Effective metastability of Halpern iterates in CAT(0) spaces

U. Kohlenbach[1], L. Leuştean[2]

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

[2] Simion Stoilow Institute of Mathematics of the Romanian Academy,
Research unit 5, P. O. Box 1-764, RO-014700 Bucharest, Romania

E-mails: kohlenbach@mathematik.tu-darmstadt.de, Laurentiu.Leustean@imar.ro.

Abstract

This paper provides an effective uniform rate of metastability (in the sense of Tao) on the strong convergence of Halpern iterations of nonexpansive mappings in CAT(0) spaces. The extraction of this rate from an ineffective proof due to Saejung is an instance of the general proof mining program which uses tools from mathematical logic to uncover hidden computational content from proofs. This methodology is applied here for the first time to a proof that uses Banach limits and hence makes a substantial reference to the axiom of choice.

MSC: 47H09, 47H10, 03F10, 53C23.

Keywords: Proof mining, Banach limits, metastability, nonexpansive mappings, CAT(0) spaces, Halpern iterations.

1 Introduction

This paper applies techniques from mathematical logic to extract an explicit uniform rate of metastability (in the sense of Tao [49, 50]) from a recent proof due to Saejung [42] of a strong convergence theorem for Halpern iterations in the context of CAT(0) spaces. The theorem in question has been established originally in the context of Hilbert spaces by Wittmann in the important paper [52] and can there be viewed as a strong nonlinear generalization of the classical von Neumann mean ergodic theorem. Indeed, Wittmann's theorem says that under suitable conditions on a sequence of scalars (λ_n) in $[0,1]$, including the case $\lambda_n := \dfrac{1}{n+1}$, the so-called Halpern iteration

$$x_0 := x, \quad x_{n+1} := \lambda_{n+1}x + (1-\lambda_{n+1})Tx_n$$

of a nonexpansive selfmapping $T : C \to C$ of a bounded closed and convex subset $C \subseteq X$ strongly converges to a fixed point of T. If T is, moreover, linear and $\lambda_n := \dfrac{1}{n+1}$, then x_n coincides with the ergodic average $\dfrac{1}{n+1}\sum_{i=0}^{n} T^i x$ from the mean ergodic theorem.

Since Wittmann's theorem does not refer to any linearity but only to a convexity structure of the underlying space X (in order to make sense of the Halpern iteration) it can be formulated in the context of hyperbolic spaces and was established by Saejung [42] for the important subclass of CAT(0) spaces which play the analogous role in the context of hyperbolic spaces as the Hilbert spaces do among all Banach spaces.

As shown in [3], even for the (linear) mean ergodic theorem, there, in general, is no computable rate of convergence for (x_n). The next best thing to achieve, therefore, is a rate of metastability, i.e. a bound $\Phi(k,g)$ such that

$$(1)\ \forall k \in \mathbb{N}\, \forall g : \mathbb{N} \to \mathbb{N}\, \exists n \leq \Phi(k,g)\, \forall i,j \in [n, n+g(n)]\ \big(\|x_i - x_j\| \leq 2^{-k}\big).$$

There are general logical metatheorems due to the first author [21] and Gerhardy and the first author [14] that guarantee the extractability of computable and highly uniform such bounds $\Phi(k,g)$ from large classes of (even highly ineffective) proofs. Moreover, these bounds have a restricted complexity depending on the principles that are used in the proof rather than merely being computable (see [22] for a comprehensive treatment).

A rate of metastability is an instance of the concept of no-counterexample interpretation that was introduced in the context of mathematical logic by Kreisel in the 50's [27, 28]: as g may be viewed as an attempt to refute the Cauchy property of (x_n), the functional $\Phi(k,g)$ in (1) provides a bound on a counterexample n to such a refutation. Note that since g may be an arbitrary number theoretic function, the seemingly weaker form

$$(2)\ \forall k \in \mathbb{N}\, \forall g : \mathbb{N} \to \mathbb{N}\, \exists n \in \mathbb{N}\, \forall i,j \in [n, n+g(n)]\ \big(\|x_i - x_j\| \leq 2^{-k}\big)$$

of the Cauchy property actually implies back the full Cauchy property, though only ineffectively so. Because of the latter point, the existence of an effective bound on (2) does not contradict the aforementioned fact that there is no effective Cauchy rate for (x_n) available.

By the uniformity of the bound Φ we refer to the fact that it is independent of the operator T, the point $x \in C$ as well as of C and X but only depends – in addition to k and g – on a bound on the diameter of C as well as – in the case of general (λ_n) – certain moduli on (λ_n).

Based on the aforementioned logical metatheorems, [3] extracted the first explicit such uniform bound Φ for the mean ergodic theorem from its usual textbook proof. Subsequently, in [24] the current authors extracted such bound for the more general class of uniformly convex Banach spaces from a proof due to G. Birkhoff. That bound – when specialized to the Hilbert space setting – even turned out to be numerically better than the one from [3].

In [23], the first author extracted – making use of a rate of asymptotic regularity due to the second author [29] – a rate of metastability of similar complexity for Wittmann's nonlinear ergodic theorem (in the Hilbert case). Wittmann's proof is based on weak compactness which, though covered by the existing proof mining machinery, in general can cause bounds of extremely poor quality. In the case at hand that could be avoided as during the logical extraction procedure the use of weak compactness turned out to be eliminable.

In the present paper, we extract a rate of metastability from Saejung's [42] generalization of Wittmann's theorem to the CAT(0)-setting. In addition to the interest of this specific result, our paper is of broader relevance in the proof mining program as it opens up new frontiers for its applicability namely to proofs that prima facie use some substantial amount of the axiom of choice. This stems from the use of Banach limits made in [42]. The existence of Banach limits is either proved by applying the Hahn-Banach theorem to l^∞ which due to the nonseparability of that space needs the axiom of choice, or via ultralimits which, again, needs choice. While weak compactness as used in Wittmann's proof at least was in principle covered by existing metatheorems mentioned above, this is not the case for Banach limits. Though it seems likely that these metatheorems can be extended to incorporate at least basic reasoning with Banach limits as we intend to discuss in a different paper, we take the route in this paper to show how to replace the use of Banach limits in the present proof by a direct arithmetical reasoning. As the way Banach limits are used in the proof at hand seems to be rather typical for other proofs in fixed point theory, our paper may also be seen as providing a blueprint for doing similar unwindings in those cases as well. Usually, a Banach limit is used to establish the almost convergence in the sense of Lorentz of some sequence (a_n) of reals towards a which – together with $\limsup\limits_{n\to\infty}(a_{n+1} - a_n) \leq 0$ – in turn implies that $\limsup\limits_{n\to\infty} a_n \leq a$. This line of reasoning goes back to Lorentz' classical paper [34] whose relevance in nonlinear ergodic theory was first realized by Reich [36]. In [44], Banach limits are used in this way to establish Wittmann's theorem for uniformly Gâteaux differentiable Banach spaces (under suitable conditions on C). This paper has subsequently been analyzed using the method developed in this paper in [25]. Other relavent papers using Banach limits in the context of nonlinear ergodic theory are [8, 41, 26].

As an intermediate step in proving our main results we also obtain in Section 6 (essentially due to the second author in [30]) a uniform effective rate of asymptotic regularity. i.e. a rate of convergence of $(d(x_n, Tx_n))$ towards 0, which holds in general W-hyperbolic spaces. As this bound, in particular, does not depend on x and T, it provides a quantitative version of the main result in [1] (see their 'Theorem 3.3').

2 Preliminaries

We shall consider hyperbolic spaces as introduced by the first author [21]. In order to distinguish them from Gromov hyperbolic spaces or from other notions of hyperbolic space that can be found in the metric fixed point theory literature (see for example [19, 15, 40]), we shall call them W-hyperbolic spaces.

A *W-hyperbolic space* (X, d, W) is a metric space (X, d) together with a mapping $W : X \times X \times [0, 1] \to X$ satisfying

$$\begin{array}{ll}(W1) & d(z, W(x, y, \lambda)) \leq (1 - \lambda)d(z, x) + \lambda d(z, y),\\ (W2) & d(W(x, y, \lambda), W(x, y, \tilde{\lambda})) = |\lambda - \tilde{\lambda}| \cdot d(x, y),\\ (W3) & W(x, y, \lambda) = W(y, x, 1 - \lambda),\\ (W4) & d(W(x, z, \lambda), W(y, w, \lambda)) \leq (1 - \lambda)d(x, y) + \lambda d(z, w).\end{array}$$

The convexity mapping W was first considered by Takahashi in [48], where a triple (X, d, W) satisfying $(W1)$ is called a *convex metric space*. We refer to [22, p. 384-387] for a detailed discussion.

The class of W-hyperbolic spaces includes normed spaces and convex subsets thereof, the Hilbert ball (see [16] for a book treatment) as well as CAT(0) spaces [5].

If $x, y \in X$ and $\lambda \in [0, 1]$, then we use the notation $(1 - \lambda)x \oplus \lambda y$ for $W(x, y, \lambda)$. It is easy to see that for all $x, y \in X$ and $\lambda \in [0, 1]$,

$$d(x, (1 - \lambda)x \oplus \lambda y) = \lambda d(x, y) \quad \text{and} \quad d(y, (1 - \lambda)x \oplus \lambda y) = (1 - \lambda)d(x, y). \tag{1}$$

Furthermore, $1x \oplus 0y = x$, $0x \oplus 1y = y$ and $(1 - \lambda)x \oplus \lambda x = \lambda x \oplus (1 - \lambda)x = x$.

For all $x, y \in X$, we shall denote by $[x, y]$ the set $\{(1 - \lambda)x \oplus \lambda y : \lambda \in [0, 1]\}$. A subset $C \subseteq X$ is said to be *convex* if $[x, y] \subseteq C$ for all $x, y \in C$. A nice feature of our setting is that any convex subset is itself a W-hyperbolic space with the restriction of d and W to C.

Let us recall now some notions on geodesic spaces. Let (X, d) be a metric space. A geodesic path, geodesic for short, in X is a map $\gamma : [a, b] \to X$ which is distance-preserving, that is

$$d(\gamma(s), \gamma(t)) = |s - t| \;\text{ for all }\; s, t \in [a, b]. \tag{2}$$

A *geodesic segment* in X is the image of a geodesic $\gamma : [a, b] \to X$, the points $x := \gamma(a)$ and $y := \gamma(b)$ being the endpoints of the segment. We say that the geodesic segment $\gamma([a, b])$ joins x and y. The metric space (X, d) is said to be a *(uniquely) geodesic space* if every two distinct points are joined by a (unique) geodesic segment. It is easy to see that any W-hyperbolic space is geodesic.

A *CAT(0) space* is a geodesic space (X, d) satisfying the so-called **CN**-inequality of Bruhat-Tits [9]: for all $x, y, z \in X$ and $m \in X$ with $d(x, m) = d(y, m) = \frac{1}{2}d(x, y)$,

$$d(z, m)^2 \leq \frac{1}{2}d(z, x)^2 + \frac{1}{2}d(z, y)^2 - \frac{1}{4}d(x, y)^2. \tag{3}$$

The fact that this definition of a CAT(0) space is equivalent to the usual definition using geodesic triangles is an exercise in [5, p. 163]. Complete CAT(0) spaces are often called Hadamard spaces. One can show that CAT(0) spaces are uniquely geodesic and that a normed space is a CAT(0) space if and only if it is a pre-Hilbert space.

CAT(0) spaces can be defined also in terms of W-hyperbolic spaces.

Lemma 2.1. *[22, p. 386-388] Let (X, d) be a metric space. The following are equivalent.*

(i) X is a CAT(0) *space.*

(ii) There exists a a convexity mapping W such that (X, d, W) is a W-hyperbolic space satisfying the **CN** *inequality (3).*

The following property of CAT(0) spaces will be very useful in the following. We refer to [11, Lemma 2.5] for a proof.

Proposition 2.2. *Let (X, d) be a* CAT(0) *space. Then for all $x, y, z \in X$ and $\lambda \in [0, 1]$.*

$$d^2((1-\lambda)x \oplus \lambda y, z) \leq (1-\lambda)d^2(x, z) + \lambda d^2(y, z) - \lambda(1-\lambda)d^2(x, y). \tag{4}$$

We recall now some terminology needed for our quantitative results. Let $(a_n)_{n\geq 1}$ be a sequence of real numbers and $a \in \mathbb{R}$. In the following $\mathbb{N} = \{0, 1, 2, \ldots\}$ and $\mathbb{Z}_+ = \{1, 2, \ldots\}$.

If the series $\sum\limits_{n=1}^{\infty} a_n$ is divergent, then a function $\gamma : \mathbb{Z}_+ \to \mathbb{Z}_+$ is called a *rate of divergence* of the series if $\sum\limits_{k=1}^{\gamma(n)} a_k \geq n$ for all $n \in \mathbb{Z}_+$.

If $\lim\limits_{n\to\infty} a_n = a$, then a function $\gamma : (0, \infty) \to \mathbb{Z}_+$ is said to be a *rate of convergence* of (a_n) if

$$\forall \varepsilon > 0\, \forall n \geq \gamma(\varepsilon)\ \ (|a_n - a| \leq \varepsilon)\,. \tag{5}$$

Assume that (a_n) is Cauchy. Then

(i) a mapping $\gamma : (0, \infty) \to \mathbb{Z}_+$ is called a *Cauchy modulus* of (a_n) if

$$\forall \varepsilon > 0\, \forall n \in \mathbb{N}\ \ \left(a_{\gamma(\varepsilon)+n} - a_{\gamma(\varepsilon)} \leq \varepsilon\right). \tag{6}$$

(ii) a mapping $\Psi : (0, \infty) \times \mathbb{N}^{\mathbb{N}} \to \mathbb{Z}_+$ is called a *rate of metastability* of (a_n) if

$$\forall \varepsilon > 0\, \forall g : \mathbb{N} \to \mathbb{N}\, \exists N \leq \Psi(\varepsilon, g)\ \forall m, n \in [N, N + g(N)]\ (|a_n - a_m| \leq \varepsilon). \tag{7}$$

Finally, we say that $\limsup\limits_{n\to\infty} a_n \leq 0$ with *effective rate* $\theta : (0, \infty) \to \mathbb{Z}_+$ if

$$\forall \varepsilon > 0\, \forall n \geq \theta(\varepsilon)\ (a_n \leq \varepsilon). \tag{8}$$

3 Halpern iterations

Let C be a convex subset of a normed space X and $T : C \to C$ nonexpansive. The so-called *Halpern iteration* is defined as follows:

$$x_0 := x, \quad x_{n+1} := \lambda_{n+1} u + (1 - \lambda_{n+1}) T x_n, \tag{9}$$

where $(\lambda_n)_{n\geq 1}$ is a sequence in $[0, 1]$, $x \in C$ is the starting point and $u \in C$ is the anchor.

If T is positively homogeneous (i.e. $T(tx) = tT(x)$ for all $t \geq 0$ and all $x \in C$), $\lambda_n = \dfrac{1}{n+1}$ and $u = x$, then

$$x_n = \frac{1}{n+1} S_n x, \quad \text{where} \quad S_0 x = x,\ \ S_{n+1} x = x + T(S_n x). \tag{10}$$

Furthermore, if T is linear, then $x_n = \dfrac{1}{n+1}\sum_{i=0}^{n} T^i x$, so the Halpern iteration could be regarded as a nonlinear generalization of the usual Cesàro average. We refer to [51, 31] for a a systematic study of the behavior of iterations given by (10).

The following problem was formulated by Reich [38] (see also [35]) and it is still open in its full generality.

Problem 3.1. *[38, Problem 6]*
Let X be a Banach space. Is there a sequence (λ_n) such that whenever a weakly compact convex subset C of X possesses the fixed point property for nonexpansive mappings, then (x_n) converges to a fixed point of T for all $x \in C$ and all nonexpansive mappings $T : C \to C$?

Different conditions on (λ_n) were considered in the literature (see also [47] for even more conditions):

$$\begin{array}{ll} (C1) & \lim_{n\to\infty} \lambda_n = 0, \\ (C2) & \sum_{n=1}^{\infty} |\lambda_{n+1} - \lambda_n| \text{ converges}, \\ (C3) & \sum_{n=1}^{\infty} \lambda_n = \infty, \\ (C4) & \prod_{n=1}^{\infty} (1-\lambda_n) = 0, \end{array}$$

and, in the case $\lambda_n > 0$ for all $n \geq 1$,

$$\begin{array}{ll} (C5) & \lim_{n\to\infty} \dfrac{\lambda_n - \lambda_{n+1}}{\lambda_{n+1}^2} = 0, \\ (C6) & \lim_{n\to\infty} \dfrac{\lambda_n - \lambda_{n+1}}{\lambda_{n+1}} = 0. \end{array}$$

For sequences λ_n in $(0, 1)$, conditions (C3) and (C4) are equivalent.

Halpern [18] initiated the study in the Hilbert space setting of the convergence of a particular case of the scheme (9). He proved that the sequence (x_n), obtained by taking $u = 0$ in (9), converges to a fixed point of T for (λ_n) satisfying certain conditions, two of which are (C1) and (C3). P.-L. Lions [32] improved Halpern's result by showing the convergence of the general (x_n) if (λ_n) satisfies (C1), (C3) and (C5). However, both Halpern's and Lions' conditions exclude the natural choice $\lambda_n = \dfrac{1}{n+1}$.

This was overcome by Wittmann [52], who obtained the most important result on the convergence of Halpern iterations in Hilbert spaces.

Theorem 3.2. *[52] Let C be a closed convex subset of a Hilbert space X and $T : C \to C$ a nonexpansive mapping such that the set $Fix(T)$ of fixed points of T is nonempty. Assume that (λ_n) satisfies (C1), (C2) and (C3). Then for any $x \in C$, the Halpern iteration (x_n) converges to the projection Px of x on $Fix(T)$.*

All the partial answers to Reich's problem require that the sequence (λ_n) satisfies (C1) and (C3). Halpern [18] showed in fact that conditions (C1) and (C3) are necessary in the sense that if, for every closed convex subset C of a Hilbert space X and every nonexpansive mappings $T : C \to C$ such that $Fix(T) \neq \emptyset$, the Halpern iteration (x_n) converges to a fixed point of T, then (λ_n) must satisfy (C1) and (C3). That (C1) and (C3) alone are not sufficient to guarantee the convergence of (x_n) was shown in [47]. Recently, Chidume and Chidume [10] and Suzuki [46] proved that if the nonexpansive mapping T in (9) is averaged, then (C1) and (C3) suffice for obtaining the convergence of (x_n).

Halpern obtained his result by applying a limit theorem for the resolvent, first shown by Browder [6]. This approach has the advantage that the result can be immediately generalized, once the limit theorem for the resolvent is generalized. This was done by Reich [37].

Theorem 3.3. *[37] Let C be a closed convex subset of a uniformly smooth Banach space X, and let $T : C \to C$ be nonexpansive such that $Fix(T) \neq \emptyset$. For each $u \in C$ and $t \in (0,1)$, let z_t^u denote the unique fixed point of the contraction mapping*

$$T_t(\cdot) = tu + (1-t)T(\cdot).$$

Then $\lim_{t \to 0^+} z_t^u$ exists and is a fixed point of T.

A similar result was obtained by Kirk [20] for CAT(0) spaces (for the Hilbert ball, which is an example of a CAT(0) space, this is already due to [16]). As a consequence of Theorem 3.3, a partial positive answer to Problem 3.1 was obtained [37] for uniformly smooth Banach spaces and $\lambda_n = \dfrac{1}{(n+1)^\alpha}$ with $0 < \alpha < 1$. Furthermore, Reich [39] proved the strong convergence of (x_n) in the setting of uniformly smooth Banach spaces that have a weakly sequentially continuous duality mapping for general (λ_n) satisfying (C1), (C3) and being decreasing. Another partial answer in the case of uniformly smooth Banach spaces was obtained by Xu [53] for (λ_n) satisfying (C1), (C3) and (C6) (which is weaker than Lions' (C5)). In [44], Shioji and Takahashi extended Wittmann's result to Banach spaces with uniformly Gâteaux differentiable norm and with the property that $\lim_{t \to 0^+} z_t^u$ exists and is a fixed point of T.

4 Main results

Let $T : C \to C$ be a nonexpansive selfmapping of a convex subset C of a W-hyperbolic space (X, d, W). We can define the Halpern iteration in this setting too:

$$x_0 := x, \quad x_{n+1} := \lambda_{n+1} u \oplus (1 - \lambda_{n+1}) T x_n, \tag{11}$$

where $x, u \in C$ and $(\lambda_n)_{n \geq 1}$ is a sequence in $[0,1]$.

The following theorem generalizes Wittman's theorem to CAT(0) spaces and was obtained by Saejung [42] (as similar result for the Hilbert ball had already been proved in [26]).

Theorem 4.1. *Let C be a closed convex subset of a complete* CAT(0) *space X and $T : C \to C$ a nonexpansive mapping such that the set $Fix(T)$ of fixed points of T is nonempty. Assume that (λ_n) satisfies (C1), (C2) and (C3). Then for any $u, x \in C$, the iteration (x_n) converges to the projection Pu of u on $Fix(T)$.*

By [20, Theorem 18], $Fix(T) \neq \emptyset$ is guaranteed to hold if C is bounded. In this paper we only consider this case and our bounds will depend on an upper bound M on the diameter d_C of C. However, similar to [23], it is not hard to adopt our bounds to the case where the condition $M \geq d_C$ is being replaced by $M \geq d(u,p), d(x,p)$ for some fixed point $p \in C$ of T.

The main results of the paper are effective versions of Theorem 4.1, obtained by applying proof mining techniques to Saejung's proof. As this proof is essentially ineffective and – as we discussed in the introduction – a computable rate of convergence does not exist, while an effective and highly uniform rate of metastability (depending only on the input data displayed in Theorems 4.2, 4.3) is guaranteed to exist (via our elimination of Banach limits from the proof) by [21, Theorem 3.7.3] (note that the conditions on α, β, θ as well as T are all purely universal while the conclusion $\exists N \forall m, n \in [N, N + g(N)]\ (d(x_n, x_m) < \varepsilon)$ can be written as a purely existential formula and that quantification over all (λ_n) in $[0,1]$ can be represented as $\forall y \leq s$ for some simple function $s : \mathbb{N}^2 \to \mathbb{N}$).

Theorem 4.2. *Assume that X is a complete* CAT(0) *space, $C \subseteq X$ is a closed bounded convex subset with diameter d_C and $T : C \to C$ is nonexpansive. Let (λ_n) satisfy (C1), (C2) and (C3).*

Then the Halpern iteration (x_n) is Cauchy.

Furthermore, let α be a rate of convergence of (λ_n), β be a Cauchy modulus of $s_n := \sum_{i=1}^{n} |\lambda_{i+1} - \lambda_i|$ and θ be a rate of divergence of $\sum_{n=1}^{\infty} \lambda_{n+1}$.

Then for all $\varepsilon \in (0,2)$ *and* $g : \mathbb{N} \to \mathbb{N}$,

$$\exists N \leq \Sigma(\varepsilon, g, M, \theta, \alpha, \beta)\ \forall m, n \in [N, N+g(N)]\ (d(x_n, x_m) \leq \varepsilon),$$

where

$$\Sigma(\varepsilon, g, M, \theta, \alpha, \beta) = \theta^+\left(\Gamma - 1 + \left\lceil \ln\left(\frac{12M^2}{\varepsilon^2}\right)\right\rceil\right) + 1 \tag{12}$$

with $M \in \mathbb{Z}_+$ *such that* $M \geq d_C$,

$$\varepsilon_0 = \frac{\varepsilon^2}{24(M+1)^2}, \quad \Gamma = \max\left\{\chi_k^*(\varepsilon^2/12) \mid \left\lceil \frac{1}{\varepsilon_0}\right\rceil \leq k \leq \widetilde{f^*}^{(\lceil M^2/\varepsilon_0^2\rceil)}(0) + \left\lceil \frac{1}{\varepsilon_0}\right\rceil\right\},$$

$$\chi_k^*(\varepsilon) = \tilde{\Phi}\left(\frac{\varepsilon}{4M(\tilde{P}_k(\varepsilon/2)+1)}\right) + \tilde{P}_k(\varepsilon/2), \quad \tilde{P}_k(\varepsilon) = \left\lceil \frac{12M^2(k+1)}{\varepsilon}\Phi\left(\frac{\varepsilon}{12M(k+1)}\right)\right\rceil,$$

$$\tilde{\Phi}(\varepsilon, M, \theta, \beta) = \theta\left(\beta\left(\frac{\varepsilon}{4M}\right) + 1 + \left\lceil \ln\left(\frac{2M}{\varepsilon}\right)\right\rceil\right) + 1,$$
$$\Phi(\varepsilon, M, \theta, \alpha, \beta) = \max\left\{\tilde{\Phi}\left(\frac{\varepsilon}{2}, M, \theta, \beta\right), \alpha\left(\frac{\varepsilon}{4M}\right)\right\},$$

$$\Delta_k^*(\varepsilon, g) = \frac{\varepsilon}{3g_{\varepsilon,k}\left(\Theta_k(\varepsilon) - \chi_k^*(\varepsilon/3)\right)}, \quad \Theta_k(\varepsilon) = \theta\left(\chi_k^*\left(\frac{\varepsilon}{3}\right) - 1 + \left\lceil \ln\left(\frac{3M^2}{\varepsilon}\right)\right\rceil\right) + 1,$$
$$g_{\varepsilon,k}(n) = n + g\left(n + \chi_k^*\left(\frac{\varepsilon}{3}\right)\right), \qquad \theta^+(n) = \max\{\theta(i) \mid i \leq n\},$$

$$f(k) = \max\left\{\left\lceil \frac{M^2}{\Delta_k^*(\varepsilon^2/4, g)}\right\rceil, k\right\} - k, \quad f^*(k) = f\left(k + \left\lceil \frac{1}{\varepsilon_0}\right\rceil\right) + \left\lceil \frac{1}{\varepsilon_0}\right\rceil, \quad \widetilde{f^*}(k) = k + f^*(k).$$

Proof. See Section 10. □

A similar result can be obtained by assuming that (λ_n) satisfies (C1), (C2) and (C4) with corresponding rates.

Theorem 4.3. *Assume that* X *is a complete* CAT(0) *space,* $C \subseteq X$ *is a closed bounded convex subset with diameter* d_C *and* $T : C \to C$ *is nonexpansive. Let* (λ_n) *satisfy (C1), (C2), (C4) and* $\lambda_n \in (0,1)$ *for all* $n \geq 2$.

Then the Halpern iteration (x_n) *is Cauchy.*

Furthermore, if α *is a rate of convergence of* (λ_n)*,* β *is a Cauchy modulus of* $s_n := \sum_{i=1}^{n} |\lambda_{i+1} - \lambda_i|$ *and* θ *is a rate of convergence of* $\prod_{n=1}^{\infty}(1-\lambda_{n+1})$ *towards* 0*, then for all* $\varepsilon \in (0,2)$ *and* $g : \mathbb{N} \to \mathbb{N}$,

$$\exists N \leq \Sigma(\varepsilon, g, M, \theta, \alpha, \beta, (\lambda_n))\ \forall m, n \in [N, N+g(N)]\ (d(x_n, x_m) \leq \varepsilon),$$

where

$$\Sigma(\varepsilon, g, M, \theta, \alpha, \beta, (\lambda_n)) := max\left\{\Theta_k(\varepsilon^2/4) \mid \left\lceil \frac{1}{\varepsilon_0}\right\rceil \leq k \leq \widetilde{f^*}^{(\lceil M^2/\varepsilon_0^2\rceil)}(0) + \left\lceil \frac{1}{\varepsilon_0}\right\rceil\right\}, \tag{13}$$

with $M \in \mathbb{Z}_+$ *such that* $M \geq d_C$,

$$0 < D \leq \prod_{n=1}^{\beta(\varepsilon/4M)}(1-\lambda_{n+1}),$$

$$
\begin{aligned}
\tilde{\Phi}(\varepsilon, M, \theta, \beta, D) &= \theta\left(\frac{D\varepsilon}{2M}\right) + 1, \\
\Phi(\varepsilon, M, \theta, \alpha, \beta, D) &= \max\left\{\theta\left(\frac{D\varepsilon}{4M}\right) + 1, \alpha\left(\frac{\varepsilon}{4M}\right)\right\}, \\
\Theta_k(\varepsilon) &= \theta\left(\frac{D_k\varepsilon}{3M^2}\right) + 1, \\
0 < D_k &\leq \prod_{n=1}^{\chi_k^*(\varepsilon/3)-1} (1 - \lambda_{n+1}),
\end{aligned}
$$

and the other constants and functionals being defined as in Theorem 4.2.

Proof. We use Proposition 6.2, Lemma 5.3 and follow the same line as in the proof of Theorem 4.2. □

One can modify Theorems 4.2, 4.3 so that only metastable versions of α, β and θ are needed. However, we refrain from doing so as the result would be rather unreadable and in the practical cases at hand – such as $\lambda_n = \dfrac{1}{n+1}$ – full rates α, β, θ are easy to compute.

Corollary 4.4. *Assume that $\lambda_n = \dfrac{1}{n+1}$ for all $n \geq 1$. Then for all $\varepsilon \in (0,1)$ and $g : \mathbb{N} \to \mathbb{N}$,*

$$
\exists N \leq \Sigma(\varepsilon, g, M)\ \forall m, n \in [N, N + g(N)]\ (d(x_n, x_m) \leq \varepsilon),
$$

where

$$
\Sigma(\varepsilon, g, M) = \left\lceil \frac{12M^2(\chi_L^*(\varepsilon^2/12) + 1)}{\varepsilon^2} \right\rceil - 1 \tag{14}
$$

with

$$
\begin{aligned}
L &= \widetilde{f^*}^{(\lceil M^2/\varepsilon_0^2 \rceil)}(0) + \left\lceil \frac{1}{\varepsilon_0} \right\rceil, \\
\tilde{P}_k(\varepsilon) &= \left\lceil \frac{12M^2(k+1)}{\varepsilon} \cdot \left(\left\lceil \frac{48M(k+1)}{\varepsilon} + \frac{2304M^4(k+1)^2}{\varepsilon^2} \right\rceil - 1 \right) \right\rceil, \\
\chi_k^*(\varepsilon) &= \left\lceil \frac{8M^2(\tilde{P}_k(\varepsilon/2) + 1)}{\varepsilon} + \frac{128M^4(\tilde{P}_k(\varepsilon/2) + 1)^2}{\varepsilon^2} \right\rceil - 1 + \tilde{P}_k(\varepsilon/2), \\
\Theta_k(\varepsilon) &= \left\lceil \frac{3M^2(\chi_k^*(\varepsilon/3) + 1)}{\varepsilon} \right\rceil - 1,
\end{aligned}
$$

while the other constants and functionals are defined as in Theorem 4.2.

Proof. Since $\displaystyle\prod_{k=1}^{n}\left(1 - \frac{1}{k+2}\right) = \frac{2}{n+2}$, we get that $\theta(\varepsilon) := \left\lceil \dfrac{2}{\varepsilon} \right\rceil - 2$ is a rate of convergence of $\displaystyle\prod_{n=1}^{\infty}\left(1 - \frac{1}{n+2}\right)$ towards 0. Furthermore, we can take $D_k := \dfrac{2}{\chi_k^*(\varepsilon/3) + 1}$ in Theorem 4.3 and – using Corollary 6.3 – $\Phi := \Psi, \tilde{\Phi} := \tilde{\Psi}$ from that corollary. We then get $P_k(\varepsilon)$, $\chi_k^*(\varepsilon)$ as above and

$$
\Theta_k(\varepsilon) \;=\; \theta\left(\frac{D_k\varepsilon}{3M^2}\right) + 1 = \left\lceil \frac{3M^2(\chi_k^*(\varepsilon/3) + 1)}{\varepsilon} \right\rceil - 1.
$$

The claim now follows by (the proof of) Theorem 4.3 using that χ_k^* increases with k. □

Despite its superficially quite different look, the bound in Corollary 4.4 has an overall similar structure as the bound extracted for the Hilbert space case in [23]: the bound results from applying a certain function $\Theta_k(\varepsilon)$ to a number $k := L$ which is the result of an iteration of a function $\tilde{f^*}$ (starting at some arbitrary value, e.g. 0), where $\tilde{f^*}(k)$ is – disregarding many details – something close to $\Theta_k(\varepsilon) + g(\Theta_k(\varepsilon))$. This is also the structure of the bound in [23, Theorem 3.3] (where Δ^* plays the role of $\tilde{f^*}$). Note that the number of iterations essentially is M^6/ε^4 while it was roughly M^4/ε^4 in the bound in [23, Theorem 3.3]. The main difference, though, is that now Θ_k is significantly more involved compared to [23] (most of its terms stemming from the remains of the original Banach-limit argument).

Remark 4.5. *(i) By replacing (X,d) by (X,d_M) with $d_M(x,y) := \frac{1}{M}d(x,y)$ one can always arrange that $1 \geq d_C$ and then apply the above bounds for 1 instead of M but with ε/M instead of ε to compensate for this rescaling. One then gets a bound in which ε and M only occur in the form ε/M and the number of iterations is (essentially) M^4/ε^4. However, in doing so M would enter the bound at many unnecessary places as well.*

(ii) The assumption on the completeness of X and the closedness of C facilitates the proofs but is not necessary in the above results. If the results would fail for an incomplete X then it is easy to show that they would fail already for the metric completion $\widehat{X}$ of X and the closure $\overline{C}$ of C in $\widehat{X}$ (since T extends to a nonexpansive operator $\widehat{T} : \overline{C} \to \overline{C}$). Alternatively, one could use directly appropriate approximate fixed points rather than fixed points in the applications of Banach's fixed point theorem in section 9 below.

(iii) Subsequently, our results have been further generalized in [43] to the case of unbounded C provided that T possesses a fixed point p. Then the above bounds hold with $M \geq diam(C)$ being replaced by $M \geq 4\max\{d(u,x), d(u,p)\}$. In [43] our method is also adapted to obtain similar bounds for more general schemes of so-called modified Halpern iterations.

5 Quantitative lemmas on sequences of real numbers

The following lemma about sequences of real numbers was proved in [2].

Lemma 5.1. *Let (s_n) be a sequence of nonnegative real numbers, (α_n) be a sequence of real numbers in $[0,1]$ with $\sum_{n=1}^{\infty}\alpha_n = \infty$, and (t_n) be a sequence of real numbers with $\limsup_{n\to\infty} t_n \leq 0$. Suppose that*

$$s_{n+1} \leq (1-\alpha_n)s_n + \alpha_n t_n \quad \text{for all } n \geq 1.$$

Then $\lim_{n\to\infty} s_n = 0$.

We prove now quantitative versions of Lemma 5.1, which also allow for an error term Δ.

Lemma 5.2. *Let $\varepsilon \in (0,2)$, $g : \mathbb{N} \to \mathbb{N}$, $M \in \mathbb{Z}_+$, $\theta : \mathbb{Z}_+ \to \mathbb{Z}_+$ and $\psi : (0,\infty) \to \mathbb{Z}_+$. Define*

$$\Theta := \Theta(\varepsilon, M, \theta, \psi) = \theta\left(\psi\left(\frac{\varepsilon}{3}\right) - 1 + \left\lceil \ln\left(\frac{3M}{\varepsilon}\right)\right\rceil\right) + 1, \tag{15}$$

$$\Delta := \Delta(\varepsilon, g, M, \theta, \psi) = \frac{\varepsilon}{3g_\varepsilon(\Theta - \psi(\varepsilon/3))}, \tag{16}$$

where $g_\varepsilon(n) = n + g(n + \psi(\varepsilon/3))$.

Assume that (α_n) is a sequence in $[0,1]$ such that $\sum_{n=1}^{\infty}\alpha_n = \infty$ with rate of divergence θ. Let (t_n) be a sequence of real numbers satisfying

$$\forall n \geq \psi(\varepsilon/3)\ (t_n \leq \varepsilon/3). \tag{17}$$

Let (s_n) be a bounded sequence with upper bound M satisfying

$$s_{n+1} \leq (1-\alpha_n)s_n + \alpha_n t_n + \Delta \qquad \text{for all } n \geq 1. \tag{18}$$

Then

$$\forall n \in [\Theta, \Theta + g(\Theta)]\ (s_n \leq \varepsilon).$$

Proof. By induction on m one shows that for all $n \geq \psi(\varepsilon/3)$ and $m \geq 1$,

$$s_{n+m} \leq \left[\prod_{j=n}^{n+m-1}(1-\alpha_j)\right] s_n + \left[1 - \prod_{j=n}^{n+m-1}(1-\alpha_j)\right]\frac{\varepsilon}{3} + m\Delta. \tag{19}$$

$m = 1$: By (18) and (17), we have that

$$\begin{aligned} s_{n+1} &\leq (1-\alpha_n)s_n + \alpha_n t_n + \Delta \leq (1-\alpha_n)s_n + \alpha_n\frac{\varepsilon}{3} + \Delta \\ &= (1-\alpha_n)s_n + (1-(1-\alpha_n))\frac{\varepsilon}{3} + \Delta. \end{aligned}$$

$m \Rightarrow m+1$: We have that

$$\begin{aligned} s_{n+m+1} &\leq (1-\alpha_{n+m})s_{n+m} + \alpha_{n+m}t_{n+m} + \Delta \\ &\leq (1-\alpha_{n+m})\left[\prod_{j=n}^{n+m-1}(1-\alpha_j)\right] s_n + (1-\alpha_{n+m})\left[1 - \prod_{j=n}^{n+m-1}(1-\alpha_j)\right]\frac{\varepsilon}{3} + \\ &\quad +(1-\alpha_{n+m})m\Delta + \alpha_{n+m}t_{n+m} + \Delta \qquad \text{by the induction hypothesis} \\ &\leq \left[\prod_{j=n}^{n+m}(1-\alpha_j)\right] s_n + (1-\alpha_{n+m})\left[1 - \prod_{j=n}^{n+m-1}(1-\alpha_j)\right]\frac{\varepsilon}{3} + \alpha_{n+m}\frac{\varepsilon}{3} + (m+1)\Delta \\ &= \left[\prod_{j=n}^{n+m}(1-\alpha_j)\right] s_n + \left[1 - \alpha_{n+m} - \prod_{j=n}^{n+m}(1-\alpha_j) + \alpha_{n+m}\right]\frac{\varepsilon}{3} + (m+1)\Delta \\ &= \left[\prod_{j=n}^{n+m}(1-\alpha_j)\right] s_n + \left[1 - \prod_{j=n}^{n+m}(1-\alpha_j)\right]\frac{\varepsilon}{3} + (m+1)\Delta. \end{aligned}$$

Using the fact that $1 - x \leq \exp(-x)$ for all $x \in [0, \infty)$, we get that

$$\prod_{j=n}^{n+m-1}(1-\alpha_j) \leq \prod_{j=n}^{n+m-1}\exp(-\alpha_j) = \exp\left(-\sum_{j=n}^{n+m-1}\alpha_j\right),$$

hence

$$s_{n+m} \leq \exp\left(-\sum_{j=n}^{n+m-1}\alpha_j\right) s_n + \frac{\varepsilon}{3} + m\Delta \leq \exp\left(-\sum_{j=n}^{n+m-1}\alpha_j\right) M + \frac{\varepsilon}{3} + m\Delta \tag{20}$$

for all $n \geq \psi(\varepsilon/3)$ and $m \geq 1$.

For simplicity, let us denote $d_{m,n} := M\exp\left(-\sum_{j=n}^{n+m-1}\alpha_j\right)$. As in [29], we get that

$$\begin{aligned} d_{m,n} \leq \frac{\varepsilon}{3} \ &\Leftrightarrow\ \exp\left(-\sum_{j=n}^{n+m-1}\alpha_j\right) \leq \frac{\varepsilon}{3M} \ \Leftrightarrow\ -\sum_{j=n}^{n+m-1}\alpha_j \leq \ln\left(\frac{\varepsilon}{3M}\right) \\ &\Leftrightarrow\ \sum_{j=n}^{n+m-1}\alpha_j \geq -\ln\left(\frac{\varepsilon}{3M}\right) = \ln\left(\frac{3M}{\varepsilon}\right) \ \Leftrightarrow\ \sum_{j=1}^{n+m-1}\alpha_j \geq \sum_{j=1}^{n-1}\alpha_j + \ln\left(\frac{3M}{\varepsilon}\right). \end{aligned}$$

Let

$$L := \Theta - \psi(\varepsilon/3) = \theta\left(\psi(\varepsilon/3) - 1 + \left\lceil \ln\left(\frac{3M}{\varepsilon}\right)\right\rceil\right) + 1 - \psi(\varepsilon/3). \tag{21}$$

Since θ is a rate of divergence of $\sum_{n=1}^{\infty} \alpha_n$ and $\alpha_n \leq 1$, it is obvious that $\theta(n) \geq n$ for all $n \geq 1$, hence $L \geq 1$. For all $m \geq L$, we have that

$$\sum_{j=1}^{\psi(\varepsilon/3)+m-1} \alpha_j \geq \sum_{j=1}^{\psi(\varepsilon/3)+L-1} \alpha_j \geq \psi(\varepsilon/3) - 1 + \left\lceil \ln\left(\frac{3M}{\varepsilon}\right)\right\rceil \geq \sum_{j=1}^{\psi(\varepsilon/3)-1} \alpha_j + \ln\left(\frac{3M}{\varepsilon}\right),$$

hence

$$d_{m,\psi(\varepsilon/3)} \leq \frac{\varepsilon}{3} \quad \text{for all } m \geq L.$$

Apply now (20) with $n := \psi(\varepsilon/3)$ to get that for all $m \geq L$,

$$s_{\psi(\varepsilon/3)+m} \leq \frac{2\varepsilon}{3} + m\Delta. \tag{22}$$

Let $n \in [\Theta, \Theta + g(\Theta)]$. Then

$$L \leq n - \psi(\varepsilon/3) \leq \Theta + g(\Theta) - \psi(\varepsilon/3) = L + g(L + \psi(\varepsilon/3)) = g_\varepsilon(L),$$

hence we can apply (22) with $m := n - \psi(\varepsilon/3)$ to get that

$$s_n \leq \frac{2\varepsilon}{3} + g_\varepsilon(L)\Delta = \varepsilon.$$

□

It is well-known that for a sequence (α_n) in $(0,1)$ we have that $\sum_{n=1}^{\infty} \alpha_n = \infty$ if and only if $\prod_{n=1}^{\infty}(1-\alpha_n) = 0$. This suggests a second quantitative version of Lemma 5.1, where, instead of a rate of divergence for $\sum_{n=1}^{\infty} \alpha_n$, we assume the existence of a rate of convergence of $\prod_{n=1}^{\infty}(1-\alpha_n)$ towards 0.

Lemma 5.3. *Let $\varepsilon > 0$, $g : \mathbb{N} \to \mathbb{N}$, $M \in \mathbb{Z}_+$, $D > 0$ and $\theta, \psi : (0, \infty) \to \mathbb{Z}_+$. Define*

$$\Theta := \Theta(\varepsilon, M, \theta, \psi, D) = \max\left\{\theta\left(\frac{D\varepsilon}{3M}\right) + 1, \psi\left(\frac{\varepsilon}{3}\right)\right\} \tag{23}$$

$$\Delta := \Delta(\varepsilon, g, M, \theta, \psi, D) = \frac{\varepsilon}{3g_\varepsilon(\Theta - \psi(\varepsilon/3))}, \tag{24}$$

where $g_\varepsilon(n) = n + g(n + \psi(\varepsilon/3))$.

Assume that (α_n) is a sequence in $(0,1)$ such that $\prod_{n=1}^{\infty}(1-\alpha_n) = 0$ with rate of convergence θ. Let (t_n) be a sequence of real numbers satisfying

$$\forall n \geq \psi(\varepsilon/3)\ (t_n \leq \varepsilon/3). \tag{25}$$

Assume furthermore that

$$D \leq \prod_{n=1}^{\psi(\varepsilon/3)-1}(1-\alpha_n). \tag{26}$$

Let (s_n) be a bounded sequence with upper bound M satisfying

$$s_{n+1} \le (1-\alpha_n)s_n + \alpha_n t_n + \Delta \qquad \text{for all } n \ge 1. \tag{27}$$

Then

$$\forall n \in [\Theta, \Theta + g(\Theta)] \ (s_n \le \varepsilon).$$

Proof. We shall denote $P_n := \prod_{j=1}^{n}(1-\alpha_j)$ for all $n \ge 1$. By convention, $P_0 = 1$. We get as in the proof of Lemma 5.2 that for all $n \ge \psi(\varepsilon/3)$ and $m \ge 1$,

$$s_{n+m} \le \left[\prod_{j=n}^{n+m-1}(1-\alpha_j)\right] s_n + \left[1 - \prod_{j=n}^{n+m-1}(1-\alpha_j)\right]\frac{\varepsilon}{3} + m\Delta. \tag{28}$$

Hence, for all $n \ge \psi(\varepsilon/3)$ and $m \ge 1$,

$$\begin{aligned} s_{n+m} &\le \left[\prod_{j=n}^{n+m-1}(1-\alpha_j)\right] s_n + \frac{\varepsilon}{3} + m\Delta \le \left[\prod_{j=n}^{n+m-1}(1-\alpha_j)\right] M + \frac{\varepsilon}{3} + m\Delta \\ &= \frac{MP_{n+m-1}}{P_{n-1}} + \frac{\varepsilon}{3} + m\Delta. \end{aligned}$$

By taking $n := \psi(\varepsilon/3)$, we get that for all $m \ge 1$,

$$s_{\psi(\varepsilon/3)+m} \le \frac{MP_{\psi(\varepsilon/3)+m-1}}{P_{\psi(\varepsilon/3)-1}} + \frac{\varepsilon}{3} + m\Delta. \tag{29}$$

Define now

$$L := \Theta - \psi(\varepsilon/3) = \max\left\{\theta\left(\frac{D\varepsilon}{3M}\right) + 1 - \psi(\varepsilon/3), 0\right\} \tag{30}$$

and take $n \in [\Theta, \Theta + g(\Theta)]$ arbitrary. Then $L \le n - \psi(\varepsilon/3) \le g_\varepsilon(L)$ and, applying (29) with $m := n - \psi(\varepsilon/3)$, it follows that

$$s_n \le \frac{MP_{n-1}}{P_{\psi(\varepsilon/3)-1}} + \frac{\varepsilon}{3} + (n - \psi(\varepsilon/3))\Delta \le \frac{MP_{\Theta-1}}{P_{\psi(\varepsilon/3)-1}} + \frac{\varepsilon}{3} + g_\varepsilon(L)\Delta \le \frac{M}{P_{\psi(\varepsilon/3)-1}} \cdot \frac{D\varepsilon}{3M} + \frac{2\varepsilon}{3},$$

as $\Theta - 1 \ge \theta\left(\frac{D\varepsilon}{3M}\right)$. By (26), we get that $s_n \le \varepsilon$. □

The above lemma turns out to be very useful to get better bounds in the case $\alpha_n = \frac{1}{n+1}$, as $\sum_{n=1}^{\infty}\frac{1}{n+1}$ has an exponential rate of divergence, while $\prod_{n=1}^{\infty}\left(1 - \frac{1}{n+1}\right)$ has a linear rate of convergence towards 0.

Corollary 5.4. *Let $\varepsilon \in (0,3)$, $g : \mathbb{N} \to \mathbb{N}$, $M \in \mathbb{Z}_+$, $\psi : (0,\infty) \to \mathbb{Z}_+$. Define*

$$\Theta := \Theta(\varepsilon, M, \psi) = \left\lceil \frac{3M\psi(\varepsilon/3)}{\varepsilon} \right\rceil + 1, \quad \Delta := \Delta(\varepsilon, g, M, \psi) = \frac{\varepsilon}{3g_\varepsilon(\Theta - \psi(\varepsilon/3))}, \tag{31}$$

where $g_\varepsilon(n) = n + g(n + \psi(\varepsilon/3))$.
Assume that (t_n) is a sequence of real numbers satisfying

$$\forall n \ge \psi(\varepsilon/3) \ (t_n \le \varepsilon/3). \tag{32}$$

Let (s_n) be a bounded sequence with upper bound M satisfying

$$s_{n+1} \le \left(1 - \frac{1}{n+1}\right) s_n + \frac{1}{n+1} t_n + \Delta \qquad \textit{for all } n \ge 1. \tag{33}$$

Then

$$\forall n \in [\Theta, \Theta + g(\Theta)] \ (s_n \le \varepsilon). \tag{34}$$

Proof. Remark that for all $n \ge 1$, we have that $\prod_{k=1}^{n} \left(1 - \frac{1}{k+1}\right) = \frac{1}{n+1}$, hence $\theta(\varepsilon) := \left\lceil \frac{1}{\varepsilon} \right\rceil$ is a rate of convergence of $\prod_{n=1}^{\infty} \left(1 - \frac{1}{n+1}\right)$ towards 0. Furthermore, we can take $D := \frac{1}{\psi(\varepsilon/3)}$ in Lemma 5.3. Since $\varepsilon \in (0,3)$, we have that $\frac{3M\psi(\varepsilon/3)}{\varepsilon} \ge \psi(\varepsilon/3)$, hence $\left\lceil \frac{3M\psi(\varepsilon/3)}{\varepsilon} \right\rceil + 1 > \psi(\varepsilon/3)$. $\square$

The proof of Lemmas 5.2, 5.3 can actually be reformulated to give a full rate of convergence for (s_n) provided that one does not have the error term Δ or that Δ can be made arbitrarily small while still keeping ψ and (17) unchanged (note that Θ – in contrast to Δ – does not depend on g). This error term stems from the fact that we have to eliminate a use of an ineffective arithmetical comprehension hidden in forming the limit z of a certain sequence of points (z_{t_k}) which is used in Saejung's proof to construct the sequence which plays the role of (t_n) in the use of Lemma 5.2 or Lemma 5.3 (see [42, (2.21)-(2.23)]). Instead of z, we take z_{t_k} where k is sufficiently large so that $d(z_{t_j}, z) < \varepsilon$ for all $j \ge k$. This error can be incorporated (also when switching from z_{t_k} to z_{t_j} for $j \ge k$) into the error already present in (17) with some ψ_k depending on k but it adds the error $\Delta_j := M^2 t_j$ (see (73) below compared to [42, (2.21)]), which we provided for in (18). The error Δ_j, however, can be made arbitrarily small by increasing j without changing ψ_k in (17) (see the proof of the main Theorem 4.2). This would give us a rate of convergence in our Theorem 4.2 provided that we had a Cauchy rate on (z_{t_k}). However, we effectively only get a rate of metastability for this sequence (see Proposition 9.3 and the discussion preceding this proposition). As a result, k and in turn ψ_k become dependent on the counterfunction g. This has the consequence that now, via ψ_k, also Θ in our application of Lemma 5.2 (in the proof of Theorem 4.2) becomes dependent on g. It is this issue which is responsible for the fact that we only get an effective rate of metastability in Theorem 4.2 (rather than a Cauchy rate), which – as discussed in the introduction – in fact is best possible.

The following quantitative lemma is the main ingredient in getting effective rates of asymptotic regularity for the Halpern iteration.

Lemma 5.5. *Let $(\lambda_n)_{n\ge1}$ be a sequence in $[0,1]$ and $(a_n)_{n\ge1}, (b_n)_{n\ge1}$ be sequences in $\mathbb{R}_+$ such that for all $n \ge 1$,*

$$a_{n+1} \le (1 - \lambda_{n+1}) a_n + b_n. \tag{35}$$

Assume that $\sum_{n=1}^{\infty} b_n$ is convergent and γ is a Cauchy modulus of $s_n := \sum_{i=1}^{n} b_i$.

(i) If $\sum_{n=1}^{\infty} \lambda_{n+1}$ is divergent with rate of divergence θ, then

$$\forall \varepsilon \in (0,2) \, \forall n \ge \Phi \ \ (a_n \le \varepsilon),$$

where

$$\Phi := \Phi(\varepsilon, M, \theta, \gamma) = \theta\left(\gamma\left(\frac{\varepsilon}{2}\right) + 1 + \left\lceil \ln\left(\frac{2M}{\varepsilon}\right)\right\rceil\right) + 1. \tag{36}$$

and $M \in \mathbb{Z}_+$ is an upper bound on (a_n).

(ii) If $\lambda_n \in (0,1)$ for all $n \geq 2$ and $\prod_{n=1}^{\infty}(1-\lambda_{n+1}) = 0$ with rate of convergence θ, then

$$\forall \varepsilon \in (0,2)\, \forall n \geq \Phi \;\; (a_n \leq \varepsilon)\,,$$

where

$$\Phi := \Phi(\varepsilon, M, \theta, \gamma, D) = \theta\left(\frac{D\varepsilon}{2M}\right) + 1, \tag{37}$$

$M \in \mathbb{Z}_+$ is an upper bound on (a_n), and

$$0 < D \leq \prod_{n=1}^{\gamma(\varepsilon/2)} (1-\lambda_{n+1}). \tag{38}$$

Proof. (i) Follow the proof of [29, Lemma 9].

(ii) The proof of (ii) is basically contained in the proof of [29, Lemma 9]. For sake of completeness we give it here. We denote $P_n := \prod_{k=1}^{n}(1-\lambda_{k+1})$ for all $n \geq 1$. Let $\varepsilon \in (0,2)$ and define

$$N := \gamma\left(\frac{\varepsilon}{2}\right) + 1. \tag{39}$$

Applying [29, Lemma 8] with $n := N$, it follows that for all $m \geq 1$,

$$\begin{aligned} a_{N+m} &\leq \left[\prod_{j=N}^{N+m-1}(1-\lambda_{j+1})\right] a_N + \sum_{j=N}^{N+m-1} b_j = \frac{P_{N+m-1}}{P_{N-1}} \cdot a_N + \left(s_{\gamma\left(\frac{\varepsilon}{2}\right)+m} - s_{\gamma\left(\frac{\varepsilon}{2}\right)}\right) \\ &\leq \frac{MP_{N+m-1}}{P_{N-1}} + \frac{\varepsilon}{2}. \end{aligned}$$

Let

$$L := \Phi - N = \theta\left(\frac{D\varepsilon}{2M}\right) + 1 - N. \tag{40}$$

Then for all $m \geq L$, we have that $N + m - 1 \geq \theta\left(\frac{D\varepsilon}{2M}\right)$, hence

$$\frac{MP_{N+m-1}}{P_{N-1}} \leq \frac{D\varepsilon}{2P_{N-1}} \leq \frac{\varepsilon}{2}.$$

This also implies that $L \geq 1$ since, otherwise,

$$1 \leq M \leq \frac{MP_{N+L-1}}{P_{N-1}} \leq \frac{\varepsilon}{2}$$

contradicting $\varepsilon \in (0,2)$. Hence the lemma follows.

$\square$

6 Effective rates of asymptotic regularity

The first step towards proving the convergence of the Halpern iterations is to obtain the so-called 'asymptotic regularity' and this can be done in the very general setting of W-hyperbolic spaces.

Asymptotic regularity is a very important concept in metric fixed-point theory, formally introduced by Browder and Petryshyn in [7]. A mapping T of a metric space (X, d) into itself is said to be asymptotically regular if $\lim_{n\to\infty} d(x_n, Tx_n) = 0$ for all $x \in X$, where $x_n := T^n x$ is the

Picard iteration starting with x. We shall say that a sequence (y_n) in X is *asymptotically regular* if $\lim_{n\to\infty} d(y_n, Ty_n) = 0$. A rate of convergence of $(d(y_n, Ty_n))_n$ towards 0 will be called a rate of asymptotic regularity.

The following two propositions provide effective rates of asymptotic regularity for the Halpern iteration. Proposition 6.1 generalizes to W-hyperbolic spaces a result obtained by the second author for Banach spaces [29]. Proposition 6.2 is new even for the case of Banach spaces.

Let (X, d, W) be a W-hyperbolic space, $C \subseteq X$ be a bounded convex subset with diameter d_C, $T : C \to C$ be nonexpansive and (x_n) given by (11).

Proposition 6.1. *Assume that (λ_n) satisfies (C1), (C2) and (C3). Then (x_n) is asymptotically regular and $\lim_{n\to\infty} d(x_n, x_{n+1}) = 0$.*

Furthermore, if α is a rate of convergence of (λ_n), β is a Cauchy modulus of $s_n := \sum_{i=1}^{n} |\lambda_{i+1} - \lambda_i|$ and θ is a rate of divergence of $\sum_{n=1}^{\infty} \lambda_{n+1}$, then for all $\varepsilon \in (0, 2)$,

$$\forall n \geq \tilde{\Phi}\ \left(d(x_n, x_{n+1}) \leq \varepsilon\right) \quad \textit{and} \quad \forall n \geq \Phi\ \left(d(x_n, Tx_n) \leq \varepsilon\right),$$

where

$$\tilde{\Phi} := \tilde{\Phi}(\varepsilon, M, \theta, \beta) \;:=\; \theta\left(\beta\left(\frac{\varepsilon}{4M}\right) + 1 + \left\lceil \ln\left(\frac{2M}{\varepsilon}\right)\right\rceil\right) + 1, \tag{41}$$

$$\Phi := \Phi(\varepsilon, M, \theta, \alpha, \beta) \;=\; \max\left\{\theta\left(\beta\left(\frac{\varepsilon}{8M}\right) + 1 + \left\lceil \ln\left(\frac{4M}{\varepsilon}\right)\right\rceil\right) + 1, \alpha\left(\frac{\varepsilon}{4M}\right)\right\}, \tag{42}$$

with $M \in \mathbb{Z}_+$ such that $M \geq d_C$.

Proof. See Section 7. $\square$

Thus, we obtain an effective rate of asymptotic regularity $\Phi(\varepsilon, M, \theta, \alpha, \beta)$ which depends only on the error ε, on an upper bound M on the diameter d_C of C, and on (λ_n) via α, β, θ. In particular, the rate Φ does not depend on u, x or T, so Proposition 6.1 provides a quantitative version of the main theorem in [1]. Note that what is called 'property I' and 'property S' in [1] has been studied under the name of 'axioms (W2) and (W4)' in [21].

Proposition 6.2. *Assume that $\lambda_n \in (0, 1)$ for all $n \geq 2$ and that (λ_n) satisfies (C1), (C2) and (C4). Then (x_n) is asymptotically regular and $\lim_{n\to\infty} d(x_n, x_{n+1}) = 0$.*

Furthermore, if α is a rate of convergence of (λ_n), β is a Cauchy modulus of $s_n := \sum_{i=1}^{n} |\lambda_{i+1} - \lambda_i|$ and θ is a rate of convergence of $\prod_{n=1}^{\infty}(1 - \lambda_{n+1}) = 0$ towards 0, then for all $\varepsilon \in (0, 2)$,

$$\forall n \geq \tilde{\Phi}\ \left(d(x_n, x_{n+1}) \leq \varepsilon\right) \quad \textit{and} \quad \forall n \geq \Phi\ \left(d(x_n, Tx_n) \leq \varepsilon\right),$$

where

$$\tilde{\Phi}(\varepsilon, M, \theta, \beta, D) \;:=\; \theta\left(\frac{D\varepsilon}{2M}\right) + 1, \tag{43}$$

$$\Phi(\varepsilon, M, \theta, \alpha, \beta, D) \;=\; \max\left\{\theta\left(\frac{D\varepsilon}{4M}\right) + 1, \alpha\left(\frac{\varepsilon}{4M}\right)\right\}, \tag{44}$$

with $M \in \mathbb{Z}_+$ such that $M \geq d_C$ and $0 < D \leq \prod_{n=1}^{\beta(\varepsilon/4M)} (1 - \lambda_{n+1})$.

Proof. Follow the proof of Proposition 6.1, applying Lemma 5.5.(ii) instead of Lemma 5.5.(i). □

That we even get full rates of convergence in Propositions 6.1, 6.2 is due to the fact that the original proof of asymptotic regularity is essentially constructive. For such proofs, the requirement of the statement to be proved to have the form $\forall x \exists y\, A_{qf}(x, y)$ with quantifier-free A_{qf}, which is crucial for ineffective proofs, is not needed (note that the Cauchy property is a $\forall\exists\forall$-statement). This is because we do not have to preprocess the proof using some negative translation (which maps proofs with classical logic into ones with constructive logic only) and can directly apply proof-theoretic techniques such as (an appropriate monotone form of) Kreisel's so-called modified realizability interpretation. Logical metatheorems covering such situations are proved in [13]. As a consequence of getting full rates of convergence in Propositions 6.1, 6.2 one then also has to strengthen the premises on the convergence of (λ_n) and $\sum\limits_{n=1}^{\infty} |\lambda_{n+1} - \lambda_n|$ by full rates of convergence α, β. If we would interpret the proof as an ineffective one using the metatheorems from [21], then one would only get a rate of metastability in the conclusion but also would only need rates of metastability for these premises (note that $\sum\limits_{n=1}^{\infty} \lambda_n = \infty$ is a $\forall\exists$-statement so that there is no difference here between a full rate and a rate of metastability).

As an immediate consequence of Proposition 6.2, for $\lambda_n = \dfrac{1}{n+1}$ we get a quadratic (in $1/\varepsilon$) rate of asymptotic regularity. For Banach spaces, this rate of asymptotic regularity was obtained by the first author in [23]. In [29], the second author obtained an exponential rate of asymptotic regularity due to the fact that he used the version for Banach spaces of Proposition 6.1, which needs a rate of divergence of $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n+1}$.

Corollary 6.3. *Assume that* $\lambda_n = \dfrac{1}{n+1}$ *for all* $n \geq 1$*. Then for all* $\varepsilon \in (0, 1)$*,*

$$\forall n \geq \tilde{\Psi}(\varepsilon, M) \ \ (d(x_n, x_{n+1}) \leq \varepsilon) \quad \textit{and} \quad \forall n \geq \Psi(\varepsilon, M) \ \ (d(x_n, Tx_n) \leq \varepsilon), \tag{45}$$

where

$$\tilde{\Psi}(\varepsilon, M) := \left\lceil \frac{2M}{\varepsilon} + \frac{8M^2}{\varepsilon^2} \right\rceil - 1 \quad \textit{and} \quad \Psi(\varepsilon, M) := \left\lceil \frac{4M}{\varepsilon} + \frac{16M^2}{\varepsilon^2} \right\rceil - 1, \tag{46}$$

with $M \in \mathbb{Z}_+$ *such that* $M \geq d_C$*.*

Proof. Obviously, $\lim\limits_{n\to\infty} \dfrac{1}{n+1} = 0$ with a rate of convergence $\alpha(\varepsilon) = \left\lceil \dfrac{1}{\varepsilon} \right\rceil - 1 \geq 1$. As we have already seen, $\theta(\varepsilon) := \left\lceil \dfrac{2}{\varepsilon} \right\rceil - 2$ is a rate of convergence of $\displaystyle\prod_{n=1}^{\infty} \left(1 - \frac{1}{n+2}\right)$ towards 0. Furthermore,

$$s_n := \sum_{k=1}^{n} \left| \frac{1}{k+2} - \frac{1}{k+1} \right| = \frac{1}{2} - \frac{1}{n+2}.$$

It follows easily that $\lim\limits_{n\to\infty} s_n = 1/2$ with Cauchy modulus $\beta(\varepsilon) := \begin{cases} \lceil 1/\varepsilon \rceil - 1 & \text{if } \varepsilon \geq 1/2 \\ \lceil 1/\varepsilon \rceil - 2 & \text{if } \varepsilon < 1/2 \end{cases}$.

Finally, $\displaystyle\prod_{n=1}^{\beta(\varepsilon/4M)} \left(1 - \frac{1}{n+2}\right) = \frac{2}{\lceil 4M/\varepsilon \rceil}$, as $\dfrac{\varepsilon}{4M} < \dfrac{1}{2}$, so we can take $D := \dfrac{2}{\lceil 4M/\varepsilon \rceil}$. Apply now Proposition 6.2 and use the fact that $\lceil x \rceil \leq x + 1$ to get the result. □

7 Proof of Proposition 6.1

The following lemma collects some useful properties of Halpern iterations that hold for unbounded C too.

Lemma 7.1. *Assume that (x_n) is the Halpern iteration starting with $x \in C$. Then*

(i) For all $n \geq 0$,

$$d(x_{n+1}, Tx_n) = \lambda_{n+1} d(Tx_n, u) \quad \text{and} \quad d(x_{n+1}, u) = (1 - \lambda_{n+1}) d(Tx_n, u). \tag{47}$$

(ii) For all $n \geq 0$,

$$\begin{aligned} d(Tx_n, u) &\leq d(u, Tu) + d(x_n, u), & (48) \\ d(x_n, Tx_n) &\leq d(x_{n+1}, x_n) + \lambda_{n+1} d(Tx_n, u), & (49) \\ d(x_{n+1}, u) &\leq (1 - \lambda_{n+1})\big(d(u, Tu) + d(x_n, u)\big), & (50) \\ d(x_{n+1}, x_n) &\leq \lambda_{n+1} d(x_n, u) + (1 - \lambda_{n+1}) d(Tx_n, x_n). & (51) \end{aligned}$$

(iii) For all $n \geq 1$,

$$d(x_{n+1}, x_n) \leq (1 - \lambda_{n+1}) d(x_n, x_{n-1}) + |\lambda_{n+1} - \lambda_n|\, d(u, Tx_{n-1}). \tag{52}$$

(iv) If (x_n) is bounded, then (Tx_n) is also bounded. Moreover, if $M \geq d(u, Tu)$ and $M \geq d(x_n, u)$ for all $n \geq 0$,

$$\begin{aligned} d(x_n, Tx_n) &\leq d(x_{n+1}, x_n) + 2M\lambda_{n+1} \text{ and} & (53) \\ d(x_{n+1}, x_n) &\leq (1 - \lambda_{n+1}) d(x_n, x_{n-1}) + 2M|\lambda_{n+1} - \lambda_n| & (54) \end{aligned}$$

for all $n \geq 1$.

Proof. (i) By (1).

(ii)

$$\begin{aligned} d(Tx_n, u) &\leq d(u, Tu) + d(Tu, Tx_n) \leq d(u, Tu) + d(x_n, u), \\ d(x_n, Tx_n) &\leq d(x_{n+1}, x_n) + d(Tx_n, x_{n+1}) = d(x_{n+1}, x_n) + \lambda_{n+1} d(Tx_n, u) \\ d(x_{n+1}, u) &= (1 - \lambda_{n+1}) d(Tx_n, u) \leq (1 - \lambda_{n+1})\big(d(u, Tu) + d(x_n, u)\big) \\ d(x_{n+1}, x_n) &\leq \lambda_{n+1} d(x_n, u) + (1 - \lambda_{n+1}) d(x_n, Tx_n) \quad \text{by (W1)}. \end{aligned}$$

(iii)

$$\begin{aligned} d(x_{n+1}, x_n) &= d(\lambda_{n+1} u \oplus (1 - \lambda_{n+1}) Tx_n, \lambda_n u \oplus (1 - \lambda_n) Tx_{n-1}) \\ &\leq d(\lambda_{n+1} u \oplus (1 - \lambda_{n+1}) Tx_n, \lambda_{n+1} u \oplus (1 - \lambda_{n+1}) Tx_{n-1}) \\ &\quad + d(\lambda_{n+1} u \oplus (1 - \lambda_{n+1}) Tx_{n-1}, \lambda_n u \oplus (1 - \lambda_n) Tx_{n-1}) \\ &\leq (1 - \lambda_{n+1}) d(Tx_n, Tx_{n-1}) + |\lambda_{n+1} - \lambda_n| d(u, Tx_{n-1}) \\ &\quad \text{by (W4) and (W2)} \\ &\leq (1 - \lambda_{n+1}) d(x_n, x_{n-1}) + |\lambda_{n+1} - \lambda_n| d(u, Tx_{n-1}). \end{aligned}$$

(iv) is an easy consequence of (ii), (iii).

□

In the following, we give the proof of Proposition 6.1.

Let us consider the sequences

$$a_n := d(x_n, x_{n-1}), \quad b_n := 2M|\lambda_{n+1} - \lambda_n|$$

By (54), we get that

$$a_{n+1} \leq (1 - \lambda_{n+1})a_n + b_n \quad \text{for all} \quad n \geq 1.$$

Moreover, $\sum_{n=1}^{\infty} \lambda_{n+1}$ is divergent with rate of divergence θ and it is easy to see that

$$\gamma : (0, \infty) \to \mathbb{Z}_+, \quad \gamma(\varepsilon) := \beta\left(\frac{\varepsilon}{2M}\right)$$

is a Cauchy modulus of $s_n := \sum_{i=1}^{n} b_i$.

Thus, the hypotheses of Lemma 5.5.(i) are satisfied, so we can apply it to get that for all $\varepsilon \in (0, 2)$ and for all $n \geq \tilde{\Phi}(\varepsilon, M, \theta, \beta)$

$$d(x_n, x_{n-1}) \leq \varepsilon, \tag{55}$$

where

$$\tilde{\Phi}(\varepsilon, M, \theta, \beta) \quad := \quad \theta\left(\beta\left(\frac{\varepsilon}{4M}\right) + 1 + \left\lceil \ln\left(\frac{2M}{\varepsilon}\right)\right\rceil\right) + 1.$$

By (53), for all $n \geq 2$,

$$d(x_{n-1}, Tx_{n-1}) \leq d(x_n, x_{n-1}) + 2M\lambda_n. \tag{56}$$

Since α is a rate of convergence of (λ_n) towards 0, we get that

$$2M\lambda_n \leq \frac{\varepsilon}{2} \quad \text{for all } n \geq \alpha\left(\frac{\varepsilon}{4M}\right). \tag{57}$$

Combining (55), (56) and (57) it follows that

$$d(x_{n-1}, Tx_{n-1}) \leq \varepsilon$$

for all $n \geq \max\left\{\tilde{\Phi}\left(\frac{\varepsilon}{2}, M, \theta, \beta\right), \alpha\left(\frac{\varepsilon}{4M}\right)\right\}$, so the conclusion of the theorem follows.

8 Elimination of Banach limits

Let us recall that a *Banach limit* [4] is a linear functional $\mu : \ell^\infty \to \mathbb{R}$ satisfying the following properties:

(i) $\mu((x_n)) \geq 0$ if $x_n \geq 0$ for all $n \geq 0$;

(ii) $\mu(\mathbf{1}) = 1$;

(iii) $\mu((x_n)) = \mu((x_{n+1}))$.

Here $\mathbf{1}$ is the sequence $(1, 1, \ldots)$ and (x_{n+1}) is the sequence $(x_1, x_2, \ldots)$.

As we have already said, to prove the existence of Banach limits one needs the axiom of choice (see, e.g., [45]). Banach limits are mainly used in Saejung's convergence proof to get the following:

Lemma 8.1. *[44] Let $(a_k) \in \ell^\infty$ and $a \in \mathbb{R}$ be such that $\mu((a_k)) \leq a$ for all Banach limits μ and $\limsup_{k\to\infty}(a_{k+1} - a_k) \leq 0$. Then $\limsup_{k\to\infty} a_k \leq a$.*

Given a sequence $(a_k)_{k\geq 1}$, consider for all $n,p \geq 1$ the following average

$$C_{n,p}((a_k)) = \frac{1}{p}\sum_{i=n}^{n+p-1} a_i. \tag{58}$$

For simplicity we shall write $C_{n,p}(a_k)$.

Lemma 8.1 is proved using a result that goes back to Lorentz [34].

Lemma 8.2. *Let $(a_k) \in \ell^\infty$ and $a \in \mathbb{R}$. The following are equivalent:*

(i) $\mu((a_k)) \leq a$ for all Banach limits μ.

(ii) For all $\varepsilon > 0$ there exists $P \geq 1$ such that $C_{n,p}(a_k) \leq a + \varepsilon$ for all $p \geq P$ and $n \geq 1$.

In fact, one only needs the implication '(i) $\Rightarrow$ (ii)' which is established in [44] using the following sublinear functional

$$q : l^\infty \to \mathbb{R}, \quad q((a_k)) := \limsup_{p\to\infty} \sup_{n\geq 1} \frac{1}{p}\sum_{i=n}^{n+p-1} a_i = \limsup_{p\to\infty} \sup_{n\geq 1} C_{n,p}(a_k).$$

Now fix $(a_k) \in l^\infty$ and use the Hahn-Banach theorem to show the existence of a linear functional $\mu : l^\infty \to \mathbb{R}$ such that $\mu \leq q$ and $\mu((a_k)) = q((a_k))$. Then μ is a Banach limit and so – by (i) – $q((a_k)) = \mu((a_k)) \leq a$ which gives (ii). Our elimination of the use of the Banach limit μ was obtained in two steps: first, the proof that – for the sequence in question in the proof from [42] – the fact $\mu((a_k)) \leq a$ holds for all Banach limits μ could be modified to directly showing this for q instead of μ. This already established the actual elimination of the use of the axiom of choice hidden in the application of the Hahn-Banach theorem (for the nonseparable space l^∞) since the existence of q follows by just using uniform arithmetical comprehension in the form of an operator $E : \mathbb{N}^\mathbb{N} \to \{0,1\}$ defined by

$$E(f) = 0 \leftrightarrow \forall n \in \mathbb{N}(f(n) = 0),$$

that is needed (and sufficient) to form both the 'sup' as well as the 'lim sup' in the definition of q (as a function in (a_k)). Using an argument due to Feferman [12], the use of E can (over the system used to formalize the overall proof) be eliminated in favor of ordinary (non-uniform) arithmetic comprehension

$$\forall f : \mathbb{N}^2 \to \mathbb{N}\, \exists g : \mathbb{N} \to \mathbb{N}\, \forall k \in \mathbb{N}\, \big(g(k) = 0 \leftrightarrow \forall n \in \mathbb{N}\, (f(k,n) = 0)\big),$$

which is covered (as a very special case of general comprehension over numbers) by the existing logical metatheorems and results in extractable bounds of restricted complexity, namely bounds that are definable by primitive recursive functionals in the extended sense of Gödel's calculus T [17] (which, however, contains the famous so-called Ackermann function), though in general not of ordinarily primitive recursive type.

In order to get a bound having the latter much more restricted complexity we – in a second step – also eliminated the use of q in favor of just elementary lemmas on the finitary objects $C_{n,p}$. In the following, rather than going through these two steps separately, we just present the resulting elementary lemmas on the averages $C_{n,p}$ which we will need later. The first lemma collects some obvious facts.

Lemma 8.3. *Let $(a_k), (b_k)$ be sequences of real numbers and $\alpha \in \mathbb{R}$.*

(i) If $a_k \leq b_k$ for all $k \geq N$, then $C_{n,p}(a_k) \leq C_{n,p}(b_k)$ for all $n \geq N$ and $p \geq 1$.

(ii) If $a_k = c \in \mathbb{R}$ for all $k \geq N$, then $C_{n,p}(a_k) = c$ for all $n \geq N$ and $p \geq 1$.

(iii) For all $n,p \geq 1$, $C_{n,p}(a_k + b_k) = C_{n,p}(a_k) + C_{n,p}(b_k)$ and $C_{n,p}(\alpha a_k) = \alpha C_{n,p}(a_k)$.

Lemma 8.4. *Let (a_k) be a sequence of real numbers, $a \in \mathbb{R}$ and $P : (0, \infty) \to \mathbb{Z}_+$ be such that*

$$\forall \varepsilon > 0 \, \forall n \geq 1 \, \big(C_{n,P(\varepsilon)}(a_k) \leq a + \varepsilon\big). \tag{59}$$

Assume that $\limsup\limits_{k\to\infty}(a_{k+1} - a_k) \leq 0$ with effective rate θ.

Then $\limsup\limits_{k\to\infty} a_k \leq a$ with effective rate ψ, given by

$$\psi(\varepsilon, P, \theta) = \theta\left(\frac{\varepsilon}{\tilde{P}+1}\right) + \tilde{P}, \tag{60}$$

where $\tilde{P} := P\left(\frac{\varepsilon}{2}\right)$.

Proof. By hypothesis,

$$C_{n,\tilde{P}}(a_k) \leq a + \frac{\varepsilon}{2} \quad \text{for all } n \geq 1,$$

and

$$a_{k+1} - a_k \leq \frac{\varepsilon}{\tilde{P}+1} \quad \text{for all } k \geq \theta\left(\frac{\varepsilon}{\tilde{P}+1}\right).$$

Let $n \geq \psi(\varepsilon, P, \theta)$. Then $n = n_0 + \tilde{P}$ for some $n_0 \geq \theta\left(\frac{\varepsilon}{\tilde{P}+1}\right)$. We get that for each $i = 0, \ldots, \tilde{P} - 1$,

$$\begin{aligned} a_n &= a_{n_0+\tilde{P}} = a_{n_0+i} + (a_{n_0+i+1} - a_{n_0+i}) + (a_{n_0+i+2} - a_{n_0+i+1}) + \ldots + (a_{n_0+\tilde{P}} - a_{n_0+\tilde{P}-1}) \\ &\leq a_{n_0+i} + \frac{(\tilde{P}-i)\varepsilon}{\tilde{P}+1}. \end{aligned}$$

By adding the inequalities, we get that

$$\begin{aligned} \tilde{P}a_n &\leq \big(a_{n_0} + a_{n_0+1} + \ldots + a_{n_0+\tilde{P}-1}\big) + \frac{(1+2+\ldots+\tilde{P})\varepsilon}{\tilde{P}+1} \\ &= \big(a_{n_0} + a_{n_0+1} + \ldots + a_{n_0+\tilde{P}-1}\big) + \frac{\tilde{P}\varepsilon}{2}, \end{aligned}$$

hence

$$a_n \leq C_{n_0,\tilde{P}}(a_k) + \frac{\varepsilon}{2} \leq a + \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = a + \varepsilon.$$

□

Lemma 8.5. *Assume that (a_k) is nonnegative and $\lim\limits_{k\to\infty} a_k = 0$. Then $\lim\limits_{p\to\infty} C_{n,p}(a_k) = 0$ uniformly in n.*

Furthermore, if φ is a rate of convergence of (a_k), then for all $\varepsilon \in (0, 2)$,

$$\forall p \geq P(\varepsilon, \varphi, L) \, \forall n \geq 1 \ \ (C_{n,p}(a_k) \leq \varepsilon),$$

where

$$P(\varepsilon, \varphi, L) = \left\lceil \frac{2L\varphi(\varepsilon/2)}{\varepsilon} \right\rceil, \tag{61}$$

with $L \in \mathbb{R}$ being an upper bound on (a_k).

Proof. Let φ, L, ε be as in the hypothesis. We shall denote $P(\varepsilon, \varphi, L)$ simply by P. Since φ is a rate of convergence of (a_k), we have that $a_k \leq \frac{\varepsilon}{2}$ for all $k \geq \varphi(\varepsilon/2)$. Furthermore,

$$\frac{L\varphi(\varepsilon/2)}{p} \leq \frac{\varepsilon}{2} \quad \text{for all } p \geq P. \tag{62}$$

Let $p \geq P$ and $n \geq 1$. We have two cases:

(i) $n \geq \varphi(\varepsilon/2)$. Then

$$C_{n,p}(a_k) = \frac{1}{p}\sum_{i=n}^{n+p-1} a_i \leq \frac{1}{p} \cdot \frac{p\varepsilon}{2} = \frac{\varepsilon}{2} < \varepsilon.$$

(ii) $n < \varphi(\varepsilon/2)$. Then

$$\begin{aligned} C_{n,p}(a_k) &\leq \frac{1}{p}\sum_{i=n}^{\varphi(\varepsilon/2)-1} a_i + \frac{1}{p}\sum_{i=\varphi(\varepsilon/2)}^{\varphi(\varepsilon/2)+p-1} a_i \leq \frac{(\varphi(\varepsilon/2)-n)\,L}{p} + \frac{\varepsilon}{2} \leq \frac{\varphi(\varepsilon/2)L}{p} + \frac{\varepsilon}{2} \\ &\leq \varepsilon. \end{aligned}$$

Thus, we have proved that $C_{n,p}(a_k) \leq \varepsilon$ for all $p \geq P$ and $n \geq 1$. □

9 Quantitative properties of an approximate fixed point sequence

In the following, X is a complete CAT(0) space, $C \subseteq X$ is a bounded convex closed subset and $T : C \to C$ is a nonexpansive mapping. We assume that C is bounded with diameter d_C and consider $M \in \mathbb{Z}_+$ with $M \geq d_C$.

For $t \in (0, 1)$ and $u \in C$, define

$$T_t^u : C \to C, \quad T_t^u(y) = tu \oplus (1-t)Ty. \tag{63}$$

It is easy to see that T_t^u is a contraction with contractive constant $L := 1-t$, so it has a unique fixed point $z_t^u \in C$ by Banach's Contraction Mapping Principle. Hence, z_t^u is the unique solution of the fixed point equation

$$z_t^u = tu \oplus (1-t)Tz_t^u. \tag{64}$$

Proposition 9.1. *Let (y_n) be a sequence in C, $u \in C$, $t \in (0, 1)$, and (z_t^u) be defined by (64). Define for all $n \geq 1$*

$$\gamma_n^t := (1-t)d^2(u, Tz_t^u) - d^2(y_n, u). \tag{65}$$

(i) For all $n \geq 1$,

$$d^2(y_n, z_t^u) \leq d^2(y_n, u) + \frac{1}{t}a_n - (1-t)d^2(u, Tz_t^u), \tag{66}$$

where

$$a_n := d^2(y_n, Ty_n) + 2Md(y_n, Ty_n). \tag{67}$$

(ii) If (y_n) is asymptotically regular with rate of asymptotic regularity φ, then for all $\varepsilon \in (0, 2)$,

$$\forall p \geq P(\varepsilon, t, M, \varphi)\, \forall m \geq 1\ \left(C_{m,p}(\gamma_n^t) \leq \varepsilon\right), \tag{68}$$

where

$$P(\varepsilon, t, M, \varphi) = \left\lceil \frac{6M^2}{t\varepsilon}\varphi\left(\frac{t\varepsilon}{6M}\right)\right\rceil. \tag{69}$$

(iii) *Assume that* (y_n) *is asymptotically regular and* $\lim\limits_{n\to\infty} d(y_n, y_{n+1}) = 0$. *Then* $\limsup\limits_{n\to\infty} \gamma_n^t \leq 0$. *Furthermore, if* φ *is a rate of asymptotic regularity of* (y_n), *and* $\tilde{\varphi}$ *is a rate of convergence of* $(d(y_n, y_{n+1}))$ *towards* 0, *then* $\limsup\limits_{n\to\infty} \gamma_n^t \leq 0$ *with effective rate* ψ, *defined by*

$$\psi(\varepsilon, t, M, \varphi, \tilde{\varphi}) = \tilde{\varphi}\left(\frac{\varepsilon}{2M(P(\varepsilon/2, t, M, \varphi) + 1)}\right) + P(\varepsilon/2, t, M, \varphi), \tag{70}$$

with P *given by (69).*

Proof. For simplicity, we shall denote z_t^u by z_t.

(i) We get that for all $n \geq 1$,

$$\begin{aligned}
d^2(y_n, z_t) &= d^2(y_n, tu \oplus (1-t)Tz_t) \\
&\leq td^2(y_n, u) + (1-t)d^2(y_n, Tz_t) - t(1-t)d^2(u, Tz_t) \quad \text{by (4)} \\
&\leq td^2(y_n, u) + (1-t)\big(d(y_n, Ty_n) + d(Ty_n, Tz_t)\big)^2 - t(1-t)d^2(u, Tz_t) \\
&\quad \text{by the triangle inequality} \\
&\leq td^2(y_n, u) + (1-t)\big(d(y_n, Ty_n) + d(y_n, z_t)\big)^2 - t(1-t)d^2(u, Tz_t) \\
&\quad \text{by the nonexpansiveness of } T \\
&= td^2(y_n, u) + (1-t)d^2(y_n, Ty_n) + 2(1-t)d(y_n, Ty_n)d(y_n, z_t) \\
&\quad +(1-t)d^2(y_n, z_t) - t(1-t)d^2(u, Tz_t) \\
&\leq td^2(y_n, u) + (1-t)d^2(y_n, Ty_n) + 2M(1-t)d(y_n, Ty_n) + (1-t)d^2(y_n, z_t) \\
&\quad -t(1-t)d^2(u, Tz_t)
\end{aligned}$$

Thus, for all $n \geq 1$,

$$\begin{aligned}
td^2(y_n, z_t) &\leq td^2(y_n, u) + (1-t)d^2(y_n, Ty_n) + 2M(1-t)d(y_n, Ty_n) - t(1-t)d^2(u, Tz_t) \\
&\leq td^2(y_n, u) + d^2(y_n, Ty_n) + 2Md(y_n, Ty_n) - t(1-t)d^2(u, Tz_t).
\end{aligned}$$

Hence, (66) follows.

(ii) Let $\varepsilon \in (0, 2)$. By (66), we get that

$$0 \leq d^2(y_n, u) + \frac{1}{t}a_n - (1-t)d^2(u, Tz_t),$$

hence $\gamma_n^t \leq \dfrac{1}{t}\, a_n$ for all $n \geq 1$. It follows by Lemma 8.3.(i) that

$$C_{m,p}(\gamma_n^t) \leq C_{m,p}\left(\frac{1}{t}\, a_n\right) \quad \text{for all } m \geq 1,\ p \geq 1.$$

Furthermore, $\lim\limits_{n\to\infty} a_n = \lim\limits_{n\to\infty} \big(d^2(y_n, Ty_n) + 2Md(y_n, Ty_n)\big) = 0$ and, given a rate of asymptotic regularity for (y_n), we can easily verify that $\varphi\left(\dfrac{\varepsilon}{3M}\right)$ is a rate of convergence of (a_n) towards 0.

Then $\varphi\left(\dfrac{t\varepsilon}{3M}\right)$ is a rate of convergence of $\dfrac{1}{t}\, a_n$ towards 0. Since $L := \dfrac{3M^2}{t}$ is an upper bound for $\left(\frac{1}{t}a_n\right)$, we can apply Lemma 8.5 for this sequence to conclude that

$$C_{m,p}\left(\frac{1}{t}a_n\right) \leq \varepsilon \quad \text{for all } p \geq P(\varepsilon, t, M, \varphi) \text{ and } m \geq 1. \tag{71}$$

(iii) We have that

$$\begin{aligned}|\gamma_{n+1}^t - \gamma_n^t| &= |((1-t)d^2(u,Tz_t) - d^2(y_{n+1},u)) - ((1-t)d^2(u,Tz_t) - d^2(y_n,u))|\\ &= |d^2(y_n,u) - d^2(y_{n+1},u)| = |d(y_n,u) + d(y_{n+1},u)| \cdot |d(y_n,u) - d(y_{n+1},u)|\\ &\leq 2Md(y_n,y_{n+1}).\end{aligned}$$

Since $\lim\limits_{n\to\infty} 2Md(y_n,y_{n+1}) = 0$, we get that

$$\limsup_{n\to\infty}(\gamma_{n+1}^t - \gamma_n^t) \leq 0.$$

with effective rate $\tilde{\varphi}\left(\frac{\varepsilon}{2M}\right)$. Apply (ii) and Lemma 8.4 to conclude that

$$\limsup_{n\to\infty} \gamma_n^t \leq 0. \tag{72}$$

with effective rate $\psi(\varepsilon, t, M, \varphi, \tilde{\varphi})$.

□

Lemma 9.2. *Let $u, x \in C$ and (x_n) be the Halpern iteration defined by (11). Then for all $t \in (0,1)$ and $n \geq 0$,*

$$d^2(x_{n+1}, z_t^u) \leq (1-\lambda_{n+1})d^2(x_n, z_t^u) + \lambda_{n+1}\Big((1-t)d^2(u,Tz_t^u) - d^2(x_{n+1},u)\Big) + M^2 t. \tag{73}$$

Proof.

$$\begin{aligned}d^2(x_{n+1}, z_t^u) &\leq \lambda_{n+1}d^2(u,z_t^u) + (1-\lambda_{n+1})d^2(Tx_n, z_t^u) - \lambda_{n+1}(1-\lambda_{n+1})d^2(u,Tx_n)\\ &\quad \text{by (4) applied to } d^2(x_{n+1},z_t^u) = d^2(\lambda_{n+1}u \oplus (1-\lambda_{n+1})Tx_n, z_t^u)\\ &\leq \lambda_{n+1}d^2(u,z_t^u) - \lambda_{n+1}(1-\lambda_{n+1})d^2(u,Tx_n) +\\ &\quad +(1-\lambda_{n+1})\Big(td^2(Tx_n,u) + (1-t)d^2(Tx_n,Tz_t^u) - t(1-t)d^2(u,Tz_t^u)\Big)\\ &\quad \text{again by (4) applied to } d^2(Tx_n,z_t^u) = d^2(Tx_n, tu \oplus (1-t)Tz_t^u)\\ &\leq \lambda_{n+1}d^2(u,z_t^u) - \lambda_{n+1}(1-\lambda_{n+1})d^2(u,Tx_n) +\\ &\quad +(1-\lambda_{n+1})\Big(td^2(Tx_n,u) + (1-t)d^2(x_n,z_t^u) - t(1-t)d^2(u,Tz_t^u)\Big)\\ &\quad \text{by the nonexpansiveness of } T\\ &= (1-\lambda_{n+1})(1-t)d^2(x_n,z_t^u) +\\ &\quad +d^2(Tx_n,u)\Big((1-\lambda_{n+1})t - \lambda_{n+1}(1-\lambda_{n+1})\Big) +\\ &\quad +\lambda_{n+1}(1-t)^2d^2(u,Tz_t^u) - (1-\lambda_{n+1})t(1-t)d^2(u,Tz_t^u)\\ &\quad \text{since } d(u,z_t^u) = (1-t)d(u,Tz_t^u)\\ &= (1-\lambda_{n+1})(1-t)d^2(x_n,z_t^u) +\\ &\quad +\lambda_{n+1}\Big((1-t)d^2(u,Tz_t^u) - (1-\lambda_{n+1})^2d^2(Tx_n,u)\Big)\\ &\quad +d^2(Tx_n,u)\Big((1-\lambda_{n+1})t - \lambda_{n+1}(1-\lambda_{n+1}) + \lambda_{n+1}(1-\lambda_{n+1})^2\Big) +\\ &\quad +d^2(u,Tz_t^u)\Big(\lambda_{n+1}(1-t)^2 - (1-\lambda_{n+1})t(1-t) - \lambda_{n+1}(1-t)\Big)\end{aligned}$$

$$
\begin{aligned}
&= (1-\lambda_{n+1})(1-t)d^2(x_n, z_t^u) + \lambda_{n+1}\Big((1-t)d^2(u, Tz_t^u) - d^2(x_{n+1}, u)\Big) \\
&\quad +d^2(Tx_n, u)\Big(t - \lambda_{n+1}t + \lambda_{n+1}^3 - \lambda_{n+1}^2\Big) + d^2(u, Tz_t^u)(t^2 - t) \\
&\quad \text{since } d(x_{n+1}, u) = (1-\lambda_{n+1})d(Tx_n, u) \\
&\leq (1-\lambda_{n+1})(1-t)d^2(x_n, z_t^u) + \lambda_{n+1}\Big((1-t)d^2(u, Tz_t^u) - d^2(x_{n+1}, u)\Big) + \\
&\quad +td^2(Tx_n, u) \\
&\leq (1-\lambda_{n+1})d^2(x_n, z_t^u) + \lambda_{n+1}\Big((1-t)d^2(u, Tz_t^u) - d^2(x_{n+1}, u)\Big) + M^2 t.
\end{aligned}
$$

□

In [6], Browder showed that for Hilbert spaces X and z_t^u defined as above one has, for $t \to 0$, the strong convergence of z_t^u towards the fixed point of T that is closest to u. Halpern [18] gave a much more elementary proof of this result. In fact, it follows from his proof that the strong convergence of $(z_{t_k}^u)_k$ holds for any nonincreasing sequence (t_k) in $(0,1)$ (while the limit in general will not be a fixed point of T unless t_k converges towards 0). In [23], the first author extracted explicit and highly uniform rates of metastability from both proofs (again effective rates of convergence are ruled out on general grounds, see [23]). In [20], Kirk showed that Halpern's proof goes through (essentially unchanged) in the context of CAT(0) spaces. Consequently, this also holds for the bound extracted from Halpern's proof in [23] (for the Hilbert ball this is already due to [16]):

Proposition 9.3. *Let (t_k) be a nonincreasing sequence in $(0,1)$. Then for all $\varepsilon > 0$ and $g : \mathbb{N} \to \mathbb{N}$ the following holds*

$$\exists K_0 \leq K(\varepsilon, g, M)\, \forall i,j \in [K_0, K_0 + g(K_0)]\ \big(d(z_{t_i}^u, z_{t_j}^u) \leq \varepsilon\big),$$

where

$$K(\varepsilon, g, M) := \tilde{g}^{(\lceil M^2/\varepsilon^2 \rceil)}(0), \tag{74}$$

with $\tilde{g}(k) := k + g(k)$.

Proof. For the case of X being a Hilbert space, Proposition 9.3 is proved in [23]. Things extend unchanged to the CAT(0)-setting with the same reasoning as in [20]. □

Remark 9.4. *(i) Reasoning as in [23], Proposition 9.3 implies the following rate of metastability for sequences (t_k) that are not necessarily nonincreasing: let $(t_k)_{k\geq 0}$ be a sequence in $(0,1)$ that converges towards 0 with rate of convergence β and $\chi : \mathbb{N} \to \mathbb{N}$ be defined by $\chi(k) = \beta\left(\frac{1}{k+1}\right)$, hence*

$$\forall k \in \mathbb{N}\, \forall i \geq \chi(k)\ \left(t_i \leq \frac{1}{k+1}\right).$$

Finally, let $h : \mathbb{N} \to \mathbb{N}$ be such that $t_k \geq \frac{1}{h(k)+1}$ for all $k \in \mathbb{N}$. Then for all $\varepsilon > 0$ and $g : \mathbb{N} \to \mathbb{N}$ the following holds

$$\exists K_0 \leq K(\varepsilon, g, M, \chi)\, \forall i,j \in [K_0, K_0 + g(K_0)]\ \big(d(z_{t_i}^u, z_{t_j}^u) \leq \varepsilon\big),$$

where

$$K(\varepsilon, g, M, \chi, h) := \chi^+\big(g_{h,\chi}^{(\lceil 4M^2/\varepsilon^2 \rceil)}(0)\big),\ \textit{with } g_{h,\chi}(k) := \max\{h(i) \mid i \leq \chi(k) + g(\chi(k))\}.$$

(ii) Instead of a rate of convergence β it suffices in '(i)' above to have a rate of metastability β_g, hence a mapping β_g such that

$$\forall k \in \mathbb{N}\, \forall i \in [\beta_g(k), \tilde{g}(\beta_g(k))]\ \left(t_i \leq \frac{1}{k+1}\right).$$

10 Proof of Theorem 4.2

Let $\varepsilon \in (0,2)$ and $g : \mathbb{N} \to \mathbb{N}$ be fixed. Let $\tilde{\Phi}, \Phi$ be as in Proposition 6.1. To make the proof easier to read, we shall omit parameters $M, \Phi, \tilde{\Phi}, \theta, \alpha, \beta$ for all the functionals which appear in the following.

Take

$$\varepsilon_0 := \frac{\varepsilon^2}{24(M+1)^2}\,. \tag{75}$$

Then $\varepsilon_0 < 1$ and

$$\varepsilon_0^2 + 2M\varepsilon_0 + M^2\varepsilon_0 \le \varepsilon_0(M+1)^2 \le \frac{\varepsilon^2}{24}. \tag{76}$$

We consider in the sequel $t_k := \dfrac{1}{k+1}$, with rate of convergence towards 0 given by $\gamma(\varepsilon) := \left\lceil \dfrac{1}{\varepsilon} \right\rceil$.

Denote $z^u_{t_k}$ simply by z^u_k and let

$$\gamma^k_n := \frac{k}{k+1} d^2(u, Tz^u_k) - d^2(x_{n+1}, u).$$

Thus, γ^k_n is defined as in (65) by taking $t := \dfrac{1}{k+1}$ and $y_n := x_{n+1}$.

We can apply Propositions 9.1.(iii) and 6.1 to conclude that $\limsup\limits_{n\to\infty} \gamma^k_n \le 0$ for each $k \ge 0$, with effective rate χ_k, given by

$$\begin{aligned}
\chi_k(\varepsilon) &= \tilde{\Phi}\left(\frac{\varepsilon}{2M(\tilde{P}_k(\varepsilon)+1)}\right) + \tilde{P}_k(\varepsilon)\,, \text{ where} \\
\tilde{P}_k(\varepsilon) &= \left\lceil \frac{12M^2(k+1)}{\varepsilon}\Phi\left(\frac{\varepsilon}{12M(k+1)}\right)\right\rceil.
\end{aligned}$$

For all $k \ge 0$, let us denote

$$\begin{aligned}
\chi^*_k(\varepsilon) &:= \chi_k(\varepsilon/2), \\
\Theta_k(\varepsilon) &:= \Theta(\varepsilon, M^2, \theta, \chi^*_k) = \theta\left(\chi^*_k(\varepsilon/3) - 1 + \left\lceil \ln\left(\frac{3M^2}{\varepsilon}\right)\right\rceil\right) + 1, \\
\Delta^*_k(\varepsilon, g) &:= \Delta(\varepsilon, g, M^2, \theta, \chi^*_k) = \frac{\varepsilon}{3g_{\varepsilon,k}\left(\Theta_k(\varepsilon) - \chi^*_k(\varepsilon/3)\right)},
\end{aligned}$$

where $g_{\varepsilon,k}(n) = n + g(n + \chi^*_k(\varepsilon/3))$, Θ is defined by (15) and Δ by (16). Now let

$$f, f^* : \mathbb{N} \to \mathbb{N},\ f(k) := \max\left\{\gamma\left(\frac{\Delta^*_k(\varepsilon^2/4, g)}{M^2}\right), k\right\} - k, \quad f^*(k) := f(k + \gamma(\varepsilon_0)) + \gamma(\varepsilon_0).$$

We can apply Proposition 9.3 for ε_0 and f^* to get the existence of $K_1 \le K(\varepsilon_0, f^*)$ such that for all $k, l \in [K_1, K_1 + f^*(K_1)]$

$$d(z^u_k, z^u_l) \le \varepsilon_0, \tag{77}$$

where K is defined by (74). Let

$$\begin{aligned}
K_0 &:= K_1 + \gamma(\varepsilon_0), \\
K^*(\varepsilon_0, f) &:= K(\varepsilon_0, f^*) + \gamma(\varepsilon_0) = \widetilde{f^*}^{(\lceil M^2/\varepsilon_0^2\rceil)}(0) + \gamma(\varepsilon_0),
\end{aligned}$$

with $\widetilde{f^*}(k) := k + f^*(k)$.

Then $\gamma(\varepsilon_0) \le K_0 \le K^*(\varepsilon_0, f)$ and it is easy to see, using (77), that

$$\forall k, l \in [K_0, K_0 + f(K_0)] \; \big(d(z_k^u, z_l^u) \le \varepsilon_0\big). \tag{78}$$

It follows that for all $k, l \in [K_0, K_0 + f(K_0)]$,

$$\begin{aligned} d^2(u, Tz_k^u) &\le (d(u, Tz_l^u) + d(Tz_l^u, Tz_k^u))^2 \le d^2(u, Tz_l^u) + d^2(Tz_l^u, Tz_k^u) + 2d(u, Tz_l^u)d(Tz_l^u, Tz_k^u) \\ &\le d^2(u, Tz_l^u) + d^2(z_l^u, z_k^u) + 2Md(z_l^u, z_k^u) \\ &\le d^2(u, Tz_l^u) + \varepsilon_0^2 + 2M\varepsilon_0. \end{aligned}$$

Let

$$J := K_0 + f(K_0) = \max\left\{\gamma\left(\frac{\Delta^*_{K_0}(\varepsilon^2/4, g)}{M^2}\right), K_0\right\} \tag{79}$$

Then for all $n \ge 1$,

$$\begin{aligned} \gamma_n^J &= \frac{J}{J+1} d^2(u, Tz_J^u) - d^2(x_{n+1}, u) \le \frac{J}{J+1}\big(d^2(u, Tz_{K_0}^u) + \varepsilon_0^2 + 2M\varepsilon_0\big) - d^2(x_{n+1}, u) \\ &\le d^2(u, Tz_{K_0}^u) - d^2(x_{n+1}, u) + \varepsilon_0^2 + 2M\varepsilon_0 \\ &= \frac{K_0}{K_0+1} d^2(u, Tz_{K_0}^u) - d^2(x_{n+1}, u) + \varepsilon_0^2 + 2M\varepsilon_0 + \frac{1}{K_0+1} d^2(u, Tz_{K_0}^u) \\ &= \gamma_n^{K_0} + \varepsilon_0^2 + 2M\varepsilon_0 + \frac{1}{K_0+1} d^2(u, Tz_{K_0}^u) \le \gamma_n^{K_0} + \varepsilon_0^2 + 2M\varepsilon_0 + \frac{1}{K_0+1} M^2 \\ &\le \gamma_n^{K_0} + \varepsilon_0^2 + 2M\varepsilon_0 + M^2\varepsilon_0 \qquad \text{as } K_0 \ge \gamma(\varepsilon_0) \\ &\le \gamma_n^{K_0} + \frac{\varepsilon^2}{24} \quad \text{by (76).} \end{aligned}$$

It follows that for all $n \ge \chi^*_{K_0}(\varepsilon^2/12)$,

$$\gamma_n^J \le \gamma_n^{K_0} + \frac{\varepsilon^2}{24} \le \frac{\varepsilon^2}{12}.$$

Applying (73) with $t := \frac{1}{J+1}$, we get that for all $n \ge 1$,

$$\begin{aligned} d^2(x_{n+1}, z_J^u) &\le (1 - \lambda_{n+1}) d^2(x_n, z_J^u) + \lambda_{n+1}\left(\frac{J}{J+1} d^2(u, Tz_J^u) - d^2(x_{n+1}, u)\right) + \frac{M^2}{J+1} \\ &= (1 - \lambda_{n+1}) d^2(x_n, z_J^u) + \lambda_{n+1}\gamma_n^J + \frac{M^2}{J+1} \\ &\le (1 - \lambda_{n+1}) d^2(x_n, z_J^u) + \lambda_{n+1}\gamma_n^J + \Delta^*_{K_0}(\varepsilon^2/4, g) \end{aligned}$$

since $J \ge \gamma\left(\frac{\Delta^*_{K_0}(\varepsilon^2/4, g)}{M^2}\right)$, hence $\frac{1}{J+1} \le \frac{\Delta^*_{K_0}(\varepsilon^2/4, g)}{M^2}$. It follows that we can apply Lemma 5.2 with $\varepsilon := \varepsilon^2/4$ to conclude that for all $n \in [N, N + g(N)]$

$$d^2(x_n, z_J^u) \le \frac{\varepsilon^2}{4}, \quad \text{hence} \quad d(x_n, z_J^u) \le \frac{\varepsilon}{2}, \tag{80}$$

where $N := \Theta_{K_0}(\varepsilon^2/4)$.

Let now

$$\begin{aligned}
\theta^+(n) &:= \max\{\theta(i) \mid i \le n\}, \\
\Gamma &:= \max\{\chi_k^*(\varepsilon^2/12) \mid \gamma(\varepsilon_0) \le k \le K^*(\varepsilon_0, f)\} \ge \chi_{K_0}^*(\varepsilon^2/12), \\
\Sigma(\varepsilon, g) &:= \theta^+\left(\Gamma - 1 + \left\lceil \ln\left(\frac{12M^2}{\varepsilon^2}\right)\right\rceil\right) + 1 \\
&\ge \theta^+\left(\chi_{K_0}^*(\varepsilon^2/12) - 1 + \left\lceil \ln\left(\frac{12M^2}{\varepsilon^2}\right)\right\rceil\right) + 1 \\
&\ge \theta\left(\chi_{K_0}^*(\varepsilon^2/12) - 1 + \left\lceil \ln\left(\frac{12M^2}{\varepsilon^2}\right)\right\rceil\right) + 1 \\
&= \Theta_{K_0}(\varepsilon^2/4) = N.
\end{aligned}$$

We get finally that $N \le \Sigma(\varepsilon, g)$ is such that for all $n, m \in [N, N + g(N)]$,

$$d(x_n, x_m) \le d(x_n, z_J^u) + d(x_m, z_J^u) \le \varepsilon. \tag{81}$$

$\square$

Acknowledgements:

Ulrich Kohlenbach has been supported by the German Science Foundation (DFG Project KO 1737/5-1). Part of his research has been carried out while visiting the Simion Stoilow Institute of Mathematics of the Romanian Academy supported by the BITDEFENDER guest professor program.

Laurenţiu Leuştean has been supported by a grant of the Romanian National Authority for Scientific Research, CNCS - UEFISCDI, project number PN-II-ID-PCE-2011-3-0383.

References

[1] K. Aoyama, K. Eshita, W. Takahashi, Iteration processes for nonexpansive mappings in convex metric spaces, in: Proceedings of the International Conference on Nonlinear and Convex Analysis, Yokohama Publishers, 2007, 31-39.

[2] K. Aoyama, Y. Kimura, W. Takahashi, M. Toyoda, Approximation of common fixed points of a countable family of nonexpansive mappings in a Banach space, Nonlinear Anal. 67 (2007), 2350-2360.

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

[4] S. Banach, Théorie des opérations linéaires, Monografie Mat., PWN, Warszawa, 1932.

[5] M. Bridson, A. Haefliger, Metric spaces of non-positive curvature, Grundlehren der Mathematischen Wissenschaften 319, Springer-Verlag, Berlin, 1999.

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

[7] F.E. Browder, W.V. Petryshyn, The solution by iteration of nonlinear functional equations in Banach spaces, Bull. Amer. Math. Soc. 72 (1966), 571–575.

[8] R.E. Bruck, S. Reich, Accretive operators, Banach limits, and dual ergodic theorems, Bull. Acad. Polon. Sci. 29 (1981), 585–589.

[9] M. Bruhat, J. Tits, Groupes réductifs sur un corps local. I. Données radicielles valuées, Inst. Hautes Études Sci. Publ. Math. 41 (1972), 5–251.

[10] C.E. Chidume, C.O. Chidume, Iterative approximation of fixed points of nonexpansive mappings, J. Math. Anal. Appl. 318 (2006), 288–295.

[11] S. Dhompongsa, B. Panyanak, On Δ-convergence theorems in CAT(0) spaces, Computers and Math. with Applications 56 (2008), 2572-2579.

[12] S. Feferman, Theories of finite type related to mathematical practice, in: J. Barwise(ed.), Handbook of Mathematical Logic, North-Holland, Amsterdam, 1977, 913-972.

[13] P. Gerhardy, U. Kohlenbach, Strongly uniform bounds from semi-constructive proofs, Ann. Pure Appl. Logic 141 (2006), 89-107.

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

[15] K. Goebel, W. A. Kirk, Iteration processes for nonexpansive mappings, in: S.P. Singh, S. Thomeier, B. Watson (eds.), Topological methods in nonlinear functional analysis. Proceedings of the special session on fixed point theory and applications held during the 86th summer meeting of the American Mathematical Society at the University of Toronto, Toronto, Ont., August 21-26, 1982, Contemp. Math. 21, Amer. Math. Soc., Providence, RI, 1983, 115–123.

[16] K. Goebel, S. Reich, Uniform convexity, hyperbolic geometry, and nonexpansive mappings, Marcel Dekker, Inc., New York and Basel, 1984.

[17] K. Gödel, Über eine bisher noch nicht benützte Erweiterung des finiten Standpunktes, Dialectica 12 (1958), 280-287.

[18] B. Halpern, Fixed points of nonexpanding maps, Bull. Amer. Math. Soc. 73 (1967), 957-961.

[19] W. A. Kirk, Krasnosel'skii iteration process in hyperbolic spaces, Numer. Funct. Anal. and Optimiz. 4 (1982), 371-381.

[20] W.A. Kirk, Geodesic geometry and fixed point theory, Seminar of Mathematical Analysis (Malaga/Seville, 2002/2003), Colecc. Abierta, 64, Univ. Seville Secr. Publ., Seville (2003), 195-225.

[21] U. Kohlenbach, Some logical metatheorems with applications in functional analysis. Trans. Amer. Math. Soc. 357 (2005), 89-128.

[22] U. Kohlenbach, Applied proof theory: Proof interpretations and their use in mathematics, Springer Monographs in Mathematics, Springer-Verlag, Berlin-Heidelberg, 2008.

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

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

[25] U. Kohlenbach, L. Leuştean, On the computational content of convergence proofs via Banach limits, to appear in: Phil. Trans. Royal Soc. A.

[26] E. Kopecká, S. Reich, A mean ergodic theorem for nonlinear semigroups on the Hilbert ball, J. Nonlinear Convex Anal. 11 (2010), 185-197.

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

[28] G. Kreisel, On the interpretation of non-finitist proofs, part II, J. Symbolic Logic 17 (1952), 43-88.

[29] L. Leuştean, Rates of asymptotic regularity for Halpern iterations of nonexpansive mappings, J. Universal Computer Science 13 (2007), 1680-1691.

[30] L. Leuştean, Proof mining in metric fixed point theory and ergodic theory, Habilitation thesis, Technische Universität Darmstadt, 2009.

[31] M. Lin, R. Wittmann, Pointwise ergodic theorems for certain order preserving mappings in L^1, in: A. Bellow, R.L. Jones (eds.), Almost every where convergence. II. Proceedings of the Second International Conference on Almost Everywhere Convergence in Probability and Ergodic Theory held at Northwestern University, Evanston, Illinois, October 16–20, 1989, Academic Press, Inc., Boston, MA, 1991, 191–207.

[32] P.-L. Lions, Approximation de points fixes de contractions, C. R. Acad. Sci. Paris Série A 284 (1977), 1357–1359.

[33] L.-S. Liu, Ishikawa and Mann iterative process with errors for nonlinear strongly accretive mappings in Banach spaces, J. Math. Anal. Appl. 194 (1995), 114-125.

[34] G.G. Lorentz, A contribution to the theory of divergent series, Acta Math. 80 (1948), 167-190.

[35] S. Reich, Some fixed point problems, Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur. 57 (1974), 194-198.

[36] S. Reich, Almost convergence and nonlinear ergodic theorems, J. Approx. Theor. 24 (1978), 269-272.

[37] S. Reich, Strong convergence theorems for resolvents of accretive operators in Banach spaces, J. Math. Anal. Appl. 75 (1980), 287–292.

[38] S. Reich, Some problems and results in fixed point theory, in: S.P. Singh, S. Thomeier, B. Watson (eds.), Topological methods in nonlinear functional analysis. Proceedings of the special session on fixed point theory and applications held during the 86th summer meeting of the American Mathematical Society at the University of Toronto, Toronto, Ont., August 21-26, 1982, Contemp. Math. 21, Amer. Math. Soc., Providence, RI, 1983, 179–187.

[39] S. Reich, Approximating fixed points of nonexpansive mappings, Panamer. Math. J. 4 (1994), 23–28.

[40] S. Reich, I. Shafrir, Nonexpansive iterations in hyperbolic spaces, Nonlinear Anal. 15 (1990), 537–558.

[41] S. Reich, A. Wallwater, Almost convergence and a dual ergodic theorem for nonlinear semigroups, J. Nonlinear Convex Anal. 11 (2010), pp. 89-99.

[42] S. Saejung, Halpern iterations in CAT(0) spaces, Fixed Point Theory and Applications, Volume 2010, Article ID 471781, 13pp.

[43] K. Schade, Proof mining for Halpern iterations in CAT(0) spaces, Master Thesis, TU Darmstadt, 2012.

[44] N. Shioji, W. Takahashi, Strong convergence of approximated sequences for nonexpansive mappings in Banach spaces, Proc. Amer. Math. Soc. 125 (1997), 3641-3645.

[45] L. Sucheston, Banach limits, Amer. Math. Monthly 74 (1967), 308-311.

[46] T. Suzuki, A sufficient and necessary condition for Halpern-type strong convergence to fixed points of nonexpansive mappings, Proc. Amer. Math. Soc. 135 (2007), 99–108.

[47] T. Suzuki, Reich's problem concerning Halpern's convergence, Arch. Math. 92 (2009), 602-613.

[48] W. Takahashi, A convexity in metric space and nonexpansive mappings I, Kodai Math. Sem. Rep. 22 (1970), 142-149.

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

[50] T. Tao, Norm convergence of multiple ergodic averages for commuting transformations, Ergodic Theory Dynam. Systems 28 (2008), 657-688.

[51] R. Wittmann, Hopf ergodic theorem for nonlinear operators, Math. Ann. 289 (1991), 239-253.

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

[53] H.-K. Xu, Iterative algorithms for nonlinear operators, J. London Math. Soc. 66 (2002), 240–256.