國立交通大學

應用數學系

數學建模與科學計算碩士班

碩 士 論 文

Monge-Ampére 方程的數值方法與其在非成像光學上的應用

Numerical Studies on the Monge-Ampére Equation and its Nonimage Optical Application

研 究 生：蔡玉麟

指導教授：吳金典 教授

中 華 民 國 九 十 九 年 七 月

Monge-Ampére 方程的數值方法與其在非成像光學上的應用
Numerical Studies on the Monge-Ampére Equation and its Nonimage Optical Application

研 究 生：蔡玉麟　　Student：Yu-Lin Tsai

指導教授：吳金典　　Advisor：Chin-Tien Wu

國 立 交 通 大 學
應用數學系數學建模與科學計算碩士班

碩 士 論 文

A Thesis

Submitted to Institute of Mathematical Modeling and Scientific Computing

Department of Applied Mathematics

College of Science

National Chiao Tung University

in partial Fulfillment of the Requirements

for the Degree of

Master

in

Mathematical Modeling and Scientific Computing

July 2010

Hsinchu, Taiwan, Republic of China

中華民國九十九年七月

Monge-Ampére 方程的數值方法與其在非成像光學上的應用

學生：蔡玉麟　　　　　　　　　　　　　　　　　　　　　指導教授：吳金典

國立交通大學應用數學系數學建模與科學計算碩士班

摘　　要

本論文介紹光學設計中自由型曲面的設計方法，我們探討了自由型曲面設計藉由偏微分方程來求得，其中的偏微分方程由 Schruben 推導而來，其偏微分方程的形式為 Monge-Ampére 方程式，我們考慮簡化型 Monge-Ampére 方程式，藉由馮教授所使用的方法，加上一個四次微分的消散項，可以使得原來的方程式是一個完全非線性的方程式轉變成類線性方程式改變了方程式的特性，使得在偏微分方程上有較好的特性，我們用有限元素法來做為我們的計算方法，挑選 BCIZ 有限元可以有效的處理四次微分項且並且可以簡單的求得曲面的曲率的計算，也可以滿足光學系統的所需要的一些特性，只需要解一個線性方程式和使用牛頓疊代法求做求解用的工具，以獲得較高的計算效率。

Numerical Studies on the Monge-Ampére Equation and its Nonimage Optical Application

student：Yu-Lin Tsai Advisors：Dr. Chin-Tien Wu

Depart of Applied Mathematics
Institute of Mathematical Modeling and Scientific Computing
National Chiao Tung University

ABSTRACT

We consider the freeform surface design problem. Fully nonlinear partial differential equations as derived by the Schruben for model. The partial differential equation is the form of well knowMonge-Ampére equation. We following Prof. Feng's idea to solve Monge-Ampére equation by adding a vanish moment biharmonic term. As a result the original fully nonlinear equation is change into quasi-linear equation. We using finite element method to solve this equation. Its well knows that the traditional BCIZ element can effectively deal with biharmonic item and compute the curvature of the solution. Which is usually required in a optical systems. We descritize the nonlinear equation by the Newton's method. The numerical studies in this thesis show that our approach is efficient and accurate.

誌　　　謝

本篇論文的完成，首先要感謝我的指導老師－吳金典教授。在老師的指點之下讓我認識了有限元素法、科學計算和數值偏微分方程等領域，帶領我如何去學習與如何做研究。除了課業之上為我的領路人之外，也讓我學到了諸多求知與考究數學的方法。在人生上，老師也指點我諸多可為不可為之事，在我人生的道路上為吾樹立了一個典範，一個值得我效仿的對象。再來我要感謝光電所－田仲豪教授，是帶領我進入光學的世界的導師，為我開啟了一扇稱之為光學的大門，讓我有深刻的體驗到何謂隔行如隔山，讓我了解了何謂光的世界，給予給我一個碩士生涯中的研究主題，也是我未來的探討主題，陪我一起悠游光的奧秘這片大海，除了在專業領域上的協助外，在觀看與了解自然界的本質上也讓我認識到除了數學外的另一種考就方式；而田教授走人生的旅途上讓我見識到另一種風采，讓我拾取另一副看待人世間的眼鏡。感謝兩位教授讓我在數值計算與光學之間，找到我的興趣與做研究的方向。除了兩位老師之外，感謝建模所的同學們，陪我度過這兩年給了我許許多多的幫助，在難過的寫程式的地獄之中也有你們的陪伴，在快樂的研讀之中也不會缺少你們的影子。感謝陳冠羽學長在數值計算上為我指點迷津，在生活中也給予我相當多的協助；感謝清大資工所徐廷使同學在寫作程式的給予我許多概念與細節上的指教，並幫助我思考程式上有新的處理手法；感謝簡明進學長與林至宏同學在光學上為我打起基礎，並不厭其煩的教導我所不足所缺少的部分，使我能在短時間學習到做研究上所需要最基本的光學知識，為我研究可以順利前行的兩位前輩。感謝吳恭儉學長以他的偏微分方程的專長在有限元素法的理論部分為我解惑。也要感謝前人所留下來的許許多多的資料。學生蔡玉麟在此致上我由衷的謝意。

在論文口試的時期，承蒙賴明治教授、田仲豪副教授以及李國明副教授的撥空審閱和提供諸多寶貴意見，令本論文得以完備與齊全，學生蔡玉麟永銘在心。

除了研究之外，我的生活之中也要感謝蘇偉碩學長、涂芳婷學姊、黃偉強學長、黃喻培學長、段俊旭同學、陳泓勳同學、碩一的學弟妹們及其田教授實驗室的同學們還有族繁不及備載的其他朋友們。感謝你們跟我在課業的討論以及努力，也讓我在休閒之刻可以有著充實的時光，交換各式各樣的心得以及意見，在我生活被快樂的時光所包圍，在互相扶持之下走過這段研究所的日子。

最後，特別感謝我的家人，在他們的支持之下，我才有機會念研究所，在他們的關心與祝福之中，我才有辦法順利的完成這個里程碑。我願分享這篇論文完成的喜悅與榮耀和我的家人，以及諸位關心我的朋友一同分享。

目　　　錄

Introduction

The Monge-Ampère equation is a important problem in differential geometry, optimal control, mass transportation, geostrophic fluid, meteorology and optimal design [1][2][3][4][5]. In this thesis, we focus on the optical free-form design problem. People study optical problems for a long time, thank for that the Mathematical fundation of the free-form design is more and more complete, we can try to solve the problem numercally. Now, what's optical free-form design problem? Given a light source and intensity in a optical system, and the illumination distribution on the target plane, the main problem of optical free-form design problem is design a optical system such that the transportantion form the light source to the target plane throught the designed system will not have energy loss. The optical system is generally consisted pf as following reflector amd refractor, etc. .

The general form of Monge-Ampère equation

$$\det(D^2u) + F\left(x, u, Du, D^2u\right) = 0 \text{ in } \Omega \tag{1.1}$$

where $D^2u = \left(\frac{\partial^2 u}{\partial x_i \partial x_j}\right)_{i,j=1,\ldots,n}$ is the Hessian of the function u at $x \in \Omega$.

Suppose coefficients in (1.19) depending on variables x, y, and the unknown function $u\left(x, y\right)$, (1.19) can be rewritten as following

$$\det(D^2u) = au_{xx} + bu_{xy} + cu_{yy} + \phi \tag{1.2}$$

The Monge-Ampere equation can be either ellip, parabilic or hyperbolic depending on the sign of $\triangle$

$$\triangle = \phi + ac + b^2 \tag{1.3}$$

If $\triangle > 0$, then the Monge-Ampère equation is of elliptic type, if $\triangle < 0$ it is of hyperbolic type and if $\triangle = 0$ it is of parabolic type. A non-linear elliptic partial differential equation. It is well know that the solution of the Monge-Ampère equation is not unique unless we confines our attention to the convex solutions. The existence and uniqueness of the convex solution of the Monge-Ampére equation satisfies version of the maximum principle, and in particular solutions with given Dirichlet condition is proved by Pogorelov in [2, 6] general result on the existence and unquenss are later obtained by Oliker and Wang etc. .

the free-form surface is the solution of Monge-Ampère equation in three dimension space. In 1972 [7], Schruben described the reflector is the solution of the Monge-Ampere equation. He derived the partial differential equation from the integral equation of the energy conservation. In 1993, Oliker and Newman also derived the Monge-Ampere equation in reflector problem. Since the Monge-Ampere equation, a fully non-linear elliptic partial differential equation, is hard to solve. So if we wanted to use it, we must add some condition such that the equation is more simplify.

until 1990, Benitez , Juan, et al. develoed of the Simltaneous Multiple Surface (SMS) method, for the design of 2D profiles of non-imaging optical devices (SMS2D). It was a breakthrough in a field dominated by bulky designs. In 2004, SMS non-imaging method generated free-form optical surfaces in 3D (SMS3D) [8], which is a major extension of SMS2D. In the SMS method, the free-form surface is constructed first by defining the incoming wavefront and outgoing wavefront, instead of the source and receiver, and then deciding the basic point and optical path length. In order to find the outgoing wavefront, one must solve the Monge-Ampère equation. So, a numerical method is desired for solving the Monge-Ampere equation.

Glowinski, Benamou etc., Gerard Awanou and Feng and Neilan consider the following Monge-Ampère equation:

$$\begin{aligned} \det\left[D^2 u\right] &= f \text{ in } \Omega && (1.4) \\ u &= g \text{ on } \partial\Omega && (1.5) \end{aligned}$$

where Ω is a convex domain with smooth boundary $\partial\Omega$ and $D^2u = \begin{bmatrix} u_{xx} & u_{xy} \\ u_{xy} & u_{yy} \end{bmatrix}$ is the Hessian of the function u at $x \in \Omega$.

Two method are employed by Benamou, Froese and Oberman [10] to solve the Monge-Ampere equation. One is an explicit finite difference method, The equ (1.4) is using discretized as following standard central difference on a uniform Cartesian grid.

$$\left(D^2_{xx}u_{ij}\right)\left(D^2_{yy}u_{ij}\right) - \left(D^2_{xy}u_{ij}\right)^2 = f_{ij} \tag{1.6}$$

where

$$\begin{aligned} D^2_{xx}u_{ij} &= \frac{1}{h^2}\left(u_{i+1,j} + u_{i-1,j} - 2u_{i,j}\right) \\ D^2_{yy}u_{ij} &= \frac{1}{h^2}\left(u_{i,j+1} + u_{i,j-1} - 2u_{i,j}\right) \\ D^2_{xy}u_{ij} &= \frac{1}{4h^2}\left(u_{i+1,j+1} + u_{i-1,j-1} - u_{i-1,j+1} - u_{i+1,j-1}\right) \end{aligned} \tag{1.7}$$

The equ (1.6) is further rewrote the a quadratic equation for $u_{i,j}$, as following

$$u_{i,j} = \frac{1}{2}\left(a_1 + a_2\right) - \frac{1}{2}\sqrt{\left(a_1 - a_2\right)^2 + \frac{1}{4}\left(a_3 - a_4\right)^2 + h^4 f_{i,j}} \tag{1.8}$$

where

$$\begin{aligned} a_1 &= (u_{i+1,j} + u_{i-1,j})/2 \\ a_2 &= (u_{i,j+1} + u_{i,j-1})/2 \\ a_1 &= (u_{i+1,j+1} + u_{i-1,j-1})/2 \\ a_1 &= (u_{i-1,j+1} + u_{i+1,j-1})/2 \end{aligned} \tag{1.9}$$

The other method employed by Benamou, Froese and Oberman is solving $u = T(u)$ by fixed point iteration where

$$T(u) = \triangle^{-1}\left(\sqrt{(\triangle u)^2 + 2\left(f - \det(D^2u)\right)}\right) \tag{1.10}$$

the itervates $u^{n+1} = T(u^n)$ is obtained by solving

$$\triangle u^{n+1} = \sqrt{(u^n_{xx})^2 + (u^n_{yy})^2 + 2\left(u^n_{xy}\right)^2 + 2f} \tag{1.11}$$

Dean and Glowinski [11, 12, 13]. They first consider the Monge–Ampére equation as a saddle-point problem where a suitable augmented Lagrangian has to he chosen. To solve this saddle-point problem, they advocate an Uzawa–Douglas–Rachford algorithm. The second approach Dean and Glowinski used is to combine non-linear least-square method and operator-splitting. A mixed finite element discretization is used in their formulation.

Feng and Neilan [15, 16] add $-\epsilon\triangle^2 u^\epsilon$ to regularize the Monge-Ampere equation. An artifical a boundary condition $\triangle u^\epsilon = \epsilon$ is introfuced on $\partial\Omega$. The quasilinear forth order pde,

$$-\epsilon\triangle^2 u^\epsilon + \det\left(D^2 u^\epsilon\right) = f, \text{ in } \Omega \tag{1.12}$$

$$u^\epsilon = g \text{ on } \partial\Omega \tag{1.13}$$

$$\triangle u^\epsilon = \epsilon \text{ on } \partial\Omega \tag{1.14}$$

is then separated into coupled second order partial difference equations system

$$\begin{aligned} \sigma^\epsilon - D^2 u^\epsilon &= 0 \\ -\epsilon\triangle tr\left(\sigma^\epsilon\right) + \det\left(\sigma^\epsilon\right) &= f \end{aligned} \tag{1.15}$$

A mixed finite element is the empolyed to solve the above equations.

Gerard Awanou [30], takes a similar approch as feng and Neilan, by adding $-\frac{\epsilon}{n}\triangle^2 u^\epsilon$ to the Monge-Ampere equation and adding a boundary condition $\triangle u^\epsilon = \epsilon^2$ on $\partial\Omega$. The corresponding variational problem is: to find $u^\epsilon \in H^2(\Omega)$, $u^\epsilon = g$, $\triangle u^\epsilon = \epsilon^2$ on $\partial\Omega$ such that

$$\epsilon\int_\Omega \triangle u^\epsilon \triangle v dx + \int_\Omega \left(cof\left(D^2 u^\epsilon\right) Du^\epsilon\right)\cdot Dv\, dx = -n\int_\Omega f v dx \;\; \forall v \in H_0^2(\Omega) \tag{1.16}$$

where

$$cof\left(D^2 u^\epsilon\right) = \begin{bmatrix} u_{yy} & -u_{xy} \\ -u_{xy} & u_{xx} \end{bmatrix} \tag{1.17}$$

Again, Awanou employ the mix finite element to approximate the partial differential equation.

In free-form design problem, we must determine the control point and the normal vector. Following this ideal, in this paper, we solve the regularized equ(1.12) which is basically a biharmonic equation with low order nonlinear term, so we solve this regularized equation direatly instead of devoupling the equation into a couple low order system as Feng and Neilan did. We employee the Newton iterative method to linearite the nonlinear part, since Newton's method is well known in finding successively better approximations to the zeros of a real-valued nonlinear function. Newton's method can often converge quickly, if the iteration have a good initial point. we choose BCIZ element. BCIZ element is one of the simplest Kirchhoff plate bending elements was presented by Bazeley, Cheung, Irons and Zienkiewicz at the 1965 Wright-Patterson Conference [17]. The "BCIZ element" is named after the authors initials. This element can be derived from the cubic interpolation which basically has 10 degrees of freedom. The variable in the element centroid is condensed out using a kinetic constraint in such a way that the curvature completeness is maintained.

The biharmonic equation, besides providing a benchmark problem for various analytical and numerical methods, arises in many particular applications. For example, the bending behaviour of a thin elastic plane.

Chapter 1
Mathematical Modeling of Optical design

In this chapter, we derive the Monge-Ampére equation follow Schruben in 1972. He consider that a point source though a reflector to target plane, he describing the free-form surface is the solution of Monge-Ampére equation in three dimension space.

The light source is assumed to have some arbitrary directional intensity distribution I and dimensions that emits are negligible compared to the fixture size. Distances are normalized such that the distance from the source to the (u, v) plane is unity. The target area on (u, v) plane that is to be illuminated.

Since the intensity of the source is directional, I may defined as a function of position on the unit sphere centered at the source. Spherical coordinates could be used, but it is preferable to employ parametric coordinates (u, v) of the unit sphere. These may be obtained as stereographic coordinates, as illustrated in Fig. 1.2, by projecting the unit sphere from its point of tangency to (x, y) plane onto the plane (u, v) plane parallel to (x, y) plane and also tangent to the sphere. The stereographic coordinates of a point on the sphere so projected are the rectangular (u, v) coordinates of the corresponding point on (u, v) plane.

Defined a function $L = L(x, y)$, which is the desired pattern of reflected illumination on the target plane. This is defined as the desired pattern of total illumination at a point (x, y) on the target plane from which has been subtracted the direct illumination of the source at (x, y) which can be obtained directly from the intensity distribution I. by energy

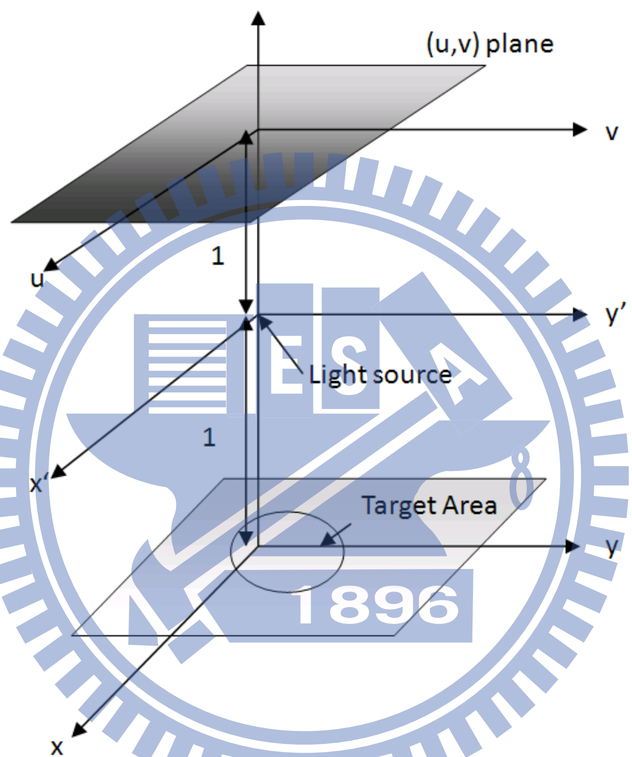


Fig. 1.1. point source of light

conservation

$$\iint_{v(\Omega)} L(x,y)\, dxdy = \iint_{\Omega} I(u,v)\, d\Omega \tag{1.18}$$

where Ω is solid angles and $v(\Omega)$ is the target area.

Define a vector mapping $\acute{x}$ that maps a point (u,v) on the u,v plane to a vector (light ray) from the orgin (light source) to a point (x',y',z') on the unit sphere.

the explicit form of this map is

$$\acute{x}(u,v) = \left(1+\frac{1}{4}w^2\right)^{-1}\left(u,v,1-\frac{1}{4}w^2\right) \tag{1.19}$$

where $w^2 = u^2 + v^2$, where the $\acute{x}(u,v)$ is not unique, under different problem we can change the coordinate. In 1993, Oliker and Newman proved that the formulation has existence and uniqueness solution.

The differential solid angle $d\Omega$ is area on the unit sphere and is related to differential area on the $uv-\ plane$ by the equation

$$d\Omega = |x_u \times x_v|\, dudv \tag{1.20}$$

Differentiation of equation (1.19) yields

$$\acute{x}_u(u,v) = \left(1+\frac{1}{4}w^2\right)^{-2}\left(1+\frac{1}{4}v^2-\frac{1}{4}u^2,-\frac{1}{2}uv,-u\right) \tag{1.21}$$

$$\acute{x}_v(u,v) = \left(1+\frac{1}{4}w^2\right)^{-2}\left(-\frac{1}{2}uv,1+\frac{1}{4}v^2-\frac{1}{4}u^2,-v\right) \tag{1.22}$$

therefore

$$d\Omega = \left(1+\frac{1}{4}w^2\right)^{-2} dudv \tag{1.23}$$

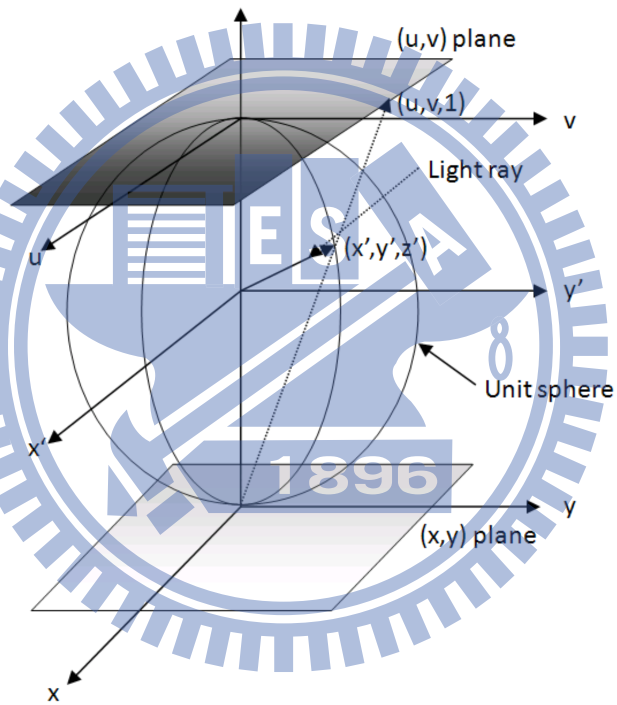


Fig. 1.2. stereographic coordinates

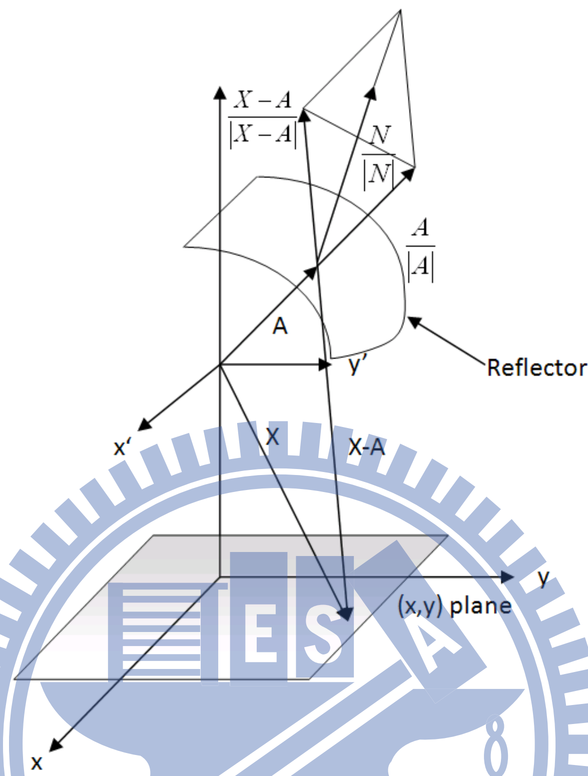


Fig. 1.3. Reflection geometry

Substituting this into equation (1.18), then

$$\iint_{v(\Omega)} L(x,y)\,dxdy = \iint_{\Omega} I(u,v)\left(1+\frac{1}{4}w^2\right)^{-2} dudv \tag{1.24}$$

We can describe the reflector by an equation $\rho = \rho(u,v)$ where ρ is the length of a ray with stereographic coordinates (u,v) from the origin to the reflecting surface. We now shall transform the integral equation (1.24) for the reflection function w to a partial differential equation for the surface function ρ.

Let A be a vector with stereographic coordinates (u,v) that strikes the reflector $\rho = \rho(u,v)$. Then $A = \rho\acute{x}$, where $\acute{x}$ is given by Eq. (1.19). Suppose this ray is reflected to the point (x, y) on xy plane, then the vector X from the source to this point has the coordinates $(x, y, -1)$ in the x', y', z' coordinate system.

The vector $N = A_u \times A_v$ is an outward normal vector to the surface $\rho = \rho(u, v)$ at A. Since $A = \rho x$,

$$N = (px + \rho x_u) \times (qx + \rho x_v) \tag{1.25}$$

where $p = \rho_u$ and $q = \rho_v$.

We find

$$N = \rho^2 \left(1 + \frac{1}{4} w^2\right)^{-2} x - p_\rho x_u - q_\rho x_v \tag{1.26}$$

As is illustrated in Figure, the law of reflection requires that

$$\frac{A}{|A|} + \frac{A - X}{|X - A|} = \frac{2N (A \cdot N)}{|A| |N|^2} \tag{1.27}$$

then

$$X = A + |X - A| \left[\acute{x} - 2N (\acute{x} \cdot N) / |N|^2 \right] \tag{1.28}$$

since $A / |A| = \acute{x}$

So the x, y can be show as

$$x = uG + 2\rho\rho_u F \tag{1.29}$$

$$y = vG + 2\rho\rho_v F \tag{1.30}$$

where

$$
\begin{aligned}
G &= \rho\left(1+\frac{1}{4}w^2\right)^{-1}+F\left[-\rho^2\left(1+\frac{1}{4}w^2\right)^{-2}\right. \\
&\quad \left.+\rho_u^2+\rho_v^2-\rho\left(\rho_u u+\rho_v v\right)\left(1+\frac{1}{4}w^2\right)^{-1}\right] \qquad (1.31)\\
F &= \frac{1+\rho\left(1-\frac{1}{4}w^2\right)\left(1+\frac{1}{4}w^2\right)^{-1}}{\left(1-\frac{1}{4}w^2\right)\left[\rho^2\left(1+\frac{1}{4}w^2\right)^{-2}-\rho_u^2-\rho_v^2\right]+2\rho\left(\rho_u u+\rho_v v\right)\left(1+\frac{1}{4}w^2\right)^{-1}} \qquad (1.32)
\end{aligned}
$$

The integration over x and y in the left-hand side of (1.24) may be transformed to integration over u and v by multiplication by the Jacobian

$$
D = \begin{vmatrix} x_u+\rho_u x_\rho+\rho_{uu}x_{\rho_u}+\rho_{uv}x_{\rho_v} & x_v+\rho_v x_\rho+\rho_{vu}x_{\rho_u}+\rho_{vv}x_{\rho_v} \\ y_u+\rho_u y_\rho+\rho_{uu}y_{\rho_u}+\rho_{uv}y_{\rho_v} & y_v+\rho_v y_\rho+\rho_{vu}y_{\rho_u}+\rho_{vv}y_{\rho_v} \end{vmatrix} \qquad (1.33)
$$

The a reflector $\rho=\rho(u,v)$ with continuous second derivatives, the integral equation (1.24) is equivalent to the partial differential equation

$$
L\left(x\left(u,v,\rho,\rho_u,\rho_v\right),y\left(u,v,\rho,\rho_u,\rho_v\right)\right)D=I\left(u,v\right)\left(1+\frac{1}{4}w^2\right)^{-2} \qquad (1.34)
$$

Expanding the jacobian

$$
\begin{aligned}
D &= J_{\rho_u\rho_v}\left(\rho_{uu}\rho_{vv}-\rho_{uv}^2\right)+\left(J_{\rho_u v}+\rho_v J_{\rho_u\rho}\right)\rho_{uu} \\
&\quad +\left(J_{\rho_v v}+J_{u\rho_u}+\rho_u J_{\rho\rho}+\rho_v J_{\rho_v\rho}\right)\rho_{uv} \\
&\quad +\left(J_{u\rho_v}+\rho_v J_{\rho_u\rho_v}\right)\rho_{vv}+J_{uv}+\rho_u J_{\rho v}+\rho_v J_{u\rho} \qquad (1.35)
\end{aligned}
$$

where

$$
J_{\alpha\beta}=x_\alpha y_\beta-x_\beta y_\alpha \quad for\ \alpha,\beta\in\{u,v,\rho,\rho_u,\rho_v\} \qquad (1.36)
$$

The leading term of the differential equation is $(\rho_{uu}\rho_{vv} - \rho_{uv}^2)$, so it easy to see the equation is Monge-Ampere type.

We consider the ideal case

$$\left(\rho_{uu}\rho_{vv} - \rho_{uv}^2\right) = f \tag{1.37}$$

in our study.

Chapter 2
Finite Element Method

The basic idea in any numerical method for a differential equation is to discretize the given continuous problem to obtain a discrete problem or system of equations with only finitely many degrees of freedom such that the differential equation can be solved by using a computer.

Finite element method start from a reformulation of the given differential equation as an equivalent variational problem. In the case of elliptic equations this variational problem in basic case is a minimization problem of the form

$$Find\ u \in V\ such\ that\ F(u) \leqq F(v)\ \ for\ all\ v \in V \tag{2.38}$$

where V is a given set of admissible functions and $F : V \to R$ is a functional. $F(v)$ is the total energy associated with v and (2.38) corresponds to an equivalent characterization of the solution of the differential equation as the function in V that minimizes the total energy of the considered system. In general the dimension of V is infinite and thus in general the problem (2.38) can't be solved exactly. To obtain a problem that can be solved on a computer the idea in the finite element method is to replace V by a set V_h consisting of simple function only depending on finitely many parameters. This leads to a finite-dimensional minimization problem of the form:

$$Find\ u_h \in V_h\ such\ that\ F(u_h) \leqq F(v)\ \ for\ all\ v \in V_h \tag{2.39}$$

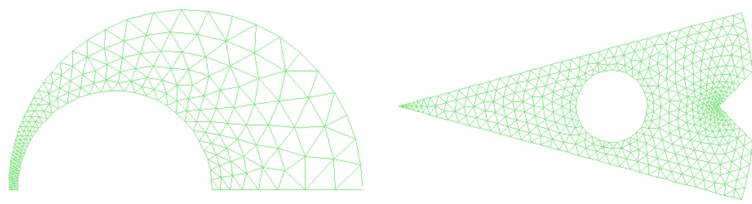


Fig. 2.4. mesh of two dimension domain

This problem is equivalent to a linear or non-linear system of equations. We hope that the solution u_h of this problem is sufficiently good approximation of the solution of the original minimization problem (2.38). Usually one chooses V_h to be a subset of V and in this case (2.39) corresponds to the classical Ritz-Galerkin method.

To solve a given differential or integral equation approximately using the finite element method, one has to go through basically the following steps:

1. Variational formulation of the given problem
2. Discretization using FEM: construction of the finite dimensional space V_h
3. Generating Mesh
4. Choose basis function
5. Assembling
6. Solve the linear system

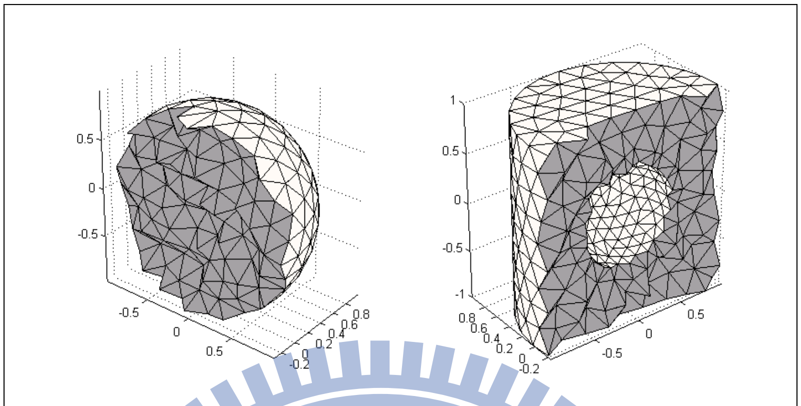


Fig. 2.5. mesh of three dimension domain

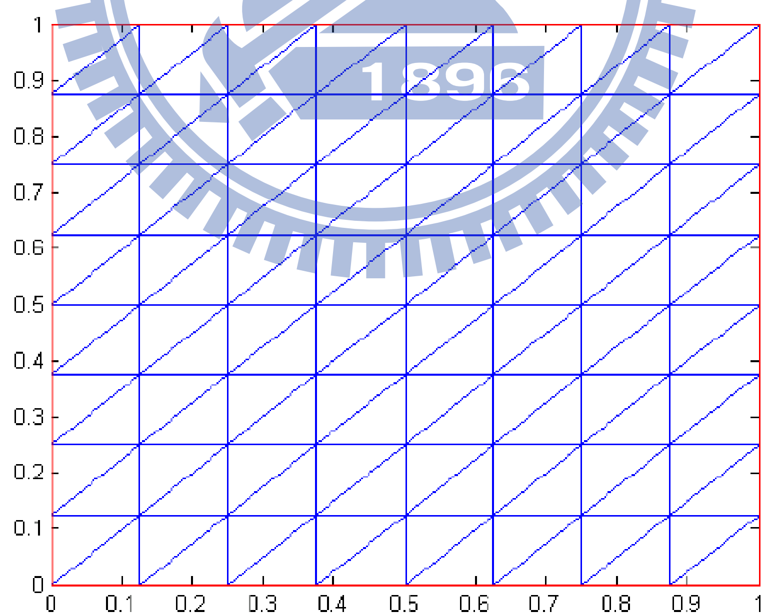


Fig. 2.6. uniform mesh

2.1 Variational formation

In this section we will give two example for variation formulation. One is Poisson equation, other is biharmonic equation.

2.1.1 Poisson Equation

Consider the following boundary value problem for the Poisson equation, the second order differential equation:

$$\begin{cases} -\nabla \cdot (A\nabla u) = f & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega \end{cases} \tag{2.40}$$

where Ω is a bounded open domain in the plane $\mathbb{R}^2$ with boundary $\partial\Omega$, A is a matrix, f is a given function and as usual,

$$\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \tag{2.41}$$

A number of problems in physics and mechanics are modelled by (2.40); u may represent for instance a temperature, an electro-magnetic potential or the displacement of an elastic membrane fixed at the boundary under a transversal load of intensity f.

We shall now give a variational formulation of problem (2.40). We shall first show that if u satisfies (2.40), then u is the solution of the following variational problem:

$$-\int_\Omega v\nabla \cdot (A\nabla u)\, dx = \int_\Omega \nabla v A \nabla u dx - \frac{\partial u}{\partial n} v|_{\partial\Omega} = \int_\Omega v f dx \tag{2.42}$$

where v is test function in $H_0^1(\Omega)$, $v = 0$ on $\partial\Omega$.

2.1.2 Biharmonic Equation

Consider the following boundary value problem for the biharmonic equation, the fourth order differential equation:

$$\begin{cases} -\Delta^2 u = f & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega \\ \frac{\partial u}{\partial n} = 0 & \text{on } \partial\Omega \end{cases} \tag{2.43}$$

where Ω is a bounded open domain in the plane $\mathbb{R}^2$ with boundary $\partial\Omega$, f is a given function and as usual,

$$\triangle^2 u = \frac{\partial^4 u}{\partial x^4} + 2\frac{\partial^4 u}{\partial x^2 \partial y^2} + \frac{\partial^4 u}{\partial y^4} \tag{2.44}$$

A number of problems in physics and mechanics are modelled by (2.43); u may represent the solution of Stokes flows or the displacement of plane bending problem.

$$\begin{aligned} -\int_\Omega v\triangle^2 u dx &= \int_\Omega \nabla v \nabla \triangle u dx - \frac{\partial \triangle u}{\partial n} v|_{\partial\Omega} \\ &= -\int_\Omega \triangle v \triangle u dx - \triangle u \frac{\partial v}{\partial n}|_{\partial\Omega} \\ &= \int_\Omega v f dx \end{aligned} \tag{2.45}$$

where v is test function in $H_0^2(\Omega)$, $v = 0$ on $\partial\Omega$, $\frac{\partial v}{\partial n} = 0$ on $\partial\Omega$.

Case in point, regularization Monge-Ampére equation, it has second order and fourth order differential term. We will introduction it in chapter 3.

2.2 Existence and Uniqueness of Solution

Definition Let H be a Hilbert space. A bilinear form $a : H \times H \to \mathbb{R}$ is called *continuous* provided there exists $C > 0$ such that

$$|a(u, v)| \le C \|u\| \|v\| \text{ for all } u, v \in H$$

A symmetric continuous bilinear form a is called *H-elliptic*, or *short elliptic* or *coercive*, provided for some $\alpha > 0$,

$$a(v, v) \ge \alpha \|v\| \text{ for all } v \in H$$

clearly, every H-elliptic bilinear form a induces a norm via

$$\|v\|_a := \sqrt{a(v, v)} \tag{2.46}$$

This is equivalent to the norm of the Hilbert space H. The norm (2.46) is called the *energy norm*.

As usual, the space of continuous linear functions on a normed linear space V will be denoted by V'.

Example Consider the boundary value problem of Poisson equation:

$$\begin{cases} -\nabla \cdot (\nabla u) = f & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega \end{cases} \tag{2.47}$$

One variational formulation for this is: Take

$$\begin{aligned} V &= H^1(\Omega) \\ a(u,v) &= \int_\Omega (\nabla u \cdot \nabla v)\, dx \\ F(v) &= (f,v) \end{aligned} \tag{2.48}$$

To prove $a(\cdot,\cdot)$ is continuous, observe that

$$|a(u,v)| \leq |(u,v)_{H^1}| \leq \|u\|_{H^1} \|v\|_{H^1} \tag{2.49}$$

The Lax-Milgram Theorem Given a Hilbert space $(V,(\cdot,\cdot))$, a continuous, coercive bilinear form $a(\cdot,\cdot)$ and a continuous linear functional $F \in V'$, there exists a unique $u \in V$ such that

$$a(u,v) = F(v) \quad \forall v \in V \tag{2.50}$$

2.3 Estimates for General Finite Element Approximation

Let u be the solution to the variational problem and u_h be the solution to the approximation problem. To estimate the error $\|u - u_h\|_V$.

Céa Lemma Suppose the bilinear form a is V-elliptic with $H_0^m(\Omega) \subset V \subset H^m(\Omega)$. In addition, suppose u and u_h are the solution of the variational problem in V and V_h, respectively, Then

$$\|u - u_h\|_V \leq \frac{C}{\alpha} \inf_{v_h \in V_h} \|u - v_h\|_V \tag{2.51}$$

where C is the continuity constant and α is the coercivity constant of $a(\cdot,\cdot)$.

2.4 Finite Element Space

Finite element have two type, conforming finite element and nonconforming finite element, in the theory of conforming finite element it is assumed that the finite element spaces lie in the function space in which the variational problem is posed. Moreover, we also require that the given bilinear form $a(\cdot,\cdot)$ can be computed exactly on the finite element spaces. The Finite element space of nonconforming finite element do not lie in function space.

Now we follow Ciarlet's definition of a finite element (Ciarlet 1978).

Definition Let

1. $K \subseteq \Omega \subseteq \mathbb{R}^n$ be a bounded closed set with non-empty interior and piecewise smooth boundary (the **element domain**),

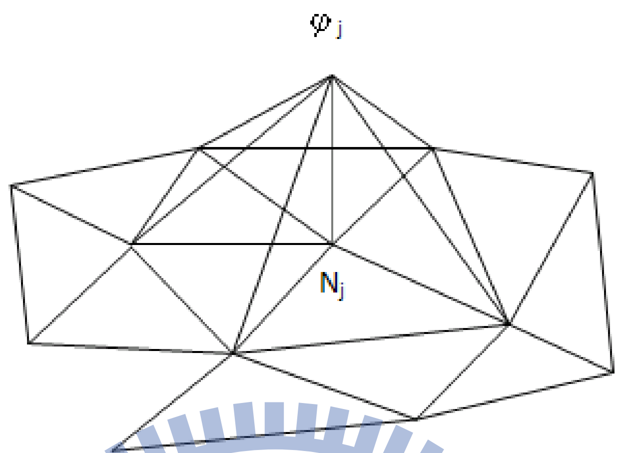
2. $\mathcal{P}$ be a finite-dimensional space of functions on K (the space of **shape function**) and

3. $\mathcal{N} = \{N_1, N_2, ..., N_k\}$ be a basis for $\mathcal{P}'$ (the set of **nodal variable**).

   Then $(K, \mathcal{P}, \mathcal{N})$ is called a **finite element.**

Definition Let $(K, \mathcal{P}, \mathcal{N})$ be a finite element. The basis $\{\varphi_1, \varphi_2, ..., \varphi_k\}$ of $\mathcal{P}$ dual to $\mathcal{N}$ is called the **nodal basis** of $\mathcal{P}$.

After generating Mesh, we construct a finite dimensional subspace V_h of the space V defined consisting of piecewise linear function. We now let V_h be the set of functions v such that v is linear on domain Ω, v is continuous on domain Ω and $v = 0$ on $\partial\Omega$.We



observe that $V_h \subset V$. As parameter to describe a function $N_j = v\left(x_j\right)$ we may choose the values $N_j = v\left(x_j\right)$ at the node points $x_j, j = 0, ..., m+1$. Let us introduction the basis function $\phi_j \in V_h, j = 0, ..., m+1$. defined by

$$\phi_j\left(x_i\right) = \delta_{ij} \tag{2.52}$$

i.e., ϕ_j is the continuous piecewise linear function that take the value 1 at node point x_j and the value 0 at other node points. A function $v \in V_h$ then has the representation

$$v\left(x\right) = \sum_{i=1}^{m} \eta_i \phi_i\left(x\right), \quad x \in \Omega \tag{2.53}$$

where $N_j = v\left(x_j\right)$, i.e., each $N_j = v\left(x_j\right)$ can be written in a unique way as a linear combination of the basis function ϕ_j. In particular it follow that V_h is a linear space of dimension m with basis $\left\{\varphi_j\right\}_{i=1}^{m}$.

We consider the shape function of K, because we need to compute the solution on computer. We give some example of **Finite Element**, and how to connect the global coordinate with local coordinate.

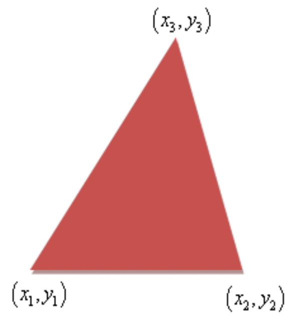


Fig. 2.7. labeled number

2.4.1 Triangular Finite Element

In two dimension domain, we can generate mesh by triangular or rectangular. We use the BCIZ triangular element to approximate the Monge-Ampére equation, more detail about BCIZ element will be introduction in Chapter 3. First of all, we introduction the relationship between two coordinates, second part is triangular finite element .

Geometry:

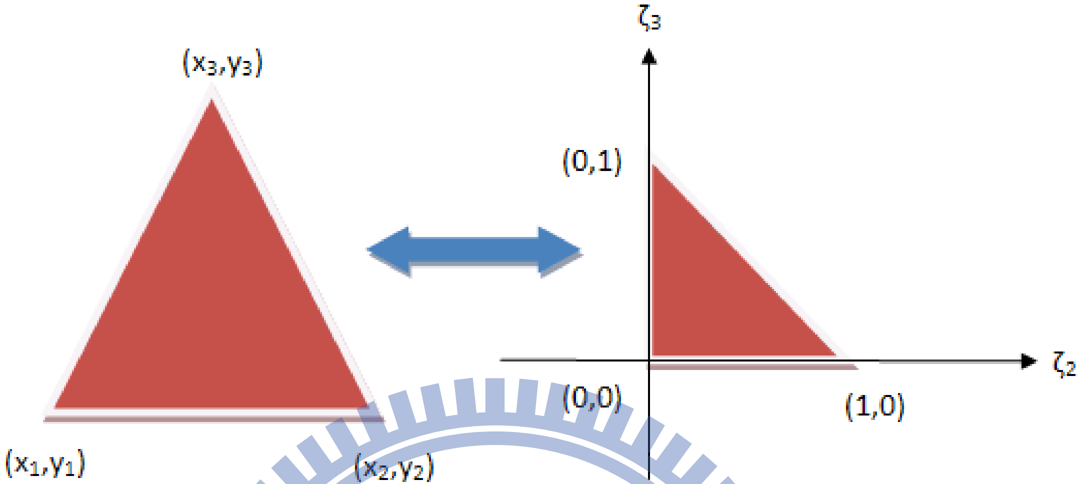
The geometry of the 3-node triangle show in Figure 2.4 is specified by the location of its three corner nodes on the $\{x, y\}$ plane. The nodes are labeled 1, 2, 3 while traversing the sides in counterclockwise fashion. The location of the corners is defined by their coordinates:

$$(x_i, y_i) \;\; i = 1, 2, 3$$

the area of triangle is denoted by $\bar{A}$ and is given by:

$$2\bar{A} = \det \begin{bmatrix} 1 & 1 & 1 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{bmatrix} = x_{21}y_{31} - x_{31}y_{21} \tag{2.54}$$

where $x_{ij} = x_i - x_j, y_{ij} = y_i - y_j$ for $i, j = 1, 2, 3 \; i \neq j$.



Properties of Triangular Coordinates:

Consider triangular on regular triangular, points of the triangle may also be located in terms of a parametric coordinate system:

$$\zeta_1, \zeta_2, \zeta_3$$

this is a local coordinate.

Represent a set of straight lines parallel to the side opposite to the i^{th} corner. See Figure. The equation of sides 12, 23 and 31 are $\varphi_1 = 0, \varphi_2 = 0$ and $\varphi_3 = 0$. respectively. The three corners have coordinates $(0, 0, 1)$, $(0, 1, 0)$ and $(1, 0, 0)$ .The three midpoints of the sides have coordinates, $\left(\frac{1}{2}, \frac{1}{2}, 0\right)$, $\left(0, \frac{1}{2}, \frac{1}{2}\right)$ and $\left(\frac{1}{2}, 0, \frac{1}{2}\right)$, the centroid $\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$, and so on. The coordinates are not independent because their sum is unity:

$$\zeta_1 + \zeta_2 + \zeta_3 = 1 \tag{2.55}$$

Coordinate Transformations:

Quantities which are closely linked with the element geometry are naturally expressed in triangular coordinates. On the other hand, quantities such as displacements, strains and stresses are often expressed in the Cartesian system x, y. We therefore need transformation equations through which we can pass from one coordinate system to the other.

Cartesian and triangular coordinates are linked by the relation

$$\begin{bmatrix} 1 \\ x \\ y \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{bmatrix} \begin{bmatrix} \zeta_1 \\ \zeta_2 \\ \zeta_3 \end{bmatrix} \tag{2.56}$$

The first equation says that the sum of the three coordinates is one. The second and third express x and y linearly as homogeneous forms in the triangular coordinates. These simply apply the linear interpolant formula to the Cartesian coordinates: $x = x_1\zeta_1 + x_2\zeta_2 + x_3\zeta_3$ and $y = y_1\zeta_1 + y_2\zeta_2 + y_3\zeta_3$.

Inversion of (2.56) yields

$$\begin{bmatrix} \zeta_1 \\ \zeta_2 \\ \zeta_3 \end{bmatrix} = \frac{1}{2\bar{A}} \begin{bmatrix} x_2y_3 - x_3y_2 & y_{23} & x_{32} \\ x_3y_1 - x_1y_3 & y_{31} & x_{13} \\ x_1y_2 - x_2y_1 & y_{12} & x_{21} \end{bmatrix} \begin{bmatrix} 1 \\ x \\ y \end{bmatrix} \tag{2.57}$$

Partial Derivatives:

From equations (2.56) and (2.57) we immediately obtain the following relations between partial derivatives:

$$\frac{\partial x}{\partial \zeta_i} = x_i, \frac{\partial y}{\partial \zeta_i} = y_i \tag{2.58}$$

$$\frac{\partial \zeta_i}{\partial x} = \frac{y_{jk}}{2\bar{A}}, \frac{\partial \zeta_i}{\partial y} = \frac{x_{kj}}{2\bar{A}} \tag{2.59}$$

where j and k denote the cyclic permutations of i. For example, if i = 3, then j = 1 and k = 2. The first derivatives of a function $w(\zeta_1, \zeta_2, \zeta_3)$ with respect to x or y follow immediately from (2.59) and application of the chain rule:

$$\frac{\partial w}{\partial x} = \frac{1}{2A}\left(\frac{\partial w}{\partial \zeta_1}y_{23} + \frac{\partial w}{\partial \zeta_2}y_{31} + \frac{\partial w}{\partial \zeta_3}y_{12}\right) \tag{2.60}$$

$$\frac{\partial w}{\partial y} = \frac{1}{2A}\left(\frac{\partial w}{\partial \zeta_1}x_{32} + \frac{\partial w}{\partial \zeta_2}x_{13} + \frac{\partial w}{\partial \zeta_3}x_{21}\right) \tag{2.61}$$

which matrix form is

$$\begin{bmatrix} \frac{\partial w}{\partial x} \\ \frac{\partial w}{\partial y} \end{bmatrix} = \frac{1}{2A}\begin{bmatrix} y_{23} & y_{31} & y_{12} \\ x_{32} & x_{13} & x_{21} \end{bmatrix}\begin{bmatrix} \frac{\partial w}{\partial \zeta_1} \\ \frac{\partial w}{\partial \zeta_2} \\ \frac{\partial w}{\partial \zeta_3} \end{bmatrix} \tag{2.62}$$

Triangular Finite Element

Let K be any triangle. Let $\mathcal{P}_k$ denote the set of all polynomials in two variables of degree $\leq k$.

1. Linear Lagrange triangle

Let $\mathcal{P} = \mathcal{P}_1$. Let $\mathcal{N}_1 = \{N_1, N_2, N_3\}$ $(\dim \mathcal{P}_1 = 3)$Note that "•" indicates the nodal

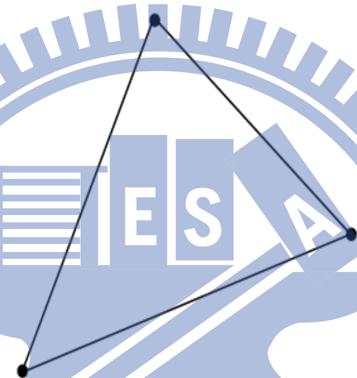


Fig. 2.8. Linear Lagrenge triangle

variable evaluation at the point where the dot is located.

2. Cubic Hermite triangle

Let $\mathcal{P} = \mathcal{P}_3$. Let $\mathcal{N}_3 = \{N_1, N_2, ..., N_{10}\}$ $(\dim \mathcal{P}_3 = 10)$Note that "•" indicates the

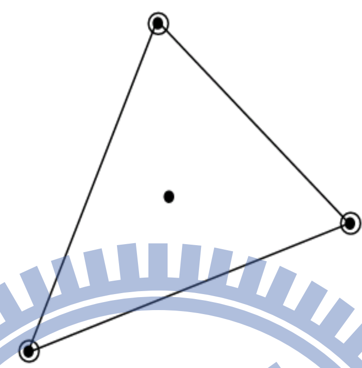


Fig. 2.9. Cubic Hermite triangle

nodal variable evaluation at the point and "◯" denote evaluation of the gradient at the center of the circle.

3. Quadratic Lagrange triangle

Let $\mathcal{P} = \mathcal{P}_2$. Let $\mathcal{N}_2 = \{N_1, N_2, ..., N_6\}$ $(\dim \mathcal{P}_2 = 6)$

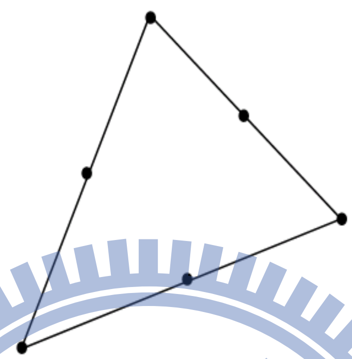


Fig. 2.10. Quadratic Lagrange triangle

4. Cubic Lagrange triangle

Let $\mathcal{P} = \mathcal{P}_3$. Let $\mathcal{N}_3 = \{N_1, N_2, ..., N_{10}\}$ $(\dim \mathcal{P}_3 = 10)$

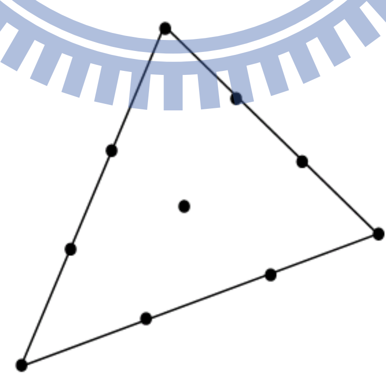


Fig. 2.11. Cubic Lagrange triangle

2.5 The Interpolant

Consider a function $w(x,y)$ that varies linearly over the triangle domain. In terms of Cartesian coordinates it may be expressed as

$$w(x,y) = a_0 + a_1 x + a_2 y \tag{2.63}$$

where a_0, a_1 and a_2 are coefficients to be determined from three conditions. In finite element work such conditions are often the nodal values taken by N at the corners:

$$N_1, N_2, N_3$$

The expression in triangular coordinates makes direct use of these three values:

$$\begin{aligned} w(\varphi_1,\varphi_2,\varphi_3) &= N_1\varphi_1 + N_2\varphi_2 + N_3\varphi_3 = \begin{bmatrix} N_1 & N_2 & N_3 \end{bmatrix} \begin{bmatrix} \varphi_1 \\ \varphi_2 \\ \varphi_3 \end{bmatrix} \\ &= \begin{bmatrix} \varphi_1 & \varphi_2 & \varphi_3 \end{bmatrix} \begin{bmatrix} N_1 \\ N_2 \\ N_3 \end{bmatrix} \end{aligned} \tag{2.64}$$

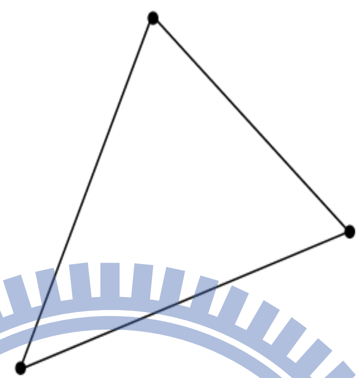
equation (2.64) is called a linear interpolant for w.

Definition Given a finite element $(K,\mathcal{P},\mathcal{N})$, let the set $\{\varphi_i : 1 \le i \le k\} \subseteq \mathcal{P}$ be the basis dual to $\mathcal{N}$. If v is a function for which all $N_i \in \mathcal{N}, i = 1,...,k$ are defined, then we define the **local interpolant** by

$$w(v) := \sum_{i=1}^{k} N_i(v)\,\varphi_i \tag{2.65}$$

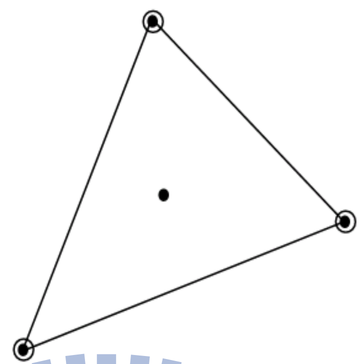
Interpolant of Triangular Finite Element

1. Linear Lagrange triangle



$$\begin{aligned}\varphi_1 &= \zeta_1\\ \varphi_2 &= \zeta_2\\ \varphi_3 &= \zeta_3\end{aligned} \tag{2.66}$$

2. Cubic Hermite triangle



$$
\begin{aligned}
\varphi_1 &= \zeta_1^2\left(\zeta_1+3\zeta_2+3\zeta_3\right)-7\zeta_1\zeta_2\zeta_3 \\
\varphi_2 &= \zeta_1^2\left(x_{21}\zeta_2-x_{13}\zeta_3\right)+\left(x_{13}-x_{21}\right)\zeta_1\zeta_2\zeta_3 \\
\varphi_3 &= \zeta_1^2\left(y_{21}\zeta_2-y_{13}\zeta_3\right)+\left(y_{13}-y_{21}\right)\zeta_1\zeta_2\zeta_3 \\
\varphi_4 &= \zeta_2^2\left(3\zeta_1+\zeta_2+3\zeta_3\right)-7\zeta_1\zeta_2\zeta_3 \\
\varphi_5 &= \zeta_2^2\left(x_{32}\zeta_3-x_{21}\zeta_1\right)+\left(x_{21}-x_{32}\right)\zeta_1\zeta_2\zeta_3 \\
\varphi_6 &= \zeta_2^2\left(y_{32}\zeta_3-y_{21}\zeta_1\right)+\left(y_{21}-y_{32}\right)\zeta_1\zeta_2\zeta_3 \\
\varphi_7 &= \zeta_3^2\left(3\zeta_1+3\zeta_2+\zeta_3\right)-7\zeta_1\zeta_2\zeta_3 \\
\varphi_8 &= \zeta_3^2\left(x_{13}\zeta_1-x_{32}\zeta_2\right)+\left(x_{32}-x_{13}\right)\zeta_1\zeta_2\zeta_3 \\
\varphi_9 &= \zeta_3^2\left(y_{13}\zeta_1-y_{32}\zeta_2\right)+\left(y_{32}-y_{13}\right)\zeta_1\zeta_2\zeta_3 \\
\varphi_{10} &= 27\zeta_1\zeta_2\zeta_3
\end{aligned}
\tag{2.67}
$$

2.6 Derive the element matrix

The variation formulation of Possin equation

$$\int_{\Omega} \nabla v \nabla u dx = \int_{\Omega} v f dx \tag{2.68}$$

let $\{K_i\}_i$ is nonoverlapping triangular domain, and $\cup_{i=1}^{n} K_i = \Omega$, then

$$\int_{\Omega} \nabla v \nabla u dx = \sum_{i=1}^{n} \int_{e_i} \nabla v \nabla u dx \tag{2.69}$$

on each element K_i the global coordinate x, y transfor into local $\zeta_1, \zeta_2, \zeta_3$ the linear Lagrange interpolation of w in the local interpolant $w = N_1\varphi_1 + N_2\varphi_2 + N_3\varphi_3 = \Phi \begin{bmatrix} N_1 \\ N_2 \\ N_3 \end{bmatrix} = \Phi\mathcal{N}$,

$\nabla = Z\nabla_{\zeta_1,\zeta_2,\zeta_3}$, where $Z = \begin{bmatrix} \frac{\partial \zeta_1}{\partial x} & \frac{\partial \zeta_2}{\partial x} & \frac{\partial \zeta_3}{\partial x} \\ \frac{\partial \zeta_1}{\partial y} & \frac{\partial \zeta_2}{\partial y} & \frac{\partial \zeta_3}{\partial y} \end{bmatrix}$

$$\int_{e_i} \nabla v \nabla u dx = \int_{e_i} V \nabla_{\xi} \Phi^T Z^T Z \nabla_{\xi,\eta} \Phi \mathcal{N} \left| J \right| d\zeta_2 d\zeta_3 \tag{2.70}$$

where $\zeta_1 = 1 - \zeta_2 - \zeta_3$.

the local element matrix is $\int_{e_i} \nabla_{\xi,\eta} \Phi^T Z^T Z \nabla_{\xi,\eta} \Phi \left| J \right| d\xi d\eta$

Chapter 3
Numerical Method of Monge-Ampére Equation

We follow the Feng's method, that is adding a vanishing biharmonic term such that the fully non-linear Monge-Ampére equation become regular. The elliptic regularization Monge-Ampére equation:

$$\begin{aligned} -\epsilon\triangle^2 u^\epsilon + \det\left(D^2 u^\epsilon\right) &= f, \text{ in } \Omega \\ u^\epsilon &= g \text{ on } \partial\Omega \\ \triangle u^\epsilon &= \epsilon \text{ on } \partial\Omega \end{aligned} \tag{3.71}$$

where Ω is a bound domain in the $\mathbb{R}^2$ with a smooth boundary $\partial\Omega$, f is a given function

3.1 Linearization Regularization Monge-Ampére equation

the function of regularization Monge-Ampére equation

$$MA[u] = -\epsilon\triangle^2 u + \det\left(D^2 u\right) \tag{3.72}$$

variation of $MA[u]$ is

$$\begin{aligned} D_u MA[u] &= (D_{yy}u)(D_{xx}) + (D_{xx}u)(D_{yy}) - 2(D_{xy}u)(D_{xy}) - \epsilon\triangle^2 \\ &= \nabla\cdot\left(cof(D^2u)\nabla\right) - \epsilon\triangle^2 \end{aligned} \tag{3.73}$$

the linearization regularization Monge-Ampére equation

$$-\epsilon\triangle^2 u + \nabla\cdot\left(cof(D^2u)\nabla u\right) = f \tag{3.74}$$

3.2 Variation formulation

the equivalent variational problem

$$-\int \epsilon\triangle^2 u^\epsilon v dx + \int \nabla\cdot\left(cof(D^2u)\nabla u\right) v dx = \int f v dx \text{ in } \Omega \tag{3.75}$$

the weak formulation of the biharmonic term $-\int \epsilon\triangle^2 u^\epsilon v dx$ and $\triangle u^\epsilon = \epsilon$

$$\begin{aligned} -\epsilon\int_\Omega \triangle^2 u^\epsilon v dx &= \epsilon\int_{\partial\Omega}\epsilon\nabla v\cdot n - \epsilon\int_\Omega \triangle u^\epsilon \triangle v dx \\ &= \epsilon\int_{\partial\Omega}\nabla\epsilon\cdot nv - \epsilon\int_\Omega \triangle u^\epsilon\triangle v dx \\ &= -\epsilon\int_\Omega \triangle u^\epsilon \triangle v dx \end{aligned} \tag{3.76}$$

where $v = 0$ on $\partial\Omega$ and the second boundary condition same as $\triangle u^\epsilon = 0$ on $\partial\Omega$.

the weak formulation of the fully non-linear term

$$\begin{aligned} \int_\Omega \nabla\left(cof\left(D^2u\right)\nabla u\right) v dx &= \int_{\partial\Omega}\left(cof\left(D^2u\right)\nabla u\right)\cdot nv dx - \int_\Omega\left(cof\left(D^2u\right)\nabla u\right)\nabla v dx \\ &= -\int_\Omega\left(cof\left(D^2u\right)\nabla u\right)\nabla v dx \end{aligned} \tag{3.77}$$

So the equivalent variational problem of equation (3.71) is

$$-\epsilon\int_\Omega \triangle u^\epsilon \triangle v dx - \int_\Omega\left(cof\left(D^2u\right)\nabla u\right)\nabla v dx = \int f v dx \text{ in } \Omega \tag{3.78}$$

3.3 Non-linear iteration

For non-linear problem, we usually use iterative method such as fixed-point iteration, Newton's iteration etc.. Iteration method can be classified by the rate of convergence, q-quadratically, q-superlinearly, and q-linearly.

3.3.1 Fixed-Point Iteration

Many non-linear equation are naturally formulated as fixed-point problem

$$x = K(x) \tag{3.79}$$

where K, the fixed-point map may be non-linear. A solution $\hat{x}$ of (3.79) called a fixed point of the map K. The fixed-point iteration is given by

$$x_{n+1} = K(x_n) \tag{3.80}$$

This iterative method is also called non-linear Richardson iteration, Picard iteration, or the method of successive substitution.

3.3.2 Newton's method

The Newton's iteration is

$$x_{n+1} = x_n - F'(x_n)^{-1} F(x_n) \tag{3.81}$$

sometimes the $F'(x_n)^{-1}$ is not easy to find, then we can consider use approximate the term, such as chord method, Shamanskii method or secant method etc..

3.3.3 Non-linear iteration of regularization Monge-Ampére equation

The regularization Monge-Ampére equation

$$F[u] = f + \epsilon \triangle^2 u - \det\left(D^2 u\right) \tag{3.82}$$

the $D_u F[u]$ is

$$D_u F[u] = -\nabla \cdot \left(cof(D^2 u)\nabla\right) + \epsilon \triangle^2 \tag{3.83}$$

Newton's iteration of F

$$u^{n+1} = u^n - D_u F\left[u^n\right]^{-1} F\left[u^n\right] \tag{3.84}$$

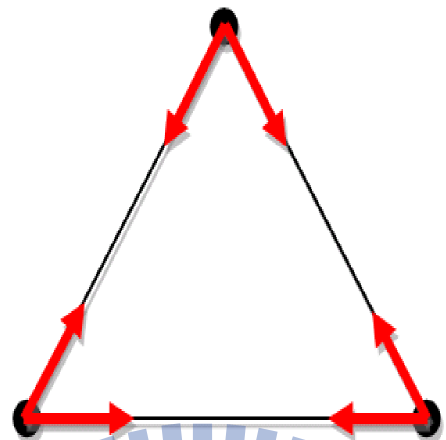


Fig. 3.12. BCIZ truangular element

3.4 Basis function of BCIZ element

To build the necessary technical tools, we shall derive and present a detailed study of the linearization of the elliptic regularization Monge-Ampere equation and its BCIZ finite element approximation. Introduction to the BCIZ element, BCIZ element is conforming element, it can calculus the curvature easily, and its approximation is very well. But the basic BCIZ element has a problem, if the mesh is non-uniform mesh, then the numerical result is lost the accuracy. So many people propose the revise BCIZ element such that numerical result has good approximation on non-uniform mesh.

BCIZ element:

Let $\mathcal{P} = \mathcal{P}_3$. Let $\mathcal{N} = \{N_1, N_2, ..., N_9\}$

In the free-form design problem, we must to consider that the first differential of the solution, so we choose BCIZ finite element approximation. It can easy the calculus the first differential and curve of each element.

The visible degree of freedom of the the element collected in v are

$$v^T = \begin{bmatrix} w_1 & \theta_{x1} & \theta_{y1} & w_2 & \theta_{x2} & \theta_{y2} & w_3 & \theta_{x3} & \theta_{y3} \end{bmatrix} \tag{3.85}$$

where the θ_x and θ_x is consider rotation, that is different from u, under Cartesian coordinate $u = w, u_y = \theta_x, u_x = -\theta_y$.

$$\begin{aligned}
\varphi_1 &= \zeta_1^2 (3 - 2\zeta_1) + 2\zeta_1\zeta_2\zeta_3 \\
\varphi_2 &= -\zeta_1^2 (y_{12}\zeta_2 + y_{13}\zeta_3) - \frac{1}{2} (y_{12} + y_{13}) \zeta_1\zeta_2\zeta_3 \\
\varphi_3 &= \zeta_1^2 (x_{12}\zeta_2 + x_{13}\zeta_3) + \frac{1}{2} (x_{12} + x_{13}) \zeta_1\zeta_2\zeta_3 \\
\varphi_4 &= \zeta_2^2 (3 - 2\zeta_2) + 2\zeta_1\zeta_2\zeta_3 \\
\varphi_5 &= -\zeta_2^2 (y_{23}\zeta_3 + y_{21}\zeta_1) - \frac{1}{2} (y_{23} + y_{21}) \zeta_1\zeta_2\zeta_3 \\
\varphi_6 &= \zeta_2^2 (x_{23}\zeta_3 + x_{21}\zeta_1) + \frac{1}{2} (x_{23} + x_{21}) \zeta_1\zeta_2\zeta_3 \\
\varphi_7 &= \zeta_3^2 (3 - 2\zeta_3) + 2\zeta_1\zeta_2\zeta_3 \\
\varphi_8 &= -\zeta_3^2 (y_{31}\zeta_1 + y_{32}\zeta_2) - \frac{1}{2} (y_{31} + y_{32}) \zeta_1\zeta_2\zeta_3 \\
\varphi_9 &= \zeta_1^2 (x_{31}\zeta_1 + x_{32}\zeta_2) + \frac{1}{2} (x_{31} + x_{32}) \zeta_1\zeta_2\zeta_3
\end{aligned} \tag{3.86}$$

where ζ_i are triangular coordinate.

Let

$$\Phi = \left[\begin{array}{ccccccccc} \varphi_1 & \varphi_2 & \varphi_3 & \varphi_4 & \varphi_5 & \varphi_6 & \varphi_7 & \varphi_8 & \varphi_9 \end{array} \right] \tag{3.87}$$

and

$$w\left(\zeta_1, \zeta_2, \zeta_3\right) = \Phi v \tag{3.88}$$

Derive the element matrix Derive the element matrix of the variation equation (3.78) with BCIZ element

3.4.1 The linearization of non-linear term and element matrix

Change coordinate from the global coordinate to the local coordinate, the relationship of $\nabla_{x,y}$ and $\nabla_{\zeta_1,\zeta_2,\zeta_3}$ is

$$\left[\begin{array}{c} \frac{\partial w}{\partial x} \\ \frac{\partial w}{\partial y} \end{array} \right] = Z \left[\begin{array}{c} \frac{\partial w}{\partial \zeta_1} \\ \frac{\partial w}{\partial \zeta_2} \\ \frac{\partial w}{\partial \zeta_3} \end{array} \right] \tag{3.89}$$

where

$$Z = \frac{1}{2A} \left[\begin{array}{ccc} -y_{32} & -y_{13} & -y_{21} \\ x_{32} & x_{13} & x_{21} \end{array} \right] \tag{3.90}$$

the w use BCIZ element to approximation, $w = \Phi v$

$$\left[\begin{array}{c} \frac{\partial w}{\partial \zeta_1} \\ \frac{\partial w}{\partial \zeta_2} \\ \frac{\partial w}{\partial \zeta_3} \end{array} \right] = \left[\begin{array}{c} \frac{\partial \Phi}{\partial \zeta_1} \\ \frac{\partial \Phi}{\partial \zeta_2} \\ \frac{\partial \Phi}{\partial \zeta_3} \end{array} \right] v = Bv \tag{3.91}$$

where

$$B = \begin{bmatrix} 2\zeta_1(3-2\zeta_1)-2\zeta_1^2+2\zeta_2\zeta_3 & 2\zeta_1\zeta_3 & 2\zeta_1\zeta_2 \\ -2\zeta_1(-y_{21}\zeta_2-y_{31}\zeta_3)+\frac{1}{2}(y_{21}+y_{31})\zeta_2\zeta_3 & \zeta_1^2 y_{21}+\frac{1}{2}(y_{21}+y_{31})\zeta_1\zeta_3 & \zeta_1^2 y_{31}+\frac{1}{2}(y_{21}+y_{31})\zeta_1\zeta_2 \\ 2\zeta_1(-x_{21}\zeta_2-x_{31}\zeta_3)-\frac{1}{2}(x_{21}+x_{31})\zeta_2\zeta_3 & -\zeta_1^2 x_{21}-\frac{1}{2}(x_{21}+x_{31})\zeta_1\zeta_3 & -\zeta_1^2 x_{31}-\frac{1}{2}(x_{21}+x_{31})\zeta_1\zeta_2 \\ 2\zeta_2\zeta_3 & 2\zeta_2(3-2\zeta_2)-2\zeta_2^2+2\zeta_1\zeta_3 & 2\zeta_1\zeta_2 \\ -\zeta_2^2 y_{21}+\frac{1}{2}(y_{32}-y_{21})\zeta_2\zeta_3 & -2\zeta_2(-y_{32}\zeta_3+y_{21}\zeta_1)+\frac{1}{2}(y_{32}-y_{21})\zeta_1\zeta_3 & \zeta_2^2 y_{32}+\frac{1}{2}(y_{32}-y_{21})\zeta_1\zeta_2 \\ \zeta_2^2 x_{21}+\frac{1}{2}(-x_{32}+x_{21})\zeta_2\zeta_3 & 2\zeta_2(-x_{32}\zeta_3+x_{21}\zeta_1)-\frac{1}{2}(x_{32}-x_{21})\zeta_1\zeta_3 & -\zeta_2^2 x_{32}-\frac{1}{2}(x_{32}-x_{21})\zeta_1\zeta_2 \\ 2\zeta_2\zeta_3 & 2\zeta_1\zeta_3 & 2\zeta_3(3-2\zeta_3)-2\zeta_3^2+2\zeta_1\zeta_2 \\ -\zeta_3^2 y_{31}-\frac{1}{2}(y_{31}+y_{32})\zeta_2\zeta_3 & -\zeta_3^2 y_{32}-\frac{1}{2}(y_{31}+y_{32})\zeta_1\zeta_3 & -2\zeta_3(y_{31}\zeta_1+y_{32}\zeta_2)-\frac{1}{2}(y_{31}+y_{32})\zeta_1\zeta_2 \\ \zeta_3^2 x_{31}+\frac{1}{2}(x_{31}+x_{32})\zeta_2\zeta_3 & \zeta_3^2 x_{32}+\frac{1}{2}(x_{31}+x_{32})\zeta_1\zeta_3 & 2\zeta_3(x_{31}\zeta_1+x_{32}\zeta_2)+\frac{1}{2}(x_{31}+x_{32})\zeta_1\zeta_2 \end{bmatrix}^T \tag{3.92}$$

so the $\nabla v = ZBv$, then the element matrix of $-\int_\Omega (cof\,(D^2u)\,\nabla u)\,\nabla v dx$ is

$$K_{e1} = -\int_e B^T Z^T C_1 ZB\,|J|\,d\zeta_1 d\zeta_2 \tag{3.93}$$

where C_1 is $cof\,(D^2u)$

3.4.2 biharmonic term

for biharmonic term $D\Delta^2 u$ where $D = \frac{c}{12(1-\nu^2)}$

$$\Delta^2 = \begin{bmatrix} \frac{\partial^2}{\partial x^2} & \frac{\partial^2}{\partial y^2} & 2\frac{\partial^2}{\partial x\partial y} \end{bmatrix} \begin{bmatrix} 1 & \nu & 0 \\ \nu & 1 & 0 \\ 0 & 0 & \frac{(1-\nu)}{2} \end{bmatrix} \begin{bmatrix} \frac{\partial^2}{\partial x^2} \\ \frac{\partial^2}{\partial y^2} \\ 2\frac{\partial^2}{\partial x\partial y} \end{bmatrix} \tag{3.94}$$

Let $C_2 = \begin{bmatrix} 1 & \nu & 0 \\ \nu & 1 & 0 \\ 0 & 0 & \frac{(1-\nu)}{2} \end{bmatrix}$

The second derivative of a function $w(\zeta_1,\zeta_2,\zeta_3)$ with respect to x or y from (2.59) and application of the chain rule:

$$\begin{aligned}\frac{\partial^2 w}{\partial x^2} &= \frac{1}{2\bar{A}}\left(\frac{\partial}{\partial\zeta_1}\frac{\partial w}{\partial x}y_{23}+\frac{\partial}{\partial\zeta_2}\frac{\partial w}{\partial x}y_{31}+\frac{\partial}{\partial\zeta_3}\frac{\partial w}{\partial x}y_{12}\right) \\ &= \frac{1}{4\bar{A}^2}\left(\frac{\partial^2 w}{\partial\zeta_1^2}y_{23}^2+\frac{\partial^2 w}{\partial\zeta_2^2}y_{31}^2+\frac{\partial^2 w}{\partial\zeta_3^2}y_{12}^2+2\frac{\partial^2 w}{\partial\zeta_1\partial\zeta_2}y_{31}y_{23}+2\frac{\partial^2 w}{\partial\zeta_1\partial\zeta_3}y_{12}y_{23}+2\frac{\partial^2 w}{\partial\zeta_2\partial\zeta_3}y_{12}y_{31}\right)\end{aligned} \tag{3.95}$$

$$\begin{aligned}\frac{\partial^2 w}{\partial y^2} &= \frac{1}{2\bar{A}}\left(\frac{\partial}{\partial\zeta_1}\frac{\partial w}{\partial y}x_{32}+\frac{\partial}{\partial\zeta_2}\frac{\partial w}{\partial y}x_{13}+\frac{\partial}{\partial\zeta_3}\frac{\partial w}{\partial y}x_{21}\right) \\ &= \frac{1}{4\bar{A}^2}\left(\frac{\partial^2 w}{\partial\zeta_1^2}x_{32}^2+\frac{\partial^2 w}{\partial\zeta_2^2}x_{13}^2+\frac{\partial^2 w}{\partial\zeta_3^2}x_{21}^2+2\frac{\partial^2 w}{\partial\zeta_1\partial\zeta_2}x_{13}x_{32}+2\frac{\partial^2 w}{\partial\zeta_1\partial\zeta_3}x_{21}x_{32}+2\frac{\partial^2 w}{\partial\zeta_2\partial\zeta_3}x_{21}x_{13}\right)\end{aligned} \tag{3.96}$$

$$\begin{aligned}\frac{\partial^2 w}{\partial x\partial y} &= \frac{1}{2\bar{A}}\left(\frac{\partial}{\partial\zeta_1}\frac{\partial w}{\partial y}y_{23}+\frac{\partial}{\partial\zeta_2}\frac{\partial w}{\partial y}y_{31}+\frac{\partial}{\partial\zeta_3}\frac{\partial w}{\partial y}y_{12}\right) \\ &= \frac{1}{4\bar{A}^2}\left(\frac{\partial^2 w}{\partial\zeta_1^2}x_{32}y_{23}+\frac{\partial^2 w}{\partial\zeta_2^2}x_{13}y_{31}+\frac{\partial^2 w}{\partial\zeta_3^2}x_{21}y_{12}+\frac{\partial^2 w}{\partial\zeta_1\partial\zeta_2}(x_{13}y_{23}+x_{32}y_{31})\right. \\ &\quad \left.+\frac{\partial^2 w}{\partial\zeta_1\partial\zeta_3}(x_{21}y_{23}+x_{32}y_{12})+\frac{\partial^2 w}{\partial\zeta_2\partial\zeta_3}(x_{21}y_{31}+x_{13}y_{12})\right)\end{aligned} \tag{3.97}$$

which matrix form is

$$\begin{bmatrix}\frac{\partial^2 w}{\partial x^2}\\ \frac{\partial^2 w}{\partial y^2}\\ 2\frac{\partial^2 w}{\partial x\partial y}\end{bmatrix} = \frac{1}{4\bar{A}^2}\begin{bmatrix} y_{23}^2 & x_{32}^2 & 2x_{32}y_{23}\\ y_{31}^2 & x_{13}^2 & 2x_{13}y_{31}\\ y_{12}^2 & x_{21}^2 & 2x_{21}y_{12}\\ 2y_{23}y_{31} & 2x_{32}x_{13} & 2(x_{32}y_{31}+x_{13}y_{23})\\ 2y_{31}y_{12} & 2x_{13}x_{21} & 2(x_{13}y_{12}+x_{21}y_{31})\\ 2y_{12}y_{23} & 2x_{21}x_{32} & 2(x_{21}y_{32}+x_{32}y_{12})\end{bmatrix}^T \begin{bmatrix}\frac{\partial^2 w}{\partial\zeta_1^2}\\ \frac{\partial^2 w}{\partial\zeta_2^2}\\ \frac{\partial^2 w}{\partial\zeta_3^2}\\ \frac{\partial^2 w}{\partial\zeta_1\partial\zeta_2}\\ \frac{\partial^2 w}{\partial\zeta_2\partial\zeta_3}\\ \frac{\partial^2 w}{\partial\zeta_1\partial\zeta_3}\end{bmatrix} \tag{3.98}$$

let

$$S = \frac{1}{4\bar{A}^2}\begin{bmatrix} y_{23}^2 & x_{32}^2 & 2x_{32}y_{23}\\ y_{31}^2 & x_{13}^2 & 2x_{13}y_{31}\\ y_{12}^2 & x_{21}^2 & 2x_{21}y_{12}\\ 2y_{23}y_{31} & 2x_{32}x_{13} & 2(x_{32}y_{31}+x_{13}y_{23})\\ 2y_{31}y_{12} & 2x_{13}x_{21} & 2(x_{13}y_{12}+x_{21}y_{31})\\ 2y_{12}y_{23} & 2x_{21}x_{32} & 2(x_{21}y_{32}+x_{32}y_{12})\end{bmatrix}^T$$

the w use BCIZ element to approximation, $w = \Phi v$

$$\begin{bmatrix} \frac{\partial^2 w}{\partial \zeta_1^2} \\ \frac{\partial^2 w}{\partial \zeta_2^2} \\ \frac{\partial^2 w}{\partial \zeta_3^2} \\ \frac{\partial^2 w}{\partial \zeta_1 \partial \zeta_2} \\ \frac{\partial^2 w}{\partial \zeta_2 \partial \zeta_3} \\ \frac{\partial^2 w}{\partial \zeta_1 \partial \zeta_3} \end{bmatrix} = \Psi v \tag{3.99}$$

where

$$\Psi = \begin{bmatrix} 6-12\zeta_1 & 0 & 0 & 2\zeta_3 & 2\zeta_1 & 2\zeta_2 \\ 2y_{21}\zeta_2 + 2y_{31}\zeta_3 & 0 & 0 & 2y_{21}\zeta_1 + \frac{1}{2}(y_{21}+y_{31})\zeta_3 & \frac{1}{2}(y_{21}+y_{31})\zeta_1 & 2y_{31}\zeta_1 + \frac{1}{2}(y_{21}+y_{31})\zeta_2 \\ -2x_{21}\zeta_2 - 2x_{31}\zeta_3 & 0 & 0 & -2x_{21}\zeta_1 - \frac{1}{2}(x_{21}+x_{31})\zeta_3 & -\frac{1}{2}(x_{21}+x_{31})\zeta_1 & -2x_{31}\zeta_1 - \frac{1}{2}(x_{21}+x_{31})\zeta_2 \\ 0 & 6-12\zeta_2 & 0 & 2\zeta_3 & 2\zeta_1 & 2\zeta_2 \\ 0 & 2y_{32}\zeta_3 - 2y_{21}\zeta_1 & 0 & -2y_{21}\zeta_2 + \frac{1}{2}(y_{32}-y_{21})\zeta_3 & 2y_{32}\zeta_2 + \frac{1}{2}(y_{32}-y_{21})\zeta_1 & \frac{1}{2}(y_{32}-y_{21})\zeta_2 \\ 0 & -2x_{32}\zeta_2 + 2x_{21}\zeta_3 & 0 & 2x_{21}\zeta_2 - \frac{1}{2}(x_{32}-x_{21})\zeta_3 & 2x_{32}\zeta_2 - \frac{1}{2}(x_{32}-x_{21})\zeta_1 & -\frac{1}{2}(x_{32}-x_{21})\zeta_2 \\ 0 & 0 & 6-12\zeta_3 & 2\zeta_3 & 2\zeta_1 & 2\zeta_2 \\ 0 & 0 & -2y_{31}\zeta_2 - 2y_{32}\zeta_2 & -\frac{1}{2}(y_{31}-y_{32})\zeta_3 & -2y_{32}\zeta_3 - \frac{1}{2}(y_{31}-y_{32})\zeta_1 & -2y_{31}\zeta_3 - \frac{1}{2}(y_{31}-y_{32})\zeta_2 \\ 0 & 0 & 2x_{31}\zeta_1 + 2x_{32}\zeta_2 & \frac{1}{2}(x_{31}+x_{32})\zeta_3 & 2x_{32}\zeta_3 + \frac{1}{2}(x_{31}+x_{32})\zeta_1 & 2x_{31}\zeta_3 + \frac{1}{2}(x_{31}+x_{32})\zeta_2 \end{bmatrix}^T \tag{3.100}$$

so the $\triangle v = S\Psi v$, then the element matrix of $-\epsilon \int_\Omega \triangle u^\epsilon \triangle v dx$ is

$$K_{e2} = -\int_e^T \Psi^T S^T C_2 S \Psi \left| J \right| d\zeta_1 d\zeta_2 \tag{3.101}$$

In this thesis, we will follow this algorithm

1. Given a initial u_0, ϵ and tolerance T

2. Fixed ϵ

3. Newton's iteration if $\|u^{n+1} - u^n\| \leq T$ then it converge, if not the iteration is diverge

4. if $\epsilon \leq h^2$ where h is mesh size then out the algorithm

5. let $\epsilon = \epsilon/c$ where c is a constant, then go to 2

Chapter 4
Numerical Study

The numerical result will be given, These are three part of this chapter: Poisson equation ,biharmonic equation and Monge-Ampére equation.

4.1 Poisson Equation

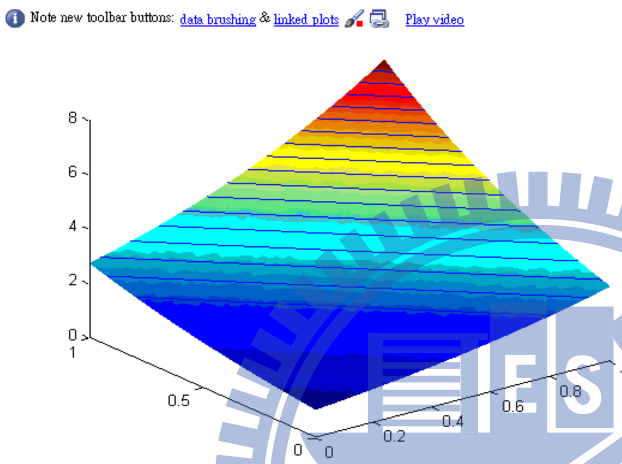
Poisson Equation:

$$\begin{aligned} \triangle u &= f \text{ in } \Omega \\ u|_{\partial\Omega} &= g \end{aligned}$$

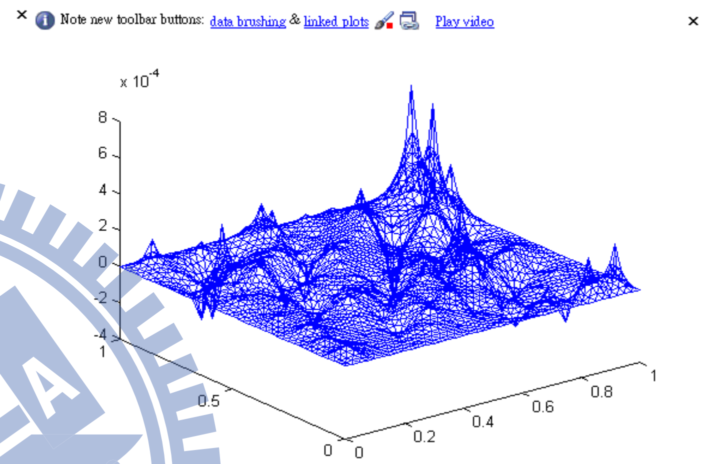
where f and g are obtained from a given analytical solution u. We use Linear element and BCIZ element to approximate the Poisson equation. We compute the Poisson equation on different mesh size. Our calculation domain is $[0,1] \times [0,1]$. The boudary condition are Dirchlet type.

4.1.1 Example:

The analytical solution $u = e^{x+y}$, $f = 2e^{x+y}$ and $g = e^{x+y}$. we use linear element to approximate the Poisson equation in this case.



Computed solution u_h



Error

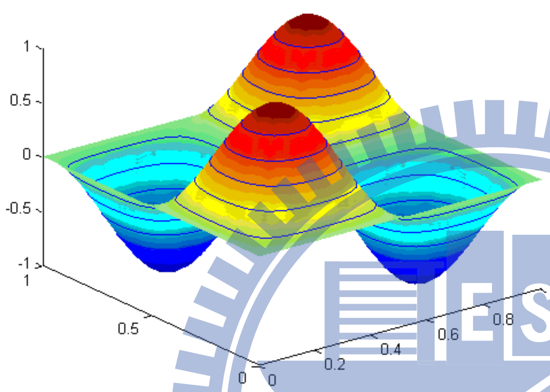
h	$\|u^0 - u_h\|_\infty$	$\|u^0 - u_h\|_{L^2}$	$\|u^0 - u_h\|_{H^1}$
$1/10$	5.97E-03	8.28E-04	2.04E-01
$1/20$	1.98E-03	2.21E-04	1.02E-01
$1/40$	6.28E-04	5.68E-05	5.10E-02
$1/80$	1.91E-04	1.44E-05	2.55E-02
$1/160$	5.65E-05	3.61E-06	1.28E-02

Table 1. Change of $\|u - u_h\|$ w.r.t. h

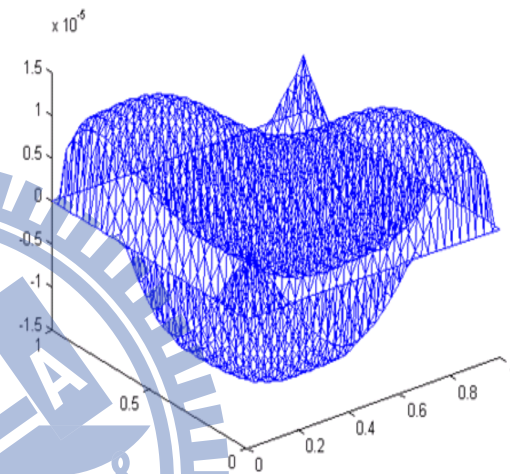
The convergence rate of L^2 norm is second order and H^1 norm is first order. This result is same as error estimates of the biharmonic equation using BCIZ element approximation.

4.1.2 Example:

The analytical solution $u = \sin(2\pi x)\sin(2\pi y)$, $f = -8\pi^2 \sin(2\pi x)\sin(2\pi y)$ and $g = 0$. we use BCIZ element to approximate the Poisson equation in this case.



Computed solution u_h



Error

h	$\|u^0 - u_h\|_\infty$	$\|u^0 - u_h\|_{L^2}$	$\|u^0 - u_h\|_{H^1}$
2^{-3}	1.23E-3	8.03E-4	4.37E-2
2^{-4}	1.67E-4	7.54E-5	4.19E-3
2^{-5}	1.32E-5	7.24E-6	5.32E-4
2^{-6}	8.83E-7	7.85E-7	1.18E-4

Table 2. Change of $\|u - u_h\|$ w.r.t. h

The convergence rate of L^2 norm is third order and H^1 norm is second order. This result is same as error estimates of the biharmonic equation using BCIZ element approximation.

4.2 Biharmonic Equation

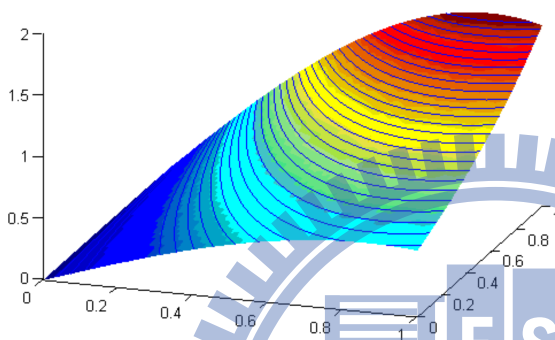
Biharmonic Equation:

$$\begin{aligned} \triangle^2 u &= f \text{ in } \Omega \\ u|_{\partial\Omega} &= g \\ \nabla u \cdot n|_{\partial\Omega} &= h \end{aligned}$$

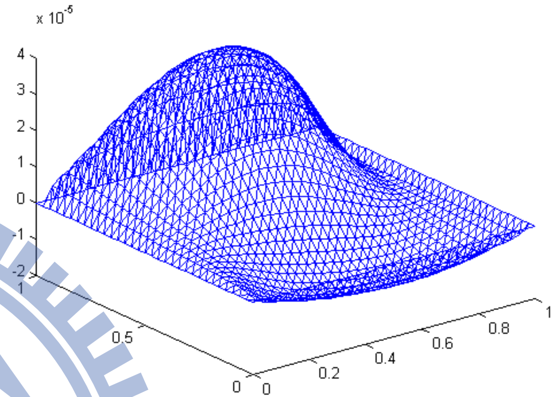
where f, g and h are obtained from a given analytical solution u. We use BCIZ element to approximate the Biharmonic equation. We compute the Biharmonic equation on different mesh size. Our calculation domain is $[0,1] \times [0,1]$. The boudary condition are Dirichlet type and Neumann type. Because of the Biharmonic equation is fourth order equation, so the approximation of linear element maybe not have a high accuracy.

4.2.1 Example:

The analytical solution $u = x\cos(x)e^y$, $f = 0$,$g = x\cos(x)e^y$ and $\nabla u = \begin{bmatrix} \cos(x)e^y - x\sin(x)e^y \\ x\cos(x)e^y \end{bmatrix}$.



Computed solution u_h



Error

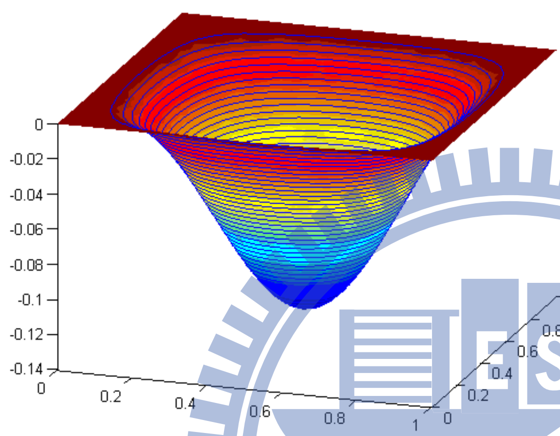
h	$\|u^0 - u_h\|_\infty$	$\|u^0 - u_h\|_{L^2}$	$\|u^0 - u_h\|_{H^2}$
2^{-2}	0.002971694	0.001206162	0.499530706
2^{-3}	0.00065049	0.000273676	0.242128734
2^{-4}	0.000155176	6.32838E-05	0.119339924
2^{-5}	3.78495E-05	1.51706E-05	0.059270777
2^{-6}	9.32505E-06	3.71417E-06	0.02954046
2^{-7}	2.31532E-06	9.1924E-07	0.014747201

Table 3. Change of $\|u - u_h\|$ w.r.t. h

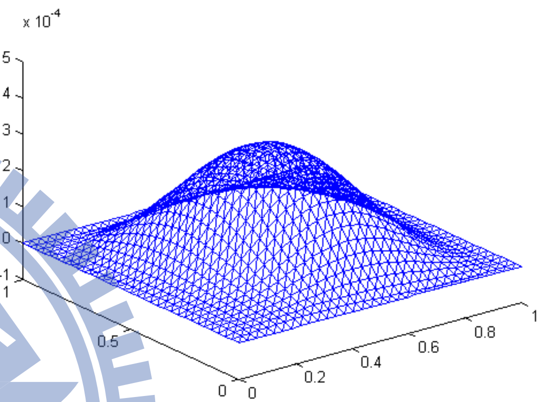
The convergence rate of L^2 norm is second order and H^2 norm is first order. This result is same as error estimates of the biharmonic equation using BCIZ element approximation.

4.2.2 Example:

The analytical solution $u = (\cos(2\pi x) - 1)(y^2 - 2y^3 + y^4)$, $f = 0$,$g = x\cos(x)e^y$ and $h = 0$.



Computed solution u_h



Error

h	$\|u^0 - u_h\|_\infty$	$\|u^0 - u_h\|_{L^2}$	$\|u^0 - u_h\|_{H^2}$
2^{-2}	0.01171965	0.004154116	0.645238157
2^{-3}	0.004269365	0.001568274	0.307797116
2^{-4}	0.001169401	0.000445123	0.150571686
2^{-5}	0.000302368	0.000116688	0.074556331
2^{-6}	7.67083E-05	2.97733E-05	0.037111559
2^{-7}	1.93091E-05	7.5141E-06	0.018516197

Table 4. Change of $\|u - u_h\|$ w.r.t. h

The convergence rate of L^2 norm is second order and H^2 norm is first order. This result is same as error estimates of the biharmonic equation using BCIZ element approximation.

4.3 Monge-Ampére

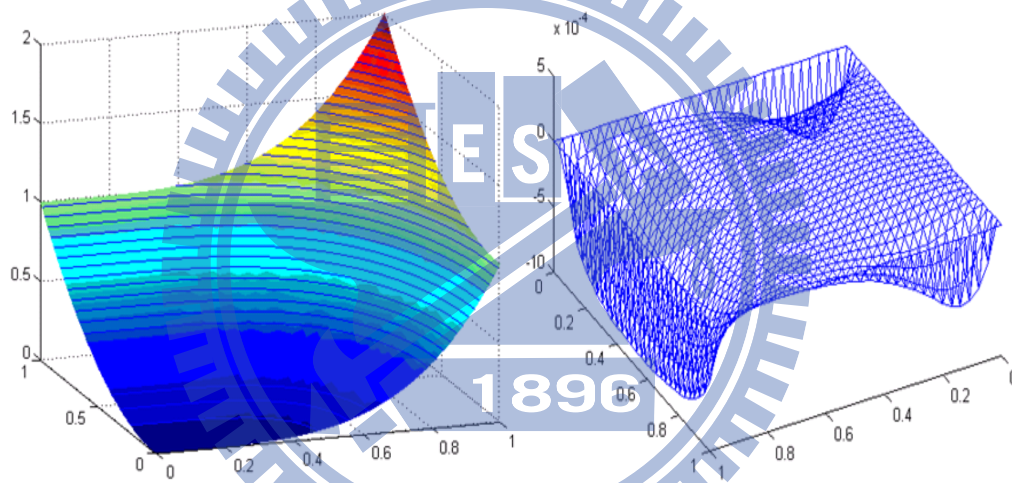
Regularization Monge-Ampére Equation:

$$
\begin{aligned}
-\epsilon\triangle^2 u^\epsilon + \det\left(D^2 u^\epsilon\right) &= f, \text{ in } \Omega \\
u^\epsilon &= g \text{ on } \partial\Omega \\
\triangle u^\epsilon &= \epsilon \text{ on } \partial\Omega
\end{aligned}
$$

where f and g are obtained from a given analytical solution u. We use BCIZ element to approximate the Monge-Ampére equation. We compute the Poisson equation on different parameter ϵ with fixed mesh size h. The boudary condition are Dirchlet type. In this section, we provide several 2-D numerical experiments of BCIZ element. And the initial condition is given by zero. The ϵ start from 1 to h^2.

4.3.1 Example:

This test, we calculus $\|u^0 - u_h^\epsilon\|$ for fixed mesh size $h = 2^{-8}$, while varying ϵ in order to approximate $\|u^0 - u^\epsilon\|$. We use BCIZ element and set to solve problem (3.71) with the analytical solution $u = x^4 + y^2$, $f = 24x^2$ and $g = x^4 + y^2$, Our calculation domain Ω is $[0,1] \times [0,1]$.



ϵ	$\|u^0 - u_h^\epsilon\|_\infty$	$\|u^0 - u_h^\epsilon\|_{L^2}$	$\|u^0 - u_h^\epsilon\|_{H^2}$	iter
1	3.05E-1	1.61E-1	5.32	6
2^{-2}	2.30E-1	1.21E-1	4.67	10
2^{-4}	1.13E-1	5.71E-2	3.63	10
2^{-6}	4.23E-2	1.90E-2	2.71	8
2^{-8}	1.45E-2	5.67E-3	1.99	8
2^{-10}	4.50E-3	1.60E-3	1.44	8
2^{-12}	1.29E-3	4.33E-4	1.03	9
2^{-14}	3.48E-4	1.13E-4	7.33E-1	10

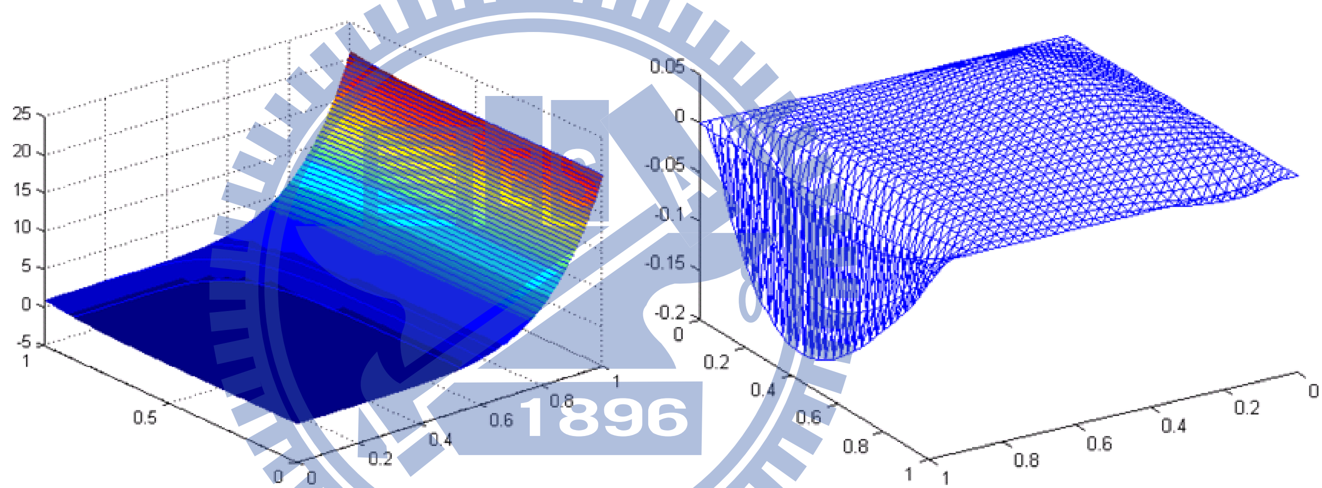
Table 5. Change of $\|u^0 - u_h^\epsilon\|$ w.r.t. ϵ $(h = 2^{-8})$

ϵ	$\frac{\left\|u^0-u_h^\epsilon\right\|_{L^2}}{\epsilon}$	$\frac{\left\|u^0-u_h^\epsilon\right\|_{H^2}}{\sqrt[4]{\epsilon}}$
1	0.160842924	5.319207667
2^{-2}	0.482036757	6.609013597
2^{-4}	0.913779843	7.268095576
2^{-6}	1.218354047	7.670044224
2^{-8}	1.450567862	7.955838131
2^{-10}	1.637202829	8.133261803
2^{-12}	1.772707203	8.235401948
2^{-14}	1.85193153	8.290262971

Table 6. Change of $\|u^0 - u_h^\epsilon\|$ w.r.t. ϵ $(h = 2^{-8})$

4.3.2 Example:

This test, we calculus $\|u^0 - u_h^\epsilon\|$ for fixed mesh size $h = 2^{-8}$, while varying ϵ in order to approximate $\|u^0 - u^\epsilon\|$. We use BCIZ element and set to solve problem (3.71) with the analytical solution $u = 20x^6 + y^6$, $f = 18000x^4y^4$ and $g = 20x^6 + y^6$, Our calculation domain Ω is $[0,1] \times [0,1]$.



solution error

ϵ	$\|u^0 - u_h^\epsilon\|_\infty$	$\|u^0 - u_h^\epsilon\|_{L^2}$	$\|u^0 - u_h^\epsilon\|_{H^2}$	iter
4	6.43	2.88	177.28	5
1	5.22	2.17	167.40	9
2^{-2}	3.31	1.19	148.99	10
2^{-4}	1.79	6.20E-1	125.31	10
2^{-6}	8.72E-1	2.57E-1	101.97	10
2^{-8}	4.02E-1	9.17E-2	82.10	10
2^{-10}	1.80E-1	3.22E-2	65.25	11
2^{-12}	7.93E-2	1.11E-2	51.21	13
2^{-14}	3.45E-2	3.74E-3	39.77	20

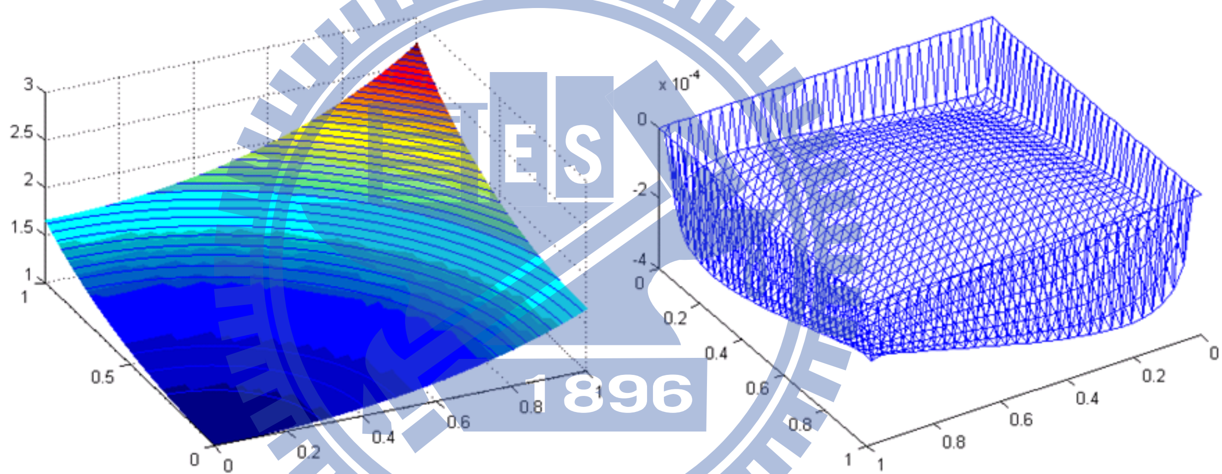
Table 7. Change of $\|u - u_h^\epsilon\|$ w.r.t. ϵ $(h = 2^{-8})$

ϵ	$\frac{\left\|u^0-u_h^\epsilon\right\|_{L^2}}{\epsilon}$	$\frac{\left\|u^0-u_h^\epsilon\right\|_{H^2}}{\sqrt[4]{\epsilon}}$
4	0.720510991	125.3567907
1	2.165121991	167.3973101
2^{-2}	4.758836554	210.7002103
2^{-4}	9.926086663	250.6286892
2^{-6}	16.4590354	288.4238466
2^{-8}	23.47587569	328.3874835
2^{-10}	32.93384144	369.1038869
2^{-12}	45.33800541	409.6430813
2^{-14}	61.24535142	449.9969415

Table 8. Change of $\|u - u_h^\epsilon\|$ w.r.t. ϵ $(h = 2^{-8})$

4.3.3 Example:

This test, we calculus $\|u^0 - u_h^\epsilon\|$ for fixed mesh size $h = 2^{-8}$, while varying ϵ in order to approximate $\|u^0 - u^\epsilon\|$. We use BCIZ element and set to solve problem (3.71) with the analytical solution $u = e^{\frac{x^2+y^2}{2}}$, $f = (1 + x^2 + y^2)\, e^{x^2+y^2}$ and $g = e^{\frac{x^2+y^2}{2}}$, Our calculation domain Ω is $[0, 1] \times [0, 1]$.



solution error

ϵ	$\|u^0 - u_h^\epsilon\|_\infty$	$\|u^0 - u_h^\epsilon\|_{L^2}$	$\|u^0 - u_h^\epsilon\|_{H^2}$	iter
1	1.78E-1	1.01E-1	3.03	29
2^{-2}	1.41E-1	8.17E-2	2.72	48
2^{-4}	7.14E-2	4.45E-2	2.13	38
2^{-6}	2.26E-2	1.56E-2	1.57	9
2^{-8}	6.30E-3	4.52E-3	1.13	8
2^{-10}	1.81E-3	1.22E-3	8.00E-1	8
2^{-12}	5.00E-4	3.16E-4	5.67E-1	8
2^{-14}	1.33E-4	8.00E-5	4.01E-1	9

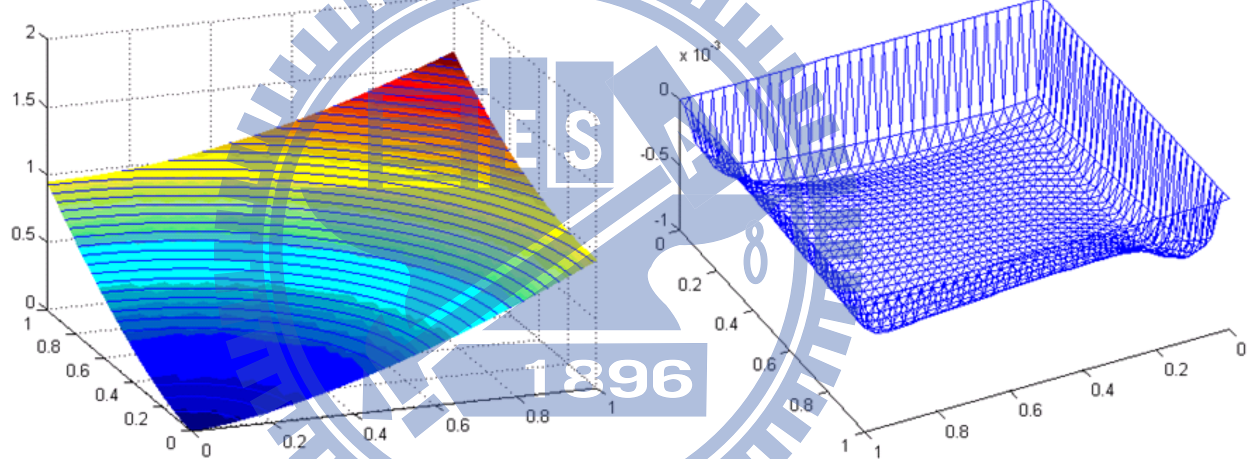
Table 9. Change of $\|u - u_h^\epsilon\|$ w.r.t. ϵ $(h = 2^{-8})$

ϵ	$\frac{\left\|u^0-u_h^\epsilon\right\|_{L^2}}{\epsilon}$	$\frac{\left\|u^0-u_h^\epsilon\right\|_{H^2}}{\sqrt[4]{\epsilon}}$
1	0.100518486	3.026105949
2^{-2}	0.32688765	3.840967403
2^{-4}	0.71238757	4.259311736
2^{-6}	0.99970787	4.434163657
2^{-8}	1.157804247	4.506118399
2^{-10}	1.24700059	4.527808719
2^{-12}	1.294427725	4.534697517
2^{-14}	1.309904805	4.53757113

Table 10. Change of $\|u - u_h^\epsilon\|$ w.r.t. ϵ ($h = 2^{-8}$)

4.3.4 Example:

This test, we calculus $\|u^0 - u_h^\epsilon\|$ for fixed mesh size $h = 2^{-8}$, while varying ϵ in order to approximate $\|u^0 - u^\epsilon\|$. We use BCIZ element and set to solve problem (3.71) with the analytical solution $u = \frac{2\sqrt{2}\left(x^2+y^2\right)^{\frac{3}{4}}}{3}$, $f = \frac{1}{\sqrt{x^2+y^2}}$ and $g = \frac{2\sqrt{2}\left(x^2+y^2\right)^{\frac{3}{4}}}{3}$, Our calculation domain Ω is $[0,1]\times[0,1]$. Where f has a singular point at $(0,0)$.



solution error

ϵ	$\|u^0 - u_h^\epsilon\|_\infty$	$\|u^0 - u_h^\epsilon\|_{L^2}$	$\|u^0 - u_h^\epsilon\|_{H^2}$	iter
1	1.41E-1	7.89E-2	2.16	5
2^{-2}	1.23E-1	6.93E-2	2.01	19
2^{-4}	7.59E-2	4.48E-2	1.64	9
2^{-6}	2.78E-2	1.81E-2	1.22	10
2^{-8}	8.01E-3	5.57E-3	8.95E-1	8
2^{-10}	2.12E-3	1.55E-3	6.51E-1	8
2^{-12}	5.57E-4	4.08E-4	4.68E-1	9
2^{-14}	1.44E-4	1.04E-4	3.34E-1	11

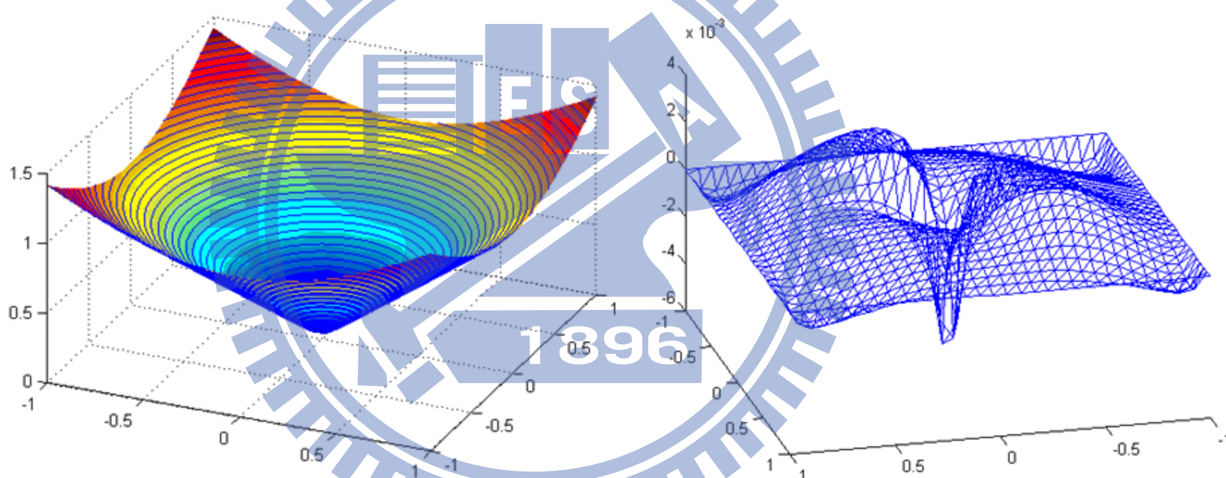
Table 11. Change of $\|u - u_h^\epsilon\|$ w.r.t. ϵ ($h = 2^{-8}$)

ϵ	$\frac{\left\|u^0-u_h^\epsilon\right\|_{L^2}}{\epsilon}$	$\frac{\left\|u^0-u_h^\epsilon\right\|_{H^2}}{\sqrt[4]{\epsilon}}$
1	0.078906702	2.164987734
2^{-2}	0.277330363	2.843363264
2^{-4}	0.717285822	3.275360039
2^{-6}	1.15590166	3.442870121
2^{-8}	1.425652095	3.57845558
2^{-10}	1.583580447	3.680240848
2^{-12}	1.671089711	3.747674023
2^{-14}	1.709884392	3.773972909

Table 12. Change of $\|u-u_h^\epsilon\|$ w.r.t. ϵ $(h=2^{-8})$

4.3.5 Example:

This test, we calculus $\|u^0 - u_h^\epsilon\|$ for fixed mesh size $h = 2^{-8}$, while varying ϵ in order to approximate $\|u^0 - u^\epsilon\|$. We use BCIZ element and set to solve problem (3.71) with the analytical solution $u = \sqrt{x^2 + y^2}$, $f = \begin{cases} 0 & \text{if } (x,y) \neq (0,0) \\ ? & \text{if } (x,y) = (0,0) \end{cases}$ and $g = \sqrt{x^2+y^2}$, Our calculation domain Ω is $[-1,1] \times [-1,1]$. Where f has a singular point at $(0,0)$, and our guess the value of f at $(0,0)$ is 3δ.



solution error

ϵ	$\|u^0 - u_h^\epsilon\|_\infty$	$\|u^0 - u_h^\epsilon\|_{L^2}$	$\|u^0 - u_h^\epsilon\|_{H^2}$	iter
1	8.17E-1	6.00E-1	8.65	5
2^{-2}	5.51E-1	3.72E-1	8.82	9
2^{-4}	2.48E-1	1.42E-1	9.27	10
2^{-6}	9.77E-2	5.31E-2	9.72	11
2^{-8}	4.151E-2	2.56E-2	10.09	14
2^{-10}	2.08E-2	1.49E-2	10.32	19
2^{-12}	9.28E-3	6.92E-3	10.51	28
2^{-13}	3.59E-3	1.95E-3	10.66	21

Table 13. Change of $\|u - u_h^\epsilon\|$ w.r.t. ϵ ($h = 1/127$)

ϵ	$\frac{\left\|u^0-u_h^\epsilon\right\|_{L^2}}{\epsilon}$	$\frac{\left\|u^0-u_h^\epsilon\right\|_{H^2}}{\sqrt[4]{\epsilon}}$
1	0.599947604	8.652542689
2^{-2}	1.48793695	12.4786026
2^{-4}	2.278636795	18.53711229
2^{-6}	3.398612852	27.50569621
2^{-8}	6.546636298	40.36922514
2^{-10}	15.30464406	58.38278827
2^{-12}	28.35360433	84.08190819
2^{-13}	16.00805235	101.4092367

Table 14. Change of $\|u-u_h^\epsilon\|$ w.r.t. ϵ $(h=1/127)$

Compared with Feng's and Oberman's result:

$$\text{case1: } u = e^{\frac{x^2+y^2}{2}}, f = (1 + x^2 + y^2)\, e^{x^2+y^2}$$

Ours		Feng's	
ϵ	$\|u - u_h^\epsilon\|_{L^2}$	ϵ	$\|u - u_h^\epsilon\|_{L^2}$
2^{-1}	0.093580805	0.5	0.038717
2^{-2}	0.081721912	0.25	0.040988
2^{-3}	0.064370429	0.1	0.032218
2^{-4}	0.044524223	0.05	0.022259
2^{-6}	0.015620435	0.0125	0.007817
2^{-9}	0.002361013	0.0025	0.001864
2^{-11}	0.00062253	0.0005	0.000405

Ours		Oberman's		
N	iter	N	M1	M2
128	69	121	31965	1205

case 2: $u = \sqrt{x^2 + y^2}$, $f = \begin{cases} 0 & \text{if } (x,y) \neq (0,0) \\ ? & \text{if } (x,y) = (0,0) \end{cases}$

Ours		Feng's	
ϵ	$\|u - u_h^\epsilon\|_{L^2}$	ϵ	$\|u - u_h^\epsilon\|_{L^2}$
2^{-1}	0.504058	none	
2^{-2}	0.371984		
2^{-3}	0.239381		
2^{-4}	0.142415		
2^{-6}	0.053103		
2^{-9}	0.014946		
2^{-11}	0.001954		

Ours		Oberman's		
N	iter	N	M1	M2
128	117	121	36396	10486

Under same accuracy, iteration number is less than other group. And the case have singularty can be achieved with high efficient computing.

Chapter 5
Conclusion

1. The error of $\|u - u_h^\epsilon\|_{L^2}$ of the Monge-Ampére is $\mathcal{O}(\epsilon)$ from test cases. The error of $\|u - u_h^\epsilon\|_{H^2}$ of the Monge-Ampére is $\mathcal{O}(\sqrt[4]{\epsilon})$ from test cases.

2. In numerical simulation of the elliptic regularization Monge-Ampére, the $\epsilon \geq h^2$, where h is mesh size.

3. In the singular case, we can shift the grid point such that the singular point locate in a element.

References

[1] L. A. Caffarelli and X. Cabre, “Fully nonlinear elliptic equations,” volume 43 of American Mathematical Society Colloquium Publications. American Mathematical Society, Providence, RI, 1995.

[2] L. A. Caffarelli and M. Milman, “Monge Ampere Equation: Applications to Geometry and Optimization, Contemporary Mathematics,” American Mathematical Society, Providence, RI, 1999.

[3] W. H. Fleming and H. M. Soner, “Controlled Markov processes and viscosity solutions,” volume 25 of Stochastic Modelling and Applied Probability. Springer, New York, second edition, 2006.

[4] D. Gilbarg and N. S. Trudinger, “Elliptic Partial Dierential Equations of Second Order, Classics in Mathematics”, Springer-Verlang, Berlin, 2001. Reprint of the 1998 edition.

[5] R. J. McCann and A. M. Oberman. “Exact semi-geostrophic flows in an elliptical ocean basin,” Nonlinearity,17(5):1891{1922, 2004.

[6] A.V. Pogorelov, "Extrinsic geometry of convex surfaces", AMS, 1973

[7] J. S. Schruben, “Formulation of a Reflector-Design Problem for a Lighting Fixture”, Journal of the Optical Society of America, V.62 N.12 (1972), 1498 − 1501.

[8] P. Benitez, J. C. Minano, J. Blen, R. Mohedano, J. Chaves, O. Dross, M. Hernandez and W. Falicoff, “ Simultaneous multiple surface optical design methof in three dimensions”, Opt. Eng. 43(7) 1489-1502.

[9] C. E. Gutierrez, “The Monge-Ampere Equation,” volume 44, Birkhauser, Boston, MA, 2001.

[10] J. D. Benamou, B. D. Froese and A. D. Oberman, “Two Numerical Method for the Elliptic Monge-Ampere Equation”, Preprint, 2009

[11] E. J. Dean and R. Glowinshi, “Numerical solution of the two-dimensional elliptic Monge–Ampere equation with Dirichlet boundary conditions: an augmented Lagrangian approach”, C. R. Acad. Sci. Paris, Ser. I 336 (2003) 779–784

[12] E. J. Dean and R. Glowinshi, “Numerical solution of the two-dimensional elliptic Monge–Ampere equation with Dirichlet boundary conditions: a least-squares approach”, C. R. Acad. Sci. Paris, Ser. I 339 (2004) 887–892

[13] E. J. Dean and R. Glowinshi, “Numerical methods for fully nonlinear elliptic equations of the Monge–Ampere type”, Comput. Methods Appl. Mech. Engrg. 195 (2006) 1344–1386

[14] P. Guan and X.J. Wang “On a Monge-Ampere equation arising in geometric optics”, J. Differential Geom, 48(1998), 205-223

[15] X. Feng and M. Neilan. “Mixed nite element methods for the fully nonlinear Monge-Amp ere equation based on the vanishing moment method.” SIAM J. Numer. Anal.,47(2),1226-1250, 2009.

[16] X. Feng and M. Neilan. “Vanishing moment method and moment solutions for fully nonlinear second order partial di erential equations.” J. Sci. Comput., 38(1),74-98, 2009.

[17] G. P. Bazeley, Y. K. Cheung, B. M. Irons and O. C. Zienkiewicz, ‘Triangular elements in plate bendingconforming and non-conforming solutions’, Proc. Conf. on Matrix Methods in Structural Mechanics, WPAFB, Ohio, 1965. pp. 547-576.

[18] G. A. Mohr and A. S. Power, “Natural Cubic Element Formulation and Infinite Domain Modelling for Potential Flow Problems”, ANZIAM J. 45(2003),133-143.

[19] C. A. Felippa and B. Haugen, “From the Individual Element Test to Finite Element Templates: Evolution the Patch test”, International Journal for Numerical Methods in Engineering, Vol. 38, 199-229 (1995)

[20] M. I. G. Bloor and M. J. Wilson, “An approximate analytic solution method for the biharmonic problem”, Proc. R. Soc. A (2006) 462, 1107-1121

[21] Julio Chaves, "Introduction to Nonimaging Optics", CRC, 271-324

[22] P. Benitez and J. C. Minano, “The Future of Illumination Design”, Optics and Photonics News, Vol. 18, Issue 5, pp. 20-25

[23] V. Oliker and E. Newman, “The Energy Conservation Equation in the Reflector Mapping Problem”, Appl. Math. Lett. Vol. 6,No. 1,pp. 91-95, 1993

[24] Daryl L. Logan. “First Course in the Finite Element Method Using Algor”, Pws Publishing, 1997.

[25] Per-Olof Persson and Gilbert Strang, "A Simple Mesh Generator in MATLAB," SIAM Review Vol. 46 (2) 2004.

[26] C. T. Kelley, “Iterative Method for Linear and nonlinear Equations”, SIAM, Philadelphia, 1995

[27] D. Braess, “Finite element” , Cambridge University Press, 2001

[28] C. Johnson, “Numerical solution of partial differential equations by the finite element method”, Cambridge University Press, 1988

[29] Susanne C. Brenner and L. Ridgway Scott, “The Mathematical Theory of Finite Element Methods”, Springer, 2002

[30] G. Awanou “Numerical Methods for Fully Nonlinear Elliptic Equations”, International Conference on Approximation Theory, 2010