Continuity of Derivatives of a Convex Solution to a Perturbed One-Laplace Equation by $p$-Laplacian

YOSHIKAZU GIGA & SHUNTARO TSUBOUCHI

Communicated by G. DAL MASO

Abstract

We consider a one-Laplace equation perturbed by $p$-Laplacian with $1 < p < \infty$. We prove that a weak solution is continuously differentiable ($C^1$) if it is convex. Note that a similar result fails to hold for the unperturbed one-Laplace equation. The main difficulty is to show $C^1$-regularity of the solution at the boundary of a facet where the gradient of the solution vanishes. For this purpose we blow-up the solution and prove that its limit is a constant function by establishing a Liouville-type result, which is proved by showing a strong maximum principle. Our argument is rather elementary, since we assume that the solution is convex. A few generalizations are also discussed.

1. Introduction

We consider a one-Laplace equation perturbed by $p$-Laplacian of the form

$$L_{b,p}u = f \quad \text{in} \quad \Omega$$

(1.1)

with

$$L_{b,p}u := -b\Delta_1 u - \Delta_p u,$$

where

$$\Delta_1 u := \text{div} \left( \nabla u / |\nabla u| \right), \quad \Delta_p u = \text{div} \left( |\nabla u|^{p-2} \nabla u \right)$$

in a domain $\Omega$ in $\mathbb{R}^n$, $\nabla u = (\partial x_1 u, \ldots, \partial x_n u)$ with $\partial x_i u = \partial u / \partial x_j$ for a function $u = u(x_1, \ldots, x_n)$, and $\text{div} X = \sum_{i=1}^n \partial_i X_i$ for a vector field $X = (X_1, \ldots, X_n)$. The constants $b > 0$ and $p \in (1, \infty)$ are given and fixed. It has been a long-standing open problem whether its weak solution is $C^1$ up to a facet, the place where the gradient $\nabla u$ vanishes, even if $f$ is smooth. This is a non-trivial question since a
weak solution to the unperturbed one-Laplace equation, that is, \(-\Delta_1 u = f\) may not be \(C^1\). This is because the ellipticity degenerates in the direction of \(\nabla u\) for \(\Delta_1 u\). Our goal in this paper is to solve this open problem under the assumption that a solution is convex.

1.1. Main theorems and our strategy

Throughout the paper, we assume that \(f \in L^q_{\text{loc}}(\Omega) (n < q \leq \infty)\), that is, \(|f|^q\) is locally integrable in \(\Omega\). Our main result is

**Theorem 1.** \((C^1\)-regularity theorem\) Let \(u\) be a convex weak solution to (1.1) with \(f \in L^q_{\text{loc}}(\Omega) (n < q \leq \infty)\). Then \(u\) is in \(C^1(\Omega)\).

Difficulty on proving regularity on gradients of solutions to (1.1) can be explained from a viewpoint of ellipticity ratio. We set a convex function \(E: \mathbb{R}^n \to [0, \infty)\) by

\[
E(z) := bE_1(z) + E_p(z) \quad \text{for } z \in \mathbb{R}^n,
\]

where \(E_s (1 \leq s < \infty)\) is defined by

\[
E_s := \frac{1}{s}|z|^s \quad \text{for } z \in \mathbb{R}^n.
\]

We rewrite (1.1) by

\[-\text{div}(\nabla z E(\nabla u)) = f \quad \text{in } \Omega.\]  

(1.2)

By differentiating (1.2) by \(x_i (i \in \{1, \ldots, n\})\), we get

\[-\text{div}\left(\nabla^2 z E(\nabla u)\nabla \partial_{x_i} u\right) = \partial_{x_i} f.\]  

(1.3)

By elementary calculations, ellipticity ratio of the Hessian \(\nabla^2 z E\) at \(z_0 \in \mathbb{R}^n \setminus \{0\}\) is given by

\[
\left(\text{ellipticity ratio of } \nabla^2 z E(z_0)\right) := \frac{\text{(the largest eigenvalue of } \nabla^2 z E(z_0))}{\text{(the lowest eigenvalue of } \nabla^2 z E(z_0))} \leq \frac{\max(p - 1, 1) + b|z_0|^{1-p}}{\min(p - 1, 1)}.\]

Since the exponent \(1 - p\) is negative, the ellipticity ratio of \(\nabla^2 z E(z_0)\) blows up as \(z_0 \to 0\). By this property, we can observe that the Equation (1.2) becomes non-uniformly elliptic near the facet. It should be noted that our problem is substantially different from the \((p, q)\)-growth problem, since for \((p, q)\)-growth equations, non-uniform ellipticity appears as a norm of a gradient blows up [26, Section 6.2]. Although regularity of minimizers of double phase functionals, including

\[
\mathcal{H}(u) := \int E_p(\nabla u) \, dx + \int a(x) E_q(\nabla u) \, dx \quad \text{with } 1 < p \leq q < \infty, a(x) \geq 0,
\]
were discussed in scalar and even in vectorial cases by Colombo and Mingione [6, 7], their results do not recover our $C^1$-regularity results. This is basically derived from the fact that, unlike $\nabla^2 \phi(z) \in \mathbb{R}^n$ with $1 < p < \infty$, the Hessian matrix $\nabla^2 \phi(z_0) (z_0 \neq 0)$ always takes 0 as its eigenvalue. In other words, ellipticity of the operator $\Delta_1 u$ degenerates in the direction of $\nabla u$, which seems to be difficult to handle analytically.

On the other hand, the ellipticity ratio of $\nabla^2 \phi(z_0)$ is uniformly bounded over $|z_0| > \delta$ for each fixed $\delta > 0$. In this sense we may regard the Equation (1.3) as locally uniformly elliptic outside the facet. To show Lipschitz bound, we do not need to study over the facet. In fact, local Lipschitz continuity of solutions to (1.1) are already established in [32]; see also [33] for a weaker result. To study continuity of derivatives, we have to study regularity up to the facet. Thus, it seems to be impossible to apply standard arguments based on De Giorgi–Nash–Moser theory. In this paper, we would like to show continuity of derivatives of convex solutions by elementary arguments based on convex analysis.

Let us give a basic strategy to prove Theorem 1. Since the problem is local, we may assume that $\Omega$ is convex, or even a ball. By $C^1$-regularity criterion for a convex function, to show $u$ is $C^1$ at $x \in \Omega$ it suffices to prove that the subdifferential $\partial u(x)$ at $x \in \Omega$ is a singleton; (1.4) see [1, Appendix D], [30, §25] and Remark 1 for more detail. Here the subdifferential of $u$ at $x_0 \in \Omega$ is defined by

$$\partial u(x_0) := \left\{ z \in \mathbb{R}^n \mid u(x) \geq u(x_0) + \langle z, x - x_0 \rangle \text{ for all } x \in \Omega \right\}.$$

Here $\langle \cdot | \cdot \rangle$ stands for the standard inner product in $\mathbb{R}^n$. For a convex function $u : \Omega \to \mathbb{R}$, we can simply express the facet of $u$ as

$$F := \left\{ x \in \Omega \mid \partial u(x) \ni 0 \right\} = \left\{ x \in \Omega \mid u(x) \leq u(y) \text{ for all } y \in \Omega \right\}.$$

By definition it is clear that the facet $F$ is non-empty if and only if a minimum of $u$ in $\Omega$ exists. By convexity of $u$, we can easily check that $F \subset \Omega$ is a relatively closed convex set in $\Omega$. We also define an open set

$$D := \Omega \setminus F = \left\{ x \in \Omega \mid u(y) < u(x) \text{ for some } y \in \Omega \right\}.$$

Our strategy to show (1.4) depends on whether $x$ is inside $F$ or not.

Remark 1. (Some properties on differentiability of convex functions) Let $v$ a real-valued convex function in a convex domain $\Omega \subset \mathbb{R}^n$, then following properties hold.

1. $v$ is locally Lipschitz continuous in $\Omega$, and therefore $v$ is almost everywhere differentiable in $\Omega$ by Rademacher’s theorem ([1, Theorem 1.19], see also [9, Theorem 3.1 and 3.2] and [30, Theorem 25.5]).

2. For $x \in \Omega$, $v$ is differentiable at $x$ if and only if the subdifferential set $\partial v(x)$ is a singleton. Moreover, if $x \in \Omega$ satisfies either of these equivalent conditions, then we have $\partial v(x) = \{ \nabla v(x) \}$ ([1, Proposition D.5], see also [30, Theorem 25.1]). In particular, Rademacher’s theorem implies that $\partial v(x) = \{ \nabla v(x) \}$ for almost everywhere $x \in \Omega$. 

3. \( v \in C^1(\Omega) \) if and only if \( \partial v \) is single-valued ([1, Remark D.3 (iii)], see also [30, Theorem 25.5]).

Throughout this paper, we use these well-known results without proofs.

We first discuss the case \( x \in D \). Our goal is to show directly that \( u \) is \( C^{1,\alpha} \) near a neighborhood of \( x \) and therefore \( \partial u(x) = \{ \nabla u(x) \} \neq \{ 0 \} \) for all \( x \in D \). This strategy roughly consists of three steps. Among them the first step, a kind of separation of \( x \in D \) from the facet \( F \), plays an important role. Precisely speaking, we first find a neighborhood \( B_r(x) \subset D \), an open ball centered at \( x \) with its radius \( r > 0 \), such that

\[
\partial_v u \geq \mu > 0 \quad \text{almost everywhere in} \quad B_r(x) \tag{1.5}
\]

for some direction \( v \) and some constant \( \mu > 0 \). In order to justify (1.5), we fully make use of convexity of \( u \) (Lemma 8 in Section A), not elliptic regularity theory. Then with the aid of local Lipschitz continuity of \( u \), the inclusion \( B_r(x) \subset \{ 0 < \mu \leq \partial_v u \leq |\nabla u| \leq M \} \) holds for some finite positive constant \( M \). Secondly, this inclusion allows us to check that \( u \) admits local \( W^{2,2} \)-regularity in \( B_r(x) \) by the standard difference quotient method. Therefore we are able to obtain the Equation (1.3) in the distributional sense. Finally, we appeal to the classical De Giorgi–Nash–Moser theory to obtain local \( C^{1,\alpha} \)-regularity at \( x \in D \), since the Equation (1.3) is uniformly elliptic in \( B_r(x) \). Here the constant \( \alpha \in (0, 1) \) we have obtained may depend on the location of \( x \in D \) through ellipticity, so \( \alpha \) may tend to zero as \( x \) tends to the facet.

It takes much efforts to prove that \( \partial u(x) = \{ 0 \} \) for all \( x \in F \). Our strategy for justifying this roughly consists of three parts; a blow-argument for solutions, a strong maximum principle, and a Liouville-type theorem. Here we describe each individual step.

We first make a blow-argument. Precisely speaking, for a given convex solution \( u : \Omega \to \mathbb{R} \) and a point \( x_0 \in \Omega \), we set a sequence of rescaled functions \( \{ u_a \}_{a > 0} \) defined by

\[
 u_a(x) := \frac{u(a(x-x_0)+x_0) - u(x_0)}{a}.
\]

We show that \( u_a \) locally uniformly converges to some convex function \( u_0 : \mathbb{R}^n \to \mathbb{R} \), which satisfies \( \partial u(x_0) \subset \partial u_0(x_0) \) by construction. Moreover, we prove that \( u_0 \) satisfies \( Lb, pu_0 = 0 \) in \( \mathbb{R}^n \) in the distributional sense. There we will face to justify almost everywhere convergence of gradients, and this is elementarily shown by regarding gradients in the classical sense as subgradients (Lemma 9 in the appendices).

Next we prove that if \( x_0 \in F \), then the convex weak solution \( u_0 \) constructed as above satisfies \( \partial u_0(x_0) = \{ 0 \} \). Moreover, we are going to prove that \( u_0 \) is constant (a Liouville-type theorem). For this purpose we establish the maximum principle.

**Theorem 2.** (Strong maximum principle) Let \( u \) be a convex weak solution to \( Lb, pu = 0 \) in a convex domain \( \Omega \subset \mathbb{R}^n \) and \( F \subset \Omega \) be the facet of \( u \). Then \( u \) is affine in each connected component of the open set \( D := \Omega \setminus F \). In particular, if \( F = \emptyset \), then \( u \) is affine in \( \Omega \).
It should be noted that this result is a kind of strong maximum principle in the sense that

\[ u \geq a \text{ in } D_0 \text{ and } u(x_0) = a(x_0) \text{ for } x_0 \in D_0 \implies u \equiv a \text{ in } D_0. \tag{1.6} \]

where \( a(x) := u(x_0) + \langle \nabla u(x_0) \mid x - x_0 \rangle \) and \( D_0 \) is a connected component of \( D \).

The affine function \( a \) clearly satisfies \( L_b \cdot pa = 0 \) in the classical sense.

In order to justify (1.6), we will face three problems. The first is a justification of the comparison principle, the second is regularity of \( u \), and the third is a construction of suitable barrier subsolutions, all of which are essentially needed in the classical proof of E. Hopf’s strong maximum principle [20]. In order to overcome these obstacles, we appeal to both classical and distributional approaches, and restrict our analysis only over regular points. For details, see Section 1.2.

Even though our strong maximum principle is somewhat weakened in the sense that this holds only on each connected component of \( D \subset \Omega \), we are able to show the following Liouville-type theorem.

**Theorem 3.** (Liouville-type theorem) Let \( u \) be a convex weak solution to \( L_b \cdot pu = 0 \) in \( \mathbb{R}^n \). Then \( F \subset \mathbb{R}^n \), the facet of \( u \), satisfies either \( F = \emptyset \) or \( F = \mathbb{R}^n \). In particular, \( u \) satisfies either of the following:

1. If \( u \) attains its minimum in \( \mathbb{R}^n \), then \( u \) is constant.
2. If \( u \) does not attain its minimum in \( \mathbb{R}^n \), then \( u \) is a non-constant affine function in \( \mathbb{R}^n \).

In the proof of the Liouville-type theorem, our strong maximum principle plays an important role. Precisely speaking, if a convex solution in the total space does not satisfy \( \emptyset \subset F \subset \mathbb{R}^n \), then Theorem 2 and the supporting hyperplane theorem from convex analysis help us to determine the shape of convex solutions. In particular, the convex solution can be classified into three types of piecewise-linear functions of one-variable. These non-smooth piecewise-linear functions are, however, no longer weak solutions, which we will prove by some explicit calculations.

By applying the Liouville-type theorem and our blow-up argument, we are able to show that subgradients at points of the facet are always 0, that is, \( \partial u(x) = \{0\} \) for all \( x \in F \), and we complete the proof of the \( C^1 \)-regularity theorem. Note that the statements in Theorem 2 and 3 should not hold for unperturbed one-Laplace equation \( -\Delta_1 u = f \), since any absolutely continuous non-decreasing function of one variable \( u = u(x_1) \) satisfies \( -\Delta_1 u = 0 \).

Finally we mention that we are able to refine our strategy, and obtain \( C^1 \)-regularity of convex solutions to more general equations. We replace the one-Laplacian \( \Delta_1 \) by another operator which is derived from a general convex functional of degree 1. This generalization requires us to modify some of our arguments, including a blow-up argument and the Liouville-type theorem. For further details, see Sections 1.4 and 6.2.

**1.2. Literature overview on maximum principles**

We briefly introduce maximum principles related to the paper. We also describe our strategy for establishing the strong maximum principle.
Maximum principles, including comparison principles and strong maximum principles, have been discussed by many mathematicians in various settings. In the classical settings, E. Hopf proved a variety of maximum principles on elliptic partial differential equations of second order, by elementary arguments based on constructions of auxiliary functions. E. Hopf’s strong maximum principle is one of the well-known results on maximum principles. In Hopf’s proof of the strong maximum principle [20], he defined an auxiliary function
\[
h(x) := e^{-\alpha|x-x_\star|^2} - e^{-\alpha R^2} \quad \text{for } x \in \mathbb{R}^n,
\]
which becomes a classical subsolution in a fixed open annulus \(E_R = E_R(x_\star) := B_R(x_\star) \setminus B_{R/2}(x_\star)\) for sufficiently large \(\alpha > 0\). An alternative function
\[
h(x) := |x - x_\star|^{-\alpha} - R^{-\alpha} \quad \text{for } x \in \mathbb{R}^n \setminus \{x_\star\}
\]
is given in [29, Chapter 2.8]. E. Hopf’s classical results on maximum principles are extensively contained in [17, Chapter 3], [28, Chapter 2] and [29, Chapter 2].

The materials [17, Chapter 8–9] and [29, Chapter 3–6] provide proofs of maximum principles, including strong maximum principles, even for distributional solutions. Among them, [29, Theorem 5.4.1] deals with a justification of the strong maximum principle for distributional supersolutions to certain quasilinear elliptic equations with divergence structures,

\[
\text{i.e., } -\text{div}(A(x, \nabla u(x))) = 0,
\]
which covers the \(p\)-Laplace equation with \(1 < p < \infty\). Even in the distributional schemes, the proof of the maximum principle [29, Theorem 5.4.1] is partially similar to E. Hopf’s classical one, in the sense that it is completed by calculating directional derivatives of auxiliary functions. The significant difference is, however, the construction of spherically symmetric subsolutions of \(C^1\) class, which is given in [29, Chapter 4], is based on Leray–Schauder’s fixed point theorem [17, Theorem 11.6]. Also it should be noted that the proofs of comparison principles [29, Theorem 2.4.1 and 3.4.1] are just based on strict monotonicity of the mapping \(A(x, \cdot): \mathbb{R}^n \to \mathbb{R}^n\), whereas Hopf’s proof appeals to direct constructions of auxiliary functions.

With our literature overview in mind, we describe our strategy for showing (1.6). A justification of comparison principles is easily obtained in the distributional schemes (see [29, Chapter 3] as a related material). However, the remaining two obstacles, the differentiability of \(u\) and the construction of subsolutions, cannot be resolved affirmatively by just imitating arguments given in [29, Chapter 4–5]. In the first place, it should be mentioned that convex weak solutions we treat in this paper are assumed to have only local Lipschitz regularity, whereas supersolutions treated in [29, Chapter 5] are required to be in \(C^1\). We recall that \(C^1\)-regularity of convex weak solutions can be guaranteed in \(D \subset \Omega\) (the outside of the facet) by the classical De Giorgi–Nash–Moser theory, and this result enables us to overcome the problem whether \(u\) is differentiable at certain points. This is the reason why Theorem 2 need to restrict on \(D\). Although the construction of distributional subsolutions is generally discussed in [29, Chapter 4], we do not appeal to this. Instead, we directly construct
a function \( v = \beta h + a \) in \( \mathbb{R}^n \setminus \{x_0 \} \), where \( \beta > 0 \) is a constant and \( h \) is defined as in (1.7) or (1.8). We will determine the constants \( \alpha, \beta > 0 \) so precisely that \( v \) satisfies \( L_{b, p} v \leq 0 \) in the classical sense over a fixed open annulus \( E_R = E_R(x_*) \). We also make \( |\nabla v| \) very close to \( |\nabla a| \equiv |\nabla u(x_0)| > 0 \) over \( E_R \), so that \( \nabla v \) no longer degenerates there. By direct calculation of \( L_{b, p} v \), we explicitly construct classical subsolutions to \( L_{b, p} u = 0 \) in \( E_R \). Finally we are able to deduce (1.6) by an indirect proof.

Another type of definitions of subsolutions and supersolutions to (1.1) in the distributional schemes can be found in F. Krügel thesis in 2013 [25]. The significant difference is that Krügel did not regard the term \( \nabla u/|\nabla u| \) as a subgradient vector field. Since monotonicity of \( \partial |\cdot| \) is not used at all, it seems that Krügel’s proof of the comparison principle [25, Theorem 4.8] needs further explanation. For details, see Remark 3.

1.3. Mathematical models and previous researches

Our problem is derived from a minimizing problem of a certain energy functional, which involves the total variation energy. The Equation (1.1) is deduced from the following Euler–Lagrange equation;

\[
 f = \frac{\delta G}{\delta u}, \quad \text{where} \quad G(u) := b \int_{\Omega} |\nabla u| \, dx + \frac{1}{p} \int_{\Omega} |\nabla u|^p \, dx.
\]

The energy functional \( G \) often appears in fields of materials science and fluid mechanics.

In [31], Spohn modeled the relaxation dynamics of a crystal surface below the roughening temperature. On \( h \) describing the height of the crystal for a two-dimensional domain \( \Omega \) is modeled as

\[
 h_t + \text{div} \, j = 0
\]

with \( j = -\nabla \mu \), where \( \mu \) is a chemical potential. In [31], its evolution is given as

\[
 \mu = \frac{\delta \Phi}{\delta h} \quad \text{with} \quad \Phi(h) = \int_{\Omega} |\nabla h| \, dx + \kappa \int_{\Omega} |\nabla h|^3 \, dx
\]

with \( \kappa > 0 \). This \( \Phi \) is essentially the same as \( G \) with \( p = 3 \). Then, the resulting evolution equation for \( h \) is of the form

\[
 bh_t = \Delta L_{b, \kappa} h \quad \text{with} \quad b = \frac{1}{3\kappa}.
\]

This equation can be defined as a limit of step motion, which is microscopic in the direction of height [23]; see also [27]. The initial value problem of this equation can be solved based on the theory of maximal monotone operators [12] under the periodic boundary condition. Subdifferentials describing the evolution are characterized by Kashima [21,22]. Its evolution speed is calculated by [21] for one dimensional setting and by [22] for radial setting. It is known that the solution...
stops in finite time [13,14]. In [27], numerical calculation based on step motion is calculated. If one considers a stationary solution, $h$ satisfies

$$\Delta L_{b, 3}h = 0.$$ 

If $L_{b, 3}h$ is a constant, our Theorem 1 implies that the height function $h$ is $C^1$ provided that $h$ is convex.

For a second order problem,

$$i.e., bh_t = L_{b, p}h,$$

its analytic formulation goes back to [4], [8, Chapter VI] for $p = 2$, and its numerical analysis is given in [19]. For the fourth order problem, its numerical study is more recent. The reader is referred to papers by [15,16,24].

Another important mathematical model for the equation (1.1) is found in fluid mechanics. Especially for $p = 2$ and $n = 2$, the energy functional $G$ appears when modeling stationary laminar incompressible flows of a material called Bingham fluid, which is a typical non Newtonian fluid. Bingham fluid reflects the effect of plasticity corresponding to $\Delta_1u$ as well as that of viscosity corresponding $\Delta_2u = \Delta u$ in (1.1). Let us consider a parallel stationary flow with velocity $U = (0, 0, u(x_1, x_2))$ in a cylinder $\Omega \times \mathbb{R}$. Of course, this is incompressible flow, that is, $\text{div } U = 0$. If this flow is the classical Newtonian fluid, then the Navier–Stokes equations become (1.1) in $\Omega$ with $b = 0$ and $f = -\partial_3 \pi$, where $\pi$ denotes the pressure. In the case that plasticity effects appears, one obtains (1.1), following [8, Chapter VI, Section 1]. There it is also mentioned that since the velocity is assumed to be uni-directional, the external force term in (1.1) is considered as constant in this laminar flow model. The significant difference is that motion of the Bingham fluid is blocked if the stress of the Bingham fluid exceeds a certain threshold. This physical phenomenon is essentially explained by the nonlinear term $b\Delta_1u$, which reflects rigidity of the Bingham fluid. For more details, see [8, Chapter VI] and the references therein.

On continuity of derivatives for solutions, less is known even for the second order elliptic case. Although Krügel gave an observation that solutions can be continuously differentiable [25, Theorem 1.2] on the boundary of a facet, mathematical justifications of $C^1$-regularity have not been well-understood. Our main result (Theorem 1) mathematically establishes continuity of gradient for convex solutions.

### 1.4. Organization of the paper

We outline the contents of the paper.

Section 2 establishes $C^{1, \alpha}$-regularity at regular points of convex weak solutions (Lemma 1). In order to apply De Giorgi–Nash–Moser theory, we will need to justify local $W^{2, 2}$-regularity by the difference quotient method. The key lemma, which is proved by convex analysis, is contained in the appendices (Lemma 8).

Section 3 provides a blow-up argument for convex weak solutions. The aim of Section 3 is to prove that $u_0: \mathbb{R}^n \to \mathbb{R}$, a limit of rescaled solutions, satisfies
$L_{b, p} u_0 = 0$ in the weak sense over the whole space $\mathbb{R}^n$ (Proposition 1). To assure this, we will make use of an elementary result on almost everywhere convergence of gradients, which is given in the appendices (Lemma 9).

Section 4 is devoted to justifications of maximum principles for the equation $L_{b, p} u = 0$. We first give definitions of sub- and supersolution in the weak sense. Section 4.1 provides a justification of the comparison principle (Proposition 2). Section 4.2 establishes an existence result of classical barrier subsolutions in an open annulus (Lemma 2). Applying these results in Section 4.1–4.2, we prove the strong maximum principle outside the facet (Theorem 2).

In Section 5, we will show the Liouville-type theorem (Theorem 3) by making use of Theorem 2, and complete the proof of our main theorem (Theorem 1).

Finally in Section 6, we discuss a few generalization of the operators $\Delta_1$ and $\Delta_p$. Since the general strategy for the proof is the same, we only indicate modification of our arguments. Among them, we especially treat with a Liouville-type theorem and a blow-up argument, since these proofs require basic facts of a general convex functional which is positively homogeneous of degree 1. These well-known facts are contained in the appendices for completeness.

2. Regularity Outside the Facet

In Section 2, we would like to show that $u$ is $C^1$ at any $x \in D$, and therefore (1.4) holds for all $x \in D$. This result will be used in the proof of the strong maximum principle (Theorem 2).

We first give a precise definition of weak solutions to $L_{b, p} u = f$ in a convex domain $\Omega \subset \mathbb{R}^n$, which is not necessarily bounded.

**Definition 1.** Let $\Omega \subset \mathbb{R}^n$ be a domain, which is not necessarily bounded, and $f \in L^q_{\text{loc}}(\Omega) (n < q \leq \infty)$. We say that a function $u \in W^{1, p}_{\text{loc}}(\Omega)$ is a weak solution to (1.1), when for any bounded Lipschitz domain $\omega \subset \Omega$, there exists a vector field $Z \in L^\infty(\omega, \mathbb{R}^n)$ such that the pair $(u, Z) \in W^{1, p}_{0}(\omega) \times L^\infty(\omega, \mathbb{R}^n)$ satisfies

$$b \int_\omega \langle Z \mid \nabla \phi \rangle \, dx + \int_\omega \left( |\nabla u|^{p-2} \nabla u \mid \nabla \phi \right) \, dx = \int_\omega f \phi \, dx$$

(2.1)

for all $\phi \in W^{1, p}_{0}(\omega)$, and

$$Z(x) \in \partial| \cdot |(\nabla u(x))$$

(2.2)

for almost everywhere $x \in \omega$. For such pair $(u, Z)$, we say that $(u, Z)$ satisfies $L_{b, p} u = f$ in $W^{-1, p'}(\omega)$ or simply say that $u$ satisfies $L_{b, p} u = f$ in $W^{-1, p'}(\omega)$. Here $p' \in (1, \infty)$ denotes the Hölder conjugate exponent of $p \in (1, \infty)$.

The aim of Section 2 is to show Lemma 1, below.
Lemma 1. Let $u$ be a convex weak solution to (1.1) in a convex domain $\Omega \subset \mathbb{R}^n$, and $f \in L^q_{\text{loc}}(\Omega)$ ($n < q \leq \infty$). If $x_0 \in D$, then we can take a small radius $r_0 > 0$, a unit vector $v_0 \in \mathbb{R}^n$, and a small number $\mu_0 > 0$ such that

$$B_{r_0}(x_0) \subset D \text{ and } \langle \nabla u(x) \mid v_0 \rangle \geq \mu_0 \text{ for almost everywhere } x \in B_{r_0}(x_0).$$

and there exists a small number $\alpha = \alpha(\mu_0) \in (0, 1)$ such that $u \in C^{1, \alpha}(B_{2r_0/2}(x_0))$. In particular, $u$ is $C^1$ in $D$, and $\partial u(x) = \{\nabla u(x)\} \neq [0]$ for all $x \in D$.

Before proving Lemma 1, we introduce difference quotients. For given $g : \Omega \to \mathbb{R}^m$ ($m \in \mathbb{N}$), $j \in \{1, \ldots, n\}$, $h \in \mathbb{R} \setminus \{0\}$, we define

$$\Delta_j, h g(x) := \frac{g(x + he_j) - g(x)}{h} \in \mathbb{R}^m \text{ for } x \in \Omega \text{ with } x + he_j \in \Omega,$$

where $e_j \in \mathbb{R}^n$ denotes the unit vector in the direction of the $x_j$-axis.

In the proof of Lemma 1, we will use Lemmas 7–8 without proofs. For precise proofs, see Section A.

Proof. For each fixed $x_0 \in D$, we may take and fix $x_1 \in \Omega$ such that $u(x_0) > u(x_1)$. We set $3\delta_0 := u(x_0) - u(x_1) > 0$, $d_0 := |x_0 - x_1| > 0$ and $v_0 := d_0^{-1}(x_0 - x_1)$. By $u \in C(\Omega)$, we may take a sufficiently small $r_0 > 0$ such that

$$u(y_0) - u(y_1) \geq \delta_0 > 0 \text{ for all } y_0 \in B_{r_0}(x_0), y_1 \in B_{r_0}(x_1). \quad (2.4)$$

From (2.4), the inclusion $B_{3r_0}(x_0) \subset D$ clearly holds. (2.4) also allows us to check that for all $y_0 \in B_{r_0}(x_0)$, $z_0 \in \partial u(y_0)$,

$$\langle z_0 \mid v_0 \rangle \geq \frac{u(y_0) - u(y_0 - d_0 v_0)}{d_0} \geq \frac{\delta_0}{d_0} =: \mu_0 > 0. \quad (2.5)$$

For the first inequality in (2.5), we have used Lemma 8, which is basically derived from convexity of $u$. Recall that $\partial u(x) = \{\nabla u(x)\}$ for almost everywhere $x \in \Omega$, and hence we are able to recover (2.3) from (2.5).

In order to obtain $C^1$-regularity in $D$, we will appeal to the classical De Giorgi–Nash–Moser theory. For preliminaries, we check that the operator $L_p \mu u$ assures uniform ellipticity in $B_{r_0}(x_0)$. Local Lipschitz continuity of $u$ implies that there exists a sufficiently large number $M_0 \in (0, \infty)$ such that

$$\text{ess sup}_{B_{r_0}(x_0)} |\nabla u| \leq M_0 \text{ and } |u(x) - u(y)| \leq M_0|x - y| \text{ for all } x, y \in B_{r_0}(x_0). \quad (2.6)$$

For notational simplicity, we write subdomains by

$$U_1 := B_{r_0}(x_0) \supseteq U_2 := B_{15r_0/16}(x_0) \supseteq U_3 := B_{7r_0/8}(x_0) \supseteq U_4 := B_{3r_0/4}(x_0) \supseteq U_5 := B_{r_0/2}(x_0).$$
It should be noted that \( E(z) := b|z| + |z|^p / p \) (\( z \in \mathbb{R}^n \)) satisfies \( E \in C^\infty(\mathbb{R}^n \setminus \{0\}) \), and there exists two constants \( 0 < \lambda(p, \mu_0, M_0) \leq \Lambda(b, p, \mu_0, M_0) < \infty \) such that

\[
\lambda|\xi|^2 \leq \left\langle \nabla_z^2 E(z_0)\xi, \xi \right\rangle \tag{2.7}
\]

\[
\left\langle \nabla_z^2 E(z_0)\xi, \omega \right\rangle \leq \Lambda|\xi||\omega| \tag{2.8}
\]

for all \( z_0, \xi, \omega \in \mathbb{R}^n \) with \( \mu_0 \leq |z_0| \leq M_0 \). We can explicitly determine \( 0 < \lambda \leq \Lambda < \infty \) by

\[
\begin{align*}
\lambda(p, \mu_0, M_0) &:= \min_{\mu_0 \leq t \leq M_0} \left( \min\{1, p - 1\}t^{p-2}\right), \\
\Lambda(b, p, \mu_0, M_0) &:= \max_{\mu_0 \leq t \leq M_0} \left( bt^{-1} + \max\{1, p - 1\}t^{p-2}\right)
\end{align*}
\]

Now we check that \( u \in W^{2,2}(U_4) \) by the difference quotient method. We refer the reader to [18, Theorem 8.1] as a related result. By [18, Lemma 8.2], it suffices to check that

\[
\sup \left\{ \int_{U_4} |\nabla(\Delta_j h u)|^2 \, dx \mid h \in \mathbb{R}, \, 0 < |h| < \frac{r_0}{16} \right\} < \infty
\]
for each \( j \in \{1, \ldots, n\} \).

Since \( u \in W^{1, p}(U_1) \) satisfies \( L_b, pu = f \) in \( W^{-1, p'}(U_1) \), we obtain

\[
\int_{U_1} \left\langle \nabla_z E(\nabla u) \mid \nabla \phi \right\rangle \, dx = \int_{U_1} f \phi \, dx \tag{2.10}
\]

for all \( \phi \in W^{1, p}_0(U_1) \). Here we note that \( \nabla u \) no longer degenerates in \( U_1 \) by (2.3). We fix a cutoff function \( \eta \in C^1(U_3) \) such that

\[
0 \leq \eta \leq 1 \text{ in } U_3, \, \eta \equiv 1 \text{ in } U_4, \, |\nabla \eta| \leq \frac{c}{r_0} \tag{2.11}
\]

for some constant \( c > 0 \). For each fixed \( j \in \{1, \ldots, n\}, h \in \mathbb{R} \) with \( 0 < |h| < r_0/16 \), we test \( \phi := \Delta_j, -h(\eta^2 \Delta_j, h u) \) into (2.10). We note that \( \phi \in W^{1, \infty}(U_1) \subset W^{1, p}(U_1) \) by (2.8), and this is compactly supported in \( U_2 \). Hence \( \phi \in W^{1, p}_0(U_2) \) is an admissible test function. By testing \( \phi \), we have

\[
0 = \int_{U_2} \left\{ \Delta_j, h(\nabla_z E(\nabla u(x))) \right\} \eta^2 \nabla(\Delta_j, h u) + 2\eta \Delta_j, h u \nabla \eta \right\} \right.
\]

\[
- \int_{U_2} f \Delta_j, h(\eta^2 \Delta_j, h u) \, dx
\]

\[
= \int_{U_2} \eta^2 \left\{ A_h(x, \nabla u(x)) \nabla(\Delta_j, h u) \right\} \nabla(\Delta_j, h u) \, dx
\]

\[
+ 2 \int_{U_2} \eta \Delta_j, h u \left\{ A_h(x, \nabla u(x)) \nabla(\Delta_j, h u) \right\} \nabla \eta \, dx
\]

\[
- \int_{U_2} f \Delta_j, h(\eta^2 \Delta_j, h u) \, dx
\]

\[
=: I_1 + I_2 - I_3. \tag{2.12}
\]
Here \( A_h = A_h(x, \nabla u(x)) \) denotes a matrix-valued function in \( U_2 \) given by
\[
A_h(x, \nabla u(x)) := \int_0^1 \nabla_x^2 E((1 - t)\nabla u(x) + t\nabla u(x + he_j)) \, dt.
\]

We note that with the aid of (2.3)–(2.6), we obtain
\[
\mu_0 \leq |(1 - t)\nabla u(x) + t\nabla u(x + he_j)| \leq M_0
\]
for almost everywhere \( x \in U_2 \) and for all \( 0 \leq t \leq 1 \). Combining this result with (2.7)–(2.8), we conclude that \( A_h \) satisfies
\[
\lambda|\xi|^2 \leq \langle A_h(x, \nabla u(x))\xi \rangle \quad \text{(2.13)}
\]
\[
\langle A_h(x, \nabla u(x))\xi , \omega \rangle \leq \Lambda|\xi||\omega| \quad \text{(2.14)}
\]
for all \( \xi, \omega \in \mathbb{R}^n \) and for almost everywhere \( x \in U_2 \). We set an integral
\[
J := \int_{U_2} \eta^2|\nabla(\Delta_j,h u)|^2 \, dx.
\]
By (2.13), it is clear that \( I_1 \geq \lambda J \). By Young’s inequality and applying a Poincaré-type inequality (Lemma 7) to \( \eta^2 \Delta_j,h u \in W_0^{1,2}(U_2) \), we obtain for any \( \varepsilon > 0 \),
\[
|I_3| \leq \frac{1}{4\varepsilon} \|f\|^2_{L^2(U_2)} + \varepsilon \int_{U_2} |\nabla(\eta^2 \Delta_j,h u)|^2 \, dx
\]
\[
\leq \frac{1}{4\varepsilon} \|f\|^2_{L^2(U_2)} + 4\varepsilon \int_{U_2} |\Delta_j,h u|^2|\nabla\eta|^2 \, dx + 2\varepsilon \int_{U_2} \eta^2|\nabla(\Delta_j,h u)|^2 \, dx.
\]
Here we have invoked the property \( 0 \leq \eta \leq 1 \) in \( U_2 \). We fix \( \varepsilon := \lambda/6 > 0 \). By (2.14) and Young’s inequality, we have
\[
|I_2| \leq 2\Lambda \int_{U_2} \eta |\nabla(\Delta_j,h u)| \cdot |\Delta_j,h u| |\nabla\eta| \, dx
\]
\[
\leq \frac{\lambda}{3} J + \frac{3\Lambda^2}{\lambda} \int_{U_2} |\Delta_j,h u|^2|\nabla\eta|^2 \, dx.
\]
It follows from (2.6) that \( \|\Delta_j,h u\|_{L^\infty(U_2)} \leq M_0 \). Therefore we obtain from (2.12),
\[
\int_{U_4} |\nabla(\Delta_j,h u)|^2 \, dx \leq J = \int_{U_2} \eta^2|\nabla(\Delta_j,h u)|^2 \, dx
\]
\[
\leq C(\lambda, \Lambda) \left( M_0^2 \|\nabla\eta\|^2_{L^2(U_2)} + \|f\|^2_{L^2(U_2)} \right).
\]
The estimate (2.9) follows from this, and therefore \( u \in W^{2,2}(U_4) \).

For each \( \psi \in C_c^\infty(U_4) \), we test \( \partial_{x_j}\psi \in C_c^\infty(U_4) \) into (2.10). Integrating by parts, we obtain
\[
\int_{U_4} \langle \nabla_x^2 E(\nabla u)\nabla \partial_{x_j} u \mid \nabla \psi \rangle \, dx = -\int_{U_4} f \partial_{x_j} \psi \, dx \quad \text{(2.15)}
\]
for all $\psi \in C_c^\infty(U_4)$. Noting that $f \in L^q(U_4) \subset L^2(U_4)$, $\partial_x j u \in W^{1,2}(U_4)$, and (2.7)–(2.8), we may extend $\psi \in W^{1,2}_0(U_4)$ by a density argument. The conditions (2.7)–(2.8) imply that $\nabla_x^2 E(\nabla u)$ is uniformly elliptic over $U_1$. Hence by [17, Theorem 8.22], there exists $\alpha = \alpha(\lambda, \Lambda, n, q) \in (0, 1)$ such that $\partial_x j u \in C^\alpha(U_3)$ for each $j \in \{1, \ldots, n\}$. This regularity result implies $\partial u(x) = \{\nabla u(x)\} \neq \{0\}$ for all $x \in D$. □

3. A Blow-Up Argument

In order to show that (1.4) holds true even for $x \in F$, we first make a blow-up argument and construct a convex weak solution in the whole space $\mathbb{R}^n$, in the sense of Definition 1.

Proposition 1. Let $\Omega \subset \mathbb{R}^n$ be a convex domain, and $f \in L^q_{\text{loc}}(\Omega) (n < q \leq \infty)$. Assume that $u$ is a convex weak solution to (1.1), and $x_0 \in \Omega$. Then there exists a convex function $u_0 : \mathbb{R}^n \to \mathbb{R}$ such that

1. $u_0$ is a weak solution to $L_{b, p} u_0 = 0$ in $\mathbb{R}^n$.
2. The inclusion $\partial u(x_0) \subset \partial u_0(x_0)$ holds. That is, if $c \in \partial u(x_0)$, then we have
   \[ u_0(x) \geq u_0(x_0) + \langle c \mid x - x_0 \rangle \text{ for all } x \in \mathbb{R}^n. \]

   In particular, if $x_0 \in F$, then the facet of $u_0$ is non-empty.

Proof. Without loss of generality, we may assume that $x_0 = 0$ and $u(x_0) = 0$. First we fix a closed ball $B_R(0) = B_R \subset \Omega$. We note that $u \in \text{Lip}(B_R)$ since $u$ is convex. Hence there exists a sufficiently large number $M \in (0, \infty)$ such that

\[ \text{ess sup}_{B_R} |\nabla u| \leq M \text{ and } |u(x) - u(y)| \leq M|y - y| \text{ for all } x, y \in B_R. \]

We take and fix a vector field $Z \in L^\infty(B_R, \mathbb{R}^n)$ such that the pair $(u, Z) \in W^{1, p}(B_R) \times L^\infty(B_R, \mathbb{R}^n)$ satisfies $L_{b, p} u = f$ in $W^{-1, p'}(B_R)$. For each $a > 0$, we define a rescaled convex function $u_a : B_{R/a} \to \mathbb{R}$ and a dilated vector field $Z_a \in L^\infty(B_{R/a}, \mathbb{R}^n)$ by

\[ u_a(x) := \frac{u(ax)}{a}, \quad Z_a(x) := Z(ax) \text{ for } x \in B_{R/a}. \]

We also set $f_a \in L^q(B_{R/a})$ by

\[ f_a(x) := af(ax) \text{ for } x \in B_{R/a}. \]

Then it is easy to check that the pair $(u_a, Z_a) \in W^{1, \infty}(B_{R/a}) \times L^\infty(B_{R/a}, \mathbb{R}^n)$ satisfies $L_{b, p} u_a = f_a$ in $W^{-1, p'}(B_{R/a})$. For each fixed $R < r < \infty$, the inclusion $B_r = B_r(0) \subset B_{R/a}$ holds for all $a \in (0, R/r)$. We also have

\[ \sup_{B_r} |u_a| \leq \frac{r}{R} M < \infty, \quad \|\nabla u_a\|_{L^\infty(B_r)} \leq M < \infty \text{ for all } a \in (0, R/r),(3.1) \]
by the definition of $u_a$. Hence by the Arzelà–Ascoli theorem and a diagonal argument, we can take a decreasing sequence $\{a_N\}_{N=1}^\infty \subset (0, \infty)$, such that $a_N \to 0$ as $N \to \infty$, and

$$u_{a_N} \to u_0 \quad \text{locally uniformly in } \mathbb{R}^n. \quad (3.2)$$

for some function $u_0: \mathbb{R}^n \to \mathbb{R}$. Clearly $u_0$ is convex in $\mathbb{R}^n$, and the inclusion $\partial u(x_0) \subset \partial u_0(x_0)$ holds true by the construction of rescaled functions $u_a$. If $x_0 \in F$, then we have $\{0\} \subset \partial u(x_0) \subset \partial u_0(x_0)$ and therefore $x_0$ lies in the facet of $u_0$. We are left to show that $u_0$ is a weak solution to $L_b, pu_0 = 0$ in $\mathbb{R}^n$. Before proving this, we note that from (3.1)–(3.2) and Lemma 9, it follows that

$$\nabla u_{a_N}(x) \to \nabla u_0(x) \text{ and } |\nabla u_0(x)| \leq M \quad \text{for almost everywhere } x \in \mathbb{R}^n \quad (3.3)$$

as $N \to \infty$. We arbitrarily fix an open ball $B_r = B_r(0) \subset \mathbb{R}^n$. We easily realize that a family of pairs $\{(u_a, Z_a)\}_{0 < a < R/r} \subset W^{1, \infty}(B_r) \times L^\infty(B_r, \mathbb{R}^n)$ satisfies

$$Z_a(x) \in \partial | \cdot |(\nabla u_a(x)) \quad \text{for almost everywhere } x \in B_r, \quad (3.4)$$

$$b \int_{B_r} (Z_a | \nabla \phi) \, dx + \int_{B_r} \left(|\nabla u_a|^{p-2}\nabla u_a \right) | \nabla \phi \right) \, dx$$

$$= \int_{B_r} f_a \phi \, dx \quad \text{for all } \phi \in W^{1, p}(B_r). \quad (3.5)$$

By definition of $f_a$, we get $\|f_a\|_{L^q(B_r)} = a^{1-n/q} \|f\|_{L^q(B_{ar})} \leq a^{1-n/q} \|f\|_{L^q(B_R)}$ for all $0 < a < R/r$. Hence by the continuous embedding $L^q(B_r) \hookrightarrow W^{-1, p'}(B_r)$, we obtain

$$f_{a_N} \to 0 \quad \text{in } W^{-1, p'}(B_r) \quad \text{as } N \to \infty. \quad (3.6)$$

By (3.1) and (3.3), we can apply Lebesgue’s dominated convergence theorem and get

$$|\nabla u_{a_N}|^{p-2}\nabla u_{a_N} \to |\nabla u_0|^{p-2}\nabla u_0 \quad \text{in } L^{p'}(B_r, \mathbb{R}^n) \quad \text{as } N \to \infty. \quad (3.7)$$

It is clear that $\|Z_a\|_{L^\infty(B_r, \mathbb{R}^n)} \leq 1$ for all $0 < a < R/r$. Hence by [5, Corollary 3.30], up to a subsequence, we may assume that

$$Z_{a_N} \overset{\ast}{\rightharpoonup} Z_{0, r} \quad \text{in } L^\infty(B_r, \mathbb{R}^n) \quad \text{as } N \to \infty \quad (3.8)$$

for some $Z_{0, r} \in L^\infty(B_r, \mathbb{R}^n)$. By lower-semicontinuity of the norm with respect to the weak* topology and (3.3)–(3.4), we get

$$\|Z_{0, r}\|_{L^\infty(B_r, \mathbb{R}^n)} \leq 1, \quad Z_{0, r}(x) = \frac{\nabla u_0(x)}{|\nabla u_0(x)|}$$

for almost everywhere $x \in B_r$ with $\nabla u_0(x) \not= 0$, which implies that

$$Z_{0, r}(x) \in \partial | \cdot |(\nabla u_0(x)) \quad \text{for almost everywhere } x \in B_r. \quad (3.9)$$
Letting \( a = a_N \) in (3.5) and \( N \to \infty \), we obtain
\[
b \int_{B_r} \langle Z_0, r \mid \nabla \phi \rangle \, dx + \int_{B_r} \left| \nabla u_0 \right|^{p-2} \nabla u_0 \cdot \nabla \phi \, dx = 0 \quad \text{for all } \phi \in W_0^{1,p}(B_r)
\]
(3.10)

by (3.5)–(3.8). Since \( B_r \subset \mathbb{R}^n \) is arbitrary, (3.9)–(3.10) means that \( u_0 \) is a weak solution to \( L_{b,p} u_0 = 0 \) in \( \mathbb{R}^n \), in the sense of Definition 1.

\[\square\]

4. Maximum Principles

In Section 4, we justify maximum principles for the equation \( L_{b,p} u = 0 \).

We first define subsolutions and supersolutions in the weak sense.

**Definition 2.** Let \( \Omega \subset \mathbb{R}^n \) be a bounded domain. A pair \( (u, Z) \in W^{1,p}(\Omega) \times L^\infty(\Omega, \mathbb{R}^n) \) is called a *weak* subsolution to \( L_{b,p} u = 0 \) in \( \Omega \), if it satisfies
\[
b \int_\Omega \langle Z \mid \nabla \phi \rangle \, dx + \int_\Omega \left| \nabla u \right|^{p-2} \nabla u \cdot \nabla \phi \, dx \leq 0 \quad \text{for all } 0 \leq \phi \in C^\infty_c(\Omega), \text{ and}
\]
\[Z(x) \in \partial | \cdot | (\nabla u(x)) \quad \text{for almost everywhere } x \in \Omega.\]
(4.1)

Similarly we call a pair \( (u, Z) \in W^{1,p}(\Omega) \times L^\infty(\Omega, \mathbb{R}^n) \) a *weak* supersolution \( L_{b,p} u = 0 \) in \( \Omega \), if it satisfies (4.2) and
\[
b \int_\Omega \langle Z \mid \nabla \phi \rangle \, dx + \int_\Omega \left| \nabla u \right|^{p-2} \nabla u \cdot \nabla \phi \, dx \geq 0 \quad \text{for all } 0 \leq \phi \in C^\infty_c(\Omega). \]
(4.3)

For \( u \in W^{1,p}(\Omega) \), we simply say that \( u \) is respectively a subsolution and a supersolution to \( L_{b,p} u = 0 \) in \( \Omega \), if there is \( Z \in L^\infty(\Omega, \mathbb{R}^n) \) such that the pair \( (u, Z) \) is a weak subsolution and a weak supersolution to \( L_{b,p} u = 0 \) in \( \Omega \).

**Remark 2.** We describe some remarks on our definitions of weak solutions, subsolutions and supersolutions.

1. By an approximation argument, we may extend the test function class of (4.1) to
\[D_+(\Omega) := \{ \phi \in W^{1,p}(\Omega) \mid \phi \geq 0 \text{ almost everywhere in } \Omega, \text{ supp } \phi \subset \Omega \}.\]

Indeed, for \( \phi \in D_+(\Omega) \) and \( 0 < \varepsilon < \text{dist(supp } \phi, \partial \Omega) \), the function,
\[\phi_\varepsilon(x) = \int_\Omega \phi(x - y) \rho_\varepsilon(y) \, dy \quad \text{for } x \in \Omega\]
satisfies \( 0 \leq \phi_\varepsilon \in C^\infty_c(\Omega) \). Here for \( 0 < \varepsilon < \infty \), \( 0 \leq \rho_\varepsilon \in C^\infty_c(B_\varepsilon(0)) \) denotes a standard mollifier so that
\[0 \leq \rho \in C^\infty_c(B_1), \quad \| \rho \|_{L^1(\mathbb{R}^n)} = 1, \quad \rho_\varepsilon(x) := \varepsilon^{-n} \rho(x/\varepsilon) \text{ for } x \in \mathbb{R}^n.\]
By testing $\phi_\varepsilon$ into (4.1) for sufficiently small $\varepsilon > 0$ and letting $\varepsilon \to 0$, we conclude that if the pair $(u, Z)$ satisfies (4.1) for all $0 \leq \phi \in C^\infty_c(\Omega)$, then (4.1) holds for all $\phi \in D_+(\Omega)$. A similar result is also valid for (4.3).

2. By Definition 1–2, if a pair $(u, Z) \in W^{1, p}(\Omega) \times L^\infty(\Omega, \mathbb{R}^n)$ satisfies $L_{b, p}u = 0$ in $W^{-1, p'}(\Omega)$, then $u$ is clearly both a subsolution and a supersolution to $L_{b, p}u = 0$ in $\Omega$ in the weak sense. Conversely, if a pair $(u, Z) \in W^{1, p}(\Omega) \times L^\infty(\Omega, \mathbb{R}^n)$ is both a weak subsolution and a weak supersolution to $L_{b, p}u = 0$ in $\Omega$, then the pair $(u, Z)$ satisfies $L_{b, p}u = 0$ in $W^{-1, p'}(\Omega)$. Indeed, by the previous remark we have already known that the pair $(u, Z)$ satisfies (4.1) and (4.3) for all $\phi \in D_+(\Omega)$, which clearly yields

$$b \int_\Omega (Z | \nabla \phi) \, dx + \int_\Omega \left( |\nabla u|^{p-2} \nabla u | \nabla \phi \right) \, dx = 0$$

(4.4)

for all $\phi \in D_+(\Omega)$. We decompose arbitrary $\phi \in C^\infty_c(\Omega)$ by $\phi = \phi_+ - \phi_-$, where $\phi_+ := \max\{\phi, 0\}$, $\phi_- := \max\{-\phi, 0\} \in D_+(\Omega)$. By testing $\phi_+, \phi_- \in D_+(\Omega)$ into (4.4), we conclude that (4.4) holds for all $\phi \in C^\infty_c(\Omega)$.

By density of $C^\infty_c(\Omega) \subset W^{1, p}_0(\Omega)$, it is clear that (4.4) is valid for all $\phi \in W^{1, p}_0(\Omega)$.

3. For a bounded domain $\Omega \subset \mathbb{R}^n$, let $u \in C^2(\overline{\Omega})$ satisfy the following two conditions (4.5)–(4.6);

$$\nabla u(x) \neq 0 \quad \text{for all } x \in \Omega, \quad (L_{b, p}u)(x) = -(b \Delta_1 u + \Delta_p u)(x) \leq 0 \quad \text{for all } x \in \Omega.$$  

(4.5)  (4.6)

Then for any fixed $0 \leq \phi \in C^\infty_c(\Omega)$, we have

$$0 \geq \int_\Omega (L_{b, p}u) \phi \, dx = b \int_\Omega \left( \frac{\nabla u}{|\nabla u|} \right) \nabla \phi \, dx + \int_\Omega \left( |\nabla u|^{p-2} \nabla u | \nabla \phi \right) \, dx,$$

with the aid of integration by parts and (4.6). We also note that

$$\partial |\cdot| (\nabla u(x)) = \left\{ \frac{\nabla u(x)}{|\nabla u(x)|} \right\} \quad \text{for all } x \in \Omega$$

by (4.5). Therefore the pair $(u, \nabla u/|\nabla u|) \in W^{1, p}(\Omega) \times L^\infty(\Omega, \mathbb{R}^n)$ satisfies (4.1)–(4.2). For such $u$, we simply say that $u$ satisfies $L_{b, p}u \leq 0$ in $\Omega$ in the classical sense.

### 4.1. Comparison principle

We justify the comparison principle, that is, for any subsolution $u^-$ and supersolution $u^+$,

$$u^- \leq u^+ \quad \text{on } \partial \Omega \quad \text{implies that } \quad u^- \leq u^+ \quad \text{in } \Omega,$$

under the condition that $u^+$ and $u^-$ admits continuity properties in $\overline{\Omega}$. 
Proposition 2. Let $\Omega \subset \mathbb{R}^n$ be a bounded domain. Assume that $u^+, u^- \in C(\overline{\Omega}) \cap W^{1,p}(\Omega)$ is a subsolution and a supersolution to $L_b, pu = 0$ in the weak sense respectively. If $u^+, u^-$ satisfies
\[ u^-(x) \leq u^+(x) \text{ for all } x \in \partial \Omega, \tag{4.7} \]
then $u^- \leq u^+$ in $\overline{\Omega}$.

Before proving Proposition 2, we recall that the mapping $A: \mathbb{R}^n \ni z \mapsto |z|^{p-2}z \in \mathbb{R}^n$ satisfies strict monotonicity,
\[ \langle A(z_2) - A(z_1) \mid z_2 - z_1 \rangle > 0 \text{ for all } z_1, z_2 \in \mathbb{R}^n \text{ with } z_1 \neq z_2. \tag{4.8} \]

Proof. We take arbitrary $\delta > 0$. By $u^+, u^- \in C(\overline{\Omega})$ and (4.7), we can take a subdomain $\Omega' \subset \Omega$ such that $u^- \leq u^+ + \delta$ in $\Omega \setminus \Omega'$. This implies that the support of the truncated non-negative function $w_\delta := (u^+ - u^- + \delta)_- \in W^{1,p}(\Omega)$ is contained in $\overline{\Omega'} \subset \Omega$ and therefore $w_\delta \in D_+(\Omega)$. Let $Z^+, Z^- \in L^\infty(\Omega, \mathbb{R}^n)$ be vector fields such that $(u^+, Z^+), (u^-, Z^-)$ satisfies (4.1)–(4.2), (4.2)–(4.3) respectively. As in Remark 2, we may test $w_\delta$ in (4.1) and (4.3). Note that $\nabla w_\delta = -\chi_\delta \nabla (u^+ - u^-)$, where $\chi_\delta$ denotes the characteristic function of $A_\delta := \{x \in \Omega \mid u^+ + \delta \leq u^-\}$. Hence, we have
\[
0 \leq -b \int_{A_\delta} \left( Z^+ - Z^- \bigm| \nabla u^+ - \nabla u^- \right) dx \\
- \int_{A_\delta} \left( |\nabla u^+|^{p-2} \nabla u^+ - |\nabla u^-|^{p-2} \nabla u^- \bigm| \nabla u^+ - \nabla u^- \right) dx \\
\leq - \int_{A_\delta} \left( |\nabla u^+|^{p-2} \nabla u^+ - |\nabla u^-|^{p-2} \nabla u^- \bigm| \nabla u^+ - \nabla u^- \right) dx.
\]

Here we have invoked (4.2) and monotonicity of the subdifferential operator $\partial | \cdot |$. From (4.8) we can easily check that $\nabla u^+ = \nabla u^-$ in $A_\delta$, and therefore $w_\delta = 0$ in $W^{1,p}_0(\Omega)$. This means that $u^- \leq u^+ + \delta$ almost everywhere in $\Omega$. By regularity assumptions $u^+, u^- \in C(\overline{\Omega})$, we conclude that $u^- \leq u^+ + \delta$ in $\overline{\Omega}$. Since $\delta > 0$ is arbitrary, this completes the proof. \hfill \qed

Remark 3. In 2013, Krügel gave another type of definitions of weak subsolutions and weak supersolutions to $L_{b,p} = a$, where $a \in \mathbb{R}$ is a constant. In Krügel’s definition [25, Definition 4.6], a function $u^- \in W^{1,p}(\Omega)$ is called a subsolution to $L_{b,p} = a$ if $u^-$ satisfies
\[
\int_{D^-} \left( \frac{\nabla u^-}{|\nabla u^-|} \bigm| \nabla \phi \right) dx + \int_{F^-} |\nabla \phi| dx \\
+ \int_{\Omega} \left( |\nabla u^-|^{p-2} \nabla u^- \bigm| \nabla \phi \right) dx \leq \int_{\Omega} a \phi dx \tag{4.9}
\]
for all $\phi \in D_+(\Omega)$. Here $F^- := \{x \in \Omega \mid \nabla u^-(x) = 0\}, D^- := \Omega \setminus F^-$. Similarly a function $u^+ \in W^{1,p}(\Omega)$ is called a supersolution to $L_{b,p} = a$ if $u^+$
satisfies
\[
\int_{D^+} \left( \frac{\nabla u^+}{|\nabla u^+|} \cdot \nabla \phi \right) \, dx + \int_{F^+} |\nabla \phi| \, dx 
+ \int_{\Omega} \left( |\nabla u^+|^{p-2} \nabla u^+ \cdot \nabla \phi \right) \, dx \geq \int_{\Omega} a \phi \, dx \tag{4.10}
\]
for all \( \phi \in D_+(\Omega) \). Here \( F^+ := \{ x \in \Omega \mid \nabla u^+(x) = 0 \} \), \( D^+ := \Omega \setminus F^+ \).

The comparison principle discussed by Krügel [25, Theorem 4.8] states that
\[
(u^- - u^+) + \in D_+(\Omega) \implies u^- \leq u^+ \text{ almost everywhere in } \Omega. \tag{4.11}
\]
By testing \((u^- - u^+) + \in D_+(\Omega)\) into (4.9)(4.10) and subtracting the two inequalities, Krügel claims that \(\nabla u^- = \nabla u^+\) over \(\Omega' := \{ x \in \Omega \mid u^-(x) \geq u^+(x) \}\) and hence \(u^- = u^+\) almost everywhere in \(\Omega'\). Despite Krügel’s comment that integrals over \(F^-\) and \(F^+\) cancel out, however, it seems unclear whether
\[
\int_{F^-} |\nabla(u^- - u^+) +| \, dx = \int_{F^+} |\nabla(u^- - u^+) +| \, dx \tag{4.12}
\]
is valid. This problem is essentially due to the fact that Krügel did not appeal to monotonicity of the subdifferential operator \(\partial |\cdot|\) and did not regard the term \(\nabla u/|\nabla u|\) as an \(L^\infty\)-vector field satisfying the property (4.2). In our proof of the comparison principle (Proposition 2), we make use of monotonicity of the operator \(\partial |\cdot|\). Compared to our argument based on monotonicity, the equality (4.12) itself seems to be too strong to hold true.

### 4.2. Construction of classical subsolutions

In Section 4.2, we construct a classical subsolution to \(L_{b,p} v = 0\) in an open annulus.

**Lemma 2.** Let \( c \in \mathbb{R}^n \setminus \{0\} \), \( m > 0 \). Then for each fixed open ball \( B_R(x_*) \subset \mathbb{R}^n \), there exists a function \( h \in C^\infty(\mathbb{R}^n \setminus \{x_*\}) \) such that
\[
\begin{align*}
    h &= 0 \text{ on } \partial B_R(x_*), \quad 0 \leq h \leq m \text{ on } \overline{E_R(x_*)}, \tag{4.13} \\
    \partial_n h &< 0 \text{ on } \partial B_R(x_*), \tag{4.14} \\
    |\nabla h| &\leq \frac{|c|}{2} \text{ in } E_R(x_*), \tag{4.15} \\
    v(x) := h(x) + \langle c \mid x \rangle &\text{ satisfies } L_{b,p} v \leq 0 \text{ in } E_R(x_*), \text{ in the classical sense.} \tag{4.16}
\end{align*}
\]

Here \( E_R(x_*) := B_R(x_*) \setminus \overline{B_{R/2}(x_*)} \) is an open annulus, and \( v \) in (4.14) denotes the exterior unit vector normal to \( B_R(x_*) \).

Before proving Lemma 2, we fix some notations on matrices. For a given \( n \times n \) matrix \( A \), we write \( \text{tr}(A) \) as the trace of \( A \). We denote \( \mathbf{1}_n \) by the \( n \times n \) unit matrix. For column vectors \( x = (x_i)_i, \ y = (y_i)_i \in \mathbb{R}^n \), we define a tensor \( x \otimes y \), which is regarded as a real-valued \( n \times n \) matrix.
Continuity of Derivatives of a Convex Solution

\[ x \otimes y := (x_i y_j)_{i, j} = \begin{pmatrix} x_1 y_1 & \cdots & x_1 y_n \\ \vdots & \ddots & \vdots \\ x_n y_1 & \cdots & x_n y_n \end{pmatrix}. \]

Assume that \( h \) satisfies (4.15). Then the triangle inequality implies that

\[ 0 < \frac{1}{2} |c| \leq |\nabla v| \leq \frac{3}{2} |c| \quad \text{in} \ E_R(x_\ast). \tag{4.17} \]

The estimate (4.17) allows us to calculate \( L_{b, p}v \) in the classical sense over \( E_R(x_\ast) \).

By direct calculations we have

\[-L_{b, p}v = + \text{div} (\nabla E(\nabla v)) \]
\[ = \sum_{i, j=1}^n \partial_i \partial_j E(\nabla v) \partial_i x_j v = \text{tr} \left( \nabla^2 E(\nabla v) \nabla^2 h \right) \quad \text{in} \ E_R(x_\ast). \]

We note that \( \nabla^2 v = \nabla^2 h \) by definition. Here we recall a well-known result on Pucci’s extremal operators. For given constants \( 0 < \lambda \leq \Lambda < \infty \) and a fixed \( n \times n \) symmetric matrix \( M \), we define

\[ \mathcal{M}^- (M, \lambda, \Lambda) := \lambda \sum_{\lambda_i > 0} \lambda_i + \Lambda \sum_{\lambda_i < 0} \lambda_i, \]

where \( \lambda_i \in \mathbb{R} \) are the eigenvalues of \( M \). The following formula is a well-known result [1, Remark 5.36];

\[ \mathcal{M}^- (M, \lambda, \Lambda) = \inf \left\{ \text{tr}(AM) \mid A \in \mathcal{A}_{\lambda, \Lambda} \right\}, \]

where \( \mathcal{A}_{\lambda, \Lambda} \) denotes the set of all symmetric matrices whose eigenvalues all belong to the closed interval \([\lambda, \Lambda] \). By (4.17) \( L_{b, p}v \) is an uniformly elliptic operator in \( E_R(x_\ast) \). This enables us to find constants \( 0 < \lambda \leq \Lambda < \infty \), depending on \( 0 < b < \infty, 1 < p < \infty, |c| > 0 \), such that \( \nabla^2 E(\nabla v) \in [\lambda, \Lambda] \) in \( E_R(x_\ast) \). Combining these results, it suffices to show that

\[ \mathcal{M}^- \left( \nabla^2 h(x), \lambda, \Lambda \right) = \lambda \sum_{\lambda_i > 0} \lambda_i (x) + \Lambda \sum_{\lambda_i < 0} \lambda_i (x) > 0 \quad \text{for all} \ x \in E_R(x_\ast), \tag{4.18} \]

where \( \lambda_i (x) \in \mathbb{R} \) denotes the eigenvalues of \( \nabla^2 h(x) \).

Now we construct classical subsolutions. Our first construction is a modification of that by E. Hopf [20].

**Proof.** Without loss of generality we may assume \( x_\ast = 0 \). We define

\[ h(x) := e^{-\alpha |x|^2} - e^{-\alpha R^2} \quad \text{for} \ x \in \mathbb{R}^n. \tag{4.19} \]
Here \( \alpha = \alpha(b, n, p, |c|, R) > 0 \) is a sufficiently large constant to be chosen later. It is clear that
\[
0 \leq h(x) \leq e^{-\alpha R^2/4} - e^{-\alpha R^2} \quad \text{in} \quad E_R(0).
\]
We first let \( \alpha > 0 \) be so large that
\[
me^{\alpha R^2} \geq e^{3\alpha R^2/4} - 1.
\] (4.20)

From (4.20), we can easily check (4.13). By direct calculation we get
\[
\nabla h(x) = -2\alpha e^{-\alpha|x|^2} x, \quad \text{and} \quad \nabla^2 h(x) = -2\alpha e^{-\alpha|x|^2} 1_n + 4\alpha^2 e^{-\alpha|x|^2} x \otimes x \quad \text{for each} \quad x \in \mathbb{R}^n.
\]

From this result, (4.14) is clear. Also, we have
\[
|\nabla h(x)| \leq 2\alpha Re^{-\alpha R^2/4} \quad \text{for all} \quad x \in E_R(0).
\]

Let \( \alpha > 0 \) be so large that
\[
\alpha e^{-\alpha R^2/4} \leq \frac{|c|}{4R},
\] (4.21)
then we can check that \( h \) satisfies (4.15). Now we prove (4.16) to complete the proof. For \( x \neq 0 \), the eigenvalues of \( \nabla^2 h(x) \) are given by
\[
\begin{align*}
\lambda_\parallel(x) &:= 4\alpha^2 |x|^2 e^{-\alpha|x|^2} - 2\alpha e^{-\alpha|x|^2}, \\
\lambda_\perp(x) &:= -2\alpha e^{-\alpha|x|^2},
\end{align*}
\]
and the geometric multiplicities are \( \left\{ \frac{1}{n-1}. \right\} \)

Assume that \( \alpha \) satisfies
\[
\alpha > \frac{2}{R^2},
\] (4.22)
so that \( \lambda_\parallel > 0 > \lambda_\perp \) in \( E_R(0) \). Therefore we get
\[
\mathcal{M}^-(\nabla^2 h(x), \lambda, \Lambda) = \lambda_\parallel(x) + (n - 1)\Lambda \lambda_\perp(x)
\]
\[
= 2\alpha e^{-\alpha|x|^2} \left[ \lambda(2\alpha|x|^2 - 1) - (n - 1)\Lambda \right]
\]
\[
\geq 2\alpha e^{-\alpha|x|^2} \left[ \lambda \left( \frac{R^2}{2}\alpha - 1 \right) - (n - 1)\Lambda \right].
\]
We can take sufficiently large \( \alpha = \alpha(|c|, m, n, R, \lambda, \Lambda) > 0 \) so that \( \alpha \) satisfies (4.18) and (4.20)–(4.22). For such constant \( \alpha > 0 \), the function \( v \) defined as in (4.19) satisfies (4.13)–(4.16).

\[\square\]

It is possible to construct an alternative function \( h \in C^\infty(\mathbb{R}^n \setminus \{x_0\}) \) which satisfies (4.13)–(4.16). We give another proof of Lemma 2, which is derived from [29, Chapter 2.8].
Proof. Without loss of generality we may assume \( x_* = 0 \). We define
\[
h(x) := \beta \left[ |x|^{-\alpha} - R^{-\alpha} \right] \quad \text{for} \quad x \in \mathbb{R}^n \setminus \{0\}.
\] (4.23)

We will later determine positive constants \( \alpha, \beta > 0 \), depending on \( b, m, n, p, |c|, R \). It is clear that \( 0 \leq h(x) \leq \beta R^{-\alpha}(2^\alpha - 1) \) in \( E_R(0) \). We first let \( \alpha, \beta > 0 \) satisfy
\[
\beta \leq \frac{m R^\alpha}{2^\alpha - 1}.
\] (4.24)

Then \( h \) satisfies (4.13). By direct calculation we get
\[
\nabla h(x) = -\frac{\alpha \beta x}{|x|^{\alpha+2}}, \quad \text{and} \quad \nabla^2 h(x) = \frac{\alpha \beta}{|x|^{\alpha+2}} \left[ (\alpha + 2) \frac{x \otimes x}{|x|^2} - I_n \right]
\]
for each \( x \in \mathbb{R}^n \setminus \{0\} \). The estimate (4.14) is clear by this result. Also, we have
\[
|\nabla h(x)| \leq \frac{\alpha \beta}{(R/2)^{\alpha+1}} \quad \text{for all} \quad x \in E_R(0).
\]

Let \( \alpha, \beta > 0 \) satisfy
\[
\beta \leq \frac{|c|(R/2)^{\alpha+1}}{2\alpha},
\] (4.25)
then we can check that \( h \) satisfies (4.15). Now we prove (4.16) to complete the proof. For \( x \neq 0 \), the eigenvalues of \( \nabla^2 h(x) \) are given by
\[
\begin{align*}
\lambda_\parallel(x) &:= (\alpha + 1)\alpha \beta |x|^{-\alpha-2}, \\
\lambda_\perp(x) &:= -\alpha \beta |x|^{-\alpha-2},
\end{align*}
\]
and the geometric multiplicities are \( \{ 1, n-1 \} \). It is clear that \( \lambda_\parallel > 0 > \lambda_\perp \) in \( \mathbb{R}^n \setminus \{0\} \), and therefore
\[
\mathcal{M}^- \left( \nabla^2 h(x), \lambda, \Lambda \right) = \alpha \beta |x|^{-\alpha-2} [(\alpha + 1)\lambda - (n-1)\Lambda]
\]
for all \( x \in E_R(0) \). We take and fix sufficiently large \( \alpha = \alpha(n, \lambda, \Lambda) > 0 \) so that \( \alpha \) satisfies (4.18). For such \( \alpha > 0 \), we choose sufficiently small \( \beta = \beta(|c|, R, \alpha) > 0 \) so that \( \beta \) satisfies (4.24)–(4.25). Then the function \( h \) defined as in (4.23) satisfies (4.13)–(4.16).

4.3. Strong maximum principle

We prove the strong maximum principle (Theorem 2).

Proof. Let \( D_0 \subset D \) be a connected component of the open set \( D \), and \( x_0 \in D_0 \).
Without loss of generality we may assume that \( x_0 = 0 \) and \( u(0) = 0 \). By Lemma 1, it is clear that \( \partial u(0) = \{ \nabla u(0) \} \neq \{0\} \). We set a vector \( c := \nabla u(0) \in \mathbb{R}^n \setminus \{0\} \) and a relatively closed set
\[
\Sigma := \{ x \in D_0 \mid u(x) = \langle c \mid x \rangle \},
\]
and we will prove that $\Sigma = D_0$. It is also clear that $0 \in \Sigma$ and hence $\Sigma \neq \emptyset$. Suppose for contradiction that $\Sigma \subsetneq D_0$. Then it follows that $\partial \Sigma \cap D_0 \neq \emptyset$, since $D_0$ is connected. We may take and fix a point $x_\ast \in D_0 \setminus \Sigma$ such that $\text{dist}(x_\ast, \Sigma) < \text{dist}(x_\ast, \partial D_0)$. By extending a closed ball centered at $x_\ast$ until it hits $\Sigma$, we can take a point $y_\ast \in D_0$ and a closed ball $\overline{B}_R(x_\ast) \subset D_0$ such that $y_\ast \in \partial B_R(x_\ast) \cap \Sigma$ and $u(x) > \langle c \mid x \rangle$ for all $x \in B_R(x_\ast)$. We note that

$$
0 = \min_{x \in \partial B_R(x_\ast)} \{ u(x) - \langle c \mid x \rangle \}, \quad \text{achieved at } y_\ast \in \partial B_R(x_\ast),
$$

and

$$
m := \min_{x \in \partial B_R/2(x_\ast)} \{ u(x) - \langle c \mid x \rangle \} > 0, \quad (4.26)
$$

by construction of $B_R(x_\ast)$. Let $h \in C^\infty(\mathbb{R}^n \setminus \{x_\ast\})$ be an auxiliary function as in Lemma 2. Then from (4.26) it is easy to check that $v := h + \langle c \mid x \rangle$ satisfies $v \leq u$ on $\partial E_R(x_\ast)$, in the sense of (4.7). By Proposition 2, we have $v \leq u$ on $E_R(x_\ast)$. Hence $0 \leq u - \langle c \mid x \rangle - h$ in $E_R(x_\ast)$. This inequality becomes equality at $y_\ast \in \partial E_R(x_\ast)$ by (4.13) and (4.26). Therefore the function $u(x) - \langle c \mid x \rangle - h(x) (x \in E_R(x_\ast))$ takes its minimum at $y_\ast \in \partial B_R(x_\ast)$. Also by $y_\ast \in \Sigma$ and the subgradient inequality

$$
u(x) \geq \langle c \mid x \rangle \quad \text{for all } x \in \Omega,$$

it is clear that the function $w(x) := u(x) - \langle c \mid x \rangle (x \in D_0)$ takes its minimum $0$ at $y_\ast \in D_0$. We note that $w, w - h \in C^1(D_0)$ by Lemma 1. By calculating classical partial derivatives at $y_\ast$ in the direction $v_0 := (y_\ast - x_\ast)/R$, we obtain

$$0 \geq \partial_{v_0}(w - h)(y_\ast) = -\partial_{v_0}h(y_\ast) > 0.$$ 

This is a contradiction, and therefore $\Sigma = D_0$. \hfill \Box

## 5. Proofs of Main Theorems

In Section 5, we give proofs of the Liouville-type theorem (Theorem 3) and the $C^1$-regularity theorem (Theorem 1).

### 5.1. Liouville-type theorem

For a preparation, we prove Lemma 3 below.

**Lemma 3.** Let $u$ be a real-valued convex function in $\mathbb{R}^n$. Assume that $u$ satisfies the following,

1. The facet of $u$, $F \subset \mathbb{R}^n$, satisfies $\emptyset \subsetneq F \subsetneq \mathbb{R}^n$.
2. $u$ attains its minimum $0$.
3. $u$ is affine in each connected component of $D := \mathbb{R}^n \setminus F$.

Then up to a rotation and a shift translation, $u$ can be expressed as either of the following three types of piecewise-linear functions.

$$u(x) = \max \{ t_1 x_1, 0 \} \quad \text{for all } x \in \mathbb{R}^n, \quad (5.1)$$

$$u(x) = \max \{ t_1 x_1, -t_2 x_1 \} \quad \text{for all } x \in \mathbb{R}^n, \quad (5.2)$$

$$u(x) = \max \{ t_1 x_1, 0, -t_2 (x_1 + l_0) \} \quad \text{for all } x \in \mathbb{R}^n. \quad (5.3)$$

Here $t_1, t_2, l > 0$ are constants.
Before starting the proof of Lemma 3, we introduce notations on affine hyperplanes. For \( c \in \mathbb{R}^n \setminus \{0\} \) and \( x_0 \in \mathbb{R}^n \), we define

\[
\begin{align*}
H_{c, x_0}^- := \{ x \in \mathbb{R}^n \mid \langle c \mid x - x_0 \rangle < 0 \}, \\
H_{c, x_0}^+ := \{ x \in \mathbb{R}^n \mid \langle c \mid x - x_0 \rangle > 0 \}, \\
\end{align*}
\]

In order to prove the Liouville-type theorem, we will make use of the supporting hyperplane theorem, which states that for any non-empty closed convex set \( C \subset \mathbb{R}^n \) and \( x_0 \in \partial C \), there exists \( c \in \mathbb{R}^n \setminus \{0\} \) such that

\[
\sup_{x \in C} \langle c \mid x \rangle \leq \langle c \mid x_0 \rangle, \quad \text{and in particular} \quad H_{c, x_0}^+ \subset \mathbb{R}^n \setminus C.
\]

For such \( c \in \mathbb{R}^n \setminus \{0\} \), a hyperplane \( H_{c, x_0} \) is often called a supporting hyperplane for \( C \) at the boundary point \( x_0 \). For the proof of the supporting hyperplane theorem, see [3, Proposition 1.5.1].

**Proof.** Since \( \mathbb{R}^n \) is connected and \( F \subset \mathbb{R}^n \) is a closed convex set, it follows that \( \partial F \neq \emptyset \). Without loss of generality we may assume that \( 0 \in \partial F \) and \( u(0) = 0 \).

By the supporting hyperplane theorem, we can take and fix a supporting hyperplane for \( F \) at the boundary point \( 0 \), which we write \( H_{e_1, 0} \subset \mathbb{R}^n \). By rotation, we may assume that \( c = e_1 \). Let \( D_1 \) be the connected component of \( D \) which contains \( H_{e_1, 0}^+ \subset \mathbb{R}^n \setminus F = D \). By the assumption 3 and \( u(0) = 0 \), it follows that there exists \( c \in \mathbb{R}^n \setminus \{0\} \) such that \( u(x) = \langle c \mid x \rangle \) for all \( x \in D_1 \). We should note that \( H_{c, 0} = H_{e_1, 0} \) and hence \( c = t_1 e_1 \) for some \( t_1 \in (0, \infty) \), since otherwise it follows that \( H_{e_1, 0}^+ \cap H_{c, 0}^- \neq \emptyset \) and \( 0 \leq u(x_0) = \langle c \mid x_0 \rangle < 0 \) for any \( x_0 \in H_{e_1, 0}^+ \cap H_{c, 0}^- \). The result \( H_{c, 0} = H_{e_1, 0} \) also implies that \( H_{e_1, 0} \subset \partial F \subset F \subset \{ x \in \mathbb{R}^n \mid x_1 \leq 0 \} = H_{e_1, 0}^- \cup H_{e_1, 0} \). Now we will deduce three possible representations of \( u \).

If \( \partial F = H_{e_1, 0} \), then we have either \( F = H_{e_1, 0}^- \cup H_{e_1, 0} \) or \( F = H_{e_1, 0} \), since the open set \( H_{e_1, 0}^- = \{ x \in \mathbb{R}^n \mid x_1 < 0 \} \) is connected. For the first case, \( u \) is clearly expressed by (5.1). For the second case, it is clear that \( D \) consists of two connected components \( D_1 = H_{e_1, 0}^+ \) and \( D_2 = H_{e_1, 0}^- \). Again, by the condition 3 and similar arguments as those above, we can determine \( u|D_2 = a(x) = \langle -t_2 e_1 \mid x \rangle \) for all \( x \in D_2 \). Here \( t_2 \in (0, \infty) \) is a constant. Hence we obtain (5.2). For the case \( H_{e_1, 0} \subset \partial F \), we take and fix \( z_0 \in \partial F \setminus H_{e_1, 0} \) and a supporting hyperplane for \( F \) at \( z_0 \), which we write by \( H_{e', z_0} \). Let \( D_2 \) be the connected component of \( D \) which contains \( H_{e', z_0}^+ \subset D \). By the assumption 3 and \( u(z_0) = 0 \), it follows that there exists \( c'' \in \mathbb{R}^n \setminus \{0\} \) such that \( u(x) = \langle c'' \mid x - z_0 \rangle \) for all \( x \in D_2 \). Completely similarly to the arguments above for showing that \( H_{c, 0} = H_{e_1, 0} \), we can easily notice that \( H_{c', z_0} = H_{c', z_0} \) and hence \( c'' = t' c' \) for some constant \( t_1' \in (0, \infty) \). Moreover, we also realize that \( c' = t_k e_1 \) for some \( t_k \in \mathbb{R} \setminus \{0\} \). Otherwise it follows that the two hyperplanes \( H_{e_1, 0} \) and \( H_{e', z_0} \) cross, and hence we get \( D = D_2 \) and \( H_{e_1, 0}^+ \cap H_{c', z_0}^- 
eq \emptyset \), which implies that there exists a point \( x_0 \in D \) such that \( u(x_0) < 0 \). This is clearly a contradiction. This result and convexity of \( u \) imply that \( D \) consists of two connected components \( D_1 = H_{e_1, 0}^+ \) and \( D_2 = H_{e_1, 0}^- \), and that \( F = \{ x \in \mathbb{R}^n \mid -l_0 \leq x_1 \leq 0 \} \). Here \( l_0 := \text{dist}(H_{e_1, 0}, H_{-e_1, z_0}) > 0 \). Finally we obtain the last possible expression (5.3). \( u \) can be expressed by either of (5.1)–(5.3). \( \square \)
Now we give the proof of Theorem 3.

**Proof.** Assume by contradiction that $F_1$, the facet of $u$, would satisfy $\emptyset \not\subset F \subset \mathbb{R}^n$. Without loss of generality, we may assume that $u$ attains its minimum 0. By the strong maximum principle (Theorem 2), the convex weak solution $u$ is affine in each connected component of $D := \mathbb{R}^n \setminus F$. Therefore we are able to apply Lemma 3. By rotation and translation, $u$ can be expressed as (5.1)–(5.3). Now we prove that $u$ is no longer a weak solution to $L_{b_p} u = 0$ in $\mathbb{R}^n$. We set open cubes $Q' := (-1, 1)^{n-1} \subset \mathbb{R}^{n-1}$ and $Q := (-d, d) \times Q' \subset \mathbb{R}^n$, where $d > 0$ is to be chosen later. We claim that $u$ does not satisfy $L_{b_p} u = 0$ in $W^{-1, p'}(Q)$. Assume by contradiction that there exists a vector field $(u, Z) \in W^{1, p}(Q) \times L^{\infty}(Q, \mathbb{R}^n)$ such that the pair $(u, Z)$ satisfies $L_{b_p} u = 0$ in $W^{-1, p'}(Q)$.

For the first case (5.1), we have

$$|Z(x)| \leq 1 \quad \text{for almost everywhere } x \in Q, \quad \text{and}$$

$$Z(x) = e_1 \quad \text{for almost everywhere } x \in Q_r := (0, d) \times Q' \subset \mathbb{R}^n. \quad (5.4)$$

by definition of $Z$. We also set another open cube $Q_l := (-d, 0) \times Q' \subset \mathbb{R}^n$. We take and fix non-negative functions $\phi_1 \in C^1_c((-d, d))$, $\phi_2 \in C^1_c(Q')$ such that

$$\phi_1 \geq 0 \text{ in } (-d, 0), \quad \max_{(-d, d)} \phi_1 = \phi_1(0) > 0, \quad \text{and } \phi_2 \neq 0. \quad (5.5)$$

We define an admissible test function $\phi \in C^1_c(Q)$ by $\phi(x_1, x') := \phi_1(x_1)\phi_2(x')$ for $(x_1, x') \in (-d, d) \times Q' = Q$. Test $\phi \in C^1_c(Q)$ into $L_{b_p} u = 0$ in $W^{-1, p'}(Q)$, and divide the integration over $Q$ into that over $Q_l$ and $Q_r$. Then (5.4) implies that

$$0 = b \int_{Q_l} (Z + |0|^{p-2}0) \nabla(\phi_1\phi_2)) \ dx + \int_{Q_r} \left( (b + t_1^{p-1}) e_1 \right) \nabla(\phi_1\phi_2) \ dx$$

$$\leq b \int_{Q_l} \phi_1' \phi_2 \ dx + b \int_{Q_l} \phi_1 |\nabla\phi_2| \ dx$$

$$+ b\phi_1(0) \int_{Q'} \phi_2(x')(e_1 | -e_1) \ dx' + t_1^{p-1}\phi_1(0) \int_{Q'} \phi_2(x')(e_1 | -e_1) \ dx'$$

$$=: I_1 + I_2 + I_3 + I_4.$$

Here we have applied the Gauss–Green theorem to the integration over $Q_r$, and the Cauchy–Schwarz inequality to the integration over $Q_l$. For the integrations $I_1$ and $I_2$, we make use of Fubini’s theorem and (5.5). Then we have

$$I_1 = \int_{Q'} \left( \int_{-d}^0 \phi_1'(x_1) dx_1 \right) \phi_2(x') \ dx'$$

$$= b\phi_1(0) \int_{Q'} \phi_2(x') \ dx' = b\phi_1(0)\|\phi_2\|_{L^1(Q')} = -I_3,$$

$$I_2 \leq b\phi_1(0) \int_{-d}^0 dx_1 \int_{Q'} |\nabla\phi_2(x')| \ dx' = bd\phi_1(0)\|\nabla\phi_2\|_{L^1(Q')}.$$
Finally we obtain
\[
0 \leq I_1 + I_2 + I_3 + I_4 \leq I_2 + I_4 \leq \phi_1(0) \left( b d \| \nabla \phi_2 \|_{L^1(Q')} - t_1^{p-1} \| \phi_2 \|_{L^1(Q')} \right).
\]
(5.6)

From (5.6), we can easily deduce a contradiction by choosing sufficiently small \(d = d(b, p, t_1, \phi_2) > 0\). Similarly we can prove that \(u\) defined as in (5.3) does not satisfy \(L_b, p u = 0\) in \(W^{-1, p'}(Q)\), since it suffices to restrict \(d < l_0\). We consider the remaining case (5.2). We have

\[
Z(x) = \begin{cases} 
  e_1 & \text{for almost everywhere } x \in Q_r, \\
  -e_1 & \text{for almost everywhere } x \in Q_l.
\end{cases}
\]

by definition of \(Z\). We test the same function \(\phi \in C^1_c(Q)\) in \(L_b, p u = 0\), then it follows that

\[
0 = \int_{Q_l} \left( - (b + t_2^{p-1}) e_1 \mid \nabla (\phi_1 \phi_2) \right) \, dx + \int_{Q_r} \left( (b + t_1^{p-1}) e_1 \mid \nabla (\phi_1 \phi_2) \right) \, dx
\]
\[
= - \left( b + t_2^{p-1} \right) \int_{Q'} \phi_1(0) \phi_2(x') \langle e_1 \mid e_1 \rangle \, dx' + \left( b + t_1^{p-1} \right) \int_{Q'} \phi_1(0) \phi_2(x') \langle e_1 \mid -e_1 \rangle \, dx'
\]
\[
= - \phi_1(0) \left( 2b + t_1^{p-1} + |t_2|^{p-1} \right) \int_{Q'} \phi_2(x') \, dx' < 0,
\]

which is a contradiction. This completes the proof. \(\square\)

**Remark 4.** The estimate (5.6) breaks for \(p = 1\), since the equation \(|0|^{p-2}0 = 0\) is no longer valid for \(p = 1\). This means that we have implicitly used differentiability of the function \(|z|^p / p\) at \(0 \in \mathbb{R}^n\). Also it should be noted that for the one-variable case, functions as in (5.1), which are in general not in \(C^1\), are one-harmonic in \(\mathbb{R}\).

### 5.2. \(C^1\)-Regularity theorem

We give the proof of Theorem 1.

**Proof.** We may assume that \(\Omega\) is convex. By [30, Theorem 25.1 and 25.5] and Lemma 1, it suffices to show that \(\partial u(x_0) = \{0\}\) for all \(x_0 \in F\). Let \(x_0 \in F\). We get a convex function \(u_0: \mathbb{R}^n \to \mathbb{R}\) as a blow-up limit as in Proposition 1. We note that the facet of \(u_0\) is non-empty by Proposition 1. Hence by the Liouville-type theorem (Theorem 3), \(u_0\) is constant and we obtain \(\partial u_0(x_0) = \{0\}\). Combining these results, we have \(\{0\} \subset \partial u(x_0) \subset \partial u_0(x_0) = \{0\}\) and therefore \(\partial u(x_0) = \{0\}\). This completes the proof. \(\square\)
6. Generalization

In Section 6, we would like to discuss $C^1$-regularity of convex weak solutions to

$$Lu := -\text{div}(\nabla_z \Psi(\nabla u)) - \text{div}(\nabla_z W(\nabla u)) = f \quad \text{in } \Omega \subset \mathbb{R}^n,$$

which covers (1.1). Precisely speaking, throughout Section 6, we make these following assumptions for $\Psi$ and $W$ on regularity and ellipticity. For regularity, we only require

$$\Psi \in C(\mathbb{R}^n) \cap C^2(\mathbb{R}^n \setminus \{0\}), \quad W \in C^1(\mathbb{R}^n) \cap C^2(\mathbb{R}^n \setminus \{0\}). \quad (6.2)$$

For $W$, we assume that for each fixed $0 < \mu \leq M < \infty$, there exist constants $0 < \gamma < \Gamma < \infty$ such that $W$ satisfies

$$\gamma |\zeta|^2 \leq \left| \nabla^2_z W(z_0)\zeta \right| \zeta, \quad (6.3)$$

$$\left| \left( \nabla^2_z W(z_0)\zeta \right| \omega \right| \leq \Gamma |\zeta||\omega| \quad (6.4)$$

for all $z_0, \zeta, \omega \in \mathbb{R}^n$ with $\mu \leq |z_0| \leq M$. Also, there is no loss of generality in assuming that

$$\nabla_z W(0) = 0. \quad (6.5)$$

Finally, we assume that $\Psi$ is positively homogeneous of degree 1. In other words, $\Psi$ satisfies

$$\Psi(\lambda z_0) = \lambda \Psi(z_0) \quad (6.6)$$

holds for all $z_0 \in \mathbb{R}^n$ and $\lambda > 0$. This clearly yields $\Psi(0) = 0$.

By modifying some of our arguments, we are able to show that

**Theorem 4.** ($C^1$-regularity theorem for general equations) *Let $\Omega \subset \mathbb{R}^n$ be a domain. Assume that $f \in L^q_{\text{loc}}(\Omega) \ (n < q \leq \infty)$ and the functionals $\Psi$ and $W$ satisfy (6.2)–(6.5). If $u$ is a convex weak solution to (6.1), then $u$ is in $C^1(\Omega)$.*

If we set

$$\Psi(z) := b|z|, \quad W(z) := \frac{|z|^p}{p}, \quad \text{where } 1 < p < \infty,$$

then the Equation (6.1) becomes (1.1). Therefore Theorem 4 generalizes Theorem 1.
6.1. Preliminaries

In Section 6.1, we mention some basic properties of \( \Psi \) and \( W \), which are derived from the assumptions (6.2)–(6.5).

For \( W \), by (6.2)–(6.3) and (6.5) it is easy to check that the continuous mapping \( A : \mathbb{R}^n \ni z \mapsto \nabla W(z) \in \mathbb{R}^n \) satisfies strict monotonicity (4.8). In particular, by (6.5) we have

\[
\langle A(z) \mid z \rangle > 0 \quad \text{for all } z \in \mathbb{R}^n \setminus \{0\}. \tag{6.7}
\]

For the proof, see Lemma 10 in the appendices.

For \( \Psi_1 \), we first note that \( \Psi_1 \) satisfies the triangle inequality

\[
\Psi_1(z_1 + z_2) \leq \Psi_1(z_1) + \Psi_1(z_2) \quad \text{for all } z_1, z_2 \in \mathbb{R}^n. \tag{6.8}
\]

We define a function \( \tilde{\Psi}_1 : \mathbb{R}^n \to [0, \infty] \) by

\[
\tilde{\Psi}_1(\zeta) := \sup \{ \langle \zeta \mid z \rangle \mid z \in \mathbb{R}^n, \Psi_1(z) \leq 1 \}.
\]

\( \tilde{\Psi}_1 \) is the support function for the closed convex set \( C_{\Psi_1} := \{ z \in \mathbb{R}^n \mid \Psi_1(z) \leq 1 \} \). By definition it is easy to check that \( \tilde{\Psi}_1 \) is convex and lower semicontinuous. Also, if \( \zeta \in \mathbb{R}^n \) satisfies \( \tilde{\Psi}_1(\zeta) < \infty \), then the following Cauchy–Schwarz-type inequality holds;

\[
\langle z \mid \zeta \rangle \leq \Psi_1(z) \tilde{\Psi}_1(\zeta) \quad \text{for all } z \in \mathbb{R}^n. \tag{6.9}
\]

If a convex function \( \Psi \) is positively homogeneous of degree 1, then the subdifferential operator \( \partial \Psi \) is explicitly given by

\[
\partial \Psi(z) = \left\{ \zeta \in \mathbb{R}^n \mid \tilde{\Psi}_1(\zeta) \leq 1, \Psi(z) = \langle z \mid \zeta \rangle \right\} \tag{6.10}
\]

for all \( z \in \mathbb{R}^n \). In particular, we have the following formula

\[
\langle \nabla_z \Psi(z_0) \mid z_0 \rangle = \Psi(z_0) \quad \text{for all } z_0 \in \mathbb{R}^n \setminus \{0\}, \tag{6.11}
\]

which is often called Euler’s identity. Also, assumptions (6.2) and (6.6) imply that

\[
\nabla \Psi(\lambda z_0) = \lambda \nabla \Psi(z_0), \quad \nabla^2 \Psi(\lambda z_0) = \lambda^{-1} \nabla \Psi(z_0) \tag{6.12}
\]

for all \( \lambda > 0 \) and \( z_0 \in \mathbb{R}^n \setminus \{0\} \). Proofs of (6.8)–(6.10) are given in Lemma 11 of the appendices for the reader’s convenience.

**Remark 5.** The results (6.11)–(6.12) give us the following basic property for \( \Psi \).

1. We set a constant

\[
K := \sup \left\{ |\nabla_z \Psi(z_0)| \mid z_0 \in \mathbb{R}^n, |z_0| = 1 \right\},
\]

which is finite. Then we have \( \partial \Psi(z_0) \subset \overline{B_K}(0) \) for all \( z_0 \in \mathbb{R}^n \). For the case \( z_0 \neq 0 \), this inclusion is clear by (6.12) and \( \partial \Psi(z_0) = \{ \nabla_z \Psi(z_0) \} \). For \( z_0 = 0 \), we take arbitrary \( w \in \partial \Psi(0) \setminus \{0\} \). Then by the subgradient inequality, Euler’s identity (6.11) and the Cauchy–Schwarz inequality, we have

\[
|w|^2 = \langle w \mid w - 0 \rangle + \Psi(0) \leq \Psi(w) = \langle \nabla_z \Psi(w) \mid w \rangle \leq K|w|.
\]

This estimate yields the inclusion \( \partial \Psi(0) \subset \overline{B_K}(0) \).
2. For $z_0 \in \mathbb{R}^n \setminus \{0\}$, the Hessian matrix $\nabla_z^2 \Psi(z_0)$ satisfies
\begin{align}
0 & \leq \frac{\nabla_z^2 \Psi(z_0) \xi}{\xi} \\
\left\| \frac{\nabla_z^2 \Psi(z_0) \omega}{\omega} \right\| & \leq \frac{C}{|z_0|} |\xi| |\omega|
\end{align}
for all $\zeta, \omega \in \mathbb{R}^n$. Here the finite constant $C$ is explicitly given by
\[
C := \sup \left\{ \left\| \frac{\nabla_z^2 \Psi(w) \zeta}{\zeta} \right\| \omega \mid z, \zeta, \omega \in \mathbb{R}^n, |w| = |\eta| = |\omega| = 1 \right\}.
\]

Lemma 4 states lower semicontinuity of a functional in the weak* topology of an $L^\infty$-space. This result is used in the justification of a blow-up argument for the Equation (6.1).

**Lemma 4.** Let $\Omega \subset \mathbb{R}^m$ be a Lebesgue measurable set, and let $\Psi : \mathbb{R}^n \to [0, \infty)$ be a convex function which satisfies (6.6). Assume that a vector field $Z \in L^\infty(\Omega, \mathbb{R}^n)$ and a sequence $\{Z_N\}_N \subset L^\infty(\Omega, \mathbb{R}^n)$ satisfy $Z_N \rightharpoonup^* Z$ in $L^\infty(\Omega, \mathbb{R}^n)$. Then we have
\[
\text{ess sup}_{x \in \Omega} \Psi(Z(x)) \leq \liminf_{N \to \infty} \text{ess sup}_{x \in \Omega} \Psi(Z_N(x)),
\]
where $\tilde{\Psi}$ denotes the support function of the closed convex set $C_\Psi := \{ z \in \mathbb{R}^n \mid \Psi(z) \leq 1 \}$.

We give an elementary proof of Lemma 4, which is based on a definition of $\tilde{\Psi}$.

**Proof.** We consider the case $C_\infty := \liminf_{N \to \infty} \left\| \tilde{\Psi}(Z_N) \right\|_{L^\infty(\Omega)} < \infty$, since otherwise (6.15) is clear. Fix arbitrary $\varepsilon > 0$. Then we may take a subsequence $\{Z_{N_j}\}_{j=1}^\infty$ such that
\[
\text{ess sup}_{x \in \Omega} \tilde{\Psi}(Z_{N_j}(x)) \leq C_\infty + \varepsilon < \infty.
\]
Take arbitrary $0 \leq \phi \in L^1(\Omega)$ and $w \in C_\Psi$. Then with the aid of (6.9), we have
\[
\langle Z_{N_j}(x) \mid w \rangle \leq C_\infty + \varepsilon
\]
for all $j \in \mathbb{N}$ and for almost everywhere $x \in \Omega$, which yields
\[
\int_\Omega \left[ C_\infty + \varepsilon - \langle Z_{N_j}(x) \mid w \rangle \right] \phi(x) \, dx \geq 0
\]
for all $j \in \mathbb{N}$. Letting $j \to \infty$, we have
\[
\int_\Omega \left[ C_\infty + \varepsilon - \langle Z(x) \mid w \rangle \right] \phi(x) \, dx \geq 0,
\]
by $Z_{N_j} \rightharpoonup^* Z$ in $L^\infty(\Omega, \mathbb{R}^n)$. Since $0 \leq \phi \in L^1(\Omega)$ is arbitrary, for each $w \in C_\Psi$, there exists an $\mathcal{L}^n$-measurable set $U_w \subset \Omega$, such that $\mathcal{L}^n(U_w) = 0$ and
\[
\langle Z(x) \mid w \rangle \leq C_\infty + \varepsilon
\]
for all $x \in \Omega \setminus U_w$. 
Here we denote $L^n$ by the $n$-dimensional Lebesgue measure. Since $C_{\Psi} \subset \mathbb{R}^n$ is separable, we may take a countable and dense set $D \subset C_{\Psi}$. We set an $L^n$-measurable set

$$ U := \bigcup_{w \in D} U_w \subset \Omega, $$

which clearly satisfies $L^n(U) = 0$. Then we conclude that

$$ \langle Z(x) \mid w \rangle \leq C_\infty + \varepsilon \quad \text{for all } x \in \Omega \setminus U, \ w \in C_{\Psi} $$

from density of $D \subset C_{\Psi}$. Hence by definition of $\tilde{\Psi}$, it is clear that

$$ \tilde{\Psi}(Z(x)) \leq C_\infty + \varepsilon \quad \text{for almost everywhere } x \in \Omega. $$

Since $\varepsilon > 0$ is arbitrary, this completes the proof of (6.15). \hfill \Box

6.2. Sketches of the proofs

We first give definitions of weak solutions to (6.1). We also define weak subsolutions, and supersolutions to an equation $Lu = 0$ in a bounded domain.

**Definition 3.** Let $\Omega \subset \mathbb{R}^n$ be a domain.

1. Let $f \in L^q_{\text{loc}}(\Omega)$ ($n < q \leq \infty$). We say that a function $u \in W^{1,\infty}_{\text{loc}}(\Omega)$ is a weak solution to (6.1), when for any bounded Lipschitz domain $\omega \subset \subset \Omega$, there exists a vector field $Z \in L^\infty(\omega, \mathbb{R}^n)$ such that the pair $(u, Z) \in W^{1,\infty}(\omega) \times L^\infty(\omega, \mathbb{R}^n)$ satisfies

$$ \int_\omega \langle Z \mid \nabla \phi \rangle \, dx + \int_\omega \langle A(\nabla u) \mid \nabla \phi \rangle \, dx = \int_\omega f \phi \, dx $$

for all $0 \leq \phi \in C^\infty_c(\omega)$, and

$$ Z(x) \in \partial \Psi(\nabla u(x)) $$

for almost everywhere $x \in \omega$. Here $A$ denotes the continuous mapping $A : \mathbb{R}^n \ni x \mapsto \nabla_z W(x) \in \mathbb{R}^n$. For such pair $(u, Z)$, we say that $(u, Z)$ satisfies $Lu = f$ in $W^{-1,\infty}(\omega)$ or simply say that $u$ satisfies $Lu = f$ in $W^{-1,\infty}(\omega)$.

2. Assume that $\Omega$ is bounded. A pair $(u, Z) \in W^{1,\infty}(\Omega) \times L^\infty(\Omega, \mathbb{R}^n)$ is called a weak subsolution to $Lu = 0$ in $\Omega$, if it satisfies

$$ \int_\Omega \langle Z \mid \nabla \phi \rangle \, dx + \int_\Omega \langle A(\nabla u) \mid \nabla \phi \rangle \, dx \leq 0 $$

for all $0 \leq \phi \in C^\infty_c(\Omega)$, and

$$ Z(x) \in \partial \Psi(\nabla u(x)) \quad \text{for almost everywhere } x \in \Omega. $$

Similarly we call a pair $(u, Z) \in W^{1, p}(\Omega) \times L^\infty(\Omega, \mathbb{R}^n)$ a weak supersolution $Lb, pu = 0$ in $\Omega$, if it satisfies (6.21) and

$$ \int_\Omega \langle Z \mid \nabla \phi \rangle \, dx + \int_\Omega \langle A(\nabla u) \mid \nabla \phi \rangle \, dx \geq 0 $$

for all $0 \leq \phi \in C_c^\infty(\Omega)$.\hfill \Box
for all $0 \leq \phi \in C_0^\infty(\Omega)$. For $u \in W^{1, p}(\Omega)$, we simply say that $u$ is respectively a subsolution and a supersolution to $Lu = 0$ in the weak sense if there is $Z \in L^\infty(\Omega, \mathbb{R}^n)$ such that the pair $(u, Z)$ is a weak subsolution and a weak supersolution to $Lu = 0$ in $\Omega$.

**Remark 6.** We describe some remarks on Definition 3.

1. In this paper we treat a convex solution, which clearly satisfies local Lipschitz regularity. Hence it is not restrictive to assume local or global $W^{1, \infty}$ regularity for solutions in Definition 3. Also it should be noted that if a vector field $Z$ satisfies (6.19), then $Z$ is in $L^\infty$ by Remark 5. Hence our regularity assumptions of the pair $(u, Z)$ involve no loss of generality.

2. Integrals in (6.18) make sense by $Z$, $\nabla u \in L^\infty(\omega, \mathbb{R}^n)$, $A \in C(\mathbb{R}^n, \mathbb{R}^n)$, and the continuous embedding $W^{1, p}_0(\omega) \hookrightarrow L^q(\omega)$.

3. For a bounded domain $\Omega \subset \mathbb{R}^n$, let $u \in C^2(\overline{\Omega})$ satisfy $\nabla u(x) \neq 0$ and $Lu(x) \leq 0$ for all $x \in \Omega$.

Then the pair $(u, \nabla z \Psi(\nabla u)) \in W^{1, p}(\Omega) \times L^\infty(\Omega, \mathbb{R}^n)$ satisfies (6.20)–(6.21). For such $u$, we simply say that $u$ satisfies $Lu \leq 0$ in $\Omega$ in the classical sense.

To prove Theorem 4, we may assume that $\Omega$ is a bounded convex domain, since our argument is local. As described in Section 1.1, we would like to prove that a convex solution $u$ to (6.1) satisfies (1.4) for all $x \in \Omega$.

For the case $x \in D$, we can show (1.4) by De Giorgi–Nash–Moser theory. This is basically due to the fact that the functional $E(z) := \Psi(z) + W(z)$ for $z \in \mathbb{R}^n$ satisfy the following property. For each fixed constants $0 < \mu \leq M < \infty$, there exists constants $0 < \lambda \leq \Lambda < \infty$ such that the estimates (2.7)–(2.8) hold for all $z_0$, $\zeta$, $\omega \in \mathbb{R}^n$ with $\mu \leq |z_0| \leq M$. In other words, the operator $L$ is **locally uniformly elliptic outside a facet**, in the sense that for a function $v$ the operator $Lv$ becomes uniformly elliptic in a place where $0 < \mu \leq |\nabla v| \leq M < \infty$ holds. This ellipticity is an easy consequence of (6.3)–(6.4) and (6.13)–(6.14). Appealing to local uniform ellipticity of the operator $L$ outside the facet and De Giorgi–Nash–Moser theory, we are able to show that a convex solution to $Lu = f$ is $C^{1, \alpha}$ near a neighborhood of each fixed point $x \in D$, similarly to the proof of Lemma 1.

For the case $x \in F$, we first make a blow-argument to construct a convex function $u_0 : \mathbb{R}^n \rightarrow \mathbb{R}$ satisfying $\partial u(x) \subset \partial u_0(x)$, and $Lu_0 = 0$ in $\mathbb{R}^n$ in the sense of Definition 3. Next we justify a maximum principle, which is described as in (1.6), holds on each connected component of $D$. This result enables us to apply Lemma 3, and thus similarly in Section 5.1, we are able to prove a Liouville-type theorem. Hence it follows that a convex solution $u_0$, which is constructed by the previous blow-argument, should be constant. Finally the inclusions $\{0\} \subset \partial u(x) \subset \partial u_0(x) \subset \{0\}$ hold, and this completes the proof of (1.4), that is, $\partial u(x) = \{0\}$.

For maximum principles on the equation $Lu = 0$, the proofs are almost similar to those in Section 4. Indeed, we first recall that the operator $A : \mathbb{R}^n \ni z_0 \mapsto \partial u(x)$. 


Continuity of Derivatives of a Convex Solution 283

\[ \nabla W(z_0) \in \mathbb{R}^n \]

satisfies strict monotonicity (4.8). Combining with monotonicity of the subdifferential operator \( \partial \Psi \), we can easily prove a comparison principle as in Proposition 2. Also, similarly to Lemma 2, we can construct classical barrier subsolutions to \( Lu = 0 \) in an open annulus, since the operator \( L \) is locally uniformly elliptic outside a facet. These results enable us to prove a maximum principle outside a facet.

We are left to justify the remaining two problems: a blow-argument and the Liouville-type theorem. To show them, we have to make use of some basic facts on a convex functional which is homogeneous of degree 1. These fundamental results are contained in Section A.3.

For a blow-up argument as in Section 3, we similarly define rescaled solutions. Existence of a limit of these rescaled functions are guaranteed by the Arzelà–Ascoli theorem and a diagonal argument. By proving Lemma 5 below, we are able to demonstrate that \( u_0 \), a limit of rescaled solutions, is a weak solution to \( Lu = 0 \) in \( \mathbb{R}^n \), and this finishes our blow-up argument.

**Lemma 5.** Let \( U \subset \mathbb{R}^n \) be a bounded domain. Assume that sequences of functions \( \{u_N\}_{N=1}^\infty \subset W^{1,\infty}(U) \) and \( \{f_N\}_{N=1}^\infty \subset L^q(U) (n < q \leq \infty) \) satisfy all of the following:

1. For each \( N \in \mathbb{N} \), \( u_N \) satisfies \( Lu_N = f_N \) in \( W^{-1,\infty}(U) \).
2. There exists a constant \( M > 0 \), independent of \( N \in \mathbb{N} \), such that
   \[ |\nabla u_N(x)| \leq M \text{ for almost everywhere } x \in U. \tag{6.22} \]
3. There exists a function \( u \in W^{1,\infty}(U) \) such that
   \[ \nabla u_N(x) \to \nabla u(x) \text{ for almost everywhere } x \in U. \tag{6.23} \]
4. \( f_N \) strongly converges to 0 in \( L^q(U) \).

Then \( u \) satisfies \( Lu = 0 \) in \( W^{-1,\infty}(U) \).

**Proof.** For each \( N \in \mathbb{N} \), there exists a vector field \( Z_N \in L^\infty(U, \mathbb{R}^n) \) such that

\[ Z_N(x) \in \partial \Psi(\nabla u(x)) \text{ for almost everywhere } x \in U, \quad \tag{6.24} \]

\[ \int_U \langle Z_N | \nabla \phi \rangle \, dx + \int_U \langle A(\nabla u_N) | \nabla \phi \rangle \, dx = \int_U f_N \phi \, dx \quad \text{for all } \phi \in W^{0,1}_0(U). \tag{6.25} \]

Combining the assumption \( f_N \to f \) in \( L^q(U) \) with the continuous embedding \( L^q(U) \to W^{-1,\infty}(U) \), we get

\[ f_N \to 0 \text{ in } W^{-1,\infty}(U). \tag{6.26} \]

By \( A \in C(\mathbb{R}^n, \mathbb{R}^n) \) and (6.22), the vector fields \( \{A(\nabla u_N)\}_{N=1}^\infty \) satisfy

\[ A(\nabla u_N(x)) \to A(\nabla u(x)) \text{ for almost everywhere } x \in U, \]

\[ |A_N(\nabla u_N(x)) - A(\nabla u(x))| \leq C \text{ for almost everywhere } x \in U, \]
Yoshikazu Giga And Shuntaro Tsubouchi

where $C$ is independent of $N \in \mathbb{N}$. From these and Lebesgue’s dominated convergence theorem, it follows that

$$A(\nabla u_N) \rightharpoonup^* A(\nabla u) \quad \text{in } L^\infty(U, \mathbb{R}^n).$$

(6.27)

As mentioned in Remark 5–6, the $\{Z_N\}_{N=1}^\infty \subset L^\infty(U, \mathbb{R}^n)$ is bounded. Hence by [5, Corollary 3.30], we may take a subsequence $\{Z_{N_j}\}_{j=1}^\infty$ so that

$$Z_{N_j} \rightharpoonup^* Z \quad \text{in } L^\infty(U, \mathbb{R}^n)$$

(6.28)

for some $Z \in L^\infty(U, \mathbb{R}^n)$. By (6.25)–(6.28) we obtain

$$\int_U \langle Z | \phi \rangle \, dx + \int_U \langle A(\nabla u) | \nabla \phi \rangle \, dx = 0 \quad \text{for all } \phi \in W^{1,1}_0(U).$$

Now we are left to prove that

$$Z(x) \in \partial \Psi(\nabla u(x)) \quad \text{for almost everywhere } x \in U.$$

By (6.10), it suffices to show that $Z$ satisfies

$$\tilde{\Psi}(Z(x)) \leq 1,$$

(6.29)

$$\Psi(\nabla u(x)) = \langle Z | \nabla u(x) \rangle$$

(6.30)

for almost everywhere $x \in U$. Similarly, it follows that for each $N \in \mathbb{N}$, the vector field $Z_N$ satisfies

$$\begin{cases}
\tilde{\Psi}(Z_N(x)) \leq 1, \\
\Psi(\nabla u_N(x)) = \langle Z_N | \nabla u(x) \rangle,
\end{cases}$$

for almost everywhere $x \in U$. Hence (6.29) is an easy consequence of Lemma 4. We recall (6.2), and thus $\partial \Psi(z_0) = \{\nabla \zeta \Psi(z_0)\}$ holds for all $z_0 \in \mathbb{R}^n \setminus \{0\}$. Combining (6.23), we can check that $Z_N(x) \to Z(x)$ for almost everywhere $x \in D := \{x \in U | \nabla u(x) \neq 0\}$. Hence (6.30) holds for almost everywhere $x \in D$. Note that (6.30) is clear for $x \in U \setminus D$, and this completes the proof.

We prove a Liouville-type theorem as in Theorem 3. In other words, for a convex solution to $Lu = 0$ in $\mathbb{R}^n$, we show that $F$, the facet of $u$, would satisfy either $F = \emptyset$ or $F = \mathbb{R}^n$. Assume by contradiction that $F$ satisfies $\emptyset \subsetneq F \subsetneq \mathbb{R}^n$. Then by Lemma 3, we may write a convex solution $u$ by either of (5.1)–(5.3). However, Lemma 6 below states that $u$ is no longer a weak solution, and this completes our proof.

**Lemma 6.** Let $u$ be a piecewise-linear function defined as in either of (5.1)–(5.3). Then $u$ is not a weak solution to $Lu = 0$ in $\mathbb{R}^n$. 
Proof. As in the proof of Theorem 3, we introduce a constant $d > 0$, and set open cubes $Q' \subset \mathbb{R}^{n-1}$ and $Q, Q_l, Q_r \subset \mathbb{R}^n$. By choosing sufficiently small $d > 0$, we show that $u$ does not satisfy $Lu = 0$ in $W^{-1, \infty}(Q)$. Assume by contradiction that there exists a vector field $Z \in L^\infty(Q, \mathbb{R}^n)$ such that the pair $(u, Z)$ satisfies $Lu = 0$ in $W^{-1, \infty}(Q)$.

We first show that a function $u$ defined as in (5.1) is not a weak solution. For this case, (6.12) implies that $Z$ satisfies $Z(x) = \nabla \psi(e_1)$ for almost everywhere $x \in Q_r$. We take and fix non-negative functions $\phi_1 \in C_c^1((-d, d)), \phi_2 \in C_c^1(Q')$ such that (5.5) holds, and define $\phi \in C_c^1(Q)$ by $\phi(x_1, x') := \phi_1(x_1)\phi_2(x')$ for $(x_1, x') \in (-d, d) \times Q' = Q$. Testing $\phi$ into $Lu = 0$ in $W^{-1, \infty}(Q)$, we have

$$0 = \int_{Q_l} (Z + A(0) | \nabla(\phi_1\phi_2)) \, dx + \int_{Q_r} (\nabla_2 \psi(e_1) + A(t_1e_1) | \nabla(\phi_1\phi_2)) \, dx$$

$$\leq \int_{Q_l} \psi(\nabla(\phi_1\phi_2)) \tilde{\psi}(Z(x)) \, dx$$

$$+ \phi_1(0) \int_{Q'} \phi_2(x') \langle \nabla_2 \psi(e_1) | -e_1 \rangle \, dx' + \phi_1(0) \int_{Q'} \phi_2(x') \langle A(t_1e_1) | -e_1 \rangle \, dx' =: I_1 + I_2 + I_3.$$ 

Here we have used the Cauchy–Schwarz-type inequality (6.5) for the integral over $Q_l$, and applied the Gauss–Green theorem to the integration over $Q_r$. For $I_1$, we make use of (6.9)–(6.8), Fubini’s theorem and (5.5). Then we have

$$I_1 \leq \int_{Q_l} \phi_1(x_1) \psi(0, \nabla_x \phi_2(x')) \, dx + \int_{Q_l} \phi_1'(x_1)\phi_2(x')\psi(e_1) \, dx$$

$$\leq \phi_1(0) \left( d \cdot \|\psi(0, \nabla_x \phi_2)\|_{L^1(Q')} + \psi(e_1)\|\phi_2\|_{L^1(Q')} \right).$$

where $\nabla_x \phi_2 := (\partial_{x_2}\phi_2, \ldots, \partial_{x_n}\phi_2)$. For $I_2$, recalling Euler’s identity (6.11), we get $I_2 = -\phi_1(0)\psi(e_1)\|\phi_2\|_{L^1(Q')}$. We set a constant $\mu := \langle A(t_1e_1) | e_1 \rangle$, which is positive by (6.7). Then we obtain

$$I_1 + I_2 + I_3 \leq \phi_1(0) \left( d \cdot \|\psi(0, \nabla_x \phi_2)\|_{L^1(Q')} - \mu \|\phi_2\|_{L^1(Q')} \right).$$

Choosing $d = d(\mu, \Psi, \phi_2) > 0$ sufficiently small, we have $0 \leq I_1 + I_2 + I_3 < 0$, which is a contradiction. Similarly we can deduce that $u$ defined as in (5.3) does not satisfy $Lu = 0$ in $W^{-1, \infty}(Q)$, since it suffices to restrict $d < l_0$. For the remaining case (5.2), we have already known that

$$Z(x) = \begin{cases} \nabla_2 \psi(e_1) & \text{for almost everywhere } x \in Q_r, \\ \nabla_2 \psi(-e_1) & \text{for almost everywhere } x \in Q_l \end{cases}$$

by definition of $Z$ and (6.12). We set two constants $\mu_1 := \langle A(t_1e_1) | e_1 \rangle, \mu_2 := \langle A(-t_2e_1) | -e_1 \rangle$, both of which are positive by (6.7). Testing the same function
\( \phi \in C^1_c(Q) \) into \( Lu = 0 \) in \( W^{-1, \infty}(Q) \), we obtain
\[
0 = \int_{Q_l} \langle \nabla \psi(-e_1) + A(-t_2e_1) | \nabla (\phi_1\phi_2) \rangle \, dx \\
+ \int_{Q_r} \langle \nabla \psi(e_1) + A(t_1e_1) | \nabla (\phi_1\phi_2) \rangle \, dx \\
= \int_{Q'} \phi_1(0)\phi_2(x') \langle \nabla \psi(-e_1) + A(-t_2e_1) | e_1 \rangle \, dx' \\
+ \int_{Q'} \phi_1(0)\phi_2(x') \langle \nabla \psi(e_1) + A(t_1e_1) | -e_1 \rangle \, dx' \\
= -\phi_1(0) (\psi(e_1) + \psi(-e_1) + \mu_1 + \mu_2) \int_{Q'} \phi_2(x') \, dx' < 0,
\]
which is a contradiction. Here we have used the Gauss–Green theorem and Euler’s identity (6.11). This completes the proof.

Acknowledgements. The first author is partly supported by the Japan Society for the Promotion of Science through grants Kiban A (No. 19H00639). Challenging Pioneering Research (Kaitaku) (No. 18H05323), Kiban A (No. 17H01091).

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

A. Proofs for a Few Basic Facts

In this section, we give proofs for a few basic facts used in this paper for completeness.

A.1. A Poincaré-Type inequality

We give a precise proof of Lemma 7, a Poincaré-type inequality for difference quotients of functions in \( W^{1, p}_0(\Omega) \) (\( 1 \leq p < \infty \)). This result is used in the proof of Lemma 1. The proof of Lemma 7 is essentially a modification of that of the Poincaré inequality for the Sobolev space \( W^{1, p}_0 \) [10, Proposition 3.10].

Lemma 7. Let \( \Omega \subset \mathbb{R}^n \) be a bounded open set and \( 1 \leq p < \infty \). For all \( u \in W^{1, p}_0(\Omega) \), \( j \in \{ 1, \ldots, n \} \), \( h \in \mathbb{R} \setminus \{ 0 \} \), we have
\[
\| \Delta_j h u \|_{L^p(\Omega)} \leq \| \nabla u \|_{L^p(\Omega)}.
\] (A.1)
Here $\Delta_j, h u$ is defined by

$$\Delta_j, h u(x) := \frac{\overline{u}(x + h e_j) - u(x)}{h} \quad \text{for } x \in \Omega.$$ 

Before the proof of Lemma 7, we note that $\Delta_j, h u(x)$ makes sense for almost everywhere $x \in \Omega$ by the zero extension of $u \in W_0^{1, p}(U)$. That is, for a given $u \in W_0^{1, p}(U)$, we set $\overline{u} \in W^{1, p}(\mathbb{R}^n)$ by

$$\overline{u}(x) := \begin{cases} u(x) & x \in U, \\ 0 & x \in \mathbb{R}^n \setminus U. \end{cases} \quad (A.2)$$

**Proof.** We fix $j \in \{1, \ldots, n\}$, $h \in \mathbb{R} \setminus \{0\}$. We first note that the operator $\Delta_j, h : W_0^{1, p}(U) \to L^p(U)$ is bounded, since for all $u \in W_0^{1, p}(U)$ we have

$$\|\Delta_j, h u\|_{L^p(U)} \leq \frac{1}{|h|} \left[ \left( \int_U |\overline{u}(x + h)|^p \, dx \right)^{1/p} + \left( \int_U |u(x)|^p \, dx \right)^{1/p} \right]$$

by the Minkowski inequality and the Poincaré inequality. Here $\overline{u} \in W^{1, p}(\mathbb{R}^n)$ is defined as in (A.2). Hence by a density argument, it suffices to check that (A.1) holds true for all $u \in C^\infty_c(U)$. Let $u \in C^\infty_c(U)$. Then for all $x \in U$, we have

$$|\overline{u}(x + h e_j) - u(x)| = \left| \int_0^1 \langle \nabla \overline{u}(x + t h e_j), h e_j \rangle \, dt \right|$$

$$\leq |h| \int_0^1 |\nabla \overline{u}(x + t h e_j)| \, dt \leq |h| \left( \int_0^1 |\nabla \overline{u}(x + t h e_j)|^p \, dt \right)^{1/p}$$

by the Cauchy-Schwarz inequality and Hölder’s inequality. From this estimate we get

$$\|\Delta_j, h u\|_{L^p(U)} \leq \int_U \int_0^1 |\nabla \overline{u}(x + t h e_j)|^p \, dt \, dx$$

$$= \int_0^1 \int_U |\nabla \overline{u}(x + t h e_j)|^p \, dx \, dt \quad \text{(by Fubini’s theorem)}$$

$$\leq \|\nabla u\|_{L^p(U)}^p.$$ 

Hence we obtain (A.1) for all $u \in C^\infty_c(U)$, and this completes the proof. \hfill \Box

### A.2. Convex analysis

Lemma 8 is used in the proof of Lemma 1 for a justification of local $W^{2, p}$-regularity of a convex weak solution outside of the facet.

**Lemma 8.** Let $u$ be a real-valued convex function in a convex domain $\Omega \subset \mathbb{R}^n$. Assume that $x_1, x_2 \in \Omega$ satisfy $x_1 \neq x_2$, and set $d := |x_2 - x_1| > 0$, $\nu := d^{-1}(x_2 - x_1)$. Then for all $z_2 \in \partial u(x_2)$, we have

$$\langle z_2 | \nu \rangle \geq \frac{u(x_2) - u(x_1)}{d}. \quad (A.3)$$
Proof. By $z_2 \in \partial u(x_2)$, we have a subgradient inequality

$$u(x) \geq u(x_2) + \langle z_2 \mid x - x_2 \rangle$$

for all $x \in \Omega$. Substituting $x := x_1 = x_2 - d\nu \in \Omega$, we obtain

$$u(x_1) \geq u(x_2) - d\langle z_2 \mid \nu \rangle,$$

which yields (A.3). □

Remark 7. Instead of subgradient inequalities, we are able to show (A.3) by monotonicity of $\partial u$. For each fixed $x_1, x_2 \in \Omega$ with $x_1 \neq x_2$, we may take and fix $x_3 := x_1 + t(x_2 - x_1)$ for some $0 < t < 1$ and $z_3 \in \partial u(x_3)$ such that

$$u(x_2) - u(x_1) = \langle z_3 \mid x_2 - x_1 \rangle,$$

with the aid of the mean value theorem for non-smooth convex functions [1, Theorem D.6]. $x_2 - x_1 = d\nu$ is clear by definitions of $d, \nu$. Noting $x_2 - x_3 = (1 - t)d\nu$, we can check that

$$\langle z_2 - z_3 \mid \nu \rangle = \frac{1}{1-t}d\langle z_2 - z_3 \mid x_2 - x_3 \rangle \geq 0$$

by monotonicity of $\partial u$. Combining these results with (A.4), we obtain

$$u(x_2) - u(x_1) = d\langle z_3 \mid \nu \rangle \leq d\langle z_2 \mid \nu \rangle,$$

which yields (A.3).

The following lemma is used in the proof of Proposition 1.

Lemma 9. Let $U \subset \mathbb{R}^n$ be a convex open set, and let $\{u_N\}_{N=1}^{\infty}$ be a sequence of real-valued convex functions in $U$. Assume that this sequence is uniformly Lipschitz. In other words, there is a constant $L > 0$ independent of $N \in \mathbb{N}$ such that

$$|u_N(x) - u_N(y)| \leq L|x - y| \text{ for all } x, y \in U. \quad \text{(A.5)}$$

If there exists a function $u_\infty : U \rightarrow \mathbb{R}$ such that

$$u_N(x) \rightarrow u_\infty(x) \text{ for all } x \in U, \quad \text{(A.6)}$$

then we have $\nabla u_N(x) \rightarrow \nabla u_\infty(x)$ for almost everywhere $x \in U$.

Remark 8. From (A.5)–(A.6), it is easy to show that $u_\infty$ is also convex, $u_N \rightarrow u_\infty$ uniformly in $U$, and

$$|u_\infty(x) - u_\infty(y)| \leq L|x - y| \text{ for all } x, y \in U.$$

Our proof of Lemma 9 is inspired by [11, Lemma A.3].

Proof. We define $\mathcal{L}^n$-measurable sets

$$P_N := \{x \in U \mid u_N \text{ is not differentiable at } x\} \text{ for } N \in \mathbb{N} \cup \{\infty\}. $$

Clearly $P_N (N \in \mathbb{N} \cup \{\infty\})$ satisfies $\mathcal{L}^n(P_N) = 0$ by Lipschitz continuity of $u_N$, and therefore the $\mathcal{L}^n$-measurable set

$$P := \bigcup_{N \in \mathbb{N} \cup \{\infty\}} P_N \subset U$$
also satisfies $L^g(P) = 0$. We claim that
\[ \nabla u_N(x_0) \to \nabla u_\infty(x_0) \quad \text{for all } x_0 \in U \setminus P. \] (A.7)

We take and fix arbitrary $x_0 \in U \setminus P$. We note that $\nabla u_N(x_0)$ exists for each $N \in \mathbb{N}$ since $x_0 \not\in P_N$, and we obtain
\[ \sup_{N \in \mathbb{N}} |\nabla u_N(x_0)| \leq L \]
with the aid of (A.5). Hence it suffices to check that, if a subsequence $\{u_{N_k}\}_k \subset \{u_N\}_N$ satisfies
\[ \nabla u_{N_k}(x_0) \to v (k \to \infty) \quad \text{for some } v \in \mathbb{R}^N, \] (A.8)
then $v = \nabla u_\infty(x_0)$. Since $x_0 \not\in P_{N_k}$ and therefore $\partial u_{N_k}(x_0) = \{\nabla u_{N_k}(x_0)\}$ for each $k \in \mathbb{N}$, we easily get
\[ u_{N_k}(x) \geq u_{N_k}(x_0) + \langle \nabla u_{N_k}(x_0) | x - x_0 \rangle \quad \text{for all } x \in U, k \in \mathbb{N}. \]

Letting $k \to \infty$, we have
\[ u_\infty(x) \geq u_\infty(x_0) + \langle v | x - x_0 \rangle \quad \text{for all } x \in U \]
by (A.6) and (A.8). This means that $v \in \partial u_\infty(x_0)$. Note again that $x_0 \not\in P_\infty$ and therefore $\partial u_\infty(x_0) = \{\nabla u_\infty(x_0)\}$, which yields $v = \nabla u_\infty(x_0)$. This completes the proof of (A.7).

\[ \square \]

### A.3. Convex functionals

We prove some basic property of convex functionals $\Psi$ and $W$ in Section 6.

**Lemma 10.** Let $W$ be a convex function which satisfies (6.2)–(6.3) and (6.5). Then the mapping $A: \mathbb{R}^n \ni z \mapsto \nabla W(z) \in \mathbb{R}^n$ satisfies strict monotonicity (4.8).

**Proof:** We take arbitrary $z_1, z_2 \in \mathbb{R}^n$ with $z_1 \neq z_2$ and define a line segment $L := \{z_1 + t(z_2 - z_1) \in \mathbb{R}^n | 0 \leq t \leq 1\}$.

We first consider the case $0 \not\in L$. Then there exist constants $0 < \mu \leq M < \infty$ such that $\mu \leq |z_0| \leq M$ holds for all $z_0 \in L$. Here we can take a constant $\gamma > 0$ such that (6.3) holds for all $z_0 \in L$. Then by $W \in C^2(\mathbb{R}^n \setminus \{0\})$, we have
\[ \langle A(z_1) - A(z_2) | z_2 - z_1 \rangle = \int_0^1 \left\langle \frac{\partial^2 W}{\partial z^2} (z_1 + t(z_2 - z_1))(z_2 - z_1) | z_2 - z_1 \right\rangle \, dt \geq \gamma |z_2 - z_1|^2 > 0. \]

To consider the remaining case $0 \in L$, it suffices to show (6.7). Indeed, the assumption $0 \in L$ allows us to write $z_1 = -l_1 v, z_2 = l_2 v$ for some unit vector $v$ and some constants $l_1, l_2 \geq 0$. Under this notation, we obtain
\[ \langle A(z_2) - A(z_1) | z_2 - z_1 \rangle = \langle A(l_2 v) | (l_1 + l_2)v \rangle + \langle A(-l_1 v) | -(l_1 + l_2)v \rangle > 0 \]
by (6.7). Here we note that at least one of $l_1, l_2$ is positive since $l_1 + l_2 = |z_2 - z_1| > 0$.

We prove (6.7) to complete the proof. Let $z \in \mathbb{R}^n \setminus \{0\}$. Then we obtain
\[ d_N := \left\langle A(z/2^{N-1}) - A(z/2^N) \right\rangle > 0 \]
for each $N \in \mathbb{N}$, since we have already shown (4.8) for the case $0 \not\in L$. By definition of $d_j (j \in \mathbb{N})$, it is clear that
\[ \left\langle A(z) - A\left(z/2^N\right) \right\rangle = d_1 + \cdots + d_N \geq d_1. \]

Letting $N \to \infty$, we obtain $\langle A(z) | z \rangle \geq d_1 > 0$ by $A \in C(\mathbb{R}^n, \mathbb{R}^n)$.

\[ \square \]
We precisely prove (6.8)–(6.10) in Lemma 11. See also [2, Section 1.3] and [30, §13] as related items.

**Lemma 11.** Let $\Psi : \mathbb{R}^n \to [0, \infty)$ be a convex function which is positively homogeneous of degree 1.

1. $\Psi$ satisfies the triangle inequality (6.8).
2. Assume that $\zeta \in \mathbb{R}^n$ satisfies $\tilde{\Psi}(\zeta) < \infty$. Then the Cauchy–Schwarz-type inequality (6.9) holds.
3. The subdifferential operator $\partial \Psi$ is given by (6.10).

**Proof.** By convexity of $\Psi$ and (6.6), $\Psi$ satisfies

\[
\frac{\Psi(z_1 + z_2)}{2} = \Psi\left(\frac{z_1 + z_2}{2}\right) \leq \frac{\Psi(z_1) + \Psi(z_2)}{2}
\]

for all $z_1, z_2 \in \mathbb{R}^n$,

which yields (6.8).

We next show the Cauchy–Schwarz inequality (6.9). Let $z \in \mathbb{R}^n$. If $\Psi(z) > 0$, then we have

\[
\langle z \mid \zeta \rangle = \Psi(z)\left(\frac{z}{\Psi(z)} \mid \frac{\zeta}{\Psi(\zeta)}\right) \leq \Psi(z)\tilde{\Psi}(\zeta)
\]

by $z/\Psi(z) \in C_{\Psi}$. For the case $\Psi(z) = 0$, we note that $\lambda z \in C_{\Psi}$ for all $\lambda > 0$. Hence it follows that

\[
\langle z \mid \zeta \rangle = \frac{\langle \lambda z \mid \zeta \rangle}{\lambda} \leq \frac{\tilde{\Psi}(w)}{\lambda}
\]

for all $\lambda > 0$. By $\tilde{\Psi}(\zeta) < \infty$, we obtain $\langle z \mid \zeta \rangle \leq 0 = \Psi(z)\tilde{\Psi}(\zeta)$. This completes the proof of (6.9).

Finally we prove (6.10). Let $z_0 \in \mathbb{R}^n$ be arbitrarily fixed. Assume that $\zeta \in \mathbb{R}^n$ satisfies $\tilde{\Psi}(\zeta) \leq 1$ and $\Psi(z_0) = \langle z_0 \mid \zeta \rangle$. Then by combining these assumptions with (6.9), we have

\[
\Psi(z) \geq \Psi(z_0) + \langle \zeta \mid z - z_0 \rangle
\]

for all $z \in \mathbb{R}^n$. Hence $\zeta \in \partial \Psi(z_0)$. Conversely, if $\zeta \in \partial \Psi(z_0)$, then we have the subgradient inequality

\[
\Psi(z) \geq \Psi(z_0) + \langle \zeta \mid z - z_0 \rangle
\]

for all $z \in \mathbb{R}^n$. Combining with (A.9), we have

\[
\Psi(z) \geq \Psi(z_0) + \langle \zeta \mid z - z_0 \rangle
\]

for all $z \in \mathbb{R}^n$. By testing $kz_0$ into (A.9), where $k \in [0, \infty)$ is arbitrary, we have

\[
(k - 1)\Psi(z_0) = \Psi(kz_0) - \Psi(z_0) \geq \langle \zeta \mid (k - 1)z_0 \rangle = (k - 1)\langle \zeta \mid z_0 \rangle.
\]

If we let $0 \leq k < 1$ so that $k - 1 < 0$, then we have $\Psi(z_0) \leq \langle \zeta \mid z_0 \rangle$. Similarly, letting $1 < k < \infty$, we have $\Psi(z_0) \geq \langle \zeta \mid z_0 \rangle$. Hence we obtain $\Psi(z_0) = \langle \zeta \mid z_0 \rangle$. Combining with (A.9), we have

\[
\langle z \mid \zeta \rangle \leq \Psi(z)
\]

for all $z \in \mathbb{R}^n$, which yields $\tilde{\Psi}(\zeta) \leq 1$ by definition of $\tilde{\Psi}$. This completes the proof of (6.10). □
References

1. Ambrosio, L., Carlotto, A., Massaccesi, A.: Lectures on elliptic partial differential equations, volume 18 of Appunti. Scuola Normale Superiore di Pisa (Nuova Serie) [Lecture Notes. Scuola Normale Superiore di Pisa (New Series)]. Edizioni della Normale, Pisa, 2018

2. Andreu-Vaillo, F., Caselles, V., Mazón, J.M.: Parabolic quasilinear equations minimizing linear growth functionals, vol. 223. Progress in Mathematics. Birkhäuser Verlag, Basel, 2004

3. Bertsekas, D.P.: Convex optimization theory. Athena Scientific, Nashua, NH (2009)

4. Brézis, H.: Monotonicity methods in Hilbert spaces and some applications to nonlinear partial differential equations. In Contributions to Nonlinear Functional Analysis (Proc. Sympos., Math. Res. Center Univ. Wisconsin, Madison, Wis., 1971), 101–156, 1971

5. Brezis, H.: Functional analysis. Sobolev spaces and partial differential equations. Universitext. Springer, New York (2011)

6. Colombo, M., Mingione, G.: Bounded minimisers of double phase variational integrals. Arch. Ration. Mech. Anal. 218(1), 219–273, 2015

7. Colombo, M., Mingione, G.: Regularity for double phase variational problems. Arch. Ration. Mech. Anal. 215(2), 443–496, 2015

8. Duvaut, G., Lions, J.-L.: Inequalities in mechanics and physics, volume 219 of Grundlehren der Mathematischen Wissenschaften. Springer, Berlin, New York, 1976. Translated from the French by C. W. John.

9. Evans, L.C., Gariepy, R.F.: Measure theory and fine properties of functions, revised. Textbooks in Mathematics. CRC Press, Boca Raton, FL (2015)

10. Giaquinta, M., Martinazzi, L.: An introduction to the regularity theory for elliptic systems, harmonic maps and minimal graphs, volume 11 of Appunti. Scuola Normale Superiore di Pisa (Nuova Serie) [Lecture Notes. Scuola Normale Superiore di Pisa (New Series)]. Edizioni della Normale, Pisa, second edition, 2012

11. Giga, M.-H., Giga, Y.: Stability for evolving graphs by nonlocal weighted curvature. Comm. Partial Differential Equations 24(1–2), 109–184, 1999

12. Giga, M.-H., Giga, Y.: Very singular diffusion equations: second and fourth order problems. Jpn. J. Ind. Appl. Math. 27(3), 323–345, 2010

13. Giga, Y., Kohn, R.V.: Scale-invariant extinction time estimates for some singular diffusion equations. Discrete Contin. Dyn. Syst. 30(2), 509–535, 2011

14. Giga, Y., Kuroda, H., MatsuoKA, H.: Fourth-order total variation flow with Dirichlet condition: characterization of evolution and extinction time estimates. Adv. Math. Sci. Appl. 24(2), 499–534, 2014

15. Giga, Y., Muszkieta, M., Rybka, P.: A duality based approach to the minimizing total variation flow in the space $H^{-s}$. Jpn. J. Ind. Appl. Math. 36(1), 261–286, 2019

16. Giga, Y., Ueda, Y.: Numerical computations of split Bregman method for fourth order total variation flow. J. Comput. Phys., 405:109114, 2020

17. Gilbarg, D., Trudinger, N.S.: Elliptic Partial Differential Equations of Second Order. Classics in Mathematics. Springer, Berlin, 2001. Reprint of the 1998 edition.

18. Giusti, E.: Direct methods in the calculus of variations. World Scientific Publishing Co., Inc, River Edge, NJ (2003)

19. Glowinski, R., Lions, J.-L., Trémolières, R.: Numerical analysis of variational inequalities, Volume 8 of Studies in Mathematics and its Applications. North-Holland Publishing Co., Amsterdam, New York, 1981. Translated from the French.

20. Hoff, E.: A remark on linear elliptic differential equations of second order. Proc. Amer. Math. Soc. 3, 791–793, 1952

21. Kashima, Y.: A subdifferential formulation of fourth order singular diffusion equations. Adv. Math. Sci. Appl. 14(1), 49–74, 2004

22. Kashima, Y.: Characterization of subdifferentials of a singular convex functional in Sobolev spaces of order minus one. J. Funct. Anal. 262(6), 2833–2860, 2012
23. Kohn, R.V.: Surface relaxation below the roughening temperature: some recent progress and open questions. In *Nonlinear partial differential equations*, Volume 7 of *Abel Symp.*, 207–221. Springer, Heidelberg, 2012.

24. Kohn, R.V., Versieux, H.M.: Numerical analysis of a steepest-descent PDE model for surface relaxation below the roughening temperature. *SIAM J. Numer. Anal.* 48(5), 1781–1800, 2010

25. Krügel, F.: A variational problem leading to a singular elliptic equation involving the 1-Laplacian. Mensch und Buch Verlag, Berlin (2013)

26. Mingione, G.: Regularity of minima: an invitation to the dark side of the calculus of variations. *Appl. Math.* 51(4), 355–426, 2006

27. Odisharia, I.V.: *Simulation and Analysis of the Relaxation of a Crystalline Surface*. New York University, 2006.

28. Protter, M.H., Weinberger, H.F.: Maximum principles in differential equations. Prentice-Hall Inc, Englewood Cliffs, N.J. (1967)

29. Pucci, P., Serrin, J.: The maximum principle, vol. 73. *Progress in Nonlinear Differential Equations and their Applications*. Birkhäuser Verlag, Basel, 2007

30. Rockafellar, R.T.: *Convex Analysis*. Princeton Landmarks in Mathematics. Princeton University Press, Princeton, NJ, 1997. Reprint of the 1970 original, Princeton Paperbacks

31. Spohn, H.: Surface dynamics below the roughening transition. *J. Phys. I* 3(1), 69–81, 1993

32. Tsubouchi, S.: Local Lipschitz bounds for solutions to certain singular elliptic equations involving the one-Laplacian. *Calc. Var. Partial Differential Equations*, 60(1):Paper No. 33, 35, 2021

33. Xu, X.: Mathematical validation of a continuum model for relaxation of interacting steps in crystal surfaces in 2 space dimensions. *Calc. Var. Partial Differential Equations* 59(5), 158, 2020

Y. Giga, S. Tsubouchi
Graduate School of Mathematical Sciences,
The University of Tokyo,
3-8-1 Komaba Meguro-ku,
Tokyo
153-8914 Japan.
e-mail: tsubos@ms.u-tokyo.ac.jp

e-mail: labgiga@ms.u-tokyo.ac.jp

(Received August 23, 2021 / Accepted February 11, 2022)
Published online 18 March 2022
© The Author(s) (2022),
corrected publication 2022