The intricate geometry of zero-determinant strategies underlying evolutionary adaptation from extortion to generosity

Xingru Chen\textsuperscript{1,2,}\textsuperscript{*}, Long Wang\textsuperscript{3} and Feng Fu\textsuperscript{2,4}\textsuperscript{*}

\textsuperscript{1} School of Sciences, Beijing University of Posts and Telecommunications, Beijing 100876, People’s Republic of China
\textsuperscript{2} Department of Mathematics, Dartmouth College, Hanover, NH 03755, United States of America
\textsuperscript{3} Center for Systems and Control, College of Engineering, Peking University, Beijing 100871, People’s Republic of China
\textsuperscript{4} Department of Biomedical Data Science, Geisel School of Medicine at Dartmouth, Lebanon, NH 03756, United States of America

\textsuperscript{*} Author to whom any correspondence should be addressed.
E-mail: xingrucz@gmail.com

Keywords: direct reciprocity, cooperation, equalizer, adaptive dynamics

Abstract

The advent of zero-determinant (ZD) strategies has reshaped the study of reciprocity and cooperation in the iterated prisoner’s dilemma games. The ramification of ZD strategies has been demonstrated through their ability to unilaterally enforce a linear relationship between their own average payoff and that of their co-player. Common practice conveniently represents this relationship by a straight line in the parametric plot of pairwise payoffs. Yet little attention has been paid to studying the actual geometry of the strategy space of all admissible ZD strategies. Here, our work offers intuitive geometric relationships between different classes of ZD strategies as well as nontrivial geometric interpretations of their specific parameterizations. Adaptive dynamics of ZD strategies further reveals the unforeseen connection between general ZD strategies and the so-called equalizers that can set any co-player’s payoff to a fixed value. We show that the class of equalizers forming a hyperplane is the critical equilibrium manifold, only part of which is stable. The same hyperplane is also a separatrix of the cooperation-enhancing region where the optimum response is to increase cooperation for each of the four payoff outcomes. Our results shed light on the simple but elegant geometry of ZD strategies that is previously overlooked.

1. Introduction

The evolution of cooperation is one of the long-standing conundrums facing researchers from diverse fields [1]. Over the past decades, a number of seminal contributions have been made to improve our understanding of cooperation [2]. Among others, cooperation can prevail under reciprocal altruism [3], which means ‘you scratch my back and I will scratch yours’. As a repeated two-player game, iterated prisoner’s dilemma (IPD) games have been the paradigm for studying direct reciprocity and cooperation [4–10]. In particular, the famous Axelrod’s tournament has ushered in an era of studying powerful IPD strategies using computer simulations in combination with analytical approaches [11–13].

A plethora of memory-\textit{n} strategies including deterministic strategies and their stochastic counterparts have been thoroughly investigated [8, 14]. IPD strategies can be as sophisticated as possible, such as finite automata [15], lookup table-based strategies [16], or strategies that are optimized by using techniques like neural networks or swarm particle intelligence [17]. On the other hand, winning IPD strategies can be surprisingly simple but powerful. The prominent tit-for-tat (TFT), for example, is such a ‘fair-minded’ strategy that is found to be the backbone of direct reciprocity [6]. In the original version, TFT responds with full cooperation after co-player’s cooperative move but always retaliates with full defection after co-player’s defection. Variants of TFT, often called ‘compliers’ [18], include the so-called generous TFT (GTFT) [7] which forgives co-player’s defection with a certain probability. These special strategies can be part of a larger class of reactive strategies that condition their responses on their co-player’s behavioral
choice (C vs D) [7, 19]. The notion of reactive strategies has enabled analytical insights into the evolution of reciprocity and cooperation in the IPD games [20].

Memory-one IPD strategies are further specified by the probability to cooperate for the initial move and the probability to cooperate conditioned on each of the four possible payoff outcomes in a single round, denoted by \([q_1, q_2, q_3, q_4]\). The ordered labels 1, 2, 3, and 4, respectively, refer to the payoff outcome \(R\) from the pair of strategy choices, \((C, C)\), \(S\) from \((C, D)\), \(T\) from \((D, C)\), and \(P\) from \((D, D)\), as described from the focal row player’s perspective via the payoff matrix

\[
\begin{pmatrix}
C & D \\
C & R & S \\
D & T & P
\end{pmatrix}.
\]

For prisoner’s dilemma (PD), we have \(T > R > P > S\). Conventional PD games also assume \(2R > T + S > 2P\), that is, mutual cooperation is better than unilateral cooperation better than mutual defection.

From the perspective of evolution, another strategy known as win–stay lose–shift (WSLS) stands out later [7]. As for payoff control and manipulation, the discovery of equalizers is worthy of note [21]. The authors describe such a simple strategy, up to a properly chosen normalization factor \(\phi\), that is able to set any co-player’s average payoff \(O\) to be the exact same amount between \(P\) and \(R\). Using an elegant approach of linear algebra [22], Press and Dyson further show that the so-called zero-determinant (ZD) strategies are able to unilaterally enforce a linear relationship between its average payoff \(s_X\) and that of the co-player’s \(s_Y\) from \((s_X - 0) = \chi (s_Y - O)\), where \(O\) is the baseline payoff, and \(\chi\) is the extortion factor. Thus, the class of equalizers becomes a limiting subset of ZD strategies as \(\chi \to \infty\) (that is, to unilaterally set co-player \(Y\)’s average payoff \(s_Y\) to be \(O\)). ZD strategies can be categorized by their intended level of generosity \(O\) [23]. The class with \(O = P\) and \(\chi > 1\) is often called extortionate ZD since these players can always ensure their advantage with an unfair surplus as \(s_X - P = \chi (s_Y - P) \geq 0\) for conventional PD games. The class with \(O = R\) and \(\chi > 1\) is called generous ZD, ensuring that their own average payoff is never greater than the co-player’s as \(s_X - R = \chi (s_Y - R) \leq 0\).

Efforts on various extensions to ZD strategies have been proven fruitful [24–27], including but not limited to multi-person games [28–31], noises or errors [32, 33], and finitely repeated games [34]. The evolution of ZD strategies has been studied in finite populations [18, 35] as well as in structured populations [36]. The overall insight is that extortionate ZD strategies are powerful yet not evolutionarily stable in the sense that they will neutralize each other’s advantage unless they adapt to be generous [37–40]. Nevertheless, they can be the catalysts for the evolution of cooperation that pave the way for the emergence of cooperation (TFT, GTFT, or more generally, generous ZD) [41]. Noteworthy, the dominance and optimality of ZD strategies depend on the underlying payoff structure, which is determined by the sign of \(T + S - 2P\) [42]. It is found that a seemingly formidable ZD strategy can actually be outperformed, for example, by WSLS if \(T + S < 2P\) and that when against fixed unbending strategies, the best response of ZD players is to offer a fair split by letting \(\chi \to 1\) [42].

Let us now turn to the superset of ZD strategies, which is a collection of memory-one strategies \([q_1, q_2, q_3, q_4]\) with three free parameters \((O, \chi, \phi)\):

\[
\begin{align*}
q_1 &= 1 - \phi(R - O)(\chi - 1), \\
q_2 &= 1 - \phi((T - O)\chi + (O - S)], \\
q_3 &= \phi[(O - S)\chi + (T - O)], \\
q_4 &= \phi(O - P)(\chi - 1).
\end{align*}
\]

For \(q_i\)’s must be within \([0, 1]\), the admissible ranges of \((O, \chi, \phi)\) are given by

\[
\left\{ \begin{array}{l}
P \leq O \leq R, \\
1 \leq \chi < \infty, \\
0 \leq \phi \leq \text{min}\left\{ \frac{1}{(O - S)\chi + (T - O)}, \frac{1}{(T - O)\chi + (O - S)} \right\},
\end{array} \right.
\]

or
The normal vector is \( \mathbf{n} = [T + S - 2O, -(R - O), -(R - O)] \). For instance, for baseline payoff \( O = P \), the normal vector of this plane is \([T + S - 2P, -(R - P), -(R - P)]\) and for \( O = R \), the normal vector is \([2R - T - S, 0, 0]\), or simply \([1, 0, 0]\).

Of particular interest, the ‘equalizer’ subset of ZD strategies (obtained by letting \( |\chi| \to +\infty\)) forms another plane (figure 1(d)), which can be written as

\[
q_3 = \frac{(T - S)q_1 - (R - S)q_2 - (T - R)}{T - R}.
\]
Figure 1. Geometry of ZD strategies. Any ZD strategy \([q_1, q_2, q_3, q_4]\) is parameterized by three parameters: the baseline payoff \(O\), the extortion factor \(\chi\), and the normalization factor \(\phi\). Therefore, they can be visualized in the cube \([0, 1]^3\) for the ordered triplet \((q_1, q_2, q_3)\) using the fact that \(q_4\) is linearly dependent on \(q_1, q_2, \) and \(q_3\), as shown in equation (3). (a), (b), and (c) Show that ZD strategies with the same baseline payoff \(O\) are on the same plane, respectively, for \(O = P, (T + S)/2, \) and \(R\). (d) Shows that all equalizer strategies are located on the same plane, indicated by BCDE, and the dashed triangle DEG indicates equalizer strategies that are inadmissible. The lines DE and DB and DC, also highlighted in (a)–(c), present equalizer strategies that are able to fix their co-player’s average payoff to be \(P, (T + S)/2\), and \(R\), respectively. On the other hand, ZD strategies with different \(O\) share the line DA as their common limit for \(\chi \to 1\). When increasing \(O\), the plane formed by ZD strategies with the same \(O\) in effect rotates around the axis DA. All ZD strategies sharing the same \(O\) and \(\chi\) but with different \(\phi\) are on the same line; for a fixed \(O\), the corresponding line approaches the limit of equalizer monotonically from both directions as \(\chi\) tends to \(+\infty\) from 1 or to \(-\infty\) from \(\min\{-[(O - S)/(O - T), -(T - O)/(O - S)]\}\), namely, from DA to DE and from DH to DE in (a), from DA to DB and from DI to DB in (b), and from DA to DC and from DF to DC in (c). For all figures in the paper, we consider the conventional payoff matrix \([R, S, T, P]\) = \([3, 0, 5, 1]\).

We note that the equation above is exactly equivalent to the denominator of the expression of \(\chi\) as a function of \((q_1, q_2, q_3)\) given in equation (5). This plane of equalizer intersects with the plane of ZD strategies with fixed \(O\), and the intersection line fixes any co-player’s payoff to the same level \(O\) (figure 1(d)).

For any given \(O\), all of the corresponding ZD strategies contain the line DA connecting \((1, 1, 0)\) and \((1, 0, 1)\). In other words, the resulting plane formed by ZD strategies, while varying \(O\) from \(P\) to \(R\), rotates with a common axis DA which is the line connecting \((1, 1, 0)\) and \((1, 0, 1)\). This unprecedented geometry of ZD strategies can also be indirectly validated by the fact that the normal vector of the plane \(\vec{n} = [T + S - 2O, -(R - O), -(T - O)/(O - S)]\) formed by ZD strategies is always orthogonal to the vector \(DA = [0, -1, 1]\) for \(P \leq O \leq R\).

To further explain the geometrical meaning of the parameters \((O, \chi, \phi)\), we now turn to specific examples of ZD strategies with \(O = P, (T + S)/2, \) and \(R\), respectively, as illustrated in figure 1. For ZD strategies on the same line that passes through any given triplet \((q_1, q_2, q_3)\) and \((1, 1, 0)\) (that is, point D), they all
Figure 2. Strategy space of admissible ZD strategies. The plane of equalizers (see also equation (7) and figure 1(d)) separates the space of admissible ZD strategies into two regions: those with negative \( \chi \) values in (a) and those with positive \( \chi \) values in (b). The shaded region in (c) below the plane formed by ZD with \( O = P \) (more specifically, triangle ADH in figure 1(a)) and containing the corner (1, 0, 0) of the cube is inadmissible as ZD strategies.

share the same \( O \) and \( \chi \). For fixed \( O \), all the lines formed by such ZD strategies are on the same plane (figures 1(a)–(c)); when varying \( \chi \), the resulting line rotates around \((1,1,0)\) toward a direction that depends on the sign of \( \chi \) but has the exact same limiting position as \( |\chi| \to +\infty \). For increasing positive \( \chi \), the angle between this line and the vector \( \mathbf{DA} = [0, -1, 1] \) monotonically increases with \( \chi \) until reaching the limit of the ‘equalizer’ plane. For decreasing negative \( \chi \), the line monotonically approaches the same ‘equalizer’ line but from the other direction. In the following, we show this for positive \( \chi \) values. The case for negative \( \chi \) can be demonstrated analogously.

The angle \( \theta \) between the vector \( \mathbf{q}_1 - 1, \mathbf{q}_2 - 1, \mathbf{q}_3 \) and \([0, -1, 1]\) is \( \phi \)-independent and given by

\[
\cos \theta = \frac{(q_1 - 1, q_2 - 1, q_3) \cdot (0, -1, 1)}{|(q_1 - 1, q_2 - 1, q_3)| \cdot |(0, -1, 1)|} = \frac{\sqrt{2}(T - S)(\chi + 1)}{2[(R - O)^2(\chi - 1)^2 + [(T - O)\chi + O - S]^2 + [(O - S)\chi + T - O]^2]^{1/2}}.
\]

Its derivative with respect to \( \chi \) is

\[
\frac{d \cos \theta}{d\chi} = \frac{\sqrt{2}(T - S)(\chi - 1)[6O^2 - 4O(T + R + S) + 2R^2 + (T + S)^2]}{2[(R - O)^2(\chi - 1)^2 + [(T - O)\chi + O - S]^2 + [(O - S)\chi + T - O]^2]^{3/2}}.
\]

Because

\[
f(O) = 6O^2 - 4O(T + R + S) + 2R^2 + (T + S)^2 \geq f\left(\frac{T + R + S}{3}\right) = \left(\frac{2R - T - S}{3}\right)^2 > 0,
\]

we have \( d \cos \theta / d\chi < 0 \), suggesting that as \( \chi \) increases, the line formed by ZD with the same \( O \) and \( \chi \) rotates around the point \((1, 1, 0)\) toward the equalizer limit line until \( \chi \to +\infty \). For negative \( \chi \), the corresponding line similarly rotates around the point \((1, 1, 0)\) toward the equalizer limit line until \( \chi \to -\infty \).

Further, the equalizer plane separates all admissible ZD strategies into two regions: those above the plane with positive \( \chi \) values whereas those below the plane with negative \( \chi \) values (figure 2).

Most strikingly, we find the nontrivial emergence of ‘equalizers’ as the equilibrium manifold appearing in the adaptive dynamics of ZD strategies explicitly in the strategy space \((p_1, p_2, p_3)\) (figure 3(a)).
Figure 3. Adaptive dynamics of ZD strategies. For clarity, (a) only shows the three-dimensional vector field plot of equation (13) near the critical equilibrium manifold, which is the plane of equalizer strategies. Part of the equilibrium manifold, indicated by BEDJ, is stable and the remaining part, CDJ, is unstable. Different colors are used to indicate the sign of the dot product in equation (15): blue for positive and red for negative. The line DJ is the intersection of the plane \( q_1 + q_2 + q_3 - 2 = 0 \) and the equalizer plane. The vector field \( \frac{dq_1}{dt}, \frac{dq_2}{dt}, \frac{dq_3}{dt} \) is orthogonal to the vector \( \{q_1 - 1, q_2 - 1, q_3\} \) that points from \((1, 1, 0)\) toward \((q_1, q_2, q_3)\). (b) Shows the cooperation-enhancing region where \( \frac{dq_i}{dt} \) is positive for \( i = 1, 2, 3, 4 \) in adaptive dynamics. In this region, it is optimal for ZD strategies to increase their probabilities to cooperate invariably, after each of the four outcomes.

Assuming ZD strategies \( \mathbf{p} = [p_1, p_2, p_3, p_4] \) vs \( \mathbf{q} = [q_1, q_2, q_3, q_4] \), both satisfying equation (3), Press and Dyson show that the average payoff \( \pi(\mathbf{p}, \mathbf{q}) \) of player using \( \mathbf{p} \) can be calculated as

\[
\pi(\mathbf{p}, \mathbf{q}) = \det \begin{vmatrix} p_1q_1 - 1 & p_1 - 1 & q_1 - 1 & R \\ p_2q_3 & p_2 - 1 & q_3 & S \\ p_3q_2 & p_3 & q_2 - 1 & T \\ p_4q_4 & p_4 & q_4 & P \end{vmatrix} \]

The adaptive dynamics of ZD strategies can be obtained by

\[
\frac{dq_i}{dt} = \frac{\partial\pi(\mathbf{p}, \mathbf{q})}{\partial p_i} \bigg|_{p=q}.
\]

After a bit of algebra, we get

\[
\begin{align*}
\frac{dq_1}{dt} &= \frac{(q_2 + q_3 - 1)(T - S)q_1 - (R - S)q_2 - (T - R)q_3 - (T - R)}{(1 - q_2 + q_3)(1 - 2q_1 + q_2 + q_3)^2}, \\
\frac{dq_2}{dt} &= \frac{1 - q_1}{q_2 + q_3 - 1} \frac{dq_1}{dt}, \\
\frac{dq_3}{dt} &= \frac{dq_2}{dt}.
\end{align*}
\]

We see that the selection gradient is zero at the equalizer plane, forming the equilibrium manifold of the corresponding adaptive dynamics (figure 3(a)).

A simple calculation shows that

\[
(q_1 - 1)\frac{dq_1}{dt} + (q_2 - 1)\frac{dq_2}{dt} + q_3\frac{dq_3}{dt} = 0.
\]
Thus, the direction of the vector field \([dq_i/dt, dq_j/dt, dq_k/dt]\) is orthogonal to the vector \([q_1 - 1, q_2 - 1, q_3]\) that points from \((1, 1, 0)\) to \((q_1, q_2, q_3)\). As a matter of fact, the vector field is formed by the concentric spheres around the corner \((1, 1, 0)\).

### 2.1. Stability of the equilibrium manifold

The normal vector of the equalizer plane is \([T - S, -(R - S), -(T - R)]\). Its dot product with the selection gradient \([dq_i/dt, dq_j/dt, dq_k/dt]\) near the equilibrium manifold would be

\[
[T - S, -(R - S), -(T - R)] \cdot \begin{bmatrix} \frac{dq_1}{dt} \\ \frac{dq_2}{dt} \\ \frac{dq_3}{dt} \end{bmatrix} = \frac{(T - S)(q_1 + q_2 + q_3 - 2) - (T - S)q_1 - (R - S)q_2 - (T - R)q_3 - (T - R)}{(1 - q_1 + q_2 + q_3)^2}.
\]

The sign of this dot product can be used to determine the local stability of the equilibrium manifold. In particular, the plane \(q_1 + q_2 + q_3 - 2 = 0\) intersects with the equalizer plane and divides it into stable and unstable parts (figure 3(a)).

### 2.2. Cooperation-enhancing region

In the full strategy space, the cooperation-enhancing region can be found by requiring \(\frac{dq_i}{dt} > 0\) for \(i = 1, 2, 3, 4\). Using equation (3), we get

\[
\frac{dq_i}{dt} = \frac{1}{2R - S - T} \left[ -(T + S - 2P) \frac{dq_1}{dt} + (R - P) \left( \frac{dq_2}{dt} + \frac{dq_3}{dt} \right) \right] = \frac{2(R - P)(1 - q_1) - (T + S - 2P)(q_2 + q_3 - 1)}{(2R - T - S)(q_2 + q_3 - 1)} \frac{dq_1}{dt}.
\]

We thus obtain the critical plane \(2(R - P)(1 - q_1) - (T + S - 2P)(q_2 + q_3 - 1) = 0\) so as to determine the sign of \(dq_i/dt\). This plane is in fact the O-fixed plane (see equation (6)) formed by ZD strategies with the specific \(O\) value satisfying \((T + S)/2 < O < R\) and

\[
\frac{2O - T - S}{R - O} = \frac{2(R - P)}{T - S - 2P}.
\]

Hence the cooperation-enhancing region (as shown in figure 3(b)) is given by

\[
q_2 + q_3 - 1 > 0,
\]

\[
(T - S)q_1 - (R - S)q_2 - (T - R)q_3 - (T - R) > 0,
\]

\[
2(R - P)(1 - q_1) - (T + S - 2P)(q_2 + q_3 - 1) > 0.
\]

### 2.3. Equal gains from switching: \(T + S = R + P\)

A necessary condition for reactive strategies to be ZD strategies is that the four payoff values need to satisfy the so-called equal gains from switching: \(T + S = R + P\). Under this condition, we may further get \(q_1 = q_3\) and \(q_2 = q_4\).

By letting \(q_1 = q_3\) and \(q_2 = q_4\), we obtain \(\phi = 1/[(R - S)\chi + T - R]\) and \(\phi = 1/[(T - P)\chi + P - S]\), respectively. Moreover, equating these two \(\phi\) values yields

\[
\frac{(T + S - R - P)(\chi - 1)}{[(R - S)\chi + T - R][(T - P)\chi + P - S]} = 0.
\]

Hence, the payoff condition \(T + S = R + P\) is required as desired. To obtain the reactive strategy \([q_1, q_2, q_3, q_4]\) properly from the superset of ZD strategies, we also need to employ the following combinations of \(O, \chi\):

\[
O = \frac{(T + S)(1 - q_1) + R(q_1 + q_2 - 1)}{1 - q_1 + q_2},
\]

\[
\chi = \frac{R - S - (T - R)(q_1 - q_2)}{(R - S)(q_1 - q_2) - (T - R)}.
\]

Under this payoff structure condition, another interesting finding is that unconditional strategies \((q_1 = q_2 = q_3 = q_4)\) are ZD strategies with negative \(\chi\). Using equation (3) and forcing \(q_1 = q_2 = q_3\), we get
2.4. Fun facts about ZD strategies

The term $2q_1 - q_2 - q_3 - 1$ is the denominator of the expression of $O$ as a function of $(q_1, q_2, q_3)$ in equation (5). For all admissible ZD strategies, we actually have

$$2q_1 - q_2 - q_3 - 1 < 0.$$  

(22)

Algebraically, using equation (2), we get $2q_1 - q_2 - q_3 - 1 = -(2R - T - S)\phi(\chi - 1) < 0$ for any admissible parameter choices. The same inequality can also be shown by utilizing the exquisite geometry of ZD strategies. As $O \to -\infty$, the $O$-fixed plane approaches the limit $q_1 = 2q_1 - q_2 - 1$, which is further below the plane for $O = P$. Likewise, as $O \to +\infty$, the $O$-fixed plane rotates further beyond the plane corresponding to $O = R$ (whose equation is $q_1 = 1$) and approaches the same limit from the other direction. Since a ZD strategy only admits $P \leq O \leq R$, it follows that $q_3 > 2q_1 - q_2 - 1$ for all ZD strategies. Hence we obtain the inequality above as desired.

3. Conclusions

In this paper, we focus on the payoff structure of conventional PD games satisfying $T + S > 2P$. It is not impossible that in some PD games $T + S < 2P$ holds, that is, the polygon connecting $(R, R), (S, T), (P, P)$, and $(T, S)$ is non-convex in the pairwise payoff plot [42]. Under this circumstance of more adversarial nature, the geometry of ZD strategies can be studied similarly. It is also worth noting that, for games satisfying ‘equal gains from switching’, $T + S = R + P$, reactive strategies (including unconditional strategies) become special cases of ZD strategies.

Scrutinizing the geometry of ZD strategies provides useful intuition about and insights into how ZD strategies relate to one another. The parameterization combination $(O, \chi, \phi)$ of ZD strategies determine the admissible range of $q = [q_1, q_2, q_3, q_4]$ that can be realized. Each subset of ZD strategies with a fixed $O$ value contains a common extreme boundary AD (figures 1(a)–(c)) that is a linear interpolation of TFT ($q = [1, 0, 1, 0]$) and ‘AllC or AllD’ ($q = [1, 1, 0, 0]$). The line formed by ZD strategies sharing the same $(O, \chi)$ approaches the limit of equalizer strategies, located on the plane BCDE (figure 1(d)), as $\chi \to +\infty$ or $\chi \to -\infty$. This equalizer plane in fact separates the strategy space $(q_1, q_2, q_3)$ of admissible ZD strategies into two regions with positive versus negative $\chi$ values (figures 2(a) and (b)). We also realize that the region of the cube under the plane expanded by ZD strategies with $O = P$ and containing the corner $(1, 0, 0)$ is inadmissible for any ZD strategies (figure 2(c)).

From the perspective of evolution, it is straightforward to use the adaptive dynamics of ZD strategies to figure out the potential optimal ZD strategies against each other. We find that the common line AD of all ZD strategies with $\chi \to 1$ emerges as a particular equilibrium. More generally, the plane of equalizers is the non-trivial equilibrium manifold, part of which is stable whereas the remaining part is not (figure 3(a)). The same plane acts as a separatrix of the cooperation-enhancing region that encourages increases in cooperation for each $q_i$ where $i = 1, 2, 3, 4$ (figure 3(b)).

In sum, this paper reveals the non-trivial and elegant geometry of ZD strategies that is previously overlooked. Our present work offers a geometrical interpretation of the ZD parameters $(O, \chi, \phi)$ and particularly the unprecedented geometrical relationships of ZD strategies in the entire strategy space of memory-one IPD strategies with $q = [q_1, q_2, q_3, q_4]$. Most interestingly, we find that the subset of equalizer strategies forms the critical plane in the strategy space that emerges as the critical equilibrium manifold in the adaptive dynamics of ZD strategies and also as a separatrix of the cooperation-enhancing region. These results highlight the previously unforeseen connection between equalizers and general ZD strategies.

$q_4 = q_1 + (T + S - R - P)/(2R - T - S)$, which leads to the same ‘equal gains from switching’, $T + S = R + P$, in order for $q_4 = q_1$. The corresponding choice of $O, \chi, \phi$ are now

$$O = (2R - T - S)q_1 + T + S - R,$$

$$\chi = -\frac{R - S}{T - R},$$

$$\phi = -\frac{T - R}{(T - S)(2R - T - S)}.$$  

(21)
Acknowledgments

XC gratefully acknowledges the generous faculty startup fund provided by BUPT (No. 505022023). LW receives financial support from the National Natural Science Foundation of China (Grant No. 62036002). FF is grateful for support from the Bill & Melinda Gates Foundation (Award No. OPP1217336), the NIH COBRE Program (Grant No. 1P20GM130454), the Neukom CompX Faculty Grant, the Dartmouth Faculty Startup Fund, and the Walter & Constance Burke Research Initiation Award.

Ethics

Ethical assessment is not required prior to conducting the research reported in this paper, as the present study does not have experiments on human subjects and animals, and does not contain any sensitive and private information.

Conflict of interest

We declare that we have no competing interests.

Data and code availability statement

All the data and analyses pertaining to this work have been included in the main text. Source code for reproducing the results is available at the Github repository (https://github.com/fufeng/ZDGeometry).

Appendix A

The symbolic coordinates of the points in all the figures are given below. Here, 

\[ d_k = (T - P)^2 + (R - P)^2 + (R - S)^2 - 2(P - S)^2 \]

\begin{align*}
A: & (1, 0, 1) & B: & \left(\frac{2(T - R)}{T - S}, 0, 1\right) & C: & \left(1, \frac{2R - T - S}{R - S}, 1\right) \\
D: & (1, 1, 0) & E: & \left(\frac{T - R}{T - P}, 0, \frac{P - S}{T - P}\right) & F: & (1, 1, 1) \\
G: & \left(\frac{T - R}{T - S}, 0, 0\right) & H: & \left(\frac{T + S - R - P}{T + S - 2P}, 0, 0\right) & I: & (0, 1, 0) \\
J: & \left(\frac{2T - R - S}{T + R - 2S}, 2R - T - S, 1\right) & K: & \left(1 - \frac{(T + S - 2P)(2R - T - S)}{d_k}, \frac{2(R - P)(2R - T - S)}{d_k}, 1\right)
\end{align*}

In particular, if we consider the conventional PD game with payoff matrix \([R, S, T, P] = [3, 0, 5, 1]\), we further get the following specific numbers for all the figures.

\begin{align*}
B: & \left(\frac{4}{5}, 0, 1\right) & C: & \left(1, \frac{1}{5}, 1\right) & E: & \left(\frac{1}{2}, 0, \frac{1}{4}\right) & G: & \left(\frac{2}{5}, 0, 0\right) \\
H: & \left(\frac{1}{3}, 0, 0\right) & J: & \left(\frac{7}{8}, \frac{1}{8}, 1\right) & K: & \left(\frac{8}{9}, \frac{4}{27}, 1\right)
\end{align*}

ORCID iDs

Xingru Chen https://orcid.org/0000-0001-9131-3798
Long Wang https://orcid.org/0000-0001-5600-8157
Feng Fu https://orcid.org/0000-0001-8252-1990

References

[1] Axelrod R and Hamilton W D 1981 The evolution of cooperation Science 211 1390–6
[2] Nowak M A 2006 Five rules for the evolution of cooperation Science 314 1560–3
[3] Trivers R L 1971 The evolution of reciprocal altruism Q. Rev. Biol. 46 35–57
[4] Boyd R and Lorberbaum J P 1987 No pure strategy is evolutionarily stable in the repeated prisoner’s dilemma game Nature 327 58–9
[5] Nowak M et al 1990 An evolutionarily stable strategy may be inaccessible J. Theor. Biol. 142 237–41
[6] Nowak M A and Sigmund K 1992 Tit for tat in heterogeneous populations Nature 355 250–3
[7] Nowak M and Sigmund K 1993 A strategy of win–stay, lose–shift that outperforms tit-for-tat in the prisoner’s dilemma game Nature 364 36–8
[8] Hauert C and Schuster H G 1997 Effects of increasing the number of players and memory size in the iterated prisoner’s dilemma: a numerical approach Proc. R. Soc. B 264 513–9
[9] Fudenberg D and Maskin E 2009 The folk theorem in repeated games with discounting or with incomplete information A Long-Run Collaboration on Long-Run Games (Singapore: World Scientific) pp 209–30
[10] Wu T, Fu F and Wang L 2018 Coevolutionary dynamics of aspiration and strategy in spatial repeated public goods games New J. Phys. 20 063007
[11] Nowak M 1990 Stochastic strategies in the prisoner’s dilemma Theor. Popul. Biol. 38 93–112
[12] Axelrod R 2012 Launching ‘the evolution of cooperation’ J. Theor. Biol. 299 21–4
[13] Jusup M et al 2022 Social physics Phys. Rep. 948 1–148
[14] Hilbe C, Martinez-Vaquero I A, Chatterjee K and Nowak M A 2017 Memory-n strategies of direct reciprocity Proc. Natl Acad. Sci. USA 114 4715–20
[15] Rubinstein A 1986 Finite automata play the repeated prisoner’s dilemma J. Econ. Theory 39 83–96
[16] Sandholm T W and Crites R H 1996 Multiagent reinforcement learning in the iterated prisoner’s dilemma BioSystems 37 147–66
[17] Harper M, Knight V, Jones M, Koutsouvelis G, Glytsi N E and Campbell O 2017 Reinforcement learning produces dominant strategies for the iterated prisoner’s dilemma PLoS One 12 e0188046
[18] Hilbe C, Nowak M A and Sigmund K 2013 Evolution of extortion in iterated prisoner’s dilemma games Proc. Natl Acad. Sci. USA 110 6913–8
[19] Back S K, Jeong H-C, Hilbe C and Nowak M A 2016 Comparing reactive and memory-one strategies of direct reciprocity Sci. Rep. 6 25676
[20] Sigmund K 2010 The calculus of selfishness The Calculus of Selfishness (Princeton, NJ: Princeton University Press)
[21] Boerlijst M C, Nowak M A and Sigmund K 1997 Equal pay for all prisoners Am. Math. Mon. 104 303–5
[22] Press W H and Dyson F J 2012 Iterated prisoner’s dilemma contains strategies that dominate any evolutionary opponent Proc. Natl Acad. Sci. USA 109 10409–13
[23] Stewart A J and Plotkin J B 2013 From extortion to generosity, evolution in the iterated prisoner’s dilemma Proc. Natl Acad. Sci. USA 110 15348–53
[24] McAvoy A and Hauert C 2016 Autocratic strategies for iterated games with arbitrary action spaces Proc. Natl Acad. Sci. USA 113 3573–8
[25] Ueda M 2021 Memory-two zero-determinant strategies in repeated games R. Soc. Open Sci. 8 202186
[26] Knight V A, Harper M, Glytsi N E and Gillard J 2019 Recognising and evaluating the effectiveness of extortion in the iterated prisoner’s dilemma (arXiv:1904.00973)
[27] Govaert A and Cao M 2020 Zero-determinant strategies in repeated multiplayer social dilemmas with discounted payoffs IEEE Trans. Autom. Control 66 4575–88
[28] Hilbe C, Wu B, Traulsen A and Nowak M A 2014 Cooperation and control in multiplayer social dilemmas Proc. Natl Acad. Sci. USA 111 16425–30
[29] Pan L, Hao D, Rong Z and Zhou T 2015 Zero-determinant strategies in iterated public goods game Sci. Rep. 5 13096
[30] Hilbe C, Wu B, Traulsen A and Nowak M A 2015 Evolutionary performance of zero-determinant strategies in multiplayer games J. Theor. Biol. 374 113–24
[31] Chen F, Wu T and Wang L 2022 Evolutionary dynamics of zero-determinant strategies in repeated multiplayer games J. Theor. Biol. 549 111209
[32] Hao D, Rong Z and Zhou T 2015 Extortion under uncertainty: zero-determinant strategies in noisy games Phys. Rev. E 91 052803
[33] Mamiya A and Ichinose G 2020 Zero-determinant strategies in iterated public goods game Phys. Rev. E 102 032115
[34] Ichinose G and Masuda N 2018 Zero-determinant strategies in finitely repeated games J. Theor. Biol. 438 61–77
[35] Hilbe C, Nowak M A and Traulsen A 2013 Adaptive dynamics of extortion and compliance PLoS One 8 e77886
[36] Szolnoki A and Perc M 2014 Evolution of extortion in structured populations Phys. Rev. E 89 022804
[37] Adami C and Hintze A 2013 Evolutionary instability of zero-determinant strategies demonstrates that winning is not everything Nat. Commun. 4 2193
[38] Chen J and Zinger A 2014 The robustness of zero-determinant strategies in iterated prisoner’s dilemma games J. Theor. Biol. 357 46–54
[39] Hilbe C, Traulsen A and Sigmund K 2015 Partners or rivals? Strategies for the iterated prisoner’s dilemma Games Econ. Behav. 92 41–52
[40] Akin E 2016 The iterated prisoner’s dilemma: good strategies and their dynamics Ergod. Theor. Adv. Dynam. Syst. 77–107
[41] Szolnoki A and Perc M 2014 Defection and extortion as unexpected catalysts of unconditional cooperation in structured populations Sci. Rep. 4 5496
[42] Chen X and Fu F 2022 Outlearning extortioners by fair-minded unbending strategies (arXiv:2201.04198)