Non Convergence Result for Conformal Approximation of Variational Problems Subject to a Convexity Constraint

Philippe Choné[1] and Hervé V.J. Le Meur[2]

Abstract : In this article, we are interested in the minimization of functionals in the set of convex functions. We investigate the discretization of the convexity through various numerical methods and find a geometrical obstruction confirmed by numerical simulations. We prove that there exist some convex functions that cannot be the limit of any conformal P_1 Finite Element sequence for a wide variety of refined meshes.

Keywords : convexity, finite elements, interpolation, conformal approximation, minimization

AMS Classification Codes : 65K99, 49M40, 49M45, 49M99, 65M30, 65M60, 26B25, 52A41, 90C20, 90C25

1 Introduction

This paper is devoted to the numerical approximation of variational problems subject to a convexity constraint, namely problems of the form

$$\begin{cases} \inf J(u), \\ u \in K, \end{cases}$$

where J is a functional and K is a subset of the cone of all convex functions on an open set Ω in $\mathbb{R}^N$. Such problems appear in various contexts, in particular in physics and economics.

One of the first problems in the calculus of variation, Newton's problem of minimal resistance, involves a concavity constraint (see the original paper [9] and the historical survey [6]). In this context, the functional J and the set K are given by

$$J(u) = \int_\Omega \frac{1}{1+|\nabla u|^2} \mathrm{d}x, \ K = \{u \in W^{1,\infty}_{\mathrm{loc}}, 0 \le u \le M, u \text{ concave } \}.$$

Newton found the minimum of J over the set of radial function

$$K' = \{u \in K, u \text{ is radial } \},$$

when Ω is a ball in $\mathbb{R}^2$. The existence of a solution for a general convex set Ω has been proved recently (see [4]). In [3], the authors prove that, when Ω is a ball,

$$0 < \min_K J(u) < \min_{K'} J(u).$$

In other words, Newton's solution does not minimize J over K. The minimizers of $\min J$ over K are *not radial* and not unique.

We now turn to a problem coming from an economic question, namely the design of a nonlinear tariff by a regulated monopolist (see [10]). In this context, the functional J is given by

$$J(u) = \int_\Omega \left(\frac{1}{2} \nabla u^T \, C \, \nabla u - x.\nabla u + (1-\alpha)u \right) \mathrm{d}x, \tag{1}$$

[1]CREST-LEI, 28 rue des Saints-Pères, 75007 Paris, FRANCE

[2]Laboratoire d'analyse numérique et EDP, CNRS and Université Paris-Sud, Bât. 425, 91.405 Orsay CEDEX, Herve.LeMeur@math.u-psud.fr

where $0 \leq \alpha \leq 1$ and C is a positive definite (2,2) matrix. The set K is given by

$$K = \{u \in H^1(\Omega), u \geq 0, u_x \geq 0, u_y \geq 0, u \text{ convex } \}. \tag{2}$$

By contrast with Newton's problem, the functional J is convex and coercive on K. It is easy to check that there exists one unique minimizer of J over K.

In [10], the authors focus on the case $\alpha = 0$ (unregulated monopolist) and $C = Id$. They give a sufficient condition on the domain Ω for the convexity constraint to be active. Typically, when Ω is a square $[a, b]^2$, there is an area where the range of the hessian matrix of u is 1 (see Figure 1). In this last area, the function depends only on $x + y$ which varies from $a + \tau_0$ and $a + \tau_1$. The value of τ_0 is given in section 3.2.3.

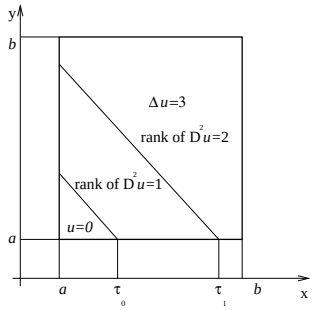

Figure 1: Solution if $\alpha = 0$

Note that, when $\alpha = 1$, the problem degenerates: the solutions are given, up to an additive constant by

$$\nabla u(x) = C^{-1}x \Rightarrow u(x) = x'C^{-1}x/2 - \text{Cst}. \tag{3}$$

In the problems described above, the convexity constraint is typically binding. It was even proved in [8] that the minimum of the Newton problem is nowhere strictly convex. Therefore any numerical method of approximation must explicitely take into account this constraint.

In this paper, we focus on *conformal* or *internal* approximations i.e. methods where the approximating sequence u_h belongs to the same set as u. The main result of the paper is a negative one: conformal P_1 Finite Element (FE) methods cannot converge to the solution of the problem although a FE discretization of any H^1 function u can be as close to u as wanted. This is essentially due to geometrical obstructions that we explain in details below. We prove also that natural extensions of the 1-D case and P_2 FE do have similar problems.

Therefore we should now turn to *non*-conformal methods. A first attempt in this direction can be found in [5], that states a convergence result. The approximated problem, however, involves a very high number of constraints (of order N^2, where N is the number of vertices in the mesh).

The paper is organized as follows. In section 2, we study the conformal approximation through non-local basis. We explain that we are not able to recover the cone structure due to a geometric obstruction. In section 3, we study the conformal P_1 finite-element approximation. We show that P_1 and P_2 Lagrange interpolation does not preserve convexity and formulate this result in a precise fashion. Section 4 presents some extensions (P_2, Argyris) and concludes.

We used the software Matlab for the minimization because we aim at providing a "not too complicate" solution to the economical problem, available for non-specialists.

2 Approximation through conformal non-local basis

The very first idea when one wants to discretize convexity is to look for an approximate cone $\mathcal{C}_h$ depending on h(h small), such that $\mathcal{C}_h$ should be a subset of the cone $\mathcal{C}$. Also, we would like

that this $\mathcal{C}_h$ could be as close as we want of $\mathcal{C}$. In order to have the structure of cone, $\mathcal{C}_h$ should be the set of all linear combinations of a finite number of convex functions, with nonnegative coefficients. As a consequence, the basis functions are non-local.

In this section, we will be interested in a specific conformal approximation for which the basis functions satisfy convexity and not only the function.

So as to test this approximation, we try it in one-dimension.

2.1 One dimensional discretization

Let $(x_i)_{i=0..N}$ be a general subdivision of $[0,1]$ where $x_0 = 0$ and $x_N = 1$. The basis we propose is composed of $N+1$ functions that satisfy the constraint (here convexity) and so, they are non-local. Their shape can be seen on Figure 2.

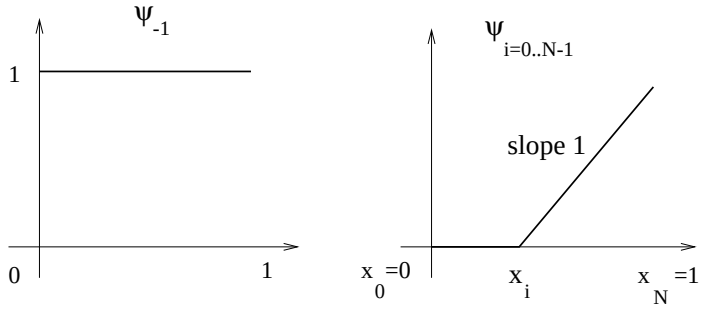

Figure 2: Shape of the 1-D basis

This basis associated to the grid on $[0,1]$ enables to state the following theorem whose proof is easy and only scheduled here so as to be referred to in the 2-D case :

THEOREM 1 *Let $f \in C^2([0,1])$. Then, for all ε, and N large enough ;*

$$\| f(x) - f(0)\Psi_{-1}(x) - f'(0)\Psi_0(x) - \sum_{1}^{N-1} [f'(i/N) - f'((i-1)/N)]\, \Psi_{i,N}(x) \|_{H^1} < \varepsilon$$

Idea of the proof of Theorem 1.

So as to find the coefficients, we write the system $f'(x_i) = \sum_{j=0}^{N} \alpha_j \Psi'_j(x_i^+)$, where $\Psi'_j(x_i^+)$ is the right derivative of Ψ_j at x_i. Thanks to the chosen functions, the matrix is triangular. Moreover, it is invertible and leads to the formulas for the coefficients. Simple estimates complete the proof.

□

This theorem shows that it is possible to approximate a 1-D convex function as closely as wanted, by a linear combination with nonnegative coefficients except two terms. The two first terms may have arbitrary sign and we recover the cone structure.

The idea of this discretization is to lift the gradients as they grow. So convexity is used, here, as a gradient increase. In 1-D, the two properties are equivalent, but not in 2-D.

Let us justify the present discretization. For that purpose, we denote $\mathcal{C}' = \mathcal{C} \cap \{u \in W^{1,\infty}(\Omega), \sup(|u|_\infty, |u'|_\infty) \leq 1\}$ a section of the cone of convex functions. The set $\mathcal{C}'$ is convex and compact in $H^1(\Omega)$.

If $\Omega = [a,b]$, the extremal points of $\mathcal{C}'$ are the functions ψ_y, $a \leq y \leq b$, where $\psi_y(x) = \sup(0, x-y)$. By the Krein-Milman theorem, $\mathcal{C}'$ coincides with the adherence in H^1 of the

convex hull of the set $\{\psi_y, a \leq y \leq b\}$, which gives another proof of Theorem 1 (the functions $\Psi_{i,N}$ of section 2.1 are clearly dense in this set).

The two-dimensional case is much more complicated since we do not know the set of the extremal points of $\mathcal{C}'$.

2.2 2-D basis-conformal discretization

Our goal is to approximate any convex function as a combination of convex functions (basis-conformal) with nonnegative coefficients in a way similar to the one of the previous subsection. We will restrict ourselves to rectangle domains ($[a,b]^2$). We choose a uniform discretization of $x : (x_i = a + i(b-a)/N)_{i=0...N-1}$ and of $y : (y_j = a + j(b-a)/N)_{j=0...N-1}$.

So as to generalize the 1-D basis, we have to conceive a family of functions that should lift the two components of the gradients and whose coefficients would be the nonnegative coefficients of the convex hull of the extremal points of a section of the cone. We use the convex piecewise linear functions :

$$f_{i,j,1}(x,y) = \sup(0, \cos(\theta)(x - x_i), \sin(\theta)(y - y_j)), \quad (4)$$

$$f_{i,j,2}(x,y) = \sup(0, \sin(\theta)(x - x_i), \cos(\theta)(y - y_j)), \quad (5)$$

for θ sufficiently small and $(i,j) \in [0, ..., N-1]$. Small θ ($\sin\theta < \delta x/(b-a)$) enable to retrieve uniqueness of the components, and to have the most natural extension of the 1-D case. Indeed, $\theta = 0$ would make our family not to be free. The shape of these functions is depicted in Figure 3 as the shape of their gradients.

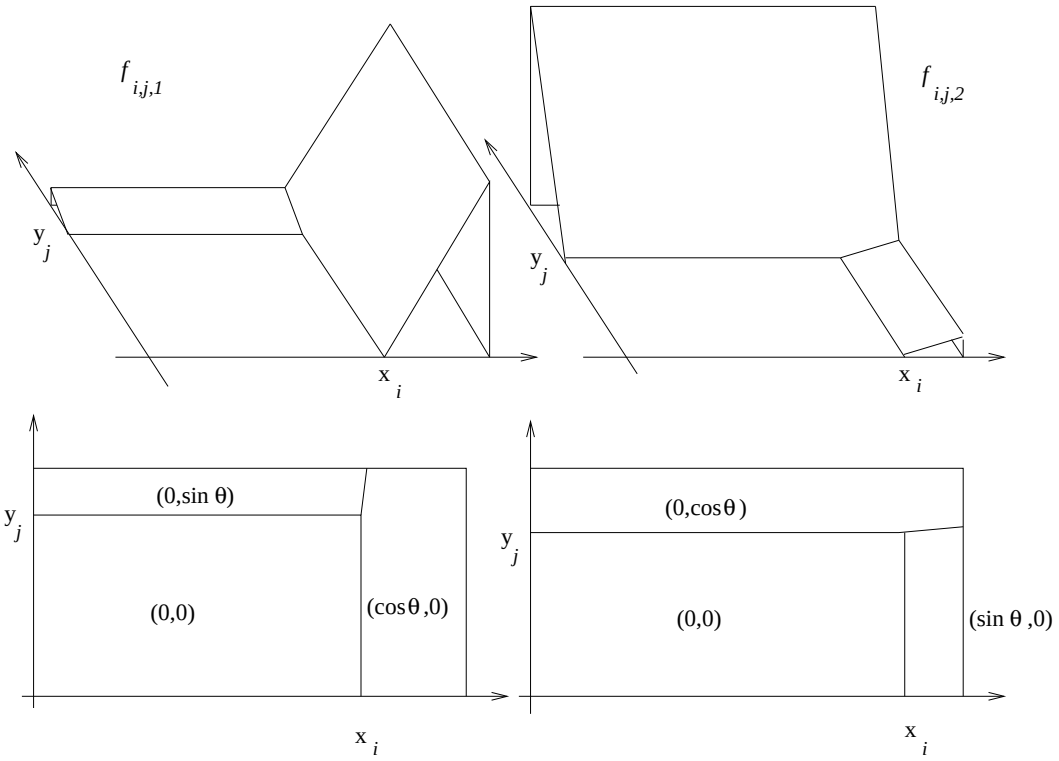

Figure 3: Shape of the 2-D basis and their gradients

In a way similar to the 1-D case, we expand :

$$f_N(x,y) = A\chi_{[a,b]^2} + \sum_{i,j=0,..,N-1} \alpha_{i,j} f_{i,j,1} + \beta_{i,j} f_{i,j,2}, \quad (6)$$

where $A, \alpha_{i,j}, \beta_{i,j}$ are constants and $\chi_{[a,b]^2}$ is the characteristic function of the set $[a,b]^2$. We denote $\alpha = ((\alpha_{i,j})_{i=0,..N-1})_{j=0,..,N-1}$ to reorder the matrix into a vector and use the same ordering for β. In order to mimic the procedure in 1-D, we define the projection f_N of f by its components $A, \alpha_{i,j}, \beta_{i,j}$:

$$\left\{ \begin{array}{rcl} f(a,a) & = & A, \\ \dfrac{\partial f}{\partial x}(x_{i_0}^+, y_{j_0}) & = & \displaystyle\sum_{i,j} \alpha_{i,j} \frac{\partial f_{i,j,1}}{\partial x}(x_{i_0}^+, y_{j_0}) + \beta_{i,j} \frac{\partial f_{i,j,2}}{\partial x}(x_{i_0}^+, y_{j_0}), \\ \dfrac{\partial f}{\partial y}(x_{i_0}, y_{j_0}^+) & = & \displaystyle\sum_{i,j} \alpha_{i,j} \frac{\partial f_{i,j,1}}{\partial y}(x_{i_0}, y_{j_0}^+) + \beta_{i,j} \frac{\partial f_{i,j,2}}{\partial y}(x_{i_0}, y_{j_0}^+), \end{array} \right. \tag{7}$$

where $\partial f / \partial x(x_{i_0}^+, y_{j_0})$ is the right derivative of f with respect to x at (x_{i_0}, y_{j_0}). A theorem similar to Theorem 1 can be stated and extended to the following :

THEOREM 2 *Let $f \in C^2([a,b]^2)$, $\theta = o(\Delta x)$ and $\Delta x \sim \Delta y$. Then there exist unique $A, \alpha_{ij}, \beta_{ij}$ given by (7) and they are such that :*

$$\| f - f_N \| \to 0 \; in H^1.$$

Moreover, for positive j ;

$$\alpha_{ij} = -\frac{\partial^2 f}{\partial x \partial y} \Delta y + o(\Delta y).$$

As a consequence, no sign of α_{ij} can be guaranteed even under the assumption that f is convex.

Proof of Theorem 2.

Let us consider the $N \times N, N^2 \times N^2$, and $N^2 \times N^2$ matrices

$$K_N = \begin{pmatrix} 1 & \dots & \dots & 1 \\ 0 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 \end{pmatrix}, D_N = \begin{pmatrix} K_N & L_N & \dots & L_N \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & L_N \\ K_N & \dots & \dots & K_N \end{pmatrix}, \; E_N = \begin{pmatrix} K_N & 0 & 0 \\ \vdots & \ddots & 0 \\ K_N & \dots & K_N \end{pmatrix},$$

where $L_N = K_N - I_N$. With these notations, and if we order $\alpha = ((\alpha_{ij})_{i=0,N-1})_{j=0,N-1}$, the system (7) may be written in an almost block-diagonal way (for θ small) :

$$\begin{pmatrix} \cos\theta D_N & \sin\theta E_N \\ \sin\theta E_N & \cos\theta D_N \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} \left(\dfrac{\partial f}{\partial x}\right)_{ij} \\ \left(\dfrac{\partial f}{\partial y}\right)_{ij} \end{pmatrix}. \tag{8}$$

Then, it is a simple exercise of linear algebra to find

$$D_N^{-1} = \begin{pmatrix} I_N & 0 & \dots & 0 & -L_N K_N^{-1} \\ -I_N & \ddots & \ddots & & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & -I_N & I_N \end{pmatrix},$$

and to get uniqueness of the components $A, \alpha_{ij}, \beta_{ij}$ thanks to the smallness of θ. Then, the proof of convergence is very similar to the one of 1-D. Last, a simple expansion gives that for $j > 0$;

$$\alpha_{ij} = \left(\frac{\partial f}{\partial x}\right)_{ij} - \left(\frac{\partial f}{\partial x}\right)_{i,j-1} + O(\theta) = \left(\frac{\partial^2 f}{\partial x \partial y}\right)_{ij} \Delta y + o(\Delta x, \Delta y).$$

□

So we exhibit a sequence of functions that tends to u, but there is no way of guaranting that the sequence should remain convex, even if u is strictly convex. As the most natural 2-D extension of the 1-D solution does not suit our requests on the structure of cone, we have left the basis-conformal approximation and tried more classical Finite Element for discretization.

3 Conformal approximation through Finite Elements P_1

We choose a triangular mesh and look for functions which are continuous and linear in each triangle. Namely we consider Lagrange P_1 Finite Elements (FE). More details can be found in [2], or [11].

3.1 Generalities

A typical basis function $\phi_i(x)$ can be found in Figure 4. Its values are 1 at the node i, zero at the other nodes, it is linear in each triangle and continuous in Ω.

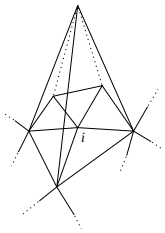

Figure 4: Shape of the function $\phi_i(x, y)$

The function f is discretized by

$$f_h(x) = \sum_{i, \text{ node of the mesh}} f_i \phi_i(x),$$

where f_i is the component of f_h and $\phi_i(x)$ is the FE basis. The overall method will be to look for a function f_h (indeed for a finite number of values f_i) that could minimize the functional ((1) with $\alpha = 1$) and satisfy the convexity condition. In that sense, the approximation will be conformal because the functions manipulated are supposed convex, although they are linear combinations of non-convex functions. We have exact solutions to which the computed solution is compared :

$$C = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \Rightarrow u(x) = \frac{1}{1-\rho^2}\left(\frac{x^2}{2} - \rho xy + \frac{y^2}{2}\right) - \text{Cst.} \tag{9}$$

We still need to have a characterization of convexity for P_1 functions. It is given by the following lemma, which proof is easy :

LEMMA 3 *A function f_h, P_1 in the rectangle $[a, b]^2$, is convex if and only if, for any pair of adjacent triangles*

$$(q_2 - q_1).n_{12} \geq 0, \tag{10}$$

where q_1 (resp. q_2) is the (constant) gradient of f_h in triangle 1 (resp. 2) and n_{12} is the unit normal pointing from triangle 1 to 2.

Proof of Lemma 3 Recall that a distribution v on Ω is a convex function if and only if, for all nonnegative smooth function ϕ with compact support in Ω, the bilinear symmetric map

$$(h,k) \to < \frac{\partial^2 v}{\partial h \partial k}, \phi >$$

is semi-definite positive (for details, see [12]). Assume $v \in P_1$ and notice that, by Green's Formula

$$\begin{aligned} < \frac{\partial^2 v}{\partial h \partial k}, \phi > &= -\sum_T < \frac{\partial v}{\partial h}, \frac{\partial \phi}{\partial k} > \\ &= \sum_e (q_2 - q_1).n_{12})(n_{12}.h)(n_{12}.k) \int_e \phi(s) \mathrm{d}s, \end{aligned}$$

where the last summation is taken over all interior edges of the mesh. The result follows from the fact that the map

$$(h,k) \to (n_{12}.h)(n_{12}.k),$$

is semi-definite positive for all vector n_{12}.

□

Notice that although convexity is a non-local property, the discretization leads to local characterization. Moreover, the function will be known at the nodes of the mesh, but the constraint makes sense only on the interior edges. The software MAPLE was used so as to have optimized symbolic formulae depending on the degrees of freedom of the unknown field.

3.2 Numerical results

In the present subsection, we use the various structured meshes depicted in Figure 5 and compare the computed results with the exact solution (3) ($\alpha = 1, \rho \in]-1,1[$).

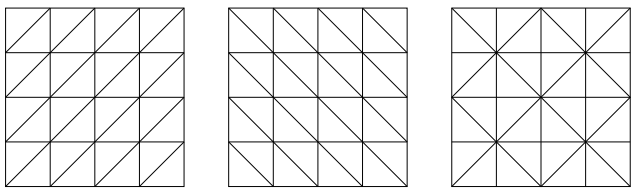

Figure 5: Three structured meshes

3.2.1 Mesh 1

It happens that if ρ is positive, the results are satisfactory. We report the L^2 norm as a function of the number N of sub-intervals in each direction and find a good convergence on the square $[4,5]^2$ (see Figure 6). Yet, for $\rho = -0.1$ (and more generally for all the non-positive ρ), we find no good convergence as can be seen on Figure 6. Before concluding, we try the next mesh 2 in the next subsubsection.

3.2.2 Mesh 2

Here, we use mesh 2 (see Figure 5). Usually, the convergence does not depend on the type of the triangles but on their size. Surprisingly, here, the results are opposite to those of mesh 1 : if ρ is negative, we reach a good convergence, while if ρ is positive, the convergence is very bad as can be seen on Figure 7. This unusual behavior highlights the crucial role of the type of the mesh on convergence.

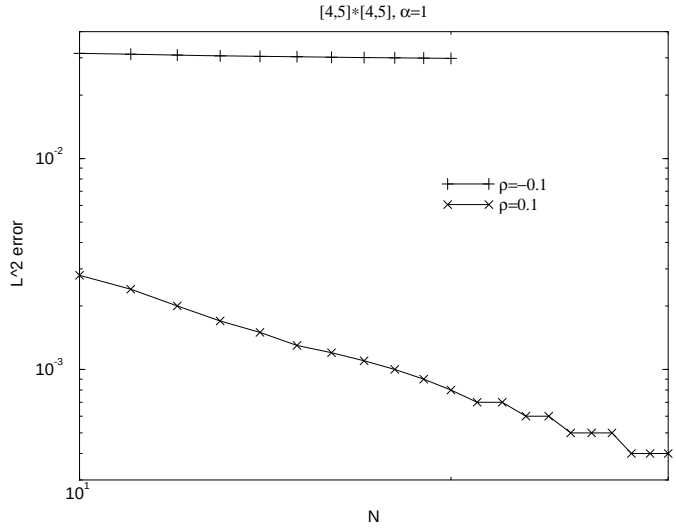

Figure 6: L^2 error for mesh 1

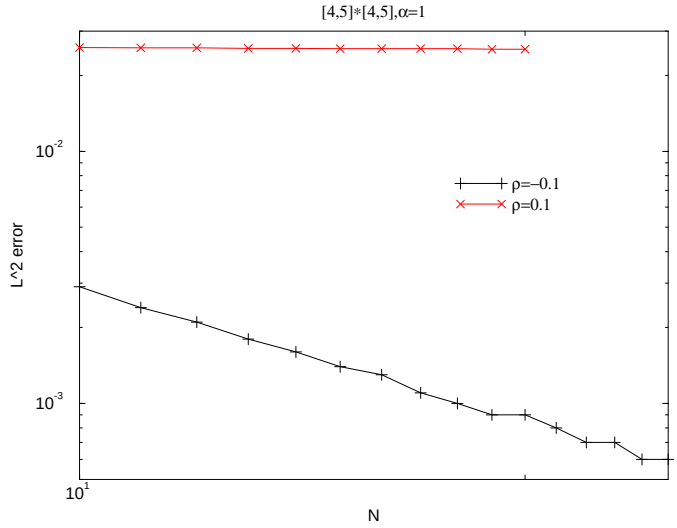

Figure 7: L^2 error for mesh 2

3.2.3 Mesh 3

As a conclusion of the two preceding tries, we use mesh 3 (see Figure 5) that seems to have both advantages of mesh 1 and mesh 2 : the direction of the edges are in alternate direction and so we hope to recover good convergence for all ρ.

Indeed, with that mesh, we have satisfactory results for both sign of ρ and even for $\rho = 0$ as can be seen on Figure 8.

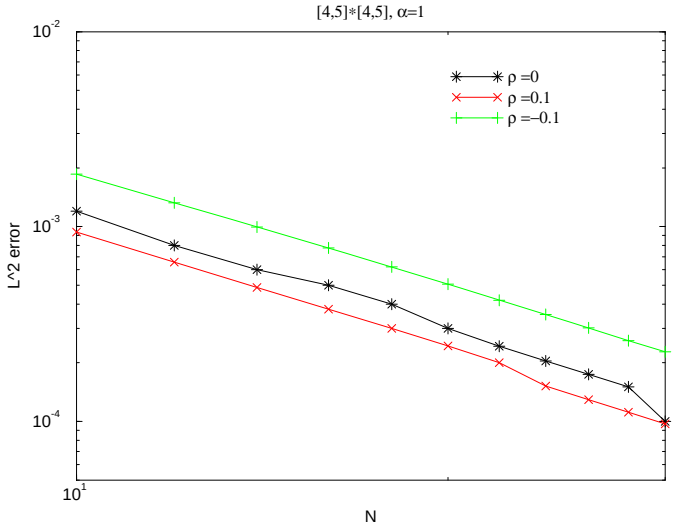

Figure 8: L^2 error for mesh 3

Moreover, we check the other properties based on the non-explicit solution (see [10]) and the agreement may look right :

1. The trace at $y = a$ for $\alpha = 0$ should give the value $\tau_0 = a/3 + \sqrt{4a^2 + 6(b-a)^2}/3$ (see Figure 1 and [10]). Here, we found numerically $\tau_0 \in [4.13, 4.16]$ (see Figure 9) while the exact value is 4.122.

2. The shape of the gradient looks very much like the one expected (see Figure 9).

3.2.4 Unstructured meshes

Various unstructured meshes were used and odd behaviour may be reported. CPU time may depend on

- reordering of the mesh : from 0 to 30 % more time for the same mesh,
- parameter : for a given unstructured mesh ($\rho = 0.4$, 170 nodes and 294 triangles), the CPU time appears to be erratic without any explanation :

α	0.94	0.92	0.90
CPU time (s)	5	1053	10

Moreover, the CPU time seems not to depend on the angles of the triangles (as is usual in most physical problem). In that case, we moved a point along so as to have some angles crossing 90^o and the CPU time remained similar.

Yet, at given number of triangles and nodes, whether the mesh is structured or not, the CPU time may be 100 times greater for the unstructured mesh ! The most striking example

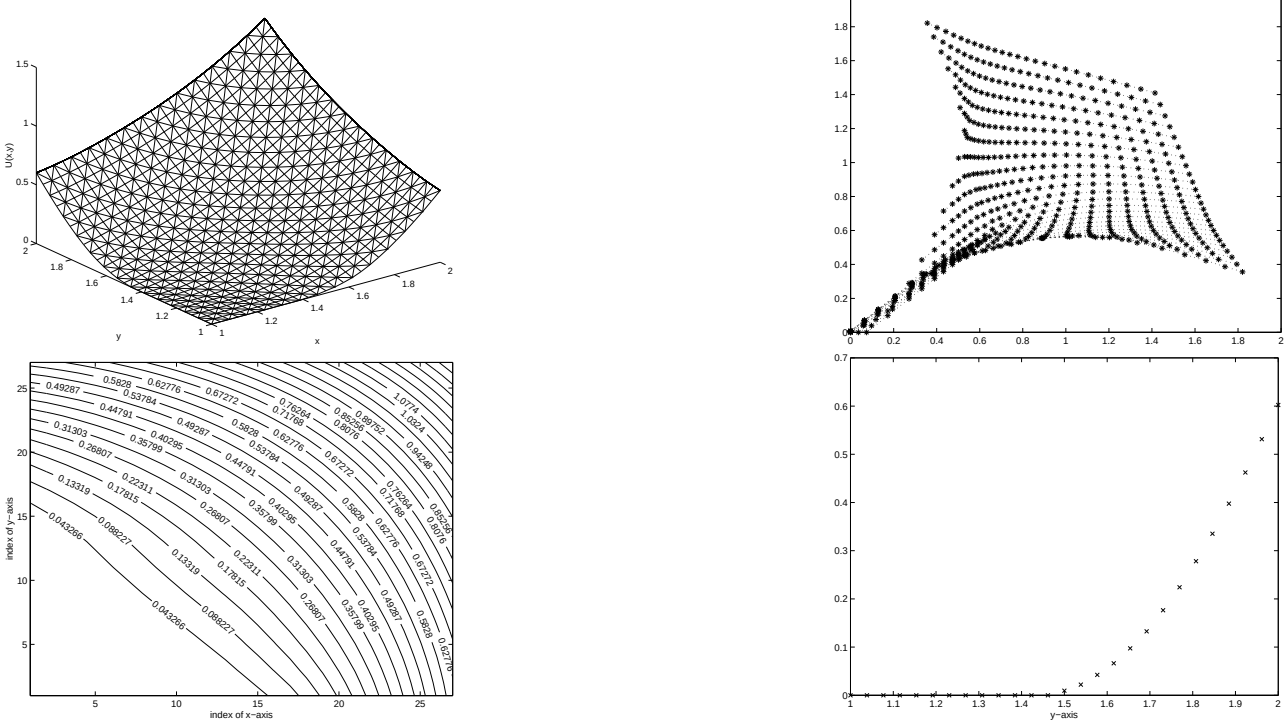

Figure 9: Graphe, gradient, contour and trace along y of the computed u

was a structured mesh to which we added only one point. The CPU time got 50 times more. This last example indicates that the odd behaviour does not come from the computation code, but from the topological properties of the mesh.

Last, the gradients are so bad that no information can be retrieved.

The study of unstructured meshes could be considered as meaningless as a structured mesh, like mesh 1 or 2 (for which P_1 FE does not work), is only a particular case of more general meshes. But the difficulty is that for a minimization problem, a numerical code will most often give a solution. Then, we have no other means than numerical analysis or good numerical experiments to test the likelihood of the solution. Both approachs are helpful and we hope that the insight given by the unstructured study might be of some help to indicate a more general obstruction.

3.3 A geometrical obstruction

In this section, we show that the geometrical structure of the mesh imply many constraints on P_1 convex functions. These constraints in turn prevent the conformal P_1 approximation to converge when the solution to the variational problem does not satisfy them and this limitation may be local.

THEOREM 4 *Assume that there are two directions h and k in a subset $\Omega' \subset \Omega$ such that*

$$(n.h).(n.k) \geq 0 \tag{11}$$

for every vector n unit normal to an edge of a triangle in the triangulation $\mathcal{T}_h \cap \Omega'$. Then, for every convex and P_1 function v, we have :

$$\frac{\partial^2 v}{\partial h \partial k} \geq 0 \tag{12}$$

in the sense of the Radon measures on the edges of $\mathcal{T}_h \cap \Omega'$

Proof of Theorem 4 Let ϕ be some nonnegative smooth function with compact support in Ω. Summing up Green's Formula for a P_1 function in every triangle yields

$$\begin{aligned} < \frac{\partial^2 v}{\partial h \partial k}, \phi > &= -\sum_T < \frac{\partial v}{\partial h}, \frac{\partial \phi}{\partial k} > \\ &= \sum_e ((q_2 - q_1).n_{12})(n_{12}.h)(n_{12}.k) \int_e \phi(s) \mathrm{d}s \end{aligned}$$

where the last summation is taken over all interior edges of the mesh. The conclusion follows from the geometric property (11) of the mesh and the convexity of v, which writes: $(q_2 - q_1).n_{12} \geq 0$ (see Lemma 3).

□

For instance, in the structured mesh 3, the normal vector are $n_1 = (0,1)$, $n_2 = (1,0)$, $n_3 = (1,1)$ $n_4 = (1,-1)$. Hence we can choose $h = (1,0)$ and $k = (1,-1)$. Therefore, for all convex and P_1 function on mesh 3, (11) is satisfied on **all** edges and so :

$$v_{xx} - v_{xy} \geq 0. \tag{13}$$

Yet, this does not hold for every convex function. If the solution u to the variational problem does not satisfy this constraint, then if the mesh remains structured, u cannot be approximated by a sequence of convex P_1 function (since a limit of functions satisfying (12) necessarily satisfies (12)). Hence we have proved the Corollary at least for structured meshes :

COROLLARY 5 *There exist convex functions that cannot be the limit of convex P_1 functions in the sense of distributions on a given family of structured meshes.*

Inded, the property (11) may occur also on a non structured mesh, and so Corollary 5 is true on a wide range of couple (mesh, refinement process).

Let Ω' be an open set in any given triangle T of $\mathcal{T}_h$. Assume the process of refinement be defined by dividing any triangle into four homothetic triangles (nodes are former nodes and mid-points). Then, the normals will **not** be enriched, whatever might be the level of refinement.

Assume the three normals of the edges of T are (up to a change of sign and change of coordinates) $(1,0), (\cos\theta_1, \sin\theta_1), (\cos\theta_2, \sin\theta_2)$ with $0 < \theta_1 < \pi/4, 3\pi/4 < \theta_2 < \pi$. Then, for $h = (\cos\theta_1, \sin\theta_1), k = (\sin\theta_2, -\cos\theta_2)$, the property (11) is satisfied and the limit of u_h will satisfy an *additional* relation in T.

Last, let us stress that the property (11) is *local* and depends not so much on the mesh than on the refinement process.

3.4 Numerical tests

So as to test the numerical validity of the preceding theorem, we use a matrix C and an exact solution more general (still with $\alpha = 1$) than (3) :

$$C = \begin{pmatrix} 1 & \rho \\ \rho & \mu \end{pmatrix}; \; u_{\rho,\mu} = \frac{1}{\mu - \rho^2} \left(\mu \frac{x^2}{2} - \rho xy + \frac{y^2}{2} \right) - \mathrm{Cst}, \tag{14}$$

where Cst is a constant such that $u_{\rho,\mu}(a,a) = 0$. Moreover, the constraint that $u_{\rho,\mu}$ should be convex gives $(\mu - \rho^2) \geq 0$. The constraint that the gradient should be positive gives $\rho \leq y/x \leq \mu/\rho$ if $\rho > 0$ and no condition if $\rho \leq 0$.

Theorem 4 (see (13)) implies that when ρ moves in such a way that

$$\frac{\partial^2 u_{\rho,\mu}}{\partial x \partial x} - \frac{\partial^2 u_{\rho,\mu}}{\partial x \partial y} = \frac{1}{\mu - \rho^2}(\mu + \rho), \tag{15}$$

crosses zero, then convergence should not be achieved. This could be found numerically by taking $\mu = 0.1, \alpha = 1$. On Figure 10, we can see the jump in the slope of the L^2 error compared to the exact solution precisely at the value predicted for $N = 20$. Moreover, this is not a problem of accuracy because we have the same results for more accurate computations with $N = 28$ as can be seen on Figure 10.

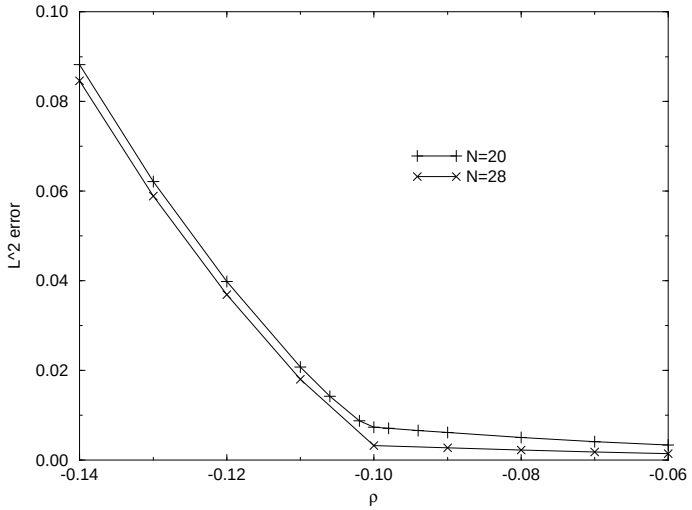

Figure 10: Convergence results depending on ρ and N ($\alpha = 1, \mu = 0.1, [4,5]^2$)

It can be concluded from Figure 10 that the convergence is not achieved for a class of convex functions.

3.5 Convexity and Lagrange interpolation

Let us consider the following partition of the square $\Omega = [0,1]^2$:

$$\Omega = T^l \cap T^u, \quad \text{with} \quad T^l = \{(x,y) \in \Omega, x + y \leq 1\} \quad \text{and} \quad T^u = \{(x,y) \in \Omega, x + y \geq 1\}.$$

We denote by P_1^1 (respectively P_1^2) the Lagrange interpolation operators to P_1 (resp. P_2) FE. Let f be the convex function defined by : $f(x,y) = \max(x,y)$ on Ω. We set $f_1^1 = P_1^1 f$ and $f_1^2 = P_1^2 f$. Then it is a simple exercise to prove the following Lemma.

LEMMA 6 *The functions f_1^1 and f_1^2 are given by*

$$f_1^1(x,y) = \min(x+y, 1) = \begin{cases} x + y & in\ T^l \\ 1 & in\ T^u \end{cases}$$

and

$$f_1^2 = \begin{cases} x + y - 2xy & in\ T^l \\ 2(x+y) - 2xy - 1 & in\ T^u. \end{cases}$$

Now we consider the regular mesh with N^2 vertices like mesh 1 (see Figure 5), $N \geq 1$ and set $h = 1/N$. We consider the Lagrange interpolation operators P_N^1 and P_N^2 and the interpolated functions $f_N^1 = P_N^1 f$ and $f_N^2 = P_N^2 f$.

Consider the squares of the mesh whose south west corners are (x_k, y_k), with $x_k = y_k = kh, k = 0, ...N-1$. Then we have on those squares

$$f_N^1 = kh + hf_1^1(\frac{x-x_k}{h}, \frac{y-y_k}{h}) \text{ and } f_N^2 = kh + hf_1^2(\frac{x-x_k}{h}, \frac{y-y_k}{h}).$$

Outside those squares, it is clear that : $f = f_N^1 = f_N^2$.

Now we can compute the second derivatives $\mathrm{D}^2 f_N^1$ and $\mathrm{D}^2 f_N^2$ in the sense of the Radon measures. We denote by $(\mathrm{D}^2 f_N^1)_-$ and $(\mathrm{D}^2 f_N^2)_-$ the negative parts of these measures. Since $(\mathrm{D}^2 f_N^1)_- = 0$ inside the triangles, its norm is given by

$$|(\mathrm{D}^2 f_N^1)_-|_{\mathcal{M}} = \sum_e \int_e (< q_2 - q_1, n_{12} >)_- \mathrm{d}s,$$

where the sum is over the edges of the mesh.

The support of $(\mathrm{D}^2 f_N^2)_-$ contains some portions of edges and a surfacic part. Its norm in the sense of Radon measure is given by

$$|(\mathrm{D}^2 f_N^2)_-|_{\mathcal{M}} = \sum_e \int_e (< q_2 - q_1, n_{12} >)_- \, \mathrm{d}s + \sum_T \lambda_-(T)\mathrm{mes}(T)$$

where T is any triangle of the mesh, $\mathrm{mes}(T) = h^2/2$ is its area, and $\lambda_-(T)$ is the negative eigenvalue of the (constant) matrix $\mathrm{D}^2 f_N^2$ in the triangle T.

The following proposition easily follows

THEOREM 7 *The norms of the negative parts of the second derivative of f_N^1 and f_N^2 are bounded away from zero. More precisely, we have:*

$$|(\mathrm{D}^2 f_N^1)_-|_{\mathcal{M}} = 2 \qquad \textit{and} \qquad |(\mathrm{D}^2 f_N^2)_-|_{\mathcal{M}} = 4 - \frac{1}{N},$$

for all $N \geq 1$.

The proof is left to the reader. We just mention that the support of $(\mathrm{D}^2 f_N^1)_-$ is the set of edges that intersect the line $x = y$ The support of the measure $(\mathrm{D}^2 f_N^2)_-$ is much more complicated, since it has a surfacic part ($\lambda_-(T) = 1/h$ in the triangles along the line $x = y$) and a part supported by the edges.

We may conclude that the Lagrange interpolate of a convex function is not necessarily convex and the distance may remain finite even asymptotically (when the size of the mesh tends to zero).

We have proved that $\mathcal{C} \cap P_1$ is not dense in $\mathcal{C}$ (Corollary 5). We conjecture that the same result is true for $\mathcal{C} \cap P_2$.

3.6 Comment on the bibliography

During the preparation of the present work, we were informed of the article of Kawohl and Schwab [7] who apply the P_1 FE to the Newton problem. In this article, the authors claim they have proved that "conforming approximations u^N converge .../... to a minimizer". Yet, as says Corollary 5 and as we checked numerically, there is no sequence of convex P_1 functions that can converge to some u.

The error is in the proof of Lemma 2.1 where is claimed that the piecewise interpolant of a convex function u is convex. The best counter-example is the function $(x, y) \mapsto \sup(x, y)$ on a square $[0, 1]^2$ divided in two triangles by the segment $x + y = 1$. The P_1 interpolant of this function is even concave.

This explains the error when the authors state that for every u convex and bounded, there exist a sequence of convex and bounded P_1 functions that converges to u in $W^{1,p}_{loc}$ for $1 \leq p < \infty$. Later this lemma is used in the step 2 of the proof of their main theorem.

Moreover, their numerical results give a symmetric solution, while the solution may not be symmetric as there exist nonsymmetric functions that decrease the energy (see [3]).

Yet, the possibility to perform non-conformal approximation of convex functions remains a good idea. It was investigated by Carlier, Lachand-Robert and Maury [5] who proved convergence by the use of a lagrange multiplier and an Uzawa method. The problem lies in the size of the discretized constraint.

The reason for non-convergence is that the surfacic second derivatives have prescribed signs not balanced by the volumic derivative. We could hope that the P_2 FE could solve this problem as there is volumic second derivative inside each triangle, even if there is still a surfacic second derivative.

4 Conclusion

In this article, we proved that basis-conformal approximation does work in 1-D, but the most natural extension to 2-D does not. The overall idea behind this is the Krein-Milman theorem that made us hope that we could, through a discrete family of extremal functions, approximate any convex function with the convex hull of a finite sub-family. We have proved that it was not possible at least with the most natural extension of 1-D case.

Then, we tried conformal Finite Element (FE) method P_1. Although some improvements in the mesh improve the apparent numerical convergence, we proved, both theoretically and numerically, that this conformal method may not converge for some limit function. We even proved that for a given convex function, the norm of the negative part of the second derivative remains finite, whatever the accuracy of discretization.

The last idea would be P_2 FE, but the same argument as in P_1 remains for P_2 : the lineic derivative of the basis functions along the edges of the mesh forces the limit function to satisfy a non-natural property. One might hope to counterbalance it with the non-zero second derivative inside the triangles. On the other hand we have exhibited one convex function f which P_2 interpolate f_N^2 has a second derivative with a negative part which remains finite, whatever N.

Last, we point out an error in an article.

Although our results are negative, we believe they might be of some help, should they help only to prevent researchers from using conformal P_1 FE in minimization of functionals under the constraint that the function should be convex.

Acknowledgements : The authors wish to thank a referee for giving helpfull comments that substantially improved the manuscript.

References

[1] J.S. Archer, Consistent matrix formulations for structural analysis using finite-element techniques. AIAA J. 3, 1910-1918 (1965).

[2] D. Braess, *Finite Elements* Cambridge University Press 1997

[3] F. Brock, V. Ferone and B. Kawohl, A symmetry problem in the calculus of variations, *Calc. Var. Partial Differ. Equ.* 4(6): 593-599, (1996)

[4] G. Buttazzo, V. Ferone and B. Kawohl, Minimum problems over sets of concave functions and related questions, *Math. Nachr.* (1995) 173:71-89

[5] G. Carlier, T. Lachand-Robert and B. Maury, A numerical approach to variational problems subject to convexity constraint, To appear in Numerische Math.

[6] H.H. Goldstine, *A history of the calculus of variations from the 17th to the 19th century* Springer-Verlag, Heidelberg, 1980

[7] B. Kawohl and C. Schwab, Convergent finite elements for a class of nonconvex variationnal problems, *IMA Journal of Numerical Analysis* (1998) **18** 133-149.

[8] T. Lachand-Robert and M.A. Peletier, An Example of Non-convex Minimization and an Application to Newton's Problem of the Body of Least Resistance, To appear in Ann. Inst. H. Poincaré.

[9] I. Newton, *Philosophiae Naturalis Principia Mathematica* (1686)

[10] J.-C. Rochet and P. Choné, Ironing, sweeping and multidimensionnal screening, *Econometrica* **66** (1998) pp. 783-826.

[11] G. Strang and G. Fix *An analysis of the finite elements method* Prenctice Hall 1973

[12] R.M. Dudley, On second derivatives of convex functions, *Mathematical Scandinavian*, (41), 159-174 (1977)