A uniform quantitative form of sequential weak compactness and Baillon's nonlinear ergodic theorem

Ulrich Kohlenbach

July 27, 2011

Abstract

We apply proof-theoretic techniques of 'proof mining' to obtain an effective uniform rate of metastability in the sense of Tao for Baillon's famous nonlinear ergodic theorem in Hilbert space. In fact, we analyze a proof due to Brézis and Browder of Baillon's theorem relative to the use of weak sequential compactness. Using previous results due to the author we show the existence of a bar recursive functional Ω^* (using only lowest type bar recursion $B_{0,1}$) providing a uniform quantitative version of weak compactness. Primitive recursively in this functional (and hence in $T_0 + B_{0,1}$) we then construct an explicit bound φ on for the metastable version of Baillon's theorem. From the type level of φ and another result of the author it follows that φ is primitive recursive in the extended sense of Gödel's T. In a subsequent paper also Ω^* will be explicitly constructed leading to the refined complexity estimate $\varphi \in T_4$.

Keywords. Nonlinear ergodic theorem, effective bounds, proof mining, metastability

1 Introduction

The direct proof-theoretic analysis (e.g. by functional interpretation) of proofs based on sequential compactness with the aim to extract new information (both effective bounds as well as new qualitative uniformity results; see [12] for the general program) is notoriously difficult as already the functional interpretation of (even simple versions) of the Bolzano-Weierstraß principle is a very intricate bar recursive construction (see [20]). Often one, therefore, applies additional proof-theoretic pre-processing to the the given proof

Ulrich Kohlenbach: Department of Mathematics, Technische Universität Darmstadt, Schlossgartenstraße 7, 64289 Darmstadt, Germany; e-mail: kohlenbach@mathematik.tu-darmstadt.de

Mathematics Subject Classification (2010): 03F10; 47H25

to eliminate the use of sequential compactness by some more arithmetical reasoning (see e.g. [12] for background information).
For proofs in abstract functional analysis that even use weak sequential compactness arguments (in the context of arbitrary – not assumed to be separable – Hilbert spaces) things are even more complicated. Only recently, we verified ([13]) that in such an abstract context the formal theories, to which logical metatheorems on the extraction of uniform bounds from [11, 6, 12] apply, suffice to establish the basic weak compactness principles for Hilbert space. As a consequence of the formalizability of weak compactness, we draw in section 2 some results on the extractability of an effective functional Ω^* solving a uniform quantitative version of weak compactness (namely its monotone variant of Gödel's functional interpretation in the sense of [12]).

The reward of studying the formalizability of proofs based on substantial uses of weak compactness is that for those proofs the above mentioned logical metatheorems guarantee a-priori the existence (together with an extraction algorithm based on monotone functional interpretation for the construction) of effective and surprisingly uniform bounds on reformulations of ordinary convergence statements (that in general – due to their intrinsic ineffective nature – exclude the possibility of such bounds): these reformulations go back to G. Kreisel's so-called no-counterexample interpretation ([18, 19]) and have in recent years been re-discovered by T. Tao under the name of 'metastability' (see [23, 24] and – for proof theoretic applications to proofs that do not yet use weak compactness – [1, 16, 17] among others).

While already these applications e.g. in [1, 16, 17] were only found using the underlying proof-theoretic machinery at least as a guiding principle, the complexity of analyzing proofs based on weak compactness is so significant that here even to find ineffective uniform bounds seems to be virtually impossible without substantial use of proof theory.

The first step towards extracting uniform metastability bounds from proofs based on weak compactness was done in [14] where proofs of two well-known theorems on the strong convergence of a certain iteration schemes for nonexpansive mappings in Hilbert space due to F.E. Browder [5] and R. Wittmann [25] resp. are analyzed. Although these proofs rely on a weak sequential compactness argument,[3] the logical analysis in the end allows one in both cases to bypass this due to the logical structure of the proofs. That paper also outlines the way one would have to follow in more general cases.

Such a case (where the use of weak compactness might be indispensable) seems to be the famous nonlinear ergodic theorem due to Baillon [2] for nonexpansive (but nonlinear) operators in Hilbert space which in itself is a weak convergence result:

[3]The paper also analyzes a different proof of Browder's theorem, due to Halpern, that already avoids any use of weak compactness.

Theorem 1.1 (Baillon [2]). *Let $C \subset X$ be a nonempty bounded closed convex subset of a Hilbert space X and let $T : C \to C$ be a nonexpansive mapping. Then for every $x \in C$, the sequence (y_n) of Cesàro averages*

$$y_n := \frac{1}{n+1} \sum_{k=0}^{n} T^k(x)$$

converges weakly to a fixed point of T.

Already as an a priori consequence of the formalizability of proofs of Baillon's theorem in suitable formal systems, the general logical metatheorems from [11, 12] guarantee that there is an effective (and – as we will see in section 2 below – even primitive recursive in the sense of Gödel's calculus T, see [7, 8]) bound $\varphi(\varepsilon, h, b)$ such that the following holds (where $\mathbb{N}^* := \{1, 2, 3, \ldots\}$):

Theorem 1.2 (Metastable version of Baillon's theorem). *Let X be a Hilbert space and $C \subset X$ be a nonempty bounded, closed and convex subset. Let $\mathbb{N}^* \ni b \geq \|x\|$ for all $x \in C$. For $\varepsilon > 0$, $h : \mathbb{N} \to \mathbb{N}$, $w \in B_1(0)$ and $T : C \to C$ be nonexpansive, the following holds*

$$\exists l \leq \varphi(\varepsilon, h, b) \left(|\langle y_l - y_{\tilde{h}(l)}, w \rangle| < \varepsilon \right),$$

where $\tilde{h}(n) := \max\{h(n), n\}$ and (y_n) as is above.

Note that this provides a highly uniform strengthening of

$$(*) \; \forall \varepsilon > 0 \, \forall h : \mathbb{N} \to \mathbb{N} \, \forall w \in B_1(0) \, \exists l \in \mathbb{N} \left(|\langle y_l - y_{\tilde{h}(l)}, w \rangle| < \varepsilon \right)$$

as the bound does not depend on X, x_0, T or w and on C only via b. As a consequence, instead of the infinite sequence (y_n) involved in $(*)$, theorem 1.2 only refers to finite initial segments of (y_n) and so can be viewed as a finitization of $(*)$ in the sense of Tao [23]. $(*)$, however, is (ineffectively) equivalent to the weak Cauchy property of (y_n) and hence (as X is a Hilbert and so a reflexive space) to the weak convergence of (y_n) which is the essence of Baillon's theorem. In this sense, theorem 1.2 is a finitization of Baillon's theorem itself.

The theorem implies the only seemingly stronger version

$$\exists l \leq \varphi(\varepsilon/2, (h^+)^M, b) \, \forall i, j \in [l; l + h(l)] \left(|\langle y_i - y_j, w \rangle| < \varepsilon \right),$$

where $(h^+)^M(n) := \max_{i \leq n}\{i + h(i)\}$ and $[l; l + h(l)] := \{l, l + 1, \ldots, l + h(l)\}$ (see corollary 3.8) which literally corresponds to the formulation of metastability as used by Tao (adapted to the case at hand).

In the final and main section of the present paper we analyze (following the strategy sketched already in section 4 of [14]) a proof of Baillon's nonlinear ergodic theorem due to Brézis and Browder [4] which is particularly easy to formalize in systems of the type used in [12].

While the actual description of the above mentioned functional Ω^* is devoted to a separate paper ([15]), we in section 3 below explicitly construct the bound φ using Ω^* as a given operator by analyzing the proof of Baillon's theorem relative to its use of weak sequential compactness.

2 The monotone functional interpretation of weak sequential compactness in Hilbert space

Throughout this paper, X will be a (real) Hilbert space and $B_1(0)$ the closed unit ball in X.

The well-known fact that every sequence (x_n) in $B_1(0)$ has a weak cluster point in $B_1(0)$ can be formalized as follows (in the language of $\mathcal{T} := \widehat{\mathrm{PA}}^{\omega}\!\upharpoonright + \text{QF-AC}+\text{CA}^0_{ar})[X, \langle\cdot,\cdot\rangle, \mathcal{C}]$, see [12] for details), where $(x_n) \subset B_1(0)$ means $\forall n^{\mathbb{N}}(\|x_n\|_X \leq_{\mathbb{R}} 1)$ and $\forall x^X$ resp. $\exists x^X$ are more conveniently written as $\forall x \in X$ resp. $\exists x \in X$ and similarly for $\mathbb{N}$)

$$(1)\ \forall (x_n) \subset B_1(0)\, \exists v \in X\, \forall w \in X \forall k \in \mathbb{N}\, \exists n \geq k\, (|\langle \tilde{v} - x_n, w\rangle| <_{\mathbb{R}} 2^{-k}),$$

where

$$\tilde{v} := \frac{v}{\max\{\|v\|, 1\}}.$$

Remark 2.1. The formulation (1) is slightly weaker than then more usual definition of 'weak cluster point', where the existence of a common $n \geq k$ for a finite set of vectors $w_1, \ldots, w_m$ is stated. However, for our quantitative interpretation of Baillon's theorem the (functional interpretation of) the version (1) is sufficient. The construction of the solution functional Ω^* of the functional interpretation of (1) from [15] referred to after Theorem 2.2 below also adapts to the more general case. In fact, as our formal framework proves the existence of a weakly convergent subsequence of (x_n) (see [13]) even that (strongest) form of weak compactness has a solution functional.

By QF-AC, (1) is equivalent (over $\mathcal{T}$) to

$$(2)\ \forall (x_n) \subset B_1(0)\, \exists v \in X\, \exists \chi\, \forall w \in X \forall k \in \mathbb{N}\, \exists n \in [k, \chi(w,k)]\, (|\langle \tilde{v} - x_n, w\rangle| <_{\mathbb{R}} 2^{-k}),$$

where χ has the type $X \to (\mathbb{N} \to \mathbb{N})$, i.e. represents a function $X \times \mathbb{N} \to \mathbb{N}$.

(2) in turn implies (and over $\mathcal{T}$ actually is equivalent to)[4]

$$(3)\ \left\{ \begin{array}{l} \forall (x_n) \subset B_1(0)\, \forall W, K\, \exists v \in X\, \exists \chi \exists n \in [K(v,\chi), \chi(W(v,\chi), K(v,\chi))] \\ \hfill \big(|\langle \tilde{v} - x_n, W(v,\chi)\rangle| <_{\mathbb{R}} 2^{-K(v,\chi)}\big), \end{array} \right.$$

where (switching for convenience implicitly to product types)

$$W : X \times (X \times \mathbb{N} \to \mathbb{N}) \to X \text{ and } K : X \times (X \times \mathbb{N} \to \mathbb{N}) \to \mathbb{N}.$$

Applying (3) to

$$\tilde{W}(v,\chi) := W(\tilde{v},\chi) \text{ and } \tilde{K}(v,\chi) := K(\tilde{v},\chi)$$

yields

$$(4)\ \left\{ \begin{array}{l} \forall (x_n) \subset B_1(0)\, \forall W, K\, \exists v \in X\, \exists \chi \exists n \in [K(\tilde{v},\chi), \chi(W(\tilde{v},\chi), K(\tilde{v},\chi))] \\ \hfill \big(|\langle \tilde{v} - x_n, W(\tilde{v},\chi)\rangle| <_{\mathbb{R}} 2^{-K(\tilde{v},\chi)}\big), \end{array} \right.$$

which in turn implies back (3) (take $v := \tilde{v}$ for v as in (4) and observe that $\tilde{v} = \tilde{\tilde{v}}$).
In the following T_0 denotes the fragment of Gödel's [7] system T of primitive recursive functionals of all finite types that is defined only using ordinary primitive recursion of type $\mathbb{N}$ (and hence defines primitive recursive functionals in the ordinary sense of Kleene). $T_0 + B_{0,1}$ denotes the functionals that are definable using such ordinary primitive recursion by means of Spector's [21] operator $B_{0,1}$ for bar recursion of type $\mathbb{N}$ (see [12] for details on all this). Functionals definable in $T_0 + B_{0,1}$ do not define total functionals in the full set-theoretic model $\mathcal{S}^\omega$ over $\mathbb{N}$ but only in the model of strongly majorizable functionals $\mathcal{M}^\omega$ due to [3] (again see [12] for details). Nevertheless, functionals of type level ≤ 2 (i.e. functionals taking only numbers and number theoretic functions as arguments and numbers as values) in $T_0 + B_{0,1}$ do define total functionals (e.g. of type $\mathbb{N}^{\mathbb{N}} \times \mathbb{N} \to \mathbb{N}$) and – as shown in [10] – define exactly those functionals (of the respective type) that are primitive recursive in the sense of Gödel's T.
In the next theorem $\mathcal{M}^{\omega,X}$ denotes the extension of Bezem's model $\mathcal{M}^\omega$ to all finite types (i.e. all function spaces) over $\mathbb{N}$ and X from [6, 12].

As shown in [13], (1) – and hence (4) – is provable in $\mathcal{T}$. In fact, $\mathcal{T}$ proves that any sequence in $B_1(0)$ contains a weakly convergent subsequence. Using the fact that $\mathcal{T}$ has a monotone functional interpretation (combined with negative translation, i.e. – using the terminology from [12] – NMD) by terms in $T_0 + B_{0,1}$ (see [12]) the following theorem holds:

Theorem 2.2. *(Uniform quantitative version of weak sequential compactness) Applying monotone functional interpretation to the proof of* (1) *from [13] yields the extractability*

[4]Note that we can replace $<_{\mathbb{R}}$ by $\leq_{\mathbb{R}}$ which is in Π^0_1.

of a closed term Ω^* in $T_0 + B_{0,1}$ such that (denoting the strong majorizability relation $\gtrsim^{0_X}$ – with the zero vector 0_X as reference point – from [6, 12] by $\gtrsim$) the following is true in the model of all strongly majorizable functionals $\mathcal{M}^{\omega,X}$

$$\begin{cases} \exists \Omega \lesssim \Omega^* \, \forall K, W \, \forall (x_n) \subset B_1(0) \\ \qquad \exists v \in X \, \exists \chi = \Omega(K, W, (x_n)) \, \exists n \in [K(\tilde{v}, \chi), \chi(W(\tilde{v}, \chi), K(\tilde{v}, \chi))] \\ \qquad\qquad \big(|\langle \tilde{v} - x_n, W(\tilde{v}, \chi)\rangle| <_{\mathbb{R}} 2^{-K(\tilde{v},\chi)}\big). \end{cases}$$

Since our final bound φ on the metastable form of Baillon's theorem is of type two it follows from the above (i.e. the fact that the proof of Baillon's theorem formalizes in $\mathcal{T}$ which has a monotone functional interpretation by terms in $T_0 + B_{0,1}$ that only define functionals of type 2 that are already definable in T) that $\varphi \in T$. The detailed construction of Ω^* carried out (for the case $C := B_1(0)$) in [15] together with the structure of the bound φ constructed in the present paper even yields that the bound is definable at the 4-th level T_4 of T.

The statement in theorem 2.2 is equivalent to

$$(6) \quad \begin{cases} \exists \Omega \lesssim \Omega^* \, \forall K, W \, \forall (x_n) \subset B_1(0) \\ \qquad \exists v \in B_1(0) \, \exists \chi = \Omega(K, W, (x_n)) \, \exists n \in [K(v, \chi), \chi(W(v, \chi), K(v, \chi))] \\ \qquad\qquad \big(|\langle v - x_n, W(v, \chi)\rangle| <_{\mathbb{R}} 2^{-K(v,\chi)}\big). \end{cases}$$

By bounded choice b-AC (which holds in $\mathcal{M}^{\omega,X}$, see [12]), (6) implies

Corollary 2.3. *The following is true in the model of all strongly majorizable functionals $\mathcal{M}^{\omega,X}$:*

$$\begin{cases} \exists \tilde{\Omega} \lesssim \Omega^* \, \forall K \, \forall l \in \mathbb{N} \, \forall (x_n) \subset B_1(0) \\ \qquad \exists v \in B_1(0) \, \exists \chi = \tilde{\Omega}(K, l, (x_n)) \, \forall w \in B_l(0) \, \exists n \in [K(v, \chi), \chi(w, K(v, \chi))] \\ \qquad\qquad \big(|\langle v - x_n, w)\rangle| <_{\mathbb{R}} 2^{-K(v,\chi)}\big). \end{cases}$$

Proof. Given $K, l, (x_n)$ define $\tilde{\Omega}(K, l, (x_n)) := \Omega(K, W_{K,l,(x_n)}, (x_n))$ (with Ω from theorem 2.2), where

$$W_{K,l,(x_n)}(v, \chi) :=$$
$$\begin{cases} \text{some } w \in B_l(0) \text{ s.t. } \neg\exists n \in [K(v, \chi), \chi(w, K(v, \chi))] \, \big(|\langle v - x_n, w\rangle| < 2^{-K(v,\chi)}\big) \\ \hfill \text{if existent,} \\ 0_X, \text{ otherwise.} \end{cases}$$

$W_{K,l,(x_n)}(v, \chi)$ exists (as a functional in $K, l, (x_n), v$ and χ) by the axiom of bounded choice. Then – by theorem 2.2 – we get that (for $\chi := \tilde{\Omega}(K, l, (x_n))$)

$$\exists v \in B_1(0) \, \exists n \in [K(v, \chi), \chi(W_{K,l,(x_n)}(v, \chi), K(v, \chi))]$$
$$\big(|\langle v - x_n, W_{K,l,(x_n)}(v, \chi)\rangle| <_{\mathbb{R}} 2^{-K(v,\chi)}\big).$$

and so – by the definition of $W_{K,l,(x_n)}$ –

$$\exists v \in B_1(0)\, \forall w \in B_l(0)\, \exists n \in [K(v,\chi), \chi(w, K(v,\chi))] \left(|\langle v - x_n, w \rangle| <_{\mathbb{R}} 2^{-K(v,\chi)} \right).$$

Moreover (using that $W_{K,l,(x_n)}$ is majorized by a suitable constant-l functional which we also denote by l)[5],

$$\forall K^*, K, l^*, l\, \forall (x_n) \subset B_1(0)\ (K^* \gtrsim K \wedge l^* \geq l \to \Omega^*(K^*, l^*) \gtrsim \tilde{\Omega}(K, l, (x_n)),$$

which implies that $\Omega^* \gtrsim \tilde{\Omega}$. □

In the above corollary, one can easily also allow the sequence (x_n) to be in $B_b(0)$ (for some $b \in \mathbb{N}^*$) instead of $B_1(0)$ if also v is allowed to be in $B_b(0)$: just applying the corollary to $x_n^b := x_n/b \in B_1(0)$ and $K^b(v,\chi) := K(v,\chi) + \lceil log(b) \rceil$ yields the result with $\tilde{\Omega}(K, b, l, (x_n)) := \tilde{\Omega}(K^b, l, (x_n^b))$. In our application in the next section we have $l = 2b$ and so from now on we simply write $\tilde{\Omega}(K, b, (x_n))$ to denote $\tilde{\Omega}(K, b, 2b, (x_n))$. Somewhat more complicated is the issue whether we may allow an arbitrary nonempty b-bounded closed and convex subset $C \subset X$ instead of $B_b(0)$. This rests on the fact whether the proof of the so-called Mazur lemma (needed to show that v again is in C) can be formalized in the weakly extensional treatment of abstract convex sets from [12] (for an extensional schematic treatment of C this is verified in [13]). While it is open whether the statement

'Every weak limit v of a sequence (x_n) in C is also in C'

formalizes in the weakly extensional setting (probably not), one easily obtains from the treatment Mazur's proof in [13] that the version:

'Every weakly convergent sequence (x_n) in C possesses a weak limit $v \in C$'

is formalizable. Note the subtle point, that while the weak limit of course is unique w.r.t. $=_X$, the weakly extensional setting does not prove

$$v \in C \wedge w =_X v \to w \in C.$$

However, the second version suffices for the proof of the above corollary for an abstract b-bounded closed and convex subset C instead of $B_b(0)$ together with the extractability of a majorant Ω^* of an operator $\tilde{\Omega}(K, b, (x_n))$ so that $\Omega^*(K^*, b)$ majorizes $\tilde{\Omega}(K, b, (x_n))$ for (x_n) in C and K^* majorizing K (K^* is of type $\mathbb{N} \times (\mathbb{N} \times \mathbb{N} \to \mathbb{N}) \to \mathbb{N}$), i.e. the majorant

[5]Officially, Ω^* has a 3rd argument $g : \mathbb{N} \to \mathbb{N}$ which is a potential majorant for (x_n). However, as we only consider sequences in $B_1(0)$ we can take this to be the constant-1 function.

does not depend on C except for the norm bound b. Instead of the previous construction $v \mapsto \tilde{v} \in B_1(0)$ one now uses $v \mapsto \tilde{v} \in C$, where

$$\tilde{v} := \begin{cases} v, \text{ if } \chi_C(v) = 0 \\ c_X, \text{ otherwise,} \end{cases}$$

where χ_C, c_X are the constants used in the formalization of abstract convex sets from [11] (see also [12]).

The combinatorial strength of a weak cluster point v of a sequence (x_n) stems from the fact that e.g. not only does exist for a given vector w and an error $\varepsilon > 0$ an index n_0 such that $|\langle v - x_{n_0}, w\rangle| < \varepsilon$ but – given another vector w' – also an index n_1 larger than some given function of v and n_0 such that $|\langle v - x_{n_1}, w'\rangle| < \varepsilon'$, where $\varepsilon' > 0$ maybe given as a function of v and n_0 (and ε). In fact, w' may (and typically will) depend on v (and maybe n_0) as a function $w' = T(v, n_0)$ etc. Of course, much more involved scenarios may happen as well (and do happen in our treatment of the proof of Baillon's theorem in the next section). Here we will as a motivating exercise show how the situation just described and stated as 'lemma 5.4' in [14] is covered by corollary 2.3 making use of a suitably chosen case distinction functional K provided that the functions involved are majorizable, i.e. exist in the model $\mathcal{M}^{\omega,X}$. This is a simpler version of a technical proof used in the next section (from now on we omit the subscript '$\mathbb{R}$' in '$\leq_{\mathbb{R}}$' etc. as it will always be clear from the context what the objects involved are):

Proposition 2.4. There exists a computable (in fact prim.-rec. in Ω^*) functional $\chi^* : \mathbb{N}^{\mathbb{N}} \times \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ such that

$$\begin{aligned} &\forall \varepsilon > 0\, \forall l \in \mathbb{N}^*\, \forall f : \mathbb{N} \to \mathbb{N}^*\, \forall (x_n) \subseteq B_1(0)\, \forall \varphi : X \times \mathbb{N} \to (0,1]\, \forall w \in B_l(0) \\ &\forall \tilde{T} : B_1(0) \times \mathbb{N} \to B_l(0)\ \Big(f \gtrsim \varphi \to \exists v \in B_1(0) \exists n_0 \leq \chi^*(f, \lceil 1/\varepsilon \rceil, l)\, \exists n_1 \geq \tfrac{2}{\varphi(v,n_0)} \\ &\qquad \Big(|\langle v - x_{n_0}, w\rangle| < \varepsilon \wedge |\langle v - x_{n_1}, \tilde{T}(v, n_0)\rangle| < \varphi(v, n_0)\Big)\Big), \end{aligned}$$

where $f \gtrsim \varphi :\equiv \forall j \in \mathbb{N}\, \forall v \in B_1(0)\, (\frac{1}{f(j)} \leq \varphi(v, j))$ for $f : \mathbb{N} \to \mathbb{N}^*$.

Proof. Fix $\varepsilon, l, f, (x_n), \varphi, w, \tilde{T}$ as in the proposition and assume that $f \gtrsim \varphi$. Define $K := K_{w,\varepsilon,\varphi,(x_n)} : X \times (X \times \mathbb{N} \to \mathbb{N}) \to \mathbb{N}$ by

$$K(v, \chi) := \begin{cases} j_\varepsilon := \lceil \log(1/\varepsilon) \rceil, \text{ if } \forall n \in [j_\varepsilon, \chi(w, j_\varepsilon)]\ \big(|\langle \tilde{v} - x_n, w\rangle| \geq 2^{-j_\varepsilon}\big) \\ \max\left(\left\lceil \frac{2}{\varphi^M(\tilde{v}, \chi(w, j_\varepsilon))} \right\rceil, \left\lceil -\log\left(\varphi^M(\tilde{v}, \chi(w, j_\varepsilon))\right) \right\rceil\right), \text{ otherwise,} \end{cases}$$

where $\varphi^M(v, m) := \min\{\varphi(v, n) : n \leq m\}$ and $\tilde{v} := \frac{v}{\max\{\|v\|, 1\}}$ as before.
Corollary 2.3 applied to $K, (x_n)$ and l yields the existence of $v \in B_1(0)$ and $\chi =$

$\tilde{\Omega}(K, l, (x_n))$ such that (reasoning in $\mathcal{M}^{\omega,X}$; note that by the construction of K^* below it is clear that K is majorizable)

$$(*)\ \forall z \in B_l(0)\, \exists n \in [K(v,\chi), \chi(z, K(v,\chi))]\ \left(|\langle v - x_n, z\rangle| < 2^{-K(v,\chi)}\right).$$

By the definition of K, $(*)$ applied to $z := w$ yields

$$\exists n_0 \in [j_\varepsilon, \chi(w, j_\varepsilon)]\ \left(|\langle v - x_{n_0}, w\rangle| < 2^{-j_\varepsilon} \leq \varepsilon\right)$$

and

$$K(v,\chi) = \max\left(\left\lceil \frac{2}{\varphi^M(v, \chi(w, j_\varepsilon))} \right\rceil, \left\lceil -\log\left(\varphi^M(v, \chi(w, j_\varepsilon))\right)\right\rceil\right).$$

Applying $(*)$ again (but this time to $w := \tilde{T}(v, n_0)$) gives (using that $n_0 \leq \chi(w, j_\varepsilon)$)

$$\exists n_1 \geq K(v,\chi) \geq \frac{2}{\varphi(v, n_0)}\ \left(|\langle v - x_{n_1}, \tilde{T}(v, n_0)\rangle| < 2^{-K(v,\chi)} \leq \varphi(v, n_0)\right).$$

Now define for $v^* \in \mathbb{N}$ and $\chi^* : \mathbb{N}^2 \to \mathbb{N}$

$$\begin{aligned}&K^*_{l,\varepsilon,f}(v^*, \chi^*) := K^*(v^*, \chi^*) := K^*(\chi^*) := \\ &\max\{j_\varepsilon, 2f(\chi^*(l, j_\varepsilon)), \lceil \log(f(\chi^*(l, j_\varepsilon)))\rceil\} = \max\{j_\varepsilon, 2f(\chi^*(l, j_\varepsilon))\}.\end{aligned}$$

Then $K^* \gtrsim K$ and so $\Omega^*(K^*, l) \gtrsim \tilde{\Omega}(K, l, (x_n)) = \chi$ (here '$\gtrsim$' for general finite type function spaces over $\mathbb{N}, X$ again is the relation denoted by '$\gtrsim^{0_X}$' in [12]).
It follows that $\chi^*(f, n, l) := \Omega^*(K^*_{l, \frac{1}{n+1}, f}, l)(l, j_\varepsilon) \geq \chi(w, j_\varepsilon)$ for $n \geq \lceil \frac{1}{\varepsilon} \rceil$. □

3 A quantitative finitization of Baillon's nonlinear ergodic theorem in Hilbert space

Let $C \subset X$ be a nonempty bounded, closed, convex subset of the Hilbert space X and $b > 0$ be such that $b \geq \|x\|$ for all $x \in C$.

Lemma 3.1. *Let (x_n) be a sequence in C and define $y_n := \frac{1}{n+1}\sum_{k=0}^{n} x_k \in C$.*

Let $\varepsilon > 0$ and $n_\varepsilon := \left\lceil \frac{2(n+1)b}{\varepsilon} \right\rceil - 1$ for $n \in \mathbb{N}$.
Then

$$\forall n \in \mathbb{N}\, \forall m \geq n_\varepsilon\, \exists u_m \in C_n^m\ \left(\|y_m - u_m\| \leq \varepsilon\right),$$

where $C_n^m := co\{x_n, x_{n+1}, \ldots, x_m\}$ is the convex closure of $\{x_n, x_{n+1}, \ldots, x_m\}$ for $m \geq n$.

Proof. For j with $n < j \leq m$ define

$$\lambda_j := \frac{1}{m+1} \text{ and } \lambda_n := 1 - \frac{m-n}{m+1} = \frac{n+1}{m+1}.$$

Then $\sum\limits_{j=n}^{m} \lambda_j = 1$ and so $\sum\limits_{j=n}^{m} \lambda_j x_j \in C_n^m$. Moreover,

$$\begin{aligned} &\|y_m - \sum_{j=n}^{m} \lambda_j x_j\| \leq \tfrac{1}{m+1} \sum_{k=0}^{n} \|x_k\| + \tfrac{n+1}{m+1}\|x_n\| \\ &\leq \tfrac{n+1}{m+1} b + \tfrac{n+1}{m+1} b = \tfrac{2(n+1)b}{m+1} \overset{m \geq n_\varepsilon}{\leq} \tfrac{2(n+1)b}{n_\varepsilon + 1} \leq \varepsilon. \end{aligned}$$

□

Lemma 3.2. *Let $(a_n), (b_n)$ be sequences in $[0, N]$, where $N > 0$.*
Define $\tilde{a}_n := \min\{a_i : i \leq n\}$ and - analogously - $\tilde{b}_n$. Then the following holds:

$$\forall \varepsilon > 0 \, \forall g : \mathbb{N} \to \mathbb{N} \, \exists n \leq \Phi(\varepsilon, g, N) \, \forall j \in [n; \tilde{g}(n)] \, \big(|\tilde{a}_n - \tilde{a}_j| < \varepsilon \wedge |\tilde{b}_n - \tilde{b}_j| < \varepsilon\big),$$

where $\tilde{g}(n) := \max\{n, g(n)\}$, $[n; m] := \{j \in \mathbb{N} : n \leq j \leq m\}$ and
$\Phi(\varepsilon, g, N) := \tilde{g}^{(\lceil \frac{2N}{\varepsilon} \rceil)}(0)$.

Proof. Since $(\tilde{a}_n)$ and $(\tilde{b}_n)$ are non-increasing sequences also (c_n) defined by $c_n := \tilde{a}_n + \tilde{b}_n \in [0, 2N]$ is non-increasing. By [12] (prop.2.27 and rem.2.29.1) we have

$$\forall \varepsilon > 0 \, \forall g : \mathbb{N} \to \mathbb{N} \, \exists n \leq \Phi(\varepsilon, g, N) \, \forall j \in [n; \tilde{g}(n)] \, \big(|c_n - c_j| < \varepsilon\big).$$

Because of $\tilde{a}_n - \tilde{a}_j, \tilde{b}_n - \tilde{b}_j \geq 0$ for all $j \geq n$ one has

$$|\tilde{a}_n - \tilde{a}_j| + |\tilde{b}_n - \tilde{b}_j| = |\tilde{a}_n - \tilde{a}_j + \tilde{b}_n - \tilde{b}_j| = |(\tilde{a}_n + \tilde{b}_n) - (\tilde{a}_j + \tilde{b}_j)| = |c_n - c_j|.$$

Hence the lemma follows. □

Lemma 3.3 (Quantitative demiclosedness principle, [14])**.** *Let (x_n) be a sequence in C, $v \in C$ and $T : C \to C$ be a nonexpansive mapping. For $\varepsilon > 0$ define $v_\varepsilon := v - \frac{\varepsilon^2}{16b^2}(Id - T)(v) \in C$. Then the following holds for all $\varepsilon > 0$ and all $j \in \mathbb{N}$:*

$$|\langle (Id - T)(v_\varepsilon), x_j - v \rangle| < \frac{\varepsilon^4}{96b^2} \wedge \|T(x_j) - x_j\| < \frac{\varepsilon^4}{96b^3} \to \|v - T(v)\| < \varepsilon.$$

Proof. For $C := B_1(0)$, this lemma is proved already in [14]. We include here for completeness the adaptation of this proof for general C : Define $\tilde{T} := Id - T$.

$$(1) \left\{ \begin{array}{l} 0 \leq \|v_\varepsilon - x_j\|^2 - \langle T(v_\varepsilon) - T(x_j), v_\varepsilon - x_j \rangle = \langle \tilde{T}(v_\varepsilon) - \tilde{T}(x_j), v_\varepsilon - x_j \rangle \\ \quad = \langle \tilde{T}(v_\varepsilon), v_\varepsilon \rangle - \langle \tilde{T}(x_j), v_\varepsilon \rangle - \langle \tilde{T}(x_\varepsilon), x_j \rangle + \langle \tilde{T}(x_j), v_j \rangle. \end{array} \right.$$

Since $\|x_j\|, \|v_\varepsilon\| \leq b$ we have that

$$b \cdot \|T(x_j) - x_j\| \geq |\langle \tilde{T}(x_j), x_j\rangle|, |\langle \tilde{T}(x_j), v_\varepsilon\rangle|.$$

Hence (1) and the assumption yield

$$\begin{array}{l} -\frac{\varepsilon^4}{32b^2} < \langle \tilde{T}(v_\varepsilon), v_\varepsilon - v\rangle = \langle \tilde{T}(v - \frac{\varepsilon^2}{16b^2}\tilde{T}(v)), -\frac{\varepsilon^2}{16b^2}\tilde{T}(v)\rangle \\ \quad = -\frac{\varepsilon^2}{16b^2}\langle \tilde{T}(v - \frac{\varepsilon^2}{16b^2}\tilde{T}(v)), \tilde{T}(v)\rangle. \end{array}$$

Thus

$$\frac{\varepsilon^2}{2} > \langle \tilde{T}(v - \frac{\varepsilon^2}{16b^2}\tilde{T}(v)), \tilde{T}(v)\rangle$$

and so (!)

$$\varepsilon^2 > \langle \tilde{T}(v), \tilde{T}(v)\rangle = \|\tilde{T}(v)\|^2, \text{ i.e. } \|\tilde{T}(v)\| < \varepsilon.$$

'!' holds since $\tilde{T} \in Lip(2)$ and $\|\tilde{T}(v)\| \leq \|v\| + \|T(v)\| \leq 2b$ and so

$$\|\tilde{T}(v) - \tilde{T}(v - \frac{\varepsilon^2}{16b^2}\tilde{T}(v))\| \leq 2\|\frac{\varepsilon^2}{16b^2}\tilde{T}(v)\| \leq \frac{\varepsilon^2}{4b},$$

which implies

$$|\langle \tilde{T}(v), \tilde{T}(v)\rangle - \langle \tilde{T}(v - \frac{\varepsilon^2}{16b^2}\tilde{T}(v)), \tilde{T}(v)\rangle| \leq \|\tilde{T}(v) - \tilde{T}(v - \frac{\varepsilon^2}{16b^2}\tilde{T}(v))\| \cdot \|\tilde{T}(v)\| \leq \frac{\varepsilon^2}{2}.$$

□

Lemma 3.4 (Asymptotic regularity)**.** *Let $T : C \to C$ be a nonexpansive mapping, $x_0 \in C$, $x_{n+1} := T(x_n)$ and $y_n := \frac{1}{n+1}\sum\limits_{k=0}^{n} x_k$. Then the following holds*

$$\forall \varepsilon > 0\, \forall n \geq \rho(\varepsilon)\ \big(\|T(y_n) - y_n\| < \varepsilon\big),$$

where $\rho(\varepsilon) := \lceil 4b^2/\varepsilon^2 \rceil$. We say that ρ is a rate of asymptotic regularity for (y_n).

Proof. From the proof of theorem 1 in [4] (for the special case where $a_{n,k} := \frac{1}{n+1}$ for $k \leq n$ and $a_{n,k} := 0$ for all $k > n$, where $k, n \in \mathbb{N}$, so that $\gamma_n := \sum_{k=0}^{\infty}(a_{n,k+1} - a_{n,k})^+ = 0$) it follows that

$$2\|y_n - T(y_n)\|^2 \leq \frac{2}{n+1} \cdot diam(C)^2 \leq \frac{8b^2}{n+1}.$$

Hence $\|y_n - T(y_n)\| \leq \frac{2b}{\sqrt{n+1}} < \varepsilon$ for all $n \geq \lceil 4b^2/\varepsilon^2 \rceil$. □

In the following, let T, x_n, y_n be as in the previous lemma and $\rho : \mathbb{R}^* \to \mathbb{N}$ be a rate of asymptotic regularity for (y_n). For $\varepsilon > 0$ and $n \in \mathbb{N}$ define $\tilde{\varepsilon} := \frac{\varepsilon^2}{144b}$ and $\delta_{\varepsilon,n} :=$

$\min\left\{1, \frac{\tilde{\varepsilon}}{(4b+1)(n+1)}\right\}$ (so that $(n+1)(4b\delta_{\varepsilon,n}+\delta^2_{\varepsilon,n}) \le \tilde{\varepsilon}$). Let $n_\varepsilon := \left\lceil \frac{2(n+1)b}{\varepsilon} \right\rceil$, $\widehat{n}_\varepsilon :=$ $\max\{n_{\tilde{\varepsilon}}, \lceil \log(2/(\tilde{\varepsilon}b))\rceil, \lceil \log(4/\varepsilon)\rceil\}$ and $\tilde{\Phi}(\varepsilon, g, b) := \tilde{g}^{\left(\left\lceil \frac{8b^2}{\tilde{\varepsilon}} \right\rceil\right)}(0)$, where for $g : \mathbb{N} \to \mathbb{N}$ we define $\tilde{g}(n) := \max\{n, g(n)\}$. Moreover, let $t_\varepsilon(n) := \rho(\delta^4_{\varepsilon,n}/96b^3)$ and – for $v \in C$ and $\delta > 0$ – $v_\delta := v - \frac{\delta^2}{16^2}(Id - T)(v) \in C$. Finally, for $h : \mathbb{N} \to \mathbb{N}$ and $w \in B_1(0)$ define

$$\chi_h(n) := \max\{n, n_{\tilde{\varepsilon}}, \chi(\widehat{n}_\varepsilon), \tilde{h}^M(\chi(\widehat{n}_\varepsilon))\} \text{ and } h_{w,v}(n) := \begin{cases} n, & \text{if } |\langle y_n - v, w\rangle| \ge \frac{\varepsilon}{2} \\ \tilde{h}(n), & \text{otherwise.} \end{cases}$$

Lemma 3.5. *Let $T : C \to C$ be nonexpansive, $w \in B_1(0)$, $h : \mathbb{N} \to \mathbb{N}$ and $\varepsilon > 0$. Then there are $u, v \in C$ and $\chi : \mathbb{N} \to \mathbb{N}$ such that*

$$\exists m \le \tilde{\Phi}(\varepsilon, \chi_h, b)\, \exists q, n, l \le \chi(\widehat{m}_\varepsilon)\, \exists k_m, \tilde{k}_m \ge t_\varepsilon\big((\chi_h)^M(\tilde{\Phi}(\varepsilon, \chi_h, b))\big)$$
$$\begin{cases} (a)\ q, n \ge m_{\tilde{\varepsilon}} \wedge \\ (b)\ |\langle y_n - u, u - v\rangle| < \frac{\tilde{\varepsilon}b}{2} \wedge \\ (c)\ |\langle y_{k_m} - u, (Id - T)(u_{\delta_{\varepsilon,(\chi_h)^M(\tilde{\Phi}(\varepsilon,\chi_h,b))}})\rangle| < \frac{\delta^4_{\varepsilon,(\chi_h)^M(\tilde{\Phi}(\varepsilon,\chi_h,b))}}{96b^2} \wedge \\ (d)\ |\langle y_{h_{w,u}(l)} - v, w\rangle| < \frac{\varepsilon}{4} \wedge \\ (e)\ |\langle y_{h_{w,u}(q)} - v, v - u\rangle| < \frac{\tilde{\varepsilon}b}{2} \wedge \\ (f)\ |\langle y_{h_{w,u}(\tilde{k}_m)} - v, (Id - T)(v_{\delta_{\varepsilon,(\chi_h)^M(\tilde{\Phi}(\varepsilon,\chi_h,b))}})\rangle| < \frac{\delta^4_{\varepsilon,(\chi_h)^M(\tilde{\Phi}(\varepsilon,\chi_h,b))}}{96b^2} \wedge \\ (g)\ \forall j \in [m; \chi_h(m)] \bigwedge\limits_{z \in \{u,v\}} \left(\left|\|\widetilde{x_m - z}\|^2 - \|\widetilde{x_j - z}\|^2\right| < \tilde{\varepsilon}\right) \end{cases}$$

Proof. $(y_n)_{n\in\mathbb{N}} \subset C$ has a weak cluster point $u \in C$, i.e.

$$(+)\ \forall z \in X\, \forall k \in \mathbb{N}\, \exists n \ge k\, \big(|\langle y_n - u, z\rangle| < 2^{-k}\big).$$

Analogously, there is a weak cluster point $v \in C$ of $(y_{h_{w,u}(n)})_{n\in\mathbb{N}}$, i.e.

$$(++)\ \forall z \in X\, \forall k \in \mathbb{N}\, \exists n \ge k\, \big(|\langle y_{h_{w,u}(n)} - v, z\rangle| < 2^{-k}\big).$$

Hence for given $\varepsilon > 0$ and $w \in B_1(0)$ there is $\chi : \mathbb{N} \to \mathbb{N}$ (using only QF-AC^{0,0}) such that

$$(*) \begin{cases} \forall k \in \mathbb{N}\, \exists q, n, l \in [k; \chi(k)] \\ \big(|\langle y_n - u, u - v\rangle| < 2^{-k} \wedge |\langle y_{h_{w,u}(q)} - v, v - u\rangle| < 2^{-k} \wedge |\langle y_{h_{w,u}(l)} - v, w\rangle| < 2^{-k}\big). \end{cases}$$

By lemma 3.2 we also have (note that $\chi_h = \widetilde{(\chi_h)}$)

$$\exists m \le \tilde{\Phi}(\varepsilon, \chi_h, b)\, \forall j \in [m; \chi_h(m)] \bigwedge_{z \in \{u,v\}} \left(\left|\|\widetilde{x_m - z}\|^2 - \|\widetilde{x_j - z}\|^2\right| < \tilde{\varepsilon}\right)$$

since $\|x_j - z\|^2 \le 4b^2$.
Let m be as above and apply $(*)$ to $k := \widehat{m}_\varepsilon$. Then (using that $m_{\tilde\varepsilon} \le \widehat{m}_\varepsilon$)

$$\begin{aligned}&\exists q, n, l \in [m_{\tilde\varepsilon}; \chi(\widehat{m}_\varepsilon)]\\ &\quad \left(|\langle y_n - u, u - v\rangle| < \tfrac{\tilde\varepsilon b}{2} \wedge |\langle y_{h_{w,u}(q)} - v, v - u\rangle| < \tfrac{\tilde\varepsilon b}{2} \wedge |\langle y_{h_{w,u}(l)} - v, w\rangle| < \tfrac{\varepsilon}{4}\right).\end{aligned}$$

Applying $(+)$ and $(++)$ again yields

$$\begin{aligned}&\exists k_m, \tilde k_m \ge t_\varepsilon\big((\chi_h)^M(\tilde\Phi(\varepsilon, \chi_h, b))\big)\\ &\quad \left(\begin{aligned}&|\langle y_{k_m} - u, (Id - T)(u_{\delta_{\varepsilon,(\chi_h)^M(\tilde\Phi(\varepsilon,\chi_h,b))}})\rangle| < \frac{\delta^4_{\varepsilon,(\chi_h)^M(\tilde\Phi(\varepsilon,\chi_h,b))}}{96b^2} \wedge\\ &|\langle y_{h_{w,u}(\tilde k_m)} - v, (Id - T)(v_{\delta_{\varepsilon,(\chi_h)^M(\tilde\Phi(\varepsilon,\chi_h,b))}})\rangle| < \frac{\delta^4_{\varepsilon,(\chi_h)^M(\tilde\Phi(\varepsilon,\chi_h,b))}}{96b^2}\end{aligned}\right),\end{aligned}$$

which concludes the proof. $\square$

In the following, we construct – using Ω^* from the previous section – a majorant χ^* for χ in lemma 3.5 that only depends on $\varepsilon > 0, b \in \mathbb{N}$ and $h : \mathbb{N} \to \mathbb{N}$ but not on X, C (except for b), the points $x \in C, w \in B_1(0)$ or the mapping $T : C \to C$. For $\psi : X \times \mathbb{N} \to \mathbb{N}$ define $(\chi \wedge \psi)(u, n) := \max\{\chi(u, n), \psi(u, n)\}$ and (for additionally $u, v \in C$)

$$\begin{aligned}&(\chi \wedge \psi)_{u,v}(n) := (\chi \wedge \psi)_{u,v,w}(n) :=\\ &\max\left((\chi \wedge \psi)(v - u, n), (\chi \wedge \psi)(u - v, n), (\chi \wedge \psi)(w, n)\right),\\ &(\chi \wedge \psi)_{u,v,h} := \big((\chi \wedge \psi)_{u,v}\big)_h.\end{aligned}$$

Now define

$$K_{v,\psi}(u, \chi) :=$$
$$\begin{cases} \widehat{m}_\varepsilon \text{ for the least } m \le \tilde\Phi(\varepsilon, (\chi \wedge \psi)_{u,v,h}, b) \text{ such that} \\ \;\neg\forall z \in B_{2b}(0)\, \exists n \in [\widehat{m}_\varepsilon, (\chi \wedge \psi)(z, \widehat{m}_\varepsilon)]\ \left(|\langle y_n - u, z\rangle| < \min\left\{\frac{\tilde\varepsilon b}{2}, \frac{\varepsilon}{4}\right\}\right) \text{ if existent,} \\ \max\left\{t_\varepsilon\left(((\chi \wedge \psi)_{u,v,h})^M(\tilde\Phi(\varepsilon, (\chi \wedge \psi)_{u,v,h}, b))\right), \log\left\lceil \frac{96b^2}{\delta^4_{\varepsilon,((\chi\wedge\psi)_{u,v,h})^M(\tilde\Phi(\varepsilon,(\chi\wedge\psi)_{u,v,h},b))}}\right\rceil\right\}, \\ \qquad\qquad\qquad\qquad \text{otherwise.}\end{cases}$$

Furthermore, define

$$\tilde K_{u,\chi}(v, \psi) :=$$
$$\begin{cases} \widehat{m}_\varepsilon \text{ for the least } m \le \tilde\Phi(\varepsilon, (\chi \wedge \psi)_{u,v,h}, b) \text{ such that} \\ \;\neg\forall z \in B_{2b}(0)\, \exists n \in [\widehat{m}_\varepsilon, (\chi \wedge \psi)(z, \widehat{m}_\varepsilon)]\ \left(|\langle y_{h_{w,u}(n)} - v, z\rangle| < \min\left\{\frac{\tilde\varepsilon b}{2}, \frac{\varepsilon}{4}\right\}\right) \text{ if existent,} \\ \max\left\{t_\varepsilon\left(((\chi \wedge \psi)_{u,v,h})^M(\tilde\Phi(\varepsilon, (\chi \wedge \psi)_{u,v,h}, b))\right), \log\left\lceil \frac{96b^2}{\delta^4_{\varepsilon,((\chi\wedge\psi)_{u,v,h})^M(\tilde\Phi(\varepsilon,(\chi\wedge\psi)_{u,v,h},b))}}\right\rceil\right\}, \\ \qquad\qquad\qquad\qquad \text{otherwise.}\end{cases}$$

Remark 3.6. It is clear that K and $\tilde{K}$ are majorizable (in fact by primitive recursive functionals; see below) and hence denote functionals in the model $\mathcal{M}^{\omega,X}$. The ineffective case distinction used in the definition of $K, \tilde{K}$ could be avoided since instead of '$\forall z \in B_{2b}(0)$' it suffices to consider '$\forall z \in \{v-u, u-v, w\}$' (see below) and '$<_{\mathbb{R}}$' can be replaced by an appropriate approximate version. However, for the construction of the majorants this anyhow does not matter as the case distinction gets replaced by taking the maximum of both cases.

Let us motivate first the definition of $K_{v,\psi}$ (and similarly of $\tilde{K}_{u,\chi}$):
for $u := V(K_{v,\psi}, 2b, (y_n))$ and $\chi := \tilde{\Omega}(K_{v,\psi}, 2b, (y_n))$ (where $\tilde{\Omega}$ is as in corollary 2.3 – extended to general b-bounded closed convex C as in the comments after corollary 2.3 – and $V(K, 2b, (y_n))$ selects a $v \in C$ as in corollary 2.3) we get from corollary 2.3 that

$$(+) \begin{cases} \forall m \leq \tilde{\Phi}(\varepsilon, (\chi \wedge \psi)_{u,v,h}, b) \\ \forall z \in B_{2b}(0)\, \exists n \in [\widehat{m}_\varepsilon, (\chi \wedge \psi)(z, \widehat{m}_\varepsilon)]\ \left(|\langle y_n - u, z\rangle| < \min\left\{\frac{\tilde{\varepsilon}b}{2}, \frac{\varepsilon}{4}\right\}\right) \end{cases}$$

since, otherwise, $K_{v,\psi}$ would be $\widehat{m}_\varepsilon$ for the least counterexample $m \leq \tilde{\Phi}(\varepsilon, (\chi \wedge \psi)_{u,v,h}, b)$ so that

$$\neg \forall z \in B_{2b}(0)\, \exists n \in [\widehat{m}_\varepsilon, (\chi \wedge \psi)(z, \widehat{m}_\varepsilon)]\ \left(|\langle y_n - u, z\rangle| < \min\left\{\frac{\tilde{\varepsilon}b}{2}, \frac{\varepsilon}{4}\right\}\right)$$

and so a fortiori – since $\chi \wedge \psi \geq \chi$ and $K_{v,\psi}(u,\chi) = \widehat{m}_\varepsilon \geq \log(2/(\tilde{\varepsilon}b), \log(4/\varepsilon)$ –

$$\neg \forall z \in B_{2b}(0)\, \exists n \in [K_{v,\psi}(u,\chi), (\chi(z, K_{v,\psi}(u,\chi))]\ \left(|\langle y_n - u, z\rangle| < 2^{-K_{v,\psi}(u,\chi)}\right)$$

contradicting corollary 2.3.
$(+)$ implies (applied to $z \in \{v-u, u-v, w\}$ and using that $(\chi \wedge \psi)_{u,v}(n) \geq (\chi \wedge \psi)(z,n)$ for those z and all n)[6]

$$(++) \begin{cases} \forall m \leq \tilde{\Phi}(\varepsilon, (\chi \wedge \psi)_{u,v,h}, b) \\ \forall z \in \{v-u, u-v, w\}\, \exists n \in [\widehat{m}_\varepsilon, (\chi \wedge \psi)_{u,v}(\widehat{m}_\varepsilon)]\ \left(|\langle y_n - u, z\rangle| < \min\left\{\frac{\tilde{\varepsilon}b}{2}, \frac{\varepsilon}{4}\right\}\right) \end{cases}$$

$(+)$ also implies (by the definition of $K_{v,\psi}$) that

$$K_{v,\psi}(u,\chi) := \max\left\{ t_\varepsilon\left(((\chi \wedge \psi)_{u,v,h})^M(\tilde{\Phi}(\varepsilon, (\chi \wedge \psi)_{u,v,h}, b))\right), \log\left\lceil \frac{96b^2}{\delta^4_{\varepsilon,((\chi\wedge\psi)_{u,v,h})^M(\tilde{\Phi}(\varepsilon,(\chi\wedge\psi)_{u,v,h},b))}} \right\rceil \right\}$$

and so – by corollary 2.3 –

$$\forall z \in B_{2b}(0) \exists n \geq t_\varepsilon\left(((\chi \wedge \psi)_{u,v,h})^M(\tilde{\Phi}(\varepsilon, (\chi \wedge \psi)_{u,v,h}, b))\right) \left(|\langle y_n - u, z\rangle| < \frac{\delta^4_{\varepsilon,((\chi\wedge\psi)_{u,v,h})^M(\tilde{\Phi}(\varepsilon,(\chi\wedge\psi)_{u,v,h},b))}}{96b^2}\right)$$

[6] For $K_{v,\psi}$ we only need $z := u - v$ while for $\tilde{K}_{u,\chi}$ we need $z := v - u$ and $z := w$.

which – in particular – can be applied to

$$z := (Id - T)\big(u_{\delta_{\varepsilon,((\chi\wedge\psi)_{u,v,h})^M(\tilde{\Phi}(\varepsilon,(\chi\wedge\psi)_{u,v,h},b))}}\big).$$

This finishes the motivation.

Now let u, χ be variables again and consider

$$\psi_{u,\chi} := \tilde{\Omega}(\tilde{K}_{u,\chi}, 2b, (y_{h_{w,u}(n)})) \text{ and } v_{u,\chi} := V(\tilde{K}_{u,\chi}, 2b, (y_{h_{w,u}(n)})),$$

where $\tilde{\Omega}$ is as in corollary 2.3 and $V(K, 2b, (y_n))$ selects a $v \in C$ as in corollary 2.3, and define

$$K(u,\chi) := K_{v_{u,\chi},\psi_{u,\chi}}(u,\chi).$$

Finally, put

$$\chi_K := \tilde{\Omega}(K, 2b, (y_n)) \text{ and } u_K := V(K, 2b, (y_n)).$$

Then

$$\chi := \big(\chi_K \wedge \psi_{u_K,\chi_K}\big)_{u,v}$$

so that

$$\chi_h = \big(\chi_K \wedge \psi_{u_K,\chi_K}\big)_{u,v,h}$$

satisfies lemma 3.5 with $u := u_K, v := v_{u_K,\chi_K}$.
χ can be easily majorized as a functional $\chi^*(\varepsilon, h, b)$ as follows: define

$$K^*_{v,\psi}(u,\chi) := \tilde{K}^*_{u,\chi}(v,\psi) := \max\left\{(\tilde{\Phi}\widehat{(\varepsilon,\alpha,b)})_\varepsilon, t_\varepsilon(\alpha(\tilde{\Phi}(\varepsilon,\alpha,b))), \log\left\lceil \frac{96b^2}{\delta^4_{\varepsilon,\alpha(\tilde{\Phi}(\varepsilon,\alpha,b))}}\right\rceil\right\},$$

where

$$\alpha := \big(((\chi\wedge\psi)(2b))_h\big)^M.$$

Then

$$\lambda v,\psi,u,\chi.K^*_{v,\psi}(u,\chi) \gtrsim \lambda v,\psi,u,\chi.K_{v,\psi}(u,\chi)$$

and

$$\lambda u,\chi,v,\psi.\tilde{K}^*_{v,\psi}(u,\chi) \gtrsim \lambda u,\chi,v,\psi.\tilde{K}_{v,\psi}(u,\chi).$$

Note that K^* and $\tilde{K}^*$ actually do not depend on u, v, w as these vectors are majorized by b and hence $u - v, v - u$ are majorized by $2b$. Likewise, $v_{u,\chi}$ is majorized by b and for

$$\psi^*_{u,\chi} := \Omega^*(\tilde{K}^*_{u,\chi}, 2b)$$

we have that $\lambda u,\chi.\psi^*_{u,\chi} \gtrsim \lambda u,\chi.\psi_{u,\chi}$. Hence for $K^*(u,\chi) := K^*_{b,\psi^*_{u,\chi}}(u,\chi)$ we have that $K^* \gtrsim K$. Finally, define (making the hidden parameters ε, h, b explicit)

$$\chi^*(\varepsilon,h,b) := \chi^*_K \wedge \psi^*_{b,\chi^*_K}, \text{ where } \chi^*_K := \Omega^*(K^*, 2b).$$

Then

$$\chi^*(\varepsilon,h,b) \gtrsim \chi(\varepsilon,h,b).$$

Note that (for rational $\varepsilon \in \mathbb{Q}^*_+$) the bound χ^* is primitive recursive (in the sense of Kleene) in Ω^* and so in total definable in $T_0 + B_{0,1}$ and hence (see the discussion before theorem 2.2) in Gödel's system T of primitive recursive functionals of finite type.

Theorem 3.7 (Effective quantitative metastable version of Baillon's nonlinear ergodic theorem)**.** *Let X be a Hilbert space and $C \subset X$ be a nonempty bounded, closed and convex subset. Let $\mathbb{N}^* \ni b \geq \|x\|$ for all $x \in C$. For $\varepsilon > 0$, $h : \mathbb{N} \to \mathbb{N}$, $w \in B_1(0)$ and $T : C \to C$ be nonexpansive, the following holds*

$$\exists l \leq \varphi(\varepsilon,h,b)\left(|\langle y_l - y_{\tilde{h}(l)}, w\rangle| < \varepsilon\right),$$

where

$$\varphi(\varepsilon,h,b) := \chi^*(\varepsilon,h,b)(\widehat{n}_\varepsilon)$$

with

$$\begin{array}{l} \widehat{n}_\varepsilon := \max\{n_{\tilde{\varepsilon}}, \lceil \log(2/(\tilde{\varepsilon}b))\rceil, \lceil \log(4/\varepsilon)\rceil\}, \textit{where} \\ n := \tilde{\Phi}(\varepsilon, (\chi^*(\varepsilon,h,b))_{h^M}, b), \\ \tilde{\varepsilon} := \frac{\varepsilon^2}{144b} \textit{ and } n_\varepsilon := \left\lceil \frac{2(n+1)b}{\varepsilon} \right\rceil. \end{array}$$

Proof. Let $u, v, \chi, m, n, q, l, k_m, \tilde{k}_m$ be as in lemma 3.5 where χ is constructed as above. We will show now that l satisfies $|\langle y_l - y_{\tilde{h}(l)}, w\rangle| < \varepsilon$. Since (by lemma 3.5) $l \leq \chi(\widehat{m}_\varepsilon)$ with $m \leq \tilde{\Phi}(\varepsilon, \chi_h, b)$ it then follows that $l \leq \varphi(\varepsilon,h,b)$ which proves the theorem.
By lemma 3.3, the conjuncts (c) and (f) in lemma 3.5 imply (using that by lemma 3.4 ρ is a rate of asymptotic regularity for (y_n) and $k_m, \tilde{k}_m \geq t_\varepsilon\big((\chi_h)^M(\tilde{\Phi}(\varepsilon,\chi_h,b))\big)$) that

$$(1)\ \|T(u) - u\| < \delta_{\varepsilon,(\chi_h)^M(\tilde{\Phi}(\varepsilon,\chi_h,b))}$$

and (since $h_{w,u}(\tilde{k}_m) \geq \tilde{k}_m$)

$$(2)\ \|T(v) - v\| < \delta_{\varepsilon,(\chi_h)^M(\tilde{\Phi}(\varepsilon,\chi_h,b))}.$$

For all $z \in C$ and $k \in \mathbb{N}$ one has

$$\begin{array}{l} \|T(z) - z\| < \delta_{\varepsilon,n} \to \|x_{k+1} - z\|^2 \leq (\|T(x_k) - T(z)\| + \|T(z) - z\|)^2 \\ < (\|T(x_k) - T(z)\| + \delta_{\varepsilon,n})^2 \leq (\|x_k - z\| + \delta_{\varepsilon,n})^2 = \|x_k - z\|^2 + 2\delta_{\varepsilon,n}\|x_k - z\| + \delta^2_{\varepsilon,n} \\ \overset{\delta_{\varepsilon,n}\leq 1}{\leq} \|x_k - z\|^2 + (4b+1)\delta_{\varepsilon,n} \leq \|x_k - z\|^2 + \tilde{\varepsilon}\cdot\frac{1}{n+1}. \end{array}$$

(1) and (2) yield (note that $(\chi_h)^M(\tilde{\Phi}(\varepsilon,\chi_h,b)) \geq \chi_h(m)$) that for $z \in \{u,v\}$

$$(3)\ \|x_j - z\|^2 - \|\widetilde{x_j - z}\|^2 < \tilde{\varepsilon}$$

for all $j \le \chi_h(m)$ and so conjunct (g) in lemma 3.5 (using that always $\|x_j - z\|^2 \geq \|\widetilde{x_j - z}\|^2$) implies

$$(4)\ \forall j \in [m; \chi_h(m)] \left(\left| \|x_m - z\|^2 - \|x_j - z\|^2 \right| < 2\tilde{\varepsilon} \right).$$

Now define (for $u, v \in C, \varepsilon > 0$ and $m \in \mathbb{N}$) a convex subset

$$K^{v,u}_{\varepsilon,m} := \left\{ z \in X : |2\langle z - v, v - u\rangle + \|x_m - v\|^2 - \|x_m - u\|^2 + \|v - u\|^2| \le 4\tilde{\varepsilon} \right\}.$$

For all $a, \tilde{a} \in X$ and $j \in \mathbb{N}$ one has

$$(5)\ \|x_j - \tilde{a}\|^2 = \|x_j - a\|^2 + \|a - \tilde{a}\|^2 + 2\langle x_j - a, a - \tilde{a}\rangle$$

and so for $(a, \tilde{a}) = (v, u)$ resp. $(a, \tilde{a}) = (u, v)$ we get together with (4)

$$C_m^{\chi_h(m)} \subseteq K^{v,u}_{\varepsilon,m}, K^{u,v}_{\varepsilon,m}.$$

By the definition of $m_{\tilde{\varepsilon}}$ we have by lemma 3.1

$$(6)\ \forall k \in [m_{\tilde{\varepsilon}}, \chi_h(m)]\, \exists u_k \in C_m^{\chi_h(m)}\ (\|y_k - u_k\| \le \tilde{\varepsilon}).$$

Hence

$$(7)\ \begin{cases} \forall k \in [m_{\tilde{\varepsilon}}, \chi_h(m)]\ (|2\langle y_k - v, v - u\rangle + \|x_m - v\|^2 - \|x_m - u\|^2 + \|v - u\|^2| \\ \le 4\tilde{\varepsilon} + 2\tilde{\varepsilon}\|u - v\| \le 8\tilde{\varepsilon} b) \end{cases}$$

and (analogously)

$$(8)\ \forall k \in [m_{\tilde{\varepsilon}}, \chi_h(m)]\ \left(\left|2\langle y_k - u, u - v\rangle + \|x_m - u\|^2 - \|x_m - v\|^2 + \|v - u\|^2\right| \le 8\tilde{\varepsilon} b \right).$$

By the conjuncts (b) and (e) in lemma 3.5 this yields (note that – using conjunct (a) in lemma 3.5 – $\chi_h(m) \ge h_{w,u}(q) \ge q \ge m_{\tilde{\varepsilon}}$ and $\chi_h(m) \ge \chi(\widehat{m}_\varepsilon) \ge n \ge m_{\tilde{\varepsilon}}$)

$$(9)\ \|x_m - v\|^2 + \|v - u\|^2 \le \|x_m - u\|^2 + 9\tilde{\varepsilon} b$$

and

$$(10)\ \|x_m - u\|^2 + \|v - u\|^2 \le \|x_m - v\|^2 + 9\tilde{\varepsilon} b$$

and so in turn

$$(11)\ \|v - u\|^2 \le 9\tilde{\varepsilon} b = \frac{\varepsilon^2}{16},\ \text{i.e. } \|v - u\| \le \frac{\varepsilon}{4}.$$

By conjunct (d) in lemma 3.5 and (11) we have that

$$(12)\ |\langle y_{h_{w,u}(l)} - u, w\rangle| < \frac{\varepsilon}{2}$$

and so (using the definition of $h_{w,u}$)

$$(13)\ |\langle y_l - u, w\rangle| < \frac{\varepsilon}{2} \wedge |\langle y_{\tilde{h}(l)} - u, w\rangle| < \frac{\varepsilon}{2}.$$

Hence

$$(14)\ \begin{cases} |\langle y_l - y_{\tilde{h}(l)}, w\rangle| = |\langle (y_l - u) - (y_{\tilde{h}(l)} - u), w\rangle| = |\langle y_l - u, w\rangle - \langle y_{\tilde{h}(l)} - u, w\rangle| \\ = |\langle y_l - u, w\rangle| + |\langle y_{\tilde{h}(l)} - u, w\rangle| < \varepsilon. \end{cases}$$

□

Theorem 3.7 has the following seemingly stronger consequence:

Corollary 3.8.

$$\forall \varepsilon > 0\, \forall h : \mathbb{N} \to \mathbb{N}\, \forall w \in B_1(0)\, \exists l \le \varphi(\varepsilon/2, (h^+)^M, b)\, \forall i, j \in [l; l+h(l)]\, \big(|\langle y_i - y_j, w\rangle| < \varepsilon\big),$$

where $h^+(n) := n + h(n)$.

Proof. Apply theorem 3.7 to

$$h^-_w(l) := l + \min i \le h(l)\, [\forall j \le h(l)\, \big(|\langle y_l - y_{l+j}, w\rangle| \le |\langle y_l - y_{l+i}, w\rangle|\big)]$$

and $\varepsilon/2$. Note that $\widetilde{h^-_w} = h^-_w$ and that $(h^+)^M$ majorizes h^-_w. Since – by construction – φ is a selfmajorizing functional we have that

$$\exists l \le \varphi(\varepsilon/2, h^-_w, b) \le \varphi(\varepsilon/2, (h^+)^M, b)\, \big(|\langle y_l - y_{h^-_w(l)}, w\rangle| < \frac{\varepsilon}{2}\big).$$

Since

$$\forall j \le h(l)\, \big(|\langle y_l - y_{l+j}, w\rangle| \le |\langle y_l - y_{h^-_w(l)}, w\rangle| < \frac{\varepsilon}{2}\big),$$

it follows that

$$\forall i, j \in [l; l + h(l)]\, \big(|\langle y_i - y_j, w\rangle| \le |\langle y_l - y_j, w\rangle| + |\langle y_l - y_i, w\rangle| < \varepsilon\big).$$

□

Final comments: While we believe that the above analysis faithfully reflects the finitary combinatorial and effective content of the proof of Baillon's theorem as given by Brézis and Browder in [4], it remains open whether maybe a different proof (e.g. Baillon's original proof from [2]) might lead to an analysis that would allow one to bypass the use of the quantitative weak compactness functional Ω^* and so might provide a primitive recursive (in the ordinary sense of Kleene) bound on the metastable version. In Baillon's proof, the weak convergence of (y_n) is established by showing that this sequence weakly converges towards the strong limit of the sequence of projections of $T^n(x_0)$ to the fixed point set $Fix(T)$ of T. Though such projection arguments are hard to unwind, we managed in [14] to do so in the case of a similar (but simpler) projection argument (namely the existence of the projection of just x_0 towards $Fix(T)$) in the proofs due to Browder [5] and Wittmann [25] (which do use additionally weak compactness) of their respective well-known theorems which resulted in an elimination of weak compactness altogether and a primitive recursive bound. Only further research will tell us whether such an elimination of weak sequential compactness might happen in the course of a logical analysis of Baillon's proof (repeated essentially also in [22]). However, even if this should turn out to be the case, it is not clear whether this will result in a simpler (and primitive recursive) bound as the existence of the sequence of projections of $T^n(x_0)$ towards $Fix(T)$ requires and instance of countable choice for a Π^0_1-formula which in itself needs (for its monotone Gödel functional interpretation) bar recursion (of lowest type) just as weak sequential compactness does.

Acknowledgments.
We are grateful to Johannes Schramm for bringing the particularly simple proof of Baillon's theorem from [4] to our attention.

The author has been supported by the German Science Foundation (DFG Project KO 1737/5-1).

References

[1] Avigad, J., Gerhardy, P., Towsner, H., Local stability of ergodic averages. Trans. Amer. Math. Soc. **362**, pp. 261-288 (2010).

[2] Baillon, J.B., Un théorème de type ergodique pour les contractions non linéaires dans un espace de Hilbert. C.R. Acad. Sci. Paris Sèr. A-B **280**, pp. 1511-1514 (1975).

[3] Bezem, M., Strongly majorizable functionals of finite type: a model for bar recursion containing discontinuous functionals. J. Symbolic Logic **50** pp. 652–660 (1985).

[4] Brézis, H., Browder, F.E., Nonlinear ergodic theorems. Bull. Amer. Math. Soc. **82**, pp. 959-961 (1976).

[5] Browder, F.E., Convergence of approximants to fixed points of nonexpansive nonlinear mappings in Banach spaces. Arch. Rational Mech. Anal. **24**, pp. 82-90 (1967).

[6] Gerhardy, P., Kohlenbach, U., General logical metatheorems for functional analysis. Trans. Amer. Math. Soc. **360**, pp. 2615-2660 (2008).

[7] Gödel, K., Über eine bisher noch nicht benützte Erweiterung des finiten Standpunktes. Dialectica **12**, pp. 280–287 (1958).

[8] Hilbert, D., Über das Unendliche. Math. Ann. **95**, pp. 161-190 (1926).

[9] Kohlenbach, U., Analysing proofs in analysis. In: W. Hodges, M. Hyland, C. Steinhorn, J. Truss, editors, *Logic: from Foundations to Applications. European Logic Colloquium* (Keele, 1993), pp. 225–260, Oxford University Press (1996).

[10] Kohlenbach, U., On the no-counterexample interpretation. J. Symbolic Logic **64**, pp. 1491-1511 (1999).

[11] Kohlenbach, U., Some logical metatheorems with applications in functional analysis. Trans. Amer. Math. Soc. **357**, no. 1, pp. 89-128 (2005).

[12] Kohlenbach, U., Applied Proof Theory: Proof Interpretations and their Use in Mathematics. Springer Monographs in Mathematics. xx+536pp., Springer Heidelberg-Berlin, 2008.

[13] Kohlenbach, U., On the logical analysis of proofs based on nonseparable Hilbert space theory. In: Feferman, S., Sieg, W. (eds.), 'Proofs, Categories and Computations. Essays in Honor of Grigori Mints'. College Publications, pp. 131-143 (2010).

[14] Kohlenbach, U.,On quantitative versions of theorems due to F.E. Browder and R. Wittmann. Advances in Mathematics **226**, pp. 2764-2795 (2011).

[15] Kohlenbach, U., Gödel functional interpretation and weak compactness. To appear in: Ann. Pure Appl. Logic.

[16] Kohlenbach, U., Leuştean, L., A quantitative mean ergodic theorem for uniformly convex Banach spaces. Ergodic Theory and Dynamical Systems **29**, pp. 1907-1915 (2009).

[17] Kohlenbach, U., Leuştean, L., Asymptotically nonexpansive mappings in uniformly convex hyperbolic spaces. Journal of the European Mathematical Society **12**, pp. 71-92 (2010).

[18] Kreisel, G., On the interpretation of non-finitist proofs, part I. J. Symbolic Logic **16**, pp.241-267 (1951).

[19] Kreisel, G., On the interpretation of non-finitist proofs, part II: Interpretation of number theory, applications. J. Symbolic Logic **17**, pp. 43-58 (1952).

[20] Safarik, P., Kohlenbach, U., On the computational content of the Bolzano-Weierstraß principle. Math. Logic Quaterly **56**, pp. 508-532 (2010).

[21] Spector, C., Provably recursive functionals of analysis: a consistency proof of analysis by an extension of principles formulated in current intuitionistic mathematics. In: Recursive function theory, Proceedings of Symposia in Pure Mathematics, vol. 5 (J.C.E. Dekker (ed.)), AMS, Providence, R.I., pp. 1–27 (1962).

[22] Takahashi, W., Nonlinear Functional Analysis. Yokohama Publishers, iv+276pp., 2000.

[23] Tao, T., Soft analysis, hard analysis, and the finite convergence principle. Essay posted May 23, 2007. Appeared in: 'T. Tao, Structure and Randomness: Pages from Year One of a Mathematical Blog. AMS, 298pp., 2008'.

[24] Tao, T., Norm convergence of multiple ergodic averages for commuting transformations. Ergodic Theory and Dynamical Systems **28**, pp. 657-688 (2008).

[25] Wittmann, R., Approximation of fixed points of nonexpansive mappings. Arch. Math. **58**, pp. 486-491 (1992).