

# Evolutionary multistability in the iterated prisoner's dilemma

Matteo Smerlak\*

Max Planck Institute for Mathematics in the Sciences,  
Inselstrasse 22, Leipzig, D-04103, Germany

## Abstract

The iterated prisoner's dilemma (IPD) is a paradigm of social dilemmas and a testing bed for ideas purporting to explain the evolution of cooperation among self-interested agents. Conventional wisdom suggests that, while selfish (or even extortionary) strategies can give an edge in head-to-head competitions, cooperative behavior tends to be more evolutionarily stable. Here I show that the memory-one IPD—where phenotypes encode the stochastic response to just four possible joint plays—contains a phase where evolutionary dynamics is in fact multistable, with long periods of cooperation followed by rapid reversions to defection, and vice versa *ad infinitum*. This surprising switch-like behavior in a strongly selected population shows that evolutionary dynamics does not always lead to evolutionarily stable states, even when phenotype space is small and easily explored.

## Introduction

Evolution through natural selection is commonly viewed as a directed process maximizing adaptation and reproductive success. Thus, in Wright's metaphor of the fitness landscape, evolving populations are pictured as climbing fitness peaks. While this view of evolution as optimization does not imply that all evolutionary change must be adaptive (some phenotypes appear to be byproducts of the evolution of other traits without being themselves adaptive), it does rely on the notion that directness and irreversibility are fundamental components of evolutionary change.

Several well-known factors complicate this naive picture. For starters, genetic drift weakens the link between selective advantage and change in allele frequencies. In many (or most) cases, mutations with no fitness effect fix by chance, i.e. populations drift

---

\*To whom correspondence should be addressed; E-mail: smerlak@mis.mpg.de

across neutral genotype networks; in the most extreme case, deleterious mutations can even replace fitter variants, as in Muller’s ratchet. Second, adaptation is not always gradual or predictable: it is well known that many species have undergone long periods of evolutionary stasis followed by short bursts of rapid adaptation; as emphasized by Gould, the outcome (and not just the timing) of these bursts is mostly contingent—“any replay of [life’s] tape would lead evolution down a pathway radically different from the road actually taken”. The extent to which these contingencies are extrinsic (e.g. meteorites) or intrinsic (e.g. self-organized) has been a topic of lively discussion.

Third, and perhaps most importantly, the widespread occurrence of frequency-dependent selection—the fact that reproductive success can depend on traits other than one’s own—significantly muddles the picture of evolution as optimization, because a phenotype capable of resisting invasion by any possible mutant may not exist. In particular, the replicator equation of evolutionary game theory does not necessarily lead to stable evolutionary attractors; simple examples, such as the rock-paper-scissors game, already give rise to periodic cycles. Nevertheless, evolutionary game theory has mostly focused on defining, identifying and characterizing evolutionarily stable strategies, ideally in terms of static payoff comparisons. But is this focus on ESS as the endpoint of evolutionary trajectories always justified?

To address some of these towering questions, it can be useful to focus on simple systems with clearly interpretable phenotypes and based on tractable mathematics. The memory-one iterated prisoner’s dilemma (IPD<sub>1</sub>) is one such system. In this game, two players X and Y cooperate (C) or defect (D) with a probability that depends on the outcome of their previous interaction. If X and Y cooperate, they both earn a reward ( $R = 3$ ); if X cooperates and Y defects, then X is a ‘sucker’ ( $S = 0$ ) and Y has fallen to temptation ( $T = 5$ ); and if X and Y both defect, they get the same punishment  $P = 1$ . This simple rule gives rise to a surprisingly rich four-dimensional phenotype space  $\Sigma_1 = [0, 1]^4$  (where the first component is the probability to cooperate after CC, the second component is the probability to cooperate after CD, etc.), comprising classical strategies such as Tit-For-Tat (TFT) and Win-Stay-Lose-Shift (WSLS). Remarkably,  $\Sigma_1$  also contains extortionary strategies which take advantage of any opponent—no matter which strategy they choose—, as well as generous strategies which guarantee that one’s opponent never does worse than oneself.

Identifying which kind of strategy is evolutionarily stable in IPD<sub>1</sub> is not just an interesting instance of the issue of directionality in evolution: it is also an important problem in itself whose solution may shed on the evolution of cooperation in biology and elsewhere. Based on the results of Axelrod’s famous tournament, it was long believed that, being simple, nice and yet retaliatory, TFT epitomizes a stable strategy in the iterated prisoner’s dilemma, and *a fortiori* in IPD<sub>1</sub>; indeed, ecological simulations run by Axelrod suggested it could rise to fixation in a field of diverse and sophisticated opponent strategies. The breakthrough discovery of so-called “zero-determinant” extortionary strategies within  $\Sigma_1$  by Press and Dyson [**press’iterated’2012**] forced a reappraisal of this conclusion [**ball’physicists’2012**]. Perhaps reassuringly, several studies found that more generous strategies can in fact replace extortionists in

large populations [adami'evolutionary'2013, stewart'extortion'2013]. Whether this scenario does indeed pan out, or defection takes root permanently, turns out depend on the mutation rate and the strength of selection [stewart'extortion'2013, iliopoulos'critical'2010]; in particular, Iliopoulos *et al.* identified a phase transition at a critical mutation rate  $\mu_c$  between two phases, one dominated by cooperation ( $\mu \leq \mu_c$ , such that opponents are predictable enough to warrant the risk of cooperation), and the other by defection ( $\mu \geq \mu_c$ , in which opponents are too unpredictable).

Here I show that which strategy is more evolutionary stable in the IPD<sub>1</sub>—which direction evolution chooses to go—is not always a well-posed question. I describe a phase of IPD<sub>1</sub>, characterized by small mutational effects, weak genetic drift and no spatial structure, in which evolutionary dynamics is *multistable* rather than convergent. This means that, instead of moving towards single attractor within  $\Sigma_1$ , evolutionary trajectories shift unpredictably between two long-lived but nevertheless transient states. The resulting dynamics resembles the burst (or ‘popcorn’) noise observed in certain electronic devices, and can be modelled mathematically by a stationary, continuous-time Markov process over few effective states with exponentially distributed sojourn times. This behavior cannot be explained with the standard tools of evolutionary game theory (adaptive dynamics or replicator-mutator equations).

## Results

**The phase diagram of evolution in the IPD.** I performed evolutionary agent-based simulations similar to those described in [iliopoulos'critical'2010]. Each agent embodies a stochastic strategy  $p = (p_1, p_2, p_3, p_4) \in [0, 1]^4 \equiv \Sigma_1$ , where  $p_1$  (resp.  $p_2, p_3, p_4$ ) is the probability to cooperate with the opponent after the joint play CC (resp. CD, DC, DD). At each generation, each agent plays a match of the infinitely iterated prisoner’s dilemma with  $T$  randomly chosen opponent in a population of size  $N$ . After each match, the payoffs of a player playing strategy  $p$  against strategy  $q$  is increased by an amount  $\pi(p, q)$  which can be computed as a  $4 \times 4$  determinant [press'iterated'2012], and similarly for the opponent. Once all agents have played all their matches, all strategies are mutated according to  $p'_i = \min(\max(p_i + \sigma\epsilon, 0), 1)$ , where  $\sigma > 0$  and  $\epsilon$  is a standard normal variable. After that, a new population of identical size  $N$  is sampled, with a weight proportional to the payoff accrued in the previous generation, and the entire process is repeated. In this formulation,  $\sigma$  measures the typical size of mutations and  $T$  the strength of selection.

Iliopoulos *et al.* describe the evolution of strategies in IPD<sub>1</sub> that are either cooperative or defective, depending on evolutionary parameters such as the mutation rate. I confirmed the existence of such a phase transition, measured here in terms of the rate of cooperation (the equilibrium probability of CC or CD plays)  $\langle C \rangle_t$  averaged over players, opponents and evolutionary time as a function of  $\sigma$  and  $T$  (Fig. XXX). As noted in Ref. [iliopoulos'critical'2010], the transition from a cooperative phase under weak mutations to a defective phase under strong mutations has a natural interpretation:

when future opponents can be reliably assumed to be similar to previous opponents, i.e. when the environment of players is sufficiently predictable, cooperation provides higher payoffs. But when mutational effects are large, there is no knowing how the next generation of opponents will play, and cooperation becomes unacceptably risky. In this phase evolution favors defection, at the cost of lower fitness for all.

Considering the variance of the mean population cooperation rate  $\text{Var}_t(C)$  over time as a function of  $\mu$  and  $T$  reveals a richer picture.

**Multistability and reduced effective dynamics.** How can an evolutionary state be stable for long stretches of time, and yet eventually yield to a phenotypically very different invader population spawn by small mutations only? Standard tools of evolutionary game theory do not help here: one checks that all critical points of the adaptive dynamics  $\dot{p} = \nabla_p \pi(p, q)|_{p=q}$ , corresponding to a strong selection weak mutation regime, all lie on the boundary of  $\Sigma_1$ .

**Strength of selection and the limits of adaptive dynamics.**

**Early-warning signals of evolutionary shifts.**

## Discussion

A large part of evolutionary theory in general, and evolutionary game theory in particular, has focused primarily on predicting the direction of evolution—or at least the phenotypes that evolution favors, be them fitness peaks or evolutionary stable states. Viewed through this angle, the problem of the evolution of cooperation is to determine whether, or under which condition, cooperative behavior is an evolutionarily stable strategy. The results presented in this letter show that Darwinian evolution through frequency-dependent selection does not always respect this paradigm: the iterated prisoner’s dilemma—the simplest and best studied setting to study the evolution of cooperation—contains a phase with multistable evolutionary dynamics.

This switch-like behavior is consistent with the observation that rapid, major evolutionary change can occur without any exogenous driver. Various explanations for such kind of “punctuated equilibrium” have been proposed. One, due to Sneppen and collaborators, relies on a simple but abstract and somewhat *ad hoc* model of self-organized critical evolutionary dynamics in a one-dimensional fitness space [sneppen’evolution’1995]. Another relates punctuations to the topology of genotype-to-phenotypes mappings: in very high-dimensional sequence spaces, it can take a long time for a population to diffuse along neutral networks until a new, fitter phenotype is found [fontana’continuity’1998, crutchfield’evolutionary’2002, bakhtin’evolution’2021]. A third explanation, sometimes called “quantum evolution”, depends on genetic drift to enable transient maladaptations before a new peak is found in the fitness landscape. In the multistable phase of  $\text{IPD}_1$ , rapid evolutionary change cannot be attributed to any of these causes.

Other aspects of the multistable behavior of  $\text{IPD}_1$  are more surprising.

Overall, our findings reveal a new mode of evolutionary dynamics in the iterated prisoner’s dilemma: one in which evolution has no attractor, no limit cycle and no preferred direction. In this phase, the evolutionary dynamics of cooperation may be compared with the reversible fluctuations of a physical system at thermal equilibrium—a process without history.

## Methods

**Simulations.** Agent-based simulations were coded and run using the Agents.jl julia package [datseris·agents·2021].