Journal of Computational Physics **174,** 903–924 (2001)
doi:10.1006/jcph.2001.6945, available online at http://www.idealibrary.com on IDEAL®

Variational Mesh Adaptation: Isotropy and Equidistribution

Weizhang Huang

Department of Mathematics, University of Kansas, Lawrence, Kansas 66045
E-mail: huang@math.ukans.edu

Received May 1, 2001; revised September 14, 2001

We present a new approach for developing more robust and error-oriented mesh adaptation methods. Specifically, assuming that a regular (in cell shape) and uniform (in cell size) computational mesh is used (as is commonly done in computation), we develop a criterion for mesh adaptation based on an error function whose definition is motivated by the analysis of function variation and local error behavior for linear interpolation. The criterion is then decomposed into two aspects, the isotropy (or conformity) and uniformity (or equidistribution) requirements, each of which can be easier to deal with. The functionals that satisfy these conditions approximately are constructed using discrete and continuous inequalities. A new functional is finally formulated by combining the functionals corresponding to the isotropy and uniformity requirements. The features of the functional are analyzed and demonstrated by numerical results. In particular, unlike the existing mesh adaptation functionals, the new functional has clear geometric meanings of minimization. A mesh that has the desired properties of isotropy and equidistribution can be obtained by properly choosing the values of two parameters. The analysis presented in this article also provides a better understanding of the increasingly popular method of harmonic mapping in two dimensions.

Key Words: mesh adaptation; variational method; equidistribution; conformity.

1. INTRODUCTION

In this article we are concerned with variational methods for generating adaptive meshes for use in the numerical solution of partial differential equations. A variational method utilizes a functional to determine the coordinate transformation needed for mesh generation. Such a functional is often formulated to measure difficulties in the numerical approximation of the physical solution and typically involves a so-called monitor function that is prescribed by the user to control the mesh adaptation. An important feature of the formulation is the well posedness of the functional, that is, the existence and uniqueness of the minimizer, and

0021-9991/01 $35.00

sometimes the maximum principle, should be assured. Generally, the necessary property of the nonsingularity of the coordinate transformation is always the most difficult thing to guarantee theoretically, although it does not seem to be a problem numerically for those systems having a convex computational domain and obeying the maximum principle.

Due to the well-posedness consideration, people usually do not directly use standard error estimates in the development of variational methods since they lead to nonconvex functionals in two and higher dimensions. Indeed, most of the existing variational methods have been developed based on other considerations such as geometric ones; e.g., see [3, 4, 9, 12, 16, 18, 19, 23, 25] and the books [10, 17, 21, 24] and references therein. For example, Brackbill and Saltzman [4] developed a very popular method by combining mesh concentration, smoothness, and orthogonality. Dvinsky [9] used the energy of harmonic mappings as his mesh adaptation functional. Knupp [16, 18] and Knupp and Robidoux [19] developed functionals based on the idea of conditioning the Jacobian matrix of the coordinate transformation. Cao *et al.* [5] studied qualitatively the effect of the monitor function on the behavior of the mesh for a class of functionals. However, the lack of direct connections with error estimates in the existing methods makes it tricky to choose a proper monitor function for some practical problems. Most of the existing methods certainly need to be better understood, and new methods that are more robust and more error-estimate–oriented are yet to be developed.

The objective of this article is to present a new approach of developing more robust and error-oriented mesh adaptation functionals. Specifically, assuming that a regular (in cell shape) and uniform (in cell size) computational mesh is used (as is commonly done in computation), we develop a criterion for mesh adaptation based on an error function whose definition is motivated by the analysis of function variation and local error behavior for linear interpolation. The criterion is further decomposed into two aspects, the isotropy (or conformity) and uniformity (or equidistribution) requirements, each of which can be easier to deal with. Functionals that satisfy these conditions approximately are then constructed using discrete and continuous inequalities. Finally, a new functional is formulated by combining the functionals corresponding to the isotropy and uniformity requirements. The features of the functional are analyzed and demonstrated by numerical results.

An outline of the article is as follows. In Section 2, several criteria for mesh adaptation are proposed and the mathematical implications of its isotropy and uniformity properties are presented. Sections 3 and 4 are devoted to the derivations of the functionals according to the isotropy and uniformity requirements, respectively. A functional that naturally combines the functionals developed in Sections 3 and 4 is derived in Section 5. Also in this section, the features of the newly developed functional are analyzed and the Euler–Lagrange equation is given. Illustrative numerical results are presented in Section 6. Finally, Section 7 contains conclusions and comments.

Throughout this article, the terms regularity, isotropy, and conformity are used for describing the shape of mesh cells, while the terms uniformity and equidistribution are used for measuring the change in cell size or volume. Thus, a uniform mesh is not necessarily regular. Only a mesh having equilateral cells can be regular.

2. CRITERIA FOR MESH ADAPTATION

In this section several criteria that are used for constructing mesh adaptation functionals are developed based on an error function whose definition is motivated by the analysis of

function variation and error estimates for linear interpolation. This error function is different from error estimates in the sense that it is used to describe the local behavior of the error rather than to provide an estimate of its magnitude.

Let Ω and Ω_c be the (simply-connected) physical and computational domains in $\Re^n$, $n = 1, 2$, or 3, respectively, and denote the coordinates for them by $\boldsymbol{x}$ and $\boldsymbol{\xi}$. For a given function $u = u(\boldsymbol{x})$ (or its approximation), we seek a coordinate transformation $\boldsymbol{x} = \boldsymbol{x}(\boldsymbol{\xi}) : \Omega_c \to \Omega$ (or its inverse mapping $\boldsymbol{\xi} = \boldsymbol{\xi}(\boldsymbol{x}) : \Omega \to \Omega_c$) such that $\hat{u}(\boldsymbol{\xi}) \equiv u(\boldsymbol{x}(\boldsymbol{\xi}))$ is easier to approximate in Ω_c. To this end, we consider the linear element or the function variation over a differential segment $d\boldsymbol{\xi}$ in Ω_c

$$ds^2 \equiv (d\hat{u})^2 + d\boldsymbol{x}^T\, d\boldsymbol{x} = d\boldsymbol{\xi}^T\, \boldsymbol{J}^T [I + \nabla u \nabla u^T]\, \boldsymbol{J}\, d\boldsymbol{\xi}, \tag{1}$$

where $\boldsymbol{J} = (\partial \boldsymbol{x})/(\partial \boldsymbol{\xi})$ is the Jacobian matrix of the coordinate transformation. Ideally, $\boldsymbol{x} = \boldsymbol{x}(\boldsymbol{\xi})$ should be chosen such that $\boldsymbol{J}^T [I + \nabla u \nabla u^T]\, \boldsymbol{J} = cI$, where I is the identity matrix and c is a constant. In this way, $\hat{u}$ has constant variation and can be well resolved on a computational mesh that is commonly chosen to be regular and uniform.

The other motivation for easy numerical approximation comes from interpolation error estimates. Let $\Pi_1 u$ be the linear interpolant of u at the vertices of an imaging mesh cell (say a triangle in two dimensions) and denote the error by $E_0(\boldsymbol{x}) = \Pi_1 u - u$. E_0 can then be expressed as a quadratic function locally, and its level surfaces form a family of ellipses with a common center $\boldsymbol{x}_c$, provided that the Hessian matrix of u, denoted by H, is positive definite. The geometric illustration of interpolation at the vertices of the mesh cell is that the circum-surface of the cell from this family of ellipses is the level surface of value zero. It is shown by D'Azevedo [6] that for $\boldsymbol{x}$ close to $\boldsymbol{x}_c$, E_0 can be written as

$$E_0(\boldsymbol{x}) \approx E_0(\boldsymbol{x}_c) - \frac{1}{2}\, d\boldsymbol{x}^T\, H\, d\boldsymbol{x},$$

where $d\boldsymbol{x} = \boldsymbol{x} - \boldsymbol{x}_c$, and H is calculated at $\boldsymbol{x} = \boldsymbol{x}_c$. Further, D'Azevedo and Simpson [7] show that the gradient of the linear interpolation error is given by

$$E_G(\boldsymbol{x}) \equiv \|\nabla(u - \Pi_1 u)\|_{l^2} \approx \sqrt{d\boldsymbol{x}^T H^T\, H\, d\boldsymbol{x}}.$$

Writing $d\boldsymbol{x} = \boldsymbol{J} d\boldsymbol{\xi}$ with $d\boldsymbol{\xi} = \boldsymbol{\xi} - \boldsymbol{\xi}_c$, we have

$$E_0(\boldsymbol{x}) \approx E_0(\boldsymbol{x}_c) - \frac{1}{2} d\boldsymbol{\xi}^T\, \boldsymbol{J}^T\, H\, \boldsymbol{J}\, d\boldsymbol{\xi},$$

$$E_G(\boldsymbol{x}) \approx \sqrt{d\boldsymbol{\xi}^T\, \boldsymbol{J}^T\, H^T\, H\, \boldsymbol{J}\, d\boldsymbol{\xi}}.$$

Once again, to resolve $\hat{u}$ by linear interpolation on a regular and uniform mesh in Ω_c, we should choose the coordinate transformation such that either $\boldsymbol{J}^T H \boldsymbol{J} = cI$ or $\boldsymbol{J}^T H^T H \boldsymbol{J} = cI$ for some constant c.

From the above formulae for the solution variation and linear interpolation error estimates, it is reasonable to assume that some error estimates can be characterized by a quadratic function. For this reason, we define a general error function as

$$E(\boldsymbol{x}) = \sqrt{d\boldsymbol{\xi}^T \boldsymbol{J}^T G \boldsymbol{J}\, d\boldsymbol{\xi}}, \tag{2}$$

where $\boldsymbol{\xi}_c$ is an arbitrary point in Ω_c, $d\boldsymbol{\xi} = \boldsymbol{\xi} - \boldsymbol{\xi}_c$, and $G = G(\boldsymbol{x})$ is an n by n symmetric and positive definite matrix that is prescribed by the user and referred to as the monitor function. Without causing confusion, we assume that, in (2) and hereafter, coordinates $\boldsymbol{x}$ and $\boldsymbol{\xi}$ are related by the coordinate transformation $\boldsymbol{x} = \boldsymbol{x}(\boldsymbol{\xi})$, and $\boldsymbol{J}^T G \boldsymbol{J}$ is calculated at $\boldsymbol{\xi}_c$. Our goal is then to find $\boldsymbol{x} = \boldsymbol{x}(\boldsymbol{\xi})$ such that

$$\boldsymbol{J}^T G \boldsymbol{J} = \frac{1}{c} I \quad \text{or} \quad \boldsymbol{J}^{-1} G^{-1} \boldsymbol{J}^{-T} = cI \tag{3}$$

for some positive constant c.

Unfortunately, (3) is unachievable in general. To see this, letting $G = M^T M$, we can rewrite (3) as

$$(\boldsymbol{J}^{-1} M^{-1})(\boldsymbol{J}^{-1} M^{-1})^T = cI.$$

It is easy to verify that

$$\boldsymbol{J}^{-1} M^{-1} = \sqrt{c} Q \tag{4}$$

for arbitrary orthogonal matrix Q. For the trivial case $G = I$ (no adaptation), (4) implies that the coordinate transformation is orthogonal irrespective of the physical and computational domains. Obviously, this is impossible. Hence, (3) can be satisfied only approximately.

To develop functionals that accommodate this criterion, we now replace it with two equivalent conditions that are easier to deal with. In fact, (3) is equivalent to requiring that the eigenvalues of $A \equiv \boldsymbol{J}^{-1} G^{-1} \boldsymbol{J}^{-T}$ are equal and the determinant is constant. Geometrically, these two conditions force the error function $E(\boldsymbol{x})$ to have an isotropic and uniform distribution. To explain this more clearly, expressing A in its eigen-decomposition $A = UDU^T$, where U is an orthogonal matrix and $D = \text{diag}(\lambda_1, \ldots, \lambda_n)$, we obtain the level surface as

$$(\boldsymbol{\xi} - \boldsymbol{\xi}_c)^T U D^{-1} U^T (\boldsymbol{\xi} - \boldsymbol{\xi}_c) = e^2 \quad \text{or} \quad \sum_i \left(\frac{\tilde{\xi}^i - \xi_c^i}{\sqrt{\lambda_i}} \right)^2 = e^2, \tag{5}$$

where $\tilde{\boldsymbol{\xi}} - \boldsymbol{\xi}_c = U^T (\boldsymbol{\xi} - \boldsymbol{\xi}_c)$ and e is a given error level. Then, the isotropy condition means that the ellipse (5) should be close to a sphere while the uniformity requirement implies that the volume of the corresponding ellipsoid should be constant with respect to location in Ω_c. Mathematically, we have

$$\text{Isotropy Criterion:} \quad \lambda_1 = \cdots = \lambda_n; \tag{6}$$

$$\text{Uniformity Criterion:} \quad \sqrt{\prod_i \lambda_i} = \text{constant}. \tag{7}$$

These are the criteria that will be used for constructing mesh adaptation functionals in the rest of the paper. Note that isotropy is a local property since it describes only the local behavior of $E(\boldsymbol{x})$, whereas uniformity is a global one because it restricts the change of the function from place to place.

It is remarked that Knupp and Robidoux [19] obtained an equation similar to (4) as a general conclusion rather than a criterion and used it to analyze a number of existing functionals, including the one for harmonic mappings.

3. ISOTROPIC ERROR DISTRIBUTION AND CONFORMITY

We now develop mesh adaptation functionals that satisfy the isotropy criterion approximately. Our basic tool is the well-known arithmetic-mean geometric-mean inequality. Its application to the eigenvalues of A gives rise to

$$\left(\prod_i \lambda_i\right)^{\frac{1}{n}} \le \frac{1}{n}\sum_i \lambda_i, \tag{8}$$

with equality if and only if $\lambda_1 = \cdots = \lambda_n$. The coordinate transformation that approximately satisfies (6) can thus be obtained by minimizing the difference

$$\frac{1}{n}\sum_i \lambda_i - \left(\prod_i \lambda_i\right)^{\frac{1}{n}}$$

or its variants.

Noticing that

$$\sum_i \lambda_i = \mathrm{tr}(A) = \sum_i (\nabla\xi^i)^T G^{-1}\nabla\xi^i,$$

$$\prod_i \lambda_i = \det(A) = \frac{1}{J^2 g},$$

where J and g are the determinants of $\boldsymbol{J}$ and G, respectively, we rewrite (8) as

$$n^n \det(A) \le (\mathrm{tr}(A))^n \tag{9}$$

or

$$\frac{n^{n/2}}{J\sqrt{g}} \le \left(\sum_i (\nabla\xi^i)^T G^{-1}\nabla\xi^i\right)^{n/2}. \tag{10}$$

Multiplication of (10) by $\sqrt{g}$ and integration over Ω gives rise to

$$n^{n/2}\int_{\Omega_c} d\boldsymbol{\xi} \le \int_\Omega \sqrt{g}\left(\sum_i (\nabla\xi^i)^T G^{-1}\nabla\xi^i\right)^{n/2} d\boldsymbol{x}.$$

Hence, the mesh adaptation functional *according to the isotropy criterion* is given by

$$I_{iso}[\boldsymbol{\xi}] = \frac{1}{2}\int_\Omega \sqrt{g}\left(\sum_i (\nabla\xi^i)^T G^{-1}\nabla\xi^i\right)^{n/2} dx. \tag{11}$$

Interestingly, functional (11) can also be derived from the concept of conformal norm that has been used in the context of the differential geometry to define a whole class of nonorthogonality measures for linear mappings; e.g., see [22]. Let $\Re^{n,n}$ be the space of

$n \times n$ real matrices. A norm N is said to be conformal if there exists a constant $\kappa_N > 0$ such that

$$\kappa_N \det(B) \le [N(B)]^n$$

for any $B \in \Re^{n,n}$ with $\det(B) > 0$, with equality if and only if B is a general orthogonal matrix. Here, a general orthogonal matrix is defined as the product of a scalar number times an orthogonal matrix. Examples of conformal norms include

$$N_p(B) = \left(\sum_{i=1}^{n} \left(\sum_{j=1}^{n} b_{ij}^2 \right)^{p/2} \right)^{1/p}$$

$$N_p^*(B) = \left(\sum_{j=1}^{n} \left(\sum_{i=1}^{n} b_{ij}^2 \right)^{p/2} \right)^{1/p}$$

for some integer $p \ge 1$. When $p = 2$, $\kappa_N = n^{n/2}$ and both $N_p(B)$ and $N_p^*(B)$ are equal to $\|B\|_F$, the Frobenius norm of B.

The definition of the conformal norm implies that the mapping that is closest to conformal mappings can be obtained by minimizing the difference $([N(B)]^n - \kappa_N \det(B))$ or its variants. Using the Frobenius norm, we obtain

$$n^{n/2} \det(B) \le \|B\|_F^n.$$

By letting $G = M^T M$ and taking $B = \boldsymbol{J}^{-1} M^{-1}$, inequality (9) follows from the relations $A = BB^T$, $\|B\|_F^2 = \mathrm{tr}(A)$, and $(\det(B))^2 = \det(A)$. Thus, the isotropy and conformity criteria lead to the same functional. In this sense, conformity and isotropy are equivalent.

We note that I_{iso} is constant in the case $n = 1$. This reflects the fact that in one dimension there is no need to impose the isotropy condition since there is only one eigen-direction. In two dimensions, I_{iso} is the energy functional for harmonic mappings. By construction, the harmonic mapping is now the closest one to conformal mappings under the given boundary correspondence. Interestingly, when $n \ge 3$, (11) is different from the currently used functionals, especially the energy of harmonic mappings and the functional studied in [5] that includes Winslow's functional as a special example.

Recall that isotropy is a local property which regulates the local behavior of the coordinate transformation. This local regulation seems to have two global impacts. First, the mapping stretch can vary from place to place in the computational domain. It means that a uniform computational mesh can sometimes fail to resolve a function $\hat{u} = u(\boldsymbol{x}(\boldsymbol{\xi}))$. The other impact is that the satisfaction of the isotropy requirement everywhere in Ω_c can lead to a very rigid mapping. Indeed, it is well known that conformal mappings are rigid in the sense that very few degrees of freedom are needed to specify them. Thus, a complete specification of the boundary correspondence, as usually done in variational mesh generation and adaptation, can adversely affect and compete with the satisfaction of the isotropy criterion. Unfortunately, in this competition, it appears that the isotropy criterion or mesh adaptation is always the loser since in most cases the satisfaction of boundary correspondence is mandatory. As a consequence, mesh adaptation is sacrificed, meaning that not enough mesh points are concentrated or the concentration is misplaced; e.g., see Fig. 4.2 of [5] and Fig. 2 in this article.

Another disadvantage of functional I_{iso} is that it cannot handle Winslow's type of monitor function $G = w(\boldsymbol{x})I$, which is already isotropic, in the sense that there is no mesh adaptation taking place with this function.

4. UNIFORM ERROR DISTRIBUTION AND EQUIDISTRIBUTION

The uniformity criterion (7) requires that the volume of the ellipsoid contained by the error level ellipse (5), or $\sqrt{\prod_i \lambda_i} = \sqrt{\det(A)} = 1/(J\sqrt{g})$, be constant. In other words, it is equivalent to

$$J\sqrt{g} = \text{constant}, \tag{12}$$

which, in fact, is a generalization of the (one-dimensional) equidistribution principle [8].

We need the following lemma for constructing the functional based on the equidistribution principle (12). The interested reader is referred to Hardy *et al.* [13] for its proof.

LEMMA 4.1. *Given a weight function $w(\boldsymbol{x})$ with $\int_\Omega w\,dx = 1$, define*

$$M_r(f) = \left(\int_\Omega w|f|^r\,d\boldsymbol{x} \right)^{1/r}$$

for arbitrary function f and real number r, with the limits that $M_0(f) = \exp(\int_\Omega w \log |f|\,dx)$ (geometric mean), $M_{+\infty} = \max|f|$, and $M_{-\infty} = \min|f|$. Then

$$M_r(f) < M_s(f) \tag{13}$$

for $-\infty \le r < s \le +\infty$ unless (a) $M_r(f) = M_s(f) = +\infty$, which can happen only if $r \ge 0$ or (b) $M_r(f) = M_s(f) = 0$, which can happen only if $s \le 0$ or (c) $f \equiv constant$.

To use this lemma, we take $w = \sqrt{g}$, $f = 1/(J\sqrt{g})$, $r = 1$, and $s = q$ for any real number $q > 1$. Then, the lemma gives rise to

$$\int_\Omega \frac{\sqrt{g}}{J\sqrt{g}}\,d\boldsymbol{x} = \int_{\Omega_c} d\boldsymbol{\xi} \le \left[\int_\Omega \frac{\sqrt{g}}{(J\sqrt{g})^q}\,d\boldsymbol{x} \right]^{1/q}, \tag{14}$$

with equality if and only if (12) holds. As in the last section, we conclude that given a boundary correspondence between Ω and Ω_c, the coordinate transformation that most closely satisfies (12) or the uniformity criterion can be obtained by minimizing the difference between both sides of (14). The mesh adaptation functional *according to the uniformity or equidistribution criterion* is thus obtained as

$$I_{ep}[\boldsymbol{\xi}] = \int_\Omega \frac{\sqrt{g}}{(J\sqrt{g})^q}\,d\boldsymbol{x}. \tag{15}$$

Taking $q = 2$, the functional becomes a least squares functional

$$I_{ep}[\boldsymbol{\xi}] = \int_\Omega \frac{\sqrt{g}}{(J\sqrt{g})^2}\,d\boldsymbol{x} = \int_\Omega \frac{1}{\sqrt{g}} \left[\det\left(\frac{\partial \boldsymbol{\xi}}{\partial \boldsymbol{x}} \right) \right]^2 d\boldsymbol{x}.$$

There exist a number of mesh generation and/or adaptation functionals related to the Jacobian J. For example, Brackbill and Saltzman [4] minimize the functional

$$\int_\Omega w(\boldsymbol{x}) J \, d\boldsymbol{x}$$

with intent that cells are small in the region where the given weight function w is large. Steinberg and Roache [23] show that the constrained discrete optimization problem

$$\min \sum_i \frac{x_i^2}{2w_i}, \quad \text{subject to } \sum x_i = 1$$

has the equidistributing solution $x_1/w_1 = \cdots = x_n/w_n$. Based on the understanding gained from this example and the analysis of a one-dimensional continuous functional, they propose to use the functional

$$\int_{\Omega_c} J^2 \, d\boldsymbol{\xi}$$

subject to the global implicit constraint

$$\int_{\Omega_c} J \, d\boldsymbol{\xi} = \int_\Omega d\boldsymbol{x},$$

intending to keep the cell volumes constant. These ideas of linking functionals to equidistribution and using global implicit constraints are carried on by Knupp and Robidoux [19]. In his recent work [1], Baines shows, using a discrete identity, that least squares minimization of the residual of the divergence of a vector field is equivalent to that of a least squares measure of equidistribution of this residual.

Equation (15) is different from the currently used functionals. Also, our method of using Lemma 4.1 to construct functionals associated with equidistribution is more straightforward and convincing. Further, the method can be used in any number of dimensions.

Unfortunately, functional I_{ep} is not coercive in two and higher dimensions. The existence and uniqueness of its minimizer are not guaranteed [11]. This makes I_{ep} hard to use in practice. In the next section, we will discuss a possible combination of functionals (11) and (15) to overcome this difficulty.

5. MESH ADAPTATION EQUATION

In the last two sections we have seen that neither functional (11) nor (15) can alone lead to a robust adaptive mesh method. On the other hand, this is not surprising since either of them represents only one side of criterion (3). Thus, it is necessary and natural to combine them together.

From (10) we have

$$n^{nq/2} \int_\Omega \frac{\sqrt{g}}{(J\sqrt{g})^q} \, d\boldsymbol{x} \le \int_\Omega \sqrt{g} \left(\sum_i (\nabla \xi^i)^T G^{-1} \nabla \xi^i \right)^{nq/2} d\boldsymbol{x}. \tag{16}$$

For a given value $\theta \in [0, 1]$, a balance of the differences between both sides of (16) and of (14) is

$$\theta\left[\int_\Omega \sqrt{g}\left(\sum_i (\nabla\xi^i)^T G^{-1}\nabla\xi^i\right)^{nq/2} d\boldsymbol{x} - n^{nq/2}\int_\Omega \frac{\sqrt{g}}{(J\sqrt{g})^q}\, d\boldsymbol{x}\right] + (1-\theta)n^{nq/2}\left[\int_\Omega \frac{\sqrt{g}}{(J\sqrt{g})^q}\, d\boldsymbol{x} - \left(\int_{\Omega_c} d\boldsymbol{\xi}\right)^q\right].$$

The first square bracket represents the isotropy requirement and the second is the uniformity or equidistribution requirement. Thus, by minimizing functional

$$I[\boldsymbol{\xi}] = \theta\int_\Omega \sqrt{g}\left(\sum_i (\nabla\xi^i)^T G^{-1}\nabla\xi^i\right)^{nq/2} d\boldsymbol{x} + (1-2\theta)n^{nq/2}\int_\Omega \frac{\sqrt{g}}{(J\sqrt{g})^q}\, d\boldsymbol{x}, \quad (17)$$

where $\theta \in [0, 1]$, we expect to find a coordinate transformation that accommodates the two requirements. Note that the two integrals on the right-hand side have the same dimension. In one dimension, $I[\boldsymbol{\xi}]$ has the form

$$I[\xi] = (1-\theta)\int_\Omega \sqrt{g}\left(\frac{1}{\sqrt{g}}\frac{\partial\xi}{\partial x}\right)^q dx.$$

Regarding well posedness, we first note that for $nq/2 \geq 1$, the first part of $I[\boldsymbol{\xi}]$ is convex, and the existence, uniqueness, and the maximum principle for its minimizer are guaranteed; e.g., see [11, 22]. It is unknown to us if this result can also apply to the whole functional. But, one can easily see that $I[\boldsymbol{\xi}]$ is coercive if $\theta \in (0, 1/2]$. Moreover, when $\theta = 1/2$, only the first part of the functional remains, viz., $I[\boldsymbol{\xi}]$ becomes

$$I_q[\boldsymbol{\xi}] = \frac{1}{2}\int_\Omega \sqrt{g}\left(\sum_i (\nabla\xi^i)^T G^{-1}\nabla\xi^i\right)^{nq/2} d\boldsymbol{x}. \quad (18)$$

Thus, it is reasonable to conjecture that the minimizer of $I[\boldsymbol{\xi}]$ exists for $\theta \in (0, 1/2]$. For this reason, we will consider values of θ only within the range $(0, 1/2]$. It is also interesting to mention that functional I_{iso} can be obtained from (17) by simply taking $q = 1$ irrespective of θ.

We now derive the Euler–Lagrange equation for functional $I[\boldsymbol{\xi}]$ for $q \geq 1$. For simplicity, (17) is rewritten as

$$I[\boldsymbol{\xi}] = \theta\int_\Omega \left(\sum_i (\nabla\xi^i)^T \bar{G}^{-1}\nabla\xi^i\right)^\gamma d\boldsymbol{x} + (1-2\theta)n^\gamma \int_\Omega \frac{\sqrt{g}}{(J\sqrt{g})^q}\, d\boldsymbol{x}, \quad (19)$$

where

$$\bar{G} = \frac{1}{g^{1/(2\gamma)}}G, \quad \gamma = \frac{nq}{2}.$$

Denote $\beta = \sum_i (\nabla\xi^i)^T \bar{G}^{-1}\nabla\xi^i$. For the purpose of well posedness, it is assumed that $n \geq 1$ and $q \geq 1$ are chosen such that $\gamma \geq 1$. With this notation, the Euler–Lagrange equation can

be written as

$$\nabla \cdot \left[\theta \gamma \beta^{\gamma-1} \bar{G}^{-1} \nabla \xi^i + \frac{(1-2\theta) q n^\gamma \sqrt{g}}{2} \left(\frac{1}{J\sqrt{g}} \right)^q \frac{\partial \boldsymbol{x}}{\partial \xi^i} \right] = 0, \quad i = 1, \ldots, n. \tag{20}$$

Practically, it is more convenient to compute $\boldsymbol{x} = \boldsymbol{x}(\boldsymbol{\xi})$ instead of its inverse $\boldsymbol{\xi} = \boldsymbol{\xi}(\boldsymbol{x})$. Interchanging the roles of the dependent and independent variables, we have the conservative form

$$\sum_j \frac{\partial}{\partial \xi^j} J (\boldsymbol{a}^i)^T \left[\theta \gamma \beta^{\gamma-1} \bar{G}^{-1} \nabla \xi^i + \frac{(1-2\theta) q n^\gamma \sqrt{g}}{2} \left(\frac{1}{J\sqrt{g}} \right)^q \frac{\partial \boldsymbol{x}}{\partial \xi^i} \right] = 0, \quad i = 1, \ldots, n \tag{21}$$

and the nonconservative form

$$\begin{aligned} &\theta \left[\sum_{ij} ((\boldsymbol{a}^i)^T \bar{G}^{-1} \boldsymbol{a}^j) \frac{\partial^2 \boldsymbol{x}}{\partial \xi^i \partial \xi^j} - \sum_i \left((\boldsymbol{a}^i)^T \sum_j \frac{\partial (\bar{G}^{-1})}{\partial \xi^j} \boldsymbol{a}^j \right) \frac{\partial \boldsymbol{x}}{\partial \xi^i} \right] \\ &\quad + \frac{\theta(\gamma-1)}{\beta} \left[2 \sum_{ij} \left((\bar{G}^{-1} \boldsymbol{a}^i)(\bar{G}^{-1} \boldsymbol{a}^j)^T \sum_k \boldsymbol{a}^k (\boldsymbol{a}^k)^T \right) \frac{\partial^2 \boldsymbol{x}}{\partial \xi^i \partial \xi^j} \right. \\ &\quad \left. - \sum_i \left(\sum_j \left((\boldsymbol{a}^i)^T \bar{G}^{-1} \boldsymbol{a}^j \sum_k (\boldsymbol{a}^k)^T \frac{\partial (\bar{G}^{-1})}{\partial \xi^j} \boldsymbol{a}^k \right) \right) \frac{\partial \boldsymbol{x}}{\partial \xi^i} \right] + \frac{(1-2\theta) q (q-1) n^\gamma \sqrt{g}}{\gamma \beta^{\gamma-1} (J\sqrt{g})^q} \\ &\quad \times \left[\sum_{ij} (\boldsymbol{a}_i (\boldsymbol{a}^j)^T) \frac{\partial^2 \boldsymbol{x}}{\partial \xi^i \partial \xi^j} + \sum_i \left(\frac{1}{\sqrt{g}} \frac{\partial \sqrt{g}}{\partial \xi^i} \right) \frac{\partial \boldsymbol{x}}{\partial \xi^i} \right] = 0, \end{aligned} \tag{22}$$

where $\boldsymbol{a}_i \equiv (\partial \boldsymbol{x})/(\partial \xi^i)$ and $\boldsymbol{a}^i \equiv \nabla \xi^i$ are the covariant and contravariant base vectors that are related by

$$\boldsymbol{a}^i = \frac{1}{J} \boldsymbol{a}_j \times \boldsymbol{a}_k \quad \text{with } (i, j, k) \text{ cyclic.}$$

6. NUMERICAL EXPERIMENTS

To demonstrate various features of the developed mesh adaptation functional, in this section we present some two-dimensional numerical results obtained mainly for function

$$u(x, y) = e^{-100((x-0.5)^2 + (y-0.5)^2)} \tag{23}$$

defined in the unit square. Ω_c is chosen to be the unit square. Since our purpose here is to explore the features of functional (17), we use a uniform boundary correspondence between Ω and Ω_c and no smoothing of the monitor function in our computations. But, we would like to emphasize that in practice, an adaptive boundary correspondence and a few sweeps of a low-pass filter for smoothing the monitor function are often necessary, and sometimes can be crucial, for accuracy and efficiency of variational methods; e.g., see [14].

The mesh equation (22) is discretized with central finite differences and solved using the moving mesh PDE approach [15]. With this approach, a derivative $(\partial \boldsymbol{x})/(\partial t)$ with respect to pseudo-time t is added to the mesh equation (22) and the resultant parabolic system is integrated using a modified backward Euler scheme with which the coefficients of terms $(\partial \boldsymbol{x})/(\partial \xi^i)$ and $(\partial^2 \boldsymbol{x})/(\partial \xi^i \partial \xi^j)$ are calculated at the previous time level. The

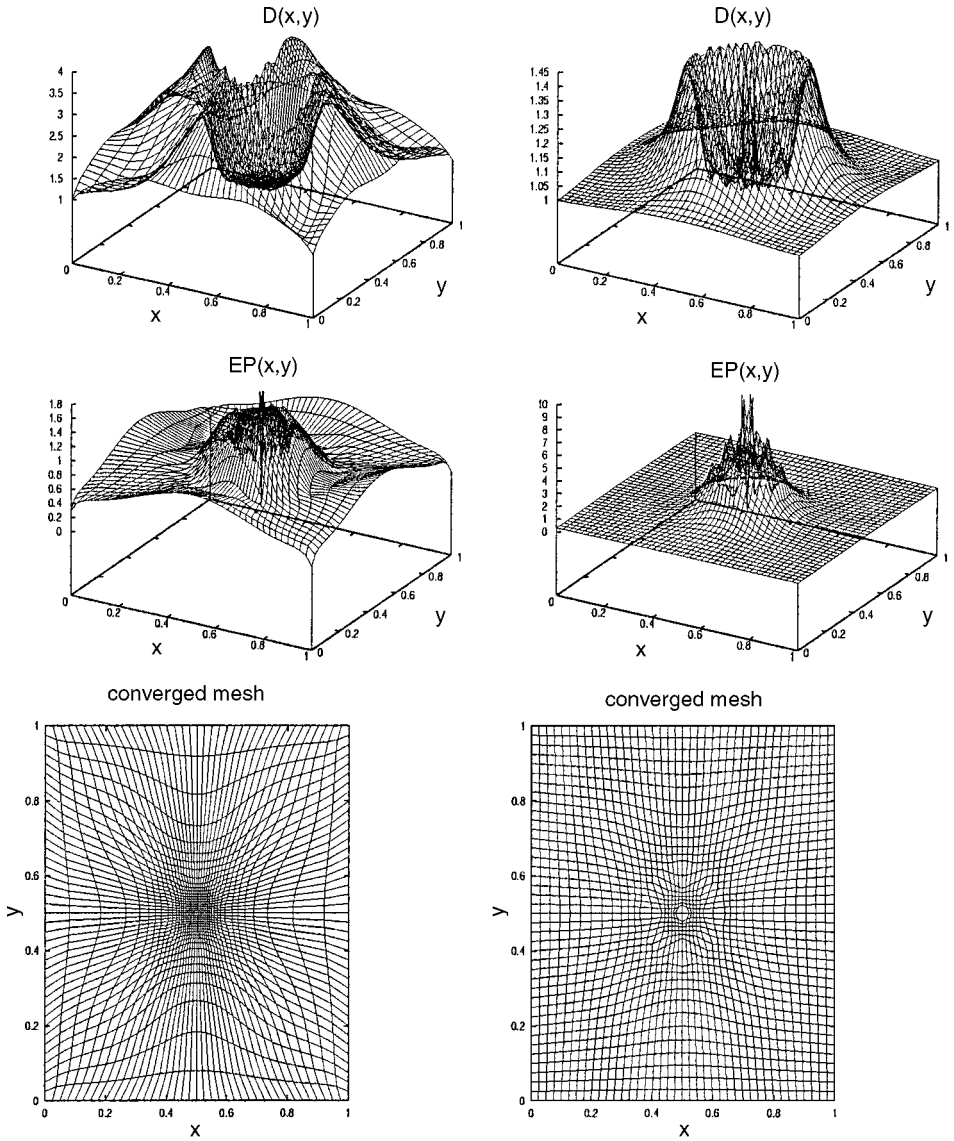

FIG. 1. Results obtained with monitor function $G_1 = I + |H|$. The left and right columns correspond to Case A: $(q, \theta) = (2, 0.1)$ and Case B: $(q, \theta) = (2, 0.5)$, respectively.

nonlinear algebraic system is solved using a preconditioned conjugate gradient method. The converged mesh is obtained when the root-mean-square norm of the residual is less than 10^{-6}. All computations start with a uniform mesh of size 41×41.

In the computations, the following four monitor functions are used,

$$G_1 = I + |H|,$$
$$G_2 = \sqrt{\det(G_1)}I,$$
$$G_3 = I + \nabla u(\nabla u)^T,$$
$$G_4 = \sqrt{\det(G_3)}I,$$

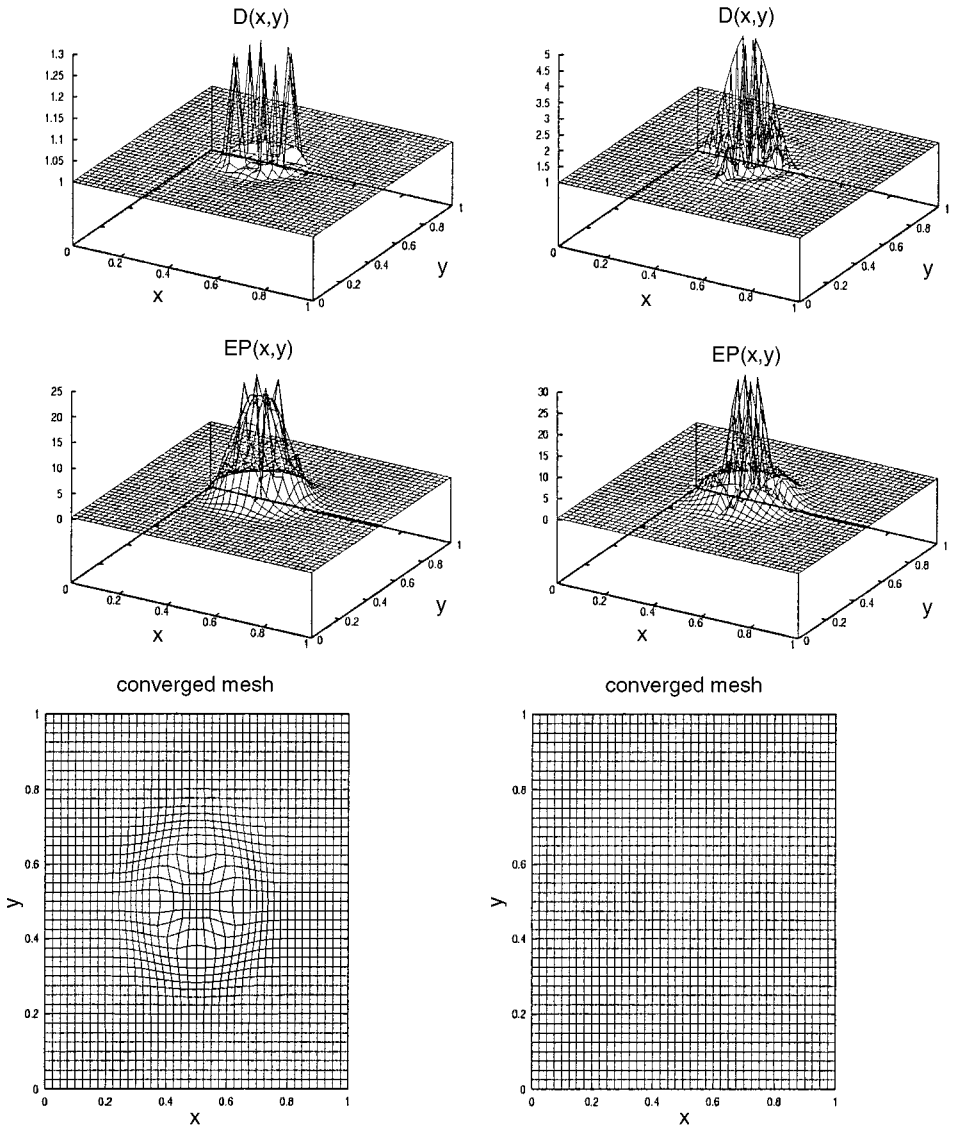

FIG. 2. (Continued from Fig. 1) Results obtained with monitor function $G_1 = I + |H|$. The left and right columns correspond to Case C: $q = 1$ (harmonic mapping) and the uniform mesh case, respectively.

where $|H| = V\text{diag}(|\mu_1|, \ldots, |\mu_n|)V^T$, assuming that the eigen-decomposition of the Hessian matrix of u, H, is given by $V\text{diag}(\mu_1, \ldots, \mu_n)V^T$. The definition of G_1 is based on the error estimate E_0 of linear interpolation. G_3 is the commonly used arc-length monitor function. G_2 and G_4 are Winslow's type monitor functions associated with the first- and second-order derivatives, respectively. Once again, we use such simply defined monitor functions only for the purpose of exploring the features of $I[\xi]$. In practice, accuracy can often be gained significantly by introducing the intensity parameter in the monitor function to control the mesh concentration. For example, G_1 can be modified as

$$G_1 = I + \alpha|H|.$$

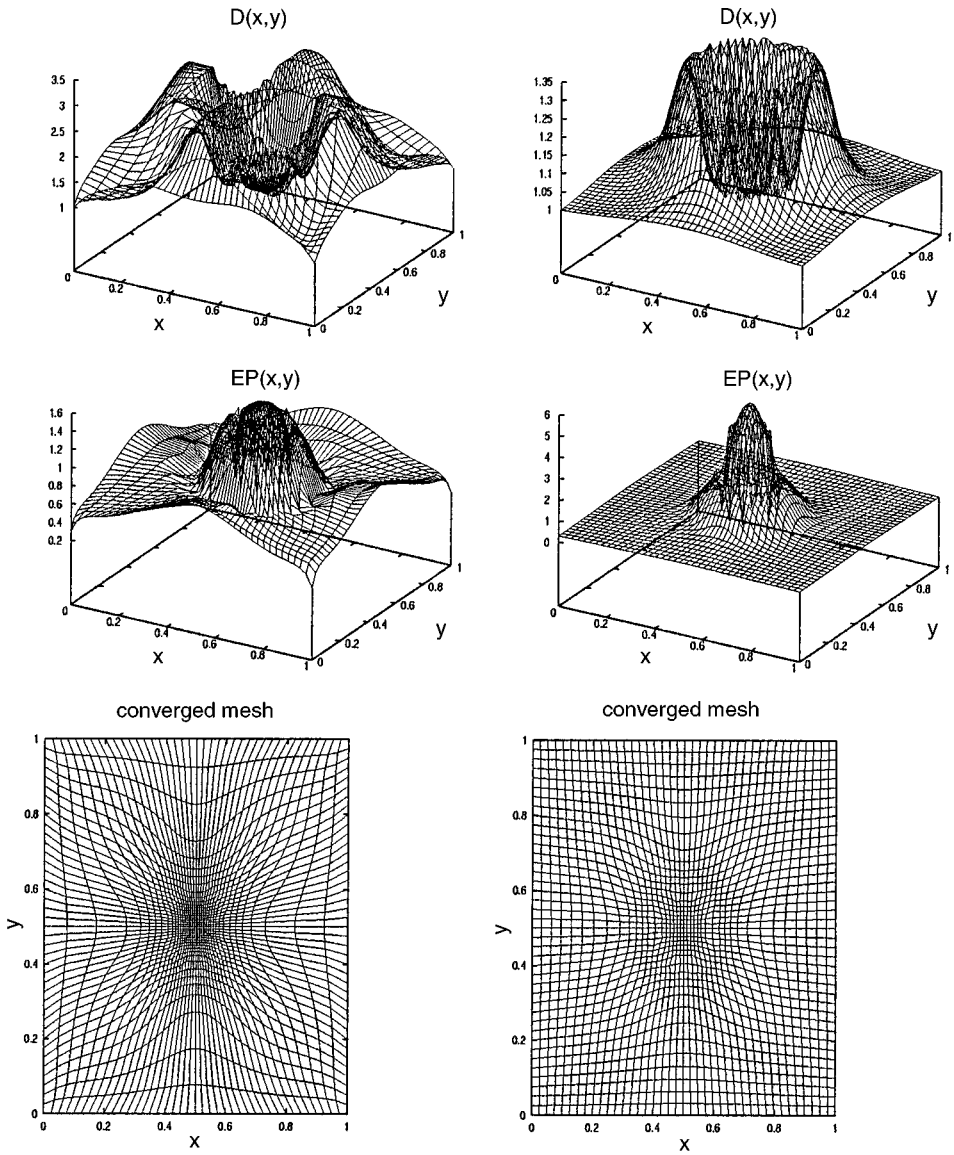

FIG. 3. Results obtained with monitor function $G_2 = \sqrt{\det(I + |H|)}I$. The left and right columns correspond to Case A: $(q, \theta) = (2, 0.1)$ and Case B: $(q, \theta) = (2, 0.5)$, respectively.

The larger α is, the more intensive the mesh adaptation can be. A proper choice of α will often lead to better accuracy. See [2, 14] for the automatic choice of this parameter in one and two dimensions.

It is noted that the monitor function can be chosen as $G = I + H^T H$ or $G = I + \sqrt{H^T H}$ according to the gradient error estimate of linear interpolation given in Section 2. While they do lead to slightly better results, these functions do not offer any new feature in the resulting meshes other than those given by G_1. For this reason and for saving space, we do not present the results obtained with these monitor functions.

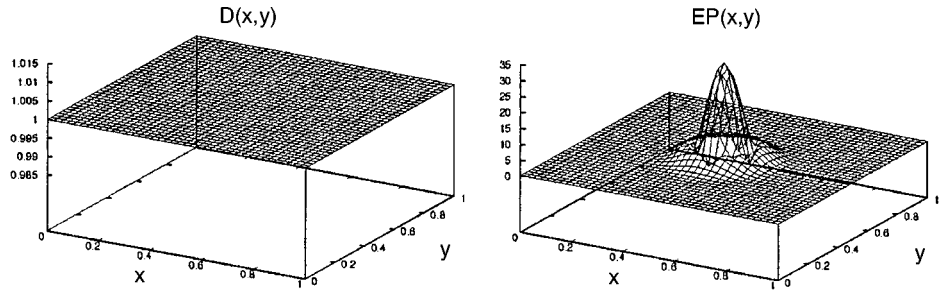

FIG. 4. (Continued from Fig. 3) Results obtained with monitor function $G_2 = \sqrt{\det(I + |H|)}I$ for the uniform mesh case.

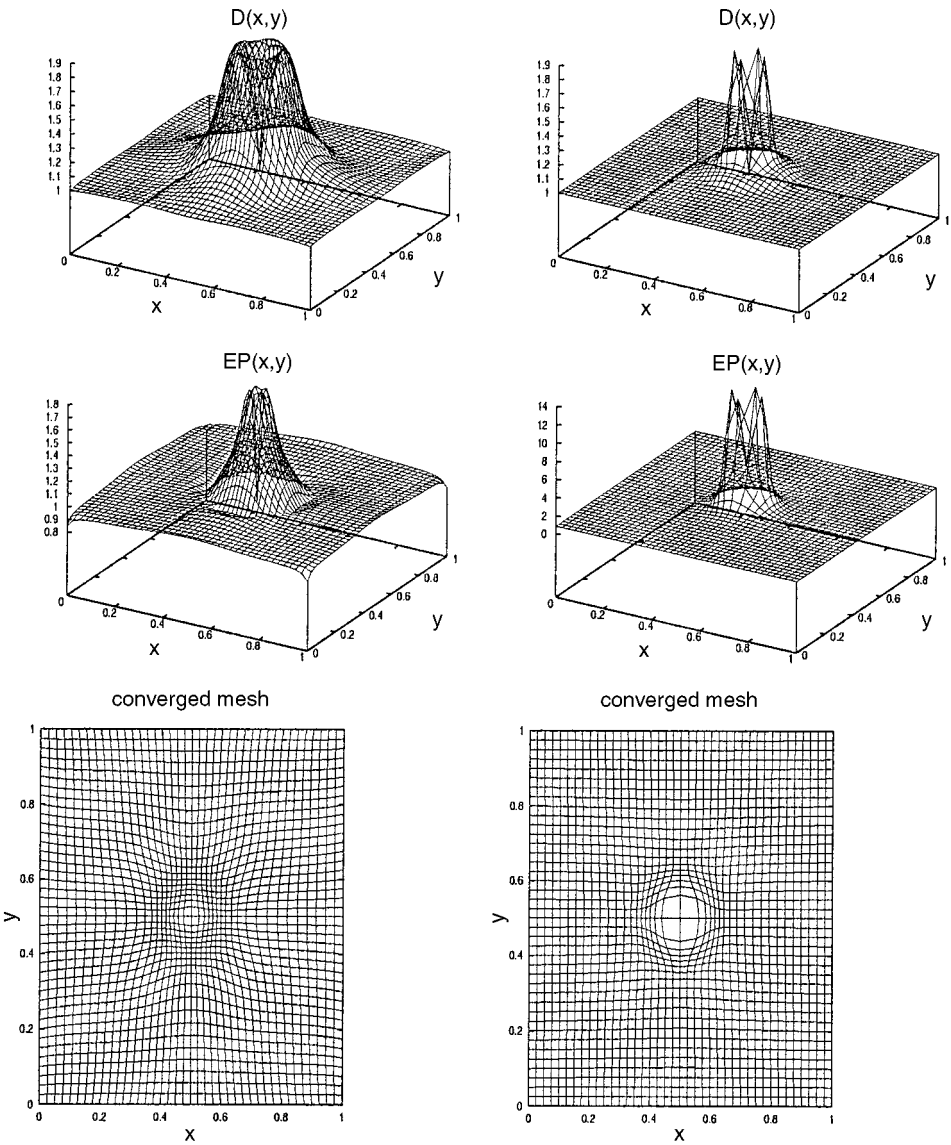

FIG. 5. Results obtained with monitor function $G_3 = I + \nabla u(\nabla u)^T$. The left and right columns correspond to Case A: $(q, \theta) = (2, 0.1)$ and Case B: $(q, \theta) = (2, 0.5)$, respectively.

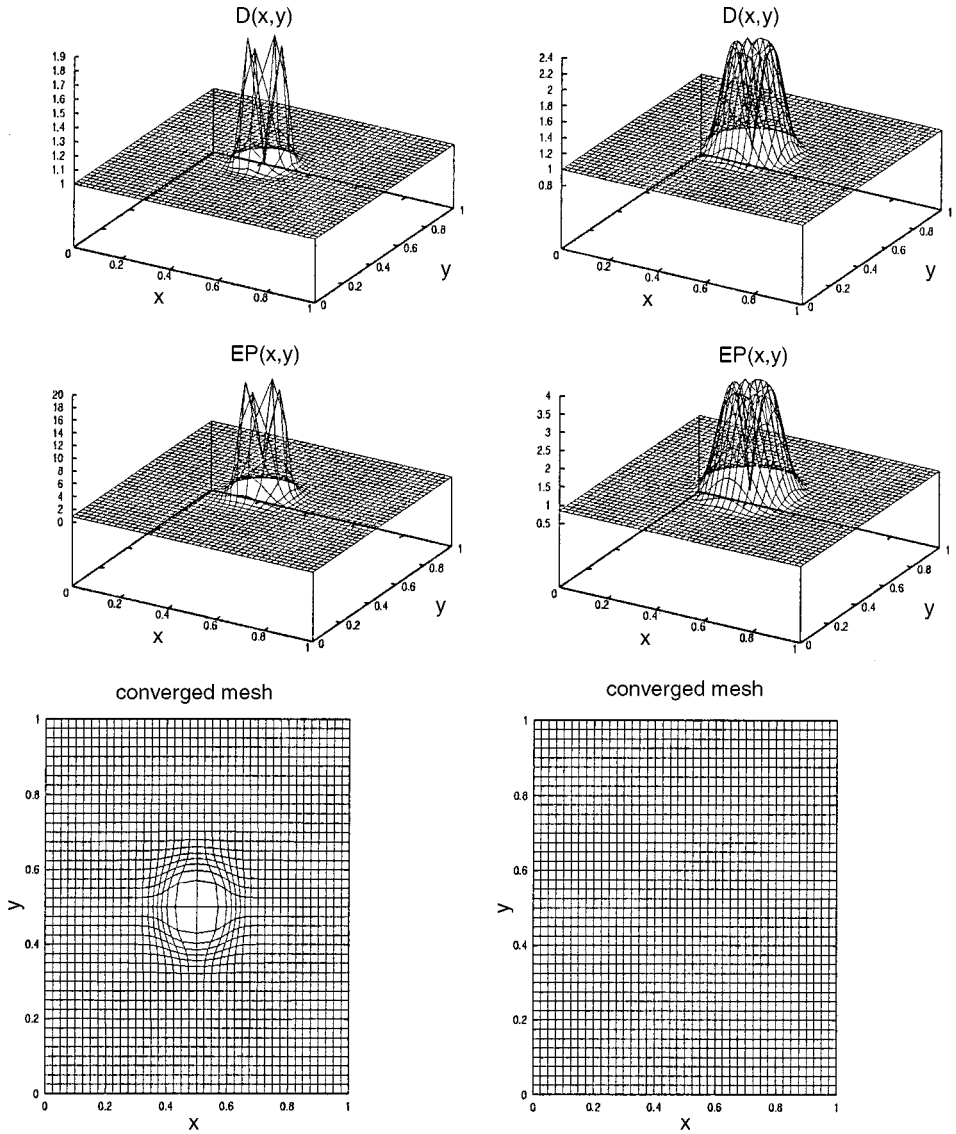

FIG. 6. (Continued from Fig. 5) Results obtained with monitor function $G_3 = I + \nabla u(\nabla u)^T$. The left and right columns correspond to Case C: $q = 1$ (harmonic mapping) and the uniform mesh case, respectively.

The numerical results will be given for the functions

$$D(\boldsymbol{x}) \equiv \frac{\mathrm{tr}(A)^{n/2}}{n^{n/2}\sqrt{\det(A)}}, \quad A = \boldsymbol{J}^{-1}G^{-1}\boldsymbol{J}^{-T}, \tag{24}$$

where G equals one of the functions G_1, G_2, G_3, and G_4, which measures the deviation from conformity and

$$EP(\boldsymbol{x}) = \frac{J\sqrt{g}}{c}, \tag{25}$$

where $c = (1/|\Omega_c|)\int_\Omega \sqrt{g}\, d\boldsymbol{x}$, that measures the deviation from equidistribution. When

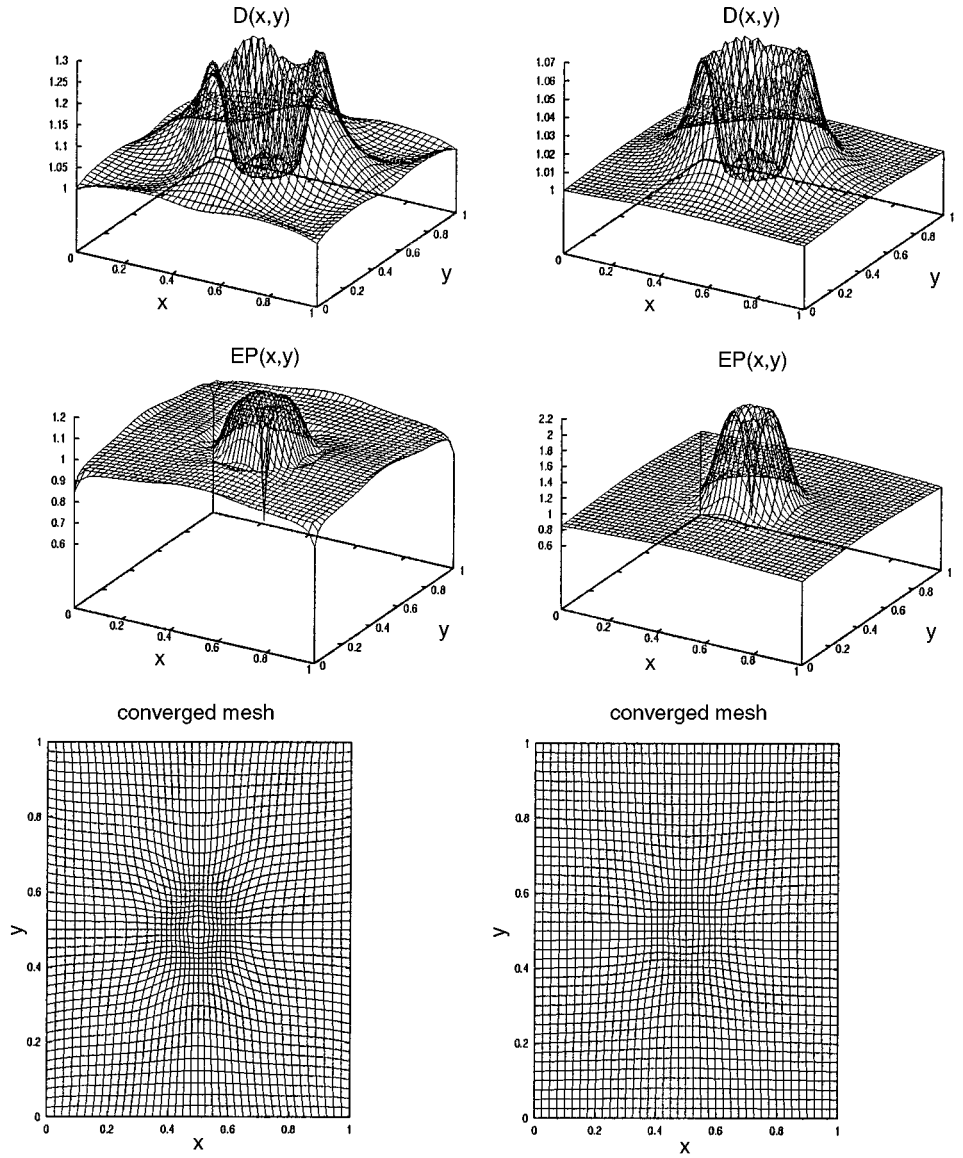

FIG. 7. Results obtained with monitor function $G_4 = \sqrt{\det(I + \nabla u(\nabla u)^T)}I$. The left and right columns correspond to Case A: $(q, \theta) = (2, 0.1)$ and Case B: $(q, \theta) = (2, 0.5)$, respectively.

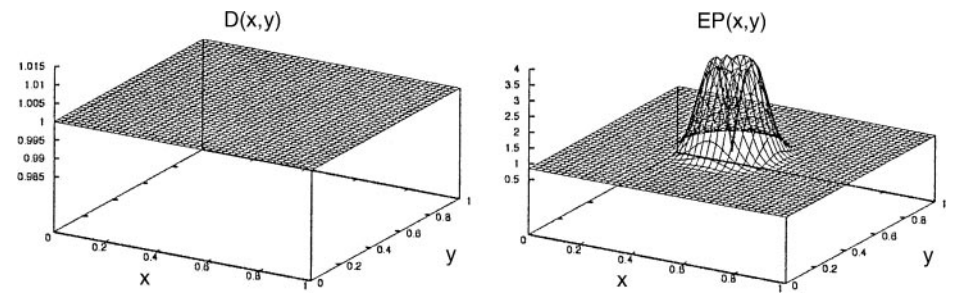

FIG. 8. (Continued from Fig. 7) Results obtained with monitor function $G_4 = \sqrt{\det(I + \nabla u(\nabla u)^T)}I$ for the uniform mesh case.

TABLE I
The Maximum Error of Linear Interpolation on the Obtained Converged Meshes

Monitor function	Case	q	θ	$\|e\|_\infty$
G_1	A	2	0.1	9.12e-3
	B	2	0.5	2.73e-2
	C	1		4.20e-2
G_2	A	2	0.1	8.90e-3
	B	2	0.5	9.71e-3
G_3	A	2	0.1	3.11e-2
	B	2	0.5	1.43e-1
	C	1		1.71e-1
G_4	A	2	0.1	2.19e-2
	B	2	0.5	2.60e-2
Uniform mesh				8.83e-1

$D(\boldsymbol{x}) \equiv 1$, the mapping becomes conformal while the equidistribution relation (12) holds if $EP(\boldsymbol{x}) \equiv 1$.

These two functions and the converged mesh obtained with monitor functions G_1, G_2, G_3, and G_4 are plotted in Figs. 1–8 for Case A: $(q, \theta) = (2, 0.1)$, Case B: $(q, \theta) = (2, 0.5)$, and Case C: $q = 1$ (the harmonic mapping case). For comparison, we also plot $D(x, y)$ and $EP(x, y)$ obtained on a uniform mesh. The maximum linear interpolation error is listed in Table I. From these results, the following observations can be made:

(a) Case C has the smallest deviation from conformity, followed by Cases B and A.

(b) On the other hand, Case A has the smallest deviation from equidistribution, the highest degree of mesh concentration, and the smallest interpolation error, followed by Case B and Case C. In particular, for Case C there are not enough mesh points concentrated in the central area and this leads to low-accuracy resolution, as shown in Figs. 2 and 6. These results are compatible with the construction, i.e., the smaller θ, the more closely the equidistribution relation (12) is satisfied and the higher degree of adaptation results.

TABLE II
The Maximum Error of Linear Interpolation on the Obtained Converged Meshes for the Second Example

Monitor function	Case	q	θ	$\|e\|_\infty$
G_1	A	2	0.1	1.23e-2
	B	2	0.5	1.75e-2
	C	1		5.76e-2
G_2	A	2	0.1	1.64e-2
	B	2	0.5	1.94e-2
G_3	A	2	0.1	1.47e-2
	B	2	0.5	8.50e-2
	C	1		1.19e-1
G_4	A	2	0.1	2.45e-2
	B	2	0.5	1.93e-2
Uniform mesh				1.95

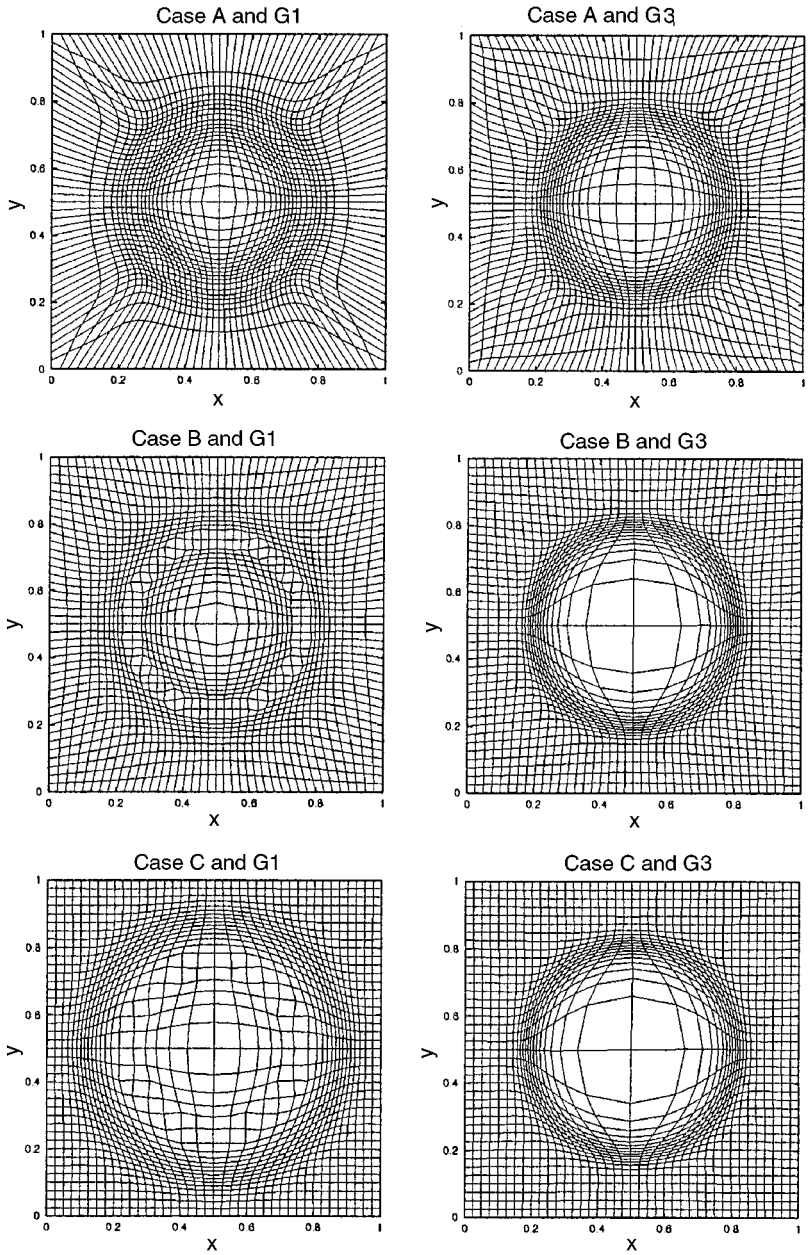

FIG. 9. Adaptive meshes obtained for the second example with monitor functions G_1 and G_3. The left column corresponds to G_1 and the right one is for G_3. The first, second, and third rows correspond to Case A, Case B, and Case C (harmonic mapping case), respectively.

Moreover, they indicate that a mesh adaptation functional should have a certain degree of equidistribution in order to produce reasonably accurate results.

(c) As mentioned in the last paragraph of Section 4, the isotropy functional I_{iso} that results in a harmonic mapping in two dimensions produces no mesh adaptation for Winslow's type monitor functions G_2 and G_4. In contrast, $I[\xi]$ with $q > 1$ works well for these monitor functions, as may be seen in Figs. 3 and 7.

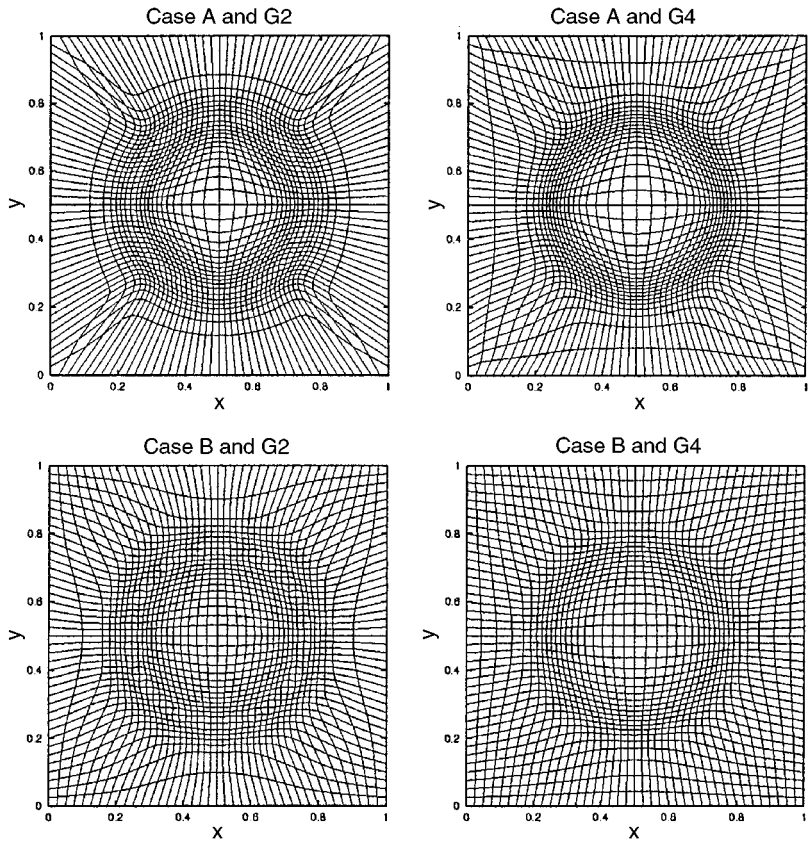

FIG. 10. (Continued from Fig. 9) Adaptive meshes obtained for the second example with monitor functions G_2 and G_4. The left column corresponds to G_2 and the right one is for G_4. The first and second rows correspond to Case A and Case B, respectively.

(d) As expected, the monitor functions based on the Hessian matrix lead to better accuracy than the arc-length monitor functions. Also, G_2 and G_4 produce more accurate results for this example but worse results in the next one (see Table II) than G_1 and G_3, respectively.

(e) Finally, all the adaptive meshes give significantly better results than a uniform mesh does with the same number of nodes.

We also show the adaptive meshes in Figs. 9 and 10 and the maximum error in Table II for the second example

$$u(x, y) = \tanh(30((x - 0.5)^2 + (y - 0.5)^2 - 1/16)).$$

The obtained results confirm the above observations.

7. CONCLUSIONS AND COMMENTS

Several criteria for mesh adaptation have been developed based on an error function whose definition is motivated by function variation and error estimates for linear interpolation. In

particular, the isotropy and uniformity criteria, (6) and (7), are shown to correspond to the practical regularity and uniformity properties of a computational mesh, respectively. It is also shown that isotropy is equivalent to conformity while uniformity is equivalent to equidistribution from the mesh adaptation point of view.

Functionals I_{iso} and I_{ep} that respectively accommodate the isotropy and equidistribution requirements are constructed using discrete and continuous inequalities. Functional $I[\xi]$ that compromises these conditions is formulated by naturally combining I_{iso} and I_{ep}. Two parameters $q \geq 1$ and $\theta \in (0, 0.5]$ are involved in the formulation. When $q = 1$, $I[\xi]$ becomes I_{iso}, which leads exactly to the widely used functional for harmonic mappings in two dimensions. When $\theta = 0.5$, $I[\xi]$ gives rise to a functional that is known to have a unique minimizer. The equidistribution functional I_{ep} can be obtained by taking $\theta = 0$. Unlike many existing mesh adaptation functionals, the geometric meaning of minimization of the developed functional $I[\xi]$ is clear by construction. That is, the smaller the value of θ, the more closely the equidistribution is satisfied and the higher degree of mesh adaptation is achieved. On the other hand, the bigger θ is or the closer to 1 the value of q, the more regular or conformal and the less adaptive the mesh. These results also provide a better understanding of the increasingly popular method of harmonic mapping in two dimensions.

The numerical results have been presented to demonstrate the features of $I[\xi]$. Particularly, they indicate that a mesh adaptation functional should have a certain degree of equidistribution in order to produce reasonably accurate solutions. Our limited experience shows that the choices for the values of q and θ are not crucial. Generally, (and as done in our numerical example), the choice of $q = 2$ and $0.1 \leq \theta \leq 0.5$ seems to work well. θ cannot be taken too close to zero, otherwise $I[\xi]$ will become nonconvex, and its minimization problem will be difficult to solve.

The presented analysis and results can be used in two ways to define a proper monitor function with the developed functional for practical problems. The simple way is through the generalized equidistribution principle (12). With it, one can choose a monitor function of Winslow's type with the weight function being large in the area where higher mesh concentration is desired. An example is to take the weight function as an estimate of the error density function. The other is to use the error function (2). For example, when considering the function variation, as shown in Section 2, we can choose $G = G_3 \equiv I + \nabla u \nabla u^T$ or $G = G_4 \equiv \sqrt{\det(G_3)}$ if a monitor function of Winslow's type is preferred. In the meantime, when the error of linear interpolation is concerned, $G = G_1 \equiv I + |H|$ or $G = G_2 \equiv \sqrt{\det(G_1)}$ will be the choice. It is worth mentioning that with the variational approach developed in this article, it is also possible to define the monitor function based directly on error estimates. A study related to this topic is currently underway.

Finally, we remark that the isotropy functional I_{iso} (11) leads to harmonic mappings in two dimensions but different ones in three dimensions. This may not be a drawback because it is unclear whether or not three dimensional harmonic mappings are invertible even if the target space (i.e., the computational domain in mesh adaptation) is Euclidean and has a convex boundary; e.g., see [20]. On the other hand, neither is it clear if a minimizer exists for the developed functional $I[\xi]$. However, if a minimizer does exist for a smaller value of θ, the analysis and numerical results given in the preceding sections suggest that the resultant coordinate transformation satisfy an approximate equidistribution

relation

$$C_1 \leq J\sqrt{g} \leq C_2$$

with some positive constants C_1 and C_2. As an immediate consequence, the Jacobian J will not vanish and the coordinate transformation is at least locally nonsingular.

ACKNOWLEDGMENTS

The author is grateful to David M. Sloan for his encouragement and useful comments on this work. This work was supported in part by the NSF under grant DMS-0074240.

REFERENCES

1. M. J. Baines, Least squares and approximate equidistribution in multidimensions, *Numer. Meth. P.D.E.* **15**, 605 (1999).
2. G. Beckett and J. A. Mackenzie, Convergence analysis of finite-difference approximations on equidistributed grids to a singularly perturbed boundary value problems, *J. Comput. Appl. Math.* **35**, 109 (2000).
3. J. U. Brackbill, An adaptive grid with directional control, *J. Comput. Phys.* **108**, 38 (1993).
4. J. U. Brackbill and J. S. Saltzman, Adaptive zoning for singular problems in two dimensions, *J. Comput. Phys.* **46**, 342 (1982).
5. W. Cao, W. Huang, and R. D. Russell, A study of monitor functions for two dimensional adaptive mesh generation, *SIAM J. Sci. Comput.* **20**, 1978 (1999).
6. E. F. D'Azevedo, Optimal triangular mesh generation by coordinate transformation, *SIAM J. Sci. Stat. Comput.* **12**, 755 (1991).
7. E. F. D'Azevedo and R. B. Simpson, On optimal triangular meshes for minimizing the gradient error, *Numer. Math.* **59**, 321 (1991).
8. C. de Boor, Good approximation by splines with variables knots ii. In *Springer Lecture Notes Series 363* (Springer-Verlag, Berlin, 1973).
9. A. S. Dvinsky, Adaptive grid generation from harmonic maps on Riemannian manifolds, *J. Comput. Phys.* **95**, 450 (1991).
10. J. E. Castillo (Ed.), *Mathematics Aspects of Numerical Grid Generation.* (Soc. for Industr. & Appl. Math., Philadelphia, 1991).
11. L. C. Evans, *Partial Differential Equations*, Graduate Studies in Mathematics (American Mathematical Society, Providence, RI, 1998), Vol. 19.
12. R. Hagmeeijer, Grid adaption based on modified anisotropic diffusion equations formulated in the parametric domain, *J. Comput. Phys.* **115**, 169 (1994).
13. G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities* (Cambridge Univ. Press, Cambridge, UK, 1934).
14. W. Huang, Practical aspects of formulation and solution of moving mesh partial differential equations. *J. Comput. Phys.* **171**, 753 (2001).
15. W. Huang and R. D. Russell, A high dimensional moving mesh strategy, *Appl. Numer. Math.* **26**, 63 (1997).
16. P. Knupp, Mesh generation using vector-fields, *J. Comput. Phys.* **119**, 142 (1995).
17. P. Knupp and S. Steinberg, *Fundationals of Grid Generation* (CRC Press, Boca Raton, FL, 1994).
18. P. M. Knupp, Jacobian-weighted elliptic grid generation, *SIAM J. Sci. Comput.* **17**, 1475 (1996).
19. P. M. Knupp and N. Robidoux, A framework for variational grid generation: conditioning the jacobian matrix with matrix norms, *SIAM J. Sci. Comput.* **21**, 2029 (2000).
20. G. Liao, On harmonic maps. In *Mathematics Aspects of Numerical Grid Generation*, edited by J. E. Castillo, (Soc. for Industr. & Appl. Math., Philadelphia, 1991), pp. 123–130.

21. V. D. Liseikin, *Grid Generation Methods* (Springer-Verlag, Berlin, 1999).
22. Yu. G. Reshetnyak, *Space Mappings with Bounded Distortion*, Translation of Mathematical Monographs (American Mathematical Society, Providence, RI, 1989), Vol. 73.
23. S. Steinberg and P. J. Roache, Variational grid generation, *Numer. Meth. P.D.E.* **2**, 71 (1986).
24. J. F. Thompson, Z. A. Warsi, and C. W. Mastin, *Numerical Grid Generation: Foundations and Applications* (North-Holland, New York, 1985).
25. A. Winslow, Numerical solution of the quasi-linear Poisson equation in a nonuniform triangle mesh, *J. Comput. Phys.* **1**, 149 (1967).