SIAM J. OPTIMIZATION
Vol. 1, No. 1, pp. 000–000, ????

TILT STABILITY OF A LOCAL MINIMUM *

R. A. POLIQUIN† AND R. T. ROCKAFELLAR‡

Abstract. The behavior of a minimizing point when an objective function is tilted by adding a small linear term is studied from the perspective of second-order conditions for local optimality. The classical condition of a positive-definite Hessian in smooth problems without constraints is found to have an exact counterpart much more broadly in the positivity of a certain generalized Hessian mapping. This fully characterizes the case where tilt perturbations cause the minimizing point to shift in a lipschitzian manner.

Key words. Variational analysis, second-order optimality, tilt stability, perturbations, sensitivity, prox-regularity, amenable functions, generalized Hessians, coderivatives, monotone mappings.

AMS subject classifications. 49A52, 58C06, 58C20; Secondary 90C30

1. Introduction.

Much of the theory of optimality conditions has proceeded from the fact that, for a $\mathcal{C}^2$ function f on $I\!R^n$, a necessary condition for a local minimum at a point $\bar{x}$ is the vanishing of the gradient $\nabla f(\bar{x})$ and the positive-semidefiniteness of the Hessian $\nabla^2 f(\bar{x})$, whereas a sufficient condition is the vanishing of $\nabla f(\bar{x})$ and the positive-definiteness of $\nabla^2 f(\bar{x})$. Analogous patterns have been found for problems with $\mathcal{C}^2$ constraints and also for problems exhibiting nonsmoothness. The goal has been to come up with necessary conditions that become sufficient with only a slight degree of strengthening. But is this the best paradigm for modern purposes?

Optimization no longer revolves around making lists of solution candidates to be checked out one by one, if it ever did. The role of optimality conditions is seen rather in the justification of numerical algorithms, in particular their stopping criteria, convergence properties and robustness. From that angle, the goal of theory could be different. Instead of focusing on the threshold between necessity and sufficiency, one might more profitably try to characterize the stronger manifestations of optimality that support computational work. Indeed, this idea has motivated much of the effort that has gone into parametric optimization—the study of how local solutions to a problem may react to shifts in data.

Here we take up this theme more narrowly, but in other respects more broadly than before. We investigate just one, fundamental mode of perturbation, but do so in a very general setting.

Any problem of optimization in n variables can be expressed as one of minimizing a function f over all of $I\!R^n$, as long as f is allowed to have values in the extended-real-line $\overline{I\!R}$ as a means of representing constraints through infinite penalties. If the aim is to minimize a function f_0 subject to a system of constraints, this is captured by defining $f(x)$ to be $f_0(x)$ when x belongs to the feasible set C, but letting $f(x) = \infty$ when

*Received by the editors ?; accepted by the editors Month, x, xxxx..

†Dept. of Math. Sciences, Univ. of Alberta, Edmonton Alberta, Canada T6G 2G1 . Work supported by the Natural Sciences and Engineering Research Council of Canada under grant OGP41983. This paper was written in part while the first author was visiting the Université de Perpignan; many thanks to D. Azé, J.-N. Corvellec, and the department of Mathematics for their generous hospitality.(rene.poliquin@math.ualberta.ca).

‡Dept. of Math., Univ. of Washington, Seattle, WA 98195 USA . Work supported by the National Science Foundation under grant DMS–9500957. (rtr@math.washington.edu).

$x \notin C$. Keeping for now to such a level of abstraction, with the idea of introducing various details of problem structure later, we are able to formulate succinctly the property we propose to analyze.

DEFINITION 1.1. *A point $\bar{x}$ will be said to give a tilt-stable local minimum of the function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ if $f(\bar{x})$ is finite and there exists $\delta > 0$ such that the mapping*

$$M : v \;\mapsto\; \operatorname*{argmin}_{|x-\bar{x}|\leq\delta} \{\, f(x) - f(\bar{x}) - \langle v, x - \bar{x}\rangle \,\},$$

is single-valued and lipschitzian on some neighborhood of $v = 0$, with $M(0) = \bar{x}$.

The lipschitzian requirement means the existence of a constant $\kappa \geq 0$ for which $|M(v'') - M(v')| \leq \kappa|v'' - v'|$ when $|v'|$ and $|v''|$ are small enough. The subtraction from f of the affine function $l(x) = f(\bar{x}) + \langle v,\, x - \bar{x}\rangle$, which agrees with f at $\bar{x}$, serves to tilt the objective in one direction or another, and the requirement is that such a perturbation, if sufficiently small, should not affect the local solution disproportionately to the size of v or threaten its uniqueness. The connection between this notion of tilt stability and the classical picture of local optimality is as follows.

PROPOSITION 1.2. *For a $\mathcal{C}^2$ function f having $\nabla f(\bar{x}) = 0$, the point $\bar{x}$ gives a tilt-stable local minimum of f if and only if $\nabla^2 f(\bar{x})$ is positive-definite.*

Proof. Sufficiency. We have $\nabla^2 f(x)$ positive-definite for all x in a convex neighborhood X of $\bar{x}$. In particular, f is strictly convex on X, so that for $\tilde{x} \in X$ and $\tilde{v} = \nabla f(\tilde{x})$ we have

$$f(x) > f(\tilde{x}) + \langle \tilde{v},\, x - \tilde{x}\rangle \ \text{ for all } x \in X,\ x \neq \tilde{x}.$$

Then for $\delta > 0$ small enough the mapping M in Definition 1.1 is single-valued on $\nabla f(X)$ with $M(0) = \bar{x}$, and it agrees on that set with the inverse of the restriction of ∇f to X. The standard implicit mapping theorem tells us that M is $\mathcal{C}^1$ around 0. Hence M is lipschitzian on some neighborhood of 0.

Necessity. From $M(0) = \bar{x}$ we have $\nabla^2 f(\bar{x})$ positive-semidefinite. Suppose that $\nabla^2 f(\bar{x})$ is not positive-definite and thus is singular. There is a vector $\bar{w} \notin \{\, \nabla^2 f(\bar{x})w \mid w \in \mathbb{R}^n \,\}$. For $\tau > 0$ so small that $\tau\bar{w}$ lies in a neighborhood where M is single-valued and lipschitzian with constant κ, we have $|M(\tau\bar{w}) - M(0)| \leq \kappa\tau|\bar{w}|$. Let $x^\nu = M(\tau^\nu\bar{w})$ for any sequence of values $\tau^\nu \downarrow 0$ (with superscript $\nu \to \infty$). The sequence of vectors $w^\nu = [x^\nu - \bar{x}]/\tau^\nu$ is bounded in norm by $\kappa|\bar{w}|$, so, by passing to a subsequence if necessary, we can assume that it converges to some $\tilde{w}$. Because $\nabla f(x^\nu) = \tau^\nu\bar{w}$ and $\nabla f(\bar{x}) = 0$, we have $\bar{w} = [\nabla f(\bar{x} + \tau^\nu w^\nu) - \nabla f(\bar{x})]/\tau^\nu \to \nabla^2 f(\bar{x})\tilde{w}$, hence $\bar{w} = \nabla^2 f(\bar{x})\tilde{w}$, which is impossible. □

Can some analogous characterization of tilt stability be given for functions f of a wider class? This question has not previously been asked. In looking for an answer in the absence of f being differentiable, we must appeal to optimality conditions based on subgradients.

For any function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ and any point $\bar{x}$ with $f(\bar{x})$ finite, a vector $\bar{v}$ is a *proximal subgradient* of f at $\bar{x}$ if there exist $\delta > 0$ and $r \geq 0$ such that

$$f(x) \geq f(\bar{x}) + \langle \bar{v},\, x - \bar{x}\rangle - \frac{r}{2}|x - \bar{x}|^2 \ \text{ when } |x - \bar{x}| \leq \delta.$$

It is a *subgradient* (in general), written $\bar{v} \in \partial f(\bar{x})$, if there is a sequence of points $x^\nu \to \bar{x}$ with $f(x^\nu) \to f(\bar{x})$ at which there exist proximal subgradients v^ν such that $v^\nu \to \bar{v}$.

When f is convex, $\partial f(\bar{x})$ is the usual subgradient set of convex analysis. When f is smooth, $\partial f(\bar{x})$ is the singleton $\{\nabla f(\bar{x})\}$. For the indicator function δ_C associated with a set C (which vanishes on C but is ∞ outside of C), the subgradients v at $\bar{x}$ are the *normal vectors* to C at $\bar{x}$; the set $\partial \delta_C(\bar{x})$ is the *normal cone* to C at $\bar{x}$, denoted by $N_C(\bar{x})$. The book [1] provides details on this topic and other aspects of variational analysis that will be important in what follows.

Two properties of f that relate to the set-valued subgradient mapping $x \mapsto \partial f(x)$ and its graph

$$\operatorname{gph} \partial f \ = \ \{ \, (x,v) \, | \, v \in \partial f(x) \, \}$$

will be important. We say that f is *subdifferentially continuous* at $\bar{x}$ for a vector $\bar{v} \in \partial f(\bar{x})$ if the function $(x,v) \mapsto f(x)$ is continuous relative to $\operatorname{gph} \partial f$ at $(\bar{x},\bar{v})$, or in other words, if for every $\delta > 0$ there exists $\epsilon > 0$ such that $|f(x) - f(\bar{x})| < \delta$ when $v \in \partial f(x)$ with $|x - \bar{x}| < \epsilon$ and $|v - \bar{v}| < \epsilon$. On the other hand, f is *prox-regular* at $\bar{x}$ for $\bar{v}$ if its epigraph is closed relative to a neighborhood of $(\bar{x}, f(\bar{x}))$ and there exist $\epsilon > 0$ and $r > 0$ such that

$$\begin{aligned} f(x') \ > \ f(x) + \langle v, \, x' - x \rangle - \frac{r}{2}|x' - x|^2 \ \text{ for } \ x' \neq x \ \text{ when} \\ |x' - \bar{x}| < \epsilon, \ |x - \bar{x}| < \epsilon, \ |f(x) - f(\bar{x})| < \epsilon, \ |v - \bar{v}| < \epsilon, \ v \in \partial f(x). \end{aligned}$$

Prox-regularity thus requires, in a certain uniform local sense with respect to $\bar{x}$ and $\bar{v}$, that all subgradients $v \in \partial f(x)$ actually be proximal subgradients.

The concepts of prox-regularity and subdifferential continuity were introduced in [2] and studied further in [3] and [4]. It was shown in those works that the class of functions exhibiting these properties is very large. It includes, in addition to all $\mathcal{C}^2$ functions f, all lsc (lower semicontinuous) proper convex functions f, as well as many of the functions that typically might be encountered in finite-dimensional optimization. For instance, if $f = f_0 + \delta_C$ for a function f_0 that is $\mathcal{C}^2$, or expressible as a max of finitely many $\mathcal{C}^2$ functions, and a set C specified by a finite family of equality and/or inequality constraints on $\mathcal{C}^2$ functions, then at any point $\bar{x}$ where the constraint system satisfies the Mangasarian-Fromovitz constraint qualification the function f is both prox-regular and subdifferentially continuous for every $\bar{v} \in \partial f(\bar{x})$.

More generally, the class of prox-regular, subdifferentially continuous functions includes all *strongly amenable* functions, which are defined as being representable locally by the composition of an lsc proper convex function with a $\mathcal{C}^2$ mapping under a constraint qualification on the domain of the convex function and the range of the linearization of the mapping; cf. [5] and also [6]–[8]. Such functions are omnipresent in applications, and they have a basic role both theoretically and computationally. For more on the use of composite function formats in optimization, see [4], [9]–[19].

The first-order condition $0 \in \partial f(\bar{x})$ is always necessary for f to have a local minimum at $\bar{x}$. Second-order conditions have been developed in various forms, but here we turn to a new form in terms of the generalized Hessian mapping introduced by Mordukhovich [20], [21]. For any point $\bar{x}$ and any subgradient $\bar{v} \in \partial f(\bar{x})$, define $\partial^2 f(\bar{x} \,|\, \bar{v}) : I\!R^n \rightrightarrows I\!R^n$ by

$$\partial^2 f(\bar{x} \,|\, \bar{v}) : w \ \mapsto \ \{ \, z \, | \, (z, -w) \in N_{\operatorname{gph} \partial f}(\bar{x}, \bar{v}) \, \}.$$

When f is $\mathcal{C}^2$ and $\bar{v} = \nabla f(\bar{x})$, $\partial^2 f(\bar{x} \,|\, \bar{v})$ reduces to the linear mapping $w \mapsto \nabla^2 f(\bar{x}) w$.

In situations where f is given in terms of other functions, it is often possible to draw upon the growing calculus of "coderivatives" to compute the generalized Hessian

$\partial^2 f(\bar{x}\,|\,\bar{v})$. For any set-valued mapping $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$, the *coderivative* at $\bar{x}$ for an element $\bar{v} \in T(\bar{x})$ is the mapping $D^*T(\bar{x}\,|\,\bar{v}) : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ whose graph is obtained from the normal cone $N_{\operatorname{gph} T}(\bar{x}, \bar{v})$ by

$$z \in D^*T\big(\bar{x}\,|\,\bar{v}\big)(w) \iff (z, -w) \in N_{\operatorname{gph} T}(\bar{x}, \bar{v}).$$

In such terms, one has $\partial^2 f(\bar{x}\,|\,\bar{v}) = D^*[\partial f](\bar{x}|\bar{v})$. For more on coderivatives, their history and calculus see [1], [22] and [23].

Our main result can now be stated. It will be proved in Section 3, following the development of a monotonicity property in Section 2. Application to the constrained minimization of a smooth function will be made in Section 4.

THEOREM 1.3. *For a function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ having $0 \in \partial f(\bar{x})$ and such that f is both prox-regular and subdifferentially continuous at $\bar{x}$ for $\bar{v} = 0$, the following are equivalent and imply the existence of $\delta > 0$ such that, for all v in some neighborhood of 0, the mapping M in Definition 1.1 has $M(v)$ as the unique $x \in (\partial f)^{-1}(v)$ with $|x - \bar{x}| \le \delta$.*

(a) The point $\bar{x}$ gives a tilt-stable local minimum of f.

(b) The mapping $\partial^2 f(\bar{x}\,|\,0)$ is positive-definite in the sense that

$$\langle z, w\rangle > 0 \quad \textit{whenever} \quad z \in \partial^2 f(\bar{x}|0)(w),\ w \neq 0.$$

(c) There exist neighborhoods X of $\bar{x}$ and V of 0 such that the mapping $x \in X \mapsto \partial f(x) \cap V$ is strongly monotone and, locally around $(\bar{x}, 0)$, is maximal monotone.

(d) There is a proper, lsc, strongly convex function h on $\mathbb{R}^n$ along with neighborhoods X of $\bar{x}$ and V of 0 such that h is finite on X with $h(\bar{x}) = f(\bar{x})$ and $0 \in \partial h(\bar{x})$, and furthermore $\operatorname{gph} \partial f \cap \big[X \times V\big] = \operatorname{gph} \partial h \cap \big[X \times V\big]$.

To understand (c), recall that a (generally set-valued) mapping $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is *monotone* if $\langle v' - v,\, x' - v\rangle \ge 0$ whenever $v \in T(x)$ and $v' \in T(x')$. It is *strongly* monotone when $T - \lambda I$ is monotone for some $\lambda > 0$. It is maximal monotone, locally around $(\bar{x}, \bar{v})$, if there exist neighborhoods X_0 of $\bar{x}$ and V_0 of $\bar{v}$ such that every monotone mapping $T' : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ with $\operatorname{gph} T' \supset \operatorname{gph} T \cap [X_0 \times V_0]$ has $\operatorname{gph} T' \cap [X_0 \times V_0] = \operatorname{gph} T \cap [X_0 \times V_0]$.

Observe in (d) that h need not agree with f on a neighborhood of $\bar{x}$. In contrast to the smooth case, tilt stability and the positive-definite generalized Hessian do not entail the local convexity of f around $\bar{x}$. An example is furnished by the function $f(x) = |1 - x^2|$ on $\mathbb{R}$. This has a tilt-stable minimum at $\bar{x} = 1$ and at $\bar{x} = -1$, although it is concave on the interval $(-1, 1)$.

2. Coderivatives of Monotone Mappings.

The argument for our characterization of tilt stability will rely on certain monotonicity properties of the subgradient mappings associated with prox-regular functions. For this purpose we must first develop a fact about coderivatives in the presence of monotonicity.

THEOREM 2.1. *If a mapping $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is maximal monotone, one has for every pair $(\bar{x}, \bar{v}) \in \operatorname{gph} T$ that*

$$\langle x', v'\rangle \ge 0 \ \textit{ when } \ v' \in D^*T(\bar{x}\,|\,\bar{v})(x').$$

The verification of this fact will utilize the following connection between monotone mappings and nonexpansive mappings. A mapping $S : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is *nonexpansive* if

$w_i \in S(z_i)$ for $i = 0, 1$ implies $|w_1 - w_0| \le |z_0 - z_1|$. On the set $\operatorname{dom} S = \{ z \mid S(z) \ne \emptyset \}$, such a mapping must actually be single-valued.

PROPOSITION 2.2. *The one-to-one linear transformation $J : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R}^n$ with $J(x, v) = (v + x,\ v - x)$ induces a one-to-one correspondence between mappings $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ and mappings $S : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ through*

$$\operatorname{gph} S = J(\operatorname{gph} T), \qquad \operatorname{gph} T = J^{-1}(\operatorname{gph} S),$$

in which T is monotone if and only if S is nonexpansive. Moreover, T is maximal monotone if and only if $\operatorname{dom} S = \mathbb{R}^n$*, or in other words, S is a single-valued mapping from all of $\mathbb{R}^n$ into itself that is globally lipschitzian with constant* 1.

Proof. See [1], 12.11 and 12.12. □

Also to be used is the well known scalarization formula for coderivatives of locally lipschitzian mappings.

PROPOSITION 2.3. *For a single-valued, locally lipschitzian mapping $S : \mathbb{R}^n \to \mathbb{R}^m$, any pair $(\bar{z}, \bar{w}) \in \operatorname{gph} S$ and any vector $y \in \mathbb{R}^m$, one has $D^*S(\bar{z} \mid \bar{w})(y) = \partial(yS)(\bar{z})$ for the locally lipschitzian function $yS : \mathbb{R}^n \to \mathbb{R}$ given by $(yS)(z) = \langle y, S(z) \rangle$.*

Proof. See [24]; or [1], 9.24(b). □

Proof of Theorem 2.1. Let S correspond to T as in Proposition 2.2, with the pair $(\bar{z}, \bar{w}) \in \operatorname{gph} S$ corresponding to the pair $(\bar{x}, \bar{v}) \in \operatorname{gph} T$. The linear transformation between the graph spaces induces a correspondence between normal cones to the graphs in question, which comes out as the property that

$$(v', -x') \in N_{\operatorname{gph} T}(\bar{x}, \bar{v}) \iff (v' - x', -v' - x') \in N_{\operatorname{gph} S}(\bar{z}, \bar{w}),$$

or equivalently

$$v' \in D^*T(\bar{x} \mid \bar{v})(x') \iff v' - x' \in D^*S(\bar{z} \mid \bar{w})(v' + x').$$

By Proposition 2.2, S is globally lipschitzian with constant 1. Therefore, we have for any y that the mapping yS is globally lipschitzian with constant $|y|$, so that for all $u \in \partial(yS)(\bar{z})$ we have $|u| \le |y|$. Applying Proposition 2.3, we see that $|u| \le |y|$ whenever $u \in D^*S(\bar{z} \mid \bar{w})(y)$. Applying this to $u = v' - x'$ and $y = v' + x'$ in the case of an arbitrary pair (x', v') with $v' \in D^*T(\bar{x} \mid \bar{v})(x')$, we obtain $|v' - x'| \le |v' + x'|$. Hence $0 \le |v' + x'|^2 - |v' - x'|^2 = 4\langle v', x' \rangle$, and we conclude that $\langle v', x' \rangle \ge 0$, as claimed. □

3. Proof of the Main Result.

In proving Theorem 1.3, we can focus on the case where $\bar{x} = 0$ and $f(\bar{x}) = 0$. These assumptions entail no real restriction and merely constitute a shift of variables. We may then assume also without loss of generality that there exists $r > 0$ with

$$f(x) > -\frac{r}{2}|x|^2 \text{ for all } x \ne 0. \tag{3.1}$$

Locally, this inequality comes out of the prox-regularity of f and the definition of a proximal subgradient. It is made into a global inequality by adding to f the indicator of a sufficiently small ball centered at 0, which is an operation having no effect on the property of tilt stability of f at $\bar{x} = 0$.

Recall that a mapping $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is *hypomonotone* around a pair $(x, v) \in \operatorname{gph} T$ if $T + \alpha I$ is monotone near (x, v) for some $\alpha \ge 0$; cf. [25].

Lemma 3.1. *Let* $T_i : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *for* $i = 1, 2$, *with* $(0,0) \in \operatorname{gph} T_1 \subset \operatorname{gph} T_2$. *Suppose* T_1 *is locally maximal monotone around* $(0,0)$, *while* T_2 *is hypomonotone around* $(0,0)$. *Then the graphs of* T_1 *and* T_2 *must coincide locally around* $(0,0)$.

Proof. Let $\overline{T}_1$ be a maximal monotone extension of T_1, as always exists (cf.[1], 12.6). By hypothesis, the graphs of $\overline{T}_1$ and T_1 agree in some neighborhood of $(0,0)$. The local hypomonotonicity of T_2 means the existence of some $\alpha \geq 0$ such that $T_2 + \alpha I$ is monotone around $(0,0)$. Because the sum of maximal monotone mappings is maximal monotone when the domain of one meets the interior of the domain of the other, cf. [26, Theorem 1], we have $\overline{T}_1 + \alpha I$ maximal monotone. Hence locally, $T_1 + \alpha I$ is maximal monotone around $(0,0)$. Since $\operatorname{gph}[T_1 + \alpha I] \subset \operatorname{gph}[T_2 + \alpha I]$, it follows that $T_1 + \alpha I = T_2 + \alpha I$ locally, and the same must therefore be true for T_1 and T_2 themselves. □

A key element in the derivation of Theorem 1.3 will be the fact that the subgradient mapping of a prox-regular function has a "localization" T that is hypomonotone; cf. [2, Theorem 3.2]. A *localization* of ∂f around $(\bar{x}, \bar{v}) \in \operatorname{gph} \partial f$ is a mapping whose graph is obtained by intersecting $\operatorname{gph} \partial f$ with a neighborhood of $(\bar{x}, \bar{v})$.

Throughout the proof, "proto-derivatives" and "second-order epi-derivatives" will be employed. A set-valued mapping $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is *proto-differentiable* at a point x for an element $v \in T(x)$, as defined in [27], if the set-valued mappings

$$\Delta_t T(x \,|\, v) : \xi \;\mapsto\; \big[T(x + t\xi) - v\big]/t \quad \text{for} \quad t > 0$$

graph-converge as $t \downarrow 0$ (i.e., one has set convergence of the graphs). If so, the limit mapping, denoted by $DT(x \,|\, v)$, is the *proto-derivative* of T at x for v. On the other hand, a function f is *twice epi-differentiable* at x for a subgradient $v \in \partial f(x)$, as defined in [10], if the second-order difference quotient functions

$$\Delta_t^2 f(x \,|\, v) : \xi \;\mapsto\; \big[f(x + t\xi) - f(x) - t\langle v, \xi\rangle\big]/\tfrac{1}{2}t^2 \quad \text{for} \quad t > 0,$$

epi-converge to a proper function as $t \downarrow 0$ (i.e., one has set convergence of the epigraphs). The epi-limit is then the *second epi-derivative* function $d^2 f(x \,|\, v) : \mathbb{R}^n \to \overline{\mathbb{R}}$.

Proof that (a) implies (d). Let $X = \{\, x \,|\, |x - \bar{x}| \leq \delta \,\}$ (where δ is given in Def. 1.1) and $g := (f + \delta_X)^*$, i.e.,

$$g(v) = \max_{x \in X} \big\{ \langle x, v\rangle - f(x) \big\}.$$

This function g is proper, lsc, convex and finite. Let

$$G(v) := \operatorname*{argmax}_{x \in X} \; \{\, \langle x, v\rangle - f(x) \,\}.$$

Under our assumptions, G is single-valued and lipschitzian on V, and one can easily verify that $G(v) \in \partial g(v)$. Moreover, the mapping G is monotone on V. Indeed, if $x_i \in G(v_i)$ with $v_i \in V$, then

$$\begin{aligned} \langle x_1 - x_2, v_1 - v_2\rangle &= \langle x_1, v_1\rangle - \langle x_2, v_1\rangle - \langle x_1, v_2\rangle + \langle x_2, v_2\rangle \\ &= g(v_1) + f(x_1) - \langle x_2, v_1\rangle - \langle x_1, v_2\rangle + g(v_2) + f(x_2) \\ &= \big[g(v_1) - \langle x_2, v_1\rangle + f(x_2)\big] + \big[g(v_2) - \langle x_1, v_2\rangle + f(x_1)\big] \geq 0. \end{aligned}$$

It follows that $\partial g(v) = G(v)$ for all $v \in V$, inasmuch as ∂g is a monotone mapping, whereas G is single-valued continuous and monotone on V, hence maximal monotone relative to V (cf. [1]; 12.7, 12.48).

Therefore, g is a convex function that is differentiable on the open neighborhood V of 0, with ∇g lipschitzian there and satisfying $\nabla g(0) = 0$ (because $G(0) = 0$), as well as $\nabla g(v) \in \operatorname{int} X$ for $v \in V$, and

$$\nabla g(v) \in [\partial f]^{-1}(v) \text{ for } v \in V. \tag{3.2}$$

Also, because $\nabla g(v)$ gives the argmax in the conjugate formula for g, we have

$$f(x) = \langle v, x\rangle - g(v) \text{ when } v \in V,\ x = \nabla g(v). \tag{3.3}$$

Denote by h the function conjugate to g; thus, $h(x) = \sup_{v\in \mathbb{R}^n}\{\, \langle x, v\rangle - g(x)\,\}$.

The mapping $\partial g : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is maximal monotone, and its inverse $[\partial g]^{-1}$, which is ∂h, is maximal monotone as well; cf. [28]. Note that $0 \in \partial h(0)$ and that $h(0) = 0$. Let T_0 be the localization of ∂h obtained by inverting ∂g relative to V; thus, T_0^{-1} is ∇g on V but empty-valued outside of V. According to (3.3), the function $(x, v) \mapsto f(x)$ is continuous on the graph of T_0. Since $\operatorname{gph} T_0 \subset \operatorname{gph} \partial f$ by (3.2), any localization T of ∂f around $(0, 0)$ must include a localization of T_0, which in turn is locally maximal monotone around $(0, 0)$. The prox-regularity of f allows us to choose the localization T to be hypomonotone; cf. [2, Theorem 3.2]. We conclude from Lemma 3.1 that the graphs of T and T_0 must agree around $(0, 0)$.

Let $\kappa > 0$ be a lipschitzian constant for ∇g on V. We will now demonstrate that T_0, or at least some localization of T_0 around $(0, 0)$, is strongly monotone with constant κ^{-1}, so that T must have this property too. This will conclude the proof of (a)⇒(d). The tool will be [2, Prop. 5.7] as applied to h. It gives the criterion that at each point $(x, v) \in \operatorname{gph} T_0$ in some neighborhood of $(0, 0)$ where the graph of the proto-derivative of T_0 is an n-dimensional subspace of $\mathbb{R}^{2n}$, the proto-derivative is strongly monotone with constant κ^{-1}.

The points $(x, v) \in \operatorname{gph} T_0$ where the graph of the proto-derivative is an n-dimensional subspace are precisely the points (v, x) where the graph of the proto-derivative of T_0^{-1} is an n-dimensional subspace. We now argue that at such a point v the function g has a symmetric Hessian matrix $\nabla^2 g(v)$. To see this, recall that the function ∇g is lipschitzian on V. Hence at points $v \in V$ where the proto-derivative of the mapping T_0^{-1} exists, the domain of the proto-derivative is the whole space (because the set-valued mapping whose graph is given by $\limsup_{t\downarrow 0} \big[\operatorname{gph} \nabla g - \big(v, \nabla g(v)\big)\big]/t$ has the whole space as its domain). Therefore, the second-order epi-derivative of the function g at the point $\big(v, \nabla g(v)\big)$ has the whole space as its domain, which means that it is purely quadratic, and that g has a "second-order expansion" at v; cf. [3, Theorem 3.1]. A convex function (or more generally any lower-C^2 function) with the above properties at the point v is necessarily twice differentiable there and has a symmetric Hessian matrix $\nabla^2 g(v)$; cf. [3, Theorem 3.2].

Based on the above observation, [2, Prop. 5.7] therefore gives the criterion that for each pair (v, x) such that $v \in V$, $x = \nabla g(v)$, and the Hessian $\nabla^2 g(v)$ exists, the inverse of the linear transformation $w \mapsto \nabla^2 g(v)w$ is strongly monotone with constant κ^{-1}. Because κ is a lipschitzian constant for ∇g, the norm of the Hessian matrix $\nabla^2 g(v)$ is bounded by κ where the Hessian exists. Then

$$\langle z, w\rangle \geq \kappa^{-1}|z|^2 \text{ when } z = \nabla^2 g(v)w, \tag{3.4}$$

as can be seen for instance through a diagonalization based on the symmetry of the Hessian. The inequality (3.4) means that the (possibly multivalued) inverse of the mapping $w \mapsto \nabla^2 g(v)w$ is strongly monotone with constant κ^{-1}, as required.

Proof that (d) implies (c). This implication is obvious.

Proof of (c) implies (b). Let T be the localization of ∂f around $(0,0)$ that is strongly monotone with constant η in addition to being maximal monotone locally. Let S be a maximal monotone extension of the monotone mapping $T - \eta I$. The mapping $S + \eta I$ is maximal monotone (again by [26, Theorem 1])and has a localization around $(0,0)$ that includes T. The local max monotonicity of T implies that the graphs of T and $S + \eta I$ must actually agree around $(0,0)$. Therefore,

$$\partial^2 f(0|0) = D^*T(0|0) = D^*S(0|0) + \eta I. \tag{3.5}$$

According to Theorem 2.1 we have $\langle x', v' \rangle \geq 0$ when $v' \in D^*S(0|0)(x')$. It follows then from (3.5) that $\langle x', v' \rangle \geq \eta |x'|^2$ when $v' \in \partial^2 f(0|0)(x')$. This yields (d), as needed.

Proof of (b) implies (a). We first show that a localization of ∂f around $(0,0)$ is strongly monotone as well as locally max monotone; this shows that $\big(\partial f\big)^{-1}$ is locally single-valued and lipschitzian near 0.

Choose a localization T of ∂f around $(0,0)$ that is not only hypomonotone but such that $\operatorname{gph} T$ is also a graphically lipschitzian manifold of dimension n around $(0,0)$. The fact that $\operatorname{gph} T$ can be chosen as such a manifold was noted in [2, Theorem 4.7]. Graphically lipschitzian manifolds were introduced in [29].

Because $\operatorname{gph} T$ is a lipschitzian manifold we know that $\operatorname{gph} T$ has a linear tangent space at almost every pair $(x,v) \in \operatorname{gph} T$ in some neighborhood of $(\bar{x}, \bar{v})$; cf. [29]. Moreover, at such points (x,v), the graph of $DT(x \,|\, v)$ is an n-dimensional subspace of $I\!R^n \times I\!R^n$, again cf. [29]. Consequently, we define $R : I\!R^n \rightrightarrows I\!R^m$ to be the set-valued mapping whose graph is given by the limsup of $\operatorname{gph} DT(x \,|\, v)$ over the set of points $(x,v) \in \operatorname{gph} T$ converging to $(0,0)$ and such that $\operatorname{gph} DT(x \,|\, v)$ is an n-dimensional subspace of $I\!R^n \times I\!R^n$.

Claim. $\operatorname{gph} R \subset \operatorname{gph} \partial^2 f(0 \,|\, 0)$.

Justification. Recall that a function $h : I\!R^n \to \overline{I\!R}$ is a *generalized quadratic function* if it is expressible in the form

$$h(x) = \begin{cases} \frac{1}{2}\langle x, Qx \rangle & \text{if } x \in N, \\ \infty & \text{if } x \notin N, \end{cases}$$

where N is a linear subspace of $I\!R^n$ and Q is a symmetric $n \times n$ matrix. According to [3, Theorem 3.9], the second-order epi-derivative is generalized quadratic at each point $(x,v) \in \operatorname{gph} T$ near $(0,0)$ where $\operatorname{gph} DT(x \,|\, v)$ is an n-dimensional subspace (simply apply [3, Theorem 3.9] to the function $\tilde{f}(x) = f(x) + (r/2)|x|^2$ (r given by (3.1)); this function meets all the requirements of that theorem). We know that the proto-derivative of the mapping T at (x,v) gives the subgradient mapping of the second-order epi-derivative, cf. [2, Theorem 6.1]. This implies that

$$DT(x \,|\, v)(w) = \begin{cases} A_{x,v}(w) + N_{x,v}^T & \text{if } w \in N_{x,v}, \\ \emptyset & \text{otherwise}, \end{cases} \tag{3.6}$$

for some symmetric matrix $A_{x,v}$ and a subspace $N_{x,v}$.

We will now verify that $\operatorname{gph} DT(x \,|\, v) \subset \operatorname{gph} \partial^2 f(x \,|\, v)$. Take $\bar{w} \in N_{x,v}$ and $\bar{u} \in N_{x,v}^T$. For $w \in N_{x,v}$ and $u \in N_{x,v}^T$ we have

$$\begin{aligned} &\big\langle \big(A_{x,v}(\bar{w}) + \bar{u}, -\bar{w}\big), \big(w, A_{x,v}(w) + u\big) \big\rangle \\ &= \big\langle A_{x,v}(\bar{w}), w \big\rangle + \big\langle \bar{u}, w \big\rangle - \big\langle \bar{w}, A_{x,v}(w) \big\rangle - \big\langle \bar{w}, u \big\rangle = 0 \end{aligned}$$

(remember here that $A_{x,v}$ is symmetric). This confirms that the vector $\big(A_{x,v}(\bar{w})+\bar{u},-\bar{w}\big)$ belongs to the normal cone to $\operatorname{gph} T$ at (x,v), or in other words that $\big(\bar{w}, A_{x,v}(\bar{w})+\bar{u}\big)$ lies in $\operatorname{gph}\partial^2 f(x\,|\,v)$. (The normal cone includes the polar of the tangent cone, which in this case is the graph of the proto-derivative $DT(x\,|\,v)$, cf. [1].)

We then have the desired result by the definition of the mapping R and the fact that limsup of $\operatorname{gph}\partial^2 f(x\,|\,v)$ over the set of points $(x,v)\in\operatorname{gph} T$ converging to $(0,0)$ is a subset of $\operatorname{gph}\partial^2 f(0,0)$, cf. [23] and [1]. This finishes the justification of the claim.

Continuing with the proof of (b)⇒(a) in Theorem 1.3, we note that, by our assumptions, one has $\langle z,w\rangle>0$ for all $z\in R(w)$ with $w\neq 0$ (the latter being true for all $z\in\partial^2 f(0\,|\,0)(w)$ with $w\neq 0$). We will show that this condition implies the existence of $\epsilon>0$ such that, for all $\delta\in(0,\epsilon)$ and $(x,v)\in\operatorname{gph} T$ close enough to $(0,0)$ with $\operatorname{gph} DT(x\,|\,v)$ an n-dimensional subspace, one has

$$\langle z,w\rangle\geq\delta|w|^2 \text{ when } z\in DT(x\,|\,v)(w) \quad\text{with}\quad w\neq 0. \tag{3.7}$$

Recall that for such (x,v), $DT(x\,|\,v)$ has a representation in the form of (3.6). Actually, in that representation we may choose the symmetric matrix $A_{x,v}$ with the property that $A_{x,v}:N_{x,v}\to N_{x,v}$; cf. [29, proof of Proposition 4.1]. (In the cited proof, it is shown that if the subgradient mapping of a closed proper convex function is a subspace then the subgradient mapping has the desired representation. Our case follows from the convex situation because the second-order epi-derivative of a prox-regular function has the property that $d^2 f(x\,|\,v)+\rho|\cdot|^2$ is convex for some $\rho>0$; cf. [1], 13.49 or [2, Theorem 6.5].) Now suppose there exist $(x_\epsilon,v_\epsilon)\in\operatorname{gph} T$ converging to $(0,0)$ as $\epsilon\to 0$ with $DT(x\,|\,v)$ an n-dimensional subspace, and $w_\epsilon\in N_{x_\epsilon,v_\epsilon}$ with $|w_\epsilon|=1$ and $|A_{x_\epsilon,v_\epsilon}(w_\epsilon)|<\epsilon$. When this is true, then by taking a cluster point we can find w of norm one with $0\in R(w)$, a contradiction. Therefore, there exists $\epsilon>0$ such that for each $(x,v)\in\operatorname{gph} T$ close to $(0,0)$ with $DT(x\,|\,v)$ an n-dimensional subspace, the subspace $N_{x,v}$ contains only the origin or the minimum eigenvalue of $A_{x,v}$ is greater than ϵ. From this we have (3.7).

Next, from (3.7) we conclude that, for each $\delta\in(0,\epsilon)$, $DT(x\,|\,v)$ is strongly monotone with constant δ. From [2, Prop. 5.7] we see that T is strongly monotone with constant δ.

To complete the proof of the theorem, we need only establish the existence of a neighborhood X of 0 such that the mapping $v\mapsto\operatorname{argmin}_{x\in X}\big\{f(x)-\langle v,x\rangle\big\}$ coincides around 0 with the mapping $v\mapsto\big(\partial f\big)^{-1}(v)\cap X$. This is accomplished through [2, Thm. 5.5] and the fact that T is strongly monotone. Using [2, Thm. 5.5] we conclude that for some $k>0$ and $\lambda>0$, $e_\lambda-k|\cdot|^2$ is convex on a neighborhood of 0. Here e_λ is the Moreau-envelope of parameter $\lambda>0$, i.e.,

$$e_\lambda(x)=\min_{x'}\Big\{f(x')+\frac{1}{2\lambda}\big|x'-x\big|^2\Big\}.$$

This shows that $e_\lambda(x)\geq k|x|^2$, which implies that $f(x)\geq k|x|^2$ on a neighborhood of 0 (because $f(x)\geq e_\lambda(x)$). From the previous fact we can find neighborhoods X and V of 0 such that, for $v\in V$, the set $\operatorname{argmin}_{x\in X}\big\{f(x)-\langle v,x\rangle\big\}$ is nonempty and is in the interior of X. For such v, $\operatorname{argmin}_{x\in X}\big\{f(x)-\langle v,x\rangle\big\}\subset\big(\partial f\big)^{-1}(v)$. The inclusion must then be an equation, because the potentially bigger set is just a singleton. ☐

Remark 3.2. According to [30, Theorem 1.2], the mapping M is hypomonotone and has the Aubin property near $(0,\bar{x})\in\operatorname{gph} M$ when $\bar{x}$ gives a tilt-stable local minimum. The Aubin property refers to a lipschitzian-like property which is formulated for mappings that might be set-valued.

4. Application to Constrained Minimization.

Let $f_0 : I\!R^n \to I\!R$ be a C^2 function and let C be a nonempty, closed subset of $I\!R^n$. As already noted, the minimization of f_0 over C corresponds to the minimization of $f = f_0 + \delta_C$ over $I\!R^n$, where δ_C is the indicator of C. In this case, $\bar{x}$ gives a tilt-stable minimum of f if and only if there exists $\delta > 0$ such that the mapping

$$M : v \;\mapsto\; \operatorname{argmin}\Big\{ f_0(x) - f_0(\bar{x}) - \langle v, x - \bar{x}\rangle \;\Big|\; x \in C,\;\; |x - \bar{x}| \le \delta \Big\} \tag{4.1}$$

is single-valued and lipschitzian on some neighborhood of $v = 0$. It is appropriate to speak then of $\bar{x}$ giving a *tilt-stable local minimum of* f_0 *relative to* C. What does Theorem 1.3 say about that property?

In this setting, we have $\partial f(\bar{x}) = \nabla f_0(\bar{x}) + \partial\delta_C(\bar{x})$. On the level of the first-order condition for optimality, therefore, we have

$$0 \in \partial f(\bar{x}) \iff -\nabla f_0(\bar{x}) \in \partial\delta_C(\bar{x}) \iff -\nabla f_0(\bar{x}) \in N_C(\bar{x}). \tag{4.2}$$

To proceed to the level of the second-order condition, we can appeal to the notion of C being prox-regular at $\bar{x}$ for a vector $\bar{v}$, which by definition is equivalent to the function δ_C being prox-regular at $\bar{x}$ for $\bar{v}$; see [2], where this class of sets is discussed in some detail.

PROPOSITION 4.1. *If $f = f_0 + \delta_C$ with f_0 of class $\mathcal{C}^2$ and C closed, one has for any $\bar{x} \in C$ and $w \in I\!R^n$ that*

$$\partial^2 f(\bar{x}\,|\,0)(w) = \nabla^2 f_0(\bar{x})w + \partial^2\delta_C(\bar{x}\,|\,\bar{v})(w) \quad \textit{for} \quad \bar{v} = -\nabla f_0(\bar{x}). \tag{4.3}$$

Furthermore, f is prox-regular and subdifferentially continuous at $\bar{x}$ if and only if C is prox-regular at $\bar{x}$ for $\bar{v}$.

Proof. By a calculus result of Mordukhovich [23], a mapping of the form $T(x) = F(x) + S(x)$ with F single-valued and of class $\mathcal{C}^1$ has $D^*T(x\,|\,v) = \nabla F(x) + D^*S(x\,|\,v - F(x))$. Here we apply this to $F = \nabla f_0$ and $S = \partial\delta_C$. Trivially, the function δ_C is subdifferentially continuous. But from [1], 13.35, the prox-regularity of a function is preserved when a $\mathcal{C}^2$ function is added; the same holds for subdifferential continuity as well. □

THEOREM 4.2. *For a $\mathcal{C}^2$ function f_0 on $I\!R^n$ and a closed set $C \subset I\!R^n$, let $\bar{x}$ be a point of C satisfying the first-order condition $-\nabla f_0(\bar{x}) \in \partial\delta_C(\bar{x})$, and suppose that C is prox-regular at $\bar{x}$ (i.e., that δ_C is prox-regular there). Then $\bar{x}$ gives a tilt-stable local minimum of f_0 relative to C if and only if, in terms of $\bar{v} = -\nabla f_0(\bar{x})$, one has*

$$\big\langle w, \nabla^2 f_0(\bar{x})w \big\rangle > -\big\langle z, w\big\rangle \quad \textit{whenever} \quad z \in \partial^2\delta_C(\bar{x}\,|\,\bar{v})(w),\;\; w \neq 0. \tag{4.4}$$

Proof. This is evident from the equivalence of (a) and (b) of Theorem 1.3 in the context of Proposition 4.1. □

When the set C is convex, the coderivative result in Theorem 2.1 can be brought into this picture.

COROLLARY 4.3. *For a $\mathcal{C}^2$ function f_0 and a closed, convex set C, let $\bar{x}$ be a point satisfying the first-order condition $-\nabla f_0(\bar{x}) \in \partial\delta_C(\bar{x})$. Then, with $\bar{v} = -\nabla f_0(\bar{x})$, a sufficient condition for $\bar{x}$ to give a tilt-stable local minimum of f_0 relative to C is that*

$$\big\langle w, \nabla^2 f_0(\bar{x})w \big\rangle > 0 \quad \textit{whenever} \quad w \in \operatorname{dom}\partial^2\delta_C(\bar{x}\,|\,\bar{v}),\;\; w \neq 0. \tag{4.5}$$

Proof. The convexity of C implies that the normal cone mapping $N_C = \partial\delta_C$ is maximal monotone; cf. [28]. Then $\langle z, w\rangle \geq 0$ by Theorem 2.1 whenever $z \in \partial^2\delta_C(\bar{x}\,|\,\bar{v})(w)$. Condition (4.5) is enough in that case to ensure that (4.4) holds. □

When C is not just convex but also polyhedral, an exact formula for the generalized Hessian $\partial^2\delta_C(\bar{x}\,|\,\bar{v})$ is available from [31]. In order to state this formula, we need to deal with the *critical cone* for C at $\bar{x}$ for a vector $\bar{v} \in \partial\delta_C(\bar{x}) = N_C(\bar{x})$; this is the polyhedral convex cone

$$K(\bar{x}, \bar{v}) = \{\, w \in T_C(\bar{x}) \,|\, w \perp \bar{v}\,\},$$

where $T_C(\bar{x})$ denotes the tangent cone to C at $\bar{x}$. Recall that a *closed face* F of a polyhedral convex cone K is a polyhedral convex cone of the form

$$F = \{\, x \in K \,|\, x \perp v\,\} \quad \text{for some} \quad v \in K^*,$$

where K^* denotes the polar of K.

PROPOSITION 4.4. *For a polyhedral convex set C, the generalized Hessian of the indicator of C at $\bar{x}$ for any $\bar{v} \in \partial\delta_C(\bar{x})$ is given, with respect to the critical cone $K(\bar{x}, \bar{v})$, by*

$$z \in \partial^2\delta_C(\bar{x}\,|\,\bar{v})(w) \iff \begin{cases} \text{there exist closed faces } F_1 \text{ and } F_2 \text{ of } K(\bar{x},\bar{v}) \\ \text{with } F_2 \subset F_1,\ w \in F_2 - F_1,\ z \in (F_1 - F_2)^*. \end{cases} \tag{4.6}$$

Proof. In [31, proof of Thm. 2], the normal cone to $\operatorname{gph} N_C = \operatorname{gph}\partial\delta_C$ at the point $(\bar{x}, \bar{v})$ is shown to be the union of products sets $K^* \times K$ over all the cones $K = F_1 - F_2$, where F_1 and F_2 are described as above. □

THEOREM 4.5. *For a $\mathcal{C}^2$ function f_0 and a polyhedral convex set C, let $\bar{x}$ be a point satisfying the first-order condition $-\nabla f_0(\bar{x}) \in \partial\delta_C(\bar{x})$. Then, with respect to $\bar{v} = -\nabla f_0(\bar{x})$ and the critical cone $K(\bar{x}, \bar{v})$, a necessary and sufficient condition for $\bar{x}$ to give a tilt-stable local minimum of f_0 relative to C is that*

$$\big\langle w, \nabla^2 f_0(\bar{x})w\big\rangle > 0 \quad \textit{for all } w \neq 0 \textit{ in } K(\bar{x},\bar{v}) - K(\bar{x},\bar{v}). \tag{4.7}$$

Proof. On the basis of Theorem 4.2 and Proposition 4.4, we know that $\bar{x}$ gives a tilt-stable local minimum if and only if $\langle w, \nabla^2 f_0(\bar{x})w\rangle > -\langle z, w\rangle$ whenever $w \neq 0$ and there exist closed faces F_1 and F_2 of $K(\bar{x}, \bar{v})$ with $F_2 \subset F_1$, such that $w \in F_2 - F_1$ and $z \in (F_1 - F_2)^*$. In those circumstances one can always take $z = 0$, so the condition comes down to having $\langle w, \nabla^2 f_0(\bar{x})w\rangle > 0$ for all $w \neq 0$ in the union of all cones $F_2 - F_1$ generated by faces of $K(\bar{x}, \bar{v})$, as described. But all such faces lie in $K(\bar{x}, \bar{v})$, which in particular is a closed face of itself. Therefore, the union is simply $K(\bar{x}, \bar{v}) - K(\bar{x}, \bar{v})$. □

Note that $K(\bar{x}, \bar{v}) - K(\bar{x}, \bar{v})$ is a subspace of $\mathbb{R}^n$, the smallest one that includes $K(\bar{x}, \bar{v})$. When the critical cone $K(\bar{x}, \bar{v})$ is itself a subspace, condition (4.7) is the same as

$$\big\langle w, \nabla^2 f_0(\bar{x})w\big\rangle > 0 \quad \text{for all } w \neq 0 \text{ in } \quad K(\bar{x},\bar{v}),$$

which is the usual second-order sufficient condition for optimality in this framework of polyhedral convexity.

Elaborations of Theorem 4.2 for other, nonpolyhedral classes of sets C must await further advances in the calculus of generalized Hessians.

REFERENCES

[1] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*, Springer-Verlag, to appear.

[2] R. A. Poliquin and R. T. Rockafellar, *Prox-regular functions in variational analysis*, Trans. Amer. Math. Soc., 348 (1996), pp. 1805–1838.

[3] ———, *Generalized Hessian properties of regularized nonsmooth functions*, SIAM J. Optimization, 6 (1996), pp. 1121–1137.

[4] ———, *Second-order nonsmooth analysis in nonlinear programming*, in Recent Advances in Nonsmooth Optimization, D. Du, L. Qi and R. Womersley, eds., World Scientific Publishers, 1995, pp. 322–350.

[5] ———, *Amenable functions in optimization*, in Nonsmooth Optimization Methods and Applications, F. Giannessi, ed., Gordon and Breach, Philadelphia, 1992, pp. 338–353.

[6] ———, *A calculus of epi-derivatives applicable to optimization*, Canadian J. Math, 45 (1993), pp. 879–896.

[7] ———, *Proto-derivatives formulas for basic subgradient mappings in mathematical programming*, Set-Valued Analysis, 2 (1994), pp. 275–290.

[8] A. Levy and R. T. Rockafellar, *Variational conditions and the proto-differentiation of partial subgradient mappings*, Nonlinear Anal. Th. Meth. Appl., 26 (1996), pp. 1951–1964.

[9] J. V. Burke, *Second order necessary and sufficient conditions for convex composite NDO*, Math. Prog., 38 (1987), pp. 287–302.

[10] R. T. Rockafellar, *First- and second-order epi-differentiability in nonlinear programming*, Trans. Amer. Math. Soc., 307 (1988), pp. 75–107.

[11] ———, *Second-order optimality conditions in nonlinear programming obtained by way of epi-derivatives*, Math.of Oper. Research, 14 (1989), pp. 462–484.

[12] R. A. Poliquin, *Proto-differentiation of subgradient set-valued mappings*, Canadian J. Math., 42 (1990), pp. 520–532.

[13] R. Cominetti, *On Pseudo-differentiability*, Trans. Amer. Math. Soc., 324 (1991), pp. 843–865.

[14] A. D. Ioffe, *Variational analysis of a composite function: a formula for the lower second order epi-derivative*, J. Math. Anal. Appl., 160 (1991), pp. 379–405.

[15] R. A. Poliquin, *Integration of subdifferentials of nonconvex functions*, Nonlinear Anal. Th. Meth.Appl., 17 (1991), pp. 385–398.

[16] J.V. Burke and R. A. Poliquin, *Optimality conditions for non-finite convex composite functions*, Math. Prog., 57 (1992), pp. 103–120.

[17] R. A. Poliquin, *An extension of Attouch's Theorem and its application to second-order epi-differentiation of convexly composite functions*, Trans. Amer. Math. Soc., 332 (1992), pp. 861–874.

[18] R. A. Poliquin, J. Vanderwerff and V. Zizler, *Convex composite representation of lower semicontinuous functions and renormings*, C.R. Acad. Sci. Paris, t. 317 Série I (1993), pp. 545–549.

[19] A. Levy, R. A. Poliquin and L. Thibault, *Partial extension of Attouch's theorem with applications to proto-derivatives of subgradient mappings*, Trans. Amer. Math. Soc., 347 (1995), pp. 1269–1294.

[20] B. S. Mordukhovich, *Sensitivity analysis in nonsmooth optimization*, in Theoretical aspects of industrial design, D. A. Field and V. Komkov, eds., SIAM Volumes in Applied Mathematics 58, 1992, pp. 32–46

[21] ———, *Lipschitzian stability of constraint systems and generalized equations*, Nonlinear Analysis Th. Meth. Appl., 22 (1994), pp. 173–206.

[22] ———, *Complete characterization of openness, metric regularity, and Lipschitzian properties of multifunctions*, Trans. Amer. Math. Soc., 340 (1993), pp. 1–36.

[23] ———, *Generalized differential calculus for nonsmooth and set-valued mappings*, J. Math. Anal. and Appl., 183 (1994), pp. 250–288.

[24] A. D. Ioffe, *Approximate subdifferentials and applications, I: the finite dimensional theory*, Trans. Amer. Math. Soc., 281 (1984), pp. 389–416.

[25] R. T. Rockafellar, *Favorable classes of Lipschitz continuous functions in subgradient optimization*, in Progress in Nondifferentiable Optimization, E. Nurminski (ed.), IIASA Collaborative Proceedings Series, International Institute of Applied Systems Analysis, Laxenburg, Austria 1982, pp. 125-144.

[26] ———, *On the maximality of sums of nonlinear monotone operators*, Trans. Amer. Math. Soc., 149 (1970), pp. 75–88.

[27] ———, *Proto-differentiability of set-valued mappings and its applications in optimization*, in Analyse Non Linéaire, H. Attouch et al. (eds.), Gauthier-Villars, Paris (1989), pp. 449–482.

[28] ———, *Convex Analysis*, Princeton University Press, Princeton, NJ, 1970.

[29] ———, *Maximal monotone relations and the second derivatives of nonsmooth functions*, Ann. Inst. H. Poincaré: Analyse Non Linéaire, 2 (1985), pp. 167–184.

[30] A. B. Levy and R. A. Poliquin, *Characterizations of the local single-valuedness of multifunctions*, Set-Valued Analysis, to appear.

[31] A. Dontchev and R. T. Rockafellar, *Characterizations of strong regularity for variational inequalities over polyhedral convex sets*, SIAM J. Optimization, 6 (1996), pp. 1087–1105.