

NECESSARY CONDITIONS  
FOR CONSTRAINED OPTIMIZATION PROBLEMS  
WITH SEMICONTINUOUS AND CONTINUOUS DATA

JONATHAN M. BORWEIN, JAY S. TREIMAN, AND QIJI J. ZHU

ABSTRACT. We consider nonsmooth constrained optimization problems with semicontinuous and continuous data in Banach space and derive necessary conditions *without* constraint qualification in terms of smooth subderivatives and normal cones. These results, in different versions, are set in reflexive and smooth Banach spaces.

1. INTRODUCTION

We consider constrained optimization problems with *semicontinuous inequality and continuous equality data* in reflexive Banach spaces or with entirely *continuous data* in smooth Banach spaces. Our necessary conditions for the optimization problem are stated in terms of smooth subderivatives and normal cones. There is a large literature on necessary conditions for constrained optimization problems, given in terms of various generalized derivatives, usually in the presence of additional constraint qualifications. The typical method used to derive necessary conditions for constrained optimization problems in nonsmooth analysis is to convert the constrained optimization problem to an unconstrained problem by using nonsmooth penalties. The problem then becomes one of expressing the necessary conditions in terms of generalized derivatives. Usually, this is achieved by using calculus rules for the appropriate generalized (sub)derivatives. Such necessary conditions have been derived in terms of Clarke's generalized gradient [6, 7], Mordukhovich's coderivative [20] and Ioffe's geometric subderivative [12, 13], among others. All of these necessary conditions require additional constraint qualifications because the underlying calculus rules demand them.

Recently, it has become clear that in spaces with a smooth renorm the three generalized derivatives mentioned above can be characterized in terms of sequential limits of smooth subderivatives. (See [1, 4, 29] for details. Such characterizations for Clarke's generalized gradient and normal cone with proximal subderivatives and proximal normals has been known since the early 1980's [6, 24]. See also [26, 27] for earlier discussion with  $\varepsilon$ -subderivatives.) Also, calculus rules for the limiting generalized (sub)derivatives can be viewed as limiting forms of the analogous calculus

---

Received by the editors September 14, 1995 and, in revised form, August 9, 1996.

1991 *Mathematics Subject Classification*. Primary 49J52; Secondary 49J40, 49J50, 58C20.

*Key words and phrases*. Constrained optimization problems, multipliers, nonsmooth analysis, subderivatives, normals, fuzzy calculus.

Research of the first author was supported by NSERC and by the Shrum Endowment at Simon Fraser University.

rules for smooth subderivatives. The constraint qualifications for the calculus rules of the various generalized (sub)derivatives are required mainly because of the limiting process. It is natural, therefore, to ask whether we can express the necessary conditions for a constrained optimization problem in terms of smooth subderivatives using only calculus rules that do not require a constraint qualification. We pursue that goal in this paper.

We establish this type of necessary condition in reflexive Banach spaces in terms of Fréchet subderivatives and Fréchet normals. The limiting forms of these necessary conditions recover some familiar necessary conditions. Since we assume only (lower semi-)continuity of the data, this result appears to be new even in finite dimensions. In deriving this necessary condition we follow the basic idea in [28]. We use the indicator functions to the level sets of the constraints as penalization functions and then use “fuzzy sum rules” for the Fréchet subderivative. The critical step is to establish the relationship between the normal cone to such a level set and the subderivative of the corresponding function. In super-reflexive Banach spaces (Hilbert spaces) similar results hold with the Fréchet subderivative and Fréchet normal replaced by the  $s$ -Hölder subderivative (proximal subderivative) and  $s$ -Hölder normal (proximal normal) respectively. We indicate the necessary modifications at the end of Section 2. We also give a full proof of the critical step in Hilbert space because it is simple and intuitive. We hope that it well communicates the idea behind the more general proofs.

We then turn to more general smoothable Banach spaces. Here we use sum rules for smooth subderivatives to deduce necessary conditions given in terms of the smooth normal cone to the graph of a mapping which has the cost and the constraint functions as its components. There are no additional constraint qualification conditions for this result. This necessary condition, however, is not expressed in terms of the individual constraint functions and the cost function. In order to derive a necessary condition given in terms of the subderivative of those component functions, we still must impose a mild constraint qualification. Whether or not a complete generalization of the reflexive Banach space result can be achieved in more general smooth Banach spaces remains an open question.

This paper is written to allow the reading of the reflexive Banach space results (Section 2) and the general smooth Banach space results (Section 3) separately.

## 2. NECESSARY CONDITIONS IN REFLEXIVE SPACES

Let  $X$  be a reflexive Banach space with closed unit ball  $B_X$  and with continuous real dual  $X^*$ . Then  $X$  has an equivalent renorm  $\|\cdot\|$  that is simultaneously *locally uniformly rotund* and *Fréchet smooth* [10]. Recall that a lower semicontinuous function  $f : X \rightarrow R \cup \{\infty\}$  is *Fréchet subdifferentiable* at  $x$  with (viscosity) *Fréchet subderivative*  $x^* \in X^*$  if  $f$  is finite at  $x$  and there exists a Fréchet smooth (*concave*) function  $g$  such that  $x^* = \nabla g(x)$  and  $f - g$  attains a local minimum at  $x$ . We denote the set of Fréchet subderivatives of  $f$  at  $x$  by  $D_F f(x)$ . For a closed subset  $C$  of  $X$  the *Fréchet normal cone* to  $C$  at  $x \in C$  is defined by  $N_F(C, x) := D_F \delta_C(x)$ . Here  $\delta_C$  is the indicator function of  $C$ .

Let  $C \subset X$  and  $g_i : X \rightarrow R \cup \{\infty\}$ ,  $i = 0, 1, \dots, N$ . Consider the following *optimization* problem:

$$\begin{aligned} \mathcal{P} : \quad & \text{minimize } g_0(x) \\ & \text{subject to } g_i(x) \leq 0, \quad i = 1, 2, \dots, M, \\ & \quad \quad \quad g_i(x) = 0, \quad i = M + 1, \dots, N, \\ & \quad \quad \quad x \in C. \end{aligned}$$

We prove a “fuzzy multiplier rule” for this problem. As usual, multipliers corresponding to the inequality constraints are nonnegative and multipliers corresponding to the equality constraints have no restriction. To simplify notation we introduce the quantities  $\tau_i, i = 0, 1, \dots, N$ . The  $\tau_i$ 's associated with the inequality constraints and the cost function are always 1, i.e.,  $\tau_i = 1, i = 0, 1, \dots, M$ . This corresponds to nonnegative multipliers. The  $\tau_i$ 's associated with the equality constraints are either 1 or  $-1$ , corresponding to multipliers with arbitrary sign, i.e.,  $\tau_i \in \{-1, 1\}, i = M + 1, \dots, N$ . We use the notation  $\tau_i, i = 0, 1, \dots, N$  throughout this section without further explanation. Our multiplier rule in reflexive spaces is:

**Theorem 2.1.** *Let  $X$  be a reflexive Banach space, let  $C$  be a closed subset of  $X$ , let  $g_i$  be lower semicontinuous for  $i = 0, 1, \dots, M$  and let  $g_i$  be continuous for  $i = M + 1, \dots, N$ . Assume that  $\bar{x}$  is a local solution of  $\mathcal{P}$ . Then, for any positive number  $\varepsilon > 0$  and any weak neighborhood  $V$  of 0 in  $X^*$ , there exist  $(x_i, g_i(x_i)) \in (\bar{x}, g_i(\bar{x})) + \varepsilon B_{X \times R}, i = 0, 1, \dots, N$ , and  $x_{N+1} \in \bar{x} + \varepsilon B_X$  such that*

$$0 \in \sum_{i=0}^N \mu_i D_F(\tau_i g_i)(x_i) + N_F(C, x_{N+1}) + V,$$

where  $\mu_i \geq 0, i = 0, 1, \dots, N$ , and  $\sum_{i=0}^N \mu_i = 1$ .

Before proving this result we compare it to necessary conditions involving limiting subderivatives. Recall the following definition:

**Definition 2.2.** [18, 21] Let  $X$  be a reflexive Banach space and let  $f : X \rightarrow R \cup \{\infty\}$  be a lower semicontinuous function. Define

$$\partial f(x) := \{w\text{-}\lim_{n \rightarrow \infty} v_n : v_n \in D_F f(x_n), (x_n, f(x_n)) \rightarrow (x, f(x))\}$$

and

$$\partial^\infty f(x) := \{w\text{-}\lim_{n \rightarrow \infty} t_n v_n : v_n \in D_F f(x_n), t_n \rightarrow 0^+, (x_n, f(x_n)) \rightarrow (x, f(x))\}$$

and call  $\partial f(x)$  and  $\partial^\infty f(x)$  the subderivative and singular subderivative of  $f$  at  $x$  respectively. Secondly, let  $C$  be a closed subset of  $X$ . Define

$$N(C, x) := \{w\text{-}\lim_{n \rightarrow \infty} v_n : v_n \in N_F(C, x_n), C \ni x_n \rightarrow x\}$$

and call  $N(C, x)$  the normal cone of  $C$  at  $x$ .

Suppose all but one of the  $g_i$ 's are (locally) Lipschitz functions and  $C = X$ , which is entirely general because we can use  $d(x, C) = 0$  to handle the abstract constraint. Application of Theorem 2.1 to problem  $\mathcal{P}$  and taking limits produces the following result which recovers, in a reflexive setting, the limiting necessary conditions of Kruger and Mordukhovich in [15].

**Corollary 2.3.** *Let  $X$  be a reflexive Banach space, let  $C = X$  and let all but one of the  $g_i, i = 0, 1, \dots, N$ , be locally Lipschitz functions. Suppose that  $\bar{x}$  is a local*

solution of problem  $\mathcal{P}$ . Then there exist  $\mu_j \geq 0$ ,  $j = 0, \dots, N$ , satisfying  $\sum_{j=0}^N \mu_j = 1$  and such that

$$0 \in \sum_{i=0}^N \mu_i \partial(\tau_i g_i)(\bar{x}).$$

*Remark 2.4.* (a) Necessary conditions for constrained optimization problems in “fuzzy” form were discussed by Kruger and Mordukhovich in [15]. Their result ([15, Theorem 2.1]) is in terms of  $\varepsilon$ -Fréchet normals to the epigraphs (for the inequality constraints) or graphs (for the equality constraints) of  $g_i$ 's and  $C$  in Banach spaces with Fréchet differentiable renorms. It is not clear how to derive a necessary condition in terms of the subderivatives of the constraints from the results in [15]. The necessary condition in Theorem 2.1 improves that in [15] in two significant directions: it expresses the necessary condition directly through subderivatives, and it uses exact Fréchet normals and Fréchet subderivatives. We note that when the  $g_i$ 's are Lipschitz functions the fuzzy necessary condition in [15] suffices to prove Corollary 2.3 (see [15, Theorem 3] and [14, Theorems 3 and 4]).

(b) Since we require the support function in the definition of the Fréchet subderivative to be concave, our concave Fréchet subderivative is potentially smaller than the usual Fréchet subderivative without the concavity requirement. Consequently the limiting subderivative and normal cone defined in Definition 2.2 are potentially smaller than the corresponding objects in [18, 21]. It is obvious that Theorem 2.1 remains valid if we use the usual Fréchet subderivative instead, and our current statement is formally stronger. Surprisingly, according to [1, Proposition 1] (which is applicable to the concave Fréchet subderivative with some notational changes) and [21, Theorem 9.5] the limiting objects in Definition 2.2 and the corresponding ones in [18, 21] coincide. There are technical advantages to using the concave Fréchet subderivative that we will discuss later.

Recall that, for a locally Lipschitz function  $f$ , the *Clarke subderivative*  $\partial_c f(x)$  of  $f$  at  $x$  is the convex closure of  $\partial f(x)$  and the *Clarke normal cone*  $N_c(C, x)$  of a closed set  $C$  at  $x \in C$  is the convex closure of  $N(C, x)$  [6, 7]. Note that  $\partial_c(-f) = -\partial_c(f)$ . We then have the following corollary, which recovers Clarke's classical necessary condition (in a reflexive Banach space).

**Corollary 2.5.** *Let  $X$  be a reflexive Banach space, let  $C$  be a closed subset of  $X$  and let  $g_i : X \rightarrow \mathbb{R}$ ,  $i = 0, 1, \dots, N$ , be locally Lipschitz functions. Suppose that  $\bar{x}$  is a local solution of problem  $\mathcal{P}$ . Then there exist  $\mu_i \geq 0$  satisfying  $\sum_{i=0}^N \mu_i = 1$  and such that*

$$0 \in \sum_{i=0}^N \mu_i \tau_i \partial_c(g_i)(\bar{x}) + N_c(C, \bar{x}).$$

When  $X$  is a finite dimensional space we can recover Mordukhovich's more precise result [19, Theorem 1(b)] (see also [20, Section 7]) after a limiting process.

**Corollary 2.6.** *Let  $X$  be a finite dimensional Banach space, let  $C$  be a closed subset of  $X$  and let  $g_i$  be lower semicontinuous for  $i = 0, 1, \dots, M$  and continuous for  $i = M + 1, \dots, N$ . Suppose that  $\bar{x}$  is a local solution of problem  $\mathcal{P}$ . Then either:*

(A1) there exist  $v_i^\infty \in \partial^\infty(\tau_i g_i)(\bar{x})$ ,  $i = 0, 1, \dots, N$ , and  $v_{N+1}^\infty \in N(C, \bar{x})$  such that

$$0 = \sum_{i=0}^{N+1} v_i^\infty$$

and  $\sum_{i=0}^{N+1} \|v_i^\infty\| = 1$ , or

(A2) there exist  $\mu_j \geq 0$ ,  $j = 0, \dots, N$ , satisfying  $\sum_{j=0}^N \mu_j = 1$  and such that

$$0 \in \sum_{j \in \{i: \mu_i > 0\}} \mu_j \partial(\tau_j g_j)(\bar{x}) + \sum_{j \in \{i: \mu_i = 0\}} \partial^\infty(\tau_j g_j)(\bar{x}) + N(C, \bar{x}).$$

*Proof.* Suppose  $\bar{x}$  is a solution to  $\mathcal{P}$ . Note that in a finite dimensional space strong and weak convergence coincide. By Theorem 2.1 we can take, for  $i = 0, 1, \dots, N+1$ ,  $x_i^n \in X$ ,  $v_i^n \in D_F(\tau_i^n g_i)(x_i^n)$ ,  $i = 0, 1, \dots, N$ ,  $v_{N+1}^n \in N_F(C, x_{N+1}^n)$  and  $\mu_i^n \geq 0$ ,  $i = 0, 1, \dots, N$ , such that  $\sum_{i=0}^N \mu_i^n = 1$ ,  $(x_i^n, g_i(x_i^n)) \rightarrow (\bar{x}, g_i(\bar{x}))$  for  $i = 1, \dots, N$ ,  $C \ni x_{N+1}^n \rightarrow \bar{x}$  and

$$(1) \quad 0 = \lim_{n \rightarrow \infty} \left( \sum_{i=0}^N \mu_i^n v_i^n + v_{N+1}^n \right).$$

Note that, for  $i = M+1, \dots, N$ ,  $\tau_i^n$  may equal either  $-1$  or  $1$ . However, passing to a subsequence if necessary, we can make  $\tau_i = \tau_i^n$  independent of  $n$ . We consider two cases:

*Case 1.* Assume the sequence  $\sum_{i=0}^N \|\mu_i^n v_i^n\| + \|v_{N+1}^n\|$  is bounded. By passing to subsequences when necessary, we may assume that  $\mu_i^n v_i^n$ ,  $i = 0, 1, \dots, N$ , and  $v_{N+1}^n$  converge to  $u_i$ ,  $i = 0, 1, \dots, N+1$ , and  $(\mu_0^n, \dots, \mu_N^n)$  converges to  $(\mu_0, \dots, \mu_N)$ . By Definition 2.2, for  $\mu_i = 0$ ,  $u_i \in \partial^\infty(\tau_i g_i)(\bar{x})$ , while for  $\mu_i > 0$ ,  $v_i := u_i / \mu_i \in \partial(\tau_i g_i)(\bar{x})$  and  $v_{N+1} \in N(C, \bar{x})$ . Thus,

$$0 \in \sum_{j \in \{i: \mu_i > 0\}} \mu_j \partial(\tau_j g_j)(\bar{x}) + \sum_{j \in \{i: \mu_i = 0\}} \partial^\infty(\tau_j g_j)(\bar{x}) + N(C, \bar{x})$$

where  $\mu_j \geq 0$  and  $\sum_{j=0}^N \mu_j = 1$ .

*Case 2.* Assume the sequence  $\sum_{i=0}^N \|\mu_i^n v_i^n\| + \|v_{N+1}^n\|$  is not bounded. We may assume that  $t_n := 1 / (\sum_{i=0}^N \|\mu_i^n v_i^n\| + \|v_{N+1}^n\|) \rightarrow 0^+$ . Note that the sequences  $\{t_n \mu_i^n v_i^n\}$ ,  $i = 0, 1, \dots, N$ , and  $\{t_n v_{N+1}^n\}$  are bounded. Again by passing to subsequences when necessary, we may assume that, for  $i = 0, 1, \dots, N$ ,  $t_n \mu_i^n v_i^n$  converges to  $v_i^\infty$ ,  $t_n v_{N+1}^n$  converges to  $v_{N+1}^\infty$  and  $(\mu_0^n, \dots, \mu_N^n)$  converges to  $(\mu_0, \dots, \mu_N)$ . Note that  $t_n \mu_i^n \rightarrow 0$  for  $i = 0, 1, \dots, N$ . By Definition 2.2  $v_i^\infty \in \partial^\infty(\tau_i g_i)(\bar{x})$ . Moreover,  $v_{N+1}^n \in N_F(C, x_{N+1}^n)$  implies that  $t_n v_{N+1}^n \in N_F(C, x_{N+1}^n)$ . Therefore,  $v_{N+1}^\infty \in N(C, \bar{x})$ . Thus, taking limits in (1) as  $n \rightarrow \infty$  yields

$$0 = \sum_{i=0}^{N+1} v_i^\infty,$$

where  $v_i^\infty \in \partial^\infty(\tau_i g_i)(\bar{x})$ ,  $i = 0, \dots, N$ , and  $v_{N+1}^\infty \in N(C, \bar{x})$ . Obviously,

$$\sum_{i=0}^{N+1} \|v_i^\infty\| = 1.$$

This completes the proof.  $\square$

*Remark 2.7.* (a) In infinite dimensional spaces one can impose “local compactness” conditions on the functions  $g_i$  and  $C$  to ensure that weak convergence of  $\{t_n \mu_i^n v_i^n\}, i = 0, 1, \dots, N$ , and  $t_n v_{N+1}^n$  implies strong convergence, so that the conclusion of Corollary 2.6 still holds. Some sufficient conditions are discussed in [3, 17, 21]. Since we are mainly concerned with the “fuzzy” form of the necessary conditions we will not go into details of these conditions. However, we must point out that we have no explicit example, and it appears quite hard to find an example to show that  $\sum_{i=0}^{N+1} \|v_i^\infty\| = 1$  in (A1) may fail in infinite dimensional spaces without such conditions.

(b) It is clear from Corollary 2.3 and Corollary 2.6 that the effect of various constraint qualification conditions is to directly ensure the necessary condition in Corollary 2.3 or to rule out (A1) in Corollary 2.6. When constraint qualifications rule out

$$(2) \quad 0 \in \sum_{i=1}^N \mu_i \partial(\tau_i g_i)(\bar{x}) + N(C, \bar{x}),$$

where  $\mu_i \geq 0$  and  $\sum_{i=1}^N \mu_i = 1$ , we get a Karush-Kuhn-Tucker type necessary condition in Corollary 2.3. The well-known Slater and Mangasarian-Fromowitz conditions are two examples of such constraint qualifications.

(c) In general, the singular subderivative part in (A2) of Corollary 2.6 cannot be eliminated. This is demonstrated by the following elementary example:

**Example 2.8.** Consider problem  $\mathcal{P}$  with  $X = C = R$ ,  $N = M = 1$ ,  $g_0(x) = x$  and  $g_1(x) = -x^{1/3}$ . Then 0 is the only solution. We can calculate directly that  $\partial g_0(0) = \{1\}$ ,  $\partial^\infty g_0(0) = \{0\}$ ,  $\partial g_1(0) = \emptyset$ ,  $\partial^\infty g_1(0) = (-\infty, 0]$ . It is clear that relation (A1) is impossible at 0 and (A2) can be satisfied at 0 only if  $\mu_0 = 1$  and  $\mu_1 = 0$ . In that case  $0 \in \partial g_0(0) + \partial^\infty g_1(0) = \{1\} + (-\infty, 0]$ .

We now turn to the proof of our main result. First we recall the following simple lemma.

**Lemma 2.9.** *Let  $X$  be a reflexive Banach space having a Gâteaux smooth and locally uniformly rotund norm  $\|\cdot\|$ . Then  $\varphi := \nabla \frac{1}{2} \|\cdot\|^2$  is one-to-one and maps  $X$  onto  $X^*$ .*

*Proof.* Let  $\xi$  be an arbitrary element of  $X^*$ . We define a function

$$x \rightarrow f(x) := -\langle \xi, x \rangle + \frac{1}{2} \|x\|^2.$$

Note that  $f$  is strict convex, weakly lower semicontinuous and coercive ( $\|x\| \rightarrow \infty$  implies that  $f(x) \rightarrow \infty$ ). We conclude that  $f$  attains its minimum at a unique point  $v \in X$ . Obviously,  $\nabla f(v) = 0$ , i.e.,  $\varphi(v) = \xi$ . Thus,  $\varphi := \nabla \frac{1}{2} \|\cdot\|^2$  is one-to-one and maps  $X$  onto  $X^*$ .  $\square$

We will also need the following geometric result. It is used to guarantee that our new subderivatives sit at appropriate points and that these subderivatives approximate the original subderivatives.

**Lemma 2.10.** *Let  $X$  be a Banach space with a norm  $\|\cdot\|$  that is Fréchet smooth and locally uniformly rotund, and let  $\varphi = \nabla \frac{1}{2} \|\cdot\|^2$ . Suppose that  $v$  is a non-zero element of  $X$  and  $\{x_\alpha\}, \alpha > 0$ , is a family of elements of  $X$  such that*

$$(3) \quad \limsup_{\alpha \rightarrow 0} \|x_\alpha\| \leq \|v\|$$

and

$$(4) \quad \limsup_{\alpha \rightarrow 0} \langle \varphi(v), v - x_\alpha \rangle \leq 0.$$

Then  $\lim_{\alpha \rightarrow 0} \|x_\alpha - v\| = 0$ .

*Proof.* We first note that  $\varphi$  is positively homogeneous and, for any  $v \in X$ ,  $\langle \varphi(v), v \rangle = \|\varphi(v)\| \|v\|$ . It is easy to see from (4) that

$$(5) \quad \liminf_{\alpha \rightarrow 0} \|x_\alpha\| \geq \|v\|.$$

Therefore,

$$(6) \quad \lim_{\alpha \rightarrow 0} \|x_\alpha\| = \|v\|.$$

Multiplying (3) by  $\|\varphi(v)\|$  yields

$$\langle \varphi(v), v \rangle \geq \limsup_{\alpha \rightarrow 0} \|\varphi(v)\| \|x_\alpha\| \geq \limsup_{\alpha \rightarrow 0} \langle \varphi(v), x_\alpha \rangle$$

or

$$(7) \quad \liminf_{\alpha \rightarrow 0} \langle \varphi(v), v - x_\alpha \rangle \geq 0.$$

Combining (4) and (7), we obtain

$$(8) \quad \lim_{\alpha \rightarrow 0} \langle \varphi(v), v - x_\alpha \rangle = 0.$$

By virtue of (6) we need only show that  $\lim_{\alpha \rightarrow 0} x_\alpha / \|x_\alpha\| = v / \|v\|$ .

We proceed by contradiction. Suppose  $\lim_{\alpha \rightarrow 0} x_\alpha / \|x_\alpha\| \neq v / \|v\|$ ; then, without loss of generality, we may assume that there exists an  $\eta > 0$  such that

$$\left\| \frac{x_\alpha}{\|x_\alpha\|} - \frac{v}{\|v\|} \right\| > \eta.$$

Since  $\|\cdot\|$  is locally uniformly rotund, there exists an  $\varepsilon > 0$  such that

$$\left\| \frac{x_\alpha}{\|x_\alpha\|} + \frac{v}{\|v\|} \right\| < 2(1 - \varepsilon).$$

Multiplying both sides of the above inequality by  $\|v\| \|\varphi(v)\| = \langle \varphi(v), v \rangle$ , we get

$$\begin{aligned} 2(1 - \varepsilon) \langle \varphi(v), v \rangle &> \|v\| \|\varphi(v)\| \left\| \frac{x_\alpha}{\|x_\alpha\|} + \frac{v}{\|v\|} \right\| \\ &\geq \|v\| \left\langle \varphi(v), \frac{x_\alpha}{\|x_\alpha\|} + \frac{v}{\|v\|} \right\rangle \\ &= \frac{\|v\|}{\|x_\alpha\|} \langle \varphi(v), x_\alpha \rangle + \langle \varphi(v), v \rangle. \end{aligned}$$

Taking the limit of the right hand side yields

$$2(1 - \varepsilon) \langle \varphi(v), v \rangle \geq 2 \langle \varphi(v), v \rangle,$$

a contradiction.  $\square$

The key to the proof of Theorem 2.1 is the following pair of results, which are of some independent interest.

**Theorem 2.11.** *Let  $X$  be a reflexive Banach space, let  $f : X \rightarrow R \cup \{\infty\}$  be a lower semicontinuous function, and suppose  $\xi \in N_F(S, \bar{x})$ , where  $S := \{x : f(x) \leq 0\}$ . Then, either*

(C1) *for any  $\varepsilon, \eta > 0$  there exists  $(x, f(x)) \in (\bar{x}, f(\bar{x})) + \eta B_{X \times R}$  such that  $D_F f(x) \cap \varepsilon B_{X^*} \neq \emptyset$*

or

(C2) for any  $\varepsilon > 0$ , there exist  $(x, f(x)) \in (\bar{x}, f(\bar{x})) + \varepsilon B_{X \times R}$ ,  $\zeta \in D_F f(x)$  and  $\lambda > 0$  such that

$$\|\lambda\zeta - \xi\| < \varepsilon.$$

*Proof.* Without loss of generality we renorm so that the norm of  $X$  is Fréchet smooth and locally uniformly rotund. As  $\lambda$  can be arbitrarily small, we need only consider the nontrivial case when  $\xi \neq 0$ . Let  $\varphi := \nabla \frac{1}{2} \|\cdot\|^2$  and  $v = \varphi^{-1}(\xi)$ . Then  $v \neq 0$  and  $\langle \xi, v \rangle = \|\xi\| \|v\|$ .

Let  $r > 0$  be a constant such that, for all  $x \in S \cap (\bar{x} + r\|v\|B_X)$ ,

$$0 \geq \langle \xi, x - \bar{x} \rangle + \omega(x),$$

where  $x \rightarrow \langle \xi, x - \bar{x} \rangle + \omega(x)$  is continuous concave and  $\omega(x) = o(\|x - \bar{x}\|)$  when  $x \rightarrow \bar{x}$ .

Since  $f$  is lower semicontinuous, on taking a smaller  $r$  if necessary, we may assume that  $f$  is bounded from below by  $-m$  for some positive constant  $m$  on  $\bar{x} + r\|v\|B_X$ . Observe that  $f(x)$  is positive on  $K := \{x : \langle \xi, x - \bar{x} \rangle + \omega(x) > 0\} \cap (\bar{x} + r\|v\|B_X)$ . For  $\alpha < \min(r/2, 1/m)$  define

$$h_\alpha(x) := \alpha^{-3} \max\{0, \|x - \bar{x} - \alpha v\| - \alpha\|v\|\}^2.$$

Note that  $h_\alpha$  is a Fréchet smooth convex function of  $x$  that has derivative 0 whenever  $h_\alpha(x) = 0$  (in particular at  $\bar{x}$ ). Consider

$$p_\alpha(z) := f(z) + h_\alpha(z) + \delta_{\{\bar{x} + r\|v\|B_X\}}(z).$$

Suppose that (C1) is not true. Then, in particular,  $0 \notin D_F f(\bar{x})$ . Therefore

$$\inf_X p_\alpha < 0.$$

Let  $e_\alpha := \min(\alpha/2, -\inf_X p_\alpha/2)$ . Note that a locally uniformly rotund norm is Kadec-Klee, and so the reflexive version of the Borwein-Preiss Smooth Variational Principle [2, Theorems 2.6 and 5.2] applies with  $p := 2$ ,  $\lambda := \sqrt{2e_\alpha/\alpha} \leq 1$  and  $\varepsilon := e_\alpha$ . Thus there exist  $y_\alpha, w_\alpha \in X$  with  $\|y_\alpha - w_\alpha\| < 2$  such that

$$p_\alpha(y_\alpha) < \inf_X p_\alpha + e_\alpha < 0$$

and such that

$$z \rightarrow p_\alpha(z) + \frac{\alpha}{2} \|z - w_\alpha\|^2$$

attains a global minimum at  $y_\alpha$ .

Since  $p_\alpha(y_\alpha) < 0$ , it follows that  $y_\alpha \in \bar{x} + r\|v\|B_X$  and

$$h_\alpha(y_\alpha) \leq h_\alpha(y_\alpha) + \frac{\alpha}{2} \|y_\alpha - w_\alpha\|^2 < -f(y_\alpha) \leq m.$$

Thus,

$$(9) \quad \|y_\alpha - \bar{x} - \alpha v\| \leq \alpha (\sqrt{m\alpha} + \|v\|)$$

and  $y_\alpha \rightarrow \bar{x}$  as  $\alpha \rightarrow 0$ . When  $\alpha$  is sufficiently small  $y_\alpha$  is in the interior of the ball  $\bar{x} + r\|v\|B_X$ , and

$$z \rightarrow f(z) + h_\alpha(z) + \frac{\alpha}{2} \|z - w_\alpha\|^2$$



attains a local minimum at  $y_\alpha$ . We will only consider such small  $\alpha$  in the sequel. Moreover,

$$f(y_\alpha) \leq p_\alpha(y_\alpha) \leq \inf_X p_\alpha + e_\alpha \leq f(\bar{x}) + e_\alpha$$

and, since  $e_\alpha \rightarrow 0$  and  $f$  is lower semicontinuous, we observe that  $f(y_\alpha) \rightarrow f(\bar{x})$  as  $\alpha \rightarrow 0$ . Furthermore,  $f(y_\alpha) < 0$  (because  $p_\alpha(y_\alpha) < 0$ ) implies that  $y_\alpha \notin K$ . That is to say,

$$(10) \quad 0 \geq \langle \xi, y_\alpha - \bar{x} \rangle + o(\|y_\alpha - \bar{x}\|).$$

Set  $x_\alpha := v + (\bar{x} - y_\alpha)/\alpha$ . Then we can rewrite relations (9) and (10) as

$$(11) \quad \|x_\alpha\| \leq \sqrt{m\alpha} + \|v\|.$$

and

$$(12) \quad 0 \geq \langle \xi, v - x_\alpha \rangle + o(\alpha\|v - x_\alpha\|)/\alpha.$$

Relations (11), (12) and Lemma 2.10 imply that  $\lim_{\alpha \rightarrow 0} x_\alpha = v$ . Since

$$z \rightarrow f(z) + h_\alpha(z) + \frac{\alpha}{2}\|z - w_\alpha\|^2$$

attains a local minimum at  $y_\alpha$  and both  $h_\alpha$  and  $\frac{\alpha}{2}\|\cdot - w_\alpha\|^2$  are convex and Fréchet smooth, the vector

$$\zeta_\alpha := k(\alpha)\varphi(\bar{x} + \alpha v - y_\alpha) + \alpha\varphi(w_\alpha - y_\alpha)$$

is a Fréchet subderivative of  $f$  at  $y_\alpha$ , where

$$k(\alpha) := \begin{cases} 2\alpha^{-3}(\|y_\alpha - \bar{x} - \alpha v\| - \alpha\|v\|)/\|y_\alpha - \bar{x} - \alpha v\| & \text{if } h_\alpha(y_\alpha) > 0, \\ 0 & \text{if } h_\alpha(y_\alpha) = 0. \end{cases}$$

Observe that  $\varphi(\bar{x} + \alpha v - y_\alpha) = \varphi(\alpha x_\alpha) = \alpha\varphi(x_\alpha)$ . Since  $x_\alpha \rightarrow v$  and  $\varphi$  is continuous, we have  $\varphi(x_\alpha) = \varphi(v) + o(1) = \xi + o(1)$  as  $\alpha \rightarrow 0$ . Therefore, we can write

$$\zeta_\alpha := k(\alpha)\alpha(\xi + o(1)) + \alpha\varphi(w_\alpha - y_\alpha)$$

Since (C1) is not true we must have that  $\liminf_{\alpha \rightarrow 0} \alpha k(\alpha)$  is greater than 0. (Otherwise there would exist a sequence  $\alpha_j \rightarrow 0$  with  $D_F f(y_{\alpha_j}) \ni \zeta_{\alpha_j} \rightarrow 0$ .) Moreover,  $\alpha\varphi(w_\alpha - y_\alpha) = O(\alpha)$ . Thus,  $\zeta_\alpha = \alpha k(\alpha)(\xi + o(1))$  as  $\alpha \rightarrow 0$ . Choose  $\alpha$  small enough so that  $(y_\alpha, f(y_\alpha)) \in (\bar{x}, f(\bar{x})) + \varepsilon B_{X \times R}$  and  $\|\zeta_\alpha/\alpha k(\alpha) - \xi\| < \varepsilon$ . Setting  $x := y_\alpha$ ,  $\lambda := 1/\alpha k(\alpha)$  and  $\zeta := \zeta_\alpha$  completes the proof.  $\square$

The next theorem is a counterpart of Theorem 2.11 for equality constraints. The proof is similar to that of Theorem 2.11. We give only a sketch of the proof to limit the length of the exposition.

**Theorem 2.12.** *Let  $X$  be a reflexive Banach space. Let  $f : X \rightarrow R$  be a continuous function and let  $\xi \in N_F(S, \bar{x})$ , where  $S := \{x : f(x) = 0\}$ . Then, either*

(D1) *for any  $\varepsilon, \eta > 0$  there exists  $(x, f(x)) \in (\bar{x}, f(\bar{x})) + \eta B_{X \times R}$  such that*

$$[D_F f(x) \cup D_F(-f)(x)] \cap \varepsilon B_{X^*} \neq \emptyset$$

or

(D2) *for any  $\varepsilon > 0$ , there exist  $(x, f(x)) \in (\bar{x}, f(\bar{x})) + \varepsilon B_{X \times R}$ ,  $\zeta \in D_F f(x) \cup D_F(-f)(x)$  and  $\lambda > 0$  such that*

$$\|\lambda\zeta - \xi\| < \varepsilon.$$

*Sketch of the Proof.* As in the proof of Theorem 2.11, we consider  $\xi \neq 0$  and set  $v := \varphi^{-1}(\xi)$ . Define  $r$  and  $m$  similarly, with  $m$  being the upper bound for  $|f|$  instead of  $-f$ , and define  $K$  and  $h_\alpha$  in the same way. Observe that  $K$  is a convex set on which  $f \neq 0$ . Since  $f$  is continuous, it has constant sign on  $K$ . Define

$$p_\alpha(z) := \begin{cases} f(z) + h_\alpha(z) + \delta_{\{\bar{x}+r\|v\|_{B_X}\}}(z) & \text{if } f \text{ is positive on } K, \\ -f(z) + h_\alpha(z) + \delta_{\{\bar{x}+r\|v\|_{B_X}\}}(z) & \text{if } f \text{ is negative on } K. \end{cases}$$

The remainder of the proof closely follows the proof of Theorem 2.11. □

*Remark 2.13.* By requiring the support function in the definition of the Fréchet subderivative to be concave we may ensure the convexity (and so connectedness) of  $K$  in the proof of Theorem 2.12. Any other condition implying that  $f$  has constant sign on  $K$  will have the same effect.

We may now prove our main result.

*Proof of Theorem 2.1.* We consider the following two cases: Case 1: one of the  $g_i, i = 0, 1, \dots, M$ , satisfies (C1) in Theorem 2.11 or one of the  $g_i, i = M + 1, \dots, N$ , satisfies (D1) in Theorem 2.12, and Case 2: all  $g_i, i = 0, 1, \dots, M$ , satisfy (C2) of the conclusion of Theorem 2.11 and all  $g_i, i = M + 1, \dots, N$ , satisfy (D2) in Theorem 2.12.

*Case 1.* Assume, say,  $g_j$  satisfies (C1) in Theorem 2.11. Then we may take  $\mu_j := 1$  (hence all the other  $\mu_i$ 's are 0) and the conclusion follows trivially.

*Case 2.* Since  $\bar{x}$  is a local solution of  $\mathcal{P}$ , it is a local minimum of  $g_0 + \delta_{C \cap (\bigcap_{i=1}^N S_i)} = g_0 + \sum_{i=1}^N \delta_{S_i} + \delta_C$ , where  $S_i := \{x : g_i(x) \leq 0\}, i = 1, \dots, M$ , and  $S_i := \{x : g_i(x) = 0\}, i = M + 1, \dots, N$ . Therefore,

$$0 \in D_F(g_0 + \sum_{i=1}^N \delta_{S_i} + \delta_C)(\bar{x}).$$

Let  $\eta \in (0, \varepsilon/2)$ , and let  $U$  be a convex weak neighborhood of 0 in  $X^*$  such that  $\eta B_{X^*} + U \in V$ . By the fuzzy sum rule in Theorem 2.10 of [5] (whose statement and proof remain valid for the concave Fréchet subderivative with a small notational change) there exist  $(x_0, g_0(x_0)) \in (\bar{x}, g_0(\bar{x})) + \eta B_{X \times R}, x_{N+1} \in \bar{x} + \eta B_X$  and  $(y_i, g_i(y_i)) \in (\bar{x}, g_i(\bar{x})) + \eta B_{X \times R}, i = 1, \dots, N$ , such that

$$0 \in D_F g_0(x_0) + \sum_{i=1}^N N_F(S_i, y_i) + N_F(C, x_{N+1}) + U.$$

Let  $u_0 \in D_F g_0(x_0), v_i \in N_F(S_i, y_i), i = 1, \dots, N$ , and  $v_{N+1} \in N_F(C, x_{N+1})$  satisfy  $0 \in u_0 + \sum_{i=1}^{N+1} v_i + U$ . Invoking Theorem 2.11 and Theorem 2.12, for each  $i = 1, \dots, N$  there exist  $(x_i, g_i(x_i)) \in (y_i, g_i(y_i)) + \eta B_{X \times R}, u_i \in D_F(\tau_i g_i)(x_i)$  and  $\lambda_i > 0$  such that

$$\|\lambda_i u_i - v_i\| < \eta/N.$$

Then

$$0 \in u_0 + \sum_{i=1}^N \lambda_i u_i + v_{N+1} + \eta B_{X^*} + U.$$

Dividing the above inclusion by  $1 + \sum_{i=1}^N \lambda_i$  and denoting  $\mu_0 := 1/(1 + \sum_{i=1}^N \lambda_i)$ ,  $\mu_i = \lambda_i/(1 + \sum_{i=1}^N \lambda_i)$ ,  $i = 1, \dots, N$ , we obtain

$$0 \in \sum_{i=0}^N \mu_i u_i + v_{N+1}/(1 + \sum_{i=1}^N \lambda_i) + \eta B_{X^*} + U.$$

Since  $v_{N+1}/(1 + \sum_{i=1}^N \lambda_i) \in N_F(C, x_{N+1})$  and  $\eta B_{X^*} + U \subset V$ , we get

$$0 \in \sum_{i=0}^N \mu_i D_F(\tau_i g_i)(x_i) + N_F(C, x_{N+1}) + V.$$

□

In several important cases one can get sharper results. We give two in the sequel. Remark 2.15 describes in what circumstances one has an  $s$ -Hölder renorm and thus the first result applies.

Recall that a lower semicontinuous function  $f : X \rightarrow R \cup \{\infty\}$  is  $s$ -Hölder subdifferentiable at  $x$ ,  $s \in (0, 1]$ , with  $s$ -Hölder subderivative  $x^* \in X^*$  if  $f$  is finite at  $x$  and there exists a function  $g$  of the form

$$g(y) = \langle x^*, y - x \rangle - \sigma \|y - x\|^{1+s},$$

where  $\sigma$  is a positive constant, such that  $f - g$  attains a local minimum at  $x$ . We denote the set of  $s$ -Hölder subderivatives of  $f$  at  $x$  by  $D_{H(s)}f(x)$ . For a closed subset  $C$  of  $X$  the  $s$ -Hölder normal cone of  $C$  at  $x \in C$  is defined by  $N_{H(s)}(C, x) := D_{H(s)}\delta_C(x)$ . When  $X$  has a power modulus of smoothness  $t^{1+s}$  we can sharpen our results by replacing the Fréchet subderivative  $D_F$  with the  $s$ -Hölder subderivative and the Fréchet normal cone  $N_F$  with the  $s$ -Hölder normal cone respectively. The changes in the proofs are mostly notational. We state the  $s$ -Hölder version of our multiplier rule with a sketch of the proof.

**Theorem 2.14.** *Let  $X$  be a (reflexive) Banach space with a norm  $\|\cdot\|$  that has a power modulus of smoothness  $t^{1+s}$ ,  $s \in (0, 1]$ . Let  $C$  be a closed subset of  $X$  and let  $g_i$  be lower semicontinuous functions for  $i = 0, 1, \dots, M$  and continuous for  $i = M + 1, \dots, N$ . Assume that  $\bar{x}$  is a local solution of  $\mathcal{P}$ . Then, for any positive number  $\varepsilon > 0$  and any weak neighborhood  $V$  of 0 in  $X^*$ , there exist  $(x_i, g_i(x_i)) \in (\bar{x}, g_i(\bar{x})) + \varepsilon B_{X \times R}$ ,  $i = 0, 1, \dots, N$ , and  $x_{N+1} \in \bar{x} + \varepsilon B_X$  such that*

$$0 \in \sum_{i=0}^N \mu_i D_{H(s)}(\tau_i g_i)(x_i) + N_{H(s)}(C, x_{N+1}) + V,$$

where  $\mu_i \geq 0$ ,  $i = 0, 1, \dots, N$ , and  $\sum_{i=0}^N \mu_i = 1$ .

*Sketch of the proof.* First, a Banach space with a norm  $\|\cdot\|$  that has a power modulus of smoothness  $t^{1+s}$ ,  $s \in (0, 1]$ , is super-reflexive. By [9, Section 5.2], there is an equivalent uniformly rotund renorm with the same modulus of smoothness. Using such a norm, we can follow the lines of the proof of Theorem 2.1. Most of the changes are notational and obvious. Two small differences are that:

1. We need to replace  $\varphi := \nabla \frac{1}{2} \|\cdot\|^2$  in Lemma 2.9, Lemma 2.10, Theorem 2.11 and Theorem 2.12 by  $\varphi := \nabla \frac{1}{1+s} \|\cdot\|^{1+s}$ .

2. In the proofs of the  $s$ -Hölder versions of Theorem 2.11 and Theorem 2.12 the function  $h_\alpha$  should be defined as

$$h_\alpha(x) := \alpha^{-2-s} \max\{0, \|x - \bar{x} - \alpha v\| - \alpha\|v\|\}^{1+s}.$$

In relation to 2, we need to modify the expression of the subderivative  $\zeta_\alpha$  in a quite obvious way. A more technical change is that now we need to show that the  $h_\alpha$  defined above is an  $s$ -Hölder smooth function of  $x$  that has derivative 0 whenever  $h_\alpha(x) = 0$ .

Obviously,  $h_\alpha$  is Fréchet smooth everywhere. When  $h_\alpha(x)$  is greater than 0, so is  $h_\alpha$  in a neighborhood of  $x$ . Therefore, for  $y$  sufficiently close to  $x$ ,

$$\nabla h_\alpha(y) = \frac{(\|y - \bar{x} - \alpha v\| - \alpha\|v\|)^s}{\|y - \bar{x} - \alpha v\|^s} \varphi(y - \bar{x} - \alpha v).$$

Since  $\varphi$  is  $s$ -Hölder continuous [2, 11] and  $y \rightarrow \frac{(\|y - \bar{x} - \alpha v\| - \alpha\|v\|)^s}{\|y - \bar{x} - \alpha v\|^s}$  is locally Lipschitz at  $x$ , we obtain that  $y \rightarrow \nabla h_\alpha(y)$  is  $s$ -Hölder continuous in a neighborhood of  $x$ . Therefore,  $h_\alpha$  is  $s$ -Hölder differentiable at  $x$ .

Finally we show that 0 is the  $s$ -Hölder derivative of  $h_\alpha$  at any point  $x$  where  $h_\alpha(x) = 0$ . If  $\|x - \bar{x} - \alpha v\| < \alpha\|v\|$ , then  $h_\alpha$  equals 0 in a neighborhood of  $x$  and the conclusion is obvious. When  $\|x - \bar{x} - \alpha v\| = \alpha\|v\|$  we have, for  $y$  close to  $x$ ,

$$\begin{aligned} 0 \leq h_\alpha(y) - h_\alpha(x) &= h_\alpha(y) \leq \alpha^{-2-s} (\|y - \bar{x} - \alpha v\| - \alpha\|v\|)^{1+s} \\ &= \alpha^{-2-s} (\|y - \bar{x} - \alpha v\| - \|x - \bar{x} - \alpha v\|)^{1+s} \leq \alpha^{-2-s} \|y - x\|^{1+s}. \end{aligned}$$

That is to say, 0 is the  $s$ -Hölder derivative of  $h_\alpha$  at  $x$ . □

*Remark 2.15.* The condition of Theorem 2.14 holds in the following settings:

(a) In  $L_p$  ( $1 < p < \infty$ ) with the original norm and  $s = \min(p - 1, 1)$ . Thus for  $p \geq 2$  the strongest possible result is achieved.

(b) In any Hilbert space with the usual norm and  $s = 1$ . In a Hilbert space,  $D_{H(1)}$  and  $N_{H(1)}$  coincide with Rockafellar’s proximal subderivative and proximal normal cone (see [24, 25]), denoted  $D_p$  and  $N_p$  respectively in the sequel.

(c) In super-reflexive spaces for some  $s \in (0, 1]$ . Super-reflexive spaces, while defined rather technically, contain most of the classical reflexive spaces and coincide with those for which there is a uniformly rotund equivalent norm. Let  $X$  be a super-reflexive Banach space. Then, in fact,  $X$  has an equivalent renorm  $\|\cdot\|$  that has a power modulus of smoothness  $t^{1+s}$  for some  $s \in (0, 1]$  [2, 23].

The Hilbert space versions of the crucial results, Theorem 2.11 and Theorem 2.12, have simpler proofs that also lead to a clearer geometric interpretation of the results. We prove, in detail, the Hilbert space version of Theorem 2.11 below in the hope that it will better communicate the idea behind the general proofs.

**Theorem 2.16.** *Let  $X$  be a Hilbert space. Let  $f : X \rightarrow R \cup \{\infty\}$  be a lower semicontinuous function and let  $\xi \in N_p(S, \bar{x})$ , where  $S := \{x : f(x) \leq 0\}$ . Then, either*

(C1) *for any  $\varepsilon, \eta > 0$  there exists  $(x, f(x)) \in (\bar{x}, f(\bar{x})) + \eta B_{X \times R}$  such that  $D_p f(x) \cap \varepsilon B_X \neq \emptyset$*

or

(C2) *for any  $\varepsilon > 0$ , there exist  $(x, f(x)) \in (\bar{x}, f(\bar{x})) + \varepsilon B_{X \times R}$ ,  $\zeta \in D_p f(x)$  and  $\lambda > 0$  such that*

$$\|\lambda \zeta - \xi\| < \varepsilon.$$

*Proof.* Suppose that (C1) is not true. Then, in particular,  $0 \notin D_p f(\bar{x})$ . Let  $r, \sigma > 0$  be constants such that, for all  $x \in S \cap (\bar{x} + r\|\xi\|B_X)$ ,

$$0 \geq \langle \xi, x - \bar{x} \rangle - \sigma \|x - \bar{x}\|^2.$$

Since  $f$  is lower semicontinuous, taking a smaller  $r$  if necessary we may assume that  $f$  is bounded from below by  $-m$  on  $\bar{x} + r\|\xi\|B_X$  for some positive constant  $m$ . Let  $\eta \in (0, r/4)$  be a constant such that  $\langle \xi, x - \bar{x} \rangle - \sigma \|x - \bar{x}\|^2$  is positive on  $(\bar{x} + 2\eta\xi + 2\eta\|\xi\|B_X) \setminus \{\bar{x}\}$  and so is  $f$ .

For  $\alpha < \min(\eta, 1/m)$  define

$$h_\alpha(x) := \alpha^{-1} \max\{0, \|x - \bar{x} - \eta\xi\| - \eta\|\xi\|\}^2.$$

Note that  $h_\alpha$  is a Lipschitz smooth function of  $x$  that has Lipschitz smooth derivative 0 at  $x$  whenever  $h_\alpha(x) = 0$  (in particular at  $\bar{x}$ ). Consider

$$p_\alpha(z) := f(z) + h_\alpha(z) + \delta_{\{\bar{x} + r\|v\|B_X\}}(z).$$

Since 0 is not a proximal subderivative of  $f$  at  $\bar{x}$ ,

$$\inf_X p_\alpha < 0.$$

Let  $e_\alpha := \min(\alpha/2, -\inf_X p_\alpha/2)$ . By the Borwein-Preiss Smooth Variational Principle with  $p := 2$ ,  $\lambda := \sqrt{2e_\alpha/\alpha} \leq 1$  and  $\varepsilon := e_\alpha$  (see [2] and, for a Hilbert space version, [17]), there exist  $y_\alpha, w_\alpha \in X$  with  $\|y_\alpha - w_\alpha\| < 2$  such that

$$p_\alpha(y_\alpha) < \inf_X p_\alpha + e_\alpha < 0$$

and such that

$$z \rightarrow p_\alpha(z) + \frac{\alpha}{2} \|z - w_\alpha\|^2$$

attains a global minimum at  $y_\alpha$ .

Since  $p_\alpha(y_\alpha) < 0$ , it follows that  $y_\alpha \in \bar{x} + r\|v\|B_X$  and

$$h_\alpha(y_\alpha) \leq h_\alpha(y_\alpha) + \frac{\alpha}{2} \|y_\alpha - w_\alpha\|^2 < -f(y_\alpha) \leq m.$$

Thus,

$$\|y_\alpha - \bar{x} - \eta\xi\| \leq \sqrt{m\alpha} + \eta\|\xi\|.$$

This is to say that  $y_\alpha$  is inside the ball  $\bar{x} + \eta\xi + (\sqrt{m\alpha} + \eta\|\xi\|)B_X$ . On the other hand,  $f(y_\alpha) < 0$  (because  $p_\alpha(y_\alpha) < 0$ ) implies that  $y_\alpha$  is outside the ball  $\bar{x} + 2\eta\xi + 2\eta\|\xi\|B_X$ . Consequently

$$\lim_{\alpha \rightarrow 0} y_\alpha = \bar{x}.$$

Moreover,

$$f(y_\alpha) \leq p_\alpha(y_\alpha) \leq \inf_X p_\alpha + e_\alpha \leq f(\bar{x}) + e_\alpha,$$

and as before  $f$  being lower semicontinuous implies that  $f(y_\alpha) \rightarrow f(\bar{x})$  as  $\alpha \rightarrow 0$ . It is easy to see that when  $\alpha$  is sufficiently small  $y_\alpha$  is in the interior of the ball  $\bar{x} + r\|v\|B_X$  and, therefore,

$$z \rightarrow f(z) + h_\alpha(z) + \frac{\alpha}{2} \|z - w_\alpha\|^2$$

attains a local minimum at  $y_\alpha$ . Consequently, the vector

$$\zeta_\alpha := k(\alpha)(\bar{x} + \eta\xi - y_\alpha) + \alpha(w_\alpha - y_\alpha)$$

is a proximal subderivative of  $f$  at  $y_\alpha$  where

$$k(\alpha) := \begin{cases} 2\alpha^{-1}(\|y_\alpha - \bar{x} - \eta\xi\| - \eta\|\xi\|)/\|y_\alpha - \bar{x} - \eta\xi\| & \text{if } h_\alpha(y_\alpha) > 0, \\ 0 & \text{if } h_\alpha(y_\alpha) = 0. \end{cases}$$

Since (C1) is not true we must have  $\liminf_{\alpha \rightarrow 0} k(\alpha)$  greater than 0. (Otherwise there would exist a sequence  $\alpha_j \rightarrow 0$  making  $D_p f(y_{\alpha_j}) \ni \zeta_{\alpha_j} \rightarrow 0$ .) Thus,  $\zeta_\alpha = \eta k(\alpha)(\xi + o(1))$  as  $\alpha \rightarrow 0$ . Choose  $\alpha$  small enough so that  $(y_\alpha, f(y_\alpha)) \in (\bar{x}, f(\bar{x})) + \varepsilon B_{X \times R}$  and  $\|\zeta_\alpha/\eta k(\alpha) - \xi\| < \varepsilon$ . Set  $x := y_\alpha$ ,  $\lambda := 1/\eta k(\alpha)$  and  $\zeta := \zeta_\alpha$  to complete the proof.  $\square$

### 3. NECESSARY CONDITIONS IN SMOOTH SPACES

In this section we explore necessary conditions for constrained optimization problems in Banach spaces admitting a Gâteaux smooth renorm.

**3.1. Notation and Terminology.** Let  $X$  be a Banach space with closed unit ball  $B_X$  and with continuous real dual  $X^*$ . Recall that a *bornology*  $\beta$  for  $X$  is a family of closed bounded subsets of  $X$  whose union is  $X$ , which is closed under multiplication by positive scalars and is directed upwards; that is, the union of any two members of  $\beta$  is contained in some member of  $\beta$  (cf. [22]). Let  $X_i, i = 1, \dots, I$ , be Banach spaces and, for each  $i$ , let  $\beta_i$  be a bornology for  $X_i$ . Then  $\{\prod_{i=1}^I V_i : V_i \in \beta_i\}$  is a bornology for the product space  $X := \prod_{i=1}^I X_i$ . We call this bornology the *product bornology* for  $\beta_i, i = 1, \dots, I$ , and denote it by  $[\beta_1, \dots, \beta_I]$ .

We will denote by  $X_\beta^*$  the dual space of  $X$  endowed with the topology of uniform convergence on  $\beta$ -sets. Similarly, when  $Y$  is a Banach space, we will denote by  $L_\beta(X, Y)$  the Banach space of bounded linear mappings from  $X$  to  $Y$  endowed with the topology of uniform convergence on  $\beta$ -sets. Examples of important bornologies are those formed by all bounded sets (Fréchet bornology denoted by  $F$ ), weak compact sets (weak Hadamard bornology denoted by  $WH$ ), compact sets (Hadamard bornology denoted by  $H$ ), finite sets (Gâteaux bornology denoted by  $G$ ) and products of these bornologies.

Given a mapping  $F : X \rightarrow Y$ , we say that  $F$  is  $\beta$ -differentiable at  $x$  and has a  $\beta$ -Jacobian  $J^\beta F(x)$  (we write  $\nabla^\beta F(x)$  when  $F$  is a function) if  $F$  is continuous at  $x$  and

$$\|t^{-1}(F(x + tu) - F(x) - t\langle J^\beta F(x), u \rangle)\| \rightarrow 0$$

as  $t \rightarrow 0$  uniformly on  $V$  for every  $V \in \beta$ . We say that a mapping  $F$  is  $\beta$ -smooth at  $x$  if  $J^\beta F : X \rightarrow L_\beta(X, Y)$  is continuous at  $x$ . We will also need the concept of (viscosity) subderivatives, which we now recall (cf. [5, 8]):

**Definition 3.1.** Let  $f$  be a lower semicontinuous function and  $f(x) < +\infty$ . We say  $f$  is  $\beta$ -viscosity subdifferentiable and  $x^*$  is a  $\beta$ -viscosity subderivative of  $f$  at  $x$  if there exists a locally Lipschitz function  $g$  such that  $g$  is  $\beta$ -smooth at  $x$ ,  $\nabla^\beta g(x) = x^*$  and  $f - g$  attains a local minimum at  $x$ . We denote the set of all  $\beta$ -viscosity subderivatives of  $f$  at  $x$  by  $D_\beta f(x)$ .

For a product space  $X = \prod_{i=1}^I X_i$  we will always use the Euclidean product norm, i.e.,  $\|\cdot\|_X = \sqrt{\sum_{i=1}^I \|\cdot\|_{X_i}^2}$ .

**Definition 3.2.** [29] Let  $X_i, i = 1, \dots, I$ , be Banach spaces with  $\beta_i$  smooth norms and let  $f : X := \prod_{i=1}^I X_i \rightarrow \bar{R}$  be a proper lower semicontinuous function with

$f(x) < +\infty$ . We say  $f$  is  $(\beta_1, \dots, \beta_I)$ -viscosity subdifferentiable and  $x^* = (x_1^*, \dots, x_I^*)$  is a  $(\beta_1, \dots, \beta_I)$ -viscosity subderivative of  $f$  at  $x = (x_1, \dots, x_I)$  if there exist locally Lipschitz functions  $g_i : X_i \rightarrow \mathbb{R}$  such that  $g_i$  is  $\beta_i$ -smooth at  $x_i$ ,  $\nabla^{\beta_i} g_i(x_i) = x_i^*$  and  $f - \sum_{i=1}^I g_i$  attains a local minimum at  $x$ . We denote the set of all  $(\beta_1, \dots, \beta_I)$ -viscosity subderivatives of  $f$  at  $x$  by  $D_{(\beta_1, \dots, \beta_I)} f(x)$ .

*Remark 3.3.* (a) Since in the definition of  $(\beta_1, \dots, \beta_I)$ -viscosity subderivative we have separated variables, it follows that

$$D_{(\beta_1, \dots, \beta_I)} f(x) \subset D_{[\beta_1, \dots, \beta_I]} f(x).$$

When  $I = 1$  and  $\beta_1 = \beta$ , Definition 3.2 recovers Definition 3.1.

(b) With a little abuse of the notations we will also use  $\beta_i$  in Definition 3.1 to represent  $H(s)$  or  $p$ .

Let  $C$  be a closed subset of  $X := \prod_{i=1}^I X_i$  and  $x \in C$ . The  $(\beta_1, \dots, \beta_I)$ -normal cone of  $C$  at  $x$  is defined by

$$N_{(\beta_1, \dots, \beta_I)}(C, x) := D_{(\beta_1, \dots, \beta_I)} \delta_C(x).$$

(See [29, Section 4.1] for a more lengthy discussion.)

**3.2. Necessary Conditions.** Let  $X$  be a  $\beta$ -smooth Banach space, let  $C$  be a closed subset of  $X$  and let  $g_i, i = 0, 1, \dots, N$ , be continuous functions. Consider the following *optimization problem*:

$$\begin{aligned} \mathcal{P} : \quad & \text{minimize } g_0(x) \\ & \text{subject to } g_i(x) \leq 0, \quad i = 1, 2, \dots, N, \\ & x \in C. \end{aligned}$$

**Theorem 3.4.** *Let  $X$  be a  $\beta$ -smooth Banach space, let  $C$  be a closed subset of  $X$  and let  $g_i, i = 0, 1, \dots, N$ , be continuous functions. Suppose that  $\bar{x}$  is a local solution to problem  $\mathcal{P}$ . Then, for any  $\varepsilon > 0$  and any weak-star neighborhood  $V$  of 0 in  $X^*$ , there exist  $x, z \in \bar{x} + \varepsilon B_X$ ,  $y \in G(\bar{x}) + \varepsilon B_{\mathbb{R}^{N+1}}$ ,  $\nu \in N_{\beta}(C, z)$  and  $\lambda \in \mathcal{T}(y)$  such that*

$$-(\nu, \lambda) \in N_{(\beta, p)}((x, G(x)), \text{graph}(G)) + V \times \varepsilon B_{\mathbb{R}^{N+1}},$$

where

$$G(x) := (g_0(x) - g_0(\bar{x}), g_1(x), \dots, g_N(x))$$

and

$$\mathcal{T}(y) := \{(\lambda_0, \dots, \lambda_N) : \sum_{i=0}^N \lambda_i = 1, \lambda_i \geq 0, \lambda_i(y_i - \max\{y_0, y_1, \dots, y_N\}) = 0\}.$$

Moreover, when  $C = X$  the conclusion holds for any  $\beta$ -neighborhood  $V$  of 0 in  $X_{\beta}^*$ .

*Proof.* See Section 3.3 below.  $\square$

In a Fréchet-smooth Banach space (or more generally an Asplund space) Mordukhovich [19] and Mordukhovich and Shao [21] define the *coderivative*  $D_m^*$  of a function  $G : X \rightarrow Y$  at  $x$  to be the multifunction from  $Y^*$  to  $X^*$  given by

$$D_m^* G(x)(y^*) := \{x^* \in X^* \mid (x^*, -y^*) \in N((x, G(x)), \text{graph}(G))\}.$$

Observe that  $N_{(F,p)}((x, G(x)), \text{graph}(G)) \subset N_F((x, G(x)), \text{graph}(G))$  and  $\mathcal{T}(y)$  is a subset of the compact set  $\{(\lambda_0, \dots, \lambda_N) : \sum_{i=0}^N \lambda_i = 1, \lambda_i \geq 0\}$  for all  $y$ . When  $C = X$ , taking the limit as  $\varepsilon \rightarrow 0$  in Theorem 3.4 yields:

**Theorem 3.5.** *Let  $X$  be a Fréchet-smooth Banach space, let  $C = X$  and let  $g_i, i = 0, 1, \dots, N$ , be continuous functions. Suppose that  $\bar{x}$  is a local solution of problem  $\mathcal{P}$ . Then, there exists  $\lambda \in \mathcal{T}(G(\bar{x}))$  such that*

$$0 \in D_m^* G(\bar{x})(\lambda).$$

Theorem 3.5 illustrates that the coderivative is a convenient device for discussing constrained minimization problems.

The necessary condition discussed in Theorem 3.4 is essentially in terms of normal cones to the graph of  $G(x) := (g_0(x) - g_0(\bar{x}), g_1(x), \dots, g_N(x))$ . We now discuss how to use it for deducing necessary conditions given more directly in terms of subderivatives of  $g_i$ 's in the presence of the following constraint qualification:

(CQ) For any  $x$  in a neighborhood of  $\bar{x}$  and any  $V \in \beta$ ,

$$\limsup_{t \rightarrow 0^+} \sup_{v \in V} \frac{\|G(x + tv) - G(x)\|^2}{t} = 0.$$

*Remark 3.6.* The condition (CQ) is satisfied in every bornology when  $G : X \rightarrow R^{N+1}$  is a locally  $s$ -Hölder continuous mapping with  $s > 1/2$ ; in particular, when each  $g_i$  is locally  $s$ -Hölder continuous ( $s > 1/2$ ) or when  $G$  is a locally Lipschitz function.

**Theorem 3.7.** *Let  $X$  be a  $\beta$ -smooth Banach space, let  $C$  be a closed subset of  $X$  and let  $g_i, i = 0, 1, \dots, N$ , be continuous functions satisfying (CQ). Suppose that  $\bar{x}$  is a local solution to problem  $\mathcal{P}$ . Then, for any  $\varepsilon > 0$  and any weak-star neighborhood  $V$  of 0 in  $X^*$ , there exist  $x_i \in \bar{x} + \varepsilon B_X, i = 0, 1, \dots, N + 1, y \in G(\bar{x}) + \varepsilon B_{R^{N+1}}$  and  $\lambda \in \mathcal{T}(y)$  such that*

$$(13) \quad 0 \in \sum_{i=0}^N D_\beta(\lambda_i g_i)(x_i) + N_\beta(C, x_{N+1}) + V.$$

Moreover, when  $C = X$  the conclusion holds for any  $\beta$ -neighborhood  $V$  of 0 in  $X_\beta^*$ .

*Proof.* By Theorem 3.4 there exist  $x, z \in \bar{x} + \varepsilon B_X, y \in G(\bar{x}) + \varepsilon B_{R^{N+1}}, \nu \in N_\beta(C, z)$  and  $\mu \in \mathcal{T}(y)$  such that

$$-(\nu, \mu) \in N_{(\beta,p)}((x, G(x)), \text{graph}(G)) + V \times \varepsilon B_{R^{N+1}}.$$

Choose  $(w, -\lambda) \in N_{(\beta,p)}((x, G(x)), \text{graph}(G))$  satisfying  $-w \in \nu + V$  and  $\lambda \in \mu + \varepsilon B_{R^{N+1}}$ . This means that there exist a  $\beta$ -smooth Lipschitz function  $h$  and a positive constant  $\sigma$  such that  $\nabla^\beta h(x) = w$  and, for all  $y$  in a neighborhood of  $x$ ,

$$\begin{aligned} & \delta_{\text{graph}(G)}(z, y) - h(z) + \langle \lambda, y \rangle + \sigma \|y - G(x)\|^2 \\ & \geq \delta_{\text{graph}(G)}(x, G(x)) - h(x) + \langle \lambda, G(x) \rangle. \end{aligned}$$

In particular, setting  $y := G(z)$  yields

$$\begin{aligned} & \delta_{\text{graph}(G)}(z, G(z)) - h(z) + \langle \lambda, G(z) \rangle + \sigma \|G(z) - G(x)\|^2 \\ & \geq \delta_{\text{graph}(G)}(x, G(x)) - h(x) + \langle \lambda, G(x) \rangle \end{aligned}$$

for  $z$  in a neighborhood of  $x$ , i.e.,

$$\langle \lambda, G(z) \rangle \geq h(z) - h(x) + \langle \lambda, G(x) \rangle - \sigma \|G(z) - G(x)\|^2$$



for all  $z$  in a neighborhood of  $x$ . Thus, on using (CQ), we have

$$w = \nabla^\beta h(x) \in D_\beta \langle \lambda, G(\cdot) \rangle (x) = D_\beta \left( \sum_{i=0}^N \lambda_i g_i \right) (x),$$

and

$$(14) \quad 0 \in D_\beta \left( \sum_{i=0}^N \lambda_i g_i \right) (x) + N_\beta(C, z) + V.$$

Then the conclusion follows from the fuzzy sum rule, Theorem 2.10 of [5].  $\square$

*Remark 3.8.* Relation (14) is more precise than (13) in Theorem 3.7.

**3.3. Proof of Theorem 3.4.** Before proving Theorem 3.4, we perform some preparatory work. First we recall the following definition of *uniform lower semicontinuity* [1, 5, 29]:

**Definition 3.9.** Let  $w$  be a lower semicontinuous function defined on  $X \times Y$  and  $E$  a closed subset of  $X$ . We say that  $w$  is **uniformly lower semicontinuous** on  $E$  if

$$\inf_{x \in E} w(x, 0) \leq \liminf_{\varepsilon \rightarrow 0} \{w(x, y) : \|y\| \leq \varepsilon, x \in E\}.$$

We say that  $w$  is **locally uniformly lower semicontinuous at  $\bar{x}$**  if  $w$  is uniformly lower semicontinuous on a closed ball centered at  $\bar{x}$ .

The following are sufficient conditions for locally uniform lower semicontinuity (their proofs are elementary) that we will use in the proof of Theorem 3.4.

**Proposition 3.10.** *Let  $X$  be a Banach space and let  $f_0, \dots, f_N$  be lower semicontinuous functions on  $X$ . Then  $w(x, z) := f_0(x) + \sum_{i=1}^N f_i(x + z_i)$  (where  $z = (z_1, \dots, z_N)$ ) is locally uniformly lower semicontinuous at  $\bar{x}$  provided one of the following conditions is satisfied:*

- (1) *at least one of  $f_0, \dots, f_N$  has locally compact level sets at  $\bar{x}$ ; or*
- (2) *all but one of  $f_0, \dots, f_N$  are locally uniformly continuous at  $\bar{x}$ .*

We also use the following variant of Sum Rule 3.4 in [29].

**Theorem 3.11.** *Let  $Y$  be a Banach space with  $\beta$ -smooth norm. Let  $f_0, \dots, f_N$  be lower semicontinuous functions on  $X := Y \times \mathbb{R}^{N+1}$ . Suppose that  $w(x, v) := f_0(x) + \sum_{i=1}^N f_i(x + v_i)$  (where  $v = (v_1, \dots, v_N)$ ) is locally uniformly lower semicontinuous at  $\bar{x} = (\bar{y}, \bar{z})$  and  $\sum_{i=0}^N f_i(x)$  attains a local minimum at  $\bar{x}$ . Then, for any  $\varepsilon > 0$ , there exist  $x_i \in \bar{x} + \varepsilon B_X$  and  $\xi_i \in D_{(\beta, p)} f_i(x_i)$ ,  $i = 0, 1, \dots, N$ , such that*

$$|f_i(x_i) - f_i(\bar{x})| < \varepsilon, \quad \text{and} \quad \left\| \sum_{i=0}^N \xi_i \right\| < \varepsilon.$$

*Proof.* Choose  $r < \varepsilon/2$  such that  $w(x_0, v) := f_0(x_0) + \sum_{i=1}^N f_i(x_0 + v_i)$  is uniformly lower semicontinuous on  $\bar{x} + rB_X$  and  $w(\bar{x}, 0) = \inf_{x_0 \in \bar{x} + rB_X} w(x_0, 0)$ . Define

$$w_t(x_0, v) := w(x_0, v) + t\|v\|^2.$$

By Lemma 3.2 of [29],

$$(15) \quad \begin{aligned} & \liminf_{t \rightarrow \infty} \{w_t(x_0, v) : x_0 \in \bar{x} + rB_X, \|v\| \leq r\} \\ &= \inf_{x_0 \in \bar{x} + rB_X} w(x_0, 0) = w(\bar{x}, 0) = w_t(\bar{x}, 0). \end{aligned}$$

Set

$$u_t(x_0, x_1, \dots, x_N) := w_t(x_0, x_1 - x_0, \dots, x_N - x_0) = \sum_{i=0}^N f_i(x_i) + t \sum_{i=1}^N \|x_0 - x_i\|^2.$$

Then (15) becomes

$$\begin{aligned} & \lim_{t \rightarrow \infty} \inf \{u_t(x_0, x_1, \dots, x_N) : x_0 \in \bar{x} + rB_X, \|x_0 - x_i\| \leq r, i = 1, \dots, N\} \\ (16) \quad & = \inf_{x_0 \in \bar{x} + rB_X} \sum_{i=0}^N f_i(x_0) = \sum_{i=0}^N f_i(\bar{x}) = u_t(\bar{x}, \dots, \bar{x}). \end{aligned}$$

Choose  $t_n \rightarrow \infty$  such that

$$(17) \quad u_{t_n}(\bar{x}, \dots, \bar{x}) < \inf \{u_{t_n}(x_0, \dots, x_N) : x_0 \in \bar{x} + rB_X, \|x_0 - x_i\| \leq r, i = 1, \dots, N\} + \frac{\varepsilon^2}{2n^2}.$$

By the Borwein-Preiss Smooth Variational Principle [2, 16], for each  $n$  there exist a  $\beta$ -smooth function  $\phi_n$  on  $X \times \dots \times X$  and a Lipschitz smooth function  $\psi_n$  on  $R^{(N+1)^2}$  and  $(x_0^n, \dots, x_N^n) = ((y_0^n, z_0^n), \dots, (y_N^n, z_N^n))$  such that

$$u_{t_n}(x_0, \dots, x_N) + \phi_n(y_0, \dots, y_N) + \psi_n(z_0, \dots, z_N)$$

attains a local minimum at  $(x_0^n, \dots, x_N^n)$ ,

$$\|\nabla^\beta \phi_n(y_0^n, \dots, y_N^n)\| < \varepsilon/n, \quad \|\nabla^{LS} \psi_n(z_0^n, \dots, z_N^n)\| < \varepsilon/n$$

(where  $\nabla^{LS}$  signifies the Lipschitz smooth derivative [2]),  $\|x_i^n - \bar{x}\| < \varepsilon/n, i = 0, 1, \dots, N$ , and

$$(18) \quad u_{t_n}(x_0^n, \dots, x_N^n) < \inf \{u_{t_n}(x_0, \dots, x_N) : x_0 \in \bar{x} + rB_X, \|x_0 - x_i\| \leq r, i = 1, \dots, N\} + \frac{\varepsilon^2}{2n^2}.$$

For each  $n$ , the function

$$\begin{aligned} & x_i \rightarrow u_{t_n}(x_0^n, \dots, x_{i-1}^n, x_i, x_{i+1}^n, \dots, x_N^n) \\ & + \phi_n(y_0^n, \dots, y_{i-1}^n, y_i, y_{i+1}^n, \dots, y_N^n) + \psi_n(z_0^n, \dots, z_{i-1}^n, z_i, z_{i+1}^n, \dots, z_N^n) \end{aligned}$$

attains a local minimum at  $x_i = (y_i^n, z_i^n)$ . Thus,

$$(19) \quad \begin{aligned} \xi_i^n & := (-\nabla_{y_i}^\beta \phi_n(y_0^n, \dots, y_N^n) + t_n \nabla^\beta \|\cdot\|_X^2(y_0^n - y_i^n), \\ & \quad -\nabla_{z_i}^{LS} \psi_n(z_0^n, \dots, z_N^n) + t_n \nabla^{LS} \|\cdot\|_{R^{N+1}}^2(z_0^n - z_i^n)) \\ & \in D_{(\beta,p)} f_i(x_i^n). \end{aligned}$$

Similarly,

$$x_0 \rightarrow u_{t_n}(x_0, x_1^n, \dots, x_N^n) + \phi_n(y_0, y_1^n, \dots, y_N^n) + \psi_n(z_0, z_1^n, \dots, z_N^n)$$

attains a local minimum at  $x_0 = x_0^n = (y_0^n, z_0^n)$ , and so

$$(20) \quad \begin{aligned} \xi_0^n := & (-\nabla_{y_0}^\beta \phi_n(y_0^n, \dots, y_N^n) - t_n \sum_{i=1}^N \nabla^\beta \|\cdot\|_X^2(y_0^n - y_i^n), \\ & -\nabla_{z_0}^{LS} \psi_n(z_0^n, \dots, z_N^n) - t_n \nabla^{LS} \sum_{i=1}^N \|\cdot\|_{R^{N+1}}^2(z_0^n - z_i^n)) \\ & \in D_{(\beta,p)} f_0(x_0^n). \end{aligned}$$

Combining (19) and (20) leads to

$$\sum_{i=0}^N \xi_i^n = \left( -\sum_{i=0}^N \nabla_{y_i}^\beta \phi_n(y_0^n, \dots, y_N^n), -\sum_{i=0}^N \nabla_{z_i}^{LS} \psi_n(z_0^n, \dots, z_N^n) \right) \in \varepsilon B_{X^*}.$$

Combining (17) and (18) yields  $|\sum_{i=0}^N f_i(x_i^n) - \sum_{i=0}^N f_i(\bar{x})| \rightarrow 0$  when  $n \rightarrow \infty$ . Since  $f_i$ ,  $i = 0, 1, \dots, N$ , are lower semicontinuous at  $\bar{x}$ , we have, for  $i = 0, 1, \dots, N$ ,  $\lim_n f_i(x_i^n) = f_i(\bar{x})$ . It remains to set  $x_i := x_i^n$ ,  $\xi_i := \xi_i^n$ ,  $i = 0, 1, \dots, N$ , for a sufficiently large integer  $n$ .  $\square$

We may now complete the proof of Theorem 3.4.

*Proof of Theorem 3.4.* Let  $\bar{x}$  be a local solution to problem  $\mathcal{P}$ . Let  $\varepsilon > 0$  and let  $V$  be a weak-star neighborhood of 0 in  $X^*$ . Fix  $r \in (0, \varepsilon/2)$  and select a finite dimensional subspace  $L$  of  $X$  containing  $\bar{x}$  such that  $L^\perp + 2rB_{X^*} \subset V$ . Then  $\bar{x}$  is a local minimum of

$$\max\{g_0(x) - g_0(\bar{x}), g_1(x), \dots, g_N(x)\} + \delta_C(x) + \delta_L(x)$$

and, consequently,  $(\bar{x}, G(\bar{x}))$  is a local minimum of

$$(x, y) \rightarrow f(y) + \delta_{\text{graph}(G)}(x, y) + \delta_C(x) + \delta_L(x),$$

where  $f(y) := \max\{y_0, \dots, y_N\}$ . Set  $f_0(x, y) := f(y)$ ,  $f_1(x, y) := \delta_{\text{graph}(G)}(x, y)$ ,  $f_2(x, y) := \delta_C(x)$  and  $f_3(x, y) := \delta_L(x)$ . Since  $f_3$  has locally compact level sets,  $(f_0, \dots, f_3)$  is locally uniformly lower semicontinuous according to Proposition 3.10. Invoking Theorem 3.11, we obtain that there exist  $x, z \in \bar{x} + rB_X \subset \bar{x} + \varepsilon B_X$ ,  $y \in G(\bar{x}) + rB_{R^{N+1}} \subset G(\bar{x}) + \varepsilon B_{R^{N+1}}$ ,  $\lambda \in D_p f(y) = \mathcal{T}(y)$  and  $\nu \in N_\beta(C, z)$  such that

$$0 \in (0, \lambda) + N_{(\beta,p)}(\text{graph}(G), (x, G(x))) + (\nu, 0) + L^\perp \times \{0\} + rB_{X^* \times R^{N+1}}.$$

This implies the general form of the conclusion of Theorem 3.4.

When  $C = X$ ,  $(\bar{x}, G(\bar{x}))$  is a local minimum of

$$(x, y) \rightarrow f(y) + \delta_{\text{graph}(G)}(x, y),$$

where  $f(y) := \max\{y_0, \dots, y_N\}$ . Since  $f(y)$  is Lipschitz ( $f(y), \delta_{\text{graph}(G)}(x, y)$ ) is locally uniformly lower semicontinuous according to Proposition 3.10. Similar argument as before yields the desired conclusion with any arbitrary  $\beta$ -neighborhood  $V$  of 0 in  $X^*$ .  $\square$

## REFERENCES

1. J. M. Borwein and A. Ioffe, Proximal analysis in smooth spaces, *Set-valued Analysis* **4** (1996), 1–24. MR **96m**:49028
2. J. M. Borwein and D. Preiss, A smooth variational principle with applications to subdifferentiability and to differentiability of convex functions, *Trans. Amer. Math. Soc.* **303** (1987), 517–527. MR **88k**:49013
3. J. M. Borwein and H. M. Strojwas, Proximal analysis and boundaries of closed sets in Banach spaces, Part 1: Theory, *Canad. J. Math.* **38** (1986), 431–452. Part 2: Applications, *Canad. J. Math.* **39** (1987), 428–472. MR **87h**:90258; MR **88f**:46034
4. J. M. Borwein and Q. J. Zhu, Variational Analysis in Non-reflexive Spaces and Applications to Control Problems with  $L^1$  Perturbations, *Nonlinear Analysis* **28** (1997), 889–915. CMP 97:05
5. J. M. Borwein and Q. J. Zhu, Viscosity solutions and viscosity subderivatives in smooth Banach spaces with applications to metric regularity, *SIAM J. Control and Optimization* **34** (1996), 1568–1591. CMP 96:17
6. F. H. Clarke, *Optimization and Nonsmooth Analysis*, John Wiley & Sons, Inc., New York 1983. MR **85m**:49002
7. F. H. Clarke, *Methods of dynamic and nonsmooth optimization*, CBMS-NSF Regional conference series in applied mathematics, SIAM, Philadelphia, 1989. MR **91j**:49001
8. R. Deville, G. Godefroy and V. Zizler, A smooth variational principle with applications to Hamilton-Jacobi equations in infinite dimensions, *J. Funct. Anal.* **111** (1993), 197–212. MR **94b**:49010
9. R. Deville, G. Godefroy and V. Zizler, *Smoothness and renormings in Banach spaces*, Pitman Monographs and Surveys in Pure and Applied Mathematics, No. 64, John Wiley & Sons, Inc., New York 1993. MR **94d**:46012
10. J. Diestel, *Geometry of Banach spaces-Selected topics*, Lecture Notes in Math., **485**, Springer-Verlag, 1975. MR **57**:1079
11. M. Fabian, J. H. Whitfield and V. Zizler, Norms with locally Lipschitz derivatives, *Israel J. Math.* **44** (1983), 262–276. MR **84i**:46028
12. A. D. Ioffe, Approximate subdifferentials and applications, Part 3, *Mathematica* **36** (1989), 1–38. MR **90g**:49012
13. A. D. Ioffe, Proximal analysis and approximate subdifferentials, *J. London Math. Soc.* **41** (1990), 175–192. MR **91i**:46045
14. A. Y. Kruger, Generalized differentials of nonsmooth functions and necessary conditions for an extremum, *Siberian Math. J.* **26** (1985), 370–379. MR **86j**:49038
15. A. Y. Kruger and B. S. Mordukhovich, Extremal points and the Euler equation in nonsmooth optimization, *Dokl. Akad. Nauk BSSR* **24** (1980), 684–687. MR **82b**:90127
16. Y. Li and S. Shi, A generalization of Ekeland's  $\varepsilon$ -variational principle and of its Borwein-Preiss smooth variant, to appear in *J. Math. Anal. Appl.*
17. P. D. Loewen, *Optimal Control via Nonsmooth Analysis*, CRM Lecture Notes Series, Amer. Math. Soc., Summer School on Control, CRM, Université de Montréal, (1992), Amer. Math. Soc., Providence, 1993. MR **94h**:49003
18. B. S. Mordukhovich, Maximum principle in problems of time optimal control with nonsmooth constraints, *J. Appl. Math. Mech.* **40** (1976), 960–969. MR **58**:7284
19. B. S. Mordukhovich, Metric approximations and necessary optimality conditions for general classes of nonsmooth extremal problems, *Soviet Math. Dokl.* **22** (1980), 526–530. MR **82b**:90104
20. B. S. Mordukhovich, *“Approximation methods in problems of optimization and control”*, Nauka, Moscow, 1988. [English transl., Wiley/Interscience, to appear]. MR **89m**:49001
21. B. S. Mordukhovich and Y. Shao, Nonsmooth sequential analysis in Asplund spaces, *Trans. Amer. Math. Soc.* **348** (1996), 1235–1280. MR **96h**:49036
22. R. R. Phelps, *Convex functions, monotone operators and differentiability*, Lecture Notes in Mathematics, No. 1364, Springer Verlag, N.Y., Berlin, Tokyo, 1988, Second Edition 1993. MR **90g**:46063; MR **94f**:46055
23. G. Pisier, Martingales with values in uniformly convex spaces, *Israel J. Math.* **20** (1975), 326–350. MR **52**:14940
24. R. T. Rockafellar, Proximal subgradients, marginal values and augmented Lagrangians in nonconvex optimization, *Math. Oper. Res.* **6** (1981), 424–436. MR **83m**:90088

25. R. T. Rockafellar, Extensions of subgradient calculus with applications to optimization, *Non-linear Analysis, TMA* **9** (1985), 665-698. MR **87a**:90148
26. J. S. Treiman, A new characterization of Clarke's tangent cone and its applications to subgradient analysis and optimization, Ph. D. Thesis, Univ. of Washington, 1983.
27. J. S. Treiman, Clarke's gradients and epsilon-subgradients in Banach spaces, *Trans. Amer. Math. Soc.* **294** (1986), 65-78. MR **87d**:90188
28. J. S. Treiman, Lagrange multipliers for nonconvex generalized gradients with equality, inequality and set constraints, preprint.
29. Q. J. Zhu, Calculus rules for subderivatives in smooth Banach spaces, preprint.

DEPARTMENT OF MATHEMATICS AND STATISTICS, SIMON FRASER UNIVERSITY, BURNABY, BC V5A 1S6, CANADA

*E-mail address:* `jborwein@cecm.sfu.ca`

DEPARTMENT OF MATHEMATICS AND STATISTICS, WESTERN MICHIGAN UNIVERSITY, KALAMAZOO, MICHIGAN 49008

*E-mail address:* `treiman@math-stat.wmich.edu`

DEPARTMENT OF MATHEMATICS AND STATISTICS, WESTERN MICHIGAN UNIVERSITY, KALAMAZOO, MICHIGAN 49008

*E-mail address:* `zhu@math-stat.wmich.edu`