Quantitative results on Fejér monotone sequences

Ulrich Kohlenbach[1], Laurenţiu Leuştean[2,3], Adriana Nicolae[4,5]

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

[2] Faculty of Mathematics and Computer Science, University of Bucharest,
Academiei 14, P.O. Box 010014, Bucharest, Romania

[3] Simion Stoilow Institute of Mathematics of the Romanian Academy,
P. O. Box 1-764, 014700 Bucharest, Romania

[4] Department of Mathematics, Babeş-Bolyai University,
Kogălniceanu 1, 400084 Cluj-Napoca, Romania

[5] Simion Stoilow Institute of Mathematics of the Romanian Academy,
Research group of the project PD-3-0152,
P. O. Box 1-764, 014700 Bucharest, Romania

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

December 21, 2015

Abstract

We provide in a unified way quantitative forms of strong convergence results for numerous iterative procedures which satisfy a general type of Fejér monotonicity where the convergence uses the compactness of the underlying set. These quantitative versions are in the form of explicit rates of so-called metastability in the sense of T. Tao. Our approach covers examples ranging from the proximal point algorithm for maximal monotone operators to various fixed point iterations (x_n) for firmly nonexpansive, asymptotically nonexpansive, strictly pseudo-contractive and other types of mappings. Many of the results hold in a general metric setting with some convexity structure added (so-called W-hyperbolic spaces). Sometimes uniform convexity is assumed still covering the important class of CAT(0)-spaces due to Gromov.

Keywords: Fejér monotone sequences, quantitative convergence, metastability, proximal point algorithm, firmly nonexpansive mappings, strictly pseudo-contractive mappings, proof mining.

1 Introduction

This paper provides in a unified way quantitative forms of strong convergence results for numerous iterative procedures which satisfy a general type of Fejér monotonicity where the convergence uses

the compactness of the underlying set. Fejér monotonicity is a key notion employed in the study of many problems in convex optimization and programming, fixed point theory and the study of (ill-posed) inverse problems (see e.g. [54, 10]). These quantitative forms have been obtained using the logic-based proof mining approach (as developed e.g. in [26]) but the results are presented here in a way which avoids any explicit reference to notions or tools from logic.
Our approach covers examples ranging from the proximal point algorithm for maximal monotone operators to various fixed point iterations (x_n) for firmly nonexpansive, asymptotically nonexpansive, strictly pseudo-contractive and other types of mappings. Many of the results hold in a general metric setting with some convexity structure added (so-called W-hyperbolic spaces in the sense of [25]). Sometimes uniform convexity is assumed still covering Gromov's CAT(0)-spaces.

For reasons from computability theory, effective rates of convergence for (x_n) in X are usually ruled out even when the space X in question and the map T used in the iteration are effective: usually (x_n) will converge to a fixed point of T but in general T will not possess a computable fixed point and even when it does (e.g. when X is $\mathbb{R}^n$ and the fixed point set is convex) the usual iterations will not converge to a computable point and hence will not converge with an effective rate of convergence (see [46] for details on all this).

The Cauchy property of (x_n) can, however, be reformulated in the equivalent form

$$(*) \quad \forall k \in \mathbb{N}\, \forall g : \mathbb{N} \to \mathbb{N}\, \exists N \in \mathbb{N}\, \forall i, j \in [N, N + g(N)] \; \left(d(x_i, x_j) \leq \frac{1}{k+1} \right)$$

and for this form, highly uniform computable bounds $\exists N \leq \Phi(k, g)$ on $\exists N$ can be obtained. $(*)$ is known in mathematical logic since 1930 as Herbrand normal form and bounds Φ have been studied in the so-called Kreisel no-counterexample interpretation (which in turn is a special case of the Gödel functional interpretation) since the 50's (see [26]). More recently, $(*)$ has been made popular under the name of 'metastability' by Terence Tao, who used the existence of uniform bounds on N in the context of ergodic theory ([52, 53]). Moreover, Walsh [55] used again metastability to show the L^2-convergence of multiple polynomial ergodic averages arising from nilpotent groups of measure-preserving transformations.

In nonlinear analysis, rates of metastability Φ for strong convergence results of nonlinear iterations have been first considered and extracted in [30, 24] (and in many other cases since then). The point of departure of our investigation is [24] which uses Fejér monotonicity and where some of the arguments of the present paper have first been used in a special context.

Let $F \subseteq X$ be a subset of X and recall that (x_n) is Fejér monotone w.r.t. F if

$$(+) \quad d(x_{n+1}, p) \leq d(x_n, p), \text{ for all } n \in \mathbb{N} \text{ and } p \in F.$$

We think of F as being the intersection $F = \bigcap_{k \in \mathbb{N}} AF_k$ of approximations $AF_{k+1} \subseteq AF_k \subseteq X$ to F, one prime example being $F := Fix(T)$ and $AF_k := \{p \in X \mid d(p, Tp) \leq 1/(k+1)\}$, where $Fix(T)$ denotes the fixed point set of some selfmap $T : X \to X$.

We say that (x_n) possesses 'approximate F-points' if

$$(**) \quad \forall k \in \mathbb{N}\, \exists n \in \mathbb{N}\, (x_n \in AF_k).$$

Finally, we define a notion of F being 'explicitly closed' w.r.t. the representation AF_k which is implied by all sets AF_k being closed and implies that F is closed (see Definition 3.3).

The main general result of our paper is a quantitative version of the following

Proposition (see Proposition 4.3) Let X be compact, F explicitly closed and (x_n) a sequence in X which is Fejér monotone w.r.t. F and possesses approximate F-points. Then (x_n) converges to a point in F.

In order to arrive at a quantitative form of this proposition we first need to enhance the assumptions used to appropriate quantitative versions ('moduli'). To be appropriate here means that

(i) these moduli are sufficient to compute a rate of metastability for (x_n) in terms of them and

(ii) it is guaranteed by general logical metatheorems that these moduli can actually be provided effectively in each concrete instantiation of the general result by analyzing given proofs of the respective properties in the case at hand.

Guided by the underlying logical methodology this leads to the following quantitative notions:

compactness	$\to$	modulus γ of total boundedness (Section 2)
explicit closedness	$\to$	moduli ω, δ of uniform closedness (Definition 3.4)
approximate F-points	$\to$	approximate F-point bound Φ (Section 5)
Fejér monotonicity	$\to$	modulus χ of uniform Fejér monotonicity (Definition 4.6).

The key notion in this paper is that of a modulus of uniform Fejér monotonicity and so will discuss this notion in more detail: it is a bound $\exists k \leq \chi(r,n,m)$ on the following uniform strengthening of 'Fejér monotone'

$$\forall r, n, m \in \mathbb{N} \exists k \in \mathbb{N} \forall p \in X \left(p \in AF_k \to \forall l \leq m \left(d(x_{n+l}, p) < d(x_n, p) + \frac{1}{r+1} \right) \right).$$

If X is compact and F is explicitly closed w.r.t. the sets AF_k, then 'Fejér monotone' and 'uniform Fejér monotone' are equivalent. However, moduli χ for uniform Fejér monotonicity can be extracted (based on results from logic) also in the absence of compactness, provided that the proof of the Fejér monotonicity is formalizable in a suitable context, and we provide such moduli χ in all our applications.

The main general quantitative theorem in our paper (Theorem 5.3) provides an explicit construction $\tilde{\Psi}$ that transforms any approximate F-point bound Φ, any modulus χ of uniform Fejér monotonicity, any modulus of total boundedness γ, and any moduli δ, ω of uniform closedness into a rate $\tilde{\Psi}(k,g) := \tilde{\Psi}(k, g, \Phi, \chi, \gamma, \delta, \omega)$ of metastability (x_n) such that all the points in the interval of metastability $[N, N+g(N)]$ belong to AF_k :

$$\forall k \in \mathbb{N}\, \forall g : \mathbb{N} \to \mathbb{N}\, \exists N \leq \tilde{\Psi}(k,g)\, \forall i, j \in [N, N+g(N)] \ \left(d(x_i, x_j) \leq \frac{1}{k+1} \text{ and } x_i \in AF_k \right).$$

This result is derived from a slightly simpler rate Ψ of metastability without the extra clause $x_i \in AF_k$ which then does not depend on δ, ω (Theorem 5.1). Let us discuss the general structure of Ψ (disregarding some inessential technical details): essentially Ψ is the $\gamma(k)$-times iterate of $\Phi \circ \chi \circ g$ (starting from 0). In particular, this yields that a rate of convergence for (x_n) (while not being computable) is effectively learnable with at most $\gamma(k)$-many mind changes and a learning strategy which - essentially - is $\Phi \circ \chi$ (see [32] for more on this). That a primitive recursive iteration of g is unavoidable follows from the fact that even for most simple cases of Fejér monotone fixed point iterations (x_n), namely Mann iterations of nonexpansive mappings in compact intervals of $\mathbb{R}$, the Cauchy property of (x_n) implies the Cauchy property of monotone sequences in $[0, C]$ (see [46]) which is equivalent to Σ^0_1-induction ([23, Corollary 5.3]). One can also obtain the Cauchy property of monotone sequences itself as an instance of our result which then gives back (up to a constant) the known optimal rate of metastability from [26], showing also that the number of iterations of the function g needed cannot be better than $\gamma(k)$ (see the example at the end of Section 5).

In all our general results we actually permit a more general form of Fejér monotonicity, where instead of $(+)$ one has

$$(++) \quad H(d(x_{n+m}, p)) \leq G(d(x_n, p)), \text{ for all } n, m \in \mathbb{N} \text{ and } p \in F$$

and $G, H : \mathbb{R}_+ \to \mathbb{R}_+$ are subject to very general conditions (this e.g. is used in the application to asymptotically nonexpansive mappings from Section 7.5).

As is typical for such quantitative 'finitizations' of noneffective convergence results, it is easy to incorporate a summable sequence (ε_n) of error terms in all the aforementioned results which covers the important concept of 'quasi-Fejér monotonicity' due to [12] (see Section 6). As a consequence of this, one can also incorporate such error terms in the iterations we are considering in this paper. However, for the sake of better readability we will not carry this out in the paper (but see [29, 33] for an applications of this to convex feasibility problems in CAT(κ)-spaces ($\kappa > 0$) and minimization problems in CAT(0)-spaces respectively).

The results mentioned so far hold for arbitrary sets $F = \bigcap_k AF_k$ provided that we have the various moduli as indicated. In the case where AF_k can be written as a purely universal formula and we have - sandwiched in between $AF_{k+1} \subseteq \tilde{AF}_k \subseteq AF_k$ - the sets $\tilde{AF}_k$ which are given by a purely existential formula (which is the case for $AF_k = \{p \in X \mid d(p, Tp) \leq 1/(k+1)\}$ with $\tilde{AF}_k = \{p \in X \mid d(p, Tp) < 1/(k+1)\}$), then the logical metatheorems from [25, 15, 26] **guarantee** the extractability of explicit and highly uniform moduli χ from proofs of (generalized) Fejér monotonicity, as well as approximate fixed point bounds or metastability rates for asymptotic regularity from proofs of the corresponding properties if these proofs can be carried out in suitable formal systems as in all our applications.

In this paper we apply this to Picard iterations of firmly nonexpansive mappings and to Ishikawa iterations of nonexpansive mappings in geodesic settings, Mann iterations of strict pseudo-contractions in Hilbert spaces, Mann iterations of Suzuki-type mappings and of asymptotically nonexpansive mappings in geodesic spaces and, finally, to the proximal point algorithm in Hilbert spaces.

As an example of how the extracted moduli look like, let us briefly consider the case of the Mann iteration

$$x_0 := x \in C, \; x_{n+1} := (1 - \lambda_n)x_n + \lambda_n T(x_n), \quad \text{where } (\lambda_n) \subseteq [0, 1]$$

for selfmappings $T : C \to C$ of convex subsets of CAT(0) spaces that satisfy a condition going back essentially to Suzuki [51]

$$(E) \; d(x, Ty) \leq \mu d(Tx, x) + d(x, y), \quad \text{where } \mu \geq 1.$$

In this case the moduli we extract are as follows:

$$\begin{aligned}\Phi(k) &= 384[(b+1)(k+1)L]^2,\\ \chi(n,m,r) &= \mu m(1-1/L)(r+1),\\ \delta(k) &= 2\mu(k+1)-1,\\ \omega(k) &= 4k+3,\end{aligned}$$

where $L \in \mathbb{N}$ is such that $(\lambda_n) \subseteq [1/L, 1-1/L]$.
Note that Φ is quadratic in the error $\varepsilon = 1/(k+1)$ as is the case in all our applications to fixed point theory when X is a CAT(0) (or Hilbert) space which is to be expected the optimal error-dependency even in the Hilbert space case (except for the real line).

Let us briefly discuss the role of 'proof mining' in this paper: as already mentioned, the definitions of the various types of quantitative moduli are instances of a general proof-theoretic transformation which then also guarantees the extractability of these moduli from given proofs. Also the construction of the bound $\tilde{\Psi}$ is obtained by logically analyzing the proof of the qualitative Proposition 4.3. Finally, in all the applications discussed in this paper, the bounds Φ are extracted with the help of this methodology (this also applies to the cases where we refer to such constructions from previous papers). The other moduli, such as χ, can be obtained more directly in our cases but e.g. in the application to convex feasibility problems in [29] it is the modulus χ whose construction requires heavy use of proof mining.
As common in the proof mining paradigm, the final proofs can be written up again in ordinary mathematical language without any use of logic which only was instrumental to find these proofs. As a consequence of this, the reader will not see any **explicit** use of logic in this paper. Let us, however, briefly comment on two logically interesting points in our treatment of the aforementioned mappings satisfying the condition (E) :

(i) The extraction of Φ proceeds via realizing that (E) implies that T satisfies a property which was first formulated in the proof mining context, namely being weakly quasi-nonexpansive (see [30] and - for a logical metatheorem designed for this class of mappings - [15]). As that property does not contribute to the quantitative bounds, it follows that Φ does not depend on the constant μ although it crucially features in the definition of (E). Note also that the extraction of Φ proceeds by extracting a more general rate of metastability for the asymptotic regularity which in turn uses a technical proof mining lemma from [30].

(ii) Although the construction of δ is quite simple in this case, there is a logically interesting point here explaining why this time we do have a dependence on μ while we do not need any uniform continuity requirement on T (as is always needed in the model-theoretic approaches to metric structures in positive bounded or continuous logic): the proof of (x_n) being Fejér monotone uses an extensionality $(=)$-axiom in the form

$$x = p \wedge p \in Fix(T) \to x \in Fix(T).$$

It is the use of such extensionality which quantitatively usually translates into uniform continuity. In this special case, however, we only need moduli $\theta, \tilde{\theta} : \mathbb{R}^*_+ \to \mathbb{R}^*_+$ such that

$$\forall \varepsilon > 0 \ \left(d(x,p) \leq \theta(\varepsilon) \wedge d(p,Tp) \leq \tilde{\theta}(\varepsilon) \to d(x,Tx) \leq \varepsilon \right),$$

which can be defined using (E) as $\theta(\varepsilon) := \varepsilon/4$, $\tilde{\theta}(\varepsilon) := \varepsilon/(2\mu)$. So, logically speaking, the somewhat ad-hoc condition (E) (in addition to implying that T is weakly quasi-nonexpansive) is nothing but a condition which guarantees a particular simple quantitative form of the instance of extensionality used in the proof.

The paper is organized as follows. In Section 2 we discuss the background from mathematical logic, i.e. so-called logical metatheorems (due to the first author in [25], see also [15, 26]) which provide tools for the extraction of highly uniform bounds from prima facie noneffective proofs of ∀∃-theorems (which covers the case of metastability statements). Since our present paper uses the context of totally bounded metric spaces we discuss this case in particular detail. Applying proof mining to a concrete proof results again in an ordinary proof in analysis and so one can read the proofs in this paper without any knowledge of logic which, however, was used by the authors to find these proofs. In Sections 3 and 4 we develop the basic definitions and facts about the sets F, AF_k, the notions of explicit and uniform closedness as well as (uniform) generalized (G, H)-Fejér monotone sequences. In Section 5 we establish our main general quantitative theorems which then will be specialized in our various applications. Section 6 generalizes these results to the case of (uniform) quasi-Fejér monotone sequences. In Section 7 we interpret our results in the case where F is the fixed point set of a selfmap T (mostly of some convex subset of X) and provide numerous applications as mentioned above: in each of these cases we provide appropriate moduli of uniform (generalized) Fejér monotonicity χ and approximate fixed point bounds Φ (usually even rates of asymptotic regularity or metastable versions thereof) so that our general quantitative theorems can be applied resulting in explicit rates of metastability for (x_n). In Section 8 we do the same for the case where F is the set of zeros of a maximal monotone operator and provide the corresponding moduli for the proximal point algorithm.

The results in this paper are based on compactness arguments. Without compactness one in general has only weak convergence for Fejér monotone sequences, but in important cases weakly convergent iterations can be modified to yield strong convergence even in the absence of compactness (see e.g. [3]). This phenomenon is known from fixed point theory, where Halpern-type variants of the weakly convergent Mann iteration yield strong convergence ([7, 20, 56]). Even when only weak convergence holds one can apply the logical machinery to extract rates of metastability for the weak Cauchy property (see e.g. [28] where this is done in the case of Baillon's nonlinear ergodic theorem). However, the bounds will be extremely complex. If, however, weak convergence is used only as an intermediate step towards strong convergence, one can often avoid the passage through weak convergence altogether and obtain much simpler rates of metastability (see e.g. [27] where this has been carried out in particular for Browder's classical strong convergence theorem of the resolvent of a nonexpansive operator in Hilbert spaces, as well as [35]). We believe that it is an interesting future research project to adapt these techniques to the context of Fejér monotone sequences.

Notations: $\mathbb{N}$ and $\mathbb{N}^*$ denote the set of natural numbers including 0 and without 0, respectively, while $\mathbb{R}_+$ are the nonnegative reals.

2 Quantitative forms of compactness

Let (X, d) be a metric space. We denote with $B(x, r)$ (resp. $\overline{B}(x, r)$) the open (resp. closed) ball with center $x \in X$ and radius $r > 0$.
Let us recall that a nonempty subset $A \subseteq X$ is *totally bounded* if for every $\varepsilon > 0$ there exists an ε-net

of A, i.e. there are $n \in \mathbb{N}$ and $a_0, a_1 \ldots, a_n \in X$ such that $A \subseteq \bigcup_{i=0}^{n} B(a_i, \varepsilon)$. This is equivalent with the existence of a $1/(k+1)$-net for every $k \in \mathbb{N}$.

Definition 2.1. *Let $\emptyset \neq A \subseteq X$. We call $\alpha : \mathbb{N} \to \mathbb{N}$ a* I-modulus of total boundedness *for A if for every $k \in \mathbb{N}$ there exist elements $a_0, a_1, \ldots, a_{\alpha(k)} \in X$ such that*

$$\forall x \in A \, \exists \, 0 \leq i \leq \alpha(k) \left(d(x, a_i) \leq \frac{1}{k+1} \right). \tag{1}$$

Thus, A is totally bounded iff A has a I-modulus of total boundedness. In this case, we also say that A is totally bounded with I-modulus α. One can easily see that any totally bounded set is bounded: given a I-modulus α and $a_0, \ldots, a_{\alpha(0)} \in X$ such that (1) is satisfied for $k = 0$, $b := 2 + \max\{d(a_i, a_j) \mid 0 \leq i, j \leq \alpha(0)\}$ is an upper bound on the diameter of A.

We now give an alternative characterization of total boundedness used in the context of proof mining first in [14]:

Definition 2.2. *Let $\emptyset \neq A \subseteq X$. We call $\gamma : \mathbb{N} \to \mathbb{N}$ a* II-modulus of total boundedness *for A if for any $k \in \mathbb{N}$ and for any sequence (x_n) in A*

$$\exists \, 0 \leq i < j \leq \gamma(k) \left(d(x_i, x_j) \leq \frac{1}{k+1} \right). \tag{2}$$

Remark 2.3. *The logarithm of the smallest possible value for a I-modulus of total boundedness is also called the $1/(k+1)$-entropy of A while the logarithm of the optimal II-modulus is called the $1/(k+1)$-capacity of A (see e.g. [41]).*

Proposition 2.4. *Let $\emptyset \neq A \subseteq X$.*

(i) If α is a I-modulus of total boundedness for A, then $\gamma(k) := \alpha(2k+1) + 1$ is a II-modulus of total boundedness for A.

(ii) If γ is a II-modulus of total boundedness for A, then $\alpha(k) := \gamma(k) - 1$ is a I-modulus of total boundedness (so, in particular, A is totally bounded).

Proof. (i) Let $a_0, \ldots, a_{\alpha(2k+1)} \in X$ be such that (1) is satisfied, hence for all $x \in A$ there exists $0 \leq i \leq \alpha(2k+1)$ such that $d(x, a_i) \leq \frac{1}{2k+2}$. Applying the pigeonhole principle to $x_0, x_1, \ldots, x_{\alpha(2k+1)+1}$, we get $0 \leq i < j \leq \alpha(2k+1) + 1$, such that x_i and x_j are in a ball of radius $\frac{1}{2k+2}$ around the same a_l with $0 \leq l \leq \alpha(2k+1)$. It follows that $d(x_i, x_j) \leq \frac{1}{k+1}$, hence (2) holds.

(ii) First, let us remark that $\gamma(k) \geq 1$ for all k, hence α is well-defined. Assume by contradiction that $\alpha(k) := \gamma(k) - 1$ is not a I-modulus of total boundedness, i.e. there exists $k \in \mathbb{N}$ such that

$$(*) \quad \forall a_0, \ldots, a_{\gamma(k)-1} \in X \, \exists x \in A \, \forall \, 0 \leq i \leq \gamma(k) - 1 \left(d(x, a_i) > \frac{1}{k+1} \right).$$

By induction on $l \leq \gamma(k)$ we show that

$$(**) \quad \exists \beta_0, \ldots, \beta_l \in A \, \forall \, 0 \leq i < j \leq l \left(d(\beta_i, \beta_j) > \frac{1}{k+1} \right),$$

which, for $l := \gamma(k)$, contradicts the assumption that γ is a II-modulus of total boundedness.
$l = 0$: Choose $\beta_0 \in A$ arbitrary.
$l \mapsto l + 1 \leq \gamma(k)$: let $\beta_0, \ldots, \beta_l$ be as in $(**)$. By $(*)$ applied to

$$a_i := \begin{cases} \beta_i & \text{if } i \leq l \\ \beta_l & \text{if } l < i \leq \gamma(k) - 1 \end{cases}$$

we get $x \in A$ such that $d(x, \beta_i) > \dfrac{1}{k+1}$ for all $i \leq l$. Then $\beta_0, \ldots, \beta_l, \beta_{l+1} := x$ satisfies $(**)$.
□

Note that the existence of a $1/(k + 1)$-net in the proof of Proposition 2.4.(ii) is noneffective. In particular, there is no effective way to compute a bound on A from a II-modulus of total boundedness. This seemingly disadvantage actually will allow us to extract bounds of greater uniformity from proofs of statements which do not explicitly refer to such a bound (see below).

2.1 General logical metatheorems for totally bounded metric spaces

In [25], the first author introduced so-called logical metatheorems for bounded metric structures (as well as for normed spaces and other classes of spaces).[1] Here systems $\mathcal{T}^\omega$ of arithmetic and analysis in the language of functionals of all finite types are extended by an abstract metric space X whose metric is supposed to be bounded by $b \in \mathbb{N}$ resulting in a system $\mathcal{T}^\omega[X, d]$. Consider now a $\mathcal{T}^\omega[X, d]$-proof of a theorem of the following form, where P is some concrete complete separable metric space and K a concrete compact metric space:[2]

$$(+) \quad \begin{cases} \forall u \in P \, \forall v \in K \, \forall x \in X \, \forall y \in X^{\mathbb{N}} \, \forall T : X \to X \\ \qquad (A_\forall(u, v, x, y, T) \to \exists n \in \mathbb{N} \, B_\exists(u, v, x, y, T)), \end{cases}$$

where $A_\forall, B_\exists$ are purely universal resp. purely existential sentences (with some restrictions on the types of the quantified variables). Then from the proof one can extract (using a method from proof theory called monotone functional interpretation, due to the first author, see [26] for details on all this) a computable uniform bound '$\exists n \leq \Phi(f_u, b)$' on '$\exists n \in \mathbb{N}$' which only depends on some representation f_u of u in P and a bound b of the metric. In particular, Φ does not depend on v, x, y, T nor on the space X (except for the bound b). In most of our applications P will be $\mathbb{N}$ or $\mathbb{N}^{\mathbb{N}}$ (with the discrete and the Baire metric, respectively) in which case $u = f_u$. In the cases $\mathbb{R}$ or $C[0, 1]$, however, f_u is some concrete fast Cauchy sequence (say of Cauchy rate 2^{-n}) of rationals representing $u \in \mathbb{R}$ resp. a pair (f, ω) with $f \in C[0, 1]$ and some modulus of uniform continuity ω for f in the case of $C[0, 1]$. f_u can always be encoded into an element of $\mathbb{N}^{\mathbb{N}}$.

Φ has some restricted subrecursive complexity which reflects the strength of the mathematical axioms from $\mathcal{T}^\omega$ used in the proof. In most applications, Φ is at most of so-called primitive recursive complexity.

As discussed in [15] and [26, Application 18.16, p. 464], the formalization of the total boundedness of X via the existence of a I-modulus of total boundedness α can be incorporated in this setting as follows: in order to simplify the logical structure of the axiom to be added it is convenient to

[1] In this discussion we focus on the case of metric spaces.
[2] For simplicity, we only consider here some special case. For results in full generality see [25, 15, 26].

combine all the individual ε-nets $a_0, \ldots, a_{\alpha(\varepsilon)}$ into one single sequence (a_n) of elements in X and to replace the quantification over $\varepsilon > 0$ by quantification over $\mathbb{N}$ via $\varepsilon := 1/(n+1)$:

The theory $\mathcal{T}^\omega[X, d, TOTI]$ of totally bounded metric spaces is obtained by adding to $\mathcal{T}^\omega[X, d]$

(i) two constants $\alpha^{\mathbb{N}\to\mathbb{N}}$ and $a^{\mathbb{N}\to X}$ denoting a function $\mathbb{N} \to \mathbb{N}$ and a sequence $\mathbb{N} \to X$, respectively, as well as

(ii) one universal axiom:[3]

$$(TOT\,I) \quad \forall k^{\mathbb{N}} \forall x^X \exists N \leq_{\mathbb{N}} \alpha(k) \left(d_X(x, a_N) \leq_{\mathbb{R}} \frac{1}{k+1} \right).$$

It is obvious that $(TOT\,I)$ implies that α is a I-modulus of total boundedness of X as defined before. Conversely, suppose α is such a modulus. Then $\alpha'(n) := \sum_{i=0}^{n}(\alpha(i)+1)$ satisfies $(TOT\,I)$ for the sequence (a_n) obtained as the concatenation of the $1/(k+1)$-nets $a_0^k, \ldots, a_{\alpha(k)}^k$, $k = 0, 1, \ldots$.

Since $(TOT\,I)$ is purely universal, its addition does not cause any problems and the only change caused by switching from $\mathcal{T}^\omega[X, d]$ to $\mathcal{T}^\omega[X, d, TOTI]$ is that the extracted bound Φ will additionally depend on α (see [26] for details).

In [15], the results from [25] are extended to the case of unbounded metric spaces. Then the bound Φ depends, instead of b, on majorizing data $x^* \gtrsim_X^p x$, $y^* \gtrsim_{\mathbb{N}\to X}^p y$, $T^* \gtrsim_{X\to X}^p T$ for x, y, T relative to some reference point $p \in X$ (which usually will be identified with x). More precisely, the p-majorizability relation $\gtrsim^p$ is defined (for the cases at hand which are special cases of a general inductive definition for all function types over $\mathbb{N}, X$ interpreted here over the full set-theoretic type structure, see [26]) as follows:

$$\begin{aligned}
&n^* \gtrsim_{\mathbb{N}}^p n := n^*, n \in \mathbb{N} \wedge n^* \geq n,\\
&\alpha^* \gtrsim_{\mathbb{N}\to\mathbb{N}}^p \alpha := \alpha^*, \alpha \in \mathbb{N}^{\mathbb{N}} \wedge \forall n^*, n(n^* \geq n \to \alpha^*(n^*) \geq \alpha^*(n), \alpha(n)),\\
&x^* \gtrsim_X^p x := x^* \in \mathbb{N}, x \in X \wedge x^* \geq d(p, x),\\
&y^* \gtrsim_{\mathbb{N}\to X}^p y := y^* \in \mathbb{N}^{\mathbb{N}}, y \in X^{\mathbb{N}} \wedge \forall n^*, n \in \mathbb{N}\, (n^* \geq n \to y^*(n^*) \geq d(p, y(n))),\\
&T^* \gtrsim_{X\to X}^p T := T^* \in \mathbb{N}^{\mathbb{N}}, T \in X^X \wedge\\
&\qquad\qquad \forall n \in \mathbb{N}\, \forall x \in X(n \geq d(p, x) \to T^*(n) \geq d(p, T(x)).
\end{aligned}$$

Note that $\gtrsim_{\mathbb{N}}^p$ and $\gtrsim_{\mathbb{N}\to\mathbb{N}}^p$ actually do not depend on p, hence we shall denote them simply $\gtrsim_{\mathbb{N}}$ and $\gtrsim_{\mathbb{N}\to\mathbb{N}}$, respectively. Whereas a majorant y^* exists for any sequence y in X, it is a genuine restriction on T to posses a majorant T^*. However, for large classes of mappings T one can construct T^*, e.g. this is the case when T is Lipschitz continuous (in the case of geodesic spaces also uniform continuity suffices) but also in general whenever T maps bounded sets to bounded sets.

It is instructive to see what happens if we take the context of **unbounded** metric spaces, i.e. - using the terminology from [15, 26] - $\mathcal{T}^\omega[X, d]_{-b}$ and add constants $\alpha : \mathbb{N} \to \mathbb{N}$ and $(a_n) : \mathbb{N} \to X$ as before. Then we need to provide majorants α^*, a^* for these objects, which in the case of α can be simply done by stipulating $\alpha^*(n) := \max\{\alpha(i) \mid \leq n\}$, whereas for (a_n) this requires - as above - a function $a^* : \mathbb{N} \to \mathbb{N}$ such that $a^* \gtrsim_{\mathbb{N}\to X}^p a$. Then the bound Φ extractable from proofs of theorems of the form considered above will additionally also depend on α^* (i.e. on α) and a^*. From these data one

[3] The bounded number quantifier can be easily eliminated by bounded collection.

can easily compute a bound b on X (e.g. we may take $b := 2 + 2a^*(\alpha^*(0))$) and, conversely, given such a bound b one can simply take $a^*(n) := b$. So, adding $(TOT\,I)$ gives in both contexts the same results w.r.t. the extractability of bounds Φ and their uniformity. This situation, however, changes if we consider the axiomatization based on the II-modulus of total boundedness in the setting of unbounded metric structures:

The theory $\mathcal{T}^\omega[X, d, TOT II]_{-b}$ of totally bounded metric spaces is obtained by adding to $\mathcal{T}^\omega[X, d]_{-b}$

(i) one constant $\gamma^{\mathbb{N}\to\mathbb{N}}$ and

(ii) one universal axiom:

$$(TOT\,II) \quad \forall k^{\mathbb{N}} \forall x^{\mathbb{N}\to X} \exists I, J \leq_{\mathbb{N}} \gamma(k) \left(I <_{\mathbb{N}} J \wedge d_X(x_I, x_J) \leq_{\mathbb{R}} \frac{1}{k+1} \right).$$

Due to the absence of the sequence (a_n) from this axiomatization, the extracted bounds will only depend on γ instead of α, a^* (or α, b). This results in a strictly greater uniformity of the bounds as the following example shows.

Consider the sequence (X, d_n) of metric spaces defined as follows:

$$X := \{0, 1\},\ d_n(0, 1) := d_n(1, 0) := n,\ d_n(0, 0) = d_n(1, 1) = 0.$$

It is easy to see that $\gamma(n) := 2$ is a common II-modulus of total boundedness for all the spaces (X, d_n) (since any sequence of 3 elements of X has to repeat some element), while the diameter of (X, d_n) tends to infinity as n does. Hence our bounds Φ will be uniform for all the spaces (X, d_n) which first might look impossible since, after all, $(TOT\,II)$ **does** imply that X is bounded, i.e.

$$(++) \quad \exists b \in \mathbb{N}\, \forall x, y \in X\, (d(x, y) < b).$$

However, $(++)$ is of the form $\exists\forall$, which is not allowed in statements of the form $(+)$ considered above. Noneffectively, $(++)$ can be equivalently reformulated as

$$(++)' \quad \forall (x_n), (y_n) \in X^{\mathbb{N}}\, \exists N \in \mathbb{N}\, (d(x_N, y_N) < N),$$

which **is** of the form $(+)$, so that the aforementioned uniform bound extraction applies (given majorants x^*, y^* for $(x_n), (y_n)$). Indeed, define recursively

$$n_0 := 0, \quad n_{k+1} := \left\lceil \max_{i,j \leq k} \{n_k, d(x_{n_i}, y_{n_j}), d(x_{n_i}, x_{n_j}), d(y_{n_i}, y_{n_j})\} + 3 \right\rceil,$$

which can easily be effectively bounded using only d and x^*, y^*.

Proposition 2.5. *For any metric space X with II-modulus of total boundedness γ we have:*

$$\exists N \leq n_{\gamma(0)}\, (d(x_N, y_N) < N).$$

Proof. Suppose that $\forall k \leq \gamma(0)(d(x_{n_k}, y_{n_k}) \geq n_k)$. Then, for all $k \leq \gamma(0)$, one of the two cases

$$(1)\ \forall i < k\, (d(x_{n_k}, x_{n_i}), d(x_{n_k}, y_{n_i}) > 1)$$

or

$$(2)\ \forall i < k\, (d(y_{n_k}, x_{n_i}), d(y_{n_k}, y_{n_i}) > 1)$$

holds since, otherwise, $d(x_{n_k}, y_{n_k}) \leq n_{k-1} + 2 < n_k$. Define a sequence $z_0, \ldots, z_{\gamma(0)}$ as follows: for $k \leq \gamma(0)$ put $z_k := x_{n_k}$, if (1) holds, and $z_k := y_{n_k}$, otherwise (which implies that (2) holds). Then $d(z_i, z_j) > 1$ whenever $0 \leq i < j \leq \gamma(0)$ which, however, contradicts the definition of γ. Hence $\exists k \leq \gamma(0)\ (d(x_{n_k}, y_{n_k}) < n_k)$. Since $n_k \leq n_{\gamma(0)}$, the claim follows. □

Remark 2.6. *As mentioned already, logical metatheorems of the form discussed above have also been established for more enriched structures such as W-hyperbolic spaces, uniformly convex W-hyperbolic spaces, $\mathbb{R}$-trees, δ-hyperbolic spaces (in the sense of Gromov) and CAT(0)-spaces as well normed spaces, uniformly convex normed spaces, complete versions of these spaces and Hilbert spaces. Most recently, also abstract L^p- and $C(K)$-spaces have been covered ([19]). In the normed case, the reference point $p \in X$ used in the majorization relation will always be the zero vector 0_X (see [15, 25, 26, 37, 38, 19] for all this). In all these cases one can add the requirement of X (or of some bounded subset in the normed case) to be totally bounded with moduli of total boundedness in the form I or II as above. Thus the applications given in this paper can be viewed as instances of corresponding logical metatheorems.*

2.2 Examples

In this subsection we give simple examples of II-moduli of total boundedness that are computed explicitly. Although some of the proofs are straightforward we include them for completeness.

Example 2.7. *Let $A = [0, C]$ be a compact interval in $\mathbb{R}$ ($C > 0$). Then $\gamma : \mathbb{N} \to \mathbb{N}$, $\gamma(k) = \lceil C \rceil (k+1)$ is a II-modulus of total boundedness for A.*

Proof. Let $k \in \mathbb{N}$ and (x_n) be a sequence in A. Divide the interval $[0, C]$ into $\lceil C \rceil (k+1)$ subintervals of length $\leq 1/(k+1)$. Applying the pigeonhole principle we obtain that there exist $0 \leq i < j \leq \lceil C \rceil (k+1)$ such that $|x_i - x_j| \leq 1/(k+1)$. □

Example 2.8. *Let A be a bounded subset of $\mathbb{R}^n$ and $b > 0$ be such that $\|a\|_2 \leq b$ for every $a \in A$. Then $\gamma : \mathbb{N} \to \mathbb{N}$, $\gamma(k) = \left\lceil 2(k+1)\sqrt{n} b \right\rceil^n$ is a II-modulus of total boundedness for A.*

Proof. Let $k \in \mathbb{N}$ and $(x_p) \subseteq A$. Denote $N = \lceil 2(k+1)\sqrt{n} b \rceil$. Clearly, A is included in the cube $[-b, b]^n$. Divide this cube into N^n subcubes of equal side lengths $2b/N$. The diameter of each subcube is $2b\sqrt{n}/N \leq 1/(k+1)$. Applying the pigeonhole principle we obtain that there exist $0 \leq i < j \leq N^n$ such that $\|x_i - x_j\|_2 \leq 1/(k+1)$. □

Example 2.9. *Let (X, d) be a metric space and $A \subseteq X$ totally bounded with II-modulus of total boundedness γ. Then the closure of A is totally bounded with II-modulus of total boundedness γ.*

Proof. Let $k \in \mathbb{N}$ and $(x_n) \subseteq \overline{A}$. Take $m \in \mathbb{N}$. Then there exists a sequence $(a_n) \subseteq A$ such that $d(x_n, a_n) \leq 1/(m+1)$. Since γ is a II-modulus of total boundedness for A, there exist $0 \leq i < j \leq \gamma(k)$ such that $d(a_i, a_j) \leq 1/(k+1)$. Thus,

$$d(x_i, x_j) \leq d(x_i, a_i) + d(a_i, a_j) + d(a_j, x_j) \leq \frac{1}{k+1} + \frac{2}{m+1}.$$

Hence, there exist $0 \leq i < j \leq \gamma(k)$ and (m_s) a strictly increasing sequence of natural numbers such that for every $s \geq 0$, $d(x_i, x_j) \leq 1/(k+1) + 2/(m_s+1)$, from where $d(x_i, x_j) \leq 1/(k+1)$. □

Example 2.10. *Let $(X, \|\cdot\|)$ be a normed space and $A \subseteq X$ totally bounded with II-modulus of total boundedness γ. Then the convex hull $co(A)$ of A is totally bounded with II-modulus of total boundedness*

$$\overline{\gamma}(k) = \left\lceil\, 2(m+1)\sqrt{n+1}\,\right\rceil^{n+1},$$

where $n = \gamma(4k+3) - 1$, $m = \lceil 2(k+1)(n+1)\left(b + 1/(4k+4)\right)\rceil - 1$ and $b > 0$ is such that $\|a\| \leq b$ for all $a \in A$.

Proof. Let $k \in \mathbb{N}$ and $(y_p) \subseteq \mathrm{co}(A)$. Denote $r_k = 1/(4k+4)$. By Proposition 2.4.(ii), there exist $a_0, \ldots, a_n \in A$ such that

$$A \subseteq \bigcup_{l=0}^{n} \overline{B}\left(a_l, r_k\right).$$

Let $p \in \mathbb{N}$. Then there exist $s(p) \in \mathbb{N}$ and for $l = 0, \ldots, s(p)$, $t_l^p \in [0,1]$ and $x_l^p \in A$ such that $\sum_{l=0}^{s(p)} t_l^p = 1$ and $y_p = \sum_{l=0}^{s(p)} t_l^p x_l^p$. We can assume that $s(p) = n$ and $x_l^p \in \overline{B}\left(a_l, r_k\right)$ for $l = 0, \ldots, n$. This can be done because we can group any two points that belong to the same ball in the following way: suppose $x_0^p, x_1^p \in \overline{B}\left(a_0, r_k\right)$. Denote

$$\overline{x}_0^p = \frac{t_0^p}{t_0^p + t_1^p} x_0^p + \frac{t_1^p}{t_0^p + t_1^p} x_1^p \in \overline{B}\left(a_0, r_k\right).$$

Then, $y_p = (t_0^p + t_1^p)\overline{x}_0^p + t_2^p x_2^p + \ldots + t_n^p x_n^p$. Note that if in this way we obtain less than $n+1$ points in the convex combination then we add the corresponding a_l's multiplied by 0.
For $p \in \mathbb{N}$, $t^p = (t_0^p, \ldots, t_n^p) \in \mathbb{R}^{n+1}$ and $\|t^p\|_2 \leq 1$. By Example 2.8, there exist $0 \leq i < j \leq \left\lceil\, 2(m+1)\sqrt{n+1}\,\right\rceil^{n+1}$ such that

$$\|t^i - t^j\|_2 \leq \frac{1}{m+1} \leq \frac{2r_k}{\left(b + r_k\right)(n+1)}.$$

Then,

$$\begin{aligned}
\|y_i - y_j\| &= \left\|\sum_{l=0}^{n}(t_l^i x_l^i - t_l^j x_l^j)\right\| \leq \left\|\sum_{l=0}^{n}(t_l^i x_l^i - t_l^i x_l^j)\right\| + \left\|\sum_{l=0}^{n}(t_l^i x_l^j - t_l^j x_l^j)\right\| \\
&\leq \sum_{l=0}^{n} t_l^i \|x_l^i - x_l^j\| + \|x_l^j\| \sum_{l=0}^{n} |t_l^i - t_l^j| \\
&\leq 2r_k \sum_{l=0}^{n} t_l^i + \left(b + r_k\right)(n+1)\|t^i - t^j\|_2 \leq 2r_k + 2r_k = \frac{1}{k+1}.
\end{aligned}$$

□

3 Approximate points and explicit closedness

In the following, (X, d) is a metric space and $F \subseteq X$ a nonempty subset. We assume that

$$F = \bigcap_{k \in \mathbb{N}} \tilde{F}_k,$$

where $\tilde{F}_k \subseteq X$ for every $k \in \mathbb{N}$ and we say that the family $(\tilde{F}_k)$ is a *representation* of F. Of course, F has a trivial representation, by letting $\tilde{F}_k := F$ for all k. Naturally, we think of more interesting choices for $\tilde{F}_k$, as we look at

$$AF_k := \bigcap_{l \leq k} \tilde{F}_l$$

as some weakened approximate form of F. A point $p \in AF_k$ is said to be a *k-approximate F-point.*

In the following we always view F not just as a set but we suppose it is equipped with a representation $(\tilde{F}_k)$ to which we refer implicitly in many of the notations introduced below.

Let (x_n) be a sequence in X.

Definition 3.1. *We say that*

(i) (x_n) *has approximate F-points if* $\forall k \in \mathbb{N}\, \exists N \in \mathbb{N}\, (x_N \in AF_k)$.

(ii) (x_n) *has the liminf property w.r.t. F if* $\forall k, n \in \mathbb{N}\, \exists N \in \mathbb{N}\ (N \geq n \text{ and } x_N \in AF_k)$.

(iii) (x_n) *is asymptotically regular w.r.t. F if* $\forall k \in \mathbb{N}\, \exists N \in \mathbb{N}\, \forall m \geq N\ (x_m \in AF_k)$.

Lemma 3.2. *Assume that $x_k \in AF_k$ for all $k \in \mathbb{N}$. Then any subsequence of (x_n) has the liminf property w.r.t. F.*

Proof. Let (x_{m_l}) be a subsequence of (x_n). Then $m_l \geq l$ and $x_{m_l} \in AF_{m_l}$ for all $l \in \mathbb{N}$. Let $k, n \in \mathbb{N}$ and take $N \geq \max\{n, k\}$. Then $x_{m_N} \in AF_{m_N} \subseteq AF_N \subseteq AF_k$. □

Definition 3.3. *We say that F is* explicitly closed *(w.r.t. the representation $(\tilde{F}_k)$) if*

$$\forall p \in X\ \left(\forall N, M \in \mathbb{N}(AF_M \cap \overline{B}\,(p, 1/(N+1)) \neq \emptyset) \rightarrow p \in F\right).$$

One can easily see that if F is explicitly closed, then F is closed. F in particular is explicitly closed if all the sets AF_k are (and so if all the sets $\tilde{F}_k$ are closed). Hence closedness of F is equivalent to explicit closedness of F w.r.t. the trivial representation. The property of being explicitly closed can be re-written (pulling also the quantifier hidden in '$p \in F$' in front) in the following equivalent form

$$\forall k \in \mathbb{N}\, \forall p \in X\, \exists N, M \in \mathbb{N}\ \left(AF_M \cap \overline{B}\,(p, 1/(N+1)) \neq \emptyset \rightarrow p \in AF_k\right).$$

This suggests the following uniform strengthening of explicit closedness:

Definition 3.4. *F is called* uniformly closed with moduli $\delta_F, \omega_F : \mathbb{N} \rightarrow \mathbb{N}$ *if*

$$\forall k \in \mathbb{N}\, \forall p, q \in X\ \left(q \in AF_{\delta_F(k)} \text{ and } d(p,q) \leq \frac{1}{\omega_F(k)+1} \rightarrow p \in AF_k\right).$$

Lemma 3.5. *Assume that F is explicitly closed, (x_n) has the liminf property w.r.t. F and that (x_n) converges strongly to $\widehat{x}$. Then $\widehat{x} \in F$.*

Proof. Let $k \in \mathbb{N}$ be arbitrary. Since F is explicitly closed, there exist $M, N \in \mathbb{N}$ such that

$$\exists q \in X\ \left(d(\widehat{x}, q) \leq \frac{1}{N+1} \text{ and } q \in AF_M\right) \rightarrow \widehat{x} \in AF_k. \tag{3}$$

As $\lim\limits_{n\to\infty} x_n = \widehat{x}$, there exists $\tilde{N} \in \mathbb{N}$ such that $d(x_n, \widehat{x}) \leq \frac{1}{N+1}$ for all $n \geq \tilde{N}$. As (x_n) has the liminf property w.r.t. F, we get that $x_K \in AF_M$ for some $K \geq \tilde{N}$. Applying (3) gives $\widehat{x} \in AF_k$. □

Lemma 3.6. *Suppose that X is compact, F is explicitly closed and that (x_n) has approximate F-points. Then the set $\{x_n \mid n \in \mathbb{N}\}$ has an adherent point $x \in F$.*

Proof. We have for each $k \in \mathbb{N}$ an $m_k \in \mathbb{N}$ such that x_{m_k} is a k-approximate F-point. Let $y_k := x_{m_k} \in AF_k$. Since X is compact, the sequence (y_k) has a convergent subsequence (y_{k_n}). Let $x := \lim_{n\to\infty} y_{k_n}$. By Lemma 3.2, (y_{k_n}) has the liminf property w.r.t. F. Apply now Lemma 3.5 to conclude that $x \in F$. ☐

4 Generalized Fejér monotone sequences

In this section we give a generalization of Fejér monotonicity, one of the most used methods for strong convergence proofs in convex optimization.
We consider functions $G : \mathbb{R}_+ \to \mathbb{R}_+$ with the property

$$(G) \quad \text{If } a_n \overset{n\to\infty}{\to} 0, \text{ then } G(a_n) \overset{n\to\infty}{\to} 0$$

for all sequences (a_n) in $\mathbb{R}_+$.
Obviously, (G) is equivalent to the fact that there exists a mapping $\alpha_G : \mathbb{N} \to \mathbb{N}$ satisfying

$$\forall k \in \mathbb{N}\, \forall a \in \mathbb{R}_+ \ \left(a \le \frac{1}{\alpha_G(k)+1} \to G(a) \le \frac{1}{k+1} \right). \tag{4}$$

We say that such a mapping α_G is a *G-modulus*.
Any continuous $G : \mathbb{R}_+ \to \mathbb{R}_+$ with $G(0) = 0$ satisfies (G) and any modulus of continuity of G at 0 is a G-modulus.
We also consider functions $H : \mathbb{R}_+ \to \mathbb{R}_+$ with the property

$$(H) \quad \text{If } H(a_n) \overset{n\to\infty}{\to} 0, \text{ then } a_n \overset{n\to\infty}{\to} 0$$

for all sequences (a_n) in $\mathbb{R}_+$, which is kind of the converse of (G).
Similarly, (H) is equivalent to the existence of an *H-modulus* $\beta_H : \mathbb{N} \to \mathbb{N}$ such that

$$\forall k \in \mathbb{N}\, \forall a \in \mathbb{R}_+ \ \left(H(a) \le \frac{1}{\beta_H(k)+1} \to a \le \frac{1}{k+1} \right). \tag{5}$$

Let (x_n) be a sequence in the metric space (X, d) and $\emptyset \ne F \subseteq X$.

Definition 4.1. *(x_n) is (G,H)-*Fejér monotone* w.r.t. F if for all $n, m \in \mathbb{N}$ and all $p \in F$,*

$$H(d(x_{n+m}, p)) \le G(d(x_n, p)).$$

Note that the usual notion of being 'Fejér monotone' is just $(id_{\mathbb{R}_+}, id_{\mathbb{R}_+})$-Fejér monotone.
The following lemma collects some useful properties of generalized Fejér monotone sequences.

Lemma 4.2. *Let (x_n) be (G,H)-Fejér monotone w.r.t. F.*

(i) If $\{x_n \mid n \in \mathbb{N}\}$ has an adherent point $\hat{x} \in F$, then (x_n) converges to $\hat{x}$.

(ii) Assume that H has the property

$$(H1) \quad \textit{If } H(a_n) \textit{ is bounded, then } (a_n) \textit{ is bounded}$$

for all sequences (a_n) in $\mathbb{R}_+$. Then (x_n) is bounded.

Proof. (i) Let $p \in \mathbb{N}$ be arbitrary and $K := K_p$ be so that

$$d(x_K, \hat{x}) \leq \frac{1}{\alpha_G(\beta_H(p)+1)+1}$$

(such a K has to exist by the assumption). Applying the fact that (x_n) is (G,H)-Fejér monotone w.r.t. F and (4), we get that for all $l \in \mathbb{N}$,

$$H(d(x_{K+l}, \hat{x})) \leq G(d(x_K, \hat{x})) \leq \frac{1}{\beta_H(p)+1}.$$

Using now (5), it follows that $d(x_{K+l}, \hat{x}) \leq \dfrac{1}{p+1}$ for all $l \in \mathbb{N}$. Hence (x_n) converges to $\hat{x}$.

(ii) Since (x_n) is (G,H)-Fejér monotone w.r.t. F we have for $p \in F$ that $H(d(x_n,p)) \leq G(d(x_0,p))$ for all $n \in \mathbb{N}$. Hence, $(H(d(x_n,p)))$ is bounded and so, by $(H1)$, $(d(x_n,p))$ is bounded.

□

As an immediate consequence of Lemma 3.6 and Lemma 4.2.(i), we get

Proposition 4.3. *Let X be a compact metric space and F be explicitly closed. Assume that (x_n) is (G,H)-Fejér monotone with respect to F and that (x_n) has approximate F-points. Then (x_n) converges to a point $x \in F$.*

Remark 4.4. *If in Proposition 4.3 we either weaken 'compact' to 'totally bounded' or drop the assumption on F being explicitly closed, then the conclusion in general becomes false, in fact (x_n) might not even be Cauchy (see Example 7.4).*

Let us recall that a metric space is said to be *boundedly compact* if every bounded sequence has a convergent subsequence. One can easily see that X is boundedly compact if and only if for every $a \in X$ and $r > 0$ the closed ball $\overline{B}(a,r)$ is compact.

Remark 4.5. *The proof of Proposition 4.3 uses the compactness property only for the sequence (x_n) and so it is enough to require that X is boundedly compact and that the sequence at hand is bounded. As we prove above, this is the case if H has the property $(H1)$ for all sequences (a_n) in $\mathbb{R}_+$.*

4.1 Uniform (G,H)-Fejér monotone sequences

Being (G,H)-Fejér monotone w.r.t. F can be logically re-written as

$$\forall n,m \in \mathbb{N}\, \forall p \in X \Bigg(\quad \forall k \in \mathbb{N}(p \in AF_k) \to \\ \forall r \in \mathbb{N}\, \forall l \leq m \left(H(d(x_{n+l},p)) < G(d(x_n,p)) + \frac{1}{r+1} \right) \Bigg),$$

hence as

$$\forall r,n,m\,\forall p\,\exists k\,\left(p\in AF_k\to\forall l\le m\left(H(d(x_{n+l},p))<G(d(x_n,p))+\frac{1}{r+1}\right)\right).$$

If $p\in AF_k$ can be written as a purely universal formula (when formalized in the language of the systems used in the logical metatheorems from [25, 15, 26]), then

$$p\in AF_k\to\forall l\le m\left(H(d(x_{n+l},p))<G(d(x_n,p))+\frac{1}{r+1}\right)$$

is (equivalent to) a purely existential formula. Hence one can use these metatheorems to extract a uniform bound on '$\exists k\in\mathbb{N}$' (and so in fact a uniform realizer as the formula is monotone in k) which - e.g. for bounded (X,d) - only depends on a bound on the metric and majorizing data of the other parameters involved but not on 'p'. This motivates the next definition:

Definition 4.6. *We say that (x_n) is* uniformly (G,H)-Fejér monotone *w.r.t. F if for all $r,n,m\in\mathbb{N}$,*

$$\exists k\in\mathbb{N}\,\forall p\in X\,\left(p\in AF_k\to\forall l\le m\left(H(d(x_{n+l},p))<G(d(x_n,p))+\frac{1}{r+1}\right)\right).$$

Any upper bound (and hence realizer) $\chi(n,m,r)$ of '$\exists k\in\mathbb{N}$' is called a modulus of (x_n) being (uniformly) (G,H)-Fejér monotone w.r.t. F.

If $G=H=id_{\mathbb{R}_+}$, we say simply that (x_n) is uniformly Fejér monotone w.r.t. F.

Remark 4.7. *(i) A standard compactness argument shows that for X compact, F explicitly closed and G,H continuous the notions '(G,H)-Fejér monotone w.r.t. F' and 'uniformly (G,H)-Fejér monotone w.r.t. F' are equivalent.*

(ii) In Corollary 5.2 we will see, as a consequence of our quantitative metastable analysis of the proof of Proposition 4.3, that the Cauchyness of (x_n) holds even if we replace 'compact' by 'totally bounded' and drop the explicit closedness of F **provided** *that we replace '(G,H)-Fejér monotone' by 'uniform (G,H)-Fejér monotone'.*

(iii) The equivalence between these notions can be proven (relative to the framework of $\mathcal{T}^\omega[X,d]$) for general bounded metric spaces X and F,G,H from the 'nonstandard' uniform boundedness principle $\exists$-UB^X studied in [26]. Though being false for specific spaces X, the use of $\exists$-UB^X in proofs of statements of the form considered in our general bound-extraction theorems is allowed and the bounds extracted from proofs in $\mathcal{T}^\omega[X,d]+\exists$-UB^X will be correct in any bounded metric space X (see [26, Theorem 17.101]).

5 Main quantitative results

In this section, (X,d) is a totally bounded metric space with a II-modulus of total boundedness γ and $\emptyset\neq F\subseteq X$. Furthermore, $G,H:\mathbb{R}_+\to\mathbb{R}_+$ satisfy $(G),(H)$ for all sequences (a_n) in $\mathbb{R}_+$, α_G is a G-modulus and β_H is an H-modulus.

Assume that (x_n) has approximate F-points. We can then define the mapping

$$\varphi_F : \mathbb{N} \to \mathbb{N}, \quad \varphi_F(k) = \min\{m \in \mathbb{N} \mid x_m \in AF_k\}. \tag{6}$$

Thus, $x_{\varphi_F(k)} \in AF_k$ for all k and φ_F is monotone nondecreasing. An *approximate F-point bound* for (x_n) is any function $\Phi : \mathbb{N} \to \mathbb{N}$ satisfying

$$\forall k \in \mathbb{N}\, \exists N \leq \Phi(k)\ (x_N \in AF_k). \tag{7}$$

If Φ is an approximate F-point bound for (x_n), then

$$\Phi^M : \mathbb{N} \to \mathbb{N}, \quad \Phi^M(k) = \max\{\Phi(m) \mid m \leq k\}$$

is monotone nondecreasing and again an approximate F-point bound for (x_n).
Thus, we shall assume w.l.o.g. that any approximate F-point bound for (x_n) is monotone nondecreasing.
Then $\Phi : \mathbb{N} \to \mathbb{N}$ is an approximate F-point bound for (x_n) if and only if Φ majorizes φ_F.

The next theorem is the main step towards a quantitative version of Proposition 4.3 (see also the discussion in [26, pp. 464-465] on the logical background behind the elimination of sequential compactness in the original proof in favor of a computational argument):

Theorem 5.1. *Assume that*

(i) (x_n) is uniformly (G, H)-Fejér monotone w.r.t. F, with modulus χ;

(ii) (x_n) has approximate F-points, with Φ being an approximate F-point bound.

Then (x_n) is Cauchy and, moreover, for all $k \in \mathbb{N}$ and all $g : \mathbb{N} \to \mathbb{N}$,

$$\exists N \leq \Psi(k, g, \Phi, \chi, \alpha_G, \beta_H, \gamma)\, \forall i, j \in [N, N + g(N)]\ \left(d(x_i, x_j) \leq \frac{1}{k+1}\right),$$

where $\Psi(k, g, \Phi, \chi, \alpha_G, \beta_H, \gamma) := \Psi_0(P, k, g, \Phi, \chi, \beta_H)$, with

$$\chi_g(n, k) := \chi(n, g(n), k), \quad \chi_g^M(n, k) := \max\{\chi_g(i, k) \mid i \leq n\},$$

$P := \gamma\left(\alpha_G\left(2\beta_H(2k+1)+1\right)\right)$ and

$$\begin{cases} \Psi_0(0, k, g, \Phi, \chi, \beta_H) := 0 \\ \Psi_0(n+1, k, g, \Phi, \chi, \beta_H) := \Phi\left(\chi_g^M\left(\Psi_0(n, k, g, \Phi, \chi, \beta_H), 2\beta_H(2k+1)+1\right)\right). \end{cases}$$

Proof. Let $k \in \mathbb{N}$ and $g : \mathbb{N} \to \mathbb{N}$. For simplicity, let us denote with φ the mapping φ_F defined by (6). Since both φ and Φ are nondecreasing and Φ majorizes φ, an immediate induction gives us that $\Psi_0(n, k, g, \varphi, \chi, \beta_H) \leq \Psi_0(n+1, k, g, \varphi, \chi, \beta_H)$, $\Psi_0(n, k, g, \Phi, \chi, \beta_H) \leq \Psi_0(n+1, k, g, \Phi, \chi, \beta_H)$ and $\Psi_0(n, k, g, \varphi, \chi, \beta_H) \leq \Psi_0(n, k, g, \Phi, \chi, \beta_H)$ for all $n \in \mathbb{N}$.
Define for every $i \in \mathbb{N}$

$$n_i := \Psi_0(i, k, g, \varphi, \chi, \beta_H). \tag{8}$$

Claim 1: For all $j \geq 1$ and all $0 \leq i < j$, x_{n_j} is a $\chi_g(n_i, 2\beta_H(2k+1)+1)$-approximate F-point.

Proof of claim: As $j \geq 1$ and

$$\begin{aligned} n_j &= \Psi_0(j,k,g,\varphi,\chi,\beta_H) = \varphi\left(\chi_g^M\left(\Psi_0(j-1,k,g,\varphi,\chi,\beta_H), 2\beta_H(2k+1)+1\right)\right) \\ &= \varphi\left(\chi_g^M\left(n_{j-1}, 2\beta_H(2k+1)+1\right)\right), \end{aligned}$$

x_{n_j} is a $\chi_g^M(n_{j-1}, 2\beta_H(2k+1)+1)$-approximate F-point. Since $0 \leq i \leq j-1$, we have that $n_i \leq n_{j-1}$. Apply now the fact that χ_g^M is nondecreasing in the first argument to get that

$$\begin{aligned} \chi_g(n_i, 2\beta_H(2k+1)+1) &\leq \chi_g^M(n_i, 2\beta_H(2k+1)+1) \\ &\leq \chi_g^M(n_{j-1}, 2\beta_H(2k+1)+1). \qquad \blacksquare \end{aligned}$$

Claim 2: There exist $0 \leq I < J \leq P$ satisfying

$$\forall l \in [n_I, n_I + g(n_I)] \left(d(x_l, x_{n_J}) \leq \frac{1}{2k+2} \right).$$

Proof of claim: By the property of γ being a II-modulus of total boundedness for X we get that there exist $0 \leq I < J \leq P$ such that

$$d(x_{n_I}, x_{n_J}) \leq \frac{1}{\alpha_G(2\beta_H(2k+1)+1)+1}$$

and so, using that α_G is a G-modulus,

$$G(d(x_{n_I}, x_{n_J})) \leq \frac{1}{2\beta_H(2k+1)+2}. \tag{9}$$

By the first claim, we have that x_{n_J} is a $\chi_g(n_I, 2\beta_H(2k+1)+1)$-approximate F-point. Applying now the uniform (G,H)-Féjer monotonicity of (x_n) w.r.t. F with $r := 2\beta_H(2k+1)+1, n := n_I, m := g(n_I)$ and $p := x_{n_J}$, we get that for all $l \leq g(n_I)$,

$$H(d(x_{n_I+l}, x_{n_J})) \quad \leq \quad G(d(x_{n_I}, x_{n_J})) + \frac{1}{2\beta_H(2k+1)+2} \leq \frac{1}{\beta_H(2k+1)+1}.$$

Since β_H is an H-modulus,

$$\forall l \leq g(n_I) \ \left(d(x_{n_I+l}, x_{n_J}) \leq \frac{1}{2k+2} \right).$$

and so the claim is proved. $\blacksquare$

It follows that

$$\forall k, l \in [n_I, n_I + g(n_I)] \ \left(d(x_k, x_l) \leq \frac{1}{k+1} \right).$$

Since $n_I = \Psi_0(I,k,g,\varphi,\chi,\beta_H) \leq \Psi_0(I,k,g,\Phi,\chi,\beta_H)$ and $I \leq P$, we get that

$$n_I \leq \Psi_0(P,k,g,\Phi,\chi,\beta_H) = \Psi(k,g,\Phi,\chi,\alpha_G,\beta_H,\gamma).$$

The theorem holds with $N := n_I$. $\square$

Corollary to the proof: One of the numbers $n_0, \ldots, n_{P-1}$ is a point of metastability.

Theorem 5.1 remarkably implies the Cauchy property of (x_n) in the absence of X being complete (and hence compact) and of F being explicitly closed which, as we remarked after Proposition 4.3, both were necessary if (x_n) only was assumed to be (G,H)-Fejér monotone rather than being uniformly (G,H)-Fejér monotone. This is a **qualitative** improvement of Proposition 4.3 whose proof is based on our quantitative analysis of metastability although the result as such does not involve metastability at all:

Corollary 5.2. *Let X be totally bounded and (x_n) be uniformly (G,H)-Fejér monotone having approximate F-points. Then (x_n) is Cauchy.*

The next theorem is a direct quantitative 'finitization' of Proposition 4.3 in the sense of Tao:

Theorem 5.3. *In addition to the assumptions of Theorem 5.1 we suppose that F is uniformly closed with moduli δ_F, ω_F. Then for all $k \in \mathbb{N}$ and all $g : \mathbb{N} \to \mathbb{N}$,*

$$\exists N \le \tilde{\Psi}\, \forall i,j \in [N, N+g(N)] \ \left(d(x_i, x_j) \le \frac{1}{k+1} \text{ and } x_i \in AF_k \right),$$

where

$$\tilde{\Psi}\,(k, g, \Phi, \chi, \alpha_G, \beta_H, \gamma, \delta_F, \omega_F) := \Psi(k_0, g, \Phi, \chi_{k,\delta_F}, \alpha_G, \beta_H, \gamma),$$

with Ψ defined as in Theorem 5.1,

$$k_0 = \max\left\{ k, \left\lceil \frac{\omega_F(k) - 1}{2} \right\rceil \right\} \text{ and } \chi_{k,\delta_F}(n,m,r) := \max\{\delta_F(k), \chi(n,m,r)\}.$$

Proof. With χ also χ_{k,δ_F} is a modulus of (x_n) being uniformly (G,H)-Fejér monotone w.r.t. F. Applying Theorem 5.1 to (k_0, χ_{k,δ_F}) we get that

$$\exists N \le \tilde{\Psi}\, \forall i,j \in [N, N+g(N)] \ \left(d(x_i, x_j) \le \frac{1}{k_0+1} \le \frac{1}{k+1} \right).$$

From the proof of Theorem 5.1, it follows that there exists $0 \le I < J \le P$ such that $N = n_I$ and x_{n_J} is a $(\chi_{k,\delta_F})_g(N, 2\beta_H(2k_0+1)+1)$-approximate F-point and

$$\forall i \in [N, N+g(N)] \ \left(d(x_i, x_{n_J}) \le \frac{1}{2k_0+2} \le \frac{1}{\omega_F(k)+1} \right).$$

Since $(\chi_{k,\delta_F})_g(N, 2\beta_H(2k_0+1)+1) = \chi_{k,\delta_F}(N, g(N), 2\beta_H(2k_0+1)+1) \ge \delta_F(k)$, it follows that x_{n_I} is a $\delta_F(k)$-approximate F-point. Hence by the definition of ω_F, we get that $x_i \in AF_k$ for all $i \in [N, N+g(N)]$. □

Notation: In our applications δ_F will be mostly $\delta_F(k) = 2k+1$. In this case we simply write χ_k instead of χ_{k,δ_F} when applying Theorem 5.3.

Remark 5.4. *Theorems 5.1 and 5.3 hold for X boundedly compact and (x_n) bounded. In this case, the bounds will depend on a II-modulus of total boundedness for the closed ball $\overline{B}(a,b)$, where $a \in X$ and $b \ge d(x_n, a)$ for all n.*

Remark 5.5. *Theorem 5.3 is a finitization of Proposition 4.3 in the sense of Tao since it only talks about a finite initial segment of* (x_n) *but trivially implies back the infinitary Proposition 4.3 for uniformly closed* F *and uniformly* (G,H)*-Fejér monotone sequences.*

Proof. Noneffectively

$$\forall k \in \mathbb{N}\, \forall g : \mathbb{N} \to \mathbb{N}\, \exists N \in \mathbb{N}\, \forall i,j \in [N, N+g(N)] \left(d(x_i, x_j) \leq \frac{1}{k+1} \text{ and } x_i \in AF_k \right)$$

implies the Cauchy property of (x_n). Since X is complete, (x_n) converges to a point $\widehat{x} \in X$. It remains to prove that $\widehat{x} \in F$. One can easily see, by taking g to be a constant function, that (x_n) has the liminf property w.r.t. F. Apply now Lemma 3.5 to conclude that $\widehat{x} \in F$. □

Assume that (x_n) is asymptotically regular w.r.t. F. A mapping $\Phi^+ : \mathbb{N} \times \mathbb{N}^{\mathbb{N}} \to \mathbb{N}$ satisfying

$$\forall k \in \mathbb{N}\, \forall g : \mathbb{N} \to \mathbb{N}\, \exists N \leq \Phi^+(k,g)\, \forall m \in [N, N+g(N)]\ (x_m \in AF_k)$$

is said to be a *rate of metastability* for the asymptotic regularity of (x_n) w.r.t. F. If Φ^+ is such a rate, then $\Phi(k) := \Phi^+(k, 0)$ is an approximate F-point bound.
A *rate* of asymptotic regularity of (x_n) w.r.t. F is a function $\Phi^{++} : \mathbb{N} \to \mathbb{N}$ with

$$\forall k \in \mathbb{N}\, \exists N \leq \Phi^{++}(k)\, \forall m \geq N\ (x_m \in AF_k)$$

which is equivalent with the fact that Φ^{++} satisfies

$$\forall k \in \mathbb{N}\, \forall n \geq \Phi^{++}(k)\ (x_n \in AF_k).$$

Obviously, if Φ^{++} is a rate of asymptotic regularity, then $\Phi^+(k,g) := \Phi^{++}(k)$ is a rate of metastability for the asymptotic regularity.
Having a rate of metastability for the asymptotic regularity one can obtain a theorem with the conclusion as in Theorem 5.3 even without the assumption of uniform closedness (not even explicit closedness is needed). This follows from a construction Ω known in logic that allows one to combine two quantitative metastability statements with resp. rates Φ^+ and Ψ into one new rate of metastability $\Omega(\Phi^+, \Psi)$ which gives a bound for an interval of metastability where both statements hold simultaneously. However, as the details of Ω are somewhat complicated to state without using notation from logic we will skip this here.

We conclude this section with a trivial but instructive example for Theorem 5.1, namely that the well-known rate of metastability for the Cauchy property of monotone bounded sequences from [26, Proposition 2.27] can be recovered (modulo a constant) from this theorem: let $X = [0,1]$ and (x_n) be a nondecreasing sequence in X. Let us take

$$F = \bigcap_{k \in \mathbb{N}} \tilde{F}_k, \quad \text{where } \tilde{F}_k := \{p \in X \mid x_k \leq p\}.$$

Then, clearly, (i) $AF_k = \tilde{F}_k$, (ii) $\Phi^{++} := id$ is a rate of asymptotic regularity and $\chi(n,m,r) := n+m$ is a modulus of the uniform Fejér monotonicity of (x_n) and we may take $\gamma(k) = \lceil C \rceil (k+1)$ (see Example 2.7). For monotone g, Theorem 5.1 now gives $\Psi(k,g) := \tilde{g}^{4\lceil C \rceil (k+1)}(0)$ with $\tilde{g}(n) := n + g(n)$, while the direct proof in this case yields the optimal rate $\Psi(k,g) := \tilde{g}^{\lceil C \rceil (k+1)}(0)$.

6 Quasi-Fejér monotone sequences

As a common consequence of arriving at a finitary quantitative version of an originally non-quantitative theorem, one can easily incorporate error terms as has been considered under the name of quasi-Fejér monotonicity (due to [12]). As pointed out in [11], quasi-Fejér monotone sequences provide a framework for the analysis of numerous optimization algorithms in Hilbert spaces.

Definition 6.1. *A sequence (x_n) in a metric space (X, d) is called quasi-Fejér monotone (of order $0 < P < \infty$) w.r.t. some set $\emptyset \neq F \subseteq X$ if*

$$\forall n \in \mathbb{N}\, \forall p \in F\, \left(d(x_{n+1}, p)^P \leq d(x_n, p)^P + \varepsilon_n\right),$$

where (ε_n) is some summable sequence in $\mathbb{R}_+$.

The appropriate generalization to general functions (G, H) then is:

Definition 6.2. *For G, H as in the definition of (G, H)-Fejér monotonicity we say that (x_n) is quasi-(G, H)-Fejér monotone w.r.t. F if*

$$\forall n, m \in \mathbb{N}\, \forall p \in F\, \left(H(d(x_{n+m}, p)) \leq G(d(x_n, p)) + \sum_{i=n}^{n+m-1} \varepsilon_i\right).$$

Note that for $G(x) := H(x) := x^P$ this covers the notion of quasi-Fejér monotonicity.
The uniform version of this notion then is:

Definition 6.3. *(x_n) is uniformly quasi-(G, H)-Fejér monotone w.r.t. F and a given representation of F via AF_k as before if*

$$\begin{aligned}&\forall r, n, m \in \mathbb{N}\, \exists k \in \mathbb{N}\, \forall p \in X\, \big(p \in AF_k \to \\ &\qquad \forall l \leq m (H(d(x_{n+l}, p)) < G(d(x_n, p)) + \textstyle\sum_{i=n}^{n+l-1} \varepsilon_i + \frac{1}{r+1})\big).\end{aligned}$$

Any function $\chi : \mathbb{N}^3 \to \mathbb{N}$ such that $\chi(r, n, m)$ provides such a k is called a modulus of (x_n) being uniformly quasi-(G, H)-Fejér monotone w.r.t. F.

Let $\xi : \mathbb{N} \to \mathbb{N}$ be a Cauchy modulus of $\sum \varepsilon_i$, i.e. $\sum\limits_{i=\xi(n)}^{\infty} \varepsilon_i < \frac{1}{n+1}$ for all $n \in \mathbb{N}$.
If (x_n) has the lim inf-property w.r.t. F we can define

$$\widehat{\varphi}_F(k, n) := \min\{m \in \mathbb{N} \mid m \geq n \wedge x_m \in AF_k\}.$$

Any monotone (in k, n) upper bound $\widehat{\Phi}$ of $\widehat{\varphi}_F$ is called a lim inf-bound w.r.t. F.

Theorem 6.4. *Assume that X is totally bounded with a II-modulus of total boundedness γ and that*

(i) (x_n) is uniformly quasi-(G, H)-Fejér monotone w.r.t. F, with modulus χ, and (ε_n) with Cauchy rate ξ for $\sum \varepsilon_i$;

(ii) (x_n) has the lim inf-property w.r.t. F, with $\widehat{\Phi}$ being a lim inf-bound w.r.t. F.

Then (x_n) *is Cauchy and, moreover, for all* $k \in \mathbb{N}$ *and all* $g : \mathbb{N} \to \mathbb{N}$,

$$\exists N \leq \widehat{\Psi}(k, g, \widehat{\Phi}, \chi, \alpha_G, \beta_H, \gamma, \xi)\, \forall i, j \in [N, N + g(N)] \ \left(d(x_i, x_j) \leq \frac{1}{k+1}\right),$$

where $\widehat{\Psi}(k, g, \widehat{\Phi}, \chi, \alpha_G, \beta_H, \gamma, \xi) := \widehat{\Psi}_0(P, k, g, \widehat{\Phi}, \chi, \beta_H, \xi)$, *with*

$$\chi_g(n, k) := \chi(n, g(n), k), \quad \chi_g^M(n, k) := \max\{\chi_g(i, k) \mid i \leq n\},$$

$P := \gamma\left(\alpha_G\left(4\beta_H(2k+1)+3\right)\right) + 1$ *and*

$$\begin{cases} \widehat{\Psi}_0(0, k, g, \widehat{\Phi}, \chi, \beta_H, \xi) := 0 \\ \widehat{\Psi}_0(n+1, k, g, \widehat{\Phi}, \chi, \beta_H, \xi) := \\ \qquad \widehat{\Phi}\left(\chi_g^M\left(\widehat{\Psi}_0(n, k, g, \widehat{\Phi}, \chi, \beta_H, \xi), 4\beta_H(2k+1)+3\right), \xi(4\beta_H(2k+1)+3)\right). \end{cases}$$

Proof. The proof is the same as the one of Theorem 5.1 up to (9) which now holds with $1/(4\beta_H(2k+1)+4)$ instead of $1/(2\beta_H(2k+1)+2)$. We then use uniform quasi-(G, H)-Fejér monotonicity as we did before without 'quasi' to now get that for all $l \leq g(n_I)$

$$H(d(x_{n_I+l}, x_{n_J})) \leq G(d(x_{n_I}, x_{n_J})) + \sum_{i=n_I}^{n_I+l-1} \varepsilon_i + \frac{1}{4\beta_H(2k+1)+4}.$$

By construction of n_I we know that $n_I \geq \xi(4\beta_H(2k+1)+3)$ (note that by the addition of '+1' to the original definition of P that was used in the proof of Theorem 5.1 I, J can now be choosen so that $0 < I < J \leq P$ rather than only $0 \leq I < J \leq P$) and so

$$\sum_{i=n_I}^{n_I+l-1} \varepsilon_i \leq \frac{1}{4\beta_H(2k+1)+4}$$

and so we get in total

$$H(d(x_{n_I+l}, x_{n_J})) \leq G(d(x_{n_I}, x_{n_J})) + \frac{1}{2\beta_H(2k+1)+2}$$

from where we can finish the proof as before. □

As it is clear from the proof above, one actually does not need a Cauchy modulus ξ of the error-sum but only a rate of metastability.

With the new bound $\widehat{\Psi}$ from Theorem 6.4 instead of Ψ all the other results of the previous section extend in the obvious way to the 'quasi'-case. As a consequence of this, we could incorporate also in the iterations considered in the rest of this paper error terms which we, however, will not carry out.

7 Application - F is $Fix(T)$

Let X be a metric space, $C \subseteq X$ a nonempty subset and $T : C \to C$ be a mapping. We assume that T has fixed points and define F as the nonempty fixed point set $Fix(T)$ of T.

One has

$$F = \bigcap_{k\in\mathbb{N}} \tilde{F}_k, \quad \text{where } \tilde{F}_k = \left\{ x \in C \mid d(x,Tx) \le \frac{1}{k+1} \right\}.$$

In this case, for all $k \in \mathbb{N}$ we have that that $AF_k = \tilde{F}_k$ and the k-approximate F-points are precisely the $1/(k+1)$-approximate fixed points of T.
Let us recall that the mapping T is uniformly continuous with modulus $\omega_T : \mathbb{N} \to \mathbb{N}$ if for all $k \in \mathbb{N}$ and all $p, q \in C$,

$$d(p,q) \le \frac{1}{\omega_T(k)+1} \ \rightarrow \ d(Tp,Tq) \le \frac{1}{k+1}.$$

One can see easily that the following properties hold.

Lemma 7.1. *Let (x_n) be a sequence in C.*

(i) (x_n) has approximate F-points if and only if for all $k \in \mathbb{N}$ there exists $N \in \mathbb{N}$ such that $d(x_N, Tx_N) \le \dfrac{1}{k+1}$. If this is the case, we say also that (x_n) has approximate fixed points.

(ii) (x_n) has the liminf property w.r.t. F if and only if $\liminf\limits_{n\to\infty} d(x_n, Tx_n) = 0$.

(iii) (x_n) is asymptotically regular w.r.t. F if and only if $\lim\limits_{n\to\infty} d(x_n, Tx_n) = 0$.

(iv) If T is continuous, then F is explicitly closed.

(v) If T is uniformly continuous with modulus ω_T, then F is uniformly closed with moduli $\omega_F(k) = \max\{4k+3, \omega_T(4k+3)\}$ and $\delta_F(k) = 2k+1$.

Remark 7.2. *Uniform closedness can be viewed as a quantitative version of the special extensionality statement $(*)$ $q \in F \wedge p =_X q \to p \in F$. Extensionality w.r.t. $p =_X q := \|p-q\|_X = 0$ is not included as an axiom in our formal framework (for reasons explained in [25]) and has to be derived (if needed) from appropriate uniform continuity assumptions (see item (v) in the lemma above). In the case of $(*)$, however, it suffices to have the moduli ω_F, δ_T which (as we will see in Section 7.4) are also available for interesting classes of in general discontinuous mappings T (where, in particular, the model theoretic approach to metastability from [2] is not applicable as it stands).*

As a consequence of Proposition 4.3 and Remark 4.5, we get

Proposition 7.3. *Let C be a boundedly compact subset of a metric space X and $T : C \to C$ be continuous with $F = Fix(T) \neq \emptyset$. Assume that (x_n) is bounded and (G,H)-Fejér monotone with respect to F and that (x_n) has approximate fixed points. Then (x_n) converges to a fixed point of T. The continuity of T can be replaced by the weaker assumption that F is explicitly closed (see Section 7.4 for a class of in general discontinuous functions for which $Fix(T)$ is uniformly closed).*

If we weaken 'boundedly compact' to 'totally bounded' or drop the assumption that T is continuous, one cannot even prove that (x_n) is Cauchy, as the following examples show.

Example 7.4. *Let $C := (0,1] \cup \{2\}$ with the metric $d(x,y) := \min\{|x-y|, 1\}$. Then C is totally bounded and the mapping*

$$T : C \to C, \ T(x) := x/2, \ \textit{if } x \in (0,1], \ T(2) := 2$$

is continuous with $F := Fix(T) = \{2\}$. Now let $x_n := T^n(1)$, for even n, and $x_n := 1$ for odd n. Then (x_n) has approximate fixed points and is Fejér monotone w.r.t. F but clearly not Cauchy.
If we drop the explicit closedness of F, we can slightly modify the above example to get a counterexample to the Cauchyness of (x_n) even for compact C: just take $C := [0,1] \cup \{2\}$ and define $T(0) := 2$.

7.1 Picard iteration for (firmly) nonexpansive mappings

Assume that T is nonexpansive. Then, obviously, T is uniformly continuous with modulus $\omega_T = id_{\mathbb{N}}$. We consider in the sequel the Picard iteration starting from $x \in C$:

$$x_n := T^n x.$$

One can see by induction that for all $n, m \in \mathbb{N}$ and $p \in C$,

$$d(x_{n+m}, p) \leq d(x_n, p) + md(p, Tp).$$

As an immediate consequence, we get that (x_n) is Fejér monotone, hence, in particular, bounded. In fact, one can easily prove more:

Lemma 7.5. *(x_n) is uniformly Fejér monotone w.r.t. F with modulus*

$$\chi(n, m, r) = m(r + 1).$$

Applying Proposition 7.3, we get

Corollary 7.6. *Let C be a boundedly compact subset of a metric space X and $T : C \to C$ be nonexpansive with $Fix(T) \neq \emptyset$. Assume that (x_n) has approximate fixed points. Then (x_n) converges to a fixed point of T.*

As (x_n) is uniformly Fejér monotone w.r.t. F and F is uniformly closed, we can apply our quantitative Theorems 5.1 and 5.3 to get the following:

Theorem 7.7. *Assume that C is totally bounded with II-modulus of total boundedness γ, $T : C \to C$ is nonexpansive with $Fix(T) \neq \emptyset$ and that (x_n) has approximate fixed points, with Φ being an approximate fixed point bound. Then for all $k \in \mathbb{N}$ and all $g : \mathbb{N} \to \mathbb{N}$,*

(i) There exists $N \leq \Sigma(k, g, \Phi, \gamma)$ such that

$$\forall i, j \in [N, N + g(N)] \left(d(x_i, x_j) \leq \frac{1}{k+1} \right),$$

where $\Sigma(k, g, \Phi, \gamma) = \Sigma_0(\gamma(4k + 3), k, g, \Phi)$, with $\Sigma_0(0, k, g, \Phi) = 0$ and

$$\Sigma_0(n + 1, k, g, \Phi) = \Phi\left((4k + 4)g^M(\Sigma_0(n, k, g, \Phi))\right).$$

(ii) There exists $N \leq \tilde{\Sigma}(k, g, \Phi, \gamma)$ such that

$$\forall i, j \in [N, N + g(N)] \ \left(d(x_i, x_j) \leq \frac{1}{k+1} \text{ and } d(x_i, Tx_i) \leq \frac{1}{k+1} \right),$$

where $\tilde{\Sigma}(k, g, \Phi, \gamma) = \tilde{\Sigma}_0(\gamma(8k + 7), k, g, \Phi)$, with $\tilde{\Sigma}_0(0, k, g, \Phi) = 0$ and

$$\tilde{\Sigma}_0(n + 1, k, g, \Phi) = \Phi\left(\max\left\{2k + 1, (8k + 8)g^M(\tilde{\Sigma}_0(n, k, g, \Phi))\right\}\right).$$

Proof. (i) With Ψ, Ψ_0 as in Theorem 5.1, $\alpha_G = \beta_H = id_{\mathbb{N}}$ and χ as in Lemma 7.5, define $\Sigma(k,g,\Phi,\gamma) = \Psi(k, g^M, \Phi, \chi, \alpha_G, \beta_H, \gamma)$ and $\Sigma_0(l,k,g,\Phi) = \Psi_0(l, k_0, g^M, \Phi, \chi, \beta_H)$.

(ii) Apply Theorem 5.3 for g^M, using that $\omega_F(k) = 4k+3$ and $\delta_F(k) = 2k+1$, by Lemma 7.1.(v). It follows that $k_0 = 2k+1$ and $(\chi_k)^M_{g^M}(n,r) = \max\{2k+1, g^M(n)(r+1)\}$.

□

We recall that a W-hyperbolic space [25] is a metric space X endowed with a convexity mapping $W : X \times X \times [0,1] \to X$ satisfying

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

for all $x,y,z,w \in X$ and all $\lambda, \tilde{\lambda} \in [0,1]$. We use in the sequel the notation $(1-\lambda)x + \lambda y$ for $W(x,y,\lambda)$.
Following [17], one can define in the setting of W-hyperbolic spaces a notion of uniform convexity. A W-hyperbolic space X is *uniformly convex with modulus* $\eta : (0,\infty) \times (0,2] \to (0,1]$ if for any $r < 0, \varepsilon \in (0,2]$ and all $a,x,y \in X$,

$$d(x,a) \leq r,\ d(y,a) \leq r,\ \text{and } d(x,y) \geq \varepsilon r \quad \text{imply} \quad d\left(\frac{1}{2}x + \frac{1}{2}y, a\right) \leq (1-\eta(r,\varepsilon))r.$$

A modulus η is said to be *monotone* if it is nonincreasing in the first argument. Uniformly convex W-hyperbolic spaces with a monotone modulus η are called UCW-hyperbolic spaces in [39]. One can easily see that CAT(0) spaces [5] are UCW-hyperbolic spaces with modulus $\varepsilon^2/8$. We refer to [39, 38, 31] for properties of UCW-hyperbolic spaces.

A very important class of nonexpansive mappings are the firmly nonexpansive ones. They are central in convex optimization because of the correspondence with maximal monotone operators due to Minty [45]. We refer to [4] for a systematic analysis of this correspondence. Firmly nonexpansive mappings were introduced by Browder [6] in Hilbert spaces and by Bruck [9] in Banach spaces, but they are also studied in the Hilbert ball [17, 34] or in different classes of geodesic spaces [48, 49, 1, 47]. Let $C \subseteq X$ be a nonempty subset of a W-hyperbolic space X. A mapping $T : C \to C$ is λ-*firmly nonexpansive* (where $\lambda \in (0,1)$) if for all $x,y \in C$,

$$d(Tx,Ty) \leq d((1-\lambda)x + \lambda Tx, (1-\lambda)y + \lambda Ty) \leq d(x,y).$$

Using proof mining methods, effective uniform rates of asymptotic regularity for the Picard iteration were obtained for UCW-hyperbolic spaces in [1] and for W-hyperbolic spaces in [47]. For a CAT(0) space X, $C \subseteq X$ a bounded subset, one gets, as an immediate consequence of [1, Theorem 7.1][4] the following rate of asymptotic regularity for the Picard iteration of a λ-firmly nonexpansive mapping $T : C \to C$:

$$\Phi^{++}(k,b,\lambda) := \left\lceil \frac{8(b+1)^2}{\lambda(1-\lambda)} \right\rceil (k+1)^2, \tag{10}$$

where $b > 0$ is an upper bound on the diameter of C.
We can, thus, apply Theorem 7.7 with $\Phi := \Phi^{++}$.

[4] Correction to [1]: In Corollary 7.4 should be '$(b+1)^2$' instead of '$(b+1)$' in the definition of $\Phi(\varepsilon,\lambda,b)$.

7.2 Ishikawa iteration for nonexpansive mappings

Assume that X is a W-hyperbolic space, $C \subseteq X$ is convex and $T : C \to C$ is nonexpansive. The *Ishikawa iteration* starting with $x \in C$ is defined as follows:

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

where $(\lambda_n), (s_n)$ are sequences in $[0,1]$. This iteration was introduced in [21] in the setting of Hilbert spaces and it is a generalization of the well-known Mann iteration [42, 18], which can be obtained as a special case of (11) by taking $s_n = 0$ for all $n \in \mathbb{N}$.

Lemma 7.8. *(i) For all $n, m \in \mathbb{N}$ and all $p \in C$,*

$$d(x_{n+1}, p) \leq d(x_n, p) + 2\lambda_n d(p, Tp) \tag{12}$$

$$d(x_{n+m}, p) \leq d(x_n, p) + 2md(p, Tp). \tag{13}$$

(ii) (x_n) is uniformly Fejér monotone w.r.t. F with modulus

$$\chi(n, m, r) = 2m(r+1).$$

Proof. (i) (12) is proved in [40, Lemma 4.3, (11)]. We get (13) by an easy induction.

(ii) follows easily from (13).

□

As in the case of the Picard iteration of a nonexpansive mapping, we can apply Proposition 7.3 to get that for boundedly compact C and $T : C \to C$ nonexpansive with $Fix(T) \neq \emptyset$, the fact that the Ishikawa iteration (x_n) has approximate fixed points implies the convergence of (x_n) to a fixed point of T. Explicit approximate fixed point bounds and rates of asymptotic regularity w.r.t. F are computed in [39] for closed convex subsets C of UCW-hyperbolic spaces X.
We shall consider in the following only the setting of CAT(0) spaces. We assume that

(i) $\sum_{n=0}^{\infty} \lambda_n(1-\lambda_n)$ is divergent with $\theta : \mathbb{N} \to \mathbb{N}$ being a nondecreasing rate of divergence, i.e. satisfying $\sum_{k=0}^{\theta(n)} \lambda_k(1-\lambda_k) \geq n$ for all n.

(ii) $\limsup_n s_n < 1$ and $L, N_0 \in \mathbb{N}$ are such that $s_n \leq 1 - \frac{1}{L}$ for all $n \geq N_0$.

Then, as a consequence of [39, Corollary 4.6] and the proof of [39, Remark 4.8], we get the following approximate fixed point bound for (x_n).

Proposition 7.9. *Let X be a $CAT(0)$ space, $C \subseteq X$ a bounded convex closed subset with diameter d_C and $T : C \to C$ nonexpansive. Then*

$$\forall k \in \mathbb{N}\, \exists N \leq \Phi(k, b, \theta, L, N_0) \left(d(x_N, Tx_N) \leq \frac{1}{k+1} \right), \tag{14}$$

where $\Phi(k, b, \theta, L, N_0) = \theta\left(4(k+1)^2 L^2 \lceil b(b+1) \rceil + N_0\right)$, with $b > 0$ being an upper bound on the diameter of C.

For the particular case $\lambda_n = \lambda$, one can take $\theta(n) = n\left\lceil \frac{1}{\lambda(1-\lambda)} \right\rceil$, hence the approximate fixed point bound Φ becomes

$$\Phi(k, b, L, N_0) = \left\lceil \frac{1}{\lambda(1-\lambda)} \right\rceil \left(4(k+1)^2 L^2 \lceil b(b+1) \rceil + N_0\right).$$

Finally, we can apply Theorem 5.3 to get for C totally bounded with II-modulus of total boundedness γ a result similar with Theorem 7.7.(ii), providing us a functional $\tilde{\Sigma} := \tilde{\Sigma}(k, g, \Phi, \gamma)$ with the property that for all $k \in \mathbb{N}$ and all $g : \mathbb{N} \to \mathbb{N}$ there exists $N \leq \tilde{\Sigma}$ such that

$$\forall i, j \in [N, N+g(N)] \ \left(d(x_i, x_j) \leq \frac{1}{k+1} \text{ and } d(x_i, Tx_i) \leq \frac{1}{k+1} \right).$$

7.3 Mann iteration for strictly pseudo-contractive mappings

Assume that X is a real Hilbert space, $C \subseteq X$ is a nonempty bounded closed convex subset with finite diameter d_C and $0 \leq \kappa < 1$.
A mapping $T : C \to C$ is a *κ-strict pseudo-contraction* if for all $x, y \in C$,

$$\|Tx - Ty\|^2 \leq \|x - y\|^2 + \kappa \|x - Tx - (y - Ty)\|^2. \tag{15}$$

This definition was given by Browder and Petryshyn in [8], where they also proved that under the above hypothesis $F = Fix(T) \neq \emptyset$. Obviously, a mapping T is nonexpansive if and only if T is a 0-strict pseudo-contraction.
In the following, T is a κ-strict pseudo-contraction. Then T is Lipschitz continuous with Lipschitz constant $L = \frac{1+\kappa}{1-\kappa}$ (see [43]), hence T is uniformly continuous with modulus $\omega_T(k) = L(k+1)$. By Lemma 7.1.(v), it follows that F is uniformly closed with moduli

$$\omega_F(k) = L(4k+4) \text{ and } \delta_F(k) = 2k+1.$$

We consider the Mann iteration associated to T which, as we remarked above, is defined by

$$x_0 := x, \quad x_{n+1} := (1-\lambda_n)x_n + \lambda_n Tx_n, \tag{16}$$

where (λ_n) is a sequence in $(0,1)$.

Lemma 7.10. *Assume that (λ_n) is a sequence in $(\kappa, 1)$ and let $b \geq d_C$. Then*

(i) For all $n, m \in \mathbb{N}$ and all $p \in C$,

$$\|x_{n+1} - p\|^2 \leq \|x_n - p\|^2 + 2b(n+3)\|p - Tp\| \tag{17}$$
$$\|x_{n+m} - p\|^2 \leq \|x_n - p\|^2 + mb(2n+m+5)\|p - Tp\|. \tag{18}$$

(ii) (x_n) is uniformly (G, H)-Fejér monotone w.r.t. F with modulus

$$\chi(n, m, r) = m(2n+m+5)(r+1)\lceil b \rceil,$$

where $G(a) = H(a) = a^2$ with G-modulus $\alpha_G(k) = \left\lceil \sqrt{k} \right\rceil$ and H-modulus $\beta_H(k) = k^2$.

Proof. (i) (17) follows from [22, Lemma 3.4.(ii)]. We prove that

$$\|x_{n+m} - p\|^2 \leq \|x_n - p\|^2 + 2b\sum_{k=0}^{m-1}(n+k+3)\|p - Tp\|$$

by induction on m.

(ii) Apply (18). Assume that $a \leq \frac{1}{\alpha_G(k)+1} \leq \frac{1}{\sqrt{k+1}}$. Then $G(a) = a^2 \leq \frac{1}{k+1+2\sqrt{k}} \leq \frac{1}{k+1}$. Assume that $H(a) = a^2 \leq \frac{1}{\beta_H(k)+1} \leq \frac{1}{(k+1)^2}$. Then $a \leq \frac{1}{k+1}$.

□

Effective rates of asymptotic regularity for the Mann iteration (x_n) are computed in [22]: if (λ_n) is a sequence in $(\kappa, 1)$ satisfying $\sum_{n=0}^{\infty}(\lambda_n - \kappa)(1-\lambda_n) = \infty$ with rate of divergence $\theta : \mathbb{N} \to \mathbb{N}$, then

$$\Phi^{++}(k, b, \theta) = \theta\left(\lceil b^2 \rceil (k+1)^2\right) \tag{19}$$

is a rate of asymptotic regularity for (x_n). Thus,

$$(\Phi^{++})^M(k, b, \theta) = \theta^M\left(\lceil b^2 \rceil (k+1)^2\right).$$

If $\lambda_n = \lambda$, one gets the following rate of asymptotic regularity for the Krasnoselskii iteration:

$$\Phi^{++}(k, b, \kappa, \lambda) = \left\lceil \frac{b^2}{(\lambda-\kappa)(1-\lambda)} \right\rceil (k+1)^2. \tag{20}$$

Thus, Theorem 5.3 can be applied now to obtain rates of metastability for the Mann iteration, in the case when C is totally bounded.

7.4 Mann iteration for mappings satisfying condition (E)

Assume that X is a W-hyperbolic space and $C \subseteq X$ is nonempty and convex. Let $T : C \to C$ and $\mu \geq 1$. The mapping T satisfies *condition* (E_μ) if for all $x, y \in C$,

$$d(x, Ty) \leq \mu d(Tx, x) + d(x, y).$$

T is said to satisfy *condition* (E) if it satisfies (E_μ) for some $\mu \geq 1$. This condition was introduced in [13] as a generalization of condition (C) studied in [51]. Note that condition (C) is a generalization of nonexpansivity and implies (E_3).

We suppose next that T is a mapping satisfying condition (E_μ) with $\mu \geq 1$. Then F is uniformly closed with moduli

$$\delta_F(k) = 2\mu(k+1) - 1 \text{ and } \omega_F(k) = 4k+3.$$

Indeed, for $k \in \mathbb{N}$, $p, q \in X$ with $d(q, Tq) \leq 1/(2\mu(k+1))$ and $d(p, q) \leq 1/(4(k+1))$ we have that

$$d(p, Tp) \leq d(p, q) + d(q, Tp) \leq 2d(p, q) + \mu d(q, Tq) \leq \frac{1}{k+1}.$$

Let (x_n) be the Mann iteration starting with $x \in C$ defined as follows:

$$x_0 = x, \quad x_{n+1} = (1-\lambda_n)x_n + \lambda_n T(x_n),$$

where $(\lambda_n) \subseteq [1/L, 1 - 1/L]$ for some $L \geq 2$.

Lemma 7.11. *(i) For all $n, m \in \mathbb{N}$ and all $p \in C$,*

$$d(x_{n+m}, p) \quad \leq \quad d(x_n, p) + \mu m(1 - 1/L)d(p, Tp). \tag{21}$$

(ii) (x_n) is uniformly Fejér monotone w.r.t. F with modulus

$$\chi(n, m, r) = \mu m(1 - 1/L)(r + 1).$$

Proof. (i) is proved by induction. When $m = 0$ this is clear. Suppose

$$d(x_{n+m}, p) \leq d(x_n, p) + \mu m(1 - 1/L)d(p, Tp).$$

Then

$$\begin{aligned} d(x_{n+m+1}, p) &= d((1 - \lambda_{n+m})x_{n+m} + \lambda_{n+m}Tx_{n+m}, p) \\ &\leq (1 - \lambda_{n+m})d(x_{n+m}, p) + \lambda_{n+m}d(Tx_{n+m}, p) \\ &\leq (1 - \lambda_{n+m})d(x_{n+m}, p) + \lambda_{n+m}(\mu d(p, Tp) + d(p, x_{n+m})) \\ &\leq d(x_{n+m}, p) + \mu(1 - 1/L)d(p, Tp) \\ &\leq d(x_n, p) + \mu(m + 1)(1 - 1/L)d(p, Tp). \end{aligned}$$

(ii) follows easily from (21).

□

We compute next a rate of metastability for the asymptotic regularity of (x_n) w.r.t. F in the setting of UCW-hyperbolic spaces.

Lemma 7.12. *Let (X, d, W) be a UCW-hyperbolic space with a monotone modulus of uniform convexity η. Let $x, p \in C$, $n \in \mathbb{N}$ and $\alpha, \beta, \delta, \nu > 0$ such that*

$$d(p, Tp) < \nu \leq \delta, \quad \alpha \leq d(x_n, p) \leq \beta, \quad \alpha \leq d(x_n, Tx_n).$$

Then

$$d(x_{n+1}, p) < d(x_n, p) + \mu\nu - 2\alpha L^{-2}\eta\left(\mu\delta + \beta, \frac{\alpha}{\mu\delta + \beta}\right). \tag{22}$$

If $\eta(r, \varepsilon) \geq \varepsilon \cdot \tilde{\eta}(r, \varepsilon)$ with $\tilde{\eta}$ increasing w.r.t. ε, then one can replace η by $\tilde{\eta}$ in (22).

Proof. Let $r_n = \mu d(p, Tp) + d(p, x_n) < \mu\delta + \beta$. Since $d(x_n, p) \leq r_n$, $d(Tx_n, p) \leq \mu d(p, Tp) + d(p, x_n) = r_n$ and $d(x_n, Tx_n) \geq \alpha > \frac{\alpha}{\mu\delta+\beta} r_n$, by uniform convexity, it follows that

$$\begin{aligned} d(x_{n+1}, p) &\leq \left(1 - 2\lambda_n(1 - \lambda_n)\eta\left(r_n, \frac{\alpha}{\mu\delta + \beta}\right)\right) r_n \quad \text{by [38, Lemma 7]} \\ &\leq \left(1 - 2\lambda_n(1 - \lambda_n)\eta\left(\mu\delta + \beta, \frac{\alpha}{\mu\delta + \beta}\right)\right) r_n \\ &\leq r_n - 2r_n L^{-2}\eta\left(\mu\delta + \beta, \frac{\alpha}{\mu\delta + \beta}\right) \\ &\leq d(p, x_n) + \mu d(p, Tp) - 2\alpha L^{-2}\eta\left(\mu\delta + \beta, \frac{\alpha}{\mu\delta + \beta}\right) \\ &< d(p, x_n) + \mu\nu - 2\alpha L^{-2}\eta\left(\mu\delta + \beta, \frac{\alpha}{\mu\delta + \beta}\right). \end{aligned}$$

The additional claim follows using α/r_n instead of $\alpha/(\mu\delta+\beta)$:

$$\begin{aligned} d(x_{n+1},p) &\leq r_n - 2r_nL^{-2}\eta\left(\mu\delta+\beta,\frac{\alpha}{r_n}\right) \\ &\leq r_n - 2\alpha L^{-2}\tilde{\eta}\left(\mu\delta+\beta,\frac{\alpha}{r_n}\right) \\ &\leq r_n - 2\alpha L^{-2}\tilde{\eta}\left(\mu\delta+\beta,\frac{\alpha}{\mu\delta+\beta}\right) \\ &< d(p,x_n)+\mu\nu - 2\alpha L^{-2}\tilde{\eta}\left(\mu\delta+\beta,\frac{\alpha}{\mu\delta+\beta}\right). \end{aligned}$$

□

Theorem 7.13. *Let (X,d,W) be a UCW-hyperbolic space with a monotone modulus of uniform convexity η. Let $x\in C$ and $b>0$ such that for any $\gamma>0$ there exists $p\in C$ with*

$$d(x,p)\leq b \quad \textit{and} \quad d(p,Tp)\leq\gamma.$$

Then for every $k\in\mathbb{N}$, $g:\mathbb{N}\to\mathbb{N}$,

$$\exists N\leq\Phi^+(k,g,L,b,\eta),\forall m\in[N,N+g(N)]\ \left(d(x_m,Tx_m)\leq\frac{1}{k+1}\right),$$

where

$$\begin{aligned} \Phi^+ &= h^M(0), \quad h(n)=g(n)+n+1, \quad M=\lceil 3(b+1)/\theta\rceil, \\ \theta &= \frac{1}{4(k+1)L^2}\eta\left(b+1,\frac{1}{4(k+1)(b+1)}\right). \end{aligned}$$

If η satisfies the extra property from Lemma 7.12, then one can replace it by $\tilde{\eta}$ in θ.

Proof. Let $k\in\mathbb{N}$, $g:\mathbb{N}\to\mathbb{N}$. Then there exists $p\in C$ such that $d(x,p)\leq b$ and $d(p,Tp)\leq 2^{-\Phi^+-2}/(3\mu)$. Take $n\leq\Phi^+$. Then $d(p,Tp)\leq 2^{-n-2}/(3\mu)$. By (21),

$$d(x_{n+1},p)\leq d(x_n,p)+\mu(1-1/L)d(p,Tp)\leq d(x_n,p)+2^{-n-2}/3.$$

Denote $a_n=d(x_n,p)$, $\alpha_0=1/6$ and $\alpha_n=\left(1-\sum_{i=0}^{n-1}2^{-i-1}\right)/6$ for $n\geq 1$. Apply [31, Proposition 6.4] with $b_n=\beta_n=\gamma_n=0$, $c_n=2^{-n-2}/3$, $B_1=B_2=C_2=0$, $A_1=b$, $A_2=1/6$, $C_1=1/6$, $\tilde{g}(n)=g(n)+1$ to get that for all $n\leq\Phi^++1$, $d(x_n,p)\leq b+1/6$ and that there exists $N=h^s(0)$ for some $s<M$ such that

$$\forall i,j\in[N,N+g(N)+1],\quad |a_i-a_j|\leq\theta,\quad |\alpha_i-\alpha_j|\leq\theta.$$

We show that

$$\forall m\in[N,N+g(N)],\quad d(x_m,Tx_m)\leq\frac{1}{k+1}.$$

Let $m\in[N,N+g(N)]$. Suppose $d(x_m,Tx_m)>1/(k+1)$. Since $m,m+1\in[N,N+g(N)+1]$ we have that

$$|d(x_{m+1},p)-d(x_m,p)|\leq\theta,\quad |\alpha_{m+1}-\alpha_m|=2^{-m-2}/3\leq\theta.$$

Assume that $d(x_m,p) \geq 1/(4(k+1))$. Note that $m \leq N+g(N) < h(N) = h^{s+1}(0) \leq h^M(0) = \Phi^+$. Hence, $d(p,Tp) \leq 2^{-\Phi^+-2}/(3\mu) < 2^{-m-2}/(3\mu) \leq 1/(3\mu)$. Apply (22) with $\alpha = 1/(4(k+1))$, $\beta = b+2/3$, $\nu = 2^{-m-2}/(3\mu)$ and $\delta = 1/(3\mu)$ to obtain that

$$d(x_{m+1},p) < d(x_m,p) + 2^{-m-2}/3 - 2\theta.$$

This yields that $2\theta < d(x_m,p) - d(x_{m+1},p) + 2^{-m-2}/3 \leq 2\theta$, a contradiction. So, $d(x_m,p) < 1/(4(k+1))$. Then

$$\begin{aligned} d(x_m,Tx_m) &\leq d(x_m,p) + d(p,Tx_m) \leq 2d(x_m,p) + \mu d(p,Tp) \\ &\leq \frac{1}{2(k+1)} + 2^{-m-2}/3 \leq \frac{1}{2(k+1)} + \theta \leq \frac{1}{k+1}. \end{aligned}$$

□

In the particular case where in the above result the mapping $g = 0$, we obtain an approximate fixed point bound for (x_n) in the context of UCW-hyperbolic spaces.

In case of CAT(0) spaces this bound is quadratic in the error since we can take $\eta(r,\varepsilon) := \varepsilon^2/8$ and so $\tilde{\eta}(r,\varepsilon) := \varepsilon/8$. We then get

$$\Phi(k,L,b) = 384\left((b+1)(k+1)L\right)^2$$

as approximate fixed point bound.

Having such an approximate fixed point bound Φ, because (x_n) is additionally uniformly Fejér monotone w.r.t. F and F is uniformly closed, we can apply Theorem 5.3 to obtain, for C convex and totally bounded with II-modulus of total boundedness γ, a result analogous to Theorem 7.7.(ii).

7.5 Mann iteration for asymptotically nonexpansive mappings

Let X be a W-hyerbolic space, $C \subseteq X$ a convex subset and (k_n) be a sequence in $[0,\infty)$ satisfying $\lim\limits_{n\to\infty} k_n = 0$.

A mapping $T : C \to C$ is said to be *asymptotically nonexpansive* [16] with sequence (k_n) if for all $x,y \in C$ and for all $n \in \mathbb{N}$,

$$d(T^n x, T^n y) \leq (1+k_n)d(x,y).$$

Let T be asymptotically nonexpansive with sequence (k_n) in $[0,\infty)$. We assume furthermore that (k_n) is bounded in sum by some $K \in \mathbb{N}$, i.e. $\sum\limits_{n=0}^{\infty} k_n \leq K$. As an immediate consequence, we get that T is Lipschitz continuous with Lipschitz constant $1+K$, hence, as in the case of strict pseudo-contractions, it follows that F is uniformly closed with moduli $\omega_F(k) = (1+K)(4k+4)$ and $\delta_F(k) = 2k+1$.

The Mann iteration starting with $x \in C$ is defined by

$$x_0 := x, \quad x_{n+1} := (1-\lambda_n)x_n + \lambda_n T^n(x_n), \tag{23}$$

where (λ_n) is a sequence in $\left[\frac{1}{L}, 1-\frac{1}{L}\right]$ for some $L \in \mathbb{N}, L \geq 2$.

Lemma 7.14. *(i) For all $n, m \in \mathbb{N}$ and all $p \in C$,*

$$d(x_{n+m}, p) \le e^K d(x_n, p) + e^K m(n+m+K) d(p, Tp).$$

(ii) (x_n) is uniformly (G,H)-Fejér monotone w.r.t. F with modulus

$$\chi(n, m, r) = m(n+m+K)\lceil e^K \rceil (r+1),$$

where $G(a) = id_{\mathbb{R}^+}$ and $H = e^K id_{\mathbb{R}_+}$. An H-modulus is given by $\beta_H(k) = \lceil e^K \rceil (k+1)$.

Proof. (i) By [24, Lemma 4.4]. (ii) Apply (i). □

Effective rates of metastability for asymptotic regularity (and, as a particular case, approximate fixed point bounds) for the Mann iteraton were obtained in [30] in the setting of uniformly convex Banach spaces and in [31] for the more general setting of UCW-hyperbolic spaces. Thus we can apply Theorems 5.1 and 5.3. The result of applying Theorem 5.1 gives essentially the rate of metastability that was first extracted in [24] in a more ad-hoc fashion and which now appears as an instance of a general schema for computing rates of metastability. In fact, [24] has been the point of departure of the present paper.

8 An application to the Proximal Point Algorithm

The proximal point algorithm is a well-known and popular method employed in approximating a zero of a maximal monotone operator. There exists an extensive literature on this topic which stems from the works of Martinet [44] and Rockafellar [50]. The method consists in constructing a sequence using successive compositions of resolvents which, under appropriate conditions, converges weakly to a zero of the considered maximal monotone operator. If imposing additional assumptions, one can even prove strong convergence. Here we show that we can apply our results to obtain a quantitative version of this algorithm in finite dimensional Hilbert spaces.

In the sequel H is a real Hilbert space and $A : H \to 2^H$ is a maximal monotone operator. We assume that the set zer$A$ of zeros of A is nonempty. For every $\gamma > 0$ let $J_{\gamma A} = (Id + \gamma A)^{-1}$ be the resolvent of γA. Then $J_{\gamma A}$ is a single-valued firmly nonexpansive mapping defined on H and zer$A = Fix(J_{\gamma A})$ for every $\gamma > 0$. We refer to [3] for a comprehensive reference on maximal monotone operators.
Let $x_0 \in H$ and (γ_n) be a sequence in $(0, \infty)$. The proximal point algorithm starting with $x_0 \in H$ is defined as follows:

$$x_{n+1} = J_{\gamma_n A} x_n.$$

Let us take $F :=$ zer$A$. One can easily see that

$$F = \bigcap_{k \in \mathbb{N}} \tilde{F}_k, \quad \text{where } \tilde{F}_k = \bigcap_{i \le k} \left\{ x \in H \mid \|x - J_{\gamma_i A} x\| \le \frac{1}{k+1} \right\}.$$

and that $AF_k = \tilde{F}_k$ for every $k \in \mathbb{N}$.
Furthermore, F is uniformly closed with moduli $\omega_F(k) = 4k+3$, $\delta_F(k) = 2k+1$.

Lemma 8.1. *(i) For all $n \in \mathbb{N}, m \in \mathbb{N}^*$ and all $p \in H$,*

$$\|x_{n+m} - p\| \leq \|x_n - p\| + \sum_{i=n}^{n+m-1} \|p - J_{\gamma_i A} p\|. \tag{24}$$

(ii) (x_n) is uniformly Fejér monotone w.r.t. F with modulus $\chi(n,m,r) = \max\{n+m-1, m(r+1)\}$.

Proof. (i) Remark that

$$\begin{aligned} \|x_{n+1} - p\| &= \|J_{\gamma_n A} x_n - p\| \leq \|J_{\gamma_n A} x_n - J_{\gamma_n A} p\| + \|J_{\gamma_n A} p - p\| \\ &\leq \|x_n - p\| + \|J_{\gamma_n A} p - p\| \end{aligned}$$

and use induction.

(ii) Apply (24) and the fact that $p \in AF_{\chi(n,m,r)}$ implies that for all $m \geq 1$ and all $l \leq m$,

$$\sum_{i=n}^{n+l-1} \|p - J_{\gamma_i A} p\| \leq \sum_{i=n}^{n+m-1} \|p - J_{\gamma_i A} p\| \leq \frac{m}{\chi(n,m,r)+1} < \frac{1}{r+1}.$$

□

In the following we consider for $n \in \mathbb{N}$,

$$u_n = \frac{x_n - x_{n+1}}{\gamma_n}.$$

The next lemma is well-known. We refer, e.g., to the proof of [3, Theorem 23.41] and [3, Exercise 23.2, p. 349].

Lemma 8.2. *(i) For every $p \in zerA$ and every $n, i \in \mathbb{N}$,*

$$\|x_{n+1} - p\|^2 \leq \|x_n - p\|^2 - \|x_n - x_{n+1}\|^2 \tag{25}$$

$$\|J_{\gamma_n A} x_n - J_{\gamma_i A} x_n\| \leq |\gamma_n - \gamma_i| \frac{\|x_n - x_{n+1}\|}{\gamma_n} \tag{26}$$

$$\|x_n - J_{\gamma_i A} x_n\| \leq \|x_n - x_{n+1}\| + |\gamma_n - \gamma_i| \frac{\|x_n - x_{n+1}\|}{\gamma_n}. \tag{27}$$

(ii) The sequence $(\|u_n\|)$ is nonincreasing.

Lemma 8.3. *Assume that $\sum_{i=0}^{\infty} \gamma_n^2 = \infty$ with a rate of divergence θ and that $b > 0$ is an upper bound on $\|x_0 - p\|$ for some $p \in zerA$. Then*

(i) $\liminf_{n\to\infty} \|x_n - x_{n+1}\| = 0$ with modulus of liminf

$$\Delta(k, L, b) := \lceil b^2 (k+1)^2 \rceil + L - 1, \text{ i.e.}$$

for every $k \in \mathbb{N}$ and $L \in \mathbb{N}$ there exists $L \leq N \leq \Delta(k,L,b)$ such that $\|x_n - x_{n+1}\| \leq 1/(k+1)$.

(ii) $\lim_{n\to\infty} u_n = 0$ *with rate of convergence* $\beta(k,\theta,b) := \theta\left(\lceil b^2(k+1)^2\rceil\right)$.

Proof. (i) Applying (25) repeatedly we get that

$$\sum_{n=0}^{\infty} \|x_n - x_{n+1}\|^2 \leq \|x_0 - p\|^2 \leq b^2.$$

Let $k, L \in \mathbb{N}$ and $\Delta := \Delta(k,L,b)$. Suppose that for every $L \leq n \leq \Delta$, $\|x_n - x_{n+1}\| > 1/(k+1)$. Then

$$(\Delta - L + 1)\,\frac{1}{(k+1)^2} < \sum_{n=L}^{\Delta} \|x_n - x_{n+1}\|^2 \leq b^2,$$

which yields $\Delta < b^2(k+1)^2 + L - 1$, a contradiction.

(ii) Let $k \in \mathbb{N}$ and $\beta := \beta(k,\theta,b)$. Since $(\|u_n\|)$ is nonincreasing, it is enough to show that there exists $0 \leq N \leq \beta$ such that $\|u_N\| \leq 1/(k+1)$. Suppose that for every $0 \leq n \leq \beta$, $\|u_n\| > 1/(k+1)$. Then

$$\begin{aligned}
\frac{1}{(k+1)^2}\left\lceil b^2(k+1)^2\right\rceil &\leq \frac{1}{(k+1)^2}\sum_{n=0}^{\beta}\gamma_n^2 < \sum_{n=0}^{\beta}\gamma_n^2\|u_n\|^2 \\
&= \sum_{n=0}^{\beta}\|x_n - x_{n+1}\|^2 \leq b^2.
\end{aligned}$$

We have obtained a contradiction.

□

Theorem 8.4. *Assume that* $\sum_{i=0}^{\infty}\gamma_n^2 = \infty$ *with a rate of divergence* θ. *Then* (x_n) *has approximate* F*-points with an approximate* F*-point bound*

$$\Phi(k, m_k, \theta, b) := \theta\left(\lceil b^2(M_k+1)^2\rceil\right)\lceil b^2(M_k+1)^2\rceil - 1,$$

where $m_k = \max_{0\leq i\leq k}\gamma_i$, $M_k = \lceil (k+1)(2+m_k)\rceil - 1$ *and* $b > 0$ *is such that* $b \geq \|x_0 - p\|$ *for some* $p \in zerA$.

Proof. Let $k \in \mathbb{N}$. By Lemma 8.3.(i), there exists $N_1 \leq \Delta(M_k, 0, b)$ such that

$$\|x_{N_1} - x_{N_1+1}\| \leq \frac{1}{M_k+1} \leq \frac{1}{(k+1)(2+m_k)}.$$

If $\gamma_{N_1} \geq 1$, it follows by (27) that for all $i \leq k$,

$$\begin{aligned}
\|x_{N_1} - J_{\gamma_i A}x_{N_1}\| &\leq \|x_{N_1} - x_{N_1+1}\| + |\gamma_{N_1} - \gamma_i|\frac{\|x_{N_1} - x_{N_1+1}\|}{\gamma_{N_1}} \\
&\leq \left(2 + \frac{\gamma_i}{\gamma_{N_1}}\right)\|x_{N_1} - x_{N_1+1}\| \\
&\leq (2+m_k)\frac{1}{(k+1)(2+m_k)} = \frac{1}{k+1}.
\end{aligned}$$

Assume that $\gamma_{N_1} < 1$. Apply again Lemma 8.3.(i) to get the existence of $N_2 \leq \Delta(M_k, N_1+1, b)$ such that $N_2 > N_1$ and $\|x_{N_2} - x_{N_2+1}\| \leq \frac{1}{M_k+1}$. If $\gamma_{N_2} \geq 1$ we use again the above argument. If $\gamma_{N_2} < 1$, we apply once more the fact that Δ is a modulus of liminf for $\|x_n - x_{n+1}\|$. Let us denote for simplicity $\beta := \theta\left(\lceil b^2(M_k+1)^2\rceil\right)$. Applying this argument β times we get a finite sequence $N_1 < N_2 < \ldots < N_\beta$ such that either $\gamma_{N_j} \geq 1$ for some j or $\gamma_{N_j} < 1$ for all $j = 1, \ldots, \beta$. In the first case, we have as above that $\|x_{N_j} - J_{\gamma_i A} x_{N_j}\| \leq \frac{1}{k+1}$ for all $i \leq k$. In the second case, since $N_\beta \geq \beta$ and $(\|u_n\|)$ is nonincreasing, an application of Lemma 8.3.(ii) for M_k gives us

$$\|u_{N_\beta}\| \leq \|u_\beta\| \leq \frac{1}{M_k+1} \leq \frac{1}{(k+1)(2+m_k)}.$$

It follows then that for all $i \leq k$,

$$\begin{aligned} \|x_{N_\beta} - J_{\gamma_i A} x_{N_\beta}\| &\leq \|x_{N_\beta} - x_{N_\beta+1}\| + |\gamma_{N_\beta} - \gamma_i| \frac{\|x_{N_\beta} - x_{N_\beta+1}\|}{\gamma_{N_\beta}} \\ &= \gamma_{N_\beta}\|u_{N_\beta}\| + |\gamma_{N_\beta} - \gamma_i|\|u_{N_\beta}\| \leq \left(2\gamma_{N_\beta} + \gamma_i\right)\|u_\beta\| \\ &\leq (2+m_k)\frac{1}{(k+1)(2+m_k)} = \frac{1}{k+1}. \end{aligned}$$

Since $N_1 \leq \Delta(M_k, 0, b) = \lceil b^2(M_k+1)^2\rceil - 1$ and for all $j = 2, \ldots, \beta$,

$$N_j \leq \Delta(M_k, N_{j-1}+1, b) = \lceil b^2(M_k+1)^2\rceil + N_{j-1}$$

we get that $N_\beta \leq \beta\lceil b^2(M_k+1)^2\rceil - 1$, which finishes the proof. □

As an immediate consequence of Proposition 4.3 and Remark 4.5 we obtain the well-known fact that in $\mathbb{R}^n$, under the hypothesis that $\sum_{i=0}^{\infty} \gamma_n^2 = \infty$, the proximal point algorithm converges strongly to a zero of the maximal monotone operator A. Furthermore, since $\|x_n\| \leq M := b + \|p\|$ (where b, p are as above) and, by Example 2.8, $\overline{B}(0, M)$ is totally bounded with II-modulus $\gamma(k) = \lceil 2(k+1)\sqrt{n}M\rceil^n$, we can apply the quantitative Theorem 5.3 to get rates of metastability for (x_n).

Acknowledgements:

Ulrich Kohlenbach was supported by the German Science Foundation (DFG Project KO 1737/5-2).

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

Adriana Nicolae was supported by a grant of the Romanian Ministry of Education, CNCS - UEFISCDI, project number PN-II-RU-PD-2012-3-0152.

References

[1] D. Ariza-Ruiz, L. Leuştean, G. Lopez-Acedo, Firmly nonexpansive mappings in classes of geodesic spaces, Trans. Amer. Math. Soc. 366 (2014), 4299–4322.

[2] J. Avigad, J. Iovino, Ultraproducts and metastability, New York J. Math. 19 (2013), 713–727.

[3] H.H. Bauschke, P.L. Combettes, Convex Analysis and Monotone Operator Theory in Hilbert Spaces, Springer, New York-Dordrecht-Heidelberg-London, 2010.

[4] H.H. Bauschke, S.M. Moffat, X. Wang, Firmly nonexpansive mappings and maximally monotone operators: correspondence and duality, Set-Valued Var. Anal. 20 (2012), 131–153.

[5] M. Bridson, A. Haefliger, Metric Spaces of Non-Positive Curvature, Springer, Berlin-Heidelberg, 1999.

[6] F.E. Browder, Convergence theorems for sequences of nonlinear operators in Banach spaces, Math. Z. 100 (1967), 201–225.

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

[8] F.E. Browder, W.V. Petryshyn, Construction of fixed points of nonlinear mappings in Hilbert spaces, J. Math. Anal. Appl. 20 (1967), 197–228.

[9] R.E. Bruck, Nonexpansive projections on subsets of Banach spaces, Pacific J. Math. 47 (1973), 341–355.

[10] P.L. Combettes, Fejér monotonicity in convex optimization, in: C.A. Floudas, P.M. Pardalos (eds.), Encyclopedia of Optimization. Second edition, Springer, 2009, 1016–1024.

[11] P.L. Combettes, Quasi-Fejérian analysis of some optimization algorithms, in: D. Butnariu, Y. Censor, S. Reich (eds.), Inherently Parallel Algorithms for Feasibility and Optimization, Elsevier, 2001, 115–152.

[12] Yu.M. Ermolév, A.D. Tuniev, Random Fejér and quasi-Fejér sequences, Theory of Optimal Solutions, Akademiya Nauk Ukrainskoi SSR Kiev 2 (1968) 76-83; translated in: American Mathematical Society Selected Translations in Mathematical Statistics and Probability 13 (1973), 143–148.

[13] J. García-Falset, E. Llorens-Fuster, T. Suzuki, Fixed point theory for a class of generalized nonexpansive mappings, J. Math. Anal. Appl. 375 (2011), 185–195.

[14] P. Gerhardy, Proof mining in topological dynamics, Notre Dame J. Form. Log. 49 (2008), 431–446.

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

[16] K. Goebel, W.A. Kirk, A fixed point theorem for asymptotically nonexpansive mappings, Proc. Amer. Math. Soc. 35 (1972), 171–174.

[17] K. Goebel, S. Reich, Uniform Convexity, Hyperbolic Geometry, and Nonexpansive Mappings, Marcel Dekker, New York-Basel, 1984.

[18] C.W. Groetsch, A note on segmenting Mann iterates, J. Math. Anal. Appl. 40 (1972), 369–372.

[19] D. Günzel, U. Kohlenbach, Logical metatheorems of abstract spaces axiomatized in positive bounded logic. To appear in: Adv. Math.

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

[21] S. Ishikawa, Fixed points by a new iteration method, Proc. Amer. Math. Soc. 44 (1974), 147–150.

[22] D. Ivan, L. Leuştean, A rate of asymptotic regularity for the Mann iteration of κ-strict pseudo-contractions, Numer. Funct. Anal. Optimiz. 36 (2015), 792-798.

[23] U. Kohlenbach, Things that can and things that cannot be done in PRA, Ann. Pure Appl. Logic 102 (2000), 223–245.

[24] U. Kohlenbach, Some computational aspects of metric fixed point theory, Nonlinear Anal. 61 (2005), 823–837.

[25] U. Kohlenbach, Some logical metatheorems with applications in functional analysis, Trans. Amer. Math. Soc. 357 (2005), 89–128.
[26] U. Kohlenbach, Applied Proof Theory: Proof Interpretations and their Use in Mathematics, Springer Monographs in Mathematics, Springer, Berlin, 2008.
[27] U. Kohlenbach, On quantitative versions of theorems due to F.E. Browder and R. Wittmann, Adv. Math. 226 (2011), 2764–2795.
[28] U. Kohlenbach, A uniform quantitative form of sequential weak compactness and Baillon's nonlinear ergodic theorem, Commun. Contemp. Math. 14 (2012), 20pp.
[29] U. Kohlenbach, On the quantitative asymptotic behavior of strongly nonexpansive mappings in Banach and geodesic spaces. To appear in: Israel Journal of Mathematics.
[30] U. Kohlenbach, B. Lambov, Bounds on iterations of asymptotically quasi-nonexpansive mappings, in: J. Garcia Falset, E. Llorens Fuster, B. Sims (eds.), International Conference on Fixed Point Theory and Applications. Proceedings of the conference held in Valencia, July 13–19, 2003, Yokohama Publ., 2004, 143–172.
[31] U. Kohlenbach, L. Leuştean, Asymptotically nonexpansive mappings in uniformly convex hyperbolic spaces, J. Eur. Math. Soc. 12 (2010), 71–92.
[32] U. Kohlenbach, P. Safarik, Fluctuations, effective learnability and metastability in analysis, Ann. Pure Appl. Logic 165 (2014), 266–304.
[33] U. Kohlenbach, G. López-Acedo, A. Nicolae, Asymptotic regularity for the composition of two mappings. Preprint 2015, submitted.
[34] E. Kopecká, S. Reich, Asymptotic behavior of resolvents of coaccretive operators in the Hilbert ball, Nonlinear Anal. 70 (2009), 3187–3194.
[35] D. Körnlein, U. Kohlenbach, Rate of metastability for Bruck's iteration of pseudocontractive mappings in Hilbert space, Numer. Funct. Anal. Optim. 35 (2014), 20–31.
[36] M.A. Krasnoselski, Two remarks on the method of successive approximation, Uspekhi Mat. Nauk 10 (1955), 123–127 (in Russian).
[37] L. Leuştean, Proof mining in $\mathbb{R}$-trees and hyperbolic spaces, Electron. Notes Theor. Comput. Sci. 165 (2006), 95–106.
[38] L. Leuştean, A quadratic rate of asymptotic regularity for CAT(0)-spaces, J. Math. Anal. Appl. 325 (2007), 386–399.
[39] L. Leuştean, Nonexpansive iterations in uniformly convex W-hyperbolic spaces, in: A. Leizarowitz, B.S. Mordukhovich, I. Shafrir, A. Zaslavski (eds.), Nonlinear Analysis and Optimization I: Nonlinear Analysis, Cont. Math. 513, Amer. Math. Soc., 2010, 193–209.
[40] L. Leuştean, An application of proof mining to nonlinear iterations, Ann. Pure Appl. Logic 165 (2014), 1484–1500.
[41] G.G. Lorentz, Metric entropy and approximation, Bull. Amer. Math. Soc. 72 (1966), 903–937.
[42] W.R. Mann, Mean value methods in iteration, Proc. Amer. Math. Soc. 4 (1953), 506–510.
[43] G. Marino, H.-K. Xu, Weak and strong convergence theorem for strict pseudo-contractions in Hilbert spaces, J. Math. Anal. Appl. 329 (2007), 336–346.
[44] B. Martinet, Régularisation dinéquations variationnelles par approximations successives, Rev. Française Informat. Recherche Opérationnelle 4 (1970), 154–158.
[45] G.J. Minty, Monotone (nonlinear) operators in Hilbert space, Duke Math. J. 29 (1962), 341-346.
[46] E. Neumann, Computational problems in metric fixed point theory and their Weihrauch degrees., arXiv:1506.05127 [math.LO]; to appear in: Logical Methods in Computer Science.
[47] A. Nicolae, Asymptotic behavior of averaged and firmly nonexpansive mappings in geodesic spaces, Nonlinear Anal. 87 (2013) 102–115.

[48] S. Reich, I. Shafrir, The asymptotic behavior of firmly nonexpansive mappings, Proc. Amer. Math. Soc. 101 (1987), 246–250.
[49] S. Reich, I. Shafrir, Nonexpansive iterations in hyperbolic spaces, Nonlinear Anal. 15 (1990), 537–558.
[50] T. Rockafellar, Monotone operators and the proximal point algorithm, SIAM J. Control Optim. 14 (1976), 877–898.
[51] T. Suzuki, Fixed point theorems and convergence theorems for some generalized nonexpansive mappings, J. Math. Anal. Appl. 340 (2008), 1088–1095.
[52] 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., Amer. Math. Soc., Providence, RI , 2008.
[53] T. Tao, Norm convergence of multiple ergodic averages for commuting transformations, Ergodic Theory Dynam. Systems 28 (2008), 657–688.
[54] V.V. Vasin, I.I. Eremin, Operators and Iterative Processes of Fejér Type. Theory and Applications, Walter de Gruyter, Berlin-New York, 2009.
[55] M. Walsh, Norm convergence of nilpotent ergodic averages, Ann. Math. 175 (2012), 1667–1688.
[56] R. Wittmann, Approximation of fixed points of nonexpansive mappings, Arch. Math. 58 (1992), 486–491.