Proof-Theoretical Aspects of Nonlinear and Set-Valued Analysis

Nicholas Pischke

Department of Mathematics
TU Darmstadt

Colloquium Logicum 2024
October 7, 2024

Proof mining in one slide

In proof mining, proof-theoretic methods are used to establish new results in core mathematics.

For this, existing proofs are analyzed to *extract* additional information.

Sometimes, such an analysis reveals superfluous premises.

In recent times, these proof-theoretic methods have also been used to produce wholly new results and notions.

Conceptually, this goes back to Georg Kreisel's program of unwinding of proofs from the 1950's, but was developed in its modern form systematically by Ulrich Kohlenbach (and his collaborators) since the 1990's.

Logical metatheorems

The core logical results substantiating proof mining are the so-called logical metatheorems.

These logical metatheorems are theorems about a corresponding logical system so that

1. the system allows for the formalization of large classes of objects and proofs in the respective area of application,
2. the metatheorem guarantees that for large classes of (ineffective) proofs carried out in the system, one can extract effective, tame and highly uniform computational information for the theorem proved thereby.

Commonly based on extended systems of arithmetic in all finite types and established using a variety of proof-theoretic devices, most notably Gödel's functional interpretation and Howard's majorizability (monotone functional interpretation).

>20 developed since the work of Kohlenbach in the 1990s in tandem with applications for various areas.

This talk

My PhD thesis, written under Ulrich Kohlenbach, is concerned with extending proof mining, both from a logical and an applied perspective, to topics from nonlinear analysis and optimization which involve set-valued operators, i.e. operators

$$A : X \to 2^Y$$

for spaces X, Y (which for the purpose of this talk will be normed vector spaces).

In the rest of the talk, I will try to give you a brief overview of one specific topic discussed in my thesis, concerned with an area of modern convex analysis called monotone operator theory.

I will try to give a high level overview of both some logical and applied aspects of that part.

In the end I briefly discuss another main area that is also treated.

As we are in the following concerned with analysis on normed vector spaces, we first need some background on proof mining for these objects.

Logical background

The basic system here is $\mathcal{A}^\omega[X, \|\cdot\|]$ for analysis over (abstract, non-separable) normed spaces (Kohlenbach 2005): arises as extension of

$$\mathcal{A}^\omega = \text{WE-PA}^\omega + \text{QF-AC} + \text{DC}$$

to an augmented set of types T^X

$$\mathbb{N}, X \in T^X, \quad \rho, \tau \in T^X \Rightarrow \rho \to \tau \in T^X$$

as well as with new constants

$$0_X, 1_X \text{ of type } X, +_X \text{ of type } X \to (X \to X), -_X \text{ of type } X \to X,$$
$$\cdot_X \text{ of type } \mathbb{N}^\mathbb{N} \to (X \to X), \|\cdot\|_X \text{ of type } X \to \mathbb{N}^\mathbb{N},$$

and axioms stating that X with these operations is a real normed vector space with 1_X such that $\|1_X\|_X =_\mathbb{R} 1$ and $-_X x$ being the additive inverse of x.

Only equality at $\mathbb{N}$ primitively. Equality at X defined by

$$x^X =_X y^X := \|x -_X y\|_X =_\mathbb{R} 0$$

using a suitable representation of reals and at higher types pointwisely.

The dual of a normed space

We main operators we want to consider are $A : X \to 2^{X^*}$ where X^* is the dual of X:

$$X^* = \{x^* : X \to \mathbb{R} \mid x^* \text{ linear and continuous}\}.$$

This becomes a normed space via setting

$$\|x^*\| = \sup_{\|x\| \leq 1} |\langle x, x^* \rangle|.$$

There are many immediate difficulties with treating this object in systems such as $\mathcal{A}^\omega[X, \|\cdot\|]$ amenable to proof mining methods:

1. The defining matrix of X^* is complex ($\geq \Pi_1^X$), making quantification over elements in X^* seen as objects of type $X \to \mathbb{N}^\mathbb{N}$ difficult and turning many essential statements about the dual into non-admissible axioms.
2. The norm of X^* involves a supremum over the (abstract, non-separable) space X and it is not clear how such a supremum can be represented in the underlying system.
3. ...

Treating the dual intensionally

The remedy, which allows for a tame treatment of these objects, is to approach the dual X^* intensionally: instead of specifying the subspace of *all* continuous and linear functionals of type $X \to \mathbb{N}^\mathbb{N}$, we introduced another base type X^* and axiomatically specify that all elements of X^*, seen as an abstract space, behave like continuous linear functionals.

For that, we need to restore the application character of elements of type X^* with an additional constant $\langle \cdot, \cdot \rangle_{X^*}$ of type $X^* \to (X \to \mathbb{N}^\mathbb{N})$ and we need to restore the linear structure on X^* with constants 0_{X^*}, 1_{X^*}, $+_{X^*}$, $-_{X^*}$ and $\cdot_{X^*}$ of suitable type and governed by suitable universal axioms (which trivialize under the functional interpretation).

The intensional approach in particular enables us to treat the associated dual norm $\|\cdot\|_{X^*}$ in the context of the monotone functional interpretation with the following norm axioms:

$$\forall {x^*}^{X^*}, x^X \left(|\langle x, x^* \rangle_{X^*}| \leq_\mathbb{R} \|x^*\|_{X^*} \|x\|_X \right),$$
$$\forall {x^*}^{X^*}, k^0 \exists x \leq_X 1_X \left(\|x^*\|_{X^*} - 2^{-k} \leq_\mathbb{R} |\langle x, x^* \rangle_{X^*}| \right).$$

Treating the dual intensionally

Crucially however: no axioms specify that this abstract space really contains *all* such functionals and instead we only have the rule

$$\frac{\begin{cases} A_0 \to \forall x^X, y^X, \alpha^1, \beta^1 \left(t(\alpha x +_X \beta y) =_{\mathbb{R}} \alpha t x + \beta t y \right) \\ A_0 \to \forall x^X \left(|tx| \leq_{\mathbb{R}} M \|x\|_X \right) \end{cases}}{A_0 \to \exists x^* \leq_{X^*} M 1_{X^*} \forall x^X \left(tx =_{\mathbb{R}} \langle x, x^* \rangle_{X^*} \right)} \qquad \text{(QF-LR)}$$

that closes this abstract space off under functionals which are provably linear and continuous.

For the resulting system $\mathcal{D}^\omega$, extending $\mathcal{A}^\omega[X, \|\cdot\|]$, a logical metatheorem can be proved using the monotone functional interpretation that allows for the extraction of computable and tame bounds from large classes of proofs that involve the dual.

Modern convex analysis: monotone operators

In particular, this system also serves as the basis for approaching a central class of set-valued operators:

Let X be a Banach space and consider so-called monotone set-valued operators $A : X \to 2^{X^*}$ (Browder 1965) which are operators that satisfy

$$\langle x - y, x^* - y^* \rangle \geq 0$$

for all $(x, x^*), (y, y^*) \in A$.

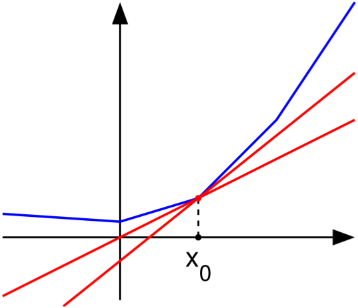
A is called maximally monotone if its graph is not strictly contained in the graph of another monotone operator.

We are interested in finding zeros of A, i.e. points x with $0 \in Ax$.

Zeros of monotone operators

Why? Consider, the canonical example of a monotone operator, the subdifferential of a convex function $h : X \to (-\infty, +\infty]$:

$$\partial h(x) := \{x^* \in X^* \mid h(x) + \langle y - x, x^* \rangle \leq h(y) \text{ for all } y \in X\}.$$



This operator is maximally monotone (Rockafellar 1966) and one in particular has $\mathrm{zer}\partial h = \min h$.

The relativized resolvent

To study such operators, one employs a special derived object, the (relativized) resolvent (Eckstein 1993):

Let A be maximally monotone. Given a (Frechét) differentiable convex function f with gradient ∇f and $\gamma > 0$: define $\mathrm{Res}^f_{\gamma A} : X \to 2^X$ with

$$\mathrm{Res}^f_{\gamma A}(x) := \left((\nabla f + \gamma A)^{-1} \circ \nabla f\right)(x).$$

Under suitable assumptions on f: single-valued and total. Also, the zeros of A are exactly the fixed points of all/any $\mathrm{Res}^f_{\gamma A}$.

All these objects can be treated in underlying systems extending $\mathcal{D}^\omega$ meanwhile allowing for metatheorems. In particular the operators can be treated intensionally via a constant χ_A of type $X \to (X^* \to \mathbb{N})$ coding its graph.

The resulting bound extraction theorems allow for various applications to convex analysis to be carried out and multiple such case studies are contained in my thesis. Let's consider one example in a bit more detail.

An exemplary application: finding zeros of monotone operators

For a maximally monotone operator A with $A^{-1}0 \neq \emptyset$, we now consider the following method for finding zeros of A (a so-called Halpern-type method):

For a given $u, x_0 \in X$, define the sequence

$$x_{n+1} = \nabla f^*(\alpha_n \nabla f u + (1 - \alpha_n) \nabla f \mathrm{Res}^f_{r_n A} x_n)$$

for scalars $\alpha_n \to 0$ as well as $\sum_{n=0}^{\infty} \alpha_n = \infty$ and $r_n \to \infty$ and where

$$f^*(x^*) = \sup_{x \in X} \left(\langle x, x^* \rangle - f(x) \right).$$

Then we get $x_n \to x \in A^{-1}0$ under suitable assumptions on f.

In particular, the iteration is asymptotically regular w.r.t. the resolvent in the sense that

$$\left\| x_n - \mathrm{Res}^f_A x_n \right\| \to 0 \quad (n \to \infty).$$

We want to extract quantitative information on this limit from the proof.

Halpern-type proximal point methods, quantitatively

Theorem

Let $b \geq \|u\|, \|x_n\|, \left\|\mathrm{Res}^f_{r_n A} x_n\right\|$. Let σ be a rate of convergence for $\alpha_n \to 0$ as $n \to \infty$. Let τ be a rate of divergence for $r_n \to \infty$ as $n \to \infty$. Let $\omega^{\nabla f}, \omega^{\nabla f^*}$ be moduli of uniform continuity on bounded sets for $\nabla f, \nabla f^*$, respectively. Let η be a modulus of uniform strict convexity for f.

Then, we can construct a mapping $\Phi_{b,\sigma,\tau,\omega^{\nabla f},\omega^{\nabla f^*},\eta}(\varepsilon)$ such that

$$\forall \varepsilon > 0 \forall n \geq \Phi_{b,\sigma,\tau,\omega^{\nabla f},\omega^{\nabla f^*},\eta}(\varepsilon) \left(\left\| x_n - \mathrm{Res}^f_A x_n \right\| < \varepsilon \right).$$

In particular, we have

$$\forall \varepsilon > 0 \forall n \geq \Phi_{b,\sigma,\tau,\omega^{\nabla f},\omega^{\nabla f^*},\eta}(\varepsilon) \exists z \in X \left(\|z\| < \varepsilon \text{ and } z \in A(\mathrm{Res}^f_A x_n) \right).$$

Monotone operators in Banach spaces, quantitatively

Theorem

The functional $\Phi_{b,\sigma,\tau,\omega^{\nabla f},\omega^{\nabla f^*},\eta}(\varepsilon)$ *can be defined as*

$$\Phi_{b,\sigma,\tau,\omega^{\nabla f},\omega^{\nabla f^*},\eta}(\varepsilon) :=$$

$$\max\left\{\varphi\left(\frac{\rho(\omega^{\nabla f}(\omega^{\nabla f^*}(\varepsilon, C(b))/2, b), b)}{4C(b)}\right), \sigma\left(\frac{\omega^{\nabla f^*}(\varepsilon, C(b)/2}{2C(b)}\right)\right.$$

$$\left.\tau\left(\frac{4C(b)b^2}{4\eta\left(\omega^{\nabla f}\left(\frac{\rho(\omega^{\nabla f}(\omega^{\nabla f^*}(\varepsilon,C(b))/2,b),b)}{8b}, b\right), b\right)}\right)\right\} + 1$$

with

$$\rho(\varepsilon, b) = 2\eta(\varepsilon, b + \varepsilon),$$

$$C(b) = \left\lceil b/\omega^{\nabla f}(1, b)\right\rceil + \|\nabla f(0)\| + 1,\ \varphi(\varepsilon) := \sigma\left(\frac{\omega^{\nabla f^*}(\varepsilon, C(b))}{2C(b)}\right).$$

Further topics from the thesis

Another major area from the thesis is concerned with so-called nonlinear semigroups and accretive operators $A : X \to 2^X$ for Banach spaces X.

These semigroups and operators are key objects in the study of differential equations: for any such operator A, the solution set of the differential equation

$$\begin{cases} u'(t) \in -Au(t),\ 0 < t < \infty \\ u(0) = x \end{cases}$$

can be described and studied through semigroup theory and many well-known differential equations can be brought into this form for a suitable operator A.

My thesis introduced the first logical systems suitable for proof mining for these objects and utilized them to carry out a selection of case studies on theorems describing the asymptotic behavior of solutions of such differential equations.

References

Proof-Theoretical Aspects of Nonlinear and Set-Valued Analysis.
PhD thesis, TU Darmstadt (2024), xvi+350pp.

Proof mining for the dual of a Banach space with extensions for uniformly Fréchet differentiable functions
Transactions of the American Mathematical Society 377(10) (2024), pp. 7475-7517

On Computational Properties of Cauchy Problems generated by Accretive Operators. w/ P. Pinto
Documenta Mathematica 28(5) (2023), pp. 1235-1274

Rates of convergence for the asymptotic behavior of second-order Cauchy problems
Journal of Mathematical Analysis and Applications 533(2) (2024), 128078, 15pp.

References

A Proof-Theoretic Metatheorem for Nonlinear Semigroups generated by an Accretive Operator and Applications
submitted, 42pp.

Effective rates for iterations involving Bregman strongly nonexpansive operators
w/ U. Kohlenbach
submitted, 42pp.

On logical aspects of extensionality and continuity for set-valued operators with applications to nonlinear analysis
submitted, 24pp.

A proof-theoretic bound extraction theorem for monotone operators in Banach spaces
submitted, 38pp.

All preprints on my website.

At last ...

Thank You!