THE PROBLEM OF OPTIMAL SMOOTHING
FOR CONVEX FUNCTIONS

MOHAMMAD GHOMI

ABSTRACT. A procedure is described for smoothing a convex function which not only preserves its convexity, but also, under suitable conditions, leaves the function unchanged over nearly all the regions where it is already smooth. The method is based on a convolution followed by a gluing. Controlling the Hessian of the resulting function is the key to this process, and it is shown that it can be done successfully provided that the original function is strictly convex over the boundary of the smooth regions.

1. INTRODUCTION

The main aim of this note is to call attention to a basic question in convex analysis: what is the most optimal way to smooth a convex function? We hope to facilitate a complete solution to this problem, and, to this end, gather here a number of observations and techniques. First, let us state the problem precisely. A function $f \colon \mathbb{R}^n \to \mathbb{R}$ is said to be *convex* if

$$f(\lambda x_1 + (1 - \lambda)x_2) \leqslant \lambda f(x_1) + (1 - \lambda)f(x_2),$$

for all x_1, $x_2 \in \mathbb{R}^n$, and $\lambda \in [0,1]$. It is well-known that the above inequality forces f to be continuous; it can even be shown that f is *Lipschitz* continuous over any compact subset [Sc, 1.5.1]. In general, however, we cannot expect to have any higher degree of regularity. To remedy this situation one usually takes a *convolution* of f with a smooth nonnegative function θ. This operation not only smooths f, but also preserves its convexity. Further, it can be achieved by an arbitrary small perturbation of f, provided θ is chosen appropriately. But there is one drawback: such convolutions do not in general fix f over the regions where it is already smooth, and hence result in what seems to be unnecessary loss of data. Thus we are led to study the following:

1.1. Problem. Suppose there exists a subset $A \subset \mathbb{R}^n$ such that $f \in C^k(\overline{A})$, $1 \leqslant k \leqslant \infty$. Find a convex function $F \in C^k(\mathbb{R}^n)$ such that $F = f$ in A.

By $f \in C^k(\overline{A})$ we mean that the partial derivatives of f, up to order k, exist and are continuous throughout an open neighborhood of $\overline{A}$, the closure of A.

In Section of 2 of this paper we solve the above problem subject to three extra conditions, and in Section 3 we discuss how these conditions can be removed or weakened. Our main result is Theorem 2.1 together with the Remarks 3.1 and 3.2.

One might expect that the above problem would have been mentioned in standard texts such as [RV] and [Ro]. Curiously, however, it does not seem to have been well studied, despite the large literature on the approximations of convex functions and bodies, see [Gr] for a survey. A solution to 1.1 would be useful in removing singularities of convex surfaces, and as such is related to constructing barriers for elliptic PDE's of Monge-Ampére type, see [Gh]. The above problem can also be phrased in terms of functions on manifolds; see [GW] for theorems and techniques in a Riemannian geometric setting.

2. Convolution and Gluing Techniques

Here we offer a solution to Problem 1.1 subject to the following conditions: suppose that there exists a subset $A' \subset \mathbb{R}^n$ such that $A \subset A'$, $f \in C^k(\overline{A'})$, and

Condition 1: the boundary of A', $\partial A'$, is compact,
Condition 2: $k \geqslant 2$,
Condition 3: the Hessian of f, $\mathrm{D}^2 f$, is nondegenrate over $\partial A'$.

By $\mathrm{D}^2 f$, we mean the matrix of the second partial derivatives of f, $[f_{ij}]$, $1 \leqslant i, j \leqslant n$. f is convex if and only if the eigenvalues of this matrix are nonnegative [Sc, 1.5.10]. Hence condition 3 implies that $\mathrm{D}^2 f$ is *positive definite*, i.e., all the eigenvalues are greater than 0. In particular, f is strictly convex over $\partial A'$. The fact that positive definite matrices form an open subset in the space of square matrices[1] together with the compactness of $\partial A'$ (Condition 1), allows us to construct F by means of a convolution followed by a gluing.

First we set up the convolution. Let $\theta = \theta_\epsilon \in C^\infty(\mathbb{R}^n)$, have *support* in a ball of radius ϵ around the origin, $\mathrm{supp}\, \theta_\epsilon \subset B_\epsilon(0)$. This means that $\theta_\epsilon = 0$ in $\mathbb{R}^n - B_\epsilon(0)$. Further, assume that $\int_{\mathbb{R}^n} \theta = 1$, and $\theta \geqslant 0$. In some places in the literature θ is called a *mollifier* [Ev, App. C.4]; in other places it may also be known as an *approximate identity*, or *probability density*. An explicit construction is as follows:

$$(2.1) \qquad \theta_\epsilon(x) := \frac{\psi_\epsilon(x)}{\int_{\mathbb{R}^n} \psi_\epsilon}, \text{ where } \quad \psi_\epsilon(x) := \begin{cases} \exp\left(\frac{1}{\|x\|^2 - \epsilon^2}\right), & \text{if } \|x\| < \epsilon; \\ 0, & \text{otherwise.} \end{cases}$$

In the above $\|x\|$ denotes the *norm* of $x = (x_1, \ldots, x_n)$, $\|x\| := \sqrt{x_1^2 + \ldots + x_n^2}$. Using θ as the *kernel*, we define a convolution for f:

$$(2.2) \qquad \qquad \tilde{f}(x) = \tilde{f}_\epsilon(x) := \int_{\mathbb{R}^n} f(x - y) \theta_\epsilon(y)\, dy.$$

The above may be regarded as a moving local average for f. Since $\theta \in C^\infty(\mathbb{R}^n)$, it follows that $\tilde{f} \in C^\infty(\mathbb{R}^n)$. Further, since $\mathrm{supp}(\theta_\epsilon) \subset B_\epsilon(0)$, and $\int_{\mathbb{R}^n} \theta = 1$, it follows that, in any compact set K where $f \in C^k(K)$, $\tilde{f}$ converges uniformly to f together with all its derivatives up to order k,

$$(2.3) \qquad \qquad \|\tilde{f} - f\|_{C^k(K)} \to 0,$$

as $\epsilon \to 0$ (These properties of $\tilde{f}$ are basic facts of convolution; see [Hi, 2.2.3]. A detailed description of the norm $\|\cdot\|_{C^k}$ may be found in [GT, Pg. 53]). Furthermore,

---

[1] By this we mean that if $[a_{ij}]$, $1 \leqslant i, j \leqslant n$, is a positive definite matrix, then there exists an $\epsilon > 0$, depending on A, such that any matrix $[b_{ij}]$ satisfying $|a_{ij} - b_{ij}| < \epsilon$ is also positive definite.

the assumption that $\theta \geqslant 0$, together with the convexity of f, easily yields that $\tilde{f}$ is convex:

$$
\begin{aligned}
\tilde{f}(\lambda x_1 + (1-\lambda)x_2) &= \int_{\mathbb{R}^n} f(\lambda x_1 + (1-\lambda)x_2 - y)\theta_\epsilon(y)\,dy \\
&= \int_{\mathbb{R}^n} f(\lambda(x_1 - y) + (1-\lambda)(x_2 - y))\theta_\epsilon(y)\,dy \\
&\leqslant \int_{\mathbb{R}^n} [\lambda f(x_1 - y) + (1-\lambda)f(x_2 - y)]\theta_\epsilon(y)\,dy \\
&= \lambda \tilde{f}(x_1) + (1-\lambda)\tilde{f}(x_2)
\end{aligned}
$$

Next, we glue $\tilde{f}$ to f. Since $\tilde{f} \in C^k(\overline{A'})$, there exists an open neighborhood Ω of $\partial A'$ such that $\tilde{f} \in C^k(\overline{\Omega})$. Further, we can choose Ω so that $\mathrm{D}^2 f$ is nondegenerate in Ω. This is possible by the compactness of $\partial A'$, and the fact that the eigenvalues of $[f_{ij}(x)]$ depend continuously on x. Now let $U := A' \cup \Omega$. Then U is an open neighborhood of $\overline{A'}$; therefore, we can set up a (*bump*) function $\phi \in C^\infty(\mathbb{R}^n)$, with $\mathrm{supp}\,\phi \in U$, and $\phi = 1$ in A' (the existence of ϕ follows from well-known theorems on partitions of unity; see, for instance, [Hi, 2.2.2]). Finally, set

$$(2.4) \qquad F := (1-\phi)\tilde{f} + \phi f.$$

It is clear that $F \in C^k(\mathbb{R}^n)$, and $F = f$ in A'. We claim that there exists an $\epsilon > 0$ such that F is convex, which would show that F is a solution to Problem 1.1. To this end, it is enough to show that $\mathrm{D}^2 F$ is *nonnegative semidefinite* (a.k.a., *positive semidefinite*) i.e., the eigenvalues are greater than or equal to zero. Since $F = f$ in A', and $F = \tilde{f}$ in $\mathbb{R}^n - U$, we need to check this only in $U - A' \subset \Omega$. By rewriting (2.4) as $F := \tilde{f} + \phi(f - \tilde{f})$ and differentiating it twice we get:

$$(2.5) \qquad F_{ij} = \tilde{f}_{ij} + \phi_{ij}(f - \tilde{f}) + \phi_i(f - \tilde{f})_j + \phi_j(f - \tilde{f})_i + \phi(f - \tilde{f})_{ij}.$$

Note that ϕ and consequently all of its derivatives are independent of ϵ. Further note that $\overline{\Omega}$ is compact; therefore, by property (2.3), $\|\tilde{f} - f\|_{C^2(\overline{\Omega})} \to 0$. Thus it follows from (2.5) that, in $\overline{\Omega}$, $|F_{ij} - \tilde{f}_{ij}| \to 0$, uniformly as $\epsilon \to 0$. But $|\tilde{f}_{ij} - f_{ij}| \to 0$, by (2.3). Hence, in $\overline{\Omega}$,

$$(2.6) \qquad |F_{ij} - f_{ij}| \to 0,$$

uniformly as $\epsilon \to 0$. Now recall that $[f_{ij}]$ is positive definite in $\overline{\Omega}$. Thus, we conclude that there exists an $\epsilon > 0$ such that $[F_{ij}]$ is positive definite in $\overline{\Omega}$ as well. So F eventually becomes convex as ϵ gets smaller. We have proved the following:

2.1. Theorem. *If Conditions 1, 2, and 3 are satisfied, then there exists an $\epsilon > 0$ such that F, as defined in (2.4), is convex. In particular, this gives a solution to Problem 1.1.*

3. Degenerate Hessians

Here we show that when $n = 1$, i.e., f is a function of a single variable, we can solve Problem 1.1 without requiring condition 3 used in the previous section. Further, in all other dimensions, this Condition may be be replaced by a weaker assumption . A key observations utilized here is that the convolution (2.2) fixes f over almost all the regions where f is linear[2].

---

[2]A similar technique has been studied in [He] in connection with smoothing convex polygons.

Let F be defined as before, i.e., by (2.4). Recall that all we need is to show that $F'' \geqslant 0$ in a neighborhood Ω of $\partial A'$. To this end, note that when $n = 1$, we can assume that $\partial A'$ is a discrete subset. Further, it follows from condition 1 that A' is finite, say $\partial A' = \{x_1, \ldots x_m\}$. By replacing A' with a slightly larger subset, if necessary, we can assume that, no x_i is an isolated degenerate point of the Hessian of f, i.e., for every $1 \leqslant i \leqslant m$, there exists an $r_i > 0$ such that throughout $B_{r_i}(x_i)$ either $f'' > 0$ or else $f'' = 0$. Let $r := \frac{1}{2} \min\{r_1, \ldots, r_m\}$, $\Omega_i := B_r(x_i)$, and $\Omega := \bigcup_i \Omega_i$. Then throughout each Ω_i, either $f'' > 0$ or else $f'' = 0$.

If $f'' = 0$ in some Ω_i, then integrating twice yields constants a and b such that
$$f(x) = ax + b,$$
for every $x \in \Omega_i$, because Ω_i is connected. Choose ϵ sufficiently small so that $f'' = 0$ in an ϵ-neighborhood of Ω_i. Then, for every $x \in \Omega_i$ and $y \in B_\epsilon(0)$, $f(x - y) = a(x - y) + b$. Consequently it follows from (2.2) that:
$$
\begin{aligned}
\tilde{f}(x) &= \int_{\mathbb{R}^n} [a(x - y) + b]\theta_\epsilon(y)\, dy \\
&= ax \int_{\mathbb{R}^n} \theta_\epsilon(y)\, dy - \int_{\mathbb{R}^n} y\theta_\epsilon(y)\, dy + b \int_{\mathbb{R}^n} \theta_\epsilon(y)\, dy \\
&= ax(1) - 0 + b(1) \\
&= f(x),
\end{aligned}
$$
for every $x \in \Omega_i$ (here we have used the requirement that $\int_{\mathbb{R}^n} \theta = 1$, and the fact that θ, as constructed in (2.1), is *even*, i.e., $\theta(y) = \theta(-y)$). Thus, for ϵ sufficiently small, $\tilde{f} = f$, and consequently $F = f$ in Ω_i. In particular, $F'' = f'' \geqslant 0$, in every component Ω_i of Ω where $f'' = 0$. Furthermore, if throughout any component Ω_i of Ω $f'' > 0$, then, by (2.6), there exists an $\epsilon > 0$ such that $F'' > 0$ in that component as well. Hence, since Ω has only finitely many components, we conclude that:

3.1. Remark. If $n = 1$, then Condition 3 in Theorem 2.1 may be removed.

In higher dimensions degeneracy of the Hessian does not in general imply linearity of the function. Thus the solution presented in this section is guaranteed to work only in dimension 1. On the other hand, if the Hessian *vanishes*, i.e., all the eigenvalues are zero, throughout some open neighborhood of $\partial A'$, then again we have linearity, and by arguing along the same lines as before it can be shown that:

3.2. Remark. In all dimensions, Condition 3 in Theorem 2.1 may be weakened as follows:

> **Condition 3′:** on each component of $\partial A'$, either the Hessian is nondegenerate, or else vanishes in an open neighborhood of that component.

To see the above let Ω_i be a neighborhood of a component C_i of $\partial A'$ where the Hessian vanishes. We can choose Ω_i so that for every $x \in \Omega_i$, the Hessian vanishes in $B_r(x)$, for some fixed r. Since C_i is compact, we can assume that $r > 0$. Now if $\epsilon \leqslant r$, then $F = f$ in Ω_i; because, by Taylor's theorem, f is linear in $B_r(x)$ for every $x \in \Omega_i$, and consequently $\tilde{f}(x) = f(x)$.

4. End Note

We do not know whether Problem 1.1 is always solvable. Of the three extra conditions imposed in Section 2, we consider the first to be a mild restriction,

perhaps it is even necessary. Condition 2 is also quite reasonable, and should cover many, if not all, of the interesting cases. The third condition, however, appears to be too restrictive, and may not be necessary even in its weakened form. If possible, it would be desirable to remove Condition 3 altogether.

Acknowledgments. The author thanks Ralph Howard for helpful comments.

REFERENCES

[Ev] Evans, L. *Partial differential equations*. Graduate Studies in Mathematics, 19. American Mathematical Society, Providence, RI, 1998.

[Gh] Ghomi, M. *Strictly Convex Submanifolds and Hypersurfaces of Positive Curvature*, to appear, preprint available at www.math.sc.edu/~ghomi.

[Gr] Gruber, P. *Aspects of approximation of convex bodies*. Handbook of convex geometry, Vol. A, B, 319–345, North-Holland, Amsterdam, 1993.

[GT] Gilbarg, D.; Trudinger, N. *Elliptic partial differential equations of second order*. Second edition. Grundlehren der Mathematischen Wissenschaften, 224. Springer-Verlag, Berlin-New York, 1983.

[GW] Greene, R.; Wu, H. C^∞ *approximations of convex, subharmonic, and plurisubharmonic functions*. Ann. Sci. cole Norm. Sup. (4) 12 (1979), no. 1, 47–84.

[He] Helms, L. *Brownian motion in a closed convex polygon with normal reflection*. Ann. Acad. Sci. Fenn. Ser. A I Math. **17** (1992), no. 2, 199–209.

[Hi] Hirsch, M. *Differential topology*. Graduate Texts in Mathematics, No. 33. Springer-Verlag, New York-Heidelberg, 1976.

[Ro] Rockafellar, R. *Convex analysis*. Princeton Mathematical Series, No. 28 Princeton University Press, Princeton, N.J. 1970

[RV] Roberts, A.; Varberg, E. *Convex functions*. Pure and Applied Mathematics, Vol. 57. Academic Press, New York-London, 1973.

[Sc] Schneider, R. *Convex bodies: the Brunn-Minkowski theory*. Encyclopedia of Mathematics and its Applications, 44. Cambridge University Press, Cambridge, 1993.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF SOUTH CAROLINA, COLUMBIA, SC 29208
E-mail address: ghomi@math.sc.edu
URL: http://www.math.sc.edu/~ghomi/