Journal of Convex Analysis
Volume 3 (1996), No.1, 117–140

A Bornological Approach to Rotundity and Smoothness Applied to Approximation[1]

John Read
Center for Experimental and Constructive Mathematics, Department of Mathematics and Statistics, Simon Fraser University, Burnaby, British Columbia V5A 1S6, Canada.
e-mail: read@cecm.sfu.ca

Received June 1, 1994
Revised manuscript received November 17, 1995

Dedicated to R. T. Rockafellar on his 60th Birthday

In this paper we are interested in examining the geometry of a bounded convex function over a Banach space via its subdifferential mapping. We will consider two concepts. The first is the single valuedness and continuity of the subdifferential mapping, and the second is the single valuedness and the continuity of the "inverse" of this mapping. The smoothness of f is important for the first concept as the convexity of f is for the second. We generalize some of the well known results on upper semi-continuity of the subdifferential mapping, and we introduce a bornological approach to convexity, which allows us to draw very nice parallels for the continuity of the inverse mapping with the coresponding concept for the well understood subdifferential mapping. The theory developed allows us to give a local Smulyan result in which the convexity at a point on the unit sphere is characterized by the uniform smoothness of the subdifferential of this point, and to give the smoothness of the primal norm at a point in terms of the convexity of the dual norm about the subdifferential of that point. As the title implies we will place special emphasis on the approximation of this convex function by a sequence of such functions and derive conditions, which ensure satisfactory approximation of the subdifferential and "inverse" mappings.

Keywords : approximation, bornology, convexity, differentiability, the subdifferential mapping and its inverse, Banach space geometry, local uniform rotundity, local Smulyan results, optimal control of vibrations

1991 Mathematics Subject Classification: 26E15, 46G05, 46A17, 46A20, 46B20, 46B25 46N10, 49J20, 49J27, 49J50, 49J52, 54C08, 58C20

1. Introduction: A Motivation

The aim of this report is to describe -in terms commonly used in Banach space geometry and convex analysis- the problem that is at the heart of approximating the norm (time) minimal distributed control of a vibrating process, which is governed by an abstract wave equation (e.g. strings, beams, membranes, plates and networks of these) in a Hilbert

[1] This paper is the first chapter of the author's dissertation submitted to the Department of Mathmatics of the Technische Hochschule Darmstadt D17. The author would like to thank Professor Hofmann of said institution for his time and effort in reading this manuscript and providing many useful comments.

ISSN 0944-6532 / \$ 2.50

space. This approach leads naturally to many questions in renorming of spaces à la [4].

One can show that the problem of finding the set of all distributed control functions (that is, the problem of determining the set of all forcing functions which bring an initial state of our system to rest within a given time $T > 0$) boils down to solving the linear equation

$$S(\omega) = c, \tag{1.1}$$

where S is a surjective bounded linear operator from the Bochner space $L^2((0,T),\ell^2)$ onto the Hilbert space $\ell^2(\mathbb{R}^2)$. The operator S is determined by the eigenelements λ_j, φ_j for $j \in \mathbb{N}$, of the system being examined. For example, the vibrating string of unit length fixed at both ends has eigenvalues $\lambda_j = (j\pi)^2$ and eigenfunctions $\varphi_j(\xi) = \sin(j\pi\xi)$ for $\xi \in [0,1]$ and all $j \in \mathbb{N}$. The right hand side of (1.1) contains information about the initial state of our problem and is an element of $\ell^2(\mathbb{R}^2)$. The solution set of (1.1) is some affine subspace of $L^2((0,T),\ell^2)$. We wish to find any (the unique) element of this affine subspace, which for some given $p \in [2,\infty]$ minimizes the norm on $L^p((0,T),\ell^2) \subset L^2((0,T),\ell^2)$.

We introduce the Hilbert adjoint operator of S, which we denote by S^*. By definition $\langle S(\omega), x\rangle = \langle \omega, S^*x\rangle$ for all $\omega \in L^2((0,T),\ell^2)$ and all $x := (x_j^1, x_j^2)_{j\in\mathbb{N}} \in \ell^2(\mathbb{R}^2)$, It is easy to see that S^* is given by $S^* = (S_j^*)_{j\in\mathbb{N}}$ where the operator S_j^* acts on $x \in \ell^2(\mathbb{R}^2)$ to give

$$(S_j^*x)(t) := x_j^1 \sin\sqrt{\lambda_j}t + x_j^2 \cos\sqrt{\lambda_j}t \quad \text{for} \quad t \in [0,T] \quad \text{a.e..} \tag{1.2}$$

After some technical details we can use (1.2) to show that any (the unique) p-norm minimal solution of (1.1) has the form (which for $p = \infty$ is called a bang-bang control) given by

$$\omega(t) := \langle c, \tilde{x}\rangle \|(S^*\tilde{x})(t)\|^{q-2}\, (S^*\tilde{x})(t) \quad \text{for} \quad t \in [0,T] \quad \text{a.e.,} \tag{1.3}$$

where q is the dual index for p, that is we have $p^{-1} + q^{-1} = 1$, and where $\tilde{x}$ is any (the unique) solution of the dual problem in $\ell^2(\mathbb{R}^2)$, which is

$$\text{to maximize} \quad \psi_c(x) := \langle c, x\rangle$$

$$\text{subject to} \quad f(x) := \|S^*x\|_q \le 1.$$

Clearly, S^* is intimately connected with the adjoint of the restriction of S to $L^p((0,T),\ell^2)$, and (even for $p = \infty$) the range of this adjoint is contained in $L^q((0,T),\ell^2)$. Much of our analysis is simplified by showing that f is actually an equivalent norm on $\ell^2(\mathbb{R}^2)$. We designate $\ell^2(\mathbb{R}^2)$ with norm f by Y, whose dual space Y' is $\ell^2(\mathbb{R}^2)$ with norm f^*.

Modulo an application of the Lagrange multiplier rule, we realize our wish if we can find any (the unique) $\tilde{x} \in Y$ such that $\tilde{x} \in \mathcal{S}_1 := \{x \in \ell^2(\mathbb{R}^2) : f(x) = 1\}$, the unit sphere in Y, and such that there exists some $\lambda > 0$ with $c/\lambda \in \partial f(\tilde{x})$, where $\partial f(\tilde{x})$ is the subdifferential of f evaluated at $\tilde{x}$. Of course, the subdifferential of f at any $x \in \mathcal{S}_1$ is contained in $\mathcal{S}'_1$, the unit sphere of Y', and is exactly the set of norming functionals for x. In this way the control problem is reduced to understanding the relationship between f and ∂f.

So far so good, but our standpoint is a practical one, and even in the simplest of cases one cannot expect to have exact knowledge of the right hand side of (1.1). More often one realizes the actual initial state of the medium (that is c) in some kind of limiting process. For instance, one might consider the initial state $c_L \in \ell^2(\mathbb{R}^2)$ arrived at by projecting c onto the space spanned by the first L fundamental harmonics for each $L \in \mathbb{N}$. That is, one truncates c as in [9]. For each c_L we find $\lambda_L > 0$ and $\tilde{x}_L \in \mathcal{S}_1$ such that $c_L/\lambda_L \in \partial f(\tilde{x}_L)$. Since the sequence of approximating control functions is defined in terms of the sequence $\{\tilde{x}_L\}$, the question of convergence of $\{\tilde{x}_L\}$ to $\tilde{x}$ arises in a most natural way. In fact, what we must understand is exactly which conditions on f ensure that the mapping

$$c \mapsto \partial f^*(c) : Y' \to Y \tag{1.4}$$

is norm to norm or at least norm to weak continuous at c, thereby ensuring $c_L \to c$ is sufficient (and necessary) for the convergence (in some sense) of the sequence of approximate control functions. Unfortunately, the dual norm f^* is only of theoretical value, in as much as it can only be approximated numerically, so we cannot examine its smoothness properties to answer this question. It turns out that the convexity properties of f at the dual solution $\tilde{x}$ guarantee the continuity of the mapping (1.4). Specifically, since Y is reflexive (even if $p = \infty$), we show that a necessary and sufficient condition for the norm to weak continuity of (1.4) is $\mathcal{G}$-convexity at $\tilde{x}$, which is something closely akin to strict convexity. We go on to show that a necessary and sufficient condition for the norm to norm continuity of (1.4) is $\mathcal{F}$-convexity at $\tilde{x}$. This condition is implied by, but not equivalent to (even in the case of reflexive Banach spaces), local uniform convexity (local uniform rotundity) as defined in [10]. See Section 4 for examples.

The next problem one encounters is the inexact knowledge of the eigenvalues themselves, which is the case if the vibrating system under consideration is approximated numerically (e.g. any finite element approximation of the vibrating system). This leads to not only a sequence of approximating right hand sides of (1.1) but also to a sequence of approximating norms $\{f_L\}$, which incorporate the inexact eigenvalues. This sequence converges uniformly on bounded sets to f. In fact, if we let $\mathcal{K}$ be the set of norms on $\ell^2(\mathbb{R}^2)$ which are equivalent to f, what we are now asking is what conditions on f ensure that the mapping

$$(f^*, c) \mapsto \partial f^*(c) : \mathcal{K} \times Y' \to Y \tag{1.5}$$

is (uniform cross) norm to norm or at least (uniform cross) norm to weak continuous at (f^*, c). That is, we topologize $\mathcal{K}$ with the topology of uniform convergence on bounded sets and give Y' the norm topology and study the continuity of (1.5) with the norm or weak topology on Y. It turns out that this setting lends itself well to the concepts of $\mathcal{G}-$ and $\mathcal{F}-$convexity and that the presence of $\mathcal{K}$ in (1.5) changes little in the analysis. Thus, we may exploit the smoothness and convexity of f to once again get that $c_L \to c$ is the (necessary and) sufficient condition for the norm to weak ($\mathcal{G}-$convexity) or norm to norm ($\mathcal{F}-$convexity) continuity of the mapping (1.5) at (f^*, c), thereby ensuring that $c_L \to c$ is sufficient (and necessary) for the convergence (in some sense) of the sequence of approximate control functions. This theory makes no demands on the smoothness of the sequence of approximating norms, but asks only that they approximate the limit norm uniformly well on some neighborhood of the point we are interested in.

In summary we apply local analysis to examine the duality between convexity in the primal space and differentiability in the dual space, as well as the duality between differentiability in the primal space and convexity in the dual space. In [5] a global theory is developed. We show that the definition of uniform convexity (rotundity) used in the global theory does not lend itself well to a satisfactory local theory. Specifically, for the case $p = \infty$ we do not have $\mathcal{G}-$convexity at each point, and so one cannot hope to apply global theory (i.e. show f is uniformly convex as in [5]) to examine the continuity of the mapping (1.4) or for that matter (1.5). By relaxing the global setting to a local one, we can apply generic types of results to show that for any c in some dense G_δ subset of Y' we in fact get the much stronger convergence.

Thus, in [13] we are able to apply the theoretical results derived here to extend two earlier works. In [8] approximation methods are employed to approximate (1.3) for $p = 2$. The geometry of the dual problem is not discussed and no convergence rate is derived. We place these arguments in a geometrical setting which allows the approximation problem to be understood properly. Exploiting the geometry involved we are able to not only give a more general sufficient condition for convergence ($c_L \to c$) than that given in [8], but we are also able to show that this condition is also necessary not only for $p = 2$ but for all $p \in [2, \infty]$. For $p \in [2, \infty)$ we show that f is uniformly convex as in [5], allowing stronger convergence results than we get for $p = \infty$. The case $p = \infty$ is treated in [9], where calculations are based on truncation and knowledge of the eigenelements. Thus, the finite dimensional subspace used in [9] is $\operatorname{span}\{\varphi_j\}_{j=1}^{M}$, which is in practice not viable. The powerful geometrical arguments iderived here allow us to weaken significantly the assumptions made (exact knowledge of the eigenelements), while at the same time strengthening the convergence obtained (norm or semi-norm convergence instead of weak*).

2. Smoothness and the Subdifferential Mapping

We consider a real Banach space Y, on which is defined a convex function f. We will further assume that f is bounded. That is, for each $\rho > 0$ there is a $\beta_\rho > 0$ such that if $\|x\| \leq \rho$, then $|f(x)| \leq \beta_\rho$. In other words the function f is bounded on any bounded set $\mathcal{V} \subset Y$. This assumption presents no real loss of generality since given any point $x \in Y$, at which a general convex function is continuous, we can construct a bounded convex function which agrees with the general function on some neighborhood of x. This assumption simplifies many of the arguments, so we will retain it throughout.

We will designate the dual of Y by Y'. The subdifferential of f at any $x \in Y$ denoted by $\partial f(x) \subset Y'$ and defined by

$$\partial f(x) := \{\xi' \in Y' : \langle \xi' | y \rangle \leq f(x+y) - f(x) \text{ for all } y \in Y\}, \tag{2.1}$$

is nonempty, bounded, convex and weak-* compact [12]. Another result to be gleaned from [12] is that for each $x, y \in Y$ we get a nice formula for the directional derivative of f at x in direction y, which we denote by

$$df(x; y) := \lim_{\alpha \to 0^+} \frac{f(x + \alpha y) - f(x)}{\alpha} = \max_{\xi' \in \partial f(x)} \langle \xi' | y \rangle.$$

For each $x, y \in Y$ we define the function $e(\,\cdot\,; x, y) : \mathbb{R} \to \mathbb{R}$ by

$$e(\alpha;x,y)=\begin{cases}\dfrac{f(x+\alpha y)-f(x)}{\alpha}-df(x;y) & \text{if}\quad \alpha\neq 0,\\ 0 & \text{if}\quad \alpha=0,\end{cases}\tag{2.2}$$

which is continuous on $\mathbb{R}$, except possibly from the left at the origin. For a fixed $x \in Y$ $\partial f(x)$ having only one element, which we write $\partial f(x) = \{\nabla f(x)\}$, is equivalent to each function $e(\,\cdot\,;x,y)$ in the family of functions indexed by $y \in Y$ being continuous at the origin. But this is the definition of Gateaux differentiability of f at $x \in Y$ as in [11].

From the linearity of $\langle \xi' | \cdot \rangle$ in (2.1) it is clear that $\alpha e(\alpha;x,y) \geq 0$ for all $\alpha \in \mathbb{R}$, and from the convexity of f it follows that $e(\,\cdot\,;x,y)$ is monotonically increasing on $\mathbb{R}$. We will in fact always choose y on the unit sphere. This choice will not restrict the generality of our arguments, since one can show $e(\alpha;x,\beta y) = \beta e(\alpha\beta;x,y)$ for all $x,y \in Y$ and all $\beta > 0$.

In our notation $\mathcal{V}+\mathcal{W} := \{x \in Y : x = v + w$ for some $v \in \mathcal{V}$ and $w \in \mathcal{W}\}$ for any $\mathcal{V},\mathcal{W} \subseteq Y$. Further, we let $\mathcal{B}_\rho(\bar{x})$ $[\mathcal{B}_\rho]$ be the open ball of radius $\rho > 0$ about $\bar{x}$ [the origin], $\bar{\mathcal{B}}_\rho(\bar{x})$ $[\bar{\mathcal{B}}_\rho]$ be the closed ball of radius $\rho \geq 0$ about $\bar{x}$ [the origin] and $\mathcal{S}_\rho(\bar{x})$ $[\mathcal{S}_\rho]$ be the sphere of radius $\rho > 0$ about $\bar{x}$ [the origin]. We will define the multiplication of a set by a scalar in the classical way $\alpha\mathcal{V} := \{x \in Y : x = \alpha y$ for some $y \in \mathcal{V}\}$.

We let $C_0(\mathbb{R})$ be the set of all continuous, strictly monotonically increasing functions on $\mathbb{R}$ which vanish at the origin.
Recall that if ϕ is a real valued, convex function on some open interval $I \subseteq \mathbb{R}$, then we can define the right- and left-hand derivatives $D^\pm\phi$ of ϕ on I, which are monotonically increasing functions, and satisfy $D^-\phi(\tau) := \min_{\partial\phi(\tau)} \leq \max_{\partial\phi(\tau)} =: D^+\phi(\tau)$ for all $\tau \in I$.

2.1. *Differentiability Concepts*

In order to give a concise yet general concept of the differentiability of a bounded, convex function f, we consider again e given by (2.2) and two nonempty sets $\mathcal{V} \subseteq Y$ and $\mathcal{W} \subseteq \mathcal{S}_1$. We will take a bornological point of view here and in the next section.

Definition 2.1. We say e admits a uniform estimate [to the right] on $\mathcal{V}$ with respect to $\mathcal{W}$ if there exists $E_\mathcal{W} \in C_0(\mathbb{R})$ such that $|e(\alpha;x,y)| \leq |E_\mathcal{W}(\alpha)|$ for all $y \in \mathcal{W}$, all $x \in \mathcal{V}$ and all $\alpha \in \mathbb{R}$ $[\alpha \geq 0]$. We also say $E_\mathcal{W}$ estimates e uniformly well [to the right] on $\mathcal{V}$ with respect to $\mathcal{W}$. If $\mathcal{V} = \{x_0\}$ we say e admits an estimate [to the right] at x_0 with respect to $\mathcal{W}$.

There are three classes of subsets of $\mathcal{S}_1$ which interest us. They are

- $\mathcal{G}$, the class of all finite subsets of $\mathcal{S}_1$,
- $\mathcal{H}$, the class of all compact subsets of $\mathcal{S}_1$, and
- $\mathcal{F}$, the class of all bounded subsets of $\mathcal{S}_1$.

Definition 2.2. If e admits a uniform estimate [to the right] on $\mathcal{V}$ with respect to each

$$\mathcal{W}\in\begin{cases}\mathcal{G}, \text{ then we say that } f \text{ is } \textit{uniformly } \mathcal{G}-[\mathcal{G}^+-]\textit{differentiable} \text{ on } \mathcal{V}.\\ \mathcal{H}, \text{ then we say that } f \text{ is } \textit{uniformly } \mathcal{H}-[\mathcal{H}^+-]\textit{differentiable} \text{ on } \mathcal{V}.\\ \mathcal{F}, \text{ then we say that } f \text{ is } \textit{uniformly } \mathcal{F}-[\mathcal{F}^+-]\textit{differentiable} \text{ on } \mathcal{V}.\end{cases}$$

If in this definition $\mathcal{V} = \{x\}$, we say that f is differentiable at x in the sense under consideration. In as much as we are considering functons which are bounded and convex, it is easy to see that in our terminology (uniform) $\mathcal{G}-$differentiability is exactly (uniform) Gateaux differentiability, and (uniform) $\mathcal{F}-$differentiability is exactly (uniform) Fréchet differentiability. We also remind the reader that due to the Lipschtz continuity of the functions under consideration $\mathcal{H} - [\mathcal{H}^+-]$differentiability is subsumed in the concept of $\mathcal{G} - [\mathcal{G}^+-]$differentiability. We will need this result later. See [1].

Since $\mathcal{S}_1 \in \mathcal{F}$, we have that if f is uniformly $\mathcal{F}-$differentiable on $\mathcal{V} \subseteq Y$, there exists some $E \in C_0(\mathbb{R})$, such that for all $x \in \mathcal{V}$, all $\alpha \in \mathbb{R}$ and all $y \in \mathcal{S}_1$, we get

$$f(x + \alpha y) \leq f(x) + \alpha\langle\nabla f(x)|y\rangle + \alpha E(\alpha).$$

Lemma 2.3. *Let f be a bounded, convex function over the Banach space Y. Then the set $\cup_{x\in\mathcal{V}}\partial f(x)$ is norm bounded if $\mathcal{V}$ is.*

We remind the reader of the following variation of the Mean Value Theorem, the proof of which can be found in [1].

Lemma 2.4. *If f is bounded and convex on the Banach space Y, then for $x, y \in Y$ there is some $\tau_0 \in [0, 1]$ and some $\xi' \in \partial f(x + \tau_0 y)$ such that $f(x + y) - f(x) = \langle\xi'|y\rangle$.*

2.2. *Differentiability and Continuity of the Subdifferential Mapping*

The set of all bounded convex functions forms a convex cone $\mathcal{K}$ in the vector space $\mathbb{R}^Y$. By fixing $\mathcal{V} \subset Y$ and $\rho > 0$, we can endow $\mathcal{K} \times Y$ with the semi-metric σ generated by the semi-norm on $\mathcal{K}$ given by $\|f\|_{\mathcal{V},\rho} := \sup_{x\in\mathcal{V}+B_\rho} |f(x)|$ and the norm on Y. Thus,

$$\sigma\left((f, x), (f_0, x_0)\right) := \max\{\|f - f_0\|_{\mathcal{V},\rho}\ ,\ \|x - x_0\|\}. \tag{2.4}$$

For each $\delta > 0$ we say $(f, x) \in \mathcal{B}_\delta((f_0, x_0))$ if $\sigma((f, x), (f_0, x_0)) < \delta$. Using this topology on $\mathcal{K} \times Y$, we want to explore the continuity of the mapping

$$(f, x) \mapsto \partial f(x) : \mathcal{K} \times Y \to \mathcal{C}, \tag{2.5}$$

where $\mathcal{C} \subset 2^{Y'}$ is the set of all non-empty, convex, weak-* compact subsets of Y'. In a subsequent paper we will consider various hyper-topologies on $\mathcal{C}$, each of which has very interesting implications for set convergence. This would take us too far afield, so in this paper we will define what continuity means directly in terms of elements of the subdifferentials. In the subsequent paper these definitions will then become conditions equivalent to the type of continuity under consideration there. Now we define the two types of [upper semi-]continuity that we are interested in. The reader should keep the example of the norm in mind. It is well known that for the norm the $\mathcal{G}-$ [uniform $\mathcal{F}-$]differentiability of the norm at a point of the unit sphere [on $\mathcal{V} \subseteq \mathcal{S}_1$] is equivalent to the norm to weak* [norm to norm uniform] continuity of any selection of the subdifferential mapping [5]. That is, if $\{x_L\} \subset \mathcal{S}_1$ and $x_L \to x_0$, then we get $\xi'_L \xrightarrow{\star} \xi'_0$ $\quad[\xi'_L \to \xi'_0$ uniformly for $x_0 \in \mathcal{V}]$, where ξ'_L is any selection from $\partial f(x_L)$ and ξ'_0 is any selection (the gradient) from $\partial f(x_0)$. We generalize these concepts in the following definitions.

Definition 2.5. Mapping (2.5) is uniformly $\mathcal{G}-[\mathcal{F}-]$upper semi-continuous on $\{f_0\}\times\mathcal{V}$ if for $\varepsilon > 0$ and $\mathcal{W} \in \mathcal{G}$ $[\mathcal{F}]$ there exists a $\delta > 0$ (independent of $x_0 \in \mathcal{V}$) such that if $x_0 \in \mathcal{V}$, if $(f, x) \in \mathcal{B}_\delta((f_0, x_0))$ and if $\xi' \in \partial f(x)$, then $\sup_{y\in\mathcal{W}}(\ \langle\xi'|y\rangle - df_0(x_0; y)\ ) < \varepsilon$.

Definition 2.6. Mapping (2.5) is uniformly $\mathcal{G}-[\mathcal{F}-]$continuous on $\{f_0\}\times\mathcal{V}$ if for each $\varepsilon > 0$ and $\mathcal{W} \in \mathcal{G}$ $[\mathcal{F}]$ there exists a $\delta > 0$ (independent of $x_0 \in \mathcal{V}$) such that if $x_0 \in \mathcal{V}$, if $(f, x) \in \mathcal{B}_\delta((f_0, x_0))$, if $\xi' \in \partial f(x)$ and if $\xi'_0 \in \partial f_0(x_0)$, then $\sup_{y\in\mathcal{W}} |\langle\xi' - \xi'_0|y\rangle| < \varepsilon$.

If $\mathcal{V} = \{x_0\}$, then we say the mapping is $\mathcal{G}-$ or $\mathcal{F}-$[upper semi-]continuous at x_0.

We remind the reader that $\mathcal{S}_1 \in \mathcal{F}$, so that one works with exactly this set in that setting. We also note that $\mathcal{G}-$ and $\mathcal{F}-$ continuity of (2.5) at x_0 imply that $\partial f_0(x_0)$ is a singleton. This should come as no suprise, as we can characterize the continuity of the mapping $(f, x) \mapsto \partial f(x) : \mathcal{K} \times Y \to \mathcal{C}$ in terms of the smoothness of f. This characterization is our next goal. One could in fact obtain a uniform version of the next result by using uniform $\mathcal{G} - [\mathcal{G}^+-]$differentiability and uniform $\mathcal{G}-$[upper semi-]continuity. Such a result is, however, of little practical interest for us, as this entails being able to estimate e uniformly in $x \in \mathcal{V}$ for each fixed $\mathcal{W} \in \mathcal{G}$. Instead, we go on to give a very useful uniform version of $\mathcal{F}-$[upper semi-]continuity. The next proof is simpler and the results are somewhat stronger than the more traditional formulation of upper semi-continuity in [11].

Theorem 2.7. *The mapping $(f, x) \mapsto \partial f(x) : \mathcal{K}\times Y \to \mathcal{C}$ is $\mathcal{G}-$[upper semi-]continuous at $(f_0, x_0) \in \mathcal{K} \times Y$ if and only if f_0 is $\mathcal{G} - [\mathcal{G}^+-]$differentiable at x_0.*

Proof. Assume first the $\mathcal{G} - [\mathcal{G}^+-]$differentiability of f_0 at x_0. We fix $\varepsilon > 0$ and $\mathcal{W} \in \mathcal{G}$. Then there exists $E_\mathcal{W} \in C_0(\mathbb{R})$ such that $|e(\alpha; x_0, y)| \leq |E_\mathcal{W}(\alpha)|$ for all $y \in \mathcal{W}$ and all $\alpha \in \mathbb{R}$ $[\alpha > 0]$. Fix $\rho > 0$, and let $\gamma \geq 1$ be a Lipschtz constant for f on $\mathcal{B}_\rho(x_0)$. Choose $\alpha \in (0, \rho)$ small enough that $E(\alpha) < \varepsilon/3$. Define $\delta := \min\{\rho - \alpha, \alpha\varepsilon/6\gamma\}$. Assume that $(f, x) \in \mathcal{B}_\delta((f_0, x_0))$ and that $\xi' \in \partial f(x)$. Fix $y \in \mathcal{W}$. Now let $\xi'_y = \nabla f_0(x_0)$ [$\xi'_y \in \partial f_0(x_0)$ satisfy $\langle\xi'_y|y\rangle = df(x; y)$ for our fixed $y \in \mathcal{W}$]. Upon rearranging (2.2) with f_0 and x_0 and using (2.1) we get the inequality

$$\begin{aligned}\langle\xi' - \xi'_y|y\rangle &\leq \frac{f(x + \alpha y) - f(x)}{\alpha} - \frac{f_0(x_0 + \alpha y) - f_0(x_0)}{\alpha} + e(\alpha; x, y)\\ &\leq \frac{f(x + \alpha y) - f_0(x + \alpha y)}{\alpha} + \frac{f_0(x) - f(x)}{\alpha} +\\ &\quad \frac{f_0(x + \alpha y) - f_0(x_0 + \alpha y)}{\alpha} + \frac{f_0(x_0) - f_0(x)}{\alpha} + E_\mathcal{W}(\alpha) \qquad (2.6)\\ &\leq 2\delta/\alpha + 2\gamma\|x - x_0\|/\alpha + E_\mathcal{W}(\alpha)\\ &< \varepsilon.\end{aligned}$$

Since this holds for each $y \in \mathcal{W}$, the definition of $\mathcal{G}-$upper semi-continuity of mapping (2.5) at x_0 is satisfied. Moreover, if f_0 is $\mathcal{G}$-differentiable at x_0, then $\partial f_0(x_0) = \{\nabla f_0(x_0)\}$, and we can use similar arguments, with $\alpha \in (-\rho, 0)$ to obtain the estimate $-\varepsilon < \langle\xi'-\xi'_y|y\rangle$ for each $y \in \mathcal{W}$, so that (2.5) is $\mathcal{G}-$continuous at x_0.

Now assume that the mapping $(f,x) \mapsto \partial f(x) : \mathcal{K} \times Y \to \mathcal{C}$ is $\mathcal{G}-$continuous at (f_0, x_0). Fix $\mathcal{W} \in \mathcal{G}$. Choose $y \in \mathcal{W}$, and let $\xi'_y \in \partial f_0(x_0)$ satisfy $\langle \xi'_y | y \rangle = df(x; y)$. Then upon rearranging (2.2) with $\langle \xi'_y | y \rangle$ replacing $df(x_0; y)$ and applying Lemma 2.4, we get

$$e(\alpha; x_0, y) = \langle \xi'_\alpha - \xi'_y | y \rangle \tag{2.7}$$

for some $\tau \in [0,1]$ and some $\xi'_\alpha \in \partial f_0(x_0 + \tau\alpha y)$. Clearly, the $\mathcal{G}-$continuity of (2.5) implies that $e(\alpha; x_0, y) = \langle \xi'_\alpha - \xi'_y | y \rangle \to 0$ as $\alpha \to 0$ for each $y \in \mathcal{W}$. Let $E : \mathbb{R} \to \mathbb{R}$ be defined by

$$E_{\mathcal{W}}(\alpha) := \operatorname{sgn}(\alpha) \cdot \max_{y \in \mathcal{W}} |e(\alpha; x_0, y)| + \alpha. \tag{2.8}$$

Then we have that $E_{\mathcal{W}}$ vanishes continuously at the origin so that $E_{\mathcal{W}} \in C_0(\mathbb{R})$, so that we get $E_{\mathcal{W}}$ estimates e at x_0 with respect to $\mathcal{W}$. Since this construction can be carried out for each $\mathcal{W} \in \mathcal{G}$, we have by definition that f_0 is $\mathcal{G}-$differentiable at x_0. Construction (2.8) gives us the $\mathcal{G}^+$-differentiability for free, since $\mathcal{W} \in \mathcal{G}$ and since for each $y \in \mathcal{W}$ we get $e(\,\cdot\,; x_0, y)$ is monotonically increasing and continuous from the right at the origin. In this case we define $E_{\mathcal{W}} : \mathbb{R} \to \mathbb{R}$ by (2.8) for $\alpha \geq 0$ and by $E_{\mathcal{W}}(\alpha) = \alpha$ for $\alpha < 0$. □

Corollary 2.8. *Mapping* (2.5) *is* $\mathcal{G}-(\mathcal{H}-)$*upper semi-continuous at any* $(f,x) \in \mathcal{K} \times Y$.

We can exploit uniformity in $y \in \mathcal{S}_1$ to extend Theorem 2.7 to get

Theorem 2.9. *Let* f_0 *be a bounded, convex function on* Y *with* $\mathcal{V} \subseteq Y$ *and* $\rho > 0$. *Consider the following conditions.*

(i) f_0 *is uniformly* $\mathcal{F} - [\mathcal{F}^+-]$*differentiable on* $\mathcal{V}$ *and* $\cup_{x \in \mathcal{V}+\mathcal{B}_\rho} \partial f_0(x)$ *is bounded.*

(ii) *Mapping* (2.5) *is uniformly* $\mathcal{F}-$*[upper semi-] continuous on* $\{f_0\} \times \mathcal{V}$.

(iii) f_0 *is uniformly* $\mathcal{F} - [\mathcal{F}^+-]$*differentiable on* $\mathcal{V}$.

Then we have (i)⇒(ii)⇒(iii)*; however,* (i) *and* (ii) *are not equivalent.*

Proof. The proof that (i) ⇒ (ii) follows by modifying the arguments leading up to the inequalities in (2.6). Using the uniform estimate in (2.3) that we get on e, we can replace $\mathcal{W} \in \mathcal{G}$ everywhere by $\mathcal{S}_1$. In (2.6) we can estimate both

$$\frac{f_0(x_0) - f_0(x)}{\alpha} \qquad \text{and} \qquad \frac{f_0(x + \alpha y) - f_0(x_0 + \alpha y)}{\alpha}$$

by using Lemma 2.4 and the boundedness of $\cup_{x \in \mathcal{V}+\mathcal{B}_\rho} \partial f_0(x)$. This gives us the inequalies in (2.6) for each $y \in \mathcal{S}_1$, so that by taking the supremum over $y \in \mathcal{S}_1$ we get the definitions (modulo the equality that the supremum might give) required in (ii). The proof that (ii)⇒(iii) is also an immediate extension of Theorem 2.7.

To see that (ii)⇏(i) in general, it suffices to consider $f : \mathbb{R} \to \mathbb{R}$ where $f(x) = x^2/2$. Then $f(x+\alpha) - f(x) = x\alpha + \alpha^2/2$ so that f is uniformly $\mathcal{F}-$differentiable on $\mathbb{R}$. The mapping $(f,x) \mapsto \partial f(x) = \{x\}$ is uniformly $\mathcal{F}-$continuous on $\mathbb{R}$, but $\cup_{x \in \mathbb{R}} \partial f(x)$ is not bounded there. Boundedness of $\mathcal{V}$ is sufficient to ensure that (i) and (ii) are equivalent.

□

3. Convexity and the "Inverted" Subdifferential Mapping

In the previous section we considered the mapping $(f, x) \mapsto \partial f(x) : \mathcal{K} \times Y \to \mathcal{C}$ and saw that various types of continuity of this mapping at the point (f, x) in fact characterize the smoothness of f at $x \in Y$. Now for a fixed $f \in \mathcal{K}$ and $\mathcal{Z} \subseteq Y$ the mapping $\partial f : \mathcal{Z} \to \mathcal{C}$ defines a binary relation

$$\begin{aligned}\mathcal{R}_{f,\mathcal{Z}} &:= \operatorname{graph}(\partial f) \cap (\mathcal{Z} \times Y') \\ &= \{(x, \xi') \in \mathcal{Z} \times Y' : x \in \mathcal{Z} \text{ and } \xi' \in Y' \text{ with } \xi' \in \partial f(x)\} \\ &= \cup_{x \in \mathcal{Z}} \quad \{x\} \times \partial f(x)\end{aligned}$$

If we define $\mathcal{Z}' := \cup_{x \in \mathcal{Z}} \partial f(x) \subseteq Y'$ and let f^* be the conjugate of f, we see that the inverse relation

$$\begin{aligned}\mathcal{R}^{-1}_{f,\mathcal{Z}} &:= \operatorname{graph}(\partial f^*) \cap (Y' \times \mathcal{Z}) \\ &= \{(\xi', x) \in \mathcal{Z}' \times \mathcal{Z} : \xi' \in \mathcal{Z}' \text{ and } x \in \mathcal{Z} \text{ with } \xi' \in \partial f(x)\} \\ &= \cup_{\xi' \in \mathcal{Z}'} \quad \{\xi'\} \times \{x \in \mathcal{Z} : \xi' \in \partial f(x)\}\end{aligned}$$

defines implicitly a multivalued function $\partial f^*_{\mathcal{Z}} : \mathcal{Z}' \to 2^{\mathcal{Z}}$. Our aim in this section is to develop a theory which describes the continuity of the mapping

$$(f, \xi') \mapsto \partial f^*_{\mathcal{Z}}(\xi') : (\mathcal{K} \times \mathcal{Z}') \to 2^{\mathcal{Z}}, \tag{3.1}$$

and which parallels the theory developed in the last section. That is, given the appropriate concepts of continuity, we want to relate certain properties of f to the continuity of (3.1) at (f, ξ') or more generally on $\{f\} \times \mathcal{V}'$ for some $\mathcal{V}' \subseteq \mathcal{Z}'$. Once again our vehicle will be equation (2.2). In order to develop a more broadly applicable theory we consider $\mathcal{Z} \subseteq Y$, defined in terms of f, to be one of three examples. These are Y, the domain of f, as well as the level and tangent sets of f at some x_0, where

Definition 3.1. The *level set* of f is $\mathcal{L}_\alpha := \{x \in Y : f(x) = \alpha\}$.

Definition 3.2. The *tangent set* of f at x_0 is $\mathcal{T}_f(x_0) := \{x \in Y : df(x_0; x - x_0) = 0\}$.

We wish to study the geometry of the set $\mathcal{R}_{f,\mathcal{Z}}$ at the point $x_0 \in \mathcal{Z}$, so we assume that x_0 and $\mathcal{Z}$ have been chosen so that

$$\rho_{x_0} := \sup\{\rho \geq 0 : \mathcal{Z} \cap \mathcal{S}_\alpha(x_0) \neq \emptyset \text{ for all } \alpha \in [0, \rho]\} \in (0, \infty]. \tag{3.2}$$

Definition 3.3. In (3.2) we call ρ_{x_0} the *diameter* of $\mathcal{Z}$ at x_0, and we write $\operatorname{dia}(\mathcal{Z}_{x_0}) = \rho_{x_0}$.

If $\mathcal{Z} = Y$ then $\rho_{x_0} = \infty$ for each $x_0 \in Y$. Further, if $\mathcal{Z} = \mathcal{T}_f(x_0)$ and $0 \notin \partial f(x_0)$, then (assuming $\dim(Y) \geq 2$) we get $\rho_{x_0} = \infty$. Finally, if $\mathcal{Z} = \mathcal{L}_{f(x_0)}$ and $\rho_{x_0} = 0$, then x_0 is the unique minimum of f. Thus, if we take $\mathcal{Z}$ to be either of these last two cases, we make the assumption that $0 \notin \partial f(x_0)$. We define the family of subsets of the unit sphere

$$\begin{aligned}\mathcal{Z}_\alpha(x_0) &:= \{y \in \mathcal{S}_1 : x_0 + \alpha y \in \mathcal{Z}\} \\ &= (1/\alpha)\,[\,(\mathcal{Z} \cap \mathcal{S}_\alpha(x_0)) + \{-x_0\}\,].\end{aligned}$$

For notational reasons let $\mathcal{Z}_\alpha(x_0) = \emptyset$ for $\alpha \geq \rho_{x_0}$. Thus, if $\mathcal{Z} = Y$, then $\mathcal{Z}_\alpha(x_0) = \mathcal{S}_1$ for all $\alpha > 0$. If $\mathcal{Z} = \mathcal{T}_f(x_0)$, then $\mathcal{Z}_\alpha(x_0) = \mathcal{T}_f(x_0) \cap \mathcal{S}_1(x_0) + \{-x_0\}$ for all $\alpha > 0$.

3.1. *Convexity Concepts and the Bregman Distance Function*

In order to give a concise yet general concept of the convexity of a bounded, convex function f, we consider again e given by (2.2) and two nonempty sets $\mathcal{V} \subseteq \mathcal{Z}$ and $\mathcal{W} \subseteq \mathcal{S}_1$. If we set $\inf \emptyset = \infty$, then for each $\mathcal{W} \subseteq \mathcal{S}_1$ and $x \in \mathcal{V}$ we can define the increasing function (which is similar to the Bregman distance function) $\mu_{x,\mathcal{W}} : [0,\infty) \to [0,\infty]$ by

$$\mu_{x,\mathcal{W}}(\alpha) := \begin{cases} 0 & \text{if} \quad \alpha = 0 \\ \inf_{\beta \geq \alpha}\{e(\beta; x, y) : y \in \mathcal{W} \cap \mathcal{Z}_\beta(x)\} & \text{if} \quad \alpha > 0 \end{cases}$$

Definition 3.4. We say that e admits a $\mathcal{Z}-$defined (lower) estimate on $\mathcal{V}$ with respect to $\mathcal{W}$ if there is some $F_\mathcal{W} \in C_0(\mathbb{R})$ such that $\mu_\mathcal{W}(\alpha) := \inf_{x\in\mathcal{V}} \mu_{x,\mathcal{W}}(\alpha) \geq F_\mathcal{W}(\alpha)$ for each $\alpha > 0$. The $\mathcal{Z}-$defined *modulus of convexity* of f on $\mathcal{V}$ is $\mu_{\mathcal{S}_1}$, which we write μ. If $\mathcal{V} = \{x\}$, we say e admits a $\mathcal{Z}-$defined (lower) estimate at x with respect to $\mathcal{W}$.

Example 3.5. If f is the norm on a Hilbert space, then the modulus of convexity at any point of the unit sphere defined in terms of the unit sphere is $\mu(\alpha) = \alpha/2$.

Definition 3.6. If e admits a $\mathcal{Z}-$defined (lower) uniform estimate on $\mathcal{V}$ with respect to each

$$\mathcal{W} \in \begin{cases} \mathcal{G}, \text{ then } f \text{ is } \textit{uniformly } \mathcal{G}\textit{-convex} \text{ on } \mathcal{V} \text{ with respect to } \mathcal{Z}. \\ \mathcal{H}, \text{ then } f \text{ is } \textit{uniformly } \mathcal{H}\textit{-convex} \text{ on } \mathcal{V} \text{ with respect to } \mathcal{Z}. \\ \mathcal{F}, \text{ then } f \text{ is } \textit{uniformly } \mathcal{F}\textit{-convex} \text{ on } \mathcal{V} \text{ with respect to } \mathcal{Z}. \end{cases}$$

If in this definition $\mathcal{V} = \{x\}$, we say that f is convex at x in the sense under consideration. Note, too, that if $\mathcal{Z} = \mathcal{T}_f(x_0)$, then we always take $\mathcal{V} = \{x\}$.

Lemma 3.7. *The $\mathcal{G}-$convexity of f at x_0 with respect to $\mathcal{Z}$ (for each $\mathcal{Z}$) is sufficient to ensure that $\mathcal{L}_{f(x_0)} \cap \mathcal{T}_f(x_0) = \{x_0\}$. It is also necessary if $\mathcal{Z} \neq Y$.*

Proof. Assume f is $\mathcal{G}-$convex at x_0 with respect to $\mathcal{Z}$. Fix $x \in \mathcal{L}_{f(x_0)} \cap \mathcal{T}_f(x_0)$ and set $\bar{\alpha} := \|x - x_0\|$. Choose $y \in \mathcal{S}_1$ such that $\bar{\alpha}y = x - x_0$. Then $x_0 + \alpha y \in \mathcal{T}_f(x_0)$ for all $\alpha \in [0, \bar{\alpha}]$. If $\bar{\alpha} > 0$, then $df(x_0; y) = 0$, and $e(\alpha; x_0, y) = 0$ for all $\alpha \in [0, \bar{\alpha}]$, contradicting the $\mathcal{G}-$convexity of f at x_0. Thus $\bar{\alpha} = 0$ and $\mathcal{L}_{f(x_0)} \cap \mathcal{T}_f(x_0) = \{x_0\}$ as required. Note that we make no use of the assumption $0 \notin \partial f(x_0)$ up to this point.
Let $\mathcal{L}_{f(x_0)} \cap \mathcal{T}_f(x_0) = \{x_0\}$. If $\mathcal{Z} = \mathcal{T}_f(x_0)$, then for each $\alpha > 0$ and $y \in \mathcal{Z}_\alpha(x_0)$ we have $f(x_0+\alpha y) > f(x_0)$, so that $\alpha e(\alpha; x_0, y) = f(x_0+\alpha y) - f(x_0) > 0$. Thus, since $\mathcal{W} \in \mathcal{G}$, we can define $F_\mathcal{W}(\alpha) := \inf_{y\in\mathcal{W}} e(\alpha; x_0, y) \in C_0(\mathbb{R})$ for $\alpha \geq 0$ and $F_\mathcal{W}(\alpha) := -\alpha$ for $\alpha < 0$. Then $F_\mathcal{W}$ estimates $\mu_{x_0,\mathcal{W}}$ (below) at x_0 with respect to $\mathcal{T}_f(x_0)$. If $\mathcal{Z} = \mathcal{L}_{f(x_0)}$, then for each $\alpha > 0$ we get $[\mathcal{Z}_\alpha(x_0)+\{x_0\}]\cap\mathcal{T}_f(x_0) = \emptyset$, so that $y \in \mathcal{Z}_\alpha(x_0)$ implies $-df(x_0; y) > 0$. Thus, $\inf_{\beta\geq\alpha}\{e(\beta; x, y) : y \in \mathcal{W} \cap \mathcal{Z}_\beta(x_0)\} = \inf_{\beta\geq\alpha}\{-df(x_0; y) : y \in \mathcal{W} \cap \mathcal{Z}_\beta(x_0)\} > 0$. Since $\mathcal{W} \in \mathcal{G}$, this is positive for each $\alpha > 0$, so that we can find some $F_\mathcal{W} \in C_0(\mathbb{R})$ which estimates $\mu_{x_0,\mathcal{W}}$ (below) at x_0 with respect to $\mathcal{T}_f(x_0)$ □

Thus, we see that the $\mathcal{G}-$ convexity of f at x_0 with respect to $\mathcal{L}_{f(x_0)}$ is equivalent to the $\mathcal{G}-$ convexity of f at x_0 with respect to $\mathcal{T}_f(x_0)$.

Note that in both cases it is essential that $\mathcal{Z} \neq \{x_0\}$ so that f not be trivially $\mathcal{G}-$convex at x_0 with respect to $\mathcal{Z}$. Thus, $0 \notin \partial f(x_0)$ could be weakened to $\mathcal{Z} \neq \{x_0\}$. Note, too, that if $0 \in \partial f(x_0)$, then $\mathcal{L}_{f(x_0)} \subseteq \mathcal{T}_f(x_0)$, so $\mathcal{L}_{f(x_0)} \neq \{x_0\}$ if and only if $\mathcal{L}_{f(x_0)} \cap \mathcal{T}_f(x_0) \neq \{x_0\}$. Of course, the conditions $0 \in \partial f(x_0)$ and $\mathcal{L}_{f(x_0)} = \{x_0\}$ together are the same as requiring that x_0 is the unique minimum of f. Thus, if $0 \in \partial f(x_0)$, then the $\mathcal{G}-$convexity of f at x_0 with respect to $\mathcal{L}_{f(x_0)}$ or $\mathcal{T}_f(x_0)$ is equivalent to f having a unique minimum at x_0. As is the case if f is a norm, if $0 \in \partial f(x_0)$, then f can be (trivially) $\mathcal{G}-$convex with respect to $\mathcal{L}_{f(x_0)}$ but not with respect to Y.

If $\mathcal{Z} = Y$ we do not expect necessity in Lemma 3.7. To see this, consider the Euclidian norm on $\mathbb{R}^2$ with $x_0 = (1, 0)$. Then $\mathcal{L}_{f(x_0)}$ is the unit sphere and $\mathcal{T}_f(x_0) = \{(1, \tau) : \tau \in \mathbb{R}\}$. Thus, we get that $\mathcal{L}_{f(x_0)} \cap \mathcal{T}_f(x_0) = \{x_0\}$. That f in this example is not $\mathcal{G}-$convex at x_0 with respect to $\mathbb{R}^2$ follows from

Lemma 3.8. *If $\mathcal{Z} = Y$, then $\mathcal{G}-$convexity at x_0 is exactly strict convexity at x_0 as expressed by the strict inequality*

$$f(\lambda x_0 + (1-\lambda)(x_0 + y)) < \lambda f(x_0) + (1-\lambda) f(x_0 + y) \tag{3.3}$$

for all $\lambda \in (0, 1)$ and all $y \in Y \backslash \{0\}$.

Proof. By setting $\alpha = 1 - \lambda$ and $\beta = \|y\|$ we see that (3.3) is equivalent to the strict inequality $e(\beta\alpha; x_0, y) < e(\beta; x_0, y)$ for all $\alpha \in (0, 1)$, for all $\beta > 0$ and for all $y \in \mathcal{S}_1$. Thus, strict convexity at x_0 implies $e(\alpha; x_0, y) > 0$ for all $\alpha > 0$, so that given $\mathcal{W} \in \mathcal{G}$ we can construct a (lower) estimate $F_{\mathcal{W}} \in C_0(\mathbb{R})$ for $\mu_{x_0,\mathcal{W}}$ just as in Lemma 3.7.

If on the other hand f is not strictly convex at x_0, then we can find some $\lambda \in (0, 1)$ and $y \neq 0$ such that equality holds in (3.3). Convexity then implies that for this y we get equality for all $\lambda \in [0, 1]$. This means that $e(\alpha; x_0, y/\|y\|) = 0$ for all $\alpha \in [0, \|y\|]$, violating the $\mathcal{G}-$convexity of f at x_0 with respect to Y. □

In analog to the previous section, it is interesting to note the equivalence of $\mathcal{G}-$ and $\mathcal{H}-$convexity with respect to $\mathcal{Z}$. We also note that the $\mathcal{Z}-$defined modulus of convexity of f on $\mathcal{V}$ is monotonically increasing, since $\mu_{x,\mathcal{S}_1}$ is monotone for $x \in \mathcal{V}$ and $\mu := \inf_{x \in \mathcal{V}} \mu_{x,\mathcal{S}_1}$.

Lemma 3.9. *The $\mathcal{Z}-$defined modulus of convexity of f on $\mathcal{V}$ is continuous at the origin.*

Proof. Note that in as much as $\mu := \inf_{x \in \mathcal{V}} \mu_{x,\mathcal{S}_1}$, it suffices to show this at a point. The results follow from monotonicity if f is not $\mathcal{F}-$convex at x_0, so we assume it is. First we let $\mathcal{Z} = Y$ or $\mathcal{Z} = \mathcal{T}_f(x_0)$. Then $\mathcal{Z}_\alpha(x_0)$ is independent of $\alpha > 0$, so we can write $0 \leq \mu_{x_0,\mathcal{S}_1}(\alpha) \leq \mu_{x_0,\{y\}}(\alpha) := e(\alpha; x_0, y)$ for any $y \in \mathcal{Z}_\alpha(x_0)$, so that the results follow by the continuity of $e(\,\cdot\,; x_0, y)$ at the origin.

Now assume $\mathcal{Z} = \mathcal{L}_{f(x_0)}$. Choose $\tilde{y} \in \mathcal{S}_1$ such that $x_0 + \tilde{y} \in \mathcal{T}_f(x_0)$. Choose $\hat{y} \in \mathcal{S}_1$ such that $x_0 + \hat{\alpha}\hat{y} \in \mathcal{Z}_{\hat{\alpha}}(x_0)$ for some $\hat{\alpha} \in (0, \rho_{x_0})$. Then for each $\alpha \in [0, \hat{\alpha}]$ we can find some $y_\alpha \in \mathcal{W} := \mathcal{S}_1 \cap \{x : x = \hat{\nu}\hat{y} + \tilde{\nu}\tilde{y}$ where $\hat{\nu}, \tilde{\nu} \geq 0\}$ such that $x_0 + \alpha y_\alpha \in \mathcal{L}_{f(x_0)}$. Since $\mathcal{W} \in \mathcal{H}$, we can find some $E_{\mathcal{W}} \in C_0(\mathbb{R})$ which estimates e (above) with respect to $\mathcal{W}$. From this we get the estimate $\mu_{x_0,\mathcal{S}_1}(\alpha) \leq df(x_0; y_\alpha) =: e(\alpha; x_0, y_\alpha) \leq E_{\mathcal{W}}(\alpha)$ which means that $\mu_{x_0,\mathcal{S}_1}$ is continuous at the origin. Note that here $0 \notin \partial f(x_0)$ is assumed. □

As a consequence of this lemma and the monotonicity of μ, if f is uniformly $\mathcal{F}-$convex on $\mathcal{V}$, then for each $\alpha \geq 0$ we can define $\hat{\mu}(\alpha) := \alpha\mu(\alpha)$, which is strictly increasing in $\alpha \in \mathbb{R}^+$ and continuous at the origin. Thus, we can define $\hat{\mu}^{-1} : \mathbb{R}^+ \to \mathbb{R}^+$ by

$$\hat{\mu}^{-1}(\beta) := \alpha \quad \text{for all} \quad \beta \in [\lim_{\gamma\to\alpha^-} \hat{\mu}(\gamma), \lim_{\gamma\to\alpha^+} \hat{\mu}(\gamma)], \tag{3.4}$$

if for notational convenience, we agree that $\lim_{\gamma\to 0^-} \mu(\gamma) = 0$. Clearly $\hat{\mu}^{-1}$ is positive on $(0,\infty)$, is monotonically increasing on $\mathbb{R}^+$ and is continuous at the origin with $\hat{\mu}^{-1}(0) = 0$.

3.2. Convexity and Continuity of the Inverse Mapping

We observe that for $x_0, x_1 \in \mathcal{Z}$ and $\xi' \in \mathcal{Z}'$ we have the equivalence of $x_0, x_1 \in \partial f^*_{\mathcal{Z}}(\xi')$ and $\xi' \in \partial f(x_0) \cap \partial f(x_1)$. Thus, the cardinality of $\partial f^*_{\mathcal{Z}}(\xi')$ being one for any $\xi' \in \partial f(x_0)$ is equivalent to the condition $\partial f(x_0) \cap \partial f(x_1) = \emptyset$ for all $x_1 \in \mathcal{Z}\backslash\{x_0\}$. This leads us to the dual equivalent to the statement that the function f is $\mathcal{G}-$differentiable at a point if and only if the subdifferential mapping is single valued there.

Theorem 3.10. *The cardinality of the inverse mapping (3.1) is one on $\{f\} \times \partial f(x_0)$ if and only if f is $\mathcal{G}$-convex at x_0 with respect to $\mathcal{Z}$.*

Proof. In the light of Lemma 3.8 this is classical for $\mathcal{Z} = Y$ [3]. Thus we consider only the case $\mathcal{Z} = \mathcal{L}_{f(x_0)}$ (and hence $\mathcal{Z} = \mathcal{T}_f(x_0)$). We assume $0 \notin \partial f(x_0)$ throughout.

First assume f is $\mathcal{G}$-convex at x_0 with respect to $\mathcal{Z} = \mathcal{L}_{f(x_0)}$. Let $\xi' \in \partial f(x_0) \cap \partial f(x)$ for some $x \in \mathcal{Z}$. Define the function ϕ by $\phi(\tau) := f(x_0 + \tau(x - x_0))$ for each $\tau \in \mathbb{R}$. Since $\phi(0) = \phi(1)$, we get $D^+\phi(0) \leq 0 \leq D^-\phi(1)$. Now, the fact that $\xi' \in \partial f(x_0) \cap \partial f(x_1)$ gives the inequalities $D^-\phi(1) \leq \langle \xi' | x - x_0 \rangle \leq D^+\phi(0)$. Thus, monotonicity of $D^+\phi$ means that $D^+\phi \equiv 0$ on $[0,1)$, so that $x \in \mathcal{T}_f(x_0)$, and hence $x = x_0$ by Lemma 3.7.

Now assume the cardinality of the inverse mapping (3.1) is one on $\{f\} \times \partial f(x_0)$. Let $x \in \mathcal{L}_{f(x_0)} \cap \mathcal{T}_f(x_0)$. Choose $\xi'_0 \in \partial f(x_0)$ such that $\langle \xi'_0 | x - x_0 \rangle = df(x_0; x - x_0) = 0$. For all $y \in Y$ we can write

$$\begin{aligned} \langle \xi'_0 | y \rangle &= \langle \xi'_0 | x + y - x_0 \rangle \\ &\leq f(x + y - x_0 + x_0) - f(x_0) \\ &= f(x + y) - f(x). \end{aligned}$$

Thus, $\xi'_0 \in \partial f(x)$ by definition. Since $x \in \mathcal{Z}$, single valuedness of (3.1) implies $x = x_0$. But this is true for all $x \in \mathcal{L}_{f(x_0)} \cap \mathcal{T}_f(x_0)$, so that by Lemma 3.7 we get that f is $\mathcal{G}-$convex at x_0 with respect to $\mathcal{Z}$. □

Unfortunately, the dual equivalent of Theorem 2.7 is not so readily availible. The temptation is to replace weak-* concepts in Y' by their weak analogs in Y, and thereby be able to characterize the $\mathcal{G}-$continuity of the inverse mapping (3.1) in terms of the $\mathcal{G}-$convexity of f at a point. It turns out that $\mathcal{G}-$convexity is too weak for this to be sucessful. We now state the definition of $\mathcal{G}-[\mathcal{F}-]$continuity of mapping (3.1) and show that this is sufficient for the $\mathcal{G}-$convexity at a point. We then go on to show that in the proper setting this is necessary, too.

Fix $f_0 \in \mathcal{K}$. Let $x_0 \in \mathcal{V}_0 \subseteq \mathcal{Z}_0$ and define $\mathcal{V}_0' := \cup_{x \in \mathcal{V}_0} \partial f_0(x)$. Similarly for some $f \in \mathcal{K}$ let $\mathcal{V} \subseteq \mathcal{Z}$ and define $\mathcal{V}' := \cup_{x \in \mathcal{V}} \partial f(x)$. We also define the families $\mathcal{G}'$ and $\mathcal{F}'$ of subsets of the unit sphere $\mathcal{S}_1'$ in Y' in the obvious way. Note that in (2.4) we make use of a set $\mathcal{V}$, which is not necessarily the same $\mathcal{V}$ as here. We will always make clear what $\mathcal{V}$ in (2.4) is.

Definition 3.11. Mapping (3.1) is uniformly $\mathcal{G}-[\mathcal{F}-]$continuous on $\{f_0\} \times \mathcal{V}_0'$ if for each $\varepsilon > 0$ and $\mathcal{W}' \in \mathcal{G}'$ $[\mathcal{F}']$ there exists a $\delta > 0$ (independent of $\xi_0' \in \mathcal{V}_0'$) such that if $\xi_0' \in \mathcal{V}_0'$, if $(f, \xi') \in \mathcal{B}_\delta((f_0, \xi_0'))$, if $x \in \partial f_{\mathcal{Z}}^*(\xi')$, if $x_0 \in \partial f_{0\,\mathcal{Z}_0}^*(\xi_0')$ and if $|f(x) - f_0(x_0)| < \delta$, then we get that $\sup_{\eta' \in \mathcal{W}'} |\langle \eta' | x - x_0 \rangle| < \varepsilon$.

Note that the Euclidian norm on $\mathbb{R}^2$ is $\mathcal{G}-$convex at any point x_0 of the unit sphere with respect to the unit sphere, but this will not guarantee the continuity of (3.1) without the condition $|f(x) - f_0(x_0)| < \delta$. As we mentioned, the $\mathcal{G}-$continuity of (3.1) ensures the $\mathcal{G}-$convexity of f. We now prove this fact.

Theorem 3.12. *If the inverse mapping* (3.1) *is (uniformly)* $\mathcal{G}-$*continuous on* $\{f\} \times \partial f(x_0)$ *for some* $x_0 \in \mathcal{Z}$*, then* f *is* $\mathcal{G}-$*convex at* x_0 *with respect to* $\mathcal{Z}$*.*

Proof. First we assume $\mathcal{Z} = \mathcal{L}_{f(x_0)}$ or $\mathcal{Z} = \mathcal{T}_f(x_0)$. Fix $x \in \mathcal{L}_{f(x_0)} \cap \mathcal{T}_f(x_0)$. Choose $\xi' \in \mathcal{S}_1'$ such that $\langle \xi' | x - x_0 \rangle = \|x - x_0\|$ and $\xi_0' \in \partial f(x_0)$ with $\langle \xi_0' | x - x_0 \rangle = 0$. We can show (as in Lemma 3.10) that $\xi_0' \in \partial f(x)$. Thus, the $\mathcal{G}-$continuity of (3.1) implies, upon choosing $\mathcal{W}' = \{\xi'\}$, that $\|x - x_0\| < \varepsilon$ for all $\varepsilon > 0$ so that $\mathcal{L}_{f(x_0)} \cap \mathcal{T}_f(x_0) = \{x_0\}$. At this point Lemma 3.7 gives the results.

Now let $\mathcal{Z} = Y$. Fix $\bar{y} \in \mathcal{S}_1$ and assume $e(\alpha; x_0, \bar{y}) = 0$ for all $\alpha \in [0, \bar{\alpha}]$, where $\bar{\alpha} \geq 0$. Let $\xi_0' \in \partial f(x_0)$ satisfy $\langle \xi_0' | \bar{y} \rangle = df(x_0; \bar{y})$. Using the definition of $\xi_0' \in \partial f(x_0)$ and of e we get the inequality

$$
\begin{aligned}
\langle \xi_0' | y \rangle &\leq f(x_0 + y) - f(x_0) \\
&= f(x_0 + \alpha\bar{y} + y - \alpha\bar{y}) - f(x_0 + \alpha\bar{y}) + f(x_0 + \alpha\bar{y}) - f(x_0) \\
&= f(x_0 + \alpha\bar{y} + y - \alpha\bar{y}) - f(x_0 + \alpha\bar{y}) + \alpha df(x_0; \bar{y}),
\end{aligned}
$$

which is valid for all $y \in Y$ and all $\alpha \in [0, \bar{\alpha}]$. Thus, if we set $\tilde{y} = y - \alpha\bar{y}$, we get by definition that $\xi_0' \in \partial f(x_0 + \alpha\bar{y})$ for each $\alpha \in [0, \bar{\alpha}]$. By $\mathcal{G}-$continuity we get $\bar{\alpha} = 0$. This gives us the $\mathcal{G}-$convexity of f at x_0 with respect to Y. □

The next example indicates that $\mathcal{G}-$convexity alone is too weak to assure the $\mathcal{G}-$continuity of mapping (3.1).

Example 3.13. Let $f : \ell^2 \to \mathbb{R}$ be given by $f(x) := \langle e_1, x \rangle + \sum_{j=1}^{\infty} \frac{1}{j}(x_j)^2$ for $x = (x_j)_{j \in \mathbb{N}} \in \ell^2$. It is easy to show that f is a bounded, convex function, which is

$\mathcal{G}-$convex at the origin with respect to $\mathcal{L}_0$. However, if we let $x_L = (\sqrt{1/4 - L^{-1/2}} - 1/2)e_1 + L^{1/4}e_L \in \mathcal{L}_0$, we see that $x_L \not\rightharpoonup 0$. This together with the fact that

$$\nabla f(x_L) = (\sqrt{1/4 - L^{-1/2}} + 1/2)e_1 + L^{-3/4}e_L \to e_1 = \nabla f(0)$$

gives us a contradiction to the sufficency of $\mathcal{G}-$convexity for the weak continuity of mapping (3.1) in the dual equivalent of Theorem 2.7 even in the Hilbert space ℓ^2. The problem here being that the level sets of f are not bounded. In fact, this alone (even on a reflexive space) is not enough as a variation on the last example shows.

Example 3.14. Let $f_L : \ell^2 \to \mathbb{R}$ be given by $f_L(x) := \max\{\sum_{j=1}^{L} \frac{1}{j}(x_j)^2, \|x\| - L\}$ and let $f_0(x) := \max\{\sum_{j=1}^{\infty} \frac{1}{j}(x_j)^2, \|x\| - 1\}$. We see that f_0 is $\mathcal{G}-$convex at the origin and that $f_L \to f_0$ uniformly on the closed unit ball. If we set $x_0 = 0$ and $x_L := L^{1/4}e_L$ for each $L \in \mathbb{N}$, then $f_L(x_L) \to 0 = f_0(x_0)$ and $\nabla f_L(x_L) = 2L^{-3/4} \to 0 = \nabla f_0(x_0)$. It is however clear that $x_L \not\rightharpoonup x_0$.

What we actually need is some condition that ensures that $\{x_L\}$ is bounded (Y reflexive) and all of its weak cluster points are in the set $\mathcal{L}_{f_0(x_0)} \cap \mathcal{T}_{f_0}(x_0)$. Then Lemma 3.7 gives the results. The following notation will be useful.

Definition 3.15. The *α-union* of f is $\mathcal{U}_\alpha := \{x \in Y : f(x) \le \alpha\}$.

One way to accomplish the goals outlined above is

Theorem 3.16. *Let f_0 be a bounded convex function with bounded level sets defined on the reflexive Banach space Y. Given a sequence of bounded convex functions $\{f_L\}$ and a null sequence $\{\delta_L\}$ such that $|f_L(x) - f_0(x)| < \delta_L$ for all $x \in \mathcal{U}_\alpha$, where $\alpha > f_0(x_0)$, assume that there is $\{x_L\}$ with $\xi'_L \in \partial f_L(x_L)$ such that $\xi'_L \to \xi'_0 \in \partial f_0(x_0)$. If $f_L(x_L) \to f_0(x_0)$, say $|f_L(x_L) - f_0(x_0)| \le \delta_L$, and if f_0 is $\mathcal{G}-$convex at x_0 with respect to $\mathcal{Z}$, then $x_L \rightharpoonup x_0$.*

Proof. Let $\alpha_L := \|x_L - x_0\|$ and choose $y_L \in \mathcal{S}_1$ such that $\alpha_L y_L = x_L - x_0$. By the boundedness of the level sets of f_0 we can let $\bar{\beta} := \sup_{x \in \mathcal{L}_\alpha} \|x - x_0\|$. If we choose $\beta_L > 0$ such that $x_0 + \beta_L y_L \in \mathcal{L}_\alpha$ then we get the inequality

$$\begin{aligned} \frac{\alpha - f_0(x_0) - 2\delta_L}{\bar{\beta}} &\le \frac{f_L(x_0 + \beta_L y_L) - f_L(x_0)}{\beta_L} \\ &\le \frac{f_L(x_L) - f_L(x_0)}{\alpha_L} \quad \to 0 \end{aligned}$$

for all $L \in \mathbb{N}$ such that $\alpha_L \ge \beta_L$. Thus, there are only finitely many L such that $x_L \notin \mathcal{U}_\alpha$.

Now let $\bar{x} \in \mathcal{U}_\alpha$ be any weak cluster point of $\{x_L\}$. Then since $|f_L(x_L) - f_0(x_L)| \le \delta_L$ for all but finitely many L, we can write that $f_0(\bar{x}) \le \liminf f_0(x_L) = f_0(x_0)$. Thus, $\bar{x} \in \mathcal{U}_{f_0(x_0)}$. This means that $\langle \xi' | \bar{x} - x_0 \rangle \le 0$ for all $\xi' \in \partial f_0(x_0)$. On the other hand, from (2.1) we get $\langle \xi'_L | x_0 - x_L \rangle \le f_L(x_0) - f_L(x_L) \to 0$. Since $\xi'_L \to \xi'_0$ and $x_L \rightharpoonup \bar{x}$ (for some subsequence), we get that $\langle \xi'_0 | \bar{x} - x_0 \rangle = 0$. Thus, we have $\bar{x} \in \mathcal{T}_f(x_0) \cap \mathcal{U}_{f(x_0)} = \mathcal{T}_f(x_0) \cap \mathcal{L}_{f(x_0)} = \{x_0\}$. This means that x_0 is the only weak cluster point of the bounded sequence $\{x_L\}$. $\square$

Corollary 3.17. *Let f_0 be a bounded convex function with bounded level sets defined on the reflexive Banach space Y. Assume $\mathcal{V}$ in* (2.4) *is $\mathcal{U}_\alpha$, where $\alpha > f_0(x_0)$. If f_0 is $\mathcal{G}-$convex at x_0 with respect to $\mathcal{Z}$, then the inverse mapping* (3.1) *is $\mathcal{G}-$continuous at (f_0, ξ_0') for each $\xi_0' \in \partial f_0(x_0)$.*

Proof. This follows from the previous result and the fact that the weak topology on Y is described by sequences. □

As a corollary to Lemma 3.7 we have the equivalence of $\mathcal{G}-$convexity at a point with respect to the level set and the tangent set. The following lemma shows this carries over to $\mathcal{F}-$convexity. This will simplify the proof of the strong analog of Theorem 3.16.

Lemma 3.18. *The convex function f is $\mathcal{F}-$convex at x_0 with respect to $\mathcal{T}_f(x_0)$ if and only if it is $\mathcal{F}-$convex at x_0 with respect to $\mathcal{L}_{f(x_0)}$.*

Proof. Assume f is not $\mathcal{F}-$convex at x_0 with respect to $\mathcal{L}_{f(x_0)}$. Then there exists $\alpha > 0$ such that the modulus of convexity of f at x_0 is zero at α, which by the monotonicity of this modulus, we can assume has been chosen small enough that f is Lipschitz on $\mathcal{B}_{2\alpha}(x_0)$. Without loss of generality, we assume that there is a sequence $\{x_L\} \subset \mathcal{B}_\alpha(x_0) \cap \mathcal{L}_{f(x_0)}$ such that $0 = \mu(\alpha) = \lim e(\alpha; x_0, y_L)$, where $\alpha y_L = x_L - x_0$. Now define the bounded convex (sublinear) function g on Y by $g(x) = df(x_0; x - x_0) = \max_{\partial f(x_0)} \langle \xi' | x - x_0 \rangle$. Thus, we get that $g(x_L) \to 0^-$. Choose $\hat{x} \in \mathcal{S}_\alpha(x_0)$ so that $g(\hat{x}) < g(x_L)$ for L large enough. For such L we choose z_L ($\neq x_0$) such that z_L, x_L and $\hat{x}$ are co-linear and such that $g(z_L) = 0$. This gives us the inequalities

$$\begin{aligned} \frac{g(x_L) - g(\hat{x})}{2\alpha} &\leq \frac{g(x_L) - g(\hat{x})}{\|x_L - \hat{x}\|} \\ &\leq \frac{g(z_L) - g(x_L)}{\|z_L - x_L\|} \end{aligned}$$

for large L. Thus, we see that $\|z_L - x_L\| \to 0$. Now we exploit the Lipschitz continuity of f to get $|f(x_L) - f(z_L)| \leq \gamma \|z_L - x_L\|$. This means that $\{z_L\} \in \mathcal{T}_f(x_0)$ and $\|z_L - x_0\| \geq \alpha/2$ for large enough L. Thus, if we let $\bar{y}_L = (z_L - x_0)/\|z_L - x_L\|$, then $e(\alpha/2; x_0, \bar{y}_L) \to 0$, so that f is not $\mathcal{F}-$convex at x_0 with respect to $\mathcal{T}_f(x_0)$. The converse of this is proved in a similar fashion. □

We will need some notation. Fix $x_0 \in Y$ and $\xi_0' \in \partial f(x_0)$. For each $\delta \in \mathbb{R}$ we define the hyperplane and the family thereof by

$$\begin{aligned} \mathcal{H}_\delta = \mathcal{H}_\delta(\xi_0') &:= \{x \in Y : \langle \xi_0' | x - x_0 \rangle = -\delta\} \\ \mathcal{H}_I = \mathcal{H}_I(\xi_0') &:= \{x \in Y : x \in \mathcal{H}_\delta \text{ for some } \delta \in I\}. \end{aligned}$$

If $I = [0, \delta]$ for some $\delta > 0$, we get $\sup\{\|x - x_0\| : x \in \mathcal{H}_I(\xi_0') \cap \mathcal{L}_{f(x_0)}\} \leq \hat{\mu}^{-1}(\delta)$ from inequality (3.4) for all $\delta > 0$ and for all $\xi_0' \in \partial f(x_0)$.

As we will be approximating a function f by a sequence of such functions $\{f_L\}$, we adopt the notation $\mathcal{L}_\alpha^L := \{x \in Y : f_L(x) = \alpha\}$ for the level sets of the approximating functions.

We let $\mathcal{U} := \cup_{\beta \leq f(x_0)} \mathcal{L}_\beta$, and similarly we let $\mathcal{U}^L := \cup_{\beta \leq f_L(x_L)} \mathcal{L}^L_\beta$. Finally, we use the notation $\mathcal{H}_{2I} := \cup_{\delta \geq \hat{\mu}(\alpha)} \mathcal{H}_\delta$ and $\mathcal{H}_I := \cup_{\delta \geq \hat{\mu}(\alpha)/2} \mathcal{H}_\delta$ to represent these half spaces.

We now consider a strong version of Theorem 3.16. We will see that $\mathcal{F}-$convexity is a much stronger condition, allowing us to relax many of the conditions needed with $\mathcal{G}-$convexity.

Theorem 3.19. *Let f_0 be a bounded convex function defined on a Banach space Y. Given a sequence of bounded convex functions $\{f_L\}$ and a null sequence $\{\delta_L\}$ such that $|f_L(x) - f_0(x)| < \delta_L$ for all $x \in \mathcal{B}_\rho(x_0)$, where $\rho > 0$, assume that there is $\{x_L\}$ with $\xi'_L \in \partial f_L(x_L)$ such that $\xi'_L \to \xi'_0 \in \partial f_0(x_0)$. If $f_L(x_L) \to f_0(x_0)$, say $|f_L(x_L) - f_0(x_0)| \leq \delta_L$, and if f_0 is $\mathcal{F}-$convex at x_0 with respect to $\mathcal{Z}$, then $x_L \to x_0$.*

Proof. We first consider the case $\mathcal{Z} = \mathcal{L}_{f_0(x_0)}$. Let $\alpha \in (0, \rho/2)$ be given. We show that there are only finitely many $\alpha_L := \|x_L - x_0\| \geq 2\alpha$. Now $\mathcal{F}-$convexity at x_0 ensures that $\mathcal{U} \cap \mathcal{S}_\alpha(x_0) \subset \mathcal{U} \cap \mathcal{H}_{2I}$. Uniform convergence assures us that there is some $L_1 \in \mathbb{N}$ such that $\mathcal{U}^L \cap \mathcal{S}_\alpha(x_0) \subset \mathcal{U}^L \cap \mathcal{H}_I$ for all $L \geq L_1$. We extend this by the convexity of f_L to $\mathcal{U}^L \cap (\mathcal{B}_\alpha(x_0))^c \subset \mathcal{U}^L \cap \mathcal{H}_I$ for all $L \geq L_1$.

Since $0 \notin \partial f_0(x_0)$, we can find a $\hat{x}_0 \in \mathcal{B}_\alpha(x_0)$ with $f_0(\hat{x}_0) < f_0(x_0)$. Uniform convergence of the sequence $\{f_L\}$ to f_0 guarantees that we can find a zero sequence $\{\tau_L\}$ and an $L_2 \geq L_1$ such that $\bar{x}_L := \tau_L(\hat{x}_0 - x_0) + x_0 \in \mathcal{L}^L_{f_L(x_L)} \cap \mathcal{B}_\alpha(x_0)$ for all $L \geq L_2$.

For each $L \in \mathbb{N}$ such that $x_L \notin \{x_0, \bar{x}_L\}$ we define $\bar{y}_L := (x_L - \bar{x}_L)/\|x_L - \bar{x}_L\|$ and $y_L := (x_L - x_0)/\alpha_L$. One can readily see that $\|\bar{y}_L - y_L\| \leq 2\|\bar{x}_L - x_0\|/\alpha_L$.

Now consider the set of all $L \geq L_2$ for which $\lambda_L := 2\alpha/\alpha_L \leq 1$, which we denote by $\mathcal{J}$. We will show this set to be finite. By the convexity of each $\mathcal{U}^L$ we have $\lambda(x_L - \bar{x}_L) + \bar{x}_L \in \mathcal{U}^L$ for any $\lambda \in [0, 1]$. We choose $\lambda = \lambda_L$ in this expression and write

$$\begin{aligned}
\lambda_L(x_L - \bar{x}_L) + \bar{x}_L &= \lambda_L(\alpha_L y_L + x_0 - \bar{x}_L) + (\bar{x}_L - x_0) + x_0 \\
&= (\alpha y_L + x_0) + [\alpha y_L + (1 - \lambda_L) \cdot (\bar{x}_L - x_0)].
\end{aligned}$$

Now the term in brackets is an element of $\mathcal{B}_\alpha(\alpha y_L)$ for all $L \in \mathcal{J}$. Thus, for such L we get the set inclusion $\lambda_L(x_L - \bar{x}_L) + \bar{x}_L \in \mathcal{U}^L \cap (\mathcal{B}_\alpha(x_0))^c \subset \mathcal{U}^L \cap \mathcal{H}_I$. If we note that $0 \leq \langle \xi'_L | x - \bar{x}_L \rangle$ for all $x \in \mathcal{L}^L_{f_L(x_L)}$ and all $L \in \mathbb{N}$, specifically for $x_L \in \mathcal{L}^L_{f_L(x_L)}$, we get the inequality $0 \leq \langle \xi'_L | \bar{y}_L \rangle$, which lets us write

$$\begin{aligned}
\hat{\mu}(\alpha)/2 &\leq - \langle \xi'_0 | \lambda_L(x_L - \bar{x}_L) + \bar{x}_L - x_0 \rangle \\
&\leq - \langle \xi'_0 | 2\alpha y_L + (1 - \lambda_L)(\bar{x}_L - x_0) \rangle \\
&\leq 2\alpha \left[\langle \xi'_L - \xi'_0 | \bar{y}_L \rangle + \langle \xi'_0 | \bar{y}_L - y_L \rangle \right] + |\langle \xi'_0 | \bar{x}_L - x_0 \rangle| \\
&\leq 2\alpha \|\xi'_L - \xi'_0\| + 3\|\xi'_0\| \cdot \|\hat{x}_0 - x_0\| \cdot |\tau_L|
\end{aligned}$$

Since $|\tau_L| \to 0$ and $\|\xi'_L - \xi'_0\| \to 0$ but $\hat{\mu}(\alpha) > 0$, it follows that $\mathcal{J}$ is finite. Thus, we have $\{x_L\} \cap (\mathcal{B}_{2\alpha(x_0)})^c$ is finite for each $\alpha \in (0, \rho/2)$, or in other words $x_L \to x_0$.

Now consider the case $\mathcal{Z} = Y$. Again, we let $\alpha_L = \|x_L - x_0\|$ and choose $y_L \in \mathcal{S}_1$ such that $\alpha_L y_L = x_L - x_0$. Now define e_L in terms of f_L as in (2.2). Then if μ is the modulus of convexity, we can write

$$\begin{aligned}\mu(\alpha_L) &\leq e_0(\alpha_L; x_0, y_L) + e_L(\alpha_L; x_0, -y_L)\\ &= \frac{f_0(x_L) - f_0(x_0) - f_L(x_L) + f_L(x_0)}{\alpha_L} - df_L(x_L; -y_L) - df_0(x_0; y_L)\\ &\leq \frac{2\delta_L}{\alpha_L} + \|\xi'_L - \xi'_0\|\end{aligned}$$

This of course implies that $x_L \to x_0$. □

It should be noted that we have shown that for any $\varepsilon > 0$ there is an L_ε such that if $L \geq L_\varepsilon$, then $\|x_L - x_0\| < \varepsilon$ and this L_ε is independent of $\xi'_0 \in \partial f(x_0)$. Under suitable conditions this theorem can be made uniform in $x \in \mathcal{V}$. Similar in spirit and proof we get

Corollary 3.20. *Let f_0 be a bounded convex function defined on the Banach space Y. Assume $\mathcal{V}$ in (2.4) is $\mathcal{B}_\rho(x_0)$, where $\rho > 0$. If f_0 is $\mathcal{F}-$convex at x_0 with respect to $\mathcal{Z}$, then the inverse mapping (3.1) is uniformly $\mathcal{F}-$continuous on $\{f_0\} \times \partial f_0(x_0)$.*

In the next section we will invert this last result in the case of norms in the process of deriving a local Smulyan theorem. That is, we will show that the bornological approach to smoothness and convexity allows for the characterization locally of one of these properties in the primal space by the dual property in the dual space.

4. Applications

We begin this section by giving a sufficient condition for $\mathcal{F}-$convexity at a point x_0. This condition reminds us of the definition of uniform convexity in [2] and [5] for the case of the norm on Y. We go on to consider the duality between differentiability and convexity.

4.1. Local Uniform Rotundity

Once again the Mean Value Theorem provides us with the machinery to obtain

Theorem 4.1. *If f is uniformly rotund at x_0 with respect to $\mathcal{Z}$, that is if there exists a modulus function $\delta \in C_0(\mathbb{R})$ such that for all $x \in \mathcal{Z}$ we have*

$$f\left(\frac{x + x_0}{2}\right) \leq \frac{f(x) + f(x_0)}{2} - \delta\left(\|x - x_0\|\right), \tag{4.1}$$

then f is $\mathcal{F}-$convex at x_0 with respect to $\mathcal{Z}$.

Proof. We consider inequality (4.1), where we assume that $x \neq x_0$. Define the modulus function $\mu_{x_0} : [0, \infty) \to [0, \infty)$ by

$$\mu_{x_0}(\alpha) := \begin{cases} 2\delta(\alpha)/\alpha & \text{if } \alpha > 0,\\ 0 & \text{if } \alpha = 0.\end{cases}$$

Since $\delta(\alpha) > 0$ for all $\alpha > 0$, we get that $\mu_{x_0}(\alpha) > 0$ for all $\alpha > 0$. We will simplify our notation by defining $\alpha := \|x - x_0\| > 0$ and $y := (x - x_0)/\alpha$.

If $\mathcal{Z} = Y$, we rearrange (4.1) and use the results of Section 2.2 to get the inequality

$$\begin{aligned}\mu_{x_0}(\alpha) &\leq \frac{f(x_0+\alpha y)-f(x_0)}{\alpha} - \frac{f(x_0+(\alpha/2)\cdot y)-f(x_0)}{\alpha/2} \\ &\leq \frac{f(x_0+\alpha y)-f(x_0)}{\alpha} - \langle \xi' | y\rangle,\end{aligned}$$

which is valid for all $\alpha > 0$, for all $\xi' \in \partial f(x_0)$ and for all $y \in \mathcal{S}_1$.

If $\mathcal{Z} = \mathcal{L}_{f(x_0)}$, let $x \in \mathcal{Z} \cap \mathcal{S}_\alpha(x_0)$. From (2.1) and (4.1) we see that for all $\xi' \in \partial f(x_0)$, we can derive the inequality

$$\begin{aligned}\left\langle \xi' \mid \frac{x-x_0}{2} \right\rangle &\leq f\left(\frac{x-x_0}{2}+x_0\right) - f(x_0) \\ &\leq -\delta\left(\|x-x_0\|\right).\end{aligned}$$

Dividing this by $\alpha/2$ and rearranging gives us $\mu_{x_0}(\alpha) \leq -\langle \xi' | y\rangle$ for all $\xi' \in \partial f(x_0)$ and all $y \in \mathcal{S}_1$ with $x_0 + \alpha y \in \mathcal{L}_{f(x_0)}$. If $\mathrm{dia}(\mathcal{L}_{f(x_0)}) \neq \infty$, we must redefine μ_{x_0} a bit but it is obvious how to do this. The proof for $\mathcal{Z} = \mathcal{T}_f(x_0)$ is similar to this proof. □

We note that Theorem 4.1 and Lemma 3.9 together show the modulus function δ is necessarily $o(\alpha)$ as $\alpha \to 0^+$. For a different proof of this see [6].

We also note that (4.1) is by no means necessary for the $\mathcal{F}-$convexity at $x_0 \in Y$ as we will see in the next example. We motivate the claim that local uniform rotundity as developed in [10] does not sufficiently capture the local duality at play between smoothness and convexity of dual norms. It is designed to measure the convexity of the norm at a point $x_0 \in \mathcal{S}_1$ in terms of points at some distance from x_0. Thus, the smoothness of the norm on a neighborhood of x_0 can play a role in measuring its convexity, and clearly for a satisfactory duality theory, differentiability should play no such role in the primal space (only in the dual space). In our definition, however, we examine the subdifferential of f at x_0, and thereby avoid the unfortunate situation of involving the smoothness of f on a neighborhood of x_0 in the measure of the convexity of f at x_0.

Example 4.2. We will introduce an equivalent norm f on ℓ^2 which is $\mathcal{F}-$convex at e_1 but not uniformly rotund there. We start by defining the family of subsets of ℓ^2

$$\mathcal{Z}_j := \{\, x := (x_i)_{i\in\mathbb{N}} \in \ell^2 : (\ |x_1|^{(j+1)/j} + |x_j|^{(j+1)/j}\ ) \leq 1\} \cap \bar{\mathcal{B}}_1$$

for each $j \geq 2$. Then we define $\mathcal{Z} := \cap_{j\geq 2}\mathcal{Z}_j$. It is not difficult to see that the set inclusion

$$\bar{\mathcal{B}}_{\sqrt{1/2}} \subset \mathcal{Z} \subset \bar{\mathcal{B}}_1 \tag{4.2}$$

is valid. Next we define f on ℓ^2 by $f(x) := \inf\{\alpha > 0 : x/\alpha \in \mathcal{Z}\}$, which, due to (4.2), induces a norm on ℓ^2 equivalent to that derived from the scalar product. If we define $x_L := (e_1 + e_L)/2 \in \ell^2$ for each $L \geq 2$, then it is easy to see that $f(x_L) = 2^{-1/(L+1)} \to 1$.

Since $f(e_j) = 1$ for all $j \in \mathbb{N}$, we get that f is not uniform rotund at e_1. Now let $\xi' \in \partial f(e_1)$. For all $j \in \mathbb{N}$ and for all $\beta \in (0,1)$ the inequalities

$$\frac{1 - f(e_1 - \beta e_j)}{\beta} \le \langle \xi' | e_j \rangle \le \frac{f(e_1 + \beta e_j) - 1}{\beta}$$

are valid. Since $f(e_1 \pm \beta e_j) = (1 + \beta^{(j+1)/j})^{j/(j+1)}$ for $j \geq 2$ and $f(e_1 \pm \beta e_1) = 1 \pm \beta$, we can use L'Hospital's rule and the above inequalities to get $\langle \xi' | e_j \rangle = \delta_{1,j}$ (Kronecker's δ), so for all $\xi' \in \partial f(e_1)$ we get $\langle \xi' | y \rangle = y_1$. We want to check the $\mathcal{F}-$convexity at e_1, so we consider any $y \in \mathcal{S}_1$ with $e_1 + \alpha y \in \mathcal{L}_1$. Since $\mathcal{L}_1 \subset \bar{\mathcal{B}}_1$, we have $y_1 < 0$, and we get some $\beta \geq 1$ such that $e_1 + \beta\alpha y \in \mathcal{S}_1$. Now, one can easily check that the norm on any Hilbert space is $\mathcal{F}-$convex at each $x_0 \in \mathcal{S}_1$ with respect to the unit sphere, and for such a norm the modulus of convexity μ_H is exactly $\mu_H(\alpha) = \alpha/2$. Putting the last few ideas together gives

$$\alpha/2 = \mu_{\ell^2}(\alpha) \le \mu_{\ell^2}(\beta\alpha) \le -\langle e_1, y \rangle = |y_1| = -\langle \xi' | y \rangle$$

for all $\xi' \in \partial f(e_1)$ and for all $y \in \mathcal{S}_1$ such that $e_1 + \alpha y \in \mathcal{L}_1$. Thus, we get that f is $\mathcal{F}-$convex at e_1, but not uniformly rotund there. □

It is interesting to note that if f is uniformly rotund at x_0 and $\mathcal{F}^+-$differentiable there, then f is also $\mathcal{F}-$convex at x_0.

Of course, one of the nice features of a uniformly convex Banach space, that is a Banach space whose norm satisfies (4.1) for all $x, x_0 \in \mathcal{S}_1$, is that if $\{x_L\} \subset \mathcal{S}_1$ and $x_L \rightharpoonup x_0 \in \mathcal{S}_1$, then $x_L \to x_0$. Theorem 4.1 gives us an elementary proof of this fact.

Corollary 4.3. *Let f be a norm, which is uniformly rotund at $x_0 \in \mathcal{S}_1$, locally uniformly convex in the sense of [10]. If $\{x_L\} \subset \mathcal{S}_1$ is any sequence such that $x_L \rightharpoonup x_0$, then $x_L \to x_0$.*

Proof. Since uniform rotundity of f at x_0 implies the $\mathcal{F}-$convexity of f there, we can define μ as in Theorem 4.1 to get $\|x_L - x_0\| \cdot \mu(\|x_L - x_0\|) \le |\langle \xi' | x_L - x_0 \rangle|$. This and $x_L \rightharpoonup x_0 \in \mathcal{S}_1$ give the norm convergence we want. □

4.2. Applications to Banach Space Geometry

We are now ready to consider more closely the special case where f is the norm on the Banach space Y. We will use both f and $\|\cdot\|$ to denote this norm (the notation f reminds us that the norm is a convex function, and $\|\cdot\|$ measures the distance between points). Likewise, we will use f^* and $\|\cdot\|$ to denote the norm on Y', the dual of Y. We assume $\mathcal{V} \subseteq \mathcal{S}_1$ and define $\mathcal{V}' := \cup_{x \in \mathcal{V}} \partial f(x) \subseteq \mathcal{S}'_1$. We remind the reader of some of the notation of Sections 2 and 3 that we will employ. In this notation

$\xi' \mapsto \partial f^*(\xi')$ is exactly the subdifferential mapping from $\mathcal{S}'_1$ to $\mathcal{S}''_1$, and

$\xi' \mapsto \partial f^*_{\mathcal{S}_1}(\xi')$ is the restriction of this of this mapping to $\mathcal{V}'$ (with range $\mathcal{S}_1$).

We will need the following generalization of a lemma which can be found in [5].

Lemma 4.4. *Let f be a norm on Y, which is uniformly $\mathcal{F}-$differentiable on $\mathcal{V} \subseteq \mathcal{S}_1$. For every $\varepsilon > 0$ there exists a $\delta > 0$ such that if $\xi' \in \mathcal{S}'_1$ and $\langle \nabla f(x) - \xi' | x \rangle < \delta$, then $\|\nabla f(x) - \xi'\| < \varepsilon$. Further, this δ is independent of $x \in \mathcal{V}$.*

Proof. Assume $\|\xi' - \nabla f(x)\| = \varepsilon$. Choose $y \in \mathcal{S}_1$ such that $\langle \xi' - \nabla f(x)|y\rangle \geq \varepsilon/2$. Define $\alpha := 4\langle \nabla f(x) - \xi'|x\rangle/\varepsilon > 0$. Since $\xi' \notin \partial f(x) = \{\nabla f(x)\}$, α is positive. Then we have

$$\begin{aligned}\langle \nabla f(x) - \xi'|x\rangle &\leq \langle \nabla f(x) - \xi'|x\rangle[4\langle \xi' - \nabla f(x)|y\rangle/\varepsilon - 1] \\ &= \langle \xi' - \nabla f(x)|x\rangle + \alpha\langle \xi' - \nabla f(x)|y\rangle \\ &= \langle \xi'|x + \alpha y\rangle - \|x\| - \alpha\langle \nabla f(x)|y\rangle \\ &\leq \|x + \alpha y\| - \|x\| - \alpha\langle \nabla f(x)|y\rangle \\ &\leq \alpha E(\alpha).\end{aligned}$$

Thus, we derive the inequality $E^{-1}(\varepsilon/4) \leq \alpha$, which gives us the inequality

$$\varepsilon E^{-1}(\varepsilon/4)/4 \leq \langle \nabla f(x) - \xi'|x\rangle. \tag{4.3}$$

If we define the function $\hat{E} \in C_0(\mathbb{R})$ implictly by $\hat{E}^{-1}(\alpha) := [\alpha E^{-1}(\alpha)]$ for all $\alpha > 0$, then inequality (4.3) yields $\|\xi' - \nabla f(x)\| \leq 4\hat{E}(\ \langle \nabla f(x) - \xi'|x\rangle\ )$, which proves our claim. □

We start by giving a second sufficient condition for $\mathcal{G} - [\mathcal{F}-]$convexity for the case of the norm on a Banach space. This is a local version of Smulyan's theorem, which states that a Banach space with a uniformly Fréchet differentiable dual norm is a uniformly convex space, that is its norm satisfies (4.1) at each $x_0 \in \mathcal{S}_1$ with a uniform modulus δ.

Theorem 4.5. *In the above notation let f^* be $\mathcal{G}$-differentiable on $\mathcal{V}'$ [uniformly $\mathcal{F}$-differentiable on $\mathcal{V}'$]. Then f is $\mathcal{G} - $[uniformly $\mathcal{F}-$]convex on $\mathcal{V}$ with respect to $\mathcal{S}_1$.*

Proof. The $\mathcal{G}$-differentability of f^* on $\mathcal{V}'$, Theorem 2.7 and the isometric embedding of Y in Y'' give the $\mathcal{G}-$ (or norm to weak) continuity of $\xi' \mapsto \partial f^*_{\mathcal{S}_1}(\xi')$ on $\mathcal{V}'$, so that the results follow by a direct application of Theorem 3.12.

If we assume uniform $\mathcal{F}$-differentiability on $\mathcal{V}'$, then we can derive μ directly from the upper estimate E that we can get for e in (2.2). To these ends we embed $\mathcal{S}_1$ in $\mathcal{S}_1''$ and use Lemma 4.4 applied to Y'. Choose $\alpha \in (0, 2)$, choose $x \in \mathcal{V}$, choose $\xi' \in \partial f(x)$ and choose $y \in \mathcal{S}_1$ such that $x + \alpha y \in \mathcal{S}_1$. By the $\mathcal{F}$-differentiability of f^* at ξ' we have $x = \nabla f^*(\xi')$, so that (4.3) together with the pairing $\langle \xi'|y\rangle = \langle y|\xi'\rangle$ gives us

$$\begin{aligned}\alpha E^{-1}(\alpha/4)/4 \ &\leq \langle \nabla f^*(\xi') - (x + \alpha y)|\xi'\rangle \\ &= \ -\alpha\langle \xi'|y\rangle,\end{aligned} \tag{4.4}$$

where $E \in C_0(\mathbb{R})$ is the uniform estimate ensured by the smoothness of f^*. Define $\mu : [0, \infty) \to \mathbb{R}$ by $\mu(\alpha) := E^{-1}(\alpha/4)/4 > 0$ for all $\alpha > 0$. Thus, inequality (4.4) shows that $\mu(\alpha) \leq \inf_{\beta \geq \alpha}\{e(\beta; x, y) : y \in \mathcal{S}_1 \cap \mathcal{Z}_\beta(x)\}$ for all $x \in \mathcal{V}$. This means that by definition f is uniformly $\mathcal{F}-$convex on $\mathcal{V}$ with respect to the unit sphere. □

We have shown not only the implications that the smoothness of the dual norm has for the convexity of the primal norm, but we also found the modulus of convexity of the primal norm in terms of the estimate on the smoothness of the dual norm. We will turn this around shortly and show how to derive an estimate of the smoothness of the primal norm in terms of the modulus of convexity of the dual norm. Our first result reiterates Corollary 3.20, while adding uniformity in $x \in \mathcal{V}$. It is actually valid for any bounded, convex function.

Lemma 4.6. *Let f be uniformly $\mathcal{F}-$convex on $\mathcal{V} \subseteq \mathcal{S}_1$. Then $\xi' \mapsto \partial f^*_{\mathcal{S}_1}(\xi')$ is norm to norm uniformly continuous on $\mathcal{V}' := \cup_{x \in \mathcal{V}} \partial f(x)$.*

Proof. Let $x \in \mathcal{S}_\alpha(x_0)$ for some $\alpha > 0$ and define $y := (x - x_0)/\alpha$. Using the moduli of convexity of f on $\mathcal{V}$ and at x we can write $\mu_x(\alpha) \leq \langle \xi' | y \rangle$ for any $\xi' \in \partial f(x)$, and we can write $\mu_{\mathcal{V}}(\alpha) \leq -\langle \xi'_0 | y \rangle$ for any $\xi'_0 \in \partial f(x_0)$. This gives

$$\begin{aligned} \mu_{\mathcal{V}}(\alpha) &\leq \mu_{\mathcal{V}}(\alpha) + \mu_x(\alpha) \\ &\leq \langle \xi' - \xi'_0 | y \rangle \\ &\leq \|\xi' - \xi'_0\|. \end{aligned} \tag{4.5}$$

Now since $\mu_{\mathcal{V}}(\alpha) > 0$ if $\alpha > 0$ and continuous at the origin, we can invert it (taking the obvious precautions if $\mu_{\mathcal{V}}$ is constant or discontiuous anywhere), so that from (4.5) we get

$$\|x - x_0\| \leq \mu_{\mathcal{V}}^{-1}(\|\xi' - \xi'_0\|) \tag{4.6}$$

for any $x \in \mathcal{S}_\alpha(x_0)$, for any $\xi'_0 \in \partial f(x_0)$ and for any $\xi' \in \partial f(x)$. Since $\mu_{\mathcal{V}}$ is the modulus of convexity for all $x_0 \in \mathcal{V}$, we get norm to norm uniform continuity that we seek. □

4.3. Local Duality and Smulyan's Theorem

Using the notation of Section 4.2, we can summarize our results. We begin with the weak equivalencies (for a reflexive space).

Theorem 4.7. *Consider the statements:*

(i) *f is $\mathcal{G}-$convex on $\mathcal{V}$ with respect to $\mathcal{S}_1$.*
(ii) *$\xi' \mapsto \partial f^*_{\mathcal{S}_1}(\xi')$ is single valued on $\mathcal{V}'$.*
(iii) *$\xi' \mapsto \partial f^*(\xi')$ is single valued on $\mathcal{V}'$.*
(iv) *f^* is $\mathcal{G}-$differentiable on $\mathcal{V}'$.*
(v) *$\xi' \mapsto \partial f^*(\xi')$ is norm to weak* continuous on $\mathcal{V}'$.*
(vi) *$\xi' \mapsto \partial f^*_{\mathcal{S}_1}(\xi')$ is norm to weak continuous on $\mathcal{V}'$.*

Then we have (i)$\Leftrightarrow$(ii)$\Leftarrow$(iii)$\Leftrightarrow$(iv)$\Leftrightarrow$(v)$\Leftrightarrow$(vi), *and all are equivalent if Y is reflexive.*

Proof. In Theorem 3.10 we demonstrated (i)$\Leftrightarrow$(ii) pointwise for any bounded, convex function. The implication (ii)$\Leftarrow$(iii) is trivial, and we get equivalency if Y is reflexive. In [11] the author shows that (iii)$\Leftrightarrow$(iv) pointwise for any convex function. Theorem 2.7 gives (iv)$\Leftrightarrow$(v) pointwise. The implication (v)$\Rightarrow$(vi) is trivial, since the weak* topology on $\mathcal{S}''_1 \cap \mathcal{S}_1$ is the same as the weak topology on $\mathcal{S}_1$. We get the equivalency by an argument similar to that showing (ii)$\Rightarrow$(iii) in the next result. Theorem 3.12 gives (vi)$\Rightarrow$(i) pointwise for any bounded, convex function with equivalency if Y is reflexive. □

Theorem 4.8. (A local Smulyan result) *The following are equivalent:*

(i) *f is uniformly $\mathcal{F}-$convex on $\mathcal{V}$ with respect to $\mathcal{S}_1$.*

(ii) $\xi' \mapsto \partial f^*_{\mathcal{S}_1}(\xi')$ *is norm to norm uniformly continuous on* $\mathcal{V}'$.
(iii) $\xi' \mapsto \partial f^*(\xi')$ *is norm to norm uniformly continuous on* $\mathcal{V}'$.
(iv) f^* *is uniformly* $\mathcal{F}-$*differentiable on* $\mathcal{V}'$.

Proof. The implication (i)⇒(ii) follows from Lemma 4.6. By arguments similar to those in [5], or by a modification of Theorem 2.9 we get (iii)⇒(iv). Theorem 4.5 gives us the implication (iv)⇒(i). That (ii)⇒(iii) can be derived by applying monotonicity of the subdifferential mapping and Bishop-Phelps in the following way. Consider some $\delta-$ball $\mathcal{B}_\delta(\xi'_0)$ about $\xi'_0 \in \mathcal{S}'_1$. We let $\hat{Y}$ be the embedding of Y in Y'' and let $\mathcal{X}''$ be the weak* closure of the convex hull of the set $\partial f^*(\mathcal{B}_\delta(\xi'_0)) \cap \hat{Y}$. Suppose there is $\eta''_0 \in \partial f^*(\mathcal{B}_\delta(\xi'_0))\backslash\mathcal{X}''$. Choose $\xi' \in \mathcal{S}'_1$ such that $\langle \eta''_0|\xi'\rangle > \alpha > \max_{\mathcal{X}''}\langle \eta''|\xi'\rangle$ for some $\alpha > 0$. Clearly, we can find a $\eta' \in \mathcal{B}_\delta(\xi'_0)$ such that $\eta'' \in \partial f^*(\eta')$. Choose $\beta > 0$ small enough that $\eta' + \beta\xi' \in \mathcal{B}_\delta(\xi'_0)$. Monotonicity of the subdifferential operator gives

$$\begin{aligned}\alpha &< \langle \eta''|\xi'\rangle \\ &\leq \max_{\partial f^*(\eta')} \langle \omega''|\xi'\rangle \\ &\leq \min_{\partial f^*(\eta'+\beta\xi')} \langle \omega''|\xi'\rangle.\end{aligned}$$

By weak* upper semi-continuity (see Theorem 2.7 or [11]) of the subdifferential mapping, there exists an open neighborhood V of $\eta' + \beta\xi'$ such that $\inf_{\partial f^*(V)}\langle \omega''|\xi'\rangle > \alpha$. But this contradicts the Bishop-Phelps theorem which says that $\partial f^*(V) \cap \mathcal{X}'' \neq \emptyset$. Thus, we have $\partial f^*(\mathcal{B}_\delta(\xi'_0)) \subseteq \mathcal{X}''$. Now, (ii) means that for every $\varepsilon > 0$ there is some $\delta > 0$ such that if $\xi' \in \mathcal{V}'$ and $x \in \partial f(\xi')$.then we get the set inclusion

$$\partial f^*_{\mathcal{S}_1}(\mathcal{B}_\delta(\xi')) \subseteq \mathcal{B}_\varepsilon(x).$$

Now we take the weak* closure of the convex hull of the above inclusion, and apply the results we just derived and we get exactly (iii). Note that the same ε and δ hold [7]. □

Next we show the impact that convexity of the dual norm has on the smoothness of the primal norm. This is the dual concept of Theorem 4.5 and allows us to derive a modulus of smoothness for the primal norm in terms of the modulus of convexity of the dual norm.

Theorem 4.9. *If f^* is $\mathcal{G}-$ [uniformly $\mathcal{F}-$] convex on $\mathcal{V}'$ with respect to the dual unit sphere, then f is $\mathcal{G}$- [uniformly $\mathcal{F}$-]differentiable on $\mathcal{V}$. For $\mathcal{G}-$convexity we need reflexivity.*

Proof. Reflexivity of Y and the $\mathcal{G}-$convexity of f^* on $\mathcal{V}'$ allows us to apply Theorem 4.7 to get this result. Likewise, we can apply Theorem 4.8 to get that the second dual norm is uniformly $\mathcal{F}-$differentiable on the restriction of $\mathcal{V}''$ back to $\mathcal{V} \subseteq \mathcal{S}_1$, giving the result. We now go on to derive a modulus of smoothness.

Assuming uniform convexity we can apply (4.6) to the proof of Theorem 2.7 to get

$$\begin{aligned}\alpha e(\alpha; x_0, y) &= f(x_0 + \alpha y) - f(x_0) - \alpha df(x_0); y) \\ &\leq \alpha\langle \xi'_\alpha - \xi'_0|y\rangle \\ &\leq |\alpha| \cdot \|\xi'_\alpha - \xi'_0\| \\ &\leq |\alpha| \cdot \mu_{\mathcal{V}}^{-1}(|\alpha|),\end{aligned}$$

where $\xi'_\alpha \in \partial f(x_0+\alpha y)$ and $\xi'_0 \in \partial f(x_0)$. Define the function $\hat{E}$ on the set $(-\delta,\delta)$ by $\hat{E}(\alpha) := \operatorname{sgn}(\alpha)\mu_{\mathcal{W}}^{-1}(|\alpha|)$ for some $\delta > 0$. Then $\hat{E}$ is continuous at the origin and satisfies $|\hat{E}(\alpha)| > 0$ if $\alpha > 0$. Thus, we can find some $E \in C_0(\mathbb{R})$ with $\alpha\hat{E}(\alpha) \leq \alpha E(\alpha)$ for all $\alpha \in [-\delta,\delta]$ and we get uniform $\mathcal{F}$-differentiability of f on $\mathcal{V}$. □

4.4. Applications to Convergence Rates

The last section puts us in a position to get a convergence rate for Theorem 3.19 in the case of the norm as follows.

Theorem 4.10. *Assume that f is uniformly $\mathcal{F}-$differentiable on $\mathcal{B}_\rho(x_0)$ for some $x_0 \in \mathcal{S}_1$ and some $\rho \in (0,1)$ and $\mathcal{F}-$convex with respect to $\mathcal{S}_1$ at x_0 with modulus μ. Let f_L be a sequence of norms on Y such that $|f_L(x) - f(x)| < \delta_L \to 0$ for all $x \in \mathcal{B}_\rho(x_0)$. Further, let $\{x_L\} \subset Y$ with $f_L(x_L) \to 1$ and $\xi'_L \to \nabla f(x_0)$ (where $\xi'_L \in \partial f_L(x_L)$), then we can derive a convergence rate in terms of $f(x_L)$, δ_L and $\|\xi'_L - \nabla f(x_0)\|$.*

Proof. By Theorem 3.19 there is some $L_0 \in \mathbb{N}$ such that $x_L \in \mathcal{B}_\rho(x_0)$ for $L \geq L_0$. For any such L we can define $\hat{x}_L := x_L/f(x_L) \in \mathcal{S}_1$. The results of Theorem 2.9 allow us to write $\|\xi'_L - \nabla f(x)\| \leq 2\sqrt{\delta_L} + E(\sqrt{\delta_L})$ for all $x \in \mathcal{B}_\rho(x_0)$. This estimate together with (4.6) and the positive homogeneity of f give

$$\begin{aligned}
\|x_L - x_0\| &\leq \|x_L - \hat{x}_L\| + \|\hat{x}_L - x_0\| \\
&\leq |f(x_L) - 1| + \mu^{-1}(\|\nabla f(x_L) - \nabla f(x_0)\|) \\
&\leq |f(x_L) - 1| + \mu^{-1}(\|\nabla f(x_L) - \xi'_L\| + \|\xi'_L - \nabla f(x_0)\|) \\
&\leq |f(x_L) - 1| + \mu^{-1}(2\sqrt{\delta_L} + E(\sqrt{\delta_L}) + \|\xi'_L - \nabla f(x_0)\|),
\end{aligned}$$

which is what we want to show. □

References

[1] F. Clarke: Optimization and Nonsmooth Analysis, John Wiley and Sons, 1983.

[2] M.M. Day: Some More Uniformly Convex Spaces, AMS Bulletin, 47 (1941) 504–507.

[3] V. Dem'yanov, L. Vasil'ev: Nondifferentiable Optimization, Opimization Software, 1986.

[4] R. Deville, G. Godefroy and V. Zizler: Smoothness and Renormings in Banach Spaces, Longman Scientific and Technical, London, 1993.

[5] J. Diestel: Geometry of Banach Spaces, Springer Verlag, Berlin, 1975.

[6] J. Diestel: Sequences and Series in Banach Spaces, Springer Verlag, Berlin, 1984.

[7] J. Giles, W. Moors: A Continuity Property Related to Kuratowski's Index of Noncompactness, Journal of Math Analysis and Applications 178 (1993) 274–268.

[8] W. Krabs, U. Lamp: On the Approximate Calculation of Time-Minimal Distributed Controls of Vibrations by Finite Element Approximations, Applied Mathematics and Optimization, 27 (1993) 179–190.

[9] W. Krabs, Z. Yan: Numerical Solution of Minimum Norm Problems of Distributed Control of Vibrations, Applied Mathematics and Optimization, 21 (1990) 243–264.

[10] A.R. Lovaglia: Locally Uniformly Convex Banach Spaces, Trans. Amer. Math. Soc., 78 (1955) 225–238.

[11] R.R. Phelps: Convex Functions, Monotone Operators, and Differentiability, Springer Verlag, Berlin, 1989.

[12] B.N. Pshenichnyi: Necessary Conditions for an Extremum, Marcel Dekker Co., New York, 1971.

[13] J. Read: The Approximation of Optimal Control of Vibrations - A Geometrical Approach, Mathematical Methods in the Applied Sciences, 19 (1996) 87–129.