
Mathematics of Operations Research

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

# Strict Constraint Qualifications and Sequential Optimality Conditions for Constrained Optimization

Roberto Andreani, José Mario Martínez, Alberto Ramos, Paulo J. S. Silva

informs
http://pubsonline.informs.org/journal/moor/

MATHEMATICS OF OPERATIONS RESEARCH
Vol. 43, No. 3, August 2018, pp. 693–717
ISSN 0364-765X (print), ISSN 1526-5471 (online)

# Strict Constraint Qualifications and Sequential Optimality Conditions for Constrained Optimization

**Roberto Andreani,[a] José Mario Martínez,[a] Alberto Ramos,[b] Paulo J. S. Silva[a]**

[a] Institute of Mathematics, Statistics and Scientific Computing, University of Campinas, Campinas, São Paulo 13083-970, Brazil;
[b] Department of Mathematics, Federal University of Paraná, Curitiba, Paraná 80060-000, Brazil

**Contact:** andreani@ime.unicamp.br (RA); martinez@ime.unicamp.br (JMM); albertoramos@ufpr.br (AR); pjssilva@ime.unicamp.br, http://orcid.org/0000-0003-1340-965X (PJSS)

**Received:** November 18, 2015
**Revised:** November 16, 2016; April 11, 2017
**Accepted:** May 14, 2017
**Published Online in Articles in Advance:** January 3, 2018

**MSC2010 Subject Classification:** Primary: 90C30; 49K99; 65K05
**OR/MS Subject Classification:** Primary: nonlinear programming; secondary: optimization

https://doi.org/10.1287/moor.2017.0879



**Abstract.** Sequential optimality conditions for constrained optimization are necessarily satisfied by local minimizers, independently of the fulfillment of constraint qualifications. These conditions support the employment of different stopping criteria for practical optimization algorithms. On the other hand, when an appropriate property on the constraints holds at a point that satisfies a sequential optimality condition, such a point also satisfies the Karush-Kuhn-Tucker conditions. Those properties will be called *strict constraint qualifications* in this paper. As a consequence, for each sequential optimality condition, it is natural to ask for its weakest strict associated constraint qualification. This problem has been solved in a recent paper for the Approximate Karush-Kuhn-Tucker sequential optimality condition. In the present paper, we characterize the weakest strict constraint qualifications associated with other sequential optimality conditions that are useful for defining stopping criteria of algorithms. In addition, we prove all the implications between the new strict constraint qualifications and other (classical or strict) constraint qualifications.

**Funding:** This work was supported by PRONEX-Optimization [PRONEX-CNPq/FAPERJ E-26/111.449/2010-APQ1], CEPID-CeMEAI [FAPESP 2013/07375-0], CNPq [Grants 304618/2013-6, 482549/2013-0, and 303013/2013-3], FAPESP [Grants 2012/20339-0 and 2013/05475-7], CAPES and National Project on Industrial Mathematics.

**Keywords:** nonlinear programming • constraint qualifications • algorithmic convergence

## 1. Introduction

We will consider finite-dimensional constrained optimization problems defined by

$$\text{Minimize } f(x) \quad \text{subject to } h(x) = 0, \;\; g(x) \leq 0, \tag{1}$$

where $f\colon \mathbb{R}^n \to \mathbb{R}$, $h\colon \mathbb{R}^n \to \mathbb{R}^m$, and $g\colon \mathbb{R}^n \to \mathbb{R}^p$ have at least continuous first-order derivatives.

It is well known that optimization problems like (1) cannot be solved efficiently in the general case. Actually, it is not even possible to characterize a global, or local, minimum using only information at a single point. Hence, in many cases, we must settle into searching for feasible points that conform to a condition that is necessary for optimality. Arguably, the most important of such conditions is the Karaush-Kuhn-Tucker condition (KKT) which is only guaranteed to hold under special Constraint Qualifications (CQs) (see Bertsekas [10], Nocedal and Wright [34]).

Algorithms to solve (1) are usually iterative and generate sequences that approximately fulfill a necessary optimality condition. Hopefully such sequences converge to points that satisfy a condition like KKT, that depends only on information at a single point. This leads directly to the definition of a *Sequential Optimality Conditions* (Martínez and Svaiter [31], Andreani et al. [2]). Sequential Optimality Conditions are properties of feasible points of (1) that are necessarily satisfied by any local minimizer $x^*$ and are formulated in terms of sequences that converge to $x^*$. Typically they are based on an inexact version of a point-based condition that gets closer and closer to being satisfied. For example, the most popular sequential optimality condition is AKKT (Approximate Karush-Kuhn-Tucker), which is satisfied by a feasible point $x^*$ if there exist sequences $x^k \to x^*$, $\{\lambda^k\} \subset \mathbb{R}^m$, and $\{\mu^k\} \subset \mathbb{R}^p_+$ such that

$$\lim_{k\to\infty}\left(\nabla f(x^k) + \sum_{i=1}^{m} \lambda_i^k \nabla h_i(x^k) + \sum_{i=1}^{p} \mu_i^k \nabla g_i(x^k)\right) = 0 \tag{2}$$

and

$$\lim_{k\to\infty} \min\{\mu_i^k, -g_i(x^k)\} = 0, \quad \text{for all } i = 1, \ldots, p. \tag{3}$$

This condition is clearly associated to KKT.

Up to our knowledge, Qi and Wei were the first to explicitly talk about approximate KKT conditions while they were analyzing the convergence of a sequential quadratic programming algorithm (Qi and Wei [35]). After this work, different sequential optimality conditions were developed, usually in association with the convergence analysis of computational methods. For example, Scaled AKKT (SAKKT) appeared in Cartis et al. [17, 18] and Wächter and Biegler [37]. The denomination "sequential optimality condition" was first used by Martínez and Svaiter [31]. This work also introduced the Approximate Gradient Projection condition (AGP). A textbook analysis of sequential optimality conditions is given in Birgin and Martínez [12]. The Linear AGP was introduced in Andreani et al. [2], where many properties of sequential optimality conditions and internal relations were elucidated. Complementarity AKKT (CAKKT) was given in Andreani et al. [4], together with the proof that, under a Lojasiewicz inequality, some Augmented Lagrangian methods generate sequences that satisfy it is as Lojasiewicz inequality (see Łojasiewicz [26] and Bolte et al. [14]).

In contrast to KKT, sequential optimality conditions like AKKT are satisfied by any local minimizer independently of the fulfillment of constraint qualifications (see the references cited previously). For instance, the KKT conditions do not hold at the minimizer of $x$ subject to $x^2=0$, but AKKT does. Therefore, it is natural to ask under which conditions it is possible to "pass into the limit" and ensure the validity of KKT at points that fulfill a sequential optimality condition. In particular, we may be interested in what condition the description of the constraints has to fulfill at a point that satisfies a sequential optimality condition to ensure that it also satisfies KKT, independently of the objective function. These conditions will be called *Strict Constraint Qualifications* (Birgin and Martínez [12]). It is important to emphasize that the validity of a strict constraint qualification would ensure the convergence to KKT points whenever an algorithm generates a sequence conforming to the respective sequential optimality condition.

Recall that a constraint qualification is a property of feasible points of the constrained optimization problem that, when satisfied by a local minimizer, implies that such a minimizer satisfies KKT. Since, on the other hand, all local minimizers satisfy sequential optimality conditions, strict constraint qualifications are, in fact, constraint qualifications. The reciprocal is not true. For instance, Abadie's constraint qualification (Abadie [1]) and Quasinormality (Bertsekas [10], Hestenes [24]) are constraint qualifications that are not strict constraint qualifications related with AKKT (see Andreani et al. [7]).

The strength, or strictness, of a sequential optimality condition is associated with the weakness, or generality, of the associated strict constraint qualifications. In fact, every sequential optimality condition implies "KKT or not-SCQ," for some strict constraint qualifications SCQ. Hence a weak, or general, SCQ corresponds to a strong, or strict, sequential optimality condition. In this sense, it is also important to ask which is the weakest strict constraint qualification associated with each sequential optimality condition. For example, in Andreani et al. [7], it has been proved that the weakest strict constraint qualification associated with AKKT is the so-called Cone Continuity Property (CCP) (see Andreani et al. [7, definition 3.1]).

In this paper, we aim to discover the weakest strict constraint qualifications associated with a number of interesting sequential optimality conditions. We also intend to provide geometrical interpretations of the strict constraint qualifications, as in the case of CCP. We hope that this type of research will be useful from a practical point of view because sequential optimality conditions are linked in a natural way to stopping criteria for numerical algorithms. For example, a stopping criterion associated with AKKT may be given by

$$\left\|\nabla f(x^k)+\sum_{i=1}^{m}\lambda_i^k\nabla h_i(x^k)+\sum_{i=1}^{p}\mu_i^k\nabla g_i(x^k)\right\|\le\varepsilon, \tag{4}$$

$$\|h(x^k)\|\le\varepsilon, \qquad \|\max\{0,g(x^k)\}\|\le\varepsilon \qquad \text{and} \qquad |\min\{\mu_i^k,g_i(x^k)\}|\le\varepsilon, \quad \text{for all } i=1,\dots,p, \tag{5}$$

where $x^k$ is the sequence generated by the algorithm under consideration and $\varepsilon$ is an error tolerance.

Many optimization algorithms use different approximations of the KKT conditions, like the AKKT condition presented above, as stopping criterion. This gives rise to an interesting algorithmic question: which is the "best" form for an approximate KKT condition? This paper starts by analyzing an alternative that became popular in the last 10 years, which is called Scaled-AKKT (see Section 2 for definitions and details). It shows that Scaled-AKKT is equivalent to KKT or not-MFCQ, which is an interesting result but not strong enough. Even more, Scaled-AKKT may hold at points that are very distant from minimizers. The strength of an approximate KKT condition, or any sequential optimality condition, is related to the constraint qualification that guarantees that its fulfillment implies first order stationarity. Usually, sequential optimality conditions are equivalent to KKT or not-CQ, for some constraint qualification CQ. Therefore a weak, or less demanding, CQ would be associated to a sequential optimality condition that is able to ensure the validity of first order stationarity in more general, degenerate, cases. The Scaled-AKKT condition is not strong enough because MFCQ is not weak enough. In this paper, we study the weakest constraint qualifications associated with AKKT, AGP, and CAKKT, which are sequential optimality conditions associated with actual numerical algorithms. We also show that these conditions are in (strict) increasing order of strength, being able to tackle different levels of degeneracy in the constraints that may appear in real problems formulations and still ensure convergence to first order stationary points. As a consequence of these results, we present an updated landscape of constraint qualifications, strict constraint qualifications, and sequential optimality conditions.

This paper is organized as follows. In Section 2, we give a motivating example where we address the only sequential optimality condition considered in this paper that is weaker than AKKT. It will be instructive to realize that the corresponding strict constraint qualification will be stronger than the strict constraint qualifications associated with other sequential optimality conditions. In Section 4, we discover the weakest strict constraint qualifications associated with AGP (Approximate Gradient Projection), CAKKT (Complementary AKKT), LAGP (Linear AGP) and SAKKT (Strong-AKKT; Haeser and Schuverdt [23]) sequential optimality conditions. In all these cases, we will stress the geometrical meaning of the strict constraint qualifications so far obtained. Section 4 will be preceded by Section 3, in which we introduce the necessary background for the rest of the paper. In Section 5, we show the relations among the new introduced strict constraint qualifications, whereas in Section 6, we establish the relations with well-known constraint qualifications. Finally, in Section 7, we state some conclusions and lines for future research.

### Notation

We will employ the standard notation of Borwein and Lewis [15], Mordukhovich [33], and Rockafellar and Wets [36]. $\mathbb{N}$ denotes the set of natural numbers, and $\mathbb{R}^n$ stands for the $n$-dimensional real Euclidean space. We denote by $\mathbb{B}$ the closed unit ball in $\mathbb{R}^n$, and by $\mathbb{B}(x,\eta) := x + \eta\mathbb{B}$ the closed ball with center $x$ and radius $\eta > 0$. $\mathbb{R}_+$ is the set of positive scalars, $\mathbb{R}_-$ is the set of negative scalars, and $a^+ = \max\{0, a\}$, the positive part of $a$. We use $\langle \cdot, \cdot \rangle$ to denote the Euclidean inner product, and $\|\cdot\|$ is the associated norm. We use $\|\cdot\|_\infty$ for the supremum norm. Given a set-valued mapping (multifunction) $F\colon \mathbb{R}^s \rightrightarrows \mathbb{R}^d$, the *sequential Painlevé-Kuratowski outer/upper limit* of $F(z)$ as $z \to z^*$ is denoted by

$$\limsup_{z\to z^*} F(z) := \{w^* \in \mathbb{R}^d\colon \exists (z^k, w^k) \to (z^*, w^*) \text{ with } w^k \in F(z^k)\} \tag{6}$$

and the inner limit by

$$\liminf_{z\to z^*} F(z) := \{w^* \in \mathbb{R}^d\colon \forall z^k \to z^* \; \exists w^k \to w^* \text{ with } w^k \in F(z^k)\}. \tag{7}$$

## 2. Example: The Scaled-AKKT Condition

The Scaled-AKKT condition provides a simple example for the type of analysis that will be done in this paper with respect to stronger sequential optimality conditions.

Let us consider feasible sets of the form

$$\{x \in \mathbb{R}^n\colon h(x) = 0, g(x) \le 0\}, \tag{8}$$

where $h\colon \mathbb{R}^n \to \mathbb{R}^m$ and $g\colon \mathbb{R}^n \to \mathbb{R}^p$ admit continuous first derivatives on $\mathbb{R}^n$.

The Scaled-AKKT condition is said to hold at a feasible point $x^*$ of (1), if there exists a sequence $\{x^k\}$ that converges to $x^*$, and sequences $\{\lambda^k\} \subset \mathbb{R}^m$ and $\{\mu^k\} \subset \mathbb{R}^p_+$, such that (3) holds and

$$\lim_{k\to\infty} \max\{1, \|\lambda^k\|_\infty, \|\mu^k\|_\infty\}^{-1} \left\| \nabla f(x^k) + \sum_{i=1}^m \lambda_i^k \nabla h_i(x^k) + \sum_{i=1}^p \mu_i^k \nabla g_i(x^k) \right\| = 0. \tag{9}$$

This property is frequently associated with stopping criteria of modern practical optimization algorithms like IPOPT (Wächter and Biegler [37]) and new algorithms that motivate interesting complexity results (Cartis et al. [17]). Clearly, AKKT implies Scaled-AKKT, so Scaled-AKKT is a sequential optimality condition. We will show that the weakest strict constraint qualification associated with Scaled-AKKT is:

$$\text{MFCQ} \qquad \text{or} \qquad \left[ \left\{ \sum_{i=1}^m \lambda_i \nabla h_i(x^*) + \sum_{g_j(x^*)=0} \mu_j \nabla g_j(x^*)\colon \lambda \in \mathbb{R}^m, \mu_j \in \mathbb{R}^p_+ \right\} = \mathbb{R}^n \right] \tag{10}$$

where MFCQ is the Mangasarian-Fromovitz Constraint Qualification (Bertsekas [10], Mangasarian and Fromovitz [27]).

First, note that (10) is a strict constraint qualification associated with Scaled-AKKT. Indeed, if $\{\sum_{i=1}^m \lambda_i \nabla h_i(x^*) + \sum_{g_j(x^*)=0} \mu_j \nabla g_j(x^*)\colon \lambda \in \mathbb{R}^m, \mu_j \in \mathbb{R}^p_+\} = \mathbb{R}^n$ it turns out that the cone generated by the gradients of active constraints at $x^*$ is the whole space $\mathbb{R}^n$. Then, $x^*$ satisfies KKT independently of the objective function. Suppose now that a feasible point $x^*$ of (1) satisfies the Scaled-AKKT condition (i.e., (3) and (9)) and MFCQ. Then, if the set $\{\lambda^k, \mu^k, k \in \mathbb{N}\}$ is bounded, KKT follows from (9) and (3) taking limits on an appropriate subsequence. If the set $\{\lambda^k, \mu^k, k \in \mathbb{N}\}$ is unbounded, by (9) and (3), we have that

$$\lim_{k\to\infty}\left[\frac{\nabla f(x^k)}{\max\{1, \|\lambda^k\|_\infty, \|\mu^k\|_\infty\}} + \sum_{i=1}^m \tilde{\lambda}_i^k \nabla h_i(x^k) + \sum_{g_i(x^*)=0} \tilde{\mu}_i^k \nabla g_i(x^k)\right] = 0,$$

where the set $\{\tilde{\lambda}^k, \tilde{\mu}^k, k \in \mathbb{N}\}$ is bounded and, for all $k$, we have that $\max\{\|\tilde{\lambda}^k\|_\infty, \|\tilde{\mu}^k\|_\infty\} = 1$. Therefore, taking an appropriate subsequence, we have that there exist $\lambda \in \mathbb{R}^m$ and $\mu \in \mathbb{R}^p_+$ with $\max\{\|\lambda\|_\infty, \|\mu\|_\infty\} = 1$, such that

$$\sum_{i=1}^m \lambda_i \nabla h_i(x^*) + \sum_{g_j(x^*)=0} \mu_j \nabla g_j(x^*) = 0. \tag{11}$$

Therefore, $x^*$ does not satisfy MFCQ. This completes the proof that (10) is a strict constraint qualification associated with Scaled-AKKT.

Let us prove now that (10) is the weakest strict constraint qualification associated with Scaled-AKKT. Assume that $x^*$ satisfies (8) and does not satisfy (10). Then, there exist $\lambda \in \mathbb{R}^m$, $\mu \in \mathbb{R}^p_+$ with $\max\{\|\lambda\|_\infty, \|\mu^k\|_\infty\} = 1$, such that (11) holds. Since $x^*$ does not satisfy (10), there exists a nonnull $c \in \mathbb{R}^n$ such that $c$ is not a linear combination of the gradients $\nabla h_i(x^*)$ and $\nabla g_j(x^*)$ for $j\colon g_j(x^*) = 0$, with nonnegative coefficients corresponding to the inequality gradients. Therefore, $x^*$ is not a KKT point of the problem (1) for $f(x) = \langle x, c\rangle$, $x \in \mathbb{R}^n$. Now take $x^k = x^*$ for all $k \in \mathbb{N}$. By (11), for all $k$ we have:

$$\nabla f(x^k) + \sum_{i=1}^m k\lambda_i \nabla h_i(x^k) + \sum_{g_i(x^*)=0} k\mu_i \nabla g_i(x^k) = \nabla f(x^k) = c.$$

So, since $\max\{\|k\lambda\|_\infty, \|k\mu\|_\infty\} = k$, we have that the Scaled-AKKT condition holds replacing $\lambda^k$ and $\mu^k$ with $k\lambda$ and $k\mu$ respectively.

## 3. Definitions and Basic Results

In this section, we review some basic concepts and results that will be used later on.

We say that $F$ is *outer semicontinuous* (osc) at $z^*$ if

$$\limsup_{z\to z^*} F(z) \subset F(z^*). \tag{12}$$

We say that $F$ is *inner semicontinuous* (isc) at $z^*$ if

$$F(z^*) \subset \liminf_{z\to z^*} F(z). \tag{13}$$

When $F$ is inner semicontinuous and outer semicontinuous at $z^*$, we say that $F$ is *continuous* at $z^*$. Given the set $S$, the symbol $z \to_S z^*$ means that $z \to z^*$ with $z \in S$. For a cone $\mathcal{K} \subset \mathbb{R}^s$, its polar (negative dual) is $\mathcal{K}^\circ = \{v \in \mathbb{R}^s \mid \langle v, k\rangle \le 0 \text{ for all } k \in \mathcal{K}\}$. We use the notation $\phi(t) \le o(t)$ for any function $\phi\colon \mathbb{R}_+ \to \mathbb{R}^s$ such that $\limsup_{t\to 0_+} t^{-1}\phi(t) \le 0$.

Given $S \subset \mathbb{R}^n$ and $z^* \in S$, define the (Bouligand-Severi) *tangent/contingent cone* to $S$ at $z^*$ by

$$T_S(z^*) := \limsup_{t\downarrow 0} \frac{S - z^*}{t} = \{d \in \mathbb{R}^n\colon \exists\, t_k \downarrow 0, d^k \to d \text{ with } z^* + t_k d^k \in S\}. \tag{14}$$

The (Fréchet) *regular normal cone* to $S$ at $z^* \in S$ is defined as

$$\hat{N}_S(z^*) := \{w \in \mathbb{R}^n\colon \langle w, z - z^*\rangle \le o(|z - z^*|) \text{ for } z \in S\}. \tag{15}$$

The (Mordukhovich) *limiting normal cone* to $S$ at $x^* \in S$ is

$$N_S(z^*) := \limsup_{z\to_S z^*} \hat{N}_S(z). \tag{16}$$

For general sets, we have the inclusion $\hat{N}_S(z^*) \subset N_S(z^*)$ for all $z^* \in S$. When $S$ is a convex set, both regular and limiting normal cones reduce to the classical normal cone of convex analysis and then the common notation $N_S(z^*)$ is used. Furthermore, there is a nice relation between the Euclidean projection and the normal cone, as the next proposition shows. Recall that the Euclidean projection onto a closed set $S$, denoted by $P_S$, is defined as, $P_S(z) := \arg\min \operatorname{Inf}\{\|z - s\|: s \in S\}$.

**Proposition 1** (Rockafellar and Wets [36, Proposition 6.17]). *Let $C$ be a nonempty convex closed set and $x \in C$. Then, $\omega \in N_C(x)$ if and only if $P_C(x + \omega) = x$.*

Now, denote by $\Omega$ the feasible set associated with (1), $\Omega := \{x \in \mathbb{R}^n \mid h(x) = 0, g(x) \leq 0\}$. Let $J(x^*)$ be the set of indices of active inequality constraints. Let $x^* \in \Omega$ be a local minimizer of (1). The geometrical first-order necessary optimality condition states that $\langle \nabla f(x^*), d\rangle \geq 0$ for all $d \in T_\Omega(x^*)$ (see for instance Rockafellar and Wets [36], Borwein and Lewis [15], Bertsekas [10]). In other words,

$$-\nabla f(x^*) \in T_\Omega(x^*)^\circ. \tag{17}$$

Associated with the tangent cone, we define the *linearized cone* $L_\Omega(x^*)$ as follows.

$$L_\Omega(x^*) := \{d \in \mathbb{R}^n \mid \langle \nabla h_i(x^*), d\rangle = 0, \forall\, i \in \{1, \dots, m\}, \langle \nabla g_j(x^*), d\rangle \leq 0, \forall\, j \in J(x^*)\}. \tag{18}$$

$L_\Omega(x^*)$ can be considered as the first-order linear approximation of the tangent cone $T_\Omega(x^*)$. If $x^* \in \Omega$ satisfies

$$T_\Omega(x^*)^\circ = L_\Omega(x^*)^\circ, \tag{19}$$

we have that, by the geometric first-order necessary optimality condition (17), the KKT conditions hold at $x^*$. The condition (19) was introduced by Guignard [22]. Gould and Tolle [21] proved that Guignard's condition (19) is the weakest constraint qualification that guarantees that a local minimizer satisfies KKT. Another well-known CQ is the Abadie's constraint qualification, which is stronger than Guignard's CQ and reads $L_\Omega(x^*) = T_\Omega(x^*)$.

Several other constraint qualifications have been proposed in the literature; for instance, we can mention CRCQ (Janin [25]), RCRCQ (Minchenko and Stakhovski [32]), CPLD (Qi and Wei [35], and Andreani et al. [3]), RCPLD (Andreani et al. [6]), Pseudonormality (Bertsekas and Ozdaglar [11]), Quasinormality (Hestenes [24]), and CRSC and CPG (Andreani et al. [8]). Recently, the Cone Continuity Property (CCP) was introduced in Andreani et al. [7], it turns out to be the weakest strict CQ associated with AKKT. CCP states the continuity of the set-valued mapping $x \in \mathbb{R}^n \rightrightarrows K(x)$ at a feasible point $x^*$, where

$$K(x) = \left\{\sum_{i=1}^{m} \lambda_i \nabla h_i(x) + \sum_{j \in J(x^*)} \mu_j \nabla g_j(x): \mu_j \in \mathbb{R}_+, \lambda_i \in \mathbb{R}\right\}. \tag{20}$$

It is worth to note that the outer semi-continuity of $K(x)$ at $x^*$ is sufficient to imply the continuity of $K(x)$ at the same point, since $K(x)$ is always inner semicontinuous at $x^*$.

## 4. Weakest Strict Constraint Qualifications Associated with Sequential Optimality Conditions

The name of the weakest strict constraint qualification associated with AKKT was motivated by its obvious geometrical meaning. However, in the case of other sequential optimality conditions, the geometrical meaning of the respective weakest strict constraint qualification is not so obvious. Therefore, we decided to name them after the corresponding sequential optimality condition. For example, if we apply this rule to the case of AKKT, we have that "AKKT-regularity" becomes an alternative denomination for CCP. If we apply the same convention to Scaled-AKKT, we say that "Scaled-AKKT regularity" holds at $x^*$ iff (10) is valid. The points that satisfy (10) should be called Scaled-AKKT-regular.

### 4.1. Weakest Strict Constraint Qualification Associated with the Approximate Gradient Projection Condition

The AGP optimality condition was introduced by Martínez and Svaiter [31]. Given a scalar $\gamma \in [-\infty, 0]$, we say that a feasible point $x^* \in \Omega$, satisfies AGP($\gamma$) for (1) if there is a sequence $\{x^k\}$ with $x^k \to x^*$ such that

$$P_{\Omega(x^k,\gamma)}(x^k - \nabla f(x^k)) - x^k \to 0, \tag{21}$$

where $P_{\Omega(x^k,\gamma)}$ is the orthogonal projection onto the closed convex set $\Omega(x^k,\gamma)$, defined as

$$\Omega(x^k,\gamma) := \left\{ z\in\mathbb{R}^n\colon \begin{array}{ll} \langle \nabla h_i(x^k), z-x^k\rangle = 0, & \text{for all } i\in\{1,\ldots,m\} \\ \langle \nabla g_j(x^k), z-x^k\rangle \le 0, & \text{if } 0\le g_j(x^k) \\ g_j(x^k)+\langle \nabla g_j(x^k), z-x^k\rangle \le 0, & \text{if } \gamma < g_j(x^k) < 0 \text{ (when } \gamma\neq 0) \end{array} \right\}. \tag{22}$$

It was shown in Martínez and Svaiter [31] that AGP($\gamma$) is independent of the parameter $\gamma$ for $\gamma\in[-\infty,0)$; that is, if AGP($\gamma$) holds for some $\gamma\in[-\infty,0)$, then AGP($\gamma'$) holds for every $\gamma'\in[-\infty,0)$. In this case, we just write AGP instead of AGP($\gamma$). AGP(0) is equivalent to the sequential optimality condition SAKKT (Haeser and Schuverdt [23]).

The set $\Omega(x^k,\gamma)$ can be considered as a linear approximation of

$$\left\{ z\in\mathbb{R}^n\colon \begin{array}{ll} h_i(z) = h_i(x^k), & \text{for all } i\in\{1,\ldots,m\} \\ g_j(z)\le g_j(x^k), & \text{if } 0\le g_j(x^k) \\ g_j(z)\le 0, & \text{if } \gamma < g_j(x^k) < 0 \text{ (when } \gamma\neq 0) \end{array} \right\}. \tag{23}$$

One of the attractiveness of AGP is that it does not involve Lagrange multipliers estimates. AGP is the natural optimality condition that fits stopping criteria for algorithms based on inexact restoration (Martínez [28], Martínez and Pilotta [30], Fischer and Friedlander [20], Bueno et al. [16], Birgin et al. [13]), and is strictly stronger than the usual AKKT condition. Consequently, the stopping criteria based on AGP are more reliable that those based on AKKT.

Note that the natural stopping criterion associated with AGP is:

$$\|h(x)\|\le \varepsilon_{\text{feas}}, \qquad \|\max\{0,g(x)\}\|\le\varepsilon_{\text{feas}} \qquad \text{and} \qquad \|P_{\Omega(x,\gamma)}(x-\nabla f(x))-x\|\le \varepsilon_{\text{opt}}, \tag{24}$$

where $\varepsilon_{\text{feas}}$ and $\varepsilon_{\text{opt}}$ are user-given tolerances.

AGP-regularity constraint qualification is defined below.

**Definition 1.** We say that AGP-regularity holds at the feasible point $x^*$ if the set-valued mapping

$$(x,\varepsilon)\in\mathbb{R}^n\times\mathbb{R}^n \rightrightarrows N_{\Omega(x,-\infty)}(x+\varepsilon) \tag{25}$$

is outer semicontinuous at $(x^*,0)$; that is,

$$\limsup_{(x,\varepsilon)\to(x^*,0)} N_{\Omega(x,-\infty)}(x+\varepsilon)\subset N_{\Omega(x^*,-\infty)}(x^*) = L_\Omega(x^*)^\circ. \tag{26}$$

Since the set $\Omega(x,-\infty)$ is defined by linear inequality and equality constraints, the normal cone $N_{\Omega(x,-\infty)}(x+\varepsilon)$ admits the geometrical interpretation given by the following proposition.

**Proposition 2.** *Every element of $N_{\Omega(x,-\infty)}(x+\varepsilon)$ has the form*

$$\sum_{i=1}^m \lambda_i\nabla h_i(x) + \sum_{j:g_j(x)\ge 0}\mu_{1j}\nabla g_j(x) + \sum_{j:g_j(x)<0}\mu_{2j}\nabla g_j(x),$$

*where $\lambda_i\in\mathbb{R}$, $\mu_{1j}\in\mathbb{R}_+$, $\mu_{2j}\in\mathbb{R}_+$,*

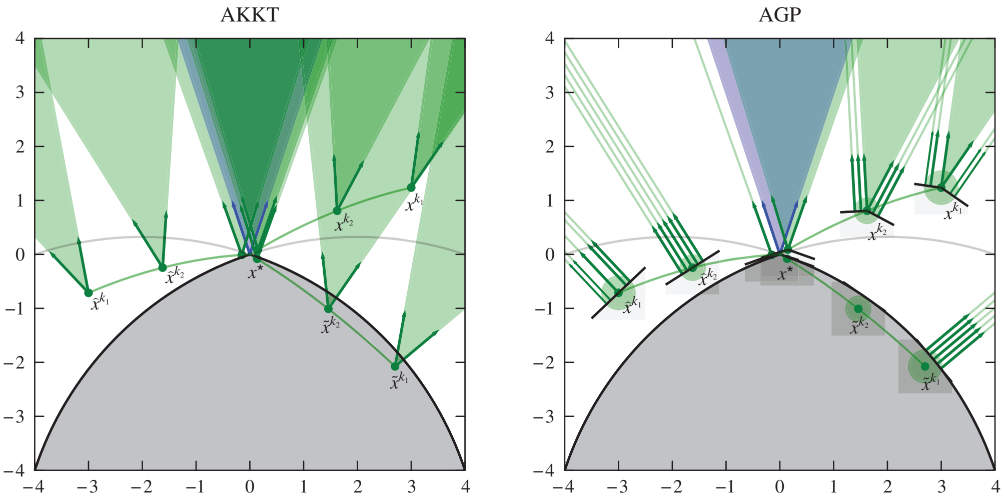
$$\mu_{1j}(\langle\nabla g_j(x),\varepsilon\rangle)=0, \quad \text{if } g_j(x)\ge 0, \qquad \text{and} \qquad \mu_{2j}(g_j(x)+\langle\nabla g_j(x),\varepsilon\rangle)=0, \quad \text{if } g_j(x)<0.$$

By the polarity theorem in Aubin and Frankowska [9, theorem 1.1.8], the outer semicontinuity of $N_{\Omega(x,-\infty)}(x+\varepsilon)$ at $(x,\varepsilon)=(x^*,0)$ is equivalent to the inner semicontinuity at $(x,\varepsilon)=(x^*,0)$ of $L_{\Omega(x,-\infty)}(x+\varepsilon)$, the tangent cone of $\Omega(x,-\infty)$ at $x+\varepsilon$. That is, for each $d\in L_\Omega(x^*)$, and for arbitrary sequences $x^k$ and $\varepsilon^k$ with $x^k\to x^*$ and $\varepsilon^k\to 0$, there exists a sequence $d^k\in L_{\Omega(x^k,-\infty)}(x^k+\varepsilon^k)$ such that $d^k\to d$. Figure 1 shows an example where AGP-regularity holds.

The next Theorem 1 shows that the outer semicontinuity of $N_{\Omega(x,-\infty)}(x+\varepsilon)$ at $(x^*,0)$ is the minimal condition to guarantee that AGP implies KKT for every objective function. Thus, AGP-regularity is the weakest strict constraint qualification associated with AGP.

**Theorem 1.** *AGP-regularity is the weakest strict constraint qualification associated with AGP.*

**Figure 1.** (Color online) Example of the cone mappings associated to the AKKT and AGP conditions.




*Notes.* The shaded area is the feasible set composed of the intersection of two circles. The point of interest is $x^* = 0$. There are three sample sequences converging to $x^*$. The sequences $\{x^k\}$, $\{\hat{x}^k\}$, and $\{\tilde{x}^k\}$ are infeasible with respect to both constraints, infeasible with respect to only one constraint, and strictly feasible respectively. AKKT-regularity and AGP-regularity basically state that the possible limits of the vectors from the respective green cones must belong to the blue cone which is the normal of the linearized cone at $x^*$. Note that the cones associated to AKKT always take into account all the active constraints at $x^*$, while the cones associated with AGP only take into account the constraints that are binding or violated. It is also interesting to observe the effect of the possible perturbations $\{\varepsilon^k\}$ allowed in AGP. Their possible values are represented by the shaded circles in the AGP figure. They allow us to take into account the gradients of constraints that will be biding at $x^*$, but for which the sequence is strictly feasible. See, for example, the point $\tilde{x}^{k_1}$ in the figure.

**Proof.** Let us show first that, under AGP-regularity, AGP implies the KKT condition for any objective function. Let $f$ be an objective function for which AGP($\gamma$) holds at $x^*$ for some $\gamma \in [-\infty, 0)$. Thus, there is a sequence $\{x^k\} \in \mathbb{R}^n$ such that $x^k \to x^*$ and $P_{\Omega(x^k,\gamma)}(x^k - \nabla f(x^k)) - x^k \to 0$. Define $y^k := P_{\Omega(x^k,\gamma)}(x^k - \nabla f(x^k))$ and $\varepsilon^k := y^k - x^k$. Thus, $y^k = x^k + \varepsilon^k$. Clearly, $\lim_{k\to\infty} \varepsilon^k = 0$.

By Proposition 1,

$$\omega^k := x^k - \nabla f(x^k) - y^k \in N_{\Omega(x^k,\gamma)}(y^k). \tag{27}$$

Since the inclusion $N_{\Omega(x^k,\gamma)}(y^k) \subset N_{\Omega(x^k,-\infty)}(y^k)$ always holds, we have that

$$\omega^k \in N_{\Omega(x^k,-\infty)}(x^k + \varepsilon^k) \qquad \text{and} \qquad \omega^k = x^k - \nabla f(x^k) - y^k = -\nabla f(x^k) - \varepsilon^k. \tag{28}$$

Taking limit in the last expression and using the continuity of the gradient of $f$, we get

$$-\nabla f(x^*) = \lim_{k\to\infty} \omega^k \in \limsup_{(x,\varepsilon)\to(x^*,0)} N_{\Omega(x,-\infty)}(x+\varepsilon) \subset N_{\Omega(x^*,-\infty)}(x^*). \tag{29}$$

Thus, $-\nabla f(x^*)$ belongs to $N_{\Omega(x^*,-\infty)}(x^*) = L_\Omega(x^*)^\circ$; that is, the KKT condition holds at $x^*$.

Now, let us prove that, if AGP implies the KKT condition for every objective function, then AGP regularity holds. Take $\omega^* \in \limsup_{(x,\varepsilon)\to(x^*,0)} N_{\Omega(x,-\infty)}(x+\varepsilon)$. Then by the definition of outer limit, there are sequences $\{x^k\}$, $\{\omega^k\}$ and $\{\varepsilon^k\}$ such that $x^k \to x^*$, $\varepsilon^k \to 0$, $\omega^k \to \omega^*$ and $\omega^k \in N_{\Omega(x^k,-\infty)}(x^k + \varepsilon^k)$. Define the objective function, $f(x) := -\langle w^*, x\rangle$ for all $x \in \mathbb{R}^n$. We will show that AGP($-\infty$) holds at $x^*$ for this choice of $f$. So, it is sufficient to show that $\lim_{k\to\infty} P_{\Omega(x^k,-\infty)}(x^k - \nabla f(x^k)) - x^k = 0$.

Define $y^k := x^k + \varepsilon^k$ and $z^k := P_{\Omega(x^k,-\infty)}(x^k - \nabla f(x^k)) = P_{\Omega(x^k,-\infty)}(x^k + \omega^*)$. Since $\omega^k$ is in $N_{\Omega(x^k,-\infty)}(y^k)$ we have $P_{\Omega(x^k,-\infty)}(\omega^k + y^k) = y^k$ (Proposition 1). Using the triangle inequality and the nonexpansivity of the Euclidean projection, we get

$$\|z^k - y^k\| = \|P_{\Omega(x^k,-\infty)}(x^k + \omega^*) - P_{\Omega(x^k,-\infty)}(\omega^k + y^k)\| \le \|\omega^* - \omega^k\| + \|y^k - x^k\|. \tag{30}$$

Taking limits in (30), we obtain $\lim_{k\to\infty} z^k - y^k = 0$, and as consequence

$$\lim_{k\to\infty} P_{\Omega(x^k,-\infty)}(x^k - \nabla f(x^k)) - x^k = \lim_{k\to\infty} z^k - x^k = \lim_{k\to\infty}(z^k - y^k) + \lim_{k\to\infty}(y^k - x^k) = 0. \tag{31}$$

Thus, AGP holds at $x^*$ and, by hypothesis, the KKT condition also holds at $x^*$; that is, $-\nabla f(x^*) = \omega^*$ belongs to $N_{\Omega(x^*,-\infty)}(x^*) = L_\Omega(x^*)^\circ$. This amounts to AGP-regularity at $x^*$ and the statement has been proved. □

### 4.2. Weakest Strict Constraint Qualification Associated with the Complementary AKKT Condition

A feasible point $x^*$ satisfies the Complementary AKKT condition (CAKKT) introduced in Andreani et al. [4] if there exist sequences $\{x^k\} \subset \mathbb{R}^n$, $\{\lambda^k\} \subset \mathbb{R}^m$, and $\{\mu^k\} \subset \mathbb{R}^p_+$, such that $x^k \to x^*$,

$$\lim_{k\to\infty} \nabla f(x^k) + \sum_{i=1}^m \lambda_i^k \nabla h_i(x^k) + \sum_{j=1}^p \mu_j^k \nabla g_j(x^k) = 0, \tag{32}$$

and

$$\lim_{k\to\infty} \sum_{i=1}^m |\lambda_i^k h_i(x^k)| + \sum_{j=1}^p |\mu_j^k g_j(x^k)| = 0. \tag{33}$$

The difference between CAKKT and AKKT is that in AKKT we require $\min\{-g_i(x^k), \mu^k\} \to 0$ for all $i = 1, \ldots, p$ instead of (33). It has been proven in Andreani et al. [4] that CAKKT is a genuine optimality condition satisfied by every local minimizer, it is strictly stronger than AKKT, and it is satisfied by every feasible limit point generated by the Augmented Lagrangian method described in Andreani et al. [5] under a weak Lojasiewicz-like assumption on the constraints.

An example in which CAKKT does not hold, but both AKKT and AGP hold at a nonoptimal point consists of minimizing $\frac{1}{2}(x_2 - 2)^2$, subject to $x_1 = 0$ and $x_1 x_2 = 0$. Clearly, $(0, 2)$ is the unique minimizer. However, every feasible point $(0, \varepsilon)$, for $\varepsilon \geq 0$, satisfies AKKT and AGP, although those points do not satisfy CAKKT. This means that algorithms which guaranteed convergence to, say, AGP points could converge to nonoptimal feasible limits; whereas algorithms with guaranteed convergence to CAKKT points could not.

The formulation (32)–(33) of CAKKT is useful because, after adding a feasibility condition like

$$\|h(x)\| \leq \varepsilon_{\text{feas}} \qquad \text{and} \qquad \|\max\{0, g(x)\}\| \leq \varepsilon_{\text{feas}},$$

it induces a natural stopping criteria to be employed in numerical methods. However, the following equivalent formulation is more adequate for mathematical proofs. We will say that a feasible point $x^*$ satisfies the CAKKT condition for the problem (1), if there exist sequences $\{x^k\} \subset \mathbb{R}^n$, $\{\lambda^k\} \subset \mathbb{R}^m$, and $\{\mu^k\} \subset \mathbb{R}^p_+$, with $\mu_j^k = 0$ for all $j \notin J(x^*)$, such that $x^k \to x^*$,

$$\lim_{k\to\infty} \nabla f(x^k) + \sum_{i=1}^m \lambda_i^k \nabla h_i(x^k) + \sum_{j\in J(x^*)} \mu_j^k \nabla g_j(x^k) = 0, \tag{34}$$

and

$$\lim_{k\to\infty} \sum_{i=1}^m |\lambda_i^k h_i(x^k)| + \sum_{j\in J(x^*)} |\mu_j^k g_j(x^k)| = 0. \tag{35}$$

In fact, (34)–(35) imply (32)–(33) defining $\mu_j^k = 0$, $j \notin J(x^*)$. The other implication follows from the fact that (33) implies that $\mu_j^k \to 0$, $j \notin J(x^*)$. Then, using the assumption that $g$ is continuously differentiable, we may conclude that

$$\lim_{k\to\infty} \sum_{j\notin J(x^*)} \mu_j^k \nabla g_j(x^k) = 0.$$

Finally, (34) follows from (32) after a simple rearrangement of terms.

For all $x \in \mathbb{R}^n$ and $r \in \mathbb{R}_+$, we define $K_C(x, r)$ by:

$$K_C(x, r) := \left\{ \sum_{i=1}^m \lambda_i \nabla h_i(x) + \sum_{j\in J(x^*)} \mu_j \nabla g_j(x) \colon \ \sum_{i=1}^m |\lambda_i h_i(x)| + \sum_{j\in J(x^*)} |\mu_j g_j(x)| \leq r, \lambda_i \in \mathbb{R}, \mu_j \geq 0 \right\}. \tag{36}$$

The set $K_C(x, r)$ is nonempty and convex, with the property $\alpha K_C(x, r) = K_C(x, \alpha r)$ for all $\alpha > 0$. Moreover, $K_C(x, \infty) = K(x)$ for all $x \in \mathbb{R}^n$ and $K_C(x, r)$ coincides with $L_\Omega(x^*)^\circ$ at $(x, r) = (x^*, 0)$, where $K(x)$ is defined by (20) and $L_\Omega(x^*)$ is defined by (18).

We can interpret $K_C(x, r)$ as a perturbation of the linearized normal cone $L_\Omega(x^*)^\circ$ around $x^*$ with the additional constraint $\sum_{i=1}^m |\lambda_i h_i(x)| + \sum_{j\in J(x^*)} |\mu_j g_j(x)| \leq r$ controlling how the complementarity condition is approximated for points $x$ approaching $x^*$.

**Definition 2.** We say that CAKKT-regularity holds at the feasible point $x^*$ if the set-valued mapping

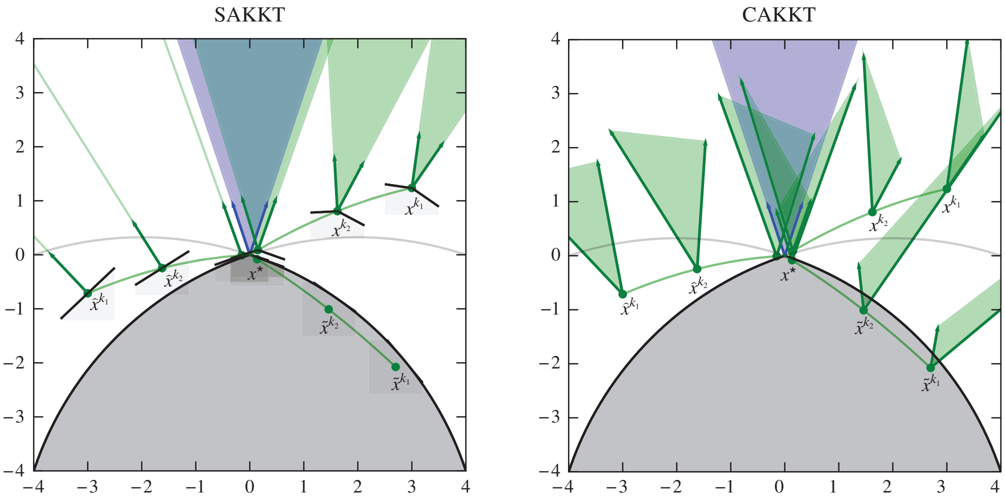
$$(x, r) \in \mathbb{R}^n \times \mathbb{R}_+ \rightrightarrows K_C(x, r)$$

is outer semicontinuous at $(x^*, 0)$; in other words, the following inclusion holds:

$$\limsup_{(x,r)\to(x^*,0)} K_C(x, r) \subset K_C(x^*, 0) = L_\Omega(x^*)^\circ. \tag{37}$$

For a graphical example, see Figure 2.

**Figure 2.** (Color online) Example of the cone mappings associated to the SKKT and CAKKT conditions using the same feasible set, $x^*$, and approximating sequences as Figure 1.




*Notes.* Once again, SKKT-regularity and CAKKT-regularity basically state that the possible limits of the vectors of the respective green cones must belong to the blue cone, which is the normal of the linearized cone at $x^*$. Note that the cones associated with SKKT always take into account only the constraints that are biding or violated and there is no perturbation $\epsilon^k$. See Figure 1 and compare. This is its main difference with respect to AGP. Moreover, the set associated with CAKKT is the cone associated with AKKT with an extra constraint that limits the size of the vectors depending on how close the respective constraint is to zero, and how large the parameter $r^k$ is. Here, $\{r^k\}$ was taken to converge 0 at a speed proportional to the speed with which the sequences approach $x^*$.

**Theorem 2.** *A feasible point $x^*$ is CAKKT-regular if and only if for every continuously differentiable objective function such that CAKKT holds at $x^*$, we have that KKT also holds.* (*That is, CAKKT-regularity is the weakest strict constraint qualification associated with CAKKT.*)

**Proof.** We start by proving that, under CAKKT-regularity, CAKKT implies KKT. Let $f$ be a smooth objective function such that CAKKT holds at $x^*$. Then, by definition, there exist sequences $\{x^k\} \subset \mathbb{R}^n$, $\{\lambda^k\} \subset \mathbb{R}^m$, $\{\mu^k\} \subset \mathbb{R}^p_+$ with $\mu^k_j = 0$ for all $j \notin J(x^*)$, $\{\zeta^k\} \subset \mathbb{R}^m$ and $\{r^k\} \subset \mathbb{R}_+$ such that $\lim_{k\to\infty} x^k = x^*$, $\zeta^k := \nabla f(x^k) + \sum_{i=1}^m \lambda_i^k \nabla h_i(x^k) + \sum_{j\in J(x^*)} \mu_j^k \nabla g_j(x^k) \to 0$ and $r^k := \sum_{i=1}^m |\lambda_i^k h_i(x^k)| + \sum_{j\in J(x^*)} |\mu_j^k g_j(x^k)| \to 0$. Define $\omega^k := \sum_{i=1}^m \lambda_i^k \nabla h_i(x^k) + \sum_{j\in J(x^*)} \mu_j^k \nabla g_j(x^k)$. Clearly, the sequence $\{\omega^k\}$ satisfies

$$\omega^k \in K_C(x^k, r^k) \qquad \text{and} \qquad \omega^k = \zeta^k - \nabla f(x^k). \tag{38}$$

Since $\zeta^k \to 0$ and $\nabla f(x^k) \to \nabla f(x^*)$ we get $\omega^k \to -\nabla f(x^*)$. From the definition of outer limit

$$-\nabla f(x^*) = \lim_{k\to\infty} \omega^k \in \limsup_{(x,r)\to(x^*,0)} K_C(x,r) \subset K_C(x^*,0) = L_\Omega(x^*)^\circ, \tag{39}$$

which implies that the KKT condition holds.

Now, we prove the implication in the other direction, that is if CAKKT implies KKT for any objective function, then CAKKT-regularity holds. Thus, our aim is to prove the inclusion $\limsup_{(x,r)\to(x^*,0)} K_C(x,r) \subset L_\Omega(x^*)^\circ$. We start by taking $\omega^* \in \limsup_{(x,r)\to(x^*,0)} K_C(x,r)$, so there are sequences $\{x^k\}$, $\{\omega^k\}$ and $\{r^k\}$, such that $x^k \to x^*$, $\omega^k \to \omega^*$, $r^k \to 0$ and $\omega^k \in K_C(x^k, r^k)$. Now, define the linear function $f(x) := -\langle w^*, x\rangle$ for all $x \in \mathbb{R}^n$. Let us see that CAKKT holds at $x^*$ with this choice of $f$. Since $\omega^k$ is in $K_C(x^k, r^k)$, there are multipliers $\{\lambda^k\} \subset \mathbb{R}^m$, $\{\mu^k\} \subset \mathbb{R}^p_+$ with $\mu^k_j = 0$ for $j \notin J(x^*)$, such that s

$$\omega^k = \sum_{i=1}^m \lambda_i^k \nabla h_i(x^k) + \sum_{j\in J(x^*)} \mu_j^k \nabla g_j(x^k) \tag{40}$$

and

$$\sum_{i=1}^m |\lambda_i^k h_i(x^k)| + \sum_{j\in J(x^*)} |\mu_j^k g_j(x^k)| \le r^k. \tag{41}$$

Since $r^k \to 0$, the expression (35) holds and from $\omega^k \to \omega^*$, $\zeta^k := \nabla f(x^k) + \omega^k = -\omega^* + \omega^k \to 0$. Thus, CAKKT holds and, due to the hypothesis, $-\nabla f(x^*) = \omega^* \in L_\Omega(x^*)^\circ = K_C(x^*, 0)$. This is sufficient for CAKKT-regularity at $x^*$ and the statement has been proved. □

### 4.3. Weakest Strict Constraint Qualification Associated with the Strong Approximate KKT Condition

We say that a feasible point $x^*$ satisfies the Strong Approximate KKT condition SAKKT if there exist sequences $x^k \to x$, $\{\lambda^k\} \subset \mathbb{R}^m$ and $\{\mu^k\} \subset \mathbb{R}^p_+$, such that (2) holds and $\mu^k_j = 0$ whenever $g_j(x^k) < 0$ (Haeser and Schuverdt [23]). Obviously, this implies (3). SAKKT strictly implies AKKT.

In spite of its strength, SAKKT does not generate practical stopping criteria for constrained optimization algorithms, because reasonable optimization algorithms may generate natural sequences for which the fulfillment of SAKKT cannot be detected. Consider, for example, the problem of minimizing $x$ subject to $-x \le 0$. A reasonable sequence generated by (say) an interior point algorithm (e.g., Chen and Goldfarb [19]) could be $x^k = 1/k$ (or any other positive sequence such that $x^k \to 0$). However, for this sequence, we have that $\nabla f(x^k) = 1$ and $g(x^k) < 0$ for all $k$. Therefore, the condition "$\mu^k < 0$ when $g(x^k) < 0$" imposes that $\mu^k = 0$ for all $k$. This means that this sequence cannot be used to detect SAKKT. In spite of this, SAKKT holds because any negative sequence that tends to zero (in particular, the constant sequence $x^k \equiv 0$) does detect SAKKT.

However, it is interesting to analyze the strict constraint qualifications under which points that satisfy SAKKT also fulfill KKT.

**Definition 3.** Let $x^*$ be a feasible point. We say that SAKKT-regularity holds at $x^*$ if the multifunction $x \in \mathbb{R}^n \rightrightarrows N_{\Omega(x,0)}(x)$ is outer semicontinuous at $x^*$; that is,

$$\limsup_{x \to x^*} N_{\Omega(x,0)}(x) \subset N_{\Omega(x^*,0)}(x^*) = L_\Omega(x^*)^\circ. \tag{42}$$

**Proposition 3.** *Let $x$ and $\varepsilon$ be elements in $\mathbb{R}^m$, such that $x + \varepsilon$ belongs to $\Omega(x, 0)$. Then, every element of $N_{\Omega(x,0)}(x + \varepsilon)$ can be written as*

$$\sum_{i=1}^{m} \lambda_i \nabla h_i(x) + \sum_{j:\, g_j(x) \ge 0} \mu_j \nabla g_j(x),$$

*where $\lambda_i \in \mathbb{R}$, $\mu_j \in \mathbb{R}_+$ for all $i, j$ and $\mu_j(\langle \nabla g_j(x), \varepsilon \rangle) = 0$, if $g_j(x) \ge 0$.*

*Also, $N_{\Omega(x,0)}(x + \varepsilon)$ is a subset of $N_{\Omega(x,0)}(x)$.*

By Proposition 3, we can rephrase SAKKT-regularity saying that it is equivalent to the outer semicontinuity of the set-valued mapping that associates to each point $x$, the linearized normal cone defined by the gradients of the equality constraints and the gradients of inequality constraints whenever the point $x$ is not in the interior of the zero-lower set defined by the corresponding inequality constraint. See an example in Figure 2.

**Theorem 3.** *Let $x^*$ be a feasible point. Then, SAKKT-regularity holds at $x^*$ if and only if for every smooth objective function such that the SAKKT condition holds at $x^*$, the KKT condition also holds at $x^*$.*

**Proof.** First, let us show that if SAKKT-regularity holds, SAKKT implies KKT independently of the objective function. Let $f$ be a function such that SAKKT holds, by the equivalence between AGP(0) and SAKKT (Haeser and Schuverdt [23, theorem 1.2.6(c)]), there is a sequence $\{x^k\} \subset \mathbb{R}^n$ such that $x^k \to x^*$ and $P_{\Omega(x^k,0)}(x^k - \nabla f(x^k)) - x^k \to 0$. Define $y^k := P_{\Omega(x^k,0)}(x^k - \nabla f(x^k))$ and $\varepsilon^k := y^k - x^k$. By Proposition 1, we have

$$\omega^k = x^k - \nabla f(x^k) - y^k \in N_{\Omega(x^k,0)}(y^k = x^k + \varepsilon^k) \subset N_{\Omega(x^k,0)}(x^k), \tag{43}$$

where the last inclusion comes from Proposition (3). Thus, the sequence $\{\omega^k\}$ satisfies

$$\omega^k \in N_{\Omega(x^k,0)}(x^k) \qquad \text{and} \qquad \omega^k = x^k - \nabla f(x^k) - y^k \to -\nabla f(x^*). \tag{44}$$

Thus, by definition of outer limit and outer semicontinuity, we can conclude

$$-\nabla f(x^*) \in \limsup_{x \to x^*} N_{\Omega(x,0)}(x) \subset N_{\Omega(x^*,0)}(x^*) = L_\Omega(x^*)^\circ, \tag{45}$$

proving that $x^*$ satisfies the KKT condition.

Now, we will prove if for any objective function, SAKKT implies KKT, then SAKKT-regularity holds at $x^*$. Take $\omega^* \in \limsup N_{\Omega(x,0)}(x)$, so, there are sequences $\{x^k\}$ and $\{\omega^k\}$ such that $x^k \to x^*$, $\omega^k \to \omega^*$ e $\omega^k \in N_{\Omega(x^k,0)}(x^k)$. Define $f(x) := -\langle w^*, x\rangle$ for all $x \in \mathbb{R}^n$. We will show that AGP(0) holds at $x^*$ for $f(x) = -\langle w^*, x\rangle$. Denote $z^k := P_{\Omega(x^k,0)}(x^k - \nabla f(x^k)) = P_{\Omega(x^k,0)}(x^k + \omega^*)$. From the nonexpansivity of the projection $P_{\Omega(x^k,0)}(x)$ and from $P_{\Omega(x^k,0)}(\omega^k + x^k) = x^k$, we have

$$\|z^k - x^k\| = \|P_{\Omega(x^k,0)}(x^k + \omega^*) - P_{\Omega(x^k,0)}(\omega^k + x^k)\| \le \|\omega^* - \omega^k\|. \tag{46}$$

The last inequality implies that $z^k - x^k \to 0$, and as consequence, AGP(0) (or equivalent SAKKT) holds at $x^*$. Thus, KKT holds at $x^*$; that is, $-\nabla f(x^*) = \omega^* \in N_{\Omega(x^*,0)}(x^*) = L_\Omega(x^*)^\circ$. This amounts to SAKKT-regularity at $x^*$ and the assertion has been proved. □

### 4.4. Weakest Strict Constraint Qualification Associated with the Linear Approximate Gradient Projection Condition

When the optimization problem (1) has linear constraints, a variation of AGP, called Linear Approximate Gradient Projection (LAGP) condition has been introduced Andreani et al. [2]. Denote by $\Omega_L$ the set defined by all the linear constraints and define $\Omega_{NL}(x^k,-\infty)$ as follows:

$$\Omega_{NL}(x^k,-\infty) := \left\{ z \in \mathbb{R}^n \colon \begin{array}{ll} \langle \nabla h_i(x^k), z - x^k\rangle = 0, & \text{for all } i \in I_1 \\ \langle \nabla g_j(x^k), z - x^k\rangle \le 0, & \text{if } 0 \le g_j(x^k),\ \ j \in J_1 \\ g_j(x^k) + \langle \nabla g_j(x^k), z - x^k\rangle \le 0, & \text{if } g_j(x^k) < 0,\ \ j \in J_1 \end{array} \right\}, \tag{47}$$

where the nonlinear constraints of (1) are defined by $\{h_i, i \in I_1\}$ and $\{g_j, j \in J_1\}$. Thus, we say that a feasible point $x^*$ satisfies the LAGP condition for the problem (1) if there is a convergent sequence $\{x^k\} \subset \Omega_L$, with limit $x^*$, such that

$$P_{\Omega_{NL}(x^k,-\infty)\cap\Omega_L}(x^k - \nabla f(x^k)) - x^k \to 0. \tag{48}$$

In Andreani et al. [2], it was shown that LAGP is stronger than AGP (and as a consequence, stronger than AKKT). Now, we will introduce the weakest strict constraint qualification associated with LAGP.

**Definition 4.** If the set-valued mapping $(x,\varepsilon) \in \mathbb{R}^n \times \mathbb{R}^n \rightrightarrows N_{\Omega_{NL}(x,-\infty)\cap\Omega_L}(x+\varepsilon)$ is outer semicontinuous relatively to $\Omega_L \times \mathbb{R}^m$ at $(x^*,0)$; that is,

$$\limsup_{(x,\varepsilon)\to(x^*,0),\, x\in\Omega_L} N_{\Omega_{NL}(x,-\infty)\cap\Omega_L}(x+\varepsilon) \subset N_{\Omega_{NL}(x^*,-\infty)\cap\Omega_L}(x^*) = L_\Omega(x^*)^\circ.$$

we say that LAGP-regularity holds at $x^* \in \Omega$.

Following the arguments of Theorem 1, we obtain

**Theorem 4.** *LAGP-regularity is the weakest strict constraint qualification associated with LAGP.*

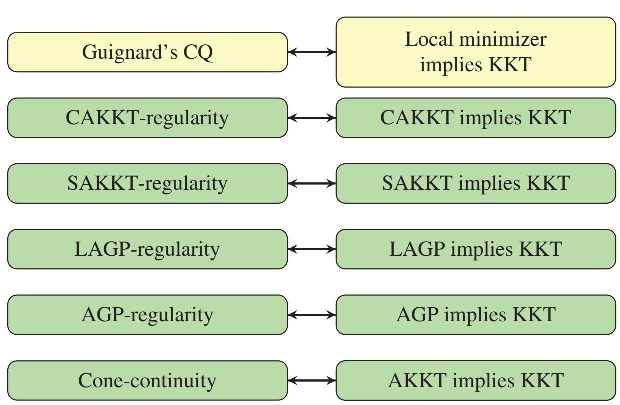
## 5. Relations Between the New Strict Constraint Qualifications

The results of Section 4, together with the equivalence result proved in Andreani et al. [7], are condensed in Figure 3, where, for completeness, we also included the equivalence between Guignard and "Local optimizer implies KKT." Moreover, by the results proved in Andreani et al. [2, 4], Haeser and Schuverdt [23] we have the following theorem.

**Theorem 5.** *The following implications hold*:

1. *CCP implies AGP-regularity*;
2. *AGP-regularity implies SAKKT-regularity*;
3. *AGP-regularity implies LAGP-regularity*;
4. *SAKKT-regularity implies CAKKT-regularity.*

**Figure 3.** (Color online) Equivalence results concerning constraint qualifications.



**Proof.** The four parts of the thesis are proved in the same way. We give one example. In Section 4 we proved that SAKKT-regularity is equivalent to "SAKKT implies KKT." In other words, SAKKT-regularity is equivalent to "SAKKT or not-KKT." Similarly, we proved that CAKKT-regularity is equivalent to "CAKKT or not-KKT." But in Haeser and Schuverdt [23], it has been proved that SAKKT implies CAKKT. Therefore, SAKKT-regularity implies CAKKT-regularity. Now, to see that AGP implies AKKT, see Andreani et al. [2]; for SAKKT implies AGP, see Haeser and Schuverdt [23]; and for LAGP implies AGP, see Andreani et al. [2]. □

The rest of this section is devoted to show that all the implications in Theorem 5 are strict.

**Example 1** (AGP-Regularity is Strictly Weaker Than the Cone Continuous Property)**.** Consider the two-dimensional Euclidean space $\mathbb{R}^2$, the point $x^* = (0,0)$ and the feasible set defined by the inequality constraints

$$g_1(x,y) = -x_1; \qquad g_2(x,y) = x_1 + x_1^3\exp(x_2^2).$$

Clearly, $x^* = (0,0)$ is feasible point and both constraints are active at $x^*$. Furthermore, by direct calculations we have

$$\nabla g_1(x_1,x_2) = (-1,0) \qquad \text{and} \qquad \nabla g_2(x_1,x_2) = \big(1+3x_1^2\exp(x_2^2), 2x_2x_1^3\exp(x_2^2)\big), \quad \forall\,(x_1,x_2)\in\mathbb{R}^2.$$

Thus, $L_\Omega(x^*)^\circ = \{\mu_1(-1,0)+\mu_2(1,0)\colon\ \mu_1,\mu_2\ge 0\} = \mathbb{R}\times\{0\}$.

***The Cone Continuous Property does not hold at $x^* = (0,0)$.*** Define $x^k := 1/k$, $y^k := 1/k$, $\mu_2^k := (2x_2x_1^3\exp(x_2^2))^{-1}$ and $\mu_1^k := \mu_2^k(1+3x_1^2\exp(x_2^2))$. Note that $(x_1^k,x_2^k)\to(0,0)$ and $\mu_1^k,\mu_2^k\ge 0$ for all $k\in\mathbb{N}$. So,

$$\omega^k := \mu_1^k(-1,0)+\mu_2^k\big(1+3x_1^2\exp(x_2^2), 2x_2x_1^3\exp(x_2^2)\big)\in K(x_1^k,x_2^k). \tag{49}$$

By direct calculations, $\omega^k = (0,1)\ \forall\,k\in\mathbb{N}$. Hence, $\lim_{k\to\infty}\omega^k = (0,1)$ is in $\limsup_{x\to x^*}K(x)$ but $(0,1)$ is not in $L_\Omega(x^*)^\circ$; thus, $K(x)$ cannot be outer semi-continuous at $x^*$.

***AGP-regularity holds at $x^*$.*** Take $\omega^* = (\omega_1,\omega_2)\in\limsup_{(x,\varepsilon)\to(x^*,0)}N_{\Omega(x,-\infty)}(x+\varepsilon)$. Then, there are sequences $\{x^k = (x_1^k,x_2^k)\}$, $\{\omega^k\}$ and $\{\varepsilon^k = (\varepsilon_1^k,\varepsilon_2^k)\}$ in $\mathbb{R}^2$ such that $x^k\to x^*$, $\varepsilon^k\to(0,0)$, $\omega^k\to\omega^*$ and $\omega^k\in N_{\Omega(x^k,-\infty)}(x^k+\varepsilon^k)$. To prove that $\omega^*\in N_{\Omega(x^*,-\infty)}(x^*) = L_\Omega(x^*)^\circ$, we must analyze all the different possible cases as $x^k$ approaches to $x^* = (0,0)$. We have the possible cases ($x_1^k>0$, $x_1^k<0$ and $x_1^k=0$ for infinitely many indices in $\mathbb{N}$).

Assume that there are infinitely many indices $k\in\mathbb{N}$ such that

1. $x_1^k>0$ holds. In this case, $g_1(x_1^k,x_2^k)<0$ and $g_2(x_1^k,x_2^k)>0$. For this case, we define two conditions

$$\text{condition } (g_1)\text{: if } g_1(x_1^k,x_2^k)+\big\langle\nabla g_1(x_1^k,x_2^k),(\varepsilon_1^k,\varepsilon_2^k)\big\rangle = 0.$$
$$\text{condition } (g_2)\text{: if } \big\langle\nabla g_2(x_1^k,x_2^k),(\varepsilon_1^k,\varepsilon_2^k)\big\rangle = 0.$$

Whether $\varepsilon^k = (\varepsilon_1^k,\varepsilon_2^k)$ satisfies the conditions $(g_1)$ and $(g_2)$ or not, we have the following subcases:

(a) $\varepsilon^k$ satisfies condition $(g_1)$ and condition $(g_2)$. Since $\varepsilon^k$ satisfies both conditions $(g_1)$ and $(g_2)$, we have

$$\varepsilon_1^k = -x_1^k \qquad \text{and} \qquad \varepsilon_1^k\big(1+3(x_1^k)^2\exp\,(x_2^k)^2\big)+\varepsilon_2^k\big(2x_2^k(x_1^k)^3\exp\,(x_2^k)^2\big) = 0. \tag{50}$$

Using (50), we get

$$-1-3(x_1^k)^2\exp\,(x_2^k)^2+\varepsilon_2^k\big(2x_2^k(x_1^k)^2\exp\,(x_2^k)^2\big) = 0. \tag{51}$$

So, if there is an infinite index set such that the expression (51) holds, we obtain a contradiction, by taking limit in an appropriate subsequence.

(b) $\varepsilon^k$ satisfies condition $(g_1)$ but not condition $(g_2)$. Since $\varepsilon^k$ does not satisfy condition $(g_2)$, we have that the multiplier associated with $\nabla g_2(x_1^k,x_2^k)$ for $\omega^k\in N_{\Omega(x^k,-\infty)}(x^k+\varepsilon^k)$ is null (see Proposition 2). Thus, $\omega^k = \mu_1^k(-1,0)\in\mathbb{R}\times\{0\}$ for some $\mu_1^k\ge 0$. Now, if there is an infinite index set in this subcase, taking limit (for an appropriate subsequence), the limit $\omega^*$ must be in $\mathbb{R}\times\{0\}$;

(c) $\varepsilon^k$ does not satisfy condition $(g_1)$ but satisfies condition $(g_2)$. In this case the multiplier associated with $\nabla g_1(x_1^k,x_2^k)$ is zero. Thus,

$$\omega^k := \mu_2^k\big(1+3x_1^2\exp(x_2^2), 2x_2x_1^3\exp(x_2^2)\big), \quad \text{for some } \mu_2^k\ge 0.$$

Moreover, by condition $(g_2)$, we have

$$\varepsilon_1^k\big(1+3(x_1^k)^2\exp\,(x_2^k)^2\big)+\varepsilon_2^k\big(2x_2^k(x_1^k)^3\exp\,(x_2^k)^2\big) = 0.$$

Now, we will show that if there is an infinite index set in this subcase, $\omega_2^*$ is zero. By contradiction, assume that $\omega_2^*$ is nonzero. For $k$ large enough,

$$2\mu_2^k |x_2^k (x_1^k)^3 \exp{(x_2^k)^2}| > (1/2)|\omega_2^*| > 0, \tag{52}$$

as consequence $x_2^k$ is a positive number. Using the expression above and the definition of $\omega_1^k$, we have

$$\omega_1^k = \mu_2^k + 3\mu_2^k (x_1^k)^2 \exp{(x_2^k)^2} > \frac{3|\omega_2^*|}{4|x_1^k x_2^k|}.$$

Taking limits in this expression for the appropriate subsequence, we obtain a contradiction, since the left side is bounded.

(d) $\varepsilon^k$ satisfies neither condition ($g_1$) nor condition ($g_2$). In this case, the multipliers associated with $\nabla g_1(x^k, y^k)$ and $\nabla g_2(x^k, y^k)$ are both zero, hence $\omega^k = (0,0) \in \mathbb{R} \times \{0\}$. Thus, if there is an infinite set of indices $k$ such that $x^k > 0$ holds, taking limit in the appropriate subsequence, we get that $\omega^* \in \mathbb{R} \times \{0\}$;

2. $x_1^k < 0$ holds. In this case, $g_1(x_1^k, x_2^k) > 0$ and $g_2(x_1^k, x_2^k) < 0$. For this case, we define two conditions

$$\begin{aligned}&\text{condition } (g_1)\text{: if } \langle \nabla g_1(x_1^k, x_2^k), (\varepsilon_1^k, \varepsilon_2^k)\rangle = 0;\\ &\text{condition } (g_2)\text{: if } g_2(x_1^k, x_2^k) + \langle \nabla g_2(x_1^k, x_2^k), (\varepsilon_1^k, \varepsilon_2^k)\rangle = 0.\end{aligned}$$

Depending if $\varepsilon^k$ satisfies the conditions above or not, we have the next subcases:

(a) $\varepsilon^k$ satisfies the condition($g_1$) and the condition ($g_2$). From these conditions we have

$$\varepsilon_1^k = 0 \qquad \text{and} \qquad x_1^k + (x_1^k)^3 \exp x_2^2 + \varepsilon_1^k \big(1 + 3x_1^2 \exp{(x_2^k)^2}\big) + \varepsilon_2^k (2x_2^k \big(x_1^k)^3 \exp{(x_2^k)^2}\big) = 0.$$

Using $\varepsilon_1^k = 0$ and dividing by $x_1^k$ in the last expression, we get

$$1 + (x_1^k)^2 \exp x_2^2 + \varepsilon_2^k \big(2x_2^k (x_1^k)^2 \exp{(x_2^k)^2}\big) = 0.$$

Thus, if there exists an infinite index set such that the expression above holds, taking limits we obtain a contradiction.

(b) $\varepsilon^k$ satisfies condition ($g_1$) but not condition ($g_2$). Since $\varepsilon^k$ satisfies condition ($g_1$), $\varepsilon_1^k = 0$, and since $\varepsilon^k$ does not satisfy condition ($g_2$), the multiplier associated with $\nabla g_2(x^k, y^k)$ is zero; then $\omega^k = \mu_1^k(-1, 0)$ for some $\mu_1^k \le 0$. Taking limit (for an appropriate subsequence) we obtain that $\omega^*$ must be in $\mathbb{R} \times \{0\}$;

(c) $\varepsilon^k$ does not satisfy condition ($g_1$) but satisfies condition ($g_2$). Since $\varepsilon^k$ does not satisfy the condition ($g_1$) the multiplier associated with $\nabla g_1(x^k, y^k)$ is zero by Proposition 2. Thus,

$$\omega^k = \mu_2^k \big(1 + 3(x_1^k)^2 \exp{(x_2^k)^2}, 2x_2^k (x_1^k)^3 \exp{(x_2^k)^2}\big)$$

and

$$x_1^k + (x_1^k)^3 \exp x_2^2 + \varepsilon_1^k \big(1 + 3(x_1^k)^2 \exp{(x_2^k)^2}\big) + \varepsilon_2^k \big(2x_2^k (x_1^k)^3 \exp{(x_2^k)^2}\big).$$

Now, if we assume that $\omega_2^*$ is not zero, we obtain for $k$ sufficiently large that

$$\omega_1^k = \mu_2^k + 3\mu_2^k (x_1^k)^2 \exp{(x_2^k)^2} > \frac{3|\omega_2^*|}{4|x_1^k x_2^k|}.$$

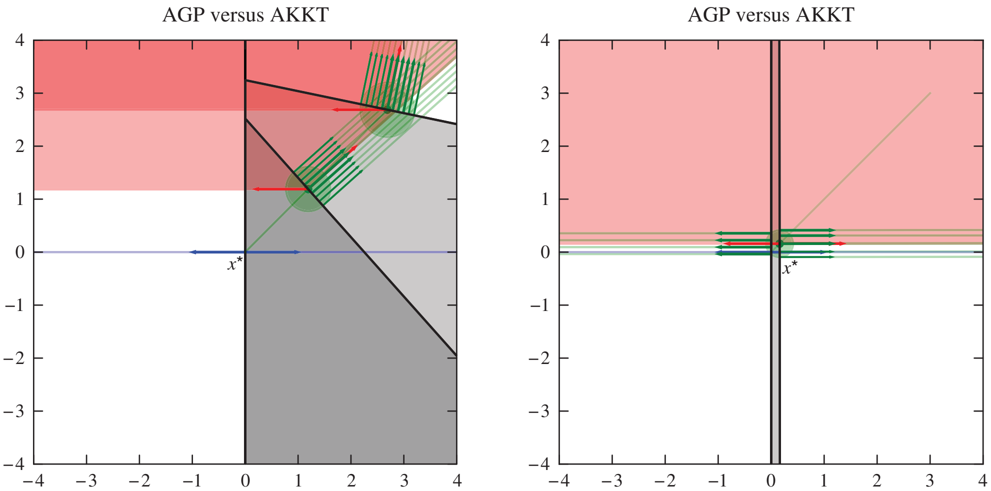
So, if there is an infinite index subset with this property, we get a contradiction, since $\omega_1^k \to \omega_1^*$ and the right-hand side blows out.

(d) $\varepsilon^k$ satisfies neither condition ($g_1$) nor condition ($g_2$). In this subcase, both multipliers associated with $\nabla g_1(x^k, y^k)$ and $\nabla g_2(x^k, y^k)$ are zero, and hence $\omega^k = (0,0)$. Therefore, if there is an infinite index set such that $x_1^k < 0$, taking limit we get that $\omega^*$ belongs to $\mathbb{R} \times \{0\}$;

3. $x_1^k = 0$ holds. For this case, we have $g_1(x_1^k, x_2^k) = 0$ and $g_2(x_1^k, x_2^k) = 0$. After some calculations, we also have $\nabla g_1(x_1^k, x_2^k) = (-1, 0)$ and $\nabla g_2(x_1^k, x_2^k) = (1, 0)$. Thus, $\omega^k = \mu_1^k \nabla g_1(x_1^k, x_2^k) + \mu_2^k \nabla g_2(x_1^k, x_2^k)$ must be in $\mathbb{R} \times \{0\}$. So, if there is an infinite index set for this subcase, the limit, $\omega^*$ must be in $L_\Omega(x^*)^\circ = \mathbb{R} \times \{0\}$.

From all the possible cases, we have that $\omega^* = \lim_{k\to\infty} \omega^k$ must be in $\mathbb{R} \times \{0\} = N_{\Omega(x^*, -\infty)}(x^*)$, in other words, $x^*$ is AGP-regular. See Figure 4.

**Figure 4.** (Color online) Picture associated with Example 1.




*Notes.* The feasible set is simply the vertical line that passes through $x^* = 0$. Once again the blue cone, which is the horizontal line, represents the normal of the linearized cone at $x^*$. The gray areas are the regions associated with the linearization of the constraints at points in $\{x^k\}$. The red regions represent the cones associated with AKKT, that appear in the definition of CCP. The green arrows are the directions that belong the cones associated to AGP. Note that far from the origin the cone associated to AKKT already have two generating directions (associated to the gradients of both constraints that are active at the origin). Since one of the directions in the AKKT cone is always not horizontal, this cone always contains vertical vector $(0,1)^t$. On the other hand, even in points close to the origin the cone associated to AGP has at most one direction at any point of interest and never contains any vertical vector.

**Example 2** (SAKKT-Regularity is Strictly Weaker Than AGP-Regularity)**.** Consider $x^* = (0,0)$ in the Euclidean space $\mathbb{R}^2$ and the feasible set defined by the inequality constraints

$$g_1(x_1, x_2) = -x_1;$$
$$g_2(x_1, x_2) = -x_1^2 - x_2^2.$$

Clearly $x^* = (0,0)$ is a feasible point where both constraints are active. By direct calculations, we get

$$\nabla g_1(x_1, x_2) = (-1, 0) \qquad \text{and} \qquad \nabla g_2(x_1, x_2) = (-2x_1, -2x_2), \quad \forall\, x = (x_1, x_2) \in \mathbb{R}^2.$$

We also have $L_\Omega(x^*)^\circ = \{\mu_1(-1,0) + \mu_2(0,0)\colon \mu_1, \mu_2 \geq 0\} = \mathbb{R}_- \times \{0\}$.

***SAKKT-regularity holds at $x^*$.*** Take $\omega^* = (\omega_1^*, \omega_2^*) \in \limsup_{x \to x^*} N_{\Omega(x,0)}(x)$; then there are sequences $\{x^k = (x_1^k, x_2^k)\}$ and $\{\omega^k\}$ in $\mathbb{R}^2$ such that $x^k \to x^*$, $\omega^k \to \omega^*$ and $\omega^k \in N_{\Omega(x^k,0)}(x^k)$. To show that $\omega^* \in N_{\Omega(x^*,-\infty)}(x^*) = L_\Omega(x^*)^\circ$ we will analyze all the possible cases. Suppose that there are infinitely many indices $k \in \mathbb{N}$ such that at least one of the following cases hold:

1. $x_1^k > 0$. In this case, we have $g_1(x_1^k, x_2^k) < 0$ and $g_2(x_1^k, x_2^k) < 0$. From Proposition 3, $\omega^k = (0,0)$, since the multipliers associated with $\nabla g_1(x_1^k, x_2^k)$ and $\nabla g_2(x_1^k, x_2^k)$ are not zero only when $g_1(x_1^k, x_2^k) \geq 0$ or $g_2(x_1^k, x_2^k) \geq 0$.
2. $x_1^k < 0$. In this case, we have $g_1(x_1^k, x_2^k) > 0$ and $g_2(x_1^k, x_2^k) < 0$. Using Proposition 3, the multipliers associated with $\nabla g_2(x_1^k, x_2^k)$ are zero. Thus, $\omega^k$ takes the form $\mu_1^k(-1,0)$ for some $\mu_1^k \geq 0$. Clearly, $\mu_1^k(-1,0) \in L_\Omega(x^*)^\circ = \mathbb{R}_- \times \{0\}$.
3. $x^k = 0$. In this case, both functions are nonnegative and, depending of the value of $x_2^k$, $g_2(x_1^k, x_2^k)$, can be strictly negative or zero. Consider the following subcases:
   (a) $x_2^k = 0$ for infinitely many indices. By direct calculations, $g_2(x_1^k, x_2^k) = 0$, so

$$\omega^k = \mu_1^k(-1,0) + \mu_2^k(-2x_1^k, -2x_2^k) = \mu_1^k(-1,0) + \mu_2^k(0,0) \in \mathbb{R}_- \times \{0\} = L_\Omega(x^*)^\circ.$$

   (b) $x_2^k \neq 0$ for infinitely many indices. In this subcase, $g_2(x_1^k, x_2^k) < 0$. From Proposition 3, the multipliers associated with $g_2(x_1^k, x_2^k)$ are zero and $\omega^k = \mu_1^k(-1,0) \in \mathbb{R}_- \times \{0\}$ for some $\mu_1^k \geq 0$.

Therefore, in all the possible cases, we obtain (taking an appropriate subsequence) that $\omega^*$ belongs to $K(x^*) = \mathbb{R}_- \times \{0\}$, This amounts to SAKKT-regularity at $x^*$ as we wanted to show.

***$x^*$ is not AGP-regular.*** For every $k \in \mathbb{N}$, define $x_1^k := 1/k$, $x_2^k := 1/k$, $\varepsilon_1^k := -x_1^k$, $\varepsilon_2^k := 0$ and multipliers $\mu_2^k := (2x_1^k)^{-1}$ and $\mu_1^k := 0$. Also, define $\omega^k := \mu_1^k(-1,0) + \mu_2^k(-2x_1^k, -2x_2^k)$. Obviously, $\varepsilon^k \to 0$. From Proposition 2, $\omega^k \in N_{\Omega((x_1^k, x_2^k),-\infty)}((x_1^k, x_2^k) + (\varepsilon_1^k, \varepsilon_2^k))$; furthermore, due to the choice of $\mu_1^k$ and $\mu_2^k$, $\omega^k = (-1,-1)\ \forall\, k \in \mathbb{N}$. Thus, $(-1,-1) \in \limsup_{(x,\varepsilon) \to (x^*,0)} N_{\Omega(x,-\infty)}(x + \varepsilon)$, but $(-1,-1)$ does not belong to $L_\Omega(x^*)^\circ = \mathbb{R}_- \times \{0\}$. As a consequence, $x^*$ is not AGP-regular.

**Example 3** (LAGP-Regularity is Strictly Weaker Than AGP-Regularity). Define $x^*=(0,0)$ and let the feasible set be given by the equality and inequality constraints

$$h(x_1,x_2)=x_1;$$
$$g(x_1,x_2)=x_1-x_1^2x_2.$$

The point $x^*=(0,0)$ is feasible and both constraints are active at $x^*$. By straight calculations, we have that

$$\nabla h(x_1,x_2)=(1,0) \quad \text{and} \quad \nabla g(x_1,x_2)=(1-2x_1x_2,-x_1^2), \quad \text{for all } x=(x_1,x_2)\in\mathbb{R}^2.$$

Furthermore, $L_\Omega(x^*)^\circ=\{\lambda(1,0)+\mu(1,0)\colon \lambda\in\mathbb{R},\mu\geq 0\}=\mathbb{R}\times\{0\}$.

***$x^*=(0,0)$ is LAGP-regular.*** First, we note that the set of linear constraint, $\Omega_L$, is given by the only equality constraint $h(x_1,x_2)=0$, so:

$$\Omega_L=\{(x_1,x_2)\in\mathbb{R}^2\colon h(x_1,x_2)=0\}=\{(x_1,x_2)\in\mathbb{R}^2\colon x_1=0\}=\{0\}\times\mathbb{R}.$$

Now, we will show that $N_{\Omega_{NL}(x,-\infty)\cap\Omega_L}(x+\varepsilon)$ is outer semicontinuous at $(x^*,0)$ relatively to $\Omega_L\times\mathbb{R}^2$. Take $\omega^*=(\omega_1,\omega_2)\in\limsup N_{\Omega_{NL}(x,-\infty)\cap\Omega_L}(x+\varepsilon)$ relatively to $\Omega_L\times\mathbb{R}^2$. From the definition of outer limit, there are sequences $\{x^k\}$, $\{\omega^k\}$ and $\{\varepsilon^k\}$ in $\mathbb{R}^2$ such that $x^k\to x^*$, $\varepsilon^k\to(0,0)$, $\omega^k\to\omega^*$ and

$$x^k\in\Omega_L, \qquad x^k+\varepsilon^k\in\Omega_{NL}(x^k,-\infty)\cap\Omega_L, \qquad \text{and} \qquad \omega^k\in N_{\Omega_{NL}(x^k,-\infty)\cap\Omega_L}(x^k+\varepsilon^k).$$

To see that $\omega^*$ belongs to $N_{\Omega_{NL}(x^*,-\infty)\cap\Omega_L}(x^*+0)=L_\Omega(x^*)^\circ$, we will analyze all the possible cases. Since $x^k\in\Omega_L$ and $x^k+\varepsilon^k\in\Omega_L$ we have $x_1^k=0$ and $\varepsilon_1^k=0$ and, as a consequence, $g(x_1^k,x_2^k)=0$ for all $k\in\mathbb{N}$. Thus, independently of the choice of $\epsilon_2^k$, the following condition holds:

$$\text{condition } (g)\colon \left\langle \nabla g(x_1^k,x_2^k),(\varepsilon_1^k,\varepsilon_2^k)\right\rangle=0.$$

This is a simple consequence of the following observation:

$$\left\langle \nabla g(x_1^k,x_2^k),(\varepsilon_1^k,\varepsilon_2^k)\right\rangle=\varepsilon_1^k(1-2x_1^kx_2^k)+\varepsilon_2^k(-(x_1^k)^2)=0.(1-0)+\varepsilon_2^k.0=0,$$

whenever $x_1^k=\varepsilon_1^k=0$. Since the condition $(g)$ is valid, there exist $\lambda^k$ and $\mu^k\geq 0$ (not necessary all zeroes) such that

$$\omega^k=\lambda^k\nabla h(x_1^k,x_2^k)+\mu^k\nabla g(x_1^k,x_2^k)\in N_{\Omega_{NL}(x^k,-\infty)\cap\Omega_L}(x^k+\varepsilon^k).$$

But, since $x^k=0$, we get $\nabla h(x_1,x_2)=(1,0)$ and $\nabla g(x_1^k,x_2^k)=(1,0)$ and, thus, $\omega^k\in\mathbb{R}\times\{0\}$ for all $k\in\mathbb{N}$, which implies $\omega^*=\lim_{k\to\infty}\omega^k\in\mathbb{R}\times\{0\}=K(x^*)$.

***$x^*=(0,0)$ is not AGP-regular.*** Define $x_1^k:=1/k$, $x_2^k:=x_1^k$, $\varepsilon_1^k:=x_2^k(x_1^k)^2$ and $\varepsilon_2^k=x_2^k(1-2x_1^kx_2^k)$. Clearly, $\varepsilon^k\to(0,0)$ and $g(x_1^k,x_2^k)=x_1^k(1-x_1^kx_2^k)>0$ (for $k$ large enough). Define multipliers $\mu^k:=((x_1^k)^2)^{-1}\in\mathbb{R}_+$ and $\lambda^k:=-\mu^k(1-2x_1^kx_2^k)$ and the sequence $\{\omega^k\}$ given by

$$\omega^k:=\lambda^k\nabla h(x_1^k,x_2^k)+\mu^k\nabla g(x_1^k,x_2^k)=\lambda^k(1,0)+\mu^k(1-2x_1^kx_2^k,-(x_1^k)^2)=(0,-1).$$

Since, for all $k\in\mathbb{N}$,

$$\langle\nabla g(x_1^k,x_2^k),(\varepsilon_1^k,\varepsilon_2^k)\rangle=\varepsilon_1^k(1-2x_1^kx_2^k)+\varepsilon_2^k(-(x_1^2)^k)=x_2^k(x_1^k)^2(1-2x_1^kx_2^k)+x_2^k(1-2x_1^kx_2^k)(-(x_1^2)^k)=0,$$

we have, from Proposition (3), that $\omega^k=(0,-1)\in N_{\Omega(x^k,-\infty)}(x^k+\varepsilon^k)$ for all $k\in\mathbb{N}$ and $\lim_{k\to\infty}\omega^k=(0,-1)$ is in $\limsup_{(x,\varepsilon)\to(x^*,0)}N_{\Omega(x,-\infty)}(x+\varepsilon)$, but $(0,-1)$ does not belong to $L_\Omega(x^*)^\circ=\mathbb{R}\times\{0\}$. So, $x^*$ is not AGP-regular.

**Example 4** (CAKKT-Regularity Does Not Imply SAKKT-Regularity). Consider $x^*=(0,0)$ and the feasible set defined by the equality and inequality constraints

$$h(x_1,x_2)=x_1; \qquad g(x_1,x_2)=x_1\exp x_2.$$

Obviously, $x^*=(0,0)$ is feasible and the inequality constraint is active. Moreover,

$$\nabla h(x_1,x_2)=(1,0) \qquad \text{and} \qquad \nabla g(x_1,x_2)=(\exp x_2,x_1\exp x_2).$$

From the last expression, we get $L_\Omega(x^*)^\circ=\{\lambda(1,0)+\mu(1,0)\colon \lambda\in\mathbb{R},\mu\geq 0\}=\mathbb{R}\times\{0\}$.

***$x^*$ is CAKKT-regular.*** Take $\omega^* = (\omega_1^*, \omega_2^*) \in \limsup_{(x,r)\to(x^*,0)} K_C(z,r)$; then, there exist sequences $\{x^k\}$ and $\{\omega^k\}$ in $\mathbb{R}^2$ and scalars $r^k \geq 0$ such that $x^k \to x^*$, $\omega^k \to \omega^*$, $\omega^k \in K_C(z^k, r^k)$ and $r^k \to 0$. Since $\omega^k \in K_C(z^k, r^k)$, there are sequences $\lambda^k$ and $\mu^k \geq 0$ such that

$$\omega^k = \lambda^k \nabla h(x^k) + \mu^k \nabla g(x^k) = \lambda^k(1,0) + \mu^k(\exp x_2^k, x_1^k \exp x_2^k) \tag{53}$$

and

$$|\lambda^k h(x^k)| + |\mu^k g(x^k)| = |\lambda^k x_1^k| + |\mu^k x_1^k \exp x_2^k| \leq r^k. \tag{54}$$

Using (53) and (54) we get $|\omega_2^k = \mu^k x_1^k \exp x_2^k| \leq r^k$ and $\omega_2^k \to 0$. From the last expression we conclude that $\omega^*$ is in $L_\Omega(x^*)^\circ = \mathbb{R}\times\{0\}$ and CAKKT-regularity holds.

***$x^*$ is not SAKKT-regular.*** Take $x_1^k := 1/k$, $x_2^k := x_1^k$, $\mu^k := (x_1^k \exp x_2^k)^{-1}$ and $\lambda^k := -\mu^k \exp x_2^k$. Since $g(x_1^k, x_2^k) > 0$, we have that

$$\omega^k := \lambda^k(1,0) + \mu^k(\exp x_2^k, x_1^k \exp x_2^k) = (0,1) \in N_{\Omega(x^k,0)}(x^k), \quad \text{for all } k \in \mathbb{N}.$$

Clearly, $\lim_{k\to\infty} \omega^k = (0,1) \in \limsup N_{\Omega(x,0)}(x)$; however $(0,1)$ is not in $\mathbb{R}\times\{0\}$. Thus, SAKKT-regularity fails.

We showed that all the implications of Theorem 5 are strict. The rest of this section is devoted to show the independence between LAGP-regularity and the conditions CAKKT-regularity and SAKKT-regularity.

The following example shows that SAKKT-regularity does not imply LAGP-regularity and, as a consequence, it does not imply CAKKT-regularity either, since CAKKT-regularity is implied by SAKKT-regularity.

**Example 5** (SAKKT-Regularity Does Not Imply LAGP-Regularity). Consider the feasible set expressed by the following equality and inequality constraints

$$\begin{aligned} &h(x_1, x_2) = x_1;\\ &g(x_1, x_2) = -x_1^2 - x_1^2 x_2^2 - x_2^2. \end{aligned}$$

Clearly, $x^* = (0,0)$ is a feasible point and both constraints are active at $x^*$. By straight calculations we have:

$$\nabla h(x_1, x_2) = (1,0) \quad \text{and} \quad \nabla g(x_1,x_2) = (-2x_1 - 2x_1x_2^2, -2x_2x_1^2 - 2x_2), \quad \text{for all } x = (x_1, x_2) \in \mathbb{R}^2$$

Moreover, $L_\Omega(x^*)^\circ = \{\lambda(1,0) + \mu(0,0) \colon \lambda \in \mathbb{R}, \mu \in \mathbb{R}_+\} = \mathbb{R}\times\{0\}$.

***$x^*$ is SAKKT-regular.*** Our aim is to show that the set-valued mapping $N_{\Omega(x,0)}(x)$ is outer semicontinuous at $x^*$. Take $\omega^* = (\omega_1^*, \omega_2^*) \in \limsup N_{\Omega(x,0)}(x)$. From the definition of outer limit, there are sequences $\{x^k\}$ and $\{\omega^k\}$ in $\mathbb{R}^2$ such that $x^k \to x^*$, $\omega^k \to \omega^*$ and $\omega^k \in N_{\Omega(x^k,0)}(x^k)$. We have two possible cases.

- There is an infinite set of indices $k$ such that $x_1^k \neq 0$. In this case, $g(x_1^k, x_2^k) = -(x_1^k)^2(1 + (x_2^k)^2) - x_2^2$ is always negative; thus, the multipliers associated with $\nabla g(x_1^k, x_2^k)$ are zero (Proposition (3)) Then, $\omega^k$ has the form $\lambda^k \nabla h(x_1^k, x_2^k) = \lambda^k(1,0) \in \mathbb{R}\times\{0\}$ for some $\lambda^k \in \mathbb{R}$. Taking the appropriate subsequence, we get $\omega^* \in \mathbb{R}\times\{0\}$;
- There is an infinite set of indices $k$ such that $x_1^k = 0$. In this case, $g(x_1^k, x_2^k) = -x_2^2$. Now, depending of the values of $x_2^k$, we obtain the following subcases:

—$x_2^k \neq 0$. In this case $g(x_1^k, x_2^k) < 0$. So, from Proposition (3), the multipliers associated with $\nabla g(x^k, y^k)$ are zero. Thus, $\omega^k = \lambda^k \nabla h(x_1^k, x_2^k) = \lambda^k(1,0) \in \mathbb{R}\times\{0\}$ for some $\lambda^k \in \mathbb{R}$. Taking limit (in an appropriate subsequence), we get $\omega^* \in \mathbb{R}\times\{0\}$;

—$x_2^k = 0$. In this case, $g(x_1^k, x_2^k) = (0,0)$ and there exist $\lambda^k \in \mathbb{R}$ and $\mu^k \in \mathbb{R}_+$ such that $\omega^k = \lambda^k(1,0) + \mu^k(-2x_1^k - 2x_1^k(x_2^k)^2, -2x_2^k(x_1^k)^2 - 2x_2^k) = \lambda^k(1,0) \in \mathbb{R}\times\{0\}$, where, in the last equality, we have used $(x_1^k, x_2^k) = (0,0)$. So, $\omega^k$ is in $\mathbb{R}\times\{0\}$ and, taking limit for an appropriate subsequence, $\omega^* \in \mathbb{R}\times\{0\}$.

In all the analyzed cases, we concluded that $\omega^*$ belongs to $\mathbb{R}\times\{0\}$. This proves the outer semicontinuity of the multifunction $N_{\Omega(x,0)}(x)$ at $x^* = (0,0)$.

***$x^*$ is not LAGP-regular.*** Since the only linear constraint is given by $h$ (an equality constraint), we have:

$$\Omega_L = \{x = (x_1, x_2) \in \mathbb{R}^2 \colon h(x_1, x_2) = 0\} = \{x = (x_1, x_2) \in \mathbb{R}^2 \colon x_1 = 0\} = \{0\}\times\mathbb{R}.$$

Now, define $x_1^k := 0$, $x_2^k := 1/k$, $\varepsilon_1^k = 0$ and $\varepsilon_2^k = -x_2^k/2$. Clearly, all these sequences go to zero. For that choice, we see that $x^k$ and $x^k + \varepsilon^k$ are in $\Omega_L$. Moreover, $g(x_1^k, x_2^k) = -(x_2^k)^2$ is a negative scalar and the following expression holds for all $k \in \mathbb{N}$:

$$g(x^k, y^k) + \langle \nabla g(x^k, y^k), (\varepsilon_1^k, \varepsilon_2^k)\rangle = 0. \tag{55}$$

By (55), we can define $\mu^k := (2x_2^k)^{-1}$, $\lambda^k := 1$, so that

$$\omega^k := \lambda^k \nabla h(x_1^k, x_2^k) + \mu^k \nabla g(x_1^k, x_2^k) = \lambda^k(1,0) + \mu^k(0, -2x_2^k) = (1,-1) \in N_{\Omega_{NL}(x^k,-\infty)\cap\Omega_L}(x^k + \varepsilon^k).$$

Thus, $\lim \omega^k = (1,-1) \in \limsup N_{\Omega_{NL}(x^k,-\infty)\cap\Omega_L}(x^k + \varepsilon^k)$ relatively to $\Omega_L \times \mathbb{R}^2$. Clearly, $(1,-1)$ is not in $L_\Omega(x^*)^\circ = \mathbb{R} \times \{0\}$. Hence, LAGP-regularity does not hold at $x^*$.

The next example shows that LAGP-regularity does not imply CAKKT-regularity and, consequently, does not imply SAKKT-regularity either.

**Example 6** (LAGP-Regularity Does Not Imply CAKKT-Regularity). Consider $x^* = (0,0)$ and the feasible set defined by the equality and inequality constraints

$$h(x_1, x_2) = x_1;$$
$$g(x_1, x_2) = x_1 - x_1^2 x_2 - x_2^2.$$

Clearly, $x^* = (0,0)$ is feasible and both constraints are active. We also have

$$\nabla h(x_1, x_2) = (-1, 0) \quad \text{and} \quad \nabla g(x_1, x_2) = (1 - 2x_1x_2, -2x_2 - x_1^2) \quad \text{for all } (x_1, x_2) \in \mathbb{R}^2.$$

Hence, $L_\Omega(x^*)^\circ = \{\lambda(1,0) + \mu(1,0)\colon \lambda \in \mathbb{R}, \mu \geq 0\} = \mathbb{R} \times \{0\}$.

***$x^*$ is LAGP-regular.*** From the equality constraint we get that $\Omega_L = \{x = (x_1, x_2) \in \mathbb{R}^2\colon x_1 = 0\} = \{0\} \times \mathbb{R}$. We will show that $N_{\Omega_{NL}(x,-\infty)\cap\Omega_L}(x + \varepsilon)$ is outer semicontinuous at $(x^*, 0)$ relatively to $\Omega_L \times \mathbb{R}^2$. Pick $\omega^* = (\omega_1, \omega_2) \in \limsup N_{\Omega_{NL}(x,-\infty)\cap\Omega_L}(x + \varepsilon)$ relatively to $\Omega_L \times \mathbb{R}^2$. Thus, there are sequences $\{x^k\}$, $\{\omega^k\}$ and $\{\varepsilon^k\}$ in $\mathbb{R}^2$ such that $x^k \to x^*$, $\varepsilon^k \to (0,0)$, $\omega^k \to \omega^*$, and

$$x^k \in \Omega_L, \qquad x^k + \varepsilon^k \in \Omega_{NL}(x^k, -\infty) \cap \Omega_L, \qquad \omega^k \in N_{\Omega_{NL}(x^k,-\infty)\cap\Omega_L}(x^k + \varepsilon^k).$$

Since $x^k \in \Omega_L$ and $x^k + \varepsilon^k \in \Omega_L$ we have $x_1^k = 0$ and $\varepsilon_1^k = 0$ and, as a consequence, $g(x_1^k, x_2^k) = -(x_2^k)^2$ for all $k \in \mathbb{N}$. To see that $\omega^*$ belongs to $N_{\Omega_{NL}(x^*,-\infty)\cap\Omega_L}(x^* + 0) = L_\Omega(x^*)^\circ$, we will analyze all the possible cases depending on the value of $x_2^k$. Assume that there is an infinite set of indices such that:

- $x_2^k \neq 0$. In this case, $g(x_1^k, x_2^k) = -(x_2^k)^2$ is strictly negative and $\omega^k = \lambda^k(1,0)$ for some $\lambda^k \in \mathbb{R}$, so, $\omega^k \in L_\Omega(x^*)^\circ = \mathbb{R} \times \{0\}$.
- $x_2^k = 0$. In this case, $g(x_1^k, x_2^k) = 0$, for any value of $\varepsilon_2^k$, $(1 - 2x_1^k x_2^k)\varepsilon_1^k + (-2x_2^k - (x_1^k)^2)\varepsilon_2^k = 0$. Then, there are $\lambda^k \in \mathbb{R}$ and $\mu^k \geq 0$ such that

$$\omega^k = \lambda^k(1,0) + \mu^k(1 - 2x_1^k x_2^k, -2x_2^k - (x_1^k)^2) = (\lambda^k + \mu^k, 0) \in L_\Omega(x^*)^\circ = \mathbb{R} \times \{0\}$$

where the last equality holds because $x_1^k = x_2^k = 0$, in this case.

Thus, for all the cases, we conclude that the limit $\omega^*$ must be in $L_\Omega(x^*)^\circ = \mathbb{R} \times \{0\}$.

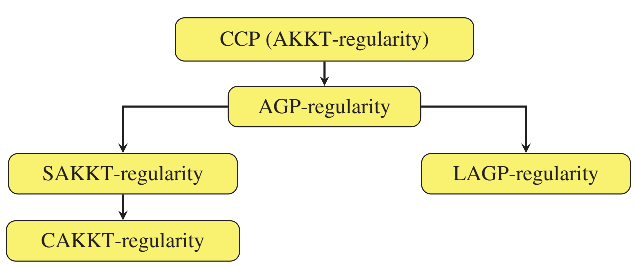
***$x^*$ is not CAKKT-regular.*** Take $x_1^k := 1/k$, $x_2^k := (x_1^k)^{1/2}$, $\mu^k := (x_2^k)^{-1}$, $\lambda^k := -\mu^k(1 - 2x_1^k x_2^k)$ and define the sequence $\{\omega^k\}$ as

$$\omega^k = \lambda^k(1,0) + \mu^k\left(1 - 2x_1^k x_2^k, -2x_2^k - (x_1^k)^2\right) = (0, -2 - (x_1^k)^{3/2}) \in K_C(x^k, r^k),$$

where $r^k := |\mu^k g(x_1^k, x_2^k)| + |\lambda^k h(x_1^k, x_2^k)| = (x_1^k)^2 + 1/2(1 - 2x_1^k x_2^k)(x_1^k)^{1/2}$. Since $x^k \to x^*$ and $r^k \to 0$, $\omega := \lim \omega^k = (0,-2)$ belongs to $\limsup_{(x,r)\to(x^*,0)} K_C(x,r)$ but not in $L_\Omega(x^*)^\circ = \mathbb{R} \times \{0\}$.

Figure 5 shows the implications between the strict constraint qualifications considered in this paper.

**Figure 5.** (Color online) Implications between strict constraint qualifications.



## 6. Relations with Other Constraint Qualifications

Recall that strict constraint qualifications are constraint qualifications. In fact, if a point is a local minimizer, it satisfies every sequential optimality condition and, if it also satisfies an associated strict constraint qualification, it necessarily fulfills KKT. Therefore, every local minimizer that satisfies a strict constraint qualification fulfills the KKT conditions. Therefore, it is natural to establish the relations between strict constraint qualifications and other constraint qualifications.

### 6.1. Strict Constraint Qualifications and Abadie's Constraint Qualification

In this subsection, we will show that both CAKKT-regularity and LAGP-regularity are strictly stronger than Abadie's constraint qualification.

Let us start with the following two auxiliary lemmas.

**Lemma 1** (Rockafellar and Wets [36, theorem 6.11]). *Let $\bar{x}$ be a feasible point. For every $y \in T^{\circ}_{\Omega}(\bar{x})$, there is a smooth function $F$ with $-\nabla F(\bar{x}) = y$ and such that $\bar{x}$ is a strict global minimizer of $F$ with respect to $\Omega$.*

**Lemma 2.** *Take $y \in T^{\circ}_{\Omega}(\bar{x})$, then there are sequences $\{x^k\} \subset \mathbb{R}^n$, $\{\lambda^k\} \subset \mathbb{R}^m$ and $\{\mu^k\} \subset \mathbb{R}^p_+$ such that:*

1. *$\{x^k\}$ converges to $\bar{x}$;*
2. *$\omega^k := \sum_{i=1}^m \lambda_i^k \nabla h_i(x^k) + \sum_{j=1}^p \mu_j^k \nabla g_j(x^k) \to y$;*
3. *For all $j \in \{1,..,p\}$, $\mu_j^k$ is proportional to $\max\{0, g_j(x^k)\}$;*
4. *$r^k := \sum_{i=1}^m |\lambda_i^k h_i(x^k)| + \sum_{j=1}^p |\mu_j^k g_j(x^k)| \to 0$.*

**Proof.** Let $y \in T^{\circ}_{\Omega}(\bar{x})$. From Lemma 1, there exists a smooth function $F$ such that $-\nabla F(\bar{x}) = y$ and $F$ attains its strictly global minimum with respect to $\Omega$ at $\bar{x}$. Pick $r > 0$, and for every $k \in \mathbb{N}$, consider the optimization problem

$$\begin{aligned} &\text{Minimize } F_k(x) := F(x) + \frac{k}{2}\left(\sum_{j=1}^p \max\{g_j(x), 0\}^2 + \sum_{i=1}^m h_i^2(x)\right) \\ &\text{subject to } x \in \mathbb{B}(\bar{x}, r). \end{aligned} \tag{56}$$

By Weierstrass' theorem, there is a solution $x^k$ for the optimization problem (56). Using penalty arguments, we get $x^k \to \bar{x}$,

$$\nabla F(x^k) + \sum_{i=1}^m k h_i(x^k) \nabla h_i(x^k) + \sum_{j=1}^p k \max\{g_j(x^k), 0\} \nabla g_j(x^k) = 0, \tag{57}$$

and

$$\sum_{i=1}^m k h_i(x^k)^2 + \sum_{j=1}^p k \max\{g_j(x^k), 0\}^2 \le F(x^*) - F(x^k). \tag{58}$$

Define $\lambda_i^k := k h_i(x^k)$ for $i \in \{1, \ldots, m\}$ and $\mu_j^k := k \max\{g_j(x^k), 0\}$ for $j \in \{1, \ldots, p\}$. Thus, we have that $\omega^k = \sum_{j=1}^p k \max\{g_j(x^k), 0\} \nabla g_j(x^k) + \sum_{i=1}^m k h_i(x^k) \nabla h_i(x^k)$. From the expression (57) and the continuity of $\nabla F$, $\omega^k \to y$. Finally, since $r^k = \sum_{i=1}^m |\lambda_i^k h_i(x^k)| + \sum_{j=1}^p |\mu_j^k g_j(x^k)|$, from the continuity of $F$ and by (58), we get $r^k \to 0$. □

The next lemma is a variation of the lemma above, useful for the analysis of LAGP-regularity.

**Lemma 3.** *Let $y$ be an element in $T^{\circ}_{\Omega}(\bar{x})$. Then, there are sequences $\{x^k\} \subset \Omega_L$ and $\{\omega^k\} \subset \mathbb{R}^m$ such that $x^k \to \bar{x}$, $\omega^k \to y$ and $\omega^k \in N_{\Omega_{NL}(x^k, -\infty) \cap \Omega_L}(x^k)$.*

**Proof.** Since $y$ belongs to $T^{\circ}_{\Omega}(\bar{x})$, we have, by Lemma 1, that there exists a smooth function $F$ such that $-\nabla F(\bar{x}) = y$ and $F$ attains its strictly global minimum with respect to $\Omega$ at $\bar{x}$. Without loss of generality, we may assume that $\{g_j\colon j \in \{1,..,p_1\}\}(p_1 \le p)$ and $\{h_i\colon i \in \{1,..,m_1\}\}(m_1 \le m)$ define the nonlinear constraints.

Take $r > 0$ and for every $k \in \mathbb{N}$, consider the optimization problem

$$\begin{aligned} &\text{Minimize } F_k(x) := F(x) + \frac{k}{2}\left(\sum_{j=1}^{p_1} \max\{g_j(x), 0\}^2 + \sum_{i=1}^{m_1} h_i^2(x)\right) \\ &\text{subject } x \in \mathbb{B}(\bar{x}, r) \cap \Omega_L. \end{aligned} \tag{59}$$

where $\Omega_L$ is the feasible set defined by the linear constraints. From Weierstrass' theorem, there is a minimizer $x^k$ for (59). Furthermore, by penalty arguments, $\{x^k\}$ converges to $\bar{x}$; thus, for $k$ large enough, $x^k \in \text{Int}(\mathbb{B}(\bar{x}, r))$. Using the geometric optimality condition (17), we get $\langle \nabla F_k(x^k), d \rangle \ge 0$ for every direction $d \in T_{\Omega_L}(x^k)$ or, equivalently,

$$-\nabla F_k(x^k) \in N_{\Omega_L}(x^k) = T_{\Omega_L}(x^k)^{\circ}.$$

Taking the derivative of $F_k$ at $x^k$, we obtain the following expression:

$$-\left((\nabla F(x^k)+\sum_{j=1}^{p_1} k\max\{g_j(x^k),0\}\nabla g_j(x^k)+\sum_{i=1}^{n_1} kh_i(x^k)\nabla h_i(x^k))\right)\in N_{\Omega_L}(x^k). \quad (60)$$

Define $\lambda_i^k := kh_i(x^k)$ for $i\in\{1,\dots,m_1\}$ and $\mu_j^k := k\max\{g_j(x^k),0\}$ for $j\in\{1,\dots,p_1\}$. We also define $\omega_1^k := \sum_{j=1}^{p_1}\mu_j^k\nabla g_j(x^k)+\sum_{i=1}^{m_1}\lambda_i^k\nabla h_i(x^k)$ and $\omega_2^k := -\nabla F(x^k)-\omega_1^k$. From the definition of $\Omega_{NL}(x^k,-\infty)$ and (60), it follows that $\omega_1^k\in N_{\Omega_{NL}(x^k,-\infty)}(x^k)$ and $\omega_2^k\in N_{\Omega_L}(x^k)$. Finally, define $\omega^k := \omega_1^k+\omega_2^k = -\nabla F(x^k)$. Clearly, $\omega^k\to -\nabla F(x^*)=y$ and

$$\omega^k=\omega_1^k+\omega_2^k\in N_{\Omega_{NL}(x^k,-\infty)}(x^k)+N_{\Omega_L}(x^k)\subset N_{\Omega_{NL}(x^k,-\infty)\cap\Omega_L}(x^k).$$

So, the sequence $\{\omega^k\}$ satisfies all the required properties. □

The fact that CAKKT-regularity implies Abadie's constraint qualification is proved in the following theorem.

**Theorem 6.** *CAKKT-regularity implies Abadie's constraint qualification.*

**Proof.** Abadie's constraint qualification says that $T_\Omega(x^*)=L_\Omega(x^*)$. Since $T_\Omega(x^*)\subset L_\Omega(x^*)$ always holds, we must show the other inclusion. In order to show the inclusion $L_\Omega(x^*)\subset T_\Omega(x^*)$ we will first show the inclusion $N_\Omega(x^*)\subset L_\Omega(x^*)^\circ$ or, equivalently, $N_\Omega(x^*)\subset K_C(x^*,0)$ (Note that for $x^*\in\Omega$, we have $K_C(x^*,0)=L_\Omega(x^*)^\circ$.)

Take $y\in N_\Omega(x^*)$; from the definition of the normal cone (16) there are sequences $\{x^k\}\subset\Omega$ and $\{y^k\}$ such that

$$x^k\to x^*,\quad y^k\to y,\quad \text{and}\quad y^k\in\hat N_\Omega(x^k)=T_\Omega^\circ(x^k).$$

Using Lemma 2, for each $y^k\in T_\Omega^\circ(x^k)$, we may find sequences with limits $x^k$ and $y^k$ such that the conclusions of the Lemma 2 holds. Hence, for each $k\in\mathbb{N}$, there is a number $j(k)\in\mathbb{N}$, scalars $r^{j(k)}$ and vector $x^{j(k)}$ and $\omega^{j(k)}$ such that

- $\|x^k-x^{j(k)}\|<1/2^k$ for all $k\in\mathbb{N}$;
- $\omega^{j(k)}=\sum_{i=1}^m\lambda_i^{j(k)}\nabla h_i(x^{j(k)})+\sum_{s=1}^p\mu_s^{j(k)}\nabla g_s(x^{j(k)})$;
- $\|y^k-w^{j(k)}\|<1/2^k$ for all $k\in\mathbb{N}$;
- $\mu_s^{j(k)}=j(k)\max\{g_s(x^{j(k)}),0\}$ for all $s\in\{1,..,p\}$;
- $r^{j(k)}=\sum_{i=1}^m|\lambda_i^{j(k)}h_i(x^{j(k)})|+\sum_{s=1}^p|\mu_s^{j(k)}g_s(x^{j(k)})|\le 1/2^k$ for all $k\in\mathbb{N}$.

Obviously, the sequences $\{r^{j(k)}\}$, $\{x^{j(k)}\}$ and $\{\omega^{j(k)}\}$ converge, respectively, to 0, $x^*$ and $y$. Furthermore, for $k$ large enough, $\omega^{j(k)}$ is in $K_C(x^{j(k)},r^{j(k)})$, since (for $k$ large ), $\mu_s^{j(k)}=j(k)\max\{g_s(x^{j(k)}),0\}=0$ for all $s\notin J(x^*)$. Summing up, $x^{j(k)}\to x^*$, $\omega^{j(k)}\to y$, $r^{j(k)}\to 0$, and $\omega^{j(k)}\in K_C(x^{j(k)},r^{j(k)})$. Thus, due to the definition of outer limit $y\in\limsup_{(x,r)\to(x^*,0)}K_C(x,r)\subset L_\Omega(x^*)^\circ$ where the last inclusion holds since CAKKT-regularity also holds at $x^*$. So, we proved the inclusion $N_\Omega(x^*)\subset L_\Omega(x^*)^\circ$, which implies

$$L_\Omega(x^*)=L_\Omega(x^*)^{\circ\circ}\subset N_\Omega(x^*)^\circ\subset T_\Omega(x^*),$$

where the last inclusion follows from Rockafellar and Wets [36, theorems 6.28(b) and 6.26]. □

For LAGP-regularity we have the following theorem.

**Theorem 7.** *LAGP-regularity implies Abadie's constraint qualification.*

**Proof.** We only need to show the inclusion $N_\Omega(x^*)\subset L_\Omega(x^*)^\circ$. Take $y\in N_\Omega(x^*)$. Then, there are sequences $\{x^k\}\subset\Omega$ and $\{y^k\}$ such that

$$x^k\to x^*,\qquad y^k\to y\qquad \text{and}\qquad y^k\in T_\Omega^\circ(x^k).$$

Using Lemma 3, for each $y^k\in T_\Omega^\circ(x^k)$, we have for each $k\in\mathbb{N}$, a number $j(k)\in\mathbb{N}$ and vectors $x^{j(k)}$ and $\omega^{j(k)}$ such that

- $\|x^k-x^{j(k)}\|<1/2^k$ for all $k\in\mathbb{N}$;
- $\omega^{j(k)}\in N_{\Omega_{NL}(x^{j(k)},-\infty)\cap\Omega_L}(x^{j(k)})$;
- $\|y^k-w^{j(k)}\|<1/2^k$ for all $k\in\mathbb{N}$;

Clearly, these sequences satisfy $\{x^{j(k)}\}\subset\Omega_L$, $\{\omega^{j(k)}\}\subset N_{\Omega_{NL}(x^{j(k)},-\infty)\cap\Omega_L}(x^{j(k)})$, $x^{j(k)}\to x^*$ and $\omega^{j(k)}\to y$. Therefore, $y\in\limsup_{(x,\varepsilon)\to(x^*,0),\,x\in\Omega_L}N_{\Omega_{NL}(x,-\infty)\cap\Omega_L}(x+\varepsilon)$ Now, by LAGP-regularity, $y\in N_{\Omega_{NL}(x^*,-\infty)\cap\Omega_L}(x^*)=L_\Omega(x^*)^\circ$, which allows us to conclude the inclusion $N_\Omega(x^*)\subset L_\Omega(x^*)^\circ$. Using Rockafellar and Wets [36, theorems 6.28(b) and 6.26], we have $L_\Omega(x^*)=L_\Omega(x^*)^{\circ\circ}\subset N_\Omega(x^*)^\circ\subset T_\Omega(x^*)$ as we wanted to prove. □

The following example shows that Abadie's constraint qualification is strictly weaker than CAKKT-regularity and LAGP-regularity.

**Example 7** (Abadie's CQ Implies Neither CAKKT-Regularity Nor LAGP Regularity)**.** Consider $x^* = (0,0)$ and the feasible set given by the inequality constraints

$$g_1(x_1,x_2) = -x_1; \qquad g_2(x_1,x_2) = -x_2 \exp x_2; \qquad g_3(x_1,x_2) = -x_1x_2.$$

The point $x^* = (0,0)$ is feasible and active for all the constraints. By direct calculations

$$\nabla g_1(x_1,x_2) = (-1,0), \qquad \nabla g_2(x_1,x_2) = (0, -\exp x_2 - x_2 \exp x_2), \qquad \text{and} \qquad \nabla g_3(x_1,x_2) = (-x_2,-x_1),$$

for all $x = (x_1,x_2) \in \mathbb{R}^2$. Furthermore, $L_\Omega(x^*)^\circ = \mathbb{R}_- \times \mathbb{R}_-$.

***Abadie's constraint qualification holds at $x^*$.*** This follows directly from $\Omega = \mathbb{R}^2_+$, the form of the gradients of the constraints at $x^* = (0,0)$, and the definition Abadie's CQ.

***CAKKT-regularity does not hold.*** We will show that $K_C(x,r)$ is not outer semicontinuous at $(x^* = (0,0), 0)$. Take $x_1^k := 1/k$, $x_2^k := -1/k$ and define $\mu_1^k := 0$, $\mu_2^k := 0$ and $\mu_3^k := k$. For that choice we obtain that

$$r^k := |\mu_1^k x_1^k| + |\mu_2^k x_2^k \exp(x_2^k)| + |\mu_3^k x_1^k x_2^k| = \frac{k}{k^2} = \frac{1}{k} \to 0$$

and

$$\omega^k := \mu_1^k \nabla g_1(x_1^k, x_2^k) + \mu_2^k \nabla g_2(x_1^k, x_2^k) + \mu_3^k \nabla g_3(x_1^k, x_2^k) = k(1/k, -1/k) = (1,-1).$$

Hence $\omega^k = (1,-1) \in K_C(x^k, r^k)\ \forall\, k \in \mathbb{N}$, $(1,-1)$ belongs to $\limsup_{(x,r)\to(x^*,0)} K_C(x,r)$ but $(1,-1)$ does not belong to $K_C((x^*,0)) = \mathbb{R}_- \times \mathbb{R}_-$. Thus, CAKKT-regularity fails.

***LAGP-regularity does not hold.*** Note that $\Omega_L = \{x = (x_1,x_2)\colon x_1 \geq 0\}$. Define $x_1^k := 1/k$, $x_2^k := -1/k$, $\varepsilon_1^k := 0$, $\varepsilon_1^k := 0$ and multipliers $\mu_1^k := 0$, $\mu_2^k := 0$ and $\mu_3^k := k$. With this choice, we have

$$\omega^k := \mu_1^k \nabla g_1(x_1^k, x_2^k) + \mu_2^k \nabla g_2(x_1^k, x_2^k) + \mu_3^k \nabla g_3(x_1^k, x_2^k) \in N_{\Omega_L \cap \Omega(x^k, -\infty)}(x^k + \varepsilon^k).$$

Clearly, $\omega^k = (1,-1)$ for all $k \in \mathbb{N}$, $(1,-1)$ belongs to $\limsup_{(x,\varepsilon)\to(x^*,0),\, x\in\Omega_L} N_{\Omega_L\cap\Omega(x,-\infty)}(x+\varepsilon)$ and does not belong to $L_\Omega(x^*)^\circ$, so LAGP fails.

### 6.2. Relations with Pseudonormality and Quasinormality

In this section, we will prove that Pseudonormality and Quasinormality do not imply and are not implied by any of the strict CQs defined in the previous section. By Theorem 5, we only need to prove that Pseudonormality and Quasinormality are independent of CAKKT-regularity and LAGP-regularity.

Let us recall the definition of Quasinormality (Hestenes [24], Bertsekas [10]). We say that the Quasinormality Constraint Qualification holds at $x^* \in \Omega$ if whenever $\sum_{j=1}^m \lambda_j \nabla h_j(x^*) + \sum_{j\in J(x^*)} \mu_j \nabla g_j(x^*) = 0$ for some $\lambda \in \mathbb{R}^m$ and $\mu_j \in \mathbb{R}_+$, $j \in J(x^*)$, there is no sequence $x^k \to x^*$ such that for every $k \in \mathbb{N}$, $\lambda_i h_i(x^k) > 0$ when $\lambda_i$ is nonzero and $g_j(x^k) > 0$ when $\mu_j > 0$. Now, if we require the nonexistence of a sequence $x^k \to x^*$ such that $\sum_{i=1}^m \lambda_j h_j(x^k) + \sum_{j\in J(x^*)} \mu_j \nabla g_j(x^k) > 0$ for all $k \in \mathbb{N}$ when $\sum_{j=1}^m \lambda_j \nabla h_j(x^*) + \sum_{j\in J(x^*)} \mu_j \nabla g_j(x^*) = 0$ for some $\lambda \in \mathbb{R}^m$ and $\mu_j \in \mathbb{R}_+$ for every $j \in J(x^*)$, we say that the Pseudonormality Constraint Qualification holds at $x^* \in \Omega$ (Bertsekas [10], Bertsekas and Ozdaglar [11]). Clearly, from the definitions, Pseudonormality is stronger than Quasinormality.

Let us start with the following example, which shows that Pseudonormality implies neither CAKKT-regularity nor LAGP-regularity.

**Example 8** (Pseudonormality Does Not Imply CAKKT-Regularity and Does Not Imply LAGP-Regularity)**.** Consider the feasible set given by the equality and inequality constraints defined by

$$h(x_1,x_2) = x_2 - x_1; \qquad g(x_1,x_2) = x_1 - x_2 \exp x_2.$$

Clearly, $x^* = (0,0)$ is a feasible point and active for both constraints. We also note that:

$$\nabla h(x_1,x_2) = (-1,1) \qquad \text{and} \qquad \nabla g(x_1,x_2) = (1, -\exp x_2 - x_2 \exp x_2), \quad \text{for all } x = (x_1,x_2) \in \mathbb{R}^2.$$

Moreover, we have $L_\Omega(x^*)^\circ = \{\lambda(-1,1) + \mu(1,-1)\colon \lambda \in \mathbb{R}, \mu \in \mathbb{R}_+\} = \mathbb{R}(-1,1)$, that is, $L_\Omega(x^*)^\circ$ is a linear subspace generated by $(-1,1)$.

***Pseudonormality is satisfied at $x^* = (0,0)$.*** First, note that since $\nabla g(x^*) = -\nabla h(x^*) = (-1,1)$, the expression $\mu\nabla g(x^*)+\lambda\nabla h(x^*)=(0,0)$ holds with nonzero $\mu\in\mathbb{R}_+$, $\lambda\in\mathbb{R}$ only if $\mu=\lambda>0$. Assume by contradiction, that there is a sequence $(x_1^k,x_2^k)\to(0,0)$, such that $\lambda h(x_1^k,x_2^k)+\mu g(x_1^k,x_2^k)>0$ for all $k\in\mathbb{N}$. Thus, if $\lambda h(x_1^k,x_2^k)+\mu g(x_1^k,x_2^k)=\mu(x_2^k-x_1^k+x_1^k-x_2^k\exp x_2^k)=\mu(x_2^k-x_2^k\exp x_2^k)>0$, then $x_2>x_2\exp x_2$ for all $k\in\mathbb{N}$, but this is impossible since there is no $x_2\in\mathbb{R}$ such that $x_2>x_2\exp x_2$. Therefore, Pseudonormality holds.

***CAKKT-regularity fails at $x^* = (0,0)$.*** Take $x_1^k := 1/k$, $x_2^k := x_1^k$, $\mu^k := -(1-\exp x_2^k - x_2^k\exp x_2^k)^{-1}$ and $\lambda^k := 2-\mu^k(-\exp x_2^k - x_2^k\exp x_2^k)$. Define

$$\omega^k := \lambda^k(-1,1)+\mu^k(1,-\exp x_2^k - x_2^k\exp x_2^k).$$

We will show that $\omega^k\to(-3,2)$, $r^k := |\lambda^k h(x_1^k,x_2^k)|+|\mu^k g(x_1^k,x_2^k)|\to 0$ and $\omega^k\in K_C(x^k,r^k)$ $\forall k\in\mathbb{N}$. After some calculations, $\omega_2^k=\lambda^k+\mu^k(-\exp x_2^k-x_2^k\exp x_2^k)=2$ and $\omega_1^k=-\lambda^k+\mu^k=-2+\mu^k(1-\exp x_2^k-x_2^k\exp x_2^k)=-3$. Thus, $\lim_{k\to\infty}\omega^k=(-3,2)$. Moreover, $r^k$ converges to zero:

$$r^k=|\lambda^k(x_1^k-x_2^k)|+|\mu^k(x_1^k-x_2^k\exp x_2^k)|=\frac{|x_2^k-x_2^k\exp x_2^k|}{|1-\exp x_2^k-x_2^k\exp x_2^k|}\to 0.$$

Thus, $\omega^k=(-3,2)\in K_C(x^k,r^k)$ $\forall k\in\mathbb{N}$, and hence, $(-3,2)\in\limsup_{(x,r)\to(x^*,0)}K_C(x,r)$ but $(-3,2)$ is not in $L_\Omega(x^*)^\circ=\mathbb{R}(-1,1)$. Thus, CAKKT-regularity does not hold.

***LAGP-regularity is not satisfied at $x^* = (0,0)$.*** First, note that $\Omega_L=\{(x_1,x_2)\in\mathbb{R}^2\colon x_1=x_2\}$. Now, define $x_1^k:=1/k$, $x_2^k:=x_1^k$, $\varepsilon_1^k:=-(x_2^k-x_2^k\exp x_2^k)(1-\exp x_2^k-x_2^k\exp x_2^k)^{-1}$, $\varepsilon_2^k:=\varepsilon_1^k$ and multipliers $\mu^k:=-(1-\exp x_2^k-x_2^k\exp x_2^k)^{-1}$ and $\lambda^k:=2-\mu^k(-\exp x_2^k-x_2^k\exp x_2^k)$. Also, define

$$\omega^k:=\lambda^k(-1,1)+\mu^k(1,-\exp x_2^k-x_2^k\exp x_2^k).$$

Let us show that $\omega^k\in N_{\Omega_{NL}(x^k,\infty)\cap\Omega_L}(x^k+\varepsilon^k)$ for all $k\in\mathbb{N}$. Clearly, $x^k$ and $x^k+\varepsilon^k$ are in $\Omega_L$, $\mu^k\ge 0$, $\varepsilon^k\to(0,0)$ and $\omega^k=(-3,2)$ $\forall k\in\mathbb{N}$. Now, we only need to show that there is no restriction for $\mu^k\ge 0$. Since $x_1^k-x_1^k\exp x_1^k<0$ for $x_1\neq 0$, we have $g(x_1^k,x_2^k)<0$ $(x_1^k=x_2^k)$. So, the multiplier associated with $\nabla g(x_1^k,x_2^k)$ is free, if $g(x_1,x_2^k)+\langle\nabla g(x_1^k,x_2^k),(\varepsilon_1^k,\varepsilon_2^k)\rangle=0$; but, for this choice of $\varepsilon^k=(\varepsilon_1^k,\varepsilon_2^k)$,

$$\begin{aligned} g(x_1,x_2^k)+\langle\nabla g(x_1^k,x_2^k),(\varepsilon_1^k,\varepsilon_2^k)\rangle &= x_2^k-x_2^k\exp x_2^k+\varepsilon_1^k+\varepsilon_2^k(-\exp x_2^k-x_2^k\exp x_2^k)\\ &= x_2^k-x_2^k\exp x_2^k+\varepsilon_1^k(1-\exp x_2^k-x_2^k\exp x_2^k)=0. \end{aligned}$$

Thus, we can choose $\mu^k=-(1-\exp x_2^k-x_2^k\exp x_2^k)^{-1}>0$ as multiplier associated with $\nabla g(x_1^k,x_2^k)$ and, thus, $\omega^k=(-3,2)=\lambda^k(-1,1)+\mu^k(1,-\exp x_2^k-x_2^k\exp x_2^k)\in N_{\Omega_{NL}(x^k,\infty)\cap\Omega_L}(x^k+\varepsilon^k)$. Clearly $(-3,2)=\lim_{k\to\infty}\omega^k$ is in $\limsup_{(x,\varepsilon)\to(x^*,0),\,x\in\Omega_L}N_{\Omega_{NL}(x,\infty)\cap\Omega_L}(x+\varepsilon)$ and $(-3,2)\notin L_\Omega(x^*)^\circ$. Hence, LAGP-regularity fails.

Since Quasinormality is implied by Pseudonormality, from the last example we have that Quasinormality implies neither CAKKT-regularity nor LAGP-regularity.

To prove that CAKKT-regularity and LAGP-regularity are independent of Pseudonormality and Quasinormality, it will be sufficient to show that CAKKT-regularity and LAGP-regularity do not imply Quasinormality. The next example meets this purpose.

**Example 9** (Neither CAKKT-Regularity nor LAGP-Regularity Imply Quasinormality). Consider the feasible set defined by the equality and inequality constraints.

$$h(x_1,x_2)=x_1;\qquad g_1(x_1,x_2)=x_1^3;\qquad g_2(x_1,x_2)=x_1\exp x_2.$$

The point $x^*=(0,0)$ is feasible and active for both constraints. Since, for all $x=(x_1,x_2)\in\mathbb{R}^2$, we have

$$\nabla h(x_1,x_2)=(1,0),\quad \nabla g_1(x_1,x_2)=(3x_1^2,0),\quad\text{and}\quad \nabla g_2(x_1,x_2)=(\exp x_2,x_1\exp x_2),$$

we obtain $L_\Omega(x^*)^\circ=\{\lambda(1,0)+\mu_1(0,0)+\mu_2(1,0),\lambda\in\mathbb{R},\mu_1\ge 0,\mu_2\ge 0\}=\mathbb{R}\times\{0\}$.

***$x^*$ is CAKKT-regular.*** Take $\omega^*\in\limsup_{(x,r)\to(x^*,0)}K_C(x,r)$, so there are sequences $\{x^k\}$, $\{r^k\}$, and $\{\omega^k\}$ with $x^k=(x_1^k,x_2^k)\to x^*=(0,0)$, $\omega^k=(\omega_1^k,\omega_2^k)\to\omega^*$ such that

$$\omega^k=\lambda^k(1,0)+\mu_1^k(3(x_1^k)^2,0)+\mu_2^k(\exp(x_2^k),x_1^k\exp x_2^k)\in K_C(x^k,r^k) \tag{61}$$

and

$$|\lambda^k x_1^k|+|\mu_1^k(x_1^k)^3|+|\mu_2^k x_1^k\exp x_2^k)|\le r^k\to 0, \tag{62}$$

for some scalars $\lambda^k,\mu_1^k,\mu_2^k$ with $\mu_1^k\ge 0$ and $\mu_2^k\ge 0$. From the expressions (61) and (62) we obtain that $|\omega_2^k=\mu_2^k x_1^k\exp x_2^k|\le r^k$ and $\omega_2^k\to 0$. Thus, $\omega^*=\lim_{k\to\infty}\omega^k$ is in $\mathbb{R}\times\{0\}$ and CAKKT-regularity holds.

***$x^*$ is LAGP-regular.*** First, we will calculate $\Omega_L$. Since the only linear constraint is defined by $h$, we have:

$$\Omega_L\left\{x=(x_1,x_2)\in\mathbb{R}^2\colon h(x)=0\right\}=\left\{x=(x_1,x_2)\in\mathbb{R}^2\colon x_1=0\right\}=\{0\}\times\mathbb{R}.$$

Let us show that $N_{\Omega_{NL}(x,-\infty)\cap\Omega_L}(x+\varepsilon)$ is outer semicontinuous at $(x^*,0)$ relative to $\Omega_L\times\mathbb{R}^2$. Take $\omega^*=(\omega_1,\omega_2)\in\limsup N_{\Omega_{NL}(x,-\infty)\cap\Omega_L}(x+\varepsilon)$ relative to $\Omega_L\times\mathbb{R}^2$, so by definition of outer limit, there are sequences $\{x^k\}$, $\{\omega^k\}$, and $\{\varepsilon^k\}$ in $\mathbb{R}^2$ such that $x^k\to x^*$, $\varepsilon^k\to(0,0)$, $\omega^k\to\omega^*$, and

$$x^k\in\Omega_L,\qquad x^k+\varepsilon^k\in\Omega_{NL}(x^k,-\infty)\cap\Omega_L,\qquad \omega^k\in N_{\Omega_{NL}(x^k,-\infty)\cap\Omega_L}(x^k+\varepsilon^k).$$

To show that $\omega^*$ belongs to $N_{\Omega_{NL}(x^*,-\infty)\cap\Omega_L}(x^*+0)=L_\Omega(x^*)^\circ$, we will analyze all the possible cases. Since $x^k\in\Omega_L$ and $x^k+\varepsilon^k\in\Omega_L$ we get $x_1^k=0$, $\varepsilon_1^k=0$, $g_1(x_1^k,x_2^k)=0$ and $g_2(x_1^k,x_2^k)=0$ for $k\in\mathbb{N}$. We also note that for any possible value of $\varepsilon_2^k$, the following expression always holds:

$$\left\langle\nabla g_1(x_1^k,x_2^k),(\varepsilon_1^k,\varepsilon_2^k)\right\rangle=0\qquad\text{and}\qquad\left\langle\nabla g_2(x_1^k,x_2^k),(\varepsilon_1^k,\varepsilon_2^k)\right\rangle=0.$$

To see this, since $x_1^k=\varepsilon_1^k=0$ we have:

$$\left\langle\nabla g_1(x_1^k,x_2^k),(\varepsilon_1^k,\varepsilon_2^k)\right\rangle=\varepsilon_1^k(3(x_1^k)^2)+\varepsilon_2^k(0)=0.(3(x_1^k)^2)+\varepsilon_2^k.0=0$$

and

$$\left\langle\nabla g_2(x_1^k,x_2^k),(\varepsilon_1^k,\varepsilon_2^k)\right\rangle=\varepsilon_1^k(\exp x_2^k)+\varepsilon_2^k(x_1^k\exp x_2^k)=0.\exp x_2^k+\varepsilon_2^k.0=0.$$

Thus, there are $\mu_1^k\geq0$, $\mu_2^k\geq0$ such that

$$\omega^k=\lambda^k\nabla h(x_1^k,x_2^k)+\mu_1^k\nabla g_1(x_1^k,x_2^k)+\mu_2^k\nabla g_2(x_1^k,x_2^k)\in N_{\Omega_{NL}(x^k,-\infty)\cap\Omega_L}(x^k+\varepsilon^k),$$

but since $x_1^k=0$, we have $\nabla h(x_1,x_2)=(1,0)$, $\nabla g_1(x_1^k,x_2^k)=(0,0)$, and $\nabla g_2(x_1^k,x_2^k)=(\exp x_2^k,0)$. Therefore, $\omega_2^k=0$ for all $k\in\mathbb{N}$ and $\omega^*=\lim_{k\to\infty}\omega^k\in\mathbb{R}\times\{0\}=K(x^*)$, as we wanted to show.

***Quasinormality does not hold at $x^*$.*** For every $k\in\mathbb{N}$, define $x_1^k:=1/k$, $x_1^k:=x_2^k$, $\lambda:=0$, $\mu_1:=1$ and $\mu_2:=0$. For these choices, we have $\lambda\nabla h(x^*)+\mu_1\nabla g_1(x^*)+\mu_2\nabla g_2(x^*)=0.(1,0)+1.(0,0)+0.(1,0)=(0,0)$ and $\mu_1 g_1(x_1^k,x_2^k)=(x_1^k)^3>0$ for all $k\in\mathbb{N}$. Thus, Quasinormality fails at $x^*$.

Figure 6 shows the major results obtained in this section. We believe that, up to the present date, this is the most complete landscape of constraint qualifications with algorithmic implications.

By the examples above, we have that neither CAKKT nor LAGP, under Pseudonormality or Quasinormality, imply the KKT conditions. We end this section with a specific example of this kind.

Consider the following optimization problem:

$$\begin{aligned}\text{Minimize }& f(x_1,x_2):=3x_1-2x_2\\ \text{s.t. }& h(x_1,x_2)=x_1-x_2=0,\\ & g(x_1,x_2)=x_1-x_2\exp(x_2)\leq0.\end{aligned}\tag{63}$$

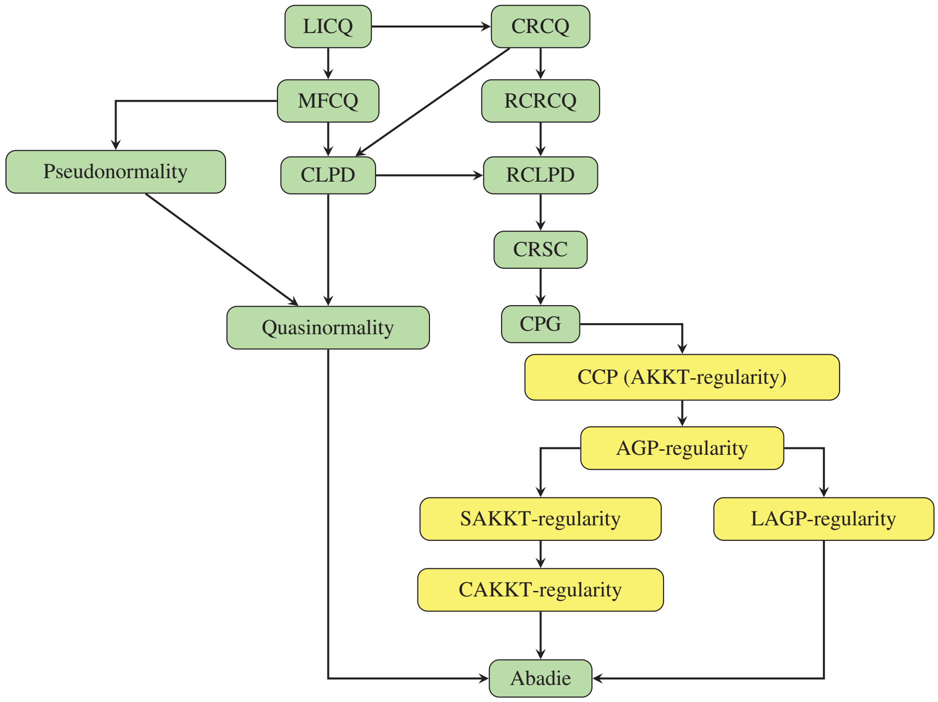
By Example 8, the constraints satisfy Quasinormality at $x^*=(0,0)$ and, thus, Abadie's CQ but neither CAKKT-regularity nor LAGP-regularity. Let us see that both CAKKT and LAGP hold for this objective function.

***CAKKT holds at $x^*=(0,0)$.*** From the Example 8, we have that for $x_1^k:=1/k$, $x_2^k:=x_1^k$, $\mu^k:=-(1-\exp x_2^k-x_2^k\exp x_2^k)^{-1}$ and $\lambda^k:=2-\mu^k(-\exp x_2^k-x_2^k\exp x_2^k)$:

$$\nabla f(x,x^k)+\lambda^k(-1,1)+\mu^k(1,-\exp x_2^k-x_2^k\exp x_2^k)\to(0,0)$$

and $r^k:=|\lambda^k h(x_1^k,x_2^k)|+|\mu^k g(x_1^k,x_2^k)|\to0$. Thus, CAKKT holds.

**Figure 6.** (Color online) An updated landscape of constraint qualifications.




*Note.* Arrows mean strict implications.

***LAGP holds at $x^* = (0,0)$.*** Take $x_1^k := 1/k$ and $x_2^k := x_1^k$ as in Example 8. Note that $(x_1^k, x_2^k)$ is in $\Omega_L = \{(x_1, x_2) \in \mathbb{R}^2: x_1 = x_2\}$. If we define $\varepsilon_1^k := -(x_2^k - x_2^k \exp x_2^k)(1 - \exp x_2^k - x_2^k \exp x_2^k)^{-1}$, $\varepsilon_2^k := \varepsilon_1^k$ and multipliers $\mu^k := -(1 - \exp x_2^k - x_2^k \exp x_2^k)^{-1}$ and $\lambda^k := 2 - \mu^k(-\exp x_2^k - x_2^k \exp x_2^k)$. We have

$$\omega^k := \lambda^k(-1,1) + \mu^k(1, -\exp x_2^k - x_2^k \exp x_2^k) \in N_{\Omega_{NL}(x^k,\infty)\cap\Omega_L}(x^k + \varepsilon^k).$$

Then, by Proposition 1, $P_{\Omega_{NL}(x^k,\infty)\cap\Omega_L}(x^k + \varepsilon^k + \omega^k) = x^k + \varepsilon^k$ Now, since $\omega^k = -\nabla f(x_1^k, x_2^k) = (-3, 2)$, $\forall\, k \in \mathbb{N}$, we conclude, from the nonexpansivity of the projection, that $P_{\Omega_{NL}(x^k,\infty)\cap\Omega_L}(x^k + \varepsilon^k + \omega^k) - x^k \to (0,0)$. Thus, the sequential optimality condition LAGP holds.

The point $x^* = (0,0)$ means nothing for the optimization problem (63). The considered point $x^*$ is neither an optimal solution nor a stationary point. But it can be attained by an algorithm that generates CAKKT points (as an augmented Lagrangian method, for instance) or by an algorithm that generates L-AGP points (like inexact restauration methods). This means that the point $(0,0)$ fulfills any sensible practical test based on CAKKT or on L-AGP (stronger than the test based on AKKT) and the algorithm will accept a point which has no relation with the optimization problem (63). This cannot happen if, instead of the Quasinormality condition, the point satisfies any constraint qualification which implies respectively CAKKT regularity and L-AGP regularity as LICQ, MFCQ, CRSC, CPG, CCP etc.

## 7. Final Remarks

The development of computers in the 20th century made it possible to solve many constrained optimization problems by means of iterative algorithms. The KKT conditions provided a theoretical basis to the definition of suitable stopping criteria for these algorithms. Approximate forms of the KKT conditions are used to declare that an iterate is satisfactory enough for the purposes of practical iterative methods since the fifties, when the first constrained optimization algorithms appeared. However, if an algorithm does not naturally provide Lagrange multipliers approximations, stopping criteria based on gradient projections may be preferred. AKKT, Scaled-AKKT, CAKKT, and SAKKT induce stopping criteria based on the KKT conditions, while AGP and LAGP are sequential optimality conditions that induce stopping criteria based on gradient projections. For the practical point of view, the fact that sequential optimality conditions are satisfied by local minimizers independently of constraint qualifications is very important, since it justifies the decision taken in every optimization software of not testing constraint qualifications at all.

Since sequential optimality conditions are genuine necessary conditions for constrained optimization, the question of their relative strength comes to be relevant. Again, this question is associated with the efficiency of methods: It can be conjectured that the efficiency of a method is linked to the strength of the optimality condition that is guaranteed to hold by the cluster points of the generated sequences. Moreover, the possible nonfulfillment of this conjecture in practical cases could reveal that the analysis of the methods under consideration should rely on alternative theoretical concepts.

Now, the strength analysis of sequential optimality conditions may be direct or indirect. The direct analysis proceeds by straight comparison of the optimality conditions, showing the implications between them and the examples in which one condition holds and other does not at a nonoptimal point. The indirect analysis asks for the constraint qualifications that must be satisfied by a point that fulfills a sequential optimality condition in order to be a KKT point. The interest of the indirect analysis relies on the fact that the constraint qualifications that guarantee that a stationary point (from the point of view of a sequential optimality condition) satisfies KKT are properties of the feasible points of a constrained optimization problem, whose geometrical meaning and consequences are instigating. In other words, this analysis provides the classification of systems of equalities and inequalities from a new point of view, which is completely independent of objective functions.

We believe that future research on strict constraint qualifications associated with sequential optimality conditions will address optimization problems of the form (1) with special characteristics on the function or the constraints (for example, in the presence of complementarity, equilibrium or cone constraints), problems of the form (1) with nonsmooth components, and optimization problems that are not given in the form (1). In the case of complementarity constraints, it is well-known that most standard constrained optimization methods may converge to nonoptimal points from which obvious descent direction emanate, a fact that motivated the definition of many alternative point-based optimality conditions whose sequential stopping-criteria counterpart have not been analyzed yet. This is also the case of bilevel optimization problems. On the other hand, optimization problems that do not attain the form (1) include multiobjective optimization problems, order-value optimization Martínez [29], semidefinite programming, PDE-constrained optimization and many other problems with engineering, economics and industrial applications. Much research on these topics should be expected in the forthcoming years.

## Acknowledgments

The authors would like to thank the referees and the editor for their constructive comments and suggestions that helped to improve the quality and presentation of the paper.

## References

[1] Abadie J (1967) On the Kuhn-Tucker theorem. Abadie J, ed. *Nonlinear Programming* (North-Holland, Amsterdam), 19–36.
[2] Andreani R, Haeser G, Martínez JM (2011) On sequential optimality conditions for smooth constrained optimization. *Optimization* 60(5):627–641.
[3] Andreani R, Martínez JM, Schuverdt ML (2005) On the relation between constant positive linear dependence condition and quasinormality constraint qualification. *J. Optim. Theory Appl.* 125(2):473–485.
[4] Andreani R, Martínez JM, Svaiter BF (2010) A new sequential optimality condition for constrained optimization and algorithmic consequences. *SIAM J. Optim.* 20(6):3533–3554.
[5] Andreani R, Birgin EG, Martínez JM, Schuverdt ML (2007) On augmented lagrangian methods with general lower-level constraints. *SIAM J. Optim.* 18(4):1286.
[6] Andreani R, Haeser G, Schuverdt ML, Silva PJS (2012) A relaxed constant positive linear dependence constraint qualification and applications. *Math. Programming* 135(1-2):255–273.
[7] Andreani R, Martínez JM, Ramos A, Silva PJS (2016) A cone-continuity constraint qualification and algorithmic consequences. *SIAM J. Optim.* 26(1):96–110.
[8] Andreani R, Haeser G, Schuverdt ML, Silva PJS (2012) Two new weak constraint qualifications and applications. *SIAM J. Optim.* 22(3):1109–1135.
[9] Aubin JP, Frankowska H (1990) *Set-Valued Analysis* (Birkhaüser, Boston).
[10] Bertsekas DP (1999) *Nonlinear Programming*, 2nd ed. (Athena Scientific, Belmont, MA).
[11] Bertsekas DP, Ozdaglar A (2002) Pseudonormality and a lagrange multiplier theory for constrained optimization. *J. Optim. Theory Appl.* 114(2):287–343.
[12] Birgin EG, Martínez JM (2014) *Practical Augmented Lagrangian Algorithms for Constrained Optimization*, Fundamental of Algorithms (SIAM Publications, Philadelphia).
[13] Birgin EG, Bueno LF, Martínez JM (2015) Assessing the reliability of general purpose inexact restoration methods. *J. Comput. Appl. Math.* 282:1–1.
[14] Bolte J, Daniilidis A, Lewis AS (2007) The Lojasiewicz inequality for nonsmooth subanalytic functions with applications to subgradient dynamical systems. *SIAM J. Optim.* 17(4):1205–1223.
[15] Borwein JM, Lewis AS (2006) *Convex Analysis and Nonlinear Optimization: Theory and Examples*, CMS Books in Mathematics (Springer, New York).
[16] Bueno LF, Haeser G, Martínez JM (2014) A flexible inexact-restoration method for constrained optimization. *J. Optim. Theory Appl.* 165(1):188–208.

[17] Cartis C, Gould NIM, Toint PL (2014) On the complexity of finding first-order critical points in constrained nonlinear optimization. *Math. Programming A* 144(1):93–106.
[18] Cartis C, Gould NIM, Toint PL (2016) Corrigendum: On the complexity of finding first-order critical points in constrained nonlinear optimization. *Math. Programming A*.161(1—2):611–626.
[19] Chen L, Goldfarb D (2006) Interior-point 2-penalty methods for nonlinear programming with strong global convergence properties. *Math. Programming* 108(1):1–26.
[20] Fischer A, Friedlander A (2010) A new line search inexact restoration approach for nonlinear programming. *Comput. Optim. Appl.* 46(2):333–346.
[21] Gould FJ, Tolle JW (1971) A necessary and sufficient qualification for constrained optimization. *SIAM J. Appl. Math.* 20(2):164–172.
[22] Guignard M (1969) Generalized Kuhn-Tucker conditions for mathematical programming problems in Banach space. *SIAM J. Control* 7(2):232–241.
[23] Haeser G, Schuverdt ML (2011) On approximate KKT condition and its extension to continuous variational inequalities. *J. Optim. Theory Appl.* 149(3):528–539.
[24] Hestenes MR (1975) *Optimization Theory: The Finite Dimensional Case* (John Wiley & Sons, New York).
[25] Janin R (1984) Direction derivative of the marginal function in nonlinear programming. *Math. Programming Stud.* 21:127–138.
[26] Łojasiewicz S (1984) Sur les trajectories du gradient d'une fonction analytique, in Seminari di Geometria 1982–1983, Università di Bologna, Istituto di Geometria, Dipartamento di Matematica 115–117.
[27] Mangasarian OL, Fromovitz S (1967) The Fritz-John necessary optimality conditions in presence of equality and inequality constraints. *J. Math. Anal. Appl.* 17(1):37–47.
[28] Martínez JM (2001) Inexact restoration method with Lagrangian tangent decrease and new merit function for nonlinear programing. *J. Optim. Theory Appl.* 111(1):39–58.
[29] Martínez JM (2012) Generalized order-value optimization. *TOP* 20(1):75–98.
[30] Martínez JM, Pilotta EA (2000) Inexact restoration algorithm for constrained optimization. *J. Optim. Theory Appl.* 104(1):135–163.
[31] Martínez JM, Svaiter BF (2003) A practical optimality condition without constraint qualifications for nonlinear programming. *J. Optim. Theory Appl.* 118(1):117–133.
[32] Minchenko L, Stakhovski S (2011) On relaxed constant rank regularity condition in mathematical programming. *Optimization* 60(4): 429–440.
[33] Mordukhovich BS (1998) *Variational Analysis and Generalized Differentiation, I. Basis Theory* (Springer, Berlin).
[34] Nocedal J, Wright S (2006) *Numerical Optimization*, 2nd ed. (Springer , New York).
[35] Qi L, Wei Z (2000) On the constant positive linear dependence condition and its application to SQP methods. *SIAM J. Optim.* 10(4): 963–981.
[36] Rockafellar RT, Wets R (1998) *Variational Analysis* (Springer, Berlin).
[37] Wächter A, Biegler LT (2006) On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Math. Programming* 106(1):25–57.