |
| | TFT-player II | C C C ... C D C D C ... |
| (c) | ALL C-player I | C C C ... C C C C ... |
| | Pavlov-player II | C C C ... C D D D ... |
| (d) | ALL C-player I | C C C ... C C C C ... |
| | TFT-player II | C C C ... C D C C ... |

Figure 3.8 The effect of an erroneous defection in the iterated Prisoner's Dilemma game. The arrow denotes the mis-implemented move in each run.

| | | | | |
|---|---|---|---|---|
| Pavlov-player I | C | C | D | D |
| Player II | C | D | C | D |
| | ↓ | ↓ | ↓ | ↓ |
| Payoff for I | R | S | T | P |
| Next move for I | C | D | D | C |

Figure 3.9 *Pavlov* as a Win-Stay, Lose-Shift strategy. After obtaining the larger payoff values $T$ and $R$, a *Pavlov* player repeats the former, successful move. After obtaining the smaller payoff values $P$ and $S$, the *Pavlov* player switches to the other move.

The strategy $S_9$ prescribes playing C if and only if, in the previous round, the co-player did the same as the other player. There is a suggestive property behind this mechanism, see figure 3.9. The strategy effectively repeats the previous move if it obtained a positive payoff (a reward, such as $b - c$, or better still the temptation $b$). It switches to the other move if the payoff was non-positive (payoff 0 if both players defected, or the sucker's payoff $-c$). This is the simplest conceivable learning mechanism, well-known to animal trainers and parents alike. *Win-Stay, Lose-Shift* is a wide-spread maxim of animal behavior.

The condition $b > 2c$ implies that *Pavlov* is not dominated by *AllD*, but that the two strategies are engaged in a bi-stable competition. The condition $b > 3c$ implies that *Pavlov* is risk-dominant.

It is interesting to consider finite populations in this context. Let us consider the two cases (a) $b = 5c/2$ and (b) $b = 4c$, a population size $M = 100$ and selection strength $s = 1/10$. Let us also assume the adiabatic case (very small innovation rates $\mu$, see section 2.17). A population consisting only of the types $S_0 = AllD$ and $S_{10} = TFT$ will be dominated by *TFT*. (In the numerical example, *TFT* occurs with 97 percent in the stationary distribution given by expression (2.90) in case (a), and with 99 percent in case (b).) This reflects the fact that *TFT* dominates *AllD*. But if $AllC = S_{15}$ is also allowed in the population, then the stationary distribution is dominated by *AllD* (64 percent in case (a) and 66 percent in case (b)). Now let us consider a population with the strategies $S_9 = Pavlov$, *AllC*, and *AllD*. If $b > 3c$, *Pavlov* risk-dominates *AllD*. This corresponds to example (b), and we see indeed that the stationary distribution consists of 90 percent *Pavlov*. If $2c < b < 3c$, *AllD* risk-dominates *Pavlov*, and we find 80 percent of defectors in the stationary distribution. This changes dramatically if we also include *TFT*: in that case, example (a) leads to 50 percent *Pavlov* (and example (b) to 95 percent). Thus *TFT* is not the winner, but can act as a king-maker—decisive for the outcome of the contest between *AllD* and *Pavlov*.

In the general Prisoner's Dilemma game, *Pavlov* acts according to a threshold separating the two better outcomes $T$ and $R$ from the two worse outcomes $P$ and $S$. A move yielding an outcome above the *aspiration level* is repeated, a move yielding an outcome below the aspiration level is not. One can consider other aspiration levels. If the aspiration level is more ambitious, content only with $T$, this leads to the strategy $S_1 = (0, 0, 0, 1)$, a bully-like strategy that only cooperates after a mutual defection. It relentlessly defects whenever it can exploit a sucker, but switches as soon as it meets a defection. This is an overly ambitious Win-Stay, Lose-Shift strategy, and it fails. Similarly, a more modest aspiration level (between $P$ and $S$) leads to $S_8 = (1, 0, 0, 0)$, which is doing rather well, especially in a population of defectors. Finally, one could also view $S_3 = (0, 0, 1, 1)$ and $S_{12} = (1, 1, 0, 0)$ as extreme forms of Win-Stay, Lose-Shift strategies. The former switches its move from one round to the next, never satisfied by any outcome. The latter never switches except by mistake, and always repeats itself, apparently content with every outcome. Obviously, it is important to have the "right" aspiration level. *Pavlov* is, in this sense, the most balanced of all Win-Stay, Lose-Shift rules. It is also doing well in the iterated Snowdrift game, which is described in matrix (3.4).

## 3.16 AUTOMATA

Memory-one strategies with deterministic reaction norms can easily be implemented by finite state automata. For instance, *Pavlov* can be implemented by an automaton with only two inner states, which we shall denote by *Same* and *Diff*. The automaton is in state *Same* if in the previous round, the player experienced payoff $R$ or $P$ (i.e., both players cooperated, or both defected), and it is in state *Diff* otherwise. In each state, the automaton prescribes the next move, i.e., C or D. The analysis of *Pavlov* can then be performed very easily by means of a directed graph, see figure 3.10. The nodes of the graph are the states *Same* and *Diff* of the player. Two directed arrows are leaving from each node, one solid and the other dashed. The solid arrow describes the transition if the player uses the move prescribed by *Pavlov* (i.e., C in *Same* and D in *Diff*). The dashed arrow describes the transition if the other move is used. In both cases, the co-player is assumed to follow the *Pavlov* strategy. Along each arrow, one can see the corresponding payoff of the player. Clearly, it is best always to follow the solid arrow, if $2R > T + P$ (or, in the case $w < 1$, if $R + wR > T + wP$). For the Donation game, this reduces to the familiar condition $b > 2c$ (resp. $w(b - c) > c$). If this condition holds, it is always best, against a *Pavlov* player, to do what the *Pavlov* rule prescribes. In a population of *Pavlov* players, it is best to follow suit. Similar graphs can be studied for all memory-one strategies. In general, this will be more complicated than for *Pavlov*, where the two players are always in the same state. But the four states $(C, C)$, $(C, D)$, $(D, C)$ and $(D, D)$ will always be enough to

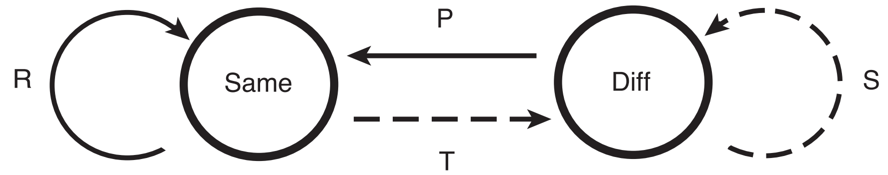


Figure 3.10 The *Pavlov* strategy described by a two-state automaton. The co-player is assumed to play *Pavlov*. The solid and dashed arrows respectively describe the transition if the player follows the *Pavlov* strategy or deviates from it. *Pavlov* is a best reply to itself if $2R > T + P$ (in the Donation game, if $b > 2c$).

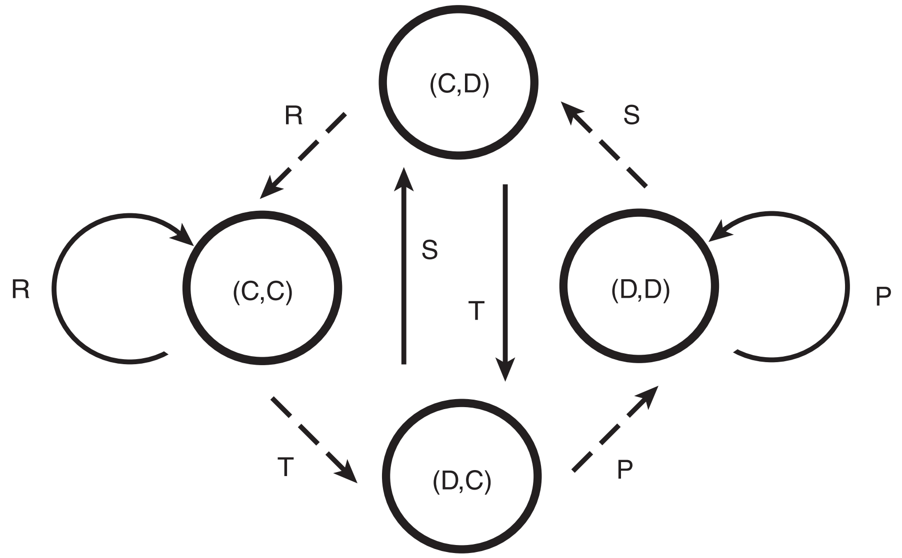


Figure 3.11 The *TFT* strategy described by a finite automaton. It is not a best reply to itself if $2R > T + S$, a condition that always holds for the Donation game.

describe the automaton (the first entry describes the move of the player, the second that of the co-player). In figure 3.11, we describe what happens when the co-player uses the strategy *TFT*. The solid arrow leaving a node describes the transition if the player, at that node, also uses the move prescribed by *TFT*, and the dashed arrow the outcome of the alternative move.

We note that *TFT* is not the best answer against itself in the Donation game. In state $(C, D)$, the best move against a *TFT* player would be to cooperate (and to reach state $(C, C)$). However, *TFT* calls for one to play D, and this locks two *TFT* players into an endless cycle of unilateral defections. The payoff per round, then, is $(b - c)/2$, which is less than the payoff per round $b - c$ obtained if the node $(C, C)$ had been reached. Of course, two *TFT* players would start out at node $(C, C)$, and in that state *TFT* prescribes the right move. But an error leading to node $(C, D)$ displays the fatal weakness of *TFT*. More generally, this strategy is not a best answer to itself if $2R > T + S$.

The method can be extended for all strategies where memory depends on the last two, or the last $N$ rounds. But we shall presently see that some very simple strategies implemented by finite automata cannot be described as strategies conditioned on a prescribed number of rounds.

## 3.17 CONTRITE TIT FOR TAT

An interesting example for this is *CTFT* (*Contrite TFT*). Imagine that a *TFT* player who mis-implemented a C move is aware of having done wrong, and accepts meekly that the co-player, in the next round, defects in retaliation. In this case, the pointless vendetta of alternating unilateral defection can be avoided and mutual cooperation resumed. To model this, let us introduce the *standing* of a player, which

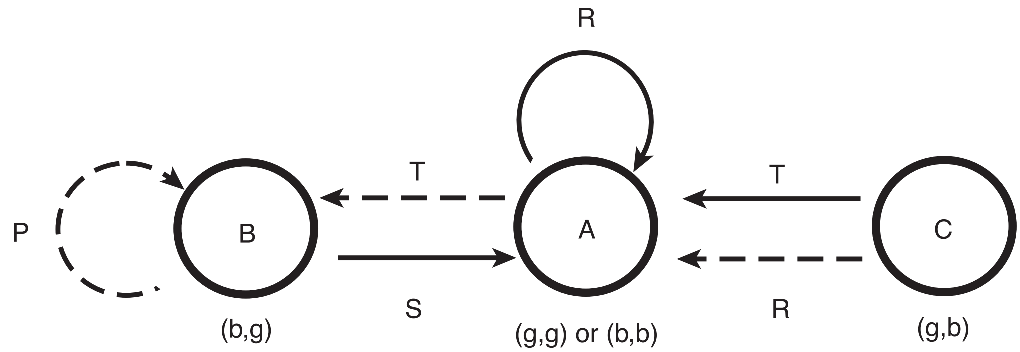


Figure 3.12 The *Contrite TFT* strategy described by a 3-state automaton. It always is a best reply to itself.

can be $g$ or $b$ ("good" or "bad"). Players start out in good standing and keep it until they commit an unjustified defection (i.e., until they play D while the co-player was in good standing). The good standing is regained by playing C. In any given round, a player can cooperate, commit a justified defection or an unjustified defection.

*Contrite TFT* is the strategy that calls for one to cooperate except when in good standing while the other player is not. This means that the player defects when provoked, but not otherwise. Thus if two *Contrite TFT* players engage in a repeated Prisoner's Dilemma game, they will always cooperate, except by mistake. After such a mistake, they will resume cooperation, and accept the co-player's retaliatory D without feeling abused.

As before, we can describe the game with a graph, and check that if the other player uses *Contrite TFT*, it is always best to also use the move prescribed by *Contrite TFT*, see figure 3.12. The nodes of the graph, i.e., the state of the game, will be $A$, $B$, and $C$. $A$ corresponds to $(g, g)$ or $(b, b)$, $B$ to $(b, g)$ (the player's standing is bad and the co-player's good), and $C$ to $(g, b)$. We note that if one player is in state $A$, so is the other, whereas if one player is in state $B$, the other is in state $C$ and vice versa. The rule for playing *Contrite TFT* calls for one to use C when in state $A$ or $B$, and to defect (i.e., use D) in state $C$ only. The corresponding graph shows immediately that it is best, against a *Contrite TFT* player, also to use the *Contrite TFT* rule.

The *Contrite TFT* strategy cannot be described as a memory-one strategy. Neither does it follow a rule that depends only on the outcome of a given number $N$ of preceding rounds. Indeed, suppose that we observe a string of mutual defections in the previous $N$ rounds. If we do not know what happened before these $N$ rounds, we cannot say who of the two players is in good, and who is in bad standing. Hence we cannot specify what a *Contrite TFT* player ought to do in the next round.

It is interesting to compare *Contrite TFT* and *Pavlov* for the Donation game. *Contrite TFT* is always a best response to itself, *Pavlov* only if $b > 2c$. Both strategies cooperate with their like, and can easily return to mutual cooperation after an accidental error in implementation. *Contrite TFT* has the huge advantage that it is as good as *TFT* at invading a population of *AllD* players; *Pavlov*, as we have seen, is hopeless at this task, and needs a retaliatory strategy to pave the way. On the other hand, in a society dominated by *Contrite TFT*, indiscriminate altruists do just as well and hence can spread by neutral drift, eventually allowing *AllD* to invade and destroy the cooperative regime. By contrast, a society of *Pavlov* players will not allow *AllC* players to spread. As soon as the first error of implementation occurs, an *AllC* player will be exploited to the hilt.

If *Pavlov* does not fare well, i.e., if $c < b < 2c$, another strategy based on standing fares as well as *Contrite TFT*: this is *Remorse*, a strategy where a player cooperates only when in bad standing, or if both players had cooperated in the previous round. After a unilateral error, two *Remorse* players defect twice. If a *Remorse* player encounters a *Pavlov* player, both obtain an average payoff of $5(b-c)/7$ per round.

## 3.18 ERRORS IN PERCEPTION

*Contrite TFT* has its Achilles heel, too. So far, we have only considered errors in implementation. What about errors in perception? In that case, players can believe themselves to be in good standing, whereas their co-player sees them in bad standing. Two *Contrite TFT* players will, in such a situation, relentlessly inflict D upon each other, both believing that their own moves are justified defections and that their co-player's moves are not. In contrast, if an error in perception occurs between two *Pavlov* players, cooperation will be smoothly resumed after the usual mutual punishment round.

In the realm of memory-one strategies, if there is a probability $\epsilon$ to mis-implement a move, then the propensity $q_R$ to play C after a round with outcome $R$ is replaced by the propensity $(1-\epsilon)q_R + \epsilon(1-q_R)$, etc., so that the "correction term"

$$\epsilon(1-2q_R, 1-2q_S, 1-2q_T, 1-2q_P) \tag{3.77}$$

has to be added to the reaction norm $(q_R, q_S, q_T, q_P)$. If the error affects the perception of the co-player's move (i.e., if the player confuses an $R$ with an $S$, or a $T$ with a $P$) then $q_R$ turns into $(1-\nu)q_R + \nu q_S$ etc., and the correction term is

$$\nu(q_S - q_R, q_R - q_S, q_P - q_T, q_T - q_P). \tag{3.78}$$

If the error $\mu$ affects the perception of the player's own move (i.e., a player confuses an $R$ with a $T$, or an $S$ with a $P$), then the correction term is

$$\mu(q_T - q_R, q_P - q_S, q_R - q_T, q_S - q_P). \tag{3.79}$$

If both types of errors in perception are admitted, then the reaction norm of *TFT*, i.e., $(1, 0, 1, 0)$, turns into $(1-\nu, \nu, 1-\nu, \nu)$, and *Pavlov* $(1, 0, 0, 1)$ is modified into $(1-(\nu+\mu), \nu+\mu, \nu+\mu, 1-(\nu+\mu))$, whereas the unconditional strategies *AllD* and *AllC* are unaffected. For $w = 1$ and the limit $\epsilon \to 0$, errors in implementation yield as payoff $(2P+2S+T)/5$ for an $S_8$ player using the *Grim* strategy $(1, 0, 0, 0)$ against an $S_2$ player using $(0, 0, 1, 0)$, whereas errors in perceiving the opponent's move yield as payoff $(S+T)/2$, etc.

Thus it is important to consider different possibilities of errors. For instance, we might make (as in section 3.10) the plausible assumption that errors occur only if one wants to implement a C, but not if one decides to play D. In this case, the stationary distribution for an $S_8(\varepsilon)$ player against an $S_{11}(\varepsilon)$ player is $(0, 0, 1/2, 1/2)$ instead of $(1/5, 0, 2/5, 2/5)$ (see section 3.14), and thus the payoff is $b/2$ instead of $(3b - c)/5$ (up to terms in $\varepsilon$). The tolerant *FirmButFair* player tries vainly, every second round, to resume cooperation. The payoff for *TFT* against itself is worse now (namely 0), but the payoff in a *Pavlov* population remains unchanged. Again, a *Pavlov* population cannot be invaded if $b > 2c$.

Even among automata with only three or four inner states, there exists a bewildering number of strategies. It seems hard to figure out which one would be selected by evolution. Individual based simulations display a lot of contingencies, and offer few robust predictions. We run up against a complexity wall. On the other hand, it seems tempting to interpret the "inner states" of automata with our emotions, such as anger at being provoked, guilt at having deviated from the norm, etc.

## 3.19 TRIGGERS AND EQUALIZERS

The so-called *folk theorem on repeated games* is a collection of results. In the simplest setup, for two players I and II engaged in an infinitely repeated Donation game, it states that any pair $(P_\mathrm{I}, P_\mathrm{II})$ of payoff values (per round) with $0 \le P_\mathrm{I}, P_\mathrm{II} \le b - c$ can be realized by a Nash equilibrium pair of strategies. The two players simply have to follow so-called *trigger strategies*: this means playing a well-specified sequence of moves leading to $(P_\mathrm{I}, P_\mathrm{II})$, but switching to a relentless, infinite sequence of D moves as soon as the co-player deviates. It is obvious, then, that the co-player has no incentive to deviate: there is no better alternative than to follow the specified sequence of moves. In fact, any pair of payoff values can be reached such that $P_\mathrm{I}$ and $P_\mathrm{II}$ are positive and $(P_\mathrm{I} P_\mathrm{II})$ in the convex hull spanned by $(0, 0)$, $(b, -c)$, $(-c, b)$, and $(b - c, b - c)$, see figure 3.13.

This result can be extended in many ways, by considering iterations of other games (the lower bound 0 will then have to be replaced by the maximin payoff, i.e., the highest payoff that players can guarantee themselves, irrespective of their co-player's strategy), by introducing a discount on future payoffs (or allowing the iteration to stop with a positive probability), by admitting the possibility that players mis-implement their moves, etc.

The concept of a trigger strategy is often criticized on the grounds that it is too stern: it is hard to imagine that players will commit themselves forever to ruinous defection, if their co-player made a mistake just once, possibly through *force majeure*. Nevertheless, trigger strategies are an essential tool for analyzing games between rational players. In evolutionary game theory, however, trigger strategies play a less conspicuous role.

It turns out that a variant of the folk theorem can easily be displayed in the context of memory-one strategies. Indeed, there exist strategies that act as *equalizers*, in the sense that co-players always obtain the same payoff, irrespective of their strategy.

For the infinitely repeated Prisoner's Dilemma game, there exist, for every value $\pi$ between $P$ and $R$, memory-one strategies $\mathbf{q} = (q_R, q_S, q_T, q_P)$ such that every op-

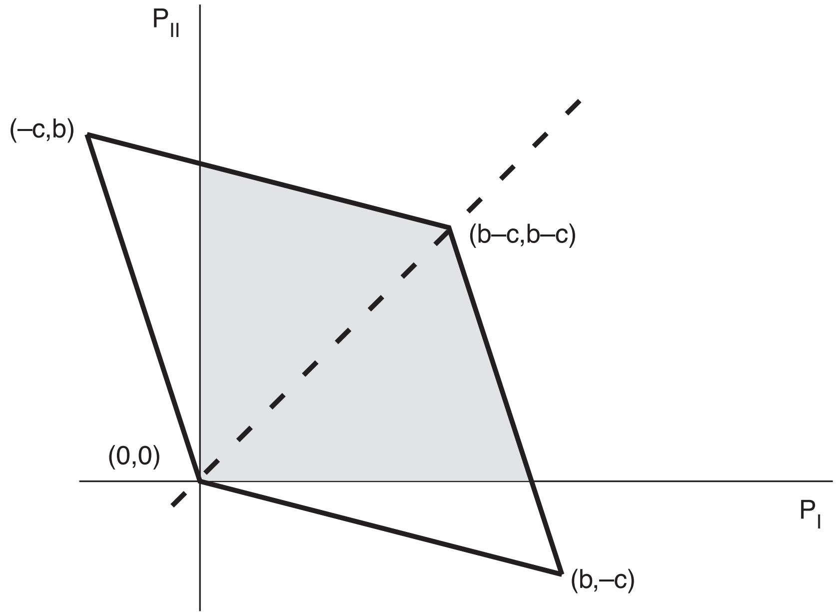


Figure 3.13 Any pair of payoff values ($P_{\mathrm{I}}$, $P_{\mathrm{II}}$) in the shaded region can be obtained if the two players I and II use the corresponding equalizer strategy for the infinitely repeated Prisoner's Dilemma game.

ponent obtains the long-run average payoff $\pi$ against a player using such a strategy. The reaction norm $\mathbf{q}$ is given by

$$(1-(R-\pi)a, 1-(T-\pi)a, (\pi-S)a, (\pi-P)a), \tag{3.80}$$

where $a > 0$ is any real number such that $\frac{1}{a} \geq \max\{T-\pi, R-\pi, \pi-S, \pi-P\}$. (The condition on $a$ guarantees that the $q_i$ are probabilities.)

Indeed, let us denote by $p_i(n)$ the conditional probability that the player II uses the move C in round $n+1$, given that the $n$-th round resulted in outcome $i$ for player I; and let $s_i(n)$ be the probability of that outcome. By conditioning on round $n$, we see that $s_R(n+1)$ is given by

$$\begin{aligned} &s_R(n)p_R(n)[1-(R-\pi)a] + s_S(n)p_S(n)[1-(T-\pi)a] \\ &\quad + s_T(n)p_T(n)(\pi-S)a + s_P(n)p_P(n)(\pi-P)a. \end{aligned} \tag{3.81}$$

Similarly, $s_S(n+1)$ is given by

$$\begin{aligned} &s_R(n)(1-p_R(n))[1-(R-\pi)a] + s_S(n)(1-p_S(n))[1-(T-\pi)a] \\ &\quad + s_T(n)(1-p_T(n))(\pi-S)a + s_P(n)(1-p_P(n))(\pi-P)a. \end{aligned} \tag{3.82}$$

Summing these equations yields the probability that player I chooses move C in round $n+1$, namely $s_R(n+1)+s_S(n+1)$. It is given by

$$s_R(n)[1-(R-\pi)a]+s_S(n)[1-(T-\pi)a]+s_T(n)(\pi-S)a+s_P(n)(\pi-P)a. \tag{3.83}$$

Hence

$$s_R(n)+s_S(n)-s_R(n+1)-s_S(n+1) \tag{3.84}$$

is given, up to the factor $a$, by

$$Rs_R(n) + Ss_T(n) + Ts_S(n) + Ps_P(n) - \pi[s_R(n) + s_S(n) + s_T(n) + s_P(n)]. \quad (3.85)$$

Since the $s_i(n)$ sum up to 1, this is just $A'(n) - \pi$, where $A'(n)$ is player II's payoff in the $n$-th round. (We must bear in mind that one player's $S$ is the other player's $T$.) Summing up for $n = 0, \dots, N$ and dividing by $N + 1$, we obtain

$$\frac{s_R(0) + s_S(0) - s_R(N) - s_S(N)}{a(N+1)} = \frac{A'(0) + \cdots + A'(N)}{N+1} - \pi, \quad (3.86)$$

which yields

$$\lim_{N\to\infty} \frac{A'(0) + \cdots + A'(N)}{N+1} = \pi. \quad (3.87)$$

## 3.20 THE ALTERNATING PRISONER'S DILEMMA

In many real-life instances of direct reciprocity, the two players alternate in their roles of donor and recipient, whereas in most of the literature, and in our treatment so far, the two players decide simultaneously. Usually, this assumption is of small effect. But in some situations, important differences can arise.

In an alternating Prisoner's Dilemma (or the alternating Donation game), to co-operate means to play C when it is one's turn to do so. This can affect strategies and payoffs. For instance, if two *TFT* players engage in an iterated Prisoner's Dilemma of the usual, simultaneous kind, and if one player defects by mistake, both players will subsequently play C and D in turn. In the alternating Prisoner's Dilemma game, if a unilateral defection occurs by mistake, the result will be a sequence of mutual defections: both players play D, see figure 3.14. The average payoff will be the same as in the simultaneous case, and in fact, the interplay between *TFT*, *AllC*, and *AllD* remains unchanged. But if two *Pavlov* players, for instance, are matched against each other, the outcome is very different. A mistaken D is answered by a D, which elicits a C, which is followed by a D in turn. Thus each player, after the erroneous defection, keeps playing two D's and one C periodically. With probability 2/3, the

(a)
TFT-player I C C C C D D ...
TFT-player II C C C D D D ...

(b)
Pavlov-player I C C C C D D C D ...
Pavlov-player II C C C D C D D ...

Figure 3.14 The effect of an erroneous defection in the alternating Prisoner's Dilemma game, (a) between two *TFT* players, (b) between two *Pavlov* players. The arrow denotes the mis-implemented move in each run.

Table 3.2 The Alternating Donation Game with Errors in Implementation.

| $*$ | $S_0$ | $S_1$ | $S_2$ | $S_3$ | $S_4$ | $S_5$ | $S_6$ | $S_7$ | $S_8$ | $S_9$ | $S_{10}$ | $S_{11}$ | $S_{12}$ | $S_{13}$ | $S_{14}$ | $S_{15}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $S_0$ | $\mathbf{0}$ | $\frac{b}{2}$ | $0$ | $\frac{b}{2}$ | $\frac{b}{3}$ | $b$ | $\frac{b}{2}$ | $b$ | $0$ | $\frac{b}{2}$ | $0$ | $\frac{b}{2}$ | $\frac{b}{2}$ | $b$ | $\frac{2b}{3}$ | $b$ |
| $S_1$ | $-\frac{c}{2}$ | $\mathbf{\frac{b-c}{3}}$ | $\frac{b-c}{2}$ | $\frac{b-c}{2}$ | $\frac{b-c}{3}$ | $b$ | $b$ | $b$ | $-\frac{c}{2}$ | $\frac{b-c}{3}$ | $\frac{b-c}{2}$ | $\frac{b-c}{2}$ | $\frac{2b-c}{4}$ | $b$ | $b$ | $b$ |
| $S_2$ | $0$ | $\frac{b-c}{2}$ | $\mathbf{0}$ | $\frac{b-c}{2}$ | $0$ | $\frac{b-c}{2}$ | $0$ | $\frac{b-c}{2}$ | $0$ | $\frac{2b-c}{3}$ | $0$ | $\frac{2b-c}{3}$ | $\frac{2b-c}{4}$ | $\frac{2b-c}{2}$ | $\frac{2b-c}{3}$ | $\frac{2b-c}{2}$ |
| $S_3$ | $-\frac{c}{2}$ | $\frac{b-c}{2}$ | $\frac{b-c}{2}$ | $\mathbf{\frac{b-c}{2}}$ | $-\frac{c}{2}$ | $\frac{b-c}{2}$ | $\frac{b-c}{2}$ | $\frac{b-c}{2}$ | $-\frac{c}{2}$ | $\frac{b-c}{2}$ | $\frac{b-c}{2}$ | $\frac{b-c}{2}$ | $\frac{b-c}{2}$ | $\frac{2b-c}{2}$ | $\frac{2b-c}{2}$ | $\frac{2b-c}{2}$ |
| $S_4$ | $-\frac{c}{3}$ | $\frac{b-c}{3}$ | $0$ | $\frac{b}{2}$ | $\mathbf{\frac{b-c}{3}}$ | $\frac{2b-c}{3}$ | $\frac{2b}{3}$ | $b$ | $-\frac{c}{3}$ | $\frac{2b-c}{5}$ | $0$ | $\frac{b}{2}$ | $\frac{3b-c}{6}$ | $b$ | $\frac{2b}{3}$ | $b$ |
| $S_5$ | $-c$ | $-c$ | $\frac{b-c}{2}$ | $\frac{b-c}{2}$ | $\frac{b-2c}{3}$ | $\mathbf{\frac{b-c}{2}}$ | $b$ | $b$ | $-c$ | $-c$ | $\frac{b-c}{2}$ | $\frac{b-c}{2}$ | $\frac{b-c}{2}$ | $\frac{2b-c}{3}$ | $b$ | $b$ |
| $S_6$ | $-\frac{c}{2}$ | $-c$ | $0$ | $\frac{b-c}{2}$ | $-\frac{2c}{3}$ | $-c$ | $\mathbf{\frac{b-c}{2}}$ | $\frac{2b-2c}{3}$ | $-\frac{c}{3}$ | $\frac{b-c}{2}$ | $0$ | $\frac{2b-c}{3}$ | $\frac{b-c}{2}$ | $\frac{4b-3c}{5}$ | $\frac{4b-2c}{5}$ | $\frac{2b-c}{2}$ |
| $S_7$ | $-c$ | $-c$ | $\frac{b-c}{2}$ | $\frac{b-c}{2}$ | $-c$ | $-c$ | $\frac{2b-2c}{3}$ | $\mathbf{\frac{2b-2c}{3}}$ | $-c$ | $-c$ | $\frac{b-c}{2}$ | $\frac{b-c}{2}$ | $\frac{2b-3c}{4}$ | $\frac{2b-2c}{3}$ | $\frac{2b-c}{2}$ | $\frac{2b-c}{2}$ |
| $S_8$ | $0$ | $\frac{b}{2}$ | $0$ | $\frac{b}{2}$ | $\frac{b}{3}$ | $b$ | $\frac{b}{3}$ | $b$ | $\mathbf{0}$ | $\frac{3b-c}{5}$ | $\frac{b-c}{3}$ | $\frac{2b-c}{3}$ | $\frac{3b-c}{6}$ | $\frac{3b-c}{3}$ | $\frac{2b-c}{3}$ | $\frac{3b-c}{3}$ |
| $S_9$ | $-\frac{c}{2}$ | $\frac{b-c}{3}$ | $\frac{b-2c}{3}$ | $\frac{b-c}{2}$ | $\frac{b-2c}{5}$ | $b$ | $\frac{b-c}{2}$ | $b$ | $\frac{b-3c}{5}$ | $\mathbf{\frac{b-c}{2}}$ | $b-c$ | $b-c$ | $\frac{b-c}{2}$ | $\frac{3b-c}{3}$ | $\frac{3b-2c}{3}$ | $\frac{2b-c}{2}$ |
| $S_{10}$ | $0$ | $\frac{b-c}{2}$ | $0$ | $\frac{b-c}{2}$ | $0$ | $\frac{b-c}{2}$ | $0$ | $\frac{b-c}{2}$ | $\frac{b-c}{3}$ | $b-c$ | $\mathbf{\frac{b-c}{2}}$ | $b-c$ | $\frac{b-c}{2}$ | $b-c$ | $\frac{2b-2c}{3}$ | $b-c$ |
| $S_{11}$ | $-\frac{c}{2}$ | $\frac{b-c}{2}$ | $\frac{b-2c}{3}$ | $\frac{b-c}{2}$ | $-\frac{c}{2}$ | $\frac{b-c}{2}$ | $\frac{b-2c}{3}$ | $\frac{b-c}{2}$ | $\frac{b-2c}{3}$ | $b-c$ | $b-c$ | $\mathbf{b-c}$ | $\frac{2b-3c}{4}$ | $b-c$ | $b-c$ | $b-c$ |
| $S_{12}$ | $-\frac{c}{2}$ | $\frac{b-2c}{4}$ | $\frac{b-2c}{4}$ | $\frac{b-c}{2}$ | $\frac{b-3c}{6}$ | $\frac{b-c}{2}$ | $\frac{b-c}{2}$ | $\frac{3b-2c}{4}$ | $\frac{b-3c}{6}$ | $\frac{b-c}{2}$ | $\frac{b-c}{2}$ | $\frac{3b-2c}{4}$ | $\mathbf{\frac{b-c}{2}}$ | $\frac{5b-3c}{6}$ | $\frac{5b-3c}{6}$ | $\frac{2b-c}{2}$ |
| $S_{13}$ | $-c$ | $-c$ | $\frac{b-2c}{2}$ | $\frac{b-2c}{2}$ | $-c$ | $\frac{b-2c}{3}$ | $\frac{3b-4c}{5}$ | $\frac{2b-2c}{3}$ | $\frac{b-3c}{3}$ | $\frac{b-3c}{3}$ | $b-c$ | $b-c$ | $\frac{3b-5c}{6}$ | $\mathbf{\frac{2b-2c}{3}}$ | $\frac{3b-2c}{3}$ | $\frac{3b-2c}{3}$ |
| $S_{14}$ | $-\frac{2c}{3}$ | $-c$ | $\frac{b-2c}{3}$ | $\frac{b-2c}{2}$ | $-\frac{2c}{3}$ | $-c$ | $\frac{2b-4c}{5}$ | $\frac{b-2c}{2}$ | $\frac{b-2c}{3}$ | $\frac{2b-3c}{3}$ | $\frac{2b-2c}{3}$ | $b-c$ | $\frac{3b-5c}{6}$ | $\frac{2b-3c}{3}$ | $\mathbf{b-c}$ | $b-c$ |
| $S_{15}$ | $-c$ | $-c$ | $\frac{b-2c}{2}$ | $\frac{b-2c}{2}$ | $-c$ | $-c$ | $\frac{b-2c}{2}$ | $\frac{b-2c}{2}$ | $\frac{b-3c}{3}$ | $\frac{b-2c}{2}$ | $b-c$ | $b-c$ | $\frac{b-2c}{2}$ | $\frac{2b-3c}{3}$ | $b-c$ | $\mathbf{b-c}$ |

next mistake will not affect this regime. Only with probability 1/3 will it redress the game to a run of mutual cooperation. The average payoff is $(b - c)/2$ in the infinitely repeated case ($w = 1$).

If players alternate in being the potential donor, then two consecutive rounds of the alternating game correspond to one round of the simultaneous game. Let us assume that the memory of each player covers the previous two rounds (i.e., one decision by each player on whether to donate or not). The outcomes will be denoted in the obvious way by $R$, $S$, $T$, and $P$, and the strategies for the infinitely iterated alternating game by the propensities $q_R$, $q_S$, etc., to cooperate after outcome $R$, $S$, etc. The transition probabilities for a $(q_R, q_S, q_T, q_P)$ player against a $(q'_R, q'_S, q'_T, q'_P)$ player are given by the matrix

$$Q = \begin{pmatrix} q_R q'_R & q_R(1-q'_R) & (1-q_R)q'_S & (1-q_R)(1-q'_S) \\ q_S q'_T & q_S(1-q'_T) & (1-q_S)q'_P & (1-q_S)(1-q'_P) \\ q_T q'_R & q_T(1-q'_R) & (1-q_T)q'_S & (1-q_T)(1-q'_S) \\ q_P q'_T & q_P(1-q'_T) & (1-q_P)q'_P & (1-q_P)(1-q'_P) \end{pmatrix}, \tag{3.88}$$

which is quite different from matrix (3.67). The payoff can be computed as before. It turns out that in the alternating Prisoner's Dilemma, *Pavlov* loses much of its appeal. As table 3.2 shows, its place is taken up by *Firm But Fair*, with reaction norm $S_{11} = (1, 0, 1, 1)$. This strategy is error-correcting and achieves the highest payoff against itself, namely $b - c$, just as $S_{14}$ and $S_{15}$ do. But in contrast to these latter two strategies, *Firm But Fair* cannot be invaded by other strategies, such as *AllD*, as long as $b > 2c$. However, the strategy $S_{14}$ can enter by neutral drift. On the other hand, *AllD* $= S_0$ can always be invaded by $S_8$ and $S_{10}$, which in turn can be invaded by $S_{11}$. If we consider only errors in implementing a cooperative move, we see as in section 3.14 that *AllD* can be subverted by many strategies through neutral drift, and that among these, $S_2$, $S_6$, $S_{10}$, and $S_{14}$ give way to *Firm But Fair*.

If we restrict attention to reactive strategies, for which $q_R = q_T = p$ and $q_S = q_P = q$, we find that the payoffs for the donation game are exactly as for the simultaneous game, although the sequence of moves can be quite different (as we have seen in the instance of two *TFT* players). Again, *Generous TFT* emerges as the winner. Within the realm of strategies given by finite automata, *Contrite TFT* is as good in the alternating as in the simultaneous Prisoner's Dilemma, and as vulnerable to errors in perception.

## 3.21 REFERENCES

Trivers (1971) introduced reciprocal altruism as a major factor in the evolution of cooperation. The use of the repeated Prisoner's Dilemma was taken up by Axelrod and Hamilton (1981) and started a huge wave of investigations, see Axelrod (1984), Axelrod and Dion (1988), Trivers (2006), and Kendall, Yao, and Chong (2007). But this had been preceded by many theoretical and empirical investigations, see e.g., Rapoport and Chammah (1965). For various views on the importance of Tit for Tat among non-human players, we refer to Dugatkin (1997), Milinski (1987), and Hammerstein (2003). The Chicken game, also known as Hawk-Dove game, played an essential role in the early development of evolutionary game theory, see Maynard Smith (1982). In the form of the Snowdrift game, its relevance for cooperation was highlighted by Sugden (1986). The treatment in sections 3.2 to 3.4 on repeated Donation games closely follows Brandt and Sigmund (2006), and that in sections 3.5 to 3.8 on reactive strategies follows Nowak and Sigmund (1990). Adaptive dynamics was introduced in Nowak and Sigmund (1989) and Hofbauer and Sigmund (1990), and has been greatly developed since, see Dieckmann and Law (1996) or Dieckmann and Metz (2009). Molander (1985) introduced *Generous TFT*, see also Nowak and Sigmund (1992). Memory-one strategies are discussed in Lindgren (1991) and Nowak, Sigmund, and El-Sedy (1995), see Hilbe (2008) for an exhaustive treatment. The success of the *Pavlov* strategy was first noted in Kraines and Kraines (1989) and Fudenberg and Maskin (1990), its role in evolutionary dynamics was pointed out by Nowak and Sigmund (1993) and, for finite populations, by Imhof, Fudenberg, and Nowak (2007). May (1987), Boyd (1989), Bendor and Swistak (1995), and Sheratt and Roberts (2001) stress the role of errors and occasional defections in stabilizing cooperation. For strategies implemented by finite automata, see Aumann (1981), Rubinstein (1986), Abreu and Rubinstein (1988), Banks and Sundaram (1990), Binmore and Samuelson (1992), and Leimar (1997). The role of errors was stressed by Boyd (1989), see also May (1987). *Contrite TFT* was proposed by Sugden (1986), see also Boerlijst, Nowak, and Sigmund (1997a). For economic experiments, see e.g. Milinski and Wedekind (1998), Kollock (1993), and Kagel and Roth (1995). From the vast literature on the folk theorem for repeated games, we refer particularly to Selten and Hammerstein (1984) and Fudenberg and Maskin (1986). The section on equalizer strategies follows Boerlijst, Nowak, and Sigmund (1997b). The alternating Prisoner's Dilemma is studied in Nowak and Sigmund (1994), Frean (1994), and Neill (2001). For the important topic of continuous investment levels, see Roberts and Sherratt (1998), Wahl and Nowak (1999), and Killingback and Doebeli (2002). Fishman, Lotem, and Stone (2001), Johnson, Stopka, and Bell (2002), and McNamara, Barta, and Houston (2004) stress the importance of variability in behavior.