On the logical analysis of proofs based on nonseparable Hilbert space theory

Ulrich Kohlenbach*

Department of Mathematics
Technische Universität Darmstadt
Schlossgartenstraße 7, 64289 Darmstadt, Germany

February 7, 2010

Dedicated to Grigori Mints for his 70th Birthday

1 Introduction

Starting in [15] and then continued in [9, 17, 24] and [18], general logical metatheorems were developed that guarantee the extractability of highly uniform effective bounds from proofs of theorems that hold for general classes of structures such as metric, hyperbolic, CAT(0), normed or Hilbert spaces. To obtain uniformity e.g. w.r.t. parameters that range over metrically bounded (but not compact) sets it is crucial to exploit the fact that the proof to be analyzed does not use any separability assumption on the underlying spaces (as e.g. the existence of uniform bounds for the very statement of separability would yield the total boundedness of bounded metric spaces, see [18] for a detailed discussion of this issue). In order to do so we developed in [15] formal systems $\mathcal{T}^\omega[X,\ldots]$ that treat such abstract spaces X as atoms added to fragments $\mathcal{T}^\omega$ of full (though only weakly extensional) analysis $\mathcal{A}^\omega$ by adding a new base type X for variables ranging over X and all the finite types built upon $\mathbb{N}$ and X. The metatheorems developed in the aforementioned papers and the book [18] were based on novel extensions of (monotone [12]) functional interpretation in the sense of Gödel and Spector. These theorems made it possible to explain some concrete proof unwindings that had been carried out in fixed point theory resulting in unexpectedly uniform bounds (see e.g. [13, 14, 21]) and paved the way for many new applications in fixed point theory, geodesic geometry, ergodic theory and topological dynamics that were guaranteed to be possible by these proof theoretic results (see e.g. [1, 3, 4, 7, 8, 16, 20, 23, 22]). For applications of related forms of 'proof mining'

*The author has been supported by the German Science Foundation (DFG Project KO 1737/5-1).

in proof theory itself see Mints [25, 26].
$\mathcal{A}^\omega$ is a very strong formal system for analysis as it contains the full axiom schemas of dependent and countable choice in all types and so full arithmetical comprehension over numbers. While the latter feature makes it possible to formalize virtually all proofs in ordinary analysis for **separable** spaces it is not clear a priori how much e.g. of Hilbert space theory for $(X, \langle\cdot,\cdot\rangle)$ can be carried out in the absence of any separability condition on X in $\mathcal{A}^\omega[X, \langle\cdot,\cdot\rangle, \mathcal{C}]$ (see [18] for the definition of this theory) or even fragments $\mathcal{T}^\omega[X, \langle\cdot,\cdot\rangle, \mathcal{C}]$.

In this paper, we show that general orthogonal projection arguments, the Riesz representation theorem and the weak sequential compactness of the unit ball in X can be proved in such systems. As an application we show that a theorem of Browder [5] (stating the convergence of a certain explicit iteration sequence) can be proved and use this to obtain (by applying a metatheorem from [18]) a highly uniform effective rate of metastability in the sense of Tao [29, 30] on that convergence.

All undefined notions and notations in this paper (including the representation of real numbers by number theoretic functions $f : \mathbb{N} \to \mathbb{N}$ and the corresponding operations $+_\mathbb{R}, -_\mathbb{R}, \ldots$ and relations $<_\mathbb{R}, \leq_\mathbb{R}, \ldots$ on these representatives) are understood as in [18] on which this paper relies. We denote the type for the natural numbers $\mathbb{N} := \{0, 1, 2, \ldots\}$ (resp. for number theoretic functions $f : \mathbb{N} \to \mathbb{N}$) by 0 (resp. 1). Since all theories used in this paper are weakly extensional in the sense of [18] we drop the prefix 'WE-' in their name and e.g. write PA^ω instead of WE-PA^ω.

2 Main results

Definition 2.1. A formula $A(x^X) \in \mathcal{L}(\mathcal{T}^\omega[X, \|\cdot\|, \ldots])$ (for some theory $\mathcal{T}^\omega$ in the language of functionals of all finite types such as PA^ω) defines

1. a nonempty subset of X if
   (a) $\forall x^X, y^X\, (A(x) \wedge x =_X y \to A(y))$,
   (b) $\exists x^X\, A(x)$;
2. a nonempty convex subset of X if in addition to (a) and (b)
$$\forall \lambda^1, x^X, y^X\ (A(x) \wedge A(y) \to A((1 -_\mathbb{R} \tilde{\lambda}) \cdot_X x +_X \tilde{\lambda} \cdot_X y)),$$
   where $\lambda \mapsto \tilde{\lambda}$ is the construction for the representation of $[0, 1]$ from [18](Definition 4.24);
3. a linear subspace of X if (a) and
$$A(0_X) \wedge \forall \alpha^1, \beta^1, x^X, y^X\ (A(x) \wedge A(y) \to A(\alpha \cdot_X x +_X \beta \cdot_X y));$$

4. a closed nonempty convex subset (resp. closed linear subspace) of X if in addition to the conditions for nonempty convex subsets (resp. linear subspaces) we have that

$$\forall x_{(\cdot)}^{0\to X}\,\big(\forall n\, A(x_n) \wedge \forall m^0, n^0, k^0\, (m, n \geq k \to \|x_m - x_n\|_X \leq_{\mathbb{R}} 2^{-k}) \to A(C(x_{(\cdot)}))\big).$$

Here C is the operator from [18] that maps each fast converging Cauchy sequence in X to its limit thereby expressing the completeness of X.

We say that A defines provably in some theory $\mathcal{T}$ whose language is contained in $\mathcal{L}(\mathrm{PA}^\omega[X, \|\cdot\|])$ resp. of $\mathcal{L}(\mathrm{PA}^\omega[X, \|\cdot\|, \mathcal{C}])$ any of the concepts above if the corresponding formulas are provable in $\mathcal{T}$. In fact, instead of PA^ω this definition can also be applied to any fragment (or extension) as long as the ingredients to formulate the respective formulas above are contained in the language.

Remark 2.2. Definition 2.1.1 resp. 2.1.2 can also be applied to formulas in $\mathcal{L}(\mathrm{PA}^\omega[X, d])$ resp. $\mathcal{L}(\mathrm{PA}^\omega[X, d, W])$, where $\mathrm{PA}^\omega[X, d]$ (resp. $\mathrm{PA}^\omega[X, d, W]$) is the extension of PA^ω by an abstract metric (resp. hyperbolic) space X (see [18] for details).

Convention on notation: From now on we will adopt a more informal notation and e.g. write $\|\cdot\|, k \in \mathbb{N}, \lambda \in \mathbb{R}, x \in X$ and $\alpha x + \beta y$ instead of $\|\cdot\|_X, k^0, \lambda^1, x^X$ and $\alpha \cdot_X x +_X \beta \cdot_X y$ and only use (partially) the latter when the precise logical form matters.

Proposition 2.3. *1.* $\mathrm{PA}^\omega[X, \|\cdot\|]$ *proves that for every nonempty subset* $S \subseteq X$ *given by a formula* A *as in the definition above (using* $\exists y \in S \ldots$, *resp.* $\forall y \in S \ldots$, *as shorthand for* $\exists y^X\, (A(y) \wedge \ldots)$ *resp.* $\forall y^X\, (A(y) \to \ldots)$*) the following holds:*

$$\forall k \in \mathbb{N} \forall x \in X \exists y \in S \forall z \in S\ (\|x - y\| < \|x - z\| + 2^{-k}). \tag{1}$$

$(\mathrm{PA}^\omega + \mathrm{AC}^{0,0})[X, \|\cdot\|]$ *proves that* $d := \inf_{z \in S}\{\|x - z\|\}$ *exists.*

2. $\mathcal{A}^\omega[X, \langle\cdot,\cdot\rangle, \mathcal{C}]$ *proves that for every closed nonempty convex subset* C *given by a formula* A *the following holds:*

$$\forall x \in X \exists! y \in C \forall z \in C\ (\|x - y\| \leq \|x - z\|). \tag{2}$$

In particular, $d := \inf_{z \in C}\{\|x - z\|\}$ *not only exists but is attained by some* $y \in C$. *The result holds a-fortiori for closed linear subspaces of* X *that are given by a formula* A.

Proof: 1. Suppose that (1) fails for some $k \in \mathbb{N}, x \in X$ i.e.

$$\forall y \in S \exists z \in S\, (\|x - y\| \geq \|x - z\| + 2^{-k}). \tag{3}$$

Let $\widehat{y} \in S$. Then by induction on $n \in \mathbb{N}$ one shows that for every n there exists a sequence, i.e. a function $f^{0\to X}$, such that

$$f(0) =_X \widehat{y} \wedge \forall i \leq n (f(i) \in S) \wedge \forall i < n\, (\|x - f(i)\| \geq \|x - f(i+1)\| + 2^{-k}). \quad (4)$$

Now let $n > \|x - \widehat{y}\| \cdot 2^k$. Then there exists a function $f^{0\to X}$ such that

$$\|x - \widehat{y}\| = \|x - f(0)\| \geq \|x - f(n)\| + n \cdot 2^{-k} > \|x - f(n)\| + \|x - \widehat{y}\|,$$

which is a contradiction.
The last claim of '1.' follows from the fact that for every $k \in \mathbb{N}$ there exists a rational number r encoded by, say, m such that r is a 2^{-k-1} rational approximation to $\|x-y\|$, where y is from (1) for $k+1$. Now $\mathrm{AC}^{0,0}$ gives a fast converging Cauchy sequence (r_k) (encoded by (m_k)) that represents d.
2. Let $x \in X$ and $\widehat{y} \in C$. Applying to (1) countable choice $\mathrm{AC}^{0,X}$ yields a sequence (y_k) in C with

$$\forall k \in \mathbb{N} \forall z \in C\, (\|x - y_k\| < \|x - z\| + 2^{-k}). \quad (5)$$

It is not hard to verify that for uniformly convex normed spaces with modulus of convexity $\eta : (0, 2] \to (0, 1]$ the function

$$\Phi(\varepsilon) := \frac{\varepsilon}{4} \cdot \eta(\varepsilon/(D + 1)),$$

where $\|x - \widehat{y}\| \leq D \in \mathbb{N}$, is a so-called modulus of uniqueness for the projection, i.e.

$$\forall \varepsilon \in (0, 1), z_1, z_2 \in C\, \Big(\bigwedge_{i=1}^{2} (\forall z \in C(\|x - z_i\| \leq \|x - z\| + \Phi(\varepsilon)) \to \|z_1 - z_2\| \leq \varepsilon \Big) \quad (6)$$

(see e.g. proposition 17.4 in [18] for – a stronger version of – this).
In the case of a Hilbert space it is well-known that (for the best modulus η)

$$\eta(\varepsilon) = 1 - \sqrt{1 - \varepsilon^2/4} \geq \varepsilon^2/8 \quad (\varepsilon \in (0, 2]).$$

Now define

$$\Psi(k) := \min l \left[2^{-l} \leq \frac{2^{-3k}}{32(D + 1)^2} \right].$$

Then – using (5) and (6) – $(y_{\Psi(k)})_k$ is a Cauchy sequence with Cauchy rate 2^{-k}. Hence by the completeness axiom $(\mathcal{C})$ of $\mathcal{A}^\omega[X, \langle\cdot,\cdot\rangle, \mathcal{C}]$ the limit y exists in X and – by the condition that A defines a closed convex subset C – also in C. It is now easy to verify that

$$\|x - y\| \leq \|x - z\| \text{ for all } z \in C.$$

The uniqueness of y immediately follows from (6). □

Remark 2.4. Proposition 2.3.2 also holds for the theory $\mathcal{A}^\omega[X, \|\cdot\|, \eta, \mathcal{C}]$ of uniformly convex Banach spaces with a modulus of uniform convexity η.

The standard proof that, in the case where X is a Hilbert space and $\mathcal{L}$ a closed linear subspace, $x-y$ for the point y as in proposition 2.3.2 is orthogonal to $\mathcal{L}$ goes through in our formal context without problems. In fact, we have the following quantitative version:

Proposition 2.5. $\widehat{\mathrm{PA}}^\omega{\restriction}[X,\langle\cdot,\cdot\rangle]$ *proves the following: if $\mathcal{L}$ is a linear subspace given by a formula A, then*

$$\begin{cases} \forall\varepsilon>0\,\forall K\geq 1\,\forall x\in X\,\forall y,z\in\mathcal{L} \\ \qquad \big(\|z\|\leq K \wedge \|x-y\|^2\leq \|x-(y+\alpha z)\|^2+\frac{\varepsilon^2}{K^2}\to |\langle x-y,z\rangle|\leq\varepsilon\big),\end{cases}$$

where

$$\alpha := \frac{\langle x-y,z\rangle}{\max\Big(\big(\varepsilon/(2\max(\|x\|,\|y\|,1))\big)^2,\|z\|^2\Big)}.$$

In particular:

$$\forall x\in X\forall y\in\mathcal{L}\,(\forall z\in\mathcal{L}\,(\|x-y\|\leq\|x-z\|)\to\forall z\in\mathcal{L}(\langle x-y,z\rangle=0)).$$

Proof: Case 1:

$$\|z\|^2\leq\big(\varepsilon/(2\max(\|x\|,\|y\|,1))\big)^2.$$

Then

$$\|z\|\leq\frac{\varepsilon}{2\max(\|x\|,\|y\|,1)}$$

and so

$$|\langle x-y,z\rangle|\leq|\langle x,z\rangle|+|\langle y,z\rangle|\leq\|x\|\cdot\|z\|+\|y\|\cdot\|z\|\leq\varepsilon.$$

Case 2:

$$\|z\|^2>\big(\varepsilon/(2\max(\|x\|,\|y\|,1))\big)^2.$$

Then

$$\alpha=\frac{\langle x-y,z\rangle}{\|z\|^2}.$$

Hence

$$\begin{array}{l}\|x-y\|^2\leq\|x-(y+\alpha z)\|^2+\frac{\varepsilon^2}{K^2}=\langle(x-y)-\alpha z,(x-y)-\alpha z\rangle+\frac{\varepsilon^2}{K^2}=\\ \|x-y\|^2-2\alpha\langle x-y,z\rangle+\alpha^2\|z\|^2+\frac{\varepsilon^2}{K^2}=\\ \|x-y\|^2-2\frac{\langle x-y,z\rangle^2}{\|z\|^2}+\frac{\langle x-y,z\rangle^2}{\|z\|^2}+\frac{\varepsilon^2}{K^2}.\end{array}$$

Thus

$$\frac{\langle x-y,z\rangle^2}{\|z\|^2}\leq\frac{\varepsilon^2}{K^2}$$

and so (using that $\|z\|\leq K$)

$$\langle x-y,z\rangle^2\leq\varepsilon^2.$$

Hence $|\langle x-y,z\rangle|\leq\varepsilon$. ☐

A particularly relevant instance of proposition 2.3 (needed below to reduce the general case of weak compactness to the separable case) is the following one where we project to a closed linear subspace

$$\mathcal{L} := \overline{\mathrm{Lin}_{\mathbb{R}}\{x_n : n \in \mathbb{N}\}}$$

generated from a sequence (x_n) in X. Clearly, $\mathcal{L}$ can (provably in $\mathcal{A}^\omega[X, \langle\cdot,\cdot\rangle, \mathcal{C}]$) be represented by a formula A_L, namely

$$A_L(x) :\equiv \begin{cases} \exists f^1 \; ((\varphi(k, f, (x_n)))_k \text{ is a Cauchy sequence in } \mathcal{L}_{\mathbb{Q}} \text{ with Cauchy rate } 2^{-k} \\ \wedge\, x =_X \lim\limits_{k\to\infty} \varphi(k, f, (x_n))), \end{cases}$$

where

$$\mathcal{L}_{\mathbb{Q}} := \bigcup_{n\in\mathbb{N}} \{r_0x_0 + \ldots + r_nx_n : (r_0, \ldots, r_n) \in \mathbb{Q}^{n+1}\},$$

$$\varphi(k, f, (x_n)) := \sum_{i=0}^{lth(f(k))-1} q_{(f(k))_i} \cdot_X x_i$$

and $(q_n)_{n\in\mathbb{N}}$ is some primitive recursive standard enumeration of $\mathbb{Q}$ (we identify q_n with its canonical embedding into $\mathbb{R}$).
In this case it is, however, easier to work directly with the countable dense subset $\mathcal{L}_{\mathbb{Q}} \subset \mathcal{L}$:

Proposition 2.6. $\mathcal{A}^\omega[X, \langle\cdot,\cdot\rangle, \mathcal{C}]$ *proves the following: for every sequence* (x_n) *in* X *and every* $x \in X$ *there exists the projection of* x *to the closed linear subspace*

$$\mathcal{L} := \overline{\mathrm{Lin}_{\mathbb{R}}\{x_n : n \in \mathbb{N}\}}$$

generated from $\{x_n : n \in \mathbb{N}\}$*, i.e.*

$$\forall x \in X \exists! y \in \mathcal{L} \forall z \in \mathcal{L}\, (\|x - y\| \le \|x - z\|).$$

Moreover, instead of the full axiom schema DC *of dependent choice (or full countable choice combined with higher induction) the proof only needs (for given* x *and* (x_n)*) a fixed instance of* Π^0_1-AC *as well as* Σ^0_1*-induction (that alone suffices to prove the approximate version 2.3.1 in this case) and hence can be carried out, in particular, in* $\mathcal{T}^\omega[X, \langle\cdot,\cdot\rangle, \mathcal{C}]$*, where* $\mathcal{T}^\omega := \widehat{\mathrm{PA}}^\omega\!\upharpoonright + \mathrm{AC}^{0,0}_{ar}$.

Proof: Since $\mathcal{L}_{\mathbb{Q}}$ is dense in $\mathcal{L}$ we can represent elements of $\mathcal{L}$ as Cauchy sequences w.r.t. $\|\cdot\|$ of elements in $\mathcal{L}_{\mathbb{Q}}$ with Cauchy rate 2^{-k}. Let $(y_k)_{k\in\mathbb{N}}$ some primitive recursive (in (x_n)) standard enumeration of $\mathcal{L}_{\mathbb{Q}}$.
The existence of an ε-projection of x to $\mathcal{L}$ (in the sense of proposition 2.3.1) can then equivalently be formulated as

$$\forall x^X, k^0 \exists n^0 \forall m^0\, (\|x - y_n\| \le \|x - y_m\| + 2^{-k})$$

which can be proved by Σ^0_1-IA.
We then use Π^0_1-AC to form a sequence $(y_{f(k)})_{k\in\mathbb{N}}$ of 2^{-k}-projections and finish the proof as in the case of proposition 2.3.2. □

Definition 2.7. A formula $A_L(x^X, y^1)$ in $\mathcal{L}(\mathcal{T}^\omega[X,\ldots])$ is said to represent a bounded linear functional $L : X \to \mathbb{R}$ of a normed space X if

$$\exists C^0 \forall x^X \, \exists y^1 \, \big(A_L(x,y) \wedge |y|_{\mathbb{R}} \leq_{\mathbb{R}} (C)_{\mathbb{R}} \cdot_{\mathbb{R}} \|x\|_X\big), \tag{7}$$

$$\forall x^X \forall y^1, y_2^1 \, \big(A_L(x,y_1) \to (A(x,y_2) \leftrightarrow y_1 =_{\mathbb{R}} y_2))\big), \tag{8}$$

$$\forall \alpha_1^1, \alpha_2^1, y_1^1, y_2^1 \, \forall x_1^X, x_2^X \begin{cases} \big(A_L(x_1,y_1) \wedge A_L(x_2,y_2) \to \\ A_L(\alpha_1 \cdot_X x_1 +_X \alpha_2 \cdot_X x_2, \alpha_1 \cdot_{\mathbb{R}} y_1 +_{\mathbb{R}} \alpha_2 \cdot_{\mathbb{R}} y_2)\big). \end{cases} \tag{9}$$

In the following we write, more informally, '$C \cdot \|x\|$' instead of '$(C)_{\mathbb{R}} \cdot_{\mathbb{R}} \|x\|_X$' etc.

Proposition 2.8. Provably in $\widehat{\mathrm{PA}}^\omega{\restriction}[X, \|\cdot\|]$, every formula A_L that represents a bounded linear functional $L : X \to \mathbb{R}$ satisfies

1. $x_1 =_X x_2 \wedge y_1 =_{\mathbb{R}} y_2 \wedge A_L(x_1,y_1) \to A_L(x_2,y_2)$,
2. $A_L(x_1,y_1) \wedge A_L(x_2,y_2) \to |y_1 -_{\mathbb{R}} y_2| \leq_{\mathbb{R}} C \cdot \|x_1 - x_2\|$, where C is as in (7).

Proof: 1. Let C by as in (7) and assume $x_1 =_X x_2$, $y_1 =_{\mathbb{R}} y_2$ and $A_L(x_1,y_1)$. By (7) we get $\exists \widehat{y}^1 A_L(x_2, \widehat{y})$ and so by (9) $A_L(x_1 -_X x_2, y_1 -_{\mathbb{R}} \widehat{y})$. Again by (7) we get the existence of a $\tilde{y}^1$ such that

$$A_L(x_1 -_X x_2, \tilde{y}) \wedge |\tilde{y}| \leq_{\mathbb{R}} C \cdot \|x_1 -_X x_2\| =_{\mathbb{R}} 0.$$

Hence $\tilde{y} =_{\mathbb{R}} 0$ and so, by (8), $y_1 - \widehat{y} =_{\mathbb{R}} 0$, i.e. $\widehat{y} =_{\mathbb{R}} y_1 =_{\mathbb{R}} y_2$. Again by (8) this gives $A_L(x_2,y_2)$.
2. Let C be as in (7) and assume $A_L(x_1,y_1)$ and $A_L(x_2,y_2)$.
Then by (9) $A_L(x_1 -_X x_2, y_1 -_{\mathbb{R}} y_2)$. From (7) we get

$$\exists y^1 \, (A_L(x_1 -_X x_2, y) \wedge |y| \leq_{\mathbb{R}} C \cdot \|x_1 -_X x_2\|).$$

(8) now yields that $y_1 -_{\mathbb{R}} y_2 =_{\mathbb{R}} y$ and so $|y_1 -_{\mathbb{R}} y_2| \leq_{\mathbb{R}} C \cdot \|x_1 -_X x_2\|$. $\square$

Proposition 2.9. *$\mathcal{A}^\omega[X, \langle\cdot,\cdot\rangle, \mathcal{C}]$ proves the Riesz representation theorem in the following schematic form: if a formula $A(x^X, y^1)$ in $\mathcal{L}(\mathcal{A}^\omega[X, \langle\cdot,\cdot\rangle, \mathcal{C}])$ represents a bounded linear functional $L : X \to \mathbb{R}$, then there exists a point v^X such that*

$$L(x) =_{\mathbb{R}} \langle v, x\rangle \quad \textit{for all } x^X.$$

Proof: We can follow essentially the standard textbook proof:
Case 1: $L \equiv 0$, i.e. $\forall x^X, y^1 \, (A_L(x,y) \to y =_{\mathbb{R}} 0)$ (or – equivalently – $\forall x^X \, \exists y^1 \, (A_L(x,y) \wedge y =_{\mathbb{R}} 0)$). Then take $v := 0_X$.
Case 2: $\exists x_0^X \, (|L(x_0)| >_{\mathbb{R}} 0)$, i.e. $\exists x_0^X \, \exists y^1 \, (A_L(x_0,y) \wedge |y| >_{\mathbb{R}} 0)$. Reasoning in $\mathcal{A}^\omega[X, \langle\cdot,\cdot\rangle, \mathcal{C}]$ one easily shows that $Kern(L) := \{x \in X : L(x) =_{\mathbb{R}} 0\}$ is a closed linear subspace of X that is given by the formula

$$A_{Kern(L)}(x) :\equiv \forall y^1 \, (A_L(x,y) \to y =_{\mathbb{R}} 0).$$

By proposition 2.3.2 there exists a (unique) point $w \in Kern(L)$ such that

$$\|x_0 - w\| = \text{dist}(x_0, Kern(L)).$$

Now consider $v_0 := x_0 - w$. Then $|L(v_0)| =_{\mathbb{R}} |L(x_0)| >_{\mathbb{R}} 0$ and so $\|v_0\| >_{\mathbb{R}} 0$ by (7). By proposition 2.5 we have

$$\forall z \in Kern(L)\ (\langle v_0, z\rangle =_{\mathbb{R}} 0).$$

Define $v_1 := \frac{v_0}{\|v_0\|}$. Then again $\langle v_1, z\rangle =_{\mathbb{R}} 0$ for all $z \in Kern(L)$. Now put $a := L(v_1)$ and, finally, $v := a \cdot_X v_1$. Using that $L\big(L(x) \cdot_X v_1 -_X L(v_1) \cdot_X x\big) =_{\mathbb{R}} 0$ we obtain

$$0 =_{\mathbb{R}} \langle v_1, L(x) \cdot_X v_1 -_X L(v_1) \cdot_X x\rangle =_{\mathbb{R}} L(x) -_{\mathbb{R}} a \cdot \langle v_1, x\rangle,$$

i.e.

$$L(x) =_{\mathbb{R}} \langle a \cdot_X v_1, x\rangle =_{\mathbb{R}} \langle v, x\rangle$$

for all $x \in X$. □

Remark 2.10.

1. Definition 2.7 can be relativized to x^X being taken from a closed linear subspace $\mathcal{L}$ of X that is given via a formula $A_{\mathcal{L}}$ in the sense of definition 2.1.4. In this sense the previous proposition also applies to linear functionals $L : \mathcal{L} \to \mathbb{R}$.

2. In the following we will need the Riesz representation theorem for bounded linear functionals $L : \mathcal{L} \to \mathbb{R}$ on **separable** closed linear subspaces of X of the form
$$\mathcal{L} := \overline{\text{Lin}_{\mathbb{R}}\{x_n : n \in \mathbb{N}\}},$$
where (x_n) is a sequence in X.
In the case of separable Banach spaces it has been shown in [2] that the Riesz representation theorem can be proved from arithmetical comprehension over the weak base system RCA_0 used in reverse mathematics (see [28]). With the same proof we can establish the Riesz representation theorem for spaces of the form $\mathcal{L}$ above and bounded linear operators $L : \mathcal{L} \to \mathbb{R}$ that are given directly as a functional (rather than a representing formula) in $(\widehat{\text{PA}}^{\omega}\!\restriction + \text{AC}^{0,0}_{ar})[X, \langle\cdot,\cdot\rangle, \mathcal{C}]$. Here the crucial observation is that if (y_k) is dense in $\mathcal{L}$ (e.g. some standard enumeration of $\mathcal{L}_{\mathbb{Q}}$) and $y \in Kern(L)$ with $L(y) = 1$ (for $Kern(L) = \{0\}$ things are trivial), then the sequence defined by $w_k := y_k - L(y_k)\cdot y$ is dense in $Kern(L)$ (see [2], p.168) so that one can reason as in the proof of proposition 2.6 with w_n instead of y_n.

Theorem 2.11. *Let $\mathcal{T}^{\omega} := \widehat{\text{PA}}^{\omega}\!\restriction + \text{AC}^{0,0}_{ar}$. Then $\mathcal{T}^{\omega}[X, \langle\cdot,\cdot\rangle, \mathcal{C}]$ proves that the closed unit ball $B_1(0)$ in X is weakly sequentially compact, i.e. that for every sequence (x_n) in $B_1(0)$ there exists a point $v \in B_1(0)$ and a subsequence $(x_{n_k})_k$ of (x_n) such $\langle x_{n_k}, w\rangle$ converges to $\langle v, w\rangle$ for a every $w \in X$.*
Instead of $B_1(0)$ and with $\mathcal{T}^{\omega} := \text{PA}^{\omega} + \text{AC}^{0,0}_{ar}$ one may have any bounded closed convex subset $C \subset X$ that is given by a formula of $\mathcal{L}(\mathcal{T}^{\omega}[X, \langle\cdot,\cdot\rangle, \mathcal{C}])$.

Proof: Let (x_n) be a sequence in $B_1(0)$. Consider again the separable closed linear subspace of X

$$\mathcal{L} := \overline{\mathrm{Lin}_{\mathbb{R}}\{x_n : n \in \mathbb{N}\}}$$

and the countable dense subset

$$\mathcal{L}_{\mathbb{Q}} := \bigcup_{n\in\mathbb{N}} \{r_0x_0 + \ldots + r_nx_n : (r_0,\ldots,r_n) \in \mathbb{Q}^{n+1}\}$$

given by some primitive recursive (in (x_n)) standard enumeration (y_k). A bounded linear functional $L : \mathcal{L} \to \mathbb{R}$ can be recovered from its restriction to $\mathcal{L}_{\mathbb{Q}}$, i.e. from $L^- : \mathbb{N} \to \mathbb{R}$ defined by $L^-(k) := L(y_k)$. In this way we can represent such functionals that are bounded by 1 as points $z = (a_n)_n$ in $\prod_{n\in\mathbb{N}} [-\|y_n\|, \|y_n\|]$ that satisfy

$$(*)\ y_k =_X r_1 \cdot_X y_i +_X r_2 \cdot_X y_j \to a_k =_{\mathbb{R}} r_1a_i +_{\mathbb{R}} r_2a_j$$

for all $i,j,k \in \mathbb{N}$ and $r_1, r_2 \in \mathbb{Q}$ (see [6]).
Using arithmetical comprehension we can show in ACA_0 (see [6] or [28]) and hence in $\mathcal{T}^\omega$ that $\prod_{n\in\mathbb{N}} [-\|y_n\|, \|y_n\|]$ as well as its subset of points satisfying $(*)$ is sequentially compact, i.e. that every sequence in this space has a convergent subsequence w.r.t. the product metric on $\prod_{n\in\mathbb{N}} [-\|y_n\|, \|y_n\|]$ (the functional interpretation of this fact can be realized by functionals involving only primitive recursion at type 0 and bar recursion of lowest type $B_{0,1}$; see [27] for an explicit construction of the solution functionals).
We now consider the following sequence (L_n) linear functionals $\mathcal{L} \to \mathbb{R}$ that are all bounded by 1 :

$$L_n(x) := \langle x_n, x\rangle.$$

Let (L_n^-) be the corresponding sequence of points in $\prod_{n\in\mathbb{N}} [-\|y_n\|, \|y_n\|]$ that satisfy $(*)$ and $(L_{n_k}^-)$ a convergent subsequence with limit L^-. Clearly, L^- also represents a linear functional $L : \mathcal{L} \to \mathbb{R}$ that is bounded by 1.
By the Riesz representation theorem applied to $\mathcal{L}$ (instead of X) L is represented by an element $v \in \mathcal{L}$ (provably in $(\widehat{\mathrm{PA}}^\omega\!\upharpoonright + \mathrm{AC}^{0,0}_{ar})[X, \langle\cdot,\cdot\rangle, \mathcal{C}]$ by remark 2.10.2) i.e.

$$\forall w \in \mathcal{L}\,(L(w) =_{\mathbb{R}} \langle v, w\rangle).$$

Hence (for all $w \in \mathcal{L}$)

$$\langle x_{n_k}, w\rangle =_{\mathbb{R}} L_{n_k}(w) \stackrel{k\to\infty}{\longrightarrow} L(w) =_{\mathbb{R}} \langle v, w\rangle.$$

Now let $x \in X$ and apply proposition 2.6 to get the (by proposition 2.5 orthogonal) projection $x_L \in \mathcal{L}$ of x onto $\mathcal{L}$. Then

$$\langle x_{n_k}, x\rangle =_{\mathbb{R}} \langle x_{n_k}, x_L\rangle \text{ and } \langle v, x\rangle =_{\mathbb{R}} \langle v, x_L\rangle.$$

Hence

$$\langle x_{n_k}, x\rangle \stackrel{k\to\infty}{\longrightarrow} \langle v, x\rangle.$$

It is easy to see in $\widehat{\mathrm{PA}}^{\omega}\!\upharpoonright\![X,\langle\cdot,\cdot\rangle]$ that $\|v\| \le 1$: Suppose that $\|v\| = 1+\varepsilon$ for some $\varepsilon > 0$. Then

$$\langle x_{n_k}, v\rangle \stackrel{k\to\infty}{\longrightarrow} \langle v, v\rangle = (1+\varepsilon)^2 \ge 1+2\varepsilon,$$

whereas $\langle x_{n_k}, v\rangle \le \|x_{n_k}\|\cdot\|v\| \le 1+\varepsilon$ for all $k\in\mathbb{N}$ which is a contradiction.
If instead of $B_1(0)$ one has a bounded closed convex subset $C \subset X$ given by a formula of the language, then one can argue essentially as above with some K satisfying $K \ge \|x\|$ for all $x\in C$ replacing 1 and $\prod_{n\in\mathbb{N}}[-\|y_n\|,\|y_n\|]$ replaced by $\prod_{n\in\mathbb{N}}[-K\cdot\|y_n\|, K\cdot\|y_n\|]$
However, to show that $v \in C$ is somewhat more involved. Usually one applies a Hahn-Banach separation theorem here to show that closed convex sets are weakly closed but in our situation there is a more elementary proof of this fact via Mazur's theorem which implies that there is a sequence (z_n) of finite convex combinations of $\{x_{n_k} : k\in\mathbb{N}\}$ which strongly converges to v. Since, by the convexity of C, we have that $z_n\in C$ for all $n\in\mathbb{N}$, the closedness of C yields that $v\in C$. The proof of Mazur's theorem for Hilbert spaces, as given e.g. in [11], can easily be formalized in $(\mathrm{PA}^{\omega}+\mathrm{AC}^{0,0}_{ar})[X,\langle\cdot,\cdot\rangle,\mathcal{C}]$. □

Theorem 2.12 (F. Browder [5]). *Let X be a Hilbert space and $U : X\to X$ be a nonexpansive mapping. Assume that there exists a nonempty bounded closed convex subset $C\subset X$ such that U maps C into itself. For $v_0\in C$ and $t\in(0,1)$ let $U_t(x) := tU(x)+(1-t)v_0$ and u_t be the unique fixed point of this strict contraction. Then (u_t) converges strongly to a fixed point $p\in C$ of U as $t\to 1$.*

Remark 2.13. *The theorem by Browder states, furthermore, that p is the unique fixed point of U in C that is closest to v_0.*

Proposition 2.14. *Browder's proof of his theorem above can, for closed bounded convex subsets C that are given by a formula A_C in the sense of definition 2.1.4, be proved in $\mathcal{A}^{\omega}[X,\langle\cdot,\cdot\rangle,\mathcal{C}]$ (under the assumption that U has a fixed point).*

Proof: We just sketch how to formalize the proof here (a detailed explicit logical analysis of the proof – together with an extraction of χ in proposition 2.16 below – has to be devoted to another paper): we first note that the fact that we may use only the weak form of extensionality is no problem as nonexpansive mappings U trivially are extensional. The proof of Browder's theorem first proceeds ('lemma 1') by forming the projection u_0 of v_0 onto the (nonempty) set F of all fixed points of U in C. It is an easy consequence of the uniform convexity of X that F is convex. Moreover, F obviously is closed and given by the formula $A_F(u^X) :\equiv (A_C(u^X)\wedge U(u) =_X u)$. Hence we can apply proposition 2.3.2 to establish the existence of $u_0\in F$.
A second – perfectly elementary – 'lemma 2' shows that every weak limit of an approximate fixed point sequence is a fixed point of U. The proof is then concluded by showing that for any sequence $(k_j)_{j\in\mathbb{N}}$ with $k_j\to 1$ one can find a subsequence of

(u_{k_j}) converging strongly to u_0 (in this proof it is used that (u_{k_j}) always possesses a weakly convergent subsequence whose weak limit $v \in C$ by 'lemma 2' even is in F). With theorem 2.11 in place, it is easily verified that the (nontrivial but – modulo the weak compactness – elementary) proof of this formalizes in our formal framework which we skip here. Thus it only remains to formalize the proof that the latter fact implies that u_t tends to u_0 for $t \to 1$. Let (k_j) be a sequence with $\lim_{j\to\infty} k_j = 1$ and assume that for some $\varepsilon > 0$

$$\forall n \exists j > n\, (\|u_{k_j} - u_0\| >_{\mathbb{R}} \varepsilon).$$

Then Σ^0_1-$\mathrm{AC}^{0,0}$ (and hence QF-$\mathrm{AC}^{0,0}$) yields a function g such that

$$\forall n\, (g(n) > n \wedge \|u_{k_{g(n)}} - u_0\| >_{\mathbb{R}} \varepsilon).$$

Then $k_{g(n)} \overset{n\to\infty}{\to} 1$, but $(u_{k_{g(n)}})_n$ does not contain a subsequence that converges to u_0. □

Remark 2.15. From a well-known theorem of Browder-Göhde-Kirk it follows that U always has fixed points. However, for our application of the logical metatheorem (corollary 6.8 in [9]) below we do not have to consider the proof of this theorem but only the proof relative to the assumption of a fixed point. In fact, the logical metatheorem we will use below allows one to convert this proof into a new one which only uses the (trivial) existence of approximate fixed points so that the need to use the Browder-Göhde-Kirk theorem disappears altogether.

Theorem 2.16. *Under the assumptions of theorem 2.12 with $C := B_1(0)$ there exists a computable functional $\chi : \mathbb{N} \times \mathbb{N}^{\mathbb{N}} \to \mathbb{N}$ (that is independent from X, U and $v_0 \in B_1(0)$) such that*

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

where $x_i := u_{t_i}$ with $t_i := 1 - \frac{1}{i+1}$ and $[n; n+m] := \{n, n+1, \ldots, n+m\}$.
Similarly, for any sequence (t_n) in $(0,1)$ that converges towards 1 where then the bound depends also on a (majorant of) rate of metastability of that convergence

$$\forall n \in \mathbb{N}\, \forall g \in \mathbb{N}^{\mathbb{N}}\, \forall i \in [\chi(g,n); \chi(g,n) + g(\chi(g,n))]\, (|1 - t_i| \leq \frac{1}{n+1})$$

and a function $h : \mathbb{N} \to \mathbb{N}$ such that $\forall n \in \mathbb{N}\, (t_n \leq 1 - \frac{1}{h(n)+1})$.

Proof: We first note that x_i can be explicitly defined as a functional in v_0, U and i in the language of $\mathcal{A}^\omega[X, \langle\cdot,\cdot\rangle, \mathcal{C}]$ using the Picard iteration of U and the completion operator C.
By proposition 2.14 we have

$$\mathcal{A}^\omega[X, \langle\cdot,\cdot\rangle, \mathcal{C}] \vdash \left\{ \begin{array}{l} \forall v_0 \in B_1(0) \forall U : X \to X\, \forall k \in \mathbb{N}\, (U \text{ n.e. } \wedge\ U(B_1(0)) \subseteq B_1(0) \\ \wedge\, Fix(U) \neq \emptyset \to \exists n \in \mathbb{N}\, \forall m \in \mathbb{N}\, (\|x_n - x_m\| <_{\mathbb{R}} 2^{-k})) \end{array} \right.$$

and hence

$$\mathcal{A}^\omega[X,\langle\cdot,\cdot\rangle,\mathcal{C}]\vdash$$
$$\begin{cases} \forall v_0\in B_1(0)\,\forall U:X\to X\,\forall k\in\mathbb{N}\,\forall g:\mathbb{N}\to\mathbb{N}\,(U \text{ n.e. } \wedge U(B_1(0))\subseteq B_1(0) \\ \wedge Fix(U)\neq\emptyset \to \exists n\in\mathbb{N}\,\forall i,j\in[n;n+g(n)]\,(\|x_i-x_j\| <_{\mathbb{R}} 2^{-k})), \end{cases}$$

where

$$\forall i,j\in[n;n+g(n)]\,(\|x_i-x_j\|<2^{-k})$$

is equivalent to a Σ^0_1-formula over already $\mathcal{T}^\omega[X,\langle\cdot,\cdot\rangle]$ with

$$\mathcal{T}^\omega := \widehat{\mathrm{PA}}^\omega\!\upharpoonright + \ \text{QF-AC}^{0,0}$$

since only the bounded collection principle for Σ^0_1-formulas is needed.
We note that the condition '$U(B_1(0))\subseteq B_1(0)$' can be written as a $\forall$-formula

$$\forall x^X\,(\|U(\check{x})\|\leq_{\mathbb{R}} 1),\ \text{where } \check{x} := \frac{x}{\max(1,\|x\|)}.$$

The claim of the theorem now follows immediately from corollary 6.8 in [9] (together with the treatment of completeness conditions from [18], pages 433-434). Here one takes v_0 in place of z and C is being treated trivially as the whole space X (i.e. by adding a universal axiom $\forall x^X\,(\chi_C(x^X)=_0 0)$), where $c_X := 0_X$, so that we can take $b:=2$. □

Remark 2.17. Since (as discussed already in remark 2.15) the above proof does not use anymore the existence of a fixed point of U, it can be used in itself as an alternative proof of the existence of a fixed point: since (x_n) satisfies the no-counterexample version of the Cauchy property it is a Cauchy sequence and its limit clearly must be a fixed point of U.

Remark 2.18. An issue to be devoted to further research is whether the closed convex set C can be treated completely abstract (via a characteristic function as in corollary 6.8 from [9]) which would mean that (in contrast to closed convex sets given by a formula in the sense of definition 2.1.4) one is allowed to use only a quantifier-free rule of extensionality

$$\frac{A_0\to s=_X t\wedge s\in C}{A_0\to t\in C}\quad (A_0 \text{ quantifier-free})$$

rather than the full extensionality condition from definition 2.1.1.(a). If the rule suffices one could extract an effective bound χ that would not depend on C except for some norm upper bound on $M\geq\|v\|$ for all $v\in C$.

Comments added in proof: Subsequent work ([19]) that we have carried out after the present paper was finished sheds new light on the logical status of Browder's proof as well as the theorem as such:

1. As a final **result** of the actual logical analysis of Browder's proof, the use of weak compactness in the end disappears and the extraction process yields even a primitive recursive (in the sense of Kleene) bound on the metastable version of Browder's theorem (see [19] for a discussion of this phenomenon). In fact, the extraction works for arbitrary bounded closed convex subsets $C \subset X$ (instead of $B_1(0)$) and the bound depends on C only via an upper bound on the diameter of C.

2. A different proof of Browder's theorem that avoids already weak compactness from the beginning can be obtained from Halpern [10]. The – much simpler – logical analysis of this proof again yields a primitive recursive rate of metastability (see again [19]).

References

[1] Avigad, J., Gerhardy, P., Towsner, H., Local stability of ergodic averages. To appear in: Trans. Amer. Math. Soc.

[2] Avigad, J., Simic, K., Fundamental notions of analysis in subsystems of second-order arithmetic. Ann. Pure Appl. Logic **139**, pp. 138-184 (2006).

[3] Briseid, E.M., A rate of convergence for asymptotic contractions. J. Math. Anal. Appl. **330**, pp. 364-376 (2007).

[4] Briseid, E.M., Fixed points of generalized contractive mappings. Journal of Nonlinear and Convex Analysis **9**, pp. 181-204 (2008).

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

[6] Brown, D.K., Functional analysis in weak subsystems of second order arithmetic. PhD Thesis. The Pennsylvania State University. University Park, PA (1987).

[7] Gerhardy, P., A quantitative version of Kirk's fixed point theorem for asymptotic contractions. J. Math. Anal. Appl. **316**, pp. 339-345 (2006).

[8] Gerhardy, P., Proof mining in topological dynamics. Notre Dame Journal of Formal Logic **49**, pp. 431-446 (2008).

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

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

[11] Hunter, J.K., Applied Analysis. World Scientific,439 pp. (2001).

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

[13] Kohlenbach, U., A quantitative version of a theorem due to Borwein-Reich-Shafrir. Numer. Funct. Anal. and Optimiz. **22**, pp. 641-656 (2001).

[14] Kohlenbach, U., Uniform asymptotic regularity for Mann iterates. J. Math. Anal. Appl. **279**, pp. 531-544 (2003).

[15] Kohlenbach, U., Some logical metatheorems with applications in functional analysis. Trans. Amer. Math. Soc. vol. 357, no. 1, pp. 89-128 (2005).

[16] Kohlenbach, U., Some computational aspects of metric fixed point theory. Nonlinear Analysis **61**, pp. 823-837 (2005).

[17] Kohlenbach, U., A logical uniform boundedness principle for abstract metric and hyperbolic spaces. Electronic Notes in Theoretical Computer Science (Proceedings of WoLLIC 2006) **165**, pp. 81-93 (2006).

[18] Kohlenbach, U., Applied Proof Theory: Proof Interpretations and their Use in Mathematics. Springer Monographs in Mathematics. xx+536pp., Springer Heidelberg-Berlin, 2008.

[19] Kohlenbach, U., On quantitative versions of a theorem due to F.E. Browder. Submitted (2009).

[20] Kohlenbach, U., Lambov, B., Bounds on iterations of asymptotically quasi-nonexpansive mappings. In: G-Falset, J., L-Fuster, E., Sims, B. (eds.), Proc. International Conference on Fixed Point Theory, Valencia 2003, pp. 143-172, Yokohama Press 2004.

[21] Kohlenbach, U., Leuştean, L., Mann iterates of directionally nonexpansive mappings in hyperbolic spaces. Abstract and Applied Analysis, vol. 2003, no.8, pp. 449-477 (2003).

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

[23] Kohlenbach, U., Leuştean, L., Asymptotically nonexpansive mappings in uniformly convex hyperbolic spaces. J. European Math. Soc. **12**, pp. 71-92 (2010).

[24] Leuştean, L., Proof mining in $\mathbb{R}$-trees and hyperbolic spaces. Electronic Notes in Theoretical Computer Science (Proceedings of WoLLIC 2006) **165**, pp. 95-106 (2006).

[25] Mints, G.E., Unwinding a non-effective cut elimination proof. In: Grigoriev, D., Harrison, J., Hirsch, E.A. (Eds.): Computer Science - Theory and Applications, First International Computer Science Symposium in Russia, CSR 2006, St. Petersburg, Russia, June 8-12, 2006, Proceedings. Springer LNCS **3967**, pp. 259-269 (2006).

[26] Mints, G.E., Proof search tree and cut elimination. In: Avron, A., Dershowitz, N., Rabinovich, A. (eds.), Pillars of Computer Science 2008, Springer LNCS 4800, pp. 521-536 (2008).

[27] Safarik, P., Kohlenbach, U., The Interpretation of the Bolzano-Weierstraß Principle Using Bar Recursion. Preprint 2009.

[28] Simpson, S.G., Subsystems of Second Order Arithmetic. Perspectives in Mathematical Logic. Springer-Verlag. xiv+445 pp. 1999.

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

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