PROOF MINING AND THE CONVEX FEASIBILITY PROBLEM: THE CURIOUS CASE OF DYKSTRA'S ALGORITHM

PEDRO PINTO
Department of Mathematics, Technische Universität Darmstadt

Abstract. In a recent proof mining application, the proof-theoretical analysis of Dykstra's cyclic projections algorithm resulted in quantitative information expressed via primitive recursive functionals in the sense of Gödel. This was surprising as the proof relies on several compactness principles and its quantitative analysis would require the functional interpretation of arithmetical comprehension. Therefore, a priori one would expect the need of Spector's bar-recursive functionals. In this paper, we explain how the use of bounded collection principles allows for a modified intermediate proof justifying the finitary results obtained, and discuss the approach in the context of previous eliminations of weak compactness arguments in proof mining.

§1. Introduction. Famously, Georg Kreisel asked '*what more* [do] *we know about a formally derived theorem F than if we merely know that F is true*?', [38, p. 110]. This question prompted the development of what is nowadays known as the *proof mining* research program: the analysis of mathematical proofs using proof-theoretical techniques as a way to obtain a deeper understanding of the proof, stripping it to its essential arguments and subsequentially obtaining new effective information. In the last twenty-five years, through the work of Ulrich Kohlenbach and his collaborators, this line of research has since been greatly developed (see [28] for a book treatment and [31] for a recent survey). In practice, the 'proof miner' begins by selecting a *prima facie* noneffective proof of a mathematical result for which one expects finitary information to be of interest. Applying some variant of Kurt Gödel's Dialectica (most frequently Kohlenbach's monotone function interpretation [25]; in some selected cases, Fernando Ferreira and Paulo Oliva's bounded functional interpretation [13, 16]), one is not only guided into the correct quantitative restatement of the theorem but is also able to reformulate the central arguments of the proof into a new proof which validates the correctness of the extracted information. The quality of these results seems to rest on two distinct aspects: first, the chosen result must be of interest to the community with its finitary information wanted; second, the extracted information must be of a simple nature if not new. So called *logical metatheorems* guarantee that, in certain general

Received: July 2, 2024.
2020 *Mathematics Subject Classification*: Primary 03F10, 03F35, Secondary 47J25, 41A65, 41A29, 90C25, 47H10.
Key words and phrases: proof mining functional interpretations, bounded collection, compactness.

doi:10.1017/S1755020325100695

situations, we can be sure that the quantitative analysis will yield effective information (e.g., [17, 26], the recent [43], and also [12]).

With many applications in nonlinear analysis and convergence statements, such quantitative information is frequently in the form of rates of convergence or in the form of rates for (what in 2008 Terence Tao [47] dubbed) the metastability property, long known to logicians as Kreisel's no-counterexample interpretation of the Cauchy property. The complexity of the extracted information is a reflection of the arguments employed in proving the mathematical result. In light of the logical metatheorems, one expects that the use of arithmetical comprehension (lurking in common place mathematical arguments, as for example whenever one uses compactness principles or considers the *infima* of positive real sequences, like in projection arguments) would require the use of Spector's bar-recursive functionals [46]. However, that is not an appealing solution, as it goes against the goal of "simple information" and would not be appreciated by the general mathematician. It so happens that in most cases the use of such comprehension principles can actually be avoided. It may happen via an 'arithmetization' of the argument, ε-weakening, or by other simplifications to the proof. This feature is relevant, not only for the quality of the extracted information, but furthermore because simpler proofs make generalizations to completely new theorems a stronger possibility. It is natural to then wonder if such phenomenon of proof-theoretical tameness is due to a selective choice of examples or to an unconscious restriction from mathematicians. The existence of a broad variety of case studies tackling many set-theoretically complicated notions, makes a strong heuristic case for the latter. Be that as it may, it rests on the logician's shoulders to argue for the existence of an elementary proof. When it is exists, a simpler proof would entail that the corresponding extracted information would be immediately expressed by primitive recursive functionals in the sense of Gödel (first defined in [18], but see also [22]).

As a follow up to the quantitative studies of Browder's and Wittmann's fixed point theorems due to Kohlenbach [30], Ferreira, Leuştean and the author developed in [15] a general approach for the elimination of sequential weak compactness via certain bounded collection principles when in the context of proof mining results. In the same paper, the argument was employed to three simple case studies (an overview of Browder and Wittmann's analysis in [30] as well as Bauschke's generalization of Wittmann's theorem, which was itself contained as a particular case of a previous quantitative study of the Hybrid Steepest Descent Method due to Körnlein [36, 37]), and has since been applied to many other cases (e.g., [6, 7, 40]). As already mentioned, the existence of an elementary proof may facilitate the generalization of the original result: an interesting example is [8], where the elimination of weak compactness allowed for the generalization from Hilbert spaces to the nonlinear setting of CAT(0) spaces. However, so far all the applications of the removal technique have a very similar feeling: the algorithm is some variation of Halpern's iteration [21] and the proof structure is reminiscent of a prevalent argument due to Wittmann in [49].

The main goal of this paper is to discuss a novel elimination of compactness arguments crucial in the convergence proof of a different kind of iteration: Dykstra's cyclic projection method. A central problem in convex optimization, known as the convex feasibility problem, is that of finding a point in the intersection of a finite number of convex closed subsets of a Hilbert space. A very useful algorithm to approximate a solution is the method of alternating projections due to von Neumann [48] and Halperin [20]. This algorithm is well understood and provides the 'optimal solution'

(i.e., the closest one to the initial guess) when in the context of vector subspaces (or more generally, translates of vector spaces with a nonempty intersection). However, if one just assumes the sets to be closed convex subsets, then by the work of Bregman [3] and Hundal [24], the method of alternating projections, in general, will only converge in the weak topology. In finite dimensional Hilbert spaces, this is not an issue since in that case the weak (inner product) and the strong (norm) topology coincide. However, already in simple examples, the limit would not necessarily be the optimal solution but just some point in the intersection. It is here that Dykstra's algorithm comes successfully into play. Consider $C_1, \dots, C_m$ to be $m \geq 2$ convex subsets of a Hilbert space with nonempty intersection. For each $n \geq 1$, take $C_n := C_{j_n}$, where $j_n := [n-1]+1$ with $[r] := r \bmod m$, i.e., we enumerate the sets cyclically. Let P_n denote the metric projection onto the nonempty convex set C_n. Dykstra's iteration is given by:

$$\begin{cases} x_0 \in X \\ q_{-(m-1)} = \cdots = q_0 = 0 \end{cases}, \forall n \geq 1 \begin{cases} x_n := P_n(x_{n-1} + q_{n-m}) \\ q_n := x_{n-1} + q_{n-m} - x_n. \end{cases} \tag{D}$$

It begins by following the alternating projection method, but then starts to incorporate correction terms which are updated after every m steps. Boyle and Dykstra proved the following result.

THEOREM 1.1 [2]. *Let $C_1, \dots, C_m$ be $m \geq 2$ closed convex subsets of a Hilbert space such that $\bigcap_{j=1}^m C_j \neq \emptyset$. For $x_0 \in X$, let (x_n) be the iteration generated by (D). Then, (x_n) converges strongly to the point in $\bigcap_{j=1}^m C_j$ closest to x_0.*

Moreover, Dykstra's method reduces to von Neumann-Halperin's alternating projection method when the C_j's are vector subspaces. Thus, Dykstra's algorithm is a proper generalization of the method of alternating projections. The convex feasibility problem is of central relevance for the mathematical community due to its relation to a broad variety of problems, from statistics to differential equations. Despite its usefulness and contrary to the alternating method, not much was known regarding the quantitative behavior of Dykstra's algorithm. Hence, this was a prime candidate for a proof mining study. The convergence proof makes a crucial use of certain arguments (sequential weak compactness, Bolzano–Weirstrass compactness, and infinite pigeonhole principle) which *a priori* hinder a simple proof-theoretical analysis. We shall discuss how it was possible in [41] to obtain a finitary version of this result where the quantitative data does not require the use of bar-recursive functionals. Furthermore we will discuss that, contrary to previous eliminations of compactness arguments, here there is no connection with Halpern-type iterations and the convergence proof is significantly different. Nevertheless, we show how the quantitative argument fits in the context of the macro developed in [15].

§2. A false ideal world. We begin by recalling the result allowing the elimination of certain weak compactness arguments in proof-theoretical studies, essentially following [15]. After a short discussion on the motivational issue, we introduce the bounded functional interpretation in a formal setting suitable for the extraction of primitive recursive (in the sense of Gödel) information from classical proofs in the context of inner product spaces. Finally, we introduce both the general principle from [15] as well as its corresponding finitary counterpart, in the particular form that we shall need for the remaining of the paper.

2.1. Wittmann's convergence theorem. Consider $(X, \langle\cdot,\cdot\rangle)$ a (real) Hilbert space with inner product $\langle\cdot,\cdot\rangle$ and the induced norm $\|x\| := \sqrt{\langle x, x\rangle}$. Let C be a nonempty closed subset which is assumed to be convex (i.e., closed for taking line segments). In [30], the quantitative analysis of the proof of Wittmann's theorem[1] resulted in a simple rate of metastability that is—for (x_n) the relevant iteration—a functional $\Phi : \mathbb{N} \times \mathbb{N}^{\mathbb{N}} \to \mathbb{N}$ such that

$$\forall k \in \mathbb{N}\, \forall f : \mathbb{N} \to \mathbb{N}\, \exists n \leq \Phi(k, f)\, \forall i, j \in [n; n + f(n)] \left(\|x_i - x_j\| \leq \frac{1}{k+1} \right),$$

which was primitive recursive in the sense of Gödel (actually primitive recursive in the sense of Kleene on the counterfunction f). This was surprising as the analysed proof relies crucially on two troublesome arguments: a projection onto the set of fixed points of a map $T : C \to C$, denoted by $\mathrm{Fix}(T) := \{x \in C : T(x) = x\}$; together with a sequential weak compactness argument. These arguments are logically justified by the presence of arithmetical comprehension, and therefore one would expect the need for functionals defined by bar-recursion. Let us give an overview of the proof-theoretical study in [30]. Regarding the projection (of a certain point u) onto $\mathrm{Fix}(T)$,

$$\exists x \in \mathrm{Fix}(T)\, \forall y \in \mathrm{Fix}(T)\, \forall k \in \mathbb{N} \left(\|u - x\| \leq \|u - y\| + \frac{1}{k+1} \right),$$

Kohlenbach showed that the weaker version where x is allowed to depend on k

$$\forall k \in \mathbb{N}\, \exists x \in \mathrm{Fix}(T)\, \forall y \in \mathrm{Fix}(T) \left(\|u - x\| \leq \|u - y\| + \frac{1}{k+1} \right)$$

is already sufficient to derive the Cauchy property of the iteration—an instance of what he called ε-weakening. This insight can be understood by the fact that the Cauchy statement follows from proving that for each ε (i.e., $\frac{1}{k+1}$) there exists some point x for which the terms of the sequence are all eventually $\varepsilon/2$-close. For this, an $\frac{1}{k+1}$-almost projection point would suffice. Note that this says that the arguments in the proof can be relaxed to their 'imperfect' versions allowing for error terms—an instance of both 'ideal' elements and infinitary arguments removal in the general sense of Hilbert's program. The relevance for the discussion is the fact that this weaker version can be proven simply by induction and so its contribution to the extracted information would be in the form of a functional recursively defined. On the other hand, the absence of contribution to the final rate from the weak compactness principle had a more unsatisfying ad hoc explanation. The proof mining analysis revealed that the use of weak compactness was actually very mild. Essentially, the role of the weak limit could be replaced by that of a point of the iteration. Overall, the proof-theoretical approach uncovered an elementary proof where arithmetical comprehension was not needed, and so the corresponding quantitative information was naturally described by primitive recursive functionals in the sense of Gödel.

Following a suggestion of Kohlenbach, in [15] Ferreira, Leuştean and the author looked again at the proof of Wittmann's convergence theorem with the goal of

[1] Actually, the work in [30] also analysed an important convergence result due to Browder. The proof itself follows similar lines to those of Wittmann, and even predates it. We choose to focus on Wittmann's theorem as it regards the convergence of the Halpern iteration which connects with several other results in our discussion.

providing a more uniform theoretical understanding of why such uses of arithmetical comprehension don't contribute towards the expected complexity of the data obtained. It turned out that this kind of argument could be bypassed by the use of certain bounded collection principles which don't contribute to the final bounding information. In [15], this was explained through the use of the bounded functional interpretation which incorporates a bounded collection principle directly into the interpretation. The same result can be obtained via Kohlenbach's monotone functional interpretation together with the generalized uniform boundedness principle $\exists\text{-UB}^X$ [19, 27]. We shall base our arguments in the framework of the bounded functional interpretation which, by infusing the bounded collection principles directly into the interpretation, appears specifically tailored to tackle this kind of issues.

2.2. A formal system for inner product spaces. Our underlying framework is that of Peano arithmetic in all finite types which is extended to a typed system with a new base type and to include an intensional (i.e., rule-governed) majorizability primitive notion. We are interested in a suitable formal system for abstract inner product spaces, $\mathrm{PA}^{\omega}_{\trianglelefteq}[X, \langle\cdot,\cdot\rangle]$. Following, in essence, [28, sec. 17.3], we treat these spaces via their characterization as the normed spaces for which the parallelogram law holds.

DEFINITION 2.1. *Let T^X be the set of all finite types with an new base type X, recursively defined by*

(i) $0, X \in \mathsf{T}^X$ (*the base types*);
(ii) *if* $\rho, \sigma \in \mathsf{T}^X$, *then* $\rho \to \sigma \in \mathsf{T}^X$.

Here X will stand for an abstract inner product space. The idea of considering new abstract types, first done in [26] for bounded metric spaces, was of great importance to the proof mining program—undoubtedly a catalyst to all current applications. Indeed, one was no longer restricted to "computable" or "representable" spaces, and it gave sense to the extraction of computable bounds in many arbitrary spaces. Furthermore, simple and natural statements can be formulated if one is unburdened by the need of working with representatives of mathematical objects and can work directly with the actual objects.

We distinguish the subset of T^X of the types where X is absent, writing T. For any $\rho \in \mathsf{T}^X$, denote by $\hat{\rho} \in \mathsf{T}$ the finite type obtained from ρ by replacing every instance of X with 0. One frequently denotes the type $0 \to 0$ simply by 1.

We consider $\mathcal{L}^{\omega,X}_{\trianglelefteq}$ a typed language in all finite types of T^X. For each type $\rho \in \mathsf{T}^X$, we have a countable number of variables $x^\rho, y^\rho, z^\rho, \ldots$. We also have the arithmetical constants 0^0 (zero), S^1 (successor) and simultaneous recursors $\underline{R}_\rho$ (extended to encompass the new types), together with the constants associated with functional completeness (i.e., the combinators extended to all the types in T^X, which ensure general λ-abstraction). Terms of $\mathcal{L}^{\omega,X}_{\trianglelefteq}$ are build up from the constants and variables via suitable term application: if t is a term of type $\rho \to \sigma$ and u is a term of type ρ, then $t(u)$ is also a term, in this case of type σ. It should be clear that by writing t^ρ we are stating that the term t is of type ρ.

The language $\mathcal{L}^{\omega,X}_{\trianglelefteq}$ includes predicate symbols $\trianglelefteq_\rho$ for each $\rho \in \mathsf{T}^X$ as well as equality at type 0, denoted by $=_0$. The atomic formulae are either of the form $t =_0 q$ for terms t, q of type 0, or of the form $t \trianglelefteq_\rho q$ for t a term of type ρ and q a term of type $\hat{\rho}$. The general formulae are build up from the atomic ones using the usual propositional logic symbols $\neg, \wedge, \vee, \to$, by (unbounded) quantifications $\forall x^\rho$, $\exists x^\rho$, and also by bounded

quantifications $\forall x \trianglelefteq_\rho t$, $\exists x \trianglelefteq_\rho t$, where x is a variable of type ρ and t is a term of type $\hat{\rho}$ where x does not occur. A formula without unbounded quantifications is called a bounded formula. Whenever $t \trianglelefteq_\rho t$, where the type ρ of t is necessarily in T, we say that the term t is monotone. A monotone universal quantification, which we abbreviate by $\tilde{\forall} x\, A(x)$, is a quantification of the form $\forall x \left(x \trianglelefteq_\rho x \to A(x)\right)$, where A is any formula of $\mathcal{L}_{\trianglelefteq}^{\omega,X}$ and $\rho \in \mathsf{T}$. Analogous for monotone existential quantifiers.

In order to discuss inner product spaces in a formal system suitable for bound extraction we include further normed space constants

- 0_X of type X, standing for the zero vector;
- $+_X$ of type $X \to (X \to X)$, standing for vector addition;
- $-_X$ of type $X \to X$, standing for the symmetric of a vector;
- $\cdot_X$ of type $X \to (1 \to X)$, standing for scalar multiplication.

We shall write simply $x -_X y$ for $x +_X (-_X y)$. The language also contains a constant $\|\cdot\|$, standing for the norm function, with type $X \to 1$. As usual, real numbers are represented in this system by type 1 terms. There are several ways to achieve a representation of the real numbers through functions $f : \mathbb{N} \to \mathbb{N}$. For example, in [28] Kohlenbach uses fast converging Cauchy sequences of rational numbers (themselves coded by natural numbers); in [11], Engrácia used the signed-digit representation. In any case, the relevant point to note is that the relations $=_{\mathbb{R}}$ and $\leq_{\mathbb{R}}$ between (representations of) real numbers are defined by Π^0_1-formulae, and the strict relation $<_{\mathbb{R}}$ by a Σ^0_1-formula. Furthermore, there is a primitive recursive operation $f \mapsto \widetilde{f}$ which converts any function $f : \mathbb{N} \to \mathbb{N}$ into function $\widetilde{f} : \mathbb{N} \to \mathbb{N}$ encoding a real number and such that $\widetilde{f}$ remains f whenever it is already representation of a real number. This can be carried out in our system and ensures that one can discuss statements about the real numbers without requiring an additional type (see [14] for that alternative approach). We assume that for every x^X, the term $\|x\|$ always represents a real number, which can be stated by a universal formula.

Let us introduce the theory $\mathrm{PA}^{\omega}_{\trianglelefteq}[X, \langle\cdot,\cdot\rangle]$ with the language $\mathcal{L}_{\trianglelefteq}^{\omega,X}$ in a classical setting. Equality at the type 0, $=_0$, is still the unique primitive equality, and is subject to the usual axioms of equality for terms of type 0. Equality between terms of type X is a defined relation given by a universal formula, namely

$$x =_X y \;:\equiv\; \|x -_X y\| =_{\mathbb{R}} 0_{\mathbb{R}},$$

where $0_{\mathbb{R}}$ is the canonical representation of the real number zero by a type 1 term. Equalities at higher types are defined in a pointwise manner. We do not include the extensionality axiom, which is known to be problematic in functional interpretations, nor do we incorporate a weak rule of extensionality. Instead, we adopt a minimal treatment of equality, as is common in the bounded functional interpretation setting (see e.g., [16]). The axioms pertaining to the arithmetical constants and the combinators are the usual ones (extended to T^X). The system includes induction for every formula. We also consider the real vector space (universal) axioms formulated using the equality $=_X$, namely stating: $+_X$ is associative and commutative, 0_X is the neutral element for $+_X$, $-_X$ gives the $+_X$-inverse, the scalar multiplication $\cdot_X$ is compatible with respect to the multiplication of (representations of) real numbers, $1_{\mathbb{R}}$ is the identity for scalar multiplication, and the two distributivity laws between scalar multiplication and the addition $+_{\mathbb{R}}$ and $+_X$. We have suitable axioms for the norm (following [26]):

$$(\mathbf{R})\quad \forall x^X\,\forall k^0\left(\|x\|(k) =_0 \widetilde{\|x\|}(k)\right)^2$$
$$(\mathbf{N}_1)\quad \forall x^X\,(\|x -_X x\| =_{\mathbb{R}} 0_{\mathbb{R}})$$
$$(\mathbf{N}_2)\quad \forall x^X\,\forall y^y\,(\|x -_X y\| =_{\mathbb{R}} \|y -_X x\|)$$
$$(\mathbf{N}_3)\quad \forall x^X\,\forall y^X\,\forall z^X\,(\|x -_X z\| \leq_{\mathbb{R}} \|x -_X y\| +_{\mathbb{R}} \|y -_X z\|)$$
$$(\mathbf{N}_4)\quad \forall \alpha^1\,\forall x^X\,\forall y^X\,(\|\alpha \cdot_X x -_X \alpha \cdot_X y\| =_{\mathbb{R}} |\tilde{\alpha}|_{\mathbb{R}} \cdot_{\mathbb{R}} \|x -_X y\|)$$
$$(\mathbf{N}_5)\quad \forall \alpha^1\,\forall \beta^1\,\forall x^X\left(\|\alpha \cdot_X x -_X \beta \cdot_X x\| =_{\mathbb{R}} |\tilde{\alpha} -_{\mathbb{R}} \tilde{\beta}|_{\mathbb{R}} \cdot_{\mathbb{R}} \|x\|\right)$$
$$(\mathbf{N}_6)\quad \forall x^X\,\forall y^X\,\forall u^X\,\forall v^X\left(\|(x +_X y) -_X (u +_X v)\| \leq \|x -_X u\| +_{\mathbb{R}} \|y -_X v\|\right)$$
$$(\mathbf{N}_7)\quad \forall x^X\,\forall y^X\left(\|(-_X x) -_X (-_X y)\| =_{\mathbb{R}} \|x -_X y\|\right)$$
$$(\mathbf{N}_8)\quad \forall x^X\,\forall y^X\,(\|x\| -_{\mathbb{R}} \|y\| \leq_{\mathbb{R}} \|x -_X y\|).$$

Easily one is able to show that $=_X$ is reflexive, symmetric and transitive. Moreover, this unfamiliar set of axioms actually ensures that $+_X, -_X, \cdot_X$ and $\|\cdot\|$ are extensional with respect to $=_X$, i.e.,

$$\begin{aligned} x =_X x' \wedge y =_X y' &\to x +_X y =_X x' +_X y' \\ x =_X x' &\to -_X x =_X -_X x' \\ \alpha =_{\mathbb{R}} \alpha' \wedge x =_X x' &\to \alpha \cdot x =_X \alpha' \cdot_X x' \\ x =_X x' &\to \|x\| =_{\mathbb{R}} \|x'\|, \end{aligned}$$

but note that, since the representation of a real numbers is not unique, $x =_X x'$ does not entail $\|x\| =_1 \|x'\|$, i.e., it does not imply $\forall k^0\,(\|x\|(k) =_0 \|x'\|(k))$. Overall, these axioms succeed in proving that $(X, +_X, -_X, 0_X)$ is a real vector space with a pseudo-norm $\|\cdot\|$ (and becomes a real norm space in the equivalence classes generated by $=_X$).

We shall denote this theory by $\mathrm{PA}^{\omega}_{\preceq}[X, \|\cdot\|]$ and the theory for inner product spaces $\mathrm{PA}^{\omega}_{\preceq}[X, \langle\cdot,\cdot\rangle]$ is obtained with the addition of the axiom for the parallelogram law,

$$(\mathrm{PL})\quad \forall x^X\,\forall y^X\left(\|x +_X y\|^2 +_{\mathbb{R}} \|x -_X y\| =_{\mathbb{R}} 2_{\mathbb{R}} \cdot_{\mathbb{R}} \left(\|x\|^2 +_{\mathbb{R}} \|y\|^2\right)\right),$$

where $(\cdot)^2$ stands for a functional of type $1 \to 1$ representing the squaring function on (the representations of) the real numbers. The inner product functional $\langle\cdot,\cdot\rangle$ of type $X \to (X \to 1)$ is defined by the polarization identity

$$\langle x, y\rangle := \left(\frac{1}{4}\right)_{\mathbb{R}} \cdot_{\mathbb{R}} \left(\|x +_X y\| -_{\mathbb{R}} \|x -_X y\|\right).$$

With the above axioms for normed vector spaces, one can see that the inner product has the usual properties. For the proofs discussed here, we do not require a formal approach to the completeness of the space. Nevertheless, we point out that it is possible to consider complete inner product spaces, i.e., Hilbert spaces, by introducing a functional C of type $(0 \to X) \to X$ which essentially assigns to each Cauchy sequence (x_n) in the pre-Hilbert space a corresponding limit point. Further details on this functional, in particular on how it can be described by a universal formula, can be found in [28, sec. 17.5]. We, moreover, direct the reader to Theorem 2.9, where it is shown that, in the presence of the characteristic principles of the interpretation (i.e., principles under which the original formula and its interpretation become equivalent), one can prove a weak form of Cauchy completeness in X.

[2] That is to say: $\|x\|$ is always a representation of a real number.

We also have axioms characterizing bounded quantifications:

$(\mathrm{B}_\forall)$ $\forall x \preceq_\rho t\, A \leftrightarrow \forall x\,(x \preceq_\rho t \to A)$
$(\mathrm{B}_\exists)$ $\exists x \preceq_\rho t\, A \leftrightarrow \exists x\,(x \preceq_\rho t \wedge A)$

The majorizability relations are an intensional version of Bezem's strong majorizability notion from [1]

(M_1) $\forall x^0\, \forall y^0\, (x \preceq_0 y \leftrightarrow x \leq_0 y)$[3]
(M_2) $\forall x^X\, \forall y^0\, \big(x \preceq_X y \to \|x\| \leq_{\mathbb{R}} (y)_{\mathbb{R}}\big)$
(M_3) $\forall x^{\rho\to\sigma}\quad \forall y^{\widehat{\rho\to\sigma}}(x \preceq_{\rho\to\sigma} y \to \forall u^\rho \quad \forall v^{\hat\rho}\,(u \preceq_\rho v \to xu \preceq_\sigma yv) \wedge \forall v'^{\hat\rho} \quad \big(v \preceq_{\hat\rho} v' \to yv \preceq_{\hat\sigma} yv'\big))$

and the reverse implication to (M_2) and (M_3) are governed by rules of majorizability

$$(\mathrm{RL}_1)\quad \frac{A_{\mathrm{bd}} \to \|p\| \leq_{\mathbb{R}} (n)_{\mathbb{R}}}{A_{\mathrm{bd}} \to p \preceq_X n},$$

$$(\mathrm{RL}_2)\quad \frac{A_{\mathrm{bd}} \wedge u \preceq_\rho v \to tu \preceq_\sigma qv \quad A_{\mathrm{bd}} \wedge v \preceq_{\hat\rho} v' \to qv \preceq_{\hat\sigma} qv'}{A_{\mathrm{bd}} \to t \preceq_{\rho\to\sigma} q},$$

where A_{bd} is a bounded formula, p is a term of type X, n is a term of type 0, t is a term of type $\rho \to \sigma$, q is a term of type $\widehat{\rho \to \sigma}$, u is of type ρ, v, v' are of type $\hat\rho$. The variables u, v, v' don't appear free in the conclusion of (RL_2). In a similar way to the axiom of full extensionality, it is essential that we work with the rules RL_1 and RL_2 as the reverse implications to (M_2) and (M_3) don't have a functional interpretation.

The next lemma follows by induction on the structural complexity of the type.

Lemma 2.2. *For every type $\rho \in \mathsf{T}^X$, the theory $PA^\omega_\preceq[X, \|\cdot\|]$ proves*:

(i) $x \preceq_\rho y \to y \preceq_{\hat\rho} y$,
(ii) $x \preceq_\rho y \wedge y \preceq_{\hat\rho} z \to x \preceq_\rho z$.

For each formula A of the language $\mathcal{L}^{\omega,X}_\preceq$, we associate the formula in which all the intensional predicates $\preceq$ are replaced by their corresponding extensional version $\leq^*$, defined by the equivalences

$$x^0 \leq^*_0 y^0 \leftrightarrow x \leq_0 y,$$
$$x^X \leq^*_X y^0 \leftrightarrow \|x\| \leq_{\mathbb{R}} (y)_{\mathbb{R}},$$
$$x^{\rho\to\sigma} \leq^*_{\rho\to\sigma} y^{\widehat{\rho\to\sigma}} \leftrightarrow \forall u^\rho\, \forall v^{\hat\rho}\, \big(u \leq^*_\rho v \to xu \leq^*_\sigma yv\big) \wedge \forall v'^{\hat\rho}\, \forall v^{\hat\rho}\, \Big(v \leq^*_{\hat\rho} v' \to yv \leq^*_{\hat\sigma} yv'\Big)$$

We denote such formula by A^* and say, following [13], that A^* is the flattening of A. The flattening of $\mathrm{PA}^\omega_\preceq[X, \|\cdot\|]$ is the theory $\mathrm{PA}^\omega_{\leq^*}[X, \|\cdot\|]$ which is an extension (by definition) of PA^ω together with the normed space constants and axioms. As any formula proved with a rule can be proved with the corresponding implication via the modus ponens rule, the following result is clear.

Proposition 2.3 (Flattening). *Let A be an arbitrary formula of $\mathcal{L}^{\omega,X}_\preceq$. We have,*

$$PA^\omega_\preceq[X, \|\cdot\|]\,(+PL) \vdash A \implies PA^\omega_{\leq^*}[X, \|\cdot\|]\,(+PL) \vdash A^*.$$

[3] Here '$x \leq_0 y$' stands as usual for the predicate defined by '$x \dot{-} y =_0 0$', where $\dot{-}$ is the primitive recursive truncated subtraction, reflecting the usual inequality between natural numbers.

The intended interpretation for $\mathrm{PA}^{\omega}_{\leq *}[X, \|\cdot\|]$ is obtained by letting the variables range over a set-theoretical type structure $S^{\omega,X} := \langle S_\rho \rangle_{\rho \in \mathsf{T}^X}$:

$$S_0 := \mathbb{N},\ S_X := X,\ \text{and}\ S_{\rho\to\sigma} := S_\sigma^{S_\rho},$$

where $(X, \|\cdot\|)$ is some (real) normed space. Besides the natural interpretation for the arithmetical constants, objects of type X are taken as vectors in the normed space X. The constant 0_X is interpreted as the null-vector. The constants $+_X$ and $-_X$ are interpreted by the vector addition and vector inversion, respectively. The constant $\cdot_X$ is interpreted as the operation that given a function $\alpha : \mathbb{N} \to \mathbb{N}$ and a vector $x \in X$, outputs the scalar multiplication of x by the (unique) real number represented by the function $\tilde{\alpha}$. The constant $\|\cdot\|$ is interpreted by a function that given any $x \in X$ selects some representation of the real number $\|x\|$. The intended interpretation of $\mathrm{PA}^{\omega}_{\leq *}[X, \langle\cdot,\cdot\rangle]$ is described similarly for X a inner product space.

We define a quantifier-free intensional notion of inequality between (representations of) real numbers, which will be useful for the sequel. First recall that, as mentioned, there exists a quantifier-free formula $A_{\mathrm{qf}}(k^0, x^1, y^1)$ such that

$$x \leq_{\mathbb{R}} y :\equiv \forall k^0\, A_{\mathrm{qf}}(k, x, y),$$

where the specification of A_{qf} will depend on our choice of representation of the real numbers via type 1 functions.

Definition 2.4. *Consider the inequality* $\trianglelefteq_{\mathbb{R}}$ *defined by*

$$x \trianglelefteq_{\mathbb{R}} y \ :\equiv\ p(x,y) \trianglelefteq_1 0^1,$$

where $0^1 := \lambda k\,.\,0^0$ *and*

$$p(x,y)k := \begin{cases} 0^0 & \text{if } A_{qf}(k,x,y) \\ 1^0 & \text{otherwise} \end{cases}.$$

Note that the flattening of the intensional inequality $x \trianglelefteq_{\mathbb{R}} y$ is the usual $x \leq_{\mathbb{R}} y$.

Lemma 2.5 [11]. *The theory* $PA^{\omega}_{\trianglelefteq}[X, \|\cdot\|]$ *proves*:

(i) $x <_{\mathbb{R}} y \to x \trianglelefteq_{\mathbb{R}} y$ *and* $x \trianglelefteq_{\mathbb{R}} y \to x \leq_{\mathbb{R}} y$,
(ii) $\trianglelefteq_{\mathbb{R}}$ *is transitive*,
(iii) $x \trianglelefteq_X y \leftrightarrow \|x\| \trianglelefteq_{\mathbb{R}} (y)_{\mathbb{R}}$.

The usefulness of $\trianglelefteq_{\mathbb{R}}$ resides in allowing us to explicitly describe the complexity present in inequalities between real numbers in a natural manner by replacing them with this intensional version in the appropriate way. Indeed, we have for all x^1, y^1:

$$x <_{\mathbb{R}} y \leftrightarrow \exists k^0 \left(x \trianglelefteq_{\mathbb{R}} y - \frac{1}{k+1} \right)$$
$$x \leq_{\mathbb{R}} y \leftrightarrow \forall k^0 \left(x \trianglelefteq_{\mathbb{R}} y + \frac{1}{k+1} \right)$$
$$x =_{\mathbb{R}} y \leftrightarrow \forall k^0 \left(|x-y| \trianglelefteq_{\mathbb{R}} \frac{1}{k+1} \right)$$

where we dropped most instances of the subscript $\mathbb{R}$ as it is clear by context. After the interpretation of the statement formalized with $\trianglelefteq_{\mathbb{R}}$, we can then without qualms return to the usual inequalities $\leq_{\mathbb{R}}$ using Proposition 2.3 or Lemma 2.5(i).

2.3. The bounded functional interpretation. We now extend the bounded functional interpretation [13] to our context. Since we are in a classical setting, we can restrict our language to $\neg, \vee, \forall$ and universal bounded quantification $\forall x \trianglelefteq t$. The remaining logical symbols are defined in the usual way: $A \wedge B :\equiv \neg(\neg A \vee \neg B)$, $A \to B :\equiv \neg A \vee B$, $\exists x A :\equiv \neg\forall x \neg A$ and $\exists x \trianglelefteq t A :\equiv \neg\forall x \trianglelefteq t \neg A$.

DEFINITION 2.6. *To each formula A of the language $\mathcal{L}^{\omega,X}_{\trianglelefteq}$, we assign formulae A^U and A_U such that A^U is of the form $\tilde{\forall}\underline{b}\ \tilde{\exists}\underline{c}\, A_U(\underline{b},\underline{c})$ and A_U is a bounded formula, according to the following clauses*:

1. *A^U and A_U are simply A, for atomic formulae A.*

Suppose that we have already interpretations for formulae A and B in $\mathcal{L}^{\omega,X}_{\trianglelefteq}$ given, respectively, by $\tilde{\forall}\underline{b}\ \tilde{\exists}\underline{c}\, A_U(\underline{b},\underline{c})$ and $\tilde{\forall}\underline{d}\ \tilde{\exists}\underline{e}\, B_U(\underline{d},\underline{e})$.

2. $(A \vee B)^U :\equiv \tilde{\forall}\underline{b},\underline{d}\ \tilde{\exists}\underline{c},\underline{e}\,\big(A_U(\underline{b},\underline{c}) \vee B_U(\underline{d},\underline{e})\big)$,
3. $(\neg A)^U :\equiv \tilde{\forall}\underline{f}\ \tilde{\exists}\underline{b}\,\Big(\tilde{\exists}\underline{b}' \trianglelefteq \underline{b}\ \neg A_U(\underline{b}', \underline{f}(\underline{b}'))\Big)$,
4. $(\forall x \trianglelefteq t\, A(x))^U :\equiv \tilde{\forall}\underline{b}\ \tilde{\exists}\underline{c}\,\big(\forall x \trianglelefteq t\, A_U(x,\underline{b},\underline{c})\big)$,
5. $(\forall x\, A(x))^U :\equiv \tilde{\forall}\tilde{x},\underline{b}\ \tilde{\exists}\underline{c}\,\big(\forall x \trianglelefteq \tilde{x}\, A_U(x,\underline{b},\underline{c})\big)$.

Above, the notation $\underline{x} \trianglelefteq \underline{y}$ abbreviates $x_1 \trianglelefteq_{\rho_1} y_1, \dots, x_n \trianglelefteq_{\rho_n} y_n$ for $\underline{x} = x_1, \dots, x_n$ and $\underline{y} = y_1, \dots, y_n$ with appropriate types; the notation $\underline{f}(\underline{x})$ abbreviates the tuple of term applications $f_1\underline{x}, \dots, f_n\underline{x}$. Note that if A_{bd} is a bounded formula, then $(A_{\mathrm{bd}})^U$ and $(A_{\mathrm{bd}})_U$ are still A_{bd}, i.e., bounded formulae are left invariant under the interpretation. In the definition of the formula $(\neg A)^U$ we have the apparently innocuous bounded quantification "$\exists\underline{b}' \trianglelefteq \underline{b}$". However, this quantification crucially changes the definition of $(\neg A)_U$ ensuring the following monotonicity property on the matrix $A_U(\underline{b},\underline{c})$:

$$(\underline{c}' \trianglelefteq \underline{c} \wedge A_U(\underline{b},\underline{c}')) \to A_U(\underline{b},\underline{c}).$$

Thus, any bound on a witness for A^U is itself a witness, which amounts to saying that the interpretation operates directly with bounds.

The interpretation has the following characteristic principles:

- Monotone Bounded Choice, $\mathsf{mAC}^X_{\mathsf{bd}}$:

$$(\mathsf{mAC}^X_{\mathsf{bd}}): \qquad \tilde{\forall}\underline{x}\ \tilde{\exists}\underline{y}\, A_{\mathrm{bd}}(\underline{x},\underline{y}) \to \tilde{\exists}\underline{f}\ \tilde{\forall}\underline{x}\ \tilde{\exists}\underline{y} \trianglelefteq \underline{f}(\underline{x})\, A_{\mathrm{bd}}(\underline{x},\underline{y}),$$

  where A_{bd} is a bounded formula of $\mathcal{L}^{\omega,X}_{\trianglelefteq}$.

The axiom of monotone choice expresses the existence of a monotone function that, instead of acting as a choice function, gives a bound on a witnessing element. Notice that all the quantifications are monotone ones.

- Bounded Collection Principle, $\mathsf{BC}^X_{\mathsf{bd}}$:

$$(\mathsf{BC}^X_{\mathsf{bd}}): \qquad \forall\underline{x} \trianglelefteq \underline{a}\ \exists\underline{y}\, A_{\mathrm{bd}}(\underline{x},\underline{y}) \to \tilde{\exists}\underline{b}\ \forall\underline{x} \trianglelefteq \underline{a}\ \exists\underline{y} \trianglelefteq \underline{b}\, A_{\mathrm{bd}}(\underline{x},\underline{y}),$$

  where A_{bd} is a bounded formula of $\mathcal{L}^{\omega,X}_{\trianglelefteq}$.

The bounded collection principle above stands out as a characteristic principle. It states that if for each x there are elements satisfying a bounded property and x is bounded, then we can already 'collect' all those witnesses below a certain bound b. Moreover,

its contrapositive allows for the conclusion of an element x (below some a) such that $\forall y A_{\mathrm{bd}}(x,y)$, from the weaker statement that such x's (below a) only exist 'locally'. We may regard such x as an *ideal* element that works *uniformly* for each b. In [13], where this Shoenfield-like version of the bounded functional interpretation was originally introduced, strong emphasis is placed on the injection of this uniformity aspect into Peano arithmetic and that is the reason in the use of the letter U for the interpretation. Further note that by having $(\mathsf{BC}^X_{\mathsf{bd}})$ stated with tuples entails that (BC^X_{Σ}) also holds, i.e., the bounded collection principle holds for Σ-formulae (even with a bounded matrix).

- Majorizability Axioms, MAJ^X:

$$(\mathsf{MAJ}^X): \qquad \forall x^{\rho}\, \exists y^{\hat{\rho}}\, \left(x \trianglelefteq_{\rho} y\right),$$

  for any $\rho \in \mathsf{T}^X$.

The majorizability axioms state that every element is intensionally majorizable. These axiom schemas characterize the interpretation in the sense of the following result, which is established by induction on the complexity of the formula.

Proposition 2.7 (Characterization). *For any formula A of $\mathcal{L}^{\omega,X}_{\trianglelefteq}$,*

$$PA^{\omega}_{\trianglelefteq}[X, \|\cdot\|] + \mathsf{mAC}^X_{\mathsf{bd}} + \mathsf{BC}^X_{\mathsf{bd}} + \mathsf{MAJ}^X \vdash A \leftrightarrow A^U$$

As pointed out already for Peano arithmetic in [13], the conjunction of our formal system with the characteristic principles of the interpretation is not set-theoretically sound, as for example it refutes the weakest form of extensionality proving the negation of the sentence $\forall \varphi^{1\to 0}\, \forall \alpha^1, \beta^1\, (\alpha =_1 \beta \to \varphi(\alpha) =_0 \varphi(\beta))$. Nevertheless, we have the following soundness theorem, akin to the usual metatheorems of proof mining.

Theorem 2.8 (Soundness). *For an arbitrary formula $A(\underline{z})$ of the language $\mathcal{L}^{\omega,X}_{\trianglelefteq}$ with free-variables $\underline{z}$, let $A(\underline{z})^U$ be $\tilde{\forall}\underline{b}\, \tilde{\exists}\underline{c}\, A_U(\underline{z},\underline{b},\underline{c})$. If*

$$PA^{\omega}_{\trianglelefteq}[X, \|\cdot\|]\,(+PL) + \mathsf{mAC}^X_{\mathsf{bd}} + \mathsf{BC}^X_{\mathsf{bd}} + \mathsf{MAJ}^X \vdash A(\underline{z}),$$

then there are closed monotone terms $\underline{t}$ of T such that

$$PA^{\omega}_{\trianglelefteq}[X, \|\cdot\|]\,(+PL) \vdash \tilde{\forall}\underline{a},\underline{b}\, \forall \underline{z} \trianglelefteq \underline{a}\, A_U(\underline{z},\underline{b},\underline{t}\,\underline{a}\,\underline{b}).$$

Moreover, the terms $\underline{t}$ can be extracted from a proof witnessing the assumption.

The full proof of this soundness result for the normed case was done in Engrácia's PhD thesis [11], and the inclusion of the universal axiom of the parallelogram law impacts no significant change. Since the complete proof is long, we shall only give a brief explanation. As it is usual in these results, the proof is by induction on the length of a derivation for the assumption. In a suitable derivation calculus for classical logic, one discuss the soundness of all the axioms and rules. In these extended systems of arithmetic, one must in particular verify the interpretation of the new axioms and rules. To this, note that all the new axioms are given by universal statements, and thus are trivially interpreted by themselves. Moreover, an essential feature in the proof of soundness for arithmetic in all finite types is that the underlying system is a majorizability theory, i.e., Howard's theorem [23] still holds: any closed term is majorizable. In our case, we require the version with the intensional predicates $\trianglelefteq$: for arithmetic, first it is a simple observation that Howard's result extends to the stronger notion of majorizability due to Bezem [1]; second, it is easy to see that the result

can be established without the reverse implication in (M_2) and using the rule instead. Therefore, the soundness for extended systems requires the verification that all the added new constants are still majorizable. This is true in our case, as shown in [11], and so the soundness theorem holds.

Proofs of soundness theorems are frequently long and their complete discussion is usually allocated to theses or to a book treatment. For this reason, they are sometimes considered 'black-boxes' in proof mining practice. This issue is worsened by the fact that the extraction of quantitative data (the closed terms $\underline{t}$ guaranteed by soundness) is frequently carried out only in a semi-formal way. Nevertheless, just the statement of the soundness theorem (as is done here) is an incomplete description of the proof-theoretical technique employed in a proof mining study. The full machinery of the proof interpretation used is actually contained in the effective way the construction of the terms $\underline{t}$ is done in the verification of the soundness theorem. In particular, the construction highlights that the complexity of the information one obtains is a reflection of the principles required in deriving the (formalized) mathematical theorem being studied.

As mentioned already, we don't include any sort of Cauchy completeness operator in our system. We finish this section with an interesting simple result stating that, in the presence of $\mathsf{BC}^X_{\mathsf{bd}}$, the formal system is nevertheless strong enough to prove a weak version of Cauchy completeness. For n^0 and a tuple $\underline{m}$ of type 0 variables, the notation $\forall \underline{m} \geq_0 n\, A$ abbreviates the following formula $\forall \underline{m}\,((n \leq_0 m_1 \wedge \cdots \wedge n \leq_0 m_r) \to A)$, and in a dual way for the existential quantifier.

THEOREM 2.9. *$PA^{\omega}_{\trianglelefteq}[X, \|\cdot\|] + \mathsf{BC}^X_{\mathsf{bd}}$ proves the convergence of any Cauchy sequence in X with a Cauchy modulus, namely it proves that for any $x^{0\to X}_{(\cdot)}$ and function f^1, if*

$$\forall k^0\, \forall i, j \geq_0 f(k)\left(\|x_i - x_j\| <_{\mathbb{R}} \frac{1}{k+1}\right),$$

then

$$\exists x^X\, \forall k^0\, \forall m \geq_0 f(k)\left(\|x_m - x\| \leq_{\mathbb{R}} \frac{1}{k+1}\right).$$

Proof. From the assumption and Lemma 2.5(i), we have

$$\forall k^0\, \forall i^0, j^0\left((f(k) \leq_0 i \wedge f(k) \leq_0 j) \to \|x_i - x_j\| \trianglelefteq_{\mathbb{R}} \frac{1}{k+1}\right).$$

Let r^0 be arbitrary Then,

$$\forall k \leq_0 r\, \forall m \leq_0 r\left(f(k) \leq_0 m \to \|x_m - x_{f^{\max}(r)}\| \trianglelefteq_{\mathbb{R}} \frac{1}{k+1}\right),$$

since $f(k) \leq f^{\max}(r)$, for $f^{\max}(r) := \max_0\{f(s) : s \leq_0 r\}$.[4] Consider N^0 to be such that $1 + \|x_{f(0)}\| <_{\mathbb{R}} (N)_{\mathbb{R}}$. Then, as $f(0) \leq_0 f^{\max}(r)$, we have

$$\|x_{f^{\max}(r)}\| \leq_{\mathbb{R}} \|x_{f^{\max}(r)} - x_{f(0)}\| + \|x_{f(0)}\| \leq_{\mathbb{R}} 1 + \|x_{f(0)}\| <_{\mathbb{R}} N.$$

[4] The function $\lambda r.(\max_0\{f(s) : s \leq_0 r\})$ is recursively defined in our system.

By Lemma 2.5(i) and (iii), we have $x_{f^{\max}(r)} \trianglelefteq_X N$ for all r^0. Thus,

$$\forall r^0\, \exists x^X \trianglelefteq_X N\, \forall k \leq_0 r\, \forall m \leq_0 r \left(f(k) \leq_0 m \to \|x_m - x\| \trianglelefteq_{\mathbb{R}} \frac{1}{k+1} \right).$$

By (the contrapositive of) $\mathsf{BC}^X_{\mathsf{bd}}$ and Lemma 2.5(i), we conclude

$$\exists x^X \trianglelefteq_X N\, \forall k^0\, \forall m^0 \left(f(k) \leq_0 m \to \|x_m - x\| \leq_{\mathbb{R}} \frac{1}{k+1} \right).$$

□

2.4. A general principle. The elimination of sequential weak compactness relies on the following result (see [15] for the discussion in metric spaces).

THEOREM 2.10. *$PA^{\omega}_{\trianglelefteq}[X, \|\cdot\|] + \mathsf{BC}^X_{\mathsf{bd}}$ proves that for any $T^{X\to X}$, $u^{0\to X}_{(\cdot)}$ and N^0 such that $\forall n^0\, (u_n \trianglelefteq_X N)$ and $\lim \|u_n - T(u_n)\| = 0$, i.e.,*

$$\forall k^0\, \exists n^0\, \forall m \geq_0 n \left(\|u_m - T(u_m)\| <_{\mathbb{R}} \frac{1}{k+1} \right),$$

and for all k^0, λ^1 and $\theta^{X\to 1}$, if

$$\forall y^X \trianglelefteq_X N \left(T(y) =_X y \to \lambda \trianglelefteq_{\mathbb{R}} \theta(y) + \frac{1}{k+1} \right)$$

then

$$\exists n^0\, \forall m \geq_0 n \left(\lambda \trianglelefteq_{\mathbb{R}} \theta(u_m) + \frac{1}{k+1} \right).$$

Proof. Let k^0, λ^1 and $\theta^{X\to 1}$ be arbitrary. From the assumption, we have

$$\forall y \trianglelefteq_X N \left(\forall m^0 \left(\|y - T(y)\| \trianglelefteq_{\mathbb{R}} \frac{1}{m+1} \right) \to \lambda \trianglelefteq_{\mathbb{R}} \theta(y) + \frac{1}{k+1} \right)$$

and hence

$$\forall y \trianglelefteq_X N\, \exists m^0 \left(\|y - T(y)\| \trianglelefteq_{\mathbb{R}} \frac{1}{m+1} \to \lambda \trianglelefteq_{\mathbb{R}} \theta(y) + \frac{1}{k+1} \right).$$

By $\mathsf{BC}^X_{\mathsf{bd}}$, there is M^0 such that

$$\forall y \trianglelefteq_X N\, \exists m \leq_0 M \left(\|y - T(y)\| \trianglelefteq_{\mathbb{R}} \frac{1}{m+1} \to \lambda \trianglelefteq_{\mathbb{R}} \theta(y) + \frac{1}{k+1} \right).$$

Since $m \leq_0 M$ entails $\frac{1}{M+1} \trianglelefteq_{\mathbb{R}} \frac{1}{m+1}$ and $\trianglelefteq_{\mathbb{R}}$ is transitive, we get

$$\forall y \trianglelefteq_X N \left(\|y - T(y)\| \trianglelefteq_{\mathbb{R}} \frac{1}{M+1} \to \lambda \trianglelefteq_{\mathbb{R}} \theta(y) + \frac{1}{k+1} \right).$$

From the assumption that $\lim \|u_n - T(u_n)\| = 0$, there is n^0 such that for $m \geq_0 n$

$$\|u_m - T(u_m)\| \trianglelefteq_{\mathbb{R}} \frac{1}{M+1} \wedge u_m \trianglelefteq_X N.$$

We thus conclude, $\exists n^0\, \forall m^0 \left(n \leq_0 m \to \lambda \trianglelefteq_{\mathbb{R}} \theta(u_m) + \frac{1}{k+1} \right)$, as desired. □

The result above can be made more general. Indeed, one may consider a finite, or even infinite, number of maps. Actually, in a more abstract way, the predicate $T(y) =_X y$ can be replaced by a predicate of the form $\forall n\, \Omega(y,n)$ in the sense of §3, provided that the premise '$\lim \|u_n - T(u_n)\| = 0$' is also replaced with '$\forall k^0\, \exists n^0\, \forall m \geq_0 n\, \Omega(u_m, k)$'. Further, note that, besides $\mathsf{BC}^X_{\mathsf{bd}}$, the argument is of a very simple logical (modus ponens-like) form. In particular, if we weaken the premise to a 'lim inf', i.e., $\forall k^0\ \forall n^0\ \exists m \geq_0 n\, \Omega(u_m, k)$, then we obtain the weaker conclusion $\forall n^0\ \exists m \geq_0 n\, \big(\lambda \trianglelefteq_{\mathbb{R}} \theta(u_m) + \frac{1}{k+1}\big)$.

Theorem 2.10 is usually employed in a more specialized form.

THEOREM 2.11. *$PA^{\omega}_{\trianglelefteq}[X, \|\cdot\|] + \mathsf{BC}^X_{\mathsf{bd}}$ proves that for any $T^{X\to X}$, $u^{0\to X}_{(\cdot)}$, and N^0 such that $\forall n^0\, (u_n \trianglelefteq_X N)$ and $\lim \|u_n - T(u_n)\| = 0$, and for all $\varphi^{X\to(X\to 1)}$, if*

$$\forall k^0\, \exists x \trianglelefteq_X N \left(T(x) =_X x \wedge \forall y \trianglelefteq_X N \left(T(y) =_X y \to \varphi(x,x) <_{\mathbb{R}} \varphi(x,y) + \frac{1}{k+1} \right) \right),$$

then

$$\forall k^0\, \exists x \trianglelefteq_X N \left(T(x) =_X x \wedge \exists n^0\, \forall m \geq_0 n \left(\varphi(x,x) \leq_{\mathbb{R}} \varphi(x,u_m) + \frac{1}{k+1} \right) \right).$$

Proof. Given arbitrary k^0, consider $x^X \trianglelefteq_X N$ such that $T(x) =_X x$ and

$$\forall y \trianglelefteq_X N \left(T(y) =_X y \to \varphi(x,x) <_{\mathbb{R}} \varphi(x,y) + \frac{1}{k+1} \right).$$

Apply Theorem 2.10 with $\lambda := \varphi(x,x)$ and $\theta(y) := \varphi(x,y)$. □

We give a quantitative version of Theorem 2.11. In the following, we use the notation $\overline{B}_N(0)$ for the closed ball in X of radius N centred at 0, and $[n;m]$ for the interval of natural numbers $[n,m] \cap \mathbb{N}$.

THEOREM 2.12. *Let $(X, \|\cdot\|)$ be a normed space and consider $N \in \mathbb{N}$, a map $T : X \to X$, a sequence $(u_n) \subseteq \overline{B}_N(0)$, and a function $\varphi : X \times X \to \mathbb{R}$. Assume that there exist monotone functionals $\alpha, \beta : \mathbb{N} \times \mathbb{N}^{\mathbb{N}} \to \mathbb{N}$ satisfying*

$$\begin{aligned}
&\text{(i)}\ \ \forall k \in \mathbb{N}\, \tilde{\forall} f : \mathbb{N} \to \mathbb{N}\, \exists n \leq \alpha(k,f)\, \forall m \in [n; f(n)]\, \big(\|u_m - T(u_m)\| \leq \tfrac{1}{k+1}\big)\\
&\text{(ii)}\ \ \forall k \in \mathbb{N}\, \tilde{\forall} f : \mathbb{N} \to \mathbb{N}\, \exists n \leq \beta(k,f)\, \exists x \in \overline{B}_N(0)\, \Big(\|x - T(x)\| \leq \tfrac{1}{f(n)+1}\\
&\qquad \wedge \forall y \in \overline{B}_N(0)\, \big(\|y - T(y)\| \leq \tfrac{1}{n+1} \to \varphi(x,x) \leq \varphi(x,y) + \tfrac{1}{k+1}\big)\Big).
\end{aligned}$$

Then,

$$\begin{aligned}
&\forall k \in \mathbb{N}\, \tilde{\forall} f : \mathbb{N} \to \mathbb{N}\, \exists n \leq \psi(k,f)\, \exists x \in \overline{B}_N(0)\\
&\left(\|x - T(x)\| \leq \frac{1}{f(n)+1} \wedge \forall m \in [n; f(n)] \left(\varphi(x,x) \leq \varphi(x,u_m) + \frac{1}{k+1} \right) \right),
\end{aligned}$$

where $\psi(k,f) := \alpha\big(\beta(k, f_\alpha), f\big)$, where $f_\alpha(r) := f(\alpha(r,f))$.

Proof. The argument is essentially as in [15, proposition 4.3]. Let $k \in \mathbb{N}$ and monotone $f : \mathbb{N} \to \mathbb{N}$ be given. By (ii), applied to k and the function f_α (which is monotone as α is monotone), there is $x \in \overline{B}_N(0)$ and $n_0 \leq \beta(k, f_\alpha)$ such that $\|x - T(x)\| \leq \frac{1}{f_\alpha(n_0)+1}$ and

$$\forall y \in \overline{B}_N(0) \left(\|y - T(y)\| \leq \frac{1}{n_0+1} \to \varphi(x,x) \leq \varphi(x,y) + \frac{1}{k+1} \right).$$

Applying (i), we obtain $n_1 \leq \alpha(n_0, f)$ such that

$$\forall m \in [n_1; f(n_1)] \left(\|u_m - T(u_m)\| \leq \frac{1}{n_0 + 1} \right).$$

Since α is monotone, we have $n_1 \leq \alpha(n_0, f) \leq \alpha(\beta(k, f_\alpha), f) =: \psi(k, f)$. By the monotonicity of f, we have $\|x - T(x)\| \leq \frac{1}{f(n_1)+1}$. Since $(u_n) \subseteq \overline{B}_N(0)$, the result follows. □

In several proof mining studies, this argument has underlined the bypass of weak compactness arguments which would otherwise require the interpretation of arithmetical comprehension and thus the use of bar-recursive functionals $\mathsf{BR}_{0,1}$. On each instance, one relies on a specific function φ tailored for the combinatorial argument at hand.

Lastly, we go on a small tangent to discuss some quantitative results regarding switching from lim-statements to the corresponding weaker lim inf-statement. Note that a rate of metastability α for $\lim \|u_n - T(u_n)\| = 0$ (as in Theorem 2.12(i)), entails a so-called lim inf-rate, i.e., a functional $\Phi : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ satisfying

$$\forall k, n \in \mathbb{N}\, \exists m \in [n; \Phi(k, n)] \left(\|u_m - T(u_m)\| \leq \frac{1}{k + 1} \right),$$

given by $\Phi(k, n) := \max\{\alpha(k, \lambda r. \max\{r, n\}), n\}$. In a similar but slightly more convoluted way, we also have the following result.

LEMMA 2.13. *Let $\psi : \mathbb{N} \times \mathbb{N}^{\mathbb{N}} \to \mathbb{N}$ be a monotone functional such that*

$$\forall k \in \mathbb{N}\, \tilde{\forall} f : \mathbb{N} \to \mathbb{N}\, \exists n \leq \psi(k, f)\, \exists x \in \overline{B}_N(0)$$
$$\left(\|x - T(x)\| \leq \frac{1}{f(n) + 1} \wedge \forall m \in [n; f(n)] \left(\varphi(x, x) \leq \varphi(x, u_m) + \frac{1}{k + 1} \right) \right),$$

then

$$\forall k \in \mathbb{N}\, \tilde{\forall} \mathcal{G} : \mathbb{N}^{\mathbb{N}} \to \mathbb{N}\, \tilde{\exists} h \leq_1 \xi(k, \mathcal{G})\, \exists x \in \overline{B}_N(0)$$
$$\left(\|x - T(x)\| \leq \frac{1}{\mathcal{G}(h) + 1} \wedge \forall n \leq \mathcal{G}(h)\, \exists m \in [n; h(n)] \left(\varphi(x, x) \leq \varphi(x, u_m) + \frac{1}{k + 1} \right) \right)$$

with $\xi(k, \mathcal{G}) := \lambda r. \max\{r, \psi(k, f_{\mathcal{G}})\}$ where $f_{\mathcal{G}} := \lambda n. \max\{\mathcal{G}(\lambda r. \max\{r, n\}), n\}$.

Proof. Let $k \in \mathbb{N}$ and a monotone $\mathcal{G}$ be given. As $f_{\mathcal{G}}$ is monotone, by the assumption, there exist $n_0 \leq \psi(k, f_{\mathcal{G}})$ and $\tilde{x} \in \overline{B}_N(0)$ such that

$$(\circ) \quad \|\tilde{x} - T(\tilde{x})\| \leq \frac{1}{f_{\mathcal{G}}(n_0) + 1} \wedge \forall m \in [n_0; f_{\mathcal{G}}(n_0)] \left(\varphi(\tilde{x}, \tilde{x}) \leq \varphi(\tilde{x}, u_m) + \frac{1}{k + 1} \right).$$

Define $h := \lambda r. \max\{r, n_0\}$ which is $\leq_1 \xi(k, \mathcal{G})$. Then, $f_{\mathcal{G}}(n_0) \geq \mathcal{G}(\lambda r. \max\{r, n_0\}) = \mathcal{G}(h)$, and so

$$\|\tilde{x} - T(\tilde{x})\| \leq \frac{1}{\mathcal{G}(h) + 1}.$$

Consider now $n \leq \mathcal{G}(h)$, and take $m = \max\{n, n_0\} = h(n)$. Then, on the one hand, $m \in [n; h(n)]$, and on the other $m \in [n_0, f_{\mathcal{G}}(n_0)]$, since

$$m = \max\{n, n_0\} \leq \max\{\mathcal{G}(h), n_0\} = f_{\mathcal{G}}(n_0).$$

The result follows from the second conjunct of $(\circ)$. □

The next result gives a quantitative version of Theorem 2.11 when $\lim \|u_n - T(u_n)\| = 0$ is replaced with $\liminf \|u_n - T(u_n)\| = 0$.

THEOREM 2.14. *Assume that there exist monotone functionals* $\Phi : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ *and* $\beta : \mathbb{N} \times \mathbb{N}^{\mathbb{N}} \to \mathbb{N}$ *satisfying*

(i) $\forall k, n \in \mathbb{N}\, \exists m \in [n; \Phi(k,n)] \left(\|u_m - T(u_m)\| \leq \frac{1}{k+1}\right)$;

(ii) $\forall k \in \mathbb{N}\, \tilde{\forall} f : \mathbb{N} \to \mathbb{N}\, \exists n \leq \beta(k,f)\, \exists x \in \overline{B}_N(0) \left(\|x - T(x)\| \leq \frac{1}{f(n)+1} \right.$
$\left. \wedge \forall y \in \overline{B}_N(0) \left(\|y - T(y)\| \leq \frac{1}{n+1} \to \varphi(x,x) \leq \varphi(x,y) + \frac{1}{k+1}\right)\right)$.

Then,

$$\forall k \in \mathbb{N}\, \tilde{\forall} \mathcal{G} : \mathbb{N}^{\mathbb{N}} \to \mathbb{N}\, \tilde{\exists} h \leq_1 \xi(k,\mathcal{G})\, \exists x \in \overline{B}_N(0)$$
$$\left(\|x - T(x)\| \leq \frac{1}{\mathcal{G}(h)+1} \wedge \forall n \in \mathbb{N}\, \exists m \in [n; h(n)] \left(\varphi(x,x) \leq \varphi(x,u_m) + \frac{1}{k+1}\right)\right),$$

where $\xi(k,f) := \lambda r.\Phi(\beta(k, f_\Phi), r)$, *where* $f_\Phi := \lambda n.\mathcal{G}(\lambda r.\Phi(n,r))$.

Proof. Let $k \in \mathbb{N}$ and a monotone functional $\mathcal{G}$ be given. Since f_Φ is monotone, by (ii) there exist $n_0 \leq \beta(k, f_\Phi)$ and $\tilde{x} \in \overline{B}_N(0)$ such that

$$(1) \quad \|\tilde{x} - T(\tilde{x})\| \leq \frac{1}{f_\Phi(n_0)+1} = \frac{1}{\mathcal{G}(\lambda r.\Phi(n_0,r))+1},$$

and

$$(2) \quad \forall y \in \overline{B}_N(0) \left(\|y - T(y)\| \leq \frac{1}{n_0+1} \to \varphi(\tilde{x},\tilde{x}) \leq \varphi(\tilde{x},y) + \frac{1}{k+1}\right).$$

By (i) with $k = n_0$, we have $\forall n \in \mathbb{N}\, \exists m \in [n; \Phi(n_0,n)] \left(\|u_m - T(u_m)\| \leq \frac{1}{n_0+1}\right)$. Since $(u_n) \subseteq \overline{B}_N(0)$, by (2) we get

$$\forall n \in \mathbb{N}\, \exists m \in [n; \Phi(n_0,n)] \left(\varphi(\tilde{x},\tilde{x}) \leq \varphi(\tilde{x},u_m) + \frac{1}{k+1}\right).$$

Therefore, the result holds with $h := \lambda r.\Phi(n_0,r)$, for which we have

$$h = \lambda r.\Phi(n_0,r) \leq_1 \lambda r.\Phi(\beta(k,f_\Phi),r) = \xi(k,\mathcal{G}). \qquad \square$$

As a consequence of a rule-type treatment regarding assumption (i), observe that the conclusion of Theorem 2.14 above is stronger, than if we were to merely apply Lemma 2.13 to the conclusion of Theorem 2.12—see the '$\forall n \in \mathbb{N}$' in the former and only the '$\forall n \leq \mathcal{G}(h)$' in the latter.

§3. Interpretation of the ε-weak metric projection. The first proof mining studies on the metric projection are due to Kohlenbach in [29] and [30]. The simple formulation that we discuss here first appeared in [15] and resulted from the study on the application of the bounded functional interpretation to proof mining in the context of the author's doctoral studies [39]. Since then, this formulation has appeared in several proof mining studies by the author, e.g., [6, 7, 9]. Recently, it was also used in [45] in the context of a study extending the quantitative analysis in [32] (see also [7]).

3.1. Interpretation. Let S be a nonempty subset of a normed space $(X, \|\cdot\|)$. We assume that S is described via a sequence of subsets (S_n) by

$$S = \bigcap_{n\in\mathbb{N}} S_n.$$

Without loss of generality, we moreover assume that (S_n) is nonincreasing, i.e., $S_{n+1} \subseteq S_n$ for all $n \in \mathbb{N}$, otherwise we can redefine $S'_n := \bigcap_{n'\leq_n} S_{n'}$. The idea is that the sets S_n approximate S, and that the membership relation $x \in S_n$ is deemed computationally simple. We are interest in studying the projection argument of an arbitrary point $u \in X$ onto the set S. Formally, we describe the approximation sets S_n via some bounded formula $\Omega(x^X, n^0)$ from $\mathcal{L}^{\omega,X}_{\trianglelefteq}$, possible with additional parameters, which similarly, without loss of generality, we can assume to be monotone in the sense that

$$\forall x^X\, \forall n^0\, (\Omega(x,n) \to \forall n' \leq_0 n\, \Omega(x,n'))\,,$$

replacing $\Omega(x,n)$ with $\forall n' \leq_0 n\, \Omega(x,n')$, if that is not the case. We will also use the informal notation '$x \in S_n$' for the predicate $\Omega(x,n)$. In this sense, '$x \in S$' corresponds to the predicate $\forall n^0 \Omega(x,n)$. Note that one indeed have a description of a nonempty subset of X whenever in our formal system,

(1) $\forall x^X, y^X\, (x =_X y \wedge x \in S \to y \in S)$ (extensionality),
(2) $\exists x^X\, (x \in S)$ (nonempty).

For any u^X, the projection of u^X onto S is stated by

$$(-) \quad \exists x^X \left(x \in S \wedge \forall y^X\, (y \in S \to \|x-u\| \leq_{\mathbb{R}} \|y-u\|)\right).$$

Moreover, one may discuss the projection of u onto S by assuming that there is some N^0 such that $x \in S \to x \trianglelefteq_X N$. Indeed, we can expand on the argument in [40]. Consider $p_0 \in S$ and let N^0 be such that $\|p_0 - u\| + \|u\| <_{\mathbb{R}} (N)_{\mathbb{R}}$, then $(-)$ is equivalent to

$$(+) \quad \exists x \trianglelefteq_X N\, (x \in S \wedge \forall y \trianglelefteq_X N\, (y \in S \to \|x-u\| \leq_{\mathbb{R}} \|y-u\|))\,.$$

Clearly, to see that $(-)$ entails $(+)$ it suffices to verify that we have $x \trianglelefteq_X N$, for x^X witnessing $(-)$. We have,

$$\|x\| \leq_{\mathbb{R}} \|x-u\| + \|u\| \leq_{\mathbb{R}} \|p_0 - u\| + \|u\| <_{\mathbb{R}} (N)_{\mathbb{R}},$$

which entails $x \trianglelefteq_X N$ by Lemma 2.5(i) and (iii). For the reverse implication, we just need to argue the second conjunct still holds for any y^X not satisfying $y \trianglelefteq_X N$. By Lemma 2.5, it follows that $\|p_0 - u\| + \|u\| <_{\mathbb{R}} \|y\|$ and, as clearly $p_0 \trianglelefteq_X N$, we have

$$\|x-u\| \leq_{\mathbb{R}} \|p_0 - u\| =_{\mathbb{R}} \|p_0 - u\| + \|u\| - \|u\| <_{\mathbb{R}} \|y\| - \|u\| \leq_{\mathbb{R}} \|y-u\|.$$

Hence, we can replace $x \in S$ by its conjunction with $x \trianglelefteq_X N$. The proof of this projection argument requires countable choice, and therefore escapes the strength of the systems discussed here. From the observation by Kohlenbach in [30], we know that the weaker statement

$$\forall k^0\, \exists x^X \left(x \in S \wedge \forall y^X \left(y \in S \to \|x-u\| <_{\mathbb{R}} \|y-u\| + \frac{1}{k+1}\right)\right),$$

stating only the existence of $\frac{1}{k+1}$-almost projection points of u onto S, which can be proved inductively in our formal system, already suffices for most situations.

Let us discuss the interpretation of

$$(++)\quad \forall k^0\,\exists x \preceq_X N\,\big(\forall m^0\,\Omega(x,m) \wedge \forall y \preceq_X N\,\big(\forall n^0\,\Omega(y,n) \to \Theta(x,y,k)\big)\big),$$

where $\Theta(x,y,k)$ denotes some bounded formula—for the case at hand, we will consider $\Theta(x,y,k) \equiv \|x-u\|^2 \preceq_{\mathbb{R}} \|y-u\|^2 + \frac{1}{k+1}$, using the squared norms for an easier connection with the inner product latter on. By classical logic, $(++)$ is equivalent to

$$\forall k^0\,\exists x \preceq_X N\,\big(\forall m^0\,\Omega(x,m) \wedge \forall y \preceq_X N\,\exists n^0\,\big(\Omega(y,n) \to \Theta(x,y,k)\big)\big)$$

and so, using (the contrapositive of) $\mathsf{BC}^X_{\mathsf{bd}}$, is also equivalent to

$$\forall k^0\,\exists x \preceq_X N\,\big(\forall m^0\,\Omega(x,m) \wedge \exists n^0\,\forall y \preceq_X N\,\exists n' \leq_0 n\,\big(\Omega(y,n') \to \Theta(x,y,k)\big)\big).$$

By the monotonicity assumption on Ω, the statement is equivalent to

$$\forall k^0\,\exists x \preceq_X N\,\big(\forall m^0\,\Omega(x,m) \wedge \exists n^0\,\forall y \preceq_X N\,\big(\Omega(y,n) \to \Theta(x,y,k)\big)\big).$$

Hence, we equivalently have

$$\forall k^0\,\exists n^0\,\exists x \preceq_X N\,\forall m^0\,\big(\Omega(x,m) \wedge \forall y \preceq_X N\,\big(\Omega(y,n) \to \Theta(x,y,k)\big)\big),$$

which, by (the contrapositive of) $\mathsf{BC}^X_{\mathsf{bd}}$ and the monotonicity of Ω, is equivalent to

$$\forall k^0\,\exists n^0\,\forall m^0\,\exists x \preceq_X N\,\big(\Omega(x,m) \wedge \forall y \preceq_X N\,\big(\Omega(y,n) \to \Theta(x,y,k)\big)\big).$$

Now by $\mathsf{mAC}^X_{\mathsf{bd}}$, we have equivalently

$$\forall k^0\,\tilde{\forall} f^1\,\exists n^0\,\forall m \leq_0 f(n)\,\exists x \preceq_X N\,\big(\Omega(x,m) \wedge \forall y \preceq_X N\,\big(\Omega(y,n) \to \Theta(x,y,k)\big)\big),$$

which by the monotonicity of Ω is equivalent to the formula

$$\forall k^0\,\tilde{\forall} f^1\,\exists n^0\,\exists x \preceq_X N\,\big(\Omega(x,f(n)) \wedge \forall y \preceq_X N\,\big(\Omega(y,n) \to \Theta(x,y,k)\big)\big).$$

The steps above naturally also hold with tuples. The soundness theorem therefore ensures that we can compute a primitive recursive bound on n. The next result is a quantitative version of the (ε-weak) projection argument.

PROPOSITION 3.1. *Consider $u \in X$ and $N \geq \|p_0 - u\| + \|u\|$ for some $p_0 \in S$. Then, for all $k \in \mathbb{N}$ and monotone function $f : \mathbb{N} \to \mathbb{N}$,*

$$\exists n \leq f^{(r-1)}(0)\,\exists x \in \overline{B}_N(0)\,\left(x \in S_{f(n)} \wedge \forall y \in \overline{B}_N(0)\left(y \in S_n \to \|x-u\|^2 \leq \|y-u\|^2 + \frac{1}{k+1}\right)\right),$$

where $r := N^2(k+1)$.

Proof. Assuming that the result does not hold, we can define $\{x_0, \dots, x_r\}$ with $x_0 = p_0$ and for all $i < r$, $x_i \in S_{f^{(r-i)}(0)} \cap \overline{B}_N(0)$ and

$$\|x_{i+1} - u\|^2 < \|x_i - u\|^2 - \frac{1}{k+1}.$$

We then obtain the contradiction

$$\|x_r - u\|^2 < \|x_0 - u\|^2 - \frac{r}{k+1} \leq N^2 - \frac{N^2(k+1)}{k+1} = 0. \qquad \square$$

3.2. Variational inequality characterization. In Hilbert spaces (or more generally in CAT(0) spaces), the metric projection of u onto a nonempty set S, say $\mathsf{P}_S(u)$, has a useful characterization via a variational inequality,

$$\forall y \in S \left(\langle \mathsf{P}_S(u) - u, \mathsf{P}_S(u) - y\rangle \leq 0\right).$$

The characterization above relies essentially on two facts:

(1) the set S is convex,

(2) $\forall x, y \in S \left(\forall t \in [0,1] \left(\|x-u\| \leq \|w_t(x,y) - u\|\right) \to \langle x-u, x-y\rangle \leq 0\right)$,

where $w_t(x,y)$ stands for $(1-t)x + ty$. Let us discuss the interpretation of (1) and (2).

The convexity property of a set $S \subseteq X$ corresponds to the following statement

$$(\dagger) \quad \forall x^X \, \forall y^X \left(x, y \in S \to \forall t \in [0,1] \left(w_t(x,y) \in S\right)\right),$$

which (viewing '$\forall t \in [0,1]$' as a bounded quantification) in the presence of MAJ^X and $\mathsf{BC}^X_{\mathsf{bd}}$ is equivalent to

$$(\ddagger) \quad \forall n^0 \, \forall N^0 \, \exists m^0 \, \forall x \trianglelefteq_X N \, \forall y \trianglelefteq_X N \left(x, y \in S_m \to \forall t \in [0,1] \left(w_t(x,y) \in S_n\right)\right).$$

When a proof of the convexity of S formalizes in (some suitable extension of) the system described for normed spaces, the soundness theorem will entail the existence of a bound for m given by a monotone closed term on inputs n and N and hence, by the monotonicity assumption on the formula Ω (i.e., on the sets S_m), a precise witness. Naturally, the existence of such proof amenable for bound extraction is centered on the specifics of the set S.

If we want to discuss a mathematical proof in which a set S (universally described via some formula Ω as above) is assumed to be convex, then we can't simply add the statement (†) to our system, since its interpretation (‡) asks for computational information that we a priori do not have. We can however study such a proof modulo the missing information regarding the convexity of the set. Namely, we can assume that the set S satisfies

$$(\star) \quad \forall n \in \mathbb{N} \, \forall N \in \mathbb{N} \, \exists m \in \mathbb{N} \, \forall x, y \in \overline{B}_N(0) \left(x, y \in S_m \to \forall t \in [0,1] \left(w_t(x,y) \in S_n\right)\right),$$

and provide the underlying formal system with a monotone constant $\gamma^{0\to(0\to 0)}$, i.e.,

$$\forall n_1^0, n_2^0 \, \forall N_1^0, N_2^0 \left((n_1 \leq_0 n_2 \wedge N_1 \leq_0 N_2) \to \gamma_{N_1}(n_1) \leq_0 \gamma_{N_2}(n_2)\right),$$

governed by the axiom

$$\forall n^0 \, \forall N^0 \, \forall x \trianglelefteq_X N \, \forall y \trianglelefteq_X N \left(x, y \in S_{\gamma_N(n)} \to \forall t \in [0,1] w_t(x,y) \in S_n\right).$$

The condition ($\star$) is an uniformization of the convexity property and, in general, more restrictive than just asking for the set to be convex—note however that the two notions coincide in the finite dimensional case. This is a consequence of the uniformity injected by the use of the $\mathsf{BC}^X_{\mathsf{bd}}$ principle (see [13]).

In frequent applications, S is the fixed point set of some map T, or more generally the set of common fixed points of a (even countably in)finite number of maps, and

in that case a function γ is indeed available. The next result, essentially due to [30], corresponds to a quantitative version of fact (1), in the case where

$$S := \bigcap_{n\in\mathbb{N}} \left\{ x \in X : \|x - T(x)\| \leq \frac{1}{n+1} \right\}$$

is the set of fixed points of a nonexpansive map T, which will suffice for our discussion.

LEMMA 3.2. *For all $n, N \in \mathbb{N}$ and $x, y \in \overline{B}_N(0)$, if*

$$\max\{\|x - T(x)\|, \|y - T(y)\|\} \leq \frac{1}{12(2N+1)(n+1)^2},^{5}$$

then $\forall t \in [0,1] \left(\|w_t(x,y) - T(w_t(x,y))\| \leq \frac{1}{n+1} \right)$.

We can discuss the proof of fact (2) in the setting of $\mathrm{PA}^{\omega}_{\lhd}[X, \langle\cdot,\cdot\rangle]$ as it only depends on simple properties of the norm and the inner-product. We have the following finitary version.

LEMMA 3.3. *For all $k, N \in \mathbb{N}$, $u \in X$ and $x, y \in \overline{B}_N(0)$, if*

$$\forall t \in [0,1] \left(\|x - u\|^2 \leq \|w_t(x,y) - u\|^2 + \frac{1}{4N^2(k+1)^2} \right),$$

then $\langle x - u, x - y\rangle \leq \frac{1}{k+1}$.

Proof. Considering the assumption, we have for all $t \in [0,1]$,

$$\begin{aligned}
\|w_t(x,y) - u\|^2 &= \langle w_t(x,y) - u, w_t(x,y) - u\rangle \\
&= \|x - u\|^2 - 2t\langle x - u, x - y\rangle + t^2\|x - y\|^2 \\
&\leq \|w_t(x,y) - u\|^2 + \frac{1}{4N^2(k+1)^2} - 2t\langle x - u, x - y\rangle + 4N^2t^2.
\end{aligned}$$

Hence, for all $t \in (0,1]$

$$\langle x - u, x - y\rangle \leq \frac{1}{8N^2(k+1)^2 t} + 2N^2 t,$$

and, in particular, for $t = \frac{1}{4N^2(k+1)}$ we conclude

$$\langle x - u, x - y\rangle \leq \frac{4N^2(k+1)}{8N^2(k+1)^2} + \frac{2N^2}{4N^2(k+1)} = \frac{1}{k+1}. \qquad \square$$

By combining Lemma 3.3 and Proposition 3.1, in the case where S is convex in the strong sense of $(\star)$, and we have a monotone function $\gamma : \mathbb{N}^2 \to \mathbb{N}$ providing

[5] That is, $\gamma_N(n) := 12(2N+1)(n+1)^2 - 1$, for all $n, N \in \mathbb{N}$. The same function holds also for when S is the set of common fixed points of a finite or an infinite number of nonexpansive maps.

computational information on its convexity, we obtain the following result corresponding to the quantitative analysis of the ε-weak version of the variational inequality characterization.

PROPOSITION 3.4. *Let $u \in X$ and consider a sequence (S_n) such that $S_{n+1} \subseteq S_n$, for all $n \in \mathbb{N}$. Assume that $S := \bigcap_{n\in\mathbb{N}} S_n \neq \emptyset$ is convex and there exists a monotone function $\gamma : \mathbb{N}^2 \to \mathbb{N}$ satisfying*

$$\forall n, N \in \mathbb{N}\, \forall x, y \in \overline{B}_N(0)\, \big(x, y \in S_{\gamma_N(n)} \to \forall t \in [0,1]\, \big(w_t(x,y) \in S_n\big)\big).$$

Consider $N \in \mathbb{N}$ such that $N \geq \|p_0 - u\| + \|u\|$, for some $p_0 \in S$. Then, for all $k \in \mathbb{N}$ and monotone function $f : \mathbb{N} \to \mathbb{N}$,

$$\exists n \leq \gamma_N\left(h^{(R-1)}_{\gamma,f}(0)\right)\, \exists x \in \overline{B}_N(0)\, \left(x \in S_{f(n)} \wedge \forall y \in \overline{B}_N(0)\left(y \in S_n \to \langle x-u, x-y\rangle \leq \frac{1}{k+1}\right)\right),$$

where $R := 4N^4(k+1)^2$ and $h_{\gamma,f}(m) := \max\{f(\gamma(m)), \gamma(m)\}$.

Proof. Apply Proposition 3.1 with the natural number $4N^2(k+1)^2 - 1$, and the monotone function $h_{\gamma,f}$, to conclude the existence of $n_0 \leq h^{(R-1)}_{\gamma,f}(0)$ and $x \in \overline{B}_N(0)$ such that $x \in S_{h_{\gamma,f}(n_0)}$ and for all $y \in \overline{B}_N(0)$,

$$y \in S_{n_0} \to \|x-u\|^2 \leq \|y-u\| + \frac{1}{4N^2(k+1)^2}.$$

Take $n := \gamma_N(n_0)$—which is bounded by $\gamma_N\left(h^{(R-1)}_{\gamma,f}(0)\right)$, since γ is monotone. Since $h_{\gamma,f}(n_0) \geq f(n)$ and by the monotonicity assumption on (S_n), we have $x \in S_{f(n)}$, as desired. For the second conjunct, take $y \in \overline{B}_N(0)$ such that $y \in S_n$, i.e., $y \in S_{\gamma_N(n_0)}$. Note that, since $h_{\gamma,f}(n_0) \geq n$ and using again the monotonicity of (S_n), we also have $x \in S_{\gamma_N(n_0)}$. Hence, the assumption on the function γ entails

$$\forall t \in [0,1]\, \big(w_t(x,y) \in S_{n_0}\big).$$

Since $w_t(x,y) \in \overline{B}_N(0)$, we conclude

$$\forall t \in [0,1]\left(\|x-u\|^2 \leq \|w_t(x,y)-u\|^2 + \frac{1}{4N^2(k+1)^2}\right).$$

The result now follows from Lemma 3.3. □

We can now instantiate the function γ with the output from Lemma 3.2 to immediately obtain the following two results. In a Hilbert space X, consider $(T_n)_{n\in\mathbb{N}}$ a countable family of nonexpansive maps and let $S := \bigcap \mathrm{Fix}(T_n)$ be the set of common fixed points which we assume to be nonempty. Then,

$$S := \bigcap_{n\in\mathbb{N}} S_n, \quad \text{where } S_n := \left\{x \in X : \forall n' \leq n\left(\|x - T_{n'}(x)\| \leq \frac{1}{n+1}\right)\right\}.$$

COROLLARY 3.5. *Let $u \in X$ and consider $N \in \mathbb{N}$ such that $N \geq \|p_0 - u\| + \|u\|$, for some $p_0 \in S$. Then, for all $k \in \mathbb{N}$ and monotone function $f : \mathbb{N} \to \mathbb{N}$,*

$$\exists n \leq 12(2N+1)\left(h_f^{(R-1)}(0)+1\right)^2 \ \exists x \in \overline{B}_N(0)$$
$$\left(\forall n' \leq f(n)\left(\|x - T_{n'}(x)\| \leq \frac{1}{f(n)+1}\right)\right.$$
$$\left.\wedge \forall y \in \overline{B}_N(0)\left(\forall n' \leq n\left(\|y - T_{n'}(y)\| \leq \frac{1}{n+1}\right) \to \langle x-u, x-y\rangle \leq \frac{1}{k+1}\right)\right),$$

where $R := 4N^4(k+1)^2$ *and* $h_f(m) := \max\{f(12(2N+1)(m+1)^2), 12(2N+1)(m+1)^2\}$.

The argument also holds for a finite number of nonexpansive maps, $T_1, \ldots, T_m$, which is the version needed in the sequel, cf. [41, proposition 2.5],[6] by considering instead

$$S_n := \left\{x \in X : \bigwedge_{j=1}^m \left(\|x - T_j(x)\| \leq \frac{1}{n+1}\right)\right\}.$$

COROLLARY 3.6. *Let $u \in X$ and consider $N \in \mathbb{N}$ such that $N \geq \|p_0 - u\| + \|u\|$, for some $p_0 \in \bigcap_{j=1}^m Fix(T_j)$. Then, for all $k \in \mathbb{N}$ and monotone function $f : \mathbb{N} \to \mathbb{N}$,*

$$\exists n \leq 12(2N+1)\left(h_f^{(R-1)}(0)+1\right)^2 \ \exists x \in \overline{B}_N(0)$$
$$\left(\bigwedge_{j=1}^m \left(\|x - T_j(x)\| \leq \frac{1}{f(n)+1}\right)\right.$$
$$\left.\wedge \forall y \in \overline{B}_N(0)\left(\bigwedge_{j=1}^m \left(\|y - T_j(y)\| \leq \frac{1}{n+1}\right) \to \langle x-u, x-y\rangle \leq \frac{1}{k+1}\right)\right),$$

where R and h_f are as before.

§4. Previous applications. To make it clear why we identify Dykstra's algorithm as '*curious*', we recall here how the sequential weak compactness argument was eliminated in the case of Wittmann's convergence proof as an application of the main result from [15], enabling an easier comparison with the study of Dykstra's algorithm discussed in §5.

4.1. The proof of Wittmann's theorem. In 1967, Halpern [21] introduced the following iterative procedure to approximate fixed points of a given nonexpansive map $T : C \to C$, with C a nonempty, bounded, closed and convex subset of a Hilbert space

$$x_0 \in C, \quad x_{n+1} := (1-\alpha_n)T(x_n) + \alpha_n u, \tag{H}$$

[6] Besides the ε-δ formulation, [41, proposition 2.5] also includes some small optimizations related to the majorizability of the initial data irrelevant for our discussion.

where u is some point in the set C, anchoring the construction of the iteration via a convex combination using the parameter sequence $(\alpha_n) \subseteq [0,1]$. Under some conditions on the sequence (α_n), [21] shows the strong convergence of (H) to a fixed point of T. However, Halpern's conditions prevented the canonical choice $\alpha_n = \frac{1}{n+1}$, which was overcome in 1992 by Wittmann in a celebrated result (considered an important nonlinear generalization of von Neumann Mean Ergodic theorem).

THEOREM 4.1 (Wittmann [49]). *If* $(\alpha_n) \subseteq [0,1]$ *satisfies*

$$\text{(i) } \lim \alpha_n = 0, \quad \text{(ii) } \sum \alpha_n = \infty, \quad \text{(iii) } \sum |\alpha_{n+1} - \alpha_n| < \infty$$

then (H) *converges strongly to a fixed point of* T*, the closest to the anchor point* u*.*

Wittmann's argument has the following structure:

(1) (x_n) is bounded;
(2) (x_n) is asymptotically regular, i.e., $\lim \|x_n - x_{n+1}\| = 0$;
(3) (x_n) is T-asymptotically regular, i.e., $\lim \|x_n - T(x_n)\| = 0$;
(4) Since $\mathrm{Fix}(T)$ is a convex nonempty subset, let P be the projection of u onto $\mathrm{Fix}(T)$, that is $P := \mathsf{P}_{\mathrm{Fix}(T)}(u)$;
(5) We have the following inequality,

$$\forall n \in \mathbb{N} \left(\|x_{n+1} - P\|^2 \leq (1-\alpha_n)\|x_n - P\|^2 + \alpha_n B_n \right),$$

where $B_n := 2(1-\alpha_n)\langle u - P, T(x_n) - P\rangle + \alpha_n b$ for some real number $b \in \mathbb{R}$;
(6) By sequential weak compactness, there is some $Q \in C$ and a subsequence (x_{ρ_n}) such that $x_{\rho_n} \rightharpoonup Q$;
(7) If necessary passing to a further subsequence, we have w.l.o.g.

$$x_{\rho_n} \rightharpoonup Q \text{ and simultaneously } \limsup B_n = \lim B_{\rho_n};$$

(8) By the demiclosedness principle (see [4]), we have $Q \in \mathrm{Fix}(T)$;
(9) From the variational inequality characterization of the metric projection, we have,

$$\limsup B_n = \lim 2\langle u - P, T(x_{\rho_n}) - P\rangle = 2\langle u - P, T(Q) - P\rangle = 2\langle u - P, Q - P\rangle \leq 0;$$

(10) Using a combinatorial argument, later distilled into the following technical lemma about sequences of real numbers by Xu [50].

LEMMA 4.2. *Let* $(\alpha_n) \subseteq [0,1]$ *and* $(r_n) \subseteq \mathbb{R}$ *be real sequences such that*

$$\sum \alpha_n = \infty \text{ and } \limsup r_n \leq 0.$$

If (s_n) *is a sequence of non-negative real numbers satisfying*

$$\forall n \in \mathbb{N} \left(s_{n+1} \leq (1-\alpha_n)s_n + \alpha_n r_n \right),$$

then $\lim s_n = 0$.

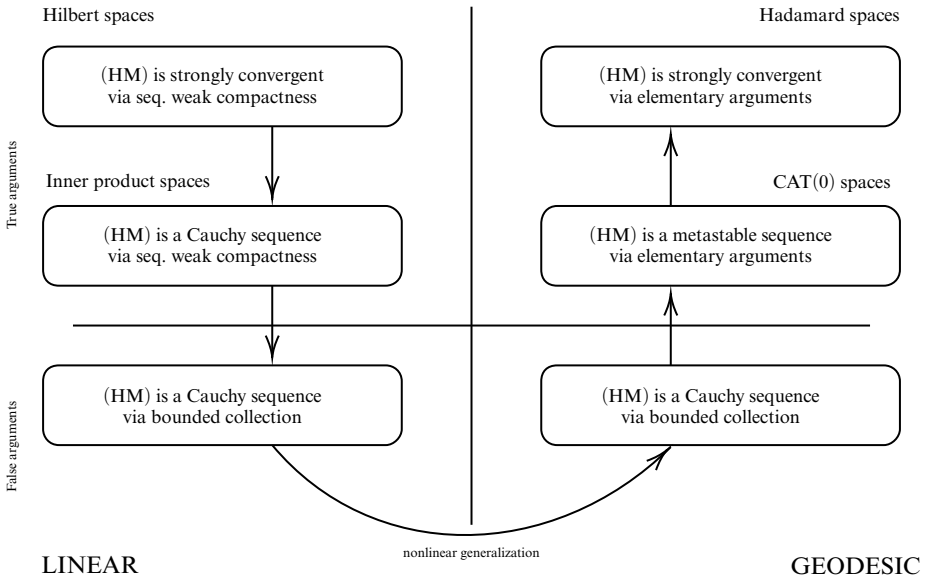
it follows that $\lim \|x_n - P\|^2 = 0$.

Hence, (x_n) converges in norm to $\mathsf{P}_{\mathrm{Fix}(T)}(u)$.

As discussed in §2.1, from a proof-theoretical point of view, the only troublesome arguments are in the use of the projection in step (4), and in the use of a sequential weak compactness argument in step (6), which require arithmetical comprehension. The finitary study in [30] showed that one can bypass (4) by using instead the simpler ε-weakening of the projection argument. Regarding step (6), it uncovered that the role of the weak limit Q could already be replaced by some term of the iteration and thus the quantitative version did not require the interpretation of the full weak compactness argument. This entails in particular that no bar-recursive functionals were needed in the construction of the metastability rate. We remark that the resulting simpler proof was not an automatic output of the proof interpretation: while employing the functional interpretation highlighted the potential for simplification, it ultimately relied on a discernment to recognize the opportunity for a more straightforward argument.

In [15], this was understood under the perspective that Wittmann's argument could be modified into a proof formalized in a system (an extension of $\mathrm{PA}^{\omega}_{\lhd}[X, \langle\cdot,\cdot\rangle]$ with some ad hoc constants tailored to the discussion of Wittmann's result) with bound extraction and where the use of weak compactness was sidestepped by an application of $\mathsf{BC}^{X}_{\mathsf{bd}}$. Concretely, weak compactness is replaced by Theorem 2.11 with the function $\varphi(x, y) := \langle x - u, T(y)\rangle$, as explained in [15, sec. 5.2]. It is this modified proof that can be analysed knowing a priori that no bar-recursive functionals will feature in the extracted quantitative information.

4.2. Further proofs with false principles. In [15] the removal of sequential weak compactness was applied to three case studies: Browder's fixed point theorem, Wittmann's convergence theorem as well as to its generalization due to Bauschke. Since then the usefulness of this approach has been substantiated with several further applications in the literature, e.g., [5–7, 9, 40]. Recently [8], in the general nonlinear setting of CAT(0) spaces, this quantitative perspective allowed for a convergence proof of a new iterative method: the alternating Halpern–Mann iteration, (HM). As explained, in Hilbert spaces, we can recursively approximate fixed points of nonexpansive maps via Halpern's procedure. Extensions to a nonlinear setting usually reduce the convergence proof to that of a Browder-like iteration and require the use of Banach limits—see the discussion by Kohlenbach and Leuştean in [33, 34] regarding a possible tame proof-theoretical treatment of the latter. The method introduced in [8] is more general than (H) and strongly approximates a common fixed point of two nonexpansive maps. The convergence proof crucially relies on a finitary formulation, and does not reduce to a Browder-like iteration nor it requires the use of Banach limits. Instead it follows from the observation that the argument bypassing sequential compactness, using the set-theoretically false principle $\mathsf{BC}^{X}_{\mathsf{bd}}$, further holds in a geodesic setting (while it is not even clear how a sequential weak compactness argument would go through in a nonlinear setting) and thus it was possible to establish the metastable property of the iteration. A posteriori one derives the full infinitary convergence result of this general method in the broad geodesic setting of Hadamard spaces. This example further strengthens the usefulness of the metastability formulation and of proof-theoretical base arguments in applications to standard mathematics.



Yet, a concern lingers in the applicability of this approach to the removal of weak compactness arguments: all applications share a common theme, namely that the recursive method is some variant of the Halpern iteration and that the proof is always structurally similar to Wittmann's argument. The study of Dykstra's algorithm breaks away from such paradigm.

§5. The study of Dykstra's algorithm.

5.1. The proof of Dykstra's convergence. Here we discuss the proof of Theorem 1.1 regarding the strong convergence of Dykstra's algorithm. We denote $C := \bigcap_{j=1}^{m} C_j$ and assume it to be nonempty, as entailed by the 'feasibility' requirement. As before, the list of closed convex sets $\{C_1, \dots, C_m\}$ is extended cyclically to a countable family $(C_n)_{n\geq 1}$ and, for each $n \geq 1$, P_n denotes the projection onto C_n. For $n \in \mathbb{N}$ and $z \in X$, we write

$$s(n, z) := 2 \sum_{k=n-m+1}^{n} \langle x_k - z, q_k \rangle,$$

where $x_{-(m-1)}, \dots, x_{-1}$ are arbitrary elements of X.

The convergence proof of Dykstra's algorithm has the following structure:

(1) Inductively, one proves that for all $n \in \mathbb{N}$,

$$\sum_{k=n-m+1}^{n} q_k = x_0 - x_n.$$

(2) By the characterization of the projections via the variational inequality and the definition of q_k, we immediately see that

$$\forall z \in X \; \forall n \in \mathbb{N} \left(z \in C \to s(n, z) \geq 0 \right).$$

(3) The main combinatorial part of the proof essentially establishes:

$$\forall z \in X\ \forall n \in \mathbb{N}\ \forall i \geq n \left(\|x_i - z\|^2 \leq \|x_n - z\|^2 + s(n,z) - s(i,z)\right).$$

(4) From (3), it follows that (x_n) is bounded, and we take $B \geq 0$ such that for all $n \in \mathbb{N}$, it holds $\|x_0 - x_n\| \leq B$.
(5) At the same time, it is established that (x_n) is asymptotically regular, i.e., that $\lim \|x_n - x_{n+1}\| = 0$.
(6) Underlying the remainder of the proof, one can infer the following implicit lemma:

LEMMA 5.1. *For all $k \in \mathbb{N}$, there exists $r \in \mathbb{N}$ such that*

$$\forall z \in C\ \forall n \in \mathbb{N} \left(\max\{\|x_n - z\|, s(n,x_n)\} \leq \frac{1}{r+1} \to \forall i \geq n \left(\|x_i - z\| \leq \frac{1}{k+1}\right)\right).$$

Proof. We have,

$$\begin{aligned} s(n,z) &= 2 \sum_{k=n-m+1}^{n} \langle x_k - x_n, q_k \rangle + 2 \sum_{k=n-m+1}^{n} \langle x_n - z, q_k \rangle \\ &\leq s(n,x_n) + 2\|x_n - z\| \cdot \|x_0 - x_n\| \\ &\leq \frac{1+2B}{r+1}. \end{aligned}$$

The result follows using (2) and (3). □

Therefore, in order to conclude the proof, it suffices to find a pair $(z,n) \in C \times \mathbb{N}$ satisfying the premise of Lemma 5.1. Indeed, that would entail that (x_n) is a Cauchy sequence and hence, by completeness of the space, a convergent sequence. We moreover remark at this point that the argument does not work if the point z is taken simply as some x_{n_0} for some large n_0, and requires a more delicate approach. Let us continue.

(7) Making use of a technical lemma regarding square-summable sequences (which goes back to the original work of Dykstra [10]), one establishes

$$\liminf \sum_{k=n-m+1}^{n} |\langle x_k - x_n, q_k \rangle| = 0.$$

(8) Consider a subsequence (x_{ρ_n}) and $z \in X$ such that
  (i) $\lim \sum_{k=\rho_n - m+1}^{\rho_n} |\langle x_k - x_{\rho_n}, q_k \rangle| = 0$ (which entails that $\lim s(\rho_n, x_{\rho_n}) = 0$),
  (ii) $x_{\rho_n} \rightharpoonup z$,
  (iii) $\exists j \in [1;m]\ \forall n \in \mathbb{N}\,(j_{\rho_n} = j)$, entailing that $(x_{\rho_n}) \subseteq C_j$,
  (iv) $(\|x_{\rho_n}\|)$ converges.
(9) Since $\lim \|x_n - x_{n+1}\| = 0$, and the sets C_j are weakly closed, we get that $z \in C$.
(10) Lastly, one shows that $\|x_{\rho_n}\| \to \|z\|$ which, together with (8)(*ii*), entails that $x_{\rho_n} \to z$. Moreover, in this step it is also established that $z = \mathrm{P}_C(x_0)$.

At this point, the proof is concluded since

$$x_{\rho_n} \to z \in C \quad \text{and} \quad s(\rho_n, x_{\rho_n}) \to 0,$$

provides us with a pair (z, n) is the required conditions, for each given $k \in \mathbb{N}$. Note that since, for each k, the point z considered is always $\mathrm{P}_C(x_0)$, one can further conclude that the limit of the sequence is indeed the optimal solution to the convex feasibility problem.

The steps (1)–(7) are proof-theoretically very simple, making use of combinatorial and inductive arguments only. It is at step (8) that a priori a primitive recursive (in the sense of Gödel) bound becomes out of reach, due to the use of sequential weak compactness (ii), of the infinite pigeonhole principle (iii), and of Bolzano–Weirstrass compactness (iv). In section 5.4, we show that via a modified proof, a pair (z, n) in the required conditions can be obtained using the principle $\mathsf{BC}^X_{\mathsf{bd}}$ and sidestepping the compactness arguments in the final steps of the proof.

It is clear that the convergence proof described above is significantly different than Wittmann's general approach. On the one hand, it doesn't require verifying the asymptotic regularity of the sequence with regards to the nonexpansive maps involved in the iteration, i.e., it does not prove P_j-asymptotic regularity for all $j \in [1; m]$. It also does not employ Xu's technical lemma (Lemma 4.2) central to all Wittmann-like proofs. Moreover, the use of compactness arguments in step (8) goes beyond sequential weak compactness and is much more convoluted. For these reasons, it was not expected that bar-recursive functionals could be absent from a proof-theoretical study of Dykstra's convergence proof.

5.2. A tailored formal system. Here we will extend the system $\mathrm{PA}^{\omega}_{\trianglelefteq}[X, \langle\cdot,\cdot\rangle]$ with some ad hoc constants tailored to the treatment of the convergence proof of Dykstra's algorithm and later show that, leveraging on the presence of the $\mathsf{BC}^X_{\mathsf{bd}}$ principle, we can prove that the algorithm is a Cauchy sequence bypassing the use of the compactness principles needed in the original proof. Naturally, one can shift all the indexes so as to start from zero. Namely, we formally consider the translated sequences $(\tilde{C}_n)$, $(\tilde{P}_n)$, $(\tilde{q}_n)$ and $(\tilde{x}_n)$ according to

$$\tilde{C}_n := C_{n+1}, \quad \tilde{P}_n := P_{n+1}, \quad \tilde{q}_n := q_{n-(m-1)}$$
$$\tilde{x}_n := \hat{x}_{n-(m-1)}, \text{ with } \hat{x}_n := \begin{cases} x_0 & \text{if } n \in [-(m-1); 0] \\ x_n & \text{otherwise} \end{cases},$$

where in $(\hat{x}_n)$ we just set all the terms $x_{-(m-1)}, \dots, x_{-1}$, originally taken arbitrarily from X, equal to the initial point x_0. In this way, we can rewrite Dykstra's algorithm as

$$\begin{cases} \tilde{x}_0 = \cdots = \tilde{x}_{m-1} \in X \\ \tilde{q}_0 = \cdots = \tilde{q}_{m-1} = 0 \end{cases} \quad \text{and } \forall n \in \mathbb{N} \begin{cases} \tilde{x}_{n+m} = \tilde{\mathrm{P}}_n(\tilde{x}_{n+m-1} + \tilde{q}_n) \\ \tilde{q}_{n+m} = \tilde{x}_{n+m-1} + \tilde{q}_n - \tilde{x}_{n+m} \end{cases}.$$

It is with this observation that we discuss the formal aspects of the convergence proof, while at the same time we dispense with writing the tildes.

To the language $\mathcal{L}^{\omega,X}_{\trianglelefteq}$ we add the following constants

- m and b of type 0
- x_0 and p of type X

- $x_{(\cdot)}$ and $q_{(\cdot)}$ of type $0 \to X$[7]
- $\chi_{(\cdot)}$ of type $0 \to (X \to 0)$
- $\mathsf{P}_{(\cdot)}$ of type $0 \to (X \to X)$,

where m will stand for the number of convex sets, b will provide the necessary information regarding the majorizability requirement of the novel constants, $\chi_{(\cdot)}$ will stand for characteristic functions for the convex sets, i.e., informally

$$\chi_j(x) =_0 1 \rightsquigarrow x \in \tilde{C}_j = C_{j+1},$$

and the $\mathsf{P}_{(\cdot)}$ for the projection maps necessary for the definition of the algorithm. We extend the system $\mathrm{PA}^{\omega}_{\trianglelefteq}[X, \langle\cdot,\cdot\rangle]$ to this language and add the following (computationally complete) axioms

(D1) $2 \leq_0 m$
(D2) $\|x_0\|, \|p\| \trianglelefteq_{\mathbb{R}} (b)_{\mathbb{R}}$[8]
(D3) $\forall j^0\, \forall x^X \left(\chi_j(x) \leq_0 1\right)$
(D4) $\forall j^0\, \forall x^X\, \forall y^X \left((\chi_j(x) =_0 1 \wedge \chi_j(y) =_0 1) \to \forall t \in [0,1](\chi_j(w_t(x,y)) =_0 1)\right)$
(D5) $\forall j^0\, \forall x^X \left(\chi_j(x) =_0 \chi_{j+m}(x)\right)$
(D6) $\forall j^0 \left(\chi_j(p) =_0 1\right)$
(D7) $\forall j^0\, \forall x^X \left(\chi_j(\mathsf{P}_j(x)) =_0 1\right)$
(D8) $\forall j^0\, \forall x^X\, \forall y^X \left(\chi_j(y) =_0 1 \to \|x - \mathsf{P}_j(x)\| \trianglelefteq_{\mathbb{R}} \|x - y\|\right)$

and regarding Dykstra's algorithm

(D9) $\forall n^0 \left(n \leq_0 m - 1 \to (x_n =_X x_0 \wedge q_n =_X 0_X)\right)$
(D10) $\forall n^0 \left(x_{n+m} =_X \mathsf{P}_n(x_{n+m-1} + q_n)\right)$
(D11) $\forall n^0 \left(q_{n+m} =_X x_{n+m-1} + q_n - x_{n+m}\right)$.

We denote the resulting system by $\mathcal{D}^{\omega}_{\trianglelefteq}$. From (D4), (D7) and (D8), using the arguments in Lemma 3.3, one shows that

$$\mathcal{D}^{\omega}_{\trianglelefteq} \vdash \forall j^0\, \forall x^X\, \forall y^X \left(\chi_j(y) =_0 1 \to \langle P_j(x) - x, P_j(x) - y\rangle \leq_{\mathbb{R}} 0_{\mathbb{R}}\right),$$

which, moreover, entails that the term $P_j(x)$ is provably the unique z^X satisfying

$$\chi_j(z) =_0 1 \wedge \forall y^X \left(\chi_j(y) =_0 1 \to \|x - z\| \leq_{\mathbb{R}} \|x - y\|\right).$$

This also implies that

$$\mathcal{D}^{\omega}_{\trianglelefteq} \vdash \forall j^0\, \forall x^X \left(\chi_j(x) =_0 1 \to P_j(x) =_X x\right)$$

and in particular, by (D6), $\forall j^0 \left(P_j(p) =_X p\right)$. Furthermore, we have that for all j^0 and x^X, y^X

$$\begin{aligned}
\|P_j(x) - P_j(y)\|^2 &=_{\mathbb{R}} \langle P_j(x) - P_j(y), P_j(x) - P_j(y)\rangle \\
&=_{\mathbb{R}} \langle P_j(x) - x, P_j(x) - P_j(y)\rangle + \langle x - y, P_j(x) - P_j(y)\rangle \\
&\quad + \langle y - P_j(y), P_j(x) - P_j(y)\rangle \\
&\leq_{\mathbb{R}} \langle x - y, P_j(x) - P_j(y)\rangle
\end{aligned}$$

[7] Note the innocuous double use of the notation x_0.
[8] One may change the reference point used in the majorizability notion from 0_X to x_0 (or p) allowing to work simply with a numerical bound on $\|x_0 - p\|$, cf. [41]. We do not care to do this here.

which entails that $\|P_j(x) - P_j(y)\| \leq_{\mathbb{R}} \|x - y\|$, i.e., P_j is provably nonexpansive in $\mathcal{D}^{\omega}_{\trianglelefteq}$. This, in turn, implies the extensionality of P_j.

Adapting the Theorem 2.8 to this setting, we have the following result stating that the formal system $\mathcal{D}^{\omega}_{\trianglelefteq}$ is suitable for bound extractions.

THEOREM 5.2 (Extraction in $\mathcal{D}^{\omega}_{\trianglelefteq}$). *For an arbitrary formula $A(\underline{z})$ of the language $\mathcal{L}(\mathcal{D}^{\omega}_{\trianglelefteq})$ with free-variables $\underline{z}$, let $A(\underline{z})^U$ be $\tilde{\forall}\underline{b}\,\tilde{\exists}\underline{c}\,A_U(\underline{z},\underline{b},\underline{c})$. If*

$$\mathcal{D}^{\omega}_{\trianglelefteq} + \mathsf{mAC}^X_{\mathsf{bd}} + \mathsf{BC}^X_{\mathsf{bd}} + \mathsf{MAJ}^X \vdash A(\underline{z}),$$

then there are closed monotone terms $\underline{t}$ of $\mathcal{L}(\mathcal{D}^{\omega}_{\trianglelefteq})$ such that

$$\mathcal{D}^{\omega}_{\trianglelefteq} \vdash \tilde{\forall}\underline{a},\underline{b}\,\forall\underline{z} \trianglelefteq \underline{a}\,A_U(\underline{z},\underline{b},\underline{t}\,\underline{a}\,\underline{b}).$$

Moreover, the terms $\underline{t}$ can be extracted from a proof witnessing the assumption.

Proof. We just need to verify the interpretation of the new axioms, and that we still have a majorizability theory, i.e., that all the introduced constants are majorizable. For the first part, note that the axioms (D1)–(D11) are given by universal sentences (possibly with bounded matrices) or even quantifier-free formulas. Therefore, they are trivially interpreted by themselves. The second requirement is equally easy. It is clear for m and b as they are of type 0 and so self-majorizable. The majorizability of x_0 and p follows from (D2) using Lemma 2.5(*iii*). By the axiom (D3), the constant $\chi_{(\cdot)}$ is majorizable by the constant $\lambda j, n.\,(1^0)$. From (D9)–(D11), the majorizability of $x_{(\cdot)}$ and $q_{(\cdot)}$ follows from the majorizability of x_0, 0_X and $P_{(\cdot)}$ (as well as from the majorizability of $+_X$ and $-_X$). Hence, we just need to see that $P_{(\cdot)}$ is majorizable. Indeed, since for any j^0 and x^X, we have

$$\begin{aligned}\|P_j(x)\| &\leq_{\mathbb{R}} \|P_j(x) - p\| + \|p\| \\ &=_{\mathbb{R}} \|P_j(x) - P_j(p)\| + \|p\| \\ &\leq_{\mathbb{R}} \|x - p\| + \|p\| \\ &\leq_{\mathbb{R}} \|x\| + \|p\| + \|p\|\end{aligned}$$

it follows from Lemma 2.5 that $P_{(\cdot)}$ is majorized by the constant $\lambda j, n.\,(n + 2b + 1)$. □

5.3. On the intensional perspective. One may wonder that, if the set C_j are *apparently* given by the universal property $\forall k^0\left(\|x - P_j(x)\| \trianglelefteq_{\mathbb{R}} \left(\frac{1}{k+1}\right)_{\mathbb{R}}\right)$, then why is it that we can formalize the projection principle

$$P_j(x) \in C_j \wedge \forall y^X\left(y \in C_j \to \|x - P_j(x)\| \leq_{\mathbb{R}} \|x - y\|\right),$$

without any concern for the issues discussed in §3. It turns out that $\chi_j(x) =_0 1$ is an imperfect description of the set C_j, and subsequently this projection is formally also weaker. We have shown that $\chi_j(x) =_0 1 \to P_j(x) =_X x$ is provable in $\mathcal{D}^{\omega}_{\trianglelefteq}$. However, our system crucially does not encompass the reverse direction. This is a required feature. In fact, we even have that

$$\mathcal{D}^{\omega}_{\trianglelefteq} + \mathsf{mAC}^X_{\mathsf{bd}} + \mathsf{BC}^X_{\mathsf{bd}} + \mathsf{MAJ}^X \nvdash P_j(x) =_X x \to \chi_j(x) =_0 1.$$

Otherwise, by Theorem 5.2, in any suitable structure for the (flattened version of) our system there would exist $k \in \mathbb{N}$ such that

$$\|P_j(x) - x\| \leq \frac{1}{k+1} \to x \in C_j,$$

which is in general false unless the C_j are taken to be the full space X. This distinction between the universal predicate $P_j(x) =_X x$ and the quantifier-free $\chi_j(x) =_0 1$, is actual connected with the absence of a prominent feature in the constant χ_j: its extensionality,

$$(\mathrm{E}_\chi) \qquad \forall j^0\, \forall x^X\, \forall y^X \left((x =_X y \wedge \chi_j(x) =_0 1) \to \chi_j(y) =_0 1\right).$$

Indeed, using (D7), we trivially have

$$\mathcal{D}^\omega_{\trianglelefteq} \vdash (\mathrm{E}_\chi) \to \forall j^0\, \forall x^X \left(P_j(x) =_X \to \chi_j(x) =_0 1\right).$$

For the reverse direction, observe that since P_j is provably extensional one derives the implication $(x =_X y \wedge \chi_j(x) =_0 1) \to P_j(y) =_X y$. We conclude that

$$\mathcal{D}^\omega_{\trianglelefteq} \vdash \forall j^0\, \forall x^X \left(P_j(x) =_X \to \chi_j(x) =_0 1\right) \to (\mathrm{E}_\chi).$$

This kind of intensional perspective have been used in proof mining before, but its tremendous usefulness was only recently shown through the work of Nicholas Pischke (see [43, 44]).

5.4. A modified proof of Dykstra's convergence. We now discuss how one can prove that Dykstra's algorithm is a Cauchy sequence in our formal system $\mathcal{D}^\omega_{\trianglelefteq}$ avoiding the compactness arguments central to the original proof by using instead $\mathsf{BC}^X_{\mathsf{bd}}$. We first remark that the steps (1)–(6) immediately formalize in our system. From the proof that (x_n) is bounded, we have $\forall n^0\, (x_n \trianglelefteq_X 3b)$. Moreover, the system proves that

$$\forall n^0\, \forall z^X \left(\bigwedge_{j=0}^{m-1} z =_X P_j(z) \to s(n,z) \geq_{\mathbb{R}} 0\right),$$

where $s(n,z)$ is as before (but now with shifted indexes):

$$s(n,z) := 2 \sum_{k=n}^{n+m-1} \langle x_k - z, q_k \rangle.$$

Consider the ε-weak projection associated with the universal formula $\forall n^0 \Omega(x,n)$, where

$$\Omega(x,n) \equiv \bigwedge_{j=0}^{m-1} \|x - P_j(x)\| \trianglelefteq_{\mathbb{R}} \frac{1}{n+1}.$$

With arguments as in Lemmas 3.2 and 3.3, one can prove an ε-weak version of the variational inequality characterization, and so for any k^0 there exists a point z^X such that $\bigwedge_{j=0}^{m-1} z =_X P_j(z)$ and

$$\forall y \trianglelefteq_X 3b \left(\bigwedge_{j=0}^{m-1} y =_X P_j(y) \to \langle z - x_0, z - y \rangle \trianglelefteq_{\mathbb{R}} \frac{1}{k+1}\right),$$

i.e.,

$$\forall y \trianglelefteq_X 3b \left(\forall r^0 \bigwedge_{j=0}^{m-1} \|y - P_j(y)\| \trianglelefteq_{\mathbb{R}} \frac{1}{r+1} \to \langle z - x_0, z - y \rangle \trianglelefteq_{\mathbb{R}} \frac{1}{k+1} \right).$$

Using $\mathsf{BC}^X_{\mathsf{bd}}$, it follows that there exists R^0 such that

$$\forall y \trianglelefteq_X 3b \left(\bigwedge_{j=0}^{m-1} \|y - P_j(y)\| \trianglelefteq_{\mathbb{R}} \frac{1}{R+1} \to \langle z - x_0, z - y \rangle \trianglelefteq_{\mathbb{R}} \frac{1}{k+1} \right).$$

While the original proof only needed to prove that the sequence (x_n) was asymptotically regular, in our modified proof we further need to argue that it is P_j-asymptotically regular for all $j \leq m-1$, i.e., $\forall j \leq_0 m-1 \left(\lim \|x_n - P_j(x_n)\| = 0\right)$, which means that

$$\forall R^0 \, \exists N^0 \, \forall n \geq_0 N \left(\bigwedge_{j=0}^{m-1} \|x_n - P_j(x_n)\| \trianglelefteq_{\mathbb{R}} \frac{1}{R+1} \right).$$

This property can be derived from the usual asymptotic regularity by leveraging the properties of P_j and using some simple combinatorial arguments (cf. [41, proposition 3.7]). Therefore, the P_j-asymptotic regularity of (x_n) is provable in $\mathcal{D}^{\omega}_{\trianglelefteq}$. We overall have shown that

$$(\star) \qquad \forall k^0 \, \exists z^X \, \exists N^0 \left(\bigwedge_{j=0}^{m-1} z =_X P_j(z) \wedge \forall n \geq_0 N \left(\langle z - x_0, z - x_n \rangle \trianglelefteq_{\mathbb{R}} \frac{1}{k+1} \right) \right).$$

Remark 5.3. *We take this moment to point out that this result can be view as an instance of Theorem 2.11 with* $\varphi(x, y) = \langle x - x_0, y \rangle$ *and the functional* T *defined for each* x^X *by* $T(x) := P_{i_0}(x)$, *where* $i_0 := \min\{j \leq m-1 : \forall j' \leq_0 m-1 \left(\|x - P_{j'}(x)\| \leq_{\mathbb{R}} \|x - P_j(x)\|\right)\}$. *Alternatively, look at the similar result but immediately stated for a finite number of maps in* [15, *proposition* 6.1].

We return to our modified proof, now arguing that $(\star)$ is sufficient for the main combinatorial argument to go through, concluding the proof. Corresponding to the technical lemma in step (7), we have that provably in $\mathcal{D}^{\omega}_{\trianglelefteq}$,

$$\liminf \sum_{k=n}^{n+m-1} |\langle x_k - x_{n-m+1}, q_k \rangle| =_{\mathbb{R}} 0_{\mathbb{R}},$$

and from $(\star)$, we derive

$$(\star\star) \; \forall k^0 \exists z^X \exists n^0 \left(\bigwedge_{j=0}^{m-1} z =_X P_j(z) \wedge \langle z - x_0, z - x_n \rangle \trianglelefteq_{\mathbb{R}} \frac{1}{k+1} \wedge s(n, x_{n-m+1}) \trianglelefteq_{\mathbb{R}} \frac{1}{k+1} \right).$$

We can consider the arguments used in Lemma 5.1 from step (6) in the original proof, and obtain

$$\forall k^0\,\exists r^0\,\forall z^X\,\forall n^0\left(\bigwedge_{j=0}^{m-1} z =_X P_j(z) \wedge \|x_{n+m-1}-z\|, s(n,x_{n+m-1}) \trianglelefteq_{\mathbb{R}} \frac{1}{r+1} \right.$$
$$\left. \to \forall i \geq n\left(\|x_{i+m-1}-z\| \trianglelefteq_{\mathbb{R}} \frac{1}{k+1}\right)\right).$$

The proof concludes using the following result.

LEMMA 5.4. *The following statement holds in* $\mathcal{D}^{\omega}_{\trianglelefteq}$

$$\forall r^0\,\exists \ell^0\,\forall z^X\,\forall n^0$$
$$\left(\bigwedge_{j=0}^{m-1} z =_X P_j(z) \wedge \langle x_0-z, x_{n+m-1}-z\rangle, s(n,x_{n+m-1}) \leq_{\mathbb{R}} \frac{1}{\ell+1}\right)$$
$$\to \|x_{n+m-1}-z\| \leq \frac{1}{r+1}.$$

Proof. We have,

$$\begin{aligned}
\|x_{n+m-1}-z\|^2 &=_{\mathbb{R}} \langle x_0-z, x_{n+m-1}-z\rangle + \langle x_{n+m-1}-x_0, x_{n+m-1}-z\rangle\\
&\leq_{\mathbb{R}} \langle x_0-z, x_{n+m-1}-z\rangle + \sum_{k=n}^{n+m-1}\langle z-x_{n+m-1}, q_k\rangle\\
&=_{\mathbb{R}} \langle x_0-z, x_{n+m-1}-z\rangle + \frac{1}{2}\left(s(n,x_{n+m-1})-s(n,z)\right)\\
&\leq_{\mathbb{R}} \langle x_0-z, x_{n+m-1}-z\rangle + \frac{1}{2}s(n,x_{n+m-1}),
\end{aligned}$$

which entails the result. □

We have therefore obtained overall

$$\forall k^0\,\exists z^X\,\exists n^0\,\forall i \geq_0 n\left(\|x_{i+m-1}-z\| \leq_{\mathbb{R}} \frac{1}{k+1}\right),$$

and so, by triangle inequality, we conclude that (x_n) is a Cauchy sequence.

It is this modified proof in the system $\mathcal{D}^{\omega}_{\trianglelefteq}$ that was analysed in [41] and resulted in quantitative data in the form of a primitive recursive functional Δ providing an uniform rates of metastability for Dykstra's algorithm.

THEOREM 5.5 [41, Theorem 3.11]. *Let $C_1,\ldots,C_m$ be $m \geq 2$ convex subsets of a Hilbert space X such that $\bigcap_{j=1}^m C_j \neq \emptyset$. Let $x_0 \in X$ and $b \in \mathbb{N}\setminus\{0\}$ be given such that $b \geq \|x_0-p\|$, for some $p \in \bigcap_{j=1}^m C_j$. Then, the sequence (x_n) generated by (D) is metastable and*

$$\forall \varepsilon \in (0,1]\,\forall f:\mathbb{N}\to\mathbb{N}\,\exists n \leq \Delta(m,b,\varepsilon,f)\,\forall i,j \in [n;n+f(n)]\left(\|x_i-x_j\| \leq \varepsilon\right).$$

Finally, note that a pair (z,n) for the premise of Lemma 5.1 can also be obtained under a metric regularity result, i.e., under the assumption that

$$(\circledast)\quad \forall \varepsilon>0\,\forall R\in\mathbb{N}\,\exists \delta>0\,\forall x \in \overline{B}_R(0)\left(\bigwedge_{j=1}^m \|x-P_j(x)\| \leq \delta \to \exists z\in C\left(\|x-z\| \leq \varepsilon\right)\right),$$

together with the fact that (x_n) is P_j-asymptotically regular for each $j \in [1;m]$. Indeed, let $k \in \mathbb{N}$ be given and consider $r \in \mathbb{N}$ as in Lemma 5.1. For $\varepsilon = \dfrac{1}{r+1}$ and with R a bound on the Dykstra's iteration, take $\delta > 0$ as guaranteed to exist in the assumption of metric regularity. By the P_j-asymptotic regularity, we have $N \in \mathbb{N}$ such that x_n satisfies the premise of $(\circledast)$, for any $n \geq N$. By step (7), we can find $n_0 \geq N$ such that $s(n_0, x_{n_0}) \leq \dfrac{1}{r+1}$. Now, by the assumption on δ, we have some $z \in C$ such that $\|x_{n_0} - z\| \leq \dfrac{1}{r+1}$. By Lemma 5.1, we conclude that

$$\forall i \geq n_0 \left(\|x_i - z\| \leq \frac{1}{k+1} \right).$$

It is known that the assumption of metric regularity is very strong and, in the case of Fejér monotone sequences, it will entail the existence of uniform rates of convergence. However, Dykstra's algorithm is not Fejér monotone and still rates of convergence were possible via the argument above. Motivated by this result (see [41, sec. 4]), a localized and relativized generalization of the usual concept of Fejér monotonicity was introduced and studied in the recent [35]. In particular, [35] provides a much simpler convergence proof of Dykstra's algorithm in the finite dimensional case. Moreover, note that despite the final output of a rate of convergence for the iteration, the study of asymptotic regularity for the sequence up to that point was only in the form of a rate for the metastable formulation. Naturally, a posteriori, the metastable rates for asymptotic regularity get upgraded into full rates of convergence, but this too strengthens the usefulness of metastablity. In fact, one should not avoid it in preference of rates of convergence at the risk of missing out on subtle hidden quantitative information.

§6. Epilogue. This paper provides a proof-theoretical justification to the recent proof mining study in [41]. There, the asymptotic behaviour of Dykstra's algorithm was analysed and metastability rates were obtained which are primitive recursive in the sense of Gödel. However, a naive interpretation of the original convergence proof would require quantitative data given by Spector's bar-recursive functionals and their absence was unexpected. We discuss the proof-theoretical arguments underlying the analysis in [41] by explaining that, in the presence of tame bounded collection principles, one can sidestep the compactness arguments crucial in the original proof which would otherwise require the functional interpretation of arithmetical comprehension.

There were previous proof mining case studies in which simplifications allow to avoid the use of certain compactness arguments (e.g., [30]), and recent applications which featured such phenomenon were supported by the theoretical justification given in [15]. However, all previous such examples were of a similar nature: they always regard an iterative method akin to the Halpern schema and the arguments are always similar to the Wittmann's proof strategy. One could therefore question the generality in practice of the approach from [15]. This paper also answers this issue. On one hand, we discuss how the convergence proof for the case of Dykstra's algorithm is significantly different than Wittmann's argument. On the other, we are still able to bypass the troublesome arguments relying on the technique from [15]. Naturally, this

does not settle the issue and further study is needed to fully understand the reach of this approach using bounded collection principles.

The ability to provide a simplified convergence proof reliant on a metastable formulation gives a strong argument for the usefulness of metastability and of proof-theoretical methods in mathematics. Like in the example discussed in §4.2, mathematical proofs that are not reliant on strong analytical principles but instead mostly consist on combinatorial arguments are easier to generalize. For the case at hand, the availability of a simpler proof allowed to establish in [42] the strong convergence of Dykstra's algorithm with Bregman projection maps in the context of general reflexive Banach spaces.

Acknowledgements. The author is grateful to Nicholas Pischke and Andrei Sipoş for comments which improved the presentation of this paper.

Funding. The author acknowledges the support of the Mathematisches Forschungsinstitut Oberwolfach in 2024 through the program "Oberwolfach Leibniz Fellows". Further, the author was supported by the German Science Foundation (DFG – Ref. PI 2070/1-1, Projektnummer 549333475).

BIBLIOGRAPHY

[1] Bezem, M. (1985). Strongly majorizable functionals of finite type: A model for bar recursion containing discontinuous functionals. *The Journal of Symbolic Logic*, **50**(3), 652–660.

[2] Boyle, J. P., & Dykstra, R. L. (1986). A method for finding projections onto the intersection of convex sets in Hilbert spaces. In Dykstra, R. L., Robertson, T., & Wright, F. T., editors. *Advances in Order Restricted Statistical Inference (Iowa City, Iowa, 1985)*, volume 37 of Lecture Notes in Statistics, Berlin: Springer, pp. 28–47.

[3] Bregman, L. (1965). The method of successive projections for finding a common point of convex sets. *Soviet Mathematics—Doklady*, **6**, 688–692.

[4] Browder, F. E. (1965). Nonexpansive nonlinear operators in a Banach space. *Proceedings of the National Academy of Sciences U.S.A.*, **54**, 1041–1044.

[5] Dinis, B., & Pinto, P. (2020). Metastability of the proximal point algorithm with multi-parameters. *Portugaliae Mathematica*, **77**(3–4), 345–381.

[6] ———. (2021). On the convergence of algorithms with Tikhonov regularization terms. *Optimization Letters*, **15**(4), 1263–1276.

[7] ———. (2021). Quantitative results on the multi-parameters proximal point algorithm. *Journal of Convex Analysis*, **28**(3), 729–750.

[8] ———. (2023). Strong convergence for the alternating Halpern-Mann iteration in CAT(0) spaces. *SIAM Journal on Optimization*, **33**(2), 785–815.

[9] ———. (2024). Effective metastability for a method of alternating resolvents. *Fixed Point Theory*, **25**(1), 61–98.

[10] Dykstra, R. L. (1983). An algorithm for restricted least squares regression. *Journal of the American Statistical Association*, **78**(384), 837–842.

[11] Engrácia, P. (2009). *Proof-Theoretical Studies on the Bounded Functional Interpretation*. PhD Thesis, Universidade de Lisboa.

[12] Engrácia, P., & Ferreira, F. (2020). Bounded functional interpretation with an abstract type. In Rezuş, A., editor. *Contemporary Logic and Computing*, volume 1 of Landscapes in Logic, London: College Publications, pp. 87–112.

[13] Ferreira, F. (2009). Injecting uniformities into Peano arithmetic. *Annals of Pure and Applied Logic*, **157**(2–3), 122–129.

[14] ———. (2021). The abstract type of the real numbers. *Archive for Mathematical Logic*, **60**(7–8), 1005–1017.

[15] Ferreira, F., Leuştean, L., & Pinto, P. (2019). On the removal of weak compactness arguments in proof mining. *Advances in Mathematics*, **354**, 106728. 55.

[16] Ferreira, F., & Oliva, P. (2005). Bounded functional interpretation. *Annals of Pure and Applied Logic*, **135**(1–3), 73–112.

[17] Gerhardy, P., & Kohlenbach, U. (2008). General logical metatheorems for functional analysis. *Transactions of the American Mathematical Society*, **360**(5), 2615–2660.

[18] Gödel, K. (1958). Über eine bisher noch nicht benützte Erweiterung des finiten Standpunktes. *Dialectica*, **12**(3–4), 280–287.

[19] Günzel, D., & Kohlenbach, U. (2016). Logical metatheorems for abstract spaces axiomatized in positive bounded logic. *Advances in Mathematics*, **290**, 503–551.

[20] Halperin, I. (1962). The product of projection operators. *Acta Universitatis Szegediensis. Acta Scientiarum Mathematicarum*, **23**, 96–99.

[21] Halpern, B. (1967). Fixed points of nonexpanding maps. *Bulletin of the American Mathematical Society*, **73**(6), 957–961.

[22] Hilbert, D. (1926). Über das Unendliche. *Mathematische Annalen*, **95**(1), 161–190.

[23] Howard, W. A. (1973). Hereditarily majorizable functionals of finite type. In Troelstra, A. S., editor. *Metamathematical Investigation of Intuitionistic Arithmetic and Analysis*, Lecture Notes in Mathematics, 344, New York, NY: Springer, pp. 454–461.

[24] Hundal, H. S. (2004). An alternating projection that does not converge in norm. *Nonlinear Analysis: Theory, Methods & Applications*, **57**(1), 35–61.

[25] Kohlenbach, U. (1996). Analysing proofs in analysis. In Hodges, W., Hyland, M., Steinhorn, C., & Truss, J., editors. *Logic: From Foundations to Applications: European Logic Colloquium (Keele, 1993)*, Oxford: Oxford University Press, pp. 225–260.

[26] ———. (2005). Some logical metatheorems with applications in functional analysis. *Transactions of the American Mathematical Society*, **357**(1), 89–128.

[27] ———. (2006). A logical uniform boundedness principle for abstract metric and hyperbolic spaces. In Mints, G., & de Queiroz, R., editors. *Proceedings of the 13th Workshop on Logic, Language, Information and Computation (WoLLIC 2006)*, volume 165 of Electronic Notes in Theoretical Computer Science, Amsterdam: Elsevier Sci. B. V., pp. 81–93.

[28] ———. (2008). *Applied Proof Theory: Proof Interpretations and their Use in Mathematics*. Springer Monographs in Mathematics, Springer Verlag Berlin Heidelberg: Springer Science & Business Media.

[29] ———. (2010). On the logical analysis of proofs based on nonseparable Hilbert space theory. In Feferman, S., & Sieg, W., editors. *Proofs, Categories and Computations*, volume 13 of Tributes, London: College Publications, pp. 131–143.

[30] ———. (2011). On quantitative versions of theorems due to F.E. Browder and R. Wittmann. *Advances in Mathematics*, **226**(3), 2764–2795.

[31] ———. (2019). Proof-theoretic Methods in Nonlinear Analysis. In Sirakov, B., de Souza, P. N., & Viana, M., editors. *Proceedings of the International Congress of Mathematicians 2018*, Vol. 2. World Scientific Publishing Co Pte Ltd, Singapore, pp. 62–82.

[32] ———. (2020). Quantitative analysis of a Halpern-type proximal point algorithm for accretive operators in Banach spaces. *Journal of Nonlinear and Convex Analysis*, **21**(9), 2125–2138.

[33] Kohlenbach, U., & Leuştean, L. (2012). Effective metastability of Halpern iterates in CAT(0) spaces. *Advances in Mathematics*, **231**(5), 2526–2556. Addendum in: Adv. Math., 250:650–651, 2014.

[34] ———. (2012). On the computational content of convergence proofs via Banach limits. *Philosophical Transactions of the Royal Society of London. Series A. Mathematical, Physical and Engineering Sciences*, **370**(1971), 3449–3463.

[35] Kohlenbach, U., & Pinto, P. (2024). Fejér monotone sequences revisited. *Journal of Convex Analysis*. ArXiv Preprint:2310.06528. To appear.

[36] Körnlein, D. (2016). *Quantitative Analysis of Iterative Algorithms in Fixed Point Theory and Convex Optimization*. PhD thesis, Technische Universität Darmstadt.

[37] ———. (2016). Quantitative strong convergence for the hybrid steepest descent method. ArXiv Preprint:1610.00517.

[38] Kreisel, G. (1976). What have we learnt from Hilbert's second problem? In Browder, F. E., editor. *Mathematical developments arising from Hilbert problems (Proc. Sympos. Pure Math., Northern Illinois Univ., De Kalb, Ill., 1974)*, volume Vol. 28 of Proceedings of Symposia in Pure Mathematics, Providence, RI: American Mathematical Society, pp. 93–130.

[39] Pinto, P. (2019). *Proof Mining with the Bounded Functional Interpretation*. PhD thesis, Universidade de Lisboa.

[40] ———. (2021). A rate of metastability for the Halpern type proximal point algorithm. *Numerical Functional Analysis and Optimization*, **42**(3), 320–343.

[41] ———. (2025). On the finitary content of Dykstra's cyclic projections algorithm. *Zeitschrift für Analysis und ihre Anwendungen*, **44**(1-2), 165–192.

[42] Pinto, P., & Pischke, N. (2024). On Dykstra's algorithm with Bregman projections. Oberwolfach Preprint: OWP-2024-04.

[43] Pischke, N. (2024). Logical metatheorems for accretive and (generalized) monotone set-valued operators. *Journal of Mathematical Logic*, **24**(2), Paper No. 2350008. 59pp.

[44] ———. (2024). Proof mining for the dual of a Banach space with extensions for uniformly Fréchet differentiable functions. *Transactions of the American Mathematical Society*, **377**(10), 7475–7517.

[45] Pischke, N., & Kohlenbach, U. (2024). Effective rates for iterations involving Bregman strongly nonexpansive operators. *Set-Valued and Variational Analysis. Theory and Applications*, 32(4), Article No. 33, 58pp.

[46] Spector, C. (1962). Provably recursive functionals of analysis: a consistency proof of analysis by an extension of principles formulated in current intuitionistic mathematics. In Dekker, F. D. E., editor. *Recursive Function Theory*, Vol. 5. Providence, RI: American Mathematical Society, pp. 1–27.

[47] Tao, T. (2008). *Structure and Randomness: Pages from Year One of a Mathematical Blog*. Providence, RI: American Mathematical Society.

[48] von Neumann, J. (1950). *Functional Operators. II. The Geometry of Orthogonal Spaces*, Annals of Mathematics Studies, 22, Princeton, NJ: Princeton University Press. Reprint of mimeographed lecture notes first distributed in 1933

[49] Wittmann, R. (1992). Approximation of fixed points of nonexpansive mappings. *Archiv der Mathematik*, **58**(5), 486–491.

[50] Xu, H.-K. (2002). Iterative algorithms for nonlinear operators. *Journal of the London Mathematical Society*, **66**(1), 240–256.

DEPARTMENT OF MATHEMATICS, TECHNISCHE
UNIVERSITÄT DARMSTADT
SCHLOSSGARTENSTRAßE 7, 64289 DARMSTADT
GERMANY
E-mail: pinto@mathematik.tu-darmstadt.de