Moduli of regularity and rates of convergence for Fejér monotone sequences

Ulrich Kohlenbach[a], Genaro López-Acedo[b], Adriana Nicolae[c,b]

April 28, 2019

[a]Department of Mathematics, Technische Universität Darmstadt, Schlossgartenstraße 7, 64289 Darmstadt, Germany
[b]Department of Mathematical Analysis - IMUS, University of Seville, Sevilla, Spain
[c]Department of Mathematics, Babeş-Bolyai University, Kogălniceanu 1, 400084 Cluj-Napoca, Romania

E-mail addresses: kohlenbach@mathematik.tu-darmstadt.de (U. Kohlenbach), glopez@us.es (G. López-Acedo), anicolae@math.ubbcluj.ro (A. Nicolae)

Abstract

In this paper we introduce the concept of modulus of regularity as a tool to analyze the speed of convergence, including the linear convergence and finite termination, for classes of Fejér monotone sequences which appear in fixed point theory, monotone operator theory, and convex optimization. This concept allows for a unified approach to several notions such as weak sharp minima, error bounds, metric subregularity, Hölder regularity, etc., as well as to obtain rates of convergence for Picard iterates, the Mann algorithm, the proximal point algorithm and the cyclic projection method. As a byproduct we obtain a quantitative version of the well-known fact that for a convex lower semi-continuous function the set of minimizers coincides with the set of zeros of its subdifferential and the set of fixed points of its resolvent.

MSC: 41A25; 41A52; 41A65; 53C23; 03F10

Keywords: Fejér monotone sequences, rates of convergence, linear convergence, finite termination, metric subregularity, Hölder regularity, weak sharp minima, bounded regularity.

1 Introduction

Many problems in applied mathematics can be brought into the following format:

Let (X, d) be a metric space and $F : X \to \overline{\mathbb{R}}$ be a function: find a zero of F,

where as usual $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$. This statement covers many equilibrium, fixed point and minimization problems. Numerical methods, e.g. those based on suitable iterative techniques, usually yield sequences (x_n) in X of $(1/n)$-approximate zeros, i.e., $|F(x_n)| < 1/n$. Based on extra assumptions (e.g., the compactness of X, the Fejér monotonicity of (x_n) and the continuity of F) one then shows that (x_n) converges to an actual zero z of F. An obvious question then concerns the speed of the convergence of (x_n) to z and whether there is an effective rate of convergence.

For general families of such problems formulated for a whole class $\mathcal{F}$ of functions F one largely has the following dichotomy:

(i) if the zero for $F \in \mathcal{F}$ is unique, then it usually is possible to give an explicit effective rate of convergence,

(ii) if $\mathcal{F}$ contains functions F with many zeros, one usually can use the non-uniqueness to define a (computable) function $F \in \mathcal{F}$ for which (x_n) does not have a computable rate of convergence.

'(i)' e.g. holds for most fixed point iterations involving functions $T : X \to X$ which satisfy some form of a contractive condition which guarantees the uniqueness of the fixed point (and hence of the zero of $F(x) = d(x, Tx)$). The obvious case, of course, is the Banach fixed point theorem, but there are also many situations where this is highly nontrivial and tools from logic were used (see [20] which in turn is based on methods from [36]) to extract explicit rates of convergence for Picard iterates, see, e.g., [2] and the references listed in [20].

'(ii)' is most strikingly exemplified in [49], where it is shown that all the usual iterations (x_n) used to compute fixed points of nonexpansive mappings already fail in general to have computable rates of convergence even for simple computable firmly nonexpansive mappings $T : [0, 1] \to [0, 1]$, i.e. there is no computable function $\alpha : \mathbb{N} \to \mathbb{N}$ s.t.

$$\forall k \in \mathbb{N} \setminus \{0\}\, \forall n, m \geq \alpha(k)\ (|x_n - x_m| < 1/k).$$

Here we refer to the standard definition of 'computable function' resulting in the 30's from the work of Gödel, Church and Turing which has been extended to functions in $C[0, 1]$ in [29] and more general function spaces in [55, 60].

Even though sometimes left implicit, the construction of an explicit rate of convergence of iterative procedures in the case of unique zeros (or fixed points) usually rests on constructing a so-called modulus of uniqueness: let (X, d) be a metric space, $F : X \to \overline{\mathbb{R}}$ with zer $F = \{z\}$ and $r > 0$.

Definition 1.1. We say that $\phi : (0, \infty) \to (0, \infty)$ is a *modulus of uniqueness* for F w.r.t. zer F and $\overline{B}(z, r)$ if for all $\varepsilon > 0$ and $x \in \overline{B}(z, r)$ we have the following implication

$$|F(x)| < \phi(\varepsilon) \ \Rightarrow\ d(x, z) < \varepsilon.$$

Suppose now that (x_n) is a sequence of $(1/n)$-approximate zeros contained in $\overline{B}(z, r)$ for some $r > 0$. If ϕ is a modulus of uniqueness for F w.r.t. zer F and $\overline{B}(z, r)$, then

$$\forall k \geq \lceil 1/\phi(\varepsilon) \rceil\, (d(x_k, z) < \varepsilon).$$

The concept of 'modulus of uniqueness' (in the case where $X := K_u$ is a compact metric space which is parametrized by elements $u \in P$ in some Polish space P) can be found in [34]. In particular, it has been applied to the uniqueness of the best uniform (Chebycheff) as well as best L^1 approximation of $f \in C[0, 1]$ by algebraic polynomials of degree $\leq n$. Proof-theoretic metatheorems applied to the nonconstructive uniqueness proofs in these cases guarantee the extractability of explicit moduli of uniqueness (of low complexity) depending only on n, ε and a modulus of uniform continuity ω of f which has been carried out in [34, 35, 41] (see also [36] for a book treatment of all this).

In this paper, we are concerned with a generalization of the concept of 'modulus of uniqueness' called 'modulus of regularity' which is applicable also in the non-unique case by considering the distance of a point to the set zer F (see Definition 3.1). Note that this concept coincides with that of a 'modulus of uniqueness' if zer F is a singleton.

Again, whenever (x_n) is a sequence of $(1/n)$-approximate zeros of F in $\overline{B}(z, r)$, where $z \in$ zer F and $r > 0$, x_k is ε-close to some zero $z_k \in$ zer F for all $k \geq \lceil 1/\phi(\varepsilon) \rceil$. A condition which converts this into a rate of convergence is that (x_n) is Fejér monotone w.r.t. zer F, i.e., for all $z \in$ zer F and $n \in \mathbb{N}$

$$d(x_{n+1}, z) \leq d(x_n, z).$$

In this case we can infer that for all $k, m \geq \lceil 1/\phi(\varepsilon) \rceil$

$$d(x_k, x_m) < 2\varepsilon.$$

So if X is complete and zer F is closed, then (x_k) converges with rate $\lceil 1/\phi(\varepsilon/2) \rceil$ to a zero of F (see Theorem 4.1).

As discussed above, in general one cannot expect to have a computable rate of convergence in the non-unique case and so to be able to explicitly write down a modulus ϕ of regularity w.r.t. zer F (in cases where it exists) will rest on very specific properties of the individual mapping F. Nevertheless, the existence of a modulus of regularity w.r.t. zer F holds if X is compact and F is continuous with zer $F \neq \emptyset$ (see Proposition 3.3). This strikingly illustrates the difference between the unique and the non-unique case: a modulus of uniqueness is a uniform version of having a unique zero which - e.g. by logical techniques - can be extracted in explicit form from a given proof of uniqueness

$$F(x) = 0 = F(z) \to x = z,$$

(see [36, Section 15.2] with Corollary 17.54 instead of Theorem 15.1 to be used in the noncompact case) whereas a modulus of regularity w.r.t. zer F is a uniform version of the trivially true property

$$F(x) = 0 \to \forall \varepsilon > 0 \, \exists z \in \text{zer}\, F \, (d(x, z) < \varepsilon),$$

which does not carry information to be used to construct an explicit modulus of regularity even in the presence of compactness.

While the concept of a modulus of regularity (and also Proposition 3.3) has been used in various special situations before (see, e.g., [1] and the literature cited there), we develop it in this paper as a general tool towards a unified treatment of a number of concepts studied in convex optimization such as weak sharp minima, error bounds, metric subregularity, Hölder regularity, etc., which can be seen as instances of moduli of regularity w.r.t. zer F for suitable choices of F. Actually, as it will be pointed out in Section 3, for minimization problems the notion of modulus of regularity is tightly related to the ones of weak sharp minima or error bounds.

After some preliminaries, we show in Section 3 how the concept of 'modulus of regularity' w.r.t. zer F can be specialized to suitable notions of 'modulus of regularity' for equilibrium problems, fixed point problems, the problem of finding a zero of a set-valued operator and minimization problems. In Theorem 3.10, we give - in terms of the respective moduli of regularity - a quantitative version of the well-known identities between minimizers of proper, convex and lower semi-continuous functions f, fixed points of the resolvent $J_{\gamma \partial f}$ of ∂f of order $\gamma > 0$ and the zeros of ∂f :

$$\text{argmin } f = \text{Fix } J_{\gamma \partial f} = \text{zer } \partial f.$$

In Section 4 we use the concept of 'modulus of regularity' to give a general convergence result, Theorem 4.1, which provides, under suitable assumptions, explicit rates of convergence for Fejér monotone sequences. Moreover, we also focus on conditions that yield finite termination or linear convergence. We then apply these general results to various iterative methods such as cyclic projections, Picard and Mann iterations, as well as the proximal point algorithm.

2 Preliminaries

We start this section with some notations and a brief account of geodesic metric spaces. More details can be found, e.g., in [19]. Let (X, d) be a metric space. For $x \in X$ and $r > 0$, we denote the *open ball* and the *closed ball* centered at x with radius r by $B(x, r)$ and $\overline{B}(x, r)$, respectively.

If C is a subset of X, the *diameter* of C is diam $C = \sup\{d(x,y) : x, y \in C\}$, the *distance* of a point $x \in X$ to C is $\operatorname{dist}(x, C) = \inf\{d(x,c) : c \in C\}$, and the *metric projection* P_C onto C is the mapping $P_C : X \to 2^C$ defined by $P_C(x) = \{y \in C : d(x,y) = \operatorname{dist}(x,C)\}$.

Having $x, y \in X$, a *geodesic* from x to y is a mapping $c : [0,l] \subseteq \mathbb{R} \to X$ such that $c(0) = x$, $c(l) = y$ and $d(c(t), c(s)) = |t - s|$ for all $t, s \in [0, l]$. The image of c is called a *geodesic segment* joining x to y and is not necessarily unique. We say that X is a *(uniquely) geodesic metric space* if every two points in it are joined by a (unique) geodesic. If X is a geodedic space, then a point $z \in X$ belongs to a geodesic segment joining x to y if and only if there exists $t \in [0,1]$ such that $d(x,z) = td(x,y)$ and $d(y,z) = (1-t)d(x,y)$ and we write $z = (1-t)x + ty$ if no confusion arises. A set C in a uniquely geodesic metric space is called *convex* if given any two points in C, the geodesic segment joining them is contained in C.

Although most of the algorithms considered in the subsequent sections are defined in Hilbert spaces, in some situations we also refer to the context of $\mathrm{CAT}(\kappa)$ spaces, $\kappa \geq 0$, which can be defined via a quadrilateral condition. We state next this condition for the case $\kappa = 0$. More precisely, a geodesic metric space (X, d) is said to be a $\mathrm{CAT}(0)$ *space* if

$$d(x,y)^2 + d(u,v)^2 \leq d(x,v)^2 + d(y,u)^2 + d(x,u)^2 + d(y,v)^2, \tag{2.1}$$

for all $x, y, u, v \in X$ (see [14]). When $\kappa > 0$, a related inequality recently given in [15] can be used to introduce $\mathrm{CAT}(\kappa)$ spaces.

Every $\mathrm{CAT}(0)$ space X is uniquely geodesic and satisfies the following condition known as *Busemann convexity*

$$d((1-t)u + tv, (1-t)x + ty) \leq (1-t)d(u,x) + td(v,y), \tag{2.2}$$

for all $u, v, x, y \in X$ and all $t \in [0,1]$. The Hilbert ball with the hyperbolic metric is a prime example of a $\mathrm{CAT}(0)$ space, see [28]. Other examples include Hilbert spaces, $\mathbb{R}$-trees, Euclidean buildings, Hadamard manifolds, and many other important spaces.

Let X be a complete $\mathrm{CAT}(\kappa)$ space (with diam $X < \pi/(2\sqrt{\kappa})$ if $\kappa > 0$) and $C \subseteq X$ nonempty, closed and convex. Then $P_C : X \to C$ is well-defined, single-valued and satisfies

$$d(P_C x, y)^2 \leq d(x,y)^2 - \beta d(x, P_C x)^2, \tag{2.3}$$

for all $x \in X$ and $y \in C$, where $\beta > 0$. In $\mathrm{CAT}(0)$ spaces, $\beta = 1$. If $\kappa > 0$, explicit values for β can be given in terms of κ and diam X (see [4, 52]).

In the following we recall definitions and properties of operators which are significant in this paper. We refer to [12] for a detailed exposition on this topic. Let (X, d) be a metric space, $C \subseteq X$ nonempty and $T : C \to X$. The *fixed point set* of T is denoted by $\operatorname{Fix} T = \{x \in C : Tx = x\}$. The mapping T is said to be *nonexpansive* if $d(T(x), T(y)) \leq d(x,y)$ for all $x, y \in C$. Likewise, T is said to be *quasi-nonexpansive* if $\operatorname{Fix} T \neq \emptyset$ and $d(T(x), z) \leq d(x,z)$ for all $z \in \operatorname{Fix} T$. Suppose next that (X, d) is a uniquely geodesic metric space. If C is closed and convex, then $\operatorname{Fix} T$ is also closed and convex, whenever T is quasi-nonexpansive. We say that T is *firmly nonexpansive* if

$$d(T(x), T(y)) \leq d((1-\lambda)x + \lambda T(x), (1-\lambda)y + \lambda T(y)),$$

for all $x, y \in C$ and $\lambda \in [0,1]$. In complete $\mathrm{CAT}(0)$ spaces, the metric projection onto nonempty, closed and convex subsets is firmly nonexpansive (see [3]). When X is a Hilbert space, there are several equivalent definitions of firm nonexpansivity, one of them being that T can be written as $T = (1/2)\mathrm{Id} + (1/2)S$, where S is nonexpansive.

Let A be a set-valued operator defined on a Hilbert space H, $A : H \to 2^H$. We say that A is *monotone* if $\langle x^* - y^*, x - y \rangle \geq 0$ for all $x, y \in H$, $x^* \in A(x)$, $y^* \in A(y)$. Suppose next that

A is monotone. The *resolvent* of A of order $\gamma > 0$ is the mapping $J_{\gamma A} = (\mathrm{Id} + \gamma A)^{-1}$ defined on $\mathrm{ran}(\mathrm{Id} + \gamma A)$, which can be shown to be single-valued and firmly nonexpansive. Denoting the set of zeros of A by $\mathrm{zer}\ A = \{x \in H : O \in A(x)\}$, it immediately follows from the definition of the resolvent that $\mathrm{Fix}\ J_{\gamma A} = \mathrm{zer}\ A$. The *reflected resolvent* is the mapping $R_{\gamma A} = 2J_{\gamma A} - \mathrm{Id}$, which is nonexpansive as $J_{\gamma A}$ is firmly nonexpansive. If the monotone operator A has no proper monotone extension, then it is called *maximal monotone*. In this case $J_{\gamma A}$ and $R_{\gamma A}$ are defined on H.

Let $f : H \to (-\infty, \infty]$ be proper. The *subdifferential* of f is the set-valued operator $\partial f : H \to 2^H$ defined by

$$\partial f(x) = \{u \in H : \langle u, y - x\rangle \leq f(y) - f(x),\ \forall y \in H\}.$$

A very important property of ∂f is its monotonicity. Denoting the *set of minimizers* of f by $\mathrm{argmin}\, f = \{x \in H : f(x) \leq f(y),\ \forall y \in H\}$, we have $\mathrm{zer}\ \partial f = \mathrm{argmin}\, f = \mathrm{Fix}\ J_{\gamma\partial f}$ for all $\gamma > 0$.

Let $C \subseteq H$ be nonempty and convex. Recall that the *indicator function* $\delta_C : H \to [0, \infty]$ is defined by

$$\delta_C(x) = \begin{cases} 0, & \text{if } x \in C, \\ \infty, & \text{otherwise} \end{cases}$$

and the *normal cone map* $N_C : H \to 2^H$ is

$$N_C(x) = \begin{cases} \{u \in H : \langle u, c - x\rangle \leq 0,\ \forall c \in C\}, & \text{if } x \in C, \\ \emptyset, & \text{otherwise.} \end{cases}$$

Then $\partial\delta_C = N_C$. Suppose now that f is additionally convex and lower semi-continuous. Then $\mathrm{int\ dom}\ f = \mathrm{cont}\ f \subseteq \mathrm{dom}\ \partial f \subseteq \mathrm{dom}\ f$ and ∂f is a maximal monotone operator. Note that if $C \subseteq H$ is nonempty, closed and convex, then δ_C is proper, convex and lower semi-continuous and N_C is maximal monotone.

The mapping $\mathrm{Prox}_f : H \to H$,

$$\mathrm{Prox}_f(x) = \underset{y \in H}{\mathrm{argmin}} \left(f(y) + \frac{1}{2}\|x - y\|^2 \right), \tag{2.4}$$

is well-defined and called the *proximal mapping* of f. Note that $J_{\gamma\partial f} = \mathrm{Prox}_{\gamma f}$ for all $\gamma > 0$ and $J_{N_C} = \mathrm{Prox}_{\delta_C} = P_C$. One can also show that

$$f(J_{\gamma\partial f}(x)) - f(y) \leq \frac{1}{2\gamma}\left(\|y - x\|^2 - \|J_{\gamma\partial f}(x) - x\|^2 - \|J_{\gamma\partial f}(x) - y\|^2 \right), \tag{2.5}$$

for every $\gamma > 0$ and $x, y \in H$ (see, e.g., [5, Lemma 3.2]).

Let X be a complete CAT(0) space and $f : X \to (-\infty, \infty]$ a proper, convex and lower semi-continuous function. In this setting, the resolvent of f was introduced by Jost [32] via the proximal mapping which is well-defined. Namely, for any $\gamma > 0$, $J_\gamma : X \to X$ is given by

$$J_\gamma(x) = \mathrm{argmin}_{y \in X} \left(f(y) + \frac{1}{2\gamma} d^2(x, y) \right).$$

Again, $\mathrm{argmin}\, f = \mathrm{Fix}\ J_\gamma$ for every $\gamma > 0$. Moreover, J_γ is a firmly nonexpansive mapping (see [3]). Actually, (2.5) holds in this setting too (considering J_γ instead of $J_{\gamma\partial f}$ and d instead of the distance induced by the norm). In particular, if $\mathrm{argmin}\, f \neq \emptyset$ (and hence $\mathrm{Fix}\ J_\gamma \neq \emptyset$), then

$$d(J_\gamma x, y)^2 \leq d(x, y)^2 - d(x, J_\gamma x)^2, \tag{2.6}$$

for all $x \in X$ and $y \in \mathrm{Fix}\ J_\gamma$.

In the rest of this section we assume that (X,d) is a metric space and C is nonempty subset of X. As the notions of Fejér monotonicity and asymptotic regularity also play an important role in our further discussion, we recall them in what follows.

Let (x_n) be a sequence in X that converges to some $x \in X$. A function $\alpha : (0,\infty) \to \mathbb{N}$ is called a *rate of convergence* for (x_n) to x if

$$\forall \varepsilon > 0 \, \forall n \geq \alpha(\varepsilon) \; (d(x_n, x) < \varepsilon) \, .$$

We also say that (x_n) *converges linearly* to x if

$$\forall n \in \mathbb{N} \, (d(x_n, x) \leq Mc^n),$$

where $M > 0$ and $c \in (0,1)$. Any linearly convergent sequence has a rate of convergence (in the sense defined above) that is logarithmic in $1/\varepsilon$ for $\varepsilon < M$.

A sequence $(x_n) \subseteq X$ is *Fejér monotone* with respect to C if

$$\forall p \in C \, \forall n \in \mathbb{N} \, (d(x_{n+1}, p) \leq d(x_n, p)).$$

Let $F : X \to \overline{\mathbb{R}}$. We say that $(x_n) \subseteq X$ is a sequence of *approximate zeros* of F if $\lim_{n\to\infty} F(x_n) = 0$. Likewise, we say that (x_n) *has approximate zeros* for F if

$$\forall \varepsilon > 0 \, \exists n \in \mathbb{N} \; (|F(x_n)| < \varepsilon) \, .$$

Of particular importance is the case $F(x) = d(x, Tx)$, where $T : X \to X$. Then a sequence (x_n) of approximate zeros of F is also said to be *asymptotically regular* and a rate of convergence of $(d(x_n, Tx_n))$ to 0 is called a *rate of asymptotic regularity* for (x_n). We also say that (x_n) *has approximate fixed points* for T if (x_n) has approximate zeros for F.

We end this subsection with a definition that will be needed later on. A function $\theta : \mathbb{N} \to \mathbb{N}$ is a *rate of divergence* for a series $\sum_{n\geq 0} \gamma_n$, where $(\gamma_n) \subseteq \mathbb{R}_+$, if $\sum_{k=0}^{\theta(n)} \gamma_k \geq n$ for all $n \in \mathbb{N}$.

3 Modulus of regularity

Let (X,d) be a metric space and $F : X \to \overline{\mathbb{R}}$ with $\text{zer}\, F \neq \emptyset$.

Definition 3.1. Fixing $z \in \text{zer}\, F$ and $r > 0$, we say that $\phi : (0,\infty) \to (0,\infty)$ is a *modulus of regularity* for F w.r.t. zer F and $\overline{B}(z,r)$ if for all $\varepsilon > 0$ and $x \in \overline{B}(z,r)$ we have the following implication

$$|F(x)| < \phi(\varepsilon) \; \Rightarrow \; \text{dist}(x, \text{zer}\, F) < \varepsilon.$$

If there exists $z \in \text{zer}\, F$ such that $\phi : (0,\infty) \to (0,\infty)$ is a modulus of regularity for F w.r.t. zer F and $\overline{B}(z,r)$ for any $r > 0$, then ϕ is said to be a *modulus of regularity* for F w.r.t. zer F.

Remark 3.2. If ϕ is a modulus of regularity for F w.r.t. zer F and $\overline{B}(z,r)$, then $|F(x)| \geq \phi(\text{dist}(x, \text{zer}\, F))$, for all $x \in \overline{B}(z,r)$. Indeed, supposing that there exists $x \in \overline{B}(z,r)$ such that $|F(x)| < \phi(\text{dist}(x, \text{zer}\, F))$, then $\text{dist}(x, \text{zer}\, F) < \text{dist}(x, \text{zer}\, F)$, a contradiction. Thus, a modulus of regularity also induces a growth condition for the function $|F|$.

Our first result shows that such a modulus always exists when the function is continuous and the closed balls are compact.

Proposition 3.3. If X is proper and F is continuous, then for any $z \in \text{zer}\, F$ and $r > 0$, F has a modulus of regularity w.r.t. zer F and $\overline{B}(z,r)$.

Proof. It is enough to prove that

$$\forall \varepsilon > 0 \, \exists n \in \mathbb{N} \setminus \{0\} \, \forall x \in \overline{B}(z,r) \left(|F(x)| < \frac{1}{n} \to \exists q \in \operatorname{zer} F \ (d(x,q) < \varepsilon) \right).$$

Assume that this is not the case. Then there exist $\varepsilon > 0$ and a sequence (x_n) in $\overline{B}(z,r)$ such that

$$\forall n \in \mathbb{N} \setminus \{0\} \left(|F(x_n)| < \frac{1}{n} \wedge \forall q \in \operatorname{zer} F \ (d(x_n,q) \geq \varepsilon) \right). \tag{3.7}$$

Let $\widehat{x}$ be a limit point of (x_n). Then, using the continuity of F, we get $F(\widehat{x}) = 0$, i.e., $\widehat{x} \in \operatorname{zer} F$. Also

$$\exists n \in \mathbb{N} \ (d(x_n, \widehat{x}) < \varepsilon).$$

Putting $q = \widehat{x}$, this contradicts the last conjunct in (3.7). □

Remark 3.4. From the above, it follows that if X is compact and F is continuous, then F has a modulus of regularity w.r.t. zer F.

The notion of modulus of regularity appears in a natural way in different relevant problems such as the following ones. We only mention that an implicit use of this notion also appears in the concept of "controlled resolvent" in the context of the computational spectral problem (see [13]) which, however, is not studied in this paper.

Equilibrium problems

Given the nonempty subsets C and D of two Hilbert spaces H_1 and H_2, respectively, and a mapping $G : C \times D \to \mathbb{R}$, the equilibrium problem associated to the mapping G and the sets C and D consists of finding an element $p \in C$ such that

$$G(p,y) \geq 0, \tag{3.8}$$

for all $y \in D$.

Suppose that the set of solutions for problem (3.8), denoted by $\mathrm{EP}(G,C,D)$, is nonempty and define $F : C \to \mathbb{R}$,

$$F(x) = \min \left\{ 0, \inf_{y \in D} G(x,y) \right\}.$$

Note that $\operatorname{zer} F = \mathrm{EP}(G,C,D)$.

Let $z \in \mathrm{EP}(G,C,D)$ and $r > 0$. A *modulus of regularity* for G w.r.t. $\mathrm{EP}(G,C,D)$ and $\overline{B}(z,r)$ is a modulus of regularity for F w.r.t. zer F and $\overline{B}(z,r)$. This modulus appears, under the name of error bound, in the study of parametric inequality systems. In [46], such an approach is used to obtain rates of convergence for the cyclic projection method employed in solving convex feasibility problems (an early work on the cyclic projection method is [30]).

The equilibrium problem covers in particular the classical variational inequality problem. Given a nonempty, closed and convex subset C of a Hilbert space H and a mapping $A : C \to H$, the classical variational inequality problem associated to A and C consists of finding an element $z \in C$ such that

$$\langle A(z), y - z \rangle \geq 0, \tag{3.9}$$

for all $y \in C$. Denote by $\mathrm{VI}(A,C)$ the set of solutions for problem (3.9) and assume that it is nonempty. In this case one considers $G : C \times C \to \mathbb{R}$ defined by $G(x,y) = \langle A(x), y - x \rangle$ and, for $z \in \mathrm{VI}(A,C)$ and $\overline{B}(z,r)$, a *modulus of regularity* for A w.r.t. $\mathrm{VI}(A,C)$ and $\overline{B}(z,r)$ is a modulus of regularity for G w.r.t. $\mathrm{EP}(G,C,C)$ and $\overline{B}(z,r)$.

Fixed point problems

Let (X,d) be a metric space, $T : X \to X$ with $\mathrm{Fix}\, T \neq \emptyset$ and define $F : X \to \mathbb{R}$ by $F(x) = d(x, Tx)$. Note that $\mathrm{zer}\, F = \mathrm{Fix}\, T$.

Let $z \in \mathrm{Fix}\, T$ and $r > 0$. A *modulus of regularity* for T w.r.t. $\mathrm{Fix}\, T$ and $\overline{B}(z,r)$ is a modulus of regularity for F w.r.t. $\mathrm{zer}\, F$ and $\overline{B}(z,r)$. In a similar way, a *modulus of regularity* for T w.r.t. $\mathrm{Fix}\, T$ is defined to be a modulus of regularity for F w.r.t. $\mathrm{zer}\, F$.

A closely related concept already appeared in the early works [53, 58]. More recent references where this notion is used are, e.g., [18] and [47, 48] where it was employed, respectively, to study the linear and Hölder local convergence for algorithms related to nonexpansive mappings.

The next result is a direct consequence of Proposition 3.3 and Remark 3.4.

Corollary 3.5. If X is proper, T is continuous, $z \in \mathrm{Fix}\, T$ and $r > 0$, then T has a modulus of regularity w.r.t. $\mathrm{Fix}\, T$ and $\overline{B}(z,r)$. If X is additionally compact, then T has a modulus of regularity w.r.t. $\mathrm{Fix}\, T$.

We give next concrete instances of moduli of regularity that are computed explicitly.

Example 3.6. (i) Let X be a complete metric space and $T : X \to X$ a contraction with constant $k \in [0,1)$. Then $\mathrm{Fix}\, T = \{z\}$ for some $z \in X$ and $\phi(\varepsilon) = (1-k)\varepsilon$ is a modulus of regularity for T w.r.t. $\mathrm{Fix}\, T$ (in fact it is even a modulus of uniqueness). Indeed, $d(x,Tx) < (1-k)\varepsilon$ yields

$$d(x,z) \leq d(x,Tx) + d(Tx,Tz) < (1-k)\varepsilon + kd(x,z),$$

hence $d(x,z) < \varepsilon$.

(ii) Let X be a complete metric space and $T : X \to X$ an orbital contraction with constant $k \in [0,1)$ (i.e., $d(Tx, T^2x) \leq kd(x,Tx)$ for all $x \in X$). If T is additionally continuous, one can show that $\phi(\varepsilon) = (1-k)\varepsilon$ is a modulus of regularity for T w.r.t. $\mathrm{Fix}\, T$. To this end let $x \in X$ with $d(x,Tx) < \phi(\varepsilon)$ and let $n, l \in \mathbb{N}$. Then

$$d(T^n x, T^{n+l}x) \leq \sum_{i=0}^{l-1} d(T^{n+i}x, T^{n+i+1}x) \leq \sum_{i=0}^{l-1} k^{n+i} d(x,Tx) \leq \frac{k^n}{1-k} d(x,Tx).$$

This shows that $(T^n x)$ is a Cauchy sequence, hence it converges to some $z \in X$. Note that since T is continuous, $z \in \mathrm{Fix}\, T$. Moreover, $d(Tx,z) \leq \frac{k}{1-k} d(x,Tx)$ and so

$$\mathrm{dist}(x, \mathrm{Fix}\, T) \leq d(x,z) \leq d(x,Tx) + d(Tx,z) \leq \frac{1}{1-k} d(x,Tx) < \varepsilon.$$

We include below an example of a continuous orbital contraction which has more than one fixed point and refer to [54] for a more detailed discussion on orbital contractions.

Let $X = \{(x,y) \in \mathbb{R}^2 : 0 \leq x \leq 1, 0 \leq y \leq 1-x\}$ with the usual Euclidean distance. Define $f : X \to X$ by

$$f(x,y) = \left(x, \frac{y+1-x}{2} \right).$$

Then f is continuous, $\|f^2(x,y) - f(x,y)\| = \|f(x,y) - (x,y)\|/2$ for all $(x,y) \in X$ and $\mathrm{Fix}\, T = \{(x, 1-x) : x \in [0,1]\}$.

(iii) Let X be a metric space and $T : X \to C \subseteq X$ be a retraction. Then $\phi(\varepsilon) = \varepsilon$ is a modulus of regularity for T w.r.t. $\mathrm{Fix}\, T$. To see this, note that $\mathrm{Fix}\, T = T(X) = C$ and $d(x, Tx) < \varepsilon$ implies $\mathrm{dist}(x, \mathrm{Fix}\, T) < \varepsilon$ since $Tx \in \mathrm{Fix}\, T$ (not even the continuity of T is needed for this). In particular, this applies to the case where T is the metric projection of X onto C if the metric projection exists as a single-valued function.

(iv) For nonempty, closed and convex subsets $C_1, C_2 \subseteq \mathbb{R}^n$ consider

$$T = R_{N_{C_2}} R_{N_{C_1}}.$$

In [18, p. 18] it is shown that if C_1, C_2 are convex semi-algebraic sets with $O \in C_1 \cap C_2$ which can be described by polynomials on $\mathbb{R}^n$ of degree greater than 1, then (in our terminology), given $r > 0$, there exist $\mu > 0$ and $\gamma \geq 1$ such that T admits the following modulus of regularity w.r.t. $\mathrm{Fix}\, T$ and $\overline{B}(O, r)$

$$\phi(\varepsilon) = 2(\varepsilon/\mu)^\gamma.$$

Minimization problems

Let (X, d) metric space and $f : X \to (-\infty, \infty]$. We consider the problem

$$\operatorname*{argmin}_{x \in X} f(x). \tag{3.10}$$

Suppose that its set of solutions S is nonempty and denote $m = \min_{x\in X} f(x)$. Define the function $F : X \to \overline{\mathbb{R}}$, $F(x) = f(x) - m$. Note that $\mathrm{zer}\, F = S$ and $F(x) = \infty$ for $x \notin \mathrm{dom}\, f$.

Given $z \in S$ and $r > 0$, a *modulus of regularity* for f w.r.t. S and $\overline{B}(z, r)$ is a modulus of regularity for F w.r.t. $\mathrm{zer}\, F$ and $\overline{B}(z, r)$. Similarly, a *modulus of regularity* for f w.r.t. S is a modulus of regularity for F w.r.t. $\mathrm{zer}\, F$. This concept is closely related to growth conditions for the function f such as the notions of sets of weak sharp minima or error bounds (see, e.g., [22, 27, 21, 16, 45] with the remark that there is a vast literature on these topics and their connection to other regularity properties). These conditions are especially used to analyze the linear convergence or the finite termination of central algorithms in optimization.

In the following S stands as above for the set of solutions of problem (3.10).

Example 3.7. (i) The set S is called a set of *ψ-global weak sharp minima* for f if

$$f(x) \geq m + \psi(\mathrm{dist}(x, S)), \tag{3.11}$$

for all $x \in X$, where $\psi : [0, \infty) \to [0, \infty)$ is a strictly increasing function satisfying $\psi(0) = 0$. In this case, $\phi : (0, \infty) \to (0, \infty)$, $\phi(\varepsilon) = \psi(\varepsilon)$, acts as a modulus of regularity for f w.r.t. S. The case $\psi(\varepsilon) = k\, \varepsilon$ with $k > 0$ was introduced in [22].

(ii) More generally, one can assume that S is a set of *ψ-boundedly weak sharp minima* for f, that is, for any bounded set $C \subseteq X$ with $C \cap S \neq \emptyset$, there exists a strictly increasing function $\psi = \psi_C : [0, \infty) \to [0, \infty)$ satisfying $\psi(0) = 0$ such that (3.11) holds for all $x \in C$. Fixing $z \in S$ and $r > 0$, a modulus of regularity for f w.r.t. S and $\overline{B}(z, r)$ can be defined by $\phi : (0, \infty) \to (0, \infty)$, $\phi(\varepsilon) = \psi_C(\varepsilon)$, where $C = \overline{B}(z, r)$.

Remark 3.8. In this regard, if ω is an increasing function satisfying $\omega(0) = 0$, an inequality of the form

$$\omega(f(x) - m) \geq \mathrm{dist}(x, S),$$

where x is contained either in X or in a bounded set, is also called an *error bound* (see [16] and the references therein).

Zeros of set-valued operators

Let X and Y be normed spaces and $A : X \to 2^Y$ be a set-valued operator such that zer $A \neq \emptyset$ and

$$\mathrm{dist}(O_Y, A(x)) = 0 \;\Rightarrow\; x \in \mathrm{zer}\ A, \tag{3.12}$$

for all $x \in X$. If $F : X \to \overline{\mathbb{R}}$ is defined by $F(x) = \mathrm{dist}(O_Y, A(x))$, then zer $F =$ zer A and $F(x) = \infty$ for $x \notin$ dom A. Note that if H is a Hilbert space and $A : H \to 2^H$ is maximal monotone, then $A(x)$ is closed for all $x \in H$, so (3.12) holds.

Given $z \in$ zer A and $r > 0$, a *modulus of regularity* for A w.r.t. zer A and $\overline{B}(z,r)$ is a modulus of regularity for F w.r.t. zer F and $\overline{B}(z,r)$. Similarly, a *modulus of regularity* for A w.r.t. zer A is a modulus of regularity for F w.r.t. zer F. We give next two instances when moduli of regularity for A w.r.t. zer A exist.

Example 3.9. (i) Let X be a Banach space and X^* its dual. The normalized duality mapping $J : X \to 2^{X^*}$ is defined as

$$J(x) = \{j \in X^* : j(x) = \|x\|^2, \|j\| = \|x\|\}.$$

An operator $A : X \to 2^X$ is called *ψ-strongly accretive*, where $\psi : [0,\infty) \to [0,\infty)$ is a strictly increasing function with $\psi(0) = 0$, if

$$\langle x^* - y^*, x - y\rangle_+ \geq \psi(\|x-y\|)\|x-y\|, \tag{3.13}$$

for all $x, y \in X$, $x^* \in A(x)$, $y^* \in A(y)$, where $\langle v, u\rangle_+ = \max\{j(v) : j \in J(u)\}$.

Assume that zer $A \neq \emptyset$ (hence it is a singleton) and let $x \in X, x \notin$ zer A. Taking in (3.13) $y \in$ zer A, we obtain

$$\|x^*\| \geq \frac{\langle x^*, x-y\rangle_+}{\|x-y\|} \geq \psi(\|x-y\|),$$

for all $x^* \in A(x)$. Then $\mathrm{dist}(O_X, A(x)) \geq \psi(\|x-y\|) > 0$, hence A satisfies (3.12) and $\phi : (0,\infty) \to (0,\infty)$, $\phi(\varepsilon) = \psi(\varepsilon)$, is a modulus of regularity for A w.r.t. zer A. Furthermore, it is actually a modulus of uniqueness, a fact that was also observed in [39, Remark 2]. If A is single-valued, then for any $\gamma > 0$, $\gamma\phi$ is a modulus of regularity for $\mathrm{Id} - \gamma A$ w.r.t. Fix $(\mathrm{Id} - \gamma A)$.

(ii) Metric subregularity has been extensively used in optimization in relation with stability problems and the linear local convergence of proximal point methods (see [26, 44]). An operator $A : X \to 2^Y$ is called *metrically subregular* at $z \in$ zer A for O_Y if there exist $k, r > 0$ such that

$$\mathrm{dist}(x, \mathrm{zer}\ A) \leq k\,\mathrm{dist}(O_Y, A(x)),$$

for all $x \in \overline{B}(z,r)$. In this case, if zer A is closed, (3.12) holds and so $\phi : (0,\infty) \to (0,\infty)$, $\phi(\varepsilon) = \varepsilon/k$, is a modulus of regularity for A w.r.t. zer A and $\overline{B}(z,r)$.

Recall that if (X_1, d_1) and (X_2, d_2) are metric spaces, a *modulus of uniform continuity* for a uniformly continuous mapping $T : X_1 \to X_2$ is a function $\rho : (0,\infty) \to (0,\infty)$ satisfying

$$\forall \varepsilon > 0\, \forall x, y \in X_1\ \left(d_1(x,y) < \rho(\varepsilon) \Rightarrow d_2(Tx, Ty) < \varepsilon\right).$$

We give below a quantitative version of the identity

$$\mathrm{argmin}\ f = \mathrm{Fix}\ J_{\gamma\partial f} = \mathrm{zer}\ \partial f$$

that holds in a Hilbert space H for a proper, convex and lower semi-continuous function $f : H \to (-\infty, \infty]$ that has minimizers and for any $\gamma > 0$. Besides, in the proof we actually convert the respective moduli of regularity into each other.

Theorem 3.10. *Let H be a Hilbert space and $f : H \to (-\infty, \infty]$ a proper, convex and lower semi-continuous function which attains its minimum. Take $z \in \operatorname{argmin} f$ and $r, r' > 0$. Consider the following statements:*

1. *The function f admits a modulus of regularity w.r.t. $\operatorname{argmin} f$ and $\overline{B}(z, r)$.*

2. *For $\gamma > 0$, the resolvent of f, $J_{\gamma\partial f}$, admits a modulus of regularity w.r.t. $\operatorname{Fix} J_{\gamma\partial f}$ and $\overline{B}(z, r)$.*

3. *The subdifferential of f, ∂f, admits a modulus of regularity w.r.t. $\operatorname{zer} \partial f$ and $\overline{B}(z, r')$.*

Then

(i) If $f|_{\overline{B}(z,r+1)}$ is additionally admitting a modulus of uniform continuity, then 1 implies 2 for all $\gamma > 0$.

(ii) If there exists $\gamma > 0$ such that 2 holds, then 1 is satisfied. Moreover, 3 holds too if $r' < r$.

(iii) If ∂f is single-valued, $r' = r$ and $(\mathrm{Id} + \gamma\partial f)|_{\overline{B}(z,r+1)}$, $\gamma > 0$, is admitting a modulus of uniform continuity, then 3 implies 2.

Proof. Recall first that

$$\operatorname{argmin} \ f = \operatorname{Fix} J_{\gamma\partial f} = \operatorname{zer} \partial f,$$

for every $\gamma > 0$.

(i) Let ϕ be a modulus of regularity for f w.r.t. $\operatorname{argmin} f$ and $\overline{B}(z, r)$, and ρ a modulus of uniform continuity for $f|_{\overline{B}(z,r+1)}$. Fix $\gamma > 0$. Define $\phi^* : (0, \infty) \to (0, \infty)$ by

$$\phi^*(\varepsilon) = \min \left\{ \rho\left(\frac{\phi(\varepsilon)}{2}\right), \frac{\gamma\phi(\varepsilon)}{2r}, 1 \right\}.$$

To see that ϕ^* is a modulus of regularity for $J_{\gamma\partial f}$ w.r.t. $\operatorname{Fix} J_{\gamma\partial f}$ and $\overline{B}(z, r)$, let $\varepsilon > 0$ and $x \in \overline{B}(z, r)$. Assume $\|x - J_{\gamma\partial f}(x)\| < \phi^*(\varepsilon)$. Then $J_{\gamma\partial f}(x) \in \overline{B}(z, r+1)$, so $f(x) - f(J_{\gamma\partial f}(x)) < \phi(\varepsilon)/2$. At the same time, since $z \in \operatorname{argmin} \ f$, by (2.5),

$$f(J_{\gamma\partial f}(x)) - m \leq \frac{1}{2\gamma} \left(\|z - x\|^2 - \|J_{\gamma\partial f}(x) - x\|^2 - \|J_{\gamma\partial f}(x) - z\|^2 \right).$$

Because

$$\|z - x\|^2 - 2\|z - x\|\|J_{\gamma\partial f}(x) - x\| + \|J_{\gamma\partial f}(x) - x\|^2 = (\|z - x\| - \|J_{\gamma\partial f}(x) - x\|)^2 \leq \|J_{\gamma\partial f}(x) - z\|^2,$$

we have

$$\begin{aligned}
\|z - x\|^2 - \|J_{\gamma\partial f}(x) - x\|^2 - \|J_{\gamma\partial f}(x) - z\|^2 &\leq 2\|J_{\gamma\partial f}(x) - x\| \left(\|z - x\| - \|J_{\gamma\partial f}(x) - x\| \right) \\
&\leq 2\|J_{\gamma\partial f}(x) - x\|\|J_{\gamma\partial f}(x) - z\| = 2\|J_{\gamma\partial f}(x) - x\|\|J_{\gamma\partial f}(x) - J_{\gamma\partial f}(z)\| \\
&\leq 2\|J_{\gamma\partial f}(x) - x\|\|x - z\| < \gamma\phi(\varepsilon).
\end{aligned}$$

Thus, $f(J_{\gamma\partial f}(x)) - m < \phi(\varepsilon)/2$, from where

$$f(x) - m = f(x) - f(J_{\gamma\partial f}(x)) + f(J_{\gamma\partial f}(x)) - m < \phi(\varepsilon).$$

Consequently, $\operatorname{dist}(x, \operatorname{Fix} J_{\gamma\partial f}) < \varepsilon$.

(ii) Let $\gamma > 0$ and ϕ be a modulus of regularity for $J_{\gamma\partial f} = \operatorname{Prox}_{\gamma f}$ w.r.t. $\operatorname{Fix} J_{\gamma\partial f}$ and $\overline{B}(z, r)$.

We prove first that $\phi^* : (0,\infty) \to (0,\infty)$,

$$\phi^*(\varepsilon) = \frac{\phi(\varepsilon)^2}{2\gamma},$$

is a modulus of regularity for f w.r.t. $\operatorname{argmin} f$ and $\overline{B}(z,r)$. To see this, let $\varepsilon > 0$ and $x \in \overline{B}(z,r)$ such that $f(x) - m < \phi^*(\varepsilon)$. Since

$$f(J_{\gamma\partial f}(x)) + \frac{1}{2\gamma}\|J_{\gamma\partial f}(x) - x\|^2 \leq f(x),$$

we get

$$\|J_{\gamma\partial f}(x) - x\|^2 \leq 2\gamma(f(x) - f(J_{\gamma\partial f}(x))) \leq 2\gamma(f(x) - m).$$

Thus, $\|J_{\gamma\partial f}(x) - x\| < \phi(\varepsilon)$, which yields $\operatorname{dist}(x, \operatorname{argmin} f) < \varepsilon$.

Define now $\phi^* : (0,\infty) \to (0,\infty)$ by

$$\phi^*(\varepsilon) = \frac{1}{\gamma}\min\left\{\phi\left(\frac{\varepsilon}{2}\right), \frac{\varepsilon}{2}, r - r'\right\}.$$

We show that ϕ^* is a modulus of regularity for ∂f w.r.t. $\operatorname{zer} \partial f$ and $\overline{B}(z,r')$. Let $\varepsilon > 0$ and $x \in \overline{B}(z,r')$. Suppose $\operatorname{dist}(O_H, \partial f(x)) < \phi^*(\varepsilon)$ and choose $y \in \partial f(x)$ such that $\|y\| < \phi^*(\varepsilon)$. Then $x + \gamma y \in (\mathrm{Id} + \gamma\partial f)(x)$, so $J_{\gamma\partial f}(x + \gamma y) = x$,

$$\|x + \gamma y - z\| \leq \|x - z\| + \gamma\|y\| \leq r' + r - r' = r,$$

and

$$\|J_{\gamma\partial f}(x + \gamma y) - (x + \gamma y)\| = \gamma\|y\| < \phi\left(\frac{\varepsilon}{2}\right).$$

It follows that $\operatorname{dist}(x + \gamma y, \operatorname{zer} \partial f) < \varepsilon/2$, hence

$$\operatorname{dist}(x, \operatorname{zer} \partial f) \leq \operatorname{dist}(x + \gamma y, \operatorname{zer} \partial f) + \gamma\|y\| < \varepsilon.$$

(iii) Note that in this case $\mathrm{Id} + \gamma\partial f : H \to H$ is also bijective (see [12, Chapter 23]). Let ρ be a modulus of uniform continuity for $(\mathrm{Id} + \gamma\partial f)|_{\overline{B}(z,r+1)}$ and ϕ a modulus of regularity for ∂f w.r.t. $\operatorname{zer} \partial f$ and $\overline{B}(z,r)$, and define $\phi^* : (0,\infty) \to (0,\infty)$ by

$$\phi^*(\varepsilon) = \min\left\{\rho(\gamma\phi(\varepsilon)), 1\right\}.$$

Let $\varepsilon > 0$ and $x \in \overline{B}(z,r)$ such that $\|J_{\gamma\partial f}(x) - x\| < \phi^*(\varepsilon)$. Note that $J_{\gamma\partial f}(x) \in \overline{B}(z,r+1)$. Because

$$\|\partial f(x)\| = \frac{1}{\gamma}\|\gamma\partial f(x)\| = \frac{1}{\gamma}\|(\mathrm{Id} + \gamma\partial f)(J_{\gamma\partial f}(x)) - (\mathrm{Id} + \gamma\partial f)(x)\| < \phi(\varepsilon),$$

it follows that $\operatorname{dist}(x, \operatorname{Fix} J_{\gamma\partial f}) < \varepsilon$. $\square$

4 Rates of convergence

The next result shows that having a function $F : X \to \overline{\mathbb{R}}$ that possesses zeros and a sequence $(x_n) \subseteq X$ that is Fejér monotone w.r.t. $\operatorname{zer} F$, a modulus of regularity for F yields a rate of convergence for (x_n) to a zero of F if (x_n) has approximate zeros for F with an approximate zero bound.

Theorem 4.1. *Let (X,d) be a metric space and $F : X \to \overline{\mathbb{R}}$ with* $\mathrm{zer}\ F \neq \emptyset$*. Suppose that (x_n) is a sequence in X which is Fejér monotone w.r.t.* $\mathrm{zer}\ F$*, $b > 0$ is an upper bound on $d(x_0, z)$ for some $z \in \mathrm{zer}\ F$ and there exists $\alpha : (0,\infty) \to \mathbb{N}$ such that*

$$\forall \varepsilon > 0\, \exists n \leq \alpha(\varepsilon)\ (|F(x_n)| < \varepsilon)\,.$$

If ϕ is a modulus of regularity for F w.r.t. $\mathrm{zer}\ F$ *and $\overline{B}(z,b)$, then (x_n) is a Cauchy sequence with Cauchy modulus*

$$\forall \varepsilon > 0\, \forall n, \tilde{n} \geq \alpha(\phi(\varepsilon/2))\ (d(x_n, x_{\tilde{n}}) < \varepsilon) \tag{4.14}$$

and

$$\forall \varepsilon > 0\, \forall n \geq \alpha(\phi(\varepsilon))\ (\mathrm{dist}(x_n, \mathrm{zer}\ F) < \varepsilon)\,. \tag{4.15}$$

Moreover,

(i) if X is complete and $\mathrm{zer}\ F$ *is closed, then (x_n) converges to a zero of F with a rate of convergence $\alpha(\phi(\varepsilon/2))$;*

(ii) if there exists $\varepsilon^* > 0$ such that

$$\forall w \in \mathbb{R},\ |w| < \varepsilon^*\ \left(F^{-1}(w) \subseteq \mathrm{zer}\ F \cup \{x \in X : \mathrm{dist}(x, \mathrm{zer}\ F) \geq \varepsilon^*\}\right), \tag{4.16}$$

then $x_n = z'$ for all $n \geq \alpha\left(\min\{\varepsilon^*, \phi(\varepsilon^*)\}\right)$, where $z' \in \mathrm{zer}\ F$.

Proof. Let $\varepsilon > 0$. Note that by Fejér monotonicity, $(x_n) \subseteq \overline{B}(z,b)$. Since there exists $N \leq \alpha(\phi(\varepsilon/2))$ such that $|F(x_N)| < \phi(\varepsilon/2)$, it follows that $\mathrm{dist}(x_N, \mathrm{zer}\ F) < \varepsilon/2$. Thus, $d(x_N, y) < \varepsilon/2$ for some $y \in \mathrm{zer}\ F$. Since (x_n) is Fejér monotone w.r.t. $\mathrm{zer}\ F$, this implies that $d(x_n, y) \leq d(x_N, y) < \varepsilon/2$ for all $n \geq \alpha(\phi(\varepsilon/2))$, so (4.14) and (4.15) hold.

(i) If X is complete, then $z' = \lim_{n\to\infty} x_n$ exists and, by the above Cauchy rate, we get that $d(x_m, z') \leq \varepsilon$ for $m \geq \alpha(\phi(\varepsilon/2))$. Hence,

$$\mathrm{dist}(z', \mathrm{zer}\ F) \leq \mathrm{dist}(x_m, \mathrm{zer}\ F) + d(z', x_m) < 3\varepsilon/2.$$

Since $\varepsilon > 0$ was arbitrary, we get $\mathrm{dist}(z', \mathrm{zer}\ F) = 0$ which yields, if $\mathrm{zer}\ F$ is closed, that $z' \in \mathrm{zer}\ F$.

(ii) Let $N \leq \alpha\left(\min\{\varepsilon^*, \phi(\varepsilon^*)\}\right)$ such that $|F(x_N)| < \min\{\varepsilon^*, \phi(\varepsilon^*)\}$. Taking $w = F(x_N)$ in (4.16), we obtain $x_N \in \mathrm{zer}\ F \cup \{x \in X : \mathrm{dist}(x, \mathrm{zer}\ F) \geq \varepsilon^*\}$. However, as $|F(x_N)| < \phi(\varepsilon^*)$, we have $\mathrm{dist}(x_N, \mathrm{zer}\ F) < \varepsilon^*$, and so $x_N = z'$ for some $z' \in \mathrm{zer}\ F$. But then, by Fejér monotonicity, $d(x_n, z') \leq d(x_N, z') = 0$ for all $n \geq N$. Hence, $x_n = z'$ for all $n \geq \alpha\left(\min\{\varepsilon^*, \phi(\varepsilon^*)\}\right)$. □

Remark 4.2. If instead of (4.16) one actually has the stronger condition

$$\exists \varepsilon^* > 0\, \forall w \in \mathbb{R},\ |w| < \varepsilon^*\ \left(F^{-1}(w) \subseteq \mathrm{zer}\ F\right), \tag{4.17}$$

then one does not need to assume the existence of a modulus of regularity for F in order to obtain the finite convergence of (x_n). In this case the corresponding rate is $\alpha(\varepsilon^*)$.

Remark 4.3. In fact in order to obtain finite termination, condition (4.16) does not need to hold for all points in $F^{-1}(w)$, but only for those also belonging to the set of values of (x_n).

The following result is in some sense a converse of Theorem 4.1 for a particular situation.

Proposition 4.4. Let X be a metric space, $T : X \to X$ nonexpansive with $\mathrm{Fix}\ T \neq \emptyset$, $z \in \mathrm{Fix}\ T$ and $b > 0$. If for every $x \in \overline{B}(z,b)$, $(T^n x)$ converges to a fixed point of T with a common rate of convergence ψ, then $\phi(\varepsilon) = \varepsilon/(2\psi(\varepsilon/2))$ is a modulus of regularity for T w.r.t. $\mathrm{Fix}\ T$ and $\overline{B}(z,b)$.

Proof. Assume on the contrary that ϕ is not a modulus of regularity for T w.r.t. Fix T and $\overline{B}(z,b)$. Then there exist $\varepsilon > 0$ and $x \in \overline{B}(z,b)$ such that $d(x,Tx) < \phi(\varepsilon)$ and $\mathrm{dist}(x, \mathrm{Fix}\ T) \geq \varepsilon$. Denote $n = \psi(\varepsilon/2)$. Then there exists $w \in \mathrm{Fix}\ T$ such that $d(T^n x, w) < \varepsilon/2$, which yields

$$\mathrm{dist}(x, \mathrm{Fix}\ T) \leq d(x,w) \leq d(w, T^n x) + \sum_{i=0}^{n-1} d(T^i x, T^{i+1}x) < \frac{\varepsilon}{2} + nd(x,Tx) < \frac{\varepsilon}{2} + n\phi(\varepsilon) = \varepsilon,$$

a contradiction. $\square$

The next result provides conditions that are sufficient for obtaining linear convergence of the sequence (x_n).

Theorem 4.5. *Let (X,d) be a metric space, $F : X \to \overline{\mathbb{R}}$ with* zer $F \neq \emptyset$*, (x_n) a sequence in X and $b > 0$ an upper bound on $d(x_0, z)$ for some $z \in$* zer F*. Suppose that there exist $p, \beta > 0$ such that*

$$\forall y \in \mathrm{zer}\ F\, \forall n \in \mathbb{N}\ (d(x_{n+1}, y)^p \leq d(x_n, y)^p - \beta |F(x_n)|^p). \tag{4.18}$$

Assume also that $\phi : (0,\infty) \to (0,\infty)$, $\phi(\varepsilon) = t\varepsilon$, where $t > 0$, is a modulus of regularity for F w.r.t. zer F *and $\overline{B}(z,b)$. Denote*

$$c = (1 + \beta t^p)^{-1/p} \in (0,1).$$

Then

$$\forall n \in \mathbb{N}\ (\mathrm{dist}(x_n, \mathrm{zer}\ F) \leq c^n \mathrm{dist}(x_0, \mathrm{zer}\ F)). \tag{4.19}$$

In addition, (x_n) is a Cauchy sequence and satisfies

$$\forall k, n \in \mathbb{N}\ (d(x_n, x_{n+k}) \leq 2c^n \mathrm{dist}(x_0, \mathrm{zer}\ F)).$$

Furthermore, if X is complete and zer F *is closed, then (x_n) converges linearly to some $z' \in$* zer F *and*

$$\forall n \in \mathbb{N}\ \left(d(x_n, z') \leq 2c^n \mathrm{dist}(x_0, \mathrm{zer}\ F)\right).$$

Proof. Let $y \in \mathrm{zer}\ F$. Note first that $(x_n) \subseteq \overline{B}(z,b)$ by (4.18). In addition, using the expression of ϕ and Remark 3.2,

$$|F(x)| \geq t\mathrm{dist}(x, \mathrm{zer}\ F), \tag{4.20}$$

for all $x \in \overline{B}(z,b)$. Then

$$\begin{aligned} \mathrm{dist}(x_{n+1}, \mathrm{zer}\ F)^p &\leq d(x_{n+1}, y)^p \leq d(x_n, y)^p - \beta |F(x_n)|^p \quad \text{by (4.18)} \\ &\leq d(x_n, y)^p - \beta t^p \mathrm{dist}(x_n, \mathrm{zer}\ F)^p \quad \text{by (4.20)} \\ &\leq d(x_n, y)^p - \beta t^p \mathrm{dist}(x_{n+1}, \mathrm{zer}\ F)^p \quad \text{by (4.18)}. \end{aligned}$$

Hence, $\mathrm{dist}(x_{n+1}, \mathrm{zer}\ F) \leq cd(x_n, y)$ for all $y \in \mathrm{zer}\ F$ and $n \in \mathbb{N}$, which shows that

$$\mathrm{dist}(x_{n+1}, \mathrm{zer}\ F) \leq c\mathrm{dist}(x_n, \mathrm{zer}\ F),$$

from where (4.19) follows. Let $y \in \mathrm{zer}\ F$ and $k, n \in \mathbb{N}$. Then

$$d(x_n, x_{n+k}) \leq d(x_{n+k}, y) + d(x_n, y) \leq 2d(x_n, y),$$

so

$$d(x_n, x_{n+k}) \leq 2\mathrm{dist}(x_n, \mathrm{zer}\ F) \leq 2c^n \mathrm{dist}(x_0, \mathrm{zer}\ F).$$

Suppose now that X is complete and zer F is closed. As (x_n) is Cauchy, it converges to some $z' \in X$. Using (4.19), we have $\mathrm{dist}(z', \mathrm{zer}\ F) = 0$ and since zer F is closed, we obtain $z' \in \mathrm{zer}\ F$. $\square$

In order to apply Theorem 4.1 to obtain rates of convergence for sequences (x_n) converging to common fixed points of finitely many self-mappings $T_1, \dots, T_n : X \to X$ of a metric space (X, d) (which e.g. is the situation for iterative procedures to solve so-called convex feasibility problems, see below) we first define the following notion.

Definition 4.6 (compare [10])**.** Let (X, d) be a metric space and $C_1, \dots, C_m, K$ be subsets of X with $C = \bigcap_{i=1}^m C_i \neq \emptyset$. We say that $C_1, \dots, C_m$ are *regular w.r.t.* K if

$$\forall \varepsilon > 0\, \exists \delta > 0\, \forall x \in K \left(\max_{i=1,\dots,m} \mathrm{dist}(x, C_i) < \delta \Rightarrow \mathrm{dist}(x, C) < \varepsilon \right).$$

We call a function $\rho : (0, \infty) \to (0, \infty)$ producing such a $\delta = \rho(\varepsilon)$ a *modulus of regularity* for $C_1, \dots, C_m$ w.r.t. K.

We say that $C_1, \dots, C_m$ are *linearly regular w.r.t.* K if

$$\exists \tau > 0\, \forall x \in K \left(\mathrm{dist}(x, C) \leq \tau \max_{i=1,\dots,m} \mathrm{dist}(x, C_i) \right). \tag{4.21}$$

Example 4.7 ([17])**.** From a result shown in [17], the following is immediate: let $C_1, \dots, C_m \subseteq \mathbb{R}^n$ be basic convex semi-algebraic sets given by

$$C_i = \{x \in \mathbb{R}^n | g_{i,j}(x) \leq 0,\, j = 1, \dots, m_i\},$$

where $g_{i,j}$ are convex polynomials on $\mathbb{R}^n$ with degree at most $d \in \mathbb{N}$. Suppose that $\bigcap_{i=1}^m C_i \neq \emptyset$. Then for any compact $K \subseteq \mathbb{R}^n$ there exists $c > 0$ such that

$$\rho(\varepsilon) = (\varepsilon/c)^\gamma/m, \text{ with } \gamma = \min\left\{ \frac{(2d-1)^n + 1}{2}, B(n-1)d^n \right\},$$

where $B(n) = \binom{n}{[n/2]}$, is a modulus of regularity for $C_1, \dots, C_m$ w.r.t. K.

As an easy consequence of Theorem 4.1 we obtain the next result.

Corollary 4.8. Let (X, d) be a complete metric space and $T_1, \dots, T_m$ be self-mappings of X with $C = \bigcap_{i=1}^m \mathrm{Fix}\, T_i$ nonempty and closed. Let (x_n) be a sequence in X which is Fejér monotone w.r.t. C and assume that $b > 0$ is an upper bound on $d(x_0, z)$ for some $z \in C$. Suppose that ϕ is a common modulus of regularity for T_i w.r.t. $\mathrm{Fix}\, T_i$ and $\overline{B}(z, b)$ for each $i = 1, \dots, m$. If $\mathrm{Fix}\, T_1, \dots, \mathrm{Fix}\, T_m$ are regular w.r.t. $\overline{B}(z, b)$ with modulus ρ and (x_n) also has common approximate fixed points for $T_1, \dots, T_m$ with $\alpha : (0, \infty) \to \mathbb{N}$ a common approximate fixed point bound, i.e.,

$$\forall \varepsilon > 0\, \exists n \leq \alpha(\varepsilon) \left(\max_{i=1,\dots,m} d(x_n, T_i x_n) < \varepsilon \right),$$

then (x_n) converges to a point in C with a rate of convergence $\alpha(\phi(\rho(\varepsilon/2)))$.

Proof. Denote $F : X \to \mathbb{R}$,

$$F(x) = \max_{i=1,\dots,m} d(x, T_i(x)).$$

Then $x \in \mathrm{zer}\ F$ if and only if $d(x, T_i x) = 0$ for all $i \in \{1, \dots, m\}$, so $\mathrm{zer}\ F = C$. Consider $\varepsilon > 0$ and $x \in \overline{B}(z, b)$ with $F(x) < \phi(\rho(\varepsilon))$. Then $d(x, T_i x) < \phi(\rho(\varepsilon))$ for all $i \in \{1, \dots, m\}$, so $\mathrm{dist}(x, \mathrm{Fix}\ T_i) < \rho(\varepsilon)$ for all $i \in \{1, \dots, m\}$. Since the sets $\mathrm{Fix}\ T_1, \dots, \mathrm{Fix}\ T_m$ are regular w.r.t. $\overline{B}(z, b)$ with modulus ρ, we get $\mathrm{dist}(x, \mathrm{zer}\ F) < \varepsilon$. Thus, $\phi \circ \rho$ is a modulus of regularity for F w.r.t. $\mathrm{zer}\ F$ and $\overline{B}(z, b)$. The result follows now from Theorem 4.1. $\square$

Recall that for retractions $T_i : X \to C_i (= \mathrm{Fix}\ T_i)$, $\phi(\varepsilon) = \varepsilon$ is a modulus of regularity for T_i w.r.t. Fix T_i. Consequently, we get the next result.

Corollary 4.9. Let $C_1, \dots, C_m \subseteq X$ be subsets of a complete metric space (X, d) with $C = \bigcap_{i=1}^m C_i$ nonempty and closed, and $T_i : X \to C_i$, $i = 1, \dots, m$, be retractions. Then under the assumptions on (x_n) and on the regularity of $C_1, \dots, C_m$ from Corollary 4.8, one has $\alpha(\rho(\varepsilon/2))$ as a rate of convergence for (x_n) to some point in C.

Note that Corollary 4.9 applies in particular to so-called convex feasibility problems which we consider in the setting of a complete CAT(0) space X (later we will also focus on CAT(κ) spaces with $\kappa > 0$ and an appropriate upper bound on its diameter). More precisely, if $C_1, \dots, C_m \subseteq X$ are nonempty, closed and convex with $C = \bigcap_{i=1}^m C_i \neq \emptyset$, the *convex feasibility problem* (CFP) consists in finding a point in C. Then we may apply Corollary 4.9 with T_i being the metric projection onto C_i and (x_n) can generated by the *cyclic projection method*, i.e., given $x_0 \in X$, $x_{n+1} = T_{\bar{n}} x_n$ for all $n \in \mathbb{N}$, where $T_{\bar{n}} = T_{n(\text{mod } m)+1}$. Due to the nonexpansivity of the metric projection, the sequence (x_n) is Fejér monotone w.r.t. C and an explicit common approximate fixed point bound can be obtained from [4, Theorem 3.2, Remark 3.1].

We study next the finite convergence of a sequence to a zero of a maximal monotone operator. To this end we assume condition (4.22) which was considered by Rockafellar in [57, Theorem 3] (see also [50]) to show that the proximal point algorithm terminates in finitely many iterations.

Corollary 4.10. Let H be a Hilbert space and $A : H \to 2^H$ maximal monotone such that

$$\exists z \in H\ \exists \varepsilon^* > 0\ \ (B(O, \varepsilon^*) \subseteq A(z)). \tag{4.22}$$

If (x_n) is a sequence in H that is a Fejér monotone w.r.t. zer A and there exists $\alpha : (0, \infty) \to \mathbb{N}$ such that

$$\forall \varepsilon > 0\, \exists n \leq \alpha(\varepsilon)\ (\mathrm{dist}(O, A(x_n)) < \varepsilon),$$

then zer $A = \{z\}$ and $x_n = z$ for all $n \geq \alpha(\varepsilon^*)$.

Proof. Define $F : H \to \overline{\mathbb{R}}$, $F(x) = \mathrm{dist}(O, A(x))$. We show that F satisfies (4.17). For $w \in \mathbb{R}$, $0 \leq w < \varepsilon^*$ and $x \in F^{-1}(w)$, we have $\mathrm{dist}(O, A(x)) = w < \varepsilon^*$, so there exists $u \in A(x)$ such that $\|u\| < \varepsilon^*$. Assume that $x \neq z$ and define

$$v_n = \frac{\varepsilon^*}{1 + 1/n} \frac{x - z}{\|x - z\|},\ n \in \mathbb{N} \setminus \{0\}.$$

Note that $\|v_n\| < \varepsilon^*$, so $v_n \in A(z)$. By the monotonicity of A, $\langle x - z, v_n \rangle \leq \langle x - z, u \rangle$ for all $n \in \mathbb{N} \setminus \{0\}$, which yields $\varepsilon^* \|x - z\| \leq \|u\| \|x - z\|$. This is a contradiction, so $F^{-1}(w) = \{z\}$, which shows in particular that zer $A = \{z\}$. By Remark 4.2, $x_n = z$ for all $n \geq \alpha(\varepsilon^*)$. ☐

We apply in the following the above results to different algorithms.

Picard iteration

Let X be a complete metric space and $T : X \to X$ a quasi-nonexpansive mapping with Fix $T \neq \emptyset$. The *Picard iteration* generates starting from $x_0 \in X$ the sequence given by

$$x_{n+1} = T x_n \quad \text{for any } n \in \mathbb{N}. \tag{4.23}$$

It follows directly from the definition of a quasi-nonexpansive mapping that (x_n) is Fejér monotone w.r.t. Fix T. Moreover, Fix T is closed. Notice also that if T is nonexpansive, a function $\alpha : (0, \infty) \to \mathbb{N}$ such that

$$\forall \varepsilon > 0 \, \exists n \leq \alpha(\varepsilon) \; (d(x_n, Tx_n) < \varepsilon)$$

is actually a rate of asymptotic regularity for (x_n) as the sequence $(d(x_n, Tx_n))$ is nonincreasing.

Let $b > 0$ be an upper bound on $d(x_0, z)$ for some $z \in \text{Fix } T$. By Fejér monotonicity, $(x_n) \subseteq \overline{B}(z, b)$. Considering $F : X \to \mathbb{R}, F(x) = d(x, Tx)$, if ϕ is a modulus of regularity for F w.r.t. zer F and $\overline{B}(z, b)$, and α is a rate of asymptotic regularity for (x_n), then, applying Theorem 4.1, we can deduce that (x_n) converges to a fixed point of T with a rate of convergence $\alpha(\phi(\varepsilon/2))$. In what follows we consider two problems where such α and ϕ can be computed explicitly.

First we focus on the problem of minimizing the distance between two nonintersecting sets applying the alternating projection method. Suppose in the sequel that X is a complete CAT(0) space and $U, V \subseteq X$ are nonempty, closed and convex with $U \cap V = \emptyset$. We aim to find best approximation pairs $(u, v) \in U \times V$ such that $d(u, v) = \text{dist}(U, V)$. This problem was studied in [7, 4] (for results in Hilbert spaces, see, e.g., [11, 42]). Denote $\rho = \text{dist}(U, V)$ and suppose that $S = \{(u, v) \in U \times V : d(u, v) = \rho\} \neq \emptyset$.

Given $x_0 \in X$, consider the sequence (x_n) defined by (4.23), where $T : X \to X$, $T = P_U \circ P_V$. Then T is nonexpansive and if $(u, v) \in S$, then $u \in \text{Fix } T$, so Fix $T \neq \emptyset$. At the same time, if $u \in \text{Fix } T$, then $(u, P_V u) \in S$ (see [4]). We take b an upper bound on $d(x_0, z)$ for some fixed $z \in \text{Fix } T$.

Lemma 4.11. For any $x \in X$ and $u \in \text{Fix } T$,

$$d(Tx, P_V x)^2 \leq \rho^2 + d(u, x)^2 - d(u, Tx)^2.$$

Proof. Apply first (2.3) with $\beta = 1$ to get $d(Tx, P_V x)^2 + d(u, Tx)^2 \leq d(u, P_V x)^2$ and $d(x, P_V x)^2 + d(P_V x, P_V u)^2 \leq d(x, P_V u)^2$. Then

$$\begin{aligned} d(Tx, P_V x)^2 + d(u, Tx)^2 &\leq d(u, P_V x)^2 + d(x, P_V u)^2 - d(x, P_V x)^2 - d(P_V x, P_V u)^2 \\ &\leq d(u, P_V u)^2 + d(u, x)^2 \quad \text{by (2.1)}. \end{aligned}$$

□

In [4] it was proved that the sequence (x_n) is asymptotically regular with a rate of asymptotic regularity

$$\alpha_s(\varepsilon) = \left\lceil \frac{s}{\varepsilon^2} \right\rceil + 1,$$

where $s \geq d(Tx_0, P_V x_0)^2$. Applying Lemma 4.11, we get $d(Tx_0, P_V x_0)^2 \leq \rho^2 + d(z, x_0)^2 \leq \rho^2 + b^2$, so we can take $s = \rho^2 + b^2$ to obtain the following rate of asymptotic regularity

$$\alpha(\varepsilon) = \left\lceil \frac{\rho^2 + b^2}{\varepsilon^2} \right\rceil + 1.$$

To obtain a modulus of regularity for T, we assume next additional conditions on the sets U and V. These notions are natural analogues of the ones given in [9] in the setting of Hilbert spaces.

Definition 4.12. Let $K \subseteq X$. We say that U and V are *regular* w.r.t. K if

$$\forall \varepsilon > 0 \, \exists \delta > 0 \, \forall x \in K \; (\text{dist}(x, U) < \delta \wedge \text{dist}(x, V) < \rho + \delta \Rightarrow \text{dist}(x, \text{Fix } T) < \varepsilon). \tag{4.24}$$

We say that U and V are *linearly regular* w.r.t. K if

$$\exists \tau > 0 \, \forall x \in K \; (\text{dist}(x, \text{Fix } T) \leq \tau \max\{\text{dist}(x, U), \text{dist}(x, V) - \rho\}). \tag{4.25}$$

Note that 4.24 - for $m = 2$ - can be viewed as a generalization of 4.6 to the case of not necessarily intersecting sets U, V.
Extensions of this concept have been introduced in the recent paper [23] to analyze the speed of convergence of a sequence defined by a family of operators.

Proposition 4.13. If U and V are additionally regular w.r.t. $\overline{B}(z,b)$, then T admits a modulus of regularity w.r.t. Fix T and $\overline{B}(z,b)$.

Proof. Let $\varepsilon > 0$. Since U, V are regular w.r.t. $\overline{B}(z,b)$, there exists $\delta = \delta(\varepsilon) > 0$ given by (4.24). We prove that $\phi : (0,\infty) \to (0,\infty)$,

$$\phi(\varepsilon) = \frac{\rho\delta}{b+\rho},$$

is a modulus of regularity for T w.r.t. Fix T and $\overline{B}(z,b)$. To this end, let $x \in \overline{B}(z,b)$ such that $d(x,Tx) < \phi(\varepsilon)$. As $Tx \in U$, $\mathrm{dist}(x,U) < \phi(\varepsilon) < \delta$. From Lemma 4.11, it follows that

$$d(Tx, P_V x)^2 \leq \rho^2 + 2d(x,Tx)d(z,x) \leq \rho^2 + 2b\phi(\varepsilon).$$

Now,

$$\mathrm{dist}(x,V) \leq d(x,P_V x) \leq d(x,Tx) + d(Tx,P_V x) < \phi(\varepsilon) + \sqrt{\rho^2 + 2b\phi(\varepsilon)}.$$

Because $b\phi(\varepsilon) = \rho(\delta - \phi(\varepsilon))$, it follows that

$$\rho^2 + 2b\phi(\varepsilon) = \rho^2 + 2\rho(\delta - \phi(\varepsilon)) < \rho^2 + 2\rho(\delta - \phi(\varepsilon)) + (\delta - \phi(\varepsilon))^2 = (\rho + \delta - \phi(\varepsilon))^2,$$

and we obtain $\sqrt{\rho^2 + 2b\phi(\varepsilon)} < \rho + \delta - \phi(\varepsilon)$, hence $\phi(\varepsilon) + \sqrt{\rho^2 + 2b\phi(\varepsilon)} < \rho + \delta$. This shows that $\mathrm{dist}(x,V) < \rho + \delta$, so $\mathrm{dist}(x, \mathrm{Fix}\, T) < \varepsilon$. $\square$

Remark 4.14. Even more, if U and V are linearly regular w.r.t. $\overline{B}(z,b)$, then one can show as above that $\phi(\varepsilon) = (\varepsilon/\tau)\rho/(b+\rho)$, where τ satisfies (4.25), is a modulus of regularity for T w.r.t. Fix T and $\overline{B}(z,b)$.

Another algorithm which fits into the scheme (4.23) is the gradient descent method employed to find minimizers of convex functions. Let H be a Hilbert space and $f : H \to \mathbb{R}$ convex, Fréchet differentiable on H and such that its gradient ∇f is L-Lipschitz. Suppose that argmin $f \neq \emptyset$. Given $x_0 \in H$, the gradient descent method with constant step size $1/L$ generates the sequence (x_n) defined by (4.23), where $T : H \to H$, $T = \mathrm{Id} - \frac{1}{L}\nabla f$. Then T is firmly nonexpansive by the Baillon-Haddad theorem (see [6]) and argmin $f \subseteq \mathrm{Fix}\, T$, so $\mathrm{Fix}\, T \neq \emptyset$. We take b to be an upper bound on $d(x_0, z)$ for some $z \in \mathrm{Fix}\, T$.

From [3], it follows that the sequence (x_n) is asymptotically regular with a rate of asymptotic regularity

$$\alpha(\varepsilon) := \left\lceil \frac{32(b+1)^2}{\varepsilon^2} \right\rceil.$$

According to Example 3.9.(i) if, in addition, ∇f is ψ-strongly accretive, then $\phi(\varepsilon) = \psi(\varepsilon)/L$ is a modulus of regularity for T w.r.t. Fix T.

In [4, 37], it is shown that in a complete CAT(κ) space X (with diam $X < \pi/(2\sqrt{\kappa})$ if $\kappa > 0$), given $x_0 \in X$, the Picard iteration (x_n) of the composition $T = T_m \circ \ldots \circ T_1$ of finitely many metric projections $T_i = P_{C_i}$ onto nonempty, closed and convex sets $C_i \subseteq X$, $i = 1, \ldots, m$, with $C = \bigcap_{i=1}^m C_i \neq \emptyset$ is asymptotically regular and has common approximate fixed points. Moreover, an explicit common approximate fixed point bound (in the sense of Corollary 4.8) is given (follows from [4, Remark 3.1] if $\kappa = 0$ and [37, Corollaries 4.17, 4.5] and the Lipschitz continuity of T_i if

$\kappa > 0$). Since metric projections in CAT(κ) spaces are quasi-nonexpansive (and so is T since the fixed points of T are precisely the common fixed points of $T_1, \ldots, T_m$, see [37]) one gets the Fejér monotonicity w.r.t. C of the sequence (x_n). Thus, our general results on rates of convergence are applicable in this situation too (see Corollary 4.9 and the comment below it).

In fact we can even obtain linear convergence if the sets are linearly regular. We first prove the following inequality.

Lemma 4.15. Let $T_i = P_{C_i}$, $i = 1 \ldots m$, and β be as in (2.3). Then

$$d(T_i \circ \ldots \circ T_1 x, y)^2 \leq d(x,y)^2 - \frac{\beta}{i^2} d(x, T_i \circ \ldots \circ T_1 x)^2,$$

for all $x \in X$, $y \in C$, and $i = 1, \ldots, m$.

Proof. Let $x \in X$ and $y \in C$. If $i = 1$, the conclusion follows directly from (2.3). Suppose next that $i \geq 2$ and denote

$$\varepsilon = d(x,y)^2 - d(T_i \circ \ldots \circ T_1 x, y)^2.$$

Using (2.3) and the inequalities

$$d(T_i \circ \ldots \circ T_1 x, y) \leq d(T_{i-1} \circ \ldots \circ T_1 x, y) \leq \ldots \leq d(T_1 x, y) \leq d(x,y),$$

we get

$$\beta d(x, T_1 x)^2 \leq d(x,y)^2 - d(T_1 x, y)^2 \leq \varepsilon$$

and

$$\beta d(T_{k-1} \circ \ldots \circ T_1 x, T_k \circ \ldots \circ T_1 x)^2 \leq d(T_{k-1} \circ \ldots \circ T_1 x, y)^2 - d(T_k \circ \ldots \circ T_1 x, y)^2 \leq \varepsilon,$$

for all $k = 2, \ldots, i$. Therefore,

$$d(x, T_i \circ \ldots \circ T_1 x) \leq d(x, T_1 x) + \sum_{k=2}^{i} d(T_{k-1} \circ \ldots \circ T_1 x, T_k \circ \ldots \circ T_1 x) \leq i \sqrt{\frac{\varepsilon}{\beta}}.$$

So,

$$d(x, T_i \circ \ldots \circ T_1 x)^2 \leq \frac{i^2}{\beta} \left(d(x,y)^2 - d(T_i \circ \ldots \circ T_1 x, y)^2 \right),$$

which proves the desired inequality. $\square$

Fix now $z \in C$ and $b > 0$ an upper bound on $d(x_0, z)$. We consider as above $T_i = P_{C_i}$, $i = 1 \ldots m$, and $T = T_m \circ \ldots \circ T_1$.

Proposition 4.16. If $C_1, \ldots, C_m$ are linearly regular w.r.t. $\overline{B}(z,b)$, then

$$\phi(\varepsilon) = \frac{\beta}{2\tau^2 m^2} \varepsilon,$$

where β and τ are given by (2.3) and (4.21), respectively, is a modulus of regularity for T w.r.t. Fix T and $\overline{B}(z,b)$.

Proof. Let $x \in \overline{B}(z,b)$ and $i \leq m$. Then

$$\begin{aligned}\operatorname{dist}(x,C_i)^2 &\leq d(x,T_i\circ\ldots\circ T_1x)^2 \leq \frac{i^2}{\beta}\left(d(x,y)^2 - d(T_i\circ\ldots\circ T_1x,y)^2\right) \quad \text{by Lemma 4.15}\\ &\leq \frac{i^2}{\beta}\left(d(x,y)^2 - d(Tx,y)^2\right) \leq \frac{2i^2}{\beta}d(x,Tx)d(x,y),\end{aligned}$$

for all $y \in C$. Thus,

$$\operatorname{dist}(x,C_i)^2 \leq \frac{2i^2}{\beta}d(x,Tx)\operatorname{dist}(x,C),$$

so

$$\operatorname{dist}(x,C)^2 \leq \tau^2 \max_{i=1,\ldots,m} \operatorname{dist}(x,C_i)^2 \leq \frac{2\tau^2m^2}{\beta}d(x,Tx)\operatorname{dist}(x,C).$$

Hence,

$$\operatorname{dist}(x,C) \leq \frac{2\tau^2m^2}{\beta}d(x,Tx),$$

from where we obtain the conclusion. □

Now we can state the linear convergence result.

Corollary 4.17. If $C_1,\ldots,C_m$ are linearly regular w.r.t. $\overline{B}(z,b)$, then (x_n) converges linearly to some point in C.

Proof. We apply Theorem 4.5 with $F(x) = d(x,Tx)$. Note that zer $F = \operatorname{Fix} T = C$, which is closed. By Lemma 4.15, (4.18) is satisfied. At the same time, Proposition 4.16 provides the required modulus of regularity. □

We finish this subsection with the following observations.

Remark 4.18. There exists a computable firmly nonexpansive mapping $T : [0,1] \to [0,1]$ such that the computable Picard iteration $x_n = T^n 0$ is convergent and does not have a computable rate of convergence.

Proof. We use a construction from [49]: let (a_n) be a so-called Specker sequence, i.e., a computable nondecreasing sequence of rational numbers in $[0,1]$ without a computable limit (which exists by [59]). Define

$$f_n : [0,1] \to [0,1],\ \ f_n(x) = \max\{x, a_n\}$$

and put

$$T(x) = \frac{1}{2}(x + f(x)),\ \text{where } f(x) = \sum_{n=0}^{\infty} 2^{-n-1} f_n(x).$$

Then $f : [0,1] \to [0,1]$ is nonexpansive with Fix $f = [a,1]$, where $a = \lim_{n\to\infty} a_n$, and so T is firmly nonexpansive with Fix $T = [a,1]$. Since $x_n \leq x_{n+1} \leq a$, (x_n) converges to a fixed point of T which must be a. If (x_n) had a computable rate of convergence, then a would be computable, which is a contradiction. □

Remark 4.19. Since T is firmly nonexpansive, the sequence (x_n) defined in Remark 4.18 is Fejér monotone w.r.t. Fix T and asymptotically regular with an explicit rate of asymptotic regularity. Thus, by Theorem 4.1 (applied to $F(x) = |x - Tx|$), T has no computable modulus of regularity w.r.t. Fix T. This shows that Proposition 3.3 is essentially nonconstructive. A characterization of the computability-theoretic status of Proposition 3.3 in terms of ‘reverse mathematics’ and Weihrauch complexity is given in [38].

Mann iteration

Let X be a uniquely geodesic space and $T : X \to X$ a selfmapping with $\mathrm{Fix}\, T \neq \emptyset$. The *Mann iteration* associated to T starting from $x_0 \in X$ is defined by

$$x_{n+1} = (1-\lambda_n)x_n + \lambda_n T x_n \quad \text{for any } n \in \mathbb{N}, \tag{4.26}$$

where the coefficients λ_n are in $[0,1]$.

Suppose next that X is a CAT(0) space, T is nonexpansive, $b > 0$ is an upper bound on $d(x_0, z)$ for some $z \in \mathrm{Fix}\, T$, and (λ_n) satisfies $\sum_{n=0}^{\infty} \lambda_n(1-\lambda_n) = \infty$ with a rate of divergence θ. By (2.2) and the nonexpansivity of T, (x_n) is Fejér monotone w.r.t. $\mathrm{Fix}\, T$.

In [43] it was proved that the sequence (x_n) is asymptotically regular with a rate of asymptotic regularity

$$\alpha(\varepsilon) = \theta\left(\left\lceil \frac{4(b+1)^2}{\varepsilon^2} \right\rceil\right).$$

In the setting of Hilbert spaces, the Mann algorithm has been used in combination with splitting methods to solve problems that can be abstracted into finding a zero of the sum of two maximal monotone operators. Let $A, B : H \to 2^H$ be two maximal monotone operators with $\mathrm{zer}(A+B) \neq \emptyset$ and let $\gamma > 0$. The Douglas-Rachford algorithm is the Mann algorithm with $T = R_{\gamma A} R_{\gamma B}$. Note that in this case, as mentioned in Section 2, T is a nonexpansive mapping defined on H. Since, by [12, Proposition 26.1], $\mathrm{zer}(A+B) = J_{\gamma B}(\mathrm{Fix}\, T)$, we have $\mathrm{Fix}\, T \neq \emptyset$. If ϕ is a modulus of regularity for T w.r.t. $\mathrm{Fix}\, T$ and $\overline{B}(z,b)$, then, by Theorem 4.1, the sequence (x_n) converges to a fixed point of T with rate of convergence $\alpha(\phi(\varepsilon/2))$.

In particular, if C_1, C_2 and T are as in Example 3.6.(iv), then we obtain a rate of convergence for (x_n) to a fixed point of T whose projection onto C_1 lies in $C_1 \cap C_2 = \mathrm{zer}(N_{C_2} + N_{C_1})$. Consequently, the sequence $(P_{C_1} x_n)$ converges to a point in $C_1 \cap C_2$ with the same rate of convergence.

The CFP can also be solved using a Mann-type iteration studied by Crombez [24, 25] which was analyzed quantitatively in [33] (see also [56, 51]). Let H be a Hilbert space and $C_1, \ldots, C_m \subseteq H$ be closed and convex subsets with $C = \bigcap_{i=1}^m C_i \neq \emptyset$. For $1 \leq i \leq m$, let $P_{C_i} : H \to C_i$ be metric projections, $T_i = \mathrm{Id} + \lambda_i(P_{C_i} - \mathrm{Id})$ with $0 < \lambda_i \leq 2$, $\lambda_1 < 2$, and put $T = \sum_{i=1}^m a_i T_i$, where $a_1, \ldots, a_m \in (0,1)$ with $\sum_{i=1}^m a_i = 1$. As shown in [33], T can be written as $T = a\mathrm{Id} + (1-a)S$ for suitable $a \in (0,1)$ and nonexpansive $S : C \to C$ which satisfies $\mathrm{Fix}\, S = C$. Let $x_0 \in H$ and $b \geq \|x_0 - z\|$ for some $z \in C$. The sequence $x_n = T^n x_0$ is Fejér monotone w.r.t. $\mathrm{Fix}\, S$ since it is the Mann iteration associated to S with constant coefficient a.

Moreover, the sequences $(\|x_n - P_{C_i} x_n\|)_n$, $i = 1, \ldots, m$, are asymptotically regular with a rate of asymptotic regularity α^* which is quartic in $1/\varepsilon$ (see [33, Corollary 4.3.(i)] where the exact expression of α^* is given). In particular, α^* is a common approximate fixed point bound for $P_{C_1}, \ldots, P_{C_m}$.

We can now apply Corollary 4.9 to obtain that the sequence (x_n) converges to a point in C with a rate of convergence $\alpha^*(\rho(\varepsilon/2))$ whenever $C_1, \ldots, C_m$ are regular w.r.t. $\overline{B}(z,b)$ with modulus ρ. This, in particular, applies to $H = \mathbb{R}^n$ and the situation of Example 4.7 with the modulus of regularity ρ given there.

Proximal point algorithm

Let H be a Hilbert space and $A : H \to 2^H$ a maximal monotone operator with $\mathrm{zer}\, A \neq \emptyset$. Note that $\mathrm{zer}\, A$ is closed. Given $x_0 \in H$ and a sequence of positive numbers (γ_n), the *proximal point algorithm* (PPA) generates the sequence defined by

$$x_{n+1} = J_{\gamma_n A} x_n \quad \text{for any } n \in \mathbb{N}. \tag{4.27}$$

Because $\operatorname{zer} A = \operatorname{Fix} J_{\gamma_n A}$ and $J_{\gamma_n A}$ is nonexpansive, it follows that (x_n) is Fejer monotone w.r.t. $\operatorname{zer} A$. Define $F : H \to \overline{\mathbb{R}}$, $F(x) = \operatorname{dist}(O, A(x))$, take $b > 0$ an upper bound of $\|x_0 - z\|$ for some $z \in \operatorname{zer} A$, and suppose that $\sum_{n=0}^{\infty} \gamma_n^2 = \infty$ with a rate of divergence θ.

Proposition 4.20. (x_n) is a sequence of approximate zeros of F and

$$\alpha(\varepsilon) = \theta\left(\left\lceil \frac{b^2}{\varepsilon^2} \right\rceil\right) + 1 \tag{4.28}$$

is a rate of convergence of $(F(x_n))$ to 0.

Proof. Denote $u_n = \frac{x_n - x_{n+1}}{\gamma_n}$ for all $n \in \mathbb{N}$. By [40, Lemma 8.3.(ii)], $\theta\left(\left\lceil \frac{b^2}{\varepsilon^2} \right\rceil\right)$ is a rate of convergence of $(\|u_n\|)$ to 0. From the definition of the resolvent it follows that $u_n \in A(x_{n+1})$, hence $F(x_{n+1}) \leq \|u_n\|$ for all $n \in \mathbb{N}$. Now the conclusion is immediate. ☐

If ϕ is a modulus of regularity for A w.r.t. $\operatorname{zer} A$ and $\overline{B}(z, b)$, then, by Theorem 4.1, (x_n) converges to some $z' \in \operatorname{zer} A$ with a rate of convergence $\alpha(\phi(\varepsilon/2, b))$. In addition, if (4.22) holds, then we can apply Corollary 4.10 to get $x_n = z'$ for all $n \geq \alpha(\varepsilon^*)$. This gives a quantitative version of a special form of [57, Theorem 3].

The PPA has been extensively applied as a method to localize a minimizer of a convex function. Let $f : H \to (-\infty, \infty]$ be proper, convex and lower semi-continuous and $S = \operatorname{argmin}\ f \neq \emptyset$. In this case, the sequence (x_n) is given by (4.27) for $A = \partial f$ and we take z and b as above.

Assuming additional conditions, we can obtain an explicit modulus of regularity for ∂f.

Proposition 4.21. If S is a set of ψ-boundedly global weak sharp minima for f and $f|_{\overline{B}(z,b+2)}$ is uniformly continuous with a modulus of uniform continuity ρ, then denoting $C = \overline{B}(z, b+1)$,

$$\phi(\varepsilon) = \min\left\{\rho\left(\frac{\psi_C(\varepsilon/2)}{2}\right), \frac{\psi_C(\varepsilon/2)}{2(b+1)}, \frac{\varepsilon}{2}, 1\right\}$$

is a modulus of regularity for ∂f w.r.t. $\operatorname{zer} \partial f$ and $\overline{B}(z, b)$.

Proof. Since S is a set of ψ-boundedly global weak sharp minima for f, using Example 3.7.(ii), $\phi^*(\varepsilon) = \psi_C(\varepsilon)$ is a modulus of regularity for f w.r.t. S and $\overline{B}(z, b+1)$. The conclusion follows by Theorem 3.10 and using the expressions of the moduli computed in its proof. ☐

The following particular situation yields finite convergence of (x_n). This gives a quantitative version of the main result in [27].

Proposition 4.22. If S is a set of ψ-global weak sharp minima for f with $\psi(\varepsilon) = k\,\varepsilon$, where $k > 0$, then $x_n = z'$ for all $n \geq \alpha(\varepsilon^*)$, where $z' \in S$ and α is given by (4.28).

Proof. We prove that $F : H \to \overline{\mathbb{R}}$, $F(x) = \operatorname{dist}(O, \partial f(x))$ satisfies (4.17). By Theorem 2 and Lemma 5 in [27], there exists $\varepsilon^* > 0$ such that if $x \in H$, $u \in \partial f(x)$ with $\|u\| \leq \varepsilon^*$, then $x \in S$. For $w \in \mathbb{R}$, $0 \leq w < \varepsilon^*$ and $x \in F^{-1}(w)$, we have $\operatorname{dist}(O, \partial f(x)) = w < \varepsilon^*$, so there exists $u \in \partial f(x)$ such that $\|u\| < \varepsilon^*$. Therefore, $x \in S$, which yields $F^{-1}(w) \subseteq \operatorname{zer} F$. Now we only need to apply Remark 4.2. ☐

We finish with a result concerning the linear convergence of the PPA in CAT(0) spaces which follows as a straightforward consequence of Theorem 4.5. This provides an extension of [44, Theorem 3.1].

Let X be a complete CAT(0) space, $f : X \to (-\infty, \infty]$ a proper, convex and lower semi-continuous function with $\operatorname{argmin} f \neq \emptyset$, $x_0 \in X$ and $b \geq d(x_0, z)$ for some $z \in \operatorname{argmin} f$. Consider the sequence $(x_n) \subseteq X$ generated by the PPA starting at x_0 and applying successively the resolvent J_γ for a fixed $\gamma > 0$, i.e.,

$$x_{n+1} = J_\gamma(x_n),$$

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

Proposition 4.23. If J_γ has a modulus of regularity ϕ w.r.t Fix J_γ and $\overline{B}(z, b)$ of the form $\phi(\varepsilon) = t\varepsilon$, where $t > 0$, then (x_n) converges linearly to a minimizer of f.

Proof. By (2.6), it follows that (4.18) holds with $F(x) = d(x, J_\gamma x)$, $p = 2$ and $\beta = 1$. Moreover, the set Fix J_γ is closed. This means that we can apply Theorem 4.5. □

Acknowledgements

The first author was supported by the German Science Foundation (DFG Project KO 1737/6-1). The second and third authors were supported by DGES (grant number MTM2015-65242-C2-1-P). The third author would also like to acknowledge the Juan de la Cierva - Incorporación Fellowship Program of the Spanish Ministry of Economy and Competitiveness.

The authors thank the referee for the comments and suggestions that helped improve the manuscript.

References

[1] R.M. Anderson, *'Almost' implies 'Near'*, Trans. Amer. Math. Soc. **296** (1986), 229–237.

[2] D. Ariza-Ruiz, E.M. Briseid, A. Jiménez-Melado, and G. López-Acedo, *Rate of convergence under weak contractiveness conditions*, Fixed Point Theory **14** (2013), 11–27.

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

[4] D. Ariza-Ruiz, G. López-Acedo, and A. Nicolae, *The asymptotic behavior of the composition of firmly nonexpansive mappings*, J. Optim. Theory Appl. **167** (2015), 409–429.

[5] M. Bačák, *Computing medians and means in Hadamard spaces*, SIAM J. Optim. **24** (2014), 1542–1566.

[6] J.B. Baillon and G. Haddad, *Quelques proprietes des operateurs angle-bornes et ncy-cliquement monotones*, Israel J. Math. **26** (1977), 137–150.

[7] S. Banert, *Backward-backward splitting in Hadamard spaces*, J. Math. Anal. Appl. **414** (2014), 656–665.

[8] H.H. Bauschke, *The approximation of fixed points of compositions of nonexpansive mappings in Hilbert space*, J. Math. Anal. Appl. **202** (1996), 150–159.

[9] H.H. Bauschke and J.M. Borwein, *On the convergence of von Neumann's alternating projection algorithm for two sets*, Set-Valued Anal. **1** (1993), 185–212.

[10] H.H. Bauschke and J.M. Borwein, *On projection algorithms for solving convex feasibility problems,* SIAM Review **38** (1996), 367–426.

[11] H.H. Bauschke and J.M. Borwein, *Dykstra's alternating projection algorithm for two sets,* J. Approx. Theory **79** (1994), 418–443.

[12] H.H. Bauschke and P.L Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces, Second Edition,* CMS Books in Mathematics, Springer, Cham, 2017.

[13] J. Ben-Artzi, A.C. Hansen, O. Nevanlinna and M. Seidel, *New barriers in complexity theory: on the solvability complexity index and the towers of algorithms,* C. R. Math. Acad. Sci. Paris, Ser. I **353** (2015), 931–936.

[14] I.D. Berg and I.G. Nikolaev, *Quasilinearization and curvature of Aleksandrov spaces,* Geom. Dedicata **133** (2008), 195–218.

[15] I.D. Berg and I.G. Nikolaev, *Characterization of Aleksandrov Spaces of Curvature Bounded Above by Means of the Metric Cauchy-Schwarz Inequality,* Michigan Math. J. **67** (2018), 289–332.

[16] J. Bolte, T.P. Nguyen, J. Peypouquet, and B.W. Suter, *From error bounds to the complexity of first-order descent methods for convex functions,* Math. Program., Ser. A **165** (2017), 471–507.

[17] J.M. Borwein, G. Li, and L.J. Yao, *Analysis of the convergence rate for the cyclic projection algorithm applied to basic semi-algebraic convex sets,* SIAM J. Optim. **24** (2014), 498–527.

[18] J.M. Borwein, G. Li, and M.K. Tam, *Convergence rate analysis for averaged fixed point iterations in common fixed point problems,* SIAM J. Optim. **27** (2017), 1–33.

[19] M.R. Bridson and A. Haefliger, *Metric spaces of non-positive curvature,* Springer-Verlag, Berlin, 1999.

[20] E.M. Briseid, *Logical aspects of rates of convergence in metric spaces,* J. Symb. Logic **74** (2009), 1401–1428.

[21] J.V. Burke and S. Deng, *Weak sharp minima revisited. I. Basic theory,* Control Cybernet. **31** (2002), 439–469.

[22] J.V. Burke and M.C. Ferris, *Weak sharp minima in mathematical programming,* SIAM J. Control Optim. **31** (1993), 1340–1359.

[23] A. Ciegielski, S. Reich, and R. Zalas, *Regular Sequences of Quasi-Nonexpansive Operators and Their Applications,* SIAM J. Optim. **28** (2018), 1508–1532.

[24] G. Crombez, *Image recovery by convex combinations of projections,* J. Math. Anal. Appl. **155** (1991), 413–419.

[25] G. Crombez, *Parallel methods in image recovery by projections onto convex sets,* Czechoslovak Math. J. **42** (1992), 445–450.

[26] A.L. Dontchev and R.T. Rockafellar, *Implicit functions and solution mappings. A view from variational analysis,* Springer Monographs in Mathematics, Springer, Dordrecht, 2009.

[27] M.C. Ferris, *Finite termination of the proximal point algorithm,* Math. Program., Ser. A **50** (1991), 359–366.

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

[29] A. Grzegorczyk, *On the definitions of computable real continuous functions,* Fund. Math. **44** (1957), 61–71.

[30] L.G. Gubin, B.T. Polyak, and E.V. Raik, *The method of projections for finding the common point of convex sets,* USSR Comput. Math. Phys. **7** (1967), 1–24.

[31] I. Halperin, *The product of projection operators,* Acta Sci. Math. (Szeged) **23** (1963), 96–99.

[32] J. Jost, *Convex functionals and generalized harmonic maps into spaces of non positive curvature,* Comment. Math. Helvetici **70** (1995), 659–673.

[33] M.A.A. Khan and U. Kohlenbach, *Quantitative image recovery theorems,* Nonlinear Anal. **106** (2014), 138–150.

[34] U. Kohlenbach, *Effective moduli from ineffective uniqueness proofs. An unwinding of de La Vallée Poussin's proof for Chebycheff approximation,* Ann. Pure Appl. Logic **64** (1993), 27–94.

[35] U. Kohlenbach, *New effective moduli of uniqueness and uniform a–priori estimates for constants of strong unicity by logical analysis of known proofs in best approximation theory.* Numer. Funct. Anal. Optim. **14** (1993), 581–606.

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

[37] U. Kohlenbach, *On the quantitative asymptotic behavior of strongly nonexpansive mappings in Banach and geodesic spaces,* Israel J. Math. **216** (2016), 215–246.

[38] U. Kohlenbach, *On the reverse mathematics and Weihrauch complexity of moduli of regularity and uniqueness,* to appear in: Computability.

[39] U. Kohlenbach and A. Koutsoukou-Argyraki, *Rates of convergence and metastability for abstract Cauchy problems generated by accretive operators,* J. Math. Anal. Appl. **423** (2015), 1089–1112.

[40] U. Kohlenbach, L. Leuştean, and A. Nicolae, *Quantitative results on Fejér monotone sequences,* Commun. Contemp. Math. **20** (2018), 1750015, 42 pp.

[41] U. Kohlenbach and P. Oliva, *Proof mining in L_1-approximation.* Ann. Pure Appl. Logic **121** (2003), 1–38.

[42] E. Kopecká and S. Reich, sl A note on alternating projections in Hilbert space, J. Fixed Point Theory Appl. **12** (2012), 41–47.

[43] L. Leuştean, *A quadratic rate of asymptotic regularity for CAT(0)-spaces,* J. Math. Anal. Appl. **325** (2007), 386–399.

[44] D. Leventhal, *Metric subregularity and the proximal point method,* J. Math. Anal. Appl. **360** (2009), 681–688.

[45] C. Li, B.S. Mordukhovich, J.H. Wang, and J.C. Yao, *Weak sharp minima on Riemannian manifolds,* SIAM J. Optim. **21** (2011), 1523–1560.

[46] G. Li, B.S. Mordukhovich, T.T.A. Nghia, and T.S. Pham, *Error bounds for parametric polynomial systems with applications to higher order stability systems and convergence rates,* Math. Program., Ser. A **168** (2018), 313–346.

[47] J. Liang, M.J. Fadili, and G. Peyré, *Local linear convergence of Forward-Backward under partial smoothness,* In: Advances in Neural Information Processing Systems (NIPS) (2014), pp. 1979–1978.

[48] J. Liang, J. Fadili, and G. Peyré, *Convergence rates with inexact non-expansive operators,* arXiv:1404.4837 [math.OC].

[49] E. Neumann, *Computational problems in metric fixed point theory and their Weihrauch degrees,* Log. Method. Comput. Sci. **11** (2015), 44pp.

[50] O. Nevanlinna and S. Reich, *Strong convergence of contraction semigroups and of iterative methods for accretive operators in Banach spaces,* Israel J. Math. **32** (1979), 44–58.

[51] A. Nicolae, *Asymptotic behavior of averaged and firmly nonexpansive mappings in geodesic spaces,* Nonlinear Anal. **87** (2013), 102–115.

[52] S.-I. Ohta, *Convexities of metric spaces,* Geom. Dedicata **125** (2007), 225–250.

[53] C.L. Outlaw, *Mean value iteration of nonexpansive mappings in a Banach space,* Pacific J. Math. **30** (1969) 747–750.

[54] A. Petruşel and I.A. Rus, *Graphic contraction principle and applications,* Preprint.

[55] M.B. Pour-El and J.I. Richards, *Computability in Analysis and Physics,* Springer-Verlag, Berlin Heidelberg, 1989.

[56] S. Reich, *A limit theorem for projections,* Linear and Multilinear Algebra **13** (1983), 281–290.

[57] R.T. Rockafellar, *Monotone operators the proximal point algorithm,* SIAM J. Control Optim. **14** (1976), 877–898.

[58] H.F. Senter and W.G. Dotson Jr., *Approximating fixed points of nonexpansive mappings,* Proc. Amer. Math. Soc. **44** (1974), 375–380.

[59] E. Specker, *Nicht konstruktiv beweisbare Sätze der Analysis,* J. Symb. Logic **14** (1949), 145–158.

[60] K. Weihrauch, *Computable Analysis,* Springer, Berlin, 2000.