BOUNDS FOR A NONLINEAR ERGODIC THEOREM FOR BANACH SPACES

ANTON FREUND AND ULRICH KOHLENBACH

ABSTRACT. We extract quantitative information (specifically, a rate of meta-stability in the sense of Terence Tao) from a proof due to Kazuo Kobayasi and Isao Miyadera, which shows strong convergence for Cesàro means of non-expansive maps on Banach spaces.

1. INTRODUCTION

Throughout this paper, we assume that X is a uniformly convex real Banach space, and that $C \subseteq X$ is non-empty, closed and convex. Furthermore, we assume that the map $T : C \to C$ has a fixed point and is nonexpansive (i.e., that we have $\|Tx - Ty\| \leq \|x - y\|$ for all $x, y \in C$). By

$$S_n x := \frac{1}{n} \sum_{i=0}^{n-1} T^i x$$

we denote the Cesàro means with respect to T. Our aim is a quantitative version of the following result due to K. Kobayasi and I. Miyadera:

Theorem 1.1 ([13]). *Consider $T : C \to C$ as above. Given $x \in C$, assume that the sequences $(\|T^n x - T^{n+i}x\|)_n$ converge uniformly in i. We then have*

$$\lim_{n \to \infty} \|y - S_n T^k x\| = 0 \quad \text{uniformly in } k,$$

for some fixed point y of T.

Let us discuss what a quantitative version of the conclusion should look like. The proof by Kobayasi and Miyadera shows that the fixed point y from the theorem is the limit of the sequence $(S_m T^m x)$. Hence we obtain

$$(1) \qquad \lim_{m,n \to \infty} \|S_m T^m x - S_n T^k x\| = 0 \quad \text{uniformly in } k,$$

$$(2) \qquad \lim_{n \to \infty} \|T^l S_n T^n x - S_n T^n x\| = 0 \quad \text{uniformly in } l.$$

To see that we get (2), note that $Ty = y$ and the fact that T is nonexpansive yield

$$\left\| T^l S_n T^n x - S_n T^n x \right\| \leq \left\| T^l S_n T^n x - T^l y \right\| + \left\| y - S_n T^n x \right\| \leq 2 \cdot \left\| y - S_n T^n x \right\|.$$

Indeed, the conjunction of (1) and (2) is equivalent to the conclusion of Theorem 1.1 together with the result that $(S_n T^n x)$ converges to a fixed point of T. With respect to (2), we note that uniformity in l is somewhat trivial in the presence of an actual fixed point y. It is less trivial when only approximate fixed points are available, and

---

this will play a role in our quantitative analysis (cf. the formulation of (5) below). Let us also point out that statements (1) and (2) entail

$$(3) \qquad \lim_{n \to \infty} \left\| T S_n T^k x - S_n T^k x \right\| = 0 \quad \text{uniformly in } k.$$

This asymptotic regularity result is due to R. Bruck [5] (in the Banach space case).

What, then, should a quantitative version of (1) assert? As a first idea, we might look for a rate of convergence, i.e., for a function $\varepsilon \mapsto N(\varepsilon)$ with

$$\left\| S_m T^m x - S_n T^k x \right\| < \varepsilon \quad \text{for all } m, n \geq N = N(\varepsilon) \text{ and all } k \in \mathbb{N}.$$

In particular, $\varepsilon \mapsto N(\varepsilon/2)$ would be a Cauchy rate for the sequence $(S_n x)$. It is known that such a rate cannot be computed (or given by a "simple closed expression") in general (see [1, Theorem 5.1]). However, we will be able to construct a rate of metastability, i.e., a map $(\varepsilon, g, h) \mapsto \Phi(\varepsilon, g, h)$ that takes an $\varepsilon > 0$ and functions $g, h : \mathbb{N} \to \mathbb{N}$ as input and ensures that we have

$$(4) \quad \left\| S_m T^m x - S_n T^k x \right\| < \varepsilon \quad \text{for some } N \leq \Phi(\varepsilon, g, h) \text{ and}$$
$$\text{all } m, n \in [N, N + g(N)] \text{ and } k \leq h(N).$$

Note that this is still as strong as (1) from above: for if the latter fails, there is an $\varepsilon > 0$ such that any N admits $m, n \geq N$ and $k \in \mathbb{N}$ with $\left\| S_m T^m x - S_n T^k x \right\| \geq \varepsilon$. If we set $g(N) := \max\{m, n\} - N$ and $h(N) := k$ for such numbers, (4) must fail. Similarly, our quantitative analysis of (2) will yield a map $(\varepsilon, g, h) \mapsto \Psi(\varepsilon, g, h)$ with

$$(5) \quad \left\| T^l S_n T^n x - S_n T^n x \right\| < \varepsilon \quad \text{for some } N \leq \Psi(\varepsilon, g, h) \text{ and}$$
$$\text{all } n \in [N, N + g(N)] \text{ and } l \leq h(N).$$

Following the structure of Kobayasi and Miyadera's proof, we will first construct Ψ and use it to define Φ. However, it will then turn out that one can switch to $\Psi := \Phi$, and that all desired properties can be satisfied simultaneously (see Theorem 1.2).

As mentioned, one cannot expect a computable rate of convergence (rather than metastability) for (1). It is not clear whether rates of convergence are available in the case of (2) or (3). The proof by Kobayasi and Miyadera [13] does not seem to yield such rates, while Bruck's [5] proof of (3) remains to be analyzed. In the case of Hilbert space, a rate of convergence for (3) is known (see [20, Lemma 3.4]).

The term "metastability" for statements such as (4) and (5) has been coined by T. Tao [31]. Even before Tao had introduced this terminology, the notion had been studied in mathematical logic, in particular in the proof mining program (see the textbook [17]), with foundational work reaching back to K. Gödel. One interesting aspect of metastability is its connection with the number of ε-fluctuations (see the results by J. Avigad and J. Rute [2, Section 5]). Both experience and general meta-theorems from logic (cf. the end of this section) show that rates of metastability can be extracted from a wide range of mathematical proofs. Specifically, the present paper complements quantitative work on nonlinear nonexpansive operators that satisfy a condition due to Wittmann [32] (cf. Section 5), notably by P. Safarik [30] (convergence of Cesàro means in Hilbert spaces) and the second author [19] (convergence of iterates of asymptotically regular maps in Banach spaces). Concerning complexity, it will be straightforward to see that our bounds are primitive recursive in the sense of S. Kleene (see e.g. [17, Section 3.4]), as in the cited results. More generally, there is a wealth of quantitative results on the convergence of various

iteration schemes in Hilbert, Banach and more general spaces (see e. g. [1, 23] for the linear and [15, 18, 20, 22, 24, 26, 29] for the nonlinear case).

As explained above, our quantitative analysis of Theorem 1.1 (due to Kobayasi and Miyadera [13]) consists in the construction of maps Φ and Ψ as in (4) and (5). In the following, we specify the quantitative data that we consider as given. First, we assume that we have a bound $b > 0$ with

$$(6) \qquad C \subseteq B_{b/2} := \left\{ x \in X \ \middle|\ \|x\| \leq \frac{b}{2} \right\}.$$

Note that such a bound exists if, and only if, C has bounded diameter (specifically (6) yields $\|x - y\| \leq b$ for $x, y \in C$). The version of Theorem 1.1 for bounded C is not actually weaker: To obtain the full theorem, recall the assumption that T has a fixed point f. For $x \in C$ and $r := \|x - f\|$, the set $D := \{x \in C \mid \|x - f\| \leq r\} \cap C$ is closed and convex with $T(D) \subseteq D \supseteq \{x, f\}$. In view of $D \subseteq B_{r+\|f\|}$, we can conclude by the seemingly weaker version of Theorem 1.1. In the context of Lemma 2.6 we will need a bound as in (6), not just a bound on the diameter.

Secondly, to witness our standing assumption on the Banach space X, we assume as given a modulus $\eta : (0, 2] \to (0, 1]$ of uniform convexity, so that we have

$$(7) \qquad \left\| \frac{x + y}{2} \right\| \leq 1 - \eta(\varepsilon) \quad \text{for } \|x\|, \|y\| \leq 1 \text{ with } \|x - y\| \geq \varepsilon.$$

Sometimes it is convenient to have η defined on $[0, \infty)$ with values in $[0, \infty)$ as a continuous function satisfying

$$(8) \qquad 0 < s < t \quad \Rightarrow \quad 0 = \eta(0) < \eta(s) < \eta(t) \text{ and } \frac{\eta(s)}{s} \leq \frac{\eta(t)}{t}.$$

This form is used in Section 2, which exhibits the quantitative content of an intermediate result by Bruck [6] and where it is shown how an arbitrary modulus as in (7) can be converted into a new modulus η' with the additional properties. Conversely, given $\eta' : [0, \infty) \to [0, \infty)$ one can simply take $\eta : (0, 2] \to (0, 1]$ with $\eta(\varepsilon) := \min\{1, \eta'(\varepsilon)\}$ (note in particular that (7) is trivial when $\varepsilon > 2$ and false or void when $\eta(\varepsilon) > 1$). The property (8) can equivalently be expressed as $\eta(s) = s \cdot \tilde{\eta}(s)$ for a non-decreasing and continuous $\tilde{\eta} : [0, \infty) \to [0, \infty)$ with $\tilde{\eta}(s) > 0$ for $s > 0$. Let us also recall that the L^p-spaces admit the natural modulus

$$\eta(\varepsilon) = \begin{cases} \frac{p-1}{8} \varepsilon^2 & \text{if } 1 < p < 2, \\ \frac{\varepsilon^p}{p \cdot 2^p} & \text{if } 2 \leq p < \infty, \end{cases}$$

for which (7) and (8) are satisfied (see [12] and also [14, Section 3]).

Thirdly, given that X is a uniformly convex Banach space, so is X^2 with the norm defined by $\|(x, y)\|_2 = (\|x\|^2 + \|y\|^2)^{1/2}$. Indeed, assume (7) holds for $(X, \|\cdot\|)$ and a modulus η that satisfies $\eta(\varepsilon) \leq \varepsilon/4$ and $\eta(\varepsilon) = \varepsilon \cdot \tilde{\eta}(\varepsilon)$ with non-decreasing $\tilde{\eta}$ (which e.g. follows if η satisfies (8)). An analysis of the proof in [7] shows that (7) remains valid when $(X, \|\cdot\|)$ and η are replaced by the space $(X^2, \|\cdot\|_2)$ and the modulus $\eta_2 : (0, 2] \to (0, 1]$ given by

$$\eta_2(\varepsilon) := \delta\left(\frac{\varepsilon}{8\sqrt{2}} \cdot \eta\left(\frac{\varepsilon}{\sqrt{2}} \right) \right) \quad \text{with} \quad \delta(\varepsilon) := \frac{\varepsilon^2}{8}.$$

Note that η_2 satisfies (8), if η does. Now the uniformly convex space $(X^2, \|\cdot\|_2)$ is, in particular, B-convex. By a characterization due to G. Pisier [28] (cf. the proof of [6, Theorem 1.1]), this means that there are $c > 0$ and $q > 1$ with the following

property: for all independent mean zero random variables $Z_1, \ldots, Z_n$ with values in X^2, the expected values satisfy

$$(9) \qquad \mathbb{E}\left(\left\|\sum_{i=1}^{n} Z_i\right\|_2^q\right) \leq c^q \cdot \sum_{i=1}^{n} \mathbb{E}\left(\|Z_i\|_2^q\right).$$

For our quantitative analysis, we assume that we are given such c and q. This additional data could be avoided, i.e., expressed in terms of our modulus η of uniform convexity (as guaranteed by the metatheorem cited below). Indeed one can explicitly construct suitable c, q in terms of η_2 via an analysis of the proof in [28]. In fact, since that proof only uses that X is uniformly nonsquare in the sense of James, it suffices to use one nontrivial value of η_2, e.g. $\eta_2(1)$. In the case of L^p-spaces one can take $q := \min\{2, p\}$, where the optimal constant c has been computed in [11] (see [27, Section 9.2]). We include nevertheless c, q among our input data, because this simplifies matters and is computationally harmless: Note that c and q depend on the space only. Also, the complexity class of our bounds does not depend on c and q, because the latter are numbers rather than functions.

Finally, for given $x \in C$ we abbreviate

$$\alpha_n^i := \left\|T^n x - T^{n+i} x\right\|.$$

Since T is nonexpansive, we always have $\alpha_n^i \geq \alpha_{n+1}^i \geq 0$, so that each of the sequences (α_n^i) converges. A central assumption of Theorem 1.1 demands that the rate of convergence (but not necessarily the limit) is independent of the number i. In terminology due to Bruck [4], this means that T is asymptotically isometric on the set $\{x\}$. As a quantitative version of this assumption, we suppose that we are given a rate of metastability, i.e., a map $(\varepsilon, g, h) \mapsto A(\varepsilon, g, h)$ that guarantees

$$(10) \quad \left|\alpha_m^i - \alpha_n^i\right| < \varepsilon \quad \text{for some } N \leq A(\varepsilon, g, h)$$
$$\text{and all } m, n \in [N, N + g(N)] \text{ and } i \leq h(N).$$

In order to apply our quantitative result, one will have to provide a map A with this property. We mention three situations where this is possible: First, assume that T satisfies Wittmann's [32] condition $\|Tx + Ty\| \leq \|x + y\|$ (which is e.g. the case when $C = -C$ and T is odd in addition to being nonexpansive) and is asymptotically regular with given rate ($\left\|T^{n+1}x - T^n x\right\| \to 0$ for $n \to \infty$, which holds e.g. for averaged maps). As shown by the second author (see [19] and the generalization in [21, Section 3]), one can then construct a rate of metastability that witnesses $\left\|T^i x - T^j x\right\| \to 0$ for $i, j \to \infty$. Such a rate is readily transformed into a map A that validates (10). Secondly, the assumption that T is asymptotically regular can be dropped in the Hilbert space case. Finally, one can satisfy (10) when $(T^n x)$ has a convergent subsequence (even in Banach space). A quantitative analysis of the second and third situation (which are mentioned by Kobayasi and Miyadera [13]) is given in Section 5 of our paper. We point out that all three constructions of A yield a rate of metastability rather than convergence. In this respect, it is also interesting to consider the beginning of Section 4, where a "limsup $\leq$ liminf"-argument from the proof of [13, Lemma 2] forces us to settle for metastability.

Based on the preceding discussion, we now state the main result of our paper, which will be proved at the end of Section 4. Note that we simultaneously satisfy statements (4) and (5) as well as the metastable version of (3). As explained above,

the original result of Kobayasi and Miyadera [13] (stated as Theorem 1.1) can be recovered as an immediate consequence of the following quantitative version.

Theorem 1.2. *Consider a uniformly convex real Banach space X, a subset $C \subseteq X$ that is non-empty, closed and convex, and a nonexpansive map $T : C \to C$ that has a fixed point. Assume we have maps A, η and numbers $b, c > 0$ and $q > 1$ that satisfy (6-9) and (10), the latter for given $x \in C$. Depending on this data only, we get a map $(\varepsilon, g, h) \mapsto \Phi(\varepsilon, g, h)$, explicitly given in Definition 4.11, such that the following holds: For any $\varepsilon > 0$ and $g, h : \mathbb{N} \to \mathbb{N}$ there is an $N \leq \Phi(2\varepsilon/5, g, h)$ with*

$$\max \left\{ \left\| S_m T^m x - S_n T^k x \right\|, \left\| T^l S_n T^n x - S_n T^n x \right\|, \left\| T^l S_n T^k x - S_n T^k x \right\| \right\} < \varepsilon$$

for all $m, n \in [N, N + g(N)]$ and $k, l \leq h(N)$.

To conclude the introduction, we return to the topic of logical metatheorems. In order to determine the precise bounds Φ and Ψ, it is of course necessary to consider the proof of Theorem 1.1 in detail. However, the fact that one can extract suitable Φ and Ψ is guaranteed in advance, by the second author's general result [16, Theorem 3.30] on uniformly convex normed linear spaces. We only sketch why the latter applies (cf. [30, Section 3] for a detailed discussion in a related case): The cited metatheorem covers, roughly speaking, results of the form "for all–exists". A convergence statement such as (1) does not have this form, as the existential claim ("there is an N") is followed by a universal quantification ("for all $m, n \geq N$"). On the other hand, the metastable version "for all ε, g, h there is a $\Phi(\varepsilon, g, h)$ as in (4)" does have the required form; here it is crucial that the quantification over k, m, n and N inside (4) "does not count", because it only refers to numbers below a given bound. Similarly, the assumption associated with (10) contains essentially no existential quantification when we treat A as a given function (which we may assume to be a majorant, namely of the function A^- providing the least N satisfying (10)). In this situation, the cited metatheorem predicts that there are computable maps Φ and Ψ that only depend on our bound b with $C \subseteq B_{b/2}$, on the given function A, and on the modulus η of uniform convexity. Furthermore, the proof of the metatheorem suggests a general strategy for the extraction of Φ and Ψ. Hence the logical background is useful in practice and interesting as a uniform explanation. At the same time, each concrete application can be presented without any reference to logic, as the following sections testify.

Remark 1.3 (For logicians)**.** Officially, the aforementioned metatheorem requires A to be a **strong** majorant for some A^- satisfying (10). However, strong majorization is only needed when dealing with proofs whose quantitative analysis requires so-called bar recursion, which is not the case here, while otherwise ordinary majorization can be used in the monotone functional interpretation proving the metatheorem (see [17], Remark 17.37).

2. Nonlinearity and convex combinations

In this section, we discuss quantitative aspects of a result due to Bruck [6]. Specifically, we construct an increasing function $\gamma : [0, \infty) \to [0, \infty)$ such that

$$(11) \qquad \gamma \left(\left\| T \left(\sum_{i=1}^{n} \lambda_i x_i \right) - \sum_{i=1}^{n} \lambda_i T x_i \right\| \right) \leq \max_{1 \leq i, j \leq n} \left(\| x_i - x_j \| - \| T x_i - T x_j \| \right)$$

holds for any convex combination $\sum \lambda_i x_i$ (i.e., we require $\lambda_i \geq 0$ and $\sum \lambda_i = 1$). We note that most quantitative information in this section is already quite explicit in Bruck's original presentation. Nevertheless, it will be important to streamline some constructions for our purpose (cf. the paragraph after Definition 2.4).

Bruck first constructs functions $\gamma = \gamma_n$ that satisfy (11) for fixed n. In a second step, he achieves independence of n by diagonalizing over these functions. A more common way to assert (11) for $n = 2$ is to say that T is of type (γ). This reveals that the case $n = 2$ coincides with [5, Lemma 1.1]. The following makes the computational information explicit. By a standing assumption from the introduction, the function $\eta : [0, \infty) \to [0, \infty)$ is a modulus of uniform convexity for X, while $b > 0$ bounds the diameter of $C \subseteq X$, the domain of our map $T : C \to C$.

Definition 2.1. Let $\gamma_2 : [0, \infty) \to [0, \infty)$ be given by $\gamma_2(t) = \min\left\{t, \frac{b}{2} \cdot \eta\left(\frac{4t}{b}\right)\right\}$.

The next lemma shows that (11) holds for $\gamma = \gamma_2$ and fixed $n = 2$. The additional properties ensure that we have a strictly increasing inverse $\gamma_2^{-1} : [0, \infty) \to [0, \infty)$ (with $\gamma_2^{-1}(t) \geq t$ due to the minimum above), as used in the proof of Lemma 2.5.

Lemma 2.2. *The function γ_2 is strictly increasing, unbounded and continuous with minimal value $\gamma_2(0) = 0$. For all $x_1, x_2 \in C$ and $\lambda \in [0,1]$ we have*

$$\gamma_2(\|T(\lambda x_1 + (1-\lambda)x_2) - (\lambda T x_1 + (1-\lambda)T x_2)\|) \leq \|x_1 - x_2\| - \|Tx_1 - Tx_2\|.$$

Proof. The first sentence of the lemma is immediate by the corresponding properties of η, which hold by a standing assumption from the introduction (note that (8) yields $\eta(t) \geq \eta(1) \cdot t$ for $t \geq 1$). For the remaining claim, we follow the proof of [5, Lemma 1.1]. As in the latter, the value of

$$(12) \quad 2\lambda(1-\lambda) \cdot \|x_1 - x_2\| \cdot \eta\left(\frac{\|\lambda T x_1 + (1-\lambda)T x_2 - T(\lambda x_1 + (1-\lambda)x_2)\|}{\lambda(1-\lambda) \cdot \|x_1 - x_2\|}\right)$$

is smaller than or equal to $\|x_1 - x_2\| - \|Tx_1 - Tx_2\|$ (unless $\lambda \in \{0, 1\}$ or $x_1 = x_2$ and the claim is trivial). Now recall the standing assumption that η is convex. More specifically, from (8) we readily get $t \cdot \eta(r/t) \leq s \cdot \eta(r/s)$ for $r \geq 0$ and $0 < s \leq t$. With $s = \lambda(1-\lambda) \cdot \|x_1 - x_2\| \leq b/4 = t$, we see that (12) is larger than or equal to

$$\frac{b}{2} \cdot \eta\left(\frac{4 \cdot \|\lambda T x_1 + (1-\lambda)T x_2 - T(\lambda x_1 + (1-\lambda)x_2)\|}{b}\right).$$

Hence the definition of γ_2 (even without the minimum) is as required.   □

We have reproduced part of the proof from [5] in order to show how the convexity of η is used. As promised in the introduction, we now recall how a convex modulus can be constructed.

Remark 2.3. Assume that $\eta_0 : (0, 2] \to (0, 1]$ is any modulus of uniform convexity for our Banach space X, which means that (7) holds with η_0 at the place of η. Define $\eta_1 : (0, 2] \to (0, 1]$ by setting

$$\eta_1(\varepsilon) = \sup\{\eta_0(\varepsilon') \mid \varepsilon' \in (0, \varepsilon]\}.$$

Then (7) does still hold with η_1 at the place of η (since $\eta_1(\varepsilon) > 1 - \|(x - y)/2\|$ entails $\eta_0(\varepsilon') > 1 - \|(x - y)/2\|$ for some $\varepsilon' \leq \varepsilon$). The point is that η_1 is increasing (not necessarily strictly). We also write $\eta_1 : [0, \infty) \to [0, 1]$ for the extension of this function by the values $\eta_1(0) = 0$ and $\eta_1(\varepsilon) = \eta_1(2)$ for $\varepsilon > 2$. Then η_1 is still

increasing and hence Riemann integrable. As in the proof of [5, Lemma 1.1], we now define $\eta : [0, \infty) \to [0, \infty)$ by

$$\eta(\varepsilon) = \frac{1}{2} \cdot \int_0^\varepsilon \eta_1(t) \, dt.$$

For $\varepsilon \in (0, 2]$ we have $\eta(\varepsilon) \leq \varepsilon \cdot \eta_1(\varepsilon)/2 \leq \eta_1(\varepsilon)$, so that (7) holds for η. Note that this extends to all $\varepsilon \in [0, \infty)$, by $\eta(0) = 0$ and the trivial reason mentioned in the introduction. Furthermore, η is strictly increasing and continuous. For $0 < s \leq t$ we can split the integral to get $2 \cdot \eta(t) \geq 2 \cdot \eta(s) + (t - s) \cdot \eta_1(s)$ and then

$$2 \cdot (s \cdot \eta(t) - t \cdot \eta(s)) \geq (t - s) \cdot (s \cdot \eta_1(s) - 2 \cdot \eta(s)) \geq 0.$$

We thus have $\eta(s)/s \leq \eta(t)/t$, as required by the convexity property from (8). In conclusion, η satisfies all standing assumptions from the introduction. To estimate the cost of these assumptions, we observe $\eta_1(\varepsilon) \leq 2 \cdot \eta(2\varepsilon)/\varepsilon$ for $\varepsilon \in (0, 2]$. As noted in the introduction, the given construction is not required in the case of L^p-spaces with $1 < p < \infty$, where the specific η given does already satisfy (7) and (8).

We will see that the following functions validate (11) for arbitrary but fixed n.

Definition 2.4. We construct functions $\gamma_n : [0, \infty) \to [0, \infty)$ by recursion on $n \geq 2$. The base case is provided by Definition 2.1, while the step is given by

$$\gamma_{n+1}(t) = \min \left\{ \gamma_n(t), \gamma_2 \left(\frac{\gamma_n(t/2)}{3} \right) \right\}.$$

Let us point out that we do not make the functions γ_n convex, as this property will not be needed beyond the proof of Lemma 2.2. This allows us to give a somewhat simpler definition than Bruck. One other point is important: The proof of [6, Lemma 2.1] suggests a recursive construction of γ_n^{-1} rather than γ_n. We will consider inverses in the verification below, but we have avoided them in the construction itself, because this gives more control on the complexity of our bounds. Indeed, a function and its inverse do not generally belong to the same complexity class. As an example for logicians, we mention that the function $F_{\varepsilon_0}^{-1}$ in [9] is primitive recursive while F_{ε_0} is not.

Lemma 2.5. *The functions γ_n are strictly increasing, unbounded and continuous with $\gamma_n(0) = 0$. For each fixed $n \geq 2$, inequality (11) is valid for $\gamma = \gamma_n$.*

Proof. A straightforward induction over n yields the first sentence of the lemma. The point is that we can now consider the inverses $\gamma_n^{-1} : [0, \infty) \to [0, \infty)$, which are strictly increasing as well. By the proof of [6, Lemma 2.1], the claim that (11) holds for any convex combination (in the domain of $T : C \to C$) reduces to

$$\gamma_{n+1}^{-1}(s) \geq \gamma_2^{-1}(s) + \gamma_n^{-1}(s + 2 \cdot \gamma_2^{-1}(s)).$$

Given that Definition 2.1 forces $\gamma_2(r) \leq r$, we inductively get $\gamma_{n+1}(r) \leq \gamma_n(r) \leq r$ and hence $r \leq \gamma_n^{-1}(r)$. This ensures that we have

$$\gamma_2^{-1}(s) + \gamma_n^{-1}(s + 2 \cdot \gamma_2^{-1}(s)) \leq 2 \cdot \gamma_n^{-1}(3 \cdot \gamma_2^{-1}(s)) =: t.$$

We can thus conclude

$$\gamma_{n+1}(\gamma_2^{-1}(s) + \gamma_n^{-1}(s + 2 \cdot \gamma_2^{-1}(s))) \leq \gamma_{n+1}(t) \leq \gamma_2 \left(\frac{\gamma_n(t/2)}{3} \right) = s.$$

Since γ_{n+1}^{-1} is increasing, this yields the open claim. $\square$

In order to obtain (11) for a function γ that is independent of n, Bruck argues that any convex combination can be approximated by one with a bounded number of summands. More specifically, this observation is applied to convex combinations of elements $(x, Tx) \in C \times C \subseteq X^2$. Here X^2 is a uniformly convex Banach space with norm given by $\|(x, y)\|_2 = (\|x\|^2 + \|y\|^2)^{1/2}$, as discussed in the introduction. For a number $p \geq 1$ and a subset $M \subseteq X^2$, we put

$$\text{co}_p(M) = \left\{ \sum_{i=1}^n \lambda_i z_i \; \middle| \; z_i \in M \text{ and } \lambda_i \geq 0 \text{ with } \sum_{i=1}^n \lambda_i = 1 \text{ for } n \leq p \right\}.$$

Let us note that we can always arrange $n = p$ by repeating some of the z_i. Also write $\text{co}(M)$ for the full convex hull (i.e., the union over all $\text{co}_p(M)$ with $p \geq 1$), and recall that $B_r = \{z \in X^2 \mid \|z\|_2 \leq r\}$ is the closed ball with radius r. By a standing assumption, we have $c > 0$ and $q > 1$ that validate (9) from the introduction. The proof of [6, Theorem 1.1] contains the following computational information.

Lemma 2.6. *We have* $\text{co}(M) \subseteq \text{co}_p(M) + B_{r \cdot \varepsilon}$ *when* $M \subseteq B_r$ *and* $2cp^{(1-q)/q} \leq \varepsilon$.

Proof. In the proof of [6, Theorem 1.1], Bruck seems to claim that the result holds for $2cp^{1/q} < \varepsilon$. This appears counterintuitive, since $p \mapsto 2cp^{1/q}$ is strictly increasing while $\text{co}_p(M)$ grows with p. We recall Bruck's proof to show that it actually yields our bound (cf. the proof of Theorem 6.2 in [8] for a similar reasoning). By a straightforward rescaling we may assume $r = 1$. Consider any convex combination

$$z = \sum_{i=1}^n \lambda_i z_i \in \text{co}(M) \subseteq B_1.$$

For p as in the lemma, consider independent X^2-valued random variables $Z_1, \ldots, Z_p$ with identical distribution given by

$$Z_1, \ldots, Z_p \stackrel{\text{iid}}{\sim} \frac{Z - z}{p} \quad \text{for} \quad Z = z_i \text{ with probability } \lambda_i.$$

Now $(2/p)^q$ bounds all possible values of $\|Z_i\|_2^q$, and hence its expectation. By (9) we can conclude

$$\mathbb{E}\left(\left\| \sum_{j=1}^p Z_j \right\|_2^q \right) \leq (2c)^q \cdot p^{1-q}.$$

In particular, the right side must bound at least one possible value of $\|\sum Z_j\|_2^q$. More explicitly, there must be some event $\omega : \{1, \ldots, p\} \to \{1, \ldots, n\}$ (under which Z_j assumes value $(z_{\omega(j)} - z)/p$) such that we have

$$\left\| z - \sum_{j=1}^p \frac{1}{p} \cdot z_{\omega(j)} \right\|_2 = \left\| \sum_{j=1}^p \frac{z_{\omega(j)} - z}{p} \right\|_2 \leq 2cp^{\frac{1-q}{q}} \leq \varepsilon.$$

This shows that our given $z \in \text{co}(M)$ lies in $\text{co}_p(M) + B_\varepsilon$, as desired. $\qquad \square$

Following the construction by Bruck, we now diagonalize over the functions γ_n, in order to achieve independence of n.

Definition 2.7. Let $\gamma : [0, \infty) \to [0, \infty)$ be given by $\gamma(0) = 0$ and, for $t > 0$,

$$\gamma(t) = \gamma_{p(t)}\left(\frac{t}{3} \right) \quad \text{with} \quad p(t) = \max\left\{ 2, \left\lceil \left(\frac{6bc}{\sqrt{2t}} \right)^{q/(q-1)} \right\rceil \right\}.$$

The next proof is very close to the one of [6, Theorem 2.1]. We provide details in order to show how the previous constructions come together.

Proposition 2.8. *The function γ is strictly increasing and validates (11) for any convex combination of elements $x_1, \ldots, x_n \in C$ with arbitrary $n \geq 1$.*

Proof. The function $t \mapsto p(t)$ is decreasing, as $q > 1$ holds by a standing assumption from the introduction. Since Definition 2.4 ensures $\gamma_{n+1}(t) \leq \gamma_n(t)$, this yields

$$s < t \quad \Rightarrow \quad \gamma(s) = \gamma_{p(s)}\left(\frac{s}{3}\right) \leq \gamma_{p(t)}\left(\frac{s}{3}\right) < \gamma_{p(t)}\left(\frac{t}{3}\right) = \gamma(t).$$

To establish (11), consider a convex combination $\sum \lambda_i x_i$ of $x_1, \ldots, x_n \in C$. We set

$$t := \left\| T\left(\sum_{i=1}^n \lambda_i x_i\right) - \sum_{i=1}^n \lambda_i T x_i \right\|$$

and assume $t > 0$, as the remaining case is trivial. The crucial idea (from the proof by Bruck [6]) is to apply Lemma 2.6 to $M = \{z_1, \ldots, z_n\}$ with $z_i = (x_i, T x_i)$. By standing assumption (6) we have $C \subseteq B_{b/2}$ in X and hence $M \subseteq C \times C \subseteq B_{b/\sqrt{2}}$ in X^2 (recall $\|(x,y)\|_2 = (\|x\|^2 + \|y\|^2)^{1/2}$). One readily checks

$$\varepsilon := \frac{\sqrt{2}t}{3b} \geq 2c \cdot p(t)^{(1-q)/q}.$$

For this ε and with $p = p(t)$, Lemma 2.6 yields a subset $\{z_{i_1}, \ldots, z_{i_p}\} \subseteq M$ and coefficients $\mu_1, \ldots, \mu_p \geq 0$ with $\sum \mu_j = 1$ such that

$$\left\| \sum_{i=1}^n \lambda_i z_i - \sum_{j=1}^p \mu_j z_{i_j} \right\|_2 \leq \frac{t}{3}.$$

In view of $\|x\|, \|y\| \leq \|(x,y)\|_2$ and since T is nonexpansive, this yields

$$\left\| T\left(\sum_{i=1}^n \lambda_i x_i\right) - T\left(\sum_{j=1}^p \mu_j x_{i_j}\right) \right\| \leq \frac{t}{3} \quad \text{and} \quad \left\| \sum_{i=1}^n \lambda_i T x_i - \sum_{j=1}^p \mu_j T x_{i_j} \right\| \leq \frac{t}{3}.$$

Using the triangle inequality, we can conclude

$$t \leq \frac{t}{3} + \left\| T\left(\sum_{j=1}^p \mu_j x_{i_j}\right) - \sum_{j=1}^p \mu_j T x_{i_j} \right\| + \frac{t}{3}.$$

Hence the remaining norm on the right must be at least $t/3$. By Lemma 2.5 we get

$$\gamma(t) = \gamma_p\left(\frac{t}{3}\right) \leq \max_{1 \leq k,l \leq p} (\|x_{i_k} - x_{i_l}\| - \|T x_{i_k} - T x_{i_l}\|).$$

This suffices to conclude, as the maximum on the right becomes larger when we admit all $i, j \in \{1, \ldots, n\}$, rather than those of the form $i = i_k$ and $j = i_l$ only. $\quad \square$

3. Nonlinearity and Cesàro means

Recall that $S_n x$ denotes the Cesàro mean $(T^0 x + \ldots + T^{n-1} x)/n$. If T is nonlinear, then it may fail to commute with S_n. This failure can be measured by

$$\beta_n^l := \left\| S_n T^{l+n} x - T^l S_n T^n x \right\|.$$

More generally, we can recover these values as $\beta_n^l = \beta_{n,n}^l$ for

$$\beta_{m,n}^l := \left\| \frac{S_m T^{l+m} x + S_n T^{l+n} x}{2} - T^l \left(\frac{S_m T^m x + S_n T^n x}{2} \right) \right\|.$$

Kobayasi and Miyadera [13, Lemma 1] show that, under the assumptions made in our Theorem 1.1 (which is their Theorem 1), we have

$$\lim_{n \to \infty} \beta_n^l = \lim_{m,n \to \infty} \beta_{m,n}^l = 0.$$

In the present section, we establish a rate of metastability for this result. Let us recall that we have only assumed a rate of metastability rather than convergence in (10). If we had assumed a rate of convergence there, then we would get one for the $\beta_{m,n}^l$ here (simply read $g(N) = \infty = h(N)$ below). However, we would then be forced to downgrade from convergence to metastability later, as explained in the introduction. In view of this fact, it is reasonable to work with a rate of metastability from the outset, as this makes for a computationally weaker assumption. Concerning the next definition, recall that our standing assumptions provide a map $(\varepsilon, g, h) \mapsto A(\varepsilon, g, h)$ that validates (10). The function γ is given by Definition 2.7.

Definition 3.1. For $\varepsilon > 0$ and $g, h : \mathbb{N} \to \mathbb{N}$ we put $B(\varepsilon, g, h) := A(\gamma(\varepsilon), g', h')$ with $g'(N) := N + 2 \cdot g(N) + h(N)$ and $h'(N) := 2 \cdot (N + g(N))$.

By analyzing the proof of [13, Lemma 1], we see the following:

Lemma 3.2. *For any $\varepsilon > 0$ and $g, h : \mathbb{N} \to \mathbb{N}$, there is an $N \leq B(\varepsilon, g, h)$ with*

$$\beta_{m,n}^l \leq \varepsilon \quad \text{for all } m, n \in [N, N + g(N)] \text{ and all } l \leq h(N).$$

In particular we get $\beta_n^l \leq \varepsilon$ under the same conditions on n and l.

To be precise, we point out that $\beta_{m,n}^l$ is only defined for $m, n > 0$. In order to avoid this condition, one can simply declare $\beta_{0,n}^l = 0 = \beta_{m,0}^l$.

Proof. We first observe that Proposition 2.8 remains valid with $T^l : C \to C$ at the place of $T : C \to C$, with the same $\gamma : [0, \infty) \to [0, \infty)$ for all $l \in \mathbb{N}$. Indeed, a glance at Definitions 2.1, 2.4 and 2.7 reveals that the relevant quantitative information depends on our Banach space X and the bounded subset $C \subseteq X$ only (cf. the standing assumptions (6-10) from the introduction). The qualitative assumption that T is nonexpansive holds of T^l as well. As in the proof of [13, Lemma 1], we now apply (11) with T^l and $m + n$ at the place of T and n, respectively, and with

$$\begin{cases} x_i = T^{m+i-1} x & \text{and} \quad \lambda_i = 1/(2m) \quad \text{for} \quad 1 \leq i \leq m, \\ x_i = T^{n+i-(m+1)} x & \text{and} \quad \lambda_i = 1/(2n) \quad \text{for} \quad m+1 \leq i \leq m+n. \end{cases}$$

One readily checks that this yields

$$\sum_{i=1}^{m+n} \lambda_i x_i = \frac{S_m T^m x + S_n T^n x}{2} \quad \text{and} \quad \sum_{i=1}^{m+n} \lambda_i T^l x_i = \frac{S_m T^{l+m} x + S_n T^{l+n} x}{2}.$$

Hence inequality (11) amounts to

$$\gamma(\beta_{m,n}^l) \leq \max\left\{\|x_i - x_j\| - \left\|T^l x_i - T^l x_j\right\| \,:\, 1 \leq i,j \leq m+n\right\}.$$

Writing $\alpha_n^i = \left\|T^n x - T^{n+i} x\right\|$ as in the introduction, we get

(13) $\qquad \gamma(\beta_{m,n}^l) \leq \max\{\alpha_k^p - \alpha_{l+k}^p \,|\, \min\{m,n\} \leq k < 2 \cdot \max\{m,n\} > p\}.$

Let g' and h' be given as in Definition 3.1. By a standing assumption from the introduction, pick an $N \leq A(\gamma(\varepsilon), g', h')$ such that (10) holds with $\gamma(\varepsilon), g'$ and h' at the place of ε, g and h, respectively. To establish the present lemma, consider arbitrary $m, n \in [N, N + g(N)]$ and $l \leq h(N)$. For all k, l, p as in (13) we have

$$N \leq k \leq k + l < 2 \cdot (N + g(N)) + h(N) = N + g'(N)$$

and $p < h'(N)$. Hence the current version of (10) yields $\alpha_k^p - \alpha_{l+k}^p < \gamma(\varepsilon)$ in all cases that are relevant for (13). The latter thus entails $\gamma(\beta_{m,n}^l) < \gamma(\varepsilon)$. Proposition 2.8 tells us that γ is strictly increasing, so that we get $\beta_{m,n}^l < \varepsilon$ as desired. $\qquad\square$

4. Cesàro means and fixed points

This section completes the quantitative analysis of Kobayasi and Miyadera's results in [13], as sketched in the introduction.

We first analyze the preliminary result from Lemma 2 of [13], which asserts that

$$\theta_n^f := \|S_n T^n x - f\|$$

converges whenever $f = Tf$ is a fixed point. From a methodological standpoint, this can be seen as the most interesting part of our quantitative analysis: it is here that we are forced to extract a rate of metastability rather than convergence. In order to make this transparent, we recall the original proof of [13, Lemma 2]. For a fixed point $f = Tf$ and arbitrary $m \geq 1$ and $n \in \mathbb{N}$, the cited proof shows

(14) $\qquad \theta_{m+n}^f \leq \theta_m^f + \dfrac{m-1}{m+n} \cdot \|x - f\| + \dfrac{1}{m+n} \cdot \displaystyle\sum_{i=0}^{m+n-1} \beta_m^{n+i},$

with $\beta_n^l = \left\|S_n T^{l+n} x - T^l S_n T^n x\right\|$ as in the previous section. The proof then concludes with a "limsup $\leq$ liminf"-argument, which can be spelled out as follows: Write θ^f for the limit inferior of the sequence (θ_n^f). According to [13, Lemma 1] we have $\lim_{n \to \infty} \beta_n^l = 0$ uniformly in l (cf. our Lemma 3.2). Given $\varepsilon > 0$, there must thus be a number $m \in \mathbb{N}$ such that we have

(i) $\theta_k^f \geq \theta^f - \varepsilon/4$ for $k \geq m$, (ii) $\beta_m^l \leq \varepsilon/4$ for $l \in \mathbb{N}$, (iii) $\theta_m^f \leq \theta^f + \varepsilon/4$.

Put $N := \max\{m, \lceil 4(m-1) \cdot \|x - f\| / \varepsilon \rceil\}$. For $j \geq N$ (think $j = m+n$) we can combine (14) and (ii) to get $\theta_j^f \leq \theta_m^f + \varepsilon/2$, which by (iii) yields $\theta_j^f \leq \theta^f + 3\varepsilon/4$. Together with (i) we get $|\theta_j^f - \theta_k^f| \leq \varepsilon$ for $j, k \geq N$, as needed to show that (θ_n^f) is a Cauchy sequence. Our focus on the Cauchy property (rather than on convergence to the limit θ^f) assimilates the argument to the quantitative analysis below.

From a computational viewpoint, the previous argument appears problematic, because we do not know how fast the limit inferior is approximated. More explicitly, there is no obvious bound on a number m that would satisfy (i) or (iii) above. Nevertheless, a modified argument does reveal quantitative information: If (ii) holds for sufficiently many l (which can be ensured by Lemma 3.2), then we obtain $\theta_j^f \leq \theta_m^f + \varepsilon/2$ as above. Without reference to θ^f, we can conclude that $|\theta_j^f - \theta_k^f| \leq \varepsilon$

holds unless we have $\theta_k^f < \theta_m^f - \varepsilon/2$. In the latter case, we repeat the argument with k at the place of m. Crucially, there can only be finitely many repetitions of this type, as θ_n^f cannot become negative. For an explicit bound, use $f = Tf$ to get

$$(15) \qquad \theta_n^f = \left\| f - \frac{1}{n} \cdot \sum_{i=0}^{n-1} T^i x \right\| \leq \frac{1}{n} \cdot \sum_{i=0}^{n-1} \left\| T^i x - T^i f \right\| \leq \|x - f\| \leq b.$$

Here b is a bound on the diameter of $C \subseteq B_{b/2}$, the domain of our nonexpansive map $T : C \to C$ (cf. standing assumption (6) from the introduction). The preceding discussion is somewhat informal, but it helps to motivate the following:

Definition 4.1. Consider a function $g : \mathbb{N} \to \mathbb{N}$. As usual, we write g^M for the monotone bound given by $g^M(n) := \max_{m \leq n} g(m)$. For $\varepsilon > 0$ we put

$$g_\varepsilon(n) := n' + g^M(n') \quad \text{with} \quad n' = \max\left\{ n, \left\lceil \frac{4nb}{\varepsilon} \right\rceil \right\}.$$

Now consider the iterates $g_\varepsilon^{(i)} : \mathbb{N} \to \mathbb{N}$ (with corrected start value) given by

$$g_\varepsilon^{(0)}(n) = \max\{1, n\} \quad \text{and} \quad g_\varepsilon^{(i+1)}(n) = g_\varepsilon\left(g_\varepsilon^{(i)}(n) \right).$$

Finally, define a map $(\varepsilon, g) \mapsto \Theta(\varepsilon, g)$ by setting

$$\Theta(\varepsilon, g) := g_\varepsilon^{(K+1)}\left(B\left(\frac{\varepsilon}{4}, g_\varepsilon^{(K+1)}, 2 \cdot g_\varepsilon^{(K+1)} \right) \right) \quad \text{with} \quad K = \left\lceil \frac{2b}{\varepsilon} \right\rceil.$$

Here B is the map from Definition 3.1.

As a systematic explanation for logicians, we point out that Θ can be seen as the combination of two rates of metastability (cf. Theorem 5.8 in the preprint version arXiv:1412.5563 of [25]): It is well known that a non-increasing sequence in $[0, b]$ admits a rate of metastability that depends on b only (see [17, Proposition 2.27]). We have a similar rate here, given that the θ_n^f are "almost" non-increasing by (14). This rate is combined with the rate B, which ensures that the last summand in (14) is indeed small (see Lemma 3.2). The fact that Θ combines two rates is made explicit in Corollary 4.4 below (note that Θ^B is a minor variant of Θ). Let us now present our metastable version of [13, Lemma 2]:

Proposition 4.2. *Let $f = Tf$ be a fixed point. For any $\varepsilon > 0$ and $g : \mathbb{N} \to \mathbb{N}$, there is an $N \leq \Theta(\varepsilon, g)$ such that $|\theta_m^f - \theta_n^f| < \varepsilon$ holds for all $m, n \in [N, N + g(N)]$.*

Proof. For $K = \lceil 2b/\varepsilon \rceil$, Lemma 3.2 yields an $N_0 \leq B(\varepsilon/4, g_\varepsilon^{(K+1)}, 2 \cdot g_\varepsilon^{(K+1)})$ with

$$(16) \qquad \beta_m^l \leq \frac{\varepsilon}{4} \quad \text{for all } m \in \left[N_0, N_0 + g_\varepsilon^{(K+1)}(N_0) \right] \text{ and } l \leq 2 \cdot g_\varepsilon^{(K+1)}(N_0).$$

We will show that some $N \in [N_0, g_\varepsilon^{(K+1)}(N_0)]$ validates the proposition. To see that we get $N \leq \Theta(\varepsilon, g)$, it suffices to observe that each iterate $g_\varepsilon^{(i)}$ is increasing. Inductively, this reduces to the same statement about g_ε, which holds by construction. Aiming at a contradiction, we now assume: for any number $N \in [N_0, g_\varepsilon^{(K+1)}(N_0)]$ there are $m, n \in [N, N + g(N)]$ with $|\theta_m^f - \theta_n^f| \geq \varepsilon$. We shall construct a sequence of numbers $n(i) \leq g_\varepsilon^{(i)}(N_0)$ with $\theta_{n(i)}^f \leq b - i\varepsilon/2$ by recursion on $i \leq K + 1$, which

contradicts $\theta^f_{n(K+1)} \geq 0$. In the base case, set $n(0) = g^{(0)}_\varepsilon(N_0)$ and note $\theta^f_{n(0)} \leq b$ due to (15). In the recursion step from $i \leq K$ to $i+1$, we consider

$$N := \max\left\{ n(i), \left\lceil \frac{4b \cdot n(i)}{\varepsilon} \right\rceil \right\}.$$

By (14) for $m = n(i)$ and $n = N - n(i) + k$ (with arbitrary k) we get

$$(17) \qquad \theta^f_{N+k} \leq \theta^f_{n(i)} + \frac{\varepsilon}{4} + \frac{1}{N+k} \cdot \sum_{j=0}^{N+k-1} \beta^{N-n(i)+k+j}_{n(i)}.$$

In order to apply (16), we verify that the indices of β lie in the appropriate interval of metastability. We anticipate that we will choose an $n(i+1)$ above $N \geq n(i)$, so that we may inductively assume $n(i) \geq n(0) \geq N_0$. As the function g_ε is increasing with $g_\varepsilon(j) \geq j$, we also have

$$n(i) \leq N \leq N + g^M(N) = g_\varepsilon(n(i)) \leq g_\varepsilon\left(g^{(i)}_\varepsilon(N_0) \right) \leq g^{(K+1)}_\varepsilon(N_0).$$

Furthermore, for $k \leq g(N)$ and $j < N + k$ we obtain

$$N - n(i) + k + j < 2 \cdot (N + g^M(N)) \leq 2 \cdot g^{(K+1)}_\varepsilon(N_0).$$

In view of these bounds, we can combine (16) and (17) to get

$$\theta^f_{N+k} \leq \theta^f_{n(i)} + \frac{\varepsilon}{2} \quad \text{for} \quad k \leq g(N).$$

On the other hand, we have $|\theta^f_m - \theta^f_n| \geq \varepsilon$ for some $m, n \in [N, N + g(N)]$, by the contradictory assumption for the present $N \in [N_0, g^{(K+1)}_\varepsilon(N_0)]$. Thus there must be a $k \leq g(N)$ with $\theta^f_{N+k} \leq \theta^f_{n(i)} - \varepsilon/2$. In order to complete the recursion step, we set $n(i+1) := N + k$ for some such k. Note that we indeed get

$$n(i+1) \leq N + g^M(N) \leq g^{i+1}_\varepsilon(N_0),$$

$$\theta^f_{n(i+1)} \leq \theta^f_{n(i)} - \frac{\varepsilon}{2} \leq b - \frac{i \cdot \varepsilon}{2} - \frac{\varepsilon}{2} = b - \frac{(i+1) \cdot \varepsilon}{2},$$

so that $n(i+1)$ retains the properties that were promised above. $\square$

We will want to bound $|\theta^f_m - \theta^f_n|$ and β^l_m simultaneously, i. e., on the same interval of metastability. In the present case, it suffices to tweak our previous construction:

Definition 4.3. Extending Definition 4.1, we set

$$\Theta^B(\varepsilon, g, h) := g^{(K+1)}_\varepsilon\left(B\left(\frac{\varepsilon}{4}, g^{(K+2)}_\varepsilon, \max\left\{ 2 \cdot g^{(K+1)}_\varepsilon, h^M \circ g^{(K+1)}_\varepsilon \right\} \right) \right),$$

still with $K = \lceil 2b/\varepsilon \rceil$ and $h^M(n) = \max_{m \leq n} h(m)$.

As promised, this yields a simultaneous bound:

Corollary 4.4. *For a fixed point $f = Tf$ and arbitrary $\varepsilon > 0$ and $g, h : \mathbb{N} \to \mathbb{N}$, there is an $N \leq \Theta^B(\varepsilon, g, h)$ such that we have both $\beta^l_{m,n} \leq \varepsilon/4$ and $|\theta^f_m - \theta^f_n| < \varepsilon$ for all $m, n \in [N, N + g(N)]$ and $l \leq h(N)$.*

Proof. From Lemma 3.2 we obtain a number

$$N_0 \leq B\left(\frac{\varepsilon}{4}, g^{(K+2)}_\varepsilon, \max\left\{ 2 \cdot g^{(K+1)}_\varepsilon, h^M \circ g^{(K+1)}_\varepsilon \right\} \right)$$

such that $\beta^l_{m,n} \leq \varepsilon/4$ holds whenever we have $m, n \in [N_0, N_0 + g^{(K+2)}_\varepsilon(N_0)]$ as well as $l \leq 2 \cdot g^{(K+1)}_\varepsilon(N_0)$ or $l \leq h^M \circ g^{(K+1)}_\varepsilon(N_0)$. The proof of Proposition 4.2 yields an $N \in [N_0, g^{(K+1)}_\varepsilon(N_0)]$ with $|\theta^f_m - \theta^f_n| < \varepsilon$ for all $m, n \in [N, N + g(N)]$. As before we see $N \leq \Theta^B(\varepsilon, g, h)$. In view of $h(N) \leq h^M \circ g^{(K+1)}_\varepsilon(N_0)$ and

$$N + g(N) \leq g_\varepsilon(N) \leq g^{(K+2)}_\varepsilon(N_0),$$

all the desired inequalities $\beta^k_{m,n} \leq \varepsilon/4$ are available as well. $\qquad\square$

In their proof of [13, Lemma 3], Kobayasi and Miyadera show that $(S_n T^n x)$ is a Cauchy sequence. We now provide a rate of metastability.

Definition 4.5. For $\varepsilon > 0$ and $g : \mathbb{N} \to \mathbb{N}$ we set

$$g'_\varepsilon(n) := \left\lceil \frac{6b \cdot (n + g(n))}{\delta(\varepsilon)} \right\rceil \quad \text{with} \quad \delta(\varepsilon) := \min\left\{ b, \frac{\varepsilon}{4}, \frac{\varepsilon}{8} \cdot \eta\left(\frac{\varepsilon}{2b}\right) \right\}.$$

We then define

$$\Delta(\varepsilon, g) := \Theta^B(\delta(\varepsilon), g + g'_\varepsilon, \mathrm{Id} + g + 2 \cdot g'_\varepsilon),$$

with $\mathrm{Id}(n) = n$ and for $\delta(\varepsilon)$ as above.

As promised, we have the following:

Proposition 4.6. *For any $\varepsilon > 0$ and $g : \mathbb{N} \to \mathbb{N}$ there is an $N \leq \Delta(\varepsilon, g)$ such that we have $\|S_m T^m x - S_n T^n x\| < \varepsilon$ for all $m, n \in [N, N + g(N)]$.*

Proof. Pick a fixed point $f = Tf \in C \subseteq B_{b/2}$, as justified by the standing assumptions from the introduction. In the conclusion of the proposition, we can replace both expressions $S_k T^k x$ by $y_k := S_k T^k x - f$. Note that we get $\|y_k\| = \theta^f_k$ in the notation from above. Corollary 4.4 yields an $N \leq \Delta(\varepsilon, g)$ with

$$(18) \qquad \left. \begin{array}{l} |\theta^f_m - \theta^f_n| < \delta(\varepsilon) \\[4pt] \text{and } \beta^l_{m,n} \leq \frac{\delta(\varepsilon)}{4} \end{array} \right\} \quad \text{when} \quad \left\{ \begin{array}{l} m, n \in [N, N + g(N) + g'_\varepsilon(N)] \\[4pt] \text{and } l \leq N + g(N) + 2 \cdot g'_\varepsilon(N). \end{array} \right.$$

We must show $\|y_m - y_n\| < \varepsilon$ for arbitrary $m, n \in [N, N + g(N)]$, say with $m \leq n$. For $\delta(\varepsilon)$ as in Definition 4.5, we put

$$k := \left\lceil \frac{6n \cdot b}{\delta(\varepsilon)} \right\rceil \leq g'_\varepsilon(N) \quad \text{and} \quad K := m + k \leq N + g(N) + g'_\varepsilon(N).$$

First assume that we have $\theta^f_K \leq \varepsilon/4$. Using (18), we then get

$$\|y_m - y_n\| \leq \|y_m\| + \|y_n\| = \theta^f_m + \theta^f_n < 2 \cdot (\theta^f_K + \delta(\varepsilon)) \leq \varepsilon.$$

From now on we assume $\theta^f_K > \varepsilon/4$, which entails

$$(19) \qquad (\theta^f_K + \delta(\varepsilon)) \cdot \left(1 - \eta\left(\frac{\varepsilon}{2b}\right)\right) < \theta^f_K - \delta(\varepsilon).$$

Normalize y_m and y_n by dividing through $\theta^f_K + \delta(\varepsilon)$, and then apply the contrapositive of (7) with $\varepsilon/(2b)$ at the place of ε. Due to $\theta^f_K, \delta(\varepsilon) \leq b$ this yields

$$\left\| \frac{y_m + y_n}{2} \right\| > (\theta^f_K + \delta(\varepsilon)) \cdot \left(1 - \eta\left(\frac{\varepsilon}{2b}\right)\right) \;\Rightarrow\; \|y_m - y_n\| < (\theta^f_K + \delta(\varepsilon)) \cdot \frac{\varepsilon}{2b} \leq \varepsilon.$$

In view of (19) it remains to show that we have $\|(y_m + y_n)/2\| \geq \theta_K^f - \delta(\varepsilon)$. As in the proof of [13, Lemma 3], we first observe that $\|y_{l+1} - y_l\|$ tends to zero: Since our standing assumption $C \subseteq B_{b/2}$ ensures $\|T^l x\| \leq b/2$, we have

$$\|y_{l+1} - y_l\| = \left\| \left(\frac{1}{l} - \frac{1}{l(l+1)} \right) \cdot \sum_{i=1}^{l+1} T^{l+i} x - \frac{1}{l} \cdot \sum_{i=0}^{l-1} T^{l+i} x \right\| \leq$$

$$\leq \frac{1}{l} \cdot \left\| T^{2l+1} x + T^{2l} x - T^l x \right\| + \frac{1}{l(l+1)} \cdot \sum_{i=1}^{l+1} \left\| T^{l+i} x \right\| \leq \frac{2b}{l}.$$

Summing up yields $\|y_{l+i} - y_l\| \leq 2b \cdot i/l$. Applied to $l = m + k = K$ and $i = n - m$, we obtain

$$\|y_{n+k} - y_K\| \leq 2b \cdot \frac{n-m}{m+k} \leq 2b \cdot \frac{n}{k} \leq \frac{\delta(\varepsilon)}{3} \leq \delta(\varepsilon).$$

We can now use the reverse triangle inequality to infer

(20) $$\|y_{m+k} + y_{n+k}\| \geq 2 \cdot \|y_K\| - \|y_{n+k} - y_K\| \geq 2 \cdot \theta_K^f - \delta(\varepsilon).$$

On the other hand, the proof of [13, Lemma 3] shows

$$\|y_{m+k} + y_{n+k}\| \leq \|y_m + y_n\| + \frac{3n}{m+k} \cdot b + \frac{2}{m+k} \cdot \sum_{i=0}^{m+k-1} \beta_{m,n}^{k+i}.$$

For $i \leq m + k - 1$ we have $k + i \leq m + 2k \leq N + g(N) + 2 \cdot g_\varepsilon'(N)$. By (18) and the choice of k we can thus infer $\|y_{m+k} + y_{n+k}\| \leq \|y_m + y_n\| + \delta(\varepsilon)$. Together with inequality (20) we can conclude

$$\|y_m + y_n\| \geq \|y_{m+k} + y_{n+k}\| - \delta(\varepsilon) \geq 2 \cdot (\theta_K^f - \delta(\varepsilon)),$$

as needed to establish the open claim. $\qquad \square$

We have just analysed the proof that $(S_n T^n x)$ is a Cauchy sequence. The limit y of this sequence is a fixed point of T, as shown by Kobayasi and Miyadera (still in the proof of [13, Lemma 3]). One might expect that a quantitative version of this result consists in (metastable) bounds on the norms $\|T S_n T^n x - S_n T^n x\|$. However, in the proof of [13, Theorem 1], Kobayasi and Miyadera use that we have $T^l y = y$ for arbitrary $l \in \mathbb{N}$ (rather than just for $l = 1$). In order to reflect this fact, we will bound $\|T^l S_n T^n x - S_n T^n x\|$ with a metastable dependency on l.

Definition 4.7. First extend Definition 4.5 by setting

$$\Delta^B(\varepsilon, g, h) := \Theta^B(\delta(\varepsilon), g + g_\varepsilon', \max\{\mathrm{Id} + g + 2 \cdot g_\varepsilon', h\}).$$

For $\varepsilon > 0$ and $g, h : \mathbb{N} \to \mathbb{N}$ we now put

$$\Psi(\varepsilon, g, h) := \Delta^B \left(\frac{\varepsilon}{4}, g_{\varepsilon,h}'', h \right) \quad \text{with} \quad g_{\varepsilon,h}''(N) := \max \left\{ g(N), \left\lceil \frac{4b \cdot h(N)}{\varepsilon} \right\rceil \right\}.$$

It is straightforward to see that Δ^B combines Proposition 4.6 with Corollary 4.4 (by the proof of the proposition and $\delta(\varepsilon) \leq \varepsilon/4$; cf. the proof of the corollary):

Corollary 4.8. *For any $\varepsilon > 0$ and $g, h : \mathbb{N} \to \mathbb{N}$ there is an $N \leq \Delta^B(\varepsilon, g, h)$ with $\beta_{m,n}^l \leq \varepsilon/16$ and $\|S_m T^m - S_n T^n x\| < \varepsilon$ for $m, n \in [N, N + g(N)]$ and $l \leq h(N)$.*

We are ready to show that Ψ validates statement (5) from the introduction:

Theorem 4.9. *For any $\varepsilon > 0$ and $g, h : \mathbb{N} \to \mathbb{N}$ there is an $N \leq \Psi(\varepsilon, g, h)$ such that we have $\|T^l S_n T^n x - S_n T^n x\| < \varepsilon$ for all $n \in [N, N + g(N)]$ and $l \leq h(N)$.*

Proof. For $g''_{\varepsilon,h}$ as in Definition 4.7, the corollary yields an $N \leq \Psi(\varepsilon,g,h)$ with

$$
\tag{21}
\left.
\begin{aligned}
\|S_m T^m x - S_n T^n x\| &< \tfrac{\varepsilon}{4} \\
\text{and } \beta_n^l &< \tfrac{\varepsilon}{4}
\end{aligned}
\right\}
\quad \text{when} \quad
\left\{
\begin{aligned}
&m, n \in [N, N + g''_{\varepsilon,h}(N)] \\
&\text{and } l \leq h(N).
\end{aligned}
\right.
$$

To establish the conclusion of the theorem, we consider arbitrary $n \in [N, N + g(N)]$ and $l \leq h(N)$. Set $m := N + g''_{\varepsilon,h}(N)$ and observe

$$
\left\|T^l S_m T^m x - T^l S_n T^n x\right\| \leq \|S_m T^m x - S_n T^n x\| < \frac{\varepsilon}{4}.
$$

Using the triangle inequality, we can conclude

$$
\left\|T^l S_n T^n x - S_n T^n x\right\| \leq \left\|T^l S_m T^m x - S_m T^m x\right\| + \frac{\varepsilon}{2}.
$$

In other words, we have reduced the claim for n to the claim for the given m (at the cost of a smaller ε). The point is that m is large, so that $S_m T^m x$ is a better approximation of the fixed point $y = Ty$ that is considered in Kobayasi and Miyadera's proof of [13, Lemma 3]. Also by the triangle inequality, we get

$$
\left\|T^l S_m T^m x - S_m T^m x\right\| \leq \left\|S_m T^{l+m} x - S_m T^m x\right\| + \beta_m^l.
$$

Due to cancellations between the Cesàro sums, we have

$$
\left\|S_m T^{l+m} x - S_m T^m x\right\| \leq \frac{1}{m} \cdot \left(\sum_{i=m}^{m+l-1} \left\|T^{m+i} x\right\| + \sum_{i=0}^{l-1} \left\|T^{m+i} x\right\| \right).
$$

By the standing assumption that $T : C \to C$ has domain $C \subseteq B_{b/2}$, this entails

$$
\left\|S_m T^{l+m} x - S_m T^m x\right\| \leq \frac{l \cdot b}{m} \leq \frac{h(N) \cdot b}{g''_{\varepsilon,h}(N)} \leq \frac{\varepsilon}{4}.
$$

Given that (21) provides $\beta_m^l < \varepsilon/4$, we can combine the previous inequalities to obtain $\left\|T^l S_n T^n x - S_n T^n x\right\| < \varepsilon$, as desired. $\square$

In order to obtain the simultaneous bounds from Corollaries 4.4 and 4.8, we have modified Θ and Δ into Θ^B and Δ^B, respectively. In the case of Ψ, we get a simultaneous bound without additional modifications, by the proof of Theorem 4.9:

Corollary 4.10. *For any $\varepsilon > 0$ and $g, h : \mathbb{N} \to \mathbb{N}$ there is an $N \leq \Psi(\varepsilon, g, h)$ such that $\left\|T^l S_n T^n x - S_n T^n x\right\| < \varepsilon$ and $\|S_m T^m x - S_n T^n x\| < \varepsilon/4$ and $\beta_n^l < \varepsilon/4$ hold for all $m, n \in [N, N + g(N)]$ and $l \leq h(N)$.*

Finally, we can specify the map Φ that validates (4) from the introduction:

Definition 4.11. First, let the map $(\varepsilon, n) \mapsto N_\varepsilon(n)$ be given by

$$
N_\varepsilon(n) := \max \left\{ n, \left\lceil \frac{6n \cdot b}{\varepsilon} \right\rceil \right\}.
$$

For $\varepsilon > 0$ and $g, h : \mathbb{N} \to \mathbb{N}$ we now set

$$
\Phi(\varepsilon, g, h) := N_\varepsilon \left(\Psi \left(\frac{\varepsilon}{2}, g', h' \right) \right)
\quad \text{with} \quad
\left\{
\begin{aligned}
g'(n) &:= N_\varepsilon(n) + g\left(N_\varepsilon(n)\right), \\
h'(n) &:= g(n) + h\left(N_\varepsilon(n)\right).
\end{aligned}
\right.
$$

As promised, we get the following:

Theorem 4.12. *For any $\varepsilon > 0$ and $g, h : \mathbb{N} \to \mathbb{N}$ there is an $N \leq \Phi(\varepsilon, g, h)$ such that $\left\|S_m T^m x - S_n T^k x\right\| < \varepsilon$ holds for all $m, n \in [N, N + g(N)]$ and $k \leq h(N)$.*

Proof. Use Corollary 4.10 to find an $M \leq \Psi\left(\frac{\varepsilon}{2}, g', h'\right)$ with

$$
(22) \quad
\left.
\begin{aligned}
\left\|T^l S_n T^n x - S_n T^n x\right\| &< \tfrac{\varepsilon}{2} \\
\text{and } \left\|S_m T^m x - S_n T^n x\right\| &< \tfrac{\varepsilon}{8} \\
\text{and } \beta_n^l &< \tfrac{\varepsilon}{8}
\end{aligned}
\right\}
\quad \text{when} \quad
\left\{
\begin{aligned}
&m, n \in [M, M + g'(M)] \\
&\text{and } l \leq h'(M),
\end{aligned}
\right.
$$

for g' and h' as in Definition 4.11. To establish the conclusion of the theorem for $N := N_\varepsilon(M) \leq \Phi(\varepsilon, g, h)$, consider arbitrary $m, n \in [N, N + g(N)]$ and $k \leq h(N)$. In view of $n \geq N \geq M$, the proof of [13, Theorem 1] yields

$$
(23) \quad \left\|S_n T^k x - S_m T^m x\right\| \leq \frac{3M}{2n} \cdot b + \frac{1}{n} \cdot \sum_{i=M}^{n-1} \left\|S_M T^{k+i} x - S_m T^m x\right\|.
$$

For each $i \in [M, n-1]$ we can write $k + i = M + l$ with $l \leq h'(M)$. Using the triangle inequality and (22), it follows that $\left\|S_M T^{k+i} x - S_m T^m x\right\|$ is smaller than

$$
\beta_M^l + \left\|T^l S_M T^M x - S_M T^M x\right\| + \left\|S_M T^M x - S_m T^m x\right\| \leq \frac{3}{4} \cdot \varepsilon.
$$

By (23) and the definition of $N = N_\varepsilon(M)$, we can conclude

$$
\left\|S_n T^k x - S_m T^m x\right\| < \frac{3M}{2 \cdot N_\varepsilon(M)} \cdot b + \frac{3}{4} \cdot \varepsilon \leq \varepsilon,
$$

just as the theorem claims. $\square$

Finally, we show that all relevant quantities can be bounded simultaneously. This yields the main result of our paper, which was stated in the introduction.

Proof of Theorem 1.2. The proof of Theorem 4.12 yields an $N \leq \Phi(2\varepsilon/5, g, h)$ such that we have

$$
\left\|S_m T^m x - S_n T^k x\right\| < 2\varepsilon/5 < \varepsilon \quad \text{and} \quad \left\|T^l S_n T^n x - S_n T^n x\right\| < \varepsilon/5 < \varepsilon
$$

for all $m, n \in [N, N + g(N)]$ and $k, l \leq h(N)$. Using the triangle inequality and the fact that T is nonexpansive, we also get

$$
\left\|T^l S_n T^k - S_n T^k\right\| \leq 2 \cdot \left\|S_n T^k x - S_N T^N x\right\| + \left\|T^l S_N T^N x - S_N T^N x\right\| < \varepsilon,
$$

still for $n \in [N, N + g(N)]$ and $k, l \leq h(N)$. $\square$

5. BOUNDS ON ASYMPTOTIC ISOMETRY

Theorem 1.1 (due to Kobayasi and Miyadera) involves the condition that T is asymptotically isometric on $\{x\}$, or explicitly: that the values $\alpha_n^i = \left\|T^n x - T^{n+i} x\right\|$ converge for $n \to \infty$ uniformly in $i \in \mathbb{N}$. On the quantitative side, this corresponds to one of our standing assumptions: throughout the previous sections, we have assumed that we are given a map $(\varepsilon, g, h) \mapsto A(\varepsilon, g, h)$ that validates statement (10). In the introduction, we have mentioned three cases in which a suitable A can be constructed. The first case is covered in a previous paper by the second author [19]: For an asymptotically regular map T satisfying Wittmann's condition on a uniformly convex Banach space, Theorem 2.2 of the cited paper (with $g'(n) := g(n) + h(n)$ and $\varepsilon' := \varepsilon/2$) yields a bound on an N such that we have

$$
|\alpha_m^i - \alpha_n^i| \leq |\alpha_m^i| + |\alpha_n^i| < \varepsilon \quad \text{for all } m, n \in [N, N + g(N)] \text{ and } i \leq h(N).
$$

Two further constructions in different cases are provided in the present section. Note that each construction yields a version of Theorems 4.9 and 4.12 in which Φ and Ψ no longer depend on A (but possibly on new data such as Γ in (26) below).

For the first part of this section, we consider a nonexpansive map $T : C \to C$ on a subset C of a Hilbert space. If T is odd on $C = -C \ni 0$, then we clearly have

$$(24) \qquad T0 = 0 \in C \quad \text{and} \quad \|Tx + Ty\| \leq \|x + y\| \text{ for all } x, y \in C.$$

In the following we do not assume that T is odd but do require that it satisfies (24). This condition has been studied by Wittmann [32] and plays an important role in several applications of proof mining (cf. [19, 21, 30]). Most standing assumptions from the introduction are not needed for the following, but we still assume $C \subseteq B_{b/2}$ (actually it suffices here to assume that $\|x\| \leq \frac{b}{2}$).

Our quantitative analysis is based on the following result, which is implicit in the proof of [3, Theorem 2] (cf. also [4, Theorem 2.3]). Let us point out that the sequence of norms $\|T^n x\|$ is non-increasing, since T is nonexpansive with $T0 = 0$. We have already seen that the sequence (α_n^i) is non-increasing for each $i \in \mathbb{N}$.

Lemma 5.1. *We have* $(\alpha_m^i)^2 - (\alpha_{m+k}^i)^2 \leq 4 \cdot \left(\|T^m x\|^2 - \left\| T^{m+k+i} x \right\|^2 \right).$

Proof. Using binomial expansion and the fact that $(\|T^n x\|)$ is non-increasing, we learn that $(\alpha_m^i)^2 - (\alpha_{m+k}^i)^2$ is bounded by

$$2 \cdot \left(\|T^m x\|^2 - \left\| T^{m+k+i} x \right\|^2 + \langle T^{m+k} x, T^{m+k+i} x \rangle - \langle T^m x, T^{m+i} x \rangle \right).$$

Hence the claim reduces to

$$\langle T^{m+k} x, T^{m+k+i} x \rangle - \langle T^m x, T^{m+i} x \rangle \leq \|T^m x\|^2 - \left\| T^{m+k+i} x \right\|^2.$$

To obtain the latter, iterate (24) to get $\left\| T^{m+k} x + T^{m+k+i} x \right\|^2 \leq \left\| T^m x + T^{m+i} x \right\|^2$. Now consider binomial expansions, and use that $(\|T^n x\|)$ is non-increasing. $\square$

In view of the standing assumption $C \subseteq B_{b/2}$, the values $\|T^n x\|^2$ form a non-increasing sequence of reals in $[0, b^2/4]$. Let us recall the known rate of metastability for such sequences. Given $g : \mathbb{N} \to \mathbb{N}$, we define $\widetilde{g} : \mathbb{N} \to \mathbb{N}$ by $\widetilde{g}(n) := n + g(n)$. The iterates $\widetilde{g}^i : \mathbb{N} \to \mathbb{N}$ are given by the recursive clauses

$$\widetilde{g}^0(n) := n \quad \text{and} \quad \widetilde{g}^{i+1}(n) := \widetilde{g}(\widetilde{g}^i(n)).$$

By [17, Proposition 2.27] (cf. also the beginning of Section 4 above) we have

$$(25) \quad \left| \|T^m x\|^2 - \|T^n x\|^2 \right| < \varepsilon \quad \text{for some } N \leq \widetilde{g}^{\lceil b^2/(4\varepsilon) \rceil}(0)$$

$$\text{and all } m, n \in [N, N + g(N)],$$

for any $\varepsilon > 0$ and $g : \mathbb{N} \to \mathbb{N}$. We will see that statement (10) from the introduction holds with the following map A_1 at the place of A.

Definition 5.2. For $\varepsilon > 0$ and $g, h : \mathbb{N} \to \mathbb{N}$ we set

$$A_1(\varepsilon, g, h) := \widetilde{g + h}^{\lceil b^2/\varepsilon^2 \rceil}(0).$$

As promised, we obtain the following:

Proposition 5.3. *Consider a nonexpansive map* $T : C \to C$ *on a set* $C \subseteq B_{b/2}$ *in Hilbert space. If* T *satisfies (24), then any* $\varepsilon > 0$ *and* $g, h : \mathbb{N} \to \mathbb{N}$ *admit a number* $N \leq A_1(\varepsilon, g, h)$ *with* $|\alpha_m^i - \alpha_n^i| < \varepsilon$ *for all* $m, n \in [N, N + g(N)]$ *and* $i \leq h(N)$.

Proof. By (25) with $\varepsilon^2/4$ and $g + h$ at the place of ε and g, respectively, we find a number $N \leq A_1(\varepsilon, g, h)$ with

$$\left| \|T^m x\|^2 - \|T^l x\|^2 \right| < \frac{\varepsilon^2}{4} \quad \text{for all } m, l \in [N, N + g(N) + h(N)].$$

Given $m \leq n$ in $[N, N + g(N)]$ and $i \leq h(N)$, combine this inequality for $l = n + i$ with Lemma 5.1 for $k = n - m$. This yields

$$(\alpha_m^i - \alpha_n^i)^2 \leq (\alpha_m^i)^2 - (\alpha_n^i)^2 \leq 4 \cdot \left(\|T^m x\|^2 - \|T^{n+i} x\|^2 \right) < \varepsilon^2$$

and hence $0 \leq \alpha_m^i - \alpha_n^i < \varepsilon$, as desired. $\qquad\qquad\square$

For the second part of this section, we return to the case of a Banach space. Our aim is to satisfy (10) when $(T^n x)$ has a convergent subsequence. The latter entails that there are $M, N \in \mathbb{N}$ as in the following lemma. We point out that the lemma is a quantitative version of a step in Bruck's proof of [4, Theorem 2.4]. Also note that the lemma establishes (10) whenever we have $N \leq A(\varepsilon, g, h) = A(\varepsilon, g)$, independently of the function $h : \mathbb{N} \to \mathbb{N}$ that provides the bound $i \leq h(N)$ in (10).

Lemma 5.4. *For $\varepsilon > 0$ and $g : \mathbb{N} \to \mathbb{N}$, assume that we are given $M, N \in \mathbb{N}$ with $M \geq N + g(N)$ and $\|T^M x - T^N x\| < \varepsilon/2$. We then have $|\alpha_m^i - \alpha_n^i| < \varepsilon$ for all numbers $m, n \in [N, N + g(N)]$ and $i \in \mathbb{N}$.*

Proof. From $\|T^M x - T^N x\| < \varepsilon/2$ we get $\|T^{M+i} x - T^{N+i} x\| < \varepsilon/2$, as T is non-expansive by a standing assumption. We recall $\alpha_N^i = \|T^N x - T^{N+i} x\|$ to conclude

$$\alpha_N^i \leq \|T^N x - T^M x\| + \|T^M x - T^{M+i} x\| + \|T^{M+i} x - T^{N+i} x\| < \varepsilon + \alpha_M^i.$$

The claim follows since α_n^i is decreasing in n (again because T is nonexpansive). $\quad\square$

The assumption that $(T^n x)$ has a convergent subsequence is satisfied when the domain C of our map $T : C \to C$ (or just the set $\{T^n x \,|\, n \in \mathbb{N}\}$) is compact or – equivalently – (closed and) totally bounded. On the quantitative side, this last property can be witnessed by a modulus of total boundedness (in the sense of P. Gerhardy [10]), i.e., by a function $\gamma : (0, \infty) \to \mathbb{N}$ that satisfies the following: for any $\varepsilon > 0$ and any sequence (x_i) in C, we have $\|x_j - x_i\| \leq \varepsilon$ for some $i < j \leq \gamma(\varepsilon)$ (see [25] for a metatheorem on uniform bound extractions for spaces given with such a modulus). We make an (apparently) weaker assumption, where γ may depend on the sequence and only the smaller index is controlled: for the following, we assume that we are given a map $(\varepsilon, g) \mapsto \Gamma(\varepsilon, g)$ that guarantees

$$(26) \qquad \left\| T^{g(j)} x - T^{g(i)} x \right\| < \varepsilon \qquad \text{for some } i \leq \Gamma(\varepsilon, g) \text{ and some } j > i.$$

Recall the notation $\widetilde{g}$ and $\widetilde{g}^i$ from the paragraph before Definition 5.2.

Definition 5.5. For $\varepsilon > 0$ and $g : \mathbb{N} \to \mathbb{N}$ we put

$$A_2(\varepsilon, g) := \widetilde{g}^K(0) \quad \text{with} \quad K := \Gamma\left(\frac{\varepsilon}{2}, l \mapsto \widetilde{g}^l(0) \right).$$

Here $(\varepsilon, g) \mapsto \Gamma(\varepsilon, g)$ is a given map that satisfies (26).

To conclude, we show that (10) holds with $(\varepsilon, g, h) \mapsto A_2(\varepsilon, g, h) := A_2(\varepsilon, g)$ at the place of A. Given that A_2 is independent of $h : \mathbb{N} \to \mathbb{N}$, the condition $i \leq h(N)$ can be dropped in the present situation.

Proposition 5.6. *Assume that $T : C \to C$ is a nonexpansive map on a totally bounded subset of a Banach space, where the total boundedness for the sequences $(T^{g(n)}x)$ is witnessed by a given map $(\varepsilon, g) \mapsto \Gamma(\varepsilon, g)$ as in (26). For any $\varepsilon > 0$ and $g : \mathbb{N} \to \mathbb{N}$ we can then find an $N \leq A_2(\varepsilon, g)$ such that $|\alpha_m^i - \alpha_n^i| < \varepsilon$ holds for all $m, n \in [N, N + g(N)]$ and $i \in \mathbb{N}$.*

Proof. Our modulus of total boundedness provides $i \leq \Gamma(\varepsilon/2, l \mapsto \widetilde{g}^l(0))$ and $j > i$ such that $M := \widetilde{g}^j(0)$ and $N := \widetilde{g}^i(0)$ satisfy $\left\| T^M x - T^N x \right\| < \varepsilon/2$. Note that the function $l \mapsto \widetilde{g}^l(0)$ is increasing because of $\widetilde{g}(n) = n + g(n) \geq n$. This allows us, first, to infer $N \leq A_2(\varepsilon, g)$ from the definition of A_2. Secondly, we learn

$$M = \widetilde{g}^j(0) \geq \widetilde{g}^{i+1}(0) = \widetilde{g}(\widetilde{g}^i(0)) = \widetilde{g}(N) = N + g(N),$$

so that we can conclude by Lemma 5.4. $\qquad\square$

Remark 5.7. In the finite dimensional case $X := \mathbb{R}^N$ endowed with the Euclidean norm, a modulus of total boundedness for $B_{b/2}$ can be taken as

$$\gamma(\varepsilon) := \left\lceil \frac{\sqrt{N}b}{\varepsilon} \right\rceil,$$

see [25, Example 2.8], and so we may take $\Gamma(\varepsilon, g) := \gamma(\varepsilon/2)$ independently of g, T, x.

Final Comment: An inspection of the proofs above shows that the quantitative results do not depend on the assumptions of X being complete or C being closed.

References

[1] Jeremy Avigad, Philipp Gerhardy, and Henry Towsner, *Local stability of ergodic averages*, Transactions of the American Mathematical Society **362** (2010), 261–288.

[2] Jeremy Avigad and Jason Rute, *Oscillation and the mean ergodic theorem for uniformly convex Banach spaces*, Ergodic Theory and Dynamical Systems **35** (2015), 1009–1027.

[3] Haïm Brézis and Felix Browder, *Nonlinear ergodic theorems*, Bulletin of the American Mathematical Society **82** (1976), no. 6, 959–961.

[4] Ronald E. Bruck, *On the almost-convergence of iterates of a nonexpansive mapping in Hilbert space and the structure of the weak ω-limit set*, Israel Journal of Mathematics **29** (1978), 1–16.

[5] ———, *A simple proof of the mean ergodic theorem for nonlinear contractions in Banach spaces*, Israel Journal of Mathematics **32** (1979), 107–116.

[6] ———, *On the convex approximation property and the asymptotic behavior of nonlinear contractions in Banach spaces*, Israel Journal of Mathematics **38** (1981), 304–314.

[7] James A. Clarkson, *Uniformly convex spaces*, Transactions of the American Mathematical Society **40** (1936), no. 3, 396–414.

[8] Stephen J. Dilworth, Ralph Howard, and James W. Roberts, *On the size of approximately convex sets in normed spaces*, Studia Mathematica **140** (2000), 213–241.

[9] Anton Freund, *Proof lengths for instances of the Paris-Harrington principle*, Annals of Pure and Applied Logic **168** (2017), 1361–1382.

[10] Philipp Gerhardy, *Proof mining in topological dynamics*, Notre Dame Journal of Formal Logic **49** (2008), no. 4, 431–446.

[11] Uffe Haagerup, *The best constants in the Khintchine inequality*, Studia Mathematica **70** (1982), 231–283.

[12] Olof Hanner, *On the uniform convexity of L^p and l^p*, Arkiv för Matematik **3** (1956), no. 3, 239–244.

[13] Kazuo Kobayasi and Isao Miyadera, *On the strong convergence of the Césaro means of contractions in Banach spaces*, Proc. Japan Acad. **56** (1980), 245–249.

[14] Ulrich Kohlenbach, *On the computational content of the Krasnoselski and Ishikawa fixed point theorems*, Computability and Complexity in Analysis (Jens Blanck, Vasco Brattka, and Peter Hertling, eds.), Lecture Notes in Computer Science, vol. 2064, Springer, 2001, pp. 119–145.

[15] ———, *Uniform asymptotic regularity for Mann iterates*, Journal of Mathematical Analysis and Applications **279** (2003), no. 2, 531–544.

[16] ———, *Some logical metatheorems with applications in functional analysis*, Transactions of the American Mathematical Society **357** (2005), no. 1, 89–128.

[17] ———, *Applied Proof Theory: Proof Interpretations and their Use in Mathematics*, Springer Monographs in Mathematics, Springer, Berlin and Heidelberg, 2008.

[18] ———, *On quantitative versions of theorems due to F.E. Browder and R. Wittmann*, Advances in Mathematics **226** (2011), no. 3, 2764–2795.

[19] ———, *On the asymptotic behavior of odd operators*, Journal of Mathematical Analysis and Applications **382** (2011), 615–620.

[20] ———, *A uniform quantitative form of sequential weak compactness and Baillon's nonlinear ergodic theorem*, Communications in Contemporary Mathematics **14** (2012), no. 1, 1250006, 20 pp.

[21] ———, *On the quantitative asymptotic behavior of strongly nonexpansive mappings in Banach and geodesic spaces*, Israel Journal of Mathematics **216** (2016), 215–246.

[22] ———, *Quantitative results on the Proximal Point Algorithm in uniformly convex Banach spaces*, Journal of Convex Analysis **28** (2021), no. 1, 11–18.

[23] Ulrich Kohlenbach and Laurenţiu Leuştean, *A quantitative mean ergodic theorem for uniformly convex Banach spaces*, Ergodic Theory and Dynamical Systems **29** (2009), 1907–1915.

[24] ———, *Asymptotically nonexpansive mappings in uniformly convex hyperbolic spaces*, Journal of the European Mathematical Society **12** (2010), 71–92.

[25] Ulrich Kohlenbach, Laurenţiu Leuştean, and Adriana Nicolae, *Quantitative results on Fejér monotone sequences*, Communications in Contemporary Mathematics **20** (2018), no. 2, article no. 1750015, 42 pp.

[26] Ulrich Kohlenbach and Andrei Sipoş, *The finitary content of sunny nonexpansive retractions*, Communications in Contemporary Mathematics **23** (2021), no. 1, 1950093, 63pp.

[27] Michel Ledoux and Michel Talagrand, *Probability in Banach spaces*, Ergebnisse der Mathematik und ihrer Grenzgebiete, Springer, Berlin and Heidelberg, 1991.

[28] Gilles Pisier, *Sur les espaces qui ne contiennent pas de l_n^1 uniformément*, Séminaire Maurey-Schwartz (1973-1974), Espaces L^p, applications radonifiantes et géométrie des espaces de Banach, Centre de Mathématiques, École Polytechnique, Paris, 1974, Exp. No. 7, 19 pp.

[29] Thomas Powell, *A new metastable convergence criterion and an application in the theory of uniformly convex Banach spaces*, Journal of Mathematical Analysis and Applications **478** (2019), no. 2, 790–805.

[30] Pavol Safarik, *A quantitative nonlinear strong ergodic theorem for Hilbert spaces*, Journal of Mathematical Analysis and Applications **391** (2012), no. 1, 26–37.

[31] Terence Tao, *Structure and Randomness: pages from year one of a mathematical blog*, American Mathematical Society, 2008.

[32] Rainer Wittmann, *Mean ergodic theorems for nonlinear operators*, Proceedings of the American Mathematical Society **108** (1990), no. 3, 781–788.

DEPARTMENT OF MATHEMATICS, TECHNICAL UNIVERSITY OF DARMSTADT, SCHLOSSGARTEN-STR. 7, 64289 DARMSTADT, GERMANY

E-mail address: {freund,kohlenbach}@mathematik.tu-darmstadt.de