ON THE RADON-NIKODYM-PROPERTY AND MARTINGALE CONVERGENCE*

G. A. Edgar
Department of Mathematics
The Ohio State University
Columbus, Ohio 43210 U.S.A.

It is well-known that there is an intimate connection between the Radon-Nikodym property and martingale convergence in a Banach space. This connection can be "localized" to a closed bounded convex subset of a Banach space. In this paper we are interested primarily in this connection for a bounded convex set which is not closed.

If C is a bounded convex set in a locally convex space, C is said to have the martingale convergence property iff every martingale with values in C converges in measure. Since C is not assumed to be metrizable, it is appropriate to use martingales indexed by an arbitrary directed set, and not restrict attention to sequential martingales. Similarly, C is said to have the Radon-Nikodym property iff every vector-valued measure defined on a probability space with average range in C has a derivative which has sufficiently strong measurability properties. The one-dimensional example of an open interval shows that the two properties are no longer equivalent. Theorem 2.4 describes the connection between the two notions.

This paper is also concerned with an ordering on the tight probability measures on a bounded convex set C. The ordering $\prec$, which has been called "comparison of experiments", "the Choquet ordering", "the dilation ordering", and many others, can be described in many equivalent ways; they are given in Theorem 2.2. For example, $\mu \prec \nu$ means $\int f d\mu \leq \int f d\nu$ for all bounded continuous convex functions f on C. Other descriptions of the ordering involve dilations and conditional expectations. Earlier versions of this theorem have been attributed to: Hardy, Littlewood and Polya, Blackwell, Stein, Sherman, Cartier and Strassen.

One other result proved here deserves mention (Corollary 2.7). If C is a separable closed bounded convex subset of a Banach space, and if each point of C admits a unique representing measure on the extreme points of C, then C has the Radon-Nikodym property.

---

* Supported in part by National Science Foundation grant MCS77-04049.

The attentive reader will notice that the assumption of convexity is hardly ever used in an essential way, so that most of what appears below has a reformulation for nonconvex sets C. I have not included such reformulations here.

The paper has two sections. The first is preliminary; the results there are either substantially known or straightforward, so most proofs are omitted. The second section contains the main results of the paper, including those mentioned above; proofs are given in this section.

1.

If T is a completely regular Hausdorff space, we will write $\mathcal{C}_b(T)$ for the Banach space of all bounded continuous real-valued functions on T. We will be interested in several subsets of the dual $\mathcal{C}_b(T)^*$ in its weak* topology. First,

$$\mathcal{P}_f(T) = \{\mu \in \mathcal{C}_b(T)^*: \langle\mu,1\rangle = 1 , \mu \geq 0\} ;$$

this can be identified with the set of finitely-additive regular probability measures on the algebra generated by the zero sets [18, p. 165]; the identification (and similar ones below) will be made whenever convenient. Note that $\mathcal{P}_f(T)$ is compact. Next,

$$\mathcal{P}_\sigma(T) = \{\mu \in \mathcal{P}_f(T): \text{ if } f_n \in \mathcal{C}_b(T)\ (n=1,2,\dots),\ f_n \downarrow 0 \text{, then } \langle\mu,f_n\rangle \to 0\} ;$$

these measures extend uniquely to the Baire sets of T. Also,

$$\mathcal{P}_\tau(T) = \{\mu \in \mathcal{P}_f(T): \text{ if } f_\alpha \in \mathcal{C}_b(T) \text{ is a net, } f_\alpha \downarrow 0 \text{, then } \langle\mu,f_\alpha\rangle \to 0\} ;$$

these measures extend uniquely to (closed-) regular measures on the Borel sets of T. These are called τ-smooth measures. Next,

$$\mathcal{P}_t(T) = \{\mu \in \mathcal{P}_\tau(T): \text{ for each } \varepsilon > 0 \text{, there is a compact set } K \subseteq T \text{ such that } \mu(K) \geq 1-\varepsilon\} ;$$

these are called tight measures (on the Baire sets) or Radon measures (on the Borel sets). Finally,

$$\mathcal{P}_s(T) = \{\sum_{i=1}^{\infty} t_i \varepsilon_{x_i} : t_i \geq 0 ,\ \Sigma t_i = 1 ,\ x_i \in T\} ,$$

$$\mathcal{P}_d(T) = \{\sum_{i=1}^{n} t_i \varepsilon_{x_i} : n \in \mathbb{N} ,\ t_i \geq 0 ,\ \Sigma t_i = 1 ,\ x_i \in T\} .$$

Note $\mathcal{P}_f \supseteq \mathcal{P}_\sigma \supseteq \mathcal{P}_\tau \supseteq \mathcal{P}_t \supseteq \mathcal{P}_s \supseteq \mathcal{P}_d$, with $\mathcal{P}_f = \mathcal{P}_t$ if T is compact.

Let E be a locally convex (Hausdorff) topological vector space, and C a subset of E . If $\mu \in \mathcal{P}_f(C)$ and $x \in E$, we say that x is the <u>resultant</u> of μ , and write $x = r(\mu)$, iff for every $f \in E^*$, we have $\langle \mu, f \rangle = f(x)$. The set C will be called d-convex [resp. s-, t-, τ-, σ-, f-convex] iff for every $\mu \in \mathcal{P}_d(C)$ [resp. $\mathcal{P}_s(C)$, etc.], there exists $r(\mu) \in C$. Note that d-convex is the same as convex and that f-convex is the same as compact and convex. We will say that C satisfies <u>condition (EC)</u> iff the closed convex hull of a compact subset of C is a compact subset of C , i.e. if $K \subseteq C$ is compact, then there is a compact convex set K_1 with $K \subseteq K_1 \subseteq C$.

The following is from [8].

1.1 PROPOSITION. (a) <u>The set</u> C <u>satisfies condition (EC) if and only if, for every measure</u> $\mu \in \mathcal{P}_f(C)$ <u>with compact support, the resultant</u> $r(\mu)$ <u>exists in</u> C .

(b) C <u>is</u> t-<u>convex if and only if</u> C <u>is</u> s-<u>convex and satisfies condition (EC)</u>.

It is easy to show that $\mathcal{P}_x(T)$ is x-convex, where $x = d, s, \sigma, f$. It is not hard to show that $\mathcal{P}_\tau(T)$ is τ-convex. Indeed, suppose $\gamma \in \mathcal{P}_\tau(\mathcal{P}_\tau(T))$. Then $\mu = r(\gamma)$ exists in $\mathcal{P}_f(T)$. If f_α is a net in $\mathcal{C}_b(T)$ with $f_\alpha \downarrow 0$, then for all $\lambda \in \mathcal{P}_\tau(T)$, we have $\langle \lambda, f_\alpha \rangle \downarrow 0$. But for each α , the function $\lambda \mapsto \langle \lambda, f_\alpha \rangle$ is in $\mathcal{C}_b(\mathcal{P}_\tau(T))$, so $\langle \mu, f_\alpha \rangle = \int \langle \lambda, f_\alpha \rangle \, d\gamma(\lambda) \to 0$ since γ is τ-smooth. Thus μ is τ-smooth.

An example of D. H. Fremlin shows that $\mathcal{P}_t(T)$ need not satisfy condition (EC) and therefore need not be t-convex. Clearly, if $\mathcal{P}_\tau(T) = \mathcal{P}_t(T)$ (such spaces T are called, variously, "universally measurable" or "semi-Radonian" [10, Theorem 2, p. 133]), then $\mathcal{P}_t(T)$ is t-convex.

Let C be a subset of a locally convex space E , and let $(\Omega, \mathfrak{F}, P)$ be a probability space. If $\varphi\colon \Omega \to C$ is Borel measurable, we define a Borel measure $\varphi(P)$ on C by $\varphi(P)(B) = P(\varphi^{-1}(B))$. We will write $L^0(\Omega, \mathfrak{F}, P; C)$ for the set of all Borel measurable functions $\varphi\colon \Omega \to C$ such that $\varphi(P) \in \mathcal{P}_t(C)$, i.e. for every Borel set $B \subseteq C$, and every $\epsilon > 0$, there is a compact set $K \subseteq B$ with $P(\varphi^{-1}(B) \setminus \varphi^{-1}(K)) < \epsilon$. For $\varphi \in L^0(\Omega; C)$, we will write $x = \int_A \varphi \, dP$ iff $f(x) = \int_A f(\varphi(\omega)) dP(\omega)$ for all $f \in E^*$; if such an element x exists for each $A \in \mathfrak{F}$, we will say that φ is Pettis integrable. (Elements φ, ψ of L^0 should be identified iff they are <u>weakly equivalent</u>, i.e. $f \circ \varphi = f \circ \psi$ a.e. for all $f \in E^*$, the exceptional set may depend on f .)

Let E be a locally convex space, let $(\Omega,\mathfrak{F},P)$ be a probability space, and let $m\colon \mathfrak{F}\to E$ be a vector-valued measure. The P-average range of m is $\{m(A)/P(A)\colon A\in\mathfrak{F},\ P(A)>0\}$. We say m is differentiable with respect to P iff there exists $\varphi\in L^0(\Omega;E)$ such that $m(A)=\int_A \varphi\,dP$ for all $A\in\mathfrak{F}$; in that case we write $\varphi = dm/dP$. A bounded subset C of E is said to have the Radon-Nikodym property iff for any probability space $(\Omega,\mathfrak{F},P)$ and any measure $m\colon\mathfrak{F}\to E$ with $m\ll P$ and average range contained in C, there exists $dm/dP\in L^0(\Omega,\mathfrak{F},P;C)$.

Here is the Radon-Nikodym theorem which will be used below. The basic form goes back to Grothendieck; the version given here can be found in [13, Theorem 4.9], except for the assertion that $dm/dP\in L^0$.

1.2 THEOREM. *Let C be a subset of a locally convex space E, let $(\Omega,\mathfrak{F},P)$ be a complete probability space, and let $m\colon\mathfrak{F}\to E$ be a vector-valued measure $\ll P$. Suppose that m almost has P-average range relatively compact in C. Then there exists $dm/dP\in L^0(\Omega,\mathfrak{F},P;C)$.*

The following corollary has been proved independently by several mathematicians (see for example [4, Theorem 3.1], [9, Théorème 4.2]).

1.3 COROLLARY. *Let T be a completely regular space. Then $\mathcal{P}_t(T)$ has the Radon-Nikodym property.*

Let E be a locally convex space, $(\Omega,\mathfrak{F},P)$ a probability space, $\varphi\in L^0(\Omega,\mathfrak{F},P;E)$. Let $\mathcal{G}$ be a sub-σ-algebra of $\mathfrak{F}$. A conditional expectation of φ given $\mathcal{G}$ is a function $\psi\in L^0(\Omega,\mathcal{G},P;E)$ such that $f\circ\psi = E[f\circ\varphi|\mathcal{G}]$ for all $f\in E^*$; we write $\psi = E[\varphi|\mathcal{G}]$. If $\varphi\in L^0(\Omega;C)$, where C is a bounded t-convex subset of E, then $\psi=E[\varphi|\mathcal{G}]$ if and only if $\int_A \psi\,dP = \int_A \varphi\,dP$ for all $A\in\mathcal{G}$. Thus, if C also has the Radon-Nikodym property, then $E[\varphi|\mathcal{G}]$ necessarily exists.

Let C be a bounded convex subset of a locally convex space E. A martingale in C consists of: a probability space $(\Omega,\mathfrak{F},P)$; a directed set J; a family $(\mathfrak{F}_\alpha)_{\alpha\in J}$ of sub-σ-algebras of $\mathfrak{F}$ indexed by J such that $\mathfrak{F}_\alpha\subseteq\mathfrak{F}_\beta$ if $\alpha\le\beta$; and a family $(\varphi_\alpha)_{\alpha\in J}$ where $\varphi_\alpha\in L^0(\Omega,\mathfrak{F}_\alpha,P;C)$ and $\varphi_\alpha = E[\varphi_\beta|\mathfrak{F}_\alpha]$ if $\alpha\le\beta$. Let $(\varphi_\alpha)_{\alpha\in J}$ be a martingale in C, and let $\varphi\in L^0(\Omega,\mathfrak{F},P;C)$. We say that φ closes (φ_α) iff $\varphi_\alpha = E[\varphi|\mathfrak{F}_\alpha]$ for all $\alpha\in J$. We say that φ_α converges in measure to φ iff, for every neighborhood U of 0 in E,

$$\lim_{\alpha\in J} P\{\omega\colon \varphi_\alpha(\omega)-\varphi(\omega)\in U\} = 1 .$$

We say that φ_α converges in mean to φ iff, for every continuous seminorm q on E,

$$\lim_{\alpha\in J}\int q(\varphi_\alpha(\omega)-\varphi(\omega))\,dP(\omega) = 0 .$$

1.4 PROPOSITION. *Let* (φ_α) *be a martingale in* C , *and let* $\varphi \in L^0(\Omega, \mathfrak{F}, P; C)$. *The following are equivalent.*

(a) φ_α *converges in measure to* φ ;

(b) φ_α *converges in mean to* φ ;

(c) φ *closes* (φ_α) *and* φ *is measurable with respect to the* σ-*algebra generated by* $\bigcup_{\alpha \in J} \mathfrak{F}_\alpha$.

1.5 PROPOSITION. *Let* (φ_α) *be a martingale in* C . *Then* (φ_α) *converges in measure if and only if* $\varphi_\alpha(P)$ *converges in* $\mathcal{P}_t(C)$.

The bounded convex set C is said to have the *J-martingale convergence property* iff every martingale in C indexed by J converges in measure. (The $\mathbb{N}$-martingale convergence property will be called the sequential martingale convergence property.) Also, C is said to have the *martingale convergence property* iff it has the J-martingale convergence property for all directed sets J . The *well-ordered martingale convergence property* and the *totally-ordered martingale convergence property* are defined in a similar fashion.

2.

Let E be a locally convex space, and let C be a bounded convex subset of E . A partial order can be defined on $\mathcal{P}_t(C)$ as follows. If $\mu, \nu \in \mathcal{P}_t(C)$, define $\mu \prec \nu$ iff $\langle\mu, f\rangle \le \langle\nu, f\rangle$ for all bounded continuous convex functions f on C . This relation is clearly reflexive and transitive; the antisymmetry of the relation follows from the following result, which essentially goes back to LeCam (see [14, p. 216], [11, Lemma 2.1], [12, Lemma 3.1]).

2.1 PROPOSITION. *Let* $(T, \mathfrak{T})$ *be a completely regular space. Suppose* $F \subseteq \mathcal{C}_b(T)$ *is a class of functions which separates points of* T *and if* $f, g \in F$, *then the pointwise maximum* $f \vee g \in F$. *Let* $\mathfrak{T}'$ *be the topology on* T *generated by* F . *If* $\mu_\alpha, \mu \in \mathcal{P}_t(T, \mathfrak{T}')$ *and*

$$\lim_\alpha \int_T f d\mu_\alpha = \int_T f d\mu \quad \text{for all} \quad f \in F ,$$

then $\mu_\alpha \to \mu$ *in the weak* topology of* $\mathcal{P}_t(T, \mathfrak{T}')$. *In particular*, if $\mu, \nu \in \mathcal{P}_t(T, \mathfrak{T})$ *and* $\int f d\mu = \int f d\nu$ *for all* $f \in F$, *then* $\mu = \nu$.

Let $\mu \in \mathcal{P}_t(C)$. A μ-*dilation* is a function

$$T \in L^0(C, \mathcal{B}^\mu(C), \mu; \mathcal{P}_t(C))$$

such that for every $h \in E^*$, we have $\langle T(x), h\rangle = h(x)$ for μ-almost every x . ($\mathcal{B}(C)$) denotes the Borel sets on C ; $\mathcal{B}^\mu(C)$ the completion with respect to μ ;

$\mathcal{P}_t(C)$ is understood to have its weak* topology.) If C is t-convex, then $r(T(x))$ exists for every x, so in that case the condition is the same as the assertion that $r \circ T$ is weakly equivalent to the identity on C. If $\nu \in \mathcal{P}_t(C)$, we write $\nu = T(\mu)$, and say that μ dilates to ν, iff $\langle \nu, f\rangle = \int \langle T(x), f\rangle \, d\mu(x)$ for all $f \in \mathcal{C}_b(C)$; that is, $\nu = \int T \, d\mu$ in $\mathcal{P}_t(C)$.

The following theorem shows that the ordering $\prec$ can be characterized in terms of dilations and in terms of conditional expectations. It goes back (in the one-dimensional case) to Hardy, Littlewood and Polya. The proof here is closest to that of V. Strassen [17].

2.2. THEOREM. Let C be bounded and convex, and let $\mu, \nu \in \mathcal{P}_t(C)$. The following are equivalent:

(a) $\mu \prec \nu$; i.e. $\langle \mu, f\rangle \leq \langle \nu, f\rangle$ for all bounded continuous convex functions f on C;

(b) μ dilates to ν; i.e. there is a μ-dilation $T\colon C \to \mathcal{P}_t(C)$ such that $T(\mu) = \nu$;

(c) There exist a probability space $(\Omega, \mathfrak{F}, P)$, a σ-algebra $\mathfrak{G} \subseteq \mathfrak{F}$, and functions $\varphi, \psi \in L^0(\Omega, \mathfrak{F}, P; C)$ with $\varphi(P) = \mu$, $\psi(P) = \nu$, and $\varphi = E[\psi \mid \mathfrak{G}]$.

Proof. (a) $\Rightarrow$ (b). If $f \in \mathcal{C}_b(C)$ define its upper envelope $f^\wedge\colon C \to \mathbb{R}$ by

$$f^\wedge(x) = \inf\{h(x) : h \text{ is bounded continuous and affine on } C \text{ and } h \geq f\} .$$

Note that for fixed f, the map $x \mapsto f^\wedge(x)$ is upper semicontinuous, hence Borel measurable. Also, for fixed x, the map $f \mapsto f^\wedge(x)$ is continuous for the uniform norm on $\mathcal{C}_c(C)$. We can also write

$$f^\wedge(x) = \inf\{h(x) : h \text{ is bounded continuous and concave on } C \text{ and } h \geq f\} .$$

Let S be the vector space of all Borel-measurable simple functions $\theta\colon X \to \mathcal{C}_b(C)$. Define $p\colon S \to \mathbb{R}$ by

$$p(\theta) = \int \theta(x)^\wedge(x) \, d\mu(x) . \tag{1}$$

This integral exists since the integrand is bounded and Borel measurable. It is easily checked that p is a sublinear functional on S. For $f \in \mathcal{C}_b(C)$ and $A \in \mathcal{B}(C)$, define $\chi_A \otimes f \in S$ by

$$(\chi_A \otimes f)(x) = \begin{cases} f & \text{if } x \in A \\ 0 & \text{if } x \notin A . \end{cases}$$

Thus $f \mapsto \chi_C \otimes f$ is an embedding of $\mathcal{C}_b(C)$ in S. Define a linear functional ℓ on the range of this embedding by

$$\ell(\chi_C \otimes f) = \langle \nu, f \rangle . \tag{2}$$

We claim that $\ell(\chi_C \otimes f) \leq p(\chi_C \otimes f)$. Let h be concave bounded and continuous on C and $h \geq f$. Then $\int f\, d\nu \leq \int h\, d\nu \leq \int h\, d\mu$ (by the assumption that $\mu \prec \nu$), so by the tightness of μ, $\int f\, d\nu \leq \int f^\wedge\, d\mu$; i.e. $\ell(\chi_C \otimes f) \leq p(\chi_C \otimes f)$.

By the Hahn-Banach theorem, the linear functional ℓ can be extended to all of S with $\ell(\theta) \leq p(\theta)$ for all $\theta \in S$. Define $m\colon \mathcal{B}(C) \to \mathcal{C}_b(C)^*$ by $\langle m(A), f \rangle = \ell(\chi_A \otimes f)$. Now if h is affine, bounded and continuous on C, then $h^\wedge = h$, so $\langle m(A), h \rangle = \ell(\chi_A \otimes h) \leq p(\chi_A \otimes h) = \int_A h^\wedge\, d\mu = \int_A h d\mu$; similarly $(-h)^\wedge = -h$, so $\langle m(A), -h \rangle \leq \int_A (-h) d\mu$; therefore

$$\langle m(A), h \rangle = \int_A h\, d\mu . \tag{3}$$

In particular $\langle m(A), 1 \rangle = \mu(A)$. Again, if $f \in \mathcal{C}_b(C)$ and $f \geq 0$, then $(-f)^\wedge \leq 0$, so $\langle m(A), f \rangle = \ell(\chi_A \otimes (-f)) \leq p(\chi_A \otimes (-f)) = \int_A (-f)^\wedge\, d\mu \leq 0$; thus $\langle m(A), f \rangle \geq 0$. Now if $f \geq 0$, then $\langle m(A), f \rangle = \langle m(C), f \rangle - \langle m(C \backslash A), f \rangle \leq \langle m(C), f \rangle = \ell(\chi_C \otimes f) = \langle \nu, f \rangle$, so $0 \leq m(A) \leq \nu$. This shows that $m(A)$ is tight, so $m(A)/\mu(A) \in \mathcal{P}_t(C)$. Thus the vector-valued measure $m\colon \mathcal{B}(C) \to \mathcal{C}_b(C)^*$ has average range in $\mathcal{P}_t(C)$.

By Corollary 1.3, there is $T\colon C \to \mathcal{P}_t(C)$, $T \in L^0(C, \mathcal{B}^\mu(C), \mu; \mathcal{P}_t(C))$, such that $\int_A T\, d\mu = m(A)$ for all $A \in \mathcal{B}(C)$. Now $\langle \nu, f \rangle = \ell(\chi_C \otimes f) = \langle m(C), f \rangle = \int \langle T(x), f \rangle\, d\mu(x)$ for all $f \in \mathcal{C}_b(C)$, i.e. $\nu = T(\mu)$. Finally, if $h \in E^*$, then for $A \in \mathcal{B}(C)$, using (3) yields $\int_A \langle T(x), h \rangle\, d\mu(x) = \langle m(A), h \rangle = \int_A h\, d\mu$, so $\langle T(x), h \rangle = h(x)$ μ-a.e. Thus T is a μ-dilation.

(b) $\Rightarrow$ (c). Suppose $\nu = T(\mu)$. Let $\Omega = C \times C$, $\mathfrak{F} = \mathcal{B}(C) \times \mathcal{B}(C)$, $\mathcal{J} = \mathcal{B}(C) \times \{C, \emptyset\}$, and define $\varphi, \psi\colon \Omega \to C$ by $\varphi(x,y) = x$, $\psi(x,y) = y$. Define P on $\mathfrak{F}$ by

$$P(D) = \int T(x)(D_x)\, d\mu(x) ,$$

where $D_x = \{y \in C\colon (x,y) \in D\}$. The integrand is $\mathcal{B}^\mu(C)$-measurable since T is a dilation. It follows that

$$\int g\, dP = \int \int g(x,y)\, d[T(x)](y)\, d\mu(x)$$

for any bounded $\mathfrak{F}$-measurable function g on Ω. Now for $A \in \mathcal{B}(C)$,

$$\begin{aligned}\varphi(P)(A) &= P(\varphi^{-1}(A)) = \int T(x)((A \times C)_x) d\mu(x) \\ &= \int_A 1 \, d\mu(x) = \mu(A) , \\ \psi(P)(A) &= P(\psi^{-1}(A)) = \int T(x)((C \times A)_x) \, d\mu(x) \\ &= \int T(x)(A) \, d\mu(x) = T(\mu)(A) = \nu(A) ,\end{aligned}$$

so $\varphi(P) = \mu$, $\psi(P) = \nu$. This shows that $\varphi, \psi \in L^0(\Omega;C)$. Also, for any $A \in \mathcal{B}(C)$ and $h \in E^*$,

$$\begin{aligned}\int_{A \times C} h \circ \psi \, dP &= \int_A \int_C h(y) dT(x)(y) \, d\mu(x) \\ &= \int_A h(x) \, d\mu(x) = \int_A \int_C h(x) dT(x)(y) \, d\mu(x) \\ &= \int_{A \times C} h \circ \varphi \, d\mathbf{P} ,\end{aligned}$$

so $\varphi = E[\psi | \mathcal{S}]$.

(c) $\Rightarrow$ (a). Suppose $\varphi = E[\psi|\mathcal{S}]$ and $\varphi(P) = \mu$, $\psi(P) = \nu$. Let f be bounded, continuous and convex on C. They by Jensen's inequality, $\langle \mu, f \rangle = \int f \circ E[\psi|\mathcal{S}] dP \leq \int E[f \circ \psi|\mathcal{S}] dP = \int f \circ \psi \, dP = \langle \nu, f \rangle$. □

Here is a lemma which is probably known, but I am unable to provide a reference. The proof is, apparently, not merely are application of Zorn's Lemma, but requires well-ordering as well.

2.3. LEMMA. <u>Let</u> $(P, \leq)$ <u>be a partially ordered set. Suppose every chain in</u> P <u>has at least upper bound. Then any subset of</u> P <u>which is directed has at least upper bound.</u>

<u>Proof.</u> Let $D \subseteq P$ be a directed subset of P. Let $\mathcal{M}$ be the collection of all $A \subseteq P$ which satisfy

(a) $A \supseteq D$,

(b) if C is a chain included in A, then $\sup C \in A$.

Notice that $P \in \mathcal{M}$, so $\mathcal{M} \neq \emptyset$. Let $M = \cap \mathcal{M}$. Then $M \in \mathcal{M}$.

Next, let $\mathcal{R}$ be the collection of all $B \subseteq P$ which satisfy

(a) $D \subseteq B \subseteq M$,

(b) B is directed.

Notice that $D \in \mathcal{R}$, so $\mathcal{R} \neq \emptyset$. Apply Zorn's Lemma to $\mathcal{R}$; let $R \in \mathcal{R}$ be maximal. We claim that $R \in \mathcal{M}$. Trivially $R \supseteq D$. Suppose (for purposes of contradiction) that not every chain in R has its sup in R. Then there is a well-ordered chain in R with sup not in R. Let ξ be the least ordinal of such a chain. Let $R' = \{\sup C\colon C \text{ is a chain of order type } \xi \text{ included in } R\} \cup R$. We claim that $R' \in \mathcal{R}$. Clearly $R' \supseteq R \supseteq D$ and $R' \subseteq M$. To prove that R' is directed, let $x, y \in R'$, say $x = \sup\{x_\gamma\colon \gamma < \xi\}$, $y = \sup\{y_\gamma\colon \gamma < \xi\}$, with $x_\gamma, y_\gamma \in R$. (If $x \in R$, let $x_\gamma = x$ for all $\gamma < \xi$.) Define $z_\gamma \in R$ inductively so that

$$z_{\gamma+1} \geq x_\gamma \,,\; z_{\gamma+1} \geq y_\gamma \,,\; z_{\gamma+1} \geq z_\gamma \tag{1}$$

and if $\gamma < \xi$ is a limit ordinal,

$$z_\gamma = \sup\{z_\beta\colon \beta < \gamma\} \quad . \tag{2}$$

Now (1) is possible since R is directed and (2) is possible by the minimality of ξ. Then $z = \sup\{z_\gamma\colon \gamma < \xi\} \in R'$ and $z \geq x$, $z \geq y$. Thus R' is directed, so $R' \in \mathcal{R}$. By the maximality of R, we have $R' = R$, so in fact every chain in R has its sup in R. Thus $R \in \mathcal{M}$. Then we have $R \supseteq M$, so $R = M$. This chows that M is directed.

Next, we claim that every subset S of D has an upper bound in M. To prove this, well-order $S = \{x_\gamma\colon \gamma < \alpha\}$, and proceed by induction on α. For $\alpha = 1$, $S = \{x_0\} \subseteq M$. If $\alpha = \beta + 1$, let $y \in M$ be an upper bound for $\{x_\gamma\colon \gamma < \beta\}$, which exists by the induction hypothesis; since M is directed, there is an upper bound for $\{y, x_\beta\}$ in M. If α is a limit ordinal, construct inductively $y_\gamma \geq x_\gamma$, $y_\gamma \in M$, y_γ increasing. Since $M \in \mathcal{M}$, $\sup\{y_\gamma\colon \gamma < \alpha\} \in M$ and is the required upper bound.

In particular, D itself has an upper bound in M, say x_0. If y is any upper bound for D in P, then $\{x \in P\colon x \leq y\} \in \mathcal{M}$, so $M \subseteq \{x \in P\colon x \leq y\}$, and hence $x_0 \leq y$. Therefore x_0 is the least upper bound of D, q.e.f. $\square$

2.4. THEOREM. Let C be a bounded convex set in a locally convex space. Consider the following conditions:

(a) C has both the Radon-Nikodym property and the sequential martingale convergence property.

(b) C has the well-ordered martingale convergence property.

(c) C has the martingale convergence property.

(b') Every well-ordered subset of $\mathcal{P}_t(C)$ has a least upper bound.

(c') Every directed subset of $\mathcal{P}_t(C)$ has a least upper bound.

Then: (b), (c), (b'), (c') are equivalent and imply (a). If C is t-convex, then all five conditions are equivalent.

Proof. (b) $\Rightarrow$ (c). Suppose C has the well-ordered martingale convergence property. Let D be a directed set, and suppose $(\varphi_\alpha)_{\alpha \in D}$ is a martingale with values in C ; the additional data is $(\Omega, \mathfrak{F}, P)$, $(\mathfrak{F}_\alpha)_{\alpha \in D}$. There is a partially ordered set $Q \supseteq D$ in which each chain has a least upper bound, namely let Q be the set of ideals of D [1, p. 113]. As in the proof of Lemma 2.3, there is a directed set M , $D \subseteq M \subseteq Q$, such that every chain in M has sup in M , and if $D \subseteq A \subseteq M$ and every chain in A has sup in A , then $A = M$. For each $\gamma \in M$, define $\mathfrak{F}_\gamma$ to be the σ-algebra generated by $\bigcup\{\mathfrak{F}_\alpha : \alpha \in D, \alpha \leq \gamma\}$. Let $A = \{\gamma \in M$: there exists $\varphi_\gamma \in L^0(\Omega, \mathfrak{F}_\gamma, P; C)$ such that $\varphi_\alpha = E[\varphi_\gamma | \mathfrak{F}_\alpha]$ for all $\alpha \in D$ with $\alpha \leq \gamma\}$. Note that such a function φ_γ (if it exists) is unique as an element of $L^0(\Omega, \mathfrak{F}_\gamma, P; C)$. Hence, if $\gamma, \gamma' \in A$, and $\gamma \leq \gamma'$, then $\varphi_\gamma = E[\varphi_{\gamma'} | \mathfrak{F}_\gamma]$.

Clearly $D \subseteq A \subseteq M$. We claim that every chain in A has sup in A . Let B be a chain included in A ; write $\gamma_0 = \sup B$. Now B has a well-ordered cofinal subset B_0 . Then $(\varphi_\gamma)_{\gamma \in B_0}$ is a well-ordered martingale, so it converges; by Proposition 1.4, its limit φ_{γ_0} satisfies $\varphi_\alpha = E[\varphi_{\gamma_0} | \mathfrak{F}_\alpha]$ for all $\alpha \in D$, $\alpha \leq \gamma_0$, so $\gamma_0 \in A$. (By the choice of Q , if $\alpha \leq \gamma_0$, then $\alpha \leq \gamma$ for some $\gamma \in B_0$.) Thus, every chain in A has sup in A , so we have $A = M$.

Now M is directed, so M has a largest element γ^* . The martingale $(\varphi_\alpha)_{\alpha \in D}$ is closed by φ_{γ^*} , so (φ_α) converges.

(c) $\Rightarrow$ (b) is trivial.

(c) $\Rightarrow$ (a). Suppose C has the martingale convergence property. Then C trivially has the sequential martingale convergence property. Let $(\Omega, \mathfrak{F}, P)$ be a probability space and $m: \mathfrak{F} \to E$ a vector-valued measure with $m \ll P$ and average range included in C . Let D be the set of finite $\mathfrak{F}$-measurable partitions of Ω , ordered by a.e. refinement. For $\alpha = \{A_1, A_2, \dots, A_n\} \in D$, let $\mathfrak{F}_\alpha$ be the σ-algebra generated by $\{A_1, \dots, A_n\}$, and define $\varphi_\alpha : \Omega \to E$ by

$$\varphi_\alpha = \sum_{A \in \alpha} \frac{m(A)}{P(A)} \chi_A ,$$

where we interpret $0/0$ as an arbitrary element of C . Then $\varphi_\alpha \in L^0(\Omega, \mathfrak{F}_\alpha, P; C)$, and if $\alpha \leq \beta$, we have $\varphi_\alpha = E[\varphi_\beta | \mathfrak{F}_\alpha]$. Thus (φ_α) is a martingale with values in C . Let $\varphi \in L^0(\Omega, \mathfrak{F}, P; C)$ be the limit of (φ_α) . Thus $\varphi_\alpha = E[\varphi | \mathfrak{F}_\alpha]$ for all $\alpha \in D$ by Proposition 1.4. Let $A \in \mathfrak{F}$. Then $\alpha = \{A, \Omega \backslash A\} \in D$, and we have

$$m(A) = \frac{m(A)}{P(A)} \cdot P(A) = \int_A \varphi_\alpha \, dP = \int_A \varphi \, dP ,$$

so $dm/dP = \varphi$. Thus C has the Radon-Nikodym property.

(a) $\Rightarrow$ (b). Let C be t-convex. Suppose C has the Radon-Nikodym property and the sequential martingale convergence property. Let $(\varphi_\alpha)_{\alpha<\xi}$ be a martingale indexed by the ordinals less than the ordinal ξ . If $\xi = \beta + 1$, then $(\varphi_\alpha)_{\alpha<\xi}$ is closed by φ_β . If ξ has countable cofinality, then convergence follows from the sequential martingale convergence property. Suppose that ξ has uncountable cofinality, that is, any countable set of ordinals $<\xi$ has an upper bound $<\xi$. Let $\mathfrak{F}_\infty$ denote the σ-algebra generated by $\bigcup\{\mathfrak{F}_\alpha\colon \alpha<\xi\}$. Since ξ has uncountable cofinality, $\mathfrak{F}_\infty = \bigcup\{\mathfrak{F}_\alpha\colon \alpha<\xi\}$. For $A \in \mathfrak{F}_\infty$, say $A \in \mathfrak{F}_\alpha$, let $m(A) = \int_A \varphi_\alpha \, dP$. This integral exists since C is t-convex. The average range of m is included in C , so by the Radon-Nikodym property, there exists $\varphi = dm/dP$. Now if $A \in \mathfrak{F}_\alpha$, $\int_A \varphi \, dP = m(A) = \int_A \varphi_\alpha dP$, so $\varphi_\alpha = E[\varphi|\mathfrak{F}_\alpha]$. Thus $(\varphi_\alpha)_{\alpha<\xi}$ converges to φ .

(b') $\Rightarrow$ (c') follows from Lemma 2.3; (b') $\Rightarrow$ (b) follows from Proposition 1.5.

(c') $\Rightarrow$ (b') is trivial.

(b) $\Rightarrow$ (b'). Suppose C has the well-ordered martingale convergence property. Let $(\mu_\gamma)_{\gamma<\xi}$ be a well-ordered subset of $\mathcal{P}_t(C)$. Fir $\alpha \leq \xi$, let Ω_α be a product of copies of C indexed by the ordinals $<\alpha$; equip Ω_α with the product σ-algebra; for $\gamma<\alpha\leq\xi$, define $\varphi_{\alpha,\gamma}\colon \Omega_\alpha \to C$ by $\varphi_{\alpha,\gamma}(\omega) = \omega(\gamma)$; and let $\mathfrak{F}_{\alpha,\gamma}$ be the σ-algebra on Ω_α generated by $(\varphi_{\alpha,\beta})_{\beta\leq\gamma}$. We will define inductively a measure P_α on Ω_α so that, for each α , $(\varphi_{\alpha,\gamma}, \mathfrak{F}_{\alpha,\gamma})_{\gamma<\alpha}$ forms a martingale and $\varphi_{\alpha,\gamma}(P_\alpha) = \mu_\gamma$. On $\Omega_1 = C$, let $P_1 = \mu_0$ so that $\varphi_{1,0}(P_1) = \mu_0$. Suppose $\alpha > 1$ and P_β has been defined for all $\beta<\alpha$.

First, suppose α is a limit ordinal. Then Ω_α is the inverse limit of $(\Omega_\beta)_{\beta<\alpha}$ and the measures P_β are consistent, so there is a unique extension P_α to Ω_α consistent with the P_β .

Next, suppose $\alpha = \beta+1$. Then $(\varphi_{\beta,\gamma})_{\gamma<\beta}$ is a martingale, hence converges, say to $\psi_\beta\colon \Omega_\beta \to C$. (If $\beta = \beta'+1$, then of course $\psi_\beta = \varphi_{\beta,\beta'}$.) Now the measure $\nu_\beta = \psi_\beta(P_\beta)$ is the least upper bound of $(\mu_\gamma)_{\gamma<\beta}$, so $\nu_\beta \prec \mu_\beta$. Choose a ν_β-dilation T_β so that $T_\beta(\nu_\beta) = \mu_\beta$. Define P_α on $\Omega_\alpha = \Omega_\beta \times C$ by:

$$P_\alpha(A) = \int_C T_\beta(x)(A_\omega)d\nu_\beta(x) \quad ,$$

where $A_\omega = \{x \in C\colon (\omega,x) \in A\}$ for $\omega \in \Omega_\beta$. Thus $\varphi_{\alpha,\beta}(P_\alpha) = \mu_\beta$ and $E[\varphi_{\alpha,\beta}|\mathfrak{F}_{\alpha,\beta'}] = \varphi_{\alpha,\beta'}$ for $\beta' \leq \beta$. This completes the inductive definition of $(P_\alpha)_{\alpha\leq\xi}$.

As before, $(\varphi_{\xi,\gamma})_{\gamma<\xi}$ is a martingale, hence converges, say to $\psi_\xi\colon \Omega_\xi \to C$. The measure $\nu_\xi = \psi_\xi(P_\xi)$ is the least upper bound of $(\mu_\gamma)_{\gamma<\xi}$. $\square$

If the condition (b') or (c') is satisfied, then (accoring to Zorn's Lemma) for every $\mu \in \mathcal{P}_t(C)$, there is a maximal $\lambda \in \mathcal{P}_t(C)$ with $\lambda \succ \mu$. This fact is relevant in Choquet-type representation theorems: see [16, p. 25], [6, p. 157].

Next is a result giving conditions under which every maximal measure is concentrated on the extreme points in the sense that $\mu(B) = 1$ for every Borel set $B \supseteq \text{ex } C$. (This is not true in general, even if C is a closed bounded convex set in a Banach space: see [6, p. 159].) Of course, part (iii) is weaker then the known Bishop-de Leeuw Theorem.

2.5 PROPOSITION. *Let C be a bounded convex subset of a locally convex space. Suppose either*

(i) C *is analytic; or*

(ii) C *is completely metrizable and* $r: \mathcal{P}_t(C) \to C$ *is open; or*

(iii) C *is compact and* r *is open.*

Then if $\mu \in \mathcal{P}_t(C)$ *is maximal, we have* $\mu(B) = 1$ *for every Borel set* $B \supseteq \text{ex } C$.

Proof. We will prove the contrapositive. Suppose $\mu(B) < 1$ for some Borel set $B \supseteq C$. Then we may assume without loss of generality that there is a compact set $K \subseteq C \setminus \text{ex } C$ with $\mu(K) = 1$. In each of the three cases, C is t-convex, so $r: \mathcal{P}_t(C) \to C$ is defined. Let $\mathcal{R} = \{\epsilon_x : x \in C\}$. Now $\mathcal{R}$ is closed in $\mathcal{P}_t(C)$ and, in case (i) r is continuous on $\mathcal{P}_t(C)\setminus\mathcal{R}$, which is analytic; in case (ii) r is continuous and open on $\mathcal{P}_t(C)\setminus\mathcal{R}$ which is completely metrizable; in case (iii) r is continuous and open on $\mathcal{P}_t(C)\setminus\mathcal{R}$ which is locally compact. Then (i) by the von Neumann selection theorem or (ii) (iii) by [7], there is a measurable weak section $T: C\setminus\text{ex } C \to \mathcal{P}_t(C)\setminus\mathcal{R}$, i.e. $r \circ T$ is weakly equivalent to the identity on $C\setminus\text{ex } C$. Define $T(x) = \epsilon_x$ for $x \in \text{ex } C$, so that T is a μ-dilation. We claim that $T(\mu) \neq \mu$; this will show that μ is not maximal.

Now $T \in L^0$, so there is a compact set $\mathcal{D} \subseteq \mathcal{P}_t(C)\setminus\mathcal{R}$ with $\mathcal{D} \subseteq T(K)$ and $\mu(T^{-1}(\mathcal{D}) \cap K) > 0$. But $\mathcal{D}$ is compact, $\mathcal{R}$ is closed, and $\mathcal{D} \cap \mathcal{R} = \emptyset$, so there is a continuous function $h: \mathcal{P}_t(C) \to \mathbb{R}$ such that $h = 0$ on $\mathcal{D}$, $h = 1$ on $\mathcal{R}$ and $0 \leq h \leq 1$. Now $\mu\{x: h(T(x)) \neq h(\epsilon_x)\} > 0$. But the topology of $\mathcal{P}_t(C)$ is generated by the set of maps $\nu \mapsto \int f\, d\nu$, where f is bounded continuous convex (Proposition 2.1). Thus there is a bounded continuous convex function f such that $\mu\{x: \langle T(x), f\rangle \neq \langle \epsilon_x, f\rangle\} > 0$. Thus $\langle T(\mu), f\rangle = \int \langle T(x), f\rangle d\mu(x) > \langle \mu, f\rangle$. Therefore $T(\mu) \neq \mu$, and thus μ is not maximal. $\square$

Condition (iii) is studied by R. C. O'Brien in [15].

Let $(\Omega, \mathcal{F}, P)$ be a probability space and C a bounded convex set. Write $\mathcal{F}^+ = \{A \in \mathcal{F}: P(A) > 0\}$. A function $\eta: \mathcal{F}^+ \to C$ is an *averaged measure* provided

$$\eta(A \cup B) = \frac{P(A)}{P(A \cup B)}\, \eta(A) + \frac{P(B)}{P(A \cup B)}\, \eta(B)$$

for disjoint $A, B \in \mathfrak{F}^+$. Clearly, if m is a measure with average range in C, then m/P is an averaged measure with values in C and conversely (countable additivity of $m = P \bullet \eta$ follows from the boundedness of C).

2.6 PROPOSITION. Let C', C be bounded convex sets in locally convex spaces. Suppose $u: C' \to C$ is continuous, bijective and affine. If C' has the Radon-Nikodym property, then C has the Radon-Nikodym property.

Proof. Let $(\Omega, \mathfrak{F}, P)$ be a probability space, m a measure with average range in C. Then $\eta = m/P$ is an averaged measure in C. Define $\eta': \mathfrak{F}^+ \to C'$ by $\eta'(A) = u^{-1}(\eta(A))$. Since u is bijective and affine, η' is an averaged measure in C'. Since C' has the Radon-Nikodym property, there is $\varphi' \in L^0(\Omega, \mathfrak{F}, P; C')$ with $\eta'(A) = P(A)^{-1} \int_A \varphi' \, dP$ for all $A \in \mathfrak{F}^+$. Let $\varphi = u \circ \varphi'$. Now u is continuous, so $\varphi \in L^0(\Omega, \mathfrak{F}, P; C)$. Also u is continuous and affine, so $\eta(A) = u(\eta'(A)) = P(A)^{-1} \int_A u \circ \varphi \, dP = P(A)^{-1} \int_A \varphi \, dP$. Thus C has the Radon-Nikodym property. □

2.7 COROLLARY. Let C be a bounded convex set in a locally convex space. Write $\operatorname{ex} C$ for the set of extreme points of C. Suppose that

(1) $\operatorname{ex} C$ is relatively t-convex in C; i.e. for every $\mu \in \mathcal{P}_t(\operatorname{ex} C)$, there exists $r(\mu) \in C$;

(2) for every $x \in C$; there is a unique $\mu \in \mathcal{P}_t(\operatorname{ex} C)$ with $r(\mu) = x$.

Then C has the Radon-Nikodym property.

Proof. First, $\mathcal{P}_t(\operatorname{ex} C)$ has the Radon-Nikodym property by Corollary 1.3. The resultant map $r: \mathcal{P}_t(\operatorname{ex} C) \to C$ is defined by (1) and bijective by (2); it is always affine and continuous. Thus by Proposition 2.6, C has the Radon-Nikodym property. □

Remarks. (1) For example, if C is a separable closed bounded convex subset of a Banach space, then $\operatorname{ex} C$ is universally measurable [2, Prop. 2.1], so $\mathcal{P}_t(\operatorname{ex} C)$ can be identified with $\{\mu \in \mathcal{P}_t(C): \mu(\operatorname{ex} C) = 1\}$. Thus, if every point of C is represented by a unique measure on $\operatorname{ex} C$, then C has the Radon-Nikodym property. This is a (very) partial converse of [5].

(2) If C is a separable closed bounded subset of a Banach space and C is a (noncompact) simplex, does it follow that a point of C can have at most one representing measure on $\operatorname{ex} C$? (The Radon-Nikodym property is not postulated, cf. [3, Theorem 1.1].)

(3) If C is a nonseparable closed bounded convex subset of a Hilbert space, the set of maximal measures on C need not have the Radon-Nikodym property (the example in [6, p. 159] exhibits this behavior), so Proposition 2.6 will not apply in this case.

Note. After this paper was written, H. von Weizsäcker kindly gave me a copy of his paper "Einige masstheoretische Formen der Sätze von Krein-Milman und Choquet". It has considerable overlap with the present paper. Among many other things, von Weizsäcker gives an example of a completely regular space T for which $\mathcal{P}_t(T)$ fails the martingale convergence property. (See von Weizsäcker's paper in this volume.)

References

1. G. Birkhoff, Lattice Theory, Third ed., American Mathematical Society, Providence, R.I., 1967.
2. R. D. Bourgin, Barycenters of measures on certain noncompact convex sets, Trans. Amer. Math. Soc. 154 (1971) 323-340.
3. R. D. Bourgin and G. A. Edgar, Noncompact simplexes in spaces with the Radon-Nikodym property, J. Functional Analysis 23 (1976) 162-176.
4. G. A. Edgar, Disintegration of measures and the vector-valued Radon-Nikodym theorem, Duke Math. J. 42 (1974) 447-450.
5. G. A. Edgar, A noncompact Choquet theorem, Proc. Amer. Math. Soc. 49 (1975) 354-358.
6. G. A. Edgar, Extremal integral representations, J. Functional Analysis 23 (1976) 145-161.
7. G. A. Edgar, Measurable weak sections, Illinois J. Math. 20 (1976) 630-646.
8. D. H. Fremlin and I. Pryce, Semiextremal sets and measure representation, Proc. London Math. Soc. (3) 29 (1974) 502-520.
9. A. Goldman, Measures cylindriques, measures vectorielles et questions de concentration cylindrique, Pacific J. Math. (to appear).
10. J. Hoffmann-Jørgensen, Weak compactness and tightness of subsets of $M(X)$, Math. Scand. 31 (1972) 127-150.
11. J. Hoffmann-Jørgensen, The strong law of large numbers and the central limit theorem in Banach spaces, Aarhus Universitet, Various Publications Series no. 24, pp. 74-99.
12. J. Hoffmann-Jørgensen and G. Pisier, The law of large numbers and the central limit theorem in Banach spaces, Ann. of Prob. 4 (1976) 587-599.
13. J. Kupka, Radon-Nikodym theorems for vector-valued measures, Trans. Amer. Math. Soc. 169 (1972) 197-217.
14. L. LeCam, Convergence in distribution of stochastic processes, Univ. of Calif. Publ. in Stat. 2 (1957) 207-236.
15. R. C. O'Brien, On the openness of the barycentre map, Math. Ann. 223 (1976) 207-212.
16. R. R. Phelps, Lectures on Choquet's Theorem, Van Nostrand, New York, 1966.
17. V. Strassen, The existence of probability measures with given marginals, Ann. of Math. Stat. 36 (1965) 423-439.
18. V. S. Varadarajan, Measures on topological spaces, Amer. Math. Soc. Transl. (2) 48 (1965) 161-228.

19. H. von Weizsäcker, Der Satz von Choquet-Bishop-de Leeuw für konvexe nicht kompakte Mengen straffer Masse über beliebigen Grundräumen, Math. Z. 142 (1975) 161-165.

The Ohio State University
Columbus, Ohio 43210 U.S.A.