WIDE STENCIL FINITE DIFFERENCE SCHEMES FOR THE
ELLIPTIC MONGE-AMPÈRE EQUATION AND FUNCTIONS
OF THE EIGENVALUES OF THE HESSIAN

ADAM M. OBERMAN

Abstract. Certain fully nonlinear elliptic Partial Differential Equations can be written as functions of the eigenvalues of the Hessian. These include: the Monge-Ampère equation, Pucci’s Maximal and Minimal equations, and the equation for the convex envelope. In this article we build convergent monotone finite difference schemes for the aforementioned equations. Numerical results are presented.

1. Introduction

Certain fully nonlinear elliptic Partial Differential Equations can be written as functions of the eigenvalues of the Hessian. These include: the Monge-Ampère equation, the Pucci maximal and minimal equations, and the equations for the convex and concave envelope of the boundary data. In this article we present convergent approximation schemes for such equations, in the form of wide stencil finite difference schemes. We prove that numerical solutions converge uniformly to the unique viscosity solution of the equation without any regularity requirement on the solution.

To keep the ideas clear, we work in two dimensions and limit ourselves to the examples listed above. Our methods generalize to other equations which are functions of the eigenvalues, to equations with lower order terms, and to higher dimensions.

1.1. Presentation of the Equations. We begin by writing down the general form of the equations, and then show how the specific equations mentioned above fit into the framework. Write $D^2u(x)$ for the Hessian of $u : \mathbb{R}^2 \to \mathbb{R}$ and write $\lambda_- [u](x) \leq \lambda_+ [u](x)$ for the smaller and larger eigenvalues of the Hessian of $u$, respectively. Let $\Omega \subset \mathbb{R}^2$ be a bounded domain, and consider the Dirichlet problem for the fully nonlinear, elliptic Partial Differential Equation

$$\begin{cases} F[u] \equiv -G(D^2u(x)) + f(x) = 0 & \text{in } \Omega \\ u = g & \text{on } \partial \Omega. \end{cases}$$

Since $F$ is elliptic, $G$ is a non-decreasing function of the eigenvalues of the Hessian. Next we show how different choices of $G$ which lead to the different equations.

*Date: June 18, 2007.*

*2000 Mathematics Subject Classification.* 65N06, 65N12, 65M06, 65M12, 35B50, 35J60, 35R35, 35K65, 49L25.

*Key words and phrases.* finite difference schemes, partial differential equations, viscosity solutions, Monge-Ampère equation, Pucci maximal equation, convex functions, convex hull.
(i) Convex functions satisfy $\lambda_- \geq 0$, and concave functions satisfy $\lambda_+ \leq 0$ in the viscosity sense [17]. Setting $f = 0$ in (PDE) and setting

$$G^{CE}(M) = \lambda_-$$

we get the equation for the convex envelope of the boundary data. Likewise setting $G(M) = \lambda_+$ we get the equation for the concave hull of the boundary data.

(ii) The Monge-Ampère equation [19, 12] is obtained when we set $G(M) \equiv \det(M) = \lambda_+ \lambda_-$ It has been the subject of much recent attention, in part because of its appearance in optimal mass transportation [11, 3]. In order for the equation to be elliptic, the additional requirement that solution be convex needs to be imposed. This condition can be imposed using $\lambda_- \geq 0$ in the viscosity sense [17]. The result is the single equation

$$G^{MA}(M) = \max(\lambda_-^0, 0) \max(\lambda_+^{0}, 0).$$

Solutions of Monge-Ampère with $f = 0$ and the convexity requirement are also solutions of $\lambda_- = 0$.

(iii) The Pucci maximal and minimal equations obtained by setting

$$G^M = \gamma \lambda_- + \Gamma \lambda_+$$

and

$$G^m = \Gamma \lambda_- + \gamma \lambda_+$$

where $0 < \gamma \leq \Gamma$. Solutions of these equations are the best underestimator and overestimators of solutions of elliptic equations with ellipticity constants related to $\gamma, \Gamma$, see [2] sec. 2.2] and [12 p. 442]. As such, they can be used to give pointwise upper and lower bounds on solutions of elliptic equations with unknown coefficients.

For each of the equations, the challenge is to discretize the nonlinear second order differential operator; adding zeroth or first order terms to the equation can be accomplished without too much difficulty, we refer to [16] for details.

1.2. Related work. We know of little early numerical work on these equations. Oliker and Prussner presented a convergent scheme for the Monge-Ampère equation in two dimensions [15]. It uses a natural discretization of the geometric Alexandrov-Bakelman interpretation of solutions.

Recently, Dean and Giowinski have investigated numerical methods for the Monge-Ampère equation [8 7 6 10] and Pucci’s equations [9]. As the authors are careful to point out, in order to converge, their method requires more regularity than can be expected for solutions.

1.3. A semi-discrete scheme. To illustrate ideas, we begin with a semi-discrete scheme. The theoretical properties we will prove hold for the semi-discrete scheme, so it can be used to sketch out the main ideas in a simpler setting.

Consider, for the sake of illustration, the equation

$$-\lambda_- [u](x) = f(x)$$

Using the classical Rayleigh-Ritz characterization of the eigenvalues, we can write

$$\lambda_- [u](x) = \min_{\theta} \frac{d^2 u}{d\theta^2}(x),$$

and

$$\lambda_+ [u](x) = \max_{\theta} \frac{d^2 u}{d\theta^2}(x),$$
where \( v_\theta = (\cos \theta, \sin \theta) \) is the unit vector in the direction of the angle \( \theta \). We substitute the standard centered difference approximation to the second derivative in the direction of \( v_\theta \), to arrive at the semi-discrete scheme

\[
\lambda^h u(x) \equiv \min_\theta \frac{u(x + hv_\theta) - 2u(x) + u(x - hv_\theta)}{h^2}.
\]

1.4. Higher Dimensions. The discretization for the eigenvalues of the Hessian presented in Section 1.3 using the classical Rayleigh-Ritz formula for the eigenvalues of a symmetric matrix. This formula simplifies in two dimensions, since the maximum or a minimum over all directions to gives the largest and smallest eigenvalues, respectively. In order to compute the range of approximate eigenvalues in dimensions higher than three, the full formula must be used. This means finding the largest eigenvalue, using the max formula, then finding the next largest by taking a maximum over (approximately) orthogonal directions, and so on.

1.5. Wide stencil finite difference schemes. Monotonicity, along with consistency and stability is sufficient for convergence of finite difference schemes to the unique viscosity solution of a nonlinear, possibly degenerate elliptic PDE [1]. For second order equations, even if they are linear, it is may not be easy to build monotone approximation schemes without making further assumptions on the structure of the equation. In fact, Motzkin and Wasow showed that for a given fixed stencil, it is possible to find a linear elliptic equation for which no second order accurate, monotone scheme can be built [13]. Their example was non-constructive, but later work by Crandall and Lions [5] makes the problem clear.

Suppose the equation is \( d^2u/dv^2 \), where \( v \) is a direction vector which doesn’t line up with the grid. Then it is impossible to find a monotone second order accurate scheme. Since nonlinear PDEs often reduce to two derivatives in a given direction, the same difficulties apply in this case. Crandall and Lions resorted to using grid with impractically large stencils. An alternative is to use a wider, but still reasonable sized grid, and resulting in lower accuracy. The idea is to replace derivatives of the form \( d^2u/dv^2 \), for general directions vectors \( v \), with \( d^2u/dw^2 \), where \( w \) is the closest direction vector to \( v \) which lines up on the grid. This incurs an error of which depends on the directional resolution \( d\theta \), which is a measure of the distance from \( v \) to \( w \). If this can be done in a monotone way, then a convergent scheme can be built.

Wide stencil schemes which have previously been used by the author to approximate the equation for Motion of Level Sets by Mean Curvature [14] and the Infinity Laplace equation [15]. These schemes required using specific properties of the equations for which a discrete analogue could be found. The result (a median scheme for Mean Curvature, and a discretely locally minimizing Lipschitz extension for Infinity Laplace) led to an expression which could be shown to be consistent and monotone. For equations which are functions of the eigenvalues, we have a natural discrete characterization of the eigenvalues, which allows schemes for a wider class of equations to be built.

The schemes converge as the spatial resolution, \( h \), and the directional resolution, \( d\theta \), go to zero. In practice, we work with reasonably small stencils (up to four times the grid spacing). The computational time for the wide stencil schemes scales no worse than linearly with the stencil size, depending on the particular choice of stencil. Wider stencil schemes have the advantage of better capturing the geometry
of the solution, and enjoy approximate versions of the properties (e.g. rotational invariance) of the exact solutions, as well as being necessary to obtain theoretical convergence. The main disadvantage is the lower accuracy, which also depends on the directional resolution, for example, $O(h^2 + d\theta^2)$, as in (5.1).

It is natural to try to extend these schemes to obtain higher order accuracy. One approach to building higher order schemes would be to pursue an analogy with WENO schemes for first order equations. Using a foundation of a monotone scheme, it might be possible to selectively use a higher order scheme when the solution appears regular enough, and fall back on the lower order monotone scheme when necessary.

1.6. Properties of the semi-discrete scheme. While the scheme above is only semi-discrete, we set aside that consideration aside for now and study its properties. Many of these properties will also hold for the fully discrete schemes which follow.

This scheme is second order accurate, by which we mean,

$$\lambda_- [\phi](x) - \lambda^h [\phi](x) = O(h^2),$$

for every twice differentiable function $\phi(x)$.

Begin by substituting (1.2) into (1.1) and solving for the reference value $u(x)$, thus rewriting the scheme in the explicit form,

$$u(x) = \frac{h^2}{2} f(x) + \min_{\theta} \frac{u(x + hv_\theta) + u(x - hv_\theta)}{2}.$$

The scheme is defined locally, but by extending this definition to all points in the domain (with possible adjustments near the boundary) we can arrive at a scheme for the equation in a domain. While we make local arguments, the properties below apply globally. More precisely, the properties below apply to the solution operator of the scheme, when it is well-defined.

The scheme is monotone: increasing the value at any of the neighboring values $u(x + hv_\theta)$ cannot decrease the value at $u(x)$.

The scheme is degenerate: changing the value at some of the neighbors might have no effect on the value at $u(x)$.

The scheme is stable, in fact it is non-expansive in the uniform norm. Consider a perturbation $\delta(x)$, write $\tilde{u}(x + hv_\theta) = u(x + hv_\theta) + \delta(x + hv_\theta)$, and call the resulting value $\tilde{u}(x)$. It follows that

$$|\tilde{u}(x) - u(x)| \leq \max_{|y - x| = h} |\tilde{u}(y) - u(y)| = \max_{|y - x| = h} |\delta(y)|.$$

The scheme is nonlinear, but the nonlinearity is of a particular kind: it consists of a minimum of linear terms.

The consistency error doesn't give a global solution error. As we observed above, the consistency error for the scheme is second order. Since the equations support solutions which are not twice differentiable, (for example, we can take $u(x, y) = |x|$ as solution for the convex envelope, or the Monge-Ampère equation), the interpolation error on an exact solution will not even be second order accurate. Thus the lack of a higher order convergence rate is not a deficiency of the scheme. Rather, the scheme is robust enough to handle non-smooth solutions.

2. Convergence Theory

The advantage of using viscosity solutions is that the theory applies to whole classes of equations; the machinery for proving existence, uniqueness and stability
results is the same for each equation. We review the corresponding theory for convergence of numerical schemes in the next sections.

2.1. Convergence of approximation schemes. One advantage of the viscosity solutions formulation is that we need only check consistency on twice continuously differentiable functions, even though we allow for non-differentiable solutions.

The fundamental result for the convergence of numerical schemes for fully nonlinear, degenerate equations is due to Barles and Souganidis.

**Theorem 1** (Convergence of Approximation Schemes [1]). Consider a degenerate elliptic equation, \((\text{PDE})\), for which there exist unique viscosity solutions. A consistent, stable approximation scheme converges uniformly on compact subsets to the viscosity solution, provided it is monotone.

Note the theorem only gives a uniform convergence rate. As mentioned above, the lack of regularity of solutions in the general setting prevents asserting convergence rates like \(O(h^2)\), despite second order consistency for schemes.

Monotonicity is an essential property which we discuss in depth below. The definition of consistency is in terms of twice differentiable test functions: it amounts to checking Taylor series computations, see Definition 2. The theorem requires only a mild form of stability: our schemes always satisfy a stronger notion of stability: they are nonexpansive in the uniform norm.

2.2. The comparison principle and monotonicity. The comparison principle is a nonlinear generalization of the maximum principle. When applied to \((\text{PDE})\) it can be written

\[
(2.1) \quad \text{if } F[u_1] \leq F[u_2] \text{ in } \Omega \text{ and } u_1 \leq u_2 \text{ on } \partial \Omega \text{ then } u_1 \leq u_2 \text{ in } \Omega
\]

Some authors prefer to avoid the minus sign in front of \(F[u]\) in \((\text{PDE})\), in which case the first inequality would be switched. A consequence of the comparison principle is the monotonicity of the solution operator. If \(u, v\) are solutions of the same equations, \((\text{PDE})\), but with different boundary conditions then

\[
(2.2) \quad u \leq v \text{ on } \partial \Omega \text{ implies } u \leq v \text{ in } \Omega.
\]

We prefer to write this property more abstractly as follows. For a given equation \(F\), as in \((\text{PDE})\), let \(S_F : C(\partial \Omega) \to C(\Omega)\) be the solution operator, which takes boundary data \(g\), to the solution \(u\). Thus

\[
\begin{align*}
S(g)(x) &= g(x) \quad \text{for } x \in \partial \Omega, \\
F[S(g)](x) &= 0 \quad \text{for } x \in \Omega.
\end{align*}
\]

The the monotonicity property \((2.2)\) becomes

\[
(2.3) \quad g_1 \leq g_2 \text{ implies } S_F(g_1) \leq S_F(g_2).
\]

In this form, monotonicity of the solution operator takes the same form for the solution operator of the numerical scheme. Embracing a duplication of notation, we write \(S_F\) for the solution operator of the corresponding numerical scheme, and take \((2.3)\) as the definition of monotonicity.
2.3. Degenerate elliptic schemes. Taking into account the Barles-Souganidis result, the problem of building convergent schemes for fully nonlinear equations is reduced to the problem of building consistent, stable, monotone schemes. However, building monotone schemes remains a challenge for many interesting equations.

While monotonicity is a global property of the solution operator, we can take advantage of local conditions for the scheme which ensure monotonicity. Degenerate elliptic schemes are defined by a local structure condition which is the discrete analogue of the degenerate ellipticity condition for PDEs. The definition follows.

A finite difference scheme consists of a grid, \( G \), and a finite difference equation, \( F \) which is defined on grid functions \( u \). The grid is a graph consisting of a set of points, \( x_i \in \Omega, i = 1, \ldots, N \), each endowed with a list of neighbors, \( N(i) \). The direction vectors \( x(j) - x(i) \) associated with the neighbors \( j \in N(i) \), defines the stencil at \( x_i \). A grid function is simply a real-valued function defined on the grid, with values \( u_i = u(x_i) \).

Writing the finite difference equation in the form
\[
F^i[u] \equiv F^i(u_i, u_i - u_j)_{j=N(i)} \equiv F^i(u_i, u_i - u_j), \quad i = 1, \ldots, N,
\]
where \( u_j \) is shorthand for the list of neighbors \( u_{j|j=N(i)} \). Using this form emphasizes that each term \( u_i - u_j \) (which usually occurs with the factor \( 1/|x_i - x_j| \)) represents the first order approximation to the derivatives in the direction of the neighbor, \( du/dv_{ij} \), where \( v_{ij} = x_j - x_i/|x_j - x_i| \),
\[
\frac{du}{dv_{ij}} = \nabla u(x_i) \cdot v_{ij} = \frac{u_i - u_j}{|x_i - x_j|} + O(|x_j - x_i|)
\]
Second derivatives in the direction \( v_{ij} \) are approximated by averaging \( du/dv_{ij} \) and \( du/dv_{ik} \), where \( v_{ik} = -v_{ij} \).
\[
\frac{du^2}{dv_{ij}^2} = v_{ij}^T D^u(x_i) v_{ij} = \frac{u_i + u_k - 2u_i}{|x_i - x_j|^2} + O(|x_j - x_i|^2)
\]

A boundary point is a grid point with no neighbors. Dirichlet boundary conditions \( u = g \) may be imposed at boundary points by setting \( F^i[u] = u_i - g(x_i) \). A solution is a grid function which satisfies \( F[u] = 0 \).

We write \( u = S_F(g) \) when we wish to specify the equation, for the solution operator, provided it is well-defined. We regard the solution operator as a mapping from the Dirichlet data on the boundary points to grid functions.

We now define degenerate elliptic schemes.

**Definition 1 ([16]).** The scheme \( F \) is degenerate elliptic if each component \( F^i \) is nondecreasing in each variable.

**Remark.** We emphasize that each component of the scheme \( F^i \) is a nondecreasing function of \( u_i \) and the differences \( u_i - u_j \) for \( j = 1, \ldots, N(i) \).

**Theorem 2 ([16]).** Under mild analytic conditions, degenerate elliptic schemes are monotone, and stable (in fact they are non-expansive in the uniform norm). They can be solved by an iterative method with time step \( dt = K^{-1} \), where \( K \) is the Lipschitz constant of the scheme, regarded as a function from \( \mathbb{R}^N \to \mathbb{R}^N \).

In addition, we can combine degenerate elliptic schemes for simple equations into schemes for more complicated ones. In particular, once we have built degenerate elliptic schemes for \( \lambda_- \), \( \lambda_+ \), we build schemes for functions of the eigenvalues by simple substitution, see [4, 2].
3. **Wide stencil finite difference schemes**

Equation (1.2) gives a semi-discrete approximation to the operator $\lambda_-[u]$. In order to fully discretize the scheme, we need to take second derivatives only in directions $\theta$ which are available on the computational grid. The limited *directional resolution*, of a finite grid requires the introduction of a second parameter $d\theta$, in addition to the spatial resolution, $h$. The spatial resolution is improved by using more grid points, the directional resolution is improved by increasing the size of the stencil.

The indices of neighboring grid points for successively wider stencil schemes are indicated in Table 1. The first two levels are displayed in Figure 1(a).

### Table 1. Neighbors of reference point (i,j) in the first quadrant.

| Level | Neighbors of (i,j) in First Quadrant |
|-------|--------------------------------------|
| 1     | (i+1, j) (i+1, j+1)                  |
| 2     | (i+1, j) (i+1, j+1) (i+2, j+1) (i+1, j+2) |
| 3     | (i+1, j) (i+1, j+1) (i+2, j+1) (i+1, j+2) |
|       | (i+1, j+3) (i+3, j+1) (i+2, j+3) (i+3, j+2) |

The neighbors in the other quadrants are given by rotation.

**Figure 1.** (a) Computational stencils for the 9 and 17 point schemes. (b) Computational stencil near the boundary.

3.1. **Boundary grid points.** Near the boundary of the domain, the stencil width is limited by available grid points. Second derivatives in a given direction are still computed, but to lower accuracy. This is achieved in two steps. First the boundary value at the intermediate point is determined, either by using the given value, if it is available, or by quadratic interpolation, using the three nearest neighbors. Next, the intermediate value is used to compute a first order accurate second derivative. To give a concrete example, in a typical calculation, we might compute

$$\frac{d^2u}{dv^2} = \frac{1}{5h^2} (u(i+2,j+i) - 2u(i,j) + u(i-2,j-1)) + O(h^2), \quad v = \frac{(2,1)}{\sqrt{5}}.$$  

Near the boundary, if the $i-2$ value is not available, we compute instead

$$\frac{d^2u}{dv^2} = \frac{4}{15h^2} (u(i+2,j+i) - 3u(i,j) + 2u(i-1,j-1/2)) + O(h).$$
Here the \( u(i-1, j-1/2) \) value is computed using quadratic interpolation at the indices \((i-1, k), k = j-1, j, j+1\) via

\[
u(i-1, j-1/2) = \frac{3}{8}u(i-1, j-1) + \frac{6}{8}u(i-1, j) - \frac{1}{8}u(i-1, j+1).\]

We note that this gives different values for the intermediate value depending on the direction of the computed derivate, but this is acceptable.

3.2. Local spatial and directional resolution. On a regular cartesian grid, let the stencil at the reference point \( x_0 \) consist of the neighbors \( x_1, \ldots, x_N \) (illustrated by open and closed circles respectively in Figure 1). Write the direction vectors in polar coordinates

\[
v_i \equiv x_i - x_0 \equiv h_i v_\theta_i\]

We assume the stencil is symmetric,

\[
(3.1) \quad \text{if } v_i \text{ is a direction vector, then so is } -v_i
\]

and define the local spatial resolution

\[
(3.2) \quad h(x_0) \equiv \max_i h_i
\]

and the local directional resolution

\[
(3.3) \quad d\theta(x_0) \equiv \max_{\theta \in [-\pi, \pi]} \min_i |\theta - \theta_i|.
\]

For the purpose of calculating the time step for iterative schemes, we also define

\[
(3.4) \quad h(x_0) \equiv \min_i h_i.
\]

4. The Schemes

In this section we present the schemes. From the basic building block of the discrete second derivative, we build monotone schemes for \( \lambda_- , \lambda_+ \) on a wide stencil uniform grid. These operators are combined to give schemes for general functions of the eigenvalues.

We show the scheme is degenerate elliptic, which allows us to establish consistency, stability, and monotonicity, and thereby prove convergence using Theorem 1.

4.1. Definition of the schemes for \( \lambda_-, \lambda_+ \). Now we can discretize the eigenvalues by the following simple formula.

\[
\lambda_-^{h,d\theta}[u](x_0) = \min_i \frac{u(x_0 + v_i) - 2u(x_0) + u(x_0 - v_i)}{|v_i|^2}
\]

and similarly for \( \lambda_+ \),

\[
\lambda_+^{h,d\theta}[u](x_0) = \max_i \frac{u(x_0 + v_i) - 2u(x_0) + u(x_0 - v_i)}{|v_i|^2}.
\]

Lemma 1. The schemes (4.1)(4.2) are degenerate elliptic.

Proof. Each discrete second derivative in the direction \( v_i \) is the average of the terms \( u(x_0 \pm v_i) - u(x_0) \) which have the form \( u_j - u_i \), so the are certainly non-decreasing in \( u_j - u_i \). Taking a minimum (or maximum) of several non-decreasing functions yields a non-decreasing function. \( \square \)
4.2. Definition of the schemes for functions of the eigenvalues. Given the schemes (4.1), (4.2) for $\lambda_-$ and $\lambda_+$, respectively, we build schemes for any non-decreasing function $G(\lambda_- , \lambda_+)$ by simple substitution

(4.3) \[ G_{h,d\theta}(\lambda_- , \lambda_+) \equiv G(\lambda_{-,h\theta} , \lambda_{+,h\theta}). \]

These schemes are clearly consistent, with local accuracy determined by substitution. In addition, they are degenerate elliptic, as we show below.

5. Convergence

Next we prove convergence of the schemes. By Theorem 1, we need only to verify that the schemes are consistent and monotone. The first property involves a Taylor series calculation. The effect of the parameter $d\theta$ which reflects the directional resolution of the grid means requires an additional argument beyond the standard Taylor series arguments used for consistency in $h$.

5.1. Monotonicity.

Lemma 2 (Monotonicity). The scheme $G_{h,d\theta}(\lambda_- , \lambda_+)$, defined by (4.1), (4.2), (4.3) is monotone, provided $G : \mathbb{R}^2 \rightarrow \mathbb{R}$ is a nondecreasing function in each variable. In particular, the schemes for $G^{MA}, G^{CE}, G^{M}, G^{m}$ are monotone. In addition, the schemes for $G - f(x)$ are monotone. More generally, if $G_2$ is a degenerate elliptic scheme, then so is $G + G_2$.

The last part of the theorem allows the schemes to be used as building blocks for more complicated equations. For example, to obtain schemes for the PDE $G - f(Du, x) = 0$ we simply combine the scheme for $G$ with a degenerate elliptic scheme for $f$.

Proof. By virtue of Lemma 1 the schemes (4.1), (4.2) are degenerate elliptic. As was already shown in [16], it follows from the properties of nondecreasing functions, that if $F_1$ and $F_2$ are degenerate elliptic finite difference schemes, then so is $F = G(F_1 , F_2)$, provided $G : \mathbb{R}^2 \rightarrow \mathbb{R}$ is a nondecreasing function. It follows from Theorem 2 that degenerate elliptic schemes are monotone.

Furthermore, combining the scheme for $G$ with the trivial scheme for $f(x_i) = f(x_i)$ does affect the ordering properties, so the result is still degenerate elliptic. The last remark follows from the fact that a sum of non-decreasing functions is still non-decreasing. \qed

5.2. Consistency.

Definition 2 (Consistency). We say the scheme $F^{h,d\theta}$ is consistent with the equation (PDE) at $x_0$ if for every twice continuously differentiable function $\phi(x)$ defined in a neighborhood of $x_0$,

\[ F^{h,d\theta}[\phi](x_0) \rightarrow F[\phi](x_0) \quad \text{as} \quad h, d\theta \rightarrow 0. \]

The global scheme defined on $\Omega$ is consistent if the limit above holds uniformly for all $x \in \Omega$. (The domain is assumed to be closed and bounded).

Note that in practice we test consistency only at grid points. We implicitly assume the existence of an interpolation function that gives the value of the scheme at points in the domain which are not grid points. Typically, we can use linear
interpolation, over simplices which partition the domain. Other types of monotone interpolation are also permitted.

Next we define the local spatial and directional resolution, and compute the consistency error of the scheme. The consistency error for the scheme gives a formal accuracy estimate for the scheme, which as we remarked above, is not a rigorous convergence rate.

**Lemma 3 (Consistency).** Let \( x_0 \) be a reference point with neighbors \( x_1, \ldots, x_N \), and direction vectors \( v_i = x_i - x_0, i = 1, \ldots, N \), arranged symmetrically so that (3.1) holds. Use the schemes for \( \lambda^- \), \( \lambda^+ \) given by (4.1), (4.2), respectively. If \( \phi(x) \) is a twice continuously differentiable function defined in a neighborhood of the grid, then

\[
\lambda_- [\phi](x_0) - \lambda_-^{h, d\theta} [\phi](x_0) = O(h^2 + (\lambda_+ - \lambda_-) d\theta^2)
\]

and likewise for \( \lambda_+ \). Consistency also holds for the scheme (4.3).

**Proof.** The standard Taylor series computation gives

\[
\frac{u(x_0 + v_i) - 2u(x_0) + u(x_0 - v_i)}{|v_i|^2} = \frac{d^2 u}{d v_i^2} + O(h_i^2).
\]

Let \( M \) be a given symmetric \( 2 \times 2 \) matrix. Without loss of generality we may choose coordinates so that \( M \) is diagonal. For the unit vector \( v_\theta \), write

\[
v_\theta^T M v_\theta = \begin{bmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{bmatrix} \begin{bmatrix} \lambda_- \\ \lambda_+ \end{bmatrix} = \lambda_- \cos^2 \theta + \lambda_+ \sin^2 \theta
\]

This last equation recovers the Rayleigh-Ritz characterization of the eigenvalues: minimizing over \( \theta \) yields \( \lambda_- \), and maximizing yields \( \lambda_+ \).

Next minimize over a discrete set \( \theta \in \{\theta_1, \ldots, \theta_N\} \), with resolution \( d\theta \), defined by (3.3). Calculate

\[
\min_{\theta \in \{\theta_1, \ldots, \theta_N\}} v_\theta^T M v_\theta = \lambda_- + (\lambda_+ - \lambda_-) O(d\theta^2),
\]

which follows from the fact that \( \cos(\pi - \theta) = -1 + O(\theta^2) \).

A similar argument works for \( \lambda_+ \), with minimum replaced by maximum.

Combining the \( h \) error with the \( d\theta \) error gives the desired result.

Inserting (5.1) into the scheme (4.3) shows that \( G^{h,d\theta}(\lambda_-, \lambda_+) \to G(\lambda_-, \lambda_+) \) as \( h, d\theta \to 0 \). For \( G^{lin} \), the accuracy is the same. For \( G^{MA} \) expanding the product shows that the accuracy is multiplied by a factor of \( (\lambda_+ + \lambda_-) \). \( \square \)

**Remark.** The factor of \( \lambda_+ - \lambda_- \) in the error term shows that the \( d\theta \) error arises differently from the \( h \) error. The spatial consistency error arises as the difference between the function and a quadratic function. The directional consistency error arises as the difference between the function an one whose Hessian is a multiple of the identity.
5.3. Convergence.

**Theorem 3** (Convergence). Suppose $G$ is a non-decreasing function of the two variables, and that unique viscosity solutions exist for the equation $G(\lambda_-, \lambda_+) = f(x)$. Then the finite difference scheme given by (4.3) converges uniformly on compacts subsets of $\Omega$ to the unique viscosity solution of the equation.

**Proof.** By virtue of Theorem 1, we need only show that the scheme is consistent and monotone. Consistency follows from Lemma 3 and monotonicity follows from Lemma 2. □

5.4. Explicit optimal time steps. The schemes yield a fully nonlinear equation defined on grid functions. In order to solve $F^{h,d\theta}\{u\} = 0$ we perform the iteration

$$u^{n+1} = u^n + dt F^{h,d\theta}\{u^n\}.$$  

By Theorem 2 this iteration will converge to a fixed point which is a solution of the equation, provided the nonlinear CFL condition is respected,

$$dt \leq K^{-1},$$

where $K$ is the Lipschitz constant of the scheme, regarding $F$ map from grid functions to grid functions. Here $dt$ is a parameter, so-named because it corresponds to the time step in the explicit Euler discretization of $u_t = F[u]$.

Next we compute the constant $K$ for each of the schemes.

In the above, we can compute $K(i)$ which is the Lipschitz constant of $F^i$ at the grid point $x_i$. We can set $K = \max_i K(i)$, and set $dt$ accordingly. Alternatively, we can update at each grid point, using a local $dt$, and choosing the ordering of the grid points for the update.

We can do this by noting that the Lipschitz constant for $\lambda^{h,d\theta}_-, \lambda^{h,d\theta}_+$, defined by (4.1) (4.2), is given by $2h^2$. We write this as,

$$K(\lambda^{h,d\theta}_-) = K(\lambda^{h,d\theta}_+) = \frac{2}{h^2},$$

where $h$ is given by (3.4). This follows since the Lipschitz constant of a maximum or minimum is bounded by the maximum of the Lipschitz constants. Finally, we have

$$K\left(\frac{u(x_0 + v_i) - 2u(x_0) + u(x_0 - v_i)}{|v_i|^2}\right) = \frac{2}{|v_i|^2}.$$ 

Taking the worst possible value for $|v_i|$, and using the definition of $h$ (3.4) gives the result.

When the equation is linear in the eigenvalues of the Hessian, as is the case for the convex envelope equation, $G^{CE}$, and the maximal and minimal Pucci equations, $G^M, G^m$, we calculate the time step as follows. Write

$$G^{lin}(\lambda^{h,d\theta}_-, \lambda^{h,d\theta}_+) = c_- \lambda^{h,d\theta}_- + c_+ \lambda^{h,d\theta}_+,$$

where $c_-, c_+ > 0$. Since $G$ is linear, we have

$$K(G^{lin}) = \frac{2}{h^2} (c_- + c_+).$$

Therefore the optimal time step is

$$dt(G^{lin}) = (c_- + c_+) \frac{h^2}{2}.$$
Table 2. Errors for the exact solution of Monge-Ampère, $u(x) = \exp(|x|^2/2)$, $f(x) = (1 + |x|^2) \exp(|x|^2)$, using the 9 and 17 point schemes.

| n    | $\|u - u_n^9\|_\infty$ | $\|u - u_n^{17}\|_\infty$ | $\|u - u_n^9\|_2$ | $\|u - u_n^{17}\|_2$ |
|------|------------------------|-------------------------------|-------------------|---------------------|
| 21   | $1.8 \times 10^{-3}$   | $5.2 \times 10^{-3}$         | $4.1 \times 10^{-3}$ |                      |
| 41   | $1.3 \times 10^{-3}$   | $1.6 \times 10^{-5}$         | $7.2 \times 10^{-6}$ |                      |
| 81   | $7.9 \times 10^{-4}$   | $3.9 \times 10^{-6}$         | $1.3 \times 10^{-6}$ |                      |
| 128  | $5.0 \times 10^{-4}$   | $1.3 \times 10^{-6}$         | $4.0 \times 10^{-7}$ |                      |

For $G^M_A(\lambda_\pm^{h,d\theta},\lambda_{+}^{h,d\theta}) = \max(\lambda_\pm^{h,d\theta},0)\max(\lambda_{+}^{h,d\theta},0)$ apply the chain rule to get

\[
K(G^M_A) = \max(\lambda_{+}^{h,d\theta},0)K(\lambda_{+}^{h,d\theta}) + \max(\lambda_\pm^{h,d\theta},0)K(\lambda_\pm^{h,d\theta})
\]

\[
= (\max(\lambda_{+}^{h,d\theta},0) + \max(\lambda_\pm^{h,d\theta},0)) \frac{2}{h^2}.
\]

Thus

\[
dt(G^M_A) = \frac{h^2}{2} (\max(\lambda_{+}^{h,d\theta},0) + \max(\lambda_\pm^{h,d\theta},0))^{-1}.
\]

The resulting allowable time step in this case can be significantly less restrictive than the one given by the dimensional scaling $O(h^4)$.

6. Computations

In the section we perform several convergence tests for the Monge-Ampère equation and the Pucci equations, using examples as a basis of comparison from the works of Dean and Glowinski. In addition, we perform computations to test the accuracy dependence on the $d\theta$ parameter.

6.1. Numerical solutions of Monge-Ampère. We begin with an exact solution to the Monge-Ampère equation. For radial functions, the equations takes the form

\[
\frac{u'u''}{\rho}, \quad \rho = |x - x_0|.
\]

An solution is given by

\[
u(x) = \exp(|x|^2/2), \quad f(x) = (1 + |x|^2) \exp(|x|^2).
\]

Numerical solutions were computed for this example, see Table 2.

A second example involved solved the Monge-Ampère equation $G^M_A[u] = 1$ with constant boundary data 1 on the unit square. Since the function is constant on the boundary, which means one of $u_{xx}$ or $u_{yy}$ is zero on the boundary, classical solutions don’t exist in this case. In particular, along the boundary, one of the eigenvalues is zero, so the larger eigenvalue must approach infinity as the boundary is approached. However the method used converges independently of the regularity of the solution.

Solutions were computed solutions using the 9 point and 17 point stencil. The difference in accuracy between the smaller and the wider stencil was not significant. See Table 3. Since the exact solution is unknown, we compared the minimum value of the solution, as well as the difference between solutions in the maximum norm. The solution is plotted in Figure 2. The figure was coarsened for the plot, since too many lines appear using the full grid. The contour plot shows that the level sets go
from circles, which develop high curvature along the diagonals until they approach a square. The computed maximum and minimum eigenvalues, ranged from on the order of .1 and 10 to .001 and 100 as $n$ increased.

Table 3. Value at the minimum for the solution of Monge-Ampère with constant boundary data and $f(x) = 1$. Difference between the solutions.
6.2. Study in the directional resolution error. The consistency error of the scheme is given by (5.1). This section studies the numerical error of the scheme, using an example designed to measure the error in $d\theta$. For this purpose, we take a quadratic solution of the form

$$u_{\alpha,\theta}(x, y) = (\cos^2 \theta + \alpha \sin^2 \theta) x^2 + (1 - \alpha) \cos \theta \sin \theta xy + (\cos^2 \theta \alpha + \sin^2 \theta) y^2$$

with $1 < \alpha$, so that $\lambda_-[u](x, y) = 1$ and $\lambda_+[u](x, y) = \alpha$. Then $u_{\alpha,\theta}$ is a solution of Monge-Ampère with $f(x, y) = \alpha$. Solutions were computed for $\theta = \pi/8, \pi/4, \ldots, \pi/20$, with $\alpha = 5$ and $n = 80$. The computed solution was numerically exact whenever the direction $\theta$ lined up with the grid. However, for the 9 point stencil, the error was large, and reached a maximum of .13 with $\theta = \pi/8$, which was midway between the grid directions. For the 17 and 33 point stencils, the error was acceptable, with a maximum of .004 for the 17 point scheme, and .002 for the 33 point scheme, again reaching a maximum when $\theta$ was between the grid directions. See Figure 3.

We can conclude that when we have a combination of a large ratio between $\lambda_-$ and $\lambda_+$ along directions which are not favorable to the grid, the $d\theta$ error can dominate. In more tame situations, the $d\theta$ error does not.

6.3. Numerical solution of the Pucci equation. We follow examples from [9]. For the equation

$$\alpha \lambda_+ + \lambda_- = 0 \text{ in } \Omega,$$

use the exact radial solution

$$\psi(x, y) = -\rho^{1-\alpha}, \quad \text{where } \rho(x, y) = \sqrt{(x + 2)^2 + (y + 2)^2}$$

on $[-1, 1]^2$. We solved for $\alpha = 2, 2.5, 3, 5$, using the 17 point stencil. The errors are displayed in Table 4.

Next we consider an example with step function boundary data,

$$u(x, y) = \begin{cases} 1 & \text{if } |x| > .5 \text{ and } |y| > .5, \\ 0 & \text{otherwise.} \end{cases}$$

We solved on $[-1, 1]^2$, with $n = 256, 128, \ldots$, for $\alpha = 2, 2.5, 3$, using the 17 point scheme with interpolation on the boundary. Each time the spatial resolution was increased, there was a large error between the previous solution near the discontinuity at the boundary, which was better resolved. The solutions are increasing as a function of $\alpha$, by the comparison principle. See Figure 4.
6.4. **Discrete Comparison Principle.** Numerical solutions of respected the discrete comparison principle, when it was true at the continuous level. For example, if $u_1$ is the solution of the convex envelope equation,

$$-\lambda_- [u_1] = 0,$$

and $u_2$ is the solution of the Pucci equation

$$-(2\lambda_- + \lambda_+) [u_2] = 0,$$

with the same boundary data $g(x)$ for each equation, then

$$-(2\lambda_- + \lambda_+) [u_1] = -\lambda_+ [u_1] \leq 0.$$

So $u_1$ is a subsolution of the Pucci equation. By the comparison principle (2.1), we conclude

$$u_1 \leq u_2.$$  

By Theorem 2, we expect that numerical solutions of the difference equations solved on the same grid respect the comparison principle as well. We verified that this held numerically, for each of the equations, and also using coefficients which were functions of $x$. For example the variable coefficient Pucci equations $(A(x)\lambda_- + \lambda_+)$, where $1 \leq A(x) \leq 2$. See Figure 5.

6.5. **Boundary continuity of solutions.** We remark that solutions need not be continuous up to the boundary. For example, if the boundary data is not convex, then there is no way to build a convex function which agrees at the boundary. Since both solutions of (MA) and (CE) are required to be convex, we can’t expect boundary continuity. However this is a feature of the equation, not the scheme, so it shows the robustness of the scheme.

6.6. **Accelerating Iterations.** The iteration (5.2) is a simple, explicit, convergent method to find the solution of the difference equation. While it may thousands of iterations to converge, on the largest grids used, solution time was a few minutes.
Figure 5. Solutions of the Pucci equation $2\lambda_- + \lambda_+ = 0$ and the convex envelope, $\lambda_- = 0$, with boundary data $x^2 - y^2$, on the unit square.

| n  | 16  | 32  | 64  | 128 | 256 |
|----|-----|-----|-----|-----|-----|
| guess | 240 | 770 | 2700 | 9620 | 47400 |
| interpolated | 170 | 430 | 920 | 1810 | 3490 |
| ratio | 1.4 | 1.8 | 2.9 | 5.3 | 13.6 |

Table 5. Number of iterations required for convergence comparing initial data given by a rough guess with the interpolated solution from a coarser grid.

A simple method to accelerate iterations was to interpolate the solution on a coarser grid. Given a grid size $n \times n$, we solved the equation on a coarser grid of $n/2 \times n/2$, and interpolated the solution onto the the finer grid. This proved to be faster than solving with a less accurate initial guess.

We compared the number of iterations to get a solution using the two methods. See Table 5 which records data from the solution of the convex envelope. The results suggest that the number of iterations using the coarse interpolant appears to $O(n)$, compared to superlinear growth using less accurate initial data. Each iteration (5.2) is an explicit application of the finite difference scheme at each grid point.

References

[1] Guy Barles and Panagiotis E. Souganidis. Convergence of approximation schemes for fully nonlinear second order equations. *Asymptotic Anal.*, 4(3):271–283, 1991.

[2] Luis A. Caffarelli and Xavier Cabrè. *Fully nonlinear elliptic equations*, volume 43 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 1995.

[3] Luis A. Caffarelli and Mario Milman, editors. *Monge Ampère equation: applications to geometry and optimization*, volume 226 of *Contemporary Mathematics*, Providence, RI, 1999. American Mathematical Society.

[4] Michael G. Crandall, Hitoshi Ishii, and Pierre-Louis Lions. User’s guide to viscosity solutions of second order partial differential equations. *Bull. Amer. Math. Soc. (N.S.)*, 27(1):1–67, 1992.

[5] Michael G. Crandall and Pierre-Louis Lions. Convergent difference schemes for nonlinear parabolic equations and mean curvature motion. *Numer. Math.*, 75(1):17–41, 1996.

[6] Edward J. Dean and Roland Glowinski. Numerical methods for fully nonlinear elliptic equations of the Monge-Ampère type. *Comput. Methods Appl. Mech.*
[7] Edward J. Dean and Roland Glowinski. Numerical solution of the two-dimensional elliptic Monge-Ampère equation with Dirichlet boundary conditions: an augmented Lagrangian approach. *C. R. Math. Acad. Sci. Paris*, 336(9):779–784, 2003.

[8] Edward J. Dean and Roland Glowinski. Numerical solution of the two-dimensional elliptic Monge-Ampère equation with Dirichlet boundary conditions: a least-squares approach. *C. R. Math. Acad. Sci. Paris*, 339(12):887–892, 2004.

[9] Edward J. Dean and Roland Glowinski. On the numerical solution of a two-dimensional Pucci’s equation with Dirichlet boundary conditions: a least-squares approach. *C. R. Math. Acad. Sci. Paris*, 341(6):375–380, 2005.

[10] Edward J. Dean, Roland Glowinski, and Tsrong-Whay Pan. Operator-splitting methods and applications to the direct numerical simulation of particulate flow and to the solution of the elliptic Monge-Ampère equation. In *Control and boundary analysis*, volume 240 of *Lect. Notes Pure Appl. Math.*, pages 1–27. Chapman & Hall/CRC, Boca Raton, FL, 2005.

[11] Lawrence C. Evans. Partial differential equations and Monge-Kantorovich mass transfer. In *Current developments in mathematics, 1997 (Cambridge, MA)*, pages 65–126. Int. Press, Boston, MA, 1999.

[12] David Gilbarg and Neil S. Trudinger. *Elliptic partial differential equations of second order*. Classics in Mathematics. Springer-Verlag, Berlin, 2001. Reprint of the 1998 edition.

[13] Theodore S. Motzkin and Wolfgang Wasow. On the approximation of linear elliptic differential equations by difference equations with positive coefficients. *J. Math. Physics*, 31:253–259, 1953.

[14] Adam M. Oberman. A convergent monotone difference scheme for motion of level sets by mean curvature. *Numer. Math.*, 99(2):365–379, 2004.

[15] Adam M. Oberman. A convergent difference scheme for the infinity Laplacian: construction of absolutely minimizing Lipschitz extensions. *Math. Comp.*, 74(251):1217–1230 (electronic), 2005.

[16] Adam M. Oberman. Convergent difference schemes for degenerate elliptic and parabolic equations: Hamilton-Jacobi equations and free boundary problems. *SIAM J. Numer. Anal.*, 44(2):879–895 (electronic), 2006.

[17] Adam M. Oberman. The convex envelope is the solution of a nonlinear obstacle problem. *Proc. Amer. Math. Soc.*, 135:1689–1694, 2007.

[18] V. I. Oliker and L. D. Prussner. On the numerical solution of the equation $((\partial^2 z/\partial x^2)\partial^2 z/\partial y^2)−((\partial^2 z/\partial x\partial y))^2 = f$ and its discretizations. I. *Numer. Math.*, 54(3):271–293, 1988.

[19] A. V. Pogorelov. *Monge-Ampère equations of elliptic type*. Translated from the first Russian edition by Leo F. Boron with the assistance of Albert L. Rabenstein and Richard C. Bollinger. P. Noordhoff Ltd., Groningen, 1964.

DEPARTMENT OF MATHEMATICS, SIMON FRASER UNIVERSITY
E-mail address: aoberman@sfu.ca