JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS **151**, 1–16 (1990)

New Sequential Compactness Results for Spaces of Scalarly Integrable Functions

ERIK J. BALDER

Mathematical Institute, University of Utrecht, Netherlands

Submitted by David Aldous

Received February 25, 1988

A new general relative sequential compactness criterion for scalarly integrable functions is presented. The novel feature of this result is the fact that it deals with a strong form of pointwise Cesaro-convergence (almost everywhere), which is stronger than the usual types of weak convergence. Applications of the main result include extensions of a classical weak compactness criterion for abstract L_1-spaces and a recent relative sequential compactness criterion of Prohorov type for transition probabilites. The main result can also be seen as an abstract version of a famous theorem of Komlós, which is also essential for its proof. © 1990 Academic Press, Inc.

1. INTRODUCTION

Let $(T, \mathscr{T}, \mu)$ be an arbitrary measure space, and let E be some topological vector space. A sequence $\{f_m\}$ of functions $f_m: T \to E$ is defined to *K-converge almost everywhere* on T to a function $f_*: T \to E$ if for every subsequence $\{f_{m_i}\}$ of $\{f_m\}$ there exists a null set $N \in \mathscr{T}$ (i.e., $\mu(N) = 0$) such that

$$\frac{1}{n} \sum_{i=1}^{n} f_{m_i}(t) \to f_*(t) \qquad \text{for every } t \in T \backslash N. \tag{1.1}$$

In case the topology (say τ) on E must be specified, we shall also speak of τ-K-convergence a.e., etc. Because of its subsequence character K-convergence a.e. is obviously a topological notion.

A famous result of Komlós [14], derived by means of a rather delicate truncation argument, involving martingale convergence, says the following:

THEOREM (Komlós). *Suppose that $\{\phi_k\}$ is a sequence of integrable functions $\phi_k: T \to \mathbb{R}$ such that*

$$\sup_k \int_T |\phi_k| \, d\mu < +\infty.$$

0022-247X/90 $3.00

Then $\{\phi_k\}$ *has a subsequence which K-converges a.e. to an integrable function* $\phi_*: T \to \mathbb{R}$.

In [1d] this result was extended in two ways to functions $\{\phi_k\}$ which take their values in a Banach space E. One of these extensions requires E to be reflexive, and in the other one the set $\{\phi_k(t)\} \subset E$ must be relatively weakly compact for almost every $t \in T$ (in both cases one deals with $\sigma(E, E')$-K-convergence a.e.). Different extensions of Komlós' result to infinite dimensions were given in [10, 16]; there one studies norm-K-convergence a.e. Obviously, conditions of a much stronger nature must then be imposed, since E must at least have the Banach–Saks property. In this connection mention should be made of a remarkable result by J. Bourgain [2, 9], which says that the validity of Komlós' result for norm-K-convergence is equivalent to L^2_E having the Banach–Saks property and also equivalent to L^1_E having the weak Banach–Saks property, but *not* equivalent to E having the Banach–Saks property (see also [15] for the latter observation).

The proofs in [1d] depend on a diagonal method, whereby the above scalar version of Komlós' theorem is applied repeatedly, plus the introduction of a countable collection of linear continuous functionals which separates the points of E. The purpose of the present paper is to show how these ideas can be used in a much more general setup, and then lead to a new, very general criterion for relative sequential compactness for K-convergence a.e. in spaces of scalarly integrable functions (Theorem 2.1). As an immediate consequence, a concomitant criterion for relative sequential "weak" compactness in those spaces is obtained in Corollary 2.2. The latter criterion is also new, but much more standard in nature.

The power of these criteria is demonstrated by four applications. The first application, Theorem 3.2, gives extensions of Komlós' theorem which go slightly further than the above-mentioned extensions in [1d]. The second application, Theorem 4.1, furnishes a new relative weak compactness criterion for L^1-spaces. Among other things, this criterion generalizes a well-known result due to Diestel [8a] (Corollary 4.2). The third application, Theorem 5.1, extends a recent criterion for relative narrow sequential compactness in a space of transition probabilities [1a, 1c]. Finally, Theorem 6.1, a fourth application of Theorem 2.1 and Corollary 2.2, gives a criterion for relative weak sequential compactness in a generalized Köthe space, which forms the sequential counterpart to a result of Castaing and Valadier [4, V.13]. For more applications along the lines of this paper, found after this paper was written, we refer to [1e].

2. Main Result

Let $(T, \mathscr{T}, \mu)$ be an arbitrary measure space, and E a convex subset of a Hausdorff topological vector space. Any function from $T \times E$ into $(-\infty, +\infty]$ is be called *integrand.* An integrand $g: T \times E \to (-\infty, +\infty]$ is said to be respectively convex, affine, (sequentially) lower semicontinuous, (sequentially) continuous, or (sequentially) inf-compact if for every $t \in T$ the function $g(t, \cdot)$ on E has the corresponding property (recall that $g(t, \cdot)$ is (sequentially) *inf-compact* on E if for every $\beta \in \mathbb{R}$ the set of all $x \in E$ such that $g(t, x) \leqslant \beta$, is (sequentially) compact). Thus, these adjectives refer only to the behavior of the integrand in its *second* variable.

Let $h: T \times E \to [0, +\infty]$ be given nonnegative convex sequentially inf-compact integrand and let $\mathscr{A}$ be a set of affine sequentially continuous integrands $a: T \times E \to \mathbb{R}$. A function $f: T \to E$ is defined to be *$\mathscr{A}$-scalarly measurable* if

$$t \to a(t, f(t)) \text{ is measurable on } T \text{ for every } a \in \mathscr{A}.$$

We require that all integrands $a \in \mathscr{A}$ have the following *property* (**B**) with respect to h: there exist $C > 0$ and an integrable function $\phi: T \to \mathbb{R}$ (these may depend on a) such that

$$\text{(B)} \quad |a(t, x)| \leqslant Ch(t, x) + \phi(t) \text{ on } T \times E.$$

A subset D of E is defined to be *countably separated* by $\mathscr{A}$ if there exists a countable subset $\{a_j\}$ of $\mathscr{A}$ such that for every $t \in T$ and every pair of points $x, y \in D$

$$a_j(t, x) = a_j(t, y) \text{ for } j = 1, 2, \ldots \text{ implies that } x = y.$$

We recall the definition of outer integration: the *outer integral* over $(T, \mathscr{T}, \mu)$ of a (possibly nonmeasurable) function $\psi: T \to (-\infty, +\infty]$ is defined by

$$\int_T^* \psi \, d\mu := \inf \left\{ \int_T \phi \, d\mu : \phi: T \to \mathbb{R} \text{ integrable}, \psi \leqslant \phi \right\}$$

(the infimum over the empty set is $+\infty$ by convention). For more details we refer to, e.g., [1a, Appendix].

Our main result, a criterion for relative sequential compactness for K-convergence a.e., can now be stated (recall below that cl co is used to denote closed convex hulls).

Theorem 2.1. *Suppose that all elements of the collection $\mathscr{A}$ of affine sequentially continuous integrands have property* (**B**) *with respect to the* *convex sequentially* inf-*compact integrand h.* Let $\{f_k\}$ *be a sequence of* $\mathscr{A}$-*scalarly measurable functions* $f_k: T \to E$ *such that*

$$\sup_k \overset{*}{\int_T} h(t, f_k(t))\, \mu(dt) < +\infty. \tag{2.1}$$

Suppose also that there exists a null set $N \in \mathscr{T}$ *such that*

$$\bigcap_{p=1}^{\infty} \operatorname{cl} \operatorname{co} \bigcup_{k \geq p} f_k(T \backslash N) \text{ is countably separated by } \mathscr{A}. \tag{2.2}$$

Then $\{f_k\}$ *has a subsequence* $\{f_m\}$ *which K-converges a.e. to an* $\mathscr{A}$-*scalarly measurable function* $f_*: T \to E$ *satisfying*

$$\overset{*}{\int_T} h(t, f_*(t))\, \mu(dt) \leq \sup_k \overset{*}{\int_T} h(t, f_k(t))\, \mu(dt) < +\infty. \tag{2.3}$$

Proof. It follows directly from the definition of outer integration that for every $k \in \mathbb{N}$ there exists an integrable function from T into $\mathbb{R}$, for convenience to be denoted—perhaps somewhat oddly—as $t \to a_0(t, f_k(t))$, such that

$$\overset{*}{\int_T} h(t, f_k(t))\, \mu(dt) = \int_T a_0(t, f_k(t))\, \mu(dt),$$

$$a_0(t, f_k(t)) \geq h(t, f_k(t)) \geq 0 \qquad \text{on } T. \tag{2.4}$$

Let $\{a_j\} \subset \mathscr{A}$ be the countable subset which figures in the definition of (2.2). By property (B) and (2.1)

$$\sup_k \int_T |a_j(t, f_k(t))|\, \mu(dt) < +\infty \qquad \text{for} \quad j = 0, 1, 2 \ldots .$$

We follow an obvious diagonal method, by successively applying Komlós' theorem to suitably chosen subsequences of $\{a_j(\cdot, f_k(\cdot)) : k \in \mathbb{N}\}$ for $j = 0, 1, 2, \ldots$. This yields the existence of a sequence $\{\phi_j\}$ of integrable functions $\phi_j : T \to \mathbb{R}$, and a subsequence $\{f_m\}$ of $\{f_k\}$ such that for every subsequence $\{f_{m_i}\}$ of $\{f_m\}$

$$\lim_n \frac{1}{n} \sum_{i=1}^{n} a_j(t, f_{m_i}(t)) = \phi_j(t) \qquad \text{for} \quad j = 0, 1, 2, \ldots \text{ a.e. in } T. \tag{2.5}$$

Denote by M the union of N and the exceptional set for (2.5) if $\{f_m\}$ itself is taken to be the subsequence of interest. Define $s_n := (1/n) \sum_{m=1}^{n} f_m$. Fix

$t\in T\backslash M$; then it follows from convexity of the function $h(t,\cdot)$ and (2.4)–(2.5) that

$$\limsup_n h(t, s_n(t)) \leqslant \limsup_n \frac{1}{n} \sum_{m=1}^{n} a_0(t, f_m(t)) \leqslant \phi_0(t) < +\infty. \tag{2.6}$$

By sequential inf-compactness of $h(t,\cdot)$ this implies that there exist $y_t \in E$ and a (t-dependent) subsequence $\{s_{n_p}(t)\}$ of $\{s_n(t)\}$ such that $s_{n_p}(t) \to y_t$. Note that this implies that $y_t \in \bigcap_p \operatorname{cl} \operatorname{co} \bigcup_{k \geqslant p} f_k(T\backslash N)$. By (2.5) and affinity and sequential continuity of the function $a_j(t,\cdot)$ on E it follows that

$$a_j(t, y_t) = \phi_j(t) \qquad \text{for} \quad j = 1, 2, \dots . \tag{2.7}$$

But at the same time this argument shows that *every* limit point y_t of $\{s_n(t)\}$ must satisfy the above identity. Suppose we had $s_n(t) \nrightarrow y_t$. Then there would be an open neighborhood V of y_t and a subsequence $\{s_{n_q}(t)\}$ of $\{s_n(t)\}$ such that $s_{n_q}(t) \notin V$ for $q = 1, 2, \dots$. But by (2.6) and sequential inf-compactness of $h(t,\cdot)$ on E the subsequence $\{s_{n_q}(t)\}$ would have a further subsequence converging to a point $z_t \in (E\backslash V) \cap [\bigcap_p \operatorname{cl} \operatorname{co} \bigcup_{k \geqslant p} f_k(T\backslash N)]$. By (2.7) we would have

$$a_j(t, z_t) = \phi_j(t) = a_j(t, y_t) \qquad \text{for} \quad j = 1, 2, \dots,$$

which, by condition (2.2), would imply that $y_t = z_t \notin V$. This contradiction shows that $s_n(t) \to y_t$. Now we define $f_*: T \to E$ by $f_*(t) := y_t$ for $t \in T\backslash M$. On M we set f_* equal to some arbitrary but fixed constant in E. This gives that f_* is $\mathscr{A}$-scalarly measurable, since for every $a \in \mathscr{A}$

$$a(t, f_*(t)) = \lim_n \frac{1}{n} \sum_{m=1}^{n} a(t, f_m(t)) \qquad \text{on} \quad T\backslash M.$$

Note that by convexity and sequential lower semicontinuity of $h(t,\cdot)$ for every $t \in T$, it follows from (2.6) that $h(t, f_*(t)) \leqslant \phi_0(t)$ a.e. By (2.5) this means that f_* satisfies (2.3).

Now let $\{f_{m_i}\}$ be an arbitrary subsequence of $\{f_m\}$. Then the argument above can be repeated for $s'_n := (1/n) \sum_{i=1}^{n} f_{m_i}$. This gives that there is a null set M' such that for every $t \in T\backslash M'$ there exists $y'_t \in \operatorname{cl} \operatorname{co} \bigcup_k f_k(T\backslash N)$ such that $s'_n(t) \to y'_t$ and

$$a_j(t, y'_t) = \phi_j(t) \qquad \text{for} \quad j = 1, 2, \dots .$$

By (2.2), (2.7) it follows that $y'_t = f_*(t)$ for every $t \in T\backslash(M \cup M')$. Thus, we have proved that $\{f_m\}$ K-converges a.e. to f_*. ∎

An immediate consequence of Theorem 2.1 is as follows. We shall say that an integrand $a \in \mathscr{A}$ has the *growth property* (G) with respect to h if for every $\varepsilon > 0$ there exists an integrable function $\phi_\varepsilon\colon T \to \mathbb{R}$ such that

$$\text{(G)} \quad |a(t, x)| \leqslant \varepsilon h(t, x) + \phi_\varepsilon(t) \qquad \text{on} \quad T \times E.$$

Slightly more generally, we define an integrand $g\colon T \times E \to (-\infty, +\infty]$ to have the *growth property* (G′) with respect to h if for every $\varepsilon > 0$ there exists an integrable function $\phi_\varepsilon\colon T \to \mathbb{R}$ such that

$$\text{(G}'\text{)} \quad \max(0, -g(t, x)) \leqslant \varepsilon h(t, x) + \phi_\varepsilon(t) \qquad \text{on} \quad T \times E,$$

COROLLARY 2.2. *The K-convergent a.e. subsequence* $\{f_m\}$ *of* $\{f_k\}$ *in Theorem* 2.1 *is such that, for every* $a \in \mathscr{A}$ *having growth property* (G),

$$\lim_m \int_T a(t, f_m(t))\, \mu(dt) = \int_T a(t, f_*(t))\, \mu(dt)$$

and, more generally,

$$\liminf_m \int_T^* g(t, f_m(t))\, \mu(dt) \geqslant \int_T^* g(t, f_*(t))\, \mu(dt) \tag{2.8}$$

for each convex sequentially lower semicontinuous integrand $g\colon T \times E \to (-\infty, +\infty]$ *having growth property* (G′).

Proof. Let α be the left side of (2.8); if α equals $+\infty$ we are done. If not, there is a subsequence $\{f_{m_i}\}$ of $\{f_m\}$ such that $\alpha = \lim_i \alpha_i$, where every $\alpha_i := \int_T^* g(t, f_{m_i}(t))\, \mu(dt)$ is finite. By convexity and sequential lower semicontinuity of $g(t, \cdot)$ it follows that

$$\liminf_n \frac{1}{n} \sum_{i=1}^n g(t, f_{m_i}(t)) \geqslant g(t, f_*(t)) \qquad \text{a.e. in } T. \tag{2.9}$$

For every $i \in \mathbb{N}$ there exists an integrable function $\phi_i\colon T \to \mathbb{R}$ such that $\phi_i(t) \geqslant g(t, f_{m_i}(t))$ on T, with $\alpha_i = \int_T \phi_i\, d\mu$. By the growth property (G) we get, for every $\varepsilon > 0$,

$$\phi_i(t) + \varepsilon h_i(t) \geqslant -\phi_\varepsilon(t) \qquad \text{for all} \quad t \in T,\ i \in \mathbb{N},$$

where we denote $h_i(t) := a_0(t, f_{m_i}(t))$ (see the proof of Theorem 2.1). Note that $h_i(t) \geqslant h(t, f_{m_i}(t)) \geqslant 0$. Let us set $\sigma := \sup_i \int_T h_i\, d\mu$ (σ is finite by (2.1)). We can apply Fatou's lemma, which gives

$$\alpha + \varepsilon\sigma = \lim_n \frac{1}{n} \sum_{i=1}^n \alpha_i + \varepsilon\sigma \geqslant \liminf_n \int_T \sum_{i=1}^n (\phi_i + \varepsilon h_i)\, d\mu$$

$$\geqslant \int_T \liminf_n \frac{1}{n} \sum_{i=1}^n (\phi_i + \varepsilon h_i)\, d\mu.$$

Since

$$\frac{1}{n}\sum_{i=1}^{n}(\phi_i(t)+\varepsilon h_i(t)) \geqslant \frac{1}{n}\sum_{i=1}^{n}\phi_i(t) \geqslant \frac{1}{n}\sum_{i=1}^{n} g(t, f_{m_i}(t)),$$

the desired inequality (2.8) follows from the above by (2.9), letting ε go to zero. For every $a \in \mathscr{A}$ satisfying (G), both a and $-a$ satisfy (G′). The inequality (2.8) therefore holds for both $g = a$ and $g = -a$, with outer integration replaced by ordinary integration. ▮

3. Applications: Extensions of Komlós' Theorem

Here we present two very direct application of Theorem 2.1, which lead to two different extensions of Komlós' theorem (cf. [1d]). In this section E is a Banach space, the norm of which we denote by $\|\cdot\|$. As before, $(T, \mathscr{T}, \mu)$ stands for an arbitrary measure space. Recall that a function $f\colon T \to E$ is said to be *strongly measurable* if it is the pointwise limit a.e. in the norm $\|\cdot\|$ of a sequence of finite-valued measurable functions; moreover, f is said to be *Bochner-integrable* if $\int_T \|f\|\, d\mu < +\infty$ [8b, 17]. The following technical lemma is important in connection with the countable separability condition (2.2) and the sequential nature of our inf-compactness condition for h.

Lemma 3.1. *Let $\{f_k\}$ be a sequence of strongly measurable functions from T into the Banach space E. Suppose that E is equipped with a Hausdorff locally convex topology τ not finer than the norm topology. Then there exists a null set $N \in \mathscr{T}$ such that the following hold for $D := \text{cl co} \bigcup_k f_k(T\backslash N)$:*

(a) *D is $\|\cdot\|$-separable.*

(b) *D is a Suslin space for the relative topology τ.*

(c) *There exists a countable subset $\{x'_j\}$ of the topological dual of (E, τ) which separates the points of D (that is $\langle x, x'_j\rangle = \langle y, x'_j\rangle$ for $j = 1, 2, \ldots$ implies that $x = y$ for every pair $x, y \in D$).*

Proof. From the definition of strong measurability it is evident that there exists a nul set $N \in \mathscr{T}$ such that the set $\bigcup_k f_k(T\backslash N)$ is separable for $\|\cdot\|$. Hence, $D := \text{cl co} \bigcup_k f_k(T\backslash N)$ is also $\|\cdot\|$ – separable. Now $(D, \|\cdot\|)$ is obviously Polish (i.e., separable and complete for $\|\cdot\|$). Therefore, the space (D, τ) is Suslin [7, III.67], for the injection from $(D, \|\cdot\|)$ into (D, τ) is continuous. Since (E, τ) is Hausdorff locally convex, the elements of its topological dual separate the points of E, and in particular the points of D. It now remains to apply a well-known result for Suslin spaces due to L. Schwartz [4, III.31]. ▮

THEOREM 3.2. *Suppose that E is the topological dual of a Banach space F, and that $\|\cdot\|$ is the corresponding dual norm. Let $\{f_k\}$ be a sequence of Bochner-integrable functions $f_k\colon T \to E$ such that*

$$\sup_k \int_T \|f_k\| \, d\mu < +\infty.$$

Then $\{f_k\}$ has a subsequence which $\sigma(E, F)$-K-converges a.e. to a Bochner-integrable function $f_*\colon T \to E$.

Proof. We apply Theorem 2.1, making the following substitutions: E is equipped with the topology $\sigma(E, F)$. By the Alaoglu–Bourbaki theorem, the function $\|\cdot\|$ is inf-compact on E (but not necessarily sequentially inf-compact). Define the integrand h has

$$h(t, x) := \begin{cases} \|x\| & \text{if} \quad x \in D, \\ +\infty & \text{otherwise.} \end{cases}$$

Here D is the $\sigma(E, F)$-closed set $\text{cl co} \bigcup_k f_k(T \backslash N)$ corresponding to $\{f_k\}$ by Lemma 3.1. Note that for every $t \in T$, $\beta \in \mathbb{R}$ the set of all $x \in E$ such that $h(t, x) \leqslant \beta$ is a $\sigma(E, F)$-compact set contained in D, and therefore metrizable [7, III.66], in view of Lemma 3.1b. Thus, h is a convex and sequentially inf-compact integrand. Note that (2.1) holds trivially. By Lemma 3.1c there exists a countable subset $\{y_j\}$ of F separating the points of D. We define $\mathscr{A}$ to consist of the integrands $(t, x) \to a(t, x) := \langle x, t \rangle$, $y \in F$ (note that they have property (B)). Then condition (2.2) holds as well. The result follows from Theorem 2.1, because the limit function $f_*\colon T \to E$ is such that $t \to \langle f_*(t), y \rangle$ is measurable on T for every $y \in F$, and satisfies $f_*(t) \in D$ a.e. in T. Hence, by Pettis' measurability theorem [8b, IV], f_* is strongly measurable. In fact, from (2.3) it follows that f_* is Bochner-integrable. ∎

Let us say that a subset K of E is *locally* sequentially τ-compact for some topology τ on E if the intersection of K with each norm-closed ball in E is sequentially τ-compact.

THEOREM 3.3. *Let τ be a Hausdorff locally convex topology on E not finer than the norm topology. Let $\{f_k\}$ be a sequence of Bochner-integrable functions $f_k\colon T \to E$ such that*

$$\sup_k \int_T \|f_k\| \, d\mu < +\infty,$$

$$\text{cl co}\{f_k(t)\} \textit{ is locally (sequentially) } \tau\textit{-compact a.e. in } T.$$

Then $\{f_k\}$ has a subsequence which τ-K-converges a.e. to a Bochner-integrable function $f_*\colon T \to E$.

Proof. We imitate the previous proof, making the following alteration: E is now equipped with τ, and the integrand h is defined as

$$h(t, x) := \begin{cases} \|x\| & \text{if } x \in D \cap \operatorname{cl} \operatorname{co}\{f_k(t)\}, \\ +\infty & \text{otherwise.} \end{cases}$$

Here $D := \operatorname{cl} \operatorname{co} \bigcup_k f_k(T \backslash N)$ again corresponds to $\{f_k\}$ as in Lemma 3.1. Note that for every $t \in T$, $\beta \in \mathbb{R}$ the set of all $x \in E$ such that $h(t, x) \leqslant \beta$ consists of the intersection of the ball $\{x \in E: \|x\| \leqslant \beta\}$ with the locally (sequentially) compact set $\operatorname{cl} \operatorname{co}\{f_k(t)\}$; hence $h(t, \cdot)$ is inf-compact and sequentially inf-compact (in a Suslin space ordinary compactness implies sequential compactness [7, III.66]). Finally, Lemma 3.1c yields the same collection $\mathscr{A}$ as used previously. ∎

4. Applications: Criteria for Weak Compactness in L_1-Spaces

In this section $(T, \mathscr{T}, \mu)$ is σ-finite, and $(E, \|\cdot\|)$ is a Banach space; the topological dual of $(E, \|\cdot\|)$ is denoted by E', and the dual norm on E' by $\|\cdot\|'$. Let $\mathscr{L}^1_E$ be the space of all Bochner-integrable functions $f: T \to E$, and define the usual seminorm

$$\|f\|_1 := \int_T \|f(t)\| \, \mu(dt).$$

It is well known that the topological dual of $(\mathscr{L}^1_E, \|\cdot\|_1)$ can be identified with the space $\mathscr{M}^{\infty}_{E'}[E]$ consisting of all E-scalarly measurable bounded functions from T into $(E', \|\cdot\|')$ [11, VII]. We call the topology $\sigma(\mathscr{L}^1_E, \mathscr{M}^{\infty}_{E'}[E])$ the *weak* topology on $\mathscr{L}^1_E$. We can already observe that by the Eberlein–Šmulian theorem (relative) weak sequential compactness and (relative) weak compactness in $\mathscr{L}^1_E$ coincide (the fact that we work in a prequotient setting does not affect this; actually, everything said below can easily be transcribed to the usual quotient setting).

In this setup we present a new criterion for (relative) sequential K-compactness, which implies a criterion for (relative) weak compactness in $\mathscr{L}^1_E$.

Theorem 4.1. *Suppose that* $h: T \times E \to [0, +\infty]$ *is a* $\sigma(E, E')$-inf-*compact integrand with the following superlinear growth property: for every* $\varepsilon > 0$ *there exists an integrable function* ψ_ε: $T \to [0, +\infty)$ *such that*

$$\frac{h(t, x)}{\|x\|} \geqslant \frac{1}{\varepsilon} \qquad \text{for all } t \in T \text{ and } x \in E \backslash \{0\} \text{ with } \|x\| \geqslant \psi_\varepsilon(t). \tag{4.1}$$

Then every subset $\mathscr{L} \subset \mathscr{L}^1_E$ for which

$$\sup_{f \in \mathscr{L}} \int_T^* h(t, f(t))\, \mu(dt) < +\infty, \tag{4.2}$$

is relatively sequentially compact for $\sigma(E, E')$-K-convergence a.e. In particular, $\mathscr{L}$ is relatively weakly compact and relatively weakly sequentially compact.

The following corollary of Theorem 4.1 generalizes a result frequently encountered under the name "Diestel's theorem" [8a, 3, 13].

COROLLARY 4.2. *Suppose that $\Gamma: T \to 2^E$ is a multifunction with convex $\sigma(E, E')$-compact values such that for some integrable function $\psi: T \to [0, +\infty)$*

$$\|x\| \leqslant \psi(t) \qquad \text{for all} \quad x \in \Gamma(t).$$

Then the set $\mathscr{S}$ of all $f \in \mathscr{L}^1_E$ such that $f(t) \in \Gamma(t)$ a.e. in T is sequentially compact for $\sigma(E, E')$-K-convergence a.e.; in particular, $\mathscr{S}$ is a weakly compact subset of $\mathscr{L}^1_E$.

Proof. Apply Theorem 4.1 with the following integrand h:

$$h(t, x) := \begin{cases} \|x\| & \text{if} \quad x \in \Gamma(t) \\ +\infty & \text{otherwise.} \end{cases}$$

Then sequential inf-compactness of $h(t, \cdot)$ follows directly, as in the proof of Theorem 3.3. Also, (4.1) and (4.2) hold trivially . Note elementarily that $\mathscr{S}$ is sequentially closed for K-convergence a.e. ∎

Proof of Theorem 4.1. As noted before, the Eberlein–Šmulian theorem ensures that it is enough to prove relative weak sequential compactness of $\mathscr{L}$. So let $\{f_k\}$ be an arbitrary sequence in $\mathscr{L}$. We apply Theorem 2.1 and Corollary 2.2 with the following substitutions. Let h be as given, and notice that, by the Eberlein–Šmulian theorem, $\sigma(E, E')$-compactness and sequential $\sigma(E, E')$-compactness coincide. Let $\mathscr{A}$ be the collection of all integrands $a^g: T \times E \to \mathbb{R}$, defined by

$$a^g(t, x) := \langle x, g(t) \rangle, \qquad g \in \mathscr{M}^\infty_{E'}[E].$$

Then by (4.1) for every $\varepsilon > 0$

$$|a^g(t, x)| \leqslant \|x\| \; \|g\|_\infty \leqslant \|g\|_\infty \, (\varepsilon h(t, x) + \psi_\varepsilon(t)),$$

where $\|\cdot\|_\infty$ stands for the supremum norm on $\mathscr{M}^\infty_{E'}[E]$. This shows that all elements of $\mathscr{A}$ have property (G). Note that $\mathscr{A}$-scalar measurability of the elements of $\mathscr{L}^1_E$ is guaranteed by their strong measurability. Since integrands $(t, x) \to \langle x, x' \rangle$, $x' \in E'$, also belong to $\mathscr{A}$, it follows from Lemma 3.1c that condition (2.2) is fulfilled. Of course, (2.1) holds by (4.2). Thus, by Theorem 2.1 there exist an $\mathscr{A}$-scalarly measurable function $f_*: T \to E$ and a sequence $\{f_m\}$ of $\{f_k\}$ such that

$$\{f_m\}\ \sigma(E, E')\text{-}K\text{-converges a.e. to } f_*,$$

$$\int_T \|f_*(t)\|\ \mu(dt) \leqslant \int_T [h(t, f_*(t)) + \psi_1(t)]\ \mu(dt) < +\infty.$$

It follows immediately from the former statement that f_* is strongly measurable; (see the proof of Theorem 3.2). Together with the latter statement this implies that f_* is Bochner-integrable, i.e., belongs to $\mathscr{L}^1_E$. Also, by Corollary 2.2

$$\lim_m \int_T \langle f_m(t), g(t) \rangle\ \mu(dt) = \int_T \langle f_*(t), g(t) \rangle\ \mu(dt) \quad \text{for every} \quad g \in \mathscr{M}^\infty_{E'}[E].$$

Thus, the subsequence $\{f_m\}$ of $\{f_*\}$ converges weakly to f_* in $\mathscr{L}^1_E$. ∎

5. Applications: Prohorov's Theorem for Transition Probabilities

In this section $(T, \mathscr{T}, \mu)$ is σ-finite. Let S be a completely regular Suslin space, and denote the set of all probability measures on $(S, \mathscr{B}(S))$ by $\mathscr{P}(S)$. We equip $\mathscr{P}(S)$ with the *narrow* topology $\sigma(\mathscr{P}(S), \mathscr{C}_b(S))$, i.e., the initial topology with respect to the functionals $\nu \to \int_S c\, d\nu$, $c \in \mathscr{C}_b(S)$. Here $\mathscr{C}_b(S)$ indicates the set of all bounded continuous real-valued functions on S. Under the above conditions for S, the space $E := \mathscr{P}(S)$ is a Suslin space; see for instance [1b, Appendix]. Note already that this entails that ordinary narrow compactness on $\mathscr{P}(S)$ implies sequential narrow compactness [7, III.66].

A *transition probability* from T into S is defined to be a function $\delta: T \to \mathscr{P}(S)$ such that for every $B \in \mathscr{B}(S)$ the function $t \to \delta(t)(B)$ is $\mathscr{T}$-measurable. It is not hard to verify, following [4, p. 103], that in this framework transition probabilities are precisely the functions from T into $\mathscr{P}(S)$ that are with respect to $\mathscr{T}$ and the narrow-Borel σ-algebra on $\mathscr{P}(S)$. The set of all transition probabilities from T into S is denoted as $\mathscr{R}(T; S)$. The narrow topology on $\mathscr{P}(S)$ has an obvious analogue on $\mathscr{R}(T; S)$: Let $\mathscr{G}_C(T; S)$ be the set of all *Carathéodory integrands* on $T \times S$, i.e., the set of all $\mathscr{T} \times \mathscr{B}(S)$-measurable continuous integrands $g: T \times S \to \mathbb{R}$ such that for some integrable function $\phi: T \to [0, +\infty)$

$$|g(t, s)| \leqslant \phi(t) \qquad \text{for all} \quad t \in T,\ s \in S. \tag{5.1}$$

The *narrow topology* on $\mathscr{R}(T; S)$ is defined as the initial topology on $\mathscr{R}(T; S)$ with respect to the functionals

$$\delta \to \int_T \left[\int_S g(t, s)\, \delta(t)(ds) \right] \mu(dt), \qquad g \in \mathscr{G}_C(T; S)$$

(by what was said above about $\mathscr{R}(T; S)$, these integrals are well-defined).

Following [1a, 1c], a subset $\mathscr{R}_0$ of $\mathscr{R}(T; S)$ is defined to be *tight* if there exists an inf-compact integrand $h' : T \times S \to [0, +\infty]$ such that

$$\sup_{\delta \in \mathscr{R}_0} \int_T^* \left[\int_S h'(t, s)\, \delta(t)(ds) \right] \mu(dt) < +\infty.$$

As shown in [12, Prop. 2.2], a remark in [1a, p. 573] can be extended as follows: the subset $\mathscr{R}_0$ of $\mathscr{R}(T; S)$ is tight if and only if for every $\varepsilon > 0$ there exists a multifunction $\Gamma_\varepsilon : T \to 2^S$ with compact values, such that

$$\sup_{\delta \in \mathscr{R}_0} \int_T^* \delta(t)(S \backslash \Gamma_\varepsilon(t))\, \mu(dt) \leqslant \varepsilon.$$

Thus, tightness extends the classical notion of tightness on $\mathscr{P}(S)$ [7]. We now derive a Prohorov-type criterion for relative narrow sequential compactness. The technical conditions imposed here on S are less restrictive than those of earlier such results, which require S to be metrizable Lusin [1a, Theorem I], or the countable union of metrizable Lusin spaces [1c, Theorem 2.1]. For a number of applications of such results we refer to [1a, 1c]. In the same setting as here, in [1f] a Prohorov-type result was obtained in terms of nonsequential narrow compactness. Together with the sequential compactness result presented below, this resolves an open question of the author (cf., e.g., [1b, p. 468]) and C. Castaing [5, p. 517].

THEOREM 5.1. *Every tight subset of $\mathscr{R}(T; S)$ is relatively sequentially compact for $\sigma(\mathscr{P}(S), \mathscr{C}_b(S))$-K-convergence a.e. In particular, it is relatively narrowly sequentially compact.*

Proof. Let $\{\delta_k\}$ be an arbitrary sequence in $\mathscr{R}(T; S)$. We apply Theorem 2.1, making the following substitutions: E is the set $\mathscr{P}(S)$, equipped with the topology $\sigma(\mathscr{P}(S), \mathscr{C}_b(S))$. To h' as in the definition of tightness we let correspond the function $h : T \times \mathscr{P}(S) \to [0, +\infty]$, given by

$$h(t, x) := \int_S h'(t, s)\, x(ds);$$

then inf-compactness of $h'(t, \cdot)$ on S implies inf-compactness and sequential inf-compactness of $h(t, \cdot)$ on $\mathscr{P}(S)$. Of course, tightness now means that (2.1) is fulfilled. Also, convexity (indeed affinity) of $h(t, \cdot)$ holds automatically. We define $\mathscr{A}$ to be the collection of continuous integrands $a^g: T \times \mathscr{P}(S) \to \mathbb{R}$ given by

$$a^g(t, x) := \int_S g(t, s)\, x(ds), \qquad g \in \mathscr{G}_C(T; S).$$

For every $g \in \mathscr{G}_C(T; S)$ there exists an integrable function $\phi: T \to \mathbb{R}$ such that (5.1) is valid; then also $|a^g(t, x)| \leqslant \phi(t)$ for all $t \in T$, $x \in \mathscr{P}(S)$, so property (G) is immediate. Further, by what was said above about $\mathscr{R}(T; S)$, it is easy to see that for every $k \in \mathbb{N}$ and $g \in \mathscr{G}_C(T; S)$

$$t \to a^g(t, \delta_k(t)) = \int_S g(t, s)\, \delta_k(t)(ds) \text{ is measurable on } T.$$

Thus, all elements of the sequence $\{\delta_k\}$ are $\mathscr{A}$-scalarly measurable. As remarked above and proven in, e.g., [1b, Appendix], $\mathscr{P}(S)$ is Suslin; by L. Schwartz's result [4, III.31] there exists a countable subset $\{c_j\}$ of $\mathscr{C}_b(S)$ which separates the points of $\mathscr{P}(S)$. Let $\psi: T \to (0, +\infty)$ be a strictly positive integrable function; then the integrands $(t, x) \to \psi(t) \int_S c_j(s)\, x(ds)$ on $T \times \mathscr{P}(S)$, $j \in \mathbb{N}$, belong to $\mathscr{A}$. This shows that condition (2.2) also holds here (take for N the empty set). Thus, we may apply Theorem 2.1: there exist a transition probability $\delta_* \in \mathscr{R}(T; S)$ and a subsequence $\{\delta_m\}$ of $\{\delta_k\}$ such that

$$\{\delta_m\}\ \sigma(\mathscr{P}(S), \mathscr{C}_b(S))\text{-}K\text{-converges a.e. to } \delta_*,$$

and in particular, by Corollary 2.2

$$\{\delta_m\} \text{ narrowly converges to } \delta_*.$$

This proves the desired results. ∎

6. Applications: Compactness Criterion for Generalized Köthe Functions

In this section we derive from Theorem 2.1 and Corollary 2.2 the sequential analogue of a relative compactness criterion in a generalized Köthe function space, due to Castaing and Valadier [4, V.13]. We suppose in this section that the measure space $(T, \mathscr{T}, \mu)$ is σ-finite. Recall that a function

$\phi: T \to \mathbb{R}$ is said to be *locally integrable* on T if it is measurable and if for every $B \in \mathscr{T}$, $\mu(B) < +\infty$,

$$\int_B |\phi| \, d\mu < +\infty.$$

Let $\mathscr{L}^1_{\text{loc}}$ be the space of all such real-valued locally integrable functions on T. Let $\mathscr{G}$ be a given nonempty subset of $\mathscr{L}^1_{\text{loc}}$; define $\mathscr{L}$ to be the vector space of all $\phi \in \mathscr{L}^1_{\text{loc}}$ such that

$$\int_T |\phi\psi| \, d\mu < +\infty \qquad \text{for all} \quad \psi \in \mathscr{G}.$$

Also, define $\mathscr{L}^*$ to be the vector space of all $\psi \in \mathscr{L}^1_{\text{loc}}$ such that

$$\int_T |\psi\phi| \, d\mu < +\infty \qquad \text{for all} \quad \phi \in \mathscr{L}.$$

Let E be a locally convex Suslin vector space; following [4, V] we define the *generalized Köthe* space $\mathscr{L}_E$ to be the vector space of all functions $f: T \to E$ such that for every $x' \in E'$ the function $t \to \langle f(t), x' \rangle$ belongs to $\mathscr{L}$ (here E' is the topological dual of E). The *weak topology* on $\mathscr{L}_E$ is defined to be the initial topology with respect to the functionals

$$f \to \int_T \psi(t) \langle f(t), x' \rangle \, \mu(dt), \qquad \psi \in \mathscr{L}^*, \qquad x' \in E'.$$

The following result forms an improved sequential analogue of [4, V.13].

THEOREM 6.1. *Suppose that $\Gamma: T \to 2^E$ is a multifunction with convex $\sigma(E, E')$-compact values such that for every $x' \in E'$ the function $\phi^{x'}$ belongs to $\mathscr{L}$, where $\phi^{x'}: T \to \mathbb{R}$ is defined by*

$$\phi^{x'}(t) := \sup_{x \in \Gamma(t)} \langle x, x' \rangle.$$

Then the set $\mathscr{L}_\Gamma$ of all functions $f \in \mathscr{L}_E$ such that $f(t) \in \Gamma(t)$ a.e., is sequentially compact for $\sigma(E, E')$-K-convergence a.e., and in particular weakly sequentially compact.

Proof. Let $\{f_k\}$ be an arbitrary sequence in $\mathscr{L}_\Gamma$. We can apply Theorem 2.1 by making the following substitutions: Define the convex sequentially $\sigma(E, E')$-inf-compact integrand h as

$$h(t, x) := \begin{vmatrix} 0 & \text{if} \quad x \in \Gamma(t) \\ +\infty & \text{otherwise} \end{vmatrix}$$

(note that by [7, III.66] ordinary $\sigma(E, E')$-compactness implies sequential $\sigma(E, E')$-compactness). Then (2.1) is fulfilled. Define $\mathscr{A}$ to be the set of all integrands $a^{x',\psi}: T \times E \to \mathbb{R}$ given by

$$a^{x',\psi}(t, x) := \psi(t)\langle x, x'\rangle, \qquad x' \in E', \qquad \psi \in \mathscr{L}^*.$$

Since

$$|a^{x',\psi}(t, x)| \leqslant \phi^{x'}(t)\, \psi(t), \tag{6.1}$$

where the right-hand side forms an integrable function, all integrands in $\mathscr{A}$ have property (G). Also, for every $x' \in E'$, $\psi \in \mathscr{L}^*$

$$t \to a^{x',\psi}(t, f_k(t)) = \psi(t)\langle f_k(t), x'\rangle \text{ is measurable,}$$

so the elements of $\{f_k\}$ are $\mathscr{A}$-scalarly measurable. Since E is locally convex Suslin, it follows by [4, III.31] that there exists a countable subset $\{x_j'\}$ of E' which separates the points of E (see also the proof of Lemma 3.1). Also, since $(T, \mathscr{T}, \mu)$ is σ-finite, there exists a sequence $\{T_i\}$ of mutually disjoint sets in $\mathscr{T}$, having T as its union, such that $\mu(T_i) < +\infty$ for all $i \in \mathbb{N}$. It is easy to see that for every $i \in \mathbb{N}$ the characteristic function $1_{T_i}: T \to \{0, 1\}$ of T_i belongs to both $\mathscr{L}$ and $\mathscr{L}^*$. Now note that the integrands $a_{i,j}$, $i, j \in \mathbb{N}$, given by

$$a_j(t, x) := 1_{T_i}(t)\langle x, x_j'\rangle,$$

form a countable collection in $\mathscr{A}$ which satisfies (2.2) (take for N the empty set). By Theorem 2.1 there exist a subsequence $\{f_m\}$ of $\{f_k\}$ and an $\mathscr{A}$-scalarly measurable function $f_*: T \to E$ such that

$$\{f_m\}\ \sigma(E, E')\text{-}K\text{-converges a.e. to } f_*,$$
$$f_*(t) \in \Gamma(t) \text{ a.e.}$$

By (6.1) and $\mathscr{A}$-scalar measurability of f_*, it follows immediately that f_* belongs to $\mathscr{L}_E$, and also to $\mathscr{L}_\Gamma$. Finally, by Corollary 2.2

$$\lim_m \int_T \psi(t)\langle f_m(t), x'\rangle\, \mu(dt) = \int_T \psi(t)\langle f_*(t), x'\rangle\, \mu(dt)$$

for every $x' \in E'$, $\psi \in \mathscr{L}^*$, so in particular $\{f_m\}$ converges weakly to f_*. ∎

References

1. (a) E. J. Balder, A general approach to lower semicontinuity and lower closure in optimal control theory, *SIAM J. Control Optim.* **22** (1984), 570–598; (b) E. J. Balder, A general denseness result for relaxed control theory, *Bull. Austral. Math. Soc.* **30**

(1984), 463–475; (c) E. J. BALDER, An extension of Prohorov's theorem for transition probabilities with applications to infinite-dimensional lower closure problems, *Rend. Circ. Mat. Palermo (2)* **34** (1985), 427–447; (d) E. J. BALDER, Infinite-dimensional extension of a theorem of Komlós, *Probab. Theory Related Fields* **81** (1989), 185–188; (e) E. J. BALDER, Unusual applications of a.e. convergence, *in* "Almost Everywhere Convergence: Proceedings of the International Conference on Almost Everywhere Convergence in Probability and Ergodic Theory, June 11–14, 1988" (G. Edgar and L. Sucheston, Eds.), pp. 31–53. Academic Press, New York, 1989; (f) E. J. BALDER, On Prohorov's theorem for transition probabilities, *Travaux Sém. Anal. Convexe* **19** (1989), 9.1–9.11.

2. J. BOURGAIN, The Komlós theorem for vector valued funçtions, unpublished paper, 1984.
3. C. L. BYRNE, Remarks on the set-valued integrals of Debreu and Aumann, *J. Math. Anal. Appl.* **62** (1978), 243–246.
4. C. CASTAING AND M. VALADIER, "Convex Analysis and Measurable Multifunctions," Lecture Notes in Mathematics, No. 580, Springer-Verlag, Berlin, 1977.
5. C. CASTAING, Compacité dans les espaces de mesures de probabilité de transition, *Travaux Sém. Anal. Convexe* **16** (1986), 5.1–5.19.
6. S. D. CHATTERJI, A subsequence principle theory. *Jber d. Dt. Math.-Verein.* **87** (1985), 91–107.
7. C. DELLACHERIE AND P.-A. MEYER, "Probabilités et Potentiel," Hermann, Paris, 1975.
8. (a) J. DIESTEL, Remarks on weak compactness in $L_1(\mu, X)$, *Glasgow Math. J.* **18** (1977), 87–91; (b) J. DIESTEL, "Sequences and Series in Banach Spaces," Springer-Verlag, Berlin, 1984.
9. J. DIESTEL AND J. J. UHL, Progress in vector measures – 1977–1983, *in* "Lecture Notes in Mathematics, No. 1033" (Belley, Dubois and Morales, Eds.), pp. 144–192, Springer-Verlag, Berlin.
10. D. J. H. GARLING, Subsequence principles for vector-valued random variables, *Math. Proc. Cambridge Philos. Soc.* **86** (1979) 301–311.
11. A. AND C. IONESCU-TULCEA, "Topics in the Theory of Lifting," Springer-Verlag, Berlin, 1969.
12. A. JAWHAR, Mesures de transition et applications, *Travaux Sém. Anal. Convexe* **14** (1984), 13.1–13.61.
13. M. A. KHAN, On extensions of the Cournot–Nash theorem, *in* "Advances in Equilibrium Theory" (C. D. Aliprantis *et al.*, Eds.), pp. 79–106, Lecture Notes in Economics and Mathematical Sciences, No. 244, Springer-Verlag, 1985.
14. J. KOMLÓS, A generalisation of a problem of Steinhaus, *Acta Math. Acad. Sci. Hungar.* **18** (1967), 217–229.
15. W. SCHACHERMEYER, The Banach–Saks property is not L^2-hereditary, *Israel J. Math.* **40** (1981) 340–344.
16. A. M. SUCHANEK, On almost sure convergence of Cesaro averages of subsequences of vector-valued functions, *J. Multivariate Anal.* **8** (1978) 589–597.
17. R. L. TAYLOR, "Stochastic Convergence of Weighted Sums of Random Elements in Linear Spaces," Lecture Notes in Mathematics, No. 672, Springer-Verlag, Berlin, 1978.