Rationalizability and Approximate Common-Knowledge

Jeffrey C. Ely*

November 20, 2001

Abstract

Given a strategic form game G, we can derive the set of rationalizable action profiles $\Lambda^\infty(G)$. We can also derive the set of all action profiles which are rationalizable in some incomplete information environment in which it is approximate common-knowledge that the payoffs are as they are in G. Call this set $\overline{\Lambda^\infty}(G)$. This paper provides a formal definition of $\overline{\Lambda^\infty}(G)$ where the concept of approximate common-knowledge is that of Monderer and Samet (1989). In general $\overline{\Lambda^\infty}(G)$ can be a proper subset of $\Lambda^\infty(G)$, but the two sets are equal for compact and continuous G.

*`ely@nwu.edu`. Financial support from NSF grant #9985462 is gratefully acknowledged.

1 Introduction

Standard solution concepts in game theory are built on the assumption, either implicit or explicit, that the structure of the game to be played is common knowledge among the players. The Nash equilibrium concept applied to games with complete information entails an implicit assumption that the payoffs are common knowledge. The weaker concept of rationalizability follows from the explicit assumption that the payoffs and the rationality of the players are both common knowledge (Tan and da Costa Werlang (1988)).

Precise knowledge of anything is already a rare condition. Common knowledge, defined as mutual knowledge of arbitrary order, is *a fortiori* a theoretical ideal. It is therefore natural that researchers have investigated whether these standard concepts truly rely on exact common knowledge, or whether their conclusions can be expected to hold (perhaps in the approximate) in environments in which rationality and payoffs are nearly common knowledge. As it pertains to equilibrium concepts, the pioneering study of this type of question is Rubinstein (1989), who shows that Nash equilibrium may produce solutions for complete information games (i.e. common knowledge of payoffs) which need not be (Nash equilibrium) solutions when those payoffs are mutual knowledge of arbitrary finite order. Subsequent work which has investigated this apparent discontinuity includes Carlsson and van Damme (1993), Kajii and Morris (1997) and Kajii and Morris (1998). An important paper in this literature is Monderer and Samet (1989) who provide a formal measure of the closeness of an information structure to exact common knowledge.

To similarly determine the robustness of the non-equilibrium concept of rationalizability would appear to be at least as important as this concept is explicitly defined as the behavioral implications of common knowledge of rationality and payoffs. That is, the set of rationalizable strategies of a complete information game are all those that can be played in a state of the world in which all players are rational and know the game's payoffs and this is common knowledge (Tan and da Costa Werlang (1988)).

This paper shows by example that in general the set of rationalizable strategies may be unacceptably sensitive to small deviations from exact common knowledge. There can be strategies which are rationalizable under every information structure arbitrarily close to common knowledge, but which are not rationalizable under exact common knowledge. [1] The example has infinitely many actions and a payoff function that is continuous in each player's action independently but has a "diagonal discontinuity." On the other hand, it is shown that when action spaces are compact and the payoff function jointly continuous, rationalizability under approximate common knowledge coincides with rationalizability under exact common knowledge. The latter result generalizes the observation made in Dekel and Gül (1997) pertaining to finite games.

Lipman (1994) concerns an issue that is closely related. The relationship between the example in Lipman (1994) and the present paper is discussed in Section 3.

2 Setup

A strategic form game G with N players is a pair (A, π), where $A = \times_{i \in N} A_i$ is the set of action profiles $a = (a_1, \ldots, a_N)$ with $a_i \in A_i$, and $\pi : A \to \mathbf{R}^N$ is the payoff function assigning a utility vector $\pi(a) = (\pi_1(a), \ldots, \pi_N(a))$ to each action profile a. We assume that each A_i is a topological space and give products of these spaces the product topology. For $K \subset A$, we denote by $\Delta_i(K)$ the set of Borel probability measures, called conjectures, over A_{-i} with support in K_{-i}. In general, we use the notation $C(\phi)$ to denote the support of a conjecture ϕ. We can extend the payoff function π_i to conjectures by writing $\pi_i(a_i, \phi_i)$ for the expectation of $\pi_i(a_i, \cdot)$ with respect to ϕ_i.

For a given $K \subset A$, the set $R_i(K)$ is defined by

$$R_i(a_i|K) = \{\phi_i \in \Delta_i(K) : a_i \in \text{argmax}_{z_i \in A_i} \pi_i(a_i, \phi_i)\}.$$

This is the set of *rationalizations* of a_i over K_{-i}. The operator Λ is defined as follows. An action profile a belongs to $\Lambda(K)$ if $R_i(a_i|K) \neq \emptyset$ for each i.

[1] Note that while Rubinstein (1989) and related examples demonstrate a possible failure of lower-hemicontinuity of the equilibrium correspondence, the violation here is of upper-hemicontinuity. Failure of upper-hemicontinuity is a potentially more serious defect when it is important for a planner to ensure that social outcomes belong to some restricted subset. Chung and Ely (????) discusses the significance for implementation theory. See also Fudenberg, Kreps, and Levine (1988) for related problems with upper-hemicontinuity of equilibrium refinements.

We set $\Lambda^0 = A$ and define inductively $\Lambda^n = \Lambda(\Lambda^{n-1})$. Finally,

$$\Lambda^\infty = \bigcap_{n=1}^{\infty} \Lambda^n$$

is the set of *rationalizable* action profiles.[2]

An *elaboration* of a strategic form game G is a Bayesian game Γ consisting of a set $T = \times_i T_i$ of type profiles; a utility function $u : A \times T \to \mathbf{R}^N$ satisfying for each t in some non-empty set of types T^0, $u(a,t) = \pi(a)$ for all $a \in A$; and finally a probability measure μ on T. We assume there exists a uniform bound B such that $u_i(a,t) - u_i(z,t) < B$ for every a, z, t and i for all u arising from elaborations. This is the only restriction on the types of elaborations considered.

We can consider the strategic form of an elaboration, with Σ_i^Γ denoting the set of strategies for player i, and calculating payoffs by taking expectations of u with respect to μ. A conjecture for player i is then a probability measure φ_i over Σ_{-i}^Γ.If $K \subset A$, and φ_i a conjecture for i, then in an abuse of notation, we denote by $\varphi_i^{-1}(K)$ the set

$$\varphi_i^{-1}(K) = \bigcap_{\sigma_{-i} \in C(\varphi)} \sigma_{-i}^{-1}(K_{-i})$$

If φ is a profile of conjectures, we write $\varphi^{-1}(K) = \cap_i \varphi_i^{-1}(K)$. The set $\varphi^{-1} \subset T$ is the set of type profiles at which according to φ, it is the players' mutual conjecture that the action profile will be in K. For any subset of strategy profiles S, Let $\Delta_i S$ be the set of conjectures for i with support included in S_{-i}, and ΔS the set of profiles of such conjectures.

For any event $E \subset T$, let $B_\Gamma^p(E)$ be the set of type profiles $t \in T$ such that $\mu(E|t_i) > p$ for all i, i.e. that the event E is mutual p-belief among the players. Then we say that E is common p-belief at a type profile t if

$$t \in C_\Gamma^p(E) := \cap_k B_\Gamma^p(E).$$

Monderer and Samet (1989) introduced the concept of common p-belief as an approximation to common-knowledge, which is common 1-belief.[3] They

[2]Bernheim (1984) defines the set of rationalizable action profiles in a different way and proves that for compact and continuous games, his definition is equivalent to the one given here. See Section 3

[3]modulo zero probability events.

provide a characterization of $C^p_\Gamma(E)$ in terms of p-evident events, a generalization of self-evident events. For our purposes, the following implication will be useful.

Proposition 1 *The event $C^p_\Gamma(E)$ is p-evident. That is, $\mu(C^p_\Gamma(E)|t_i) > p$ for each $t_i \in C^p_\Gamma(E)$.*

Write $C^{1-\varepsilon}_\Gamma$ for the set of type profiles in Γ at which T^0 is common $1-\varepsilon$ belief. We record the following simple fact about the operator B^p_Γ.

Lemma 1 $B^p_\Gamma(E) \cap B^p_\Gamma(F) \subset B^{2p-1}_\Gamma(E \cap F)$

We now define the operator Λ_ε. An action profile a belongs to $\Lambda_\varepsilon(K)$ iff there exists an elaboration Γ of G, a type profile t, and a profile of conjectures φ such that $t \in B^{(1-\varepsilon)}_\Gamma(\varphi^{-1}(K) \cap T^0)$ and a is an interim best-reply for t against φ. In words, Λ_ε is the set of action profiles that can be rationalized when it is mutual $1-\varepsilon$ belief that the payoffs are as in T^0 and that the action profile will be from K.

We now inductively define $\Lambda^0_\varepsilon = A$ and $\Lambda^n_\varepsilon = \Lambda_\varepsilon(\Lambda^{n-1}_\varepsilon))$. Finally,

$$\Lambda^\infty_\varepsilon = \bigcap_n \Lambda^n_\varepsilon$$

is the set of ε-rationalizable action profiles.

Lemma 2 *For any $\varepsilon \geq 0$, $\varepsilon' > \varepsilon$ and n, $\Lambda^n_\varepsilon \subset \Lambda^n_{\varepsilon'}$.*

Proof: Because $B^{(1-\varepsilon)}_\Gamma(E) \subset B^{1-\varepsilon'}_\Gamma(E)$ for any E, it follows immediately from the definition that $\Lambda_\varepsilon(K) \subset \Lambda_{\varepsilon'}(K)$ for any $K \subset A$. Moreover, for any $K \subset K'$, and profile of conjectures φ, $\varphi^{-1}(K) \subset \varphi^{-1}(K')$ and hence $\Lambda_\varepsilon(K) \subset \Lambda_\varepsilon(K')$ for any $\varepsilon \geq 0$. The first claim implies that $\Lambda^1_\varepsilon = \Lambda_\varepsilon(A) \subset \Lambda_{\varepsilon'}(A) = \Lambda^1_{\varepsilon'}$. Suppose now the lemma is true for $n-1$. Then

$$\Lambda^n_\varepsilon = \Lambda_\varepsilon(\Lambda^{n-1}_\varepsilon) \subset \Lambda_\varepsilon(\Lambda^{n-1}_{\varepsilon'}) \subset \Lambda_{\varepsilon'}(\Lambda^{n-1}_{\varepsilon'}) = \Lambda^n_{\varepsilon'}$$

■

The following proposition clarifies why we are interested in $\Lambda^\infty_\varepsilon$.

Proposition 2

1. *$\Lambda^\infty = \Lambda^\infty_0$*
2. *If $\sigma \in \Lambda^\infty(\Gamma)$ for some Γ, then $\sigma(t) \in \Lambda^\infty_\varepsilon$ for all $\varepsilon > 0$ and $t \in C^{1-\frac{\varepsilon}{2}}_\Gamma$.*

Proof: The first part is straightforward.

For the second part, we prove the following claim by induction.

$$\sigma \in \Lambda_\Gamma^n \implies C_\Gamma^{1-\frac{\varepsilon}{2}} \subset \sigma^{-1}(\Lambda_\varepsilon^n) \quad \forall\, \varepsilon > 0$$

It will follow from this that $\sigma \in \Lambda^\infty(\Gamma) \subset \Lambda_\Gamma^n$ for all $n \implies C_\Gamma^{1-\frac{\varepsilon}{2}} \subset \sigma^{-1}(\Lambda_\varepsilon^n)$ for all $n \implies C_\Gamma^{1-\frac{\varepsilon}{2}} \subset \cap_n \sigma^{-1}(\Lambda_\varepsilon^n) = \sigma^{-1}(\cap_n \Lambda_\varepsilon^n) = \sigma^{-1}(\Lambda_\varepsilon^\infty)$, which would prove the proposition.

Since $\Lambda_\varepsilon^0 = A$ and $\Lambda_\Gamma^0 = \Sigma_\Gamma$, the claim is trivial for $n = 0$. Suppose now the claim is true for $n-1$, and let $\sigma \in \Lambda_\Gamma^n$. Then there exists a profile of conjectures $\varphi \in \Delta\Lambda_\Gamma^{n-1}$ such that $\sigma \in \beta(\varphi)$.

Let $t \in C_\Gamma^{1-\frac{\varepsilon}{2}}$. Then $t \in B_\Gamma^{1-\frac{\varepsilon}{2}}(T^0)$. And since $C_\Gamma^{1-\frac{\varepsilon}{2}}$ is $1-\frac{\varepsilon}{2}$-evident, $t \in B_\Gamma^{1-\frac{\varepsilon}{2}}(C_\Gamma^{1-\frac{\varepsilon}{2}})$. By the induction hypothesis, $C_\Gamma^{1-\frac{\varepsilon}{2}} \subset \tilde{\sigma}^{-1}(\Lambda_\varepsilon^{n-1})$ for every $\tilde{\sigma} \in \Lambda_\Gamma^{n-1}$, and since $C(\varphi_i) \subset [\Lambda_\Gamma^{n-1}]_{-i}$ for each i, we have $C_\Gamma^{1-\frac{\varepsilon}{2}} \subset \varphi^{-1}(\Lambda_\varepsilon^{n-1})$. Thus $t \in B_\Gamma^{1-\frac{\varepsilon}{2}}(T^0) \cap B_\Gamma^{1-\frac{\varepsilon}{2}}(\varphi^{-1}(\Lambda_\varepsilon^{n-1}))$. By lemma 1,

$$t \in B_\Gamma^{(1-\varepsilon)}(T^0 \cap \varphi^{-1}(\Lambda_\varepsilon^{n-1}))$$

and since $\sigma(t)$ is an interim best-reply for t, we conclude that $\sigma(t) \in \Lambda_\varepsilon(\Lambda_\varepsilon^{n-1}) = \Lambda_\varepsilon^n$. ■

The *closure* of the rationalizability solution concept, applied to G is then defined as follows.

$$\overline{\Lambda^\infty}(G) = \bigcap_{\varepsilon>0} \Lambda_\varepsilon^\infty$$

Requiring a strategy profile to be ε-rationalizable for every $\varepsilon > 0$ may not *a priori* seem the only natural way to define rationalizability with approximate common-knowledge. Two equally natural alternative limit sets are:

$$\liminf_{\varepsilon\to 0} \Lambda_\varepsilon^\infty = \bigcup_{\varepsilon>0} \bigcap_{0<\varepsilon'<\varepsilon} \Lambda_{\varepsilon'}^\infty$$

and

$$\limsup_{\varepsilon\to 0} \Lambda_\varepsilon^\infty = \bigcap_{\varepsilon>0} \bigcup_{0<\varepsilon'<\varepsilon} \Lambda_{\varepsilon'}^\infty$$

It turns out that the monotonicity of of Λ_ε implies that these three alternative definitions are equivalent.

Proposition 3 $\bigcap_{\varepsilon>0} \Lambda_\varepsilon^\infty = \liminf_{\varepsilon\to 0} \Lambda_\varepsilon^\infty = \limsup_{\varepsilon\to 0} \Lambda_\varepsilon^\infty$

Proof: First observe that $\sigma \in \limsup \Lambda_\varepsilon^\infty$ iff for every $\varepsilon > 0$ there exists $\varepsilon' \in (0,\varepsilon)$ such that $\sigma \in \Lambda_{\varepsilon'}^\infty$, and $\sigma \in \liminf \Lambda_\varepsilon^\infty$ iff there exists $\varepsilon > 0$ such that $\sigma \in \Lambda_{\varepsilon'}^\infty$ for all $\varepsilon' \in (0,\varepsilon)$. Clearly $\liminf \Lambda_\varepsilon^\infty \subset \limsup \Lambda_\varepsilon^\infty$. It is also obvious that $\cap_\varepsilon \Lambda_\varepsilon^\infty \subset \liminf \Lambda_\varepsilon^\infty$. Thus, it suffices to show $\limsup \Lambda_\varepsilon^\infty \subset \cap_\varepsilon \Lambda_\varepsilon^\infty$.

Let $\varepsilon > 0$. If $\sigma \in \limsup \Lambda_\varepsilon^\infty$, then there exists $\varepsilon' \in (0,\varepsilon)$ such that $\sigma \in \Lambda_{\varepsilon'}^\infty$. Since $\Lambda_{\varepsilon'}^\infty = \cap_n \Lambda_{\varepsilon'}^n$ and $\Lambda_{\varepsilon'}^n \subset \Lambda_\varepsilon^n$ for every n by Lemma 2, we have $\Lambda_{\varepsilon'}^\infty \subset \cap_n \Lambda_\varepsilon^n = \Lambda_\varepsilon^\infty$. Thus, $\sigma \in \Lambda_\varepsilon^\infty$. ■

We show by example below that there can be strategies that are rationalizable in all nearby elaborations, and yet not rationalizable in G. For such games, the rationalizability loses some of its force. Fortunately, when G is compact and continuous, these sets are equal, i.e. no action which is ruled out by rationalizability can be rationalized in nearby incomplete information environments, where "nearby" is in the sense of Monderer and Samet (1989) approximate common-knowledge of payoffs.

Theorem 1 *If G is compact and continuous, then $\overline{\Lambda^\infty}(G) = \Lambda^\infty$.*

Let ΔA_{-i} be the compact space of Borel probability measures on A_{-i} with the topology of weak convergence. For any action a_i and conjecture $\phi \in \Delta A_{-i}$ define

$$f_i(a,\phi) = \max_{z_i \in A_i} \pi_i(z_i,\phi) - \pi_i(a_i,\phi).$$

By the compactness of ΔA_{-i} and the continuity of π_i this function well-defined and continuous.

Lemma 3 *$a \in \Lambda_\varepsilon(K)$ for some $K \subset A$ if and only if for each i, there is a conjecture $\phi_i \in \Delta K_{-i}$ such that*

$$f_i(a,\phi_i) \le \frac{\varepsilon B}{1-\varepsilon}$$

Proof: For sufficiency, we construct an elaboration Γ. There are two types t_i^0 and $\tilde{t}_i$ for each i. The prior distribution is such that conditional on being type t_i^0, each player assigns probability $(1-\varepsilon)$ to the true state being t^0, and the remaining probability to the event that all other players j are of type $\tilde{t}_j$. We set $u(\cdot, t^0) = \pi(\cdot)$ and for any type profile t other than t^0,

$$u_i(z_i, a_{-i}, t) = B\left[1 - \frac{\max\{0, \pi_i(z_i, \phi_i) - \pi_i(a_i, \phi_i)\}}{f_i(a_i, \phi_i)}\right]$$

for each z_i, a_{-i}. This payoff function satisfies the boundedness restriction and the set $T^0 = \{t^0\}$ is $(1-\varepsilon)$ evident and hence $t^0 \in C_\Gamma^{1-\varepsilon}$. We will show that action a_i is interim rationalizable for type t_i^0 in this elaboration. For each $a_{-i} \in K_{-i}$, let σ_{-i}^a denote the strategy profile in Γ where $\sigma_j(t_j^0) = \sigma_j(\tilde{t}_j) = a_j$ for each $j \neq i$. Consider the conjecture φ_i for i that his opponents choose strategies from $\{\sigma_{-i}^a : a_{-i} \in K_{-i}\}$ according to the measure ϕ_i. The difference in expected utility to type t_i^0 from choosing a_i over some alternative action z_i is

$$\begin{aligned}
&(1-\varepsilon)\left[\pi(a_i, \phi_i) - \pi(z_i, \phi_i)\right] + \varepsilon B \frac{\max\{0, \pi_i(z_i, \phi_i) - \pi_i(a_i, \phi_i)\}}{f_i(a_i, \phi_i)} \\
&\geq (1-\varepsilon)\left[\pi(a_i, \phi_i) - \pi(z_i, \phi_i) + \max\{0, \pi_i(z_i, \phi_i) - \pi_i(a_i, \phi_i)\}\right] \\
&\geq 0
\end{aligned}$$

Thus a_i is an interim best-reply to φ_i.

To show necessity, suppose for some i that $f_i(a_i, \phi_i) > \frac{\varepsilon B}{1-\varepsilon}$ for all $\phi_i \in \Delta K_{-i}$. Consider any elaboration Γ, type profile t in $C_\Gamma^{1-\varepsilon}$, and profile of conjectures φ such that $t \in B_\Gamma^p(T^0 \cap \varphi^{-1}(K))$. Let $\phi_i \in \Delta K_{-i}$ be the conditional probability over K_{-i} derived from φ_i conditional on type t_i and the event $T^0 \cap \varphi^{-1}(K)$. For any $z_i \in \mathrm{argmax} \pi_i(\cdot, \phi_i)$, the expected utility difference for t_i from choosing z_i over a_i is at least $(1-\varepsilon)(\pi_i(z_i, \phi_i) - \pi_i(a_i, \phi_i)) - \varepsilon B = (1-\varepsilon) f_i(a_i, \phi_i) - \varepsilon B > 0$. Therefore, a_i is not an interim best-reply for type t_i and thus $a \notin \Lambda_\varepsilon(K)$. ■

It follows from Lemma 3 that $\Lambda_\varepsilon(K)$ is compact when K is, and hence that Λ_ε^n is compact for every n.

Lemma 4 *Suppose $a \notin \Lambda_0^n$. Then there exists $\varepsilon > 0$ such that $a \notin \Lambda_\varepsilon^n$.*

Proof: The lemma is proven by induction. Note that Λ_ε^1 is just $\Lambda_\varepsilon(A)$. Since A_{-i} is compact, it follows from lemma 3 that if $a \notin \Lambda_0^1$ then there exists $\delta > 0$ such that $f_i(a_i, \phi_i) > \delta$ for some i and $\phi_i \in \Delta A_{-i}$. Therefore, for $\varepsilon = \frac{\delta}{B+\delta}$, we have $a \notin \Lambda_\varepsilon^1$.

For the inductive step, suppose $a_i \notin [\Lambda_0^n]_i$ and consider the compact subset $R_i := R_i(a_i, A)$ of rationalizations of a_i in G. If $\phi_i \in R_i$, then $C(\phi_i) \not\subset [\Lambda_0^n]_{-i}$, else $\phi_i \in \Delta_i \Lambda_0^{n-1}$ which would imply $a \in [\Lambda_0^n]_i$. Let $z_{-i} \in C(\phi_i) \cap \neg[\Lambda_0^n]_{-i}$. By the inductive hypothesis, there exists $\varepsilon > 0$ such that $z_{-i} \notin [\Lambda_\varepsilon^n]_{-i}$, and by Lemma 3 there is a neighborhood U of z_{-i} such that $U \subset \neg[\Lambda_\varepsilon^n]_{-i}$. But this implies that $z_{-i} \notin C(\tilde{\varphi}_i)$ for any $\tilde{\varphi}_i \in \Delta_i \Lambda_\varepsilon^{n-1}$, and since $z_{-i} \in C(\phi_i)$, we conclude $\phi_i \notin \Delta_i \Lambda_\varepsilon^{n-1}$.

Consider the following family of open sets:

$$\mathcal{U} = \{\neg\Delta_i \Lambda_\varepsilon^{n-1} : \varepsilon > 0\}.$$

The argument of the previous paragraph amounts to the statement that $\mathcal{U}$ covers R_i. And by lemma 3 this is an open covering. Hence, there exists a finite set $\{\varepsilon_1, \ldots, \varepsilon_k\}$ such that $\{\neg\Delta_i \Lambda_{\varepsilon_j}^{n-1}\}_{j=1}^k$ covers R_i. Let $\tilde{\varepsilon} = \min_j \varepsilon_j$. Note that $\neg\Delta_i \Lambda_{\varepsilon_j}^{n-1} \subset \neg\Delta_i \Lambda_{\tilde{\varepsilon}}^{n-1}$ and therefore

$$R_i \subset \neg\Delta_i \Lambda_{\tilde{\varepsilon}}^{n-1}.$$

Hence if $\phi \in \Delta_i \Lambda_{\tilde{\varepsilon}}^{n-1}$, then $a_i \notin \operatorname{argmax} \pi_i(\cdot, \phi)$, i.e. $f_i(a_i, \phi) > 0$. Now by the compactness of $\Delta_i \Lambda_{\tilde{\varepsilon}}^{n-1}$, there exists $\delta > 0$ such that $f_i(a_i, \phi) > \delta$ for all $\phi \in \Delta_i \Lambda_{\tilde{\varepsilon}}^{n-1}$. If we set $\varepsilon = \frac{\delta}{B+\delta} > 0$, we have $f_i(a_i, \phi) > \frac{\varepsilon B}{1-\varepsilon}$ and hence $a \notin \Lambda_\varepsilon^n$. ■

Proof of Theorem 1 By Lemma 2 and the first part of Proposition 2, $\Lambda^\infty \subset \overline{\Lambda^\infty}$. To show the opposite inclusion, suppose $\sigma \notin \Lambda^\infty = \Lambda_0^\infty$. Then $\sigma \notin \Lambda_0^n$ for some n. By Lemma 4, $\sigma \notin \Lambda_\varepsilon^n$ for some $\varepsilon > 0$. Since $\Lambda_{\varepsilon'}^n \subset \Lambda_\varepsilon^n$ for all $\varepsilon' < \varepsilon$ by Lemma 2, we have $\sigma \notin \Lambda_{\varepsilon'}^n$ and hence $\sigma \notin \Lambda_{\varepsilon'}^\infty$ for all $\varepsilon' \in (0, \varepsilon)$. This implies that $\sigma \notin \limsup \Lambda_\varepsilon^\infty$ and hence $\sigma \notin \overline{\Lambda^\infty}$ by Proposition 3. ■

3 A Discontinuous Counterexample

In this section I show that Theorem 1 can fail for discontinuous games. There is an example in Lipman (1994) which has a similar motivation but makes a different point. To illustrate the difference, it will help to introduce some additional terminology. Let us say for the purposes of this section that the set $R(G)$ of rationalizable strategies for a game G is the set of strategies that could be played at some state of the world in which the payoffs of G as well as the rationality of the players are common-knowledge. In fact, this is how Bernheim (1984) formally defined rationalizability and was also the motivation for the definition in Pearce (1984). It was proven in Bernheim (1984) that $R(G) = \Lambda^\infty(G)$ for games with continuous payoffs and compact action spaces. Since $\Lambda^\infty(G)$ is equal to the "limit" as $N \to \infty$ of $\cap_{n \le N} \Lambda^n(G)$ and $\cap_{n \le N} \Lambda^n(G)$ is equal to the set of strategies which could be played in a state of the world in which the payoffs and rationality of the players are mutually known to N-levels, the Bernheim (1984) result could be thought of as stating that in continuous games, rationalizability (i.e. $R(G)$) is well approximated when there is approximate common-knowledge in the sense of many, but finitely many, levels of mutual knowledge. [4]

Lipman (1994) shows that compactness and continuity are necessary for this result. In his example, $R(G)$ is a proper subset of $\Lambda^\infty(G)$. However, there always exists a sufficiently large ordinal number α such that $\Lambda^\alpha(G) = R(G)$ regardless of the continuity properties of the payoff function. That is, rationalizability is well-approximated by approximate common-knowledge in the sense of sufficiently, and possibly trans-finitely many, levels of mutual knowledge of rationality.

The example below demonstrates something different. In the game G I present below, $R(G) = \Lambda^\infty(G)$. Also, for the elaborations $\Gamma(\varepsilon)$ used in the example, $R(\Gamma(\varepsilon)) = \Lambda^\infty(\Gamma(\varepsilon))$. Thus, rationalizability for complete information games is not well-approximated by *rationalizability itself* for games with near-complete information.

Consider the following two-player symmetric game. The action sets are $A_i = \mathbb{Z} \cup \{-\infty, \infty\}$ for $i = 1, 2$ where $\mathbb{Z}$ is the set of integers. These are compact sets when we take ∞ to be the one-point compactification of $\mathbb{Z}_+$ and $-\infty$ the one-point compactification of $\mathbb{Z}_-$. The payoff function is symmetric: $\pi_1(a_1, a_2) = \pi_2(a_2, a_1) = \pi(a_1, a_2)$ and π is defined as follows. First $\pi(-\infty, \cdot) = 0$, $\pi(0, \cdot) = 1$, and $\pi(\infty, y)$ is 1 for $y < 0$ and 0 for $y \ge 0$. For all other cases,

$$\pi(x, y) = \begin{cases} 1 - \frac{1}{x} \text{ if } y > 0 \text{ and } x = y \\ 1 \text{ if } y > 0 \text{ and } x = -y \\ 1 - \frac{1}{x} \text{ if } y < 0,\, x > 0 \\ 0 \text{ otherwise} \end{cases}$$

[4] Note that Theorem 1 above establishes something different: rationalizability is well-approximated by approximate common-knowledge in the sense of Monderer and Samet (1989).

With this payoff function, each player's payoff is continuous in his own action and in the action of his opponent as the reader can easily verify. The only discontinuities are on the "diagonal:" for all $0 < x < \infty$, $\pi(-x, x) = 1$, but $\pi(-\infty, \infty) = 0$, and $\lim_{x\to\infty} \pi(x, x) = 1 \neq 0 = \pi(\infty, \infty)$. This diagonal discontinuity is enough to invalidate Theorem 1.

There is a unique rationalizable action profile, $\{0, 0\}$, and this action profile can be obtained by iterative elimination of actions that are never best-replies, i.e. $\Lambda^\infty = \{0, 0\}$. In the first round, $-\infty$, together with all $0 < x < \infty$ are eliminated. Each of these are strongly dominated by 0. In the second round, each $x < 0$ is dominated by 0 and hence eliminated, and ∞ is eliminated in the third round.

Consider the following class of elaborations $\Gamma(\varepsilon)$ parameterized by $\varepsilon \in (0, 1/2)$. Each i has two types, $T_i = \{t^0, \tilde{t}\}$ for $i = 1, 2$. The prior μ is given by $\mu((t^0, t^0)) = \mu((\tilde{t}, \tilde{t})) = \frac{1-\varepsilon}{2}$ and $\mu((\tilde{t}, t^0)) = \mu((t^0, \tilde{t})) = \varepsilon/2$. Thus, $T^0 = \{(t^0, t^0)\}$ is $(1 - \varepsilon)$-evident and hence common $(1 - \varepsilon)$-belief when it occurs. The payoffs are as π in state (t^0, t^0), but in all other type profiles t, $u_j((x, x), t) = 2$ for $j = 1, 2$, for all actions x. with all other payoffs unchanged from π.

Let σ^x denote the strategy which plays x independent of type. In $\Gamma(\varepsilon)$, a positive integer $x > \frac{1-\varepsilon}{\varepsilon}$ is an interim best-reply for both types of player i to the strategy σ^x. To see this, note that under π, the maximum payoff against x is 1. Since the only action whose payoff against x has been affected in the elaboration is x itself, it is enough to show that the interim expected payoff from playing x against σ^x exceeds 1. This is obvious for type $\tilde{t}$. We calculate the interim expected payoff for type t^0 to action x against σ^x as

$$\begin{aligned}
\mathbf{E}u(x, \sigma^x | t^0) &= (1 - \varepsilon)u((x, x), (t^0, t^0)) + \varepsilon u((x, x), (t^0, \tilde{t})) \\
&= (1 - \varepsilon)\pi(x, x) + \varepsilon 2 \\
&= (1 - \varepsilon)(1 - \frac{1}{x}) + 2\varepsilon \\
&> (1 - \varepsilon)(1 - \frac{\varepsilon}{1 - \varepsilon}) + 2\varepsilon \\
&= 1
\end{aligned}$$

Since for each $x > \frac{1-\varepsilon}{\varepsilon}$, σ^x is an interim best-reply to itself, it belongs to $\Lambda^\infty_{\Gamma(\varepsilon)}$. Next we will show that $\sigma^{-x} \in \Lambda^\infty_{\Gamma(\varepsilon)}$ for for all such x. Consider the the strategy $\sigma(t^0) = x$, $\sigma(\tilde{t}) = -x$. If we can show that σ^{-x} is an interim best-reply to σ, then it will follow that both σ and σ^{-x} are rationalizable since $\sigma(t^0) = \sigma^x(t^0)$ is an interim best-reply to σ^x and and also $\sigma(\tilde{t}) = \sigma^{-x}(\tilde{t})$ is an interim best-reply to σ.

Since $-x$ is a best-reply to x in state (t^0, t^0), and $-x$ is a best-reply to $-x$ in state $(t^0, \tilde{t})$, $-x$ is an interim best-reply to σ for type t^0. The interim expected payoff for type $\tilde{t}$ against σ is

$$\begin{aligned}\mathbf{E}u(-x, \sigma|\tilde{t}) &= (1-\varepsilon)u((-x,-x),(\tilde{t},\tilde{t})) + \varepsilon u((-x,x),(\tilde{t},t^0)) \\ &= 2(1-\varepsilon) + \varepsilon\pi(-x,x) \\ &= 2(1-\varepsilon) + \varepsilon \\ &= 2-\varepsilon\end{aligned}$$

which is greater than 1, the interim expected payoff to action 0 and is greater than the interim expected payoff to x

$$\begin{aligned}\mathbf{E}u(x, \sigma|\tilde{t}) &= (1-\varepsilon)u((x,-x),(\tilde{t},\tilde{t})) + \varepsilon u((x,x),(\tilde{t},t^0)) \\ &= (1-\varepsilon)\pi(x,-x) + 2\varepsilon \\ &< (1-\varepsilon) + 2\varepsilon \\ &= 1+\varepsilon\end{aligned}$$

since $\varepsilon < 1/2$. All other actions are strictly worse as can easily be verified. Thus, for $x > \frac{1-\varepsilon}{\varepsilon}$, $-x \in \Lambda^{\infty}_{\Gamma(\varepsilon)}$.

Finally, we show that $\sigma^{\infty} \in \Lambda^{\infty}_{\Gamma(\varepsilon)}$ for all $\varepsilon \in (0, 1/2)$. For any $-\infty < x < 0$, consider the strategy $\tilde{\sigma}(t^0) = x$ and $\tilde{\sigma}(\tilde{t}) = \infty$. Since ∞ is a best-reply to x in state (t^0, t^0) and ∞ is a best-reply to ∞ in state $(t^0, \tilde{t})$, we conclude that ∞ is an interim best-reply to $\tilde{\sigma}$ for type t^0. To check type $\tilde{t}$, we calculate

$$\begin{aligned}\mathbf{E}u(\infty, \tilde{\sigma}|\tilde{t}) &= (1-\varepsilon)u((\infty,\infty),(\tilde{t},\tilde{t})) + \varepsilon u((\infty,x),(\tilde{t},t^0)) \\ &= 2(1-\varepsilon) + \varepsilon\end{aligned}$$

which is greater than both $\mathbf{E}u(0, \tilde{\sigma}|\tilde{t}) = 1$ and $\mathbf{E}u(x, \tilde{\sigma}|\tilde{t}) = 2\varepsilon$. Since all other actions are strictly worse, ∞ is an interim best-reply for type $\tilde{t}$. It follows that for every $\varepsilon \in (0, 1/2)$, both $\tilde{\sigma}$ and σ^{∞} belong to $\Lambda^{\infty}_{\Gamma(\varepsilon)}$ given that $x < \frac{1-\varepsilon}{\varepsilon}$.

For every $\varepsilon \in (0, 1/2)$, we have constructed an elaboration $\Gamma(\varepsilon)$ and a $\sigma \in \Lambda^{\infty}_{\Gamma(\varepsilon)}$ such that $\sigma(t) = \infty$ and $t \in T^0$. We conclude that $\infty \in \overline{\Lambda^{\infty}}$, and hence that $\overline{\Lambda^{\infty}} \neq \Lambda^{\infty}$.

References

BERNHEIM, D. (1984): "Rationalizable Strategic Behavior," *Econometrica*, 52, 1007–1028.

CARLSSON, H., AND E. VAN DAMME (1993): "Global Games and Equilibrium Selection," *Econometrica*, 61, 989–1018.

CHUNG, K.-S., AND J. C. ELY (????): "Implementation with Near Complete Information," Northwestern University CMS-EMS Discussion Paper #1332.

DEKEL, E., AND F. GÜL (1997): "Knowledge and Rationality in Game Theory," in *Advances in Economic Theory: Proceedings of the Seventh World Congress of the Econometric Society*, ed. by D. M. Kreps, Cambridge. Cambridge University Press.

FUDENBERG, D., D. M. KREPS, AND D. K. LEVINE (1988): "On the Robustness of Equilibrium Refinements," *Journal of Economic Theory*, 44, 354–380.

KAJII, A., AND S. MORRIS (1997): "The Robustness of Equilibria to Incomplete Information," *Econometrica*, 65, 1283–1309.

——— (1998): "Payoff Continuity in Incomplete Information Games," *Journal of Economic Theory*, 82, 267–276.

LIPMAN, B. L. (1994): "A Note on the Implications of Common Knowledge of Rationality," *Games and Economic Behavior*, 6, 114–129.

MONDERER, D., AND D. SAMET (1989): "Approximating Common Knowledge with Common Beliefs," *Games and Economic Behavior*, 1, 170–190.

PEARCE, D. (1984): "Rationalizable Strategic Behavior and the Problem of Perfection," *Econometrica*, 52, 1029–1050.

RUBINSTEIN, A. (1989): "The Electronic Mail Game: Strategic Behavior under Almost Common Knowledge," *American Economic Review*, 79, 385–391.

TAN, T. C.-C., AND S. R. DA COSTA WERLANG (1988): "The Bayesian Foundations of Solution Concepts of Games," *Journal of Economic Theory*, 45(2).