Personal reflections on becoming an applied proof theorist, and thoughts for the future

Thomas Powell
University of Bath

International Conference on Applied Proof Theory 2022
(on the occassion of Ulrich Kohlenbach's 60th birthday)

Pescara
29 August – 2 September 2022

These slides will be available at
`https://t-powell.github.io/talks`

Outline

My first scientific note (never submitted):

A note on proof interpretations and Dialectica categories

Thomas Powell

written early 2010

Introduction

In most applications of functional interpretations, the interpretation is a means to an end, a syntactic translation that extracts witnesses from proofs. As a consequence, on the whole proof theorists pay little attention to the structural properties of functional interpretations, though the use of interpretations is central to their work.

This is a short note in which we discuss and bring together several works which view functional interpretations from a more abstract perspective. Our ultimate aim is to construct a general abstract framework in which a range of interpretations can be compared and better understood.

A rich variety of functional interpretations have been developed since Gödel invented his prototype, ranging from early examples used to prove foundational theorems to more exotic modern varieties tailored specifically for the purpose of proof mining. It is natural, then, to ask whether we can isolate the key features of functional interpretations and develop a unifying framework in which they can be compared, either on a syntactic or a semantic level.

The question of unifying proof interpretations has been separately considered from each of these perspectives, by Oliva and de Paiva respectively. de Paiva used the language of categorical logic to gain a better semantic understanding of the Dialectica interpretation - constructing and studying the *Dialectica category* [7]. This yielded some interesting results, notably that the Dialectica interpretation itself behaves rather badly - and that the best that can be achieved in terms of a categorical semantics is a model of linear logic. However, an interpretation of the linear modality ! via a comonad on the category produced an elegant model of a variant of the Dialectica interpretation - the Diller-Nahm interpretation.

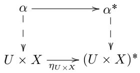
Over a decade later, in his work on unifying functional interpretations [5], Oliva introduced, on a syntactic level, a parametrised functional interpretation with a uniform soundness proof, from which a large family of familiar interpretations could be retrieved.

This note attempts to combine these ideas in the construction of a uniform semantic framework for functional interpretations, based on de Paiva's Dialectica category. Studying interpretations in this way yields insights into their structure that may appear hidden in a more syntactic presentation. The idea is that many different interpretations can be modelled in an abstract way via comands on the Dialectica category. While this has been observed before, by Biering in [1] for instance, we show that

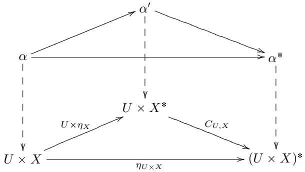
How I first viewed proof interpretations:

$\mathbb{P}_\star(U \times TX)$ consists of those formulas in $\mathbb{P}(U \times TX)$ such that $\beta(u, \mathbf{x})$ if and only if $\beta(u, \langle x \rangle)$ for all $x \in \mathbf{x}$. The map $\mathbb{P}_\star(U \times TX) \to \mathbb{P}(U \times X)$ induced by re-indexing along $U \times \eta_X$ has right adjoint, sending $\alpha(u, x)$ to $\forall x \in \mathbf{x} \alpha(u, x)$.

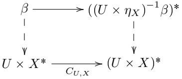
The construction given in [7] is in many ways more elegant that the one given here, in that it has a more abstract formulation. This is in part due to the fact that in de Paiva's setting the free monad extends to a fibred monad, and we get the following useful property, that



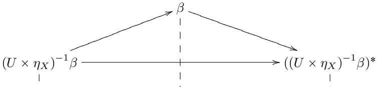
By properties of strength, we can decompose this as



If we define the preorder $\mathbb{P}_\star(U \times TX)$ to be those objects in $\mathbb{P}(U \times TX)$ that for which



then the following diagram commutes



But I was already skeptical…

in **Dial** , where the comonads ! correspond directly to Oliva's generalised bounded quantifier $x \sqsubset t$. In this way we provide categorical models for a number of familiar interpretations, and a general *semantic* framework in which they can be compared.

Our work highlights the fact that, despite the rather peculiar features of the Dialectica interpretation, by enriching the interpreting system with some kind of bounded quantifier we obtain variants that posses excellent structural properties: by interpreting the contraction axiom in a canonical manner we gain a model of the $\rightarrow$, $\wedge$, $\bot$ fragment of logic that identifies proofs that are equivalent under normalisation, which is not the situation with the messy definition by case functionals required for the Dialectica interpretation.

A nice feature of the work begun by de Paiva is its natural link to linear logic and in particular the categorical semantics of linear logic, where the rather mysterious model of Seely is given a concrete illustration by the Dialectica category. We have already referenced the work of Biering [1], which demonstrates that there is certainly potential for the cateogorical semantics of the Dialectica interpretation to be explored further.

However, while it is always important to be able to step back and see things from an abstract perspective, the key significance of functional interpretations today lies in what they are capable of as tools in logic. While the main features of functional interpretations can be expressed in the language of category theory, many of their more interesting aspects lie outside of our framework.

References

[1] B. Biering. *Dialectica Interpretations: a Categorical Analysis*. PhD thesis, IT-University of Copenhagen, 2008.

[2] F. Ferreira and P. Oliva. Bounded functional interpretation. *Annals of Pure and Applied Logic*, 135:73–112, 2005.

[3] J.-Y. Girard. Linear logic. *Theoretical Computer Science*, 50(1):1–102, 1987.

[4] J. M. E. Hyland. Proof theory in the abstract. *Annals of Pure and Applied Logic*, 114:43–78, 2002.

[5] P. Oliva. Unifying functional interpretations. *Notre Dame Journal of Formal Logic*, 47(2):263–290, 2006.

[6] P. Oliva. An analysis of Gödel's dialectica interpretation via linear logic. *dialectica*, 62(2):269–290, 2008.

[7] V. C. V. de Paiva. The Dialectica categories. Technical Report 213, Computer Laboratory, University of Cambridge, Jan 1991.

[8] R. A. G. Seely. Linear logic, *-autonomous categories and cofree coalgebras. *Contemporary Math-*

The moment I got hooked on applied proof theory (Dini's theorem)

Let Φ_n and Φ be closed terms of E-HA$^\omega$ representing (uniformly) continuous functions $[0,1] \to \mathbb{R}$. Then

$$\text{E-HA}^\omega + \text{stuff} \vdash \forall k \in \mathbb{N}\, \forall x \in [0,1]\, \exists n \in \mathbb{N}\, \forall m \geq n(|\Phi_m(x) - \Phi(x)| < 2^{-k})$$

implies that

$$\text{E-HA}^\omega + \text{stuff} \vdash \forall k \in \mathbb{N}\, \exists n \in \mathbb{N}\, \forall x \in [0,1] \forall m \geq n(|\Phi_m(x) - \Phi(x)| < 2^{-k})$$

Moreover, in the special case that pointwise convergence is *monotone* (and thus purely existential), we can extract moduli of uniform convergence even when the proof of pointwise convergence is ineffective. Based on a combination of:

- Proof interpretations,
- Majorizability,
- Representation of compact spaces.

For full details see [Kohlenbach, 2008, pages 112–113].

What I value the most about applied proof theory

- It allows one to be a logician, a computer scientist, and a mathematician, at the same time.
- It encourages you to read widely.
- The field is now established and highly respected (with an excellent textbook), but there is huge potential for the future.

Outline

Set-valued accrective operators

Let X be a real Banach space and X^* its dual. The normalized duality mapping $J : X \to 2^X$ is defined by

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

A set valued operator $A : X \to 2^X$ is called *accretive* if for all $u \in Ax$ and $v \in Ay$ there exists $j \in J(x - y)$ such that

$$\langle u - v, j \rangle \geq 0$$

This generalises the notion of A being *monotone*.

A stronger property: It is ϕ-uniformly accretive for a continuous function $\phi : [0, \infty) \to [0, \infty)$ with $\phi(0) = 0$ and ϕ positive on $(0, \infty)$, if instead

$$\langle u - v, j \rangle \geq \phi(\|x - y\|)$$

Zeros of set-valued operators: The point q is a zero of A if $0 \in Aq$.

Question. Can we compute zeroes of accretive operators?

Computing zeroes of (strong) accretive operators I

- **Space:** X Banach, $J(x) := \{j \in X^* : \langle x, j \rangle = \|x\|^2 = \|j\|^2\}$.
- **Mapping:** $u \in Ax \wedge v \in Ay \implies \exists j \in J(x-y)(\langle u - v, j\rangle \geq \phi(\|x-y\|))$.
- **Target:** $q \in X$ such that $0 \in Aq$.
- **Algorithm:** $x_{n+1} := x_n - \alpha_n u_n$ for $u_n \in Ax_{n+1}$, with $\alpha_n \geq 0$ and $\sum_{i=0}^{\infty} \alpha_i = \infty$.

Lemma (type A)

Suppose that $\|x_n - q\| \leq K$ *for some* $K > 0$*. Then*

$$\|x_{n+1} - q\| \leq \|x_n - q\| - \alpha_n \phi(\|x_{n+1} - q\|)/K$$

Proof. We calculate:

$$\begin{aligned}
\|x_{n+1} - q\|^2 &= \langle x_{n+1} - q, j\rangle \quad \text{for any } j \in J(x_{n+1} - q) \\
&= \langle x_n - q, j\rangle - \alpha_n \langle u_n, j\rangle \quad \text{since } x_{n+1} = x_n - \alpha_n u_n \\
&\leq \langle x_n - q, j\rangle - \alpha_n \phi(\|x_{n+1} - q\|) \quad \text{by strong accretivity} \\
&\leq \|x_n - q\| \cdot \|x_{n+1} - q\| - \alpha_n \phi(\|x_{n+1} - q\|) \quad \text{since } \|j\| = \|x_{n+1} - q\|
\end{aligned}$$

Computing zeroes of (strong) accretive operators II

Lemma (type B)

Let $\{\mu_n\}$ be a sequence of nonnegative reals satisfying

$$\mu_{n+1} \leq \mu_n - \alpha_n \psi(\mu_{n+1})$$

where $\psi : [0, \infty) \to [0, \infty)$ is a continuous function with $\psi(0)$ and ψ positive on $(0, \infty)$, and $\{\alpha_n\}$ is a sequence of positive reals with $\sum_{i=0}^{\infty} \alpha_i = \infty$. Then $\mu_n \to 0$.

Theorem (type A + B)

Let X be a Banach space, $A : X \to 2^X$ a set-valued ψ-uniformly accretive operator, and q a zero of A. Define

$$x_{n+1} = x_n - \alpha_n u_n \text{ for } u_n \in Ax_{n+1}$$

where $\{\alpha_n\}$ is a sequence of positive reals with $\sum_{i=0}^{\infty} \alpha_i = \infty$. Then $x_n \to q$.

Proof. Setting $\mu_n := \|x_n - q\|$, we have $\mu_{n+1} \leq \mu_n - \alpha_n \psi(\mu_{n+1})$ for $\psi(t) := \phi(t)/K$ (by type A lemma). Therefore $x_n \to q$ by type B lemma.

Main results

Lemma (type B quantitative)

Let $\{\mu_n\}$ be a sequence of nonnegative reals satisfying

$$\mu_{n+1} \leq \mu_n - \alpha_n \psi(\mu_{n+1}) + \gamma_n$$

where $\psi : [0, \infty) \to [0, \infty)$ is a nondecreasing function with $\psi(0) = 0$ and ψ positive on $(0, \infty)$. Suppose in addition that $\sum_{i=0}^{\infty} \alpha_i = \infty$ with rate of divergence r, and $\gamma_n / \alpha_n \to 0$ with rate of convergence σ. Then $\mu_n \to 0$ with rate

$$\Phi(\varepsilon) := r\left(\sigma\left(\frac{\psi(\varepsilon)}{2}\right), \frac{2K}{\psi(\varepsilon)}\right) + 1$$

where $K > 0$ is any upper bound on $\{\mu_n\}$.

From this result we can derive, in a uniform way, a number of quantitative "type A + B" results, by analysing different type A lemmas. These include:

- An implicit scheme $x_{n+1} = x_n - \alpha_n u_n$ for $u_n \in A_n x_{n+1}$, and $\{A_n\}$ a sequence of operators that converge to A w.r.t. Hausdorff distance (here we introduced a new abstract predicate $H^*[P, Q, a]$ for dealing with Hausdorff distance);
- Ishikawa type schemes for uniformly continuous operators and in uniformly smooth spaces.

Example: Quantitative version of a result in [Alber et al., 2002]

Theorem ([Kohlenbach and Powell, 2020] – type A + B quantitative)

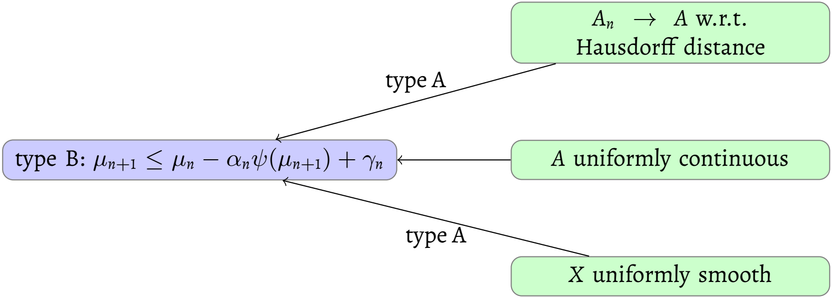
Let $A : D \to 2^X$ with $0 \in Aq$ be uniformly accretive at zero with modulus Θ, and $A_n : D \to 2^X$ be a sequence of operators which uniformly approximates A with rate μ. Let $\{\alpha_i\}$ be a sequence of nonnegative reals such that $\sum_{i=0}^{\infty} \alpha_i = \infty$ with modulus of divergence r, and suppose that $\{x_n\}$ and $\{u_n\}$ are sequences satisfying $x_n \in D$ and

$$x_{n+1} = x_n - \alpha_n u_n, \quad u_n \in A_n x_{n+1}$$

for all $n \in \mathbb{N}$. Finally, suppose $K, K' \in (0, \infty)$ satisfy $\|x_n - q\| < K$ for all $n \in \mathbb{N}$ and $\|q\| < K'$. Then $\|x_n - q\| \to 0$ with rate of convergence

$$\Phi_{\Theta,\mu,r,K,K'}(\varepsilon) := r\left(\mu_{K+K'}\left(\frac{\Theta_K(\varepsilon)}{2K}\right), \frac{K^2}{\Theta_K(\varepsilon)}\right) + 1.$$

Summary: Structure of Kohlenbach/P. 2020



Outline

1. Personal reflections
2. My paper with Ulrich: Set-valued accretive operators
3. Contractive mappings (jww Franziskus Wiesnet)
4. Quantitative recursive inequalities (jww Morenikeji Neri)
5. Thoughts for the future

ψ-contractive mappings

Let X be a real Banach space. A mapping $T : X \to X$ is called *contractive* if for all $x, y \in X$ we have

$$x \neq y \implies \|Tx - Ty\| < \|x - y\|$$

A stronger property: It is ψ-weakly contractive for a continuous function $\psi : [0, \infty) \to [0, \infty)$ with $\psi(0) = 0$ and ψ positive on $(0, \infty)$ if instead

$$\|Tx - Ty\| \leq \|x - y\| - \psi(\|x - y\|)$$

Examples:

- Contraction mappings are ψ-weakly contractive for $\psi(t) := k \in (0, 1]$
- sin is ψ-weakly contractive for $\psi(t) = t^3/8$.

Fixpoints. The point q is a fixpoint of T if $Tq = q$.

Question. Can we compute fixpoints of ψ-weakly contractive operators?

Computing fixpoints of ψ-weakly contractive mappings I

- **Space:** X Banach.
- **Mapping:** $\|Tx - Ty\| \leq \|x - y\| - \psi(\|x - y\|)$.
- **Target:** $q \in X$ such that $Tq = q$.
- **Algorithm:** $x_{n+1} := (1 - \alpha_n)x_n + \alpha_n Tx_n$ with $\alpha_n \geq 0$ and $\sum_{i=0}^{\infty} \alpha_i = \infty$.

Lemma (type A)

We have

$$\|x_{n+1} - q\| \leq \|x_n - q\| - \alpha_n \psi(\|x_n - q\|)$$

Proof. We calculate:

$$\begin{aligned}
\|x_{n+1} - q\| &= \|(1 - \alpha_n)x_n + \alpha_n Tx_n - q\| \quad \text{def. of } \{x_n\} \\
&\leq (1 - \alpha_n)\|x_n - q\| + \alpha_n \|Tx_n - q\| \\
&= (1 - \alpha_n)\|x_n - q\| + \alpha_n \|Tx_n - Tq\| \quad q \text{ fixpoint} \\
&\leq (1 - \alpha_n)\|x_n - q\| + \alpha_n(\|x_n - q\| - \psi(\|x_n - q\|)) \quad \text{property of } T \\
&= \|x_n - q\| - \alpha_n \psi(\|x_n - q\|)
\end{aligned}$$

Computing fixpoints of ψ-weakly contractive mappings II

Lemma (type B)

Let $\{\mu_n\}$ be a sequence of nonnegative real satisfying

$$\mu_{n+1} \leq \mu_n - \alpha_n \psi(\mu_n)$$

where $\psi : [0, \infty) \to [0, \infty)$ is a continuous function with $\psi(0)$ and ψ positive on $(0, \infty)$, and $\{\alpha_n\}$ is a sequence of positive reals with $\sum_{i=0}^{\infty} \alpha_i = \infty$. Then $\mu_n \to 0$.

Theorem (type A + B)

Let X be a Banach space, $T : X \to X$ a ψ-weakly contractive mapping, and q a fixpoint of T. Define

$$x_{n+1} := (1 - \alpha_n)x_n + \alpha_n T x_n$$

where $\{\alpha_n\}$ is a sequence of positive reals with $\sum_{i=0}^{\infty} \alpha_i = \infty$. Then $x_n \to q$.

Proof. Setting $\mu_n := \|x_n - q\|$, we have $\mu_{n+1} \leq \mu_n - \alpha_n \psi(\mu_n)$ (by type A lemma). Therefore $x_n \to q$ by type B lemma.

Main results

Lemma (type B quantitative)

Let $\{\mu_n\}$ be a sequence of nonnegative reals satisfying

$$\mu_{n+1} \leq \mu_n - \alpha_n \psi(\mu_n) + \gamma_n$$

where $\psi : [0, \infty) \to [0, \infty)$ is a nondecreasing function with $\psi(0) = 0$ and ψ positive on $(0, \infty)$. Suppose in addition that $\sum_{i=0}^{\infty} \alpha_i = \infty$ with rate of divergence r, $\alpha \in (0, \alpha]$, and $\gamma_n / \alpha_n \to 0$ with rate of convergence σ. Then $\mu_n \to 0$ with rate

$$\Phi(\varepsilon) := r\left(\sigma\left(\frac{1}{2}\min\left\{\psi\left(\frac{\varepsilon}{2}\right), \frac{\varepsilon}{\alpha}\right\}\right), 2\int_{\varepsilon/2}^{c} \frac{dt}{\psi(t)}\right)$$

and c is an upper bound for $\{\mu_n\}$.

Quantitative "type A + B" results include:

- Mann schemes for asymptotic versions of weakly contractive mappings;
- d-weakly contractive mappings in uniformly smooth spaces;
- perturbed Mann schemes.

Example: Quantitative generalisation of several results in the literature

Theorem ([Powell and Wiesnet, 2021] – type A + B quantitative + rate conversion)

Suppose that $\{A_n\}$ is quasi asymptotically ψ-weakly contractive w.r.t. q and σ, and that the sequence $\{x_n\}$ satisfies

$$x_{n+1} = (1 - \alpha_n)x_n + \alpha_n A_n x_n$$

for $\{\alpha_n\}$ a sequence of nonnegative reals such that $\sum_{n=0}^{\infty} \alpha_n = \infty$. Then whenever $\|x_n - q\|$ is bounded above by some $c > 0$, we have $\|x_n - q\| \to 0$, with rate of convergence

$$\|x_n - q\| \leq F^{-1}\left(2\Psi(c) - \sum_{i=0}^{n-2} \alpha_i\right)$$

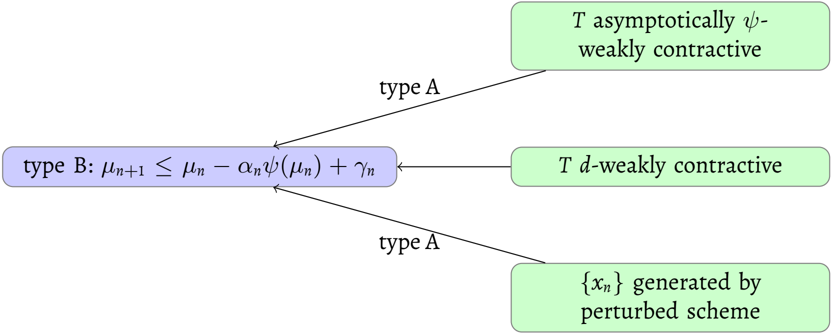
where $F : (0, \infty) \to \mathbb{R}$ is any strictly increasing and continuous function satisfying

$$F(\varepsilon) \geq 2\Psi\left(\frac{\varepsilon}{2}\right) - \alpha \cdot \sigma\left(\frac{1}{2}\min\left\{\psi\left(\frac{\varepsilon}{2}\right), \frac{\varepsilon}{\alpha}\right\}, c\right)$$

and Ψ is given by

$$\Psi(s) := \int^{s} \frac{dt}{\psi(t)}$$

Summary: Structure of P./Wiesnet 2021



Outline

Starting point

Kohlenbach/P. 2020 is based on the recursive inequality:

$$\mu_{n+1} \leq \mu_n - \alpha_n \psi(\mu_{n+1}) + \gamma_n$$

P./Wiesnet 2021 is based on the recursive inequality:

$$\mu_{n+1} \leq \mu_n - \alpha_n \psi(\mu_n) + \gamma_n$$

There are many other papers in applied proof theory based on similar inequalities.

Both of the above are instances of the following more general scheme:

$$\mu_{n+1} \leq \mu_n - \alpha_n \beta_n + \gamma_n$$

What can we say in general about sequences of real numbers that satisfy this recursive inequality?

More concrete challenge:

Suppose that $\{\mu_n\}$, $\{\alpha_n\}$, $\{\beta_n\}$ and $\{\gamma_n\}$ are sequences of nonnegative reals satisfying

$$\mu_{n+1} \leq \mu_n - \alpha_n \beta_n + \gamma_n$$

where, in addition:

- $\sum_{i=0}^{\infty} \alpha_i = \infty$ (intuition: $\{\alpha_n\}$ represent step sizes)
- $\gamma_n \to 0$ as $n \to \infty$ (intuition: $\{\gamma_n\}$ represent error terms)

We aimed to tackle the following questions:

1. Under which additional conditions can we prove that $\{\mu_n\}$ and $\{\beta_n\}$ converge? Are these conditions also necessary for convergence? (**real analysis**)
2. Under what circumstances can we extract rates of convergence? When not, what about rates of metastability? (**computability theory**, **proof theory**)
3. Are there new areas of application? (**applied proof theory**)
4. Can we implement some of this in a proof assistant? (**formalization**)

Overview of results

1. *Under which additional conditions can we prove that $\{\mu_n\}$ and $\{\beta_n\}$ converge? Are these also necessary for convergence?* (**real analysis**)
   - $\sum_{i=0}^{\infty} \gamma_i < \infty$. Then $\{\mu_n\}$ converges to some limit (existing result, rates of metastability already given in e.g. [Kohlenbach and Lambov, 2004]) and
   $$\beta_n \to 0 \iff \limsup_{N\to\infty} \left\{ \beta_n - \beta_m - \theta \sum_{i=n}^{m-1} \alpha_i \mid N \leq n \leq m \right\} \leq 0$$
   - $\gamma_n/\alpha_n \to 0$. Then
   $$\mu_n \to 0 \iff \forall \varepsilon > 0, n \, \exists \delta > 0 (\beta_n \leq \delta \implies \mu_n \leq \varepsilon)$$

2. *Can we always extract rates of convergence? When not, what about rates of metastability?* (**computability theory**, **proof theory**)

   Specker phenomena shown to exist unless we have strong quantitative version of our premises e.g. rate of conversion for $\sum_{i=0}^{\infty} \gamma_i < \infty$ rather than just a bound. These explain lack of rates of convergence in certain papers.

3. *Are there new areas of application?* (**applied proof theory**)

   Yes: Generalised gradient descent methods (in case $\sum_{i=0}^{\infty} \gamma_i < \infty$).

4. *Can we implement some of our lemmas in a proof assistant?* (**formalization**)

   We're developing a Lean library:
   `https://github.com/mneri123/Proof-mining-`

There is now a preprint (https://arxiv.org/abs/2207.14559):

A computational study of a class of recursive inequalities

Morenikeji Neri and Thomas Powell

July 29, 2022

Abstract

We examine the convergence properties of sequences of nonnegative real numbers that satisfy a particular class of recursive inequalities, from the perspective of proof theory and computability theory. We first establish a number of results concerning rates of convergence, setting out conditions under which computable rates are possible, and when not, providing corresponding rates of metastability. We then demonstrate how the aforementioned quantitative results can be applied to extract computational information from a range of proofs in nonlinear analysis. Here we provide both a new case study on subgradient algorithms, and survey a number of recent results which each involve an instance of our main recursive inequality. All of the relevant concepts from both proof theory and mathematical analysis are defined and motivated within the paper itself, and as such, we hope that this work also forms an accessible overview of aspects of current research in applied proof theory.

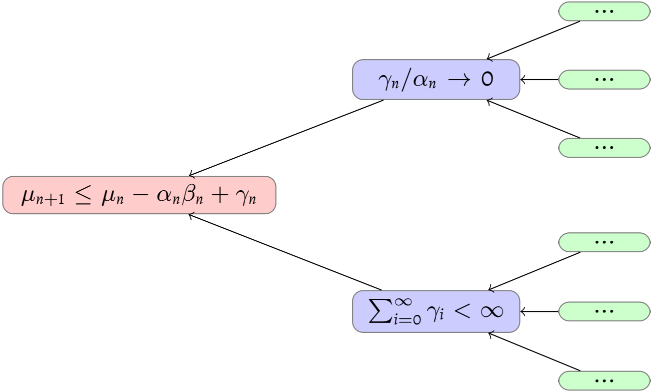
1 Introduction

Recursive inequalities on sequences of nonnegative real numbers play an important role in functional analysis. They can be used to establish convergence properties of algorithms in a very general setting, and often yield explicit *rates* of convergence in addition. A simple example of this phenomenon is represented by the inequality

$$\mu_{n+1} \leq c\mu_n \tag{1}$$

for $c \in [0, 1)$. Here, any sequence $\{\mu_n\}$ of nonnegative reals that satisfies (1) converges to zero, and moreover an effective rate of convergence for $\mu_n \to 0$ is given by $\mu_n \leq c^n \mu_0$. The inequality (1) is associated most famously with the Banach fixed point theorem: Suppose that (X, d) is a metric space and $T : X \to X$ a contractive mapping with constant c i.e.

Summary: Structure of Neri/P. 2022



While we were writing the paper:

ARTICLE IN PRESS

Annual Reviews in Control xxx (xxxx) xxx

Contents lists available at ScienceDirect

Annual Reviews in Control

journal homepage: www.elsevier.com/locate/arcontrol

Review article

Convergence of sequences: A survey ☆

Barbara Franci [a,*], Sergio Grammatico [b]

[a] *Department of Data Science and Knowledge Engineering, Maastricht University, Maastricht, The Netherlands*
[b] *Delft Center for Systems and Control, Delft University of Technology, Delft, The Netherlands*

ARTICLE INFO

Keywords:
Convergence

ABSTRACT

Convergent sequences of real numbers play a fundamental role in many different problems in system theory, e.g., in Lyapunov stability analysis, as well as in optimization theory and computational game theory. In this survey, we provide an overview of the literature on convergence theorems and their connection with Féjer monotonicity in the deterministic and stochastic settings, and we show how to exploit these results.

1. Introduction

Why Are Convergence Theorems Necessary?
The answer to this "naive" question is not simple.
cit. Boris T. Polyak, 1987 (Polyak, 1987, Section 1.6.2).

While the answer may have become clearer through the years, since many problems in applied mathematics rely on convergence theorems, it is still not simple. Besides the theoretical investigation, in fact, one fundamental aspect is how convergence theorems can be of practical use, i.e., if the assumptions are plausible for a variety of applications, for instance, in systems theory. Moreover, convergence theorems may also give qualitative information, e.g., if convergence is guaranteed for any initial point and in what sense (strongly, weakly, almost surely, in probability), which affects the range of application. The aim of this paper is to collect these results toward a complete

control in traffic networks (Duvocelle, Meier, Staudigl, & Vuong, 2019) and in modeling the prosumer behavior in smart power grids (Franci & Grammatico, 2020a; Franci et al., 2020; Kannan, Shanbhag, & Kim, 2013; Yi & Pavel, 2019).

1.1. Lyapunov decrease and Féjer monotonicity

In the mathematical literature, many convergence results hold for sequences of numbers while in system and control theory, the state and decision variables are usually *vectors* of real numbers. It is therefore important to understand the deep connection between the two theories. The bridging idea is to associate a real number to the state vector, i.e., via a function, and then prove convergence exploiting the properties of such a function. The most common example of this approach is that of Lyapunov theory where a suitable Lyapunov function is shown to be decreasing along the evolution of the state variable, thus

While we were writing the paper:

ARTICLE IN PRESS

Table 1
Convergence results for Féjer monotone sequences, deterministic sequences of real numbers and with variable metric (separated by the horizontal lines, respectively). For the applications, MI stands for Monotone Inclusion, VI for variational inequalities, NE for Nash Equilibrium problems, LYAP for Lyapunov analysis and NC for nonconvex optimization.

Result	Reference	Application	Reference
Proposition 3.1	Bauschke et al. (2011, Proposition 5.4)		
Theorem 3.2	Combettes (2001b, Theorem 3.8)		
Lemma 3.3	Opial et al. (1967) (Opial)	MI - Theorem 6.1	Malitsky and Tam (2020, Theorem 2.5)
		VI - Theorem 6.4	Malitsky (2020, Theorem 1)
Lemma 3.4	Combettes (2001b, Lemma 3.1)	NC - Theorem 6.9	Di Lorenzo and Scutari (2016, Theorem 3)
Corollary 3.5	Scutari and Sun (2019, Lemma 9)		
Lemma 3.6	Bauschke et al. (2011, Lemma 5.31)	VI - Theorem 6.4	Malitsky (2020, Theorem 1)
Corollary 3.7	Malitsky (2015, Lemma 2.8)	VI - Theorem 6.5	Malitsky (2015, Theorem 3.2)
		LYAP - Theorem 6.8	Polyak (1987, Theorem 1.4.1)
Corollary 3.8	Polyak (1987, Lemma 2.2.2)		
Lemma 3.9	Polyak (1987, Lemma 2.2.3)	NE - Theorem 6.7	Kannan and Shanbhag (2012, Theorem 2.4)
Lemma 3.10			
Lemma 3.11	Xu (2003, Lemma 2.1)		
Lemma 3.12	Extension of Xu (2002, Lemma 2.5)	NE - Theorem 6.6	Duvocelle et al. (2019, Theorem 3.1)
Corollary 3.13	Lei, Shanbhag and Chen (2020, Proposition 3)		
Corollary 3.14	Qin, Shang, and Su (2008, Lemma 1.1)		
Corollary 3.15	Xu (1998, Lemma 3)	MI - Theorem 6.3	Dadashi and Postolache (2019, Theorem 3.1)
Proposition 3.16	Alber, Iusem, and Solodov (1998, Proposition 2)		
Lemma 3.17	He and Yang (2013, Lemma 7)		
Lemma 3.18	Maingé (2008, Lemma 2.2)		
Lemma 3.19	Malitsky and Tam (2018, Lemma 2.7)	MI - Theorem 6.2	Malitsky and Tam (2020, Theorem 2.9)
Proposition 5.1	Combettes and Vũ (2013, Proposition 3.2)	MI - Theorem 8.1	Vũ (2013, Theorem 3.1)
Theorem 5.2	Combettes and Vũ (2013, Theorem 3.3)	MI - Theorem 8.1	Vũ (2013, Theorem 3.1)
Corollary 5.3	Combettes and Vũ (2013, Proposition 4.1)		

constructed sequence from such set can be analyzed anyways. On the contrary, in Lyapunov stability analysis, the target set is usually known a priori.

By exploiting the relation between the iterations and a suitable distance-like function, we show in this paper that convergence theorems represent a key ingredient for a wide variety of system-theoretic problems in fixed-point theory, game theory and optimization (Bauschke, Combettes, et al., 2011; Combettes, 2001b; Eremin & Popov, 2009; Facchinei & Pang, 2007; Polyak, 1987). In many cases, the study of iterative algorithms allows for a systematic analysis that follows from the concept of Féjer monotone sequence. The basic idea behind Féjer monotonicity is that at each step, each iterate is closer to the target set than the previous one. In a sense, the distance used for Féjer sequences can be seen as a specific class of Lyapunov function and Féjer monotonicity shows that it is decreasing along the iterates. The

1.2. What this survey is about

In this survey, we present a number of convergence theorems for sequences of real (random) numbers. We show how they can be used in combination with (quasi) Féjer monotone sequences or Lyapunov functions to obtain convergence of an iterative algorithm, essentially a discrete-time dynamical system, to a desired solution. Moreover, we present some applications to show how they can be adopted in a variety of settings. Specifically, we present convergence results for both deterministic and stochastic sequences of real numbers and we also include some results on Féjer monotone sequences and with variable metric. We show that these results help proving not only convergence of an iterative algorithm but also the Law of Large Numbers, with applications in model predictive control (Lee & Nedić, 2015) and opinion dynamics (Shi et al., 2013) among others.

We report in Tables 1 and 2 the results for deterministic and

Outline

The analysis of further recursive inequalities

A comprehensive survey paper on recursive inequalities for applied proof theory would certainly be valuable! But there are also plenty of new directions to look at.

Particularly interesting would be **stochastic algorithms**. These rely heavily on things like the *Robbins-Siegmund lemma* (which in turn relies on Martingale theory):

Lemma (Robbins-Siegmund 1971)

Let $\{\mu_n\}$, $\{\delta_n\}$, $\{\varepsilon_n\}$ *and* $\{\theta_n\}$ *be sequences of nonnegative reals such that* $\sum_{i=0}^{\infty} \varepsilon_i < \infty$, $\sum_{i=0}^{\infty} \delta_i < \infty$

$$E[\mu_{n+1} \mid \mathcal{F}_n] \leq (1 + \delta_n)\mu_n + \varepsilon_n - \theta_n \;\; a.s.$$

for some filtration $\{\mathcal{F}_n\}$. *Then* $\sum_{i=0}^{\infty} \theta_i < \infty$ *and* $\{\mu_n\}$ *converges a.s.*

- Can we give results of this kind a computational interpretation?
- Are there applications in stochastic optimization?

Formalizing applied proof theory (and nonlinear analysis!)

I'm aware of two projects on developing libraries for "proof mining":

- H. Cheval: `https://github.com/hcheval`
- M. Neri: `https://github.com/mneri123/Proof-mining-`

Building a library on convergence results for sequences of reals (along with rates of convergence/metastability) would be extremely useful:

- Many results in both areas reduce to lemmas on recursive inequalities. Formalizing these provide a solid base for more extensive formalization work.
- This would not need to rely on advanced libraries: It's enough to have the basic theory of real numbers, infinite series etc.
- Could be given to good students for projects.

Some initial progress:

```
lemma abstract_lemma1 (θ : nnseq )(α : nnseq)(K: {x: ℝ // x > 0}) (r: ℕ → {x: ℝ // x > 0} → ℕ )
(N : {x:ℝ // x > 0 }  → ℕ ) (ϕ : {x:ℝ // x > 0 }  → {x:ℝ // x > 0 })
(h1 : ∀(n:ℕ), (θ.1 n) < K) (h2: RoD r α)
(h3: ∀ ε : {x:ℝ // x > 0 } , ∀ n ≥ N(ε), (ε:ℝ) < θ.1 (n + 1)  → θ.1 (n + 1) ≤ θ.1 n - (α.1 n)*ϕ(ε)):
RoC (λ ε : {x: ℝ // x > 0},  (r (N ε) ⟨K/(ϕ ε), div_pos K.2 (ϕ ε).2 ⟩+1)) θ  :=
begin
have
H1 : ∀ ε : {x:ℝ // x > 0 },∀ n ≥ N(ε), θ.1 n ≤ ε  → θ.1 (n + 1) ≤ ε,
by_contradiction p1,
push_neg at p1,
cases p1 with ε p2,
cases p2 with n p3,
have p5 : ε < ε,
calc ε.1 < θ.1 (n+1): (p3.2).2
... ≤ θ.1 n - (α.1 n)*ϕ(ε): h3 ε n p3.1 (p3.2).2
... ≤ θ.1 n : sub_le_self (θ.1 n) (mul_nonneg (α.2 n) (le_of_lt (ϕ ε).2 ))
... ≤ ε :(p3.2).1,
exact (lt_self_iff_false ε).mp p5,
have H2 : ∀ ε : {x :ℝ // x > 0}, ∃ n ∈ finset.Ico (N ε) ((r (N ε) ⟨↑K / ↑(ϕ ε), _⟩)+1), θ.1 (n + 1) ≤
by_contradiction,
push_neg at h,
```

Automating the reduction to (quantitative) lemmas

- The reduction of e.g. $\{\|x_n - q\|\}$ to some recursive inequality (i.e. type A lemmas) often use little more than routine calculations and properties of mapping and space.
- Can we develop algorithms for automating this procedure?
- Are there new logics for reasoning about abstract spaces that would be helpful?
- This could also then automate bound extraction.

Conclusion

Three possible directions for future research that each reinforce the other:

1. The proof theoretic analysis of new recursive inequalities, particularly in the stochastic setting.
2. A formal library of lemmas on convergent sequences of real numbers.
3. Automating the reduction of concrete algorithms to recursive inequalities.

THANK YOU!

Alber, Y., Reich, S., and Shoikhet, D. (2002).
Iterative approximations of null points of uniformly accretive operatos with estimates of the convergence rate.
Communications in Applied Analysis, 6(1):89–104.

Kohlenbach, U. (2008).
Applied Proof Theory: Proof Interpretations and their Use in Mathematics.
Springer Monographs in Mathematics. Springer.

Kohlenbach, U. and Lambov, B. (2004).
Bounds on iterations of asymptotically quasi-nonexpansive mappings.
In *Proceedings of the International Conference on Fixed Point Theory and Applications*, pages 143–172. Yokohama Publishers.

Kohlenbach, U. and Powell, T. (2020).
Rates of convergence for iterative solutions of equations involving set-valued accretive operators.
Computers and Mathematics with Applications, 80:490–503.

Powell, T. and Wiesnet, F. (2021).
Rates of convergence for asymptotically weakly contractive mappings in normed spaces.
Numerical Functional Analysis and Optimization, 42(15):1802–1838.