A quantitative Mean Ergodic Theorem for uniformly convex Banach spaces

U. Kohlenbach[1], L. Leuştean[1,2]

[1] Department of Mathematics, Technische Universität Darmstadt,
Schlossgartenstrasse 7, 64289 Darmstadt, Germany

[2] Institute of Mathematics "Simion Stoilow" of the Romanian Academy,
Calea Griviţei 21, P.O. Box 1-462, Bucharest, Romania

E-mails: kohlenbach,leustean@mathematik.tu-darmstadt.de

October 17, 2008

Abstract

We provide an explicit uniform bound on the local stability of ergodic averages in uniformly convex Banach spaces. Our result can also be viewed as a finitary version in the sense of T. Tao of the Mean Ergodic Theorem for such spaces and so generalizes similar results obtained for Hilbert spaces by Avigad, Gerhardy and Towsner [1] and T. Tao [11].

1 Introduction

In the following $\mathbb{N} := \{1, 2, 3, \ldots\}$.
Let X be a Banach space and $T : X \to X$ be a selfmapping of X. The *Cesaro mean* starting with $x \in X$ is the sequence $(x_n)_{n\geq 1}$ defined by $x_n := \frac{1}{n}\sum_{i=0}^{n-1} T^i x$.

In 1939, Garrett Birkhoff proved the following generalization of von Neumann's Mean Ergodic Theorem.

Theorem 1.1. *[2] Let X be a uniformly convex Banach space and $T : X \to X$ be a linear operator with $\|Tx\| \leq \|x\|$ for all $x \in X$. Then for any $x \in X$, the Cesaro mean (x_n) is convergent.*

In [1], Avigad, Gerhardy and Towsner address the issue of finding an effective rate of convergence for (x_n) in Hilbert spaces. They show that even for the separable Hilbert space L_2 there are simple computable such operators T and computable points $x \in L_2$ such that there is no computable rate of convergence of (x_n). In such a situation the best one can hope for is an effective bound on the following reformulation of the Cauchy property of (x_n) which in logic is called the Herbrand normal form of the latter:

$$\forall \varepsilon > 0\, \forall g : \mathbb{N} \to \mathbb{N}\, \exists n \in \mathbb{N}\, \forall i, j \in [n, n+g(n)]\ (\|x_i - x_j\| < \varepsilon). \tag{1}$$

It is trivial to see that (1) is implied by the Cauchy property. However, ineffectively, also the converse implication holds. The mathematical relevance of this reformulation of convergence was recently pointed out by T. Tao ([10, 11]), who also uses the term 'metastability'. In [6] (and refined in [4]) a general logical metatheorem is proved that guarantees (given a proof of (1)) the extractability of an effective bound $\Phi(\varepsilon, g, b)$ on '$\exists n$' in (1) that is highly uniform in the sense that it only depends on g, ε and an upper bound $\mathbb{N} \ni b \geq \|x\|$ but otherwise is independent from x, X and T. In fact, by a simple renorming argument one can always achieve to have the bound to depend on b, ε only via b/ε. The proof of this metatheorem, which is based on a recent extension and refinement of a technique from logic called (monotone) Gödel functional interpretation, provides an algorithm for extracting an explicit such Φ from a given proof (for a book treatment of all this see [8]). Moreover, the proof of the metatheorem (more precisely the 'soundness theorem' for the monotone functional interpretation) implies that in analyzing any other proof of some theorem that uses the Mean Ergodic Theorem as a lemma, one will only need to know this effective uniform bound about the Mean Ergodic Theorem in extracting a corresponding bound for the theorem in question and so does not have to analyze again some proof of the Mean Ergodic Theorem in such applications of the latter ('modularity of functional interpretation'). For this it is crucial that the bound on (1) holds for **all** functions $g : \mathbb{N} \to \mathbb{N}$. In the context of metric fixed point theory, this feature of functional interpretation has been used already in many application (see [7] for a survey) and can also be seen in a simple form internally in the proof of the main result of this paper: the proof of the Mean Ergodic Theorem uses the convergence of monotone bounded sequences of reals as a lemma. In order to extract a bound on the Herbrand normal form (1) of the former we only need from the latter a corresponding quantitative form (Lemma 3.1) which, however, in order to yield a bound on (1) for g needs to be applied to the more complicated function h in Theorem 2.1.

Guided by this proof theoretic approach sketched above, Avigad, Gerhardy and Towsner [1] extract such a bound from a standard textbook proof of von Neumann's Mean Ergodic Theorem. A less direct proof for the existence of a bound with the above mentioned uniformity features is – for a particular finitary dynamical system – also given by T. Tao [11] as part of his proof of a generalization of the von Neumann Mean Ergodic Theorem to commuting families of invertible measure preserving transformations $T_1, \ldots, T_l$.

In this note we apply the same methodology to Birkhoff's proof of Theorem 1.1 and extract an even easier to state bound for the more general case of uniformly convex Banach spaces. In this setting, the bound additionally depends on a given modulus of uniform convexity for X. Despite of our result being significantly more general then the Hilbert space case treated in [1], the extraction of our bound is considerably more easy compared to [1] and even numerically better.

Notation: Let $h : \mathbb{N} \to \mathbb{N}$ be a number-theoretic function. Then for $K \in \mathbb{N} \cup \{0\}$ the function h^K denotes the K-th iterative of h, i.e.

$$h^0(n) := n, \text{ and } h^{K+1}(n) := h(h^K(n)).$$

2 Main results

Uniformly convex Banach spaces were introduced in 1936 by Clarkson in his seminal paper [3].

A Banach space X is called *uniformly convex* if for all $\varepsilon \in (0, 2]$ there exists $\delta \in (0, 1]$ such that for all $x, y \in X$,

$$\|x\| \leq 1, \quad \|y\| \leq 1 \text{ and } \|x - y\| \geq \varepsilon \text{ imply } \left\| \frac{1}{2}(x + y) \right\| \leq 1 - \delta. \tag{2}$$

A mapping $\eta : (0, 2] \to (0, 1]$ providing such a $\delta := \eta(\varepsilon)$ for given $\varepsilon \in (0, 2]$ is called a *modulus of uniform convexity.*

Since the condition (2) is empty for $\varepsilon > 2$ we can simply extend any such η to all strictly positive real numbers by stipulating $\eta'(\varepsilon) := \eta(\min(2, \varepsilon))$ if η is not already defined for $\varepsilon > 2$. We will make free use of this without further mentioning.

An example of a modulus of uniform convexity is Clarkson's *modulus of convexity* [3], defined for any Banach space X as the function $\delta_X : [0, 2] \to [0, 1]$ given by

$$\delta_X(\varepsilon) = \inf \left\{ 1 - \left\| \frac{x + y}{2} \right\| : \|x\| \leq 1, \|y\| \leq 1, \|x - y\| \geq \varepsilon \right\}. \tag{3}$$

It is easy to see that $\delta_X(0) = 0$ and that δ_X is nondecreasing. A well-known result is the fact that a Banach space X is uniformly convex if and only if $\delta_X(\varepsilon) > 0$ for $\varepsilon \in (0, 2]$. Note that for uniformly convex spaces X, δ_X is the largest modulus of uniform convexity.

The main result of our paper is a quantitative version of Birkhoff's generalization to uniformly convex Banach spaces of von Neumann's Mean Ergodic Theorem.

Theorem 2.1. *Assume that X is a uniformly convex Banach space, η is a modulus of uniform convexity and $T : X \to X$ is a linear operator with $\|Tx\| \leq \|x\|$ for all $x \in X$. Let $b > 0$. Then for all $x \in X$ with $\|x\| \leq b$,*

$$\forall \varepsilon > 0 \, \forall g : \mathbb{N} \to \mathbb{N} \, \exists P \leq \Phi(\varepsilon, g, b, \eta) \, \forall i, j \in [P, P + g(P)] \, \big(\|x_i - x_j\| < \varepsilon \big). \tag{4}$$

where (x_n) is the Cesaro means starting with x and

$$\Phi(\varepsilon, g, b, \eta) := M \cdot \tilde{h}^K(1), \tag{5}$$

with

$$
\begin{array}{l}
M := \left\lceil \dfrac{16b}{\varepsilon} \right\rceil, \gamma := \dfrac{\varepsilon}{16}\eta\left(\dfrac{\varepsilon}{8b}\right), \quad K := \left\lceil \dfrac{b}{\gamma} \right\rceil, \\
h,\, \tilde{h} : \mathbb{N} \to \mathbb{N},\ h(n) := 2(Mn + g(Mn)), \quad \tilde{h}(n) := \max\limits_{i \le n} h(i).
\end{array}
$$

If $\eta(\varepsilon)$ can be written as $\varepsilon \cdot \tilde{\eta}(\varepsilon)$ with $0 < \varepsilon_1 \le \varepsilon_2 \to \tilde{\eta}(\varepsilon_1) \le \tilde{\eta}(\varepsilon_2)$, then we can replace η by $\tilde{\eta}$ and the constant '16' by '8' in the definition of γ in the bound above.

Remark 2.2. *Note that our bound Φ is independent from T and depends on the space X and the starting point $x \in X$ only via the modulus of convexity η and the norm upper bound $b \ge \|x\|$. Moreover, it is easy to see that the bound depends on b and ε only via b/ε.*

As an immediate consequence of our theorem we get a quantitative version of von Neumann's Mean Ergodic Theorem.

Corollary 2.3. *Assume that X is a Hilbert space and $T : X \to X$ is a linear operator with $\|Tx\| \le \|x\|$ for all $x \in X$. Let $b > 0$. Then for all $x \in X$ with $\|x\| \le b$,*

$$
\forall \varepsilon > 0\, \forall g : \mathbb{N} \to \mathbb{N}\, \exists P \le \Phi(\varepsilon, g, b)\, \forall i, j \in [P, P + g(P)]\, \big(\|x_i - x_j\| < \varepsilon\big). \tag{6}
$$

where (x_n), Φ are defined as above, but with $K := \left\lceil \dfrac{512b^2}{\varepsilon^2} \right\rceil$.

Proof. It is well-known that as a modulus of uniform convexity of a Hilbert space X we can take $\eta(\varepsilon) := \varepsilon^2/8$ with $\tilde{\eta}(\varepsilon) := \varepsilon/8$ satisfying the requirements in the last claim of the theorem. $\square$

We get a similar result for L_p-spaces ($2 < p < \infty$), using the fact that $\eta(\varepsilon) = \dfrac{\varepsilon^p}{p2^p}$ is a modulus of uniform convexity for L_p (see e.g. [5]). Note that $\dfrac{\varepsilon^p}{p2^p} = \varepsilon \cdot \tilde{\eta}_p(\varepsilon)$ with $\tilde{\eta}_p(\varepsilon) = \dfrac{\varepsilon^{p-1}}{p2^p}$ satisfying the monotonicity condition in the theorem above.

Remark 2.4. *The bound extracted in* [1] *for Hilbert spaces is the following one*

$$
\begin{array}{l}
\Phi(\varepsilon, g, b) = h^K(1) \text{ where } h(n) = n + 2^{13}\rho^4\tilde{g}((n+1)\tilde{g}(2n\rho)\rho^2) \text{ with} \\
\rho = \left\lceil \frac{b}{\varepsilon} \right\rceil,\ K = 512\rho^2, \text{ and } \tilde{g}(n) = \max\limits_{i \le n}(i + g(i)).
\end{array}
$$

Note that the number of iterations K in both this bound and in our bound in Corollary 2.3 coincide (disregading the different placement of '$\lceil\cdot\rceil$') whereas the function h being iterated in our bound is much simper than that occurring in the above bound from [1].
The latter paper has an improved bound (roughly corresponding to our bound

for T being linear and nonexpansive) only in the special case where the linear operator T is an isometry. For this case, Avigad et al. [1] *show that one can take h as*

$$h(n) = n + 2^{13}\rho^4\tilde{g}((n+1)\tilde{g}(1)\rho^2)$$

which still is somewhat more complicated than the function h in our bound for the general case of T being nonexpansive.
From this, Avigad et al. [1] *obtain in the isometric case that $\Phi(\varepsilon, g, b) = 2^{O(\rho^2 \log\rho)}$ (with $\rho := \lceil \frac{b}{\varepsilon} \rceil$) for linear functions g, i.e. $g = O(n)$.*
Our bound in Corollary 2.3 generalizes this complexity upper bound on Φ to T being nonexpansive rather than being an isometry.

3 Technical lemmas

Lemma 3.1. *Let $(a_n)_{n\geq 0}$ be a sequence of nonnegative real numbers. Then*

(i) $\forall\varepsilon > 0\, \forall g : \mathbb{N} \to \mathbb{N}\, \exists N \leq \Theta(b, \varepsilon, g)\, \big(a_N \leq a_{g(N)} + \varepsilon\big)$,
where $\Theta(b, \varepsilon, g) := \max\limits_{i\leq K} g^i(1)$, $b \geq a_0$, $K := \left\lceil \dfrac{b}{\varepsilon} \right\rceil$. *Moreover,* $N = g^i(1)$ *for some* $i < K$.

(ii) $\forall\varepsilon > 0\, \forall g : \mathbb{N} \to \mathbb{N}\, \exists N \leq h^K(1)\, \forall m \leq g(N)\big(a_N \leq a_m + \varepsilon\big)$,
where $h(n) := \max\limits_{i\leq n} g(i)$ *and* b, K *are as above.*

Proof. (i) See e.g. [9, Lemma 6.3]

(ii) Let $\varepsilon > 0$, $g : \mathbb{N} \to \mathbb{N}$ and define

$$\tilde{g} : \mathbb{N} \to \mathbb{N}, \quad \tilde{g}(n) := \text{ the least } i \leq g(n) \text{ satisfying } a_i = \min\{a_j \mid j \leq g(n)\}.$$

Then, for all $n \in \mathbb{N}$ and for all $m \leq g(n)$, we have that $a_m \geq a_{\tilde{g}(n)}$. Applying now (i) for ε and $\tilde{g}$, we get that there exists $N \leq \Theta(b, \varepsilon, \tilde{g})$ such that $a_N \leq a_{\tilde{g}(N)} + \varepsilon \leq a_m + \varepsilon$ for all $m \leq g(N)$. Let us now define $h : \mathbb{N} \to \mathbb{N}$, $h(n) = \max\limits_{i\leq n} g(i)$. Then h is nondecreasing and $h(n) \geq g(n) \geq \tilde{g}(n)$ for all $n \in \mathbb{N}$. It is easy to see $h^i(n) \geq \tilde{g}^i(n)$ and $h^i(1) \geq h^{i-1}(1)$ for all $i, n \in \mathbb{N}$. Hence, $h^K(1) = \max\limits_{i\leq K} h^i(1) \geq \max\limits_{i\leq K} \tilde{g}^i(1) = \Theta(b, \varepsilon, \tilde{g}) \geq N$.

□

Lemma 3.2. *Let X be a uniformly convex Banach space and η be a modulus of uniform convexity. Define $u_\eta : (0, 2] \to (0, 1]$, $u_\eta(\varepsilon) = \dfrac{\varepsilon}{2} \cdot \eta(\varepsilon)$. Then for all $\varepsilon > 0$ and for all $x, y \in X$*

$$\|x\| \leq \|y\| \leq 1 \text{ and } \|x - y\| \geq \varepsilon \quad \text{imply} \quad \left\| \frac{1}{2}(x+y) \right\| \leq \|y\| - u_\eta(\varepsilon). \tag{7}$$

We use the notation u_X for u_{δ_X}, where δ_X is the modulus of convexity.
If $\eta(\varepsilon)$ can be written as $\varepsilon \cdot \tilde{\eta}(\varepsilon)$ with $0 < \varepsilon_1 \leq \varepsilon_2 \to \tilde{\eta}(\varepsilon_1) \leq \tilde{\eta}(\varepsilon_2)$, then we can replace u_η by $\tilde{u}_\eta(\varepsilon) := \varepsilon \cdot \tilde{\eta}(\varepsilon)$.

Proof. We have that $\frac{\|x\|}{\|y\|} \leq \frac{\|y\|}{\|y\|} = 1$ and $\frac{1}{\|y\|}\|x-y\| \geq \frac{\varepsilon}{\|y\|} \geq \varepsilon$, since $\|y\| \leq 1$. Applying the fact that η is a modulus of uniform convexity, we get that $\frac{1}{\|y\|}\left\|\frac{1}{2}(x+y)\right\| \leq 1-\eta(\varepsilon)$, hence

$$\left\|\frac{1}{2}(x+y)\right\| \leq \|y\| - \|y\|\eta(\varepsilon) \leq \|y\| - u_\eta(\varepsilon),$$

since $\|y\| \geq \frac{1}{2}(\|x\|+\|y\|) \geq \frac{1}{2}\|x-y\| \geq \frac{\varepsilon}{2}$.
The last claim follows from

$$\left\|\frac{1}{2}(x+y)\right\| \leq \|y\| - \|y\|\eta(\varepsilon/\|y\|) = \|y\| - \varepsilon\cdot\tilde{\eta}(\varepsilon/\|y\|) \leq \|y\| - \varepsilon\cdot\tilde{\eta}(\varepsilon).$$

□

The following lemma collects some facts already remarked by Birkhoff in his paper [2]. For completeness, we give the proofs here.

Lemma 3.3. *[2] Let X be a Banach space, $T : X \to X$ be linear and (x_n) be the Cesaro mean starting with x.*

(i) For all $n, k \in \mathbb{N}$,

$$x_{n+k} = \frac{n}{n+k}x_n + \frac{1}{n+k}\sum_{i=0}^{k-1} T^{n+i}x, \tag{8}$$

$$x_{kn} = \frac{1}{k}\sum_{i=0}^{k-1} T^{in}x_n, \tag{9}$$

$$x_{2kn} = \frac{1}{k}\sum_{i=0}^{k-1} \frac{1}{2}T^{in}\left(x_n + T^{kn}x_n\right). \tag{10}$$

(ii) Assume moreover that T satisfies $\|Tx\| \leq \|x\|$ for all $x \in X$. Then for all $n, k \in \mathbb{N}$,

$$\|x_{n+k} - x_n\| \leq \frac{2k\|x\|}{n+k}, \tag{11}$$

$$\|x_{kn} - x_n\| \leq \max_{i=0,\ldots,k-1} \|T^{in}x_n - x_n\| \tag{12}$$

Proof. (i) (8) is obvious, (9) and (10) are obtained by grouping terms:

$$\begin{aligned}
x_{kn} &= \frac{1}{kn}\sum_{j=0}^{nk-1} T^j x = \frac{1}{kn}\sum_{i=0}^{k-1}\left(T^{in}x + T^{in+1}x + \ldots + T^{in+(n-1)}x\right) \\
&= \frac{1}{k}\sum_{i=0}^{k-1} T^{in}\left(\frac{1}{n}(x+\ldots+T^{n-1}x)\right) = \frac{1}{k}\sum_{i=0}^{k-1} T^{in}x_n \\
x_{2kn} &= \frac{1}{2kn}\sum_{j=0}^{2nk-1} T^j x = \frac{1}{2kn}\sum_{i=0}^{k-1}\left(\sum_{j=0}^{n-1} T^{in+j}x + \sum_{j=0}^{n-1} T^{(k+i)n+j}x\right) \\
&= \frac{1}{k}\sum_{i=0}^{k-1}\frac{1}{2}T^{in}\left(\frac{1}{n}\sum_{j=0}^{n-1} T^j x + T^{kn}\left(\frac{1}{n}\sum_{j=0}^{n-1} T^j x\right)\right) \\
&= \frac{1}{k}\sum_{i=0}^{k-1}\frac{1}{2}T^{in}\left(x_n + T^{kn}x_n\right).
\end{aligned}$$

(ii) By assumption we have that $\|Ty\| \leq \|y\|$ for all $y \in X$, so $\|T^n x\| \leq \|x\|$ for all $n \in \mathbb{N}$ and, moreover, $\|x_n\| \leq \frac{1}{n}\sum_{i=0}^{n-1}\|T^i x\| \leq \|x\|$ for all $n \in \mathbb{N}$.

$$\begin{aligned}
\|x_{n+k} - x_n\| &\overset{(8)}{=} \left\|\left(\frac{n}{n+k}x_n + \frac{1}{n+k}\sum_{i=0}^{k-1} T^{n+i}x\right) - x_n\right\| \\
&= \left\|\frac{-k}{n+k}x_n + \frac{1}{n+k}\sum_{i=0}^{k-1} T^{n+i}x\right\| \\
&\leq \frac{2k}{n+k}\|x\|.
\end{aligned}$$

$$\begin{aligned}
\|x_{kn} - x_n\| &\overset{(9)}{=} \left\|\frac{1}{k}\sum_{i=0}^{k-1} T^{in}x_n - x_n\right\| = \left\|\frac{1}{k}\sum_{i=0}^{k-1}(T^{in}x_n - x_n)\right\| \\
&\leq \frac{1}{k}\sum_{i=0}^{k-1}\|T^{in}x_n - x_n\| \leq \max_{i=0,\ldots,k-1}\|T^{in}x_n - x_n\|.
\end{aligned}$$

□

4 Proof of Theorem 2.1

Let $x \in X$, $\varepsilon > 0$ and $g : \mathbb{N} \to \mathbb{N}$ be arbitrary and $\Phi, b, M, \gamma, K, h, \tilde{h}$ as in the hypotheses. Then $M \geq \frac{16b}{\varepsilon}$, that is $\frac{2b}{M} \leq \frac{\varepsilon}{8}$.

Let N be obtained by applying Lemma 3.1(ii) for the sequence $(\|x_n\|)_{n\geq 1}$ and the above γ and h. It follows that $0 < N \leq \tilde{h}^K(1)$ exists satisfying

$$\forall m \leq h(N)(\|x_N\| \leq \|x_m\| + \gamma). \tag{13}$$

Denote for all $k \in \mathbb{N}$,

$$y_k := \|T^{kN}x_N - x_N\|. \tag{14}$$

Claim: For all $k \leq \dfrac{h(N)}{2N}$, we have that $y_k \leq \dfrac{\varepsilon}{8}$.

Proof of claim: If $y_k = 0$, then it is obvious, so we can assume in the sequel that $y_k \neq 0$. We get that for all $k \in \mathbb{N}$

$$\left\|\frac{1}{b}T^{kN}x_N\right\| \leq \left\|\frac{1}{b}x_N\right\| \leq \frac{\|x\|}{b} \leq 1 \quad \text{and}$$

$$\frac{y_k}{b} = \left\|\frac{1}{b}(T^{kN}x_N - x_N)\right\| \leq \frac{1}{b}(\|T^{kN}x_N\| + \|x_N\|) \leq 2\frac{\|x\|}{b} \leq 2.$$

Thus, applying Lemma 3.2, we get that

$$\left\|\frac{1}{2b}(T^{kN}x_N + x_N)\right\| \leq \frac{1}{b}\|x_N\| - u_X\left(\frac{y_k}{b}\right), \tag{15}$$

that is

$$\left\|\frac{1}{2}(T^{kN}x_N + x_N)\right\| \leq \|x_N\| - bu_X\left(\frac{y_k}{b}\right) \tag{16}$$

for all $k \in \mathbb{N}$.

Using now (10) of Lemma 3.3, we obtain

$$\begin{aligned}
\|x_{2kN}\| &= \left\|\frac{1}{k}\sum_{i=0}^{k-1}\frac{1}{2}T^{iN}\left(x_N + T^{kN}x_N\right)\right\| \leq \frac{1}{k}\sum_{i=0}^{k-1}\left\|T^{iN}\left(\frac{1}{2}(x_N + T^{kN}x_N)\right)\right\| \\
&\leq \frac{1}{k}\sum_{i=0}^{k-1}\left\|\frac{1}{2}(x_N + T^{kN}x_N)\right\| = \left\|\frac{1}{2}(x_N + T^{kN}x_N)\right\| \\
&\leq \|x_N\| - bu_X\left(\frac{y_k}{b}\right).
\end{aligned}$$

On the other hand, applying (13), we get for $k \leq \dfrac{h(N)}{2N}$

$$\|x_{2kN}\| \geq \|x_N\| - \gamma.$$

Thus we must have that

$$bu_X\left(\frac{y_k}{b}\right) \leq \gamma \quad \text{for all} \quad k \leq \frac{h(N)}{2N}. \tag{17}$$

Assume that $y_k > \dfrac{\varepsilon}{8}$. Then, since δ_X is nondecreasing and $\delta_X \geq \eta$, we get that

$$bu_X\left(\frac{y_k}{b}\right) = b\cdot\frac{y_k}{2b}\cdot\delta_X\left(\frac{y_k}{b}\right) > \frac{\varepsilon}{16}\delta_X\left(\frac{\varepsilon}{8b}\right) \geq \frac{\varepsilon}{16}\eta\left(\frac{\varepsilon}{8b}\right) = \gamma, \tag{18}$$

that is a contradiction with (17). Hence, we must have $y_k \le \frac{\varepsilon}{8}$ for all $k \le \frac{h(N)}{2N}$. This finishs the proof of the claim.

Using the claim it follows that for all $0 < m \le \frac{h(N)}{2N}$ and $0 \le i < N$, we get that

$$\|x_{mN+i} - x_N\| \le \frac{2b}{m} + \frac{\varepsilon}{8}, \tag{19}$$

since

$$\begin{array}{rcll} \|x_{mN+i} - x_N\| & \le & \|x_{mN+i} - x_{mN}\| + \|x_{mN} - x_N\| & \\ & \le & \frac{2ib}{mN+i} + \|x_{mN} - x_N\|, & \text{by (11) and the fact that } \|x\| \le b \\ & < & \frac{2b}{m} + \|x_{mN} - x_N\|, & \text{since } 0 \le i < N \text{ implies} \frac{2i}{mN+i} < \frac{2}{m} \\ & \le & \frac{2b}{m} + \max\limits_{j=0,\ldots,m-1} y_j, & \text{by (12)} \\ & \le & \frac{2b}{m} + \frac{\varepsilon}{8} & \text{by the above claim.} \end{array}$$

Let us define $P := MN \le \Phi(\varepsilon, g, b, \eta)$ and take $j \in [P, P + g(P)]$. Then there are $q \in \mathbb{N}_0, 0 \le i < N$ such that $j - P = Nq + i$; moreover $Nq \le j - P \le g(P) = g(MN)$, so $q \le \frac{g(MN)}{N}$. It follows that

$$\begin{array}{rcl} \|x_j - x_P\| & = & \|x_{MN+Nq+i} - x_{MN}\| = \|x_{N(M+q)+i} - x_{MN}\| \\ & \le & \|x_{N(M+q)+i} - x_N\| + \|x_{MN} - x_N\| \\ & < & \frac{2b}{M+q} + \frac{\varepsilon}{8} + \frac{2b}{M} + \frac{\varepsilon}{8} \le \frac{\varepsilon}{4} + \frac{4b}{M} \le \frac{\varepsilon}{2}, \end{array}$$

since $M \le M + q \le M + \frac{g(MN)}{N} = \frac{h(N)}{2N}$, so we can apply (19) with $m := M$ and $m := M + q$.

It follows immediately that for all $j, l \in [P, P + g(P)]$, we have that

$$\|x_j - x_l\| \le \|x_j - x_P\| + \|x_l - x_P\| < \varepsilon.$$

The last claim of the theorem follows using the last claim in Lemma 3.2 with $\gamma := \frac{\varepsilon}{8}\tilde{\eta}\left(\frac{\varepsilon}{8b}\right)$ and $\tilde{u}_\eta$ instead of u_X. Then (18) needs to be replaced by

$$b \cdot \tilde{u}_\eta\left(\frac{y_k}{b}\right) = y_k \cdot \tilde{\eta}\left(\frac{y_k}{b}\right) > \frac{\varepsilon}{8}\tilde{\eta}\left(\frac{\varepsilon}{8b}\right) = \gamma.$$

Final remark on the extraction of the bound: The only ineffective principle used in Birkhoff's original proof is the fact that any sequence (a_n) of positive real numbers has an infimum (GLB). In our analysis we first replaced this analytical existential statement by a purely arithmetical one, namely

$$(\mathrm{GLB}_{ar}):\ \forall \varepsilon > 0\, \exists n \in \mathbb{N}\, \forall m \in \mathbb{N}\ (a_n \le a_m + \varepsilon).$$

This principle still is ineffective as there (in general) is no computable bound on '$\exists n \in \mathbb{N}$' (even for computable (a_n)). We then carried out (informally) a version of Gödel's functional interpretation by which (GLB_{ar}) gets replaced in the proof by the quantitative form provided in Lemma 3.1. For the general underlying facts from logic that guarantee this to be possible see [8].

References

[1] J. Avigad, P. Gerhardy, H. Towsner, Local stability of ergodic averages. To appear in: Trans. Amer. Math. Soc.

[2] G. Birkhoff, The mean ergodic theorem. Duke Math. J. 5 (1939), no. 1, 19-20.

[3] J.A. Clarkson, Uniformly convex spaces. Trans. Amer. Math. Soc. 40 (1936), 396-414.

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

[5] U. Kohlenbach, Uniform asymptotic regularity for Mann iterates. J. Math. Anal. Appl. 279 (2003), 531-544.

[6] U. Kohlenbach, Some logical metatheorems with application in functional analysis. Trans. Am. Math. Soc. 357 (2005), 89-128.

[7] Kohlenbach, U., Effective uniform bounds from proofs in abstract functional analysis. In: Cooper, B., Löwe, B., Sorbi, A. (eds.), New Computational Paradigms: Changing Conceptions of What is Computable. Springer Publisher, pp. 223-258 (2008).

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

[9] U. Kohlenbach, L. Leuştean, Asymptotically nonexpansive mappings in uniformly convex hyperbolic spaces. To appear in: J. of the European Math. Soc.

[10] T. Tao, Soft analysis, hard analysis, and the finite convergence principle. Essay posted May 23, 2007. Available at: http://terrytao.wordpress.com/2007/05/23/soft-analysis-hard-analysis-and-the-finite-convergence-principle/.

[11] T. Tao, Norm convergence of multiple ergodic averages for commuting transformations. Ergodic Theory Dynam. Systems 28 (2008), no. 2, 657-688.