

TECHNICAL REPORT SERIES

RULES FOR CHANGING THE RULES

by

Spyros Vassilakis

Technical Report No. 32

THE ECONOMICS SERIES

STANFORD INSTITUTE FOR THEORETICAL ECONOMICS

FOURTH FLOOR, ENCINA HALL

STANFORD UNIVERSITY

STANFORD, CALIFORNIA



## RULES FOR CHANGING THE RULES

SPYROS VASSILAKIS

November 14, 1990

I would like to thank, without implicating them, the Referee, an Associate Editor, Richard Boylan, Don Brown, Leo Hurwicz, Ernest Manes, Douglas Gale, Paul Milgrom, Al Roth, Ariel Rubinstein, and Shmuel Zamir; seminar participants at Stanford, Minnesota and Stony Brook; Kevin Sontheimer for his help in obtaining financial assistance from the Scaife Foundation, which is gratefully acknowledged. Many thanks goes to Lauree Graham who typed a difficult manuscript with marvelous efficiency. Special thanks to Fadi Alameddine who read the paper carefully and made numerous helpful suggestions. Finally, Ed Green's help and encouragement cannot be overemphasized.

---

## Table of Contents

1. Introduction
  2. Outline of the main constructions
    - 2.1 Solution concepts
    - 2.2 Universal Games
    - 2.3 Coordination - Proof Solution Concepts.
  3. The main categories
    - 3.1 The base category  $D$
    - 3.2 Fibrations and Sections
    - 3.3 The category  $G$  of games
    - 3.4 The category  $M$  of mechanisms
    - 3.5 Preferences over lotteries
    - 3.6 The category  $E$  of solution concepts
  4. Universal games
  5. Coordination - proof solution concepts
  6. Conclusions
- Appendix 1
- Appendix 2

...while we may not be too restrictive in the behavioral assumptions we make, we make enormously strong ad hoc assumptions in the range of actions we allow the actors in our models.

D. Kreps ((1990b), p. 166)

## I. INTRODUCTION

Many interesting problems in economics involve recursive constructions that cannot be studied with the familiar tools of analysis and fixed point theory; in fact, many of these constructions have been considered to be impossible to handle, because they seem to lead to an infinite regress. Crawford (1985), for example, states (p. 823): "The second observation concerns my assumption that the rules of bargaining over mechanisms cannot be altered. It is easy to imagine more general specifications, in which agents can amend the rules, amend the rules for amending the rules, and so on. But it is clear on reflection that unless this sequence comes to a halt at some level, there is no place to anchor a game-theoretic analysis ... given that exogenous specification cannot be avoided, it seems a sensible research strategy to assume that agents are free to bargain over mechanisms, but they cannot amend the rules of bargaining."

In a companion paper, Vassilakis (1990), I claimed that category theory, and in particular the theory of functorial fixed points, is the appropriate tool for handling problems of infinite regress of the sort discussed by Crawford; the application that I presented to illustrate my claim was an easy (re)construction of the Mertens-Zamir (1985) universal beliefs space. In this paper I will show that in the appropriate category, every game generates a universal game, that is a game whose strategy spaces include the strategy spaces of the original game and, in addition, all possible proposals to modify the rules, including the rules of the universal game. Secondly, I will show how each solution concept with certain properties can be "completed" so as to be coordination-proof, i.e., such that the prediction of the solution concept on the universal game, when projected on the original game, is identical with the prediction of the solution concept on the original game. Both universal games and coordination-proof solution concepts will be defined as fixed points of functors, in the same way the universal beliefs space was defined as a fixed point of (a translation of) the probability functor.

Finally, I will show that, for each solution concept, the corresponding coordination-proof solution concept defines equilibrium institutions (mechanisms, rules).

Why are such constructions relevant? To answer this question, recall that noncooperative game theory generates predictions for each particular situation in a three-step process: first, a game form is specified; secondly, payoffs are assigned to strategy vectors; and finally, a particular vector of strategies is singled out by some criterion (the solution concept), and is called the equilibrium of the game. It is implicitly assumed that (a) the game form is an adequate description of the opportunities available to the players, and (b) that this three-step method can be applied to any economic problem to generate predictions.

This method is subject to some limitations, in particular in the applied areas of bargaining, oligopoly and mechanism design. The limitations can be summarized in one question: "Where do the rules come from?", where by rules, we mean, in this paper, game forms. From the point of view of noncooperative game theory, this question is nonsense; the theory takes the rules as exogenous and observable. In each application at hand, therefore, the theory recommends that we should identify the exogenous and observable quantities and codify them in the game form; the possible values of exogenous but unobservable quantities are to form the space of types; and an application of the solution concept to the game form constructed should generate a prediction for each possible valuation of the outcomes and each possible value of the unobservables. In what follows I argue that in some applied areas the rules that can reasonably be considered both exogenous and observable are sometimes either far less detailed than the rules actually employed to generate predictions, or ignored by players, or both. These arguments motivate the construction of universal games.

To apply bargaining theory to predict the outcome of some particular bargaining situation we have to specify the order and possible content of offers and counteroffers, whether they are binding or not after they have been accepted, the time it takes to respond to an offer, and many other details described in Kreps (1990, Chapter 15). To quote Kreps again, "in most negotiations these things, which constitute the protocol of negotiations, are not really specified with the precision we require for an extensive form game" (ibid, p. 376).

The same problem applies in oligopoly theory: to apply it in some particular market we have to specify the strategic variables (prices, quantities, supply functions, ...) the timing of decisions, the communication

allowed among rival firms, the possibilities for mergers or takeovers, the precise form of uncertainty about various cost or demand parameters, and so on. Again, an applied problem will rarely come equipped with so much detail. What is more, most of these "details" of the specification of the game form cannot be reasonably considered as exogenous: they are under the control of players, and they make a difference in the outcome.

A different but related problem limits the area of possible applications of the theory of mechanism design. This problem is best described in the context of the market for medical interns studied in a series of papers by Roth (1984, 1990) and Roth and Sotomayor (1990). The players are students and hospitals. A typical mechanism asks students to submit a list of hospitals in decreasing order of preference, and hospitals to submit a similarly ordered list of students. The outcome function of such a mechanism is a matching algorithm that assigns students to hospitals taking into account their submitted lists. Mechanisms differ in the matching algorithms they employ. In some cases, when algorithms were "unstable" in the sense of Roth, players either ignored the mechanism or colluded in order to manipulate it (see Roth, (1980) pp. 9-11). Players ignore the mechanism when they don't submit lists but arrange for residencies by direct communication. Players collude to manipulate a mechanism when, for example, a hospital promises to rank a student first if the student returns the favor. Notice that the mechanism might be incentive compatible (in the formal sense) and still be ignored or manipulated by collusion: no single student or hospital might have an incentive to lie or cheat, but in the presence of an alternative mechanism coordinating actions, everybody does.

Another example is in Holstrom and Tirole (1989, Chapters 5.1 and 5.5.), who question the applicability of mechanism design theory to the theory of the firm. The view they oppose is that of the owner of the firm (or some principal in general) designing an incentive compatible mechanism for the firm's employees. While (the formal definition of) incentive compatibility allows for the fact that employees can unilaterally lie and cheat, it does not allow for the possibility of coordinated lying and cheating (multilateral contracting in Holstrom and Tirole's terminology). Given that our theory is noncooperative, however, successful coordinations are those that are equilibria of the game induced by some incentive compatible mechanism. Hence, ignoring multilateral contracting amounts to ignoring the possibility that employees can propose mechanisms. Is this a serious omission? It seems to be important in the theory of the firm, as it is

argued in Holstrom and Tirole (1989) and their references. But then, one argument goes, if this omission is truly important, why can't we simply enlarge the strategy spaces of employees to allow them to propose mechanisms? To be sure, the set of incentive-compatible mechanisms the principal can choose from might shrink, but the theory would still be the same. This argument misses the following points. First, if everybody can propose mechanisms, the identity of the principal(s) is endogenous. Secondly, the theory would not be the same because the incentive compatibility of the principal's mechanism is defined as the property that no incentive compatible attempt to induce coordinated lying and cheating is successful, and this is a self-referential statement. Thirdly, the new strategy spaces will have to capture all the opportunities for coordination available to the agents, including coordination on these new strategy spaces themselves; this is again a self-referential statement.

A possible argument against this approach is that, in all known cases, rules for changing the rules do not go beyond a small number of steps up in the hierarchy. For example, some argue, while it is true that decisions of the lower courts can be appealed at higher-level courts, the number of steps one can go up in the hierarchy of courts is very small. Hence, for practical purposes, we can ignore infinite regress and universal games: in each particular case, we can just observe the actual rules; find out the level of the hierarchy in which they belong; and describe the situation at hand by a game that truncates the hierarchy at this particular level. The fact that the position of observed rules in the hierarchy of rules is an endogenous variable (to be determined in equilibrium) undercuts the force of this argument. Comparative statics based on this argument are unreliable because they miss the fact that different values of the parameters might change the position of equilibrium rules in the hierarchy. Explanations of, say, the long-run stability of the position of some actual rule in the hierarchy of rules become impossible, since this stability is simply assumed.

What is the theoretical problem here? It is not that the rules we take as exogenous and employ to generate predictions are too detailed to be really exogenous, as it was the case in the previous examples; the optimal incentive-compatible mechanism is clearly endogenous. It is rather the fact that the theory predicts that the observed outcome will be the equilibrium of the optimal incentive compatible mechanism, or some

incentive compatible mechanism, and ignores the fact that subsets of players can coordinate their actions by employing different mechanisms.

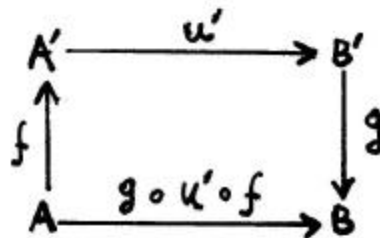
There seems to be a common thread in the previous arguments: there are applications in which the game form employed to generate predictions does not capture all the opportunities open to players either because it takes as exogenous some "details" that players can, and want to, vary, or because it ignores the ability of each player to propose, as well as participate in, mechanisms. In all cases, the actions omitted from the action sets of players include (proposals of) rules. Examples of such proposals: "let's bargain à la Rubinstein" or "let's vote about how to bargain; I vote for Rubinstein;" "let's compete à la Cournot," or "let the biggest firm decide on how we are going to compete;" "let's ostracize rate busters," or "let the shop stewards decide for us on how to treat rate busters," etc., ...

The natural way to proceed from here is to recognize that a game form can be the mathematical representation of two different kinds of rules. Physical rules have the same methodological status as production sets in competitive analysis: they are exogenous and describe the technological constraints faced by players. Artificial rules are game forms like Cournot's or Rubinstein's. They can be utilized in two ways, either as proxies for the physical rules, because the physical rules are too complex; or as actions (proposals) included in the agents' action sets. In this paper it is artificial rules in the sense of possible proposals that are of interest. The methodological status of such rules is dual: on the one hand, they define a game form on which the solution concept generates predictions; on the other hand, being endogenous, they are themselves equilibrium outcomes of some other game. The game form of this new game is certainly artificial, because its strategy spaces include proposed rules and its outcome function maps rules into rules. Being artificial, it is itself the equilibrium outcome of some other game ...;infinite regress raises its ugly head.

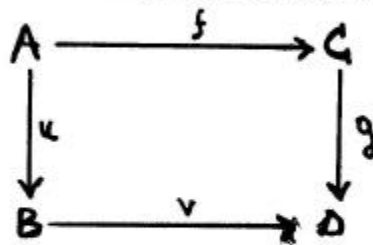
The rest of the paper is organized as follows: Section 2 describes the main concepts, namely solution concepts, universal games and coordination-proof solution concepts, without technical details; section 3 describes the main categories and functors of interest, and examines their properties: existence and characterization of various colimits for categories, continuity properties for functors; section 4 describes in detail the construction of universal games; section 5 describes the construction of coordination-proof solution

concepts; section 6 discusses the limitations of the constructions and comments on how they can be overcome by further research. The appendix contains proofs of peripheral interest. The reader of this paper is assumed to be familiar with the facts from category theory contained in Vassilakis (1990).

A few words about notation: the composite of two morphisms  $f: A \rightarrow B$  and  $g: B \rightarrow C$  is denoted by  $g \circ f$  or simply  $gf$ : the value of this composite at a point  $a$  in  $A$  is denoted by  $g(f(a))$ ,  $(g \circ f)(a)$ , or  $gfa$ , depending on context and intended emphasis. If  $A, B$  are objects in a category, then  $[A \rightarrow B]$  denotes the set of morphisms from  $A$  to  $B$ ; whenever  $A, B$  are ordered sets,  $[A \rightarrow B]$  will also be an ordered set, with functions from  $A$  to  $B$  ordered pointwise. If  $f: A \rightarrow A'$  is a morphism in some category and  $B$  is an object in the same category, then  $\text{---} \circ f: [A' \rightarrow B] \rightarrow [A \rightarrow B]$  denotes the morphism that assigns to each  $u': A' \rightarrow B$  the composite  $u' \circ f: A \rightarrow B$ . If  $g: B' \rightarrow B$  is a morphism in the same category, then  $g \circ \text{---} \circ f: [A' \rightarrow B'] \rightarrow [A \rightarrow B]$  assigns to each  $u': A' \rightarrow B'$  the composite morphism  $g \circ u' \circ f: A \rightarrow B$ ; the following diagram clarifies this:



Finally, recall that a diagram commutes if all the paths with the same beginning and the same end are equal. (i.e., they denote the same morphism). For example, to say that the following diagram commutes



means that  $g \circ f = v \circ u$

## 2. OUTLINE OF THE MAIN CONSTRUCTIONS

To keep the discussion as simple as possible, categories in this section will be defined only by their objects, neglecting morphisms; functors will be defined only on objects; and terms that are not necessary for an intuitive understanding of the constructions but have to be mentioned for future reference will not be defined. Only normal form, complete information games are considered.

### 2.1 Solution concepts and mechanisms

Recall that a solution concept is a rule that assigns to each game a set of probability measures on its aggregate strategy space: for example, the Nash solution concept assigns to each game the set of "mixed strategies" that are fixed points of its best reply correspondence. Given that the method of selecting a solution concept adopted in this paper is to require the solution concept to be a fixed point of some functor that maps solution concepts into solution concepts, we need a precise mathematical expression for solution concepts and the category in which they live: in other words, solution concepts are both "rules" and "points" in a "space" (in fact, objects in a category). To achieve this, we need the following:

A (normal form complete information) game  $g = (A, u)$  is an  $n$ -tuple  $A = \langle A_1, \dots, A_n \rangle$  of strategy spaces and a vector  $u: \Pi(A) \rightarrow \mathbb{R}^n$  of payoff functions. Each strategy space  $A_i$  is an object of some category  $D$ , to be specified in section 3; hence, the  $n$ -tuple  $A$  belongs to  $D^n$ , the  $n$ -fold product of  $D$ . The category  $G$  of games has objects of the form  $g = (A, u)$ . The functor  $P_1: G \rightarrow D^n$  assigns to each game  $g = (A, u)$  the  $n$ -tuple  $A$ ; the functor  $\Pi: D^n \rightarrow D$  assigns to each  $n$ -tuple  $A$  its product  $\Pi(A)$ , while the functor  $\Delta: D \rightarrow D$  assigns to each space  $X$  the space  $\Delta(X)$  of probability measures on  $X$ . The composite functor  $\Delta \Pi P_1: G \rightarrow D$  assigns to each game  $g = (A, u)$  the space  $\Delta(\Pi(A))$  of probability measures on its aggregate strategy space  $\Pi(A)$ .

Let  $g$  be a game and  $p$  a probability measure in  $\Delta(\Pi(P_1(g)))$ ; then  $p$  is a candidate solution of the game  $g$ . This motivates the following definition:

1. *Definition:* The category of fibrations of the functor  $\Delta \circ \Pi \circ P_1$  has objects the pairs  $(g, p)$  where  $g$  is in  $G$  and  $p \in \Delta \Pi P_1(g)$ . Call this category  $\text{Fib}(\Delta \Pi P_1)$ .

A solution concept, then, is a rule that assigns to each game  $g$  a probability measure  $p$  in  $\Delta\Pi P_1(g)$ . This motivates the following

2. *Definition:* The category  $\text{Sect}(\Delta\Pi P_1)$  of sections of the functor  $\Delta\Pi P_1$  has objects all the functors  $\sigma: G \rightarrow \text{Fib}(\Delta\Pi P_1)$  such that  $\sigma(g) = (g, p)$  for some  $p \in \Delta\Pi P_1(g)$ . We write  $P_1\sigma(g) = g$ ,  $P_2\sigma(g) = p$ ; the projection symbol  $P_1$  utilized here should not be confused with the projection functor  $P_1: G \rightarrow D^n$  that appears in  $\Delta\Pi P_1$ ; the meaning of  $P_1$  will be clear from its context.

The category of solution concepts will be a subcategory  $\Sigma$  of  $\text{Sect}(\Delta\Pi P_1)$ : its objects will be functors that satisfy a "weak revelation principle" and are  $\omega$ -cocontinuous, i.e., preserve  $\omega$ -colimits; precise definitions are in the next section.

Note that this definition achieves its objective: solution concepts are rules that assign to each game  $g$  a probabilistic "prediction"  $p$ ; at the same time they are objects in some category, so that functors can be defined on, and map into, them. Note also that this definition is very restrictive, because it allows for single-valued solution concepts only: each game is assigned exactly one probability measure. What is gained by this restriction is the ability to address at a relatively simple conceptual level, and in full generality as far as all other aspects of the problem are concerned, the methodological and mathematical issues raised by the open problems mentioned in the introduction. What is lost is the ability to build on already existing solution concepts, all of which are multi-valued, and, therefore, the ability to construct interesting concrete examples of coordination-proof solution concepts and equilibrium institutions, as opposed to the ability to prove general existence theorems (on the other hand, trivial explicit examples are known). The Harsanyi-Selten theory might suggest itself at this point, but recall that it applies only to finite games: even if the original game considered is finite, the universal game generated by it is not. All in all, this seems to be an acceptable tradeoff, given that these constructions were considered impossible up to now, and given that a (nontrivial) extension of the present methods, as outlined in the last section, seems to be a promising way of handling multivalued solution concepts.

Economic agents can seek to coordinate their actions on a game  $g = (A, u)$  by proposing mechanisms on  $g$ , i.e., triples  $(S, R, f)$  where  $S, R$  are in  $D^n$  and  $f: \Pi(S) \rightarrow \Delta(\Pi(R))$  is in  $D$ ;  $S$  is an  $n$ -tuple of report spaces

and  $R$  an  $n$ -tuple of recommendation spaces; if agents report a vector  $s \in \Pi(S)$ , the mechanism will recommend a vector in a subset  $E$  of  $\Pi(R)$  with probability  $f(s)(E)$ . This ought to motivate the following

3. *Definition:* The functor  $M: D^n \times D^n \rightarrow D$  is given by  $M(S, R) = [\Pi(S) \rightarrow \Delta(\Pi(R)) =$  morphisms in  $D$  from  $\Pi(S)$  to  $\Delta(\Pi(R))$ .

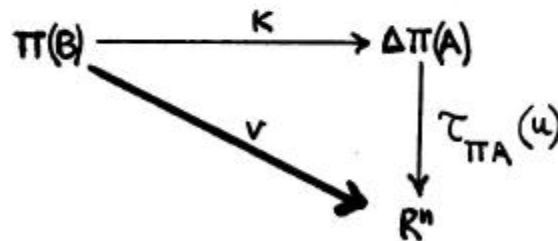
4. *Definition:* The category  $\text{Fib}(M)$  of fibrations of the functor  $M$  has objects triples,  $(S, R, f)$ ,  $f \in M(S, R)$ ; hence, fibrations of  $M$  are just mechanisms.

When a mechanism  $m = (S, R, f)$  is utilized to coordinate actions on a game  $g = (A, u)$  a new game  $T(m, g) = (B, v)$  is played. The strategy spaces of the new game,

$$B_i = S_i \times [R_i \rightarrow A_i] \quad i=1, \dots, n,$$

consist of the space  $S_i$  of possible reports and the space  $[R_i \rightarrow A_i]$  of possible deviation strategies. If agents play  $(s, \delta) \in \Pi(B)$ , they induce a probability measure  $k(s, \delta) = f(s) \circ \delta^{-1}$  on  $\Pi(A)$ , where, for each Borel subset  $E$  of  $\Pi(A)$ ,  $\delta^{-1}(E) = \{x \in \Pi(R) : \delta(x) \in E\}$ . Hence, to define the payoff function  $v$  of the new game,  $u: \Pi(A) \rightarrow \mathbb{R}^n$  has to be extended from  $\Pi(A)$  to  $\Delta(\Pi(A))$ ; let  $\tau_{\Pi(A)}(u): \Delta(\Pi(A)) \rightarrow \mathbb{R}^n$  be this extension. If preferences over lotteries were to satisfy the expected utility axioms, then  $\tau_{\Pi(A)}(u)(p) = \int_{\Pi(A)} u dp$ ,  $p \in \Delta \Pi A$ .

The following diagram, then, defines  $v$ :



$$v(s, \delta) = \tau_{\Pi(A)}(u)(k(s, \delta)) = \tau_{\Pi(A)}(u)(f(s) \circ \delta^{-1}).$$

## 2.2 Universal Games

Given that solution concepts in  $\Sigma$  satisfy the weak revelation principle, and that only games of complete information are considered, direct mechanisms with singleton message spaces can attain each utility vector attained by general mechanisms (a proof of this fact is in section 3.6). Each  $A$  in  $D^n$ , then, generates new strategy spaces  $R(A)$ , defined by

$$R_i(A) = \Delta(\Pi(A)) \times [A_i^n - A_i], \quad i=1, \dots, n$$

In words, given a game with strategy spaces  $A$ , agent  $i$  can propose any mechanism  $p_i \in (\Delta(\Pi(A)))$  and adopt any deviation strategy  $\delta_i: A_i^n \rightarrow A_i$ ; the domain of  $\delta$  is  $A_i^n$  because  $i$  will be receiving recommendations in  $A_i$  from the mechanism proposed by each agent  $j=1, \dots, n$ . Given that agents propose mechanisms  $p_1, \dots, p_n$ , and adopt deviation strategies  $\delta_1, \dots, \delta_n$  with what probability is each point in  $\Pi(A)$  going to be played? The answer is given by a standard morphism  $\xi_A: \Pi(R(A)) \rightarrow \Delta(\Pi(A))$  whose definition is, again, postponed for the next section.

5. *Definition:*  $F: G \rightarrow G$  is a functor defined on objects  $g = (A, u)$  by  $F(g) = (R(A), v(g))$ , where

$$\begin{array}{ccc} \Pi R A & \xrightarrow{v(g)} & R^n \\ \downarrow \xi_A & \nearrow \tau_{\Pi A}(u) & \\ \Delta \Pi(A) & & \end{array}$$

In words,  $F(g)$  is the game (generated by  $g$ ) whose strategy spaces include all possible proposals to coordinate on  $g$  and all possible deviation strategies, while  $v(g)$  assigns payoffs to each "vector" of proposed mechanisms and deviation strategies by first computing the probability measure on  $\Pi(A)$  induced by these mechanisms and deviation strategies, and then assigning utility to this measure by extending  $u$  from  $\Pi(A)$  to  $\Delta(\Pi(A))$ .

Note that it is not necessary to assume that agents can commit to the mechanisms they announce: mechanisms and deviation strategies are announced simultaneously, and without prior knowledge of the mechanisms and deviation strategies announced by other players. In other words, mechanisms and deviation strategies are treated as ordinary actions in ordinary games, in agreement with the principle we have adopted that the choice of mechanisms will be thought of as a game-theoretic rather than a decision-theoretic problem (of the principal, that is). On the other hand, the ability to commit to mechanisms is an important issue, and the formalism should be easily adaptable to the case where there is a commitment technology, i.e., the case where there is a cost associated with the announcement of each mechanism and each deviation strategy. In fact, it is not even necessary to change the notation: this cost could be directly incorporated into the payoff function  $v(g)$  of the game  $F(g)$  and all the results would go through exactly as they are now; in other words, the precise specification of the payoffs of  $F(g)$  does not matter as far as the feasibility of the instructions is concerned (as long as  $F$  is a continuous functor that is).

Recall that all three applied problems in the introduction involved an infinite regress. To see this in terms of the formalism developed here, let  $F^{t+1}(g) = F(F^t(g))$ ,  $t \geq 0$ . Then, if it is worthwhile to coordinate actions in  $F^t(g)$  by proposing mechanisms on  $F^t(g)$  (i.e., by taking actions in the strategy spaces of  $F^{t+1}(g)$ ), then it is also worthwhile, in general, to coordinate actions in  $F^{t+1}(g)$  by proposing mechanisms on  $F^{t+1}(g)$  (i.e. actions in the strategy spaces of  $F^{t+2}(g)$ ). But hold: doesn't the revelation principle ensure that only mechanisms in  $F(g)$  are needed to achieve any vector of payoffs attainable by general mechanisms? Isn't it, therefore, useless to consider higher levels of the hierarchy  $F^t(g)$ ,  $t \geq 0$ ? To answer this, recall that the revelation principle (and the fact that direct mechanisms suffice for coordination) have already been adopted. This fact in and by itself, however, does not solve the problem of choice of (direct) mechanism: different agents will in general want different (direct) mechanisms to be adopted, because the mechanism adopted will make a difference in the equilibrium payoff attained. What is more, agents are now allowed to try to influence the choice of mechanism by proposing mechanisms, rather than just choosing whether to participate or not in the mechanism proposed by some principal. Why can't we then apply the solution concept to the game  $F(g)$ , compute the equilibrium mechanisms, and get done with it? For the same reason that this procedure is

unsatisfactory when applied to the original game  $g$ . The solution concept might generate "incentives to coordinate" on  $F(g)$  as it did on  $g$ . In other words, the solution concept might predict that the equilibrium of  $F^2(g)$  has some desirable features that the equilibrium of  $F(g)$  does not, in the same way that the equilibrium of  $F(g)$  might have some features that the equilibrium of  $g$  does not. We have, therefore, to examine the equilibria of all levels of the hierarchy  $F^t(g)$  to find out which mechanisms will be implemented and which deviation strategies will be adopted.

We are now at last in a position to tackle the problem of infinite regress. Let  $X(g) = (B, w)$  be the universal game generated by  $g$ : how should it be defined? The obvious answer is to let  $B_i = \bigcup_{t=0}^{\infty} R_i^t(A)$ : this union seems to include all the possible actions that  $i$  could take in any of his strategy spaces  $R_i^t(A)$  in the corresponding games  $F^t(g)$ . There are, however, two problems with this specification of  $B$ :  $B$  contains junk elements and, at the same time, might be incomplete.

To see why  $B$  contains junk elements, we invoke a fact from section 3: (the strategy spaces of)  $g$  can be embedded into (the strategy spaces of)  $F(g)$ , i.e., all the strategy vectors available to players in  $g$  are also available to them (possibly under a different name) in  $F(g)$ : call the embedding  $\phi: g \rightarrow F(g)$  where  $\phi = \langle \phi_1, \dots, \phi_n \rangle$  and  $\phi_i: A_i \rightarrow R_i(A)$  for all  $i$ . Then, consider the following diagram

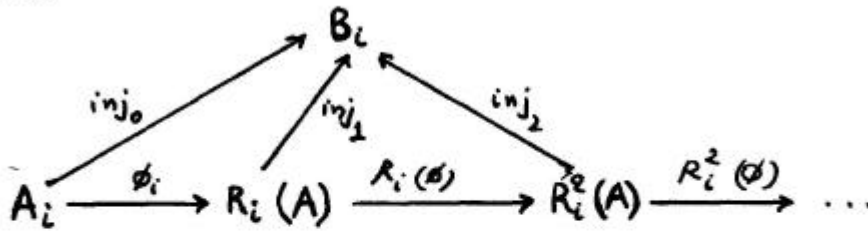
$$A_i \xrightarrow{\phi_i} R_i(A) \xrightarrow{R_i(\phi)} R_i^2(A) \xrightarrow{R_i^2(\phi)} \dots$$

which shows that  $R_i^t(A)$  can be embedded in  $R_i^{t+1}(A)$  for all  $t \geq 0$  (these embeddings are quite trivial, although

at this point we lack the notation necessary to define them explicitly). Let  $b_i$  be an element of  $B_i = \bigcup_{t=0}^{\infty} R_i^t(A)$ : then there is some  $t$  such that  $b_i$  belongs to  $R_i^t(A)$ . But then  $R_i^t(\phi)(b_i)$  represents the same action as  $b_i$ , given that all  $R_i^t(\phi)$  does is to embed  $b_i$  into  $R_i^{t+1}(A)$ : in words  $b_i$  and  $R_i^t(\phi)(b_i)$  are different names for the same action. Hence, they should be identified; they are, however, two distinct elements of  $B$ ; one of them is junk, because it could be eliminated without affecting the opportunities available to the players.

How can we remedy this problem? To answer this, we need to express the fact that  $B$  is too large in diagrammatic terms. Let  $\text{inj}_i: R_i^t(A) \rightarrow B_i$  be the function that injects  $R_i^t(A)$  into  $B_i$ , defined by  $\text{inj}_i(x) = x, x \in$

$R_i^t(A)$ , where  $B_i$  is still defined as  $\bigcup_{t=0}^{\infty} R_i^t(A)$ . Then, by the preceding discussion, the following diagrams do not commute



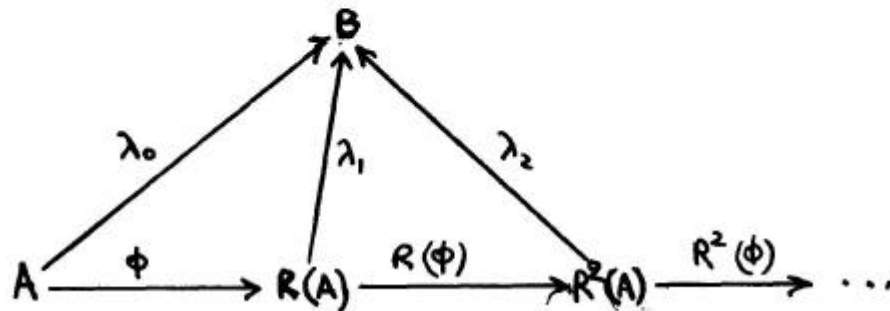
For example, the inequality  $B_i \neq R_i^t(\phi_0(b))$  can be written as

$$\text{inj}_i(b) \neq \text{inj}_{i+1}(R_i^t(\phi_0)b), \text{ or } \\ \text{inj}_i \neq \text{inj}_{i+1} \circ R_i^t(\phi).$$

Hence, if  $B$  is to be the vector of strategy spaces of the universal game  $X(g)$ , it has to satisfy commutativity conditions that fulfill the "no junk" condition: all names of strategies in the sets  $R_i^t(A)$  that represent the same strategy ought to be amalgamated into a single name. At the same time,  $B$  should not miss any strategies available to the players, so it should include (a copy of) each  $R^t(A)$ : this is the "no waste" condition. These two conditions form the content of

**Requirement 1:** There is a natural transformation  $\lambda = \langle \lambda_t: R_t(A) \rightarrow B, t \geq 0 \rangle$ ; each  $\lambda_t$  is an embedding (i.e., a continuous injection).

Four comments about this requirement: first,  $\lambda_t = \langle \lambda_t(1), \dots, \lambda_t(n) \rangle$ , where  $\lambda_t(i): R_t(A) \rightarrow B_i$ , all  $i$ . Secondly, the naturality of  $\lambda$  means that the following diagrams commute (i.e., no junk):

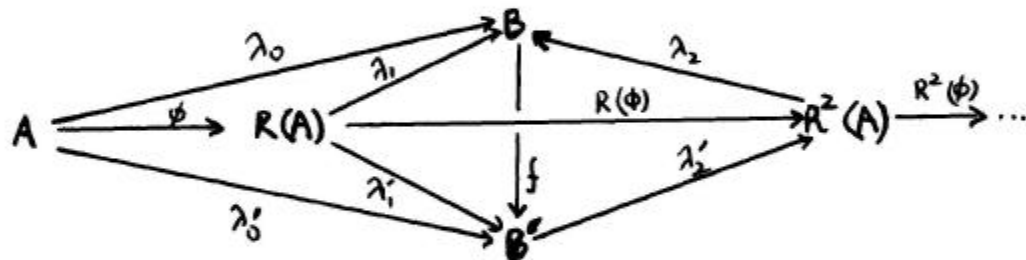


Thirdly, we have required for all  $t$ ,  $\lambda_t = \lambda_{t+1} \circ R_t(\phi)$  rather than  $\text{inj}_t = \text{inj}_{t+1} \circ R_t(\phi)$  because  $B$  needs to be defined only up to isomorphism (the names of the strategies don't matter). Fourthly, the no waste condition is captured by the fact that each  $\lambda_t$  is an embedding.

There are going to be, in general, many vectors of strategy spaces  $B$  that satisfy requirement 1. Some of them will exaggerate the opportunities available to the players because they will contain "extraneous" or "irrelevant" elements, i.e., elements that are not included in some  $R^t(A)$ . To exclude these elements, we impose a "no exaggeration" condition, which forms the content of.

**Requirement 2:**  $B$  should be the smallest vector of strategy spaces that satisfies requirement 1. Hence, if  $\langle \lambda'_t: R^t(A) \rightarrow B', t \geq 0 \rangle$  is a natural transformation and if each  $\lambda'_t$  is an embedding, there is a unique embedding  $f: B \rightarrow B'$  such that  $f \circ \lambda_t = \lambda'_t$ , all  $t$ .

This requirement is captured by the following commuting diagrams:



What do the commutativity conditions mean? If  $x$  is a vector of strategies in  $R^t(A)$ ,  $\lambda'_t(x)$  is the "name" these strategies obtain when they are embedded in  $B'$ , and  $\lambda_t(x)$  is their "name" when embedded in  $B$ . The commutativity conditions guarantee that there is a unique way (namely  $f$ ) to map the name of  $x$  in  $B$  into the name of  $x$  in  $B'$ ; the fact that  $f$  is an embedding means that  $B$  is smaller than  $B'$ .

In the terminology of the companion paper,  $B$  is a colimit (in the appropriate category) of the functor of iterations of  $R$ . Hence, general existence theorems about colimits in this category can, and will, be utilized to show that it is possible to construct a  $B$  that fulfills both requirements.

What about the completeness requirement?  $B_i$  is incomplete if it is not isomorphic to  $R_i(B)$ , i.e., if  $R_i(B)$  contains "more strategies" than  $B_i$ ; in this case  $B_i$  is not truly universal, since it does not include all the opportunities available to player  $i$ . Hence, we impose

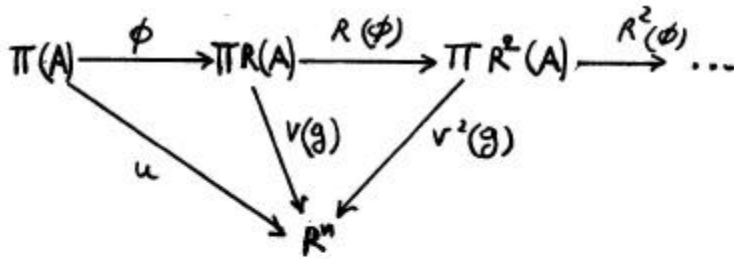
Requirement 3:  $B_i = \Delta \prod (B) \times [B_i^n \rightarrow B_i]$ , all  $i$ . This "equation," however, has no solution in the category of sets, unless  $B_i$  is singleton for all  $i$ ; by requirements 1 and 2, though,  $B_i$  cannot be singleton, because it includes all  $R^t(A)$ ,  $t \geq 0$ . The reason this equation has no non-singleton solution in the category of sets is that the cardinality of  $[B_i^n \rightarrow B_i]$  is always larger than the cardinality of  $B_i$ ; see Manes and Arbib ((1986), exercise 13.3.6) for a simple proof using Cantor's diagonal argument. In this sense, Crawford (1985) was right. To satisfy requirement 3, therefore, we have to work in a different category, where there are enough restrictions on morphisms to "reduce the size" of the function space  $[B_i^n \rightarrow B_i]$ . How should we understand these restrictions? They are similar to the continuity restrictions that are imposed as a matter of course in economics:  $B_i$  is usually equipped with a topology, and functions from  $B_i^n$  to  $B_i$  are required to be continuous with respect to this topology. This is precisely what section 3 does: the topology on each  $B_i$  is a remarkable one, due to Scott ((1980) chapter 3); the space of Scott-continuous functions  $[B_i^n \rightarrow B_i]$  is then sufficiently "small" to ensure that requirement 3 can be fulfilled. Given that the Scott topology is, and ought to be thought of, as different from the usual Hausdorff topologies on strategy spaces, we ought to be able to show how strategy spaces can be endowed with this topology in a "nondistorting" way, i.e., in a way that does not change the structure of the payoffs; one way to do this is developed in section 3.3.

There is another, equivalent but more technical-sounding, way to view the restrictions we impose on the morphisms. Recall that theorem 28 in the companion paper establishes that if the functor  $R$  preserves  $\omega$ -colimits, the colimit of its functor of iterations is a fixed point of  $R$ , i.e., that any  $B$  that satisfies the first two requirements will automatically satisfy the third one. What we buy by changing the category in which we work (by imposing those restrictions on the morphisms) is the fact that  $R$  does preserve  $\omega$ -colimits in this category (while it does not in the category of sets): this is shown in section 4.

Requirements 1, 2 and 3 determine the strategy spaces  $B$  of the universal game uniquely up to isomorphism. To determine the payoff functions  $w$ , note that, by construction, the games  $F^t(g) = (R^t(A), v^t(g))$  satisfy the following property

$$v^{t+1}(g) \circ R^t(\phi) = v^t(g), \text{ all } t \geq 0$$

In other words, the following diagrams commute



The meaning of these conditions is that the same vector of strategies will generate the same vector of utilities no matter where the vector of strategies is embedded into: in other words,  $v^{t+1}(g)$  is an extension of  $v^t(g)$  for all  $t \geq 0$ . This fact, together with requirement 2, implies that there exists a unique  $w: \Pi(B) \rightarrow R^n$  such that for all  $t$ ,  $w \circ \Pi(\lambda_t) = v^t(g)$ . In words, this condition fulfills:

Requirement 4:  $w$  is the unique extension of all  $v^t(g)$ ,  $t \geq 0$ .

The importance of the unique extension property is that no extraneous value judgements are needed to assign payoffs to the strategy vectors of the universal game  $X(g)$ : the payoffs of the original game suffice.

All four requirements taken together amount to a single one: the universal game  $X(g)$  is the colimit of the functor of iterations of  $F$  beginning from the original game  $g$ , namely of the functor

$$g \xrightarrow{\phi} F(g) \xrightarrow{F(\phi)} F^2(g) \rightarrow \dots;$$

$X(g)$  is in addition, a fixed point of  $F$ . Given the fact that  $F$  preserves  $\omega$ -colimits (shown in section 4) and given theorem 28 of the companion paper, we have motivated:

6. *Definition*: The universal game generated by  $g$  is the least fixed point of  $F$  that includes  $g$ ; call this game  $X(g)$ .

### 2.3 Coordination - Proof Solution Concepts.

Let

$$\gamma(g): \Delta \Pi P_1 X(g) \rightarrow \Delta \Pi P_1(g)$$

be the standard morphism that reduces probability measures on the aggregate strategy space of the universal game  $X(g)$  to probability measures on the aggregate strategy space of the original game  $g$ .

7. *Definition:* A solution concept  $\sigma \in \Sigma$  is coordination - proof if for all games  $g$

$$P_2 \sigma(g) = \gamma(g)(P_2 \sigma(X(g)))$$

i.e., if the prediction of  $\sigma$  on the original game  $g$  coincides with the projection on the original game of the prediction of  $\sigma$  on the universal game  $X(g)$ . The intuitive idea behind the definition is that agents will attempt to coordinate their actions on  $g$ , and that the solution concept predicts the outcome of these attempts to coordinate: whatever this outcome is, when projected on the original game it should be identical to the prediction of  $\sigma$  on the original game. Otherwise  $\sigma$  is self-contradictory. It should be stressed that coordination-proofness does not in general imply that the "cooperative" or "efficient" outcome will prevail; it is a consistency, not an optimality, condition.

Finally, we show how coordination-proof solution concepts can be obtained as fixed point of some functor. Let

$$\Psi: G \times \Sigma \rightarrow \text{Fib}(\Delta \Pi P_1)$$

be defined by

$$\Psi(g, \sigma) = (g, \gamma(g)(P_2 \sigma(X(g))))$$

and let

$$\Psi: \Sigma \rightarrow [G \rightarrow \text{Fib}(\Delta \Pi P_1)]$$

be defined by

$$\Psi(\sigma)(g) = \Psi(g, \sigma) \quad \sigma \in \Sigma, \quad g \in G$$

Then,  $\sigma$  is coordination-proof if  $\sigma(g) = \Psi(g, \sigma)$ , for all  $g$ , i.e., if

$$\sigma = \psi(\sigma)$$

It will turn out that  $\psi(\sigma): G \rightarrow \text{Fib}(\Delta\text{IIP})$  is  $\omega$ -cocontinuous and satisfies the weak revelation principle if  $\sigma$  has the same properties. Hence  $\psi$  maps  $\Sigma$  into  $\Sigma$ , and  $\sigma = \psi(\sigma)$  is a functorial fixed point equation.

Section 5 shows that each solution concept  $\sigma$  in  $\Sigma$  can be "completed into" a coordination-proof solution concept  $\hat{\sigma}$ , obtained by taking the colimit of the functor  $\sigma \rightarrow \psi(\sigma) \rightarrow \psi^2(\sigma) \rightarrow \dots$

A possible objection to the requirement that solution concepts be coordination-proof is that  $g$  and  $X(g)$  are different games, and that either  $g$  or  $X(g)$  is played: hence, it is not that unreasonable to expect a violation of coordination-proofness; after all, players have more opportunities for coordination in  $X(g)$  than in  $g$ , so the prediction of the solution concept on  $X(g)$  should be different from its prediction on  $g$ . To overcome this objection, recall the motivating examples: while  $g$  describes the physical rules of the game,  $X(g)$  describes all the opportunities for coordination available to the players. If an economic environment is described by  $g$ , then  $X(g)$  is played, given that players will not miss any opportunities to coordinate. Recall that the solution concept assigns to each game  $g$  a prediction  $p$  that answers the question "what is going to be played in  $g$ ?" Recall again the motivating examples: the solution concept applied to the universal game  $X(g)$  predicts which cooperative agreements will be successful (i.e., implemented, carried out) in an environment described by the original game  $g$ : this means that it predicts which mechanisms will be actually proposed and which deviation strategies will be actually adopted. When this prediction is projected on  $g$ , it produces the vector of strategies prescribed for  $g$  by these cooperative agreements. But given that these cooperative agreements are successful, their prescribed strategies on  $g$  ought to be played. Hence, the answer to the question "what is going to be played in game  $g$ ?" ought to be "the projection on  $g$  of the prediction of the solution concept on  $X(g)$ ." But hold: it is the prediction of the solution concept on  $g$  that answers the question "what is going to be played in  $g$ ?", by definition. Hence coordination-proofness.

This argument might raise a second objection, that of triviality. After all, the argument goes, all we are doing is apply a trivial three-step process: (a) take a solution concept (say Nash, to fix ideas, and forgetting the multiplicity problem) (b) apply it only to the universal game generated by each game, and (c) project the prediction of the solution concept on the original game. Then, we have a new solution concept that is

coordination-proof (and agrees with the original solution concept on universal games.) This would indeed be trivial, but remember that we impose two requirements on solution concepts: continuity and satisfaction of the weak revelation principle. It is by no means obvious that the new solution concept obtained by this three-step process satisfies these requirements, and this is the nontrivial part of coordination-proofness.

This argument might raise a third objection, that of relevance. Why are continuity and the (weak) revelation principle properties that solution concepts should satisfy? In fact, why should solution concepts be functors, let alone continuous functors? Roughly, the revelation principle restricts the kind of mechanisms needed for coordination, and is therefore crucial for the construction of universal games. Functoriality and continuity will take a little longer to motivate. Recall that the Nash solution concept is not just an amorphous collection of answers to questions of the form "what is going to happen in game  $g$ ?": these answers are obtained "in the same way" in all games, i.e., by computing the fixed points of the best reply correspondence of each game. This property is captured abstractly by the requirement that a solution concept be a functor into a category of fibrations, because of the form of morphisms on the category of fibrations (section 3.2). But why a continuous functor (i.e., a functor that preserves  $\omega$ -colimits?). To answer this, recall that continuity means that if a game  $g$  is obtained as a colimit of other games  $g_i$ , then the equilibrium of the game  $g$  is obtained as the colimit of the equilibria of the games  $g_i$ . A similar continuity property is enjoyed by (some refinements of) Nash equilibria: see Hillas ((1990), pp. 1384-1385). While in Hillas' case continuity is defined in terms of a topology on the space of best-reply correspondences from a given product of strategy spaces to itself, in our case continuity has to be defined categorically, since we allow the strategy spaces to vary, and since the collection of all games is not a set. An analogy with a computer science application might be useful here: operations like addition, multiplication, forming lists of elements of sets, are polymorphic: they can be defined on many different domains (reals, integers, positive reals, etc.) but on each domain they operate "in the same way." One way to model polymorphic operations abstractly is to represent them by continuous sections, in exactly the same way we have defined solution concepts; see Coquand, Gunter and Winskel (1989).

### 3. THE MAIN CATEGORIES

#### 3.1 The Base Category D

Which categories are appropriate for carrying out the constructions of section 2? There is no unique answer to this question (Gunter (1989)), but it seems that the least complicated is the category D of domains and continuous functions studied in Manes and Arbib ((1986, Ch. 13). The objects of D are domains, i.e., partially ordered sets with a least (bottom) element  $\perp$  and with least upper bounds of all their increasing sequences. The morphisms of D are monotonic functions between domains that preserve least upper bounds of increasing sequences; call them continuous functions.

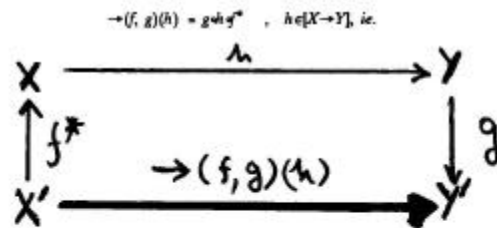
An extremely useful subcategory of D is  $D_E$ , the category of domains and embeddings. An embedding is a morphism  $f: x \rightarrow y$  of D such that there is  $g: y \rightarrow x$  that is monotonic and satisfies

$$(1) \quad g(f(x)) = x \quad \text{all } x$$

$$(2) \quad f(g(y)) \leq y \quad \text{all } y$$

It turns out that there is at most one such  $g$ , call it  $f^*$ , and that if  $f$  is an embedding it is injective and maps bottom elements into bottom elements, while  $f^*$  is surjective and maps bottom elements into bottom elements. The composite of two embeddings is also an embedding. Both D and  $D_E$  have all  $\omega$ -colimits;  $\omega$ -colimits in  $D_E$  are also  $\omega$ -colimits in D. If  $T: \omega \rightarrow D_E$  is a functor and  $\lambda: T \rightarrow X$  a natural transformation (where X is a domain in  $D_E$ ) then  $(X, \lambda)$  is a colimit of T if and only if  $\bigvee_{n=1}^{\infty} \lambda_n(\lambda_n^*(x)) = x$  for all  $x$  in X, where  $\bigvee$  denotes the operation of taking the least upper bound of a sequence. If X and Y are domains, then so is  $X \times Y$ , ordered coordinate-wise, and  $[X \rightarrow Y]$ , the set of continuous functions from X to Y, ordered pointwise ( $f \leq g$  if  $f(x) \leq g(x)$  for all  $x$  in X.) Finally, functors that preserve  $\omega$ -colimits will be called just continuous (rather than  $\omega$ -cocontinuous).

The function space functor  $\rightarrow: D_E \times D_E \rightarrow D_E$  is defined by  $\rightarrow(X, Y) = [X \rightarrow Y] =$  all continuous functions from X to Y (not only all embeddings from X to Y). On morphisms, the functor is defined as follows: let  $f: X \rightarrow X'$ ,  $g: Y \rightarrow Y'$  be embeddings; then  $\rightarrow(f, g): [X \rightarrow Y] \rightarrow [X' \rightarrow Y']$  is given by



Scott (1980), Chapter 4) has shown that  $\rightarrow$  is continuous (in a closely related category); proofs can also be found in Smyth and Plotkin (1982) and Manes and Arbib ((1986), Ch. 13).

To define the probability functor  $\Delta: D_E \rightarrow D_E$  we need Borel sets on domains, hence a topology on domains induced by the order. For each domain  $X$ , the Scott topology consists of all subsets  $U$  of  $X$  that satisfy the following two properties:

- $x \in U$  and  $x \leq y$  imply  $y \in U$ .
- If  $\langle x_n \rangle$  is an increasing sequence and if  $\forall x_n \in U$  then there is some  $n$  such that  $x_n \in U$ .

In  $\bar{\mathbb{R}}$  with the usual order, the Scott opensets are all the sets of the form  $(\alpha, \infty]$ , i.e., the open half-rays. A function  $f: X \rightarrow Y$  is Scott-continuous if it is continuous in the Scott topology; it turns out that  $f: X \rightarrow Y$  is Scott continuous if and only if it is monotonic and preserves least upper bounds of increasing sequences. A function  $f: [a, b] \rightarrow \bar{\mathbb{R}}$ , for example, is Scott-continuous if for each  $\alpha \in \mathbb{R}$  there is a  $\beta \in [\alpha, b)$  such that

$$f^{-1}((\alpha, \infty]) = (\beta, b]$$

For each domain  $X$ ,  $\Delta(X)$  is the set of all Borel probability measures on  $X$ . To make  $\Delta(X)$  into a domain, define an order on it by  $p \leq q$  iff  $p(U) \leq q(U)$  for all Scott-open subsets  $U$  of  $X$ . If  $X = \bar{\mathbb{R}}$ , for example, the Scott-open sets are all sets of the form  $(\alpha, \infty]$ , and  $p \leq q$  if  $p$  is stochastically dominated by  $q$ . With this ordering on  $\Delta(X)$ , its bottom element is the measure that assigns unit mass to the bottom element of  $X$ .

On morphisms,  $\Delta$  is defined as follows: to each  $f: X \rightarrow Y$  in  $D_E$ ,  $\Delta$  assigns  $\Delta(f): \Delta(X) \rightarrow \Delta(Y)$  given by  $\Delta(f)(p) = p \circ f^{-1}$ . The continuity of  $\Delta$  is proven in the appendix. More information on  $\Delta$  is in Graham (1989) and Saheb-Djahromi (1980).

### 3.2 Fibrations and Sections

The concepts of fibration and section will be used in defining the categories of games, mechanisms, and solution concepts, so it seems appropriate to give their general definitions (they can be found in p. 131 of Coquand, Gunter and Winskel (1989)). Let  $F: C \rightarrow E$  be a functor from a category  $C$  to a category of categories  $E$ . The category  $\text{Fib}(F)$  of fibrations of  $F$  consists of

- Objects: all pairs  $(X, t_x)$ , where  $X \in C$  and  $t_x \in F(X)$
- Morphisms  $(X, t_x) \rightarrow (Y, t_y)$  are pairs  $(f, \alpha)$  where  $f: X \rightarrow Y$  and  $\alpha: F(f)(t_x) \rightarrow t_y$
- Composition: if  $(f, \alpha): (X, t_x) \rightarrow (Y, t_y)$  and  $(g, \beta): (Y, t_y) \rightarrow (Z, t_z)$  then  $(g, \beta) \circ (f, \alpha) = (g \circ f, \beta \circ F(g)(\alpha))$
- Identity:  $(id_x, id_{t_x}): (X, t_x) \rightarrow (X, t_x)$

The projection  $p: \text{Fib}(F) \rightarrow C$  is a functor that takes  $(X, t_x)$  to  $X$  and  $(f, \alpha)$  to  $f$ . A section of  $F$  is a functor  $s: G \rightarrow \text{Fib}(F)$  such that  $\text{pos} = id_C$  i.e., such that  $s(X) = (X, t_x)$ ,  $t_x \in F(X)$ . A morphism  $v: s \rightarrow s'$  between two sections is a natural transformation with the additional property  $p(v_x) = id_x$  for all  $X$  in  $C$ . Hence each  $v_x: (X, t_x) \rightarrow (X, t'_x)$  is equal to  $(id_x, \alpha_x)$  for some  $\alpha_x: F(id_x)(t_x) \rightarrow t'_x$  in  $F(X)$ . Let  $\text{Sect}(F)$  be the category of sections of  $F$ . Notice that  $D$  and  $D_E$  can be taken to be categories of categories, since each domain can be considered as a category with objects its elements and morphisms  $x \rightarrow x'$  if  $x \leq x'$ . Then, for any functor  $F: C \rightarrow D_E$ ,  $\text{Sect}(F)$  is a partial order, in the sense that there is at most one morphism between any two sections  $\sigma$  and  $\sigma'$  of  $F$ . To see this, notice that if  $v: \sigma \rightarrow \sigma'$  is such a morphism, there is a morphism  $\alpha_x: t_x \rightarrow t'_x$  in  $F(X)$  for each  $X$  in  $D_E$ ; but  $F(X)$  is a domain, so  $\alpha_x: t_x \rightarrow t'_x$  means  $t_x \leq t'_x$ , and so  $v_x$  is uniquely determined by the order on  $F(X)$ , for all  $X$ .

If  $F: C \rightarrow D_E$  is a continuous functor and if  $C$  has all  $\omega$ -colimits, then  $\omega$ -colimits in  $\text{Fib}(F)$  assume a simple form. Let  $T: \omega \rightarrow \text{Fib}(F)$  be a functor; let  $T_n = (X_n, t_n)$ ,  $t_n \in F(X_n)$ , and  $T(n \rightarrow n+1) = (f_n, \alpha_n)$  where  $f_n: X_n \rightarrow X_{n+1}$  is in  $C$  and  $\alpha_n: F(f_n)(t_n) \rightarrow t_{n+1}$  is in  $F(X_{n+1})$ . Let  $(X, \lambda)$  be the colimit in  $C$  of the functor

$$X_1 \xrightarrow{f_1} X_2 \xrightarrow{f_2} X_3 \rightarrow \dots$$

Then, let  $t = \varinjlim_{n \rightarrow \infty} F(\lambda_n)(t_n)$  and let  $\beta_n: F(\lambda_n)(t_n) \rightarrow t$  in  $F(X)$  (such a morphism exists in  $F(X)$  because  $F(\lambda_n)(t_n) \leq t$  for all  $n$ ). The colimit of  $T$  in  $\text{Fib}(F)$  is then  $((X, t), (\lambda, \beta))$ .

Under the same conditions on  $F$ , a continuous section  $\sigma: C \rightarrow \text{Fib}(F)$  can be easily characterized. If  $T: \omega \rightarrow C$  is a functor with colimit  $(X, \lambda)$ , then  $\sigma$  preserves it, i.e.,  $(\sigma(X), \sigma(\lambda))$  is a colimit of  $\sigma \circ T$ . To see the meaning of this, let  $T(n \rightarrow n+1) = f_n: T_n \rightarrow T_{n+1}$ .

Then, let

$$\begin{aligned} \sigma(T_n) &= (T_n, t_n) & \sigma(X) &= (X, t) \\ \sigma(f_n) &= (f_n, \alpha_n) & \sigma(\lambda_n) &= (\lambda_n, \beta_n) \end{aligned}$$

where  $\alpha_n: F(f_n)(t_n) \rightarrow t_{n+1}$ ,  $\beta_n: F(\lambda_n)(t_n) \rightarrow t$ , i.e.,  $F(f_n)(t_n) \leq t_{n+1}$ ,  $F(\lambda_n)(t_n) \leq t$ . By the form of colimits in  $\text{Fib}(F)$ ,  $\sigma$  is continuous iff  $t = \varinjlim_{n \rightarrow \infty} F(\lambda_n)(t_n)$ , given that  $(X, \lambda)$  is already a colimit of  $T$ .

### 3.3 The Category $G$ of games

Let  $F_1: D_E^0 \rightarrow D_E$  be the functor defined by

$$\begin{aligned} \bullet F_1(A) &= [\Pi(A) \rightarrow \bar{R}^n] & A \in D_E^0 \\ \bullet F_1(\lambda)(u) &= u \circ \Pi(\lambda^*) & \lambda: A \rightarrow A' \text{ in } D_E, u \in F_1(A) \end{aligned}$$

$G$  is then the category of fibrations of  $F_1$ , with objects all pairs  $(A, u)$ ,  $u \in F_1(A)$ , and morphisms  $(A, u) \rightarrow (A', u')$  all pairs  $(\lambda, \alpha)$  such that  $\lambda: A \rightarrow A'$  is in  $D_E^0$  and  $\alpha: F_1(\lambda)(u) \rightarrow u'$  in  $F_1(A)$ , i.e.,  $u \circ \Pi(\lambda^*) \leq u'$ :

$$\begin{array}{ccc} A & & \Pi(A) \\ \downarrow \lambda & & \uparrow \alpha \\ A' & & \Pi(A') \end{array} \quad \begin{array}{ccc} \Pi(A) & \xrightarrow{u} & \bar{R}^n \\ \Pi(A^*) & \leq & \\ \Pi(A') & \xrightarrow{u'} & \bar{R}^n \end{array}$$

The continuity of  $F_1$  and the fact that  $D_2^{\omega}$  has all  $\omega$ -colimits guarantee that  $G$  has all  $\omega$ -colimits as well.

The requirements of this definition seem to be very strong: for example, finite games do not come equipped with orders on their strategy spaces; games like Cournot's, Bertrand's, or those reported in Milgrom and Roberts (1989), do come equipped with orders, but their payoff functions are not monotonic with respect to these orders; finally, strategic market games like those in Dubey, Mas-Colell and Shubik (1980), Green (1984) or Khan and Sun (1990), come equipped with topologies on their strategy spaces, and payoff functions continuous in these topologies, but it is not clear whether these topologies are compatible with some order structure on strategy spaces.

All of these games, however, can be faithfully embedded into the category  $G$  considered in the present paper. Let  $(B, v)$  be a game, not necessarily in  $G$ ; assume that for each  $b \in \Pi(B)$ , and each  $i=1,2, \dots, n$ ,  $v_i(b) > -\infty$ . Then,  $(B, v)$  can be embedded in  $G$  as follows:

Step 1: Define  $A_i = B_i \cup \{\perp\}$ , where  $\perp$  is an element that does not belong to any of the  $B_i$ 's.

Step 2: Define an order on  $A_i$ , the flat order, by  $a_i \leq a_i'$  if  $a_i = \perp$  or  $a_i = a_i'$ .

Step 3: Define  $u_i: \Pi(A_i) \rightarrow \bar{R}$  by

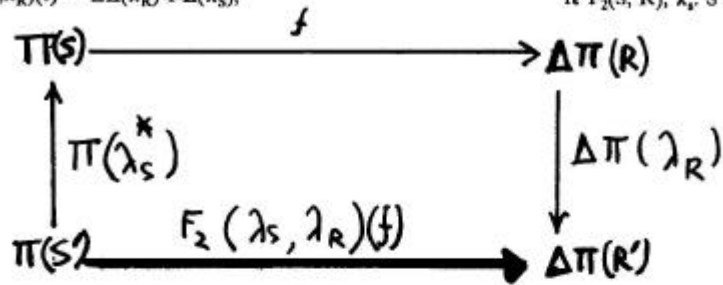
$$u_i(a) = \begin{cases} v_i(a) & a \in \Pi(B) \\ -\infty & \text{otherwise} \end{cases}$$

It is easy to show that  $(A_i, u_i)$  belongs to  $G$ ; notice that  $u_i$  is monotonic in the flat order, while it might not be monotonic with respect to the order that  $B$  originally carried (if any); similarly,  $u_i$  is continuous with respect to the flat order, while  $v_i$  might be continuous with respect to the topology carried by  $B$  (if any), and this topology could very well be incompatible with any order.

### 3.4 The Category of Mechanisms M

Let  $F_2: D_E^0 \times D_E^0 \rightarrow D_E$  be defined by

- $F_2(S, R) = [\Pi(S) \rightarrow \Delta(\Pi(R))]$ ,  $S, R \in D_E^0$
- $F_2(\lambda_S, \lambda_R)(f) = \Delta\Pi(\lambda_R) \circ f \circ \Pi(\lambda_S^*)$ ,  $f \in F_2(S, R), \lambda_S: S \rightarrow S', \lambda_R: R \rightarrow R' \text{ in } D_E^0$



The category  $M = \text{Fib}(F_2)$  of mechanisms has objects all triples  $(S, R, f), f \in F_2(S, R)$ , and morphisms  $(S, R, f) \rightarrow (S', R', f')$  all triples  $(\lambda_S, \lambda_R, \alpha), \lambda_S: S \rightarrow S', \lambda_R: R \rightarrow R', \text{ in } D_E^0$  and  $\alpha: F_2(\lambda_S, \lambda_R)(f) \rightarrow f' \text{ in } F_2(S', R')$  i.e.,

$$\Delta\Pi(\lambda_R) \circ f \circ \Pi(\lambda_S^*) \leq f'$$

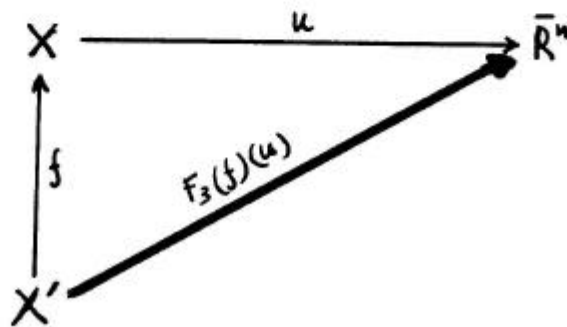
The continuity of  $F_2$  and the fact that  $D_E^0 \times D_E^0$  has all  $\omega$ -colimits guarantee that  $M$  has all  $\omega$ -colimits as well.

### 3.5 Preferences over Lotteries

Let  $F_3: D^{op} \rightarrow D, F_4: D^{op} \rightarrow D$  be defined by

$$F_3(X) = [X \rightarrow \bar{R}^n]$$

$$F_3(f)(u) = u \circ f, \quad u \in F_3(X), f: X' \rightarrow X \text{ in } D.$$



$$F_4(X) = [\Delta(X) - \bar{R}^n]$$

$$F_4(f)(u) = u \circ \Delta(f) \quad , u \in F_4(X), f: X' \rightarrow X \text{ in } D$$

$$\begin{array}{ccc} \Delta(X) & \xrightarrow{u} & \bar{R}^n \\ \uparrow \Delta(f) & \nearrow F_4(f)(u) & \\ \Delta(X') & & \end{array}$$

Both  $F_3$  and  $F_4$  are contravariant functors when defined on  $D$ : for example, if  $f: X' \rightarrow X$  then  $F_3(f)$  maps  $F_3(X)$  into  $F_3(X')$ ; hence they are ordinary functors when defined on  $D^{op}$ , the opposite category of  $D$ .

Preferences over lotteries are then defined by a natural transformation  $\tau: F_3 \rightarrow F_4$ , i.e. a collection of morphisms in  $D$ ,  $\langle \tau_x: F_3(X) \rightarrow F_4(X), X \in D \rangle$ , such that if  $f: X' \rightarrow X$  then the following diagram commutes

$$\begin{array}{ccccc} X & [X \rightarrow \bar{R}^n] & \xrightarrow{\tau_x} & [\Delta(X) \rightarrow \bar{R}^n] & \\ \uparrow f & \downarrow F_3(f) & & \downarrow F_4(f) & \\ Y & [Y \rightarrow \bar{R}^n] & \xrightarrow{\tau_y} & [\Delta(Y) \rightarrow \bar{R}^n] & \end{array}$$

Hence, for each  $u: X \rightarrow \bar{R}^n$ ,

$$\tau_x(u) \circ \Delta(f) = \tau_y(u \circ f)$$

This is simply a change-of-variable formula; to see this, suppose that  $\tau$  is generated by expected utility axioms. Then

$$\tau_x(u)(p) = \int_X u dp \quad \text{and the naturality condition reduces to, for all } q \in \Delta(Y)$$

$$\tau_x(u)(\Delta(f)(q)) = \tau_y(u \circ f)(q), \text{ i.e.,}$$

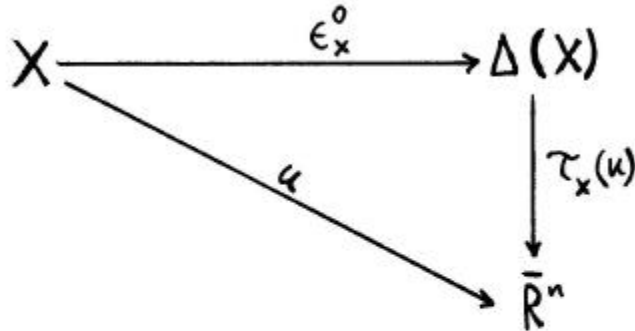
$$\int_X u d(q \circ f^{-1}) = \int_Y (u \circ f) dq$$

which is precisely the change-of-variable formula in, say, p. 163 of Halmos (1950).

We will require two more conditions on  $\tau$ ; the first is simply that  $\tau_x(u)$  is an extension of  $u: X \rightarrow \bar{R}^n$  in the sense that  $\tau_x(u)$  assigns to probability measures with unit mass on a point  $x$  in  $X$  the same utility assigned by  $u$  to  $x$ . To express this, let  $\epsilon_x^X: X \rightarrow \Delta(X)$  be the (Scott-continuous) map that assigns to each  $x$  in  $X$  the Dirac measure  $\epsilon_x^X(x)$ , defined by

$$e_x^0(x)(U) = \begin{cases} 1 & x \in U \\ 0 & x \notin U \end{cases}$$

for each open  $U$  in  $X$ . Then, we require the following diagram to commute for all  $X$  in  $D$ , all  $u \in F_3(X)$

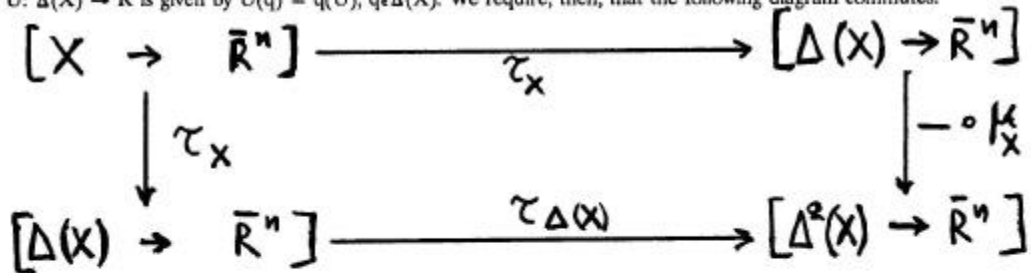


This condition is trivially satisfied by expected utility.

The second requirement involves the (Scott-continuous) function  $\mu_x: \Delta^2(X) \rightarrow \Delta(X)$  that reduces compound to simple lotteries; if  $p \in \Delta^2(X)$ ,  $\mu_x(p) \in \Delta(X)$  is defined by

$$\mu_x(p)(U) = \int_{\Delta(X)} \hat{U} dp \quad , \quad U \text{ open in } X$$

where  $\hat{U}: \Delta(X) \rightarrow \bar{R}$  is given by  $\hat{U}(q) = q(U)$ ,  $q \in \Delta(X)$ . We require, then, that the following diagram commutes:



ie, that for each  $u: X \rightarrow \bar{R}^n$  and each  $p \in \Delta^2(X)$

$$\tau_X(u)(\mu_X(p)) = \tau_{\Delta(X)}(\tau_X(u))(p)$$

In words, a compound lottery  $p \in \Delta^2(X)$  can be evaluated in two ways; the first way involves reducing  $p$  to a simple lottery  $\mu_X(p)$  and then evaluating  $\mu_X(p)$  by extending  $u$  from  $X$  to  $\Delta(X)$ ; the second way involves evaluating  $p$  directly by extending  $\tau_X(u)$  from  $\Delta(X)$  to  $\Delta^2(X)$ ; both ways assign to  $p$  the same vector of utilities.

### 3.6 The category $\Sigma$ of solution concepts

A solution concept will assign to each game  $g$  a probability distribution on its aggregate strategy space  $\Pi P_1(g)$ ; formally, it will be a section  $\sigma: G \rightarrow \text{Fib}(\Delta \Pi P_1)$  of the functor  $\Delta \Pi P_1: G \rightarrow D_G$ ; let  $\text{Sect}(\Delta \Pi P_1)$  be the category of these sections.

Not every section will be considered as a legal solution concept; attention will be restricted to continuous sections that satisfy the weak revelation principle. Given the continuity of  $\Delta \Pi P_1$  and the fact that  $G$  has all  $\omega$ -colimits, continuity of a section  $\sigma$  means that if  $T: \omega \rightarrow G$ ,  $T_n = g_n = (\Lambda_n, u_n)$ ,  $T(n \rightarrow n+1) = f_n: g_n \rightarrow g_{n+1}$ , is a functor with colimit  $(g, \lambda)$ , and

$$\text{if } \sigma(g) = (g, p), \sigma(g_n) = (g_n, p_n), \text{ then } p = \varinjlim_{n \rightarrow 1} \Delta \Pi P_1(\lambda_n)(p_n).$$

To define solution concepts that satisfy the weak revelation principle, recall that the functor  $T: \text{MxG} \rightarrow G$  assigns to each mechanism  $m = (S, R, f)$  and to each game  $g = (A, u)$  a new game  $T(m, g) = (B, v)$ , where  $B_i = S_X [R_i \rightarrow A_i]$  and  $v(s, \delta) = \tau_{\Pi(A)}(u)(f(s) \circ \delta^{-1})$ ; if  $(\lambda_S, \lambda_R): m \rightarrow m'$  and  $\lambda_A: g \rightarrow g'$ , then  $T(\lambda_S, \lambda_R, \lambda_A): T(m, g) \rightarrow T(m', g')$  is given by  $T(\lambda_S, \lambda_R, \lambda_A)(s, \delta) = (\lambda_S(s), \lambda_A \circ \delta \circ \lambda_R^{-1})$ . Let  $\sigma(T(m, g)) = (T(m, g), q)$ , where  $q \in \Delta(\Pi(B))$ ; define  $f': 1 \rightarrow \Delta(\Pi(A))$  by  $f'(1) = \mu_A(\Delta(k)(q))$

$$\begin{array}{ccc}
 1 & \xrightarrow{q} & \Delta \Pi(B) \\
 \downarrow f' & & \downarrow \Delta(\kappa) \\
 \Delta \Pi(A) & \xleftarrow{\mu_A} & \Delta^2 \Pi(A)
 \end{array}
 \qquad
 \begin{array}{c}
 \Pi(B) \\
 \downarrow \kappa \\
 \Delta \Pi(A)
 \end{array}$$

$f'(1)$  is the measure on  $\Delta(\Pi(A))$  induced by  $q$ ; recall that  $k(s, \delta) = f(s) \circ \delta^{-1}$ ,  $\Delta(k)(v) = \kappa v^{-1}$  for each  $v \in \Delta(\Pi(B))$ , and that  $\mu_A$  reduces compound to simple lotteries. Let  $m' = (1, A, f')$  be the direct mechanism on  $g = (A, u)$  that recommends actions with the same probability that  $\sigma$  predicts that actions will be taken in  $T(m, g)$ . A solution concept  $\sigma$  satisfies the weak revelation principle if for all mechanisms  $m$  and games  $g$ ,  $P_{\sigma} T(m', g) \geq \text{id}_A$ . The revelation principle proper would require that  $P_{\sigma} T(m', g) = \text{id}_A$ , i.e., that whenever  $\sigma$  predicts that equilibrium actions in  $T(m, g)$  will induce a probability distribution

q on the aggregate strategy space of g, then a mechanism m' that recommends actions in g with probability law q ought to be obeyed with probability one.

The category  $\Sigma$  of solution concepts is, then, the full subcategory of  $\text{sect}(\Delta\Pi_1)$  that consists of continuous sections of  $\Delta\Pi_1$  satisfying the weak revelation principle.

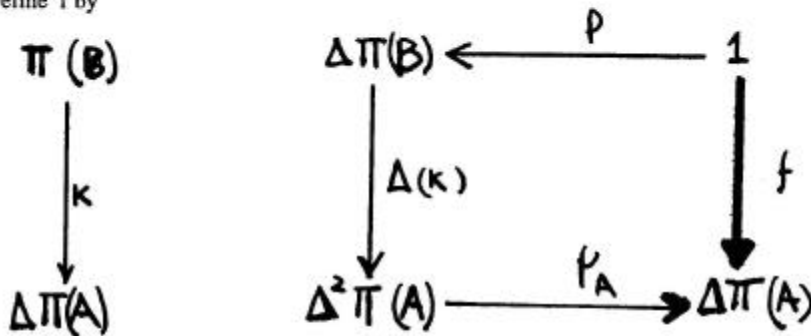
As in Myerson (1982), the weak revelation principle implies that direct mechanisms suffice for coordination on any game g. To show this, say that  $(S, R) \in D_{\mathbb{R}^n} \times D_{\mathbb{R}^n}$  is  $\sigma$ -adequate for coordination on  $A \in D_{\mathbb{R}^n}$  if for any mechanism  $m = (X, Y, h)$  and any payoff function  $u: \Pi(A) \rightarrow \mathbb{R}^n$  there exists a D-morphism  $f: \Pi(S) \rightarrow \Delta(\Pi(R))$  such that coordination on  $g = (A, u)$  with  $m = (X, Y, h)$  generates less equilibrium utility, according to  $\sigma$ , than coordination on  $g$  with  $m' = (S, R, f)$ . In other words, if  $T(m, g) = (B, v)$ ,  $T(m', g) = (B', v')$ ,  $Pr_2 \sigma(T(m, g)) = p$ ,  $Pr_2 \sigma(T(m', g)) = p'$ , then  $\tau_{\Pi_B}(v)(p) \leq \tau_{\Pi_{B'}}(v')(p')$ .

**Theorem:** The message-recommendation spaces  $(1, A)$  are  $\sigma$ -adequate for coordination on  $A$ , for all  $\sigma \in \text{Sect}(\Delta\Pi_1)$  that satisfy the weak revelation principle, and all  $A \in D_{\mathbb{R}^n}$ .

*Proof:* Let  $m = (X, Y, h)$ ,  $u: \Pi(A) \rightarrow \mathbb{R}^n$ ,  $S = 1$ ,  $R = A$ ; we have to construct  $f: 1 \rightarrow \Delta(\Pi(A))$  such that if  $m' = (1, A, f)$ ,  $T(m, g) = (B, v)$ ,  $T(m', g) = (B', v')$ ,  $Pr_2 \sigma(T(m, g)) = p$ ,  $Pr_2 \sigma(T(m', g)) = p'$ , then

$$\tau_{\Pi_B}(v)(p) \leq \tau_{\Pi_{B'}}(v')(p').$$

Define  $f$  by



i.e.,  $f(1) = \mu_A(\Delta(\kappa)(p))$

By the weak revelation principle  $p' = Pr_2 \sigma(T(m', g)) \geq id_A$ . Then:

$$\begin{aligned}
\tau_{\Pi_B}(\nu')(p) &\geq && \text{(Monotonicity of } \tau_{\Pi_B}(\nu')\text{)} \\
\tau_{\Pi_B}(\nu')(id_A) &= && \text{(property of } \tau \text{ on dirac measures)} \\
&= \nu'(id_A) = && \text{(definition of } \nu'\text{)} \\
\tau_{\Pi_A}(u)(f(1) \circ id_A^{-1}) &= \\
&= \tau_{\Pi_A}(u)(f(1)).
\end{aligned}$$

At the same time,

$$\begin{aligned}
\tau_{\Pi_B}(\nu)(p) &= && \text{(definition of } \nu\text{)} \\
\tau_{\Pi_B}(\tau_{\Pi_A}(u) \circ k)(p) &= && \text{(naturality of } \tau\text{)} \\
\tau_{\Delta \Pi_A}(\tau_{\Pi_A}(u))(\Delta(k)(p)) &= && \text{(property of } \tau \text{ on compound lotteries)} \\
\tau_{\Pi_A}(u)(\mu_A \circ \Delta(k)(p)) &= && \text{(definition of } f(1)\text{)} \\
\tau_{\Pi_A}(u)(f(1)). &&& \text{QED.}
\end{aligned}$$

#### 4. UNIVERSAL GAMES

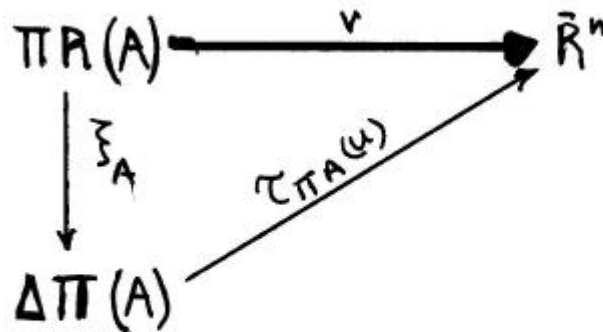
The functor  $R_i: D_E^n \rightarrow D_E$  is defined by  $R_i(A) = \Delta\Pi(A) \times [A_i^n \rightarrow A_i]$ , and for each  $\lambda: A \rightarrow A'$  in  $D_E^n$ ,  $R_i(\lambda) = \Delta\Pi(\lambda) \times (\lambda_i \circ - \circ \lambda_i^*)$ . In words, it assigns to each vector of strategy spaces  $A$  the space of strategies that agent  $i$  can take to "change the rules" of the game:  $i$  can propose a (direct) mechanism in  $\Delta\Pi(A)$  and adopt a deviation strategy  $\delta_i: A_i^n \rightarrow A_i$ . Let  $R: D_E^n \rightarrow D_E^n$  be the functor  $R(A) = (R_1(A), \dots, R_n(A))$ ,  $R(\lambda) = (R_1(\lambda), \dots, R_n(\lambda))$ .

The D-morphism  $\xi_A: \Pi(R(A)) \rightarrow \Delta(\Pi(A))$  assigns to each vector  $(p, \delta)$  of strategies in  $\Pi(R(A))$  the probability measure they induce on  $\Delta\Pi(A)$ ; to calculate this measure, take each  $p_i \in \Delta(\Pi(A))$  and find its marginal  $p_{ij} \in \Delta(A_j)$ ; the product measure  $p_{i1} \times \dots \times p_{in}$  in  $\Delta(A_i^n)$  indicates the probability with which agent  $j$  will receive recommendations in each Borel subset of  $A_j^n$ ; then,  $(p_{i1} \times \dots \times p_{in}) \circ \delta_j^{-1}$  is a probability measure on  $\Delta(A_j)$  that indicates the probability that agent  $j$  will take actions in each Borel subset of  $A_j$  given that agents play  $(p, \delta)$ . Hence

$$\xi_A(p, \delta) = \prod_{j=1}^n (p_{ij} \times \dots \times p_{in}) \circ \delta_j^{-1}$$

is the measure induced on  $\Delta(\Pi(A))$  by  $(p, \delta)$ . It is easy to show that  $\xi: \Pi R \rightarrow \Delta\Pi$  is a natural transformation.

The functor  $F: G \rightarrow G$  assigns to each game  $g = (A, u)$  the game  $F(g) = (R(A), v)$ , where



To each morphism  $\lambda: g \rightarrow g'$  in  $G$ ,  $F$  assigns  $F(\lambda) = R(\lambda)$ . To verify that  $R(\lambda)$  is a morphism in  $G$ , we have to show that  $v \circ \Pi R(\lambda)^* \leq v'$ , given that  $u \circ \Pi(\lambda)^* \leq u'$ . To see this, consider the diagram

$$\begin{array}{ccccc}
 A & & \text{TR}(A) & \xrightarrow{\xi_A} & \Delta\text{TR}(A) & \xrightarrow{\tau_{\text{TR}(A)}} & \bar{R}^n \\
 \downarrow \lambda & & \downarrow \text{TR}(\lambda) & & \downarrow \Delta\text{TR}(\lambda) & \geq & \\
 A' & & \text{TR}(A') & \xrightarrow{\xi_{A'}} & \Delta\text{TR}(A') & \xrightarrow{\tau_{\text{TR}(A')}} & \bar{R}^n
 \end{array}
 \tag{1}$$

The square commutes by the naturality of  $\xi$ ; to show that  $\tau_{\text{TR}(A')}(u') \circ \Delta\text{TR}(\lambda) \geq \tau_{\text{TR}(A)}(u)$ , notice that, by the naturality of  $\tau$ , the following square commutes

$$\begin{array}{ccc}
 A & [\text{TR}(A) \rightarrow \bar{R}^n] & \xrightarrow{\tau_{\text{TR}(A)}} & [\Delta\text{TR}(A) \rightarrow \bar{R}^n] \\
 \downarrow \lambda & \uparrow \circ \text{TR}(\lambda) & & \uparrow \circ \Delta\text{TR}(\lambda) \\
 A' & [\text{TR}(A') \rightarrow \bar{R}^n] & \xrightarrow{\tau_{\text{TR}(A')}} & [\Delta\text{TR}(A') \rightarrow \bar{R}^n]
 \end{array}$$

i.e., for  $u': \text{TR}(A') \rightarrow \bar{R}^n$ ,

$$\tau_{\text{TR}(A)}(u' \circ \text{TR}(\lambda)) = \tau_{\text{TR}(A)}(u') \circ \Delta\text{TR}(\lambda)
 \tag{2}$$

The inequalities  $u \circ \text{TR}(\lambda^*) \leq u'$ ,  $\lambda^* \circ \lambda = \text{id}$  imply  $u \leq u' \circ \text{TR}(\lambda)$ ; by the monotonicity of  $\tau_{\text{TR}(A)}$

$$\tau_{\text{TR}(A)}(u) \leq \tau_{\text{TR}(A)}(u' \circ \text{TR}(\lambda)).
 \tag{3}$$

By (2) and (3),  $\tau_{\text{TR}(A)}(u) \leq \tau_{\text{TR}(A')}(u') \circ \Delta\text{TR}(\lambda)$ . To show  $v' \geq v \circ \text{TR}(\lambda^*)$  notice that  $\lambda \circ \lambda^* \leq \text{id}_{A'}$  implies

$$\begin{aligned}
v' &\geq v \circ \Pi R(\lambda) \circ \Pi R(\lambda^*) = && \text{(def. of } v') \\
&= \tau_{\Pi A}(u') \circ \xi_A \circ \Pi R(\lambda) \circ \Pi R(\lambda^*) = && \text{(by (1))} \\
&= \tau_{\Pi A}(u') \circ \Delta \Pi(\lambda) \circ \xi_A \circ \Pi R(\lambda^*) \geq && \text{(by (1))} \\
&\geq \tau_{\Pi A}(u) \circ \xi_A \circ \Pi R(\lambda^*) = && \text{(def. of } v) \\
&= v \circ \Pi R(\lambda^*).
\end{aligned}$$

Next, we show that  $F$  is a continuous functor (this is the basic step in the construction of a universal game). Notice that  $R$  is continuous, since its constituent functors  $\Delta$ ,  $\Pi$ ,  $\times$  and  $\rightarrow$  are all continuous; see the appendix for  $\Delta$  and Chapter 13 of Manes and Arbib for the rest. Let  $T: \omega \rightarrow G$  be a functor with colimit  $(g, \lambda)$ ; let  $T_n = g_n = (A_n, u_n)$ ,  $g = (A, u)$ ,  $F(g_n) = (R(A_n), v_n)$ ,  $F(g) = (R(A), v)$  and  $T(n \rightarrow n+1) = f_n: A_n \rightarrow A_{n+1}$ . It has to be shown that

- $(R(A), R(\lambda))$  is a colimit of the functor

$$R(A_1) \xrightarrow{R(f_1)} R(A_2) \xrightarrow{R(f_2)} R(A_3) \rightarrow \dots, \text{ and}$$

$$v = \bar{\bigvee}_{n=1} F_1 F(\lambda_n)(v_n) = \bar{\bigvee}_{n=1} F_1 R(\lambda_n)(v_n).$$

The first part is simply the definition of continuity of  $R$ . For the second part, first simplify notation by setting  $\xi_n = \xi_{A_n}$ ,  $\tau_n = \tau_{\Pi A_n}$ ,  $\tau_{\Pi A} = \tau_A$ , and then recall that

$$F_1(f_n)(u_n) \leq u_{n+1}, \quad u = \bar{\bigvee}_{n=1} F(\lambda_n)(u_n),$$

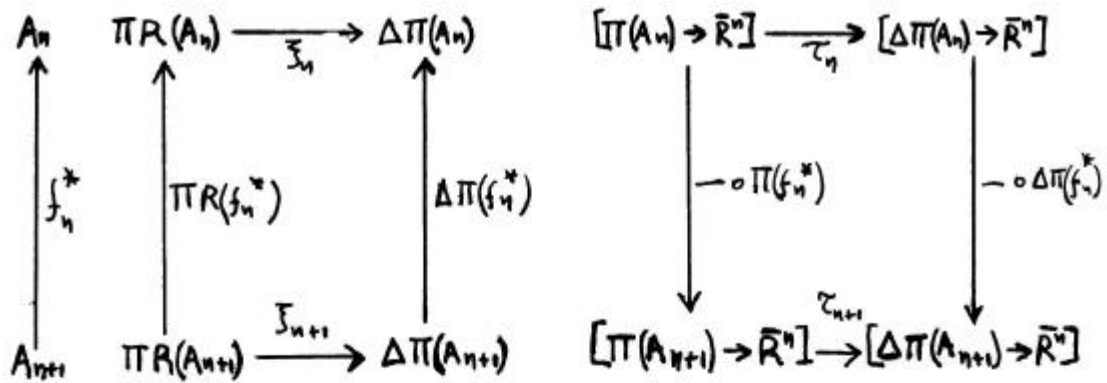
$$v_n = \tau_n(u_n) \circ \xi_n, \quad v = \tau_A(u) \circ \xi_A,$$

$$\lambda_n \circ \lambda_n^* \leq \lambda_{n+1} \circ \lambda_{n+1}^*, \quad \bar{\bigvee}_{n=1} \lambda_n \circ \lambda_n^* = id_A,$$

$$f_n = \lambda_{n+1}^* \circ \lambda_n, \quad \lambda_n \circ \lambda_n^* \leq id_{A_n}, \quad \lambda_n^* \circ \lambda_n = id_{A_n}$$

Step 1:  $\bar{\bigvee}_{n=1} F_1 F(\lambda_n)(v_n)$  exists.

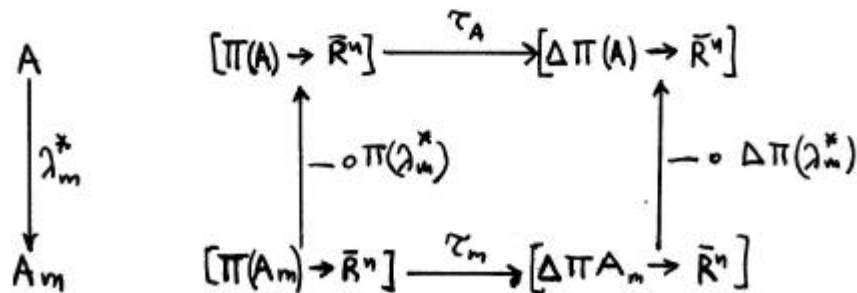
It suffices to show that, for each  $n$ ,  $F_1 F(\lambda_n)(v_n) \leq F_1 F(\lambda_{n+1})(v_{n+1})$ . The proof will rest on the naturality of  $\lambda$ ,  $\xi$ , and  $\tau$  and the naturality arguments that follow can be read off the following diagrams.

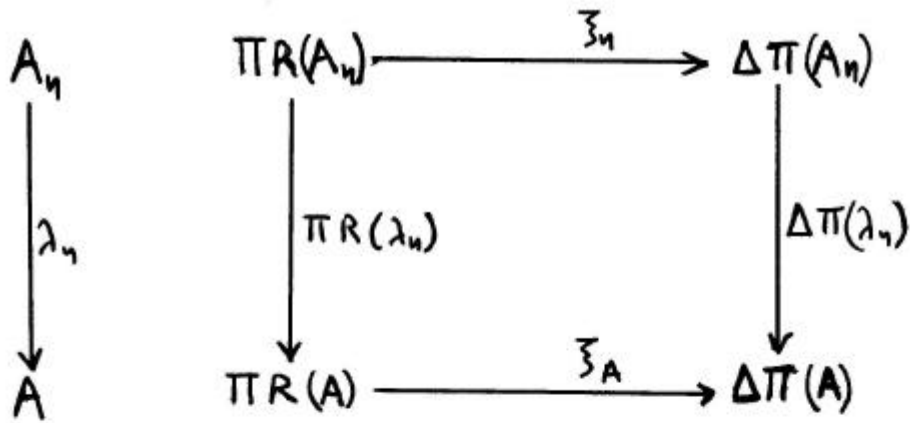


$$\begin{aligned}
 F_1 F(\lambda_n)(v_n) &= && \text{(def of } F_1, F) \\
 = v_n \circ \Pi R(\lambda_n) &= && \text{(def. of } v_n) \\
 = \tau_n(u_n) \circ \xi_n \circ \Pi R(\lambda_n) &= && \text{(naturality of } \lambda) \\
 = \tau_n(u_n) \circ \xi_n \circ \Pi R(f_n^*) \circ \Pi R(\lambda_{n+1}^*) &= && \text{(naturality of } \xi) \\
 = \tau_n(u_n) \circ \Delta \Pi(f_n^*) \circ \xi_{n+1} \circ \Pi R(\lambda_{n+1}^*) &= && \text{(naturality of } \tau) \\
 = \tau_{n+1}(u_n \circ \Pi(f_n^*)) \circ \xi_{n+1} \circ \Pi R(\lambda_{n+1}^*) &\leq && \text{(property of } u_n) \\
 \\
 \leq \tau_{n+1}(u_{n+1}) \circ \xi_{n+1} \circ \Pi R(\lambda_{n+1}^*) &= && \text{(def. of } v_{n+1}) \\
 = v_{n+1} \circ \Pi R(\lambda_{n+1}^*) &= && \text{(def. of } F_1, F) \\
 = F_1 F(\lambda_{n+1})(v_{n+1}) &= && \text{QED.}
 \end{aligned}$$

Step 2:  $v = \prod_{n=1}^{\infty} \prod_{m=1}^{\infty} (\tau_n(u_m) \circ \Delta \Pi(\lambda_m^*) \circ \Delta \Pi(\lambda_n) \circ \xi_n \circ \Pi R(\lambda_n^*))$

The proof will rest on the naturality of  $\tau$  and  $\xi$  as expressed in the following diagrams





( $\Pi R(\lambda_n)$  is a colimit natural transf.)

$$\begin{aligned}
 v &= v \circ id = && \text{(continuity of } v) \\
 &= v \circ \bigvee_{n=1}^{\infty} (\Pi R(\lambda_n) \circ \Pi R(\lambda_n^*)) = && \text{(def. of } v) \\
 &= \bigvee_{n=1}^{\infty} (v \circ \Pi R(\lambda_n) \circ \Pi R(\lambda_n^*)) = && \text{(Property of } u) \\
 &= \bigvee_{n=1}^{\infty} (\tau_A(u) \circ \xi_A \circ \Pi R(\lambda_n) \circ \Pi R(\lambda_n^*)) = && \text{(naturality of } \tau) \\
 &= V_n V_m (\tau_A(u_m \circ \Pi(\lambda_m^*)) \circ \xi_A \circ \Pi R(\lambda_n) \circ \Pi R(\lambda_n^*)) = \\
 &= V_n V_m (\tau_n(u_m) \circ \Pi \Delta(\lambda_m^*) \circ \xi_A \circ \Pi R(\lambda_n) \circ \Pi R(\lambda_n^*)) = && \text{(naturality of } \xi) \\
 &= V_n V_m (\tau_n(u_m) \circ \Pi \Delta(\lambda_m^*) \circ \Pi \Delta(\lambda_n) \circ \xi_n \circ \Pi R(\lambda_n^*)).
 \end{aligned}$$

Step 3: If  $n \geq m$ ,  $f_{mn} = f_{n-1} \circ f_{n-2} \circ \dots \circ f_m$ , then  $\lambda_m^* \circ \lambda_n \leq f_{mn}$ .

Recall that, by the naturality of  $\lambda$ ,  $f_n = \lambda_{n+1}^* \circ \lambda_n$  for each  $n$ . Then,

$$\begin{aligned}
f_{m0} \circ \lambda_m^* \circ \lambda_0 &= f_{0-1} \circ f_{0-2} \circ \dots \circ f_{m-1} \circ \lambda_m^* \circ \lambda_0 = \\
&= \lambda_m^* \lambda_{m-1} \lambda_{m-2}^* \lambda_{m-3} \dots \lambda_{1-1} \lambda_{1-2}^* \lambda_{1-3} \lambda_{1-4} \dots \lambda_{0-1} \lambda_{0-2}^* \lambda_{0-3} \lambda_{0-4} \dots \\
&\leq \lambda_m^* id \dots id \lambda_0 = \\
&= \lambda_m^* \lambda_0 \leq id.
\end{aligned}$$

Hence,  $f_{m0} \circ \lambda_m^* \circ \lambda_0 \leq id$ ; composing both sides with the monotonic function  $f_{m0}$ ,  $f_{m0} \circ f_{m0} \circ \lambda_m^* \lambda_0 \leq f_{m0}$ .

and therefore,  $\lambda_m^* \lambda_0 \leq f_{m0}$ .

Step 4: If  $B \subset A \subset P$ , where  $P$  is a poset, and if  $h = \bigvee A$ , and if for each  $a \in A$  there is  $b \in B$  such that  $a \leq b$ , then  $h = \bigvee B$ . ( $\bigvee$  is the operation of taking the supremum of a set)

Step 5: Set  $P = [\text{IRR}(A) \rightarrow \bar{R}^n]$

$$\begin{aligned}
A &= \{ \tau_n(u_m) \circ \Pi \Delta(\lambda_m^*) \circ \Pi \Delta(\lambda_0) \circ \xi_n \circ \text{IRR}(\lambda_0^*); m, n \in \omega \} \\
B &= \{ \tau_n(u_n) \circ \xi_n \circ \text{IRR}(\lambda_0^*); n \in \omega \} = \\
&= \{ F_1 F \lambda_n(v_n); n \in \omega \}.
\end{aligned}$$

Then  $B \subset A \subset P$  and for each  $a \in A$  there is  $b \in B$  such that  $a \leq b$ .

Proof: Let  $a = \tau_n(u_m) \circ \Pi \Delta(\lambda_m^*) \circ \Pi \Delta(\lambda_0) \circ \xi_n \circ \text{IRR}(\lambda_0^*)$ . If  $n = m$ ,  $a = \tau_n(u_n) \circ \xi_n \circ \text{IRR}(\lambda_0^*) \in B$  and there is nothing to prove. If  $n > m$ , let  $b = \tau_n(u_n) \circ \xi_n \circ \text{IRR}(\lambda_0^*)$ . Then

$$\begin{aligned}
a &= \\
\tau_n(u_m) \circ \Pi \Delta(\lambda_m^*) \circ \Pi \Delta(\lambda_0) \circ \xi_n \circ \text{IRR}(\lambda_0^*) &= \text{(naturality of } \tau) \\
&= \tau_n(u_m) \circ \Pi \Delta(\lambda_m^*) \circ \xi_n \circ \text{IRR}(\lambda_0^*) && (11) \\
&\leq \tau_n(u_m) \circ \Pi \Delta(\lambda_m^*) \circ \xi_n \circ \text{IRR}(\lambda_0^*) && (\text{step 3}) \\
&\leq \tau_n(u_m) \circ \xi_n \circ \text{IRR}(\lambda_0^*) && (\text{property of } u_m, u_n) \\
&\leq \tau_n(u_n) \circ \xi_n \circ \text{IRR}(\lambda_0^*) = b.
\end{aligned}$$

The naturality arguments in the first two equalities can be read of the following commuting diagram

$$\begin{array}{ccc}
A_m & [\Pi A_m \rightarrow \bar{R}^n] & \xrightarrow{\tau_m} & [\Delta \Pi A_m \rightarrow \bar{R}^n] \\
\uparrow \lambda_m^* & \downarrow - \circ \Pi(\lambda_m^*) & & \downarrow - \circ \Delta \Pi(\lambda_m^*) \\
A & [\Pi A \rightarrow \bar{R}^n] & \xrightarrow{\tau} & [\Delta \Pi A \rightarrow \bar{R}^n]
\end{array}$$

$$\begin{array}{ccc}
 A & [\Pi A \rightarrow \bar{R}^n] & \xrightarrow{\tau} & [\Delta \Pi A \rightarrow \bar{R}^n] \\
 \uparrow \lambda_n & \downarrow - \circ \Pi(\lambda_n) & & \downarrow - \circ \Delta \Pi(\lambda_n) \\
 A_n & [\Pi A_n \rightarrow \bar{R}^n] & \xrightarrow{\tau_n} & [\Delta \Pi A \rightarrow \bar{R}^n]
 \end{array}$$

If  $n < m$ , set  $b = \tau_m(u_m) \circ \xi_m \circ \Pi R(\lambda_m^*)$ . Then,

$$a = \quad \text{(naturality of } \xi)$$

$$\tau_m(u_m) \Pi \Delta(\lambda_m) \Delta \Pi(\lambda_n) \Delta \Pi(\lambda_n^*) \xi_A \leq$$

$$\leq \tau_m(u_m) \Pi \Delta(\lambda_m^*) \xi_A = \quad \text{(naturality of } \xi)$$

$$= \tau_m(u_m) \xi_m \Pi R(\lambda_m^*) = b$$

The naturality arguments are read off the following diagram

$$\begin{array}{ccccc}
 A_n & & \Pi R A_n & \xrightarrow{\xi_n} & \Delta \Pi A_n \\
 \uparrow \lambda_n^* & & \uparrow \Pi R(\lambda_n^*) & & \uparrow \Delta \Pi(\lambda_n^*) \\
 A & & \Pi R A & \xrightarrow{\xi_A} & \Delta \Pi A \\
 \downarrow \lambda_m^* & & \downarrow \Pi R(\lambda_m^*) & & \downarrow \Delta \Pi(\lambda_m^*) \\
 A_m & & \Pi R A_m & \xrightarrow{\xi_m} & \Delta \Pi A_m
 \end{array}$$

Step 6:  $v = \bar{V} F_1 F(\lambda_n v_n)$

$v = VA$  by step 2, and

$$\bar{V} F_1 F(\lambda_n)(v_n) = VB.$$

Hence, by steps 4 and 5,  $VA = VB$ .

We have shown the following

*Theorem 1:*  $F: G \rightarrow G$  is continuous.

Recall that our objective is to construct, for each game  $g$ , a universal game  $X(g)$  that includes  $g$ , i.e., a colimit of the functor

$$g \xrightarrow{\phi_g} F(g) \xrightarrow{F(\phi_g)} F^2(g) \xrightarrow{F^2(\phi_g)} \dots$$

The continuity of  $F$  will guarantee that this colimit is a fixed point of  $F$ , i.e., the universal game  $X(g)$ ; we still have to determine the nature of the morphisms  $\phi_g: g \rightarrow F(g)$ . If  $g = (A, u)$ , then  $F(g) = (R(A), v)$ ; define  $\phi_g^i: A_i \rightarrow R_i(A)$  by  $\phi_g^i(a_i) = (\perp, \hat{a}_i)$  where  $\perp$  is the bottom element of  $\Delta\Pi(A)$  and  $\hat{a}_i: A_i^0 \rightarrow A_i$  is the function "constant at  $a_i$ ." It is easy to show that  $\phi_g^i$  is in  $D_{\mathbb{B}}$ , with adjoint  $\phi_g^{i*}: R_i(A) \rightarrow A_i$  defined by  $\phi_g^{i*}(p_i, \delta_i) = \delta_i(\perp)$  where  $\perp$  is the bottom element of  $A_i^0$ . Let  $\phi_g: A \rightarrow R(A)$  be  $\langle \phi_g^1, \dots, \phi_g^n \rangle$ ; to show that  $\phi_g: g \rightarrow F(g)$  we need to have  $u \circ \Pi(\phi_g^*) \leq v$ . Given that  $\phi_g \circ \phi_g^* \leq \text{id}$ , it suffices to show that  $u \leq v \circ \Pi(\phi_g)$ . For each  $a \in \Pi(A)$

$$\begin{aligned} v(\Pi(\phi_g)(a)) &= \\ &= v(\phi_g^1(a_1), \dots, \phi_g^n(a_n)) = \\ &= v((\perp, \hat{a}_1), \dots, (\perp, \hat{a}_n)) = \\ &= \tau_{\Pi A}(u) \xi_A((\perp, \hat{a}_1), \dots, (\perp, \hat{a}_n)) = \\ &= \tau_{\Pi A}(u) \circ e_{\Pi A}^*(a) = u(a). \end{aligned}$$

(Recall that  $e_{\Pi A}^*(a)$  is the Dirac measure concentrated on  $a$ , and that  $u$  and  $\tau_{\Pi A}(u)$  agree on Dirac measures). Hence  $\phi_g: g \rightarrow F(g)$  is a morphism in  $G$ . In fact  $\phi: I \rightarrow F$  is a natural transformation, since for each  $\lambda: g \rightarrow g'$  the following diagram commutes

$$\begin{array}{ccc} A & \xrightarrow{\phi_g} & R(A) \\ \downarrow \lambda & & \downarrow R(\lambda) \\ A' & \xrightarrow{\phi_{g'}} & R(A') \end{array}$$

For each  $a \in A$ ,

$$R(\lambda)\phi_g(a) = R(\lambda)(\perp, \bar{a}) = (\Delta\Pi(\lambda)(\perp), \lambda \circ \delta \circ \lambda^{**}) = (\perp, \lambda(a)) = \phi_g(\lambda(a)).$$

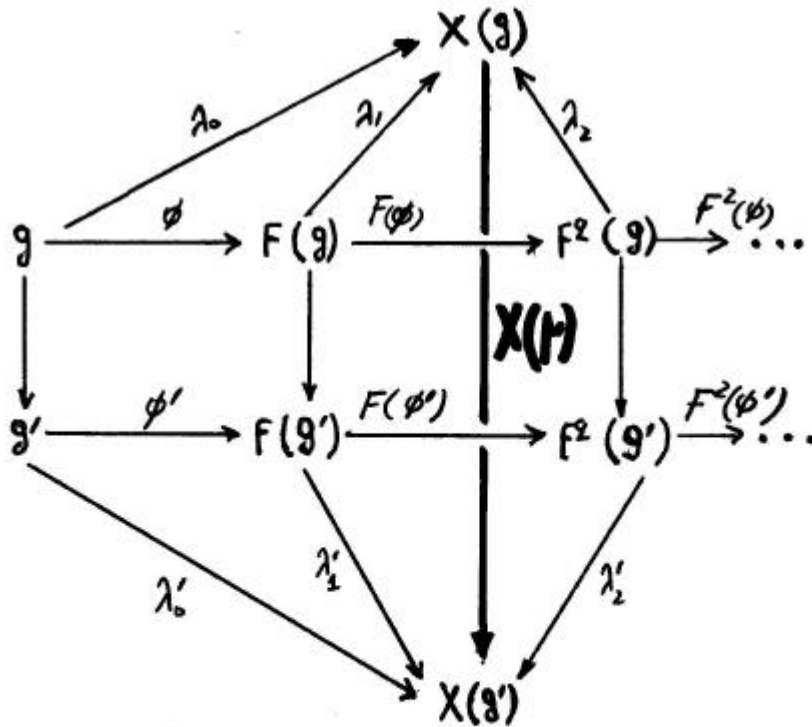
*Theorem 2:* For each  $g \in G$ , there is a game  $X(g)$  in  $g$  that is the colimit object of the functor

$$g \xrightarrow{\phi} F(g) \xrightarrow{F(\phi)} F^2(g) \rightarrow \dots$$

and satisfies  $X(g) = F(X(g))$ .  $X(g)$  is the universal game generated by  $g$ ; it is unique up to isomorphism.

*Proof:*  $F$  is continuous,  $G$  has all  $\omega$ -colimits and for each  $g$  there is a morphism  $\phi_g: g \rightarrow F(g)$ . Then, the theorem follows by (a simple corollary of) theorem 28 in the companion paper, or by the construction in p. 766-767 of Smyth and Plotkin (1982).

*Remark.* The operation  $g \rightarrow X(g)$  can be extended (uniquely) to morphisms  $\mu: g \rightarrow g'$ ; let  $\lambda, \lambda'$  be the colimiting natural transformations corresponding to  $g, g'$  respectively, and let  $\phi = \phi_g, \phi' = \phi_{g'}$ . Then the following diagram commutes



and so there is a unique morphism, call it

$X(\mu): X(g) \rightarrow X(g')$  such that  $X(\mu) \circ \lambda_t = \lambda'_t \circ F^t(\mu)$  for all  $t \geq 0$ .

In fact, by theorems 21 (and 29) in Vassilakis (1990), due to Lehmann and Smyth (1981),  $X$  is a continuous functor.

### 5. Coordination - Proof Solution Concepts

For each game  $g$ , there is a D-morphism  $\gamma(g): \Delta \Pi P_1 X(g) \rightarrow \Delta \Pi P_1(g)$  that projects probability measures on the aggregate strategy space of the universal game  $X(g)$  generated by  $g$ , to probability measures on the aggregate strategy space of  $g$  itself;  $\gamma(g)$  will be defined by exploiting the colimit property of  $X(g)$ .

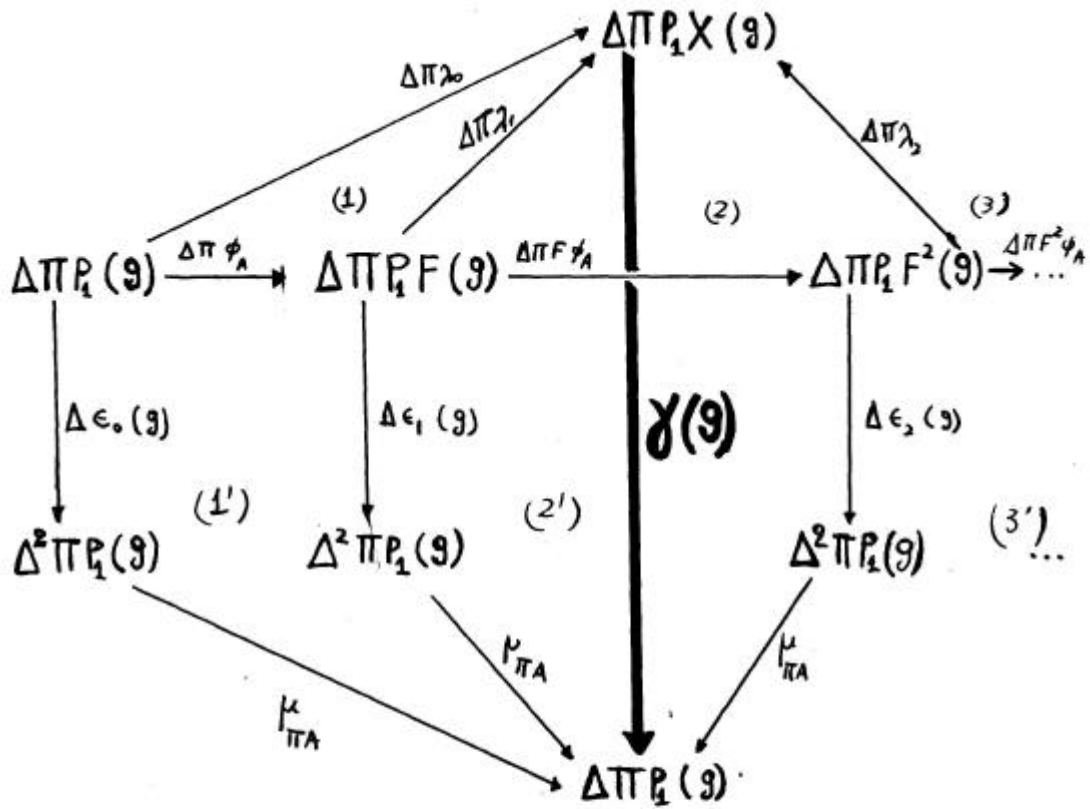
First, define morphisms,  $e_t(g): \Pi P_1 F^t(g) \rightarrow \Delta \Pi P_1(g)$  as follows:  $e_t(g)$  is the morphism assigning to each element in  $\Pi P_1(g)$  the probability measure concentrated on this element;  $e_1(g) = \xi_{P_1(g)}$  (recall that  $P_1 F = R$ ); for  $t \geq 2$ ,  $e_t(g)$  is defined by the following diagram.

$$\begin{array}{ccc}
 \Pi P_1 F^t(g) & \xrightarrow{\sum_{P_1 F^{t-1}(g)}} & \Delta \Pi P_1 F^{t-1}(g) \\
 \downarrow \varepsilon_t(g) & & \downarrow \Delta(\varepsilon_{t-1}(g)) \\
 \Delta \Pi P_1(g) & \xrightarrow{\mu_{\Pi P_1(g)}} & \Delta^2 \Pi P_1(g)
 \end{array}$$

In words,  $e_t(g)$  assigns to each aggregate strategy vector in the  $t$ -level game generated by  $g$  the probability measure on the aggregate strategy space of  $g$  induced by this strategy vector.

Let  $g = (A, u)$  ( $P_1 g = A$ ) and let  $\phi_g$  denote the morphism  $\phi_g: g \rightarrow F(g)$ ; recall that  $\phi_A: A \rightarrow R(A) = P_1 F(g)$ . Let  $\lambda_t: F^t(g) \rightarrow X(g)$ ,  $t \geq 0$ , be the colimit natural transformation associated with the

definition of  $X(g)$  as a colimit. Consider the following diagram



The top panels (1), (2), (3)... commute because  $\lambda$  is natural, i.e. because  $\lambda_{t+1} \circ F(\phi_A) = \lambda_t$ ,  $t \geq 0$ . The bottom panels (1') (2') (3')... commute by the definitions of  $\epsilon$  and  $\mu$ . The continuity of the functor  $\Delta\Pi P_1$  implies that  $\Delta\Pi P_1 \lambda_t$ ,  $t \geq 0$ , is also a colimit natural transformation. Then, the commutativity of the panels (1') (2')... and the definition of a colimit imply that there is a unique  $\gamma(g)$  in  $D$  such that, for all  $t \geq 0$ ,

$$\gamma(g) \circ \Delta \Pi(\lambda) = \mu_{\Pi \lambda} \circ \Delta(\epsilon_t(g)).$$

The next step is to show that, for each  $t \geq 0$

$$\begin{array}{ccc} & \xrightarrow{\pi_{P_t} F^t} & \\ G & \xrightarrow{\quad \downarrow \epsilon_t \quad} & D \\ & \xrightarrow{\Delta \pi_{P_t}} & \end{array}$$

is a natural transformation. It is clear that  $e_t$  (the Dirac measures) and  $e_t$  (the  $\xi$ -morphisms) are indeed natural. For the induction step, suppose  $e_t$  is natural; let  $v: g \rightarrow g'$  in  $G$ . Then we need to show that  $\Delta \pi_{P_t}(v) \circ e_{t+1}(g) = e_{t+1}(g') \circ \pi_{P_t} F^t(v)$ .

To see this, let

$$\xi_t = \xi_{\pi_{P_t} F^t(g)}, \quad \xi'_t = \xi_{\pi_{P_t} F^t(g')}, \quad \mu = \mu_{\pi_{P_t} \mu'} = \mu_{\pi_{P_t} g'}.$$

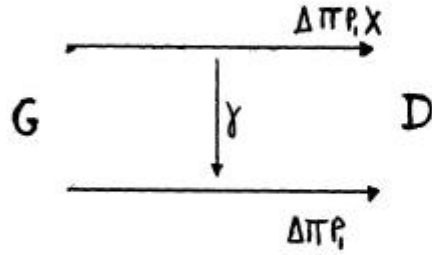
Then the following diagram commutes, by the naturality of  $\xi$ ,  $\mu$ , and  $e_t$  and the definition of  $e_{t+1}$

$$\begin{array}{ccc} \pi_{P_t} F^t(g) & \xrightarrow{\pi_{P_t} F^t(v)} & \pi_{P_t} F^t(g') \\ \downarrow \epsilon_{t+1}(g) & \begin{array}{c} (?) \\ \Delta \pi_{P_t}(v) \end{array} & \downarrow \epsilon_{t+1}(g') \\ \Delta \pi_{P_t}(g) & \xrightarrow{\Delta \pi_{P_t}(v)} & \Delta \pi_{P_t}(g') \\ \downarrow \xi_t & \Delta \pi_{P_t} F^t(v) & \downarrow \xi'_t \\ \Delta \pi_{P_t} F^t(g) & \xrightarrow{\Delta \pi_{P_t} F^t(v)} & \Delta \pi_{P_t} F^t(g') \\ \downarrow \Delta \epsilon_t(g) & & \downarrow \Delta \epsilon_t(g') \\ \Delta^2 \pi_{P_t}(g) & \xrightarrow{\Delta^2 \pi_{P_t}(v)} & \Delta^2 \pi_{P_t}(g') \\ \mu \nearrow & & \mu' \nearrow \\ g & \xrightarrow{v} & g' \end{array}$$

Then

$$\begin{aligned}
 \Delta \Pi_{P_1}(v) \circ \epsilon_{*,1}(g) &= && \text{(definition of } \epsilon_{*,1} \text{)} \\
 \Delta \Pi_{P_1}(v) \circ \mu \circ \Delta \epsilon_i(g) \circ \xi_i &= && \text{(naturality of } \mu \text{)} \\
 = \mu' \circ \Delta^2 \Pi_{P_1}(v) \circ \Delta \epsilon_i(g) \circ \xi_i &= && \text{(naturality of } \epsilon_i \text{)} \\
 = \mu' \circ \Delta \epsilon_i(g') \circ \Delta \Pi_{P_1 F}(v) \circ \xi_i &= && \text{(naturality of } \xi_i \text{)} \\
 = \mu' \circ \Delta \epsilon_i(g') \circ \xi_i' \circ \Pi_{P_1 F}(v) &= && \text{(definition of } \epsilon_{*,1} \text{)} \\
 = \epsilon_{*,1}(g') \circ \Pi_{P_1 F}(v).
 \end{aligned}$$

Next, we show that



is a natural transformation, i.e. that for each  $v: g \rightarrow g'$  in  $G$ , the following holds:

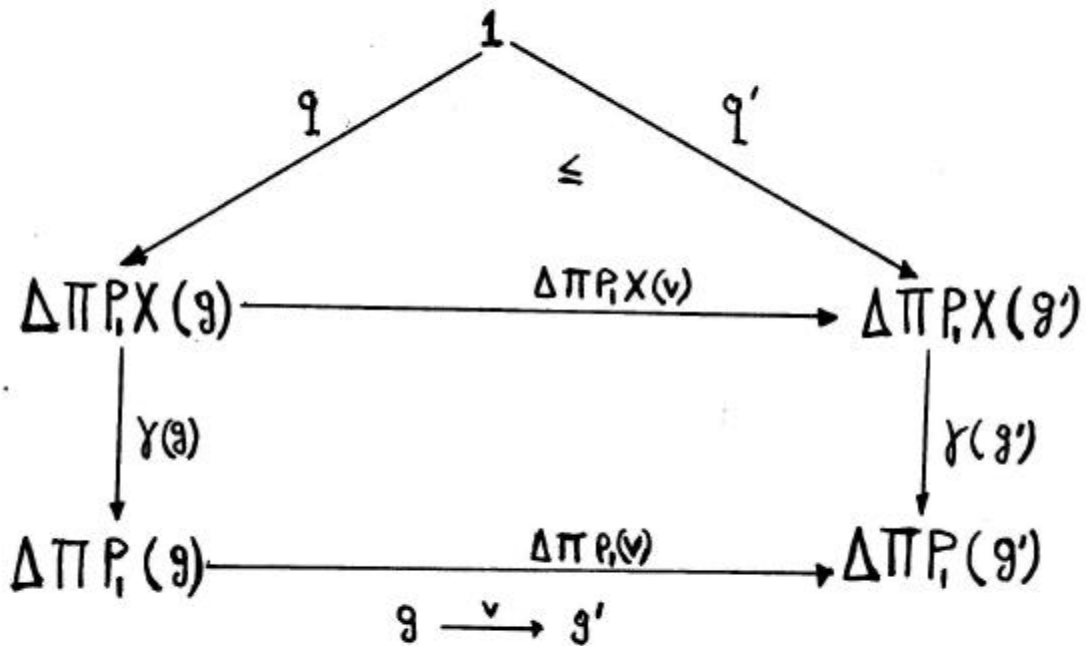


$$\begin{aligned}
&g(\gamma') \circ \Delta \Pi_{P_1, X}(v) \circ \Delta \Pi(\lambda_t) = (\text{definition of } X(v)) \\
&= g(\gamma') \circ \Delta \Pi(\lambda) \circ \Delta \Pi_{P_1, F^c}(v) = (\text{definition of } \gamma(g')) \\
&= \mu' \circ \Delta e_t(g') \circ \Delta \Pi_{P_1, F^c}(v) = (\text{naturality of } e_t) \\
&= \mu' \circ \Delta^2 \Pi_{P_1}(v) \circ \Delta e_1(g) = (\text{naturality of } \mu) \\
&= \Delta \Pi_{P_1}(v) \circ \mu \circ \Delta e_2(g) = (\text{definition of } \gamma) \\
&= \Delta \Pi_{P_1}(v) \circ \gamma(g) \circ \Delta \Pi(\lambda_t), \text{ for all } t \geq 0.
\end{aligned}$$

Then, by the joint surjectivity of colimit natural transformations (proven in appendix 2),  $\gamma(g') \circ \Delta \Pi_{P_1, X}(v) = \Delta \Pi_{P_1}(v) \circ \gamma(g)$ .

Recall that if  $\sigma(g) = (g, p)$ , we write  $P_1 \sigma(g) = g, P_2 \sigma(g) = p$ . We now define the functor  $\Psi: \Sigma \rightarrow [G \rightarrow \text{Fib}(\Delta \Pi_{P_1})]$  as follows: for each  $\sigma \in \Sigma$ ,  $\Psi(\sigma)$  is a functor; on objects it is defined by  $\Psi(\sigma)(g) = (g, \gamma(g)(P_2 \sigma(X(g))))$ ,  $g \in G$ .

On morphisms  $v: g \rightarrow g'$ , we define  $\Psi(\sigma)(v): \Psi(\sigma)(g) \rightarrow \Psi(\sigma)(g')$  by  $\Psi(\sigma)(v) = (v, \alpha)$ ,  $\alpha: \Delta \Pi_{P_1}(v)(\gamma(g)(q)) \rightarrow \gamma(g')(q')$ , where  $P_2 \sigma(X(g)) = q, P_2 \sigma(X(g')) = q'$ . We now show that such an  $\alpha$  exists, i.e., that  $\Delta \Pi_{P_1}(v)(\gamma(g)(q)) \leq \gamma(g')(q')$ . To see this, consult the following diagram



The square commutes by the naturality of  $\gamma$ , while the inequality of the triangle,  $\Delta \Pi_{P_1, X}(v)(q) \leq q'$ ,

follows from the fact that  $\sigma: G \rightarrow \text{Fib}(\Delta\Pi P_1)$  is a section of  $\Delta\Pi P_1$ : for then  $\sigma(X(v)) = (X(v), \beta): (X(v),$

$q) \rightarrow ((g'), q')$  in  $\text{Fib}(\Delta\Pi P_1)$ , and by the definition of morphisms in categories of fibrations,

$\beta: \Delta\Pi P_1 X(v)(q) \rightarrow q'$ , i.e.

$$\begin{aligned} \Delta\Pi P_1 X(v)(q) &\leq q'. \text{ Then,} \\ \Delta\Pi P_1(v)\gamma(g)(q) &= (\text{square}) \\ &= \gamma(q')\Delta\Pi P_1 X(v)(q) \leq (\text{triangle}) \\ &\leq \gamma(g')(q'). \end{aligned}$$

To complete the definition of  $\Psi$ , let  $\Psi$  assign to each morphism  $\beta: \sigma \rightarrow \sigma'$  in  $\Sigma$  the natural transformation  $\Psi(\beta): \Psi(\sigma) \rightarrow \Psi(\sigma')$  defined by  $\Psi(\beta) = \beta \circ X$ , i.e.

$$\Psi(\beta)(g) = \beta(X(g)): \Psi(\sigma)(g) \rightarrow \Psi(\sigma')(g).$$

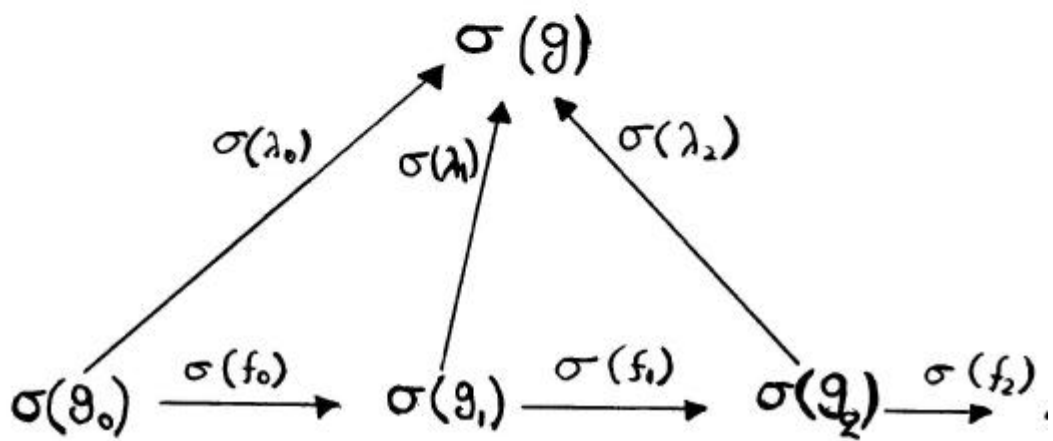
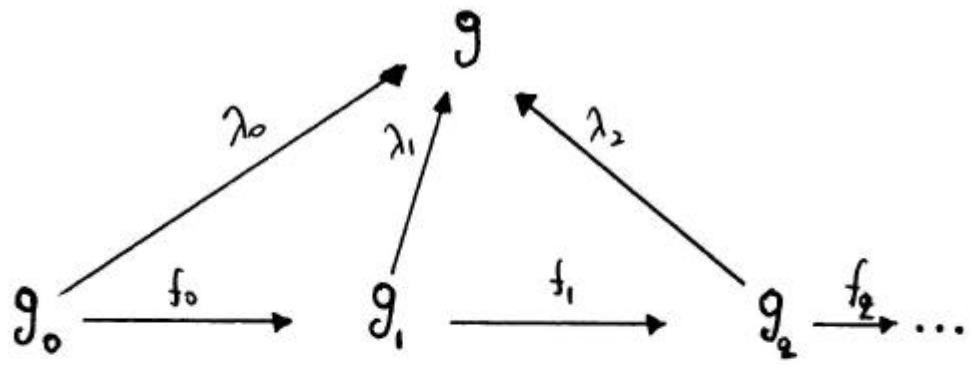
It is easy to verify that  $\Psi$  is a functor, i.e. it preserves composition and identity.

Recall that we are looking for a solution concept  $\sigma$  that satisfies  $\sigma \simeq \Psi(\sigma)$ . To do this, we need to show that  $\Psi$  maps  $\Sigma$  into  $\Sigma$ , i.e. that if  $\sigma \in \Sigma$  then  $\Psi(\sigma)$  is continuous and satisfies the weak revelation principle; that  $\Sigma$  has all  $\omega$ -colimits and that  $\Psi$  preserves them; and that, for some  $\sigma \in \Sigma$ , there is a morphism  $\beta: \sigma \rightarrow \Psi(\sigma)$  in  $\Sigma$ . Then we can apply the fixed point theorem to obtain coordination-proof solution concepts.

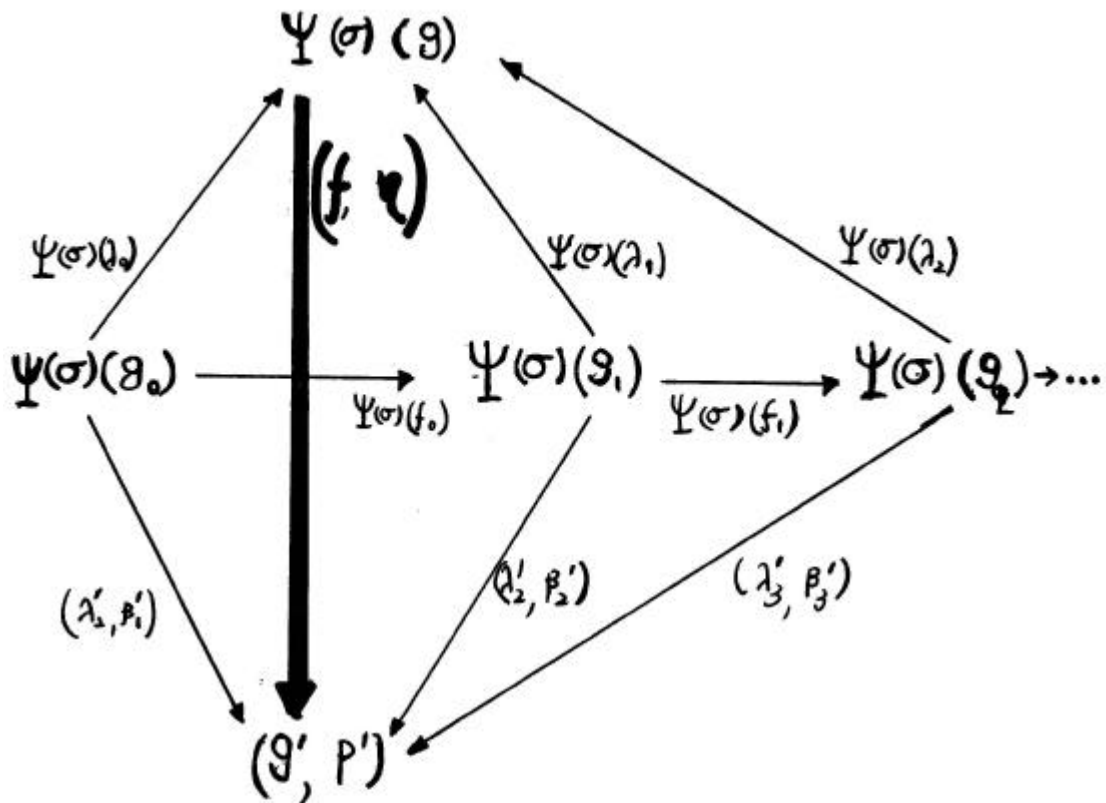
*Lemma 1:* For each  $\sigma \in \Sigma$ ,  $\Psi(\sigma): G \rightarrow \text{Fib}(\Delta\Pi P_1)$  is continuous.

*Proof:* Let  $T: \omega \rightarrow G$  be a functor with colimit  $(g, \lambda)$ . We need to show that  $(\Psi(\sigma)(g), \Psi(\sigma)(\lambda))$  is a colimit of  $\Psi(\sigma) \circ T$ . To see what this means, let  $T_n = g_n = (A_n, u_n)$ ,  $g = (A, u)$ ,  $T(n \rightarrow n+1) = f_n: g_n \rightarrow g_{n+1}$ ,  $\lambda_n: g_n \rightarrow g$ .

Then the following are colimit natural transformations ( $\sigma \in \Sigma$  is continuous)



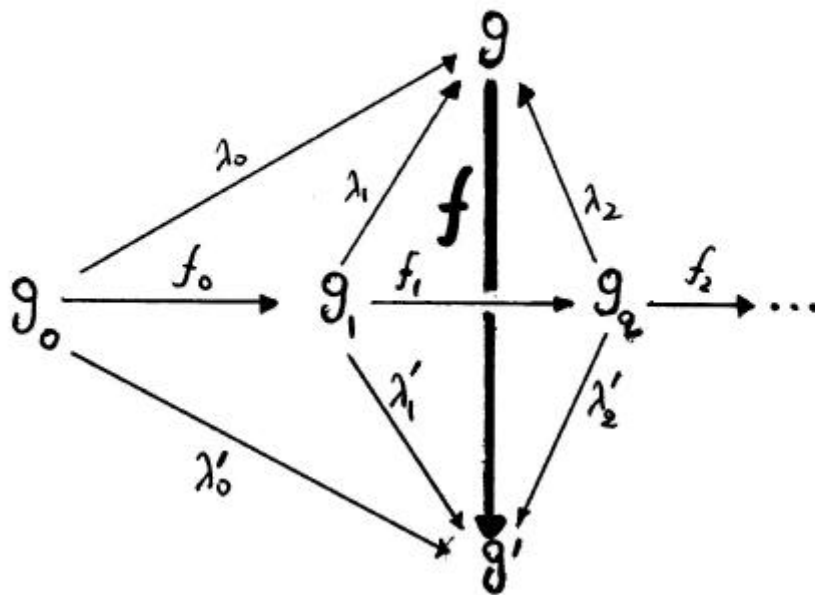
Now consider the following diagram



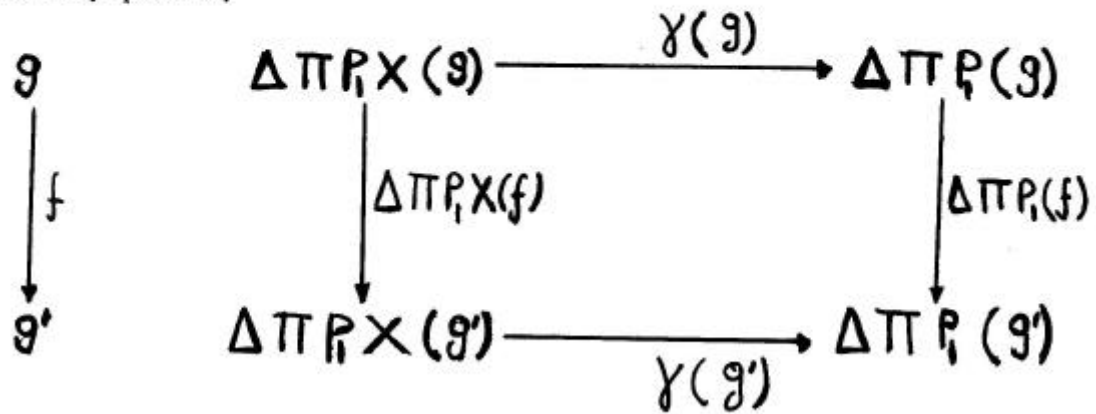
The top triangles commute because  $\lambda$  is natural; the bottom triangles commute because  $(\lambda'_n, \beta'_n)$ :  $\Psi(\sigma)(g_n) \rightarrow (g', P')$ ,  $n \geq 0$ , is taken to be an arbitrary natural transformation in  $\text{Fib}(\Delta\text{IP1})$ . We have to show that there is a unique morphism  $(f, \eta): \Psi(\sigma)(g) \rightarrow (g', p')$  such that for all  $n \geq 0$ ,  $(f, \eta) \circ \Psi(\sigma)(\lambda_n) = (\lambda'_n, \beta'_n)$ . To show this, we need the following notation:  $\sigma(g_n) = (g_n, p_n)$ ,  $\sigma(g) = (g, p)$ ,  $\sigma(X(g_n)) = (X(g_n), q_n)$ ,  $\sigma(X(g)) = (X(g), q)$ ,  $\Psi(\sigma)(g_n) = (g_n, \gamma(g_n)(q_n))$ ,  $\Psi(\sigma)(g) = (g, \gamma(g)(q))$ ,  $\sigma(\lambda_n) = (\lambda_n, \alpha_n)$ ,  $\alpha_n: \Delta\text{IP}_1(\lambda_n) (p_n) \rightarrow p$ ,  $\Psi(\sigma)(\lambda_n) = (\lambda_n, \beta_n)$ ,  $\beta_n: \Delta\text{IP}_1(\lambda_n) (\gamma(g_n)(q_n)) \rightarrow \gamma(g)(q)$ ,  $\Psi(\sigma)(f_n) = (f_n, \theta_n)$ ,  $\theta_n: \Delta\text{IP}_1(f_n) (\gamma(g_n)(q_n)) \rightarrow \gamma(g_{n+1})(q_{n+1})$ . Then,  $(f, \eta) \circ \Psi(\sigma)(\lambda_n) = (f, \eta) \circ (\lambda_n, \beta_n) = (f \circ \lambda_n, \eta \circ \Delta\text{IP}_1(f)(\beta_n))$ .

The existence and uniqueness of  $f$  follows directly from the following diagram, recalling that  $(g, \lambda)$

is a colimit:



To show that  $\gamma$  exists, it suffices to show that  $\Delta \Pi_{P_1}(f)(\gamma(g)(q)) \leq p'$ ; for this, we will need the naturality squares of  $\gamma$



$$\begin{array}{ccccc}
 \mathfrak{g}' & & \Delta \Pi P_1 X(\mathfrak{g}') & \xrightarrow{\gamma(\mathfrak{g}')} & \Delta \Pi P_1(\mathfrak{g}') \\
 \uparrow \lambda'_n & & \uparrow \Delta \Pi P_1 X(\lambda'_n) & & \uparrow \Delta \Pi P_1(\lambda'_n) \\
 \mathfrak{g}_n & & \Delta \Pi P_1 X(\mathfrak{g}_n) & \xrightarrow{\gamma(\mathfrak{g}_n)} & \Delta \Pi P_1(\mathfrak{g}_n)
 \end{array}$$

and the (in) equalities:

$$q = \varinjlim_{n \geq 0} \Delta \Pi P_1 X(\lambda_n)(q_n) \quad (1)$$

$$f = \varinjlim_{n \geq 0} (\lambda'_n \circ \lambda_n) \quad (2)$$

$$\Delta \Pi P_1(\lambda'_n)(\gamma(\mathfrak{g}_n)(q_n)) \leq p', \text{ all } n \geq 0. \quad (3)$$

To prove equation (1), we write  $(X(\mathfrak{g}), q) = \sigma(X(\mathfrak{g})) = \sigma(X(\text{colim } \mathfrak{g}_n)) = \text{colim } \sigma(X(\mathfrak{g}_n)) = \text{colim } (X(\mathfrak{g}_n), q_n)$ .

Then, the form of colimits in  $\text{Fib}(\Delta \Pi P_1)$  directly implies (1); notice the essential use of continuity of  $\sigma$  and  $X$  in the argument.

To prove equation (2), recall that  $\lambda$  is a colimit natural transformation and therefore

$$\varinjlim_{n \geq 0} (\lambda_n \circ \lambda'_n) = f \circ \text{id}_q; \quad \text{then } f = f \circ \text{id}_q = \text{fo}(\varinjlim_{n \geq 0} (\lambda_n \circ \lambda'_n)) = \text{V}(\text{fo} \lambda_n \circ \lambda'_n) = \text{V}(\lambda'_n \circ \lambda_n), \text{ where}$$

we have used the continuity of  $f$  and the fact that  $\text{fo} \lambda_n = \lambda'_n$ , the defining property of  $f$ .

Finally, (3) follows directly from the fact that  $(\lambda'_n, \beta'_n): \Psi(\sigma(\mathfrak{g}_n)) \rightarrow (\mathfrak{g}', p')$  in  $\text{Fib}(\Delta \Pi P_1)$ ; the form of morphisms in this category implies that  $\beta'_n: \Delta \Pi P_1(\lambda'_n)(\gamma(\mathfrak{g}_n)(q_n)) \rightarrow p'$ , which is equivalent to (3).

We can now prove that  $\eta$  exists:

$$\begin{aligned}
& \Delta \Pi P_1(\mathcal{F})(\gamma(g)(q)) = \\
& = \gamma(g') \circ \Delta \Pi P_1 X(\mathcal{F})(q) = && \text{(naturality of } \gamma) \\
& = \gamma(g') \circ \Delta \Pi P_1 X(V_n(\lambda'_n \circ \lambda_n^*) (V_n \Delta \Pi P_1 X(\lambda_n) q_n)) && \text{((1) and (2))} \\
& = \bar{V}_n \gamma(g') \Delta \Pi P_1 X(\lambda'_n \lambda_n^* \lambda_n) q_n \\
& = V_n \gamma(g') \Delta \Pi P_1(X)(\lambda'_n)(q_n) = && \text{(naturality of } \gamma) \\
& = V \Delta \Pi P_1(\lambda'_n) \circ \gamma(g_n)(q_n) \leq && \text{by (3)} \\
& \leq p'.
\end{aligned}$$

The reduction of the double to a single supremum in step 4 above is identical to steps 4 and 5 in the proof of the continuity of  $F$  in section 4, and is omitted here.

*Lemma 2:* For each  $\sigma \in \Sigma$ ,  $\sigma \leq \Psi(\sigma)$ .

*Proof:* We need to show that for any game  $g$ ,  $\sigma(g) \leq \Psi(\sigma)(g)$ .

Let  $\sigma(g) = (g, p)$ ,  $\sigma(X(g)) = (X(g), q)$ ; then it has to be shown that  $p \leq \gamma(g)(q)$ . Recall that  $(X(g), \lambda)$  is a colimit of the functor

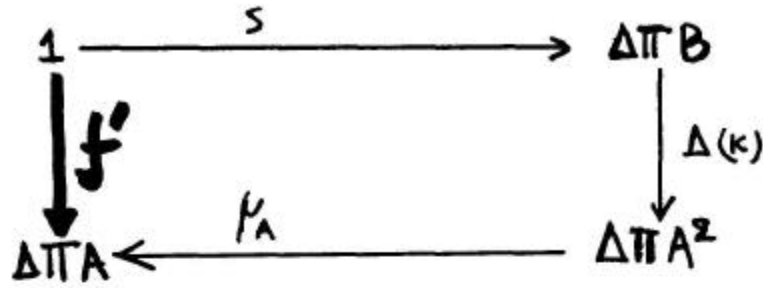
and that  $\sigma$  is a continuous functor; then if  $q_t = P_2 \sigma(F^t(g))$ ,  $t \geq 0$ , we have  $q = \bar{V}_{t=0} \Delta \Pi P_1(\lambda_t)(q_t)$ , and,

in particular,  $q_0 = p$ . Hence

$$\begin{aligned}
& \gamma(g)(q) = \\
& = \bar{V}_{t=0} (\gamma(g) \circ \Delta \Pi P_1(\lambda_t))(q_t) = && \text{(continuity of } \gamma(g)) \\
& = \bar{V}_{t=0} (\mu_{\Pi A} \circ \Delta \epsilon_t(g))(q_t) = && \text{(defining property of } \gamma) \\
& = \mu_{\Pi A} \left( \bar{V}_{t=0} \Delta \epsilon_t(g)(q_t) \right) \geq && \text{(continuity of } \mu_{\Pi A}) \\
& \geq \mu_{\Pi A}(\Delta \epsilon_0(g)(q_0)) = && (\mu_{\Pi A} \circ \Delta \epsilon_t(g) = \text{id}_{\Pi A}) \\
& = q_0 = p.
\end{aligned}$$

*Lemma 3:* For each  $\sigma \in \Sigma$ ,  $\Psi(\sigma)$  satisfies the weak revelation principle.

*Proof:* Let  $m = (S, R, f)$  be a mechanism,  $g = (A, u)$  be a game,  $P_2\Psi(\sigma)(T(m, g)) = s$  and  $m' = (1, A, f')$



It has to be shown that

$$P_2\Psi(\sigma)(T(m', g)) \geq id_A.$$

Let  $P_2\sigma(T(m, g)) = p$ ,  $P_2\sigma T(m, g) = q$ ; then by definition,  $s = \gamma(T(m, g))(q)$  and by lemma 2 :

$$(1) \quad s \geq p.$$

Let  $m'' = (1, A, (\mu_A \circ \Delta(k))(p))$ ; by the weak revelation principle for  $\sigma$ ,

$$(2) \quad P_2\sigma T(m'', g) \geq id_A$$

By (1)

$$(3) \quad (\mu_A \circ \Delta(k))(s) \geq (\mu_A \circ \Delta(k))(p)$$

By (3) and the definitions of  $m'$ ,  $m''$  there is a morphism in  $M$

$$(4) \quad (id_1, id_A, \alpha): m'' \rightarrow m'$$

Hence, there is a morphism in  $M \times G$

$$(5) \quad (id_1, id_A, \alpha, id_g)(m'', g) \rightarrow (m', g)$$

By the functoriality of  $\sigma$ ,  $T$ :

$$(6) \quad \sigma T(id_1, id_A, \alpha, id_g): \sigma T(m'', g) \rightarrow \sigma T(m', g)$$

By (6)

$$(7) \quad P_2\sigma T(m'', g) \leq P_2\sigma T(m', g)$$

By (2) and (7)

$$(8) \quad P_2\sigma T(m', g) \geq id_A$$

By lemma (2),

$$(9) \quad P_2 \Psi(\sigma)(T(m', g)) \geq P_2 \sigma T(m', g)$$

By (9) and (8)

$$(10) \quad P_2 \Psi(\sigma)(T(m', g)) \geq \text{id}_A \quad \text{QED}$$

*Lemma 4:*  $\Sigma$  has all  $\omega$ -colimits

Let  $H: \omega \rightarrow \Sigma$  be a functor, with  $H(t) = \sigma_t \in \Sigma$  and  $H(t \rightarrow t+1) = \alpha_t: \sigma_t \rightarrow \sigma_{t+1}$ , i.e.,  $\sigma_t \leq \sigma_{t+1}$  for all  $t$ . Define  $\sigma \in \Sigma$  and  $\lambda: H \rightarrow \sigma$  as follows

$$\bullet \sigma(g) = (g, \bar{\bigvee}_{t=0} P_2 \sigma_t(g));$$

the supremum exists because  $\sigma_t(g) \leq \sigma_{t+1}(g)$  for all  $t, g$ .

$$\bullet \text{If } \beta: g \rightarrow g', \sigma(\beta) = (\beta, \theta) \text{ where } \theta: \Delta \Pi P_1(\beta) (\bar{\bigvee}_{t=0} P_2 \sigma_t(g)) \rightarrow \bar{\bigvee}_{t=0} P_2 \sigma_t(g')$$

$$\bullet \lambda_t: \sigma_t \rightarrow \sigma \text{ exists because for each } g, \sigma_t(g) \leq \sigma(g)$$

To see that  $\theta$  exists, note that as  $\beta: g \rightarrow g'$  and  $\sigma_t: G \rightarrow \text{Fib}(\Delta \Pi P_1)$  is a functor,  $\sigma_t(\beta): \sigma_t(g) \rightarrow \sigma_t(g')$  is a morphism in  $\text{Fib}(\Delta \Pi P_1)$

$$\Delta \Pi P_1(\beta) (P_2 \sigma_t(g)) \leq P_2 \sigma_t(g');$$

taking suprema on both sides and exploiting the continuity of  $\Delta \Pi P_1(\beta)$

$$\Delta \Pi P_1(\beta) (\bar{\bigvee}_{t=0} P_2 \sigma_t(g)) \leq \bar{\bigvee}_{t=0} P_2 \sigma_t(g'), \text{ i.e., } \theta \text{ exists.}$$

By the form colimits take in  $\text{Sect}(\Delta \Pi P_1)$ ,  $(\sigma, \lambda)$  is a colimit of  $H$ . To check whether  $\sigma: G \rightarrow \text{Fib}(\Delta \Pi P_1)$  is continuous, let  $K: \omega \rightarrow G$  be a functor with colimit  $(g, \xi)$ ; it is shown that  $(\sigma(g), \sigma(\xi))$  is a colimit of  $\sigma \circ K$ . It suffices to show that  $P_2 \sigma(g) =$

$$= \bar{\bigvee}_{n=0} \Delta \Pi P_1(\sigma(\xi_n)) (P_2 \sigma(g_n)), \text{ or,}$$

equivalently, that

$$\bar{V}_{t=0} P_2 \sigma_t(g) = \bar{V}_{n=0} \bar{V}_{t=0} \Delta \Pi P_1(\xi_n)(P_2 \sigma_t(g_n))$$

By the continuity of each  $\sigma_t$

$$P_2 \sigma_t(g) = \bar{V}_{n=0} \Delta \Pi P_1(\xi_n)(P_2 \sigma_t(g_n))$$

Taking suprema over  $t$  on both sides yields the desired equality. Finally to show that  $\sigma$  satisfies the weak revelation principle, let  $m = (X, Y, h)$  be a mechanism,  $g = (A, u)$  a game,  $P_2 \sigma(T(m, g)) = q$ ,  $f' = \mu_A \Delta(k)(q)$  and  $m' = (1, A, f')$ . It has to be shown that  $P_2 \sigma(T(m', g)) \geq id_A$ . To see this let  $q_t = P_2 \sigma_t(T(m, g))$ ,  $f'_t = \mu_A \Delta(k)(q_t)$ ,  $m'_t = (1, A, f'_t)$ . Then

$$(1) q_t \leq q_{t+1} \leq \dots \quad \text{because } \sigma_0 \leq \sigma_1 \leq$$

$$(2) q = \bar{V}_{t=0} q_t$$

(definition of  $\sigma$ )

$$(3) m' = \bar{V}_{t=0} m'_t$$

Equation (3) follows from the fact that  $q_t \leq q_{t+1} \leq \dots \leq q$  implies  $f'_t \leq f'_{t+1} \leq \dots \leq f'$  and the form of colimits in  $M$ . Finally

$$(4) P_2 \sigma_t(T(m'_t, g)) \geq id_A$$

(weak revelation principle).

We can now write

$$P_2 \sigma(T(m', g)) =$$

(definition of  $\sigma$ )

$$= \bar{V}_{t=0} P_2 \sigma_t(T(m'_t, g)) =$$

(by (3))

$$= \bar{V}_{t=0} \bar{V}_{s=0} P_2 \sigma_t(T(m'_s, g)) =$$

$$= \bar{V}_{t=0} P_2 \sigma_t(T(m'_t, g)) \geq$$

(by (4))

$$\geq id_A. QED.$$

The third step, a reduction of a double to a single supremum, is done in the same way as steps

4 and 5 in the proof of continuity of F in section 4.

*Lemma 5:*  $\Psi: \Sigma \rightarrow \Sigma$  is a continuous functor.

*Proof* by lemmas 1 and 3  $\Psi(\sigma)$  is in  $\Sigma$  for each  $\sigma \in \Sigma$ ; to show  $\Psi$  is continuous, it suffices to show that if  $H: \omega \rightarrow \Sigma$  is a functor with colimit  $(\sigma, \lambda)$ , then  $(\Psi(\sigma), \Psi(\lambda))$  is a colimit of  $\Psi \circ H$ . By lemma 4, if  $H(t)$

$$= \sigma_t \text{ we have } \sigma_t \leq \sigma_{t+1} \text{ for all } t \text{ and } P_2 \sigma(g) = \varinjlim_{t=0} P_2 \sigma_t(g), \quad \text{and we need to show that } P_2 \Psi(\sigma) \leq P_2 \Psi(\sigma_{t+1}) \text{ and that } P_2 \Psi(\sigma)(g) = \varinjlim_{t=0} P_2 \Psi(\sigma_t)(g) \quad \text{In fact,}$$

$$\begin{aligned} P_2 \Psi(\sigma) &= \gamma(g)(P_2 \sigma_t(X(g))) \leq \\ &\leq \gamma(g)(P_2 \sigma_{t+1}(X(g))) = P_2 \Psi(\sigma_{t+1}) \text{ and } P_2 \Psi(\sigma)(g) = \gamma(g)(P_2 \sigma(X(g))) = \\ &= \gamma(g) \left( \varinjlim_{t=0} P_2 \sigma_t(X(g)) \right) = \\ &= \varinjlim_{t=0} \gamma(g) (P_2 \sigma_t(X(g))) = \\ &= \varinjlim_{t=0} P_2 \Psi(\sigma_t)(g). \end{aligned}$$

The main theorem of this section can now be stated

*Theorem 1:* for each  $\sigma \in \Sigma$ , there is a coordination-proof solution concept  $\hat{\sigma} = \Psi(\hat{\sigma})$  obtained by

$$\hat{\sigma} = \varinjlim_{t=0} \Psi^t(\sigma)$$

*Proof:* By lemma 3,  $\sigma \leq \Psi(\sigma)$ : by lemma 5,  $\sigma \leq \Psi(\sigma) \leq \Psi^2(\sigma) \leq \dots$ ; hence  $\hat{\sigma} = \varinjlim_{t=0} \Psi^t(\sigma)$  exists. Then,

$$\Psi(\hat{\sigma}) =$$

$$\Psi \left( \varinjlim_{t=0} \Psi^t(\sigma) \right) = \quad \text{(continuity of } \Psi \text{)}$$

$$= \varinjlim_{t=0} (\Psi^{t+1}(\sigma)) = \varinjlim_{t=0} \Psi^t(\sigma) = \hat{\sigma}.$$

It is suggestive to say that  $\sigma$  is completed into a coordination-proof solution concept  $\hat{\sigma}$  by the process described in theorem 1.

The prediction  $q = P_2 \sigma(X(g))$  of a solution concept  $\sigma$  on the universal game  $X(g)$  generated by a game  $g$  is also a (probabilistic) prediction on  $F(X(g)) \approx X(g)$ , the game whose strategy space include all mechanisms on  $X(g)$ . This prediction defines (stochastically) equilibrium institutions, and equilibrium strategies of players in these institutions.

## 6. Conclusions

What has been accomplished in the present paper? To answer this, recall a statement of Kreps ((1990, p. 388): "The ground rules in noncooperative game theory are that the description of the game should include all relevant opportunities for the players." The introduction demonstrated why in general the attempt to capture all relevant opportunities will lead to an infinite regress; section 2 that it is impossible to resolve this infinite regress in the category of sets; and section 4 that the infinite regress is resolved in the category of domains and Scott-continuous functions: "all the relevant opportunities" are captured by universal games, and universal games are fixed points of the functor whose iterations generate the infinite regress.

Secondly, section 5 has demonstrated the possibility of constructing solution concepts whose predictions are outcomes immune to successful attempts to coordinate, where "success" is defined by the solution concept itself.

Thirdly, the last paragraph of section 5 has described how to define equilibrium institutions, if we accept the identification of institutions with mechanisms. This is the most interesting, and least developed, theme of the paper.

Fourthly, the whole paper is an example of applications of categorical techniques to economics, and hopefully an indication of their usefulness and relevance. These techniques were useful both in the formulation of problems (try thinking about infinite regress without bringing in functorial fixed points) and in the proofs of theorems (there is a naturality argument, and a commutative diagram, in almost every page of sections 4 and 5).

The research reported in the previous sections is still in its infancy. The lack of interesting examples is particularly vexing: the only coordination-proof solution concept I can concretely describe at this moment is the trivial one that assigns unit mass to the bottom element of each game. This shortcoming can be overcome by further research. To explain this, recall that the theory provides an algorithm that will take any admissible solution concept  $\sigma \in \Sigma$  and transform it into a coordination-proof solution concept  $\hat{\sigma} \in \Sigma$  (Theorem 5.1). The fact that each  $\sigma$  in  $\Sigma$  is restricted to be single-valued implies that no existing solution concept can be utilized as an initial value in the algorithm that constructs coordination-proof solution

concepts; the Harsanyi-Selten theory might suggest itself at this point, but recall that it applies to finite games, while we need single-valued predictions for  $g$  and  $X(g)$ : even if  $g$  is finite,  $X(g)$  is not.

How can further research overcome this problem? A possible answer proceeds along the following lines. First, extend the methods of the present paper to multivalued solution concepts; the main difficulty here is technical, namely the construction of a continuous power set functor so that the fixed-point constructions go through. An appropriate adaptation of powerdomains, a concept developed for computer science purposes by Plotkin (1976) might resolve this difficulty. Secondly, apply the (extended) methods twice: once to refine multivalued solution concepts "as much as possible," and a second time to render them coordination-proof. (Refinements result from allowing agents to propose mechanisms that recommend which equilibrium is to be played, mechanisms that recommend which mechanisms to propose, ... etc.) This second step will define an algorithm  $\psi$  whose inputs and outputs are solution concepts, similar to the one constructed by Theorem 5.1.

Finally, define the initial point of the algorithm to be your favorite solution concept  $\sigma$ , and let the algorithm compute beginning from  $\sigma$ : the outcome of the computation will be a solution concept

$$\delta = \bigvee_{i=0}^{\infty} \psi^i(\sigma) \quad \text{which is coordination-proof in a double sense: all possible attempts to refine, and move away}$$

from inefficient, equilibria will have been taken into account by the process of step 2; in addition, your favorite concept of self-enforcement will have been taken into account by the initial point  $\sigma$ . Interesting examples can then be constructed by choosing the game  $g$  of interest and the initial solution concept  $\sigma$ . All this is will be developed in a sequel to the present paper.

Current work in progress that involves functorial fixed points is Brown (1989) on the problem of defining statements of the form "knowledge operators are common knowledge;" Boylan (1990) on the problem of learning how to learn; and Alameddine (1991) on Gauthier's bargaining approach to social choice. Recent work related to universal games is in Lagunoff (1988).

## Appendix 1

*Theorem:* The probability functor  $\Delta: D_E \rightarrow D_E$  is continuous, i.e., it preserves  $\omega$ -colimits. The proof will proceed in several steps.

*Step 1:* If  $f: X \rightarrow Y$  is an embedding, and if  $U$  is a Scott-open subset of  $Y$ , then

$$(f^*)^{-1}(U) \in U$$

*Proof:*

$$y \in (f^*)^{-1}(U) \Rightarrow$$

$$(f^*y) \in U \Rightarrow$$

$$y \in U.$$

The last step follows from two facts:

$(f^*y) \leq y$  (definition of  $f^*$ ), and  
 $x \in U, x \leq y$  imply  $y \in U$  ( $U$  is Scott open)

*Step 2:* If  $f: X \rightarrow Y$  is an embedding, then so is  $\Delta(f): \Delta(X) \rightarrow \Delta(Y)$ ; hence,  $\Delta$  maps  $D_E$  into  $D_E$ .

*Proof:* It is shown that  $\Delta(f^*)$  is an adjoint of  $\Delta(f)$ , i.e., that

$$\Delta(f^*) \circ \Delta(f) = id_{\Delta(X)}. \tag{1}$$

$$\Delta(f) \circ \Delta(f^*) \leq id_{\Delta(Y)}. \tag{2}$$

(1) follows from the fact that  $\Delta$  is a functor:

$$\begin{aligned} \Delta(f^*) \circ \Delta(f) &= \\ &= \Delta(f^* \circ f) = \\ &= \Delta(id_X) = \\ &= id_{\Delta(X)}. \end{aligned}$$

To establish (2), it suffices to show that for each  $p \in \Delta(Y)$  and each Scott open subset  $U$  of  $Y$ ,

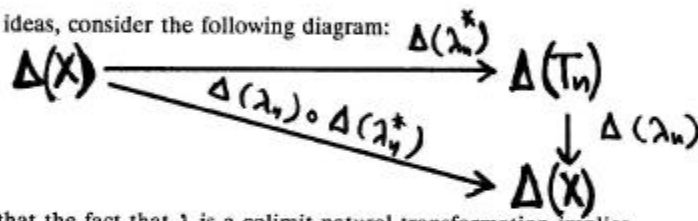
$$\begin{aligned}
& \Delta(f) \circ \Delta(f^*)(p(U)) \leq p(U). \text{ In fact,} \\
& \Delta(f) \circ \Delta(f^*)(p(U)) = \\
& = \Delta(f \circ f^*)(p(U)) = \\
& = p((f \circ f^*)^{-1}(U)) \leq \quad \text{(step1)} \\
& \leq p(U). -
\end{aligned}$$

We now show that  $\Delta$  preserves  $\omega$ -colimits, i.e., that if  $T: \omega \rightarrow D_{\mathbb{E}}$  is a functor with colimit  $(X, \lambda)$ , then the composite functor  $\Delta \circ T$  has colimit  $(\Delta(X), \Delta(\lambda))$ : it suffices to show that

$$\Delta(\lambda_n) \circ \Delta(\lambda_n^*) \leq \Delta(\lambda_{n+1}) \circ \Delta(\lambda_{n+1}^*), \text{ all } n \quad (3)$$

$$\bigvee_{n=0} \Delta(\lambda_n) \circ \Delta(\lambda_n^*) = id_{\Delta(X)}. \quad (4)$$

To fix ideas, consider the following diagram:



Finally, recall that the fact that  $\lambda$  is a colimit natural transformation implies

$$\lambda_n \circ \lambda_n^* \leq \lambda_{n+1} \circ \lambda_{n+1}^*, \text{ all } n \quad (5)$$

$$\bigvee_{n=0} (\lambda_n \circ \lambda_n^*) = id_X. \quad (6)$$

*Step 3:* If  $U$  is a Scott-open subset of  $X$  and  $U_n = (\lambda_n \circ \lambda_n^*)^{-1}(U)$ , then for all  $n$ :  $U_n \subseteq U_{n+1}$  and  $U_n \subseteq U$ .

*Proof:*

$$\begin{aligned}
y \in U_n &= \\
(\lambda_n \circ \lambda_n^*)(y) \in U &= \\
(\lambda_{n+1} \circ \lambda_{n+1}^*)(y) \in U &= \\
y \in U_{n+1} &\bullet \text{ Hence, } U_n \subseteq U_{n+1}. \\
y \in U_n &= \\
(\lambda_n \circ \lambda_n^*)(y) \in U &= \quad (\lambda_n \circ \lambda_n^* \leq id_X; U \text{ is Scott-open}) \\
y \in U. &\quad \text{ Hence } U_n \subseteq U.
\end{aligned}$$

Step 4: Under the notation of step 3,

$$U = \bigcup_{n=0}^{\infty} U_n$$

Proof: If  $x \in U$ , then  $x = \bigvee_{n=0}^{\infty} \lambda_n \circ \lambda_n^*(x)$ , by (6). But  $U$  is Scott-open; hence, there is some  $n$  such that  $\lambda_n \circ \lambda_n^*(x) \in U$ , i.e.,  $x \in U_n$ . QED.

Step 5: (3) holds true.

Proof: Let  $p \in \Delta(X)$ ,  $U$  a Scott-open subset of  $X$ . Then

$$\begin{aligned}
\Delta(\lambda_n) \circ \Delta(\lambda_n^*)(p)(U) &= \\
\Delta(\lambda_n \circ \lambda_n^*)(p)(U) &= \\
p((\lambda_n \circ \lambda_n^*)^{-1}(U)) &= \\
p(U_n) &\leq \text{step(3)} \\
P(U_{n+1}) &= \\
= \Delta(\lambda_{n+1}) \circ \Delta(\lambda_{n+1}^*)(p)(U). &\quad \text{QED.}
\end{aligned}$$

Step 6: (4) holds true.

$$\begin{aligned}
& \varinjlim_{n \rightarrow 0} (\Delta(\lambda_n) \circ \Delta(\lambda_n^*)) (p)(U) = \\
& \varinjlim_{n \rightarrow 0} (\Delta(\lambda_n \circ \lambda_n^*)) (p)(U) = \\
& = \varinjlim_{n \rightarrow 0} (p((\lambda_n \circ \lambda_n^*)^{-1})(U)) = \\
& = \varinjlim_{n \rightarrow 0} (p(U_n)) = \quad (\text{steps 3 and 4}) \\
& = p(U). \quad \text{QED.}
\end{aligned}$$

## Appendix 2

*Theorem:* Let  $(X, \lambda)$  be a colimit of the functor  $T: \omega \rightarrow K$ , where  $K$  is an arbitrary category. Then  $\lambda$  is jointly surjective, i.e., if  $f$  and  $g$  are morphisms in  $K$  such that for all  $n$   $f \circ \lambda_n = g \circ \lambda_n$ , then  $f = g$ .

Proof: Let  $f, g: X \rightarrow Y$  be morphisms in  $K$ , and let  $\lambda'_n: T_n \rightarrow Y$  be defined by

$$f \circ \lambda_n = \lambda'_n = g \circ \lambda_n.$$

Then,  $\lambda': T \rightarrow Y$  is a natural transformation; hence, by the definition of colimits, there is a *unique* morphism  $h: X \rightarrow Y$  such that  $h \circ \lambda_n = \lambda'_n$  for all  $n$ ; but then  $f = h$  and  $g = h$ , i.e.,  $f = g$ .

## References

- Alameddine, F. (1991): Some Problems of Infinite Regress in Social Choice Models. *Federal Reserve Bank of Cleveland* working paper no. 9016.
- Boylan, R. (1990): unpublished notes.
- Brown, D. (1989): unpublished notes.
- Coquand, J., Gunter, C., and Winskel, G. (1989): Domain- Theoretic Models of Polymorphism. *Information and Computation*, 81:123-167.
- Crawford, V. (1985): Efficient and Durable Decision Rules: A Reformulation. *Econometrica*, 53(4): 817-835.
- Dubey, P., Mas-Colell, A., and Shubik, M. (1980): Efficiency Properties of Strategic Market Games: An Axiomatic Approach. *JET*, 22(4): 339-362.
- Gauthier, D. (1985): *Morals by Agreement*.
- Graham, S. K. (1988): Closure Properties of a Probabilistic Domain Construction, in M. Main (ed.): *Math. Foundations of Programming Language Semantics*. LNCS, v. 298, Springer.
- Green, E. (1984): Continuum and Finite — Player Noncooperative Models of Competition. *Econometrica*, 52(4): 915-993.
- Gunter, C. (1989): Comparing Categories of Domains, in M. Main (ed.): *Math. Foundations of Programming Language Semantics*. LNCS, v. 298, Springer.
- Halmos, P. (1950): *Measure Theory*. Springer.
- Hillas, J. (1990): On the Definition of Strategic Stability of Equilibria. *Econometrica*, 58(6) 1365-1390.
- Holmstrom, B. and Tirole, J. (1989): Theory of the Firm; in R. Schmalensee and R. Willig (eds.): *Handbook of Industrial Organization*. North Holland.
- Khan, A. and Sun, Y. (1990): On a Reformulation of Cournot-Nash Equilibria. *Journal of Math Analysis and Applications*, forthcoming.
- Kreps, D. (1990): *A Course in Microeconomic Theory*. Princeton, UP.
- \_\_\_\_\_ (1990b). The Clarendon Lectures; forthcoming, *Oxford University Press*.
- Lagunoff, R. (1988): Fully endogenous mechanism selection. Mimeo.

- Lehmann, D. and Smyth, M. (1981): Algebraic Specification of Data Types. *Mathematical Systems Theory*, 14, 97-139.
- Manes, E. and Arbib, M. (1986): *Algebraic Approaches to Program Semantics*. Springer.
- Mertens, J. F. and Zamir, S. (1985): Formulation of Bayesian Analysis for Games of Incomplete Information. *Int. J. of Game Theory*, 1-31.
- Milgrom, P. and Roberts, J. (1989): Rationalizability, Learning and Equilibrium in Games with Strategic Complementarities. *Econometrica* 58(6) 1255-1278.
- Myerson, R. (1982): Optimal Coordination Mechanisms in Generalized Principal-Agent Models. *Journal of Mathematical Economics*, 10, 67-81.
- Plotkin, G. (1976): A Powerdomain Construction. *Siam Journal of Computing*, 5, 452-488.
- Roth, A. (1984): The Evolution of the Labor Market for Medical Interns and Residents. *Journal of Political Economy*, 92, 991-1016.
- Roth, A. (1990): A Natural Experiment in the Organization of Entry Level Labor Markets: Regional Markets for New Physicians in the U. K. *University of Pittsburgh Working Paper # 245*.
- Roth, A. and Sotomayor, M. (1990): *Two-Sided Matching*. Cambridge, U. P.
- Saheb-Djahromi, N. (1980): CPO's of Measures for Nondeterminism. *Theoretical Computer Science*, 12, 19-37.
- Scott, D., et. al. (1980): *A Compendium of Continuous Lattices*. Springer.
- Smyth, M. and Plotkin, G. (1982): The Category-Theoretic Solution of Recursive Domain Equations. *SIAM J. of Computing* 11(4): 761-783.
- Vassilakis, S. (1990): Functorial Fixed Points. Mimeo.