CONVEX FUNCTIONS WITH UNBOUNDED GRADIENT

OLIVER C. SCHNÜRER

Abstract. We show that domains, that allow for convex functions with unbounded gradient at their boundary, are convex.

In this paper, we prove the following

Theorem 1. *Let $\Omega \subset \mathbb{R}^n$ be any domain. If there exists a convex function $u : \Omega \to \mathbb{R}$, such that ∇u becomes unbounded near $\partial\Omega$, then Ω is convex.*

This problem arises in the context of affine hypersurfaces [1], where a similar statement is proven.

Acknowledgement: We want to thank Neil Trudinger, who brought this problem to our attention and encouraged us to provide a geometric proof. We also want to thank Klaus Ecker and Free University Berlin for support during the preparation of this paper and Ann Björner for converting our sketch into a beautiful illustration.

We will use the following notions of convexity.

Definition and Remark 2. Let $\Omega \subset \mathbb{R}^n$ be a given domain.

A function $u : \Omega \to \mathbb{R}$ is called convex, if

$$u(\tau x + (1-\tau)y) \le \tau u(x) + (1-\tau)u(y)$$

for all $x,\, y \in \Omega$ and all $\tau \in (0,\, 1)$ such that $\tau x + (1-\tau)y \in \Omega$.

A convex function is locally Lipschitz continuous and thus differentiable almost everywhere. We obtain

$$u(y) \ge u(x) + \langle \nabla u(x),\, y - x\rangle$$

for all $y \in \Omega$ and all $x \in \Omega$, where u is differentiable.

A function u is called locally convex, if

$$u(\tau x + (1-\tau)y) \le \tau u(x) + (1-\tau)u(y)$$

for all $x,\, y \in \Omega$ and $\tau \in (0,\, 1)$ such that $tx + (1-t)y \in \Omega$ for all $t \in (0,\, 1)$.

For $u \in C^2(\Omega)$, local convexity is equivalent to positive semi-definiteness of its Hessian in Ω.

Note especially that locally convex functions do not need to be convex, unless their domain of definition is a convex set, i. e. $x,\, y \in \Omega$, $\tau \in (0,\, 1) \Longrightarrow \tau x + (1-\tau)y \in \Omega$.

In order to strengthen Theorem 1, we fix $x_0 \in \Omega$ and define $\Omega_1 = \Omega_1(x_0)$ by

$$\Omega_1 := \{x \in \Omega : tx + (1-t)x_0 \in \Omega \text{ for } 0 \le t \le 1\}$$

Date: April 2005, revised September 2005.

2000 *Mathematics Subject Classification.* 52A99.

Key words and phrases. Convex function, convex domain.

The author is a member of SFB 647/3B "Raum – Zeit – Materie".

as the largest subdomain of Ω that is star shaped with respect to x_0. Note that Ω is convex if and only if $\Omega = \Omega_1(x)$ for all $x \in \Omega$.

For convenience, we will assume from now on that x_0 is the origin. So the domain $\Omega_1 = \Omega_1(0)$ is star shaped with respect to the origin.

We define

$$\partial_2\Omega_1 := \{x \in \partial\Omega_1 : \exists \varepsilon > 0 : tx \in \partial\Omega_1 \forall t \in (1-\varepsilon,\, 1+\varepsilon)\}$$

and $\partial_1\Omega_1 := \partial\Omega_1 \setminus \partial_2\Omega_1$.

For Ω as in Theorem 1, we deduce that ∇u becomes unbounded near $\partial_1\Omega_1$.

Thus the following result generalizes Theorem 1 and [1, Lemma 2.4].

Theorem 3. *Let $\Omega \subset \mathbb{R}^n$ be open and star shaped with respect to $0 \in \Omega$. If there exists a convex function $u : \Omega \to \mathbb{R}$, such that*

$$\|\nabla u\|_{L^\infty(B_\delta(x)\cap\Omega)} = \infty$$

for any $x \in \partial_1\Omega$ and any $\delta > 0$, then Ω is convex.

Here, the set $\partial_1\Omega$ is defined as above. We consider ∇u only at those points, where u is differentiable.

Proof. We argue by contradiction. Assume that there exist $\tilde{y}_0$ and $\tilde{y}_1$ in Ω and $\tilde{\tau} \in (0,\, 1)$ such that

$$\tilde{\tau}\tilde{y}_0 + (1-\tilde{\tau})\tilde{y}_1 \notin \Omega.$$

Let $\lambda > 0$ be the supremum over all positive numbers such that

$$t\lambda\tilde{y}_0 + (1-t)\lambda\tilde{y}_1 \in \Omega \quad \text{for all } 0 \le t \le 1.$$

As 0 is an interior point of Ω, λ is positive. Define $y_0 := \lambda\tilde{y}_0$ and $y_1 := \lambda\tilde{y}_1$. Let $\tau \in (0,\, 1)$ be such that $\tau y_0 + (1-\tau)y_1 \in \partial\Omega$. According to the definition of λ,

$$\mu(\tau y_0 + (1-\tau)y_1) \in \Omega \quad \text{for all } 0 \le \mu < 1.$$

So we deduce that $\tau y_0 + (1-\tau)y_1 \in \partial_1\Omega$. Thus, there exist $x_i \in \Omega$, $i \in \mathbb{N}$, such that $x_i \to \tau y_0 + (1-\tau)y_1$ for $i \to \infty$, ∇u exists at x_i for all i, and $|\nabla u(x_i)| \to \infty$ as $i \to \infty$. This contradicts the gradient bounds that we will prove in the following.

As Ω is open and star shaped, there exists $\varepsilon > 0$ such that $B_{3\varepsilon}(y_k) \subset \Omega$, $k = 0,\, 1$, and $B_{3\varepsilon}(0) \subset \Omega$. Assume that the balls $B_{3\varepsilon}(y_0)$, $B_{3\varepsilon}(y_1)$, $B_{3\varepsilon}(\tau y_0 + (1-\tau)y_1)$, and $B_{3\varepsilon}(0)$ are disjoint. By scaling u, we may therefore arrange that $|u| \le 1$ in $B_{2\varepsilon}(y_0) \cup B_{2\varepsilon}(y_1) \cup B_{2\varepsilon}(0)$.

Assume furthermore that, after an appropriate rotation, $y_1 - y_0$ is a positive multiple of $e_1 = (1,\, 0,\, \ldots,\, 0)$.

We will now use the convexity of u and boundedness of u in balls around y_0 and y_1 to prove explicit bounds on ∇u near $\tau y_0 + (1-\tau)y_1$. Set $u_k := \frac{\partial u}{\partial x^k}$, $1 \le k \le n$.

The first step is to bound u in $B_\varepsilon(\tau y_0 + (1-\tau)y_1)$. Let $z \in B_\varepsilon(0)$ be such that $\tau y_0 + (1-\tau)y_1 + z \in \Omega$. Convexity of u and $u \le 1$ in $B_\varepsilon(y_k)$, $k = 0,\, 1$, imply that

$$\begin{aligned} u(\tau y_0 + (1-\tau)y_1 + z) =& u(\tau(y_0+z) + (1-\tau)(y_1+z)) \\ \le& \tau u(y_0+z) + (1-\tau)u(y_1+z) \\ \le& \tau \cdot 1 + (1-\tau)\cdot 1 = 1. \end{aligned}$$

For a lower bound on u, we compare u at $\tau y_0 + (1-\tau)y_1 + z =: p$, at $2\varepsilon\frac{p}{|p|}$, and at the origin. We get $|p| \le |\tau y_0 + (1-\tau)y_1| + \varepsilon$,

$$2\varepsilon\frac{p}{|p|} = tp + (1-t)0$$

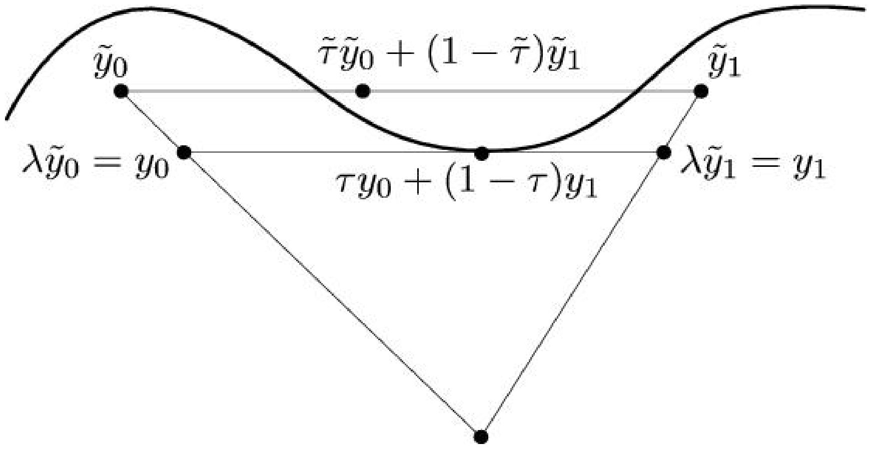

FIGURE 1. Geometric situation

with $t = \frac{2\varepsilon}{|p|}$, and

$$|p| \geq |\tau y_0 + (1-\tau)y_1| - |z| \geq 6\varepsilon - \varepsilon,$$

as $B_{3\varepsilon}(\tau y_0 + (1-\tau)y_1) \cap B_{3\varepsilon}(0) = \emptyset$. Thus

$$u\left(2\varepsilon\frac{p}{|p|}\right) \leq tu(p) + (1-t)u(0)$$

and we obtain that

$$\begin{aligned} u(p) \geq& \frac{1}{t}\left(u\left(2\varepsilon\frac{p}{|p|}\right) - (1-t)u(0)\right) \\ \geq& \frac{|\tau y_0 + (1-\tau)y_1| + \varepsilon}{2\varepsilon}(-1-1) \\ =& -\frac{1}{\varepsilon}(|\tau y_0 + (1-\tau)y_1| + \varepsilon). \end{aligned}$$

Thus $|u| \leq 1 + \frac{1}{\varepsilon}|\tau y_0 + (1-\tau)y_1| \equiv C_0$ in $B_\varepsilon(\tau y_0 + (1-\tau)y_1)$.

The next step is to bound $|u_1|$ in $B_\varepsilon(\tau y_0 + (1-\tau)y_1)$. Remember that $y_1 - y_0$ is a positive multiple of e_1. As balls of radius 3ε as chosen above are disjoint, we get

$$|(1-\tau)(y_0 - y_1)| = |y_0 - (\tau y_0 + (1-\tau)y_1)| \geq 6\varepsilon$$

and

$$|\tau(y_0 - y_1)| = |(\tau y_0 + (1-\tau)y_1) - y_1| \geq 6\varepsilon.$$

Thus we obtain

$$1 - \tau \geq \frac{6\varepsilon}{|y_0 - y_1|} \quad \text{and} \quad \tau \geq \frac{6\varepsilon}{|y_0 - y_1|}. \tag{1}$$

Let $z \in B_\varepsilon(0)$ and $\tau y_0 + (1-\tau)y_1 + z =: p$ be an arbitrary point in $B_\varepsilon(\tau y_0 + (1-\tau)y_1)$, where u is differentiable. We deduce that

$$p + (1-\tau)(y_0 - y_1) \in B_\varepsilon(y_0)$$

and

$$p+\tau(y_1-y_0)\in B_\varepsilon(y_1).$$

Convexity and boundedness of u in $B_\varepsilon(y_0)\cup B_\varepsilon(y_1)\cup B_\varepsilon(\tau y_0+(1-\tau)y_1)$ imply that

$$\begin{aligned}\tau\,|y_1-y_0|\,\langle\nabla u(p),\,e_1\rangle =&\langle\nabla u(p),\,\tau(y_1-y_0)\rangle\\ \le& u(p+\tau(y_1-y_0))-u(p)\\ \le& 1+C_0\end{aligned}$$

and

$$\begin{aligned}(1-\tau)\,|y_1-y_0|\,\langle\nabla u(p),\,-e_1\rangle =&\langle\nabla u(p),\,(1-\tau)(y_0-y_1)\rangle\\ \le& u(p+(1-\tau)(y_0-y_1))-u(p)\\ \le& 1+C_0.\end{aligned}$$

We insert the bounds (1) and deduce that

$$|\langle\nabla u(p),\,e_1\rangle|=|u_1|\,(p)\le\frac{1}{6\varepsilon}(1+C_0).$$

The last step is to bound u_k in $B_\varepsilon(\tau y_0+(1-\tau)y_1)$ for $2\le k\le n$. Consider an arbitrary vector $\xi\in\mathbb{R}^n$ such that $|\xi|=1$ and $\langle\xi,\,e_1\rangle=0$. As above, let $p\in B_\varepsilon(\tau y_0+(1-\tau)y_1)\cap\Omega$ be chosen arbitrarily such that u is differentiable in p. Then we have

$$p+(1-\tau)(y_0-y_1)\pm\varepsilon\xi\in B_{2\varepsilon}(y_0).$$

We remark that considering $p+\tau(y_1-y_0)\pm\varepsilon\xi\in B_{2\varepsilon}(y_1)$ in the following yields a similar estimate. As u is convex, we get

$$\langle\nabla u(p),\,(1-\tau)(y_0-y_1)\pm\varepsilon\xi\rangle\le u(p+(1-\tau)(y_0-y_1)\pm\varepsilon\xi)-u(p).$$

So we deduce

$$\begin{aligned}\varepsilon\langle\nabla u(p),\,\pm\xi\rangle \le& u(p+(1-\tau)(y_0-y_1)\pm\varepsilon\xi)-u(p)+\langle\nabla u(p),\,(1-\tau)(y_1-y_0)\rangle\\ \le&\left(1+\frac{|y_0-y_1|}{6\varepsilon}\right)(1+C_0).\end{aligned}$$

Thus $|\nabla u|$ is bounded in $B_\varepsilon(\tau y_0+(1-\tau)y_1)$, wherever u is differentiable. This contradicts the assumption that $\tau y_0+(1-\tau)y_1\in\partial_1\Omega$. The theorem follows. □

Remark 4. In Theorem 1, we don't have to assume Ω being connected. If y_0 and y_1 lie in different components of Ω, we find $0<\tau<1$ such that $\tau y_1+(1-\tau)y_0\in\partial\Omega$. Choose $0<t<1$, $t\ne\tau$, such that $ty_1+(1-t)y_0\in\Omega$. Assume that $t>\tau$. There exists $0<\sigma<1$ such that

$$ty_1+(1-t)y_0=\sigma(\tau y_1+(1-\tau)y_0)+(1-\sigma)y_1.$$

Thus

$$u(ty_1+(1-t)y_0)\le\sigma u(\tau y_1+(1-\tau)y_0)+(1-\sigma)u(y_1)$$

and $u(\tau y_1+(1-\tau)y_0)$ is bounded below. A similar argument bounds u from below in a neighborhood (relative to Ω) of $\tau y_1+(1-\tau)y_0$. Following the lines of the proof of Theorem 3, we obtain an upper bound on u near $ty_1+(1-t)y_0$ and gradient bounds. As ∇u is bounded near $\tau y_1+(1-\tau)y_0$, we obtain a contradiction. We conclude that Ω has to be connected.

Remark 5. Let Ω be any convex domain in $\mathbb{R}^n$ and $d := \operatorname{dist}(\cdot, \partial\Omega)$. As d is concave in Ω, $-\sqrt{d}$ is a convex function in Ω with unbounded gradient along $\partial\Omega$. Thus, according to Theorem 1, such functions exist precisely on convex domains.

Reference

[1] Neil S. Trudinger and Xu-Jia Wang, *Affine complete locally convex hypersurfaces*, Invent. Math. **150** (2002), no. 1, 45–60.

Oliver Schnürer, FU Berlin, Arnimallee 2-6, 14195 Berlin, Germany
E-mail address: Oliver.Schnuerer@math.fu-berlin.de