A MEASURABLE "MEASURABLE CHOICE" THEOREM

JEAN-FRANÇOIS MERTENS
CORE, Université Catholique de Louvain
Louvain-la-Neuve, Belgium

Abstract. We prove here a measurable version of the measurable choice theorem (a.o., basically of Lyapunov's theorem), in the sense that the measurable selection (the set) can be chosen in a measurable way as a function of the underlying probability measure, of the integral (measure) desired, and of the correspondence itself.

1. Introduction

We prove here a measurable version of the measurable choice theorem (among others, basically of Lyapunov's theorem), in the sense that the measurable selection (the set) can be chosen in a measurable way as a function of the underlying probability measure, of the desired integral (measure), and of the correspondence itself. We also show that the integral of the correspondence is itself a measurable function of those parameters.

More precisely, those quantities are parametrized in a measurable way by an auxiliary measurable space, and it is shown that a selection exists which depends measurably on this auxiliary parameter – i.e., which is measurable on the product. Such a result is basic in a number of optimal control-type problems, like the existence of equilibria in discounted stochastic games [7]. We have to assume that the underlying measurable space is separable, and we deal only with compact-valued correspondences, where measurability is understood as the measurability of the corresponding mapping to compact subsets of $\mathbb{R}^n$. Such assumptions were sufficient for the above-mentioned applications. Note that, by Corollary 8'.1 of the Appendix, this measurability assumption is essentially necessary to insure the existence of measurable selections. Besides, by the results of **A.2** of the Appendix, just the measurability of the graph already implies the existence of selections which are bianalytically measurable on the product (hence measurable if both measurable spaces are Blackwell spaces, by the first separation theorem for analytic sets). Those results of the first paragraphs of the Appendix are also used, in the above-mentioned application, to reduce the general problem to our present separable setup. Finally, some results in the Appendix are presented in somewhat more generality than required here, because of other uses in the above-mentioned application.

In the theorem below, (2) essentially asserts that the integral we define coincides, for each fixed $e \in E$, with the usual integral; (1b) asserts the measurability of this integral as a function of the parameters and (3) yields the measurable selection.

The "propositions" in the proof refer to the Appendix.

2. Theorem

Let $P(d\omega \mid e)$ be a bounded $\mathbb{R}^k$-valued kernel from $(E, \mathcal{E})$ to $(\Omega, \mathcal{A})$, two measurable spaces (i.e., $\forall e \in E, P(\cdot \mid e)$ is a bounded $\mathbb{R}^k$-valued measure on $(\Omega, \mathcal{A})$, and $\forall A \in \mathcal{A}, P(A \mid \cdot)$ is $\mathcal{E}$-measurable). Let N be a measurable map from $(E \times \Omega, \mathcal{E} \otimes \mathcal{A})$ to $\mathcal{K}^*(\mathbb{R}^\ell)$, which is P-integrable in the sense that for any measurable selection f from N, $f(\omega, e)$ is $P(d\omega \mid e)$-absolutely integrable for any $e \in E$. Define $\int N\ d\ P$ as the map from E to subsets of $\mathbb{R}^{k.\ell}$ (the tensor product of $\mathbb{R}^k$ and $\mathbb{R}^\ell$) defined by $\int N\ d\ P : e \to \{\int f(\omega, e) d\ P(d\omega \mid e) \mid f$ is an $\mathcal{E} \otimes \mathcal{A}$-measurable selection from $N\}$, and denote its graph by $(F, \mathcal{F}) \subseteq (E \times \mathbb{R}^{k.\ell}, \mathcal{E} \otimes$ Borel sets). Assume that $\mathcal{A}$ is separable, except for (1a) and (2). Then

1. *(a) $(\int N\ d\ P)(e) \in \mathcal{K}^*(\mathbb{R}^{k\ell}) \quad \forall e \in E$.*
   (b) $(\int N dP)$ is an $\mathcal{E}$-measurable map to $\mathcal{K}^*(R^{k\ell})$, and F is measurable in $E \times \mathbb{R}^{k\ell}$.
2. *Fix $e \in E$, and let $\mu(\cdot) = P(\cdot \mid e)$: if f is a μ-measurable selection from the convex hull of $N(e, \cdot)$, such that f belongs to $N(e, \cdot)$ μ-a.e. on the atoms of μ, then $\int f d\mu \in (\int N\ d\ P)(e)$.*
3. *There exists a measurable, $\mathbb{R}^\ell$-valued function f on $(F \times \Omega, \mathcal{F} \otimes \mathcal{A})$ such that $f(e, x, \omega) \in N(e, \omega)$ and $x = \int f(e, x, \omega) P(d\omega \mid e)$.*

For the notation $\mathcal{K}^*(X)$, cf. Appendix **A.2.**

The proof follows more or less classical lines (e.g., [1], [3], [5], [6]), taking care at each step to do things in a measurable way.

Proof. Let us first show that $(\int N\ d\ P)(e_0) = \{\int f(\omega) P(d\omega \mid e_0) \mid f$ is an $\mathcal{A}$-measurable selection from $N(e_0, \omega)\}$. Clearly, the second set includes the first. Assume thus that f is an $\mathcal{A}$-measurable selection from $N(e_0, \omega)$ for some $e_0 \in E$. Then $\{e, \omega) \mid f(\omega) \in N(e, \omega)\} \in \mathcal{E} \otimes \mathcal{A}$, as the inverse image by a measurable map of the closed set $\{(x,K) \in \mathbb{R}^\ell \times \mathcal{K}(\mathbb{R}^\ell) \mid x \in K\}$. Let $\tilde{f}(e,\omega) = f(\omega)$ on this set, and be a fixed measurable selection from N (Proposition 7.3) on the complement: $\tilde{f}$ is a measurable selection from N, with $\tilde{f}(e_0,\omega) = f(\omega)$. Hence our equality. It follows that, for (1a) and (1b), we can assume that E is a singleton.

There is clearly no loss in assuming that the function f in (2) is $\mathcal{A}$-measurable: indeed the μ-equivalence class of f contains an $\mathcal{A}$-measurable element $\tilde{f}$, and $\{\omega \mid \tilde{f}(\omega) \in \hat{N}(\omega)\} \in \mathcal{A}$, being equal to $\{\omega \mid \{\tilde{f}(\omega)\} \cap \hat{N}(\omega) \neq \emptyset\}$ (use Proposition 6.1, Proposition 6.5, and Proposition 6.3 – $\hat{N}(\omega)$ denotes the convex hull of $N(\omega)$); therefore by Proposition 7.3, we can construct an $\mathcal{A}$-measurable element $\bar{f}$ of the equivalence class of f which belongs everywhere to $\hat{N}$. Note also that, in (1b), the measurability of F is a consequence of the measurability of $\int N dP$ (Proposition 9.2) – so we will concern ourselves only with the former.

Let us first show that we can assume that $\tilde{Q} = \Sigma_i |P_i|$ is a nonnegative kernel, where P_i denotes the i-th coordinate of P, and $|P_i| = P_i^+ + P_i^-$ (Jordan decomposition). Thus we have to show $\mathcal{E}$-measurability – say of P_i^+ – except in cases (1a) and (2), where only one $e \in E$ is involved, so that one can assume $E = \{e\}$. Thus we can assume $\mathcal{A}$ to be separable: $\mathcal{E}$-measurability follows then because $\mu^+(A) = \sup\{\mu(A \cap B) \mid B$ varies over a countable algebra generating $\mathcal{A}\}$ for any measure μ and $A \in \mathcal{A}$.

Thus $\{e \mid \tilde{Q}(\Omega \mid e) = 0\} \in \mathcal{E}$; (1) and (2) are trivial on this set; for (3), define $f(e,x,\omega)$ on this set to be a fixed $\mathcal{A}$-measurable selection from N (Proposition 7.3,). Thus we can assume $\tilde{Q}(\Omega \mid e) > 0$. Let then $Q(A \mid e) = \tilde{Q}(A \mid e)/\tilde{Q}(\Omega \mid e)$: Q is a transition probability. Let $f_i(\omega, e)$ denote a Radon-Nikodym derivative of $P_i(d\omega \mid e)$ with respect to $Q(d\omega \mid e)$. Except in cases (1a) and (2), we want to show that we can select f_i to be $\mathcal{E} \otimes \mathcal{A}$-measurable. We use Doob's old trick [4] in this generalized situation: since $\mathcal{A}$ is separable, let $\mathcal{A}_n$ denote an increasing sequence of finite sub-σ-fields of $\mathcal{A}$ such that $\cup_n \mathcal{A}_n$ generates $\mathcal{A}$. Denote the elements of the partition corresponding to $\mathcal{A}_n$ by $(A_1^n, A_2^n, \dots, A_{j_n}^n)$. Let $g_i^n(e,\omega) = \Sigma_{j=1}^{j_n} \mathrm{I}_{A_j^n}(\omega) r\big(P_i(A_j^n|e), Q(A_j^n|e)\big)$, where $r(x,y) = x/y$ when $y \neq 0$, and $r(x,y) = 0$ when $y = 0$. As a composition of measurable functions, g_i^n is $\mathcal{E} \otimes \mathcal{A}$-measurable. For each e and i, the $g_i^n(e,\cdot)$ form a bounded martingale under $Q(d\omega \mid e)$, converging $Q(d\omega \mid e)$ – a.s. to $f_i(e,\omega)$ by the martingale convergence theorem. Let thus $g_i(e,\omega) = \lim\ \inf_{n\to\infty} g_i^n(e,\omega)$; g_i is $\mathcal{E} \otimes \mathcal{A}$-measurable, real-valued because $\Sigma \mid g_i(e,\omega)| \leqslant \lim\ \sup_{n\to\infty} \Sigma_i\ |g_n^i(e,\omega)|$ and $\Sigma_i |g_i^n(e,\omega)| \leqslant \max_j r(\tilde{Q}(A_j^n \mid e), Q(A_j^n \mid e)) \leqslant \tilde{Q}(\Omega \mid e)$, and $\forall\ e \in E$, $g(e,\omega)$ is a Radon-Nikodym derivative of $P(d\omega \mid e)$ w.r.t. $Q(d\omega \mid e)$. We can keep all those properties and change $g_1(e,\omega)$ on a null set, such as to have in addition $\Sigma_i\ |g_i(e,\omega)| = \tilde{Q}(\Omega \mid e) > 0$ everywhere. Now let $\tilde{N}(e,\omega) = g(e,\omega) \star N(e,\omega)$ (the tensor product, in $\mathbb{R}^{k\ell}$) : $\tilde{N}$ is an $\mathcal{E} \otimes \mathcal{A}$-measurable map to $\mathbb{R}^{k\ell}$ (since the tensor product, from $\mathbb{R}^k \times \mathcal{K}^*(\mathbb{R}^\ell)$ to $\mathcal{K}^*(\mathbb{R}^{k\ell})$, is continuous) and Q is a transition probability from $(E, \mathcal{E})$ to $(\Omega, \mathcal{A})$. If f is a measurable selection from N, clearly $g \star f$ is a measurable selection from $\tilde{N}$, and $\int g \star f dQ = \int f dP$. Conversely, if $h(e,\omega)$ is a measurable selection from $\tilde{N}$, $h(e,\omega) \in \tilde{N}(e,\omega)$ means $h_{i,j}(e,\omega) = g_i(e,\omega) f_j(e,\omega)$, with $f(e,\omega) \in N(e,\omega), (f = (f_j))$. Since $\Sigma_i \, |g_i(e,\omega)| > 0$, this determines f uniquely, and this unique f is measurable by the measurability of g and h. Hence $\int h_{ij}(e,\omega) dQ(\omega \mid e) = \int g_i(e,\omega) f_j(e,\omega) dQ(\omega \mid e) = \int f_j(e,\omega) P_i(d\omega \mid e)$ belongs to $(\int N dP)(e)$, since f is a measurable selection from N. Thus $\int \tilde{N} dQ = \int N dP$; in particular, $\tilde{N}$ is Q-integrable. It immediately follows then that (1) and (2) will hold for (N, P) if they hold for $(\tilde{N}, Q)$; and for (3), assume that $\tilde{f}(e,x,\omega) \in \tilde{N}(e,\omega)$, is measurable, and such that $x = \int \tilde{f}(e,x,\omega) Q(d\omega \mid e)$. As seen above, this implies that $\tilde{f}_{i,j}$ can be written uniquely as $g_i(e,\omega) f_j(e,x,\omega)$, with $f(e,x,\omega) \in N(e,\omega)$ and f will be measurable by the same argument as above. We then obtain that $x = \int f(e,x,\omega) P(d\omega \mid e)$: (3) also follows. Thus it is sufficient to prove the theorem in case P is a transition probability from $(E, \mathcal{E})$ to $(\Omega, \mathcal{A})$.

We now need a "well-known" lemma of independent interest:

Lemma 1 *If f_n is a pointwise bounded sequence of vector-valued measurable functions on $(\Omega, \mathcal{A}, \mu)$, and f is a weak limit of f_n (in L_1), then (for an appropriate version of f) f is $\mathcal{A}$-measurable, $f(\omega) \in (\mathrm{Co}\,\mathrm{Lim} f_i)(\omega)$ everywhere, and $f(\omega) \in (\mathrm{Lim} f_i)(\omega)$ on the atoms of μ.*

Proof. Clearly $f(\omega) \in (\mathrm{Lim} f_i)(\omega)$ μ-a.e. on the atoms, by definition of the weak topology. Select for f an $\mathcal{A}$-measurable element of its equivalence class. f belongs to the weak closure of the convex hull of $\{f_i \mid i \geqslant n\}$ for all n – by the Hahn-Banach theorem, this is the strong closure; i.e., for all n there exists $\alpha_i^n \geqslant 0$ – where $\Sigma_i \alpha_i^n = 1, \alpha_i^n = 0$ for $i \leqslant n$ and $\alpha_i^n > 0$ only for finitely many i's – such that $\|\Sigma_i \alpha_i^n f_i - f\|_1 \leqslant 2^{-n}$. Since 2^{-n} is summable, this implies that $\Sigma_i \alpha_i^n f_i$ converges to f μ-a.e.: thus, μ-a.e., $f(\omega)$ belongs, for all n, to the closed convex hull of $\{f^j(\omega) \mid j \geqslant n\}$. Denote by C_n the (compact, nonempty, by pointwise boundedness) closure of $\{f^j(\omega) \mid j \geqslant n\}$, and let $C = \cap_k C_k : C = (\mathrm{Lim} f_i)(\omega)$. By Proposition 10.3, the convex hulls $\hat{C}_n$ are compact and converge in $\mathcal{K}^*(\mathbb{R}^\ell)$ to $\hat{C}$. Thus $f(\omega) \in \hat{C}_n$ for all n implies that $f(\omega) \in \hat{C}$: μ-a.e., $f(\omega) \in \mathrm{Co}(\mathrm{Lim} f_i)(\omega)$.

By Proposition 10.1, $\omega \mapsto (\mathrm{Lim} f_i)(\omega)$ is measurable, hence, by Proposition 10.3, $\omega \mapsto \mathrm{Co}(\mathrm{Lim} f_i)(\omega)$ is also. Thus $\{\omega \mid f(\omega) \notin \mathrm{Co}(\mathrm{Lim} f_i)(\omega)\} \in \mathcal{A}$, e.g., by Proposition 6.1, and Proposition 6.5. By Proposition 7.3, f can be modified on this null set belonging to $\mathcal{A}$, so as to have $f(\omega) \in \mathrm{Co}(\mathrm{Lim} f_i)(\omega)$ everywhere, and still f $\mathcal{A}$-measurable. This proves Lemma 1.

Lemma 1 implies that (1a) will follow from (2) since any integrably bounded sequence has weak limit points. For (2) it is sufficient to consider a nonatomic probability P on $(\Omega, \mathcal{A})$, since on the atoms $f(\omega)$ already belongs to $N(\omega)$. (We omit the argument e which is fixed.) By Proposition 10.2, there exist measurable functions $\lambda_1(\omega) \ldots \lambda_k(\omega)$ and $g^1(\omega) \ldots g^k(\omega)$ such that $\lambda_i(\omega) \geqslant 0, \Sigma_i \lambda_i(\omega) = 1, g^i(\omega) \in N(\omega)$, $\Sigma \lambda^i(\omega) g^i(\omega) = f(\omega)$: let $\mu_{ij}(A) = \int_A g^i_j(\omega dP(\omega)$. If we can find another measurable function $\tilde{\lambda}(\omega)$, whose values are extreme points of the simplex, and such that $\int \tilde{\lambda}^i(\omega)\mu^i_j(d\omega)$ $= \int \lambda^i(\omega)\mu^i_j(d\omega)$, we have finished. We can rank all positive and negative parts of all measures μ^i_j into one long vector $\nu_\ell(d\omega)$ – and further normalize those ν_ℓ if so desired – and it will then be sufficient to have $\int \tilde{\lambda}^i(\omega) d\nu_\ell(\omega) = \int \lambda^i(\omega) d\nu_\ell(\omega)$ for all i and ℓ. Since ν is a nonatomic probability vector, the existence of $\tilde{\lambda}$ is guaranteed by the Dvoretzky-Wald-Wolfowitz theorem – or by the construction below, which implies a "measurable version" of this theorem.

In the remaining cases [(1b) and (3)], $(\Omega, \mathcal{A})$ is separable. Let us show that we can then also assume that $(E, \mathcal{E})$ is separable, and that $N(e, \omega)$ is independent of $e \in E$, i.e., is a measurable map $N(\omega)$ on $(\Omega, \mathcal{A})$. If $A_i \in \mathcal{A}$ enumerates a Boolean algebra that generates the σ-algebra $\mathcal{A}$, the sub-σ-field $\mathcal{E}_0$ of $\mathcal{E}$ generated by the $P(A_i \mid e)$ is separable, and P is a kernel from $(E, \mathcal{E}_0)$ to $(\Omega, \mathcal{A})$. Since the Borel σ-field on $\mathcal{K}^*(\mathbb{R}^\ell)$ is separable, N is a measurable map for some separable sub σ-field $\mathcal{F}$ of $\mathcal{E} \otimes \mathcal{A}$. Let $F_i \in \mathcal{F}$ be a sequence generating the σ-field $\mathcal{F}$; each $F_i \in \mathcal{E} \otimes \mathcal{A}$ is generated as a measurable set by a sequence of product sets $E^i_n \times A^i_n, E^i_n \in \mathcal{E}, A^i_n \in \mathcal{A}$, by definition of the product σ-field. Denote by $\mathcal{E}_1$ the separable sub-σ-field of $\mathcal{E}$ generated by $\mathcal{E}_0$ and by all $E^i_n : N(e, \omega)$ is $\mathcal{E}_1 \otimes \mathcal{A}$-measurable, and P is a kernel from $(E, \mathcal{E}_1)$ to $(\Omega, \mathcal{A})$. It is then clearly sufficient to prove the theorem with $\mathcal{E}_1$ instead ot $\mathcal{E}$ – i.e., we can assume $\mathcal{E}$ separable. P can also be viewed as a kernel from $(E, \mathcal{E})$ to $(\Omega', \mathcal{A}') = (E \times \Omega, \mathcal{E} \times \mathcal{A})$. Since also N is a measurable map on $(\Omega', \mathcal{A}')$, and since this space is separable, we can replace $(\Omega, \mathcal{A})$ by $(\Omega', \mathcal{A}')$. Our definition of $\int N dP$ is unaffected, and if $f'(e, x, \omega')$ is the measurable map of (3) for this transformed problem, then $f(e, x, \omega) = f'(e, x, (e, \omega))$ will solve the original problem. Thus we can always assume that $(E, \mathcal{E})$ is separable, and that N is a measurable map defined on $(\Omega, \mathcal{A})$.

Having proved (1a) and (2), (1b) is now an immediate corollary of Proposition 8.1. (For 8.1, a weaker sufficient assumption is mentioned in the beginning of the proof, and is immediate to check. But the stronger assumption of the proposition is satisfied too, by Lemma 1 and (2).)

So there remains to prove (3). Let $(A_i) \in \mathcal{A}$ be a sequence that generates A; map $(\Omega, \mathcal{A})$ into $\{0, 1\}^\infty \times \mathcal{K}^*(\mathbb{R}^\ell)$ by the mapping $\varphi : \omega \to$

$[(I_{A^i}(\omega))_{i=1}^{\infty}, N(\omega)]$. Then $\mathcal{A} = \varphi^{-1}(\mathcal{B})$, when $\mathcal{B}$ denotes the Borel σ-field on the image space; so $(\Omega, \mathcal{A})$ can be identified by φ with this image. The identification does not affect $\int N dP$, and any solution f after identification can be composed with φ to yield a solution of the original problem. Hence $(\Omega, \mathcal{A})$ is a subset of $\{0,1\}^\infty \times \mathcal{K}^*(R^\ell)$. Denote by $\bar{\Omega}$ its closure, and let N still denote the (continuous) projection from $\bar{\Omega}$ to $\mathcal{K}^*(\mathbb{R}^\ell)$. For each $e \in E, P(d\omega \mid e)$ is a (regular) Borel measure on $\bar{\Omega}$, for which Ω has outer measure 1. The Borel functions on Ω are the restrictions to Ω of the Borel functions on $\bar{\Omega}$ – this is true by definition for indicator functions of open sets, hence it follows for the indicator functions of all Borel sets, hence of all real-valued Borel step functions, hence for all real-valued Borel functions, hence for $\mathbb{R}^\ell$-valued Borel functions also.

Since $N(\omega)$ is Borel on $\bar{\Omega}$, it has Borel selections on $\bar{\Omega}$ (Proposition 7.3). Therefore the Borel selections of N on Ω are the restrictions to Ω of the Borel selections of N on $\bar{\Omega}$. Since Ω has outer measure 1 under $P(d\omega \mid e)$, any two Borel functions with the same restriction to Ω will have the same integral under $P(d\omega \mid e)$, which is also the integral of their restriction of Ω under the restriction of $P(d\omega \mid e)$ to Ω. This implies first that $P(d\omega \mid e)$ is also a transition probability from $(E, \mathcal{E})$ to $\bar{\Omega}$, and that the integral $\int N dP$ is the same computed over Ω or over $\bar{\Omega}$. Hence if we prove (3) on $\bar{\Omega}$, the restriction of the function f to Ω will prove (3) on Ω.

Thus it is sufficient to prove (3) on $\bar{\Omega}$ which is locally compact with a countable basis as a closed subspace of such a space. To make sure $\bar{\Omega}$ has the power of the continuum, replace it by its disjoint union with $[0,1]$, defining $N(\omega) = \{0\}$ on this additional part – it is clearly sufficient to prove the statement on the enlarged space.

Now $(\Omega, \mathcal{A})$ is Borel-isomorphic to $[0,1]$ with the Borel sets – we will identify them. $(E, \mathcal{E})$ can now also be made identifiable with $[0,1]$: we have seen, when reducing to the case of separable $\mathcal{E}$, that we can identify $(E, \mathcal{E})$ with a set of probability measures on $(\Omega, \mathcal{A})$, endowed with the coarsest σ-field for which the mappings $P \to P(A)$ $(A \in \mathcal{A})$ are measurable (since N is now independent of E). Consider the set M of all probability measures on $(\Omega, \mathcal{A}) = [0,1]$, endowed with the weak*-topology, and the Borel σ-field $\mathcal{B}$. Since $P \to \int f dP$ for f continuous is continuous, it is $\mathcal{B}$-measurable. By taking pointwise limits this thus remains true for bounded Borel f. Conversely, let $\mathcal{G}$ be the coarsest σ-field on M for which $P \to P(A)$ (all A Borel) is measurable. $\mathcal{G}$ is separable since the Borel σ-field is countably generated, and by the above argument $\mathcal{G} \subseteq \mathcal{B}$. For $\mathcal{G}$ the mappings $P \to \int f dP$ are measurable, by the usual argument, for all bounded Borel f – in particular for f continuous. Since those mappings form a family of continuous functions on M that separates points, the Stone-Weierstrass theorem (compactness of M) implies that the closed algebra generated by those functions is the space of all continuous functions on M: any continuous function on M is $\mathcal{G}$-measurable, hence since M is metrizable, any Borel function on M: $\mathcal{G} = \mathcal{B}$. (Using instead of this argument the separability of $\mathcal{G}$ and the first separation theorem would have given the same result for any Souslin space Ω.) Thus we can identify $(E, \mathcal{E})$ with a subspace of $(M, \mathcal{B})$. Consider the functions $\bar{x}_i(\omega) = \max\{x_i \mid x \in N(\omega)\}$: they are measurable, as a composition of a measurable and a continuous function (Proposition 9.1). By Proposition 9.3, there exists a measurable selection f^i of $N(\omega)$, such that $f^i_i(\omega) = \bar{x}_i(\omega)$. Hence the integrability of N implies the integrability of $\bar{x}_i$ – and similarly of $\underline{x}_i(\omega) = \min\{x_i \mid x \in N(\omega)\}$. Hence it implies the integrability of $x(\omega) = \max_i\left(\max(\bar{x}_i(\omega), -\underline{x}_i(\omega))\right)$ - and clearly the integrability of x will conversely imply the integrability of N.

Since x is Borel and nonnegative, $\int x(\omega)dP(\omega)$ is a well-defined Borel function from $(M, \mathcal{B})$ to $[0, +\infty]$. Thus $\tilde{E} = \left\{P \in M \mid \int x dP < \infty\right\}$ is a Borel subset of $(M, \mathcal{B})$, and is the set of P's for which N is integrable.

Therefore $(E, \mathcal{E})$ is identifiable with a subspace of $(\tilde{E}, \mathcal{B})$. Thus it is sufficient to prove the theorem in case $E = \tilde{E}$ – by restricting again at the end the function f obtained to the subset E. Since $\tilde{E}$ is a Borel subset of the compact metric space M, and since it has clearly the power of the continuum (it contains all probabilities with finite support on $[0, 1]$), it too is Borel-isomorphic to $[0, 1]$. Thus, from now on, we assume that both $(E, \mathcal{E})$ and $(\Omega, \mathcal{A})$ are $[0, 1]$ with the Borel sets, and that P is a transition probability.

Note that the graph F of the integral, being a Borel subset of $E \times \mathbb{R}^\ell$, is itself Borel-isomorphic to $[0, 1]$. Further, P can be thought of as being a transition probability from F to Ω, and the projection x from F to $\mathbb{R}^\ell$ is then a Borel selection from the graph of the integral $\left(\int N dP\right)(f)$, $f \in F$. Writing thus E for this F, we can assume furthermore that some Borel selection $x(e)$ from the graph of $\int N dP$ is given, and it will be sufficient to construct some Borel function f on $(E \times \Omega)$ such that $f(e, \omega) \in N(\omega)$ and $\int f(e, \omega)P(d\omega \mid e) = x(e)$. We will freely use in the sequel the equivalence of those two formulations.

Having dispensed with the more trivial pathologies, we can start the core of the proof.

Denote by $F(\omega, e)$ the cumulative distribution function of $P(d\omega \mid e)$: $F(\omega, e) = P\left(\{\omega' \mid \omega' \leqslant \omega\} \mid e\right)$. Since P is a transition probability, F is $\mathcal{E}$-measurable for each fixed ω, and is nondecreasing, right-continuous, with $F \geqslant 0, F(1, e) = 1$. Let $F_n(\omega, e) = F(\omega_n, e)$, with $\omega_n = \min\{k \cdot 2^{-n} \mid k \cdot 2^{-n} \geqslant \omega\}$. F_n is $\mathcal{A} \otimes \mathcal{E}$-measurable, since each $F(k \cdot 2^{-n}, e)$ is $\mathcal{E}$-measurable, and since $\omega \to \omega_n$ is $\mathcal{A}$-measurable. Further, since F is nondecreasing and right-continuous, F_n decreases pointwise to F. Hence F is $\mathcal{E} \otimes \mathcal{A}$-measurable. Similarly, the left-continuous version $F'(\omega, e) = P\left(\{\,\omega' \mid \omega' < \omega\,\} \mid e\right)$ is $\mathcal{E} \otimes$

$\mathcal{A}$-measurable.

Now let $B_n = \{(\omega, e) \mid 1/n \geqslant F(\omega, e) - F'(\omega, e) > 1/(n+1)\}$ and $A_n(e) = \{\omega \mid (\omega, e) \in B_n\}$: B_n is $\mathcal{E} \otimes \mathcal{A}$-measurable, and has compact sections ($\#A_n(e) \leqslant n+1$): by Proposition 1, A_n is a $\mathcal{B}(\mathcal{E})$-measurable map from E to $\mathcal{K}_\Omega$. Since $(E, \mathcal{E})$ is a standard Borel space, the first separation theorem for analytic sets implies that $\mathcal{B}(\mathcal{E}) = \mathcal{E}$. Since the map A_n is $\mathcal{E}$-measurable, Proposition 7.3 yields the existence of a Borel map f_0^n from the (Borel) projection of B_n on E to Ω, whose graph is included in B_n. Subtracting this graph from B_n yields a new Borel set B'_n and mapping A'_n, with $\#A'_n(e) \leqslant n$. Hence we obtain a new selection f_1^n from B'_n, and so on: B_n is the disjoint union of the graphs of Borel maps $f_0^n, f_1^n, \dots, f_n^n$ (defined on Borel subsets of E). Rank all functions f_j^n into one sequence f_i: we have a sequence of Borel maps f_i from Borel subsets E_i of E to Ω, with disjoint graphs, such that the union of the graphs is the set of atoms of the measures $P(d\omega \mid e)$.

For our purposes it will be more convenient to have the f_i's defined on the whole of E, even if the union of the graphs becomes too big. Thus, let $n_{-1}(e) = -\infty, n_k(e) = \min(\{+\infty\} \cup \{i \mid e \in E_i, i > n_{k-1}(e)\})$. Since the sets E_i are Borel, the n_k's are Borel too by induction. Let thus $g_k(e) = f_{n_k(e)}(e)(k = 0, 1, 2, \dots)$ on the Borel set $F_k = \{e \mid n_k(e) < +\infty\}$: the sequence (g_k, F_k) has all properties of the sequence (f_i, E_i), but in addition $F_k \supseteq F_{k+1}$. Define now inductively h_k by $h_k = g_k$ on F_k, $h_k(e) = \max\{i^{-1} \mid i^{-1} \neq h_j(e) \; \forall \; j < k\}$ for $e \notin F_k$. One checks immediately by induction that the h_k's are Borel. Thus we have a sequence h_i of Borel functions from E to Ω, with disjoint graphs, such that those graphs cover all atoms of the measures $P(d\omega \mid e)$.

Add now $\mathbb{N} \times (E, \mathcal{E})$ to Ω, obtaining thus $(\Omega', \mathcal{A}')$. Define N' on Ω' by $N'(\omega) = N(\omega)$ for $\omega \in \Omega$, $N'(k, e) = N(h_k(e))$, and $P'(d\omega' \mid e)$ by $P'(A \mid e) = P_{na}(A \mid e)$ for $A \subseteq \Omega$, denoting by P_{na} the nonatomic part of P, and by $P'(\{k\} \times \tilde{E} \mid e) = I_{\tilde{E}}(e) P(\{h_k(e)\} \mid e\} \; \forall \; \tilde{E} \in \mathcal{E}$.

Since the h_k are Borel, it is clear that N' is still a Borel map from $(\Omega', \mathcal{A}')$ to $\mathcal{K}^*(\mathbb{R}^\ell)$, since P is also a transition probability to $(E \times \Omega, \mathcal{E} \otimes \mathcal{A})$, and since the graph of h_k is Borel in this space, it follows also that $I_{\tilde{E}}(e) \cdot P(\{h_k(e)\} \mid e)$ is $\mathcal{E}$-measurable. Thus, to show that P' is also a transition probability from $(E, \mathcal{E})$ to $(\Omega', \mathcal{A}')$, there only remains to show that $P_{na}(A \mid e)$ is $\mathcal{E}$-measurable for $A \in \mathcal{A}$: this follows from $P_{na}(A \mid e) = P(A \mid e) - \Sigma_{k=0}^\infty I_{h_k^{-1}(A)}(e) P(\{h_k(e)\} \mid e)$, as we have seen that each term is $\mathcal{E}$-measurable.

It is clear that $\int N dP = \int N' dP'$, since any Borel function, modified arbitrarily on a countable set, remains Borel. Thus, it will be sufficient to prove the theorem for Ω', N' and P': indeed, if f' solves (3) for those data, $f(e, \omega)$ defined for $\omega = h_k(e)$ by $f'(e, (k, e))$ and for $\omega \notin \cup_k \{h_k(e)\}$ by $f'(e, \omega)$ will be Borel since the graph of each h_k is a Borel set and since the restriction of f to each of those and to their complement is Borel. It is now clear that f will solve the original problem.

Let $A_\infty \in \mathcal{A}'$ denote the original space Ω, and A_k the subset $\{k\} \times E$. Let $C(e) = \int_{A_\infty \cup (\cup_{i \geqslant 1} A_i)} N'(\omega')P'(d\omega' \mid e)$, $K(e) = \int_{A_0} N'(\omega')P'(d\omega' \mid e)$ $[= N(h_0(e))P(h_0(e) \mid e)]$. By (1), C and K are measurable maps to $\mathcal{K}^*(\mathbb{R}^\ell)$, and $C(e) + K(e) = (\int N' dP')(e)$. Thus, by Proposition 6.6, $C(e) \times K(e)$ is a measurable map to $\mathcal{K}^*(\mathbb{R}^\ell \times \mathbb{R}^\ell)$, and taking for ψ the sum from $\mathbb{R}^\ell \times \mathbb{R}^\ell$ to $\mathbb{R}^\ell$, Proposition 9.3 yields us measurable maps φ_C and φ_K from the graph F of $\int N dP$ to $\mathbb{R}^\ell$ such that $\varphi_C(e,x) + \varphi_K(e,x) = x$ and $\varphi_C(e,x) \in C(e), \varphi_K(e,x) \in K(e)$. Let $\psi_C(e) = \varphi_C(e, x(e))$, $\psi_K(e) = \varphi_K(e, x(e))$: we have $\psi_C(e) \in C(e)$, $\psi_K(e) \in K(e)$, $\psi_C(e) + \psi_K(e) = x(e)$. Let $f(e) = \psi_K(e)/P'(A_0 \mid e)$ on the Borel set $\{e \mid P'(A_0 \mid e) > 0\}$, and define $f(e)$ as an arbitrary Borel selection of $N'[(0,e)]$ on the complement (Proposition 7 3). Then $f(e)$ is Borel, $f(e) \in N'[(0,e)]$, and $\psi_C(e) + f(e)P'(\{(0,e)\} \mid e) = x(e)$. Thus if N' is modified to N'_0, by setting $N'[(0,e)] = \{f(e)\}, N'_0 = N'$ on the complement of A_0, then $N'_0 \subseteq N', N'_0$ is still a Borel map to $\mathcal{K}^*(\mathbb{R}^\ell)$, and $x(e) \in (\int N'_0 dP')(e)$. By induction, we can thus define a decreasing sequence of Borel maps $N'_k \subseteq N'$ from $(\Omega', \mathcal{A}')$ to $\mathcal{K}^*(\mathbb{R}^\ell)$, such that for all $k, x(e) \in (\int N'_k dP')(e)$ and such that $\omega' \in A_i \Rightarrow N'_k(\omega')$ is a singleton for $k \geqslant i$. Let $N'_\infty = \cap_k N'_k$: this is still Borel, is a singleton on each A_k, and one still has $x(e) \in (\int N'_\infty dP')(e)$ by Lemma 1 and (2). Since also $N'_\infty \subseteq N'$, it is sufficient to prove the result for N'_∞. Hence it is sufficient to prove the result for the restriction of N'_∞ and P' to A_∞ – i.e., for the original N on the original $(\Omega, \mathcal{A})$, with the kernel P_{na} – with, as new function x, the function $x(e) - \int_{\cup_k A_k} N'_\infty(\omega')P'(d\omega' \mid e)$, since this integral is a singleton for each e, and a Borel map. Finally, $\{e \mid P_{na}(\Omega \mid e) = 0\}$ being a Borel set, on which the selection f can be defined arbitrarily by Proposition 7.3, it suffices to do the proof on its complement, which is still Borel-isomorphic to $[0,1]$.

As before, we can then renormalize, dividing $x(e)$ and $P_{na}(A \mid e)$ by $P_{na}(\Omega \mid e)$. Thus we have reduced the problem to the case where P is a nonatomic transition probability from $[0,1]$ to itself.

Note that the above argument proves the following lemma:

Lemma 2 *Let $(\Omega, \mathcal{A})$ be a separable measurable space, such that $\mathcal{A}$ separates points of Ω. Denote by $(P, \mathcal{P})$ the space of bounded measures on $(\Omega, \mathcal{A})$, endowed with the coarsest σ-field for which the mappings $p \to p(A), A \in \mathcal{A}$ are measurable. Then the set of nonatomic measures belongs to $\mathcal{P}$, and on its complement there exists a $(\mathcal{P}, \mathcal{A})$-measurable map φ to Ω, such that, for all p, $\varphi(p)$ is an atom of p of maximal measure. (Removing this atom from p thus yields a measurable map to $(P, \mathcal{P})$, on which φ can be applied again to remove a second atom, and so on.)*

Let $X = \{(x,K) \in \mathbb{R}^\ell \times \mathcal{K}^*(\mathbb{R}^\ell) \mid x \in K,\ K \text{ is convex}\}$: X is locally compact with a countable basis, as a closed subspace of such a space.

We now prove:

Lemma 3

1. *The map $\varphi\colon X \to \mathcal{K}^*(\mathbb{R}^\ell)$ defined by $\varphi(x,K) = \{p \in \mathbb{R}^\ell \mid \|p\| \leqslant 1, \langle p,x\rangle = \min\{\langle p,y\rangle \mid y \in K\}\}$ is u.s.c. and convex-valued.*
2. *The map $\chi : (p,K) \to \{x \in K \mid p \in \varphi(x,K)\}$ is u.s.c.*

Proof. Since always $0 \in \varphi(x,K)$, it is clear that φ has nonempty, compact convex values. Since the values of φ are included in a fixed compact set – the unit ball – it is sufficient to check that φ has a closed graph. This is an immediate verification. It also implies (2).

Lemma 4 *There exist Borel maps on X - $\psi_F, \psi_D, k, \psi_1, \psi_2, \alpha$ – such that*

1. *$\psi_F(x,K) \in \mathcal{K}^*(R^\ell)$ is the minimal face of K containing x.*
2. *$\psi_D(x,K) \in (\mathbb{R}^\ell)^k$ is a list of $k = k(x,K)$ linearly independent vectors $(p_1,\dots,p_k)$ in $\mathbb{R}^\ell$ – when $\psi_F(x,K)$ is $(\ell-k)$-dimensional – such that $\psi_F(x,K)$ is the set of points in K that lexicographically minimize $(p_1,\dots,p_k)$.*
3. *$\psi_1, \psi_2 \in \mathbb{R}^\ell$ belong to the relative boundary of $\psi_F(x,K)$ if x is not an extreme point of K, and equal x otherwise.*
4. *$0 < \alpha < 1$ is such that $x = \alpha(x,K)\psi_1(x,K) + (1-\alpha(x,K))\psi_2(x,K)$.*

Proof. Denote by h a Borel selection from $\mathcal{K}^*(\mathbb{R}^\ell)$ (Proposition 7.3). Let $\bar\varphi(x,K;p_1,\dots,p_k) = \varphi(x,K) \cap \{p \mid \|p\| = 1, \langle p,p_i\rangle = 0, i = 1,\dots,k\}$ for $k \geqslant 0$: $\bar\varphi$ is Borel as an intersection of two u.s.c. maps (Lemma 3.1, Propositions 6.3 and 5.2). Define inductively Borel maps $K_0(x,K) = K$, and, for $0 \leqslant i < i_0$, $p_{i+1}(x,K) = h \circ \bar\varphi[x, K_i(x,K); p_1(x,K), p_2(x,K), \dots, p_i(x,K)]$ and $K_{i+1}(x,K) = \chi[p_{i+1}(x,K), K_i(x,K)]$, where $i_0(x,K)$ is the first index i such that $\bar\varphi[x, K_i(x,K); p_1(x,K), \dots, p_i(x,K)] = \emptyset$. Since $p_{i+1} \in \varphi(x,K_i)$ and $x \in K_0$ we get by induction that $x \in K_i$ for all i.

By induction $K_i(x,K)$ is the face of K where the linear functionals $p_1(x,K),\dots,p_i(x,K)$ are lexicographically minimized; and $K_{i+1} \subseteq K_i$. Since the vectors $p_1,\dots,p_i$ are thus constant on K_i, $\bar\varphi(x,K_i;p_1,\dots,p_i) = \emptyset$ only when $\varphi(x,K_i)$ is included in the vector space spanned by $p_1,\dots,p_i$ – i.e., K_i is the minimal face containing x and $p_1,\dots,p_i$ are a basis of equations for the affine space spanned by this face – a basis, because the p_i's are mutually orthogonal and of norm 1, which also implies that $i_0(x,K) \leqslant \ell$. The map i_0 is clearly Borel too, since $\bar\varphi$ is Borel and $\{\emptyset\}$ is Borel in $\mathcal{K}(\mathbb{R}^\ell)$.

Thus (1) follows since $\psi_F(x,K) = K_{i_0(x,K)}(x,K)$, and (2) also since we can take $\psi_D(x,K) = (p_1(x,K), p_2(x,K), \ldots, p_{i_0(x,K)}(x,K))$. As for (3), denote by $f(x,K)$ a Borel selection (Proposition 7.3 and Proposition 5.2) from the u.s.c. (like Proposition 7.4) map $(x,K) \to \{y \in K \mid d(x,y) = \max_{z\in K} d(x,z)\}$. Since $\{(x,K) \mid x$ is an extreme point of $K\} = \{(x,K) \mid i_0(x,K) = \ell\}$ is Borel, we just have to define ψ_1 and ψ_2 on the complement. Let $\psi_1(x,K)$ denote the Borel map $f(x,\psi_F(x,K))$, and denote by $\psi_2(x,K)$ the endpoint opposite to $\psi_1(x,K)$ of the line segment which is the intersection of K with the straight line joining x to $\psi_1(x,K)$: since $\psi_1(x;K) \neq x$, the straight line (viewed in the one-point compactification of $\mathbb{R}^\ell$ to preserve compactness) is a continuous function of x and ψ_1, hence its intersection with K is an u.s.c. (Proposition 6.3) – hence Borel (Proposition 5.2) – function of x, ψ_1 and K, and its extremity opposite to ψ_1 is clearly a continuous function of this segment and of ψ_1: hence ψ_2 is Borel too. $\alpha(x,K)$ is then uniquely determined (set $\alpha = \frac{1}{2}$ if x is an extreme point of K), and necessarily Borel – this proves the lemma.

Lemma 5 *If $g(\omega)$ is a Borel function, integrable for each $P(d\omega \mid e)$, there exists a Borel set B in $E \times [0,1] \times \Omega$ such that identically $\int_{B_{e,\alpha}} g(\omega)P(d\omega \mid e) = \alpha \int g(\omega)P(d\omega \mid e)$ with $B_{e,\alpha} = \{\omega \mid (e,\alpha,\omega) \in B\}$ and such that $B_{e,\alpha}$ is monotone in α.*

Proof. Replace each coordinate of g by its positive and negative parts, and ℓ by (2ℓ): it is clearly sufficient to consider this case. Thus we assume that all coordinates of g are nonnegative. Since for each coordinate g_i of g, $\{e \mid \int g_i(\omega)P(d\omega \mid e) = 0\}$ is Borel, we get a Borel partition of E, such that on each partition element, for each i $\int g_i(\omega)P(d\omega \mid e)$ is either identically zero or everywhere positive. It is clearly sufficient to prove the theorem on each partition element separately. We can then delete the coordinates i for which $\int g_i(\omega)P(d\omega \mid e)$ is identically zero, and thus assume that $\int g_i(\omega)P(d\omega \mid e) > 0$ everywhere. Now let $P_i(A \mid e) = \left[\int_A g_i(\omega)P(d\omega \mid e)\right] / \left[\int g_i(\omega)P(d\omega \mid e)\right]$: $P = (P_0, P_1, \ldots, P_{\ell-1})$ is a vector of nonatomic transition probabilities from $(E,\mathcal{E})$ to $(\Omega,\mathcal{A})$, both of those spaces are copies of $[0,1]$ with the Borel sets (enlarge E if necessary), and we want that $P_i(B_{e,\alpha} \mid e) = \alpha$ identically. We can further add, w.l.o.g., the average of the P_i's as an additional P_ℓ, and thus assume that all P_i's are absolutely continuous w.r.t. P_ℓ.

We prove the existence of such a Borel set B by induction on ℓ. For $\ell = 0$, let $F(\omega,e) = P_0(\{\omega' \mid \omega' \leqslant \omega\} \mid e)$ be the cumulative distribution function of P_0: F is monotone and continuous in ω, with $F(0,e) = 0$, $F(1,e) = 1$, and, as we have seen previously, F is measurable on $E \times \Omega$. Thus $B = \{(e,\alpha,\omega) \mid F(e,\omega) \leqslant \alpha\}$ satisfies our requirements.

For the induction step, we can assume that we have a set B such that $P_i(B_{e,\alpha} \mid e) = \alpha$, $\forall i = 1, \ldots, \ell$. Now let $F(e,x) = P_0(B_{e,x} \mid e)$. F is measurable on $E \times [0,1]$, since P_0 can be viewed as a transition probability from $E \times [0,1]$ to $\Omega' = E \times [0,1] \times \Omega$, and B is Borel in the latter. Further, F is monotone in x, since $B_{e,x}$ is so and $P_0(\cdot \mid e)$ is a probability, and it is continuous in x, with $F(e,0) = 0$, $F(e,1) = 1$ because $P_0(\cdot \mid e)$ is absolutely continuous with respect to $P_\ell(\cdot \mid e)$, and $P_\ell(B_{e,x} \mid x) = x$.

Let $G(e,x) = F(e, x + \frac{1}{2}) - F(e,x) - \frac{1}{2}$ for $0 \leqslant x \leqslant \frac{1}{2}$. G is still measurable on $E \times [0, \frac{1}{2}]$, and continuous in x. Further, $G(e,0) + G(e, \frac{1}{2}) = F(e,1) - F(e,0) - 2 \cdot \frac{1}{2} = 0$. Thus, if we set $X(e) = \{x \in [0, \frac{1}{2}] \mid G(e,x) = 0\}$, then $X(e)$ is a map with Borel graph (measurability of G) and with compact (continuity of G in x), nonempty ($G(e,0) + G(e, \frac{1}{2}) = 0$, Rolle's theorem) values. Hence by Proposition 1, X is a $\mathcal{B}(\mathcal{E})$-measurable map to $\mathcal{K}([0, \frac{1}{2}])$ – hence $\mathcal{E}$-measurable by the first separation theorem, since $(E, \mathcal{E})$ is a standard Borel space. Thus by Proposition 7.3, $X(e)$ has a Borel selection $x(e)$: we have a Borel map $x(e)$ from $(E, \mathcal{E})$ to $[0, \frac{1}{2}]$ such that $F(e, x(e) + \frac{1}{2}) - F(e, x(e)) = \frac{1}{2}$ identically. Then let $\tilde{B} = \big\{(e,\omega) \mid (e, x(e) + \frac{1}{2}, \omega) \in B\big\} \setminus \big\{(e,\omega) \mid (e, x(e), \omega) \in B\big\}$: the two sets, and hence $\tilde{B}$, are Borel as inverse images of a Borel set B by a Borel mapping. Further, as B is monotone in α, the second set $\tilde{B}_2$ is included in the first $\tilde{B}_1$. Since also $P_i(\tilde{B}_2(e) \mid e) = x(e)$ and $P_i(\tilde{B}_1(e) \mid e) = x(e) + \frac{1}{2}$ identically for $i = 1, \ldots, \ell$, (property of B), and since $P_0(\tilde{B}_2(e) \mid e) = F(e, x(e)), P_0(\tilde{B}_1(e) \mid e) = F(e, x(e) + \frac{1}{2})$, we obtain for all i, $0 \leqslant i \leqslant \ell$ that $P_i(\tilde{B}(e) \mid e) = \frac{1}{2}$.

Denote this Borel set in $E \times \Omega$ by $\tilde{B}_{\frac{1}{2}}$. Let also $\tilde{B}_0 = \emptyset, \tilde{B}_1 = E \times \Omega$. Since, after multiplication by 2, the restriction of $P_0, \ldots, P_\ell$ to $\tilde{B}_{\frac{1}{2}} \setminus \tilde{B}_0$ and to $\tilde{B}_1 \setminus \tilde{B}_{\frac{1}{2}}$ are also a vector of $(\ell+1)$ nonatomic transition probabilities from E to Ω such that each one is absolutely continuous with respect to the last one, we can similarly find Borel sets C and D in $E \times \Omega$ which cut those into 2. Letting then $\tilde{B}_{1/4} = \tilde{B}_0 \cup \left[C \cap \tilde{B}_{1/2}\right], \tilde{B}_{3/4} = \left[D \cup \tilde{B}_{1/2}\right] \cap \tilde{B}_1$ we obtain two new Borel sets, such that $\tilde{B}_0 \subseteq \tilde{B}_{1/4} \subseteq \tilde{B}_{1/2} \subseteq \tilde{B}_{3/4} \subseteq \tilde{B}_1$, and such that $P_i(\tilde{B}_r(e) \mid e) = r \ \forall\, i = 0, 1, 2, \ldots, \ell$ and $\forall\, r = k \cdot 2^{-2} (0 \leqslant k \leqslant 2^2)$. By induction, we obtain thus Borel sets $\tilde{B}_r$ in $E \times \Omega$ for all dyadic r, such that $r \leqslant s$ implies $\tilde{B}_r \subseteq \tilde{B}_s$, and such that $P_i(\tilde{B}_r(e) \mid e) = r$ for all i and r. Now let B_r denote the Borel set $\tilde{B}_r \times [r, 1]$ in $E \times \Omega \times [0,1]$, and B the Borel set $\cup_r B_r$: we have $B_{e,\alpha} = \{\omega \mid \exists r \leqslant \alpha\colon (e,\omega) \in \tilde{B}_r\} = \cup_{r \leqslant \alpha} \tilde{B}_r(e)$, and since the $\tilde{B}_r$ are monotone, it follows that $P_i(B_{e,\alpha} \mid e) = \sup_{r \leqslant \alpha} P_i(\tilde{B}_r(e) \mid e) = \sup\{r \mid r \leqslant \alpha\} = \alpha$, for all i, α and e. Since clearly B is also monotone in α, this finishes the induction step, and hence the proof of Lemma 5.

To finish the proof of the theorem, we construct the Borel function

$f(e,x,\omega) \in N(\omega)$ on $F \times \Omega$ inductively on the Borel sets $F_n \times \Omega$, where $F_n = \{(e,x) \in F \mid$ the face of $(\int NdP)(e)$ spanned by x is n-dimensional $\} = \{(e,x) \in F \mid k\big(x, (\int NdP)(e)\big) = \ell - n\}$ (by Lemma 4, k and hence F_n are Borel). Write for short the different functions in Lemma 4 as functions of e and x instead of $(\int NdP)(e)$ and x, hence as Borel functions on F, since $\int NdP$ is a measurable map (1.b). On F_0, let $f(e,x,\omega)$ denote the lexicographic minimizer of $p_1(e,x), p_2(e,x), \ldots, p_\ell(e,x)$ in $N(\omega)$ – it is a singleton because $p_1, \ldots, p_\ell$ are linearly independent. This will be a Borel selection of N because the $p_i(e,x)$ are Borel (Lemma 4.2) and $N(\omega)$ is, and because the lexicographic minimizer of $(p_1, \ldots, p_\ell)$ in K is a Borel function of $(p_1, \ldots, p_\ell; K)$ – by inductive use of Lemma 3.2 (plus Proposition 5.2 and Proposition 6.1). Clearly $\int f(e,x,\omega)P(d\omega \mid e) = x$, since x is the lexicographic minimizer of $(p_1, \ldots, p_\ell)$ in $(\int NdP)(e)$.

Assume now by induction that f is well defined and satisfies our requirements on $\cup_{i<n} F_i$. For $(e,x) \in F_n$ we have $\psi_1(e,x)$ and $\psi_2(e,x) \in \cup_{i<n} F_i$. Let thus $g_i(e,x,\omega) = f(e, \psi_i(e,x), \omega)$ $(i = 1,2)$: those are Borel functions on $\Omega' = F_n \times \Omega$ – which is Borel-isomorphic to $[0,1]$ – and P can be viewed as a transition probability from F_n to Ω', denoted $P(d\omega' \mid e,x)$. Let $g(\omega') = g_2(\omega') - g_1(\omega')$, and consider the Borel set B in $F_n \times [0,1] \times \Omega'$ given by Lemma 5. For $u \in F_n$ and letting $\varphi(u,\alpha,\omega') = g_2(\omega')I_B(u,\alpha,\omega') + g_1(\omega')I_{B^c}(u,\alpha,\omega')$ we get a Borel function φ on $F_n \times [0,1] \times \Omega'$ that satisfies (by adding g_1) $\int \varphi(u,\alpha,\omega')P(d\omega' \mid u) = \alpha \int g_2(\omega')P(d\omega' \mid u) + (1-\alpha)\int g_1(\omega')P(d\omega' \mid u)$. Hence, with $f(e,x,\omega) = \varphi((e,x), \alpha(e,x), (e,x,\omega))$, ($\alpha(e,x)$ being given by Lemma 4.4), f will be Borel, will satisfy $f(e,x,\omega) \in N(\omega)$, since $g_i(e,x,\omega) \in N(\omega)$, and also $\int f(e,x,\omega)P(d\omega \mid e) = \alpha(e,x)$ $\int f(e,\psi_2(e,x),\omega)P(d\omega \mid e) + (1-\alpha(e,x))\int f(e,\psi_1(e,x),\omega)P(d\omega \mid e)$ which equals $\alpha(e,x)\psi_2(e,x) + (1-\alpha(e,x))\psi_1(e,x)$, since f satisfies our requirements on $\cup_{i<n} F_i$, and hence, by Lemma 4.4, $\int f(e,x,\omega)P(d\omega \mid e) = x$. Thus f satisfies our requirements on $\cup_{i\leqslant n} F_i$. This proves the induction; thus f is constructed so as to satisfy our conditions on the whole of $F \times \Omega$: the theorem is proved.

Appendix

A.1. Let $(\Omega, \mathcal{A})$ be a measurable space. Define the class $\mathscr{A}$ – or $\mathscr{A}(\mathcal{A})$ – of analytic subsets as the projections on $(\Omega, \mathcal{A})$ of measurable sets of the product of $(\Omega, \mathcal{A})$ with a standard Borel space (i.e., a measurable space isomorphic to the unit interval with the Borel sets).

Clearly $\mathcal{A} \subseteq \mathscr{A}$, and $\mathscr{A}$ is stable under countable unions and countable intersections: for unions, it follows immediately from the fact that a union of projections is the projection of the union; for intersections, let $A_i = \operatorname{proj}(M_i)$, M_i measurable in $\Omega \times S_i$, S_i standard Borel. Then $\cap_i A_i = \operatorname{proj}\left(\cap_i [M_i \times \Pi_{j \neq i} S_i]\right)$, where the intersection is taken in the product of Ω with the standard Borel space $S = \Pi_j S_j$. Further, the projection on Ω of any analytic subset $\tilde{A}$ of $\Omega \times S$ – say $\tilde{A} = \operatorname{proj}(B)$, B-measurable in $\Omega \times S \times S'$ – is analytic in Ω, being the projection of B. Finally, it is immediate that inverse images of analytic sets by measurable maps are analytic. $\mathscr{A}^c$ will denote the coanalytic sets, i.e., sets whose complement is analytic. Denote by $\mathcal{B}$ – or $\mathcal{B}(\mathcal{A})$ – the class of bianalytic sets – i.e., $\mathcal{B} = \mathscr{A} \cap \mathscr{A}^c$. Since $\mathscr{A}$ is stable under countable unions and intersections, $\mathcal{B}$ is a σ-field containing $\mathcal{A}$. Similarly, any measurable map between two spaces is also measurable for their bianalytic σ-fields. Stability of analytic sets under projections implies that $\mathcal{B}(\mathcal{B}(\mathcal{A})) = \mathcal{B}(\mathcal{A})$.

A.2. The Hausdorff topology on the space $\mathcal{K}_E$ (or $\mathcal{K}(E)$) of compact subsets of a Hausdorff topological space E is defined by the basis of open sets $\{K \mid K \subseteq \cup_{i=1}^n O_i, K \cap O_i \neq \emptyset \ \forall i\}$ for all finite families of open sets O_i in E. We let $\mathcal{K}_E^* = \mathcal{K}_E \setminus \emptyset$.

It is immediate to verify that $\mathcal{K}_E$ inherits many topological properties of E – e.g., if E is Polish, or locally compact, so is $\mathcal{K}_E$ (check first for compact spaces, e.g., using Alexander's subbase theorem – and second that, if $E' \subseteq E$, $\mathcal{K}_{E'}$ is a subspace of $\mathcal{K}_E$).

Proposition 1 *Let Γ be a measurable compact-valued correspondence from $(\Omega, \mathcal{A})$ to a metrizable Lusin space E (i.e., Γ is a measurable (or just bianalytic) subset of the product, such that $\Gamma_\omega \in \mathcal{K}_E$ for all ω).*

Then $\omega \to \Gamma_\omega$ is measurable from $(\Omega, \mathcal{B})$ to $\mathcal{K}_E$, and is, when E is compact, the limit of a decreasing sequence of measurable step functions from $(\Omega, \mathcal{B})$ to $\mathcal{K}_E$.

Note. Extension to the case where the values of Γ are K_σ's instead of compact would require a different proof: one should use directly the Arsenin theorem mentioned in Rogers et al. at the end of page 255, and use on subsets of E the σ-field generated by $\{X \subseteq E \mid X \cap F \neq \emptyset\}$ where F varies over the closed subsets of E. (Or better, to get a measurable structure on

$\{\Gamma_\omega\}$ which depends only on the topology of the sets Γ_ω and on the Borel structure of E, consider the σ-field generated by $\{\Gamma_\omega \mid \Gamma_\omega \cap F \neq \emptyset\}$ when F varies over all Borel sets of E such that $\forall\omega, F \cap \Gamma_\omega$ is a K_σ.) Metrizability of E can then also be dispensed with (Proposition 4).

Proof. We follow essentially the proof of Dellacherie ([8], pp. 218–225), indicating just the required modifications. E is metrizable, thus homeomorphic to a Borel subset of a compact metric space. If Γ is a bianalytic subset of $\Omega \times E$, there exist two measurable subsets G_1 and G_2 of $\Omega \times E \times S$ (standard Borel), whose projections are respectively Γ and Γ^c. Each set G_i is in the σ-field generated by a sequence of product sets $A_n^i \times E_n^i \times S_n^i$. Denote by $\mathcal{U}$ the separable sub σ-field of $\mathcal{A}$ generated by the sets A_n^i: there exists a $\mathcal{U}$-measurable map φ from Ω to the Cantor set $C = \{0,1\}^{\mathbb{N}}$, such that $F = \varphi(\Omega)$ is dense in C and $\mathcal{U}$ is the σ-field generated by φ and the Borel sets. Γ can then also be viewed as a bianalytic subset of $F \times E$. Since in addition φ is in particular measurable for the bianalytic σ-fields on Ω and F, it will be sufficient to prove the theorem when $(\Omega, \mathcal{A})$ is a dense subspace F of the Cantor space C, with the Borel sets.

Similarly, since E is a Borel subset of a compact metric space $\bar{E}$ (which is itself a subspace of $[0,1]^{\mathbb{N}}$), Γ is still a bianalytic subset of $F \times \bar{E}$. Since also $\mathcal{K}_E$ is a subspace of $\mathcal{K}_{\bar{E}}$, it follows that we can assume $E = [0,1]^{\mathbb{N}}$.

Note first that, if $(\Omega, \mathcal{A})$ is compact metric (or Polish, or ...) with its Borel sets, our above definition of analytic sets yields the usual definition (e.g., continuous images of $\mathbb{N}^{\mathbb{N}}$): indeed, Borel subsets of $(\Omega \times [0,1])$ are known to be analytic, and the projection is continuous; conversely, $\mathbb{N}^{\mathbb{N}}$ is homeomorphic with the irrationals in $[0,1]$, so the graph of the continuous mapping becomes a Borel subset of $(\Omega \times [0,1])$. This proves our claim – assuming the identity between the Borel σ-fields on $(\Omega \times [0,1])$ and the product of the Borel σ-fields to be known. Γ being analytic in $E \times F$, there exists a Borel set in $E \times F \times [0,1]$ with projection Γ. This is the trace on $E \times F \times [0,1]$, of a Borel subset of $E \times C \times [0,1]$, whose analytic projection A on $E \times C$ has Γ as trace on $E \times F$. Similarly for the complement of Γ, there exists an analytic subset $\tilde{A}$ of $E \times C$ with Γ^c as trace on $E \times C$.

Let U_n denote a basis of open sets in E, and let C_n be the coanalytic complement of $\tilde{D}_n = \mathrm{proj}_C[(U_n \times C) \cap A]$. Let L_n be the coanalytic set $U_n \times C_n$. Since Γ has compact values, $(\cup_{n \geqslant 1} L_n) \cap (E \times F) = \Gamma^c$. Let also L_0 be the complement of $\tilde{A}$ (hence $L_0 \cap (E \times F) = \Gamma$). Then $\cup_n L_n$ covers $E \times F$. By the second separation theorem (loc cit Theorem 28, p. 252), there exists a sequence M_n of disjoint coanalytic sets, $M_n \subseteq L_n$, with $\cup_n M_n = \cup_n L_n$. Now let B denote the complement of $\cup_n M_n$, and $\bar{M}_n = M_n \cup B$: $\bar{M}_n = \cap_{k \neq n}(M_k^c)$ is analytic, hence also its projection D_n on C. Since $\cup_n M_n = \cup L_n$ covers $E \times F$, $\mathrm{proj}_C(B) \subseteq C \setminus F$, hence $D_n \cap F = (\mathrm{proj}_C M_n) \cap F \subseteq$

$(\mathrm{proj}_C L_n) \cap F = C_n \cap F$. Thus C_n and the complement of D_n are coanalytic sets that cover F. By the same second separation theorem, there exists a bianalytic subset B_n of F, such that $D_n \cap F \subseteq B_n \subseteq C_n \cap F$. Then $\cup_{n\geqslant 1}[U_n \times B_n] \supseteq \cup_{n\geqslant 1}[U_n \times D_n]_{|F} \supseteq (\cup_{n\geqslant 1} M_n)_{|F} = (C \times E \setminus \bar{M}_0)_{|F} \supset \Gamma^c$, and $\cup_{n\geqslant 1}[U_n \times B_n] \subseteq \cup_{n\geqslant 1}[U_n \times C_n]_{|F} = \Gamma^c$ as seen previously. Thus $\cup_{n\geqslant 1}[U_n \times B_n] = \Gamma^c$. Thus $\Gamma = \cap_{n\geqslant 1}[U_n^c \times F \cup E \times B_n^c]$.

Let $\Gamma_k = \cap_{n\geqslant 1}^k [U_n^c \times F \cup E \times B_n^c]$: then there exists a finite, bianalytic partition of F on each element of which Γ_k is a constant compact subset of E: Γ_k is a decreasing sequence of measurable step functions from $(\Omega, \mathcal{B})$ to $\mathcal{K}_E$, with intersection Γ. Since a decreasing sequence in $\mathcal{K}_E$ is convergent, it follows that Γ is measurable, as a pointwise limit of measurable functions. This finishes the proof.

In the course of the proof we have also shown

Proposition 2 *Given any countable family of analytic subsets of products of $(\Omega, \mathcal{A})$ with other measurable spaces, there exists a separable sub σ-field of $\mathcal{A}$ for which those sets are still analytic.*

We turn finally to transition probabilities. For a collection C of subsets, we denote by $\sigma(C)$ the σ-field it generates.

Proposition 3 *Let P denote a transition probability from $(E, \mathcal{E})$ to $(F, \mathcal{F})$. Then*

1. *(a) P is a transition probability from $(E, \mathcal{B}(\mathcal{E}))$ to $(E \times F, \mathcal{B}(\mathcal{E} \otimes \mathcal{F}))$;*
   (b) for any $\mathcal{B}(\mathcal{E} \otimes \mathcal{F})$-measurable function g, and any $e \in E$, $g(e, \cdot)$ is $\mathcal{B}(\mathcal{F})$-measurable.
2. *P is a transition probability from $(E, \sigma(\mathscr{A}(\mathcal{E})))$ to $[E \times F, \sigma(\mathscr{A}(\mathcal{E} \otimes \mathcal{F}))]$.*
3. *$\forall A \in \mathscr{A}(\mathcal{F}), \forall \alpha, \{e \in E \mid P_e(A) \geqslant \alpha\} \in \mathscr{A}(\mathcal{E})$.*

(It is well known that $A \in \mathscr{A}(\mathcal{F}) \Rightarrow A$ is μ-measurable for any probability μ on $(F, \mathcal{F})$.)

Remark : (2) also implies that P will still be a transition probability when both E and $E \times F$ are equipped with the smallest σ-field $\mathcal{G}$ containing resp. $\mathcal{E}$ and $\mathcal{E} \otimes \mathcal{F}$ and such that $\mathscr{A}(\mathcal{G}) = \mathcal{G}$.

Proof. It is well known that P is also a transition probability to $(E \times F, (\mathcal{E} \otimes \mathcal{F}))$. Let us rename this space $(F, \mathcal{F})$: it will be sufficient to prove (1) and (2) with F (resp. $\mathcal{F}$) instead of $E \times F$ (resp. $\mathcal{E} \otimes \mathcal{F}$). It is clear that (3) immediately implies (1a) (remember $\mathscr{A}$ is stable under countable unions and intersections). It is sufficient to prove (1b) for indicator functions – i.e., if $B \in \mathcal{B}(\mathcal{E} \otimes \mathcal{F})$, then $B_e \in \mathcal{B}(\mathcal{F})$. This will follow, by considering both B and its complement, if we show that $A \in \mathscr{A}(\mathcal{E} \otimes \mathcal{F}) \Rightarrow A_e \in \mathscr{A}(\mathcal{F})$. Let

$\tilde{A}$ be a measurable set in $E \times F \times [0,1]$ with A as projection. Then $\tilde{A}_e$ is measurable in $F \times [0,1]$, and has A_e as projection.

To show that (2) will also follow from (3), we first claim that the Boolean algebra generated by $\mathcal{A}$ consists of the sets of the form

$$X_n = (A_1 \setminus A_2) \cup (A_3 \setminus A_4) \cup (A_5 \setminus A_6) \cup \ldots \cup (A_{2n-1} \setminus A_{2n})$$

where the A_i's are a decreasing sequence in $\mathcal{A}$.

For this, check first stability under unions: it will be sufficient to show that the union of $B_1 \setminus B_2$ with X_n has the same form:

$$(B_1 \setminus B_2) \cup (A_1 \setminus A_2) = (B_1 \cup A_1) \setminus (B_2 \cup A_1) \cup (A_1 \setminus A_2) \cup (B_1 \cap A_2) \setminus (B_2 \cap A_2),$$

and the same decomposition can now be applied to the last term and $A_3 \setminus A_4$, and so on.

Note now that the complement of $A_1 \setminus A_2$ is the disjoint union of $\Omega \setminus A_1$ and of A_2. Hence the complement of X_n is a disjoint union of intersections of coanalytic and analytic sets, i.e., of sets of the type $A_1 \setminus A_2$. The stability under unions implies therefore that the collection is a Boolean algebra.

(3) implies that, for any analytic set A, $P_e(A)$ is $\sigma(\mathcal{A}(\mathcal{E}))$-measurable. Thus so is $P_e(A_1 \setminus A_2)$, as the difference of measurable functions, and thus also $P_e(X_n)$, as a sum of measurable functions. Since $\sigma(\mathcal{A}(\mathcal{F}))$ is the monotone class generated by the sets X_n, (2) follows.

To prove (3), first use Proposition 2 to reduce the problem to the case of separable $\mathcal{F}$. Next, as in the proof of Proposition 1, we can assume that F is a (dense) subset of the Cantor set C, with the Borel σ-field; and $A = \tilde{A} \cap F$, with $\tilde{A}$ analytic in C. Thus P can be viewed as a transition probability to C, such that any Borel set disjoint from F has probability zero. Since the probabilities on C form a compact metric space M in the weak*-topology, the Borel σ-field is the σ-field generated by the continuous linear functionals $\int f dp$, with f continuous on C (Stone-Weierstrass). Hence P is a measurable mapping from $(E, \mathcal{E})$ to M with the Borel sets. Denote by M_F the subspace of M consisting of those probabilities where any Borel set disjoint from F has zero weight.

Since the analytic set A is measurable for any measure μ on F, there exist two Borel sets B_μ and B'_μ in F, with $B_\mu \subseteq A \subseteq B'_\mu$, and $\mu(B_\mu) = \mu(B'_\mu) = \mu(A)$. Let $\bar{B}_\mu$ and $\bar{B}'_\mu$ denote two Borel sets in C with $B_\mu = \bar{B}_\mu \cap F, B'_\mu = \bar{B}'_\mu \cap F$. If necessary, rename their intersection as $\bar{B}_\mu$ and their reunion as $\bar{B}'_\mu$, to have in addition $\bar{B}_\mu \subseteq \bar{B}'_\mu$. Note that for any Borel set B in C, such as $\bar{B}'_\mu \setminus \bar{B}_\mu$, its measure $\bar{\mu}(B)$ under the extension of μ as a measure on C is defined to be equal to $\mu(B \cap F)$. Thus $\bar{\mu}(\bar{B}'_\mu \setminus \bar{B}_\mu) = 0$. Consider the set $\tilde{A} \setminus \bar{B}'_\mu$: being analytic, it is $\bar{\mu}$-measurable, hence its $\bar{\mu}$-measure is equal to the $\bar{\mu}$-measure of some Borel set contained in it: since this Borel set is disjoint from F, it has probability zero, hence $\bar{\mu}(\tilde{A}\setminus\bar{B}'_\mu) = 0$. Similarly $\bar{\mu}(\bar{B}_\mu \setminus \tilde{A}) = 0$, so that finally $\mu(A) = \bar{\mu}(\tilde{A})$, for any measure μ on F and any analytic subset $\tilde{A}$ of C with $\tilde{A} \cap F = A$.

Thus, if we prove that $\bar{A} = \{\mu \in M \mid \mu(\tilde{A}) \geqslant \alpha\}$ is analytic, it follows that $\bar{A} \cap M_F = \{\mu \in M_F \mid \mu(A) \geqslant \alpha\}$ is analytic, and hence (as noted in the beginning of this appendix), so will its inverse image in $(E, \mathcal{E})$ under the measurable mapping P from $(E, \mathcal{E})$ to $M_F : \{e \in E \mid P_e(A) \geqslant \alpha\} \in \mathcal{A}(\mathcal{E})$, and the theorem will be proved.

Let B denote a Borel subset of $C \times [0,1]$ with projection $\tilde{A}$. Let also $\bar{M}$ denote the space of probabilities on $C \times [0,1]$, with the weak*-topology, and let $\bar{B} = \{\mu \in \bar{M} \mid \mu(B) \geqslant \alpha\}$. Since $\mu(B)$ is a Borel function on $\bar{M}$ – it is in the monotone class generated by the functions $\int f d\mu$ for f continuous – $\bar{B}$ is a Borel set in $\bar{M}$. Let $\varphi : \bar{M} \to M$ map any measure in $\bar{M}$ to its marginal on C: φ is a continuous map, hence $\varphi(\bar{B})$ is analytic (as the projection of the intersection of the (closed) graph of φ with the Borel set $\bar{B} \times M$ – recall that $\bar{M}$ is a standard Borel space).

Clearly $\varphi(\bar{B}) \subseteq \bar{A}$. (If $\mu \in \bar{B}$, $\varphi(\mu)(\bar{A}) = \varphi(\mu)(\bar{D})$ for some Borel set D containing $\bar{A}$, since $\bar{A}$ is $\varphi(\mu)$-measurable; but by definition $\varphi(\mu)(D) = \mu(D \times [0,1]) \geqslant \mu(B)$.) To prove the converse, choose $\nu \in \bar{A}$. Use the measurable selection theorem (any version) to select a ν-measurable function f from C to $[0,1]$, such that ν-a.e. on $\tilde{A}$, $(x, f(x)) \in B$. Define the measure ν_f on $C \times [0,1]$ for any Borel set D by $\nu_f(D) = \nu\{x \mid (x, f(x)) \in D\}$: we have $\nu_f(B) = \nu(\tilde{A}) \geqslant \alpha$, and $\varphi(\nu_f) = \nu$. This proves the theorem.

A.3.

Proposition 4 *On every Lusin space, there is a weaker metrizable topology, which can further be chosen so as to leave any given open set open.*

Note: Lusin is not used: we only use the fact that it is a regular Hausdorff space S, which is the continuous image of a separable metric space $\tilde{S}$.

Proof. S – and any subspace of S – is a Lindelöf space, as a continuous image of the Lindelöf space $\tilde{S}$. Since it is also regular, it is paracompact (e.g., [2] § 4, Ex. 23(c)). Fix an open set O: each $x \in O$ has (regularity) an open neighborhood V_x with closure $\bar{V}_x \subseteq O$. By the Lindelöf property of O, there exists a sequence x_i such that the V_{x_i} cover O. Then the $\bar{V}_{x_i}$ cover O - hence every open set is an F_σ: S is perfectly normal (ibidem, Ex. 7), since it is also normal as a paracompact space. By Part (a) of the same Ex. 7, there exists for every closed set F a continuous function f_F with values in $[0,1]$ such that $F = f_F^{-1}(\{0\})$. By the same argument as above, the complement of the diagonal Δ in $S \times S$ is also Lindelöf. For any two distinct points x and y there exists (Hausdorff property) disjoint open neighborhoods O_x and $O_y : (O_x \times O_y) \cap \Delta = \emptyset$. By the Lindelöf property, we can find a sequence

(x_i, y_i) such that $\cup_i (O_{x_i} \times O_{y_i}) = (S \times S)\setminus\Delta$. Since the space is perfectly normal, there exist continuous functions f_i with values in $[0,1]$ such that $O_{x_i} = \{x \mid f_i(x) > 0\}$. Let $d(x,y) = |f_F(x) - f_F(y)| + \Sigma_i 2^{-i}|f_i(x) - f_i(y)|$: d is a continuous pseudo-metric, for which the function f_F is continuous, so that F is closed under d, and for any $x \neq y$, we have $(x,y) \in O_{x_i} \times O_{y_i}$ for some i – since those sets cover the complement of the diagonal – hence $f_i(x) > 0, f_i(y) = 0$: $d(x,y) > 0$, and d is a metric.

A.4. Recall that a Polish space is a space homeomorphic to a complete separable metric space – equivalently it is homeomorphic to a G_δ (a countable intersection of open sets) in a compact metric space.

Proposition 5 *Let X be a separable metric space. Then*

1. *$\mathcal{K}_X$ is a separable metric space; if X is Polish, so is $\mathcal{K}_X$.*
2. *An uppersemicontinuous (u.s.c.) map φ from a topological space Y to $\mathcal{K}_X$ is Borel. (φ u.s.c. means that $\{y \mid \varphi(y) \subseteq O\}$ is open for every open set O in X.)*
3. *Even with X arbitrary, φ induces an u.s.c. map from $\mathcal{K}_Y$ to $\mathcal{K}_X$.*

Proof. (1) X can be embedded in a compact metric space $\bar{X}$ (e.g., if x_i is a dense sequence in X, the map $x \to (\min(1, d(x,x_i)))_{i=1}^\infty$ is an embedding in $[0,1]^\infty = \tilde{S}$). Since, as noted in the beginning of **A.2** of this appendix, $\mathcal{K}_X$ is a subspace of $\mathcal{K}_{\bar{X}}$ (cfr. Proposition 6.7 below), and since $\mathcal{K}_{\bar{X}}$ is metrizable, the result follows. If X is open in $\bar{X}$ (i.e., locally compact) then so is $\mathcal{K}_X$ in $\mathcal{K}_{\bar{X}}$. Hence if X is a G_δ, so is $\mathcal{K}_X$.

(2) For any open set $O \subseteq X, \{z \mid \varphi(z) \subseteq O\}$ is open, hence for any closed set $F, \{z \mid \varphi(z) \cap F \neq \emptyset\}$ is closed. Since any open set U is an F_σ (a countable union of closed sets), $\{z \mid \varphi(z) \cap U \neq \emptyset\}$ is an F_σ. Thus, for any basic open set $V = \{K \mid K \subseteq O, K \cap U_i \neq \emptyset, i = 1, \ldots, n\}$ of $\mathcal{K}$, $\{z \mid \varphi(z) \in V\}$ is Borel. $\mathcal{K}_X$ is separable and metrizable (cf. (1)). Hence any open set is a countable union of basic open sets. Thus φ is Borel.

(3) Let $(O_i)_{i\in I}$ be an open covering of $\varphi(K)$ for $K \in \mathcal{K}_Y$. For $y \in K$, let I_y be a finite subset of I such that $\varphi(y) \subseteq \cup_{i\in I_y} O_i = O_y$, and let U_y be an open neighborhood of y such that $y' \in U_y \Rightarrow \varphi(y') \subseteq O_y$: the U_y form an open covering of K; let the $U_{y_j}, j \in J$, form a finite subcovering. Then $\{O_i \mid i \in \cup_{j\in J} I_{y_j}\}$ is a finite subcovering of $\varphi(K)$: hence $\varphi(K)$ is compact, and φ is indeed a map from $\mathcal{K}_Y$ to $\mathcal{K}_X$. Now let O be an open set in X. Then $U = \{y \mid \varphi(y) \subseteq O\}$ is open in Y. Hence $\{K \mid \varphi(K) \subseteq O\} = \{K \mid K \subseteq U\}$ is open in $\mathcal{K}_Y : \varphi$ is u.s.c. from $\mathcal{K}_Y$ to $\mathcal{K}_X$.

Proposition 6 *Let X, Y be topological spaces, f a continuous map from X to Y.*

1. *$x \to \{x\}$ is a homeomorphism from X into $\mathcal{K}_X$.*
2. *The union, from $\mathcal{K}_X \times \mathcal{K}_X$ to $\mathcal{K}_X$, is continuous.*
3. *The intersection, from $\mathcal{K}_X \times \mathcal{K}_X$ to $\mathcal{K}_X$, is u.s.c. when X is Hausdorff.*
4. *$\hat{f} : \mathcal{K}_X \to \mathcal{K}_Y : K \to f(K)$ is continuous.*
5. *When X is separable metric, the union and the intersection are measurable when $\mathcal{K}_X$ is endowed with the Borel σ-field and $\mathcal{K}_X \times \mathcal{K}_X$ with the product σ-field.*
6. *$\mathcal{K}_X \times \mathcal{K}_Y$ is a subspace of $\mathcal{K}_{X\times Y}$, closed if X and Y are Hausdorff.*
7. *If X is a subspace of Y, $\mathcal{K}_X$ is a subspace of $\mathcal{K}_Y$.*

Proof. (1), (2), (4), (6) and (7) are obvious. For (3), consider the map φ from $X \times X$ to $\mathcal{K}_X$ defined by $\varphi(x, y) = \emptyset$ if $x \neq y$, $\varphi(x, x) = \{x\}$. φ is u.s.c., by (1) and the Hausdorff character of X. Hence (5.3), φ induces an u.s.c. map from $\mathcal{K}_{X\times X}$ to $\mathcal{K}_X$, whose restriction to the subspace (cf. (6)) $\mathcal{K}_X \times \mathcal{K}_X$ is the intersection.

(5) follows from (2) and (3) because, by Proposition 5.1, the product σ-field on $\mathcal{K}_X \times \mathcal{K}_X$ is the Borel σ-field, and by Proposition 5.2, u.s.c. mappings are Borel.

Proposition 7 *Let X be a separable metric space. Then*

1. *If the measurable space $(E, \mathcal{E})$ is separable, and φ is a measurable mapping to $\mathcal{K}_X$, there exists a sequence of measurable mappings h_i from $(E, \mathcal{E})$ to X, such that $h_i(e) \in \varphi(e)$ $\forall e$, and such that the smallest class of functions that contains the h_i's and that contains, with every pointwise convergent sequence in the class, its limit, coincides with the class of all measurable selections from φ.*
2. *In particular (i.e., taking $(E, \mathcal{E})$ to be $\mathcal{K}_X$ with its Borel sets, and φ the identity), there exists a sequence f_i of Borel mappings from $\mathcal{K}_X$ to X, such that $\forall K \in \mathcal{K}_X$, the sequence $f_i(K)$ is a dense sequence in K.*
3. *In particular, there exists a Borel mapping f from $\mathcal{K}_X$ to X such that $f(K) \in K, \forall K \in \mathcal{K}_X$.*
4. *The mapping $(x, K) \to K_x = \{y \in K \mid d(x, y) = d(x, K)\}$ from $X \times \mathcal{K}_X$ to $\mathcal{K}_X$ is u.s.c.*

Proof. (4) is an immediate verification. For the other points, since $\mathcal{K}_X$ is a topological subspace of $\mathcal{K}_{\bar{X}}$ (Propositions 6 and 7), where $\bar{X}$ is a metrizable compactification of X, we can assume X compact metric.

(3) By (4) and Proposition 5.2, $\forall x_0 \in X$ the map $\varphi_0 : K \to K_{x_0}$ is Borel from $\mathcal{K}_X$ to itself. Let (x_i) be a dense sequence in X: since compositions of Borel mappings are Borel, the maps φ_n defined by $\varphi_n(K) = (\varphi_{n-1}(K))_{x_n}$ are Borel. Since the $\varphi_n(K)$ form a decreasing sequence of compact sets, they converge in $\mathcal{K}_X$, say to $\varphi_\infty(K)$: φ_∞ is still Borel, as a pointwise limit of Borel functions. Any two points x and y in $\varphi_\infty(K)$ are at the same

distance from each x_i; finding a subsequence x_{i_n} that converges to y yields $d(x,y) = \lim d(x, x_{i_n}) = \lim d(y, x_{i_n}) = d(y,y) = 0 : x = y$, so $\varphi_\infty(K)$ is a singleton, say $\{f(K)\}$. By Proposition 6.1, f is Borel. This proves (3).

For (2), let O_i be a countable basis of open sets in X, and denote their closures by $\bar{O}_i$. $C_i = \{K \mid K \cap \bar{O}_i \neq \emptyset\}$ is closed, and by Proposition 6.5, $K \to K \cap \bar{O}_i$ is Borel on C_i: let $\varphi_i(K) = K \cap \bar{O}_i$ for $K \in C_i$, and $\varphi_i(K) = K$ otherwise; then $\varphi_i : \mathcal{K}_X \to \mathcal{K}_X$ is Borel. Let $f_i = f \circ \varphi_i$, with f as in (3): clearly the sequence f_i fills the bill.

For (1), let $\mathcal{E}_0$ be a countable subalgebra of $\mathcal{E}$ that generate $\mathcal{E}$. Let $\mathcal{F}$ denote the set of bounded positive, integer-valued functions g on E, such that $\forall n, g^{-1}(n) \in \mathcal{E}_0$: $\mathcal{F}$ is still countable. For $g \in \mathcal{F}$, let $h_g(e) = f_{g(e)}(\varphi(e))$, where the sequence f_i is given by (2). Since compositions of measurable maps are measurable, the h_g's clearly form a sequence of measurable selections from φ. One checks immediately that, by repeated pointwise limit operations, one obtains from $\mathcal{F}$ first all measurable functions with finitely many integer values, next all integer-valued measurable functions. Thus, by repeated pointwise limit operations, we obtain from the h_g all functions $f_{n(e)}(\varphi(e))$, for any integer-valued measurable n. Given now an arbitrary measurable selection h from φ, define $n_k(e) = \min\{i \mid d\big(h(e), f_i(\varphi(e))\big) \leqslant k^{-1}\}$: the $n_k(e)$ are everywhere well defined, since the $f_i(K)$ are dense in K for any k, and are clearly measurable. Since h is the (uniform) limit of the $f_{n_k(e)}(\varphi(e))$, point (1) is also proved.

Proposition 8 *For $i = 1,2$, let X_i be a separable metric space, $(E_i, \mathcal{E}_i)$ a measurable space, with $\mathcal{E}_2$ separable, and Γ_i a mapping from E_i to $\mathcal{K}_{X_i}$ with Γ_2 measurable. Let Φ be a mapping that assigns to every measurable selection from Γ_2 a measurable selection from Γ_1, such that for each $e \in E_1$, and for every sequence of measurable selections f_i from Γ_2, any limit point of the sequence $[\Phi(f_i)](e)$ is in the closure of the set $\{[\Phi(g)](e) \mid g$ measurable selection from $(Lim\ f_i)\}$. For $e \in E_1$, let $\Gamma(e)$ denote the closure in X_1 of $\{[\Phi(f)](e) \mid f$ measurable selection from $\Gamma_2\}$. Then*

1. *Γ is a measurable mapping to $\mathcal{K}_{X_1}$, denoted $\Phi(\Gamma_2)$.*
2. *There exists a sequence f_i of measurable selections from Γ_2 such that, $\forall e \in E$, the $[\Phi(f_i)](e)$ are dense in $\Gamma(e)$.*
3. *For any pointwise convergent sequence of measurable mappings C_i from $(E_2, \mathcal{E}_2)$ to $\mathcal{K}_{X_2}$ satisfying $C_i \subseteq \Gamma_2$, one has $\Phi[\lim_i C_i] = \lim_i \Phi(C_i)$.*

Proof. We start with (1) and (2). The assumption clearly implies (3) when the $C_i(e)$'s are singleton's, i.e., the C_i's are mappings f_i to X_2, which is all we need for (1) and (2).

By this continuity property of Φ, $\Gamma(e)$ is the closure of $\{g_i(e) \mid i = 1,2,\dots\}$, where the g_i's are the measurable selections from Γ_1, which are the images by Φ of the sequence h_i of selections of Γ_2 furnished by Proposition 7.1. By Proposition 6.1 the mappings $\bar{g}_i$ from $(E_1, \mathcal{E}_1)$ to $\mathcal{K}_{X_1}$ defined by $\bar{g}_i(e) = \{g_i(e)\}$ are measurable. By Proposition 6.5 it follows that $\varphi_n(e) = \cup_{i=1}^n \bar{g}_i(e)$ is measurable, as a composition of measurable functions. Since $\Gamma(e)$, being closed in $\Gamma_1(e)$, is compact, and since the $\varphi_n(e)$ form an increasing sequence of compact subsets whose union is dense in $\Gamma(e)$, φ_n converges pointwise $(\mathcal{K}_{X_1})$ to Γ, hence Γ is measurable.

As for (3), let $C_0 = \lim C_i, K_i = [\Phi(C_i)](e)$. Since all K_i are included in the compact set $\Gamma(e)$, it is sufficient to show that any limit point x of a sequence $x_i \in K_i$ belongs to K_0, and that any $x \in K_0$ is the limit of some sequence $x_i \in K_i$. Let us begin with the first point: let $x_i = [\Phi(f_i)](e)$, f_i selection from C_i. By the assumption, $x \in [\Phi\ (\mathrm{Lim}\ f_i)](e)$, hence to $[\Phi(C_0)](e) = K_0$ (the argument, if spelled out, does not use the measurability of (Lim f_i), which is established anyway, independently of this, in Proposition 10). As for the second point, choose a measurable selection f from C_0. By Proposition 7.4 and Proposition 5.2, $e \to \{x \in C_i(e) \mid d(x, f(e)) = d(C_i(e), f(e))\}$ are a sequence of measurable maps to $\mathcal{K}_{X_2}$ – by composition – so they admit measurable selections $f_i(e)$ by Proposition 7.3. Then $f_i(e)$ converges pointwise to $f(e)$, so, by the remark in the beginning of this proof, $\Phi(f_i)$ converges pointwise to $\Phi(f)$. Since the points $[\Phi(f)](e)$ are dense in K_0, this proves (3).

Corollary 8'

1. *Let Γ be a map from $(\Omega, \mathcal{A})$ to $\mathcal{K}_X^*$, where X is a separable metric space. If the conclusion of Proposition 7.2 is true (i.e., if there exists a sequence of measurable maps f_i from $(\Omega, \mathcal{A})$ to X, such that $\{f_i(\omega)\}$ is dense in $\Gamma(\omega)$ for each ω), then Γ is measurable, for the Borel σ-field on $\mathcal{K}_X^*$.*
2. *Let X_1, X_2 be separable metric spaces. Let Γ be a measurable map from a measurable space $(E, \mathcal{E})$ to $\mathcal{K}_{X_1}$. Let f be a measurable map from the graph of Γ, endowed with the product σ-field, to X_2, such that $f(e, x_1)$ is continuous w.r.t. x_1 $\forall\ e \in E$. Denote by $F(e)$ the graph of $f(e, \cdot)$ in $X_1 \times X_2$. Then F is a measurable map to $\mathcal{K}_{X_1 \times X_2}$.*

Proof. (1) is the conclusion from the argument in Proposition 8.1.

(2) For a measurable selection g from Γ, let $[\Phi(g)](e) = \big(g(e), f(e, g(e))\big)$. Apply then Proposition 8.1.

Proposition 9 *Let X and Y be separable metric spaces, ψ an u.s.c. map from X to $\mathcal{K}_Y$, and Γ a measurable mapping from a measurable space $(E, \mathcal{E})$ to $\mathcal{K}_X$. Let $\Gamma'(e) = \{y \in Y \mid \exists x \in X \colon x \in \Gamma(e), y \in \psi(x)\}$. Then*

1. *Γ' is a measurable map to $\mathcal{K}_Y$.*
2. *(1) implies that the graph G' of Γ' is measurable in $(E \times Y, \mathcal{E} \otimes \mathcal{B})$, when $\mathcal{B}$ is the Borel σ-field on Y.*
3. *There exists an $(\mathcal{E} \otimes \mathcal{B})$-measurable map φ from G' to X such that $\varphi(e, y) \in \Gamma(e)$ and $y \in \psi(\varphi(e, y))$ everywhere.*

Proof. Let i_x denote the homeomorphism from X into $\mathcal{K}_X$ (Proposition 6.1), and j the embedding of $\mathcal{K}_X \times \mathcal{K}_Y$ into $\mathcal{K}_{X \times Y}$. Then $(i_x \times \psi)$ is u.s.c. from X to $\mathcal{K}_X \times \mathcal{K}_Y$, and the latter being a closed subspace of $\mathcal{K}_{X \times Y}$ (Proposition 6 7), $j \circ (i_x \times \psi)$ is u.s.c. from X to $\mathcal{K}_{X \times Y}$, hence induces an u.s.c. Borel map, say Γ_1, from $\mathcal{K}_X$ to $\mathcal{K}_{X \times Y}$ (Proposition 5.2 and Proposition 5.3). Replacing now Γ by the (Borel) composition $\Gamma_1 \circ \Gamma$, we are reduced to the case where X is a product $X_1 \times X_2$, and ψ the projection to X_1: we need a measurable φ from the graph of $\psi \circ \Gamma$ to X_2 such that $(x_1, \varphi(e, x_1)) \in \Gamma(e)$ everywhere.

There is again no loss in assuming X_1 and X_2 compact, by embedding.

(1) follows, by composition, from Proposition 6.4, applied to the projection.

(2) For each $k > 0$, consider a finite covering O_i of X_1 by open balls of radius k^{-1}, and their complements K_i with $K_0 = X_1$. Let $\varphi_k : \mathcal{K}_{X_1} \to \mathcal{K}_{X_1} : K \to \cap_{K_i \supseteq K} K_i$. Since $\{K \mid K \subseteq K_i\}$ is closed, φ_k is Borel, and φ_k converges pointwise (uniformly even) to the identity on $\mathcal{K}_{X_1}$. Then clearly the graph G_k of $\varphi_k \circ \Gamma'$ is measurable, and $\cap_k G_k = G'$.

(3) For $g \in G'$, denote its projections by $e(g)$ and $x(g) - e(\cdot)$ and $x(\cdot)$ are measurable. The map $x \to \bar{x} = x \times X_2 \in \mathcal{K}_{X_1 \times X_2}$ is clearly measurable too (continuous). Thus, by Proposition 6.5, $\Phi(g) = [(\Gamma \circ e)(g)] \cap [\bar{x}(g)]$ is a measurable map from G' to $\mathcal{K}_{X_1 \times X_2}$ (composition). Taking the composition of Φ with the measurable projection (Proposition 6.4) from $\mathcal{K}_{X_1 \times X_2}$ to $\mathcal{K}_{X_2}$, we get that the map $g\colon (e, x_1) \mapsto \{x_2 \in X_2 \mid (x_1, x_2) \in \Gamma(e)\} \in \mathcal{K}_{X_2}$ is measurable. Composing with the map of Proposition 7.3 yields now the result.

Proposition 10

1. *Let f_i be measurable functions from a measurable space $(E, \mathcal{E})$ to a separable metric space Y. Assume $\forall e \in E$ the sequence $f_i(e)$ is relatively compact in Y. Define $(Lim_i f_i)(e) = \{y \in Y \mid y$ is a limit point of $f_i(e)\}$. Then $Lim_i f_i$ is a measurable map to $\mathcal{K}_Y^*$.*
2. *If $Y = \mathbb{R}^\ell$, and L is a measurable map to $\mathcal{K}_Y^*$, there exist measurable functions on $E \times Y, g_0, g_1, \ldots, g_\ell$, such that $g_j(e, y) \in L(e)$ and such that $g_j(e, y) = y$ for $y \in L(e)$, and measurable functions on $E \times Y$, $\alpha_0, \alpha_1, \ldots, \alpha_\ell$ such that $\alpha_j \geqslant 0$ and $\Sigma\alpha_j = 1$ everywhere, and such that $\Sigma_{j=0}^{\ell} \alpha_j(e, y) g_j(e, y) = y$ for any y in the convex hull $\hat{L}(e)$ of $L(e)$.*

3. The map $C \to \hat{C}$ from $\mathcal{K}^*_{R^\ell}$ to itself is continuous.

Proof. (1) Let $C_k(e)$ denote the closure of $\{f_i(e) \mid i \geqslant k\}$. The proof of Proposition 8 shows that C_k is a measurable map to $\mathcal{K}^*_Y$. Since $(\mathrm{Lim} f_i)(e) = \cap_k C_k(e)$, C_k converges pointwise in $\mathcal{K}^*_Y$ to $(\mathrm{Lim} f_i)$. Hence the measurability of the latter.

(2) follows from Proposition 9 by using for X the space $Y^{\ell+1} \times \Delta$, where Δ is the ℓ-dimensional simplex, for Γ the map $e \to [L(e)]^{\ell+1} \times \Delta$, and for ψ the map $(y_0, y_1, \ldots, y_\ell; \alpha_0, \alpha_1, \ldots, \alpha_\ell) \to \{\Sigma_0^\ell \alpha_j y_j\}$. Indeed, by Caratheodory's theorem, every point of the convex hull of a set S in $\mathbb{R}^\ell$ can be written as $\Sigma_{j=0}^\ell \alpha_j y_j$, with $\alpha \in \Delta$ and $y_j \in S$. Finally, since $L(e)$ has measurable graph (Proposition 9.2), the $g_j(e, y)$ can be changed to y for $y \in L(e)$.

(3) is immediate from Caratheodory's result.

References

1. Aumann, R.J. (1965) Integrals of set-valued functions, *Journal of Mathematical Analysis and Applications* **12**, 1–12.
2. Bourbaki, N. (1958) *Eléments de mathematiques, Livre III: Topologie générale. Chapitre IX: Utilisation des nombres réels en topologie générale*, 2nd edition, Hermann, Paris.
3. Debreu, G. (1966) Integration of correspondences, *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press, Berkeley, pp. 351–372.
4. Doob, J. (1953) *Stochastic Processes*, John Wiley, New York.
5. Dvoretzky, A., Wald, A. and Wolfowitz, J. (1951) Relations among certain ranges of vector measures, *Pacific Journal of Mathematics* **1**, 59–74.
6. Lyapunov, A. (1940) Sur les fonctions-vecteurs complétement additives, *Bull. Acad. Sci. URSS* **6**, 465–478.
7. Mertens, J.-F. and Parthasarathy, T. (1987) Equilibria for discounted stochastic games, CORE Discussion Paper 8750, Université Catholique de Louvain, Louvain-la-Neuve, Belgium (Chapter 10 in this volume).
8. Rogers, C.A., Jayne, J.E., Dellacherie, C., Topsøe, F., Hoffman-Jørgensen, J., Martin, D.A., Kechris, A.S. and Stone, A.H. (1980) *Analytic Sets*, Academic Press, London.