THE SECOND DERIVATIVE OF A CONVEX FUNCTION

Daniel M. Oberlin

Department of Mathematics, Florida State University

May 2004

Suppose f is a convex function on an open interval I. The following facts are well known and easy to verify:

(a) the second (distributional) derivative of f is a nonnegative locally finite Borel measure on I, and any such measure is the second derivative of a convex function f which is unique up to the addition of an affine function;

(b) it follows from (a) that

$$\langle f'', \phi\rangle \geq 0 \tag{1}$$

whenever $\phi \in C_c^\infty(I)$ is nonnegative; conversely, if f is a distribution on I which satisfies (1), then f is a convex function.

The purpose of this note is to describe certain analogues of these facts for functions of several variables and then to prove Theorem 2 below. Thus let U be a convex open subset of $\mathbb{R}^n$. Suppose f is a convex function on U. Then, as a continuous function, f defines a distribution on U. The derivatives of f mentioned below are to be interpreted in that sense. In particular, the second derivative $D^2 f$ is the Hessian matrix of distributions on U whose entries are the second-order partial derivatives $f_{x_i x_j}$ of f. Let $\mathcal{B}_0(U)$ be the collection of all Borel sets $E \subset U$ having compact closures contained in U. An analogue of (a) is

Theorem 1. *Suppose f is convex on U. For $1 \leq i, j \leq n$ there are real-valued set functions $\mu_{ij} : \mathcal{B}_0(U) \to \mathbb{C}$ which are complex measures on any compact subset of U. For all $\phi \in C_0^\infty(U)$ and $k = 1, \cdots, n$, the μ_{ij} satisfy the equations*

$$\langle f_{x_i x_j}, \phi\rangle = \langle \mu_{ij}, \phi\rangle, \tag{2}$$

$$\langle \mu_{ij}, \phi_{x_k}\rangle = \langle \mu_{ik}, \phi_{x_j}\rangle. \tag{3}$$

Further, the symmetric matrix

$$\mu(E) = \Big(\mu_{ij}(E)\Big)$$

is positive definite for all $E \in \mathcal{B}_0(U)$. Conversely, if

$$\mu = \Big(\mu_{ij}\Big)$$

is such a (positive definite matrix valued) set function, then there is a convex function f on U with $D^2 f = \mu$. Such an f is unique up to the addition of an affine function.

The following analogue of (b) is a corollary of Theorem 1.

Corollary. *Suppose f is a distribution on U satisfying*

$$\sum_{i,j=1}^{n} \langle f_{x_i x_j}, \phi_i \phi_j \rangle \geq 0$$

whenever $\phi_i \in C_0^\infty(U)$. Then f is a convex function on U.

More interestingly, there is a lower bound on the Hausdorff dimension of the support of $D^2 f$ for nontrivial convex functions f:

Theorem 2. *With $D^2 f = \mu$ as above, if μ is supported on a Borel set having Hausdorff dimension less than $n-1$, then f is affine.*

To begin the proofs, suppose that η is a nonnegative and radial C^∞ function supported in $B(0,1)$ and having integral 1. With $\eta_\epsilon(x) = \epsilon^{-n}\eta(x/n)$ for $\epsilon > 0$ and for ρ a distribution on U, let $\rho^\epsilon = \eta_\epsilon * \rho$ be the usual ϵ-mollification of ρ, a distribution acting on those $\phi \in C_0^\infty(U)$ whose support is distant at least ϵ from ∂U.

Proof of Theorem 1: Suppose that f is convex on U. Fix an open $V \subseteq U$ such that V has compact closure contained in U and choose $\epsilon_0 > 0$ such that $V + B(0, 3\epsilon_0) \subseteq U$. Then if $0 < \epsilon < \epsilon_0$, f^ϵ is convex and C^∞ on V. Suppose $\psi \in C_0^\infty$ is nonnegative, identically 1 on V, and supported in $V + B(0, 2\epsilon_0)$. Since

$$0 \leq \int_V f^\epsilon_{x_i x_i}(x)\ dx \leq \int f^\epsilon_{x_i x_i}(x)\psi(x)dx =$$

$$\int f^\epsilon(x)\psi_{x_i x_i}(x)\ dx \longrightarrow \int f(x)\psi_{x_i x_i}(x)\ dx$$

as $\epsilon \to 0$, the nonnegative measures

$$\chi_V(x) f^\epsilon_{x_i x_i}(x)\ dx$$

have uniformly bounded total variations on V. Since

$$|f^\epsilon_{x_i x_j}| \leq |f^\epsilon_{x_i x_i}|^{1/2} |f^\epsilon_{x_j x_j}|^{1/2} \leq \frac{|f^\epsilon_{x_i x_i}| + |f^\epsilon_{x_j x_j}|}{2},$$

the same is true for the complex measures

$$\chi_V(x) f^\epsilon_{x_i x_j}(x)\ dx.$$

Taking weak* limits gives complex measures μ_{ij} on V satisfying (2) and (3) if $\phi \in C_0^\infty(V)$. Letting $V \longrightarrow U$ then yields the set functions μ_{ij} whose existence is the first claim of Theorem 2. (The statement about the positive definite character of $\mu(E)$ follows from

$$\sum_{i,j=1}^{n} \int_V f^{\epsilon}_{x_i x_j}(x)\phi_i(x)\phi_j(x)\ dx \geq 0$$

and a limit argument.)

Suppose now that $\mu = (\mu_{ij})$ is as in Theorem 2. Suppose that V and ϵ_0 are as above and suppose additionally that ∂V is C^1. Fix $i,j = 1,\ldots,n$. For $\epsilon < \epsilon_0$, the C^∞ functions μ^{ϵ}_{ij} have uniformly bounded $L^1(V + B(0,\epsilon))$ norms (since the set functions μ_{ij} are complex measures on $V + B(0, 2\epsilon_0)$). The conditions (3) and the symmetry of μ imply that there are C^∞ functions f_ϵ on $V + B(0,\epsilon_0)$ with

$$\frac{\partial^2 f_\epsilon}{\partial x_i x_j} = \mu^{\epsilon}_{ij}.$$

The Poincaré-Sobolev inequalities and the uniform bounds on

$$\|\mu^{\epsilon}_{ij}\|_{L^1(V+B(0,\epsilon_0))}$$

show that there is $q = q(n) > 1$ such that the functions f_ϵ can be chosen to have

$$\|f_\epsilon\|_{L^q(V+B(0,\epsilon_0))} \leq\ C$$

for some positive C independent of ϵ. Passing to a subsequence which converges weakly in $L^q(V + B(0,\epsilon_0))$ shows that there is $f \in L^q(V + B(0,\epsilon_0))$ satisfying

$$\frac{\partial^2 f}{\partial x_i x_j} = \mu_{ij} \tag{4}$$

(in the sense of distributions on $V + B(0,\epsilon_0)$). To check that f is equal a.e. to a convex function on V, we begin by noting that, for $0 < \epsilon < \epsilon_0$, (4) and the positive definite assumption on μ show that the mollifications f^ϵ are convex on V. The formula

$$g^\delta(0) = \frac{1}{2}\int_{\Sigma_{n-1}}\int_0^\infty \eta(r\sigma)\int_{-\delta r}^{\delta r} \frac{d^2}{dt^2} g(t\sigma) r^{n-1} dr\ d\sigma + g(0)$$

for the δ-mollification of a twice differentiable function g at 0 shows that δ-mollifications of twice differentiable convex functions decrease pointwise as $\delta \longrightarrow 0$. In particular, if $B(x,\delta) \subseteq V$ and $0 < \delta < \delta_0$, this applies to $(f^\epsilon)^\delta(x)$ for each ϵ and so, in the limit as $\epsilon \to 0$, to $f^\delta(x)$. If

$$\lim_{\delta\to 0} f^\delta(x) = -\infty$$

held for any $x \in V$, it would follow from convexity that

$$\lim_{\delta\to 0} \|f^\delta\|_{L^q(V)} \to \infty.$$

Thus the decreasing limit

$$\lim_{\delta\to 0} f^\delta(x)$$

gives a realization of f as a convex function on V. Now suppose that $\widetilde{V}$ satisfies the same hypotheses as V and is such that $V \subseteq \widetilde{V}$. Then the argument above yields a convex function $\widetilde{f}$ on $\widetilde{V}$ satisfying (4) on $\widetilde{V}$. Such an $\widetilde{f}$ can be adjusted by the addition of an affine function to agree with f on V. Considering a sequence $V_1 \subseteq V_2 \subseteq \cdots$ of such open sets with $\cup V_n = U$ furnishes a convex f on U satisfying $D^2 f = \mu$ on U. Since such an f is clearly unique up to the addition of an affine function, the proof of Theorem 1 is complete.

Proof of Corollary: If f is a distribution on U satisfying the hypothesis, then it is clear that each $f_{x_i x_i}$ is nonnegative and therefore realizable as a nonnegative Borel measure μ_{ii} on U. The hypothesis also implies that

$$\sum_{i,j=1}^{n} \langle f_{x_i x_j}, \phi^2 \rangle \zeta_i \, \zeta_j \geq 0$$

for all real numbers $\zeta_1, \ldots, \zeta_n$ whenever $\phi \in C_0^\infty(U)$. It follows that

$$|\langle f_{x_i x_j}, \phi^2 \rangle| \leq \frac{\langle f_{x_i x_i}, \phi^2 \rangle + \langle f_{x_j x_j}, \phi^2 \rangle}{2} = \frac{\langle \mu_{ii}, \phi^2 \rangle + \langle \mu_{jj}, \phi^2 \rangle}{2}$$

for $\phi \in C_0^\infty(U)$. This implies that each $f_{x_i x_j}$ is realizable as a complex Borel measure μ_{ij} on each open and bounded $V \subseteq U$. Thus the desired result follows from Theorem 1.

Proof of Theorem 2: It follows from the proof of Theorem 1 that if f is convex on U, each of the distributions f_{x_j} is realizable as a function on U. Thus Theorem 2 can be proved by applying the following result to each f_{x_j}:

Lemma 1. *Suppose the locally integrable function g on U has the property that each of the distributions g_{x_j} is realized by a Borel measure μ_j supported on a set of Hausdorff dimension less than $n-1$. Then g is equal a.e. to a constant function.*

The proof depends on a simple fact:

Lemma 2. *Suppose $E \subseteq \mathbb{R}^n$ is a set of Hausdorff dimension less than $n-1$. Then the n-dimensional Lebesgue measure of*

$$\cup_{t>0} \big(tE\big)$$

is 0.

Proof of Lemma 2: The mapping $\pi : x \mapsto \frac{x}{|x|}$ is Lipschitz on

$$E_\delta \doteq E \cap \{\delta < |x| < 1/\delta\}$$

for each $\delta > 0$. Thus $\pi(E_\delta)$ has $(n-1)$-dimensional measure 0 in Σ_{n-1} and the desired result follows by integrating in polar coordinates.

Proof of Lemma 1: Assume without loss generality that $0 \in U$ and that

$$\lim_{\epsilon \to 0} g^\epsilon(0) = g(0).$$

This is possible since the local integrability of g implies that

$$\lim_{\epsilon\to 0} g^\epsilon(x) = g(x)$$

for a.e. $x \in U$. If h is C^1 on U and $x \in U$, then

$$h(x) = \int_0^1 \sum_{j=0}^n x_j h_{x_j}(xt)\ dt + h(0)$$

leads to

$$\int_U h(x)\phi(x)\ dx = h(0)\int_U \phi(x)\ dx + \int_U \int_\delta^1 \phi(\frac{x}{t}) \sum_{j=0}^n x_j h_{x_j}(x) \frac{dt}{t^{n+1}} dx$$

for $\phi \in C_0^\infty(U)$ and $\delta = \delta(\phi) > 0$. Thus

$$\langle g, \phi\rangle = \lim_{\epsilon\to 0}\langle g^\epsilon, \phi\rangle =$$

$$\lim_{\epsilon\to 0}\big[g^\epsilon(0)\int_U \phi(x)\ dx + \int_U\int_\delta^1 \phi(\frac{x}{t})\sum_{j=0}^n x_j g^\epsilon_{x_j}(x)\frac{dt}{t^{n+1}}dx\big].$$

Letting $\phi_t(x) = \phi(x/t)$, we have

$$\lim_{\epsilon\to 0}\int_U g^\epsilon_{x_j}(x) x_j \phi_t(x)\ dx = \langle g_{x_j}, x_j\phi_t\rangle$$

uniformly for $\delta \le t \le 1$ and so

$$\langle g, \phi\rangle = \big(\lim_{\epsilon\to 0} g^\epsilon(0)\big)\int_U \phi(x)\ dx + \int_U\int_\delta^1 \phi(\frac{x}{t})\sum_{j=0}^n x_j g_{x_j}(x)\frac{dt}{t^{n+1}}dx. \tag{5}$$

If E is a set of Hausdorff dimension less than $n-1$ which supports each of the measures g_{x_j}, then (5) implies that

$$\langle g, \phi\rangle = \big(\lim_{\epsilon\to 0} g^\epsilon(0)\big)\int_U \phi(x)\ dx + \int_U \phi(x)\ d\nu(x)$$

where the measure ν is supported on the set

$$\cup_{t>0}\big(tE\big).$$

By Lemma 2, that set has Lebesgue measure 0. Since g is a locally integrable function, it must be the case that $\nu = 0$, and Lemma 1 follows.