UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS

Proof mining with the bounded functional interpretation

" Documento Definitivo"

Doutoramento em Matemática

Especialidade de Álgebra, Lógica e Fundamentos

Pedro Pinto

Tese orientada por:

Prof. Doutor Fernando Jorge Inocêncio Ferreira

Documento especialmente elaborado para a obtenção do grau de doutor

2019

UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS

Proof mining with the bounded functional interpretation

Doutoramento em Matemática

Especialidade de Álgebra, Lógica e Fundamentos

Pedro Pinto

Tese orientada por:

Prof. Doutor Fernando Jorge Inocêncio Ferreira

Júri:

Presidente:

- Doutora Maria Teresa de Lemos Monteiro Fernandes, Professora Catedrática e Presidente do Departamento de Matemática, Faculdade de Ciências da Universidade de Lisboa

Vogais:

- Doutor Paulo Borges Oliva, Reader in Mathematical Logic, School of Engineering and Computer Science Queen Mary da University of London, Londres
- Doutor Daniel Silva Graça, Professor Auxiliar com Agregação, Faculdade de Ciências e Tecnologia da Universidade do Algarve
- Doutora Gilda Maria Saraiva Dias Ferreira, Professora Auxiliar, Departamento de Ciências e Tecnologia da Universidade Aberta
- Doutor Fernando Jorge Inocêncio Ferreira, Professor Catedrático, Faculdade de Ciências da Universidade de Lisboa, Orientador
- Doutor Pedro Miguel Nunes da Rosa Dias Duarte, Professor Associado com Agregação, Faculdade de Ciências da Universidade de Lisboa

Documento especialmente elaborado para a obtenção do grau de doutor

Suportado por FCT – Fundação para a Ciência e a Tecnologia

2019

Acknowledgements

I want to express my deep gratitude to my advisor Prof. Dr. Fernando Ferreira. I am very grateful for all of his guidance and valuable suggestions, for always finding the time and patience to answer my questions and for all the encouragement throughout my time as a PhD student. I am also very grateful to Prof. Dr. Laurenţiu Leuştean, who hosted me in Bucharest in February 2018, where most of the work in Chapter 6 was carried out, for his valuable insight into the actual practice of proof mining. I would like to thank my friend and colleague Bruno Dinis for all the interesting and lengthy discussions on proof mining and for pushing me to finish this thesis.

This work would have been impossible without financial support and I thank FCT for the doctoral grant PD/BD/52645/2014, under the program Lisbon Mathematics PhD, and to CMAFcIO under the FCT project UID/MAT/04561/2019.

Finally, to Andreia, for all of her patience, constant support and for always making everything easier, I give a special and loving thank you.

Pedro Pinto
Lisboa, June 17, 2019

Abstract

In this doctoral thesis, we will see how the bounded functional interpretation of Ferreira and Oliva [13] can be used and contribute to the Proof Mining program, a program which aims to extract computational information from mathematical theorems using proof-theoretic techniques. We present a method for the elimination of sequential weak compactness arguments from the quantitative analysis of certain mathematical results. This method works as a "macro" and allowed us to obtain quantitative versions of important results of F. E. Browder [6], R. Wittmann [51] and H. H. Bauschke [2] in fixed point theory in Hilbert spaces. Although Browder's and Wittmann's theorems were previously analyzed by Kohlenbach using the monotone functional interpretation, it was not clear why such analyses did not require the use of functionals defined by bar recursion. This phenomenon is now fully understood, by a theoretical justification for the elimination of sequential weak compactness in the context of the bounded functional interpretation. Bauschke's theorem is an important generalization of Wittmann's theorem and its original proof is also analyzed here. The analyses of these results also required a quantitative version of a projection argument which turned out to be simpler when guided by the bounded functional interpretation than when using the monotone functional interpretation.
In the context of the theory of monotone operators, results due to Boikanyo/Moroşanu [5] and Xu [52] for the strong convergence of variants of the proximal point algorithm were analyzed and bounds on the metastablility property of these iterations obtained.
These results are the first applications of the bounded functional interpretation to the proof mining of concrete mathematical results.

Keywords: bounded functional interpretation, majorants, metastability, weak compactness, fixed points

Resumo

Nesta tese de doutoramento, iremos ver como a interpretação funcional limitada de Ferreira e Oliva [13] pode contribuir para o programa de *Proof Mining*, um programa que tem como objetivo a extração de nova informação de teoremas matemáticos usando ferramentas da teoria da demonstração. Apresentamos um método para a eliminação de argumentos de compacidade fraca sequencial da análise quantitativa de certos resultados matemáticos. Este método funciona como uma "macro" muito geral e a sua aplicação permitiu obter versões quantitativas de importantes resultados de F. E. Browder [6], R. Wittmann [51] e H. H. Bauschke [2] no âmbito da teoria dos pontos fixos em espaços de Hilbert. Apesar de análises dos teoremas de Browder e Wittmann já terem sido anteriormente obtidas por Kohlenbach usando a interpretação funcional monótona, não era claro o motivo por que em tais análises não é necessária a utilização de funcionais definidos por *bar recursion*. A justificação da eliminação de compacidade fraca sequencial no contexto da interpretação funcional limitada vem por completo clarificar este fenómeno. O teorema de Bauschke é uma importante generalização do teorema de Wittmann e a sua demonstração original foi também aqui analisada. A análise destes resultados levou-nos também a uma versão quantitativa de um argumento de projeção que se veio a revelar mais simples quando guiada pela interpretação funcional limitada do que seguindo a interpretação funcional monótona.
Já no contexto da teoria dos operadores monótonos, foram analisados resultados devidos a Boikanyo/Moroşanu [5] e a Xu [52] relativamente à convergência forte de variantes do *proximal point algorithm* tendo-se obtido majorações na propriedade metaestável desses algoritmos. Estes resultados constituem os primeiros exemplos de aplicação da interpretação funcional limitada no *proof mining* de resultados matemáticos concretos.

Palavras-chave: interpretação funcional limitada, majoração, metaestabilidade, compacidade fraca, pontos fixos

Resumo alargado

Nesta tese de doutoramento, iremos ver como a interpretação funcional limitada de Ferreira e Oliva [13] pode contribuir para o programa de *Proof Mining*, um programa que tem como objetivo a extração de nova informação de demonstrações matemáticas. Para esse efeito, são usadas ferramentas da teoria da demonstração, nomeadamente interpretações funcionais sendo a interpretação funcional monótona de Kohlenbach [21] a mais frequentemente utilizada. Enquanto que outras interpretações procuram extrair informações computacionais exatas, tanto a interpretação monótona como a limitada tentam extrair majorantes. Esta distinção é essencial para permitir a análise de uma maior classe de resultados matemáticos. Naturalmente, a extração de majorantes exige uma noção adequada de majoração e, com esse propósito, estas interpretações consideram a relação de majoração forte de Bezem [3]. No entanto, enquanto que a interpretação monótona apenas relaxa a extração na fase final da interpretação, a interpretação limitada combina a majoração com a própria tradução das fórmulas. Isto levanta a questão de saber se, usando a interpretação funcional limitada, poderão existir situações em que a informação computacional extraída possa ser diferente ou possa ser obtida de forma mais eficiente. O trabalho desenvolvido por Patrícia Engrácia na sua tese de doutoramento preparou o caminho para o uso desta interpretação em casos concretos de *Proof Mining*. Esta tese de doutoramento apresenta essas primeiras aplicações práticas.

Os teoremas aqui analisados focam-se na convergência forte de certas iterações (x_n) num espaço de Hilbert. Na sua análise quantitativa, olhamos para a afirmação equivalente de que (x_n) é uma sucessão de Cauchy,

$$\forall k \in \mathbb{N}\, \exists N \in \mathbb{N}\, \forall i, j \geq N \left(\|x_i - x_j\| \leq \frac{1}{k+1} \right),$$

que, do ponto de vista lógico, tem uma complexidade mais simples de que a afirmação da convergência. Em geral, não é possível extrair informação computacional para "$\exists N$" na propriedade de Cauchy. Em vez disso, olhamos para a versão metaestável,

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

que corresponde à interpretação da propriedade de Cauchy e que designamos de *propriedade metaestável* de (x_n) – no sentido de Terence Tao [46]. Aqui o intervalo $[N, fN]$ denota o conjunto $\{N, N+1, \cdots, fN\}$.
Estas duas propriedades são (de forma não-efetiva) equivalentes mas, enquanto que a propriedade de Cauchy é dada por uma fórmula $\forall\exists\forall$, a propriedade metaestável é (vista como) uma afirmação $\forall\exists$. Para fórmulas desta complexidade, o *soundness* da interpretação garante que, de uma demonstração de que (x_n) é uma sucessão de Cauchy, conseguimos extrair um majorante para N que apenas depende de k e f. Isto é, a análise quantitativa de tais demonstrações permite a extração de uma função $\phi : \mathbb{N} \times \mathbb{N}^{\mathbb{N}} \to \mathbb{N}$ tal que

$$\forall k \in \mathbb{N}\, \forall f \in \mathbb{N}^{\mathbb{N}}\, \exists N \leq \phi(k, f)\, \forall i, j \in [N, fN] \left(\|x_i - x_j\| \leq \frac{1}{k+1} \right).$$

As versões quantitativas dos resultados analisados nesta tese são na forma de extrações explícitas de funções ϕ como acima, i.e. majorações na propriedade metaestável de uma iteração (x_n) e, no final, a maquinaria lógica subjacente aparece oculta, sendo por isso os resultados quantitativos passíveis de ser apreciados tanto por lógicos como por não-lógicos.

No âmbito da teoria dos pontos fixos, iremos obter versões quantitativas de importantes teoremas devidos a F. E. Browder [6] e a R. Wittmann [51]. Análises dos teoremas de Browder e de Wittmann já tinham sido anteriormente realizadas por Kohlenbach usando a interpretação funcional monótona. Apresentamos também uma versão quantitativa de uma generalização do teorema de Wittmann devida a H. H. Bauschke [2]. Apesar de uma posterior generalização do teorema de Bauschke ter sido analisada na tese doutoral de Daniel Körnlein [35], aqui vamos focar-nos pela primeira vez na demonstração original de Bauschke. Dado que as demonstrações destes resultados se centram num argumento de compacidade fraca sequencial, não era claro o motivo por que tais análises não exigem a utilização de funcionais definidos por *bar recursion* [43], como seria expectável (de acordo com a teoria lógica). Neste sentido, as análises anteriores dos teoremas de Browder e de Wittmann tinham uma explicação teórica incompleta. A justificação da eliminação de compacidade fraca sequencial no contexto da interpretação funcional limitada vem por completo clarificar este fenómeno.

Apresentamos um método geral para a eliminação de argumentos de compacidade fraca sequencial da análise quantitativa de certos resultados matemáticos. Este método funciona como uma macro que pode ser aplicada em situações muito gerais para eliminar compacidade fraca sequencial em resultados quantitativos. Este resultado permitiu obter as versões quantitativas dos teoremas de Browder, de Wittmann e de Bauschke.
Seja X um espaço de Hilbert, C um subconjunto limitado, u_0 um ponto em C e $U : X \to X$ uma função não-expansiva (i.e. $\|U(x) - U(y)\| \leq \|x - y\|$, para $x, y \in X$) com $U[C] \subseteq C$. Nos casos em que nos focamos, a compacidade fraca sequencial é necessária para concluir

que $\limsup\langle\tilde{x}-u_0,\tilde{x}-x_n\rangle \leq 0$, onde $\tilde{x}$ é o ponto de projecção de u_0 sobre o conjunto dos pontos fixos de U em C. Significa portanto que para um certo ponto fixo $\tilde{x}$ se tem

$$\forall k \in \mathbb{N}\, \exists N \in \mathbb{N}\, \forall n \geq N \left(\langle \tilde{x}-u_0, \tilde{x}-x_n \rangle \leq \frac{1}{k+1} \right).$$

Em vez deste facto, mostramos que para quaisquer $k \in \mathbb{N}$ e função $f : \mathbb{N} \to \mathbb{N}$ existem $N \in \mathbb{N}$ e $x \in C$ tais que

$$\|U(x)-x\| \leq \frac{1}{f(N)+1} \text{ e } \forall n \in [N, fN] \left(\langle x-u_0, x-x_n \rangle \leq \frac{1}{k+1} \right).$$

Mostramos que esta afirmação é já suficiente para concluir os resultados de convergência que analisamos. O facto notável é que este resultado pode ser demonstrado usando um argumento de coleção característico da interpretação funcional limitada e evitando o uso de compacidade fraca sequencial. Esse argumento pode ser visto como uma aplicação de compacidade Heine-Borel, nomeadamente

$$\forall x \in C\, \exists k \in \mathbb{N}\ (x \in \Omega_k) \to \exists n \in \mathbb{N}\, \forall x \in C\, \exists k \leq n\ (x \in \Omega_k),$$

onde (Ω_n) é uma família contável de conjuntos abertos. Em suma, conclui-se que, em certas circunstâncias, o uso de compacidade Heine-Borel, que em geral é uma propriedade falsa do espaço, é uma maneira perfeitamente aceitável de mostrar a convergência de iterações em espaços de Hilbert.

A interpretação destes resultados exigiu também a análise quantitativa de um argumento de projeção. Um ponto interessante é que a interpretação funcional limitada da projeção é significativamente mais simples que a análise mais elaborada obtida anteriormente pela interpretação funcional monótona em [27].
O argumento de projeção pode ser escrito na forma

$$\exists x \in C\, \forall k \in \mathbb{N} \left[U(x) = x \wedge \forall y \in C \left(U(y) = y \to \|x-u_0\| \leq \|y-u_0\| + \frac{1}{k+1} \right) \right].$$

No entanto, a demonstração deste resultado requer o uso de escolha contável o que, em termos da informação que pode ser obtida, significa que uma vez mais teríamos de considerar funcionais definidos por *bar recursion*. Acontece que o enfraquecimento

$$\forall k \in \mathbb{N}\, \exists x \in C \left[U(x) = x \wedge \forall y \in C \left(U(y) = y \to \|x-u_0\| \leq \|y-u_0\| + \frac{1}{k+1} \right) \right],$$

que se demonstra com um simples argumento indutivo, é já suficiente para concluir que (x_n) é uma sucessão de Cauchy. Neste caso, a informação obtida é da forma de funcionais recursivos primitivos no sentido de Gödel (sendo por isso mais simples e desejável) e será suficiente para obter um majorante na propriedade de metaestabilidade de (x_n).

Já no contexto da teoria dos operadores monótonos, foram analisados resultados devido a Boikanyo e Moroşanu em [5] e devido a Xu em [52] sobre a convergência forte de uma variante do *proximal point algorithm*. A variante que aqui nos interessa é o designado *Halpern type proximal point algorithm*,

$$x_{n+1} := \alpha_n u + (1 - \alpha_n) J_{\beta_n}(x_n),$$

onde, para $\beta > 0$, J_β é a função resolvente de um operador monótono maximal, $(\alpha_n) \subset [0, 1]$, u é um ponto do espaço e $(\beta_n) \subset \mathbb{R}^+$.
Em ambos os casos, foram obtidas majorações na propriedade de metaestabilidade sendo que na versão quantitativa do teorema de Boikanyo e de Moroşano mostrámos ser possível converter uma majoração ϕ_{Br} para a versão metaestável do teorema de Browder numa majoração ϕ para a metaestabilidade da iteração.

Os resultados presentes nesta tese constituem os primeiros exemplos de aplicação da interpretação funcional limitada no *proof mining* de resultados matemáticos concretos.

Contents

Chapter 1

Introduction

In the 1950's, George Kreisel advocated the concept of "*unwinding of proofs*", under the idea that new useful information was hidden inside mathematical proofs. The validity of a mathematical theorem (or the validity of its proof) was only a fraction of the information given by mathematical reasoning. He proposed that such new mathematical content could be obtained by the use of proof-theoretic methods which were first developed in the context of Hilbert's consistency program. This idea was developed by Ulrich Kohlenbach and his collaborators over the last twenty five years and applied in a systematic way to various areas of mathematics, specially in the field of nonlinear analysis. The term "*proof mining*" – suggested by Dana Scott – eventually replaced "unwinding of proofs" and nowadays refers to the modern program focused on the analysis of mathematical proofs based on Kreisel's ideas. For a good overview of Kreisel's original program of "*unwinding of proofs*" and how it shaped the current proof mining enterprise, see [30].

The main technique used in proof mining are functional interpretations. In essence, a functional interpretation is a mapping from a certain formal theory $\mathcal{A}$ into a formal theory $\mathcal{B}$ that gives a recursive translation of formulas and of proofs in $\mathcal{A}$ to formulas and proofs in $\mathcal{B}$. This is done in such a way that theorems and corresponding proofs are associated between $\mathcal{A}$ and $\mathcal{B}$ by the interpretation (soundness). Furthermore, these interpretations disclose information that is hidden behind the use of quantifiers and which cannot be read directly from the original proof. Such new information can then be effectively *extracted* from the translated proof. Naturally, this technique allows for (relative) consistency results: if lower complexity formulas are left unchanged namely, having the formula $\perp$ fixed by the interpretation, we see that if $\mathcal{B}$ is consistent then so must be $\mathcal{A}$. In fact, functional interpretations were first introduced by Kurt Gödel in 1958 with his *Dialectica* interpretation [16] as a way to show the consistency of PA relative to the consistency of his system T of primitive recursive functionals of finite type. Also noteworthy was Clifford Spector extension with bar recursive functionals [43] as a way to prove the consistency of full second-order arithmetic, i.e. the consistency of analysis.

Introduced in 1996 by Kohlenbach, the standard such functional interpretation used in proof mining is the so-called monotone functional interpretation [21]. The translation of formulas and proofs is the same as in the *Dialectica* interpretation but, in the last step, the focus is placed, not on exact witnessing terms, but instead on majorants for those terms. This requires an appropriate notion of majorant that combines well with the interpretation, and Howard's [19] (or Bezem's [3]) majorizability relation is used here. This weakening in the goal of the functional interpretation may seem unproductive, yet this shift from precise witnesses to majorants is of essential relevance in the proof mining practice. By only requiring majorizing terms, the monotone functional interpretation can deal with many non-constructive principles, in contrast with the *Dialectica* in which any auxiliary lemma must be restricted to an universal formula. Most notably, it allows one to consider proofs that make use of Weak Kőnig's Lemma, WKL. Thus, e.g. many mathematical proofs based on Heine-Borel compactness can be analyzed using this inexact version of *Dialectica*. Another important feature for the analysis of mathematical proofs was the introduction of typed formal systems with an additional ground type X for abstract (metric, normed, Hilbert, hyperbolic, CAT(0), etc.) spaces. With this additional type, one is no longer restricted to working with "computable" spaces only and it makes sense to extract computable bounds for abstract spaces. In [15][22], general logic metatheorems were proved that guarantee that the existence of (uniform) bounds can be obtained from a large class of theorems and proofs. For a deeper look into the proof mining program see [33], [24] and [26], more recently [29] and [31], and Kolenbach's book [25].

In 2005, a different interpretation, which gives a completely new translation of formulas, was introduced by Fernando Ferreira and Paulo Oliva, the bounded functional interpretation [13]. This interpretation relies on Bezem's majorizability notion and this time the search for bounds is considered at every step of translation. In fact, the bounded functional interpretation uses an *intensional* variant of Bezem's majorizability in the sense that it is partly governed by a rule. In contrast with the monotone interpretation, which only considers the majorizability notion after the interpretation of formulas, here the majorizability is infused into the interpretation. This is problematic as the usual majorizability is not given by a quantifier-free formula (hence, it is not computationally empty) and would require additional information to ensure a soundness theorem. As it is well-known, functional interpretations extract less information from rules than from provable implications, e.g. the *Dialectica* cannot interpret the extensionality axiom but has no problem with an extensionality rule. In a similar way, by working in an intensional setting using a majorizability partly governed by a rule, we essentially deactivate the interpretation at the level of the majorizability predicates. In her PhD thesis [8], Patrícia Engrácia considered the bounded functional interpretation with new abstract types X (for the case of normed real spaces). This generalization laid the groundwork for future applications of the interpretation to analyses of concrete mathematical proofs, applications which were until now absent. In the bounded interpretation, the focus is placed entirely on the majorants and bounded data is considered to be computationally empty. By considering majorants at an early stage of the interpretation, one may wonder if there are instances where the use of this technique happens to produce final bounds in a clear and faster way than the monotone functional interpretation.

This doctoral thesis is centered on the practical application of the bounded functional interpretation to the analysis of mathematical proofs and the main goal is to show that it is a valuable and helpful technique to the proof mining program. This will be achieved by doing quantitative analyses of some mathematical results while having the bounded functional interpretation as an underlying guiding principle. We will look at proofs based on classical logic and, instead of considering a negative translation, we will be using a Shoenfield-like bounded functional interpretation which operates directly in classical systems, introduced in [11]. Another main point of this work is that these analyses can be carried out in restricted formal theories such that the extracted information is given by primitive recursive functionals in the sense of Gödel, avoiding the use of Spector's bar recursive functionals.

By suggestion of Kohlenbach, the first proof analyzed by the bounded functional interpretation was of the following theorem due to F. E. Browder,

Theorem 1.1 (Browder [6]). *Let X be a real Hilbert space and $U : X \to X$ a nonexpansive mapping. Assume that C is a bounded closed convex subset of X, that $v_0 \in C$, and that U maps C into itself. For each natural number n, define*

$$U_n(x) := \left(1 - \frac{1}{n+1}\right) U(x) + \frac{1}{n+1} v_0 \tag{1.1}$$

and consider u_n to be the unique fixed point of this strict contraction. Then the sequence (u_n) converges strongly to a fixed point of U in C (the closest one to v_0).

A previous quantitative analysis of this result was already carried out by Kohlenbach using the monotone functional interpretation [27] and the idea was to do a step-by-step comparison between the use of the monotone and the bounded functional interpretations.
Browder's original argument begins by seeing that the set of fixed points is non-empty. While the proof used Zorn's lemma (which corresponds to an application of choice) in the context of functional interpretations one can simply add a new constant to the language and the universal axiom saying that it denotes a fixed point. Furthermore, in the context of the bounded functional interpretation, this is not even necessary. From the easy fact that U has almost fixed points, i.e.

$$\forall k \in \mathbb{N} \exists x \in C \left(\|x - U(x)\| \leq \frac{1}{k+1} \right),$$

it immediately follows that the theory contains one such fixed point. In section 4.2.1, this is explained as a simple application of a collection argument. In general, by being able to introduce "uniformities" in the theory, this type of reasoning by collection has the potential to avoid the need of these type of *ideal* elements whose justification may sometimes be problematic.

The proof then considers a projection argument of v_0 over the set of fixed points of U:

$$\exists x \in C \ (U(x) = x \wedge \forall y \in C \ (U(y) = y \to \|x - v_0\| \leq \|y - v_0\|)),$$

which can be equivalently written as

$$\exists x \in C \forall k \in \mathbb{N} \left[U(x) = x \wedge \forall y \in C \left(U(y) = y \to \|x - v_0\| \leq \|y - v_0\| + \frac{1}{k+1} \right) \right].$$

As commented by Kohlenbach in [27], the following weaker statement (what Kohlenbach called the "ε-version") is already sufficient to carry out Browder's theorem:

$$\forall k \in \mathbb{N} \exists x \in C \left[U(x) = x \wedge \forall y \in C \left(U(y) = y \to \|x - v_0\| \leq \|y - v_0\| + \frac{1}{k+1} \right) \right].$$

While this weaker $\forall\exists$ statement can be shown using induction only, the original $\exists\forall$ projection argument required countable choice. The complexity of the extracted information follows directly from the strength of the logical principles required for the proof. In this case, by considering this weaker statement, the extracted bounds are defined by recursion in Gödel's T and we can avoid the use of bar recursive functionals, which one would need to interpret the original projection statement.

The final piece in Browder's proof is a sequential weak compactness argument. As before, the interpretation of this step needs the use of bar recursion. Nevertheless, in the quantitative version, as analyzed by Kohlenbach, Browder's theorem didn't require bar recursive functionals and, in fact, the weak compactness argument is absent.

There are two relevant questions in the analysis of this result. First, 'was there any significant difference between the use of the monotone and the bounded functional interpretations?'. Second, since Browder's proof made use of strong principles (namely sequential weak compactness) for which the interpretation is only guaranteed by Spector's bar recursive functionals, 'why were the extracted information given by primitive recursive functional from Gödel's T?'. To the first question, one can say that the main distinction was found in the treatment of the (weaker) projection argument. While the monotone interpretation resulted in a quantitative version equivalent to the original statement, the computational content in the context of the bounded interpretation, although weaker, is easier to extract and still enough to carry out the full quantitative analysis of Browder's theorem. Essentially, the nested bounded quantifications could be dealt swiftly by the bounded interpretation and the analysis still results in a quantitative bound sufficient for our purposes. This is because

the use of certain choice functions required by the monotone functional interpretation, which forcibly raise the types in the formula, can be replaced by collection arguments. The second question was first commented by Kohlenbach, e.g. in [28]: he attributed the simpler analysis to the structure of the original proof and to a mild use of weak compactness. However, a deeper look at the original argument showed that it was possible to modify Browder's original proof and obtain an "intermediate" proof which could be formalizable in the setting of the bounded functional interpretation without bar recursion. Further down the line, this idea was extended to define a "macro" which could be employed to avoid certain sequential weak compactness arguments in proof mining [12]. This general procedure for avoiding weak compactness not only justifies the second point in Browder's analysis, but was employed with the same goal in quantitative analyses of Wittmann's theorem [51] and, with a slight adaptation, of Bauschke's theorem [2].
Wittmann's theorem is an important result in nonlinear analysis that, under some conditions, shows the strong convergence of the so-called *Halpern iteration*,

$$u_0 \in X, \quad u_{n+1} := \lambda_n u_0 + (1-\lambda_{n+1})U(u_n), \tag{1.2}$$

where $(\lambda_n) \subset [0,1]$ is a sequence of real numbers, u_0 is some initial point in the Hilbert space and U is a nonexpansive map. This theorem extends previous results in the sense that it considers conditions which are satisfied by the natural choice $\lambda_n = \frac{1}{n+1}$ and, in this sense, can be seen as a nonlinear extension of von Neumann's mean ergodic theorem. Bauschke's theorem considers instead a finite number of nonexpansive maps and is still able to prove the strong convergence to a common fixed point for all the maps. The conditions considered by Bauschke are the natural extension of Wittmann's conditions and the converging sequence can also be seen as an adaptation of Halpern iterations to a finite number of maps. In fact, Wittmann's strong convergence result is obtained as the particular case of Bauschke's theorem when we only work with one nonexpansive map. Although Bauschke's theorem was given a quantitative version before by Daniel Körnlein in his PhD thesis [35], Bauschke's original proof was for the first time analyzed here.

In a second part of this thesis, following a suggestion of Laurenţiu Leuştean, we looked at quantitative versions of some results on the strong convergence of variants of the proximal point algorithm. R. T. Rockafellar's famous proximal point algorithm [42] is a useful iteration used to find zeros of maximal monotone operators, which in turn relates to many problems in optimization theory and nonlinear analysis. However these iterations in general fail to strongly converge and many variants were devised in an attempt to ensure such strong convergence. One of those variants is the *Halpern type proximal point algorithm*, first considered by Shoji Kamimura and Wataru Takahashi in [20] and, independently, by Hong-Kun Xu in [52],

$$x_{n+1} := \alpha_n u + (1-\alpha_n)J_{\beta_n}(x_n), \tag{HPPA}$$

where, for $\beta > 0$, J_β is the resolvent function of a maximal monotone operator, $(\alpha_n) \subset]0,1[$, u is some "anchor" point of the space and $(\beta_n) \subset \mathbb{R}^+$.
The motivation behind these types of iteration resides in the success of the results due to Halpern and Wittmann in fixed point theory. In addition, the strong convergence results of HPPA considered here are proved in a very similar manner to Wittmann's theorem and thus, it was not expected that their quantitative analysis should pose any additional problems. In the end, metastability results (in the sense of Terence Tao [46]) for theorems due to Boikanyo and Moroşano in [5] and due to Xu in [53] were obtained. Again, whereas the original arguments rely on sequential weak compactness, we avoided the use of bar recursive functionals and the quantitative analysis can be seen as an application of the "macro" obtained before for results in the context of fixed point theory.

1.1 Outline of this thesis

We start in the first two chapters by presenting the theoretical background of the bounded functional interpretation. In Chapter 2, we introduce the framework of finite type arithmetic which will serve as the underlying system for all the formal systems considered in this thesis. Then, we discuss the intensional majorizability notion and define the bounded functional interpretation directly into Peano arithmetic in all finite types. In Chapter 3 we briefly explain how the interpretation can be extended to formal systems with a new base type that are suitable for formalizing the proofs considered here. This is done by focusing on the context of bounded metric spaces, while describing how to proceed in the general case. We explain how the real numbers can be represented via the signed-digit representation and remark some of the benefits when compared to the usual Cauchy-sequence representation. We end the chapter with some logical theorems in the same spirit as the metatheorems for the monotone functional interpretation. This chapter serves as a theoretical background to the extraction of computational information in the following chapters but its goal is not to give a complete description of the appropriate formal theories neither of the corresponding metatheorems.
Chapter 4 is centered on the proof mining program and the use of the bounded functional interpretation to the analysis of mathematical proofs. First, we discuss some isolated topics: we look at the metastability property; we show some quantitative results related to the infinite convergence principle; we look at a way to avoid the use of $\limsup$ – which in general requires arithmetical comprehension – by replacing it with rational approximations; and we make some considerations regarding the analysis of proofs that follow a discussion by cases. Secondly, we explain the general idea behind the elimination of certain sequential weak compactness arguments from proof mining practice. This is first motivated by showing an "intermediate" proof for Browder's theorem in the context of the bounded functional interpretation. Then, we abstract these ideas and prove a general result which can be ap-

plied to the quantitative minings of many mathematical results (as shown in the following chapters). Finally, we end the chapter with the analysis of the projection argument guided by the bounded functional interpretation.
Chapters 5 and 6 are reserved for the analysis of concrete cases. In Chapter 5, we look at the direct application of the "macro" obtained in the previous chapter to produce quantitative versions of Browder, Wittmann and Bauschke's theorems. In Chapter 6, we look at the proof mining of the results on the strong convergence of variants of the proximal point algorithm. We begin with some introductory remarks related to the theory of monotone operators and to the proximal point algorithm. Then we look at some quantitative technical results that will be useful in the analyses that follow. The chapter ends with the proof of the metastable versions of the strong convergence of Halpern type proximal point algorithm.

We finish the thesis with some remarks and considerations on possible future work.

Chapter 2

Logical framework

In this chapter, we introduce the basic formal framework that serves as the base for all the theories in this thesis. We start by describing the theory of arithmetic in all finite types and present the standard structure and the model of strongly majorizable functionals, where a notion of majorizability between functionals plays an essential role. After that, we extend this theory with a notion of majorizability that is partially governed by a rule. Finally, we introduce the (Shoenfield-like) Bounded Functional Interpretation of Peano Arithmetic as done in [11].

2.1 Arithmetic in all finite types

Let $\mathcal{T}$ be the set of all finite types, which is defined inductively by:

(i) $0 \in \mathcal{T}$ (the ground or base type);

(ii) if $\rho, \sigma \in \mathcal{T}$, then $\rho \to \sigma \in \mathcal{T}$.

The standard interpretation is that objects of type 0 are natural numbers and the objects of type $\rho \to \sigma$ are the (total) functions from objects of type ρ to objects of type σ. Usually, one denotes the type $0 \to 0$ by 1 and in general, for any natural number n, $n+1$ denotes the type $n \to 0$. These are called pure types.

The language of finite type arithmetic, $\mathcal{L}^\omega$, is a many-sorted language with variables $x^\rho, y^\rho, z^\rho, \cdots$ and quantifiers $\forall x^\rho$, $\exists x^\rho$ for every finite type $\rho \in \mathcal{T}$. We have the following constant symbols: 0^0 (zero), S^1(successor), the combinators $\Pi_{\rho,\sigma}$ of type $\rho \to (\sigma \to \rho)$ and $\Sigma_{\rho,\sigma,\tau}$ of type $(\rho \to (\sigma \to \tau)) \to ((\rho \to \sigma) \to (\rho \to \tau))$, and simultaneous recursors $\underline{R}_{\underline{\rho}}$ of type $0 \to (\underline{\rho} \to ((\underline{\rho} \to (0 \to \underline{\rho})) \to \underline{\rho}))$. Furthermore, the only primitive predicate is $=_0$ for the equality between natural numbers (we follow the minimal treatment of equality as described by A. S. Troelstra in [48]).

The terms of type ρ in $\mathcal{L}^\omega$ are the constants of type ρ, the variables of type ρ or given by the application of terms of type $\sigma \to \rho$ and type σ, i.e. if t is a term of type $\sigma \to \rho$ and u is a term of type σ, then tu is a term of type ρ. For term application, if $t, u_1, \cdots, u_k$ are terms, by $tu_1u_2 \cdots u_k$ we mean the term resulting from the consecutive applications $(\cdots((tu_1)u_2)\cdots)u_k$, meaning that whenever we hide the parenthesis we are associating to the left. As usual, $\underline{t}$ denotes a tuple of terms $t_1, \cdots, t_k$, more precisely, $\underline{t}^{\underline{\rho}}$ denotes a (possibly empty) tuple of terms $t_1^{\rho_1}, \cdots, t_k^{\rho_k}$.

The atomic formulas of $\mathcal{L}^\omega$ are of the form $t =_0 q$, with t, q terms of type 0. Formulas are constructed recursively:

(i) atomic formulas are formulas;

(ii) if A, B are formulas, then $A \wedge B$, $A \vee B$ and $A \to B$ are also formulas;

(iii) if A is a formula, then for all $\rho \in T$, $\forall x^\rho A$ and $\exists x^\rho A$ are also formulas.

As usual, $\neg A$ is $A \to \perp$, where $\perp \equiv 0 =_0 1$ and $A \leftrightarrow B$ is $(A \to B) \wedge (B \to A)$.

Pointwise equality for higher types, $=_\rho$, can be defined recursively with

$$t =_{\rho\to\sigma} q \;\text{ being }\; \forall x^\rho \,(tx =_\sigma qx)\,,$$

for t, q terms of type $\rho \to \sigma$ and with x a variable of type ρ that does not occur in t, q.

In the theory of finite type arithmetic we have the following axioms:

(a) **Equality Axioms:**
$n =_0 n$;
$n =_0 m \wedge A[n/w] \to A[m/w]$,
where A is an atomic formula with one distinguished type 0 variable w and $A[t/w]$ is obtained from A by replacing all free occurrences of w by the term t;

(b) **Axioms for Successor:**
$0 \neq_0 S(n)$;
$S(n) =_0 S(m) \to n =_0 m$;

(c) **Axioms for Recursors:**
$A[R_\rho(0, y, z)/w] \leftrightarrow A[y/w]$;
$A[R_\rho(Sx, y, z)/w] \leftrightarrow A[z(R_\rho(x, y, z), x)/w]$,
where A is an atomic formula with a distinguished variable w of type ρ;

(d) **Axioms for $\Pi_{\rho,\sigma}$:**
$A[\Pi_{\rho,\sigma}xy/w] \leftrightarrow A[x/w]$,
where A is an atomic formula with a distinguished variable w of type ρ;

(e) **Axioms for $\Sigma_{\rho,\sigma,\tau}$:**
$A[\Sigma_{\rho,\sigma,\tau}xyz/w] \leftrightarrow A[(xz)(yz)/w]$,
where A is an atomic formula with an distinguished variable w of type ρ;

(f) **Induction Scheme:**
$(A(0) \wedge \forall x^0(A(x) \to A(S(x)))) \to \forall x^0 A(x)$,
where A is an arbitrary formula of $\mathcal{L}^\omega$.

It is easy to show that $=_0$ is symmetric and transitive and to see that axioms for equality, recursors, projector and combinators extend to A an arbitrary formula.

The theory HA^ω is called *Heyting Arithmetic in all finite types* and is the above described theory where the underlying logic is intuitionistic. The *Peano Arithmetic in all finite types*, PA^ω, is its classical counterpart and results from HA^ω by adding the law of excluded middle (LEM).

The combinators Π and Σ are crucial to ensure the combinatorial completeness that allows for definitions by λ-abstraction.

Theorem 2.1. *For every term $t[x^\rho]^\sigma$, it is possible to construct a term $q^{\rho\to\sigma}$ of $\mathcal{L}^\omega$ satisfying:*

(i) $FV(q) = FV(t) \setminus \{x\}$ and

(ii) $\mathsf{HA}^\omega \vdash A[t[s/x]/w] \leftrightarrow A[qs/w]$, for every atomic formula A with a distinguished variable w of type σ.

This result extends to all formulas of the language, provided there is no clash of variables. The term q is usually denoted by $\lambda x.\, t$ and, then, the result says that the term $t[s/x]$ can be substituted by $(\lambda x.\, t)s$ in any formula.

Using the recursor R_0, one may construct a closed term for each description of a primitive recursive function satisfying the respective conditions of the description. Hence, HA^ω contains all primitive recursive functions – and so we can see HA as a subsystem of HA^ω. Moreover, using the recursors $\underline{R}_{\underline{\rho}}$ in all generality, we can define functions beyond the primitive recursive ones, e.g. the canonical example, the Ackermann function (see [47] for details).

In HA^ω, it can be shown that all the quantifier-free formulas have a characteristic function – this property plays an essential role in the proof of the soundness theorem for the Dialectica and the monotone functional interpretations.

Proposition 2.2. *Let $A_{qf}(\underline{x})$ be a quantifier-free formula of $\mathcal{L}^\omega$ whose free variables are all among $\underline{x}$. Then it is possible to construct a closed term t of appropriate type such that*

$$\mathsf{HA}^\omega \vdash \forall x\big(t\underline{x} =_0 0 \leftrightarrow A_{qf}(\underline{x})\big).$$

Since in HA^ω it can be shown that $\forall n^0\,(n =_0 0 \vee n \neq_0 0)$, we have LEM restricted to quantifier-free formulas:

Corollary 2.3. *Let A_{qf} be a quantifier-free formula of $\mathcal{L}^\omega$. Then*

$$\mathsf{HA}^\omega \vdash A_{qf} \vee \neg A_{qf}.$$

In finite type arithmetic, one can define the usual less or equal relation between natural numbers, $\leq_0$, and the usual term $\max_0$ of type $0 \to (0 \to 0)$, that gives the maximum of any two natural numbers. Some basic properties are easily deduced.

Lemma 2.4. *We have in HA^ω, with x, y, z variables of type 0,*

(i) $x \leq_0 x$;

(ii) $x \leq_0 y \wedge y \leq_0 z \to x \leq_0 z$;

(iii) $x \leq_0 \max_0(x, y) \wedge y \leq_0 \max_0(x, y)$;

(iv) $x' \leq_0 x \wedge y' \leq_0 y \to \max_0(x', y') \leq_0 \max_0(x, y)$.

A less or equal relation for higher types, $\leq_\rho$, can be defined recursively in a pointwise fashion,

$$t \leq_{\rho\to\sigma} q :\equiv \forall u^\rho(tu \leq_\sigma qu)$$

Models of finite type arithmetic

Now we present the standard model for finite-type arithmetic, $\mathcal{S}^\omega$, and the strongly majorizable functionals model, $\mathcal{M}^\omega$.

The standard model

Let S_0 be the set of natural numbers $\mathbb{N}$ and $S_{\rho\to\tau}$ the set of all functions from S_ρ to S_τ, i.e., $S_\tau^{S_\rho}$. Define $\mathcal{S}^\omega := \langle S_\sigma \rangle_{\sigma\in\mathcal{T}}$. With the natural interpretations, it is easy to see that $\mathcal{S}^\omega$ is a model of $\mathsf{E}\text{-}\mathsf{PA}^\omega$, i.e., PA^ω together with the axiom of full extensionality

$$\mathsf{E} : \forall z^{\rho\to\tau}\forall x^\rho, y^\rho(x =_\rho y \to zx =_\tau zy).$$

The model $\mathcal{S}^\omega$ is usually called the *standard structure for finite type arithmetic.*

The model of majorizable functionals

We now display the *model of strongly majorizable functionals* introduced by Bezem in [3] which relies on a strong majorizability notion that is a variant of Howard's majorizability relation ([19]).
For every finite-type $\rho \in \mathcal{T}$, we define a set M_ρ and a majorizability relation $\leq^*_\rho$, in the following recursive way:

$M_0 := \mathbb{N}$ is the set of natural numbers and $\leq^*_0$ is the usual "less or equal" relation between natural numbers.
Given ρ, σ finite types and x, y elements of $M_\sigma^{M_\rho}$, we define

$$x \leq^*_{\rho\to\sigma} y \ \text{ :iff } \ \forall u, v \in M_\rho \Big(u \leq^*_\rho v \to x(u) \leq^*_\sigma y(v) \wedge y(u) \leq^*_\sigma y(v) \Big),$$

and say that x is (strongly) majorized by y.

Finally define $M_{\rho\to\sigma} := \{x \in M_\sigma^{M_\rho} \mid \exists y \in M_\sigma^{M_\rho}(x \leq^*_{\rho\to\sigma} y)\}$.

Then, $\mathcal{M}^\omega := \langle M_\rho \rangle_{\rho\in\mathcal{T}}$ with the natural interpretations for constants and for application of functionals is a model for $\mathsf{E}\text{-}\mathsf{PA}^\omega$.
For details see [10] and [25].

We have a easy result regarding this majorizability relation,

Lemma 2.5. *The following proprieties are true,*

(1) $x \leq^*_\rho y \to y \leq^*_\rho y$

(2) $x \leq^*_\rho y \wedge y \leq^*_\rho z \to x \leq^*_\rho z$

(3) $x \leq_\rho y \wedge y \leq^*_\rho z \to x \leq^*_\rho z$

(4) For all $\rho = \rho_1 \to (\cdots \to (\rho_k \to \sigma)\cdots)$ *and* $x, y : M_{\rho_1} \to (\cdots \to (M_{\rho_k} \to M_\sigma)\cdots)$ *we have:*

$$x \leq^*_\rho y \leftrightarrow \forall u_1, v_1, \cdots, u_k, v_k \Big(\bigwedge_{i=1}^{k} u_i \leq^*_{\rho_i} v_i \to x\underline{u},\ y\underline{u} \leq^*_\sigma y\underline{v} \Big)$$

If x is majorized by y, then by (1), y is self-majorizing implying that not only x is in $\mathcal{M}^\omega$ but the majorizing functional y is also in the model – this is relevant as this relation is not reflexive. In (2) we have the transitivity propriety of $\leq^*$. From (3), we see that the majorizability is $\leq$-downward preserved. Finally, to better see the relevance of the propriety (4), first observe that any type ρ can be written as $\rho_1 \to (\cdots \to (\rho_k \to 0)\cdots)$, i.e. ending at the base type. Thus, (4) allows one to reduce $\leq^*$ at the ρ level to the base type relation.

We have $S_0 = M_0$ and, since for every $f \in \mathbb{N}^{\mathbb{N}}$, $f \leq_1^* f^M$, where $f^M(n) := \max_{k \leq n} f(k)$, we get $M_1 = S_1$. However, the two models start diverging right at level $2 := (0 \to 0) \to 0$, meaning that $M_2 \subsetneq S_2$. This next example is from [10].
Consider the functional $\Sigma \in S_2$ that gives the first natural number where a function is nonzero, formally defined by:

$$\Sigma(f) := \begin{cases} n & \text{if } \ f(n) \neq 0 \ \text{ and } \ \forall k < n(f(k) = 0), \\ 0 & \text{if } \ \forall k(f(k) = 0). \end{cases}$$

If one assumes that $M_2 = S_2$, then there must exist a type 2 functional Ψ such that $\Sigma \leq_2^* \Psi$. In particular, for one such majorizing functional Ψ, we will have

$$\forall f \in S_1\big(f \leq_1^* 1^1 \to \Sigma(f) \leq \Psi(1^1)\big),$$

where 1^1 is the constant function 1. This gives a contradiction as $\Psi(1^1)$ is giving a bound on where the first nonzero term of a function $\leq^* 1$ can appear. In fact, as $\Psi(1^1)$ is a fixed natural number, we can define the type 1 function:

$$f(n) := \begin{cases} 0 & \text{if } n < \Psi(1^1) + 1, \\ 1 & \text{if } n \geq \Psi(1^1) + 1. \end{cases}$$

Then $f \leq_1^* 1^1$ and $\Sigma(f) = \Psi(1^1) + 1$ which gives the contradiction. This means that Σ cannot be majorizable and so $M_2 \subsetneq S_2$.

2.2 An intensional majorizability notion

Kohlenbach's monotone functional interpretation is similar to the Dialectica interpretation but relaxed the need of precise witnesses and instead only asks for bounds on those witnesses. This shift from precise terms to bounds allowed for the analysis of proofs using additional lemmas, most notably Weak Kőnig's Lemma. It is then natural that the first results from the proof mining program were quantitative studies of proofs using Heine-Borel compactness in the form of Weak Kőnig's Lemma. Obviously, to be able to talk about bounding terms it is required to have a suitable notion of majorizability. The monotone functional interpretation combines the Dialectica interpretation with Howard's majorizability notion [19], a predecessor of Bezem's notion. For the bounded functional interpretation we will need the stronger notion introduced by Bezem for his model of strongly majorizable functionals, like in the previous section – the monotone interpretation also holds if one uses this notion.

It will be useful to formalize the Bezem's majorizability notion inside our theory of finite type arithmetic.
For every finite-type $\rho \in \mathcal{T}$, the (strong) majorizability formulas "$x \leq^*_\rho y$" are then defined by:

$$\begin{cases} x \leq^*_0 y & :\equiv \; x \leq_0 y, \\ x \leq^*_{\rho\to\sigma} y & :\equiv \; \forall u^\rho, v^\rho \big(u \leq^*_\rho v \to xu \leq^*_\sigma yv \wedge yu \leq^*_\sigma yv\big). \end{cases}$$

In that situation, we say that x is majorized by y. We say that a term t is monotone if it is self-majorizing, i.e., $t \leq^*_\rho t$.

Easy properties can be shown to hold in the theory of finite type arithmetic:

Lemma 2.6. *For each finite-type $\rho \in \mathcal{T}$, we have:*

(i) $\mathsf{HA}^\omega \vdash x \leq^*_\rho y \to y \leq^*_\rho y$,

(ii) $\mathsf{HA}^\omega \vdash x \leq^*_\rho y \wedge y \leq^*_\rho z \to x \leq^*_\rho z$,

(iii) $\mathsf{HA}^\omega \vdash x \leq_\rho y \wedge y \leq^*_\rho z \to x \leq^*_\rho z$.

The next result will be necessary for the functional interpretations that we will present.

Theorem 2.7 (Howard)**.** *For each closed term t, there is a closed term q such that*

$$\mathsf{HA}^\omega \vdash t \leq^* q.$$

Notice that, trivially the majorizability relations formalized here coincide with the previous relations when interpreted in the model $\mathcal{M}^\omega$, and furthermore the model $\mathcal{M}^\omega$ satisfies the majorizability axioms MAJ^ω:

$$\mathsf{MAJ}^\omega : \quad \forall x \exists y (x \leq^* y).$$

For the purpose of the bounded functional interpretation, presented in the next section, it is crucial that we work with an *intensional notion of majorizability.* This new notion, the counterpart of the extensional majorizability relation defined above, is called intensional in the sense that it is now partially governed by a rule.

We defined a new language, $\mathcal{L}^\omega_\trianglelefteq$, as the extension of $\mathcal{L}^\omega$ with new atomic predicate symbols between terms of type ρ, $\trianglelefteq_\rho$, for each finite type $\rho \in \mathcal{T}$. We also consider *bounded quantifiers*, $\forall x \trianglelefteq_\rho t$ and $\exists x \trianglelefteq_\rho t$, where x does not occur in t.

We now have new atomic formulas that are of the form $t \trianglelefteq_\rho q$, where t and q are terms of type ρ, as well as formulas build up using the bounded quantifiers. A formula is called *bounded* if all of its quantifiers are bounded ones.

Peano arithmetic in all finite types with intensional majorizability, $\mathsf{PA}^\omega_\trianglelefteq$ is obtained from PA^ω by extending the axioms for the recursors, for the Π and Σ terms to the new atomic formulas and the induction scheme axiom to all the new formulas of $\mathcal{L}^\omega_\trianglelefteq$. Regarding the new predicate symbols and the bounded quantifiers we add:

(a) **Axioms for the bounded universal quantifiers:**
$\forall x \trianglelefteq_\rho t\, A(x) \leftrightarrow \forall x\, (x \trianglelefteq_\rho t \to A(x))$,

(b) **Axioms for the bounded existential quantifiers:**
$\exists x \trianglelefteq_\rho t\, A(x) \leftrightarrow \exists x\, (x \trianglelefteq_\rho t \wedge A(x))$,

where x and t are of type ρ, x does not occur in t and A is an arbitrary formula of $\mathcal{L}^\omega_\trianglelefteq$.

(c) **Axioms for $\trianglelefteq_\rho$, $\rho \in \mathcal{T}$:**
$n \trianglelefteq_0 m \leftrightarrow n \leq_0 m$,
$x \trianglelefteq_{\rho\to\sigma} y \to \forall u \forall v\, (u \trianglelefteq_\rho v \to xu \trianglelefteq_\sigma yv \wedge yu \trianglelefteq_\sigma yv)$,

where n,m are variables of type 0, x, y are variables of type $\rho \to \sigma$ and u and v are variables of type ρ. Notice that in the second axiom of (c) we only ask for the direct implication instead of the equivalence as in the definition of the extensional $\leq^*$. In its place we consider the following rule.

(d) **Majorizability rule:**

$$\mathsf{RL}_\trianglelefteq : \quad \frac{A_{bd} \wedge u \trianglelefteq_\rho v \to tu \trianglelefteq_\sigma qv \wedge qu \trianglelefteq_\sigma qv}{A_{bd} \to t \trianglelefteq_{\rho\to\sigma} q}$$

where A_{bd} is a bounded formula, t, q are terms of type $\rho \to \sigma$ and u, v are of type ρ and don't appear free in the conclusion.

A term t is *monotone* if $t \trianglelefteq t$. A *monotone universal quantification* is a quantification of the form $\forall x\, (x \trianglelefteq x \to A(x))$, where A is any formula, and that we abbreviate by $\widetilde{\forall} x\, A(x)$. Analogous for *monotone existential quantifiers*. Note that, in general, monotone quantifications are not bounded.

Proposition 2.8. *We have,*

1. $\mathsf{PA}^\omega_\trianglelefteq \vdash x \trianglelefteq_\rho y \to y \trianglelefteq_\rho y$*;*
2. $\mathsf{PA}^\omega_\trianglelefteq \vdash x \trianglelefteq_\rho y \wedge y \trianglelefteq_\rho z \to x \trianglelefteq_\rho z$*.*

Flattening: returning to $\leq^*$

To have a soundness theorem for the bounded functional interpretation, for reasons similar to those that prevent the Dialectica from interpreting full extensionality, we must consider the majorizability notion partially governed by a rule. Essentially, this rule will deactivate the computational content of the majorizability relation with respect to the bounded functional interpretation.
The need of this intensionality implies that in the theory $\mathsf{PA}^\omega_\trianglelefteq$ the deduction theorem fails, which is sometimes viewed as an unattractive feature. However, this property actually highlights the subtle distinction between postulates – on the left-hand side of the provability sign – and implicative premises – on the right-hand side –, showing that this difference can have an impact on the actual application of functional interpretations in proof mining.

Nevertheless, it is possible to return to an extensional realm by replacing all the intensional symbols of majorizability with their extensional counterparts. Following [11], we call this procedure "flattening".
Consider the language $\mathcal{L}^\omega_{\leq^*}$ to be the extension of $\mathcal{L}^\omega$ obtained by adding the symbols $\leq^*_\rho$ for every finite type ρ. Consider the theory $\mathsf{PA}^\omega_{\leq^*}$ which is an extension by definitions of the theory PA^ω and where $\leq^*$ are the extensional majorizability relations.

Definition 2.9. Let A be a formula of $\mathcal{L}^\omega_\trianglelefteq$. The flattening of A, A^*, is the formula in $\mathcal{L}^\omega_{\leq^*}$ obtained from A by replacing all instances of intensional majorizability relations $\trianglelefteq_\rho$ with their extensional versions $\leq^*_\rho$.
Furthermore, if Γ is a set of sentences in $\mathcal{L}^\omega_\trianglelefteq$, then the set Γ^*, the flattened version of Γ, is the set of sentences in $\mathcal{L}^\omega_{\leq^*}$ obtained from flattening the sentences of Γ.

The theory $\mathsf{PA}^\omega_{\leq^*}$ is the flattened version of $\mathsf{PA}^\omega_\trianglelefteq$ and, since any formula proved using the rule $\mathsf{RL}_\trianglelefteq$ can be proved using the implication, we have the following result.

Lemma 2.10 (Flattening)**.** *Let A be a formula of the language $\mathcal{L}^\omega_\trianglelefteq$ and Γ a set of sentences. If $\mathsf{PA}^\omega_\trianglelefteq + \Gamma \vdash A$, then $\mathsf{PA}^\omega_{\leq^*} + \Gamma^* \vdash A^*$.*

2.3 The bounded functional interpretation

We will now present the bounded functional interpretation in a classical context.
This can be done by defining the interpretation in an intuitionistic setting and using a negative translation to jump to the classical logic. We will instead procede in a Shoenfield-like manner and introduce the bounded functional interpretation directly in $\mathsf{PA}^\omega_\trianglelefteq$. We followed the original paper [11]. In fact, in [14] Jaime Gaspar showed that, similar to the Shoenfield interpretation – which is Dialectica after Krivine [1], [44] –, Ferreira's Shoenfield-like bounded functional interpretation can be seen as Ferreira and Oliva's bounded functional

interpretation of $\mathsf{HA}^\omega_\trianglelefteq$ [13] after Krivine's negative translation [38], [45].
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 2.11. To each formula A of $\mathcal{L}^\omega_\trianglelefteq$ we assign formulas $(A)^U$ and A_U so that $(A)^U$ is of the form $\widetilde{\forall}\underline{b}\widetilde{\exists}\underline{c}A_U(\underline{b},\underline{c})$ with $A_U(\underline{b},\underline{c})$ a bounded formula, according to the following clauses:

1. $(A)^U$ and A_U are simply A, for prime formulas A.

Suppose we already have interpretations for A and B given, respectively, by $\widetilde{\forall}\underline{b}\widetilde{\exists}\underline{c}A_U(\underline{b},\underline{c})$ and $\widetilde{\forall}\underline{d}\widetilde{\exists}\underline{e}B_U(\underline{d},\underline{e})$, then we define:

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

Note that if A_{bd} is a bounded formula, then $(A_{\mathrm{bd}})^U$ and $(A_{\mathrm{bd}})_U$ are simply A_{bd}, i.e. bounded formulas are left invariant under the interpretation. In the definition of the formula $(\neg A)^U$ appears the apparently innocuous bounded quantification "$\widetilde{\exists}b' \trianglelefteq b$". However, this quantification changes the definition of the matrix $(\neg A)_U$, which in turn ensures the following crucial monotonicity propriety: the matrix $A_U(\underline{b},\underline{c})$ is monotonous in the existentially quantified variables. Thus, any bound on a witness for A_U is itself a witness.

Lemma 2.12 (Monotonicity of U)**.** *For each formula A of the language $\mathcal{L}^\omega_\trianglelefteq$ we have*

$$\mathsf{PA}^\omega_\trianglelefteq \vdash \widetilde{\forall}\underline{b}\left(\widetilde{\exists}\underline{\tilde{c}}\,\widetilde{\exists}\underline{c} \trianglelefteq \underline{\tilde{c}}\,A_U(\underline{b},\underline{c}) \to A_U(\underline{b},\underline{\tilde{c}})\right).$$

Proof. The proof is by induction on the complexity of the formula A. The base case is trivially true, since in A^U the tuples $\underline{b}$ and $\underline{c}$ are empty. Every case follows easily from the induction hypothesis with the exception of the negation's case. Consider that the Lemma holds for a formula A and we want to see that it still holds for $\neg A$. We have $(\neg A)^U \equiv \widetilde{\forall}\underline{f}\widetilde{\exists}\underline{b}(\neg A)_U(\underline{f},\underline{b})$, with $(\neg A)_U(\underline{f},\underline{b}) \equiv \widetilde{\exists}\underline{b}' \trianglelefteq \underline{b}\neg A_U(\underline{b}',\underline{f}\underline{b}')$. Fix an arbitrary monotone $\underline{f}$ and assume that there is a bound $\underline{\tilde{b}}$ such that

$$\widetilde{\exists}\underline{b} \trianglelefteq \underline{\tilde{b}}\,(\neg A)_U(\underline{f},\underline{b}) \text{ which is } \widetilde{\exists}\underline{b} \trianglelefteq \underline{\tilde{b}}\widetilde{\exists}\underline{b}' \trianglelefteq \underline{b}\,\neg A_U(\underline{b}',\underline{f}\underline{b}').$$

Then by the transitivity property of $\trianglelefteq$ – property 2 of 2.8 – we conclude

$$\widetilde{\exists}\underline{b}' \trianglelefteq \underline{\tilde{b}}\,\neg A_U(\underline{b}',\underline{f}\underline{b}') \text{ which is } (\neg A)_U(\underline{f},\underline{\tilde{b}}).$$

□

Characteristic Principles

We now show the three principles that have a crucial role in the interpretation.

1. Bounded Collection Principle, $\mathsf{BC}^\omega_{\mathsf{bd}}$:

$$\mathsf{BC}^\omega_{\mathsf{bd}}: \quad \forall \underline{x} \trianglelefteq \underline{a} \exists \underline{y}\, A_{\mathrm{bd}}(\underline{x}, \underline{y}) \to \widetilde{\exists} \underline{b}\, \forall \underline{x} \trianglelefteq \underline{a} \exists \underline{y} \trianglelefteq \underline{b}\, A_{\mathrm{bd}}(\underline{x}, \underline{y}),$$

   where A_{bd} is a bounded formula of $\mathcal{L}^\omega_{\trianglelefteq}$.

2. Monotone Bounded Choice, $\mathsf{mAC}^\omega_{\mathsf{bd}}$:

$$\mathsf{mAC}^\omega_{\mathsf{bd}}: \quad \widetilde{\forall} \underline{x} \widetilde{\exists} \underline{y}\, A_{\mathrm{bd}}(\underline{x}, \underline{y}) \to \widetilde{\exists} \underline{f} \widetilde{\forall} \underline{x} \widetilde{\exists} \underline{y} \trianglelefteq \underline{f}\underline{x}\, A_{\mathrm{bd}}(\underline{x}, \underline{y}),$$

   where A_{bd} is a bounded formula of $\mathcal{L}^\omega_{\trianglelefteq}$.

3. Majorizability Axioms, MAJ^ω:

$$\mathsf{MAJ}^\omega: \quad \forall x \exists y\, (x \trianglelefteq y).$$

The majorizability axioms state that every element is intensionally majorizable – since by the context there is never the risk of confusion, we use the same notation for both the intensional and extensional versions of this principle. 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. Finally and maybe more interesting, we look at the collection 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 bellow a certain bound b. Furthermore, its contra-positive allows for the conclusion of an element x (bellow a) such that $\forall y \neg A_{\mathrm{bd}}(x, y)$, from the weaker statement that such x's (bellow a) only exist "locally". We may regard such x as an ideal element that works uniformly for each b. The original paper where this interpretation was introduced [11] placed emphasis on this uniformity aspect and that is the reason the letter U is used for the interpretation.

The theory $\mathsf{PA}^\omega_{\trianglelefteq}$ together with these three principles is not set-theoretically sound. For example, it refutes the weakest form of extensionality. More specifically, it can be shown, with the use of the principle $\mathsf{BC}^\omega_{\mathsf{bd}}$, that it proves the negation of the sentence

$$\forall \Phi^2 \forall \alpha^1, \beta^1 \big(\forall k^0 (\alpha k = \beta k) \to \Phi\alpha = \Phi\beta\big).$$

These principles characterize the interpretation in the following sense.

Theorem 2.13 (Characterization)**.** *Let A be an arbitrary formula of $\mathcal{L}^{\omega}_{\trianglelefteq}$. Then,*

$$\mathsf{PA}^{\omega}_{\trianglelefteq} + \mathsf{mAC}^{\omega}_{\mathsf{bd}} + \mathsf{BC}^{\omega}_{\mathsf{bd}} + \mathsf{MAJ}^{\omega} \vdash A \leftrightarrow (A)^{U}.$$

The result is shown by an easy induction on the structure of the formula A.

Soundness and Extraction

Like in the case of the monotone interpretation, the bounded functional interpretation can consider proofs that make use of auxiliary lemmas. Let Δ be a set of sentences of the form

$$\widetilde{\forall}\underline{a}\exists\underline{b} \trianglelefteq \underline{ra}\forall\underline{c}B_{\mathrm{bd}}(\underline{a},\underline{b},\underline{c}),$$

where $\underline{r}$ is a given tuple of closed terms and B_{bd} a bounded formula.

For each such set, define Δ_{w}, the weakening of Δ consisting of sentences of the form

$$\widetilde{\forall}\underline{a},\underline{a}'\exists\underline{b} \trianglelefteq \underline{ra}\forall\underline{c} \trianglelefteq \underline{a}'B_{\mathrm{bd}}(\underline{a},\underline{b},\underline{c}),$$

each corresponding to a sentence of Δ.
The formulas of Δ are those that contain all the information required by the interpretation, just as long as we have the corresponding Δ_{w} formula in the verifying theory.
The soundness theorem bellow guarantee the consistence of $\mathsf{PA}^{\omega}_{\trianglelefteq}$ together with the characterizing principles relative to Peano arithmetic. To prove the soundness theorem is crucial the fact that the Howard's theorem is still true for the intensional majorizability relation.

Lemma 2.14. *For each closed term t, there is a closed term q such that $\mathsf{PA}^{\omega}_{\trianglelefteq} \vdash t \trianglelefteq q$.*

The relevant observation is that Howard's construction only requires the use of the rule $\mathsf{RL}_{\trianglelefteq}$ and there is no need for the absent implication (cf. [13]).

Theorem 2.15 (Soundness)**.** *Let $A(\underline{a})$ be an arbitrary formula of $\mathcal{L}^{\omega}_{\trianglelefteq}$, with free variables $\underline{a}$. If*

$$\mathsf{PA}^{\omega}_{\trianglelefteq} + \mathsf{mAC}^{\omega}_{\mathsf{bd}} + \mathsf{BC}^{\omega}_{\mathsf{bd}} + \mathsf{MAJ}^{\omega} + \Delta \vdash A(\underline{a}),$$

then there exists a tuple of closed monotone terms $\underline{t}$ of $\mathcal{L}^{\omega}_{\trianglelefteq}$, which can be extracted from a proof of $A(\underline{a})$, such that

$$\mathsf{PA}^{\omega}_{\trianglelefteq} + \Delta_{\mathsf{w}}, \vdash \widetilde{\forall}\underline{w}\forall\underline{a} \trianglelefteq \underline{w}\widetilde{\forall}\underline{x}(A(\underline{a}))_{U}(\underline{x},\underline{twx}).$$

We could have written

$$\mathsf{PA}^{\omega}_{\trianglelefteq} + \Delta_{\mathsf{w}} \vdash \widetilde{\forall}\underline{w}\forall\underline{a} \trianglelefteq \underline{w}\widetilde{\forall}\underline{x}\widetilde{\exists}\underline{y} \trianglelefteq \underline{twx}(A(\underline{a}))_U(\underline{x},\underline{y}),$$

however, by the monotonicity propriety, the two are equivalent.

In this short overview of the bounded functional interpretation we opt by not writing the proof of the soundness theorem. The absent proof can be found in [11]. The proof is done by induction on the length of the formal derivation of $A(\underline{a})$. For the verifying theory we considered $\mathsf{PA}^{\omega}_{\trianglelefteq}$, however, $\mathsf{HA}^{\omega}_{\trianglelefteq} + \mathrm{LEM}_{\mathsf{bd}}$ also works, since we will only be required to have Law of Excluded Middle for bounded formulas – $\mathrm{LEM}_{\mathsf{bd}}$.

A particular instance of the soundness theorem is the following extraction theorem:

Theorem 2.16 (Extraction)**.** *Suppose*

$$\mathsf{PA}^{\omega}_{\trianglelefteq} + \mathsf{mAC}^{\omega}_{\mathsf{bd}} + \mathsf{BC}^{\omega}_{\mathsf{bd}} + \mathsf{MAJ}^{\omega} + \Delta \vdash \forall x \exists y A_{\mathrm{bd}}(x,y),$$

where A_{bd} is a bounded formula with free variables x, y, then it is possible to extract a closed monotone term t from a proof of the hypothesis such that

$$\mathsf{PA}^{\omega}_{\trianglelefteq} + \Delta_{\mathsf{w}} \vdash \widetilde{\forall} w \forall x \trianglelefteq w \exists y \trianglelefteq tw A_{\mathrm{bd}}(x,y).$$

Proof. This result is a direct application of the soundness theorem by computing the interpretation of $\forall x \exists y A_{\mathrm{bd}}(x,y) \equiv \forall x \neg \forall y \neg A_{\mathrm{bd}}(x,y)$:

$$\begin{aligned}
(\neg A_{\mathrm{bd}}(x,y))^U &\equiv \neg A_{\mathrm{bd}}(x,y) \\
(\forall y \neg A_{\mathrm{bd}}(x,y))^U &\equiv \widetilde{\forall} a \forall y \trianglelefteq a \neg A_{\mathrm{bd}}(x,y) \\
(\neg \forall y \neg A_{\mathrm{bd}}(x,y))^U &\equiv \widetilde{\exists} a \widetilde{\exists} a' \trianglelefteq a \neg \forall y \trianglelefteq a' \neg A_{\mathrm{bd}}(x,y) \\
(\forall x \neg \forall y \neg A_{\mathrm{bd}}(x,y))^U &\equiv \widetilde{\forall} w \widetilde{\exists} a \forall x \trianglelefteq w \widetilde{\exists} a' \trianglelefteq a \neg \forall y \trianglelefteq a' \neg A_{\mathrm{bd}}(x,y),
\end{aligned}$$

which is $\widetilde{\forall} w \widetilde{\exists} a \forall x \trianglelefteq w \widetilde{\exists} a' \trianglelefteq a \exists y \trianglelefteq a' A_{\mathrm{bd}}(x,y)$. By the soundness theorem, from a proof of the hypothesis a closed monotone term t can be extracted satisfying

$$\mathsf{PA}^{\omega}_{\trianglelefteq} + \Delta_{\mathsf{w}} \vdash \widetilde{\forall} w \forall x \trianglelefteq w \widetilde{\exists} a' \trianglelefteq tw \exists y \trianglelefteq a' A_{\mathrm{bd}}(x,y),$$

which, using the transitivity of $\trianglelefteq$, yields the result. □

At this point, we can notice that the characterization theorem ensures that we are not missing any principles in the soundness theorem. Suppose to the contrary that we had the soundness theorem with a new principle P. Then, in particular, we would have, $\mathsf{PA}^{\omega}_{\trianglelefteq} + \mathsf{mAC}^{\omega}_{\mathsf{bd}} + \mathsf{BC}^{\omega}_{\mathsf{bd}} + \mathsf{MAJ}^{\omega} + \mathsf{P} \vdash \mathsf{P}$ and then, by that new soundness theorem, we would have

$\mathsf{PA}^\omega_\trianglelefteq \vdash \mathsf{P}^U$ and, in particular,$\mathsf{PA}^\omega_\trianglelefteq + \mathsf{mAC}^\omega_{\mathsf{bd}} + \mathsf{BC}^\omega_{\mathsf{bd}} + \mathsf{MAJ}^\omega \vdash \mathsf{P}^U$. By the characterization theorem, it now follows that the principle P is redundant in the theory.

One of the interesting points of the monotone functional interpretation compared to the Dialectica is that it can consider a wider class of auxiliary lemmas in its soundness theorem. Most notably, Kohlenbach showed that the Weak Kőnig's Lemma (WKL) could be considered as one such lemma. As in the monotone interpretation's case, the bounded functional interpretation can interpret Weak Kőnig's Lemma, however we don't even need to consider WKL as a Δ-sentence since the theory $\mathsf{PA}^\omega_\trianglelefteq$ together with the characteristic principles can actually prove it and thus, its presence in the theory is superfluous. In fact, bounded collection suffices in proving this principle.

The Weak Kőnig's Lemma can be stated as,

$$\mathsf{WKL}: \quad \forall \mathrm{T} \trianglelefteq_1 1 \left(\mathrm{Tree}_\infty(\mathrm{T}) \to \exists \alpha \trianglelefteq_1 1 \forall k^0\, \mathrm{T}(\overline{\alpha}k) = 0\right),$$

where $\mathrm{Tree}_\infty(\mathrm{T})$ says that T is an infinite binary tree and $\bar{\alpha}k$ denotes the code of the binary sequence $\langle \alpha(0), \cdots, \alpha(k-1)\rangle$.
$\mathrm{Tree}_\infty(\mathrm{T})$ is the conjunction of

T is closed for initial segments:	$\forall s^0, r^0(\mathrm{T}(s) = 0 \wedge r \preccurlyeq s \to \mathrm{T}(r) = 0)$
T is a binary tree:	$\forall s^0(\mathrm{T}(s) = 0 \to Seq_2(s))$
T is infinite:	$\forall n^0 \exists s^0(\mathrm{T}(s) = 0 \wedge \lvert s\rvert = n)$,

where $r \preccurlyeq s$ says that the sequence coded by r is an initial segment of the sequence coded by s, $Seq_2(s)$ is a predicate stating that s is the code of a binary sequence and $\lvert s\rvert$ is the length of the sequence coded by s.

Lemma 2.17. $\mathsf{PA}^\omega_\trianglelefteq + \mathsf{BC}^\omega_{\mathsf{bd}} \vdash \mathsf{WKL}$

Proof. Assume $\mathrm{Tree}_\infty(\mathrm{T})$. For each n, by the infinity property of T, there exists s a code of a binary sequence satisfying $\mathrm{T}(s) = 0$ and $\lvert s\rvert = n+1$. Define the functional α of type 1 that is obtained by extending the binary sequence coded by s, $\langle s_0, \cdots, s_n\rangle$, with zeros,

$$\alpha(m) := \begin{cases} s_m & \text{if } m \leq n+1 \\ 0 & \text{if } m > n+1 \end{cases}.$$

Since $\langle s_0, \cdots, s_n\rangle$ is a binary sequence, we have $\alpha \trianglelefteq 1$. Now, using the fact that T is closed for initial segments and $\mathrm{T}(s) = 0$, we conclude

$$\forall n^0 \exists \alpha \trianglelefteq_1 1 \forall k \leq n\, \mathrm{T}(\overline{\alpha}k) = 0.$$

Finally, using the contra-positive of $\mathsf{BC}^{\omega}_{\mathsf{bd}}$ we conclude

$$\exists \alpha \trianglelefteq_1 1 \forall k^0 \, \mathrm{T}(\overline{\alpha} k) = 0.$$

$\square$

It is even possible to argue the extensional version,

$$\mathsf{WKL}^* : \quad \forall \mathrm{T} \leq_1 1 \left(\mathrm{Tree}_{\infty}(\mathrm{T}) \to \exists \alpha \leq_1 1 \forall k^0 \, \mathrm{T}(\overline{\alpha} k) = 0 \right),$$

by the simple remark that if $\mathrm{T} \leq_1 1$, then $\min(\mathrm{T}, 1^1) =_1 T$ and we have $\min(\mathrm{T}, 1^1) \trianglelefteq_1 1$. We can then apply the intensional version of WKL to $\min(\mathrm{T}, 1^1)$ and conclude the extensional one.

With the Flattening Lemma (2.10) employed after the Extraction Theorem we get:

Corollary 2.18. *Suppose*

$$\mathsf{PA}^{\omega}_{\trianglelefteq} + \mathsf{mAC}^{\omega}_{\mathsf{bd}} + \mathsf{BC}^{\omega}_{\mathsf{bd}} + \mathsf{MAJ}^{\omega} + \Delta \vdash \forall x \exists y A_{\mathrm{qf}}(x, y),$$

where A_{qf} is a quantifier-free formula of $\mathcal{L}^{\omega}$ with free variables x, y, then it is possible to extract a closed monotone term t from a proof of the hypothesis such that

$$\mathsf{PA}^{\omega} + (\Delta_w)^* \vdash \widetilde{\forall} w \forall x \leq^* w \exists y \leq^* tw A_{\mathrm{qf}}(x, y).$$

Thus, the conjunction of the flattening argument with the Extraction Theorem for the bounded functional interpretation yields an extensional version of the theorem which no longer uses the intensional majorizability symbols – notice that the monotone quantification $\widetilde{\forall} w$ is also extensional.

Chapter 3

The extended bounded functional interpretation

This chapter is devoted to the justification of applying the bounded functional interpretation to the analysis of concrete mathematical proofs. The application of this proof-theoretical technique require us to have a formal proof of the theorem that we which to analyze. A formal proof, in turn, entails the need of a suitable formal setting in which the proof can be formalized and where the functional interpretation can be applied.
We will start by talking about the signed-digit representation of the real numbers and its benefits when compared to the usual Cauchy-sequence representation. Next, as an example of theories to which the bounded functional interpretation can be extended, we will see a system tailored to treat results in the context of bounded metric spaces. Finally, we present a general result which ensures the extraction procedure of quantitative information from proofs in that system. This example serves as a good illustration of the formal systems used in the applications in chapters 4 and 5.

3.1 The real numbers

In order to describe ordinary mathematical proofs we must first explain how the real numbers are represented in our formal setting. There are several ways of representing the real numbers like Cauchy sequences of rational numbers, Dedekind cuts and binary representations. Here we present the signed-digit representation and explain its advantages compared to the more usual, and maybe more intuitive, Cauchy-sequence representation – used for example by Kohlenbach [25]. All these representations are equivalent in the sense that the resulting structures are isomorphic.

A first step in the representation for the real numbers is to define the rational numbers.

As so, we start by doing this in the usual and natural way via a coding $j(n,m)$ of pairs of natural numbers. The code $j(n,m)$ represents the rational number $\frac{n'}{m+1}$ if $n = 2n'$, and $-\frac{n'}{m+1}$ if $n = 2n' - 1$. Thus every rational number can be represented by one such code of a certain (not unique) pair of natural numbers. The equality $=_{\mathbb{Q}}$, the inequalities $<_{\mathbb{Q}}$, $\leq_{\mathbb{Q}}$ and the operations $+_{\mathbb{Q}}$, $\cdot_{\mathbb{Q}}$ between representants of rational numbers are defined in the usual way.
Whenever we talk about rational numbers we must always use the coding j. However, just to ease the reading, we usually write the rational number instead.

In the signed-digit representation (see [50] for details), real numbers are represented by tuples (n, α), where n is a natural number and α is a sequence of numbers in $\{0, 1, 2\}$. The pair (n, α) represents the real number

$$int(n) + \sum_{i=1}^{\infty} (\alpha_i - 1)\frac{1}{2^i},$$

where $int(n)$ is the rational number coded by $j(n, 0)$. In this particular instance of the function j, we have that

$$int(n) \text{ represents } \begin{cases} m & \text{if } n =_0 2m \\ -m & \text{if } n =_0 2m - 1 \end{cases}.$$

where m and $-m$ are seen as rational numbers.
On the other hand, the sequence $(\alpha_i - 1)$ is a sequence of numbers in $\{-1, 0, 1\}$ and each of these sequences α represents the real number in $[-1, 1]$ given by $\sum_{i=1}^{\infty}(\alpha_i - 1)\frac{1}{2^i}$.

Thus to represent a real number we consider type one objects f, where $f(0)$ is any natural number and for $n \geq 1$, $f(n)$ is a number in $\{0, 1, 2\}$. We can then extend this to any function f of type 1 via the transformation $f \mapsto \widetilde{f}$:

$$\widetilde{f}(0) := f(0) \text{ and for } n \geq 1,\ \widetilde{f}(n) := \begin{cases} 0 & \text{if } f(n) = 0 \\ 1 & \text{if } f(n) \text{ odd} \\ 2 & \text{if } f(n) \text{ even} \end{cases}.$$

$\widetilde{f}$ represents an unique real number as described above.
The statement $f^1 \in \mathbb{R}$ is an universal one: $\forall n \in \mathbb{N}\,(f(n+1) \in \{0, 1, 2\})$. Clearly, for each function f of type 1, $\widetilde{f} \in \mathbb{R}$.

The predicates for equality and inequality for the signed-digit representation can be written as:

$$
\begin{aligned}
f =_{\mathbb{R}} g &:\equiv \forall i \in \mathbb{N} \left(\left| int(f(0)) - int(g(0) + \sum_{k=1}^{i+2} (\widetilde{f}(k) - \widetilde{g}(k)) \cdot 2^{-k} \right| <_{\mathbb{Q}} \frac{1}{2^i} \right) \\
f \leq_{\mathbb{R}} g &:\equiv \forall i \in \mathbb{N} \left(int(f(0)) + \sum_{k=1}^{i+2} \widetilde{f}(k) \cdot 2^{-k} <_{\mathbb{Q}} int(g(0)) + \sum_{k=1}^{i+2} \widetilde{g}(k) \cdot 2^{-k} + \frac{1}{2^i} \right) \\
f <_{\mathbb{R}} g &:\equiv \exists i \in \mathbb{N} \left(int(f(0)) + \sum_{k=1}^{i+2} \widetilde{f}(k) \cdot 2^{-k} + \frac{1}{2^i} \leq_Q int(g(0)) + \sum_{k=1}^{i+2} \widetilde{g}(k) \cdot 2^{-k} \right).
\end{aligned}
$$

Note that these relations have the correct complexity: $=_{\mathbb{R}}$ and $\leq_{\mathbb{R}}$ are Π^0_1 statements while the sentence $<_{\mathbb{R}}$ is Σ^0_1.

If n is a natural number, then we see n and $-n$ as real numbers represented by

$$
\begin{aligned}
(n)_{\mathbb{R}} &:= \langle 2n, 1, 1, \cdots \rangle, \\
(-n)_{\mathbb{R}} &:= \langle 2n-1, 1, 1 \cdots \rangle.
\end{aligned}
$$

The signed-digit representation is, perhaps, less intuitive than other representations, like the usual Cauchy-sequence representation. Nevertheless, this representation is preferred here since it mixes well with the notion of majorizability. In order to extend the bounded functional interpretation to new base types, done in the next section, it is important to choose a representation that satisfies the following property: there exists a function $g : \mathbb{N} \to \mathbb{N}$ such that, for any representation f of a real number in $[-n, n]$, we have $f(i) \leq g(n)$, for all $i \in \mathbb{N}$. This property is not satisfied if one represents the real numbers by Cauchy sequences of rationals due to the fact that a representation of a rational may be very large. (It is possible to circumvent this problem by bounding the representation of rationals – see definition 4.24 in [25].)

Proposition 3.1. *Consider the function $g : \mathbb{N} \to \mathbb{N}$ defined by $g(n) = 2n+3$ for any $n \in \mathbb{N}$. Then, HA^ω proves that, for any $n \in \mathbb{N}$, if f is a representation of a real number in $[-n, n]$ then $\forall i \in \mathbb{N}\,(f(i) \leq g(n))$.*

Proof. Since f is a representation of a real number, $\forall i \in \mathbb{N}\,(f(i+1) \leq 2)$. Hence it suffices to see that $f(0) \leq 2n+3$. From $f \leq_{\mathbb{R}} n$ and $-n \leq_{\mathbb{R}} f$, it follows

$$
\begin{aligned}
int(f(0)) + \sum_{k=1}^{2} (f(k) - 1) \cdot 2^{-k} &<_{\mathbb{Q}} n + 1 \\
int(f(0)) + \sum_{k=1}^{2} (f(k) - 1) \cdot 2^{-k} &>_{\mathbb{Q}} -n - 1,
\end{aligned}
$$

which implies,

$$int(f(0)) \leq_{\mathbb{Q}} int(f(0)) + \sum_{k=1}^{2}(f(k)-1)\cdot 2^{-k} + 1 <_{\mathbb{Q}} n+2$$
$$int(f(0)) \geq_{\mathbb{Q}} int(f(0)) + \sum_{k=1}^{2}(f(k)-1)\cdot 2^{-k} - 1 >_{\mathbb{Q}} -n-2.$$

If $int(f(0)) \geq_{\mathbb{Q}} 0$, then (seeing $int(f(0))$ as a natural number since it is a positive integer) $f(0) = 2int(f(0)) <_0 2n+4$ and if $int(f(0)) <_{\mathbb{Q}} 0$, then $int(f(0)) = -k$, with $k \in \mathbb{N}$ and $f(0) = 2k-1 <_0 2n+3$, concluding the proof. □

Due to this majorizability property, we choose to adopt the signed-digit representation for the real numbers. Still, it is easier to define arithmetical operations and to compute them using the Cauchy-sequence representation. The next result, which we do not prove (see [8, theorem 27]) states that there is an effective way to switch back-and-forth between these two representations. Thus, to do computations with real numbers represented by the signed-digit representation one can translate them to the Cauchy-sequence representation, make the computations and then translate them back to the signed-digit representation. Furthermore, the usual properties for real numbers showed for the Cauchy-sequence representation (e.g. [25, section 4.1]) still hold true for the signed-digit representation.

Proposition 3.2. HA^ω *proves that there is an effective translation between the signed-digit representation and the Cauchy-sequence representation. Furthermore, the arithmetic relations* $=$, $<$ *and* $\leq$ *are provably preserved by the translation.*

3.2 Formal theories

To apply functional interpretations to concrete case analyses is necessary to have a formal setting in which, in principle, an ordinary mathematical proof could be formalized.
Several metatheorems exist in the context of the monotone functional interpretation that guarantee *a priori* the extractability of certain computational bounds if the analyzed proof is formalized in a certain axiomatic system ([22], [15]). The formal systems usually considered $\mathcal{A}^\omega[\cdots]$, of classical analysis augmented with certain mathematical structures are very strong and consist of PA^ω together with dependent choice for all types, which in turn implies countable choice and gives full arbitrary comprehension over the natural numbers. Thus, by identifying the subsets of $\mathbb{N}$ with their characteristic functions, full second order arithmetic is contained in those systems. For the interpretation of the axiom schema of dependent choice, one must go beyond the primitive recursive functionals of Gödel's T and make use of Spector's bar recursors [43].

However, since higher forms of comprehension and strong principles of choice are not used in an essential way in most proofs analyzed (so far, the analysis of Baillon's theorem [28] seems to be the exception), bar recursive functionals rarely appear in proof mining.
For the applications shown here, there will be no need to consider bar-recursion and, in fact, the main point of the next chapter is to give a theoretical reason behind the elimination of such expectation in proofs that make use of a certain weak sequential compactness argument. Thus in the theories described below we exclude the axiom of dependent choice. For an explanation on how the bounded functional interpretation can be extended to full second order arithmetic see the doctoral thesis of Patrícia Engrácia [8].

We start by extending our finite types to include a new base type X that stands for an abstract space.

Definition 3.3. Let $\mathcal{T}^{\mathrm{X}}$ be the set of all finite types with additional base type X, which is defined inductively by:

(i) $0, X \in \mathcal{T}^{\mathrm{X}}$ (the ground or base types);

(ii) if $\rho, \sigma \in \mathcal{T}$, then $\rho \to \sigma \in \mathcal{T}$.

Here we will only focus on the cases where X is a bounded metric space, but one could be more specific with X standing for normed, Hilbert, hyperbolic, CAT(0) spaces, etc..
The idea of considering new abstract types, first done in [22] for bounded metric spaces, was of great importance to the proof mining program. This allows to state theorems no longer restricted to "computable" or "representable" spaces only and gave sense to the extraction of computable bounds for many arbitrary spaces. Furthermore, better quantitative statements can be formulated if one is unburdened by the need of working with the representatives of the mathematical objects and works with the actual objects.

The finite types in $\mathcal{T}$, i.e. the ones in which X does not appear, are called arithmetic types. For $\rho \in \mathcal{T}^{\mathrm{X}}$, the finite type denoted by $\widehat{\rho} \in \mathcal{T}$ is the arithmetic type obtained by replacing all the occurrences of X in ρ by 0. Obviously, arithmetic types are left unchanged by this transformation.

For $\rho \in \mathcal{T}^{\mathrm{X}}$, the majorizability relations $x \trianglelefteq_\rho y$, are now between elements x of type ρ and elements y of type $\widehat{\rho}$, in particular majorants are always of an arithmetic type. The concrete description of these notions is given below and extends the previous one in the sense that they coincide for finite types in $\mathcal{T}$.

We will extend our formal setting by introducing axioms characterizing the abstract space (possibly after considering suitable quantitative *moduli* functionals). In order to ensure that a extraction theorem still holds true for the resulting extended theory there are some requirements:

(a) the axioms used to axiomatize X (possibly with additional quantitative moduli) have a functional interpretation;

(b) all the introduced functionals have effective majorants.

Condition (a) follows from the fact that all the axioms are given by universal statements (with bounded matrices)

The theory $\mathsf{PA}^{\omega}_{\trianglelefteq}[X, d_X, a_X, b_X]$:

To axiomatize the structure of bounded metric spaces (X, d), we denote by $\mathcal{L}^{\omega,\mathrm{X}}_{\trianglelefteq}$ the natural extension of the language $\mathcal{L}^{\omega}_{\trianglelefteq}$ to include the new types in $\mathcal{T}^{\mathrm{X}}$ together with three new constants: d_X of type $X \to (X \to 1)$, a_X of type X and b_X of type 0. The main idea is that d_X stands for the metric function on X, a_X is a reference point of the abstract space X and b_X is a bound on the diameter of the space.

We can extend the notion of majorizability to the new types $\mathcal{T}^{\mathrm{X}}$ with the following axioms and rules:

(M1) $\forall n^0 \, \forall m^0 \; (n \trianglelefteq_0 m \leftrightarrow n \leq_0 m)$,

(M2) $\forall x^X \, \forall n^0 \; (x \trianglelefteq_X n \to d_X(x, a_X) \leq_{\mathbb{R}} (n)_{\mathbb{R}})$,

(M3) For each $\rho, \sigma \in \mathcal{T}^{\mathrm{X}}$, $\forall x^{\rho\to\sigma} \, \forall y^{\widehat{\rho\to\sigma}}$
$x \trianglelefteq_{\rho\to\sigma} y \to \forall u^{\rho} \forall v^{\widehat{\rho}} \, (u \trianglelefteq_{\rho} v \to xu \trianglelefteq_{\sigma} yv) \wedge \forall v^{\widehat{\rho}}, v'^{\widehat{\rho}} \, (v \trianglelefteq_{\widehat{\rho}} v' \to yv \trianglelefteq_{\widehat{\sigma}} yv')$

and finally the majorizability rules,

$$\mathsf{RL}_1 : \quad \frac{A_{\mathrm{bd}} \to d_X(p, a_X) \leq_{\mathbb{R}} (n)_{\mathbb{R}}}{A_{\mathrm{bd}} \to p \trianglelefteq_X n}$$

and

$$\mathsf{RL}_2 : \quad \frac{A_{\mathrm{bd}} \wedge u \trianglelefteq_{\rho} v \to tu \trianglelefteq_{\sigma} qv \qquad A_{\mathrm{bd}} \wedge v \trianglelefteq_{\widehat{\rho}} v' \to qv \trianglelefteq_{\widehat{\sigma}} qv'}{A_{\mathrm{bd}} \to t \trianglelefteq_{\rho\to\sigma} q}$$

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 and v, v' are of type $\widehat{\rho}$. The variables u, v, v' don't appear free in the conclusion.

Note that, in (M3) and RL_2, the type $\widehat{\rho \to \sigma}$ is the same as the type $\widehat{\rho} \to \widehat{\sigma}$. The majorizability notion at the base type X is usually defined in terms of the structure one is considering (using the metric function when it is a metric space, the norm when it is a normed space, etc.). However, since the definition is usually given by an universal formula on must divide it into an axiom for the direct implication and a rule for the reverse one, similar to the arrow types $\rho \to \sigma$, in view of the reverse implication not having a bounded functional interpretation.

By also considering the following axioms, we obtained the theory $\mathsf{PA}^{\omega}_{\trianglelefteq}[X, d_X, a_X, b_X]$ as an extention of PA^{ω}:

(R) $\forall x^X \, \forall y^X \; (d_X(x, y) \in \mathbb{R})$,

(1) $\forall x^X \; (d_X(x, x) =_{\mathbb{R}} 0_{\mathbb{R}})$,

(2) $\forall x^X \, \forall y^X \; (d_X(x, y) =_{\mathbb{R}} d_X(y, x))$,

(3) $\forall x^X \, \forall y^X \, \forall z^X \; (d_X(x, y) \leq_{\mathbb{R}} d_X(x, z) +_{\mathbb{R}} d_X(z, y))$,

(4) $\forall x^X \; (d_X(x, a_X) \leq_{\mathbb{R}} (b_X)_{\mathbb{R}})$.

Regarding the treatment of equality, we still only consider primitive equality at type 0 and the equality $=_X$ between elements of type X is defined by

$$x =_X y \;:\equiv\; d_X(x, y) =_{\mathbb{R}} 0_{\mathbb{R}}.$$

With equality for higher types defined in a pointwise manner, we promptly deal with extensionality by considering a bounded extensionality rule, BD-ER:

$$\mathsf{BD\text{-}ER}: \quad \frac{A_{\mathrm{bd}} \to s =_{\rho} t}{A_{\mathrm{bd}} \to r[s/x^{\rho}] =_{\sigma} r[t/x^{\rho}]}$$

where A_{bd} is a bounded formula and s^{ρ}, t^{ρ}, r^{σ} are terms of arbitrary types ρ, $\sigma \in \mathcal{T}^{\mathrm{X}}$.
While the axiom of full extensionality is problematic, as we saw before, the interpretation of this weak extensionality rule will follow trivially from the fact that both the premise and the conclusion are (equivalent to) universal formulas with bounded matrices. Furthermore, BD-ER suffices to proving the rule

$$\frac{A_{\mathrm{bd}} \to s =_{\rho} t}{A_{\mathrm{bd}} \to (B[s/x^{\rho}] \to B[t/x^{\rho}])}$$

with B an arbitrary formula and s^{ρ}, t^{ρ} terms free for x^{ρ} in B.
This treatment of extensionality is similar to the one of the monotone functional interpretation where a quantifier-free extensionality rule is considered – see chapters 3 and 17 of [25].

The axioms for bounded universal and bounded existential quantifiers are extended from the ones in $\mathsf{PA}^{\omega}_{\trianglelefteq}$ to all the types in $\mathcal{T}^{\mathrm{X}}$. The terminology of bounded formulas is still used for formulas in which no unbounded quantification appear. Notice that, by (4) and the rule RL_X we have that $\mathsf{PA}^{\omega}_{\trianglelefteq}[X, d_X, a_X, b_X] \vdash \forall x^X \, (x \trianglelefteq_X d_X)$. Thus for any formula A,

$$\forall x^X \, A \leftrightarrow \forall x \trianglelefteq_X b_X \, A,$$

and similarly for the existential quantifier.
For that reason, we can think of quantifications over elements of the bounded space X as bounded ones. In fact, we could have replaced (M2) and RL_1 simply by the equivalence $x \trianglelefteq_X n \leftrightarrow 0 =_0 0$ and all the theoretical results shown here for the metric bounded space would still be true – additionally, the reference point a_X would be pointless. We chose this more elaborate notion of majorizability at the base type X, to ensure a uniformity between bounded and unbounded spaces. In [8, section 4.1.2], Patrícia Engrácia defined a similar extended notion of majorizability for normed spaces where the reference point is the zero element of the space.

Other theories, like of the unbounded metric space – $\mathsf{PA}^{\omega}_{\trianglelefteq}[X, d_X, a_X]$ – and of the normed space – $\mathsf{PA}^{\omega}_{\trianglelefteq}[X, \|\cdot\|_X]$ – , can be given a similar treatment to this. For example, the formal theory used when dealing with unbounded metric spaces is exactly the same minus axiom (4). Although, for the applications in the following sections, we in fact rely on the existence of those formal systems as well, we will not write them here. It is common practice in proof mining to proceed only in a semi-formal way when formalizing the proof being analyzed. Thus, in the end, all this formal machinery is hidden and the end quantitative results can be read by non-logicians. Small observations will still be made regarding formal aspects of the analyzed proofs.

3.3 Metatheorems

The notion of majorizability extended to $\mathcal{T}^{\mathrm{X}}$ still has the basic properties:

Lemma 3.4. *The theory $\mathsf{PA}^{\omega}_{\trianglelefteq}[X, d_X, a_X, b_X]$ proves, for every $\rho \in \mathcal{T}^{\mathrm{X}}$:*

(i) $x \trianglelefteq_\rho y \to y \trianglelefteq_{\widehat{\rho}} y$,

(ii) $x \trianglelefteq_\rho y \wedge y \trianglelefteq_{\widehat{\rho}} z \to x \trianglelefteq_\rho z$.

Proof. The proof in both cases is by induction on the structure of the type ρ. In (i), the base cases 0 and X follow from the reflexivity of $\leq_0$. The induction step follows from (M3) and RL_2. In (ii), the base case $\rho = 0$ reduces to the transitivity of $\leq_0$. For the case X, the result follows from $y \leq_0 z \to (y)_\mathbb{R} \leq_\mathbb{R} (z)_\mathbb{R}$, (M2), the transitive propriety of $\leq_\mathbb{R}$ and finally the application of RL_1. For the induction step, by (M3),

$$x \trianglelefteq_{\rho\to\sigma} y \wedge u \trianglelefteq_\rho v \to xu \trianglelefteq_\sigma yv$$

and

$$y \trianglelefteq_{\widehat{\rho\to\sigma}} z \wedge v \trianglelefteq_{\widehat{\rho}} v' \wedge v' \trianglelefteq_{\widehat{\rho}} v'' \to yv \trianglelefteq_{\widehat{\sigma}} zv' \wedge zv' \trianglelefteq_{\widehat{\sigma}} zv''.$$

By (i), $u \trianglelefteq_\rho v \to v \trianglelefteq_{\widehat{\rho}} v$ and so

$$y \trianglelefteq_{\widehat{\rho\to\sigma}} z \wedge u \trianglelefteq_\rho v \wedge v \trianglelefteq_{\widehat{\rho}} v' \to yv \trianglelefteq_{\widehat{\sigma}} zv \wedge zv \trianglelefteq_{\widehat{\sigma}} zv'.$$

Thus, using the induction hypothesis we conclude,

$$x \trianglelefteq_{\rho\to\sigma} y \wedge y \trianglelefteq_{\widehat{\rho\to\sigma}} z \wedge u \trianglelefteq_\rho v \wedge v \trianglelefteq_{\widehat{\rho}} v' \to xu \trianglelefteq_\sigma zv \wedge zv \trianglelefteq_{\widehat{\sigma}} zv',$$

and the result follows with the rule RL_2. □

As we saw the inequalities between real numbers $\leq_\mathbb{R}$ and $<_\mathbb{R}$ are given by universal and existential formulas, respectively. It will be useful to have an intensional inequality between real numbers that mixes well with the majorizability notion. In the signed-digit representation $x \leq_\mathbb{R} y$ is given by a formula of the form $\forall n^0 A_{\mathrm{qf}}(n, x, y)$, where $A_{\mathrm{qf}}(n, x, y)$ is the quantifier-free formula $int(x(0)) + \sum_{k=1}^{n+2}(xk) \cdot 2^{-k} <_\mathbb{Q} int(y(0)) + \sum_{k=1}^{n+2}(yk) \cdot 2^{-k} + 2^{-n}$.

Definition 3.5. We define the inequality $\trianglelefteq_\mathbb{R}$ between real numbers by

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

where

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

Notice that the flattening of the quantifier-free formula "$x \trianglelefteq_\mathbb{R} y$" is the universal "$x \leq_\mathbb{R} y$", which justifies saying that $\trianglelefteq_\mathbb{R}$ is the intensional version of $\leq_\mathbb{R}$. Next, we give some simple properties relating the relations $<_\mathbb{R}$, $\leq_\mathbb{R}$, $\trianglelefteq_\mathbb{R}$ and $\trianglelefteq_X$.

Lemma 3.6. *For all n^0, z^X and real numbers x, y, the theory $\mathsf{PA}^\omega_\trianglelefteq[X, d_X, a_X, b_X]$ proves:*

(i) $\trianglelefteq_\mathbb{R}$ *is transitive*

(ii) $x <_\mathbb{R} y \to x \trianglelefteq_\mathbb{R} y$ *and* $x \trianglelefteq_\mathbb{R} y \to x \leq_\mathbb{R} y$

(iii) $z \trianglelefteq_X n \leftrightarrow d_X(z, a_X) \trianglelefteq_\mathbb{R} (n)_\mathbb{R}$

In light of this lemma, in the analysis of mathematical proofs we can circumvent the high complexity of the inequalities between real numbers by replacing them with this intensional version in the appropriate way. In fact, we have for all $x, y \in \mathbb{R}$:

$$\begin{array}{lcl} x <_{\mathbb{R}} y & \leftrightarrow & \exists k^0 \left(x \trianglelefteq_{\mathbb{R}} y - \dfrac{1}{k+1} \right) \\ x \leq_{\mathbb{R}} y & \leftrightarrow & \forall k^0 \left(x \trianglelefteq_{\mathbb{R}} y + \dfrac{1}{k+1} \right) \\ x =_{\mathbb{R}} y & \leftrightarrow & \forall k^0 \left(|x-y| \trianglelefteq_{\mathbb{R}} \dfrac{1}{k+1} \right) \end{array}$$

This way, since the inner formulas are quantifier-free, all the relevant information is placed in the quantifiers. One can then carry out the relevant translations and extractions using this intensional inequality. Finally, in the quantitative result, it is possible to return to the usual inequalities either by using the Lemma 3.6(ii) or by flattening.

The constants a_X and b_X are trivially majorized. Furthermore, the extended constants Π, Σ and $\underline{R}$ are still majorizable and so, to have the condition (b), it just remains to see that the functional d_X is majorized.

From the axioms $(2)-(4)$, we conclude $\forall x^X, y^X \ (d_X(x,y) \leq_{\mathbb{R}} 2(b_X)_{\mathbb{R}})$. From axiom (R) and Proposition 3.1, we conclude

$$\forall i^0 \ (d_X(x,y)i \leq_0 4b_X + 3) \, .$$

Then, by applying the rule RL_2 three times, we conclude $d_X \trianglelefteq \lambda n, m, i \cdot (4b_X + 3)$.

In fact, even in the case of unbounded metric spaces, i.e. without axiom (4), it is easy to see by the same arguments that the function $\lambda n, m, i \cdot (2(n+m)+3)$ majorizes the functional d_X.

Now with the majorizability notion for all finite types in $\mathcal{T}^{\mathrm{X}}$, we extend the bounded functional interpretation by considering the inductive definition 2.11 for all formulas A in $\mathcal{L}^{\omega,\mathrm{X}}_{\trianglelefteq}$. With this extended interpretation it is possible to see that a monotonicity property still holds for the formulas A_U as in Lemma 2.12. Furthermore, the characteristic principles are also extended to encompass the new types. Notice that monotone elements are always of arithmetic type.
Let $\mathsf{PA}^{\omega}_{\trianglelefteq}[X, d_X, a_X, b_X]^+$ denote the theory plus the characteristic principles.

Theorem 3.7. *Let $A(\underline{a})$ be an arbitrary formula of $\mathcal{L}^{\omega,\mathrm{X}}_{\trianglelefteq}$, with free variables $\underline{a}$. If*

$$\mathsf{PA}^{\omega}_{\trianglelefteq}[X, d_X, a_X, b_X]^+ \vdash A(\underline{a}),$$

then there exists a tuple of closed monotone terms $\underline{t}$ of $\mathcal{L}_{\trianglelefteq}^{\omega,\mathrm{X}}$, which can be extracted from a proof of $A(\underline{a})$, such that

$$\mathsf{PA}^{\omega}_{\trianglelefteq}[X, d_X, a_X, b_X] \vdash \widetilde{\forall} \underline{w} \forall \underline{a} \trianglelefteq \underline{w} \widetilde{\forall} \underline{x} (A(\underline{a}))_U(\underline{x}, \underline{twx}).$$

The proof is done by induction on the length of the derivation of the hypothesis. In fact, it is essentially the same proof as in the original soundness theorem for bounded functional interpretation, with the exception that one does need to verify the interpretation of the new axioms and rules. Since the axioms governing the extended majorizability are (equivalent to) universal formulas they have a trivial bounded functional interpretation. The interpretation of the rules RL_1, RL_2 are derivable by applying those same rules. Also, one easily sees that the axioms $(1)-(4)$ and (R) are universal sentences, and thus are trivially interpreted in $\mathsf{PA}^{\omega}_{\trianglelefteq}[X, d_X, a_X, b_X]$.

As the intended structures for this theory are the bounded metric spaces, we have that the (flattening of the) conclusion will hold true in any nonempty bounded metric space (X, d), where d_X is interpreted by the metric function d, b_X by some natural number bounding the diameter of the space and a_X by some point of the space. Notice that being majorized is independent on the choice of the reference point a_X, although the actual majorants may depend on this choice.

As a corollary, we have the following extraction theorems that operate at the level of types usually encountered during concrete case analyses:

Corollary 3.8. *Let A be a existential formula of $\mathcal{L}_{\trianglelefteq}^{\omega,\mathrm{X}}$. If*

$$\mathsf{PA}^{\omega}[X, d_X, a_X, b_X]^{+} \vdash \forall k^0 \widetilde{\forall} f^1 \exists n^0 A(k, f, n),$$

then there is a closed monotone term ϕ of type $0 \to (1 \to 0)$ such that

$$\mathsf{PA}^{\omega}_{\trianglelefteq}[X, d_X, a_X, b_X] \vdash \forall k \widetilde{\forall} f \exists n \leq \phi(k, f)\, A(k, f, n).$$

Sometimes we may wish to consider some premises and, in those cases, the interpretation of such implication may reveal stronger results by weakening the implicative premise. The next result covers this case.

Corollary 3.9. *Let A be an existential formula and B an universal formula of $\mathcal{L}_{\trianglelefteq}^{\omega,\mathrm{X}}$. If*

$$\mathsf{PA}^{\omega}[X, d_X, a_X, b_X]^{+} \vdash \forall k^0 \widetilde{\forall} f^1 \left(\forall m^0 B(k, f, m) \to \exists n^0 A(k, f, n)\right),$$

then there are closed monotone terms ϕ and ψ of type $0 \to (1 \to 0)$ such that

$$\mathsf{PA}^{\omega}_{\trianglelefteq}[X, d_X, a_X, b_X] \vdash \forall k \widetilde{\forall} f \left(\forall m \leq \psi(k, f)\, B(k, f, m) \to \exists n \leq \phi(k, f)\, A(k, f, n)\right).$$

Chapter 4

Proof mining

Proof mining is a research program focused on using proof-theoretical techniques to analyze ordinary mathematical proofs in order to obtain new information. This idea can be traced back to the 1950's, when Georg Kreisel suggested to *unwind proofs* [9][36][37] under the general question

What more do we know if we have proved a theorem by restricted means than if we merely know that it is true?'

Eventually the term "unwinding of proofs" evolved into the more pleasing "proof mining", by suggestion of Dana Scott. In the last twenty five years, the proof mining program has been substantially developed, mainly by the work of Kohlenbach and his collaborators, with a great number of results analyzing proofs from various areas of mathematics such as approximation theory, ergodic theory, fixed point theory, optimization theory and the theory of partial differential equations. In the proof mining program, functional interpretations are used as tools to extract constructive (i.e. computational) information from given ordinary mathematical proofs. While the standard interpretation used in proof mining is Kohlenbach's monotone functional interpretation, one of the objectives of this doctoral dissertation is to investigate the application of the bounded functional interpretation in concrete cases of proof mining. Thus, all the quantitative results obtained were guided by this functional interpretation.

For the results that follow, the author was guided by an unofficial and simple procedure of analyzing proofs:

1. Understanding the proof:

In this first step it is required that one completely understands all the inner works of the proof. This is essential since not only it may allow to carry out simplifications but also because the original argument usually guides the proof of the final quantitative result. At this point, it is important to look for the logical principles used in the original proof. While the common mathematician is most of the time oblivious to the logical strength of the principles used in his proofs, for a logician these are of great relevance and for someone carrying out a proof mining they provide an insight of the relevant theories in which such a proof could be formalized. Furthermore, in this step, one can *a priori* know the functional complexity of the bounds that will be eventually extracted.

2. Structure and formula interpretation:
While the first step already provides a good understanding of the structure of the proof, it is important to clearly layout the structure of the proof in a lemma-by-lemma formulation. Having this, one can start interpreting the relevant formulas. To this effect, it is necessary to have in mind a suitable formal setting that includes all the premises of the theorem in a quantitative form and in which all of the proof can be formalized – the theory in Chapter 3 is an example of one such theory.

3. Extraction and final version:
By having the correct interpretation of the relevant formulas, it is now clear which quantitative information can be extracted. The original proof is then adapted to show the validity of the extracted bounds. In fact, these two steps – extraction and verification – are usually intertwined as one goes back-and-forth in the verification of the quantitative bounds to see what they should be. In the end, there must be a clear formulation of the quantitative information and of its verification.

In the end of the mining, all the proof-theoretical machinery used can be concealed to ensure that the final quantitative result is just a piece of ordinary mathematics. This may be relevant depending on the prospective audience and allows such results to reach both logicians and non-logicians. Notice that in step 2, we only think of a formal setting instead of actually doing the work of completely formalize the proof as we can work only with a semi-formal proof and avoid all the trivial but tiresome formalizations.

The soundness theorem and the extraction theorem for the bounded functional interpretation refers to the formula interpretation given by definition 2.11. However, sometimes a clever use of the characteristic principles together with logical and easy mathematical simplifications can result in better $\widetilde{\forall}\exists$ - formulas. Formulas which are better in the sense that the relevant variables are of lower types. However one can then be left wondering what difference does exist between the extracted information and the *official* quantitative information of the formal interpretation. Of course, if such equivalence is derived in the theory then having the quantitative information for one implies the possibility of having the quantitative information for the other; it is just a question of going through the quantitative proof of the equivalence. However, more can be said. By looking at the proof of the soundness theorem, it is clear that the characteristic principles are given a trivial interpretation and logical and easy mathematical simplifications still hold in the verifying theory. So we are free to use the characteristic principles as a replacement to the formal interpretation of a formula. In the end, the extracted information will only differ from the information obtained through the formal interpretation by trivial computations. We exemplify this with the following examples.

Consider A and B two formulas in the language of arithmetic in all finite types. Assume that we wish to give a bounded functional interpretation to the conjunction formula $A \wedge B$ and already know that A^U is $\widetilde{\forall} x \widetilde{\exists} y A_U(x, y)$ and B^U is $\widetilde{\forall} w \widetilde{\exists} z B_U(w, z)$ – we omit the fact that the variables may be tuples.
If we replace the interpretation of A and of B directly in the conjunction $A \wedge B$ and simply use logic we arrive at:

$$
\begin{gathered}
A \wedge B \\
\left(\widetilde{\forall} x \widetilde{\exists} y A_U(x, y)\right) \wedge \left(\widetilde{\forall} w \widetilde{\exists} z B_U(w, z)\right) \\
\widetilde{\forall} x, w \widetilde{\exists} y, z \ (A_U(x, y) \wedge B_U(w, z))
\end{gathered}
$$

and an extraction for such a formula would yield monotone terms t_1 and t_2 such that

$$
\widetilde{\forall} x, w \left(A_U(x, t_1(x, w)) \wedge B_U(w, t_2(x, w))\right). \tag{4.1}
$$

However, if we follow the formal interpretation of the formula we have to recall that the conjunction is a defined logical symbol and the interpretation goes like this:

$$
\begin{gathered}
A \wedge B \\
\neg \left(\neg A \vee \neg B\right) \\
\neg \left(\neg \widetilde{\forall} x \widetilde{\exists} y A_U(x, y) \vee \neg \widetilde{\forall} w \widetilde{\exists} z B_U(w, z)\right) \\
\neg \left(\widetilde{\forall} f \widetilde{\exists} x \widetilde{\exists} x' \trianglelefteq x \neg A_U(x', fx') \vee \widetilde{\forall} g \widetilde{\exists} w \widetilde{\exists} w' \trianglelefteq w \neg B_U(w', gw')\right) \\
\neg \left(\widetilde{\forall} f, g \widetilde{\exists} x, w \left(\widetilde{\exists} x' \trianglelefteq x \neg A_U(x', fx') \vee \widetilde{\exists} w' \trianglelefteq w \neg B_U(w', gw')\right)\right) \\
\widetilde{\forall} \mathcal{H}_1, \mathcal{H}_2 \widetilde{\exists} f, g \widetilde{\exists} f' \trianglelefteq f \widetilde{\exists} g' \trianglelefteq g \ \left(\widetilde{\forall} x' \trianglelefteq \mathcal{H}_1(f', g') A_U(x', f'x') \wedge \widetilde{\forall} w' \trianglelefteq \mathcal{H}_2(f', g') B_U(w', g'w'))\right)
\end{gathered}
$$

and an extraction for such a formula would give monotone terms T_1 and T_2 such that

$$
\begin{aligned}
&\widetilde{\forall} \mathcal{H}_1, \mathcal{H}_2 \widetilde{\exists} f \trianglelefteq T_1(\mathcal{H}_1, \mathcal{H}_2) \widetilde{\exists} g \trianglelefteq T_2(\mathcal{H}_1, \mathcal{H}_2) \\
&\qquad \left(\widetilde{\forall} x \trianglelefteq \mathcal{H}_1(f, g) A_U(x, fx) \wedge \widetilde{\forall} w \trianglelefteq \mathcal{H}_2(f, g) B_U(w, gw))\right)
\end{aligned} \tag{4.2}
$$

where we used the monotonicity of the matrix in the existentially quantified variables and replaced f', g', x' and w' by f, g, x and w, respectively.

Having the monotone terms t_1 and t_2 satisfying the formula 4.1 it is not hard to see that the terms $T_1 := \lambda\mathcal{H}_1, \mathcal{H}_2, x.\, t_1(x, \mathcal{O}_w)$ and $T_2 := \lambda\mathcal{H}_1, \mathcal{H}_2, w.\, t_2(\mathcal{O}_x, w)$ satisfy 4.2, where $\mathcal{O}_x$ and $\mathcal{O}_w$ denote some monotone term of the same type of x and of type w, respectively – e.g. the functional of the appropriate type that at the final base type yields 0. The other way is also easy as we can take the particular instance of 4.2 when the functionals $\mathcal{H}_1$ and $\mathcal{H}_2$ are constant to conclude:

$$\widetilde{\forall} x, w \widetilde{\exists} f' \trianglelefteq T_1(\lambda f, g.\, x, \lambda f, g.\, w) \widetilde{\exists} g' \trianglelefteq T_2(\lambda f, g.\, x, \lambda f, g.\, w) \\ \Big(\widetilde{\forall} x' \trianglelefteq x A_U(x', f'x') \wedge \widetilde{\forall} w' \trianglelefteq w B_U(w', g'w'))\Big)$$

and in particular,

$$\widetilde{\forall} x, w \widetilde{\exists} f' \trianglelefteq T_1(\lambda f, g.\, x, \lambda f, g.\, w) \widetilde{\exists} g' \trianglelefteq T_2(\lambda f, g.\, x, \lambda f, g.\, w)\ \ (A_U(x, f'x) \wedge B_U(w, g'w)))\,.$$

By monotonicity of the formulas A_U and B_U in the second variable we obtain:

$$\widetilde{\forall} x, w\ \ (A_U(x, T_1(\lambda f, g.\, x, \lambda f, g.\, w)x) \wedge B_U(w, T_2(\lambda f, g.\, x, \lambda f, g.\, w)w))\,,$$

showing that 4.1 is satisfied with the monotone terms $t_1(x, w) := T_1(\lambda f, g.\, x, \lambda f, g.\, w)x$ and $t_2(x, w) := T_2(\lambda f, g.\, x, \lambda f, g.\, w)w$.

We will now look at another example, the interpretation of the formula $(A_1 \vee A_2) \to B$. Assume that we already have the interpretation of the formulas A_1, A_2 and B, given respectively by $\widetilde{\forall} x_1 \widetilde{\exists} y_1 (A_1)_U(x_1, y_1)$, $\widetilde{\forall} x_2 \widetilde{\exists} y_2 (A_2)_U(x_2, y_2)$ and $\widetilde{\forall} w \widetilde{\exists} z B_U(w, z)$.
If we simply input the interpretations of the formulas into the implication and use the characteristic principles we obtain:

$$(A_1 \vee A_2) \to B$$

$$\Big(\widetilde{\forall} x_1 \widetilde{\exists} y_1 (A_1)_U(x_1, y_1)\ \vee\ \widetilde{\forall} x_2 \widetilde{\exists} y_2 (A_2)_U(x_2, y_2)\Big) \to \widetilde{\forall} w \widetilde{\exists} z B_U(w, z)$$

By using monotone choice twice and the monotonicity property, we get

$$\Big(\widetilde{\exists} f_1 \widetilde{\forall} x_1\, (A_1)_U(x_1, f_1 x_1)\ \vee\ \widetilde{\exists} f_2 \widetilde{\forall} x_2 (A_2)_U(x_2, f_2 x_2)\Big) \to \widetilde{\forall} w \widetilde{\exists} z B_U(w, z)$$

By classical logic, this is equivalent to

$$\widetilde{\forall} f_1, f_2, w \widetilde{\exists} x_1, x_2, z\ (((A_1)_U(x_1, f_1 x_1)\ \vee\ (A_2)_U(x_2, f_2 x_2)) \to B_U(w, z))\,.$$

The extraction for this formula would yield monotone terms t_1, t_2 and t_3 such that

$$\widetilde{\forall} f_1, f_2, w \widetilde{\exists} x_1 \trianglelefteq t_1(f_1, f_2, w) \widetilde{\exists} x_2 \trianglelefteq t_2(f_1, f_2, w) \\ ((A_1)_U(x_1, f_1 x_1)\ \vee\ (A_2)_U(x_2, f_2 x_2)) \to B_U(w, t_3(f_1, f_2, w)). \tag{4.3}$$

On the other hand, to carry out the interpretation using the clauses in 2.11 first we must recall that the implication is a defined notion. The interpretation is as follows:

$$
\begin{gathered}
(A_1 \vee A_2) \to B \\
\neg(A_1 \vee A_2) \vee B \\
\neg\left(\widetilde{\forall} x_1 \widetilde{\exists} y_1 (A_1)_U(x_1, y_1) \vee \widetilde{\forall} x_2 \widetilde{\exists} y_2 (A_2)_U(x_2, y_2)\right) \vee \widetilde{\forall} w \widetilde{\exists} z B_U(w, z) \\
\neg\left(\widetilde{\forall} x_1, x_2 \widetilde{\exists} y_1, y_2 (A_1)_U(x_1, y_1) \vee (A_2)_U(x_2, y_2)\right) \vee \widetilde{\forall} w \widetilde{\exists} z B_U(w, z)
\end{gathered}
$$

$$
\begin{aligned}
&\widetilde{\forall} \mathcal{H}_1, \mathcal{H}_2 \widetilde{\exists} x_1, x_2 \\
&\qquad \left(\widetilde{\exists} x_1' \trianglelefteq x_1 \widetilde{\exists} x_2' \trianglelefteq x_2 \neg\left((A_1)_U(x_1', \mathcal{H}_1(x_1', x_2')) \vee (A_2)_U(x_2', \mathcal{H}_2(x_1', x_2'))\right)\right) \vee \widetilde{\forall} w \widetilde{\exists} z B_U(w, z) \\
&\widetilde{\forall} \mathcal{H}_1, \mathcal{H}_2, w \widetilde{\exists} x_1, x_2, z \\
&\qquad \left(\widetilde{\exists} x_1' \trianglelefteq x_1 \widetilde{\exists} x_2' \trianglelefteq x_2 \neg\left((A_1)_U(x_1', \mathcal{H}_1(x_1', x_2')) \vee (A_2)_U(x_2', \mathcal{H}_2(x_1', x_2'))\right) \vee B_U(w, z)\right)
\end{aligned}
$$

Using the definition of the implication, this can be written as:

$$
\begin{aligned}
&\widetilde{\forall} \mathcal{H}_1, \mathcal{H}_2, w \widetilde{\exists} x_1, x_2, z \\
&\qquad \left(\widetilde{\exists} x_1' \trianglelefteq x_1 \widetilde{\exists} x_2' \trianglelefteq x_2 \left((A_1)_U(x_1', \mathcal{H}_1(x_1', x_2')) \vee (A_2)_U(x_2', \mathcal{H}_2(x_1', x_2'))\right) \to B_U(w, z)\right).
\end{aligned}
$$

The extraction for this formula gives monotone terms T_1, T_2 and T_3 satisfying

$$
\begin{aligned}
&\widetilde{\forall} \mathcal{H}_1, \mathcal{H}_2, w \widetilde{\exists} x_1 \trianglelefteq T_1(\mathcal{H}_1, \mathcal{H}_2, w) \widetilde{\exists} x_2 \trianglelefteq T_2(\mathcal{H}_1, \mathcal{H}_2, w) \\
&\qquad ((A_1)_U(x_1, \mathcal{H}_1(x_1, x_2)) \vee (A_2)_U(x_2, \mathcal{H}_2(x_1, x_2))) \to B_U(w, T_2(\mathcal{H}_1, \mathcal{H}_2, w)),
\end{aligned}
\tag{4.4}
$$

where we used the monotonicity of the matrix in the existentially quantified variables and replaced the x_1' and x_2' by x_1 and x_2, respectively.

With the monotone terms t_1, t_2 and t_3 satisfying the formula 4.3 it is possible to define terms satisfying 4.4. To that effect, for given functional $\mathcal{H}_1$ and $\mathcal{H}_2$ define the functional $f_1 := \lambda x.\, \mathcal{H}_1(x, \mathcal{O}_{x_2})$ and $f_2 := \lambda x.\, \mathcal{H}_2(\mathcal{O}_{x_1}, x)$. It is not hard to see that the monotone terms $T_1 := \lambda \mathcal{H}_1, \mathcal{H}_2, w.\, t_1(f_1, f_2, w)$, $T_2 := \lambda \mathcal{H}_1, \mathcal{H}_2, w.\, t_2(f_1, f_2, w)$ and $T_3 := \lambda \mathcal{H}_1, \mathcal{H}_2, w.\, t_3(f_1, f_2, w)$ satisfy 4.4.

For the other direction, we argue as before by considering a particular instance of 4.4. Consider the case where the functional $\mathcal{H}_1$ and $\mathcal{H}_2$ are constant. Then,

$$
\begin{aligned}
&\widetilde{\forall} \mathcal{H}_1, \mathcal{H}_2, w \widetilde{\exists} x_1 \trianglelefteq T_1(\mathcal{H}_1, \mathcal{H}_2, w) \widetilde{\exists} x_2 \trianglelefteq T_2(\mathcal{H}_1, \mathcal{H}_2, w) \\
&\qquad ((A_1)_U(x_1, \mathcal{H}_1(x_1, x_2)) \vee (A_2)_U(x_2, \mathcal{H}_2(x_1, x_2))) \to B_U(w, T_2(\mathcal{H}_1, \mathcal{H}_2, w)),
\end{aligned}
$$

Let us now make a brief comment on the choice of interpreting a piece of implicative mathematics either by formalizing it as a rule of the system or as a provable implication. Assume that in some mathematical proof of a theorem the author states "if $\widetilde{A}$, ..., then $\widetilde{B}$". Are we to look at formalizing a proof of $\widetilde{A} \to \widetilde{B}$ or to add a quantitative version of $\widetilde{A}$ as a postulate to our theory and show that $\widetilde{B}$ is derivable? This choice has an impact on the quantitative information one can extract. Assume that the interpretations of the formulas are, respectively, of the form $\widetilde{\forall} x \widetilde{\exists} y A(x,y)$ and $\widetilde{\forall} w \widetilde{\exists} z B(w,z)$, where A and B are bounded formulas. Then, we can interpret the implication in the following way:

$$\widetilde{A} \to \widetilde{B}$$
$$\widetilde{\forall} x \widetilde{\exists} y A(x,y) \to \widetilde{\forall} w \widetilde{\exists} z B(w,z)$$
$$\widetilde{\forall} f, w \left(\widetilde{\forall} x \widetilde{\exists} y \trianglelefteq f x A(x,y) \to \widetilde{\exists} z B(w,z) \right)$$

An extraction for this formula would yield monotone functionals ϕ_1 and ϕ_2 such that

$$\widetilde{\forall} f, w \left(\widetilde{\forall} x \trianglelefteq \phi_1(f,w) \widetilde{\exists} y \trianglelefteq f x A(x,y) \to \widetilde{\exists} z \trianglelefteq \phi_2(f,w) B(w,z) \right).$$

This, in turn, yields the weaker statement

$$\widetilde{\forall} f \left(\widetilde{\forall} x \widetilde{\exists} y \trianglelefteq f x A(x,y) \to \widetilde{\forall} w \widetilde{\exists} z \trianglelefteq \phi_2(f,w) B(w,z) \right).$$

On the other hand, if we only consider $\widetilde{A}$ as a postulate, instead of as an implicative premise, we would have that, for some functional ψ

$$\mathsf{T} + \widetilde{\forall} x \widetilde{\exists} y \trianglelefteq \psi(x) A(x,y) \vdash \widetilde{B},$$

and we could extract from the proof a monotone term Φ – in the extended language with ψ – satisfying

$$\widetilde{\forall} w \widetilde{\exists} z \trianglelefteq \Phi(w) B(w,z).$$

Notice, that even if we could go back to an implicative formulation at this point we would only get

$$\widetilde{\forall} x \widetilde{\exists} y \trianglelefteq \psi(x)\, A(x,y) \to \widetilde{\forall} w \widetilde{\exists} y \trianglelefteq \Phi(w)\, B(w,z),$$

which contains less information than the quantitative version of the implication, as we lost the information exhibited by ϕ_1.

Since the proof analyzed is the same, notice that we have $\Phi(w) = \phi_2(\psi, w)$. Although in most cases the property being proved can be formalized in an implication form, if in the end we are only interested in the quantitative information for $\widetilde{B}$, then the interpretation of the rule is a better avenue. Nevertheless, whenever possible, interpreting the implication would provide a more comprehensive quantitative information of the proof being analyzed. In the interpretation of the implication, we obtain a quantitative result stating how much of (the quantitative version of) $\widetilde{A}$ is needed to conclude a particular instance of (the quantitative version of) $\widetilde{B}$. There are however cases where we are left without the choice of using the implicative option. For example, when dealing with a discontinous functional f, the implication "$a = b \to f(a) = f(b)$" cannot be derived in the theory. If it were, its quantitative analysis would force continuity on the functional f. This phenomenon is not specific to the bounded functional interpretation and, in fact, this was already commented by Kohlenbach (e.g., at the end of [30]) regarding the monotone functional interpretation. There are several instances of proof mining where it was only possible to use the rule option (e.g., [15, 22, 25]).

4.1 Simple arguments in proof mining

In this section, we will look at four simple but interesting arguments frequently used in mathematical proofs and present their quantitative version. The first one is the simple conjunction of two properties that hold true after a certain point – e.g. the conjunction of two convergence statements. After, we will look at a variation of the statement that any bounded and decreasing sequence of real numbers is a Cauchy sequence. Then, we see a quantitative treatment of lim sup by using rational approximations. Finally we will see how to give a correct quantitative interpretation of a discussion by cases.

4.1.1 Metastability

In general, it is not possible to guarantee the extractability of quantitative information bounding the existential quantifier in a provable Π^0_3-formula (with bounded matrix). For example, consider the formula $\forall n \exists m\, (T(n,n,m) \vee \forall p \neg T(n,n,p))$, where T is the primitive recursive Kleene T-predicate. This statement is equivalent to the provable Π^0_3-formula $\forall n \exists m \forall p\, (T(n,n,m) \vee \neg T(n,n,p))$ and is clear that no bounding information can be obtained on m, since it would entail the decidability of the halting problem.

However, by the soundness theorem, it is possible to extract information for the translated formula. We will now show how the bounded functional interpretation is applied to Π_3 formulas.

Let $A_{\mathrm{bd}}(x, y, z)$ be a bounded formula. Guided by the characterizing principles, we have:

$$\forall x \exists y \forall z A_{\mathrm{bd}}(x, y, z),$$

by the MAJ^ω this is equivalent to

$$\forall x \widetilde{\exists} y \exists y' \trianglelefteq y \widetilde{\forall} z \forall z' \trianglelefteq z A_{\mathrm{bd}}(x, y', z').$$

Using the contra-positive of $\mathsf{BC}^\omega_{\mathsf{bd}}$ we obtain

$$\forall x \widetilde{\exists} y \widetilde{\forall} z \exists y' \trianglelefteq y \widetilde{\forall} z' \trianglelefteq z \forall z'' \trianglelefteq z' A_{\mathrm{bd}}(x, y', z''),$$

which in turn is logically equivalent to

$$\forall x \widetilde{\exists} y \widetilde{\forall} z \exists y' \trianglelefteq y \forall z' \trianglelefteq z A_{\mathrm{bd}}(x, y', z').$$

If we abbreviate $B_{\mathrm{bd}}(x, y, z) \equiv \exists y' \trianglelefteq y \forall z' \trianglelefteq z A_{\mathrm{bd}}(x, y', z')$ then, by the contra-positive of $\mathsf{mAC}^\omega_{\mathsf{bd}}$ together with MAJ^ω, we get

$$\widetilde{\forall} x \widetilde{\forall} f \forall x' \trianglelefteq x \widetilde{\exists} y \widetilde{\forall} z \trianglelefteq f y B_{\mathrm{bd}}(x', y, z),$$

and, lastly, by the $\mathsf{BC}^\omega_{\mathsf{bd}}$ we obtain the final formula

$$\widetilde{\forall} x \widetilde{\forall} f \widetilde{\exists} y \forall x' \trianglelefteq x \widetilde{\exists} y' \trianglelefteq y \widetilde{\forall} z \trianglelefteq f y' B_{\mathrm{bd}}(x', y', z)$$

In essence, overlooking the quantifications "$\forall x' \trianglelefteq x \widetilde{\exists} y' \trianglelefteq y$", the interpretation makes the inner universal quantifier "$\forall z$" become bounded by considering a functional of higher type. We call this final statement the metastable version of the original property.
In the interpretation above, all the generality considered brings additional troublesome steps which in most cases do not appear. The additional quantifiers only appear if one is working with non-monotone quantifications at the start, which is not the case in most situations. Frequently, we consider Π^0_3 formulas. There we quantify over natural numbers and so, over monotone objects. This simplifies the interpretation as then there is no need to consider the axiom MAJ^ω (as well as $\mathsf{BC}^\omega_{\mathsf{bd}}$), which is responsible for the appearance of new quantifiers at several steps of the translation.

To exemplify how the translation is in fact very easy in practice and how the interpretation operates in a more common formula, we consider the Cauchy property.
Let (x_n) be a Cauchy sequence of real numbers. Then,

$$\forall k \in \mathbb{N} \exists n \in \mathbb{N} \forall i, j \geq n \left(|x_i - x_j| \leq \frac{1}{k+1} \right).$$

We can replace the inequality between reals by the intensional $\trianglelefteq_{\mathbb{R}}$ and work with a quantifier-free matrix, instead of an universal one. With this consideration, we can apply the contra-positive of the $\mathsf{mAC}^\omega_{\mathsf{bd}}$ and then go back to the extensional inequality to obtain the metastable version of the Cauchy property:

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

where $[n, fn]$ denotes the set of natural numbers $\{n, n+1, \cdots, fn\}$. Notice, that in the case that for some natural number n, $fn < n$, the result is trivially true.
We call the metastable version of the Cauchy property of (x_n), simply the metastablility property of (x_n). The interval of natural numbers $[n, fn]$ is a region of metastability for (x_n). In this sense, the Cauchy property is translated into a statement about regions of metastability for the sequence. In fact, due to the generality of the function f, the metastability of a sequence (x_n) is equivalent to its Cauchy property. It is trivial that the Cauchy property implies metastability. To recover the Cauchy property from the metastability of the sequence one must rely on classical logic and the axiom of choice, which makes the proof ineffective. This is consistent with the fact that we are always guaranteed to be able to extract a bound on n in the metastability of (x_n), but in general we cannot convert it into a bound for the Cauchy property (even for recursive sequences).

The metastability property was already known to logicians as Kreisel's no-counterexample interpretation. The term "*metastability*" was introduced by Terence Tao in his blog and is nowadays used to designate these properties.

Bounds on the existential quantifier of $\forall k \exists n \forall m \geq n A_{\mathrm{bd}}(k, n)$ are called *rates of convergence*, and bounds on the existential quantifier of $\forall k \widetilde{\forall} f \exists n \forall m \in [n, fn] A_{\mathrm{bd}}(k, m)$ are called bounds on the metastability. To simplify the analyses, we always assume that these functions are monotone, instead of considering *ad hoc* modifications.

- A monotone function $\chi : \mathbb{N} \to \mathbb{N}$ is a rate of convergence, if:

$$\forall k \forall m \geq \chi(k)\, A_{\mathrm{bd}}(k, m)$$

- A monotone function $\phi : \mathbb{N} \times \mathbb{N}^{\mathbb{N}} \to \mathbb{N}$ is a bound on the metastability, if:

$$\forall k \widetilde{\forall} f \exists n \leq \phi(k, f) \forall m \in [n, fn]\, A_{\mathrm{bd}}(k, m)$$

While carrying out a quantitative analysis, it is common to encounter an argument that takes in the conjunction of two convergence statements. If at that point the quantitative information extracted was in the form of rates of convergence, then by simply considering its maximum we obtain a rate for the conjunction:
If $\chi_1, \chi_2 : \mathbb{N} \to \mathbb{N}$ are such that

$$\forall k \forall m \geq \chi_1(k)\, A_1(k, m) \text{ and } \forall k \forall m \geq \chi_2(k)\, A_2(k, m),$$

then

$$\forall k \forall m \geq \chi(k)\, (A_1(k, m) \wedge A_2(k, m)),$$

where $\chi(k) := \max\{\chi_1(k), \chi_2(k)\}$.

However, as we said, in most cases such extractions are not possible and we must instead work with bounds on the metastability. We will now see how to consider a rate of convergence together with a bound on the metastability.

Proposition 4.1. *Let $\chi : \mathbb{N} \to \mathbb{N}$ and $\phi : \mathbb{N} \times \mathbb{N}^{\mathbb{N}} \to \mathbb{N}$ be monotone functions such that*

(i) $\forall k \forall m \geq \chi(k)\, A_1(k, m)$ *and*

(ii) $\forall k \widetilde{\forall} f \exists n \leq \phi(k, f) \forall m \in [n, fn]\, A_2(k, m)$,

where A_1 and A_2 are bounded formulas. Then,

$$\forall k \widetilde{\forall} f \exists n \leq \Phi(k, f) \forall m \in [n, fn]\ (A_1(k, m) \wedge A_2(k, m)),$$

where $\Phi(k, f) := \max\{\chi(k)\, ,\, \phi(k, \lambda m.\, (f(\max\{\chi(k), m\})))\}$.

Proof. Let $k \in \mathbb{N}$ and a monotone function $f : \mathbb{N} \to \mathbb{N}$ be given.
Apply (ii) to k and the monotone function $g(m) := f(\max\{\chi(k), m\})$ to conclude the existence of $n_1 \leq \phi(k, g)$ such that

$$\forall m \in [n_1, gn_1]\, A_1(k, m).$$

With $n := \max\{\chi(k), n_1\}$, we have $n \leq \Phi(k, f)$ and $m \in [n, fn]$ implies $m \geq \chi(k)$ and $m \in [n_1, gn_1]$. Hence,

$$\forall m \in [n, fn]\ (A_1(k, m) \wedge A_2(k, m)).$$

□

The next result deals with the general conjunction of two bounds of metastability.

Proposition 4.2. *Let $\phi_1, \phi_2 : \mathbb{N} \times \mathbb{N}^{\mathbb{N}} \to \mathbb{N}$ be monotone functions such that*

(i) $\forall k \widetilde{\forall} f \exists n \leq \phi_1(k, f) \forall m \in [n, fn]\, A_1(k, m)$ *and*

(ii) $\forall k \widetilde{\forall} f \exists n \leq \phi_2(k, f) \forall m \in [n, fn]\, A_2(k, m)$,

where A_1 and A_2 are bounded formulas. Then,

$$\forall k \widetilde{\forall} f \exists n \leq \Phi(k, f) \forall m \in [n, fn]\ (A_1(k, m) \wedge A_2(k, m)),$$

where $\Phi(k, f) := \max\{\phi_1(k, f_1)\, ,\, \phi_2(k, f_2[\phi_1(k, f_1)])\}$ *with* $f_1(m) := f(\max\{m, \phi_2(k, f_2[m])\})$ *and* $f_2[m](n) := f(\max(m, n))$.

Proof. Let $k \in \mathbb{N}$ and a monotone function $f : \mathbb{N} \to \mathbb{N}$ be given. Then the functions f_1 and f_2 are also monotone.
By (i) applied to k and f_1, there is $n_1 \leq \phi_1(k, f_1)$ such that

$$\forall m \in [n_1, f_1(n_1)]\, A_1(k, m).$$

Now, by (ii) applied to k and $f_2[n_1]$, there is $n_2 \leq \phi_2(k, f_2[n_1])$ satisfying

$$\forall m \in [n_2, f_2[n_1](n_2)]\, A_2(k, m).$$

We will now check that, $n := \max(n_1, n_2)$ satisfies the desired conclusion.
First of all, we have $n \leq \max\{\phi_1(k, f_1)\,,\, \phi_2(k, f_2[\phi_1(k, f_1)])\} = \Phi(k, f)$.
Secondly, by the definition of n and the monotonicity of the functions f and ϕ_2, we have

$$\begin{aligned} &[n, fn] \subset [n_1, f(\max\{n_1, \phi_2(k, f_2[n_1])\})] = [n_1, f_1(n_1)] \text{ and} \\ &[n, fn] \subset [n_2, fn] = [n_2, f_2[n_1](n_2)] \end{aligned}$$

Thus, for $m \in [n, fn]$, we have $A_1(k, m) \wedge A_2(k, m)$ as we wanted. □

In fact, since we could have changed the order of (i) and (ii), a better bound is the minimum of those two possibilities.

Corollary 4.3. *Let $\phi_1, \phi_2 : \mathbb{N} \times \mathbb{N}^{\mathbb{N}} \to \mathbb{N}$ be monotone functions such that*

(i) $\forall k \widetilde{\forall} f \exists n \leq \phi_1(k, f) \forall m \in [n, fn]\, A_1(k, m)$ *and*

(ii) $\forall k \widetilde{\forall} f \exists n \leq \phi_2(k, f) \forall m \in [n, fn]\, A_2(k, m)$,

where A_1 and A_2 are bounded formulas. Then,

$$\forall k \widetilde{\forall} f \exists n \leq \Phi(k, f) \forall m \in [n, fn]\, \left(A_1(k, m) \wedge A_2(k, m)\right),$$

where $\Phi(k, f) := \min\{\Phi_1(k, f), \Phi_2(k, f)\}$, with:

$$\begin{gathered} \Phi_1(k, f) := \max\{\phi_1(k, f_1)\,,\, \phi_2(k, f_2[\phi_1(k, f_1)])\}, \\ \Phi_2(k, f) := \max\{\phi_2(k, f_3)\,,\, \phi_1(k, f_2[\phi_2(k, f_3)])\}, \\ f_1(m) := f(\max\{m, \phi_2(k, f_2[m])\}),\ f_2[m](n) := f(\max(m, n)) \\ \textit{and } f_3(m) := f(\max\{m, \phi_1(k, f_2[m])\}). \end{gathered}$$

A rate of convergence is a particular bound on the metastability that does not depend on the function f. Hence, the conjunction of a rate of convergence with a bound on the metastability is a particular case of the conjunction of two bounds on metastability. In fact, Proposition 4.1 can be derived from Corollary 4.3, by assuming that one of the bounds on the metastability does not depend on f: in that case, $\Phi_1(k, f) = \Phi_2(k, f)$ and coincide with the bound in Proposition 4.1. Furthermore, if both bounds on the metastability are rates of convergence, i.e. if $\phi_1(k, f) = \chi_1(k)$ and $\phi_2(k, f) = \chi_2(k)$, then we recover the simplest bound, $\max\{\chi_1(k), \chi_2(k)\}$.

4.1.2 Monotone convergent sequences

The term metastability was introduced by Terence Tao when studing the *infinite convergence principle* from a finitary perspective [46, Section 1.3].

Theorem 4.4 (Infinite Convergence Principle)**.** *Every decreasing sequence of non-negative real numbers (s_n) is convergent.*

Instead of dealing with a convergence statement, which requires an explicit reference to the $\inf\{s_n\}$, we consider the simpler statement that (s_n) is a Cauchy sequence.
As we saw, in its metastable version, the Cauchy property is

$$\forall k \in \mathbb{N} \widetilde{\forall} f : \mathbb{N} \to \mathbb{N} \exists N \in \mathbb{N}\, \forall i, j \in [N, fN] \left(|s_i - s_j| \leq \frac{1}{k+1} \right).$$

A quantitative version of this result was previously obtained by Kohlenbach [25]. A slight adaptation of that result to the setting of the bounded functional interpretation is shown next:

Proposition 4.5. *Let (s_n) be a decreasing sequence of real numbers and $D \in \mathbb{N}$ such that for all $n \in \mathbb{N}$, $0 \leq s_n \leq D$. Then,*

$$\forall k \in \mathbb{N} \widetilde{\forall} f : \mathbb{N} \to \mathbb{N} \exists N \leq f^{(D(k+1))}(0) \forall i, j \in [N, fN] \left(|s_i - s_j| \leq \frac{1}{k+1} \right),$$

where $f^{(r)}$ stands for the consecutive iteration, i.e. $f^{(0)}(n) = n$ and $f^{(r+1)}(n) = f\left(f^{(r)}(n)\right)$.

We can even consider non-monotone functions f, in which case the bound $f^{(D(k+1))}(0)$ is replaced by $\max\{f^{(r)}(0) \,|\, r \leq D(k+1)\}$. Notice also that, beside k and f, the dependence of the bound on the sequence is only in the upper bound D, not on the sequence itself.

Obviously, this result extends to bounded sequences which are only decreasing after a certain order, i.e. for (s_n) a bounded sequence we have

$$\exists M \in \mathbb{N} \forall m \geq M \;(s_{m+1} \leq s_m) \;\to\; \forall k \in \mathbb{N} \exists N \in \mathbb{N} \forall i, j \geq N \left(|s_i - s_j| \leq \frac{1}{k+1} \right).$$

The analysis of this statement is similar to Proposition 4.5 and we have the following quantitative version.

Proposition 4.6. *Let (s_n) be a sequence of real numbers and $D \in \mathbb{N}$ such that for all $n \in \mathbb{N}$, $0 \leq s_n \leq D$. Then,*

$$\begin{aligned}&\forall k \in \mathbb{N} \forall M \in \mathbb{N} \widetilde{\forall} f : \mathbb{N} \to \mathbb{N} \\ &\quad\left(\forall m \geq M \,(s_{m+1} \leq s_m) \;\to\; \exists N \leq \phi(k, M, f)\, \forall i, j \in [N, fN] \left(|s_i - s_j| \leq \frac{1}{k+1} \right) \right),\end{aligned}$$

with $\phi(k, M, f) := \max\{M, f^{(D(k+1))}(M)\}$.

One can prove the result above following the same arguments used in proving Proposition 4.5. However, this result also follows directly from the Proposition 4.5. In fact, let $k, M \in \mathbb{N}$ and a monotone function f be given. The relevant case is when $f(M) \geq M$. If the premise of the implication in Proposition 4.6 holds, then the sequence $w_n := s_{n+M}$ satisfy the conditions of Proposition 4.5. In particular, we can consider the (monotone) function $g(n) := f(n+M)-M$ and conclude

$$\exists N' \leq g^{(D(k+1))}(0)\, \forall i, j \in [N', g(N')] \left(|s_{i+M} - s_{j+M}| \leq \frac{1}{k+1} \right).$$

Finally the result follows for $N := N' + M$ by showing $g^{(D(k+1))}(0) + M = f^{(D(k+1))}(M)$.

Clearly, the statement that the sequence is eventually decreasing can be equivalently written as

$$\exists M \in \mathbb{N} \forall k \in \mathbb{N} \forall m \geq M \left(s_{m+1} \leq s_m + \frac{1}{k+1} \right).$$

In an attempt to weaken this condition we could think of considering

$$\forall k \in \mathbb{N} \exists M \in \mathbb{N} \forall m \geq M \left(s_{m+1} \leq s_m + \frac{1}{k+1} \right). \tag{4.5}$$

However, since we are considering the "decreasing property" with a possibility of error $\frac{1}{k+1}$, having the property stated only for m and $m+1$ we leave open the possibility that the sequence fluctuates too much, in which case the conclusion may not hold. An example is the following sequence. For each $n \in \mathbb{N}$, take $k \in \mathbb{N}$ and $i \leq k+1$ such that $n = \frac{(k+3)k}{2} + i$ and define $s_n = \frac{i}{k+1}$. This sequence alternates indefinitely between 0 and 1 and so it is not a Cauchy sequence, even though it satisfies (4.5).

Instead, a similar notion of decreasing may be used here,

$$\forall k \in \mathbb{N} \exists M \in \mathbb{N} \forall n, m \geq M \left(n < m \to s_m \leq s_n + \frac{1}{k+1} \right). \tag{4.6}$$

In this way, we ask that not only the successor be below the previous term, with error $\frac{1}{k+1}$, but also all the terms after that. Any bounded sequence of non-negative real numbers (s_n) under this assumption is a Cauchy sequence, i.e.

$$\begin{aligned} &\forall k \in \mathbb{N} \exists M \in \mathbb{N} \forall n, m \geq M \left(n < m \to s_m \leq s_n + \frac{1}{k+1} \right) \\ &\to \forall k \in \mathbb{N} \exists N \in \mathbb{N} \forall i, j \geq N \left(|s_i - s_j| \leq \frac{1}{k+1} \right). \end{aligned}$$

This statement is more general as it additionally guarantees that certain non-decreasing sequences, e.g. the sequence defined by $s_n = 1 + \frac{(-1)^n}{n+1}$, are Cauchy sequences.
It is even possible argue the instance $\frac{1}{k+1}$ of the Cauchy property if for some $k' \in \mathbb{N}$ the premise holds for the instance $\frac{1}{k+1} - \frac{1}{k'+1}$.

Proposition 4.7. *Let (s_n) be a bounded sequence of non-negative real numbers. Then, for all $k \in \mathbb{N}$, if*

$$\exists k' \in \mathbb{N} \exists M \in \mathbb{N} \forall n, m \geq M \left(n < m \to s_m \leq s_n + \frac{1}{k+1} - \frac{1}{k'+1} \right),$$

then

$$\exists N \in \mathbb{N} \forall i, j \geq N \left(|s_i - s_j| \leq \frac{1}{k+1} \right).$$

Proof. Let $k \in \mathbb{N}$ be given such that for some k', $M \in \mathbb{N}$ we have

$$\forall n, m \geq M \left(n < m \to s_m \leq s_n + \frac{1}{k+1} - \frac{1}{k'+1} \right). \tag{4.7}$$

Assume, towards a contradiction, that for all $N \in \mathbb{N}$,

$$\exists i, j \geq N \left(i < j \wedge |s_i - s_j| > \frac{1}{k+1} \right). \tag{4.8}$$

By induction we can show

$$\forall n \widetilde{\exists} f, g \left(f(n) < g(n) < f(n+1) \wedge s_{f(n)} > s_{g(n)} + \frac{1}{k+1} \geq s_{f(n+1)} + \frac{1}{k'+1} \right). \tag{4.9}$$

Apply (4.8) to $N = M$ to obtain $i_0, j_0 \geq M$ such that $i_0 < j_0$ and $|s_{i_0} - s_{j_0}| > \frac{1}{k+1}$.
We must have $s_{i_0} \geq s_{j_0}$, otherwise

$$s_{j_0} > s_{i_0} + \frac{1}{k+1} > s_{i_0} + \frac{1}{k+1} - \frac{1}{k'+1},$$

contradicting (4.7). Hence $s_{i_0} > s_{j_0} + \frac{1}{k+1}$.
Apply (4.8) to $N = j_0 + 1$ to obtain $i_1, j_1 > j_0$ such that $i_1 < j_1$ and $|s_{i_1} - s_{j_1}| > \frac{1}{k+1}$. Since $M \leq j_0 < i_1$, by (4.7), we have

$$s_{i_1} \leq s_{j_0} + \frac{1}{k+1} - \frac{1}{k'+1}.$$

Thus $s_{j_0} + \frac{1}{k+1} \geq s_{i_1} + \frac{1}{k'+1}$ and we can take $f(0) = i_0$, $f(1) = i_1$, $g(0) = j_0$ and $g(1) = j_1$.
The inductive step is argued in an identical way and concludes the proof of (4.9).
From (4.9) it follows

$$\forall n \widetilde{\exists} f : \mathbb{N} \to \mathbb{N} \left(s_{f(0)} > s_{f(n)} + \frac{n}{k'+1} \right),$$

contradicting the fact that (s_n) is a bounded sequence. □

We will give a quantitative version of Proposition 4.7. We start by explaining the interpretation. Saying that (s_n) is a bounded sequence of non-negative real numbers is not an universal statement. Instead, we work with an actual natural number D such that $0 \leq s_n \leq D$, for all $n \in \mathbb{N}$. On the other hand, Proposition 4.7 can be equivalently expressed by using the intensional (quantifier-free) inequality between real numbers $\trianglelefteq_{\mathbb{R}}$:
For all $k \in \mathbb{N}$, if

$$\exists k', M \in \mathbb{N}\, \forall n, m \geq M \left(n < m \to s_m \trianglelefteq_{\mathbb{R}} s_n + \frac{1}{k+1} - \frac{1}{k'+1} \right),$$

then

$$\exists N \in \mathbb{N}\, \forall i, j \geq N \left(|s_i - s_j| \trianglelefteq_{\mathbb{R}} \frac{1}{k+1} \right).$$

By monotone choice, the conclusion of the implication above is equivalent to

$$\widetilde{\forall} f : \mathbb{N} \to \mathbb{N}\, \exists N \in \mathbb{N}\, \forall i, j \in [N, fN] \left(|s_i - s_j| \trianglelefteq_{\mathbb{R}} \frac{1}{k+1} \right).$$

Then, by classical logic, the statement is equivalent to

$$\begin{aligned} \forall k, k', M \in \mathbb{N} \widetilde{\forall} f : \mathbb{N} \to \mathbb{N} \Bigg[\forall n, m \geq M \left(n < m \to s_m \trianglelefteq_{\mathbb{R}} s_n + \frac{1}{k+1} - \frac{1}{k'+1} \right) \\ \to \exists N \in \mathbb{N} \forall i, j \in [N, fN] \left(|s_i - s_j| \trianglelefteq_{\mathbb{R}} \frac{1}{k+1} \right) \Bigg] \end{aligned}$$

We are then guaranteed to be able to extract bounds on the values of n, m and N in terms of k, k', M and f, by analyzing the proof of Proposition 4.7. By flattening we return to the usual extensional inequalities between real numbers and have the next quantitative result.

Proposition 4.8. *Let (s_n) be a sequence of real numbers and $D \in \mathbb{N}$ such that for all $n \in \mathbb{N}$, $0 \leq s_n \leq D$. Then, for all $k, k', M \in \mathbb{N}$ and monotone function $f : \mathbb{N} \to \mathbb{N}$, if*

$$\forall n, m \in [M, \phi_1(k, k', M, f)] \left(n < m \to s_m \leq s_n + \frac{1}{k+1} - \frac{1}{k'+1} \right) \tag{4.10}$$

then

$$\exists N \leq \phi_2(k, k', M, f)\, \forall i, j \in [N, fN] \left(|s_i - s_j| \leq \frac{1}{k+1} \right), \tag{4.11}$$

where $\phi_1(k, k', M, f) := f(\phi_2(k, k', M, f))$ and $\phi_2(k, k', M, f) := \max\{M, f^{(D(k''+1))}(M)\}$ with $k'' := \max\{k, k'\}$.

Proof. Let k, k', $M \in \mathbb{N}$ and a monotone function be given. Assume $f(M) \geq M$, otherwise the result is trivially true with $N = M$. Notice that, from the monotonicity of f, we have $f^{(i+1)}(M) \geq f^{(i)}(M)$, for each $i \in \mathbb{N}$. Assume that (4.10) holds. Then, we also have:

$$\forall n, m \in [M, \phi_1(k, k', M, f)] \left(n < m \to s_m \leq s_n + \frac{1}{k+1} - \frac{1}{k''+1} \right). \tag{4.12}$$

Suppose towards a contradiction that (4.11) is false. Then, for all $N \leq f^{(D(k''+1))}(M)$,

$$\exists i, j \in [N, fN] \left(i < j \wedge |s_i - s_j| > \frac{1}{k+1} \right). \tag{4.13}$$

Write $R := D(k''+1)$. We define sequences $(i_n)_{n \leq R}$ and $(j_n)_{n \leq R}$ in the following way.

<u>i_0 and j_0:</u>
By (4.13) with $N = M$, there are $i, j \in [M, f(M)]$ such that

$$i < j \wedge |s_i - s_j| > \frac{1}{k+1}.$$

Consider i_0, j_0 to be one such pair of i, j. Notice that we must have $s_{i_0} \geq s_{j_0}$, otherwise

$$s_{j_0} > s_{i_0} + \frac{1}{k+1} > s_{i_0} + \frac{1}{k+1} - \frac{1}{k''+1},$$

contradicting (4.10). We conclude $s_{i_0} > s_{j_0} + \frac{1}{k+1}$.

<u>i_{r+1} and j_{r+1}, with $r < R$:</u>
Assume, for the inductive definition, that we have i_r,$j_r \in [f^{(r)}(M), f^{(r+1)}(M)]$ satisfying

$$s_{i_r} > s_{j_r} + \frac{1}{k+1}.$$

Applying (4.13) with $N = f^{(r+1)}(M)$, we have $i_{r+1}, j_{r+1} \in [f^{(r+1)}(M), f^{(r+2)}(M)]$ satisfying

$$i_{r+1} < j_{r+1} \wedge |s_{i_{r+1}} - s_{j_{r+1}}| > \frac{1}{k+1},$$

and, in a similar way to before, we conclude $s_{i_{r+1}} > s_{j_{r+1}} + \frac{1}{k+1}$.

Since for all $r < R$, by the definition, $j_r \leq i_{r+1}$ and $\frac{1}{k+1} - \frac{1}{k''+1} \geq 0$, we conclude

$$s_{i_{r+1}} \leq s_{j_r} + \frac{1}{k+1} - \frac{1}{k''+1}.$$

If $j_r = i_{r+1}$ the inequality is trivial, and if $j_r < i_{r+1}$ it follows from (4.10).
Hence, for all $r < R$, $s_{j_r} + \frac{1}{k+1} \geq s_{i_{r+1}} + \frac{1}{k''+1}$.
By the definition, we conclude for all $r < R$, $s_{i_r} > s_{i_{r+1}} + \frac{1}{k''+1}$. In particular, it follows

$$s_{i_0} > s_{i_R} + \frac{R}{k''+1} \geq \frac{R}{k''+1} = D,$$

contradicting the condition on D. $\square$

We can recover the Proposition 4.5, by noticing that if (s_n) is a decreasing sequence, then the premise of Proposition 4.8 holds with $k' = k$ and $M = 0$ in which case, the bound ϕ_2 from Proposition 4.8 is the same as in Proposition 4.5. In a similar way, we can recover Proposition 4.6. The result is in fact slightly stronger, as it specifies till when we need to have the "decreasing property" to guarantee the (k, f)-instance of the metastability property.

4.1.3 Working with $\limsup$

When analyzing mathematical proofs, it is not uncommon to encounter the need to consider the $\limsup x_n$ for some bounded sequence of real numbers (x_n). However, the existence of $\limsup$ for a bounded sequence (x_n) in general is only guaranteed by using arithmetical comprehension. Thus, an analysis of a mathematical proof making use of a $\limsup$ cannot be formalized in our restricted formal theories and its interpretation, in general, would require the use of bar recursive functionals. Nevertheless, in many practical cases the reference to that *ideal* real number can be replaced by approximations that are still good enough in the sense of still allowing to carry out the main arguments of the proof. Thus, in those situations we can modify the original proof to circumvent the need of arithmetical comprehension and obtain a proof formalizable in the restrictive setting of our formal theories. The computational information extracted from this modified proof will then be given by functionals of Gödel's T. In this section, we will argue how the existence of $\limsup$ can be replaced by a combinatorial argument using rational numbers. In the modified proof, we will be replacing the $\limsup$ by a "good enough" (in the sense of usefulness to the proof) approximation of rational numbers.

Let $N \in \mathbb{N}$ and let (x_n) be a non-negative sequence of real numbers contained in the interval $[0, N]$. The existence of $\limsup x_n$ can be stated as

$$\begin{aligned}&\exists d \in \mathbb{R}\, \forall k \in \mathbb{N}\\ &\left(\forall n \in \mathbb{N} \exists m \geq n \left(x_m \geq d - \frac{1}{k+1}\right) \wedge \exists n' \in \mathbb{N} \forall m' \geq n' \left(x_{m'} \leq d + \frac{1}{k+1}\right)\right).\end{aligned} \tag{4.14}$$

The relevant point is that we can weaken this statement by switching the outermost quantifiers,

$$\forall k \in \mathbb{N}\, \exists d \in \mathbb{R} \left(\forall n \in \mathbb{N} \exists m \geq n \left(x_m \geq d - \frac{1}{k+1} \right) \wedge \exists n' \in \mathbb{N} \forall m' \geq n' \left(x_{m'} \leq d + \frac{1}{k+1} \right) \right). \tag{4.15}$$

We will see that

$$\forall k \in \mathbb{N}\, \exists p < N(k+1) \left(\forall n \in \mathbb{N} \exists m \geq n \left(x_m \geq \frac{p}{k+1} \right) \wedge \exists n' \in \mathbb{N} \forall m' \geq n' \left(x_{m'} \leq \frac{p+1}{k+1} \right) \right), \tag{4.16}$$

which implies (4.15) with $d = \frac{p}{k+1}$. The previous statement (4.16) can be understood in the following way: with $k \in \mathbb{N}$ given, let $p < N(k+1)$ be such that $\frac{p}{k+1} \leq \limsup x_n \leq \frac{p+1}{k+1}$; then, the middle point of $\left[\frac{p}{k+1}, \frac{p+1}{k+1}\right]$ witnesses d in (4.15) for $2k+1$ (and thus, also for k). We will avoid the mention of $\limsup$ by an easy combinatorial argument. We have the following result.

Lemma 4.9. *Let $N \in \mathbb{N}$ and (x_n) be a sequence of real numbers such that for all $n \in \mathbb{N}$, $0 \leq x_n \leq N$. Then, for all $k, n \in \mathbb{N}$ and monotone function $f : \mathbb{N} \to \mathbb{N}$,*

$$\exists p < N(k+1) \left(\exists m \in [n, n+fn] \left(x_m \geq \frac{p}{k+1} \right) \wedge \forall m' \in [n, n+fn] \left(x_{m'} \leq \frac{p+1}{k+1} \right) \right). \tag{4.17}$$

Proof. Suppose towards a contradiction that (4.17) does not hold. Then there exist natural numbers k, n and a monotone function f such that for all $p < N(k+1)$ it holds that

$$\forall m \in [n, n+fn] \left(x_m < \frac{p}{k+1} \right) \vee \exists m' \in [n, n+fn] \left(x_{m'} > \frac{p+1}{k+1} \right). \tag{4.18}$$

In order to additionally see that this result only requires induction for bounded formulas, note that (4.18) implies

$$\forall p < N(k+1) \left(A(p+1) \to A(p) \right),$$

where $A(p)$ is the bounded formula $\forall m \in [n, n+fn] \left(x_m \trianglelefteq_{\mathbb{R}} \frac{p}{k+1} \right)$. Then, one shows by induction that

$$\forall M \in \mathbb{N} \left(\forall p \leq M \left(A(p+1) \to A(p) \right) \to \left(A(M+1) \to A(0) \right) \right).$$

Hence, with $M = N(k+1) - 1$ we conclude that

$$\forall m \in [n, n+fn] \left(x_m \trianglelefteq_{\mathbb{R}} N \right) \to \forall m \in [n, n+fn] \left(x_m \trianglelefteq_{\mathbb{R}} 0 \right).$$

By the first disjunct of (4.18) with $p = N(k+1)-1$, we obtain $\forall m \in [n, n+fn]\,(x_m \trianglelefteq_{\mathbb{R}} N)$. Thus, $\forall m \in [n, n+fn]\,(x_m \trianglelefteq_{\mathbb{R}} 0)$. However, by the second disjunct of (4.18) when $p = 0$, we have $\frac{1}{k+1} \trianglelefteq_{\mathbb{R}} x_{m'}$ for some $m' \in [n, n+fn]$, which gives a contradiction. We conclude that (4.17) holds. $\square$

In particular, from the previous result, we have

$$\begin{aligned} &\forall k, n \in \mathbb{N} \widetilde{\forall} f : \mathbb{N} \to \mathbb{N} \exists p < N(k+1) \\ &\qquad \left(\exists m \geq n \left(x_m \geq \frac{p}{k+1} \right) \wedge \exists n' \in \mathbb{N} \forall m' \in [n', n' + fn'] \left(x_{m'} \leq \frac{p+1}{k+1} \right) \right). \end{aligned} \tag{4.19}$$

By a collection argument, we conclude

$$\begin{aligned} &\forall k \in \mathbb{N} \exists p < N(k+1) \forall n \in \mathbb{N} \widetilde{\forall} f : \mathbb{N} \to \mathbb{N} \\ &\qquad \left(\exists m \geq n \left(x_m \geq \frac{p}{k+1} \right) \wedge \exists n' \in \mathbb{N} \forall m' \in [n', n' + fn'] \left(x_{m'} \leq \frac{p+1}{k+1} \right) \right). \end{aligned} \tag{4.20}$$

In fact, this collection argument is justified by an inductive reasoning proving:

$$\begin{aligned} &\forall r \leq N(k+1)\, \exists n \widetilde{\exists} f : \mathbb{N} \to \mathbb{N}\, \forall p < r \\ &\qquad \left(\forall m \geq n \left(x_m < \frac{p}{k+1} \right) \vee \forall n' \in \mathbb{N} \exists m' \in [n', n' + fn'] \left(x_{m'} > \frac{p+1}{k+1} \right) \right), \end{aligned}$$

under the assumption that (4.20) does not hold. Then, if (4.19) holds, then (4.20) must also hold. Note that the reverse implication is obviously true. By monotone choice, (4.20) is equivalent to the statement (4.16).

We will now see that, with these rational approximations of the lim sup, we can still argue a version of a useful property of the lim sup. With d denoting the $\limsup x_n$, the property one is concerned is

$$\forall k, M, \ell \in \mathbb{N}\, \exists m \geq M \forall n \geq m \left(x_{m+\ell} \geq d - \frac{1}{k+1} \wedge x_n \leq d + \frac{1}{k+1} \right). \tag{4.21}$$

This property is easily shown by recalling the defining properties of d:

(I) $\forall k \in \mathbb{N} \forall n \in \mathbb{N} \exists m \geq n \left(x_m \geq d - \frac{1}{k+1} \right)$;

(II) $\forall k \in \mathbb{N} \exists n \in \mathbb{N} \forall m \geq n \left(x_m \leq d + \frac{1}{k+1} \right)$.

Let $k, M, \ell \in \mathbb{N}$ be given. By (II) applied to k, there is $n_0 \in \mathbb{N}$ such that

$$\forall m \geq n_0 \left(x_m \leq d + \frac{1}{k+1} \right).$$

Now, consider the property (I) with $k \in \mathbb{N}$ and with $n = \max\{M + \ell, n_0 + \ell\}$. There is $m_0 \geq n$ such that $x_{m_0} \geq d - \frac{1}{k+1}$. Define $m_1 = m_0 - \ell$. Since $m_1 \geq n_0$, we also have

$$\forall n \geq m_1 \left(x_n \leq d + \frac{1}{k+1} \right).$$

On the other hand, we have $m_1 \geq M$ and

$$x_{m_1+\ell} = x_{m_0} \geq d - \frac{1}{k+1}.$$

Putting these two statements together, we conclude (4.21).
In the version where d is replaced by rational approximations, it then makes sense that a similar result can be obtained by two applications of Lemma 4.9. The first application would correspond to the instance of (II) to find n_0 and the second application to the use of (I) applied to $n = \max\{M + \ell, n_0 + \ell\}$. However, between applications of Lemma 4.9, we may get different values for the natural number p. This problem can be overcome by a more elaborate argument. The next result is a quantitative version of (4.21).

Lemma 4.10. *Let $N \in \mathbb{N}$ and (x_n) be a sequence of real numbers such that for all $n \in \mathbb{N}$, $0 \leq x_n \leq N$. Let $k, M, \ell \in \mathbb{N}$ and $f : \mathbb{N} \to \mathbb{N}$ be a monotone function. Denote $\mathrm{P} := N(k+1)$. For $i \leq \mathrm{P}$ define $n_i = M + i\ell$ and*

$$r_i := \begin{cases} 0, & i = \mathrm{P} \\ \ell + r_{i+1} + f(n_{i+1} + r_{i+1}), & i < \mathrm{P}. \end{cases}$$

Then

$$\exists p < \mathrm{P} \exists m \in [M, \theta] \forall n \in [m, m + fm] \left(x_{m+\ell} \geq \frac{p}{k+1} \wedge x_n \leq \frac{p+1}{k+1} \right), \tag{4.22}$$

where $\theta = \theta(k, M, \ell, f) := M + (\mathrm{P} - 1)\ell + r_0$.

Proof. Let natural numbers k, M, ℓ and a monotone function f be given. We define, for each $i \leq \mathrm{P}$, the constant functions $g_i := \lambda m.\, r_i$. For each $i \leq \mathrm{P}$, we apply Lemma 4.9 with $k = k$, $f = g_i$ and $n = n_i$. Then, we find, for each $i \leq \mathrm{P}$, $m_i \in [n_i, n_i + r_i]$ and $p_i < \mathrm{P}$ such that $x_{m_i} \geq \frac{p_i}{k+1}$ and $\forall n \in [n_i, n_i + r_i] \left(x_n \leq \frac{p_i+1}{k+1} \right)$. Now, there exists $i_0 \leq \mathrm{P}$ such that $p_{i_0} \leq p_{i_0+1}$, otherwise there would be a sequence of length $\mathrm{P} + 1$ of natural numbers such that $p_B < p_{B-1} < \cdots < p_1 < p_0 < \mathrm{P}$, which is absurd. Define the natural numbers $m := m_{i_0+1} - \ell$ and $p := p_{i_0+1}$. We have that $x_{m+\ell} \geq \frac{p}{k+1}$. To conclude the result it is enough to show that $[m, m+fm] \subseteq [n_{i_0}, n_{i_0} + r_{i_0}]$. Indeed, we would get, for $n \in [m, m+fm]$ that $x_n \leq \frac{p_{i_0}+1}{k+1} \leq \frac{p_{i_0+1}+1}{k+1} = \frac{p+1}{k+1}$. We have that $m = m_{i_0+1} - \ell \geq n_{i_0+1} - \ell = n_{i_0}$ and, since f is monotone, $m + fm \leq m_{i_0+1} + f(m_{i_0+1}) \leq n_{i_0+1} + r_{i_0+1} + f(n_{i_0+1} + r_{i_0+1}) = n_{i_0} + \ell + r_{i_0+1} + f(n_{i_0+1} + r_{i_0+1}) = n_{i_0} + r_{i_0}$. Hence $[m, m+fm] \subseteq [n_{i_0}, n_{i_0} + r_{i_0}]$, which concludes the proof. $\square$

In this thesis, there will be no applications of the results of this subsection, as these results are still part of ongoing work [7]. Thus we restrain from making any comments on how "good enough" these rational approximations are in the sense of an hypothetical modified proof. Recently, in [34], Kohlenbach and Sipoş gave a similar treatment of the lim sup and were able to obtain a quantitative version of Reich's theorem [41] analyzing, via the monotone functional interpretation, a modified proof using what they called ε-lim sup's.

4.1.4 Discussion by cases

Sometimes mathematical proofs follow a discussion by cases and it is important to know how to analyze this type of argument. As we discussed before, when comparing the rule versus its implicative counterpart, it is sometimes relevant to decide whether to do the analysis either by adding the assumptions as axioms to the formal theory (in a correct form) or by consider them as implicative assumptions. In a discussion by cases, we will show that, in non-trival cases, it is necessary to choose the latter. This example gives a small contribution to the understanding of the role of postulates and of implicative assumptions in proof mining.

Assume that, for some statements A and B, we have simultaneously

$$A \to B \tag{4.23}$$

$$\neg A \to B. \tag{4.24}$$

Then, by classical logic, one can derive B. In its quantitative form, the analyses of $A \to B$ and of $\neg A \to B$, should provide us with a quantitative version for B. This is what we will show next.

Suppose that $\widetilde{\forall} x \widetilde{\exists} y A_U(x, y)$ is the interpretation of A and $\widetilde{\forall} w \widetilde{\exists} z B_U(w, z)$ is the interpretation of B. Then the quantitative version of (4.23) can easily be computed:

$$\begin{aligned}
&A \to B\\
&\widetilde{\forall} x \widetilde{\exists} y\, A_U(x,y) \to \widetilde{\forall} w \widetilde{\exists} z\, B_U(w,z)\\
&\widetilde{\exists} f \widetilde{\forall} x \widetilde{\exists} y \trianglelefteq fx\, A_U(x,y) \to \widetilde{\forall} w \widetilde{\exists} z\, B_U(w,z)\\
&\widetilde{\forall} w \widetilde{\forall} f\, \left(\widetilde{\forall} x \widetilde{\exists} y \trianglelefteq fx\, A_U(x,y) \to \widetilde{\exists} z\, B_U(w,z)\right)
\end{aligned} \tag{4.25}$$

The universal quantifier in the antecedent and the existential quantifier in the consequent can be placed outside the implication as existential quantifiers but, we choose to leave the formula written like this.

From a proof of the implication $A \to B$ we would extract monotone bounds ϕ_1 and ϕ_2 satisfying

$$\widetilde{\forall} w \widetilde{\forall} f\, \left(\widetilde{\forall} x \trianglelefteq \phi_1(w,f) \widetilde{\exists} y \trianglelefteq fx\, A_U(x,y) \to \widetilde{\exists} z \trianglelefteq \phi_2(w,f)\, B_U(w,z)\right). \tag{4.26}$$

As for the interpretation of $\neg A \to B$, we have:

$$\begin{aligned}
&\neg A \to B\\
&\neg\widetilde{\forall} x \widetilde{\exists} y\, A_U(x,y) \to \widetilde{\forall} w \widetilde{\exists} z\, B_U(w,z)\\
&\widetilde{\exists} x \widetilde{\forall} y\, \neg A_U(x,y) \to \widetilde{\forall} w \widetilde{\exists} z\, B_U(w,z)\\
&\widetilde{\forall} w \widetilde{\forall} x\, \left(\widetilde{\forall} y\, \neg A_U(x,y) \to \widetilde{\exists} z\, B_U(w,z)\right),
\end{aligned} \tag{4.27}$$

and from a proof of $\neg A \to B$ we can extract monotone bounds ψ_1 and ψ_2 such that

$$\widetilde{\forall} w \widetilde{\forall} x\, \left(\widetilde{\forall} y \trianglelefteq \psi_1(w,x)\, \neg A_U(x,y) \to \widetilde{\exists} z \trianglelefteq \psi_2(w,x)\, B_U(w,z)\right). \tag{4.28}$$

Then, from (4.26) and (4.28), we can compute a bound Θ such that

$$\widetilde{\forall} w \widetilde{\exists} z \trianglelefteq \Theta(w)\, B_U(w,z). \tag{4.29}$$

In fact, $\Theta(w) := \max\{\phi_2(w,f_0)\,,\, \psi_2(w, \phi_1(w,f_0))\}$, where $f_0 := \lambda x.\,(\psi_1(w,x))$, is a monotone bound satisfying (4.29):

Take monotone w arbitrary.

Since ψ_1 is monotone, the same is true for f_0. Apply (4.26) to w and f_0 in order to conclude

$$\widetilde{\forall} x \trianglelefteq \phi_1(w,f_0) \widetilde{\exists} y \trianglelefteq f_0 x\, A_U(x,y) \to \exists z \trianglelefteq \phi_2(w,f_0)\, B_U(w,z).$$

If $\widetilde{\forall} x \trianglelefteq \phi_1(w,f_0) \widetilde{\exists} y \trianglelefteq f_0 x\, A_U(x,y)$ holds, then we conclude the result since $\Theta(w) \geq \phi_2(w,f_0)$. Otherwise, there is a monotone $x_0 \trianglelefteq \phi_1(w,f_0)$ such that

$$\begin{aligned}
&\widetilde{\forall} y \trianglelefteq f_0 x_0\; \neg A_U(x_0,y),\\
&\text{i.e. } \widetilde{\forall} y \trianglelefteq \psi_1(w,x_0)\, \neg A_U(x_0,y).
\end{aligned} \tag{4.30}$$

In this case, we can apply (4.28) to w and x_0 to conclude

$$\widetilde{\exists} z \trianglelefteq \psi_2(w, x_0)\, B_U(w, z).$$

Since ψ_2 is a monotone function and $x_0 \trianglelefteq \phi_1(w, f_0)$, we have $\psi_2(w, x_0) \trianglelefteq \psi_2(w, \phi_1(w, f_0))$ and the result follows.

Notice that from (4.26), we have the weaker statement

$$\widetilde{\forall} x \widetilde{\forall} f \left(\widetilde{\forall} x \widetilde{\exists} y \trianglelefteq f x\, A_U(x, y) \to \widetilde{\exists} z \trianglelefteq \phi_2(w, f)\, B_U(w, z) \right).$$

Since w no longer appears in the antecedent, we have equivalently

$$\widetilde{\forall} f \left(\widetilde{\forall} x \widetilde{\exists} y \trianglelefteq f x\, A_U(x, y) \to \widetilde{\forall} w \widetilde{\exists} z \trianglelefteq \phi_2(w, f)\, B_U(w, z) \right). \tag{4.31}$$

This statement just means that from an analyses of the implications $A \to B$, we have all the information given by the analysis of its "rule counterpart". In fact, by adding a monotone bound α and the axiom $\widetilde{\forall} x \widetilde{\exists} y \trianglelefteq \alpha x\, A_U(x, y)$ to an appropriate theory where the proof of the implication $A \to B$ is formalizable and we can apply the functional interpretation, we could extract a bound θ such that $\widetilde{\forall} w \widetilde{\exists} z \trianglelefteq \theta(w)\, B_U(w, z)$. Of course such a bound could depend on the introduced α and, if the extraction argument is the same, we would have $\theta(w) = \phi_2(w, \alpha)$.
In the same way, we have for the other implication

$$\widetilde{\forall} x \left(\widetilde{\forall} y\, \neg A_U(x, y) \to \widetilde{\forall} w \widetilde{\exists} z \trianglelefteq \psi_2(w, x)\, B_U(w, z) \right). \tag{4.32}$$

and a similar argument could be made regarding the analyses via its rule counterpart: by adding an element u satisfying $\widetilde{\forall} y\, \neg A_U(u, y)$, we could find $\theta(w)$ as $\psi_2(w, u)$.

However, it is not clear how one could join these two analyses together as formally they are working on two contradicting formal theories. Additionally, notice that the bound Θ depends not only on ϕ_2 and ψ_2, but also on the bounds ϕ_1 and ψ_1 which are absent from a "rule" analysis. This is an indication that the mining of a discussion by cases may only be possible via its implicative form.

4.2 Weak compactness

In this section, we will devise a general method for bypassing certain sequential weak compactness arguments in proof mining analyses. We will obtain a quantitative version of this method that works as a blueprint for minings where the method is applicable.

If one uses a functional interpretation to analyze a mathematical proof, a first step is to look for an appropriate formal theory where the proof can be formalized. In order to use sequential weak compactness, it is however necessary to consider systems much stronger than the one exhibit in chapter 3. Sadly the interpretation of those theories goes beyond the primitive recursive functionals of Gödel's T and C. Spector bar recursive functionals are needed [43]. It so happens that, in most concrete cases, it is possible to avoid these strong theories by consider instead modified proofs which rely on weaker logical principles and thus are interpretable using functionals from Gödel's T. For the sequential weak compactness arguments pertinent for our discussion, we will show that it is possible to bypass them by instead using a *false* Heine/Borel compactness principle. This argument is explained in the formal setting of the bounded functional interpretation where it corresponds to an application of the axiom $\mathsf{BC}^{\omega}_{\mathsf{bd}}$. In the proof of the quantitative results, the characteristic principles of the interpretation disappear and, with them, the application of Heine/Borel compactness.

The motivation for this section came from previous studies on a strong convergence result by F.E. Browder carried out by Kohlenbach using the monotone functional interpretation. Even though the proof being analyzed used a sequential weak compactness, in the end of the mining no bar recursive functionals were needed to construct the final bounds. This raised the question of why the troublesome argument was vanishing from the quantitative version of the proof. It turns out that the weak compactness used in the mining was very mild and had a trivial solution. It was not necessary to use any real strength of sequential weak compactness and so no need to use bar recursive functionals.
The work on this section gives a theoretical explanation for this elimination and abstracts a general method for removing sequential weak compactness from proof mining. The results shown here were obtained with Fernando Ferreira and Laurenţiu Leuştean and can be viewed in detail in [12].

4.2.1 Modified Browder's proof

Let X be a Banach space and U a mapping of X into X. The map U is nonexpansive if it does not increase the distance between points

$$\forall x, y \in X \; (\|U(x) - U(y)\| \leq \|x - y\|),$$

and is a strict contraction if for some $k < 1$,

$$\forall x, y \in X \; (\|U(x) - U(y)\| \leq k\|x - y\|).$$

For strict contractions, Banach's fixed-point theorem states that the Picard's iteration, $x_{n+1} = U(x_n)$, converges strongly to the unique fixed point of U. However, for nonexpansive mappings, this sequence doesn't even have to converge. In [6], Browder proved the following strong convergence result for a different iteration:

Theorem 4.11 (Browder [6]). *Let X be a real Hilbert space and $U : X \to X$ a nonexpansive mapping. Assume that C is a bounded closed convex subset of X, that $v_0 \in C$, and that U maps C into itself. For each natural number n, define*

$$U_n(x) := \left(1 - \frac{1}{n+1}\right) U(x) + \frac{1}{n+1} v_0 \tag{4.33}$$

and consider u_n to be the unique fixed point of this strict contraction. Then the sequence (u_n) converges strongly to a fixed point of U in C (the closest one to v_0).

We adapt Browder's proof so that it can be formalized in a theory similar to the one shown in Chapter 3. The main point is that, in this formal setting, we can replace the sequential weak compactness argument in the original proof by an application of the axiom $\mathsf{BC}^{\omega}_{\mathsf{bd}}$, which can be seen as an application of a *false* Heine/Borel compactness principle. An additional troublesome step in the proof was a projection argument which from a logical perspective requires the use of countable choice. However, as was already commented by Kohlenbach in [27], we can instead consider a weaker statement and still be able to prove the Cauchy property of the iteration. While the original projection argument could not be justified by our formal setting, this weaker statement can be proved using only induction. We will see in detail the treatment of the projection argument using the bounded functional interpretation in a following section.

In order to analyze (a modified) Browder's proof we consider a formal theory $\mathcal{T}_B$, which we will now explain.
The theory $\mathcal{T}_B$ follows a similar construction as the theory $\mathsf{PA}^{\omega}[X, d_X, a_X, b_X]$ in Chapter 3. Again we are considering a typed language with two base types 0 and X, where X stands for an abstract Hilbert space. At this point we rely on the established literature and treat inner product spaces as a special case of normed spaces where the parallelogram law holds, as in section 17.3 of [25].
The language includes vector space constants $0_X, +_X, -_X$ and $\cdot_X$ of types X, $X \to (X \to X)$, $X \to X$ and $1 \to (X \to X)$, respectively. They stand, respectively, for the zero vector, the vector addition, symmetric vector and scalar multiplication. Notice that for $\gamma : \mathbb{N} \to \mathbb{N}$ and $x \in X$, the scalar multiplication $\gamma \cdot_X x$ is the scalar multiplication of the real number $\gamma_{\mathbb{R}}$ with the vector x, where $\gamma_{\mathbb{R}}$ is the real number represented by the function γ via the signed-digit representation (Section 3.1). We also add a constant standing for the norm, $\|\cdot\|$, of type $X \to 1$.
We can deal with Browder's proof in a more straightforward way by adding some *ad hoc* constants. We consider a constant v_0 of type X for the given point stated in the theorem, a constant of type $X \to 0$ for the characteristic function of the bounded closed convex subset C and a constant b of type 0 for an upper (positive) bound on the diameter of C. The sequence of fixed points for the strict contractions U_n are given by a constant u of type

$0 \to X$, with $u(n)$ giving the unique fixed point of U_n. To simplify, we write $x \in C$ instead of $C(x) =_0 0$ and u_n instead of $u(n)$.
We follow a treatment of real normed spaces, via an universal axiomatization, as in the doctoral dissertation of Patrícia Engrácia [8]. We can state that the type 1 functional $\|x\|$ is always a representation of a real number by considering the axiom $\forall n^0\, (\|x\|(n) =_0 \|x\|_{\mathbb{R}}(n))$.
With equality between elements of X, $x =_X y$, defined by the universal formula

$$\|x - y\| =_{\mathbb{R}} 0,$$

it was shown in [8] that $=_X$ is indeed an equivalence relation and that it is congruent with respect to the vector space notions.
The inner product functional $\langle \cdot, \cdot \rangle$, of type $X \to (X \to 1)$, is defined by

$$\langle x, y \rangle := \frac{1}{4} \left(\|x + y\|^2 - \|x - y\|^2 \right).$$

We are using a more informal notation above and will do so whenever it is convenient. The axiomatization of normed vector spaces with the so-called paralellogram law,

$$\forall x^X, y^X \, \left(\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2) \right),$$

entails the usual properties for the inner product.

The axioms of $\mathcal{T}_B$ related to its specification to the mining of Browder's proof are the following:

$$\begin{array}{l}
v_0 \in C \\
\forall x, y \in C \ (\|x - y\| \leq (b)_{\mathbb{R}}) \\
\forall x, y \in C \forall \gamma \in [0, 1] \ ((1 - \gamma) \cdot x + \gamma \cdot y \in C) \\
\forall x \in C \ (U(x) \in C) \\
\forall x, y \in X \ (\|U(x) - U(y)\| \leq \|x - y\|) \\
\forall n \in \mathbb{N} \ (u_n \in C) \\
\forall n \in \mathbb{N} \left((1 - \dfrac{1}{n+1}) U(u_n) + \dfrac{1}{n+1} v_0 = u_n \right)
\end{array}$$

We remark that the quantifications $\forall x \in C\, (\cdots)$ stand for $\forall x^X\, (C(x) =_0 0 \to \cdots)$ and dually for $\exists x \in C\, (\cdots)$. Since we have a bound on the diameter of the C, these quantifications are equivalent to bounded quantifications and treated as such.

The majorizability notion $\trianglelefteq_\rho$ is defined in a similar way to Section 3.2, but using instead the norm function of X – with the obvious choice of 0_X for the reference point. All the clauses stay the same with the exception of (M2), which now becomes

$$x \trianglelefteq_X n \to \|x\| \leq_{\mathbb{R}} (n)_{\mathbb{R}}.$$

As before, we denote by $\mathcal{T}_B^+$ the theory with the characteristic principles extended to the new types. It is not hard to see that the conditions (a) and (b) from Chapter 3 are satisfied by $\mathcal{T}_B$. The described $\mathcal{T}_B$ is thus a theory appropriate for the mining of Browder's proof and, in the context of the bounded functional interpretation, the extraction theorems in Section 3.3 still hold true with $\mathsf{PA}^\omega_\trianglelefteq[X, d_X, a_X, b_X]$ replaced by $\mathcal{T}_B$.

We will now recall the arguments in Browder's original proof. Browder starts by showing that the set of fixed points of the map U, $F = Fix(U)$, is nonempty, convex and closed. Formally, we can only speak of sets via their characteristic functions. The closeness condition is not required for the mining. Browder's argument that F is convex is easily formalizable in the theory. The existence of a fixed point in C can be added as an axiom to the theory, since its proof makes use of Zorn's Lemma and is not formalizable in $\mathcal{T}_B$. Formally, this would be a matter of extending the language with a constant c of type X and postulating the universal statement $U(c) = c$ together with $c \in C$. This works as these formulas are in the appropriate form. However, in this case, there is an easier approach. In $\mathcal{T}_B^+$, it is possible to prove the existence of a fixed point in C. This result follows from the fact that U has "almost fixed points" and by using $\mathsf{BC}^\omega_{\mathsf{bd}}$. The argument is the following. First we see that, in $\mathcal{T}_B$, one can prove

$$\forall n \in \mathbb{N} \left(\|U(u_n) - u_n\| \leq \frac{b}{n+1} \right). \tag{4.34}$$

This entails,

$$\forall n \in \mathbb{N}\, \exists x \in C \forall n' \leq n \left(\|U(x) - x\| < \frac{b}{n'+1} \right),$$

and, by (ii) of Lemma 3.6,

$$\forall n \in \mathbb{N}\, \exists x \in C \forall n' \leq n \left(\|U(x) - x\| \trianglelefteq_\mathbb{R} \frac{b}{n'+1} \right).$$

Finally, by an application of the (contrapositive of) $\mathsf{BC}^\omega_{\mathsf{bd}}$,

$$\exists x \in C \forall n \in \mathbb{N} \left(\|U(x) - x\| \trianglelefteq_\mathbb{R} \frac{1}{n+1} \right),$$

and, by (ii) of Lemma 3.6, we conclude

$$\exists x \in C\ (U(x) = x)\,.$$

Instead of this back-and-forth between the extensional and intensional inequalities, we could just observe that $\mathsf{BC}^\omega_{\mathsf{bd}}$ is trivially true for existential formulas and use its contrapositive directly.

Next, Browder's proof considers a projection argument to justify the existence of an unique point of F closest to v_0:

Let $\lambda := \inf_{x \in F} \|x - v_0\|$. By the definition of the infimum:

$$\forall k \in \mathbb{N} \exists x \in F \left(\|x - v_0\| \leq \lambda + \frac{1}{k+1} \right).$$

It is possible to write the previous statement without referencing the existence of this infimum.

$$\forall k \exists x \in C \left(U(x) = x \wedge \forall y \in C \left(U(y) = y \to \|x - v_0\| \leq \|y - v_0\| + \frac{1}{k+1} \right) \right). \tag{4.35}$$

Using countable choice, one then considers a sequence (x_n) of fixed points of U in C such that for all $n \in \mathbb{N}$,

$$\forall y \in C \left(U(y) = y \to \|x_n - v_0\| \leq \|y - v_0\| + \frac{1}{k+1} \right).$$

It is then shown that the sequence (x_n) is a Cauchy sequence and so it converges (to the point of F closest to v_0). However, as was commented by Kohlenbach in [27], (4.35) is already sufficient to carry out Browder's theorem and the mentioned application of countable choice is not needed. In $\mathcal{T}_B$, with a simple inductive argument, it is possible to prove (4.35).

Browder's proof continues with two technical facts:

(I) $\forall k \in \mathbb{N} \exists x \in C \ \big(U(x) = x \wedge \forall y \in C \big(U(y) = y \to \langle x - v_0, x - y \rangle < \frac{1}{k+1}\big)\big)$;

(II) $\forall n \in \mathbb{N} \forall x \in C \ (U(x) = x \to \|u_n - x\|^2 \leq \langle x - v_0, x - u_n \rangle)$.

Both (I) and (II) can be proven in $\mathcal{T}_B$. Fact (I) is derived from the projection argument and the convexity property of C. Fact (II) relies on simple computations and is the main combinatorial core of Browder's proof.
From (I) and (II) it is possible to derive that (u_n) is a Cauchy sequence.
In fact, let $k \in \mathbb{N}$ be given. From (I), consider $\tilde{x}$ such that $U(\tilde{x}) = \tilde{x}$ and

$$\forall y \in C \left(U(y) = y \to \langle \tilde{x} - v_0, \tilde{x} - y \rangle < \frac{1}{k+1} \right). \tag{4.36}$$

By (II) applied to $\tilde{x}$, it is enough to see that for some $n \in \mathbb{N}$ it holds

$$\forall i \geq n \left(\langle \tilde{x} - v_0, \tilde{x} - u_i \rangle < \frac{1}{k+1} \right). \tag{4.37}$$

Suppose not. Then,

$$\forall n \in \mathbb{N} \exists i \geq n \left(\langle \tilde{x} - v_0, \tilde{x} - u_i \rangle \geq \frac{1}{k+1} \right). \tag{4.38}$$

Take (v_n) a subsequence of (u_n) satisfying

$$\forall n \in \mathbb{N} \left(\langle \tilde{x} - v_0, \tilde{x} - v_n \rangle \geq \frac{1}{k+1} \right).$$

The sequential weak compactness is used at this point in the proof: Take (w_n) a subsequence of (v_n) weakly convergent to some $y \in C$. Since (w_n) is a sequence of almost fixed points (since (u_n) is) that is weakly convergent to $y \in C$, the demi-closeness principle states that such y is a fixed point. Furthermore, since (w_n) is a subsequence of (v_n), one concludes, $\langle \tilde{x} - v_0, \tilde{x} - y \rangle \geq \frac{1}{k+1}$. However, since $y \in F$, this is contradictory with (4.36).
Hence, (4.37) must hold for some $n \in \mathbb{N}$. By (II) applied to $\tilde{x}$ and one such n, we conclude $\|u_i - \tilde{x}\|^2 \leq \frac{1}{k+1}$. The Cauchy property now follows from this argument with $4(k+1)^2 - 1$ and triangle inequality.

The problem with the above argument resides in the fact that sequential weak compactness is not formalizable in $\mathcal{T}_B^+$. We now show how it can be bypassed by replacing it with Heine/Borel compactness, which corresponds to an application of $\mathsf{BC}^\omega_{\mathsf{bd}}$.

Given m, since (u_n) is a sequence of almost fixed points, by the assumption (4.38) we can take $y = u_i$ for a suitable large enough i in order to conclude

$$\forall m \in \mathbb{N} \exists y \in C \left(\|U(y) - y\| \leq \frac{1}{m+1} \wedge \langle \tilde{x} - v_0, \tilde{x} - y \rangle \geq \frac{1}{k+1} \right).$$

This is clearly equivalent to

$$\forall m \in \mathbb{N} \exists y \in C \forall i \leq m \left(\|U(y) - y\| \leq \frac{1}{i+1} \wedge \langle \tilde{x} - v_0, \tilde{x} - y \rangle \geq \frac{1}{k+1} \right). \tag{4.39}$$

Now, by the contrapositive of $\mathsf{BC}^\omega_{\mathsf{bd}}$, we conclude

$$\exists y \in C \left(U(y) = y \wedge \langle \tilde{x} - v_0, \tilde{x} - y \rangle \geq \frac{1}{k+1} \right). \tag{4.40}$$

This contradicts (4.36).
Notice that the matrix of (4.39) is not a bounded formula. Again, with the easy observation that $\mathsf{BC}^\omega_{\mathsf{bd}}$ extends to existential formulas (with bounded matrix) we are justified in this application of the (contrapositive of the) $\mathsf{BC}^\omega_{\mathsf{bd}}$. It is also possible to replace the inequalities

between real numbers with its intensional quantifier-free version $\trianglelefteq_{\mathbb{R}}$, and also make the application of $\mathsf{BC}^{\omega}_{\mathsf{bd}}$ valid. Since this argument, circumventing sequential weak compactness, is formalizable in $\mathcal{T}_B^+$, we can mine it and obtain a quantitative version of this modified proof. All the extracted quantitative information is now guaranteed to be given by primitive recursive functionals in Gödel's T.

Finally, we make some remarks. The application of $\mathsf{BC}^{\omega}_{\mathsf{bd}}$ above is an application of Heine/Borel compactness, where the relevant open sets Ω_i are

$$\Omega_i := \{x \in X \mid \|U(y) - y\| > \frac{1}{i+1}\} \cup \{x \in X \mid \langle \tilde{x} - v_0, \tilde{x} - y\rangle < \frac{1}{k+1}\}.$$

Assume (4.40) does not hold. Then $C \subseteq \bigcup_{i\in\mathbb{N}} \Omega_i$. By Heine/Borel compactness, there is some $m \in \mathbb{N}$ such that $C \subseteq \bigcup_{i\leq m} \Omega_i$. This contradicts (4.39).

We have shown that (u_n) is a Cauchy sequence and thus, convergent. Since (u_n) is a sequence of almost fixed points, it follows that (u_n) must converge to a fixed point. The important point is that the use of a false Heine/Borel compactness principle, in the context of $\mathcal{T}_B^+$, is an acceptable way of proving the convergence of sequences in Hilbert spaces.
The need of sequential weak compactness was replaced with an argument that makes use of the bounded collection axiom, which is one of characteristic principles of the bounded functional interpretation. However, the argument could also be made in the context of the monotone functional interpretation using the so-called *generalized uniform boundness principle* $\exists$-UB^X (see [23] or sections 17.7 and 17.8 of [25]). Still, while the bounded functional interpretation trivializes the (countable) Heine/Borel compactness – since it corresponds to a particular instance of $\mathsf{BC}^{\omega}_{\mathsf{bd}}$ –, the monotone functional interpretation can only interpret it by adding to the verifying theory an axiom which is true in the structure of the strongly majorizable functionals (extended to the new base type X). In other words, in the bounded functional interpretation we have a conservation result, whereas in the monotone functional interpretation we only have a true quantitative statement. For proof mining purposes (at the lower types usually considered) this makes no difference. In the context of the bounded functional interpretation it was easier to unearth this phenomenon via its bounded collection principles, but it still remains to see if there are applications with these principles which cannot be obtained using $\exists$-UB^X instead.

4.2.2 A general principle

In this section we isolate the Heine/Borel argument that allows the elimination of sequential weak compactness from mining. We will work in the theory of bounded metric spaces shown in Section 3.2. Then, we will prove its quantitative version. This general framework allows for the quantitative principle shown here to be applicable in many concrete situations.

Proposition 4.12. *The theory* $\mathsf{PA}^{\omega}_{\trianglelefteq}[X, d_X, a_X, b_X]^{+}$ *proves the following. Let* U *be a map from* X *to* X *and* $(u_n)_{n\in\mathbb{N}}$ *a sequence of elements of* X *such that* $\lim_n d(U(u_n), u_n) = 0$. *Consider* $F := Fix(U)$. *Given* $k \in \mathbb{N}$, $\lambda \in \mathbb{R}$ *and* $\theta : X \to \mathbb{R}$, *if*

$$\forall y \in F \left(\lambda \trianglelefteq_{\mathbb{R}} \theta(y) + \frac{1}{k+1}\right) \tag{4.41}$$

then, for n *sufficiently large,*

$$\lambda \trianglelefteq_{\mathbb{R}} \theta(u_n) + \frac{1}{k+1} \tag{4.42}$$

Proof. By hypothesis and the definition of fixed point

$$\forall y \in X \left(\forall m \in \mathbb{N} \left(d(U(y), y) \trianglelefteq_{\mathbb{R}} \frac{1}{m+1}\right) \to \lambda \trianglelefteq_{\mathbb{R}} \theta(y) + \frac{1}{k+1}\right).$$

Hence,

$$\forall y \in X \exists m \in \mathbb{N} \left(d(U(y), y) \trianglelefteq_{\mathbb{R}} \frac{1}{m+1} \to \lambda \trianglelefteq_{\mathbb{R}} \theta(y) + \frac{1}{k+1}\right).$$

By $\mathsf{BC}^{\omega}_{\mathsf{bd}}$,

$$\exists l \in \mathbb{N} \forall y \in X \exists m \leq l \left(d(U(y), y) \trianglelefteq_{\mathbb{R}} \frac{1}{m+1} \to \lambda \trianglelefteq_{\mathbb{R}} \theta(y) + \frac{1}{k+1}\right).$$

Take one such $l = l_0$. Clearly

$$\forall y \in X \left(d(U(y), y) \trianglelefteq_{\mathbb{R}} \frac{1}{l_0+1} \to \lambda \trianglelefteq_{\mathbb{R}} \theta(y) + \frac{1}{k+1}\right).$$

Since $\lim_n d(U(u_n), u_n) = 0$, the result follows. ☐

A more concrete result is needed for the arguments in which we intend to apply this principle.

Proposition 4.13 (General principle)**.** *The theory* $\mathsf{PA}^{\omega}_{\trianglelefteq}[X, d_X, a_X, b_X]^{+}$ *proves the following. Let* U *be a map from* X *to* X, φ *a map from* $X \times X$ *to* $\mathbb{R}$, *and* $(u_n)_{n\in\mathbb{N}}$ *be a sequence of elements of* X *such that* $\lim_n d(U(u_n), u_n) = 0$. *Consider* $F := Fix(U)$. *If*

$$\forall k \in \mathbb{N}\, \exists x \in F\, \forall y \in F \left(\varphi(x, x) \trianglelefteq_{\mathbb{R}} \varphi(x, y) + \frac{1}{k+1}\right)$$

then

$$\forall k \in \mathbb{N}\, \exists x \in F\, \exists n \in \mathbb{N}\, \forall m \geq n \left(\varphi(x, x) \trianglelefteq_{\mathbb{R}} \varphi(x, u_m) + \frac{1}{k+1}\right).$$

Proof. Given $k \in \mathbb{N}$ take, by hypothesis, $\tilde{x} \in F$ such that $\forall y \in F\, (\varphi(\tilde{x}, \tilde{x}) \trianglelefteq_{\mathbb{R}} \varphi(\tilde{x}, y) + \frac{1}{k+1})$. Let $\lambda := \varphi(\tilde{x}, \tilde{x})$ and $\theta(y) := \varphi(\tilde{x}, y)$. Now apply Proposition 4.12. ☐

When $\trianglelefteq_{\mathbb{R}}$ is replaced by $\leq_{\mathbb{R}}$, these results are set-theoretically false and we can give a counterexample: Take X as the unit ball of the normed space ℓ^1 (the space of real-value sequences whose series is absolutely convergent). Let U be the shift operator $U(x_0, x_1, \cdots) := (0, x_0, x_1, \cdots)$ and $\theta(x_0, x_1, x_2, \cdots) := -\sum_{i=0}^{\infty} |x_i|$, i.e., $\theta(x)$ is the symmetric of the norm of x. The only fixed point of U is the zero vector. Let u_n be the vector $(\frac{1}{n+1}, \cdots, \frac{1}{n+1}, 0, 0, \cdots)$, where there are $n+1$ nonzero entries. Clearly, we have always $\theta(u_n) = -1$ and $\|U(u_n) - u_n\| = \frac{2}{n+1}$. Then, Proposition 4.12 fails with $\lambda = 0$ (for any $k \in \mathbb{N}$). By considering $\varphi(x, y) = -\|y\|$, it is possible to see with this example that the general principle is also false.

Now we show a quantitative version of Proposition 4.13. This quantitative version can then be applied to concrete cases of minings to achieve the correct quantitative version of the argument replacing sequential weak compactness. This is exemplified with the analyses of theorems due to Browder and due to Wittmann in the following Section 4.4.
The condition (a) is the metastable version of the assumption that $\lim_n d(U(u_n), u_n) = 0$. Condition (b) is the interpretation of the main hypothesis of Proposition 4.13 and, in the applications, corresponds to the interpretation of the corresponding fact (I) of Browder's proof. The conclusion is again a metastable version, this time of the conclusion of Proposition 4.13.

Theorem 4.14 (Quantitative version of the general principle)**.** *Let (X, d) be a metric space. Let U be a map from X to X, φ a map from $X \times X$ to $\mathbb{R}$ and $(u_n)_{n\in\mathbb{N}}$ be a sequence of elements of X. Suppose that there are monotone functionals α and β from $\mathbb{N} \times \mathbb{N}^{\mathbb{N}}$ to $\mathbb{N}$ satisfying:*

(a) $\forall k \in \mathbb{N}\, \tilde{\forall} f \in \mathbb{N}^{\mathbb{N}}\, \exists N \leq \alpha(k, f)\, \forall n \in [N, f(N)] \left(d(U(u_n), u_n) \leq \dfrac{1}{k+1} \right)$;

(b) $\forall k \in \mathbb{N}\, \tilde{\forall} f \in \mathbb{N}^{\mathbb{N}}\, \exists N \leq \beta(k, f) \exists x \in X$
$$\left(d(U(x), x) \leq \frac{1}{f(N)+1} \wedge \forall y \in X \left(d(U(y), y) \leq \frac{1}{N+1} \to \varphi(x, x) \leq \varphi(x, y) + \frac{1}{k+1} \right) \right).$$

Then, for every $k \in \mathbb{N}$ and any monotone function $f \in \mathbb{N}^{\mathbb{N}}$, there are a natural number N with $N \leq \psi(k, f)$ and $x \in X$ such that

$$d(U(x), x) \leq \frac{1}{f(N)+1} \wedge \forall n \in [N, f(N)] \left(\varphi(x, x) \leq \varphi(x, u_n) + \frac{1}{k+1} \right), \tag{4.43}$$

where $\psi(k, f) := \alpha\left(\beta\left(k, \widehat{f}\right), f\right)$, with $\widehat{f}(m) := f(\alpha(m, f))$.

Proof. Take $k \in \mathbb{N}$ and a monotone function $f \in \mathbb{N}^{\mathbb{N}}$. By (b), applied to k and $\widehat{f}$ there are $N_1 \leq \beta\left(k, \widehat{f}\right)$ and $\tilde{x} \in X$ such that

$$d(U(\tilde{x}),\tilde{x}) \leq \frac{1}{\widehat{f}(N_1)+1} \text{ and}$$

$$\forall y \in X \left(d(U(y),y) \leq \frac{1}{N_1+1} \to \varphi(\tilde{x},\tilde{x}) \leq \varphi(\tilde{x},y) + \frac{1}{k+1} \right). \tag{4.44}$$

Apply (a) to N_1 and f to get $N \leq \alpha(N_1, f)$ satisfying

$$\forall n \in [N, f(N)] \left(d(U(u_n),u_n) \leq \frac{1}{N_1+1} \right). \tag{4.45}$$

We have $N \leq \alpha(N_1, f) \leq \alpha\left(\beta\left(k, \widehat{f}\right), f\right) = \psi(k, f)$ and, by the monotonicity of f,

$$d(U(\tilde{x}),\tilde{x}) \leq \frac{1}{\widehat{f}(N_1)+1} = \frac{1}{f\left(\alpha(N_1,f)\right)+1} \leq \frac{1}{f(N)+1}.$$

Also, for $n \in [N, f(N)]$, by (4.44) and (4.45), we have

$$\varphi(\tilde{x},\tilde{x}) \leq \varphi(\tilde{x},u_n) + \frac{1}{k+1}.$$

$\square$

Provided that the space is bounded, the form of Theorem 4.14 is the quantitative version of Proposition 4.13 obtained using the bounded functional interpretation. However, in the end, we did not require the boundedness of the metric space (X, d). As can be seen from the proof, this hypothesis is not necessary. Proposition 4.14 is just a simple mathematical fact. In our applications, however, the given monotone functionals α and β depend on the bound of the metric space (as well as the concluding bounding functional ψ). A similar situation also happens in the forthcoming Proposition 5.9.

4.3 The projection argument

As we discussed in Subsection 4.2.1, the projection argument requires the use of countable choice and thus its interpretation goes beyond the functionals of Gödel's T. In contrast the weaker projection statement (4.35) can be proved using induction only and is already sufficient to carry on Browder's argument. In this section, we interpret and mine the proof of this weaker projection statement using the bounded functional interpretation. In section 3 of [12], a "dance" between the inequalities $<_{\mathbb{R}}$ and $\leq_{\mathbb{R}}$ is used to fully justify the interpretation of (4.35). There the full aparathus of the bounded functional interpretation was not introduced, which required that *ad hoc* treatment. Here, however, it is possible to switch to the intensional quantifier-free inequality $\trianglelefteq_{\mathbb{R}}$ and get a correct explanation of the interpretation. We show the analysis of the (weaker) projection statement. By the context, there is no problem in writing $\trianglelefteq$ in place of $\trianglelefteq_{\mathbb{R}}$. The statement (4.35) is equivalent to

$$\forall k \exists x \in C$$
$$\left[U(x) = x \wedge \forall y \in C \left(\forall N \left(\|U(y) - y\| \trianglelefteq \frac{1}{N+1}\right) \to \|x - v_0\|^2 \trianglelefteq \|y - v_0\|^2 + \frac{1}{k+1}\right)\right]$$

and, hence, equivalent to

$$\forall k \exists x \in C$$
$$\left(U(x) = x \wedge \forall y \in C\, \exists N \left(\|U(y) - y\| \trianglelefteq \frac{1}{N+1} \to \|x - v_0\|^2 \trianglelefteq \|y - v_0\|^2 + \frac{1}{k+1}\right)\right).$$

The innocuous change to the squares is just to make it easier to relate with the inner product functional. Since the formula after '$\exists N$' is quantifier-free, by $\mathsf{BC}^{\omega}_{\mathsf{bd}}$ we easily get

$$\forall k \exists x \in C$$
$$\left(U(x) = x \wedge \exists N \forall y \in C \left(\|U(y) - y\| \trianglelefteq \frac{1}{N+1} \to \|x - v_0\|^2 \trianglelefteq \|y - v_0\|^2 + \frac{1}{k+1}\right)\right)$$

or, equivalently,

$$\forall k \exists N \exists x \in C \forall m \left(\|U(x) - x\| \trianglelefteq \frac{1}{m+1} \right.$$
$$\left. \wedge\ \forall y \in C \left(\|U(y) - y\| \trianglelefteq \frac{1}{N+1} \to \|x - v_0\|^2 \trianglelefteq \|y - v_0\|^2 + \frac{1}{k+1}\right)\right).$$

In turn, by the contrapositive of $\mathsf{BC}^{\omega}_{\mathsf{bd}}$, this is equivalent to

$$\forall k \exists N \forall m \exists x \in C \left(\|U(x) - x\| \trianglelefteq \frac{1}{m+1} \right.$$
$$\left. \wedge \forall y \in C \left(\|U(y) - y\| \trianglelefteq \frac{1}{N+1} \to \|x - v_0\|^2 \trianglelefteq \|y - v_0\|^2 + \frac{1}{k+1}\right)\right).$$

Finally, using $\mathsf{mAC}^{\omega}_{\mathsf{bd}}$, we obtain

$$\forall k \widetilde{\forall} f \exists N \exists x \in C \left(\|U(x) - x\| \trianglelefteq \frac{1}{f(N)+1} \right.$$
$$\left. \wedge \forall y \in C \left(\|U(y) - y\| \trianglelefteq \frac{1}{N+1} \to \|x - v_0\|^2 \trianglelefteq \|y - v_0\|^2 + \frac{1}{k+1}\right)\right).$$

The formula above is the translation of (4.35) using the bounded functional interpretation and is provable in $\mathcal{T}_B^+$. By the results of Section 3.3 (for $\mathcal{T}_B$), we are guaranteed to be able to find a monotone functional bounding N in terms of k and f. Then, by flattening, we can go back to the extensional inequalities. The reader should compare the end formula above with the corresponding formula given by the monotone functional interpretation in p. 2772 of [27]. Since (4.35) is proved by induction, it is not a surprise that the bound on N is defined by recursion:

Proposition 4.15. *For any natural number k and monotone function $f : \mathbb{N} \to \mathbb{N}$, there are $N \leq f^{(r)}(0)$ and $x \in C$ such that*

$$\|U(x)-x\| \leq \frac{1}{f(N)+1} \wedge \forall y \in C\left(\|U(y)-y\| \leq \frac{1}{N+1} \to \|x-v_0\|^2 \leq \|y-v_0\|^2 + \frac{1}{k+1}\right),$$

where $r := b^2(k+1)$ and $f^{(r)}$ is the r-th fold composition of f.

Proof. Assume that the result is not true. Then there are $k \in \mathbb{N}$ and a monotone function $f : \mathbb{N} \to \mathbb{N}$, such that for all $N \leq f^{(r)}(0)$ and $x \in C$

$$\|U(x)-x\| \leq \frac{1}{f(N)+1} \to \exists y^C \left(\|U(y)-y\| \leq \frac{1}{N+1} \wedge \|y-v_0\|^2 < \|x-v_0\|^2 - \frac{1}{k+1}\right). \tag{4.46}$$

First of all, note that the r-sequence given by the expression $f^{(r)}(0)$ is monotone (because f is). We define a finite sequence $x_0, x_1, \ldots, x_r, x_{r+1}$ of elements of C as follows:
$\underline{x_0}$:
By (4.34), let x_0 be such that

$$\|U(x_0)-x_0\| \leq \frac{1}{f^{(r+1)}(0)+1}.$$

$\underline{x_{j+1}, \text{ for } j \leq r}$:
Assume that we have x_j such that $\|U(x_j)-x_j\| \leq \frac{1}{f^{(r-j+1)}(0)+1}$. By (4.46) applied to $N = f^{(r-j)}(0)$ and to $x = x_j$, we conclude that there is $y \in C$ satisfying

$$\|U(y)-y\| \leq \frac{1}{f^{(r-j)}(0)+1} \wedge \|y-v_0\|^2 < \|x_j-v_0\|^2 - \frac{1}{k+1}.$$

Let x_{j+1} be one such y.

By the definition, for all $j \leq r$,

$$\|x_{j+1}-v_0\|^2 < \|x_j-v_0\|^2 - \frac{1}{k+1},$$

which implies the contradiction

$$\|x_{r+1} - v_0\|^2 < \|x_0 - v_0\|^2 - \frac{r+1}{k+1} \leq b^2 - \frac{b^2(k+1)+1}{k+1} < 0.$$

□

The fact (I) of Subsection 4.2.1 is a refinement of the projection argument. Its bounded functional interpretation is similar to the interpretation of (4.35):

$$\forall k \widetilde{\forall} f \exists N \, \exists x \in C$$
$$\left(\|U(x) - x\| \trianglelefteq \frac{1}{f(N)+1} \wedge \forall y \in C \left(\|U(y) - y\| \trianglelefteq \frac{1}{N+1} \to \langle x - v_0, x - y \rangle \trianglelefteq \frac{1}{k+1} \right) \right),$$

and the (flattening of) the corresponding mined result gives:

Proposition 4.16. *For any natural number k and monotone function $f : \mathbb{N} \to \mathbb{N}$, there are $N \leq 12b(\check{f}^{(R)}(0)+1)^2$ and $x \in C$ such that*

$$\|U(x) - x\| \leq \frac{1}{f(N)+1} \wedge \forall y \in C \left(\|U(y) - y\| \leq \frac{1}{N+1} \to \langle x - v_0, x - y \rangle \leq \frac{1}{k+1} \right),$$

with $R := b^4(k+1)^2$ and $\check{f}(m) := \max\{f(12b(m+1)^2),\, 12b(m+1)^2\}$.

As we will explain, this mining can be obtained from Proposition 4.15 and the following two estimates, essentially due to Kohlenbach in [27].
In the following, we write $w_\gamma(u, v) := (1-\gamma)u + \gamma v$, for $\gamma \in [0, 1]$.

Lemma 4.17. *For all $k \in \mathbb{N}$ and $x_1, x_2 \in C$,*

$$\bigwedge_{j=1}^{2} \left(\|U(x_j) - x_j\| \leq \frac{1}{12b(k+1)^2} \right) \to \forall \gamma \in [0,1] \left(\|U(w_\gamma(x_1, x_2)) - w_\gamma(x_1, x_2)\| \leq \frac{1}{k+1} \right).$$

The lemma above is a quantitative version of the statement that the set of fixed points in C is a convex set.

Lemma 4.18. *For all $k \in \mathbb{N}$ and $x, y \in C$,*

$$\forall \gamma \in [0,1] \left(\|x - v_0\|^2 \leq \|w_\gamma(x, y) - v_0\|^2 + \frac{1}{b^2(k+1)^2} \right) \to \langle x - v_0, x - y \rangle \leq \frac{1}{k+1}.$$

Notice that the extracted information on these two lemmas does not depend on the points in C but only on the bound on the diameter of C. Via the bounded functional interpretation this is made clear by the use of the bounded collection principle and the fact that quantifications over C are treated as bounded.
Using Lemma 4.17, we have the following intermediate result:

Corollary 4.19. *For any natural number k and monotone function $f : \mathbb{N} \to \mathbb{N}$, there are $N \le 12b(\check{f}^{(r)}(0)+1)^2$ and $x \in C$ for which the following two properties hold:*

$$\|U(x) - x\| \le \frac{1}{f(N)+1}$$

and

$$\forall y \in C \left(\|U(y) - y\| \le \frac{1}{N+1} \to \forall \gamma \in [0,1] \left(\|x - v_0\|^2 \le \|w_\gamma(x,y) - v_0\|^2 + \frac{1}{k+1} \right) \right),$$

with $r := b^2(k+1)$ and $\check{f}(m) := \max\{f(12b(m+1)^2),\, 12b(m+1)^2\}$.

Proof. Let k and monotone f be given. By Proposition 4.15, there exist $x \in C$ and $N' \le \check{f}^{(r)}(0)$ with

$$\|U(x) - x\| \le \frac{1}{\check{f}(N')+1}$$

and

$$\forall y \in C \left(\|U(y) - y\| \le \frac{1}{N'+1} \to \|x - v_0\|^2 \le \|y - v_0\|^2 + \frac{1}{k+1} \right), \tag{4.47}$$

where $r = b^2(k+1)$. Define $N := 12b(N'+1)^2$. Clearly, $N \le 12b(\check{f}^{(r)}(0)+1)^2$. This entails that

$$\|U(x) - x\| \le \frac{1}{f(N)+1}$$

because $f(N) = f(12b(N'+1)^2) \le \check{f}(N')$. Now, take $y \in C$ such that $\|U(y) - y\| \le \frac{1}{N+1}$. Hence $\|U(y) - y\| \le \frac{1}{12b(N'+1)^2}$. On the other hand, we also have

$$\|U(x) - x\| \le \frac{1}{\check{f}(N')+1} \le \frac{1}{12b(N'+1)^2}.$$

By Lemma 4.17, we get $\|U(w_\gamma(x,y)) - w_\gamma(x,y)\| \le \frac{1}{N'+1}$.
The result then follows from (4.47) for $w_\gamma(x,y)$. □

Lemma 4.18 corresponds to the mining of the following result:

$$\forall \gamma \in [0,1] \big(\|x - v_0\|^2 \le \|w_\gamma(x,y) - v_0\|^2 \big) \to \langle x - v_0, x - y \rangle \le 0.$$

This result is implicit in Browder's proof [6] and is needed to show (I) of Subsection 4.2.1. Proposition 4.16 is an immediate consequence of Lemma 4.18 and Corollary 4.19 by instantiating k with $b^2(k+1)^2 - 1$.

Chapter 5

Removing weak compactness

In this chapter we shall obtain, using the quantitative general principle, quantitative versions of three theorems. We discuss theorems by F.E. Browder [6] and by R. Wittmann [51] and prove quantitative metastable versions. Later, by expanding this method, we will obtain a metastable version of an extension of Wittmann's theorem due to Bauschke [2]. We will be following section 5 of [12].
By an analysis of the proofs of these results, one can see that they finish with a simple argument that relies on an application of *modus ponens* (and the triangle inequality). The next result isolates this argument in the general framework of bounded metric spaces.

Lemma 5.1. *The theory* $\mathsf{PA}^{\omega}_{\trianglelefteq}[X, d_X, a_X, b_X]$ *proves the following. Let* U *be a map from* X *to* X*,* φ *a map from* $X \times X$ *to* $\mathbb{R}$ *and* (u_n) *be a sequence of elements of* X*. Consider* $F := Fix(U)$*. Suppose that*

$$\forall k \in \mathbb{N}\, \exists x \in F\, \exists n \in \mathbb{N}\, \forall m \geq n \left(\varphi(x, x) \trianglelefteq_{\mathbb{R}} \varphi(x, u_m) + \frac{1}{k+1} \right)$$

and that there is a monotone function $\delta : \mathbb{N} \to \mathbb{N}$ *such that, for all* $k \in \mathbb{N}$ *and* $x \in F$*,*

$$\exists n \in \mathbb{N}\, \forall m \geq n \left(\varphi(x, x) \trianglelefteq_{\mathbb{R}} \varphi(x, u_m) + \frac{1}{\delta(k)+1} \right) \to \exists M \in \mathbb{N}\, \forall m \geq M \left(d(u_m, x) \trianglelefteq_{\mathbb{R}} \frac{1}{k+1} \right).$$

Then, (u_n) *is a Cauchy sequence.*

Proof. Let $k \in \mathbb{N}$ be given. By the first assumption, applied to $\delta(2k+1)$, we have for some $\tilde{x} \in F$,

$$\exists n \in \mathbb{N}\, \forall m \geq n \left(\varphi(\tilde{x}, \tilde{x}) \trianglelefteq_{\mathbb{R}} \varphi(\tilde{x}, u_m) + \frac{1}{\delta(2k+1)+1} \right).$$

By the second assumption, we conclude that

$$\exists M \in \mathbb{N}\, \forall m \geq M \left(d(u_m, \tilde{x}) \trianglelefteq_{\mathbb{R}} \frac{1}{2(k+1)} \right).$$

Using triangle inequality, it follows that (u_n) is a Cauchy sequence. $\square$

It is not difficult to find the form that the quantitative version of the above lemma must take under the bounded functional interpretation. The interpretation of the first assumption is the statement in the conclusion of Theorem 4.14. The second assumption can be written as

$$\forall k \forall x \in X \left[\left(\forall r \left(d(U(x), x) \trianglelefteq \frac{1}{r+1} \right) \wedge \exists n \forall i \geq n \left(\varphi(x,x) \trianglelefteq \varphi(x, u_i) + \frac{1}{\delta(k)+1} \right) \right) \right.$$
$$\left. \to \exists M \forall m \geq M \left(d(u_m, x) \trianglelefteq \frac{1}{k+1} \right) \right].$$

By the contrapositive of $\mathsf{mAC}^{\omega}_{\mathsf{bd}}$, the conclusion of the implication is equivalent to its metastable version

$$\widetilde{\forall} f \exists M \forall m \in [M, f(M)] \left(d(u_m, x) \trianglelefteq \frac{1}{k+1} \right).$$

Therefore, the second assumption is equivalent to

$$\forall k, n \widetilde{\forall} f \forall x \in X \exists r \exists i \geq n \exists M \left[\left(d(U(x), x) \trianglelefteq \frac{1}{r+1} \wedge \varphi(x,x) \trianglelefteq \varphi(x, u_i) + \frac{1}{\delta(k)+1} \right) \right.$$
$$\left. \to \forall m \in [M, f(M)] \left(d(u_m, x) \trianglelefteq \frac{1}{k+1} \right) \right].$$

By $\mathsf{BC}^{\omega}_{\mathsf{bd}}$, we have

$$\forall k, n \widetilde{\forall} f \exists r, i, M \forall x \in X \exists r' \leq r \exists j \in [n, i] \exists M' \leq M \left[\left(d(U(x), x) \trianglelefteq \frac{1}{r'+1} \right. \right.$$
$$\left. \left. \wedge\, \varphi(x,x) \trianglelefteq \varphi(x, u_j) + \frac{1}{\delta(k)+1} \right) \to \forall m \in [M', f(M')] \left(d(u_m, x) \trianglelefteq \frac{1}{k+1} \right) \right]$$

Therefore

$$\forall k, n \widetilde{\forall} f \exists r, i, M \forall x \in X \left[\left(d(U(x), x) \trianglelefteq \frac{1}{r+1} \wedge \forall j \in [n, i] \left(\varphi(x,x) \trianglelefteq \varphi(x, u_j) + \frac{1}{\delta(k)+1} \right) \right) \right.$$
$$\left. \to \exists M' \leq M \forall m \in [M', f(M')] \left(d(u_m, x) \trianglelefteq \frac{1}{k+1} \right) \right].$$

We can then look at the *modus ponens* argument justifying Lemma 5.1 and extract a bound ϕ on the metastability of (u_n). Since the proof is essentially an application of *modus ponens*, by the proof of the soundness theorem, it is expected that the computed bound be defined with a function composition. By flattening, we can return to the usual extensional inequalities between real numbers. The next result is the quantitative version of Lemma 5.1.

Proposition 5.2 (Quantitative version of 5.1)**.** *Let (X, d) be a metric space. Let U be a map from X to X, φ a map from $X \times X$ to $\mathbb{R}$ and (u_n) be a sequence of elements of X. Suppose that there are monotone functions $\delta, \psi, \gamma, \eta$ and σ satisfying:*

(i) $\forall k \in \mathbb{N}\, \widetilde{\forall} f \in \mathbb{N}^{\mathbb{N}}\, \exists N \leq \psi(k, f)$
$$\exists x \in X \left(d(U(x), x) \leq \frac{1}{f(N)+1} \wedge \forall n \in [N, f(N)] \left(\varphi(x,x) \leq \varphi(x, u_n) + \frac{1}{k+1} \right) \right) \quad \textit{and}$$

(ii) $\forall k, n \in \mathbb{N}\, \widetilde{\forall} f \in \mathbb{N}^{\mathbb{N}}\, \forall x \in X$
$$\Big[d(U(x), x) \leq \frac{1}{\gamma(k,n,f)+1} \wedge \forall i \in [n, \eta(k,n,f)] \left(\varphi(x,x) \leq \varphi(x, u_i) + \frac{1}{\delta(k)+1} \right)$$
$$\rightarrow \exists M \leq \sigma(k,n,f)\, \forall m \in [M, f(M)] \left(d(u_m, x) \leq \frac{1}{k+1} \right) \Big].$$

Then
$$\forall k \in \mathbb{N}\, \widetilde{\forall} f \in \mathbb{N}^{\mathbb{N}}\, \exists M \leq \phi(k,f)\, \forall m, n \in [M, f(M)] \left(d(u_m, u_n) \leq \frac{1}{k+1} \right), \tag{5.1}$$
where $\phi(k,f) := \sigma\Big(2k+1, \psi\big(\delta(2k+1), \overline{f}\big), f\Big)$, where the auxiliary function $\overline{f}$ is defined by $\overline{f}(m) := \max\{\gamma(2k+1, m, f),\, \eta(2k+1, m, f)\}$.

Proof. Let $k \in \mathbb{N}$ and monotone $f \in \mathbb{N}^{\mathbb{N}}$ be given. Notice that, since γ and η are monotone, the function $\overline{f}$ is also monotone. We apply condition (i) to $\delta(2k+1)$ and $\overline{f}$ in order to get $N_1 \leq \psi(\delta(2k+1), \overline{f})$ and $\tilde{x} \in X$ such that
$$d(U(\tilde{x}), \tilde{x}) \leq \frac{1}{\overline{f}(N_1)+1} \quad \text{and}$$
$$\forall n \in [N_1, \overline{f}(N_1)] \left(\varphi(\tilde{x}, \tilde{x}) \leq \varphi(\tilde{x}, u_n) + \frac{1}{\delta(2k+1)+1} \right). \tag{5.2}$$

Now apply (ii) to $2k+1, N_1, f$ and $\tilde{x} \in X$ and obtain
$$\begin{aligned} &\Bigg(d(U(\tilde{x}), \tilde{x}) \leq \frac{1}{\gamma(2k+1, N_1, f)+1} \\ &\wedge \forall i \in [N_1, \eta(2k+1, N_1, f)] \left(\varphi(\tilde{x}, \tilde{x}) \leq \varphi(\tilde{x}, u_i) + \frac{1}{\delta(2k+1)+1} \right) \Bigg) \\ &\qquad \rightarrow \exists M \leq \sigma(2k+1, N_1, f)\, \forall m \in [M, f(M)] \left(d(u_m, \tilde{x}) \leq \frac{1}{2k+2} \right). \end{aligned} \tag{5.3}$$

Since $\gamma(2k+1, N_1, f)$, $\eta(2k+1, N_1, f) \leq \overline{f}(N_1)$, by (5.2) we have the antecedent of (5.3). Therefore
$$\exists M \leq \sigma(2k+1, N_1, f)\, \forall m \in [M, f(M)] \left(d(u_m, \tilde{x}) \leq \frac{1}{2k+2} \right)$$

Finally, we have $M \leq \sigma(2k+1, N_1, f) \leq \sigma\big(2k+1, \psi\big(\delta(2k+1), \overline{f}\big), f\big) = \phi(k,f)$ and, by the triangle inequality, the result follows. $\square$

5.1 Browder's theorem

In the following, we are in the hypotheses of Theorem 4.11. Thus, X is a real Hilbert space, C is a bounded closed convex subset of X, $U : X \to X$ is a nonexpansive mapping that maps C into itself, $v_0 \in C$ and the sequence (u_n) is defined as in Theorem 4.11.

We use the quantitative general principle (Proposition 4.14) and Proposition 5.2 for the bounded metric space C with the metric induced by the Hilbert space norm and for the mapping $\varphi(x, y) := \langle x - v_0, y\rangle$. Let $b \in \mathbb{N}$ be a positive upper bound on the diameter of C. Let us define first the following functions:

$$r : \mathbb{N} \to \mathbb{N}, \quad r(k) = b^4(k+1)^2. \tag{5.4}$$

and, for every $g : \mathbb{N} \to \mathbb{N}$,

$$\omega_g : \mathbb{N} \to \mathbb{N}, \quad \omega_g(m) = \max\{g(12b(m+1)^2),\, 12b(m+1)^2\}. \tag{5.5}$$

As an immediate consequence of (4.34) of Subsection 4.2.1, we get that condition (a) of Proposition 4.14 is fulfilled with

$$\alpha(k, f) := \alpha(k) := b(k+1).$$

From the analysis of the projection argument, namely Proposition 4.16, condition (b) of Proposition 4.14 is satisfied with

$$\beta(k, f) := 12b\left(\omega_f^{(r(k))}(0) + 1\right)^2.$$

Therefore we can apply the quantitative version of the general principle (Proposition 4.14) to get that for every $k \in \mathbb{N}$ and any monotone function $f \in \mathbb{N}^{\mathbb{N}}$, there exists $N \leq \psi(k, f)$ and $x \in C$ such that

$$d(U(x), x) \leq \frac{1}{f(N)+1} \wedge \forall n \in [N, f(N)] \left(\varphi(x, x) \leq \varphi(x, u_n) + \frac{1}{k+1}\right), \tag{5.6}$$

where

$$\begin{aligned} \psi(k, f) &= \alpha\left(\beta(k, \widehat{f}), f\right) = b\left(\beta\left(k, \widehat{f}\right) + 1\right) = b\left(12b\left(\omega_{\widehat{f}}^{(r(k))}(0) + 1\right)^2 + 1\right) \\ &= 12b^2\left(\omega_{\widehat{f}}^{(r(k))}(0) + 1\right)^2 + b. \end{aligned}$$

with

$$\widehat{f}(m) = f(\alpha(m, f)) = f(b(m+1)).$$

Thus, condition (i) of Proposition 5.2 is satisfied with this functional ψ.
We now need to show that hypothesis (ii) of Proposition 5.2 holds and determine the required bounds. This follows from the mining of fact (II) of Subsection 4.2.1. It can be read from Kohlenbach's computations of lemma 2.11 of [27] that, for all $x \in C$ and $k, n \in \mathbb{N}$,

$$\left(\|U(x) - x\| \leq \frac{1}{2b(n+1)(k+1)^2} \wedge \langle x - v_0, x - u_n \rangle \leq \frac{1}{2(k+1)^2}\right) \to \|u_n - x\| \leq \frac{1}{k+1}. \tag{5.7}$$

Therefore, condition (ii) of Proposition 5.2 holds with

$$\gamma(k,n,f) := 2b(f(n)+1)(k+1)^2 - 1,\ \delta(k) := 2(k+1)^2 - 1,\ \eta(k,n,f) := f(n) \text{ and } M := \sigma(k,n,f) := n.$$

Finally, the conclusion of Proposition 5.2 yields a quantitative version of Browder's theorem and a bound ϕ_{Br} on the metastability of the sequence (u_n) is extracted.

Theorem 5.3 (Quantitative Browder)**.** *Under the conditions of Browder's theorem, let $b \in \mathbb{N}$ be a positive upper bound on the diameter of C. Then, for all $k \in \mathbb{N}$ and every monotone function $f : \mathbb{N} \to \mathbb{N}$,*

$$\exists N \leq \phi_{\mathrm{Br}}(k,f)\, \forall i,j \in [N, f(N)] \left(\|u_i - u_j\| < \frac{1}{k+1}\right),$$

where

$$\phi_{\mathrm{Br}}(k,f) := 12b^2 \left(h^{(R)}(0) + 1\right)^2 + b,$$

with $R := 64b^4(k+1)^4$
and $h(m) := \max\{8b(f(12b^2(m+1)^2 + b) + 1)(k+1)^2 - 1, 12b(m+1)^2\}$.

Proof. Apply Proposition 5.2 and see that

$$\begin{aligned}
\overline{f}(m) &= 8b(f(m)+1)(k+1)^2 - 1, \\
\phi_{\mathrm{Br}}(k,f) &= \sigma\left(2k+1, \psi\left(\delta(2k+1), \overline{f}\right), f\right) = \psi\left(\delta(2k+1), \overline{f}\right) = \\
&= 12b^2 \left(\omega_{\widehat{\overline{f}}}^{(r(\delta(2k+1)))}(0) + 1\right)^2 + b = 12b^2 \left(h^{(R)}(0) + 1\right)^2 + b.
\end{aligned}$$

□

5.2 Wittmann's theorem

In an attempt to find fixed point for nonexpansive maps, Benjamin Halpern went in a different direction than Browder. He introduced the so-called *Halpern iterations* which are defined recursively by

$$u_0 := u \in X, \quad u_{n+1} := \lambda_n u_0 + (1 - \lambda_{n+1})U(u_n), \tag{5.8}$$

where $(\lambda_n) \subset [0, 1]$ is a sequence of real numbers and u is some point in the Hilbert space.

Initially with $u_0 = 0$, in [18], Halpern showed the strong convergence of (u_n) under certain conditions for (λ_n). However, the conditions considered by Halpern, left out the natural choice $\lambda_n = \frac{1}{n+1}$. In 1992, Wittmann proved in [51] the following important result,

Theorem 5.4 (Wittmann [51])**.** *Let X be a Hilbert space, C be a nonempty closed convex bounded subset of X and $U : C \to C$ be a nonexpansive mapping. Assume that (λ_n) is a sequence in $(0, 1)$ satisfying*

$$(C1) \quad \lim \lambda_n = 0, \quad (C2) \quad \sum_{n=1}^{\infty} \lambda_n = \infty, \quad (C3) \quad \sum_{n=1}^{\infty} |\lambda_n - \lambda_{n+1}| < \infty.$$

If (u_n) is a sequence defined by (5.8) with $u_0 = u$ for some point $u \in C$, then (u_n) converges strongly to a fixed point of U in C (the closest one to u).

One can easily see that $\lambda_n := \frac{1}{n+1}$ satisfies conditions $(C1)$-$(C3)$. Notably, in the particular case when U is linear and $\lambda_n := \frac{1}{n+1}$, the Halpern iteration becomes the usual ergodic average. Thus, Wittmann's result is a nonlinear generalization of the von Neumann mean ergodic theorem and, in this sense, (u_n) can be seen as a nonlinear ergodic average.

In this section we show how, using Propositions 4.14 and 5.2, a quantitative version of Wittmann's theorem can be obtained. Here, for simplicity, we will only consider the case $\lambda_n := \frac{1}{n+1}$. The general case can, however, be argued as a particular instance of Baushcke's theorem analyzed in Section 5.4.

As in the case of Browder's theorem, we work with the bounded metric space C with the metric induced by the Hilbert space norm. Let $b \in \mathbb{N}$ be a positive upper bound on the diameter of C. This time, we will be using the function

$$\varphi(x, y) := \langle x - u_0, U(y) \rangle.$$

This different function ψ, requires us to adjust Proposition 4.16. For every $k \in \mathbb{N}$ and every $g : \mathbb{N} \to \mathbb{N}$, let

$$\gamma_{k,g} : \mathbb{N} \to \mathbb{N} \quad \text{with} \quad \gamma_{k,g}(m) = \max\{g(m), 2b(k + 1)\}. \tag{5.9}$$

Proposition 5.5. *For any natural number k and monotone function $f : \mathbb{N} \to \mathbb{N}$, there are $N \leq 12b(\omega_{\gamma_{k,f}}^{(r(2k+1))}(0)+1)^2$ and $x \in C$ satisfying*

$$\|U(x)-x\| \leq \frac{1}{f(N)+1} \wedge \forall y \in C \left(\|U(y)-y\| \leq \frac{1}{N+1} \to \langle x-u_0, U(x)-U(y)\rangle \leq \frac{1}{k+1} \right),$$

where $\omega_{\gamma_{k,f}}$ is defined by (5.5) and r is defined by (5.4).

Proof. Let $k \in \mathbb{N}$ and $f : \mathbb{N} \to \mathbb{N}$ be monotone. Applying Proposition 4.16 to $2k+1$ and to the monotone function $\gamma_{k,f}$, we get the existence of $N \leq 12b(\omega_{\gamma_{k,f}}^{(r(2k+1))}(0)+1)^2$ and $x \in C$ such that

$$\|U(x)-x\| \leq \frac{1}{\gamma_{k,f}(N)+1}$$

and

$$\forall y \in C \left(\|U(y)-y\| \leq \frac{1}{N+1} \to \langle x-u_0, x-y\rangle \leq \frac{1}{2(k+1)} \right). \tag{5.10}$$

Since $\gamma_{k,f}(N) \geq f(N)$, we have that

$$\|U(x)-x\| \leq \frac{1}{f(N)+1}. \tag{5.11}$$

Consider now $y \in C$ satisfying $\|U(y)-y\| \leq \frac{1}{N+1}$. As U is nonexpansive, we have that $\|U(U(y))-U(y)\| \leq \frac{1}{N+1}$. Thus, we can apply (5.10) to $U(y)$ and conclude that

$$\langle x-u_0, x-U(y)\rangle \leq \frac{1}{2(k+1)}.$$

We have

$$\begin{aligned} \langle x-u_0, U(x)-U(y)\rangle &\leq \langle x-u_0, U(x)-x\rangle + \langle x-u_0, x-U(y)\rangle \\ &\leq b\|U(x)-x\| + \frac{1}{2(k+1)} \leq \frac{b}{\gamma_{k,f}(N)+1} + \frac{1}{2(k+1)} \end{aligned}$$

Therefore, since we also have $\gamma_{k,f}(N) \geq 2b(k+1)$, we conclude

$$\langle x-u_0, U(x)-U(y)\rangle \leq \frac{1}{k+1}. \tag{5.12}$$

By (5.11) and (5.12), the result follows. □

Hence, condition (b) of Proposition 4.14 holds with

$$\beta(k,f) := 12b\left(\omega_{\gamma_{k,f}}^{(r(2k+1))}(0)+1\right)^2.$$

The bound α for (a) of Proposition 4.14 was computed in [27, Lemma 3.1]:

$$\alpha(k,f) := \alpha(k) := 4b(k+1)\big(4b(k+1)+2\big) = 16b^2(k+1)^2 + 8b(k+1).$$

From the quantitative version of the general principle (Proposition 4.14), we get that the condition (i) of Proposition 5.2 holds when

$$\begin{aligned}\psi(k,f) &= \alpha\left(\beta\left(k,\widehat{f}\right),f\right) = 16b^2\left(\beta\left(k,\widehat{f}\right)+1\right)^2 + 8b\left(\beta\left(k,\widehat{f}\right)+1\right)\\ &= 16b^2\left(12b\left(\omega^{(r(2k+1))}_{\gamma_{k,\widehat{f}}}(0)+1\right)^2+1\right)^2 + 8b\left(12b\left(\omega^{(r(2k+1))}_{\gamma_{k,\widehat{f}}}(0)+1\right)^2+1\right),\end{aligned}$$

where $\widehat{f}(m) := f(\alpha(m,f)) = f(16b^2(m+1)^2+8b(m+1))$.

For the condition (ii) of Proposition 5.2, we rely on the following result, which is an immediate consequence of the more general Proposition 5.16 (proved in section 5.4).

Proposition 5.6. *For all $x \in C$ and all $k,n,p \in \mathbb{N}$,*

$$\|U(x)-x\| \le \frac{1}{9b(k+1)^2(p+1)} \ \wedge\ \forall i \in [n,p]\left(\langle x-u_0, U(x)-U(x_i)\rangle \le \frac{1}{12(k+1)^2}\right)$$
$$\to\ \forall m \in [\sigma'(k,n),p]\left(\|u_m - x\| \le \frac{1}{k+1}\right),$$

where $\sigma'(k,n) := \exp\left(\tilde{n}+1+\lceil \ln(3b^2(k+1)^2)\rceil\right)$, *with* $\tilde{n} := \max\{n,\, 6b^2(k+1)^2\}$.

Applying Proposition 5.6 with $p := f(\sigma'(k,n))$, we get condition (ii) of Proposition 5.2 with the following monotone functions:

$$\gamma(k,n,f) := 9b(k+1)^2(f(\sigma'(k,n))+1)-1, \quad \delta(k) := 12(k+1)^2-1,$$
$$\eta(k,n,f) := f(\sigma'(k,n)) \quad \text{and} \quad \sigma(k,n,f) := \sigma'(k,n).$$

Finally, we can apply Proposition 5.2 to get the following quantitative version of Wittmann's theorem.

Theorem 5.7 (Quantitative Wittmann)**.** *Under the conditions of Wittmann's theorem, let $b \in \mathbb{N}$ be a positive upper bound on the diameter of C. Then, for all $k \in \mathbb{N}$ and every monotone function $f : \mathbb{N} \to \mathbb{N}$,*

$$\exists N \le \phi_{\mathrm{W}}(k,f)\, \forall i,j \in [N, f(N)]\left(\|u_i - u_j\| \le \frac{1}{k+1}\right),$$

where

$$\phi_{\mathrm{W}}(k,f) := \sigma'\left(2k+1, \psi\left(48(k+1)^2-1, \overline{f}\right)\right),$$

with σ' and ψ defined above and $\overline{f}(m) = 36b(k+1)^2(f(\sigma'(2k+1,m))+1)-1$.

5.3 Adapting the general principle

In this section, Bauschke's original proof of a generalization of Wittmann's theorem to a finite family of nonexpansive mappings [2] is analyzed. Bauschke's proof has the same structure of Wittmann's proof and depends on a sequential weak compactness argument in the same way. Because of this, first we adapt the general principle to deal with a finite number of maps and then we will use this extended principle to bypass the sequential weak compactness argument in the mining of Bauschke's theorem. We will be following section 6 of [12].

Throughout this section, we fix a natural positive number ℓ. Let (X, d) be a bounded metric space and let $U_0, \ldots, U_{\ell-1}$ be mappings from X to X. Consider also mappings $\varphi_0, \ldots, \varphi_{\ell-1}$ from $X \times X$ to $\mathbb{R}$ and (u_n) a sequence of elements of X such that $\lim_n d(U_i(u_n), u_n) = 0$, for all $i < \ell$. We denote by

$$F := \bigcap_{i=0}^{\ell-1} Fix(U_i)$$

the set of common fixed points of the mappings $U_0, \ldots, U_{\ell-1}$.
The next result is an adaptation of the general principle that consider a finite number of maps.

Proposition 5.8. *The theory $\mathsf{PA}^{\omega}_{\trianglelefteq}[X, d_X, a_X, b_X]^+$ proves the following. Assume that*

$$\forall k \in \mathbb{N}\, \exists x \in F\, \forall y \in F\, \forall i < \ell \left(\varphi_i(x,x) \trianglelefteq \varphi_i(x,y) + \frac{1}{k+1} \right). \tag{5.13}$$

Then

$$\forall k \in \mathbb{N}\, \exists x \in F\, \exists n \in \mathbb{N}\, \forall m \geq n\, \forall i < \ell \left(\varphi_i(x,x) \trianglelefteq \varphi_i(x,u_m) + \frac{1}{k+1} \right). \tag{5.14}$$

Proof. Let $k \in \mathbb{N}$ be arbitrary.
By (5.13), there exists $\tilde{x} \in F$ such that $\forall y \in F \forall i < \ell \left(\varphi_i(\tilde{x}, \tilde{x}) \trianglelefteq \varphi_i(\tilde{x}, y) + \frac{1}{k+1} \right)$.
By the definition of F, we get that

$$\forall y \in X \left(\forall i < \ell\, \forall r \in \mathbb{N} \left(d(U_i(y), y) \trianglelefteq \frac{1}{r+1} \right) \to \forall i < \ell \left(\varphi_i(\tilde{x}, \tilde{x}) \trianglelefteq \varphi_i(\tilde{x}, y) + \frac{1}{k+1} \right) \right).$$

Therefore

$$\forall y \in X\, \exists r \in \mathbb{N} \left(\forall i < \ell \left(d(U_i(y), y) \trianglelefteq \frac{1}{r+1} \right) \to \forall i < \ell \left(\varphi_i(\tilde{x}, \tilde{x}) \trianglelefteq \varphi_i(\tilde{x}, y) + \frac{1}{k+1} \right) \right).$$

We can apply $\mathsf{BC}^{\omega}_{\mathsf{bd}}$ to obtain

$$\exists r \in \mathbb{N}\, \forall y \in X \left(\forall i < \ell \left(d(U_i(y), y) \trianglelefteq \frac{1}{r+1}\right) \to \forall i < \ell \left(\varphi_i(\tilde{x}, \tilde{x}) \trianglelefteq \varphi_i(\tilde{x}, y) + \frac{1}{k+1}\right)\right).$$

Take r_0 to be one such r. Since $\lim_n d(U_i(u_n), u_n) = 0$ for all $i < \ell$, we have that

$$\forall i < \ell\, \forall r \in \mathbb{N}\, \exists n \in \mathbb{N}\, \forall m \geq n \left(d(U_i(u_m), u_m) \trianglelefteq \frac{1}{r+1}\right).$$

Note that the bounded quantification "$\forall i < \ell$" stands for a finite conjunction because ℓ is a fixed natural number. We really have

$$\forall r \in \mathbb{N} \bigwedge_{i<\ell} \exists n \in \mathbb{N}\, \forall m \geq n \left(d(U_i(u_m), u_m) \trianglelefteq \frac{1}{r+1}\right).$$

Therefore, we can easily obtain

$$\forall r \in \mathbb{N} \exists n \in \mathbb{N} \bigwedge_{i<\ell} \forall m \geq n \left(d(U_i(u_m), u_m) \trianglelefteq \frac{1}{r+1}\right).$$

That is,

$$\forall r \in \mathbb{N} \exists n \in \mathbb{N} \forall m \geq n \forall i < \ell \left(d(U_i(u_m), u_m) \trianglelefteq \frac{1}{r+1}\right).$$

The result now follows by instantiating r by r_0. $\square$

When $\ell = 1$, the above proposition reduces to Proposition 4.13 and so it is a proper generalization of the general principle. Moreover, since the definition of the new set F only differs from the original fixed point set by the finite conjunction '$\forall i < \ell$', it is easy to obtain the interpretation of the statements in Proposition 5.8 in a similar way to before. The next result is the quantitative version of Proposition 5.8 and can be argued essentially in the same manner as in the proof of Proposition 4.14.

Proposition 5.9 (Quantitative version of 5.8)**.** *Suppose that there are monotone functionals α and β from $\mathbb{N} \times \mathbb{N}^{\mathbb{N}}$ to $\mathbb{N}$ satisfying:*

(a) $\forall k \in \mathbb{N}\, \widetilde{\forall} f \in \mathbb{N}^{\mathbb{N}}\, \exists N \leq \alpha(k, f)\, \forall n \in [N, f(N)]\, \forall i < \ell \left(d(U_i(u_n), u_n) \leq \dfrac{1}{k+1}\right)$*;*

(b) $\forall k \in \mathbb{N}\, \widetilde{\forall} f \in \mathbb{N}^{\mathbb{N}}\, \exists N \leq \beta(k, f)\, \exists x \in X \left(\forall i < \ell \left(d(U_i(x), x) \leq \dfrac{1}{f(N)+1}\right)\right.$
$\left.\wedge\ \forall y \in X \left(\forall i < \ell \left(d(U_i(y), y) \leq \dfrac{1}{N+1}\right) \to \forall i < \ell \left(\varphi_i(x, x) \leq \varphi_i(x, y) + \dfrac{1}{k+1}\right)\right)\right).$

Then, for every $k \in \mathbb{N}$ and any monotone function $f \in \mathbb{N}^{\mathbb{N}}$, there are $N \leq \psi(k, f)$ and $x \in X$ such that

$$\forall i < \ell \left(d(U_i(x), x) \leq \frac{1}{f(N)+1} \right) \wedge \forall n \in [N, f(N)] \forall i < \ell \left(\varphi_i(x, x) \leq \varphi_i(x, u_n) + \frac{1}{k+1} \right),$$

where $\psi(k, f)$ is defined as in Proposition 4.14.

In analogy with Lemma 5.1, a '*modus ponens*' lemma is also true in this case:

Lemma 5.10. *The theory $\mathsf{PA}^{\omega}_{\trianglelefteq}[X, d_X, a_X, b_X]$ proves the following. Suppose that*

$$\forall k \in \mathbb{N}\, \exists x \in F\, \exists n \in \mathbb{N}\, \forall m \geq n\, \forall i < \ell \left(\varphi_i(x, x) \trianglelefteq \varphi_i(x, u_m) + \frac{1}{k+1} \right)$$

and that there is a monotone function $\delta : \mathbb{N} \to \mathbb{N}$ such that, for all $k \in \mathbb{N}$ and $x \in F$,

$$\exists n \in \mathbb{N}\, \forall m \geq n\, \forall i < \ell \left(\varphi_i(x, x) \trianglelefteq \varphi_i(x, u_m) + \frac{1}{\delta(k)+1} \right)$$
$$\to \exists M \in \mathbb{N}\, \forall m \geq M \left(d(u_m, x) \trianglelefteq \frac{1}{k+1} \right).$$

Then, (u_n) is a Cauchy sequence.

By the same reasoning of the proof of Proposition 5.2, we get the following quantitative version of Lemma 5.10:

Proposition 5.11 (Quantitative version of 5.10)**.** *Suppose that there are monotone functions $\delta, \psi, \gamma, \eta$ and σ satisfying:*

(i) $$\forall k \in \mathbb{N}\, \widetilde{\forall} f \in \mathbb{N}^{\mathbb{N}}\, \exists N \leq \psi(k, f)\, \exists x \in X \left(\forall i < \ell \left(d(U_i(x), x) \leq \frac{1}{f(N)+1} \right) \right.$$
$$\left. \wedge\ \forall n \in [N, f(N)]\, \forall i < \ell \left(\varphi_i(x, x) \leq \varphi_i(x, u_n) + \frac{1}{k+1} \right) \right)$$

and

(ii) $$\forall k, n \in \mathbb{N}\, \widetilde{\forall} f \in \mathbb{N}^{\mathbb{N}}\, \forall x \in X \left[\left(\forall i < \ell \left(d(U_i(x), x) \leq \frac{1}{\gamma(k, n, f)+1} \right) \right. \right.$$
$$\left. \wedge\ \forall m \in [n, \eta(k, n, f)]\, \forall i < \ell \left(\varphi_i(x, x) \leq \varphi_i(x, u_m) + \frac{1}{\delta(k)+1} \right) \right)$$
$$\left. \to \exists M \leq \sigma(k, n, f)\, \forall m \in [M, f(M)] \left(d(u_m, x) \leq \frac{1}{k+1} \right) \right]$$

Then

$$\forall k \in \mathbb{N}\, \widetilde{\forall} f \in \mathbb{N}^{\mathbb{N}}\, \exists M \leq \phi(k,f)\, \forall m,n \in [M, f(M)] \left(d(u_m, u_n) \leq \frac{1}{k+1} \right), \tag{5.15}$$

where $\phi(k,f)$ is defined as in Proposition 5.2.

Note that the above quantitative versions are also true when the metric space is unbounded.

5.4 Bauschke's theorem

In the following, X is a Hilbert space, C is a nonempty closed convex bounded subset of X, $b \in \mathbb{N}$ is a positive upper bound on the diameter of C and $T_0, \ldots, T_{\ell-1}$ are nonexpansive selfmappings of C. Let F be the set of common fixed points of the mappings $T_0, \ldots, T_{\ell-1}$. For each $n \in \mathbb{N}$, define the mapping

$$U_n := T_{n \bmod \ell}. \tag{5.16}$$

Obviously, $U_i = T_i$ for all $i < \ell$ and $F = \bigcap_{i=0}^{\ell-1} Fix(U_i) = \bigcap_{n\in\mathbb{N}} Fix(U_n)$.
Let (λ_n) be a sequence in $(0,1)$ satisfying the conditions

$$(C1) \quad \lim \lambda_n = 0, \quad (C2) \quad \sum_{n=1}^{\infty} \lambda_n = \infty, \quad (C3[\ell]) \quad \sum_{n=1}^{\infty} |\lambda_n - \lambda_{n+\ell}| < \infty.$$

Given $u \in C$, we define the sequence (u_n) by

$$u_0 := u, \quad u_{n+1} := \lambda_{n+1} u_0 + (1 - \lambda_{n+1}) U_{n+1}(u_n), \tag{5.17}$$

which is a generalization of the Halpern iteration to the sequence of maps (U_n).

The following theorem was proved by Heinz Bauschke in [2].

Theorem 5.12 (Bauschke)**.** *With the above assumptions, suppose furthermore that*

$$F = Fix(T_{\ell-1} \cdots T_1 T_0) = Fix(T_0 T_{\ell-1} \cdots T_1) = \cdots = Fix(T_{\ell-2} \cdots T_0 T_{\ell-1}). \tag{5.18}$$

Then (u_n) converges strongly to a common fixed point of $T_0, \ldots, T_{\ell-1}$ (the closest one to u).

Condition $(C3[1])$ is the same as $(C3)$. Thus, when $\ell = 1$, we get Wittmann's theorem as a particular case of Theorem 5.12.

We remark first that (5.18) is equivalent to

$$F = Fix(U_{m+\ell} \cdots U_{m+1}) \quad \text{for all } m \in \mathbb{N}. \tag{5.19}$$

The left-to-right inclusion is obvious. Therefore, (5.19) holds if, and only if,

$$\forall m \in \mathbb{N} \, (F \supseteq Fix(U_{m+\ell} \cdots U_{m+1}))\,.$$

The above statement can be rewritten as

$$\begin{aligned}&\forall m \in \mathbb{N} \, \forall x \in C \\ &\left(\forall r \in \mathbb{N} \left(\|x - U_{m+\ell} \cdots U_{m+1}(x)\| \trianglelefteq \frac{1}{r+1}\right) \to \forall i < \ell \, \forall k \in \mathbb{N} \left(\|x - U_i(x)\| \trianglelefteq \frac{1}{k+1}\right)\right).\end{aligned}$$

Since the mappings U_m are defined cyclically, the quantification "$\forall m \in \mathbb{N}$" above can be seen as bounded, i.e. it can be replaced by a bounded quantification. Therefore, we get in our formal setting, by using $\mathsf{BC}^{\omega}_{\mathsf{bd}}$, that

$$\begin{aligned}&\forall k \in \mathbb{N} \, \exists r \in \mathbb{N} \, \forall m \in \mathbb{N} \, \forall x \in C \\ &\left(\|x - U_{m+\ell} \cdots U_{m+1}(x)\| \trianglelefteq \frac{1}{r+1} \to \forall i < \ell \left(\|x - U_i(x)\| \trianglelefteq \frac{1}{k+1}\right)\right).\end{aligned} \tag{5.20}$$

Hence, for the quantitative version of (5.20), we ask for a monotone function $\tau : \mathbb{N} \to \mathbb{N}$ satisfying

$$\begin{aligned}&\forall k \in \mathbb{N} \, \forall m \in \mathbb{N} \, \forall x \in C \\ &\left(\|x - U_{m+\ell} \cdots U_{m+1}(x)\| \leq \frac{1}{\tau(k)+1} \to \forall i < \ell \left(\|x - U_i(x)\| \leq \frac{1}{k+1}\right)\right).\end{aligned} \tag{5.21}$$

For the quantitative versions of the conditions $(C1)$, $(C2)$ and $(C3[\ell])$ on the sequence (λ_n) we will assume the existence of monotone functions $\mu, \nu, \xi : \mathbb{N} \to \mathbb{N}$ satisfying:

1. μ is a rate of convergence for (λ_n) towards zero, that is

$$(C1_q) \quad \forall k \in \mathbb{N} \, \forall n \geq \mu(k) \left(\lambda_n \leq \frac{1}{k+1}\right);$$

2. ν is a rate of divergence for $\sum_n \lambda_n$, that is

$$(C2_q) \quad \forall k \in \mathbb{N} \left(\sum_{j=0}^{\nu(k)} \lambda_j \geq k\right);$$

3. ξ is a Cauchy modulus for the series $\sum_n |\lambda_n - \lambda_{n+\ell}|$, that is

$$(C3[\ell]_q) \quad \forall k \in \mathbb{N} \, \forall n \in \mathbb{N} \left(\sum_{j=\xi(k)+1}^{\xi(k)+n} |\lambda_j - \lambda_{j+\ell}| \leq \frac{1}{k+1}\right).$$

Note that $\nu(k+1) \geq k$. In the sequel, we prove some useful properties of the sequence (u_n). First, let us remark that, for all $n \in \mathbb{N} \setminus \{0\}$ and all $m \in \mathbb{N}$,

$$\|u_{n+m+\ell} - u_{n+m}\| \leq b \cdot \sum_{j=n}^{n+m} |\lambda_{j+\ell} - \lambda_i| + \|u_{n+\ell-1} - u_{n-1}\| \cdot \prod_{j=n}^{n+m} (1 - \lambda_{j+\ell}). \tag{5.22}$$

The proof is an easy induction on m (see the proof of [2, Theorem 3.1]).

Lemma 5.13. *For each $k \in \mathbb{N}$, the following holds:*

(i) $\forall n \geq \mu(b(k+1)) \left(\|u_{n+1} - U_{n+1}(u_n)\| \leq \frac{1}{k+1}\right)$

(ii) $\forall n \geq \chi(k) \left(\|u_{n+\ell} - u_n\| \leq \frac{1}{k+1}\right)$,
where $\chi(k) := \nu(\xi(2b(k+1)) + 1 + \ell + \lceil \ln(2b(k+1)) \rceil)$.

(iii) $\forall n \geq \widetilde{\alpha}(k) \left(\|u_n - U_{n+\ell} \cdots U_{n+1}(u_n)\| \leq \frac{1}{k+1}\right)$,
where $\widetilde{\alpha}(k) := \max\{\mu(2\ell b(k+1)), \chi(2k+1)\}$.

(iv) $\forall n \geq \widehat{\alpha}(k) \, \forall i < \ell \left(\|u_n - U_i(u_n)\| < \frac{1}{k+1}\right)$,
where $\widehat{\alpha}(k) := \widetilde{\alpha}(\tau(k))$*, with* τ *satisfying* (5.21).

Proof. (i) Since $u_{n+1} = \lambda_{n+1} u_0 + (1 - \lambda_{n+1}) U_{n+1}(u_n)$, for $n \geq \mu(b(k+1))$ we have

$$\|u_{n+1} - U_{n+1}(u_n)\| = \lambda_{n+1} \|u_0 - U_{n+1}(u_n)\| \leq \lambda_{n+1} b \leq \frac{1}{k+1}.$$

(ii) Let $N := \xi(2b(k+1)) + 1$. Applying (5.22) with $n := N$ and using $(C3[\ell]_q)$ and the fact that $1 - x \leq \exp(-x)$ for $x \geq 0$, we get that for all $m \in \mathbb{N}$,

$$\|u_{N+m+\ell} - u_{N+m}\| \leq \frac{1}{2(k+1)} + b \cdot \exp\left(-\sum_{j=N}^{N+m} \lambda_{j+\ell}\right) \tag{5.23}$$

Let $M := \chi(k) - N = \nu(N + \ell + \lceil \ln(2b(k+1)) \rceil) - N$. By $(C2_q)$, it follows that for all $m \geq M$,

$$\sum_{i=0}^{N+m+\ell} \lambda_i \;\geq\; \sum_{i=0}^{N+M} \lambda_i \;\geq\; N + \ell + \lceil \ln(2b(k+1)) \rceil \;\geq\; \sum_{i=0}^{N+\ell-1} \lambda_i + \ln(2b(k+1)).$$

Therefore, $\sum\limits_{i=N}^{N+m} \lambda_{i+\ell} = \sum\limits_{i=N+\ell}^{N+m+\ell} \lambda_i \geq \ln(2b(k+1))$, which yields

$$b \cdot \exp\left(-\sum_{i=N}^{N+m} \lambda_{i+\ell}\right) \leq \frac{1}{2(k+1)}. \tag{5.24}$$

Now, apply (5.23) and (5.24) to get (ii).

(iii) Let $n \geq \widetilde{\alpha}(k)$ be arbitrary. For every $1 \leq i \leq \ell$, let $S_i := U_{n+i} \cdots U_{n+1}$. We get

$$\begin{aligned} \|u_n - U_{n+\ell} \cdots U_{n+1}(u_n)\| &= \|u_n - S_\ell(u_n)\| \leq \|u_n - u_{n+\ell}\| + \|u_{n+\ell} - S_\ell(u_n)\| \\ &\leq \frac{1}{2(k+1)} + \|u_{n+\ell} - S_\ell(u_n)\| \end{aligned}$$

The inequality is explained by (ii), given that $n \geq \chi(2k+1)$.

Remark that

$$\begin{aligned} \|u_{n+\ell} - S_\ell(u_n)\| &\leq \|u_{n+\ell} - U_{n+\ell}(u_{n+\ell-1})\| + \|U_{n+\ell}(u_{n+\ell-1}) - S_\ell(u_n)\| \\ &\leq \|u_{n+\ell} - U_{n+\ell}(u_{n+\ell-1})\| + \|u_{n+\ell-1} - S_{\ell-1}(u_n)\|, \end{aligned}$$

since $U_{n+\ell}$ is nonexpansive. Reasoning in the same way, it follows that

$$\|u_{n+\ell} - S_\ell(u_n)\| \leq \sum_{i=1}^{\ell} \|u_{n+i} - U_{n+i}(u_{n+i-1})\| \leq \frac{\ell}{2\ell(k+1)} = \frac{1}{2(k+1)},$$

by (i), given that $n \geq \mu(2\ell b(k+1))$.

(iv) Just apply (iii) and (5.21).

$\square$

We show now that the quantitative version of Bauschke's theorem can be obtained by applying the quantitative Propositions 5.9 and 5.11. The relevant functionals φ_i in this application are,

$$\varphi_i(x, y) := \langle x - u_0, U_i(y) \rangle, \quad \text{for } i < \ell. \tag{5.25}$$

Note that, as an immediate consequence of Lemma 5.13.(iv), the functional

$$\alpha : \mathbb{N} \times \mathbb{N}^{\mathbb{N}} \to \mathbb{N}, \quad \alpha(k, f) := \widehat{\alpha}(k) \tag{5.26}$$

satisfies condition (a) of Proposition 5.9.

In the sequel, we show how to compute a functional β satisfying condition (b) of Proposition 5.9. We consider the projection onto a different set F than the one in Section 4.3. Since now F is $\{x \in C \,|\, \forall i < \ell \,(U_i(x) = x)\}$, the only difference to the analysis of the projection argument is in the innocuous addition of the finite conjunction "$\forall i < \ell$." We get, using similar arguments to the ones used in the proof of Proposition 4.16, the following result:

Proposition 5.14. *For any natural number k and monotone function $f : \mathbb{N} \to \mathbb{N}$, there are $N \leq 12b(\omega_{\gamma_{k,f}}^{(r(k))}(0) + 1)^2$ and $x \in C$ such that*

$$\forall i < \ell \left(\|U_i(x) - x\| \leq \frac{1}{f(N) + 1} \right)$$

and
$$\forall y \in C\left(\forall i < \ell\left(\|U_i(y) - y\| \le \frac{1}{N+1}\right) \to \langle x - u_0, x - y\rangle \le \frac{1}{k+1}\right),$$
where r is defined by (5.4) *and and $\omega_{(\cdot)}$ is defined by* (5.5).

We must change the conclusion of the implication to be compatible with our functions φ_i. I.e., we must replace the conclusion $\langle x - u_0, x - y\rangle \le \frac{1}{k+1}$ by
$$\forall i < \ell\left(\langle x - u_0, U_i(x) - U_i(y)\rangle \le \frac{1}{k+1}\right).$$
This is done, in two steps, in the proposition below.

Proposition 5.15. *Let $k \in \mathbb{N}$ and $f : \mathbb{N} \to \mathbb{N}$ be monotone.*

(i) *There exist $N_0 \in \mathbb{N}$ with $N_0 \le \beta_0(k, f)$ and $x \in C$ such that*

(a_0) $\|U_i(x) - x\| \le \frac{1}{f(N_0)+1}$ *for all $i < \ell$, and*

(b_0) *for all $y \in C$, $\forall i < \ell\left(\|U_i(y) - y\| \le \frac{1}{N_0+1}\right) \to \forall i < \ell\left(\langle x - u_0, U_i(x) - y\rangle \le \frac{1}{k+1}\right)$;*

(ii) *There exist $N \in \mathbb{N}$ with $N \le \beta(k, f)$ and $x \in C$ such that*

(a) $\|U_i(x) - x\| \le \frac{1}{f(N)+1}$ *for all $i < \ell$, and*

(b) *for all $z \in C$,*
$\forall i < \ell\left(\|U_i(z) - z\| \le \frac{1}{N+1}\right) \to \forall i < \ell\left(\langle x - u_0, U_i(x) - U_i(z)\rangle \le \frac{1}{k+1}\right),$

where
$$\beta_0(k, f) := 12b(\omega_{\gamma_{k,f}}^{(r(2k+1))}(0) + 1)^2 \quad \textit{and} \quad \beta(k, f) = 3\beta_0(k, g) + 2, \tag{5.27}$$
with $g(m) = f(3m + 2)$.

Proof. 1. Applying Proposition 5.14 to $2k + 1$ and to the monotone function $\gamma_{k,f}$, we get $N_0 \le \beta_0(k, f)$ and $x \in C$ such that $\|U_i(x) - x\| \le \frac{1}{\gamma_{k,f}(N_0)+1}$ for all $i < \ell$ and, for all $y \in C$,
$$\forall i < \ell\left(\|U_i(y) - y\| \le \frac{1}{N_0 + 1}\right) \to \langle x - u_0, x - y\rangle \le \frac{1}{2(k+1)}. \tag{5.28}$$

By the definition of $\gamma_{k,f}$, we have that, for all $i < \ell$, $\|U_i(x) - x\| \le \frac{1}{\gamma_{k,f}(N_0)+1} \le \frac{1}{f(N_0)+1}$. Thus, ($a_0$) holds. Let now $y \in C$ be such that the premise of the implication in (b_0)

holds and let $i < \ell$ be arbitrary. It follows that

$$\begin{aligned}
\langle x - u_0, U_i(x) - y\rangle &= \langle x - u_0, U_i(x) - x\rangle + \langle x - u_0, x - y\rangle \\
&\leq b \cdot \|U_i(x) - x\| + \frac{1}{2(k+1)} \quad \text{by (5.28)} \\
&\leq \frac{1}{2(k+1)} + \frac{1}{2(k+1)} = \frac{1}{k+1},
\end{aligned}$$

since $\|U_i(x) - x\| \leq \frac{1}{\gamma_{k,f}(N)+1} \leq \frac{1}{2b(k+1)}$. Hence, (b_0) holds too.

2. Apply (i) for k and g to get $N_0 \leq \beta_0(k, g)$ and $x \in C$ satisfying (a_0) for g and (b_0). Let $N := 3N_0 + 2 \leq 3\beta_0(k, g) + 2 = \beta(k, f)$. Then, for all $i < \ell$, we have that $\|U_i(x) - x\| \leq \frac{1}{g(N_0)+1} = \frac{1}{f(N)+1}$, so (a) holds. In order to prove (b), assume that $z \in C$ is such that $\forall i < \ell \left(\|U_i(z) - z\| \leq \frac{1}{N+1}\right)$.

   For all $i, j < \ell$, we have that

$$\begin{aligned}
\|U_i(U_j(z)) - U_j(z)\| &\leq \|U_i(U_j(z)) - U_i(z)\| + \|U_i(z) - z\| + \|z - U_j(z)\| \\
&\leq \|U_j(z) - z\| + \frac{2}{N+1} \leq \frac{3}{N+1} = \frac{1}{N_0+1}.
\end{aligned}$$

   Thus, we can apply (b_0) for $y := U_j(z)$, with $j < \ell$ arbitrary, and conclude

$$\forall i < \ell \, \forall j < \ell \left(\langle x - u_0, U_i(x) - U_j(z)\rangle \leq \frac{1}{k+1}\right).$$

   Take $j := i$ above to get (b).

□

Thus, we can apply Proposition 5.9 to get, for every $k \in \mathbb{N}$ and any monotone function $f \in \mathbb{N}^{\mathbb{N}}$, an $N \in \mathbb{N}$ with $N \leq \psi(k, f)$ and $x \in C$ such that

$$\forall i < \ell \left(\|U_i(x) - x\| \leq \frac{1}{f(N)+1}\right)$$

and

$$\forall n \in [N, f(N)] \, \forall i < \ell \left(\langle x - u_0, U_i(x) - U_i(u_n)\rangle \leq \frac{1}{k+1}\right),$$

where

$$\psi(k, f) := \alpha\left(\beta\left(k, \widehat{f}\right), f\right) = \widehat{\alpha}\left(\beta\left(k, \widehat{f}\right)\right), \quad \text{with } \widehat{f}(m) := f(\alpha(m, f)) = f(\widehat{\alpha}(m)). \tag{5.29}$$

Hence, condition (i) of Proposition 5.11 is satisfied with ψ as above.
Next, we present the quantitative result of the main combinatorial step in Bauschke's proof, slightly adapted to fit into the general principle. This will allow us to see that (ii) of Proposition 5.11, with appropriate bounds, also holds.

Proposition 5.16. *Assume that $k, n, p \in \mathbb{N}$ and $x \in C$ satisfy*

$$\forall i < \ell \left(\|U_i(x) - x\| \leq \frac{1}{9b(k+1)^2(p+1)} \right)$$

and

$$\forall r \in [n,p]\, \forall i < \ell \left(\langle x - u_0, U_i(x) - U_i(u_r) \rangle \leq \frac{1}{12(k+1)^2} \right).$$

Then

$$\forall m \in [\sigma'(k,n), p] \left(\|u_m - x\| \leq \frac{1}{k+1} \right),$$

where $\sigma'(k,n) := \nu(\widetilde{n} + 1 + \lceil \ln(3b^2(k+1)^2) \rceil)$ with $\widetilde{n} := \max\{n,\, \mu(6b^2(k+1)^2)\}$.

Proof. First, let us remark that for all $r \in \mathbb{N}$ and $x \in C$,

$$\begin{aligned}
\|u_{r+1} - x\|^2 &= \|\lambda_{r+1}u_0 + (1-\lambda_{r+1})U_{r+1}(u_r) - x\|^2 \\
&= \|\lambda_{r+1}(u_0 - x) + (1-\lambda_{r+1})(U_{r+1}(u_r) - x)\|^2 \\
&= \lambda_{r+1}^2\|u_0 - x\|^2 + 2\lambda_{r+1}(1-\lambda_{r+1})\langle u_0 - x, U_{r+1}(u_r) - x\rangle \\
&\quad +(1-\lambda_{r+1})^2\|U_{r+1}(u_r) - x\|^2 \\
&= \lambda_{r+1}^2\|u_0 - x\|^2 + 2\lambda_{r+1}(1-\lambda_{r+1})(\langle x - u_0, x - U_{r+1}(x)\rangle \\
&\quad + \langle x - u_0, U_{r+1}(x) - U_{r+1}(u_r)\rangle \\
&\quad +(1-\lambda_{r+1})^2\|(U_{r+1}(u_r) - U_{r+1}(x)) + (U_{r+1}(x) - x)\|^2 \\
&\leq \lambda_{r+1}^2 b^2 + 2b\lambda_{r+1}(1-\lambda_{r+1})\|x - U_{r+1}(x)\| \\
&\quad +2\lambda_{r+1}(1-\lambda_{r+1})\langle x - u_0, U_{r+1}(x) - U_{r+1}(u_r)\rangle \\
&\quad +(1-\lambda_{r+1})^2(\|u_r - x\|^2 + 2\|u_r - x\|\|U_{r+1}(x) - x\| + \|U_{r+1}(x) - x\|^2) \\
&\leq \lambda_{r+1}^2 b^2 + 2b\lambda_{r+1}(1-\lambda_{r+1})\|x - U_{r+1}(x)\| \\
&\quad +2\lambda_{r+1}(1-\lambda_{r+1})\langle x - u_0, U_{r+1}(x) - U_{r+1}(u_r)\rangle \\
&\quad +3b(1-\lambda_{r+1})^2\|x - U_{r+1}(x)\| + (1-\lambda_{r+1})^2\|u_r - x\|^2 \\
&\leq \lambda_{r+1}^2 b^2 + 2\lambda_{r+1}(1-\lambda_{r+1})\langle x - u_0, U_{r+1}(x) - U_{r+1}(u_r)\rangle \\
&\quad +(2b\lambda_{r+1}(1-\lambda_{r+1}) + 3b(1-\lambda_{r+1})^2)\|x - U_{r+1}(x)\| + (1-\lambda_{r+1})^2\|u_r - x\|^2 \\
&\leq \lambda_{r+1}(\lambda_{r+1}b^2 + 2\langle x - u_0, U_{r+1}(x) - U_{r+1}(u_r)\rangle) \\
&\quad +3b(1-\lambda_{r+1})\|x - U_{r+1}(x)\| + (1-\lambda_{r+1})\|u_r - x\|^2
\end{aligned}$$

Fix $k, n, p \in \mathbb{N}$ and $x \in C$, and assume that they satisfy the hypothesis of the theorem. Take $r \in \mathbb{N}$ with $r \in [\tilde{n}, p] \subseteq [n, p]$. Then $\lambda_{r+1} \leq \frac{1}{6b^2(k+1)^2}$, since $r + 1 > \tilde{n} \geq \mu(6b^2(k+1)^2)$ and μ satisfies $(C1)_q$. Moreover, $\langle x - u_0, U_{r+1}(x) - U_{r+1}(u_r)\rangle \leq \frac{1}{12(k+1)^2}$ by hypothesis. Hence,

$$\lambda_{r+1}b^2 + 2\langle x - u_0, U_{r+1}(x) - U_{r+1}(u_r)\rangle \leq \frac{1}{3(k+1)^2}.$$

Furthermore, as $\|x - U_{r+1}(x)\| \leq \frac{1}{9b(k+1)^2(p+1)}$, we get that

$$\|u_{r+1} - x\|^2 \leq \lambda_{r+1}\frac{1}{3(k+1)^2} + (1-\lambda_{r+1})\frac{1}{3(k+1)^2(p+1)} + (1-\lambda_{r+1})\|u_r - x\|^2.$$

By induction on m, we can prove that for $m \in [\tilde{n}+1, p]$,

$$\|u_m - x\|^2 \leq \frac{1}{3(k+1)^2} + A_m\frac{1}{3(k+1)^2(p+1)} + B_m\|u_{\tilde{n}} - x\|^2 \tag{5.30}$$

where $A_m = \sum_{j=\tilde{n}}^{m-1}\prod_{i=j}^{m-1}(1-\lambda_{i+1})$ and $B_m = \prod_{i=\tilde{n}}^{m-1}(1-\lambda_{i+1})$.
Let $m \in [\sigma'(k,n), p]$ be arbitrary. Since $m \leq p$, we have that $A_m \leq m - \tilde{n} < p+1$. As $\sigma'(k,n) \geq \tilde{n}+1$, it follows that $m \in [\tilde{n}+1, p]$, so we can apply (5.30) and get

$$\|u_m - x\|^2 \leq \frac{1}{3(k+1)^2} + \frac{1}{3(k+1)^2} + B_m\|u_{\tilde{n}} - x\|^2 \tag{5.31}$$

Now, because $m \geq \sigma'(k,n)$, we get

$$\sum_{j=0}^{m}\lambda_j \geq \sum_{j=0}^{\sigma'(k,n)}\lambda_j \geq \tilde{n} + 1 + \ln(3b^2(k+1)^2) \geq \sum_{j=0}^{\tilde{n}}\lambda_j + \ln(3b^2(k+1)^2).$$

Therefore, $\sum_{j=\tilde{n}}^{m} -1\lambda_{j+1} = \sum\limits_{j=\tilde{n}+1}^{m}\lambda_j \geq \ln(3b^2(k+1)^2)$. This, in turn, implies

$$B_m\|u_{\tilde{n}} - x\|^2 \leq b^2\exp\left(-\sum_{j=\tilde{n}}^{m-1}\lambda_{j+1}\right) \leq \frac{1}{3(k+1)^2}. \tag{5.32}$$

The conclusion follows. $\square$

Note that Proposition 5.6 is the particular case of the above proposition. One can see this by putting $\ell = 1$ and $\lambda_n = \frac{1}{n+1}$, and taking into account that $\mu(n) = n$ is a rate of convergence towards 0 for the sequence $\left(\frac{1}{n+1}\right)$ and that $\nu(n) = \exp(n)$ is a rate of divergence for $\sum_n \frac{1}{n+1}$.
We are now in position to apply Proposition 5.16 with $p := f(\sigma'(k,n))$ in order to obtain condition (ii) of Proposition 5.11. Just let

$$\gamma(k,n,f) := 9b(k+1)^2(f(\sigma'(k,n))+1) - 1, \quad \delta(k) := 12(k+1)^2 - 1,$$
$$\eta(k,n,f) := f(\sigma'(k,n)) \quad \text{and} \quad M := \sigma(k,n,f) := \sigma'(k,n).$$

Finally, we apply Proposition 5.11 to obtain the metastable version of Bauschke's theorem.

Theorem 5.17. *Let X be a Hilbert space, C be a nonempty closed convex bounded subset of X, $b \in \mathbb{N}$ be a positive upper bound on the diameter of C and $T_0, \ldots, T_{\ell-1}$ be nonexpansive selfmappings of C.*
For each $n \in \mathbb{N}$, let U_n be the mapping defined by (5.16) and assume that $\tau : \mathbb{N} \to \mathbb{N}$ is a monotone function satisfying (5.21). Consider a sequence (λ_n) in $(0,1)$ and monotone functions $\mu, \nu, \xi : \mathbb{N} \to \mathbb{N}$ such that $(C1_q), (C2_q)$ and $(C3[\ell]_q)$ hold. Let $u_0 \in C$ be given and (u_n) be the iteration defined by (5.17).
Then, for all $k \in \mathbb{N}$ and every monotone function $f : \mathbb{N} \to \mathbb{N}$,

$$\exists N \leq \phi_{\mathrm{Bau}}(k,f)\, \forall i,j \in [N, f(N)] \left(\|u_i - u_j\| \leq \frac{1}{k+1} \right), \tag{5.33}$$

where

$$\phi_{\mathrm{Bau}}(k,f) := \sigma'\left(2k+1, \psi\left(48(k+1)^2 - 1, \overline{f}\right)\right),$$

with σ' defined in Proposition 5.16, ψ defined by (5.29) and the function $\overline{f}$ defined by $\overline{f}(m) = 36b(k+1)^2(f(\sigma'(2k+1,m))+1) - 1$.

Condition $(C2)$, $\sum_n \lambda_n = \infty$, is equivalent to the condition

$$(C4) \qquad \prod_{n\to\infty} (1-\lambda_n) = 0.$$

Hence, one can obtain general quantitative results by using, instead of a rate of divergence ν for $\sum_n \lambda_n$, the quantitative version of $(C4)$, asserting the existence of a rate of convergence θ for $\prod_{n\to\infty}(1-\lambda_n)$:

$$(C4_q) \qquad \forall k \in \mathbb{N} \left(\prod_{i=1}^{\theta(k)} (1-\lambda_i) \leq \frac{1}{k+1} \right).$$

This was done in [32], where Kohlenbach and Leuştean obtained rates of metastability for the generalization of Wittmann's theorem to CAT(0) spaces using both $(C2_q)$ and $(C4_q)$. As Kohlenbach remarked in [27], for $\lambda_n = \frac{1}{n+1}$, one has an exponential ν and a linear θ, so one gets, by using $(C4_q)$, a quadratic rate of asymptotic regularity for the Halpern iteration (see [27, Lemma 3.1]), significantly improving the exponential bound obtained in [39], where $(C2_q)$ is used. As a consequence, better rates of metastability for Wittmann's theorem are obtained in [27, 32] compared to our Theorem 5.7.
One can replace $(C2_q)$ with $(C4_q)$ also in the quantitative analysis of Bauschke's theorem and prove corresponding versions of Proposition 5.16 and Theorem 5.17 having as a consequence, for $\ell = 1$, a metastable version of Wittmann's theorem with bounds similar to the ones computed in [27, 32].

Next, we show how to adapt the previous results to use the condition $(C4_q)$ instead of $(C2_q)$. The equivalence between conditions $(C2)$ and $(C4)$ is only guaranteed if the sequence (λ_n) is strictly bounded below 1[1]. In turn, that condition implies $\prod_{i=1}^{n}(1-\lambda_i) > 0$, for all $n \in \mathbb{N}$. We then assume the existence of a monotone function $\epsilon : \mathbb{N} \to \mathbb{N}$ such that

$$\forall n \in \mathbb{N} \left(\prod_{i=1}^{n}(1-\lambda_i) \geq \frac{1}{\epsilon(n)+1} \right). \tag{5.34}$$

In order to compute a new function α, we must change Lemma 5.13. The next result replaces item (ii) in that Lemma. The property (i) still holds true and the remaining properties (iii) and (iv) hold with χ replaced by the function χ' defined below.

Lemma 5.18. *Let $b \in \mathbb{N}$ be a positive upper bound on the diameter of C. Let ξ, θ and ϵ be monotone functions satisfying $(C3[\ell]_q)$, $(C4_q)$ and (5.34), respectively. Then,*

$$\forall k \in \mathbb{N} \forall n \geq \chi'(k) \left(\|u_{n+\ell} - u_n\| \leq \frac{1}{k+1} \right). \tag{5.35}$$

where $\chi'(k) := \max\{\theta(2b(\epsilon(N+\ell-1)+1)(k+1)) - \ell, N\}$, with $N := \xi(2b(k+1)) + 1$.

Proof. Let $k \in \mathbb{N}$ be given and define $N := \xi(2b(k+1)) + 1$. By (5.22) with $n := N$ and using $(C3[\ell]_q)$, we get for all $m \in \mathbb{N}$,

$$\begin{aligned} \|u_{N+m+\ell} - u_{N+m}\| &\leq b \left(\sum_{j=N}^{N+m} |\lambda_{j+\ell} - \lambda_j| + \prod_{j=N}^{N+m} \lambda_{j+\ell} \right) \\ &\leq b \left(\frac{1}{2b(k+1)} + \prod_{j=N+\ell}^{N+m+\ell}(1-\lambda_j) \right). \end{aligned}$$

For $m \in \mathbb{N}$, we write $P_m := \prod_{j=1}^{m}(1-\lambda_j)$. Hence,

$$\|u_{N+m+\ell} - u_{N+m}\| \leq b \left(\frac{1}{2b(k+1)} + \frac{P_{N+m+\ell}}{P_{N+\ell-1}} \right),$$

Define $M := \max\{\theta(2b(\epsilon(N+\ell-1)+1)(k+1), N+\ell\} - N - \ell$. Then, for any $m \geq M$, we have $N+m+\ell \geq \theta(2b(\epsilon(N+\ell-1)+1)(k+1))$, which by $(C4_q)$ gives

$$P_{N+m+\ell} \leq \frac{1}{2b(\epsilon(N+\ell-1)+1)(k+1)}.$$

[1] The reason we wrote the product in $(C4_q)$ starting at $i = 1$, is just to be able to make some remarks for $\lambda_n = \frac{1}{n+1}$ below – note that for that sequence $\lambda_0 = 1$.

On the other hand, we have $\frac{1}{P_{N+\ell-1}} \leq \epsilon(N+\ell-1)+1$. Hence, for any $m \geq M$,

$$\|u_{N+m+\ell} - u_{N+m}\| \leq b\left(\frac{1}{2b(k+1)} + \frac{\epsilon(N+\ell-1)+1}{2b(\epsilon(N+\ell-1)+1)(k+1)}\right) = \frac{1}{k+1}.$$

Finally, see that $N + M := \chi'(k)$. □

The properties (*iii*) and (*iv*) from Lemma 5.13 hold, respectively, with the functions $\widetilde{\alpha}'$ and $\widehat{\alpha}'$ defined with this new function χ':

$$\widetilde{\alpha}'(k) := \max\{\mu(2\ell b(k+1)), \chi'(2k+1)\},$$

$$\widehat{\alpha}'(k) := \widetilde{\alpha}'(\tau(k)),$$

where τ satisfies (5.21).
Hence, condition (*a*) of Proposition 5.9 is satisfied by the functional

$$\alpha : \mathbb{N} \times \mathbb{N}^{\mathbb{N}} \to \mathbb{N}, \qquad \alpha(k, f) := \widehat{\alpha}'(k).$$

Now, by Proposition 5.9, the condition (i) of Proposition 5.11 is satisfied with ψ defined by

$$\psi(k, f) := \widehat{\alpha}'\left(\beta\left(k, \widehat{f}\right)\right), \text{ with } \widehat{f}(m) := f(\widehat{\alpha}'(m)). \tag{5.36}$$

In order to see that (ii) of Proposition 5.11 also holds when we consider $(C4_q)$ instead of $(C2_q)$, we adapt the argument in Proposition 5.16.

Proposition 5.19. *Assume that $k, n, p \in \mathbb{N}$ and $x \in C$ satisfy*

$$\forall i < \ell \left(\|U_i(x) - x\| \leq \frac{1}{9b(k+1)^2(p+1)}\right)$$

and

$$\forall r \in [n, p]\, \forall i < \ell \left(\langle x - u_0, U_i(x) - U_i(u_r)\rangle \leq \frac{1}{12(k+1)^2}\right).$$

Then

$$\forall m \in [\sigma''(k, n), p] \left(\|u_m - x\| \leq \frac{1}{k+1}\right),$$

where $\sigma''(k, n) := \max\{\theta(3b^2(\epsilon(\widetilde{n})+1)(k+1)^2 - 1), \widetilde{n}+1\}$ *with* $\widetilde{n} := \max\{n,\ \mu(6b^2(k+1)^2)\}$.

Proof. Following the arguments used in the proof of Proposition 5.16 we conclude that, for $m \in [\widetilde{n}+1, p]$,

$$\|u_m - x\|^2 \leq \frac{2}{3(k+1)^2} + B_m\|u_{\widetilde{n}} - x\|^2,$$

with $B_m = \prod_{j=\widetilde{n}+1}^{m}(1-\lambda_j)$.
For all $m \in \mathbb{N}$, let $P_m := \prod_{j=1}^{m}(1-\lambda_j)$. Then, for $m \in [\widetilde{n}+1, p]$,

$$\|u_m - x\|^2 \leq \frac{2}{3(k+1)^2} + b^2 \frac{P_m}{P_{\widetilde{n}}}. \tag{5.37}$$

Consider $m \in [\sigma''(k,n), p]$. Then, $m \geq \theta(3b^2(\epsilon(\widetilde{n})+1)(k+1)^2-1)$ and, by $(C4_q)$, we have

$$P_m \leq P_{\theta(3b^2(\epsilon(\widetilde{n})+1)(k+1)^2-1)} \leq \frac{1}{3b^2(\epsilon(\widetilde{n})+1)(k+1)^2}.$$

By the hypothesis on the function ϵ, $P_{\widetilde{n}} \geq \frac{1}{\epsilon(\widetilde{n})+1}$. It follows,

$$b^2 \frac{P_m}{P_{\widetilde{n}}} \leq \frac{b^2(\epsilon(\widetilde{n})+1)}{3b^2(\epsilon(\widetilde{n})+1)(k+1)^2} = \frac{1}{3(k+1)^2}. \tag{5.38}$$

Since $\sigma''(k,n) \geq \widetilde{n}+1$, we can consider (5.37) and (5.38) together, which concludes the proof. $\square$

We obtain the bounds satisfying condition (ii) of Proposition 5.11 by instantiating with $p = f(\sigma''(k,n))$,

$$\gamma(k,n,f) := 9b(k+1)^2(f(\sigma'(k,n))+1)-1, \quad \delta(k) := 12(k+1)^2-1,$$
$$\eta(k,n,f) := f(\sigma''(k,n)) \quad \text{and} \quad M := \sigma(k,n,f) := \sigma''(k,n).$$

Notice that the only change from before is in the functions η and σ, which now use the function σ'' from Proposition 5.19 instead of the function σ' from Proposition 5.16. By applying Proposition 5.11, we obtain a metastable version of Bauschke's theorem using the condition $(C4_q)$.

Theorem 5.20. *Under the assumptions of Theorem 5.12, assume that $\tau : \mathbb{N} \to \mathbb{N}$ is a monotone function satisfying* (5.21)*. Consider monotone functions $\mu, \xi, \theta : \mathbb{N} \to \mathbb{N}$ such that $(C1_q), (C3[\ell]_q)$ and $(C4_q)$ hold and $\epsilon : \mathbb{N} \to \mathbb{N}$ a monotone function satisfying* (5.34)*. Let $u_0 \in C$ be given and (u_n) be the iteration defined by* (5.17)*.*
Then, for all $k \in \mathbb{N}$ and every monotone function $f : \mathbb{N} \to \mathbb{N}$,

$$\exists N \leq \phi_{\mathrm{Bau'}}(k,f)\, \forall i,j \in [N, f(N)] \left(\|u_i - u_j\| \leq \frac{1}{k+1} \right), \tag{5.39}$$

where

$$\phi_{\mathrm{Bau'}}(k,f) := \sigma''\left(2k+1, \psi\left(48(k+1)^2-1, \overline{f}\right)\right),$$

with σ'' defined in Proposition 5.19, ψ defined by (5.36) *and the function $\overline{f}$ defined by $\overline{f}(m) = 36b(k+1)^2(f(\sigma''(2k+1,m))+1)-1$.*

When $\lambda_n = \frac{1}{n+1}$, we can take the functions μ, ξ, θ and ϵ to be the identity function and they will satisfy the required conditions $(C1_q)$, $(C3[\ell]_q)$, $(C4_q)$ and (5.34). In the particular case of Bauschke's theorem when $\ell = 1$ we recover Wittmann's theorem and can take τ to also be the identity function. Furthermore, when $\ell = 1$, in Proposition 5.15 we may have the function β of (ii) to instead be the same as β_0 from (i) – notice that β_0 is the same bound that was consider in the analysis of Wittmann's theorem. With these considerations, a quantitative version of Bauschke's theorem for the particular sequence $\lambda_n = \frac{1}{n+1}$ and a quantitative version of Wittmann's theorem using condition $(C4_q)$, for general sequence (λ_n) and for $\lambda_n = \frac{1}{n+1}$, are easily derived.

Chapter 6

Proximal point algorithm

Let X be a real Hilbert space and consider a multi-valued function $T : X \to 2^X$.
The operator T is said to be monotone if

$$\langle x - x', y - y' \rangle \geq 0, \tag{6.1}$$

for all $x, x' \in X$, $y \in T(x)$ and $y' \in T(x')$. A monotone operator T is said to be *maximal monotone* if additionally the graph of T,

$$\text{graph}(T) = \{(x, y) \in X \times X \mid x \in X, y \in T(x)\},$$

is not strictly contained in the graph of any monotone operator. Assume that T is maximal monotone and let $S := \{x \in X \mid 0 \in T(x)\}$ be the set of zeros of T. It is well-known that the set S is closed and convex, and we will henceforth assume it to be *non-empty*. One of the major problems in the theory of maximal operators is how to find a point $x \in S$. The relevance of this search for zeros derives from the fact that many problems in nonlinear analysis and optimization theory can be formulated as a question of finding a zero for specific maximal monotone operators.
For any positive real number $\beta > 0$, the function J_β defined by

$$J_\beta(x) := \{y \in X \mid x \in y + \beta T(y)\}$$

is called the *resolvent function* of βT. The resolvent functions associated with a maximal monotone operator are single-valued [40], in which case we just write $J_\beta(x) = y$, and nonexpansive mappings in X. Furthermore, an easy observation is that, for every $\beta > 0$, the set of fixed points of J_β coincide with the set of zeros of T,

$$\forall \beta > 0 \; (\text{Fix}(J_\beta) = S)\,.$$

An important tool in finding zeros of maximal monotone operators is the proximal point algorithm (PPA): given an initial guess x_0, a (regularization) sequence of positive real numbers

(β_n) and an error sequence $(e_n) \subset X$, the proximal point algorithm is inductively defined by

$$x_{n+1} := J_{\beta_n}(x_n) + e_n. \tag{PPA}$$

This definition, with an error term e_n, is sometimes called the *inexact* proximal point algorithm in contrast with the exact version when $e_n \equiv 0$.

In [42], Rockafellar showed that, if $(\|e_n\|)$ is a summable sequence and (β_n) is bounded away from 0, then the sequence (x_n) is weakly convergent to a zero of T. However, a strong convergence result is not obtained. In fact, Güler [17] presented a counterexample proving that in general (PPA) may fail to have strong convergence.
Several modifications to the (PPA) were considered in an attempt to guarantee a strong convergence result. Here, we will look at one such alternative iteration, the Halpern type proximal point algorithm. Motivated by the success of Halpern iterations for nonexpansive mappings, this algorithm was introduced independently by Kamimura and Takahashi in [20] and by Xu in [52].
Let $(\alpha_n) \subset]0,1[$ be a sequence of real numbers, $x_0 \in X$ be an initial guess, $u \in X$ an "anchor" point and (β_n) a sequence of positive real numbers. Then (x_n) is an exact Halpern type proximal point algorithm (with anchor point u and initial guess x_0) if it is defined inductively by

$$x_{n+1} := \alpha_n u + (1-\alpha_n) J_{\beta_n}(x_n). \tag{HPPA}$$

More generally, with $(e_n) \subset X$ an error sequence, we can also consider the inexact Halpern type proximal point algorithm:

$$x_{n+1} := \alpha_n u + (1-\alpha_n) J_{\beta_n}(x_n) + e_n \tag{HPPA$_1$}$$

$$x_{n+1} := \alpha_n u + (1-\alpha_n) \left(J_{\beta_n}(x_n) + e_n\right) \tag{HPPA$_2$}$$

Although (HPPA$_1$) is equivalent to (HPPA$_2$), by considering these two definitions, we can state the results with their original conditions. We will look at two strong convergence results of these iterations under certain conditions.

The relevant conditions on the parameters of the iterations are the following:

(C1) $\lim \alpha_n = 0$;

(C2) $\sum \alpha_n = \infty$;

(C3) $\lim \frac{|\alpha_{n+1}-\alpha_n|}{\alpha_n^2} = 0$;

(C4) $\lim \beta_n = \beta$, for some $\beta > 0$;

(C5) $\lim \beta_n = \infty$;

(C6) $\sum \|e_n\| < \infty$;

(C7) $\lim \frac{\|e_n\|}{\alpha_n} = 0$.

The first result, which we will give a quantitative version in section 6.3, is due to Boikanyo and Moroşanu.

Theorem 6.1 ([5, Theorem 2]). *Consider a sequence (x_n) given by* (HPPA$_1$) *and satisfying conditions* (C1)–(C4) *and* (C6) *or* (C7)*. Then (x_n) is strongly convergent to a zero of T, the closest one to u.*

A strong convergence of the Halpern type proximal point algorithm under different conditions was proved by Xu.

Theorem 6.2 ([52, Theorem 5.1]). *Consider a sequence (x_n) given by* (HPPA$_2$) *with $x_0 = u$ and satisfying conditions* (C1)*,* (C2)*,* (C5) *and* (C6)*. Then (x_n) is strongly convergent to a zero of T, the closest one to x_0.*

Both of these results use a projection argument and a sequential weak compactness argument. In the first result, sequential weak compactness is needed simply to establish the convergence of the Browder iteration which, as we saw, can be bypassed. In the analysis of the second result, it will also be possible to eliminate the need of the sequential weak compactness argument. This elimination can be seen as an application of the general principle in Chapter 4. Regarding the projection argument, we can again consider the weaker $\forall\exists$-version, avoiding the countable choice principle, and still prove the metastability of the iterations. In section 4.3, we explained how the bounded functional interpretation can deal with the weaker projection statement. However, there, a boundedness condition was essential in simplifying the translation of the statement. There is no obvious boundedness condition here. In section 6.2, under the assumption that $S \neq \emptyset$, we can prove that the iterations considered are bounded and carry out a similar quantitative analysis. The quantitative metastable version of the theorems above are obtained in sections 6.3 and 6.4, respectively.

We start with some useful technical lemmas regarding some properties on sequences of real numbers which correspond to the main combinatorial arguments of the theorems 6.1 and 6.2.

6.1 Technical Lemmas

In [53], it was shown that if $\sum \|e_n\|$ converges, then the sequence (x_n) given by (HPPA$_1$) is bounded. The same was shown to be true if instead we have a bound for the sequence $\left(\frac{\|e_n\|}{\alpha_n}\right)$, [4]. Next we show a quantitative version of those arguments, where a bound on (x_n) is computed from a bound on $(\sum \|e_n\|)$ or on $\left(\frac{\|e_n\|}{\alpha_n}\right)$.

Lemma 6.3. *Let p_0 be a zero of T. Given $(\alpha_n) \subset]0,1[$, $(\beta_n) \subset]0,+\infty[$, $(e_n) \subset X$ and $u, x_0 \in X$, consider a sequence (x_n) defined by* (HPPA$_1$)*. Then, for any $D \in \mathbb{N}$ we have,*

(a) $\forall n \in \mathbb{N} \left(\sum\limits_{i=0}^{n} \|e_i\| \leq D \right) \to \forall n \in \mathbb{N} \left(\|x_n - p_0\| \leq d_1(D) \right)$;

(b) $\forall n \in \mathbb{N} \left(\dfrac{\|e_n\|}{\alpha_n} \leq D \right) \to \forall n \in \mathbb{N} \left(\|x_n - p_0\| \leq d_2(D) \right)$,

with

$$\begin{aligned} d_1(D) &:= \lceil \max\{\|u - p_0\|, \|x_0 - p_0\|\} \rceil + D \\ \textit{and } d_2(D) &:= \lceil \max\{2(\|u - p_0\| + D), \|x_0 - p_0\|\} \rceil. \end{aligned}$$

Proof. Since $p_0 \in S$, p_0 is a fixed-point of J_{β_n} for all $n \in \mathbb{N}$.
For (a), see that

$$\begin{aligned} \|x_{n+1} - p_0\| &= \|\alpha_n u + (1-\alpha_n) J_{\beta_n}(x_n) + e_n - p_0\| = \\ &= \|\alpha_n(u - p_0) + (1-\alpha_n)(J_{\beta_n}(x_n) - J_{\beta_n}(p_0)) + e_n\| \\ &\leq \alpha_n \|u - p_0\| + (1-\alpha_n)\|x_n - p_0\| + \|e_n\| \end{aligned} \tag{6.2}$$

By induction on n, we see that

$$\forall n \in \mathbb{N} \left(\|x_n - p_0\| \leq \max\{\|u - p_0\|, \|x_0 - p_0\|\} + \sum_{i=0}^{n-1} \|e_i\| \right).$$

The base case $n = 0$ is trivial. The induction step $n+1$ follows from (6.2) and the induction hypothesis:

$$\begin{aligned} \|x_{n+1} - p_0\| &\leq \alpha_n \|u - p_0\| + (1-\alpha_n)\|x_n - p_0\| + \|e_n\| \leq \\ &\leq \alpha_n \|u - p_0\| + (1-\alpha_n) \left(\max\{\|u - p_0\|, \|x_0 - p_0\|\} + \sum_{i=0}^{n-1} \|e_i\| \right) + \|e_n\| \leq \\ &\leq \alpha_n \max\{\|u - p_0\|, \|x_0 - p_0\|\} + (1-\alpha_n) \left(\max\{\|u - p_0\|, \|x_0 - p_0\|\} \right) + \sum_{i=0}^{n} \|e_i\| = \\ &= \max\{\|u - p_0\|, \|x_0 - p_0\|\} + \sum_{i=0}^{n} \|e_i\|. \end{aligned}$$

The implication (a) then follows from the assumption on D and the definition of d_1.
For (b), with $M := \max\{\|u - p_0\| + D, \frac{\|x_0 - p_0\|}{2}\}$, we have $d_2(D) = 2M$.

The result is shown by induction on n. Again the base case is trivial. For the induction step, notice that, from $\|x+y\|^2 \leq \|x+y\|^2 + \|x\|^2$, one easily derives

$$\forall x, y \in X \ \left(\|x+y\|^2 \leq \|y\|^2 + 2\langle x, x+y\rangle\right).$$

Hence,

$$\begin{aligned}
\|x_{n+1} - p_0\| &= \|\alpha_n u + (1-\alpha_n)J_{\beta_n}(x_n) + e_n - p_0\|^2 = \\
&= \left\|\alpha_n\left(u - p_0 + \frac{e_n}{\alpha_n}\right) + (1-\alpha_n)\left(J_{\beta_n}(x_n) - J_{\beta_n}(p_0)\right)\right\|^2 \leq \\
&\leq (1-\alpha_n)^2\|J_{\beta_n}(x_n) - J_{\beta_n}(p_0)\|^2 + 2\alpha_n\langle u - p_0 + \frac{e_n}{\alpha_n}, x_{n+1} - p_0\rangle \leq \\
&\leq (1-\alpha_n)^2\|x_n - p_0\|^2 + 2\alpha_n\left(\|u - p_0\| + \frac{\|e_n\|}{\alpha_n}\right)\|x_{n+1} - p_0\| \leq \\
&\leq (1-\alpha_n)^2 d_2(D)^2 + 2M\alpha_n\|x_{n+1} - p_0\|.
\end{aligned}$$

From this we conclude,

$$(\|x_{n+1} - p_0\| - M\alpha_n)^2 \leq (1-\alpha_n)^2 d_2(D)^2 + M^2\alpha_n^2. \tag{6.3}$$

Thus $\|x_{n+1} - p_0\| \leq M\alpha_n + \sqrt{(1-\alpha_n)^2 d_2(D)^2 + M^2\alpha_n^2}$.

Since, for non-negative a, b it holds $\sqrt{a+b} \leq \sqrt{a} + \sqrt{b}$,

$\|x_{n+1} - p_0\| \leq M\alpha_n + \sqrt{(1-\alpha_n)^2 d_2(D)^2 + M\alpha_n^2} \leq M\alpha_n + (1-\alpha_n)d_2(D) + M\alpha_n = d_2(D)$.

□

Usually we consider stronger information on the sequences $(\sum \|e_i\|)$ and $(\frac{\|e_n\|}{\alpha_n})$, like rates of convergence or on the Cauchy property. Of course a trivial bound on the sequence can then be computed from the given information. Notice that, from the lemma above, we also derive a bound on the sequence (x_n):

$$\forall n \in \mathbb{N}\left(\sum_{i=0}^{n}\|e_i\| \leq D\right) \to \forall n \in \mathbb{N}\left(\|x_n\| \leq D_1\right)$$

$$\forall n \in \mathbb{N}\left(\frac{\|e_n\|}{\alpha_n} \leq D\right) \to \forall n \in \mathbb{N}\left(\|x_n\| \leq D_2\right),$$

with $D_1 := d_1(D) + \mathrm{P}$ and $D_2 := d_2(D) + \mathrm{P}$, where $\mathrm{P} \in \mathbb{N}$ is a bound on the norm of some zero of T.

For a quantitative analysis of the theorems 6.1 and 6.2, it is important to work with the conditions on the parameters in their quantitative form. The next conditions (Q1)–(Q7) are the quantitative versions of (C1)–(C7), respectively.

(Q1) $\mathrm{a} : \mathbb{N} \to \mathbb{N}$ is a rate of convergence towards zero for (α_n), i.e.

$$\forall k \in \mathbb{N} \forall n \geq \mathrm{a}(k) \left(\alpha_n \leq \frac{1}{k+1} \right);$$

(Q2) $\mathrm{A} : \mathbb{N} \to \mathbb{N}$ is a rate of divergence for $\sum \alpha_n$, i.e.

$$\forall k \in \mathbb{N} \left(\sum_{i=0}^{\mathrm{A}(k)} \alpha_i \geq k \right);$$

(Q3) $\mathrm{c} : \mathbb{N} \to \mathbb{N}$ is a a rate of convergence towards zero for $\left(\frac{|\alpha_{n+1} - \alpha_n|}{\alpha_n^2} \right)$, i.e.

$$\forall k \in \mathbb{N} \forall n \geq \mathrm{c}(k) \left(\frac{|\alpha_{n+1} - \alpha_n|}{\alpha_n^2} \leq \frac{1}{k+1} \right);$$

(Q4) $\mathrm{B} : \mathbb{N} \to \mathbb{N}$ that is a rate of convergence towards $\beta > 0$ for (β_n), i.e.

$$\forall k \in \mathbb{N} \forall n \geq \mathrm{B}(k) \left(|\beta_n - \beta| \leq \frac{1}{k+1} \right),$$

and $\ell \in \mathbb{N}$ is such that $\beta \geq \frac{1}{\ell+1}$;

(Q5) $\mathrm{B} : \mathbb{N} \to \mathbb{N}$ is a rate of divergence for (β_n), i.e.

$$\forall k \in \mathbb{N} \forall n \geq \mathrm{B}(k) \ (\beta_n \geq k);$$

(Q6) $\mathrm{E} : \mathbb{N} \to \mathbb{N}$ is a *Cauchy rate* for $\sum \|e_n\|$, i.e.,

$$\forall k \in \mathbb{N} \forall n \in \mathbb{N} \left(\sum_{i=\mathrm{E}(k)+1}^{\mathrm{E}(k)+n} \|e_i\| \leq \frac{1}{k+1} \right);$$

(Q7) $\mathrm{E} : \mathbb{N} \to \mathbb{N}$ is a rate of convergence towards zero for $\left(\frac{\|e_n\|}{\alpha_n} \right)$, i.e.

$$\forall k \in \mathbb{N} \forall n \geq \mathrm{E}(k) \left(\frac{\|e_n\|}{\alpha_n} \leq \frac{1}{k+1} \right).$$

A useful result regarding sequences of real numbers was proved by Xu in [52, Lemma 2.5].

Lemma 6.4. *Let (s_n) be a bounded sequence of non-negative real numbers and assume that for any $n \in \mathbb{N}$*

$$s_{n+1} \leq (1-\alpha_n)s_n + \alpha_n r_n + \gamma_n, \tag{6.4}$$

where $(\alpha_n) \subset]0,1[$, (r_n) and $(\gamma_n) \subset \mathbb{R}_0^+$ are sequences of real numbers such that (C2) *holds,* $\limsup r_n \leq 0$ *and* $\sum \gamma_n < \infty$. *Then* $\lim s_n = 0$.

The assumption of (s_n) being a bounded sequence can be dropped as it follows from (6.4) and the assumptions on the sequences (α_n), (r_n) and (γ_n). In fact, one can easily see by induction that, with $\mathcal{R}$ an upper bound on (r_n), for all $n \in \mathbb{N}$

$$s_n \leq \max\{s_0, \mathcal{R}\} + \sum_{i=0}^{n-1} \gamma_i,$$

and a bound on (s_n) can be effectively computed from quantitative information on the assumptions. Nevertheless, for simplicity we consider this universal assumption instead of working with a computed value.
The condition "$\limsup r_n \leq 0$" can be expressed by

$$\forall k \in \mathbb{N} \exists N \in \mathbb{N} \forall n \geq N \left(r_n \leq \frac{1}{k+1} \right),$$

and by asking for a monotone function $\mathrm{R} : \mathbb{N} \to \mathbb{N}$ such that, for each $k \in \mathbb{N}$, $\mathrm{R}(k)$ witnesses N above, we have a quantitative version of "$\limsup r_n \leq 0$". In these quantitative versions, we are asking for the existential information on $\forall\exists\forall$-formulas. However, more general results could be considered if instead we worked with bounds on metastable versions. From a practical point a view, in most cases we can consider this stronger quantitative information since, in the particular sequences that one is usually interested in, it is in fact possible to ascertain such witnessing rates.
Lemma 6.4 above contains the main combinatorial part of the proofs that we want to analyze and thus, we will need to give it a quantitative version. The particular case of this lemma when $\gamma_n \equiv 0$, was already given a quantitative version by Kohlenbach and Leuştean in [32]. The next results are quantitative versions of Lemma 6.4 for general sequences (γ_n).

Lemma 6.5 (Quant.version I of Lemma 6.4)**.** *Let (s_n) be a bounded sequence of non-negative real numbers and $D \in \mathbb{N}$ a positive upper bound on (s_n). Consider sequences of real numbers $(\alpha_n) \subset]0,1[$, (r_n) and $(\gamma_n) \subset \mathbb{R}_0^+$ and assume the existence of monotone functions* A, R, $\mathrm{G} : \mathbb{N} \to \mathbb{N}$ *such that* A *satisfies* (Q2), R *is such that*

$$\forall k \in \mathbb{N}\, \forall n \geq \mathrm{R}(k) \left(r_n \leq \frac{1}{k+1} \right),$$

and G *is a Cauchy rate for* $\sum \gamma_n$*, i.e.* $\forall k \in \mathbb{N}\, \forall n \in \mathbb{N} \left(\sum_{i=G(k)+1}^{G(k)+n} \gamma_i \leq \dfrac{1}{k+1} \right)$.

If for any $n \in \mathbb{N}$

$$s_{n+1} \leq (1-\alpha_n)s_n + \alpha_n r_n + \gamma_n,$$

then

$$\forall k \in \mathbb{N}\, \forall n \geq \theta_1[\mathrm{A}, \mathrm{R}, \mathrm{G}, D](k) \left(s_n \leq \frac{1}{k+1} \right),$$

with $\theta_1[\mathrm{A}, \mathrm{R}, \mathrm{G}, D](k) := \mathrm{A}\left(N + \lceil \ln(3D(k+1)) \rceil\right) + 1$,
where $N := \max\{\mathrm{R}(3k+2), \mathrm{G}(3k+2)+1\}$.

With the observation above regarding the boundedness of (s_n), we can have D replaced by $\max\{s_0, \mathcal{R}\} + \mathcal{E}$, where $\mathcal{R} := \max_{n \leq \mathrm{R}(0)}\{1, r_n\}$ is a bound on (r_n) and $\mathcal{E} := 1 + \lceil \sum_{i=0}^{\mathrm{G}(0)} \gamma_i \rceil$ is a bound on $(\sum^n \gamma_i)$.
The chosen notation "$\theta_1[\mathrm{A}, \mathrm{R}, \mathrm{G}, D]$" may not be very pleasing, however it will allow us to apply this result for particular choices of parameters A, R, G and D without risk of confusion.

Proof. Let $k \in \mathbb{N}$ be given and consider N as defined in the lemma. By induction, we show that, for all $m \in \mathbb{N}$

$$s_{N+m+1} \leq \left(\prod_{i=N}^{N+m} (1-\alpha_i) \right) s_N + \left(1 - \prod_{i=N}^{N+m} (1-\alpha_i) \right) \frac{1}{3(k+1)} + \sum_{i=N}^{N+m} \gamma_i. \tag{6.5}$$

From the assumption and the fact that $N \geq \mathrm{R}(3k+2)$, we can argue the base case $m = 0$,

$$\begin{aligned} s_{N+1} &\leq (1-\alpha_N)s_N + \alpha_N r_N + \gamma_N \leq \\ &\leq (1-\alpha_N)s_N + (1-(1-\alpha_N))\frac{1}{3(k+1)} + \gamma_N. \end{aligned}$$

For the induction step $m+1$, by using the assumption and the induction hypothesis, we get the following

$$\begin{aligned} s_{N+m+1+1} &\leq (1-\alpha_{N+m+1})s_{N+m+1} + \alpha_{N+m+1} r_{N+m+1} + \gamma_{N+m+1} \leq \\ &\leq (1-\alpha_{N+m+1}) \left[\left(\prod_{i=N}^{N+m} (1-\alpha_i) \right) s_N + \left(1 - \prod_{i=N}^{N+m} (1-\alpha_i) \right) \frac{1}{3(k+1)} + \sum_{i=N}^{N+m} \gamma_i \right] \\ &\quad + \alpha_{N+m+1} \frac{1}{3(k+1)} + \gamma_{N+m+1} \leq \\ &\leq \left(\prod_{i=N}^{N+m+1} (1-\alpha_i) \right) s_N + \left(1 - \prod_{i=N}^{N+m+1} (1-\alpha_i) \right) \frac{1}{3(k+1)} + \sum_{i=N}^{N+m+1} \gamma_i, \end{aligned}$$

where at the second inequality we used the induction hypothesis and the condition on R. This concludes the induction.

Since $N \geq G(3k+2)+1$,

$$\sum_{i=N}^{N+m} \gamma_i \leq \sum_{i=G(3k+2)+1}^{N+m} \gamma_i \leq \frac{1}{3(k+1)}.$$

Hence from (6.5), we get for all $m \in \mathbb{N}$

$$s_{N+m+1} \leq \left(\prod_{i=N}^{N+m} (1-\alpha_i) \right) D + \frac{2}{3(k+1)}. \tag{6.6}$$

Since for all $n \in \mathbb{N}$, $\alpha_n \leq 1$, we have for all $j \in \mathbb{N}$, $A(j+1) \geq j$. Also, as $D \geq 1$, we have $\lceil \ln(3D(k+1)) \rceil \geq 1$. From these two facts, we can see that the expression $M := A\left(N + \lceil \ln(3D(k+1)) \rceil\right) - N$ gives a natural number. We have, for any $m \geq M$,

$$\begin{aligned} \sum_{i=0}^{N+m} \alpha_i \geq \sum_{i=0}^{N+M} \alpha_i = \sum_{i=0}^{A(N+\lceil \ln(3D(k+1)) \rceil)} \alpha_i \geq \\ \geq N + \ln(3D(k+1)) \geq \sum_{i=0}^{N-1} \alpha_i + \ln(3D(k+1)), \end{aligned}$$

from which we conclude $\sum_{i=N}^{N+m} \alpha_i \geq \ln(3D(k+1))$.

Recall that for non-negative x, $1-x \leq \exp(-x)$. Thus, we obtain for all $m \geq M$,

$$\left(\prod_{i=N}^{N+m} (1-\alpha_i) \right) D \leq \exp\left(-\sum_{i=N}^{N+m} \alpha_i \right) D \leq \frac{D}{3D(k+1)} = \frac{1}{3(k+1)}. \tag{6.7}$$

From (6.6) and (6.7) we get, for all $n \geq N+M+1 = \theta_1[\mathrm{A}, \mathrm{R}, \mathrm{G}, D](k)$,

$$s_n \leq \frac{1}{k+1}.$$

□

Consider the condition

$$(\mathrm{C2}') \quad \prod (1-\alpha_n) = 0.$$

It is well-known that, for $(\alpha_n) \subset]0,1[$, the condition (C2) is equivalent to this new condition (C2'). Hence, it makes sense to consider a quantitative version of (C2'):

$$(\mathrm{Q2}') \quad \mathrm{A} : \mathbb{N} \to \mathbb{N} \text{ is a rate of convergence towards zero for } \prod(1-\alpha_n), \text{ i.e.}$$

$$\forall k \in \mathbb{N} \forall n \in \mathbb{N} \left(\prod_{i=0}^{\mathrm{A}(k)+n} (1-\alpha_i) \leq \frac{1}{k+1} \right).$$

We can also prove the previous Lemma 6.5 with condition (Q2) replaced by condition (Q2'). The usefulness of this result is in the fact that, for particular sequences (α_n), a function satisfying (Q2) may have different complexity than a function satisfying (Q2') and so we can choose to work with the simpler one in order to obtain better bounds. We already saw an example of this at the end of Chapter 5, where rate of convergence towards zero for $\prod(1-\frac{1}{n+2})$ is linear, while we only have an exponential rate of divergence for $\sum \frac{1}{n+2}$.

Lemma 6.6 (Quant.version II of Lemma 6.4). *Let (s_n) be a bounded sequence of non-negative real numbers and $D \in \mathbb{N}$ a positive upper bound on (s_n). Consider sequences of real numbers $(\alpha_n) \subset]0,1[$, (r_n) and $(\gamma_n) \subset \mathbb{R}_0^+$ and assume the existence of monotone functions* A, R, $\mathrm{G} : \mathbb{N} \to \mathbb{N}$ *such that* A *satisfies* (Q2'), R *is such that*

$$\forall k \in \mathbb{N}\, \forall n \geq \mathrm{R}(k) \left(r_n \leq \frac{1}{k+1} \right),$$

and G *is a Cauchy rate for $\sum \gamma_n$.*
If for any $n \in \mathbb{N}$

$$s_{n+1} \leq (1-\alpha_n)s_n + \alpha_n r_n + \gamma_n,$$

then

$$\forall k \in \mathbb{N}\, \forall n \geq \theta_2[\mathrm{A}, \mathrm{R}, \mathrm{G}, D, \ell](k) \left(s_n \leq \frac{1}{k+1} \right),$$

with $\theta_2[\mathrm{A}, \mathrm{R}, \mathrm{G}, D, \ell](k) := \max\{\mathrm{A}(3D(k+1)(\ell+1)-1),\, N\}+1$,
where $N := \max\{\mathrm{R}(3k+2), \mathrm{G}(3k+2)+1\}$ and $\ell \in \mathbb{N}$ is such that $\prod_{i=0}^{N-1}(1-\alpha_i) \geq \frac{1}{\ell+1}$.

Proof. Consider $k \in \mathbb{N}$ given and N as defined in the lemma. Notice that $N \geq 1$. Denote $P_n := \prod_{i=0}^{n}(1-\alpha_i)$. As in the proof of the previous lemma we conclude, for any $m \in \mathbb{N}$

$$s_{N+m+1} \leq \left(\prod_{i=N}^{N+m} (1-\alpha_i) \right) s_N + \frac{2}{3(k+1)} \leq D\frac{P_{N+m}}{P_{N-1}} + \frac{2}{3(k+1)}. \tag{6.8}$$

Define the natural number $M := \max\{\mathrm{A}\left(3D(k+1)(\ell+1)-1\right), N\} - N$. By the condition on A, and since $(1-\alpha_i) \leq 1$, for any $m \geq M$,

$$P_{N+m} = \prod_{i=0}^{N+m}(1-\alpha_i) \leq \prod_{i=0}^{N+M}(1-\alpha_i) \leq \frac{1}{3D(k+1)(\ell+1)}.$$

On the other hand, by the definition of ℓ, $P_{N-1} \geq \frac{1}{\ell+1}$.
Thus,

$$D\frac{P_{N+m}}{P_{N-1}} \leq \frac{(\ell+1)D}{3D(k+1)(\ell+1)} = \frac{1}{3(k+1)}. \tag{6.9}$$

From (6.8) and (6.9), we conclude, for any $n \geq N+M+1 = \theta_2[\mathrm{A},\mathrm{R},\mathrm{G},D,\ell](k)$,

$$s_n \leq \frac{1}{k+1}.$$

$\square$

Sometimes the sequence (s_n) is defined using an ideal element whose existence cannot be guaranteed in our restrictive formal setting, e.g. using some projection point. Instead we work with approximations of that ideal object. However, in that case, the inequality (6.4) may fail and we have a weaker version that at the right-hand side in place of s_n only has $s_n + v_n$, for (v_n) some possible sequence of errors. With arguments similar to those of Proposition 5.16, we prove the next result.

Lemma 6.7. *Let (s_n) be a bounded sequence of non-negative real numbers and $D \in \mathbb{N}$ a positive upper bound on (s_n). Consider sequences of real numbers $(\alpha_n) \subset]0,1[$, (r_n), (v_n) and $(\gamma_n) \subset \mathbb{R}_0^+$ and assume the existence of a monotone function A satisfying (Q2). For natural numbers k, N and p assume*

$$\forall n \in [N,p]\left(v_n \leq \frac{1}{4(k+1)(p+1)} \land r_n \leq \frac{1}{4(k+1)}\right),$$

$$\forall n \in \mathbb{N}\left(\sum_{i=N}^{N+n}\gamma_i \leq \frac{1}{4(k+1)}\right)$$

and for all $n \in \mathbb{N}$,

$$s_{n+1} \leq (1-\alpha_n)(s_n+v_n) + \alpha_n r_n + \gamma_n.$$

Then,

$$\forall n \in [\sigma_1(k,N),p]\left(s_n \leq \frac{1}{k+1}\right),$$

with $\sigma_1(k,N) := \mathrm{A}\left(N + \lceil \ln(4D(k+1)) \rceil\right) + 1$.

Proof. Consider k, N and p such that the premises of the lemma hold. We may assume $p \geq \mathrm{A}\left(N + \lceil \ln(4D(k+1)) \rceil\right) + 1$, otherwise the result is trivially true. Then $p \geq N$. By induction, we see that for all $m \leq p - N$,

$$s_{N+m+1} \leq \left(\prod_{i=N}^{N+m}(1-\alpha_i)\right) s_N + \frac{1}{4(k+1)(p+1)} \sum_{j=N}^{N+m} \prod_{i=j}^{N+m} (1-\alpha_i) + \frac{1}{4(k+1)} + \sum_{i=N}^{N+m} \gamma_i. \tag{6.10}$$

The base case $m = 0$ follows from the assumptions of the lemma. For the induction step $m + 1 \leq p - N$, we have

$$\begin{aligned}
&s_{N+m+2} \leq (1-\alpha_{N+m+1})(s_{N+m+1} + v_{N+m+1}) + \alpha_{N+m+1} r_{N+m+1} + \gamma_{N+m+1} \leq \\
&\leq (1-\alpha_{N+m+1}) \left[\left(\prod_{i=N}^{N+m}(1-\alpha_i)\right) s_N + \frac{1}{4(k+1)(p+1)} \sum_{j=N}^{N+m} \prod_{i=j}^{N+m} (1-\alpha_i) + \frac{1}{4(k+1)} \right. \\
&\qquad \left. + \sum_{i=N}^{N+m} \gamma_i \right] + (1-\alpha_{N+m+1}) v_{N+m+1} + \alpha_{N+m+1} \frac{1}{4(k+1)} + \gamma_{N+m+1} \leq \\
&\leq \left(\prod_{i=N}^{N+m+1}(1-\alpha_i)\right) s_N + \frac{1}{4(k+1)(p+1)} \sum_{j=N}^{N+m+1} \prod_{i=j}^{N+m+1} (1-\alpha_i) + \frac{1}{4(k+1)} + \sum_{i=N}^{N+m+1} \gamma_i,
\end{aligned}$$

using the induction hypothesis and the fact that, since $N+m+1 \in [N, p]$, $r_{N+m+1} \leq \frac{1}{4(k+1)}$. This concludes the induction.

For $m \leq p - N$, we have

$$\sum_{j=N}^{N+m} \prod_{i=j}^{N+m} (1-\alpha_i) \leq m + 1 \leq p + 1,$$

hence,

$$\frac{1}{4(k+1)(p+1)} \sum_{j=N}^{N+m} \prod_{i=j}^{N+m} (1-\alpha_i) \leq \frac{1}{4(k+1)}.$$

Since $\sum_{i=N}^{N+m} \gamma_i \leq \frac{1}{4(k+1)}$, by (6.10), we conclude for all $m \leq p - N$

$$s_{N+m+1} \leq \left(\prod_{i=N}^{N+m}(1-\alpha_i)\right) D + \frac{3}{4(k+1)}. \tag{6.11}$$

Define the natural number $M := \mathrm{A}\left(N + \lceil \ln(4D(k+1)) \rceil\right) - N$ and conclude with the same arguments as in the proof of Lemma 6.5 that for $m \geq M$,

$$D \prod_{i=N}^{N+m} (1-\alpha_i) \leq \frac{1}{4(k+1)}. \tag{6.12}$$

Finally, with (6.11) and (6.12), we obtain for $m \in [M, p-N]$

$$s_{N+m+1} \leq \frac{1}{k+1}$$

and thus, for $n \in [N+M+1, p] = [\sigma_1(k,N), p]$, we have $s_n \leq \frac{1}{k+1}$. $\square$

We can again change the final steps in the previous proof to conclude a similar result but when the function A satisfies instead (Q2′).

Lemma 6.8. *Let (s_n) be a bounded sequence of non-negative real numbers and $D \in \mathbb{N}$ a positive upper bound on (s_n). Consider sequences of real numbers $(\alpha_n) \subset]0,1[$, (r_n), (v_n) and $(\gamma_n) \subset \mathbb{R}_0^+$ and assume the existence of a monotone function* A *satisfying* (Q2′). *Let $h : \mathbb{N} \to \mathbb{N}^*$ be a monotone function satisfying $\prod_{i=0}^{n}(1-\alpha_i) \geq \frac{1}{h(n)}$, for all $n \in \mathbb{N}$. For natural numbers k, N and p assume*

$$\forall n \in [N,p] \left(v_n \leq \frac{1}{4(k+1)(p+1)} \wedge r_n \leq \frac{1}{4(k+1)} \right),$$

$$\forall n \in \mathbb{N} \left(\sum_{i=N}^{N+n} \gamma_i \leq \frac{1}{4(k+1)} \right)$$

and for all $n \in \mathbb{N}$,

$$s_{n+1} \leq (1-\alpha_n)(s_n + v_n) + \alpha_n r_n + \gamma_n.$$

Then,

$$\forall n \in [\sigma_2(k,N), p] \left(s_n \leq \frac{1}{k+1} \right),$$

with $\sigma_2(k,N) := \max\{\mathrm{A}\,(4D(k+1)\hbar(N-1)-1)+1, N\}+1$, with $\hbar(-1) := 1$ and for $n \in \mathbb{N}$, $\hbar(n) := h(n)$.

Proof. Denote $P_n := \prod_{i=0}^{n}(1-\alpha_i)$. Following the proof of the previous lemma we conclude that for all $m \leq p-N$,

$$s_{N+m+1} \leq \left(\prod_{i=N}^{N+m}(1-\alpha_i) \right) D + \frac{3}{4(k+1)} = D\frac{P_{N+m}}{P_{N-1}} + \frac{3}{4(k+1)}$$

Define the natural number $M := \max\{\mathrm{A}\,(4D(k+1)\hbar(N-1)-1)+1, N\} - N$.
On one hand, by (Q2′) and since $1-\alpha_i \leq 1$, we have for all $m \geq M$,

$$P_{N+m} = \prod_{i=0}^{N+m}(1-\alpha_i) \leq \prod_{i=0}^{N+M}(1-\alpha_i) \leq \frac{1}{4D(k+1)\hbar(N-1)}.$$

On the other hand, by the definition of h, $\hbar$ and with the usual convention $P_{-1} = 1$, we have $P_{N-1} \geq \dfrac{1}{\hbar(N-1)}$. Hence,

$$D\frac{P_{N+m}}{P_{N-1}} \leq \frac{D\hbar(N-1)}{4D(k+1)\hbar(N-1)} = \frac{1}{4(k+1)}.$$

This shows that for $n \in [N+M+1, p] = [\sigma_2(k,N), p]$, we have $s_n \leq \frac{1}{k+1}$. $\square$

The next two lemmas are specially tailored to the quantitative analysis of theorem 6.2 in section 6.4.

Lemma 6.9. *Let Ω be a bounded subset of X. Let $(\alpha_n) \subset]0,1[$ be a given sequence. For each $x \in \Omega$, consider the sequences of real numbers $(s_{n,x})$, $(v_{n,x})$, $(r_{n,x})$ and $(\gamma_{n,x})$ with $(s_{n,x})$, $(\gamma_{n,x}) \subset \mathbb{R}_0^+$ and such that, for all $x \in \Omega$,*

$$\forall n \in \mathbb{N} \ (s_{n+1,x} \leq (1-\alpha_i)(s_{n,x} + v_{n,x}) + \alpha_n r_{n,x} + \gamma_{n,x}).$$

For a natural number $D \in \mathbb{N}^*$ and monotone functions $\mathrm{A}, \mathrm{G} : \mathbb{N} \to \mathbb{N}$ and $\psi : \mathbb{N} \times \mathbb{N}^{\mathbb{N}} \to \mathbb{N}$, suppose:

(1) A *satisfies condition* (Q2)*;*

(2) *For all $x \in \Omega$, D is a positive upper bound on $(s_{n,x})$;*

(3) *For all $x \in \Omega$, G is a Cauchy rate on $\sum \gamma_{n,x}$;*

(4) $\forall k \in \mathbb{N} \widetilde{\forall} f : \mathbb{N} \to \mathbb{N} \exists x \in \Omega \, \exists N \leq \psi(k,f) \forall n \in [N, fN] \ \left(v_{n,x} \leq \frac{1}{f(N)+1} \wedge r_{n,x} \leq \frac{1}{k+1} \right)$.

Then, for any natural number k and monotone function $f : \mathbb{N} \to \mathbb{N}$ there are $x \in \Omega$ and $N \leq \Theta_1[A, \psi, G, D](k,f)$ such that

$$\forall n \in [N, fN] \ \left(s_{n,x} \leq \frac{1}{k+1} \right),$$

where $\Theta_1[A, \psi, G, D](k,f) := \mathrm{A}\left(N_0(\psi(4k+3, g))\right) + 1$ with
$N_0(m) = \max\{m, \mathrm{G}(4k+3)+1\} + \lceil \ln(4D(k+1)) \rceil$ and $g(m) := 4(k+1)\left(f(\mathrm{A}(N_0(m)) + 1) + 1\right)$.

Proof. Let $k \in \mathbb{N}$ and a monotone function f be given.
By (4), consider $\tilde{x} \in \Omega$ and $N_1 \leq \psi(4k+3, g)$ such that for $n \in [N_1, g(N_1)]$

$$v_{n,\tilde{x}} \leq \frac{1}{g(N_1)+1} \ \text{ and } \ r_{n,\tilde{x}} \leq \frac{1}{4(k+1)}.$$

Define $N_2 := \max\{N_1, \mathrm{G}(4k+3)+1\}$. By (3), for all $n \in \mathbb{N}$, $\sum_{i=N_2}^{N_2+n} \gamma_{n,\tilde{x}} \leq \frac{1}{4(k+1)}$.
We have $N_1 \leq N_2$ and

$$g(N_1) \geq f(\mathrm{A}(N_0(N_1))+1) = f(\sigma_1(k, N_2)),$$

where σ_1 is as in Lemma 6.7. Hence, for $n \in [N_2, f(\sigma_1(k, N_2))]$,

$$v_{n,\tilde{x}} \leq \frac{1}{4(k+1)(f(\sigma_1(k,N_2))+1)} \wedge r_{n,\tilde{x}} \leq \frac{1}{4(k+1)}.$$

We are in the conditions of the Lemma 6.7 with $N = N_2$ and $p = f(\sigma_1(k, N_2))$, and so

$$\forall n \in [\sigma_1(k,N_2), f(\sigma_1(k,N_2))] \left(s_{n,\tilde{x}} \leq \frac{1}{k+1} \right).$$

Noticing that, by the monotonicity of A, we have $\sigma_1(k, N_2) \leq \Theta_1[A, \psi, G, D](k, f)$, we conclude the proof. □

Once again, we have a similar result with the condition (Q2′) instead.

Lemma 6.10. *Let Ω be a bounded subset of X. Let $(\alpha_n) \subset]0,1[$ be a given sequence and $h : \mathbb{N} \to \mathbb{N}^*$ be a monotone function satisfying $\prod_{i=0}^{n}(1-\alpha_i) \geq \frac{1}{h(n)}$, for all $n \in \mathbb{N}$. For each $x \in \Omega$, consider the sequences of real numbers $(s_{n,x})$, $(v_{n,x})$, $(r_{n,x})$ and $(\gamma_{n,x})$ with $(s_{n,x})$, $(\gamma_{n,x}) \subset \mathbb{R}_0^+$ and such that, for all $x \in \Omega$,*

$$\forall n \in \mathbb{N}\ (s_{n+1,x} \leq (1-\alpha_i)(s_{n,x}+v_{n,x}) + \alpha_n r_{n,x} + \gamma_{n,x}).$$

For a natural number $D \in \mathbb{N}^*$ and monotone functions A, h, G $: \mathbb{N} \to \mathbb{N}$ *and* $\psi : \mathbb{N} \times \mathbb{N}^{\mathbb{N}} \to \mathbb{N}$, *suppose:*

1. A *satisfies condition* (Q2′)*;*
2. *For all $x \in \Omega$, D is a positive upper bound on $(s_{n,x})$;*
3. *For all $x \in \Omega$,* G *is a Cauchy rate on $\sum \gamma_{n,x}$;*
4. $\forall k \in \mathbb{N} \widetilde{\forall} f : \mathbb{N} \to \mathbb{N} \exists x \in \Omega\, \exists N \leq \psi(k,f) \forall n \in [N, fN] \left(v_{n,x} \leq \frac{1}{f(N)+1} \wedge r_{n,x} \leq \frac{1}{k+1} \right)$.

Then, for any natural number k and monotone function $f : \mathbb{N} \to \mathbb{N}$ there are $x \in \Omega$ and $N \leq \Theta_2[A, \psi, G, D, h](k, f)$ such that

$$\forall n \in [N, fN] \left(s_{n,x} \leq \frac{1}{k+1} \right),$$

where $\Theta_2[A, \psi, G, D, h](k, f) := \max\{\mathrm{A}\left(4D(k+1)\hbar(M-1)-1\right)+1, M\}$
with $g(m) := 4(k+1)\left(f(\max\{\mathrm{A}\left(4D(k+1)\hbar(M-1)-1\right)+1, N_0(m)\})+1\right)$,
$N_0(m) := \max\{m, \mathrm{G}(4k+3)+1\}$, $M := N_0(\psi(4k+3, g))$ *and* $\hbar$ *as in Lemma 6.8.*

Proof. The proof is similar to before but uses instead Lemma 6.8 and the function σ_2. One has to additionally note that

$$\prod_{i=0}^{N_2-1}(1-\alpha_i) \geq \prod_{i=0}^{M-1}(1-\alpha_i) \geq \frac{1}{\hbar(M-1)}.$$

□

We finish this section with the well-known identity for resolvent functions.

Lemma 6.11 (Resolvent Identity). *For any $x \in X$ and $\gamma, \beta > 0$,*

$$J_\beta(x) = J_\gamma\left(\frac{\gamma}{\beta}x + (1-\frac{\gamma}{\beta})J_\beta(x)\right). \tag{6.13}$$

6.2 A review of the projection argument

In Section 4.3, we looked at the quantitative analysis of the projection argument onto the set $F = \{x \in C \,|\, U(x) = x\}$, where C is a given nonempty bounded subset of a normed space and U is a nonexpansive mapping of C into itself. Using Kohlenbach's observation that the weaker statement (4.35) is enough to carry out the proof of a metastability property, we proceed in using the characteristic principles of the bounded functional interpretation to obtain its quantitative form. In the interpretation it was essential that the set C was bounded as it ensured that quantifications over C could be treated as bounded quantifications. However, sometimes the condition that C is a nonempty bounded set is replaced by the hypothesis that the set F is nonempty, without any condition of boundedness.
In this section, we will see how the projection argument onto F can be analyzed under the assumption that F is nonempty and without asking for the set C to be bounded. The crucial observation is that, when looking for a fixed point closer to some point, say, $v_0 \in C$, it suffices to work inside a ball that already encompass at least one fixed point. That fixed point may not be the closest fixed point to v_0, but any fixed point even further away clearly isn't and so, fixed points outside a certain radius can be dropped from the argument. We now formalize this statement.

Let X be a normed space and C a (possible unbounded) subset of X. Consider $v_0 \in C$ and p_0 some point in F, the set of fixed points of U in C. Consider a natural number $\widetilde{b} \geq \|p_0 - v_0\| + \|v_0\| + 1$ and denote $\mathcal{B}_\trianglelefteq$ to be the ball with (intensional) radius $\widetilde{b}$,

$$\mathcal{B}_\trianglelefteq := B_\trianglelefteq(\widetilde{b}) := \left\{x \in X \,|\, \|x\| \trianglelefteq_\mathbb{R} \widetilde{b}\right\}.$$

Consider

$$\forall k \in \mathbb{N}\, \exists x \in C \cap \mathcal{B}_{\trianglelefteq} \\ \left(U(x) = x \wedge \forall y \in C \cap \mathcal{B}_{\trianglelefteq} \left(U(y) = y \to \|x - v_0\| \trianglelefteq_{\mathbb{R}} \|y - v_0\| + \frac{1}{k+1}\right)\right), \tag{6.14}$$

which is the (weaker) projection of v_0 over $F \cap \mathcal{B}_{\trianglelefteq}$ and, since the set $F \cap \mathcal{B}_{\trianglelefteq}$ is nonempty (we have $\|p_0\| \trianglelefteq_{\mathbb{R}} \widetilde{b}$), (6.14) is clearly true.
We can replace the weaker projection statement over F with this one over $F \cap \mathcal{B}_{\trianglelefteq}$. To see this, given $k \in \mathbb{N}$, assume there is $p \in C$ such that

$$U(p) = p \wedge \forall y \in C \left(U(y) = y \to \|p - v_0\| \trianglelefteq_{\mathbb{R}} \|y - v_0\| + \frac{1}{k+2}\right).$$

Since $\frac{1}{k+2} \trianglelefteq_{\mathbb{R}} \frac{1}{k+1}$, we just have to see that $p \in \mathcal{B}_{\trianglelefteq}$. Since $U(p_0) = p_0$, we have

$$\|p\| \leq \|p - v_0\| + \|u\| \trianglelefteq_{\mathbb{R}} \|p_0 - v_0\| + \frac{1}{k+2} + \|v_0\| < \widetilde{b},$$

and it follows $\|p\| \trianglelefteq_{\mathbb{R}} \widetilde{b}$.

In the other direction, for an arbitrary $k \in \mathbb{N}$, assume we have $p \in C \cap \mathcal{B}_{\trianglelefteq}$ such that

$$U(p) = p \wedge \forall y \in C \cap \mathcal{B}_{\trianglelefteq} \left(U(y) = y \to \|p - v_0\| \trianglelefteq_{\mathbb{R}} \|y - v_0\| + \frac{1}{k+1}\right)$$

We just have to see that the second conjunct still holds outside $\mathcal{B}_{\trianglelefteq}$. Take $y \in F \setminus \mathcal{B}_{\trianglelefteq}$. Then $\|y\| \geq \widetilde{b}$. Since $p_0 \in F \cap \mathcal{B}_{\trianglelefteq}$ we have

$$\begin{aligned} \|p - v_0\| \leq \|p_0 - v_0\| + \frac{1}{k+1} &= \|p_0 - v_0\| + \|v_0\| - \|v_0\| + \frac{1}{k+1} < \\ < \widetilde{b} - \|v_0\| + \frac{1}{k+1} &\leq \|y\| - \|v_0\| + \frac{1}{k+1} \leq \|y - v_0\| + \frac{1}{k+1}. \end{aligned}$$

Hence $\|p - v_0\| < \|y - v_0\| + \frac{1}{k+1}$, which implies $\|p - v_0\| \trianglelefteq_{\mathbb{R}} \|y - v_0\| + \frac{1}{k+1}$.
Thus, the weak projection statement over F (4.35) is equivalent to the weaker projection statement over $F \cap \mathcal{B}$ (6.14). In the case where we don't have a boundedness condition on the set C, but instead the hypothesis that F is nonempty, we can consider the projection argument restricted to the ball $\mathcal{B}_{\trianglelefteq}$ and its interpretation will be the same as in Section 4.3.
We can go back to the extensional inequalities by flattening. For any $n \in \mathbb{N}$, write

$$B_{\leq}(n) := \{x \in X \mid \|x\| \leq_{\mathbb{R}} n\}.$$

We have the following quantitative result of the projection argument.

Proposition 6.12. *Let X be a normed space, U a nonexpansive mapping and C a subset of X which maps C into itself. Let $v_0 \in C$ and $\widetilde{b} \in \mathbb{N}$ such that $\widetilde{b} \geq \|p_0 - v_0\| + \|v_0\| + 1$ for p_0 some fixed point of U in C. We abbreviate $\mathcal{B}_{\leq} := B_{\leq}(\widetilde{b})$ and $b := 2\widetilde{b}$.*
For any natural number k and monotone function $f : \mathbb{N} \to \mathbb{N}$, there are $N \leq f^{(r)}(0)$ and $x \in C \cap \mathcal{B}_{\leq}$ such that

$$\|U(x) - x\| \leq \frac{1}{f(N) + 1}$$

and

$$\forall y \in C \cap \mathcal{B}_{\leq} \left(\|U(y) - y\| \leq \frac{1}{N+1} \to \|x - v_0\|^2 \leq \|y - v_0\|^2 + \frac{1}{k+1} \right),$$

where $r := b^2(k+1)$ and $f^{(r)}$ is the r-th fold composition of f.

Proof. The proof is essentially the same as in Proposition 4.15 with the observations that $b = 2\widetilde{b}$ is a bound on the diameter of $C \cap \mathcal{B}_{\leq}$ and that for the definition of x_0 we can take a fixed point p_0 guaranteed to exist by the definition of $\widetilde{b}$. □

After carrying out the interpretation and the extraction, we can extend the result to the original set C:

Proposition 6.13. *Let X be a normed space, U a nonexpansive mapping and C a subset of X which maps C into itself. Let $v_0 \in C$ and $\widetilde{b} \in \mathbb{N}$ such that $\widetilde{b} \geq \|p_0 - v_0\| + \|v_0\| + 1$ for p_0 some fixed point of U in C. Write $b := 2\widetilde{b}$.*
For any natural number k and monotone function $f : \mathbb{N} \to \mathbb{N}$, there are $N \leq f^{(r)}(0)$ and $x \in C$ such that

$$\|U(x) - x\| \leq \frac{1}{f(N)+1} \wedge \forall y \in C \left(\|U(y) - y\| \leq \frac{1}{N+1} \to \|x - v_0\|^2 \leq \|y - v_0\|^2 + \frac{1}{k+1} \right),$$

where $r := b^2(k+1)$ and $f^{(r)}$ is the r-th fold composition of f.

Proof. Given $k \in \mathbb{N}$ and a monotone function f, apply Proposition 6.12. We just have to see that the second conjunct holds outside $\mathcal{B}_{\leq}$.
Let $y \in C \setminus \mathcal{B}_{\leq}$ be such that $\|U(y) - y\| \leq \frac{1}{N+1}$.
Since $p_0 \in F \cap \mathcal{B}_{\leq}$ and $0 < \widetilde{b} - \|v_0\| \leq \|y\| - \|v_0\| \leq \|y - v_0\|$, we have

$$\begin{aligned} \|x - v_0\|^2 &\leq \|p_0 - v_0\|^2 + \frac{1}{k+1} \leq (\|p_0 - v_0\| + \|v_0\| - \|v_0\|)^2 + \frac{1}{k+1} \leq \\ &\leq (\widetilde{b} - \|v_0\|)^2 + \frac{1}{k+1} \leq \|y - v_0\|^2 + \frac{1}{k+1}. \end{aligned}$$

□

Let X now be a Hilbert space and assume that the set C is convex. Lemma 4.17 can be adapted in order to deal with an unbounded set C. Formally, the idea is to consider the argument restricted to $C \cap B_{\trianglelefteq}(n)$ for an arbitrary $n \in \mathbb{N}$, which justifies the interpretation. By flattening, we return to the extensional inequalities – and to $B_{\leq}(n)$ – and, since $diam(C \cap B_{\leq}(n)) \leq 2n$, we have the following result.

Lemma 6.14. *For all $k, n \in \mathbb{N}$ and $x_1, x_2 \in C \cap B_{\leq}(n)$,*

$$\bigwedge_{j=1}^{2}\left(\|U(x_j) - x_j\| \leq \frac{1}{12(2n)(k+1)^2}\right) \to \forall \gamma \in [0,1]\left(\|U(w_\gamma(x_1,x_2)) - w_\gamma(x_1,x_2)\| \leq \frac{1}{k+1}\right).$$

We make a small comment before procceding. It is essential in the proof of Lemmas 4.17 and 6.14 that the space have a *modulus of uniform convexity*, η:

$$\forall x, y \in B_{\leq}(1)\, \forall \varepsilon > 0 \left(\left\|\frac{x+y}{2}\right\| > 1 - \eta(\varepsilon) \to \|x - y\| < \varepsilon \right).$$

When X is a Hilbert space, Kohlenbach showed that we can consider $\eta(\varepsilon) = \frac{\varepsilon^2}{8}$. One could work with an arbitrary uniformly convex space and derived a corresponding Lemmas 4.17 and 6.14, just as long as a modulus of uniform convexity η is given. In that case, the extracted bound would additionally depend on the function η. For a detailed proof of these lemmas, and the relation to uniform convexity, see [27, Section 2]. We also remark that, in the case of uniformity convexity, it does not looks to be possible to work only with a version where $\frac{1}{k+1}$ replaces ε. Namely, in [27, Lemma 2.2], the proof would fail since the role of K cannot be appropriately captured by a natural number.

Similarly, we have a version of Lemma 4.18 adapted in order to deal with unbounded C.

Lemma 6.15. *For all $k, n \in \mathbb{N}$ and $x, y \in C \cap B_{\leq}(n)$,*

$$\forall \gamma \in [0,1]\left(\|x - v_0\|^2 \leq \|w_\gamma(x,y) - v_0\|^2 + \frac{1}{(2n)^2(k+1)^2}\right) \to \langle x - v_0, x - y\rangle \leq \frac{1}{k+1}.$$

Using these adapted lemmas and Proposition 6.12 we obtain:

Proposition 6.16. *Let X be a Hilbert space, U a nonexpansive mapping and C a subset of X which maps C into itself. Let $v_0 \in C$ and $\widetilde{b} \in \mathbb{N}$ such that $\widetilde{b} \geq \|p_0 - v_0\| + \|v_0\| + 1$ for p_0 some fixed point of U in C. Write $\mathcal{B}_{\leq} := B_{\leq}(\widetilde{b})$ and $b := 2\widetilde{b}$.*
For any $k \in \mathbb{N}$ and monotone function $f : \mathbb{N} \to \mathbb{N}$, there are $N \leq 12b(\check{f}^{(R)}(0)+1)^2$ and $x \in C \cap \mathcal{B}_{\leq}$ such that

$$\|U(x) - x\| \leq \frac{1}{f(N)+1} \wedge \forall y \in C \cap \mathcal{B}_{\leq}\left(\|U(y) - y\| \leq \frac{1}{N+1} \to \langle x - v_0, x - y\rangle \leq \frac{1}{k+1}\right),$$

with $R := b^4(k+1)^2$ and $\check{f}(m) := \max\{f(12b(m+1)^2),\, 12b(m+1)^2\}$.

Since, one cannot extend the Lemmas 6.14 and 6.15 to points with unbounded norm, it does not seem to be possible to extend Proposition 6.16 to the original set C. Nevertheless, this result still suffices in proving a metastability property. In the end, we only work with provable bounded iterations and so all the arguments will still hold true inside $C \cap \mathcal{B}_{\leq}$ for $\widetilde{b}$ big enough.

6.3 Metastability of HPPA1

In this section, we obtain the quantitative version of Theorem 6.1. The main point is to compare the relevant sequence (x_n) to a certain sequence defined in the way of Browder:

Given a point $u \in X$, a sequence $(\alpha_n) \subset]0,1]$ and a nonexpansive map $U : X \to X$. For each $n \in \mathbb{N}$, consider the function $U_n : X \to X$ defined by

$$U_n(x) := \alpha_n u + (1 - \alpha_n)U(x).$$

Then, for all $n \in \mathbb{N}$, U_n is a strict contraction with contraction constant $1 - \alpha_n < 1$ and so, by Banach's contraction principle, U_n has an unique fixed point z_n. Thus, z_n is defined implicitly by

$$z_n = \alpha_n u + (1 - \alpha_n)U(z_n).$$

We call such iteration (z_n) the *Browder type sequence* associated with u, (α_n) and the nonexpansive map U.

In the proof of Theorem 6.1, the authors first show, under some conditions, the strong convergence of the Browder type sequence for a particular map U and latter prove that $\|x_n - z_n\|$ converges to zero. In our quantitative analysis, we extract a bound on the metastability of (z_n), and then show how a bound on the metastability of (z_n) effectively implies a bound on the metastability of (x_n).
Essentially by the same arguments as in Chapter 4, we start by showing that there exists a monotone function $\phi_{\mathrm{Br}'}$ that is a bound for the metastability of the Browder type sequences for an arbitrary sequence (α_n) converging to zero:

$$\forall k \in \mathbb{N} \widetilde{\forall} f : \mathbb{N} \to \mathbb{N} \exists N \leq \phi_{Br'}(k, f) \forall i, j \in [N, fN] \left(\|z_i - z_j\| \leq \frac{1}{k+1} \right).$$

We compute the value of $\phi_{\mathrm{Br}'}$ in several steps.

We make a theoretical comment first. In subsection 4.4.1, the boundedness condition on C implied the existence of a fixed point in C from the fact that almost fixed points existed. Here without that assumption we must consider a quantitative version of such existential statement. It is possible to work only with a radius N such that the ball $B_N(0)$ contains almost fixed points for any level of accuracy. This comment follows from the interpretation of the existence of a fixed point in C:

$$\exists x \in C\ (U(x) = x)$$
$$\exists x \in C \forall k \in \mathbb{N} \left(\|U(x) - x\| \trianglelefteq_{\mathbb{R}} \frac{1}{k+1} \right)$$
$$\exists N \in \mathbb{N} \exists x \in C \cap B_{\trianglelefteq}(N)\, \forall k \in \mathbb{N} \left(\|U(x) - x\| \trianglelefteq_{\mathbb{R}} \frac{1}{k+1} \right)$$
$$\exists N \in \mathbb{N} \forall k \in \mathbb{N} \exists x \in C \cap B_{unlhd}(N) \left(\|U(x) - x\| \trianglelefteq_{R} \frac{1}{k+1} \right),$$

where we used $\mathsf{BC}^{\omega}_{\mathsf{bd}}$ at the last step. An even more general stance would be to consider a bound on the metastable version

$$\widetilde{\forall} f : \mathbb{N} \to \mathbb{N} \exists N \in \mathbb{N} \exists x \in C \cap B_{\trianglelefteq}(N) \left(\|U(x) - x\| \trianglelefteq_{\mathbb{R}} \frac{1}{f(N)+1} \right).$$

However, to simplify the computations, one may work with a bound on the norm of some fixed point in C.

As commented in Chapter 4, Browder's proof makes use of sequential weak compactness, which can be bypassed in quantitative analysis. This can be done by applying the general principle with the function $\varphi(x, y) := \langle x - v_0, y \rangle$. Condition (b) of Proposition 4.14 will correspond to an instance of Proposition 6.16. To see condition (a) of Proposition 4.14, we will see that (z_n) is a bounded sequence and compute a rate of asymptotic regularity, i.e., a monotone function $\chi : \mathbb{N} \to \mathbb{N}$ such that

$$\forall k \in \mathbb{N} \forall n \geq \chi(k) \left(\|U(z_n) - z_n\| \leq \frac{1}{k+1} \right).$$

These properties of (z_n) are shown in the next two lemmas.

Lemma 6.17. *For X a real Hilbert space and C a subset (not necessarily bounded), consider a sequence $(\alpha_n) \subset]0, 1[$, $U : C \to C$ a nonexpansive mapping, v_0 a point in C and let (z_n) be the associated Browder type sequence. If p_0 is some fixed point of U, then for all $n \in \mathbb{N}$,*

$$\|z_n - p_0\| \leq 2\|v_0 - p_0\|. \tag{6.15}$$

Proof. We have the following, for any $n \in \mathbb{N}$

$$\begin{aligned}
\|z_n - p_0\|^2 &\leq \|\alpha_n v_0 + (1 - \alpha_n)U(z_n) - p_0\|^2 = \|\alpha_n(v_0 - p_0) + (1 - \alpha_n)(U(z_n) - p_0)\|^2 \leq \\
&\leq (1 - \alpha_n)^2 \|U(z_n) - p_0\|^2 + 2\langle \alpha(v_0 - p_0), z_n - p_0 \rangle = \\
&= (1 - \alpha_n)^2 \|U(z_n) - U(p_0)\|^2 + 2\langle \alpha(v_0 - p_0), z_n - p_0 \rangle \leq \\
&\leq (1 - \alpha_n)^2 \|z_n - p_0\|^2 + 2\alpha_n \|v_0 - p_0\| \|z_n - p_0\|,
\end{aligned}$$

where at the second inequality we used $\|x+y\|^2 \leq \|y\|^2 + 2\langle x, x+y\rangle$.
This implies

$$\left(1-(1-\alpha_n)^2\right)\|z_n - p_0\|^2 \leq 2\alpha_n \|v_0 - p_0\|z_n - p_0\|,$$

from which follows (6.15)

$$\|z_n - p\| \leq \frac{2}{2-\alpha_n}\|v_0 - p_0\| \leq 2\|v_0 - p_0\|.$$

$\square$

In particular, (z_n) and $(U(z_n))$ are bounded and, with p_0 some fixed point of U, we have, for all $n \in \mathbb{N}$

$$\|z_n\| \leq 2\|v_0 - p_0\| + \|p_0\|,$$
$$\|U(z_n) - v_0\| \leq 3\|v_0 - p_0\|,$$
$$\|U(z_n)\| \leq 2\|v_0 + p_0\| + \|p_0\|.$$

Next we give a rate of asymptotic regularity for the Browder type sequence.

Lemma 6.18. *Let* $\mathrm{a} : \mathbb{N} \to \mathbb{N}$ *be a rate of convergence towards zero for* (α_n)*, and* $b \in \mathbb{N}$ *a positive bound on* $\|U(z_n) - v_0\|$*. Then*

$$\forall k \in \mathbb{N} \forall n \geq \mathrm{a}(b(k+1)) \left(\|z_n - U(z_n)\| \leq \frac{1}{k+1}\right). \tag{6.16}$$

Proof. For a given $k \in \mathbb{N}$, take an arbitrary $n \geq \mathrm{a}(b(k+1))$. Then,

$$\|z_n - U(z_n)\| = \alpha_n \|v_0 - U(z_n)\| \leq \frac{b}{b(k+1)} = \frac{1}{k+1}.$$

$\square$

Next we turn to the projection argument. In order to apply Proposition 6.16 from the previous section we need to consider some natural number $\widetilde{b} \geq \|v_0 - p_0\| + \|v_0\| + 1$. It will be important that for all $n \in \mathbb{N}$, $\|z_n\| \leq \widetilde{b}$ and $\|U(z_n) - v_0\| \leq 2\widetilde{b}$ and so we consider a natural number $\widetilde{b} \geq \max\{\|v_0 - p_0\| + \|v_0\| + 1, 2\|v_0 - p_0\| + \|p_0\|\}$. Define $\mathcal{B}_{\leq} := B_{\leq}(\widetilde{b})$ and $b := 2\widetilde{b}$. Notice that b satisfies the assumption on Lemma 6.18. Recall the functions of section 5.1:

$$r : \mathbb{N} \to \mathbb{N}, \quad r(k) = b^4(k+1)^2. \tag{6.17}$$

and, for every $g : \mathbb{N} \to \mathbb{N}$,

$$\omega_g : \mathbb{N} \to \mathbb{N}, \quad \omega_g(m) = \max\{g(12b(m+1)^2),\, 12b(m+1)^2\}. \tag{6.18}$$

Then from Proposition 6.16, we have

Proposition 6.19. *For any natural number k and monotone function $f : \mathbb{N} \to \mathbb{N}$, there are $N \leq 12b(\omega_f^{(r(k))}(0)+1)^2$ and $x \in C \cap \mathcal{B}_{\leq}$ such that*

$$\|U(x)-x\| \leq \frac{1}{f(N)+1} \wedge \forall y \in C \cap \mathcal{B}_{\leq} \left(\|U(y)-y\| \leq \frac{1}{N+1} \to \langle x - v_0, x - y \rangle \leq \frac{1}{k+1} \right).$$

Hence, in Proposition 4.14, the condition (a) is satisfied by the function

$$\bar{\alpha}(k) = \mathrm{a}(b(k+1))$$

and condition (b) by the function

$$\bar{\beta}(k,f) = 12b(\omega_f^{(r(k))}(0)+1)^2,$$

with b a natural number as stated. We write the functions under a bar to avoid confusion with the parameters of the Halpern type proximal point algorithm.
Therefore we can apply the quantitative Proposition 4.14 to obtain that for every $k \in \mathbb{N}$ and any monotone function $f \in \mathbb{N}^{\mathbb{N}}$, there exists $N \leq \psi(k,f)$ and $x \in C \cap \mathcal{B}_{\leq}$ such that

$$\|U(x)-x\| \leq \frac{1}{f(N)+1} \wedge \forall n \in [N, f(N)] \left(\langle x - v_0, x - z_n \rangle \leq +\frac{1}{k+1} \right), \tag{6.19}$$

where

$$\begin{aligned} \psi(k,f) &= \bar{\alpha}\left(\bar{\beta}(k,\widehat{f})\right) = \mathrm{a}\left(b\left(\beta\left(k,\widehat{f}\right)+1\right)\right) = \mathrm{a}\left(b\left(12b\left(\omega_{\widehat{f}}^{(r(k))}(0)+1\right)^2+1\right)\right) \\ &= \mathrm{a}\left(12b^2\left(\omega_{\widehat{f}}^{(r(k))}(0)+1\right)^2+b\right). \end{aligned}$$

with

$$\widehat{f}(m) = f(\bar{\alpha}(m,f)) = f(\mathrm{a}(bm+b))).$$

This functional ψ satisfies condition (i) of Proposition 5.2 and we are only missing the second condition corresponding to the combinatorial part of the argument. Similarly to section 5.1, but with the general sequence (α_n) in place of the sequence $(\frac{1}{n+1})$, we have for all $x \in C \cap \mathcal{B}_{\leq}$ and $k, n, \ell \in \mathbb{N}$, if

$$\|U(x)-x\| \leq \frac{1}{2b(\ell+1)(k+1)^2} \wedge \langle x - v_0, x - z_n \rangle \leq \frac{1}{2(k+1)^2} \wedge \alpha_n \geq \frac{1}{\ell+1}$$

then, $\|u_n - x\| < \dfrac{1}{k+1}$.
Following Kohlenbach's arguments in [27, Lemma 2.11] (with $s_j = 1 - \alpha_j$) we arrive at

$$\alpha_n \|z_n - x\|^2 \leq \alpha_n \langle x - v_0, x - z_n \rangle + \frac{1}{2(\ell+1)(k+1)^2},$$

and the result is obtained by noticing that $\frac{1}{(\ell+1)\alpha_n} \leq 1$.

Therefore, condition (ii) of Proposition 5.2 holds with

$$\gamma(k,n,f) := 2b(\ell+1)(k+1)^2 - 1,\ \delta(k) := 2(k+1)^2 - 1,\ \eta(k,n,f) := f(n) \text{ and}$$
$$\sigma(k,n,f) := n, \text{ where } \ell := \lceil \tfrac{1}{\alpha_{f(n)}} \rceil - 1 \text{ ensures that } \alpha_{f(n)} \geq \tfrac{1}{\ell+1}.$$

Finally, the conclusion of Proposition 5.2 yields the desired quantitative result.

Theorem 6.20 (Quant.Browder II)**.** *Let X be a real Hilbert space and C a non-empty convex subset of X (not necessarily bounded). Let $U : C \to C$ be a nonexpansive mapping, v_0 a point in C and $\widetilde{b} \in \mathbb{N}$ such that $\widetilde{b} \geq \max\{\|v_0 - p_0\| + \|v_0\| + 1, 2\|v_0 - p_0\| + \|p_0\|\}$, for p_0 some fixed point of U in C, and write $b = 2\widetilde{b}$. Consider a sequence $(\alpha_n) \subset]0,1[$ converging to zero with rate of convergence* $\mathrm{a} : \mathbb{N} \to \mathbb{N}$*. Let $h : \mathbb{N} \to \mathbb{N}$ be a monotone function such that for all $n \in \mathbb{N}$, $\alpha_n \geq \dfrac{1}{h(n)+1}$. Then, for (z_n) the corresponding Browder type sequence, we have*

$$\forall k \in \mathbb{N} \widetilde{\forall} f : \mathbb{N} \to \mathbb{N} \exists N \leq \phi_{Br'}(k,f) \forall i,j \in [N, fN] \left(\|z_i - z_j\| \leq \frac{1}{k+1} \right),$$

where

$$\phi_{\mathrm{Br}'}(k,f) := \mathrm{a}\left(12b^2 \left(g^{(R)}(0)+1\right)^2 + b\right),$$

with $R := 64b^4(k+1)^4$, $g(m) := \max\{\overline{f}(\mathrm{a}(12b^2(m+1)^2 + b)), 12b(m+1)^2\}$ and $\overline{f}(m) := \max\{8b\,(\,h(f(m))+1)\,(k+1)^2 - 1, f(m)\}$.

Lets now look at how a bound on the metastability of (x_n) can be computed from a bound on the metastability of a certain Browder type sequence.

Theorem 6.21. *Let T be a maximal monotone operator, S the set of zeros of T and p_0 a point in S. Consider sequences $(\alpha_n) \subset]0,1[$, $(\beta_n) \subset \mathbb{R}^+$ and $(e_n) \subset X$. With $x_0, u \in X$, let (x_n) be the corresponding Halpern type proximal point iteration inductively defined by* (HPPA_1).
Assume the existence of monotone functions A, B, c, E $: \mathbb{N} \to \mathbb{N}$*, $\beta \in \mathbb{R}^+$ and $\ell \in \mathbb{N}$ such that the conditions* (Q2) *to* (Q4) *and* (Q6) *hold. Let (z_n) be the Browder type iteration associated to the sequence (α_n), the nonexpansive map J_β and to the point u. If a monotone function $\phi : \mathbb{N} \times \mathbb{N}^{\mathbb{N}} \to \mathbb{N}$ is a bound on the metastability of (z_n), i.e.*

$$\forall k \in \mathbb{N} \widetilde{\forall} f : \mathbb{N} \to \mathbb{N} \, \exists N \leq \phi(k,f) \, \forall i,j \in [N, f(N)] \left(\|z_i - z_j\| \leq \frac{1}{k+1} \right),$$

and $D \in \mathbb{N}$ is a positive upper bound on $\|J_\beta(z_n) - u\|$, then

$$\forall k \in \mathbb{N} \widetilde{\forall} f : \mathbb{N} \to \mathbb{N} \, \exists N \leq \Phi_1(k,f) \, \forall i,j \in [N, f(N)] \left(\|x_i - x_j\| \leq \frac{1}{k+1} \right),$$

where $\Phi_1(k,f) := \max\{\theta_1[\mathrm{A},\mathrm{R},\mathrm{E},D_0]\left(\overline{k}\right),\ \phi\left(\overline{k},h[\overline{k},f]\right)\}$, *with*

$$\begin{aligned}
&\theta_1 \ \textit{is as in Lemma 6.5},\\
&\mathrm{R}(k) := \max\{\mathrm{B}(2D(\ell+1)(k+1)-1),\, \mathrm{c}(2D(k+1)-1)\},\\
&D_0 := d_1\left(1+\lceil\textstyle\sum_{i=0}^{\mathrm{E}(0)}\|e_i\|\rceil\right)+\lceil\|p_0-u\|\rceil+D,\\
&d_1 \ \textit{as in Lemma 6.3},\\
&\overline{k} := 3k+2 \ \textit{and} \ h[k,f](m) := f\left(\max\{\theta_1[\mathrm{A},\mathrm{R},\mathrm{E},D_0](k),m\}\right).
\end{aligned}$$

Proof. From condition (Q6) it follows,

$$\forall n\in\mathbb{N}\left(\sum_{i=0}^{n}\|e_i\|\le\mathcal{E}\right),$$

with $\mathcal{E}=1+\lceil\sum_{i=0}^{\mathrm{E}(0)}\|e_i\|\rceil$.

From Lemma 6.3 we get that (x_n) is bounded and

$$\forall n\in\mathbb{N}\ (\|x_n-p_0\|\le d_1(\mathcal{E}),$$

where $d_1(k):=\lceil\max\{\|u-p_0\|,\|x_0-p_0\|\}\rceil+k$.

We have, for all $n\in\mathbb{N}$, since $u-z_n=(1-\alpha_n)(u-J_\beta(z_n)$,

$$\begin{aligned}
\|x_n-z_n\| &\le \|x_n-p_0\|+\|p_0-u\|+\|u-z_n\| =\\
&= \|x_n-p_0\|+\|p_0-u\|+(1-\alpha_n)\|u-J_\beta(z_n)\|\le\\
&\le d_1(\mathcal{E})+\|p_0-u\|+D,
\end{aligned}$$

and define $D_0:=d_1(\mathcal{E})+\lceil\|p_0-u\|\rceil+D$.
Following [5], we have,

$$\begin{aligned}
\|x_{n+1}-z_n\| &\le (1-\alpha_n)\|J_{\beta_n}(x_n)-J_\beta(z_n)\|+\|e_n\|\le\\
&\le (1-\alpha_n)\|J_{\beta_n}(x_n)-J_{\beta_n}(z_n)\|+\|J_{\beta_n}(z_n)-J_\beta(z_n)\|+\|e_n\|\le\\
&\le (1-\alpha_n)\|x_n-z_n\|+\frac{|\beta-\beta_n|}{\beta}\|z_n-J_\beta(z_n)\|+\|e_n\|=\\
&= (1-\alpha_n)\|x_n-z_n\|+\alpha_n\frac{|\beta-\beta_n|}{\beta}\|u-J_\beta(z_n)\|+\|e_n\|\le\\
&\le (1-\alpha_n)\|x_n-z_n\|+\alpha_n\frac{|\beta-\beta_n|}{\beta}D+\|e_n\|, \qquad (6.20)
\end{aligned}$$

using the fact that the functions J_{β_n} are nonexpansive and, by Lemma 6.11 (resolvent identity), $J_\beta(z_n) = J_{\beta_n}\left(\frac{\beta_n}{\beta}z_n + (1-\frac{\beta_n}{\beta})J_\beta(z_n)\right)$.

On the other hand, from

$$\begin{aligned}\|z_n - z_{n+1}\| &\leq \|(\alpha_n - \alpha_{n+1})(u - J_\beta(z_{n+1})) + (1-\alpha_n)(J_\beta(z_n) - J_\beta(z_{n+1}))\| \leq \\ &\leq |\alpha_{n+1} - \alpha_n|\|u - J_\beta(z_{n+1})\| + (1-\alpha_n)\|z_n - z_{n+1}\| \leq \\ &\leq |\alpha_{n+1} - \alpha_n|D + (1-\alpha_n)\|z_n - z_{n+1}\|,\end{aligned}$$

we conclude,

$$\|z_n - z_{n+1}\| \leq \frac{|\alpha_{n+1} - \alpha_n|}{\alpha_n}D. \tag{6.21}$$

Hence we conclude, from (6.20) and (6.21), that for all $n \in \mathbb{N}$

$$\|x_{n+1} - z_{n+1}\| \leq (1-\alpha_n)\|x_n - z_n\| + \alpha_n\frac{|\beta - \beta_n|}{\beta}D + \|e_n\| + \frac{|\alpha_{n+1} - \alpha_n|}{\alpha_n}D. \tag{6.22}$$

Then, with $r_n := D\left(\frac{|\beta-\beta_n|}{\beta} + \frac{|\alpha_{n+1}-\alpha_n|}{\alpha_n^2}\right)$, for all $n \in \mathbb{N}$

$$\|x_{n+1} - z_{n+1}\| \leq (1-\alpha_n)\|x_n - z_n\| + \alpha_n r_n + \|e_n\|,$$

and we can apply Lemma 6.5 with $s_n = \|x_n - z_n\| \leq D_0$, $\gamma_n := \|e_n\|$, $\mathrm{G} = \mathrm{E}$ and

$$\mathrm{R}(k) := \max\{\mathrm{B}(2D(\ell+1)(k+1)-1),\, \mathrm{c}(2D(k+1)-1)\}. \tag{6.23}$$

This R satisfies the condition of Lemma 6.4. In fact, for any $k \in \mathbb{N}$ and $n \geq \mathrm{R}(k)$,

$$\begin{aligned}r_n &\leq D\left(\frac{|\beta - \beta_n|}{\beta} + \frac{|\alpha_{n+1} - \alpha_n|}{\alpha_n^2}\right) \leq D\left(\frac{1}{\beta 2D(k+1)(\ell+1)} + \frac{1}{2D(k+1)}\right) \leq \\ &\leq D\left(\frac{\ell+1}{2D(k+1)(\ell+1)} + \frac{1}{2D(k+1)}\right) = \frac{1}{k+1}.\end{aligned}$$

Hence, we conclude

$$\forall n \geq \theta_1[\mathrm{A}, \mathrm{R}, \mathrm{E}, D_0](k)\ \left(\|x_n - z_n\| \leq \frac{1}{k+1}\right).$$

Now, by the assumption on ϕ and by Proposition 4.1,

$$\forall k \in \mathbb{N}\widetilde{\forall} f : \mathbb{N} \to \mathbb{N}\, \exists N \leq \widetilde{\Phi_1}(k, f) \forall i, j \in [N, fN]\ \left(\|x_i - z_i\| \leq \frac{1}{k+1} \wedge \|z_i - z_j\| \leq \frac{1}{k+1}\right),$$

where $\widetilde{\Phi_1}(k,f) := \max\{\theta_1[\mathrm{A},\mathrm{R},\mathrm{E},D_0](k), \phi(k,h[k,f])\}$, with R as in (6.23) and the monotone function $h[k,f] : \mathbb{N} \to \mathbb{N}$ defined by

$$h[k,f](m) := f\left(\max\{\theta_1[\mathrm{A},\mathrm{R},\mathrm{E},D_0](k), m\}\right).$$

For given $k \in \mathbb{N}$ and monotone $f : \mathbb{N} \to \mathbb{N}$, there is $N \le \widetilde{\Phi_1}(3k+2,f)$ such that for any $i,j \in [N, fN]$

$$\|x_i - z_i\| \le \frac{1}{3(k+1)} \text{ and } \|z_i - z_j\| \le \frac{1}{3(k+1)}.$$

It follows for $i,j \in [N, fN]$,

$$\|x_i - x_j\| \le \|x_i - z_i\| + \|z_i - z_j\| + \|z_j - x_j\| \le \frac{1}{k+1}.$$

The proof concludes with the observation that $\widetilde{\Phi_1}(3k+2,f) = \Phi_1(k,f)$. $\square$

We now show the quantitative version when one instead have the condition (Q7).

Theorem 6.22. *Under the same assumptions as in Theorem 6.21, but with* (Q6) *replaced by* (Q7)*, we have*

$$\forall k \in \mathbb{N}\, \widetilde{\forall} f : \mathbb{N} \to \mathbb{N}\, \exists N \le \Phi_2(k,f)\, \forall i,j \in [N, f(N)] \left(\|x_i - x_j\| \le \frac{1}{k+1}\right),$$

where $\Phi_2(k,f) := \max\{\theta_1[\mathrm{A},\mathrm{R},\mathbf{0},D_0](\overline{k}), \phi(\overline{k}, h[\overline{k},f])\}$*, with*

θ_1 *is as in Lemma 6.5,*
$\mathbf{0} := \lambda m.\, 0$ *is the zero constant function,*
$\mathrm{R}(k) := \max\{\mathrm{B}(3D(\ell+1)(k+1)-1), \mathrm{c}(3D(k+1)-1), \mathrm{E}(3k+2)\}$,
$D_0 := d_2\left(\lceil \max_{i<\mathrm{E}(0)}\{\frac{\|e_i\|}{\alpha_i^2}, 1\}\rceil\right) + \lceil \|p_0 - u\| \rceil + D$,
d_2 *as in Lemma 6.3,*
$\overline{k} := 3k+2$ *and* $h[k,f](m) := f\left(\max\{\theta_1[\mathrm{A},\mathrm{R},\mathbf{0},D_0](k), m\}\right).$

Proof. From condition (Q7),

$$\forall n \in \mathbb{N}\left(\frac{\|e_n\|}{\alpha_n^2} \le \mathcal{E}'\right),$$

with $\mathcal{E}' = \left\lceil \max_{i<\mathrm{E}(0)}\left\{\frac{\|e_i\|}{\alpha_i^2}, 1\right\}\right\rceil$.

From Lemma 6.3 we conclude

$$\forall n \in \mathbb{N}\left(\|x_n - p_0\| \le d_2(\mathcal{E}'\right),$$

where $d_2(k) := \lceil \max\{2\left(\|u - p_0\| + k\right), \|x_0 - p_0\|\}\rceil$.

Similar to before we conclude, for all $n \in \mathbb{N}$,

$$\|x_n - z_n\| \leq d_2(\mathcal{E}') + \|p_0 - u\| + D,$$

and define $D_0 := d_2(\mathcal{E}') + \lceil \|p_0 - u\|\rceil + D$.
Like in the previous proof, we conclude that for all $n \in \mathbb{N}$,

$$\|x_{n+1} - z_{n+1}\| \leq (1 - \alpha_n)\|x_n - z_n\| + \alpha_n \frac{|\beta - \beta_n|}{\beta} D + \|e_n\| + \frac{|\alpha_{n+1} - \alpha_n|}{\alpha_n} D. \tag{6.24}$$

Now, with $r_n := D\left(\frac{|\beta-\beta_n|}{\beta} + \frac{|\alpha_{n+1}-\alpha_n|}{\alpha_n^2}\right) + \frac{\|e_n\|}{\alpha_n}$, for all $n \in \mathbb{N}$

$$\|x_{n+1} - z_{n+1}\| \leq (1 - \alpha_n)\|x_n - z_n\| + \alpha_n r_n,$$

and we can apply Lemma 6.5 with $s_n = \|x_n - z_n\| \leq D_0$, $\gamma_n := \|e_n\|$, $\mathrm{G} = \mathbf{0}$ (zero function) and

$$\mathrm{R}(k) := \max\{\mathrm{B}(3D(\ell+1)(k+1) - 1),\, \mathrm{c}(3D(k+1) - 1), \mathrm{E}(3k+2)\}. \tag{6.25}$$

This R now satisfies the condition of Lemma 6.4, since for any $k \in \mathbb{N}$ and $n \geq \mathrm{R}(k)$,

$$\begin{aligned}
r_n = D\left(\frac{|\beta - \beta_n|}{\beta} + \frac{|\alpha_{n+1} - \alpha_n|}{\alpha_n^2}\right) + \frac{\|e_n\|}{\alpha_n^2} \leq \\
\leq D\left(\frac{1}{\beta 3D(k+1)(\ell+1)} + \frac{1}{3D(k+1)}\right) + \frac{1}{3(k+1)} \leq \\
\leq D\left(\frac{\ell+1}{3D(k+1)(\ell+1)} + \frac{1}{3D(k+1)}\right) + \frac{1}{3(k+1)} = \frac{1}{k+1}.
\end{aligned}$$

Hence, we conclude

$$\forall n \geq \theta_1[\mathrm{A}, \mathrm{R}, \mathbf{0}, D_0](k) \left(\|x_n - z_n\| \leq \frac{1}{k+1}\right).$$

Now, by the assumption on ϕ and by Proposition 4.1,

$$\forall k \in \mathbb{N} \widetilde{\forall} f : \mathbb{N} \to \mathbb{N}\, \exists N \leq \widetilde{\Phi_2}(k, f) \forall i, j \in [N, fN] \left(\|x_i - z_i\| \leq \frac{1}{k+1} \wedge \|z_i - z_j\| \leq \frac{1}{k+1}\right),$$

where $\widetilde{\Phi_2}(k, f) := \max\{\theta_1[\mathrm{A}, \mathrm{R}, \mathbf{0}, D_0](k), \phi(k, h[k, f])\}$, with R as in (6.23) and the monotone function $h[k, f] : \mathbb{N} \to \mathbb{N}$ defined by

$$h[k, f](m) := f\left(\max\{\theta_1[\mathrm{A}, \mathrm{R}, \mathbf{0}, D_0](k), m\}\right).$$

The result now follows by triangle inequality as previously and $\Phi_2(k, f) = \widetilde{\Phi_2}(3k+2, f)$. $\square$

We can also consider the quantitative condition (Q2′). The proofs are identical to those of the previous two results but, instead of Lemma 6.5 and the function θ_1, one makes use of Lemma 6.6 and the function θ_2.

Theorem 6.23. *Under the same assumptions as in Theorem 6.21, but with* (Q2) *replaced by* (Q2′)*. For each $k \in \mathbb{N}$, let $\ell_k \in \mathbb{N}$ be a natural number satisfying $\prod_{i=0}^{n_k-1}(1-\alpha_i) \geq \frac{1}{\ell_k+1}$, where $n_k := \max\{\mathrm{R}(3k+2), \mathrm{E}(3k+2)+1\}$ with* R *as below. We have*

$$\forall k \in \mathbb{N}\, \widetilde{\forall} f : \mathbb{N} \to \mathbb{N}\, \exists N \leq \Phi_3(k,f)\, \forall i,j \in [N, f(N)] \left(\|x_i - x_j\| \leq \frac{1}{k+1} \right),$$

where $\Phi_3(k,f) := \max\{\theta_2[\mathrm{A}, \mathrm{R}, \mathrm{E}, D_0, \ell_k](\overline{k}), \phi(\overline{k}, h[\overline{k}, f])\}$*, with*

θ_2 *is as in Lemma 6.6,*
$\mathrm{R}(k) := \max\{\mathrm{B}(2D(\ell+1)(k+1)-1),\, \mathrm{c}(2D(k+1)-1)\}$,
$D_0 := d_1\left(1 + \lceil \sum_{i=0}^{\mathrm{E}(0)} \|e_i\| \rceil\right) + \lceil \|p_0 - u\| \rceil + D$,
d_1 *as in Lemma 6.3,*
$\overline{k} := 3k+2$ *and* $h[k,f](m) := f\left(\max\{\theta_2[\mathrm{A}, \mathrm{R}, \mathrm{E}, D_0, \ell_k](k), m\}\right)$.

Theorem 6.24. *Under the same assumptions as in Theorem 6.21, but with* (Q2) *replaced by* (Q2′) *and* (Q6) *replaced by* (Q7)*. For each $k \in \mathbb{N}$, let $\ell_k \in \mathbb{N}$ be a natural number satisfying $\prod_{i=0}^{n_k-1}(1-\alpha_i) \geq \frac{1}{\ell_k+1}$, where $n_k := \mathrm{R}(3k+2)$ with* R *as below. We have*

$$\forall k \in \mathbb{N}\, \widetilde{\forall} f : \mathbb{N} \to \mathbb{N}\, \exists N \leq \Phi_4(k,f)\, \forall i,j \in [N, f(N)] \left(\|x_i - x_j\| \leq \frac{1}{k+1} \right),$$

where $\Phi_4(k,f) := \max\{\theta_2[\mathrm{A}, \mathrm{R}, \mathbf{0}, D_0, \ell_k](\overline{k}), \phi(\overline{k}, h[\overline{k}, f])\}$*, with*

θ_2 *is as in Lemma 6.6,*
$\mathrm{R}(k) := \max\{\mathrm{B}(3D(\ell+1)(k+1)-1),\, \mathrm{c}(3D(k+1)-1), \mathrm{E}(3k+2)\}$,
$D_0 := d_2\left(\lceil \max_{i<\mathrm{E}(0)}\{\frac{\|e_i\|}{\alpha_i^2}, 1\} \rceil\right) + \lceil \|p_0 - u\| \rceil + D$,
d_2 *as in Lemma 6.3,*
$\overline{k} := 3k+2$ *and* $h[k,f](m) := f\left(\max\{\theta_2[\mathrm{A}, \mathrm{R}, \mathbf{0}, D_0, \ell_k](k), m\}\right)$.

Hence, under these conditions, a bound for the metastability of (x_n) is obtained by instantiating the function ϕ in these last four theorems with the function $\phi_{\mathrm{Br}'}$ from Theorem 6.20.

An elementary proof of Browder's theorem for the case $C = B_1(0)$ and $v_0 = 0$, due to Halpern [18], already does not use weak compactness. In [27], Kohlenbach generalized this proof to a general closed convex set C and point v_0, and strong convergence of (z_n) is obtained when (α_n) is any decreasing sequence in $]0,1[$ even if it does not converges to zero. Analyzing this proof with the monotone functional interpretation, Kohlenbach showed a different quantitative version of Browder's theorem. The next result is theorem 4.2 from [27], slightly adapted to our context.

Theorem 6.25 (Quant.Browder III). *Let X be a real Hilbert space, $d \in \mathbb{N}^*$ and $C \subset X$ be a bounded closed and convex subset with $b \geq diam(C)$. Let $U : C \to C$ be a nonexpansive mapping and $v_0 \in C$. Let (α_n) be a sequence in $]0,1[$ that converges to zero and $h : \mathbb{N} \to \mathbb{N}$ be a monotone function such that for all $n \in \mathbb{N}$, $\alpha_n \geq \frac{1}{h(n)+1}$. Let $\mathrm{a} : \mathbb{N} \to \mathbb{N}$ be a monotone function satisfying* (Q1)*. Then, for (z_n) the corresponding Browder type sequence, we have*

$$\forall k \in \mathbb{N} \widetilde{\forall} f : \mathbb{N} \to \mathbb{N} \exists N \leq \phi_{\mathrm{Br}''}(k,f) \forall i,j \in [N, fN] \left(\|z_i - z_j\| \leq \frac{1}{k+1} \right),$$

where

$$\phi_{\mathrm{Br}''}(k,f) := \mathrm{a}\left(f_{h,\mathrm{a}}^{(r(k))}(0) \right),$$

with $f_{h,\mathrm{a}}(n) := h(a(n) + f(a(n))$ and $r(k) := 4b^2(k+1)^2$.
Instead of C being bounded it suffices to have (z_n) bounded and the bound above also holds if $b \geq \|z_n - v_0\|$, for all $n \in \mathbb{N}$.
If (α_n) is an decreasing sequence in $]0,1[$ (not necessarily converging to 0), then the bound can be simplified to $\phi_{\mathrm{Br}''}(k,f) := \breve{f}^{(b^2(k+1)^2)}(0)$, where $\breve{f}(n) := n + f(n)$.

Proof. The bounds above follow directly from [27, theorem 4.2] with some observations. There $\alpha_n = 1 - s_n$, $z_n = \breve{u}_n$ and $\frac{1}{k+1} = \varepsilon$. We are considering the function $\mathrm{a} : \mathbb{N} \to \mathbb{N}$ to be a rate of convergence towards 0 for (α_n) instead of a "quasi-rate of convergence" – which is a stronger condition than the existence of Kohlenbach's function χ_g. By monotonicity of a, the function $(\mathrm{a})^M$ is just a. The definition of $f_{h,\mathrm{a}}$ also simplifies using the monotonicity of the function h. Finally, instead of adapting the bounds to conclude the result directly to the interval $[N, fN]$, we apply Kohlenbach's theorem and, since trivially $[N, fN] \subset [N, N + fN]$, the result follows. ☐

We can also apply theorems 6.21-6.24 with ϕ instantiated by $\phi_{\mathrm{Br}''}$ to obtain a bound on the metastability of (x_n).

6.4 Metastability of HPPA2

In this section, we are considering the sequence (x_n) defined by (HPPA_2) with $u = x_0$.
Xu's original proof of of Theorem 6.2 begins by seeing that the sequence (x_n) is bounded. Then, using sequential weak compactness together with the projection argument, he concludes that $\limsup \langle \tilde{x} - x_0, \tilde{x} - x_n \rangle \leq 0$, with $\tilde{x}$ the projection point of x_0 onto S. Finally,

the strong convergence is obtained by applying Lemma 6.4. We will obtain a metastable version of Theorem 6.2 by following the quantitative form of those same arguments. The projection argument's treatment is as in section 6.2, and the sequential weak compactness can be bypassed by using the general principle from Chapter 4.

Let p_0 be a point in S. By an inductive argument similar to that of Lemma 6.3, one sees that for all $n \in \mathbb{N}$,

$$\|x_n - p_0\| \leq \|x_0 - p_0\| + \sum_{i=0}^{n-1} \|e_i\|.$$

Hence, (x_n) is bounded and, if E is a Cauchy rate for $\sum \|e_n\|$, we have for all $n \in \mathbb{N}$,

$$\begin{aligned} \|x_n - p_0\| &\leq \|x_0 - p_0\| + \mathcal{E}, \\ \|x_n\| &\leq \|x_0 - p_0\| + \|p_0\| + \mathcal{E}, \\ \|x_n - x_0\| &\leq 2\|x_0 - p_0\| + \mathcal{E}, \end{aligned}$$

with $\mathcal{E} := 1 + \lceil \sum_{i=0}^{E(0)} \|e_i\| \rceil$.

Now, we will compute a rate of asymptotic regularity for the sequence (x_n) that will correspond to the function α in Proposition 4.14 when $\varphi(x, y) := \langle x - x_0, y \rangle$. We do this in two steps. First, the following rate of convergence is easily derived from the original proof of Theorem 6.2.

Lemma 6.26. *Consider monotone functions* $\mathrm{a}, \mathrm{E} : \mathbb{N} \to \mathbb{N}$ *satisfying* (Q1) *and* (Q6) *and let* $b \in \mathbb{N}$ *be a positive bound on* $\|x_n - p_0\|$*, for some* $p_0 \in S$*. Define* $\xi(k) := \max\{\mathrm{a}(4b(k+1) - 1), \mathrm{E}(2k+1) + 1\}$*. Then,*

$$\forall k \in \mathbb{N} \forall n \geq \xi(k) \left(\|x_{n+1} - J_{\beta_n}(x_n)\| \leq \frac{1}{k+1} \right).$$

Proof. Let $k \in \mathbb{N}$ be given. Consider $n \geq \xi(k)$. Then, from condition (Q6), we have in particular

$$\|e_n\| \leq \frac{1}{2(k+1)}.$$

Using the fact that J_{β_n} is nonexpansive and p_0 is a fixed point of J_{β_n}, we get

$$\begin{aligned} \|x_{n+1} - J_{\beta_n}(x_n)\| &= \|\alpha_n x_0 + (1 - \alpha_n)(J_{\beta_n}(x_n) + e_n) - J_{\beta_n}(x_n)\| \leq \\ &\leq \alpha_n \|x_0 - J_{\beta_n}(x_n)\| + \|e_n\| \leq \\ &\leq \alpha_n(\|x_0 - p_0\| + \|p_0 - J_{\beta_n}(x_n)\|) + \|e_n\| \leq \\ &\leq \alpha_n(\|x_0 - p_0\| + \|p_0 - x_n\|) + \|e_n\| \leq \alpha_n(2b) + \|e_n\| \leq \\ &\leq \frac{2b}{4b(k+1)} + \frac{1}{2(k+1)} = \frac{1}{k+1}. \end{aligned}$$

$\square$

Next we compute a rate of asymptotic regularity for the sequence (x_n) in relation to a resolvent function J_γ.

Proposition 6.27. *Consider a real number $\gamma > 0$ and monotone functions* $\mathrm{a}, \mathrm{B}, \mathrm{E} : \mathbb{N} \to \mathbb{N}$ *satisfying* (Q1)*,* (Q5) *and* (Q6) *and let $b \in \mathbb{N}$ be a positive bound on $\|x_n - p_0\|$, for some $p_0 \in S$.*
Define $\chi_\gamma(k) := \max\{\xi(4k+3), \mathrm{B}(8b(k+1)\lceil\gamma\rceil - 1)\} + 1$*, where ξ is as in Lemma 6.26. Then,*

$$\forall k \in \mathbb{N} \forall n \geq \chi_\gamma(k) \left(\|x_n - J_\gamma(x_n)\| \leq \frac{1}{k+1} \right).$$

Proof. First notice that, by the monotonicity of a and E, we have

$$\chi_\gamma(k) \geq \mathrm{a}(8b(k+1) - 1) + 1 \text{ and } \xi(k) \geq E(4k+3) + 2.$$

For $n + 1 \geq \chi_\gamma(k)$, using the resolvent identity,

$$\begin{aligned}
&\|x_{n+1} - J_\gamma(x_{n+1})\| \leq \alpha_n\|x_0 - J_\gamma(x_{n+1})\| + \|J_{\beta_n}(x_n) - J_\gamma(x_{n+1})\| + \|e_n\| \leq \\
&\leq 2b\alpha_n + \left\| J_\gamma\left(\frac{\gamma}{\beta_n}x_n + (1 - \frac{\gamma}{\beta_n})J_{\beta_n}(x_n)\right) - J_\gamma(x_{n+1})\right\| + \|e_n\| \leq \\
&\leq 2b\alpha_n + \left\|\frac{\gamma}{\beta_n}x_n + (1 - \frac{\gamma}{\beta_n})J_{\beta_n}(x_n) - x_{n+1}\right\| + \|e_n\| \leq \\
&\leq 2b\alpha_n + \frac{\gamma}{\beta_n}\|x_n - x_{n+1}\| + \left|1 - \frac{\gamma}{\beta_n}\right| \|J_{\beta_n}(x_n) - x_{n+1}\| + \|e_n\| \leq \\
&\leq 2b\alpha_n + \frac{\gamma}{\beta_n}2b + \left|1 - \frac{\gamma}{\beta_n}\right| \|J_{\beta_n}(x_n) - x_{n+1}\| + \|e_n\| \leq \\
&\leq \frac{2b}{8b(k+1)} + \frac{2b\gamma}{8b(k+1)\gamma} + \frac{1}{4(k+1)} + \frac{1}{4(k+1)} = \frac{1}{k+1},
\end{aligned}$$

which concludes the result. $\square$

We now turn to the projection argument. It will be useful to consider a natural number $\widetilde{b}$ that is big enough to guarantee the arguments of section 6.2, to be a bound on the sequence $\|x_n\|$ and such that $2\widetilde{b}$ satisfies the assumption on b in the previous result. It is not hard to see, that a natural number

$$\widetilde{b} \geq \|x_0 - p_0\| + \max\{\|x_0\|, \|p_0\|\} + \mathcal{E}, \tag{6.26}$$

with $\mathcal{E} := 1 + \lceil\sum_{i=0}^{E(0)} \|e_i\|\rceil$ and p_0 some zero of T, satisfies all those requirements.
In the sequel, we consider $\widetilde{b}$ with this condition and define $\mathcal{B} := B_{\widetilde{b}}(0)$ and $b := 2\widetilde{b}$. By Proposition 6.16, with $\widetilde{b}$ in the condition above and with $J = J_1$ the resolvent function $(Id + T)^{-1}$, we have

Proposition 6.28. *For any natural number k and monotone function $f : \mathbb{N} \to \mathbb{N}$, there are $N \leq 12b(\omega_f^{(r(k))}(0)+1)^2$ and $x \in C \cap \mathcal{B}$ such that*

$$\|J(x) - x\| \leq \frac{1}{f(N)+1} \wedge \forall y \in C \cap \mathcal{B} \left(\|J(y) - y\| \leq \frac{1}{N+1} \to \langle x - x_0, x - y \rangle \leq \frac{1}{k+1} \right),$$

with $r(k)$ and ω_g as in (6.17) *and* (6.18).

The condition (a) in Proposition 4.14 (with $\varphi(x, y) := \langle x - x_0, y \rangle$ and the map J) is satisfied by the function χ_1 and condition (b) by the function

$$\bar{\beta}(k, f) = 12b(\omega_f^{(r(k))}(0) + 1)^2,$$

with b a natural number as explained above.
Therefore, by the general principle of Chapter 4, we obtain

Proposition 6.29. *For every $k \in \mathbb{N}$ and any monotone function $f \in \mathbb{N}^{\mathbb{N}}$, there exists $N \leq \psi(k, f)$ and $x \in C \cap \mathcal{B}$ such that*

$$\|J(x) - x\| \leq \frac{1}{f(N)+1} \wedge \forall n \in [N, fN] \left(\langle x - x_0, x - x_n \rangle \leq \frac{1}{k+1} \right),$$

where $\psi(k, f) = \chi_1(\bar{\beta}(k, \widehat{f}))$, with $\widehat{f}(m) = f(\chi_1(m))$.

Since χ_1 is a rate of convergence, we even have the Proposition 6.29 with $\forall n \geq N$ in place of $\forall n \in [N, fN]$.

At this point in our analysis, we have gave a quantitative version of the relevant (weak) projection argument and removed the original sequential weak compactness argument. In order to obtain the metastable version of Theorem 6.2, we will argue that the conditions of Lemmas 6.9 and 6.10 are satisfied.

Notice that in Proposition 6.28, we only analyzed the projection onto the set of fixed points of the map J. However, there is no problem in focusing on that particular set since all the sets of fixed points of resolvent functions coincide. This last statement, requires a quantitative version, that we show below.

Lemma 6.30. *For all $k, n \in \mathbb{N}$ and $x \in X$,*

$$\|J(x) - x\| \leq \frac{1}{\delta(k, n) + 1} \to \|J_{\beta_n}(x)\| \leq \frac{1}{k+1},$$

where $\delta(k, n) := (k+1)(1 + \max_{j \leq n}\{\lceil |1 - \beta_j| \})- 1$.

Proof. Assume that, for given $k, n \in \mathbb{N}$ and $x \in X$, we have $\|J(x) - x\| \leq \frac{1}{\delta(k,n)+1}$. Then,

$$\|J(x) - x\| \leq \frac{1}{(k+1)(1+|1-\beta_n|)}.$$

By the definition of the resolvent function, we have with $\|e\| \leq \frac{1}{(k+1)(|\beta_n-1|+1)}$,

$$\begin{aligned} &J(x) = x + e \Leftrightarrow x \in x + e + T(x+e) \Leftrightarrow \\ &\Leftrightarrow -e \in T(u+e) \Leftrightarrow x + (1-\beta_n)e \in x + e + \beta_n T(x+e) \Leftrightarrow \\ &\Leftrightarrow J_{\beta_n}(x + (1-\beta_n)e) = x + e \end{aligned}$$

From this we conclude,

$$\begin{aligned} \|J_{\beta_n}(x) - x\| &\leq \|J_{\beta_n}(x) - J_{\beta_n}(x + (1-\beta_n)e)\| + \|J_{\beta_n}(x + (1-\beta_n)e) - x\| \leq \\ &\leq \|x - (x + (1-\beta_n)e)\| + \|e\| \leq \|e\|(1+|1-\beta_n|) \leq \frac{1}{k+1}. \end{aligned}$$

□

The previous proof show that to ensure that x is an almost fixed point of J_{β_n} with error $\frac{1}{k+1}$, it is enough to have x be an almost fixed point of J with error $\frac{1}{(k+1)(1+|1-\beta_n|)}$. The $\max_{j \leq n}$ in the definition of δ is only to ensure monotonicity of the bound. From this monotonicity property of δ, follows

$$\|J(x) - x\| \leq \frac{1}{\delta(k,n)+1} \to \forall j \leq n \left(\|J_{\beta_j}(x) - x\| \leq \frac{1}{k+1} \right). \tag{6.27}$$

This lemma is enough for our quantitative analysis, however one can replaced the functions J and J_{β_n} by any two resolvent functions. In fact, by the same argument, for any $\alpha, \beta > 0$ and $k \in \mathbb{N}$,

$$\|J_\alpha(x) - x\| \leq \frac{1}{(k+1)(1+|1-\frac{\beta}{\alpha}|)} \to \|J_\beta(x) - x\| \leq \frac{1}{k+1}.$$

From Proposition 6.29 and Lemma 6.30, we derive the following result.

Proposition 6.31. *For every $k \in \mathbb{N}$ and any monotone function $f : \mathbb{N} \to \mathbb{N}$, there exists $N \leq \psi(k, \sigma_f)$ and $x \in C \cap \mathcal{B}$ such that*

$$\forall n \in [n, fN] \left(\|J_{\beta_n}(x) - x\| \leq \frac{1}{f(n)+1} \wedge \langle x - x_0, x - x_n \rangle \leq \frac{1}{k+1} \right),$$

where $\psi(k, f)$ is as in Proposition 6.29 and $\sigma_f(m) = \delta(f(m), f(m))$.

Proof. Let $k \in \mathbb{N}$ and a monotone function $f : \mathbb{N} \to \mathbb{N}$ be given. Consider the monotone function σ_f defined as in the statement of the proposition. Notice that by the definition of the function δ, for all $m \in \mathbb{N}$, $\sigma_f(m) \geq f(m)$.
By Proposition 6.29, we see that there are $N \leq \psi(k, \sigma_f)$ and $x \in C \cap \mathcal{B}$ such that

$$\|J(x) - x\| \leq \frac{1}{\sigma_f(N) + 1}$$

and

$$\forall n \in [N, \sigma_f(N)] \left(\langle x - x_0, x - x_n \rangle \leq \frac{1}{k+1} \right).$$

By (6.27), we conclude, for $n \leq f(N)$

$$\|J_{\beta_n}(x) - x\| \leq \frac{1}{f(N) + 1},$$

which gives the first conjunct.
On the other hand, since $[N, fN] \subset [N, \sigma_f(N)]$, the second conjunct is also true. □

Now define, for any $n \in \mathbb{N}$ and $x \in X$, the sequences

$$v_{n,x} := \|J_{\beta_n}(x) - x\|^2 + 2\|J_{\beta_n}(x) - x\|\|x_n - x\|,$$

and

$$r_{n,x} := 2\langle x - x_0, x - x_{n+1} \rangle.$$

The next result is a direct application of Proposition 6.31 and will correspond to condition (4) of Lemmas 6.9 and 6.10.

Proposition 6.32. *For every $k \in \mathbb{N}$ and any monotone function $f : \mathbb{N} \to \mathbb{N}$, there exists $N \leq \Psi(k, f)$ and $x \in C \cap \mathcal{B}$ such that*

$$\forall n \in [n, fN] \left(v_{n,x} \leq \frac{1}{f(n) + 1} \wedge r_{n,x} \leq \frac{1}{k+1} \right),$$

where $\Psi(k, f) := \psi(2k + 1, \sigma_g)$, with $g(m) = (1 + 2b)(f(m) + 1) - 1$.

Proof. By Proposition 6.31, there are $N \leq \psi(2k + 1, \sigma_g)$ and $x \in C \cap \mathcal{B}$ such that for all $n \in [N, gN]$

$$\|J_{\beta_n}(x) - x\| \leq \frac{1}{g(N) + 1} \tag{6.28}$$

and

$$\langle x - x_0, x - x_n \rangle \leq \frac{1}{2(k+1)}. \tag{6.29}$$

If $n \in [N, fN]$, then $n+1 \in [N, f(N)+1] \subset [N, gN]$. Hence, by (6.29), for $n \in [N, fN]$,

$$\langle x - x_0, x - x_{n+1} \rangle \leq \frac{1}{2(k+1)},$$

which implies $r_{n,x} \leq \frac{1}{k+1}$.

For $n \in [N, fN]$, and noticing that $b \geq \|x_n - x\|$ and $\|J_{\beta_n}(x) - x\| \leq 1$, we have

$$\begin{aligned} v_{n,x} &= \|J_{\beta_n}(x) - x\| \left(\|J_{\beta_n}(x) - x\| + 2\|x_n - x\|\right) \leq \|J_{\beta_n}(x) - x\| \left(1 + 2b\right) \leq \\ &\leq \frac{1+2b}{g(N)+1} = \frac{1+2b}{(1+2b)(f(N)+1)} = \frac{1}{f(N)+1}, \end{aligned}$$

which concludes the proof. □

Next we will argue that, for any $n \in \mathbb{N}$ and $x \in X$,

$$\|x_{n+1} - x\|^2 \leq (1 - \alpha_n) \left(\|x_n - x\|^2 + v_{n,x}\right) + \alpha_n r_{n,u} + \gamma_{n,u}$$

with $v_{n,u}$ and $r_{n,x}$ as before and with $\gamma_{n,u} := \|e_n\|(\|e_n\| + 2\|J_{\beta_n}(x_n) - x\|)$.

This inequality is obtained by using the subdifferential inequality,

$$\|x + y\|^2 \leq \|x\|^2 + 2\langle y, x + y \rangle,$$

in the following way,

$$\begin{aligned} \|x_{n+1} - x\|^2 &= \|(1 - \alpha_n)(J_{\beta_n}(x_n) + e_n - x) + \alpha_n(x_0 - x)\|^2 \leq \\ &\leq (1 - \alpha_n)^2 \|J_{\beta_n}(x_n) + e_n - x\|^2 + \alpha_n(2\langle x_0 - x, x_{n+1} - x \rangle) \leq \\ &\leq (1 - \alpha_n) \left(\|J_{\beta_n}(x_n) - x\| + \|e_n\|\right)^2 + \alpha_n r_{n,x} \leq \\ &\leq (1 - \alpha_n)\|J_{\beta_n}(x_n) - x\|^2 + \alpha_n r_{n,x} + \|e_n\|(\|e_n\| + 2\|J_{\beta_n}(x_n) - x\|) \leq \\ &\leq (1 - \alpha_n) \left(\|J_{\beta_n}(x_n) - J_{\beta_n}(x)\| + \|J_{\beta_n}(x) - x\|\right)^2 + \alpha_n r_{n,x} + \gamma_{n,x} \leq \\ &\leq (1 - \alpha_n)\|x_n - x\|^2 + (1 - \alpha_n) v_{n,x} + \alpha_n r_{n,u} + \gamma_{n,x}. \end{aligned}$$

In order to apply the Lemmas 6.9 and 6.10, we only need to compute a Cauchy rate G for the sequence of partial sums $(\sum_{i=0}^{n} \gamma_{i,x})$, which is easily derived from the function E.

Lemma 6.33. *Consider a monotone function* E *satisfying* (Q6). *Define the monotone function* $\mathrm{G} : \mathbb{N} \to \mathbb{N}$ *by*

$$\mathrm{G}(k) := \mathrm{E}((k+1)(1+2b) - 1).$$

Then,

$$\forall k \in \mathbb{N} \forall n \in \mathbb{N} \forall x \in \mathcal{B} \left(\sum_{i=\mathrm{G}(k)+1}^{\mathrm{G}(k)+n} \gamma_{i,x} \leq \frac{1}{k+1} \right).$$

Proof. Let $k, n \in \mathbb{N}$ and $x \in \mathcal{B}$ be given.
For all $j \in \mathbb{N}$, since $\|x\|, \|p_0\| \leq \frac{b}{2}$ and $\|x_j - p_0\| \leq b$, we get

$$\|J_{\beta_j}(x_j) - x\| \leq \|x_j - p_0\| + \|p_0 - x\| \leq 2b.$$

Also see that, from condition (Q6), for any $j \geq \mathrm{E}(k)+1$ one has $\|e_j\| \leq \frac{1}{k+1}$. Hence, by the definition of G,

$$\forall j \geq \mathrm{G}(k)+1 \; (\|e_j\| \leq 1)\,.$$

Therefore,

$$\begin{aligned}\sum_{j=\mathrm{G}(k)+1}^{\mathrm{G}(k)+n} \gamma_{j,x} = \sum_{j=\mathrm{G}(k)+1}^{\mathrm{G}(k)+n} \|e_j\|(\|e_j\| + \|J_{\beta_j}(x_j) - x\|) \leq \\ \leq (1+2b) \sum_{j=\mathrm{G}(k)+1}^{\mathrm{G}(k)+n} \|e_j\| \leq \frac{1+2b}{(k+1)(1+2b)} = \frac{1}{k+1}.\end{aligned}$$

□

We are now ready to give the quantitative version of Theorem 6.2.

Theorem 6.34. *Let T be a maximal monotone operator. Consider sequences $(\alpha_n) \subset]0,1[$, $(\beta_n) \subset \mathbb{R}^+$ and $(e_n) \subset X$. With $x_0 \in X$, let (x_n) be the corresponding Halpern type proximal point iteration inductively defined by* (HPPA_2) *with $u = x_0$.*
Assume the existence of monotone functions a,A,B,E : $\mathbb{N} \to \mathbb{N}$ *such that the conditions* (Q1), (Q2), (Q5) *and* (Q6) *hold. Consider $\widetilde{b} \in \mathbb{N}$ such that* (6.26) *holds and define $b := 2\widetilde{b}$. Then*

$$\forall k \in \mathbb{N}\, \widetilde{\forall} f : \mathbb{N} \to \mathbb{N}\, \exists N \leq \Phi_1(k,f)\, \forall i,j \in [N, f(N)] \left(\|x_i - x_j\| \leq \frac{1}{k+1}\right),$$

where $\Phi_1(k,f) := \Theta_1[\mathrm{A}, \Psi, \mathrm{G}, b^2](4(k+1)^2 - 1, f)$, with Θ_1 as in Lemma 6.9, Ψ as in Proposition 6.32 and G *as in Lemma 6.33.*

Proof. Let $k \in \mathbb{N}$ and a monotone function f be given.
Apply Lemma 6.9, with $s_{n,x} := \|x_n - x\|$ and $\Omega = \mathcal{B}$, to $4(k+1)^2 - 1$ and f. Then, there are $x \in \mathcal{B}$ and $N \leq \Theta_1[\mathrm{A}, \Psi, \mathrm{G}, b^2](4(k+1)^2 - 1, f)$ such that for $n \in [N, fN]$,

$$\|x_n - x\|^2 \leq \frac{1}{4(k+1)^2}.$$

Hence, for $n \in [N, fN]$,

$$\|x_n - x\| \leq \frac{1}{2(k+1)},$$

and for $i, j \in [N, fN]$,

$$\|x_i - x_j\| \leq \|x_i - x\| + \|x_j - x\| \leq \frac{1}{k+1}.$$

$\square$

We can also consider the condition (Q2) replaced by (Q2′).

Theorem 6.35. *Assume the hypothesis of Theorem 6.34 with* (Q2) *replaced by the condition* (Q2′) *and the existence of a monotone function* $h : \mathbb{N} \to \mathbb{N}^*$ *such that for all* $n \in \mathbb{N}$, $\prod_{i=0}^{n}(1 - \alpha_i) \geq \frac{1}{h(n)}$. *We have*

$$\forall k \in \mathbb{N}\, \widetilde{\forall} f : \mathbb{N} \to \mathbb{N}\, \exists N \leq \Phi_2(k, f)\, \forall i, j \in [N, f(N)] \left(\|x_i - x_j\| \leq \frac{1}{k+1} \right),$$

where $\Phi_2(k, f) := \Theta_2[\mathrm{A}, \Psi, \mathrm{G}, b^2, h](4(k+1)^2 - 1, f)$, *with* Θ_2 *as in Lemma 6.10,* Ψ *as in Proposition 6.32 and* G *as in Lemma 6.33.*

Proof. Similar to before, considering $s_{n,x} := \|x_n - x\|^2$, we apply instead Lemma 6.10 to $4(k+1)^2 - 1$ and f and then use triangle inequality to conclude the result. $\square$

Epilogue

We have shown quantitative versions of known mathematical theorems in the context of fixed point theory and the theory of monotone operators, while being guided by the bounded functional interpretation. A general method for bypassing sequential weak compactness arguments in proof mining was obtained and clarifies previous quantitative analysis of Browder's and Wittmann's theorems. We also saw that the bounded functional interpretation of the projection argument is easier than the previous analysis obtained using the monotone functional interpretation. These results are the first look at proof mining guided by the bounded functional interpretation and show it to be a useful proof theoretic technique to the proof mining program. This functional interpretation was shown to be a valid option that can be used to carry out quantitative analysis.

Future work concerning the use of the bounded functional interpretation should be additional analyses of mathematical proofs. Additionally, one can be attentive to possible instances where its use may prove simpler than using other techniques (as it happened in the projection argument). Currently, together with Bruno Dinis, a quantitative analysis of a theorem due to Yao and Noor in [54] is being done. This result is concerned with the strong convergence of a multi-parameters proximal point algorithm and its proof requiers the existence of a certain $\limsup$ to be assumed. The change of the $\limsup$ to rational approximations, as explained in section 4.1.3, is of paramount importance in restricting the extracted information to Gödel's T. Another quantitative analysis that is being considered is that of a theorem due to Wang and Cui in [49]. Its original proof follows a discussion by (non-trivial) cases and it is expected that section 4.1.4 can help shed some light on the proper way to carry out the mining. Other results that are natural good candidates for a quantitative analysis are further theorems in fixed point theory and further results related to the proximal point algorithm. Another topic that was not discussed in this thesis but may have merit is to see how do the complexity of bounds extracted using the bounded functional interpretation compares to information extracted by other means, e.g. when using the monotone functional interpretation. This

could lead to a deeper understanding of the differences between the bounded functional interpretation and other methods.
Regarding the weak compactness, at the moment, it is not expected that weak compactness arguments can be removed from discussions that end up proving weak convergence results, e.g. as in the analysis of Baillon's theorem [28]. It will be interesting to see if this is in fact true or if some adaptation of the method shown here for removing weak compactness could be used in those cases.
These ideas for future work show that the use of the bounded functional interpretation has a lot to offer to the thriving proof mining program and that there is still much to understand.

Bibliography

[1] J. Avigad. A variant of the double-negation translation. Technical report, Carnegie Mellon Technical Report CMU-PHIL 179, 2006.

[2] H. H. Bauschke. The approximation of fixed points of compositions of nonexpansive mappings in Hilbert space. *Journal of Mathematical Analysis and Applications*, 202(1):150–159, 1996.

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

[4] O. A. Boikanyo and G. Moroşanu. A proximal point algorithm converging strongly for general errors. *Optimization Letters*, 4(4):635–641, 2010.

[5] O. A. Boikanyo and G. Moroşanu. Inexact Halpern-type proximal point algorithm. *Journal of Global Optimization*, 51(1):11–26, 2011.

[6] F. E. Browder. Convergence of approximants to fixed points of nonexpansive nonlinear mappings in Banach spaces. *Archive for Rational Mechanics and Analysis*, 24(1):82–90, 1967.

[7] B. Dinis and P. Pinto. Metastability of the proximal point algorithm with multi-parameters. *arXiv preprint 1906.09129*, 2019.

[8] P. Engrácia. *Proof-theoretical studies on the bounded functional interpretation.* PhD thesis, Universidade de Lisboa, 2009.

[9] S. Feferman. Kreisel's "unwinding" program. In P. Odifreddi, editor, *Kreiseliana*, pages 247–273. A K Peters, Wellesley, MA, 1996.

[10] F. Ferreira. Proof interpretations and majorizability. In F. Delon, U. Kohlenbach, P. Maddy, and F. Stephan, editors, *Logic Colloquium 2007*, volume 35, pages 32–81. Cambridge University Press, 2007.

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

[12] F. Ferreira, L. Leuştean, and P. Pinto. On the removal of weak compactness arguments in proof mining. *Advances in Mathematics*, 354, 2019.

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

[14] J. Gaspar. Factorization of the Shoenfield-like bounded functional interpretation. *Notre Dame Journal of Formal Logic*, 50(1):53–60, 2009.

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

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

[17] O. Güler. On the convergence of the proximal point algorithm for convex minimization. *SIAM Journal on Control and Optimization*, 29(2):403–419, 1991.

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

[19] W. Howard. Hereditarily majorizable functionals of finite type. In A. S. Troelstra, editor, *Metamathematical Investigation of Intuitionistic Arithmetic and Analysis*, pages 454–461. Lecture Notes in Math., Vol. 344. Springer, New York, 1973.

[20] S. Kamimura and W. Takahashi. Approximating solutions of maximal monotone operators in Hilbert spaces. *Journal of Approximation Theory*, 106(2):226–240, 2000.

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

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

[23] U. Kohlenbach. A logical uniform boundedness principle for abstract metric and hyperbolic spaces. *Electronic Notes in Theoretical Computer Science*, 165:81–93, 2006.

[24] U. Kohlenbach. Proof interpretations and the computational content of proofs in mathematics. *Bulletin of the European Association for Theoretical Computer Science. EATCS*, (93):143–175, 2007.

[25] U. Kohlenbach. *Applied Proof Theory: Proof Interpretations and their Use in Mathematics*. Springer Science & Business Media, 2008.

[26] U. Kohlenbach. Effective uniform bounds from proofs in abstract functional analysis. In B. Cooper, B. Loewe, and A. Sorbi, editors, *New Computational Paradigms: Changing Conceptions of What is Computable*, pages 223–258. Springer, New York, 2008.

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

[28] U. Kohlenbach. A uniform quantitative form of sequential weak compactness and Baillon's nonlinear ergodic theorem. *Communications in Contemporary Mathematics*, 14(01):1250006, 2012.

[29] U. Kohlenbach. Recent progress in proof mining in nonlinear analysis. *Journal of Applied Logics – IfCoLoG Journal of Logics and their Applications*, 4(10):3361–3410, 2017.

[30] U. Kohlenbach. Kreisel's 'shift of emphasis' and contemporary proof mining. Chapter of forthcoming book *Intuitionism, Computation, and Proof: Selected themes from the research of G. Kreisel*, 2018.

[31] U. Kohlenbach. Proof-theoretic Methods in Nonlinear Analysis. In B. Sirakov, P. N. de Souza, and M. Viana, editors, *Proceedings of the International Congress of Mathematicians 2018*, volume 2, pages 62–82. World Scientific, 2019.

[32] U. Kohlenbach and L. Leuştean. Effective metastability of Halpern iterates in CAT(0) spaces. *Advances in Mathematics*, 231(5):2526–2556, 2012.

[33] U. Kohlenbach and P. Oliva. Proof mining: a systematic way of analyzing proofs in mathematics. *Tr. Mat. Inst. Steklova*, 242(Mat. Logika i Algebra):147–175, 2003.

[34] U. Kohlenbach and A. Sipos. The finitary content of sunny nonexpansive retractions. *arXiv preprint 1812.04940*, 2018.

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

[36] G. Kreisel. On the interpretation of non-finitist proofs–Part I. *The Journal of Symbolic Logic*, 16(4):241–267, 1951.

[37] G. Kreisel. On the interpretation of non-finitist proofs–Part II. *The Journal of Symbolic Logic*, 17(1):43–58, 1952.

[38] J.-L. Krivine. Opérateurs de mise en mémoire et traduction de Gödel. *Archive for Mathematical Logic*, 30(4):241–267, 1990.

[39] L. Leuştean. Rates of asymptotic regularity for Halpern iterations of nonexpansive mappings. *Journal of Universal Computer Science*, 13(11):1680–1691, 2007.

[40] G. J. Minty. Monotone (nonlinear) operators in Hilbert space. *Duke Mathematical Journal*, 29(3):341–346, 1962.

[41] S. Reich. Strong convergence theorems for resolvents of accretive operators in Banach spaces. *Journal of Mathematical Analysis and Applications*, 75(1):287–292, 1980.

[42] R. T. Rockafellar. Monotone operators and the proximal point algorithm. *SIAM Journal on Control and Optimization*, 14(5):877–898, 1976.

[43] C. Spector. Provably recursive functionals of analysis: a consistency proof of analysis by an extension of principles formulated in current intuitionistic mathematics. In F. D. E. Dekker, editor, *Recursive Function Theory*, volume 5, pages 1–27. American Mathematical Society, Providence, Rhode Island, 1962.

[44] T. Streicher and U. Kohlenbach. Shoenfield is Gödel after Krivine. *Mathematical Logic Quarterly*, 53(2):176–179, 2007.

[45] T. Streicher and B. Reus. Classical logic, continuation semantics and abstract machines. *Journal of Functional Programming*, 8(6):543–572, 1998.

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

[47] A. S. Troelstra. *Metamathematical Investigation of Intuitionistic Arithmetic and Analysis*, volume 344. Springer Science & Business Media, 1973.

[48] A. S. Troelstra. Introductory note to 1958 and 1972. In S. Feferman et al., editors, *Kurt Gödel Collected Works*, volume II, pages 217–241. Oxford University Press, New York, 1990.

[49] F. Wang and H. Cui. On the contraction-proximal point algorithms with multi-parameters. *Journal of Global Optimization*, 54(3):485–491, 2012.

[50] K. Weihrauch. *Computable Analysis: an Introduction.* Springer Science & Business Media, 2012.

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

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

[53] H.-K. Xu. A regularization method for the proximal point algorithm. *Journal of Global Optimization*, 36(1):115–125, 2006.

[54] Y. Yao and M. A. Noor. On convergence criteria of generalized proximal point algorithms. *Journal of Computational and Applied Mathematics*, 217(1):46–55, 2008.