EVOLUTION IN ISOLATED POPULATIONS

- PRELIMINARY -

TYMON TATUR
BONN

Abstract. We consider a class of stochastic evolutionary dynamics in which a finite, spatially structured population plays a symmetric 2-player normal-form game with a finite set of actions. The selection process is stochastic: the higher the payoff of a player at a given location, the higher the chance that some offspring of that player will replace players at other locations. In addition, each period, with probability ε a player "mutates" and starts using a strategy that is randomly drawn according to some fixed distribution λ.

In this standard framework, the long run properties of the dynamic depend on the distribution λ, even if we restrict attention to λ with full support and consider the limit as ε goes to zero. We therefore consider an extended framework where there are many populations identical to the one described above. Populations are isolated in the sense that only with a tiny probability η a player from one population replaces a player from another population. In this framework, we impose a natural evolutionary requirement on λ. Theorem 1 states that a distribution λ satisfying our criterion always exists.

For the case where selection is weak, Theorem 2 provides a precise characterization of the strategies used in the long run if λ satisfies our criterion. A single parameter captures population structure.

As an application, we consider the theory of finitely repeated games and obtain results that are in contrast with standard predictions based on Nash Equilibrium and sub-game perfect equilibrium.

Email: tatur@uni-bonn.de. Mailing address: Bonn University, Department of Economics, Adenauerallee 24-42, 53113 Bonn, Germany.

1. Introduction

We consider a class of stochastic evolutionary dynamics in which a finite, spatially structured population P plays some symmetric 2-player game G with a finite set of strategies. Individuals who receive higher payoffs have on average more offspring. Each period, with some probability ε one of the players in the population "mutates" and switches to a new strategy which is randomly drawn according to some distribution λ.

The assumption that natural selection is stochastic is natural in a model where the population is finite. In general this, however, leads to a problem. The long run behavior of the dynamic will depend on the distribution λ. If natural selection is truly stochastic, this will unfortunately remain true even if we consider the limit as ε goes to zero.[1]

As a result, a number of recent papers in evolutionary biology (see Tarnita et al. 2009, Antal et al. 2009, Tarnita et al. 2011) just assumes that λ is the uniform distribution which puts equal weight on each of the n strategies of the game G and calls a strategy s favored by the evolutionary dynamic if, in the long run, for this particular λ, the strategy s is used with probability greater than $\frac{1}{n}$.

We see three major reasons why this approach is problematic. Firstly, the fact that a strategy is used by a fraction of the population which is larger than $\frac{1}{n}$ has weak predictive power. For example, for the case of three strategies, it could very well be that the "favored strategy" is used by 34% of the population while the other two strategies are each used by 33% of the population.[2] Secondly, the whole approach is based on the idea that λ puts exactly equal probabilities on all strategies, independently of the payoff structure. This is hard to justfy.[3] Thirdly, results may depend strongly on modeling choices that in most game-theoretic models would be irrelevant.[4]

---

[1] This is different than in models like those proposed by Young (1993), Kandori et al. (1993), and Ellison (1993) in which the best response dynamic - which can be seen as the analog of the selection process - is essentially deterministic.

[2] This is not just a theoretical possibility. As a matter of fact, under the limit where selection gets weak, the long-run expected fraction of the population using a strategy s will converge to $\lambda(s)$, so for a λ which is uniform, each strategy in the limit will be played with probability $\frac{1}{n}$.

[3] Tarnita et al. (2009) justify the fact that λ puts probability $\frac{1}{2}$ on both strategies (they consider 2x2 games) as follows: "Relabeling the two strategies and correspondingly swapping the entries of the payoff matrix must yield symmetric dynamics. This assumption is entirely natural. It means that the difference between A and B is fully captured by the payoff matrix, while the population structure and update rule do not introduce and additional difference between A and B." We full-heartedly agree that a relabaling of strategies and swapping the entries of the payoff matrix must yield symmetric dynamics. However, this only implies that λ puts probability $\frac{1}{2}$ on both strategies if λ is not allowed to depend on the payoff matrix. Tarnita et al. (2009) provide no explanation why it is natural to assume that λ does not depend on the payoff matrix.

[4] For example, if we substitute a strategy with two different strategies that are absolutely identical this may affect which strategy is favored.

In this paper we explore a different approach. We ask what distributions λ will be selected by evolution if all λ are technologically feasible. To address this question, we consider an extended model where our population is one out of many similar. Populations are isolated in the sense that only with a small probability η a player from one population replaces a player from another population. In this framework we impose a natural evolutionary criterion on λ for the case where the the mutation rate ε is small and the probability η even smaller. We see this model as appropriate when thinking about evolution of species which live in isolated small populations. Examples would be species who live in isolated sub-populations because of natural barriers, for example, species living in lakes, puddles, islands, isolated patches of forest, next to dispersed human settlements, etc. Perhaps the model could also apply to hunter-gatherer societies whose members live in small isolated groups since larger groups are unsustainable in terms of food gathering technology.

Theorem 1, our first result, shows that strategies satisfying our requirement always exist for the broad class of natural selection processes and spatially structured populations considered in this paper. From the literature which assumes an exogenously given λ, it is well known that the spacial structure can effect the strategies played in the long run. This is also the case in our framework. To get sharp predictions we therefore consider the case where selection is weak.[5] In Theorem 2 we prove that in this case the long run probability that a player uses strategy s does not depend on his location and is given by a probability distribution $\nu \in \Delta S$ such that that for any strategy $s' \in S$:

$$\sum_{s \in S} \nu(s) \cdot (k \cdot (u(s,s) - u(s',s')) + (u(s,s') - u(s',s))) \geqslant 0,$$

where u is the payoff function of the game and $k \geqslant 0$ is a constant that depends on the local interaction structure of the population but not the considered game. Note, that this result holds for arbitrary symmetric 2-player games with a finite set of strategies.

Since symmetric 2x2 games with two strict Nash equilibria have been widely studied in the literature let us consider the implication of Theorem 2 for such games. Consider a fixed population structure and let k be the constant from Theorem 2 for that population structure. If we denote the two strategies in the symmetric 2x2 game by s^1 and s^2 then Thoerem 2 predicts that if

$$k \cdot (u(s^1,s^1) - u(s^2,s^2)) + (u(s^1,s^2) - u(s^2,s^1))) > 0,$$

---

[5] Weak selection means that individuals have very similar fitness. See Wild and Traulsen (2007) for a discussion of different ways to introduce weak selection and conditions under which these are equivalent. As we explain in section 4, in our framework weak selection can also be seen as a provider of a link between evolutionary game theory in which payoffs correspond to "fitness" and standard game theory in which affine transformations of payoffs do not affect predictions.

then strategy s^1 is played with probability 1 as selection gets weak and the mutation rate goes to zero. If the inequality goes in the other direction, i.e. the left hand side is smaller than zero, then s^2 is played with probability 1 as selection gets weak and the mutation rate goes to zero. Let us contrast this with Tarnita et al. (2009) who consider 2x2, also allow general population structures, and also consider the limit as selection becomes weak, but assume that λ is exogenously given and puts probability $\frac{1}{2}$ on both strategies. In their model, in the limit as selection gets weak, both strategies s^1 and s^2 are used with probability $\frac{1}{2}$. This is true both for a fixed mutation rate ε and $\varepsilon \to 0$. Their main result is that, for any ε, if selection is sufficiently weak, the strategy for which

$$\sigma \cdot (u(s^1, s^1) - u(s^2, s^2)) + (u(s^1, s^2) - u(s^2, s^1))) > 0$$

(holds will be played with probability larger than $\frac{1}{2}$, where σ is a parameter which plays an analogous role as the parameter k in our model.[6] Note that since, in their framework, in the limit as selection gets weak both strategies are used with probability $\frac{1}{2}$ and in our framework the probability with which s^1 is used converges either to 0 or 1 for generic games the two models yield predictions which are very different.

We use Theorem 2 to show how our model can be applied to the study of finitely repeated games. We first consider the finitely repeated prisoner's dilemma. The T-times repeated prisoner's dilemma is interesting since it is a game where classic game theory is able to make a clear prediction: the only outcome consistent with Nash equilibrium is one where both players deviate in each of the T periods.

This clear cut prediction is in contrast with the prediction of our model. Of course, the exact long run distribution will depend on the spatial structure of the population which is captured by the parameter k. However, we show that, for any fixed interaction structure, our model predicts that average payoffs in the T-times repeated prisoner's dilemma as a fraction of payoffs which players would achieve if they cooperated each period, converges to 1 as T goes to infinity. Theorem 3 generalizes this observation to more general games. Thus, also for finitely repeated games where folk theorems (see, for example, Benoit and Krishna 1986) predict that "all is possible" our model predicts behavior that, in a sense[7], is close to efficiency.

The paper is organized as follows. In section 2 we introduce the model. In section 3 we introduce the notion of an optimal mutation strategy and prove existence. Section 4 considers the limit where selection is weak. Section 5 considers 2x2 games as an example. Section 6 applies the model to finitely repeated games. Section 7 concludes. All proofs are in the Appendix.

---

[6] Since they consider fixed mutation rates ε, their k will depends on the mutation rate ε.

[7] While the average payoff as a fraction of the highest possible payoff goes to 1, the average payoff is still bounded away from the highest possible payoff.

2. BASIC MODEL

Let G be a symmetric 2-player normal form game with a finite strategy set S and payoff function u.

There are N individuals living in locations numbered from 1 to N. In any time period t the payoff of the individual living in location i is given by:

$$U_i^t = \sum_j w_{ij} \cdot u(s_i^t, s_j^t),$$

where $s_j^t \in S$ is the strategy used by the individual in location j in time period t and w_{ij} are weights that are non-negative and have the property that for each location $i \in \{1, .., n\}$ there are two other locations j and k such that $w_{ij} > 0$ and $w_{ik} > 0$.

Each period $t = 1, 2, 3, ...$, the individual at location j is replaced with an offspring of the individual at location i with probability

$$p_{i,j}(U_1^t, ..., U_N^t),$$

where $p_{i,j} : \mathbb{R}^n \to [0, 1]$ are continuously differentiable functions such that (i) $\frac{\delta p_{i,j}}{\delta U_i} > 0$, (ii) $\frac{\delta p_{i,j}}{\delta U_j} < 0$, and (iii)$\frac{\delta p_{i,j}}{\delta U_k} \leqslant 0$ for $k \neq i, j$,[8] and, for any $(U_1, .., U_N) \in \mathbb{R}^n$,

$$\sum_i p_{ij}(U_1, .., U_N) = 1.$$

All offspring of a player uses the same strategy as that player. Of course, we can think about reproduction being either cultural or genetic. The assumptions (i)$\frac{\delta p_{i,j}}{\delta U_i} > 0$, (ii) $\frac{\delta p_{i,j}}{\delta U_j} < 0$, and (iii)$\frac{\delta p_{i,j}}{\delta U_k} \leqslant 0$ correspond to the idea that an individual in location j is more likely to be replaced by the offspring of an individual at location i if (i) the payoff of the individual at location i is high (ii) the payoff of the individual at location j low and (iii) the payoff of other individuals competing for the spot is low.[9]

In addition, after reproduction took place, a coin is thrown and with probability $\varepsilon \in (0, 1)$ a player at a randomly drawn[10] location "mutates". A player who "mutates" switches to a new strategy that is drawn according to a fixed probability distribution $\lambda \in \Delta S$. Since λ specifies what strategies a mutant chooses we will call it the *mutant strategy*. The probability ε will be called the *mutation rate*.

---

[8] This requirement can be weakened. In particular, it is enough to assume that (i) for any two locations$i \neq j$, either $p_{i,j} \equiv 0$ or $p_{i,j} : \mathbb{R}^n \to [0, 1]$ is continuously differentiable with $\frac{\delta p_{i,j}}{\delta U_i} > 0$, $\frac{\delta p_{i,j}}{\delta U_j} < 0$,$\frac{\delta p_{i,j}}{\delta U_k} \leqslant 0$ and (ii) for any two locations $i \neq j$, there exists a sequence $k_1, ..., k_l$ with $k_1 = i$ and $k_l = j$ such that $p_{k_r, k_{r+1}} \not\equiv 0$ for $r = 1, 2, .., l - 1$.

[9] Note that the assumption that p_{ij} is strictly increasing in U_i implicitly implies that $p_{ij} > 0$.

[10] We assume that the location is drawn according to some fixed distribution. We do not need to assume that the distribution has full support or puts equal probability on all locations.

Consider a fixed mutant strategy λ and mutation rate ε. Clearly, the profile of strategies $s^t = (s_1^t, ..., s_N^t)$ follows a time-homogenous Markov chain with state space S^N. Since the strategy set S of the game G is finite so is S^N. We will denote the transition matrix of this Markov chain by $(P^{\varepsilon}_{ss'})_{s,s' \in S^N}$. From the fact that all the probabilities $p_{i,j}(U_1^t, ..., U_N^t)$ are greater than zero for any $(U_1, .., U_N) \in \mathbb{R}^n$ and that the probability ε is also greater than zero, we can conclude that the process has a single recurrent class equal to $Support(\lambda)^N \subset S^N$. The restriction of the process to this recurrent class is irreducible and aperiodic. Therefore the process possesses a unique stationary distribution $\mu^{\varepsilon}_{\lambda}$.

Remark 1. Note that our environment is rather general. In particular, it allows for situations where players live on some arbitrary graph and their payoffs are only affected by the behavior of their neighbors. It also allows for situations in which sending offspring to locations k may be easier from location j than from some other location i, for example, the birth-death and death-birth processes on graphs considered in Ohtsuki et al. (2006) and generalizations of the Moran process in the spirit of Taylor et al. (2004) or Novak et al. (2004). Finally, it allows for different locations to differ in quality. For example, it can be that one location i is better than location j in the sense that $p_{ki} > p_{kj}$ for all k.

Remark 2. We introduced our dynamic as one involving natural selection and mutations. An alternative interpretation is one where we thing about players imitating each other (player j imitates i with a probability p_{ij}) and occasionally experimenting in which case they choose a strategy drawn randomly according to λ.

Consider a distribution λ with full support. The family of Markov chains $(P^{\varepsilon}_{ss'})_{s,s' \in S^N}$ for different mutation rates ε is a regular perturbed Markov process in the sense, in which this term is used by Young (1998) and, therefore, the limit

$$\mu_{\lambda} = \lim_{\varepsilon \to 0} \mu^{\varepsilon}_{\lambda}$$

exists. The support of μ_{λ} can only contain states that are in a recurrent class of the unperturbed dynamic $(P^0_{ss'})_{s,s' \in S^N}$. Note that the recurrent classes of $(P^0_{ss'})_{s,s' \in S^N}$ correspond exactly to singletons $\{(s, s, ..., s)\}$ where the state $s \in S$. Note that this means that the probability of a player at a given location i playing a given strategy $s \in S$ does not depend on the location i.

Since only a single mutation is needed to go from one such state to another, standard tree-surgery methods (see Young (1993,1998)) immediately imply that

$$Support(\mu_{\lambda}) = \{(a, ..., a) \in S^N \ : \ a \in S\}.$$

Indeed, this follows from the observation that only a single mutation is needed to move from one recurrent class of the unperturbed process to another.

In other words, as long as λ has full support, all strategies $a \in S$ will be used in the long run with positive probabilities, even in the limit as ε becomes small.[11]

Given that, for λ with full support, all strategies are used both for fixed ε and $\varepsilon > 0$ some recent papers concerned with 2x2 games (see, for example, Tarnita et al 2009) assume that λ puts equal weight on both strategies and a strategy is said to be favored by the evolutionary dynamic if, in the long run, it is used with probability greater $\frac{1}{2}$.[12] This approach has also been generalized to the case of $n > 2$ strategies: λ is assumed to put equal probability on each of the n strategies and a strategy is said to be favored by selection if in the long run it is used by a random individual with probability greater than $\frac{1}{n}$ and otherwise said to be opposed by selection (see Antal et al. 2009, Tarnita et al. 2011).

We see three major reasons why this approach is problematic. Firstly, the fact that a strategy is used by the majority of the population has weak predictive power. Even for the case of two strategies it could very well be that the "favored strategy" is used by 51% of the population while the other strategy is used by 49% of the population. For more strategies predictions become even weaker since the majority of the population may use different strategies than the favored one. Secondly, the whole approach is based on the idea that λ puts equal probability on all strategies. This may be appearing if the game is very symmetric in its structure but otherwise often appears unnatural. Thirdly, results depend strongly on modeling choices that in most game theoretic models are irrelevant. For example, if we substitute a strategy with two different strategies that are identical in terms of payoffs and effects on others (but perhaps differ in some characteristics which are natural to include in a larger context) this may affect which strategy is favored.

This paper considers a fundamentally different approach. The main idea is to endogenize λ and thereby get clear cut predictions, in particular, obtain the exact distribution of strategies that are used as the mutation rate ε becomes small.

3. Extended Model

Consider the following extension of the basic model from the last section. There is now a countable set of populations, each of which is identical to the population P from the previous model. The set of all populations will be denoted by $\mathscr{P}$. In addition, the behavior of each player in period t is now characterized by a pair (s_i^t, λ_i^t) where $s_i^t \in S$ is the strategy used when interacting with others and $\lambda_i^t \in \Delta S$ is the distribution from which strategies are drawn when mutating.

---

[11] A stronger statement can be easily shown: for any $\mu \in \Delta S$ there is a $\lambda \in \Delta S$ such that $\mu = \mu_\lambda$.

[12] For certain spatial structures, like for example players living on a circle and interacting only with closest neighbours limits as the population size gets large can also be naturally formed.

As before, each period, the probabilities p_{ij} are used to determine whether the player at location j gets replaced by an offspring of player the player at location i. An offspring inherits both s_i^t and λ_i^t from his parent. As before, with probability ε a mutation takes place in which case a new strategy is drawn, but now the individual at location i used his individual mutation strategy λ_i^t. In addition, after that, each period a coin is thrown and with probability η two populations P and P' are randomly drawn (both according to some fixed distribution $G \in \Delta\mathscr{P}$ with full support) and a randomly selected player in population P' is replaced by the offspring of a randomly selected player from population P.[13]

Note that the above defines a time-homogenous Markov process with state space $(S \times \Delta S)^{\mathscr{P} \times N}$.

Fix ε and η and consider an initial state of the dynamic in which all players use the same mutation strategy $\lambda^* \in \Delta S$ except for one player who uses some $\lambda' \neq \lambda^*$. Let $E^{\varepsilon,\eta}$ be the probability that individuals using λ' become extinct, i.e. that there exists a time period t in which all players use λ^*.[14]

We will say that a mutation strategy $\lambda^* \in \Delta S$ is *evolutionary optimal* or *evolutionary stable* if and only if

$$\varliminf_{\varepsilon \to 0} \varliminf_{\eta \to 0} E^{\varepsilon,\eta} = 1$$

for any $\lambda' \neq \lambda^*$ and any initial state of the above form.

Remark 3. Note that the model presented here is quite different from "haystack models" of group selection in which individuals form groups, live in those groups for a while, then are all mixed, then groups are formed again, and so on. We consider a situation in which individuals living in different populations (perhaps on different islands) rarely interact - there is no periodic "mixing" and formation of new "haystacks".

Theorem 1. *The set of evolutionary optimal mutation strategies is non-empty.*

Proof. See appendix. □

Now, that we have a reasonable criterion to select mutation strategies we can return once more to the basic model and ask what happens for λ which are evolutionary optimal.

Let $\Lambda(G)$ be the set of evolutionary optimal mutation strategies for a given game G and let

$$\Pi(G) = \{\mu_{\lambda^*} : \lambda^* \in \Lambda(G)\}$$

---

[13] The probability distribution used to select a player in populations P and P' does not have to be uniform but the same distribution must be used in both populations.

[14] Note that if in some time period t all players use λ^* then also in all $t' > t$ all players will use λ^*.

be the set of stationary distributions for $\lambda \in \Lambda(G)$. Unfortunately, the set $\Pi(G)$ depends in a non-trivial way both on the population structure (the functions $p_{i,j}$) and the payoffs of the game G.[15] However, in the following section, we will obtain a very simple characterization under an additional limit.

4. Characterization for the Case where Selection is Weak

Given the generality of the model, it seems impossible to characterize evolutionary optimal λ (and the corresponding stationary distributions over strategies employed depending on location) for general games in a way which would not depend strongly on the parameters of the model.

We will therefore consider the case where selection becomes weak, i.e. where evolutionary forces are small. Weak selection has been considered as a natural limit in evolutionary biology in situations where the decision have only a small effect on overall fitness. Here, it is also a natural way to connect evolutionary game theory where payoffs correspond to fitness and where results may be affected by affine transformations with standard game theory where payoffs correspond to expected utilities and where predictions are unaffected by affine transformations.

For any game G and $\gamma > 0$ let $\gamma \cdot G$ be the game in which all payoffs are multiplied by γ.

Theorem 2. *There exists a parameter $k > 0$ that does not depend on payoffs of the game G such that*\

$$\limsup_{\gamma \to 0} \Pi(\gamma \cdot G)$$

is contained in the set of probability distributions $\mu \in \Delta S$ such that the support of μ contains only strategy profiles in which all players use the same strategy s and

$$\sum_{s \in S} \mu(s, s, .., s) \cdot (k \cdot (u(s, s) - u(s', s')) + (u(s, s') - u(s', s))) \geqslant 0$$

for any $s' \in S$.

Proof. See appendix. □

Remark 4. Note that the characterization in Theorem 2 is invariant under affine transformations of the payoffs of the game G.

Remark 5. Theorem 2 characterizes the stationary distributions (which also describe the long run behavior of the dynamics) for λ which satisfy our criterion. The reader might ask what the evolutionary optimal mutation strategies are. Note that, under the weak selection limit, $\mu_\lambda((s, s, ...s))$ converges to $\lambda(s)$ in this sense, under that limit, Theorem 2 also characterizes the evolutionary optimal λ.

---

[15] Lemma 2 in the Appendix describes this dependence.

Given that the parameter depends on the interaction structure one may wonder whether the above result can be used to make interesting predictions independently of the population structure. In section 6 we will apply our theory to finitely repeated games and show that this is indeed the case. Before we do so, however, we will quickly consider the implications of our theory for a class of games which has been studied a lot in evolutionary biology.

5. EXAMPLE: 2X2 GAMES

Since symmetric 2x2 games with two strict Nash equilibria have been widely studied in the literature we will briefly look at the implications of Theorem 2 for such games. Let G be the symmetric game given by the payoff matrix

	A	B
A	a	c
B	d	b

Consider a fixed population structure. Theorem 2 immediately predicts that if

$$k \cdot (a-b) + (c-d) > 0$$

then strategy A is played with probability 1 as selection gets weak and the mutation rate goes to zero. If

$$k \cdot (a-b) + (c-d) < 0$$

then B is played with probability 1 under the same limit.

This allows us to make predicitions which hold for all k and thus for all population structures. For example, for fixed a and b, strategy A will be selected whenever B is "sufficiently more risky than A" in the sense that $c-d$ is sufficiently large. In particular, if a is equal to b and one equilibrium strictly risk dominates the other, then the risk dominant equilibrium will be played with probability 1.

Note also that, for generic a, b, c, d, one strategy will be played with probability 1 under the limit where selection gets weak and the mutation rate goes to zero.

Let us contrast this with Tarnita et al. (2009) who consider 2x2 games, also allow general population structures, and also consider the limit as selection becomes weak, but assume that λ is exogenously given and puts probability $\frac{1}{2}$ on both strategies. In their model, in the limit as selection gets weak, both strategies A and B are used with probability $\frac{1}{2}$, independently of the values a, b, c, and d. This is true both for a fixed mutation rate ε and $\varepsilon \to 0$. Their main result is that, for any ε, if selection is sufficiently weak, strategy A will be played with probability larger than $\frac{1}{2}$ if

$$\sigma \cdot (a-b) + (c-d) > 0$$

and strategy B will be played with probability larger than $\frac{1}{2}$ if

$$\sigma \cdot (a - b) + (c - d) < 0,$$

where σ is a parameter which plays an analogous role as the parameter k in our model.[16] Note that since, in their framework, in the limit as selection gets weak both strategies are used with probability $\frac{1}{2}$ independently of a, b, c, d and in our framework the probability with which s^1 is used converges either to 0 or 1 for generic games as selection gets weak the two models yield very different predictions that could be tested empirically.

6. APPLICATION: FINITELY REPEATED GAMES

Let us first look at G^T, the T-times repeated prisoners dilemma where the stage game is given by

	C	D
C	1	-1
D	2	0

This game is an interesting benchmark since the only outcome consistent with Nash equilibrium is (D,D) in each period.

Let s^{grim} be the "grim-trigger" strategy in the T-times repeated PD, i.e. s^{grim} prescribes C if nobody so far played D and D otherwise. Clearly, $u(s^{grim}, s^{grim}) = 1 \cdot T$ and, for any s, $u(s, s^{grim}) - u(s^{grim}, s) \leqslant 3$.

The condition from Theorem 2 was that for any $s' \in S$:

$$\sum_{s \in S} \mu((s, ..., s)) \cdot (k \cdot (u(s, s) - u(s', s')) + (u(s, s') - u(s', s))) \geqslant 0$$

Therefore, for $s' = s^{grim}$:

$$\sum_{s \in S} \mu((s, .., s)) \cdot k \cdot (u(s, s) - T) + 3) \geqslant 0.$$

Average per-period payoffs in the T-times repeated PD thus satisfy:

$$\sum_{s \in S} \mu((s, ..., s)) \cdot \frac{u(s, s)}{T} \geqslant 1 - \frac{3}{k \cdot T}.$$

In particular, as T becomes large the average per-period payoffs converge to 1, the cooperative payoff of the prisoner's dillemma.

The above reasoning generalizes giving the following result.

Theorem 3. *Assume that, for each location i, there exist two locations j and k such that $j \neq k$ and both w_{ij} and w_{ik} are positive. Let G be any symmetric 2-player stage game with a finite action space A and payoff function g. Consider the finitely repeated game G^T with perfect monitoring consisting of T repetitions of G where payoff of each player is the undiscounted sum of payoffs from the T stages. Then, for the case where selection is weak and*

---

[16] Since they consider fixed mutation rates ε, their σ will, however, depends on the mutation rate ε.

mutations rare, the long run behavior consistent with maximization of evolutionary success by mutants has the property that the average per-period-payoff in the game G^T converges to $\max_{a \in A} g(a,a)$, as the number of repetitions T goes to infinity.

Proof. Straightforward generalization of the argument that was made for the T-times repeated prisoner's dilemma.[17] ☐

Of course, a similar argument could be used to show that, no matter what the population structure and the corresponding value of k, our model predicts significant cooperation in the centipede game as long as that game is sufficiently large. In that sense, the model could also potentially be helpful explaining experimental results for such games (see, for example, McKelvey and Palfrey (1992)), the traveling salesman dilemma (see , for example, Rubinstein (2005), Goree and Holt (2001)) and other games with a similar structure. Note that, if in reality, different players are part of different subpopulations with different population structures and different corresponding parameters k that could be helpful in explaining the observed heterogeneity in games like the traveling salesman dilemma where in experimental settings part of the participants chooses the Nash equilibrium outcome (consistent with small k close to zero) and part uses high strategies (consistent with a sufficiently large k).

7. CONCLUSION

In this paper, we motivate endogenizing λ with the fact that a currently fashionable approach to deal with models where the population is finite and natural selection is stochastic, suffers from many problems. We then propose a natural criterion of evolutionary optimality and show that this criterion leads to sharp predictions in a large class of models and for arbitrary symmetric n-strategy games. (The fact that even under weak selection we require a parameter k is not surprising, it is well understood that local structure matters.) As an application, we show that evolution - or at least our model - favors significant cooperation in finitely repeated games if the number of repetition is large.

We would like to end this paper by pointing out a different way of looking at the approach proposed in this paper. Most evolutionary papers dealing with finite populations introduce some kind of persistent shocks which, depending on the framework are usually called mutations or experiments. Mathematically, this is convenient since it gives the process ergodic properties. However, the source leading to those mutations or experiments is not studied, they are just assumed and their exact form exogenously given. This is done, even though it is often clear, that - in the considered model - it

---

[17] Fis a $a \in \arg\max_{a \in A} g(a,a)$. Use as s' a strategy that requires a player to play a after histories in which only a was played in the past and after all other histories play a strategy that maximizes the expected value of $g(b,c) - g(c,b)$ given the distribution of actions c the opponent will use after that history.

would be extremely hard to argue that the assumed form of experimentation and mutation is in the interest of the individuals who do it. (Consider for example the influential class of models started by Young (1993), Kandori et al. (1993) and Ellison (1993) applied to 2x2 games with a risk dominant equilibrium. There it is quite clear that experimenting players would hardly have a reason to experiment with the riskdominated strategy, but the model still requires them to randomize over both strategies when "experimenting".) As a result, mutations and experiments have sometimes been interpreted as errors (see, for example, van Damme and Weibull (2002)). Note however, that there is a huge difference between a trembling hand leading to occasional mistakes (which we view as a natural concept) and occasional permanent switches to another strategy. All of this suggests that it would be interesting to better understand why players experiment or mutate and how exactly they do it. We hope that the approach proposed in this paper may perhaps help shed some additional light on this issue.

Appendix

Proof of Theorem 1. Consider first the base model for some λ and $\varepsilon = 0$. For any $s', s'' \in S$ define $r_{s's''} > 0$ to be the probability that the dynamic reaches a state where all players use s'' from a state where all but one player use s' and one player uses s'', where the location of the player using s'' was drawn randomly with the same probability distribution that is used to determine at which location a mutation will occur. Note that, since the mutation rate ε is equal zero, this definition will not depend on λ.

Lemma 1. *Let $\lambda \in \Delta S$. Then the support of μ_λ is equal to $\{(s, s, ..., s) \in S^N : s \in Support(\lambda)^N\} \subset S^N$. Moreover, μ_λ satisfies*

$$\mu_\lambda((s, ..., s)) = \sum_{s' \in S} \mu_\lambda((s', ..., s')) \cdot \lambda(s) \cdot r_{s's}.$$

Proof. Note that, for $\varepsilon > 0$, the Markov chain P^ε has a unique recurrent class equal to $Support(\lambda)^N \subset S^N$. The restriction of the chain to that recurrent class is a regular Markov process in the sense, in which this term is used by Young (1998) and, therefore, the limit

$$\mu_\lambda = \lim_{\varepsilon \to 0} \mu_\lambda^\varepsilon$$

exists and the support of μ_λ can only contain states that are in a recurrent class of the unperturbed dynamic $(P^0_{ss'})_{s,s' \in S^N}$. Since the recurrent classes of $(P^0_{ss'})_{s,s' \in S^N}$ correspond exactly to singletons $\{(s, s, ..., s)\}$ where the state $s \in S$ the statement it follows that the support of μ_λ is contained in $\{(s, s, ..., s) \in S^N : s \in Support(\lambda)^N\} \subset S^N$. Then, standard tree-surgery methods (see Young (1993,1998)) immediately imply that

$$Support(\mu_\lambda) = \{(a, ..., a) \in S^N \ : \ a \in S\}$$

since moving between any two recurrent classes of the unperturbed Markov chain requires only a single mutation. The formula

$$\mu_\lambda((s, ..., s)) = \sum_{s' \in S} \mu_\lambda((s', ..., s')) \cdot \lambda(s) \cdot r_{s's}$$

now simply follows from the observation that, as $\varepsilon \to 0$ the probability that a mutation occurs while the dynamic is in a state in $Support(\mu_\lambda)$ converges to one and that the probability of another mutation occuring before the dynamic goes to a state where again all players use the same strategy converges to zero. ☐

Lemma 2. *Define $r_{s's''}$ for $s', s'' \in S$ as above. Consider an initial state where all players use $\lambda^* \in \Delta S$ except for a single player in some population P who uses $\lambda' \in \Delta S$. Then*

$$\sum_{s,s' \in S} \mu_{\lambda^*}((s, s, ..., s)) \cdot \mu_{\lambda'}((s', .., s')) \cdot r_{ss'} \leqslant \sum_{s,s' \in S} \mu_{\lambda'}((s, .., s) \cdot \mu_{\lambda^*}((s', s', ..., s')) \cdot r_{ss'} \tag{7.1}$$

implies

$$\underline{\lim}_{\varepsilon \to 0} \underline{\lim}_{\eta \to 0} E^{\varepsilon, \eta} = 1.$$

Proof. Consider a fixed ε and η for a moment. Let $T_1, T_2, ...$ be the random times in which players were transfered between populations. Let C^k be the set of populations at time $T_k - 1$ in which at least one player used λ'. We will now inductively characterize

$$\Pr(C^{k+1} = B | C^k = A)$$

under the considered double limit for finite $A \subset \mathscr{P}$.

Since the transfer between populations at time T_k is the only transfer that occured between periods $T_k - 1$ and $T_{k+1} - 1$, it is clear that

$$\Pr(C^{k+1} = B | C^k = A)$$

converges to zero under our double limit if B none of the following is true: (i) $A = B$, (ii) B contains exactly one more population then A, (iii) A contains exactly one more population than A. Trivially, if $A = \emptyset$, then $\Pr(C^{k+1} = \emptyset | C^k = A) = 1$. Let us now consider the case where A is non-empty.

For $A \neq \emptyset$ and $P' \notin A$, consider

$$\Pr(C^{k+1} = A \cup \{P'\} | C^k = A).$$

This is the probability that the strategy λ' is used by populations in A and population P' at time $T_{k+1} - 1$ if it was used by populations in A at time T_k. Note that, for that to be possible, the only transfer between populations that happened between times $T_k - 1$ and $T_{k+1} - 1$ had to involve a player from some population in A who used λ' being transfered to the population P' and then surviving there rather than being dying out before time T_{k+1}. This probability under the considered limit therefore converges to

$$G(A) \cdot G(P') \cdot \sum_{s,s' \in S} \mu_{\lambda'}(s) \cdot \mu_{\lambda^*}(s') \cdot r_{ss'}.$$

Note that the speed of convergence is the same for all P'. Suming up the above over all P' we get that

$$\Pr(|C^{k+1}| = |C^k| + 1 | C^k = A)$$

converges to

$$G(A) \cdot (1 - G(A)) \cdot \sum_{s,s' \in S} \mu_{\lambda^*}(s) \cdot \mu_{\lambda'}(s) \cdot r_{ss'}.$$

Now consider the probability

$$\Pr(C^{k+1} = A - \{P'\} | C^k = A)$$

where $P' \in A$. Using an analogous reasoning as above this probability converges to

$$(1 - G(A)) \cdot G(P') \cdot \sum_{s,s' \in S} \mu_{\lambda'}(s) \cdot \mu_{\lambda^*}(s') \cdot r_{ss'}.$$

Summing up over all P' we obtain that

$$\Pr(|C^{k+1}| = |C^k| - 1 | C^k = A)$$

converges to

$$(1 - G(A)) \cdot G(A) \cdot \sum_{s,s' \in S} \mu_{\lambda'}(s) \cdot \mu_{\lambda^*}(s') \cdot r_{ss'}.$$

Now let k_l be the random sequence of values k for which C^k changed. Formally, k_l is defined inductively by

$$k_1 = 1$$

and

$$k_{l+1} = \min\{k | k \geqslant k_l, |C^k| \neq |C^{k_l}|\}.$$

Note that, immediately from the definition of k_l

$$\Pr(|C^{k_{l+1}}| = |C^{k_l}| + 1 | C^{k_l} = B) = \Pr(|C^{k_{l+1}}| = |C^{k_{l+1}-1}| + 1 \mid |C^{k_{l+1}}| \neq |C^{k_{l+1}-1}|, C^{k_{l+1}-1} = B)$$

Thus the above probability under our double limit to

$$\frac{G(B) \cdot (1 - G(B)) \cdot \sum_{s,s' \in S} \mu_{\lambda^*}(s) \cdot \mu_{\lambda'}(s) \cdot r_{ss'}}{(1 - G(B)) \cdot G(B) \cdot \sum_{s,s' \in S} \mu_{\lambda^*}(s) \cdot \mu_{\lambda'}(s') \cdot r_{ss'} + G(B) \cdot (1 - G(B)) \cdot \sum_{s,s' \in S} \mu_{\lambda'}(s) \cdot \mu_{\lambda^*}(s') \cdot r_{ss'}} =$$

$$= \frac{\sum_{s,s' \in S} \mu_{\lambda^*}(s) \cdot \mu_{\lambda'}(s) \cdot r_{ss'}.}{\sum_{s,s' \in S} \mu_{\lambda^*}(s) \cdot \mu_{\lambda'}(s') \cdot r_{ss'} + \sum_{s,s' \in S} \mu_{\lambda'}(s) \cdot \mu_{\lambda^*}(s') \cdot r_{ss'}},$$

which does not depend on B. Let

$$\alpha = \frac{\sum_{s,s' \in S} \mu_{\lambda^*}(s) \cdot \mu_{\lambda'}(s) \cdot r_{ss'}.}{\sum_{s,s' \in S} \mu_{\lambda^*}(s) \cdot \mu_{\lambda'}(s') \cdot r_{ss'} + \sum_{s,s' \in S} \mu_{\lambda'}(s) \cdot \mu_{\lambda^*}(s') \cdot r_{ss'}}.$$

Note that (7.1) implies that $\alpha \leqslant \frac{1}{2}$.

Note that for fixed L:

$$\Pr(C^{k_L} = B | C^{k_1} = \{P\}) = \sum \prod_{i=2}^{L-1} \Pr(C^{k_i} = A^i | C^{k_{i-1}} = A^{i-1}),$$

where the sum is over all sequences $A^1, A^2, ..., A^l$ of length l, such $A^1, ..., A^l$ are sets of populations, that $A^1 = \{P\}$, $A^l = B$ and $||A^{j+1}| - |A^j|| = 1$ for $j = 1, 2, ..., l-1$. Since the sum is finite and we showed that each of the terms appeaaring in th product converges under our double limit it is also the case that $\Pr(C^{k_l} = B)$ converges to some value under our double limit. In particular, for any fixed l, $\Pr(C^{k_l} = \emptyset)$ converges to some value under our double limit.

Let $S_l = 1 + \sum_{i=1}^{l} X_i$ be a random walk where X_i are iid random variables with support $\{-1, 1\}$ such that $\Pr(X_i = 1) = \alpha$. Define q_l be the probability that the random walk returns to zero at time k or earlier. Since for each l' $\Pr(C^{k_{l'+1}} = |C^{k_{l'}}| + 1 | C^{k_l} = B)$ converges to α independently of B and $\Pr(C^{k_{l+1}} = |C^{k_l}| - 1 | C^{k_l} = B)$ converges to $1 - \alpha$ independently of B it follows that, for any fixed l, $\Pr(C^{k_l} = \emptyset)$ converges to q_l.

We want to show that

$$\underline{\lim}_{\varepsilon\to 0}\underline{\lim}_{\eta\to 0}E^{\varepsilon,\eta} = 1.$$

Assume that this is not the case, i.e.

$$\underline{\lim}_{\varepsilon\to 0}\underline{\lim}_{\eta\to 0}E^{\varepsilon,\eta} < a < 1 \tag{7.2}$$

for some number $a < 1$. For any l, ε, and η it is the case that whenever players using λ' are extinct at time $T_{k_l} - 1$ they will stay forever extinct. Therefore, for any l, ε, and η it is the case that

$$\Pr(C^{k_l} = \emptyset) \leqslant E^{\varepsilon,\eta}.$$

Togethere with (7.2) this implies that, for any l,

$$\underline{\lim}_{\varepsilon\to 0}\underline{\lim}_{\eta\to 0}\Pr(C^{k_l} = \emptyset) < a < 1.$$

Since q_l converges to 1 we can find an l such that $q_l > a$. For this l, however, it cannot be that $\Pr(C^{k_l} = \emptyset)$ under the double limit converges to $q_l > a$ and stays below a. The contradiction shows that

$$\underline{\lim}_{\varepsilon\to 0}\underline{\lim}_{\eta\to 0}E^{\varepsilon,\eta} = 1.$$

QED. □

Lemma 3. *Define $r_{s's''}$ for $s', s'' \in S$ as above. Consider an initial state where all players use $\lambda^* \in \Delta S$ except for a single player in some population P who uses $\lambda' \in \Delta S$. Then*

$$\underline{\lim}_{\varepsilon\to 0}\underline{\lim}_{\eta\to 0}E^{\varepsilon,\eta} = 1 \tag{7.3}$$

implies

(7.4)

$$\sum_{s,s'\in S} \mu_{\lambda^*}((s,s,...,s))\cdot\mu_{\lambda'}((s',..,s'))\cdot r_{ss'} \leqslant \sum_{s,s'\in S} \mu_{\lambda'}((s,..,s)\cdot\mu_{\lambda^*}((s',s',...,s'))\cdot r_{ss'}$$

Proof. We will prove the result by contradiction. Assume that (7.4) is not true. Since (7.3) holds there exists a sequence $(\varepsilon_m, \eta_m) \to (0,0)$ such that $\frac{\eta_m}{\varepsilon_m} \to 0$ and

$$E^{\varepsilon_m,\eta_m} \to 1.$$

Define T_k as in the previous lemma. Let D^k be the set of populations at time $T_k - 1$ in which at all players used λ'. In addition let P_k be the population from which an offspring is transfered in period T_k and P'_k be the population to which the offspring is transfered in period T_k. For any state z, let D_z be the set of population in z for which it is the case that all players play λ'.

Fix m for a moment and consider the dynamic for (ε_m, η_m). Fix $P \neq P'$ and consider

$$\Pr(D^2 = \emptyset | D^1 = \{P'\}, P'_1 = P', P_1 = P, z^0 = z)$$

for a state z such that all players not in P use λ^* and everybody in P uses either λ' or λ^*. Note that, for any such state z the above probability converges as $m \to \infty$. Note also, that for each such state z the limit must be smaller equal to $\sum_{s,s' \in S} \mu_{\lambda'}(s) \cdot \mu_{\lambda^*}(s') \cdot r_{ss'}$. Finally note that, as far as the state z is concerned the above probabiliuty only depends on what players in P and P' do in the state z - what z prescribes in all other populations is irrelevant. For fixed m and parameters (ε_m, η_m) let

$$q_m = \sup_z \Pr(D^2 = \emptyset | D^1 = \{P'\}, P'_1 = P', P_1 = P, z^0 = z)$$

where the supremum is taken over all states z of the above form. The last two observations imply that $q_m \to \sum_{s,s' \in S} \mu_{\lambda'}(s) \cdot \mu_{\lambda^*}(s') \cdot r_{ss'}$. Note also that the above definition of q_m does not depend on the initial choice of the populations P and P' as long as $P \neq P'$.

Note that, whenever P, P' are both in D^k or P, P' are both not in D^k, then with probability 1 it is the case that $D^k \subset D^{k+1}$. The only possibility for D^{k+1} not to contain D^k is if P is not in D^k but P' is in D^k. In this case, note that $|D_{k+1}| - |D_k| \geqslant -1$. The above definition of q_m implies that for any k, populations P, P' such that $P \neq P'$ and subset of populations A:
(7.5)

$$q_m \geqslant \Pr(|D^{k+1}| - |D^k| = -1 | P' \in D^k, P \notin D^k, P_k = P, P'_k = P', D_k = A)$$

if we consider our dynamic for (ε_m, η_m). Using an analogous construction we conclude that there exists a sequence r_m such that$r_m \to \sum_{s,s' \in S} \mu_{\lambda^*}(s) \cdot \mu_{\lambda'}(s') \cdot r_{ss'}$ and for any k, populations P, P' such that $P \neq P'$ and subset of populations A:
(7.6)

$$r_m \leqslant \Pr(|D_{k+1}| - |D_k| \geqslant 1 | P \in D_k, P' \notin D_k, P_k = P, P'_k = P', D_k = A)$$

if we consider our dynamic for (ε_m, η_m).

If (7.4) is not true, then $\sum_{s,s' \in S} \mu_{\lambda^*}(s) \cdot \mu_{\lambda'}(s') \cdot r_{ss'} > \sum_{s,s' \in S} \mu_{\lambda'}(s) \cdot \mu_{\lambda^*}(s') \cdot r_{ss'}$ and, therefore, $r_m > q_m$ for all sufficiently large m. Fix an M large enough so that for $m > M$ it is the case that $r_m / q_m > \alpha/(1-\alpha)$ for some $\alpha > \frac{1}{2}$. Define k_l as in lemma 2. Comparing $|D_{k_l}|$ with an asymmetric random walk as in lemma 2 we obtain from Inequalities (7.5) and (7.6) that $|D_k|$ will have a probability of never returning to zero which is at least as much as that of the assymetric random walk $S_l = 1 + \sum_{i=1}^{l} X_i$ where X_i are iid random variables with support $\{-1, 1\}$ such that $\Pr(X_i = 1) = \alpha$. However it cannot be that the probability that $|D_k|$ never returns to zero is bounded away for all $m > M$ given that $E^{\varepsilon_m, \eta_m} \to 1$. The contradiction proves that our assumption that (7.4) is not true was incorrect and therefore (7.4) holds. QED ☐

Note that lemmas 2 and 3 imply that

$$\underline{\lim}_{\varepsilon \to 0} \underline{\lim}_{\eta \to 0} E^{\varepsilon, \eta} = 1$$

is equivalent to

$$\sum_{s,s' \in S} \mu_{\lambda^*}((s, s, ..., s)) \cdot \mu_{\lambda'}((s', .., s')) \cdot r_{ss'} \leqslant \sum_{s,s' \in S} \mu_{\lambda'}((s, .., s) \cdot \mu_{\lambda^*}((s', s', ..., s')) \cdot r_{ss'}$$

Theorem 1 therefore follows immediately from the following lemma.

Lemma 4. *Define $r_{s's''}$ for $s', s'' \in S$ as above. There exists λ such that μ_λ*

$$\sum_{s\in S} \mu_{\lambda^*}((s,s,...,s)) \cdot r_{ss'} \leqslant \sum_{s\in S} \mu_{\lambda^*}((s,s,...,s)) \cdot r_{s's}$$

for $s' \in S$ and for s' is in the support of λ^*

$$\sum_{s\in S} \mu_{\lambda^*}((s,s,...,s)) \cdot r_{ss'} = \sum_{s\in S} \mu_{\lambda^*}((s,s,...,s)) \cdot r_{s's}.$$

Proof. Consider a symmetric 2-player zero-sum game with strategy space S and payoff function

$$g(s,s') = r_{s's} - r_{ss'}.$$

Let $\mu \in \Delta S$ be a symmetric NE of this game. Note that since the game is zero-sum the value of the game is well defined and since it is symmetric the value must be zero. Thus,

$$\sum_{s\in S} \mu(s) \cdot r_{ss'} \leqslant \sum_{s\in S} \mu(s) \cdot r_{s's}$$

for $s' \in S$ and for s' is in the support of μ

$$\sum_{s\in S} \mu(s) \cdot r_{ss'} = \sum_{s\in S} \mu(s) \cdot r_{s's}.$$

Define $\lambda \in \Delta S$ by

$$\lambda(s) = \frac{\mu(s)}{\sum_{s'\in S} \mu(s') \cdot r_{s's}}.$$

Note that, therefore, by lemma 1 it is the case that $\mu_\lambda((s,...s)) = \mu(s)$. Note that, the formula defining λ implies that μ and λ have identical supports. From the above inequalities and equalities for μ we therefore obtain that

$$\sum_{s\in S} \mu_{\lambda^*}((s,s,...,s)) \cdot r_{ss'} \leqslant \sum_{s\in S} \mu_{\lambda^*}((s,s,...,s)) \cdot r_{s's}$$

for $s' \in S$ and for s' is in the support of λ^*

$$\sum_{s\in S} \mu_{\lambda^*}((s,s,...,s)) \cdot r_{ss'} = \sum_{s\in S} \mu_{\lambda^*}((s,s,...,s)) \cdot r_{s's}.$$

QED ☐

Proof of Theorem 2.

Lemma 5. *Define $r_{s's''}$ for $s', s'' \in S$ as at the beginning of this appendix. Then $r_{s's''}$ as a function of the payoffs of the game are continuously differentiable.*

Proof. Consider the definition of $r_{s's''}$ in the base model for some λ and $\varepsilon = 0$. $r_{s's''}$ is a finite weighted sum of $r^i_{s's''}$ where $r^i_{s's''}$ is defined to be the probability that the dynamic reaches a state where all players use s'' from a state where all except the player at location i uses s' and the player at location i uses s''. It is therefore enough to show that $r^i_{s's''}$ are continuously differentiable.

Let Z be the set of states in which only the strategies s' and s'' are used. For any state $z \in Z$ define a_z to be the probability that the dynamic eventually reaches a state where all players play s'' . Notice that to show that $r^i_{s's''}$ are continuously- differentiable it is enough to show that a_z are.

Note that a_z is a solution to the system:

$$a_z = \sum_{z' \in Z} P_{zz'} \cdot a_{z'}$$

for all $z \in Z - \{(s', ..., s'), (s'', ..., s'')\}$,

$$a_{(s',s',...,s')} = 0$$

$$a_{(s'',s'',...,s'')} = 1.$$

Clearly, since $p_{ij} > 0$ for all $i, j \in \{1, 2, ..., N\}$ it has to be the case that for any $z \neq (s', ..., s'), (s'', ..., s'')$, $z' \in Z$ the probability of the dynamic moving from z to z' in a fnite number of periods is positive. It follows that $a_z > 0$ for all $z \in Z - \{(s', ..., s'), (s'', ..., s'')\}$.

We will now argue, that a_z is the unique solution to the above system. Indeed, assume b_z is a different solution solving the same system of equations. Then $a + \alpha \cdot (b - a)$ also solves the same system of equations for any real α. Since $a \neq b$, there either exists a z so that $b_z - a_z < 0$ or there exists a z so that $b_z - a_z > 0$. Assume first there exists a z so that $b_z - a_z < 0$ for some z, we will deal with the second case in a moment. Let α be the largest value for which $a_z + \alpha \cdot (b_z - a)$ are non-negative for all $z \in Z - \{(s', ..., s'), (s'', ..., s'')\}$. Since α is the largest such value, $a_z + \alpha \cdot (b_z - a_z)$ must be equal to zero for some state $z \in Z - \{(s', ..., s'), (s'', ..., s'')\}$. Given that for any $z \neq (s', ..., s'), (s'', ..., s'')$, $z' \in Z$ the probability of the dynamic moving from z to z' in a fnite number of periods is positive this cannot be the case. For the case where there exists a z so that $b_z - a_z > 0$ the argument is analogous except that the vector $a + \alpha \cdot (a - b)$ is considered.

We have shown that our system of $|S|^N$ equations with $|S|^N$ unknowns has exactly one solution. Rewrite this system as $A \cdot a = b$. Since the system has exactly one solution the determinant of A is not equal to zero. Note that the above reasoning did not rely on the payoffs of the stage game and therefore holds for all games G. Using Cramer's rule we obtain a representation of a_z in terms of $P_{zz'}$ which shows that a_z are continuously differentiable for states $z \in Z$. (Cramer's rule represents the solutions as rational functions as long as the determinant of A is not equal to zero.) Since $r^i_{s's''} = a_z$ for a particular state z it follows that $r^i_{s's''}$ are continuously differentiable as a function of the game payoffs. $\square$

From lemma 3 we know that if $\lambda^* \in \Delta S$ is an evolutionary optimal mutation strategy then

$$\sum_{s\in S} \mu_{\lambda^*}((s,s,...,s)) \cdot r_{ss'} \leqslant \sum_{s\in S} \mu_{\lambda^*}((s,s,...,s)) \cdot r_{s's}$$

for all $s' \in S$ and for s' is in the support of λ^*

$$\sum_{s\in S} \mu_{\lambda^*}((s,s,...,s)) \cdot r_{ss'} = \sum_{s\in S} \mu_{\lambda^*}((s,s,...,s)) \cdot r_{s's}.$$

Rewrite the above as

$$\sum_{s\in S} \mu_{\lambda^*}((s,s,...,s)) \cdot (r_{ss'} - r_{s's}) \leqslant 0$$

for all $s' \in S$ and for s' is in the support of λ^*

$$\sum_{s\in S} \mu_{\lambda^*}((s,s,...,s)) \cdot (r_{ss'} - r_{s's}) = 0.$$

Note that $r_{ss'}$ only depends on the payoffs $u(s,s'), u(s',s), u(s,s), u(s',s')$. Since the function $r_{ss'}$ is differentiable according to lemma 5

$$r_{ss'} = A \cdot u(s',s') \cdot \gamma - B \cdot u(s,s) \cdot \gamma + C \cdot u(s',s) \cdot \gamma - D \cdot u(s,s') \cdot \gamma + O(\gamma^2)$$

moreover the assumptions we made on p_{ij} imply that $B, C \geqslant 0$ and $A, D > 0$. (To see $A, D > 0$ consider states where exactly one player uses s'.) Plugging into the above equations and taking the limes as γ goes to zero we get

$$\sum_{s\in S} \mu_{\lambda^*}((s,s,...,s)) \cdot ((A+B) \cdot (u(s'',s') - u(s,s'')) + (C+D) \cdot (u(s'',s'') - u(s,s))) \leqslant 0$$

for all $s' \in S$. Since $D > 0$ and $C \geqslant 0$ we can divide by $C + D$ obtaining

$$\sum_{s\in S} \mu_{\lambda^*}((s,s,...,s)) \cdot (k \cdot (u(s'',s') - u(s,s'')) + (u(s'',s'') - u(s,s))) \leqslant 0$$

for $k = \frac{A+B}{C+D}$.

References

[1] T. Antal, A. Traulsen, H. Ohtsuki, C.E. Tarnita and M.A. Nowak, (2009). Mutation-selection equilibrium in games with multiple strategies. *Journal of Theoretical Biology* , 258, 614-622

[2] Bergin, J. and B. Lipman (1996) "Evolution with State-Dependent Mutations," *Econometrica* 64, pp 943-956.

[3] Goree and Holt (2001): "Ten Little Treasures of Game Theory and Ten Intuitive Contradictions," *American Economic Review*, Vol 91: 1402-1422.

[4] McKelvey, R. and T. Palfrey (1992): "An experimental study of the centipede game," *Econometrica*, 60: 803-836.

[5] Nowak M., Sasaki A, Taylor C, Fudenberg D (2004) Emergence of cooperation and evolutionary stability in finite populations. *Nature* 428:646–650

[6] Ohtsuki, H. Hauert, C., Liebermann, E., Nowak, M.A. (2006): "A simple rule for the evolution of cooperation on graphs and social networks.", *Nature* 441, 502-505

[7] Rosenthal, R. (1982): "Games of Perfect Information, Predatory Pricing, and the Chain Store Paradox", *Journal of Economic Theory*, 25.

[8] Rubinstein, A. (2005): “Dilemmas of an Economic Theorist,” working paper.

[9] Tarnita, C, H. Ohtsuki, T. Antal, F. Fu, and M. Nowak. (2009): “Strategy selection in structured populations,” *Journal of Theoretical Biology* 259: 570-581.

[10] Tarnita C.,N. Wage, M. Nowak (2011): “Multiple strategies in structured populations.” Proc Natl Acad Sci 108:2334–2337.

[11] Taylor C, Fudenberg D, Sasaki A, Nowak MA (2004) Evolutionary game dynamics in finite populations. Bull Math Biol 66:1621–1644.

[12] van Damme, E. and J. Weibull (2002): “Evolution in Games with Endogenous Mistake Probabilities,” *Journal of Economic Theory* 106 : 2, 296-315

[13] Wild G. and A. Traulsen (2007): “The different limits of weak selection and the evolutionary dynamics of finite populations,” *Journal of Theoretical Biology* 247, 382–390

[Wu] Wu, B., C. Gokhale, L. Wang, A. Traulsen (2012): “How small are small mutation rates?,” *Journal of Mathematical Biology* 64_5, 803-827

[14] Young, P. (1998): “Individual Strategy and Social Structure: An evolutionary Theory of Institutions” Princeton University Press, Princeton, NJ

[15] Young, P. (1998b): “Conventional Contracts,” *Review of Economic Studies* 65, 773-792.