A Survey of Examples of Convex Functions and Classifications of Normed Spaces

JONATHAN M. BORWEIN, SHRUM PROFESSOR OF SCIENCE *

December 1, 1994, January 22, 1995

Abstract

This paper represents a slightly extended version of the eponymous talk given at the VII Colloque Franco-Allemand d’Optimisation. My aim is to illustrate the tight connection between the sequential properties of a Banach space and the corresponding properties of the convex functions and sets which may or may not be defined on that space.

Keywords: convex functions, sequential properties, differentiability, nearest points, classical Banach spaces, Haar-null sets.


*Department of Mathematics and Statistics, Simon Fraser University, Burnaby, BC V5A 1S6, Canada; jborwein@ecm.sfu.ca (Research supported by NSERC and the Shrum Endowment at Simon Fraser University.)
1 Introduction.

This paper represents a slightly extended version of the eponymous talk given at the VII Colloque Franco-Allemand d’Optimisation. My aim is to illustrate the tight connection between the sequential properties of a Banach space and the corresponding properties of the convex functions and sets which may or may not be defined on that space. While for the most part we will give no full proofs we will often sketch the underlying ideas, relying on [2] and [5] for the full arguments. The central question we address is: what convexity properties characterize the most classical classes of Banach spaces? Thus what prevails in

1. Finite Dimensions?
2. Reflexive Spaces?
3. Separable Spaces?
4. Asplund spaces and spaces not containing $\ell_1(\mathbb{N})$?
5. Spaces containing $c_0(\mathbb{N})$?

Throughout $X$ is a real Banach space with continuous dual space $X^*$ and $f : X \to [-\infty, \infty]$ is lower semicontinuous convex and proper (somewhere finite). If $f$ is everywhere finite then $f$ is continuous — since a Banach space is barreled and so is a Baire space. This is one of the few significant analytic properties which hold in a large class of incomplete normed spaces. By contrast, in [5] we show how completeness is characterized by the non-emptiness or maximality of subgradients. We recall that the convex subgradient is defined by

$$\partial f(x) := \{x^* \in X^* : \langle x^*, h \rangle \leq f(x + h) - f(x), \forall h \in X\}.$$ 

In the sequel sets are usually closed and convex and $B$ denotes the open unit ball, $B := \{x : \|x\| < 1\}$. In general our notation and terminology are consistent with [17]. Moreover, if at any point the references or discussion herein seem incomplete, the trail will be found in [5] or [2].

We complete the preliminaries by recalling some derivative notions discussed in more detail in [2], [5] and [17]. Let $\beta$ denote a bornology of bounded sets. We write $$x^* \in \partial^\beta f(x) \Leftrightarrow \forall B \in \beta, \quad \forall \epsilon > 0, \quad \exists \delta > 0 \text{ with}$$

$$\langle x^*, h \rangle \leq \frac{f(x + th) - f(x)}{t} + \epsilon$$

for $0 < t < \delta$, $h \in B$.

It is conventional to identify the following bornologies:

- points $\Leftrightarrow$ Gâteaux ($G$)
(norm) compacts ⇔ Hadamard (H) 
weak compacts ⇔ weak Hadamard (WH) 
bounded ⇔ Fréchet (F)

Then $\partial^H f(x) = \partial^G f(x)$ for locally Lipschitz $f$, while $\partial^F f(x) = \partial^WH f(x)$ for $X$ reflexive. With this language we may define the $\beta$-derivative of $f$ at $x$ by

$$\{\nabla^\beta f(x)\} = \partial^\beta f(x) \cap -\partial^\beta (-f)(x)$$

so that

$$\{\nabla^\beta f(x)\} = \partial^\beta f(x) \quad \text{for concave } f.$$

# 2 Finite Dimensions.

We begin with a compendium of standard and relatively easy results whose proofs may be pieced together from many sources.

**Theorem 2.1** The following are equivalent:

(i) $X$ is finite dimensional.

(ii) Every linear functional on $X$ is continuous.

(iii) Every convex function $f : X \to \mathbb{R}$ is continuous.

(iv) The closed unit ball in $X$ is (pre-) compact.

(v) The weak and norm topologies coincide on $X$.

(vi) The weak-star and norm topologies coincide on $X^*$.

(vii) Every (closed) convex set in $X$ has non-empty relative interior.

(viii) $A \cap R = \emptyset$, $A$ closed and convex, $R$ a ray $\Rightarrow A$ and $R$ are separated by a continuous linear functional.

In essence this result says “don’t trust finite dimensionally derived intuitions”. By comparison, a much harder and less well known set of results is:

**Theorem 2.2** The following are equivalent:

(i) $X$ is finite dimensional.

(ii) Weak-star and norm convergence agree for sequences in $X^*$.

(iii) Every continuous convex $f : X \to \mathbb{R}$ is bounded on bounded sets.
(iv) For every continuous convex \( f : X \to \mathbb{R} \), \( \partial f \) is bounded on bounded sets.

(v) For every continuous convex \( f : X \to \mathbb{R} \), any point of Gâteaux differentiability is a point of Fréchet differentiability.

Proof Sketch. (i) \( \Rightarrow \) (iii) or (v) is clear; (iii) \( \Rightarrow \) (iv) is easy. To see (v) \( \Rightarrow \) (ii) and (iii) \( \Rightarrow \) (ii) we proceed as follows.

Let \( \|x_n^*\| = 1 \) and \( 0 < \alpha_n \downarrow 0 \). Define

\[
f(x) := \sup_{n \in \mathbb{N}} (x_n^*, x) - \alpha_n \tag{2-a}
\]

Then \( f \) is convex and continuous and is:

Gâteaux differentiable at \( 0 \Leftrightarrow x_n^* \rightharpoonup 0 \)

and

Fréchet differentiable at \( 0 \Leftrightarrow x_n^* \rightrightarrows 0 \).

Thus (v) \( \Rightarrow \) (ii). (See [2].) Now consider

\[
f(x) := \sum_n \varphi_n((x_n^*, x)) \tag{2-b}
\]

where \( \varphi_n(t) := n (|t| - \frac{1}{2})^+ \). Then \( f \) is

finite (continuous) \( \Leftrightarrow x_n^* \rightharpoonup 0 \)

and is

bounded on bounded sets \( \Leftrightarrow x_n^* \rightrightarrows 0 \).

Thus (iii) \( \Rightarrow \) (ii). (See [5].) \( \square \)

Note that the sequential coincidence of weak and norm topologies characterizes the Schur spaces (such as \( l_1(\mathbb{N}) \); see [12]) while the sequential coincidence of weak and weak–star topologies characterizes the Grothendieck spaces (reflexive spaces and non-reflexive spaces such as \( l_\infty(\mathbb{N}) \); see [12]).

The statements of Theorem 1.2 are equivalent in the strong sense that they are easily interderivable while no “easy proof” is known of (ii) \( \Rightarrow \) (i). This is the Josephson-Nissenzweig Theorem first established in 1975, see [12]. For example, (ii) \( \Rightarrow \) (iii) follows from:

Lemma 2.1 ([5]) Suppose that \( f : X \to \mathbb{R} \) is continuous and convex and that \( \{x_n\} \) is bounded while \( f(x_n) \to \infty \). Then

\[
x_n^* \in \partial f(x_n) \Rightarrow \psi_n := \frac{x_n^*}{\|x_n^*\|} \rightharpoonup 0.
\]

Thus each such function yields a Josephson-Nissenzweig sequence of unit vectors \( w^* \)-convergent to 0. \( \square \)
3 Reflexive spaces.

We begin with the traditional “James-Eberlein-Smulian” characterizations of reflexivity (see [10] or [15]):

**Theorem 3.1** The following are equivalent:

(i) $X$ is reflexive.

(ii) The unit ball on $X$ is weak compact.

(iii) Every continuous linear functional on $X$ achieves its norm.

(iv) If $\{C_n\}$ are non-empty nested, closed convex bounded sets, then $\cap_{n \in N} C_n \neq \emptyset$.

One may add the less traditional:

(v) Fenchel conjugacy is Mosco continuous for closed convex functions ([11]).

We will say that $f$ is coercive if $f(x)/\|x\| \to \infty$ when $\|x\| \to \infty$. A corresponding set of subgradient characterizations given in [5] is:

**Theorem 3.2** The following are equivalent:

(i) $X$ is reflexive.

(ii) $\text{Range}(\partial f) = X^*$ for some (or all) coercive continuous convex $f : X \to \mathbb{R}$.

(iii) $\text{Int} \ \text{Range}(\partial f)$ is convex for all (coercive) continuous convex $f : X \to \mathbb{R}$.

[Similar statements hold for maximal monotone operators.]

**Proof Sketch.** The “key” is the construction of

$$f(x) := \max \left\{ 1/2 \|x\|^2, \|x - p\| - 12 \pm \langle p^*, x \rangle \right\} \quad (3-a)$$

where $\|p\| = 5$ and $p^* \in \partial_{1/2} \|p\|^2$. Now $\text{Int} \ \text{Range}(\partial f)$ is non-convex. Indeed, using James Theorem one may show that it contains $B(X^*) \pm p^*$ and that $\frac{1}{2} B(X^*)$ lies in $\text{conv} \{\text{Int} \ \text{Range}(\partial f)\}$ but not in $\text{int} \ \text{Range}(\partial f)$.

Note that in any normed space $\text{int} \ \text{dom}(\partial f)$ is convex. The easiest explicit example for (iii) of the previous result lies in the space $c_0(\mathbb{N})$ of null sequences endowed with the supremum norm. One may use

$$f(x) := \|x - e_1\|_{\infty} + \|x + e_1\|_{\infty} \quad (3-b)$$
where $e_1$ is first unit vector. Then
\[ \text{Int Range}(\partial f) = \{ B(t_1) + e_1 \} \cup \{ B(t_1) - e_1 \} \]
which is far from convex. \( \square \)

We pause to indicate some relations with nearest points. Here as always we consider a distance function
\[ d_C(x) := \inf_{c \in C} \| c - x \| \]
and we say $C$ admits a nearest point if $d_C(x)$ is attained for some $x \notin C$. By James theorem $(X, \| \|)$ is reflexive iff every closed convex set in $X$ admits nearest points. A norm is (sequentially) **Kadec-Klee** if
\[ x_n \rightharpoonup x \text{ and } \| x_n \| \to \| x \| \Rightarrow \| x_n - x \| \to 0. \]

This is the most significant renorming property for results related to best approximation as is illustrated by:

**Theorem 3.3** ([Lau-Konyagin, 1976]) Every closed subset $C$ of $X$ densely (or generically) admits nearest points iff $(X, \| \|)$ is reflexive and has the Kadec-Klee property.

All reflexive spaces can be renormed to be Kadec-Klee. A fundamental open isometric question is:

“When $(X, \| \|)$ is reflexive must every non-trivial closed subset admit at least one nearest point?”

This question is open even for arbitrary renorms of Hilbert space! Counter-examples are necessarily unbounded, must fail to be weakly closed, and must lie in highly non Kadec-Klee spaces. (See [3] for details on all these matters relating to best approximations.)

Continuing to look at reflexivity, we consider a striking recent characterization in a slightly specialized form. It provides a remarkable liaison between norm compactness, openness, separability and reflexivity.

**Theorem 3.4** ([16]) Suppose $X$ is separable. The following are equivalent.

(i) $X$ is not reflexive.

(ii) $X$ contains a closed convex set $\tilde{C}$ with empty interior but such that every norm compact $K$ lies in some additive translate of $\tilde{C}$. 


Thus in any separable non-reflexive space there is a closed convex set \( C \) with empty interior which is not \textbf{Haar null}: meaning that no Borel probability measure can vanish on all translates of \( C \) (see [9]). This was motivated by a conjecture in [7] but leaves open the following tantalizing question:

“In a reflexive space is every closed convex set with empty interior Haar null?”

This is clearly the case in finite dimensions. As a consequence of Theorem 2.4 we also have a significant limiting example for Fréchet differentiability of Lipschitz functions.

\textbf{Corollary 3.1} ([7]) \textit{Let} \( X \) \textit{be separable}. Let

\[ d_C(x) := \inf_{c \in C} \| x - c \| \]  

(3-c)  

\textit{If} \( \text{int} C = \emptyset \) \textit{then} \( d_C \) \textit{fails to be Fréchet differentiable at all points of} \( C \). \textit{In particular, for} \( \hat{C} \) \textit{as in Theorem 2.4,} \( \hat{C} \) \textit{is not Haar-null and so} \( d_{\hat{C}} \) \textit{is not Fréchet Haar almost everywhere.}

We note that if \( X^* \) is not separable there is actually a nowhere Fréchet differentiable continuous convex function on \( X \). Since the Haar null sets are the largest class of reasonable null sets in a separable Banach space, Corollary 2.5 “rules out” studying Fréchet differentiability by measure-like techniques. By contrast, measure theoretic techniques work very well for studying Gateaux differentiability in a separable setting (see [7]).

\section{Separable spaces}

Many separable results continue to hold for spaces \( X \) with an (infinite dimensional) \textbf{separable quotient} \( Y \). That is, there exists a continuous linear surjection \( T: X \to Y \). There is no known instance of a space without a separable quotient. For example \( \ell_2(\mathbb{N}) \) is a separable quotient of \( \ell_\infty(\mathbb{N}) \).

\textbf{Theorem 4.1} ([5]) \textit{Suppose} \( X \) \textit{has a separable quotient}. \textit{Then there exist proper lower semi-continuous convex functions} \( f \) \textit{and} \( g \) \textit{with the properties that}

\begin{enumerate}
  \item \( \text{dom}(f) = \text{dom}(g) \) \textit{is dense in} \( X \),
  \item \( \partial f \) \textit{and} \( \partial g \) \textit{are both at most singleton (“almost Gateaux”)},
  \item \( \text{dom}(\partial f) \cap \text{dom}(\partial g) = \emptyset \).
\end{enumerate}
Proof sketch. From the existence of a separable quotient one argues that without loss \( X \) is separable. Let \( \{x_n, x_m^\dagger\} \) be an **M-basis**: meaning that \( \langle x_n, x_m^\dagger \rangle = \delta_{n,m} \) for \( n, m \in \mathbb{N} \), and \( \overline{\operatorname{span}} \{x_n\} = X \).

Then we may use

\[
f(x) := \sum_{n \in \mathbb{N}} (n \langle x^\dagger_n, x \rangle)^2 \quad \text{and} \quad g(x) := f(x - y)
\]

where \( y := \sum_{n \in \mathbb{N}} n^{-7/4} x_n \).

Before continuing we recall that the **quasi-relative interior** of \( C \) is given by

\[
\operatorname{qri}(C) := \{ x \in C : T_C(x) \text{ is linear} \}
\]

Here \( T_C(x) \) is the closed convex tangent cone generated by \( C \) at \( x \). Equivalently, \( x \in \operatorname{qri}(C) \) iff \( x \) is a non-support point of \( C \) in the sense that \( \varphi \in X^* \) and \( \langle \varphi, x \rangle = \operatorname{inf}_C \varphi \Rightarrow \langle \varphi, x \rangle = \operatorname{sup}_C \varphi \).

In finite dimensions it is easy to show that “\( \operatorname{qri} \)” = “rel-int”, while if \( X \) is separable every closed convex set \( C \) has non-empty quasi-relative interior; that is \( C \) has a non-support point. Indeed, let \( \{c_n : n \in \mathbb{N}\} \) be dense in \( C \) and consider

\[
\hat{c} := \sum_{n \in \mathbb{N}} \lambda_n c_n \quad \text{where the coefficients} \quad \sum_{n \in \mathbb{N}} \lambda_n = 1, \ \lambda_n > 0
\]

are chosen to ensure convergence of \( \hat{c} \). Then

\[
\langle \varphi, \hat{c} \rangle = \operatorname{inf}_C \varphi \Rightarrow \langle \varphi, \hat{c} \rangle = \langle \varphi, c_n \rangle = \operatorname{sup}_C \varphi.
\]

All of this is detailed in [6]. In short, the quasi-relative interior provides a useful surrogate for the relative interior which, by Theorem 1.1, must be empty for some closed convex set as soon as \( X \) is infinite dimensional. It is conjectured that “the converse holds” to Theorem 4.1 in the sense that in any non-separable space there is a closed convex set with empty quasi-relative interior. We detail some recent partial results in this direction:

**Theorem 4.2** ([3]) \( X \) contains a closed convex set consisting only of support points if either

1. (a) \( X = Y^* \) is non-separable or (b) \( X^* \) is not weak-star separable;

2. \( X \) contains an uncountable biorthogonal sequence;

or

3. \( X = C(\Gamma) \) where \( \Gamma \) is compact and Hausdorff and \( \Gamma \) contains a closed subset which is either non-separable or not a \( G_\delta \).
Proof Sketch. With sufficient work (1) follows from (2) and actually (1)(a) is considerably deeper than 1(b). (See [14].)

To see (2), let \( \{x_\alpha, x_\beta^*\} \) be biorthogonal for \( \alpha, \beta < \Omega \) (the first uncountable ordinal). Then

\[
C_\Omega := \overline{\text{conv}}\{x_\alpha : \alpha < \Omega\}
\]

is a support set: that is it contains only support points. Indeed,

\[
x_0 \in C_\Omega \Rightarrow x_0 \in \overline{\text{conv}}\{x_\alpha : \alpha < \beta\}
\]

for some \( \beta < \Omega \). Then \( x_\beta^* \) properly supports \( C_\Omega \) at \( x_0 \).

In (3) the harder case uses a closed non-\( G_\delta \) subset \( F \subset \Gamma \). Then

\[
C_F := \{f \in C(\Gamma) : f \geq 0, f|F = 0\}
\]

is a support set because

\[
f|F = 0 \Rightarrow f(t) = 0 \text{ for some } t \notin F.
\]

We may now use the Tietze extension theorem to build a function \( g \) in \( C_F \) with \( g(t) > 0 \). Then \( \delta_t \) supports \( C_F \) at \( t \).

\( \square \)

In the presence the Continuum Hypothesis it is shown in [8] that (2) and (3) are mutually distinct conditions. No other way of building support sets is known.

Thus, the continuous function spaces for which the converse remains open form a subclass of the non-metrizable \( \Gamma \) which are both hereditarily separable and hereditarily normal.

Another related open question now suggests itself:

“Does every infinite dimensional space contain a closed densely spanning convex set with at least one non-support point (a quasi-relative interior point) in its boundary?”

If \( Y \) is separable (and infinite dimensional) the answer is “yes”. Indeed, let \( \{y_n : n \in \mathbb{N}\} \) be dense in the unit sphere in \( Y \). The set

\[
C := \overline{\text{conv}}\{\pm 2^{-n}y_n : n \in \mathbb{N}\}
\]

is compact. Thus \( C \) has empty interior and so \( 0 \in bd(C) \); but also \( 0 \) is a non-support point of \( C \). As another illustration of the use of separable quotients, if \( X \) has a separable quotient \( Y \) with quotient map \( T \), then \( T^{-1}(C) \) “lifts” the example to \( X \).
5 Asplund spaces and spaces containing $\ell_1$

Recall that $X$ is an Asplund space if separable subspaces have separable duals as is the case for reflexive spaces. Equivalently, convex functions are generically Fréchet differentiable (see [11],[12],[17]). Recall also that Mackey convergence in $X^*$ is uniform convergence on weak compact convex subsets of $X$ and coincides with the norm topology on $X^*$ iff $X$ is reflexive.

Theorem 5.1 ([2],[5]) The following are equivalent:

(i) The space of absolutely summable sequences $\ell_1(\mathbb{N}) \subset X$ (isomorphically).

(ii) Mackey and norm convergence agree sequentially in $X^*$ (X is “sequentially reflexive”).

(iii) Every continuous convex $f : X \to \mathbb{R}$ which is bounded on weakly compact sets is bounded on bounded sets.

(iv) For every continuous convex $f : X \to \mathbb{R}$, any point of weak Hadamard differentiability is a point of Fréchet differentiability.

Proof sketch. The hard step (i) $\Leftrightarrow$ (ii) is a version of the wonderful “Rosenthal $\ell_1$ theorem” (1974) given in [12].

The remainder is analogous to our finite dimensional results. As before let

$$f(x) := \sup_{n \in \mathbb{N}} \langle x_n^*, x \rangle - \alpha_n$$

Then $f$ is convex and continuous and is:

- Gâteaux differentiable at $0 \iff x_n^* \rightharpoonup 0$
- weak Hadamard differentiable at $0 \iff x_n^* \overset{\text{weakly}}{\rightharpoonup} 0$
- Fréchet differentiable at $0 \iff x_n^* \rightharpoonup 0$.

Thus in any Asplund space, somewhat surprisingly, for convex functions one need only establish weak Hadamard differentiability rather than the ostensibly stronger Fréchet differentiability. In contrast in any non-reflexive space there is a non-convex distance function with a point of weak Hadamard differentiability that is not a point of Fréchet differentiability, [5].

Example. $C(\Gamma)$ is Asplund iff $\ell_1 \not\subset C(\Gamma)$ but generally the Asplund class is much smaller. Correspondingly $L^1(\mu)$, $\mu$ $\sigma$-finite, admits a weak Hadamard smooth renorm (see [4]). This is useful in applications to control or optimization problems since $L^1(\mu)$ is not Asplund but convex functions are, nonetheless, generically $W^H$ differentiable. Moreover, any separable space with a non-separable dual with a weak Hadamard renorm, must by Theorem 5.1, contain a copy of $\ell_1(\mathbb{N})$. 
Also $X$ is Asplund iff $X^*$ has the **Radon–Nikodym property** (RNP): every norm closed bounded convex set in $X^*$ has a strongly exposed point (equivalently in a dual space an extreme point). Reflexive spaces have the RNP as do separable dual spaces such as $\ell_1$. $\square$

**Theorem 5.2** ([5]) The following are equivalent:

(i) $X$ has the Radon–Nikodym property.

(ii) The range of the subgradient, $\text{Range}(\partial f)$, is a generic set in $X^*$ for each coercive lower semicontinuous convex function $f : X \to ]-\infty, \infty]$.

### 6 Spaces containing $c_0$

The sequence space $c_0$ is the prototype of an Asplund space without the Radon–Nikodym property and is the home to many useful examples. For instance:

**Theorem 6.1** Let $f$ and $g$ be lower semicontinuous convex coercive proper convex functions.

(1) Suppose $X$ has the Radon–Nikodym property. Then the infimal convolution

$$f \Box g(x) := \inf_y \{ f(y) - g(x - y) \}$$

is attained for some $x$.

(2) This fails if $X$ contains a copy of $c_0$.

**Proof.** (1) is not published but is fairly simple to establish while (2) is due to Edelstein and Thompson ([13]). $\square$

Indeed $c_0$ contains two closed norm balls, $\overline{B}_1$ and $\overline{B}_2$, such that $\overline{B}_1 + \overline{B}_2$ is open. Equivalently $\|\cdot\|$ admits no nearest points in $\overline{B}_2$ and conversely. Such pairs are called anti-proximinal (see [13]). It is easy to show that anti-proximinal pairs can not be found in a space with the RNP, or more generally in a space with the slightly less arduous convex point of continuity property. The prototype of a space without the point of continuity property but which fails to admit copies of $c_0$ is the space of Lebesgue integrable functions $L_1[0,1]$. Thus we finish with yet another open question:

"Does an anti-proximinal pair exist in $L_1"$?

It is easy to show that neither of the sought balls can be the original norm ball.
References


