

**Thermodynamics of evolutionary games**

Christoph Adami\*

*Department of Microbiology & Molecular Genetics, Department of Physics & Astronomy, and BEACON Center for the Study of Evolution in Action, Michigan State University, East Lansing, Michigan 48824, USA*

Arend Hintze

*Department of Computer Science & Engineering, Department of Integrative Biology, and BEACON Center for the Study of Evolution in Action, Michigan State University, East Lansing, Michigan 48824, USA*

(Received 21 June 2017; revised manuscript received 3 April 2018; published 20 June 2018)

How cooperation can evolve between players is an unsolved problem of biology. Here we use Hamiltonian dynamics of models of the Ising type to describe populations of cooperating and defecting players to show that the equilibrium fraction of cooperators is given by the expectation value of a thermal observable akin to a magnetization. We apply the formalism to the public goods game with three players and show that a phase transition between cooperation and defection occurs that is equivalent to a transition in one-dimensional Ising crystals with long-range interactions. We then investigate the effect of punishment on cooperation and find that punishment plays the role of a magnetic field that leads to an “alignment” between players, thus encouraging cooperation. We suggest that a thermal Hamiltonian picture of the evolution of cooperation can generate other insights about the dynamics of evolving groups by mining the rich literature of critical dynamics in low-dimensional spin systems.

DOI: [10.1103/PhysRevE.97.062136](https://doi.org/10.1103/PhysRevE.97.062136)**I. INTRODUCTION**

Cooperation is a particularly interesting phenomenon in the context of evolution. Evolution acts on short-term benefits, which makes cooperators vulnerable to exploitation in the form of cheating or “defection” even if cooperation is a strategy with higher payoffs in the long term, creating what is known as the “dilemma of cooperation.” It is often stated that because of the dilemma, the expected outcome of evolution should be defection, rendering the plethora of examples for cooperators in nature mysterious. However, there are a number of different mechanisms that nevertheless enable cooperation [1–4] suggesting that, contrary to the naive expectation, cooperation is after all the natural outcome of evolution when mechanisms enabling assortment (such as discrimination via communication) are available [5]. These results have been obtained using mathematics as well as computational-simulation tools. The mathematical results in particular provide insight into the evolutionary dynamics that give rise to cooperation from inspecting closed-form solutions, but such solutions are hard to come by when populations are finite, are not well mixed, or are subject to significant mutation [5]. Recently, progress was made in understanding the evolutionary dynamics of games played on arbitrary grids [6], but closed-form solutions predicting the “critical point” for the transition between cooperation and defection still do not exist. Here, we use methods borrowed from statistical physics that show the path to such general formulas.

Prior investigations of several standard evolutionary games [7–12] revealed that the evolutionary process often critically depends on a single parameter that causes an abrupt change in the winning strategy. In some cases it is possible

to move the parameter beyond the critical point without triggering the transition—the hallmark of hysteresis [11]. These results suggest that there is an underlying analogy between evolutionary game dynamics and the statistical description of phase transitions. Indeed, Szabó and Hauert [13,14] applied mathematical methods that are used to describe critical phase transitions like those found in the celebrated Ising model [15] to evolutionary games on a lattice and showed (via numerical simulation, as well as the pair-approximation on square lattices) that the prisoner’s dilemma (PD) game dynamics on random regular lattices fall into the directed percolation class of phase transitions.

Here we take a different approach, by explicitly constructing Hamiltonians for game dynamics inspired by Ising-type models and studying games on finite regular lattices analytically (albeit only in one dimension). It might at first appear odd to consider thermal game theory, because temperature plays no role in evolutionary dynamics. In physics, thermal effects are due to fluctuations in energy, but payoffs in evolutionary games can fluctuate as well, for a number of different reasons. For example, a finite evolving population is subject to drift and thus to a random element in the payoffs. Mutations that change strategies can play a similar role. In evolutionary games, we can summarize the effect of fluctuations by introducing a parameter that controls the *strength of selection* in the game, using the “strategy adoption” mode of selection (see Ref. [14] and below). While the dynamics under this rule is not precisely the same as the “strategy inheritance” mode of Darwinian selection, the differences (also discussed in Ref. [14]) are irrelevant for our purposes. The relationship between game dynamics and Ising-type models has been reviewed recently [12].

To introduce our method and notation, we first study the prisoner’s dilemma Hamiltonian at finite temperature and

\*adami@msu.edu

recover well-known results. We then apply the method to the public goods game without punishment, which turns out to be equivalent to an Ising model with long-range interactions, but without a magnetic field. We then add punishment to the public goods game, leading to an Ising model with magnetic field (and corresponding hysteresis effects) that we solve exactly.

## II. PRISONER'S DILEMMA

The prisoner's dilemma is a game played between two individuals in which both players have to make a decision about whether to cooperate or to defect. After both players have made their choice—to cooperate (C) or to defect (D)—their actions are revealed and players receive a payoff according to a payoff matrix (note that the values in the matrix correspond to the payoff given to the “row” player)

$$E = \begin{array}{cc} & \begin{array}{c} C \\ D \end{array} \\ \begin{array}{c} C \\ D \end{array} & \begin{pmatrix} R & S \\ T & P \end{pmatrix} \end{array} \quad (1)$$

The payoffs in that matrix define the type of game to be played. To obtain a prisoner's dilemma, we must have [2]  $T > R > P > S$ . If the game is played repeatedly it becomes the *iterated* prisoner's dilemma (IPD), a variant not considered here. Evolutionary game theory focuses on determining what strategies are evolutionarily stable in a population of strategies. In the simplest case, competition is between two unconditional deterministic strategies: one that always cooperates and one that always defects. A population starts out as a mix of both strategies, and players interact with a defined number of neighbors. Each player's performance is evaluated by accumulating all payoffs received in that round. To model evolution, randomly picked players (called focal players) can now either maintain their strategy or adopt the strategy of a competitor. Over time this process will lead to the spread of successful strategies and thus to evolution. This process of probabilistic strategy adoption is similar to the dynamics of strongly interacting spins described by Glauber [16]. In such a model of ferromagnetism, adjacent particles interact so that their spins will predominantly align (a spin adopting the state of its neighbor), giving rise to an overall magnetization that depends on the temperature of the system. In the following, we explore this analogy more deeply.

We first derive the thermodynamics of the prisoner's dilemma with a payoff matrix where we set the reward  $R = b - c$  (the benefit of cooperation minus the cost), while the temptation payoff  $T = b$  (obtaining the benefit without bearing the cost). At the same time, the so-called “sucker payoff”  $S = -c$  due to paying the cost without any benefit, while  $P = 0$  is the “punishment” for both players mistrusting each other. In all of the following, we assume  $c \geq 0$  as well as  $b - c \geq 0$ , so that the net benefit  $r = b - c \geq 0$ , ensuring that a dilemma exists. Indeed, even though the benefit outweighs the cost ( $r > 0$ ), the Nash equilibrium and evolutionarily stable strategy is known to be defection, not cooperation. The payoff matrix in terms of these values then becomes

$$E = \begin{pmatrix} b - c & -c \\ b & 0 \end{pmatrix}.$$

To define a Hamiltonian (an operator that describes the total energy for this system) we can transform the payoffs into an energy by subtracting the payoff from its largest possible value. However, because this only adds a global constant it will cancel in observables, so to understand the population dynamics in terms of thermodynamics we can keep the payoff as is. A Hamiltonian is an operator that acts on a vector space (Hilbert space). A basis for the Hilbert space is spanned by the cooperative strategy C and the defecting strategy D by the vectors

$$C = |0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

and

$$D = |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

In analogy to Ising spin systems, the Hamiltonian for the PD game can then be written in terms of the energy matrix  $E$  and the projectors

$$P_0 = |0\rangle\langle 0| = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

and

$$P_1 = |1\rangle\langle 1| = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

as

$$H = \sum_{i=1}^N \sum_{m,n=0}^1 E_{mn} P_m^{(i)} \otimes P_n^{(i+1)}, \quad (2)$$

where the sum over  $i$  goes over all the sites in this one-dimensional (1D) “spin chain.”

We proceed by calculating the thermal partition function of the system by writing ( $\beta = 1/T$  is the inverse of the temperature, which the reader will not confuse with the temptation payoff)

$$Z = \text{Tr} e^{-\beta H} = \sum_x \langle x | e^{-\beta H} | x \rangle, \quad (3)$$

where  $|x\rangle = |m_1 m_2 \cdots m_N\rangle$  is a circular chain so that the  $N$ th site is adjacent to the first site. It is then easy to see that

$$\begin{aligned} Z &= \sum_{m_1 \cdots m_N} e^{-\beta(E_{m_1 m_2} + E_{m_2 m_3} + \cdots + E_{m_N m_1})} \\ &= \sum_{m_1 \cdots m_N} U_{m_1 m_2} U_{m_2 m_3} \cdots U_{m_N m_1} \\ &= \text{Tr} U^N, \end{aligned} \quad (4)$$

where  $U_{ij} = e^{-\beta E_{ij}}$ .

To determine the equilibrium population composition, we define an *order parameter* given by the fraction of cooperators minus the fraction of defectors. For spin chains this is equal to the *magnetization* of the chain, defined using a spin operator  $J_z$  for which  $\langle 0 | J_z | 0 \rangle = 1$  and  $\langle 1 | J_z | 1 \rangle = -1$ . This can be achieved, e.g., with ( $\sigma_z$  is a Pauli matrix)

$$J_z = \sigma_z = P_0 - P_1. \quad (5)$$

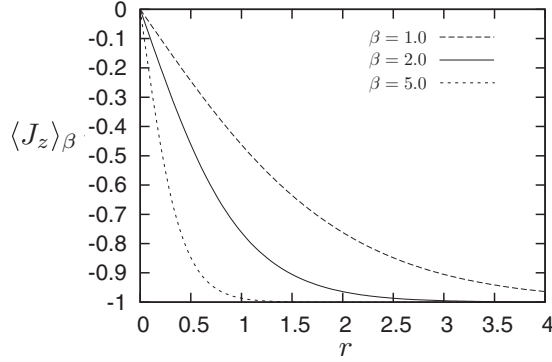


FIG. 1. Order parameter  $\langle J_z \rangle_\beta$  as a function of the net reward  $r = b - c$ , for three different temperatures. As opposed to the game in two dimensions [14], the phase transition occurs at  $r = 0$ .

We will understand this operator to act on the “row” player (that is, the first spin of the pair). For a chain of length  $N$ ,

$$J_z = \sum_i^N (P_0^{(i)} - P_1^{(i)}), \quad (6)$$

so that

$$\sum_x \langle x | J_z e^{-\beta H} | x \rangle = N \text{Tr} (U' U^{N-1}) \quad (7)$$

due to the cyclic property of the trace. Here we introduced the matrix  $U'_{ij} = (-1)^i U_{ij}$ . An explicit calculation shows that (recall that  $r = b - c$ )

$$Z = \text{Tr} U^N = (1 + e^{-\beta r})^N, \quad (8)$$

while since  $U' U = (1 + e^{-\beta r}) U'$  and  $\text{Tr} U' = -1 + e^{-\beta r}$ ,

$$\text{Tr} (U' U^{N-1}) = (1 + e^{-\beta r})^{N-1} (-1 + e^{-\beta r}), \quad (9)$$

so that finally the thermal expectation value of the magnetization is

$$\begin{aligned} \langle J_z \rangle_\beta &= \frac{1}{Z} \sum_x \langle x | J_z e^{-\beta H} | x \rangle \\ &= -N \tanh(\beta r / 2). \end{aligned} \quad (10)$$

We show the magnetization *per player* [Eq. (10) divided by  $N$ ] as a function of the critical parameter  $r$  in Fig. 1 and see that, at low temperatures (high  $\beta$ ), the population will consist mostly of defectors (negative magnetization) because this is the Nash equilibrium. We note that the parameter  $r$  plays the same role as the interaction strength  $J$  in the standard Ising model. The phase transition (vanishing magnetization) occurs at  $r = 0$  (the “boundary” of the parameter values), which is expected from the general arguments of van Hove [17] and of Landau [18] that forbid phase transitions in one-dimensional systems. Thus, we do not observe cooperation in the one-dimensional Prisoner’s dilemma, as is of course well known.

### III. PUBLIC GOODS GAME IN ONE DIMENSION

The PD game we just described turns out to be the two-player version of the more general public goods (PG) game. The PG game is a staple of evolutionary game theory as well as experimental economics [19–21] and has been used

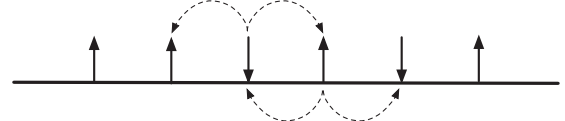


FIG. 2. One-dimensional string of players in a public goods game that interact with two nearest neighbors. Because players interact with more than one nearest neighbor, effectively next to nearest neighbors interact.

to understand the *Tragedy of the Commons* [22], a social dilemma that can lead to the overuse of public resources (for example, overfishing) because of selfish behavior. In the PG game, payoffs are defined for cooperators and defectors via

$$\Pi_C = \frac{r}{(k+1)} (N_C + 1) - 1, \quad (11)$$

$$\Pi_D = \frac{r N_C}{(k+1)}, \quad (12)$$

where  $\Pi_C$  is the payoff for a cooperator ( $\Pi_D$  for a defector).  $N_C$  is the number of cooperators in the neighborhood (not counting the focal player, so it is the number of cooperators in the player’s periphery), and  $r$  is the reward multiplier (synergy factor). These are the rules for a game with  $k + 1$  players in a group. In the following, we treat the game in one dimension (so  $k = 2$ ).

The rules (11) and (12) imply a payoff matrix

$$\Pi_C = \begin{array}{cc} & \begin{array}{c} C \\ D \end{array} \\ \begin{array}{c} C \\ D \end{array} & \begin{pmatrix} r-1 & \frac{2}{3}r-1 \\ \frac{2}{3}r-1 & \frac{1}{3}r-1 \end{pmatrix} \end{array} \quad (13)$$

for cooperators, where the matrix elements indicate the states of spins in the periphery of the focal player. For example,  $r - 1$  is the payoff for a cooperator surrounded by two cooperators. The payoff matrix for defectors is simply  $\Pi_D = \Pi_C - (\frac{1}{3}r - 1)$ .

We now construct a Hamiltonian to solve this evolutionary model exactly in two cases: one where the dynamics maximize the mean payoff of the population, and one in which the payoff of an individual is maximized. Naturally, we expect a correspondence with the evolutionary scenario only in the latter case. In this one-dimensional game, the population is arranged linearly so that each player forms a group with its left and right neighbor ( $k = 2$ ); see Fig. 2.

As mentioned earlier, we can create matrices for energies that should be minimized (rather than payoffs that need to be maximized) by subtracting the payoffs from the maximal payoff (here,  $r - 1$ ), leading to a ground state that has zero energy. Strictly speaking, the Hamiltonian for this system should be written as an interaction of three spins, but we will often write it in terms of a two-spin interaction matrix *conditional* on the state of the focal spin. For example, we can write

$$E^{(C)} = \begin{pmatrix} 0 & \frac{1}{3}r \\ \frac{1}{3}r & \frac{2}{3}r \end{pmatrix}, \quad E^{(D)} = \frac{1}{3}r - 1 + E^{(C)}. \quad (14)$$

We write a Hamiltonian for cooperators using these energies and the projectors previously defined:

$$H_C^{(i)} = \sum_{i=0}^N \sum_{m,n=0}^1 E_{mn}^{(C)} P_m^{i-1} \otimes P_n^{i+1}, \quad (15)$$

and similarly for  $H_D^{(i)}$ . The total Hamiltonian is (recall that  $P_0$  projects onto a cooperator, so that  $P_0|0\rangle = |0\rangle$  while  $P_0|1\rangle = 0$ )

$$H = \sum_{i=1}^N H_C^{(i)} P_0^{(i)} + H_D^{(i)} P_1^{(i)}. \quad (16)$$

Using the spin operator (6) and the methods outlined earlier, we obtain after a somewhat tedious calculation

$$\langle J_z \rangle_\beta = \frac{1}{Z} \text{Tr}(J_z e^{-\beta H}) = N \tanh \frac{\beta}{2} (r - 1), \quad (17)$$

suggesting a phase transition at  $r = 1$ , in contradiction with the standard result [11] that suggests a transition at  $r = 3$  (see below). The reason for this discrepancy is not difficult to find: Hamiltonian dynamics minimize the energy of the entire spin chain, which is equivalent to maximizing population fitness as a whole. Darwinian evolution, however, does not optimize population fitness, but rather maximizes the fitness of a single individual *within* a population.

We can implement the latter dynamic by dropping the sum over sites in Eq. (15) and consider only the contribution to the energy from a single spin with its two neighbors. In that case (we take the middle site to be the focal site whose energy is minimized)

$$\begin{aligned} Z &= \sum_{m_1 m_2 m_3} \langle m_1 m_2 m_3 | e^{-\beta H_{m_2}} | m_1 m_2 m_3 \rangle \\ &= \sum_{m_1 m_3} (U_{m_1 m_3} + V_{m_1 m_3}), \end{aligned} \quad (18)$$

where  $U$  is the ‘‘cooperative’’ matrix  $U = e^{-\beta H_0}$  while the defector matrix  $V = e^{-\beta H_1} = e^{\beta(r/3-1)} U$  because defector energies differ by  $r/3 - 1$  from cooperator energies; see Eq. (14). Then,

$$\begin{aligned} Z &= \sum_{m_1 m_3} U_{m_1 m_3} (1 + e^{-\beta(r/3-1)}) \\ &= (1 + e^{-\beta \frac{r}{3}})^2 (1 + e^{-\beta(\frac{r}{3}-1)}). \end{aligned} \quad (19)$$

Using the spin operator defined in Eq. (5) we obtain (again for a single focal player in the middle position)

$$\begin{aligned} &\sum_{m_1 m_2 m_3} \langle m_1 m_2 m_3 | J_z e^{-\beta H_{m_2}} | m_1 m_2 m_3 \rangle \\ &= (P_0 - P_1) \sum_{m_1 m_3} (P_0 U_{m_1 m_3} + P_1 V_{m_1 m_3}) \\ &= (1 - e^{-\beta(r/3-1)}) (1 + e^{-\beta \frac{r}{3}})^2, \end{aligned} \quad (20)$$

which allows us to calculate the order parameter as

$$\langle J_z \rangle_\beta = \frac{1}{Z} \text{Tr}(e^{-\beta H} J_z) = \tanh \frac{\beta}{2} \left( \frac{1}{3} r - 1 \right). \quad (21)$$

This function is plotted in Fig. 3 and suggests that a phase transition with an interior critical point is possible in this game

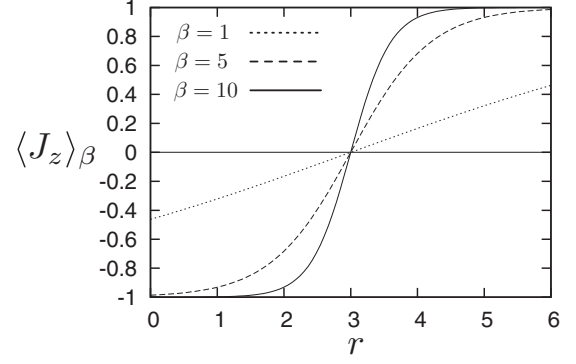


FIG. 3. Exact solution for the order parameter  $\langle J_z \rangle_\beta$  as a function of synergy parameter  $r$  for three different temperatures, from Hamiltonian dynamics.

even though the game is one dimensional, seemingly violating van Hove’s theorem [17]. However, the theorem forbidding internal critical points in one dimension only holds for short-range interactions, while the interaction between three players studied here is not of that kind.

To test the accuracy of our theoretical result, we now simulate the public goods game using agent-based methods [5,11,23]. In the agent-based simulations we use a population of 1024 players that either cooperate or defect, arranged in a one-dimensional chain just as in Fig. 2. Which of the two moves an agent chooses is determined by a genome (here a single locus) that evolves. At every update, players have a chance to change their strategy by probabilistically adopting the strategy of a competitor (Glauber dynamics; see, e.g., Refs. [7,14]) using the rule (here  $x$  is the focal player while  $y$  is an alternative strategy)

$$p(x \leftarrow y) = \frac{1}{1 + e^{-\beta(w_x - w_y)}}, \quad (22)$$

where  $\beta$  is related to the strength of selection and  $w$  is the fitness of each player defined by the payoff the player receives. In the case of rejection (i.e., nonadoption) the focal player retains its strategy.

We define an order-parameter-like function that indicates to what extent the population is in a cooperative or a defective regime. This parameter depends on the fraction of players in the population cooperating ( $P_C$ ) and the fraction defecting ( $P_D$ ) and is defined as

$$M = \frac{P_C - P_D}{P_C + P_D}. \quad (23)$$

The agent-based simulations confirm that the fate of an evolving population depends critically on the synergy factor  $r$  (see Fig. 4), and changes from negative (defection) to positive (cooperation) at  $r = 3$ , in accordance with the critical  $r_c = k + 1$  for strategies to evolve cooperative behavior in the public goods game [11]. In particular, the simulations confirm the theoretical results with high accuracy.

#### IV. PUBLIC GOODS GAME WITH PUNISHMENT

Cooperation evolves in the PG game if the synergy  $r$  is at least as large as the group’s size  $k + 1$ . However, it is

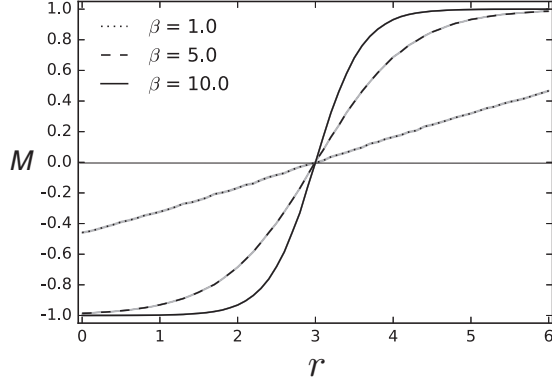


FIG. 4. Fraction of cooperators for a chain of length  $2^{10}$ , as a function of the synergy parameter  $r$  for three different selection strengths defined by  $\beta = 1/T$ . Each data point (each  $r$ , increments of  $\Delta r = 0.1$ ) is the average over 100 replicate agent-based simulations using strategy adoption for  $2 \times 10^6$  updates. Barely visible gray bands represent standard error.

unlikely that in nature cooperation would ever create such a high synergy factor, implying that cooperation cannot evolve in this type of game. It has previously been suggested that punishment is one way to promote cooperation [23–29]. By introducing punishment, players can now not only choose between cooperation and defection, but can do this in conjunction with deciding whether to punish cheaters. This introduces two more strategies: a “moralist” M who cooperates and punishes, and an “immoralist” I who defects but also punishes [23]. For every player punished for defecting, each punishing player must pay a cost  $\gamma$ , and every player that is punished in such a way suffers a fine  $\epsilon$ , thus extending the rules (11) and (12) to (here, we show the special 1D case  $k = 2$ , for the general case see, for example, Ref. [11])

$$\Pi_C = \frac{r}{3}(N_C + N_M + 1) - 1, \quad (24)$$

$$\Pi_D = \frac{r(N_C + N_M)}{3} - \epsilon \left( \frac{N_M + N_I}{2} \right), \quad (25)$$

$$\Pi_M = \Pi_C - \gamma \left( \frac{N_D + N_I}{2} \right), \quad (26)$$

$$\Pi_I = \Pi_D - \gamma \left( \frac{N_D + N_I}{2} \right), \quad (27)$$

where  $N_i$  is the number of players in the immediate neighborhood of the focal player with strategy  $i$ ,  $\epsilon$  parametrizes the effect of punishment, while  $\gamma$  stands for the cost of punishment (see Refs. [11,23]).

We now study this model thermodynamically, but in order to compare with the evolutionary dynamics we study the regime where the energy of a single site is minimized. To account for the additional strategies (beyond cooperator and defector), we extend the Hilbert space by allowing for a site-dependent magnetization  $|i\rangle \rightarrow |i\rangle|j\rangle$ , so that each strategy is defined by a product of spin vectors. If we define punishment as  $|0\rangle$  and

nonpunishment as  $|1\rangle$ , we can write the states of the punishing and nonpunishing cooperator as

$$M = |0\rangle|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad (28)$$

$$C = |0\rangle|1\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}. \quad (29)$$

The payoffs (24)–(27) can be written in terms of a Hamiltonian for each of the four strategies as

$$H = P_{00}H_M + P_{01}H_C + P_{10}H_D + P_{11}H_I, \quad (30)$$

with projectors  $P_{ij}$  on the respective states (with  $\sum_{ij=0}^1 P_{ij} = 1$ ). Each Hamiltonian  $H_k$  ( $k = C, D, M, I$ ) is written in terms of an energy matrix  $F^{(k)}$  just as in Eq. (15),

$$F^{(C)} = \begin{matrix} & \begin{matrix} C & D & M & I \end{matrix} \\ \begin{matrix} C \\ D \\ M \\ I \end{matrix} & \begin{pmatrix} 0 & \frac{r}{3} & 0 & \frac{r}{3} \\ \frac{r}{3} & \frac{2}{3}r & \frac{r}{3} & \frac{2}{3}r \\ 0 & \frac{r}{3} & 0 & \frac{r}{3} \\ \frac{r}{3} & \frac{2}{3}r & \frac{r}{3} & \frac{2}{3}r \end{pmatrix} \end{matrix} = \begin{pmatrix} E^{(C)} & E^{(C)} \\ E^{(C)} & E^{(C)} \end{pmatrix}. \quad (31)$$

Similarly,

$$F^{(D)} = \frac{r}{3} - 1 + \begin{pmatrix} E^{(C)} & E^{(C)} + \frac{\epsilon}{2} \\ E^{(C)} + \frac{\epsilon}{2} & E^{(C)} + \epsilon \end{pmatrix}, \quad (32)$$

$$F^{(M)} = \begin{pmatrix} E^{(C)} & E^{(C)} + \frac{\gamma}{2} \\ E^{(C)} + \frac{\gamma}{2} & E^{(C)} + \gamma \end{pmatrix}, \quad (33)$$

$$F^{(I)} = \frac{r}{3} - 1 + \begin{pmatrix} E^{(C)} & E^{(C)} + \frac{\gamma + \epsilon}{2} \\ E^{(C)} + \frac{\gamma + \epsilon}{2} & E^{(C)} + \gamma + \epsilon \end{pmatrix}. \quad (34)$$

We can now calculate the partition function

$$Z = \text{Tr}(e^{-\beta H}) = Z_C + Z_D + Z_M + Z_I \quad (35)$$

on account of the decomposition (30), where

$$Z_C = \text{Tr}(e^{-\beta H_C}) = 4(1 + e^{-\beta \frac{r}{3}})^2, \quad (36)$$

or four times the contribution from each  $E^{(C)}$ . Similarly,

$$Z_D = e^{-\beta(\frac{1}{3}r-1)}(1 + e^{-\beta \frac{r}{3}})^2(1 + e^{-\beta \frac{\epsilon}{2}})^2, \quad (37)$$

$$Z_M = 4(1 + e^{-\beta(\frac{r}{3} + \frac{\gamma}{2})})^2, \quad (38)$$

$$Z_I = e^{-\beta(\frac{1}{3}r-1)}(1 + e^{-\beta(\frac{r}{3} + \frac{\gamma}{2})})^2(1 + e^{-\beta \frac{\epsilon}{2}})^2. \quad (39)$$

Finally, we obtain the order parameter that measures the degree of cooperation (the fraction of C and M players minus the fraction of D and I players), which turns into the surprisingly simple expression (see Fig. 5)

$$\langle J_z \rangle_\beta = \frac{1 - \cosh^2\left(\beta \frac{\epsilon}{4}\right) e^{-\beta(\frac{r}{3} + \frac{\epsilon}{2} - 1)}}{1 + \cosh^2\left(\beta \frac{\epsilon}{4}\right) e^{-\beta(\frac{r}{3} + \frac{\epsilon}{2} - 1)}}. \quad (40)$$

Note that the order parameter only depends on the effect of punishment  $\epsilon$  but *not* the cost  $\gamma$  and reduces to expression (21) in the limit  $\epsilon \rightarrow 0$ .



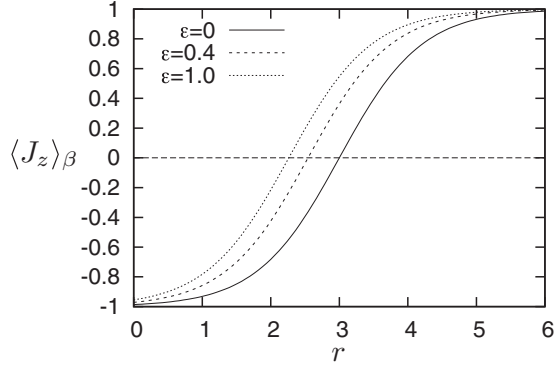


FIG. 5. Exact results for the order parameter  $\langle J_z \rangle_\beta$  for the public goods game with punishment, as a function of synergy parameter  $r$  for three different punishment fines  $\epsilon$ , at a constant temperature ( $\beta = 5$ ).

To check the theory, we can extend the agent-based model described above by including the two new strategies I and M. As before, we use 1024 players in a population that is arranged linearly (see Methods), and games are played in groups of three. Again, when we evolve this population using strategy adoption, we see the dependence of the critical point on the synergy factor  $r$  and the selection strength  $\beta = 1/T$ . Since the game now includes two more strategies, we have to modify the function  $M$  that describes the fraction of cooperators in the game to contain all four strategies as the fraction of contributing (cooperating) strategies:

$$M = \frac{(P_C + P_M) - (P_D + P_I)}{P_C + P_D + P_M + P_I}. \quad (41)$$

Evolving these populations using different fines  $\epsilon$  and costs  $\gamma$ , we find that the critical point now only depends on  $\epsilon$  (see Fig. 6), and moves the critical point in such a manner that the punishment fine reduces the critical synergy for cooperation [11]. It turns out that the closed-form solution Eq. (40) reproduces the agent-based simulations shown in Fig. 6 to a remarkable extent, confirming the unintuitive finding that the critical point only depends on the effect, but not on the cost, of punishment. The Hamiltonian model also

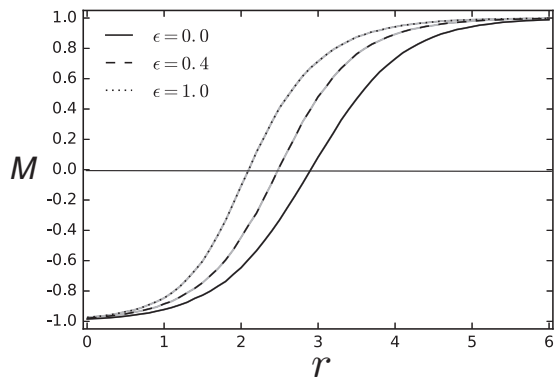


FIG. 6. Fraction of cooperators  $M$  for a chain of length  $2^{10}$  as a function of the synergy parameter  $r$  for three different fines  $\epsilon$  (at fixed  $\beta = 5$ ). Each data point (increments of  $\delta r = 0.1$ ) is the average over 100 replicates running the agent-based simulation with Glauber dynamics for  $2 \times 10^6$  updates.

clarifies that punishment indeed acts like a magnetic field that encourages alignment of spins, thus explaining why, in agent-based simulations, punishment induces hysteresis as a population is subjected to an adiabatically varying  $r$  [11]. Further work using the Hamiltonian model of cooperation with punishment may elucidate other aspects of the critical dynamics, in particular for games in higher dimensions, with more players per group, or even on irregular lattices.

## V. DISCUSSION

Evolutionary game theory is a mathematical framework that has been eminently successful at unraveling the numerous elements that impact decisions, and to work out the consequences of decisions. While both mathematics and computational simulations have influenced this field (see, for example, the review [5], along with commentaries), the relationship between game theory and physics has been explored less. We know that in realistic scenarios, decisions must be made under uncertainty; either due to unpredictable environments, or due to inherent noise. For evolutionary dynamics in particular, noise is unavoidable. After all, high reproductive potential does not guarantee survival, but only biases future outcomes. A standard result of population genetics, for example, predicts that a gene that confers a ten percent advantage in reproductive rate only has a twenty percent chance of being represented in future generations. The branch of science best equipped to tackle the impact of chance on dynamics is physics, with a well-developed corpus of results in statistical mechanics and thermodynamics. A growing literature has found success in mining these well-established methods, from harnessing the Fokker–Planck equation to describe the effect of chance due to drift in small populations [30] to using tools from statistical mechanics to study the universality class of phase transitions in the spatial prisoner’s dilemma [14]. Here, we tapped a different set of well-established tools from statistical physics; namely, the thermodynamics of spin systems. The analogy between the critical dynamics of spin systems and game theory is not difficult to see. After all, the correspondence between Eigen and Schuster’s model for the evolution of macromolecules [31] and two-dimensional (2D) Ising models was pointed out over thirty years ago [32] (see also section 11.4 in Ref. [33]) but we have not, as yet, seen a concerted effort to marshal the considerable machinery developed to tackle low-dimensional condensed-matter structures to aid in understanding evolutionary game theory.

It may seem odd, at first sight, that a thermodynamic approach to game theory is possible at all, given that thermodynamics relies on the assumption that the system tends towards equilibrium, whereas in many game-theoretic situations (in particular, those that are of the rock-paper-scissors type) the system appears to be maintained out of equilibrium. Fortunately, it is possible to show that even in systems out of equilibrium, detailed balance can be assured as long as microscopic reversibility is guaranteed [34,35]. While this result depends on the nature of the boundary condition (it holds under “normal” boundary conditions; that is, boundary conditions in which the probability distribution vanishes at the boundary), there are strong reasons to believe that at least in the limit of large systems and low mutation rates, detailed balance

can always be achieved for these games. Investigating this issue more deeply is left for forthcoming work.

The Hamiltonian approach we described here leads to important new insights about the dynamics of evolving populations at fixed strength of selection (and thus, to some extent, fixed temperature). First, we have shown that the standard statistical approach in which the energy of the entire ensemble is minimized does not correspond to the evolutionary scenario, giving rise instead to a transition at  $r = 1$ . That result would imply that a dilemma is absent, and indeed this is precisely what we would expect if groups of organisms, rather than individuals, are selected. Second, the treatment of the public goods game with punishment revealed that punishment plays the role that a local magnetic field plays when interacting with a system that can display spontaneous magnetization. An extensive literature in the area of spin-glasses of the Sherrington–Kirkpatrick type [36] suggests that local magnetic fields can give rise to spontaneous symmetry breaking, and that their mean-field solutions are similar to those of local spins interacting with a global magnetic field. These insights immediately suggest looking for effects such as hysteresis (as seen, for example, in Ref. [11]), but also interactions between hysteresis and impurities, for example. Indeed, it is not unreasonable to imagine that including a third player strategy such as “abstaining” [37–39] can be viewed as impurities that can dramatically alter the critical dynamics we observe, for example by “pinning” the interfaces between domains [40]. While the dynamics that includes abstaining may give rise to intransitive dominance dynamics [41] (such as the rock-paper-scissors game), the arguments given above that out-of-equilibrium dynamics still gives rise to stationary equilibrium distributions as long as microscopic reversibility holds, suggest that such games can also be solved using Hamiltonian dynamics.

We should caution, however, that extending the present results to games in higher dimensions will be difficult. For example, while the Ising model can be solved in two dimensions, there is no solution for the model in two dimensions with a magnetic field, as it is related to the three-dimensional model for which a closed-form solution does not exist. Nevertheless, we expect that the tools developed here will be useful because if

the analogy between evolutionary game dynamics and phase transitions in spin systems is established, other results from the rich literature of critical phenomena in spin systems may inform us about the dynamics of cooperation in groups. In particular, an extension of the calculation shown here to two dimensions may produce an exact solution along the lines of Onsager’s, which would allow us to move beyond pair-approximations for games on a 2D regular lattice. We hope that the simple results derived here (validated via computational simulation) can serve as a seed for the future development of this field.

## METHODS

The computational evolutionary model instantiates a population of 1024 random agents in a circular configuration. At each update a single agent is randomly selected and its payoff computed by playing the strategy against its left and right neighbors. At the same time, the payoff of a strategy to potentially replace the agent is computed. In the case of a two-player game (C and D) the only other alternative strategy is used, in the case of four players (C, D, M, and I) one alternative strategy is chosen at random. Instead of the evolutionary updating of the population described in Refs. [5,11], here the likelihood to replace the strategy of the selected agent with the alternative is given by Eq. (22). In each replicate run, we updated strategies 2 million times (roughly 2000 updates per site), then calculated the order parameter.

The code as well as the analysis scripts to create all figures can be found in Ref. [43].

## ACKNOWLEDGMENTS

We thank Nathaniel Pasmanter for collaboration in the early stages of this work, as well as Claus Wilke for discussions. This work was supported in part by National Science Foundation’s BEACON Center for the Study of Evolution in Action, under Contract No. DBI-0939454. We wish to acknowledge the support of the Michigan State University High Performance Computing Center and the Institute for Cyber-Enabled Research.

- 
- [1] J. Maynard Smith, *Evolution and the Theory of Games* (Cambridge University Press, Cambridge, 1982).
  - [2] R. Axelrod, *The Evolution of Cooperation* (Basic Books, New York, 1984).
  - [3] J. Hofbauer and K. Sigmund, *Evolutionary Games and Population Dynamics* (Cambridge University Press, Cambridge, 1998).
  - [4] M. Nowak, *Evolutionary Dynamics* (Harvard University Press, Cambridge, 2006).
  - [5] C. Adami, J. Schossau, and A. Hintze, Evolutionary game theory using agent-based methods, *Phys. Life Rev.* **19**, 1 (2016).
  - [6] B. Allen, G. Lippner, Y.-T. Chen, B. Fotouhi, N. Momeni, S.-T. Yau, and M. A. Nowak, Evolutionary dynamics on any population structure, *Nature (London)* **544**, 227 (2017).
  - [7] G. Szabó and G. Fáth, Evolutionary games on graphs, *Phys. Rep.* **446**, 97 (2007).
  - [8] A. Szolnoki, M. Perc, and G. Szabó, Phase diagrams for three-strategy evolutionary prisoner’s dilemma games on regular graphs, *Phys. Rev. E* **80**, 056104 (2009).
  - [9] D. Iliopoulos, A. Hintze, and C. Adami, Critical dynamics in the evolution of stochastic strategies for the iterated prisoner’s dilemma, *PLoS Comput. Biol.* **6**, e1000948 (2010).
  - [10] C. Adami, J. Schossau, and A. Hintze, Evolution and stability of altruist strategies in microbial games, *Phys. Rev. E* **85**, 011914 (2012).
  - [11] A. Hintze and C. Adami, Punishment in public goods games leads to meta-stable phase transitions and hysteresis, *Phys. Biol.* **12**, 046005 (2015).

- [12] G. Szabó and I. Borsos, Evolutionary potential games on lattices, *Phys. Rep.* **624**, 1 (2016).
- [13] G. Szabó and C. Hauert, Phase Transitions and Volunteering in Spatial Public Goods Games, *Phys. Rev. Lett.* **89**, 118101 (2002).
- [14] C. Hauert and G. Szabó, Game theory and physics, *Am. J. Phys.* **73**, 405 (2005).
- [15] E. Ising, Beitrag zur Theorie des Ferromagnetismus, *Z. Phys.* **31**, 253 (1925).
- [16] R. J. Glauber, Time-dependent statistics of the Ising model, *J. Math. Phys.* **4**, 294 (1963).
- [17] L. van Hove, Sur l'intégrale de configuration pour les systèmes de particules à une dimension, *Physica* **16**, 137 (1950).
- [18] L. D. Landau and E. M. Lifschitz, *Statistische Physik Teil 1* (Akademie-Verlag, Berlin, 1987).
- [19] M. Olson, *The Logic of Collective Action: Public Goods and the Theory of Groups* (Harvard University Press, Cambridge, 1971).
- [20] D. D. Davis and C. A. Holt, *Experimental Economics* (Princeton University Press, Princeton, 1993).
- [21] J. O. Ledyard, Public goods: A survey of experimental research, in *Handbook of Experimental Economics*, edited by John H. Kagel and Alvin E. Roth (Princeton University Press, Princeton, 1995), pp. 111–194.
- [22] G. Hardin, The tragedy of the commons, *Science* **162**, 1243 (1968).
- [23] D. Helbing, A. Szolnoki, M. Perc, and G. Szabó, Evolutionary establishment of moral and double moral standards through spatial interactions, *PLoS Comput. Biol.* **6**, e1000758 (2010).
- [24] E. Fehr and S. Gächter, Altruistic punishment in humans, *Nature (London)* **415**, 137 (2002).
- [25] E. Fehr and U. Fischbacher, The nature of human altruism, *Nature (London)* **425**, 785 (2003).
- [26] C. F. Camerer and E. Fehr, When does “economic man” dominate social behavior?, *Science* **311**, 47 (2006).
- [27] K. Sigmund, C. Hauert, and M. A. Nowak, Reward and punishment, *Proc. Natl. Acad. Sci. U. S. A.* **98**, 10757 (2001).
- [28] R. Boyd, H. Gintis, S. Bowles, and P. J. Richerson, The evolution of altruistic punishment, *Proc. Natl. Acad. Sci. U. S. A.* **100**, 3531 (2003).
- [29] H. Brandt, C. Hauert, and K. Sigmund, Punishment and reputation in spatial public goods games, *Proc. R. Soc. London, Ser. B* **270**, 1099 (2003).
- [30] A. Traulsen and Ch. Hauert, Stochastic evolutionary game dynamics, in *Reviews of Nonlinear Dynamics and Complexity*, edited by H.-G. Schuster (Wiley-VCH, Weinheim, 2009), Vol. 2, pp. 25–63.
- [31] M. Eigen and P. Schuster, *The Hypercycle—A Principle of Natural Self-Organization* (Springer-Verlag, Berlin, 1979).
- [32] I. Leuthäusser, Statistical mechanics of Eigen’s evolution model, *J. Stat. Phys.* **48**, 343 (1987).
- [33] C. Adami, *Introduction to Artificial Life* (Springer Verlag, New York, 1998).
- [34] R. Graham and H. Haken, Generalized thermodynamical potential for Markoff systems in detailed balance and far from thermal equilibrium, *Z. Phys. A: Hadrons Nucl.* **243**, 289 (1971).
- [35] H. Risken, Solutions of the Fokker-Planck equation in detailed balance, *Z. Phys. A: Hadrons Nucl.* **251**, 231 (1972).
- [36] D. Sherrington and S. Kirkpatrick, Solvable Model of a Spin-Glass, *Phys. Rev. Lett.* **35**, 1792 (1975).
- [37] J. H. Fowler, Altruistic punishment and the origin of cooperation, *Proc. Natl. Acad. Sci. U. S. A.* **102**, 7047 (2005).
- [38] H. Brandt, C. Hauert, and K. Sigmund, Punishing and abstaining for public goods, *Proc. Natl. Acad. Sci. U. S. A.* **103**, 495 (2006).
- [39] C. Hauert, A. Traulsen, H. Brandt, M. A. Nowak, and K. Sigmund, Public goods with punishment and abstaining in finite and infinite populations, *Biol. Theory* **3**, 114 (2008).
- [40] Note that, for this analogy to hold, the fraction of abstaining players (the density of impurities) must be kept constant. In general, these impurities can move through the crystal, but were kept fixed in numerical simulations of this effect [42].
- [41] C. Hauert, S. De Monte, J. Hofbauer, and K. Sigmund, Volunteering as Red Queen mechanism for cooperation in public goods games, *Science* **296**, 1129 (2002).
- [42] M. H. Vainstein and J. J. Arenzon, Disordered environments in spatial games, *Phys. Rev. E* **64**, 051905 (2001).
- [43] [https://github.com/Hintzelab/TD\\_EGT](https://github.com/Hintzelab/TD_EGT).