The Maximum Principle for Viscosity Solutions of Fully Nonlinear Second Order Partial Differential Equations

ROBERT JENSEN

Communicated by C. M. DAFERMOS

Abstract

We prove that viscosity solutions in $W^{1,\infty}$ of the second order, fully nonlinear, equation $F(D^2u, Du, u) = 0$ are unique when (i) F is degenerate elliptic and decreasing in u or (ii) F is uniformly elliptic and nonincreasing in u. We do not assume that F is convex. The method of proof involves constructing nonlinear approximation operators which map viscosity subsolutions and supersolutions onto viscosity subsolutions and supersolutions, respectively. This method is completely different from that used in LIONS [8, 9] for second order problems with F convex in D^2u and from that used by CRANDALL & LIONS [3] and CRANDALL, EVANS & LIONS [2] for fully nonlinear first order problems.

0. Introduction

This paper considers the uniqueness of solutions of nonlinear second order elliptic partial differential equations,

$$F(D^2u, Du, u, x) = 0 \quad \text{in } \Omega; \tag{0.1}$$

with Dirichlet boundary condition

$$u = g \quad \text{on } \partial\Omega. \tag{0.2}$$

We implicitly assume throughout this paper that Ω is a bounded domain in R^n, g is continuous on $\partial\Omega$, and solutions of (0.1) and (0.2) are always in $C(\overline{\Omega})$. Uniqueness for such problems depends both on the properties of the function F and on the space in which solutions are taken. For example, it is easily shown by use of the classical maximum principle that if F is C^1, uniformly elliptic, and decreasing in u, then all solutions of (0.1) and (0.2) in $C^2(\Omega)$ are equal. The primary deficiency in this result is the assumption that solutions are in $C^2(\Omega)$. The best general existence and regularity theorems of which I am aware yield solutions of (0.1) and

(0.2) in $C^2(\Omega)$ only under the additional assumption that F is convex (or concave) in D^2u (see L. C. EVANS [6], for example). Another deficiency in the previous uniqueness result is the regularity imposed on F. An improved uniqueness theorem can be achieved by using the extension of the classical maximum principle in J. M. BONY [1], whence it follows that if F is Lipschitz continuous, uniformly elliptic, and decreasing in u, then all solutions of (0.1) and (0.2) in $W^{2,p}(\Omega)$ are equal. This result is still restrictive. Until recently, however, the most general definition of solution of (0.1) for fully nonlinear F, *i.e.*, for F nonlinear in D^2u, required that the solution have two Sobolev derivatives and that (0.1) hold almost everywhere. The difficulty in broadening the concept of solution in the way it is usually done for linear equations is that for fully nonlinear F's it is generally impossible to integrate $F(D^2u, Du, u, x) \cdot \varphi(x)$ by parts and remove the derivatives of u from the resulting integral. Consequently, a new approach to the idea of weak solution for (0.1) is required.

In 1983 M. G. CRANDALL & P. L. LIONS [3] introduced the definition of *viscosity solution* as a notion of weak solution for nonlinear first order partial differential equations,

$$H(Du, u, x) = 0 \quad \text{in } \Omega. \tag{0.3}$$

Under assumptions more general than previously known, they were able to establish global uniqueness and existence of viscosity solutions. Further, they showed that classical solutions are always viscosity solutions. A closely related idea had appeared several years earlier. In L. C. EVANS [4] and [5] an L^∞ Minty technique was developed. This technique was (in another guise) an implicit use of the defining property of viscosity solutions. Finally in P.-L. LIONS [8] the definition of viscosity solution was extended to second order problems, *i.e.*, to (0.1). Under some assumptions of regularity on F, which include convexity, it was proved that viscosity solutions are unique. In fact, what P.-L. LIONS actually proved is that the viscosity solution is the value function for some associated stochastic optimization problem.

In this paper we prove a maximum principle for viscosity solutions which implies the following theorem.

Theorem. *If F is continuous, does not depend on x, and either*
(i) *F is degenerate elliptic and uniformly decreasing in u*
or
(ii) *F is uniformly elliptic and nonincreasing in u*
then all viscosity solutions of (0.1) *and* (0.2) *in* $W^{1,\infty}(\Omega)$ *are equal.*

The only other uniqueness theorems I know of are given in P.-L. LIONS [8] and [9]. Next, we compare the present result with those in [8] and [9]. Both [8] and [9] make the assumptions that F is convex (or concave) in (D^2u, Du, u) and that F grows linearly in (D^2u, Du, u). These assumptions appear to be necessary for the method of proof employed there. Additionally both [8] and [9] assume that F is uniformly decreasing in u, which is crucial for degenerate elliptic F's but probably not necessary if F is uniformly elliptic. Finally both [8] and [9] assume a technical condition which is closely related to having $(F(D^2u, Du, u, x) - F(0, 0, 0, x))$

uniformly Lipschitz continuous. On the other hand, spatial dependence in F is allowed in [8] and [9]. Our approach, however, does not allow spatial dependence in x although several subsequent papers by different authors on this are now in varying degrees of preparation. Furthermore, the results in [8] apply to all continuous viscosity solutions not just those in $W^{1,\infty}(\Omega)$ (this, however, does require additional regularity of F).

The techniques we use in this paper are new and novel. So far, all attempts to use modifications of the techniques in [3] to prove general uniqueness results for (0.1) and (0.2) have failed. A heuristic analysis of the techniques in [3] suggests that they are inherently inadequate for second order problems. If w and v are two viscosity solutions of (0.3), the maximum of $v - w$ is studied in [3] with the aid of the auxiliary function

$$Q^\varepsilon(x, y) = v(x) - w(y) - \frac{|x - y|^2}{\varepsilon}. \tag{0.4}$$

If $(x_\varepsilon, y_\varepsilon)$ are points where the maximum of Q^ε occurs and if $(x_{\varepsilon_i}, y_{\varepsilon_i}) \to (x_0, y_0)$ for a sequence $\varepsilon_i \searrow 0$, then $x_0 = y_0$ and $v - w$ has a maximum at x_0. The basic idea in [3] is to use the definition of viscosity solution and the auxiliary function, Q^ε, to preserve the information that is carried in $Dv(x)$ and $Dw(x)$ when v and w are smooth. In fact, if v and w are smooth then $Dv(x_\varepsilon) = Dw(y_\varepsilon)$ and consequently for a sequence $\varepsilon_i \searrow 0$ such that $(x_{\varepsilon_i}, y_{\varepsilon_i}) \to (x_0, x_0)$ we see $Dv(x_0) = Dw(x_0)$ (a fact which also follows from knowing that $v - w$ has a maximum at x_0). Unfortunately, the appropriate information from the second derivatives is not carried by Q^ε. All that we can conclude is that $D^2v(x_\varepsilon) - D^2w(y_\varepsilon) - \frac{2}{\varepsilon} I$ is negative semidefinite. For sequences $\varepsilon_i \searrow 0$ and $(x_{\varepsilon_i}, y_{\varepsilon_i}) \to (x_0, x_0)$ we obtain no information about $D^2v(x_0) - D^2w(x_0)$ from the Q^ε's. My opinion is that the techniques of [3] cannot be adapted because they study the structure of the viscosity solutions at isolated points (namely, the maxima of Q^ε). Thus they are too local to account for second derivatives. The approach I employ incorporates the structure of the viscosity solutions on neighborhoods of points.

I have organized the material in this paper into three sections. In the first section we construct two approximation operators $A_\varepsilon^+[u] \equiv u_\varepsilon^+ \geqq u \geqq u_\varepsilon^- \equiv A_\varepsilon^-[u]$. The operators are constructed from the distance function in the ambient space of graph(u). The range of A_ε^+ is the space of semiconcave functions and the range of A_ε^- is the space of semiconvex functions. These operators are nonlinear and do not map into $C^\infty(\Omega)$. The important point, however, is that A_ε^+ takes viscosity subsolutions of (0.1) into viscosity subsolutions and A_ε^- takes viscosity supersolutions of (0.1) into viscosity supersolutions. This property is in fact the main result of Section 2. In order to prove this we use one of the equivalent formulations of viscosity solution found in M. G. Crandall, L. C. Evans, & P.-L. Lions [2], L. C. Evans [4] and [5], and P.-L. Lions [8]. In Section 3 we combine results on differentiation of semiconvex and semiconcave functions and modifications of ideas from J. M. Bony [1] and C. Pucci [10]. The result is obtained by purely analytic techniques.

Finally, let me also mention that I have endeavored to keep this paper selfcontained. As a result I have given proofs to some results which may already be known. The proofs, however, may be new.

In closing I wish to thank E. N. BARRON and L. C. EVANS for their suggestions and advice. Their help has been greatly appreciated and has certainly helped make this paper readable.

1. Construction of Approximations

The purpose of this section is to construct special, nonstandard, regularized approximations to a continuous function, u, defined on a bounded, open domain Ω in R^n; and to develop some of the unique properties that these approximations possess. This construction and development cause us to frequently alternate in viewing the ambient space as either R^n or R^{n+1}. Some of our notation and definitions are used in either space; the meaning of these cases is clear from the context, *e.g.*, $B(z, r)$ could denote the ball in R^n or R^{n+1}, and the proper choice is made by knowing the space in which z lies. The letters x and ξ are used exclusively for points in R^n and we often denote points in R^{n+1} as ordered pairs (x, y) with x in R^n and y in R. The derivatives of a smooth function $\phi(z)$ are denoted by $D\phi(z)$, $D^2\phi(z), \ldots$; although we also frequently view $D\phi(z)$ as the vector valued function $\left(\frac{\partial\phi}{\partial z_1}(z), \ldots, \frac{\partial\phi}{\partial z_m}(z)\right)$. If λ is a unit vector then $D_\lambda\phi(z) = D\phi(z)(\lambda)$; if $z = (x, y)$ then $D_x\phi(x, y) = D\phi(x, y) \circ i$ and $D_y\phi(x, y) = D\phi(x, y) \circ j$ where $i: R^n \to R^{n+1}$ is the injection $i(\xi) = (\xi, 0)$ and $j: R \to R^{n+1}$ is the injection $j(\eta) = (0, \eta)$.

We shall now begin the construction of approximations to a function $u \in C(\overline{\Omega})$.

Definition 1.1. Given $u \in C(\overline{\Omega})$ define $Q \in C(R^{n+1})$ by

$$Q(x, y) = \text{distance}\,((x, y), \text{graph}\,(u));$$

define the sets Ω_ε and $\mathcal{O}$ by

$$\Omega_\varepsilon = \{x \in R^n \mid \text{distance}\,(x, (R^n \setminus \Omega)) > \varepsilon\};$$

and

$$\mathcal{O} = \{(x, y) \in R^{n+1} \mid x \in \Omega_\varepsilon \text{ and } Q(x, y) \leqq \varepsilon\}.$$

The function Q has the following important properties.

Lemma 1.2. *Q is Lipschitz continuous and satisfies the following inequalities:*

$$\left.\begin{array}{l} |DQ| \leqq 1 \text{ a.e.;} \\ D_\lambda^2 Q(x, y) \leqq \dfrac{1}{Q(x, y)} \quad \text{in } R^{n+1} \setminus \text{graph}\,(u) \\ \text{for any direction } \lambda \text{ (in the sense of distributions).} \end{array}\right\} \tag{1.3}$$

Proof. That Q is Lipschitz continuous follows easily after characterizing Q as

$$Q(x, y) = \inf_{z \in \text{graph}(u)} d_z(x, y) \tag{1.4}$$

where $d_z(\zeta) = |\zeta - z|$. In fact, Q is the limit of smooth approximations $Q_p \geqq Q$ given by

$$Q_p(x, y) = \left[\int_\Omega \{d_{(\xi,u(\xi))}(x, y)\}^{-p} \, d\xi\right]^{-\frac{1}{p}}, \quad \text{for } (x, y) \notin \text{graph}\,(u). \tag{1.5}$$

We next calculate

$$DQ_p(x, y) = (Q_p(x, y))^{p+1} \int_\Omega |(x - \xi, y - u(\xi))|^{-p-2} (x - \xi, y - u(\xi)) \, d\xi.$$

An elementary estimate yields

$$\begin{aligned} |DQ_p(x, y)| &\leqq (Q_p(x, y))^{p+1} \int_\Omega |(x - \xi, y - u(\xi))|^{-p-1} \, d\xi \\ &\leqq (Q_p(x, y))^{p+1} \left[\int_\Omega \{d_{(\xi,u(\xi))}(x, y)\}^{-p} \, d\xi\right] (Q(x, y))^{-1} \\ &\leqq Q_p^{p+1} \cdot Q_p^{-p} \cdot Q^{-1}. \end{aligned}$$

Therefore $|DQ_p(x, y)| \leqq \dfrac{Q_p(x, y)}{Q(x, y)}$. Since $Q_p(x, y) \to Q(x, y)$ uniformly on compact subsets of R^{n+1}, we conclude $|DQ_p(x, y)| \leqq 1 + o(1)$ in $R^{n+1} \setminus$ graph (u). (Thus, the first inequality in 1.3) is established by letting $p \to \infty$.

In order to complete the proof of Lemma 1.2, we calculate $D_\lambda^2 Q_p(x, y)$ for an arbitrary direction λ. We find

$$\begin{aligned} D_\lambda^2 Q_p(x, y) &= (p + 1) (Q_p(x, y))^{2p+1} \\ &\times \left[\int_\Omega |(x - \xi, y - u(\xi))|^{-p-2} (\lambda \cdot (x - \xi, y - u(\xi))) \, d\xi\right]^2 \\ &- (p + 2) (Q_p(x, y)^{p+1} \left[\int_\Omega |(x - \xi, y - u(\xi))|^{-p-4} (\lambda \cdot (x - \xi, y - u(\xi)))^2 \, d\xi\right] \\ &+ (Q_p(x, y))^{p+1} \int_\Omega |(x - \xi, y - u(\xi))| \, d\xi.^{-p-2} \end{aligned}$$

An application of Hölder's inequality to the first term on the right together with (1.4) and (1.5) yields

$$\begin{aligned} D_\lambda^2 Q_p(x, y) &\leqq (p + 1) (Q_p(x, y))^{2p+1} \left[\int_\Omega |(x - \xi, y - u(\xi))|^{-p} \, d\xi\right] \\ &\times \left[\int_\Omega |(x - \xi, y - u(\xi))|^{-p-4} (\lambda \cdot (x - \xi, y - u(\xi)))^2 \, d\xi\right] \\ &- (p + 2) (Q_p(x, y))^{p+1} \left[\int_\Omega |(x - \xi, y - u(\xi))|^{-p-4} (\lambda \cdot (x - \xi, y - u(\xi)))^2 \, d\xi\right] \\ &+ (Q_p(x, y))^{p+1} \int_\Omega |(x - \xi, y - u(\xi))|^{-p-2} \, d\xi \\ &\leqq (Q_p(x, y))^{p+1} \int_\Omega |(x - \xi, y - u(\xi))|^{-p-2} \, d\xi \\ &\leqq (Q_p(x, y))^{p+1} (Q(x, y))^{-2} \int_\Omega |(x - \xi, y - u(\xi))|^{-p} \, d\xi, \end{aligned} \tag{1.6}$$

and consequently

$$D_\lambda^2 Q_p(x, y) \leqq \left(\frac{Q_p(x, y)}{Q(x, y)}\right) \frac{1}{Q(x, y)}.$$

We conclude that

$$D_\lambda^2 Q_p(x, y) \leqq \frac{1}{Q(x, y)} + o(1) \quad \text{in } R^{n+1} \setminus \text{graph}(u).$$

This last result clearly implies the second inequality in (1.3).

QED

Definition 1.7. Define subsets, $\mathcal{O}^+$ and $\mathcal{O}^-$, of $\mathcal{O}$ by

$$\mathcal{O}^+ = \{(x, y) \in \mathcal{O} \mid y > u(x)\};$$
$$\mathcal{O}^- = \{(x, y) \in \mathcal{O} \mid y < u(x)\}.$$

The next lemma is used to prove that in the appropriate regions the level sets of Q are graphs of Lipschitz functions.

Lemma 1.8. *If* $Du \in L^\infty(\Omega; R^n)$ *then*

$$D_y Q \geqq (1 + \|Du\|_{L^\infty}^2)^{-\frac{1}{2}} \quad \text{a.e. in } \mathcal{O}^+;$$
$$D_y Q \leqq -(1 + \|Du\|_{L^\infty}^2)^{-\frac{1}{2}} \quad \text{a.e. in } \mathcal{O}^-.$$

Proof. Consider $D_y Q(x, y)$ for $(x, y) \in \mathcal{O}^+$. Since $(x, y) \in \mathcal{O}^+ \subset \mathcal{O}$

$$B((x, y); \varepsilon) \subset \Omega \times R, \quad \text{where } \varepsilon = Q(x, y). \tag{1.9}$$

Let $(\hat{x}, \hat{y}) \in \text{graph}(u)$ be a point such that $|(x - \hat{x}, y - \hat{y})| = \varepsilon$. It follows from (1.9) and the Lipschitz bound on u that

$$\frac{y - \hat{y}}{|x - \hat{x}|} \geqq (\|Du\|_{L^\infty})^{-1}. \tag{1.10}$$

Indeed, by implicit differentiation of $(\hat{y} - y)^2 + |\hat{x} - x|^2 = \varepsilon^2$

$$D_{\hat{x}} \hat{y} = -\frac{\hat{x} - x}{\hat{y} - y}$$

and so

$$\|Du\|_{L^\infty} \geqq \frac{|x - \hat{x}|}{y - \hat{y}}.$$

The definition of Q then shows that

$$Q(x, y - h) \leqq (|x - \hat{x}|^2 + (y - \hat{y} - h)^2)^{\frac{1}{2}}.$$

Rademacher's Theorem implies that Q is differentiable almost everywhere; thus we may assume without loss of generality that Q is differentiable at x, y and so from the last inequality obtain

$$D_yQ(x, y) \geqq (y - \hat{y})\,(|x - \hat{x}|^2 + (y - \hat{y})^2)^{-\frac{1}{2}}$$
$$\geqq \left(\frac{|x - \hat{x}|^2}{(y - \hat{y})^2} + 1\right)^{-\frac{1}{2}}.$$

Using (1.10) we conclude that

$$D_yQ(x, y) \geqq (1 + \|Du\|^2_{L^\infty})^{-\frac{1}{2}}.$$

It is clear that the second half of this lemma can be proved by an argument symmetric to the one just employed.

QED

The next theorem completes the construction of the approximations.

Theorem 1.11. *If* $Du \in L^\infty(\Omega; R^n)$ *then there is an* $\varepsilon_0 > 0$ *such that for* $\varepsilon \in (0, \varepsilon_0)$ *there exist functions* $u_\varepsilon^+, u_\varepsilon^- \in C(\overline{\Omega}_\varepsilon))$ *with the following properties:*

$$u_\varepsilon^- < u < u_\varepsilon^+ \quad \text{in } \Omega_\varepsilon,$$
$$Q(x, u_\varepsilon^\pm(x)) = \varepsilon \quad \text{for } x \in \Omega_\varepsilon,$$
$$\|Du_\varepsilon^\pm\|_{L^\infty} \leqq \|Du\|_{L^\infty}, \tag{1.12}$$
$$D_\lambda^2 u_\varepsilon^+ \geqq -\frac{c_0}{\varepsilon} \quad \text{(in the sense of distributions);}$$
$$D_\lambda^2 u_\varepsilon^- \leqq \frac{c_0}{\varepsilon} \quad \text{(in the sense of distributions),}$$

where the constant c_0, *appearing above, depends only on* $\|Du\|_{L^\infty}$.

Proof. Let ψ be a smooth function on R^{n+1} such that

$$\psi \geqq 0, \ \operatorname{supp}(\psi) \subset B(0, 1) \quad \text{and} \quad \int_{R^{n+1}} \psi(z)\,dz = 1.$$

Define ψ_δ by $\psi_\delta(z) = \delta^{-n-1}\,\psi\left(\frac{z}{\delta}\right)$ and define Q_δ by

$$Q_\delta(z) = \psi_\delta * Q(z) = \int_{R^{n+1}} \psi_\delta(z - \zeta)\,Q(\zeta)\,d\zeta.$$

If $\mathcal{O}_\delta^+ = \{(x, y) \in \mathcal{O}^+ \mid \text{distance}((x, y), R^{n+1} \setminus \mathcal{O}^+) > \delta\}$, then by using Lemma 1.8 we conclude

$$D_yQ_\delta \geqq (1 + \|Du\|^2_{L^\infty})^{-\frac{1}{2}} \quad \text{in } \mathcal{O}_\delta^+. \tag{1.13}$$

On the other hand, Lemma 1.2 yields

$$\left.\begin{aligned} &|DQ_\delta| \leqq 1 \quad \text{in } \mathcal{O}_\delta^+; \\ &D_\lambda^2 Q_\delta(x, y) \leqq \frac{1}{Q(x, y) - \delta} \quad \text{in } \mathcal{O}_\delta^+. \end{aligned}\right\} \tag{1.14}$$

We now apply the classical Implicit Function Theorem to Q_δ and so obtain a function $u^+_{\varepsilon,\delta} \in C^\infty(\Omega_{\varepsilon+\delta})$ such that

$$Q_\delta(x, u^+_{\varepsilon,\delta}(x)) = \varepsilon \quad \text{for } x \in \Omega_{\varepsilon+\delta}.$$

Calculating $Du^+_{\varepsilon,\delta}$, we find

$$Du^+_{\varepsilon,\delta}(x) = -\frac{D_x Q_\delta(x, u^+_{\varepsilon,\delta}(x))}{D_y Q_\delta(x, u^+_{\varepsilon,\delta}(x))};$$

and using (1.13) and the first inequality in (1.14), we obtain

$$\| Du^+_{\varepsilon,\delta} \|_{L^\infty} \leqq \| Du \|_{L^\infty}. \tag{1.15}$$

Calculating $D^2_\lambda u^+_{\varepsilon,\delta}$ with $\bar{\lambda} = (\lambda, 0)$ gives us

$$\begin{aligned} D^2_\lambda u^+_{\varepsilon,\delta} &= -(D_y Q_\delta)^{-1} [D^2_{\bar\lambda} Q_\delta + 2(D_y D_{\bar\lambda} Q_\delta)(D_y u^+_{\varepsilon,\delta}) + (D^2_y Q_\delta)(D_y u^+_{\varepsilon,\delta})^2] \\ &= -(D_y Q_\delta)^{-1} [D^2 Q_\delta((\lambda, D_y, u^+_{\varepsilon,\delta}), (\lambda, D_y u^+_{\varepsilon,\delta}))] \\ &\geqq -\frac{c_0}{\varepsilon - \delta}. \end{aligned}$$

By letting $\delta \searrow 0$ we obtain all the desired results for u^+_ε where $u^+_\varepsilon = \lim_{\delta \searrow 0} u^+_{\varepsilon,\delta}$.

The corresponding results for u^-_ε can be obtained by an argument symmetric to the one just employed.

QED

We have now constructed the approximations to the Lipschitz continuous function u. We shall eventually show that if u is the solution (in the sense made precise in the next section) to an elliptic partial differential equation then u^+_ε is a subsolution and u^-_ε is a supersolution to the same equation. In order to do that we need some additional properties of these constructions.

Definition 1.16. Given an open subset G of R^n and $\varphi \in C^\infty(G)$ we define ν: $G \to R^{n+1}$ by

$$\nu(x) = (1 + |D\varphi(x)|^2)^{-\frac{1}{2}} (D\varphi(x), -1).$$

Definition 1.17. Given G and φ as before, we define η_φ by

$$\eta_\varphi = \sup \{\eta \geqq 0 \mid B((x, \varphi(x)) + \eta\nu(x); \eta) \cap \text{graph}(\varphi) = \emptyset \text{ for all } x \in G\}.$$

Definition 1.18. Given G and φ as before, define $\mathscr{G}$ by

$$\mathscr{G} = \{(x, y) \mid (x, y) = (\xi, \varphi(\xi)) + \eta\nu(\xi)) \text{ for some } (\eta, \xi) \in [0, \eta_\varphi) \times G\}.$$

Lemma 1.19. *Let G be an open subset of R^n and let $\varphi \in C^\infty(G)$. Assume $\eta_\varphi > 0$ and define $\Phi(x, y)$ as $\Phi(x, y) =$ distance $((x, y), \text{graph}(\varphi))$. Then*

$$\Phi|_{\mathscr{G}} \in C^\infty(\mathscr{G}). \tag{1.20}$$

Proof. By the definition of η_φ it is clear that

$$\Phi((x, \varphi(x)) + \eta\nu(x)) = \eta \quad \text{if } (\eta, x) \in [0, \eta_\varphi) \times \mathscr{G}G. \tag{1.21}$$

Fix an $x_0 \in G$ and consider the map $H: [0, \eta_0) \times B(x_0; \eta_0) \to R^{n+1}$ given by $H(\eta, x) = (x, \varphi(x)) + \eta\nu(x)$. It is easily seen that H is smoothly invertible on $[0, \eta_0) \times B(x_0; \eta_0)$ if η_0 is sufficiently small. Thus Φ is smooth on $\mathscr{N} \cap \mathscr{G}$ for some neighborhood $\mathscr{N}$ of graph (φ).

We now proceed to prove that Φ is smooth on $\mathscr{G} \setminus \mathscr{N}$. Let (η_0, x_0) be a point in $[0, \eta_\varphi) \times \mathscr{G}$ and consider

$$\tilde{\Phi}(\tilde{x}, \tilde{y}) = \Phi(T(\tilde{x}, \tilde{y})),$$

where $T: R^{n+1} \to R^{n+1}$ is a rotation such that

$$T((0, -1)) = \nu(x_0). \tag{1.22}$$

Let $p: R^{n+1} \to R^n$ and $q: R^{n+1} \to R$ be the projections such that $p \circ i = id_{R^n}$ and $q \circ j = id_R$ (where i and j are the injections $i(x) = (x, 0)$ and $j(y) = (0, y)$). In order to prove that Φ is smooth near $(x_0, \varphi(x_0)) + \eta_0\nu(x_0)$ it is sufficient to prove that $\tilde{\Phi}$ is smooth near $T^{-1}((x_0, \varphi(x_0)) + \eta_0\nu(x_0))$. To do this we consider the map

$$S(x) = p \circ T^{-1}(x, \varphi(x)). \tag{1.23}$$

We claim that S is smoothly invertible near x_0. Indeed,

$$DS(x_0)(h) = p \circ T^{-1}(h, D\varphi(x_0)(h));$$

since $(h, D\varphi(x_0)(h))$ is orthogonal to $\nu(x_0)$ and T is a rotation, it follows from (1.22) that

$$T^{-1}(h, D\varphi(x_0)(h)) = (DS(x_0)(h), 0).$$

The last equation implies that $DS(x_0)$ is nonsingular, which proves our claim.

Define $\tilde{\varphi}(\tilde{x})$ for $\tilde{x}$ near $\tilde{x}_0 = S(x_0)$ by

$$\tilde{\varphi}(\tilde{x}) = q \circ T^{-1}(S^{-1}(\tilde{x}), \varphi(S^{-1}(\tilde{x}))). \tag{1.24}$$

We claim that

$$T(\text{graph}(\tilde{\varphi})) \subset \text{graph}(\varphi). \tag{1.25}$$

Indeed, since a point of graph $(\tilde{\varphi})$ can be written as $(\tilde{x}, \tilde{\varphi}(\tilde{x}))$, for $x = S^{-1}(\tilde{x})$ we see that

$$\begin{aligned} T(\tilde{x}, \tilde{\varphi}(\tilde{x})) &= T(S(x), q \circ T^{-1}(x, \varphi(x))) \\ &= T(p \circ T^{-1}(x, \varphi(x)), q \circ T^{-1}(x, \varphi(x))) \\ &= (x, \varphi(x)). \end{aligned}$$

This establishes our second claim.

It is clear that $\eta_{\tilde{\varphi}} \geqq \eta_\varphi$ since T is a rotation. Let $\tilde{\nu}(\tilde{x})$ be defined by (1.16) for $\tilde{\varphi}$. Since $\tilde{\nu}(\tilde{x})$ is normal to graph $(\tilde{\varphi})$ it follows from (1.25) that $T(\tilde{\nu}(\tilde{x}))$ is normal to graph (φ) and in fact

$$T(\tilde{\nu}(\tilde{x})) = \nu(S^{-1}(\tilde{x})). \tag{1.26}$$

It is also clear that

$$\tilde{\Phi}((\tilde{x}, \tilde{\varphi}(\tilde{x})) + \eta\tilde{\nu}(\tilde{x})) = \eta .$$

Consider the map Ψ defined by

$$\Psi(\tilde{x}, \eta) = (\tilde{x}, \tilde{\varphi}(\tilde{x})) + \eta\tilde{\nu}(\tilde{x}). \tag{1.27}$$

In order to prove that $\tilde{\Phi}$ is smooth near $T^{-1}((x_0, \varphi(x_0)) + \eta_0\nu(x_0))$ it suffices to show that Ψ is smoothly invertible near $(\tilde{x}_0, \eta_0)$.

By use of (1.22)–(1.24) we find

$$D\tilde{\varphi}(\tilde{x}_0) \circ DS(x_0)(h) = 0 \quad \text{for all } h \in R^n,$$

and therefore

$$D\tilde{\varphi}(\tilde{x}_0) = 0.$$

Using this fact, we obtain the following equation

$$D\Psi(\tilde{x}_0, \eta_0)(h, k) = (h, 0) + (0, k) + \eta_0 D\tilde{\nu}(\tilde{x}_0)(h).$$

To prove that $D\Psi(\tilde{x}_0, \eta_0)$ is nonsingular (and implicitly that Ψ is smoothly invertible near $(\tilde{x}_0, \eta_0)$) it is sufficient to show

$$h + p \circ D\tilde{\nu}(\tilde{x}_0)(h) \neq 0 \quad \text{if } h \neq 0. \tag{1.28}$$

We take the inner product of the left side of (1.28) with h and so conclude that a sufficient condition for (1.28) to hold is

$$I(h) = |h|^2 + \eta_0((h, 0) \cdot D\tilde{\nu}(\tilde{x}_0)(h)) > 0 \quad \text{if } h \neq 0.$$

Since $D\tilde{\varphi}(\tilde{x}_0) = 0$ we find that

$$I(h) = |h|^2 - \eta_0(D^2\tilde{\varphi}(\tilde{x}_0)(h, h)),$$

and because $\eta_0 < \eta_{\tilde{\varphi}}$ we conclude that

$$I(h) \geqq |h|^2 - \frac{\eta_0}{\eta_{\tilde{\varphi}}}|h|^2;$$

$$I(h) \geqq \left(1 - \frac{\eta_0}{\eta_{\tilde{\varphi}}}\right)|h|^2.$$

By our chain of reasoning this implies Φ that is smooth near $(x_0, \varphi(x_0)) + \eta_0\nu(x_0)$ and so completes the proof of this lemma.

QED

Lemma 1.29. *Let G be an open subset of R^n and let $\varphi \in C^\infty(G)$. Assume $\eta_\varphi > 0$ and let Φ be the distance function as defined in Lemma 1.19. Then for every $\eta \in [0, \varphi_\varphi)$ there is an open set G_η and $\varphi_\eta \in C^\infty(G_\eta)$ such that*

$$\begin{gathered}
\Phi(x, \varphi_\eta(x)) = \eta, \\
\varphi_\eta(x + \eta p \circ \nu(x)) = \varphi(x) + q \circ \nu(x) \quad \text{for all } x \in G, \\
D\varphi_\eta(x + \eta p \circ v(x)) = D\varphi(x) \quad \text{for all } x \in G; \\
D^2_\lambda\varphi_\eta(x + \eta p \circ \nu(x)) \leqq D^2_\lambda\varphi(x) \quad \text{for all } x \in G \text{ and all directions } \lambda.
\end{gathered} \tag{1.30}$$

Proof. As noted in the proof of Lemma 1.19

$$\Phi((x,\varphi(x))+\eta\nu(x))=\eta \quad \text{for all } (\eta,x)\in[0,\eta_\varphi)\times G.$$

We differentiate this equation and obtain

$$D\Phi((x,\varphi(x))+\eta\nu(x)\,(\nu(x)))=1.$$

Since $|D\Phi(\xi,\zeta)|\leqq 1$ in $\mathscr{G}$, we conclude

$$\begin{aligned} D_x\Phi((x,\varphi(x))+\eta\nu(x)) &= p(\nu(x)),\\ D_y\Phi((x,\varphi(x))+\eta\nu(x)) &= q(\nu(x)); \end{aligned} \tag{1.31}$$

where p and q are the projections that are used in the proof of Lemma 1.19. Since $q(\nu(x))\to 0$ for all $x\in G$, we may apply the Implicit Function Theorem and conclude the existence of a function $\varphi_\eta(x)$ defined on

$$G_\eta=\{x \mid x=\xi+p(\eta\nu(\xi)) \text{ for some } \xi\in G\}, \tag{1.32}$$

such that

$$\Phi(x,\varphi_\eta(x))=\eta. \tag{1.33}$$

We conclude

$$\varphi_\eta(\xi+p(\eta\nu(\xi)))=\varphi(\xi)+q(\nu(\xi)) \quad \text{for all } \xi\in G.$$

Differentiating (1.33) with respect to x, we obtain

$$D_x\Phi(x,\varphi_\eta(x))+D_y\Phi(x,\varphi_\eta(x))\,D\varphi_\eta(x)=0.$$

Solving for $D\varphi_\eta$, we there obtain

$$D\varphi_\eta(x)=-(D_y\Phi(x,\varphi_\eta(x)))^{-1}\,D_x\Phi(x,\varphi_\eta(x)).$$

From this equation and (1.31) we calculate that

$$D\varphi_\eta(\xi+p(\eta\nu(\xi)))=-\frac{p(\nu(\xi))}{q(\nu(\xi))} \quad \text{for all } \xi\in G.$$

Using the definition of $\nu(\xi)$, we obtain

$$D\varphi_\eta(\xi+p(\eta\nu(\xi)))=D\varphi(\xi) \quad \text{for all } \xi\in G.$$

We now proceed to prove the last inequality of (1.30). Let $z=(x,y)$ for the following calculations. By (1.31) we see that $|D\Phi(z)|^2=1$ on $\mathscr{G}$. Differentiate this equation twice in a direction λ to obtain the inequality

$$D\Phi D(D_\lambda^2\Phi)\leqq 0 \quad \text{in } \mathscr{G}.$$

Using (1.31), we find that

$$\nu(\xi)\cdot D(D_\lambda^2\Phi((\xi,\varphi(\xi))+\eta\nu(\xi)))\leqq 0 \quad \text{for all } \eta\in[0,\eta_\varphi).$$

The inequality above may be rewritten as

$$\frac{d}{d\eta}(D_\lambda^2\Phi((\xi,\varphi(\xi))+\eta\nu(\xi)))\leqq 0 \quad \text{for all } \eta\in[0,\eta_\varphi).$$

This differential inequality implies that for $(\eta, \xi) \in [0, \eta_\varphi) \times G$

$$D_\lambda^2 \Phi(\xi, \varphi(\xi)) \geqq D_\lambda^2 \Phi((\xi, \varphi(\xi)) + \eta\nu(\xi)). \tag{1.34}$$

Let $\bar{\lambda} = (\lambda, 0)$ and differentiate (1.33) twice in the direction λ to obtain

$$D_{\bar{\lambda}}^2 \Phi + 2(D_{\bar{\lambda}} D_y \Phi)(D_\lambda \varphi_\eta) + (D_y^2 \Phi)(D_\lambda \varphi_\eta) + D_y \Phi \, D_\lambda^2 \varphi_\eta = 0.$$

Solving for $D_\lambda^2 \varphi_\eta$ yields

$$D_\lambda^2 \varphi_\eta(x) = -(D_y \Phi(x, \varphi_\eta(x)))^{-1} D^2 \Phi(x, \varphi_\eta(x)) \langle (\lambda, D_\lambda \varphi_\eta(x)), (\lambda, D_\lambda \varphi_\eta(x)) \rangle.$$

Replacing x by $\xi + p(\eta\nu(\xi))$ leads to

$$D_\lambda^2 \varphi_\eta(\xi + p(\eta\nu(\xi))) = -\frac{1}{q(\nu(\xi))} \left| \left(\lambda, -\lambda \cdot \frac{p(\nu(\xi))}{q(\nu(\xi))} \right) \right| \cdot D_{\hat{\lambda}}^2 \Phi((\xi, \varphi(\xi)) + \eta\nu(\xi)),$$

where $\hat{\lambda} = \left| \left(\lambda, -\lambda \cdot \frac{p(\nu(\xi))}{q(\nu(\xi))} \right) \right|^{-1} \left(\lambda, -\lambda \cdot \frac{p(\nu(\xi))}{q(\nu(\xi))} \right)$. From the equation above and (1.34) we conclude

$$D_\lambda^2 \varphi_\eta(\xi) \geqq D_\lambda^2 \varphi_\eta(\xi + p(\eta\nu(\xi))) \quad \text{for all } \xi \in G \text{ and } \eta \in [0, \eta_\varphi).$$

This completes the proof of the lemma.

QED

2. Viscosity Solutions for Fully Nonlinear Equations

We begin by recalling the definition of a fully nonlinear elliptic partial differential operator. Next we recall the definition of viscosity solution of a fully nonlinear elliptic partial differential equation. The main objective of this section is to show that under reasonable assumptions the approximations, $u_\varepsilon^+ - \varepsilon$ and $u_\varepsilon^- + \varepsilon$, of a viscosity solution u of a fully nonlinear elliptic partial differential equation are viscosity subsolutions and supersolutions, respectively. This fact will be exploited in the next section when we prove a maximum principle for viscosity solutions of fully nonlinear elliptic partial differential equations.

The set of $n \times n$ real valued symmetric matrices will be denoted by $\mathcal{S}(n)$. These matrices admit the partial ordering $\succ$ where $M \succ N$ if $M \neq N$ and $M - N$ is positive semidefinite. A fully nonlinear partial differential operator $\mathcal{F}[\cdot]$ is defined by

$$\mathcal{F}[\varphi](x) = F(D^2\varphi(x), D\varphi(x), \varphi(x)) \quad \text{for all } \varphi \in C^\infty(\Omega), \tag{2.1}$$

where $F \in C(\mathcal{S}|(n) \times R^n \times R)$.

Definition 2.2. The operator $\mathcal{F}[\cdot]$ is degenerate elliptic if

$$F(M, p, t) \geqq F(N, p, t) \quad \text{for all } M \succ N \text{ and } (p, t) \in R^n \times R. \tag{2.3}$$

The operator $\mathcal{F}[\cdot]$ is uniformly elliptic if there are constants c_0 and c_1 such that

$$F(M, p, t) - F(N, q, t) \geqq c_0 \operatorname{trace}(M - N) - c_1 |p - q| \tag{2.4}$$
$$\text{for all } M \succ N, \text{ and } (p, q, t) \in R^n \times R^n \times R.$$

Definition 2.5. The operator $\mathscr{F}[\cdot]$ is nonincreasing if

$$F(M, p, t) \leqq F(M, p, s) \text{ for all } t > s, \text{ and } (M, p) \in \mathscr{S}|(n) \times R^n. \quad (2.6)$$

The operator $\mathscr{F}[\cdot]$ is decreasing if there is a constant $k_0 > 0$ such that

$$F(M, p, t) - F(M, p, s) \leqq k_0(s - t) \text{ for all } t > s, \text{ and } (M, p) \in \mathscr{S}|(n) \times R^n. \quad (2.7)$$

We shall now restate the definition of viscosity solution for the nonlinear partial differential equation

$$\mathscr{F}[w] = 0 \text{ in } \Omega, \quad (2.8)$$

as formulated in P.-L. Lions [8]. In order to proceed we need first to define the superdifferential and subdifferential of a function $w \in C(\Omega)$.

Definition 2.9. Let $w \in C(\Omega)$. The superdifferential, $D^+w(x)$, is defined as the set

$$D^+w(x) = \{(p, M) \in R^n \times \mathscr{S}(n) \mid w(x + z) \leqq w(x) + p \cdot z + \tfrac{1}{2} M(z, z) + o(|z|^2)\}. \quad (2.10)$$

The subdifferential, $D^-w(x)$, is defined as the set

$$D^-w(x) = \{(p, M) \in R^n + \mathscr{S}(n) \mid w(x \times z) \geqq w(x) + p \cdot z + \tfrac{1}{2} M(z, z) - o(|z|^2)\}. \quad (2.11)$$

Definition 2.12. $w \in C(\Omega)$ is a viscosity supersolution of (2.8) if

$$F(M, p, w(x)) \leqq 0 \text{ for all } (p, M) \in D^-w(x), \text{ and for all } x \in \Omega. \quad (2.13)$$

$w \in C(\Omega)$ is a viscosity subsolution of (2.8) if

$$F(M, p, w(x)) \geqq 0 \text{ for all } (p, M) \in D^+w(x), \text{ and for all } x \in \Omega. \quad (2.14)$$

$w \in C(\Omega)$ is a viscosity solution of (2.8) if it is both a viscosity supersolution and a viscosity subsolution.

The next lemma is a useful necessary and sufficient condition for w to be a viscosity supersolution of (2.8). The set G which appears in the lemma is an arbitrary open subset of Ω.

The following lemmas have been established in one form or another in P.-L. Lions' collective work, in particular [8].

Lemma 2.15. *Let* $w \in C(\Omega)$. *The following are equivalent:*

(i) *w is a viscosity supersolution of* (2.8).

(ii) $F(D^2\varphi(x_0), D\varphi(x_0), \varphi(x_0)) \leqq 0$ for all $(x_0, \varphi) \in \Omega \times C^\infty(G)$ such that $x_0 \in G$, $w(x) \geqq \varphi(x)$ for all $x \in G$; $w(x_0) = \varphi(x_0)$.

Lemma 2.20. *Let* $w \in C(\Omega)$. *The following are equivalent:*

(i) *w is a viscosity subsolution of* (2.8).

(ii) $F(D^2\varphi(x_0), D\varphi(x_0), \varphi(x_0)) \geqq 0$ for all $(x_0, \varphi) \in \Omega \times C^\infty(G)$ such that $x_0 \in G$, $w(x) \leqq \varphi(x)$ for all $x \in G$; $w(x_0) = \varphi(x_0)$.

Recall the definitions of Ω_ε, u_ε^+, u_ε^- and ε_0 from Section 1. We are now ready to relate the approximations u_ε^+ and u_ε^- with viscosity subsolutions and supersolutions, respectively.

Theorem 2.21. *Assume* $u \in C(\bar{\Omega}) \cap W^{1,\infty}(\Omega)$ *is a viscosity subsolution of* (2.8). *If* $\mathscr{F}[\cdot]$ *is degenerate elliptic and nonincreasing then* $u_\varepsilon^+ - \varepsilon$ *is also a viscosity subsolution of* (2.8) *for all* $\varepsilon \in [0, \varepsilon_0)$.

Proof. We shall prove this theorem by showing that (ii) of Lemma 2.20 holds for $u_\varepsilon^+ - \varepsilon$ on Ω_ε. Let $(x_0, \varphi) \in \Omega \times C^\infty(G)$ such that $x_0 \in G$, $u_\varepsilon^+(x) - \varepsilon \leqq \varphi(x)$ for all $x \in G$ and $u_\varepsilon^+(x_0) \in - \varepsilon = \varphi(x_0)$. Assume for now that in fact, $u_\varepsilon^+(x) - \varepsilon \leqq \varphi(x) - \delta\,|x - x_0|^2$ for some $\delta > 0$. Define $\hat{\varphi}(x) = \varphi(x) + \varepsilon - \delta\,|x - x_0|^2$ and note that $u_\varepsilon^+(x) \leqq \hat{\varphi}(x)$ for all $x \in G$ and $u_\varepsilon^+(x_0) = \hat{\varphi}(x_0)$. Let $\hat{\nu}(x)$ be the normal to graph $(\hat{\varphi})$ at $(x, \hat{\varphi}(x))$ as defined by Definition 1.16. We claim

$$(x_0, \hat{\varphi}(x_0)) + \varepsilon\hat{\nu}(x_0) \in \text{graph}\,(u). \tag{2.22}$$

Indeed, take $(x', u(x'))$ to be a closest point to $(x_0, \hat{\varphi}(x_0)) = (x_0, u_\varepsilon^+(x))$. Thus $|(x' - x_0, u(x') - \hat{\varphi}(x_0))| = \varepsilon$ and since $u_\varepsilon^+(x) \leqq \hat{\varphi}(x)$ for all $x \in G$ we see that

$$|(x' - x, u(x') - \hat{\varphi}(x))| \geqq \varepsilon \quad \text{for all } x \in G. \tag{2.23}$$

Consider the constrained minimization problem

$$\text{minimize } h(x, y) = |(x - x', y - u(x'))|^2;$$
$$\hat{\varphi}(x) - y = 0.$$

Since $(x_0, \hat{\varphi}(x_0))$ is a solution, the theory of Lagrange multipliers shows that there is a $\lambda_0 < 0$ such that

$$(x_0 - x', \hat{\varphi}(x_0) - u(x')) = \lambda_0(D\hat{\varphi}(x_0), -1).$$

Thus for some $\bar{\lambda}_0 > 0$ we have

$$(x' - x_0, u(x') - \hat{\varphi}(x_0)) = \hat{\lambda}_0\hat{\nu}(x_0);$$

and by taking the norms of both sides of this equation we conclude $\varepsilon = \hat{\lambda}_0$. This proves our claim.

Let $\nu(x)$ be the normal to graph $(\varphi + \varepsilon)$ at $(x, \varphi(x) + \varepsilon)$ as defined by Definition 1.16. Note that $\nu(x_0) = \hat{\nu}(x_0)$ and $\hat{\varphi}(x_0) = \varphi(x_0) + \varepsilon$. Thus by (2.22) we see that

$$(x_0, \varphi(x_0) + \varepsilon) + \varepsilon\nu(x_0) \in \text{graph}\,(u). \tag{2.24}$$

We claim (for a possibly smaller G) that $\eta_{\varphi+\varepsilon} > \varepsilon$. Indeed, by (2.23) we see that

$$B((x_0, \hat{\varphi}(x_0)) + \varepsilon\hat{\nu}(x_0); \varepsilon) \cap \text{graph}(\hat{\varphi}) = \emptyset.$$

It follows by the definition of $\hat{\varphi}$ that for some $\eta' > \varepsilon$

$$B((x_0, \varphi(x_0) + \varepsilon) + \eta'\nu(x_0); \eta') \cap \text{graph}(\varphi + \varepsilon) = \emptyset.$$

By the continuity of $D\varphi$ and $D^2\varphi$ we conclude that if G is a sufficiently small neighborhood of x_0 then $\eta_{\varphi+ve} > \varepsilon$.

Apply Lemma 1.29 to G and $(\varphi + \varepsilon) \in C^\infty(G)$. The conclusion is that there is an open set G_ε and a function $\varphi_\varepsilon \in C^\infty(G_\varepsilon)$ such that

$$\begin{aligned} \varphi_\varepsilon(x + \varepsilon p \circ \nu(x)) &= \varphi(x) + \varepsilon + q \circ \nu(x) \quad \text{for all } x \in G, \\ D\varphi_\varepsilon(x + \varepsilon p \circ \nu(x)) &= D\varphi(x) \quad \text{for all } x \in G; \\ D^2_\lambda\varphi_\varepsilon(x + \varepsilon p \circ \nu(x)) &\leqq D^2_\lambda\varphi(x) \quad \text{for all } x \in G \text{ and all directions } \lambda. \end{aligned} \tag{2.25}$$

Since $u_\varepsilon^+(x) \leqq \varphi(x) + \varepsilon$ for all x in G and since distance (graph $(\varphi + \varepsilon)$, graph (φ_ε)) $= \varepsilon$ we conclude that

$$u(x) \leqq \varphi_\varepsilon(x) \quad \text{for all } x \in G_\varepsilon.$$

By (2.24) we see that

$$u(x_0 + \varepsilon p \circ \nu(x_0)) = \varphi_\varepsilon(x_0 + \varepsilon p \circ \nu(x_0)).$$

Set $x' = x_0 + \varepsilon p \circ \nu(x_0)$ and note that since u is a viscosity subsolution of (2.8) we have by Lemma 2.20

$$F(D^2\varphi_\varepsilon(x'), D\varphi_\varepsilon(x'), \varphi_\varepsilon(x')) \geqq 0. \tag{2.26}$$

By (2.25) we see that

$$\begin{gathered} \varphi_\varepsilon(x') \geqq \varphi(x_0), \\ D\varphi_\varepsilon(x') = D\varphi(x_0); \\ D^2\varphi(x_0) \succ D^2\varphi_\varepsilon(x'). \end{gathered}$$

Since $\mathscr{F}[\cdot]$ is degenerate elliptic and nonincreasing we conclude

$$F(D^2\varphi(x_0), D\varphi(x_0), \varphi(x_0)) \geqq F(D^2\varphi_\varepsilon(x'), D\varphi_\varepsilon(x'), \varphi_\varepsilon(x')).$$

Combining this inequality with (2.26) we obtain

$$F(D^2\varphi(x_0), D\varphi(x_0), \varphi(x_0)) \geqq 0.$$

We now only have to remove the restriction $u_\varepsilon^+(x) \leqq \varphi(x) + \varepsilon - \delta\,|x - x_0|^2$ for some $\delta > 0$. This, however, is relatively easy. Indeed for any $\varphi \in C^\infty(G)$ such that $u_\varepsilon^+(x) \leqq \varphi(x) + \varepsilon$ and $u_\varepsilon^+(x_0) = \varphi(x_0) + \varepsilon$ we can apply our previous results to $\varphi'(x) = \varphi(x) + \delta\,|x - x_0|^2$ for any $\delta > 0$. Thus

$$F(D^2\varphi(x_0) + 2\delta I, D\varphi(x_0), \varphi(x_0)) \geqq 0 \quad \text{for any } \delta > 0.$$

By the continuity of F we conclude

$$F(D^2\varphi(x_0), D\varphi(x_0), \varphi(x_0)) \geqq 0.$$

Lemma 2.20 now shows that $u_\varepsilon^+ - \varepsilon$ is a viscosity subsolution of (2.8). This concludes the proof of the theorem.

QED

The following dual theorem to Theorem 2.21 can of course be proved in an analogous manner.

Theorem 2.27. *Assume* $u \in C(\bar{\Omega}) \cap W^{1,\infty}(\Omega)$ *is a viscosity supersolution of* (2.8). *If* $\mathscr{F}[\cdot]$ *is degenerate elliptic and noncinreasing then* $u_\varepsilon^- + \varepsilon$ *is also a viscosity supersolution of* (2.8) *for all* $\varepsilon \in [0, \varepsilon_0)$.

3. The Maximum Principle for Viscosity Solutions

The purpose of this section and the fundamental result of this paper is a proof of the following maximum principle for viscosity solutions of (2.8).

Theorem 3.1. *Let* $u, v \in C(\bar{\Omega}) \cap W^{1,\infty}(\Omega)$. *Assume u is a viscosity supersolution of* (2.8) *and v is a viscosity subsoluiotn of* (2.8). *If either*

(α) $\mathscr{F}[\cdot]$ *is degenerate elliptic and decreasing,*

or

(β) $\mathscr{F}[\cdot]$ *is uniformly elliptic and nonincreasing,*

then

$$\sup_{\partial\Omega} (v - u)^+ \geqq \sup_{\Omega} (v - u)^+. \tag{3.2}$$

The proof of Theorem 3.1 will be deferred until sufficient machinery has been developed to work with the approximations $u_\varepsilon^- + \varepsilon$ and $v_\varepsilon^+ + \varepsilon$. We begin our development with a fundamental lemma on semiconcave functions.

Lemma 3.3. *Let* $w \in C(\bar{\Omega}) \cap W^{1,\infty}(\Omega)$ *and assume*

$$D_\lambda^2 w \geqq -K_0 \text{ in } \Omega \textit{ (in the sense of distributions) for all directions } \lambda. \tag{3.4}$$

Then there is a function $M \in L^1(\Omega; (\mathscr{S}(n))$ *and a matrix-valued measure* $\Gamma \in \mathscr{M}(\Omega; \mathscr{S}(n))$ *such that*

(i) $D^2 w = M + \Gamma$ *(in the sense of distributions).*
(ii) Γ *is singular with respect to Lebesgue measure.*
(iii) $\Gamma(S)$ *is positive semidefinite for all Borel subsets, S, of* Ω.
(iv) $M(x)(\xi, \xi) \geqq -K_0 |\xi|^2$ *for all* $\xi \in R^n$; *for almost all* $x \in \Omega$.

Proof. By considering $w(x) + K_0 |x|^2$ we can assume without loss of generality that $K_0 = 0$ and so

$$D_\lambda^2 w \geqq 0 \text{ in } \Omega \text{ (in the sense of distributions)}. \tag{3.5}$$

Let $\{e_1, \ldots, e_n\}$ be the standard basis for R^n and denote the Borel σ-algebra of subsets Ω by $\mathscr{B}$. By (3.5)

$$D^2_{e_i} w \geqq 0 \text{ in } \Omega \text{ (in the sense of distributions) for } i = 1, 2, \ldots, n.$$

It is well known that a nonnegative distribution is a measure and so we conclude that for $i = 1, \ldots, n$ there are measures γ_{ii} on $\mathscr{B}$ such that for any $\varphi \in C_0^\infty(\Omega)$

$$\int_\Omega w(x)\, D^2_{e_i}\varphi(x)\, dx = \int_\Omega \varphi(x)\, d\gamma_{ii}(x). \tag{3.6}$$

Applying (3.5) with $\lambda^+_{ij} = \frac{1}{\sqrt{2}}(e_i + e_j)$ and $\lambda^-_{ij} = \frac{1}{\sqrt{2}}(e_i - e_j)$ allows us to conclude, for $1 \leqq i, j \leqq n$,

$$D^2_{e_i} w \pm 2 D_{e_i} D_{e_j} w + D^2_{e_j} w \geqq 0 \quad \text{in } \Omega \text{ (in the sense of distributions)}.$$

Using this and (3.6), we find that for $1 \leqq i,\ j \leqq n$, $\varphi \geqq 0$ and $\varphi \in C_0^\infty(\Omega)$

$$\left| \int_\Omega w(x)\, D_{e_i} D_{e_j}\varphi(x)\, dx \right| \leqq \tfrac{1}{2}\left(\int_\Omega \varphi(x)\, d\gamma_{ii}(x) + \int_\Omega \varphi(x)\, d\gamma_{jj}(x) \right).$$

Let $T_{ij}(\varphi) = \int_\Omega w(x)\, D_{e_i} D_{e_j}\varphi(x)\, dx$ and note that the preceding inequality implies that T_{ij} has a continuous extension to $C_0(\Omega)$. Measure theory then implies the existence of a signed measure γ_{ij} on $\mathscr{B}$ such that for any $\varphi \in C_0^\infty(\Omega)$

$$\int_\Omega w(x)\, D_{e_i} D_{e_j}\varphi(x)\, dx = \int_\Omega \varphi(x)\, d\gamma_{ij}(x). \tag{3.7}$$

Each γ_{ij} for $1 \leqq i, j \leqq n$ can be decomposed as

$$\gamma_{ij} = \tilde{\gamma}_{ij} + \hat{\gamma}_{ij},$$

where $\tilde{\gamma}_{ij}$ is absolutely continuous with respect to Lebesgue measure and $\hat{\gamma}_{ij}$ is singular with respect to Lebesgue measure. We can also assume without loss of generality that $\tilde{\gamma}_{ij} = \tilde{\gamma}_{ji}$ and $\hat{\gamma}_{ij} = \hat{\gamma}_{ji}$ for all $1 \leqq i, j \leqq n$.

The Radon-Nikodym Theorem then shows that there is an L^1 function m_{ij} for each $1 \leqq i, j \leqq n$ such that

$$\int_S m_{ij}(x)\, dx = \tilde{\gamma}_{ij}(S) \quad \text{for all } S \in \mathscr{B}.$$

Again we can assume without loss of generality that $m_{ij} = m_{ji}$ for all $1 \leqq i,$ $j \leqq n$. Property (i) is now established by taking $M = (m_{ij})$ and $\Gamma = (\tilde{\gamma}_{ij})$. Property (ii) is also established by this construction.

In order to prove property (iii) let E be any closed subset contained by the support of Γ. Thus meas $(E) = 0$. Let φ_k be a sequence of monotone decreasing functions in $C_0^\infty(\Omega)$ such that

$$\varphi_k(x) = 1 \quad \text{if } x \in E,$$

$$\varphi_k(x) \searrow 0 \quad \text{if } x \in E.$$

Given a direction $\lambda = (\lambda_1, \ldots, \lambda_n)$ we have by (3.5) for each k

$$\int_\Omega w(x) \sum_{i,j=1}^n \lambda_i \lambda_j D_{e_i} D_{e_j} \varphi_k(x)\, dx \geqq 0.$$

From our representation we conclude

$$\int_\Omega \sum_{i,j=1}^n m_{ij}(x)\, \lambda_i \lambda_j \varphi_k(x)\, dx + \sum_{i,j=1}^n \int_\Omega \varphi_k(x)\, \lambda_i \lambda_j\, d\tilde{\gamma}_{ij}(x) \geqq 0.$$

Letting $k \to \infty$, we conclude that

$$\sum_{i,j=1}^\infty \int_E \lambda_i \lambda_j\, d\tilde{\gamma}_{ij}(x) \geqq 0$$

and consequently

$$\Gamma(E)\,(\lambda, \lambda) \geqq 0.$$

From this, property (iii) easily follows.

An argument similar to the one above establishes

$$\left(\int_E M(x)\, dx\right)(\xi, \xi) \geqq 0 \quad \text{for any closed subset, } E,$$

of Ω which is disjoint from the support of Γ.

Since the support of Γ has Lebesgue measure zero the last inequality implies that property (iv) holds.

QED

The next lemma examines the size of the set of points with small derivatives for a function satisfying (3.4) and having an interior maximum. It is based on the ideas in C. PUCCI [10].

Definition 3.8. Let $w \in C(\bar{\Omega}) \cap W^{1,\infty}(\Omega)$ and define $\mathcal{G}_\delta$ by

$$\mathcal{G}_\delta = \{x \in \Omega \mid \text{for some } p \in \overline{B(0;\delta)},\ w(z) \leqq w(x) + p \cdot (z - x) \text{ for all } z \in \Omega\}. \tag{3.9}$$

Lemma 3.10. *Assume* $w \in C(\bar{\Omega}) \cap W^{1,\infty}(\Omega)$ *and that* (3.4) *holds. If* w *has an interior maximum then there are constants* $c_0 > 0$ *and* $\delta_0 > 0$ *such that*

$$\operatorname{meas}(\mathcal{G}_\delta) \geqq c_0\, \delta^n \quad \text{for all } \delta < \delta_0. \tag{3.11}$$

Proof. Let ψ be a smooth function on R^n such that

$$\psi \geqq 0 \quad \text{and} \quad \operatorname{supp}(\psi) \subset B(0;1);$$

$$\int_{R^n} \psi(x)\, dx = 1.$$

For $\eta > 0$ set $\psi_\eta(x) = \eta^{-n}\, \psi\left(\frac{x}{\eta}\right)$ and set $w_\eta(x) = \int_\Omega \psi_\eta(x - \xi)\, w(\xi)\, d\xi$ for

$x \in \Omega_\eta$. Define $\mathscr{G}_\delta^\eta$ analogously to $\mathscr{G}_\delta$. We claim that for any sequence $\eta_i \searrow 0$ that

$$\operatorname{meas}\left(\left[\limsup_{i\to\infty} \mathscr{G}_\delta^{\eta_i}\right] \setminus \mathscr{G}_\delta\right) = 0. \tag{3.12}$$

Indeed, for almost all $x \in \limsup_{i\to\infty} \mathscr{G}_\delta^{\eta_i}$ we have

$$\begin{aligned} w_{\eta_i}(x) &\to w(x); \\ Dw_{\eta_i}(x) &\to Dw(x). \end{aligned} \tag{3.13}$$

We can assume without loss of generality that $x \in \mathscr{G}_\delta^{\eta_i}$ for all i. By the definition of $\mathscr{G}_\delta^{\eta_i}$ we know

$$w_{\eta_i}(z) \leqq w_{\eta_i}(x) + Dw_{\eta_i}(x)\,(z-x) \quad \text{for all } z \in \Omega; \quad Dw_{\eta_i}(x) \in \overline{B(0;1)}.$$

We may also assume without loss of generality that $w_{\eta_i}(x) \to w(x)$ and $Dw_{\eta_i}(x) \to Dw(x)$. We therefore conclude that

$$w(z) \leqq w(x) + Dw(x)\,(z-x) \quad \text{for all } z \in \Omega; \quad Dw(x) \in \overline{B(0;1)}.$$

Thus $x \in \mathscr{G}_\delta$ and our claim is verified.

In order to complete a proof of this lemma it will now suffice to prove (3.11) for the sets $\mathscr{G}_\delta^\eta$ with c_0 independent of η. Since w has an interior maximum it follows that w_η has an interior maximum for η sufficiently small. In fact, there must be an $\eta_0 > 0$ such that

$$\gamma_\eta = \left(\sup_{\Omega_\eta} w_\eta - \sup_{\partial\Omega_\eta} w_\eta\right) \geqq \tfrac{1}{2}\left(\sup_{\Omega} w - \sup_{\partial\Omega} w\right) - \frac{\gamma_0}{2}.$$

Since Ω is bounded it now follows that for some $\delta_0 = \delta_0\left(\frac{\gamma_0}{2}, \Omega\right)$

$$Dw_\eta(\mathscr{G}_\delta^\eta) = \overline{B(0;\delta)} \quad \text{if } \delta < \delta_0 \text{ and } \eta < \eta_0. \tag{3.14}$$

Consider the integral

$$I = \int_{\mathscr{G}_\delta^\eta} |\det(D^2 w_\eta(x))|\, dz;$$

note that by the change of variables $y = Dw_\eta(z)$ and (3.14)

$$I \geqq \int_{B(0;\delta)} dy = \operatorname{meas}(B(0;\delta)). \tag{3.15}$$

On the other hand, by (3.4) and since w_η is concave down at every point $z \in \mathscr{G}_\delta^\eta$ we obtain

$$|\det(D^2 w_\eta(z))| \leqq K_0^n \quad \text{for all } z \in \mathscr{G}_\delta^\eta.$$

Combining this inequality with (3.15) we conclude

$$\operatorname{meas}(B(0;\delta)) \leqq K_0^n \operatorname{meas}(\mathscr{G}_\delta^\eta) \quad \text{for all } \delta < \delta_0 \text{ and } \eta < \eta_0.$$

It is clear that there is a constant $c_0 > 0$ such that

$$\text{meas}\,(\mathscr{G}_\delta^\eta) \geqq c_0\,\delta^n \quad \text{for all } \delta < \delta_0 \text{ and } \eta < \eta_0.$$

This completes and concludes the proof of the lemma.

QED

The next lemma deals with superdifferentials and subdifferentials.

Lemma 3.15. *Let* $w \in C(\overline{\Omega}) \cap W^{1,\infty}(\Omega)$ *and assume* (3.4) *holds. If* $D^2w = M + \Gamma$ *is the decomposition given in Lemma 3.3 then for almost all* $x \in \Omega$

$$w(x + z) - w(x) - Dw(x)\,(z) - \tfrac{1}{2}\,M(x)\,(z, z) = o(|z|^2). \tag{3.16}$$

Proof. Our proof makes use of the following differentiation theorem

$$\lim_{\varepsilon \searrow 0} \left(\varepsilon^{-1}\left\{\int_{B(x;\varepsilon)} |x - y|^{-n+1}\, M(x) - M(y)|\, dy \right.\right.$$
$$\left.\left. + \int_{B(x;\varepsilon)} |x - y|^{-n+1}\, d\,|\Gamma(y)|\right\}\right) = 0 \quad \text{for a.e. } x \in \Omega, \tag{3.17}$$

where $|\Gamma| = \sum_{i,j=1}^{n} |\gamma_{ij}|$ and $|M(x) - M(y)| = \sum_{i,j=1}^{n} |m_{ij}(x) - m_{ij}(y)|$.

Let x_0 be a point at which $Dw(x)$ exists and (3.17) holds. We claim that (3.16) holds at x_0.

Without loss of generality we can assume $w(x_0) = 0$, $Dw(x_0) = 0$, $M(x_0) = 0$, and $x_0 = 0$. Let w_η be as defined in the proof of Lemma 3.10. For any $E \in \mathscr{B}$ let

$$I_\varepsilon^\eta(E) = \varepsilon^{-n-2} \int_{B(0;2\varepsilon)} \chi_E(x)\, w_\eta(x)\, dx,$$

where χ_E is the characteristic function of E. We change coordinates and integrate by parts to obtain

$$I_\varepsilon^\eta(E) = \varepsilon^{-n-2} \int_{S^{n-1}} \int_0^{2\varepsilon} \chi_E(rx)\, r^{n-1} w_\eta(rx)\, dr\, dS_x$$
$$= \varepsilon^{-n-2} \int_{S^{n-1}} \int_0^{2\varepsilon} r^{n-1} \chi_E(rx) \left(\int_0^r \left(\int_0^\varrho D^2 w_\eta(\sigma x)\,(x, x)\, d\sigma\right.\right.$$
$$\left.\left. - Dw_\eta(0)\,(x)\right) d\varrho - w_\eta(0)\right) dr\, dS_x$$
$$= \varepsilon^{-n-2} \int_{S^{n-1}} \int_0^{2\varepsilon} \left(\int_\sigma^{2\varepsilon} \int_\varrho^{2\varepsilon} r^{n-1}\, \chi_E(rx)\, dr\, d\varrho\right) D^2 w_\eta(\sigma x)\,(x, x)\, d\sigma\, dS_x$$
$$- \mu_1(\varepsilon, E) \cdot Dw_\eta(0) - \mu_2(\varepsilon, E\; w_\eta(0).$$

Set $\mathscr{I}_E(\sigma, x) = \int\limits_{\sigma}^{2\varepsilon} \int\limits_{\varrho}^{2\varepsilon} \varepsilon^{-n-1} r^{n-1} \chi_E(rx)\, dr\, d\varrho$; then

$$I_\varepsilon^\eta(E) = \varepsilon^{-1} \int\limits_{S^{n-1}} \int\limits_0^{2\varepsilon} \mathscr{I}_E(\sigma, x)\, D^2 w_\eta(\sigma x)(x, x)\, d\sigma\, dS_x$$

$$- \mu_1(\varepsilon, E) \cdot Dw_\eta(0) - \mu_2(\varepsilon, E)\, w_\eta(0).$$

Changing coordinates again gives

$$I_\varepsilon^\eta(E) = \varepsilon^{-1} \int\limits_{B(0;2\varepsilon)} \mathscr{I}_E\left(|x|, \frac{x}{|x|}\right) |x|^{-n+1} D^2 w_\eta(x) \left(\frac{x}{|x|}, \frac{x}{|x|}\right) dx$$
$$- \mu_1(\varepsilon, E) \cdot Dw_\eta(0) - \mu_2(\varepsilon, E)\, w_\eta(0).$$

Using the definition of w_η and Lemma 3.3, we obtain

$$\lim_{\eta \searrow 0} I_\varepsilon^\eta(E) = \varepsilon^{-1}\left\{ \int\limits_{B(0;2\varepsilon)} \mathscr{I}_E\left(|x|, \frac{x}{|x|}\right) |x|^{-n+1} M(x) \left(\frac{x}{|x|}, \frac{x}{|x|}\right) dx \right.$$
$$\left. + \int\limits_{B(0;2\varepsilon)} \mathscr{I}_E\left(|x|, \frac{x}{|x|}\right) |x|^{-n+1} d\Gamma(x) \left(\frac{x}{|x|}, \frac{x}{|x|}\right)\right\},$$

where $d\Gamma(x) \left(\frac{x}{|x|}, \frac{x}{|x|}\right) = \sum\limits_{i,j=1}^{n} \frac{x_i x_j}{|x|^2}\, d\gamma_{ij}(x).$

On the other hand,

$$\lim_{\eta \searrow 0} I_\varepsilon^\eta(E) = \varepsilon^{-n-2} \int\limits_{B(0;2\varepsilon)} \chi_E(x)\, w(x)\, dx,$$

and so we conclude

$$\varepsilon^{-n-2} \int\limits_{B(0;2\varepsilon)} \chi_E(x)\, w(x)\, dx = \varepsilon^{-1}\left\{ \int\limits_{B(0;2\varepsilon)} \mathscr{I}_E\left(|x|, \frac{x}{|x|}\right) |x|^{-n+1} M(x) \left(\frac{x}{|x|}, \frac{x}{|x|}\right) dx \right.$$
$$\left. + \int\limits_{B(0;2\varepsilon)} \mathscr{I}_E\left(|x|, \frac{x}{|x|}\right) |x|^{-n+1} d\Gamma(x) \left(\frac{x}{|x|}, \frac{x}{|x|}\right)\right\}.$$

Finally, we obtain the inequality

$$\left|\varepsilon^{-n-2} \int\limits_{B(0;2\varepsilon)} \chi_E(x)\, w(x)\, dx\right| \leqq c_1 \varepsilon^{-1} \left\{ \int\limits_{B(0;2\varepsilon)} |x|^{-n+1} |M(x)|\, dx \right. \tag{3.18}$$
$$\left. + \int\limits_{B(0;2\varepsilon)} |x|^{-n+1}\, d\,|\Gamma|\,(x)\right\}.$$

Now let $m_\varepsilon = \max \left\{0, \sup\limits_{B(0;\varepsilon)} w\right\}$ and let y_ε be a point in $\overline{B(0, \varepsilon)}$ such that $w(y_\varepsilon) = m_\varepsilon$. For some constant $K_1 \geqq 0$ we see that $\tilde{w}(z) = w(z) + K_1 |x|^2$ is convex. Let $p_\varepsilon \in R^n$ be a vector such that

$$\tilde{w}(z) \geqq \tilde{w}(y_\varepsilon) + p_\varepsilon(z - y_\varepsilon) \quad \text{for all } z \in \Omega.$$

For w we have

$$w(y_\varepsilon + z) \geqq m_\varepsilon + (p_\varepsilon - 2K_1 y_\varepsilon) \cdot z - K_1 |z|^2.$$

Let $H_\varepsilon = \left\{ y_\varepsilon + z \mid |z|^2 \leqq \min \left\{ \varepsilon^2, \left(\frac{m_\varepsilon}{2K_1} \right) \right\} \text{ and } (p_\varepsilon - 2K_1 y_\varepsilon) \cdot z \geqq 0 \right\}$. We see that $w(x) \geqq \frac{m_\varepsilon}{2}$ in H_ε and that

$$\int_{B(0;2\varepsilon)} \chi_{H_\varepsilon}(x)\, w(x)\, dx \geqq c m_\varepsilon^{\frac{n+2}{2}}.$$

Combining this inequality with (3.18) we conclude

$$\left(\frac{m_\varepsilon}{\varepsilon^2} \right)^{\frac{n+2}{2}} \leqq c_2 \varepsilon^{-1} \left\{ \int_{B(0;2\varepsilon)} |x|^{-n+1} |M(x)|\, dx + \int_{B(0;2\varepsilon)} |x|^{-n+1}\, d|\Gamma|(x) \right\};$$

and consequently by (3.17)

$$m_\varepsilon^2 = o(\varepsilon^2).$$

This proves that

$$w(z) \leqq o(|z|^2). \tag{3.19}$$

We now consider $l_\varepsilon = \min \left\{ 0, \inf_{B(0;\varepsilon)} w \right\}$; let E_ε be the set

$$E_\varepsilon = \left\{ x \in B(0; 2\varepsilon) \mid w(x) \leqq \frac{l_\varepsilon}{2} \right\}.$$

Using inequality (3.18) we obtain

$$\frac{|l_\varepsilon|}{2} \operatorname{meas}(E_\varepsilon) = o(\varepsilon^{n+2}).$$

We now prove by contradiction that $l_\varepsilon = o(\varepsilon^2)$. Indeed, if $l_\varepsilon \neq o(\varepsilon^2)$, then there is a sequence $\varepsilon_i \searrow 0$ such that dor $c_0 > 0$, $l_{\varepsilon_i} \leqq -c_0 \varepsilon_i^2$ for all $i \in Z^+$. Now this implies

$$\operatorname{meas}(E_{\varepsilon_i}) = o(\varepsilon_i^n).$$

Let y_i be a point such that $w(y_i) = l_{\varepsilon_i}$ and $y_i \in \overline{B(0; \varepsilon_i)}$. Since $\operatorname{meas}(E_{\varepsilon_i}) = o(\varepsilon_i^n)$ it follows that there is a sequence of points $z_i \in B(0; \varepsilon_i)$ such that

$$w(z_i) \geqq + \frac{l_{\varepsilon_i}}{2};$$

$$|y_i - z_i| = o(\varepsilon_i).$$

From this it follows that there are points $\tilde{y}_i \in B(0; \varepsilon_i)$ such that

$$|Dw(\tilde{y}_i)| = \frac{\varepsilon_i}{o(1)} \quad \text{as } i \to \infty.$$

For some $K_1 > 0$ we see that $\tilde{w}(z) = w(z) + K_1 |z|^2$ is convex and

$$|D\tilde{w}(\tilde{y}_i)| = \frac{\varepsilon_i}{o(1)} + 2K_1 y_i$$

$$= \frac{\varepsilon_i}{o(1)}.$$

Since $\tilde{w}$ is convex

$$\tilde{w}(z) \geqq \tilde{w}(\tilde{y}_i) + D\tilde{w}(\tilde{y}_i)(z - \tilde{y}_i) \quad \text{for all } z \in \Omega,$$

and so

$$\tilde{w}(z) \geqq 0 + D\tilde{w}(\tilde{y}_i)(z - \tilde{y}_i) \quad \text{for all } z \in \Omega.$$

Therefore,

$$w(z) \geqq -K_1 |z|^2 + D\tilde{w}(\tilde{y}_i)(z - \tilde{y}_i) \quad \text{for all } z \in \Omega.$$

Taking $\tilde{z}_i = \tilde{y}_i + \varepsilon_i \dfrac{D\tilde{w}(\tilde{y}_i)}{|D\tilde{w}(\tilde{y}_i)|}$ we find that $\tilde{z}_i \in B(0; 2\varepsilon_i)$ and

$$w(\tilde{z}_i) \geqq -K_1\varepsilon_i^2 + |D\tilde{w}(\tilde{y}_i)|\,\varepsilon_i$$

$$\geqq -K_1\varepsilon_i^2 + \frac{\varepsilon_i^2}{o(1)}$$

$$\geqq \frac{\varepsilon_i^2}{o(1)}.$$

This contradicts (3.19) and so proves

$$w(z) \geqq -o(|z|^2).$$

This and (3.19) show that (3.16) holds for all points at which $Dw(x)$ exists and (3.17) holds. The set of such points includes almost all points of Ω and so the lemma is proved.

QED

The next lemma lets us estimate the ratio of trace $(M(x))$ and $|Dw(x)|$ on appropriate subsets of Ω.

Lemma 3.20. *Let* $w \in C(\overline{\Omega}) \cap W^{1,\infty}(\Omega)$ *and assume* (3.4) *holds. If* w *has an interior maximum then there is a constant* $\delta_0 > 0$ *such that for* $D^2w = M + \Gamma$ *(as given in Lemma 3.2)*

$$\int_{\mathscr{G}_\delta} \left[\frac{(\text{trace}\,(M(x)))^-}{|Dw(x)|}\right]^n dx = \infty \quad \text{for } 0 < \delta < \delta_0. \tag{3.21}$$

Proof. Let w_η be the approximation used in the proof of Lemma 3.10 and let $\mathscr{G}_\delta^\eta$ be the analogs for $\mathscr{G}_\delta$. Let

$$I_\delta^\eta = \int_{\mathscr{G}_\delta^\eta \setminus \mathscr{G}_{\delta/2}^\eta} |\Delta w_\eta(x)|\,|\det D^2 w_\eta(x)|^{\frac{n-1}{n}}\,dx;$$

we see that just as in the proof of Lemma 3.10 there is a $\delta_0 > 0$ such that

$$I_\delta^\eta = \int_{\mathcal{G}_\delta^\eta \setminus \mathcal{G}_{\delta/2}^\eta} |\det D^2 w_\eta(x)|\, dx = \int_{B(0;\delta)\setminus B(0;\delta/2)} dx \quad \text{for } 0 < \delta < \delta_0.$$

Obviously

$$I_\delta^\eta \geqq c_0\, \delta^n \quad \text{for } 0 < \delta < \delta_0;$$

and by Hölder's Inequality applied to the definition of I_δ^η we have

$$I_\delta^\eta \leqq \left(\int_{\mathcal{G}_\delta^\eta \setminus \mathcal{G}_{\delta/2}^\eta} |\Delta w_\eta(x)|^n\, dx \right)^{\frac{1}{n}} \left(\int_{\mathcal{G}_\delta^\eta \setminus \mathcal{G}_{\delta/2}^\eta} |\det D^2 w_\eta(x)|\, dx \right)^{\frac{n-1}{n}}$$

$$\leqq \left(\int_{\mathcal{G}_\delta^\eta \setminus \mathcal{G}_{\delta/2}^\eta} [(\operatorname{trace}(D^2 w_\eta(x)))^-]^n\, dx \right)^{\frac{1}{n}} c_1\, \delta^{n-1};$$

since $\Delta w_\eta(x) = \operatorname{trace}(D^2 w_\eta(x)) \leqq 0$ for all $x \in \mathcal{G}_\delta^\eta$.

These inequalities allow us to conclude for some $c_2 > 0$

$$\int_{\mathcal{G}_\delta^\eta \setminus \mathcal{G}_{\delta/2}^\eta} [(\operatorname{trace}(D^2 w_\eta(x)))^-]^n\, dx \geqq c_2\, \delta^n. \tag{3.22}$$

Now we claim that, for almost all $\delta \in (0, \delta_0)$

$$\operatorname{meas}\left((\mathcal{G}_{2^{-j+1}\delta}^\eta \setminus \mathcal{G}_{2^{-j+1}\delta}) \cup (\mathcal{G}_{2^{-j+1}\delta} \setminus \mathcal{G}_{2^{-j+1}\delta}^\eta)\right) \to 0 \text{ as } \eta \searrow 0 \tag{3.23}$$

for any positive integer, j.

Indeed, we have already seen in the proof of Lemma 3.10 that $\operatorname{meas}(\mathcal{G}_\delta^\eta \setminus \mathcal{G}_\delta) \to 0$ as $\eta \searrow 0$ for all $\delta \in (0, \delta_0)$. By the definition of $\mathcal{G}_\delta^\eta$ and w_η it is not difficult to see that

$$\operatorname{meas}(\mathcal{G}_{\delta'} \setminus \mathcal{G}_\delta^\eta) \to 0 \quad \text{as } \eta \searrow 0 \quad \text{for all } 0 < \delta' < \delta < \delta_0.$$

Clearly

$$\operatorname{meas}(\mathcal{G}_\delta \setminus \mathcal{G}_\delta^\eta) \to 0 \quad \text{as } \eta \searrow 0 \text{ if } \operatorname{meas}\left(\mathcal{G}_\delta \setminus \bigcup_{\delta' < \delta} \mathcal{G}_{\delta'}\right) = 0.$$

Since the family $\mathcal{G}_\delta$ is a monotone decreasing family of sets $\operatorname{meas}\left(\mathcal{G}_\delta \setminus \bigcup_{\delta' < \delta} \mathcal{G}_{\delta'}\right) = 0$ for almost all $\delta \in (0, \delta_0)$. Thus

$$\operatorname{meas}(\mathcal{G}_\delta \setminus \mathcal{G}_\delta^\eta) \to 0 \quad \text{as } \eta \searrow 0 \text{ for almost all } \delta \in (0, \delta_0).$$

Let $V_j = \{\delta \in (0, \delta_0) \mid \operatorname{meas}(\mathcal{G}_{2^{-j}\delta} \setminus \mathcal{G}_{2^{-j}\delta}^\eta) \not\to 0 \text{ as } \eta \searrow 0\}$. Each set V_j is of measure zero so $V = \bigcup_{j=0}^{\infty} V_j$ is also of measure zero. For each $\delta \in (0, \delta_0) \setminus V$ we see

$$\operatorname{meas}(\mathcal{G}_{2^{-j+1}\delta} \setminus \mathcal{G}_{2^{-j+1}\delta}^\eta) \to 0 \quad \text{as } \eta \searrow 0 \text{ for all } j \in Z^+.$$

This proves our claim.

The differentiation theorem expressed by, (3.17) shows that $D^2 w_\eta(x) \to M(x)$ for almost all $x \in \Omega$. Since

$$0 \geqq \operatorname{trace}(D^2 w_\eta(x)) \geqq -K_0 \quad \text{for all } x \in \mathscr{G}_\delta^\eta$$

and since (3.22) and (3.23) hold we see that

$$\int_{\mathscr{G}_{2^{-j+1}\delta} \setminus \mathscr{G}_{2^{-j}\delta}} [(\operatorname{trace}(M(x)))^-]^n \, dx \geqq c_2 (2^{-j+1}\delta)^n$$

$$\text{for all } j \in Z^+ \text{ for almost all } \delta \in (0, \delta_0).$$

Thus

$$\int_{\mathscr{G}_{2^{-j+1}\delta \setminus 2^{-j}\delta}} \left[\frac{(\operatorname{trace}(M(x)))^-}{2^{-j+1}\delta}\right]^n dx \geqq c_2$$

$$\text{for all } j \in Z^+; \text{ for almost all } \delta \in (0, \delta_0).$$

On the set $\mathscr{G}_{2^{-j+1}\delta \setminus 2^{-j}\delta}$ we see that $|Dw(x)| \leqq 2^{-j+1}\delta$ almost everywhere and so

$$\int_{\mathscr{G}_{2^{-j+1}\delta} \setminus \mathscr{G}_{2^{-j}\delta}} \left[\frac{(\operatorname{trace}(M(x)))^-}{|Dw(x)|}\right]^n dx \geqq c_2$$

$$\text{for all } j \in Z^+; \text{ for almost all } \delta \in (0, \delta_0).$$

By adding these inequalities over $j \in Z^+$ we obtain

$$\int_{\mathscr{G}_\delta} \left[\frac{(\operatorname{trace}(M(x)))^-}{|Dw(x)|}\right]^n dx = \infty \text{ for almost all } \delta \in (0, \delta_0).$$

This of course gives the result claimed by the lemma.

QED

We are now ready to prove Theorem 3.1.

Proof of Theorem 3.1. This will be a proof by contradiction. Assume the theorem is false; then

$$m_0 = \sup_{\Omega} (v - u)^+ - \sup_{\partial\Omega} (v - u)^+ > 0.$$

Let $\tilde{v} = v_\varepsilon^+ - \varepsilon$ and $\tilde{u} = u_\varepsilon^- + \varepsilon$; note that if ε is sufficiently small then

$$\sup_{\Omega_\varepsilon} (\tilde{v} - \tilde{u})^+ - \sup_{\partial\Omega_\varepsilon} (\tilde{v} - \tilde{u}) \geqq \frac{m_0}{2}.$$

By Theorems 2.21 and 2.27 we see that $\tilde{u}$ is a viscosity supersolution of (2.8) and $\tilde{v}$ is a viscosity subsolution of (2.8). By Theorem 1.11 for any direction λ

$$D_\lambda^2 \tilde{u} \leqq \frac{K}{\varepsilon} \quad \text{and} \quad D_\lambda^2 \tilde{v} \geqq -\frac{K}{\varepsilon} \quad \text{in } \Omega_\varepsilon \text{ (in the sense of distributions)}. \tag{3.24}$$

Letting $\tilde{w} = \tilde{v} - \tilde{u}$, we see that w is in $C(\bar{\Omega}) \cap W^{1,\infty}(\Omega)$ and it satisfies (3.4). With $\tilde{\mathscr{G}}_\delta$ defined by (3.9) with $w = \tilde{w}$ we see by Lemma 3.10 that for some $\delta_0 > 0$

$$\operatorname{meas}(\mathscr{G}_\delta) \geqq c_0\,\delta^n \quad \text{for all } \delta < \delta_0.$$

By Lemma 3.3

$$D^2\tilde{v} = \tilde{M}^+ + \tilde{\Gamma}^+;$$

$$D^2\tilde{u} = \tilde{M}^- + \tilde{\Gamma}^-.$$

Then $\tilde{M} = \tilde{M}^+ - \tilde{M}^-$ and $\tilde{\Gamma} = \tilde{\Gamma}^+ - \tilde{\Gamma}^-$ give a representation for $D^2\tilde{w}$. By Lemma 3.15, for almost all $x \in \tilde{\mathscr{G}}_\delta$, $D\tilde{u}(x)$ and $D\tilde{v}(x)$ exist and

$$\tilde{v}(x+z) - \tilde{v}(x) - D\tilde{v}(x)(z) - \tfrac{1}{2}\tilde{M}^+(x)(z,z) \leqq o(|z|^2);$$

$$\tilde{u}(x+z) - \tilde{u}(z)x - D\tilde{u}(x)(z) - \tfrac{1}{2}\tilde{M}^-(x)(z,z) \geqq -o(|z|^2).$$

Furthermore, by the definition of $\tilde{\mathscr{G}}_\delta$

$$\tilde{M}^-(x) > \tilde{M}^+(x) \quad \text{for almost all } x \in \tilde{\mathscr{G}}_\delta.$$

Applying the definitions of viscosity subsolution and supersolution we find that

$$F(\tilde{M}^+(x), D\tilde{v}(x], \tilde{v}(x)) \geqq F(\tilde{M}^-(x), D\tilde{u}(x), \tilde{u}(x)) \text{ for almost all } x \in \mathscr{G}_\delta. \tag{3.25}$$

By Lemma 3.3 and (3.24) we see

$$\frac{K}{\varepsilon} I > \tilde{M}^-(x) > \tilde{M}^+(x) > -\frac{K}{\varepsilon} I \quad \text{for almost all } x \in \tilde{\mathscr{G}}_\delta. \tag{3.26}$$

If (α) holds then for almost all $x \in \tilde{\mathscr{G}}_\delta$

$$F(\tilde{M}^+(x), D\tilde{v}(x), \tilde{v}(x)) \leqq F(\tilde{M}^-(x), D\tilde{v}(x), \tilde{u}(x)) - k_0(\tilde{v}(x) - \tilde{u}(x)),$$

and by the continuity of F and (3.26) there is a continuous increasing function $\sigma(t)$ such that $\sigma(0) = 0$ and

$$F(\tilde{M}^+(x), D\tilde{v}(x), \tilde{v}(x)) \leqq F(\tilde{M}^-(x), D\tilde{u}(x), \tilde{u}(x)) - k_0(\tilde{v}(x) - \tilde{u}(x)) + \sigma(\delta)$$

$$\text{for almost all } x \in \tilde{\mathscr{G}}_\delta.$$

Choosing δ sufficiently small yields

$$F(\tilde{M}^+(x), D\tilde{v}(x), \tilde{v}(x)) \leqq F(\tilde{M}^-(x), D\tilde{u}(x), \tilde{u}(x)) - \frac{k_0}{4} m_0 + \sigma(\delta)$$

$$\text{for almost all } x \in \tilde{\mathscr{G}}_\delta.$$

This clearly contradicts (3.25) if δ is sufficiently small.

If (β) holds, then for almost all $x \in \tilde{\mathscr{G}}_\delta$

$$F(\tilde{M}^+(x), D\tilde{v}(x), \tilde{v}(x)) \leqq F(\tilde{M}^-(x), D\tilde{u}(x), \tilde{u}(x))$$
$$- (c_0 \operatorname{trace}((\tilde{M}^- - \tilde{M}^+)(x)) - c_1 |D\tilde{u}(x) - D\tilde{v}(x)|).$$

Note that $\tilde{M}(x) = \tilde{M}^+(x) - \tilde{M}^-(x)$ and $D\tilde{w}(x) = D\tilde{v}(x) - D\tilde{u}(x)$ for almost all $x \in \Omega_\varepsilon$. By Lemma 3.20 there is a set $\tilde{E}$ with positive measure in $\mathscr{G}_\delta$ such that

$$(\text{trace}\,(\tilde{M}(x)))^- \geqq 2\frac{c_1}{c_0}\,|D\tilde{w}(x)| \quad \text{for almost all } x \in \tilde{E}.$$

Consequently

$$F(\tilde{M}^+(x), D\tilde{v}(x), \tilde{v}(x)) < F(\tilde{M}^-(x), D\tilde{u}(x), \tilde{u}(x)) \text{ for almost all } x \in \tilde{E} \subset \tilde{\mathscr{G}}_\delta.$$

This also contradicts (3.25) and so completes the proof of Theorem 3.1.

QED

The research reported here was supported in part by grants from the Alfred P. Sloan Foundation and the National Science Foundation.

Bibliography

1. J. M. Bony, Principe du maximum dans les espaces de Sobolev. C. R. Acad. Sci. Paris **265** (1967), 333–336.
2. M. G. Crandall, L. C. Evans, & P.-L. Lions, Some properties of viscosity solutions of Hamilton-Jacobi equations. Trans. Amer. Math. Soc. **282** (1984), 487–502.
3. M. G. Crandall & P.-L. Lions, Viscosity solutions of Hamilton-Jacobi equations. Trans. Amer. Math. Soc. **277** (1983), 1–42.
4. L. C. Evans, A convergence theorem for solutions of nonlinear second order elliptic equations. Indiana Univ. Math. J. **27** (1978), 875–887.
5. L. C. Evans, On solving certain nonlinear partial differential equations by accretive operator methods. Israel J. Math. **36** (1980), 225–247.
6. L. C. Evans, Classical solutions of fully nonlinear, convex, second order elliptic equations. Comm. Pure Appl. Math. **25** (1982), 333–363.
7. P.-L. Lions, Optimal control of diffusion processes and Hamilton-Jacobi-Bellman equations. Part 1. Dynamic Programming Principle and applications. Comm. Partial Differential Equations **8** (1983), 1101–1134.
8. P.-L. Lions, Optimal control of diffusion processes and Hamilton-Jacobi-Bellman equations. Part 2. Viscosity solutions and uniqueness. Comm. Partial Differential Equations **8** (1983), 1229–1276.
9. P.-L. Lions, Optimal control of diffusion processes and Hamilton-Jacobi-Bellman equations. Part 3. Regularity of the optimal cost function. Collège de France Seminar, Vol. V, Pitman, London, 1983.
10. C. Pucci, Limitazioni per soluzioni di equazioni ellitiche. Ann. Mat. Pura Appl. **74** (1966), 15–30.
11. E. M. Stein, Singular Integrals and Differentiability Properties of Functions. Princeton University Press, Princeton, New Jersey, 1970.

Department of Mathematical Sciences
Loyola University of Chicago
Chicago, Illinois 60626

(Received June 26, 1987)