Available online at www.sciencedirect.com
SCIENCE DIRECT

J. Math. Anal. Appl. 310 (2005) 609–625

Journal of MATHEMATICAL ANALYSIS AND APPLICATIONS
www.elsevier.com/locate/jmaa

The optimal form of selection principles for functions of a real variable

Vyacheslav V. Chistyakov [*,1]

Department of Mathematics, State University Higher School of Economics, Bol'shaya Pechërskaya Street 25, Nizhny Novgorod 603600, Russia

Received 15 June 2004

Available online 24 March 2005

Submitted by D. Waterman

Abstract

Let T be a nonempty set of real numbers, X a metric space with metric d and X^T the set of all functions from T into X. If $f \in X^T$ and n is a positive integer, we set $\nu(n, f) = \sup \sum_{i=1}^{n} d(f(b_i), f(a_i))$, where the supremum is taken over all numbers $a_1, \dots, a_n, b_1, \dots, b_n$ from T such that $a_1 \leqslant b_1 \leqslant a_2 \leqslant b_2 \leqslant \cdots \leqslant a_n \leqslant b_n$. The sequence $\{\nu(n, f)\}_{n=1}^{\infty}$ is called the *modulus of variation of* f in the sense of Chanturiya. We prove the following pointwise selection principle: *If a sequence of functions* $\{f_j\}_{j=1}^{\infty} \subset X^T$ *is such that the closure in* X *of the set* $\{f_j(t)\}_{j=1}^{\infty}$ *is compact for each* $t \in T$ *and*

$$\lim_{n\to\infty}\left(\frac{1}{n}\limsup_{j\to\infty}\nu(n, f_j)\right) = 0, \tag{$*$}$$

then there exists a subsequence of $\{f_j\}_{j=1}^{\infty}$, *which converges in* X *pointwise on* T *to a function* $f \in X^T$ *satisfying* $\lim_{n\to\infty}\nu(n, f)/n = 0$. We show that condition $(*)$ is optimal (the best possible) and that all known pointwise selection theorems follow from this result (including Helly's theorem). Also, we establish several variants of the above theorem for the almost everywhere convergence and weak pointwise convergence when X is a reflexive separable Banach space.

* Current address: Institute of Mathematics, Technical University of Łódź, ul. Wólczańska 215, 93-005 Łódź, Poland.

E-mail address: czeslaw@mail.ru.

1 Supported by the Russian Foundation for Basic Research Grant 03-01-00473 and J. Mianowski Fund/ Foundation of Polish Science (2004).

0022-247X/$ – see front matter
doi:10.1016/j.jmaa.2005.02.028

Keywords: Modulus of variation; Selection principle; Pointwise convergence; Proper function; Generalized variation

1. Main result

We begin with reviewing certain definitions and facts needed for our results. Throughout the paper we assume $T \subset \mathbb{R}$ to be a nonempty set, X a metric space with metric d and X^T the set of all functions $f : T \to X$ mapping T into X. Given a sequence $\{f_j\} = \{f_j\}_{j=1}^{\infty} \subset X^T$ and $f \in X^T$, we write $f_j \to f$ on T to denote the *pointwise* (or *everywhere*) *convergence* of f_j to f as $j \to \infty$, i.e., $\lim_{j\to\infty} d(f_j(t), f(t)) = 0$ for all $t \in T$. A sequence $\{f_j\} \subset X^T$ is said to be *pointwise precompact* (on T) provided the sequence $\{f_j(t)\}$ is precompact in X (i.e., its closure in X is compact) for all $t \in T$.

Let $\mathrm{M}(T;\mathbb{R}) = \{f \in \mathbb{R}^T \mid f$ is nondecreasing and bounded$\}$. Helly's theorem states that *a uniformly bounded sequence of functions from* $\mathrm{M}(T;\mathbb{R})$ *contains a pointwise convergent subsequence* ([20], and also [21, II.8.9–10], [25, VIII.4.2] if T is a closed interval $[a, b]$ and [13, Theorem 1.3] if T is arbitrary). This theorem implies a number of selection principles for functions of various types of bounded (generalized) variations having real values [1,24, 29] as well as values from a metric or Banach space ([2, 1.3.5], [3,6–13,15,16,26][2]). As an example, *a pointwise precompact sequence* $\{f_j\} \subset X^T$ *of uniformly bounded* (*Jordan*) *variation contains a subsequence which converges pointwise on* T *to a function from* X^T *of bounded* (*Jordan*) *variation* [3,13]. Such (Helly type) selection principles have numerous applications in analysis (cf. [13] and references therein) since they provide efficient tools for proving existence theorems (see also [18] where Helly's theorem has been generalized to monotone functions between linearly ordered sets).

The aim of this paper is to present a unified approach to the diverse selection principles mentioned above without invoking the uniform boundedness of variations of any kind. Our main result (Theorem 1 below) gives a sufficient condition for extracting a pointwise convergent subsequence, but it turns out to be (almost necessary and) the best possible in the sense to be made precise (see Lemma 4(a2), (b)). In order to formulate it, we need a definition.

Given $n \in \mathbb{N}$, $f \in X^T$ and $\emptyset \neq E \subset T$, we set

$$\nu(n, f, E) = \sup\left\{ \sum_{i=1}^{n} d\big(f(b_i), f(a_i)\big) \;\middle|\; \{a_i\}_{i=1}^n, \{b_i\}_{i=1}^n \subset E \text{ such that } a_1 \leqslant b_1 \leqslant a_2 \leqslant b_2 \leqslant \cdots \leqslant a_{n-1} \leqslant b_{n-1} \leqslant a_n \leqslant b_n \right\}.$$

The sequence $\nu(\cdot, f, E) : \mathbb{N} \to [0, \infty]$ is called the *modulus of variation* of f on E. This notion was first considered by Chanturiya in [4] and [5] (see also [19, Section 11.3.7]) for

[2] I have not seen book [26] in its original form: in [2, Remark 3.2 on p. 60] the authors refer to [26] where a selection principle is established which is originally due to Foias.

$E = T = [a, b]$ and $X = \mathbb{R}$ in connection with convergence problems from the theory of Fourier series. It will play an important role in our considerations as well.

For a sequence $\mu : \mathbb{N} \to \mathbb{R}$ we employ Landau's notation $\mu(n) = o(n)$ to denote the condition $\lim_{n\to\infty} \mu(n)/n = 0$. Note at once (cf. Lemma 3 in Section 2) that if X is complete, then a function $f : [a, b] \to X$ has left and right limits at all points of $[a, b]$ if and only if $\nu(n, f, [a, b]) = o(n)$. Thus, the modulus of variation characterizes functions with simple discontinuities rather than functions of bounded variation of any type.

The following theorem is a *pointwise selection principle* for metric space valued functions of a real variable in terms of modulus of variation.

Theorem 1. *Let $\emptyset \ne T \subset \mathbb{R}$ and (X, d) be a metric space. Suppose that $\{f_j\} \subset X^T$ is a pointwise precompact sequence such that*

$$\mu(n) \equiv \limsup_{j\to\infty} \nu(n, f_j, T) = o(n). \tag{1}$$

Then there exists a subsequence of $\{f_j\}$, which converges pointwise on T to a function $f \in X^T$ satisfying $\nu(n, f, T) \leqslant \mu(n)$, $n \in \mathbb{N}$.

In order to see how this theorem implies all the above mentioned selection principles, let us recall three classical notions of bounded (generalized) variation.

Let $\varphi : \mathbb{R}^+ = [0, \infty) \to \mathbb{R}^+$ be a *φ-function*, that is, φ is nondecreasing, continuous, $\varphi(\rho) = 0$ if and only if $\rho = 0$, and $\lim_{\rho\to\infty} \varphi(\rho) = \infty$. We say that $f \in X^T$ is of *bounded φ-variation* in the sense of Wiener and Young (e.g., [11,13,14,16,23,24]) and write $f \in \mathrm{BV}_\varphi(T; X)$ if

$$V_\varphi(f, T) = \sup\left\{ \sum_{i=1}^{m} \varphi\big(d\big(f(t_i), f(t_{i-1})\big)\big) \;\middle|\; m \in \mathbb{N},\ t_{i-1} \leqslant t_i,\ i = 1, \ldots, m \right\} < \infty.$$

If $\varphi(\rho) = \rho$, $V_\varphi(f, T)$ is the classical variation of f in the sense of Jordan, which we denote by $V(f, T)$, and write $\mathrm{BV}(T; X)$ instead of $\mathrm{BV}_\varphi(T; X)$. Note that if φ is superadditive (i.e., $\varphi(\rho_1) + \varphi(\rho_2) \leqslant \varphi(\rho_1 + \rho_2)$ for all $\rho_1, \rho_2 \in \mathbb{R}^+$), then $\mathrm{BV}(T; X) \subset \mathrm{BV}_\varphi(T; X)$; in addition, if φ is convex and $\lim_{\rho\to 0} \varphi(\rho)/\rho = 0$, then $\mathrm{BV}(T; X)$ is a strict subset of $\mathrm{BV}_\varphi(T; X)$.

Let $\Lambda = \{\lambda_i\}_{i=1}^{\infty} \subset (0, \infty)$ be a nondecreasing sequence such that $\sum_{i=1}^{\infty} 1/\lambda_i = \infty$. A function $f \in X^T$ is said to be of *Λ-bounded variation* in the sense of Waterman ([28, 29], [19, Section 11.3]), in symbols $f \in \Lambda\mathrm{BV}(T; X)$, provided

$$V_\Lambda(f, T) = \sup \sum_{i=1}^{m} \frac{d(f(b_i), f(a_i))}{\lambda_{\sigma(i)}} < \infty,$$

where the supremum is taken over all $m \in \mathbb{N}$, $\{a_i\}_{i=1}^m, \{b_i\}_{i=1}^m \subset T$ such that $a_1 \leqslant b_1 \leqslant a_2 \leqslant b_2 \leqslant \cdots \leqslant a_m \leqslant b_m$ and all permutations $\sigma : \{1, \ldots, m\} \to \{1, \ldots, m\}$. Note that if Λ is an unbounded sequence, then $\mathrm{BV}(T; X)$ is a strict subset of $\Lambda\mathrm{BV}(T; X)$.

Given $n \in \mathbb{N}$, the following relations hold:

$$\nu(n, f, T) = \sup_{t\in T} f(t) - \inf_{t\in T} f(t), \qquad f \in \mathrm{M}(T; \mathbb{R});$$

$$\nu(n, f, T) \leqslant V(f, T) = \lim_{n\to\infty} \nu(n, f, T), \qquad f \in \mathrm{BV}(T; X); \tag{2}$$

if φ is a convex φ-function, then (it admits the continuous inverse φ^{-1} and)

$$\nu(n, f, T) \leqslant n\varphi^{-1}\left(\frac{V_\varphi(f, T)}{n}\right), \quad f \in \mathrm{BV}_\varphi(T; X) \ \big(\text{cf. [4]}\big);$$

$$\nu(n, f, T) \leqslant \frac{n}{\sum_{i=1}^{n} 1/\lambda_i} V_\Lambda(f, T), \quad f \in \Lambda\,\mathrm{BV}(T; X) \ \big(\text{cf. [19, Theorem 11.17]}\big).$$

Now, let $\mathrm{BV}_*(T; X)$ denote one of the sets $\mathrm{BV}(T; X)$, $\mathrm{BV}_\varphi(T; X)$ with convex φ-function φ (the case of general φ will be treated in Example 7 of Section 3) or $\Lambda\,\mathrm{BV}(T; X)$ and $V_*(f, T)$ designate the variation in the corresponding set: $V(f, T)$, $V_\varphi(f, T)$ or $V_\Lambda(f, T)$. If a pointwise precompact sequence $\{f_j\} \subset \mathrm{BV}_*(T; X)$ is such that $\sup_{j\in\mathbb{N}} V_*(f_j, T) = C < \infty$ (the usual assumption of the uniform boundedness of variations), then the inequalities above yield: $\sup_{j\in\mathbb{N}} \nu(n, f_j, T) = o(n)$. By Theorem 1, a subsequence of $\{f_j\}$ (denoted as the whole sequence) converges pointwise on T to a function $f \in X^T$. Since the functional $V_*(\cdot, T)$ is sequentially lower semi-continuous with respect to the pointwise convergence in X^T, we have $V_*(f, T) \leqslant \liminf_{j\to\infty} V_*(f_j, T) \leqslant C$, and so, the pointwise limit f is in $\mathrm{BV}_*(T; X)$.

The paper is organized as follows. In Section 2 we establish properties of the modulus of variation and prove Theorem 1. Section 3 contains various examples illustrating the optimality of Theorem 1. A selection principle for the almost everywhere convergence of an extracted subsequence is treated in Section 4. In the final Section 5 we prove a selection principle including weak pointwise convergence and weak almost everywhere convergence when values of functions under consideration lie in a reflexive separable Banach space.

2. Pointwise selection principle

It follows from the definition of the value $\nu(n, f, E)$ that it is finite for each $n \in \mathbb{N}$, and so, $\nu(\cdot, f, E) : \mathbb{N} \to \mathbb{R}^+$, if and only if f is bounded on E (i.e., $\sup_{t,s\in E} d(f(t), f(s)) < \infty$). In what follows all functions $f \in X^T$ under consideration are assumed to be bounded.

The straightforward properties of the modulus of variation, needed for our purposes, are gathered in the following

Lemma 2. *Given $f \in X^T$ and $\emptyset \neq E \subset T$, we have*:

(a) *the sequence $\{\nu(n, f, E)\}_{n=1}^{\infty}$ is nondecreasing* [4];
(b) $\nu(n+m, f, E) \leqslant \nu(n, f, E) + \nu(m, f, E)$ *for all $n, m \in \mathbb{N}$* [4];
(c) $\nu(n+1, f, E) \leqslant \nu(n, f, E) + \frac{\nu(n+1, f, E)}{n+1}$ *for all $n \in \mathbb{N}$* [5, Lemma];
(d) $\nu(n, f, E') \leqslant \nu(n, f, E)$ *for all $\emptyset \neq E' \subset E$ and $n \in \mathbb{N}$*;
(e) $\nu(n, f, E) \leqslant \liminf_{j\to\infty} \nu(n, f_j, E)$ *for all $\{f_j\} \subset X^T$ such that $f_j \to f$ on E and all $n \in \mathbb{N}$*;
(f) $d(f(t), f(s)) + \nu(n, f, (-\infty, s] \cap E) \leqslant \nu(n+1, f, (-\infty, t] \cap E)$ *for all $s, t \in E$ such that $s \leqslant t$ and all $n \in \mathbb{N}$.*

As a consequence of Lemma 2(c), for any bounded function $f \in X^T$ the sequence $\{\nu(n, f, E)/n\}_{n=1}^{\infty}$ is nonincreasing, and so, the following limit always exists: $\lim_{n\to\infty} \nu(n, f, E)/n \in \mathbb{R}^+$.

Two more modes of convergence of $\{f_j\} \subset X^T$ to $f \in X^T$ will be of significance: *uniform*, $f_j \rightrightarrows f$ on T, that is, $\lim_{j\to\infty} \sup_{t\in T} d(f_j(t), f(t)) = 0$; and *almost everywhere*, $f_j \to f$ a.e. on T, that is, $f_j \to f$ on $T \setminus E$ for some set $E \subset T$ of Lebesgue measure zero, $\mathcal{L}(E) = 0$.

A function $f \in X^{[a,b]}$ is said to be *proper* if it satisfies the Cauchy condition at every point of $[a, b]$, i.e., $d(f(t), f(s)) \to 0$ as $t, s \to \tau - 0$ for each point $a < \tau \leqslant b$ and $d(f(t), f(s)) \to 0$ as $t, s \to \tau + 0$ for each point $a \leqslant \tau < b$. If X is complete, then f is, by virtue of the Cauchy criterion, proper if and only if at each point $a < \tau \leqslant b$ the left limit $f(\tau - 0) \in X$ exists (i.e., $d(f(t), f(\tau - 0)) \to 0$ as $t \to \tau - 0$) and at each point $a \leqslant \tau < b$ the right limit $f(\tau + 0) \in X$ exists (and so, $d(f(t), f(\tau + 0)) \to 0$ as $t \to \tau + 0$).

The following illustrative result was first stated in [4, Theorem 5] for $X = \mathbb{R}$ without proof.

Lemma 3. *A function $f \in X^{[a,b]}$ is proper if and only if $\nu(n, f, [a, b]) = o(n)$.*

Proof. *Sufficiency.* Given $n \in \mathbb{N}$, we set $\nu_n(t) = \nu(n, f, [a, t])$, $t \in [a, b]$. By Lemma 2(d), $\nu_n : [a, b] \to \mathbb{R}^+$ is nondecreasing and, hence, proper. Let $a < \tau \leqslant b$ and $\nu_n(\tau - 0)$ be the corresponding left limit. If $a \leqslant s \leqslant t < \tau$, by Lemma 2(f), (c), (d), we have:

$$\begin{aligned} d\big(f(t), f(s)\big) &\leqslant \nu_{n+1}(t) - \nu_n(s) \leqslant \nu_n(t) + \frac{\nu_{n+1}(t)}{n+1} - \nu_n(s) \\ &\leqslant \big|\nu_n(t) - \nu_n(\tau - 0)\big| + \frac{\nu(n+1, f, [a, b])}{n+1} + \big|\nu_n(\tau - 0) - \nu_n(s)\big|. \end{aligned}$$

For $\varepsilon > 0$ choose and fix $n = n(\varepsilon) \in \mathbb{N}$ such that $\nu(n + 1, f, [a, b])/(n + 1) \leqslant \varepsilon/3$. Let $0 < \delta = \delta(\varepsilon) < \tau - a$ be such that if $\tau - \delta \leqslant t < \tau$, then $|\nu_n(t) - \nu_n(\tau - 0)| \leqslant \varepsilon/3$. It follows that if $t, s \in [\tau - \delta, \tau)$, then $d(f(t), f(s)) \leqslant \varepsilon$, which proves that $d(f(t), f(s)) \to 0$ as $t, s \to \tau - 0$. The case when $a \leqslant \tau < b$ and $d(f(t), f(s)) \to 0$ as $t, s \to \tau + 0$ is treated similarly.

Necessity. Being proper, the function f is the uniform limit on $[a, b]$ of a sequence $\{f_j\} \subset X^{[a,b]}$ of step functions (e.g., [17, (7.6.1)]; recall that $f_j \in X^{[a,b]}$ is called a *step function* if there exists a partition $a = t_0 < t_1 < \cdots < t_{m-1} < t_m = b$ of $[a, b]$ such that f_j takes a constant value on each interval (t_{i-1}, t_i), $i = 1, \ldots, m$). Since step functions belong to $\mathrm{BV}([a, b]; X)$, the equality in (2) implies $\nu(n, f_j, [a, b]) = o(n)$ for all $j \in \mathbb{N}$. Now the result follows from the uniform convergence of f_j to f and the estimate:

$$\frac{\nu(n, f, T)}{n} \leqslant \frac{\nu(n, f_j, T)}{n} + 2\sup_{t\in T} d\big(f_j(t), f(t)\big), \quad T = [a, b],\ n, j \in \mathbb{N}. \tag{3}$$

In fact, given $\varepsilon > 0$, there exists $j = j(\varepsilon) \in \mathbb{N}$ such that, for all $t \in T$, $d(f_j(t), f(t)) \leqslant \varepsilon/3$, and there exists $n_0 = n_0(\varepsilon) \in \mathbb{N}$ such that $\nu(n_0, f_j, T)/n_0 \leqslant \varepsilon/3$. Therefore, $\nu(n, f, T)/n \leqslant \varepsilon$ for all $n \geqslant n_0$, which was to be proved. □

Lemma 3 implies, in particular, that all functions belonging to $\mathrm{BV}([a, b]; X)$, $\mathrm{BV}_\varphi([a, b]; X)$ and $\Lambda\mathrm{BV}([a, b]; X)$ are proper.

It is known (e.g., [17, (7.6.3)]) that the image $f([a,b]) = \{f(t) \mid t \in [a,b]\}$ of a proper function $f \in X^{[a,b]}$ is totally bounded in X (precompact if X is complete). This is also true for proper multifunctions with compact values with respect to the Hausdorff metric (cf. [12, Lemma 11 and its proof]).

Items (a2) and (b) in the next lemma may be considered as *partial converses* of Theorem 1 showing at the same time the optimality of condition (1).

Lemma 4.

(a) *Suppose* $\{f_j\} \subset X^T$, $f \in X^T$ *and* $f_j \rightrightarrows f$ *on* T. *We have*:
  (a1) $\lim_{j\to\infty} \nu(n, f_j, T) = \nu(n, f, T)$ *for all* $n \in \mathbb{N}$;
  (a2) *if* $\nu(n, f, T) = o(n)$, *then* $\lim_{j\to\infty} \nu(n, f_j, T) = o(n)$; *however, it may happen that* $\nu(n, f_j, T) \neq o(n)$ *for all* $j \in \mathbb{N}$;
  (a3) *if* $\nu(n, f_j, T) = o(n)$ *for all* $j \in \mathbb{N}$, *then* $\nu(n, f, T) = o(n)$.
  Assertions (a1)–(a3) *are wrong for the pointwise convergence.*

(b) *If* T *is a measurable set with finite Lebesgue measure* $\mathcal{L}(T)$, $\{f_j\} \subset X^T$ *is a sequence of measurable functions,* $f \in X^T$, $\nu(n, f, T) = o(n)$, *and* $f_j \to f$ *a.e. on* T (*or* $f_j \to f$ *on* T), *then for each* $\varepsilon > 0$ *there exists a measurable set* $E = E(\varepsilon) \subset T$ *with* $\mathcal{L}(E) \leqslant \varepsilon$ *such that* $\lim_{j\to\infty} \nu(n, f_j, T \setminus E) = o(n)$.

Proof. (a1) Passing to the limit superior as $j \to \infty$ in the inequality (cf. (3))

$$\nu(n, f_j, T) \leqslant \nu(n, f, T) + 2n \sup_{t\in T} d\big(f(t), f_j(t)\big), \quad n, j \in \mathbb{N},$$

we get $\limsup_{j\to\infty} \nu(n, f_j, T) \leqslant \nu(n, f, T)$, $n \in \mathbb{N}$, by virtue of the uniform convergence of f_j to f, and it remains to take into account Lemma 2(e).

(a2) The first part is a consequence of (a1). As for the second part, see Example 6 in Section 3.

(a3) Replace $[a,b]$ by T in the necessity part of the proof of Lemma 3.

That (a1)–(a3) are wrong for pointwise convergence, see Examples 4 and 5 in Section 3.

(b) By the assumptions and Egorov's theorem (e.g., [27, Theorem 3.2.7]), for each $\varepsilon > 0$ there exists a measurable set $E = E(\varepsilon) \subset T$ with $\mathcal{L}(E) \leqslant \varepsilon$ such that $f_j \rightrightarrows f$ on $T \setminus E$. Since $\nu(n, f, T) = o(n)$, Lemma 2(d) implies $\nu(n, f, T \setminus E) = o(n)$. Our assertion follows from Lemma 4(a2). □

Now we are in a position to prove our main result.

Proof of Theorem 1. (1) First, making use of the standard diagonal process we show that there are a subsequence of $\{f_j\}$ (for which, without loss of generality, we use the notation of the original sequence) and a nondecreasing sequence $\gamma : \mathbb{N} \to \mathbb{R}^+$ such that

$$\lim_{j\to\infty} \nu(n, f_j, T) = \gamma(n) \leqslant \mu(n) \quad \text{for all } n \in \mathbb{N}. \tag{4}$$

Set $\gamma(1) = \mu(1)$. Since $\limsup_{j\to\infty} \nu(1, f_j, T) = \mu(1)$, there exists a subsequence $\{f_j^{(1)}\}_{j=1}^\infty$ of $\{f_j\}$ such that $\lim_{j\to\infty} \nu(1, f_j^{(1)}, T) = \gamma(1)$. Inductively, if $n \geqslant 2$ and a subsequence $\{f_j^{(n-1)}\}_{j=1}^\infty$ of $\{f_j\}$ is already chosen, we set $\gamma(n) = \limsup_{j\to\infty} \nu(n, f_j^{(n-1)}, T)$

and, since $\gamma(n) \leqslant \mu(n)$, we pick a subsequence $\{f_j^{(n)}\}_{j=1}^\infty$ of $\{f_j^{(n-1)}\}_{j=1}^\infty$ such that $\lim_{j\to\infty} \nu(n, f_j^{(n)}, T) = \gamma(n)$. Then the diagonal sequence $\{f_j^{(j)}\}_{j=1}^\infty$, which we denote by $\{f_j\}$, enjoys properties (4).

(2) Let us show that there exists a subsequence of $\{f_j\}$ from (4) (which we again will denote as the whole sequence $\{f_j\}$) and for each $n \in \mathbb{N}$ there exists a function $\nu_n \in \mathrm{M}(T; \mathbb{R}^+)$ such that

$$\lim_{j\to\infty} \nu\big(n, f_j, (-\infty, t] \cap T\big) = \nu_n(t) \quad \text{for all } n \in \mathbb{N} \text{ and } t \in T. \tag{5}$$

Given $n \in \mathbb{N}$, by Lemma 2(d) the function $\eta(n, f_j, t) = \nu(n, f_j, (-\infty, t] \cap T)$ is nondecreasing in $t \in T$, and it follows from the equality in (4) that there exists a constant $C(n) \in \mathbb{R}^+$ such that $\nu(n, f_j, T) \leqslant C(n)$ for all $j \in \mathbb{N}$. Again, we apply the diagonal process. The sequence $\{\eta(1, f_j, \cdot)\}_{j=1}^\infty \subset \mathrm{M}(T; \mathbb{R}^+)$ is uniformly bounded by $C(1)$, and so, by Helly's theorem, there exists a subsequence $\{f_j^{(1)}\}_{j=1}^\infty$ of $\{f_j\}$ and a function $\nu_1 \in \mathrm{M}(T; \mathbb{R}^+)$ such that $\eta(1, f_j^{(1)}, t) \to \nu_1(t)$ as $j \to \infty$ for all $t \in T$. If $n \geqslant 2$ and a subsequence $\{f_j^{(n-1)}\}_{j=1}^\infty$ of $\{f_j\}$ is already chosen, by Helly's theorem applied to the sequence $\{\eta(n, f_j^{(n-1)}, \cdot)\}_{j=1}^\infty \subset \mathrm{M}(T; \mathbb{R}^+)$, which is uniformly bounded by $C(n)$, we find a subsequence $\{f_j^{(n)}\}_{j=1}^\infty$ of $\{f_j^{(n-1)}\}_{j=1}^\infty$ such that $\eta(n, f_j^{(n)}, \cdot)$ converges pointwise on T as $j \to \infty$ to a function $\nu_n \in \mathrm{M}(T; \mathbb{R}^+)$. It follows that the diagonal sequence $f_j = f_j^{(j)}$, $j \in \mathbb{N}$, satisfies (5).

(3) Denote by Q an at most countable dense subset of T (so that $Q \subset T \subset \bar{Q}$) and note that any point $t \in T$, which is not a limit point for T, belongs to Q. Since ν_n is monotone, the set $Q_n \subset T$ of its points of discontinuity is at most countable. We set $S = Q \cup \bigcup_{n=1}^\infty Q_n$. Then S is at most countable dense subset of T and, if $T \setminus S \neq \emptyset$,

$$\text{each function } \nu_n \text{ is continuous at points } t \in T \setminus S,\ n \in \mathbb{N}. \tag{6}$$

Since the set $\{f_j(t)\}$ is precompact in X for all $t \in T$ and $S \subset T$ is at most countable, without loss of generality we may assume (again applying the diagonal process and passing to a subsequence of $\{f_j\}$ if necessary) that $f_j(s)$ converges in X as $j \to \infty$ to a point denoted by $f(s) \in X$ for all $s \in S$. If $T = S$, the proof is complete.

(4) Suppose $T \neq S$. Let us prove that, given $t \in T \setminus S$, the sequence $\{f_j(t)\}$ converges in X. For this, we fix arbitrary $\varepsilon > 0$. By the assumption, $\mu(n)/n \to 0$ as $n \to \infty$, so we choose and fix $n = n(\varepsilon) \in \mathbb{N}$ such that $\mu(n+1)/(n+1) \leqslant \varepsilon/15$. By virtue of (4), there exists $j_1 = j_1(\varepsilon, n) \in \mathbb{N}$ such that $\nu(n+1, f_j, T) \leqslant \gamma(n+1) + (\varepsilon/15)$ for all $j \geqslant j_1$. The definition of S and (6) imply that the point t is a limit point for T and a point of continuity of ν_n, so by the density of S in T there exists $s = s(\varepsilon, t, n) \in S$ such that $|\nu_n(t) - \nu_n(s)| \leqslant \varepsilon/15$. Property (5) yields the existence of $j_2 = j_2(\varepsilon, t, s, n) \in \mathbb{N}$ such that if $j \geqslant j_2$, then

$$\big|\nu\big(n, f_j, (-\infty, t] \cap T\big) - \nu_n(t)\big| \leqslant \frac{\varepsilon}{15} \quad \text{and}$$
$$\big|\nu\big(n, f_j, (-\infty, s] \cap T\big) - \nu_n(s)\big| \leqslant \frac{\varepsilon}{15}.$$

Assuming (without loss of generality) that $s < t$ and applying items (f), (c) and (d) of Lemma 2 and (4), we get for all $j \geqslant \max\{j_1, j_2\}$:

$$\begin{aligned}
d\big(f_j(t), f_j(s)\big) &\leqslant \nu\big(n+1, f_j, (-\infty, t] \cap T\big) - \nu\big(n, f_j, (-\infty, s] \cap T\big) \\
&\leqslant \nu\big(n+1, f_j, (-\infty, t] \cap T\big) - \nu\big(n, f_j, (-\infty, t] \cap T\big) \\
&\quad + \big|\nu\big(n, f_j, (-\infty, t] \cap T\big) - \nu_n(t)\big| + \big|\nu_n(t) - \nu_n(s)\big| \\
&\quad + \big|\nu_n(s) - \nu\big(n, f_j, (-\infty, s] \cap T\big)\big| \\
&\leqslant \frac{\nu(n+1, f_j, (-\infty, t] \cap T)}{n+1} + \frac{\varepsilon}{15} + \frac{\varepsilon}{15} + \frac{\varepsilon}{15} \\
&\leqslant \frac{\gamma(n+1)}{n+1} + \frac{\varepsilon}{15(n+1)} + \frac{3\varepsilon}{15} \\
&\leqslant \frac{\mu(n+1)}{n+1} + \frac{4\varepsilon}{15} \leqslant \frac{\varepsilon}{3}.
\end{aligned}$$

Since $\{f_j(s)\}$ is convergent, it is a Cauchy sequence, and so, there exists $j_3 = j_3(\varepsilon, s) \in \mathbb{N}$ such that $d\big(f_j(s), f_{j'}(s)\big) \leqslant \varepsilon/3$ for all $j, j' \geqslant j_3$. It follows that $j_4 = \max\{j_1, j_2, j_3\}$ depends on ε only and for all $j, j' \geqslant j_4$ we have:

$$d\big(f_j(t), f_{j'}(t)\big) \leqslant d\big(f_j(t), f_j(s)\big) + d\big(f_j(s), f_{j'}(s)\big) + d\big(f_{j'}(s), f_{j'}(t)\big) \leqslant \varepsilon.$$

Thus, $\{f_j(t)\}$ is a Cauchy sequence in X and, since it is precompact in X, it is convergent to a point denoted by $f(t) \in X$.

(5) The function $f \in X^T$ defined at the end of steps (3) and (4) is the pointwise limit on T of the sequence $\{f_j\}$ (which is a subsequence of the original sequence). Applying Lemma 2(e), we conclude that

$$\nu(n, f, T) \leqslant \liminf_{j\to\infty} \nu(n, f_j, T) \leqslant \limsup_{j\to\infty} \nu(n, f_j, T) \leqslant \mu(n), \quad n \in \mathbb{N}. \qquad \square$$

Clearly, in Theorem 1 we have $\nu(n, f, T) = o(n)$ for the limit function f, although we did not suppose for $j \in \mathbb{N}$ that $\nu(n, f_j, T) = o(n)$. Cf. also Examples 3 and 6 in Section 3.

Applying Theorem 1 and the diagonal process over expanding intervals, we get the following *local* version of Theorem 1:

Corollary 5. *If $\{f_j\} \subset X^T$ is a pointwise precompact sequence such that*

$$\limsup_{j\to\infty} \nu\big(n, f_j, [a, b] \cap T\big) = o(n) \quad \textit{for all } a, b \in T,\ a \leqslant b,$$

then a subsequence of $\{f_j\}$ converges pointwise on T to a function $f \in X^T$ satisfying $\nu(n, f, [a, b] \cap T) = o(n)$ for all $a, b \in T$, $a \leqslant b$.

3. Examples

All assumptions in Theorem 1 are essential for its validity as the following examples show.

Example 1. If $(X, \|\cdot\|)$ is a finite-dimensional normed vector space, then, by virtue of the inequalities,

$$\sup_{t\in T}\|f_j(t)\| \leqslant \|f_j(t_0)\| + \nu(1, f_j, T) \leqslant \|f_j(t_0)\| + C(1),$$

where $C(1)$ is the constant $C(n)$ from step (2) of the proof of Theorem 1 corresponding to $n = 1$, condition "$\{f_j\} \subset X^T$ is pointwise precompact" in Theorem 1 may be replaced by an equivalent condition "$\{f_j\} \subset X^T$ and $\{f_j(t_0)\}$ is bounded for some $t_0 \in T$." In contrast to this, for an infinite-dimensional Banach space X the precompactness of $\{f_j(t)\}$ at all points $t \in T$ cannot be replaced by their boundedness and closedness even at a single point. In fact, let $T = [0, 1]$, $X = \ell^1 \equiv \{x : \mathbb{N} \to \mathbb{R}$ such that $\|x\| = \sum_{i=1}^{\infty} |x(i)| < \infty\}$ and, if $j \in \mathbb{N}$, let the element $e_j = \{e(i)\}_{i=1}^{\infty} \in \ell^1$ be given by $e(i) = 0$ if $i \neq j$ and $e(j) = 1$. Define $f_j \in X^T$ by $f_j(0) = e_j$ and $f_j(t) = 0$ if $0 < t \leqslant 1$, $j \in \mathbb{N}$. We have: $\{f_j(0)\} = \{e_j\}$ is bounded and closed, $\{f_j(t)\} = \{0\}$ is compact if $0 < t \leqslant 1$, $\nu(n, f_j, T) = V(f_j, T) = 1$ for all $n, j \in \mathbb{N}$, and no subsequence of $\{f_j\}$ converges in X at the point $t = 0$.

Example 2. Continuity of the sequence $\{f_j\}$ is not preserved in the limit procedure of Theorem 1: $\{f_j\} \subset \mathbb{R}^{[0,2]}$ where $f_j(t) = t^j$ if $0 \leqslant t \leqslant 1$ and $f_j(t) = (2-t)^j$ if $1 \leqslant t \leqslant 2$.

Example 3. In general, absent condition (1) Theorem 1 is wrong. It is well known that the sequence $\{f_j\} \subset \mathbb{R}^{[0,2\pi]}$ defined by $f_j(t) = \sin(jt)$, $t \in T = [0, 2\pi]$, has no subsequence convergent at all points of T. Given $n, j \in \mathbb{N}$, a straightforward calculation shows that

$$\nu(n, f_j, T) = \begin{cases} 2n & \text{if } 1 \leqslant n < 2j, \\ 2n - 1 = 4j - 1 & \text{if } n = 2j, \\ 4j = V(f_j, T) & \text{if } n \geqslant 2j + 1. \end{cases} \tag{7}$$

It follows that $\lim_{j\to\infty} \nu(n, f_j, T) = 2n \neq o(n)$. In view of Lemma 3, this example also ensures that condition (1) in Theorem 1 cannot be replaced by "$\nu(n, f_j, T) = o(n)$ for all $j \in \mathbb{N}$," and that one cannot interchange the limits $\lim_{n\to\infty}$ and $\limsup_{j\to\infty}$ in this condition.

Example 4. In this example we will show that: (i) condition (1) in Theorem 1 is *not necessary*, although we have $\nu(n, f, T) = o(n)$ and $\nu(n, f_j, T) = o(n)$ for all $j \in \mathbb{N}$, and (ii) assertions (a1) and (a2) in Lemma 4 are not valid for pointwise convergence, and the inequality in Lemma 2(e) may be strict.

Define the sequence $\{f_j\} \subset \mathbb{R}^{[0,2\pi]}$ by

$$f_j(t) = \begin{cases} \sin(j^2 t) & \text{if } 0 \leqslant t \leqslant 2\pi/j, \\ 0 & \text{if } 2\pi/j \leqslant t \leqslant 2\pi, \end{cases} \qquad j \in \mathbb{N}.$$

Clearly, f_j converges pointwise on $T = [0, 2\pi]$ to $f \equiv 0$. The graph of f_j on $[0, 2\pi/j]$ "looks like" the graph of $t \mapsto \sin(jt)$ on $[0, 2\pi]$ and, in particular, we have $j = j^2/j$ flattened copies of graphs of the ordinary sine function on its period and $V(f_j, [0, 2\pi]) = V(f_j, [0, 2\pi/j]) = 4j$. Thus, the modulus of variation of our sequence is given by (7), and so,

$$0 = \nu(n, f, T) < \lim_{j\to\infty} \nu(n, f_j, T) = 2n \neq o(n).$$

Example 5. Here we will see that Lemma 4(a3) is wrong for the pointwise convergence. Let f be the Dirichlet function, i.e., the characteristic function of the rationals $\mathbb{Q}$: $f(t)=1$ if $t\in\mathbb{Q}$ and $f(t)=0$ if $t\in\mathbb{R}\setminus\mathbb{Q}$. We set

$$f_j(t)=\lim_{m\to\infty}\bigl(\cos(j!\pi t)\bigr)^{2m}=\begin{cases}1 & \text{if } j!t\in\mathbb{Z},\\ 0 & \text{if } j!t\in\mathbb{R}\setminus\mathbb{Z},\end{cases}\qquad t\in\mathbb{R},\ j\in\mathbb{N},$$

where $\mathbb{Z}$ stands for the set of all integers. It is well known that f_j converges pointwise on $\mathbb{R}$ to f. Given an interval $[a,b]\subset\mathbb{R}$, any function f_j is equal to zero on $[a,b]$ outside a finite number of points, so it is proper and, according to Lemma 3, $\nu(n,f_j,[a,b])=o(n)$ for all $j\in\mathbb{N}$. Let $T=[0,1]$. Since the Jordan variation of f_j on $[0,1]$ is equal to $2\cdot j!$, we have:

$$\nu(n,f_j,T)=\begin{cases}n & \text{if } n<2\cdot j!,\\ 2\cdot j! & \text{if } n\geqslant 2\cdot j!,\end{cases}\qquad n,j\in\mathbb{N}.$$

Thus, $\lim_{j\to\infty}\nu(n,f_j,T)=n$. Note that $\nu(n,f,T)=n$, as well.

Example 6. Let $f_j\in\mathbb{R}^{[0,1]}$ be defined by $f_j(t)=f(t)/j$, $j\in\mathbb{N}$, where $f\in\mathbb{R}^{[0,1]}$ is the Dirichlet function on $[0,1]$. Clearly, $f_j\rightrightarrows 0$ on $[0,1]$ and $\nu(n,f_j,[0,1])=n/j\neq o(n)$ for all $j\in\mathbb{N}$. So, Theorem 1 is applicable to $\{f_j\}$, but none of the more classical selection principles applies (since $V_*(f_j,[0,1])$ is infinite for all $j\in\mathbb{N}$). Also, condition $\sup_{j\in\mathbb{N}}\nu(n,f_j,T)=o(n)$ is too restrictive as compared to condition (1).

Example 7. Let φ be a φ-function (not necessarily convex). We are going to show that if $\{f_j\}\subset\mathrm{BV}_\varphi(T;X)$ and $C=\sup_{j\in\mathbb{N}}V_\varphi(f_j,T)<\infty$, then $\sup_{j\in\mathbb{N}}\nu(n,f_j,T)=o(n)$, and so, condition (1) is satisfied in this general case as well. In particular, if $f\in\mathrm{BV}_\varphi(T;X)$, then $\nu(n,f,T)=o(n)$, i.e., f is proper by Lemma 3 (for $T=[a,b]$).

The function $\varphi_+^{-1}:\mathbb{R}^+\to\mathbb{R}^+$ defined by $\varphi_+^{-1}(r)=\sup\{\rho\in\mathbb{R}^+\mid\varphi(\rho)\leqslant r\}$, $r\in\mathbb{R}^+$, is called the *right inverse* of φ. Recall (cf. [22, Section 1.2]) that φ_+^{-1} is nondecreasing, continuous from the right, $\varphi_+^{-1}(r)=0$ if and only if $r=0$, and $\lim_{r\to\infty}\varphi_+^{-1}(r)=\infty$; moreover, the following relations hold: $\varphi(\varphi_+^{-1}(r))=r$ if $r\in\mathbb{R}^+$, $\varphi_+^{-1}(\varphi(\rho))\geqslant\rho$ if $\rho\in\mathbb{R}^+$, and $\varphi_+^{-1}(\varphi(\rho)-\varepsilon)\leqslant\rho$ if $\rho>0$ and $0<\varepsilon<\varphi(\rho)$. If, in addition, the φ-function φ is convex, then it is strictly increasing and its usual inverse φ^{-1} coincides with the right inverse φ_+^{-1}.

Let $n\in\mathbb{N}$ and $\{a_i\}_{i=1}^n$, $\{b_i\}_{i=1}^n\subset T$ be arbitrary such that $a_1\leqslant b_1\leqslant a_2\leqslant b_2\leqslant\cdots\leqslant a_n\leqslant b_n$. By the definition of V_φ, for $j\in\mathbb{N}$ and $i\in\{1,\dots,n\}$ we have

$$\varphi\bigl(d\bigl(f_j(b_i),f_j(a_i)\bigr)\bigr)\leqslant V_\varphi\bigl(f_j,[a_i,b_i]\cap T\bigr)\equiv c_i(j),$$

so that, taking the right inverse φ_+^{-1}, we get $d(f_j(b_i),f_j(a_i))\leqslant\varphi_+^{-1}(c_i(j))$. Summing over $i=1,\dots,n$, we find

$$\sum_{i=1}^n d\bigl(f_j(b_i),f_j(a_i)\bigr)\leqslant\sum_{i=1}^n\varphi_+^{-1}\bigl(c_i(j)\bigr),$$

where, by virtue of the semi-additivity of V_φ (e.g., [16, (P3)], [24, 1.17]),

$$\sum_{i=1}^n c_i(j)=\sum_{i=1}^n V_\varphi\bigl(f_j,[a_i,b_i]\cap T\bigr)\leqslant V_\varphi(f_j,T)\leqslant C,\qquad j\in\mathbb{N}.$$

Due to the arbitrariness of $\{a_i\}_{i=1}^n$ and $\{b_i\}_{i=1}^n$ and the definition of $\nu(n, f_j, T)$, the last two inequalities yield:

$$\nu(n, f_j, T) \leqslant \sup\left\{\sum_{i=1}^n \varphi_+^{-1}(c_i) \;\middle|\; \{c_i\}_{i=1}^n \subset \mathbb{R}^+ \text{ and } \sum_{i=1}^n c_i \leqslant C\right\}. \tag{8}$$

Denote by $\xi(n)$ the right-hand side in (8). Since $\xi(n)$ is independent of j, we have $\sup_{j\in\mathbb{N}} \nu(n, f_j, T) \leqslant \xi(n)$ for all $n \in \mathbb{N}$. Let us show that $\xi(n) = o(n)$.

Given $\varepsilon > 0$, the (right) continuity of φ_+^{-1} at 0 implies the existence of $r_0 = r_0(\varepsilon) > 0$ such that $\varphi_+^{-1}(r) \leqslant \varepsilon/2$ for all $0 \leqslant r \leqslant r_0$. Set $n_0 = n_0(\varepsilon) = [C/r_0] + 1$, where $[u] = \max\{k \in \mathbb{Z} \mid k \leqslant u\}$. Clearly, $n_0 \in \mathbb{N}$ and $n_0 > C/r_0$. Now, let $n \geqslant n_0$ and $\{c_i\}_{i=1}^n \subset \mathbb{R}^+$ be arbitrary such that $\sum_{i=1}^n c_i \leqslant C$. We denote by $I_1(n)$ the set of those $i \in \{1, \ldots, n\}$ for which $c_i \leqslant r_0$ and by $I_2(n)$ the set of those $i \in \{1, \ldots, n\}$ for which $c_i > r_0$ and note that the number of elements in $I_2(n)$ is $\leqslant n_0$. If $n \geqslant n_1(\varepsilon) = \max\{n_0, 2n_0\varphi_+^{-1}(C)/\varepsilon\}$, then

$$\begin{aligned}\sum_{i=1}^n \varphi_+^{-1}(c_i) &= \sum_{i\in I_1(n)} \varphi_+^{-1}(c_i) + \sum_{i\in I_2(n)} \varphi_+^{-1}(c_i) \leqslant \sum_{i\in I_1(n)} \frac{\varepsilon}{2} + \sum_{i\in I_2(n)} \varphi_+^{-1}(C)\\ &\leqslant n\frac{\varepsilon}{2} + n_0\varphi_+^{-1}(C) \leqslant n\frac{\varepsilon}{2} + n\frac{\varepsilon}{2} = n\varepsilon,\end{aligned}$$

and so, $\xi(n)/n \leqslant \varepsilon$ for all $n \geqslant n_1(\varepsilon)$, which was to be proved.

Given a φ-function φ, a function $f \in X^T$ is said to be of *generalized bounded φ-variation* (cf. [14,24]) if there exists a constant $\lambda > 0$ such that $V_{\varphi_\lambda}(f, T) < \infty$, where $\varphi_\lambda(\rho) = \varphi(\rho/\lambda)$, $\rho \in \mathbb{R}^+$. Theorem 1 and the above considerations imply the following result, which generalizes Theorem 1.3 from [24] and Theorem 1.3 from [13]: *If $\{f_j\} \subset X^T$ is a pointwise precompact sequence and there is a constant $\lambda > 0$ such that $\sup_{j\in\mathbb{N}} V_{\varphi_\lambda}(f_j, T) < \infty$, then a subsequence of $\{f_j\}$ converges pointwise on T to a function $f \in X^T$ satisfying $V_{\varphi_\lambda}(f, T) < \infty$.*

4. Almost everywhere convergence

Theorem 1 implies immediately that if $\{f_j\} \subset X^T$ is pointwise precompact and $\limsup_{j\to\infty} \nu(n, f_j, T \setminus E) = o(n)$ for some $E \subset T$ with $\mathcal{L}(E) = 0$, then a subsequence of $\{f_j\}$ converges a.e. on T to a function $f \in X^T$ such that $\nu(n, f, T \setminus E) = o(n)$.

The following theorem, which is a *selection principle for almost everywhere convergence* in terms of the modulus of variation, is more subtle and is subsequence-converse to Lemma 4(b).

Theorem 6. *Suppose $\emptyset \neq T \subset \mathbb{R}$, (X, d) is a metric space and $\{f_j\} \subset X^T$ is an almost everywhere (or pointwise) on T precompact sequence satisfying the condition: for each $\varepsilon > 0$ there exists a measurable set $E = E(\varepsilon) \subset T$ with $\mathcal{L}(E) \leqslant \varepsilon$ such that $\limsup_{j\to\infty} \nu(n, f_j, T \setminus E) = o(n)$. Then a subsequence of $\{f_j\}$ converges a.e. on T to a function $f \in X^T$ having the property: for each $\varepsilon > 0$ there exists a measurable set $E' = E'(\varepsilon) \subset T$ with $\mathcal{L}(E') \leqslant \varepsilon$ such that $\nu(n, f, T \setminus E') = o(n)$.*

Proof. Let $T_0 \subset T$ be a (possibly empty) set of measure zero such that the sequence $\{f_j(t)\}$ is precompact in X for all $t \in T \setminus T_0$. We employ Theorem 1 and the diagonal process. By the assumption, there exists a measurable set $E_1 \subset T$ with $\mathcal{L}(E_1) \leqslant 1$ such that $\limsup_{j\to\infty} \nu(n, f_j, T \setminus E_1) = o(n)$. The sequence $\{f_j\}$ is pointwise precompact on $T \setminus (T_0 \cup E_1)$ and, by Lemma 2(d),

$$\limsup_{j\to\infty} \nu\bigl(n, f_j, T \setminus (T_0 \cup E_1)\bigr) \leqslant \limsup_{j\to\infty} \nu(n, f_j, T \setminus E_1) = o(n).$$

Applying Theorem 1, we find a subsequence $\{f_j^{(1)}\}_{j=1}^{\infty}$ of $\{f_j\}$ and a function $f^1 : T \setminus (T_0 \cup E_1) \to X$ satisfying $\nu(n, f^1, T \setminus (T_0 \cup E_1)) = o(n)$ such that $f_j^{(1)} \to f^1$ on $T \setminus (T_0 \cup E_1)$. If $k \geqslant 2$ and a subsequence $\{f_j^{(k-1)}\}_{j=1}^{\infty}$ of $\{f_j\}$ is already chosen, there exists a measurable set $E_k \subset T$ with $\mathcal{L}(E_k) \leqslant 1/k$ such that $\limsup_{j\to\infty} \nu(n, f_j, T \setminus E_k) = o(n)$. The sequence $\{f_j^{(k-1)}\}_{j=1}^{\infty}$ is pointwise precompact on $T \setminus (T_0 \cup E_k)$ and, again by Lemma 2(d),

$$\begin{aligned}\limsup_{j\to\infty} \nu\bigl(n, f_j^{(k-1)}, T \setminus (T_0 \cup E_k)\bigr) &\leqslant \limsup_{j\to\infty} \nu\bigl(n, f_j^{(k-1)}, T \setminus E_k\bigr) \\ &\leqslant \limsup_{j\to\infty} \nu(n, f_j, T \setminus E_k) = o(n),\end{aligned}$$

and so, by Theorem 1, there exists a subsequence $\{f_j^{(k)}\}_{j=1}^{\infty}$ of $\{f_j^{(k-1)}\}_{j=1}^{\infty}$ and a function $f^k : T \setminus (T_0 \cup E_k) \to X$ satisfying $\nu(n, f^k, T \setminus (T_0 \cup E_k)) = o(n)$ such that $f_j^{(k)} \to f^k$ pointwise on $T \setminus (T_0 \cup E_k)$.

Setting $E = T_0 \cup \bigcap_{k=1}^{\infty} E_k$, we have: E is measurable, $\mathcal{L}(E) = 0$ and

$$T \setminus E = \bigcup_{k=1}^{\infty} \bigl(T \setminus (T_0 \cup E_k)\bigr).$$

Define the function $f : T \setminus E \to X$ as follows: given $t \in T \setminus E$, there exists $k \in \mathbb{N}$ such that $t \in T \setminus (T_0 \cup E_k)$, and so, we set $f(t) = f^k(t)$. The definition of f is correct, i.e., $f(t)$ is independent of k: in fact, if $k_1 \in \mathbb{N}$ and $t \in T \setminus (T_0 \cup E_{k_1})$, then $k \leqslant k_1$ (with no loss of generality), so that $\{f_j^{(k_1)}\}_{j=1}^{\infty}$ is a subsequence of $\{f_j^{(k)}\}_{j=1}^{\infty}$ and, therefore,

$$f^{k_1}(t) = \lim_{j\to\infty} f_j^{(k_1)}(t) = \lim_{j\to\infty} f_j^{(k)}(t) = f^k(t) \quad \text{in } X.$$

Let us show that the diagonal sequence $f_j^{(j)}$ (which, of course, is a subsequence of $\{f_j\}$) converges to f pointwise on $T \setminus E$. In fact, if $t \in T \setminus E$, then $t \in T \setminus (T_0 \cup E_k)$ for some $k \in \mathbb{N}$, and so, $f(t) = f^k(t)$. Since $\{f_j^{(j)}\}_{j=k}^{\infty}$ is a subsequence of $\{f_j^{(k)}\}_{j=1}^{\infty}$, we have:

$$\lim_{j\to\infty} f_j^{(j)}(t) = \lim_{j\to\infty} f_j^{(k)}(t) = f^k(t) = f(t) \quad \text{in } X.$$

We extend f arbitrarily from $T \setminus E$ to the whole T and denote this extension again by f. Given $\varepsilon > 0$, choose $k \in \mathbb{N}$ such that $1/k \leqslant \varepsilon$ and set $E' = E'(\varepsilon) = T_0 \cup E_k$. Then we have: $\mathcal{L}(E') = \mathcal{L}(E_k) \leqslant 1/k \leqslant \varepsilon$, $f = f^k$ on $T \setminus (T_0 \cup E_k) = T \setminus E'$ and $\nu(n, f, T \setminus E') = \nu(n, f^k, T \setminus (T_0 \cup E_k)) = o(n)$. $\square$

5. Weak pointwise selection principle

The aim of this Section is to prove a weak variant of Theorem 1 using some specific features when the values of functions under consideration lie in a Banach space (see Theorem 7 below).

Let $(X, \|\cdot\|)$ be a normed linear space (over the field $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$) and X^* be its *dual*, i.e., the space $L(X; \mathbb{K})$ of all continuous linear functionals on X. Recall that X^* is a Banach space under the norm $\|x^*\| = \sup\{|x^*(x)| \mid x \in X \text{ and } \|x\| \leqslant 1\}$, $x^* \in X^*$. The natural duality between X and X^* is determined by the bilinear functional $\langle\cdot,\cdot\rangle : X \times X^* \to \mathbb{K}$ defined by $\langle x, x^*\rangle = x^*(x)$, $x \in X$, $x^* \in X^*$. Recall also that if a sequence $\{x_j\} \subset X$ *converges weakly* in X to $x \in X$, in symbols, $x_j \xrightarrow{w} x$ in X (i.e., $\lim_{j\to\infty}\langle x_j, x^*\rangle = \langle x, x^*\rangle$ for all $x^* \in X^*$), then $\|x\| \leqslant \liminf_{j\to\infty} \|x_j\|$.

The notion of the modulus of variation $\nu(n, f, T)$ for $f \in X^T$ is introduced as in Section 1 with respect to the natural metric $d(x, y) = \|x - y\|$, $x, y \in X$.

Theorem 7. *Let $\emptyset \neq T \subset \mathbb{R}$ and $(X, \|\cdot\|)$ be a reflexive separable Banach space with separable dual X^*. Suppose the sequence $\{f_j\} \subset X^T$ is such that*

(i) $\sup_{j\in\mathbb{N}} \|f_j(t)\| < \infty$ *for all $t \in T$, and*
(ii) $\mu(n) \equiv \limsup_{j\to\infty} \nu(n, f_j, T) = o(n)$.

Then there exist a subsequence of $\{f_j\}$ (still denoted as the whole sequence) and a function $f \in X^T$ satisfying $\nu(n, f, T) \leqslant \mu(n)$ for all $n \in \mathbb{N}$ such that $f_j(t) \xrightarrow{w} f(t)$ in X for all $t \in T$.

Proof. (1) We set $C(t) = \sup_{j\in\mathbb{N}} \|f_j(t)\|$, $t \in T$. Given $j \in \mathbb{N}$ and $x^* \in X^*$, by virtue of (i) we have

$$\bigl|\bigl\langle f_j(t), x^*\bigr\rangle\bigr| \leqslant \bigl\|f_j(t)\bigr\| \cdot \|x^*\| \leqslant C(t)\|x^*\|, \quad t \in T, \tag{9}$$

and since $\nu(n, \langle f_j(\cdot), x^*\rangle, T) \leqslant \nu(n, f_j, T)\|x^*\|$, condition (ii) implies

$$\mu_{x^*}(n) \equiv \limsup_{j\to\infty} \nu\bigl(n, \bigl\langle f_j(\cdot), x^*\bigr\rangle, T\bigr) \leqslant \mu(n)\|x^*\|. \tag{10}$$

Applying Theorem 1 to the sequence $\{\langle f_j(\cdot), x^*\rangle\} \subset \mathbb{K}^T$ for any given $x^* \in X^*$, we find a subsequence $\{f_{j,x^*}\}$ of $\{f_j\}$ (generally depending on x^*) and a function $y_{x^*} \in \mathbb{K}^T$ satisfying $\nu(n, y_{x^*}, T) \leqslant \mu_{x^*}(n) \leqslant \mu(n)\|x^*\|$, $n \in \mathbb{N}$, such that $\langle f_{j,x^*}(t), x^*\rangle \to y_{x^*}(t)$ in $\mathbb{K}$ for all $t \in T$.

(2) Making use of the diagonal process, we will get rid of the dependence of $\{f_{j,x^*}\}$ on the element $x^* \in X^*$. Let $\{x_k^*\}_{k=1}^\infty$ be a countable dense subset of X^*. From step (1), for $x^* = x_1^*$ we get a subsequence $\{f_j^{(1)}\} = \{f_{j,x_1^*}\}$ of $\{f_j\}$ and a function $y_{x_1^*} \in \mathbb{K}^T$ satisfying $\nu(n, y_{x_1^*}, T) \leqslant \mu(n)\|x_1^*\|$ such that $\langle f_j^{(1)}(t), x_1^*\rangle \to y_{x_1^*}(t)$ in $\mathbb{K}$ for all $t \in T$. If $k \geqslant 2$ and a subsequence $\{f_j^{(k-1)}\}_{j=1}^\infty$ of $\{f_j\}$ is already chosen, by (9) and (10) we have:

$|\langle f_j^{(k-1)}(t), x_k^*\rangle| \leqslant C(t)\|x_k^*\|$ for all $j \in \mathbb{N}$ and $t \in T$ and

$$\limsup_{j\to\infty} \nu\bigl(n, \bigl\langle f_j^{(k-1)}(\cdot), x_k^*\bigr\rangle, T\bigr) \leqslant \limsup_{j\to\infty} \nu\bigl(n, \bigl\langle f_j(\cdot), x_k^*\bigr\rangle, T\bigr) \leqslant \mu(n)\bigl\|x_k^*\bigr\|,$$

and so, by Theorem 1, there exist a subsequence $\{f_j^{(k)}\}_{j=1}^\infty$ of $\{f_j^{(k-1)}\}_{j=1}^\infty$ and a function $y_{x_k^*} \in \mathbb{K}^T$ satisfying $\nu(n, y_{x_k^*}, T) \leqslant \mu(n)\|x_k^*\|$, $n \in \mathbb{N}$, such that $\langle f_j^{(k)}(t), x_k^*\rangle \to y_{x_k^*}(t)$ in $\mathbb{K}$ for all $t \in T$. Then the diagonal subsequence $\{f_j^{(j)}\}_{j=1}^\infty$, which we again denote by $\{f_j\}$, satisfies the condition:

$$\lim_{j\to\infty} \bigl\langle f_j(t), x_k^*\bigr\rangle = y_{x_k^*}(t) \quad \text{for all } t \in T \text{ and } k \in \mathbb{N}. \tag{11}$$

(3) If $x^* \in X^*$ is arbitrary and $t \in T$, let us show that $\{\langle f_j(t), x^*\rangle\} \subset \mathbb{K}$ is a Cauchy sequence. Given $\varepsilon > 0$, by the density of $\{x_k^*\}_{k=1}^\infty$ in X^*, there exists $k = k(\varepsilon) \in \mathbb{N}$ such that $\|x^* - x_k^*\| \leqslant \varepsilon/(4C(t)+1)$, and from (11) we find $j_0 = j_0(\varepsilon) \in \mathbb{N}$ such that $|\langle f_j(t), x_k^*\rangle - \langle f_{j'}(t), x_k^*\rangle| \leqslant \varepsilon/2$ for all $j, j' \geqslant j_0$. It follows that

$$\begin{aligned}
\bigl|\bigl\langle f_j(t), x^*\bigr\rangle - \bigl\langle f_{j'}(t), x^*\bigr\rangle\bigr| &\leqslant \bigl\|f_j(t) - f_{j'}(t)\bigr\| \cdot \bigl\|x^* - x_k^*\bigr\| \\
&\quad + \bigl|\bigl\langle f_j(t), x_k^*\bigr\rangle - \bigl\langle f_{j'}(t), x_k^*\bigr\rangle\bigr| \\
&\leqslant 2C(t)\bigl\|x^* - x_k^*\bigr\| + \bigl|\bigl\langle f_j(t), x_k^*\bigr\rangle - \bigl\langle f_{j'}(t), x_k^*\bigr\rangle\bigr| \\
&\leqslant 2C(t)\frac{\varepsilon}{4C(t)+1} + \frac{\varepsilon}{2} \leqslant \varepsilon, \quad j, j' \geqslant j_0.
\end{aligned}$$

Hence, there exists an element $y_{x^*}(t) \in \mathbb{K}$ such that $\langle f_j(t), x^*\rangle \to y_{x^*}(t)$ in $\mathbb{K}$. In other words, we have shown that for each $x^* \in X^*$ there exists a function $y_{x^*} \in \mathbb{K}^T$ satisfying (cf. Lemma 2(e) and (10))

$$\nu(n, y_{x^*}, T) \leqslant \liminf_{j\to\infty} \nu\bigl(n, \bigl\langle f_j(\cdot), x^*\bigr\rangle, T\bigr) \leqslant \mu(n)\|x^*\|, \quad n \in \mathbb{N},$$

such that

$$\lim_{j\to\infty} \bigl\langle f_j(t), x^*\bigr\rangle = y_{x^*}(t) \quad \text{in } \mathbb{K} \text{ for all } t \in T \text{ and } x^* \in X^*. \tag{12}$$

(4) Given $t \in T$, let us show that $f_j(t)$ converges weakly in X. By the reflexivity of X, we have $f_j(t) \in X = X^{**} = L(X^*; \mathbb{K})$ for all $j \in \mathbb{N}$. Defining the functional $Y_t : X^* \to \mathbb{K}$ by $Y_t(x^*) = y_{x^*}(t)$, $x^* \in X^*$, we get from (12):

$$\lim_{j\to\infty} \bigl\langle f_j(t), x^*\bigr\rangle = y_{x^*}(t) = Y_t(x^*) \quad \text{for all } x^* \in X^*,$$

i.e., the sequence $\{f_j(t)\} \subset L(X^*; \mathbb{K})$ converges pointwise on X^* to the operator $Y_t : X^* \to \mathbb{K}$. By the uniform boundedness principle, $Y_t \in L(X^*; \mathbb{K}) = X$ and $\|Y_t\| \leqslant \liminf_{j\to\infty} \|f_j(t)\|$. Setting $f(t) = Y_t$, $t \in T$, we find that $f \in X^T$ and

$$\lim_{j\to\infty} \bigl\langle f_j(t), x^*\bigr\rangle = Y_t(x^*) = \langle Y_t, x^*\rangle = \bigl\langle f(t), x^*\bigr\rangle, \quad x^* \in X^*,\ t \in T, \tag{13}$$

that is, $f_j(t) \xrightarrow{w} f(t)$ in X for all $t \in T$.

(5) It remains to prove that $\nu(n, f, T) \leqslant \mu(n)$, $n \in \mathbb{N}$. Given $a, b \in T$, condition (13) yields $f_j(b) - f_j(a) \xrightarrow{w} f(b) - f(a)$, and so, $\|f(b) - f(a)\| \leqslant \liminf_{j\to\infty} \|f_j(b) - f_j(a)\|$. Now, if $n \in \mathbb{N}$ and $\{a_i\}_{i=1}^n$, $\{b_i\}_{i=1}^n \subset T$ are arbitrary such that $a_1 \leqslant b_1 \leqslant a_2 \leqslant b_2 \leqslant \cdots \leqslant a_n \leqslant b_n$, then, by (ii),

$$\begin{aligned}\sum_{i=1}^n \big\| f(b_i) - f(a_i)\big\| &\leqslant \sum_{i=1}^n \liminf_{j\to\infty} \big\| f_j(b_i) - f_j(a_i)\big\| \\ &\leqslant \liminf_{j\to\infty} \sum_{i=1}^n \big\| f_j(b_i) - f_j(a_i)\big\| \\ &\leqslant \liminf_{j\to\infty} \nu(n, f_j, T) \leqslant \mu(n),\end{aligned}$$

and so, $\nu(n, f, T) \leqslant \mu(n)$, which completes the proof. □

Remarks.

(1) Condition (i) in Theorem 7 can be replaced by: $\sup_{j\in\mathbb{N}} \|f_j(t_0)\| \leqslant C_0$ for some $t_0 \in T$ and $C_0 \in \mathbb{R}^+$. In fact, since $\limsup_{j\to\infty} \nu(1, f_j, T) = \mu(1)$, we have $\sup_{j\in\mathbb{N}} \nu(1, f_j, T) \leqslant C_1$ for some $C_1 \in \mathbb{R}^+$, and so, for any $j \in \mathbb{N}$ and $t \in T$,
$$\big\| f_j(t)\big\| \leqslant \big\| f_j(t) - f_j(t_0)\big\| + \big\| f_j(t_0)\big\| \leqslant \nu(1, f_j, T) + C_0 \leqslant C_1 + C_0.$$
(2) If in Theorem 7 instead of condition (i) we assume that the sequence $\{f_j(t)\}$ is precompact in X for all $t \in T$, then, by Theorem 1, a subsequence of $\{f_j\}$ can be chosen to converge pointwise on T strongly in X. In this case X may be any normed linear space.
(3) In step (5) of the proof of Theorem 7 we have shown that if $f_j(t) \xrightarrow{w} f(t)$ in X for all $t \in T$, then $\nu(n, f, T) \leqslant \liminf_{j\to\infty} \nu(n, f_j, T)$, $n \in \mathbb{N}$.
(4) If in Theorem 7 condition (ii) is replaced by $\sup_{j\in\mathbb{N}} V_\varphi(f_j, T) < \infty$, then the weak limit function f will belong to $\mathrm{BV}_\varphi(T; X)$. To see this, it suffices to apply arguments similar to step (5) in the proof of Theorem 7 and note that if $a, b \in T$, then $\varphi(\|f(b) - f(a)\|) \leqslant \liminf_{j\to\infty} \varphi(\|f_j(b) - f_j(a)\|)$. In this case Theorem 7 with $\varphi(\rho) = \rho$ and $T = [a, b]$ gives a result from [2, Chapter 1, Theorem 3.5].

A similar conclusion holds if (ii) is replaced by $\sup_{j\in\mathbb{N}} V_\Lambda(f_j, T) < \infty$.

The following theorem can be proved along the same lines as Theorem 6 by applying Theorem 7 instead of Theorem 1.

Theorem 8. *Let $\emptyset \neq T \subset \mathbb{R}$ and $(X, \|\cdot\|)$ be a reflexive separable Banach space with separable dual X^*. Suppose that $\{f_j\} \subset X^T$ satisfies the conditions*:

(i) $\sup_{j\in\mathbb{N}} \|f_j(t)\| < \infty$ *for almost all $t \in T$, and*
(ii) *for each $\varepsilon > 0$ there exists a measurable set $E = E(\varepsilon) \subset T$ with $\mathcal{L}(E) \leqslant \varepsilon$ such that $\limsup_{j\to\infty} \nu(n, f_j, T \setminus E) = o(n)$.*

Then there exists a subsequence of $\{f_j\}$ (*denoted as the whole sequence*) *such that* $f_j(t) \xrightarrow{w} f(t)$ *in* X *for almost all* $t \in T$, *where* $f \in X^T$ *is a function with the property*: *for each* $\varepsilon > 0$ *there exists a measurable set* $E' = E'(\varepsilon) \subset T$ *with* $\mathcal{L}(E') \leqslant \varepsilon$ *such that* $\nu(n, f, T \setminus E') = o(n)$.

Acknowledgments

The author is grateful to Patrick J. Muldowney (Magee College, University of Ulster, Londonderry, Northern Ireland) for stimulating discussions on the results of this paper. My visit to the School of International Business (February 11–18, 2004) was made possible by Networking Grant 226 (Title of Project: Financial Risk Management in Energy Markets), InvestNI and University of Ulster.

References

[1] J. Banaś, W.G. El-Sayed, Functions of generalized bounded variation, Zeszyty Nauk. Politech. Rzeszowskiej Mat. Fiz. 85 (1991) 91–109.
[2] V. Barbu, Th. Precupanu, Convexity and Optimization in Banach Spaces, second ed., Reidel, Dordrecht, 1986.
[3] S.A. Belov, V.V. Chistyakov, A selection principle for mappings of bounded variation, J. Math. Anal. Appl. 249 (2000) 351–366; and comments on this paper: J. Math. Anal. Appl. 278 (2003) 250–251.
[4] Z.A. Chanturiya, The modulus of variation of a function and its application in the theory of Fourier series, Dokl. Akad. Nauk SSSR 214 (1974) 63–66 (in Russian). English translation: Soviet Math. Dokl. 15 (1974) 67–71.
[5] Z.A. Chanturiya, Absolute convergence of Fourier series, Mat. Zametki 18 (1975) 185–192 (in Russian). English translation: Math. Notes 18 (1975) 695–700.
[6] V.V. Chistyakov, On mappings of bounded variation, J. Dynam. Control Systems 3 (1997) 261–289.
[7] V.V. Chistyakov, On the theory of multivalued mappings of bounded variation of one real variable, Mat. Sb. 189 (1998) 153–176 (in Russian). English translation: Sb. Math. 189 (1998) 797–819.
[8] V.V. Chistyakov, Mappings of bounded variation with values in a metric space: Generalizations, in: Pontryagin Conference, 2, Nonsmooth Analysis and Optimization, Moscow, 1998, J. Math. Sci. (New York) 100 (2000) 2700–2715.
[9] V.V. Chistyakov, Generalized variation of mappings with applications to composition operators and multifunctions, Positivity 5 (2001) 323–358.
[10] V.V. Chistyakov, Metric space-valued mappings of bounded variation, in: Functional Analysis, 8, J. Math. Sci. (New York) 111 (2002) 3387–3429.
[11] V.V. Chistyakov, Mappings of generalized variation and composition operators, in: Dynamical systems, 10, J. Math. Sci. (New York) 110 (2002) 2455–2466.
[12] V.V. Chistyakov, On multi-valued mappings of finite generalized variation, Mat. Zametki 71 (2002) 611–632 (in Russian). English translation: Math. Notes 71 (2002) 556–575.
[13] V.V. Chistyakov, Selections of bounded variation, J. Appl. Anal. 10 (2004) 1–82.
[14] V.V. Chistyakov, Lipschitzian Nemytskii operators in the cones of mappings of bounded Wiener φ-variation, Folia Math. 11 (2004) 15–39.
[15] V.V. Chistyakov, O.E. Galkin, On maps of bounded p-variation with $p > 1$, Positivity 2 (1998) 19–45.
[16] V.V. Chistyakov, O.E. Galkin, Mappings of bounded Φ-variation with arbitrary function Φ, J. Dynam. Control Systems 4 (1998) 217–247.
[17] J. Dieudonné, Foundations of Modern Analysis, Academic Press, New York, 1960.
[18] S. Fuchino, Sz. Plewik, On a theorem of E. Helly, Proc. Amer. Math. Soc. 127 (1999) 491–497.
[19] C. Goffman, T. Nishiura, D. Waterman, Homeomorphisms in Analysis, Math. Surveys Monogr., vol. 54, Amer. Math. Soc., Providence, RI, 1997.

[20] E. Helly, Über lineare Funktionaloperationen, Sitzungsber. Naturwiss. Kl. Kaiserlichen Akad. Wiss. Wien 121 (1912) 265–297.
[21] T.H. Hildebrandt, Introduction to the Theory of Integration, Academic Press, New York, 1963.
[22] M.A. Krasnosel'skii, Ya.B. Rutickii, Convex Functions and Orlicz Spaces, Fizmatgiz, Moscow, 1958 (in Russian). English translation: Noordhoff, Groningen, The Netherlands, 1961.
[23] R. Leśniewicz, W. Orlicz, On generalized variations (II), Studia Math. 45 (1973) 71–109.
[24] J. Musielak, W. Orlicz, On generalized variations (I), Studia Math. 18 (1959) 11–41.
[25] I.P. Natanson, Theory of Functions of a Real Variable, third ed., Nauka, Moscow, 1974 (in Russian). English translation: Ungar, New York, 1965.
[26] M. Nicolescu, Analiză matematică, vol. II, III, Editura tehnică, Bucureşti, 1958, 1960.
[27] M.M. Rao, Measure Theory and Integration, Wiley, New York, 1987.
[28] D. Waterman, On convergence of Fourier series of functions of generalized bounded variation, Studia Math. 44 (1972) 107–117.
[29] D. Waterman, On Λ-bounded variation, Studia Math. 57 (1976) 33–45.