TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 352, Number 11, Pages 5231–5249
S 0002-9947(00)02550-2
Article electronically published on June 9, 2000

LOCAL DIFFERENTIABILITY OF DISTANCE FUNCTIONS

R. A. POLIQUIN, R. T. ROCKAFELLAR, AND L. THIBAULT

ABSTRACT. Recently Clarke, Stern and Wolenski characterized, in a Hilbert space, the closed subsets C for which the distance function d_C is continuously differentiable everywhere on an open "tube" of uniform thickness around C. Here a corresponding local theory is developed for the property of d_C being continuously differentiable outside of C on some neighborhood of a point $x \in C$. This is shown to be equivalent to the prox-regularity of C at x, which is a condition on normal vectors that is commonly fulfilled in variational analysis and has the advantage of being verifiable by calculation. Additional characterizations are provided in terms of d_C^2 being locally of class C^{1+} or such that $d_C^2 + \sigma|\cdot|^2$ is convex around x for some $\sigma > 0$. Prox-regularity of C at x corresponds further to the normal cone mapping N_C having a hypomonotone truncation around x, and leads to a formula for P_C by way of N_C. The local theory also yields new insights on the global level of the Clarke-Stern-Wolenski results, and on a property of sets introduced by Shapiro, as well as on the concept of sets with positive reach considered by Federer in the finite dimensional setting.

1. INTRODUCTION

The distance function d_C for a closed subset C of a Hilbert space H gives for each $u \in H$ the distance $d_C(u) = \inf\{|u - x| \,|\, x \in C\}$. To what extent is d_C Fréchet or Gâteaux differentiable, or continuously differentiable (the Gâteaux case then automatically implying the Fréchet sense)? This is of considerable interest in variational analysis, not only for its connection to the geometry of C and the projection mapping P_C (giving for each u the set points of C nearest to u) but also for its applications in optimization. The distance to the feasible set in a problem of constrained minimization, for instance, can be used as a penalty in setting up a computationally equivalent unconstrained problem. For convex C, the differentiability of d_C everywhere outside of C is well known, but for nonconvex C, less has been understood, apart from results on generic differentiability as in Borwein and Giles [1].

Received by the editors June 17, 1997 and, in revised form, June 10, 1998.
1991 *Mathematics Subject Classification.* Primary 49J52, 58C06, 58C20; Secondary 90C30.
Key words and phrases. Variational analysis, distance functions, single-valued projections, proximal normals, prox-regularity, proximal smoothness, primal-lower-nice functions, hypomonotone mappings, monotone mappings.
This work was supported in part by the Natural Sciences and Engineering Research Council of Canada under grant OGP41983 for the first author, by the National Science Foundation under grant DMS–9500957 for the second author, and by NATO under grant CRG 960360 for the third author. The authors are grateful for useful discussions with C. Combari, and for helpful comments made by the referee.

Clarke, Stern and Wolenski [2] recently made headway by studying, as a generalization of convex sets, the *proximally smooth* sets, which they defined to be the closed sets $C \subset H$ such that d_C is (norm-to-norm-) continuously differentiable on an open "tube" of the type

$$U_C(r) := \{ u \in H \,|\, 0 < d_C(u) < r \} \tag{1.1}$$

for some $r > 0$. They characterized such sets in several interesting ways. In particular, they showed that C is proximally smooth if and only if there exists $r > 0$ such that, for all $u \in U_C(r)$, the projection $P_C(u)$ is nonempty and each of its elements x belongs also to $P_C(x + v)$ for $v = r[u - x]/|u - x|$; cf. [2, Theorem 4.1(d)]. Since the vectors v of the form $v = \lambda[u - x]/|u - x|$ for some $u \in P_C^{-1}(x)$ and $\lambda > 0$ are by definition the nonzero *proximal normals* to C at x, they spoke of the latter as meaning that "every nonzero proximal normal v to C can be realized by an r-ball"; an equivalent statement is that

$$0 \;\geq\; \left\langle \frac{v}{|v|}, x' - x \right\rangle - \frac{1}{2r}|x' - x|^2, \quad \forall x' \in C. \tag{1.2}$$

Sets that satisfy (1.2) have appeared elsewhere in the literature under several names. We refer the reader to [3] and the references therein for more information.

Beyond the appeal of this global property on a tube, there is a need for local information on the behavior of d_C around a point $\bar{x} \in C$, because applications are often of this character and do not require global considerations. What characterizations can be given for the existence of an open neighborhood O of $\bar{x}$ such that d_C is continuously differentiable on $O \backslash C$ (relative complement)? It might be imagined that local results could be obtained by invoking global results about proximal smoothness in the case of $C \cap B$ for some closed ball B centered at $\bar{x}$, but this runs into serious difficulty over what happens at the points where the boundary of B meets C. From another angle, the trouble can be seen in the fact that the tube concept in (1.1) is hard to coordinate with that of a neighborhood of a point $\bar{x}$ because of the way it depends also on other points of C near to $\bar{x}$.

There is a need also for better understanding of how local properties of d_C correspond to those of P_C. It is well known that a closed convex set C has its projection mapping P_C globally single-valued and nonexpansive (Lipschitz continuous with modulus 1). For nonconvex sets C, where a distinction has to be made between strong and weak closure, Clarke, Stern and Wolenski [2] showed that a weakly closed set C is proximally smooth if and only if P_C is single-valued on a tube $U_C(r)$. Another result was obtained by Shapiro [4] on the local level. He showed, for a strongly closed set C and a point $\bar{x} \in C$, that P_C is single-valued on a neighborhood of $\bar{x}$ if the following property holds: there is a constant $k > 0$ along with a neighborhood O of $\bar{x}$ such that

$$d_{T_C(x)}(x' - x) \leq k|x' - x|^2 \quad \text{for all } x, x' \in C \cap O, \tag{1.3}$$

where $T_C(x)$ denotes the general tangent cone (contingent cone) to C at x. We'll refer to this condition as the *Shapiro property* of C at $\bar{x}$. (Shapiro actually introduced in [4] a more general condition of C being what he called $O(m)$-*convex* at $\bar{x}$, for which this is the case of $m = 2$.) The single-valuedness of the projection mapping on a neighborhood of $\bar{x}$ was used by Federer to define sets with positive reach near $\bar{x}$. In the finite dimensional setting, Federer [5] established, among other results, that the square of d_C is continuously differentiable near $\bar{x}$ whenever C has positive reach near $\bar{x}$.

In taking up the challenge of a local theory of differentiability of the distance function d_C and its consequences for the projection mapping P_C in the Hilbert space setting, we rely on a different property of C at a point $\bar{x}$, namely prox-regularity. This property has so far been considered only in the finite-dimensional case, where it was introduced by Poliquin and Rockafellar [6]; see also [7]–[9]. In defining it, we denote by $N_C(\bar{x})$ the general cone of normals to C at a point $\bar{x} \in C$; a vector $v \neq 0$ belongs to $N_C(\bar{x})$ if and only if there is a sequence of points $x_k \to \bar{x}$ in C at which there are proximal normals v_k converging weakly to v. (Along with such vectors $v \neq 0$, the cone $N_C(\bar{x})$ is defined to contain $v = 0$.)

Definition 1.1. *A closed set C is* prox-regular at $\bar{x}$ for $\bar{v}$, *where $\bar{x} \in C$ and $\bar{v} \in N_C(\bar{x})$, if there exist $\varepsilon > 0$ and $\rho > 0$ such that whenever $x \in C$ and $v \in N_C(x)$ with $|x-\bar{x}| < \varepsilon$ and $|v-\bar{v}| < \varepsilon$, then x is the unique nearest point of $\{x' \in C \,|\, |x'-\bar{x}| < \varepsilon\}$ to $x + \rho^{-1}v$. It is* prox-regular at $\bar{x}$ *(without mention of a particular $\bar{v}$) if this property holds for every vector $\bar{v} \in N_C(\bar{x})$.*

Poliquin and Rockafellar [6], developed prox-regularity more broadly, as a property of functions and their subgradients, rather than sets and their normals. The set version was obtained by specializing to indicator functions. Although we deal here only with sets, the tie to functions is important because a number of fundamental results in variational analysis revolve around prox-regularity in that context. For instance, prox-regularity is the key to connections between generalized second-order derivatives of f and graphical derivatives of its subgradient mapping ∂f, and thus in the indicator case it is the key to such derivatives of the mapping N_C. By putting prox-regularity of C at the center of our discussion, we provide access not only to that larger framework but also to the many examples of prox-regularity in the literature.

In concentrating on sets, we will find it helpful to have an alternative description of prox-regularity alongside of Definition 1.1.

Proposition 1.2. *A closed set C is prox-regular at $\bar{x}$ if and only if it is prox-regular at $\bar{x}$ for the vector $\bar{v} = 0$. This is equivalent to the existence of $\varepsilon > 0$ and $\rho > 0$ such that whenever $x \in C$ and $v \in N_C(x)$ with $|x - \bar{x}| < \varepsilon$ and $|v| < \varepsilon$, one has*

$$(1.4) \qquad 0 \;\geq\; \langle v, x' - x\rangle - \frac{\rho}{2}|x' - x|^2 \quad \textit{for all } x' \in C \textit{ with } |x' - \bar{x}| < \varepsilon.$$

Proof. Obviously if C is prox-regular at $\bar{x}$ for every $\bar{v} \in N_C(\bar{x})$, it is prox-regular at $\bar{x}$ for $\bar{v} = 0$. To prove the converse, assume that C is prox-regular at $\bar{x}$ for the vector 0 with constants $\varepsilon > 0$ and $\rho > 0$. Take $\bar{v} \in N_C(\bar{x})$ with $\bar{v} \neq 0$, and let $\varepsilon' := \min\{\varepsilon/2, |\bar{v}|/2\}$. For $x \in C$ and $v \in N_C(x)$ with $|x - \bar{x}| < \varepsilon'$ and $|v - \bar{v}| < \varepsilon'$ we have

$$(\varepsilon/2|\bar{v}|)|v| \leq (\varepsilon/2|\bar{v}|)|v - \bar{v} + \bar{v}| \leq (\varepsilon/4) + (\varepsilon/2) < \varepsilon.$$

By the choice of ε this implies that x is the unique closest point of $\{x' \in C \,|\, |x'-\bar{x}| < \varepsilon\}$ to $x + \rho^{-1}(\varepsilon/2|\bar{v}|)v$. From this we conclude that C is prox-regular at $\bar{x}$ for $\bar{v}$ with constants ε' and $\rho' := \varepsilon^{-1}2\rho|\bar{v}|$.

For the second claim of the proposition, note that the inequality in (1.4) can be made strict for $x' \neq x \in C$ by replacing ρ with $\rho' > \rho$. With the inequality in (1.4) now strict, (1.4) is equivalent to saying that x is the unique closest point of $\{x' \in C \,|\, |x' - \bar{x}| < \varepsilon\}$ to $x + {\rho'}^{-1}v$. Therefore C is prox-regular at $\bar{x}$ for $\bar{v} = 0$ if and only if there exist $\varepsilon > 0$ and $\rho > 0$ such that whenever $x \in C$ and $v \in N_C(x)$ with $|x - \bar{x}| < \varepsilon$ and $|v| < \varepsilon$, one has (1.4). $\square$

A special virtue of prox-regularity is that it can be established in many situations by checking whether a constraint qualification is satisfied. Poliquin and Rockafellar in [6] gave a number of examples of sets exhibiting prox-regularity in finite dimensions. In particular they showed that, under natural assumptions, a set C enjoying a smooth constraint representation around a point $x \in C$ is prox-regular at x for any $v \in N_C(x)$; see Section 2.

We are ready to state our main result. In this theorem we work with the mapping $N_C^r : H \rightrightarrows H$ defined for $r > 0$ by

$$N_C^r(x) = \begin{cases} N_C(x) \cap \operatorname{int} \mathbb{B}\big(0, r\big) & \text{if} \quad x \in C, \\ \emptyset & \text{if} \quad x \notin C. \end{cases}$$

(Here $\mathbb{B}\big(0, r\big)$ denotes the closed ball of center 0 and radius r.) A mapping $T : H \rightrightarrows H$ is *hypomonotone* on a subset O of X if there exists $\sigma > 0$ such that $T + \sigma I$ is monotone on O; this corresponds to having

$$\langle v_1 - v_2, x_1 - x_2\rangle \geq -\sigma|x_1 - x_2|^2 \quad \text{whenever } v_i \text{ are in } T(x_i) \text{ and } x_i \text{ are in } O.$$

Theorem 1.3. *For a closed set $C \subset H$ and any point $\bar{x} \in C$, the following properties are equivalent:*

(a) C is prox-regular at $\bar{x}$.

(b) d_C is continuously differentiable on $O \backslash C$ for some open neighborhood O of $\bar{x}$.

(c) d_C is Fréchet differentiable on $O \backslash C$ for some open neighborhood O of $\bar{x}$.

(d) d_C is Gâteaux differentiable on $O \setminus C$ for some open neighborhood O of $\bar{x}$, and P_C is nonempty-valued on O.

(e) d_C^2 is C^{1+} on an open neighborhood O of $\bar{x}$. i.e., Fréchet differentiable on O with the derivative mapping $D(d_C^2)(x) : H \rightrightarrows H$ depending Lipschitz continuously on x.

(f) There exist $r > 0$ and a neighborhood O of $\bar{x}$ such that every nonzero proximal normal to C at any x in $C \cap O$ can be realized by an r-ball.

(g) For some $r > 0$ and neighborhood O of $\bar{x}$, the truncated mapping N_C^r is hypomonotone on O.

(h) There exists $\lambda > 0$ such that

$$\left.\begin{aligned} x = P_C(u),\ x \neq u \\ 0 < |u - \bar{x}| < \lambda \end{aligned}\right\} \implies x = P_C(u') \text{ for } u' = x + \lambda\frac{u - x}{|u - x|}. \tag{1.5}$$

(i) P_C is single-valued and strongly-weakly continuous (i.e., from the strong topology in the domain to the weak topology in the range) on a neighborhood of $\bar{x}$.

(j) C has the Shapiro property at $\bar{x}$.

Then there is a neighborhood O of $\bar{x}$ on which P_C is single-valued, monotone and Lipschitz continuous with $P_C = (I + N_C^r)^{-1}$ on O for some $r > 0$, whereas $D(d_C) = [I - P_C]/d_C$ on $O \setminus C$. Here $I : H \to H$ denotes the identity mapping.

If the set C is weakly closed relative to a (strong) neighborhood of $\bar{x}$ (which is always the case when the space H is finite-dimensional), then one can add the following to the set of equivalent properties:

(k) P_C is single-valued around $\bar{x}$.

In the equivalence in Theorem 1.3 between the prox-regularity property (a) and the Shapiro property (j), the implication from (j) to (g) could be seen already in Shapiro's paper [4] (written before prox-regularity was developed in [6]). By providing the reverse implication along with the other equivalences, we place the Shapiro property in a much stronger light. Shapiro also proved (in the same paper) that a set with the Shapiro property has a (locally) Lipschitz continuous projection mapping. This property was also noted by Federer in [5] for sets with positive reach in finite dimensions.

Other aspects of Theorem 1.3 are worth noting as well. We see that Fréchet differentiability of d_C is sufficient to ensure the Lipschitz continuity of its derivative (locally). We have a criterion for P_C to be single-valued, monotone and Lipschitz continuous around $\bar{x}$, with an exact formula for P_C in terms of a truncation of the normal cone mapping N_C. Moreover the hypomonotonicity of this truncation characterizes prox-regularity.

Our paper is organized as follows. In Section 2 we discuss the relationship between p.l.n. (i.e., *primal-lower-nice*) functions and prox-regular sets. The results obtained in this section enable us to conclude that (a) is equivalent to (g) in Theorem 1.3. Section 3 is devoted to the remainder of the proof of Theorem 1.3 and of the statement and proof of its corollaries. There too, we establish that C is prox-regular at $\bar{x}$ if and only if there exists $\sigma > 0$ such that the function $d_C^2 + \sigma|\cdot|^2$ is convex on some open neighborhood O_σ of $\bar{x}$. In Section 4 we use the techniques developed in Sections 2 and 3 to obtain results similar to Theorem 1.3, but on the global level of proximally smooth sets. Several results of Clarke-Stern-Wolenski [2] are rederived in this way, and others are added.

2. P.L.N. Functions

In finite-dimensional spaces, the equivalence between (a) and (g) in Theorem 1.3 can be derived in a much more general context, namely that of a prox-regular function; see [6]. In the setting of an indicator function, it is actually a consequence of earlier work on p.l.n. (primal-lower-nice) functions; for more on p.l.n. functions, see [10]–[14]. Recall that a lower semicontinuous function $f : H \to \overline{\mathbb{R}}$ is p.l.n. at $\bar{x}$, a point where f is finite, if there exist $t_0 > 0$, $c > 0$ and $\varepsilon > 0$ with the property that

$$f(x') \geq f(x) + \langle v, x' - x\rangle - \frac{t}{2}\left|x' - x\right|^2 \tag{2.1}$$

whenever $t > t_0$, $\left|v\right| < ct$, $v \in \partial_p f(x)$, $\left|x' - \bar{x}\right| < \varepsilon$, and $|x - \bar{x}| < \varepsilon$. Here $\partial_p f(x)$ is the set of proximal subgradients to f at x, i.e., $v \in \partial_p f(x)$ if there exist $t \geq 0$ such that (2.1) is verified in a neighborhood of x (for more on proximal subgradient see [2] and [9]). We will denote by $\partial f(x)$ the set of weak-limiting proximal subgradients to f at x; thus $v \in \partial f(x)$ if there exists x_k converging strongly to x with $f(x_k)$ converging to $f(x)$ and v_k converging weakly to v with $v_k \in \partial_p f(x_k)$. Note that for a closed set C and any point $x \in C$ we have $N_C(x) = \partial\delta_C(x)$, and the cone of proximal normals to C at x is equal to $\partial_p\delta_C(x)$; see [9] for more details.

The fact that a function is p.l.n. has powerful consequences. For example, if the function is p.l.n. at $\bar{x}$, then for all x in a neighborhood of $\bar{x}$ we have $\partial f(x) = \partial_p f(x)$, and this set is closed and convex; see [13, Theorem 2.4]. The connection between prox-regular sets and p.l.n. functions will now be established.

Proposition 2.1. *The set C is prox-regular at $\bar{x} \in C$ if and only if the indicator of C is p.l.n. at $\bar{x}$.*

Proof. When the indicator of C is p.l.n. at $\bar{x}$, then (as noted above) $N_C(x)$ agrees with the cone of proximal normals to C at x for all x in a neighborhood of $\bar{x}$. From this we easily establish that C is prox-regular at $\bar{x}$ for $\bar{v} = 0$, and therefore that C is prox-regular at $\bar{x}$, according to Proposition 1.2.

Now assume that C is prox-regular at $\bar{x}$ for $\bar{v} = 0$. By Proposition 1.2, there then exist $\rho > 0$ and $\varepsilon > 0$ such that

$$\delta_C(x') \geq \delta_C(x) + \langle v, x' - x \rangle - (\rho/2)|x' - x|^2$$

whenever $|x - \bar{x}| < \varepsilon$, $|x' - \bar{x}| < \varepsilon$, $|v| < \varepsilon$ with $v \in N_C(x)$. Let $c = \varepsilon/\rho$. If $v \in N_C(x)$ and $|v| \leq ct$, then $(\rho/t)v \in N_C(x)$ with $|(\rho/t)v| \leq \varepsilon$. This implies that

$$\delta_C(x') \geq \delta_C(x) + (\rho/t)\langle v, x' - x \rangle - (\rho/2)|x' - x|^2 \tag{2.2}$$

whenever $t > 0$, $|x - \bar{x}| < \varepsilon$, $|x' - \bar{x}| < \varepsilon$, $|v| < ct$ with $v \in N_C(x)$. Note that (2.2) is equivalent to

$$\delta_C(x') \geq \delta_C(x) + \langle v, x' - x \rangle - (t/2)|x' - x|^2. \tag{2.3}$$

This shows that δ_C is p.l.n. at $\bar{x}$. □

As a consequence of Proposition 2.1 we get the following piece of Theorem 1.3.

Corollary 2.2. *Let C be a closed subset of H and let $\bar{x} \in C$. The set C is prox-regular at $\bar{x}$ if and only if for some $r > 0$ and some neighborhood O of $\bar{x}$, N_C^r is hypomonotone on O. In that case there exists an open neighborhood O of $\bar{x}$ such that for all $x \in O \cap C$ the normal cone $N_C(x)$ is closed and convex, with every $v \in N_C(x)$ actually being a proximal normal to C at x.*

Proof. This follows from [13, Cor. 2.3 and Thm. 2.4]. To use [13, Cor. 2.3], simply note (as in the proof of Proposition 2.1) that the hypomonotonicity of N_C^r on some neighborhood O of $\bar{x}$ is equivalent to the existence of $c > 0$, $t_0 > 0$ and $\varepsilon > 0$ with the property that

$$\langle v_1 - v_2, x_1 - x_2 \rangle \geq -t|x_1 - x_2|^2$$

whenever $v_i \in N_C(x_i)$, $|v_i| \leq ct$ and $|x_i - \bar{x}| \leq \varepsilon$. □

In [6], a major class of sets enjoying prox-regularity locally was developed in terms of constraint representations. It was shown that $C \subset \mathbb{R}^n$ is prox-regular at $\bar{x}$ if there is an open neighborhood O of $\bar{x}$ such that

$$C \cap O = \big\{ x \in O \,\big|\, F(x) \in D \big\} \tag{2.4}$$

for a C^2 mapping $F : O \to \mathbb{R}^m$ and a closed, convex set $D \subset \mathbb{R}^m$ satisfying the constraint qualification that the only vector $y \in N_D\big(F(\bar{x})\big)$ with $\nabla F(\bar{x})^* y = 0$ is $y = 0$. (The Jacobian matrix for F at $\bar{x}$ is denoted here by $\nabla F(\bar{x})$, and its adjoint by $\nabla F(\bar{x})^*$.) Because D is convex, this constraint qualification is equivalent to having

$$\mathbb{R}_+\big[D - F(\bar{x})\big] - \nabla F(\bar{x})\mathbb{R}^n \;=\; \mathbb{R}^m.$$

Provided we adopt an extended version of the alternate form of the constraint qualification, this example carries forward to the setting of an infinite-dimensional Hilbert space. In formulating the next result, we denote by $DF(x)$ the Fréchet derivative mapping associated with F at x.

Proposition 2.3. *For a closed set $C \subset H$ and a point $\bar{x} \in C$, assume that (2.4) holds for an open neighborhood O of $\bar{x}$, a closed, convex set D in a Banach space E, and a mapping $F : O \to E$ that is Fréchet differentiable and such that $DF(x)$ depends Lipschitz continuously on $x \in O$. If*

$$\mathbb{R}_+\big[D - F(\bar{x})\big] - DF(\bar{x})(H) = E,$$

then C is prox-regular at $\bar{x}$.

Proof. Apply [10, Theorem 2.4] to conclude that δ_C is p.l.n. at $\bar{x}$, and then invoke Proposition 2.1 of the present paper. □

We will now show that if every nonzero proximal normal to a set C at any point x of C can be realized by an r-ball, then C is uniformly prox-regular in the following sense.

Definition 2.4. *A closed set C is uniformly prox-regular with constant $\rho > 0$ if whenever $x \in C$ and $v \in N_C(x)$ with $|v| < 1$, then x is the unique nearest point of C to $x + \rho^{-1}v$.*

At first glance it might seem obvious that if every nonzero proximal normal to a set C at any point x of C can be realized by some r-ball then C is uniformly prox-regular, but in the definition of uniform prox-regularity, *all* normal vectors $v \in N_C(x)$ with $|v| < 1$ are involved (not just the proximal normals). Although it is true that every normal vector is a weak limit of proximal normal vectors, one cannot control the norms of these proximal normal vectors. We get around these difficulties by showing, with the help of the following proposition and Corollary 2.2, that for a proximally smooth set C every vector $v \in N_C(x)$ must be a proximal normal vector.

Proposition 2.5. *Assume there exist $r > 0$ and an open neighborhood O of $\bar{x} \in C$ such that every nonzero proximal normal to C at any x in $C \cap O$ can be realized by an r-ball. Then N_C^r is hypomonotone on O.*

Proof. Let v be a nonzero proximal normal to C at $x \in C \cap O$. We know that v can be realized by an r-ball. Therefore, as we observed in the introduction, this implies that

$$-\big\langle v, x' - x\big\rangle \geq -\frac{|v|}{2r}|x' - x|^2, \quad \forall x' \in C.$$

So, for $i = 1, 2$, let v_i be a proximal normal to C at x_i with v_i nonzero and $x_i \in O$. Then

$$-\big\langle v_1, x_2 - x_1\big\rangle \geq -\frac{|v_1|}{2r}|x_2 - x_1|^2,$$

and

$$-\big\langle v_2, x_1 - x_2\big\rangle \geq -\frac{|v_2|}{2r}|x_1 - x_2|^2,$$

which yields (even if $v_i = 0$)

$$\big\langle v_1 - v_2, x_1 - x_2\big\rangle \geq -\frac{1}{2r}\Big[|v_1| + |v_2|\Big]|x_1 - x_2|^2.$$

Therefore if $|v_i| < r$, then $\big\langle v_1 - v_2, x_1 - x_2\big\rangle \geq -|x_1 - x_2|^2$, which shows that S_C^r is hypomonotone on O with constant $\sigma = 1$. Here $S_C(x)$ is the set of proximal normals to C at x. From this and from [13, Theorem 2.4] we deduce that $S_C(x) = N_C(x)$ for all $x \in O$, and that N_C^r is hypomonotone on O with constant $\sigma = 1$. $\square$

Corollary 2.6. *If every nonzero proximal normal to C at any point x of C can be realized by an r-ball, then C is uniformly prox-regular with constant $1/r'$ for every $0 < r' < r$.*

Proof. Proposition 2.5 and Corollary 2.2 can be combined to show that C is prox-regular at every $x \in C$, and that every vector $v \in N_C(x)$ for $x \in C$ is actually a proximal normal vector. Let $0 < r' < r$. It follows from (1.2) that for every $x \in C$ and $v \in N_C(x)$ with $|v| < 1$, the point x is the unique closest point of C to $x + r'v$, which shows that C is uniformly prox-regular with constants $1/r'$. $\square$

The converse of Corollary 2.6 will be established later in Theorem 4.1. We will further show in Theorem 4.1 that a set C is proximally smooth with associated tube $U_C(r)$ if and only if the set C is uniformly prox-regular with constant $1/r'$ for every $0 < r' < r$.

3. Proof of the Main Theorem plus Corollaries

The proof of Theorem 1.3 is divided into several parts. The combination of Corollary 2.2 and Proposition 2.5 with the coming 3.1, 3.4–3.6 will yield it in full.

A crucial step in showing that the distance function is continuously differentiable on $O \setminus C$ for some open neighborhood O of $\bar{x}$ is that for some $\sigma > 0$, the function $d_C^2 + \sigma|\cdot|^2$ is convex on a neighborhood of $\bar{x}$. This property of d_C^2 can be obtained by noticing that d_C^2 is the Moreau-Yosida regularization of the indicator function δ_C with the norm square, and then applying the results of [6] in the finite-dimensional case and [14] in a general Hilbert space. However this property of d_C^2 can easily be established here without a direct appeal to those papers. Once we show that $d_C^2 + \sigma|\cdot|^2$ is convex on a neighborhood of $\bar{x}$, we will know that d_C^2 has proximal subgradients at all points in a neighborhood of $\bar{x}$. This will tell us in particular that the projection mapping is nonempty-valued. The implication from (g) to (e) in Theorem 1.3 will thereby be validated.

Proposition 3.1. *Assume that C is prox-regular at $\bar{x}$. Then*

(i) P_C is single-valued around $\bar{x}$.

(ii) d_C^2 is C^{1+} around $\bar{x}$.

(iii) For every $\sigma > 0$, there is a convex neighborhood O_σ of $\bar{x}$ on which the function $d_C^2 + \sigma|\cdot|^2$ is convex.

Moreover there is a neighborhood O of $\bar{x}$ such that P_C is monotone and Lipschitz continuous with $P_C = (I + N_C^r)^{-1}$ on O for some $r > 0$, while $D(d_C) = [I - P_C]/d_C$ on $O \setminus C$.

In the proof of Proposition 3.1 we employ Fréchet subgradients. Recall that v is a Fréchet subgradient to a function f at x, denoted $v \in \partial_F f(x)$, provided

$$\liminf_{y \to 0} \frac{f(x+y) - f(x) - \langle v, y \rangle}{|y|} \geq 0.$$

Proof. Assume that C is prox-regular at $\bar{x}$. According to Corollary 2.2, N_C^r is hypomonotone on some neighborhood O of $\bar{x}$ for some $r > 0$. Therefore there exist $\rho > 0$ and $\varepsilon > 0$ such that

$$\langle v_1 - v_2, x_1 - x_2 \rangle \geq -\rho |x_1 - x_2|^2 \tag{3.1}$$

whenever $v_i \in N_C(x_i)$ with $|v_i| < \varepsilon$ and $|x_i - \bar{x}| < \varepsilon$, $i = 1, 2$ (just pick $\varepsilon < r$ with $\operatorname{int} \mathbb{B}(\bar{x}, \varepsilon) \subset O$). We may also assume that (3.1) holds when $v_i \in \partial_F \delta_C(x_i)$ with $|v_i| < \varepsilon$ and $|x_i - \bar{x}| < \varepsilon$, $i = 1, 2$. This is because the set $\partial_F \delta_C(x)$ is always included in the closure of the convex hull of $N_C(x)$, which is the same as $N_C(x)$ in a neighborhood of $\bar{x}$ (according to Corollary 2.2).

We first show that in a neighborhood of $\bar{x}$, P_C is single-valued and Lipschitz continuous relative to its domain.

Claim. *Let $0 < \lambda \leq \rho$ with $\lambda < 2$. For $i = 1, 2$, let $x_i' \in P_C(x_i)$, where $|x_i - \bar{x}| < \lambda\varepsilon/2\rho$. Then*

$$|x_1' - x_2'| \leq \left(\frac{2}{2 - \lambda}\right) |x_1 - x_2|,$$

and

$$\langle x_1 - x_2, x_1' - x_2' \rangle \geq \big[1 - (\lambda/2)\big] |x_1' - x_2'|^2.$$

Proof of the Claim. It follows that $|x_i' - x_i| < \lambda\varepsilon/2\rho$ and that $|x_i' - \bar{x}| < \lambda\varepsilon/\rho \leq \varepsilon$. So, as $(2\rho/\lambda)(x_i - x_i')$ is a proximal normal to C at x_i' with $|(2\rho/\lambda)(x_i - x_i')| < \varepsilon$ we have

$$\langle (2\rho/\lambda)(x_1 - x_1') - (2\rho/\lambda)(x_2 - x_2'), x_1' - x_2' \rangle \geq -\rho |x_1' - x_2'|^2.$$

Therefore

$$\langle x_1 - x_2, x_1' - x_2' \rangle - |x_1' - x_2'|^2 \geq -(\lambda/2) |x_1' - x_2'|^2,$$

which means that $\langle x_1 - x_2, x_1' - x_2' \rangle \geq \big[1 - (\lambda/2)\big] |x_1' - x_2'|^2$. From this we conclude that

$$|x_1 - x_2| \geq \big[1 - (\lambda/2)\big] |x_1' - x_2'|.$$

This is the same as $|x_1' - x_2'| \leq \big(2/(2 - \lambda)\big) |x_1 - x_2|$. ☐

For $0 < \lambda \leq \rho$, with $\lambda < 2$, let x_1 and x_2 be two points of $\operatorname{int} \mathbb{B}\big(\bar{x}, \lambda\varepsilon/2\rho\big)$ (the open ball of radius $\lambda\varepsilon/2\rho$ around $\bar{x}$). Assume that the Fréchet subdifferential of d_C^2 is nonempty at x_1 and x_2. From [1, Theorem 11], we know that P_C is nonempty-valued at those points, and is in fact single-valued according to the Claim. Let $x_i' = P_C(x_i)$. We deduce from [15, Lemma 3.6] that

$$\partial_F \big(d_C^2\big)(x_i) \subset \partial_F \delta_C(x_i') \cap \big\{2(x_i - x_i')\big\}, \tag{3.2}$$

i.e., $\partial_F \big(d_C^2\big)(x_i) = 2(x_i - x_i')$ with $2(x_i - x_i') \in \partial_F \delta_C(x_i')$. From the Claim we have

$$\begin{aligned} 2\langle (x_1 - x_1') - (x_2 - x_2'), x_1 - x_2 \rangle &= 2|x_1 - x_2|^2 - 2\langle x_1' - x_2', x_1 - x_2 \rangle \\ &\geq 2|x_1 - x_2|^2 - 2|x_1' - x_2'||x_1 - x_2| \\ &\geq 2|x_1 - x_2|^2 - 2\left(\frac{2}{2 - \lambda}\right) |x_1 - x_2|^2 \\ &= -\left(\frac{2\lambda}{2 - \lambda}\right) |x_1 - x_2|^2. \end{aligned}$$

On the basis of [15, Theorem 3.8] we conclude that $d_C^2 + \big(\lambda/(2 - \lambda)\big) |\cdot|^2$ is convex on $\operatorname{int} \mathbb{B}\big(\bar{x}, \lambda\varepsilon/2\rho\big)$. This shows (iii), and it also implies that $\partial_F d_C^2(x) = \partial d_C^2(x)$ for all $x \in \operatorname{int} \mathbb{B}\big(\bar{x}, \lambda\varepsilon/2\rho\big)$. This in turn implies that $\partial_F d_C^2$ is nonempty-valued on $\operatorname{int} \mathbb{B}\big(\bar{x}, \lambda\varepsilon/2\rho\big)$, which shows that for all x in $\operatorname{int} \mathbb{B}\big(\bar{x}, \lambda\varepsilon/2\rho\big)$ the set $P_C(x)$ is nonempty. The Claim can then be applied at all such points to conclude that P_C

is single-valued, monotone, and Lipschitz continuous. From (3.2) and the fact that $d_C^2 + \big(\lambda/(2-\lambda)\big)|\cdot|^2$ is convex on $\operatorname{int}\mathbb{B}\big(\bar{x}, \lambda\varepsilon/2\rho\big)$ we conclude that the Gâteaux derivative mapping of d_C^2 on $\operatorname{int}\mathbb{B}\big(\bar{x}, \lambda\varepsilon/2\rho\big)$ is $2(I-P_C)$. This shows that d_C^2 is C^{1+} on $\operatorname{int}\mathbb{B}\big(\bar{x}, \lambda\varepsilon/2\rho\big)$.

Fix $\bar{\lambda} > 0$ with $\bar{\lambda} < \min\{\rho, 1\}$. Let $T(x) = N_C(x) \cap \operatorname{int}\mathbb{B}\big(0, \bar{\lambda}\varepsilon/2\rho\big)$ for $x \in C \cap \operatorname{int}\mathbb{B}\big(\bar{x}, \bar{\lambda}\varepsilon/2\rho\big)$, and $T(x) = \emptyset$ otherwise. There only remains to show that $(I+T)^{-1}(x) = P_C(x)$ when $x \in \operatorname{int}\mathbb{B}\big(\bar{x}, \bar{\lambda}\varepsilon/2\rho\big)$. It can easily be verified that $P_C(x) \subset (I+T)^{-1}(x)$ for the x's in question. We know that $P_C(x)$ is nonempty when $x \in \operatorname{int}\mathbb{B}\big(\bar{x}, \bar{\lambda}\varepsilon/2\rho\big)$; therefore the desired equality will be obtained once we show that $(I+T)^{-1}(x)$ is at most a singleton. For $i = 1, 2$, let $x_i' \in (I+T)^{-1}(x)$ with $x \in \operatorname{int}\mathbb{B}\big(\bar{x}, \bar{\lambda}\varepsilon/2\rho\big)$. It follows that $(x - x_i') \in T(x_i')$. By the choice of T we have $x_i' \in C \cap \operatorname{int}\mathbb{B}\big(\bar{x}, \bar{\lambda}\varepsilon/2\rho\big)$ with $2\rho|x - x_i'| < \bar{\lambda}\varepsilon < \varepsilon$ (because $\bar{\lambda} < 1$). With the help of (3.1) we have

$$-2\rho|x_1' - x_2'|^2 = \langle 2\rho(x - x_1') - 2\rho(x - x_2'), x_1' - x_2' \rangle \geq -\rho|x_1' - x_2'|^2,$$

which implies that $x_1' = x_2'$.

The formula for the derivative of d_C follows immediately from the formula for the derivative of d_C^2. □

In Theorem 1.3, (e) obviously implies (b), and that in turn implies (c). Before going any further we will need to show that if the distance function is Fréchet differentiable, then the projection mapping is strongly continuous.

Lemma 3.2. *Let C be a nonempty closed subset of H. If d_C^2 is Fréchet differentiable on some open set O, or equivalently d_C is Fréchet differentiable on $O \setminus C$, then P_C is (single-valued and) strongly continuous on O.*

Proof. As we saw in the proof of Proposition 3.1, the Fréchet derivative of d_C^2 at the point u is $2(u - P_C(u))$. The function $-d_C^2$ is equal to a convex function minus the norm square. Indeed, as observed by Asplund [16], one has

$$-d_C^2(u) = \sup_{x \in C}\big\{-|u-x|^2\big\} = -|u|^2 + \sup_{x \in C}\big\{2\langle u, x\rangle - |x|^2\big\}.$$

On the other hand, we know that the derivative of a convex function is strongly-weakly continuous—see [17] for example. The preceding observation therefore implies that the derivative of d_C^2, and hence P_C, is strongly-weakly continuous on O. Let x_k converge strongly to x, where $x \in O$. We have that $D(d_C^2)(x_k)$ converges weakly to $D(d_C^2)(x)$. We also have that $|D(d_C^2)(x_k)| = 2d_C(x_k)$ converges to $|D(d_C^2)(x)| = 2d_C(x)$ (because d_C is continuous—in fact it is Lipschitz). Thus, we have weak convergence and convergence of the norms; this implies strong convergence. □

We will also need the following fact.

Lemma 3.3. *Assume that d_C is Fréchet differentiable on a neighborhood of a point $\bar{u} \notin C$. Then there exists $\delta > 0$ such that whenever $u \in \operatorname{int}\mathbb{B}\big(\bar{u}, \delta\big)$ and $P_C(u) = x$, there exists $t > 0$ such that the point $u_t := u + t(u - x)$ likewise has $P_C(u_t) = x$.*

Proof. By Lemma 3.2, there exists $\varepsilon > 0$ such that P_C is single-valued and continuous on $\operatorname{int}\mathbb{B}\big(\bar{u}, 2\varepsilon\big)$, with d_C Fréchet differentiable there as well. Let $\sigma = \sup\big\{d_C(u)\,\big|\, u \in \operatorname{int}\mathbb{B}(\bar{u}, \varepsilon)\big\}$. Then for all $u \in \operatorname{int}\mathbb{B}(\bar{u}, \varepsilon)$ we have $\varepsilon \leq d_C(u) \leq \sigma$, and as long as $t \in (0, \varepsilon/\sigma)$ the point $u_t = u + t(u - P_C(u))$ lies in $\operatorname{int} \mathbb{B}(\bar{u}, 2\varepsilon)$; indeed,

$$\begin{aligned} |u_t - \bar{u}| &= |(u - \bar{u}) + t(u - P_C(u))| \leq |u - \bar{u}| + t|u - P_C(u)| \\ &\leq |u - \bar{u}| + t d_C(u) < \varepsilon + [\varepsilon/\sigma]\sigma = 2\varepsilon. \end{aligned}$$

Fix $\delta \in (0, \varepsilon)$ and $s \in (0, \delta/\sigma)$ (thus $s < 1$) such that

$$|P_C(u_s) - P_C(u)| < d_C(u) \text{ for all } u \in \operatorname{int} \mathbb{B}(\bar{u}, \delta), \tag{3.3}$$

which is possible by the continuity of d_C and P_C because $d_C(u) - |P_C(u_s) - P_C(u)| \to d_C(\bar{u}) > 0$ as $s \searrow 0$ and $u \to \bar{u}$. Then for all $u \in \operatorname{int} \mathbb{B}(\bar{u}, \delta)$ we have $s d_C(u) < \delta$, and moreover $d_C(u_s) > d_C(u)$, since by (3.3)

$$\begin{aligned} d_C(u_s) &= |u_s - P_C(u_s)| = |u + s(u - P_C(u)) - P_C(u_s)| \\ &= |(1 + s)(u - P_C(u_s)) + s(P_C(u_s) - P_C(u))| \\ &\geq (1 + s)|u - P_C(u_s)| - s|P_C(u_s) - P_C(u)| \\ &> (1 + s)d_C(u) - s d_C(u) = d_C(u). \end{aligned}$$

Consider now any $u \in \operatorname{int} \mathbb{B}(\bar{u}, \delta)$, and let $D = \{w \mid d_C(w) \geq d_C(u_s)\}$. In particular, $u_s \in D$. We know that there is a sequence of points u_k converging to u with $P_D(u_k) \neq \emptyset$; see [17]. Since $d_C(u_s) > d_C(u)$, we eventually have $d_C(u_s) > d_C(u_k)$, so $u_k \notin D$. Let $w_k \in P_D(u_k)$. Then $u_k - w_k$ is a nonzero proximal normal to D at w_k, and w_k must therefore be a boundary point of D and have $d_C(w_k) = d_C(u_s)$. Furthermore, for k sufficiently large we have w_k in the ball $\operatorname{int} \mathbb{B}(\bar{u}, 2\delta) \subset \operatorname{int} \mathbb{B}(\bar{u}, 2\varepsilon)$, because u_k eventually belongs to $\operatorname{int} \mathbb{B}(u, \delta)$ and

$$|u_k - w_k| = d_D(u_k) \leq |u_k - u_s| \to |u - u_s| = s|u - P_C(u)| = s d_C(u) < \delta. \tag{3.4}$$

In particular, then, $d_C(w_k) > \varepsilon$ and d_C is Fréchet differentiable at w_k with derivative $D(d_C)(w_k)$ given by $(w_k - P_C(w_k))/d_C(w_k)$, which has norm 1.

In view of the constraint representation of D in its definition, the half-space $H_k := \{v \mid \langle -D(d_C)(w_k), v \rangle \leq 0\}$ then gives the general tangent cone (contingent cone) to D at w_k, and since proximal normals must lie in the polar of this tangent cone, the vector $u_k - w_k$ must be a nonnegative scalar multiple of the normal vector $-(w_k - P_C(w_k))/d_C(w_k)$ to H_k. In fact we must have

$$u_k - w_k = -\lambda_k (w_k - P_C(w_k))/d_C(w_k) \quad \text{with } \lambda_k = |u_k - w_k| = d_D(u_k) > 0,$$

where eventually $\lambda_k < \delta < \varepsilon$ by (3.4); hence $\lambda_k < \varepsilon \leq d_C(u) < d_C(u_s) = d_C(w_k)$. In terms of $r_k = d_D(u_k)/d_C(w_k)$ we then have $r_k \in (0, 1)$ and $u_k = (1 - r_k)w_k + r_k P_C(w_k)$. Thus, u_k belongs to the line segment joining w_k with $P_C(w_k)$, and in consequence we have $P_C(u_k) = P_C(w_k)$ and $\lambda_k = d_C(w_k) - d_C(u_k) = d_C(u_s) - d_C(u_k)$. This gives us

$$w_k = u_k + t_k(u_k - P_C(u_k))$$

with

$$t_k := \frac{r_k}{1 - r_k} = \frac{d_D(u_k)}{d_C(w_k) - d_D(u_k)} = \frac{d_C(u_s) - d_C(u_k)}{d_C(u_k)}.$$

Since t_k converges to $t := (d_C(u_s) - d_C(u))/d_C(u) > 0$, we obtain that w_k converges to u_t and $P_C(w_k)$ converges to $P_C(u_t)$. But $P_C(w_k) = P_C(u_k) \to P_C(u)$. Hence for this t we have $P_C(u_t) = P_C(u)$, as desired. $\square$

And now we show (part (i) below) that (c) implies (h) in Theorem 1.3. Note also that we could add part (ii) of Proposition 3.4 to the list of equivalent properties in Theorem 1.3.

Proposition 3.4. *Assume d_C is Fréchet differentiable on $O \setminus C$ for some open neighborhood O of $\bar{x}$. Then there exists $\lambda > 0$ such that:*

(i)

$$\left.\begin{array}{r} x = P_C(u),\ x \neq u \\ 0 < |u - \bar{x}| < \lambda \end{array}\right\} \implies x = P_C(u') \text{ for } u' = x + \lambda \frac{u - x}{|u - x|}. \tag{3.5}$$

(ii) For $D = \{y \,|\, d_C(y) \geq \lambda\}$ and for any $u \in \operatorname{int}\mathbb{B}(\bar{x}, \lambda) \setminus C$ one has $d_C(u) + d_D(u) = \lambda$.

Proof. By Lemma 3.2 we may assume that there exists $\lambda > 0$ such that d_C^2 is Fréchet differentiable on $\operatorname{int}\mathbb{B}(\bar{x}, 2\lambda)$ while P_C is single-valued and strongly continuous there. Let $x = P_C(u)$ with $|u - \bar{x}| < \lambda$ and $u \notin C$. It follows that $0 < |x - u| < \lambda$. Since d_C is Fréchet differentiable on a neighborhood of u, we can apply Lemma 3.3 to get the existence of $s > 0$, with $s < \lambda$, such that for all $t \in (0, s)$ we have $P_C(u_t) = x$, where $u_t := u + t(u - x)/|u - x|$. Note that for all such t, one has $u_t \notin C$. Let λ_0 be the supremum over all $t \in [0, \lambda]$ such that $P_C(u_t) = x$. The continuity of P_C over $\operatorname{int}\mathbb{B}(\bar{x}, 2\lambda)$ (note that $|u_t - \bar{x}| < 2\lambda$) implies that the supremum is attained, and since $u_t \in \operatorname{int}\mathbb{B}(\bar{x}, 2\lambda)$ for $t \in [0, \lambda]$ one has $u_{\lambda_0} \in \operatorname{int}\mathbb{B}(\bar{x}, 2\lambda)$. We cannot have $\lambda_0 < \lambda$, because when we apply Lemma 3.3 with u_{λ_0} in place of u we arrive at a contradiction. Note that

$$u_t = x + (|u - x| + t)\frac{(u - x)}{|u - x|},$$

and since $\lambda_0 = \lambda$ we obtain (3.5). Let $u' := x + \lambda(u - x)/|u - x|$. Since $x \in C$ $(x = P_C(u))$ and $u' \in D := \{y \,|\, d_C(y) \geq \lambda\}$, we have

$$d_C(u) + d_D(u) \leq |u - x| + |u - u'| = \lambda. \tag{3.6}$$

On the other hand, $d_C(y) \geq \lambda$ for any $y \in D$, which implies that

$$|y - u| \geq d_C(y) - d_C(u) \geq \lambda - d_C(u). \tag{3.7}$$

The combination of (3.6) and (3.7) yields (ii). □

It is clear in Theorem 1.3 that (h) implies (f) . But property (f) implies, by combining Corollary 2.2 with Proposition 2.5, that C is prox-regular at $\bar{x}$. We therefore have the equivalence between (a), (b), (c), (e), (f), (g) and (h). We now turn our attention to adding (d), (i), and (k) to the list.

Proposition 3.5. *Consider a closed set $C \subset H$, a point $\bar{x} \in C$ and a neighborhood O of $\bar{x}$. The following properties are equivalent:*

(i) d_C is continuously differentiable on $O \setminus C$.

(ii) d_C is Fréchet differentiable on $O \setminus C$.

(iii) d_C is Gâteaux differentiable on $O \setminus C$ and P_C is non-empty on O.

(iv) P_C is single-valued and strongly-weakly continuous on O.

If the set C is weakly closed relative to O, then one can add the following to the set of equivalent properties:

(v) P_C is single-valued on O.

Proof. First recall that the Fréchet derivative of d_C^2 at an arbitrary point u is $2(u - P_C(u))$ (see the proof of Proposition 3.1). From this and Lemma 3.2 we conclude that (i) is equivalent to (ii), and that (ii) implies (iii). Let

$$f(u) = \sup_{x \in C} \left\{ 2\langle u, x \rangle - |x|^2 \right\}.$$

The function f is convex, and we saw in the proof of Lemma 3.2 that $f(\cdot) + d_C^2(\cdot) = |\cdot|^2$. From Hiriart-Urruty [24] we know that

$$P_C(u) \subset (\tfrac{1}{2})\partial f(u) \quad \text{for any} \quad u \in H. \tag{3.8}$$

Since the derivative of a convex function is strongly-weakly continuous and equals its subdifferential, we conclude that (iii) implies (iv) (under (iii), the Gâteaux derivative of $(\frac{1}{2})f$ on $O \setminus C$ is P_C). To show that (iv) implies (i), first use Phelps [25, Lemma 7.7] to conclude that P_C is maximal monotone on O. Since the subdifferential of a convex function is monotone (see for example [25, Theorem 3.24]), we get from (3.8) that the Gâteaux derivative of $(\frac{1}{2})f$ equals $P_C(u)$ for any $u \in O$. This shows that d_C^2 is continuously differentiable on O and that d_C is continuously differentiable on $O \setminus C$ (strong-weak continuity of the Gâteaux derivative of d_C^2 implies continuous differentiability—see the proof of Lemma 3.2).

Finally, it is an easy exercise to show that when a set is weakly closed and has single-valued projections then its projection mapping is strongly-weakly continuous. Therefore when the set C is weakly closed (iv) and (v) are equivalent. $\square$

To complete the proof of Theorem 1.3 we need only show that prox-regularity is equivalent to the Shapiro property.

Proposition 3.6. *A closed subset C of H is prox-regular at $\bar{x}$ if and only if C has the Shapiro property at $\bar{x}$.*

Proof. Assume that the set C is prox-regular at $\bar{x}$. There exist $r > 0$ and a neighborhood O of $\bar{x}$ such that every proximal normal to C at x in $C \cap O$ can be realized by an r-ball. This means that for every unit normal v to C at x in $C \cap O$ we have $(1/2r)|x' - x|^2 \geq \langle v, x' - x \rangle$ for every $x' \in C$. From this we conclude that C has the Shapiro property at $\bar{x}$, since in this context the cones $T_C(x)$ and $N_C(x)$ are polar to each other, and therefore, by Fenchel duality, we have

$$\sup_{v \in N_C(x), |v|=1} \langle v, x' - x \rangle = d_{T_C(x)}(x' - x).$$

Now assume that C satisfies the Shapiro property at $\bar{x}$ with constant k and neighborhood O. As in Shapiro [4, Lemma 2.1] we conclude that $\langle v_1 - v_2, x_1 - x_2 \rangle \geq -2k|x_1 - x_2|^2$ whenever v_i is a proximal normal to C at x_i with x_i in $O \cap C$ and $|v_i| \leq 1$. As in Proposition 2.5 we deduce from [13, Theorem 2.4] that the set of proximal normals to C at $x \in O \cap C$ is equal to the normal cone $N_C(x)$. Therefore N_C^1 is hypomonotone on O, and we conclude from Corollary 2.2 that C is prox-regular at $\bar{x}$. $\square$

Now that the entire proof of Theorem 1.3 has been put together, we turn to a couple of consequences of this theorem which give further characterizations of prox-regular sets.

Corollary 3.7 (of Theorem 1.3)**.** *For a closed set C of H, the following are equivalent:*

(a) The set C is prox-regular at $\bar{x}$.

(b) For all $\sigma > 0$, the function $d_C^2 + \sigma|\cdot|^2$ is convex on a convex neighborhood O_σ of $\bar{x}$.

(c) For some $\sigma > 0$, the function $d_C^2 + \sigma|\cdot|^2$ is convex on a convex neighborhood of $\bar{x}$.

Proof. We already observed in Proposition 3.1 that for all $\sigma > 0$, the function $d_C^2 + \sigma|\cdot|^2$ is convex on some open neighborhood O_σ of $\bar{x}$ when the set is prox-regular at $\bar{x}$. On the other hand, if for some $\sigma > 0$ the function $d_C^2 + \sigma|\cdot|^2$ is convex on a neighborhood of $\bar{x}$, then the function d_C^2 has Fréchet subgradients at all points in a neighborhood of $\bar{x}$; but this is also true of $-d_C^2$ since we saw in the proof of Lemma 3.2 that $-d_C^2 + |\cdot|^2$ is a convex function. Therefore d_C^2 is Fréchet differentiable on a neighborhood of $\bar{x}$, so C is prox-regular at $\bar{x}$ by Theorem 1.3. □

Corollary 3.8 (of Theorem 1.3)**.** *For a closed set C of H, the following are equivalent:*

(a) C is prox-regular at $\bar{x}$.

(b) $\partial_p d_C(x)$ is nonempty at all points x in a neighborhood of $\bar{x}$.

(c) $\partial_F d_C(x)$ is nonempty at all points x in a neighborhood of $\bar{x}$.

Proof. Assume that C is prox-regular at $\bar{x}$. From Theorem 1.3(e) we have that d_C is C^{1+} on $O \setminus C$ for some open neighborhood O of $\bar{x}$. This implies that $\partial_p d_C(x)$ is nonempty at all points x of $O \setminus C$. On the other hand, 0 is always a proximal subgradient to d_C at points x in C. This shows that (b) follows from (a).

Obviously (b) implies (c).

We will show that (c) implies (a) by verifying that d_C^2 is Fréchet differentiable near $\bar{x}$, which implies that d_C is Fréchet differentiable on $O \setminus C$ for some open neighborhood O of $\bar{x}$. This is easily established with the help of Lemma 3.9 below. Indeed, according to Lemma 3.9, $\partial_F d_C^2$ is nonempty-valued on a neighborhood of $\bar{x}$. On the other hand, we know that for all $x \in H$, $\partial_F(-d_C^2)(x)$ is nonempty (see the proof of Lemma 3.2). From this we conclude that d_C^2 is Fréchet differentiable on a neighborhood of $\bar{x}$. □

Lemma 3.9. *If $v \in \partial_F d_C(u)$, then $2d_C(u)v \in \partial_F d_C^2(u)$.*

Proof. For each $\varepsilon > 0$, we have (by the definition of a Fréchet subgradient) that

$$\langle v, x-u\rangle \le d_C(x) - d_C(u) + \varepsilon|x-u|$$

for all x in a neighborhood of u. Therefore

$$\begin{aligned}\langle 2d_C(u)v, x-u\rangle &\le 2d_C(u)d_C(x) - 2d_C(u)d_C(u) + 2d_C(u)\varepsilon|x-u| \\ &= d_C(x)^2 - d_C(u)^2 - \big(d_C(x) - d_C(u)\big)^2 + 2d_C(u)\varepsilon|x-u| \\ &\le d_C(x)^2 - d_C(u)^2 + 2d_C(u)\varepsilon|x-u|.\end{aligned}$$

From this we conclude that $2d_C(u)v \in \partial_F d_C^2(u)$. □

4. Proximally Smooth Sets

The local theory that has been developed so far will now be applied to the global setting of Clarke, Stern and Wolenski [2] to obtain certain of their characterizations, along with some new ones.

Theorem 4.1. *Let C be a closed subset of H and let $r > 0$. The following properties are equivalent:*

(a) C is uniformly prox-regular with constant $1/r'$ for every $0 < r' < r$.

(b) d_C is continuously differentiable on $U_C(r)$.

(c) d_C is Fréchet differentiable on $U_C(r)$.

(d) d_C is Gâteaux differentiable on $U_C(r)$, and P_C is nonempty-valued on $U_C(r)$.

(e) d_C^2 is C^{1+} on $U_C(r)$, i.e., differentiable with locally Lipschitz continuous derivative mapping (in fact with Lipschitz continuous derivative on $U_C(\rho)$ for each positive $\rho < r$).

(f) Every nonzero proximal normal to C at any point x of C can be realized by an r-ball.

(g) Whenever $x_i \in C$ and $v_i \in N_C^r(x_i)$, one has

$$\langle v_1 - v_2, x_1 - x_2\rangle \geq -|x_1 - x_2|^2.$$

If $u \in U_C(r)$ and $x = P_C(u)$, then $x = P_C(u')$ for $u' = x + r(u - x)/|u - x|$.

(i) P_C is single-valued and strongly-weakly continuous on $U_C(r)$.

(j) $d_{T_C(x)}(x' - x) \leq \frac{1}{2r}|x' - x|^2$ whenever x', x are in C (global Shapiro property).

Then P_C is (single-valued) monotone on $U_C(r)$ and Lipschitz continuous on $U_C(\rho)$ for any $\rho \in (0, r)$, with $P_C = (I + N_C^r)^{-1}$ on $U_C(r)$. Moreover, $D(d_C) = [I - P_C]/d_C$ on $U_C(r)$.

If C is weakly closed (which is always the case when the space H is finite-dimensional), then one can add the following to the list of equivalent properties:

(k) P_C is single-valued on $U_C(r)$.

In this theorem, the equivalence between (b) (i.e., the definition of proximal smoothness), (d), (e), and (k) (when the set is weakly closed), along with the fact that P_C is single-valued, monotone and Lipschitz continuous under these equivalent assumptions, was shown by Clarke, Stern and Wolenski [2]. They also proved that proximal smoothness is equivalent to (f) under the extra assumption that $P_C(u) \neq \emptyset$ for each $u \in U_C(r)$. The addition of (a), (c), (f), (g), (h), (i) and (j) to the list of equivalent properties is new. Also new is the formula for P_C in terms of a truncation of the normal cone mapping N_C. Our arguments are quite different than those of [2] and provide an easier way of obtaining the equivalence between (b) and (f).

The following will be used in the proof of Theorem 4.1.

Lemma 4.2. *Let C be a closed subset of H, and let $0 < \rho < r < \infty$. Assume that*

$$\langle v_1 - v_2, x_1 - x_2\rangle \geq -|x_1 - x_2|^2$$

whenever $x_i \in C$, $v_i \in N_C^r(x_i)$. Then:

(i) For x_i in $P_C(u_i)$ with $u_i \in U_C(\rho)$, one has

$$|x_1 - x_2| \leq \big(r/(r - \rho)\big)|u_1 - u_2|$$

and

$$\langle u_1 - u_2, x_1 - x_2\rangle \geq \big(1 - (\rho/r)\big)|x_1 - x_2|^2.$$

(ii) P_C is single-valued and monotone on $U_C(r)$ and Lipschitz continuous on $U_C(\rho)$. Moreover, $P_C = \big(I + N_C^r\big)^{-1}$ on $U_C(r)$.

(iii) d_C^2 is C^{1+} on $U_C(r)$, and the derivative of d_C is equal to $\big(I - P_C\big)/d_C$ on $U_C(r)$.

(iv) The function $d_C^2 + \big(\rho/(r - \rho)\big)|\cdot|^2$ is convex on any convex subset included in $U_C(\rho)$.

Proof. (i) The vector $(r/\rho)(u_i - x_i)$ is a proximal normal to C at x_i, and it satisfies $|(r/\rho)(u_i - x_i)| < r$. Therefore the assumptions ensure that

$$\langle (r/\rho)(u_1 - x_1) - (r/\rho)(u_2 - x_2), x_1 - x_2 \rangle \geq -|x_1 - x_2|^2,$$

and

$$\langle u_1 - u_2, x_1 - x_2 \rangle - |x_1 - x_2|^2 \geq -(\rho/r)|x_1 - x_2|^2,$$

which means that $\langle u_1 - u_2, x_1 - x_2 \rangle \geq \big(1 - (\rho/r)\big)|x_1 - x_2|^2$. From this we conclude that $|u_1 - u_2| \geq \big(1 - (\rho/r)\big)|x_1 - x_2|$. This can also be written as $|x_1 - x_2| \leq \big(r/(r - \rho)\big)|u_1 - u_2|$. This finishes the proof of (i).

(ii)–(iv) As in the proof of Proposition 3.1, we deduce from (i) that $d_C^2 + \big(\rho/(r - \rho)\big)|\cdot|^2$ is convex on any convex subset of $U_C(\rho)$. This implies that the Fréchet subdifferential of d_C^2 is nonempty on $U_C(r)$. From this we conclude that $P_C(u)$ is nonempty for every $u \in U_C(r)$. Part (i) can then be used (as in Proposition 3.1) to show that P_C is single-valued, monotone, and locally Lipschitz continuous on $U_C(r)$ (in fact, Lipschitz continuous on $U_C(\rho)$). We also have that d_C^2 is C^{1+} on $U_C(r)$.

If $x = P_C(u)$ for $u \in U_C(r)$, then one easily shows that $x \in \big(I + N_C^r\big)^{-1}(u)$. Therefore the desired equality $\big(I + N_C^r\big)^{-1}(u) = P_C(u)$ will be obtained once we show that $\big(I + N_C^r\big)^{-1}(u)$ is at most a singleton. For $i = 1, 2$, let $x_i \in (I + N_C^r)^{-1}(u)$ with $u \in U_C(r)$. It follows that $(u - x_i) \in N_C^r(x_i)$ and that $|u - x_i| < r$. Thus, there exist $s > 1$ such that $s|u - x_i| < r$ (and we still have $s(u - x_i) \in N_C(x_i)$). We therefore have

$$-s|x_1 - x_2|^2 = s\langle (u - x_1) - (u - x_2), x_1 - x_2 \rangle \geq -|x_1 - x_2|^2,$$

which implies that $x_1 = x_2$. □

Proof of Theorem 4.1. From Lemma 4.2 we conclude that (g) implies (e). This lemma also gives us that P_C is single-valued, monotone on $U_C(r)$ and Lipschitz continuous on $U_C(\rho)$ for any $\rho \in (0, r)$ with $P_C = (I + N_C^r)^{-1}$ on $U_C(r)$, whereas $D(d_C) = [I - P_C]/d_C$ on $U_C(r)$.

Obviously (e) implies (b), which in turn is equivalent (by Proposition 3.5) to (c), (d), and (i).

(c) implies (h): By Lemma 3.2 we have that P_C is single-valued and strongly continuous on $U_C(r)$. Let $x = P_C(u)$ with $u \in U_C(r)$. Since the function d_C is Fréchet differentiable on a neighborhood of u, we can apply Lemma 3.3 to get the existence of $s > 0$ such that $P_C(u_t) = x$, where $u_t := u + t(u - x)/|u - x|$ and $0 < t < s$. Let λ_0 be the supremum over all $t \in [0, (r - d_C(u))]$ such that $P_C(u_t) = x$. The continuity of P_C on $U_C(r)$ (note that $u_t \in U_C(r)$) implies that the supremum is attained. We cannot have $\lambda_0 < (r - d_C(u))$, because this would contradict Lemma 3.3. Note that $u_t = x + (d_C(u) + t)(u - x)/|u - x|$. Since $\lambda_0 = (r - d_C(u))$, we obtain (h).

(h) obviously implies (f).

(f) implies (g): This follows from Proposition 2.5.

We now know that (b)–(i) are equivalent. Property (j) can also be added to this list of equivalent properties, since one can easily show, as in the proof of Proposition 3.6, that (f) implies (j) and that (j) implies (g).

The fact that (f) implies (a) was noted in Corollary 2.6. Now assume that C is uniformly prox-regular with constant $1/r'$ for every $0 < r' < r$; we will show that

(f) is fulfilled. According to the definition of uniform prox-regularity, we have for all $x \in C$ and $v' \in N_C(x)$ with $|v'| < 1$ that x is the unique nearest point of C to $x + r'v'$. This means that $|x' - (x + r'v')| \geq r'|v'|$ for every $x' \in C$. If we fix $v \in N_C(x)$ with $|v| = 1$ and take the limit as (r', v') converges to (r, v), we obtain that $|x' - (x + rv)| \geq r$ for every $x' \in C$. This gives (f).

When the set C is weakly closed and has single-valued projections on $U_C(r)$, we obtain, as in Proposition 3.5, that P_C is strongly-weakly continuous on $U_C(r)$. This completes the proof of Theorem 4.1. ☐

Remark. Another way to show that the Fréchet differentiability of d_C on $U_C(r)$ (for some $r > 0$) is sufficient for the proximal smoothness of C is to invoke [21, Theorem 3, Corollaries 2 and 3]. Indeed, as was observed by Asplund ([16, page 236]), the Fréchet differentiability of $|\cdot|^2 - d_C^2$ at a point x is equivalent to the (norm-to-norm-) continuous differentiability of this same function at the point x. From this we can conclude that d_C is continuously differentiable on the tube $U_C(r)$.

The proof of the following corollary parallels that of Corollary 3.8.

Corollary 4.3 ([2, Thm. 4.1])**.** *Let C be a closed subset of H and r a positive number. The following properties are equivalent:*

(a) C is proximally smooth with associated tube $U_C(r)$.

(b) $\partial_p d_C(x)$ is nonempty at all points x of $U_C(r)$.

(c) $\partial_F d_C(x)$ is nonempty at all points x of $U_C(r)$.

Another new characterization of proximally smooth sets comes next.

Corollary 4.4. *The set C is proximally smooth if and only if there exist some $\sigma > 0$ and some tube around C such that $d_C^2 + \sigma|\cdot|^2$ is convex on any convex subset of this tube.*

Proof. When the set is proximally smooth with associated tube $U_C(r)$, we saw in Lemma 4.2 that $d_C^2 + \big(\rho/(r - \rho)\big)|\cdot|^2$ is convex on any convex subset of $U_C(\rho)$ for $0 < \rho < r$. The rest of the proof parallels that of Corollary 3.7, and is omitted. ☐

Theorem 4.1 enables us to recover two well-known results concerning Chebyshev sets, i.e., sets C for which P_C is single-valued everywhere. Part (a) of the following Corollary 4.5 was originally proved by Motzkin [22] in the finite-dimensional case and by Klee [23] in the infinite-dimensional case, and it was re-derived by Clarke, Stern and Wolenski in [2]. Part (b) of Corollary 4.5 is due to Asplund [16]. For a thorough discussion of the "Chebyshev problem", see Hiriart-Urruty [24].

Corollary 4.5. *(a) A nonempty, weakly closed set $C \subset H$ is convex if and only if its projection mapping P_C is single-valued on H.*

(b) A closed set $C \subset H$ is convex if and only if its projection mapping P_C is single-valued and strongly-weakly continuous on H.

Proof. It is well known that if C is nonempty, convex and closed (which is the same as being weakly closed under convexity), then P_C is single-valued and continuous (in fact nonexpansive, i.e., Lipschitz continuous with constant 1). On the other hand (under the assumption that C is weakly closed), if P_C is single-valued on H, we get from Theorem 4.1(g) that $\langle v_1 - v_2, x_1 - x_2 \rangle \geq 0$ when $v_i \in N_C(x_i)$. This, according to [15, Thm. 3.8], shows that C is convex. Under the assumptions that

P_C is single-valued and strongly-weakly continuous on H we can also conclude from Theorem 4.1(g) and [15, Thm. 3.8] that C is convex. □

The following was established in [2, Cor. 4.15] in the finite-dimensional setting.

Corollary 4.6. *If C is proximally smooth, then at every point $x \in C$ the normal cone $N_C(x)$ is closed and convex, with every $v \in N_C(x)$ actually being a proximal normal.*

Proof. Theorem 4.1 shows that C is prox-regular at any $x \in C$. We can then apply Corollary 2.2 to obtain the desired conclusion. □

References

[1] J. M. Borwein and J. R. Giles, *The proximal normal formula in Banach space*, Trans. Amer. Math. Soc. **302** (1987), 371–381. MR **88m:**49013

[2] F. H. Clarke, R. J Stern and P. R. Wolenski, *Proximal smoothness and the lower-C^2 property*, J. Convex Analysis, **2** (1995), 117–144. MR **96j:**49014

[3] J.-P. Vial, *Strong and weak convexity of sets and functions*, Math. Ops. Res. **8** (1983), 231–259. MR **84m:**90107

[4] A. S. Shapiro, *Existence and differentiability of metric projections in Hilbert spaces*, SIAM J. Optimization **4** (1994), 130–141. MR **94m:**90111

[5] H. Federer, *Curvature measures*, Trans. Amer. Math. Soc. **93** (1959), 418–491. MR **22:**961

[6] R. A. Poliquin and R. T. Rockafellar, *Prox-regular functions in variational analysis*, Trans. Amer. Math. Soc. **348** (1996), 1805–1838. MR **96h:**49039

[7] R. A. Poliquin and R. T. Rockafellar, *Generalized Hessian properties of regularized nonsmooth functions*, SIAM J. Optimization **6** (1996), 1121–1137. MR **97j:**49025

[8] R. A. Poliquin and R. T. Rockafellar, *Second-order nonsmooth analysis in nonlinear programming*, Recent Advances in Nonsmooth Optimization (D. Du, L. Qi and R. Womersley, eds.), World Scientific Publishers, 1995, pp. 322–350. MR **98e:**49043

[9] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*, Springer-Verlag, 1997. MR **98m:**49001

[10] C. Combari, A. Elhilali Alaoui, A. Levy, R. A. Poliquin and L. Thibault, *Convex composite functions in Banach spaces and the primal-lower-nice property*, Proc. Amer. Math. Soc. **126** (1998), 3701–3708. MR **99b:**49016

[11] R. A. Poliquin, *Integration of subdifferentials of nonconvex functions*, Nonlinear Anal. Th. Meth. Appl. **17** (1991), 385–398. MR **92j:**49008

[12] 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), 861–874. MR **93a:**49013

[13] 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), 1269–1294. MR **95k:**49035

[14] L. Thibault and D. Zagrodny, *Integration of subdifferentials of lower semicontinuous functions*, J. Math. Anal. Appl. **189** (1995), 33–58. MR **95i:**49032

[15] R. Correa, A. Jofré and L. Thibault, *Characterization of lower semicontinuous convex functions*, Proc. Amer. Math. Soc. **116** (1992), 61–72. MR **92k:**49027

[16] E. Asplund, *Čebyšev sets in Hilbert spaces*, Trans. Amer. Math. Soc. **144** (1969), 235–240. MR **40:**6238

[17] J.-P. Aubin and I. Ekeland, *Applied Nonlinear Analysis*, Wiley-Interscience, 1984. MR **87a:**58002

[18] K.-S. Lau, *Almost Chebyshev subsets in reflexive Banach spaces*, Indiana Univ. Math. J. **27** (1978), 791–795. MR **58:**23286

[19] J. M. Borwein and H. M. Strojwas, *Proximal analysis and boundaries of closed sets in Banach spaces II: Applications*, Canad. J. Math. **39** (1987), 428–472. MR **88f:**46034

[20] F. H. Clarke, *Optimization and Nonsmooth Analysis*, SIAM Publications, Philadelphia, 1990 (originally published in 1983). MR **85m:**49002; MR **91e:**49001

[21] E. Asplund and R. T. Rockafellar, *Gradients of convex functions*, Trans. Amer. Math. Soc. **121** (1968), 31–47. MR **39:**1968

[22] T. S. Motzkin, *Sur quelques propriétés caractéristiques des ensembles convexes*, Att. R. Acad. Lincei, Rend. **21** (1935), 562–567.
[23] V. Klee, *Convexity of Chebyshev sets*, Math. Annalen **142** (1961), 292–304. MR **22:**12367
[24] J.-B. Hiriart-Urruty, *Ensembles de Tchebychev vs. ensembles convexes: l'état de la situation vu par l'analyse convexe non lisse*, Ann. Sci. Math. Que. **22** (1998), 47–62. MR **99f:**49018
[25] R. R. Phelps, *Convex Functions, Monotone Operators and Differentiability*, Lecture Notes in Mathematics, Vol. 1364 2nd edition, Springer-Verlag, 1993. MR **94f:**46055

Department of Mathematical Sciences, University of Alberta, Edmonton, Alberta, Canada T6G 2G1
E-mail address: rene.poliquin@ualberta.ca

Department of Mathematics 354350, University of Washington, Seattle, Washington 98195-4350
E-mail address: rtr@math.washington.edu

Laboratoire d'Analyse Convexe, Université Montpellier II, 34095 Montpellier, France
E-mail address: thibault@math.univ-montp2.fr