Invasion Dynamics of the Finitely Repeated Prisoner's Dilemma

Martin A. Nowak

Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, United Kingdom

AND

Karl Sigmund

Institut für Mathematik, Universität Wien, Strudlhofgasse 4, A-1090 Vienna, Austria

Computer simulations have shown that mutation–selection processes frequently lead to the establishment of cooperation in the repeated prisoner's dilemma. To simplify the mathematical analysis, it has usually been assumed that the interaction is repeated infinitely often. Here, we consider the finitely repeated case. Using renewal equations, we derive analytic results on the adaptive dynamics of monomorphic populations evolving in trait-space, describe the cooperation-rewarding zone and specify when unconditional defectors can invade. Tit for tat plays an essential, but transient, role in the evolution of cooperation. A large part of the paper considers the case when players make their moves not simultaneously, but alternatingly. *Journal of Economics Literature* Classification Numbers C70, C72. © 1995 Academic Press, Inc.

1. Introduction

Starting with Trivers' classical paper on reciprocal altruism (1971), socio-biologists have used the prisoner's dilemma (PD) game to focus on the main bugbear besetting mutual aid: the danger arising from exploiters who do not reciprocate the support of others and thereby undermine cooperation (see, e.g., Axelrod and Hamilton, 1981; Trivers, 1985; May, 1987; Dawkins, 1991; and, for a review, Axelrod and Dion, 1988). Today, the PD approach has become the dominant (but not the only) theme of the theory of recipro-

364

cal altruism; it motivated a wealth of theoretical investigations and even a handful of empirical studies on animal behavior.

The prisoner's dilemma encapsulates the problem of defection in a minimalistic setup. Two players choose either to play C (i.e., to cooperate) or to play D (i.e., to defect). If both defect, they receive a payoff P (punishment) which is smaller than the payoff R (reward), which they obtain if both cooperate. But if one player defects and the other cooperates, the defector receives a payoff T (temptation) which is larger than R, while the cooperator receives only S (the sucker's payoff), which is even smaller than P. In addition to

$$T > R > P > S \tag{1}$$

one also assumes

$$2R > T + S \tag{2}$$

so that the joint payoff for two players is larger if both cooperate than if one of them defects unilaterally. By (1), the strategy D dominates in the sense that it yields the higher payoff, no matter whether the coplayer opts for C or D. As a result, two rational players obtain only payoff P.

The dilemma can also be viewed in a population setting. Let us assume a population of players interacting with randomly chosen opponents and interpret payoff as reproductive success. In this case, D players will always have more success than C players and therefore spread in the population. Defectors will eventually take over.

The situation changes if one assumes that the prisoner's dilemma game is played repeatedly between two players. Let us assume that there exists a constant probability w for a further round. In this case, if A_n is a player's payoff in the nth round, his total expected payoff is $\sum A_n w^n$ (we note that w^n is the probability for the occurrence of the nth round and that $(1 - w)^{-1}$ is the expected number of rounds). There are enormously many strategies for the iterated PD, and Axelrod (1984) has shown that if w is sufficiently large, no strategy dominates: its success now depends crucially on the strategy of the coplayer, i.e., on the composition of the population. It is true that a player using the *AllD* strategy (always defect) will never do worse than his coplayer: but this coplayer can have cooperative interactions with *other* members of the population which make up for the losses due to the *AllD* player.

There are many ways of exploring the evolutionary dynamics of the amazingly complex world of the iterated PD. A straightforward approach would be to study its replicator dynamics: according to this dynamics, the increase (or decrease) of a strategy's frequency is given by how much better

(or worse) it does than the population average (see, for instance, Hofbauer and Sigmund, 1988). However, as soon as one admits a sizeable number of competing strategies, this dynamics becomes too complex to analyze. We are obliged, therefore, to restrict our attention to the case of two competing strategies. More precisely, we investigate when one strategy can invade another. We shall assume that the population is very large and essentially homogeneous, all individuals playing the same strategy E. A dissident individual playing strategy E' interacts only with E players in this case. The question which we shall address is then, when can this dissident spread under the influence of replicator dynamics? In Nowak (1990a,b), this question has been completely solved for reactive strategies (described in Section 2) in the case $w = 1$. In our present paper, we consider larger classes of strategies and assume $w < 1$. Because we cannot fully solve the problem, we restrict our attention to two aspects: (a) When can an E population be invaded by close-by mutants? (b) When can an E population be invaded by $AllD$?

To deal with the first question, we shall use the notion of adaptive dynamics, which singles out the optimal directional change for a deviation from the common strategy E. The rationale is that the homogeneous population is surrounded by close-by mutants and that myopic selection leads in the direction of the most prospering one. Of course this dynamics is an extremely rough caricature of an evolutionary process. Real-life populations are not homogeneous; mutants can easily miss the most promising directions; even if they do invade a population, this need not mean that they can completely take over; and the assumption that no further mutations will mess up the picture before this takeover is completed is also wildly optimistic. Nevertheless, one can view adaptive dynamics as a helpful tool for understanding evolutionary change (see Hofbauer and Sigmund, 1991; Metz et al., 1992; Rand et al., 1994; Dieckmann and Law, 1995).

If x is a parameter describing a phenotypic trait—or, in the present context, a strategy E—then x changes, according to the adaptive dynamics, by

$$\dot{x} = \frac{\partial P(E, E')}{\partial x},\tag{3}$$

where $P(E, E')$ is the payoff for an E-strategist in a population of E'-players and the right-hand side is evaluated at $E' = E$. This adaptive dynamics describes at least *locally* how phenotypic traits change in a monomorphic population. It points toward the local best response. (For a derivation of (3), see Hofbauer and Sigmund, 1991.)

Since it seems almost impossible to investigate the game in its full generality, we shall restrict our attention to conditional strategies where the propen-

sity to cooperate depends on the outcome of the previous round only. These strategies have been investigated in Nowak and Sigmund (1993a,b, 1994a,b), but only for the infinitely iterated PD, where the probability w for a further round is equal to 1. In this case, the total payoff is given by the limit in the mean of the payoffs per round, i.e., by $\lim(1/n)(A_1 + \cdots + A_n)$, where A_n, as before, is the payoff in round n. This allows for considerable computational simplifications, but can, of course, be challenged for its lack of realism.

For starters, we briefly describe the special case of reactive strategies, where a player's decision depends only on what the coplayer did in the previous round. This situation, which is fairly well understood, helps us get acquainted with the notion of adaptive dynamics and lays the groundwork for the more complex case, where a player takes also the *own* previous move into account. In the next stage, we deal with the alternating PD, where a relation between the payoff values simplifies the analysis considerably. (It will be argued that the alternating case is interesting in many instances of mutual help.) Only then do we deal with the full case of stochastic memory-one strategies. In the final discussion, we compare this with numerical simulations of the replicator dynamics (occasionally introducing new mutant strategies) and emphasize the many problems which are left open in this context.

It should be stressed that we do not discuss the problem of evolutionary stability. Roughly speaking, a strategy is an ESS (evolutionarily stable strategy) if, whenever a population adopts this strategy, no mutant can invade under the action of selection. We are considering sets of strategies, however, which are so large that to any strategy E, one can find strategies E' which do just as well against E and against E', as E does. There exists no ESS in such a case. Moreover, even if, in a restricted set of strategies, an ESS exists, it can happen that evolution leads away from it (see Nowak, 1990b).

2. REACTIVE STRATEGIES

Let us start by considering the case of the infinitely iterated PD, i.e., $w = 1$, while restricting our attention to strategies defined by the two conditional probabilities p and q of playing **C** after a **C**, resp. after a **D**, of the coplayer in the previous round.

The set of all (p, q) strategies is the unit square. The corner $(0, 0)$ corresponds to the strategy *AllD* which always defects the southeast corner $(1, 0)$ to the tit-for-tat (TFT) strategy, the center $(\frac{1}{2}, \frac{1}{2})$ to the fully random strategy, etc. We assume that all strategies are subject to some noise and hence restrict our attention to the (p, q) strategies in the interior of the

unit square. This has the advantage that the resulting game is described by a Markov chain which is ergodic: every state (i.e., every pair of moves by the two opponents) occurs with a positive probability. Instead of $(1, 0)$, for instance, we shall assume that the tit-for-tat player occasionally makes a mistake in implementing his strategy, and we consider a strategy given by $(1 - \varepsilon, \varepsilon)$, where ε is some (small) error probability.

The payoff for a (p, q)-player against a (p', q')-player is easily computed to be

$$(R - T + P - S)\sigma\sigma' + (S - P)\sigma + (T - P)\sigma' + P, \qquad (4)$$

where

$$\sigma = \frac{q - uq'}{1 - uu'} \qquad (5)$$

(with $u = q - p$ and $u' = q' - p'$) is the asymptotic probability for the (p, q) player to play C in the nth round (for $n \to \infty$), and σ' is the corresponding expression for the coplayer.

If we compute the adaptive dynamics according to (3), with p and q as parameters x, we find

$$\dot{p} = \frac{q}{1 + u} G \qquad (6)$$

$$\dot{q} = \frac{(1 - p)}{1 + u} G, \qquad (7)$$

with

$$G = \frac{1}{(1 + u)^2}\left[(R - S - T + P)q - (1 + u)\frac{(T - P)u + (P - S)}{(1 - u)}\right]. \qquad (8)$$

It follows that the propensities p and q to cooperate are either both increasing or both decreasing, depending on the sign of G. The region where this expression is positive defines the so-called *cooperation-rewarding zone*. Within this zone, any sufficiently small mutation which increases the probability σ to cooperate can invade, while any small mutation which decreases this probability cannot.

Similarly, we can also specify the so-called *defector-proof zone*, the set of all strategies E which cannot be invaded by *AllD*. This is given by the inequality
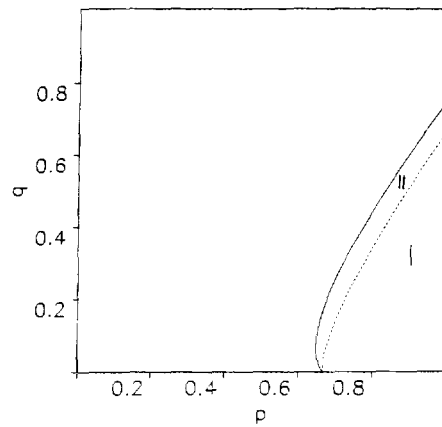
FIG. 1. The defector-proof zone (I) and the cooperation-rewarding zone (I + II) for $T = 5$, $R = 4$, $P = 2$, and $S = 0$ (see text).

$$(R - S - T + P)q - (1 + u)((T - P)u + (P - S)) > 0. \qquad (9)$$

We note that for most E, $AllD$ cannot be viewed as a small mutation. In Fig. 1, we sketch the two zones for $T = 5$, $R = 4$, $P = 2$, and $S = 0$. We see that in this case, both the cooperation-rewarding zone and the defector-proof zone are neighborhoods of tit for tat, i.e., of the southeast corner $(1, 0)$ of the space of reactive strategies. The latter zone is (for the given payoff values) a proper subset of the former. A strategy E belonging to one but not to the other zone can be invaded by $AllD$, but not by any mutant E' differing from E by only a slightly smaller value of p or q.

The vector field defined by the adaptive dynamics is, at any point (p, q), orthogonal to the line joining (p, q) to the tit-for-tat corner $(1, 0)$, suggesting that tit for tat is the pivot rather than the ultimate goal of the evolution toward cooperation.

This holds for reactive strategies and $w = 1$ (see Nowak, 1990a; Nowak and Sigmund, 1990; Molander, 1985). In Nowak and Sigmund (1990), we also discussed reactive strategies for $w < 1$. We now turn to strategies depending on the previous moves by both players, and $w < 1$. Again, we shall look for the cooperation-rewarding and the defector-proof zones. These two concepts, while far from yielding a complete understanding of the evolution, seem to yield at least a certain insight into its structure.

3. THE ALTERNATING PRISONER'S DILEMMA

In Nowak and Sigmund (1994) it has been shown that many instances of reciprocal altruism ought to be modeled by an alternating form of the PD (see also Boyd, 1988; Frean, 1994). In this setup, one assumes that in

each round, chance decides which of the two players is the leader (or donor) and which is the recipient. The leader then chooses between two options C and D. Option C yields a points to the donor and b points to the recipient, whereas option D yields c points to the donor and d points to the recipient. We shall assume that in a single round, playing C rather than D entails a cost to the donor which is smaller than the benefit that this action brings to the recipient. Since the cost is $c - a$ and the benefit $b - d$, this means that

$$0 < c - a < b - d. \tag{10}$$

Now let us consider two rounds for which the players are donors in turn. If both play C, both earn $a + b$, which we denote by R; if both play D, both earn $c + d$, which we denote by P; if one plays C and the other D while leader, the cooperator earns $a + d = S$ and the defector earns $c + b = T$. Condition (10) implies (1) and (2), i.e., the usual conditions for the payoff for the simultaneous PD. Moreover, we have

$$T + S = R + P, \tag{11}$$

which greatly simplifies Eqs. (4), (8), and (9).

It is argued in Boyd (1988), Nowak and Sigmund (1994), and Frean (1994) that mutual aid is often given alternatingly: a good turn to a partner in the hope of an eventual return. A vampire bat feeding a hungry fellow bat is obviously performing a cooperative move; this can be repaid, not simultaneously, but with a time lag. The same holds for a wolf that joins a fight to help its fellow, a bird emitting a warning call, or a monkey scratching another monkey's back. It needs no help right now; it may need it later. A similar principle operates in economic exchanges within simple social groups, for instance in the bartering of goods and services in households and among neighbors.

The outcome of one round of the repeated alternating PD is completely specified by the payoff obtained by one of the players; this can be a, b, c, or d. We denote these outcomes by 1 to 4 (in this order), noting that one player's a (or c) is the other player's b (resp. d). We restrict our attention to players whose strategy is determined by the outcome of the previous round only, i.e., given by a quadruple $\mathbf{p} = (p_1, p_2, p_3, p_4)$, where p_i denotes the propensity to play C after outcome i, and in addition by the propensity y to play C in the first round. These propensities are independent of the random decision about who is going to be the leader (we shall always assume the corresponding probability to be $\frac{1}{2}$). If a $\mathbf{p}$ player is matched against a $\mathbf{p}'$ player, then the transition probability from one round to the next is given by

$$T = \frac{1}{2} \begin{pmatrix} p_1 & p_2' & (1-p_1) & (1-p_2') \\ p_2 & p_1' & (1-p_2) & (1-p_1') \\ p_3 & p_4' & (1-p_3) & (1-p_4') \\ p_4 & p_3' & (1-p_4) & (1-p_3') \end{pmatrix}. \tag{12}$$

The total payoff of a player using strategy $E = (y, \mathbf{p})$ against a player using strategy $E' = (y', \mathbf{p}')$ is given by

$$P(E, E') = \tfrac{1}{2}[yA + (1-y)C + y'B + (1-y')D] \tag{13}$$
$$= \tfrac{1}{2}[C + D + y(A - C) + y'(B - D)],$$

where A, B, C, and D are the expected total payoffs for the E player, given that the first round of the game resulted in a, b, c resp. d. We have

$$A = a + \frac{w}{2}[p_1 A + p_2' B + (1 - p_1)C + (1 - p_2')D] \tag{14}$$

since w is the probability that the game proceeds to the second round, where the situation renews itself: $\tfrac{1}{2}$ is the probability that the E player will be leader in this round, p_1 is the probability that he will play $\mathbf{C}$ (remembering that he just experienced an a), etc. Equation (14) and the corresponding equations for B, C, and D can be written as

$$M \begin{pmatrix} A \\ B \\ C \\ D \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}, \tag{15}$$

where

$$M = Id - w\mathbf{T} \tag{16}$$

has full rank, so that we can compute the payoff. Writing out (15), we obtain

$$C + D + p_1(A - C) + p_2'(B - D) + \frac{2}{w}(a - A) = 0 \tag{17}$$

$$C + D + p_2(A - C) + p_1'(B - D) + \frac{2}{w}(b - B) = 0 \tag{18}$$

$$C + D + p_3(A - C) + p_4'(B - D) + \frac{2}{w}(c - C) = 0 \qquad (19)$$

$$C + D + p_4(A - C) + p_3'(B - D) + \frac{2}{w}(d - D) = 0. \qquad (20)$$

Subtracting (17) from (19), and setting

$$v = w(p_3 - p_1), \qquad u = w(p_4 - p_2), \qquad t = w(p_3 + p_4), \qquad (21)$$

and similarly for v', u', and t', we obtain

$$(2 + v)(A - C) + u'(B - D) = 2(a - c).$$

Similarly substracting (18) from (20), we obtain

$$u(A - C) + (2 + v')(B - D) = 2(b - d).$$

Cramer's rule yields

$$A - C = 2\frac{(a - c)(2 + v') - (b - d)u'}{(2 + v)(2 + v') - uu'} \qquad (22)$$

and

$$B - D = 2\frac{(b - d)(2 + v) - (a - c)u}{(2 + v)(2 + v') - uu'}. \qquad (23)$$

Adding (19) and (20) yields

$$C + D = \frac{1}{1 - w}[c + d + \frac{t}{2}(A - C) + \frac{t'}{2}(B - D)]. \qquad (24)$$

THEOREM 1. *The total payoff $P(E, E')$ for an E player against an E' player is given by (13), where $B - D$, $A - C$, and $C + D$ are as in (22), (23), and (24), respectively.*

For the sake of completeness, we add that in the limiting case $w = 1$, where the total payoff is defined as the limit in the mean of the payoff per round, we have

$$2P(E, E') = c + d + \frac{(a - c)[t(2 + v') - t'u] + (b - d)[t'(2 + v) - tu']}{(2 + v)(2 + v') - u'},$$

which turns out to be the same expression as (24), except for the factor $(1 - w)^{-1}$. (It is obvious that the initial probabilities y and y', which occur in (13), play no role in this case.)

We now turn to the adaptive dynamics underlying the alternating PD. In our case, the trait x in Eq. (3) can be $y, p_1, \ldots, p_4$. Rather than differentiating $P(E, E')$, which yields very cumbersome computations, we shall use the implicit function theorem. We write (15) as

$$f_i(p_1, p_2, p_3, p_4, p_1', p_2', p_3', p_4', A, B, C, D) = 0$$

(for $i = 1$ to 4) and see that

$$\frac{\partial(f_1, f_2, f_3, f_4)}{\partial(A, B, C, D)} = M$$

and that

$$\frac{\partial(f_1, f_2, f_3, f_4)}{\partial(p_1, p_2, p_3, p_4)} = \frac{w}{2} \operatorname{diag}(C - A).$$

It follows from the implicit function theorem that

$$\frac{\partial(A, B, C, D)}{\partial(p_1, p_2, p_3, p_4)} = -\left[\frac{\partial(f_1, f_2, f_3, f_4)}{\partial(A, B, C, D)}\right]^{-1} \frac{\partial(f_1, f_2, f_3, f_4)}{\partial(p_1, p_2, p_3, p_4)} \tag{25}$$

$$= -\frac{w}{2}(C - A)G,$$

where $G = M^{-1}$. We note that since

$$G = (Id - w\mathbf{T})^{-1} = Id + w\mathbf{T} + (w\mathbf{T})^2 + \cdots$$

all elements g_{ij} of G are strictly positive. The formulae for these expressions are rather unpleasant, but they combine in a decent way. For future use, we note

$$\det M = \frac{1 - w}{4}(2 + v + u)(2 + v - u) \tag{26}$$

and, as one can easily deduce from the special form of M,

$$g_{11} = g_{22}, \quad g_{12} = g_{21}, \quad g_{13} = g_{24}, \quad g_{14} = g_{23}, \qquad (27)$$

as well as

$$g_{31} = g_{42}, \quad g_{32} = g_{41}, \quad g_{33} = g_{44}, \quad g_{34} = g_{43}. \qquad (28)$$

A straightforward computation shows that

$$g_{11} + g_{12} = g_{21} + g_{22} = \frac{2w}{\det M} (2 + v - u)\left(\frac{2}{w} - 2 + p_3 + p_4\right) \qquad (29)$$

and

$$g_{13} + g_{14} = g_{23} + g_{24} = \frac{2w}{\det M} (2 + v - u)(2 - p_1 - p_2) \qquad (30)$$

and hence

$$g_{11} + g_{12} - g_{13} - g_{14} = \frac{2w}{\det M} (2 + v - u)$$
$$\times \left(\frac{2}{w} - 4 + p_1 + p_2 + p_3 + p_4\right). \qquad (31)$$

Now (13) and (4) yield

$$\dot{y} = \frac{1}{2} (A - C),$$

$$\dot{p}_1 = \frac{1}{2}\left[y\left(\frac{\partial A}{\partial p_1} + \frac{\partial B}{\partial p_1}\right) + (1 - y)\left(\frac{\partial C}{\partial p_1} + \frac{\partial D}{\partial p_1}\right)\right],$$

etc. By (25), this yields

$$\dot{p}_1 = \frac{w}{4} (A - C)[y(g_{11} + g_{12}) + (1 - y)(g_{13} + g_{14})].$$

We note that the last term on the right-hand side is always strictly positive, so that the sign of p_1, etc., is always that of $A - C$, an expression which has been computed in (22). (Since we now evaluate our expressions for $E = E'$, we have $v = v'$, etc.) Together with (29) and (31), we obtain

THEOREM 2. *The adaptive dynamics for the randomly alternating PD is given by*

$$\dot{y} = \frac{1}{2}(A - C) \tag{32}$$

$$
\dot{p}_1 = \dot{p}_2 = \frac{4w^2}{1 - w}(2 + v + u)^{-1}
$$
$$
\times \left[2 - p_1 - p_2 + y\left(\frac{2}{w} - 4 + p_1 + p_2 + p_3 + p_4\right) \right](A - C) \tag{33}
$$

$$
\dot{p}_3 = \dot{p}_4 = \frac{4w^2}{1 - w}(2 + v + u)^{-1}
$$
$$
\times \left[\frac{2}{w} - p_1 - p_2 + y\left(-\frac{2}{w} + p_1 + p_2 + p_3 + p_4\right) \right](A - C), \tag{34}
$$

where

$$A - C = 2\frac{(a - c)(2 + v) - (b - d)u}{(2 + v + u)(2 + v - u)}. \tag{35}$$

The corresponding expressions for the limiting case $w = 1$ is given by

$$\dot{p}_1 = \dot{p}_2 = \frac{t[(a - c)(2 + v) - (b - d)u]}{2(2 + v + u)^2(2 + v - u)} \tag{36}$$

and an analoguous expression for $\dot{p}_3 = \dot{p}_4$, with $t = p_3 + p_4$ replaced by $2 - p_1 - p_2$.

We note in every case (i) that $\dot{p}_1 = \dot{p}_2$ and $\dot{p}_3 = \dot{p}_4$, which implies that the optimal adaptation depends only on whether there was a **C** or a **D** in the previous round and not on who actually implemented it and (ii) that $\dot{y}$ and all the $\dot{p}_i$ are of the same sign. This sign is positive in the cooperation-rewarding zone, which is defined by

$$(a - c)(2 + v) > (b - d)u \tag{37}$$

and hence independent of y. Setting

$$\alpha = \frac{b - d}{c - a}$$

(which is just $(T - P)/(T - R)$, and always larger than 1), we see that the cooperation-rewarding zone is given by

$$\frac{2}{w} < p_1 - p_3 + \alpha(p_2 - p_4). \tag{38}$$

This zone is non-empty, therefore, if and only if it contains the tit-for-tat strategy, which is given by $(p_1, p_2, p_3, p_4) = (1, 1, 0, 0)$, i.e., if and only if

$$\alpha > \frac{2 - w}{w}. \tag{39}$$

We note that in the limiting case $w = 1$, this condition is always satisfied. Cooperation is easier to achieve the larger the zone, i.e., the smaller the "temptation" $T - R$ to defect unilaterally, compared with the gain $R - P$ obtained by mutual cooperation. In particular, if

$$\alpha > \frac{2}{w}, \tag{40}$$

then the cooperation-rewarding zone contains $(1, 1, 1, 0)$, the strategy which is always ready to cooperate except if it has been played for a sucker, i.e., if it has experienced a d in the last round. In this case, a $(1, 1, 1, 0)$ player defects if he is leader in the next round, but he defects only once; if he is leader in the subsequent round, or if the other player resumes cooperation, he will switch back to C. We note that in the limiting case $w = 1$, condition (40) simply means that the cost to the donor $c - a$ is at least twice as large as the benefit to the recipient $b - d$.

It is easy to find the best consensus strategy which is immune to defection. If all members of the population adopt this strategy, then exploiters with a lower propensity to cooperate cannot invade, and the overall payoff for the population is maximal (subject to this noninvadability condition). This payoff is given by

$$P(E, E) = \frac{\alpha - 1}{2 + v + u}\left(2y + \frac{t}{1 - w}\right) + \frac{c + d}{1 - w}.$$

THEOREM 3. *The cooperation-rewarding zone is non-empty if and only if (39) holds. If*

$$\frac{2 - w}{w} < \alpha < \frac{2}{w},$$

then the consensus strategy E with the highest payoff in this zone is given by

$$(y, p_1, p_2, p_3, p_4) = \left(1, 1, 1, 1 + \alpha - \frac{2}{w}, 0\right)$$

and if (40) holds, it is given by

$$(y, p_1, p_2, p_3, p_4) = \left(1, 1, 1, 1, 1 - \frac{2}{\alpha w}\right).$$

Next, we ask for which values of x a homogeneous population of players using $E = (1, 1, 1, 1, x)$ is defector-proof, i.e., cannot be invaded by $AllD$. The average payoff between two E players is given by $(R/2)(1 + w + \cdots)$, i.e., by $(R/2)(1 - w)^{-1}$. Let μ be the payoff for an $AllD$ player against an E player. $AllD$ players can invade if and only if

$$\mu > \frac{R}{2(1 - w)}. \tag{41}$$

We have

$$\mu = \tfrac{1}{2}(L + N),$$

where L and N are the payoff for the $AllD$ player, given that he (resp. his opponent) makes the first move. Clearly, we have

$$N = b + w\mu \tag{42}$$

since b is what the $AllD$ player receives in the first round, and then everything proceeds as from scratch. Also,

$$L = c + \frac{w}{2}[xN + (1 - x)(d + w\mu)]. \tag{43}$$

Indeed, c is the $AllD$ player's payoff in the first round, and in the next round, he and the other player have equal chances to be the leader. We used that the game proceeds as from scratch if the E-player retaliates by a **D**. By (42) and (43), we can compute μ. (41) implies the following result.

THEOREM 4. *$AllD$ can invade a $(1, 1, 1, 1, x)$ population if and only if*

$$x > 1 - \frac{2}{w}\left(\frac{T - R/w}{R - S}\right). \tag{44}$$

For $w = 1$, this reduces to

$$x > 1 - \frac{2}{\alpha}. $$

4. THE SIMULTANEOUS PD

The same methods can be used to compute the payoff and the adaptive dynamics for the simultaneous prisoner's dilemma. The outcome of each round can be described by the payoff R, S, T, or P for one of the two players. We shall again label these outcomes by 1 to 4 (in this order). We restrict our attention again to strategies which depend only on the outcome of the previous round. If we denote by p_i the probability that the player plays C, given that he experienced outcome i in the previous round, then the player's strategy E is completely specified by $\mathbf{p} = (p_1, p_2, p_3, p_4)$ and by the probability y that he plays C in the first round. The transition matrix, from one round to the next, will again be denoted by $\mathbf{T}$. It is now given by

$$\mathbf{T} = \begin{pmatrix} p_1 p_1' & p_1(1 - p_1') & (1 - p_1)p_1' & (1 - p_1)(1 - p_1') \\ p_2 p_3' & p_2(1 - p_3') & (1 - p_2)p_3' & (1 - p_2)(1 - p_3') \\ p_3 p_2' & p_3(1 - p_2') & (1 - p_3)p_2' & (1 - p_3)(1 - p_2') \\ p_4 p_4' & p_4(1 - p_4') & (1 - p_4)p_4' & (1 - p_4)(1 - p_4') \end{pmatrix}. \tag{45}$$

We can use the same renewal argument as before to compute the overall payoff. We denote by r the total payoff if the player experiences an R in the first round and define s, t, and p in a similar way. Since

$$r = R + w[p_1 p_1' r + p_1(1 - p_1')s + (1 - p_1)p_1' t + (1 - p_1)(1 - p_1')p],$$

etc., we get

$$M \begin{pmatrix} r \\ s \\ t \\ p \end{pmatrix} = \begin{pmatrix} R \\ S \\ T \\ P \end{pmatrix}, \tag{46}$$

where

$$M = Id - w\mathbf{T}.$$

The total payoff for a player using strategy $E = (y, \mathbf{p})$ against a player using strategy $E' = (y', \mathbf{p}')$ is now given by

$$P(E, E') = yy'r + y(1 - y')s + (1 - y)y't + (1 - y)(1 - y')p \quad (47)$$

and can be computed by Eq. (46). The explicit formula, however, is a good deal more complicated than in the case of the alternating PD.

Using the same methods for the adaptive dynamics, we find that

$$\dot{y} = y(r - t) + (1 - y)(s - p),$$

that $\dot{p}_1$ is a strictly positive multiple of

$$p_1(r - t) + (1 - p_1)(s - p),$$

and that similarly $\dot{p}_2, \dot{p}_3$, and $\dot{p}_4$ are strictly positive multiples of $p_3(r - t) + (1 - p_3)(s - p)$, $p_2(r - t) + (1 - p_2)(s - p)$, and $p_4(r - t) + (1 - p_4)(s - p)$, respectively. In general, these expressions will not always have the same sign, because $r - s - t + p$ will in general be distinct from 0. This holds even if (11) is valid, except for the special case that

$$p_1 + p_4 = p_2 + p_3. \quad (48)$$

This condition is satisfied if the strategies are self-determined, i.e., depend only on what the player himself did in the previous round (meaning that $p_1 = p_2$ and $p_3 = p_4$) and in the case where strategies are reactive, and hence depend only on what the opponent did in the previous round (meaning that $p_1 = p_3$ and $p_2 = p_4$; discussed in Section 2).

In the general case, the cooperation-rewarding zone (i.e., that region in the parameter space $(y, p_1, \ldots, p_4)$, where $\dot{y}, \dot{p}_1, \ldots, \dot{p}_4$ are all positive) is difficult to describe. But we can easily compute the signs of $r - t$ and $s - p$ for those *deterministic* strategies where all p_i are equal to 0 or 1. These strategies are characterized by transition rules defined by quadruples $(p_1, \ldots, p_4)$ of zeros and ones. There are 16 such quadruples and hence 16 transition rules which we denote by S_i, for $i = 0, 1, \ldots, 15$, in such a way that $i = p_4 + 2p_3 + 2^2 p_2 + 2^3 p_1$.

Let us normalize the payoff values such that $R - P = 1$ (this is no loss of generality) and set $\beta = P - S$ and $\gamma = T - R$ (both β and γ are positive: they are equal if (11) holds). We find that $r - t$ is always negative except

for S_2, S_3, S_8, S_9, S_{10}, and S_{11}, when the sign is that of $-\gamma + \beta w - w^2$, $-\gamma + \beta w$, $-\gamma + w(1 + \gamma)$, $-\gamma + w$, $-\gamma + w(1 + \beta)$, and $-\gamma + w(1 + \beta)$, respectively. Similarly, the sign of $s - p$ is always negative, except for S_2, S_3, S_6, S_{10}, S_{11}, and S_{14}, when the sign is that of $-\beta + w(1 + \gamma)$, $-\beta + w\gamma$, $-\beta + w$, $-\beta + w(1 + \gamma)$, $-\beta + w\gamma - w^2$, and $-\beta + w(1 + \beta)$, respectively. From this the next result follows easily.

THEOREM 5. *The only deterministic strategies which can lie in the cooper-ation-rewarding zone are those with the tit-for-tat transition rule S_{10}. They belong to the zone if and only if*

$$w > \max\left[\frac{T - R}{R - S}, \frac{P - S}{T - P}\right]. \tag{49}$$

This emphasizes once more the role of tit for tat in twisting evolution toward cooperation. We conjecture that the zone of cooperation is non-empty if and only if (49) holds, i.e., that it has to contain tit for tat if it contains any strategy at all.

Extensive computer simulations show that populations evolve very frequently to a cooperative strategy with $y = 1$ and $\mathbf{p} = (1, 0, 0, x_0)$ for a certain value x_0 which depends on the payoff values R, S, T, P and on w (for $w = 1$, we refer to Nowak and Sigmund, 1993b; for $w < 1$ see Section 5 below). To better understand this result, we note, first of all, that the transition rule $(1, 0, 0, 1)$ means that the player cooperates if and only if he experienced an R or a P in the previous round, i.e., if he opted for the same move as his coplayer. To put it another way, he repeats his previous move if and only if he experienced an R or a T and consequently had reason to be satisfied with his payoff. (For this reason, a strategy with $\mathbf{p} = (1, 0, 0, 1)$ has been termed Pavlov by Kraines and Kraines, 1988.) In contrast to tit for tat, this strategy (a) is error-correcting against itself and (b) exploits unconditional cooperators. Indeed, we see that (a), if, in a match between two Pavlov-players, one commits a mistake and defects, then both players defect in the next round but subsequently resume mutual cooperation; and (b) if a Pavlov-player notices that an inadvertent D against his coplayer meets no retaliation, he will continue to defect. In a Pavlov-dominated society, suckers cannot spread. However, in a society dominated by players who always defect, Pavlov fares poorly, as it tries every second time to resume cooperation.

A strategy with transition rule $(1, 0, 0, x)$ has—like Pavlov—the property to resume cooperation, against its like, after an erroneous D, and to exploit players who do not retaliate, but it is less naive than Pavlov against all-out defectors. Let us compute the payoff μ for a player with strategy $AllD$ (i.e., always defecting) against a player who starts with a C move and then uses
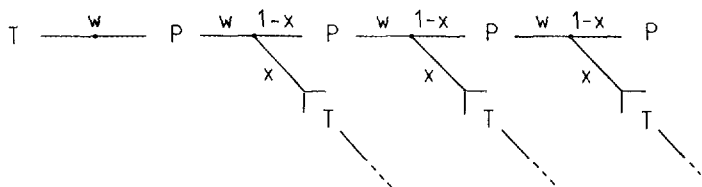
Fig. 2. The event tree for the game between two $(1, 0, 0, x)$ players (see text).

$(1, 0, 0, x)$. In the first round, the *AllD* player receives a T and in the second round (which occurs with probability w), he receives a P. In the third round (which occurs with probability w^2), there are two alternatives: (i) with probability x, the other player cooperates: in this case, it is as if the game started anew, and the future payoff, for the *AllD* player, is again μ; (ii) with probability $(1 - x)$, the other player continues to defect, so that the *AllD* player receives P as his payoff in this round and finds himself in exactly the same situation as after the previous round (see Fig. 2). This means that if a further round occurs (for which the probability is w), he faces again the two alternatives (i) and (ii). Iterating this argument, we see that

$$\mu = T + wP + w^2[(1 - x)P + x\mu] + w^3(1 - x)[(1 - x)P + x\mu] + \cdots$$
$$= T + (wP + w^2x\mu)[1 + w(1 - x) + w^2(1 - x)^2 + \cdots]$$

and hence that

$$\mu = T + w\frac{P + \mu wx}{1 - w(1 - x)}.$$

It follows that

$$\mu\left(1 - \frac{w^2x}{1 - w(1 - x)}\right) = T + \frac{wP}{1 - w(1 - x)}$$

so that

$$\mu = \frac{T - w(1 - x)T + wP}{(1 - w)(1 + wx)}. \tag{50}$$

The corresponding expression for $w = 1$ reduces to

$$\mu = \frac{P + xT}{1 + x}.$$

Returning to $w < 1$, we see that in a population of players all using $y = 1$ and $\mathbf{p} = (1, 0, 0, x)$, the average payoff is $(1 - w)^{-1} R$. $AllD$ players can invade if and only if

$$\mu > \frac{R}{1 - w}. \tag{51}$$

This implies the following result.

THEOREM 6.   *In a population of players all using* $y = 1$ *and* $\mathbf{p} = (1, 0, 0, x)$, $AllD$ *players can invade if and only if*

$$x > \alpha - \frac{1}{w} := x_0. \tag{52}$$

Here we used again $\alpha = (T - P)/(T - R)$. In particular, $AllD$ can invade Pavlov (where $x = 1$) if and only if

$$w < \frac{T - R}{R - P}. \tag{53}$$

Thus if $2R > T + P$, Pavlov cannot be invaded by any strategy. However, this does not mean that Pavlov is an evolutionarily stable strategy in the sense of Maynard Smith (1982). As stressed (in the context of the tit-for-tat strategy) in Selten and Hammerstein (1984), there are many strategies which do just as well, against themselves and against Pavlov, and which can spread through neutral drift. This is not just a mathematical detail: the spread of some such strategies can render the population vulnerable to subsequent invasions by defectors (see also Young and Foster, 1991).

One can perform a computation similar to the previous one with the generous TFT strategy, which has transition rule $(1, x, 1, x)$ and starts with a C. This rule forgives a coplayer's D with probability x. Against its like, its payoff is again $R(1 - w)^{-1}$. The payoff μ of an $AllD$ player against generous TFT is

$$\mu = T + w[x\mu + (1 - x)P] + w^2[x\mu + (1 - x)P] + \cdots$$
$$= T + w\frac{xT + (1 - x)P}{1 - w}. \tag{54}$$

From this it follows easily that *AllD* can invade a generous TFT population if and only if

$$x > 1 - \frac{1}{\alpha w}. \tag{55}$$

A simple exercise shows that in a population of generous TFT players with rule $(1, x, 1, x)$, a one-sided defection is overcome after (on average) $1/x - 1$ rounds: after that both players resume cooperation. Similarly, such a defection is overcome after (on average) $2/x - 1$ rounds in the case of a population with Pavlovian transition rule $(1, 0, 0, x)$.

5. DISCUSSION

We have stressed in the Introduction that the adaptive dynamics can be expected only to give a rough idea of the *local* dynamics in the parameter space. We do not believe that an analysis of its asymptotic behavior would be very fruitful. If the population in the cooperation-rewarding zone is close to the boundary, for instance, then a mutation could lead to the establishment of a population which is outside of this zone. It could also lead to a mixed population to which the adaptive dynamics (which relies on a monomorphic population) no longer applies.

For a more realistic picture, we have to use computer simulations describing the evolution of the frequencies of different strategies under a mutation–selection dynamics. Selection alone is described by the replicator dynamics (see, e.g., Hofbauer and Sigmund, 1988). If we assume that E_i (for $i = 1, \ldots , n$) are the different strategies in the population and $x_i(t)$ their frequencies at time t, then this dynamics is given by

$$\dot{x}_i = x_i(f_i - \bar{f}), \tag{56}$$

where f_i is the expected payoff for an E_i-player (which depends on the frequencies $x_1, \ldots , x_n$ of the different strategies in the population), whereas $\bar{f} = \Sigma \ x_i f_i$ is the average payoff in the population. As mentioned already, even with as few as four or five strategies present, this dynamics can be extremely complicated, exhibiting all kinds of regular or irregular oscillations. We have studied computer simulations with hundreds of strategies present: it seems hopeless to expect analytical results in this context, given the present state of knowledge. For a realistic picture of the evolution, one also has to add a mutation term introducing, from time to time, small amounts of new strategies. Such dynamics were described in Nowak and

Sigmund (1993b, 1994a) for the case $w = 1$. The corresponding simulations for $w < 1$ lead to analogous results. The main conclusions are:

(a) The average payoff in the population is usually close to the values R or P; the shifts from one regime to the other are rare and sharp and somewhat analogous to phase transitions. The later in the run, the longer the periods of stasis, and the more likely is a regime of cooperation.

(b) If the average payoff is close to R, the population is dominated by a few strategies sharing two properties: (a) while a **D** can occasionally occur in a game between two players using such a strategy, it is quickly overcome and mutual cooperation is resumed; and (b) they relentlessly keep playing **D** if the coplayer does not retaliate.

Figure 3, for the case of the alternating PD, is a typical run leading to cooperation. The outcome is close to the consensus strategy defined in Section 3. Figures 4 and 5 apply to the case of the simultaneous PD. More precisely, in these simulations, we use a discrete game dynamics: if x_i is the frequency of strategy E_i in one generation, its frequency in the next generation is given by $x_i' = (\bar{f})^{-1}x_i f_i$. Every hundredth generation, we introduce a new randomly chosen strategy (y, p_1, p_2, p_3, p_4) with a small initial frequency. Each p_i is chosen according to the U-shaped probability density $[\pi x(1 - x)]^{-1/2}$ from the unit interval $[0, 1]$ (or more precisely from the interval $[0.001, 0.999]$, because we assume a minimal amount of noise). Strategies with frequencies below 0.001 are removed (and the other frequencies re-adjusted to sum up to 1). Usually, there are 1 to 10 strategies in the population. We extend each simulation to cover 10^7 generations, thereby sampling 10^5 strategies in each run, giving special attention to those near to the corners (this is why we chose the U-shaped density).

We see in Fig. 4 a typical run: after some time, the parameters y and p_1 are close to 1, the parameters p_2 and p_3 are close to 0, and p_4 slowly edges closer to the value x_0 given by (52). We see from Fig. 5 that Pavlov is the most frequent outcome as soon as $R > 3$, which again agrees well with (52).

How robust are these simulations? If one changes the starting point of the simulation, the arrival interval of mutations, the payoff values, the average length of the game, or the frequencies below which the strategies are considered to be "extinct" and removed from the population, or if one varies the minimal amount of noise or the distribution of the mutants, nothing much changes in the overall picture. There is a single parameter which is critical: this is the initial frequency of the mutant. If this is too low, there will be no emergence of a cooperative regime. This emphasizes a result well known since the pioneering work of Axelrod and Hamilton: there has to be a minimal clustering of retaliators.

The adaptive dynamics emphasizes the role of the cooperation-rewarding zone, where evolution leads toward strategies more and more ready to
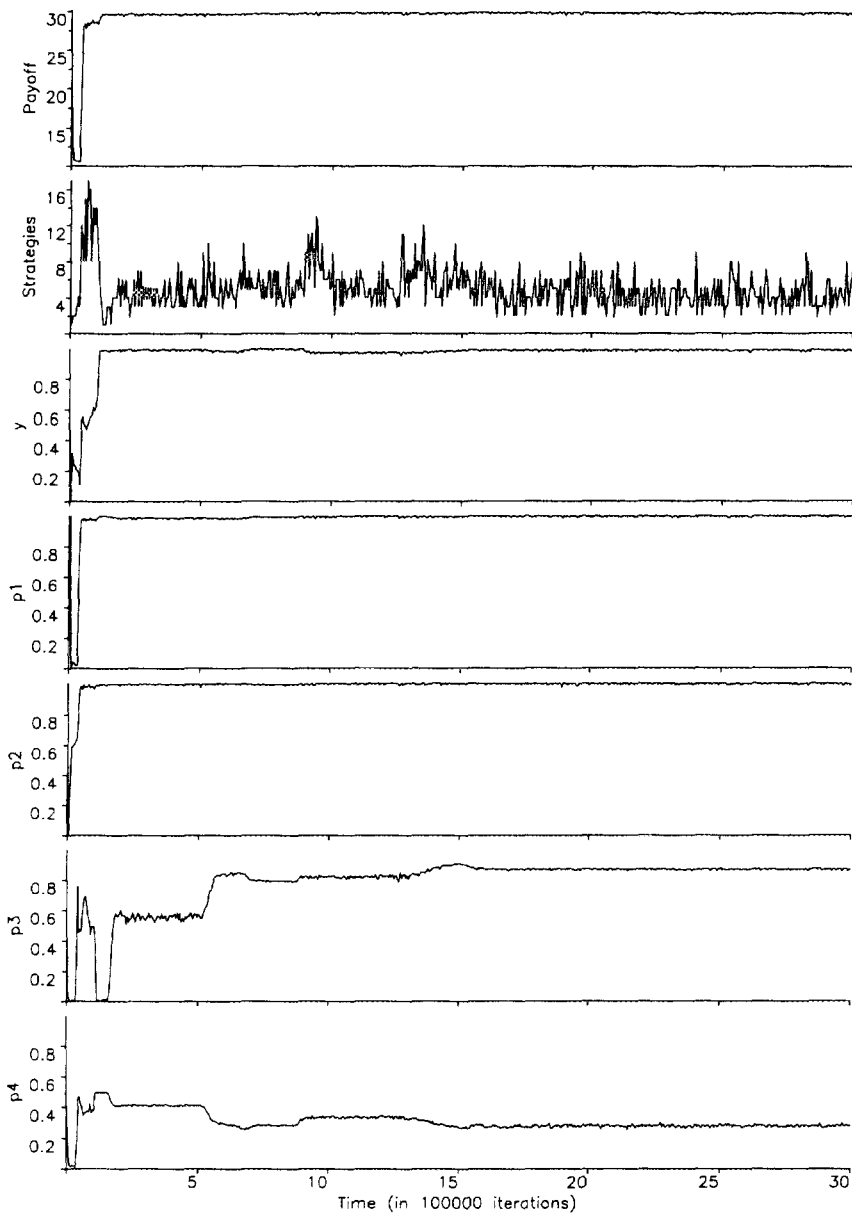
FIG. 3. Evolutionary simulation of the stochastically alternating, finitely repeated prisoner's dilemma with stochastic strategies that only consider the previous round of the iterated game. Each strategy is given by five probabilities (y, p_1, p_2, p_3, p_4). The probability to cooperate in the first move is denoted by y. The probabilities to cooperate after one's own C, the

cooperate. The main question in this context is, how can a population of defectors be coaxed into this zone? In the case of reactive strategies, the answer is clear (see Nowak and Sigmund 1992): one needs a cluster of tit-for-tat players. Tit for tat is, in this case, almost specified by the property that it can invade an *AllD*-population in minimal cluster size. In the more general case, a careful look at the numerical simulations shows that again a very stern, highly provokable strategy has to invade. However, this need not necessarily be tit for tat: in the case of the simultaneous PD, it is more frequently a strategy close to grim S_R, which starts with a **C** and then applies the transition rule $(1, 0, 0, 0)$. After a defection, grim never cooperates again. We note that grim is very vulnerable to errors, whereas Pavlov is not: strategies with transition rule $(1, 0, 0, x)$ lie in between: the larger the x, the quicker such a defection will be overcome, but the lower the payoff against *AllD*-players.

In contrast to the case $w = 1$, the initial move of a strategy is now of importance. For a non-forgiving rules like TFT or grim, it is essential that this first move is a **C**, since otherwise the whole future will be poisoned. For a rule like Pavlov, this need no longer be the case. On the contrary, it can be advantageous now to start with a **D**.

In papers by Rubinstein (1986), Abreu and Rubinstein (1988), Binmore

---

opponent's C, one's own D, and the opponent's D are given by p_1, p_2, p_3, and p_4, respectively. The population averages of these quantities, together with the total number of strategies in the population and the average payoff of the population, are shown. The simulation was performed with the usual game dynamical difference equation $x'_i = x_i f_i / f$, where x_i denotes the frequency of strategy i, f_i its payoff in the population, and $\bar{f} = \Sigma\, x_i f_i$ the average payoff of the population. Thus strategies spread according to their success. Payoff equals fitness. We start with a homogeneous population using strategy $(0.5, 0.5, 0.5, 0.5, 0.5)$. Every 100 iterations (on average) a new mutant is introduced. New mutants are chosen at random; the probabilities y and p_i are taken from the U-shaped distribution $[\pi x(1 - x)]^{-1/2}$ to get more bias toward the boundaries of the five-dimensional strategy space (because relevant strategies are usually close to the boundary). We only admit probabilities between 0.001 and 0.999, thereby including a minimum amount of intrinsic stochasticity. If the frequency of a strategy drops below 0.001 it is removed from the population. New strategies are introduced at an initial frequency of 0.01. It is important that this initial frequency is not too low, as otherwise *AllD* populations could never be invaded by *TFT*-like strategies and cooperation would never emerge. This simulation shows a rapid emergence of cooperation. There are ensembles of strategies with y, p_1, and p_2 all close to one. The two other probabilities, p_3 and p_4, show more fluctuations. Interestingly at the end a strategy emerges which is given approximately by the consensus strategy defined in the text. We performed 40 such simulations. Cooperation was reached in 15 runs and was always based on strategies around $y = p_1 = p_2 = 1$, p_3 varying between 0.5 and 1, and p_4 varying between 0 and 0.4. Payoff values: $a = 2$, $b = 1$, $c = 3$, $d = -2$, corresponding to $R = 3$, $S = 0$, $T = 4$, and $P = 1$. Discount of the future $w = 0.95$, implying an average length of the repeated prisoner's dilemma of 20 rounds.
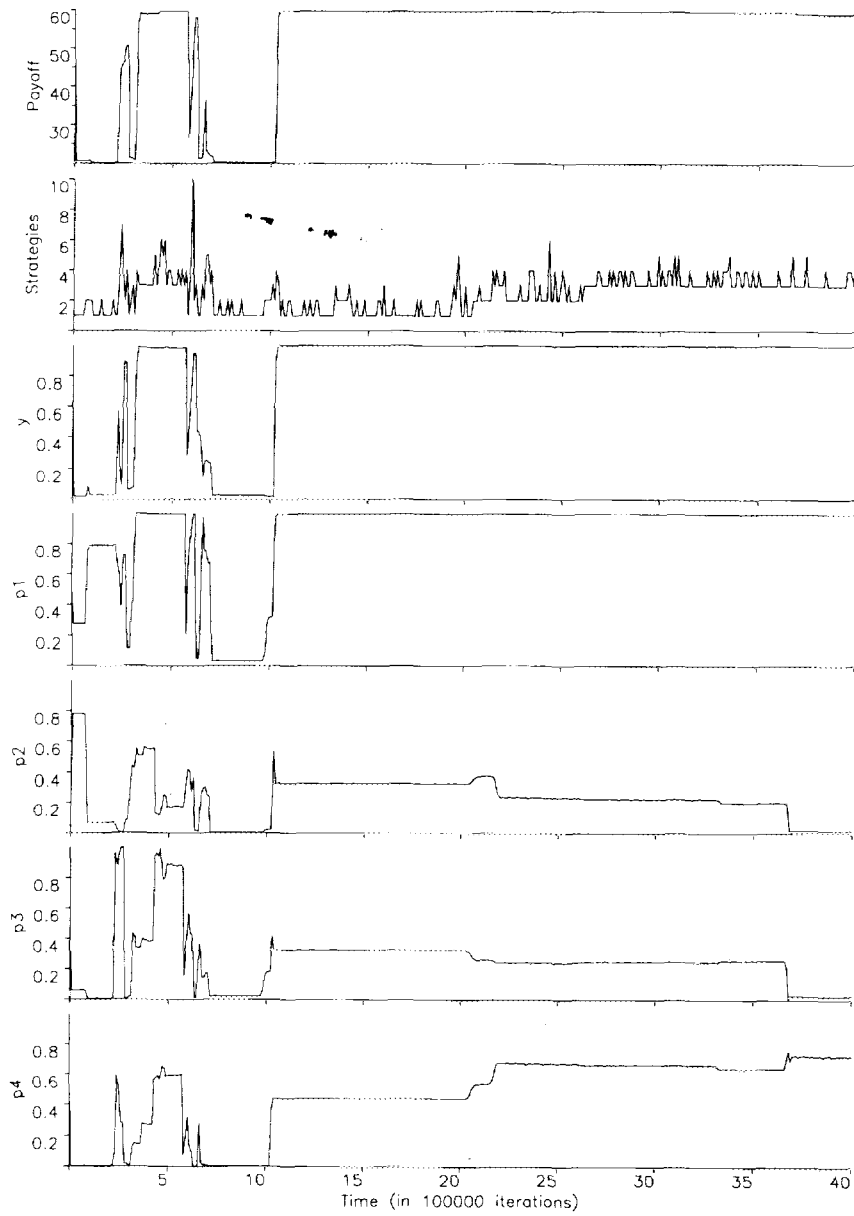
FIG. 4. An evolutionary run as before, but this time for the simultaneous PD. The strategies are given the five probabilities: the probability y to cooperate in the first move, and four probabilities (p_1, p_3, p_3, p_4) to cooperate after having received payoff R, S, T, or P. We used the population values $R = 3$, $S = 0$, $T = 5$, and $P = 1$.
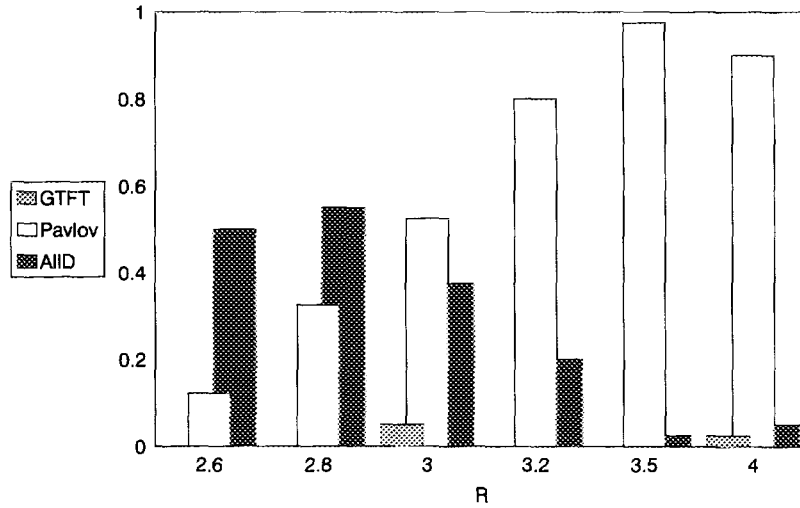
FIG. 5. Statistics for evolutionary runs for the simultaneous PD. We used the parameter values $S = 0$, $P = 1$, $T = 5$, and R as indicated. For each value of R we performed 40 *independent simulations as described in the legend to Fig. 4, each generating a total of* 10^5 different mutant strategies. The fractions of runs dominated by generous tit for tat, Pavlov, and *AllD* after 10^7 iterations of the game dynamical difference equation are shown.

and Samuelson (1992), Papadimitriou (1992), and Probst (1993) and Schlag (1994), we can find investigations of the infinitely iterated PD for deterministic strategies based on finite automata (see also Binmore, 1992). Lindgren (1991) investigates mutation–selection dynamics of such strategies, adding small error terms. He considers strategies with longer memories, whereas we allow for stochastic strategies which are not necessarily close to deterministic strategies. Other papers dealing with the effect of stochastic perturbations on repeated games are by Farrell and Ware (1988) and Fudenberg and Maskin (1990), who deal with the ESS approach, as well as Foster and Young (1990), Kandori et al. (1991), and Young (1993), who use another approach based on adaptive play (not to be confused with adaptive dynamics).

Here we have investigated the finitely repeated PD ($w < 1$), allowing for stochastic strategies. Using renewal arguments, we computed payoffs and dealt with questions of invadability. Among the problems left open, we mention (a) a characterization of those strategies best able to invade populations of defectors and to lead the population into the cooperation-rewarding zone and (b) the extension of this work to strategies with longer memories.

REFERENCES

ABREU, D., AND RUBINSTEIN, A. (1988). "The Structure of Nash Equilibrium in Repeated Games with Finite Automata," *Econometrica* **56**, 1259–1282.

AXELROD, R. (1984). The Evolution of Cooperation. New York: Basic Books; reprinted, 1989. Harmondsworth: Penguin.

AXELROD, R., AND HAMILTON, W. D. (1981). "The Evolution of Cooperation," *Science* **211** 1390–1396.

AXELROD, R., AND DION, D. (1988). "The Further Evolution of Cooperation," *Science* **242**, 1385–1390.

BINMORE, K. G. (1992). *Fun and Games: A text on Game Theory*. Lexington, MA: Heath.

BINMORE, K. G., AND SAMUELSON, L. (1992). "Evolutionary Stability in Repeated Games Played by Finite Automata," *J. Econ. Theory* **57**, 278–305.

BOYD, R. (1988), "Is the Repeated Prisoner's Dilemma a Good Model of Reciprocal Altruism?" *Ethology and Sociobiology* **9**, 211–222.

DAWKINS, R. (1991). *The Selfish Gene*, 2nd ed. London: Oxford Univ. Press.

DIECKMANN, U., AND LAW, R. (1995). "The Dynamical Theory of Coevolution: A Derivation from Stochastic Ecological Processes, *J. Math. Biol.*, to appear.

FARRELL, J., AND WARE, R. (1988). "Evolutionary Stability in the Repeated Prisoner's Dilemma Game," *Theor. Pop. Biol.* **36**, 161–166.

FOSTER, D., AND YOUNG, P. (1990). "Stochastic Evolutionary Game Dynamics," *Theor. Pop. Biol.* **38**, 219–232.

FREAN, M. (1994). "The Prisoner's Dilemma without Synchrony," *Proc. Royal Soc. Series B* **257**, 75–79.

FUDENBERG, G. AND MASKIN, E. (1990). "Evolution and Cooperation in Noisy Repeated Games," *Amer. Econ. Rev.* **80**, 274–279.

HOFBAUER, J., AND SIGMUND, K. (1988). *The Theory of Evolution and Dynamical Systems*. London: Cambridge Univ. Press.

HOFBAUER, J., AND SIGMUND, K. (1991). "Adaptive Dynamics and Evolutionary Stability," *Appl. Math. Lett.* **3**, 75–79.

KANDORI, M., MAILATH, G., AND ROB, R., (1991). "Learning, Mutation and Long Run Equilibria in Games," *Econometrica* **61**, 29–56.

KRAINES, D., AND KRAINES, V. (1988). "Pavlov and the Prisoner's Dilemma," *Theory and Decision* **26**, 47–79.

LINDGREN, K. (1991). "Evolutionary Phenomena in Simple Dynamics," in *Artificial Life, II* (C. G. Langton *et al.*, Ed.), pp. 295–312, Santa Fe Institute for Studies in the Sciences of Complexity, Vol. X, Redwood City, CA: Addison-Wesley.

MAY, R. M. (1987). "More Evolution of Cooperation," *Nature* **327**, 15–17.

MAYNARD SMITH, J. (1982). *Evolution and the Theory of Games*, London: Cambridge Univ. Press.

METZ, J. A. J., NISBET, R. M., AND GERITZ, S. A. H. (1992). "How Should We Define Fitness for General Ecological Scenarios?" *TREE* **7**, 198–202.

MOLANDER, P. (1985). "The Optimal Level of Generosity in a Selfish, Uncertain Environment," *J. Conflict Resolution* **29**, 611–618.

NOWAK, M. A. (1990a). "Stochastic Strategies in the Prisoner's Dilemma," *Theor. Pop. Biol.* **38**, 93–112.

Nowak, M. A. (1990b). "An Evolutionarily Stable Strategy May Be Inaccessible," *J. Theor. Biol.* **142,** 237–241.

Nowak, M. A., and Sigmund, K. (1990). "The Evolution of Reactive Strategies in Iterated Games," *Acta Appl. Math.* **20,** 247–265.

Nowak, M. A., and Sigmund, K. (1992). "Tit for Tat in Heterogeneous Populations," *Nature* **355,** 250–252.

Nowak, M. A., and Sigmund, K. (1993a). "Chaos and the Evolution of Cooperation," *Proc. Nat. Acad. Science USA* **90,** 5091–5094.

Nowak, M. A., and Sigmund, K. (1993b). "Win-Stay, Lose-Shift Outperforms Tit-for-Tat," *Nature* **364,** 56–58.

Nowak, M. A., and Sigmund, K. (1994). "The Alternating Prisoner's Dilemma," *J. Theor. Biol.* **168,** 219–226.

Nowak, M. A., and Sigmund, K. (1995). "Automata, Repeated Games, and Noise," *J. Math. Biol.* **33,** 703–722.

Papadimitriou, C. (1992). "On Players with a Bounded Number of States," *Games Econ. Behav.* **4,** 122–131.

Probst, D. A. (1993). "Evolution, Automata and the Repeated Prisoner's Dilemma," Discussion paper, Wirtschaftstheoretische Abteilung III, University of Bonn.

Rand, D., Wilson, H. B., and McGlade, J. M. (1994). "Dynamics and Evolution: Evolutionarily Stable Attractors, Invasion Exponents and Phenotypic Dynamics," preprint, Warwick.

Rubinstein, A. (1986). "Finite Automata Play the Repeated Prisoner's Dilemma," *J. Econ. Theory* **39,** 83–96.

Schlag, K. (1994). "Dynamic Stability in the Repeated Prisoner's Dilemma Played by Finite Automata, "Discussion paper, Wirtschaftstheoretische Abteiling III, University of Bonn.

Selten, R., and Hammerstein, P. (1984). "Gaps in Harley's Argument on Evolutionarily Stable Learning Rules and in the Logic of Tit for Tat," *Behav. Brains Sci.* **7,** 115, 116.

Trivers, R. (1971). "The Evolution of Reciprocal Altruism," *Quart. Rev. Biol.* **46,** 35–57.

Trivers, R. (1985). *Social Evolution,* Menlo Park, CA: Cummings.

Young, H. P., and Foster, D. (1991). "Cooperation in the Short and in the Long Run," *Games Econ. Behav.* **3,** 145–156.

Young, H. P. (1993). "The Evolution of Conventions," *Econometrica* **61,** 57–84.