

On the Continuity and Regularity of Convex Extensions

Orest Bucicovschi*

*Department of Mathematics, University of California at San Diego,
La Jolla, CA 92093-0112, USA
obucicov@math.ucsd.edu*

Jiří Lebl†

*Department of Mathematics, University of Wisconsin-Madison,
Madison, WI 53706, USA
lebl@math.wisc.edu*

Received: October 21, 2011

Revised manuscript received: February 14, 2013

We study continuity and regularity of convex extensions of functions from a compact set C to its convex hull $K = \text{co}(C)$. We show that if C contains the relative boundary of K , and f is a continuous convex function on C , then f extends to a continuous convex function on K using the standard convex roof construction. In fact, a necessary and sufficient condition for f to extend from any set to a continuous convex function on the convex hull is that f extends to a continuous convex function on the relative boundary of the convex hull. We give examples showing that the hypotheses in the results are necessary. In particular, if C does not contain the entire relative boundary of K , then there may not exist any continuous convex extension of f . Finally, when the boundary of K and f are C^1 we give a necessary and sufficient condition for the convex roof construction to be C^1 on all of K . We also discuss an application of the convex roof construction in quantum computation.

1. Introduction

Extending convex functions is a problem with applications from economics [4] to quantum computing [6, 7, 8] (see also § 6) and in a wide variety of other optimization problems. A classical theorem of Gale, Klee, and Rockafellar [1] is that given a convex domain, then every bounded convex function defined on the interior extends continuously to the boundary if and only if the domain is a polyhedron. In this paper we look at the opposite problem of extending a function from the boundary to the interior. For the set of extreme points this was also studied by Lima [2].

Given a convex function defined on a set, there exists a common construction called the *convex roof* that defines the largest convex extension to the convex hull of the original set. In this paper we study the continuity and regularity properties of this construction. Background in convex analysis is taken from the book [5].

*The first author was supported in part by DARPA QuEST grant N66001-09-1-2025.

†The second author was supported in part by NSF grant DMS 0900885.

For a convex set $K \subset \mathbb{R}^d$, let $\text{ri}(K)$ denote the *relative interior* of K . That is, the topological interior of K inside the affine hull of K . Then let $\text{bd}(K) = \overline{K} \setminus \text{ri}(K)$ denote the *relative boundary*. If $f: C \rightarrow \mathbb{R}$ is a function, where $C \subset \mathbb{R}^d$ is not necessarily convex, we say that f is *convex* if

$$f(x) \leq \sum_{j=1}^k t_j f(x_j), \quad (1)$$

for any finite collection $x, x_1, \dots, x_k \in C$ and $t_1, \dots, t_k \in [0, 1]$, such that $\sum_j t_j = 1$ and $\sum_j t_j x_j = x$.

Theorem 1.1. *Let $K \subset \mathbb{R}^d$ be a compact convex set and $C \subset K$ be a closed subset such that $\text{bd}(K) \subset C$. If $f: C \rightarrow \mathbb{R}$ is continuous and convex then f extends using the convex roof construction to a continuous convex function on K .*

Examples show that the theorem is optimal in the sense that we cannot simply take any other smaller natural subset of the boundary (such as the extreme points) and always expect the convex roof construction to be continuous. Lima [2] proves that a continuous function defined on a closed subset of the set of extreme points extends to a continuous convex function of the convex hull. In particular, every continuous function on the set of extreme points extends to a continuous convex function on the convex hull if and only if the set of extreme points is closed. In this paper we are more interested in convex extensions from sets that are larger than the extreme points. We are in particular interested in the convex roof construction and its regularity.

The theorem gives us a necessary and sufficient condition for a continuous function defined on a closed compact set C to extend continuously to the convex hull.

Corollary 1.2. *Let $C \subset \mathbb{R}^d$ be a compact set and let $f: C \rightarrow \mathbb{R}$ be a continuous convex function. Then f extends to a continuous convex function on the convex hull $K = \text{co}(C)$ if and only if f extends to a continuous convex function on $\text{bd}(K)$.*

Example 5.4 shows that it is not always possible to extend a convex function continuously to a convex function of the boundary.

Let us give a rather interesting corollary of the main theorem. We will also give a simple independent proof of this result. A set K is *strictly convex* if the relative boundary $\text{bd}(K)$ does not contain any intervals. Notice that any function defined on the boundary of a strictly convex set is automatically convex.

Corollary 1.3. *Let $K \subset \mathbb{R}^d$ be a compact strictly convex set. If $f: \text{bd}(K) \rightarrow \mathbb{R}$ is continuous, then f extends using the convex roof construction to a continuous convex function on K .*

It is natural to ask about regularity. Example 5.5 shows a C^∞ function f on a strictly convex $\text{bd}(K)$ such that no convex extension of f is Lipschitz on K ; the derivative must blow up when approaching the boundary.

However, the derivative blowing up at a boundary point is the worst that can happen. We say that a convex function $f: \text{bd}(K) \rightarrow \mathbb{R}$ has a *nonvertical supporting hyperplane* at $p \in \text{bd}(K)$ if there exists an affine function $A: \mathbb{R}^d \rightarrow \mathbb{R}$ such that $A(x) \leq f(x)$ for all $x \in \text{bd}(K)$ and such that $A(p) = f(p)$. This is equivalent to $A(x) = L(x) + f(p)$, where L is a subgradient of f at p , that is, L is an element of $\partial f(p)$. If the convex roof $\hat{f}$ is differentiable at $p \in \text{bd}(K)$, then f has a nonvertical supporting hyperplane at p (the tangent hyperplane to the graph of $\hat{f}$ at p). It turns out this is also a sufficient condition for $\hat{f}$ to be continuously differentiable on K .

Theorem 1.4. *Let $K \subset \mathbb{R}^d$ be a compact convex set, $E \subset \text{bd}(K)$ a closed subset such that $\text{bd}(K)$ is a C^1 manifold near every point of $\text{bd}(K) \setminus E$. Suppose that $f: \text{bd}(K) \rightarrow \mathbb{R}$ is a convex function bounded from below that is C^1 on $\text{bd}(K) \setminus E$. Then*

- (i) $\hat{f}: K \rightarrow \mathbb{R}$ is C^1 on $\text{ri}(K) \setminus \text{co}(E)$.
- (ii) If f has a nonvertical supporting hyperplane at $p \in \text{bd}(K) \setminus \text{co}(E)$, then $\hat{f}$ is differentiable at p . In fact there exists a convex neighborhood U of p and a convex function $g: U \rightarrow \mathbb{R}$, $g|_{U \cap K} = \hat{f}|_{U \cap K}$, and $g|_{U \cap K}$ is C^1 .

By $\text{bd}(K)$ being a C^1 manifold near some $p \in \text{bd}(K)$ we mean that there exists a neighbourhood U of p such that $\text{bd}(K) \cap U$ is an embedded C^1 submanifold of U .

It is possible to construct examples showing that $\hat{f}$ need not be differentiable at points of $\text{co}(E)$. The construction of the g in the proof together with the compactness of K yields the following corollary.

Corollary 1.5. *Let $K \subset \mathbb{R}^d$ be a compact convex set and suppose that $\text{bd}(K)$ is a C^1 manifold. Suppose that $f: \text{bd}(K) \rightarrow \mathbb{R}$ is a C^1 convex function that has a nonvertical supporting hyperplane at every point of the boundary. Then there exists a proper convex function $g: \mathbb{R}^d \rightarrow \mathbb{R}$ such that $g|_K = \hat{f}$ and $\hat{f}: K \rightarrow \mathbb{R}$ is a C^1 function on all of K .*

Surprisingly, Example 5.6 shows that this theorem does not extend to C^2 or higher regularity. We also remark that the g we construct in the proof is the smallest convex extension to all of $\mathbb{R}^d \setminus K$, and this g need not be differentiable everywhere.

2. Convex roof extension

Let I be nonempty set, not necessarily finite. Let $\mathbb{R}^{(I)}$ be the real vector space of I -tuples of real numbers indexed by I , with only finitely many nonzero components:

$$\mathbb{R}^{(I)} = \{t = \{t_i\}_{i \in I} \in \mathbb{R}^I : \text{supp}(t) \text{ is finite}\}, \quad (2)$$

where $\text{supp}(t)$ denotes the set $\{i \in I : t_i \neq 0\}$. Denote by $|\text{supp}(t)|$ the cardinality of $\text{supp}(t)$.

Let $\{x_i\}_{i \in I}$ a family of elements of the vector space $\mathbb{R}^d$ and $x \in \text{co}(\{x_i\})$. Consider the (nonempty by hypotheses) set $\mathcal{C}(x)$ of all writings of x as a convex combination

of $\{x_i\}$:

$$\mathcal{C}(x) := \left\{ \{t_i\} \in [0, 1]^{(I)} : \sum t_i = 1 \text{ and } x = \sum t_i x_i \right\}. \quad (3)$$

Of course, $\mathcal{C}(x)$ depends on x , I , and $\{x_i\}_{i \in I}$

Definition 2.1. Let $C \subset \mathbb{R}^d$, and let $f: C \rightarrow \mathbb{R}$. Let $I = C$, and $\{x_i\}_{i \in I} = C$ where $x_i = i$. We define the *convex roof* of f to be the function $\widehat{f}: \text{co}(C) \rightarrow \mathbb{R} \cup \{-\infty\}$ given by

$$\widehat{f}(x) := \inf_{t \in \mathcal{C}(x)} \sum_{i \in I} t_i f(x_i). \quad (4)$$

To make the above definition workable we need the classical lemma of Carathéodory (see e.g. [5]). Note that $\mathcal{C}(x)$ is a convex subset of the real vector space $\mathbb{R}^{(I)}$.

Lemma 2.2 (Carathéodory). *Every element of $\mathcal{C}(x)$ is a convex combination of elements of $\mathcal{C}(x)$ with supports of cardinality at most $d + 1$.*

A sketch of the proof: For every $r \geq 1$, consider the set

$$\mathcal{C}_r(x) := \left\{ \{t_i\} \in [0, 1]^{(I)} : |\text{supp}(t_i)| \leq r, \sum t_i = 1 \text{ and } x = \sum t_i x_i \right\}. \quad (5)$$

One shows using the standard method (see e.g. [5]) that for every $r \geq d + 1$, any element in $\mathcal{C}_{r+1}(x)$ is an average of two elements in $\mathcal{C}_r(x)$.

In particular, the set $\mathcal{C}_{d+1}(x)$ is nonempty. An easy computation obtains the following standard corollary.

Corollary 2.3. *Let $C \subset \mathbb{R}^d$, and $f: C \rightarrow \mathbb{R}$ be bounded. As before, let $I = C$, and $\{x_i\}_{i \in I} = C$ where $x_i = i$. Then*

$$\widehat{f}(x) = \inf_{t \in \mathcal{C}_{d+1}(x)} \sum_{i \in I} t_i f(x_i). \quad (6)$$

The following facts are standard and not difficult to prove.

Proposition 2.4. *Let $C \subset \mathbb{R}^d$ be a compact set and let $K = \text{co}(C)$. If $f: C \rightarrow \mathbb{R}$ is bounded from below and convex, then*

- (i) $\widehat{f}$ is bounded from below (in particular real valued).
- (ii) if f is also bounded from above, then $\widehat{f}$ is bounded.
- (iii) $f = \widehat{f}|_C$.
- (iv) $\widehat{f}$ is convex.
- (v) If $g: K \rightarrow \mathbb{R}$ is convex and $g|_C \leq f$, then $g \leq \widehat{f}$.

Lemma 2.5. *Let $C \subset \mathbb{R}^d$ be a compact set, and $K = \text{co}(C)$. If $f: C \rightarrow \mathbb{R}$ is bounded, lower semicontinuous, and convex, then $\widehat{f}$ is lower semicontinuous.*

The lemma follows easily from the results in [5], however, the idea in the following proof will be useful and so we include it.

Proof. Take a sequence $\{x_i\}$ in K that converges to $x \in K$. Using Carathéodory's lemma, for each x_i , find $d+1$ points $x_i^1, \dots, x_i^{d+1} \in C$ and $t_i^1, \dots, t_i^{d+1} \in [0, 1]$ such that $\sum_j t_i^j = 1$ and such that $x_i = \sum_j t_i^j x_i^j$. By compactness of K and the corollary to the Carathéodory's lemma we assume that

$$\widehat{f}(x_i) = \sum_{j=1}^{d+1} t_i^j f(x_i^j). \quad (7)$$

For any subsequence of $\{x_i\}$ we can take a further subsequence where x_i^j and t_i^j converge to x^j and t^j respectively (as K is compact).

$$\liminf_{i \rightarrow \infty} \widehat{f}(x_i) = \liminf_{i \rightarrow \infty} \sum_{j=1}^{d+1} t_i^j f(x_i^j) \geq \sum_{j=1}^{d+1} t^j f(x^j) \geq \widehat{f}(x). \quad (8)$$

□

3. Continuity of the extension

A point $p \in K$ (K convex) is said to be *extreme* if $p = tx + (1-t)y$, $t \in [0, 1]$, $x, y \in K$ implies that either $p = x$ or $p = y$. We always get continuity of the extension at the extreme points, even without requiring that f be defined on the entire relative boundary.

Lemma 3.1. *Let $C \subset \mathbb{R}^d$ be a compact convex set and let $K = \text{co}(C)$. If $f: C \rightarrow \mathbb{R}$ is continuous and convex and p is an extreme point of K , then $\widehat{f}$ is continuous at p .*

Proof. As in the proof of Lemma 2.5 we take a sequence $\{x_i\}$ in K that converges to p . Using Carathéodory's lemma and its corollary, the compactness of K , and taking subsequence of a subsequence we assume that $x_i = \sum_{j=1}^{d+1} t_i^j x_i^j$, where

$$\widehat{f}(x_i) = \sum_{j=1}^{d+1} t_i^j f(x_i^j). \quad (9)$$

Also assume that x_i^j converges to x^j and t_i^j converges to t^j . Thus,

$$\lim_{i \rightarrow \infty} \widehat{f}(x_i) = \lim_{i \rightarrow \infty} \sum_{j=1}^{d+1} t_i^j f(x_i^j) = \sum_{j=1}^{d+1} t^j f(x^j). \quad (10)$$

We also obtain that $p = \sum t^j x^j$. As p is an extreme point this means that $x^j = p$ (or $t^j = 0$) for all j , therefore $\lim \widehat{f}(x_i) = \sum t^j f(x^j) = f(p) = \widehat{f}(p)$. □

To obtain continuity of the extension at nonextreme points we require the following lemma.

Lemma 3.2. *Let $K \subset \mathbb{R}^d$ be a compact convex set and let $f: K \rightarrow \mathbb{R}$ be a convex function lower semicontinuous at $p \in \text{bd}(K)$, such that $f|_{\text{bd}(K)}$ is continuous at p . Then f is continuous at p .*

Proof. Suppose for contradiction that f is not continuous at $p \in \text{bd}(K)$. As f is lower semicontinuous at p , there must exist a $\delta > 0$ and a sequence $x_j \in K$ converging to p such that $f(x_j) \geq f(p) + \delta$ for all j . We pick a fixed point $y \in \text{ri}(K)$. Let $y_j \in \text{bd}(K)$ be the unique point in $\text{bd}(K)$ on the line through y and x_j such that x_j lies on the line segment $[y_j, y]$. It is clear that $\lim y_j = p$. As $f|_{\text{bd}(K)}$ is continuous at p , then for large enough j we have $f(y_j) \leq f(p) + \frac{\delta}{2}$.

As f is convex, then for $t \geq 1$ we have

$$(1-t)f(y_j) + tf(x_j) \leq f((1-t)y_j + tx_j), \quad (11)$$

whenever $(1-t)y_j + tx_j \in K$. Let $t_j > 1$ be such that $y = (1-t_j)y_j + t_jx_j$. We have $t_j \rightarrow \infty$ as $\lim x_j = \lim y_j = p$. For all j we have

$$(1-t_j)f(y_j) + t_jf(x_j) \leq f(y). \quad (12)$$

For large enough j we have $f(y_j) \leq f(p) + \frac{\delta}{2}$ and $f(x_j) \geq f(p) + \delta$. Hence

$$f(p) + \frac{\delta}{2} + t_j \frac{\delta}{2} = (1-t_j) \left(f(p) + \frac{\delta}{2} \right) + t_j (f(p) + \delta) \leq f(y). \quad (13)$$

Now $t_j \rightarrow \infty$ obtains a contradiction. $\square$

Although we will not need it, let us remark that the lemma is really local. That is, compactness of K is not necessary. We can apply it to any closed convex set K by taking a closed ball B centered at p and applying the theorem to $B \cap K$.

We now have all the tools to prove the main theorem.

Proof of Theorem 1.1. We know that $\widehat{f}(x) = f(x)$ for $x \in C$ by Proposition 2.4 and by Lemma 2.5, $\widehat{f}$ is lower semicontinuous. As $\widehat{f}$ is convex, it is standard that $\widehat{f}$ is continuous on $\text{ri}(K)$. Then we apply Lemma 3.2. $\square$

4. Regularity of the convex roof

Let $f: K \rightarrow \mathbb{R}$ be a convex function. Define a *subgradient* at $p \in K$ to be a linear mapping L such that $L(x-p) + f(p) \leq f(x)$ for all $x \in K$. We need some classical results about derivatives of convex functions.

Theorem 4.1 (See e.g. Theorems 25.1 and 25.5 in [5]). *Let $U \subset \mathbb{R}^d$ be an open convex set and let $f: U \rightarrow \mathbb{R}$ be a convex function.*

- (i) *f has a unique subgradient at p if and only if f is differentiable at p .*
- (ii) *f is differentiable on a dense subset $D \subset U$.*
- (iii) *The mapping $x \mapsto \nabla f(x)$ is continuous on D .*

Before we prove Theorem 1.4, let us note the following observation (see also [7], although Uhlmann only considers continuous f).

Proposition 4.2. *Let $K \subset \mathbb{R}^d$ be a compact convex set, $f: \text{bd}(K) \rightarrow \mathbb{R}$ a convex function bounded from below, and $p \in \text{ri}(K)$. There exists a closed convex set $W \subset K$, with $p \in W$ and $W = \text{co}(W \cap \text{bd}(K))$ with the following property. If A is any affine function with $A(p) = \widehat{f}(p)$ and $A(x) \leq \widehat{f}(x)$ for all $x \in K$, then*

$$W \subset \{x \in K : A(x) = \widehat{f}(x)\}. \quad (14)$$

In particular, the proposition says that there exists a line $\ell \subset \mathbb{R}^d$ through p such that for any subgradient L of $\widehat{f}$ at p , we get that

$$L(x - p) + \widehat{f}(p) = \widehat{f}(x) \quad (15)$$

for all $x \in \ell \cap K$.

Proof. Let $p \in \text{ri}(K)$. By definition of convex roof and the lemma of Carathéodory we can find sequences $t_i^1, t_i^2, \dots, t_i^{d+1} \in [0, 1]$, and $x_i^1, x_i^2, \dots, x_i^{d+1} \in \text{bd}(K)$ with $\sum_j t_i^j = 1$ and $p = \sum_j t_i^j x_i^j$, and such that

$$\widehat{f}(p) = \lim_{i \rightarrow \infty} \sum_{j=1}^{d+1} t_i^j f(x_i^j). \quad (16)$$

By compactness of K we can assume that there exist $t^1, \dots, t^{d+1} \in [0, 1]$ and $x^1, \dots, x^{d+1} \in \text{bd}(K)$ such that t_i^j and x_i^j converge to t^j and x^j respectively as i goes to infinity. Furthermore, the sequence $\{f(x_i^j)\}_{i=1}^\infty$ must be bounded above and as f is also bounded from below and K is compact, we can assume that the limit of $f(x_i^j)$ exists.

Write

$$w_j = \lim_{i \rightarrow \infty} f(x_i^j). \quad (17)$$

We see that

$$\widehat{f}(p) = \sum_{j=1}^{d+1} t^j w_j. \quad (18)$$

Let $C = \text{co}(\{x^1, x^2, \dots, x^{d+1}\})$. If A is an affine function such that $A(x) \leq \widehat{f}(x)$ and $A(p) = \widehat{f}(p)$, then $A(x^j) \leq w_j$. We can now without loss of generality assume that $t^j \neq 0$, and therefore $p \in \text{ri}(C)$. As $A(p) = \widehat{f}(p)$ we see that if $x = \sum_j s^j x^j$ (with $x \in \text{ri}(C)$) where $\sum s^j = 1$ then

$$A(x) = \sum_{j=1}^{d+1} s^j w_j. \quad (19)$$

For any $\epsilon > 0$ then for large enough i we have

$$\widehat{f}\left(\sum_j s^j x_i^j\right) \leq \sum_{j=1}^{d+1} s^j f(x_i^j) \leq \widehat{f}\left(\sum_j s^j x^j\right) + \epsilon. \quad (20)$$

As $\widehat{f}$ is continuous in $\text{ri}(K)$, then

$$\widehat{f}(x) = A(x) \quad (21)$$

for all $x \in \text{ri}(K)$, and we are done. $\square$

We can now prove the first part of Theorem 1.4. For this statement we can safely drop the continuity hypothesis for the derivative. The C^1 hypothesis is necessary to extend differentiability to the boundary. On the other hand we automatically obtain C^1 differentiability in the interior by Theorem 4.1.

Lemma 4.3. *Let $K \subset \mathbb{R}^d$ be a compact convex set, $E \subset \text{bd}(K)$ a closed set such that $\text{bd}(K)$ is a differentiable manifold near every point of $\text{bd}(K) \setminus E$. Suppose that $f: \text{bd}(K) \rightarrow \mathbb{R}$ is a convex function bounded from below that is differentiable on $\text{bd}(K) \setminus E$. Then $\widehat{f}: K \rightarrow \mathbb{R}$ is C^1 on $\text{ri}(K) \setminus \text{co}(E)$.*

Proof. Let us assume that we work in the affine hull of K and therefore without loss of generality assume that the relative interior $\text{ri}(K)$ is in fact equal to the topological interior.

Let $p \in \text{ri}(K) \setminus \text{co}(E)$. Suppose that L and $\widetilde{L}$ are subgradients of $\widehat{f}$ at p . By Proposition 4.2 there is a line $\ell \subset \mathbb{R}^d$ (with $p \in \ell$) such that

$$\widehat{f}(x) = L(x - p) + \widehat{f}(p) = \widetilde{L}(x - p) + \widehat{f}(p) \quad (22)$$

for all $x \in \ell$. We can also assume that there exists a $q \in \ell \cap (\text{bd}(K) \setminus E)$ as p was not in the convex hull of E . We notice that ℓ is not tangent to $\text{bd}(K)$ at q as K is convex; if ℓ were tangent then $\ell \cap K \subset \text{bd}(K)$, which contradicts the fact that $p \in \text{ri}(K)$.

Let T denote the $(d-1)$ -dimensional affine manifold in $\mathbb{R}^d \times \mathbb{R}$ through the point $(p, f(p))$, tangent to the graph of f

$$\Gamma_f = \{(x, f(x)) \in \mathbb{R}^d \times \mathbb{R} : x \in \text{bd}(K)\}. \quad (23)$$

Note that near $(p, f(p))$, Γ_f is a C^1 manifold. And define the line $\widehat{\ell} \in \mathbb{R}^d \times \mathbb{R}$ as

$$\widehat{\ell} = \{(x, L(x - p) + \widehat{f}(p)) = (x, \widehat{f}(x)) \in \mathbb{R}^d \times \mathbb{R} : x \in \ell\}. \quad (24)$$

Let W be the affine span of T and $\widehat{\ell}$ (that is, the smallest affine space containing both T and $\widehat{\ell}$). As ℓ was not tangent to $\text{bd}(K)$, then $\widehat{\ell}$ is not contained in T and therefore W is d -dimensional.

Let Γ_L be the graph of the mapping $x \mapsto L(x - p) + \widehat{f}(p)$, that is

$$\Gamma_L = \{(x, y) \in \mathbb{R}^d \times \mathbb{R} : y = L(x - p) + \widehat{f}(p)\}. \quad (25)$$

Similarly let $\Gamma_{\widetilde{L}}$ be the graph of the mapping $x \mapsto \widetilde{L}(x - p) + \widehat{f}(p)$. As for $x \in \text{bd}(K)$ we have $\widetilde{L}(x - p) + \widehat{f}(p) \leq f(x)$ and $L(x - p) + \widehat{f}(p) \leq f(x)$ we see that T is a subset of $\Gamma_L \cap \Gamma_{\widetilde{L}}$. Furthermore, $\widehat{\ell} \subset \Gamma_L \cap \Gamma_{\widetilde{L}}$, by definition of ℓ . Therefore, $W \subset \Gamma_L \cap \Gamma_{\widetilde{L}}$.

As Γ_L , $\Gamma_{\tilde{L}}$, and W are all d -dimensional affine subspaces, we see that

$$\Gamma_L = \Gamma_{\tilde{L}} = W. \quad (26)$$

So $L = \tilde{L}$ and $\hat{f}$ is differentiable at p by Theorem 4.1. $\square$

Let us prove the second part of the theorem.

Lemma 4.4. *Let $K \subset \mathbb{R}^d$ be a compact convex set, $E \subset \text{bd}(K)$ a closed set such that $\text{bd}(K)$ is a C^1 manifold near every point of $\text{bd}(K) \setminus E$. Suppose that $f: \text{bd}(K) \rightarrow \mathbb{R}$ is a convex function bounded from below that is C^1 on $\text{bd}(K) \setminus E$ and such that f has a nonvertical supporting hyperplane at $p \in \text{bd}(K) \setminus \text{co}(E)$.*

Then $\hat{f}$ is differentiable at p . In fact there exists a convex neighborhood U of p and a convex function $g: U \rightarrow \mathbb{R}$, such that $g|_{U \cap K} = \hat{f}|_{U \cap K}$, and $g|_{U \cap K}$ is C^1 .

Proof. Again without loss of generality assume that $\text{ri}(K)$ is an open set in $\mathbb{R}^d$. We pick a point $p \in \text{bd}(K) \setminus \text{co}(E)$. Near p , the graph of f above $\text{bd}(K)$ is a C^1 manifold. As f has a nonvertical supporting hyperplane at p , then it has such a hyperplane at all points near p ; for example, we can simply take tangent hyperplanes to some C^1 manifold of dimension d in $\mathbb{R}^d \times \mathbb{R}$ that contains the graph of f and is tangent to the supporting hyperplane at p .

Take a small convex neighborhood U of p , such that $U \cap \text{bd}(K)$ is a connected C^1 manifold and such that f has a nonvertical supporting hyperplane at each $q \in U \cap \text{bd}(K)$. At each $q \in U \cap \text{bd}(K)$ we take the unique supporting hyperplane given by an affine function $L_q: \mathbb{R}^d \rightarrow \mathbb{R}$ such that L_q minimizes the derivative in the normal direction to $\text{bd}(K)$. The magnitude of the gradient of L_q must be uniformly bounded for $q \in U \cap \text{bd}(K)$ by the same argument that guaranteed nonvertical supporting hyperplanes. Define

$$g(x) := \begin{cases} \sup\{L_q(x) : q \in U \cap \text{bd}(K)\} & \text{if } x \in U \setminus K, \\ \hat{f}(x) & \text{if } x \in U \cap K. \end{cases} \quad (27)$$

It is not hard to check that g must be a convex function. Furthermore, g has a unique subgradient at p (and in fact at all $q \in U \cap \text{bd}(K)$). To see this fact, note that any other possible subgradient must be tangent to the graph of f over $\text{bd}(K)$ and hence would give a candidate for L_p , which is unique. $\square$

We now prove Corollary 1.5. We look at how the function g was constructed above. We assume that the boundary $\text{bd}(K)$ is C^1 and the function f is C^1 on all of $\text{bd}(K)$. As $\text{bd}(K)$ is compact, the magnitude of the gradient of L_q must be uniformly bounded for all $q \in \text{bd}(K)$. As g was constructed by taking a supremum over L_q , this supremum must be bounded on all of $\mathbb{R}^d$.

The function g constructed above is the smallest possible extension to $\mathbb{R}^d \setminus K$, because every convex extension must lie above the subgradients. To see that g need not be differentiable take the function whose graph is a union of lines all going through a single point $p \notin K$. We can arrange such a function to be convex and C^1 on K . It is then obvious that the function is equal to the g above and we can arrange it to not be differentiable at p .

5. Examples

Example 5.1 (Punctured tomato can). Defining f on only a subset of the boundary (including the extremal set) is not enough to guarantee continuity of the convex roof. Let $(x, y, z) \in \mathbb{R}^3$ be our coordinates. Let

$$\begin{aligned} K_1 &= \{(x, y, z) \mid x = -1, y^2 + z^2 = 1\}, \\ K_2 &= \{(x, y, z) \mid x = 1, y^2 + z^2 = 1\}, \\ K_3 &= \{(0, 0, 1)\}. \end{aligned} \tag{28}$$

The convex hull of the union is the cylinder

$$\text{co}(K_1 \cup K_2 \cup K_3) = \{(x, y, z) \mid -1 \leq x \leq 1, y^2 + z^2 \leq 1\}. \tag{29}$$

Define f to be identically 1 on K_1 and K_2 and let f be zero on K_3 . Then f is continuous and convex on $K_1 \cup K_2 \cup K_3$. Lemma 2.5 tells us that $\widehat{f}$ is lower semicontinuous.

The convex roof $\widehat{f}$ must be identically 1 on the lines $\{(x, y, z) \mid -1 \leq x \leq 1, y = y_0, z = z_0\}$ where $y_0^2 + z_0^2 = 1$ are fixed and $y_0 \neq 0$. On the other hand, at the point $(0, 0, 1)$, which lies on the line $\{(x, y, z) \mid -1 \leq x \leq 1, y = 0, z = 1\}$, the function $\widehat{f}$ must be 0. Hence $\widehat{f}$ cannot be upper semicontinuous at $(0, 0, 1)$.

In this case, note that a continuous extension does exist. For example the function x^2 . We have only shown that the convex roof construction is not continuous.

Example 5.2. As the set of extreme points may not be closed, when we define a function only on the set of extreme points, it is natural to require that the function be uniformly continuous such that it extends to the closure of the extreme points. In this case however the extension to the closure can fail to be convex. For example take again $(x, y, z) \in \mathbb{R}^3$ be our coordinates and let

$$\begin{aligned} K_1 &= \{(-1, 0, 0)\}, \\ K_2 &= \{(x, y, z) \mid x = 0, (y - 1)^2 + z^2 = 1, y \neq 0\}, \\ K_3 &= \{(1, 0, 0)\}. \end{aligned} \tag{30}$$

Define the function to be 0 at K_1 and K_3 , and let it be identically 1 on K_2 . The function is convex on $K_1 \cup K_2 \cup K_3$ and uniformly continuous, however the continuous extension to $K_1 \cup K_2 \cup K_3$ fails to be convex because the function will be 1 at $(1, 0, 0)$, while it will be 0 at $(-1, 0, 0)$ and $(1, 0, 0)$.

A natural question to ask is what happens then when the function is defined, continuous, and convex on the closure of the extreme points. Lima [2] shows that then any continuous function on the set of extreme points extends to a continuous convex function on the convex hull if and only if the set of extreme points is closed. Let us therefore see an example where the set of extreme points is not closed. Let us combine examples 5.2 and 5.1.

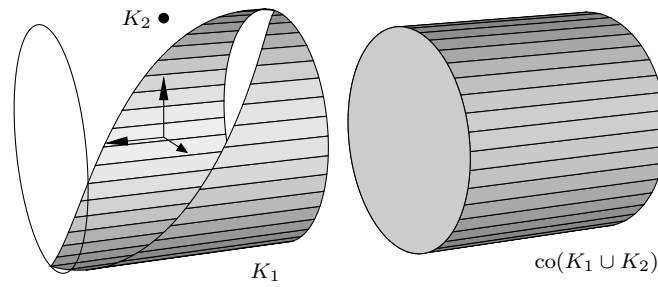


Figure 5.1: The sets K_1 and K_2 and their convex hull.

Example 5.3. Take again $(x, y, z, w) \in \mathbb{R}^4$ be our coordinates and let

$$\begin{aligned} K_1 &= \{(x, y, z, w) \mid x = -1, y^2 + (z - 1)^2 = 1, w = 0\}, \\ K_2 &= \{(x, y, z, w) \mid x = 0, y^2 + (w - 1)^2 = 1, z = 0\}, \\ K_3 &= \{(x, y, z, w) \mid x = 1, y^2 + (z - 1)^2 = 1, w = 0\}. \end{aligned} \quad (31)$$

Define the function f to be identically 1 on K_1 and K_3 , and let it be identically 0 on K_2 . Let $K = \text{co}(K_1 \cup K_2 \cup K_3)$. The set of extreme points of K is $K_1 \cup K_2 \cup K_3 \setminus \{(0, 0, 0, 0)\}$. The function f is continuous, convex, and defined precisely on the closure of the extreme points.

As $\{w = 0\}$ is a supporting hyperplane, the convex roof construction done in $\{w = 0\}$ is equal to the convex roof construction done in $\mathbb{R}^4$ and restricted to $\{w = 0\}$. On $\{w = 0\}$ we are in the situation of Example 5.1, and thus $\hat{f}$ is not continuous.

In the above two examples, there always existed some continuous convex extension. However, the following modification of Example 5.1 shows that this is not always true either.

Example 5.4. Let $(x, y, z) \in \mathbb{R}^3$ be our coordinates. Let

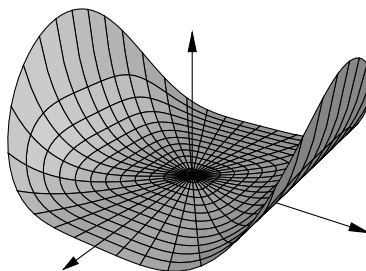
$$\begin{aligned} K_1 &= \{(x, y, z) \mid -1 \leq x \leq -z, y^2 + z^2 = 1\} \cup \{(x, y, z) \mid x = 1, y^2 + z^2 = 1\}, \\ K_2 &= \{(0, 0, 1)\}. \end{aligned} \quad (32)$$

See Figure 5.1. The convex hull of the union is again cylinder

$$\text{co}(K_1 \cup K_2) = \{(x, y, z) \mid -1 \leq x \leq 1, y^2 + z^2 \leq 1\}. \quad (33)$$

Define f to be identically 1 on K_1 , and let f be zero on K_2 . Then f is obviously continuous. Also f is convex as it is the restriction of the convex roof construction from Example 5.1.

On a line $\{(x, y, z) \mid -1 \leq x \leq 1, y = y_0, z = z_0\}$, $y_0 \neq 0$, any convex extension must be identically 1, since f is 1 at the endpoints and identically 1 for all $x \in [-1, -z_0]$. Therefore, any convex extension will fail to be continuous at $(0, 0, 1)$.

Figure 5.2: $\hat{f}(x, y) = 1 - \sqrt{1 - y^4}$ on $x^4 + y^4 \leq 1$.

Example 5.5. Let us construct a C^∞ function on the set where $x^4 + y^4 = 1$ whose convex roof extension is not Lipschitz on the convex hull.

Let $K = \{(x, y) : x^4 + y^4 \leq 1\}$. Define $f: \text{bd}(K) \rightarrow \mathbb{R}$ by $f(x, y) = 1 - \sqrt{1 - y^4}$. The function f is C^∞ ; the only issue arises at $(0, \pm 1)$, by parametrizing $\text{bd}(K)$ near $(0, 1)$ by $(x, (1 - x^4)^{1/4})$, we notice that on $\text{bd}(K)$ near $(0, 1)$, f is given by $1 - x^2$.

It is also not hard to see that $\hat{f}(x, y) = 1 - \sqrt{1 - y^4}$. This is because for a fixed y , being constant is the largest that $\hat{f}$ can be and $\hat{f}$ is the largest convex extension. But that means that the derivative of $\hat{f}$ goes to infinity when we approach $(0, \pm 1)$ from the inside of the disc, and hence $\hat{f}$ is not Lipschitz on K . See Figure 5.2.

As $\hat{f}$ is the largest convex extension, it is not hard to see that any other convex extension of f is also not Lipschitz (the derivative must also blow up at $(0, \pm 1)$).

Example 5.6. It is rather surprising that even if f is C^∞ on the boundary, the convex roof need not be C^2 on the interior (though it must be C^1). Let $K = \{(x, y) : x^2 + y^2 \leq 1\}$. Let $f: \text{bd}(K) \rightarrow \mathbb{R}$ be defined by $(x + 1)x^2$. Obviously f is C^∞ (in fact real-analytic) on $\text{bd}(K)$. By Theorem 1.4 $\hat{f} \in C^1(K)$.

We now note that on the convex hull of the points $(0, 1)$, $(0, -1)$, and $(-1, 0)$, the convex roof $\hat{f}$ must be identically zero. On the other hand, by a similar argument as above, for $x \geq 0$, we have $\hat{f}(x, y) = (x + 1)x^2$. This means that for example on the line $y = 0$, the function $\hat{f}$ is identically zero for $x \leq 0$ and $(x + 1)x^2$ for $x \geq 0$. Therefore $\hat{f}$ is not C^2 at the origin (and in fact along the line $x = 0$).

6. Application to quantum computing

The convex roof extension was introduced by Uhlmann [7] as a way of defining a different measure of entanglement on the set of density operators.

Given a quantum system there is associated to it a complex Hilbert space $\mathcal{H}$ so that all the possible states of that quantum system are in 1-1 correspondence with the set of 1-dimensional subspaces of $\mathcal{H}$, that is, the projective space $\mathbb{P}(\mathcal{H})$. By assigning the orthogonal projection onto a 1-dimensional subspace we obtain a bijection between such orthogonal projectors and $\mathbb{P}(\mathcal{H})$. In particular if ρ is such a projector it has trace 1.

Mixed states of a quantum system are probabilistic averages of pure states. The mixed states of a quantum system are in 1-1 correspondence with the set of positive semidefinite operators on $\mathcal{H}$ of trace 1. This set is a compact convex subset with nonempty interior of the set of Hermitian operators of trace 1, which in turn is an affine subspace of the set of complex linear operators on $\mathcal{H}$.

Any positive semidefinite operator of trace 1 is a convex combination of projectors of rank 1. Hence the set extreme points of the convex set of states (the mixed states) consists of the orthogonal projectors of rank 1, the so-called pure states. Note also that for any positive semidefinite operator ρ of trace 1 we have the inequality $\text{Trace}(\rho^2) \leq 1$ with equality if and only if ρ is a projector of rank 1. Therefore the pure states are the relative boundary of the set of all states. The set of pure states can be identified with the set of nonzero vectors in $\mathcal{H}$ modulo multiplication by nonzero complex scalars.

Suppose that $\mathcal{H}$ is a tensor product of two Hilbert spaces $\mathcal{H}_1, \mathcal{H}_2$,

$$\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2. \quad (34)$$

A pure state given by the vector v is called a product state (or unentangled) if v is a tensor product of two vectors $v_i \in \mathcal{H}_i$

$$v = v_1 \otimes v_2. \quad (35)$$

This decomposition does not depend on the representative chosen; a vector is a tensor product then any multiple of it is again a tensor product.

We call a pure state entangled if (any) representative vector is *not* a tensor product. We consider functions defined on the set of pure states that are zero on the set of unentangled states and larger than zero on the set of entangled states. This is called a measure of entanglement.

The question posed by Uhlmann is as follows: How to extend a measure of entanglement from pure states to mixed states? In other words, how to extend a function from the (relative) boundary of the set of states to the whole set. The method proposed is the convex roof construction.

This was later applied to extend particular measures of entanglement, for instance the Von Neumann entropy (see Nielsen and Chuang [3]) or linear entropy.

Let us analyse a bit the common structure for these measures of entanglement. Recall (also see [3]) that for every vector v there exists an integer $r \geq 1$, strictly positive numbers $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r$, and orthonormal systems $e_1, \dots, e_r$ in $\mathcal{H}_1$ and $f_1, \dots, f_r$ in $\mathcal{H}_2$ such that

$$v = \sum_{i=1}^r \sqrt{\lambda_i} e_i \otimes f_i \quad (\text{the Schmidt decomposition}). \quad (36)$$

The integer r and the λ_j are uniquely determined. The vector v is a tensor product if and only if $r = 1$.

Assuming $\|v\| = 1$ we can measure how “far” is r from 1 by, for example, considering a function ϕ with some extra properties ($\phi(0) = 0$, $\phi \geq 0$, etc. . .) and then defining $f(v) = \sum \phi(\lambda_i)$, or a function of this sum. For example, the linear entropy is defined by

$$f(v) = \sqrt{1 - \sum_{i=1}^r \lambda_i^2}. \quad (37)$$

The function f is zero if and only if exactly one of the numbers λ_i is 1. One also notices that f is continuous. In the case when $\mathcal{H}_1$ and $\mathcal{H}_2$ are both of dimension 2, there exists a closed formula for the convex roof extension of the linear entropy due to Wootters [8]. Note that such a closed formula also describes the set of separable mixed states, that is, mixed states that are convex combination of pure product states.

However, in the study of measures of entanglement there is an acute lack of explicit formulas for the convex roof extensions. A consequence of our results is that even if explicit formulas are unknown, still one concludes that all the measures of entanglements defined using the convex roof construction are continuous if the measure is continuous on the set of pure states.

Theorem 6.1. *Let f be a continuous measure of entanglement defined on the pure states, then the convex roof extension $\hat{f}$ is a continuous function of all states.*

Acknowledgements. The authors would like to acknowledge David Meyer for inspiring suggestions and Jon Grice for useful discussions. We are grateful to David Ullrich for pointing out a subtle typo.

References

- [1] D. Gale, V. Klee, R. T. Rockafellar: Convex functions on convex polytopes, Proc. Amer. Math. Soc. 19 (1968) 867–873.
- [2] A. Lima: On continuous convex functions and split faces, Proc. Lond. Math. Soc., III. Ser. 25 (1972) 27–40.
- [3] M. A. Nielsen, I. L. Chuang: Quantum Computation and Quantum Information, Cambridge University Press, Cambridge (2000).
- [4] H. J. M. Peters, P. P. Wakker: Convex functions on non-convex domains, Econ. Lett. 22 (1987) 251–255.
- [5] R. T. Rockafellar: Convex Analysis, Princeton Mathematical Series 28, Princeton University Press, Princeton (1970).
- [6] A. Uhlmann: Entropy and optimal decompositions of states relative to a maximal commutative subalgebra, Open Syst. Inf. Dyn. 5(3) (1998) 209–228.
- [7] A. Uhlmann: Roofs and convexity, Entropy 12(7) (2010) 1799–1832.
- [8] W. K. Wootters: Entanglement of formation of an arbitrary state of two qubits, Phys. Rev. Lett. 80 (1998) 2243–2248.