The Structure of Non-zero-sum Stochastic Games

Robert Samuel Simon

7 December 2005

CDAM Research Report LSE-CDAM-2005-19

London School of Economics
Department of Mathematics
Houghton Street
London WC2A 2AE

This research was supported financially by the German Science Foundation (Deutsche Forschungsgemeinschaft) and the Center for High Performance Computing of the Technical University (Dresden). Discussions on Conjecture 1 with Michael Voit were helpful.

Abstract

Strategies in a stochastic game are $\delta > 0$ perfect if the induced one-stage games have certain δ equilibrium properties. Sufficient conditions are proven for the existence of δ perfect strategies for all $\delta > 0$ implying the existence of ϵ equilibria for every $\epsilon > 0$. Using this approach we prove the existence of ϵ equilibria for every $\epsilon > 0$ for a special class of quitting games. The important technique of the proof belongs to algebraic topology and reveals that more general proofs for the existence of ϵ equilibria in stochastic games must involve the topological structure of how the equilibria of one-stage games are related to changes in the payoffs.

Key words: Stochastic Games, Equilibria, Orbits of Dynamic Systems (Discrete Time), Martingales, Markov Chains, Total Variation, The Structure Theorem (Game Theory)

1 Introduction

A *stochastic* game is played on a state space. The present state and the present behavior of all players determines stochastically the transition to a new state. All players have complete knowledge of the past history of play and the present state. A priori there is no bound on the number of stages of play.

We define a stochastic game to be *normal* if
(1) there are countably many states,
(2) there are finitely many players and at any state the action sets for all players are finite,
(3) the payoffs defined in the game are uniformly bounded,
(4) the payoffs are functions on the histories of play that are measurable with respect to the Borel σ-algebra defined by the finite stages of the game. This fourth property will be made more precise later.

For any $\epsilon \geq 0$, an ϵ *equilibrium* in a game is a set of strategies, one for each player, such that no player can gain in payoff by more than ϵ by choosing a different strategy, given that all the other players do not change their strategies. An equilibrium is a 0 equilibrium. We say that approximate equilibria exist if for every $\epsilon > 0$ there exists an ϵ equilibrium.

It is not known whether all normal stochastic games have approximate equilibria. This question is arguably the most important open question of game theory today. Advantageous for approximate equilibria in stochastic games is the common knowledge by the players of the past history of play and the present options and their consequences. The only uncertainty concerns what the other players will do in the present and in the future. If the stochastic game has finitely many stages then equilibria exist, a consequence of the original Nash proof (Nash [9]). Disadvantageous for approximate equilibria are the infinite number of stages of play.

A stochastic game is a *limit average* game when for every player n the payoff is between $\lim_{i\to\infty}\inf$ and $\lim_{i\to\infty}\sup$ of the average $\frac{1}{i}\sum_{k=0}^{i-1} w^n_{s_k}(a^1_k, \ldots, a^m_k)$ where m is the number of players and for every state $s \in S$ w^n_s is a real function defined on the collections $(a^1, a^2, \ldots, a^m)$ of actions, once for each player, at the state s.

So far, most positive results of profound generality have concerned two-player limit average stochastic games. Mertens and Neyman [7] proved that every zero-sum limit average and normal stochastic game played on a finite state space has approximate equilibria. Maitra and Sudderth [5] extended this result to zero-sum stochastic games with countably many states and Martin [6] extended this result further to payoff functions defined on the infinite paths of play that are Borel with respect to the σ algebra generated by the finite stage truncations.

Concerning two-player non-zero-sum limit average and normal stochastic games the central result was accomplished by Vieille [14]; he proved that all such stochastic game with finitely many states have approximate equilibria. For two-player non-zero-sum normal games with countably many states the question is still open.

One approach to non-zero-sum normal stochastic games is to break down the game into infinitely many one-stage games. Given a strategy for each player and any given present state and past history of play, one can look at the one-stage game that starts with this present state and ends with reaching the state on the following stage. One assumes that the players will act according to the given strategies on all the future stages but on the present stage they are free to choose and the payoff consequences for their choices of actions at the present state are determined accordingly. As was shown in Simon [10], the existence of approximate equilibrium implies a property known as *perfection*, which concerns ϵ equilibrium conditions for the one stage games.

This break down of a stochastic game to its one-stage games also breaks down the problem of the existence of approximate equilibria to two separate questions:
a) does some perfection property hold, and
b) can this perfection property imply the existence of approximate equilibria?
In many situations the perfection property allows one to construct approximate equilibrium through statistical testing and punishment in response to statistical deviation. We establish some conditions (Theorems 1 and 2) for which perfection implies the existence of approximate equilibria.

The general model for normal non-zero-sum stochastic games is given in the second section. Basic results showing that some forms of perfection properties will imply the existence of approximate equilibria is the subject of the third section. This approach was inspired by the Vieille proof [14], which uses a special case of Theorem 2.

In the fourth section we investigate a special class of stochastic games called *quitting games*. Examples of quitting games were studied first by Flesch, Thuijsman, and Vrieze [2] but defined in generality by Solan and Vieille [13]. For quitting games the perfection property does imply the existence of approximate equilibria, and we prove this below.

In the fifth section we define a class of quitting games called escape games. Using algebraic topology we prove that all escape games have the perfection property, hence have approximate equilibria.

In the Conclusion, we discuss the broader question of whether all normal stochastic games have approximate equilibria.

2 The Model

2.1 Normal Stochastic Games

For every finite or countable set A let $\Delta(A)$ stand for the set of all probability distributions on A. If A is finite then $\Delta(A)$ is a finite dimensional simplex. If $x \in \Delta(A)$ and $a \in A$ then the a coordinate of x will be represented as $x(a)$ (the probability given to a by x).

There is a countable or finite state space S and a finite set N of players. For every player $n \in N$ and every $s \in S$ there is a finite set A^n_s of actions. For every $s \in S$ and every $a \in A_s := \prod_{n \in N} A^n_s$ (a choice of action for each player) there will be a transition law $p^s_a \in \Delta(S)$ governing the motion to states at the next stage of play after a visit to s.

We assume that the game starts at an initial state $\hat{s} \in S$. (If one prefers to start with a distribution on all the states in S one can add an initial state $\hat{s}$ that occurs only at the start of the game and such that every player has only one action at this state.) Define $\mathcal{H}_\infty := \{(\hat{s} = s_0, a_0, s_1, a_1, \ldots) \mid \forall i \geq 0 \ a_i \in A_{s_i}, \ p^{s_i}_{a_i}(s_{i+1}) > 0\}$, the set of possible infinite histories of play. Define $\mathcal{H}^{\hat{s}}_0 := \{(\hat{s})\}$, and for every $i \geq 1$ let $\mathcal{H}^s_i$ be the set of truncations of $\mathcal{H}_\infty$ of the form $(\hat{s} = s_0, a_0, s_1, a_1, \ldots, s_{i-1}, a_{i-1}, s_i = s)$ (leaving out the actions at stage i). Let $\mathcal{H}_i$ be the union $\cup_{s \in S} \mathcal{H}^s_i$ and let $\mathcal{H}^s$ be the union $\cup_{i=0}^\infty \mathcal{H}^s_i$. Let $\mathcal{H}_\omega$ be the union $\cup_{i=0}^\infty \mathcal{H}_i = \cup_{s \in S} \mathcal{H}^s$. If $h \in \mathcal{H}_\omega$ is also in $\mathcal{H}^s$ then we say that

h *terminates* at s. The i stage truncation of either an infinite history in $\mathcal{H}_\infty$ or of a finite history in $\mathcal{H}_j$ for $j \geq i$ is the canonical projection to $\mathcal{H}_i$.

A payoff for a player $n \in N$ in a normal stochastic game is a function $\mathcal{V}^n$ on $\mathcal{H}_\infty$ that is uniformly bounded and measurable with respect to the Borel σ-algebra generated by the partitions of $\mathcal{H}_\infty$ induced by the discrete partitions of $\mathcal{H}_i$. A two-player game is *zero-sum* if $\mathcal{V}^1(h) + \mathcal{V}^2(h) = 0$ for all $h \in \mathcal{H}_\infty$ (where without loss of generality we assume that $N = \{1, 2\}$). Let $M \geq 1$ be a positive real number larger than the maximal difference between all payoffs in the game.

2.2 Strategies and Equilibria

A strategy σ^n of Player $n \in N$ is a collection of functions $(\sigma^n_s \mid s \in S)$ such that for every $s \in S$ σ^n_s is a function from $\mathcal{H}^s$ to $\Delta(A^n_s)$. For every profile $\sigma = (\sigma^n \mid n \in N)$ of strategies, one strategy for each player, probability distributions $\mu_{\sigma,i}$ are induced on the $\mathcal{H}_i$ in the natural way. We start at the initial history $(\hat{s}) \in \mathcal{H}^{\hat{s}}_0$ with $\mu_{\sigma,0}(\{(\hat{s})\}) = 1$. Given that $\mu_{\sigma,i}(h_i)$ is positive for some $h_i \in \mathcal{H}^{s_i}_i$ and $h_{i+1} \in \mathcal{H}_{i+1}$ is a history such that the i stage truncation of h_{i+1} is equal to $h_i \in \mathcal{H}^{s_i}_i$ with $h_{i+1} = (h_i, a_i, s_{i+1})$ and $a_i = (a^n_i \mid n \in N)$ we define inductively $\mu_{\sigma,i+1}(h_{i+1}) := \mu_{\sigma,i}(h_i) p^{s_i}_{a_i}(s_{i+1}) \prod_{n \in N} \sigma^n_{s_i}(h_i)(a^n_i)$. A regular Borel probability distribution μ_σ is induced on $\mathcal{H}_\infty$ in the natural way, by the $\mu_{\sigma,i}$ and Kolmogorov's Extension Theorem. For every player $n \in N$ and every strategy profile σ the distribution μ_σ generates a payoff $\mathcal{V}^n(\sigma)$ for player n as the expected value of the function $\mathcal{V}^n$ on $\mathcal{H}_\infty$, determined by the probability distribution μ_σ.

For any profile $\sigma = (\sigma^n \mid n \in N)$ of strategies, an alternative profile $\tilde{\sigma} = (\tilde{\sigma}^n \mid n \in N)$ and a player $k \in N$ define $\sigma|\tilde{\sigma}^k$ to be the strategy profile such that $\tilde{\sigma}^k$ is the strategy for player k but if $n \neq k$ then σ^n is the strategy for player n. An ϵ equilibrium is a strategy profile $\sigma = (\sigma^n \mid n \in N)$ such that for any alternative strategy profile $(\tilde{\sigma}^n \mid n \in N)$ and every player $n \in N$ it holds that $\mathcal{V}^n(\sigma|\tilde{\sigma}^n) \leq \epsilon + \mathcal{V}^n(\sigma)$. A zero-sum game has the value $r \in \mathbf{R}$ for a designated first player if for every positive ϵ there is an ϵ equilibrium whose expected payoff for the first player is within ϵ of r.

2.3 Perfection

For any profile σ of strategies, a player $n \in N$, and a stage i of play define $v_\sigma^n : \mathcal{H}_i \to \mathbf{R}$ by $v_\sigma^n(h_i)$ equaling the expected value of $\mathcal{V}^n(\sigma)$ conditioned on reaching h_i on the ith stage, with $v_\sigma^n(h_i)$ defined to be any quantity bounded within the payoffs defining the game if the probability of reaching h_i is zero. Extend to a definition of $v_\sigma^n : \mathcal{H}_\omega \to \mathbf{R}$ in the natural way.

For every player n define $\chi^n : S \to \mathbf{R}$ so that $\chi^n(s)$ is the min-max value for player n at the state s, the upper bound for what player n can obtain from a start at s in response to all strategy choices of the other players. Formally $\chi^n(s)$ equals $\inf_\sigma \sup_{\tilde{\sigma}^n} \mathcal{V}_s^n(\sigma|\tilde{\sigma}^n)$ where the payoff function $\mathcal{V}_s^n$ is defined by the game for which s is the initial state. The importance of the function χ^n is that it represents the ability of the players to punish player n with predetermined strategies (for example as part of an approximate equilibrium).

For every $a^n \in A_s^n$ and $\hat{a} \in \prod_{k \neq n} A_s^k$ let $(\hat{a}, a^n)$ be the corresponding member of $A_s = \prod_{k \in N} A_s^k$, with $\hat{a}^k$ the corresponding action of Player k for all $k \neq n$.

For any function $f : \mathcal{H}_\omega \to \mathbf{R}$, state $s \in S$, finite history $h \in \mathcal{H}^s$, action $a^n \in A_s^n$ and strategy profile σ define $w_\sigma^f(h)(a^n)$ to be the expected value of f on the next stage after h, conditioned on the use of a^n by Player j and the use of $\sigma_s^k(h)$ by all the other players $k \neq j$. This means that

$$w_\sigma^f(h)(a^n) = \sum_{t \in S} \sum_{\hat{a} \in \prod_{k \neq n} A_s^k} f(h, (\hat{a}, a^n), t) \prod_{k \in N \setminus \{n\}} \sigma_s^k(h)(\hat{a}^k) p_{(\hat{a}, a^n)}^s(t).$$

Define $w_\sigma^n(h)(a^n)$ to be $w_\sigma^{v_\sigma^n}(h)(a^n)$. For any σ and player n the functions w_σ^n and v_σ^n have the property that for every $h \in \mathcal{H}_\omega$ the value $v_\sigma^n(h)$ is equal to the expectation of $w_\sigma^n(h)(\cdot)$ taken over all the actions of player n and $w_\sigma^n(h)(a^n)$ is equal to the expectation of v_σ^n on the next stage following h, conditioned on the event that a^n was chosen with positive probability.

For every player $n \in N$ and strategy profile σ, define the *jump* function $j_\sigma^n : \mathcal{H}_\omega^s \to \mathbf{R}$ by

$$j_\sigma^n(h) = \max_{a^n \in A_s^n} \sum_{t \in S} \chi^n(t) \sum_{\hat{a} \in \prod_{k \neq n} A_s^k} \prod_{k \in N \setminus \{n\}} \sigma_s^k(h)(\hat{a}^k) p_{(\hat{a}, a^n)}^s(t),$$

namely the maximal expected value of χ^n on the next stage following s. Extend this definition to $j_\sigma^n : \mathcal{H}_\omega \to \mathbf{R}$ in the natural way. With the definition of χ^n extended to a function on $\mathcal{H}_\omega$ so that if h terminates at t then $\chi^n(h)$ is equal to $\chi^n(t)$, for all $h \in \mathcal{H}_s$ we have $j^n_\omega(h) = \max_{a^n \in A^n_s} w^{\chi^n}_\sigma(h)(a^n)$.

Definitions: A strategy profile σ of a stochastic game is ϵ *perfect* if for every player $n \in N$ there exists a function $r^n : \mathcal{H}_\omega \to \mathbf{R}$ and a subset $\mathcal{B} \subseteq \mathcal{H}_\omega$ such that the probability of reaching $\mathcal{H}_\omega \backslash \mathcal{B}$ with the strategies σ does not exceed ϵ and for all players $n \in N$ and all finite histories $h \in \mathcal{B}$,
$r^n(h) \geq j^n_\sigma(h) - \epsilon$,
$|r^n(h) - v^n_\sigma(h)| \leq \epsilon$, and
for all actions a^n chosen with positive probability by σ^n at h
$|w^{r^n}_\sigma(h)(a^n) - r^n(h)| \leq \epsilon$.
A stochastic game is ϵ self-perfect if for all players n the function r^n is equal to the function v_σ. A stochastic game is *perfect* if there exists an ϵ perfect strategy profile for every positive ϵ and *self perfect* if there is an ϵ self-perfect strategy profile for every positive ϵ.

The following theorem was proven in Simon [10]: A normal stochastic game with approximate equilibria is also perfect.

3 From Perfection to Approximate Equilibria

3.1 The Basic Result

For a normal stochastic game it is easy to define a topology on the infinite histories $\mathcal{H}_\infty$ of the game. For each finite stage i there will be only countably many histories in $\mathcal{H}_i$. A member of the base of open sets is a set of the form $\{O_{h_i} := \{h \mid$ the ith stage truncation of h is $h_i\}$ for any finite history $h_i \in \mathcal{H}_\omega$. Given a strategy profile σ the μ_σ is a regular Borel probability distribution, meaning that for every Borel measurable subset $A \subseteq \mathcal{H}_\infty$ and every $\epsilon > 0$ there is a closed subset C of infinite histories contained in A and an open subset O of infinite histories containing A such that the measure of the open set $O \backslash C$ is no more than ϵ.

A collection $f = (f^n : \mathcal{H}_\omega \to \mathbf{R} \mid n \in N)$ of functions is called *viable* if for every $\epsilon > 0$ and finite history h that terminates at s from a start at the state s there are strategies $\sigma = (\sigma^n \mid n \in N)$ by the players such that no player n can receive more than $f^n(h) + \epsilon$ from any choice of an alternative strategy $\hat{\sigma}^n$ played against the strategies $(\sigma^k \mid k \neq n)$ of the other players. If additionally for every player $n \in N$ the expected payoff from σ were within ϵ of $f^n(h_0)$ with $h_0 = (\hat{s})$ then we would be describing a 2ϵ equilibrium of the game. With only two players viability means exactly that for every state s the function gives to each player at any history terminating at s at least her min-max value for the state s, as both players can hold down the other player simultaneously to their min-max value plus any arbitrary $\epsilon > 0$. However with three or more players viability is more complex.

We define a strategy profile σ to be δ-*viable* if there are viable functions $f = (f^n \mid n \in N)$ such that $v^n_\sigma(h) \geq f^n(h) - \delta$ for every history h in $\mathcal{H}_\omega$.

For every player n, strategy profile σ, and finite history $h = (s_0, a_1, \ldots, s_i) \in \mathcal{H}_\omega$ define $W^n_\sigma(h) = \sum_{j=0}^{i-1}(w^n_\sigma(h_j)(a^n_j) - v^n_\sigma(h_j))$, where h_j is the j stage truncation of h.

Theorem 1: If $0 < \epsilon \leq 1$ and σ is an ϵ self-perfect and ϵ-viable strategy profile of a normal stochastic game such that for every player n with probability no more than ϵ some history h is reached with $W^n_\sigma(h) > \epsilon$, then the game has a $3(M|N| + 5)\epsilon$ equilibrium.

Proof: For every player $j \in N$ define $v^j_\sigma : \mathcal{H}_\infty \to \mathbf{R}$ by $v^j_\sigma(h) = \lim_{i\to\infty} \sup v^j_\sigma(h_i)$. By the Martingale Convergence Theorem this limit equals the lim inf of the same expression almost everywhere. Furthermore v^j_σ equals $\mathcal{V}^j$ almost everywhere, as their integrals are equal on all open sets of positive measure and the distribution μ_σ induced on $\mathcal{H}_\infty$ is Borel.

Due to the regularity of μ_σ there is an open subset $\mathcal{O}$ of $\mathcal{H}_\infty$ of measure no more than $\epsilon/(2|N|M)$ that contains all infinite histories where some finite truncation is outside of $\mathcal{B}$ (defining the perfection property) and where for all $n \in N$ the function v^n_σ does not equal the function $\mathcal{V}^n$ on $\mathcal{H}_\infty$. We extend this to an open set $\mathcal{A}$ of $\mathcal{H}_\infty$ of probability no more than $(1 + |N|)\epsilon$ that contains all infinite histories with finite truncations h where $W^n_\sigma(h) > \epsilon$ for some $n \in N$.

Define the following strategies of the players. If any player n chooses an action that was not given positive probability by σ then on the next following stage all other players hold player n down to an expectation of no more than $\chi^n(s) + \epsilon$ for the rest of the game, where s is the state on the following stage. (If two players do this simultaneously then punishment follows according to any predetermined ordering of the players.) If h is the first finite history reached which implies that any infinite extension of h must be in $\mathcal{A}$ yet no player had chosen an action given zero probability (well defined by the definition of the topology) then the players perform according to strategies holding down each player n to a future expectation of no more than $f^n(h)+\epsilon$ where the f^n are the viable functions with $v^n_\sigma \geq f^n - \epsilon$. Otherwise the players follow the strategies σ. Let $\hat{\sigma}$ stand for this strategy profile . Due to the unlikelihood of reaching the set $\mathcal{A}$ we have $v^n_{\hat{\sigma}}(h_0) \geq v^n_\sigma(h_0) - (2 + |N|)M\epsilon$ for every player n, (where $h_0 \in \mathcal{H}_0$ is the initial history).

Define $\overline{\sigma}$ to be the strategy profile where Player n chooses some alternative strategy $\overline{\sigma}^n$ and the other players stay with their strategies as defined by $\hat{\sigma}$. Define a stop rule t on $\mathcal{H}_\omega$ by $t(h)$ being the first stage where all future infinite histories must belong to the open set $\mathcal{A}$ or the next stage following the first stage when player n chooses a strategy given zero probability. Otherwise if neither occurs let $t(h)$ be infinite. For player n define two functions $\tilde{g}^n_i, g^n_i : \mathcal{H}_i \rightarrow \mathbf{R}$ by $g^n_i(h_i) = v^n_\sigma(h_i) - W^n(h_i)$ and $\tilde{g}^n_i(h_i) = v^n_\sigma(h_i)$ if $t > i$ and otherwise $g^n_i(h) = f^n(h_{t(h)}) - W^n(h_{t(h)}) - \epsilon$ and $\tilde{g}^n_i(h) = f^n(h_{t(h)}) + \epsilon$ if $t(h)$ is the first stage implying that $\mathcal{A}$ must be reached in the future (but no player had chosen an action given zero probability) or $g^n_i(h) = \chi^n(s_{t(h)}) - W^n(h_{t(h)}) - \epsilon$ and $\tilde{g}^n_i(h) = \chi^n(s_{t(h)}) + \epsilon$ if $t(h)$ is the first stage when player n had chosen an action given zero probability. The function g^n defines a sub-martingale with respect to the distribution $\mu_{\overline{\sigma}}$. The function $\tilde{g}_n$ is never 4ϵ more than g_n. Both functions g^n_i and $\tilde{g}^n_i$ converge everywhere to Borel measurable functions $g^n : \mathcal{H}_\infty \rightarrow \mathbf{R}$ and $\tilde{g}^n : \mathcal{H}_\infty \rightarrow \mathbf{R}$, the former because a sub-martingale is defined and the latter because the stop time t is defined using the open set $\mathcal{A}$ that covers all points in $\mathcal{H}_\infty$ where the limit v^n_σ doesn't exist or doesn't equal the payoff function $\mathcal{V}^n$. Furthermore, the expectation $\mathcal{V}^n(\overline{\sigma})$ does not exceed the expectation of $\tilde{g}^n$. As the g^n_i is a sub-martingale we must conclude that the expectation of g^n does not exceed that of $v^n_\sigma(h_0)$, and that concludes the proof. $\Box$.

Question 1: Does Theorem 1 hold if viability is dropped?

The difficulty of Question 1 lies with determining whom is to be punished. Given that $\mathcal{A}$ is an open subset of infinite histories covering all of the finite histories which should trigger punishment, the relevant question is "who is responsible for steering the game toward the set $\mathcal{A}$" (and therefore should be punished)? Without the viability of the strategies, Player n may want to steer the game toward $\mathcal{A}$ in such a way that often player k is held responsible for entering the set $\mathcal{A}$, although player k was adhering faithfully to her part of the proscribed strategy profile σ. If the stochastic game is structured in a sufficiently simple way, viability may be unnecessary. As we will see below, the viability property can be dropped for quitting games.

3.2 Discrete Decision Processes

Let X be a countable or finite set. For every $x \in X$ let Y_x be a countable or finite set, with $Y := \cup_{x \in X} Y_x$. For every $x \in S$ there is a transition law $p^x \in \Delta(Y_x)$ and for every $y \in Y_x$ there is a transition law $p^y \in \Delta(X)$. The process starts at some fixed $\hat{x} \in X$ and on the even stages $i = 0, 2, 4, \ldots$ the process is in X and on the odd stages the process is in Y. There is a function $v : X \cup Y \to \mathbf{R}$ such that for every $y \in Y_x$ $v(y)$ is the expectation of $v(x)$ on the next stage following y and $v(x)$ is the expectation of $v(y)$ on the next stage following x. We assume that v is uniformly bounded, with $M \geq 1$ a bound on the greatest difference between any two values of v.

A Markov chain with a function as described in the above paragraph is called a *discrete decision process*.

The interpretation of a discrete decision process is as follows. There is an agent choosing the actions in Y_x. The agent receives as a payoff the lim-sup of the function v on the path of states in X. Given that the agent chooses elements in Y according to the time independent Markovian strategy defined by the p^x at any state the function v will represent the agent's future expected payoff (since by the uniform bound for v and the Martingale Convergence Theorem there will be convergence almost everywhere). We presume that the agent will follow the given strategy, but we will image what could happen if the agent chose to follow a different strategy.

The connection to stochastic games is direct. Let j be a player in a normal stochastic game. Given any strategy profile σ a discrete decision process for Player j is defined by extending the state space so that $X = \mathcal{H}_\omega$. Define Y_h to be only those actions in A^j_s (h terminating at s) chosen with positive probability. Because every state in the new expanded state space is encountered at most once, time indendendent Markovian strategies are well defined, in addition to a function v derived from the v^j_σ on the set $X = \mathcal{H}_\omega$

of finite histories and from the $w^j_\sigma(h)$ on the actions in Y_h. Also a discrete decision process for Player j may be defined by any partition of the finite histories that is equal to or finer than the partition $\{\mathcal{H}^x \mid x \in S\}$ such that $v^j_\sigma(h) = v^j_\sigma(h')$ and $\sigma^j(h) = \sigma^j(h')$ for all h, h' in the same partition member. When this occurs the discrete decision process for Player j is *generated* by the stochastic game and the strategy profile σ. We get the following corollary of Theorem 1.

Corollary 1: If for a normal stochastic game there were a strategy profile σ that is ϵ self-perfect, ϵ-viable, and for every player j a discrete decision process is generated such that the probability that there is an l with $\sum_{i=0,2,\ldots}^{l}(v(y_{i+1}) - v(x_i)) \geq \epsilon$ does not exceed ϵ then this stochastic game has a $3\epsilon(M|N| + 5)$ equilibrium.

For any given path $p = (x_0, y_1, x_2, y_3, \ldots)$ of a discrete decision process define $\overline{w}(p)$ to be $\sum_{i=0,2,\ldots} |v(y_{i+1}) - v(x_i)|$.

Notice that discrete decision processes and the functions $\overline{w}$ involve no loss of generality from Markov chains and the total variation of a martingale function defined on them. Given a Markov chain, we could define Y_x so that there is a bijection between Y_x and the states that follow x with positive probability, and then for every $y \in Y_x$ define the distribution p^y to be the appropriate Dirac mass.

Definition: A discrete decision process is is ϵ *balanced* if for all states $x \in X$ and $y \in Y_x$ it follows that $|v(y) - v(x)| \leq \epsilon$.

Proposition 1: Assume that a discrete decision process is δ balanced, that the expectation of $\overline{w}(p)$ does not exceed a finite $B > 0$ and δ is less than or equal to $\epsilon^2\rho/B$ for some positive $\rho > 0$. Then the probability that there exists an l with $|\sum_{i=0,2,\ldots}^{l}(v(y_{i+1}) - v(x_i))| \geq \epsilon$ does not exceed ρ.

Proof: The Doob sub-martingale inequality states that if $(S_i \mid i = 0, 1, \ldots, n)$ is a martingale with zero expectation then for every $n \geq 0$, positive value $c > 0$ and exponent $p \geq 1$ the probability that $\max_{i \leq n} |S_i| > c$ does not exceed $\mathbf{E}(|S_n|^p)/c^p$ (Williams [15], Section 14.6). Since the martingale property implies that $\mathbf{E}(S_n^2)$ is equal to the sum over all the stages $1 \leq i \leq n$ of the $E(s_i^2)$ where $s_i = S_i - S_{i-1}$ is the change in value between the $i-1$st stage and the ith stage, we can re-write as

$$\text{Probability } \Big(\max_{i\leq n} |S_i| > \epsilon \Big) < \frac{1}{\epsilon^2} \mathbf{E} \Big(\sum_{i=1}^{n} s_i^2 \Big).$$

In the context of a discrete decision process, for every $i = 0, 2, \ldots$ define the random variable r_i to be $v(y_{i+1}) - v(x_i)$, and for every $i = 0, 2, \ldots$ let R_i be the sum of the r_k for the even $k \leq i$. The process R_i is a martingale with zero expectation and so for every even and non-negative integer Q and even non-negative even integer i less than or equal to Q

$$\text{Probability } \Big(\max_{i\leq Q} |R_i| > \epsilon \Big) < \frac{1}{\epsilon^2} \mathbf{E} \Big(\sum_{i=0,2,\ldots Q} r_i^2 \Big).$$

By taking the limit as Q goes to infinity and $\delta \leq |r_i|$ we get

$$\text{Probability } \Big(\max_{i<\infty} |R_i| > \epsilon \Big) < \frac{1}{\epsilon^2} \mathbf{E} \Big(\sum_{i<\infty} r_i^2 \Big) \leq$$

$$\delta \frac{1}{\epsilon^2} \mathbf{E} \Big(\sum_{i<\infty} |r_i| \Big) \leq \delta B / \epsilon^2$$

The conclusion follows from the size of δ. □

For all even i and path $p = (x_0, y_1, \ldots)$ in a discrete decision process, either infinite or finite going at least to some x_i, define
$W_i(p) := \sum_{j=0,2,\ldots,i} (v(y_{i+1} - v(x_i)).$

3.3 Rank

Given a discrete decision process, for any subset $A \subseteq X$, $x \in A$ and any $y \in Y_x$ define $r^A(y) \in \Delta(A)$ to be the distribution on A determined by the location of the next visit to A. Formally, consider the process such that x is the initial state, the action y is chosen at the initial stage at x, and at all subsequent stages the actions are chosen according to the given time independent distributions. Let q_y^A be the probability that this process returns to the set A at some stage after the initial stage and for every $z \in A$ let $q_y^{A,z}$ be the probability that this return occurs and first at the state z. If q_y^A is positive then define $r_y^A \in \Delta(A)$ by $r_y^A(\{z\}) = q_y^{A,z}/q_y^A$ and otherwise define r_y^A to be anything.

Definitions: A state $x \in X$ of a discrete decision process is *varied* if there exists some $y \in Y_x$ such that $v(y) \neq v(x)$. The *rank* of a discrete decision process is the minimal number n such that the varied states can be partitioned into n subsets $A_1, \dots, A_n$ with the property that for every $k = 1, 2, \dots, n$ and every $x \in A_k$ there is a distribution $r_x^{A_k} \in \Delta(A_k)$ such that for every $y \in Y_x$ the distribution $r_y^{A_k}$ is equal to $r_x^{A_k}$.

Proposition 2: If a discrete decision process has rank n then the expectation of $\overline{w}$ does not exceed $2nM$.

Proof: For every subset $A = A_k$ and $x \in A$ let l_x be the probability that the last visit to A occurs at x, let m_x be the probability that there is no return to A from a start at x, and let $n_i(x)$ be the probability that x is the state on the ith stage. We have $l_x = \sum_i n_i(x) m_x$, $\sum_{x \in A} l_x \leq 1$, and $m_x = \sum_{y_i n Y_x} (1 - q_y^A) p^x(y)$

Define v_x to be the expected value of v conditioned on starting at x and returning to the set A (with v_x defined to be anything if this occurs with zero probability). Because the distributions are Markovian and time independence and there is a constant r_x^A for the distributions r_y^A for all $y \in Y_x$ we have $v(y) = q_y^A v_x + (1 - q_y^A) s_y$ for some s_y whose difference from v_x does not exceed M, and therefore $|v(y) - v_x| \leq M(1 - q_y^A)$.

Next consider the quantity $|v_x - v(x)|$. The equality $v(x) = \sum_{y \in Y_x} p^x(y)((1 - q_y^A) s_y + q_y^A v_x)$ implies $v(x) - v_x = \sum_{y \in Y_x} p^x(y)(1 - q_y^A)(s_y - v_x)$ and $|v(x) - v_x| \leq M m_x$.

By the triangle inequality $|v(y) - v(x)| \leq |v(y) - v_x| + |v_x - v(x)|$. By $m_x = \sum_{y_i n Y_x} (1 - q_y^A) p^x(y)$ the contribution to $\overline{w}$ in the set $A = A_k$ does not exceed $2M \sum_i \sum_{x \in A} n_i(x) m_x$, which is no more than $2M$. □

The following example shows that the conclusion of Proposition 2 must be dependent on the rank or on some similar concept.

Example 1: The set X has $2n+1$ states, namely $x_{-n}, x_{-n+1}, \dots, x_0, \dots, x_{n-1}, x_n$. Assume that $|Y_{x_{-n}}| = |Y_{x_n}| = 1$ and that the state on the next stage following any visit to x_{-n} is again x_{-n} and the same holds for the state x_n. Define $v(-n)$ to be -1 and $v(n)$ to be 1. The process starts at x_0 and for every i strictly between $-n$ and n there are two elements of Y_{x_i}, namely L and R. If L is chosen then the process moves to the state x_{i-1} with certainty and if R is chosen then the process moves to the state x_{i+1} with certainty. At every state strictly between x_{-n} and x_n the actions L and R are both chosen with 1/2 probability. Extend v to a function $v : X \to [0,1]$ that defines a Martingale; it follows that $v(x_i) = \frac{i}{n}$ for all $-n \leq i \leq n$. Given any small $\delta > 0$, one can make n large enough so that δ is less than $\frac{1}{n}$. However from a start at the position 0 the probability that W_i will reach 1 for some i will be exactly 1/2. By Kolmogorov's inequality with probability at least 1/2 the process will avoid $-n$ and n for at least $\frac{n^2}{2}$ stages, implying that the expectation of $\overline{w}$ will be at least $\frac{n}{2}$.

3.4 Chain Reduction

We would like to exploit Proposition 2 in combination with Corollary 1. We look for any way to reduce our discrete decision process to that of fewer states so that the rank could go down but the distribution of $\max_i W_i(p)$ does not change significantly.

A subset B of non-varied states of a discrete decision process is *removable* if from any visit to a state in B the probability of leaving the set B at some future stage is one.

In some cases the decisions made in a subset A of states can be represented equivalently as decisions made at a single state. This happens if there is some special state s such that the first visit to the subset $A \cup \{s\}$ is always at the state s and there is a finite m such that from any state in A before m stages occur, regardless of the choice of actions, the process leaves the set A. The probability distributions on the Y_x for all the $x \in A \cup \{s\}$ can be represented by probability distributions on the set $Y'_s = \times_{z \in A \cup \{s\}} Y_z$ (Kuhn [4]).

However a reduction of such a subset $A \cup \{s\}$ to the single state s could present problems for applying Corollary 1 to stochastic games. If a player should be punished for striving to attain a higher payoff, should that player be punished for the actions actually made or for the actions in Y'_s? The Y'_s may define counter-factual behavior, meaning that many different "actions" in Y'_s may generate the same seen behavior. On the other hand, the variance of the functions $W_i(p)$ may be considerably higher with the actions from the original discrete decision process than from such a reduction.

Define two disjoint subsets S and T to be *chain reducible* if T is removable, for every $s \in S$ there is a finite subset $A_s \subseteq X \backslash T$ not containing s such that every visit to the set $A_s \cup \{s\}$ starts at s, there is a positive integer m such that from any start in A_s before m stages the subset A_s is left, and furthermore for every x in $A_s \backslash T$ and any $y \in Y_x$ either with probability one the first state reached in $X \backslash T$ is not in A_s or there is a state $n(y)$ in $A_s \backslash T$ such that if y is chosen then with probability one $n(y)$ is reached first before any other state in $X \backslash T$. An action $y \in A_x$ with $x \in A_s$ such that with probability one the next state in $X \backslash T$ is not in A_s is called a *completing* action. If S and T are chain reducible then define a new discrete decision process with the new state space $X \backslash (T \cup_{s \in S} A_s)$ and for every $s \in S$ the action space $\overline{Y}_s$ is defined to be $\{(y_0, y_1, \ldots, y_k) \mid y_0 \in Y_s, \forall \ \ 0 \leq i < k \ \ y_{i+1} \in Y_{n(y_i)}$ and y_k is completing$\}$. The new discrete decision process is called a *chain reduction* of the original discrete decision process.

Lemma 1: If a chain reduction of a discrete decision process is $\delta > 0$ balanced then for every $\epsilon > 0$ the probability of $\sup_i W_i$ of the original discrete decision process exceeding $\epsilon + \delta$ is not greater than the probability of the same expression exceeding ϵ for the chain reduction.

Proof: Let $x_0, y_1, x_2, \ldots, x_i, y_{i+1}$ be any sequence in the original discrete decision process. It can be broken down to $(x_0, \ldots, y_{n_1 - 1}), (x_{n_1}, y_{n_1+1}, \ldots y_{n_2-1}), \ldots,$ $(x_{n_k}, \ldots x_i, y_{i+1})$ where the $x_0, x_{n_1}, \ldots, x_{n_k}$ are states in the chain reduction. Given any partial sequence $x_{n_l}, \ldots y_{n_{l+1}-1}$ and $y = (y_{n_l+1}, \ldots, y_{n_{l+1}-1}) \in \overline{Y}_{x_{n_l}}$ by the probability one properties of the chain reduction $\sum_{j=n_l,\ldots,n_{l+1}-2}(v(y_{j+1}) - v(x_j)) = v(y) - v(x_{n_l})$. Therefore it suffices for any sequence $x_0, y_1, \ldots, x_l, y_{l+1}$ with $x_0 \in S$ and $x_1, \ldots, x_l \in A_{x_0}$ that $\sum_{j=0,2,\ldots,l}(v(y_{j+1}) - v(x_j)) \leq \delta$. Complete $x_0, \ldots, y_{l+1}$ to any $x_0, \ldots, y_{l+1}, x_{l+2}, \ldots, x_k, y_{k+1}$ satisfying $x_i = n(y_{i-1})$ for all even $l \leq i \leq k$, y_{k+1} is completing, and $\sum_{j=l+2,l+4,\ldots,k} (v(y_{j+1}) - v(x_j)) \geq 0$. By the δ balanced property it follows that $\sum_{j=0,2,\ldots,l}(v(y_{j+1}) - v(x_j)) \leq \delta$. □

Theorem 2: If for a normal stochastic game there is a number k such that for every $0 < \epsilon \leq 1$ there is a strategy profile σ that is ϵ self-perfect, ϵ viable and for every player there is a generated discrete decision process with a chain reduction that is $\epsilon^3/(3kM)$ balanced of rank k then the game has approximate equilibria.

Proof: It follows from Corollary 1, Proposition 1, Proposition 2, and Lemma 1.

A chain reduction can result in a very significant drop in the expectation of $\overline{w}$.

Example 2: The set X consists of $x_0, x_1, \ldots, x_n$ and b. At the states b and x_n the sets Y_b and Y_{x_n} have only one element and the result of this one action is return to these respective states with certainty. Define $v(b)$ to be 1 and $v(x_n)$ to be -1. At every $0 \leq i \leq n-1$ the set Y_{x_i} consists of two elements L and R. If L is chosen then with probability $\frac{1}{2^n-1}$ there is motion to the state b and with probability $\frac{2^n-2}{2^n-1}$ there is motion to the state x_0. If R is chosen then there is motion with certainty to the state x_{i+1}. At every state x_i with $0 \leq i \leq n-1$ the actions L and R are chosen both with 1/2 probability. By induction one can prove that $v(x_i) = \frac{1-2^i}{2^n-1}$ (with $v(x_0) = 0$). The discrete decision process can be chain reduced to the three states x_0, x_n, and b with $A_{x_0} = \{x_1, \ldots, x_{n-1}\}$. The probability of not returning to x_0 from a start at x_0 would be $\frac{1}{2^{n-1}}$, with half of this probability resulting in a move to a and the other half to a move to x_n. The expectation on $\overline{w}$ in the chain reduction would be 1, as the number of expected visits to x_0 would be 2^{n-1} and at each visit to x_0 there would be a probability of $\frac{1}{2^{n-1}}$ of the function v changing by exactly a value of 1. However in the original discrete decision process the expected number of visits to x_i is $2^{-i}2^{n-1}$ and the expected change in v from one visit to the state x_i is $\frac{2^i}{2^n-1}$ (and hence $\frac{2^{n-1}}{2^n-1}$ from all visits to x_i). This implies that the expectation on $\overline{w}$ exceeds $\frac{n}{2}$.

3.5 Markov Chains and Total Variation

Let X be the finite state space of a Markov chain and $v : X \times \{0, 1, \ldots\} \rightarrow [0, 1]$ a function such that for every x on stage i the value $v(x, i)$ is the expectation of $v(\cdot, i+1)$ on stage $i+1$. For any infinite path $p = (x_0, x_1, \ldots)$ in X define the quantity $\overline{w}(p) = \sum_{i=0}^{\infty} |v(x_{i+1}, i+1) - v(x_i, i)|$.

Lemma 2: If the Markov chain is time homogeneous then the expected value of the function $\overline{w}$ is no more than $|X|$.

Proof: For every $x \in X$ define q_x to be the probability that starting at x the process will not return to x in the future. The contribution to $\overline{w}$ at the state x will not exceed q_x times the number of expected visits to x. The number of expected visits to x does not exceed $1 + (1 - q_x) + (1 - q_x)^2 + \ldots = 1/q_x$.

Conjecture 1: Without the time homogeneous assumption the expected value of the function $\overline{w}$ is no more than $|X|$.

The Markovian property is critical to Conjecture 1; it is easy to find counter-examples if the transitions and the function are dependent on the past history. The main difficulty with Conjecture 1 lies with the lack of a state identity that transcends the stages. We would be satisfied if the expectation of $\overline{w}$ does not exceed $f(n)$ for any function $f : \{1, 2, \ldots\} \to \mathbf{R}$ that is independent of the choice of Markov chain.

4 Quitting Games

4.1 The Definition

In a quitting game each player has only two action s, c for continue and q for quit. As soon as one or more of the players at any stage chooses q, the game stops and the players receive their payoffs, which are determined by the subset of players that choose simultaneously the action q. As long as no player has stopped the game, all players receive a payoff of zero.

Let N be the set of players. A strategy profile for the players is a sequence of probability vectors $(p_i \mid i = 0, 1, 2, \ldots)$ such that for every stage i $p_i \in [0, 1]^N$. p_i^j stands for the probability that Player j will stop the game (with the action q) at stage i conditioned on the event that stage i is reached. With $\underline{0}$ standing for the origin, $\underline{0} \in [0, 1]^N$ means that all players choose the action c with certainty.

The payoffs are defined as follows. For every non-empty subset $A \subseteq N$ of players there is a payoff vector $v(A) \in \mathbf{R}^N$. At the first stage that any player chooses the action q and A is the non-empty subset of players that choose q at this stage, the players receive the payoff $v(A)$. This means that Player i receives $v(A)^i \in \mathbf{R}$. If nobody chooses the action q throughout all stages of play, then all players receive 0.

A quitting game is a normal stochastic game. Let $\hat{x}$ be the state at the start and at any stage such that at all previous stages all players had chosen c. We could define $2^{|N|-1}$ additional states corresponding to the non-empty subsets of N such that once any of these states is reached then no matter what the players do the game remains at this state forever and the players receive the corresponding payoffs. Equivalently we could choose $|N|+1$ affinely independent vectors in $\mathbf{R}^N$ whose convex hull contains all the payoffs defined in the game – then a subset A of players quitting at the same time causes an appropriate probability distribution on the $|N|+1$ states. Let $M \geq 1$ be an upper bound on the difference between all payoffs.

4.2 Correspondences and Orbits

By a *correspondence* $F : X \to\to Y$ we mean any subset F of $X \times Y$. If X_0 is a subset of X then $F \cap (X_0 \times Y)$ is called the restriction of F to X_0 and denoted by $F|X_0$. For every $x \in X$ define $F(x) := \{y \mid (x, y) \in F\}$. It is not assumed a priori that $F(x) \neq \emptyset$ for all or any particular $x \in X$. The domain of a correspondence F is the subset $\{x \mid F(x) \neq \emptyset\}$ and the image of F is the subset $\{y \mid y \in F(x) \text{ for some } x \in X\}$.

If $F : X \to\to X$ is a correspondence then an *infinite orbit* of the correspondence F is an infinite sequence $(x_0, x_1, \ldots)$ of points of X such that for every non-negative integer $n \geq 0$ we have $(x_n, x_{n+1}) \in F$. A finite orbit is a finite sequence $(x_0, \ldots, x_l)$ with $(x_n, x_{n+1}) \in F$ for all $0 \leq n < l$. An *extended orbit* of F is a sequence $(s_j \mid 0 \leq j < L)$ of sequences $s_j = (x_{j,i} \mid 0 \leq i < n_j)$, possibly with $L = \infty$ or $n_j = \infty$ for some or all $j < L$, such that for every $i + 1 < n_j$ $x_{j,i+1} \in F(x_{j,i})$ and if $n_j = \infty$ then $\lim_{k\to\infty} x_{j,k} = x_{j+1,0}$ and otherwise $x_{j,n_j-1} = x_{j+1,0}$. The extended orbit has bounded total variation if $\sum_{j<L} \sum_{1 \leq i < n_j} ||x_{j,i} - x_{j,i-1}|| < \infty$, and otherwise it has unbounded total variation.

For every $r \in \mathbf{R}^N$ let Γ_r be the one stage game where Player $j \in N$ receives the payoff r^j if all players choose the action c.

For every $r \in \mathbf{R}^N$ and $p \in [0,1]^N$, let $a^j(p)$ be the expected payoff for Player j if she chooses q simultaneously with the strategies $(p^k \mid k \neq j)$ and let $b^j(p, r)$ be the expected payoff for Player j from the action c in the game Γ_r, given that the other players choose the strategies $(p^k \mid k \neq j)$, meaning that she will receive the payoff r^j if everyone chooses the action c. One can calculate $a^j(p)$ and $b^j(p,r)$ easily. We have

$$a^j(p) = \sum_{A \subseteq N \setminus \{j\}} v(A \cup \{j\})^j \prod_{k \in A} p^k \prod_{k \neq j,\ k \notin A} (1 - p^k)$$

and

$$b^j(p,r) = r^j \prod_{k \neq j} (1 - p^k) + \sum_{\emptyset \neq A \subseteq N \setminus \{j\}} v(A)^j \prod_{k \in A} p^k \prod_{k \neq j,\ k \notin A} (1 - p^k).$$

Every strategy profile $p = (p_i \mid i = 0, 1, 2, \ldots)$ defines payoff vectors $r_i(p) \in \mathbf{R}^N$ for all $i = 0, 1, 2, \ldots$ such that r_i^j is the expected payoff for player j from the strategy profile $(p_i, p_{i+1}, \ldots)$, equivalent to the payoff conditioned on all players choosing c before the stage i given that no player chooses q with certainty before this stage.

Define a function $q : [0,1]^N \to [0,1]$ by $q(p) := 1 - \prod_{j \in N}(1 - p^j)$. The function q is the probability that at least one player chooses the action q.

We will consider the correspondences generated by moving backward from some stage $i+1$ to stage i through an approximate equilibrium of the one stage game. For any $\epsilon \geq 0$ we define correspondences $E_\epsilon \subseteq \mathbf{R}^N \times [0,1]^N$ and $F_\epsilon \subseteq \mathbf{R}^N \times \mathbf{R}^N$:

$$E_\epsilon(r) := \{p \in [0,1]^N \mid p^j > 0 \Rightarrow a^j(p) \geq b^j(p,r) - \epsilon,$$

$$p^j < 1 \Rightarrow b^j(p,r) \geq a^j(p) - \epsilon\}.$$

For every $r \in \mathbf{R}^N$ and $p \in [0,1]^N$ define a new member of $\mathbf{R}^N$, namely

$$f(r,p) := r \prod_{j \in N} (1 - p^j) + \sum_{\emptyset \neq A \subset N} v(A) \prod_{j \in A} p^j \prod_{j \notin A} (1 - p^j),$$

the expected payoffs in the game Γ_r when the players choose p. We define F_ϵ by $F_\epsilon(r) := \{f(r,p) \mid p \in E_\epsilon(r)\}$.

4.3 Normal Players, Instant and Stationary Equilibria

Definitions: A vector $r \in \mathbf{R}^N$ is *feasible* if it is in the convex combination of $\{v(A) \mid \emptyset \neq A \subseteq N\} \cup \{\underline{0}\}$. For every player $n \in N$ define χ^j to be $\chi^j(\hat{x})$ where $\hat{x}$ is the initial state. A vector $r \in \mathbf{R}^N$ is ϵ*-rational* for any positive ϵ

if $r^n \geq \chi^n - \epsilon$ for all $n \in N$. A player $n \in N$ is *normal* if $v(\{n\})^n \geq \chi^n$. The vector $v \in \mathbf{R}^N$ is defined by $v^i := v(\{i\})^i$ for every player i.

Lemma 3: If j is an abnormal player then $v^j < 0$ and $v(\{i\})^j \geq \chi^j$ for every $i \neq j$.

Proof: Consider what happens when every other player chooses c with certainty at every stage. Player j could respond by choosing q at any stage. $v^j < \chi^j$ implies that responding by never choosing q must be at least as good as χ^j, meaning that $v^j < \chi^j \leq 0$.

Let $\delta > 0$ be given, and consider what happens when Player i chooses q with a probability of δ at every stage of play (and all other players choose c with certainty). By quitting at any stage Player j would receive no more than $v^j + \delta M$, and for small enough δ this would be worse than χ^j. It follows that choosing c at all stages would be the much better choice for Player j, implying that $v(\{i\})^j \geq \chi^j$. □

Lemma 4: Let $p_1, p_2, \ldots, p_k$ be a sequence of one-stage strategies in $[0,1]^N$ such that $0 < \rho = 1 - \prod_{i=1}^k (1 - q(p_i)) < 1$ is the probability that some player chooses q on some stage and let $s_0, \ldots, s_k$ be a sequence of vectors in $\mathbf{R}^N$ such that $s_i = f(s_{i-1}, p_i)$ for each $1 \leq i \leq k$. If $||s_0 - s_k|| \leq \delta$ then the strategy profile $p = (p_k, p_{k-1}, \ldots, p_1, p_k, \ldots, p_1, \ldots) = \overline{p}_0, \overline{p}_1, \ldots,$ generates a sequence $r_i(p)$ for $i = 0, 1, \ldots$ such that $||r_i(p) - s_{nk-i}|| \leq \frac{\delta}{\rho}$ for all $(n-1)k < i \leq nk$ and if $p_i \in E_\epsilon(s_{i-1})$ then $\overline{p}_i \in E_{\epsilon + \frac{\delta}{\rho}}(r_i(p))$.

Proof: Define $r \in \mathbf{R}^N$ to be the payoffs to the players conditioned on the event that some player chose q from the strategies $p_k, \ldots, p_1$ (starting with p_k). We have assumed that $s_k = (1-\rho)s_0 + \rho r$. With $||s_0 - s_k|| < \delta$ and $\rho(r - s_0) = s_k - s_0$ we have $||r - s_0|| \leq \delta/\rho$. With $r_{nk}(p) = r$ for all multiples nk of k and by the definition of the function $f(r,p)$ we have $||r_i(p) - s_{nk-i}|| \leq \delta/\rho$ for all $(n-1)k < i \leq nk$. The last claim follows from $|b^n(r,p) - b^n(s,p)| \leq ||r - s||$ for all vectors $r, s \in \mathbf{R}^N$ and players $n \in N$.

Definitions: A quitting game has *stationary approximate equilibria* if for every $\epsilon > 0$ there is a $p \in [0,1]^N$ such that $(p, p, p, \ldots)$ is an ϵ equilibrium. A quitting game has *instant approximate equilibria* if for every $\epsilon > 0$ there is a $p \in [0,1]^N$ with $p^j = 1$ for some player $j \in N$ and such that a 2ϵ equilibrium is described by the behavior p on the first stage followed by punishment of Player j on the second stage (given that she didn't quit) yielding to Player

j no more than $\chi^j + \epsilon$.

Lemma 5: If a quitting game does not have stationary approximate equilibria or instant approximate equilibria then
1) $v^j > 0$ for some normal player and for every normal player j there is another normal player k such that $v(\{j\})^k < v^k$,
2) there is an $\rho > 0$ small enough so that if $r \in \mathbf{R}^N$ is a ρ-rational vector within a distance of 1 of a feasible vector, $p \in E_\rho(r)$ and $y = f(r)$ then
a) $\rho q(p) \leq ||x - y||$ and
b) $q(p) \leq 1 - \rho$.

Proof: 1) If the first claim didn't hold for some player, normal or abnormal, then all players choosing c on all stages would be an equilibrium; and by Lemma 3 this player must be normal. Furthermore, if there were not such a second player k, normal or abnormal, then with $\epsilon > 0$ fixed player j choosing q at every stage with probability small enough would describe an ϵ equilibrium; and by Lemma 3 this player must be normal.

2) Let $((r_i, s_i, p_i) \mid i = 1, 2, \ldots)$ be a sequence such that all the r_i are within a distance of 1 of a feasible vector and p_i is a member of $E_{1/i}(r_i)$ with $s_i = f(x_i, p_i)$, $r_i^j \geq \chi^j - 1/i$ for all $j \in N$ and either $\frac{q(p_i)}{i} > ||r_i - s_i||$ or $q(p_i) \geq 1 - \frac{1}{i}$. Let $0 < \epsilon \leq 1$ be fixed and choose i large enough so that $i\epsilon^{|N|} > (3M)^{|N|}$. Let $\hat{r}_i$ be the vector in $\mathbf{R}^N$ representing the expected payoffs of the players from the stationary strategy profile $p = (p_i, p_i, p_i, \ldots)$. If $\frac{q(p)}{i} > ||r_i - s_i||$ it holds by Lemma 4 that $||\hat{r}_i - r_i|| \leq \epsilon/2$ and therefore p_i is in $E_\epsilon(\hat{r}_i)$. Now assume that $q(p_i) \geq 1 - \frac{1}{i}$. For some player n the quantity p^n is at least $1 - \epsilon/(3M)$. Define $\hat{p}$ to be the the strategy profile such that $\hat{p}^j = p^j$ if $j \neq n$ and otherwise $\hat{p}^n = 1$. It follows that $\hat{p}$ along with punishment of Player n for not quitting does describe an ϵ equilibrium. □

4.4 Equivalences

Proposition 3 and Theorem 3 are generalizations of some results of Solan and Vieille [13].

Proposition 3: Let $0 < \epsilon \leq 1$ be given and let positive δ be less than $\frac{\epsilon^4}{2M^3}$. A cyclic strategy profile $p = (p_0, \ldots, p_{k-1}, p_k = p_0, \ldots)$ with all $r_i(p)$ ϵ-rational, $q(p_i)$ positive for some $0 \leq i \leq k-1$ and $r_i(p) \in F_\delta(r_{i+1}(p))$ for all $i = 0, 1, \ldots$ generates a 3ϵ equilibrium.

Proof: For all $i \geq 1$ and players $n \in N$ define u_i^n to be the summation $\sum_{k=0}^{i-1}(b^n(r_k^n(p), p_{k-1}) - r_{k-1}^n(p))$, the cumulative advantage in expectation that Player n has obtained by choosing the action c on all stages up to but not including the stage i, conditioned on the event that no other player has chosen q. For every player n define i_n^* to be the first stage i where u_i^n is at least ϵ. For every player $n \in N$ and stage $i \geq 1$ define c_i^n to be $\prod_{k=0}^{i-1}(1 - p_k^n)$, Define $i_n^\sharp$ to be the first stage i such that c_i^n is no more than $\frac{\epsilon}{M}$.

We must determine whom to punish and when. Define $\hat{i}$ to be $\min_{n \in N}(i_n^*, i_n^\sharp)$. If $\hat{i}$ is equal to $i_n^\sharp$ for some $n \in N$, then define $\hat{n} \in N$ to be any $n \in N$ with $\hat{i} = i_n^\sharp$. Otherwise if $\hat{i}$ is less than $i_n^\sharp$ for all $n \in N$ then define $\hat{n}$ to be any $n \in N$ such that $\hat{i}$ is equal to i_n^*. Before the stage $\hat{i}$ the players perform according to p. If the game reaches stage $\hat{i}$ then player $\hat{n}$ will be punished such that the expected future payoff for this player is no more than $\chi^{\hat{n}} + \epsilon/10$.

Because the decision to choose q terminates the game immediately, and the one stage advantage by doing so never exceeds $\epsilon^4/(2M^3)$, the only deviant strategy we need to consider is the repetitive decision to choose c by a player. Due to $\hat{i} \leq i_{\hat{n}}^*$ and that the vectors are ϵ rational, there is no advantage beyond 2ϵ for Player $\hat{n}$ to choose c repetitively. Likewise from $\hat{i} \leq i_m^*$ for all players $m \in N$ we need only consider the advantage to a player $m \neq \hat{n}$ from the punishment of player $\hat{n}$ at stage $\hat{i}$. It suffices to show that even if Player m never chooses q the stage $\hat{i}$ (for punishing $\hat{n}$) is reached with a probability of no more than ϵ/M.

Case 1; $\hat{i} < i_{\hat{n}}^\sharp$ and $\hat{i} = i_{\hat{n}}^*$: We will look at the discrete decision process for Player $\hat{n}$ generated by the stochastic game and the given profile of strategies. Let s_i be the state representing the history where the ith stage is reached and so far every player has chosen c at every stage up until i. Notice that from s_i the distribution on the next visit to $\{s_0, s_1, \ldots\}$ is the same for both actions, namely total weight given to the state s_{i+1}, (for the action q this holds because there is a zero probability of returning to the set). Therefore for player $\hat{n}$ the generated discrete decision process has rank 1. By Propositions 1 and 2 the probability does not exceed $\epsilon^2/(M^2)$ that the stage $\hat{i}$ is reached. As $c_{\hat{i}}^m \geq \epsilon/M$ if Player m never quits the probability of reaching $\hat{i}$ is still no more than ϵ/M.

Case 2; $\hat{i} = i_{\hat{n}}^\sharp$: Whether or not Player m or any other player other than $\hat{n}$ refuses to choose q the probability of Player $\hat{n}$ not choosing q before stage

$\hat{i}$ does not exceed ϵ/M. □

Theorem 3: For a quitting game without stationary approximate equilibria or instant approximate equilibria the following are equivalent:

(i) the game has approximate equilibria,

(ii) for every $\epsilon > 0$ there is a cyclic strategy profile $p = (p_0, \ldots, p_{k-1}, p_k = p_0, \ldots)$ with $r_i(p) \in F_\epsilon(r_{i+1}(p))$ for all $i = 0, 1, \ldots$, all the r_i are ϵ-rational, and $q(p_i)$ is positive for some $0 \leq i \leq k-1$,

(iii) for every $\epsilon > 0$ and every $B > 1$ there is a finite orbit of F_ϵ of ϵ-rational vectors within a distance of 1 of the feasible vectors with a total variation of at least B,

(iv) for every $\epsilon > 0$ there is an infinite orbit of F_ϵ of ϵ-rational vectors with unbounded total variation,

(v) for every $\epsilon > 0$ there is an infinite extended orbit of F_ϵ of ϵ-rational vectors with unbounded total variation.

Proof: (ii) implies (i) is the content of Proposition 3. **(ii) implies (iii), (iv) and (v)** is trivial and **(iv) implies (v)** is also trivial.

(iv) implies (iii): As the orbit has unbounded variation, any cluster point of the orbit must be feasible.

(iii) implies (ii): Let $\epsilon > 0$ be fixed. By the fact that the feasible and $\epsilon/5$ rational vectors form a compact set there will be a B be large enough so that any finite sequence of $F_{\epsilon/5}$ of total variation at least B will have two vectors s_i and s_j with $i < j$ in the sequence separated by a total variation of at least $2M$ such that $||s_i - s_j|| < \epsilon/5$. As they are also separated by strategies $p_j, p_{j-1}, \ldots, p_{i+1}$ giving a probability of at least 1/2 that q was chosen, Lemma 4 suffices for (ii).

(i) implies (iii): Assume that there does exist a positive δ and a bound $B > 0$ such that every orbit of F_δ of vectors that are feasible and δ rational has a total variation less than B. Without loss of generality we assume that δ is less than the ρ given by Lemma 5. By Lemma 5 this case can be reformulated: there exists a positive $\theta > 0$ such that any orbit of F_δ of vectors that are feasible and δ rational is created from a strategy profile where the probability that all players choose c on all stages is at least θ. Also by Lemma 5 there is a $0 < d \leq 1$ such that $v^j \geq d$ for some normal player j. Assume that $p = (p_0, p_1, \ldots)$ is a $\delta\theta\rho d/(5M)$ equilibrium. As some player can obtain at least d by quitting alone, we must assume that the probability according to p of no player ever quitting is no more than $\delta\theta\rho/(4M)$. This means that there must be a stage i where the probability of no player ever quitting before reaching i is between $\theta\rho/3$ and $\theta/3$. Since we have a $\delta\theta\rho d/(5M)$ equilibrium the steps from stage 0 to stage i generate a finite orbit of F_δ of δ-rational and feasible vectors, a contradiction.

(v) implies (ii): We assume the existence of an extended orbit $((x_{l,j} \mid 0 \leq l < Q), j < n_l)$ of $F_{\epsilon/3}$ with unbounded total variation in $\{x \mid \forall j \ x^j \geq \chi^j - \epsilon/3\}$. Let $(p_{l,i})$ be the corresponding strategies in $[0,1]^N$) (such that $x_{l,i+1} = f(x_{l,i}, p_{l,i})$).

Case 1; There is a sequence $(x_{l,0}, x_{l,1}, \ldots)$ such that $\sum_{i=0}^{\infty} ||x_{l,i} - x_{l,i+1}|| = \infty$:

This implies (iii), and we have proven already that (iii) implies (ii).

Case 2; $\sum_{i=0}^{\infty} ||x_{l,i} - x_{l,i+1}|| < \infty$ for every $l < \infty$:

The argument is essentially the same as (iii) implies (ii). Let x be any cluster point of the sequence $(x_{0,0}, x_{1,0}, \ldots)$. Let $x_{m,0}$ and $x_{n,0}$ be any two points in this sequence such that both are within $\epsilon/5$ of x and $\prod_{m<l<n} \prod_{i<n_l} (1 - q(p_{l,i})) < \frac{\epsilon}{30M}$. We can assume without loss of generality that for every l the total variation in the lth orbit is $T_l > 0$. For every $m \leq i \leq n-1$ define k_i large enough so that the total variation from x_{i,k_i} to $x_{i+1.0}$ does not exceed $\rho\epsilon 2^{-i} T_i$. Lemma 4 implies that (ii) holds with the cyclic strategy profile obtained from reversing the probabilities to $p_{n-1,k_{n-1}-1}, \ldots, p_{n-1,0}, p_{n-2,k_{n-2}-1}, \ldots, p_{1,0}, p_{0,k_0-1}, \ldots, p_{0,0}$, and then repeating. □

Lemma 6: Assume that all players are normal, that there are neither stationary nor instant approximate equilibria, and $s \in F_{\epsilon^2/(2M)}(r)$. If $r^n \geq \chi^n - 3\epsilon$ then $s^n \geq \chi^n - 3\epsilon$ and if $r^n < \chi^n - 3\epsilon$ then $s^n \geq r^n + \epsilon^2/(2M)$.

Proof: For the sake of contradiction assume the contrapositive. In either case it must hold that $s^n < \chi^n - 3\epsilon + \epsilon^2/(2M)$, meaning also that $b^n(p, r)$ and $a^n(p)$ are less than $\chi^n - 3\epsilon + \epsilon^2/M$. As Player n can get at least χ^n from quitting alone we must also assume that the total probability that other players are quitting according to p must exceed $2\epsilon/M$. But then by Lemma 4 the strategies $(p^k \mid k \neq n)$ would be a way to hold the payoff of Player n down to $\chi^n - \epsilon$, something impossible. □

Corollary 2: If all players are normal and there are neither instant approximate equilibria nor stationary approximate equilibria then the game has approximate equilibria if and only if there is an orbit $r_0, r_1, \ldots$ of F_δ of unbounded total variation.

5 Escape Games

5.1 The Definition

Define the set $W_j := \{r \mid r^j \leq v^j\}$ and define $W := \cup_{j \in N} W_j = \{r \mid r^j \leq v^j$ for some $j \in N\} = \mathbf{R}^N \backslash \{r \mid r^j > v^j$ for all $j \in N\}$.

A quitting game is an escape game if

1) every player is normal ($v^n = v(\{n\})^n \geq \chi^n$ for all $n \in N$),

and there is a closed subset Q of $\mathbf{R}^N$ and a positive $\bar{\epsilon} > 0$ with the following existence and closure properties:

2) $Q \cap \partial W \neq \emptyset$ and for every $x \in Q \cap \partial W$ there is a y with $y^j > v^j$ for all $j \in N$ such that the closed line segment from x to y is in the set Q,

3) if $x \in Q \backslash W$ then any payoff vector $y \in \mathbf{R}^N$ in $F_0(x)$ with $y \neq x$ satisfies $y^j > v^j + \bar{\epsilon}$ for all $j \in N$,

4) if $x \in Q$ and $y \in F_{\bar{\epsilon}}(x)$ then $y \in Q$.

The name "escape" reflects the assumption that once one has left the set $\{x \mid x^j \leq v^j + \bar{\epsilon}$ for some $j\}$ with the correspondence $F_{\bar{\epsilon}}|Q$ then one has also "escaped" this set for good.

5.2 The Spanning Property

We use a property for correspondences called the "spanning" property, defined in Simon, Spiez, and Torunczyk [12]. The homology used in that article is the Cech homology with coefficients in a non-trivial compact Abelian group. This approach was chosen because the Cech homology is defined using approximations (Eilenberg and Steenrod [1]) and hence many properties are preserved when passing to limits. Because the approximation arguments of this article are made explicit, we could use instead the more conventional homology groups defined by continuous maps from simplicies to the topological spaces (and with integer coefficients).

An n dimensional compact manifold with boundary is a topological space such that every point is contained in a subset of the space topologically equivalent to the n dimensional disk D^n with this point in the center or on the boundary of this disk.

If C is an n-dimensional compact manifold with boundary in $\mathbf{R}^n$ then by $[\partial C]$ we denote the generator element of the reduced homology group $\tilde{H}_{n-1}(\partial C)$ according to any orientation. (The reduced homology group differs from the non-reduced only in dimension 0.) For example, $[\partial C]$ could be generated by any subdivision of C into parts topologically equivalent to D^n with the boundary map applied to functions from the n dimensional simplex to these parts of the subdivision. Let U be a non-empty open bounded subset of E. A compact (correspondence) $F \subseteq \mathbf{R}^n \times Y$ is said to have the *spanning property* for U if there exists a z in the reduced homology group $\tilde{H}_{n-1}(F|\partial\overline{U})$ such that the images of z in $\tilde{H}_{n-1}(\partial\overline{U})$ and $\tilde{H}_{n-1}(F)$ are $[\partial\overline{U}]$ and 0, respectively, where the first map is that induced by the canonical projection of $F|\partial\overline{U}$ to $\partial\overline{U}$ and the second map is that induced by the inclusion of $F|\partial\overline{U}$ in the set F (and $\overline{A}$ stands for the topological closure of A). If the compact correspondence $F \subseteq \mathbf{R}^n \times Y$ has the spanning property for a non-empty open set U then we say that is has the spanning property for the closure of U. If F has the spanning property for an open set U then $F(x) \neq \emptyset$ for every point x in U (proven in Simon, Spiez, and Torunczyk [11]). This property is the origin for the term "spanning".

We demonstrate some of the power of the spanning property. In a usual proof of Brouwer's Fixed Point Theorem, if a continuous function $g : D^n \rightarrow D^n$ didn't have a fixed point then there would be a continuous function $f : D^n \rightarrow S^{n-1}$ such that $f(x) = x$ for all $x \in S^{n-1}$, and from looking at the induced homology groups we see that this is not possible. The spanning property goes further: for any continuous function from D^n to D^n such that for all $x \in S^{n-1}$ it holds that $f(x) = x$ the image of f must cover all of D^n. Indeed, one can go further. Let $f : D^n \rightarrow \mathbf{R}^n \times \mathbf{R}^n$ be any continuous function such that $f(x) = (x, x)$ for all $x \in S^{n-1}$. $F :=$ image (f) is contractible, and hence has zero reduced homology groups for all dimensions. Considering either the first or second copy of S^{n-1} to be the domain of the correspondence F, the embedding of S^{n-1} into $S^{n-1} \times S^{n-1}$ yields an appropriate element of the $n-1$ reduced homology group of the correspondence over S^{n-1}, implying the spanning property for D^n. The spanning property applied to either D^n

in the first or second position in $\mathbf{R}^n \times \mathbf{R}^n$ implies that both the image of F and the domain of F must contain D^n.

An important aspect of the spanning property is that it is preserved by restrictions of the correspondence to subsets. Lemma 2 of Simon, Spiez, and Torunczyk [12] states that if F is spanning for an open A and D is any open subset of A then $F|\overline{D}$ is also spanning for $\overline{D}$.

Lemma 7: If a correspondence F has the spanning property for an open and bounded set $U \subseteq \mathbf{R}^n$ and C is a connected and compact subset of U then for every pair $x, y \in C$ there is some $z_1 \in F(x)$ and some $z_2 \in F(y)$ such that (x, z_1) and (y, z_2) are in the same connected component of $F|C$.

Proof: Let U_i be a decreasing sequence of open, bounded and connected subsets of U converging to C, meaning that $C = \cap_{i=1}^{\infty} \overline{U_i}$. Since the U_i are connected by Lemma 2 of Simon, Spiez and Torunczyk [12] there are connected and compact subsets Z_i of F such that the correspondence Z_i is spanning for U_i. Due to Simon, Spiez, and Torunczyk [11] for every i there are pairs (x, a_i) and (y, b_i) in Z_i. Because the U_i is a decreasing sequence of sets, again due to Lemma 2 of Simon, Spiez, and Torunczyk ([12]) we can assume without loss of generality that the Z_i is also a decreasing sequence of sets. Define Z to be the intersection of the Z_i. Because the Z_i are connected and compact, Z is also connected and compact. By its compactness Z contains a pair (x, a) and (y, b) for some a and b as limits, respectively, of some subsequences of the a_i and b_i. □

5.3 The Structure Theorem

A *homotopy* is a continuous map $h : X \times [0, 1] \to Y$, where X and Y are topological spaces. If Y can be embedded in a convex space then the homotopy $h : X \times [0, 1] \to Y$ is a *straight line* homotopy if for every $x \in X$ and $t \in [0, 1]$ $h(x, t) = t\ h(x, 1) + (1 - t)h(x, 0)$.

There is a strong connection between quitting games and another area of game theory usually not associated with stochastic games – structure theorems used to establish stability properties of one-shot games. We remind the readers of the main theorem of Kohlberg and Mertens, [3]. Let N be a finite player set, $(A^j \mid j \in N)$ the finite sets of actions for the players, X the space of all $|A^1| \times \ldots \times |A^{|N|}|$ matrices with vector payoff entries from $\mathbf{R}^N$. For any $x \in X$ let G_x be the one stage game defined by the matrix x. Let $\tilde{A}$ be $\prod_{j \in N} \Delta(A^j)$, the strategy space, (where $\Delta(A^j)$ is the simplex of probability distributions on A^j). Let $E \subseteq X \times \tilde{A}$ be the correspondence defined by $E(x) := \{y \in \tilde{A} \mid y \text{ is an equilibrium of the game } G_x\}$. Let $\pi : X \times \tilde{A} \to X$ be the canonical projection. The structure theorem of Kohlberg and Mertens states that there is a straight line homotopy $H(\cdot, \cdot)$ from $X \times [0,1]$ to $X \times \tilde{A}$ such that $\pi \circ H(x,0) = x$ for all $x \in X$, the image of $H(\cdot, 1)$ is exactly the correspondence E, and the function H can be extended continuously to the one-point compactification of X (meaning that for every compact set $C \subseteq X$ there is an $R > 0$ large enough that if the norm $||x||$ exceeds R then for all $t \in [0,1]$ the point $H(x,t)$ does not lie over C). Here we have slightly modified the statement of the structure theorem, using the fact that $\tilde{A}$ is convex.

5.4 Finitely Repeated Quitting Games

For every $k \geq 0$ and vector $x \in \mathbf{R}^N$ let Γ^k_x be the k stage game such that at the conclusion of k stages the players receive the payoff x given that all players chose c on all stages. A strategy in Γ^k_x is a sequence $p = (p_0, p_1, \ldots, p_{k-1}) \in ([0,1]^N)^k$ representing the probabilities that the players would quit on the various stages. Let $E^k \subseteq \mathbf{R}^N \times ([0,1]^N)^k$ be the equilibrium correspondence of the games Γ^k, meaning that $E^k(x)$ are the equilibria of Γ^k_x. For all $k \geq 1$ define the function $f^k : \mathbf{R}^N \times ([0,1]^N)^k$ by $f^1(x,p) = f(x,p)$ and $f^k(x,p) = f(f^{k-1}(x,(p_1, \ldots, p_{k-1})), p_0)$. Define the correspondences $F^k \subseteq \mathbf{R}^N \times \mathbf{R}^N$ by $F^k(x) = \{f^k(x,p) \mid p \in E^k(x)\}$. F^k is also the k iteration of the correspondence F_0.

Lemma 8: Let a quitting game be fixed. If $k \geq 1$ and x and y belong to a connected and compact subset D of $\mathbf{R}^N$ such that no equilibrium on the set D involves some player quitting with certainty then there is a pair $p^x = (p^x_1, p^x_2, \ldots, p^x_k)$ and $p^y = (p^y_1, p^y_2, \ldots, p^y_k)$ in $([0,1]^N)^k$ with $(x, p^x) \in E^k$ and $(y, p^y) \in E^k$ such that (x, p^x) and (x, p^y) belong to the same connected component of $E^k|D$, the equilibrium correspondence lying over D.

Proof: We represent the k repeated quitting game Γ^k_r as a game with the original set N of players and finitely many actions. Let k, the number of stages, be fixed, and let each player n have the finite set $A^n_k = \{c, q_1, \ldots, q_k\}$ of actions. The action q_j means that the player will choose c on all stages up to stage j and then choose q on stage j. If no player chooses q_j for any $1 \leq j \leq k$ then the players will receive the payoff r. Otherwise, let j be the first stage such that some player chooses q_j and the payoff to the players will be $v(A)$ where A is the set of players who choose q_j.

The actions A^n_k define games with variable payoff matrices. As above, define X to be the space of all $k+1 \times k+1 \times \ldots \times k+1$ payoff matrices, G_x the corresponding game for every $x \in X$, $\tilde{A}$ the space of mixed strategies, and $E \subseteq X \times \tilde{A}$ the equilibrium correspondence. Let $H(\cdot, \cdot)$ from $X \times [0,1]$ to $\tilde{A}$ be the straight line homotopy (Kohlberg and Mertens [3]) as described above such that the image of $H(\cdot, 1)$ is the equilibrium correspondence E. Define $\hat{i} : \mathbf{R}^n \to X$ so that the $(c, c, \ldots, c)$ coordinate of $\tilde{i}(r)$ is equal to r and otherwise the other coordinates of $\tilde{i}(r)$ are independent of the choice of r and correspond to the appropriate $v(A)$ defining the quitting game where A is the set of players choosing q_l where l is the smallest number such that no player chose q_i for all $i < l$. Let $\tilde{D}$ be the image $\tilde{i}(D)$. Let $R > 0$ be large enough so that if $||r||$ exceeds R then for all $t \in [0,1]$ the point $H(r,t)$ projected to X does not lie in $\tilde{D}$. Define a function $b_R : X \to [0,1]$ by $b_R(r) = 0$ if $||r|| \geq R+1$, $b_R(r) = 1$ if $||r|| \leq R$, and otherwise $b_R(r) = R + 1 - ||r||$ if $R \leq ||r|| \leq R+1$. Define a continuous function $h : X \to X \times \tilde{A}$ by $h(r) = H(r, b_R(r))$. The correspondence $h(\{r \mid ||r|| \leq R+2\}) \subseteq X \times \tilde{A}$ has the spanning property for $\{r \mid ||r|| \leq R+2\}$ (because the projection to X of h on $\{r \mid ||r|| = R+2\}$ is the identity function). By our choice of R this same correspondence $h(\{r \mid ||r|| \leq R+2\})$ over the set $\tilde{D}$ is the equilibrium correspondence E over the set $\tilde{D}$. As D is compact and therefore there is a maximal probability $\rho < 1$ that any player quits in any equilibrium E^k over D, $E|\tilde{D}$ is topologically equivalent to $E^k|D$. The rest follows by Lemma 7. □

5.5 Escape Games Have Approximate Equilibria

We fix an escape game that does not have stationary approximate equilibria nor instant approximate equilibria, and let $\rho > 0$ be a quantity defined by Lemma 5 and let $\epsilon > 0$ be strictly smaller than either the $\bar{\epsilon} > 0$ defining the escape game or ρ. All claims that follow refer to this game.

Lemma 9: There is a quantity $B > 0$ so large that if $x^j \geq B$ for all $j \in N$ then there is only one equilibrium in E^k, namely $\underline{0}$, the equilibrium where no player chooses q with positive probability on any stage.

Proof: By induction, it suffices to prove this for E_0. Since every equilibrium involves a probability of at least ρ that no player chose the action q, it suffices that B is larger than $\frac{M+1}{\rho} + \max_{n \in N} \chi^n$. □

Define the positive quantity δ to be $\frac{\epsilon}{10M|N|}$. Define $T := \{x \mid v^j \leq x^j \leq v^j + \epsilon \text{ for some } j \in N\} \cap \{x \mid x^j \geq v^j \text{ for all } j \in N\}$.

Define the correspondence $\tilde{F}_{j,\delta}$ to be $\{(x,y) \mid x \in T,\ x^j \leq v^j + \epsilon,\ y = f(x,p) \text{ for some } p \text{ satisfying } 0 \leq p^j \leq \delta \text{ and } p^k = 0 \text{ for all } k \neq j\}$. Define $\tilde{F}_\delta := F_0 \cup_{j \in N} \tilde{F}_{j,\delta}$.

Lemma 10: $\tilde{F}_\delta \subseteq F_\epsilon$ and if an extended orbit of $\tilde{F}_\delta$ starts at a point in $\{x \mid x^j \geq \chi^j - \epsilon\}$ then it remain in this set. If the extended orbit of $\tilde{F}_\delta$ started at a point in Q then it remains in Q, and if it starts in $Q \backslash (W \cup T)$ then it remains in $Q \backslash (W \cup T)$.

Proof: Assume that $x \in T$ with $x^j \leq v^j + \epsilon$ and $p \in \tilde{F}_{j,\delta}$. By quitting alone Player j gets a payoff of v^j and by not quitting a payoff of x^j. By not quitting any other player $k \neq j$ gets a payoff of at least $v^k - \delta M$ and by quitting a payoff no better than $v^k + \delta M \leq v^k + \epsilon/10$. This completes the proof of $\tilde{F}_\delta \subseteq F_{\epsilon/3}$. Staying in the set $\{x \mid x^j \geq \chi^j - \epsilon\}$ is the result of Lemma 6.

Containment in Q follows by the containment of $\tilde{F}_\delta$ in $F_{\epsilon/3}$, the definition of an escape game, and the closure of Q.

We assumed that ϵ is smaller than the $\overline{\epsilon} > 0$ defining the escape game properties. Assume that $x, y \in Q$ with $x \notin W \cup T$ and $y \in \tilde{F}_\delta(x)$. Since x is already outside of T we know that $y \in F_0(x)$. By the definition of an escape game either $y = x$ or $y^j > v^j + \overline{\epsilon}$ for all $j \in N$. Since $\overline{\epsilon}$ is larger than ϵ a convergence to a point in T is not possible. □

Define an $x \in \partial W$ to be *critical* if there exists a pair of player j, k in N such that $x^j = v^j$, $x^k = v^k$, and $v(\{j\})^k < v^k$.

Lemma 11: From any start at a point in T there is a finite orbit of $\tilde{F}_\delta$ staying in T that ends at a critical point.

Proof: It follows directly from Lemma 5 and the intermediate value theorem. □

Theorem 4: All escape games have approximate equilibria.

Proof: It suffices to prove the claim for escape games without stationary or instant approximate equilibria, and therefore for an escape game with the assumptions made above. By Theorem 3 and Lemma 11 it suffices to show that starting at any critical point $x \in Q \cap \partial W$ either there is an extended orbit of $\tilde{F}_\delta$ in $(W \cup T) \cap Q$ with unbounded total variation or there is an orbit of $\tilde{F}_\delta$, finite or infinite, ending or converging to some member of $T \cap Q$ with total variation of at least $\delta\rho/3$.

Given a critical point $x \in Q \cap \partial W$ with $x^j = v^j$, $x^i = v^i$ and $v(\{j\})^i < v^i$ let $y = f(x, p)$ with $p^j = \delta$ and $p^k = 0$ for all $k \neq j$. Lemma 10 implies that y is in Q.

Case 1; there is an infinite orbit of F_0 starting at y and contained in $W \cup T$ that does not converge: By the definition of an escape game the orbit is in Q and non-convergence implies unbounded total variation.

Case 2; there is an infinite orbit of F_0 starting at y and contained in $W \cup T$ that does converge: Convergence to a point z in the interior of W is impossible, since a distance of $t > 0$ from the boundary of W implies that any equilibrium of Γ_z involves a probability of quitting of at least t/M, and by Lemma 5 this would also mean a motion of at least $\rho t/M$ away from this point. With the assumption that the orbit converges to a point in T, a total variation of at least $\rho\delta$ is obtained in the motion from x to y. The convergence point is in Q because Q is closed.

Case 3; there is no infinite orbit of F_0 starting at y and contained in $W \cup T$:

There must be a k such that $F^k(y)$ is contained in the complement of $W \cup T$, since otherwise by the closure of the correspondence F_0 the existence of a finite orbit of F_0 of length k contained in $W \cup T$ for every k would imply the existence an infinite orbit of F_0 in $W \cup T$, (an easy exercise, also see McGehee [8]).

Let $p = (p_0, \ldots, p_{k-1})$ be any equilibrium in $E^k(y)$. Let B be a positive quantity given by Lemma 9 and let $\overline{x}$ be a point satisfying $\overline{x}^j > v^j$ for all $j \in N$ such that the closed line segment between x and $\overline{x}$ is in Q. Consider three line segments, that from y to x, that from x to $\overline{x}$, and that from $\overline{x}$ to the point $z := (B, B, \ldots, B)$; define D to be the union of these three line segments. Define a complete ordering on D in the natural way so that

$z > x > y$. By Lemma 8 $(z, \underline{0})$ and (y, p) must be in same connected component of $E^k|D$ (as Lemma 9 implies that $\underline{0}$ is the only member of $E^k(z)$). Let $\tilde{x} \in D$ be any point with $x < \tilde{x} \leq \overline{x}$. If $(\tilde{x}, \hat{p}) \in E^k$ and with $\hat{p}$ there is a positive probability that some player chooses q then from the definition of an escape game this probability is at least $\overline{\epsilon}/M$. Furthermore there will be a positive constant $c > 0$ such that if $\hat{x} \in D$ satisfies $\overline{x} \leq \tilde{x} \leq z$ then the distance from $\hat{x}$ to W is at least c, and therefore if $\hat{p} \in E^k(\hat{x})$ then the probability that some player chooses q will be at least c/M. Let $d = \frac{1}{2M} \min\{c, \epsilon\}$.

It suffices to show that there is a finite orbit of F_0 of length k starting at $\hat{y} \in D$ with $y \leq \hat{y} \leq x$ and ending at some $\hat{z}$ with $\hat{z}^j = v^j + \epsilon$ for some $j \in N$. By Lemma 10 all points in this orbit are in $Q \cap (W \cup T)$ and by Lemma 11 there is a return to the set $\partial W \cap Q$. A total variation of at least ϵ is obtained on the return to T.

Suppose for the sake of contradiction that there is no finite orbit of F_0 of length k starting at any $\hat{y} \in D$ with $y \leq \hat{y} \leq x$ and ending at some $\hat{z}$ with $\hat{z}^j = v^j + \epsilon$ for some $j \in N$. Define open subsets O_1, O_2 of $D \times ([0,1]^N)^k$ by $O_1 := \{(\tilde{x}, \tilde{p}) \mid \tilde{x} > x,\ \sum_{i=0}^{k-1} q(\tilde{p}_i) < d\}$ and $O_2 := \{(\tilde{x}, \tilde{p}) \mid \tilde{x} < \overline{x},\ \min_{n \in N} (f^k(\tilde{x}, \tilde{p}))^n - v^n < \epsilon\}$. Let O be $O_1 \cup O_2$. By assumption ∂O, the boundary of the open set O, contains no members of $E^k|D$, (here $\epsilon < \overline{\epsilon}$ is used). The point $(z, \underline{0})$ is in O and the point (y, p) is in $(D \times ([0,1]^N)^k) \backslash \overline{O}$, which means that $(z, \underline{0})$ and (y, p) lie in two distinct connected components of $E^k|D$, a contradiction. □

6 Conclusion

How could one extend the proof of Theorem 4 to a proof for all quitting games? What would a quitting game counter-example look like (to approximate equilibria) if one existed? Infinite total variation is an analytic property, orbit existence an algebraic property, and it seems to be a coincidence that there was a synthesis for a proof of Theorem 4. On the other hand, any counter-example must fail to be an escape game, which means that over some x outside of W there are equilibria involving a positive probability of quitting. Since for any point x outside of W the set $F_0(x)$ contains at least x from the equilibrium $\underline{0}$ and generically $F_0(x)$ has an odd number of points we must presume that for many such x the set $F_0(x)$ contains at least two points. Any analysis of orbits for a candidate counter-example must involve multiple choices for some of the vectors reached.

Looking beyond quitting games, the situation doesn't look any better for finding a counter-example. Stochastic games are played on infinitely many stages, and therefore in general the game trees branch wildly. Quitting games are designed to prevent rapidly growing ways that a player could respond to the past behavior of the other players. In our opinion, to find a counter-example one would fair a better chance staying with quitting games that are not escape games.

The step from perfection to approximate equilibria is well founded for quitting games, however in general the scope of Theorem 2 is very limited. It seems that a proof for the existence of approximate equilibria in all normal stochastic games must tackle this problem – given an $\epsilon > 0$ how can some kind of stage-for-stage $\delta > 0$ equilibria translate to the existence of $\epsilon > 0$ equilibrium? Example 1 is highly discouraging, and perhaps an integration of a variation of this example into a game with a multitude of players could be the basis of a counter-example.

7 References

[1] EILENBERG, S. and STEENROD, N. (1952). *Foundations of Algebraic Topology*, Princeton University Press.

[2] FLESCH, J., THUIJSMAN, F., and VRIEZE, O.J. (1997). "Cyclic Markov Equilibria in a Cubic Game", *International Journal of Game Theory*, **26**, pp. 303-314.

[3] KOHLBERG, E. and MERTENS, J.-F. (1986). "On the Strategic Stability of Equilibria", *Econometrica,* **54**, No. 5, 1003-1037.

[4] KUHN, H. W. "Extensive Games and the Problem of Information", in Contribution to the Theory of Games II, *Annals of Mathematics Studies* **28** (1953), pp. 193-216.

[5] MAITRA, A. and SUDDERTH, W. (1991). "An Operator Solution of Stochastic Games", *Israel Journal of Mathematics*, **78**, pp. 33-49.

[6] MARTIN, D. (1998). "The Determinacy of Blackwell Games", *The Journal of Symbolic Logic,* **63**, pp. 1565-1581.

[7] MERTENS, J.F. and NEYMAN, A. (1981). "Stochastic Games", *International Journal of Game Theory*, **10**, 53-66.

[8] McGEHEE, R. (1992). "Attractors for Closed Relations on Compact Hausdorff Spaces", *Indiana University Mathematics Journal*, **41**, No. 4, pp. 1165-1209.

[9] NASH, J. (1950). " Equilibrium Points in n-Person Games", Proceedings of the National Academy of Sciences, **36**, pp. 48-49.

[10] SIMON, R., (2006). "Value and Perfection in Stochastic Games", to appear *Israel Journal of Mathematics.*

[11] SIMON, R., SPIEŻ, S., TORUŃCZYK, H. (1995). "The Existence of Equilibria in Certain Games, Separation for Families of Convex Functions and a Theorem of Borsuk-Ulam Type", *Israel Journal of Mathematics*, **92**, pp. 1-21.

[12] SIMON, R., SPIEŻ, S., and TORUŃCZYK, H. (2002). "Equilibrium Existence and Topology in Some Repeated Games of Incomplete Information", *Transactions of the American Mathematical Society*, **354**, No. 12, pp. 5005-5026.

[13] SOLAN, E. and VIEILLE, N. (2001). "Quitting Games", *Mathematics of Operations Research*, **26**, pp. 265-285.

[14] VIEILLE, N. (2000). "Two-Player Stochastic Games I: A Reduction", *Israel Journal of Mathematics*, **119**, pp. 55-91, "Two-Player Stochastic Games II: The Case of Recursive Games", *Israel Journal of Mathematics*, **119**, pp. 92-126, "Small Perturbations and Stochastic Games", *Israel Journal of Mathematics*, **119**, pp. 127-142.

[15] WILLIAMS, D. (1991). *Probability with Martingales*, Cambridge University Press.