Journal of Convex Analysis
Volume 25 (2018), No. 3, 789–815

Conglomerability and Representations

Gianluca Cassese
Dept. of Economics, Statistics and Management, Università Milano Bicocca, Via Bicocca degli Arcimboldi 8, 20126 Milano, Italy
gianluca.cassese@unimib.it

Received: August 27, 2016
Accepted: May 6, 2017

We prove results concerning the representation of linear functionals as integrals of a given random quantity X. The existence of such representation is related to the notion of conglomerability, originally introduced by de Finetti and Dubins. We show that this property has interesting applications in probability and in analysis. These include a version of Skorohod theorem, a proof that Brownian motion assumes whatever family of finite dimensional distributions upon a change of the probability measure and a version of the extremal representation theorem of Choquet.

Keywords: Choquet integral representation, conglomerability, Riesz representation, Skhorohod representation, vector lattice.

2010 Mathematics Subject Classification: Primary 28A25; Secondary 46A22, 52A41.

1. Introduction

In this paper we study the classical problem of the integral representation of linear functionals in situations of considerable generality in which a direct application of classical techniques may not be feasible. Conglomerability is then necessary and sufficient to conveniently transform the original problem into one in which such a representation is possible, although in an indirect way. Although the fields in which our results may be fruitfully applied are disparate, we were motivated by the problem of *existence of companions* that arises in several places in probability and statistics.

To be more explicit, let S and Ω be given, non empty sets and $\mathscr{H}$ a family of real valued functions on S. Fix a pair (X, m), with X a mapping of Ω into S and m a positive, finitely additive set function on some ring of subsets of Ω and such that $\{h(X) : h \in \mathscr{H}\} \subset L^1(m)$. Borrowing terminology from Dubins and Savage [14], a pair (X', μ) on some given set Ω' is a companion to (X, m), relatively to the family $\mathscr{H}$, if it solves the equation

$$h(X') \in L^1(\mu) \quad \text{and} \quad \int h(X)dm = \int h(X')d\mu, \quad h \in \mathscr{H}. \tag{1}$$

The collection $\mathscr{H}$ is interpreted as a model of the information available.

ISSN 0944-6532 / $ 2.50

Finding a correct statistical model X' for some given experimental sample is a problem fitting into (1): set $S = \Omega = \mathbb{R}$, let X be the identity, m the sample distribution and each $h \in \mathscr{H}$ a statistic. Given a predictive marginal m on an algebra $\mathscr{A}$, a similar problem in Bayesian statistics is that of finding a parametric family $\mathcal{Q} = \{Q_\theta : \theta \in \Theta\}$ of probabilities and a prior λ on the parameter space Θ such that

$$m(A) = \int_\Theta Q_\theta(A) d\lambda, \quad A \in \mathscr{A}. \tag{2}$$

Dubins [15] proved long ago that the existence of a disintegration formula similar to (2) is equivalent to *conglomerability*, a notion originally due to de Finetti [11] that has then remained undeservedly neglected outside a limited number of distinguished authors (which include Schervisch et al. [26], Hill and Lane [20] or Zame [27]). The conglomerability property, we believe, may be formulated in more general terms than those in which it was originally stated and it may be applied to more ambitious problems in probability and in analysis than those for which it had been originally devised.

We found it useful to write problem (1) in more abstract terms, replacing the integral on left hand side with a linear functional and modeling the action of X' on $\mathscr{H}$ as a linear transformation. We solve this version of our problem in Theorem 3.3 obtaining a special integral representation for a conglomerative linear functional on an arbitrary vector space. The absence of any structure on the underlying space, save linearity, makes the claim significantly more general than classical Riesz representation theorems. In this functional analytic formulation, conglomerability may be nicely restated as a geometric property. In Corollary 3.4 we show that if Φ and Ψ are two sets of positive functionals on a vector lattice then Φ is Ψ-conglomerative if and only if each $\phi \in \Phi$ is the barycenter of a measure supported by Ψ. In Corollary 3.5 we obtain a generalization of the original theorem of Choquet [7].

Theorem 3.3 admits a large number of implications in probability and statistics, the most immediate of which is the existence of companions. Further to existence we establish conditions for the representing measure μ to possess additional properties, such as countable additivity or absolute continuity with respect to some given, reference set function. As a corollary we obtain that, relatively to continuous functions, *any* X is companion to a normally distributed random quantity and Brownian motion can assume whatever family of finite dimensional distributions on $\mathbb{R}$ upon an appropriate choice of the underlying probability. We also provide applications to the classical Skhorohod representation theorem in the case in which S is separable.

In the closing section we prove some results concerning the representation of convex functions as integrals. We show that any convex function on $\mathbb{R}$ decomposes into the sum of a piecewise linear component and an integral part, a representation curiously near to the one popular in mathematical finance as a model for option prices.

All proofs are quite simple and, despite a natural interest for countable additivity, they are obtained by exploiting the theory of the finitely additive integral in which the measurability constraint is much less burdensome. We hope to disprove thus, at least partially, the harsh judgment of Bourgin [5, p. 173] that *"an integral representation theory based on finitely additive measures is virtually useless"*.

2. Notation and preliminaries

Throughout the paper the symbol $\mathfrak{F}(\Omega, S)$ (resp. $\mathfrak{F}(\Omega)$) denotes the family of functions mapping Ω into S (resp. into $\mathbb{R}$) and $\mathfrak{F}$ is replaced with $\mathfrak{L}$, $\mathscr{C}$ or $\mathscr{C}_K$ when restricting to linear, continuous or continuous functions with compact support, respectively. A collection $\{f_y : y \in Y\} \subset \mathfrak{F}(X, S)$ is also written as a function $f \in \mathfrak{F}(X \times Y, S)$ with $f(x, y) = f_y(x)$ and viceversa. If $f \in \mathfrak{F}(\Omega, S)$ and $A \subset \Omega$ the symbols $f|A$ and $f[A]$ designate the restriction of f to A and the image of A under f. A sublattice $\mathscr{H}$ of $\mathfrak{F}(S)$ is *Stonean* if $h \in \mathscr{H}$ implies $h \wedge 1 \in \mathscr{H}$, where $1 \in \mathfrak{F}(S)$ indicates the function constantly equal to 1.

If $\mathscr{A}$ is a ring of subsets of Ω, then $\mathscr{S}(\mathscr{A})$ and $\mathfrak{B}(\mathscr{A})$ denote the families of $\mathscr{A}$ simple functions and its closure in the topology of uniform convergence, respectively. $fa(\mathscr{A})$ (resp. $fa(\Omega)$) is the space of real valued, finitely additive set functions on $\mathscr{A}$ (resp. 2^{Ω}) and $ba(\mathscr{A})$ the subspace of set functions of bounded variation. A pair $(\mathscr{A}, \lambda)$ with $\mathscr{A}$ a ring of subsets of Ω and $\lambda \in fa(\mathscr{A})_+$ is a measure structure on Ω.

We recall a few definitions and facts relative to the finitely additive integral (see [4] and [17]) given a measure structure $(\mathscr{A}, \lambda)$ on Ω[1]. $X \in \mathfrak{F}(\Omega)$ is λ-*measurable* if there exists a sequence $\langle X_n \rangle_{n\in\mathbb{N}}$ in $\mathscr{S}(\mathscr{A})$ that λ-converges to X, i.e. such that

$$\lim_n \lambda^*(|X_n - X| > c) = 0 \quad \text{for every} \quad c > 0 \tag{3}$$

where the set function λ^* and its conjugate λ_* are defined (with the convention $\inf \varnothing = \infty$) as

$$\lambda^*(E) = \inf_{\{A\in\mathscr{A}:E\subset A\}} \lambda(A) \quad \text{and} \quad \lambda_*(E) = \sup_{\{B\in\mathscr{A}:B\subset E\}} \lambda(B), \qquad E \subset \Omega. \tag{4}$$

If S is a topological space, X is λ-*tight* if for all $\varepsilon > 0$ there exists $K \subset S$ compact such that $\lambda^*(X \notin K) < \varepsilon$. X is λ-*integrable*, $X \in L^1(\lambda)$, if there is a sequence $\langle X_n \rangle_{n\in\mathbb{N}}$ in $\mathscr{S}(\mathscr{A})$ that λ-converges to X and is Cauchy in $L^1(\lambda)$; we

[1] To be formal, we depart from the classical theory of Dunford and Schwartz which has an extended real valued set function on an algebra of sets as its starting point. Our notion of a simple function is obtained from theirs after restricting to the family of sets of finite measure, a ring, and coincides therefore with the notion of *integrable simple functions* of Dunford and Schwartz. Thus, our notion of measurability is more restrictive than that of *total measurability* given in [17, III.2.10] although integrable functions are defined by Dunford and Schwartz as being measurable in our restrictive sense.

then write $\int X d\lambda$ or $\int_\Omega X(\omega)d\lambda(\omega)$. We notice that if $A, B \subset \Omega$ and $f \in L^1(\lambda)$, then

$$\mathbb{1}_A \leq f \leq \mathbb{1}_B \quad \text{implies} \quad \lambda^*(A) \leq \int f d\lambda \leq \lambda_*(B). \tag{5}$$

The following collections are important:

$$D(X,\lambda) = \big\{t > 0 : \lim_n \lambda_*(X > t - 2^{-n}) = \lim_n \lambda_*(X > t + 2^{-n})\big\}, \tag{6a}$$

$$\mathscr{R}_0(X,\lambda) = \big\{\{X > t\} : t \in D(X,\lambda)\big\} \cup \big\{\{-X > u\} : u \in D(-X,\lambda)\big\}, \tag{6b}$$

$$\mathscr{A}(\lambda) = \big\{E \subset \Omega : \lambda^*(E) = \lambda_*(E) < \infty\big\}. \tag{6c}$$

There is clearly just one extension of λ to $\mathscr{A}(\lambda)$ and X is λ-measurable if and only if it is measurable with respect to such extension, which we shall denote, accordingly, again by λ. A sequence $\langle X_n\rangle_{n\in\mathbb{N}}$ in $L^1(\lambda)$ converges to X in norm if and only if it λ-converges to X and is Cauchy in the norm of $L^1(\lambda)$, [17, III.3.6].

We shall use the following results on measurability and integrability of a positive function.

Lemma 2.1. *Let $X \geq 0$. X is λ-measurable if and only if it is λ-tight and either (i) $\infty > \lambda_*(X > s) \geq \lambda^*(X \geq t)$ for all $0 < s < t$, (ii) $\mathscr{R}_0(X,\lambda) \subset \mathscr{A}(\lambda)$, or (iii) the set $\big\{t > 0 : \{X > t\} \in \mathscr{A}(\lambda)\big\}$ is dense in $\mathbb{R}_+$.*

Proof. If X is λ-measurable it is λ-tight, [22, p. 190]. Choose $\langle X_k\rangle_{k\in\mathbb{N}}$ in $\mathscr{S}(\mathscr{A})$ λ-convergent to X, fix $s, \eta > 0$ and $A_k^\eta \in \mathscr{A}$ such that $\{|X - X_k| \geq \eta\} \subset A_k^\eta$ and $\lambda(A_k^\eta) \leq \lambda^*(|X - X_k| \geq \eta) + 2^{-k}$.

$$\{X \geq s + 2\eta\} \subset \{X_k \geq s + \eta\} \cup A_k^\eta \subset \{X > s\} \cup A_k^\eta$$

so that $\lambda^*(X \geq s + 2\eta) \leq \lambda\big(\{X_k \geq s + \eta\} \cup A_k^\eta\big) \leq \lambda_*(X > s) + \lambda\big(A_k^\eta\big)$ and $\lambda^*(X \geq s + 2\eta) < \infty$. Assume (i). If $t \in D(X,\lambda)$ then $\infty > \lambda_*(X > t) = \lim_n \lambda_*(X > t - 2^{-n}) \geq \lambda^*(X \geq t) \geq \lambda^*(X > t)$. $(ii)\Rightarrow(iii)$ is obvious. Assuming (iii), choose $\{0 = t_0^n \leq t_1^n \leq \ldots \leq t_{I_n}^n \leq t_{I_n+1}^n = 2^n\}$ such that $\{X > t_i^n\} \in \mathscr{A}(\lambda)$ for $i = 1, \ldots, I_n$ and $\sup_{0\leq i\leq I_n} |t_i^n - t_{i+1}^n| < 2^{-n}$. Define

$$X_n = \sum_{i=1}^{I_n - 1} t_i^n \mathbb{1}_{\{t_i^n < X \leq t_{i+1}^n\}} \in \mathscr{S}\big(\mathscr{A}(\lambda)\big). \tag{7}$$

Then $\{|X - X_n| \geq 2^{-n}\} \subset \{X > 2^{n-1}\}$ so that X_n λ-converges to X whenever X is λ-tight. $\square$

Lemma 2.2. *Let $X \geq 0$. $X \in L^1(\lambda)$ if and only if $\int_0^\infty \lambda_*(X > t)dt = \int_0^\infty \lambda^*(X > t)dt < \infty$. Then,*

$$\int X d\lambda = \int_0^\infty \lambda_*(X > t)dt. \tag{8}$$

Proof. Assume $\int \lambda_*(X > t)dt = \int \lambda^*(X > t)dt < \infty$. Then X is λ-tight and $\{t \in \mathbb{R} : \{X > t\} \in \mathscr{A}(\lambda)\}$ is dense in $\mathbb{R}_+$ so that X is λ-measurable. As in (7) we can construct an increasing sequence $\langle X_n \rangle_{n\in\mathbb{N}}$ in $\mathscr{S}(\mathscr{A}(\lambda))$ such that $0 \le X_n \le X$ and λ-converges to X. But then,

$$\infty > \int_0^\infty \lambda_*(X > t)dt \ge \lim_n \int_0^\infty \lambda(X_n > t)dt = \lim_n \int X_n d\lambda = \int X d\lambda \quad (9)$$

as $\langle X_n \rangle_{n\in\mathbb{N}}$ is Cauchy in $L^1(\lambda)$. Assume conversely that $X \in L^1(\lambda)$ and take $b > a > \varepsilon > 0$. If $\langle X_n \rangle_{n\in\mathbb{N}}$ in $\mathscr{S}(\mathscr{A})$ converges to X in $L^1(\lambda)$, then

$$\begin{aligned}\int_{a+\varepsilon}^{b+\varepsilon} \lambda^*(X > t)dt &\le \int_a^b \lambda(X_n > t)dt + (b-a)\lambda^*(|X - X_n| > \varepsilon)\\ &\le \int_{a-\varepsilon}^{b-\varepsilon} \lambda_*(X > t)dt + 2(b-a)\lambda^*(|X - X_n| > \varepsilon)\end{aligned}$$

by [4, 3.2.8.(iii)]. Thus, $\int_a^b \lambda_*(X > t)dt = \int_a^b \lambda^*(X > t)dt$ and

$$\begin{aligned}\int_a^b \lambda^*(X > t)dt = \lim_n \int_a^b \lambda(X_n > t)dt &= \lim_n \int (b \wedge X_n - a)^+ d\lambda\\ &= \int (b \wedge X - a)^+ d\lambda\,.\end{aligned}$$

Thus $\int_0^\infty \lambda_*(X > t)dt = \int_0^\infty \lambda^*(X > t)dt = \int X d\lambda < \infty$ and (8) holds. $\square$

Proving uniqueness of the set function generating a given class of integrals requires to identify a minimal measure structure associated with a given family of functions. This we do by writing, for two measurable structures $(\mathscr{A}, \lambda)$ and $(\mathscr{B}, \xi)$ on the same underlying space,

$$(\mathscr{A}, \lambda) \preceq (\mathscr{B}, \xi) \quad \text{whenever} \quad \mathscr{A} \subset \mathscr{B}(\xi) \quad \text{and} \quad \xi|\mathscr{A} = \lambda. \quad (10)$$

Lemma 2.3. *Let $\mathscr{H}$ be a convex cone in $\mathfrak{F}(\Omega)_+$ with the property that $f, g \in \mathscr{H}$ implies $f \wedge g$, $f \wedge 1 \in \mathscr{H}$ and let $\phi \in \mathfrak{L}(\mathscr{H})$. The family $\mathfrak{M}(\phi)$ of measure structures $(\mathscr{A}, \lambda)$ on Ω satisfying*

$$\mathscr{H} \subset L^1(\lambda) \quad \textit{and} \quad \int h d\lambda = \phi(h) \qquad h \in \mathscr{H}, \quad (11)$$

is either empty or contains a minimal element, $(\mathscr{R}_\phi, \lambda_\phi)$.

Proof. Assume that $(\mathscr{A}, \lambda) \in \mathfrak{M}(\phi)$ and denote by $\mathscr{R}_\phi$ the smallest ring containing

$$\mathscr{R}_{0,\phi} = \big\{\{h > t\} : h \in \mathscr{H},\ t \in D(h, \lambda)\big\}. \quad (12)$$

Suppose that $(\mathscr{B},\xi)\in\mathfrak{M}(\phi)$. Fix $h\in\mathscr{H}$ and consider the classical inequality

$$\mathbb{1}_{\{h>a\}}\geq\frac{h\wedge b-h\wedge a}{b-a}\geq\mathbb{1}_{\{h\geq b\}}\,,\quad h\in\mathscr{H},b>a>0. \tag{13}$$

As the inner term belongs to the linear span of $\mathscr{H}$, $\infty>\lambda_*(h>a)\geq\xi^*(h\geq b)$, by (5). Choosing a and b conveniently and interchanging λ with ξ we establish that $D(h,\lambda)=D(h,\xi)$ and that

$$\lambda^*(h\geq t)=\xi^*(h\geq t)=\xi_*(h>t)=\lambda_*(h>t)\qquad t\in D(h,\lambda).$$

Thus, $\mathscr{R}_{0,\phi}\subset\mathscr{B}(\xi)$ and λ and ξ coincide on $\mathscr{R}_{0,\phi}$ and therefore on the collection

$$\mathscr{E}=\big\{E\subset\Omega:\mathbb{1}_E\in\mathscr{S}(\mathscr{R}_{0,\phi})\big\}.$$

To show that $\mathscr{R}_{0,\phi}$ is closed with respect to intersection, for $i=1,2$ pick $h_i\in\mathscr{H}$ and $t_i\in D(h_i,\lambda)$. Fix $t_1\wedge t_2\geq\eta>0$, define $h_\eta=\big(h_1-(t_1-\eta)\big)^+\wedge\big(h_2-(t_2-\eta)\big)^+$ and observe that $h_\eta=(h_1+h_2\wedge(t_2-\eta))\wedge(h_2+h_1\wedge(t_1-\eta))-$ $-(h_1\wedge(t_1-\eta)+h_2\wedge(t_2-\eta))\in\operatorname{span}(\mathscr{H})$. Since the sets $D(h_\eta,\lambda)$ are dense in $\mathbb{R}_+$, choose

$$\delta\in(0,t_1\wedge t_2]\cap\mathbb{Q}\cap\bigcap_{\eta\in\mathbb{Q}\cap(0,t_1\wedge t_2]}D(h_\eta,\lambda).$$

Then $\delta\in D(h_\delta,\lambda)$, $h_\delta\in\mathscr{H}$ and $\{h_1>t_1\}\cap\{h_2>t_2\}=\{h_\delta>\delta\}$. But then $\mathscr{E}$ too is closed with respect to intersection and this fact together with the linear structure of $\mathscr{S}(\mathscr{R}_{0,\phi})$ imply in turn that $\mathscr{E}$ is also closed with respect to set difference and, from $\mathbb{1}_{E_1\cup E_2}=\mathbb{1}_{E_1}+\mathbb{1}_{E_2\setminus E_1}$, to union as well. In other words, λ and ξ coincide on the ring $\mathscr{E}$ which contains $\mathscr{R}_{0,\phi}$ and *a fortiori* on $\mathscr{R}_\phi$. Let $h\in\mathscr{H}$, $t>s>0$ and $\lambda_\phi=\lambda|\mathscr{R}_\phi$. Then, h is λ_ϕ-tight because $h\in L^1(\lambda)$ and there are $t',s'\in D(h,\lambda)$ with $t>t'>s'>s$ and therefore such that $\lambda_{\phi*}(h>s)\geq\lambda_\phi(h>s')\geq\lambda_\phi(h>t')\geq\lambda^*_\phi(h\geq t)$. By Lemma 2.1 h is thus λ_ϕ-measurable and therefore $\int hd\lambda_\phi=\int hd\lambda$. □

Although the minimal structure $(\mathscr{R}_\phi,\lambda_\phi)$ will generally depend on ϕ, the generated σ ring corresponds to the usual notion, as $D(h,\lambda)$ is dense.

Lemma 2.4. *Let $g\in\mathfrak{F}(\Omega)_+$ be λ-measurable and define the ring $\mathscr{R}_g=\big\{A\in\mathscr{A}(\lambda):g\mathbb{1}_A\in L^1(\lambda)\big\}$. There exists a unique $\lambda_g\in fa(\mathscr{R}_g)_+$ such that*

$$\int f\lambda_g=\int fgd\lambda,\quad f\in\mathfrak{B}(\lambda),\ fg\in L^1(\lambda). \tag{14}$$

Proof. (14) implies $\lambda_g(A)=\int\mathbb{1}_Agd\lambda$ for every $A\in\mathscr{R}_g$ and thus uniqueness. In proving (14) we may assume $f\in\mathfrak{B}(\lambda)_+$. Let $\langle f_n\rangle_{n\in\mathbb{N}}$ be an increasing sequence in $\mathscr{S}(\mathscr{A}(\lambda))$ such that $0\leq f_n\leq f$ and f_n converges to f uniformly, obtained as in (7). Then f_n is λ- and λ_g-convergent to f. Moreover, f_n and f_ng are Cauchy sequence in $L^1(\lambda_g)$ and $L^1(\lambda)$. □

3. Integral representation of linear functionals

First we make the notion of conglomerability precise.

Definition 3.1. Let $\mathscr{H}$ be a vector space. Then $\phi \in \mathfrak{L}(\mathscr{H})$ is said to be *conglomerative with respect to* $T \in \mathfrak{F}\big(\mathscr{H}, \mathfrak{F}(\Omega)\big)$ (or T*-conglomerative*) if $\phi(h) < 0$ implies $\inf_\omega (Th)(\omega) < 0$ for all $h \in \mathscr{H}$.

T-conglomerative linear functionals form a $\mathscr{H}$-closed convex cone in $\mathfrak{L}(\mathscr{H})$, i.e. in the topology induced by $\mathscr{H}$ on $\mathfrak{L}(\mathscr{H})$. Another key property is the following:

Definition 3.2. Let $\mathscr{H}$ be a vector space. A map $T \in \mathfrak{F}\big(\mathscr{H}, \mathfrak{F}(\Omega)\big)$ is said to be *directed* if:

$$\forall h \in \mathscr{H},\ \exists h' \in \mathscr{H} \quad \text{such that} \quad |Th| \le Th'. \tag{15}$$

Property (15) holds in two special cases of interest: (α) when $\mathscr{H}$ is a vector lattice and T is positive and (β) when $T[\mathscr{H}] \subset \mathfrak{B}(\Omega)$ and $\sup_h \inf_\omega (Th)(\omega) > 0$ – e.g. if $T[\mathscr{H}]$ contains the constants.

Most of the results in this paper follow from the next claim.

Theorem 3.3. *Let $\mathscr{H}$ be a vector space, $\phi \in \mathfrak{L}(\mathscr{H})$ and let $T \in \mathfrak{L}\big(\mathscr{H}, \mathfrak{F}(\Omega)\big)$ be directed. Write $L = \big\{f \in \mathfrak{F}(\Omega) : |f| \le Th \text{ for some } h \in \mathscr{H}\big\}$. ϕ is T-conglomerative if and only if there exist (i) $F^\perp \in \mathfrak{L}(L)_+$ with $F^\perp[L \cap \mathfrak{B}(\Omega)] = \{0\}$ and (ii) a measure structure $(\mathscr{R}, \mu)$ on Ω with*

$$L \subset L^1(\mu) \quad \text{and} \quad \phi(h) = F^\perp\big(Th\big) + \int Th d\mu, \quad h \in \mathscr{H}. \tag{16}$$

Moreover, $(\mathscr{R}, \mu)$ can be chosen such that:

(a) $\|\mu\| = 1$ *if and only if* $\inf_\omega (Th)(\omega) \le \phi(h)$ *for all* $h \in \mathscr{H}$,

(b) μ *is countably additive if* $\limsup_n \phi(h_n) \le 0$ *for all sequences* $\langle (h_n, f_n) \rangle_{n \in \mathbb{N}}$ *in* $\mathscr{H} \times L$ *satisfying*

$$1 \ge f_n \downarrow 0 \text{ and } \limsup_n \ \sup\big\{\phi(g) : g \in \mathscr{H},\ Tg \le |f_n - Th_n|\big\} \le 0, \tag{17}$$

(c) μ *is L_0-maximal for given $L_0 \subset L \cap \mathfrak{B}(\Omega)$, i.e. if $(\mathscr{R}', \mu')$ is another measurable structure satisfying* (16) *and if $\int f d\mu' \ge \int f d\mu$ for each $f \in L_0$ then $\int f d\mu = \int f d\mu'$ for all $f \in L_0$.*

Proof. $T[\mathscr{H}]$ is a majorizing subspace of the vector lattice L, by (15). If ϕ is T-conglomerative

$$F\big(Th\big) = \phi(h), \quad h \in \mathscr{H} \tag{18}$$

implicitly defines a positive linear functional on $T[\mathscr{H}]$. By [1, Theorem 1.32], F extends as a positive linear functional (still denoted by F) to the whole of L. For each $\alpha \subset \mathscr{H}$ finite, let $h_\alpha \in \mathscr{H}$ be such that $Th_\alpha \geq \bigvee_{h \in \alpha} |Th|$, $\Omega_\alpha = \{Th_\alpha \neq 0\}$ and define $I_\alpha \in \mathfrak{F}\big(L, \mathfrak{F}(\Omega_\alpha)\big)$ by letting

$$I_\alpha(f)(\omega) = \frac{f(\omega)}{Th_\alpha(\omega)}, \quad f \in L,\ \omega \in \Omega_\alpha .$$

Let also

$$L_\alpha = \{f \in L : |f| \leq c\, Th_\alpha \text{ for some } c > 0\} \quad \text{and} \quad H_\alpha = I_\alpha[L_\alpha]. \tag{19}$$

H_α is a sublattice of $\mathfrak{B}(\Omega_\alpha)$ containing the constants; $f, g \in L_\alpha$ and $I_\alpha(f) \geq I_\alpha(g)$ imply $f \geq g$. Thus, upon writing

$$U_\alpha\big(I_\alpha(f)\big) = F(f), \quad f \in L_\alpha \tag{20}$$

we obtain yet another positive, linear functional U_α on H_α. [6, Theorem 1] implies

$$U_\alpha\big(I_\alpha(f)\big) = \int I_\alpha(f) d\bar{m}_\alpha, \quad f \in L_\alpha \tag{21}$$

for some $\bar{m}_\alpha \in ba(\Omega_\alpha)_+$. Let $m_\alpha(A) = \bar{m}_\alpha(A \cap \Omega_\alpha)$ for each $A \subset \Omega$. By Lemma 2.4, we can write (with the convention $0/0 = 0$)

$$F(f) = \int \frac{f}{Th_\alpha} \mathbb{1}_{\Omega_\alpha} dm_\alpha = \int f d\bar{\mu}_\alpha, \quad f \in L_\alpha \cap \mathfrak{B}(\Omega) \tag{22}$$

with $\bar{\mu}_\alpha = m_{\alpha,g}$ defined as in (14) with $g = \mathbb{1}_{\Omega_\alpha}/Th_\alpha$. Given that $L_\alpha \cap \mathfrak{B}(\Omega)$ is a Stonean lattice, we deduce from Lemma 2.3 the existence of a minimal measure structure $(\mathscr{R}_\alpha, \mu_\alpha)$ supporting the representation (22). Define $\mathscr{R} = \bigcup_\alpha \mathscr{R}_\alpha$ and $\mu(A) = \lim_\alpha \mu_\alpha(A)$ for all $A \in \mathscr{R}$. $\alpha \subset \alpha'$ implies $L_\alpha \subset L_{\alpha'}$, $(\mathscr{R}_\alpha, \mu_\alpha) \preceq (\mathscr{R}_{\alpha'}, \mu_{\alpha'})$ as well as the martingale restriction

$$\mu_\alpha = \mu_{\alpha'} | \mathscr{R}_\alpha = \mu | \mathscr{R}_\alpha, \quad \alpha \subset \alpha'. \tag{23}$$

But then for each $f \in L_\alpha$ with $f \geq 0$,

$$\begin{aligned} F(f) &= \lim_k F(f \wedge k) + \lim_k F\big((f-k)^+\big) \\ &= \lim_k \int (f \wedge k) d\mu + F^\perp(f) = \int f d\mu + F^\perp(f) \end{aligned} \tag{24}$$

where we have set $F^\perp(f) = \lim_k F\big((f-k)^+\big)$ and the inequality $\mu^*(f > k) \leq k^{-1} \int f \wedge k d\mu \leq k^{-1} F(f)$ induces the conclusion that $f \wedge k$ is μ-convergent to f and is Cauchy in $L^1(\mu)$. $\int |f| d\mu \leq F(|f|)$ follows from (24) and implies $L \subset L^1(\mu)$. (16) is a consequence of (18) and (24). Necessity is obvious as the right hand side of (16) defines a positive linear functional on L.

(a). Suppose that $\phi(h) < a < \inf_\omega Th(\omega)$ for some $a \in \mathbb{R}$ and $h \in \mathscr{H}$. Then, by (16) and properties of $F^\perp$, $a > \int Th d\mu \geq a\|\mu\|$ which is contradictory if

$\|\mu\| = 1$. Conversely, define $\hat{\phi} \in \mathfrak{L}(\mathbb{R}\times\mathscr{H})$ and $\hat{T} \in \mathfrak{L}\big(\mathbb{R}\times\mathscr{H}, \mathfrak{F}(\Omega)\big)$ implicitly by letting

$$\hat{\phi}(r,h) = r + \phi(h) \quad \text{and} \quad \hat{T}(r,h) = r + T(h) \qquad (r,h) \in \mathbb{R}\times\mathscr{H}. \tag{25}$$

By assumption $\hat{\phi}$ is $\hat{T}$-conglomerative and thus admits a pair $\hat{F}^{\perp}$ and $\hat{\mu}$ as above. Therefore, for every $(r,h) \in \mathbb{R}\times\mathscr{H}$,

$$r + \phi(h) = \hat{F}^{\perp}(r + Th) + \int (r + Th) d\hat{\mu} = \hat{F}^{\perp}(Th) + \int (r + Th) d\hat{\mu}. \tag{26}$$

Letting $h = \emptyset$ we deduce $\|\hat{\mu}\| = 1$ and, from this, $\phi(h) = \hat{F}^{\perp}(Th) + \int Th d\hat{\mu}$ for every $h \in \mathscr{H}$.

(b). Fix a sequence $\langle f_n\rangle_{n\in\mathbb{N}}$ as in (17). By [1, Theorem 1.33] the extension of F from $T[\mathscr{H}]$ to L constructed above may be chosen such that we obtain $\inf_{h\in\mathscr{H}} F\big(|f - Th|\big) = 0$ for every $f \in L$. Let $F^{\perp}$ and μ be the corresponding components of F according to (16). Thus, for each $n \in \mathbb{N}$, let $h_n \in \mathscr{H}$ be such that $F(|f_n - Th_n|) \le 2^{-n}$. If $g \in \mathscr{H}$ and $Tg \le |f_n - Th_n|$, then

$$\phi(g) = F(Tg) \le F(|f_n - Th_n|) \le 2^{-n}.$$

Thus $\langle (h_n, f_n)\rangle_{n\in\mathbb{N}}$ satisfies (17) and, by assumption, $\limsup_n \phi(h_n) \le 0$. The inequality $\int f_n d\mu = F(f_n) \le \phi(h_n) + F(|f_n - Th_n|)$ then proves that the functional $f \to \int f d\mu$ is a Daniel integral on the Stonean lattice $L \cap \mathfrak{B}(\Omega)$ and it may thus be represented by some countably additive measure structure $(\mathscr{R}, \hat{\mu})$ on Ω. To prove that $\hat{\mu}$ agrees with μ over the whole of L it is enough to remark that when $f \in L$ and $f \ge 0$, then

$$\hat{\mu}^*(f > k) \le k^{-1} \int f \wedge k d\hat{\mu} = k^{-1} \int f \wedge k d\mu \le k^{-1} F(f)$$

and therefore $f \wedge k$ converges to f in $L^1(\hat{\mu})$.

(c). For each α in a directed set $\mathfrak{A}$, let $F_\alpha \in \mathfrak{L}(L)_+$ be such that $F_\alpha(Th) = \phi(h)$ for each $h \in \mathscr{H}$. Given that F_α is conglomerative with respect to the identity on L, it is of the form

$$F_\alpha(f) = F_\alpha^{\perp}(f) + \int f d\mu_\alpha, \quad f \in L \tag{27}$$

with $F_\alpha^{\perp}[L \cap \mathfrak{B}(\Omega)] = \{0\}$ and $(\mathscr{R}_\alpha, \mu_\alpha)$ a measure structure on Ω such that $L \subset L^1(\mu_\alpha)$. Observe that if $f \in L$ then there exists $h \in \mathscr{H}$ such that $|f| \le Th$ and thus such that $F_\alpha(|f|) \le \phi(h)$. The net $\langle F_\alpha\rangle_{\alpha\in\mathfrak{A}}$ admits then a subnet (still indexed by α for convenience) such that

$$F(f) = \lim_\alpha F_\alpha(f), \quad f \in L.$$

Since F is positive we write it as $F(f) = F^\perp(f) + \int f d\mu$. If the net $\langle \mu_\alpha \rangle_{\alpha \in \mathfrak{A}}$ is increasing on $L_0 \subset L \cap \mathfrak{B}(\Omega)$ then

$$\lim_\alpha \int f d\mu_\alpha = \lim_\alpha F_\alpha(f) = F(f) = \int f d\mu, \quad f \in L_0.$$

It is clear that $\int f d\mu \geq \int f d\mu_\alpha$ for each $\alpha \in \mathfrak{A}$ and $f \in L_0$. By Zorn lemma this proves the existence of a representing measure μ which is L_0-maximal. □

The idea that some linear functionals may be representable as integrals only after a convenient transformation seems to have been first formulated and studied by Choquet in [8] and [9] with the aim of extending the classical Riesz-Markoff theorem to situations in which it cannot be applied directly. Choquet assumes that Ω is a compact topological space, $\mathscr{H}$ a positively generated linear space of extended real-valued, continuous functions on Ω and takes T to be of the simple form $T(h) = h/g$ with g a positive, continuous function on Ω with $\{g = \pm\infty\}$ a nowhere dense subset of Ω. This construction permits to characterize positive linear functionals on $\mathscr{H}$ as sums of a summable family of submeasures [9, Theorem 42].

Theorem 3.3 is more strictly related to another, well known result of Choquet, the extremal representation theorem, that was originally proved in [7] and later variously extended and reformulated (see, [10], [23] or [24] for an overview of this literature). To see this connection it is enough to fix $\Omega = \Psi \subset \mathfrak{L}(\mathscr{H})$ and to define $T \in \mathfrak{L}(\mathscr{H}, \mathfrak{F}(\Psi))$ as the map associating each $h \in \mathscr{H}$ with the (restriction to Ψ of the) corresponding evaluation e_h on $\mathfrak{F}(\mathscr{H})$. It is then easily seen that conglomerability may be nicely restated in geometric terms as the condition

$$\phi \in \overline{\text{con}}^{\mathscr{H}}(\Psi), \tag{28}$$

i.e. ϕ being an element of the closed, conical hull of Ψ, the closure being in the $\mathscr{H}$ topology. Likewise, the inequality $\inf_\omega Th(\omega) \leq \phi(h)$ for all $h \in \mathscr{H}$ is equivalent to the condition $\phi \in \overline{\text{co}}^{\mathscr{H}}(\Psi)$.

In the light of these remarks the following result becomes obvious.

Corollary 3.4. *Let $\mathscr{H}$ be a vector lattice and $\Psi \subset \mathfrak{L}(\mathscr{H})_+$. $\phi \in \overline{\text{con}}^{\mathscr{H}}(\Psi)$ if and only if there exist (i) $\phi^\perp \in \mathfrak{L}(\mathscr{H})$ with $\phi^\perp(h) \geq 0$ when $\inf_\psi \psi(h) > -\infty$ and (ii) a measure structure $(\mathscr{R}, \mu)$ on Ψ such that*

$$e_h|\Psi \in L^1(\mu) \quad \textit{and} \quad \phi(h) = \phi^\perp(h) + \int_\Psi \psi(h) d\mu, \quad h \in \mathscr{H}. \tag{29}$$

Moreover, one can choose μ to be a probability if and only if $\phi \in \overline{\text{co}}^{\mathscr{H}}(\Psi)$.

The order structure of $\mathscr{H}$ guarantees that the map T defined above is directed, as in (α).

To compare this result with the classical extremal or barycentrical representation, we remark that the conical structure and the choice of the $\mathscr{H}$ topology make conglomerability a very weak restriction not requiring compactness nor boundedness and *a fortiori* not relying on the existence of extreme points. The first to obtain a proof of Choquet's theorem without compactness was Edgar [18, Theorem p. 355] who considered a bounded, closed, convex, separable subset of a Banach space possessing the Radon Nikodym property based on norm convergence of vector valued martingales.

Another version of Choquet's theorem is constructed starting from condition (β) for directedness of T and requires boundedness.

Corollary 3.5. *Let $\mathscr{H} \subset \mathfrak{F}(S)$ be a vector subspace, $\phi \in \mathfrak{L}(\mathscr{H})$ and let $V \subset S$ satisfy* $\sup_{v\in V} |h(v)| < \infty$ *for all $h \in \mathscr{H}$. Then,*

$$\phi(h) \geq \inf_{v\in V} h(v), \quad h \in \mathscr{H} \tag{30}$$

if and only if there exists a probability structure $(\mathscr{R}, \mu)$ on V such that

$$h \in L^1(\mu) \quad \textit{and} \quad \phi(h) = \int_V h(v) d\mu, \quad h \in \mathscr{H}. \tag{31}$$

Proof. Consider the vector space $\mathscr{H} \times \mathbb{R}$ as acting on S via $(h, r)(s) = h(s)+r$. In the notation of Theorem 3.3, let $\Omega = V$, $T(h, r) = (h, r)|V$ and $\hat{\phi}(h, r) = \phi(h) + r$. By (β), T is directed as $T[\mathscr{H} \times \mathbb{R}]$ is a subset of $\mathfrak{B}(\Omega)$ containing the constants. (30) is equivalent to $\hat{\phi}(h, r) \geq \inf_v (Th)(v)$, i.e. to the representation of $\hat{\phi}$ in the form (16) form some probability structure $(\mathscr{R}, \mu)$ and with $F^\perp = 0$ as $\hat{T}[\mathscr{H} \times \mathbb{R}] \subset \mathfrak{B}(V)$. (31) follows upon restricting to elements of the form $(h, 0)$. The converse implication is obvious. □

A clear example in which (30) holds is the one in which $\mathscr{H}$ consists of affine functions and $\phi(h) = h(u)$ for some $u \in \overline{\text{co}}^{\mathscr{H}}(V)$. We highlight that Corollary 3.5 does not require the elements of $\mathscr{H}$ to be continuous. A drawback in our approach is the difficulty in characterizing the mapping $u \to \mu_u$. In the case in which $\mathscr{H}$ is a Stonean sublattice, however, the minimality property is enough to imply that $u \in V$ if and only if μ_u is the point mass measure at u.

The classical proof of Choquet's theorem takes advantage of the duality between Radon measures and continuous functions on a compact topological space and of its implication that regular, countably additive probabilities form a compact set in the corresponding topology. In the present setting no such compactness property can be invoked in proving Theorem 3.3 but it is rather one of its implications.

Corollary 3.6. *Let L be a sublattice of $\mathfrak{F}(\Omega)$, $\chi \in \mathfrak{F}(L)$ and choose $L_0 \subset L \cap \mathfrak{B}(\Omega)$. The set*

$$\mathscr{M} = \left\{\mu \in fa(\Omega)_+ : L \subset L^1(\mu) \textit{ and } \int f d\mu \leq \chi(f) \textit{ for all } f \in L\right\} \tag{32}$$

is compact in the topology induced by $L \cap \mathfrak{B}(\Omega)$. Moreover, $\mathscr{M}$ contains a L_0-maximal element.

Proof. Extract from a net in $\mathscr{M}$ a subnet $\langle \mu_\alpha \rangle_{\alpha \in \mathfrak{A}}$ such that $\phi(f) = \lim_\alpha \int f d\mu_\alpha$ exists for every $f \in L$. Clearly, $\phi \in \mathfrak{L}(L)_+$ so that, by Theorem 3.3, it is of the form (16). Compactness follows from $F^\perp[L \cap \mathfrak{B}(\Omega)] = \{0\}$. From a net which converges on L_0 it is then possible to extract a subnet which converges on $L \cap \mathfrak{B}(\Omega)$ to some element of $\mathscr{M}$. Any family in $\mathscr{M}$ which is linearly ordered by L_0 admits then an upper bound in $\mathscr{M}$ with respect to such order. The existence of a maximal element follows from Zorn lemma. □

The non trivial implication is the inclusion $L \subset L^1(\mu)$ for any cluster point μ of $\mathscr{M}$.

4. Finitely additive companions

In this section we return to the problem of the existence of companions. To make the connection with the work of Dubins [15] more transparent we start by establishing the following version of the problem considered by him:

Corollary 4.1. (Dubins) *Let $(\mathscr{B}, \lambda)$ be a measure structure on $\Omega \times S$ and $\mathscr{H}$ a Stonean sublattice of $L^1(\lambda)$. Let $\{\sigma_\omega : \omega \in \Omega\} \subset \mathfrak{L}(\mathscr{H})_+$ be such that $\sup_\omega \sigma_\omega(h^+ \wedge 1) < \infty$ for all $h \in \mathscr{H}$. Then,*

$$\int h d\lambda < 0 \quad \textit{implies} \quad \inf_\omega \sigma_\omega(h) < 0, \quad h \in \mathscr{H} \tag{33}$$

is necessary and sufficient for the existence of a measure structure $(\mathscr{R}, \mu)$ on Ω such that

$$\sigma_h \in L^1(\mu) \quad \textit{and} \quad \int h d\lambda = \int \sigma_\omega(h) d\mu, \quad h \in \mathscr{H}. \tag{34}$$

Proof. Apply Theorem 3.3 with $T \in \mathfrak{F}\big(\mathscr{H}, \mathfrak{F}(\Omega)\big)$ defined as $Th(\omega) = \sigma_\omega(h)$. We get the representation (16) with $F^\perp$ vanishing on $\sigma[\mathscr{H}] \cap \mathfrak{B}(\Omega)$. But then, since $T(h^+ \wedge n) \in \mathfrak{B}(\Omega)$, we conclude:

$$\int h d\lambda = \lim_n \int (h \wedge n) d\lambda = \lim_n \int \sigma_\omega(h \wedge n) d\mu \leq \int \sigma_\omega(h) d\mu$$

and the claim follows from the linearity of $\mathscr{H}$. □

Dubins considered the case $\mathscr{H} = \mathfrak{B}(S)$ and Ω a partition of S. The family σ indexed by Ω is a *strategy* in his terminology and a probability λ admitting the representation (34) is called *strategic*. Theorem 1 in [15] states that λ is σ-strategic if and only if it is σ-conglomerative.

For the rest of this section, fix a measure structure $(\mathscr{A}, m)$ on a set Ω.

Theorem 4.2. *Let $X \in \mathfrak{F}(\Omega, S)$ and $\mathscr{H}$ a Stonean vector sublattice of $\mathfrak{F}(S)$ such that $\{h(X) : h \in \mathscr{H}\} \subset L^1(m)$. Let $X' \in \mathfrak{F}(\Omega', S)$. There is equivalence between the condition*

$$\int h(X)dm < 0 \quad \textit{implies} \quad \inf_{\omega' \in \Omega'} h\big(X'(\omega')\big) < 0, \quad h \in \mathscr{H} \tag{35}$$

and the existence of a minimal measure structure $(\mathscr{R}, \mu)$ on Ω' satisfying

$$h(X') \in L^1(\mu) \quad \textit{and} \quad \int h(X)dm = \int h(X')d\mu, \quad h \in \mathscr{H}. \tag{36}$$

If $X'[\Omega']$ is closed in the topological space S and $\mathscr{H} \subset \mathscr{C}(S)$ then μ is countably additive if either (i) $\mathscr{H} \subset \mathscr{C}_K(S)$, (ii) X' is μ-tight or (iii) X is m-tight and $m_*\big(X \notin X'[\Omega']\big) = 0$.

Proof. (35) is equivalent to ϕ being T-conglomerative with $\phi(h) = \int h(X)dm$ and $Th = h(X')$ for every $h \in \mathscr{H}$. Thus, (36) follows from (16) after noting that, in the present setting, $\phi(h) = \lim_k \phi(h \wedge k)$ for every $h \in \mathscr{H}_+$.

Let $X'[\Omega']$ be closed and $\langle h_n \rangle_{n\in\mathbb{N}}$ a sequence in $\mathscr{H} \subset \mathscr{C}(S)$ with $h_n(X')$ decreasing to 0, i.e. h_n decreasing to 0 on $X'[\Omega']$. We claim that (i), (ii) or (iii) imply $\lim_n \int h_n(X')d\mu = 0$. If $\mathscr{H} \subset \mathscr{C}_K(S)$, then in computing such limit one may replace S with some compact subset so that (i) follows from (ii). Fix $\varepsilon > 0$. Under (ii) there exists $K' \subset S$ compact and $B'^c \in \mathscr{R}$ such that $B' \subset \{X' \in K'\}$ and

$$\int h_n(X')d\mu \le \int h_n(X')\mathbb{1}_{B'}d\mu + \varepsilon, \quad n \in \mathbb{N}.$$

But then $\lim_n \sup_{\omega'\in B'} h_n(X') \le \lim_n \sup_{s\in X'[\Omega']\cap K'} h_n(s) = 0$, by Dini's theorem. Under (iii), we can find an extension $\bar{m}$ of m to the minimal ring containing the set $F = \{X \notin X'[\Omega']\}$ such that $\bar{m}(F) = 0$. We can also find $K \subset S$ compact and $B^c \in \mathscr{A}$ such that $B \subset \{X \in K\}$ and that

$$\int h_n(X')d\mu = \int h_n(X)dm = \int h_n(X)d\bar{m} \le \int h_n(X)\mathbb{1}_{B\backslash F}d\bar{m} + \varepsilon, \quad n \in \mathbb{N}$$

so that again $\lim_n \sup_{\omega\in B\backslash F} h_n(X) \le \lim_n \sup_{s\in X'[\Omega']\cap K} h_n(s) = 0$. In either case the positive linear functional $\int h(X')d\mu$ on the Stonean lattice $\mathscr{H}[X']$ is a Daniell integral and it may be represented via a countably additive set function. Since μ is minimal, it must then be countably additive. □

In the absence of restrictions on μ, the existence of companions is guaranteed under a weak condition such as (35), namely if X is X'-conglomerative. An obvious companion to any X is the identity map on $\Omega' = S$. Given that being companion (relatively to the one given family $\mathscr{H}$) is a transitive property, the problem in Theorem 4.2 may be simplified with no loss of generality by assuming that X is the identity map on $\Omega = S$. In this case, if m consists of sample frequencies, then the condition $m^*(X'[\Omega']) = 0$ sufficient for m to be X'-conglomerative means that all the observations in the given sample must belong to the range of X'.

The existence of a countably additive companion was proved under (ii) by Dubins and Savage [14, p. 190], for the case $\Omega = \Omega' = S = \mathbb{R}$, and has then been revived and extended to the case $S = \mathbb{R}^n$ by Karandikar, [21] and [22], who used it in the proof of finitely additive limit theorems. The conditions for the existence of a countably additive companion obtained in Theorem 4.2 may be employed to refine the results of the preceding section. In particular if the set Ψ in Corollary 3.4 is $\mathscr{H}$-compact then in (29) one has $\phi^\perp = 0$ and μ can be chosen to be countably additive.

An interesting issue is how to construct an auxiliary state space on which the existence of a countably additive companion is guaranteed. The key point is to obtain the closed range condition.

Theorem 4.3. *Let S be a metric space, $X \in \mathfrak{F}(\Omega, S)$ and $\mathscr{H} = \big\{h \in \mathscr{C}(S) : h(X) \in L^1(m)\big\}$. There exist a set $\tilde{\Omega}$, an embedding κ of Ω into $\tilde{\Omega}$, a countably additive measure structure $(\mathscr{R}, \mu)$ on $\tilde{\Omega}$, and a map $\tilde{X} \in \mathfrak{F}(\tilde{\Omega}, S)$ such that (a) $\tilde{X} \circ \kappa = X$ and (b) $(\tilde{X}, \mu)$ and (X, m) are companions relatively to $\mathscr{H} \cap \mathscr{C}_K(S)$ (or to $\mathscr{H}$, if X is m-tight). If m is a probability then so is μ and (c) for all $h \in \mathscr{C}(S)$ $h(X) \in L^1(m)$ and $h(\tilde{X}) \in L^1(\mu)$ are equivalent conditions.*

In particular, if X is a sequence in $\mathfrak{F}(\Omega)$ with $\sup_n |X_n|$ m-tight and $(\tilde{X}, \mu)$ its countably additive companion relatively to $\mathscr{H}$, then if X_n m-converges (resp. converges in $L^1(m)$) to 0 then $\tilde{X}_n$ μ-converges (resp. converges in $L^1(\mu)$) to 0. If m is a probability, then the converse implication holds as well.

Proof. Let $\tilde{\Omega}$ consist of all sequences $\tilde{\omega} = \langle \omega_k \rangle_{k \in \mathbb{N}}$ in Ω along which X converges and define

$$\tilde{X}(\tilde{\omega}) = \lim_k X(\omega_k), \qquad \tilde{\omega} = \langle \omega_k \rangle_{k \in \mathbb{N}} \in \tilde{\Omega}. \tag{37}$$

The map $\kappa \in \mathfrak{F}(\Omega, \tilde{\Omega})$ defined by letting $\kappa(\omega)_n = \omega$ for each $\omega \in \Omega$ and $n \in \mathbb{N}$ is an embedding. Moreover, $\tilde{X} \circ \kappa = X$, $\tilde{X}[\tilde{\Omega}]$ is closed and contains $X[\Omega]$, so that X is $\tilde{X}$-conglomerative relatively to $\mathscr{H}$. The claim follows from Theorem 4.2.(i) (or from Theorem 4.2.(iii), if X is m-tight).

For the last claim, let $S = \mathbb{R}^{\mathbb{N}}$. By assumption for each ε there exists $a > 0$ such that

$$m^*\big(|X| \notin [0, a]^{\mathbb{N}}\big) \le m^*\Big(\sup_n |X_n| > a\Big) \le \varepsilon$$

so that X is m-tight. It then admits a countably additive companion $\tilde{X} \in \mathfrak{F}(\tilde{\Omega}, S)$, by the first part of this proof. We notice that the properties of tightness, convergence in measure or in $L^1(m)$ are unchanged if m is replaced with some positive extension to the ring $\{A \subset \Omega : m^*(A) < \infty\}$. In other words, we can assume that $(\tilde{X}, \mu)$ is a countably additive companion to $(X, \bar{m})$ relatively to $\bar{\mathscr{H}} = \{g \in \mathscr{C}(S) : g(X) \in L^1(\bar{m})\}$. For $b > a$ there exists $g \in \mathscr{C}(\mathbb{R})$ such that $\mathbb{1}_{\{b<x\}} \leq g(x) \leq \mathbb{1}_{\{a<x\}}$. Fix $h_n(X_1, X_2, \ldots) = g(|X_n|)$. If X_n is m-convergent to X_1, then $h_n \in \bar{\mathscr{H}}$ so that

$$\begin{aligned} m^*(|X_n| > a) &\geq \int h_n(X_1, X_2, \ldots) d\bar{m} = \int h_n(\tilde{X}_1, \tilde{X}_2, \ldots) d\mu \\ &\geq \mu^*(|\tilde{X}_n| > b) \end{aligned} \tag{38}$$

so that $\tilde{X}_n$ is μ-convergent to 0. Likewise, if convergence takes place in $L^1(m)$ and $h'_n(x_1, x_2, \ldots) = |x_n|$ then $h'_n \in \bar{\mathscr{H}}$ so that $\int |X_n| dm = \int |\tilde{X}_n| d\mu$. Assume eventually that m is a probability. Then so is $\bar{m}$ so that $h_n(X_1, X_2, \ldots) \in \bar{\mathscr{H}}$ and we can invert (38) to obtain

$$\mu^*(|\tilde{X}_n| > a) \geq \int h_n(\tilde{X}_1, \tilde{X}_2, \ldots) d\mu = \int h_n(X_1, X_2, \ldots) d\bar{m} \geq m^*(|X_n| > b)$$

as well as

$$\int |\tilde{X}_n| d\mu = \lim_k \int \big[|\tilde{X}_n| \wedge k\big] d\mu = \lim_k \int \big[|X_n| \wedge k\big] d\bar{m} = \int |X_n| d\bar{m}$$

since the sequence $\langle Y_k \rangle_{k \in \mathbb{N}}$, with $Y_k = |X_n| \wedge k$, is Cauchy in $L^1(m)$ – because $\tilde{X}_n \in L^1(\mu)$ by assumption – and m-converges to $|X_n|$ since m-tight. □

Theorem 4.3 may be used to understand when does pointwise convergence imply convergence in measure under finite additivity, that is when Egoroff theorem is not available and it may not be possible to exploit uniform convergence. We establish that a condition weaker than pointwise uniform convergence may be assumed.

Corollary 4.4. *Let m be a probability and $\langle X_n \rangle_{n \in \mathbb{N}}$ a sequence in $\mathfrak{F}(\Omega)$ converging pointwise to X_0. Assume that for every sequence $\langle \omega_k \rangle_{k \in \mathbb{N}}$ in Ω for which $\lim_k X_n(\omega_k)$ exists for all $n \in \mathbb{N}$ one has*

$$\lim_n \lim_k X_n(\omega_k) = \lim_k X_0(\omega_k), \tag{39}$$

the existence of the intervening limits being part of the claim. Then, X_n m-converges to X_0.

Proof. Since pointwise convergence and m-convergence are not altered when replacing m with some of its extensions to Ω, by Theorem 4.3 the sequence $Y = \langle Y_n \rangle_{n \in \mathbb{N}}$, with $Y_n = |X_n - X_0| \wedge 1$ is such that $(Y, \bar{m})$ admits a countably additive companion $(\tilde{Y}, \mu)$ relatively to the family $\{h \in \mathscr{C}(\mathbb{R}^{\mathbb{N}}) : h(Y_1, Y_2, \ldots) \in L^1(m)\}$. Then,

$$\lim_n \tilde{Y}_n(\tilde{\omega}) = \lim_n \lim_k Y_n(\omega_k) = \lim_k \lim_n Y_n(\omega_k) = 0$$

so that $\lim_n \int Y_n dm = \lim_n \int \tilde{Y}_n d\mu = 0$ and thus that X_n m-converges to X_0. $\square$

In Theorem 4.2 the set function μ is completely unrestricted. A possible mitigation is to require that μ vanishes on some suitable, given collection $\mathscr{N}$ of subsets of Ω.

Theorem 4.5. *Let $\mathscr{H}$, X and m be as in Theorem 4.2, and $\mathscr{N}$ an ideal of subsets of Ω'. The condition*

$$\int h(X)dm < 0 \quad \textit{implies} \quad \sup_{N \in \mathscr{N}} \inf_{\omega' \in N^c} h\big(X'(\omega')\big) < 0, \quad h \in \mathscr{H} \tag{40}$$

is equivalent to the existence of a measure structure $(\mathscr{R}, \mu)$ on Ω' which satisfies $\mathscr{N} \subset \mathscr{R}$,

$$\mu[\mathscr{N}] = \{0\},\ h(X') \in L^1(\mu) \ \textit{and} \int h(X)dm = \int h(X')d\mu, \quad h \in \mathscr{H}. \tag{41}$$

Moreover, if m is a countably additive set function on a σ ring, $\mathscr{N}$ is closed with respect to countable unions and $m_*\big(X \notin X'[N^c]\big) = 0$ for all $N \in \mathscr{N}$, then μ is countably additive.

Proof. Since $\mathscr{N}$ is an ideal, the binary relation $\succeq$ on $\mathfrak{F}(\Omega')$ defined by letting

$$f \succeq g \quad \text{if and only if} \quad \sup_{N \in \mathscr{N}} \inf_{\omega' \in N^c} (f - g)(\omega') \geq 0, \quad f, g \in \mathfrak{F}(\Omega') \tag{42}$$

is a partial order and $f \geq g$ implies $f \succeq g$. Moreover, $f_i \succeq g_i$ for $i = 1, 2$ implies $f_1 \vee f_2 \succeq g_1 \vee g_2$. In fact, $f_1 \vee f_2 \succeq f_i \succeq g_i$ i.e. $f_1 \vee f_2 \geq g_i - \varepsilon$ outside of some $N_i \in \mathscr{N}$. Thus, $f_1 \vee f_2 \geq g_1 \vee g_2 - \varepsilon$ outside of $N_1 \cup N_2 \in \mathscr{N}$ which, by (42), is equivalent to $f_1 \vee f_2 \succeq g_1 \vee g_2$. It is easy to see that, relatively to pointwise ordering, the set

$$\mathscr{F} = \big\{f \in \mathfrak{F}(\Omega') : f \sim h(X') \text{ for some } h \in \mathscr{H}\big\} \tag{43}$$

is a Stonean vector sublattice of $\mathfrak{F}(\Omega')$. Writing

$$\phi(f) = \int h(X)dm, \quad f \sim h(X'),\ h \in \mathscr{H} \tag{44}$$

implicitly defines, via (40), a positive linear functional on $\mathscr{F}$ so that, by Corollary 3.4, we conclude that there exists a minimal measurable structure $(\mathscr{R}, \mu)$ satisfying

$$f \in L^1(\mu) \quad \text{and} \quad \phi(f) = \int f d\mu, \quad f \in \mathscr{F}. \tag{45}$$

If $N \in \mathscr{N}$ then $\mathbb{1}_N \sim 0$, $\mathbb{1}_N \in \mathscr{F}$ and $\mu(N) = 0$. Then (41) follows while the converse implication, is obvious.

The last claim will be proved, once again, by showing that under the stated conditions the functional ϕ defined in (44) is a Daniell integral over $\mathscr{F}$. In fact, let $\langle f_n \rangle_{n\in\mathbb{N}}$ be sequence in $\mathscr{F}$ decreasing pointwise to 0. Let $h_n \in \mathscr{H}$ be such that $f_n \sim h_n(X')$, $n = 1, 2, \ldots$. Define $g_n = \bigwedge_{1\le j\le n} h_j$ and $g = \lim_n g_n$. As shown above, $f_n \sim g_n(X') \succeq g(X')$ so that, by the countable union property, $\{g(X') > \varepsilon\} \subset \bigcup_n \{g(X') \ge f_n + \varepsilon\} \in \mathscr{N}$ and $\{g > \varepsilon\} \subset X'[\{g(X') \le \varepsilon\}]^c$. By assumption then, $m(g(X) > \varepsilon) = 0$ and so $\lim_n \int f_n d\mu = \lim_n \int g_n(X')d\mu = \lim_n \int g_n(X)dm = \int g(X)dm = 0$. □

Example 4.6. Let $(\Omega', \mathscr{A}, P)$ be a classical probability space, $S = \mathbb{R}$ and let X' be a normally distributed random quantity on Ω'. Fix $m \in fa(\mathscr{B}(\mathbb{R}))_+$ arbitrarily and let $\mathscr{H} = \mathscr{C}(\mathbb{R}) \cap L^1(m)$. Given that $P(X' \in B) > 0$ for every B open, we conclude that m is X'-conglomerative relatively to $\mathscr{H}$. In other words a normally distributed random quantity can assume any arbitrary distribution (relatively to the continuous functions) upon an accurate choice of the reference measure.

In addition, let $\mathscr{N}$ consist of all P null sets and observe that $\operatorname{cl} X'[N^c] = \mathbb{R}$ for each $N \in \mathscr{N}$ – so that $X'[N]$ has empty interior. Therefore, for all $h \in \mathscr{C}(\mathbb{R})$

$$\sup_{N\in\mathscr{N}} \inf_{\omega\in N^c} h(X'(\omega)) = \sup_{N\in\mathscr{N}} \inf_{s\in X'[N^c]} h(s) = \sup_{N\in\mathscr{N}} \inf_{s\in \operatorname{cl} X'[N^c]} h(s) = \inf_{s\in\mathbb{R}} h(s).$$

Property (40) then holds for every $m \in fa(\mathscr{B}(\mathbb{R}))_+$ with $\mathscr{H} = \mathscr{C}(\mathbb{R})$. One may then find μ vanishing on $\mathscr{N}$ and such that (X', μ) and m are companions. Assume moreover that m is countably additive and does not charge closed sets with empty interior. Then $m_*(X'[N^c]^c) = 0$ for every $N \in \mathscr{N}$. In this further special case, we can take $\mu \ll P$. Of course the same conclusion holds upon replacing X' with any variable possessing a strictly positive density over the whole of $\mathbb{R}$. Observe also that in the preceding special case with m countably additive, then exploiting the fact that the indicator of each open subset B of $\mathbb{R}$ is the pointwise limit of an increasing, $\langle f_n \rangle_{n\in\mathbb{N}}$, and a decreasing, $\langle g_n \rangle_{n\in\mathbb{N}}$, sequence of continuous functions, we conclude

$$\mu(X' \in B) \le \lim_n \int g_n(X')d\mu = \lim_n \int g_n dm = m(B)$$

and

$$\mu(X' \in B) \ge \lim_n \int f_n(X')d\mu = \lim_n \int f_n dm = m(B).$$

5. Applications to statistics and probability

We return now to the Bayesian problem described in the introduction and prove the following:

Theorem 5.1. *Let* $(\mathscr{A}, m)$ *be a measure structure on* Ω *and* $X \in \mathfrak{F}(\Omega, S)$*. The following properties are equivalent: (i) there exist* $\{Q_\theta : \theta \in \Theta\} \subset fa(\mathscr{A})_+$ *and* $G \in \mathfrak{F}(\Theta, S)$ *injective satisfying:*

$$\sup_{\theta\in\Theta} Q_\theta(A) < \infty, \quad A \in \mathscr{A}, \tag{46a}$$

$$\int hdm < 0 \quad \textit{implies} \quad \inf_{\theta\in\Theta} \int hdQ_\theta < 0, \quad h \in \mathscr{S}(\mathscr{A}), \tag{46b}$$

$$Q^*_\theta\big(A \cap \{X \neq G(\theta)\}\big) = 0 \quad A \in \mathscr{A},\ \theta \in \Theta; \tag{46c}$$

(ii) there exist $K \in \mathfrak{F}(\mathscr{A} \times S)$ *and a measurable structure* $(\mathscr{R}, \mu)$ *on* Ω *which satisfy the following properties valid for all* $A \in \mathscr{A}$ *and all* $s \in S$ *and* $E \subset S$*:*

$$K_s \in fa(\mathscr{A})_+ \quad \textit{with} \quad \sup_{s\in S} K_s(A) < \infty, \tag{47a}$$

$$K_A \circ X \in L^1(\mu) \quad \textit{and} \quad m(A) = \int K(A, X)d\mu, \tag{47b}$$

$$A \cap \{X \in E\} \in \mathscr{A}(K_s) \quad \textit{and} \quad K\big(A \cap \{X \in E\}; s\big) = K\big(A; s\big)\mathbb{1}_E(s). \tag{47c}$$

Proof. $(i)\Rightarrow(ii)$. Given that G is injective we may define $K \in \mathfrak{F}(\mathscr{A} \times S)$ by letting

$$K(A; s) = \sum_{\theta\in\Theta} Q_\theta(A)\mathbb{1}_{G(\theta)}(s), \quad A \in \mathscr{A},\ s \in S. \tag{48}$$

Clearly (47a) is satisfied. By (46c), $\inf_\omega K(h; X(\omega)) \leq \inf_\theta Q_\theta(h)$ for every $h \in \mathscr{S}(\mathscr{A})$ so that, letting $(Th)(\omega) = K(h; X(\omega))$ in Theorem 3.3, we conclude that T is directed and m is T-conglomerative. There exists then a measure structure $(\mathscr{R}, \mu)$ on Ω such that

$$K_h \circ X \in L^1(\mu) \quad \text{and} \quad \int hdm = \int K(h, X)d\mu, \quad h \in \mathscr{S}(A).$$

If $A \in \mathscr{A}$ and $E \subset S$, then either $Q^*_\theta(A \cap \{X \in E\}) = 0$ (if $G(\theta) \notin E$) or $Q^*_\theta(A \cap \{X \in E^c\}) = 0$. In either case $A \cap \{X \in E\} \in \mathscr{A}(K_s) \cap \mathscr{A}(Q_\theta)$ and

$$\begin{aligned} K\big(A \cap \{X \in E\}; s\big) &= \sum_{\theta\in\Theta} Q_\theta\big(A \cap \{X \in E\}\big)\mathbb{1}_{G(\theta)}(s) = \sum_{\theta\in G^{-1}(E)} Q_\theta(A)\mathbb{1}_{G(\theta)}(s) \\ &= \mathbb{1}_E(s) \sum_{\theta\in\Theta} Q_\theta(A)\mathbb{1}_{G(\theta)}(s) = K(A; s)\mathbb{1}_E(s). \end{aligned}$$

$(ii)\Rightarrow(i)$. Take $\Theta = S$, $Q_\theta = K_s$ and G the identity. Then, (46a) follows from (47a) and (46c) from (47c). To deduce (46b) from (47b) it is enough to remark, via Theorem 4.2, that the identity on S is trivially a companion to X (relatively to the whole of $L^1(\mu)$). □

The function $K(A;s)$ in Theorem 5.1, a kernel, plays a prominent role in statistical problems in which it is interpreted as the prevision of A conditional on the occurrence of $X = s$. Its existence is generally deduced from that of regular conditional expectation and requires some classical properties, such as countable additivity of μ, measurability of X and some regularity of the space S, such as being a Blackwell space. In Theorem 5.1 instead the existence of K follows, via (46c), from that of a random quantity X that strictly separates priors, so that each $\theta \in \Theta$ may be interpreted as a corresponding hypothesis concerning X.

The following is an example of (46c) in the classical setting.

Example 5.2. Let $X_1, X_2, \dots$ be m-measurable random quantities on Ω. Define implicitly the map

$$F(\omega, t) = \lim_k \liminf_n \frac{1}{n} \sum_{j=1}^{n} \mathbb{1}_{\{X_i \le t+2^{-k}\}}(\omega) \tag{49}$$

of Ω into the set $\mathscr{X}$ of increasing, right continuous, $[0,1]$-valued functions on $\mathbb{R}$, the limiting empirical distribution. For each $\theta \in \Theta$, let $G(\theta)$ be a candidate distribution. In the classical case, with each Q_θ countably additive and the sequence $X_1, X_2, \dots$ independently and identically distributed under each Q_θ, condition (46c), with $X = F$, is a simple consequence of the strong law of large numbers, examined by Doob [13]. In order to guarantee that the inverse of G is Borel measurable Doob assumes that G is Borel measurable and that Θ is a subset of a complete and separable metric space, see also [12]. □

We pass now to the classical problem of Skhorohod which has been studied by a number of authors too large to give exact references. We have been influenced by the work of Berti, Pratelli and Rigo [3]. The starting point is the construction of a universal representation for the case of a separable space.

Corollary 5.3. *Let $U \in \mathfrak{F}(\Omega)$ with $\operatorname{cl} U[\Omega]$ having non empty interior and let S be a separable, topological space. There exists a Borel function $H \in \mathfrak{F}(\mathbb{R}, S)$ with countable range and such that $X' = H(U)$ is companion to any pair (X, m) relatively to $\mathscr{H} = \{h \in \mathscr{C}(S) : h(X) \in L^1(\mu)\}$.*

Proof. By the remarks following Theorem 4.2 we can assume with no loss of generality that X is the identity. Given that $[a,b] \subset \operatorname{cl} U[\Omega]$ for some $a, b \in \mathbb{R}$ then, upon replacing U with a suitable continuous transformation, we can assume that $\operatorname{cl} U[\Omega] = [0,1]$. Let S_0 be a countable, dense subset of S and $\iota \in \mathfrak{F}(\mathbb{N}, S_0)$ an enumeration of S_0. Define,

$$G(x) = \inf\big\{n \in \mathbb{N} : 1 - 2^{-n} \ge x\big\}, \quad x \in (0,1) \ \text{ and } \ H = \iota \circ G. \tag{50}$$

Since $G^{-1}(n) = (1 - 2^{-(n-1)}, 1 - 2^{-n}]$, H is a Borel function mapping $(0,1)$ onto S_0. If $h \in \mathscr{H}$ and $\int h dm < 0$ then $\{h < 0\}$ is an open, non empty subset of S

and as such it contains some element $\iota(n_h)$ of S_0. The set $B_h = \{U \in G^{-1}(n_h)\}$ is non empty (as $\mathrm{cl}\, U[\Omega] = [0,1]$) and coincides with $\{X' = \iota(n_h)\}$. Thus, $B_h \subset \{h(X') < 0\}$ so that m is X'-conglomerative relatively to $\mathscr{H}$. $\square$

Corollary 5.3 extends to the case of finite additivity and of a separable state space the classical idea of generating a random quantity with given distribution by applying to a uniformly distributed random quantity the inverse of the corresponding cumulative density function. Interestingly, we obtain that the *same* function X represents *all* possible distributions relatively to the class of continuous functions and for some suitable set function μ. Let us also mention the possibility of dropping the condition that S is separable by assuming that m is supported by a measurable, separable subset of S.

We highlight the advantage of doing without measurability. Constructing a function such as U in Corollary 5.3 is a rather trivial exercise as long as Ω has the right cardinality. Requiring that U is uniformly distributed on the unit interval under some classical probability measure P, as in the following Theorem 5.4, requires, in contrast, additional assumptions. The following result is inspired by [3, Theorem 3.1].

Theorem 5.4. *Let S be a normal, separable topological space, Σ a ring of subsets of S and $(\Omega, \mathscr{A}, P)$ a classical probability space supporting a random quantity U uniformly distributed on $(0,1)$. Let either $m \in fa(\Sigma)_+$ be countably additive or S be compact and write $\mathscr{H} = \mathscr{C}(S) \cap L^1(m)$. There exists a Borel function $g \in \mathfrak{F}((0,1), S)$ such that $X = g(U)$ is supported by $(\Omega, \mathscr{A}, P)$ and*

$$\int hdm = \int h(X)dP, \quad h \in \mathscr{H}. \tag{51}$$

Proof. If S is compact then the restriction of m to the minimal ring $\mathscr{R}_{\mathscr{H}}$ is countably additive. Let H be the map defined in (50). Then, as was shown in the proof of Corollary 5.3, m is H-conglomerative relatively to $\mathscr{C}(S)$ so that, by Theorem 4.2,

$$\int hdm = \int h(H)d\mu, \quad h \in \mathscr{H} \tag{52}$$

for some measure structure $(\mathscr{R}, \mu)$ on $(0,1)$. We claim that $\sigma\mathscr{R} = \mathscr{B}((0,1))$. Recall that $\sigma\mathscr{R}$ is generated by sets of the form $\{h(H) > t\}$ which are Borel since h is continuous and H is Borel. Conversely, if $0 \le a \le b \le 1$ then the set $H[(a,b)]$ is a finite subset of S – and therefore closed. Since S is normal, for any other finite subset F of $H[(a,b)^c]$ we can find a function $f \in \mathfrak{F}(S,[0,1])$ such that $f = 1$ on $H[(a,b)]$ and $f = 0$ on F. Thus $(a,b) \subset \{f(H) \ge 1\} \in \sigma\mathscr{R}$. Since $H[(0,1)]$ is countable we find a sequence $\langle f_n \rangle_{n\in\mathbb{N}}$ of such functions each vanishing on a finite subset of $H[(a,b)^c]$ so that the intersection $\bigcap_n \{f_n(H) \ge 1\}$ is again an element of $\sigma\mathscr{R}$ and coincides with (a,b). In other words, we can assume that μ is defined on the Borel subsets of $(0,1)$. From the classical Skhorohod theorem, we deduce the existence of an S valued random quantity

Z supported by $((0,1), \mathscr{B}((0,1)), \Lambda)$ (with Λ the Lebesgue measure on $(0,1)$) and admitting μ as its distribution. On its turn, Λ is the distribution of U under P. A repeated application of Theorem 4.2 with $g = H \circ Z$ and $X = g(U)$ gives

$$\int hdm = \int h(H)d\mu = \int h(g)d\Lambda = \int h(X)dP, \quad h \in \mathscr{H}.$$

Thus the random quantity X is supported by $(\Omega, \mathscr{A}, P)$ and represents m relatively to $\mathscr{C}(S)$. □

We generalize the preceding Example 4.6 to Brownian motion. Let for the rest of this section $\mathcal{I}$ be the family of finite subsets of $\mathbb{R}_+$ and, for each $\alpha = \{t_1, \ldots, t_n\} \in \mathcal{I}$, let π_α be the projection

$$\pi_\alpha(s) = (s_{t_1}, s_{t_2}, \ldots, s_{t_n}) \qquad s \in \mathfrak{F}(\mathbb{R}_+). \tag{53}$$

If $X = (X_t : t \in \mathbb{R}_+)$, we write $X_\alpha = (X_t : t \in \alpha)$.

Corollary 5.5. *Let $X' = (X'_t : t \in \mathbb{R}_+)$ be Brownian motion on some classical probability space $(\Omega, \mathscr{F}, P)$ and $(m_\alpha : \alpha \in \mathcal{I})$ a projective family of probabilities (namely $m_\alpha \in fa(\mathscr{B}(\mathbb{R}^\alpha))_+$ is the marginal of m_β whenever $\alpha \subset \beta$). There exists a probability structure $(\mathscr{A}, \mu)$ on Ω such that $h(X'_\alpha) \in L^1(\mu)$ and*

$$\int hdm_\alpha = \int h(X'_\alpha)d\mu, \quad \alpha \in \mathcal{I}, \; h \in \mathscr{C}(\mathbb{R}^\alpha) \cap L^1(m_\alpha). \tag{54}$$

If m_α is countably additive, then

$$m_\alpha(B) = \mu(X'_\alpha \in B), \quad B \in \mathscr{B}(\mathbb{R}^\alpha). \tag{55}$$

Proof. As usual, a projective family of probabilities induces a unique probability on the algebra $\Sigma = \left\{\pi_\alpha^{-1}(B) : \alpha \in \mathcal{I}, B \in \mathscr{B}(\mathbb{R}^\alpha)\right\}$ of finite dimensional cylinders obtained by letting

$$m\big(\pi_\alpha^{-1} A\big) = m_\alpha(A), \quad A \in \mathscr{B}(\mathbb{R}^\alpha), \; \alpha \in \mathcal{I}. \tag{56}$$

If $g \in \mathfrak{F}(\mathbb{R}^\alpha)$ and $h = g \circ \pi_\alpha$ then $\{h > t\} = \pi_\alpha^{-1}(\{g > t\})$ so that from Lemma 2.2 we conclude

$$\int hdm = \int gdm_\alpha$$

whenever either side is well defined. Let $\mathscr{H} = \{g \circ \pi_\alpha : g \in \mathscr{C}(\mathbb{R}^\alpha), \; \alpha \in \mathcal{I}\} \cap L^1(m)$. If $h \in \mathscr{H}$ and $\int hdm < 0$, then $\int h_\alpha dm_\alpha < 0$ for some $\alpha = \{t_1 < \ldots < t_n\} \in \mathcal{I}$. Since $\{h_\alpha < 0\}$ is open and non empty, there exist open, non empty sets $B_1, \ldots, B_n \subset \mathbb{R}$ such that $x_i - x_{i-1} \in B_i$ for $i = 1, \ldots, n$ (and $x_0 = 0$) implies $h_\alpha(x_1, \ldots, x_n) < 0$. Therefore, $P(X'_{t_1}, \ldots, X'_{t_n} \in \{h_\alpha < 0\}) \geq \prod_{i=1}^n P(X'_{t_i} - X'_{t_{i-1}} \in B_i) > 0$ so that $\inf_\omega h_\alpha(X'_\alpha) < 0$ and m is X'-conglomerative. The second claim, as in Example 4.6, follows from metric spaces being normal. □

Corollary 5.5 is related to [16, Theorem 1] and, in Dubins' peculiar terminology, it asserts that Brownian motion is *cousin* to any stochastic process. Dubins main finding is a necessary and sufficient condition for the existence of cousins with almost all paths in a given class. His claim is an easy corollary of our previous results.

Corollary 5.6. (Dubins) *Let X be a stochastic process on a probability space $(\Omega, \mathscr{A}, m)$. Let $\mathbb{Y} \subset \mathfrak{F}(\mathbb{R}_+)$ satisfy:*

$$\forall(\omega,\alpha) \in \Omega \times \mathcal{I}, \quad \exists Y \in \mathbb{Y} \quad \textit{such that} \quad Y(t) = X(\omega,t), \qquad t \in \alpha. \tag{57}$$

There is a process X' on a probability space (Ω', Σ, μ) with μ-a.a. paths in $\mathbb{Y}$ such that for all $\alpha \in \mathcal{I}$, $g \in \mathfrak{F}(\mathbb{R}^\alpha)$, $g(X_\alpha) \in L^1(m)$

$$g(X'_\alpha) \in L^1(\mu) \quad \textit{and} \quad \int g(X_\alpha)dm = \int g(X'_\alpha)d\mu. \tag{58}$$

Proof. Write $\mathscr{H} = \big\{g \circ \pi_\alpha : \alpha \in \mathcal{I},\ g \in \mathfrak{F}(\mathbb{R}^\alpha),\ g(X_\alpha) \in L^1(m)\big\}$, $\Omega' = \mathfrak{F}(\mathbb{R}_+)$, and define $T \in \mathfrak{L}\big(\mathscr{H}, \mathfrak{F}(\mathbb{Y})\big)$ by letting $T(h)(Y) = h(Y)$ for each $h \in \mathscr{H}$ and $Y \in \mathbb{Y}$. Then, T is directed and, by (57), the linear functional $\phi(h) = \int h(X)dm$ is T-conglomerative. By Theorem 3.3 there exists a minimal measure structure $(\mathscr{R}_0, \mu_0)$ on $\mathbb{Y}$ such that

$$\int h(X)dm = \int hd\mu_0, \quad h \in \mathscr{H}.$$

Since m is a probability, then $\mathscr{R}_0$ is an algebra and μ_0 a probability. Let

$$\Sigma = \big\{A \subset \Omega' : A \cap \mathbb{Y} \in \mathscr{R}_0\big\} \quad \text{and} \quad \mu(A) = \mu_0(A \cap \mathbb{Y}), \quad A \in \Sigma.$$

Then, Σ is an algebra of subsets of Ω', μ a probability on Σ with $\mu(\mathbb{Y}^c) = 0$ and $X'(w,t) = w(t)$ a stochastic process on (Ω', Σ, Q) with $X'_w = w$. Moreover, $\int hd\mu_0 = \int h(X')d\mu$ for all $h \in \mathscr{H}$. □

Dubins deduces from this result that any stochastic process admits *cousins* having continuous or polynomial or stepwise linear paths.

6. Convex functions

Eventually, we turn attention to convex functions. For $f \in \mathfrak{F}(\mathbb{R})$ we denote by D^+f and D^-f the right and left derivatives and by $f(x+)$ and $f(x-)$ the right and left limits at x, provided such quantities exist. We also set conventionally $D^+f(\infty) = D^-f(\infty) = \lim_{x\to\infty} D^+f(x)$ and $D^+f(-\infty) = D^-f(-\infty) = \lim_{x\to-\infty} D^+f(x)$.

Theorem 6.1. *Let $\varphi \in \mathfrak{F}(\mathbb{R})$ and $x_0 \in \operatorname{argmin}_{x\in\overline{\mathbb{R}}} \varphi(x)$. For each $u, v \in \mathbb{R}$ define*

$$h_u^v(x) = (v - x \vee u)^+ \mathbb{1}_{\{x > x_0\}} - (v \wedge x - u)^+ \mathbb{1}_{\{x \le x_0\}}, \quad x \in \mathbb{R}. \tag{59}$$

Let $\mathscr{N}$ be an ideal of subsets of Ω and $X \in \mathfrak{F}(\Omega)$. The following properties are mutually equivalent:

(i) *φ is convex and $\{u < X < v\} \in \mathscr{N}$ implies $D^-\varphi(v) \le D^+\varphi(u)$;*

(ii) *there exist $y_0^+, y_0^- \in \mathbb{R}$ and a measure structure $(\mathscr{R}, \lambda)$ on Ω such that*

(a) $\mathscr{N} \subset \mathscr{R}$ and $\lambda[\mathscr{N}] = \{0\}$, (b) $\lim_n \lambda^*(|X - x_0| < 2^{-n}) = 0$,

(c) $\{h_u^v(X) : v, u \in \mathbb{R}\} \subset L^1(\lambda)$ and

$$\varphi(v) - \varphi(u) = y_0^+ h_u^v(x_0+) + y_0^- h_u^v(x_0) + \int h_u^v(X) d\lambda, \quad v \ge u; \tag{60}$$

(iii) there exist $y_0^+, y_0^- \in \mathbb{R}$ and $\nu \in fa(\mathscr{B}(\mathbb{R}))_+$ countably additive such that (a) $\nu(A) = 0$ for A open and $X^{-1}(A) \in \mathscr{N}$, (b) $\nu^*(\{x_0\}) = 0$, (c) $\{h_u^v : v, u \in \mathbb{R}\} \subset L^1(\nu)$ and

$$\varphi(v) - \varphi(u) = y_0^+ h_u^v(x_0+) + y_0^- h_u^v(x_0) + \int h_u^v d\nu, \quad v \ge u. \tag{61}$$

Proof. $(i)\Rightarrow(ii)$. By convexity, $D^+\varphi(x_0), D^-\varphi(x_0) \in \mathbb{R}$. Write $\mathcal{D} = \big\{t : D^-\varphi(t) = D^+\varphi(t)\big\} \cup \{x_0\}$ and define $A_u = \{u < X \le x_0\}$, $A^v = \{x_0 < X \le v\}$ and

$$\mathscr{R}_0 = \Big\{\big(A_u \cap N_u^c\big) \cup \big(A^v \cap N_v^c\big) \cup N : u, v \in \mathcal{D},\ N_u, N_v, N \in \mathscr{N}\Big\}. \tag{62}$$

It is clear that $\mathscr{R}_0$ contains $\mathscr{N}$ (upon taking $u = v = x_0$) as well as $\{A_u, A^v : u, v \in \mathcal{D}\}$. Moreover, it is routine to verify that $\mathscr{R}_0$ is closed with respect to union and intersection with

$$H_1 \cup H_2 = \big(A_{u_1 \wedge u_2} \cap N_u^c\big) \cup \big(A^{v_1 \vee v_2} \cap N_v^c\big) \cup N \tag{63a}$$

$$H_1 \cap H_2 = \big(A_{u_1 \vee u_2} \cap \hat{N}_u^c\big) \cup \big(A^{v_1 \wedge v_2} \cap \hat{N}_v^c\big) \cup \hat{N} \tag{63b}$$

whenever $H_i = \big(A_{u_i} \cap N_{u_i}^c\big) \cup \big(A^{v_i} \cap N_{v_i}^c\big) \cup N_i \in \mathscr{R}_0$ for $i = 1, 2$. Write $F(x) = \big[D^+\varphi(x \vee x_0) - D^+\varphi(x_0)\big] + \big[D^-\varphi(x \wedge x_0) - D^-\varphi(x_0)\big]$ and

$$\lambda_0(H) = F(v \vee x_0) - F(u \wedge x_0) \text{ when } H = (A_u \cap N_u^c) \cup (A^v \cap N_v^c) \cup N \in \mathscr{R}_0. \tag{64}$$

To see that λ_0 is well defined observe that if $u_1 \wedge x_0 < u_2 \wedge x_0$ and

$$\big(A_{u_1} \cap N_{u_1}^c\big) \cup \big(A^{v_1} \cap N_{v_1}^c\big) \cup N_1 = \big(A_{u_2} \cap N_{u_2}^c\big) \cup \big(A^{v_2} \cap N_{v_2}^c\big) \cup N_2 \in \mathscr{R}_0$$

then $\{u_1 \wedge x_0 < X \le u_2 \wedge x_0\} \in \mathscr{N}$. Thus by (i) and the fact that $u_1, u_2 \in \mathcal{D}$ and that $u_1 < x_0$,

$$D^-\varphi(u_1 \wedge x_0) = D^-\varphi(u_2 \wedge x_0) \quad \text{i.e.} \quad F(u_1 \wedge x_0) = F(u_2 \wedge x_0)$$

and likewise $F(v_1 \vee x_0) = F(v_2 \vee x_0)$. In other words $\lambda_0 \in fa(\mathscr{R}_0)_+$ with $\lambda[\mathscr{N}] = \{0\}$. Moreover, if $H_1, H_2 \in \mathscr{R}_0$ then by (63)

$$\begin{aligned}\lambda_0(H_1) + \lambda_0(H_2) &= F(v_1 \vee x_0) + F(v_2 \vee x_0) - F(u_1 \wedge x_0) - F(u_2 \wedge x_0)\\ &= F(v_1 \vee v_2 \vee x_0) + F((v_1 \wedge v_2) \vee x_0) - F((u_1 \vee u_2) \wedge x_0) - F(u_1 \wedge u_2 \wedge x_0)\\ &= \lambda_0(H_1 \cup H_2) + \lambda_0(H_1 \cap H_2)\end{aligned}$$

i.e. λ_0 is strongly additive on $\mathscr{R}_0$. It follows from [4, 3.1.6 and 3.2.4] that λ_0 admits a unique extension $\lambda_1 \in fa(\mathscr{R}_1)_+$ to the generated ring $\mathscr{R}_1$. Let I be an interval with endpoints in $\mathbb{R} \cup \{x_0\}$. Given that $\mathcal{D}$ is dense in $\mathbb{R} \cup \{x_0\}$, $\lambda^*(X \in I) < \infty$. By [4, 3.4.1 and 3.4.4] we obtain a further extension $\lambda \in fa(\mathscr{R})_+$ to the ring $\mathscr{R} = \{A \subset \Omega : \lambda_1^*(A) < \infty\}$. Then $\{X \in I\} \in \mathscr{R}$ and $X\mathbb{1}_I(X)$ is λ-measurable whenever I is as above, by Lemma 2.1. Therefore, setting $y_0^+ = D^+\varphi(x_0)$ and $y_0^- = D^-\varphi(x_0)$,

$$\begin{aligned}\int_{u\vee x_0}^{v\vee x_0} D^+\varphi(t)dt &= y_0^+ h_u^v(x_0+) + \int_{u\vee x_0}^{v\vee x_0} \mathbb{1}_{\mathcal{D}}[D^+\varphi(t) - y_0^+]dt\\ &= y_0^+ h_u^v(x_0+) + \int_u^v \mathbb{1}_{\mathcal{D}}\lambda_1(x_0 < X \le t)dt\\ &= y_0^+ h_u^v(x_0+) + \int_u^v \lambda(x_0 < X \le t)dt\\ &= y_0^+ h_u^v(x_0+) + \int_{x_0}^{\infty} (v - u \vee X)^+ d\lambda \qquad \text{(by Lemma 2.2)}\end{aligned}$$

and similarly $\int_{u\wedge x_0}^{v\wedge x_0} D^+\varphi(t)dt = y_0^- h_u^v(x_0) - \int_{-\infty}^{x_0} (v \wedge X - u)^+ d\lambda$. We conclude

$$\begin{aligned}\varphi(v) - \varphi(u) &= \int_{u\vee x_0}^{v\vee x_0} D^+\varphi(t)dt + \int_{u\wedge x_0}^{v\wedge x_0} D^-\varphi(t)dt\\ &= y_0^+ h_u^v(x_0+) + y_0^- h_u^v(x_0) + \int h_u^v(X)d\lambda\,.\end{aligned}$$

Fix an increasing $\langle u_n\rangle_{n\in\mathbb{N}}$ and a decreasing $\langle v_n\rangle_{n\in\mathbb{N}}$ sequence in $\mathcal{D}$ converging to x_0, with $u_n < u_{n+1} < x_0$ if $x_0 > -\infty$ and $v_n > v_{n+1} > x_0$ if $x_0 < \infty$. Then,

$$\lim_n \lambda^*(u_n < X < v_n) \le \lim_n D^+\varphi(v_n) - D^+\varphi(x_0) - D^-\varphi(u_n) + D^-\varphi(x_0) = 0$$

so that $\lim_n \lambda^*(|X - x_0| < 2^{-n}) = 0$.

(ii)$\Rightarrow$*(iii)*. With u_n and v_n defined as above, define the function

$$h_u^v(x;n) = \begin{cases} h_u^v(x), & \text{if } x \notin (u_n, v_n],\\ h_u^v(u_n)\frac{u_{n+1}-x}{u_{n+1}-u_n}, & \text{if } x \in (u_n, u_{n+1}],\\ h_u^v(v_n)\frac{x-v_{n+1}}{v_n-v_{n+1}}, & \text{if } x \in (v_{n+1}, v_n].\end{cases}$$

Then, $h_u^v(\cdot;n)$ is a continuous function vanishing outside of the interval $[u \wedge v_{n+1}, v \vee u_{n+1}]$. Moreover: (a) $\{|h_u^v(x;n) - h_u^v(x)| > c\} \subset (u_n, v_n]$ so that $h_u^v(X;n)$ is λ-convergent to $h_u^v(X)$, (b) $|h_u^v(x;n)| \le |h_u^v(x;n+1)| \le |h_u^v(x)|$, (c) $\lim_n h_u^v(x;n) = h_u^v(x)$ for all $x \neq x_0$ and (d) $h_u^v(X;n)$ is λ-measurable and therefore an element of $L^1(\lambda)$. Let (X', ν), with $\Omega' = \mathbb{R}$ and X' the identity, be the countably additive companion of (X, λ) relatively to the family $\{h(X) : h \in \mathscr{C}_K(\mathbb{R})\}$. It follows that

$$\int h_u^v(X) d\lambda = \lim_n \int h_u^v(X;n) d\lambda = \lim_n \int h_u^v(x;n) d\nu = \int h_u^v d\nu.$$

Observe that if $x_0 \in \mathbb{R}$ and $g_n \in \mathscr{C}_K(\mathbb{R})$ is such that $\mathbb{1}_{(u_n, v_n]} \ge g_n \ge \mathbb{1}_{(u_{n+1}, v_{n+1}]}$, then

$$\nu^*(\{x_0\}) \le \lim_n \int g_n(X) d\lambda \le \lim_n \lambda(u_n < X \le v_n) = 0.$$

Let $I \subset \mathbb{R}$ be an open interval with $X^{-1}(I) \in \mathscr{N}$ and $\langle g_n \rangle_{n \in \mathbb{N}}$ a sequence of non negative, continuous functions which increases to $\mathbb{1}_I$. It is then obvious that

$$0 = \lim_n \int g_n(X) d\lambda = \lim_n \int g_n d\nu = \nu(I).$$

The conclusion extends to open sets.

$(iii) \Rightarrow (i)$. If φ satisfies (61) it is clearly convex since the function $v \to h_u^v(x)$ is convex for every $u \le v$. Assume that $u < v$ and $\{u < X < v\} \in \mathscr{N}$. Then, $\nu((u,v)) = 0$ so that, for arbitrary $u < t < v$

$$\frac{\varphi(v) - \varphi(u)}{v - u} = \begin{cases} D^+\varphi(x_0) + \nu([x_0, t)), & \text{if } v > u \ge x_0 \\ D^-\varphi(x_0) + \nu([t, x_0)), & \text{if } x_0 \ge v > u \\ D^+\varphi(x_0), & \text{if } v > x_0 > u \end{cases} \tag{65}$$

and (i) follows. □

If, e.g., φ is differentiable at x_0, then (60) simplifies into:

$$\varphi(v) = \varphi(x_0) + \int_{\{v < X \le x_0\}} (X - v) d\lambda + \int_{\{x_0 < X \le v\}} (v - X) d\lambda. \tag{66}$$

The above result can be stated in a slightly different way:

Corollary 6.2. *Let $X \in \mathfrak{F}(\Omega)$ with $\operatorname{cl} X[\Omega] = \mathbb{R}$, $\varphi \in \mathfrak{F}(\mathbb{R})$ and define x_0, $D^+\varphi(x_0)$, $D^-\varphi(x_0)$ and h_u^v as in Theorem 6.1. φ is convex if and only if $D^+\varphi(x_0), D^-\varphi(x_0) \in \mathbb{R}$ and there exists a measure structure $(\mathscr{R}, \lambda)$ on Ω such* *that (a)* $\lambda(u < X < v) = 0$ *whenever* $D^+\varphi(v) \leq D^-\varphi(u)$*, (b)* $\{h_u^v(X) : v \geq u\} \subset L^1(\lambda)$ *and, for all* $v \geq u$

$$\varphi(v) - \varphi(u) = D^+\varphi(x_0)h_u^v(x_0+) + D^-\varphi(x_0)h_u^v(x_0-) + \int h_u^v(X)d\lambda. \tag{67}$$

Proof. Define $\mathscr{N} = \big\{\{u < X < v\} : u, v \in \mathbb{R}, D^+\varphi(v) \leq D^-\varphi(u)\big\}$. From $\operatorname{cl} X[\Omega] = \mathbb{R}$ follows that $\{u < X < v\} \in \mathscr{N}$ if and only if $D^+\varphi(v) \leq D^-\varphi(u)$ and that $\mathscr{N}$ is an ideal of sets. Then (67) follows from Theorem 6.1.(*iii*). □

References

[1] C. D. Aliprantis, O. Burkinshaw: Positive Operators, Springer Publ., Dordrecht (2006).

[2] P. Berti, L. Pratelli, P. Rigo: Asymptotic behaviour of the empirical process for exchangeable data, Stoch. Proc. Appl. 116 (2006) 337–344.

[3] P. Berti, L. Pratelli, P. Rigo: Skorohod representation on a given probability space, Probab. Theory Relat. Fields 137 (2007) 277–288.

[4] K. P. S. Bhaskara Rao, M. Bhaskara Rao: Theory of Charges, Academic Press, London (1993).

[5] R. D. Bourgin: Geometric Aspects of Convex Sets with the Radon-Nikodým Property, Lect. Notes Math. 993, Springer-Verlag, Berlin (1983).

[6] G. Cassese: Sure wins, separating probabilities and the representation of linear functionals, J. Math. Anal. Appl. 354 (2009) 558–563.

[7] G. Choquet (1956): Existence et unicité des représentations intégrales au moyen des points extrémaux dans les cônes convexes, Sém. Bourbaki 139 (1956) 33–47.

[8] G. Choquet: Determination and study of positive forms on spaces of functions: I, J. Approx. Th. 7 (1973) 325–333.

[9] G. Choquet: Determination and study of positive forms on spaces of functions: II, J. Approx. Th. 10 (1974) 358–378.

[10] G. Choquet: Représentation intégrale, in: Measure Theory and its Applications, Lect. Notes Math. 1033 (1983) 114–143.

[11] B. de Finetti: Sulla proprietà conglomerativa delle probabilità subordinate, Rend. R. Istit. Lomb. Sc. Lett. 63 (1930) 3–7.

[12] P. Diaconis, D. Freedman: On the consistency of Bayes estimates, Ann. Probab. 14 (1986) 1–26.

[13] J. L. Doob: Application of the theory of martingales, Actes Coll. Int. CNRS, Paris (1948) 23–27.

[14] L. E. Dubins, L. J. Savage: How to Gamble if You Must, McGraw-Hill, New York (1965).

[15] L. E. Dubins: Finitely additive conditional probability, conglomerability and disintegrations, Ann. Probab. 3 (1975) 89–99.

[16] L. E. Dubins: Paths of Finitely Additive Brownian Motion Need not be Bizarre, in: J. Azema, M. Eméry, M. Ledoux, M. Yor (eds.), Séminaire de Probabilités 32, Lect. Notes Math. 1709, Springer-Verlag, Berlin (1999).

[17] N. Dunford, J. T. Schwartz: Linear Operators. Part I, Wiley, New York (1988).

[18] G. A. Edgar: A noncompact Choquet theorem, Proc. Amer. Math. Soc. 49 (1975) 354–358.

[19] G. A. Edgar: Extremal integral representations, J. Func. Anal. 23 (1976) 145–161.

[20] B. M. Hill, D. Lane: Conglomerability and countable additivity, in: Bayesian Inference and Decision Techniques. Essays in Honor of Bruno de Finetti, P. K. Goel, A. Zellner (eds.), North-Holland, Amsterdam (1986).

[21] R. Karandikar (1982), A general principle for limit theorems in finitely additive probability, Trans. Amer. Math. Soc. 273 (1982) 541–550.

[22] R. Karandikar: A general principle for limit theorems in finitely additive probability: the dependent case, J. Multivariate Anal. 24 (1988) 189–206.

[23] J. Lukeš, J. Malý, I. Netuka, J. Spurný: Integral Representation Theory. Applications to Convexity, Banach Spaces and Potential Theory, de Gruyter, Berlin (2010).

[24] R. R. Phelps: Lectures on Choquet's Theorem, Lect. Notes Math. 1757, Springer-Verlag, Berlin-Heidelberg (2000).

[25] H. P. Rosenthal: On the Choquet Representation Theorem, in: Functional Analysis, E. W. Odell Jr., H. P. Rosenthal (eds.), Lect. Notes Math. 1332, Springer-Verlag, Berlin-Heidelberg (1988).

[26] M. J. Schervish, T. Seidenfeld, J. B. Kadane: The extent of non-conglomerability of finitely additive probabilities, Z. Wahrsch. verw. Geb. (1984) 205–226.

[27] A. Zame: Conglomerability and finite partitions, Proc. Amer. Math. Soc. 102 (1988) 165–168.