SETS OF DIFFERENTIALS AND SMOOTHNESS
OF CONVEX FUNCTIONS

Wee-Kee Tang

Approximation by smooth convex functions and questions on the Smooth Variational Principle for a given convex function \( f \) on a Banach space are studied in connection with majorising \( f \) by \( C^1 \)-smooth functions.

It is known that a weakly compactly generated (WCG) Banach space admits an equivalent Fréchet differentiable norm if it admits a Fréchet differentiable bump function (see, for example [5]). However, there are nonseparable spaces that admit Fréchet differentiable bump functions but admit no equivalent Fréchet differentiable norm (see, for example [3, Chapter VII]). If the space \( X \) admits an equivalent norm with modulus of smoothness of power type 2, then every convex continuous function on \( X \) has points of Lipschitz smoothness (see, for example [3, Chapter IV]). The purpose of this note is to localise these results. We prove that any convex Lipschitz function \( f \) that is defined on a WCG Banach space \( X \) can be uniformly approximated by Fréchet differentiable convex functions if \( f \) is majorised on \( X \) by a Fréchet smooth convex function. If, moreover, span\( \{\partial f(x) : x \in X\} \) is a subspace of \( X^* \) that admits a norm with modulus of rotundity of power type 2, then there is a convex function \( \psi \) with \( \psi' \) Lipschitz on \( X \) such that \( \psi \geq f \) on \( X \) and \( \psi(x) = f(x) \) for some \( x \in X \). Thus in particular, \( f \) has points of Lipschitz smoothness. A separable version of these problems was studied in [10]. We use standard notation in this note (see for example [3]), and refer to [6, 7, 9] and [3] for some unexplained notions and results used in this note.

**Theorem 1.** Let \( f \) be a convex Lipschitz function defined on a WCG Banach space \( X \). Then the following are equivalent.

1. The function \( f \) can be uniformly approximated on \( X \) by a Fréchet differentiable convex function.
2. There exists a Fréchet differentiable convex function \( \phi \) defined on \( X \) such that \( \phi \geq f \) on \( X \).

Received 15th October, 1994

The author was supported by NSERC (Canada). The author is indebted to Professor V. Zizler for his guidance and suggestions, and to the referee for his/her helpful comments.

Copyright Clearance Centre, Inc. Serial-fee code: 0004-9729/95 $A2.00+0.00.
PROOF: Clearly (1) $\implies$ (2). The proof (2) $\implies$ (1) is divided into a few steps.

PROPOSITION 2. Let $X$ be a WCG Banach space, $\phi$ be a Fréchet differentiable convex function defined on $X$ and let $Y := \overline{\text{span}}\{\phi'(x) : x \in X\}$. Then there exists a projectional resolution of the identity (PRI) $\{P_\alpha : \omega \leq \alpha \leq \mu\}$ on $X$ such that

1. $P_\mu = I$, $\|P_\alpha\| = 1$ for all $\alpha$.
2. $P_\alpha^* P_\beta = P_\beta^* P_\alpha = P_{\min(\alpha, \beta)}$.
3. $P_\alpha Y \subset Y$ for all $\alpha$.
4. $\text{dens}(P_\alpha^* Y) \leq \alpha$ for all $\alpha \leq \mu$.
5. $P_\alpha^* Y = \overline{\text{span}} \bigcup_{\beta < \alpha} P_{\beta + 1} Y$ for all $\alpha \leq \mu$.

PROOF: Using standard techniques for constructing projectional resolutions of the identity (see for example [3, Chapter VI]), we only need to show $P_\alpha^* Y \subset Y$ and $P_\alpha^* Y = \overline{\text{span}} \bigcup_{\beta < \alpha} P_{\beta + 1} Y$ for all $\alpha$. The proof of this is contained in Lemmas 3 to 5.

LEMMA 3. In notation as above, we can construct a PRI $\{P_\alpha : \omega \leq \alpha \leq \mu\}$ so that $P_\alpha^* \phi'(y) = \phi'(y)$ for all $y \in P_\alpha X$.

PROOF: See Lemma 5 in [5].

LEMMA 4. With notation as above, $P_\alpha^* Y = \overline{\text{span}} \{\phi'(x) : x \in P_\alpha X\}$.

PROOF: To see $P_\alpha^* Y \supset \overline{\text{span}} \{\phi'(x) : x \in P_\alpha X\}$, we let $x \in P_\alpha X$, and show that $\phi'(x) \in P_\alpha^* Y$. Since $\phi$ is $C^1$-smooth, given $\varepsilon > 0$, there exists an $x_\beta \in P_{\beta + 1} X$ for some $\beta < \alpha$, such that $\|\phi'(x) - \phi'(x_\beta)\| < \varepsilon$. By Lemma 3, $\phi'(x_\beta) = P_{\beta + 1} \phi'(x_\beta)$. Therefore $\phi'(x_\beta) \in P_{\beta + 1}^* Y \subset P_\alpha^* Y$. As $P_\alpha^* Y$ is closed, $\phi'(x) \in P_\alpha^* Y$. For the converse inclusion, we follow the idea in [4]. Let $\phi'(x) \in Y$. Clearly $g(\cdot) = \phi(\cdot) - \phi'(x)(\cdot)$ is a continuous bounded below function on $X$. Hence its restriction $g|_{P_\alpha X}$ is also continuous and bounded below. By Ekeland's variational principle, given $\varepsilon > 0$, there exists $x_\alpha \in P_\alpha X$ such that for every $w \in B_{P_\alpha X}$, $t > 0$, we have $g(x_\alpha + tw) \geq g(x_\alpha) - ct$, thus, $\phi'(x)(w) \leq (\phi(x_\alpha + tw) - \phi(x_\alpha))/t - \varepsilon$. Hence, by taking limits, we have $\phi'(x)(w) - \phi'(x_\alpha)(w) \leq \varepsilon$. Therefore $\sup\{\|\phi'(x)(v) - \phi'(x_\alpha)(v)\| : v \in B_{P_\alpha X}\} \leq \varepsilon$. Given any $h \in B_X$, we have $(h, P_\alpha^* \phi'(x) - \phi'(x_\alpha)) = (h, P_{\beta + 1}^* \phi'(x) - P_\alpha^* \phi'(x_\alpha)) = (P_\alpha h, \phi'(x) - \phi'(x_\alpha)) \leq \varepsilon$. Therefore $P_\alpha^* \phi'(x) - \phi'(x_\alpha) \leq \varepsilon$. Finally, since $Y$ is the closed linear span of the derivatives of $\phi$ and $P_\alpha$ is bounded, the assertion follows.

LEMMA 5. $P_\alpha^* Y = \overline{\text{span}} \bigcup_{\beta < \alpha} P_{\beta + 1} Y$ for every $\alpha \leq \mu$.

PROOF: Clearly $P_\alpha^* Y \supset \overline{\text{span}} \bigcup_{\beta < \alpha} P_{\beta + 1} Y$. The converse inclusion follows from Lemma 4 and the continuity of $\phi'$.

PROOF OF THEOREM 1: Since $f \leq \phi$, using Ekeland's variational principle as in Lemma 4 we show that $\text{dom} f^* \subset Y$. Using Lemma 5, and the classical Troyanski's
construction (see for example [3, Chapter VII]) we obtain a dual norm \( \| \cdot \|^{*} \) in \( X^{*} \) such that its restriction on \( Y \) is locally uniformly rotund (LUR). Define a sequence of functions \( \{ h_{n} \} \) on \( X^{*} \) by \( h_{n}(x^{*}) = f^{*}(x^{*}) + \| x^{*} \|^{2}/(4n^{4}) \). Clearly, \( \text{dom } h_{n} = \text{dom } f^{*} \). Define \( g_{n} := f \Box n^{4} \| \cdot \|^{2} \), where \( \Box \) denotes the infimal convolution. Note that \( g_{n} \) is convex and continuous on \( X \) and \( g_{n}^{*} = h_{n} \) for all \( n \). Given \( n \in \mathbb{N} \), \( z \in X \) and \( y \in \partial g_{n}(z) \), note that \( h_{n} \) is rotund at \( y \) with respect to \( z \) in the sense of [1], that is, for every \( \varepsilon > 0 \), there exist \( \delta > 0 \) such that \( \{ v : h_{n}(y + v) - h_{n}(y) - (x, v) \leq \delta \} \subset \varepsilon B_{X} \) (see, [10]). By [1, Proposition 4], \( g_{n} \) is Fréchet differentiable at \( z \) with the derivative \( y \). One can also show that \( \lim g_{n} = f \) uniformly on \( X \) (see for example [8, Lemma 2.4]).

Since the function \( f \) can be quite “flat” in Theorem 1, there is a difficulty in applying the techniques of Smooth Variational Principles (see, [3, Chapter I]) in this situation. However, under more restrictive assumptions we can use the Stegall-Fabian variational principle and obtain our variational result by duality. We shall say that \( z \in X \) is a point of Lipschitz smoothness of a convex function \( f \) if \( f(z + h) + f(z - h) - 2f(z) = O(h^{2}) \).

**Lemma 6.** Let \( f \) be a convex continuous function on a Banach Space \( X \) and \( g \) be its dual function. Suppose there exists a constant \( C \) such that for any \( z \in X \), \( y \in \partial f(z) \), and for any \( \varepsilon > 0 \), we have

\[
\{ v : g(y + v) - g(y) - (z, v) \leq C\varepsilon^{2} \} \subset \varepsilon B_{X}^{*}.
\]

Then \( f \) is Fréchet differentiable and \( f' \) is Lipschitz on \( X \).

**Proof:** By taking polars, we have \( \varepsilon^{-1} B_{X} \subset \{ v : g(y + v) - g(y) - (z, v) \leq C\varepsilon^{2} \}^{\circ} \). According to Proposition 3 of [1], \( \{ u : f(z + u) - f(z) - (y, u) \leq C\varepsilon^{2} \}^{\circ} \subset C^{-1} \varepsilon^{-2} \{ u : f(z + u) - f(z) - (y, u) \leq C\varepsilon^{2} \} \), that is, for any \( u \in \varepsilon CB_{X} \), \( f(z + u) + f(z - u) - 2f(z) \leq 2/C(\varepsilon C)^{2} \). Thus \( f' \) exists at \( z \) and we have that \( f' \) is Lipschitz on \( X \) (see, for example [3, Lemma V.3.5]).

**Theorem 7.** Let \( f \) be a Lipschitz convex function on a Banach space \( X \) and \( Y = \overline{\text{span} \| \cdot \|^{*}} \{ \partial f(z) : z \in X \} \). Suppose that \( Y \) admits an equivalent norm with modulus of convexity of power type 2. Then \( f \) can be majorised by a convex function \( \psi \) that has a Lipschitz derivative and \( \psi(z) = f(z) \) for some \( z \in X \). In particular, \( f \) has points of Lipschitz smoothness.

**Proof:** Let \( \| \cdot \| \) be an equivalent norm on \( X^{*} \) such that its restriction on \( Y \) has modulus of convexity of power type 2 (see, for example [3, Lemma II.8.1]). We note that \( Y \) is \( w^{*} \)-closed. Indeed, since \( Y \) is reflexive, \( B_{Y} \) is compact in the weak topology...
of $X^*$ and thus $B_Y$ is $w^*$-compact in $X^*$. By the Banach-Dieudonné theorem, $Y$ is $w^*$-closed. Assume that $f(0) = 0$, and thus we have $f^* \geq 0$ on $X^*$. Let

$$h(x^*) = \begin{cases} \frac{1}{2} \| x^* \|^2 - \frac{1}{2} m^2 & \text{if } x^* \in Y \\ \infty & \text{otherwise,} \end{cases}$$

where $m = \text{Lip}(f)$. Since $Y$ is $w^*$-closed, $h$ is $w^*$-lower semicontinuous and $h = (h^*|^X)$. We show that $h$ satisfies the condition on the function $g$ given in Lemma 6. Indeed, by the modulus of rotundity of $\| \cdot \|$, there exists $L > 0$ such that for any $y_1, y_2 \in Y$, we have

$$\frac{1}{2} \left( \| y_1 \|^2 + \| y_2 \|^2 \right) - \| \frac{y_1 + y_2}{2} \|^2 \geq L \| y_1 - y_2 \|^2 \quad (\star)$$

(see, for example [2, Lemma 5.1.4]). Assume that for every $k \in \mathbb{N}$ there exist $\epsilon_k > 0$, $x_k \in X$, $y_k \in \partial h^*_X(x_k)$ and $v_k \in X^*$, $\| v_k \| > \epsilon_k$, such that $h(y_k + v_k) - h(y_k) - v_k(x_k) \leq \epsilon_k^2 / k$. Then $\| y_k + v_k \|^2 / 2 - \| y_k \|^2 / 2 - (x_k, v_k) \leq \epsilon_k^2 / k$ for all $k$. From the definition of a subdifferential, we have $- (x_k, v_k) \geq \| y_k \|^2 - \| y_k + v_k / 2 \|^2$. Therefore,

$$\left( \| y_k \|^2 + \| y_k + v_k \|^2 \right) / 2 - \| y_k + v_k / 2 \|^2 \leq \epsilon_k^2 / k \leq \| v_k \|^2 / k,$n

which contradicts $(\star)$. Now, for each $x^* \in \text{dom } f^* \subset mbX^*$, we have $h(x^*) \leq 0 \leq f(x^*)$. Therefore $f^* - h$ is a lower semicontinuous and convex function on $\text{dom } f^*$ that is bounded below. Note that $f^* - h \geq \| \cdot \| - m$. By the Stegall-Fabian result (see, for example [9, Corollary 5.22]), there exists $\widehat{x} \in Y^*$ such that $f^* - h - \widehat{x}$ attains its minimum in $\text{dom } f^*$, that is, there is a $x^* \in \text{dom } f^*$ such that $f^*(x^*) - h(x^*) - \widehat{x}(x^*) = \alpha \leq f^*(y^*) - h(y^*) - \widehat{x}(y^*)$ for all $y^* \in \text{dom } f^*$. Therefore we have $h(\cdot) + \widehat{x}(\cdot) + \alpha \leq f^*(\cdot)$ on $\text{dom } f^*$ and the equality holds at $x^*$. Since $Y$ is reflexive, there exists $x \in X$ such that $y^*(x) = \widehat{x}(y^*)$ for each $y^* \in Y$. Let $k : X^* \rightarrow \mathbb{R}$ be a function defined by $k(\cdot) = h(\cdot) + x(\cdot) + \alpha$. Then $k$ is a convex function such that $k \leq f^*$ and $k(x^*) = f^*(x^*)$. Put $l = k|_Y$. The function $l$ is continuous and convex on $Y$. Let $\widehat{y} \in \partial l(x^*) \subset Y^*$. As $Y$ is reflexive, there exists $y \in X$ such that $\widehat{y}(y^*) = y^*(y)$ for each $y^* \in Y$. We claim that $y \in \partial k(x^*)$. Indeed, let $z^* \in X^*$. If $z^* \in Y$, $y(z^* - z^*) = \widehat{y}(z^* - z^*) \leq k(z^*) - k(x^*)$. If $z^* \notin Y$, then $y(z^* - z^*) < k(z^*) - k(x^*) = \infty$ Hence $y \in \partial k(x^*)$. Since $k(x^*) = f^*(x^*)$, we have $y \in \partial f^*(z^*)$. Thus $k^*(y) + k(x^*) = (z^*, y) = f^*(z^*) + f(y)$. Therefore $f(y) = k^*(y)$. Since $f^* \geq k$, we have $k^* \geq f$. Put $\psi = k^*|^X$. The function $\psi$ has a Lipschitz derivative and is our required function. Indeed, $k^* = (h^* + x^* + \alpha)^* = (h + x)^* - \alpha = h^*(\cdot) \Delta \delta_x(\cdot) - \alpha = h^*(\cdot - x) - \alpha$ (where $\delta_x$ is the indicator function of the singleton $\{x\}$) and $h^*$ has the desired differentiability by Lemma 6. Finally, since $f(y) = k^*(y) = \psi(y)$ and $f \leq \psi$ on $X$, we have $f(y + v) + f(y - v) - 2f(y) \leq \psi(y + v) + \psi(y - v) - 2\psi(y) \leq C \| v \|^2$, for some constant $C$. Therefore the function $f$ is Fréchet differentiable at $y$ and $f'$ is Lipschitz at $y$. \[\square\]
Corollary 8. Let \( f \) be a Lipschitz convex function on a Banach space \( X \) and \( Y = \overline{\text{span}} \{ \partial f(x) : x \in X \} \). If \( Y \) is reflexive, then \( f \) can be majorised on \( X \) by a convex function \( \phi \) that is Fréchet differentiable and \( \phi(x) = f(x) \) for some \( x \in X \).

Under the assumptions in Theorem 7, the techniques in Theorem 1 may be applied to obtain approximation by functions with Lipschitz derivatives.

Theorem 9. Let \( X, Y \) and \( f \) be as in Theorem 7. Then \( f \) can be uniformly approximated on \( X \) by convex functions that have a Lipschitz derivative.

Proof: As in the proof of Theorem 7, let \( \| \cdot \| \) be an equivalent norm of \( X^* \) such that its restriction on \( Y \) is LUR. Let \( h = \| \cdot \|^2/2 \) and \( g := h + f^* \) on \( X^* \). The function \( g \) is \( \psi^* \)-lower semicontinuous on \( X^* \). Let \( k \) be a convex function on \( X \) such that \( k^* = g \). We claim that there exists a constant \( C \) such that for any \( \varepsilon > 0 \), \( x \in X \) and \( y \in \partial k(x) \), we have \( \{ v : g(v + y) - g(y) - (x,v) \leq C\varepsilon \} \subseteq \epsilon B_{X^*} \). Since \( g(u) = \infty \) whenever \( u \notin Y \), we only need to consider points in \( Y \). Let \( v \in Y \), then \( (g(y) + g(y + v))/2 - g((2y + v)/2) \geq (h(y) + h(y + v))/2 - h((2y + v)/2) \) for any \( y \in Y \). Using \((*)\), we have \( (g(y) + g(y + v))/2 - g((2y + v)/2) \geq L\| v \|^2 \) for any \( v \in X \) and for any \( y \in Y \). Following the same idea as in the proof of Theorem 7, we can complete the proof of the claim. By Lemma 6, \( k \) is Fréchet differentiable and \( k' \) is Lipschitz. For each \( n \in \mathbb{N} \) define \( g_n := f^* + h/(2n^4) \) and \( k_n \) such that \( k_n^* = g \). By the above argument, the function \( k_n \) is Fréchet differentiable and \( k_n' \) is Lipschitz for each \( n \in \mathbb{N} \). By [8, Lemma 2.1], \( \lim g_n = f \) uniformly on \( X \).

References


https://doi.org/10.1017/S0004972700014477 Published online by Cambridge University Press


Department of Mathematics  
University of Alberta  
Edmonton  
Canada T6G 2G1