Rates of convergence for splitting algorithms*

Jacqueline Treusch and Ulrich Kohlenbach
Department of Mathematics
Technische Universität Darmstadt
Schlossgartenstraße 7, 64289 Darmstadt, Germany
jacqueline.treusch@gmx.de, kohlenbach@mathematik.tu-darmstadt.de

August 26, 2024

Abstract

We study the convergence of so-called splitting algorithms that compute zeros of sums $\mathsf{A}+\mathsf{B}$ of monotone operators from a quantitative point of view under suitable uniform monotonicity assumptions which guarantee their strong convergence. More precisely, we apply logic-based techniques from proof mining to construct rates of the convergence for (i) Tseng's algorithm, (ii) the forward-backward splitting algorithm, (iii) the Douglas-Rachford splitting algorithm as well as its limiting case given by (iv) the Peaceman-Rachford algorithm. In the latter case, we use a recent result due to Liu et al. together with a quantitative form of strong nonexpansivity. The rates of convergence depend on moduli of uniform monotonicity for A and/or B or, in the case of Tseng's algorithm, just for $\mathsf{A}+\mathsf{B}$ and are (at least) as general as the original strong convergence results.

Keywords: Maximally monotone operators, splitting methods, Douglas-Rachford, Peaceman-Rachford, Tseng's algorithm.
Mathematics Subject Classification (2010): 03F10, 47H04, 47H05, 47J05.

1 Introduction

This paper studies the convergence of so-called splitting algorithms that compute zeros of sums of monotone operators from a quantitative point of view under assumptions which guarantee their strong convergence. More precisely, we apply logic-based techniques from proof mining (see e.g. [6, 9]) to the convergence proofs of four different splitting methods presented in [1, Chapter 26], namely, (i) Tseng's algorithm, (ii) the forward-backward splitting algorithm, (iii) the Douglas-Rachford splitting algorithm as well as its limiting case given by (iv) the Peaceman-Rachford algorithm. The algorithms (i)-(iii) construct sequences (x_n) for which, with the help of an auxiliary sequence (z_n), it is proven that $x_n - z_n \to 0$ (asymptotic regularity). This is then used to show the weak convergence of (x_n) towards a solution $x \in$zer$(A+B)$. In a third step, these results are used to establish under the additional assumption that one of the operators A, B is uniformly monotone, that (x_n) strongly converges towards the then unique solution $x \in$zer$(A+B)$. This last type of result is particularly amenable to proof mining as from uniqueness statements one usually can extract so-called moduli of uniqueness which in many cases can be combined with asymptotic regularity results to obtain effective and highly uniform rates of convergence (see e.g. [6]). As discussed in [10], such moduli of uniqueness correspond, in the cases at hand, to quantitative moduli for (generalized forms of) the uniform monotonicity of A or B. The construction of rates of convergence depending on such moduli is, however, not trivial as even the strong convergence proofs for the algorithms (i)-(iii) at least prima facie refer to the weak convergence results proved in the more general case of monotone operators and - in some cases - due to the lack of monotonicity of the sequence $(\|x_n - x\|)$. Nevertheless, we will be able to construct full rates of convergence for the algorithms (i)-(iii) in given moduli of uniform monotonicity for A or B (in the case of (ii),(iii)) or just for $A+B$ (in the case of (i)). Interestingly, the strong convergence proof for (iv) as given in [1] - which does not proceed via a weak convergence result - is inherently noneffective and - when logically analyzed - only gives rise to a so-called rate of metastability (for

*This paper grew out of a Bachelor thesis [16] of the first author written under the supervision of the 2nd author except for section 5.2 which is due to the 2nd author.

$x_n \to x$) in the sense of Tao [14, 15]. Only when **both** operators A, B are assumed to posses moduli of uniform monotonicity, very recent results from [12], together with previous proof mining results due to the 2nd author in [7], can be utilized to obtain a full rate of convergence for the Peaceman-Rachford algorithm. For each of the algorithms treated, we will give brief discussions on how recent logical metatheorems of proof mining in the context of (maximally) monotone operators in Hilbert space (see [13, 11] which in turn extend [5, 4]) explain our rates of convergence from a qualitative point of you (w.r.t. the data used) as instances of general logical phenomena.
While there are numerous results in the literature on rates of convergence for splitting methods in special cases and under additional assumptions (see e.g. [1, Prop. 26.16] or Chapter 10 of [3]) our results are as general as the original strong convergence theorems.

1.1 Basic Notions and Propositions

We first introduce some basic notation which will be used throughout this paper. In the following, $(\mathcal{H}, \|\cdot\|)$ always describes a real Hilbert space with scalar product $\langle\cdot,\cdot\rangle$ and $2^{\mathcal{H}}$ is the power set of $\mathcal{H}$. The natural numbers are denoted by $\mathbb{N}$ as usual and include 0, i.e., $\mathbb{N} = \{0, 1, 2, \ldots\}$.
The theorems that we consider involve set-valued operators. A set-valued operator $\mathsf{A} : \mathcal{H} \to 2^{\mathcal{H}}$ is an operator that maps every point x in $\mathcal{H}$ to a subset $\mathsf{A}x$ of $\mathcal{H}$ and is characterized by its *graph*

$$\mathsf{gra}(\mathsf{A}) := \{(x, u) \in \mathcal{H} \times \mathcal{H} : \ u \in \mathsf{A}x\}.$$

Other useful notions for such operators are

- the *domain* of A, i.e., $\mathsf{dom}(\mathsf{A}) := \{x \in \mathcal{H} : \ \mathsf{A}x \neq \emptyset\}$,
- the *range* of A, i.e., $\mathsf{ran}(\mathsf{A}) := \mathsf{A}(\mathcal{H})$,
- and the set of *zeros* of A, i.e., $\mathsf{zer}(\mathsf{A}) := \mathsf{A}^{-1}0 = \{x \in \mathcal{H} : \ 0 \in \mathsf{A}x\}$.

Given the operators $\mathsf{A} : \mathcal{H} \to 2^{\mathcal{H}}$ and $\mathsf{B} : \mathcal{H} \to 2^{\mathcal{H}}$, the sum $\mathsf{A} + \mathsf{B}$ is defined by

$$\mathsf{A} + \mathsf{B} : \mathcal{H} \to 2^{\mathcal{H}}, x \mapsto \mathsf{A}x + \mathsf{B}x = \{s + t : s \in \mathsf{A}x, t \in \mathsf{B}x\}.$$

Obviously $\mathsf{gra}(\mathsf{A} + \mathsf{B}) = \{(x, u + v) : (x, u) \in \mathsf{gra}(\mathsf{A}), (x, v) \in \mathsf{gra}(\mathsf{B})\}$ and $\mathsf{dom}(\mathsf{A} + \mathsf{B}) = \mathsf{dom}(\mathsf{A}) \cap \mathsf{dom}(\mathsf{B})$.
An important class of set-valued operators are those which are monotone.

Definition 1.1 (20.1 in [1])**.** *An operator* $\mathsf{A} : \mathcal{H} \to 2^{\mathcal{H}}$ *is* monotone *if*

$$\forall (x, u), (y, v) \in \mathsf{gra}(\mathsf{A}) \ \Big(\langle x - y, u - v\rangle \geq 0\Big)$$

Definition 1.2 (20.20 in [1])**.** *A monotone operator* $\mathsf{A} : \mathcal{H} \to 2^{\mathcal{H}}$ *is* maximally monotone *if for every* $(x, u) \in \mathcal{H} \times \mathcal{H}$

$$(x, u) \in \mathsf{gra}(\mathsf{A}) \ \Leftrightarrow \ \forall (y, v) \in \mathsf{gra}(\mathsf{A})\Big(\langle x - y, u - v\rangle \geq 0\Big).$$

Remark 1.3. *If* $\mathsf{A} : \mathcal{H} \to 2^{\mathcal{H}}$ *is (maximally) monotone and* $\gamma \in (0, \infty)$*, then* $\gamma\mathsf{A}$ *is again (maximally) monotone.*

We can also strengthen the notion of monotone operators to obtain uniformly monotone operators. Instead of only requiring that the scalar product is greater than or equal to zero, we actually demand the scalar product to have a proper distance to zero. This distance has to be quantifiable by a modulus. In [1], we have the following definition:

Definition 1.4 (22.1(iii) in [1])**.** *An operator* $\mathsf{A} : \mathcal{H} \to 2^{\mathcal{H}}$ *is* uniformly monotone *with function* $\phi : [0, \infty) \to [0, \infty]$ *if* ϕ *is increasing, vanishes only at 0 and*

$$\forall (x, u), (y, v) \in \mathsf{gra}(\mathsf{A}) \ \Big(\langle x - y, u - v\rangle \geq \phi(\|x - y\|)\Big).$$

This notion of uniform monotonicity can be naturally localized to a subset $\mathcal{C}$ of the domain. Then the last property of ϕ in Definition 1.4 only has to hold for $x, y \in \mathcal{C}$.
One advantage of uniformly monotone operators is that they actually have at most one zero. In fact, this holds for strictly monotone operators (see Proposition 23.35 in [1]) which is a weaker assumption on the operator than uniform monotonicity.

Proposition 1.5. *Let* $\mathsf{A} : \mathcal{H} \to 2^{\mathcal{H}}$ *be uniformly monotone. Then* $\mathsf{zer}(\mathsf{A})$ *is at most a singleton.*

For the extraction of rates of convergence later on, we need to convert such a given function ϕ using results from [10] into a modulus Θ which is defined next. The following definition is based on Definition 9 in [10]. The latter paper only treats operators that are uniformly monotone at zero. In this paper, on the other hand, we are going to need such a modulus also for arbitrary pairs of arguments. We can extend the definition naturally as follows:

Definition 1.6. *An operator* $\mathsf{A} : \mathcal{H} \to 2^{\mathcal{H}}$ *is* uniformly monotone *with modulus of uniform monotonicity* $\Theta_{(\cdot)}(\cdot) : \mathbb{N} \times \mathbb{N}\setminus\{0\} \to \mathbb{N}$ *if*

$$\forall k \in \mathbb{N}\, \forall K \in \mathbb{N}\setminus\{0\} \forall (x,u),(y,v) \in \mathsf{gra}(\mathsf{A}) \ \Big(\|x-y\| \in [2^{-k}, K] \to \langle x-y, u-v\rangle \geq 2^{-\Theta_K(k)}\Big).$$

For the quantitative the analysis of Tseng's algorithm we only need to assume the aforementioned weaker notion (if $\mathsf{zer}(\mathsf{A}) \neq \emptyset$) of a modulus of uniform monotonicity at zero from [10]:

Definition 1.7 ([10])**.** A *is uniformly monotone at* $x \in \mathsf{zer}(\mathsf{A})$ *with modulus* $\Theta_{(\cdot)}(\cdot) : \mathbb{N} \times \mathbb{N}\setminus\{0\} \to \mathbb{N}$ *if*

$$\forall k \in \mathbb{N}\ \forall K \in \mathbb{N}\setminus\{0\}\, \forall (y,u) \in \mathsf{gra}(\mathsf{A}+\mathsf{B}) \ \Big(\|x-y\| \in [2^{-l}, K] \to \langle y-x, u\rangle \geq 2^{-\Theta_K(k)}\Big). \tag{2.2.1}$$

If A is uniformly monotone at some zero x, then this x is the only zero of A and so A is uniformly monotone at any zero. Hence we can simply say 'uniformly monotone at zero' instead of 'uniformly monotone at zero x' (provided that $\mathsf{zer}(\mathsf{A}) \neq \emptyset$).
Let A be uniformly monotone with function φ, then

$$\Theta(l) := \min_{n\in\mathbb{N}} \big\{2^{-n} \leq \varphi\big(2^{-l}\big)\big\}.$$

is a modulus for A being uniformly monotone (and a-fortiori a modulus for A being uniformly monotone at zero) which does not depend on K (using that φ is increasing).
In this paper, we will actually use our moduli always for some fixed K. This is because we will mostly treat operators that are uniformly monotone on bounded subsets for which we can compute a suitable bound K.

Definition 1.8 (cf. 23.1 in [1])**.** *Let* $\mathsf{A} : \mathcal{H} \to 2^{\mathcal{H}}$*. The* resolvent *of* A *is defined to be the set-valued mapping* $J_{\mathsf{A}} = (\mathsf{Id} + \mathsf{A})^{-1}$.

Some useful statements regarding the resolvent, which follow directly from the definition, are summarized in the next proposition.

Proposition 1.9 (23.2 in [1])**.** *Let* $\mathsf{A} : \mathcal{H} \to 2^{\mathcal{H}}$*,* $\gamma \in (0,\infty)$ *and* $x, p \in \mathcal{H}$*. Then the following hold:*

(i) $\mathsf{dom}(\mathsf{J}_{\gamma\mathsf{A}}) = \mathsf{ran}(\mathsf{Id} + \gamma\mathsf{A})$ *and* $\mathsf{ran}(\mathsf{J}_{\gamma\mathsf{A}}) = \mathsf{dom}(\mathsf{A})$.

(ii) $p \in \mathsf{J}_{\gamma\mathsf{A}}\, x \Leftrightarrow x \in p + \gamma\mathsf{A}p \Leftrightarrow x - p \in \gamma\mathsf{A}p \Leftrightarrow (p, \gamma^{-1}(x-p)) \in \mathsf{gra}(\mathsf{A})$.

Definition 1.10 (4.1(i)+(ii) in [1])**.** *Let* $\mathsf{D} \subseteq \mathcal{H}$ *be nonempty. An operator* $\mathcal{T} : \mathsf{D} \to \mathcal{H}$ *is*

(i) nonexpansive *if it is Lipschitz-continuous with Lipschitz constant 1, i.e.,*

$$\forall x, y \in \mathsf{D} \ \Big(\|\mathcal{T}x - \mathcal{T}y\| \leq \|x-y\|\Big).$$

(ii) firmly nonexpansive *if*

$$\forall x, y \in \mathsf{D} \ \Big(\|\mathcal{T}x - \mathcal{T}y\|^2 + \|(\mathsf{Id} - \mathcal{T})x - (\mathsf{Id} - \mathcal{T})y\|^2 \leq \|x-y\|^2\Big).$$

Note that every firmly nonexpansive operator is in particular nonexpansive.

Proposition 1.11 (23.10 in [1])**.** *Let* $\mathsf{A} : \mathcal{H} \to 2^{\mathcal{H}}$ *be monotone. Then* $J_{\mathsf{A}} = (\mathsf{Id} + \mathsf{A})^{-1} : \mathsf{ran}(\mathsf{Id} + \mathsf{A}) \to \mathsf{dom}(\mathsf{A})$ *is a single valued firmly nonexpansive mapping. If A is maximally monotone, then* $\mathsf{ran}(\mathsf{Id} + \mathsf{A}) = \mathcal{H}$ *and so* J_{A} *is a total mapping.*

For maximally monotone operators we are going to consider another single-valued operator based on the resolvent. The operator is the so-called reflected resolvent and is defined as follows:

Definition 1.12. *For maximally monotone* $\mathsf{A} : \mathcal{H} \to 2^{\mathcal{H}}$ *and arbitrary* $\gamma \in (0, \infty)$ *the* reflected resolvent $\mathrm{R}_{\gamma\mathsf{A}}$ *is defined by*

$$\mathrm{R}_{\gamma\mathsf{A}} : \mathcal{H} \to \mathcal{H}, x \mapsto 2\,\mathrm{J}_{\gamma\mathsf{A}}\, x - x$$

$\mathrm{R}_{\gamma\mathsf{A}}$ is nonexpansive (see Corollary 23.11(ii) in [1]).

Definition 1.13 (4.10 and 4.33 in [1])**.** *Let D be a nonempty subset of $\mathcal{H}$, and let $\mathcal{T} : D \to \mathcal{H}$.*

a) Let $\alpha \in (0,1)$ and let $\mathcal{T}$ be nonexpansive. Then $\mathcal{T}$ is α-averaged *if there exists a nonexpansive operator* $\mathsf{R} : D \to \mathcal{H}$ *such that* $\mathcal{T} = (1-\alpha)\mathsf{Id} + \alpha\mathsf{R}$.

b) Let $\beta \in (0, \infty)$. Then $\mathcal{T}$ is β-cocoercive *if*

$$\forall x, y \in D \ \left(\langle x - y, \mathcal{T}x - \mathcal{T}y\rangle \geq \beta\|\mathcal{T}x - \mathcal{T}y\|^2\right).$$

Definition 1.14. *Let $D \neq \emptyset$ be a subset of $\mathcal{H}$ and let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{H}$. Then $(x_n)_{n\in\mathbb{N}}$ is* Fejér-monotone *with respect to D if*

$$\forall x \in D\, \forall n \in \mathbb{N}\ (\|x_{n+1} - x\| \leq \|x_n - x\|).$$

Theorem 1.15 (Groetsch, 5.15 in [1])**.** *Let D be a nonempty closed convex subset of $\mathcal{H}$, let $\mathcal{T} : D \to D$ be a nonexpansive operator such that $\mathsf{Fix}(\mathcal{T}) \neq \emptyset$, let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $[0,1]$ such that $\sum_{n\in\mathbb{N}} \lambda_n(1-\lambda_n) = \infty$, and let $x_0 \in D$. Set*

$$\forall n \in \mathbb{N}\ \ \big(x_{n+1} = x_n + \lambda_n(\mathcal{T}x_n - x_n)\big).$$

Then the following hold:

(i) $(x_n)_{n\in\mathbb{N}}$ is Fejér-monotone with respect to $\mathsf{Fix}(\mathcal{T})$.

(ii) $(\mathcal{T}x_n - x_n)_{n\in\mathbb{N}}$ converges strongly to 0.

In addition to Theorem 1.15 itself, we will also need some facts used in its proof: For the proof of *(i)* the nonexpansiveness of $\mathcal{T}$ is used to show statement (5.16) in [1] which asserts that for all $y \in \mathsf{Fix}(\mathcal{T})$ we have

$$\forall n \in \mathbb{N}\ \ \Big(\|x_{n+1} - y\|^2 \leq \|x_n - y\|^2 - \lambda_n(1-\lambda_n)\|\mathcal{T}x_n - x_n\|^2\Big). \tag{1.15.1}$$

This already proves *(i)* since the subtracted term on the right is always positive.
The inequality (1.15.1) is then used in the proof of *(ii)* to establish

$$\sum_{n\in\mathbb{N}} \lambda_n(1-\lambda_n)\|\mathcal{T}x_n - x_n\|^2 \leq \|x_0 - y\|^2. \tag{1.15.2}$$

By assumption, we have $\sum_{n\in\mathbb{N}} \lambda_n(1-\lambda_n) = \infty$ and therefore $\liminf_{n\to\infty} \|\mathcal{T}x_n - x_n\| = 0$. By (5.17) in [1]

$$\forall n \in \mathbb{N}\ \ \Big(\|\mathcal{T}x_{n+1} - x_{n+1}\| \leq \|\mathcal{T}x_n - x_n\|\Big) \tag{1.15.3}$$

holds which then yields $\mathcal{T}x_n - x_n \to 0$.

2 Tseng's Splitting Algorithm

The first theorem we are going to analyze is the convergence of Tseng's Splitting Algorithm. This algorithm finds zeros of the sum of two operators where one is maximally monotone and the other is single-valued and Lipschitz-continuous on a suitable subset. Each iteration consists of four computation steps: two forward, one backward and one projection step.

Theorem 2.1 (Tseng's Algorithm, 26.17 in [1]). *Let $\mathsf{D} \subseteq \mathcal{H}$ be nonempty, let $\mathsf{C} \subseteq \mathsf{D}$ be closed and convex, and let $\beta \in (0, \infty)$. Assume that $A : \mathcal{H} \to 2^{\mathcal{H}}$ is a maximally monotone operator with $\mathsf{dom}(\mathsf{A}) \subseteq \mathsf{D}$ and $B : \mathcal{H} \to 2^{\mathcal{H}}$ is a monotone operator which is single-valued on D and $\frac{1}{\beta}$-Lipschitz continuous relative to $C \cup \mathsf{dom}(\mathsf{A})$ and that $A + B$ is maximally monotone. Now suppose $C \cap \mathsf{zer}(\mathsf{A} + \mathsf{B}) \neq \emptyset$ and let $x_0 \in \mathsf{C}$ and $\gamma \in (0, \beta)$. For $n \in \mathbb{N}$ we set*

$$\begin{aligned} y_n &= x_n - \gamma \mathsf{B} x_n, \\ z_n &= J_{\gamma \mathsf{A}} y_n, \\ r_n &= z_n - \gamma \mathsf{B} z_n, \\ x_{n+1} &= \mathsf{P}_{\mathsf{C}}\big(x_n - y_n + r_n\big), \end{aligned}$$

where P_{C} is the metric projection onto C.
Then the following hold:

(i) $(x_n - z_n)_{n \in \mathbb{N}}$ converges strongly to 0.

(ii) Suppose that A or B is uniformly monotone on every nonempty bounded subset of $\mathsf{dom}(\mathsf{A})$. Then $(x_n)_{n \in \mathbb{N}}$ and $(z_n)_{n \in \mathbb{N}}$ converge strongly to the unique point in $\mathsf{C} \cap \mathsf{zer}(\mathsf{A} + \mathsf{B})$.

Theorem 26.17 in [1] has the assumption that $\mathsf{A} + \mathsf{B}$ is maximally monotone. As we will see, we do not need this assumption to prove the statements in Theorem 2.1. It is used in [1] to show the additional assertion that $(x_n)_{n \in \mathbb{N}}$ and $(z_n)_{n \in \mathbb{N}}$ converge weakly toward a point in $\mathsf{C} \cap \mathsf{zer}(\mathsf{A} + \mathsf{B})$. This weak convergence is then used in the proof of Theorem 2.1*(ii)* to show the boundedness of a specific subset of $\mathsf{dom}(\mathsf{A})$. We will show this boundedness in another way in order to quantify it and avoid the assumption of maximal monotonicity of $\mathsf{A} + \mathsf{B}$. Furthermore, we will avoid the definition of the subset D in the following theorems. We can instead claim B to be single-valued and $\frac{1}{\beta}$-Lipschitzian on $\mathsf{C} \cup \mathsf{dom}(\mathsf{A})$ since we are only going to use the property on this subset of D.

2.1 Extracting Rates of Convergence and Metastability

We first consider the situation in which A or B is additionally uniformly monotone on every bounded subset of $\mathsf{dom}(\mathsf{A})$. Then $\mathsf{A} + \mathsf{B}$ is also uniformly monotone on those subsets as we will see in the next remark. Therefore, we can apply Proposition 1.5 and get a unique $x \in \mathsf{C} \cap \mathsf{zer}(\mathsf{A} + \mathsf{B})$.

Remark 2.2 (Modulus of Uniform Monotonicity). *Let $\mathsf{A}, \mathsf{B} : \mathcal{H} \to 2^{\mathcal{H}}$ be monotone operators defined as in Tseng's Algorithm and C be a closed and convex subset of $\mathcal{H}$ such that $\mathsf{C} \cap \mathsf{zer}(\mathsf{A} + \mathsf{B}) \neq \emptyset$ holds. We consider an arbitrary subset $S \subseteq \mathsf{dom}(\mathsf{A})$. First, we want to argue that if A or B is uniformly monotone on S then $\mathsf{A} + \mathsf{B}$ is as well. Let ϕ be the function from the definition of uniform monotonicity for A or B on S as described in Definition 1.4. Let $s, t \in S$ and $(s, u), (t, v) \in \mathsf{gra}(\mathsf{A} + \mathsf{B})$. Then, by definition of $\mathsf{A} + \mathsf{B}$, there are $u_a \in \mathsf{A}(s), v_a \in \mathsf{A}(t)$ and $u_b \in \mathsf{B}(s), v_b \in \mathsf{B}(t)$ such that $u_a + u_b = u$ and $v_a + v_b = v$ hold. We obtain*

$$\langle s - t, u - v \rangle = \langle s - t, u_a + u_b - v_a - v_b \rangle = \langle s - t, u_a - v_a \rangle + \langle s - t, u_b - v_b \rangle \geq \phi(\|s - t\|)$$

since one of the scalar products is $\geq \phi(\|s - t\|)$ (depending on which operator is uniformly monotone) and the other one is ≥ 0 (because both are monotone). This shows that $\mathsf{A} + \mathsf{B}$ is uniformly monotone on S with the same function ϕ.

For the convergence of $(x_n)_{n \in \mathbb{N}}$ in Tseng's Algorithm we actually only need the uniform monotonicity of $\mathsf{A} + \mathsf{B}$ on one specific bounded subset of $\mathcal{H}$. By the comment made after Definition 1.7, we may assume the existence of some modulus Θ of uniform monotonicity with property (2.2.1). In the premise of the next theorem we only demand such a Θ with a suitably chosen K. In the proof of the theorem we will then show that this K is actually suitable.

Theorem 2.3. *Let $\mathsf{C} \subseteq \mathcal{H}$ be nonempty, closed and convex and let $\beta \in (0, \infty)$. Assume that $\mathsf{A} : \mathcal{H} \to 2^{\mathcal{H}}$ is a maximally monotone operator and $\mathsf{B} : \mathcal{H} \to 2^{\mathcal{H}}$ a monotone operator which is single-valued and $\frac{1}{\beta}$-Lipschitz continuous on $\mathsf{C} \cup \mathsf{dom}(\mathsf{A})$. Let C be such that $\mathsf{C} \cap \mathsf{zer}(\mathsf{A} + \mathsf{B}) \neq \emptyset$. Take $x_0 \in \mathsf{C}$, $\gamma \in (0, \beta)$ and $x \in \mathsf{C} \cap \mathsf{zer}(\mathsf{A} + \mathsf{B})$. Let the sequences $(x_n)_{n \in \mathbb{N}}$ and $(z_n)_{n \in \mathbb{N}}$ be defined as in Theorem 2.1. Assume that $\mathsf{b} \in \mathbb{N}$ is such that $\beta \leq \mathsf{b}$*

and $\mathsf{k} \in \mathbb{N}, \mathsf{k} \geq 1$ *is such that* $\frac{1}{\mathsf{k}} \leq \gamma \leq \beta - \frac{1}{\mathsf{k}}$ *(witnessing the strict inequalities in* $\gamma \in (0, \beta)$*). Let* $\mathsf{d}, \mathsf{M} \in \mathbb{N}$ *be such that*

$$\|x_0 - x\| \leq \mathsf{d} \ \ \wedge \ \ \|x_0 - z_0\| \leq \mathsf{M}.$$

Assume we have some modulus Θ *of uniform monotonicity at the zero* x *for the operator* $\mathsf{A} + \mathsf{B}$ *on the bounded subset* $\mathsf{D} := \{z_n : n \in \mathbb{N}\}$*, i.e.*

$$(\Diamond_1) \qquad \forall l \in \mathbb{N}\, \forall y \in \mathsf{D}\, \forall (y, u) \in \mathsf{gra}(\mathsf{A} + \mathsf{B}) \ \Big(\|x - y\| \in [2^{-l}, \mathsf{M} + 5\mathsf{d}] \to \langle y - x, u \rangle \geq 2^{-\Theta(l)}\Big).$$

Then (x_n) *converges to* x *with the following rate of convergence.*

$$\forall l \in \mathbb{N}\, \forall n \geq \varphi(\mathsf{d}, \mathsf{b}, \mathsf{k}, \Theta, \mathsf{M}, l) \ \Big(\|x_n - x\| < 2^{-l}\Big),$$

where

$$\varphi(\mathsf{d}, \mathsf{b}, \mathsf{k}, \Theta, \mathsf{M}, l) := \left\lceil \frac{\mathsf{d}^2\mathsf{b}^2\mathsf{k}^2}{3(\varepsilon(\mathsf{d}, \mathsf{k}, \Theta, \mathsf{M}, l))^2} \right\rceil,$$
$$\varepsilon(\mathsf{d}, \mathsf{k}, \Theta, \mathsf{M}, l) := \min\left\{2^{-(l+1)}, \frac{2^{-\Theta(l+1)}}{2\mathsf{k} \cdot (\mathsf{M} + 5\mathsf{d})}\right\}.$$

Proof. Let $l \in \mathbb{N}$ be arbitrary. Because $\mathsf{d}, \mathsf{b}, \mathsf{k}, \mathsf{M}, l$ and Θ are fixed throughout the proof, we will write φ for $\varphi(\mathsf{d}, \mathsf{b}, \mathsf{k}, \Theta, \mathsf{M}, l)$ and ε for $\varepsilon(\mathsf{d}, \mathsf{k}, \Theta, \mathsf{M}, l)$ from now on. Utilizing the bounds for γ and β, we obtain the following estimate:

$$\frac{\gamma^2}{\beta^2} \leq \frac{(\beta - \frac{1}{\mathsf{k}})^2}{\beta^2} = \frac{\beta^2 - \frac{2\beta}{\mathsf{k}} + \frac{1}{\mathsf{k}^2}}{\beta^2} = 1 - \frac{2}{\beta\mathsf{k}} + \frac{1}{\mathsf{k}^2\beta^2}$$

and so

$$(2.3.1) \qquad 1 - \frac{\gamma^2}{\beta^2} \geq 1 - \Big(1 - \frac{2}{\mathsf{k}\beta} + \frac{1}{\mathsf{k}^2\beta^2}\Big) = \frac{2}{\mathsf{k}\beta} - \frac{1}{\mathsf{k}^2\beta^2} = \frac{2\mathsf{k}\beta - 1}{\mathsf{k}^2\beta^2} \geq \frac{4\mathsf{k} \cdot \frac{1}{\mathsf{k}} - 1}{\mathsf{k}^2\mathsf{b}^2} = \frac{3}{\mathsf{k}^2\mathsf{b}^2} > 0.$$

The following inequality is shown in (26.71) in [1]:

$$(2.3.2) \qquad \forall n \in \mathbb{N} \ \Big(\|x_{n+1} - x\|^2 \leq \|x_n - x\|^2 - (1 - \frac{\gamma^2}{\beta^2})\|x_n - z_n\|^2\Big).$$

Since $\big(1 - \frac{\gamma^2}{\beta^2}\big) > 0$ holds by (2.3.1), one of the consequences of this statement is the monotonicity of $(\|x_n - x\|)_{n\in\mathbb{N}}$:

$$(2.3.3) \qquad \forall n \in \mathbb{N}\, \forall m \leq n \ (\|x_n - x\| \leq \|x_m - x\| \leq \|x_0 - x\|).$$

We now show that D is bounded and compute a bound N such that $\|x - y\| \leq N$ holds for all $y \in \mathsf{D}$. Since $z_n = \mathrm{J}_{\gamma\mathsf{A}}\, y_n \in \mathsf{dom}(\mathsf{A})$ holds for all $n \in \mathbb{N}$ and $x \in \mathsf{dom}(\mathsf{A})$ is true, D is a subset of $\mathsf{dom}(\mathsf{A})$. We use that A, and therefore also $\gamma\mathsf{A}$, is maximal monotone. Thus, by Proposition 1.11, $\mathrm{J}_{\gamma\mathsf{A}} : \mathcal{H} \to \mathcal{H}$ is nonexpansive. Since B is Lipschitz-continuous on $\{x_n\}_{n\in\mathbb{N}} \subseteq \mathsf{C}$, we get for all $n \in \mathbb{N}$

$$\begin{aligned}
\|z_n - x\| &= \|J_{\gamma\mathsf{A}} y_n - x\| = \|\,\mathrm{J}_{\gamma\mathsf{A}}\, y_n - \mathrm{J}_{\gamma\mathsf{A}}\, y_0 + \mathrm{J}_{\gamma\mathsf{A}}\, y_0 - x_0 + x_0 - x\| \\
&\leq \|\,\mathrm{J}_{\gamma\mathsf{A}}\, y_n - \mathrm{J}_{\gamma\mathsf{A}}\, y_0\| + \underbrace{\|\,\mathrm{J}_{\gamma\mathsf{A}}\, y_0 - x_0\|}_{=\|z_0 - x_0\|} + \|x_0 - x\| \\
(2.3.5) \qquad &\leq \|y_n - y_0\| + \mathsf{M} + \mathsf{d} = \|x_n - \gamma\mathsf{B}x_n - x_0 + \gamma\mathsf{B}x_0\| + \mathsf{M} + \mathsf{d} \\
&\leq \|x_n - x_0\| + \gamma\|\mathsf{B}x_n - \mathsf{B}x_0\| + \mathsf{M} + \mathsf{d} \leq (1 + \frac{\gamma}{\beta})\|x_n - x_0\| + \mathsf{M} + \mathsf{d} \\
&\leq 2\|x_n - x + x - x_0\| + \mathsf{M} + \mathsf{d} \leq 2\|x_n - x\| + \mathsf{M} + 3\mathsf{d} \\
&\overset{(2.3.3)}{\leq} 2\|x_0 - x\| + \mathsf{M} + 3\mathsf{d} \leq \mathsf{M} + 5\mathsf{d}.
\end{aligned}$$

So we have shown that $\mathsf{M}+5\mathsf{d}$ was suitably chosen in the property $(\diamondsuit_1)$ of the modulus Θ.
Another consequence of (2.3.2) is the following statement:

$$\forall n \in \mathbb{N}\, \forall m \leq n \left(\sum_{i=m}^{n} \|x_i - z_i\|^2 \leq \frac{\|x_0 - x\|^2}{1-\frac{\gamma^2}{\beta^2}} \right). \tag{2.3.4}$$

We now can finally confirm φ as the rate of convergence. First, we want to prove that there is some $\bar{n} \leq \varphi$ such that $\|x_{\bar{n}} - z_{\bar{n}}\| < \varepsilon$ holds. Assume on the contrary that $\|x_i - z_i\| \geq \varepsilon$ holds for all $i \leq \varphi$. Then

$$\sum_{i=0}^{\varphi} \|x_i - z_i\|^2 \geq (\varphi+1)\varepsilon^2 > \varphi\varepsilon^2 = \left\lceil \frac{\mathsf{d}^2\mathsf{b}^2\mathsf{k}^2}{3\varepsilon^2} \right\rceil \cdot \varepsilon^2 \geq \frac{\mathsf{d}^2\mathsf{k}^2\mathsf{b}^2}{3} \geq \frac{\|x_0 - x\|^2}{\frac{3}{\mathsf{k}^2\mathsf{b}^2}} \overset{(2.3.1)}{\geq} \frac{\|x_0 - x\|^2}{1-\frac{\gamma^2}{\beta^2}}$$

which contradicts (2.3.4). Hence, we can fix some $\bar{n} \leq \varphi$ such that $\|x_{\bar{n}} - z_{\bar{n}}\| < \varepsilon$ holds. We define $u_n := \gamma^{-1}(x_n - z_n) + (\mathsf{B}z_n - \mathsf{B}x_n)$ and utilize the Lipschitz-continuity of B to find a bound for $\|u_n\|$. By the definition of $(x_n)_{n\in\mathbb{N}}$ and $(z_n)_{n\in\mathbb{N}}$ we have $\{x_n\}_{n\in\mathbb{N}} \subseteq \mathsf{C}$ and $\{z_n\}_{n\in\mathbb{N}} \subseteq \operatorname{ran} \mathsf{J}_{\gamma\mathsf{A}} = \operatorname{dom}(\mathsf{A})$. Therefore, B is $\frac{1}{\beta}$-Lipschitz continuous on $\{x_n\}_{n\in\mathbb{N}} \cup \{z_n\}_{n\in\mathbb{N}}$ and we deduce:

$$\forall n \in \mathbb{N} \quad \left(\|u_n\| = \|\gamma^{-1}(x_n - z_n) + (\mathsf{B}z_n - \mathsf{B}x_n)\| \leq \left(\frac{1}{\gamma} + \frac{1}{\beta}\right)\|x_n - z_n\| \right). \tag{2.3.6}$$

By assumption, we have that $0 \in (\mathsf{A}+\mathsf{B})(x)$. Moreover, $\gamma^{-1}(x_n - z_n) - \mathsf{B}x_n = \gamma^{-1}(x_n - \gamma\mathsf{B}x_n - z_n) = \gamma^{-1}(y_n - z_n)$ holds for all $n \in \mathbb{N}$ by definition as well. We can apply Proposition 1.9 *(ii)* and obtain $\gamma^{-1}(y_n - z_n) = \gamma^{-1}(y_n - \mathsf{J}_{\gamma\mathsf{A}}\, y_n) \in \mathsf{A}z_n$ for all $n \in \mathbb{N}$. Hence, we have $u_n = \gamma^{-1}(x_n - z_n) - \mathsf{B}x_n + \mathsf{B}z_n \in (\mathsf{A}+\mathsf{B})z_n$ for all $n \in \mathbb{N}$. Now we can use (2.3.6) and the modulus Θ to estimate $\|z_{\bar{n}} - x\|$. Using the assumptions on ε, γ and β, we first obtain:

$$\|u_{\bar{n}}\| \overset{(2.3.6)}{\leq} \left(\frac{1}{\gamma} + \frac{1}{\beta}\right)\|x_{\bar{n}} - z_{\bar{n}}\| < \left(\frac{1}{\gamma} + \frac{1}{\beta}\right)\varepsilon \leq 2\gamma^{-1} \cdot \frac{2^{-\Theta(l+1)}}{2\mathsf{k}\cdot(\mathsf{M}+5\mathsf{d})} \leq 2\mathsf{k} \cdot \frac{2^{-\Theta(l+1)}}{2\mathsf{k}\cdot(\mathsf{M}+5\mathsf{d})} = \frac{2^{-\Theta(l+1)}}{\mathsf{M}+5\mathsf{d}}$$

and, therefore,

$$(\mathsf{M}+5\mathsf{d}) \cdot \|u_{\bar{n}}\| < 2^{-\Theta(l+1)}.$$

Applying the Cauchy-Schwartz inequality, we get

$$\langle z_{\bar{n}} - x, u_{\bar{n}} \rangle \leq \|z_{\bar{n}} - x\|\|u_{\bar{n}}\| \overset{(2.3.5)}{\leq} (\mathsf{M}+5\mathsf{d}) \cdot \|u_{\bar{n}}\| < 2^{-\Theta(l+1)}.$$

This implies

$$\|z_{\bar{n}} - x\| < 2^{-(l+1)}$$

by $(\diamondsuit_1)$ since $\|z_{\bar{n}} - x\| > \mathsf{M}+5\mathsf{d}$ is not possible by (2.3.5). Putting everything together, we get

$$\|x_{\bar{n}} - x\| \leq \|x_{\bar{n}} - z_{\bar{n}}\| + \|z_{\bar{n}} - x\| < \varepsilon + 2^{-(l+1)} \leq 2^{-(l+1)} + 2^{-(l+1)} = 2^{-l}.$$

For $n \geq \varphi$ it follows now easily that $\|x_n - x\| \leq \|x_{\bar{n}} - x\| < 2^{-l}$ holds by (2.3.3) since we have $n \geq \varphi \geq \bar{n}$. ☐

Corollary to the proof: If we take as an additional input $L > 1$ with $\frac{1}{L} \leq 1 - \frac{\gamma^2}{\beta^2}$ we can rewrite φ as $\varphi = \left\lceil \frac{\mathsf{d}^2 L}{(\varepsilon(\mathsf{d},\mathsf{k},\Theta,\mathsf{M},l))^2} \right\rceil$ which is numerically better than our previous rate if β is large compared to γ.

We can now use this result to also obtain a rate of convergence for $(x_n - z_n)_{n\in\mathbb{N}}$. Note that this rate heavily depends on the modulus Θ of uniform monotonicity and, therefore, does not hold for the general case of Theorem 2.1 *(i)*.

Corollary 2.4. *Under the assumptions of Theorem 2.3 the following holds*

$$\forall l \in \mathbb{N}\, \forall n \geq \varphi(\mathsf{d},\mathsf{b},\mathsf{k},\Theta,\mathsf{M},m(l,\mathsf{k},\mathsf{b})) \left(\|x_n - z_n\| < 2^{-l} \right) \text{ with } \varphi \text{ as before and}$$

where

$$m(l,\mathsf{k},\mathsf{b}) := \lceil l + \log_2(\mathsf{kb}) - \log_2(\sqrt{3}) \rceil.$$

Proof. Let $n \geq \psi(\mathsf{d},\mathsf{b},\mathsf{k},\Theta,\mathsf{M},m(l,\mathsf{k},\mathsf{b}))$ be arbitrary. By Theorem 2.3 and definition of ψ we know that $\|x_n - x\| < 2^{-m(l,\mathsf{k},\mathsf{b})}$ holds. We can apply (2.3.2) to obtain

$$\|x_n - z_n\| \overset{(2.3.2)}{\leq} \frac{\|x_n - x\|}{(1-\frac{\gamma^2}{\beta^2})^{\frac{1}{2}}} \overset{(2.3.1)}{\leq} \frac{\|x_n - x\|}{\left(\frac{3}{\mathsf{k}^2\mathsf{b}^2}\right)^{\frac{1}{2}}} < 2^{-\lceil l+\log_2(\mathsf{kb})-\log_2(\sqrt{3})\rceil} \cdot \frac{\mathsf{kb}}{\sqrt{3}} \leq 2^{-l} \cdot 2^{-\log_2(\mathsf{kb})} \cdot 2^{\log_2(\sqrt{3})} \cdot \frac{\mathsf{kb}}{\sqrt{3}} = 2^{-l}.$$

□

Corollary to the proof: Again, if we use an additional input $L > 1$ with $\frac{1}{L} \leq 1 - \frac{\gamma^2}{\beta^2}$ we can replace $m(l,\mathsf{k},\mathsf{b}))$ by $m(l,L) := \lceil l + \sqrt{L}\rceil$.

Remark 2.5. *(i) In Theorem 2.3 we can assume the inequality*

$$(2.3.2) \qquad \forall n \in \mathbb{N} \ \left(\|x_{n+1} - x\|^2 \leq \|x_n - x\|^2 - (1 - \frac{\gamma^2}{\beta^2})\|x_n - z_n\|^2\right).$$

instead of assuming that B *is monotone. This is based on the fact that, once we require* $\mathsf{A}+\mathsf{B}$ *to be uniformly monotone with modulus* Θ*, the monotonicity of* B *is only used to show this inequality.*

(ii) We have now found a rate of convergence for $(x_n)_{n\in\mathbb{N}}$ *in Theorem 2.3. In Theorem 2.1(ii) there is also mention of the convergence of the sequence* $(z_n)_{n\in\mathbb{N}}$ *toward the same point* x*. We can now easily obtain a rate of convergence for this by putting together Theorem 2.3 and Corollary 2.4. Let* $\varphi(\mathsf{d},\mathsf{b},\mathsf{k},\Theta,\mathsf{M},l)$ *be defined as in Theorem 2.3 and* $\psi(\mathsf{d},\mathsf{b},\mathsf{k},\Theta,\mathsf{M},l)$ *be defined as in Corollary 2.4. Then by an easy triangle inequality we obtain*

$$\forall l \in \mathbb{N}\, \forall n \geq \widetilde{\varphi}(\mathsf{d},\mathsf{b},\mathsf{k},\Theta,\mathsf{M},l) \ \left(\|z_n - x\| < 2^{-l}\right)$$

for $\widetilde{\varphi}(\mathsf{d},\mathsf{b},\mathsf{k},\Theta,\mathsf{M},l) := \max\{\varphi(\mathsf{d},\mathsf{b},\mathsf{k},\Theta,\mathsf{M},l+1), \psi(\mathsf{d},\mathsf{b},\mathsf{k},\Theta,\mathsf{M},l+1)\}$.

We now want to take another look at the convergence of $(x_n - z_n)_{n\in\mathbb{N}}$. As mentioned, the rate of convergence in Corollary 2.4 for this sequence heavily depends on the modulus Θ of uniform monotonicity, which we assumed for $\mathsf{A}+\mathsf{B}$. By Theorem 2.1*(i)*, however, the convergence is shown without this additional assumption using the inequality (2.3.2). More precisely, the derived statement (2.3.4) shows that for arbitrary $\varepsilon > 0$ there can only be finitely many $i \in \mathbb{N}$ such that $\|x_i - z_i\| \geq \varepsilon$ holds. However, it does not indicate for which $i \in \mathbb{N}$ this happens for the last time. If $(\|x_n - x\|)_{n\in\mathbb{N}}$ converges towards 0, we can proceed as in Corollary 2.4 to obtain a rate of convergence for $(x_n - z_n)_{n\in\mathbb{N}}$. However, without the uniform monotonicity of $\mathsf{A}+\mathsf{B}$ we can in general not show the convergence of $(x_n)_{n\in\mathbb{N}}$. Since $(x_n - z_n)_{n\in\mathbb{N}}$ itself is in general not monotone, we can only extract a rate of metastability for this sequence instead of a rate of convergence in that case. The concept of 'metastability' for the convergence of $\|x_n - z_n\| \to 0$ refers to the (nonconstructively equivalent) reformulation as

$$\forall l \in \mathbb{N}\, \forall g : \mathbb{N} \to \mathbb{N}\ \exists n \in \mathbb{N}\ \forall i \in [n, n+g(n)] \ \left(\|x_i - z_i\| < 2^{-l}\right),$$

where $[n, n+g(n)] := \{i \in \mathbb{N} : n \leq i \leq n+g(n)\}$.
By a rate of metastability we mean a bound on '$\exists n$'. Since the metastable form of convergence implies the convergence only noneffectively, it in general is not possible to effectively convert a rate of metastability into a rate of convergence. For the history of the concept of metastability, which goes back to G. Kreisel in 1951, was used in proof mining since 2004 and was rediscovered by T. Tao in 2007 under the name 'metastability', see [9]. The following theorem proves a rate of metastability for $(x_n - z_n)_{n\in\mathbb{N}}$.

Theorem 2.6. *Let* $\mathsf{C} \subseteq \mathcal{H}$ *nonempty, closed and convex and let* $\beta \in (0,\infty)$*. Assume that* $\mathsf{A} : \mathcal{H} \to 2^{\mathcal{H}}$ *is a maximally monotone operator and* $\mathsf{B} : \mathcal{H} \to 2^{\mathcal{H}}$ *a monotone operator which is single-valued and* $\frac{1}{\beta}$*-Lipschitz continuous on* $\mathsf{C}\cup\mathsf{dom}(\mathsf{A})$*. Let* C *be such that* $\mathsf{C}\cap\mathsf{zer}(\mathsf{A}+\mathsf{B}) \neq \emptyset$*. Take* $x_0 \in \mathsf{C}$, $\gamma \in (0,\beta)$ *and* $x \in \mathsf{C}\cap\mathsf{zer}(\mathsf{A}+\mathsf{B})$ *and let* k, b, d, M *be as before.*
Then $(x_n - z_n)_{n\in\mathbb{N}}$ *converges strongly to 0 with the following rate of metastability*

$$\forall l \in \mathbb{N}\, \forall g : \mathbb{N} \to \mathbb{N}\ \exists n \leq \widetilde{g}^{\left(\lceil \frac{\mathsf{d}^2\mathsf{k}^2\mathsf{b}^2\cdot 2^{2l}}{3}\rceil\right)}(0)\ \forall i \in [n, n+g(n)] \ \left(\|x_i - z_i\| < 2^{-l}\right),$$

where $\tilde{g}(n) := n + g(n) + 1$.

Proof. By (2.3.3) and the assumption that $\|x_0 - x\| \leq \mathsf{d}$ holds, $(\|x_n - x\|^2)_{n\in\mathbb{N}}$ is a nondecreasing sequence in $[0, \mathsf{d}^2]$. We can therefore use Proposition 2.27 and Remark 2.29.1) in [6] to obtain the following statement:

$$(\dagger) \qquad \forall \delta > 0\, \forall g : \mathbb{N} \to \mathbb{N}\; \exists n \leq \phi(g, \delta, \mathsf{d})\, \forall i \in [n, n+g(n)] \left(|\|x_i - x\|^2 - \|x_{i+1} - x\|| < \delta \right)$$

with
$$\phi(g, \delta, \mathsf{d}) := \widetilde{g}^{(\lceil \mathsf{d}^2 \cdot \delta^{-1} \rceil)}(0) \text{ and } \widetilde{g}(n) := n + g(n) + 1.$$

Moreover, we can use (2.3.1) and (2.3.2) to estimate

$$\|x_n - z_n\|^2 \leq \frac{\mathsf{k}^2\mathsf{b}^2}{3}\left(\left| \|x_n - x\|^2 - \|x_{n+1} - x\|^2 \right| \right)$$

for all $n \in \mathbb{N}$. The rate of metastability in (†) can now be converted into one for $(\|x_n - z_n\|)_{n\in\mathbb{N}}$ by setting $\delta := \frac{3\cdot 2^{-2l}}{\mathsf{k}^2\mathsf{b}^2}$. More precisely, we get for all $l \in \mathbb{N}$ that there is an $m \leq \widetilde{g}^{\left(\lceil \frac{\mathsf{d}^2\mathsf{k}^2\mathsf{b}^2 \cdot 2^{2l}}{3} \rceil\right)}(0)$ such that

$$\|x_i - z_i\|^2 \leq \frac{\mathsf{k}^2\mathsf{b}^2}{3} |\|x_i - x\|^2 - \|x_{i+1} - x\|^2| < \frac{\mathsf{k}^2\mathsf{b}^2}{3}\delta = 2^{-2l}$$

holds for all $i \in [m, m + g(m)]$. □

2.2 Discussion of the Results (for logicians)

As mentioned in the introduction, general logical metatheorems can be used to guarantee in advance the possibility of extracting uniform effective bounds from proofs, in particular, in the context of nonlinear analysis and to explain their dependence on rather few data. We briefly indicate this for Theorem 2.3 (the other results can be treated similarly). Metatheorems for nonexpansive and Lipschitzian operators in Hilbert space and abstract convex subsets have been developed in [5, 4] (see also [6]) and have been extended to cover maximally monotone set-valued operators and their sums in [13, 11] (in the case at hand we do not need the bound ξ used in [11] to treat $\mathsf{A}+\mathsf{B}$ as B - essentially - is single valued). For Tseng's algorithm we also need a constant $\mathcal{P}_\mathsf{C}$ of type $\mathcal{H} \to \mathcal{H}$ for the metric projection onto C and add the axioms

$$(i) \qquad \forall x, y \in \mathcal{H} \; \left(\chi_\mathsf{C}(y) = 0 \to \|\mathcal{P}_\mathsf{C}(x) - x\| \leq \|y - x\| \right),$$

$$(ii) \qquad \forall x \in \mathcal{H} \; \left(\chi_\mathsf{C}(\mathcal{P}_\mathsf{C}(x)) = 0 \right).$$

Such a projector is majorizable as described in the introduction of [8] and it is justified in the intended interpretation if C additionally is assumed to be closed (which we could but do not need to formalize as it is only used in connection with $\mathcal{P}_\mathsf{C}$). The operator B is formalized as a single-valued operator on $\mathcal{H}$ which is assumed to be $(1/\beta)$-Lipschitzian on $\mathsf{C} \cup \mathsf{dom}(\mathsf{A})$. Given the operator in the theorem its interpretation is

$$\bar{\mathsf{B}} : \mathcal{H} \to \mathcal{H}, \bar{\mathsf{B}}(x) = \begin{cases} \mathsf{B}x & \text{if } x \in \mathsf{C} \cup \mathsf{dom}(\mathsf{A}), \\ 0 & \text{otherwise.} \end{cases}$$

This operator $\bar{\mathsf{B}}$ is in general not monotone but the monotonicity of B is used only to prove the inequality (2.3.2) and to make sure that $\mathsf{A}+\mathsf{B}$ is uniformly monotone if A is uniformly monotone. The modulus Θ of uniform monotonicity with property $(\diamondsuit_1)$ ensures the uniform monotonicity of $\mathsf{A}+\mathsf{B}$ directly. Therefore, we actually do not need to incorporate the monotonicity of B if we assume (2.3.2) to be true. This means that we can introduce B by a constant $\mathcal{B}$ for a single-valued operator instead of introducing the set-valued operator B whose Lipschitz-property is only specified on $\mathsf{C} \cup \mathsf{dom}(\mathsf{A})$ and whose intended interpretation is given by $\bar{\mathsf{B}}$. Lastly, the sum of A and B can be incorporated with the help of [11] by adding constants and axioms as described in section 2(d). In the following, let $\chi_{\mathsf{A}+\mathsf{B}}(x, \cdot)$ be the constant for the characteristic function of $(\mathsf{A}+\mathsf{B})(x)$ for $x \in \mathcal{H}$.
Now we can express Theorem 2.3 with the help of the introduced constants in a form where we can apply the logical metatheorems. These metatheorems require the axioms in the premise of a statement to be universal. Therefore, we note that the axiom $(\diamondsuit_1)$ in Theorem 2.3 is actually of the right form (provided that we reformulate it in an inessential way by replacing $[2^{-l}, \mathsf{M} + 5\mathsf{d}]$ by $(2^{-l}, \mathsf{M} + 5\mathsf{d} + 1)$). Thus, we can ensure the uniform monotonicity of $\mathsf{A}+\mathsf{B}$ in the following expression by including this axiom in the premise which in the formal

setting of the aforementioned logical metatheorems reads as follows:

$$(\bigstar)\begin{cases}\forall l,\mathsf{b},\mathsf{k}\in\mathbb{N},\Theta\in\mathbb{N}^{\mathbb{N}}\,\forall\gamma,\beta\in[0,\mathsf{b}]\forall x,x_0\in\mathcal{H}\Big(\mathsf{k}\geq 1\wedge\frac{1}{\mathsf{k}}\leq\gamma\leq\beta-\frac{1}{\mathsf{k}}\wedge\chi_{\mathsf{A}+\mathsf{B}}(x,0)=0\wedge\chi_{\mathsf{C}}(x)=0\wedge(\diamondsuit_1)\\ \wedge\forall z,z',v,v'\in\mathcal{H}\Big(\Big(\chi_{\mathsf{C}}(z)=0\vee\chi_{\mathsf{A}}(z,v)=0\Big)\wedge\Big(\chi_{\mathsf{C}}(z')=0\vee\chi_{\mathsf{A}}(z',v')=0\Big)\to\|\mathcal{B}z-\mathcal{B}z'\|\leq\frac{1}{\beta}\|z-z'\|\Big)\\ \wedge\forall n\in\mathbb{N}\Big(\|x_{n+1}-x\|^2\leq\|x_n-x\|^2-\Big(1-\frac{\gamma^2}{\beta^2}\Big)\|x_n-z_n\|^2\Big)\wedge\chi_{\mathsf{C}}(x_0)=0\longrightarrow\exists m\in\mathbb{N}\ \ \|x_m-x\|<2^{-l}\Big),\end{cases}$$

where the constructions of the sequences $(x_n),(y_n),(z_n),(r_n)$ are explicitly hardwired into the formal system using the recursor constants.
Note that the existence of a bound on $\exists m\in\mathbb{N}$ is enough to show that the sequence converges since $(\|x_n-x\|)_{n\in\mathbb{N}}$ is monotone.
The expressions in the premise of $(\bigstar)$ are all universal and the conclusion is existential and so the aforementioned logical metatheorems can be applied for the extraction of an effective uniform bound on '$\exists m\in\mathbb{N}$' which only depends on $l,\mathsf{b},\mathsf{k},\Theta$, a norm bound $\mathsf{d}\geq\|x-x_0\|,\|x_0\|$ and majorizing data for $\mathrm{J}_{\gamma\mathsf{A}}$ and $\gamma\mathsf{B}$ (in fact for B which then via the upper bound b on γ also yields a majorant for $\gamma\mathsf{B}$). As shown in [13] and [5], resp., such majorants can be obtained from upper bounds on the displacements $\|\mathrm{J}_{\gamma\mathsf{A}}(s)-s\|$ and $\|\gamma\mathsf{B}(t)-t\|$ of $\mathrm{J}_{\gamma\mathsf{A}}$ and $\gamma\mathsf{B}$ in some points s,t (and hence - by the Lipschitz properties of these mappings - in all points). As our Theorem 2.3 shows we, in the special case at hand, only need this in the weak form of the bound $\mathsf{M}\geq\|x_0-z_0\|$ which is definable in such data and d via

$$\|x_0-z_0\|=\|\mathrm{J}_{\gamma\mathsf{A}}(x_0-\gamma\mathsf{B}x_0)-x_0\|\leq\|\mathrm{J}_{\gamma\mathsf{A}}(x_0-\gamma\mathsf{B}x_0)-(x_0-\gamma\mathsf{B}x_0)\|+\|\gamma\mathsf{B}x_0-x_0\|+\|x_0\|.$$

Moreover, given M, we only need an upper bound on $\|x-x_0\|$ but not an additional upper bound on $\|x_0\|$.
In the previous part we did not only give rates of convergence and metastability for Tseng's Algorithm but also had some noteworthy results regarding the premises of the algorithm. More precisely, some of the assumptions in the theorem can be chosen weaker and we still get the same (strong) convergence statements. We had already seen that we do not need the operator $\mathsf{A}+\mathsf{B}$ to be maximally monotone to show the strong convergence statements of Tseng's Algorithm (Theorem 26.17 in [1]). Another observation, which we already mentioned briefly in Remark 2.5(i), regards the monotonicity of the operator B. Instead of the assumption that B is monotone, we can actually suppose that the inequality

$$\forall n\in\mathbb{N}\ \ \Big(\|x_{n+1}-x\|^2\leq\|x_n-x\|^2-\Big(1-\frac{\gamma^2}{\beta^2}\Big)\|x_n-z_n\|^2\Big)\tag{2.3.2}$$

holds. Note that in the proof of Theorem 26.17 in [1] it is shown that this inequality holds in particular if B is monotone.
Finally, there is one last observation we can make. For the strong convergence of the sequence $(x_n)_{n\in\mathbb{N}}$ toward the unique point in $\mathsf{C}\cap\mathsf{zer}(\mathsf{A}+\mathsf{B})$ Theorem 2.1 assumed additionally that either A or B are uniformly monotone on every bounded subset of $\mathsf{dom}(\mathsf{A})$. We have seen that we can instead require the sum $\mathsf{A}+\mathsf{B}$ of these operators to be uniformly monotone with modulus Θ with property $(\diamondsuit_1)$ on the specific subset $\{z_n\}_{n\in\mathbb{N}}$ of $\mathsf{dom}(\mathsf{A})$ which by Remark 2.2 is a consequence of the original assumption.

3 Forward-Backward Splitting Algorithm

Next we are going to consider an algorithm which works in a very similar setting as Tseng's Splitting Algorithm. We impose a cocoercivity condition on one of the operators which allows us to have less computations in each iteration. The resulting algorithm is known as a forward-backward algorithm that alternates an explicit step using the cocoercive operator with an implicit resolvent step involving the second operator.

Theorem 3.1 (Forward-Backward Algorithm, 26.14 in [1]). *Let $\beta\in(0,\infty)$ and $\gamma\in(0,2\beta)$ and set $\delta=2-\frac{\gamma}{2\beta}$. Assume that $\mathsf{A}:\mathcal{H}\to 2^{\mathcal{H}}$ is a maximally monotone operator and $\mathsf{B}:\mathcal{H}\to\mathcal{H}$ is a β-cocoercive operator and suppose $\mathsf{zer}(\mathsf{A}+\mathsf{B})\neq\emptyset$. Furthermore, let $(\lambda_n)_{n\in\mathbb{N}}\subseteq[0,\delta]$ be a sequence such that $\sum_{n\in\mathbb{N}}\lambda_n(\delta-\lambda_n)=\infty$ and let $x_0\in\mathcal{H}$. For $n\in\mathbb{N}$ we set*

$$y_n=x_n-\gamma\mathsf{B}x_n,$$
$$x_{n+1}=x_n+\lambda_n(\mathrm{J}_{\gamma\mathsf{A}}\,y_n-x_n).$$

Then the following hold:

(i) Let $x \in \mathsf{zer}(\mathsf{A}+\mathsf{B})$. Then $(\mathsf{B}x_n)_{n\in\mathbb{N}}$ converges strongly to the unique dual solution $\mathsf{B}x$.

(ii) Suppose that one of the following holds:

a) A is uniformly monotone on every nonempty bounded subset of $\mathsf{dom}(\mathsf{A})$.

b) B is uniformly monotone on every nonempty bounded subset of $\mathcal{H}$.

Then $(x_n)_{n\in\mathbb{N}}$ converges strongly to the unique point in $\mathsf{zer}(\mathsf{A}+\mathsf{B})$.

As for Tseng's Algorithm, we omit a statement which is proven in Theorem 26.14 in [1] regarding the weak convergence of $(x_n)_{n\in\mathbb{N}}$ towards a point in $\mathsf{zer}(\mathsf{A}+\mathsf{B})$. This weak convergence is used in the proof of Theorem 3.1*(ii)* afterwards to show the boundedness of specific subsets. We will show this again in another way to be able to quantify it and to avoid using this weak convergence.

3.1 Extracting Rates of Convergence

First, we want to find a rate of convergence for $(\mathsf{B}x_n - \mathsf{B}x)_{n\in\mathbb{N}}$. Looking at the proof for Theorem 3.1*(i)* in [1], we observe that the convergence of this sequence is shown with the help of another sequence which involves an operator T defined by $\mathsf{T} := \mathsf{J}_{\gamma\mathsf{A}} \circ (\mathsf{Id} - \gamma\mathsf{B})$. We construct a rate of convergence for $(\mathsf{T}x_n - x_n)_{n\in\mathbb{N}}$ and convert it into one for $(\mathsf{B}x_n - \mathsf{B}x)_{n\in\mathbb{N}}$ afterwards. The following theorem gives such a rate for $(\mathsf{T}x_n - x_n)_{n\in\mathbb{N}}$.

Theorem 3.2. *Let $\beta \in (0,\infty)$, $\gamma \in (0, 2\beta)$ and set $\delta = 2 - \frac{\gamma}{2\beta}$. Assume that $\mathsf{A} : \mathcal{H} \to 2^{\mathcal{H}}$ is a maximally monotone operator and $\mathsf{B} : \mathcal{H} \to \mathcal{H}$ is a β-cocoercive operator and suppose $\mathsf{zer}(\mathsf{A}+\mathsf{B}) \neq \emptyset$. Furthermore, let $(\lambda_n)_{n\in\mathbb{N}} \subseteq [0,\delta]$ be a sequence such that $\sum_{n\in\mathbb{N}} \lambda_n(\delta - \lambda_n) = \infty$. Define $\mathsf{T} := \mathsf{J}_{\gamma\mathsf{A}} \circ (\mathsf{Id} - \gamma\mathsf{B})$ and take $x_0 \in \mathcal{H}$ and $x \in \mathsf{zer}(\mathsf{A}+\mathsf{B})$. Let the sequences $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ be defined as in Theorem 3.1. Assume that $\mathsf{b},\mathsf{k},\mathsf{d} \in \mathbb{N}, \mathsf{k} \geq 1$ are such that*

$$\beta \leq \mathsf{b} \;\wedge\; \frac{1}{\mathsf{k}} \leq \gamma \leq 2\beta - \frac{1}{\mathsf{k}} \;\wedge\; \|x_0 - x\| \leq \mathsf{d}.$$

Furthermore, let $m : \mathbb{N} \to \mathbb{N}$ be a rate of divergence for $\sum_{n\in\mathbb{N}} \lambda_n(\delta - \lambda_n) = \infty$, i.e.

$$\forall L \in \mathbb{N} \;\left(\sum_{n=0}^{m(L)} \lambda_n(\delta - \lambda_n) \geq L \right).$$

Then the following holds

$$\forall \varepsilon > 0 \,\forall n \geq \varphi(\varepsilon, \mathsf{d}, \mathsf{b}, \mathsf{k}, m) \;\Big(\|\mathsf{T}x_n - x_n\| < \varepsilon \Big),$$

where

$$\varphi(\varepsilon, \mathsf{d}, \mathsf{b}, \mathsf{k}, m) := \widetilde{m}\left(\frac{\mathsf{d}^2 + 1}{(1 + \frac{1}{2\mathsf{kb}})^2 \cdot \varepsilon^2} \right),\; \widetilde{m}(s) := m\left(\left\lceil \left(2 - \frac{1}{2\mathsf{kb}}\right)^2 \cdot s \right\rceil \right).$$

Proof. With the bounds for γ and β we get the following estimates for δ:

$$\delta = 2 - \frac{\gamma}{2\beta} \leq 2 - \frac{\frac{1}{\mathsf{k}}}{2\mathsf{b}} = 2 - \frac{1}{2\mathsf{bk}} < 2, \tag{3.2.1}$$

$$\delta = 2 - \frac{\gamma}{2\beta} \geq 2 - \frac{2\beta - \frac{1}{\mathsf{k}}}{2\beta} = 2 - 1 + \frac{1}{2\beta\mathsf{k}} \geq 1 + \frac{1}{2\mathsf{bk}} > 1. \tag{3.2.2}$$

We define the operator $\mathsf{R} := (1-\delta)\mathsf{Id} + \delta\mathsf{T}$ on $\mathcal{H}$ and the sequence $(\mu_n)_{n\in\mathbb{N}} := (\frac{1}{\delta}\lambda_n)_{n\in\mathbb{N}} \subset [0,1]$. As in [1], one shows that Theorem 1.15 can be applied Theorem to R and $(\mu_n)_{n\in\mathbb{N}}$ using that $x_n + \mu_n(\mathsf{R}x_n - x_n) = x_{n+1}$. Since we have $x \in \mathsf{zer}(\mathsf{A}+\mathsf{B})$ by assumption and $\mathsf{zer}(\mathsf{A}+\mathsf{B}) = \mathsf{Fix}(\mathsf{R})$ (which follows from Prop. 26.1(iv)(a) in [1]), we can use (1.15.2) to obtain

$$\forall m \in \mathbb{N} \;\left(\sum_{n=0}^{m} \mu_n(1-\mu_n)\|\mathsf{R}x_n - x_n\|^2 \leq \|x_0 - x\|^2 \right). \tag{3.2.3}$$

Moreover, by (1.15.3) we know that $(\|\mathsf{R}x_n - x_n\|)_{n\in\mathbb{N}}$ is monotone:

$$\forall n \in \mathbb{N} \ \Big(\|\mathsf{R}x_{n+1} - x_{n+1}\| \leq \|\mathsf{R}x_n - x_n\|\Big). \tag{3.2.4}$$

We also note that we have by definition:

$$\forall n \in \mathbb{N} \ \Big(Rx_n - x_n = (1-\delta)x_n + \delta\mathsf{T}x_n - x_n = \delta(\mathsf{T}x_n - x_n)\Big). \tag{3.2.5}$$

Now we have everything in place to confirm $\varphi(\varepsilon, \mathsf{d}, \mathsf{b}, \mathsf{k}, m)$ as the rate of convergence for $(\mathsf{T}x_n - x_n)_{n\in\mathbb{N}}$. Let $\varepsilon > 0$ be arbitrary. To prove our statement let us assume that for all $n \leq \varphi(\varepsilon, \mathsf{d}, \mathsf{b}, \mathsf{k}, m)$ we have $\|\mathsf{T}x_n - x_n\| \geq \varepsilon$. Because $\mathsf{d}, \mathsf{k}, \mathsf{b}, m$ and ε are fixed throughout the proof, we will write φ for $\varphi(\varepsilon, \mathsf{d}, \mathsf{b}, \mathsf{k}, m)$ from now on. We first show the following:

$$\forall s \geq 0 \Bigg(\sum_{n=0}^{\widetilde{m}(s)} \mu_n(1-\mu_n) \geq s \Bigg). \tag{3.2.6}$$

Using (3.2.1), we get

$$\begin{aligned}
\sum_{n=0}^{\widetilde{m}(s)} \mu_n(1-\mu_n) &= \sum_{n=0}^{\widetilde{m}(s)} \frac{1}{\delta}\lambda_n(1-\frac{1}{\delta}\lambda_n) = \frac{1}{\delta^2}\sum_{n=0}^{\widetilde{m}(s)} \lambda_n(\delta - \lambda_n) \\
&\geq \frac{1}{\big(2-\frac{1}{2\mathsf{bk}}\big)^2} \sum_{n=0}^{\widetilde{m}(s)} \lambda_n(\delta-\lambda_n) \overset{\text{Def. } \widetilde{m}}{\geq} \frac{\left\lceil \big(2-\frac{1}{2\mathsf{bk}}\big)^2 \cdot s \right\rceil}{\big(2-\frac{1}{2\mathsf{bk}}\big)^2} \geq s.
\end{aligned}$$

This proves (3.2.6). By assumption, we have $\|\mathsf{R}x_n - x_n\| \overset{(3.2.5)}{=} \delta\|\mathsf{T}x_n - x_n\| \overset{(3.2.2)}{\geq} (1+\frac{1}{2\mathsf{bk}})\|\mathsf{T}x_n - x_n\| \geq (1+\frac{1}{2\mathsf{bk}})\cdot\varepsilon$ for all $n \leq \varphi$. Using the definition of φ, we get the following estimate

$$\begin{aligned}
\sum_{n=0}^{\varphi} \mu_n(1-\mu_n)\|\mathsf{R}x_n - x_n\|^2 &\geq \Big(1+\frac{1}{2\mathsf{bk}}\Big)^2 \cdot \varepsilon^2 \sum_{n=0}^{\varphi} \mu_n(1-\mu_n) \\
&\overset{(3.2.6)}{\geq} \Big(1+\frac{1}{2\mathsf{bk}}\Big)^2 \cdot \varepsilon^2 \cdot \frac{(\mathsf{d}^2+1)}{\big(1+\frac{1}{2\mathsf{kb}}\big)^2 \cdot \varepsilon^2} = \mathsf{d}^2 + 1 \\
&> \mathsf{d}^2 \geq \|x_0 - x\|^2
\end{aligned}$$

which contradicts (3.2.3).
Thus, there has to be some $\bar{n} \leq \varphi$ such that $\|\mathsf{T}x_{\bar{n}} - x_{\bar{n}}\| < \varepsilon$ holds. Now let $n > \bar{n}$ be arbitrary. We apply the monotonicity of $(\|\mathsf{R}x_n - x_n\|)_{n\in\mathbb{N}}$ to obtain

$$\|\mathsf{T}x_n - x_n\| \overset{(3.2.5)}{=} \frac{1}{\delta}\|\mathsf{R}x_n - x_n\| \overset{(3.2.4)}{\leq} \frac{1}{\delta}\|\mathsf{R}x_{\bar{n}} - x_{\bar{n}}\| = \|\mathsf{T}x_{\bar{n}} - x_{\bar{n}}\| < \varepsilon.$$

In particular, $\|\mathsf{T}x_n - x_n\| < \varepsilon$ holds for all $n \geq \varphi$. □

This rate of convergence can now be converted into a rate of convergence for $(\mathsf{B}x_n - \mathsf{B}x)_{n\in\mathbb{N}}$ by applying a fact from [1].

Theorem 3.3. *Under the same assumptions as in Theorem 3.2 the following holds*

$$\forall \varepsilon > 0 \, \forall n \geq \psi(\varepsilon, \mathsf{d}, \mathsf{b}, \mathsf{k}, m) \ \Big(\|\mathsf{B}x_n - \mathsf{B}x\| < \varepsilon\Big),$$

where

$$\psi(\varepsilon, \mathsf{d}, \mathsf{b}, \mathsf{k}, m) := \varphi\Big(\frac{\varepsilon^2}{3\mathsf{dk}^2}, \mathsf{d}, \mathsf{b}, \mathsf{k}, m\Big), \ \textit{with } \varphi \textit{ as in Theorem 3.2.}$$

Proof. Again consider $\mathsf{T} := \mathsf{J}_{\gamma\mathsf{A}} \circ (\mathsf{Id} - \gamma\mathsf{B})$. The following inequality is proven as (26.56) in [1]:

$$\forall n \in \mathbb{N} \ \Big(\|\mathsf{B}x_n - \mathsf{B}x\|^2 \leq \frac{3}{\gamma\beta}\|x_0 - x\|\|\mathsf{T}x_n - x_n\| \Big). \tag{3.3.1}$$

Let $\varepsilon > 0$ and $n \geq \psi(\varepsilon, \mathsf{d}, \mathsf{b}, \mathsf{k}, m)$ be arbitrary. We have $\|\mathsf{T}x_n - x_n\| < \frac{\varepsilon^2}{3\mathsf{d}\mathsf{k}^2}$ by Theorem 3.2. Using (3.3.1), we conclude

$$\|\mathsf{B}x_n - \mathsf{B}x\|^2 \leq \frac{3}{\gamma\beta}\|x_0 - x\|\|\mathsf{T}x_n - x_n\| \leq \frac{3\mathsf{d}}{\frac{1}{\mathsf{k}} \cdot \frac{1}{\mathsf{k}}} \cdot \|\mathsf{T}x_n - x_n\| < 3\mathsf{d}\mathsf{k}^2 \cdot \frac{\varepsilon^2}{3\mathsf{d}\mathsf{k}^2} = \varepsilon^2.$$

□

Theorem 3.4. *Under the same assumptions as in Theorem 3.2 and $M \geq \|x_0 - Tx_0\|$ we obtain the following:*

(a) Assume we have some modulus Θ_A of uniform monotonicity for the operator A on the bounded subset $\mathsf{D} := \{\mathsf{J}_{\gamma\mathsf{A}}\, y_n\}_{n\in\mathbb{N}} \cup \{x\}$, i.e., such that

$$(\diamondsuit_2) \qquad \forall l \in \mathbb{N}\, \forall y \in \mathsf{D} \forall (x,u), (y,v) \in \mathsf{gra}(\mathsf{A}) \ \Big(\|x - y\| \in [2^{-l}, \mathsf{M} + \mathsf{d}] \to \langle x - y, u - v\rangle \geq 2^{-\Theta_\mathsf{A}(l)} \Big).$$

Then the following holds

$$\forall l \in \mathbb{N}\, \forall n \geq \tau_\mathsf{A}(l, \Theta_\mathsf{A}, \mathsf{d}, \mathsf{b}, \mathsf{k}, \mathsf{M}, m) \ \Big(\|x_n - x\| < 2^{-l} \Big),$$

where

$$\tau_\mathsf{A}(l, \Theta_\mathsf{A}, \mathsf{d}, \mathsf{b}, \mathsf{k}, \mathsf{M}, m) := \max\left\{ \psi\Big(\frac{2^{-\Theta_\mathsf{A}(l+1)}}{2(\mathsf{M}+\mathsf{d})}, \mathsf{d}, \mathsf{b}, \mathsf{k}, m\Big), \ \varphi\Big(\frac{2^{-\Theta_\mathsf{A}(l+1)}}{2\mathsf{k}(\mathsf{M}+\mathsf{d})}, \mathsf{d}, \mathsf{b}, \mathsf{k}, m\Big), \ \varphi\big(2^{-(l+1)}, \mathsf{d}, \mathsf{b}, \mathsf{k}, m\big) \right\}$$

with ψ, φ as in Theorem 3.3.

(b) Assume we have some modulus Θ_B of uniform monotonicity for the operator B on the bounded subset $\widetilde{\mathsf{D}} := \{x_n\}_{n\in\mathbb{N}} \cup \{x\}$, i.e., such that

$$(\diamondsuit_2') \qquad \forall l \in \mathbb{N}\, \forall y \in \widetilde{\mathsf{D}} \ \Big(\|x - y\| \in [2^{-l}, \mathsf{d}] \to \langle x - y, \mathsf{B}x - \mathsf{B}y\rangle \geq 2^{-\Theta_\mathsf{B}(l)} \Big).$$

Then the following holds

$$\forall l \in \mathbb{N}\, \forall n \geq \tau_\mathsf{B}(l, \Theta_\mathsf{B}, \mathsf{d}, \mathsf{b}, \mathsf{k}, m) \ \Big(\|x_n - x\| < 2^{-l} \Big),$$

where

$$\tau_\mathsf{B}(l, \Theta_\mathsf{B}, \mathsf{d}, \mathsf{b}, \mathsf{k}, m) := \psi\Big(\frac{2^{-\Theta_\mathsf{B}(l)}}{\mathsf{d}}, \mathsf{d}, \mathsf{b}, \mathsf{k}, m\Big)$$

with ψ as in Theorem 3.3.

Proof. We observe that, as in the proof of Theorem 3.2, we can argue that Theorem 1.15 is applicable to $\mathsf{R} := (1-\delta)\mathsf{Id} + \delta\mathsf{T}$ and $(\mu_n)_{n\in\mathbb{N}} := (\frac{1}{\delta}\lambda_n)_{n\in\mathbb{N}}$. This gives us the Fejér-monotonicity of $(x_n)_{n\in\mathbb{N}}$ with respect to $\mathsf{Fix}(\mathsf{R}) = \mathsf{zer}(\mathsf{A}+\mathsf{B})$. By $x \in \mathsf{zer}(\mathsf{A}+\mathsf{B})$, we, in particular, have

$$\forall n \in \mathbb{N} \ \Big(\|x_{n+1} - x\| \leq \|x_n - x\| \Big). \tag{3.4.1}$$

Furthermore, by (3.2.4) and (3.2.5) we get the monotonicity of the sequence $(\|\mathsf{T}x_n - x_n\|)_{n\in\mathbb{N}}$:

$$\forall n \in \mathbb{N} \ \Big(\|\mathsf{T}x_{n+1} - x_{n+1}\| \leq \|\mathsf{T}x_n - x_n\| \Big). \tag{3.4.2}$$

(a): We define $z_n := \mathsf{J}_{\gamma\mathsf{A}}\, y_n$ and note that $z_n = \mathsf{J}_{\gamma\mathsf{A}}(x_n - \gamma\mathsf{B}x_n) = \mathsf{T}x_n$ holds for all $n \in \mathbb{N}$. We want to show that D is bounded first. For arbitrary $n \in \mathbb{N}$ we obtain

$$\begin{aligned} \|z_n - x\| &= \|\mathsf{T}x_n - x_n + x_n - x\| \leq \|\mathsf{T}x_n - x_n\| + \|x_n - x\| \\ &\leq \|\mathsf{T}x_0 - x_0\| + \|x_0 - x\| \leq \mathsf{M} + \mathsf{d}. \end{aligned} \tag{3.4.3}$$

This shows, in particular, the boundedness of D and that we chose $\mathsf{M}+\mathsf{d}$ appropriately in the property $(\diamondsuit_2)$ of the modulus Θ_{A}.
We apply the Cauchy-Schwartz inequality to obtain the following:

$$\begin{aligned}\langle z_n - x, \gamma^{-1}(x_n - z_n) - \mathsf{B}x_n - (-\mathsf{B}x)\rangle &\leq \|z_n - x\|\|\gamma^{-1}(x_n - z_n) - \mathsf{B}x_n + \mathsf{B}x\| \\ &\overset{(3.4.3)}{\leq} (\mathsf{M}+\mathsf{d})\big(\gamma^{-1}\|x_n - \mathsf{T}x_n\| + \|\mathsf{B}x_n - \mathsf{B}x\|\big) \\ &\leq \mathsf{k}(\mathsf{M}+\mathsf{d})\|x_n - \mathsf{T}x_n\| + (\mathsf{M}+\mathsf{d})\|\mathsf{B}x_n - \mathsf{B}x\|.\end{aligned}$$

From now on let $l \in \mathbb{N}$ be arbitrary and assume $n \geq \tau_{\mathsf{A}}(l, \Theta_{\mathsf{A}}, \mathsf{d}, \mathsf{b}, \mathsf{k}, \mathsf{M}, m)$. By Theorem 3.2 and Theorem 3.3 respectively, we know that

$$\|x_n - \mathsf{T}x_n\| < \frac{2^{-\Theta_{\mathsf{A}}(l+1)}}{2\mathsf{k}(\mathsf{M}+\mathsf{d})} \text{ and } \|\mathsf{B}x_n - \mathsf{B}x\| < \frac{2^{-\Theta_{\mathsf{A}}(l+1)}}{2(\mathsf{M}+\mathsf{d})}$$

hold. Thus, we obtain

$$\langle z_n - x, \gamma^{-1}(x_n - z_n) - \mathsf{B}x_n - (-\mathsf{B}x)\rangle < \mathsf{k}(\mathsf{M}+\mathsf{d}) \cdot \frac{2^{-\Theta_{\mathsf{A}}(l+1)}}{2\mathsf{k}(\mathsf{M}+\mathsf{d})} + (\mathsf{M}+\mathsf{d}) \cdot \frac{2^{-\Theta_{\mathsf{A}}(l+1)}}{2(\mathsf{M}+\mathsf{d})} = 2^{-\Theta_{\mathsf{A}}(l+1)}. \tag{3.4.4}$$

By assumption, it holds that $x \in \mathsf{zer}(\mathsf{A}+\mathsf{B})$ which yields $-\mathsf{B}x \in \mathsf{A}(x)$. Moreover, we have $x_n - \gamma\mathsf{B}x_n - z_n = y_n - \mathrm{J}_{\gamma\mathsf{A}}\, y_n$ for all $n \in \mathbb{N}$. Proposition 1.9*(ii)* now implies $y_n - \mathrm{J}_{\gamma\mathsf{A}}\, y_n \in \gamma\mathsf{A}(\mathrm{J}_{\gamma\mathsf{A}}\, y_n)$ and thus $\gamma^{-1}(x_n - z_n) - \mathsf{B}x_n \in \mathsf{A}z_n$ for all $n \in \mathbb{N}$. Therefore, we can use the property $(\diamondsuit_2)$ of our modulus Θ_{A} on (3.4.4) and conclude $\|z_n - x\| < 2^{-(l+1)}$. By Theorem 3.2 and the definition of $\tau_{\mathsf{A}}(l, \Theta_{\mathsf{A}}, \mathsf{d}, \mathsf{b}, \mathsf{k}, \mathsf{M}, m)$ we also have that $\|\mathsf{T}x_n - x_n\| < 2^{-(l+1)}$ holds. Hence, we can conclude for $n \geq \tau_{\mathsf{A}}(l, \Theta_{\mathsf{A}}, \mathsf{d}, \mathsf{b}, \mathsf{k}, \mathsf{M}, m)$

$$\|x_n - x\| \leq \|x_n - z_n\| + \|z_n - x\| = \|\mathsf{T}x_n - x_n\| + \|z_n - x\| < 2^{-(l+1)} + 2^{-(l+1)} = 2^{-l}.$$

(b): By (3.4.1) it holds that

$$\forall n \in \mathbb{N}\ (\|x_n - x\| \leq \|x_0 - x\| \leq \mathsf{d}).$$

Hence $\tilde{D}$ is bounded and d was chosen suitably in the property $(\diamondsuit_2')$ of the modulus Θ_{B}. By Cauchy-Schwartz, we get

$$\langle x_n - x, \mathsf{B}x_n - \mathsf{B}x\rangle \leq \|x_n - x\|\|\mathsf{B}x_n - \mathsf{B}x\| \leq \mathsf{d}\|\mathsf{B}x_n - \mathsf{B}x\|.$$

From now on let $l \in \mathbb{N}$ be arbitrary and $n \geq \tau_{\mathsf{B}}(l, \Theta_{\mathsf{B}}, \mathsf{d}, \mathsf{b}, \mathsf{k}, m)$. By Theorem 3.3, we know that $\|\mathsf{B}x_n - \mathsf{B}x\| < \frac{2^{-\Theta_{\mathsf{B}}(l)}}{\mathsf{d}}$ holds. Thus,

$$\langle x_n - x, \mathsf{B}x_n - \mathsf{B}x\rangle < \mathsf{d} \cdot \frac{2^{-\Theta_{\mathsf{B}}(l)}}{\mathsf{d}} = 2^{-\Theta_{\mathsf{B}}(l)}.$$

Finally, we apply the property $(\diamondsuit_2')$ of the modulus Θ_{B} to conclude $\|x_n - x\| < 2^{-l}$. □

Remark 3.5. *Note that, if the operator* B *is uniformly monotone, the rate of convergence* τ_{B} *does not depend on* M. *This means we do not actually need an estimate for* $\|\mathsf{T}x_0 - x_0\|$ *in that case.*

3.2 Discussion of the Results (for logicians)

We now analyze our results for the Forward-Backward Algorithm like we did for the results for Tseng's Algorithm. We are going to discuss Theorem 3.4 exemplary. The other results can be treated analogously again. The treatment of A, B and $\mathsf{A}+\mathsf{B}$ is similar to the case of Tseng's algorithm. The uniform monotonicity of the operator A or B respectively will again be ensured by a universal premise. This universal premise is $(\diamondsuit_2)$ if A is uniformly monotone or $(\diamondsuit_2')$ if B is uniformly monotone (again, we have to reformulate these premises - in an obviously inessential way - by replacing e.g. $\|x-y\| \in [2^{-l}, \mathsf{M}+\mathsf{d}]$ by $\|x-y\| \in (2^{-l}, \mathsf{M}+\mathsf{d}+1)$ in order to make them purely universal). We express Theorem 3.4(a) similar to the expression of Tseng's Algorithm in the following way:

$$\begin{cases}\forall l, \mathsf{b}, \mathsf{k} \in \mathbb{N}\ \forall m, \Theta \in \mathbb{N}^{\mathbb{N}}\ \forall \beta \in [0, \mathsf{b}]\ \forall \gamma \in [0, 2\mathsf{b}]\ \forall x, x_0 \in \mathcal{H}\ \forall \delta \in [0, 2]\ \forall (\lambda_n)_{n\in\mathbb{N}} \in [0, 2]^{\mathbb{N}} \\ \Bigg(\mathsf{k} \geq 1 \wedge \frac{1}{\mathsf{k}} \leq \gamma \leq 2\beta - \frac{1}{\mathsf{k}} \wedge \delta =_{\mathbb{R}} 2 - \frac{\gamma}{2\beta} \wedge \chi_{\mathsf{A}+\mathsf{B}}(x, 0) = 0 \wedge (\diamondsuit_2) \wedge \forall n \in \mathbb{N}\ \big(0 \leq_{\mathbb{R}} \lambda_n \leq_{\mathbb{R}} \delta\big) \wedge \\ \forall L \in \mathbb{N} \left(\sum_{n=0}^{m(L)} \lambda_n(\delta - \lambda_n) \geq L \right) \wedge \forall x, y \in \mathcal{H} \left(\langle x - y, \mathcal{B}x - \mathcal{B}y \rangle \geq \beta \|\mathcal{B}x - \mathcal{B}y\|^2 \right) \to \exists m \in \mathbb{N}\ \ \|x_m - x\| < 2^{-l} \Bigg),\end{cases}$$

where, again, the definition of (x_n) is explicitly hardwired via the recursor constant in our system.
By the aforementioned logical metatheorems, we can extract an effective uniform bound on '$\exists m$' which only depends on $l, \mathsf{b}, \mathsf{k}, m, \Theta_{\mathsf{A}}$, some $\mathsf{d} \geq \|x - x_0\|, \|x_0\|$ and majorizing data for $\mathrm{J}_{\gamma\mathsf{A}}$ and B.[1] Analogously, we could express Theorem 3.4(b) by exchanging $(\diamondsuit_2)$ with $(\diamondsuit_2')$. As in the case of Tseng's Algorithm, the existence of a bound on some $m \in \mathbb{N}$, such that $\|x_m - x\| < 2^{-l}$ holds, is enough to show the convergence of the sequence $(x_n)_{n\in\mathbb{N}}$ since we have the monotonicity shown in (3.4.1). Also, similarly to the case of Tseng's Algorithm, the majorizing data for $\mathrm{J}_{\gamma\mathsf{A}}$ and B turned out - in the special situation at hand - to be only needed in the weak form of an upper bound $\mathsf{M} \geq \|x_0 - \mathsf{T}x_0\|$ and it has been sufficient to assume that $\mathsf{d} \geq \|x - x_0\|$.

4 Douglas-Rachford Splitting Algorithm

The following algorithm works in a setting that is a bit different to the ones for Tseng's Algorithm and the Forward-Backward Algorithm. We now demand both operators to be maximally monotone and use their resolvents in the iterations. However, we do not impose any further constraints like Lipschitz-continuity or cocoercivity on the operators which both may be set-valued.

Theorem 4.1 (Douglas-Rachford Algorithm, 26.11 in [1])**.** *Assume that* $\mathsf{A} : \mathcal{H} \to 2^{\mathcal{H}}$ *and* $\mathsf{B} : \mathcal{H} \to 2^{\mathcal{H}}$ *are maximally monotone operators such that* $\mathsf{zer}(\mathsf{A} + \mathsf{B}) \neq \emptyset$*. Let* $(\lambda_n)_{n\in\mathbb{N}} \subseteq [0, 2]$ *be a sequence such that* $\sum_{n\in\mathbb{N}} \lambda_n(2 - \lambda_n) = \infty$*. Moreover, let* $\gamma \in (0, \infty)$ *and* $y_0 \in \mathcal{H}$*. For* $n \in \mathbb{N}$ *we set*

$$\begin{aligned} x_n &= \mathrm{J}_{\gamma\mathsf{B}}\, y_n, \\ z_n &= \mathrm{J}_{\gamma\mathsf{A}}(2x_n - y_n), \\ y_{n+1} &= y_n + \lambda_n(z_n - x_n). \end{aligned}$$

Then the following hold:

(i) $(x_n - z_n)_{n\in\mathbb{N}}$ *converges strongly to* 0*.*

(ii) Suppose that one of the following holds:

- *a)* A *is uniformly monotone on every nonempty bounded subset of* $\mathsf{dom}(\mathsf{A})$*.*
- *b)* B *is uniformly monotone on every nonempty bounded subset of* $\mathsf{dom}(\mathsf{B})$*.*

Then $(x_n)_{n\in\mathbb{N}}$ *and* $(z_n)_{n\in\mathbb{N}}$ *converge strongly to the unique point in* $\mathsf{zer}(\mathsf{A} + \mathsf{B})$*.*

There are some weak convergences which are additionally shown in Theorem 26.11 in [1]. They are used in the proof of Theorem 4.1*(ii)* to establish the boundedness of certain sets. As in the sections before, we will quantify these boundedness statements without using these weak convergences.

4.1 Extracting Rates of Convergence

In a first step, we are going to find a rate of convergence for $(x_n - z_n)_{n\in\mathbb{N}}$. Let $x \in \mathsf{zer}(\mathsf{A} + \mathsf{B})$. By Proposition 26.1(iii)(b) in [1] we have that $\mathsf{zer}(\mathsf{A} + \mathsf{B}) = \mathrm{J}_{\gamma\mathsf{B}}(\mathsf{Fix}(\mathrm{R}_{\gamma\mathsf{A}}\, \mathrm{R}_{\gamma\mathsf{B}}))$ and so there is some $y \in \mathsf{Fix}(\mathrm{R}_{\gamma\mathsf{A}}\, \mathrm{R}_{\gamma\mathsf{B}})$ such that $x = \mathrm{J}_{\gamma\mathsf{B}}\, y$ holds.

[1]The bound ξ required in [11] in the treatment of $\mathsf{A} + \mathsf{B}$ is again not needed since B is single-valued.

Theorem 4.2. *Assume that* $\mathsf{A} : \mathcal{H} \to 2^{\mathcal{H}}$ *and* $\mathsf{B} : \mathcal{H} \to 2^{\mathcal{H}}$ *are maximally monotone operators such that* $\mathsf{zer}(\mathsf{A}+\mathsf{B}) \neq \emptyset$*. Let* $(\lambda_n)_{n\in\mathbb{N}} \subseteq [0,2]$ *be a sequence such that* $\sum_{n\in\mathbb{N}} \lambda_n(2-\lambda_n) = \infty$*. Moreover, let* $\gamma \in (0,\infty)$ *and take* $y_0 \in \mathcal{H}$ *and* $x \in \mathsf{zer}(\mathsf{A}+\mathsf{B})$*. Let the sequences* $(x_n)_{n\in\mathbb{N}}$*,* $(y_n)_{n\in\mathbb{N}}$ *and* $(z_n)_{n\in\mathbb{N}}$ *be defined as in Theorem 4.1. Assume that* $y \in \mathsf{Fix}(\mathrm{R}_{\gamma\mathsf{A}}\,\mathrm{R}_{\gamma\mathsf{B}})$ *is such that* $x = \mathrm{J}_{\gamma\mathsf{B}}\, y$ *and* $\mathsf{d} \in \mathbb{N}$ *is such that*

$$\|y_0 - y\| \leq \mathsf{d}.$$

Furthermore, let $m : \mathbb{N} \to \mathbb{N}$ *be such that*

$$\forall L \in \mathbb{N} \ \left(\sum_{n=0}^{m(L)} \lambda_n(2-\lambda_n) \geq L \right).$$

Then the following holds

$$\forall \varepsilon > 0 \, \forall n \geq \varphi(\varepsilon, \mathsf{d}, m) \ \Big(\|x_n - z_n\| < \varepsilon \Big), \ \textit{where } \varphi(\varepsilon, \mathsf{d}, m) := \widetilde{m}\left(\frac{\mathsf{d}^2+1}{4\varepsilon^2} \right), \ \widetilde{m}(s) := m\Big(\lceil 4s \rceil\Big).$$

Proof. We define $\mathsf{T} := \mathrm{J}_{\gamma\mathsf{A}}\,\mathrm{R}_{\gamma\mathsf{B}} + \mathsf{Id} - \mathrm{J}_{\gamma\mathsf{B}}$. Note that by definition we then have for all $n \in \mathbb{N}$:

$$z_n - x_n = \mathrm{J}_{\gamma\mathsf{A}}(2x_n - y_n) - x_n = \mathrm{J}_{\gamma\mathsf{A}}(2\,\mathrm{J}_{\gamma\mathsf{B}}\, y_n - y_n) - \mathrm{J}_{\gamma\mathsf{B}}\, y_n = \mathsf{T}y_n - y_n. \tag{4.2.1}$$

From [1][Prop.4.31] we now that $\mathsf{R} := \mathrm{R}_{\gamma\mathsf{A}}\,\mathrm{R}_{\gamma\mathsf{B}} = 2\mathsf{T} - \mathsf{Id}$ and R is nonexpansive. As in [1], one shows that Theorem 1.15 is applicable to $\mathsf{R}, (\mu_n) := (\frac{1}{2}\lambda_n)$ and (y_n) since $y_0 \in \mathcal{H}$ and

$$y_n + \mu_n(\mathsf{R}y_n - y_n) = y_n + \frac{1}{2}\lambda_n((2\mathsf{T}y_n - y_n - y_n) = y_n + \lambda_n(\mathsf{T}y_n - y_n) \overset{(4.2.1)}{=} y_n + \lambda_n(z_n - x_n) = y_{n+1}$$

holds for all $n \in \mathbb{N}$. Because of $y \in \mathsf{Fix}(\mathrm{R}_{\gamma\mathsf{A}}\,\mathrm{R}_{\gamma\mathsf{B}}) = \mathsf{Fix}(\mathsf{R})$, we can use (1.15.2) to obtain:

$$\forall m \in \mathbb{N} \ \left(\sum_{n=0}^{m} \mu_n(1-\mu_n)\|\mathsf{R}y_n - y_n\|^2 \leq \|y_0 - y\|^2 \right). \tag{4.2.2}$$

Furthermore, by (1.15.3) we get the monotonicity of $(\|\mathsf{R}y_n - y_n\|)_{n\in\mathbb{N}}$:

$$\forall n \in \mathbb{N} \ \Big(\|\mathsf{R}y_{n+1} - y_{n+1}\| \leq \|\mathsf{R}y_n - y_n\| \Big). \tag{4.2.3}$$

Because of $\mathsf{R} = 2\mathsf{T} - \mathsf{Id}$, we have

$$\forall n \in \mathbb{N} \ \Big(\mathsf{R}y_n - y_n = 2\mathsf{T}y_n - y_n - y_n = 2(\mathsf{T}y_n - y_n) \Big). \tag{4.2.4}$$

Let $\varepsilon > 0$ be arbitrary. From now on we write φ for $\varphi(\varepsilon, \mathsf{d}, m)$ since the parameters are fixed throughout the proof. Assume that for all $n \leq \varphi$ we have $\|\mathsf{T}y_n - y_n\| \geq \varepsilon$. For arbitrary $s \geq 0$, we get

$$\sum_{n=0}^{\widetilde{m}(s)} \mu_n(1-\mu_n) = \sum_{n=0}^{\widetilde{m}(s)} \frac{1}{2}\lambda_n(1 - \frac{1}{2}\lambda_n) = \frac{1}{4}\sum_{n=0}^{m(\lceil 4s \rceil)} \lambda_n(2-\lambda_n) \geq \frac{\lceil 4s \rceil}{4} \geq s.$$

Moreover, for all $n \leq \varphi$ we have $\|\mathsf{R}y_n - y_n\| = 2\|\mathsf{T}y_n - y_n\| \geq 2\varepsilon$. Combining these things, we get

$$\sum_{n=0}^{\varphi} \mu_n(1-\mu_n)\|\mathsf{R}y_n - y_n\|^2 \geq 4\varepsilon^2 \sum_{n=0}^{\varphi} \mu_n(1-\mu_n) \overset{\text{Def. } \varphi}{\geq} 4\varepsilon^2 \cdot \frac{\mathsf{d}^2+1}{4\varepsilon^2} = \mathsf{d}^2 + 1 > \mathsf{d}^2 \geq \|y_0 - y\|^2$$

which contradicts (4.2.2). Hence, there has to be some $\bar{n} \leq \varphi$ such that $\|\mathsf{T}y_{\bar{n}} - y_{\bar{n}}\| < \varepsilon$ holds. We obtain for $n > \bar{n}$ with the monotonicity of $(\|\mathsf{R}y_n - y_n\|)_{n\in\mathbb{N}}$

$$\|\mathsf{T}y_n - y_n\| \overset{(4.2.4)}{=} \frac{1}{2}\|\mathsf{R}y_n - y_n\| \overset{(4.2.3)}{\leq} \frac{1}{2}\|\mathsf{R}y_{\bar{n}} - y_{\bar{n}}\| = \|\mathsf{T}y_{\bar{n}} - y_{\bar{n}}\| < \varepsilon.$$

Thus, $\|z_n - x_n\| \overset{(4.2.1)}{=} \|\mathsf{T}y_n - y_n\| < \varepsilon$ holds for all $n \geq \varphi$. $\square$

Now let us consider again the additional assumption that either A or B is uniformly monotone on a suitable bounded subset of its domain:

Theorem 4.3. *Let* $\mathsf{A}, \mathsf{B}, (\lambda_n), \gamma, y_0, y, x, (x_n), (y_n), (z_n), \mathsf{k}, \mathsf{d}, m$ *be as in Theorem 4.2 and set* $\mathsf{T} := \mathrm{J}_{\gamma\mathsf{A}} \mathrm{R}_{\gamma\mathsf{B}} + \mathsf{Id} - \mathrm{J}_{\gamma\mathsf{B}}$*. Let* $\mathsf{M} \in \mathbb{N}$ *be such that* $\|\mathsf{T}y_0 - y_0\| \leq \mathsf{M}$.

(a) *Assume that we have some modulus* Θ_{A} *of uniform monotonicity for the operator* A *on the bounded subset* $\mathsf{D} := \{z_n\}_{n\in\mathbb{N}} \cup \{x\}$*, i.e., such that*

$$(\diamondsuit_3) \qquad \forall l \in \mathbb{N}\, \forall y \in \mathsf{D}\, \forall (x,u),(y,v) \in \mathsf{gra}(\mathsf{A}) \ \Big(\|x-y\| \in [2^{-l}, \mathsf{M}+\mathsf{d}] \to \langle x-y, u-v \rangle \geq 2^{-\Theta_{\mathsf{A}}(l)} \Big).$$

Then the following holds

$$\forall l \in \mathbb{N}\, \forall n \geq \psi(l, \Theta_{\mathsf{A}}, \mathsf{k}, \mathsf{d}, \mathsf{M}, m) \ \Big(\|z_n - x\| < 2^{-l} \Big),$$

where

$$\psi(l, \Theta_{\mathsf{A}}, \mathsf{k}, \mathsf{d}, \mathsf{M}, m) := \varphi\Big(\frac{2^{-\Theta_{\mathsf{A}}(l)}}{\mathsf{k}(\mathsf{M}+\mathsf{d})}, \mathsf{d}, m \Big),$$
$$\varphi(\varepsilon, \mathsf{d}, m) := \widetilde{m}\left(\frac{\mathsf{d}^2+1}{4\varepsilon^2} \right), \ \ \widetilde{m}(s) := m\Big(\lceil 4s \rceil \Big).$$

(b) *Assume that we have some modulus* Θ_{B} *of uniform monotonicity for the operator* B *on the bounded subset* $\widetilde{\mathsf{D}} := \{x_n\}_{n\in\mathbb{N}} \cup \{x\}$*, i.e., such that*

$$(\diamondsuit_3') \qquad \forall l \in \mathbb{N}\, \forall y \in \widetilde{\mathsf{D}}\, \forall (x,u),(y,v) \in \mathsf{gra}(\mathsf{B}) \ \Big(\|x-y\| \in [2^{-l}, \mathsf{d}] \to \langle x-y, u-v \rangle \geq 2^{-\Theta_{\mathsf{B}}(l)} \Big).$$

Then the following holds

$$\forall l \in \mathbb{N}\, \forall n \geq \psi(l, \Theta_{\mathsf{B}}, \mathsf{k}, \mathsf{d}, \mathsf{M}, m) \ \Big(\|x_n - x\| < 2^{-l} \Big),$$

where ψ *is as in (a).*

Proof. As in the proof of Theorem 4.2, we can argue that Theorem 1.15 is applicable to $\mathsf{R} := \mathrm{R}_{\gamma\mathsf{A}} \mathrm{R}_{\gamma\mathsf{B}} = 2\mathsf{T} - \mathsf{Id}$ and $(\mu_n)_{n\in\mathbb{N}} := (\frac{1}{2}\lambda_n)_{n\in\mathbb{N}}$. This gives us the Fejér-monotonicity of $(y_n)_{n\in\mathbb{N}}$ with respect to $\mathsf{Fix}(\mathsf{R}) = \mathsf{Fix}(\mathrm{R}_{\gamma\mathsf{A}} \mathrm{R}_{\gamma\mathsf{B}})$. Hence, for arbitrary $y \in \mathsf{Fix}(\mathrm{R}_{\gamma\mathsf{A}} \mathrm{R}_{\gamma\mathsf{B}})$ we have

$$\forall n \in \mathbb{N} \ \Big(\|y_{n+1} - y\| \leq \|y_n - y\| \Big). \tag{4.3.1}$$

Moreover, by (4.2.3) and (4.2.4) we get the monotonicity of $(\|\mathsf{T}y_n - y_n\|)_{n\in\mathbb{N}}$:

$$\forall n \in \mathbb{N} \ \Big(\|\mathsf{T}y_{n+1} - y_{n+1}\| \leq \|\mathsf{T}y_n - y_n\| \Big). \tag{4.3.2}$$

By the proof of (26.10) in [1], we conclude that $(x, \gamma^{-1}(x-y)) \in \mathsf{gra}(\mathsf{A})$ and $(x, \gamma^{-1}(y-x)) \in \mathsf{gra}(\mathsf{B})$. We define $u := \gamma^{-1}(y - \mathrm{J}_{\gamma\mathsf{B}}\, y)$ and, hence, get $(x,u) \in \mathsf{gra}(\mathsf{B})$ and $(x,-u) \in \mathsf{gra}(\mathsf{A})$. Furthermore, we define $w_n := \gamma^{-1}(2x_n - y_n - z_n)$ and $u_n := \gamma^{-1}(y_n - x_n)$ for all $n \in \mathbb{N}$. By definition, we know that $z_n = \mathrm{J}_{\gamma\mathsf{A}}(2x_n - y_n)$ and $x_n = \mathrm{J}_{\gamma\mathsf{B}}\, y_n$ hold for all $n \in \mathbb{N}$. We can apply Proposition 1.9*(ii)* to obtain $(z_n, w_n) = (z_n, \gamma^{-1}(2x_n - y_n - z_n)) \in \mathsf{gra}(\mathsf{A})$ and $(x_n, u_n) = (x_n, \gamma^{-1}(y_n - x_n)) \in \mathsf{gra}(\mathsf{B})$ for all $n \in \mathbb{N}$.
(a): We first want to show that D is actually a bounded subset. For this we consider the following estimate for all $n \in \mathbb{N}$:

$$\begin{aligned} \|z_n - x\| &\leq \|z_n - x_n\| + \|x_n - x\| \overset{(4.2.1)}{=} \|\mathsf{T}y_n - y_n\| + \|\mathrm{J}_{\gamma\mathsf{B}}\, y_n - \mathrm{J}_{\gamma\mathsf{B}}\, y\| \\ &\leq \|\mathsf{T}y_0 - y_0\| + \|y_n - y\| \leq \mathsf{M} + \|y_0 - y\| \leq \mathsf{M} + \mathsf{d}. \end{aligned}$$

Therefore, D is bounded. It also shows that $\mathsf{M}+\mathsf{d}$ is suitably chosen in the property $(\diamondsuit_3)$ of the modulus Θ_{A}. Since $y \in \mathsf{Fix}(\mathrm{R}_{\gamma\mathsf{A}}\,\mathrm{R}_{\gamma\mathsf{B}})$ holds, we can estimate:

$$\begin{aligned} \|w_n + \gamma^{-1}(y - x_n)\| &= \gamma^{-1}\|2x_n - y_n - z_n + y - x_n\| = \gamma^{-1}\|x_n - z_n + y - y_n\| \\ &\overset{(4.2.1)}{\leq} \gamma^{-1}\Big(\|\mathsf{T}y_n - y_n\| + \|y_n - y\|\Big) \overset{(4.3.1),(4.3.2)}{\leq} \gamma^{-1}\Big(\|\mathsf{T}y_0 - y_0\| + \|y_0 - y\|\Big) \\ &\overset{(4.3.1)}{\leq} \mathsf{k}(\mathsf{M}+\mathsf{d}). \end{aligned} \tag{4.3.4}$$

Moreover, we have the following equality for all $n \in \mathbb{N}$ which is shown in (26.43) in [1].

$$\langle z_n - x, w_n + u\rangle + \langle x_n - x, u_n - u\rangle = \langle z_n - x_n, w_n + \gamma^{-1}(y - x_n)\rangle. \tag{4.3.5}$$

Using the monotonicity of B and the Cauchy-Schwartz inequality, we obtain the following

$$\begin{aligned} \langle z_n - x, w_n - (-u)\rangle &\leq \langle z_n - x, w_n + u\rangle + \langle x_n - x, u_n - u\rangle \overset{(4.3.5)}{=} \langle z_n - x_n, w_n + \gamma^{-1}(y - x_n)\rangle \\ &\leq \|z_n - x_n\|\|w_n + \gamma^{-1}(y - x_n)\| \overset{(4.3.4)}{\leq} \mathsf{k}(\mathsf{M}+\mathsf{d})\|z_n - x_n\|. \end{aligned}$$

From now on let $l \in \mathbb{N}$ be arbitrary and $n \geq \psi(l, \Theta_{\mathsf{A}}, \mathsf{k}, \mathsf{d}, \mathsf{M}, m)$. By Theorem 4.2 and the definition of ψ, we know that $\|z_n - x_n\| < \frac{2^{-\Theta_{\mathsf{A}}(l)}}{\mathsf{k}(\mathsf{M}+\mathsf{d})}$ holds and therefore also

$$\langle z_n - x, w_n + u\rangle < \mathsf{k}(\mathsf{M}+\mathsf{d}) \cdot \frac{2^{-\Theta_{\mathsf{A}}(l)}}{\mathsf{k}(\mathsf{M}+\mathsf{d})} = 2^{-\Theta_{\mathsf{A}}(l)}.$$

Since $\|z_n - x\| \leq \mathsf{M}+\mathsf{d}$ holds, we can use $(\diamondsuit_3)$ to obtain $\|z_n - x\| < 2^{-l}$.
(b): The proof works rather similar to (a). The boundedness of $\widetilde{D}$ follows from the following estimate:

$$\|x_n - x\| = \|\,\mathrm{J}_{\gamma\mathsf{B}}\, y_n - \mathrm{J}_{\gamma\mathsf{B}}\, y\| \leq \|y_n - y\| \overset{(4.3.1)}{\leq} \|y_0 - y\| \leq \mathsf{d}.$$

Again this also shows that d is chosen suitably in the property $(\diamondsuit_3')$ for the modulus Θ_{B}. Moreover, we can estimate

$$\|u_n + \gamma^{-1}(z_n - y)\| = \gamma^{-1}\|y_n - x_n + z_n - y\| = \gamma^{-1}\|z_n - x_n + y_n - y\| \overset{(4.3.4)}{\leq} \mathsf{k}(\mathsf{M}+\mathsf{d}). \tag{4.3.6}$$

Similar to (4.3.5), we have the following equality for all $n \in \mathbb{N}$ (cf. (26.44) in [1]):

$$\langle x_n - x, u_n - u\rangle + \langle z_n - x, w_n + u\rangle = \langle x_n - z_n, u_n + \gamma^{-1}(z_n - y)\rangle. \tag{4.3.7}$$

Now, using the monotonicity of A and the Cauchy-Schwartz inequality, we obtain

$$\begin{aligned} \langle x_n - x, u_n - u\rangle &\leq \langle x_n - x, u_n - u\rangle + \langle z_n - x, w_n + u\rangle \overset{(4.3.7)}{=} \langle x_n - z_n, u_n + \gamma^{-1}(z_n - y)\rangle \\ &\leq \|z_n - x_n\|\|u_n + \gamma^{-1}(z_n - y)\| \overset{(4.3.6)}{\leq} \mathsf{k}(\mathsf{M}+\mathsf{d})\|z_n - x_n\|. \end{aligned}$$

Let $l \in \mathbb{N}$ be arbitrary and $n \geq \psi(l, \Theta_{\mathsf{B}}, \mathsf{k}, \mathsf{d}, \mathsf{M}, m)$. We know that $\|z_n - x_n\| < \frac{2^{-\Theta_{\mathsf{B}}(l)}}{\mathsf{k}(\mathsf{M}+\mathsf{d})}$ holds by Theorem 4.2 and the definition of ψ. Hence,

$$\langle x_n - x, u_n - u\rangle < \mathsf{k}(\mathsf{M}+\mathsf{d})\frac{2^{-\Theta_{\mathsf{B}}(l)}}{\mathsf{k}(\mathsf{M}+\mathsf{d})} = 2^{-\Theta_{\mathsf{B}}(l)}.$$

Since $\|x_n - x\| \leq \mathsf{d}$ holds, we can apply the property $(\diamondsuit_3')$ to obtain $\|x_n - x\| < 2^{-l}$. □

Remark 4.4. *In the case where* A *is uniformly monotone on every bounded subset of* $\mathsf{dom}(\mathsf{A})$*, we have now shown a rate of convergence for* $(z_n)_{n\in\mathbb{N}}$*. In Theorem 4.1(ii), it is mentioned that* $(x_n)_{n\in\mathbb{N}}$ *converges to the same point* $x \in \mathsf{zer}(\mathsf{A}+\mathsf{B})$*. We can obtain a rate of convergence for that by using the triangle inequality to obtain* $\|x_n - x\| \leq \|x_n - z_n\| + \|z_n - x\|$ *and then combining Theorem 4.2 and Theorem 4.3(a).*
Similarly, we get a rate of convergence for $(z_n)_{n\in\mathbb{N}}$ *in the case that* B *is uniformly monotone on every bounded subset of* $\mathsf{dom}(\mathsf{B})$ *by using the triangle inequality and then combining Theorem 4.2 and Theorem 4.3(b).*

4.2 Discussion of the Results (for logicians)

The treatment of the operators and their resolvents is similar to the previous algorithms. However, this time we cannot argue via the monotonicity of ($\|x_m - x\|$). In fact, we need to apply metatheorems twice to get, combined, an explanation for Theorem 4.3. First, we apply (a suitable version) of the logical metatheorems to

$$\forall l \in \mathbb{N}\, \forall m, \Theta_{\mathsf{A}} \in \mathbb{N}^{\mathbb{N}}\, \forall x, y_0 \in \mathcal{H}\, \forall (\lambda_n)_{n\in\mathbb{N}} \in [0,2]^{\mathbb{N}}\, \forall \gamma \in [0,\infty)\, \forall k \in \mathbb{N}\setminus\{0\}$$
$$\Bigg(\gamma \geq_{\mathbb{R}} \tfrac{1}{k} \wedge \forall n \in \mathbb{N}\left(0 \leq_{\mathbb{R}} \lambda_n \leq_{\mathbb{R}} 2\right) \wedge \forall L \in \mathbb{N} \left(\sum_{n=0}^{m(L)} \lambda_n(2-\lambda_n) \geq L \right) \wedge \chi_{\mathsf{A}+\mathsf{B}}(x,0) = 0$$
$$\to \exists m \in \mathbb{N}\quad \|z_m - x_m\| < 2^{-l} \Bigg),$$

to extract an effective bound $\varphi(l)$ on '$\exists m$' which, by the monotonicity of ($\|z_m - x_m\|$), is a rate of convergence for $\|z_m - x_m\| \to 0$. Here, again, the definitions of $(x_n), (y_n), (z_n)$ are explicitly hardwired by a recursor constant into the formal system. We then extract - under the additional assumption of ($\diamondsuit_3$) (which again is stated with the open interval $(2^{-l}, \mathsf{M}+\mathsf{d}+1)$ to make it purely universal) - from the proof that

$$\|z_m - x_m\| = 0 \to \|z_m - x\| = 0,$$

which prenexes into

$$\forall l \in \mathbb{N}\, \exists j \in \mathbb{N}\, (\|z_m - x_m\| \leq 2^{-j} \to \|z_m - x\| < 2^{-l}),$$

a bound $\alpha(l)$ (and hence a witness) for $\exists j$ depending on l (but not on m as the proof only uses that $(z_m, w_m) \in \mathsf{gra}(\mathsf{A})$ and $(x_m, u_m) \in \mathsf{gra}(\mathsf{B})$ and certain boundedness facts). If now $n \geq \alpha(\varphi(l))$, then $\|z_n - x\| < 2^{-l}$. This rate of convergence is a-priorily guaranteed (by the aforementioned metatheorems) to depend only on l, m, Θ_A, k, some upper bound $\mathsf{d} \geq \|x\|, \|x - y_0\|$, an upper bound for γ and the bound ξ from the treatment of $\mathsf{A}+\mathsf{B}$ in [11] as well as majorizing data for A, B. In the special case at hand, the use of ξ, the majorizing data for A, B and the upper bound on γ and d is only made implicitly in stipulating that we have bounds $\mathsf{d} \geq \|y_0 - y\|$ (for some y as in the quantitative version of the theorem which does not occur in the theorem itself) and $\mathsf{M} \geq \|\mathsf{T}y_0 - y_0\|$ which can be computed in these data using - in the case of $\mathsf{d} \geq \|y_0 - y\|$ - the bound ξ on $(x, x-y) \in \mathsf{gra}(\gamma\mathsf{A})$ and $(x, y-x) \in \mathsf{gra}(\mathsf{B})$ and $(x, 0) \in \mathsf{gra}(\gamma\mathsf{A}+\mathsf{B})$ (together with [1][(26.10)]) and - for M - the estimate

$$\begin{aligned} \|\mathsf{T}y_0 - y_0\| &= \|\,\mathrm{J}_{\gamma\mathsf{A}}(\mathrm{R}_{\gamma\mathsf{B}}\, y_0) + y_0 - \mathrm{J}_{\gamma\mathsf{B}}\, y_0 - y_0\| \leq \|\,\mathrm{J}_{\gamma\mathsf{A}}(\mathrm{R}_{\gamma\mathsf{B}}\, y_0) - y_0\| + \|\,\mathrm{J}_{\gamma\mathsf{B}}\, y_0 - y_0\| \\ &\leq \|\,\mathrm{J}_{\gamma\mathsf{A}}(\mathrm{R}_{\gamma\mathsf{B}}\, y_0) - \mathrm{R}_{\gamma\mathsf{B}}\, y_0\| + \|\,\mathrm{R}_{\gamma\mathsf{B}}\, y_0 - y_0\| + \|\,\mathrm{J}_{\gamma\mathsf{B}}\, y_0 - y_0\| \\ &= \|\,\mathrm{J}_{\gamma\mathsf{A}}(\mathrm{R}_{\gamma\mathsf{B}}\, y_0) - \mathrm{R}_{\gamma\mathsf{B}}\, y_0\| + 3\|\,\mathrm{J}_{\gamma\mathsf{B}}\, y_0 - y_0\|. \end{aligned}$$

Theorem 4.3(b) is treated similarly.

5 Peaceman-Rachford Splitting Algorithm

The last algorithm we are going to treat is the so-called Peaceman-Rachford Algorithm. It is regarded as a kind of limiting case of the Douglas-Rachford Algorithm where $\lambda_n = 2$ holds for all $n \in \mathbb{N}$.

Theorem 5.1 (Peaceman-Rachford Algorithm, 26.13 in [1])**.** *Assume that* $\mathsf{A} : \mathcal{H} \to 2^{\mathcal{H}}$ *and* $\mathsf{B} : \mathcal{H} \to 2^{\mathcal{H}}$ *are maximally monotone operators such that* B *is uniformly monotone and* $\mathsf{zer}(\mathsf{A}+\mathsf{B}) \neq \emptyset$. *Let* $\gamma \in (0,\infty)$, $y_0 \in \mathcal{H}$, *and* $x \in \mathsf{zer}(\mathsf{A}+\mathsf{B})$ *be the unique zero of* $\mathsf{zer}(\mathsf{A}+\mathsf{B})$. *For* $n \in \mathbb{N}$ *we set*

$$\begin{aligned} x_n &= \mathrm{J}_{\gamma\mathsf{B}}\, y_n, \\ z_n &= \mathrm{J}_{\gamma\mathsf{A}}(2x_n - y_n), \\ y_{n+1} &= y_n + 2(z_n - x_n). \end{aligned}$$

Then $(x_n)_{n\in\mathbb{N}}$ *converges strongly to* x.

Note that $y_{n+1} = \mathrm{R}_{\gamma\mathsf{A}}\,\mathrm{R}_{\gamma\mathsf{B}}\, y_n$.

5.1 Extracting a Rate of Metastability

The convergence of $(x_n)_{n\in\mathbb{N}}$ is proven in as (26.52) in [1] with the following inequality:

$$\forall n \in \mathbb{N} \ \left(\|y_{n+1}-y\|^2 \leq \|y_n - y\|^2 - 4\gamma\phi(\|x_n - x\|)\right)$$

for $y \in \mathsf{Fix}\,\mathrm{R}_{\gamma\mathsf{A}}\,\mathrm{R}_{\gamma\mathsf{B}}$. We use (a suitable reformulation of) this inequality to prove a statement which shows that $\|x_n - x\| \geq \varepsilon$ can only be true finitely many times for arbitrary $\varepsilon > 0$. This does, however, not give any information on when this will happen for the last time. The inequality above does not show whether or not the sequence $(\|x_n - x\|)_{n\in\mathbb{N}}$ itself is monotone. Hence, instead of a rate of convergence we can only extract a rate of metastability for the Peaceman-Rachford Algorithm from the proof given in [1]. We will do this in the next theorem.

Again, we need some $y \in \mathsf{Fix}(\mathrm{R}_{\gamma\mathsf{A}}\,\mathrm{R}_{\gamma\mathsf{B}})$ in the following theorem. As for Theorem 4.2, we can argue that this exists since $\mathsf{zer}(\mathsf{A}+\mathsf{B})$ is nonempty by assumption and $\mathsf{zer}(\mathsf{A}+\mathsf{B}) = \mathrm{J}_{\gamma\mathsf{B}}(\mathsf{Fix}(\mathrm{R}_{\gamma\mathsf{A}}\,\mathrm{R}_{\gamma\mathsf{B}}))$ holds by Proposition 26.1(iii)(b) in [1]. For an $x \in \mathsf{zer}(\mathsf{A}+\mathsf{B})$ we, therefore, have some $y \in \mathsf{Fix}(\mathrm{R}_{\gamma\mathsf{A}}\,\mathrm{R}_{\gamma\mathsf{B}})$ such that $x = \mathrm{J}_{\gamma\mathsf{B}}\,y$ holds.

Theorem 5.2. *Assume that $\mathsf{A} : \mathcal{H} \to 2^{\mathcal{H}}$ and $\mathsf{B} : \mathcal{H} \to 2^{\mathcal{H}}$ are maximally monotone operators. Let $\gamma \in (0,\infty)$, $y_0 \in \mathcal{H}$ and $x \in \mathsf{zer}(\mathsf{A}+\mathsf{B})$. Let the sequence $(x_n)_{n\in\mathbb{N}}$ be defined as in Theorem 5.1. Assume $y \in \mathsf{Fix}(\mathrm{R}_{\gamma\mathsf{A}}\,\mathrm{R}_{\gamma\mathsf{B}})$ is such that $x = \mathrm{J}_{\gamma\mathsf{B}}\,y$ and $\mathsf{k},\mathsf{d} \in \mathbb{N}, \mathsf{k} \geq 1$ are such that*

$$\gamma \geq \frac{1}{\mathsf{k}} \ \wedge \ \|y_0 - y\| \leq \mathsf{d}.$$

Furthermore, let Θ be a modulus of uniform monotonicity for the operator B, i.e., such that

$$(\Diamond_4) \qquad \forall l \in \mathbb{N}\,\forall (s,u),(t,v) \in \mathsf{gra}(\mathsf{B}) \ \left(\|s-t\| \in [2^{-l},\mathsf{d}] \to \langle s-t, u-v\rangle \geq 2^{-\Theta(l)}\right).$$

Then we get a rate of metastability for the sequence $(x_n)_{n\in\mathbb{N}}$, i.e.,

$$\forall l \in \mathbb{N}\,\forall g : \mathbb{N}\to\mathbb{N}\,\exists n \leq \phi(l,g,\mathsf{d},\mathsf{k},\Theta)\,\forall i \in [n, n+g(n)] \ \left(\|x_n - x\| < 2^{-l}\right),$$

where $\quad \phi(l,g,\mathsf{d},\mathsf{k},\Theta) := \widetilde{g}^{(\lceil p(l,\mathsf{d},\mathsf{k},\Theta)\rceil)}(0)$ *with* $\widetilde{g}(n) := n + g(n) + 1$, $\ p(l,\mathsf{d},\mathsf{k},\Theta) := \dfrac{\mathsf{d}^2\mathsf{k}}{4}\cdot 2^{\Theta(l)}$.

Proof. Let $l \in \mathbb{N}$ and $g : \mathbb{N}\to\mathbb{N}$ be arbitrary.

For all $n \in \mathbb{N}$ we can use the (proof of the) estimate (26.52) in [1] and obtain the following

$$\begin{aligned}
\|y_{n+1}-y\|^2 &\leq \|y_n - y\|^2 - 4\langle \mathrm{J}_{\gamma\mathsf{B}}\,y_n - \mathrm{J}_{\gamma\mathsf{B}}\,y, y_n - y\rangle + 4\|\mathrm{J}_{\gamma\mathsf{B}}\,y_n - \mathrm{J}_{\gamma\mathsf{B}}\,y\|^2 \\
&= \|y_n - y\|^2 - 4\langle \mathrm{J}_{\gamma\mathsf{B}}\,y_n - \mathrm{J}_{\gamma\mathsf{B}}\,y, y_n - y\rangle + 4\langle \mathrm{J}_{\gamma\mathsf{B}}\,y_n - \mathrm{J}_{\gamma\mathsf{B}}\,y, \mathrm{J}_{\gamma\mathsf{B}}\,y_n - \mathrm{J}_{\gamma\mathsf{B}}\,y\rangle \\
&= \|y_n - y\|^2 - 4\langle \mathrm{J}_{\gamma\mathsf{B}}\,y_n - \mathrm{J}_{\gamma\mathsf{B}}\,y, (y_n - \mathrm{J}_{\gamma\mathsf{B}}\,y_n) - (y - \mathrm{J}_{\gamma\mathsf{B}}\,y)\rangle \\
&= \|y_n - y\|^2 - 4\gamma\langle \mathrm{J}_{\gamma\mathsf{B}}\,y_n - \mathrm{J}_{\gamma\mathsf{B}}\,y, \gamma^{-1}(y_n - \mathrm{J}_{\gamma\mathsf{B}}\,y_n) - \gamma^{-1}(y - \mathrm{J}_{\gamma\mathsf{B}}\,y)\rangle.
\end{aligned} \tag{5.2.1}$$

Furthermore, by the same estimate and the monotonicity of B (together with Proposition 1.9*(ii)*) we obtain the monotonicity of $(\|y_n - y\|)_{n\in\mathbb{N}}$, i.e.,

$$\forall n \in \mathbb{N} \ \left(\|y_{n+1}-y\| \leq \|y_n - y\|\right). \tag{5.2.2}$$

Since we have $\|y_0 - y\| \leq \mathsf{d}$ by assumption, $(\|y_n - y\|^2)_{n\in\mathbb{N}}$ is a nondecreasing sequence in $[0,\mathsf{d}^2]$. We can apply Proposition 2.27 and Remark 2.29.1) of [6] and obtain the following statement:

$$(\dagger) \qquad \forall \delta > 0\,\forall g : \mathbb{N}\to\mathbb{N}\,\exists n \leq \psi(g,\delta,\mathsf{d})\,\forall i \in [n, n+g(n)] \ \left(\left|\|y_i - y\|^2 - \|y_{i+1}-y\|^2\right| < \delta\right)$$

with $\quad \psi(g,\delta,\mathsf{d}) := \widetilde{g}^{(\lceil \mathsf{d}^2\cdot\delta^{-1}\rceil)}(0).$

In particular, for $\bar{\delta} := 2^{-\Theta(l)}\cdot\frac{4}{\mathsf{k}}$ there is an $m \leq \widetilde{g}^{(\lceil \mathsf{d}^2\cdot 2^{\Theta(l)}\cdot\frac{\mathsf{k}}{4}\rceil)}(0) = \phi(l,g,\mathsf{d},\mathsf{k},\Theta)$ such that

$$\forall i \in [m, m+g(m)] \ \left(\left|\|y_i - y\|^2 - \|y_{i+1}-y\|^2\right| < 2^{-\Theta(l)}\cdot\frac{4}{\mathsf{k}}\right) \tag{5.2.3}$$

holds. For all $i \in [m, m+g(m)]$ we can estimate

$$\langle \mathrm{J}_{\gamma \mathsf{B}}\, y_i - \mathrm{J}_{\gamma \mathsf{B}}\, y, \gamma^{-1}(y_i - \mathrm{J}_{\gamma \mathsf{B}}\, y_i) - \gamma^{-1}(y - \mathrm{J}_{\gamma \mathsf{B}}\, y)\rangle \overset{(5.2.1)}{\leq} \frac{1}{4\gamma}\Big(\|y_i - y\|^2 - \|y_{i+1} - y\|^2\Big)$$
$$\leq \frac{\mathsf{k}}{4}\big|\|y_i - y\|^2 - \|y_{i+1} - y\|^2\big| \tag{5.2.4}$$
$$\overset{(5.2.3)}{<} \frac{\mathsf{k}}{4} \cdot 2^{-\Theta(l)} \cdot \frac{4}{\mathsf{k}} = 2^{-\Theta(l)}.$$

Recall that we have $(\mathrm{J}_{\gamma \mathsf{B}}\, y_i, \gamma^{-1}(y_i - \mathrm{J}_{\gamma \mathsf{B}}\, y_i)) \in \mathsf{gra}(\mathsf{B})$ for all $i \in [m, m+g(m)]$ and $(\mathrm{J}_{\gamma \mathsf{B}}\, y, \gamma^{-1}(y - \mathrm{J}_{\gamma \mathsf{B}}\, y)) \in \mathsf{gra}(\mathsf{B})$ by Proposition 1.9 *(ii)*. Furthermore, we can estimate

$$\| \mathrm{J}_{\gamma \mathsf{B}}\, y_n - \mathrm{J}_{\gamma \mathsf{B}}\, y\| \leq \|y_n - y\| \overset{(5.2.2)}{\leq} \|y_0 - y\| \leq \mathsf{d}$$

for all $n \in \mathbb{N}$. Therefore, we can apply the property $(\diamondsuit_4)$ of Θ to (5.2.4). In conclusion, we found an $m \leq \phi(l, g, \mathsf{d}, \mathsf{k}, \Theta)$ such that for all $i \in [m, m+g(m)]$ it holds that

$$\|x_i - x\| = \| \mathrm{J}_{\gamma \mathsf{B}}\, y_i - \mathrm{J}_{\gamma \mathsf{B}}\, y\| < 2^{-l}.$$

□

In contrast to the Douglas-Rachford Algorithm, we can only find a rate of metastability for the Peaceman-Rachford Algorithm from the proof in [1]. This is even though the iterations of both algorithms are defined almost identically. However, we are now in the degenerate case of the Krasnsoselski-Mann iteration to which Groetsch's theorem can no longer be applied. The different proof strategy in [1], however, is inherently noneffective as the convergence of the sequence $(\|y_n - y\|)_{n\in\mathbb{N}}$ is established by appealing to the monotone convergence principle which only admits a metastable quantitative version.

5.2 Extracting a Rate of Convergence

In this section we show how recent results of [12] combined with results due to the 2nd author can be used to obtain an effective rate of convergence for the Peaceman-Rachford algorithm if one additionally assumes that the operator A is uniformly monotone as well with a modulus of uniform monotonicity.

Definition 5.3 ([2]). *Let X be a Banach space and $S \subseteq X$. A nonexpansive mapping $T : S \to X$ is called strongly nonexpansive (SNE) if for all sequences $(x_n), (y_n)$ in S the following implication is true:*

$$\textit{if}\, (x_n - y_n) \textit{ is bounded and } \|x_n - y_n\| - \|Tx_n - Ty_n\| \to 0, \textit{then}\, (x_n - y_n) - (Tx_n - Ty_n) \to 0.$$

Lemma 5.4 ([7]). *A mapping $T : S \to X$ is strongly nonexpansive iff T satisfies*

$$(*) \begin{cases} \forall c \in \mathbb{N}\setminus\{0\}, k \in \mathbb{N}\ \exists n \in \mathbb{N}\ \forall x, y \in S \\ \qquad \big(\|x - y\| \leq c \wedge \|x - y\| - \|Tx - Ty\| < 2^{-n} \to \|(x - y) - (Tx - Ty)\| < 2^{-k}\big). \end{cases}$$

Definition 5.5 ([7]). *A function $\omega : \mathbb{N}^2 \to \mathbb{N}$ witnessing '$\exists n$' in $(*)$ above, i.e.*

$$(**) \begin{cases} \forall c \in \mathbb{N}\setminus\{0\}, k \in \mathbb{N}\ \forall x, y \in S \\ \big(\|x - y\| \leq c \wedge \|x - y\| - \|Tx - Ty\| < 2^{-\omega(c,k)} \to \|(x - y) - (Tx - Ty)\| < 2^{-k}\big), \end{cases}$$

is called an SNE-modulus of T.

By the above lemma, $T : S \to X$ is strongly nonexpansive iff it possesses an SNE-modulus.
The next lemma is a quantitative version of [12][Theorem 8.1(i)(c)]:

Lemma 5.6. *Let $\mathcal{H}$ be a Hilbert space and $\mathsf{A}, \mathsf{B} : \mathcal{H} \to 2^{\mathcal{H}}$ be maximally monotone operators which are uniformly monotone with a common modulus of uniform monotonicity Θ_K and $\gamma \geq 2^{-l}$. Then $\mathrm{R}_{\gamma \mathsf{A}}\, \mathrm{R}_{\gamma \mathsf{B}} : \mathcal{H} \to \mathcal{H}$ is strongly nonexpansive with*

$$\omega(l, c, k) := \Theta_c(k+2) + \lceil \log_2 c \rceil + l - 1$$

as a modulus of strong nonexpansivity.

Proof. First note that $\gamma\mathsf{A}$ is uniformly monotone with modulus $\Theta_{l,c}(k) := \Theta_c(k) + l$.
Claim: $-\mathrm{R}_{\gamma\mathsf{A}}$ (and analogously $-\mathrm{R}_{\gamma\mathsf{B}}$) is strongly nonexpansive with modulus

$$\omega'(l,c,k) := \Theta_c(k+1) + \lceil \log_2 c \rceil + l - 1.$$

Proof of Claim: By '(4)' in [12] we have that

$$(+)\ \mathrm{gra}(\gamma\mathsf{A}) = \left\{ \frac{1}{2}(x + \mathrm{R}_{\gamma\mathsf{A}}\, x, x - \mathrm{R}_{\gamma\mathsf{A}}\, x) : x \in \mathcal{H} \right\}.$$

Now assume that $c \geq 1$ and for $x, y \in \mathcal{H}$ that

$$\|x - y\| \leq c \text{ and } \|x - y\| - \|(-\mathrm{R}_{\gamma\mathsf{A}})x - (-\mathrm{R}_{\gamma\mathsf{A}})y\| < 2^{-\omega'(l,c,k)}.$$

Then (using that $a \mapsto a^2$ is Lipschitz continuous on $[0, c]$ with Lipschitz constant $2c$ and the nonexpansivity of $\mathrm{R}_{\gamma\mathsf{A}}$)

$$\|x - y\|^2 - \|\mathrm{R}_{\gamma\mathsf{A}}\, x - \mathrm{R}_{\gamma\mathsf{A}}\, y\|^2 < 2c \cdot 2^{-\omega'(l,c,k)} \leq 2^{-\Theta_{l,c}(k+1)+2}.$$

Thus

$$\begin{aligned}
&\langle \tfrac{1}{2}(x + \mathrm{R}_{\gamma\mathsf{A}}\, x) - \tfrac{1}{2}(y + \mathrm{R}_{\gamma\mathsf{A}}\, y), \tfrac{1}{2}(x - \mathrm{R}_{\gamma\mathsf{A}}\, x) - \tfrac{1}{2}(y - \mathrm{R}_{\gamma\mathsf{A}}\, y)\rangle \\
&= \tfrac{1}{4}\langle (x - y) + (\mathrm{R}_{\gamma\mathsf{A}}\, x - \mathrm{R}_{\gamma\mathsf{A}}\, y), (x - y) - (\mathrm{R}_{\gamma\mathsf{A}}\, x - \mathrm{R}_{\gamma\mathsf{A}}\, y)\rangle \\
&= \tfrac{1}{4}\left(\tfrac{\|2(x-y)\|^2}{4} + \tfrac{\|2(\mathrm{R}_{\gamma\mathsf{A}}\, x - \mathrm{R}_{\gamma\mathsf{A}}\, y\|^2}{4} \right) \\
&= \tfrac{1}{4}\left(\|x - y\|^2 - \|\mathrm{R}_{\gamma\mathsf{A}}\, x - \mathrm{R}_{\gamma\mathsf{A}}\, y\|^2 \right) < 2^{-\Theta_{l,c}(k+1)}.
\end{aligned}$$

Hence, by (+) and the definition of $\Theta_{l,c}$

$$\frac{1}{2}\|(x - y) - ((-\mathrm{R}_{\gamma\mathsf{A}})x - (-\mathrm{R}_{\gamma\mathsf{A}})y)\| = \left\| \frac{1}{2}(x + \mathrm{R}_{\gamma\mathsf{A}}\, x) - \frac{1}{2}(y + \mathrm{R}_{\gamma\mathsf{A}}\, y) \right\| < 2^{-k-1}$$

using that

$$\left\| \frac{1}{2}(x + \mathrm{R}_{\gamma\mathsf{A}}\, x) - \frac{1}{2}(y + \mathrm{R}_{\gamma\mathsf{A}}\, y) \right\| \leq \frac{1}{2}\|x - y\| + \frac{1}{2}\|\mathrm{R}_{\gamma\mathsf{A}}\, x - \mathrm{R}_{\gamma\mathsf{A}}\, y\| \leq \|x - y\| \leq c.$$

Thus

$$\|(x - y) - ((-\mathrm{R}_{\gamma\mathsf{A}})x - ((-\mathrm{R}_{\gamma\mathsf{A}})y)\| < 2^{-k}$$

which concludes the proof of the Claim.
Now consider $x, y \in \mathcal{H}$ with $\|x - y\| \leq c$ and

$$\begin{aligned}
&\|x - y\| - \|\mathrm{R}_{\gamma\mathsf{A}}\, \mathrm{R}_{\gamma\mathsf{B}}\, x - \mathrm{R}_{\gamma\mathsf{A}}\, \mathrm{R}_{\gamma\mathsf{B}}\, y\| = \\
&\underbrace{\|x - y\| - \|(-\mathrm{R}_{\gamma\mathsf{B}})x - (-\mathrm{R}_{\gamma\mathsf{B}})y\|}_{\geq 0} + \underbrace{\|\mathrm{R}_{\gamma\mathsf{B}}\, x - \mathrm{R}_{\gamma\mathsf{B}}\, y\| - \|(-\mathrm{R}_{\gamma\mathsf{A}})\,\mathrm{R}_{\gamma\mathsf{B}}\, x - (-\mathrm{R}_{\gamma\mathsf{A}})\,\mathrm{R}_{\gamma\mathsf{B}}\, y\|}_{\geq 0} \\
&< 2^{-\omega(l,c,k)} = 2^{-\omega'(l,c,k+1)}.
\end{aligned}$$

Then by the Claim

$$\|(x - y) + (\mathrm{R}_{\gamma\mathsf{B}}\, x - \mathrm{R}_{\gamma\mathsf{B}}\, y)\| = \|(x - y) - ((-\mathrm{R}_{\gamma\mathsf{B}})x - (-\mathrm{R}_{\gamma\mathsf{B}})y)\| < 2^{-k-1}$$

and likewise

$$\|(\mathrm{R}_{\gamma\mathsf{B}}\, x - \mathrm{R}_{\gamma\mathsf{B}}\, y) + (\mathrm{R}_{\gamma\mathsf{A}}\, \mathrm{R}_{\gamma\mathsf{B}}\, x - \mathrm{R}_{\gamma\mathsf{A}}\, \mathrm{R}_{\gamma\mathsf{B}}\, y)\| < 2^{-k-1}$$

and so

$$\begin{aligned}
&\|(x - y) - (\mathrm{R}_{\gamma\mathsf{A}}\, \mathrm{R}_{\gamma\mathsf{B}}\, x - \mathrm{R}_{\gamma\mathsf{A}}\, \mathrm{R}_{\gamma\mathsf{B}}\, y)\| \\
&\leq \|(x - y) + (\mathrm{R}_{\gamma\mathsf{B}}\, x - \mathrm{R}_{\gamma\mathsf{B}}\, y)\| + \|(\mathrm{R}_{\gamma\mathsf{B}}\, x - \mathrm{R}_{\gamma\mathsf{B}}\, y) + (\mathrm{R}_{\gamma\mathsf{A}}\, \mathrm{R}_{\gamma\mathsf{B}}\, x - \mathrm{R}_{\gamma\mathsf{A}}\, \mathrm{R}_{\gamma\mathsf{B}}\, y)\| < 2^{-k}
\end{aligned}$$

which finishes the proof of the lemma. □

Together with [7][Theorem 2.8] one obtains

Proposition 5.7. *Let* $y_{n+1} := \mathrm{R}_{\gamma\mathsf{A}}\, \mathrm{R}_{\gamma\mathsf{B}}\, y_n$ *and* $\mathsf{d} \geq \|y_0 - y\|, \mathsf{d} \geq 1$ *for some* $y \in \mathsf{Fix}(\mathrm{R}_{\gamma\mathsf{A}}\, \mathrm{R}_{\gamma\mathsf{B}})$. *Then*

$$\forall k \in \mathbb{N}\ \forall n \geq 2^{\omega(l,\mathsf{d},k) + \lceil \log_2 \mathsf{d} \rceil} \left(\|y_{n+1} - y_n\| < 2^{-k} \right).$$

Let

$$\mathsf{T} := \mathrm{J}_{\gamma \mathsf{A}} \mathrm{R}_{\gamma \mathsf{A}} + \mathsf{Id} - \mathrm{J}_{\gamma \mathsf{B}} = \mathrm{J}_{\gamma \mathsf{A}}(2\,\mathrm{J}_{\gamma \mathsf{B}} - \mathsf{Id}) + \mathsf{Id} - \mathrm{J}_{\gamma \mathsf{B}}\,.$$

Then by [1][Proposition 4.31(i)]

$$2\mathsf{T} - \mathsf{Id} = (2\,\mathrm{J}_{\gamma \mathsf{A}} - \mathsf{Id})(2\,\mathrm{J}_{\gamma \mathsf{B}} - \mathsf{Id}) = \mathrm{R}_{\gamma \mathsf{A}} \mathrm{R}_{\gamma \mathsf{B}},\ \text{i.e. } \mathsf{T} = \frac{\mathrm{R}_{\gamma \mathsf{A}} \mathrm{R}_{\gamma \mathsf{B}} + \mathsf{Id}}{2}$$

and so

$$\| \mathrm{R}_{\gamma \mathsf{A}} \mathrm{R}_{\gamma \mathsf{B}}\, y - y\| < 2^{-k+1} \to \|\mathsf{T}y - y\| = \frac{1}{2}\| \mathrm{R}_{\gamma \mathsf{A}} \mathrm{R}_{\gamma \mathsf{B}}\, y - y\| < 2^{-k}.$$

Thus we obtain the

Corollary 5.8. *Let $y_0 \in \mathcal{H}$, $\gamma \in (0,\infty)$ and $\mathsf{A}, \mathsf{B} : \mathcal{H} \to 2^{\mathcal{H}}$ be maximally monotone operators which are uniformly monotone with a common modulus Θ_K of uniform monotonicity, where K can be taken as the fixed number d, where $\mathsf{d} \geq \|y_0 - y\|, \mathsf{d} \geq 1$ for some $y \in \mathsf{Fix}(\mathrm{R}_{\gamma \mathsf{A}} \mathrm{R}_{\gamma \mathsf{B}})$. Let $(x_n), (y_n), (z_n)$ be defined as in the Peaceman-Rachford algorithm.*
Then

$$\forall k \in \mathbb{N}\ \forall n \geq 2^{\omega(l,\mathsf{d},k-1)+\lceil \log_2 \mathsf{d}\rceil} = 2^{\omega'(l,\mathsf{d},k)+\lceil \log_2 \mathsf{d}\rceil}\ \left(\|z_n - x_n\| < 2^{-k}\right).$$

Proof. This follows from $z_n - x_n = \mathsf{T}y_n - y_n$, $y_{n+1} = \mathrm{R}_{\gamma \mathsf{A}} \mathrm{R}_{\gamma \mathsf{B}}\, y_n$ and the comment above via Proposition 5.7. □

Theorem 5.9. *Let $\mathsf{A}, \mathsf{B} : \mathcal{H} \to 2^{\mathcal{H}}$ be maximally monotone operators which are uniformly monotone with a common modulus Θ_K of uniform monotonicity (where K can be taken as d below). Let $\gamma \geq 2^{-l}, y_0 \in \mathcal{H}$ and $x \in \mathsf{zer}(\mathsf{A}+\mathsf{B})$ and $(x_n), (y_n)$ be defined as in the Peaceman-Rachford algorithm. Let $y \in \mathsf{Fix}(\mathrm{R}_{\gamma \mathsf{A}} \mathrm{R}_{\gamma \mathsf{B}})$ be such that $x = \mathrm{J}_{\gamma \mathsf{B}}\, y$ and $\|y_0 - y\| \leq \mathsf{d}$ and $\|\mathsf{T}y_0 - y_0\| \leq \mathsf{M}$, where $\mathsf{d}, \mathsf{M} \geq 1$. Then we have the following rate of convergence for (x_n) towards x :*

$$\forall k \in \mathbb{N}\ \forall n \geq \Psi(\mathsf{k}, \mathsf{d}, l, \Theta_{\mathsf{d}}, \mathsf{M})\ \left(\|x_n - x\| < 2^{-k}\right),$$

where

$$\Psi(\mathsf{k}, \mathsf{d}, l, \Theta_{\mathsf{d}}, \mathsf{M}) := 2^{\omega'(l,\mathsf{d},\Theta_{\mathsf{d}}(\mathsf{k})+l+\lceil \log_2(\mathsf{M}+\mathsf{d})\rceil)+\lceil \log_2 \mathsf{d}\rceil}.$$

Proof. The result follows from the proof of the quantitative convergence result of the Douglas-Rachford algorithm given in Theorem 4.3(b) when stipulating that $\lambda_n := 2$ for all $n \in \mathbb{N}$ observing that the condition made there that $\sum \lambda_n(2-\lambda_n) = \infty$ (which rules out this choice) is only used to obtain the rate φ of convergence for $\|z_n - x_n\| \to 0$ in Theorem 4.2 which - by Corollary 5.8 - we can replace by the rate given in that corollary. □

5.3 Discussion of the Results (for logicians)

The treatment of the operators $\mathsf{A}, \mathsf{B}, \mathsf{A}+\mathsf{B}$ is as in the case of the Douglas-Rachford algorithm. However, in the case of Theorem 5.2 neither would it be sufficient to just extract a bound on $\exists n\, (\|x_n - x\| < 2^{-l})$ as in the case of Tseng's and the forward-backward algorithm (due to the lack of monotonicity) nor can we decompose the proof into two separate statements of the form $\forall\exists$ to which metatheorems then are applicable as in the case of the Douglas-Rachford algorithm (due to the missing asymptotic regularity result $\|x_n - z_n\| \to 0$). Since the convergence statement is of the form $\forall\exists\forall$ via

$$(*)\ \forall k \in \mathbb{N}\ \exists n \in \mathbb{N}\ \forall m \geq n\, (\|x_m - x\| \leq 2^{-k}),$$

which is not allowed in the logical metatheorems (for theories based on full classical logic), we have to replace $(*)$ by its (noneffectively) equivalent metastable formulation

$$\forall k \in \mathbb{N}\ \forall g : \mathbb{N} \to \mathbb{N}\ \exists n \in \mathbb{N}\ \forall i \in [n, n+g(n)]\ \left(\|x_n - x\| < 2^{-k}\right)$$

which (disregarding the bounded quantifier '$\forall i \in [n, n+g(n)]$') is of the form $\forall\exists$.

The logical analysis in section 5.2 follows directly that of the Douglas-Rachford algorithm except that the rate of asymptotic regularity for the Krasnoselski-Mann iteration used in the latter is now replaced by an asymptotic regularity result for strongly nonexpansive mappings in the former case (whose logical underpinning is explained in [7]).

References

[1] Heinz H. Bauschke and Patrick L. Combettes. Convex Analysis and Monotone Operator Theory in Hilbert Spaces. Canadian Mathematical Society, 2011.

[2] Ronald E. Bruck, and Simeon Reich. Nonexpansive projections and resolvents of accretive operators in Banach spaces. Houston J. Math. **3**, pp. 459-470 (1977). Israel Journal of Mathematics 22, pp. 81-86 (1975).

[3] Christian Clason and Tuomo Valkonen. Introduction to Nonsmooth Analysis and Optimization. http://arxiv.org/pdf/2001.00216.

[4] Philipp Gerhardy and Ulrich Kohlenbach. General logical metatheorems for functional analysis. Transactions of the American Mathematical Society vol. 360, pp. 2615-2660 (2008).

[5] Ulrich Kohlenbach. Some logical metatheorems with applications in functional analysis. Transactions of the American Mathematical Society vol. 357, pp. 89-128 (2005).

[6] Ulrich Kohlenbach. Applied Proof Theory: Proof Interpretations and Their Use in Mathematics. Springer Science & Business Media, 2008.

[7] Ulrich Kohlenbach. On the quantitative asymptotic behavior of strongly nonexpansive mappings in Banach and geodesic spaces. Israel Journal of Mathematics **216**, pp. 215-246 (2016).

[8] Ulrich Kohlenbach. A polynomial rate of asymptotic regularity for compositions of projections in Hilbert space. Foundations of Computational Mathematics vol. 19, pp. 83-99 (2019).

[9] Ulrich Kohlenbach. Proof-theoretic methods in nonlinear analysis. In: B. Sirakov, P. Ney de Souza, M. Viana (eds.), Proceedings of the International Congress of Mathematicians 2018 (ICM 2018), Vol. 2, pp. 61-82, World Scientific, Singapore, 2019.

[10] Ulrich Kohlenbach and Angeliki Koutsoukou-Argyraki. Rates of convergence and metastability for abstract cauchy problems generated by accretive operators. Journal of Mathematical Analysis and Applications vol. 423, pp. 1089-1112 (2015).

[11] Ulrich Kohlenbach and Nicholas Pischke. Proof theory and nonsmooth analysis. Philosophical Transactions of the Royal Society A, 381:20220015, 21 pages (2023).

[12] Leon Liu, Walaa M. Moursi and Jon Vanderwerff. Strongly nonexpansive mappings revisited: uniform monotonicity and operator splitting. SIAM Journal on Optimization vol. 33, pp. 2570-2597 (2023).

[13] Nicholas Pischke. Logical metatheorems for accretive and (generalized) monotone set-valued operators. Journal of Mathematical Logic vol. 24, 2350008, 59 pages (2024).

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

[15] Terence Tao. Norm convergence of multiple ergodic averages for commuting transformations. Ergodic Theory and Dynamical Systems **28**, pp. 657-688 (2008).

[16] Jacqueline Treusch. Logical extraction of rates of convergence for algorithms computing zeros of monotone operators. Bachelor Thesis, TU Darmstadt, 2024.