ON THE SETWISE CONVERGENCE OF SEQUENCES OF MEASURES

JEAN B. LASSERRE
LAAS-CNRS, 7 Av. du Colonel Roche
31077 Toulouse Cedex, France
lasserre@laas.fr

(Received April, 1996; Revised July, 1996)

We consider a sequence \( \{\mu_n\} \) of (nonnegative) measures on a general measurable space \((X, \mathcal{B})\). We establish sufficient conditions for setwise convergence and convergence in total variation.

Key words: Measures, Convergence of Sequences of Measures.
AMS subject classifications: 28A33, 28C15.

1. Introduction

Consider a sequence \( \{\mu_n\} \) of (nonnegative) measures on a measurable space \((X, \mathcal{B})\) where \( X \) is some topological space. Setwise convergence of measures, as opposed to weak* or weak convergence, is a highly desirable and strong property. If proved, some important properties can be derived (for instance, the Vitali-Hahn-Saks Theorem) and thus sufficient conditions ensuring this type of convergence are of interest. However, as noted in [2], in contrast to weak or weak* convergence (for instance, in metric spaces), it is in general difficult to exhibit such a property, except if e.g., \( \mu_n \) is an increasing or decreasing sequence (e.g. [2], [4]).

The present paper provides two simple sufficient conditions. Thus, for instance, in a locally compact Hausdorff space, an order-bounded sequence of probability measures that is vaguely or weakly convergent is in fact setwise convergent.

We also establish a sufficient condition for the convergence in total variation norm that is even a stronger property.

2. Notations and Definitions

Let \((X, \mathcal{B})\) be a measurable space and let \( \mathcal{M}b(X) \) denote the Banach space of all bounded measurable real-valued functions on \( X \) equipped with the sup-norm. Let \( S \) be the positive (convex) cone in \( \mathcal{M}b(X) \).

Let \( \mathcal{M}b(X)' \) be the (Banach) topological dual of \( \mathcal{M}b(X) \) with the duality bracket \( \langle \cdot, \cdot \rangle \) between \( \mathcal{M}b(X) \) and \( \mathcal{M}b(X)' \). \( \mathcal{M}b(X)' \) is equipped with the dual norm \( |\varphi| := \sup \|f\| = 1 \langle \varphi, f \rangle \). Let \( S' \in \mathcal{M}b(X)' \) be the positive cone in \( \mathcal{M}b(X)' \), i.e.,
the dual cone of $S \in \mathcal{M}(X)$. Convergence in the weak* topology of $\mathcal{M}(X)'$ is denoted by $\overset{w^*}{\rightarrow}$.

If $X$ is a topological space, then $C(X)$ denotes the Banach space of all real-valued bounded continuous functions on $X$, and if $X$ is locally compact Hausdorff, $C_0(X)$ ($C_c(X)$ resp.) denotes the Banach space of real-valued continuous functions that vanish at infinity (with compact support, resp.).

In the sequel, the term *measure* will stand for a nonnegative $\sigma$-additive measure and a set function on $\mathcal{B}$ with the finite-additivity property (and not necessarily the $\sigma$-additivity property) will be referred to as a *finitely additive measure*. Let $M(X)$ be the Banach space of signed measures on $(X, \mathcal{B})$ equipped with the total variation norm $\| \cdot \|_{TV}$, simply denoted $| \cdot |$.

Note that $M(X) \subset \mathcal{M}(X)'$ and for every $f \in \mathcal{M}(X)$, $\mu \in M(X)$, $\int f \, d\mu = \langle \mu, f \rangle$ when $\mu$ is considered to be an element of $\mathcal{M}(X)'$.

Also note that any element $\varphi \in \mathcal{S}'$ can be associated with a finitely additive nonnegative measure (also denoted $\varphi$) $\varphi(A) := \langle \varphi, 1_A \rangle$, $\forall A \in \mathcal{B}$, so that $\varphi(A \cup B) = \varphi(A) + \varphi(B)$, $\forall A, B \in \mathcal{B}$ with $A \cap B = \emptyset$. Thus, $\varphi(A) \leq \varphi(X) = | \varphi |$, $\forall A \in \mathcal{B}$ (see e.g., [3]).

For a topological space $X$, by analogy with sequences of probability measures in a metric space, a sequence of measures $\{ \mu_n \}$ in $M(X)$ is said to converge *weakly* to $\mu \in M(X)$, iff

$$\int f \, d\mu_n \rightarrow \int f \, d\mu, \quad \forall f \in C(X).$$

This type of convergence is denoted $\mu_n \overset{w}{\rightarrow} \mu$.

Similarly, and again, by analogy with sequences of probability measures in a metric space, if $X$ is a locally compact Hausdorff space, a sequence of measures $\{ \mu_n \}$ in $M(X)$ is said to converge *vaguely* to $\mu \in M(X)$, iff

$$\int f \, d\mu_n \rightarrow \int f \, d\mu, \quad \forall f \in C_0(X),$$

and this type of convergence is denoted $\mu_n \overset{\text{vaguely}}{\rightarrow} \mu$. In fact, because the topological dual of $C_0(X)$ is $M(X)$ (see e.g., [1]), the vague convergence is simply the weak* convergence in $M(X)$.

A sequence $\{ \mu_n \}$ in $M(X)$ is said to converge *setwise* to $\mu \in M(X)$ iff

$$\mu_n(B) \rightarrow \mu(B), \quad \forall B \in \mathcal{B},$$

and this convergence is denoted $\mu_n \overset{\text{setwise}}{\rightarrow} \mu$.

Finally, a sequence $\{ \mu_n \}$ in $M(X)$ converges to $\mu \in M(X)$ in *total variation* (or convergences strongly (or in norm) to $\mu$) iff $| \mu_n - \mu | \rightarrow 0$ as $n \rightarrow \infty$. This convergence is denoted by $\mu_n \overset{TV}{\rightarrow} \mu$.

3. Preliminaries

In this section, we present some results that we will repeatedly use in the sequel.

For a nonnegative finitely additive measure $\mu$, proceeding as in [6], let:

$$\Gamma(\mu) := \{ \nu \in M(X) \mid 0 \leq \nu \leq \mu \}, \quad \Delta(\mu) := \{ \nu \in M(X) \mid \mu \leq \nu \},$$
where by $\nu \leq \mu$ we mean $\nu(A) \leq \mu(A), \forall A \in \mathcal{B}$.

Given two ($\sigma$-additive) measures $\varphi$ and $\psi$, let

$$\sup(\varphi, \psi) = \varphi \lor \psi = \frac{|\varphi - \psi| + \varphi + \psi}{2}, \quad \inf(\varphi, \psi) = \varphi \land \psi = \frac{\varphi + \psi - |\varphi - \psi|}{2},$$

where for a signed measure $\gamma$, $|\gamma|$ is its corresponding total variation measure. With the partial ordering $\leq$, $M(X)$ is a complete Banach lattice (see e.g., [5]).

**Lemma 3.1:** Let $\mu$ be (nonnegative) finitely additive measure. Then,

(i) $\Gamma(\mu)$ has a maximal element $\varphi \in M(X)^+$. 

(ii) If $\Delta(\mu) \neq \emptyset$, then $\mu$ is $\sigma$-additive.

**Proof:** To prove that $\Gamma(\mu)$ has a maximal element we use arguments similar to those in [6]. Let $\delta := \sup_{\nu \in \Gamma(\mu)} \nu(X)$. Of course, we have $\delta \leq \mu(X) = 1$. Thus, consider a sequence $\{\nu_n\} \in M(X)^+$, with $\nu_n(X) \leq c$. Define

$$\varphi_n \equiv \nu_1 \lor \nu_2 \lor \ldots \lor \nu_n. \quad (3.1)$$

$\{\varphi_n\}$ is an increasing sequence in $\Gamma(\mu)$. Indeed, for any two measures $\tau$ and $\chi$ in $\Gamma(\mu)$, $(\tau \lor \chi)(A) \leq \mu(A) \forall A \in \mathcal{B}$.

Since $\varphi_n(A) \leq \mu(A), \forall A \in \mathcal{B}$, and $\varphi_n$ is increasing, it converges setwise to an element $\varphi \leq \mu$. That $\varphi$ is a ($\sigma$-additive) measure follows from the fact that $\{\varphi_n\}$ is an increasing sequence (see e.g., [2]). It follows that $\varphi \in \Gamma(\mu)$ and $\varphi(X) = \delta$. We now prove that $\varphi$ is a maximal element of $\Gamma(\mu)$.

Consider any element $\chi \in \Gamma(\mu)$. Assume that there is some $A \in \mathcal{B}$ such that $\chi(A) > \varphi(A)$. Let $\tau := \chi \lor \varphi$. From the Hahn-Jordan decomposition of $(\chi - \varphi)$, $\exists X_1, X_2$ with $X_1 \cup X_2 = X, X_1 \cap X_2 = \emptyset$ so that

$$\tau(A) = \chi(A \cap X_1) + \varphi(A \cap X_2), \quad A \in \mathcal{B}.$$ 

Thus, $\chi(X_1) > \varphi(X_1)$ and, therefore,

$$\tau(X) = \chi(X_1) + \varphi(X_2) > \varphi(X_1 \cup X_2) = \delta,$$

is a contradiction with $\tau \in \Gamma(\mu)$ and $\delta = \max\{\nu(X) | \nu \in \Gamma(\mu)\}$. Hence, $\chi \leq \varphi$. In fact, $\varphi$ is the $\sigma$-additive part in the decomposition of $\mu$ into a $\sigma$-additive part $\mu_c$ and a purely finitely additive part $\mu_p$, with $\mu = \mu_c + \mu_p$ (see [6]).

To prove (ii), note that if $0 \leq \mu \leq \psi$, where $\psi$ is $\sigma$-additive, then $\mu$ is $\sigma$-additive (see e.g., [3], [6]). Indeed, for any decreasing sequence of sets $\{A_n\}$ in $\mathcal{B}$ with $A_n \downarrow \emptyset$, we have $0 \leq \mu(A_n) \leq \psi(A_n) \downarrow 0$ which implies $\mu(A_n) \downarrow 0$, i.e., $\mu$ is $\sigma$-additive. $\square$

**Lemma 3.2:** Let $\{\nu_n\}$ be a sequence of (nonnegative) $\sigma$-additive measures on $(X, \mathcal{B})$ with $\sup_n \nu_n(X) < \infty$. Then,

$$O - \lim \inf \nu_n = \bigvee_{k \geq 1} \bigwedge_{n \geq k} \nu_n \text{ is a (finite) } \sigma\text{-additive measure.} \quad (3.2)$$

If $\nu_n \leq \nu(\in \Gamma(M(X))) \forall n$, then,

$$O - \lim \sup \nu_n = \bigwedge_{k \geq 1} \bigvee_{n \geq k} \nu_n \text{ is a (finite) } \sigma\text{-additive measure.} \quad (3.3)$$

**Proof:** Let $\varphi_{kn} := \mu_k \land \mu_{k+1} \land \ldots \land \mu_n$. $\{\varphi_{kn}\}_{n \geq 0}$ is a decreasing sequence so that it converges setwise to a (finite) $\sigma$-additive measure $\varphi_k$ (see e.g., [4]). In turn, $\{\varphi_k\}$ is an increasing sequence with $\sup_k \varphi_k(X) < \infty$. Hence, $\varphi_k$ converges setwise to a (finite) $\sigma$-additive measure and (3.2) follows from
\( \forall k \geq 1 \ \land \ \mu_n = \lim_{k \to \infty} \varphi_k \).

(3.3) follows from \( 0 \leq \mu_n \leq \nu \) for all \( n \) and the fact that \( M(X) \) is a complete Banach lattice (see e.g., [5]).

In the sequel, to avoid confusion, the reader should be careful in distinguishing

\[(O - \liminf_{n \to \infty} \mu_n)(B) \text{ from } \liminf_{n \to \infty} \mu_n(B) \quad B \in \mathcal{B},\]

for we have in fact,

\[(O - \liminf_{n \to \infty} \mu_n)(B) \leq \liminf_{n \to \infty} \mu_n(B) \quad B \in \mathcal{B}.\]

4. Setwise Convergence

We now give sufficient conditions for setwise convergence of an \( F \)-converging sequence, where \( F \) is a subset of \( \mathcal{M}(X) \) separating points of \( M(X) \) and such that \( \mu \in M(X) \) and \( 0 \leq \int f d\mu \ \forall f \in F \) imply \( \mu \geq 0 \). Typical examples of such \( F \) are \( C(X) \) for a topological space \( X \) and \( C_c(X) \) or \( C_0(X) \) for a locally compact Hausdorff space \( X \).

**Lemma 4.1:** Let \( F \) be a subspace of \( \mathcal{M}(X) \) separating points in \( M(X) \) and such that for every \( \mu \in M(X) \), \( 0 \leq \int f d\mu \ \forall f \in F \) yields that \( \mu \geq 0 \). Let \( \{ \mu_n \} \) be a sequence of (nonnegative) measures on \( (X, \mathcal{B}) \) with \( \sup_n \mu_n(X) < \infty \). Assume that \( \mu_n \to \mu \in M(X) \), i.e.,

\[\int f d\mu_n \to \int f d\mu \ \forall f \in F. \quad (4.1)\]

(i) If \( (O - \liminf_{n \to \infty} \mu_n)(X) = \mu(X) \) then \( \mu_n \to \mu \text{ setwise} \).

(ii) If for some \( \nu \in M(X) \), \( \mu_n \leq \nu \ \forall n \), then \( \mu_n \to \mu \text{ setwise} \).

(iii) If \( (O - \liminf_{n \to \infty} \mu_n)(X) = \mu(X) = (O - \limsup_{n \to \infty} \mu_n)(X) \) then \( \mu_n \to \mu \text{ TV} \).

**Proof:** (i) Since \( \sup_n \mu_n(X) < \infty \), the sequence \( \{ \mu_n \} \) is in a weak* compact set in \( \mathcal{M}(X)' \). Thus, there is a directed set \( D \) and a subnet (not a subsequence in general) \( \{ \mu_n, \alpha \in D \} \) that converges to some \( 0 \leq \varphi \in \mathcal{M}(X)' \) for the weak* topology in \( \mathcal{M}(X)' \), and \( \varphi \) is a finitely additive measure. From (3.2) in Lemma 3.2, \( O - \liminf_{n \to \infty} \mu_n \) exists and \( \varphi \geq O - \liminf_{n \to \infty} \mu_n \). Now, \( \varphi \) has a unique decomposition into a \( \sigma \)-additive (nonnegative) part \( \varphi_c \) and a purely finitely additive (nonnegative) part \( \varphi_p \) with \( \varphi = \varphi_c + \varphi_p \) (see e.g., [6]).

From \( \sigma \)-additivity of \( O - \liminf_{n \to \infty} \mu_n \) and \( O - \liminf_{n \to \infty} \mu_n \leq \varphi_c \) since \( \varphi_c \) is a maximal element of \( \Gamma(\varphi) \) (see the proof of Lemma 3.1). Therefore, \( \varphi_c(X) \geq (O - \liminf_{n \to \infty} \mu_n)(X) = \mu(X) \). In addition,

\[\langle \mu, f \rangle = \langle \varphi_c, f \rangle + \langle \varphi_p, f \rangle \ \forall f \in F.\]

In particular, for \( f \geq 0 \) in \( F \),

\[\int f d(\mu - \varphi_c) \geq 0 \ \forall 0 \leq f \in F.\]

Thus, \( \mu \geq \varphi_c \) and \( \mu(X) = \varphi_c(X) \) which in turn implies \( \varphi_c = \mu \) and \( \varphi_p(X) = 0 \), i.e., \( \varphi \) is \( \sigma \)-additive and \( \varphi = \mu \). As \( \varphi \) was an arbitrary weak* accumulation point, all the weak* accumulation points are identical to \( \mu \), i.e., \( \mu_n \to \mu \) for the weak*
topology in $\mathcal{M}(X)'$ and, in particular, $\mu_n \xrightarrow{\text{setwise}} \mu$.

(ii) As $\mu_n \leq \nu$ for all $n$, the sequence $\{\mu_n\}$ is in a weakly sequentially compact set of $M(X)$. Indeed, for all $n$, $\mu_n$ are norm-bounded and the $\sigma$-additivity of $\mu_n$ is uniform in $n$ (if $A_k = 0$, $\nu(A_k) = 0$ so that $\mu_n(A_k) \leq \nu(A_k) = 0$ uniformly). Therefore, from Theorem 2, p. 306 in [3], $\{\mu_n\}$ forms a weakly sequentially compact set in $M(X)$. Hence, there is a subsequence $\{\mu_{n_k}\}$ that converges weakly to some $\varphi \in M(X)$, and, in particular,

$$\int f \, d\mu_{n_k} \xrightarrow{\text{weakly}} \int f \, d\varphi \quad \forall f \in \mathcal{M}(X),$$

so that, from the $F$-convergence of $\mu_n$ to $\mu$,

$$\int f \, d\mu = \int f \, d\varphi \quad \forall f \in F.$$  

As both $\varphi$ and $\mu$ are in $M(X)$ and $F$ separates points in $M(X)$, (4.3) implies $\mu = \varphi$. As $\varphi$ was an arbitrary weak-limit point of $\{\mu_n\}$ in $M(X)$, we also conclude that all the weak-limit points are all equal to $\mu$. In other words, $\mu_n \xrightarrow{\text{setwise}} \mu$.

(iii) From (i) we conclude that $O\liminf_n\mu_n = \mu$ and with similar arguments, $O\limsup_n\mu_n = \mu$, i.e., the sequence $\{\mu_n\}$ in $M(X)$ has an $O$-limit $\mu$, or equivalently (see e.g., [5]), there exists $\{w_n\}$ in $M(X)$ such that

$$|\mu_n - \mu| \leq w_n \text{ with } w_n \downarrow 0.$$ 

Clearly, this implies convergence in total variation since

$$|\mu_n - \mu|(X) \leq w_n(X) \text{ with } w_n(X) \downarrow 0. \square$$

Lemma 4.1 applies to the following situations

- $X$ is a locally Hausdorff space and $\mu_n \xrightarrow{\text{vaguely}} \mu \in M(X)$, i.e.,

$$\int f \, d\mu_n \xrightarrow{\text{vaguely}} \int f \, d\mu \quad \forall f \in C_0(X) = : F$$

or if

$$\int f \, d\mu_n \xrightarrow{\text{weakly}} \int f \, d\mu \quad \forall f \in C_c(X) = : F$$

- $X$ is a topological space and $\mu_n \xrightarrow{\text{weakly}} \mu \in M(X)$, i.e.,

$$\int f \, d\mu_n \xrightarrow{\text{weakly}} \int f \, d\mu \quad \forall f \in C(X) = : F.$$ 

As a consequence of Lemma 4.1 we also get:

**Corollary 4.2**: Let $X$ be a locally compact Hausdorff space and $\lambda$ a $\sigma$-finite measure on $(X, \mathcal{B})$. Consider a sequence of probability densities $\{f_n\} \in L_1(\lambda)$ with almost everywhere limit $f \in L_1(\lambda)$. Let

$$\mu(B) = \int_B f \, d\lambda, \quad \mu_n(B) = \int_B f_n \, d\lambda \quad B \in \mathcal{B}, \ n = 1, \ldots.$$  

Assume that $\mu_n \xrightarrow{\text{vaguely}} \mu$. If $\mu_n \leq \nu$ for some $\nu \in M(X)$ then:
\[ \int |f_n - f| \, d\lambda \to 0 \text{ as } n \to \infty \]  
(4.8)

and \( \mu_n^{TV} \). In addition, if \( \lambda \) is finite, the family \( \{f_n\} \) is uniformly integrable.

**Proof:** Since the \( \mu_n \)'s are order-bounded, from Lemma 4.1(ii) with \( F = C_0(X) \), we conclude that \( \mu_n \to \mu \) \( \text{setwise} \). In particular, this implies that

\[ \int f_n \, d\lambda \to \int f \, d\lambda \text{ as } n \to \infty, \]

and by Scheffe's Theorem,

\[ \int |f_n - f| \, d\lambda \to 0 \text{ as } n \to \infty, \]

which yields (4.8). That \( \mu_n^{TV} \to \mu \) follows from the \( L_1 \) convergence of \( f_n \) to \( f \), i.e., (4.8). If \( \lambda \) is finite, the uniform integrability of the family \( \{f_n\} \) follows from [2], p. 155.

Note that if instead of \( \mu_n \leq \nu \), we had \( |f_n| \leq g \in L_1(\lambda) \) then by the Dominated Convergence Theorem, \( \int f_n \, d\lambda \to \int f \, d\lambda \) and (4.8) would follow from Scheffe’s Theorem. However, note that the condition \( \mu_n \leq \nu \) does not require \( \nu \) to have a density.

**References**