

A GENERAL DENSENESS RESULT FOR RELAXED CONTROL THEORY

E.J. BALDER

A result by the author on the elimination of randomization (or relaxation) for variational problems is partially extended and then used to obtain a very general result on the denseness of the set of original control functions in the set of relaxed control functions. Also, a slight extension of Aumann's theorem on the integrals of multifunctions is shown to follow directly from the elimination result.

1. Introduction

It is well-known that the set of all original (or nonrandomized) control functions is a dense subset - for the usual topology - of the set of all relaxed (or randomized) control functions, in case the underlying measure space is nonatomic and the control space is metrizable and compact [16, IV.2.6], [12], [15, V]; see also [8, V.12 (Remark)] for a more abstract result of this kind and [9] for a vector version in the same spirit. Using the device of an Alexandrov (one point) compactification, Berliocchi and Lasry gave a density result for the noncompact case [6, Proposition II.7]. Inspired by their approach to the subject, the present author has recently expanded the domain of relaxed control theory, basing himself on a Hilbert cube compactification of the control space - which has to be a metrizable Lusin space as a rule - and a notion of tightness for

Received 2 August 1984.

Copyright Clearance Centre, Inc. Serial-fee code: 0004-9727/84
\$A2.00 + 0.00.

sets of relaxed control functions. The latter notion extends the classical tightness concept in topological measure theory, and the main (relative) compactness results for sets of relaxed control functions are thus seen to follow from a generalization of Prohorov's theorem; for example, see [2], [3]. The main purpose of this note is to show how this work also leads to an extremely general version of the above denseness result, which merely requires the control space to be completely regular and Suslin.

Our starting point will be a result on the elimination of randomization, which was obtained in [3, Lemma III]. This result generalizes similar results obtained in [6, II] and is of some interest in its own right, since it forms a one-sided analog - and a generalization - of a similar result by Dvoretzky, Wald and Wolfowitz [11] (see [4]) as well as a generalization of Aumann's important result on the integrals of multifunctions [1, Theorem 3], [13], as we shall show below. The elimination result of [3], which requires the control space to be metrizable Lusin, is here partially extended to the case where the control space is a completely regular Suslin space (nonmetrizable), and then used to obtain the announced denseness result.

2. Main results

Let (T, \mathcal{T}, μ) be a nonatomic finite measure space and S a completely regular Suslin space (Appendix A), equipped with the Borel σ -algebra $\mathcal{B}(S)$. An *original control function* is a $(T, \mathcal{B}(S))$ -measurable function from T into S . A *relaxed control function* is a transition probability with respect to (T, \mathcal{T}) and $(S, \mathcal{B}(S))$ [14, III]. The set of all original [relaxed] control functions is denoted by $M(T; S)$ [$R(T; S)$]. For any $u \in M(T; S)$ we define the *relaxation* $\varepsilon_u \in R(T; S)$ of u by

$$\varepsilon_u(t) \equiv \text{Dirac probability measure at } u(t) .$$

A *normal integrand* on $T \times S$ is a $T \times \mathcal{B}(S)$ -measurable function $g : T \times S \rightarrow (-\infty, +\infty]$ such that the function $g(t, \cdot)$ is lower semi-continuous on S for every $t \in T$ [8]; the set of all [nonnegative] normal integrands on $T \times S$ is denoted by $G(T; S)$ [$G^+(T; S)$]. A *Carathéodory integrand* on $T \times S$ is a function $g \in G(T; S)$ for which $g(t, \cdot)$ is continuous on S for every $t \in T$, such that there exists a

function $\phi \in L_1(T, \mathcal{T}, \mu)$, the set of all μ -integrable functions from T into \mathbb{R} , with

$$|g(t, s)| \leq \phi(t) \text{ for all } t \in T, s \in S.$$

We denote the set of all Caratheodory integrands on $T \times S$ by $G_C(T; S)$.

An *inf-compact* normal integrand on $T \times S$ is a function $h \in G^+(T; S)$ such that for every $t \in T$ the function $h(t, \cdot)$ is inf-compact on S . The set of all inf-compact normal integrands on $T \times S$ will be denoted by $H(T; S)$.

For any $g \in G(T; S)$, $\delta \in R(T; S)$ we define

$$I_g(\delta) \equiv \int_T \left[\int_S g(t, s) \delta(t)(ds) \right] \mu(dt),$$

provided that the integral makes sense at least as a quasi-integral [14].

As a consequence of this definition we find for relaxations

$$I_g(u) \equiv I_g(\varepsilon_u) = \int_T g(t, u(t)) \mu(dt),$$

for any $g \in G(T; S)$ and $u \in M(T; S)$, again with the provision that the integral makes sense. The *weak* topology on $R(T; S)$ is defined as the weakest topology for which all functionals $I_g : R(T; S) \rightarrow \mathbb{R}$,

$g \in G_C(T; S)$, are continuous. Note that in case S is a metrizable

compact space, this topology coincides with the usual one of relaxed control theory [16]. Note also that if T is a singleton the set $R(T; S)$

coincides with the set $P(S)$ of all probability measures on $(S, \mathcal{B}(S))$.

In that case $G_C(T; S)$ coincides with the set $C_b(S)$ of all bounded

continuous functions from S into \mathbb{R} , and the weak topology coincides

with its classical namesake [7], [10], as does the following notion: a

subset R_0 of $R(T; S)$ is defined to be *tight* if there exists

$h \in H(T; S)$ with

$$\sup_{\delta \in R_0} I_h(\delta) < +\infty$$

(see [3]). Only a passing role will be played by the following growth

relation: for any $g \in G(T; S)$ and $h \in H(T; S)$ we write $g^- \ll h$,

where $g^- \equiv \max(-g, 0)$ if for every $\varepsilon > 0$ there exists $\phi_\varepsilon \in L_1(T, T, \mu)$ such that

$$(1) \quad g^-(t, s) \leq \varepsilon h(t, s) + \phi_\varepsilon(t) \quad \text{for all } t \in T, s \in S.$$

Let us observe that this growth relation holds automatically for h and g if there exists $\phi \in L_1(T, T, \mu)$ with

$$g(t, s) \geq \phi(t) \quad \text{for all } t \in T, s \in S.$$

The following result [3, Lemma III], is our starting point. In [3] it was derived from Prohorov's theorem in $R(T; S)$, Lyapunov's theorem and certain extreme point considerations for relaxed control functions.

THEOREM 1. *Suppose that S is a metrizable Lusin space. If for $\delta_* \in R(T; S)$, $h \in H(T; S)$ and $\{g_1, \dots, g_m\} \subset G(T; S)$, $m \in \mathbb{N}$, the following conditions hold:*

$$(2) \quad I_h(\delta_*) < +\infty,$$

$$(3) \quad \bar{g}_j \ll h, \quad j = 1, \dots, m,$$

then there exists an original control function $u_ \in M(T; S)$ such that*

$$I_{g_j}(u_*) \leq I_{g_j}(\delta_*), \quad j = 1, \dots, m.$$

We shall now partially extend this result to the general situation where S is a completely regular Suslin space.

THEOREM 1'. *If for $\{g_1, \dots, g_m\} \subset G(T; S)$, $m \in \mathbb{N}$, there exists $\phi \in L_1(T, T, \mu)$ with*

$$(4) \quad g_j(t, s) \geq \phi(t) \quad \text{for all } t \in T, s \in S, j = 1, \dots, m,$$

then there exists for every $\delta_ \in R(T; S)$ an original control function $u_* \in M(T; S)$ such that*

$$(5) \quad I_{g_j}(u_*) \leq I_{g_j}(\delta_*), \quad j = 1, \dots, m.$$

In deriving Theorem 1' from Theorem 1 we shall need the following result on "automatic tightness" of singletons in $R(T; S)$.

LEMMA 2. *For every $\delta_* \in R(T; S)$ there exists $h \in H(T; S)$ such*

that

$$I_h(\delta_*) < +\infty ;$$

as a matter of fact, h can be chosen so as to depend only on the variable s .

Proof. Let ν be the finite measure on $(S, \mathcal{B}(S))$, defined as follows:

$$\nu(B) \equiv \int_T \delta_*(t)(B) \mu(dt) , \quad B \in \mathcal{B}(S) ;$$

see [14, III.2]. Since S is Suslin, the measure ν is tight [10, III.69]. As was shown in [3, Example 2.5], this implies the existence of an inf-compact function $h : S \rightarrow [0, +\infty]$ with

$$\int_S h(s) \nu(ds) < +\infty . \quad \square$$

Proof of Theorem 1'. As a first step, we prove (5) in case S is a metrizable Lusin space. The result then follows *a fortiori* from Theorem 1, since (2) holds by Lemma 2 and (3) by condition (4); see our comments following (1). As our second step, we now consider the general case and show how it can be reduced to the situation of the first step. By Appendix A there exist a Polish space P (which is certainly metrizable Lusin) and a continuous surjection $\pi : P \rightarrow S$. We define $\{g_1^\pi, \dots, g_m^\pi\} \subset G(T; P)$ by

$$(6) \quad g_j^\pi(t, p) \equiv g_j(t, \pi(p)) .$$

Our claim is that there exists $\delta^* \in \mathcal{R}(T; P)$ such that for μ almost everywhere, $t \in T$,

$$(7) \quad \delta^*(t) \text{ is the image under } \pi \text{ of the measure } \delta_*(t) ;$$

see [10, II.11]. Suppose for a moment that (7) holds; then it is easy to finish the proof: by Lemma 2 there exists $h \in H(T; P)$ such that $I_h(\delta^*) < +\infty$. By (4) and (6) we see that the situation considered in the first step obtains; as a consequence, there exists a function $u^* \in M(T; P)$ with

$$I_{g_j^\pi}(u^*) \leq I_{g_j^\pi}(\delta^*) , \quad j = 1, \dots, m .$$

By a well-known formula on integration under a change of variables [10, II.12] we have for every $t \in T$, $j = 1, \dots, m$,

$$\int_P g_j^\pi(t, p) \delta^*(t)(dp) = \int_S g_j(t, s) \delta_*(t)(ds) ,$$

by virtue of (6)-(7). Setting $u_* \equiv \pi \circ u^*$, we obtain (5). Hence, the crux of the proof lies in showing the validity of (7). Let Γ be the multifunction from T into the set $\mathcal{P}(P)$ of all probability measures on $(P, \mathcal{B}(P))$, defined by

$$\Gamma(t) \equiv \{ \nu \in \mathcal{P}(P) : \text{the image of } \nu \text{ under } \pi \text{ is } \delta_*(t) \} .$$

By applying [10, III.45] in the same way as was done in the proof of [10, III.69], we see that the values of Γ are nonempty. Also, since $\mathcal{P}(S)$, equipped with the weak topology, is a Suslin space (Appendix A) and the set $C_b(S)$ separates the elements of $\mathcal{P}(S)$ [10, III.54] (here we use the fact that S is completely regular), there exists by [δ , III.31] a countable subset D of $C_b(S)$ which also separates the elements of $\mathcal{P}(S)$. Hence, the desired measurability of the graph of Γ follows directly by [14, III.2] and [δ , III.14], since now

$$\Gamma(t) = \left\{ \nu \in \mathcal{P}(P) : \int_S c d\delta_*(t) = \int_P c \circ \pi d\nu \text{ for all } c \in D \right\}$$

for every $t \in T$, and $\nu \mapsto \int_P c \circ \pi d\nu$, $c \in D$, is clearly continuous on the Polish (by [10, III.60]) space $\mathcal{P}(P)$. Thus, the stage has been set for an application of the von Neumann-Aumann measurable selection theorem [δ , III.22]: there exists a $(T, \mathcal{B}(\mathcal{P}(P)))$ -measurable function $\delta^* : T \rightarrow \mathcal{P}(P)$ such that $\delta_*(t) \in \Gamma(t)$ for μ almost everywhere $t \in T$ (here we use a standard μ almost everywhere modification argument to convert the result of [δ , III.22]). This shows that (7) is true, which is all that remained to be done. \square

It is perhaps illuminating to point out why the approach follows here in proving Theorem 1' is bound to fail if one tries to extend other results of [3] to the present situation with a completely regular Suslin control

space. Here we had an automatic tightness result at our disposal (Lemma 2). In general, this is not so: if we set out with a tightness condition for a general subset R_0 of $R(T; S)$ (not finite) - to be more precise, the condition that

$$\sup_{\delta_* \in R_0} I_h(\delta_*) < +\infty$$

for some $h \in H(T; S)$, then we do *not* have that the function h^π , defined by

$$h^\pi(t, p) \equiv h(t, \pi(p))$$

belongs to $H(T; P)$. Hence the original tightness of the transition probabilities δ_* does not carry over to the transition probabilities δ^* which correspond in the sense of (7).

An argument, similar to the proof of Theorem 1', can also be found in the Appendix to [5].

As an immediate consequence of Theorem 1', we have the following denseness result for original control functions.

COROLLARY 3. *If $\{g_1, \dots, g_m\} \subset G(T; S)$, $m \in \mathbb{N}$, satisfies (4), then we have, for every $\{\alpha_1, \dots, \alpha_m\} \subset \mathbb{R}$,*

$$R_0 = \text{weak closure of } \{\varepsilon_u : u \in M(T; S), I_{g_j}(u) \leq \alpha_j, j = 1, \dots, m\},$$

where

$$R_0 \equiv \{\delta \in R(T; S) : I_{g_j}(\delta) \leq \alpha_j, j = 1, \dots, m\}.$$

In particular, we have

$$R(T; S) = \text{weak closure of } \{\varepsilon_u : u \in M(T; S)\}$$

and for every multifunction Δ from T into S , having $T \times \mathbb{B}(S)$ -measurable graph and closed values

$R_\Delta = \text{weak closure of}$

$$\{\varepsilon_u : u \in M(T; S), u(t) \in \Delta(t) \text{ for } \mu \text{ almost every } t \in T\},$$

where

$$R_{\Delta} \equiv \{ \delta \in R(T; S) : \delta(t) (\Delta(t)) = 1 \text{ for } \mu \text{ almost every } t \in T \} .$$

Proof. Let $\delta_{*} \in R_0$ be arbitrary, and let N be an arbitrary neighborhood of δ_{*} . By definition of the weak topology, there exist $\{l_1, \dots, l_n\} \subset G_C(T; S)$, $n \in \mathbb{N}$, such that, for every $\delta \in R(T; S)$,

$$(8) \quad |I_{l_i}(\delta) - I_{l_i}(\delta_{*})| < 1, \quad i = 1, \dots, n, \text{ implies } \delta \in N .$$

Now apply Theorem 1' to

$$\{g_1, \dots, g_m, l_1, \dots, l_n, -l_1, \dots, -l_n\} \subset G(T; S) .$$

It follows that there exists $u_{*} \in M(T; S)$ with

$$(9) \quad \begin{aligned} I_{l_i}(u_{*}) &= I_{l_i}(\delta_{*}), \quad i = 1, \dots, n, \\ I_{g_j}(u_{*}) &\leq I_{g_j}(\delta_{*}) \leq \alpha_j, \quad j = 1, \dots, m. \end{aligned}$$

By (8)-(9) we have $\varepsilon_{u_{*}} \in N$, which proves the main statement. The specializations follow by taking $m = 1$, $\alpha_1 = 0$ and $g_1 = 0$ for the first case, and for the second one

$$g_1(t, s) \equiv \begin{cases} 0 & \text{if } s \in \Delta(t), \\ +\infty & \text{if } s \notin \Delta(t). \end{cases} \quad \square$$

This denseness result generalizes similar results by Warga [16, IV.2.6, IV.3.10 (second part)] and Sainte-Beuve [15, V] (her results cover more than just denseness). These results apply only to a metrizable compact control space S . Further, Corollary 3 also generalizes a denseness result by Berliocchi and Lasry [6, Proposition II.7], which requires the control space to be locally compact and countable at infinity.

Note that (9) proves our earlier claim that Theorems 1, 1' form one-sided generalizations of a well-known result by Dvoretzky, Wald and Wolfowitz [11, §4] on the elimination of randomization (they also require the control (or action) space to be metrizable and compact).

We finish by showing how Theorem 1 (or 1') directly implies a slight extension of Aumann's theorem on the integrals of multifunctions [1, Theorem 3].

COROLLARY 4. If for $\{g_1, \dots, g_m\} \subset G(T; \mathbb{R}^n)$, $m, n \in \mathbb{N}$, and the multifunction F from T into \mathbb{R}^n ,

(10) $g_j(t, \cdot)$ is concave on $\text{co } F(t)$ for every $t \in T$, $j = 1, \dots, m$,

and if there exist $\phi, \psi \in L_1(T, \mathcal{T}, \mu)$ such that

(11) $g_j(t, s) \geq \phi(t)$ for all $t \in T$, $s \in F(t)$,

(12) $s \geq \psi(t)$ (coordinatewise) for all $t \in T$, $s \in F(t)$,

then, denoting by I_F the set of all μ -integrable selectors of F , that is

$$I_F \equiv \left\{ f \in L_1^n(T, \mathcal{T}, \mu), f(t) \in F(t) \text{ for } \mu \text{ almost every } t \in T \right\},$$

we have for every $\{\alpha_1, \dots, \alpha_m\} \subset \mathbb{R}$,

$$(13) \left\{ \int_T f d\mu : f \in I_F, I_{g_j}(f) \leq \alpha_j, j = 1, \dots, m \right\} \\ = \left\{ \int_T f d\mu : f \in I_{\text{co}F}, I_{g_j}(f) \leq \alpha_j, j = 1, \dots, m \right\},$$

where $I_{\text{co}F}$ denotes the set of all μ -integrable selectors of the multifunction $t \mapsto \text{co } F(t) \equiv [\text{convex hull of } F(t)]$.

Proof. One inclusion is trivial. Thus, if the set following the equality sign in (13) is empty, the proof is finished. Otherwise, let $q \in \mathbb{R}^n$ be such that, for some $f \in I_{\text{co}F}$,

(14) $I_{g_j}(f) \leq \alpha_j$, $j = 1, \dots, m$, and $\int_T f d\mu = q$.

By a standard application of Carathéodory's theorem and the von Neumann-Aumann theorem (see [8, pp. 101-102]) there exist $(T, \mathcal{B}(\mathbb{R}^n))$ -measurable selectors f_1, \dots, f_{n+1} of F and $(T, \mathcal{B}(\mathbb{R}))$ -measurable functions $\alpha_1, \dots, \alpha_{n+1} : T \rightarrow [0, 1]$ such that

$$f(t) = \sum_{i=1}^{n+1} \alpha_i(t) f_i(t) ,$$

(15)

$$\sum_{i=1}^{n+1} \alpha_i(t) = 1 \text{ for } \mu \text{ almost everywhere } t \in T .$$

By μ -integrability of f and (12), (15) the functions f_1, \dots, f_{n+1} must also be μ -integrable. Now for $\delta_* \in R(T; \mathbb{R}^n)$, defined by

$$\delta_*(t) \equiv \sum_{i=1}^{n+1} \alpha_i(t) \varepsilon_{f_i}(t) ,$$

we have by (10), (14) and (15),

$$I_{g_j}(\delta_*) = \sum_{i=1}^{n+1} \int_T \alpha_i(t) g_j(t, f_i(t)) \mu(dt) \leq \alpha_j , \quad j = 1, \dots, m ,$$

$$\int_T \left[\int_{\mathbb{R}^n} s \delta_*(t) (ds) \right] \mu(dt) = q .$$

(17)

Let s^i denote the i th component and $|s|$ the Euclidean norm of any $s \in \mathbb{R}^n$. Denote by $\bar{\phi} \in L_1(T, T, \mu)$ the function $t \mapsto \max_{1 \leq i \leq n+1} |f_i(t)|$

and define

$$g_0(t, s) \equiv \begin{cases} 0 & \text{if } s \in \{f_1(t), \dots, f_{n+1}(t)\} , \\ +\infty & \text{otherwise,} \end{cases}$$

$$l_i(t, s) \equiv \max\{s^i, -\bar{\phi}(t)\} , \quad i = 1, \dots, n ,$$

$$l_{n+i}(t, s) \equiv \max\{-s^i, -\bar{\phi}(t)\} , \quad i = 1, \dots, n .$$

We may apply Theorem 1 (or 1') to

$$\{g_0, g_1, \dots, g_m, l_1, \dots, l_{2n}\} \subset G(T; \mathbb{R}^n) .$$

It follows that there exists $u_* \in M(T; \mathbb{R}^n)$ with

$$I_{g_0}(u_*) \leq I_{g_0}(\delta_*) = 0 , \quad I_{g_j}(u_*) \leq I_{g_j}(\delta_*) \leq \alpha_j , \quad j = 1, \dots, m ,$$

$$I_{L_i}(u_*) \leq q^i, \quad I_{L_{n+i}}(u_*) \leq -q^i, \quad i = 1, \dots, n.$$

This shows u_* to be a μ -integrable selector of F , in view of the above definition of g_0, L_1, \dots, L_{2n} , with $\int_T u_* d\mu = q$. We conclude that

(13) holds. \square

Appendix

We prove here the well-known fact that the set $P(S)$ of all probability measures on $(S, \mathcal{B}(S))$ is a Suslin space in case the space S is completely regular and Suslin.

A Hausdorff topological space S is said to be *Suslin* if there exist a Polish (separable metrizable and complete) space P and a continuous surjection $\pi : P \rightarrow S$ [10, III.67]. Now the set $P(P)$, equipped with the weak topology, is a Polish space by [10, III.60]. We define the function $\bar{\pi} : P(P) \rightarrow P(S)$ by

$$\bar{\pi}(v) \equiv \text{image of the measure } v \text{ under } \pi;$$

see [10, II.11]. By [10, III.45] the function $\bar{\pi}$ is a surjection from $P(P)$ into $P(S)$, as is seen by imitating the argument used to prove [10, III.69]. Finally, since, for all $c \in C_b(S)$, $v \in P(P)$,

$$\int_S c d(\bar{\pi}(v)) = \int_P c \circ \pi dv,$$

we see that $\bar{\pi}$ is also continuous from $P(P)$ into $P(S)$. This proves the result, since $P(S)$ is a Hausdorff space by virtue of the complete regularity of S [10, III.54].

References

- [1] Robert J. Aumann, "Integrals of set-valued functions", *J. Math. Anal. Appl.* 12 (1965), 1-12.
- [2] E.J. Balder, "On a useful compactification for optimal control problems", *J. Math. Anal. Appl.* 72 (1979), 391-398.
- [3] E.J. Balder, "A general approach to lower semicontinuity and lower closure in optimal control theory", *SIAM J. Control Optim.* 22 (1984), 570-598.

- [4] E. J. Balder, "Elimination of randomization in statistical decision theory reconsidered", *J. Multivariate Anal.* (to appear).
- [5] E. J. Balder, "Mathematical foundations of statistical decision theory: a modern viewpoint", *Statistics and decisions* (to appear).
- [6] Henri Berliocchi and Jean-Michel Lasry, "Intégrales normales et mesures paramétrées en calcul des variations", *Bull. Soc. Math. France* 101 (1973), 129-184.
- [7] Patrick Billingsley, *Convergence of probability measures* (John Wiley & Sons, New York, London, Sydney, 1968).
- [8] C. Castaing, M. Valadier, *Convex analysis and measurable multifunctions* (Lecture Notes in Mathematics, 580. Springer-Verlag, Berlin, Heidelberg, New York, 1977).
- [9] Phan Van Chuong, "Vector versions of a density theorem and applications to problems of control theory", *Travaux du Séminaire d'Analyse Convexe*, Montpellier 1981, Exposé 19, 19.1-19.19.
- [10] Claude Dellacherie, Paul-André Meyer, *Probabilités et potentiel* (Hermann, Paris, 1975; English Transl., North-Holland, Amsterdam, 1978).
- [11] A. Dvoretzky, A. Wald and J. Wolfowitz, "Elimination of randomization in certain statistical decision procedures and zero-sum two-person games", *Ann. Math. Statist.* 22 (1951), 1-21.
- [12] A. Ghouila-Houri, "Sur la généralisation de la notion de commande d'un système guidable", *Rev. Franc. d'Inf. Rech. Opér.* 4 (1967), 7-32.
- [13] Werner Hildenbrand, *Core and equilibria of a large economy* (Princeton University Press, Princeton, 1974).
- [14] Jacques Neveu, *Bases mathématiques du calcul des probabilités* (Masson, Paris, 1964; English Transl. Holden-Day San Francisco, London, Amsterdam, 1965).
- [15] M.-F. Sainte-Beuve, "Some topological properties of vector measures with bounded variation and its applications", *Ann. Mat. Pura Appl.* (4) 116 (1978), 317-379.

- [16] J. Warga, *Optimal control of differential and functional equations*
(Academic Press, New York, London, 1972).

Mathematical Institute,
University of Utrecht,
PO Box 80.010,
3508 TA Utrecht,
The Netherlands.