**Electronic Transactions on Numerical Analysis.**
**Volume 22, pp. 71-96, 2006.**

**ISSN 1068-9613.**

**ETNA**
**Kent State University**
**etna@mcs.kent.edu**

# AN AUGMENTED LAGRANGIAN APPROACH TO THE NUMERICAL SOLUTION OF THE DIRICHLET PROBLEM FOR THE ELLIPTIC MONGE-AMPÈRE EQUATION IN TWO DIMENSIONS*

E. J. DEAN† AND R. GLOWINSKI‡

**Abstract.** In this article, we discuss the numerical solution of the Dirichlet problem for the real elliptic Monge-Ampère equation, in two dimensions, by an augmented Lagrangian based iterative method. To derive the above algorithm, we take advantage of a reformulation of the Monge-Ampère problem as a saddle-point one, for a well-chosen augmented Lagrangian functional over the product of appropriate primal and dual sets. The convergence of the finite element approximation and of the iterative methods described in this article still has to be proved, however, on the basis of numerical experiments reported in this article, it is safe to say that: (i) The augmented Lagrangian methodology discussed here provides a sequence converging to a solution of the Monge-Ampère problem under consideration, if such a solution exists in the space $H^2(\Omega)$. (ii) If, despite the smoothness of its data, the above problem has no solution, the above augmented Lagrangian method solves it in a least-squares sense, to be precisely defined in this article.

**Key words.** elliptic Monge-Ampère equation, augmented Lagrangian algorithms, mixed finite element approximations

**AMS subject classifications.** 35J60, 65F10, 65N30

**1. Introduction.** These last years have been witnessing a surge of interest for real Monge-Ampère equations. Evidences of this interest can be found, for example, in the fact that at the last ICM meeting (Beijing 2002) there was a plenary lecture on these topics (and closely related ones) by L. A. Caffarelli (ref. [1]) and an invited one by Y. Brenier (ref. [2]). The interest of the authors of this article is stemming from their interaction with David Bao, a colleague at University of Houston, one of the goals of this collaboration being to solve, ultimately, some Monge-Ampère related equations coming from Riemannian and Finsler Geometries (refs.[3] and [4]), and with L. A. Caffarelli at U. T. Austin. This article is dedicated to the numerical solution of the Dirichlet problem for the real two-dimensional elliptic Monge-Ampère equation (called E-MAD (!) problem in the sequel). The computational methodology used for the solution of the above problem is based on the combination of a mixed finite element approximation, reminiscent of those used in, e.g., [5]-[12], for the solution of linear and nonlinear bi-harmonic problems, with a reformulation of (E-MAD) as a saddle-point problem, for a well-chosen augmented Lagrangian functional over an appropriate product of primal and dual sets. The convergence of the finite element approximations and of the iterative methods described in this article still has to be proved, but on the basis of the various numerical experiments which have been performed we conjecture that the solution methods to be described in the following sections can find the solutions of the Monge-Ampère equation which have the $H^2$-regularity, and least-squares solutions if $H^2$ solutions do not exist, while the data are smooth enough. The mathematical analysis of real Monge-Ampère and related equations has generated a fairly large literature body; let us mention the following references (among many others and in addition to [1]-[3]): [13], [14], [15, Chapter 4], [16]-[21]. Applications to Mechanics and Physics can be found in [22]-[25], [28] (see also the references therein). As far as we know, Monge-Ampère equations have motivated very few

*Received September 15, 2004. Accepted for publication February 24, 2005. Recommended by R. Lehoucq. This work was supported in part by the National Science Foundation under grant DMS-0412267.

†Department of Mathematics, University of Houston, 4800 Calhoun Road, Houston, TX 77204-3008 (dean@math.uh.edu).

‡Department of Mathematics, University of Houston, 4800 Calhoun Road, Houston, TX 77204-3008 (roland@math.uh.edu).

computationally oriented publications, but we hope that this will change in the future. Among those few references with a strong numerical flavor let us mention [22], [25]-[28]; the method discussed in [25], [26], [28] is very geometrical in nature, in contrast with the method in the present article (introduced in [27]; see also [29]), which is of the variational type and can be applied to the solution of other fully nonlinear elliptic equations. An interesting - and important - feature of the method described in the present article is that in those cases where the Monge-Ampère equation under consideration has no smooth solution, despite the smoothness of its data, it provides a "generalized solution" whose nature will be made more precise in Sections 4 and 6. A concise description of the method described in this article can be found in refs. [27] and [29], which contain also some preliminary numerical results.

This article is structured as follows: In Section 2, we introduce the Dirichlet problem for the elliptic Monge-Ampère equation in two-space dimensions; we identify there a simple situation where smooth data do not imply the existence of a smooth solution. In Section 3, assuming that the data are sufficiently smooth, we reformulate (E-MAD) as a constrained optimization problem in $H^2(\Omega)$, to which we associate a well chosen augmented Lagrangian functional $\mathcal{L}_r$ whose saddle-points will provide us with solutions to (E-MAD). In Section 4, we discuss an Uzawa-Douglas-Rachford algorithm for the computation of the saddle-points of $\mathcal{L}_r$. The mixed finite element implementation of the above algorithm is discussed in Section 5, while related numerical results are shown in Section 6.

**2. Formulation of the Dirichlet problem for the elliptic Monge-Ampère equation in two dimensions. Preliminary remarks.** Let $\Omega$ be a bounded domain of $\mathbf{R}^2$; we denote by $\Gamma$ the boundary of $\Omega$. The two-dimensional E-MAD problem reads as follows:

$$\det D^2\psi = f \text{ in } \Omega,\ \psi = g \text{ on } \Gamma, \qquad \text{(E-MAD)}$$

where, in (E-MAD), $D^2\psi$ is the Hessian of $\psi$, i.e., $D^2\psi = (\frac{\partial^2\psi}{\partial x_i \partial x_j})_{1\leq i,j\leq 2}$ and where $f$ and $g$ are two given functions, with $f > 0$. Unlike the closely related (see Remark 2.1) Dirichlet problem for the Laplace operator, (E-MAD) may have multiple solutions (actually, two at most; cf., e.g., [15, Chapter 4]), and the smoothness of the data does not imply the existence of a smooth solution. Concerning the last property, suppose that $\Omega = (0,1)\times(0,1)$ and consider the special case where (E-MAD) is defined by

$$\frac{\partial^2\psi}{\partial x_1^2}\frac{\partial^2\psi}{\partial x_2^2} - \left|\frac{\partial^2\psi}{\partial x_1 \partial x_2}\right|^2 = 1 \text{ in } \Omega,\ \psi = 0 \text{ on } \Gamma. \tag{2.1}$$

Problem (2.1) can not have smooth solutions since, for those solutions, the boundary condition $\psi = 0$ on $\Gamma$ implies that the product $\frac{\partial^2\psi}{\partial x_1^2}\frac{\partial^2\psi}{\partial x_2^2}$ and the cross-derivative $\frac{\partial^2\psi}{\partial x_1\partial x_2}$ vanish at the boundary, implying in turn that $\det D^2\psi$ is strictly less than one in some neighborhood of $\Gamma$.

The above (non-existence) result is not a consequence of the non-smoothness of $\Gamma$, since a similar non-existence property holds if in (2.1) one replaces the above $\Omega$ by the ovoïd-shaped domain whose $C^\infty$-boundary is defined by $\Gamma = \bigcup_{i=1}^4 \Gamma_i$, with $\Gamma_1 = \{x\,|\,x = \{x_1,x_2\},\ x_2 = 0,\ 0 \leq x_1 \leq 1\}$, $\Gamma_3 = \{x\,|\,x = \{x_1,x_2\},\ x_2 = 1,\ 0 \leq x_1 \leq 1\}$, $\Gamma_2 = \{x\,|\,x = \{x_1,x_2\},\ x_1 = 1 + \ln 4/\ln x_2(1-x_2),\ 0 \leq x_2 \leq 1\}$, $\Gamma_4 = \{x\,|\,x = \{x_1,x_2\},\ x_1 = -\ln 4/\ln x_2(1-x_2),\ 0 \leq x_2 \leq 1\}$. Actually, for the above two $\Omega$s the non-existence of solutions for problem (2.1) follows from the non-strict convexity of these domains.

**Remark 2.1.** We claimed, just above, that (E-MAD) and the Poisson-Dirichlet problem are closely related. To justify this statement, let us denote by $\lambda^+$ and $\lambda^-$ the two (necessary real) eigenvalues of matrix $D^2\psi$, with $\lambda^+ \geq \lambda^-$; (E-MAD) and the Poisson-Dirichlet problem

$$-\triangle\psi = f \text{ in } \Omega,\ \psi = g \text{ on } \Gamma, \tag{2.2}$$

read also as follows:

$$\lambda^+ \times \lambda^- = f \text{ in } \Omega,\ \psi = g \text{ on } \Gamma, \tag{2.3}$$

and

$$-(\lambda^+ + \lambda^-) = f \text{ in } \Omega,\ \psi = g \text{ on } \Gamma, \tag{2.4}$$

respectively. Equation (2.3) (resp., (2.4) ) shows the link between the Monge-Ampère (resp., Poisson-Dirichlet) problem and the geometric mean (resp., arithmetic mean) of the eigenvalues of the Hessian matrix $D^2\psi$.

It is worth noticing that, using appropriate iterative methods and mixed finite element approximations, we will be able to reduce the solution of (E-MAD) to the solution of sequences of discrete variants of Poisson-Dirichlet problems such as (2.2), (2.4).

**Remark 2.2.** Consider $x_0$ belonging to $\mathbf{R}^2$ and denote $|x - x_0|$ by $\rho$. Suppose that $u$ is a function of $\rho$ only; we have then

$$\det D^2 u = \frac{u' u''}{\rho} \tag{2.5}$$

where $u' = \frac{du}{d\rho}$, $u'' = \frac{d^2u}{d\rho^2}$. It follows from equation (2.5) that the function $u$ defined by

$$u = 2^{3/2}(m+4)^{-1}(m+2)^{-1/2}\rho^{\frac{m}{2}+2} + p, \tag{2.6}$$

where, in (2.6), $p$ is an arbitrary polynomial of degree $\leq 1$, is a solution of the Monge-Ampère equation

$$\det D^2 u = \rho^m.$$

Similarly, one can easily show that $u = e^{\rho^2/2} + p$, with $p$ as above, is a solution of the Monge-Ampère equation $\det D^2 u = (1+\rho^2)e^{\rho^2}$. We shall take advantage of these exact solutions to validate the (E-MAD) solvers to be described in the following sections.

**Remark 2.3.** Suppose that $\Omega$ is simply connected; let us define a vector-valued function $\mathbf{u}$ by $\mathbf{u} = \{\frac{\partial\psi}{\partial x_2}, -\frac{\partial\psi}{\partial x_1}\}\ (=\ \{u_1, u_2\})$; problem E-MAD takes then the equivalent formulation

$$\begin{cases} \det \boldsymbol{\nabla}\mathbf{u} = f, \text{ in } \Omega,\ \boldsymbol{\nabla}\cdot\mathbf{u} = 0 \text{ in } \Omega, \\ \mathbf{u}\cdot\mathbf{n} = \frac{dg}{ds} \text{ on } \Gamma, \end{cases} \tag{2.7}$$

where, in (2.7), $\mathbf{n}$ denotes the outward unit vector normal at $\Gamma$, and $s$ is a counter-clockwise curvilinear abscissa. Once $\mathbf{u}$ is known, one obtains $\psi$ via the solution of the following Poisson-Dirichlet problem

$$-\triangle\psi = \frac{\partial u_2}{\partial x_1} - \frac{\partial u_1}{\partial x_2} \text{ in } \Omega,\ \psi = g \text{ on } \Gamma. \tag{2.8}$$

System (2.7), (2.8) has clearly an incompressible fluid flow flavor, $\psi$ playing here the role of a stream function.

Relations (2.7), (2.8) can be used to solve problem E-MAD but this approach will not be further investigated here.

**3. A saddle-point formulation of problem E-MAD .** The simplest Hilbert space where to solve problem E-MAD is clearly $H^2(\Omega)$. This leads us to introduce

$$V_g = \{\varphi | \varphi \in H^2(\Omega),\, \varphi = g \text{ on } \Gamma\}; \tag{3.1}$$

if $g \in H^{3/2}(\Gamma)$, the (affine) space $V_g$ is non-empty. Assuming that (E-MAD) has solutions in $V_g$, it makes sense to consider the following problem from Calculus of Variations:

$$\begin{cases} \psi \in E_{fg}, \\ J(\psi) \leq J(\varphi),\, \forall\, \varphi \in E_{fg}, \end{cases} \tag{3.2}$$

where, in (3.2),

$$J(\varphi) = \frac{1}{2}\int_\Omega |\triangle\varphi|^2\, dx$$

and

$$E_{fg} = \{\varphi|\varphi \in V_g, \det D^2\varphi = f\}.$$

Replacing $|\triangle\varphi|^2$ by $|D^2\varphi|^2$ in $J(\cdot)$ would work as well (above, $|D^2\varphi|$ is the Frobenius norm of $D^2\varphi$, i.e., $|D^2\varphi| = (\sum_{1\leq i,j\leq 2} |\frac{\partial^2\varphi}{\partial x_i \partial x_j}|^2)^{1/2}$). Motivated by previous work on nonlinear bi-harmonic problems (see, e.g., refs. [6]-[12]), we introduce the symmetric tensor-valued functions $\mathbf{p} = D^2\psi, \mathbf{q} = D^2\varphi$ and the related minimization problem (equivalent to (3.2)):

$$\begin{cases} \{\psi, \mathbf{p}\} \in \mathcal{E}_{fg}, \\ j(\psi, \mathbf{p}) \leq j(\varphi, \mathbf{q}),\, \forall \{\varphi, \mathbf{q}\} \in \mathcal{E}_{fg}, \end{cases} \tag{3.3}$$

where, in (3.3),

$$j(\varphi, \mathbf{q}) = \frac{1}{2}\int_\Omega |\triangle\varphi|^2\, dx$$

and

$$\mathcal{E}_{fg} = \{\{\varphi, \mathbf{q}\}|\, \varphi \in V_g,\, \mathbf{q} \in \mathbf{Q},\, \mathbf{q} = D^2\varphi,\, \det \mathbf{q} = f\},$$

with

$$\mathbf{Q} = \{\mathbf{q}|\, \mathbf{q} = (q_{ij})_{1\leq i,j\leq 2},\, q_{21} = q_{12},\, q_{ij} \in L^2(\Omega)\}. \tag{3.4}$$

Let $r$ be a positive parameter; we associate to problem (3.3) the following saddle-point problem

$$\begin{cases} \{\{\psi, \mathbf{p}\}, \boldsymbol{\lambda}\} \in (V_g \times \mathbf{Q}_f) \times \mathbf{Q}, \\ \mathcal{L}_r(\psi, \mathbf{p}; \boldsymbol{\mu}) \leq \mathcal{L}_r(\psi, \mathbf{p}; \boldsymbol{\lambda}) \leq \mathcal{L}_r(\varphi, \mathbf{q}; \boldsymbol{\lambda}), \\ \quad \forall \{\{\varphi, \mathbf{q}, \}\boldsymbol{\mu}\} \in (V_g \times \mathbf{Q}_f) \times \mathbf{Q}, \end{cases} \tag{3.5}$$

where

$$\mathbf{Q}_f = \{\mathbf{q}|\, \mathbf{q} \in \mathbf{Q},\, \det \mathbf{q} = f\}$$

ETNA
Kent State University
etna@mcs.kent.edu



and

$$\mathcal{L}_r(\varphi, \mathbf{q}; \boldsymbol{\mu}) = \frac{1}{2}\int_\Omega |\triangle\varphi|^2\,dx + \frac{r}{2}\int_\Omega |D^2\varphi - \mathbf{q}|^2\,dx + \int_\Omega \boldsymbol{\mu} : (D^2\varphi - \mathbf{q})\,dx,$$

with $\mathbf{S} : \mathbf{T} = \sum_{1\leq i,j\leq 2} s_{ij}t_{ij}$ if $\mathbf{S} = (s_{ij})$ and $\mathbf{T} = (t_{ij})$; if $f \in L^1(\Omega)$, then $\mathbf{Q}_f \neq \emptyset$. Suppose that problem (3.5) has a solution $\{\{\psi, \mathbf{p}\}, \boldsymbol{\lambda}\}$, then $\{\psi, \mathbf{p}\}$ is also a solution of problem (3.3), in turn implying that $\psi$ solves problem E-MAD. The computation of the saddle-points of the augmented Lagrangian functional $\mathcal{L}_r$ will be addressed in Section 4.

**Remark 3.1.** When the authors of this article started investigating the numerical solution of the elliptic Monge-Ampère equation in dimension two, it never crossed their minds that after all this equation could be the Euler-Lagrange equation of some problem from Calculus of Variations. The fact that it is actually the case was pointed to them by B. Dacorogna during a visit of the second author at EPFL in January 2003. Indeed, if one denotes by $\operatorname{cof} D^2\varphi$ the matrix

$$\begin{pmatrix} \frac{\partial^2\varphi}{\partial x_2^2} & -\frac{\partial^2\varphi}{\partial x_1\partial x_2} \\ -\frac{\partial^2\varphi}{\partial x_1\partial x_2} & \frac{\partial^2\varphi}{\partial x_1^2} \end{pmatrix}$$

it is a relatively simple integration by parts exercise to show that any sufficiently smooth stationary point in $V_g$ of the functional $I(\cdot)$ defined by

$$I(\varphi) = \int_\Omega (\operatorname{cof} D^2\varphi)\boldsymbol{\nabla}\varphi\cdot\boldsymbol{\nabla}\varphi\,dx + 6\int_\Omega f\varphi\,dx$$

is a solution of the E-MAD problem; this result seems to belong to the "folklore" of partial differential equations and is reported in, e.g., [15, Chapter 4] ).

Suppose that $\psi$ is one of these stationary points; the condition $I'(\psi) = 0$ leads to the following weak formulation of (E-MAD):

$$\begin{cases} \psi \in V_g, \\ \int_\Omega (\operatorname{cof} D^2\psi)\boldsymbol{\nabla}\psi\cdot\boldsymbol{\nabla}\varphi\,dx + 2\int_\Omega f\varphi\,dx = 0, \forall\varphi \in \mathcal{D}(\Omega), \end{cases} \tag{3.6}$$

where $\mathcal{D}(\Omega) = \{\varphi | \, \varphi \in C^\infty(\bar{\Omega}), \varphi \text{ has a compact support in } \Omega\}$.

It follows from the above results that (E-MAD) can be formulated as a nonlinear variational problem closely related to those problems whose numerical solution has been addressed in, e.g., [30], [31]. Actually, there is more since relation

$$< I''(\psi)\varphi, \theta >= 6\int_\Omega (\operatorname{cof} D^2\psi)\boldsymbol{\nabla}\varphi\cdot\boldsymbol{\nabla}\theta\,dx, \forall\,\varphi,\theta \in \mathcal{D}(\Omega),$$

implies that functional $I(\cdot)$ is either convex or concave in the neighborhood of a solution of (E-MAD), if such a solution does exist. This last result is quite important from a methodological point of view since it links (E-MAD) to Convex Analysis, a well investigated area from both analytical and computational points of view. This suggests, in particular, that Newton's and conjugate gradient methods may be well-suited to the solution of (E-MAD) through formulation (3.6). Among the reasons we did not give priority to the above approach let us mention:

(i) Formulation (3.6) combines the difficulties of both harmonic and bi-harmonic problems, making the approximation a delicate matter, albeit solvable.

(ii) If (E-MAD) has no solution we can expect the divergence of the Newton and conjugate gradient algorithms mentioned above. On the other hand, the method discussed in Section 4, solving (E-MAD) in a least-squares sense, will find automatically a solution if such a solution exists in $V_g$ and a best possible solution if (E-MAD) has no solution albeit neither $V_g$ nor $\mathbf{Q}_f$ are empty.

(iii) From a partial differential equation point of view the method discussed in Section 4 is easy to implement since it requires no more than access to a Poisson-Dirichlet solver, an interesting achievement considering the seeming complexity of problem E-MAD. This method applies also to the solution of those variants of (E-MAD) which are not Euler-Lagrange equations (or which have not been identified as such).

Having said that, we intend to investigate the solution of (E-MAD), via formulation (3.7), in a near future.

**Remark 3.2.** It follows from the above remark that a natural Neumann boundary condition for the Monge-Ampère equation is given by

$$(\operatorname{cof} D^2\psi)\boldsymbol{\nabla}\psi\cdot\mathbf{n} = g \text{ on } \Gamma, \tag{3.7}$$

with, in (3.7), $\mathbf{n}$ the outward unit vector normal at $\Gamma$.

**Remark 3.3.** Neither $E_{fg}$ nor $\mathcal{E}_{fg}$ are convex; moreover, there is no simple way to find elements belonging to these sets (this would imply that we have in turn a simple way to solve problem E-MAD). The key idea in this article is to decouple $\mathbf{p}$ and $D^2\psi$, via an augmented Lagrangian formulation, in order to approximate $\{\psi, \mathbf{p}\}$ by a sequence $\{\psi^n, \mathbf{p}^n\}_n$ whose elements are external to $\mathcal{E}_{fg}$, avoiding thus the above difficulty. The above observation applies also to the finite element analogues (to be defined in Section 5) of the sets $E_{fg}$ and $\mathcal{E}_{fg}$.

**Remark 3.4.** The correct notion of weak solutions for Monge-Ampère equations (and related fully nonlinear elliptic equations) is provided by the viscosity solutions, as shown in e.g., [40], [41], [42], [13], [14] (see also the references therein).

## 4. Iterative solution of the saddle-point problem (3.5).

**4.1. Description of the algorithm.** The iterative solution of saddle-point problems such as (3.5) has been addressed in, e.g., [10], [11], [30]. In order to solve problem (3.5) we advocate (among other algorithms and for its simplicity) the algorithm called ALG2 in the above references. This algorithm is in fact an Uzawa-Douglas-Rachford algorithm and it reads as follows when applied to the solution of problem (3.5):

$$\{\psi^{-1}, \boldsymbol{\lambda}^0\} \text{ is given in } V_g \times \mathbf{Q}; \tag{4.1}$$

*then, for* $n \geq 0$, $\{\psi^{n-1}, \boldsymbol{\lambda}^n\}$ *being known in* $V_g \times \mathbf{Q}$, *solve*

$$\begin{cases} \mathbf{p}^n \in \mathbf{Q}_f, \\ \mathcal{L}_r(\psi^{n-1}, \mathbf{p}^n; \boldsymbol{\lambda}^n) \leq \mathcal{L}_r(\psi^{n-1}, \mathbf{q}; \boldsymbol{\lambda}^n),\ \forall\, \mathbf{q} \in \mathbf{Q}_f, \end{cases} \tag{4.2}$$

$$\begin{cases} \psi^n \in V_g, \\ \mathcal{L}_r(\psi^n, \mathbf{p}^n; \boldsymbol{\lambda}^n) \leq \mathcal{L}_r(\varphi, \mathbf{p}^n; \boldsymbol{\lambda}^n),\ \forall\, \varphi \in V_g, \end{cases} \tag{4.3}$$

and update $\boldsymbol{\lambda}^n$ by

$$\boldsymbol{\lambda}^{n+1} = \boldsymbol{\lambda}^n + r(D^2\psi^n - \mathbf{p}^n). \tag{4.4}$$

**Remark 4.1.** Concerning the initialization of algorithm (4.1)-(4.4) we advocate $\boldsymbol{\lambda}^0 = \mathbf{0}$ and $\psi^{-1}$ defined as the solution of the Poisson-Dirichlet problem

$$-\triangle\psi^{-1} = f^{1/2} \text{ in } \Omega,\ \psi^{-1} = g \text{ on } \Gamma. \tag{4.5}$$

It will make sense (see Remark 2.1) to replace the right hand side of the Laplace equation in (4.5) by $2f^{1/2}$; however, this substitution does not improve significantly the convergence of algorithm (4.1)-(4.4).

**4.2. On the convergence of algorithm (4.1)-(4.4).** Suppose that, in (4.2), we replace $\mathbf{Q}_f$ by $\mathbf{K}$, a closed convex non-empty subset of $\mathbf{Q}$. Suppose also that the functional $\mathcal{L}_r(\cdot,\cdot;\cdot)$ has a saddle-point in $(V_g \times \mathbf{K}) \times \mathbf{Q}$ and that instead of (4.4) we use

$$\boldsymbol{\lambda}^{n+1} = \boldsymbol{\lambda}^n + \rho(D^2\psi^n - \mathbf{p}^n)$$

to update the multiplier $\boldsymbol{\lambda}^n$. Assume that these assumptions hold and that

$$0 < \rho < r(1+\sqrt{5})/2;$$

then, it has been proved in, e.g., [10], [11] and [30] that the sequence $\{\psi^n, \mathbf{p}^n; \boldsymbol{\lambda}^n\}_n$ converges to a saddle-point of $\mathcal{L}_r$ in $(V_g \times \mathbf{K}) \times \mathbf{Q}$ (a classical choice for $\rho$ being $\rho = r$). The non-convexity of $\mathbf{Q}_f$ makes things more complicated for problem (3.5); however, on the basis of past experiences where algorithms such as (4.1)-(4.4) have been applied to the solution of nonlinear problems in Finite Elasticity (see refs. [10], [11]), we expect the convergence of algorithm (4.1)-(4.4), assuming that $r$ is sufficiently large. The numerical results reported in Section 6 justify this prediction.

**Remark 4.2.** If the continuous E-MAD problem has no smooth solution (like it is the case for problem (2.1)), we expect the discrete analogues of set $\mathcal{E}_{fg}$ to be empty (or "pathological") and, consequently, the divergence of algorithm (4.1)-(4.4). Strictly speaking this is what happens, however we shall see in Section 6, that even in these "desperate" situations, algorithm (4.1)-(4.4) behaves constructively, in the sense that albeit sequence $\{\boldsymbol{\lambda}_h^n\}_n$ diverges (arithmetically), sequence $\{\{\psi_h^n, \mathbf{p}_h^n\}\}_n$ converges (geometrically) to a limit which does not vary much with $h$ as soon as $h$ is small enough (all this supposes that neither $V_g$ nor $\mathbf{Q}_f$ are empty ). We suspect that what we have captured here is a pair $\{\psi_h, \mathbf{p}_h\}$ which, at the limit when $h \to 0$, provides a pair $\{\psi, \mathbf{p}\}$ which minimizes (locally or globally) the functional $\{\varphi, \mathbf{q}\} \to \|D^2\varphi - \mathbf{q}\|_{L^2}$ over the (closed) set $V_g \times \mathbf{Q}_f$. Assuming that the above convergence result is true (when set $\mathcal{E}_{fg}$ is empty), then algorithm (4.1)-(4.4) has solved problem E-MAD in a least squares sense. To justify our guess concerning the behavior of algorithm (4.1)-(4.4) let us consider the following finite dimensional problem (a Kuhn-Tucker's system):

$$\begin{cases} \mathbf{Ax} + \mathbf{B}^{\mathbf{t}}\boldsymbol{\lambda} = \mathbf{b} \\ \mathbf{Bx} = \mathbf{c}, \end{cases} \tag{4.6}$$

where:

- $\mathbf{A}$ and $\mathbf{B}$ are $N \times N$ and $M \times N$ real matrices, respectively, $\mathbf{A}$ being symmetric and positive definite.
- $\mathbf{b} \in \mathbf{R}^N, \mathbf{c} \in \mathbf{R}^M$.

Suppose that $\mathbf{c} \in \mathbf{R}(\mathbf{B})$ (the range of matrix $\mathbf{B}$), then problem (4.6) has a solution. If $\mathbf{B}$ is onto (i.e., $\mathbf{B}$ is a rank $M$ matrix) the above solution is unique. If rank of $\mathbf{B}$ is less than $M$, system (4.6) has an infinity of solutions sharing all the same $\mathbf{x}$. This is not surprising since system (4.6) characterizes $\mathbf{x}$ as the solution (necessarily unique from the properties of $\mathbf{A}$) of the following minimization problem:

$$\begin{cases} \mathbf{x} \in H, \\ J(\mathbf{x}) \leq J(\mathbf{y}), \, \forall \, \mathbf{y} \in H, \end{cases} \tag{4.7}$$

with functional $J(\cdot)$ and set $H$ defined (with obvious notation) by

$$J(\mathbf{y}) = \frac{1}{2}\mathbf{A}\mathbf{y}\cdot\mathbf{y} - \mathbf{b}\cdot\mathbf{y} \tag{4.8}$$

and

$$H = \{\mathbf{y}|\mathbf{y} \in \mathbf{R}^N, \mathbf{B}\mathbf{y} = \mathbf{c}\}.$$

An "old fashioned" iterative method for the solution of problem (4.6) is provided by the following algorithm (of the Uzawa's type ):

$$\boldsymbol{\lambda}^0 \textit{is given in } \mathbf{R}^N; \tag{4.9}$$

*for* $n \geq 0$, $\boldsymbol{\lambda}^n$*being known, compute* $\mathbf{x}^n$*and* $\boldsymbol{\lambda}^{n+1}$ *as follows*

$$\mathbf{A}\mathbf{x}^n = \mathbf{b} - \mathbf{B}^t\boldsymbol{\lambda}^n, \tag{4.10}$$

$$\boldsymbol{\lambda}^{n+1} = \boldsymbol{\lambda}^n + \rho(\mathbf{B}\mathbf{x}^n - \mathbf{c}). \tag{4.11}$$

It is well-known (see, e.g., refs. [10], [11]) that, if $\mathbf{c} \in \mathbf{R}(\mathbf{B})$ and if

$$0 < \rho < 2/\mu_N, \tag{4.12}$$

with $\mu_N$ the largest eigenvalue of matrix $\mathbf{A}^{-1}\mathbf{B}^t\mathbf{B}$, then

$$\lim_{n\to\infty} \{\mathbf{x}^n, \boldsymbol{\lambda}^n\} = \{\mathbf{x}, \boldsymbol{\lambda}\}$$

geometrically, the pair $\{\mathbf{x}, \boldsymbol{\lambda}\}$ being the unique solution of problem (4.6) such that, $\boldsymbol{\lambda} - \boldsymbol{\lambda}^0 \in \mathbf{R}(\mathbf{B})$. In order to speed up the convergence of algorithm (4.9)-(4.11), and avoid the computation of $\mu_N$, we advocate using conjugate gradient variants of the above algorithm, like those discussed in, e.g., refs. [10], [11]. Let us consider now the - in some sense - more interesting case where $\mathbf{c} \notin \mathbf{R}(\mathbf{B})$; it is clear that in that case system (4.6) has no solution. Suppose, however, that for the sake of curiosity (or because one has not realized that $\mathbf{c} \notin \mathbf{R}(\mathbf{B})$) one applies algorithm (4.9)-(4.11) to the solution of (the now ill-posed) system (4.6). A kind of miracle takes place since

THEOREM 4.1. *Suppose that* $\mathbf{c} \notin \mathbf{R}(\mathbf{B})$ *but that condition (4.12) still holds. Then, for any* $\boldsymbol{\lambda}^0 \in \mathbf{R}^M$ *we have:*

$$\text{the arithmetic divergence of sequence } \{\boldsymbol{\lambda}^n\}_n, \tag{4.13}$$

*but*

$$\lim_{n\to+\infty} \mathbf{x}^n = \mathbf{x} \text{ geometrically}, \tag{4.14}$$

*where* $\mathbf{x}$ *is the unique solution of the minimization problem*

$$\begin{cases} \mathbf{x} \in \widetilde{H}, \\ J(\mathbf{x}) \le J(\mathbf{y}), \forall\, \mathbf{y} \in \widetilde{H}, \end{cases} \tag{4.15}$$

*with* $J(\cdot)$ *still defined by (4.8) and*

$$\widetilde{H} = \{\mathbf{y} |\, \mathbf{y} \in \mathbf{R}^N, \mathbf{B}^t(\mathbf{B}\mathbf{y} - \mathbf{c}) = \mathbf{0}\}. \tag{4.16}$$

*Proof*. Since the affine subspace $\widetilde{H}$ is a non-empty closed convex subspace of $\mathbf{R}^N$, it follows from the properties of matrix $\mathbf{A}$ that problem (4.15) has a unique solution $\mathbf{x}$ characterized by the existence of $\mathbf{\Lambda} \in \mathbf{R}^N$ such that

$$\begin{cases} \mathbf{A}\mathbf{x} + \mathbf{B}^t\mathbf{B}\mathbf{\Lambda} = \mathbf{b}, \\ \mathbf{B}^t\mathbf{B}\mathbf{x} = \mathbf{B}^t\mathbf{c}. \end{cases} \tag{4.17}$$

Denote $\mathbf{B}\mathbf{\Lambda}$ by $\boldsymbol{\lambda}$; we have then

$$\begin{cases} \mathbf{A}\mathbf{x} + \mathbf{B}^t\boldsymbol{\lambda} = \mathbf{b}, \\ \mathbf{B}^t\mathbf{B}\mathbf{x} = \mathbf{B}^t\mathbf{c}. \end{cases} \tag{4.18}$$

Denote $\mathbf{x}^n - \mathbf{x}$ and $\boldsymbol{\lambda}^n - \boldsymbol{\lambda}$ by $\bar{\mathbf{x}}^n$ and $\bar{\boldsymbol{\lambda}}^n$, respectively. Comparing relations (4.18) to (4.10) and (4.11), we obtain

$$\mathbf{A}\bar{\mathbf{x}}^n = -\mathbf{B}^t\bar{\boldsymbol{\lambda}}^n \tag{4.19}$$

and

$$\mathbf{B}^t\bar{\boldsymbol{\lambda}}^{n+1} = \mathbf{B}^t\bar{\boldsymbol{\lambda}}^n + \rho\mathbf{B}^{\mathbf{t}}\mathbf{B}\bar{\mathbf{x}}^n. \tag{4.20}$$

Combining the above relations yields:

$$\bar{\mathbf{x}}^n \in \mathbf{R}(\mathbf{A}^{-1}\mathbf{B}^t\mathbf{B}), \forall\, n \ge 0, \tag{4.21}$$

$$\bar{\mathbf{x}}^{n+1} = (\mathbf{I} - \rho\mathbf{A}^{-1}\mathbf{B}^t\mathbf{B})\bar{\mathbf{x}}^n. \tag{4.22}$$

Let us denote by $\mu_i$ the non-zero eigenvalues of matrix $\mathbf{A}^{-1}\mathbf{B}^t\mathbf{B}$ and by $w_i$ the corresponding eigenvectors; we clearly have $\mu_i > 0$. We order these eigenvalues so that $\mu_i \le \mu_{i+1}$, with $i = N - rank(\mathbf{B}) + 1, \ldots, N$. Taking into account relation (4.21) and the fact that the above eigenvectors $w_i$ form a vector basis of subspace $\mathbf{R}(\mathbf{A}^{-1}\mathbf{B}^t\mathbf{B})$, one can easily show (by expanding relation (4.22) on the above vector basis) that $\bar{\mathbf{x}}^n \to 0,\ \forall\, \boldsymbol{\lambda}^0$, when $n \to +\infty$, *if and only if*

$$|1 - \rho\mu_i| < 1, \forall\, i = N - rank(\mathbf{B}) + 1, \ldots, N, \tag{4.23}$$

or, equivalently,

$$0 < \rho < 2/\mu_N. \tag{4.24}$$

Since $\bar{\mathbf{x}}^n = \mathbf{x}^n - \mathbf{x}$ we have definitely proved that the conditions (4.12), (4.24) implies the convergence result (4.14). Actually, using relations (4.23), we can easily show that modulo a multiplicative constant, $\|\mathbf{x}^n - \mathbf{x}\|$ converges to 0 at least as fast as $\max(|1-\rho\mu_{I_{\mathbf{B}}}|^n, |1-\rho\mu_N|^n)$ (with $I_{\mathbf{B}} = N - rank(\mathbf{B}) + 1$ ). From relations (4.23) we can also show that the optimal value of $\rho$ (i.e., the one leading, generically, to the fastest convergence) is $2/(\mu_N + \mu_{I_{\mathbf{B}}})$; for this value of $\rho$, $\|\mathbf{x}^n - \mathbf{x}\|$ converges to 0 at least as fast as $\left[\frac{\mu_N - \mu_{I_{\mathbf{B}}}}{\mu_N + \mu_{I_{\mathbf{B}}}}\right]^n$. Proving the divergence result (4.13) is trivial. $\square$

The above theorem deserves further comments:

**(i)** It follows from Theorem 4.1 that, if $\mathbf{c} \notin \mathbf{R}(\mathbf{B})$, algorithm (4.9)-(4.11) will, in some sense, substitute automatically the normal equation $\mathbf{B}^t\mathbf{B}\mathbf{x} = \mathbf{B}^t\mathbf{c}$ (which has always a solution) to the solution-less equation $\mathbf{B}\mathbf{x} = \mathbf{c}$, a most remarkable property indeed. We can also say that if problem (4.6), (4.7) has no solution, due to $\mathbf{c} \notin \mathbf{R}(\mathbf{B})$, algorithm (4.9)-(4.11) solves it in a least squares sense.

**(ii)** From a computational point of view there are no practical difficulties associated to the divergence of sequence $\{\boldsymbol{\lambda}^n\}_n$ when $\mathbf{c} \notin \mathbf{R}(\mathbf{B})$. Indeed, the divergence being arithmetic is pretty slow, implying that in practice, sequence $\{\mathbf{x}^n\}_n$ being geometrically convergent will reach its (practical) limit long before $\|\boldsymbol{\lambda}^n\|$ becomes dangerously large.

**(iii)** If rank$(\mathbf{B}) < M$, the property $\mathbf{c} \in \mathbf{R}(\mathbf{B})$ may be lost due to round-off errors. Theorem 4.1 guarantees that an approximation to the solution of problem (4.6), (4.7) will be found by algorithm (4.9)-(4.11).

**(iv)** We do not know of generalizations of the above results to nonlinear and/or non-convex situations, like the one encountered, for example, with problem (3.5). On the basis of the numerical results shown in Section 6 it seems that the Uzawa-Douglas-Rachford algorithm (4.1)-(4.4) behaves essentially like algorithm (4.9)-(4.11), albeit being much more complicated.

**4.3. Solution of sub-problems (4.2).** Problem (4.2) can be solved point-wise; taking the symmetry of tensors $\mathbf{p}^n$ and $\mathbf{q}$ into account, to obtain $\mathbf{p}^n$ from $\psi^{n-1}$ and $\boldsymbol{\lambda}^n$ we have to minimize, point-wise over $\Omega$, a three-variable polynomial of the following type:

$$\mathbf{z} = \{z_i\}_{i=1}^3 \to \frac{r}{2}(z_1^2 + z_2^2 + 2z_3^2) - \mathbf{b}_n(x)\cdot\mathbf{z} \tag{4.25}$$

over the set defined by

$$z_1 z_2 - z_3^2 = f(x). \tag{4.26}$$

The above minimization problem is a generalized eigenvalue problem which can be solved by Newton's method. To derive the above generalized eigenvalue problem we associate to relations (4.25) and (4.26) the following Lagrangian functional

$$L(\mathbf{z},\mu) = \frac{r}{2}(z_1^2 + z_2^2 + 2z_3^2) - \mathbf{b}_n(x)\cdot\mathbf{z} - \mu(z_1 z_2 - z_3^2 - f(x)). \tag{4.27}$$

**ETNA**
Kent State University
etna@mcs.kent.edu



Suppose that $\{\mathbf{p}, \lambda\}$ is a stationary point of $L(\cdot, \cdot)$ over $\mathbf{R}^4$. We have then, with obvious notation,

$$\begin{cases} rp_1 = \lambda p_2 + b_{1n}(x) \\ rp_2 = \lambda p_1 + b_{2n}(x) \\ 2rp_3 = -2\lambda p_3 + b_{3n}(x) \\ p_1 p_2 - p_3^2 = f(x). \end{cases} \tag{4.28}$$

The generalized eigenvalue $\lambda$ is clearly a Lagrange multiplier associated to relation (4.26). Solving the nonlinear system (4.28) is routine (cf. [29] and the references therein).

**Remark 4.3.** In practice (see Section 5 for details) we shall have to solve systems like (4.28) "only" at the vertices or grid points of a finite element or finite difference mesh.

**4.4. Solution of sub-problems (4.3).** Problem (4.3) is equivalent to a well-posed linear variational problem which reads as follows (with $V_0 = H^2(\Omega) \cap H_0^1(\Omega)$):

$$\begin{cases} \psi^n \in V_g, \\ \int_\Omega \triangle\psi^n \triangle\varphi \, dx + r \int_\Omega D^2\psi^n : D^2\varphi \, dx = L_n(\varphi), \, \forall \, \varphi \in V_0, \end{cases} \tag{4.29}$$

with functional $L_n(\cdot)$ linear and continuous over $V_0$. Problem (4.29) can be solved by a conjugate gradient algorithm operating in $V_g$ and $V_0$ equipped with the scalar product $\{v, w\} \to \int_\Omega \triangle v \triangle w \, dx$ which, if $\Gamma$ is smooth and/or $\Omega$ convex, defines on $V_0$ a norm equivalent to the $H^2$ one (the solution of linear variational problems in Hilbert spaces-such as (4.29) - by conjugate gradient algorithms is discussed in, e.g., [28, Chapter 3]; see also the references therein). Applying the results in the above reference leads to the following algorithm:

$$\psi^{n,0} \textit{ is given in } V_g \textit{ (a natural choice being } \psi^{n,0} = \psi^{n-1}\textit{);} \tag{4.30}$$

*solve the following bi-harmonic problem*

$$\begin{cases} g^{n,0} \in V_0, \\ \int_\Omega \triangle g^{n,0} \triangle\varphi \, dx = \int_\Omega \triangle\psi^{n,0} \triangle\varphi \, dx + r \int_\Omega D^2\psi^{n,0} : D^2\varphi \, dx \\ \quad -L_n(\varphi), \, \forall \, \varphi \in V_0, \end{cases} \tag{4.31}$$

*and set*

$$w^{n,0} = g^{n,0}. \tag{4.32}$$

*For* $k \geq 0$ *assuming that* $\psi^{n,k}, g^{n,k}$, *and* $w^{n,k}$ *are known with the last two different from* $0$, *compute* $\psi^{n,k+1}, g^{n,k+1}$ *and, if necessary,* $w^{n,k+1}$ *as follows:*

*Solve the bi-harmonic problem*

$$\begin{cases} \bar{g}^{n,k} \in V_0, \\ \int_\Omega \triangle\bar{g}^{n,k} \triangle\varphi \, dx = \int_\Omega \triangle w^{n,k} \triangle\varphi \, dx + r \int_\Omega D^2 w^{n,k} : D^2\varphi \, dx, \\ \quad \forall \, \varphi \in V_0, \end{cases} \tag{4.33}$$

*then, compute*

$$\rho_{n,k} = \frac{\int_\Omega |\triangle g^{n,k}|^2 \, dx}{\int_\Omega \triangle\bar{g}^{n,k} \triangle w^{n,k} \, dx}, \tag{4.34}$$

and set

$$\psi^{n,k+1} = \psi^{n,k} - \rho_{n,k} w^{n,k}, \tag{4.35}$$

$$g^{n,k+1} = g^{n,k} - \rho_{n,k} \bar{g}^{n,k}. \tag{4.36}$$

*If* $\frac{\int_\Omega |\triangle g^{n,k+1}|^2\,dx}{\int_\Omega |\triangle g^{n,0}|^2\,dx} \leq \epsilon$ *take* $\psi^{n+1} = \psi^{n,k+1}$; *else compute*

$$\gamma_{n,k} = \frac{\int_\Omega |\triangle g^{n,k+1}|^2\,dx}{\int_\Omega |\triangle g^{n,k}|^2\,dx}, \tag{4.37}$$

$$w^{n,k+1} = g^{n,k+1} + \gamma_{n,k} w^{n,k}. \tag{4.38}$$

*Do* $k = k + 1$ *and return to* (4.33). □

From a practical point of view algorithm (4.30)-(4.38) is not particularly difficult to implement; indeed, after an appropriate space discretization, each iteration will require the solution of two discrete Poisson- Dirichlet problems in order to solve the discrete analogues of the bi-harmonic problems (4.31) and (4.33) (see Section 5 for details).

**4.5. Further remark.** One of the main reasons explaining the rising popularity of Monge-Ampère equation related problems is without any doubt the links existing with the problem of Monge-Kantorovich (see, e.g., refs. [33],[34] for details and references). Actually, one may find in [34] a most interesting discussion of a method for the numerical solution of a particular Monge-Kantorovich problem, via a Fluid Dynamics interpretation. The solution method relies on an Uzawa-Douglas-Rachford algorithm (a particular case of ALG2, a general one, discussed in, e.g., [10], [11], [30]), quite different of the one used in the present article, the corresponding Lagrangian functionals being themselves quite different. However, it is interesting to observe that the augmented Lagrangian algorithms discussed in [10], [11] and [30] have enough generality to be able to solve a large variety of difficult problems, including some which have attracted the attention of the Scientific Computing community only recently.

## 5. Finite element approximation of problem (E-MAD).

**5.1. Generalities.** Considering the highly variational flavor of the methodology discussed in Sections 3 and 4, it makes sense to look for finite element based methods for the approximation of (E-MAD). In order to avoid the complications associated to the construction of finite element subspaces of $H^2(\Omega)$, we will employ a mixed finite element approximation (closely related to those discussed in, e.g., [5]-[12], [31], [35], [36] for the solution of linear and nonlinear bi-harmonic problems). Following this approach, it will be possible to solve (E-MAD) employing approximations commonly used for the solution of second order elliptic problems (piecewise linear and globally continuous over a triangulation of $\Omega$, for example).

**5.2. A mixed finite element approximation of problem (E-MAD).** For simplicity, we suppose that $\Omega$ is a bounded polygonal domain of $\mathbf{R}^2$. Let us denote by $\mathcal{T}_h$ a finite element triangulation of $\Omega$ (like those discussed in, e.g., [35], [37], [30, Appendix 1]). From $\mathcal{T}_h$ we approximate spaces $L^2(\Omega)$, $H^1(\Omega)$ and $H^2(\Omega)$ by the finite dimensional space $V_h$ defined by

$$V_h = \{v |\, v \in C^0(\bar{\Omega}),\, v|_T \in P_1,\, \forall T \in \mathcal{T}_h\}, \tag{5.1}$$

with $P_1$ the space of the two-variable polynomials of degree $\leq 1$. A function $\varphi$ being given in $H^2(\Omega)$ we denote $\frac{\partial^2 \varphi}{\partial x_i \partial x_j}$ by $D^2_{ij}(\varphi)$. It follows from Green's formula that

$$\int_\Omega \frac{\partial^2 \varphi}{\partial x_i^2} v \, dx = -\int_\Omega \frac{\partial \varphi}{\partial x_i} \frac{\partial v}{\partial x_i} \, dx, \; \forall \, v \in H^1_0(\Omega), \; \forall \, i = 1, 2, \tag{5.2}$$

$$\int_\Omega \frac{\partial^2 \varphi}{\partial x_1 \partial x_2} v \, dx = -\frac{1}{2} \int_\Omega \left[ \frac{\partial \varphi}{\partial x_1} \frac{\partial v}{\partial x_2} + \frac{\partial \varphi}{\partial x_2} \frac{\partial v}{\partial x_1} \right] dx, \; \forall \, v \in H^1_0(\Omega). \tag{5.3}$$

Consider now $\varphi \in V_h$; taking advantage of relations (5.2) and (5.3) we define the discrete analogues of the differential operators $D^2_{ij}$ by:

$$\begin{cases} \forall \, i = 1, 2, \; D^2_{hii}(\varphi) \in V_{0h}, \\ \int_\Omega D^2_{hii}(\varphi) v \, dx = -\int_\Omega \frac{\partial \varphi}{\partial x_i} \frac{\partial v}{\partial x_i} \, dx, \; \forall \, v \in V_{0h}, \end{cases} \tag{5.4}$$

$$\begin{cases} D^2_{h12}(\varphi) \in V_{0h}, \\ \int_\Omega D^2_{h12}(\varphi) v \, dx = -\frac{1}{2} \int_\Omega \left[ \frac{\partial \varphi}{\partial x_1} \frac{\partial v}{\partial x_2} + \frac{\partial \varphi}{\partial x_2} \frac{\partial v}{\partial x_1} \right] dx, \; \forall \, v \in V_{0h}, \end{cases} \tag{5.5}$$

in (5.4) and (5.5) the space $V_{0h}$ is defined by

$$V_{0h} = V_h \cap H^1_0(\Omega) \; (= \{ v | \, v \in V_h, \, v = 0 \text{ on } \Gamma \}). \tag{5.6}$$

The functions $D^2_{hij}(\Omega)$ are uniquely defined by relations (5.4) and (5.5). However, in order to simplify the computation of the above discrete second order partial derivatives we will use the trapezoidal rule to evaluate the integrals in the left hand sides of (5.4) and (5.5). Owing to their practical importance, let us detail these calculations:

**(i)** First we introduce the set $\Sigma_h$ of the vertices of $\mathcal{T}_h$ and $\Sigma_{0h} = \{ P \, | \, P \in \Sigma_h, \, P \notin \Gamma \}$. Next, we define the integers $N_h$ and $N_{0h}$ by $N_h = \text{Card}(\Sigma_h)$ and $N_{0h} = \text{Card}(\Sigma_{0h})$. We have then $\dim V_h = N_h$ and $\dim V_{0h} = N_{0h}$. We suppose that $\Sigma_{0h} = \{P_k\}_{k=1}^{N_{0h}}$ and $\Sigma_h = \Sigma_{0h} \cup \{P_k\}_{k=N_{0h}+1}^{N_h}$.

**(ii)** To $P_k \in \Sigma_h$ we associate the function $w_k$ uniquely defined by

$$w_k \in V_h, \; w_k(P_k) = 1, \; w_k(P_l) = 0, \; \text{ if } l = 1, \cdots N_h, \, l \neq k. \tag{5.7}$$

It is well known (see, e.g., [35], [37], [30, Appendix 1]) that the sets $\mathcal{B}_h = \{w_k\}_{k=1}^{N_h}$ and $\mathcal{B}_{0h} = \{w_k\}_{k=1}^{N_{0h}}$ are vector bases of $V_h$ and $V_{0h}$, respectively.

**(iii)** Let us denote by $A_k$ the area of the polygonal domain which is the union of those triangles of $\mathcal{T}_h$ which have $P_k$ as a common vertex. Applying the trapezoidal rule to the integrals in the left hand side of relations (5.4) we obtain:

$$\begin{cases} \forall \, i = 1, 2, \; D^2_{hii}(\varphi) \in V_{0h}, \\ D^2_{hii}(\varphi)(P_k) = -\frac{3}{A_k} \int_\Omega \frac{\partial \varphi}{\partial x_i} \frac{\partial w_k}{\partial x_i} \, dx, \; \forall \, k = 1, 2, \cdots, N_{0h}, \end{cases} \tag{5.8}$$

$$\begin{cases} D^2_{h12}(\varphi) (= D^2_{h21}(\varphi)) \in V_{0h}, \\ D^2_{h12}(\varphi)(P_k) = -\frac{3}{2A_k} \int_\Omega \left[ \frac{\partial \varphi}{\partial x_1} \frac{\partial w_k}{\partial x_2} + \frac{\partial \varphi}{\partial x_2} \frac{\partial w_k}{\partial x_1} \right] dx, \\ \qquad \forall \, k = 1, 2, \cdots, N_{0h}. \end{cases} \tag{5.9}$$

Computing the integrals in the right hand sides of (5.8) and (5.9) is quite simple since the first order derivatives of $\varphi$ and $w_k$ are piecewise constant.

Taking the above relations into account, approximating (E-MAD) is now a fairly simple issue. Assuming that the boundary function $g$ is continuous over $\Gamma$, we approximate the affine space $V_g$ by

$$V_{gh} = \{\varphi | \, \varphi \in V_h, \, \varphi(P) = g(P), \, \forall \, P \in \Sigma_h \cap \Gamma\}, \tag{5.10}$$

and then (E-MAD) by

$$(\text{E-MAD})_h \qquad \begin{cases} \textit{Find } \psi_h \in V_{gh} \textit{ such that} \\ D^2_{h11}(\psi_h)(P_k)\, D^2_{h22}(\psi_h)(P_k) - (D^2_{h12}(\psi_h)(P_k))^2 \\ \quad = f_h(P_k), \, \forall \, k = 1, 2, \cdots, N_{0h}; \end{cases}$$

above, $f_h$ is a continuous approximation of function $f$. The iterative solution of problem (E-MAD)$_h$ will be discussed in the following paragraph.

**Remark 5.1.** Suppose that $\Omega = (0,1)^2$ and that triangulation $\mathcal{T}_h$ is like the one shown on Figure 5.1.

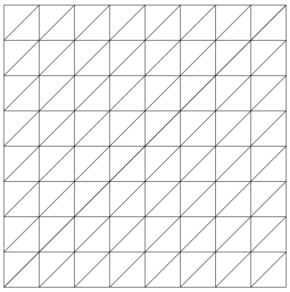

FIG. 5.1. *A uniform triangulation of* $\Omega = (0,1)^2$

Suppose that $h = \frac{1}{I+1}$, $I$ being a positive integer greater than 1. In this particular case, the sets $\Sigma_h$ and $\Sigma_{0h}$ are given by

$$\begin{cases} \Sigma_h = \{P_{ij} | \, P_{ij} = \{ih, jh\}, \, 0 \le i, j \le I+1\}, \\ \Sigma_{0h} = \{P_{ij} | \, P_{ij} = \{ih, jh\}, \, 1 \le i, j \le I\}, \end{cases} \tag{5.11}$$

implying that $N_h = (I+2)^2$ and $N_{0h} = I^2$. It follows then from relations (5.8), (5.9) that (with obvious notation):

$$D^2_{h11}(\varphi)(P_{ij}) = \frac{\varphi_{i+1,j} + \varphi_{i-1,j} - 2\varphi_{ij}}{h^2}, \; 1 \le i, j \le I, \tag{5.12}$$

$$D^2_{h22}(\varphi)(P_{ij}) = \frac{\varphi_{i,j+1} + \varphi_{i,j-1} - 2\varphi_{ij}}{h^2}, \; 1 \le i, j \le I, \tag{5.13}$$

$$\begin{aligned} D^2_{h12}(\varphi)(P_{ij}) = & \frac{(\varphi_{i+1,j+1} + \varphi_{i-1,j-1} + 2\varphi_{ij})}{2h^2} \\ & - \frac{(\varphi_{i+1,j} + \varphi_{i-1,j} + \varphi_{i,j+1} + \varphi_{i,j-1})}{2h^2}, \; 1 \le i, j \le I. \end{aligned} \tag{5.14}$$

The finite difference formulas (5.12)-(5.14) are exact for the polynomials of degree $\le$ 2. Also, as expected,

$$D^2_{h11}(\varphi)(P_{ij}) + D^2_{h22}(\varphi)(P_{ij}) = \frac{\varphi_{i+1,j} + \varphi_{i-1,j} + \varphi_{i,j+1} + \varphi_{i,j-1} - 4\varphi_{ij}}{h^2}; \tag{5.15}$$

we have recovered, thus, the well-known 5-point discretization formula for the finite difference approximation of the Laplace operator.

**5.3. Iterative solution of problem (E-MAD)$_h$.** Inspired by Sections 3 and 4, we will discuss now the solution of (E-MAD)$_h$ by a discrete variant of algorithm (4.1)-(4.4). A first step in this direction is to approximate the saddle-point problem (3.5). To achieve such a goal we approximate the set $(V_g \times \mathbf{Q}_f) \times \mathbf{Q}$ by $(V_{gh} \times \mathbf{Q}_{fh}) \times \mathbf{Q}_h$, with

$$\mathbf{Q}_h = \{\mathbf{q} | \, \mathbf{q} = \{q_{ij}\}, \, 1 \le i, j \le 2, \, q_{21} = q_{12}, \, q_{ij} \in V_{0h}\} \tag{5.16}$$

and

$$\mathbf{Q}_{fh} = \{\mathbf{q} | \, \mathbf{q} \in \mathbf{Q}_h, \, \det \mathbf{q}(P_k) = f_h(P_k), \, \forall\, k = 1, 2, \cdots, N_{0h}\}. \tag{5.17}$$

Next, we approximate the augmented Lagrangian $\mathcal{L}_r$ by $\mathcal{L}_{rh}$ defined as follows:

$$\begin{aligned}\mathcal{L}_{rh}(\varphi, \mathbf{q}; \boldsymbol{\mu}) = \frac{1}{2}\|\triangle_h \varphi\|_h^2 + \frac{r}{2}\|\mathbf{D}_h^2 \varphi - \mathbf{q}\|_h^2 + ((\boldsymbol{\mu}, \mathbf{D}_h^2(\varphi) - \mathbf{q}))_h,\\ \forall\, \varphi \in V_h, \, \mathbf{q} \in \mathbf{Q}_h, \, \boldsymbol{\mu} \in \mathbf{Q}_h,\end{aligned} \tag{5.18}$$

with

$$\triangle_h \varphi = D_{h11}^2(\varphi) + D_{h22}^2(\varphi), \tag{5.19}$$

$$\mathbf{D}_h^2(\varphi) = (D_{hij}^2(\varphi)), \, 1 \le i, j \le 2, \tag{5.20}$$

$$(v, w)_h = \frac{1}{3}\Sigma_{k=1}^{N_{0h}} A_k v(P_k) w(P_k), \, \forall\, v, w \in V_{0h}, \tag{5.21}$$

$$((\mathbf{s}, \mathbf{t}))_h = \frac{1}{3}\Sigma_{k=1}^{N_{0h}} A_k \mathbf{s}(P_k) : \mathbf{t}(P_k), \, \forall\, \mathbf{s}, \mathbf{t} \in \mathbf{Q}_h, \tag{5.22}$$

and

$$\|v\|_h = (v, v)_h^{1/2}, \, \forall\, v \in V_{0h}, \, \|\mathbf{s}\|_h = ((\mathbf{s}, \mathbf{s}))_h^{1/2}, \, \forall\, \mathbf{s} \in \mathbf{Q}_h. \tag{5.23}$$

From the above relations, we approximate problem (3.5) by the discrete saddle-point problem

$$\begin{cases} \{\{\psi_h, \mathbf{p}_h\}, \boldsymbol{\lambda}_h\} \in (V_{gh} \times \mathbf{Q}_{fh}) \times \mathbf{Q}_h, \\ \mathcal{L}_{rh}(\psi_h, \mathbf{p}_h; \boldsymbol{\mu}) \le \mathcal{L}_{rh}(\psi_h, \mathbf{p}_h; \boldsymbol{\lambda}_h) \le \mathcal{L}_{rh}(\varphi, \mathbf{q}; \boldsymbol{\lambda}_h), \\ \quad \forall\, \{\{\varphi, \mathbf{q}\}, \boldsymbol{\mu}\} \in (V_{gh} \times \mathbf{Q}_{fh}) \times \mathbf{Q}_h, \end{cases} \tag{5.24}$$

To solve the saddle-point problem (5.24) we shall use the following iterative method, a discrete variant of algorithm (4.1)-(4.4):

$$\{\psi_h^{-1}, \boldsymbol{\lambda}_h^0\} \text{ is given in } V_{gh} \times \mathbf{Q}_h; \tag{5.25}$$

*then, for* $n \ge 0$, $\{\psi_h^{n-1}, \boldsymbol{\lambda}_h^n\}$ *being known in* $V_{gh} \times \mathbf{Q}_h$, *solve*

$$\begin{cases} \mathbf{p}_h^n \in \mathbf{Q}_{fh}, \\ \mathcal{L}_{rh}(\psi_h^{n-1}, \mathbf{p}_h^n; \boldsymbol{\lambda}_h^n) \le \mathcal{L}_{rh}(\psi_h^{n-1}, \mathbf{q}; \boldsymbol{\lambda}_h^n), \, \forall\, \mathbf{q} \in \mathbf{Q}_{fh}, \end{cases} \tag{5.26}$$

$$\begin{cases} \psi_h^n \in V_{gh}, \\ \mathcal{L}_{rh}(\psi_h^n, \mathbf{p}_h^n; \boldsymbol{\lambda}_h^n) \le \mathcal{L}_{rh}(\varphi, \mathbf{p}_h^n; \boldsymbol{\lambda}_h^n), \, \forall\, \varphi \in V_{0h}, \end{cases} \tag{5.27}$$

*and update* $\boldsymbol{\lambda}_h^n$ *by*

$$\boldsymbol{\lambda}_h^{n+1} = \boldsymbol{\lambda}_h^n + r(\mathbf{D}_h^2 \psi_h^n - \mathbf{p}_h^n). \tag{5.28}$$

**ETNA**
Kent State University
etna@mcs.kent.edu



Concerning the initialization of algorithm (5.25)-(5.28), we will use the following discrete variant of the strategy advocated in Section 4.1, Remark 4.1:

*Take* $\boldsymbol{\lambda}_h^0 = \mathbf{0}$ *and* $\psi_h^{-1}$ *defined by*

$$\begin{cases} \psi_h^{-1} \in V_{gh}, \\ \int_\Omega \nabla\psi_h^{-1}\cdot\nabla\varphi\, dx = \frac{1}{3}\Sigma_{k=1}^{N_{0h}} A_k(f_h(P_k))^{1/2}\varphi(P_k),\ \forall\,\varphi \in V_{0h}. \end{cases} \tag{5.29}$$

**5.4. Solution of sub-problems (5.26).** Any sub-problem (5.26) is in fact a system of $N_{0h}$ fully decoupled minimization problems of the following kind:

$$\begin{aligned} &\textit{Minimize the functional } \mathbf{z}(=\{z_i\}_{i=1}^3) \to \frac{r}{2}(z_1^2 + z_2^2 + 2z_3^2) - \mathbf{b}_{nk}\cdot\mathbf{z} \\ &\textit{over the set defined by } z_1z_2 - z_3^2 = f_h(P_k),\ k = 1, 2, \cdots, N_{0h}. \end{aligned} \tag{5.30}$$

The solution of minimization problems of the same type than the ones in (5.30) has been addressed in Section 4.3.

**5.5. Solution of sub-problems (5.27).** We follow here Section 4.4. Any sub-problem (5.27) is equivalent to a well-posed finite dimensional linear variational problem which reads as follows:

$$\begin{cases} \psi_h^n \in V_{gh}, \\ (\triangle_h\psi_h^n, \triangle_h\varphi)_h + r((\mathbf{D}_h^2(\psi_h^n), \mathbf{D}_h^2(\varphi)))_h = L_{nh}(\varphi),\ \forall\,\varphi \in V_{0h}, \end{cases} \tag{5.31}$$

with functional $L_{hn}(\cdot)$ linear over $V_{0h}$. Problem (5.31) can be solved by a conjugate gradient algorithm operating in spaces $V_{0h}$ and $V_{gh}$ equipped with the scalar product $\{v, w\} \to (\triangle_h v, \triangle_h w)_h$. This algorithm, which is a discrete variant of (4.30)-(4.38), reads as follows:

$$\psi_h^{n,0} \textit{ is given in } V_{gh} \textit{ (a natural choice being } \psi_h^{n,0} = \psi_h^{n-1}\textit{ );} \tag{5.32}$$

*solve the following discrete bi-harmonic problem*

$$\begin{cases} g_h^{n,0} \in V_{0h}, \\ (\triangle_h g_h^{n,0}, \triangle_h\varphi)_h = (\triangle_h\psi_h^{n,0}, \triangle_h\varphi)_h + r((\mathbf{D}_h^2(\psi_h^{n,0}), \mathbf{D}_h^2(\varphi)))_h \\ \quad -L_{nh}(\varphi),\ \forall\,\varphi \in V_{0h}, \end{cases} \tag{5.33}$$

*and set*

$$w_h^{n,0} = g_h^{n,0}. \tag{5.34}$$

*For* $k \geq 0$*, assuming that* $\psi_h^{n,k}, g_h^{n,k}$ *and* $w_h^{n,k}$*, are known with the last two different from* $0$*, compute* $\psi_h^{n,k+1}, g_h^{n,k+1}$ *and, if necessary,* $w_h^{n,k+1}$ *as follows:*

*Solve the discrete bi-harmonic problem*

$$\begin{cases} \bar{g}_h^{n,k} \in V_{0h}, \\ (\triangle_h\bar{g}_h^{n,k}, \triangle_h\varphi)_h = (\triangle_h w_h^{n,k}, \triangle_h\varphi)_h + r((\mathbf{D}_h^2(w_h^{n,k}), \mathbf{D}_h^2(\varphi)))_h, \\ \ \forall\,\varphi \in V_{0h}, \end{cases} \tag{5.35}$$

*then compute*

$$\rho_{n,k} = \frac{\|\triangle_h g_h^{n,k}\|_h^2}{(\triangle_h \bar{g}_h^{n,k}, \triangle_h w_h^{n,k})_h} \tag{5.36}$$

*and set*

$$\psi_h^{n,k+1} = \psi_h^{n,k} - \rho_{n,k} w_h^{n,k}, \tag{5.37}$$

$$g_h^{n,k+1} = g_h^{n,k} - \rho_{n,k} \bar{g}_h^{n,k}. \tag{5.38}$$

*If* $\frac{\|\triangle_h g_h^{n,k+1}\|_h}{\|\triangle_h g_h^{n,0}\|_h} \leq \epsilon$ *take* $\psi_h^{n+1} = \psi_h^{n,k+1}$; *else, compute*

$$\gamma_{n,k} = \frac{\|\triangle_h g_h^{n,k+1}\|_h^2}{\|\triangle_h g_h^{n,k}\|_h^2}, \tag{5.39}$$

$$w_h^{n,k+1} = g_h^{n,k+1} + \gamma_{n,k} w_h^{n,k}. \tag{5.40}$$

*Do* $k = k + 1$ *and return to* (5.35).

**Remark 5.2.** Each iteration of algorithm (5.32)-(5.40) requires the solution of a discrete bi-harmonic problem of the following type:

$$\begin{cases} \textit{Find } u_h \in V_{0h}(\textit{ or } V_{gh}) \\ \textit{such that } (\triangle_h u_h, \triangle_h v)_h = L_h(v), \; \forall\, v \in V_{0h}, \end{cases} \tag{5.41}$$

functional $L_h(\cdot)$ in (5.41) being linear. Let us denote $-\triangle_h u_h$ by $\omega_h$; it follows then from (5.4), (5.19) and (5.21) that problem (5.41) is equivalent to the following system of two coupled discrete Poisson-Dirichlet problems:

$$\begin{cases} \omega_h \in V_{0h}, \\ \int_\Omega \boldsymbol{\nabla}\omega_h \cdot \boldsymbol{\nabla} v \, dx = L_h(v), \; \forall \in V_{0h}, \end{cases} \tag{5.42}$$

$$\begin{cases} u_h \in V_{0h}(\textit{ or } V_{gh}), \\ \int_\Omega \boldsymbol{\nabla} u_h \cdot \boldsymbol{\nabla} v \, dx = (\omega_h, v)_h, \; \forall \in V_{0h}. \square \end{cases} \tag{5.43}$$

Via algorithm (5.25)-(5.28) we have thus reduced the solution of (E-MAD)$_h$ to the solution of:

(i) A sequence of discrete (linear) Poisson-Dirichlet problems.

(ii) A sequence of finite dimensional minimization problems, each of them equivalent to $N_{0h}$ uncoupled systems of four nonlinear equations in four variables (one per internal grid point).

Numerical results obtained using algorithm (5.25)-(5.28) will be presented in Section 6.

**6. Numerical experiments.** We are going to apply the methodology discussed in Sections 3, 4 and 5 to the solution of four E-MAD test problems. For all these test problems, we shall assume that $\Omega = (0,1) \times (0,1)$ and that $\mathcal{T}_h$ is a uniform triangulation like the one in Figure 5.1, with $h$ the length of the edges of $\mathcal{T}_h$ adjacent to the right angles. The **first test problem** takes its inspiration in Remark 2.2; it can be expressed as follows:

$$\det D^2\psi = (1 + |x|^2)e^{|x|^2} \text{ in } \Omega, \psi = g \text{ on } \Gamma, \tag{6.1}$$

**ETNA**
Kent State University
etna@mcs.kent.edu



with $|x| = \sqrt{x_1^2 + x_2^2}$ and function $g$ given by $g(x) = e^{\frac{1}{2}x_1^2}$ on $\{x|\, 0 < x_1 < 1,\, x_2 = 0\}$, $g(x) = e^{\frac{1}{2}x_2^2}$ on $\{x|\, x_1 = 0,\, 0 < x_2 < 1\}$, $g(x) = e^{\frac{1}{2}(x_1^2+1)}$ on $\{x|\, 0 < x_1 < 1,\, x_2 = 1\}$, and $g(x) = e^{\frac{1}{2}(1+x_2^2)}$ on $\{x|\, x_1 = 1,\, 0 < x_2 < 1\}$. An exact solution to problem (6.1) is given by $\psi(x_1, x_2) = e^{\frac{1}{2}|x|^2}$. We have discretized problem (6.1) relying on the mixed formulation and approximation method discussed in Sections 3 and 5. Triangulation $\mathcal{T}_h$ being uniform, we have used fast Poisson solvers to solve the elliptic problems encountered at each iteration of algorithm (5.25)-(5.28), taking advantage thus of the decomposition properties mentioned in Remark 5.2. Using as initial guess the approximate solutions of the Poisson-Dirichlet problem $-\triangle\varphi = e^{\frac{1}{2}|x|^2}\sqrt{1+|x|^2}$ in $\Omega$, $\varphi = g$ on $\Gamma$, quite accurate approximations of the exact solution are obtained (taking $r = 1$ in algorithm (5.25)-(5.28)). After 100 iterations, we obtained the results summarized in Table 6.1, (where $\|\cdot\|_{0,\Omega} = \|\cdot\|_{L^2(\Omega)}$ or $\|\cdot\|_{(L^2(\Omega))^4}$ ):

TABLE 6.1
*First test problem: Convergence of the approximate solutions*

| $h$ | $\|\psi_h^c - \psi\|_{0,\Omega}$ | $\|\mathbf{D}_h^2\psi_h^c - \mathbf{p}_h^c\|_{0,\Omega}$ |
|---|---|---|
| $1/32$ | $2.6 \times 10^{-5}$ | $2.8 \times 10^{-5}$ |
| $1/64$ | $6.7 \times 10^{-6}$ | $4.1 \times 10^{-5}$ |
| $1/128$ | $1.8 \times 10^{-6}$ | $4.8 \times 10^{-5}$ |

It follows from the above table (where $\psi_h^c$ denotes the computed approximate solution, $\mathbf{D}_h^2\psi_h^c$ the corresponding approximate Hessian and $\mathbf{p}_h^c$ the computed approximation of tensor $\mathbf{p}$) that the $L^2(\Omega)$-approximation error is clearly $O(h^2)$, which is the best we can hope, generically, when using piecewise linear finite element approximations (without post-processing) to solve a second order elliptic problem. The graph of $\psi_h^c$ obtained, via algorithm (5.25)-(5.28), for $h = 1/128$, has been visualized on Figure 6.1.

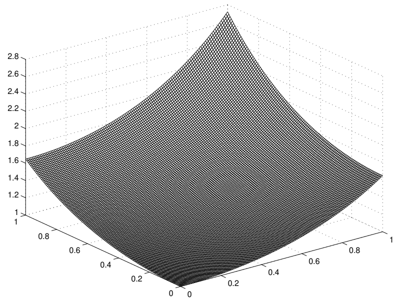


FIG. 6.1. *First test problem: Graph of $\psi_h^c$ ($h = 1/128$)*

The **second test problem** is defined as follows:

**(i)** $f(x) = \frac{R^2}{(R^2-|x|^2)^2}$, $\forall\, x \in \Omega$, with $R \geq \sqrt{2}$ and $|x|$ as above.

**(ii)** $g(x) = \sqrt{R^2 - |x|^2}$, $\forall\, x \in \Gamma$.

If the above data prevail, function $\psi$ given by

$$\psi(x) = \sqrt{R^2 - |x|^2} \tag{6.2}$$

is a solution to the corresponding E-MAD problem. The graph of function $\psi$ is clearly a piece of the sphere of radius $R$ centered at $\{0, 0, 0\}$. If $R > \sqrt{2}$ we have $\psi = C^\infty(\bar{\Omega})$; on

**ETNA**
Kent State University
etna@mcs.kent.edu



the other hand, if $R = \sqrt{2}$ we have no better than $\psi \in W^{1,p}(\Omega)$ with $p \in [1, 4)$, implies that in that particular case $\psi$ does not have the $H^2$-regularity. When applying the computational methods discussed in Sections 3, 4 and 5 to the solution of the above problem (with $r = 1$ in algorithm (4.1)-(4.4)), we obtain if $R = 2$, and after 78 iterations of the discrete variant of the above algorithm (namely, algorithm (5.25)-(5.28)), the results displayed in Table 6.2.

TABLE 6.2
*Results for the $2^{nd}$ test problem ($R = 2$).*

| $h$ | $\|\|\psi_h^c - \psi\|\|_{0,\Omega}$ | $\|\|\mathbf{D}_h^2\psi_h^c - \mathbf{p}_h^c\|\|_{0,\Omega}$ |
|---|---|---|
| $1/32$ | $4.45 \times 10^{-6}$ | $9.48 \times 10^{-7}$ |
| $1/64$ | $1.14 \times 10^{-6}$ | $1.35 \times 10^{-6}$ |
| $1/128$ | $2.97 \times 10^{-7}$ | $1.58 \times 10^{-6}$ |

In the above table, $\psi_h^c$ is the computed approximate solution, $\mathbf{D}_h^2\psi_h^c$ is the corresponding discrete Hessian, and $\mathbf{p}_h^c$ is the computed approximation of tensor $\mathbf{p}$. The above results strongly suggest second order accuracy, which is-once again-optimal considering the type of finite element approximations we are using. If we take $R = \sqrt{2}$, our methodology which has been designed to solve (E-MAD) in $H^2(\Omega)$ is unable to capture any solution of the above problem, the corresponding algorithm (5.25)-(5.28) being divergent for any value of $r$. The same troubles persist if one takes $R = \sqrt{2} + 10^{-2}$; on the other hand, if one takes $R = \sqrt{2} + 10^{-1}$, things are back to normal since, using again $r = 1$, we obtain after 117 iterations of algorithm (5.25)-(5.28) the results summarized in Table 6.3:

TABLE 6.3
*Results for the $2^{nd}$ test problem ($R = \sqrt{2} + 10^{-1}$).*

| $h$ | $\|\|\psi_h^c - \psi\|\|_{0,\Omega}$ | $\|\|\mathbf{D}_h^2\psi_h^c - \mathbf{p}_h^c\|\|_{0,\Omega}$ |
|---|---|---|
| $1/32$ | $2.20 \times 10^{-5}$ | $9.68 \times 10^{-7}$ |
| $1/64$ | $5.51 \times 10^{-6}$ | $1.54 \times 10^{-6}$ |
| $1/128$ | $1.37 \times 10^{-6}$ | $2.04 \times 10^{-6}$ |

The above results show that second order accuracy holds again. However, the second order derivatives of $\psi$ being larger for $R = \sqrt{2} + 10^{-1}$ than for $R = 2$, the corresponding approximations errors are also larger. On Figures 6.2 to 6.5 we have visualized, respectively:
**(i)** The graph of $\psi$ given by (6.2) when $R = \sqrt{2}$; the singularity of $\boldsymbol{\nabla}\psi$ at $\{1, 1\}$ appears clearly on Fig. 6.2.
**(ii)** The graph of $\psi_h^c$ corresponding to $h = 1/128$ and $R = 2$.
**(iii)** The graph of $\psi_h^c$ corresponding to $h = 1/128$ and $R = \sqrt{2} + 10^{-1}$.
**(iv)** The graph of $f$ when $R = \sqrt{2} + 10^{-1}$.

**Remark 6.1.** When computing the approximate solutions of the second test problem for $h = 1/32$, we stopped the iterations of algorithm (5.25)-(5.28) as soon as

$$\|\mathbf{D}_h^2\psi_h^n - \mathbf{p}_h^n\|_{0,\Omega} \leq 10^{-6}. \tag{6.3}$$

The corresponding number of iterations is 78 for $R = 2$, and 117 for $R = \sqrt{2} + 10^{-1}$ (we did not try to find the optimal value of $r$, or to use a variable $r$ strategy). Next, when computing the approximate solutions for $h = 1/64$ and $1/128$, we stopped iterating once the iteration numbers associated to $h = 1/32$ were reached (actually, using (6.3) as stopping criteria for $h = 1/64$ and $1/128$ did not change much the approximation errors shown in Tables 6.2 and 6.3). A similar approach was used for the other test problems. For a given

**ETNA**
Kent State University
etna@mcs.kent.edu



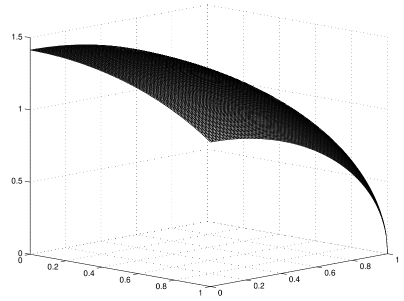

FIG. 6.2. *Second test problem: graph of $\psi$ when $R = \sqrt{2}$.*

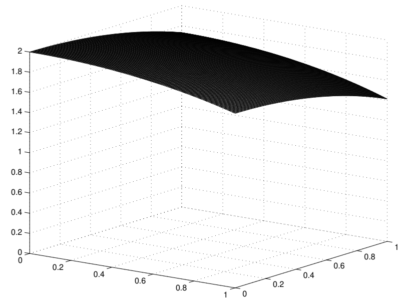

FIG. 6.3. *Second test problem: graph of $\psi_h^c$ when $R = 2$ and $h = 1/128$.*

tolerance in the stopping criterion the number of iteration necessary to achieve convergence increases (slightly) with $1/h$, implying in turn that for a given number of iterations, the residual $\|\mathbf{D}_h^2\psi_h^n - \mathbf{p}_h^n\|_{0,\Omega}$ increases (slightly) with $1/h$ (like $1/\sqrt{h}$, roughly), explaining thus the results observed in the last column of Tables 6.2 and 6.3; the above observation applies also to the other test problems.

The **third test problem** is defined as follows:

**(i)** $f(x) = 1/|x|$, $\forall\, x \in \Omega$.

**(ii)** $g(x) = \frac{(2|x|)^{\frac{3}{2}}}{3}$, $\forall\, x \in \Gamma$.

It follows from Remark 2.2 that, with these data, a solution to (E-MAD) is the function $\psi$ defined by

$$\psi(x) = \frac{(2|x|)^{\frac{3}{2}}}{3}, \forall x \in \Omega. \tag{6.4}$$

One can easily check that $\psi \notin C^2(\bar{\Omega})$; however, since $\psi \in W^{2,p}(\Omega)$ for all $p \in [1,4)$, it has, in principle, enough regularity so that we can apply algorithms (4.1)-(4.4) and (5.25)-(5.28) to the solution of the corresponding problem (E-MAD). Indeed, despite the singularity of function $f$ at $\{0,0\}$ (see Fig. 6.6), algorithm (5.25)-(5.28), with $r = 1$, provides after 160 iterations the results summarized in the Table 6.4.

From the table results, we can infer that second order accuracy still holds. On Figure 6.6 (resp., 6.7) we have visualized function $f$ (resp., the computed approximate solution obtained with $h = 1/128$).

The **fourth test problem** is-in some sense-the more interesting since we consider this

**ETNA**
Kent State University
etna@mcs.kent.edu



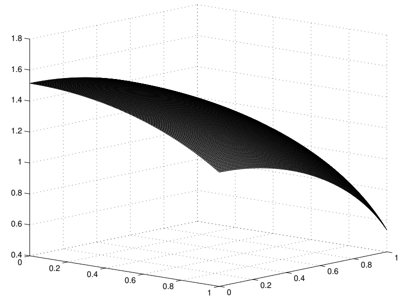

FIG. 6.4. *Second test problem: graph of* $\psi_h^c$ *when* $R = \sqrt{2} + 10^{-1}$ *and* $h = 1/128$.

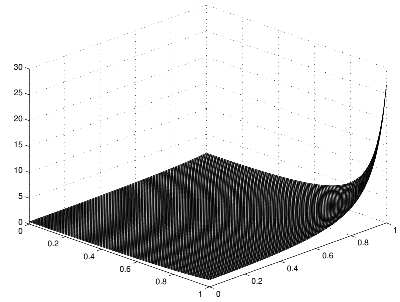

FIG. 6.5. *Second test problem: graph of* $f$ *when* $R = \sqrt{2} + 10^{-1}$.

time the solution of problem (2.1), namely:

$$\frac{\partial^2 \psi}{\partial x_1^2}\frac{\partial^2 \psi}{\partial x_2^2} - |\frac{\partial^2 \psi}{\partial x_1 \partial x_2}|^2 = 1 \text{ in } \Omega,\ \psi = 0 \text{ on } \Gamma,\ \text{with } \Omega = (0,1)\times(0,1).$$

Despite the smoothness of its data, the above problem has no smooth solution. The trouble is coming from the non-strict convexity of $\Omega$. When applying algorithm (4.1)-(4.4) (in fact its discrete variant (5.25)-(5.28)) to the solution of problem (2.1) we observe the following phenomena:

**(i)** For $r$ sufficiently small ($r = 1$ here) sequence $\{\psi_h^n, \mathbf{p}_h^n\}_{n\geq 0}$ converges geometrically (albeit slowly) to a limit $\{\psi_h^c, \mathbf{p}_h^c\}$ while sequence $\{\boldsymbol{\lambda}_h^n\}_{n\geq 0}$ diverges arithmetically.

**(ii)** A close inspection of the numerical results show that the curvature of the graph of $\psi_h^c$ becomes negative close to the corners, in violation of the Monge-Ampère equation; actually, as expected, it is also violated along the boundary, since $\|\mathbf{D}_h^2\psi_h^c - \mathbf{p}_h^c\|_{0,\Omega} = 1.8\times 10^{-2}$ if $h = 1/32$, $3.3\times 10^{-2}$ if $h = 1/64$, $4.2\times 10^{-2}$ if $h = 1/128$, while $\|\mathbf{D}_h^2\psi_h^c - \mathbf{p}_h^c\|_{0,\Omega_1} = 2.7\times 10^{-4}$ if $h = 1/32$, $4.1\times 10^{-4}$ if $h = 1/64$, $4.9\times 10^{-4}$ if $h = 1/128$, and $\|\mathbf{D}_h^2\psi_h^c - \mathbf{p}_h^c\|_{0,\Omega_2} = 4.4\times 10^{-5}$ if $h = 1/32$, $4.9\times 10^{-5}$ if $h = 1/64$, $5.1\times 10^{-5}$ if $h = 1/128$, where $\Omega_1 = (1/8, 7/8)\times(1/8, 7/8)$ and $\Omega_2 = (1/4, 3/4)\times(1/4, 3/4)$. These results suggest that $\det D^2\psi = 1$ is "almost" verified in $\Omega_2$.

The graph of $\psi_h^c$ obtained with $h = 1/64$ has been shown on Figure 6.8, while the intersections of this graph with the planes $x_1 = 1/2$ and $x_1 = x_2$ have been shown on Figures 6.9 and 6.10, respectively, for $h = 1/32$, $1/64$, and $1/128$. Since $\psi_h^c$ does not vary very much with $h$, we suspect that according to Section 4.2, what we have here is a (good)

**ETNA**
Kent State University
etna@mcs.kent.edu



TABLE 6.4
*Results for the $3^{rd}$ test problem.*

| $h$ | $\|\|\psi_h^c - \psi\|\|_{0,\Omega}$ | $\|\|\mathbf{D}_h^2\psi_h^c - \mathbf{p}_h^c\|\|_{0,\Omega}$ |
|---|---|---|
| $1/32$ | $5.56 \times 10^{-5}$ | $9.91 \times 10^{-7}$ |
| $1/64$ | $1.50 \times 10^{-5}$ | $1.60 \times 10^{-6}$ |
| $1/128$ | $3.94 \times 10^{-6}$ | $2.02 \times 10^{-6}$ |

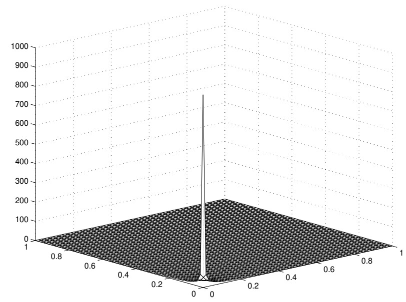


FIG. 6.6. *Third test problem: graph of $f$.*

approximation of one of those functions of $H^2(\Omega) \cap H_0^1(\Omega)$ whose Hessian is at a minimal $L^2$ - distance (global or local) from the set $\mathbf{Q}_f$ defined in Section 3. Assuming that the above is true (a convincing evidence will be given in Section 7) we can claim that the solution-less problem (2.1) has been solved in a least squares sense in the functional space $H^2(\Omega)$, leading to a concept of generalized solution (may be not so novel in general, but possibly new in the Monge-Ampère "culture").

**Remark 6.2.** When applying algorithm (5.25)-(5.28) to the solution of (E-MAD) we have to solve, at each iteration, the linear variational problem (5.27). To solve this finite dimensional problem, we have used the preconditioned conjugate gradient algorithm (5.32)-(5.40); this algorithm, which converged typically in 5 to 7 iterations in all the applications we used it, requires the solution of two discrete Poisson-Dirichlet problems per iteration for preconditioning purpose. As already mentioned, our triangulations being uniform we have used Fast Poisson Solvers to achieve preconditioning.

**7. Further comments.**

- From a geometrical point of view (E-MAD) has solutions in $H^2(\Omega)$ if $D^2V_g$ and $\mathbf{Q}_f$ (these two subsets of space $\mathbf{Q}$ have been introduced in Section 3) intersect; such a situation has been visualized on Figure 7.1. Figure 7.2 corresponds to a situation where (E-MAD) has no solution in $H^2(\Omega)$, albeit neither $D^2V_g$ nor $\mathbf{Q}_f$ are empty; a generalized solution in the sense of Remark 4.2 has been visualized on this figure. The above observation suggests the following least squares approach for the solution of (E-MAD):

$$
\begin{cases} \text{Find } \{\psi, \mathbf{p}\} \in V_g \times \mathbf{Q}_f \text{ such that} \\ j(\psi, \mathbf{p}) \le j(\varphi, \mathbf{q}),\ \forall \{\varphi, \mathbf{q}\} \in V_g \times \mathbf{Q}_f, \end{cases} \tag{7.1}
$$

with $j(\varphi, \mathbf{q}) = \frac{1}{2}\int_\Omega |\mathbf{q} - D^2\varphi|^2\,dx$. We conjecture (see Remark 4.2)that it is precisely to a solution of problem (7.1) that algorithm (4.1)-(4.4) "converges" when $D^2V_g \cap \mathbf{Q}_f = \emptyset$, with neither $V_g$ nor $\mathbf{Q}_f$ empty.

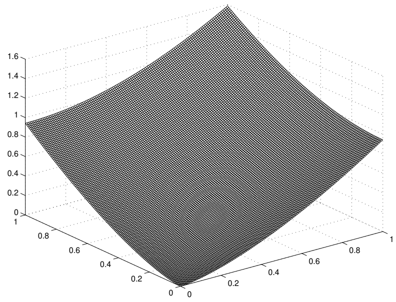


FIG. 6.7. *Third test problem: graph of* $\psi_h^c$ *(*$h = 1/128$*).*

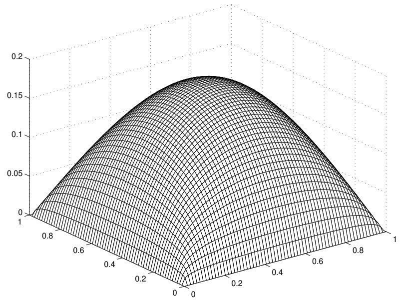


FIG. 6.8. *Fourth test problem: Graph of the computed solution (h=1/64).*

The solution of (E-MAD), via the (nonlinear) least squares formulation (7.1), is discussed in [38]; when applied to the solution of the first three test problems, the least squares methodology reproduces the results shown in Section 6. The good news is that it also reproduces the results obtained in the above section when applied to the solution of the $4^{th}$ test problem since, for this problem, $\|\psi_h^{LS} - \psi_h^{AL}\|_{L^2(\Omega)}$ of the order of $10^{-5}$ with $\psi_h^{LS}$ (respectively $\psi_h^{AL}$) the approximate solution of (E-MAD) computed via the above least squares approach (respectively, by the augmented Lagrangian methodology discussed in this article).

- It follows from [40], [41] that the viscosity solutions of Monge-Ampère, and related equations, can be approximated uniformly on the compact subsets of $\Omega$ by family of regularized viscosity solutions in the sense of Jensen (see [14] and [40] for details.)
- The saddle-point formulation (3.5) of problem (E-MAD) is a mixed variational formulation where the "nonlinearity burden" has been transferred from $\psi$ to $\mathbf{p}$, making it purely algebraic. Actually, this approach (this is even truer for the discrete analogues of problems (E-MAD) and (3.5); see Section 5 for details) provides a solution method where instead of solving (E-MAD) directly (i.e.,without introducing additional functions, like $\mathbf{p}$), we solve it via its associated Pfaff system (see, e.g., [39, Chapter AV]), namely:

$$
\begin{aligned}
& d\psi - u_1 dx_1 - u_2 dx_2 = 0 \text{ in } \Omega, \\
& du_1 - p_{11} dx_1 - p_{12} dx_2 = 0 \text{ in } \Omega, \\
& du_2 - p_{12} dx_1 - p_{22} dx_2 = 0 \text{ in } \Omega, \\
& p_{11} p_{22} - p_{12}^2 = f,
\end{aligned}
\tag{7.2}
$$

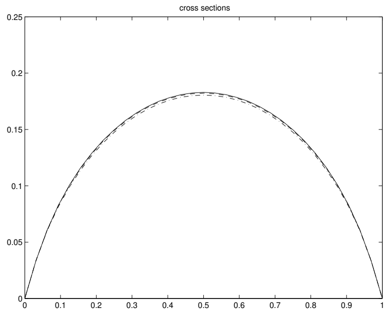


FIG. 6.9. *Fourth test problem: Graphs of the computed solutions restricted to the plane* $x_1 = 1/2$.

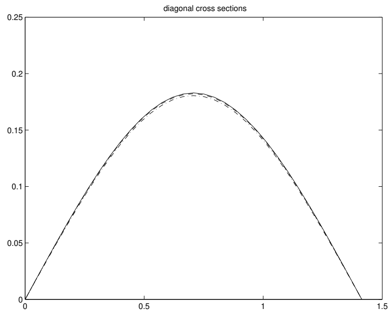


FIG. 6.10. *Fourth test problem: Graphs of the computed solutions restricted to the plane* $x_1 = x_2$.

completed by the boundary condition

$$\psi = g \text{ on } \Gamma.$$

System (7.2) provides clearly a mixed formulation of (E-MAD).

**Acknowledgments.** The authors would like to thank D. Bao, J. D. Benamou, Y. Brenier, L. Caffarelli, P. G. Ciarlet, B. Dacorogna, P. L. Lions, and the two reviewers, for helpful comments and suggestions. They also thank S. Mas-Gallic for her interest in (E-MAD).

REFERENCES

[1] L. A.CAFFARELLI, *Nonlinear elliptic theory and the Monge-Ampère equation*, in Proceedings of the International Congress of Mathematicians, Beijing 2002, August 20-28, Vol. I, Higher Education Press, Beijing, 2002, pp. 179–187.

[2] Y. BRENIER, *Some geometric PDEs related to Hydrodynamics and Electrodynamics*, in Proceedings of the International Congress of Mathematicians, Beijing 2002, August 20-28, Vol. III, Higher Education Press, Beijing, 2002, pp. 761–771.

[3] S-Y. A. CHANG AND P. C. YANG, *Nonlinear partial differential equations in conformal geometry*, in Proceedings of the International Congress of Mathematicians, Beijing 2002, August 20-28, Vol. I, Higher Education Press, Beijing, 2002, pp. 189–207.

[4] D. BAO, S.-S. CHERN, AND Z. SHEN, *An Introduction to Riemann-Finsler Geometry*,Graduate Texts in Mathematics, 200. Springer-Verlag, New York, 2000.

[5] R. GLOWINSKI AND O. PIRONNEAU, *Numerical methods for the first biharmonic equation and for the two-dimensional Stokes problem*, SIAM Rev., 17 (1979), pp. 167-212.

[6] R. GLOWINSKI, J. L. LIONS, AND R. TREMOLIÉRES, *Numerical Analysis of Variational Inequalities*, Translated from the French. Studies in Mathematics and its Applications, 8. North-Holland Publishing Co., Amsterdam-New York, 1981.

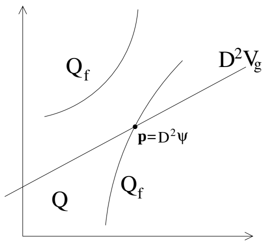


FIG. 7.1. *Problem (E-MAD) has a solution in $H^2(\Omega)$.*

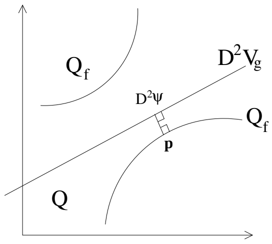


FIG. 7.2. *Problem (E-MAD) has no solution in $H^2(\Omega)$.*

[7] L. REINHART, *On the numerical analysis of the Von Kármán equations: Mixed finite element approximation and continuation techniques*, Numer. Math., 39 (1982), pp. 371–404.

[8] R. GLOWINSKI, H. B. KELLER, AND L. REINHART, *Continuation-conjugate gradient methods for the least-squares solution of nonlinear boundary value problems*, SIAM J. Sci. Stat. Comp., 4 (1985), pp. 793–832.

[9] R. GLOWINSKI, D. MARINI, AND M. VIDRASCU, *Finite element approximation and iterative solution of a fourth-order elliptic variational inequality*, IMA J. Numer. Anal., 4 (1984), pp. 127–167.

[10] M. FORTIN AND R. GLOWINSKI, *Augmented Lagrangian Methods: Application to the Numerical Solution of Boundary Value Problems*, North-Holland, Amsterdam, 1983.

[11] R. GLOWINSKI AND P. LE TALLEC, *Augmented Lagrangians and Operator-Splitting Methods in Nonlinear Mechanics*, SIAM, Philadelphia, PA, 1989.

[12] E. J. DEAN, R. GLOWINSKI, AND D. TREVAS, *An approximate factorization / least-squares solution method for a mixed finite element approximation of the Cahn- Hilliard equation*, Japan J. Industr. Appl. Math., 13 (1996), pp. 495–517.

[13] L. A. CAFFARELLI AND X. CABRÉ, *Fully Nonlinear Elliptic Equations*, American Mathematical Society, Providence, RI, 1995.

[14] X. CABRÉ, *Topics in regularity and qualitative properties of solutions of nonlinear elliptic equations*,Current developments in partial differential equations (Temuco, 1999). Discrete Contin. Dyn. Syst. 8 (2002), pp. 289– 302.

[15] R. COURANT AND D. HILBERT, *Methods of Mathematical Physics*, Vol. II, Wiley Interscience, New York, NY, 1989.

[16] T. AUBIN, *Nonlinear Analysis on Manifolds, Monge-Ampère Equations*, Springer-Verlag, Berlin, 1982.

[17] T. AUBIN, *Some Nonlinear Problems in Riemannian Geometry*, Springer-Verlag, 1998.

[18] P. L. LIONS, *Une méthode nouvelle pour l'existence de solutions réguliéres de l'équation de Monge-Ampère réelle*, C. R. Acad. Sci. Paris, Série I, 293 (1981), pp. 589–592.

[19] L. A. CAFFARELLI, L. NIRENBERG, AND J. SPRUCK, *The Dirichlet problem for nonlinear second order elliptic equations, (I) Monge-Ampère equation*, Comm. Pure Appl. Math., 37 (1984), pp. 369–402.

[20] L. A. CAFFARELLI, J. J. KOHN, L. NIRENBERG, AND J. SPRUCK, *The Dirichlet problem for nonlinear second order elliptic equations, (II) Complex Monge-Ampère and uniformly elliptic equations*, Comm. Pure Appl. Math., 38 (1985), pp. 209–252.

[21] L. A. CAFFARELLI AND Y. Y. LI, *A Liouville theorem for solutions of the Monge-Ampère equation with periodic data*, Ann. Inst. H. Poincaré Anal. Non Linéaire, 21 (2004), pp. 97–120.

[22] F. X. LE DIMET AND M. OUBERDOUS, *Retrieval of balanced fields: an optimal control method*, Tellus, 45A (1993), pp.449–461.

[23] L. A. CAFFARELLI AND M. MILMAN (EDS.), *Monge-Ampère Equation: Application to Geometry and Optimization*, American Math. Society, Providence, RI, 1999.

[24] J. R. OCKENDON, S. HOWISON, A. LACEY, AND A. MOVCHAN (EDS.), *Applied Partial Differential Equations*, Oxford University Press, 1999, Oxford, UK.
[25] L. A. CAFFARELLI, S. A. KOCHENKGIN, AND V. I. OLICKER, *On the numerical solution of reflector design with given far-field scattering data*, in Monge Ampère Equation: Applications to Geometry and Optimization, L. A. Caffarelli and M. Milman, eds., American Mathematical Society, Providence, RI, 1999, pp. 13–32.
[26] V. I. OLIKER AND L. D. PRUSSNER, *On the numerical solution of the equation $z_{xx}z_{yy} - z_{xy}^2 = f$ and its discretization*, Numer. Math., 54 (1988), pp. 271–293.
[27] E. J. DEAN AND R. GLOWINSKI, *Numerical solution of the two-dimensional elliptic Monge-Ampère equation with Dirichlet boundary conditions: an augmented Lagrangian approach*, C. R. Acad. Sci. Paris, Série I, 336 (2003), pp. 779–784.
[28] S. V. KOCHENGIN AND V. I. OLIKER, *Determination of reflector surfaces from near-field scattering data*, Numer. Math., 79 (1998), pp. 553–568.
[29] E. J. DEAN, R. GLOWINSKI, AND T. W. PAN, *Operator-splitting methods and application to the direct numerical simulation of particulate flow and to the solution of the elliptic Monge-Ampère equation*, in Control and Boundary Analysis, J. Cagnol and J. P. Zolésio, eds, Marcel Dekker, New York, NY, 2005.
[30] R. GLOWINSKI, *Numerical Methods for Nonlinear Variational Problems*, Springer-Verlag, New York, NY, 1984.
[31] R. GLOWINSKI, *Finite element methods for incompressible viscous flow*, in Handbook of Numerical Analysis, Volume IX, P.G. Ciarlet and J. L. Lions, eds., North-Holland, Amsterdam, 2003, pp. 3–1176.
[32] J. E. DENNIS, AND R. B. SCHNABEL, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, SIAM, Philadelphia, PA, 1996.
[33] L. C. EVANS, *Partial differential equations and Monge-Kantorovich mass transfer*, in Current Developments in Mathematics, R. Bott, ed., International Press, Cambridge, MA, 1997, pp.26–78.
[34] J. D. BENANOU AND Y. BRENIER, *A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem*, Numer. Math., 84 (2000), pp. 375–393.
[35] P. G. CIARLET, *The Finite Element Method for Elliptic Problems*, North-Holland, Amsterdam, 1978.
[36] F. BREZZI AND M. FORTIN, *Mixed and Hybrid Finite Element Methods*, Springer- Verlag, New-York, NY, 1991.
[37] P. G. CIARLET, *Basic error estimates*, in Handbook of Numerical Analysis, Volume II, P. G. Ciarlet and J. L. Lions eds., North-Holland, Amsterdam, 1991, pp.17–351.
[38] E. J. DEAN AND R. GLOWINSKI, *Numerical solution of the two-dimensional elliptic Monge-Ampère equation with Dirichlet boundary conditions: a least squares approach*, C. R. Acad. Sci. Paris, Série I, 339 (2004), pp.887–892.
[39] J. DIEUDONNÉ, *Panorama des Mathématiques Pures: Le Choix Bourbachique*, Editions Jacques Gabay, Paris, 2003.
[40] R. JENSEN, *The maximum principle for viscosity solutions of fully nonlinear second order partial differential equations*, Arch. Rat. Mech. Anal., 101 (1988), pp. 1–27.
[41] J. I. E. URBAS, *Regularity of generalized solutions of Monge-Ampère equations*, Math. Z., 197 (1988), pp. 365–393.
[42] M. G. CRANDALL, H. ISHII, AND P. L. LIONS, *User's guide to viscosity solutions of second order partial differential equations*, Bull. Amer. Math. Soc., 27 (1992), pp. 1–67.