NEW CONSTRAINT QUALIFICATIONS FOR OPTIMIZATION PROBLEMS IN BANACH SPACES BASED ON ASYMPTOTIC KKT CONDITIONS[*]

EIKE BÖRGENS[†], CHRISTIAN KANZOW[†], PATRICK MEHLITZ[‡], AND
GERD WACHSMUTH[†]

Abstract. Optimization theory in Banach spaces suffers from a lack of available constraint qualifications. There exist very few constraint qualifications, and these are often violated even in simple applications. This is very much in contrast to finite-dimensional nonlinear programs, where a large number of constraint qualifications is known. Since these constraint qualifications are usually defined using the set of active inequality constraints, it is difficult to extend them to the infinite-dimensional setting. One exception is a recently introduced sequential constraint qualification based on *asymptotic* KKT conditions. This paper shows that this so-called asymptotic KKT regularity allows suitable extensions to the Banach space setting in order to obtain new constraint qualifications. The relation of these new constraint qualifications to existing ones is discussed in detail. Their usefulness is also shown by several examples as well as an algorithmic application to the class of augmented Lagrangian methods.

1. Introduction. We consider the *Banach space optimization problem*

$$(P) \qquad \underset{x \in C}{\text{minimize}} \quad f(x) \quad \text{subject to} \quad G(x) \in K,$$

where X and Y are (real) Banach spaces, $f\colon X \to \mathbb{R}$ and $G\colon X \to Y$ are continuously Fréchet differentiable mappings, and $C \subset X$ and $K \subset Y$ are nonempty, closed, convex sets. The feasible set of (P) will be denoted by $\mathcal{F}$, i.e., we use

$$\mathcal{F} := \{x \in C \,|\, G(x) \in K\}.$$

In many cases, K is actually a cone. Moreover, the abstract constraints represented by the set C may not be present, i.e., $C = X$ is possible. Problems akin to (P) have long been identified as a suitable framework for generic optimization covering models from standard nonlinear programming, conic programming, inverse optimization, and optimal control; see, e.g., [12] for more details.

A central role for both the theoretical investigation and the numerical solution of optimization problems like (P) is played by *constraint qualifications* (CQs). The validity of such CQs at a local minimizer of (P) implies that the so-called Karush–Kuhn–Tucker (KKT) conditions associated with this program hold at the latter point.

[†]Institute of Mathematics, University of Würzburg, 97074 Würzburg, Germany (eike.boergens@mathematik.uni-wuerzburg.de, kanzow@mathematik.uni-wuerzburg.de).
[‡]Institute of Mathematics, Brandenburgische Technische Universität Cottbus–Senftenberg, 03046 Cottbus, Germany (mehlitz@b-tu.de, gerd.wachsmuth@b-tu.de).

Unfortunately, there is a major gap between finite- and infinite-dimensional optimization problems regarding available CQs.

Some of the weaker CQs, such as the Abadie and the Guignard CQs (see [1] and [18], respectively) can easily be defined for optimization problems in Banach spaces as well, but, as in the finite-dimensional setting, these conditions are quite abstract and, thus, difficult to check in practice. Moreover, they are usually too weak to yield meaningful consequences for convergence theory associated with optimization algorithms which can be used to tackle (P). There exist very few stronger CQs, but they are often not satisfied in practical applications. The most prominent example is probably *Robinson's constraint qualification* (RCQ) [23, 33, 41], which is already violated in the very simple situation where two-sided (pointwise) box constraints in Lebesgue spaces are considered. On the other hand, there exist numerous CQs in the finite-dimensional context, but these typically depend on the notion of active constraints and, therefore, cannot be translated directly to the infinite-dimensional setting. Here, RCQ is an exception since it boils down to the well-known Mangasarian–Fromovitz CQ in finite dimensions.

Our aim is therefore to introduce new CQs for optimization problems in Banach spaces which are weaker than RCQ and, consequently, have a chance to be satisfied for a significantly larger class of problems. These new CQs are so-called sequential CQs and, thus, closely related to the notion of the asymptotic KKT conditions (AKKT). Our study is motivated by a recent series of papers (dealing with finite-dimensional standard nonlinear programs) on the so-called AKKT regularity or cone continuity property; see, e.g., [4, 5, 7, 8]. This paper is based on the seemingly simple, but important, observation that this AKKT regularity can be formulated without using the notion of active constraints, and therefore it allows an extension to optimization problems in Banach spaces. Nevertheless, the generalization of the AKKT regularity to Banach spaces requires some care due to the difference between weak and strong convergence. On the other hand, the opportunity of distinguishing between strong and weak convergence in primal and dual spaces gives us some freedom to define not only one but several types of AKKT regularity. Specifically, we will define three kinds of AKKT regularity which turn out to be satisfied in different situations and have significantly different applications.

The organization of the paper is as follows: Section 2 recalls some basic definitions and preliminary results. Section 3 introduces the notion of AKKT conditions and essentially shows that, in a reflexive Banach space, every local minimizer of (P) satisfies these AKKT conditions under mild additional assumptions on the problem's initial data. This result does not require any constraint qualification. We then show in section 4 that these AKKT conditions reduce to the usual KKT conditions if and only if suitable sequential CQs hold; we call this *AKKT regularity* in our context. The relation between the introduced types of AKKT regularity as our new CQs and existing CQs in Banach spaces is discussed in section 5. Section 6 is devoted to the investigation of particular constraint systems where the introduced types of AKKT regularity are inherent or can be checked with the aid of reasonable conditions. Particularly, we consider linear and nonlinear equality constraints as well as two-sided box constraints in Lebesgue spaces. An application of our results to the convergence of augmented Lagrangian methods is presented in section 7. We close the paper with some final remarks in section 8.

2. Preliminaries. We mainly use standard notation in this paper; see [12]. The tools from variational analysis which we exploit here are taken from [9, 14, 27].

Throughout the paper, we denote strong, weak, and weak* convergence of sequences by $\to$, $\rightharpoonup$, and $\rightharpoonup^*$, respectively. Let X be a Banach space equipped with norm $\|\cdot\|_X$. The associated dual pairing will be represented by $\langle \cdot, \cdot \rangle_X \colon X^* \times X \to \mathbb{R}$. For $x \in X$ and $r > 0$, $B_r(x) \subset X$ is used for the closed ball with center x and radius r. If $S \subset X$ is a nonempty subset of X, we denote by $d_S = \mathrm{dist}(\cdot, S) \colon X \to \mathbb{R}$ the distance function associated with S w.r.t. the underlying norm. In the case where X is a Hilbert space, we use $(\cdot, \cdot)_X \colon X \times X \to \mathbb{R}$ in order to represent the associated inner product. If S is a nonempty, closed, convex subset of the Hilbert space X, we write $P_S \colon X \to X$ for the projection map onto S.

Given Banach spaces X and Y, a mapping $T \colon X \to Y$ is called *weak-to-weak* sequentially continuous* if it maps weakly convergent sequences to weak* convergent sequences, and is called *completely continuous* if it maps weakly convergent sequences to strongly convergent sequences. It is well known that, given a Fréchet differentiable, completely continuous operator T, the Fréchet derivative $T'(x) \in \mathbb{L}(X, Y)$ is a completely continuous (or compact) linear operator for all $x \in X$; see [15, Thm. 1.5.1]. It is also possible (but slightly more involved) to give sufficient conditions for the complete continuity of the derivative mapping $T' \colon X \to \mathbb{L}(X, Y)$; see [29]. Above, $\mathbb{L}(X, Y)$ denotes the Banach space of all bounded, linear operators mapping from X to Y. For the sake of brevity, the norm in $\mathbb{L}(X, Y)$ will be denoted by $\|\cdot\|$ since the underlying spaces X and Y will be clear from the context.

For a Banach space X and sets $A \subset X$ and $B \subset X^*$, we define

$$A^\circ := \{v \in X^* \,|\, \forall x \in A \colon \langle v, x \rangle_X \leq 0\}, \qquad B^\circ := \{x \in X \,|\, \forall v \in B \colon \langle v, x \rangle_X \leq 0\},$$

which will be referred to as the *polar cones* of A and B, respectively. Clearly, A° and B° are both closed, convex cones.

For a closed, convex set $S \subset X$, the *recession cone* of S is denoted by

$$S_\infty := \{d \in X \,|\, \{d\} + S \subset S\}.$$

It can be easily seen that S_∞ is a closed, convex cone. Next, fix a reference point $\bar{x} \in S$. We denote by

$$\mathcal{R}_S(\bar{x}) := \{\alpha(x - \bar{x}) \,|\, \alpha \geq 0, \ x \in S\}, \qquad \mathcal{N}_S(\bar{x}) := \mathcal{R}_S(\bar{x})^\circ$$

the *radial cone* (also called the *cone of feasible directions*) and the *normal cone* (in the sense of convex analysis) to S at $\bar{x}$, respectively. For points $\tilde{x} \notin S$, we set $\mathcal{N}_S(\tilde{x}) := \varnothing$.

The following lemma relates the normal cone and the recession cone of a closed, convex set.

LEMMA 2.1. *Let X be a Banach space, and let $S \subset X$ be a nonempty, closed, convex set. Then $\{v \in X^* \,|\, \sup_{x \in S} \langle v, x \rangle_X < \infty\} \subset (S_\infty)^\circ$ holds. Particularly, $\mathcal{N}_S(\bar{x}) \subset (S_\infty)^\circ$ is valid for all $\bar{x} \in S$.*

Proof. Let $v \in X^*$ be a point with $\langle v, x \rangle_X \leq c$ for some $c \in \mathbb{R}$ and all $x \in S$. Fix $d \in S_\infty$ and choose an arbitrary element $\bar{x} \in S$. Then $\bar{x} + td \in S$ is valid for all $t > 0$ since S_∞ is a cone. Thus $\langle v, \bar{x} + td \rangle_X \leq c$ has to hold for all $t > 0$. Clearly, this implies $\langle v, d \rangle_X \leq 0$. Since $d \in S_\infty$ was arbitrarily chosen, $v \in (S_\infty)^\circ$ follows. $\square$

Given a possibly nonconvex, closed set $Q \subset X$ and an element $\bar{x} \in Q$, we call

$$\mathcal{T}_Q(\bar{x}) := \{d \in X \,|\, \exists \{x^k\} \subset Q \, \exists \{t^k\} \subset \mathbb{R} \colon x^k \to \bar{x}, \ t^k \searrow 0, \ (x^k - \bar{x})/t^k \to d\},$$
$$\mathcal{T}_Q^w(\bar{x}) := \{d \in X \,|\, \exists \{x^k\} \subset Q \, \exists \{t^k\} \subset \mathbb{R} \colon x^k \to \bar{x}, \ t^k \searrow 0, \ (x^k - \bar{x})/t^k \rightharpoonup d\},$$

$$\mathcal{T}_Q^{\mathrm{C}}(\bar{x}) := \left\{ d \in X \ \middle| \ \begin{array}{l} \forall \{x^k\} \subset Q \, \forall \{t^k\} \subset \mathbb{R} \text{ such that } x^k \to \bar{x}, \, t^k \searrow 0 \\ \exists \{d^k\} \subset X \colon d^k \to d, \, x^k + t^k d^k \in Q \, \forall k \in \mathbb{N} \end{array} \right\}$$

the *(Bouligand) tangent cone* or *contingent cone*, the *weak tangent cone*, and the *Clarke tangent cone* to Q at $\bar{x}$, respectively. By definition of these cones, the inclusions $\mathcal{T}_Q^{\mathrm{C}}(\bar{x}) \subset \mathcal{T}_Q(\bar{x}) \subset \mathcal{T}_Q^w(\bar{x})$ are always satisfied. In contrast to the tangent and the Clarke tangent cones, which are always closed, the weak tangent cone does not necessarily possess this property. Furthermore, we note that the Clarke tangent cone is always convex. If Q is convex, then all these tangent cones coincide with $\mathrm{cl}\, \mathcal{R}_Q(\bar{x})$.

Next, we assume that X is a reflexive Banach space. For the nonempty, closed set $Q \subset X$ and a reference point $\bar{x} \in Q$, we define by

$$\widehat{\mathcal{N}}_Q(\bar{x}) := \{v \in X^* \,|\, \forall x \in Q \colon \langle v, x - \bar{x} \rangle_X \le o(\|x - \bar{x}\|_X)\},$$

$$\mathcal{N}_Q^{\mathrm{L}}(\bar{x}) := \left\{ v \in X^* \ \middle| \ \exists \{x^k\} \subset Q \, \exists \{v^k\} \subset X^* \colon x^k \to \bar{x}, \, v^k \rightharpoonup v, \, v^k \in \widehat{\mathcal{N}}_Q(x^k) \, \forall k \in \mathbb{N} \right\}$$

the *Fréchet normal cone* (or *regular normal cone*) and the *limiting normal cone* (or *Mordukhovich normal cone*) to Q at $\bar{x}$, respectively. We always have $\widehat{\mathcal{N}}_Q(\bar{x}) = \mathcal{T}_Q^w(\bar{x})^{\circ}$, which is why the Fréchet normal cone is always closed and convex. On the contrary, the limiting normal cone does not possess any of these properties in general. Note that the above representation of the limiting normal cone only holds in the setting of reflexive Banach spaces. A more general definition can be found in [27]. If Q is convex, it holds that $\mathcal{N}_Q(\bar{x}) = \widehat{\mathcal{N}}_Q(\bar{x}) = \mathcal{N}_Q^{\mathrm{L}}(\bar{x}) = \{v \in X^* \,|\, \langle v, x - \bar{x} \rangle \le 0 \text{ for all } x \in Q\}$, i.e., all these cones coincide with the classical normal cone from convex analysis.

Let $\bar{x} \in \mathcal{F}$ be a feasible point of the optimization problem (P). Then

$$(2.1) \qquad \mathcal{L}_{\mathcal{F}}(\bar{x}) := \{d \in \mathcal{T}_C(\bar{x}) \,|\, G'(\bar{x})d \in \mathcal{T}_K(G(\bar{x}))\}$$

is called the *linearization cone* to $\mathcal{F}$ at $\bar{x}$. Note that the definition of this cone heavily depends on the precise (nonlinear) description of the set $\mathcal{F}$ via C, K, and G. One can easily check that the inclusion $\mathcal{T}_{\mathcal{F}}^w(\bar{x}) \subset \mathcal{L}_{\mathcal{F}}(\bar{x})$ is generally valid; see, e.g., [17, proof of Lem. 4.2]. In order to guarantee validity of the converse inclusion, a CQ is necessary in general; see section 5.

We now turn to the optimality conditions of the optimization problem (P). To this end, we define the *Lagrange function* or *Lagrangian* $L \colon X \times Y^* \to \mathbb{R}$ of the problem as

$$\forall x \in X \ \forall \lambda \in Y^* \colon \quad L(x, \lambda) := f(x) + \langle \lambda, G(x) \rangle_Y.$$

This function occurs quite prominently in the KKT conditions of (P).

DEFINITION 2.2. *A feasible point $\bar{x} \in \mathcal{F}$ of (P) is called a* KKT *point if there exists $\bar{\lambda} \in Y^*$ such that*

$$-L_x'(\bar{x}, \bar{\lambda}) \in \mathcal{N}_C(\bar{x}) \quad and \quad \bar{\lambda} \in \mathcal{N}_K(G(\bar{x})).$$

In this case, $\bar{\lambda}$ is called a (Lagrange) multiplier *of (P) associated with $\bar{x}$.*

3. The asymptotic KKT conditions. The following is the central definition of this section. It generalizes the known definitions of asymptotic or approximate KKT conditions from the finite-dimensional setting (see [4, 5, 7, 8, 10]) to our optimization problem in Banach spaces (P). Note that there exist different possibilities for such a generalization, but we found the following one particularly useful (this definition is essentially taken from the Ph.D. thesis [34]).

DEFINITION 3.1. *A sequence* $\{(x^k, \lambda^k)\} \subset C \times Y^*$ *is called a* strong asymptotic KKT sequence (s-AKKT sequence) *if there are sequences* $\{\varepsilon^k\} \subset X^*$ *and* $\{r^k\} \subset [0, +\infty)$ *such that*

$$(3.1) \quad \forall k \in \mathbb{N}: \quad \varepsilon^k - L'_x(x^k, \lambda^k) \in \mathcal{N}_C(x^k) \quad and \quad \langle \lambda^k, y - G(x^k) \rangle_Y \leq r^k \quad \forall y \in K,$$

with $\varepsilon^k \to 0$ *and* $r^k \searrow 0$. *We call* $\{(x^k, \lambda^k)\} \subset C \times Y^*$ *a* weak* asymptotic KKT sequence (w-AKKT sequence) *if* (3.1) *holds with* $\varepsilon^k \rightharpoonup^* 0$ *and* $r^k \searrow 0$.

Note that the first condition simplifies to $\varepsilon^k - L'_x(x^k, \lambda^k) = 0$ if $C = X$, whereas the second condition implies $\lambda^k \in K^\circ$ for all $k \in \mathbb{N}$ if K is a cone. To see the latter statement, fix $k \in \mathbb{N}$. Then, exploiting the fact that K is a cone, the condition $\langle \lambda^k, y - G(x^k) \rangle_Y \leq r^k$ for all $y \in K$ can be written as $\langle \lambda^k, \alpha y - G(x^k) \rangle_Y \leq r^k$ for all $y \in K$ and all $\alpha > 0$. Dividing this expression by $\alpha > 0$ yields $\langle \lambda^k, y - \frac{G(x^k)}{\alpha} \rangle_Y \leq \frac{r^k}{\alpha}$ for all $y \in K$ and all $\alpha > 0$. Taking the limit $\alpha \to \infty$ therefore implies $\langle \lambda^k, y \rangle_Y \leq 0$ for all $y \in K$. Hence, $\lambda^k \in K^\circ$ holds; see Remark 4.1 as well.

The main idea of (strong or weak) AKKT sequences is, obviously, the existence of a sequence which satisfies the KKT conditions inexactly using a certain measure of inexactness. As already observed, e.g., in [4], there exist different ways to measure the degree of inexactness, and these measures are not necessarily equivalent, even in finite dimensions. As an example, consider the usual complementarity condition $\lambda \geq 0, g(x) \leq 0, \lambda g(x) = 0$ associated with an inequality constraint $g(x) \leq 0$ which is induced by a function $g \colon \mathbb{R}^n \to \mathbb{R}$. This complementarity condition can be rewritten as $\min(-g(x), \lambda) = 0$; hence the condition $|\min(-g(x), \lambda)| \leq r$ is a very natural criterion for an inexact satisfaction of the complementarity condition. Alternatively, one might use a condition such as $\lambda \geq 0, g(x) \leq 0, -\lambda g(x) \leq r$. These two conditions, however, are not equivalent, e.g., for $n = 1$, take $g(x) := -x$ and consider the sequences defined by $r^k := 1/k, \lambda^k := k^2$, and $x^k := 1/k$ for each $k \in \mathbb{N}$. Then the first condition holds for all $k \in \mathbb{N}$, whereas the second is violated, in fact, $-\lambda^k g(x^k) \to \infty$. This should be kept in mind because the concise definition of a (weak or strong) AKKT sequence plays a crucial role. In this paper, we take advantage of Definition 3.1, but alternative definitions might also be useful in other contexts. The above example also depicts that our definition of a (weak or strong) AKKT sequence is slightly different from the one stated in [7, Definition 1.2] for standard nonlinear problems in finite dimensions.

Note that Definition 3.1 does not require any convergence or weak convergence of the sequence $\{(x^k, \lambda^k)\}$. By forcing weak or strong convergence of the primal sequence $\{x^k\}$, we obtain the following definition. Notice that this definition still does not assume any convergence or boundedness of the dual sequence of multipliers $\{\lambda^k\}$.

DEFINITION 3.2. *Let* $\bar{x} \in \mathcal{F}$ *be a feasible point of* (P). *Then we call* $\bar{x}$ *a*

(a) weak asymptotic KKT point (w-AKKT point) *if there exists a w-AKKT sequence* $\{(x^k, \lambda^k)\}$ *such that* $x^k \rightharpoonup \bar{x}$.
(b) strong asymptotic KKT point (s-AKKT point) *if there exists an s-AKKT sequence* $\{(x^k, \lambda^k)\}$ *such that* $x^k \to \bar{x}$.

In principle, there exist four possible definitions of AKKT points due to the four possible combinations of strong and weak convergence of the underlying sequences $\{x^k\}$ and $\{\varepsilon^k\}$. Taking this into account, a more precise terminology for w- and s-AKKT points would be ww-AKKT and ss-AKKT points. In addition, one could therefore also define sw-AKKT and ws-AKKT points, and we do not exclude the possibility that these additional notions might be useful in certain situations. For our

purposes, however, the above two definitions are sufficient, and to avoid an overkill in the terminology, we simply talk about w-AKKT and s-AKKT points. In particular, the notion of w-AKKT points is motivated by the fact that suitable methods generate sequences which admit weak accumulation points, not necessarily strong ones.

The following statement, inspired by the corresponding finite-dimensional result in [10, Thm. 3.1] and the Hilbert space result in [34, Prop. 3.50], essentially shows that every local minimizer of (P) is an s-AKKT point, and therefore, in particular, a w-AKKT point. Recall that a local minimizer is not necessarily a KKT point, and hence the concept of (strong or weak) AKKT points is more general than the notion of KKT points.

PROPOSITION 3.3. *Let $\bar{x}$ be a local minimizer of (P). Assume that X is reflexive, and suppose that f is weakly sequentially lower semicontinuous in a neighborhood of $\bar{x}$. Moreover, we assume that*

$$(3.2) \quad \forall \{x^k\} \subset C \, \forall x \in C \colon \qquad x^k \rightharpoonup x \quad and \quad d_K(G(x^k)) \to 0 \quad \implies \quad G(x) \in K$$

holds. Then $\bar{x}$ is an s-AKKT point and, thus, also a w-AKKT point of (P).

Proof. Let $\epsilon > 0$ be such that $\bar{x}$ minimizes f on $B_\epsilon(\bar{x}) \cap \mathcal{F}$. Noting that f is continuous, we find some $r \in (0, \epsilon)$ such that f is bounded from below and weakly sequentially lower semicontinuous on $B_r(\bar{x}) \cap C$. For $k \in \mathbb{N}$, consider the problem

$$(3.3) \quad \minimize_{x \in X} \quad f(x) + \|x - \bar{x}\|_X^2 + k d_K^2(G(x)) \quad \text{subject to} \quad x \in B_r(\bar{x}) \cap C.$$

An application of Ekeland's variational principle [9, Thm. 3.3.1] in the complete metric space X yields a point $x^k \in B_r(\bar{x}) \cap C$ which minimizes

$$(3.4) \quad \begin{aligned} \minimize_{x \in X} \quad & f(x) + \|x - \bar{x}\|_X^2 + k d_K^2(G(x)) + \frac{1}{k} \|x - x^k\|_X \\ \text{subject to} \quad & x \in B_r(\bar{x}) \cap C. \end{aligned}$$

By using $x = \bar{x}$, we obtain

$$(3.5) \quad f(x^k) + \|x^k - \bar{x}\|_X^2 + k d_K^2(G(x^k)) \leq f(\bar{x}) + \frac{1}{k} \|\bar{x} - x^k\|_X.$$

Since $\{x^k\}$ is bounded, we have w.l.o.g. $x^k \rightharpoonup \hat{x}$ for some $\hat{x} \in B_r(\bar{x}) \cap C$. From (3.5) we obtain

$$\begin{aligned} f(\bar{x}) &\geq \limsup_{k \to \infty} \big(f(x^k) + \|x^k - \bar{x}\|_X^2 + k \, d_K^2(G(x^k)) \big) \\ &\geq \liminf_{k \to \infty} f(x^k) + \limsup_{k \to \infty} \|x^k - \bar{x}\|_X^2 + \liminf_{k \to \infty} k \, d_K^2(G(x^k)) \\ &\geq f(\hat{x}) + \limsup_{k \to \infty} \|x^k - \bar{x}\|_X^2 + \liminf_{k \to \infty} k \, d_K^2(G(x^k)). \end{aligned}$$

This shows that $d_K(G(x^k)) \to 0$ holds (along a subsequence). Hence, (3.2) implies validity of $G(\hat{x}) \in K$. Thus, $\hat{x} \in \mathcal{F}$. Now, the above inequality implies

$$f(\bar{x}) \geq f(\hat{x}) + \limsup_{k \to \infty} \|x^k - \bar{x}\|_X^2 \geq f(\bar{x}) + \limsup_{k \to \infty} \|x^k - \bar{x}\|_X^2.$$

Thus, $x^k \to \bar{x}$ and, therefore, $\hat{x} = \bar{x}$.

Recalling that x^k is a solution of (3.4) for each $k \in \mathbb{N}$ while $x^k \to \bar{x}$ holds, $\|x^k - \bar{x}\|_X < r$ is valid for sufficiently large $k \in \mathbb{N}$. Thus, we obtain the existence of sequences $\{\varepsilon^k\} \subset X^*$ and $\{\xi^k\} \subset Y^*$ such that $-\varepsilon^k \in \partial\|\cdot\|_X^2(x^k - \bar{x}) + \frac{1}{k}\partial\|\cdot\|_X(0)$ and $\xi^k \in \partial d_K(G(x^k))$ as well as

$$(3.6) \qquad 0 \in \left\{ f'(x^k) - \varepsilon^k + 2k\, d_K(G(x^k))\, G'(x^k)^*\xi^k \right\} + \mathcal{N}_C(x^k)$$

hold. Here, we exploited the Fermat, sum, and chain rules for Clarke's generalized derivative, see [14, Props. 2.3.2 and 2.3.3; Thms. 2.3.9 and 2.3.10]. Similarly, Clarke's chain rule implies $\partial\|\cdot\|_X^2(x^k - \bar{x}) = 2\|x^k - \bar{x}\|_X\, \partial\|\cdot\|_X(x^k - \bar{x})$. Moreover, it is well known that the inclusion $\partial\|\cdot\|_X(x) \subset B_1(0)$ holds in X^* for all $x \in X$. Hence, we obtain $\varepsilon^k \to 0$ in X^*. By definition of the convex function d_K and its subdifferential at $G(x^k)$, we have

$$0 = d_K(y) \geq d_K(G(x^k)) + \langle \xi^k, y - G(x^k) \rangle_Y \geq \langle \xi^k, y - G(x^k) \rangle_Y$$

for all $y \in K$. Thus, setting $\lambda^k := 2k\, d_K(G(x^k))\, \xi^k$, we have $\langle \lambda^k, y - G(x^k) \rangle_Y \leq 0$ for all $y \in K$. Combining this with (3.6), we easily see that $\bar{x}$ is an s-AKKT point. $\qquad\square$

Let us briefly note that the continuity property (3.2) obviously holds whenever the composition $d_K \circ G$ is weakly sequentially lower semicontinuous. In particular, this property is inherent whenever G is an affine mapping induced by a bounded linear operator since then the composition $d_K \circ G$ is a continuous, convex mapping.

Remark 3.4. If X is a Hilbert space (or if $\partial\|\cdot\|_X^2$ is strongly monotone) and f' is locally Lipschitz continuous, then the weak sequential lower semicontinuity of f can be omitted by using the problem

$$\underset{x \in X}{\text{minimize}} \quad f(x) + \gamma\|x - \bar{x}\|_X^2 + kd_K^2(G(x)) \quad \text{subject to} \quad x \in B_r(\bar{x}) \cap C$$

instead of (3.3). Above, γ is sufficiently large such that $X \ni x \mapsto f(x) + \gamma\|x - \bar{x}\|_X^2 \in \mathbb{R}$ is locally convex and, thus, weakly sequentially lower semicontinuous (this statement can be verified, e.g., by exploiting the equivalence of monotonicity of the gradient and convexity of the underlying function).

We close this section with two illustrative examples. The first considers an optimization problem which possesses a minimizer where the KKT conditions are violated, but this minimizer is an s-AKKT point.

Example 3.5. Consider the optimization problem (P) with

$$X := \mathbb{R} \times L^2(0,1), \qquad Y := L^2(0,1), \qquad\qquad K := \{0\} \subset Y,$$
$$C := \mathbb{R} \times \{u \in L^2(0,1) \mid -1 \leq u \leq 1\}, \qquad G(\alpha, u) := \alpha \cdot q - u,$$

where $q \in L^2(0,1) \setminus L^\infty(0,1)$ is a fixed function.

We argue that $\mathcal{F} = \{0\}$ holds. Indeed, the constraint $G(\alpha, u) \in K$ implies $u = \alpha\, q$. If $\alpha \neq 0$, u is unbounded and, therefore, $(\alpha, u) \notin C$. Thus, $(\alpha^*, u^*) = (0,0)$ is the global minimizer of (P) for any objective function f. In particular, $f'(\alpha^*, u^*)$ can be an arbitrary vector in $X^* = \mathbb{R} \times L^2(0,1)$.

It is easy to check the validity of $\mathcal{T}_C(0,0) = \mathbb{R} \times L^2(0,1)$, $\mathcal{T}_K(0) = \mathcal{R}_K(0) = \{0\}$, $\mathcal{N}_C(0,0) = \{(0,0)\}$, and $\mathcal{N}_K(0) = Y^*$. Hence, $(0,0)$ is a KKT point if and only if

$$f'(0,0) = -G'(0,0)^*\lambda = \begin{pmatrix} -\langle q, \lambda \rangle_{L^2(0,1)} \\ \lambda \end{pmatrix}$$

holds for some $\lambda \in L^2(0,1)$. Thus, there exist linear functionals f such that $(0,0)$ is not a KKT point, e.g., $f := (-1,0) \in X^*$.

Let us show that $(0,0)$ is an s-AKKT point for the linear functional $f = (-1,0)$. Of course, this follows from Proposition 3.3, but it is instructive to construct the corresponding s-AKKT sequence explicitly. To this end, we fix the function q via $q(t) = t^{-1/4}$. Now, we set $\alpha^k := 1/k$ and $u^k(t) := P_{[-1,1]}(\alpha^k q(t))$, i.e.,

$$u^k(t) = \begin{cases} 1 & \text{for } t \leq k^{-4}, \\ \frac{t^{-1/4}}{k} & \text{for } t > k^{-4}. \end{cases}$$

Next, we choose $\lambda^k \in L^2(0,1)$ which is supported on $[0, k^{-4}]$ with $\langle q, \lambda^k \rangle_{L^2(0,1)} = 1$, e.g., $\lambda^k = \frac{3}{4} k^3 \chi_{[0,k^{-4}]}$. Thus, $\mu^k := (0, \lambda^k)$ is an element of $\mathcal{N}_C(\alpha^k, u^k)$. Moreover, we have

$$f'(\alpha^k, u^k) + G'(\alpha^k, u^k)^* \lambda^k + \mu^k = \begin{pmatrix} -1 \\ 0 \end{pmatrix} + \begin{pmatrix} \langle q, \lambda^k \rangle_{L^2(0,1)} \\ -\lambda^k \end{pmatrix} + \begin{pmatrix} 0 \\ \lambda^k \end{pmatrix} = 0$$

and

$$\langle \lambda^k, 0 - G(\alpha^k, u^k) \rangle_{L^2(0,1)} = -(4k)^{-1} \leq 0.$$

Thus, $\{((\alpha^k, u^k), \lambda^k)\}$ is an s-AKKT sequence. Due to $\|\lambda^k\|_{L^2(0,1)} = \frac{3}{4} k$, $\{\lambda^k\}$ is unbounded in $L^2(0,1)$.

Note that, for this choice of f, the point $(0,0)$ is not a KKT point and not even a Fritz–John point of the associated problem (P).

It might be possible to construct a similar example involving ℓ^2 by using an idea of [30, Ex. 2.5]. The second example indicates that Proposition 3.3 may not hold without the reflexivity of the underlying space.

Example 3.6. Consider again the optimization problem (P) with the data

$$X := \ell^1, \qquad\qquad Y := \ell^2, \qquad\qquad C := \ell^1,$$

$$K := \{0\} \subset \ell^2, \qquad G(x) := x, \qquad f(x) := \sum_{i=1}^{\infty} a_i x_i$$

for some given sequence $a \in \ell^\infty \setminus c_0$. Clearly, $\bar{x} := 0$ is the only feasible point and therefore is also optimal.

We argue by contradiction. Let us assume that $\{(x^k, \lambda^k)\}$ is an s-AKKT sequence. Then the convergence

$$f'(x^k) + G'(x^k)^* \lambda^k = a + \lambda^k \to 0$$

has to hold in $X^* = \ell^\infty$. However, $\lambda^k \in \ell^2 \subset c_0$ holds for each $k \in \mathbb{N}$, and c_0 is a closed subspace of ℓ^∞. On the other hand, $a \notin c_0$ holds by assumption. This contradiction shows that the reflexivity assumption in Proposition 3.3 is essential.

4. Asymptotic KKT regularity. The previous section was devoted to the notion of (strong or weak) AKKT points. In particular, under fairly mild assumptions, we proved that any local minimizer of (P) is a (strong) AKKT point. Hence, $\bar{x}$ being an AKKT point is a necessary optimality condition which, in particular, does not require any constraint qualification. Therefore, the natural question arises under

which assumption such an AKKT point is already a KKT point. By generalization of the corresponding finite-dimensional theory from [4, 7] to our setting, this leads to three types of AKKT regularity which turn out to be CQs. More precisely, we will see that, in some sense, they are the weakest possible CQs which guarantee that a given (strong or weak) AKKT point of (P) is already a KKT point. The relation of these types of AKKT regularity to some existing CQs will be discussed in section 5.

Motivated by the definition of (strong or weak) AKKT points, let us introduce

$$(4.1) \qquad \mathcal{M}(x,r) := \left\{ G'(x)^*\lambda + \mu \in X^* \left| \begin{array}{l} \lambda \in Y^*, \, \mu \in \mathcal{N}_C(x), \\ \langle \lambda, y - G(x)\rangle_Y \leq r \, \forall y \in K \end{array} \right. \right\}$$

for $x \in X$ and $r \geq 0$. Note that we have $\mathcal{M}(\tilde{x}, r) = \varnothing$ for all $\tilde{x} \notin C$ and $r \geq 0$ by definition of the normal cone. If $\bar{x} \in X$ is feasible to (P), it holds that

$$(4.2) \qquad \mathcal{M}(\bar{x}, 0) = \{G'(\bar{x})^*\lambda + \mu \in X^* \mid \lambda \in \mathcal{N}_K(G(\bar{x})), \, \mu \in \mathcal{N}_C(\bar{x})\}.$$

It follows that the condition $-f'(\bar{x}) \in \mathcal{M}(\bar{x}, 0)$ is equivalent to $\bar{x}$ being a KKT point of (P); see Definition 2.2.

Let us first state two simple observations regarding the set $\mathcal{M}(x, r)$.

Remark 4.1. If K is a cone, we have the equivalence

$$\forall y \in K \colon \langle \lambda, y - G(x)\rangle_Y \leq r \quad \iff \quad \lambda \in K^\circ, \, -\langle \lambda, G(x)\rangle_Y \leq r$$

for each $\lambda \in Y^*$, cf. the corresponding discussion after Definition 3.1. This yields the representation

$$\mathcal{M}(x, r) = \{G'(x)^*\lambda + \mu \in X^* \mid \lambda \in K^\circ, \, \mu \in \mathcal{N}_C(x), \, -\langle \lambda, G(x)\rangle_Y \leq r\}$$

for all $x \in X$ and $r \geq 0$.

Remark 4.2. For a general convex set K, the condition

$$\forall y \in K \colon \quad \langle \lambda, y - G(x)\rangle_Y \leq r$$

from the definition of $\mathcal{M}(x, r)$ implies that $\sup_{y \in K} \langle \lambda, y\rangle_Y < +\infty$, and hence $\lambda \in (K_\infty)^\circ$, by Lemma 2.1. This can be interpreted as a sign property of the Lagrange multiplier λ, as it was introduced in [4] in the finite-dimensional setting.

Before we introduce the three already advertised versions of AKKT regularity, let us define the following Painlevé–Kuratowski-type outer/upper limits,

$$\limsup_{\substack{x \to \bar{x} \\ r \searrow 0}} \mathcal{M}(x, r) := \left\{ \bar{v} \in X^* \left| \begin{array}{l} \exists \{x^k\} \subset X \, \exists \{r^k\} \subset [0, +\infty) \, \exists \{v^k\} \subset X^* : \\ x^k \to \bar{x}, \, r^k \searrow 0, \, v^k \to \bar{v}, \\ v^k \in \mathcal{M}(x^k, r^k) \, \forall k \in \mathbb{N} \end{array} \right. \right\},$$

$$\text{w*-}\limsup_{\substack{x \to \bar{x} \\ r \searrow 0}} \mathcal{M}(x, r) := \left\{ \bar{v} \in X^* \left| \begin{array}{l} \exists \{x^k\} \subset X \, \exists \{r^k\} \subset [0, +\infty) \, \exists \{v^k\} \subset X^* : \\ x^k \to \bar{x}, \, r^k \searrow 0, \, v^k \rightharpoonup^* \bar{v}, \\ v^k \in \mathcal{M}(x^k, r^k) \, \forall k \in \mathbb{N} \end{array} \right. \right\},$$

$$\text{w*-}\limsup_{\substack{x \rightharpoonup \bar{x} \\ r \searrow 0}} \mathcal{M}(x, r) := \left\{ \bar{v} \in X^* \left| \begin{array}{l} \exists \{x^k\} \subset X \, \exists \{r^k\} \subset [0, +\infty) \, \exists \{v^k\} \subset X^* : \\ x^k \rightharpoonup \bar{x}, \, r^k \searrow 0, \, v^k \rightharpoonup^* \bar{v}, \\ v^k \in \mathcal{M}(x^k, r^k) \, \forall k \in \mathbb{N} \end{array} \right. \right\}$$

of the set-valued map $\mathcal{M}\colon X \times [0, +\infty) \rightrightarrows X^*$ w.r.t. some point $\bar{x} \in \mathcal{F}$. Recall that $\mathcal{M}(x^k, r^k)$ is empty for all $x^k \notin C$; hence, requiring the existence of a sequence $\{x^k\} \subset X$ in the previous definitions is equivalent to assuming that this sequence belongs to the set C.

These definitions give rise to the following observation.

LEMMA 4.3. *Let $\bar{x} \in \mathcal{F}$ be a feasible point of (P).*

(a) *Then $\bar{x}$ is an s-AKKT point if and only if $-f'(\bar{x}) \in \limsup_{\substack{x \to \bar{x} \\ r \searrow 0}} \mathcal{M}(x, r)$.*

(b) *Suppose that $f'\colon X \to X^*$ is weak-to-weak*-sequentially continuous. Then $\bar{x}$ is a w-AKKT point if and only if $-f'(\bar{x}) \in \text{w*-}\limsup_{\substack{x \to \bar{x} \\ r \searrow 0}} \mathcal{M}(x, r)$.*

The proof follows immediately from the definitions. We already observed that $\bar{x}$ is a KKT point if and only if $-f'(\bar{x}) \in \mathcal{M}(\bar{x}, 0)$. This motivates the next definition.

DEFINITION 4.4. *Let $\mathcal{M}(x, r)$ be defined as in (4.1), and let $\bar{x} \in \mathcal{F}$ be any feasible point of (P). Then $\bar{x}$ is*

(a) strongly AKKT regular *(s-AKKT regular) if*

$$\limsup_{\substack{x \to \bar{x} \\ r \searrow 0}} \mathcal{M}(x, r) \subset \mathcal{M}(\bar{x}, 0);$$

(b) strongly weakly AKKT regular *(sw-AKKT regular) if*

$$\text{w*-}\limsup_{\substack{x \to \bar{x} \\ r \searrow 0}} \mathcal{M}(x, r) \subset \mathcal{M}(\bar{x}, 0);$$

(c) weakly AKKT regular *(w-AKKT regular) if*

$$\text{w*-}\limsup_{\substack{x \to \bar{x} \\ r \searrow 0}} \mathcal{M}(x, r) \subset \mathcal{M}(\bar{x}, 0).$$

Note that these three conditions are equivalent in finite dimensions. In the infinite-dimensional setting, however, we only have

$$\text{w-AKKT regular} \implies \text{sw-AKKT regular} \implies \text{s-AKKT regular}.$$

While the role of the s- and w-AKKT regularities is motivated by the definition of s-AKKT and w-AKKT points, respectively, we will see in section 5 that sw-AKKT regularity possesses reasonable relationships to classical CQs in infinite-dimensional programming. Let us also emphasize that Definition 4.4 is stated for feasible points only. Hence, whenever we use one of the above AKKT regularity conditions in our subsequent analysis, there is the implicit assumption that $\bar{x}$ is feasible to (P) even if this might not be stated explicitly.

In the above definitions of the three types of AKKT regularity, we required that $r \geq 0$. This, however, was just done to make the definition more concise. If one desires to allow $r < 0$, all results presented in what follows are applicable using $\tilde{r} := \max(0, r)$.

Remark 4.5. Let $\bar{x} \in \mathcal{F}$ be a feasible point of (P). Then, if any of the conditions in Definition 4.4 is satisfied, we automatically have equality in the defining property, e.g., s-AKKT regularity implies

$$\limsup_{\substack{x \to \bar{x} \\ r \searrow 0}} \mathcal{M}(x, r) = \mathcal{M}(\bar{x}, 0).$$

Particularly, this yields that $\mathcal{M}(\bar{x}, 0)$ is closed.

Noting that $\mathcal{M}(\bar{x}, 0)$ is convex, we see that $\bar{x}$ being s-AKKT regular implies the weak closedness of $\mathcal{M}(\bar{x}, 0)$. Whenever X is reflexive, this coincides with weak* closedness of $\mathcal{M}(\bar{x}, 0)$.

We next discuss the question of under which condition a strong or weak AKKT point is already a KKT point. We first state a "strong" formulation in the following result.

THEOREM 4.6. *Let $\bar{x} \in \mathcal{F}$ be a feasible point of (P). Then the following statements hold:*

(a) *If $\bar{x}$ is an s-AKKT point being s-AKKT regular, then $\bar{x}$ is a KKT point.*
(b) *Conversely, if for every continuously differentiable function f, the implication "$\bar{x}$ is an s-AKKT point $\implies$ $\bar{x}$ is a KKT point" holds, then $\bar{x}$ is s-AKKT regular.*

Proof. (a) If $\bar{x}$ is an s-AKKT point of (P), Lemma 4.3(a) and the s-AKKT regularity imply

$$-f'(\bar{x}) \in \limsup_{\substack{x \to \bar{x} \\ r \searrow 0}} \mathcal{M}(x, r) \subset \mathcal{M}(\bar{x}, 0).$$

Hence, $\bar{x}$ is a KKT point.

(b) Conversely, assume that the s-AKKT conditions at $\bar{x}$ imply the KKT conditions for every continuously differentiable objective function. Then take an arbitrary element $\bar{v} \in \limsup_{\substack{x \to \bar{x} \\ r \searrow 0}} \mathcal{M}(x, r)$. By definition, there exist sequences $\{x^k\} \subset X$, $\{r^k\} \subset [0, +\infty)$, and $\{v^k\} \subset X^*$ such that $x^k \to \bar{x}$, $r^k \searrow 0$, and $v^k \to \bar{v}$ as well as $v^k \in \mathcal{M}(x^k, r^k)$ for all $k \in \mathbb{N}$. Hence, there exist $\lambda^k \in Y^*$ with $\langle \lambda^k, y - G(x^k) \rangle_Y \le r^k$ for all $y \in K$, and $\mu^k \in \mathcal{N}_C(x^k)$ such that $v^k = G'(x^k)^* \lambda^k + \mu^k$ for all $k \in \mathbb{N}$. We emphasize that $\{x^k\} \subset C$ holds by definition of the normal cone. Let us define the particular objective function $f(x) := -\langle \bar{v}, x \rangle_X$ for all $x \in X$. Then $\varepsilon^k := f'(x^k) + v^k = -\bar{v} + v^k \to 0$, and we have

$$\begin{aligned}
\varepsilon^k - L_x'(x^k, \lambda^k) &= v^k + f'(x^k) - f'(x^k) - G'(x^k)^* \lambda^k \\
&= v^k - G'(x^k)^* \lambda^k \\
&= \mu^k \in \mathcal{N}_C(x^k)
\end{aligned}$$

for all $k \in \mathbb{N}$. Hence, $\bar{x}$ is an s-AKKT point of the associated problem (P). By assumption, it follows that $\bar{v} = -f'(\bar{x}) \in \mathcal{M}(\bar{x}, 0)$. This clearly shows the inclusion $\limsup_{\substack{x \to \bar{x} \\ r \searrow 0}} \mathcal{M}(x, r) \subset \mathcal{M}(\bar{x}, 0)$, i.e., $\bar{x}$ is s-AKKT regular. $\qquad\square$

This theorem is similar to [7, Thm. 3.2], which characterizes a AKKT-regularity-type property in the setting of standard finite-dimensional nonlinear programming. However, note that the phrase "that attains a minimum at x^*" has to be deleted from the statement of [7, Thm. 3.2], since otherwise this result would characterize the Guignard CQ, despite the fact that the minimizer property does not hold for the function constructed in the presented proof. Taking together the adjusted theorem from [7] and Theorem 4.6, our types of AKKT regularity from Definition 4.4 are clearly closely related to the cone continuity property from [7] in the setting of standard nonlinear programming in finite dimensions. However, as we already mentioned in section 3, our definition of an AKKT sequence differs slightly from the one in [7] in

this setting, so it remains an open problem to check whether the concepts of the cone continuity property and AKKT regularity actually coincide. In contrast to [7], our set $\mathcal{M}(x,r)$ is not necessarily a cone, which is why we abstain from referring to the CQs from Definition 4.4 as *cone continuity properties*. Let us point out that the term *AKKT regularity* was later also exploited by the authors of [7] in order to refer to CQs of this type; see, e.g., [4].

The previous theorem is quite similar to the statement that Guignard's CQ is the weakest CQ which ensures that local minimizers are KKT points; see [17].

Using essentially the same technique of proof, we obtain the following "weak" counterpart of Theorem 4.6.

THEOREM 4.7. *Let $\bar{x} \in \mathcal{F}$ be a feasible point of (P). Then the following statements hold:*

(a) *Suppose that $f' : X \to X^*$ is weak-to-weak*-sequentially continuous. If $\bar{x}$ is a w-AKKT point being w-AKKT regular, then $\bar{x}$ is a KKT point.*

(b) *Conversely, if for every continuously differentiable function f, the implication "$\bar{x}$ is a w-AKKT point $\Longrightarrow$ $\bar{x}$ is a KKT point" holds, then $\bar{x}$ is w-AKKT regular.*

The previous results imply that the strong, strong-weak, and weak AKKT regularities are CQs under appropriate assumptions on the initial problem data of (P). In fact, given a local minimum $\bar{x}$ of (P) which is w-, sw-, or s-AKKT regular, it follows from Proposition 3.3 (under the assumptions stated there) that $\bar{x}$ is an s-AKKT point. Noting that w- or sw-AKKT regularity implies s-AKKT regularity, the first statement of Theorem 4.6 can be used to infer that $\bar{x}$ is already a KKT point of (P). We summarize these observations in the following corollary.

COROLLARY 4.8. *Let X be reflexive, and let (3.2) be satisfied. Then w-, sw-, and s-AKKT regularity are CQs for (P) in the following sense: For every objective f which is continuously differentiable and weakly sequentially lower semicontinuous, local optimality of $\bar{x}$ implies that $\bar{x}$ is a KKT point.*

Using the terminology from [7], any condition which guarantees that an AKKT point is already a KKT point is called a *strict constraint qualification*. The previous results therefore show that our AKKT-regularity-type conditions are strict CQs, and that they are the weakest possible.

Finally, we want to present sufficient criteria for the sw- and w-AKKT regularities. To this end, recall that, given any sequences $\{x^k\} \subset X$, $\{r^k\} \subset [0, +\infty)$, and $\{v^k\} \subset X^*$ with $v^k \in \mathcal{M}(x^k, r^k)$ for all $k \in \mathbb{N}$, it follows that there exist corresponding sequences $\{\lambda^k\} \subset Y^*$ and $\{\mu^k\} \subset X^*$ such that $v^k = G'(x^k)^* \lambda^k + \mu^k$ holds for all $k \in \mathbb{N}$. In general, these sequences of multipliers might be unbounded even if $\{x^k\}$, $\{r^k\}$, and $\{v^k\}$ converge. The following result discusses the situation where the sequences $\{\lambda^k\}$ and $\{\mu^k\}$ can be chosen as bounded ones.

LEMMA 4.9. *Let a feasible point $\bar{x} \in \mathcal{F}$ of (P) be given.*

(a) *Assume that for all sequences $\{x^k\} \subset X$, $\{r^k\} \subset [0, +\infty)$, and $\{v^k\} \subset X^*$ which satisfy $x^k \to \bar{x}$, $r^k \searrow 0$, $v^k \rightharpoonup^* \bar{v}$, and $v^k \in \mathcal{M}(x^k, r^k)$ for all $k \in \mathbb{N}$, there exist bounded sequences of multipliers $\{\lambda^k\} \subset Y^*$ and $\{\mu^k\} \subset X^*$ such that $\mu^k \in \mathcal{N}_C(x^k)$ and $\langle \lambda^k, y - G(x^k) \rangle_Y \leq r^k$ for all $y \in K$ and $k \in \mathbb{N}$ as well as $v^k = G'(x^k)^* \lambda^k + \mu^k \rightharpoonup^* \bar{v}$. Then $\bar{x}$ is sw-AKKT regular.*

(b) *Assume that G and G' are completely continuous and $C = X$. Assume further that for all sequences $\{x^k\} \subset X$, $\{r^k\} \subset [0, +\infty)$, and $\{v^k\} \subset X^*$ which*

satisfy $x^k \rightharpoonup \bar{x}$, $r^k \searrow 0$, $v^k \rightharpoonup^* \bar{v}$, and $v^k \in \mathcal{M}(x^k, r^k)$ for all $k \in \mathbb{N}$, there exists a bounded sequence of multipliers $\{\lambda^k\} \subset Y^*$ such that $\langle \lambda^k, y - G(x^k)\rangle_Y \leq r^k$ for all $y \in K$ and $k \in \mathbb{N}$ as well as $v^k = G'(x^k)^*\lambda^k \rightharpoonup^* \bar{v}$. Then $\bar{x}$ is w-AKKT regular.

Proof. (a) Let $\bar{v} \in w^*\text{-}\limsup_{\substack{x \to \bar{x} \\ r \searrow 0}} \mathcal{M}(x, r)$ be given. Then there exist sequences $\{x^k\} \subset X$, $\{r^k\} \subset [0, +\infty)$, and $\{v^k\} \subset X^*$ such that $x^k \to \bar{x}$, $r^k \searrow 0$, $v^k \rightharpoonup^* \bar{v}$, and $v^k \in \mathcal{M}(x^k, r^k)$ for all $k \in \mathbb{N}$. By assumption, there exist bounded sequences $\{\lambda^k\} \subset Y^*$ and $\{\mu^k\} \subset X^*$ with $\mu^k \in \mathcal{N}_C(x^k)$, $\langle \lambda^k, y - G(x^k)\rangle_Y \leq r^k$ for all $y \in K$ and $k \in \mathbb{N}$, and $v^k = G'(x^k)^*\lambda^k + \mu^k \rightharpoonup^* \bar{v}$. By the Banach–Alaoglu–Bourbaki theorem, the sequences $\{\lambda^k\}$ and $\{\mu^k\}$ possess weak* convergent subnets, indexed by $k(i)$, $i \in I$, where I is a directed set. The associated weak* limits are denoted by λ and μ, respectively. We get

$$\forall y \in K: \quad \langle \lambda, y - G(\bar{x})\rangle_Y \leftarrow \langle \lambda^{k(i)}, G(\bar{x}) - G(x^{k(i)})\rangle_Y + \langle \lambda^{k(i)}, y - G(\bar{x})\rangle_Y$$
$$= \langle \lambda^{k(i)}, y - G(x^{k(i)})\rangle_Y \leq r^{k(i)} \to 0.$$

Note that the first limit uses the boundedness of the net $\{\lambda^{k(i)}\}$, and this follows from the boundedness of the sequence $\{\lambda^k\}$. Thus, $\lambda \in \mathcal{N}_K(G(\bar{x}))$ holds. Similarly,

$$\forall x \in C: \quad \langle \mu, x - \bar{x}\rangle_X = \lim_{i \in I}\langle \mu^{k(i)}, x - x^{k(i)}\rangle_X \leq 0$$

implies $\mu \in \mathcal{N}_C(\bar{x})$. Again, we used the boundedness of $\{\mu^{k(i)}\}$ which follows from the boundedness of $\{\mu^k\}$. Hence, it holds that $G'(\bar{x})^*\lambda + \mu \in \mathcal{M}(\bar{x}, 0)$. Using again the boundedness of $\{\lambda^{k(i)}\}$ and $G'(x^{k(i)}) \to G'(\bar{x})$, we have

$$\langle G'(x^{k(i)})^*\lambda^{k(i)}, x\rangle_X = \langle \lambda^{k(i)}, G'(x^{k(i)})\, x\rangle_Y \to \langle \lambda, G'(\bar{x})\, x\rangle_Y = \langle G'(\bar{x})^*\lambda, x\rangle_Y$$

for all $x \in X$. This shows $G'(x^{k(i)})^*\lambda^{k(i)} \rightharpoonup^* G'(\bar{x})^*\lambda$. Due to the convergence $v^{k(i)} = G'(x^{k(i)})^*\lambda^{k(i)} + \mu^{k(i)} \rightharpoonup^* G'(\bar{x})^*\lambda + \mu$, we obtain from the uniqueness of the weak* limit point that $\bar{v} = G'(\bar{x})^*\lambda + \mu \in \mathcal{M}(\bar{x}, 0)$, i.e., $\bar{x}$ is sw-AKKT regular.

(b) This follows by almost the same proof. Note that the complete continuities of G and G' imply $G(x^{k(i)}) \to G(\bar{x})$ and $G'(x^{k(i)}) \to G'(\bar{x})$, respectively. $\square$

Unfortunately, the condition $C = X$ needed in the proof for the second statement regarding w-AKKT regularity is quite restrictive but cannot be omitted as long as X is infinite dimensional. Otherwise, by reprising the above proof strategy, we get bounded nets $\{x^{k(i)}\} \subset X$ and $\{\mu^{k(i)}\} \subset X^*$ satisfying $x^{k(i)} \rightharpoonup \bar{x}$, $\mu^{k(i)} \rightharpoonup^* \mu$, and $\mu^{k(i)} \in \mathcal{N}_C(x^{k(i)})$ for all $i \in I$. However, this is not enough to conclude that $\mu \in \mathcal{N}_C(\bar{x})$; see also Example 5.4.

5. Relations to other constraint qualifications. As pointed out in the previous section, weak, strong-weak, and strong AKKT regularity are CQs for (P) under some additional assumptions on the problem data. Therefore, the natural question arises of how these new CQs are related to existing ones. The most prominent case is discussed in subsection 5.1 where we show that Robinson's CQ implies sw-AKKT regularity. As it turns out, Robinson's CQ even implies w-AKKT regularity under some additional assumptions. Afterwards, we consider the relationship between sw-AKKT regularity and Abadie's CQ in subsection 5.2. Finally, we conclude that s-AKKT regularity implies that at least Guignard's constraint qualification holds at the reference point whenever X is reflexive and separable in subsection 5.3.

5.1. Robinson constraint qualification. The most common CQ in Banach spaces is Robinson's CQ which dates back to the seminal paper [33], where it was used to characterize variational stability of perturbed nonlinear systems in Banach spaces. The interpretation of this condition as a CQ in Banach space programming is due to [41]. The aim of this section is to show that Robinson's CQ is stronger than sw-AKKT regularity. As we will see later in subsection 6.3, it is strictly stronger than sw-AKKT regularity in general. To proceed, we first recall the definition of Robinson's constraint qualification.

DEFINITION 5.1. *We say that* Robinson's constraint qualification *(RCQ) holds at a feasible point* $\bar{x} \in \mathcal{F}$ *of* (P) *if the condition*

$$(5.1) \qquad Y = G'(\bar{x}) \mathcal{R}_C(\bar{x}) - \mathcal{R}_K(G(\bar{x}))$$

is valid.

In the theorem below, we show that validity of RCQ at some reference point always guarantees sw-AKKT regularity.

THEOREM 5.2. *Assume that RCQ is satisfied at a feasible point* $\bar{x} \in \mathcal{F}$ *of* (P). *Then* $\bar{x}$ *is sw-AKKT regular (and thus s-AKKT regular).*

Proof. We check that the assumption of Lemma 4.9(a) is satisfied. To this end, let sequences $\{x^k\} \subset X$, $\{r^k\} \subset [0, +\infty)$, and $\{v^k\} \subset X^*$ with $x^k \to \bar{x}$, $r^k \searrow 0$, $v^k \rightharpoonup^* \bar{v}$, and $v^k \in \mathcal{M}(x^k, r^k)$ for all $k \in \mathbb{N}$ be given. By definition of $\mathcal{M}(x^k, r^k)$, this implies the existence of sequences $\{\lambda^k\} \subset Y^*$ and $\{\mu^k\} \subset X^*$ with $\mu^k \in \mathcal{N}_C(x^k)$ and $\langle \lambda^k, y - G(x^k) \rangle_Y \le r^k$ for all $y \in K$ and $k \in \mathbb{N}$, as well as with $G'(x^k)^* \lambda^k + \mu^k = v^k \rightharpoonup^* \bar{v}$. It suffices to verify the boundedness of $\{\lambda^k\}$ and $\{\mu^k\}$.

Under assumption (5.1), we can apply the generalized open mapping theorem [41, Thm. 2.1] and obtain the existence of $M > 0$, such that for all $z \in Y$ with $\|z\|_Y \le 1$, there exist $w \in C \cap B_1(\bar{x})$ and $y \in K \cap B_1(G(\bar{x}))$ such that

$$-\frac{z}{M} = G'(\bar{x})(w - \bar{x}) - (y - G(\bar{x})).$$

We fix an arbitrary $z \in Y$ with $\|z\|_Y \le 1$ and the corresponding vectors w and y from above. Then let us write

$$-\frac{z}{M} = G'(x^k)(w - x^k) - (y - G(x^k)) + \delta^k$$

with $\delta^k := G'(\bar{x})(w - \bar{x}) - G'(x^k)(w - x^k) + G(\bar{x}) - G(x^k)$. We have the estimate

$$\|\delta^k\|_Y \le \|G'(\bar{x})(x^k - \bar{x})\|_Y + \|G'(\bar{x}) - G'(x^k)\| \, \|w - x^k\|_X + \|G(\bar{x}) - G(x^k)\|_Y$$
$$\le \|G'(\bar{x})(x^k - \bar{x})\|_Y + \|G'(\bar{x}) - G'(x^k)\| \, (1 + \|\bar{x} - x^k\|_X) + \|G(\bar{x}) - G(x^k)\|_Y$$
$$=: s^k,$$

where $s^k \to 0$ holds by continuity of G and G'. Note that s^k is independent of z. We have

$$\left\langle \lambda^k, \frac{z}{M} \right\rangle_Y = -\langle G'(x^k)^* \lambda^k, w - x^k \rangle_X + \langle \lambda^k, y - G(x^k) \rangle_Y - \langle \lambda^k, \delta^k \rangle_Y$$
$$\le \langle \mu^k - v^k, w - x^k \rangle_X + r^k + \|\delta^k\|_Y \, \|\lambda^k\|_{Y^*}$$
$$\le \|v^k\|_{X^*} (1 + \|\bar{x} - x^k\|_X) + r^k + s^k \, \|\lambda^k\|_{Y^*}.$$

In the last inequality, we used $\langle \mu^k, w - x^k \rangle_X \leq 0$ due to $\mu^k \in \mathcal{N}_C(x^k)$ and $w \in C$. Since $z \in Y$ with $\|z\|_Y \leq 1$ was arbitrary, and since the right-hand side in the above inequality is independent of z, this implies

$$\|\lambda^k\|_{Y^*} \leq M \left(\|v^k\|_{X^*} (1 + \|\bar{x} - x^k\|_X) + r^k + s^k \|\lambda^k\|_{Y^*} \right).$$

Due to $s^k \to 0$, we can conclude that

$$\|\lambda^k\|_{Y^*} \leq 2 M \left(\|v^k\|_{X^*} (1 + \|\bar{x} - x^k\|_X) + r^k \right)$$

for all large enough $k \in \mathbb{N}$. Since $\{v^k\}$ and $\{x^k\}$ are bounded, the boundedness of $\{\lambda^k\}$ follows. Finally, due to $\mu^k = v^k - G'(x^k)^* \lambda^k$, $\{\mu^k\}$ is bounded as well. □

In the theorem below, we investigate situations where validity of RCQ already implies w-AKKT regularity.

THEOREM 5.3. *Assume that RCQ is satisfied at a feasible point $\bar{x} \in \mathcal{F}$ of (P). Further, assume $C = X$ and that G and G' are completely continuous. Then $\bar{x}$ is w-AKKT regular.*

Proof. For the verification of Lemma 4.9(b), we can transfer the proof of Theorem 5.2 to the situation at hand. We have to use the boundedness of weakly and weak* convergent sequences. Moreover, the complete continuity of G and G' as well as the compactness of $G'(\bar{x})$ have to be used to conclude $s^k \to 0$, where s^k is defined as above. □

By means of an example, we show that RCQ does not imply w-AKKT regularity in general.

Example 5.4. We consider $X := \ell^2$ and its unit ball $C := \{x \in \ell^2 \mid \|x\|_{\ell^2} \leq 1\}$. We identify X^* with X. Furthermore, we assume that no further constraints are present, i.e., $Y = K = \{0\}$ and $G \colon X \to Y$ is the zero mapping. Then $\bar{x} := e^1/\sqrt{2}$ is an interior point of C and, consequently, (5.1) is satisfied at $\bar{x}$. We define the sequence $\{x^k\} \subset C$ by means of $x^k := (e^1 + e^{k+1})/\sqrt{2}$ for each $k \in \mathbb{N}$. Above, $e^n \in \ell^2$ denotes the nth unit sequence in ℓ^2 for each $n \in \mathbb{N}$. For $k \in \mathbb{N}$, it is easy to see that we have $x^k \in \mathcal{N}_C(x^k)$. In particular, we obtain $x^k \in \mathcal{M}(x^k, 0)$ for all $k \in \mathbb{N}$. From $x^k \rightharpoonup \bar{x}$, we infer

$$\bar{x} \in \text{w*-}\limsup_{\substack{x \to \bar{x} \\ r \searrow 0}} \mathcal{M}(x, r),$$

but $\bar{x} \notin \mathcal{M}(\bar{x}, 0) = \mathcal{N}_C(\bar{x}) = \{0\}$, where the first equality holds since the feasible set is defined by abstract constraints only, whereas the second equality exploits the fact that $\bar{x}$ is an interior point of C. Hence, $\bar{x}$ is not w-AKKT regular.

It is clear that similar examples can be constructed in all infinite-dimensional Hilbert spaces.

5.2. Abadie constraint qualification. In the finite-dimensional setting, it was shown in [7] that in the case of its presence, AKKT regularity implies validity of Abadie's CQ which originates from [1]. Here, we want to generalize this observation to the infinite-dimensional situation. Let us first state an appropriate notion of Abadie's CQ which applies to the general situation discussed in this paper.

DEFINITION 5.5. *Let $\bar{x} \in \mathcal{F}$ be a feasible point of (P). We say that* Abadie's *constraint qualification (ACQ) is valid at $\bar{x}$ if*

$$(5.2) \qquad \mathcal{T}_{\mathcal{F}}(\bar{x}) = \mathcal{L}_{\mathcal{F}}(\bar{x})$$

holds and $\mathcal{M}(\bar{x}, 0)$ is weak* closed.

Recall that $\mathcal{L}_{\mathcal{F}}(\bar{x})$ from (2.1) denotes the linearization cone to $\mathcal{F}$ at $\bar{x}$, and that $\mathcal{M}(\bar{x}, 0)$ can be used to characterize the KKT conditions; see (4.2), which shows that the relation

$$\mathcal{M}(\bar{x}, 0) = G'(\bar{x})^* \mathcal{N}_K(G(\bar{x})) + \mathcal{N}_C(\bar{x})$$

holds at the given feasible point $\bar{x} \in \mathcal{F}$. We note that demanding $\mathcal{M}(\bar{x}, 0)$ to be weakly* closed is, in general, indispensable in the definition of ACQ in order to guarantee that it is a CQ in the narrower sense. Indeed, the polarity relation

$$(5.3) \qquad \mathcal{L}_{\mathcal{F}}(\bar{x})^\circ = G'(\bar{x})^* \mathcal{N}_K(G(\bar{x})) + \mathcal{N}_C(\bar{x}) = \mathcal{M}(\bar{x}, 0),$$

which holds automatically in the context of standard finite-dimensional nonlinear programming, only holds if $\mathcal{M}(\bar{x}, 0)$ is weakly* closed; see [23, sect. 2] or [39, Lem. 1]. Observe that in the case where X is reflexive, the closedness of $\mathcal{M}(\bar{x}, 0)$ already yields its weak* closedness; see Remark 4.5 as well. The above arguments yield the following well-known result, which is included for the reader's convenience.

PROPOSITION 5.6. *Let $\bar{x} \in \mathcal{F}$ be a local minimizer of (P) where ACQ holds. Then $\bar{x}$ is a KKT point of (P).*

Proof. Since $\bar{x} \in \mathcal{F}$ is a local minimizer of (P), it holds that $f'(\bar{x})d \geq 0$ for all directions $d \in \mathcal{T}_{\mathcal{F}}^w(\bar{x})$, i.e., $-f'(\bar{x}) \in \mathcal{T}_{\mathcal{F}}^w(\bar{x})^\circ$. Recalling that $\mathcal{T}_{\mathcal{F}}(\bar{x}) \subset \mathcal{T}_{\mathcal{F}}^w(\bar{x}) \subset \mathcal{L}_{\mathcal{F}}(\bar{x})$ holds in general, the validity of ACQ guarantees $\mathcal{T}_{\mathcal{F}}^w(\bar{x}) = \mathcal{L}_{\mathcal{F}}(\bar{x})$. Thus, the above arguments show $\mathcal{T}_{\mathcal{F}}^w(\bar{x})^\circ = \mathcal{M}(\bar{x}, 0)$, i.e., $-f'(\bar{x}) \in \mathcal{M}(\bar{x}, 0)$ follows. Hence, $\bar{x}$ is a KKT point of (P). □

From this proof it is clear that (5.2) could be weakened to $\mathcal{T}_{\mathcal{F}}^w(\bar{x}) = \mathcal{L}_{\mathcal{F}}(\bar{x})$ in the definition of ACQ.

In order to prove the main result of this section, we need two technical preliminaries, which are provided below.

LEMMA 5.7. *Suppose that X^* is separable. Then there is an equivalent norm on X which is continuously Fréchet differentiable in $X \setminus \{0\}$.*

Proof. Combine the results from [16, Cor. 8.5 and Thm. 8.19]. □

In the remaining parts of this section, we will assume that the space X is equipped with the norm from the above lemma. Recall that reflexivity and separability of X together imply separability of X^*.

LEMMA 5.8. *Suppose that X is reflexive and separable. Fix a feasible point $\bar{x} \in \mathcal{F}$ of (P). For every $v \in \widehat{\mathcal{N}}_{\mathcal{F}}(\bar{x})$ there is a continuously Fréchet differentiable function $h \colon X \to \mathbb{R}$ with $h'(\bar{x}) = -v$ such that h restricted to $\mathcal{F}$ achieves a unique global minimum at $\bar{x}$. Moreover, the function h is weakly sequentially lower semicontinuous.*

Proof. It is well known that every differentiable convex function is already continuously differentiable; cf. [31, Cor. of Prop. 2.8]. The above assertion therefore follows directly from [27, Thm. 1.30(ii)], since X admits a continuous differentiable renorm by Lemma 5.7. □

Now, we can transfer the proof of [7, Thm. 4.4] to the infinite-dimensional setting.

THEOREM 5.9. *Let X be reflexive and separable. Let us assume that $\bar{x}$ is an sw-AKKT regular feasible point $\bar{x} \in \mathcal{F}$ of (P). Furthermore, assume that condition* (3.2) *holds. Then ACQ is valid at $\bar{x}$.*

Proof. In a preliminary step, we first verify the inclusion $\mathcal{N}_{\mathcal{F}}^{\mathrm{L}}(\bar{x}) \subset \mathcal{M}(\bar{x}, 0)$. Let $v \in \mathcal{N}_{\mathcal{F}}^{\mathrm{L}}(\bar{x})$ be given. Then, due to reflexivity of X, there exist sequences $\{x^k\} \subset \mathcal{F}$ and $\{v^k\} \subset X^*$ such that $x^k \to \bar{x}$, $v^k \rightharpoonup v$, and $v^k \in \widehat{\mathcal{N}}_{\mathcal{F}}(x^k)$ for all $k \in \mathbb{N}$. Invoking Lemma 5.8 for each $k \in \mathbb{N}$, we see that there exists a function $h^k \colon X \to \mathbb{R}$ such that x^k is the constrained minimizer of h^k over the feasible set $\mathcal{F}$ and that $(h^k)'(x^k) = -v^k$. An inspection of the corresponding proof shows that Proposition 3.3 guarantees the existence of sequences $\{x^{k,\ell}\} \subset C$ and $\{v^{k,\ell}\} \subset X^*$ such that

$$x^{k,\ell} \to x^k, \qquad v^{k,\ell} \to v^k, \qquad v^{k,\ell} \in \mathcal{M}(x^{k,\ell}, 0) \quad \forall \ell \in \mathbb{N}.$$

Thus, we can pick diagonal sequences $\{x^{k,\ell(k)}\}$ and $\{v^{k,\ell(k)}\}$ with

$$x^{k,\ell(k)} \to \bar{x}, \quad v^{k,\ell(k)} \rightharpoonup v.$$

Naturally, we have $v^{k,\ell(k)} \in \mathcal{M}(x^{k,\ell(k)}, 0)$ for each $k \in \mathbb{N}$. Now, sw-AKKT regularity yields $v \in \mathcal{M}(\bar{x}, 0)$. This shows $\mathcal{N}_{\mathcal{F}}^{\mathrm{L}}(\bar{x}) \subset \mathcal{M}(\bar{x}, 0)$.

To verify the statement of the theorem, note that Remark 4.5 and the sw-AKKT regularity imply the weak* closedness of $\mathcal{M}(\bar{x}, 0)$ since X is assumed to be reflexive. As pointed out above, this yields the polar relationship $\mathcal{M}(\bar{x}, 0) = \mathcal{L}_{\mathcal{F}}(\bar{x})^\circ$; cf. (5.3). Using our preliminary step, we therefore have $\mathcal{N}_{\mathcal{F}}^{\mathrm{L}}(\bar{x}) \subset \mathcal{M}(\bar{x}, 0) = \mathcal{L}_{\mathcal{F}}(\bar{x})^\circ$. Taking polars yields $\mathcal{N}_{\mathcal{F}}^{\mathrm{L}}(\bar{x})^\circ \supset \mathcal{L}_{\mathcal{F}}(\bar{x})$. Furthermore, [27, Thm. 3.57] guarantees $\mathcal{N}_{\mathcal{F}}^{\mathrm{L}}(\bar{x})^\circ = \mathcal{T}_{\mathcal{F}}^{\mathrm{C}}(\bar{x})$. Hence, we get the chain of inclusions

$$\mathcal{L}_{\mathcal{F}}(\bar{x}) \subset \mathcal{T}_{\mathcal{F}}^{\mathrm{C}}(\bar{x}) \subset \mathcal{T}_{\mathcal{F}}(\bar{x}) \subset \mathcal{T}_{\mathcal{F}}^{w}(\bar{x}) \subset \mathcal{L}_{\mathcal{F}}(\bar{x}),$$

and this finishes the proof. $\qquad \square$

The following corollary can be distilled from the proof of Theorem 5.9.

COROLLARY 5.10. *Let X be reflexive and separable. Let $\bar{x} \in \mathcal{F}$ be a feasible sw-AKKT regular point of (P). Finally, assume that condition* (3.2) *holds. Then we have the equalities*

$$\mathcal{L}_{\mathcal{F}}(\bar{x}) = \mathcal{T}_{\mathcal{F}}^{\mathrm{C}}(\bar{x}) = \mathcal{T}_{\mathcal{F}}(\bar{x}) = \mathcal{T}_{\mathcal{F}}^{w}(\bar{x}).$$

5.3. Guignard constraint qualification. Let us first recall the definition of Guignard's CQ which can be traced back to [18].

DEFINITION 5.11. *Let $\bar{x} \in \mathcal{F}$ be a feasible point of (P). We say that* Guignard's constraint qualification *(GCQ) is valid at $\bar{x}$ if*

$$\mathcal{T}_{\mathcal{F}}^{w}(\bar{x})^\circ = \mathcal{M}(\bar{x}, 0)$$

holds.

Recall that our definition of GCQ is not necessarily equivalent to the requirement $\mathcal{T}_{\mathcal{F}}^{w}(\bar{x})^\circ = \mathcal{L}_{\mathcal{F}}(\bar{x})^\circ$. In fact, this is true only if $\mathcal{M}(\bar{x}, 0)$ is weak* closed; cf. (5.3). By polarity, Definition 5.11 immediately implies that $\mathcal{M}(\bar{x}, 0)$ is weak* closed, whereas this does not follow from the alternative definition that $\mathcal{T}_{\mathcal{F}}^{w}(\bar{x})^\circ = \mathcal{L}_{\mathcal{F}}(\bar{x})^\circ$.

Inspecting the proof of Proposition 5.6, it is clear that GCQ is indeed a constraint qualification. Furthermore, under the assumption that X is reflexive, GCQ is

the weakest CQ which ensures that $\bar{x}$ is a KKT point for all functions $\hat{f}$ which are differentiable at $\bar{x}$ and for which $\bar{x}$ is a local minimizer of $\hat{f}$ over $\mathcal{F}$; see [17, Cor. 3.4]. We show that s-AKKT regularity implies GCQ if X is additionally separable.

THEOREM 5.12. *Let X be reflexive and separable. Let us assume that $\bar{x} \in \mathcal{F}$ is a feasible s-AKKT regular point of* (P). *Furthermore, assume that condition* (3.2) *holds. Then GCQ is valid at $\bar{x}$.*

Proof. From $\mathcal{T}_{\mathcal{F}}^w(\bar{x}) \subset \mathcal{L}_{\mathcal{F}}(\bar{x})$ we obtain $\mathcal{T}_{\mathcal{F}}^w(\bar{x})^\circ \supset \mathcal{L}_{\mathcal{F}}(\bar{x})^\circ = \mathcal{M}(\bar{x}, 0)$ by taking polars since $\mathcal{M}(\bar{x}, 0)$ is closed by the postulated s-AKKT regularity; cf. Remark 4.5. For $v \in \mathcal{T}_{\mathcal{F}}^w(\bar{x})^\circ = \widehat{\mathcal{N}}_{\mathcal{F}}(\bar{x})$, we can argue similarly to the proof of Theorem 5.9: Lemma 5.8 yields a continuously differentiable function $h : X \to \mathbb{R}$ such that h restricted to $\mathcal{F}$ achieves a unique global minimum at $\bar{x}$ and $h'(\bar{x}) = -v$. Now, we can apply Corollary 4.8 to the objective h and obtain $v = -h'(\bar{x}) \in \mathcal{M}(\bar{x}, 0)$. This finishes the proof. $\square$

In Figure 5.1, we summarize the relations between all types of AKKT regularity from Definition 4.4 as well as RCQ, ACQ, and GCQ in the context of a reflexive Banach space X. In light of Proposition 3.3 and Example 3.6, the reflexivity assumption on X is indispensable whenever sequential CQs are under consideration.
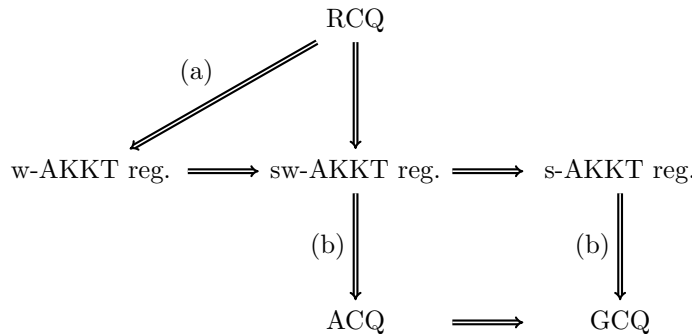

FIG. 5.1. *Relations between CQs in the setting where X is reflexive. Relations with labeled arrows only hold under the additional assumptions:* (a) *requires complete continuity of G and G' as well as $C = X$,* (b) *holds whenever X is separable and* (3.2) *holds.*

We close this section by pointing out one open problem, namely whether s-AKKT regularity already implies ACQ. The previous technique of proof does not yield this implication, on the other hand, we also were not able to find a counterexample.

6. AKKT regularity in exemplary settings. In this section, we present three practically relevant settings where w-AKKT regularity, or at least sw-AKKT regularity, is inherent or can be checked via evaluation of reasonable conditions. First, we prove in subsection 6.1 that w-AKKT regularity is automatically satisfied in the setting where the feasible set $\mathcal{F}$ is defined via affine equality constraints which are induced by a bounded linear operator with a closed range. Subsection 6.2 discusses the situation where we have nonlinear equality constraints which model $\mathcal{F}$. Finally, in subsection 6.3, we will investigate the important setting of two-sided (pointwise) box constraints in Lebesgue spaces. It will be shown that sw-AKKT regularity holds in this situation as well (whereas RCQ is known to be violated for this class of problems).

6.1. Affine equality constraints. Let X and Y be Banach spaces such that X is reflexive. Furthermore, fix an operator $A \in \mathbb{L}(X, Y)$ with closed range and some

vector $b \in Y$. We consider the affine equality constraint

$$Ax = b.$$

In this situation, it holds that $K := \{0\}$, $C := X$, and $G(x) := Ax - b$ for all $x \in X$. Clearly, RCQ holds for this constraint system if and only if A is surjective, i.e., if $\operatorname{ran} A = Y$.

By $K = \{0\}$, it clearly holds that

$$\forall x \in X \ \forall r \in [0, +\infty): \quad \mathcal{M}(x, r) = \{A^* \lambda \,|\, \lambda \in Y^*, \, -\langle \lambda, Ax - b \rangle_Y \leq r\},$$

i.e., for each point $\bar{x} \in X$ satisfying $A\bar{x} = b$, we obtain

$$\forall r \in [0, +\infty): \quad \mathcal{M}(\bar{x}, r) = A^* Y^*.$$

Particularly, $\mathcal{M}(x, r) \subset \mathcal{M}(\bar{x}, 0)$ is obtained for arbitrary $x \in X$ and $r \in [0, +\infty)$. Obviously, $\mathcal{M}(\bar{x}, 0) = A^* Y^*$ is convex. Furthermore, $\mathcal{M}(\bar{x}, 0)$ is closed by closedness of AX and the closed range theorem. Since X is reflexive, the same holds true for X^*, and, consequently, $\mathcal{M}(\bar{x}, 0)$ is a weakly* closed subset of X^*. Particularly, we obtain w-AKKT regularity of each point which satisfies the affine constraint.

In view of Proposition 3.3 and Theorem 4.6, the above observation recovers the classical result [23, Thm. 4.2].

6.2. Nonlinear equality constraints. We consider the special constraint system

$$(6.1) \qquad\qquad G(x) = 0,$$

i.e., we fix $C := X$ and $K := \{0\}$. In contrast to subsection 6.1, here this equality constraint is allowed to be nonlinear. Clearly, RCQ holds at a feasible point $\bar{x} \in X$ of the constraint system (6.1) if and only if $G'(\bar{x})$ is surjective.

The next result utilizes the *reduced minimum modulus* introduced in [22, Chap. IV, sect. 5]. For a bounded linear operator $T \in \mathbb{L}(X, Y)$, it is defined via

$$\gamma(T) := \inf\big\{\|T x\|_Y \,\big|\, x \in X, \, \operatorname{dist}(x, \ker T) = 1\big\}.$$

It is well known that the range of T is closed if and only if $\gamma(T) > 0$; see [22, Thm. IV.5.2]. Hence, $\gamma(T)$ can be used as a "quantitative measure of closedness" of the range of T. Moreover, $\gamma(T) = \gamma(T^*)$ holds, and this is a quantitative version of the closed range theorem; see [22, Thm. IV.5.13]. Finally, we have the inequality

$$(6.2) \qquad \forall x \in X: \quad \operatorname{dist}(x, \ker T) = \inf_{v \in \ker T} \|x - v\|_X \leq \frac{1}{\gamma(T)} \|T x\|_Y$$

in the case when $\gamma(T) > 0$. In finite dimensions, $\gamma(T)$ coincides with the reciprocal of the norm of the Moore–Penrose inverse of the matrix T.

PROPOSITION 6.1. *Let $\bar{x} \in \mathcal{F}$ be a feasible point of the constraint system* (6.1). *Furthermore, suppose that there are a neighborhood $U \subset X$ of $\bar{x}$ and some $\beta > 0$ such that*

$$(6.3) \qquad\qquad \forall x \in U: \quad \gamma(G'(x)) \geq \beta$$

is valid. Then $\bar{x}$ is sw-AKKT regular.

Proof. For the proof, we are going to exploit arguments similar to those used in the validation of Lemma 4.9. Therefore, we fix sequences $\{x^k\} \subset X$, $\{r^k\} \subset [0, +\infty)$, and $\{v^k\} \subset X^*$ with $x^k \to \bar{x}$, $r^k \searrow 0$, $v^k \rightharpoonup^* \bar{v}$, and $v^k \in \mathcal{M}(x^k, r^k)$ for all $k \in \mathbb{N}$. Due to $K = \{0\}$, we find a sequence $\{\lambda^k\} \subset Y^*$ such that $v^k = G'(x^k)^* \lambda^k$ and $-\langle \lambda^k, G(x^k) \rangle_Y \leq r^k$ hold for all $k \in \mathbb{N}$. By $v^k \rightharpoonup^* \bar{v}$, $\{G'(x^k)^* \lambda^k\}$ is bounded. Due to the assumptions of the proposition and (6.2), we find some constant $c > 0$ such that

$$\inf_{\mu^k \in \ker G'(x^k)^*} \|\lambda^k - \mu^k\|_{Y^*} \leq \frac{1}{\beta} \|G'(x^k)^* \lambda^k\|_{X^*} \leq c$$

for large enough $k \in \mathbb{N}$ since $x^k \to \bar{x}$ holds. Thus, we find a bounded sequence $\{\hat{\lambda}^k\} \subset Y^*$ with $\lambda^k - \hat{\lambda}^k \in \ker G'(x^k)^*$, i.e., $v^k = G'(x^k)^* \hat{\lambda}^k$ for all $k \in \mathbb{N}$. Due to the boundedness, we obtain a subnet with $\hat{\lambda}^{k(i)} \rightharpoonup^* \hat{\lambda}$ for some $\hat{\lambda} \in Y^*$. Exploiting the continuity of G' and the boundedness of $\{\hat{\lambda}^{k(i)}\}$, $G'(x^{k(i)})^* \hat{\lambda}^{k(i)} \rightharpoonup^* G'(\bar{x})^* \hat{\lambda}$ holds; see the proof of Lemma 4.9(a). The uniqueness of weak* limits yields $G'(\bar{x})^* \hat{\lambda} = \bar{v}$, and this shows $\bar{v} \in \mathcal{M}(\bar{x}, 0)$. Hence, we have shown w*-$\limsup_{\substack{x \to \bar{x} \\ r \to 0}} \mathcal{M}(x, r) \subset \mathcal{M}(\bar{x}, 0)$, i.e., $\bar{x}$ is sw-AKKT regular. □

Unfortunately, the mapping $T \mapsto \gamma(T)$ is, in general, only upper semicontinuous; e.g., for

$$T_\varepsilon := \begin{pmatrix} 1 & 0 \\ 0 & \varepsilon \end{pmatrix} \in \mathbb{R}^{2 \times 2}$$

we have $\gamma(T_0) = 1$ but $\gamma(T_\varepsilon) = \varepsilon^{-1}$ for $\varepsilon > 0$. In finite dimensions, γ is continuous on the set of matrices with constant rank, and this justifies that (6.3) can be interpreted as some kind of *constant rank constraint qualification*. In this regard, let us mention that Proposition 6.1 generalizes the results of [11] since we do not need the fact that the kernel and range of $G'(\bar{x})$ are complemented.

In infinite dimensions, conditions ensuring continuity of γ can be found in [40, sect. 3] and [38, sect. 6.1]. To formulate these conditions, we introduce the gap between subspaces $U, V \subset X$ via

$$\delta(U, V) := \sup\{\operatorname{dist}(x, V) \mid x \in U, \|x\|_X = 1\},$$

with the convention $\delta(\{0\}, V) = 0$. We have

$$\gamma(T^k) \geq \gamma(T) \frac{1 - \delta(\ker T, \ker T^k)}{1 + \delta(\ker T, \ker T^n)} - \|T - T^k\|$$

for operators $T, T^k \in \mathbb{L}(X, Y)$. If the range of T^k is closed, we also have

$$\gamma(T^k) \geq \gamma(T) \frac{1 - \delta(\operatorname{ran} T^k, \operatorname{ran} T)}{1 + \delta(\operatorname{ran} T^k, \operatorname{ran} T)} - \|T - T^k\|.$$

These two estimates can be found in [38, Prop. 6.1.5]. The first estimate is, in particular, applicable if $\ker T = \ker T^k$ since this implies $\delta(\ker T, \ker T^k) = 0$. Further, we have the trivial implications

$$\ker T = \{0\} \quad \implies \quad \delta(\ker T, \ker T^k) = 0,$$
$$\operatorname{ran} T = Y \quad \implies \quad \delta(\operatorname{ran} T^k, \operatorname{ran} T) = 0,$$

and under any of the left-hand side requirements, we have $\gamma(T^k) \geq \gamma(T) - \|T - T^k\|$. Moreover, it holds that

$$|\gamma(T) - \gamma(T^k)| \leq \|T - T^k\|$$

if T has closed range and if any of the following assumptions is satisfied:

1. $\dim \ker T = \dim \ker T^k < \infty$,
2. $\dim \operatorname{ran} T = \dim \operatorname{ran} T^k < \infty$,
3. $\operatorname{ran} T^k$ is closed and $\dim \operatorname{coker} T = \dim \operatorname{coker} T^k < \infty$;

see [38, Prop. 6.1.6].

6.3. Box constraints in Lebesgue spaces. Let $\Omega \subset \mathbb{R}^d$ be a bounded open set. For functions $u_a, u_b \in L^2(\Omega)$ such that $u_a \leq u_b$ holds almost everywhere on Ω, we consider the box constraints

$$u_a \leq u \leq u_b \quad \text{a.e. on } \Omega.$$

Here u is a function from $L^2(\Omega)$. Using $X := L^2(\Omega)$, $Y := L^2(\Omega)^2$,

$$K := \{(y_1, y_2) \in L^2(\Omega)^2 \mid y_1, y_2 \geq 0 \text{ a.e. on } \Omega\},$$

and $G(u) := (u_b - u, u - u_a)$ for each $u \in L^2(\Omega)$, the above box constraints can be described equivalently via $G(u) \in K$. For completeness, let us set $C := L^2(\Omega)$. It is easy to see that RCQ does not hold in this setting; see [35, section 6.1.2]. However, we will show below that all associated feasible points are sw-AKKT regular. To this end, we will make use of the strategy proposed in Lemma 4.9.

The associated set-valued mapping $\mathcal{M} \colon L^2(\Omega) \times [0, +\infty) \rightrightarrows L^2(\Omega)$ is given by

$$\mathcal{M}(u, r) = \left\{ \lambda_a - \lambda_b \,\middle|\, \begin{array}{l} \lambda_a, \lambda_b \in L^2(\Omega), \ \lambda_a, \lambda_b \leq 0 \text{ a.e. on } \Omega, \\ -\langle \lambda_b, u_b - u \rangle_{L^2(\Omega)} - \langle \lambda_a, u - u_a \rangle_{L^2(\Omega)} \leq r \end{array} \right\}$$

for all $u \in L^2(\Omega)$ and $r \in [0, +\infty)$; see Remark 4.1. Fix some point $\bar{u} \in L^2(\Omega)$ satisfying $G(\bar{u}) \in K$. Furthermore, let $\{u^k\} \subset L^2(\Omega)$, $\{r^k\} \subset [0, +\infty)$, and $\{\lambda^k\} \subset L^2(\Omega)$ be sequences such that $\lambda^k \in \mathcal{M}(u^k, r^k)$ for each $k \in \mathbb{N}$ as well as $u^k \to \bar{u}$ in $L^2(\Omega)$, $r^k \searrow 0$ in $\mathbb{R}$, and $\lambda^k \rightharpoonup \bar{\lambda}$ in $L^2(\Omega)$ for some $\bar{\lambda} \in L^2(\Omega)$. By definition of $\mathcal{M}$, we find sequences $\{\lambda_a^k\}, \{\lambda_b^k\} \subset L^2(\Omega)$ satisfying $\lambda^k = \lambda_a^k - \lambda_b^k$, $\lambda_a^k, \lambda_b^k \leq 0$ almost everywhere on Ω, and

$$(6.4) \qquad -\langle \lambda_b^k, u_b - u^k \rangle_{L^2(\Omega)} - \langle \lambda_a^k, u^k - u_a \rangle_{L^2(\Omega)} \leq r^k$$

for all $k \in \mathbb{N}$. Next, let us set $\tilde{\lambda}_a^k := \min(\lambda^k, 0)$ and $\tilde{\lambda}_b^k := -\max(\lambda^k, 0)$ where min and max have to be interpreted in a pointwise sense. Clearly, it holds that $\lambda^k = \tilde{\lambda}_a^k - \tilde{\lambda}_b^k$ for all $k \in \mathbb{N}$. By construction, we additionally have $0 \geq \tilde{\lambda}_a^k \geq \lambda_a^k$ as well as $0 \geq \tilde{\lambda}_b^k \geq \lambda_b^k$ for all $k \in \mathbb{N}$. Some rearrangements in (6.4) as well as $u_a - u_b \leq 0$ almost everywhere on Ω yield

$$
\begin{aligned}
r^k &\geq \langle \lambda_b^k - \lambda_a^k, u^k \rangle_{L^2(\Omega)} - \langle \lambda_b^k, u_b \rangle_{L^2(\Omega)} + \langle \lambda_a^k, u_a \rangle_{L^2(\Omega)} \\
&= \langle \lambda_b^k - \lambda_a^k, u^k \rangle_{L^2(\Omega)} + \langle \lambda_a^k - \lambda_b^k, u_b \rangle_{L^2(\Omega)} + \langle \lambda_a^k, u_a - u_b \rangle_{L^2(\Omega)} \\
&\geq \langle \tilde{\lambda}_b^k - \tilde{\lambda}_a^k, u^k \rangle_{L^2(\Omega)} + \langle \tilde{\lambda}_a^k - \tilde{\lambda}_b^k, u_b \rangle_{L^2(\Omega)} + \langle \tilde{\lambda}_a^k, u_a - u_b \rangle_{L^2(\Omega)} \\
&= \langle \tilde{\lambda}_b^k - \tilde{\lambda}_a^k, u^k \rangle_{L^2(\Omega)} - \langle \tilde{\lambda}_b^k, u_b \rangle_{L^2(\Omega)} + \langle \tilde{\lambda}_a^k, u_a \rangle_{L^2(\Omega)} \\
&= -\langle \tilde{\lambda}_b^k, u_b - u^k \rangle_{L^2(\Omega)} - \langle \tilde{\lambda}_a^k, u^k - u_a \rangle_{L^2(\Omega)}
\end{aligned}
$$

$$(6.5)$$

for each $k \in \mathbb{N}$. As a consequence, we can exploit the sequences $\{\tilde{\lambda}_a^k\}$ and $\{\tilde{\lambda}_b^k\}$ in order to represent $\lambda^k \in \mathcal{M}(u^k, r^k)$. Due to $\lambda^k \rightharpoonup \bar{\lambda}$ in $L^2(\Omega)$, the sequence $\{\lambda^k\}$ is bounded in $L^2(\Omega)$. Thus, the trivial estimates $\|\tilde{\lambda}_a^k\|_{L^2(\Omega)} \leq \|\lambda^k\|_{L^2(\Omega)}$ and $\|\tilde{\lambda}_b^k\|_{L^2(\Omega)} \leq \|\lambda^k\|_{L^2(\Omega)}$ show that $\{\tilde{\lambda}_a^k\}$ and $\{\tilde{\lambda}_b^k\}$ are bounded in $L^2(\Omega)$, too. Thus, these sequences possess weakly convergent subsequences. We assume w.l.o.g. that the convergences $\tilde{\lambda}_a^k \rightharpoonup \lambda_a$ and $\tilde{\lambda}_b^k \rightharpoonup \lambda_b$ hold true. By weak sequential closedness of $\{v \in L^2(\Omega) \,|\, v \leq 0 \text{ a.e. on } \Omega\}$, we have $\lambda_a, \lambda_b \leq 0$ almost everywhere on Ω. Since it holds that $\lambda^k \rightharpoonup \bar{\lambda}$ and $\lambda^k = \tilde{\lambda}_a^k - \tilde{\lambda}_b^k$, uniqueness of the weak limit yields $\bar{\lambda} = \lambda_a - \lambda_b$. Taking the limit $k \to \infty$ in (6.5) implies

$$0 \geq -\langle \lambda_b, u_b - \bar{u} \rangle_{L^2(\Omega)} - \langle \lambda_a, \bar{u} - u_a \rangle_{L^2(\Omega)},$$

where we used the convergences $u^k \to \bar{u}$, $\tilde{\lambda}^k \rightharpoonup \lambda_a$, and $\tilde{\lambda}_b^k \rightharpoonup \lambda_b$ in $L^2(\Omega)$. Due to $\bar{\lambda} = \lambda_a - \lambda_b$, we have shown $\bar{\lambda} \in \mathcal{M}(\bar{u}, 0)$. Particularly, $\bar{u}$ is sw-AKKT regular.

7. Application to safeguarded augmented Lagrangian methods. In this section, we want to show that the safeguarded augmented Lagrangian method (ALM), applied to (P), generates a w-AKKT sequence under appropriate assumptions, and we deduct consequences for the convergence behavior from this observation. The safeguarded ALMs have become popular for finite- and infinite-dimensional optimization problems, see [2, 3, 4, 10] for the finite-dimensional perspective and [13, 21] for the infinite-dimensional view. Since ALMs are at their core Hilbert space methods, we presume that the constraint space Y from (P) is densely embedded in a Hilbert space H such that we have the Gelfand triple structure $Y \overset{d}{\hookrightarrow} H = H^* \overset{d}{\hookrightarrow} Y^*$. Further, we require that the constraint is well represented in the Hilbert space; i.e., we assume that there is a closed convex set $\mathcal{K} \subset H$ such that $e^{-1}(\mathcal{K}) = K$, where e represents the dense embedding $Y \overset{d}{\hookrightarrow} H$. Thus, problem (P) is equivalent to

$$(P_H) \qquad \underset{x \in C}{\text{minimize}} \quad f(x) \quad \text{subject to} \quad e(G(x)) \in \mathcal{K}.$$

For better readability, the embedding e will be omitted in what follows.

We now turn to the multiplier-penalty method for the optimization problem (P). To this end, we define the augmented Lagrangian of (P_H) as follows.

DEFINITION 7.1. *For $\rho > 0$, the* augmented Lagrange function *or* augmented Lagrangian *of (P_H) is the function $L_\rho \colon X \times H \to \mathbb{R}$ defined by*

$$(7.1) \qquad \forall x \in X \,\forall \lambda \in H: \quad L_\rho(x, \lambda) := f(x) + \frac{\rho}{2} d_{\mathcal{K}}^2\left(G(x) + \frac{\lambda}{\rho}\right) - \frac{\|\lambda\|_H^2}{2\rho}.$$

Note that there are other variants of L_ρ in the literature. However, these differ from (7.1) only by an additive constant (w.r.t. x).

For the construction of our algorithm, we will need a means of controlling the penalty parameter. To this end, we define the utility function

$$\forall x \in X \,\forall \lambda \in H \,\forall \rho > 0: \quad V(x, \lambda, \rho) := \left\| G(x) - P_{\mathcal{K}}\big(G(x) + \rho^{-1}\lambda\big) \right\|_H.$$

This definition enables us to formulate our algorithm as follows.

ALGORITHM 7.2 (ALM for constrained optimization). *Let $(x^0, \lambda^0) \in X \times H$, $\rho^0 > 0$, and a nonempty, bounded set $B \subset H$ be given. Furthermore, fix $\gamma > 1$ and $\tau \in (0, 1)$, and set $k := 0$.*

Step 1. *If (x^k, λ^k) satisfies a suitable termination criterion: STOP.*

Step 2. *Choose $w^k \in B$ and compute an approximate solution x^{k+1} of the problem*

$$(7.2) \qquad \underset{x \in C}{\text{minimize}}\ L_{\rho^k}(x, w^k).$$

Step 3. *Update the vector of multipliers to*

$$\lambda^{k+1} := \rho^k \left[G(x^{k+1}) + \frac{w^k}{\rho^k} - P_{\mathcal{K}}\left(G(x^{k+1}) + \frac{w^k}{\rho^k} \right) \right].$$

Step 4. *Let $V^{k+1} := V(x^{k+1}, w^k, \rho^k)$, and set*

$$\rho^{k+1} := \begin{cases} \rho^k & \text{if } k = 0 \text{ or } V^{k+1} \le \tau V^k, \\ \gamma \rho^k & \text{otherwise.} \end{cases}$$

Step 5. *Set $k \leftarrow k + 1$ and go to Step 1.*

Note that Algorithm 7.2 differs from the classical ALM by the introduction of the bounded sequence $\{w^k\}$. The classical method is obtained by replacing w^k by λ^k everywhere. Hence, both methods coincide whenever one takes $w^k = \lambda^k$ as long as λ^k remains bounded, say $\lambda^k \in B$ for the user-specified set B from Algorithm 7.2. For an unbounded sequence $\{\lambda^k\}$, however, the global convergence properties of the above (safeguarded) ALM are stronger than for the classical ALM; cf. the example given in [19].

Let us stress that the sequence $\{\lambda^k\}$ generated by Algorithm 7.2 belongs to the Hilbert space H but will usually be viewed as a sequence in the (larger) space Y^* since boundedness of this sequence is usually easier to verify in this dual space than in H itself (note that, despite the boundedness of $\{w^k\}$, the sequence $\{\lambda^k\}$ might still be unbounded). The reader might therefore wonder why we introduce the Gelfand structure $Y \hookrightarrow H \hookrightarrow Y^*$ with a Hilbert space H. The reason for that is twofold. On the one hand, the ALM is mainly a Hilbert space technique due to the fact that we compute projections (and Y itself might not be a Hilbert space). On the other hand, the presence of H gives us more freedom for the design of the actual method. If Y itself is a Hilbert space, it seems very natural at first glance to take $H := Y$. However, if $Y := H_0^1(\Omega)$ were to be taken as the Hilbert space H, we would have to compute projections w.r.t. the norm in $H_0^1(\Omega)$, and these projections are expensive to calculate. In this case, it is a nearby idea to embed the Sobolev space Y into $H := L^2(\Omega)$, where projections are usually much cheaper to compute.

So far, we have not specified what constitutes an "approximate solution" in Step 2 of Algorithm 7.2. Clearly, there are multiple possibilities when solving the subproblem (7.2); for instance, we could look for global minima or KKT points. Here, we are only interested in the case where the subproblems are solved by computing inexact KKT points. To this end, we state the following (natural) assumption regarding the quality by which the subproblems are solved in each inner iteration of Algorithm 7.2.

ASSUMPTION 7.3. *We assume that there is a sequence $\{\varepsilon^k\} \subset X^*$ with $\varepsilon^k \rightharpoonup^* 0$ such that $x^{k+1} \in C$ and $\varepsilon^{k+1} - (L_{\rho^k})_x'(x^{k+1}, w^k) \in \mathcal{N}_C(x^{k+1})$ hold for all $k \in \mathbb{N}$.*

The next lemma verifies the first part of the w-AKKT sequence property stated in Definition 3.1. A proof of this result can be found in [20, Lem. 5.2].

LEMMA 7.4. *Let $\{(x^k, \lambda^k)\} \subset X \times H$ be a sequence generated by Algorithm 7.2. Then there is a null sequence $\{r^k\} \subset [0, +\infty)$ such that $(\lambda^k, y - G(x^k))_H \le r^k$ holds for all $y \in \mathcal{K}$ and $k \in \mathbb{N}$.*

By the last lemma and $e(K) \subset \mathcal{K}$, we also have $\langle \lambda^k, y - G(x^k) \rangle_Y \leq r^k$ for all $y \in K$ and $k \in \mathbb{N}$. This shows that one of the two defining properties in the definition of w-AKKT sequences is satisfied for problem (P) (viewed as an optimization problem with the primal-dual pair (x^k, λ^k) generated in $X \times Y^*$, whereas the optimization problem (P_H) would have to view this sequence as belonging to the space $X \times H$). The second part from the definition of a w-AKKT sequence is a requirement in X^* and therefore identical for (P) and (P_H). A verification of this second condition yields the following result.

THEOREM 7.5. *Suppose that the subproblems in* (7.2) *are solved such that Assumption* 7.3 *holds. Then Algorithm* 7.2 *generates a w-AKKT sequence* $\{(x^k, \lambda^k)\} \subset X \times H$ *of* (P_H) *(and, particularly, due to* $\{(x^k, \lambda^k)\} \subset X \times Y^*$, *of* (P), *too).*

Proof. First, we obtain

$$(L_{\rho^k})'_x(x, w^k) = f'(x) + \rho^k G'(x)^* \left[G(x) + \frac{w^k}{\rho^k} - P_{\mathcal{K}}\left(G(x) + \frac{w^k}{\rho^k} \right) \right].$$

Thus, we deduce from the definition of λ^{k+1} that $(L_{\rho^k})'_x(x^{k+1}, w^k) = L'_x(x^{k+1}, \lambda^{k+1})$. Consequently, Assumption 7.3 yields $\varepsilon^{k+1} - L'_x(x^{k+1}, \lambda^{k+1}) \in \mathcal{N}_C(x^{k+1})$ for all $k \in \mathbb{N}$, and we additionally have the convergence $\varepsilon^k \rightharpoonup^* 0$. Together with Lemma 7.4, the claim follows. □

Note that Algorithm 7.2 is a kind of penalty method and therefore suffers from the drawback that (weak or strong) limit points generated by this method may not be feasible for (P_H). The general convergence theory for this method presented in [13] shows, however, that we usually get (at least) a KKT point of the program

$$\underset{x \in C}{\text{minimize}} \ (d_{\mathcal{K}}^2 \circ G)(x).$$

Hence, these limit points can be interpreted in a suitable way, namely as stationary points of the constraint violation. In the following result, we can therefore concentrate on the situation where we have a feasible weak limit point of $\{x^k\}$ at hand. This is the situation where w-AKKT regularity can be applied in order to obtain the following convergence result which puts the corresponding theorem from [13], where validity of RCQ was assumed, in a different light.

COROLLARY 7.6. *Suppose that the subproblems in* (7.2) *are solved such that Assumption* 7.3 *holds. Assume that* f' *is weak-to-weak* continuous. Suppose that the sequence* $\{x^k\}$ *generated by Algorithm* 7.2 *admits a feasible weak accumulation point* $\bar{x} \in \mathcal{F}$ *that is w-AKKT regular. Then* $\bar{x}$ *is a KKT point of* (P).

Proof. This follows immediately from Theorems 4.7 and 7.5. □

8. Final remarks. In this paper, we have shown that certain types of AKKT regularity may serve as constraint qualifications (CQs) which apply to abstract optimization problems in Banach spaces. We discussed the relation of these new CQs to already existing ones and demonstrated by means of some examples that there are several practically relevant situations where AKKT regularity conditions apply while Robinson's CQ (RCQ) is generally violated. It has been shown that AKKT regularity may also be helpful for the convergence analysis associated with optimization algorithms. In the future, it remains to be seen to what extent these new CQs enrich the theory on optimization in Banach spaces and optimal control. Particularly, more situations have to be identified where AKKT regularity can be applied in a profitable

way. Let us point out that even in the case where X and Y are finite-dimensional, there seems to be lots of potential hidden in AKKT regularity properties since our generalized version from Definition 4.4 applies to the abstract model (P), and the latter covers, e.g., semidefinite or second-order cone programs.

In the finite-dimensional setting, AKKT regularity has been successfully generalized to the setting of *mathematical programs with complementarity constraints* (MPCCs); see [6, 32]. Recently, the notion of MPCCs has been studied in the rather general context of Banach spaces in [24, 26, 36, 37]. Therein, some reasonable problem-tailored notions of stationarity have been defined, and associated CQs have been investigated. However, due to the limited number of CQs addressing (P), only a few problem-tailored MPCC-CQs have been derived in these papers. Motivated by the results from [6, 32], we aim to study applicability of the concept of AKKT regularity (or some problem-tailored counterparts) in the setting of MPCCs in Banach spaces in the future.

One can easily check that MPCCs can be modeled in the form (P) whenever the sets K and C are allowed to be nonconvex. Thus, the question arises of whether one can extend the theory of this paper to this even more general setting. Observing that our analysis is based on sequences, we see that this naturally would lead to the involvement of the limiting normal cone. A brief investigation of the proof of [32, Theorem 3.3] reveals that this should work fine in the finite-dimensional setting. In the infinite-dimensional situation, it is expected that additional assumptions on the data have to be satisfied, e.g., in order to obtain a counterpart of Proposition 3.3 or to apply the calculus rules of limiting variational analysis. This observation is reflected by the results from [28, section 5.1.2], where KKT-type necessary optimality conditions for optimization problems of type (P) with nonconvex sets K and C are derived in the presence of suitable CQs and additional so-called *sequential compactness* assumptions on the data K or C which are quite restrictive in the function space setting; see [25]. However, these assumptions might be interpreted as the price one has to pay for the absence of convexity. Indeed, the theory in this paper does not require any sequential compactness of the data.

REFERENCES

[1] J. M. Abadie, *Problèmes d'optimisation*, Institut Blaise Pascal, Paris, 1965.

[2] R. Andreani, E. G. Birgin, J. M. Martínez, and M. L. Schuverdt, *On augmented Lagrangian methods with general lower-level constraints*, SIAM J. Optim., 18 (2007), pp. 1286–1309, https://doi.org/10.1137/060654797.

[3] R. Andreani, E. G. Birgin, J. M. Martínez, and M. L. Schuverdt, *Augmented Lagrangian methods under the constant positive linear dependence constraint qualification*, Math. Program., 111 (2008), pp. 5–32, https://doi.org/10.1007/s10107-006-0077-1.

[4] R. Andreani, N. S. Fazzio, M. L. Schuverdt, and L. D. Secchin, *A sequential optimality condition related to the quasi-normality constraint qualification and its algorithmic consequences*, SIAM J. Optim., 29 (2019), pp. 743–766, https://doi.org/10.1137/17M1147330.

[5] R. Andreani, G. Haeser, and J. M. Martínez, *On sequential optimality conditions for smooth constrained optimization*, Optimization, 60 (2011), pp. 627–641, https://doi.org/10.1080/02331930903578700.

[6] R. Andreani, G. Haeser, L. D. Secchin, and P. J. S. Silva, *New sequential optimality conditions for mathematical programs with complementarity constraints and algorithmic consequences*, SIAM J. Optim., 29 (2019), pp. 3201–3230, https://doi.org/10.1137/18M121040X.

[7] R. Andreani, J. M. Martínez, A. Ramos, and P. J. S. Silva, *A cone-continuity constraint qualification and algorithmic consequences*, SIAM J. Optim., 26 (2016), pp. 96–110, https://doi.org/10.1137/15M1008488.

[8]   R. ANDREANI, J. M. MARTÍNEZ, AND B. F. SVAITER, *A new sequential optimality condition for constrained optimization and algorithmic consequences*, SIAM J. Optim., 20 (2010), pp. 3533–3554, https://doi.org/10.1137/090777189.

[9]   J.-P. AUBIN AND H. FRANKOWSKA, *Set-Valued Analysis*, Modern Birkhäuser Classics, reprint of the 1990 edition, Birkhäuser Boston, Inc., Boston, MA, 2009, https://doi.org/10.1007/978-0-8176-4848-0,

[10] E. G. BIRGIN AND J. M. MARTÍNEZ, *Practical Augmented Lagrangian Methods for Constrained Optimization*, Fundam. Algorithms 10, SIAM, Philadelphia, 2014, https://doi.org/10.1137/1.9781611973365.

[11] J. BLOT, *Rank theorem in infinite dimension and Lagrange multipliers*, Appl. Math. Optim., 81 (2020), pp. 815–822, https://doi.org/10.1007/s00245-018-9520-y.

[12] J. F. BONNANS AND A. SHAPIRO, *Perturbation Analysis of Optimization Problems*, Springer Ser. Oper. Res., Springer, New York, 2000, https://doi.org/10.1007/978-1-4612-1394-9.

[13] E. BÖRGENS, C. KANZOW, AND D. STECK, *Local and global analysis of multiplier methods for constrained optimization in Banach spaces*, SIAM J. Control Optim., 57 (2019), pp. 3694–3722, https://doi.org/10.1137/19M1240186.

[14] F. H. CLARKE, *Optimization and Nonsmooth Analysis*, Wiley, New York, 1983.

[15] S. V. EMELYANOV, S. K. KOROVIN, N. A. BOBYLEV, AND A. V. BULATOV, *Homotopy of Extremal Problems: Theory and Applications*, De Gruyter Ser. Nonlinear Anal. Appl. 11, Walter de Gruyter & Co., Berlin, 2007, https://doi.org/10.1515/9783110893014.

[16] M. FABIAN, P. HABALA, P. HÁJEK, V. MONTESINOS SANTALUCÍA, J. PELANT, AND V. ZIZLER, *Functional Analysis and Infinite-Dimensional Geometry*, CMS Books Math./Ouvrages Math. SMC 8, Springer, New York, 2001, https://doi.org/10.1007/978-1-4757-3480-5.

[17] F. J. GOULD AND J. W. TOLLE, *Optimality conditions and constraint qualifications in Banach space*, J. Optim. Theory Appl., 15 (1975), pp. 667–684, https://doi.org/10.1007/bf00935506.

[18] M. GUIGNARD, *Generalized Kuhn–Tucker conditions for mathematical programming problems in a Banach space*, SIAM J. Control, 7 (1969), pp. 232–241, https://doi.org/10.1137/0307016.

[19] C. KANZOW AND D. STECK, *An example comparing the standard and safeguarded augmented Lagrangian methods*, Oper. Res. Lett., 45 (2017), pp. 598–603, https://doi.org/10.1016/j.orl.2017.09.005.

[20] C. KANZOW AND D. STECK, *Quasi-variational inequalities in Banach spaces: Theory and augmented Lagrangian methods*, SIAM J. Optim., 29 (2019), pp. 3174–3200, https://doi.org/10.1137/18M1230475.

[21] C. KANZOW, D. STECK, AND D. WACHSMUTH, *An augmented Lagrangian method for optimization problems in Banach spaces*, SIAM J. Control Optim., 56 (2018), pp. 272–291, https://doi.org/10.1137/16M1107103.

[22] T. KATO, *Perturbation Theory for Linear Operators*, Classics Math. 132, reprint of the 1980 edition, Springer-Verlag, Berlin, 1995, https://doi.org/10.1007/978-3-642-66282-9.

[23] S. KURCYUSZ, *On the existence and nonexistence of Lagrange multipliers in Banach spaces*, J. Optim. Theory Appl., 20 (1976), pp. 81–110, https://doi.org/10.1007/BF00933349.

[24] P. MEHLITZ, *Contributions to Complementarity and Bilevel Programming in Banach Spaces*, Ph.D. thesis, Technische Universität Bergakademie Freiberg, 2017, https://tubaf.qucosa.de/api/qucosa%3A23128/attachment/ATT-0/.

[25] P. MEHLITZ, *On the sequential normal compactness condition and its restrictiveness in selected function spaces*, Set-Valued Var. Anal., 27 (2019), pp. 763–782, https://doi.org/10.1007/s11228-018-0475-6.

[26] P. MEHLITZ AND G. WACHSMUTH, *Weak and strong stationarity in generalized bilevel programming and bilevel optimal control*, Optimization, 65 (2016), pp. 907–935, https://doi.org/10.1080/02331934.2015.1122007.

[27] B. S. MORDUKHOVICH, *Variational Analysis and Generalized Differentiation* I*: Basic Theory*, Grundlehren Math. Wiss. 330, Springer, Berlin, 2006, https://doi.org/10.1007/3-540-31247-1.

[28] B. S. MORDUKHOVICH, *Variational Analysis and Generalized Differentiation* II*: Applications*, Grundlehren Math. Wiss. 331, Springer, Berlin, 2006, https://doi.org/10.1007/3-540-31246-3.

[29] K. J. PALMER, *On the complete continuity of differentiable mappings*, J. Austral. Math. Soc., 9 (1969), pp. 441–444, https://doi.org/10.1017/s1446788700007370.

[30] J.-P. PENOT, *On the existence of Lagrange multipliers in nonlinear programming in Banach spaces*, in Optimization and Optimal Control, A. Auslender, W. Oettli, and J. Stoer, eds., Lecture Notes in Control and Inform. Sci. 30, Springer, Berlin, 1981, pp. 89–104, https://doi.org/10.1007/bfb0004508.

[31] R. R. PHELPS, *Convex Functions, Monotone Operators and Differentiability*, 2nd ed., Lecture Notes in Math. 1364, Springer-Verlag, Berlin, 1993, https://doi.org/10.1007/978-3-540-46077-0.

[32] A. RAMOS, *Mathematical programs with equilibrium constraints: A sequential optimality condition, new constraint qualifications and algorithmic consequences*, Optim. Methods Software, 2019, pp. 1–37 (published online Dec. 26, 2019), https://doi.org/10.1080/10556788.2019.1702661.

[33] S. M. ROBINSON, *Stability theory for systems of inequalities, part* II: *Differentiable nonlinear systems*, SIAM J. Numer. Anal., 13 (1976), pp. 497–513, https://doi.org/10.1137/0713043.

[34] D. STECK, *Lagrange Multiplier Methods for Constrained Optimization and Variational Problems in Banach Spaces*, Ph.D. thesis, Institute of Mathematics, Universität Würzburg, 2018, https://opus.bibliothek.uni-wuerzburg.de/opus4-wuerzburg/frontdoor/deliver/index/docId/17444/file/Steck_Daniel_Dissertation.pdf.

[35] F. TRÖLTZSCH, *Optimal Control of Partial Differential Equations: Theory, Methods and Applications* , Grad. Stud. Math. 112, American Mathematical Society, Providence, RI, 2010, https://doi.org/10.1090/gsm/112.

[36] G. WACHSMUTH, *Mathematical programs with complementarity constraints in Banach spaces*, J. Optim. Theory Appl., 166 (2015), pp. 480–507, https://doi.org/10.1007/s10957-014-0695-3.

[37] G. WACHSMUTH, *Strong stationarity for optimization problems with complementarity constraints in absence of polyhedricity: With Applications to Optimization with Semidefinite and Second-Order-Cone Complementarity Constraints*, Set-Valued Var. Anal., 25 (2017), pp. 133–175, https://doi.org/10.1007/s11228-016-0370-y.

[38] Y. XUE, *Stable Perturbations of Operators and Related Topics*, World Scientific, Singapore, 2012, https://doi.org/10.1142/8365.

[39] C. ZĂLINESCU, *A generalization of the Farkas lemma and applications to convex programming*, J. Math. Anal. Appl., 66 (1978), pp. 651–678, https://doi.org/10.1016/0022-247X(78)90260-3.

[40] C. ZHU, J. CAI, AND G.-L. CHEN, *Perturbation analysis for the reduced minimum modulus of bounded linear operator in Banach spaces*, Appl. Math. Comput., 145 (2003), pp. 13–21, https://doi.org/10.1016/s0096-3003(02)00434-4.

[41] J. ZOWE AND S. KURCYUSZ, *Regularity and stability for the mathematical programming problem in Banach spaces*, Appl. Math. Optim., 5 (1979), pp. 49–62, https://doi.org/10.1007/BF01442543.