A PROOF-THEORETIC BOUND EXTRACTION THEOREM FOR MONOTONE OPERATORS IN BANACH SPACES

NICHOLAS PISCHKE

Department of Mathematics, Technische Universität Darmstadt,
Schlossgartenstraße 7, 64289 Darmstadt, Germany,
E-mail: pischke@mathematik.tu-darmstadt.de

Abstract. We utilize a proof-theoretically tame approach to the dual of an abstract Banach space in systems amenable to methods from proof mining, as recently introduced by the author, to provide similar such systems with accompanying logical metatheorems on the extraction of uniform quantitative information from proofs pertaining to the theory of monotone set-valued operators on Banach spaces as introduced by Browder. With that, we finally extend proof mining methods to this important class of objects which are at the heart of many seminal results from nonlinear functional analysis and the metatheorems presented here in particular provide the first logical basis for a range of recent applications of proof mining methods to this branch of mathematics. Further, we provide a characterization of the extensionality principle for these operators using the analytical notion of maximality, extending previous analogous results for accretive operators on Banach spaces and monotone operators on Hilbert spaces, and with that further illustrate the central importance and special position of extensionality issues in proof mining applications dealing with set-valued operators.

Keywords: Proof mining; Metatheorems; Monotone operators; Banach spaces
MSC2010 Classification: 03F10; 47H04; 47H05; 47H09

1. Introduction

The quest for describing the computational content and strength of a mathematical theorem and developing tools to exhibit such content has been and is one of the main driving interests of proof theory. The program of proof mining, which emerged as a subfield of mathematical logic in the 1990s through the work of Ulrich Kohlenbach and his collaborators (rooting in Georg Kreisel's program of "unwinding of proofs", see in particular [51, 52]), aims at answering exactly this question and so is concerned with exhibiting the computational content of a mathematical theorem by analyzing its proofs as they are found in the mainstream mathematical literature using proof-theoretic tools. As such proofs are prima facie noneffective, e.g. involving full use of classical logic as well as many non-computational principles, this is a highly non-trivial task. However, relying on a comprehensive theoretical underpinning developed using a wide range of methods from proof theory like functional interpretations and majorizability, this research program has been very successful in a wide variety of areas of mathematics, and even more noticeably so in the area of nonlinear analysis and optimization. We refer to [36] for the comprehensive monograph on proof mining and its applications until 2008 and we refer to the survey article [47] for details on the earlier development of proof mining as well as the more recent survey articles [37, 39] for more current developments.

While the case studies of this research program are presented using the usual mathematical means of the area of application in question and without any apparent use of logical methods, they nevertheless all rely on central logical results explaining and governing these extractions.

Date: August 27, 2024.

These substrata of proof mining, often called *logical metatheorems* or *bound extraction theorems*, utilize a composition of various proof-theoretic devices like negative translations (see e.g. [54]), Kreisel's *modified realizability* [53] or Gödel's *functional (Dialectica) interpretation* [24] and their variants to provide both a logical system and a general theorem, the so-called logical metatheorem, about that system so that

(1) the system tethered to the metatheorem is suitably designed so that it allows for the formalization of large classes of objects and proofs from the core mathematical literature of the intended area of application;
(2) the associated logical metatheorem guarantees that for large classes of proofs (e.g. potentially involving a wide range of non-computational 'ideal principles' and using classical logic) carried out in this system, one can extract effective, tame and highly uniform computational information for the theorem proved thereby.

Even further, the proof of the logical metatheorems provides an algorithm to (in principle) extract this information from a given proof. Also, the complexity of this extracted information will faithfully represent the logical strength of the key principles used in the (formal) proof.

On the basis of these metatheorems, the research program of proof mining has lead to hundreds of new results in the respective areas of application over the last three decades, in particular substantiating item (1) above that the systems designed in the context of such metatheorems really are practically usable. Examples of such metatheorems can in particular be found in [21, 23, 25, 35, 36, 46, 55, 56, 65, 66, 69, 72, 73, 84].

As mentioned before, one main focus for the applications in the proof mining program in the recent years has been on proofs from nonlinear analysis and optimization that utilize so-called monotone set-valued operators in Hilbert spaces. These operators go back to the fundamental work of Minty [63, 64] and are today one of the main cornerstones of a modern operator-theoretic approach to convex analysis on Hilbert spaces as is particularly well illustrated by the seminal monograph of Bauschke and Combettes [6]. Notable examples of proof mining case studies concerned with these operators in particular are [38],[1] providing a low-complexity and highly uniform rate of asymptotic regularity in the context of Bauschke's solution [2] to the zero displacement conjecture, and the many case studies on (variants of) the influential proximal point algorithm (going back to the seminal work of Martinet [62] and Rockafellar [80]) as e.g. treated in [18, 19, 41, 42, 43, 45, 57, 58, 67], among many others. If one moves down from Hilbert to Banach spaces, the notion of a monotone operator generalizes in distinct ways based on the multiple variants of equivalently phrasing monotonicity in the context of Hilbert spaces. One such generalization leads to the notion of accretive operators which go back to the seminal work of Kato [30] and are central e.g. to the study of differential equations and semigroup theory (with a range of proof mining case studies dedicated to this class of operators already carried out, where we in particular want to mention the works [22, 44, 49, 68, 75] besides some of the papers on the proximal point algorithm mentioned before that are actually phrased in terms of these accretive operators). The other central avenue for generalizing the concept of monotonicity leads to the notion of a monotone operator in Banach spaces in the sense of Browder [14, 15] which are central e.g. to many parts of convex analysis over Banach spaces.

[1] We refer to [48] for a discussion of the logical details of this proof mining application and to [85, 86] for further extensions of that work.

While the accretive operators on Banach spaces and the monotone operators on Hilbert spaces have previously received a logical treatment in terms of associated metatheorems in the style of proof mining in [72], thereby providing the first logical footing for all of the previously mentioned case studies, the monotone operators on Banach spaces in the sense of Browder have so far remained elusive from such a treatment, which is in particular due to the heavy use made in their theory of the dual of the associated Banach space, an object that itself evaded any type of proof-theoretically tame[2] treatment in the style of proof mining until it was recently treated in [73].

It is now the aim of this paper to provide such a treatment in the style of proof mining for the monotone operators on Banach spaces in the sense of Browder by defining associated formal systems and proving a corresponding logical metatheorem, and in that way paving the way for many further applications of methods from proof mining to this central area of modern convex analysis on Banach spaces. This is achieved by carefully combining and extending the ideas of the systems presented in [72] for the treatment of monotone operators in Hilbert spaces with those presented in [73] for the treatment of the dual of a Banach space. In that way, the present paper also further elucidates the naturalness and applicability of the methods developed in the above mentioned papers to treat set-valued operators of various types as well as to treat the dual of a Banach space and that these approaches are indeed as flexible as hoped for.

Besides treating these operators, this paper also provides a proof-theoretic treatment of other central surrounding objects that are crucial for developing their theory, like in particular the so-called resolvents of such operators defined relative to a convex function f as first considered in the seminal work of Eckstein [20] in the context of finite-dimensional spaces and then highlighted as a central tool of monotone operators in general Banach spaces in particular in the work of Bauschke, Borwein and Combettes [5]. These resolvents can then in particular be used for a characterization of maximally monotone operators similar as in Minty's theorem which will be crucial for the systems devised in this paper for these operator. Further, they also are the essential objects for many central applications of these operators to convex analysis like in approaches to iterative methods used in mathematical programming. In that way, to illustrate the usability of the defined systems for these objects, we show that the main properties of the operator and the associated relativized resolvents are provable therein. Also, we show that the equivalence between maximality of the operator and extensionality, which was one of the central theoretical results established in [72] for monotone operators on Hilbert spaces and which illustrates some theoretical limitations of such systems, extends to these new objects.

The applicability of the systems defined in the present paper is then in particular confirmed from the practical perspective through the recent new case studies given in [71, 76],[3] situated in the area of convex analysis on Banach spaces. However, we expect that these metatheorems allow for many further new case studies to be carried out in the areas discussed above and for that, we want to also in particular mention the works [1, 3, 7, 13, 16, 17, 29, 50, 83,

[2]The notion of *proof-theoretic tameness* is here understood in the sense of [40], i.e. referring to the fact that although an area of mathematics could be subject to well-known Gödelian phenomena, they nevertheless "seem to be tame in the sense of allowing for the extraction of bounds of rather low complexity" (as phrased in [40]). We refer also to [59, 60] for further discussions of these types of phenomenas and their implications for logic and mathematics.

[3]We also want to note the work [70] here where the formal treatment of monotone operators on Banach spaces presented here was instrumental for introducing novel analogous notions in a nonlinear setting.

90] as promising future applications since, by inspection of the proofs, they also seem to be formalizable in (suitable extensions of) the systems introduced here.

2. Logical and mathematical preliminaries

In this section, we briefly survey both the mathematical and logical objects and results that will play a key role in developing the new formal systems of this paper.

2.1. **Monotone operators on Banach spaces and relativized resolvents.**

We begin by discussing the (very minimal) mathematical background of the theory of monotone operators in Banach spaces as well as of related objects relevant for this paper. Further definitions will be given throughout the sections as needed and we refer to the standard works [79, 81, 89] for further results and references.

2.1.1. *A primer on convex analysis.* Throughout, if not specified otherwise, let X be a Banach space with norm $\|\cdot\|$ and dual space X^* with the associated dual norm (also denoted by) $\|\cdot\|$, defined via

$$\|x^*\| := \sup_{\|x\|\leqslant 1} \langle x, x^*\rangle,$$

for $x^* \in X^*$, where we write $\langle x, x^*\rangle$ for the function value of x^* applied to $x \in X$.

Let $f : X \to (-\infty, +\infty]$ be a given function with extended real values. In the following, we will always at least assume three properties for f:

(1) f is proper, i.e.
$$\mathrm{dom} f := \{x \in X \mid f(x) < +\infty\} \neq \varnothing,$$
(2) f is lower-semicontinuous, i.e.
$$\forall x \in \mathrm{dom} f \forall y < f(x) \exists \delta > 0 \forall z \in B_\delta(x) \left(f(z) > y\right),$$
(3) f is convex, i.e.
$$\forall x, y \in \mathrm{dom} f \forall \lambda \in [0,1] \left(f\left(\lambda x + (1-\lambda)\, y\right) \leqslant \lambda f(x) + (1-\lambda)\, f(y)\right).$$

Crucial to the study of convex functions in nonlinear analysis and optimization is the use of so-called subgradients.

Definition 2.1 (Subdifferential)**.** Let $x \in \mathrm{intdom} f$, where $\mathrm{intdom} f$ denotes the interior of $\mathrm{dom} f$. We define $\partial f(x)$, the subdifferential of f at x, via

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

A subgradient of f at x, i.e. a point $x^* \in \partial f(x)$, can in many places be substituted for a gradient of the function when it is not differentiable. In this work however, the focus will be on convex functions which are actually differentiable in the following ways:

Definition 2.2 (Gâteaux and Fréchet differentiability)**.** The function f is called Gâteaux differentiable at $x \in \mathrm{intdom} f$ if there exists a $\nabla f x \in X^*$ such that

$$\langle h, \nabla f x\rangle = \lim_{t \to 0^+} \frac{f(x+th) - f(x)}{t}$$

for any $h \in X$. The function f is called Fréchet differentiable at $x \in \mathrm{intdom} f$ if there exists a $\nabla f x \in X^*$ such that

$$\lim_{\|h\| \to 0} \frac{|f(x+h) - f(x) - \langle h, \nabla f x\rangle|}{\|h\|} = 0.$$

The function f is called Fréchet (or Gâteaux) differentiable on a set $D \subseteq X$ if it is Fréchet (or Gâteaux) differentiable at every point $x \in D$ and f is called uniformly Fréchet differentiable on D if the above limit is attained uniformly in $x \in D$.

Crucial for a formal approach to such derivatives later on will be an equivalent characterization of Gâteaux and Fréchet derivatives presented in the following proposition that does not rely on the use of limits per se.

Proposition 2.3 (see e.g. [89]). *Let $x \in \mathrm{intdom} f$. Then, the following are equivalent:*

(1) f is Fréchet (respectively Gâteaux) differentiable at x.
(2) There exists a selection of ∂f that is norm-to-norm (respectively norm-to-weak) continuous at x.

Further, it holds that if f is Gâteaux differentiable at x, then $\partial f(x) = \{\nabla f(x)\}$.

Further, also the uniform Fréchet differentiability connects to a continuity property of the gradient:

Proposition 2.4 (essentially [77]). *If f is uniformly Fréchet differentiable on bounded sets and ∇f is bounded on bounded sets, then ∇f is uniformly norm-to-norm continuous on bounded sets.*

The main object for the duality theory of a convex function f in Banach spaces is the Fenchel conjugate $f^* : X^* \to (-\infty, +\infty]$, concretely defined by

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

Crucial for a formal treatment of this derived object later on will be the following result that characterizes when f^* is bounded on bounded sets.

Proposition 2.5 (see e.g. [4]). *Call f supercoercive if*

$$\lim_{\|x\| \to +\infty} \frac{f(x)}{\|x\|} = +\infty.$$

Then, the following are equivalent:

(1) f is supercoercive.
(2) f^* is bounded on bounded subsets.

In particular, both imply that $\mathrm{dom} f^* = X^*$.

We now introduce the main class of functions considered in this paper, first in the way as it is commonly introduced in the convex analysis literature:

Definition 2.6 ([4]). A function f is called:

(1) essentially smooth if ∂f is locally bounded and single-valued on its domain,
(2) essentially strictly convex if $(\partial f)^{-1}$ is locally bounded and f is strictly convex on every convex subset of $\mathrm{dom} \partial f$,
(3) Legendre if it is both essentially smooth and essentially strictly convex.

Over reflexive spaces, the class of Legendre functions can be equivalently recognized via a particularly nice differentiability property for both f and its conjugate f^* which serves as the central definition of this class of functions in this paper.

Proposition 2.7 ([4]). *If X is reflexive, then f is Legendre if, and only if both*

(1) It holds that $\mathrm{intdom} f \neq \emptyset$, *that f is Gâteaux differentiable on* $\mathrm{intdom} f$, *and* $\mathrm{dom} \nabla f = \mathrm{intdom} f$.

(2) *It holds that* $\mathrm{intdom}f^* \neq \emptyset$, *that* f^* *is Gâteaux differentiable on* $\mathrm{intdom}f^*$, *and* $\mathrm{dom}\nabla f^* = \mathrm{intdom}f^*$.

Also crucial for the study of such functions and related objects, in particular regarding analytical aspects of associated iterations, is the notion of the Bregman distance $D_f : \mathrm{dom}f \times \mathrm{intdom}f \to [0, +\infty)$ as introduced in the seminal work [13] and which is defined by

$$D_f(x, y) := f(x) - f(y) - \langle x - y, \nabla f(y)\rangle.$$

Dual to this distance notion is the function $W_f : \mathrm{dom}f \times \mathrm{dom}f^* \to [0, +\infty)$ defined by

$$W_f(x, x^*) := f(x) - \langle x, x^*\rangle + f^*(x^*).$$

For this function, we will rely later on a few further properties: If f is Legendre and if X is reflexive, one in particular has that

$$W_f(x, \nabla f(y)) = D_f(x, y)$$

for all $x \in \mathrm{dom}f$ and $y \in \mathrm{intdom}f$ as well as that W_f is convex in its right argument and satisfies the inequality

$$W_f(x, x^*) \leqslant W_f(x, x^* + y^*) - \langle \nabla f^*(x^*) - x, y^*\rangle$$

for any $x \in \mathrm{dom}f$ and any $x^*, y^* \in \mathrm{dom}f^*$ (see [61]).

2.1.2. *Monotone operators, maximality and relativized resolvents.* We now introduce one of the two main objects of concern for this paper, the so-called monotone set-valued operators. Operator-theoretic monotonicity arose in the 1960's with the work of Minty [63, 64] as a crucial concept in the context of convex analysis on Hilbert spaces. Minty's notion was subsequently extended to Banach space by Browder [14, 15], where the use of the inner product in the characterizing condition was replaced by the use of functionals from the dual space, and it has since then become one the prime notions in convex analysis on Banach spaces (recall the discussion in the introduction).

Definition 2.8 ([14, 15])**.** Let $A : X \to 2^{X^*}$ be a set-valued operator. The operator A is called monotone if

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

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

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

Monotone operators are generally studied through the use of a certain derived object called the resolvent which in the case of a set-valued monotone operator $A : X \to 2^X$ on a Hilbert space X commonly takes the form of

$$\mathrm{Res}_A := (\mathrm{Id} + A)^{-1}.$$

In the context of monotone operators $A : X \to 2^{X^*}$ on Banach spaces X, where the image space makes crucial use of the dual X^* of X, one is lead to replacing the use of the identity in the above definition by a suitable function mapping from X to X^*. In the context of smooth Banach spaces, such a function is often taken to be the (in this context single-valued) normalized duality map

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

with the resolvent then defined as

$$\mathrm{Res}_A := (J + A)^{-1} \circ J.$$

Abstracting from this concrete choice, one can consider a notion where J is replaced by the gradient ∇f of a (mostly) general convex function f. This even has impact already in finite-dimensional spaces where such resolvents where originally considered by Eckstein [20] as the gradient of a suitably chosen function f may ease the computations required to evaluate the resolvent. Subsequently, these relativized resolvents were also studied by Bauschke, Borwein and Combettes [5] under the name of "D-resolvents" and Reich and Sabach [78] under the name of "resolvents of A relative to f".

Concretely, the f-resolvents of A (as we will also call them in this paper) are now defined as follows.

Definition 2.9 ([5, 20])**.** Let $A : X \to 2^{X^*}$ be a set-valued operator. If f is Gâteaux differentiable on $\mathrm{intdom} f$, the f-resolvent $\mathrm{Res}^f_A : X \to 2^X$ of A is the operator defined by

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

for $x \in \mathrm{intdom} f$ and by $\mathrm{Res}^f_A(x) := \varnothing$ otherwise.

The following properties are essential for the resolvent relative to f:

Proposition 2.10 ([5])**.** *Let f be Gâteaux differentiable and strictly convex on* $\mathrm{intdom} f$ *and let A be a monotone operator such that* $\mathrm{intdom} f \cap \mathrm{dom} A \neq \varnothing$*. Then following statements hold:*

(1) $\mathrm{dom}(\mathrm{Res}^f_A) \subseteq \mathrm{intdom} f$ *and* $\mathrm{ran}(\mathrm{Res}^f_A) \subseteq \mathrm{intdom} f$*,*
(2) Res^f_A *is single-valued on its domain,*
(3) $\mathrm{Fix}(\mathrm{Res}^f_A) = \mathrm{intdom} f \cap A^{-1}0$ *where* $\mathrm{Fix}(\mathrm{Res}^f_A)$ *is the set of fixed points of* Res^f_A*,*
(4) Res^f_A *is Bregman firmly nonexpansive on its domain, i.e. for any $x, y \in$* $\mathrm{dom}(\mathrm{Res}^f_A)$*:*

$$\langle \mathrm{Res}^f_A x - \mathrm{Res}^f_A y, \nabla f \mathrm{Res}^f_A x - \nabla f \mathrm{Res}^f_A y \rangle \leqslant \langle \mathrm{Res}^f_A x - \mathrm{Res}^f_A y, \nabla f x - \nabla f y \rangle.$$

Further, the classical result for monotone operators in Hilbert spaces established by Minty [64] that maximal monotonicity is equivalent to the totality of the resolvents extends to these resolvents relative to f under suitable assumptions on f:

Proposition 2.11 ([8])**.** *Let X be reflexive. Let A be monotone and assume that $f : X \to \mathbb{R}$ is Gâteaux differentiable, strictly convex and cofinite (i.e.* $\mathrm{dom}\ f^* = X^*$*). Then A is maximally monotone if and only if* $\mathrm{ran}(A + \nabla f) = X^*$*.*

Important for the study of resolvents are their corresponding Yosida approximates defined by

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

for a given $\gamma > 0$.

It follows essentially by the definitions of $\mathrm{Res}^f_{\gamma A}$ and A^f_γ (see e.g. [78]) that

$$(\mathrm{Res}^f_{\gamma A} x, A^f_\gamma x) \in A$$

for any $\gamma > 0$ and any $x \in \mathrm{dom}(\mathrm{Res}^f_{\gamma A})$.

2.2. **A formal system for a normed space and its dual as well as uniformly Frechét differentiable functions and conjugates.** We now move to the logical preliminaries and in this subsection discuss the central systems for proof mining relevant for the paper. The central system for all of these considerations will be the system $\mathcal{D}^\omega$ as introduced in [73] that allows for a simultaneous treatment of a normed space and its dual, as well as its extensions as also discussed in [73] which allow for the treatment of a convex function together with its conjugate and the associated Fréchet derivatives. In the following, we therefore sketch the key aspects of these systems together with some of the other underlying formal background.

2.2.1. *A system for classical analysis in all finite types.* At first, as is the case for essentially all other modern systems of proof mining, also the system $\mathcal{D}^\omega$ and its extensions rely on a suitably strong system $\mathcal{A}^\omega = \text{WE-PA}^\omega + \text{QF-AC} + \text{DC}$ for classical analysis in all finite types. In the following, we first give a brief overview of this system and its main features relevant to this paper but otherwise refer to [23, 35, 36, 88] for any further details. Further, we in particular follow [36] in regards to our conventions for denoting types.

At first, the underlying set of types T is defined recursively via

$$0 \in T, \quad \xi, \tau \in T \Rightarrow \tau(\xi) \in T$$

and the finite types are stratified by their degrees defined by

$$\deg(0) := 0, \quad \deg(\tau(\xi)) := \max\{\deg(\tau), \deg(\xi) + 1\}.$$

Pure types, i.e. types τ that either are 0 or are of the form $\tau = 0(\rho)$ where ρ is another pure type, are abbreviated as usual by natural numbers via recursively defining

$$n + 1 := 0(n).$$

The language of the finite type systems $\text{WE-PA}^\omega/\mathcal{A}^\omega$ is then a many-sorted language, with variables and quantifiers for all finite types, that is extended with constants 0 for zero, S for successor, $\Sigma_{\xi,\tau}, \Pi_{\delta,\xi,\tau}$ for the so-called *combinators* (which originate from Schönfinkel [82] as well as Curry and Howard, see e.g. [28] for the latter) as well as constants $\underline{R}_{\underline{\xi}} = (R_1)_{\underline{\xi}}, \dots, (R_k)_{\underline{\xi}}$ for simultaneous primitive recursion in the sense of Gödel [24] and Hilbert [26] (see also [36]) for tuples of types $\underline{\xi}$.

The only relation symbol is $=_0$ for equality at type 0 and new terms are formed only via application: if t is a term of type $\tau(\xi)$ and s a term of type ξ, then $t(s)$ is a term of type τ. Higher type equality is treated as a defined notion via

$$s =_\xi t := \forall y_1^{\xi_1}, \dots, y_k^{\xi_k} (s y_1 \dots y_k =_0 t y_1 \dots y_k)$$

for terms s, t of type $\xi = 0(\xi_k) \dots (\xi_1)$.

The system WE-PA^ω then arises from a usual finite-type variant of Peano arithmetic (see [36, 88]) together with only the following weak *rule of quantifier-free extensionality*

$$\text{(QF-ER)} \qquad \frac{F_{qf} \to s =_\xi t}{F_{qf} \to r[s/x^\xi] =_\tau r[t/x^\xi]}$$

where F_{qf} is a quantifier-free *formula*, s, t are terms of type ξ, r is a term of type τ and $r[s/x^\xi]$ denotes the simultaneous substitution of s for all occurrences of x in r (and similarly with t).

Remark 2.12. As F_{qf} in the formulation can actually be a formula with free variables, the rule (QF-ER) actually allows one to derive the seemingly stronger rule

$$(\Sigma_1\text{-ER}) \qquad \frac{\exists y^\sigma F_{qf}(y) \to s =_\xi t}{\exists y^\sigma F_{qf}(y) \to r[s/x^\xi] =_\tau r[t/x^\xi]}$$

with F_{qf}, s, t, ξ, τ as before and where σ is an additional finite type and y is a variable of that type that is not free in r, s, t (see also [72]).

As is well-known, the combinators can be used to define λ-abstractions in the sense that for any term t of type τ and any variable x^ξ of type ξ, there is a term $\lambda x^\xi.t$ of type $\tau(\xi)$ such that provably

$$(\lambda x^\xi.t)(s^\xi) =_\tau t[s/x].$$

The system $\mathcal{A}^\omega$ then extends WE-PA$^\omega$ by the quantifier-free version of the axiom of choice in finite types, i.e.

$$(\text{QF-AC}) \qquad \forall \underline{x} \exists \underline{y} F_{qf}(\underline{x}, \underline{y}) \to \exists \underline{Y} \forall \underline{x} F_{qf}(\underline{x}, \underline{Y}\,\underline{x})$$

where F_{qf} is quantifier-free and the $\underline{x}, \underline{y}$ may be of arbitrary type, and the axiom of dependent choice $\mathrm{DC} = \{\mathrm{DC}^{\underline{\xi}} \mid \underline{\xi} \subseteq T\}$ where

$$(\mathrm{DC}^{\underline{\xi}}) \qquad \forall x^0, \underline{y}^{\underline{\xi}} \exists \underline{z}^{\underline{\xi}} F(x, \underline{y}, \underline{z}) \to \exists \underline{f}^{\underline{\xi}(0)} \forall x^0 F(x, \underline{f}(x), \underline{f}(S(x)))$$

and where F is now of arbitrary complexity.

In the language of WE-PA$^\omega/\mathcal{A}^\omega$, rational and real numbers are represented as usual by objects of type 0 and 1, respectively, where we follow the conventions of [36] and here only briefly present the facts crucial to the presentation of the present paper. For one, it will be convenient to fix a pairing function (for the coding of rationals as pairs of natural numbers) which we do, following the conventions of [36], by setting

$$j(n^0, m^0) := \begin{cases} \min u \leqslant_0 (n+m)^2 + 3n + m[2u =_0 (n+m)^2 + 3n + m] & \text{if existent,} \\ 0^0 & \text{otherwise.} \end{cases}$$

Using such codes, the operations $+_\mathbb{Q}, \cdot_\mathbb{Q}, (\cdot)^{-1}_\mathbb{Q}$ and relations $=_\mathbb{Q}$, $<_\mathbb{Q}$ are easily definable, with the latter being given by quantifier-free formulas. Reals are then coded via fast converging Cauchy sequences of rational numbers with a fixed convergence rate 2^{-n} (see [36]). On the level of that representation, one can then similarly define Π^0_1/Σ^0_1-formulas $=_\mathbb{R}/<_\mathbb{R}$ on type 1 objects which define the corresponding relations of the real numbers represented by the inputs. Further, one also easily defines closed terms $+_\mathbb{R}, \cdot_\mathbb{R}, |\cdot|_\mathbb{R}$ representing the usual operations of real arithmetic on these type 1 objects. This also holds true to some degree for the reciprocal $(\cdot)^{-1}$ on the reals, which crucially features in the later theory of monotone operators but which however is slightly more delicate to deal with as there in fact is no closed term of type 1(1) in WE-PA$^\omega$ which represents r^{-1} correctly for all $r \neq 0$. Following [34], we handle this by using a binary term $(\cdot)^{-1}_{\cdot}$ of type 1(1)(0) such that $(r)^{-1}_l$ correctly represents r^{-1} whenever $|r| > 2^{-l}$. Note that extensionality of the operations $+_\mathbb{R}$, $\cdot_\mathbb{R}$ and $|\cdot|_\mathbb{R}$ w.r.t. $=_\mathbb{R}$ can be proved in (weak fragments of) WE-PA$^\omega$ and in the case of $(\cdot)^{-1}$, extensionality can be shown for all l and r, s such that $|r|_\mathbb{R}, |s|_\mathbb{R} >_\mathbb{R} 2^{-l}$ (see [34]). Lastly, any natural n or rational q can be easily seen as a real defined just via the constant n- or q-sequence and we write n or q, respectively, for that type 1 representation as well. In the following, we omit the subscripts of the arithmetical operations for $\mathbb{R}$ to avoid notational overload.

In the context of the bound extraction theorems established later, we associate a canonical type 1 representation $(r)_\circ$ with any real $r \in \mathbb{R}$.[4] For this, we essentially follow the approach of [35] (see also [23, 36]) where such an association is presented for non-negative reals r. The definition given here is an immediate extension of that approach to the negative reals.

Concretely, we introduce $(r)_\circ \in \mathbb{N}^\mathbb{N}$ as

$$(r)_\circ(n) := j(2k_0, 2^{n+1} - 1),$$

where

$$k_0 := \max k \left[\frac{k}{2^{n+1}} \leqslant r \right]$$

for $r \geqslant 0$ and we set

$$(r)_\circ(n) = j(2\bar{k}_0 \dot{-} 1, 2^{n+1} - 1)$$

where $n \dot{-} m := \max\{n - m, 0\}$ and

$$\bar{k}_0 := \max k \left[\frac{k}{2^{n+1}} \leqslant |r| \right]$$

when $r < 0$. In that latter case, it rather immediately follows that $(r)_\circ(n) = -_\mathbb{Q}(|r|)_\circ(n)$ and this allows us to derive the following lemma containing exactly the properties that we later need for this notion to be useful in the context of majorizability (trivially extending Lemma 2.10 from [35] to the full reals):

Lemma 2.13 (see e.g. [73]). *Let $r \in \mathbb{R}$. Then:*

(1) $(r)_\circ$ is a representation of r in the sense of the above (see again e.g. [36]*).*
(2) For $s \in [0, \infty)$, if $|r| \leqslant s$, then $(r)_\circ \leqslant_1 (s)_\circ$.
(3) $(r)_\circ$ is nondecreasing (as a type 1 function).

Lastly, in the context of the metatheorems, we will write r_α for the real represented by some type 1 functional α similar as in [36].

2.2.2. *A system for abstract normed spaces.* In this section, we briefly discuss the extension $\mathcal{A}^\omega[X, \|\cdot\|]$ of $\mathcal{A}^\omega$ as introduced in [23, 35] which provides a suitable system for the treatment of abstract normed spaces, extended later to treat corresponding dual spaces and notions from convex analysis. All these approaches crucially utilize the seminal paradigm of abstract types introduced in [35] that is prevalent in all modern approaches to proof mining. Namely, while the first logical metatheorems in proof mining relied on bare systems of arithmetic in all finite types and consequently only covered applications involving representable Polish metric spaces, these additional abstract types allow for the treatment of spaces which are not separable and thus not representable in the (pure) language of finite type arithmetic.

As such, the system $\mathcal{A}^\omega[X, \|\cdot\|]$ operates over an extended set of types T^X defined by

$$0, X \in T^X, \quad \xi, \tau \in T^X \Rightarrow \tau(\xi) \in T^X$$

where the additional abstract type X can then be utilized to represent an abstract and generic (in general non-separable) normed space. For that, we add additional constants to the resulting extended language to induce a linear and normed structure on X. Concretely, we add the constants $0_X, 1_X$ of type X, $+_X$ of type $X(X)(X)$, $-_X$ of type $X(X)$, $\cdot_X$ of type $X(X)(1)$

[4]Such an association will be non-effective but it will suffice for all intents and purposes that it behaves nice enough w.r.t. majorization.

and $\|\cdot\|_X$ of type $1(X)$. It should be noted that $=_0$ is still the only primitive relation and in particular, identity on X is treated as a defined predicate via

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

with equality at higher types defined similar to before. The theory $\mathcal{A}^\omega[X, \|\cdot\|]$ then arises from $\mathcal{A}^\omega$ by formulating the latter over the resulting extended language by extending the constants (if appropriate) to take arguments and produce values in those new types and by trivially extending the axiom schemes and rules to allow formulas from the new language and then adding the above new constants related to X together with the relevant defining 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 (see [23, 35, 36] for further details on all of this). Note that the extensionality of all those operations is provable in $\mathcal{A}^\omega[X, \|\cdot\|]$.

2.2.3. *Systems for dual spaces, uniformly Fréchet differentiable functions and conjugates.* The system $\mathcal{D}^\omega$ introduced in [73] now extends the system $\mathcal{A}^\omega[X, \|\cdot\|]$ with an additional abstract type X^* used to intensionally and abstractly specify the dual of the space represented by X. We refer to [73] for a discussion of the underlying intuition and here only briefly present the main technical aspects of the theory as they are relevant to the present paper. At first, the system $\mathcal{D}^\omega$ operates over a further extended set of types T^{X,X^*} defined by

$$0, X, X^* \in T^{X,X^*}, \quad \xi, \tau \in T^{X,X^*} \Rightarrow \tau(\xi) \in T^{X,X^*}$$

and the resulting extended finite-type language is further enriched by the following selection of constants restoring the linear and normed structure of X^* similar as with X: Concretely, using this new type, we add constants $+_{X^*}$ of type $X^*(X^*)(X^*)$, $-_{X^*}$ of type $X^*(X^*)$, $\cdot_{X^*}$ of type $X^*(X^*)(1)$, 0_{X^*} and 1_{X^*} of type X^* and $\|\cdot\|_{X^*}$ of type $1(X^*)$ as before and, crucially, another constant $\langle\cdot,\cdot\rangle_{X^*}$ of type $1(X)(X^*)$ which serves as an abstract application functional, restoring the application character of X^* relative to X (similar as in the theory of topological vector spaces, see [73] for a further discussion). We do not add further relations and so, still, equality is a defined notion at types other than 0, with $=_X$ defined as before, $=_{X^*}$ defined via

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

and the higher-type equality defined similar to before. The system $\mathcal{D}^\omega$ then results from $\mathcal{A}^\omega[X, \|\cdot\|]$ as formulated over that extended language by adding the following additional axioms and rules: The first crucial group of axioms that we consider is the following pair

$$(*)_{1,2} \qquad \begin{cases} \forall x^{*X^*}, x^X \left(|\langle x, x^*\rangle_{X^*}| \leqslant_{\mathbb{R}} \|x^*\|_{X^*} \|x\|_X\right), \\ \forall x^{*X^*}, k^0 \exists x \leqslant_X 1_X \left(\|x^*\|_{X^*} - 2^{-k} \leqslant_{\mathbb{R}} |\langle x, x^*\rangle_{X^*}|\right), \end{cases}$$

which specify the norm of X^* as coded by $\|\cdot\|_{X^*}$ to be the actual supremum of $\langle x, x^*\rangle_{X^*}$ over the unit ball.[5]

[5]These axioms actually arise through a general approach to treating suprema over certain bounded sets as introduced in [73] and we refer to that paper for a further discussion on the particularities and practicalities of these axioms.

For a second crucial group of axioms, we further specify the application functional to be a bilinear map:

$$(*)_3 \qquad \begin{cases} \forall x^X, x^{*X^*}, y^{*X^*}, \alpha^1, \beta^1 \left(\langle x, \alpha x^* +_{X^*} \beta y^* \rangle_{X^*} =_{\mathbb{R}} \alpha \langle x, x^* \rangle_{X^*} + \beta \langle x, y^* \rangle_{X^*} \right), \\ \forall x^X, x^{*X^*}, y^{*X^*}, \alpha^1, \beta^1 \left(\langle x, \alpha x^* -_{X^*} \beta y^* \rangle_{X^*} =_{\mathbb{R}} \alpha \langle x, x^* \rangle_{X^*} - \beta \langle x, y^* \rangle_{X^*} \right), \end{cases}$$

$$(*)_4 \qquad \begin{cases} \forall x^X, y^X, x^{*X^*}, \alpha^1, \beta^1 \left(\langle \alpha x +_X \beta y, x^* \rangle_{X^*} =_{\mathbb{R}} \alpha \langle x, x^* \rangle_{X^*} + \beta \langle y, x^* \rangle_{X^*} \right), \\ \forall x^X, y^X, x^{*X^*}, \alpha^1, \beta^1 \left(\langle \alpha x -_X \beta y, x^* \rangle_{X^*} =_{\mathbb{R}} \alpha \langle x, x^* \rangle_{X^*} - \beta \langle y, x^* \rangle_{X^*} \right). \end{cases}$$

Also, as a third group of axioms $(*)_5$, we add the vector space axioms for X^* formulated with the operations $+_{X^*}$, $-_{X^*}$, $\cdot_{X^*}$, 0_{X^*}, 1_{X^*} w.r.t. $=_{X^*}$ and we also axiomatically specify $\|1_{X^*}\|_{X^*} =_{\mathbb{R}} 1$.

The last additions to the system concern the relationship between the bounded linear functionals represented in $1(X)$ (where elements of the dual of X, if naively specified, live) and the elements of X^*. While it will of course not be permissible meanwhile aiming for bound extraction theorems to include an axiom that guarantees the existence of a representing element in X^* for every element of $1(X)$ which is a continuous linear functional (which is why the dual X^* was intensionally approached in the first place, see [73] for a further discussion), we resort to the next best thing available in this situation by including a rule guaranteeing that at least all terms of type $1(X)$ which provably are continuous linear functionals indeed are represented by a corresponding element of type X^* and thus consider the rule

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

where F_{qf} is a quantifier-free formula and where t and M are terms of type $1(X)$ and 1, respectively. Also, to ease formal development of the theory of X^*, we further axiomatically populate X^* by adding the crucial consequence of the Hahn-Banach theorem that the normalized duality map J discussed previously has non-empty values, i.e. that $J(x) \neq \emptyset$ for $x \in X$, which we achieve by the following axiom:

$$(*)_6 \qquad \forall x^X \exists x^* \leqslant_{X^*} \|x\|_X 1_{X^*} \left(\langle x, x^* \rangle_{X^*} =_{\mathbb{R}} \|x\|_X^2 =_{\mathbb{R}} \|x^*\|_{X^*}^2 \right).$$

The system $\mathcal{D}^\omega$ for the abstract dual space of an abstract normed space is now officially defined as the extension of $\mathcal{A}^\omega[X, \|\cdot\|]$, formulated over the extended language using the types T^{X,X^*}, by the above constants, axioms and rules.

Various essential properties of the space X^* are immediately provable in the system (including simple but crucial aspects like that $\|\cdot\|_{X^*}$ is indeed a norm) and we refer to [73] for further details.

Over this system, convex functions are then introduced as objects of type $1(X)$ and we in particular consider a system where a constant f of type $1(X)$ is added to the language which is axiomatically specified to be convex via the axiom[6]

$$(f)_1 \qquad \forall x^X, y^X, \lambda^1 \left(f\left(\widetilde{\lambda} x +_X \left(1 - \widetilde{\lambda}\right) y \right) \leqslant_{\mathbb{R}} \widetilde{\lambda} f(x) + \left(1 - \widetilde{\lambda}\right) f(y) \right).$$

In that way, for the system which we are in the process of specifying, and throughout this paper for that matter, we will only consider convex functions that are total, i.e. where $\mathrm{dom} f = X$. We refer to [73] for some discussions on how non-total convex functions might be approached

[6]Here, $\widetilde{\cdot}$ is defined as e.g. in [36], allowing for implicit quantification over $[0, 1]$.

formally.

Gradients of uniformly Fréchet differentiable functions are then introduced via introducing suitably continuous selection functions of the subgradient, utilizing the characterization result presented in Propositions 2.3 and 2.4. Concretely, we in that way treat a uniformly Fréchet differentiable f with a (total) gradient by adding another constant ∇f of type $X^*(X)$ to the system together with the axioms that ∇f is a selection of ∂f, i.e.

$$(\nabla f)_1 \qquad \forall x^X, y^X \left(f(x) + \langle y -_X x, \nabla f(x)\rangle_{X^*} \leqslant_{\mathbb{R}} f(y) \right),$$

and that ∇f is uniformly continuous on bounded subsets, i.e.

$$(\nabla f)_2 \qquad \forall x^X, y^X, b^0, k^0 \Big(\|x\|_X, \|y\|_X <_{\mathbb{R}} b \wedge \|x -_X y\|_X <_{\mathbb{R}} 2^{-\omega^{\nabla f}(k,b)} \to \|\nabla f(x) -_{X^*} \nabla f(y)\|_{X^*} \leqslant_{\mathbb{R}} 2^{-k} \Big),$$

where $\omega^{\nabla f}$ is another additional constant of type $0(0)(0)$ representing the modulus of uniform continuity.

We denote the theory resulting from $\mathcal{D}^\omega$ by extending it with the previous constants and axioms by $\mathcal{D}^\omega[f, \nabla f]$.

Crucially, we now also want to treat the associated Fenchel conjugate f^* and if this map is to be treated in any way amenable to the metatheorems, it has to be majorizable which means that it in particular has be bounded on bounded sets. By the result collected in Proposition 2.5, this is exactly the case if, and only if, f is supercoercive in which case f^* is total. In that case however, as shown in [73], if f is supercoercive, then the supremum defining it is already attained over a suitably large ball around on the origin and that corresponding radius can be computed in terms of a given modulus witnessing the supercoercivity of f quantitatively:

Lemma 2.14 (essentially [73])**.** *Let $\alpha : \mathbb{N} \to \mathbb{N}$ be a modulus of supercoercivity, i.e.*

$$\forall K \in \mathbb{N}, x \in X \left(\|x\| > \alpha(K) \to f(x)/\|x\| \geqslant K \right)$$

and let $c \geqslant |f(0)|$. Then for $x^* \in X^*$ with $\|x^*\| \leqslant b$, we have

$$f^*(x^*) = \sup_{x \in \overline{B}_{r(\alpha,c,b)}(0)} \left(\langle x, x^* \rangle - f(x) \right)$$

where $r(\alpha, c, b) = \max\{\alpha(b+1)+1, c+1\}$.

This in turn allows one to utilize the abstract approach developed in [73] on the formal treatment of suprema over certain bounded sets in systems amenable for proof mining to introduce corresponding axioms that, given another constant f^* of type $1(X^*)$, specify that this constant is really the conjugate corresponding to f (we again refer to [73] for a further discussion of this): The first axiom specifies that f supercoercive with modulus α^f via

$$(f)_2 \qquad \forall K^0, x^X \left(\|x\|_X >_{\mathbb{R}} \alpha^f(K) \to f(x)/\|x\|_X \geqslant_{\mathbb{R}} K \right)$$

where α^f is an additional constant of type 1. The two further axioms specify that f^* is actually given by the respective supremum which is achieved by first specifying that f^* is a pointwise upper bound for all affine functionals $g_x(x^*) = \langle x, x^* \rangle - f(x)$ via

$$(f^*)_1 \qquad \forall x^{*X^*}, x^X \left(\langle x, x^* \rangle_{X^*} - f(x) \leqslant_{\mathbb{R}} f^*(x^*) \right)$$

and another axioms specifying that f^* is indeed the pointwise supremum of all these affine functionals via

$$(f^*)_2 \qquad \forall x^{*X^*}, b^0, k^0 \exists x^X \leqslant_X \max\{\alpha^f(b+1)+1, [|f(0)|](0)+2\}1_X$$
$$\left(\|x^*\|_{X^*} <_{\mathbb{R}} b \to \left(f^*(x^*) - 2^{-k} \leqslant_{\mathbb{R}} \langle x, x^* \rangle_{X^*} - f(x)\right)\right).$$

Here, $\max\{\alpha^f(b+1)+1, [|f(0)|](0)+2\}$ immediately arises as a bound restricting the defining supremum of f^* via Lemma 2.14 since $[|f(0)|](0)+1 \geqslant |f(0)|$.

If f^* is uniformly Fréchet differentiable on bounded sets as well, its gradient can now be introduced as before by adding a constant ∇f^* of type $X(X^*)$ together with the axioms that ∇f^* is a selection of ∂f^*, i.e.

$$(\nabla f^*)_1 \qquad \forall x^{*X^*}, y^{*X^*} \left(f^*(x^*) + \langle \nabla f^*(x^*), y^* -_{X^*} x^* \rangle_{X^*} \leqslant_{\mathbb{R}} f^*(y^*)\right),$$

and that ∇f^* is uniformly continuous on bounded subsets, i.e.

$$(\nabla f^*)_2 \qquad \forall x^{*X^*}, y^{*X^*}, b^0, k^0 \Big(\|x^*\|_{X^*}, \|y^*\|_{X^*} <_{\mathbb{R}} b \wedge \|x^* -_{X^*} y^*\|_{X^*} <_{\mathbb{R}} 2^{-\omega^{\nabla f^*}(k,b)}$$
$$\to \|\nabla f^*(x^*) -_X \nabla f^*(y^*)\|_X \leqslant_{\mathbb{R}} 2^{-k} \Big),$$

where $\omega^{\nabla f^*}$ is another additional constant of type $0(0)(0)$, coding the respective modulus of uniform continuity.

As outlined in the previous section, in the context of the above axioms whereby the functions f and f^* are Fréchet differentiable, the function f is a Legendre function and thus the gradients ∇f and ∇f^* are inverses of each other. Correspondingly, we add this fact as a last axiom

$$(\mathrm{L}) \qquad \forall x^X, x^{*X^*} \left(\nabla f \nabla f^*(x^*) =_{X^*} x^* \wedge \nabla f^* \nabla f(x) =_X x\right).$$

We write $\mathcal{D}^\omega[f, \nabla f, f^*, \nabla f^*]$ for the system $\mathcal{D}^\omega[f, \nabla f]$ extended with the above constants and axioms regarding the function f^* and its gradient ∇f^*.

We finally just quote a few properties that are immediately provable in this system:

Lemma 2.15 ([73]). *The theory $\mathcal{D}^\omega[f, \nabla f]$ proves:*

(1) f is uniformly Fréchet differentiable on bounded subsets, i.e.

$$\forall b^0, k^0 \exists j^0 \forall x^X, y^X \Bigg(\|x\|_X <_{\mathbb{R}} b \wedge 0 <_{\mathbb{R}} \|y\|_X <_{\mathbb{R}} 2^{-j}$$
$$\to \frac{|f(x+y) - f(x) - \langle y, \nabla f(x) \rangle_{X^*}|}{\|y\|_X} \leqslant_{\mathbb{R}} 2^{-k} \Bigg),$$

where in fact one can choose

$$j = \omega^{\nabla f}(k, b+1).$$

(2) ∇f is bounded on bounded subsets, i.e.

$$\forall b^0 \exists c^0 \forall x^X \left(\|x\|_X <_{\mathbb{R}} b \to \|\nabla f(x)\|_{X^*} \leqslant_{\mathbb{R}} c\right),$$

where in fact one can choose

$$c = C(b) = b2^{\omega^{\nabla f}(0,b)} + [\|\nabla f(0)\|_{X^*}](0) + 2.$$

(3) f is uniformly continuous on bounded subsets, i.e.

$$\forall k^0, b^0 \exists j^0 \forall x^X, y^X \left(\|x\|_X, \|y\|_X <_{\mathbb{R}} b \wedge \|x -_X y\|_X \leqslant_{\mathbb{R}} 2^{-j} \to |f(x) - f(y)| \leqslant_{\mathbb{R}} 2^{-k} \right),$$

where in fact one can choose

$$j = \omega^f(k, b) = k + C(b).$$

(4) f is bounded on bounded sets, i.e.

$$\forall b^0 \exists d^0 \forall x^X \left(\|x\|_X <_{\mathbb{R}} b \to |f(x)| \leqslant_{\mathbb{R}} d \right),$$

where in fact one can choose

$$d = D(b) = b2^{\omega^f(0,b)} + [|f(0)|](0) + 2.$$

Correspondingly, the theory $\mathcal{D}^\omega[f, \nabla f, f^*, \nabla f^*]$ proves similar properties appropriately formulated for f^* and ∇f^*. Also, this latter system proves that f^* is convex.

Also, clearly, the Bregman distance

$$D_f(x, y) = f(x) - f(y) - \langle x - y, \nabla f(y) \rangle$$

and its associated dual

$$W_f(x, x^*) = f(x) - \langle x, x^* \rangle + f^*(x^*).$$

can now just be given via closed terms in the language of $\mathcal{D}^\omega[f, \nabla f, f^*, \nabla f^*]$ and this system immediately proves the three and four point identities:

Lemma 2.16 ([73])**.** *The system $\mathcal{D}^\omega[f, \nabla f, f^*, \nabla f^*]$ proves the three and four point identities (see e.g. [5]):*

(1) $$\begin{cases} \forall x^X, y^X, z^X \big(D_f(x, y) + D_f(y, z) - D_f(x, z) \\ \qquad =_{\mathbb{R}} \langle x -_X y, \nabla f(z) -_{X^*} \nabla f(y) \rangle_{X^*} \big). \end{cases}$$

(2) $$\begin{cases} \forall x^X, y^X, z^X, w^X \big(D_f(y, x) - D_f(y, z) - D_f(w, x) + D_f(w, z) \\ \qquad =_{\mathbb{R}} \langle y -_X w, \nabla f(z) -_{X^*} \nabla f(x) \rangle_{X^*} \big). \end{cases}$$

3. Logical systems for operators and their resolvents

We now turn to the main aim of this paper which is to extend the previously discussed systems for convex functions, their derivatives and conjugates over Banach spaces with their duals with a treatment for monotone operators on Banach spaces in the sense of Browder as introduced before, together with the relativized resolvents in the style of Eckstein, both of which are approach similarly as monotone operators and their usual resolvents in Hilbert spaces as developed in [72].

For that, as mentioned before already, we will assume that the function f is total and Fréchet differentiable everywhere with a gradient that is uniformly continuous on bounded sets and where additionally f is also supercoercive and f^* is similarly Fréchet differentiable everywhere with a gradient that is uniformly continuous on bounded sets. Presumably, many of the considerations made here could be extended mutatis mutandis to the case where we only consider partial convex functions f with an intensional treatment of the domain as discussed in [73] but we do not discuss this any further.

Also, while the treatment for monotone operators over Hilbert spaces presented in [72] was divided on whether the resolvents are partial or total, we in the following will only consider systems for monotone operators on Banach spaces where the (relativized) resolvents are all total. In particular, we thereby exactly treat maximally monotone operators by Proposition 2.11 which is hence analogous to the approach to maximal monotone operators in Hilbert spaces taken in [72] which in that way leads to a proof-theoretically tame system for such objects while direct considerations of maximality are in general not feasible (as will be explored further in Section 4 later on). Also, we want to remark that if one would want to treat operators with partial resolvents, then a similar approach as presented in [72] could presumably also be followed here.

In the following, we will omit types of variables whenever convenient and omit types in proofs almost always to make everything more readable.

3.1. **Further considerations on convex functions.** The basic system for all extensions considered here will be $\mathcal{D}^\omega[f, \nabla f, f^*, \nabla f^*]$ from the preceding section, treating the dual of the abstract normed space together with a convex function, its Fenchel-conjugate and their uniformly continuous gradients. It will be convenient to slightly extend this system so that the theory of monotone operators can be developed smoothly. Concretely, it will be convenient to axiomatically include a few more properties of a convex function and its Fenchel conjugate into the previous system. Namely, by the Fenchel–Moreau theorem (see e.g. [10]), we know that if f is proper, lower-semicontinuous and convex, then f^* is proper and $f = f^{**}$ where we define $f^{**} : X \to (-\infty, +\infty]$ by

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

With this definition, we follow one particular approach to biconjugates as e.g. outlined in [10]. In other works, one finds f^{**} introduced as $(f^*)^*$ acting on X^{**} and thus on X by its embedding into X^{**}, which anyhow coincides with X in the context of reflexivity. As the spaces considered in the context of $\mathcal{D}^\omega[f, \nabla f, f^*, \nabla f^*]$ are (super-)reflexive by the results of [11, 12], these different approaches yield the same object but the above formulation will also influence the types of objects considered later.

Naturally, a function f as axiomatized by the system $\mathcal{D}^\omega[f, \nabla f, f^*, \nabla f^*]$ satisfies the assumptions of the Fenchel-Moreau theorem and the resulting fact that $f = f^{**}$ is crucial for the development of the theory of monotone operators, so we need to deal with it formally. However, we do not treat this fact by analyzing the corresponding proof and verifying that it indeed can be carried out in $\mathcal{D}^\omega[f, \nabla f, f^*, \nabla f^*]$ and instead hardwire this fact into the system akin to how f^* is treated in $\mathcal{D}^\omega[f, \nabla f, f^*, \nabla f^*]$.

In more detail, note that, in the context of the above systems, $f = f^{**} = (f^*)^*$ is bounded on bounded sets and therefore f^* is supercoercive by Proposition 2.5. So $f = f^{**}$ can be wired into the system by using a modulus of supercoercivity α^{f^*} for f^* together with the following axioms which are analogous to the axioms $(f^*)_1, (f^*)_2$ from the preceding section (and which similarly instantiate the axioms schemes introduced in [73] for treating suprema over certain bounded sets) where the corresponding bound in the axiom $(f^{**})_2$ again is derived via Lemma 2.14:

$(f^*)_3$ f^* is supercoercive with modulus α^{f^*}, i.e.

$$\forall K^0, x^{*X^*} \left(\|x^*\|_{X^*} >_{\mathbb{R}} \alpha^{f^*}(K) \to f^*(x^*)/\|x^*\|_{X^*} \geqslant_{\mathbb{R}} K \right).$$

Here, α^{f^*} is an additional constant of type 1.

$(f^{**})_1$ f is the pointwise upper bound for all affine functionals $g_{x^*}(x) = \langle x, x^* \rangle - f^*(x^*)$, i.e.

$$\forall x^X, x^{*X^*} \left(\langle x, x^* \rangle_{X^*} - f^*(x^*) \leqslant_{\mathbb{R}} f(x) \right).$$

$(f^{**})_2$ f is indeed the pointwise supremum of these affine functionals, i.e.

$$\forall x^X, b^0, k^0 \exists x^{*X^*} \leqslant_{X^*} \max\{\alpha^{f^*}(b+1)+1, [|f^*(0)|](0)+2\}1_{X^*}$$
$$\left(\|x\|_X <_{\mathbb{R}} b \to \left(f(x) - 2^{-k} \leqslant_{\mathbb{R}} \langle x, x^* \rangle_{X^*} - f^*(x^*) \right) \right).$$

With $\mathcal{D}^\omega_{f,f^*}[\mathrm{FM}]$ we abbreviate the system that arises from $\mathcal{D}^\omega[f, \nabla f, f^*, \nabla f^*]$ by adding these constants and axioms. Later, we will see that all these (new) axioms are indeed admissible while aiming for bound extraction theorems in the style of proof mining (following similar reasoning as in [73]).

One crucial aspect of the theory of the conjugate function f^* is that it provides alternative characterizations of continuity properties of ∇f in terms of convexity properties of f^*. These play a crucial role in the context of the theory of monotone operators in Banach spaces and so, before moving on to these, we first formally establish some of these corresponding convexity properties of f^*. As a byproduct, we also formally provide further quantitative properties of the gradient ∇f.

Lemma 3.1. *The system $\mathcal{D}^\omega[f, \nabla f, f^*, \nabla f^*]$ proves:*

(1) The "Fenchel-Young equality"[7] for any subgradient u^* of f at x, i.e.

$$\forall x^X, u^{*X^*} (\forall y^X \left(f(y) \geqslant_{\mathbb{R}} f(x) + \langle y -_X x, u^* \rangle_{X^*} \right) \to f(x) + f^*(u^*) =_{\mathbb{R}} \langle x, u^* \rangle_{X^*}).$$

(2) The "Fenchel-Young equality" for $\nabla f x$, i.e.

$$\forall x^X \left(f(x) + f^*(\nabla f x) =_{\mathbb{R}} \langle x, \nabla f x \rangle_{X^*} \right).$$

(3) Approximate subgradients of f are close to the gradient of f, i.e.

$$\forall b^0, k^0 \exists j^0 \forall x^X, x^{*X^*} (\|x\|_X <_{\mathbb{R}} b \wedge \forall y^X \left(\langle y -_X x, x^* \rangle_{X^*} + f(x) \leqslant_{\mathbb{R}} f(y) + 2^{-j} \right)$$
$$\to \|x^* -_{X^*} \nabla f x\|_{X^*} \leqslant_{\mathbb{R}} 2^{-k}),$$

where in fact we can take

$$j = k + 4 + \omega^{\nabla f}(k+3, b+1).$$

(4) The "Fenchel-Young equality" characterizes gradients of f, i.e.

$$\forall x^X, x^{*X^*} \left(f(x) + f^*(x^*) =_{\mathbb{R}} \langle x, x^* \rangle_{X^*} \to x^* =_{X^*} \nabla f x \right),$$

where in fact it moreover holds that

$$\forall b^0, k^0 \exists j^0 \forall x^X, x^{*X^*} (\|x\|_X <_{\mathbb{R}} b \wedge f(x) + f^*(x^*) - \langle x, x^* \rangle_{X^*} \leqslant_{\mathbb{R}} 2^{-j}$$
$$\to \|x^* -_{X^*} \nabla f x\|_{X^*} \leqslant_{\mathbb{R}} 2^{-k})$$

where we can take

$$j = k + 4 + \omega^{\nabla f}(k+3, b+1).$$

(5) $\nabla f x$ is the unique subgradient of f at x, i.e.

$$\forall x^X, u^{*X^*} \left(\forall y^X \left(f(y) \geqslant_{\mathbb{R}} f(x) + \langle y -_X x, u^* \rangle_{X^*} \right) \to u^* =_{X^*} \nabla f x \right).$$

[7] By this expression, we mean in the following that the Fenchel-Young inequality is not strict, i.e. is satisfied with equality.

(6) f^* is uniformly strictly convex on bounded subsets, i.e.

$$\begin{aligned}&\forall k^0, i^0, b^0 \exists j^0 \forall x^{*X^*}, y^{*X^*}, t^1 (\|x^*\|_{X^*}, \|y^*\|_{X^*} <_{\mathbb{R}} b \wedge 2^{-i} \leqslant_{\mathbb{R}} t \leqslant_{\mathbb{R}} 1 - 2^{-i} \\ &\quad \wedge tf^*(x^*) + (1-t)f^*(y^*) - f^*(tx^* +_{X^*} (1-t)y^*) \leqslant_{\mathbb{R}} 2^{-j} \\ &\qquad \to \|x^* -_{X^*} y^*\|_{X^*} \leqslant_{\mathbb{R}} 2^{-k})\end{aligned}$$

where we in fact can choose

$$j = (k + 4 + \omega^{\nabla f}(k+4, F(b)+2)) + i$$

where F is a modulus for ∇f^* being bounded on bounded sets (which can be constructed similar to Lemma 2.15).

(7) ∇f^* is uniformly strictly monotone on bounded subsets, i.e.

$$\begin{aligned}&\forall k^0, b^0 \exists j^0 \forall x^{*X^*}, y^{*X^*} (\|x^*\|_{X^*}, \|y^*\|_{X^*} <_{\mathbb{R}} b \wedge \|x^* -_{X^*} y^*\|_{X^*} >_{\mathbb{R}} 2^{-k} \\ &\qquad \to \langle \nabla f^* x^* -_X \nabla f^* y^*, x^* -_{X^*} y^* \rangle_{X^*} \geqslant_{\mathbb{R}} 2^{-j}).\end{aligned}$$

where we in fact can choose

$$j = k + 5 + \omega^{\nabla f}(k+4, F(b)+2).$$

with all other constants as in (6).

Proof. (1) Let u^* be such that $\forall y\, (f(y) \geqslant f(x) + \langle y - x, u^* \rangle)$, i.e.

$$\langle y, u^* \rangle - f(y) \leqslant \langle x, u^* \rangle - f(x)$$

for all y. Using $(f^*)_2$, we get that for any j, there exists a y_j such that

$$f^*(u^*) - (\langle x, u^* \rangle - f(x)) \leqslant f^*(u^*) - (\langle y_j, u^* \rangle - f(y_j)) \leqslant 2^{-j}$$

and thus we have $f^*(u^*) \leqslant \langle x, u^* \rangle - f(x)$. Using axiom $(f^*)_1$, we get $\langle x, u^* \rangle - f(x) \leqslant f^*(u^*)$ and combined this gives the result.

(2) Follows immediately from (1) and $(\nabla f)_1$.

(3) Let x^* be such that

$$\forall y \left(\langle y - x, x^* \rangle + f(x) \leqslant f(y) + 2^{-j}\right)$$

for j defined as above. This yields

$$\begin{aligned}\langle y - x, x^* - \nabla f x \rangle &= \langle y - x, x^* \rangle - \langle y - x, \nabla f(x) \rangle \\ &\leqslant f(y) - f(x) - \langle y - x, \nabla f(x) \rangle + 2^{-j}.\end{aligned}$$

Using Lemma 2.15, we get that for $\|y - x\| < 2^{-\omega^{\nabla f}(l, b+1)}$:

$$f(y) - f(x) - \langle y - x, \nabla f(x) \rangle \leqslant 2^{-l} \|y - x\|$$

and so $\langle y - x, x^* - \nabla f x \rangle \leqslant 2^{-j} + 2^{-l} \|y - x\|$ for all such y which in particular yields

$$\langle z, x^* - \nabla f x \rangle \leqslant 2^{-j} + 2^{-l} \|z\|$$

for all z with $\|z\| < 2^{-\omega^{\nabla f}(l, b+1)}$ given any l. For the given k, we now show that there is a z_k such that $0 < \|z_k\| \leqslant 1$ and

$$\langle z_k, x^* - \nabla f x \rangle \leqslant 2^{-(k+2)} \to \|x^* - \nabla f x\| \leqslant 2^{-k}.$$

To see this, suppose $\|x^* - \nabla f x\| > 2^{-k} = 2^{-(k+1)} + 2^{-(k+1)}$. Using axiom $(*)_2$, we get that there exists a z_k with $\|z_k\| \leqslant 1$ and such that $|\langle z_k, x^* - \nabla f x \rangle| > 2^{-(k+2)}$, which also in particular implies $\|z_k\| > 0$.

Now, using this z_k, define

$$\widehat{z_k} = 2^{-(\omega^{\nabla f}(k+3,b+1)+1)} z_k.$$

Clearly $\|\widehat{z_k}\| < 2^{-\omega^{\nabla f}(k+3,b+1)}$ and thus

$$\langle \widehat{z_k}, x^* - \nabla f x\rangle \leqslant 2^{-j} + 2^{-(k+3)} \|\widehat{z_k}\|$$

which yields by definition of j that

$$\begin{aligned}
\langle z_k, x^* - \nabla f x\rangle &= 2^{(\omega^{\nabla f}(k+3,b+1)+1)}\langle \widehat{z_k}, x^* - \nabla f x\rangle \\
&\leqslant 2^{(\omega^{\nabla f}(k+3,b+1)+1)}(2^{-j} + 2^{-(k+3)} \|\widehat{z_k}\|) \\
&= 2^{(\omega^{\nabla f}(k+3,b+1)+1)}2^{-j} + 2^{-(k+3)}2^{(\omega^{\nabla f}(k+3,b+1)+1)} \|\widehat{z_k}\|) \\
&\leqslant 2^{-(k+4+\omega^{\nabla f}(k+3,b+1))}2^{(\omega^{\nabla f}(k+3,b+1)+1)} + 2^{-(k+3)} \\
&= 2^{-(k+2)}
\end{aligned}$$

which implies $\|x^* - \nabla f x\| \leqslant 2^{-k}$ by the properties of z_k.

(4) Let x^* be such that $f(x) + f^*(x^*) - \langle x, x^*\rangle_{X^*} \leqslant 2^{-j}$ with j defined as above. Then we get $f^*(x^*) \leqslant 2^{-j} + \langle x, x^*\rangle - f(x)$ which yields through $(f^*)_1$ that

$$\langle y, x^*\rangle - f(y) \leqslant 2^{-j} + \langle x, x^*\rangle - f(x)$$

for all y. This is equivalent to $\langle y - x, x^*\rangle + f(x) \leqslant f(y) + 2^{-j}$ for all y. Then item (3) yields the result.

(5) This follows immediately from (3).

(6) Suppose

$$|tf^*(x^*) + (1-t)f^*(y^*) - f^*(tx^* + (1-t)y^*)| \leqslant_{\mathbb{R}} 2^{-j}$$

for j as above. Then write $z^* = tx^* + (1-t)y^*$ and pick $x = \nabla f^* z^*$, i.e. $\nabla f x = z^*$ by (L). Then by item (2), the extensionality of f^* (recall Lemma 2.15) and the extensionality of $\langle\cdot,\cdot\rangle$, we get

$$\begin{aligned}
0 &= f(x) + f^*(z^*) - \langle x, z^*\rangle \\
&\geqslant f(x) + tf^*(x^*) + (1-t)f^*(y^*) - 2^{-j} - \langle x, z^*\rangle,
\end{aligned}$$

i.e. we have

$$2^{-j} \geqslant t(f(x) + f^*(x^*) - \langle x, x^*\rangle) + (1-t)(f(x) + f^*(y^*) - \langle x, y^*\rangle)$$

and thus, using $t, 1-t \geqslant 2^{-i}$ and that $f(x) + f^*(x^*) - \langle x, x^*\rangle \geqslant 0$ as well as $f(x) + f^*(y^*) - \langle x, y^*\rangle \geqslant 0$ by the Fenchel-Young inequality (which follows from axiom $(f^*)_1$), we get

$$2^{-j}2^{i} \geqslant f(x) + f^*(x^*) - \langle x, x^*\rangle, f(x) + f^*(y^*) - \langle x, y^*\rangle$$

By definition of j, we get

$$2^{-(k+4+\omega^{\nabla f}(k+4,F(b)+2))} \geqslant f(x) + f^*(x^*) - \langle x, x^*\rangle, f(x) + f^*(y^*) - \langle x, y^*\rangle.$$

Noting that $\|z^*\| \leqslant t\|x^*\| + (1-t)\|y^*\| < b$ and thus $\|x\| < F(b) + 1$, item (4) implies that

$$\|x^* - \nabla f x\|, \|y^* - \nabla f x\| \leqslant 2^{-(k+1)}$$

which yields $\|x^* - y^*\| \leqslant 2^{-k}$.

(7) Using item (6), note that for $t = 1/2$, we have

$$\begin{aligned} f^*\left(\frac{x^*+y^*}{2}\right) &\leqslant 1/2 f^*(y^*) + 1/2 f^*(x^*) - 2^{-j} \\ &= f^*(x^*) + 1/2(f^*(y^*) - f^*(x^*)) - 2^{-j} \end{aligned}$$

if $\|x^* - y^*\| > 2^{-k}$. As

$$\langle \nabla f^* w^*, z^* \rangle \leqslant \frac{f^*(w^* + \alpha z^*) - f^*(w^*)}{\alpha},$$

for any $\alpha > 0$ (using $(\nabla f^*)_1$), we get

$$\langle \nabla f^* x^*, y^* - x^* \rangle \leqslant f^*(y^*) - f^*(x^*) - 2 \cdot 2^{-j}.$$

Similarly, we get

$$\langle \nabla f^* y^*, x^* - y^* \rangle \leqslant f^*(x^*) - f^*(y^*) - 2 \cdot 2^{-j}$$

and this implies

$$\langle \nabla f^* y^* - \nabla f^* x^*, x^* - y^* \rangle \leqslant -4 \cdot 2^{-j}$$

which gives the claim.

□

Now, the additional axioms in $\mathcal{D}^\omega_{f,f^*}[\mathrm{FM}]$ can be used to carry out the above proof with the roles of f and f^* exchanged. We collect this in the following lemma.

Lemma 3.2. *The system $\mathcal{D}^\omega_{f,f^*}[\mathrm{FM}]$ proves:*

(1) The "Fenchel-Young equality" for any subgradient u of f^* at x^*, i.e.

$$\begin{aligned} &\forall x^{*X^*}, u^X (\forall y^{*X^*} \left(f^*(y^*) \geqslant_{\mathbb{R}} f^*(x^*) + \langle u, y^* -_{X^*} x^* \rangle_{X^*}\right) \\ &\qquad \to f^*(x^*) + f(u) =_{\mathbb{R}} \langle u, x^* \rangle_{X^*}). \end{aligned}$$

(2) The "Fenchel-Young equality" for $\nabla f^* x$, i.e.

$$\forall x^{*X^*} \left(f^*(x^*) + f(\nabla f^* x^*) =_{\mathbb{R}} \langle \nabla f^* x^*, x^* \rangle_{X^*}\right).$$

(3) Approximate subgradients of f^* are close to the gradient of f^*, i.e.

$$\begin{aligned} &\forall b^0, k^0 \exists j^0 \forall x^{*X^*}, x^X (\|x^*\|_{X^*} <_{\mathbb{R}} b \\ &\qquad \wedge \forall y^{*X^*} \left(\langle x, y^* -_{X^*} x^* \rangle_{X^*} + f^*(x^*) \leqslant_{\mathbb{R}} f^*(y^*) + 2^{-j}\right) \\ &\qquad \to \|x -_X \nabla f^* x^*\|_X \leqslant_{\mathbb{R}} 2^{-k}), \end{aligned}$$

where in fact we can take

$$j = k + 4 + \omega^{\nabla f^*}(k+3, b+1).$$

(4) The "Fenchel-Young equality" characterizes gradients of f^*, i.e.

$$\forall x^{*X^*}, x^X \left(f^*(x^*) + f(x) =_{\mathbb{R}} \langle x, x^* \rangle_{X^*} \to x =_X \nabla f^* x^*\right),$$

where in fact it moreover holds that

$$\begin{aligned} &\forall b^0, k^0 \exists j^0 \forall x^{*X^*}, x^X (\|x^*\|_{X^*} <_{\mathbb{R}} b \wedge f^*(x^*) + f(x) - \langle x, x^* \rangle_{X^*} \leqslant_{\mathbb{R}} 2^{-j} \\ &\qquad \to \|x -_X \nabla f^* x^*\|_X \leqslant_{\mathbb{R}} 2^{-k}) \end{aligned}$$

where we can take

$$j = k + 4 + \omega^{\nabla f^*}(k+3, b+1).$$

(5) $\nabla f^* x^*$ *is the unique subgradient of* f^* *at* x^*, *i.e.*

$$\forall x^{*X^*}, u^X \left(\forall y^{*X^*} \left(f^*(y^*) \geqslant_{\mathbb{R}} f^*(x^*) + \langle u, y^* -_{X^*} x^* \rangle_{X^*}\right) \to u =_X \nabla f^* x^*\right).$$

(6) f *is uniformly strictly convex on bounded subsets, i.e.*

$$\begin{aligned}\forall k^0, i^0, b^0 \exists j^0 \forall x^X, y^X, t^1 (&\|x\|_X, \|y\|_X <_{\mathbb{R}} b \wedge 2^{-i} \leqslant_{\mathbb{R}} t \leqslant_{\mathbb{R}} 1 - 2^{-i} \\ &\wedge tf(x) + (1-t)f(y) - f(tx +_X (1-t)y) \leqslant_{\mathbb{R}} 2^{-j} \\ &\to \|x -_X y\|_X \leqslant_{\mathbb{R}} 2^{-k})\end{aligned}$$

where we in fact can choose

$$j = (k + 4 + \omega^{\nabla f^*}(k+4, C(b)+2)) + i$$

where C *is a modulus for* ∇f *being bounded on bounded sets (which can be constructed as in Lemma 2.15).*

(7) ∇f *is uniformly strictly monotone on bounded subsets, i.e.*

$$\begin{aligned}\forall k^0, b^0 \exists j^0 \forall x^X, y^X (&\|x\|_X, \|y\|_X <_{\mathbb{R}} b \wedge \|x -_X y\|_X >_{\mathbb{R}} 2^{-k} \\ &\to \langle x -_X y, \nabla f x -_{X^*} \nabla f y \rangle_{X^*} \geqslant_{\mathbb{R}} 2^{-j}).\end{aligned}$$

where we in fact can choose

$$j = k + 5 + \omega^{\nabla f^*}(k+4, C(b)+2).$$

with all other constants as in (6).

In particular, in the system $\mathcal{D}^\omega_{f,f^*}[\mathrm{FM}]$ we can now formally establish some of the central properties of Bregman distances used extensively throughout the applications given in [76] and which are also crucial for the development of the theory of the relativized resolvents. We begin with the fact that $W_f(x, \nabla f(y)) = D_f(x,y)$:

Lemma 3.3. *The system* $\mathcal{D}^\omega_{f,f^*}[\mathrm{FM}]$ *proves:*

$$\forall x^X, y^X \left(D_f(x,y) =_{\mathbb{R}} f(x) + f^*(\nabla f y) - \langle x, \nabla f y \rangle_{X^*}\right).$$

Proof. By Lemma 3.1, (2), we have

$$f^*(\nabla f y) = \langle y, \nabla f y \rangle - f(y)$$

and thus

$$\begin{aligned}f(x) + f^*(\nabla f y) - \langle x, \nabla f y \rangle &= f(x) - f(y) - \langle x - y, \nabla f y \rangle \\ &= D_f(x,y).\end{aligned}$$

□

Lemma 3.4. *The system* $\mathcal{D}^\omega_{f,f^*}[\mathrm{FM}]$ *proves that* D_f *is uniformly bounded in the sense of* [76], *i.e.*

$$\forall b^0, \alpha^0 \exists o^0 \forall x^X, y^X \left(\|x\|_X <_{\mathbb{R}} b \wedge D_f(x,y) <_{\mathbb{R}} \alpha \to \|y\|_X \leqslant_{\mathbb{R}} o\right)$$

and o *can be realized by*

$$o = o(\alpha, b) = F(\alpha^{f^*}(\alpha + D(b) + b) + 1)$$

where D, F *are moduli of* f, ∇f^* *being bounded on bounded sets, respectively, and* α^{f^*} *is a modulus of supercoercivity for* f^* *as before.*

Proof. First, note that $f^*(x^*) - \langle x, x^*\rangle$ is also supercoercive. For this, let $\|x\| < b$. If $\|x^*\| > \alpha^{f^*}(K+b)$, from axiom $(f^*)_3$ we derive

$$\frac{f^*(x^*) - \langle x, x^*\rangle}{\|x^*\|} \geqslant \frac{f^*(x^*)}{\|x^*\|} - \|x\| \geqslant K.$$

Now, we have

$$f^*(\nabla f y) - \langle x, \nabla f y\rangle = D_f(x,y) - f(x) < \alpha + D(b)$$

using the above Lemma 3.3. Therefore, we derive

$$\|\nabla f y\| \leqslant \alpha^{f^*}(\alpha + D(b) + b)$$

and thus we get $\|y\| = \|\nabla f^* \nabla f y\| \leqslant F(\alpha^{f^*}(\alpha + D(b) + b) + 1)$. □

3.2. **Monotone operators and their relativized resolvents.** We now move to the central new objects, the monotone operators in the sense of Browder and their relativized resolvents. Initially, generic set-valued operators of the form $A : X \to 2^{X^*}$ are, in similarity to [72], modeled via a constant for their characteristic function. In the context of the system $\mathcal{D}^\omega_{f,f^*}[\mathrm{FM}]$, we in that way add a constant χ_A of type $0(X^*)(X)$ and write $x^* \in Ax$, $(x,x^*) \in A$ or $(x,x^*) \in \mathrm{gra}A$ for $\chi_A x x^* =_0 0$. The first natural axiom is

$$(I)^* \qquad \forall x^X, x^{*X^*}(\chi_A x x^* \leqslant_0 1)$$

which witnesses that χ_A is a characteristic function.

Also, the treatment of the resolvent is conceptually similar to that work. For this, let A be monotone (in the sense of Browder, recall Section 2.1) and recall Definition 2.9 for the f-resolvents of such monotone operators whereby $\mathrm{Res}^f_{\gamma A} : X \to 2^X$ is defined by

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

for any $x \in X$ and $\gamma > 0$. It follows by our assumptions on f and Proposition 2.10 that this map is single-valued, satisfies $\mathrm{Fix}(\mathrm{Res}^f_{\gamma A}) = A^{-1}0$ (noting that $\mathrm{dom} f = X$ in this paper) and that it is Bregman firmly nonexpansive.

So, for treating an operator A with total relativized resolvents, we add a constant Res^f_A of type $X(X)(1)$ and write $\mathrm{Res}^f_{\gamma A}$ for $\mathrm{Res}^f_A \gamma$. The natural axiom for the resolvent now can be derived similar as to [72]: If seen as a set-valued operator, the resolvent satisfies

$$\begin{aligned} p \in \mathrm{Res}^f_{\gamma A} x &\Leftrightarrow p \in (\nabla f + \gamma A)^{-1} \nabla f(x) \\ &\Leftrightarrow \nabla f(x) \in \nabla f(p) + \gamma A p \\ &\Leftrightarrow \gamma^{-1}\left(\nabla f(x) - \nabla f(p)\right) \in Ap. \end{aligned}$$

This naturally leads us to consider the axiom scheme

$$(II)^* \qquad \forall \gamma^1, x^X \left(\gamma >_{\mathbb{R}} 0 \to \gamma^{-1}(\nabla f x -_{X^*} \nabla f(\mathrm{Res}^f_{\gamma A} x)) \in A(\mathrm{Res}^f_{\gamma A} x)\right)$$

in similarity to axiom (II) considered in [72], stating that $\mathrm{Res}^f_{\gamma A}$ is a selection map of the (a priori set-valued) relativized resolvent. In fact, this axiom (together with the axiom specified previously and the two axioms specified further below) will be sufficient for developing all the main parts of the theory of relativized resolvents (e.g. immediately entailing the uniqueness and Bregman firm nonexpansivity of the above selection map) as we will later discuss.

Remark 3.5. As in the context of the systems from [72], note that also here, the above axiom $(II)^*$ is actually an abbreviation for the following sentence where the dependence of γ^{-1} on a lower bound of γ is made explicit:

$$\forall \gamma^1, x^X, k^0 \left(\gamma >_{\mathbb{R}} 2^{-k} \to (\gamma)^{-1}_k (\nabla f x -_{X^*} \nabla f(\mathrm{Res}^f_{\gamma A} x)) \in A(\mathrm{Res}^f_{\gamma A} x) \right).$$

However, also similar to [72], we will still in general employ the above shorthand style of writing where the parameter k is omitted as in this paper, the context will always make it clear how a given statement in that style has to be expanded to yield a proper formal variant so that no issues arise.

Also the monotonicity of A is easily specified by a universal axiom:

$$(III)^* \qquad \forall x^X, y^X, x^{*X^*}, y^{*X^*} \left((x, x^*), (y, y^*) \in A \to \langle x -_X y, x^* -_{X^*} y^* \rangle_{X^*} \geqslant_{\mathbb{R}} 0\right).$$

Lastly, all uses of the relativized resolvent in proof mining applications presented so far (see in particular [76]) are made in the context of the assumption that $A^{-1}0 \neq \varnothing$ and we will also assume this here as it in particular will allow us to majorize the resolvent rather immediately. For this, we add a constant p_X of type X together with a corresponding axiom stating that p_X is a zero of A:

$$(IV)^* \qquad 0_{X^*} \in Ap_X.$$

This leads us to the following system:

Definition 3.6. The theory $\mathcal{B}^\omega$ is defined as the extension of the theory $\mathcal{D}^\omega_{f,f^*}[\mathrm{FM}]$ with the above constants and corresponding axioms $(I)^*$ - $(IV)^*$.

Now, in similarity to the systems from [72], also $\mathcal{B}^\omega$ is sufficient for formalizing the first main aspects of the theory of monotone operators in Banach spaces and their f-resolvents as the following proposition shows.

Lemma 3.7. *The system $\mathcal{B}^\omega$ proves:*

(1) $\mathrm{Res}^f_{\gamma A}$ *is unique for any $\gamma > 0$, i.e.*

$$\forall \gamma^1, p^X, x^X \left(\gamma >_{\mathbb{R}} 0 \wedge \gamma^{-1}(\nabla f x -_{X^*} \nabla f p) \in Ap \to p =_X \mathrm{Res}^f_{\gamma A} x \right).$$

(2) $\mathrm{Res}^f_{\gamma A}$ *is Bregman firmly nonexpansive for any $\gamma > 0$, i.e.*

$$\begin{aligned}\forall \gamma^1, x^X, y^X (\gamma >_{\mathbb{R}} 0 \to \langle \mathrm{Res}^f_{\gamma A} x -_X \mathrm{Res}^f_{\gamma A} y, \nabla f \mathrm{Res}^f_{\gamma A} x -_{X^*} \nabla f \mathrm{Res}^f_{\gamma A} y \rangle_{X^*} \\ \leqslant_{\mathbb{R}} \langle \mathrm{Res}^f_{\gamma A} x -_X \mathrm{Res}^f_{\gamma A} y, \nabla f x -_{X^*} \nabla f y \rangle_{X^*}).\end{aligned}$$

(3) $\mathrm{Res}^f_{\gamma A}$ *satisfies the alternative notion of Bregman firm nonexpansivity for any $\gamma > 0$, i.e.*

$$\begin{aligned}\forall \gamma^1, x^X, y^X (\gamma >_{\mathbb{R}} 0 \to D_f(\mathrm{Res}^f_{\gamma A} x, \mathrm{Res}^f_{\gamma A} y) + D_f(\mathrm{Res}^f_{\gamma A} y, \mathrm{Res}^f_{\gamma A} x) \\ \leqslant_{\mathbb{R}} D_f(\mathrm{Res}^f_{\gamma A} x, y) + D_f(\mathrm{Res}^f_{\gamma A} y, x) - D_f(\mathrm{Res}^f_{\gamma A} x, x) - D_f(\mathrm{Res}^f_{\gamma A} y, y)).\end{aligned}$$

(4) $A^{-1}0 \subseteq \mathrm{Fix}(\mathrm{Res}^f_{\gamma A})$ *for any $\gamma > 0$, i.e.*

$$\forall p^X, \gamma^1 \left(\gamma >_{\mathbb{R}} 0 \wedge 0 \in Ap \to p =_X \mathrm{Res}^f_{\gamma A} p \right).$$

Proof. (1) Suppose that $\gamma > 0$ and that $\gamma^{-1}(\nabla fx - \nabla fp) \in Ap$. Axiom $(II)^*$ gives $\gamma^{-1}(\nabla fx - \nabla f\mathrm{Res}^f_{\gamma A}x) \in A(\mathrm{Res}^f_{\gamma A}x)$. Axiom $(III)^*$ then implies that

$$\begin{aligned}0 &\leqslant \langle \mathrm{Res}^f_{\gamma A}x - p, \gamma^{-1}(\nabla fx - \nabla f\mathrm{Res}^f_{\gamma A}x) - \gamma^{-1}(\nabla fx - \nabla fp)\rangle\\ &= \langle \mathrm{Res}^f_{\gamma A}x - p, \gamma^{-1}(\nabla fp - \nabla f\mathrm{Res}^f_{\gamma A}x)\rangle\end{aligned}$$

where we have used extensionality of $\langle\cdot,\cdot\rangle$ and of the arithmetical operations in X^*. In particular, since $\gamma^{-1} > 0$ as $\gamma > 0$, we get that

$$\langle \mathrm{Res}^f_{\gamma A}x - p, \nabla f\mathrm{Res}^f_{\gamma A}x - \nabla fp\rangle \leqslant 0.$$

Thus, as ∇f is provably strictly monotone (Lemma 3.2), we get $\left\|\mathrm{Res}^f_{\gamma A}x - p\right\| = 0$, i.e. $\mathrm{Res}^f_{\gamma A}x = p$.

(2) Let $\gamma > 0$. Axiom $(II)^*$ gives

$$\gamma^{-1}(\nabla fx - \nabla f\mathrm{Res}^f_{\gamma A}x) \in A(\mathrm{Res}^f_{\gamma A}x) \text{ and } \gamma^{-1}(\nabla fy - \nabla f\mathrm{Res}^f_{\gamma A}y) \in A(\mathrm{Res}^f_{\gamma A}y).$$

Axiom $(III)^*$ and $\gamma^{-1} > 0$ gives

$$\langle \mathrm{Res}^f_{\gamma A}x - \mathrm{Res}^f_{\gamma A}y, \nabla fx - \nabla fy - (\nabla f\mathrm{Res}^f_{\gamma A}x - \nabla f\mathrm{Res}^f_{\gamma A}y)\rangle \geqslant 0$$

which implies

$$\langle \mathrm{Res}^f_{\gamma A}x - \mathrm{Res}^f_{\gamma A}y, \nabla fx - \nabla fy\rangle \geqslant \langle \mathrm{Res}^f_{\gamma A}x - \mathrm{Res}^f_{\gamma A}y, \nabla f\mathrm{Res}^f_{\gamma A}x - \nabla f\mathrm{Res}^f_{\gamma A}y\rangle.$$

(3) By the provability of the three-point identity for D_f (Lemma 2.16), we get

$$\begin{aligned}&\langle \mathrm{Res}^f_{\gamma A}x - \mathrm{Res}^f_{\gamma A}y, \nabla f\mathrm{Res}^f_{\gamma A}x - \nabla f\mathrm{Res}^f_{\gamma A}y\rangle\\ &\quad= D_f(\mathrm{Res}^f_{\gamma A}x, \mathrm{Res}^f_{\gamma A}y) + D_f(\mathrm{Res}^f_{\gamma A}y, \mathrm{Res}^f_{\gamma A}x) - D_f(\mathrm{Res}^f_{\gamma A}x, \mathrm{Res}^f_{\gamma A}x)\\ &\quad= D_f(\mathrm{Res}^f_{\gamma A}x, \mathrm{Res}^f_{\gamma A}y) + D_f(\mathrm{Res}^f_{\gamma A}y, \mathrm{Res}^f_{\gamma A}x).\end{aligned}$$

Further, by the provability of the four-point identity for D_f (Lemma 2.16), we get

$$\begin{aligned}&\langle \mathrm{Res}^f_{\gamma A}x - \mathrm{Res}^f_{\gamma A}y, \nabla fx - \nabla fy\rangle\\ &\quad= D_f(\mathrm{Res}^f_{\gamma A}x, y) - D_f(\mathrm{Res}^f_{\gamma A}x, x) - D_f(\mathrm{Res}^f_{\gamma A}y, y) + D_f(\mathrm{Res}^f_{\gamma A}y, x).\end{aligned}$$

Thus, using item (2), we get the claimed inequality.

(4) Let p be such that $0 \in Ap$. Then provably with the only assumption being $\gamma > 0$, we have $\gamma^{-1}(\nabla fp - \nabla fp) = 0$ and thus, using Σ_1-ER (recall Remark 2.12), we have that $0 \in Ap$ implies

$$\gamma^{-1}(\nabla fp - \nabla fp) \in Ap.$$

Using item (1), we get $p = \mathrm{Res}^f_{\gamma A}p$.

□

Also the boundedness and continuity properties of maps that are Bregman firmly nonexpansive, already emphasized in [76] as crucial for the practice of proof mining for these objects, can now be formally replicated in the context of the system $\mathcal{B}^\omega$ (where we here formulate these properties just for the resolvents):

Proposition 3.8. *The system $\mathcal{B}^\omega$ proves:*

(1) $\mathrm{Res}^f_{\gamma A}$ *is bounded on bounded sets for any* $\gamma > 0$*, i.e.*

$$\forall \gamma^1, b^0 \exists e^0 \forall x^X \left(\gamma >_{\mathbb{R}} 0 \land \|p_X\|_X, \|x\|_X <_{\mathbb{R}} b \to \left\|\mathrm{Res}^f_{\gamma A} x\right\|_X \leqslant_{\mathbb{R}} e\right),$$

where in fact one can choose

$$e = E(b) = o(2D(b) + 2bC(b), b)$$

where C, D *are moduli witnessing that* $\nabla f, f$ *are bounded on bounded sets, respectively, and* o *is defined as in Lemma 3.4.*

(2) $\mathrm{Res}^f_{\gamma A}$ *is uniformly continuous on bounded sets for any* $\gamma > 0$*, i.e.*

$$\forall \gamma^1, k^0, b^0 \exists j^0 \forall x^X, y^X (\gamma >_{\mathbb{R}} 0 \land \|p_X\|_X, \|x\|_X, \|y\|_X <_{\mathbb{R}} b$$
$$\land \|x -_X y\|_X <_{\mathbb{R}} 2^{-j} \to \left\|\mathrm{Res}^f_{\gamma A} x -_X \mathrm{Res}^f_{\gamma A} y\right\|_X \leqslant_{\mathbb{R}} 2^{-k})$$

where in fact one can choose

$$j = \varpi(k, b) = \omega^{\nabla f}(\widehat{k} + 1 + E(b), b)$$

for $\widehat{k} = k + 5 + \omega^{\nabla f^*}(k + 4, C(b) + 2)$ *with* C *being a modulus witnessing that* ∇f *is bounded on bounded sets and where* E *is defined as in (1).*

Proof. For item (1), note that by Lemma 3.7, (3) and (4), and with $p = p_X$ from axiom $(IV)^*$, we have (using the extensionality of D_f which follows from that of $f, \nabla f$ and $\langle \cdot, \cdot \rangle$):

$$\begin{aligned} D_f(\mathrm{Res}^f_{\gamma A} x, p) + D_f(p, \mathrm{Res}^f_{\gamma A} x) &\leqslant D_f(\mathrm{Res}^f_{\gamma A} x, p) + D_f(p, x) - D_f(\mathrm{Res}^f_{\gamma A} x, x) - D_f(p, p) \\ &\leqslant D_f(\mathrm{Res}^f_{\gamma A} x, p) + D_f(p, x) \end{aligned}$$

and thus

$$D_f(p, \mathrm{Res}^f_{\gamma A} x) \leqslant D_f(p, x) < 2D(b) + 2bC(b).$$

Thus, by Lemma 3.4, we get

$$\left\|\mathrm{Res}^f_{\gamma A} x\right\| \leqslant o(2D(b) + 2bC(b), b).$$

For item (2), by Lemma 3.7, (2), we have

$$\begin{aligned} &\langle \mathrm{Res}^f_{\gamma A} x - \mathrm{Res}^f_{\gamma A} y, \nabla f \mathrm{Res}^f_{\gamma A} x - \nabla f \mathrm{Res}^f_{\gamma A} y \rangle \\ &\qquad \leqslant \langle \mathrm{Res}^f_{\gamma A} x - \mathrm{Res}^f_{\gamma A} y, \nabla f x - \nabla f y \rangle \\ &\qquad \leqslant \left\|\mathrm{Res}^f_{\gamma A} x - \mathrm{Res}^f_{\gamma A} y\right\| \|\nabla f x - \nabla f y\| \\ &\qquad \leqslant 2E(b) \|\nabla f x - \nabla f y\|, \end{aligned}$$

using also additionally the above item (1). So, for $\|x - y\| < 2^{-j}$ with the j defined above, we have

$$\|\nabla f x - \nabla f y\| \leqslant 2^{-(\widehat{k} + 1 + E(b))}$$

and thus

$$\langle \mathrm{Res}^f_{\gamma A} x - \mathrm{Res}^f_{\gamma A} y, \nabla f \mathrm{Res}^f_{\gamma A} x - \nabla f \mathrm{Res}^f_{\gamma A} y \rangle < 2^{-\widehat{k}}.$$

Thus by Lemma 3.2, (7), we get

$$\left\|\mathrm{Res}^f_{\gamma A} x - \mathrm{Res}^f_{\gamma A} y\right\| \leqslant 2^{-k}.$$

□

Notice that therefore the system $\mathcal{B}^\omega$ proves that $\mathrm{Res}^f_{\gamma A}$ is extensional.

4. Maximality and extensionality

Naturally, the system $\mathcal{B}^\omega$ can not prove the extensionality of A since we also deal with potentially discontinuous operators A semantically. As mentioned in the introduction, a central theoretical result from [72] connects this extensionality of A with the maximality statement for A in the setting of monotone operators on Hilbert spaces. We can now extend this result to the monotone operators over Banach spaces. In particular, as all the results are considered in the context of a Legendre function where f and f^* are Fréchet differentiable with gradients that are uniformly continuous on bounded sets, we find by Proposition 2.11 that the totality of the resolvent (as encoded in the system $\mathcal{B}^\omega$) implies that the operators A which are considered semantically are maximally monotone and so this maximality can then not be provable due to this equivalence either.

Theorem 4.1. *Over $\mathcal{B}^\omega$, the following are equivalent:*

(1) Extensionality of A, i.e.

$$\forall x^X, x^{*X^*}, y^X, y^{*X^*} \left(x =_X y \wedge x^* =_{X^*} y^* \wedge x^* \in Ax \to y^* \in Ay\right).$$

(2) The strong resolvent axiom, i.e.

$$\forall x^X, p^X, \gamma^1 \left(\gamma >_{\mathbb{R}} 0 \wedge p =_X \mathrm{Res}^f_{\gamma A} x \to \gamma^{-1}(\nabla f x -_{X^*} \nabla f p) \in Ap\right).$$

(3) Maximal monotonicity of A, i.e.

$$\forall x^X, x^{*X^*} \left(\forall y^X, y^{*X^*} \left(y^* \in Ay \to \langle x -_X y, x^* -_{X^*} y^*\rangle_{X^*} \geqslant_{\mathbb{R}} 0\right) \to x^* \in Ax\right).$$

Proof. For the direction (1) $\Rightarrow$ (3), let x, x^* be such that

$$\langle x - y, x^* - y^*\rangle \geqslant 0 \text{ for all } (y, y^*) \in A.$$

We consider $z = \nabla f^*(x^* + \nabla f x)$. Then

$$1^{-1}(\nabla f z - \nabla f \mathrm{Res}^f_A z) \in A(\mathrm{Res}^f_A z).$$

by axiom $(II)^*$. Thus by the assumption on x, x^*, axiom (L) and the extensionality of $\langle\cdot,\cdot\rangle$ we get

$$\begin{aligned} 0 &\leqslant \langle x - \mathrm{Res}^f_A z, x^* - (\nabla f z - \nabla f \mathrm{Res}^f_A z)\rangle \\ &= \langle x - \mathrm{Res}^f_A z, \nabla f \mathrm{Res}^f_A z - \nabla f x\rangle \end{aligned}$$

which is equivalent to

$$\langle x - \mathrm{Res}^f_A z, \nabla f x - \nabla f \mathrm{Res}^f_A z\rangle \leqslant 0$$

and this yields $x = \mathrm{Res}^f_A z$ as ∇f is (provably) strictly monotone. Further, we have

$$1^{-1}\left(\nabla f z - \nabla f \mathrm{Res}^f_A z\right) = x^* + \nabla f x - \nabla f x = x^*$$

using (L) and the extensionality of ∇f and thus the extensionality of A yields $x^* \in Ax$.

For the direction (3) $\Rightarrow$ (2), assume that $\gamma > 0$ and $p = \mathrm{Res}^f_{\gamma A} x$. Then at first

$$\gamma^{-1}(\nabla f x - \nabla f \mathrm{Res}^f_{\gamma A} x) \in A(\mathrm{Res}^f_{\gamma A} x)$$

by axiom $(II)^*$. By monotonicity (axiom $(III)^*$) together with the extensionality of $\langle\cdot,\cdot\rangle$ and ∇f, we get

$$\forall (y, y^*) \in A(\langle p - y, \gamma^{-1}(\nabla f x - \nabla f p) - y^*\rangle \geqslant 0).$$

By (3), we get

$$\gamma^{-1}(\nabla fx - \nabla fp) \in A(p).$$

For (2) $\Rightarrow$ (1), let $x = y$ and $x^* = y^*$ with $x^* \in Ax$. Define

$$z = \nabla f^*(y^* + \nabla fy).$$

By $(II)^*$, we get

$$1^{-1}(\nabla fz - \nabla f\mathrm{Res}^f_A z) \in A(\mathrm{Res}^f_A z).$$

Axiom $(III)^*$ together with the extensionality of $\langle\cdot,\cdot\rangle$ and ∇f^* as well as using (L) yields

$$\begin{aligned} 0 &\leqslant \langle x - \mathrm{Res}^f_A z, x^* - (\nabla fz - \nabla f\mathrm{Res}^f_A z)\rangle \\ &= \langle y - \mathrm{Res}^f_A z, \nabla f\mathrm{Res}^f_A z - \nabla fy\rangle \end{aligned}$$

and this is equivalent to

$$\langle y - \mathrm{Res}^f_A z, \nabla fy - \nabla f\mathrm{Res}^f_A z\rangle \leqslant 0$$

which yields $y = \mathrm{Res}^f_A z$ by provable strict monotonicity of ∇f. Using (2), we have

$$1^{-1}(\nabla fz - \nabla fy) \in Ay$$

which yields by the quantifier-free extensionality rule that $y^* \in Ay$ as $1^{-1}(\nabla fz - \nabla fy) = y^*$ holds without any additional assumptions. □

Similar to [72] however (see Theorem 3.2 therein), the system $\mathcal{B}^\omega$ does support a weakened, so-called intensional, maximality principle as in the following theorem:

Theorem 4.2. *The system $\mathcal{B}^\omega$ proves the following intensional maximality principle:*

$$\begin{aligned} &\forall x^X, x^{*X^*}\left(\forall y^X, y^{*X^*}\, (y^* \in Ay \to \langle x -_X y, x^* -_{X^*} y^*\rangle_{X^*} \geqslant_{\mathbb{R}} 0)\right. \\ &\qquad \left.\to \exists x'^X, x'^{*X^*}\left(x =_X x' \wedge x^* =_{X^*} x'^* \wedge x'^* \in Ax'\right)\right). \end{aligned}$$

Proof. As in the proof of the direction (1) $\Rightarrow$ (3) from the above Theorem 4.1, we get that

$$1^{-1}(\nabla fz - \nabla f\mathrm{Res}^f_A z) \in A(\mathrm{Res}^f_A z)$$

together with $x = \mathrm{Res}^f_A z$ and $1^{-1}(\nabla fz - \nabla f\mathrm{Res}^f_A z) = x^*$ for $z = \nabla f^*(x^* + \nabla fx)$ without any use of extensionality. This gives the claim by setting $x' = \mathrm{Res}^f_A z$ and $x'^* = 1^{-1}(\nabla fz - \nabla f\mathrm{Res}^f_A z)$. □

We end this section by presenting a result analogous to Theorem 4.1 but phrased specifically for the zero set of the operator A. At first, recall Lemma 3.7, (4), by which the system $\mathcal{B}^\omega$ proves

$$\forall x^X\left(0 \in Ax \to \forall \gamma^1\left(\gamma >_{\mathbb{R}} 0 \to \mathrm{Res}^f_{\gamma A} x =_X x\right)\right).$$

Further, we even provably have

$$\forall x^X, z^{X^*}\left(z \in Ax \wedge z =_{X^*} 0 \to \forall \gamma^1\left(\gamma >_{\mathbb{R}} 0 \to \mathrm{Res}^f_{\gamma A} x =_X x\right)\right).$$

To see this, note that by the quantifier-free extensionality rule we have from $z \in Ax$ that $\gamma^{-1}(\nabla f\nabla f^*(\gamma z + \nabla fx) - \nabla fx) \in Ax$ and thus

$$x =_X \mathrm{Res}^f_{\gamma A}(\nabla f^*(\gamma z + \nabla fx)) =_X \mathrm{Res}^f_{\gamma A} x$$

using uniqueness and extensionality of the resolvent as well as $z = 0$.

As we will see now, the converse assertions are connected to the extensionality of the set of zeros of A.

Theorem 4.3. *Over $\mathcal{B}^\omega$, the following are equivalent:*

(1) $\forall x^X, z^{X^*}\left(\mathrm{Res}^f_A x =_X x \wedge z =_{X^*} 0 \to z \in Ax\right)$,

(2) $\forall x^X, z^{X^*}\left(\forall \gamma^1\left(\gamma >_{\mathbb{R}} 0 \to \mathrm{Res}^f_{\gamma A} x =_X x\right) \wedge z =_{X^*} 0 \to z \in Ax\right)$,

(3) $\forall x^X, y^X, z^{X^*}, z'^{X^*}\left(x =_X y \wedge z =_{X^*} z' =_{X^*} 0 \wedge z \in Ax \to z' \in Ay\right)$.

Proof. The implication from (1) to (2) is clear. For (2) $\Rightarrow$ (3), let $x = y$ and let $z = z' = 0$ with $z \in Ax$. Then using (2) and the provability of

$$\forall x^X, z^{X^*}\left(z \in Ax \wedge z =_{X^*} 0 \to \forall \gamma^1\left(\gamma >_{\mathbb{R}} 0 \to \mathrm{Res}^f_{\gamma A} x =_X x\right)\right)$$

as established above, we get

$$\begin{aligned} z \in Ax &\leftrightarrow \forall \gamma\left(\gamma > 0 \to \mathrm{Res}^f_{\gamma A} x = x\right) \\ &\leftrightarrow \forall \gamma\left(\gamma > 0 \to \mathrm{Res}^f_{\gamma A} y = y\right) \\ &\leftrightarrow z' \in Ay \end{aligned}$$

by the extensionality of $\mathrm{Res}^f_{\gamma A}$ (recall Proposition 3.8). Lastly, for (3) $\Rightarrow$ (1), assume that $\mathrm{Res}^f_A x = x$ and $z = 0$. Then by axiom $(II)^*$, we get

$$1^{-1}(\nabla f x - \nabla f(\mathrm{Res}^f_A x)) \in A(\mathrm{Res}^f_A x).$$

By (3) and $1^{-1}(\nabla f x - \nabla f(\mathrm{Res}^f_A x)) = 0 = z$, we get $z \in Ax$. □

Further, as we will see now, this form of the extensionality of the zero set of A is even equivalent to a corresponding fragment of the maximality principle:

Theorem 4.4. *Over $\mathcal{B}^\omega$, the following are equivalent:*

(1) $\forall x^X, y^X, z^{*X^*}, z'^{*X^*}\left(x =_X y \wedge z^* =_{X^*} z'^* =_{X^*} 0 \wedge z^* \in Ax \to z'^* \in Ay\right)$,

(2) $\forall x^X, z^{*X^*}\left(\forall y^X, y^{*X^*}\left(y^* \in Ay \to \langle x -_X y, -_{X^*} y^*\rangle_{X^*} \geqslant_{\mathbb{R}} 0\right) \wedge z^* =_{X^*} 0 \to z^* \in Ax\right)$.

Proof. For the direction (1) $\Rightarrow$ (2), assume $\forall (y, y^*) \in A\,(\langle x - y, -y^*\rangle \geqslant 0)$. By axiom (II^*), we have $1^{-1}(\nabla f x - \nabla f(\mathrm{Res}^f_A x)) \in A(\mathrm{Res}^f_A x)$. Therefore, we have

$$\langle x - \mathrm{Res}^f_A x, \nabla f x - \nabla f(\mathrm{Res}^f_A x)\rangle \leqslant 0$$

and since ∇f is strictly monotone (recall Lemma 3.2), we get $x = \mathrm{Res}^f_A x$ as well as $1^{-1}(\nabla f x - \nabla f(\mathrm{Res}^f_A x)) = 0$. Thus (1) yields $z^* \in Ax$ for any $z^* = 0$.

Conversely, for (2) $\Rightarrow$ (1), assume that $x = y$ and $z^* = z'^* = 0$ as well as $z^* \in Ax$. Then by $(III)^*$, we have for all $(y, y^*) \in A$:

$$\langle x - y, -y^*\rangle = \langle x - y, z^* - y^*\rangle \geqslant 0.$$

Thus (2) yields that $z'^* \in Ay$. □

Naturally, also the extensionality of the zero set of A as formulated above can not be provable in $\mathcal{B}^\omega$ by virtue of the upcoming bound extraction theorems.

Remark 4.5. Similar results as the above for the zero set of the operator A naturally also hold for the systems and operators considered in [72].

5. A bound extraction theorem

We now prove the central results of this paper, the bound extraction theorem for the theory $\mathcal{B}^\omega$. As discussed before, this bound extraction theorem arises from and extends those for $\mathcal{D}^\omega$ and its extensions proven in [73] as well as those proven in [72] for the systems dealing with monotone operators in Hilbert spaces. In that vein, we keep the proofs short and only briefly discuss the key ingredients, it being implied that all considerations thus abbreviated and that regard the dual space can be made similar to [73] and those that regard the operator A can be made similar to [72].

The high-level outline of the proof of the metatheorem for $\mathcal{B}^\omega$ given here is now very much standard, following that of essentially all other such metatheorems: At first, one extracts realizers from proofs of (essentially) $\forall\exists$-theorems using a combination of Gödel's functional interpretation with a negative translation. The resulting realizers have types from T^{X,X^*} and we then use majorizability to construct bounds with types from T for these realizers, depending only on majorants of the parameters, which are validated in a model based on $\mathcal{M}^{\omega,X,X^*}$, the structure of all strongly majorizable functionals defined using a Banach space X and its dual X^*. In a final step, we can then recover to the truth in a model based on the usual full set-theoretic structure $\mathcal{S}^{\omega,X,X^*}$ if the types occurring in the axioms and the theorem are "low enough" (later called admissible).

We now discuss the first major ingredient of this approach, the functional interpretation and its corresponding soundness result. Originally introduced in Gödel's seminal paper [24], the functional interpretation assigns to every formula $F(\underline{a})$ another formula $F^D(\underline{a}) = \exists\underline{x}\forall\underline{y}F_D(\underline{x},\underline{y},\underline{a})$, where $F_D(\underline{x},\underline{y},\underline{a})$ is quantifier-free (and the length of the tuples $\underline{x},\underline{y}$ and their types depend on the structure of F), via the following recursive definition (where we for simplicity omit the free variables $\underline{a}$):

Definition 5.1 ([24])**.** The Dialectica interpretation $F^D = \exists\underline{x}\forall\underline{y}F_D(\underline{x},\underline{y})$ of a formula F in the language of $\mathcal{A}^\omega[X,\|\cdot\|]$ (and its extensions) is defined via the following recursion on the structure of the formula:

(1) $F^D := F_D := F$ for F being a prime formula.

If $F^D = \exists\underline{x}\forall\underline{y}F_D(\underline{x},\underline{y})$ and $G^D = \exists\underline{u}\forall\underline{v}G_D(\underline{u},\underline{v})$, we set

(2) $(F\wedge G)^D := \exists\underline{x},\underline{u}\forall\underline{y},\underline{v}(F\wedge G)_D$
where $(F\wedge G)_D(\underline{x},\underline{u},\underline{y},\underline{v}) := F_D(\underline{x},\underline{y})\wedge G_D(\underline{u},\underline{v})$,

(3) $(F\vee G)^D := \exists z^0,\underline{x},\underline{u}\forall\underline{y},\underline{v}(F\vee G)_D$
where $(F\vee G)_D(z^0,\underline{x},\underline{u},\underline{y},\underline{v}) := (z=0\to F_D(\underline{x},\underline{y}))\wedge(z\neq 0\to G_D(\underline{u},\underline{v}))$,

(4) $(F\to G)^D := \exists\underline{U},\underline{Y}\forall\underline{x},\underline{v}(F\to G)_D$
where $(F\to G)_D(\underline{U},\underline{Y},\underline{x},\underline{v}) := F_D(\underline{x},\underline{Y}\underline{x}\underline{v})\to G_D(\underline{U}\underline{x},\underline{v})$,

(5) $(\exists z^\tau F(z))^D := \exists z,\underline{x}\forall\underline{y}(\exists z^\tau F(z))_D$
where $(\exists z^\tau F(z))_D(z,\underline{x},\underline{y}) := F_D(\underline{x},\underline{y},z)$,

(6) $(\forall z^\tau F(z))^D := \exists\underline{X}\forall z,\underline{y}(\forall z^\tau F(z))_D$
where $(\forall z^\tau F(z))_D(\underline{X},z,\underline{y}) := F_D(\underline{X}z,\underline{y},z)$.

As mentioned before, we will consider this interpretation in combination with a negative translation that serves to translate from a classical system to its intuitionistic counterpart over which the Dialectica interpretation can then be applied. For that, we fix the following negative translation first defined by Kuroda:

Definition 5.2 ([54])**.** The negative translation of F is defined by $F' := \neg\neg F^*$ where F^* is defined by the following recursion on the structure of F:

(1) $F^* := F$ for prime F;
(2) $(F \circ G)^* := F^* \circ G^*$ for $\circ \in \{\wedge, \vee, \rightarrow\}$;
(3) $(\exists x^\tau F)^* := \exists x^\tau F^*$;
(4) $(\forall x^\tau F)^* := \forall x^\tau \neg\neg F^*$.

The main result for the combination of both of these interpretations is then the following soundness result:

Lemma 5.3 (essentially [35])**.** *Let $\mathcal{P}$ be a set of universal sentences and let $F(\underline{a})$ be an arbitrary formula in the language of $\mathcal{A}^\omega[X, \|\cdot\|]$, the latter with only the variables $\underline{a}$ free. Then the rule*

$$\begin{cases} \mathcal{A}^\omega[X, \|\cdot\|] + \mathcal{P} \vdash F(\underline{a}) \Rightarrow \\ \mathcal{A}^\omega[X, \|\cdot\|]^- + (\mathrm{BR}) + \mathcal{P} \vdash \forall \underline{a}, \underline{y} (F')_D(\underline{t}\underline{a}, \underline{y}, \underline{a}) \end{cases}$$

holds where $\underline{t}$ is a tuple of closed terms of $\mathcal{A}^\omega[X, \|\cdot\|]^- + (\mathrm{BR})$ which can be extracted from the respective proof, (BR) *is the schema of* (simultaneous) bar-recursion *of Spector* [87]*, here extended to the additional abstract types (see e.g.* [36]*) and $\mathcal{A}^\omega[X, \|\cdot\|]^-$ is the respective system* without *the axiom schemes* QF-AC *and* DC.

This result extends to any suitable extension of the language of $\mathcal{A}^\omega[X, \|\cdot\|]$ (e.g. by any kind of new types and constants) together with any number of additional universal axioms in that (extended) language.

We now introduce the essential type and formula classes for the formulation of the metatheorem (essentially following [23, 35]): We call a type ξ *of degree* n if $\xi \in T$ and it has degree $\leqslant n$ in the usual sense (see e.g. [36]). We call ξ *small* if it is of the form $\xi = \xi_0(0) \dots (0)$ for $\xi_0 \in \{0, X, X^*\}$ (including $0, X, X^*$) and call it *admissible* if it is of the form $\xi = \xi_0(\tau_k) \dots (\tau_1)$ where each τ_i is small and $\xi_0 \in \{0, X, X^*\}$ (also including $0, X, X^*$). Correspondingly, a formula F is called a $\forall$-formula (respectively $\exists$-formula) if $F = \forall \underline{a}^{\underline{\xi}} F_{qf}(\underline{a})$ (respectively $F = \exists \underline{a}^{\underline{\xi}} F_{qf}(\underline{a})$) with F_{qf} quantifier-free and all types ξ_i in $\underline{\xi} = (\xi_1, \dots, \xi_k)$ are admissible.

As mentioned before, besides Gödel's functional interpretation, the other central notion used in the bound extraction result is that of majorizability, going back to the fundamental work of Howard [27]. Besides of Howard's notion, we here also in particular rely on the subsequently extended notion of strong majorizability due to Bezem [9] which crucially, as shown by Bezem, provides a model of bar-recursion. The corresponding type structure of all strongly majorizable functionals, suitably extended to new abstract types, also forms the basis for the modern bound extraction theorems of proof mining as developed in [23, 35] and subsequent papers and thus also takes a central role here. We therefore discuss this model explicitly for our theory $\mathcal{B}^\omega$ based on the types T^{X,X^*} (essentially defined as in [73], but practically just derived from [23, 35]): At first, majorants of objects with types from T^{X,X^*} are objects with types from T according to the following projection:

Definition 5.4 ([73], essentially [23])**.** Define $\widehat{\tau} \in T$, given $\tau \in T^{X,X^*}$, by recursion on the structure via

$$\widehat{0} := 0,\ \widehat{X} := 0,\ \widehat{X^*} := 0,\ \widehat{\tau(\xi)} := \widehat{\tau}(\widehat{\xi}).$$

The majorizability relation for the types T^{X,X^*} and the accompanying structure $\mathcal{M}^{\omega,X,X^*}$ of all (strongly) majorizable functionals over a given normed space X with dual X^* is then defined recursively as follows:

Definition 5.5 ([73], essentially [23, 35])**.** Let $(X, \|\cdot\|)$ be a non-empty normed space with dual X^*. The structure $\mathcal{M}^{\omega,X,X^*}$ and the majorizability relation $\gtrsim_\rho$ are defined by

$$\begin{cases} M_0 := \mathbb{N}, n \gtrsim_0 m := n \geqslant m \wedge n, m \in \mathbb{N}, \\ M_X := X, n \gtrsim_X x := n \geqslant \|x\| \wedge n \in M_0, x \in M_X, \\ M_{X^*} := X^*, n \gtrsim_{X^*} x^* := n \geqslant \|x^*\| \wedge n \in M_0, x^* \in M_{X^*}, \\ f \gtrsim_{\tau(\xi)} x := f \in M_{\hat{\tau}}^{M_{\hat{\xi}}} \wedge x \in M_\tau^{M_\xi} \\ \qquad\qquad\qquad \wedge \forall g \in M_{\hat{\xi}}, y \in M_\xi (g \gtrsim_\xi y \to fg \gtrsim_\tau xy) \\ \qquad\qquad\qquad \wedge \forall g, y \in M_{\hat{\xi}} (g \gtrsim_{\hat{\xi}} y \to fg \gtrsim_{\hat{\tau}} fy), \\ M_{\tau(\xi)} := \left\{ x \in M_\tau^{M_\xi} \mid \exists f \in M_{\hat{\tau}}^{M_{\hat{\xi}}} \left(f \gtrsim_{\tau(\xi)} x \right) \right\}. \end{cases}$$

Correspondingly, the full set-theoretic type structure $\mathcal{S}^{\omega,X,X^*}$ is defined via $S_0 := \mathbb{N}$, $S_X := X$, $S_{X^*} := X^*$ and

$$S_{\tau(\xi)} := S_\tau^{S_\xi}.$$

These structures later turn into models of the system $\mathcal{B}^\omega$ when they are equipped with suitable corresponding interpretations for the additional constants.

This, for modern proof mining essential and fundamental, idea of combining the Dialectica interpretation with the notion of majorizability goes back to [32] and was subsequently formally encapsulated through a novel type of functional interpretation, the so-called monotone functional interpretation, in [33]. In that way, also the present approach is thus "in spirit" of the monotone functional interpretation, where we write "in spirit" as we actually do not use a monotone variant of the Dialectica interpretation but treat the functional interpretation part and the majorization part of the combined interpretation separately, following the formal approach to proof mining metatheorems of [23, 35].

In any way, as is crucial for the applicability of this approach, many highly nontrivial mathematical statements nevertheless have a very simple monotone functional interpretation and one class of such principles is the so-called class Δ as originally introduced in [31, 32] (and then lifted to abstract types in [25]): a formula of type Δ is any formula of the form

$$\forall \underline{a}^{\underline{\delta}} \exists \underline{b} \leqslant_{\underline{\sigma}} \underline{r}\underline{a} \forall \underline{c}^{\underline{\gamma}} F_{qf}(\underline{a}, \underline{b}, \underline{c})$$

where F_{qf} is quantifier-free, the types in $\underline{\delta}$, $\underline{\sigma}$ and $\underline{\gamma}$ are admissible, $\underline{r}$ is a tuple of closed terms of appropriate type and where $\leqslant$ is defined by recursion on the type via

1. $x \leqslant_0 y := x \leqslant_0 y$,
2. $x \leqslant_X y := \|x\|_X \leqslant_\mathbb{R} \|y\|_X$,
3. $x^* \leqslant_{X^*} y^* := \|x^*\|_{X^*} \leqslant_\mathbb{R} \|y^*\|_{X^*}$,
4. $x \leqslant_{\tau(\xi)} y := \forall z^\xi (xz \leqslant_\tau yz)$,

and with $\underline{x} \leqslant_{\underline{\sigma}} \underline{y}$ being an abbreviation for $x_1 \leqslant_{\sigma_1} y_1 \wedge \cdots \wedge x_k \leqslant_{\sigma_k} y_k$ where $\underline{x}$, $\underline{y}$ and $\underline{\sigma}$ are k-tuples of terms and types, respectively, such that x_i, y_i are of type σ_i.

Since we treat the majorization part separately from the functional interpretation part, we need a slightly modified approach to the treatment of statements of type Δ (where we essentially follow the approach given in [25], see also the recent [73]) and treat axioms of type Δ by employing a construction that converts a theory with axioms of such a type into a theory using only additional purely universal axioms formulated using the Skolem functions of these axioms. Concretely, we proceed as follows: Write $\widehat{\mathcal{B}}^\omega$ for $\mathcal{B}^\omega$ without any of its axioms of type Δ

(which in particular covers all of the previous non-universal axioms $(*)_2$, $(*)_6$, $(f^*)_2$, $(f^**)_2$) and without the rule (QF-LR) (which produces conclusions of type Δ). Given a set $\beth$ of sentences of type Δ, we transform $\mathcal{B}^\omega + \beth$ into a new theory $\overline{\mathcal{B}}^\omega_\beth$ by adding to $\widehat{\mathcal{B}}^\omega$ the Skolem functionals $\underline{B}$ for any axiom of type Δ, say of the form

$$\forall \underline{a}^{\underline{\delta}} \exists \underline{b} \leqslant_{\underline{\sigma}} \underline{r}\underline{a} \forall \underline{c}^{\underline{\gamma}} F_{qf}(\underline{a}, \underline{b}, \underline{c}),$$

as new constants to the language and adding its 'instantiated Skolem normal form', i.e. the sentence

$$\underline{B} \leqslant_{\underline{\sigma}(\underline{\delta})} \underline{r} \wedge \forall \underline{a}^{\underline{\delta}} \forall \underline{c}^{\underline{\gamma}} F_{qf}(\underline{a}, \underline{B}\underline{a}, \underline{c})$$

as a new axiom. Further, we do the same with all conclusions of the rule (QF-LR): for any provable premise

$$\mathcal{B}^\omega + \beth \vdash F_{qf} \to \left(\forall x^X, y^X, \alpha^1, \beta^1\, (t(\alpha x +_X \beta y) =_{\mathbb{R}} \alpha t x + \beta t y) \wedge \forall x^X\, (|tx| \leqslant_{\mathbb{R}} M \left\| x \right\|_X)\right)$$

with terms t and M, we add a new constant $x^*_{t,M}$ of type X^* to the language of $\overline{\mathcal{B}}^\omega_\beth$ together with the corresponding axiom

$$\left\| x^*_{t,M} \right\|_{X^*} \leqslant_{\mathbb{R}} M \wedge \left(F_{qf} \to \forall x^X \left(tx =_{\mathbb{R}} \langle x, x^*_{t,M} \rangle_{X^*}\right)\right).$$

Now, note that this new theory $\overline{\mathcal{B}}^\omega_\beth$ extends $\mathcal{A}^\omega[X, \|\cdot\|]$ only by new types, constants and universal axioms. Therefore, Lemma 5.3 extends to this theory $\overline{\mathcal{B}}^\omega_\beth$ where the conclusion is proved in $\overline{\mathcal{B}}^{\omega-}_\beth$ + (BR) where $\overline{\mathcal{B}}^{\omega-}_\beth$ is again the same theory with the principles (QF-AC) and (DC) removed.

The main result on the majorizability part of the metatheorem is now the following model-theoretic lemma:

Lemma 5.6. *Let $\beth$ be a set of additional sentences of type Δ. Let $(X, \|\cdot\|)$ be a (nontrivial) Banach space with its dual X^*. Let f be a convex, supercoercive and Fréchet differentiable function $f : X \to \mathbb{R}$ where ∇f, ∇f^* are uniformly continuous on bounded subsets and where f^* is supercoercive. Let A be a monotone operator $A : X \to 2^{X^*}$ with $A^{-1}0 \neq \emptyset$ such that the corresponding resolvents $\mathrm{Res}^f_{\gamma A}$ are total for $\gamma > 0$.*

Then $\mathcal{M}^{\omega,X,X^*}$ is a model of $\overline{\mathcal{B}}^{\omega-}_\beth$ + (BR), provided $\mathcal{S}^{\omega,X,X^*} \models \beth$ (with $\mathcal{M}^{\omega,X,X^*}$ and $\mathcal{S}^{\omega,X,X^*}$ defined via suitable interpretations of the additional constants). Moreover, for any closed term t of $\overline{\mathcal{B}}^{\omega-}_\beth$ + (BR), one can construct a closed term t^* of $\mathcal{A}^\omega$ + (BR) such that

$$\mathcal{M}^{\omega,X,X^*} \models \forall \omega^{0(0)(0)}, \omega'^{0(0)}, n^0 \Big(\omega \gtrsim \omega^{\nabla f}, \omega^{\nabla f^*} \wedge \omega' \gtrsim \alpha^f, \alpha^{f^*}$$
$$\wedge\, n \geqslant_{\mathbb{R}} |f(0)|, \left\|\nabla f(0)\right\|_{X^*}, |f^*(0)|, \left\|\nabla f^*(0)\right\|_X, \left\|p_X\right\|_X \to t^*(\omega, \omega', n) \gtrsim t\Big).$$

Proof. The proof is completely analogous to that of Lemma 7.7 in [73] (α^{f^*} can be treated as α^f in [73]) after making the necessary additions for the treatment of A and its resolvents and zero which can be made similar as in the proof of Lemma 5.3 in [72]. But, for a more comprehensive and self-contained proof, we briefly detail these additions here: Concretely, we have to give interpretations of the additional constants χ_A, Res^f_A and p_X concerning A that were added to the system $\mathcal{D}^\omega_{f,f^*}$[FM] to form $\mathcal{B}^\omega$ and then verify that these interpretations indeed have majorants to check that they belong to $\mathcal{M}^{\omega,X,X^*}$. The interpretations are immediately defined as follows (writing $\mathcal{M}$ as a shorthand for $\mathcal{M}^{\omega,X,X^*}$):

(1) $[\chi_A]_{\mathcal{M}} := \lambda x \in X, x^* \in X^*. \begin{cases} 0 & \text{if } x^* \in Ax, \\ 1 & \text{otherwise,} \end{cases}$

(2) $[\mathrm{Res}^f_A]_{\mathcal{M}} := \lambda\alpha \in \mathbb{N}^{\mathbb{N}}, x \in X. \begin{cases} \mathrm{Res}^f_{r_\alpha A} x & \text{if } r_\alpha > 0, \\ 0 & \text{otherwise,} \end{cases}$

(3) $[p_X]_{\mathcal{M}} := p$,

where $p \in A^{-1}0$ is some zero of A as fixed in the formulation of the lemma. Take n, ω and ω' such that $\omega \gtrsim \omega^{\nabla f}, \omega^{\nabla f^*}$ and $\omega' \gtrsim \alpha^f, \alpha^{f^*}$ as well as $n \geqslant |f(0)|, \|\nabla f(0)\|, |f^*(0)|, \|\nabla f^*(0)\|, \|p\|$. Corresponding majorants for (1) and (3) are then easily given via

(1)' $\lambda x^0, y^0.1^0 \gtrsim [\chi_A]_{\mathcal{M}}$,

(3)' $n \gtrsim [p_X]_{\mathcal{M}}$.

We now focus on item (2). For this, take $b \gtrsim x$ as well as $\alpha \gtrsim \gamma$ (although the majorant will be independent of α here). By Proposition 3.8, we have

$$\left\|\mathrm{Res}^f_{\gamma A} x\right\| \leqslant o(2D(b_n) + 2b_n C(b_n), b_n)$$

where $b_n = \max\{b, n\} + 1$ for the fixed n, where

$$o(\alpha, b) := F(\alpha^{f^*}(\alpha + D(b) + b) + 1),$$

is defined as in Lemma 3.4 and

$$\begin{aligned} C(b) &:= b2^{\omega^{\nabla f}(0,b)} + [\|\nabla f(0)\|](0) + 2, \\ F(b) &:= b2^{\omega^{\nabla f^*}(0,b)} + [\|\nabla f^*(0)\|](0) + 2, \\ D(b) &:= b2^{\omega^f(0,b)} + [|f(0)|](0) + 2, \end{aligned}$$

are defined as in Lemma 2.15. Since we have

$$C(b), F(b), D(b) \leqslant b2^{\omega(0,b)} + n + 3 =: E(b)$$

it immediately follows that

$$\begin{aligned} \left\|\mathrm{Res}^f_{\gamma A} x\right\|_X &\leqslant o(2D(b_n) + 2bC(b_n), b_n) \\ &\leqslant E(\omega'(3E(b_n) + 2b_n E(b_n) + b_n) + 1) \\ &\leqslant E(\omega'(6b_n E(b_n)) + 1) \end{aligned}$$

and therefore we obtain

$$\lambda b^0.E(\omega'(6b_n E(b_n)) + 1) \gtrsim [\mathrm{Res}^f_A]_{\mathcal{M}}.$$

□

Theorem 5.7. *Let τ be admissible, δ be of degree 1 and s be a closed term of $\mathcal{B}^\omega$ of type $\sigma(\delta)$ for admissible σ. Let $\beth$ be a set of sentences of type Δ. Let $F_\forall(x, y, z, u)/G_\exists(x, y, z, v)$ be $\forall$-/$\exists$-formulas of $\mathcal{B}^\omega$ with only $x, y, z, u/x, y, z, v$ free. If*

$$\mathcal{B}^\omega + \beth \vdash \forall x^\delta \forall y \leqslant_\sigma s(x) \forall z^\tau \left(\forall u^0 F_\forall(x, y, z, u) \to \exists v^0 G_\exists(x, y, z, v)\right),$$

then one can extract a partial functional $\Phi : \mathbb{N}^{\mathbb{N}\times\mathbb{N}} \times \mathbb{N}^{\mathbb{N}} \times \mathbb{N} \times S_\delta \times S_{\hat{\tau}} \rightharpoonup \mathbb{N}$ which is total and (bar-recursively) computable on $\mathbb{N}^{\mathbb{N}\times\mathbb{N}} \times \mathbb{N}^{\mathbb{N}} \times \mathbb{N} \times M_\delta \times M_{\hat{\tau}}$ such that for all $x \in S_\delta$, $z \in S_\tau$, $z^* \in S_{\hat{\tau}}$ with $z^* \gtrsim z$ and for all $\omega \in \mathbb{N}^{\mathbb{N}\times\mathbb{N}}$, $\omega' \in \mathbb{N}^{\mathbb{N}}$, $n \in \mathbb{N}$ with $\omega \gtrsim \omega^{\nabla f}, \omega^{\nabla f^*}$, $\omega' \gtrsim \alpha^f, \alpha^{f^*}$ and $n \geqslant |f(0)|, \|\nabla f(0)\|, |f^*(0)|, \|\nabla f^*(0)\|, \|p\|$:

$$\begin{aligned} \mathcal{S}^{\omega,X,X^*} \models \forall y \leqslant_\sigma s(x)(&\forall u \leqslant_0 \Phi(\omega, \omega', n, x, z^*) F_\forall(x, y, z, u) \\ &\to \exists v \leqslant_0 \Phi(\omega, \omega', n, x, z^*) G_\exists(x, y, z, v)) \end{aligned}$$

holds whenever $\mathcal{S}^{\omega,X,X^*} \models \beth$ for $\mathcal{S}^{\omega,X,X^*}$ defined (via a suitable interpretation of the additional constants similar to [73] *and* [72], *recall Lemma 5.6) using any (nontrivial) reflexive Banach* *space* $(X, \|\cdot\|)$ *with its dual* X^* *and using a convex, supercoercive (with modulus* α^f*) and Fréchet differentiable function* $f : X \to \mathbb{R}$ *where* ∇f, ∇f^* *are uniformly continuous on bounded subsets (with moduli* $\omega^{\nabla f}, \omega^{\nabla f^*}$*, respectively) and where* f^* *is supercoercive (with modulus* α^{f^*}*) and using a monotone operator* $A : X \to 2^{X^*}$ *with* $p \in A^{-1}0 \neq \varnothing$ *such that the corresponding resolvents* $\mathrm{Res}^f_{\gamma A}$ *are all total for* $\gamma > 0$.

Further:

1. *If* $\widehat{\tau}$ *is of degree* 1*, then* Φ *is a total computable functional.*
2. *We may have tuples instead of single variables* x, y, z, u, v *and a finite conjunction instead of a single premise* $\forall u^0 F_\forall(x, y, z, u)$.
3. *If the claim is proved without* DC*, then* τ *may be arbitrary and* Φ *will be a total functional on* $\mathbb{N}^{\mathbb{N}\times\mathbb{N}} \times \mathbb{N}^{\mathbb{N}} \times \mathbb{N} \times S_\delta \times S_{\widehat{\tau}}$ *which is primitive recursive in the sense of Gödel* [24] *and Hilbert* [26]*. In that case, also plain majorization can be used instead of strong majorization.*

Proof. The proof now is completely analogous to that of Theorem 7.8 in [73] and we just briefly sketch it: Assuming the premise, one first realizes that the same statement is also provable in $\overline{\mathcal{B}}^\omega_{\daleth}$. Using the functional interpretation, i.e. Lemma 5.3, for this system, one extracts realizers for (the $\forall\exists$-prenexiation of) this statement which are then majorized in the model based on $\mathcal{M}^{\omega,X,X^*}$ through Lemma 5.6. Utilizing that the types in the statement are admissible, one can then in turn conclude that the resulting bounds are also valid in the model based on $\mathcal{S}^{\omega,X,X^*}$. The additional statements (1) – (3) can also be concluded similarly as in [73]. □

Acknowledgments: I want to thank Ulrich Kohlenbach for many helpful discussions on the topic of this paper. The results of this paper form the main part of Chapter 10 of my doctoral dissertation [74] written under his supervision. The author was supported by the 'Deutsche Forschungsgemeinschaft' Project DFG KO 1737/6-2.

References

[1] Y.I. Alber. Metric and generalized projections in banach spaces: properties and applications. In A.G. Kartsatos, editor, *Theory and Applications of Nonlinear Operators of Accretive and Monotone Type*, pages 15–50. Marcel Dekker, New York, 1996.

[2] H.H. Bauschke. The composition of projections onto closed convex sets in Hilbert space is asymptotically regular. *Proceedings of the American Mathematical Society*, 131:141–146, 2003.

[3] H.H. Bauschke and J.M. Borwein. Legendre functions and the method of random Bregman projections. *Journal of Convex Analysis*, 4(1):27–67, 1997.

[4] H.H. Bauschke, J.M. Borwein, and P.L. Combettes. Essential smoothness, essential strict convexity, and Legendre functions in Banach spaces. *Communications in Contemporary Mathematics*, 3:615–647, 2001.

[5] H.H. Bauschke, J.M. Borwein, and P.L. Combettes. Bregman monotone optimization algorithms. *SIAM Journal on Control and Optimization*, 42:596–636, 2003.

[6] H.H. Bauschke and P.L. Combettes. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces.* CMS Books in Mathematics. Springer, Cham, 2nd edition, 2017.

[7] H.H. Bauschke and A.S. Lewis. Dykstra's algorithm with Bregman projections: A convergence proof. *Optimization*, 48(4):409–427, 2000.

[8] H.H. Bauschke, X. Wang, and L. Yao. General resolvents for monotone operators: characterization and extension. In Y. Censor, M. Jiang, and G. Wang, editors, *Biomedical Mathematics: Promising Directions in Imaging, Therapy Planning and Inverse Problems*, pages 57–74. Medical Physics Publishing, Madison, WI, USA, 2010.

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

[10] R.I. Boţ, S.-M. Grad, and G. Wanka. *Duality in Vector Optimization.* Vector Optimization. Springer Berlin, Heidelberg, 2009.

[11] J.M. Borwein, A.J. Guirao, P. Hájek, and J. Vanderwerff. Uniformly convex functions on Banach spaces. *Proceedings of the American Mathematical Society*, 137(3):1081–1091, 2009.

[12] J.M. Borwein and J. Vanderwerff. Fréchet-Legendre functions and reflexive Banach spaces. *Journal of Convex Analysis*, 17(3–4):915–924, 2010.

[13] L.M. Bregman. The relaxation method for finding the common point of convex sets and its application to the solution of problems in convex programming. *USSR Computational Mathematics and Mathematical Physics*, 7:200–217, 1967.

[14] F.E. Browder. Nonlinear monotone operators and convex sets in Banach spaces. *Bulletin of the American Mathematical Society*, 71(5):780–785, 1965.

[15] F.E. Browder. Nonlinear maximal monotone operators in Banach space. *Mathematische Annalen*, 175:89–113, 1968.

[16] Y. Censor and T. Elfving. A multiprojection algorithm using Bregman projections in a product space. *Numerical Algorithms*, 8:221–239, 1994.

[17] Y. Censor and S. Reich. The Dykstra algorithm with Bregman projections. *Communications in Applied Analysis*, 2(3):407–420, 1998.

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

[19] B. Dinis and P. Pinto. Quantitative results on the multi-parameters proximal point algorithm. *Journal of Convex Analysis*, 28(3), 2021. 23 pp.

[20] J. Eckstein. Nonlinear Proximal Point Algorithms Using Bregman Functions, with Applications to Convex Programming. *Mathematics of Operations Research*, 18(1):202–226, 1993.

[21] F. Ferreira, L. Leuştean, and P. Pinto. On the removal of weak compactness arguments in proof mining. *Advances in Mathematics*, 354, 2019. 106728.

[22] R. Findling and U. Kohlenbach. Rates of convergence and metastability for Chidume's algorithm for the approximation of zeros of accretive operators in Banach spaces. *Numerical Functional Analysis and Optimization*, 45(3):216–233, 2024.

[23] P. Gerhardy and U. Kohlenbach. General logical metatheorems for functional analysis. *Transactions of the American Mathematical Society*, 360:2615–2660, 2008.

[24] K. Gödel. Über eine bisher noch nicht benützte Erweiterung des finiten Standpunktes. *Dialectica*, 12:280–287, 1958.

[25] D. Günzel and U. Kohlenbach. Logical metatheorems for abstract spaces axiomatized in positive bounded logic. *Advances in Mathematics*, 290:503–551, 2016.

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

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

[28] W.A. Howard. The formulae-as-types notion of construction. In J.R. Hindley and J.P. Seldin, editors, *To H.B. Curry*, pages 479–490. Academic, London, 1980.

[29] S. Kamimura, F. Kohsaka, and W. Takahashi. Weak and strong convergence theorems for maximal monotone operators in a Banach space. *Set-Valued Analysis*, 12:417–429, 2004.

[30] T. Kato. Nonlinear semigroups and evolution equations. *Journal of the Mathematical Society of Japan*, 19:508–520, 1967.

[31] U. Kohlenbach. *Theory of majorizable and continuous functionals and their use for the extraction of bounds from non-constructive proofs: effective moduli of uniqueness for best approximations from ineffective proofs of uniqueness.* PhD thesis, Goethe-Universität Frankfurt am Main, 1990.

[32] U. Kohlenbach. Effective bounds from ineffective proofs in analysis: an application of functional interpretation and majorization. *The Journal of Symbolic Logic*, 57:1239–1273, 1992.

[33] 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*, pages 225–260. Oxford University Press, 1996.

[34] U. Kohlenbach. Arithmetizing proofs in analysis. In J.M. Larrazabal, D. Lascar, and G. Mints, editors, *Logic Colloquium '96*, volume 12 of *Lecture Notes in Logic*, pages 115–158. Cambridge University Press, 1998.

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

[36] U. Kohlenbach. *Applied Proof Theory: Proof Interpretations and their Use in Mathematics.* Springer Monographs in Mathematics. Springer-Verlag Berlin Heidelberg, 2008.

[37] U. Kohlenbach. Recent progress in proof mining in nonlinear analysis. *IFCoLog Journal of Logics and their Applications*, 10(4):3361–3410, 2017.

[38] U. Kohlenbach. A polynomial rate of asymptotic regularity for compositions of projections in Hilbert space. *Foundations of Computational Mathematics*, 19:83–99, 2019.

[39] U. Kohlenbach. Proof-theoretic Methods in Nonlinear Analysis. In B. Sirakov, P. Ney de Souza, and M. Viana, editors, *Proc. ICM 2018*, volume 2, pages 61–82. World Scientific, 2019.

[40] U. Kohlenbach. Local formalizations in nonlinear analysis and related areas and proof-theoretic tameness. In P. Weingartner and H.-P. Leeb, editors, *Kreisel's Interests. On the Foundations of Logic and Mathematics*, volume 41 of *Tributes*, pages 45–61. College Publications, 2020.

[41] U. Kohlenbach. Quantitative analysis of a Halpern-type Proximal Point Algorithm for accretive operators in Banach spaces. *Journal of Nonlinear and Convex Analysis*, 21(9):2125–2138, 2020.

[42] U. Kohlenbach. Quantitative results on the Proximal Point Algorithm in uniformly convex Banach spaces. *Journal of Convex Analysis*, 28(1):11–18, 2021.

[43] U. Kohlenbach. On the Proximal Point Algorithm and its Halpern-type variant for generalized monotone operators in Hilbert space. *Optimization Letters*, 16:611–621, 2022.

[44] U. Kohlenbach and A. Koutsoukou-Argyraki. Rates of convergence and metastability for abstract Cauchy problems generated by accretive operators. *Journal of Mathematical Analysis and Applications*, 423:1089–1112, 2015.

[45] U. Kohlenbach, L. Leuştean, and A. Nicolae. Quantitative Results on Fejér Monotone Sequences. *Communications in Contemporary Mathematics*, 20(2):42pp., 2018.

[46] U. Kohlenbach and A. Nicolae. A proof-theoretic bound extraction theorem for CAT(κ)-spaces. *Studia Logica*, 105:611–624, 2017.

[47] U. Kohlenbach and P. Oliva. Proof Mining: A Systematic Way of Analysing Proofs in Mathematics. *Proceedings of the Steklov Institute of Mathematics*, 242:136–164, 2003.

[48] U. Kohlenbach and N. Pischke. Proof theory and nonsmooth analysis. *Philosophical Transactions of the Royal Society A*, 381, 2023. 20220015, 21pp.

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

[50] F. Kohsaka and W. Takahashi. Strong convergence of an iterative sequence for maximal monotone operators in a Banach space. *Abstract and Applied Analysis*, 3:239–249, 2004.

[51] G. Kreisel. On the Interpretation of Non-Finitist Proofs–Part I. *The Journal of Symbolic Logic*, 16(4):241–267, 1951.

[52] G. Kreisel. On the Interpretation of Non-Finitist Proofs–Part II. Interpretation of Number Theory. Applications. *The Journal of Symbolic Logic*, 17(1):43–58, 1952.

[53] G. Kreisel. Interpretation of analysis by means of constructive functionals of finite types. In A. Heyting, editor, *Constructivity in Mathematics*, Lecture Notes in Mathematics, pages 101–128. North-Holland Publishing Company, Amsterdam, 1959.

[54] S. Kuroda. Intuitionistische Untersuchungen der formalistischen Logik. *Nagoya Mathematical Journal*, 2:35–47, 1951.

[55] L. Leuştean. Proof Mining in $\mathbb{R}$-trees and Hyperbolic Spaces. *Electronic Notes in Theoretical Computer Science*, 165:95–106, 2006.

[56] L. Leuştean. An application of proof mining to nonlinear iterations. *Annals of Pure and Applied Logic*, 165(9):1484–1500, 2014.

[57] L. Leuştean, A. Nicolae, and A. Sipoş. An abstract proximal point algorithm. *Journal of Global Optimization*, 72(3):553–577, 2018.

[58] L. Leuştean and P. Pinto. Quantitative results on the Halpern type proximal point algorithm. *Computational Optimization and Applications*, 79(1):101–125, 2021.

[59] A. Macintyre. The mathematical significance of proof theory. *Philosophical Transactions of the Royal Society A*, 363:2419–2435, 2005.

[60] A. Macintyre. The impact of Gödel's incompleteness theorems on mathematics. In M. Baaz, C.H. Papadimitriou, H.W. Putnam, D.S. Scott, and jr. C.L. Harper, editors, *Kurt Gödel and the foundations of mathematics: Horizons of truth*, pages 3–25. Cambridge University Press, 2011.

[61] V. Martín-Márquez, S. Reich, and S. Sabach. Iterative methods for approximating fixed points of Bregman nonexpansive operators. *Discrete and Continuous Dynamical Systems-Series S*, 6(4):1043–1063, 2013.

[62] B. Martinet. Régularisation d'inéquations variationnelles par approximations successives. *Revue française d'informatique et de recherche opérationnelle*, 4:154–158, 1970.
[63] G.J. Minty. Monotone networks. *Proceedings of the Royal Society A*, 257:194–212, 1960.
[64] G.J. Minty. Monotone (nonlinear) operators in Hilbert spaces. *Duke Mathematical Journal*, 29:341–346, 1962.
[65] M. Neri and N. Pischke. Proof mining and probability theory. 2024. 47pp., submitted manuscript available at `https://arxiv.org/abs/2403.00659`.
[66] L. Paunescu and A. Sipoş. A proof-theoretic metatheorem for tracial von Neumann algebras. 2022. To appear in Mathematical Logic Quarterly.
[67] P. Pinto. A rate of metastability for the Halpern type Proximal Point Algorithm. *Numerical Functional Analysis and Optimization*, 42(3):320–343, 2021.
[68] P. Pinto and N. Pischke. On Computational Properties of Cauchy Problems generated by Accretive Operators. *Documenta Mathematica*, 28(5):1235–1274, 2023.
[69] N. Pischke. A proof-theoretic metatheorem for nonlinear semigroups generated by an accretive operator and applications. 2023. 42pp., submitted manuscript available at `https://arxiv.org/abs/2304.01723`.
[70] N. Pischke. Duality, Fréchet differentiability and Bregman distances in hyperbolic spaces. *Israel Journal of Mathematics*, 2024. To appear.
[71] N. Pischke. Generalized Fejér monotone sequences and their finitary content. *Optimization*, 2024. to appear, 68pp., doi:10.1080/02331934.2024.2390114.
[72] N. Pischke. Logical Metatheorems for Accretive and (Generalized) Monotone Set-Valued Operators. *Journal of Mathematical Logic*, 24(2), 2024. 2350008, 59pp.
[73] N. Pischke. Proof mining for the dual of a Banach space with extensions for uniformly Fréchet differentiable functions. *Transactions of the American Mathematical Society*, 2024. to appear, 43pp., doi:10.1090/tran/9226.
[74] N. Pischke. *Proof-Theoretical Aspects of Nonlinear and Set-Valued Analysis*. PhD thesis, TU Darmstadt, 2024. Thesis available at `https://doi.org/10.26083/tuprints-00026584`.
[75] N. Pischke. Rates of convergence for the asymptotic behavior of second-order Cauchy problems. *Journal of Mathematical Analysis and Applications*, 533(2), 2024. 128078.
[76] N. Pischke and U. Kohlenbach. Effective rates for iterations involving Bregman strongly nonexpansive operators. 2024. 42pp., submitted manuscript available at `https://www2.mathematik.tu-darmstadt.de/~kohlenbach/`.
[77] S. Reich and S. Sabach. A strong convergence theorem for a proximal-type algorithm in reflexive Banach spaces. *Journal of Nonlinear and Convex Analysis*, 10:471–485, 2009.
[78] S. Reich and S. Sabach. Two strong convergence theorems for a proximal method in reflexive Banach spaces. *Numerical Functional Analysis and Optimization*, 31(1):22–44, 2010.
[79] R.T. Rockafellar. On the maximal monotonicity of subdifferential mappings. *Pacific Journal of Mathematics*, 33(1):209–216, 1970.
[80] R.T. Rockafellar. Monotone operators and the proximal point algorithm. *SIAM Journal on Control and Optimization*, 14:877–898, 1976.
[81] R.T. Rockafellar and R.J.-B. Wets. *Variational Analysis*. Springer, New York, 1998.
[82] M. Schönfinkel. Über die Bausteine der mathematischen Logik. *Mathematische Annalen*, 92:305–316, 1924.
[83] Y. Shehu. Convergence Results of Forward-Backward Algorithms for Sum of Monotone Operators in Banach Spaces. *Results in Mathematics*, 74:Article number 138, 2019.
[84] A. Sipoş. Proof mining in L^p spaces. *The Journal of Symbolic Logic*, 84(4):1612–1629, 2019.
[85] A. Sipoş. Quantitative inconsistent feasibility for averaged mappings. *Optimization Letters*, 16(6):1915–1925, 2022.
[86] A. Sipoş. The computational content of super strongly nonexpansive mappings and uniformly monotone operators. *Israel Journal of Mathematics*, 2023. to appear, 17pp., arXiv:2303.02768.
[87] C. Spector. Provably recursive functionals of analysis: a consistency proof of analysis by an extension of principles formulated in current intuitionistic mathematics. In J.C.E. Dekker, editor, *Recursive Function Theory*, volume 5 of *Proceedings of Symposia in Pure Mathematics*, pages 1–27. American Mathematical Society, 1962.
[88] A.S. Troelstra, editor. *Metamathematical Investigation of Intuitionistic Arithmetic and Analysis*, volume 344 of *Lecture Notes in Mathematics*. Springer, New York, 1973.
[89] C. Zălinescu. *Convex analysis in general vector spaces*. World scientific, 2002.

[90] H. Zegeye. Strong convergence theorems for maximal monotone mappings in Banach spaces. *Journal of Mathematical Analysis and Applications*, 343(2):663–671, 2008.