האוניברסיטה העברית בירושלים
THE HEBREW UNIVERSITY OF JERUSALEM

ITERATED EXPECTATIONS, COMPACT SPACES, AND COMMON PRIORS

By

ZIV HELLMAN

Discussion Paper # 522 October 2009

מרכז לחקר הרציונליות

CENTER FOR THE STUDY OF RATIONALITY

Feldman Building, Givat-Ram, 91904 Jerusalem, Israel
PHONE: [972]-2-6584135 FAX: [972]-2-6513681
E-MAIL: ratio@math.huji.ac.il
URL: http://www.ratio.huji.ac.il/

ITERATED EXPECTATIONS, COMPACT SPACES, AND COMMON PRIORS

Ziv Hellman
The Department of Mathematics and the Centre for the Study of Rationality, The Hebrew University of Jerusalem

Abstract. Extending to infinite state spaces that are compact metric spaces a result previously attained by D. Samet solely in the context of finite state spaces, a necessary and sufficient condition for the existence of a common prior for several players is given in terms of the players' present beliefs only. A common prior exists if and only if for each random variable it is common knowledge that all Cesàro means of iterated expectations with respect to any permutation converge to the same value; this value is its expectation with respect to the common prior. It is further shown that compactness is a necessary condition for some of the results.

1. Introduction

The common prior assumption, ever since it was introduced into the study of games with incomplete information by Harsányi (1967-8), posits that all women and men are 'created equal' with respect to probability assessments in the absence of information hence the term common prior and all differences in probabilities should, in principle, be traced to asymmetries in information received over time. The idea has become very pervasive, and in most applications of type spaces to economics it is assumed that players' beliefs can indeed be derived from a common prior by Bayesian updating. A prior probability can be interpreted as the beliefs of a player in a previous period. In many models, however, any previous period is either fictional or irrelevant to the matter being studied. It is also clear that there are many plausible models of type spaces in which it is impossible for the players to have arrived at their current beliefs via updating from a common prior. This leads naturally to the question of whether a criterion can be identified by which one can tell, through the current beliefs of the players, that they have a common prior.

Aumann (1976), in his celebrated agreeing-to-disagree theorem, presented a necessary condition for the existence of a common prior in terms of present beliefs: if there is a common prior, then it is impossible to have common knowledge of difference in the beliefs of any given event. Numerous authors extended this result and applied it to interactions between agents in various situations. The typical result is a 'no-bet' or 'no-trade' theorem (for example, Milgrom & Stokey (1982), or Sebenius & Geanakoplos (1983)) agents who start with common prior distributions will never agree to engage in speculative trade based on differences in private

---

Date: This Version: 20 August 2009.

information that they subsequently receive. As soon as it becomes common knowledge that they wish to trade, their expectations for the value of assets in question become identical.

Aumann (1987) grants further importance to the common prior assumption by showing that rational players with a common prior will play according to a correlated equilibrium distribution. Questions relating to the common prior assumption under incomplete information became especially prominent in a series of exchanges between Gul (1998) and Aumann (1998). Gul (1998) criticised the common prior assumption, especially when it leans on a supposed 'dynamic story' the view that players assessing differing expectations of events do so solely because of differences in the private information they possess respectively, because in some hypothetical past they shared a common prior, with their current beliefs posterior to a, perhaps distant, past of shared probabilities. Gul (1998) argues that 'since there never was a prior stage, the prior distribution is meaningless'.

Aumann (1998), in reply, essentially restates the position of Aumann (1987), which includes the assertion that 'people with different information may legitimately entertain different probabilities, but there is no rational basis for people who have always been fed precisely the same information to do so'. Highlighting 'differences in information' as the sole bearer of distinction, Aumann (1998) postulates that 'if the beliefs at an "actual" prior stage are different and not commonly known, then there must be differential information already at that stage', and then argues that analysis must proceed to a further earlier stage until all differences in information have been purged and a common primeval prior can be identified. He is even willing to go so far as to say 'if one sets forth all relevant information in sufficient detail, then in principle, there should be no room for differing probabilities'.

Aumann (1998) does, however, agree with Gul on at least one point, stating 'we agree with Gul that the above development of common priors is essentially dynamic And though we consider this perfectly legitimate and intuitive, we also agree that it would be desirable to characterize common priors directly in terms of the "current", posterior probabilities'. Efforts to obtain such a characterisation of common priors were undertaken in the 1990s. Several researchers, by extending the notion of disagreement to differences in the expectation of a general random variable, were able to show that the impossibility of there being common knowledge of disagreement is also a sufficient condition for the existence of a common prior. Morris (1994) proved this result, in the context of a finite state space, by considering how the absence of common priors can affect the willingness to conduct trade in various trading environments. Feinberg (2000), utilising techniques closer to pure game theory, showed that both in the context of finite state spaces and of infinite but compact state spaces, a lack of common priors implies the existence of at least one common bet for which each player will subjectively assess that he or she has positive expectation and showed that compactness is necessary for that result. Samet (1998b) proved the same result, with a finite state space, using separating hyperplane techniques. The extension of this result to compact spaces, with separating hyperplane techniques, was attained in Heifetz (2006).

As Samet (1998a) pointed out, this criterion, based on disagreements, satisfactorily solves the question of how one can tell when players have a common prior, but it fails to express the common prior in a meaningful way; the fact that a disagreement cannot be common knowledge may guarantee the existence of a common prior, but it tells us nothing about this common prior. He then proceeded, in that same paper, to present a very different necessary and sufficient condition for the existence of a common prior that not only identifies the common prior when it exists, but also provides an epistemically meaningful interpretation to it.

This condition is expressed intuitively in Samet (1998a) in a colourful story. Imagine that Adam and Eve who have both excelled in their studies at the same school of economics are asked what return they expect on IBM stock. Having been exposed to different sources of information, we oughtnt be surprised if the two provide different answers. But we can then go on to ask Eve what she thinks Adam's answer was. Being a good Bayesian, she can compute the expectation of various answers and come up with Adam's expected answer. Likewise, Adam can provide us with what he expects was Eve's answer to that question. This process can continue, moving back and forth between Eve and Adam, theoretically forever. There are, in this example, two possible infinite sequences of alternating expectations, one that starts with Eve and one that starts with Adam. Samet calls this process 'obtaining an iterated expectation', and shows that, when the relevant state space is finite, there exists a common prior if and only if both of these sequences converge to the same limit.

He achieves this result by representing Adam's beliefs[1] by a type matrix M_1 and Eve's beliefs by a type matrix M_2. These then form two permutation matrices, $M_{\sigma_1} = M_2 M_1$, which is intended to be used for the process of obtaining iterated expectations starting with Adam, and $M_{\sigma_2} = M_1 M_2$ which does the same for the iterated expectations starting with Eve. It then turns out to be the case that both M_{σ_1} and M_{σ_2} are ergodic Markov matrices, and hence by standard results in Markov chain theory, each of them has a unique invariant probability measure, which may be labelled respectively p_1 and p_2. It is then shown that if $p_1 \neq p_2$, Adam and Eve cannot share a common prior. On the other hand, if $p_1 = p_2$, then not only is there a common prior, it has positively been identified – it is precisely $p := p_1 = p_2$.

We can term the criterion by which a common prior is ascertained to exist by the identities of the invariant probability measures associated with permutation matrices *Samet's criterion.* Samet (1998a), however, proves it only in the context of finite state spaces. Given the results in Feinberg (2000) and Heifetz (2006), which extend the other major criterion for a common prior to compact state spaces, it is natural to wonder whether an analogue of Samets characterisation can also be shown to hold in compact state spaces.

It is the goal of this paper to show that there is an affirmative answer to that question. The significance of such a result is clear, given that there are many models of interest which involve infinite state spaces and cannot be reduced to a finite space we therefore extend the application of the Samet criterion to many models to which it previously could not be applied.

It is also shown here, by way of an example, that compactness is necessary in the sense that if one does not assume compactness, the infinite dimensional analogue

[1] For the sake of simplicity here, we will make the mild technical assumption that the entire relevant state space is the meet of the Adam and Eve type space.

of the permutation matrix M_σ may not have a well-defined invariant probability measure and without that, the subsequent propositions do not follow, and indeed in that case there is no guarantee that the Samet criterion for checking the existence of a common prior can even be applied intelligibly, as there may not be invariant probability measures that can be compared against each other.

It should also be noted that the Samet criterion is significant because it provides, in principle, a way of calculating a common prior given a type space. In the finite state space context, one can form the type matrices and apply numerical solutions for calculating invariant probability measures in Markov chains a subject of active research in order to ascertain whether or not there is a common prior and if one exists, to identify it. Similarly, with the extension here of the Samet criterion to the more general compact spaces, it now becomes possible, given knowledge of the players type spaces, in principle to estimate the expected values of random variables by use of numerical solutions, such as those appearing in e.g. Hernández-Lerma & Lasserre (2003).

The following rough correspondences exist between results in this paper and those that appear in Samet (1998a), save for the fact that the results in that paper are strictly limited to finite state spaces, whereas that restriction is lifted here: Proposition 1 here is (roughly) an infinite state space version of Proposition 4 of Samet (1998a); Proposition 2 here corresponds to Proposition 5 of Samet (1998a); and similarly Proposition 3 corresponds to Proposition 2' and Proposition 4 to Theorem 1'.

1.1. **Outline of Paper.**

Definitions relating to type spaces and priors in general appear in Section 2. Section 2.1 lists the major definitions and results from Markov chain theory that will be used in the sequel. The connection between Markov chain theory and type spaces is made explicit in Section 3, which also introduces our definition of an 'everywhere mutually positive' type space. The main results of the paper are in Sections 4 and 5. Section 6 shows that compactness is necessary for the conclusion of Proposition 1, which is needed for the proofs of the subsequent propositions.

2. Preliminary Definitions and Results

A *type space* for a set of players is a tuple $< I, \Omega, \mathcal{F}, (\Pi_i, t_i)_{i \in I} >$. The set of players is denoted by $I = \{1, ..., n\}$, where $n \geq 2$. Ω is a measurable space of arbitrary cardinality, whose elements are called states. $\mathcal{F}$ is a σ-field of events (subsets of Ω). For each player $i \in I$, Π_i is a partition of Ω, which may be termed player i's knowledge partition, and $t_i(\cdot, \omega)$ denotes a belief – or probability distribution on $(\Omega, \mathcal{F})$ – associated with each player i at each state. We further assume that each element of each partition Π_i is an element of $\mathcal{F}$, so that the atoms of the knowledge partitions of each player are $\mathcal{F}$-measurable. K_i will denote the sub σ-field of $\mathcal{F}$ generated by Π_i.

The probability distributions $t_i(\cdot, \omega)$ for each player i and each state ω are required to satisfy:

1. $t_i(\Pi_i(\omega)|\omega) = 1$
2. For all $\omega' \in \Pi_i(\omega)$, $t_i(A, \omega') = t_i(A, \omega)$

The *meet* of $(\Pi_i)_{i\in I}$ is the partition Π of Ω which is the finest among all partitions that are coarser than Π_i for each i. For each ω, $\Pi(\omega)$ denotes the element of the meet containing ω. A somewhat more constructive way to define the elements of the meet utilises the concept of reachability. A state ω' is *reachable* from ω if there exists a sequence $\omega_0 = \omega, \omega_1, \omega_2, \ldots, \omega_m = \omega'$ such that for each $k \in \{0, 1, \ldots, m-1\}$, there exists a player i_k such that $\Pi_{i_k}(\omega_k) = \Pi_{i_k}(\omega_{k+1})$. It is well-known that $\omega' \in \Pi(\omega)$ iff ω' is reachable from ω, so that the relation of reachability can be used to define the partition Π, and this characterisation of the meet will be used in proofs in the body of this paper. For an event A, the event that A is *common knowledge* is the union of all the elements of Π contained in A.

A *random variable* is a real-valued function on Ω. For a probability measure ν and a random variable f on Ω, the expectation of f with respect to ν is $\nu f := \int_\Omega f(\omega) d\nu(\omega)$. For each player i and random variable f, i's expectation of f, denoted $E_i f$ is the random variable

$$(E_i f)(\omega) := \int_\Omega f(\overline{\omega}) dt_i(\overline{\omega}|\omega).$$

Given a type space, one can ask whether the space might have come to exist, in its current state, from a space with no information at all, by the players acquiring new information over time and updating their beliefs in a Bayesian manner. Each players possible initial belief on the no-information primeval space is called a prior. In general, given player is current type, there will not be a single prior there will be a set of possible priors. A main question is then whether or not the players have a common prior.

More formally, a probability measure μ over $(\Omega, \mathcal{F})$ is a prior for player i if for every event $A \in \mathcal{F}$

$$\mu(A) = \int_\Omega t_i(A|\omega) d\mu(\omega).$$

In words, μ is a prior for player i if i's types $t_i(\omega)$ are the posteriors of μ conditional on i's information function t_i. A probability distribution $P \in \Delta('O)$ is a common prior if it is a prior for each of the players $i \in I$.

2.1. **Markov Transitions.**

When working with a finite state space, a Markov chain is typically represented by a series of random variables $\{X_1, X_2, \ldots\}$ along with a transition matrix M, such that the (i,j)-th element of M is the probability that $X_{n+1} = j$ given that $X_n = i$.

In transferring this idea to a more general state space, we cannot always expect to measure the probability that the value of a random variable in a successive time period will be a specific state, but we can ask what the probability is that it will be in an event. In formulae, if $(\Omega, \mathcal{F})$ is a measurable space, $(X, \mathcal{B}, \mathcal{P})$ a probability space, E an event in $\mathcal{F}$, and $\{\zeta_1, \zeta_2, \ldots\}$ is a sequence of Ω-valued random variables defined on X, our analogue of the transition matrix will be given by $M(E|\omega) := \mathcal{P}(\zeta_{n+1} \in E | \zeta_n = \omega)$. This motivates the standard definition of a general Markov transition probability function:

A *stochastic kernel* or *Markov transition probability function* on $(\Omega, \mathcal{F})$ is a function M such that

(1) $M(\cdot|\omega)$ is a probability measure for each fixed $\omega \in \Omega$
(2) $M(E|\cdot)$ is an $\mathcal{F}$-measurable function on Ω for each fixed event $E \in \mathcal{F}$

One of the most important aspects of finite state Markov transitions is the interpretation of the n-th power of a Markov transition matrix as representing the n-th step of iterating the transition probabilities encoded in the matrix. The analogue in general state spaces iterates a Markov transition probability function M by the following recursive definition:

$$M^n(E|\omega) = \int_\Omega M^{n-1}(E|\omega')dM(\omega'|\omega) = \int_\Omega M(E|\omega')dM^{n-1}(\omega'|\omega)$$

for all $E \in \mathcal{F}$ and $\omega \in \Omega$.

In the rest of this section, fix a Markov transition probability function M.

Let $\Delta(\Omega)$ denote the space of probability measures on Ω, with this space naturally outfitted with the induced weak* topology. It is possible to regard M as a function from $\Delta(\Omega)$ to $\Delta(\Omega)$, as follows: For each $\nu \in \Delta(\Omega)$, let

$$(\nu M)(E) := \int_\Omega M(E|\omega)d\nu(\omega)$$

Then M acts on $\Delta(\Omega)$ by way of $\nu \mapsto \nu M$. Using this notation, a probability measure ν is *invariant* with respect to M if $\nu = \nu M$. If such a measure exists, M is said to admit an invariant probability measure.

The transition probability function M can also be considered as operating on bounded functions in the following way. For each bounded integrable function $f : \Omega \to \mathbb{R}$, let Mf be the bounded function

$$Mf(\omega) := \int_\Omega f(\widehat{\omega})dM(\widehat{\omega}|\omega).$$

If ν is an invariant probability measure with respect to M, then M can also be considered to be a linear operator on $L_1(\nu) := L_1(\Omega, \mathcal{F}, \nu)$ into itself. We can then define, for any k and $f \in L_1(\nu)$

$$M^k f(\omega) := \int_\Omega f(\widehat{\omega})dM^k(\widehat{\omega}|\omega).$$

We have in addition the concept of the Cesàro mean, defined as

$$M^{(n)}f(\omega) := \frac{1}{n}\sum_{k=0}^{n-1} M^k f(\omega).$$

If Ω has a topology τ, denote the class of bounded continuous functions with respect to τ from Ω to $\mathbb{R}$ by $C(\Omega)$. Then M satisfies the *weak Feller property* if M maps $C(\Omega)$ to $C(\Omega)$.

Let φ be a non-trivial σ-finite measure for the space Ω. A Markov transition function M is *φ-irreducible* if

$$\sum_{n=1}^{\infty} M^n(E|\omega) > 0$$

for all $\omega \in \Omega$ whenever $\varphi(E) > 0$ for $E \in \mathcal{F}$.

We will make use of the following important theorems from the theory of Markov chains. These three theorems appear, in Hernández-Lerma & Lasserre (2003) respectively as Theorem 7.2.3, Proposition 4.2.2, and an amalgam of Theorem 2.3.4, Proposition 2.4.2 and Proposition 2.4.3:

THEOREM (Existence of invariant probability measure). Let Ω be a compact metric space, and let M be a Markov transition function on Ω. Then M admits an invariant probability measure.

THEOREM (Uniqueness of invariant probability measure). Let M be a φ-irreducible Markov transition function and suppose that M admits an invariant probability measure ν. Then ν is the unique invariant probability measure for M.

THEOREM (Birkhoff's Ergodic Theorem for Markov processes). Let M be a Markov transition function that admits an invariant probability measure ν. For every $f \in L_1(\nu)$ there exists a function $f^* \in L_1(\nu)$ such that

$$P^{(n)} f \to f^* \nu \text{ almost everywhere.}$$

and

$$\int f^* d\nu = \int f d\nu$$

In addition, if ν is the unique invariant probability measure of M, then f^* is constant ν-almost everywhere, and $f^* = \int f d\nu$, ν-almost everywhere, so that

$$\text{the time-average} \lim_{n\to\infty} M^{(n)} f = \text{the space-average} \int f d\nu,\ \nu\text{-a.e.}$$

3. Type Spaces Within the Markov Framework

3.1. **Relating Type Spaces to Markov Processes.**

In this section, we relate the concepts of type spaces and Markov processes (similarly to the way this is accomplished in Samet (2000)).

First, note that by definition, the probability measure $t_i(\cdot|\cdot)$ associated with each player i satisfies the conditions for being a Markov transition probability function, hence we can relate to it as such. We will relabel $t_i(\cdot|\cdot)$ as M_i in the sequel when we wish to emphasise we are treating it as a Markov transition.

In general, given any two probability measures P_1 and P_2, one can define a new probability measure $P_2 P_1(E|\omega)$ by

$$P_2 P_1(E|\omega) = \int_\Omega P_2(E|\widehat{\omega}) dP_1(\widehat{\omega}|\omega)$$

This obviously can be iterated any number of times. In particular, given a measure P, we can construct an infinite sequence of measures $\{P^n(\cdot|\omega)_{n\geq 1}\}$. In our specific context, given any two players i and j and a measurable event E, the probability distribution $t_i t_j(E|\omega)$ based on t_i and t_j is similarly defined by

$$t_i t_j(E|\omega) = \int_\Omega t_i(E|\widehat{\omega}) dt_j(\widehat{\omega}|\omega)$$

In particular, given an element σ in $Sym(I)$, the set of all permutations of the elements of I, define

$$t_\sigma := t_{\sigma(1)} \dots t_{\sigma(n)}$$

iteratively, by using the above to define $t_{\sigma(n-1)}t_{\sigma n}$, then $t_{\sigma n-2}(t_{\sigma(n-1)}t_{\sigma n})$, and so forth.

We can now re-interpret various notions relating to a type space within the Markov framework. First, note that for any function f on the state space, $M_i f$ is precisely the expectation of f in player is estimation (cf. Samet (2000)). This is the primal case (Samet (1998a)), in the sense that the expectation is what is usually considered of economic significance and importance, as players choose their actions by comparing the relative expectations of functions.

Second, dual to this is, an invariant probability measure ν with respect to the Markov chain M_i is precisely a prior of player i. A common prior is a probability measure that is simultaneously invariant with respect to all $\{M_i\}_{i \in I}$.

A sequence $s = (i_1, i_2, \ldots)$ of elements of I is called an I-sequence if for each player j, $i_k = j$ for infinitely many ks. The *iterated expectation* of a random variable f with respect to the I-sequence s is the sequence of random variables $\{M_{i_k} \ldots M_{i_1} f\}_{k=1}^{\infty}$.

Given the identification of $E_i f$ with $M_i f$, we can write, given a permutation σ of I,

$$M_\sigma := E_\sigma := t_\sigma := E_{\sigma_1} \ldots E_{\sigma_n} = M_{\sigma_1} \ldots M_{\sigma_n} = t_{\sigma_1} \ldots t_{\sigma_n}$$

and term this a *permutation chain*.

The *iterated expectation of f with respect to σ* is the sequence $\{E^k_{\sigma f}\}_{k=1}^{\infty}$, and the *Cesàro iterated expectation of f with respect to σ* is $\{E^{(k)}_{\sigma f}\}_{k=1}^{\infty}$. The iterated expectation of f with respect to σ is the iterated expectation of f with respect to the I-sequence

$$\sigma_1, \ldots \sigma_n, \sigma_1, \ldots, \sigma_n, \ldots$$

as defined above.

It should be noted here that the results in this paper do not extend all the results of Samet (1998a) to compact metric spaces. To be precise, the claims of that paper, in the finite type space context, show that the existence of a common prior implies that for each random variable f it is common knowledge in each state that *all* the iterated expectations of f, with respect to all I-sequences s, converge to the same limit. The claims of this paper show that, in the context of a compact e.m.p. type space (as defined in the next section), the existence of a common prior implies that for each random variable f it is common knowledge in each state that the Cesàro iterated expectations of f *with respect to each permutation* converge to the same limit, but not with respect to all I-sequences.

3.2. **Everywhere Mutually Positive Type Space.**

A type space $< I, \Omega, \mathcal{F}, (\Pi_i, t_i)_{i \in I} >$ with a topology τ over Ω will be termed *everywhere mutually positive (e.m.p.)* if it satisfies the conditions:

(1) For each state $\omega \in \Omega$ there is an event $A(\omega) \ni \omega$ such that, for all $i \in I$, $A(\omega) \subseteq \Pi_i(\omega)$ and $t_i(A(\omega)|\omega) > 0$
(2) The correspondence $\omega \mapsto t_i(\cdot|\omega)$ is continuous with respect to the topology τ and the weak topology on $\Delta(\Omega)$

Condition (1) can be paraphrased by 'at each state ω, there is common knowledge of an event $A(\omega)$ to which all players ascribe non-zero probability'. Note that when Ω is finite and Π is positive, meaning that $t_i(\omega|\omega) > 0$ for all i and all ω, the corresponding type space trivially satisfies the conditions of being everywhere mutually positive, because the condition $\forall\omega\forall i\ t_i(\omega|\omega) > 0$ is equivalent to stating that at each state ω there is common knowledge that the event $\{\omega\}$ itself has non-zero probability. Also note that from previous definitions it follows that

$$\int_\Omega f(\widehat{\omega})dt_i(\widehat{\omega}|\omega)$$

is always continuous in ω for every $f \in C(\Omega)$.

If in addition to the above conditions, (Ω, τ) is compact metric space, the type space $< I, \Omega, \tau, \mathcal{F}, (\Pi_i, t_i)_{i\in I} >$ will be called a *compact e.m.p. space* for short. Nearly all the results in this paper will henceforth assume a compact e.m.p. type space. For notational ease, $< I, \Omega, \tau, \mathcal{F}, (\Pi_i, t_i)_{i\in I} >$ will be written simply as (Ω, τ).

4. Common Priors and Compact E.M.P. Type Spaces

Given any $Q \in \Pi$, the restriction of M_i to Q, for any player i, will be written as M_i^Q. Given a permutation σ in $Sym(I)$, the restriction of M_σ to Q is similarly denoted by M_σ^Q.

Lemma 1. *Given a type space $< I, \Omega, \mathcal{F}, (\Pi_i, t_i)_{i\in I} >$ satisfying property (1) of e.m.p. type spaces, for any permutation σ of I and player i, for any arbitrary pair of states $\omega, \widehat{\omega} \in \Pi_{\sigma(i)}(\omega)$, there is an event $A(\omega)$ such that $t_\sigma(A(\omega)|\widehat{\omega}) > 0$.*

Proof. Let $\omega, \widehat{\omega} \in \Pi_{\sigma(i)}(\omega)$. By property (1) of e.m.p. type spaces, there exist events $A(\omega)$ and $A(\widehat{\omega})$, both of them in every players partition that contains ω, such that for for $1 < j \leq n$, $t_{\sigma(j)}(A(\widehat{\omega})|\widehat{\omega}) > 0$; for $1 \leq k < i$, $t_{\sigma(k)}(A(\omega)|\omega) > 0$; and because $t_{\sigma(i)}(A(\widehat{\omega})|\widehat{\omega}) > 0$ and $t_{\sigma(i)}(A(\omega)|\widehat{\omega}) = t_{\sigma(i)}(A(\widehat{\omega})|\widehat{\omega})$ (as $\omega, \widehat{\omega} \in \Pi_{\sigma(i)}(\omega)$), it follows that $t_{\sigma(i)}(A(\omega)|\widehat{\omega}) > 0$.

We now unravel the recursive definition of $t_{\sigma(1)} \dots t_{\sigma(n)}$. For $i + 1 < j \leq n$, suppose that $t_{\sigma(j)} \dots t_{\sigma(n)}(A(\widehat{\omega})|\widehat{\omega}) > 0$ (which is certainly true when $j = n$). Then

$$t_{\sigma(j-1)}t_{\sigma(j)} \dots t_{\sigma(n)}(A(\widehat{\omega})|\widehat{\omega}) = \int_\Omega t_{\sigma(j-1)}(A(\widehat{\omega})|\omega')d(t_{\sigma(j)} \dots t_{\sigma(n)})(\omega'|\widehat{\omega})$$

But the facts that $t_{\sigma(j)} \dots t_{\sigma(n)}(A(\widehat{\omega})|\widehat{\omega}) > 0$, that $A(\widehat{\omega}) \subseteq \Pi_{\sigma(j-1)}(\widehat{\omega})$, and that $t_{\sigma(j-1)}(A(\widehat{\omega})|\omega') = t_{\sigma(j-1)}(A(\widehat{\omega})|\widehat{\omega}) > 0$ for all $\omega' \in \Pi_{\sigma(i)}(\widehat{\omega})$, taken all together, imply that $t_{\sigma(j-1)} \dots t_{\sigma(n)}(A(\widehat{\omega})|\widehat{\omega}) > 0$.

Similar reasoning can be applied at the transition point from $t_{\sigma(i+1)} \dots t_{\sigma(n)}$ to $t_{\sigma(i)} \dots t_{\sigma(n)}$, and the transition point from $t_{\sigma(k)} \dots t_{\sigma(n)}$ to $t_{\sigma(k-1)} \dots t_{\sigma(n)}$ for $1 \leq k < i$, to conclude that $t_\sigma(A(\widehat{\omega})|\omega) > 0$. ∎

Proposition 1. *For any permutation σ of I, and for any element Q of the meet of a compact e.m.p. type space (Ω, τ), M_σ^Q has a unique invariant probability measure π_σ^Q.*

Proof. By the assumed properties of an e.m.p. type space, $M_{\sigma(i)}^Q$, for any i, satisfies the weak Feller property. The weak Feller property of M_σ^Q follows readily from the concatenation formation via $M_{\sigma(1)}^Q \dots M_{\sigma(n)}^Q$. The compactness of the metric topology τ then guarantees the existence of at least one invariant probability measure for M_σ^Q, π_σ^Q, by application of a theorem cited in Section 2.1.

Next, select an arbitrary event $E \subseteq Q$, such that $\pi_\sigma^Q(E) > 0$. By definition of a prior, we can readily select a state $\omega' \in E$ such that $t_{\sigma(1)}(E|\Pi_{\sigma(1)}(\omega')) > 0$ (otherwise there would be a contradiction to the assumption that $\pi_\sigma^Q(E) > 0$).

Let $\omega \in Q$ be selected arbitrarily. Due to the fact that ω and ω' share the same element of the meet, ω' is reachable from ω. This means that there exists a sequence $\{\omega = \omega_0, \omega_1, \omega_2, \ldots, \omega_m = \omega'\}$ such that for each $k \in \{0, 1, \ldots, m-1\}$, there exists a player i_k such that $\Pi_{i_k}(\omega_k) = \Pi_{i_k}(\omega_{k+1})$.

We can now define the following iterative process: by definition, there is a player i_0 such that $\Pi_{i_0}(\omega_0) = \Pi_{i_0}(\omega_1)$. At step 0 of the iterative process, we conclude from the lemma that $t_\sigma(A(\omega_1)|\omega_0) > 0$. At step 1, there is a player i_1 such that $\Pi_{i_1}(\omega_1) = \Pi_{i_1}(\omega_2)$, hence (again by applying the lemma) from $t_\sigma(A(\omega_2)|\omega_1) > 0$ and step 0, we arrive (from the definition of t_σ^2) at $t_\sigma^2(A(\omega_2)|\omega_0) > 0$.

At step j, there is a player i_j such that $\Pi_{i_j}(\omega_j) = \Pi_{i_j}(\omega_{j+1})$, hence from $t_\sigma(A(\omega_{j+1})|\omega_j) > 0$ and step $j-1$, we derive (from the definition of t_σ^j) the fact that $t_\sigma^{j+1}(A(\omega_{j+1})|\omega_0) > 0$.

At the end of the process, the conclusion is $t_\sigma^m(A(\omega_m)|\omega_0) > 0$. Finally, by a slight tweaking of the proof of Lemma 1, from the fact that $t_{\sigma(1)}(E|\omega' = o_m) > 0$ and that for $k \in I$, $t_{\sigma(k)}(A(\omega_m)|\omega_m))$, we can show that $t_\sigma(E|\omega') > 0$, so that $t_\sigma^{m+1}(E|\omega_0) > 0$. We thus conclude that M_σ^Q is π_σ^Q-irreducible, hence from the uniqueness of invariant probability measure theorem, π_σ^Q is unique. ∎

Proposition 2. *For a compact e.m.p. type space Ω, the following are equivalent.*

(1) *π is a common prior on Ω*
(2) *π is an invariant probability measure of M_i for each $i \in I$*
(3) *π is an invariant probability measure of M_σ for every permutation σ*

Proof. This is the compact-space equivalent to Proposition 5 of Samet (1998a), and the proof is nearly identical.

Almost immediately from the definitions, 1) and 2) are equivalent. That 2) implies 3) is quite readily seen – if $\pi t_i = \pi$ for each player, then one can successively calculate $\pi(t_{\sigma(1)} \ldots t_{\sigma(n)}) = \pi(t_{\sigma(2)} \ldots t_{\sigma(n)}) = \ldots = \pi t_{\sigma(n)} = \pi$, for any permutation σ.

It remains to show that 3) implies 2). Suppose 3), and let π be the invariant probability measure. Thus

$$\pi(t_1 t_2 \ldots t_n) = \pi$$

Multiplying from the right by t_1 gives

$$\pi(t_1 t_2 \ldots t_n t_1) = \pi t_1$$

So πt_1 is an invariant probability measure of $t_2 \ldots t_n t_1$. But by 3), π is an invariant probability measure of $M_2 M_n \ldots M_1$, and by the previous proposition, $M_2 M_n \ldots M_1$ has a unique invariant probability measure. Thus, $\pi M_1 = \pi$, and by entirely similarly arguments $\pi M_i = \pi$ for all i. ∎

Corollary 1. *For each $Q \in \Pi$, there exists at most one common prior on Q.* ∎

5. Permutations, Iterated Expectations and Common Priors

5.1. Main Results.

Proposition 3. *Given a compact e.m.p. type space Ω, for each random variable f on Ω and permutation σ, $\lim_{n\to\infty} E_\sigma^{(n)} f$ exists, and on each element Q in Π it is constant and is equal to $\pi_\sigma^Q f$, π_σ^Q-almost everywhere.*

Proof. This follows from the previous propositions and Birkhoff's Ergodic Theorem, cited in Section 2.1. ∎

Proposition 4. *Given a compact e.m.p. type space Ω, with $\Pi = \Omega$, a common prior π exists iff for each random variable f, the elements of $\{\lim_{n\to\infty} E_\sigma^{(n)} f | \sigma \in Sym(I)\}$ converge π_σ-almost everywhere to the same limit. Moreover, if π is the common prior, then this limit is πf, π-almost everywhere.*

Proof. As above, $\lim_{n\to\infty} E_\sigma^{(n)} f$ is constantly $\pi_\sigma f$, π_σ-almost everywhere, where π_σ is the unique invariant probability measure of M_σ on Ω. Thus, for each f, the limits for all σ are respectively π_σ-a.e. equal to each other iff for each f, $\pi_\sigma f$ are π_σ-a.e. constantly equal to the same value for all σ.

Clearly, if there is a probability measure π such that $\pi_\sigma = \pi$ for all σ, then $\pi_\sigma f$ are all equal to each other. In the other direction, if in particular for each $A \in \mathcal{F}$, $\pi_\sigma \chi_A$ are all equal, then there is a probability measure π such that $\pi_\sigma = \pi$ for all σ. This amounts, given previous propositions, to saying that π is a common prior. ∎

We can summarise these results as follows:

Theorem 1. *Given a compact e.m.p. type space Ω such that $\Pi = \Omega$, for each random variable f and permutation σ of the players, the Cesàro iterated expectation of f with respect to σ converges and the value of its limit is common knowledge. Moreover, there exists a common prior if and only if for each random variable it is common knowledge that all its Cesàro iterated expectations with respect to all permutations converge to the same value.*

5.2. On the use of Cesàro Means.

In Samet (1998a), in the finite state space case, results are stated in terms of iterated expectations, i.e. $\lim_{n\to\infty} E_\sigma^n f$, whereas the results here work with Cesàro limits, i.e. $\lim_{n\to\infty} E_\sigma^{(n)} f$. It may be natural to inquire what is gained and/or lost in this distinction.

At the intuitive level, returning to the story of Eve and Adam in Samet (1998a), we again have the iterated operations of Eve computing the expectation of Adam's expectation of Eve's expectation ... and so forth. But now the sequences we concentrate on, $E_\sigma^n f$, are the running averages of these iterated expectations, rather than the expectations themselves, and the question is whether or not these averages converge to the same value.

From one perspective, Proposition 4 can be regarded as pointing to a 'test' for establishing whether a common prior exists – in words, check if $\lim_{n\to\infty} E_\sigma^{(n)} f$ converges a.e. to the same value for each $\sigma \in Sym(I)$ and each random variable f. But because the operation of taking Cesàro means preserves convergent sequences and their limits – i.e. if $\lim_{n\to\infty} E^n_\sigma f = a$ then certainly $\lim_{n\to\infty} E^{(n)}_\sigma f = a$ – in one direction it suffices to replace the Cesàro iterated expectation $\lim_{n\to\infty} E^{(n)}_\sigma f$ with the simple iterated expectation $\lim_{n\to\infty} E^n_\sigma f$, and from this point of view we have an 'exact' extension of the finite Samet result. However, if one has identified a random variable f and $\sigma \in Sym(I)$ such that $\lim_{n\to\infty} E^n_\sigma f$ diverges, that is not sufficient to conclude, in the infinite state space case, that there is no common prior, because in that case one needs to check in addition whether $\lim_{n\to\infty} E^{(n)}_\sigma f$ also diverges.

Similarly, Proposition 4 asserts that if there is a common prior, then for each random variable f and each $\sigma \in Sym(I)$, $\lim_{n\to\infty} E^{(n)}_\sigma f$ converges a.e. to the same value – but from this it cannot be concluded that the same can be said of the iterated expectation $\lim_{n\to\infty} E^n_\sigma f$, because it is possible for the latter to diverge when the Cesàro sums converge.

Never the less, this result may still be useful for certain applications. To take one example, consider the model of utilitarian preference aggregation under incomplete information presented in Nehring (2004), in which social preferences amongst a set of agents I is calculated as $E_\mu(\sum_{i\in I} U^f_i)$, where the random variable U^f_i is agent i's utility derived from 'social act' f (represented as a random variable over a finite set of states Ω) and E_μ denotes the expectation with respect to a common prior μ. Without going into details here, the key point of the model in that paper of interest here is a result that asserts that act f is 'socially preferred' to g, written $f \succ_I g$ if and only if $E_\mu(\sum_{i\in I} U^f_i) > E_\mu(\sum_{i\in I} U^g_i)$, where the common prior μ is assumed to exist. The common prior therefore plays the role of a 'group' valuation. Nehring (2004) seeks to characterise this group valuation in terms of the beliefs of the individual agents, and appeals to Samet's result to do so.

With the results of this paper, extending Nehring's model to an infinite compact e.m.p. type space Ω with a common prior μ, it can be shown that the group valuation may be related to the (potentially finite iterations of) beliefs of individual agents. In Nehring (2004), $f \succ_I g$ if and only if for some finite sequence $\{i_1, \ldots, i_k\}$, it is common knowledge that $E_{i_k \ldots i_1}(\sum_{i\in I} U^f_i) > E_{i_k \ldots i_1}(\sum_{i\in I} U^g_i)$. Using the Cesàro mean approach in the infinite compact case, we can recapitulate this result in 'if' direction. In place of Nehring's supposition of common knowledge that $E_{i_k \ldots i_1}(\sum_{i\in I} U^f_i) > E_{i_k \ldots i_1}(\sum_{i\in I} U^g_i)$ for some finite sequence, suppose (in the infinite compact e.m.p. case) that for some finite k and permutation $\sigma \in Sym(I)$, there is common knowledge amongst the agents in I that $E^k_\sigma(\sum_{i\in I} U^f_i) > E^k_\sigma(\sum_{i\in I} U^g_i) + \epsilon$, where $\epsilon > 0$ is arbitrary. This last inequality may be rephrased as

$$E^k_\sigma(\sum_{i\in I} U^f_i - \sum_{i\in I} U^g_i)(\omega) > \epsilon.$$

It then follows from the definitions that there is common knowledge amongst the agents that $E^n_\sigma(\sum_{i\in I} U^f_i - \sum_{i\in I} U^g_i) > \epsilon$ for all integers $n > k$. But from this it readily follows that

$$\lim_{n\to\infty} E^{(n)}_\sigma(\sum_{i\in I} U^f_i) > \lim_{n\to\infty} E^{(n)}_\sigma(\sum_{i\in I} U^g_i) + \epsilon,$$

or

$$E_\mu(\sum_{i\in I} U_i^f) > E_\mu(\sum_{i\in I} U_i^g) + \epsilon,$$

since by Proposition 4 and the assumption of the existence of a common prior μ, $E_\mu(\sum_{i\in I} U_i^f)$ is given by $\lim_{n\to\infty} E_\sigma^{(n)}(\sum_{i\in I} U_i^f)$ for each permutation σ. We can then conclude that $f \succ_I g$.

6. The Necessity of Compactness

In this section we demonstrate that the conclusion of Proposition 1 above, namely that each M_σ has an invariant probability measure, does not hold when the assumption of compactness is relaxed. As the proofs of the propositions subsequent to Proposition 1 are ultimately dependent on the conclusion of Proposition 1, they cannot be conducted without compactness.

The example we use here is a variant of the famous 'electronic mail' games. Consider two individuals, Anna and Ben, and a denumerable state space $\Omega = \{1,2,3,\dots\}$. Anna's partition is $\{\{1\},\{2,3\},\{4,5\},\dots\}$ and Ben's partition is $\{\{1,2\},\{3,4\},\{5,6\},\dots\}$. The meet in this case is all of Ω.

Ben's beliefs are always equal probabilities to the two states in each of his partition members. Anna's beliefs are also equal probabilities to the two states in her partition members, save for the probability 1 which is necessary for the single partition containing one state.

We can depict the beliefs of each of the two players in the form of infinite matrices:

$$\text{Anna} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & \dots \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & \dots \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & \dots \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & \dots \\ \cdot & \cdot & \cdot & \cdot & \cdot & \frac{1}{2} & \dots \\ \cdot & \cdot & \cdot & \cdot & \cdot & \frac{1}{2} & \dots \\ \cdot & \cdot & \cdot & \cdot & \cdot & 0 & \dots \end{bmatrix}$$

$$\text{Ben} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 & \dots \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & \dots \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & \dots \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & \dots \\ \cdot & \cdot & \cdot & \cdot & 0 & 0 & \dots \\ \cdot & \cdot & \cdot & \cdot & 0 & 0 & \dots \end{bmatrix}$$

Form the permutation matrix $M_\sigma := \text{Ben} \times \text{Anna}$

$$\text{Ben } \times \text{Anna } = \begin{bmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} & 0 & 0 & \dots \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} & 0 & 0 & \dots \\ 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \dots \\ 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \dots \\ 0 & 0 & 0 & \frac{1}{4} & \frac{1}{4} & \dots \\ 0 & 0 & 0 & \frac{1}{4} & \frac{1}{4} & \dots \\ \cdot & \cdot & \cdot & 0 & 0 & \dots \\ \cdot & \cdot & \cdot & 0 & 0 & \dots \end{bmatrix}$$

and note that it forms the following pattern: letting O stand for the set of positive odd integers, and regarding M_σ as a mapping on the domain $\mathbb{N} \times \mathbb{N}$, we start with $M_\sigma(1,1) = \frac{1}{2}$, $M_\sigma(2,1) = \frac{1}{2}$, and for each $j \in O$, $\frac{1}{4} = M_\sigma(j, j+1) = M_\sigma(j+1, j+1) = M_\sigma(j+2, j+1) = M_\sigma(j+3, j+1) = M_\sigma(j, j+2) = M_\sigma(j+1, j+2) = M_\sigma(j+2, j+2) = M_\sigma(j+3, j+2)$. For all other values of k and l, $M_\sigma(k,l) = 0$.

Suppose now that there is an invariant probability distribution π with respect to M_σ. Let $\pi(1) = \alpha$. Then by the definition of invariant probability, it must also be the case that $\pi(2) = \alpha$, because $\pi(1) = \frac{1}{2}(\pi(1) + \pi(2))$. Similar reasoning leads to the conclusion that $\pi(3) = \alpha$, $\pi(4) = \alpha$, $\dots, \pi(k) = \alpha, \dots$.

Now, $\alpha \in [0,1]$, so that either $\sum_{k=1}^{\infty} \pi(k) = 0$, or $\sum_{k=1}^{\infty} \pi(k) = \infty$. In either case, π cannot be a normalised probability.

7. Conclusion

As stated in the introduction, in this paper we have extended most of the results of Samet (1998a) to compact e.m.p. type spaces and shown that compactness is necessary for the proofs of the results. As noted in Section 3, whether our results on compact e.m.p. type spaces also apply with respect to all I-sequences remains an open question.

References

Aumann, R. J. (1976), Agreeing to disagree, *Ann. Statist.*, 4(6), 1236–1239.

Aumann, R. J. (1987), Correlated Equilibrium as an Expression of Bayesian Rationality, *Econometrica*, 55, 1–18.

Aumann, R. J. (1998), Common Priors: A Reply to Gul, *Econometrica*, 66, 929–938.

Feinberg, Y.(2000), Characterizing Common Priors in the Form of Posteriors, *Journal of Economic Theory*, 91(2), 127-179.

Gul, F. (1998), A Comment on Aumann's Bayesian View, *Econometrica*, 66, 923–927.

Harsányi , J. C. (1967-8), Games with Incomplete Information Played by Bayesian Players , *Management Science*, 14, 159–182, 320–334, 486–502.

Heifetz, A. (2006), The Positive Foundation of the Common Prior Assumption, *Games and Economic Behavior*, 56, 105-120.

Hernández-Lerma, O., and Lasserre, J. B. (2003), *Markov Chains and Invariant Probabilities*. Birkhauser-Verlag, Basel.

Milgrom, P. and Stokey, N. (1982), Information, Trade and Common Knowledge, *Journal of Economic Theory*, 26, 17-27.

Morris, S. (1994), Trade with Heterogeneous Prior Beliefs and Asymmetric Information, *Econometrica*, 62, 1327-1347.

Nehring, K. (2004), The Veil of Public Ignorance, *Journal of Economic Theory*, 119, 247-270.

Samet, D. (1998), Iterated expectations and common priors, *Games and Economic Behavior*, 24, 131-141.

Samet, D. (1998), Common Priors and the Separation of Convex Sets, *Games and Economic Behavior*, 24, 173–175.

Samet, D. (2000), Quantified Beliefs and Believed Quantities, *Journal of Economic Theory*, 95, 169-185.

Sebenius, J., and Geanakoplos, J. (1983), Dont bet on it: Contingent agreements with asymmetric information, *Journal of the American Statistical Association*, 78, 424-426.

The Department of Mathematics and The Centre for the Study of Rationality, The Hebrew University of Jerusalem.