Optimization under second order constraints: are the finite element discretizations consistent?

Hervé Le Meur

Optimization under second order constraints: are the finite element discretizations consistent ?

Hervé V.J. Le Meur

CNRS, Laboratoire de Mathématiques d'Orsay,
Orsay cedex, F-91405; univ Paris-Sud, Orsay
cedex, F-91405.

Abstract

It is proved in [9] that the problem of minimizing a Dirichlet-like functional of the function u_h discretized with P_1 Finite Elements, under the constraint that u_h be convex, cannot converge to the right solution at least on a very wide range of meshes. In this article, we first improve this result by proving that non-convergence is due to a geometrical obstruction and remains local. Then, we investigate the consistency of various natural discretizations (P_1 and P_2) of second order constraints (subharmonicity and convexity). We also discuss various other methods that have been proposed in the literature.

1 Introduction

This paper is devoted to the numerical discretization of optimization with constraints on the second order derivative, namely problems of the form

$$\begin{cases} \inf J(u), \\ \mathbf{D}^2 u \in K, \end{cases} \tag{1}$$

where J is a functional and K is a subset of the symmetric matrices set. Such problems appear in various contexts, in particular in physics and economics.

Since Newton, the shape of minimal resistance has been a topic of interest. It is called the Newton's problem. With some additional assumptions, it amounts to looking for a concave function that minimizes a nonlinear functional. See for instance the original book [17], the historical survey in [11], more recent theoretical results in [3, 13, 15] or a numerical result [12].

The underlying problem of discretizing convex functions or bodies is indeed wider than expected. For instance, the Alexandrov's problem (see [4] and [12]), the Cheeger's constant [12, 5] and the Newton's problem [12] can be numerically studied.

In economics, it suffices to remember that utility functions are concave to see the very wide applicability of these problems. For instance, in [18], the authors are interested in finding the minimum of a convex and quadratic functional J over the set of convex functions

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

where $0 \leq \alpha \leq 1$ and $\mathbf{C}$ is a (2,2) symmetric positive definite matrix and K is given by

$$K = \{u \in H^1(\Omega), u \geq 0, u_x \geq 0, u_y \geq 0, u \text{ convex }\}. \tag{3}$$

So J is strictly convex and the set K is convex. It is then easy to check that there exists a unique minimizer of J over K. The regularity of the solutions is studied in [6].

Note that, when $\alpha = 1$, the problem degenerates: an exact solution exists, up to an additive constant by a convex function

$$\nabla u(\mathbf{x}) = \mathbf{C}^{-1}\mathbf{x} \Rightarrow u(\mathbf{x}) = \mathbf{x}'\mathbf{C}^{-1}\mathbf{x}/2 - \text{Cst}. \tag{4}$$

In [9], Choné and Le Meur extensively use this explicit solution to prove both theoretically and numerically an obstruction to the convergence of the mere *discretization* of this problem through conformal P_1 Finite Element (FE) method on a wide range of meshes. Here conformal means that the *discretized* function is supposed to satisfy exactly the continuous constraint. As a consequence of this result, no optimization process that would use such conformal P_1 discretization of both the functional and the constraint may converge for any solution on these meshes.

A possible approach to circumvent the mesh problem is, first, to test whether a sample of values on *given points* (and not *mesh*) may be associated to a convex function or body. Then one must construct the associated mesh (so, function dependent !) so as to interpolate. This was done in [14] for C^1 FE by Leung and Renka. But such a regularity is too restrictive for us. This article [14] reviews other papers, some of them being commented as false. It proves that the issue is not so simple.

The H_0^1 projection of a given function is addressed in [7] through saddle-point problems. In this article, the authors even weight the convexity constraint to enable the existence of a saddle-point. The computations appear to be more robust than expected by theory. An attempt of explanation is given in [16].

In [8], the authors describe and implement an algorithm that optimizes not in the set of discretized *and* convex (i.e. conformal) functions but in the set of the convex functions after discretization (so that they may be non-convex). More precisely they characterize, for a specific structured set of cartesian points, the image through the P_1 discretization of a (continuous) convex function. This yields such a huge number of constraints on the function that it may not be recommended.

Last in [12], Lachand-Robert and Oudet address the problem of discretizing a convex body. They use the parametric generalization in order to discretize convex function's graphs. In the functional, they isolate the dependence on the point x, the unit normal ν at x and the signed distance $\phi = \nu.x$. Although they notice that the three variables are "somehow redundant", they implement a gradient like method, based on the variation of only $\phi = x.\nu$. They address both the Alexandrov's problem, the Cheeger's sets and the Newton's problem. This was revisited since in [5].

Separately, Aguilera and Morin [1] prove convergence of a Finite Difference (FD) "approximation using positive semidefinite programs and discrete Hessians". In [2], the same authors prove even convergence of a Finite Element (FE) discretization of the weak Hessian.

Since FD are included in FE for convenient meshes, these two articles seem to contradict both [9] and the present article. We will discuss them below (section 5).

More recently, in [10], Ekeland and Moreno propose a non-local discretization that relies on the representation of any convex function as the supremum of affine functions (its minorants). So their "discrete" representation of continuous functions is conformal but non-local. The major drawback is that the complexity increases drastically with dimension, due to the nonlocality, but it works.

The present article deals with two main types of non-convergence results. On the one hand, the non-convergence in L^2 for conformal P_1 Finite Element (FE) is revisited after [9] from a theoretical point of view. On the other hand, the consistency of discretized linear optimization over second order constraints is investigated for various discretizations (P_1 and P_2) and various constraints: linear (subharmonicity) or nonlinear (convexity).

In Section 2, first we restate some already known results, then we prove that non-convergence is purely local. In Section 3, we investigate the consistency of various P_1 FE discretizations of our model problem. Section 4 is devoted to the consistency of P_2 FE discretizations (strong or weak convexity). We discuss the articles [1, 2] that could seem to contradict [9] in Section 5 and we conclude in Section 6.

2 The P_1 FEM

In the present section, we first recall some already known results and make them more explicit. Symbolic computations were done through a MAPLE worksheet available on the web page of the author.

2.1 Already known results

In [9], various results are proved. Since the goal in the present subsection is to extend this study, we need to remind the reader of these results. The Lemma 3 of [9] states:

Lemma 1 *A function u_h, P_1 in the rectangle $[a,b]^2$, is convex if and only if, for any pair of adjacent triangles*

$$(\mathbf{q}_2 - \mathbf{q}_1).\mathbf{n} \geq 0, \tag{5}$$

where $\mathbf{q}_1$ (resp. $\mathbf{q}_2$) is the (constant) gradient of u_h inside triangle 1 (resp. 2) and $\mathbf{n}$ is the unit normal pointing from triangle 1 to 2.

As reminded in [9], convexity of u can be defined dually as $- < \nabla u \otimes \nabla v > \succeq 0$ where $\mathbf{A} \succeq 0$ means the matrix $\mathbf{A}$ is semidefinite positive (SDP) and v is a test function in $\mathcal{C}_0^\infty$. This definition is strong since the test functions are in $\mathcal{C}_0^\infty$. One can easily check that the strong definition of convexity (with $\mathcal{C}_0^\infty$ test functions) implies the weak definition of convexity (with only the basis functions in P_1 associated to interior points as test functions). Yet, they are not equivalent. Such a weak P_1 convexity is the same as the "FE-convexity" defined by [2] (for P_1 FE).

3

The proof of Lemma 1 is based on the following formula written for u_h, a P_1 discretization of u:

$$< \frac{\partial^2 u_h}{\partial \mathbf{a} \partial \mathbf{b}}, \varphi >= \sum_e (\mathbf{q}_2 - \mathbf{q}_1).\mathbf{n} \,(\mathbf{n}.\mathbf{a})(\mathbf{n}.\mathbf{b}) \int_e \varphi(s)\mathrm{d}s, \tag{6}$$

where the summation is taken over all interior edges e of the mesh, $\mathbf{a}$ and $\mathbf{b}$ are two unit vectors, φ is a C^∞ function with compact support in Ω, $\mathbf{q}_1$ and $\mathbf{q}_2$ designate the two (constant) gradients of u_h in the two triangles that share the edge e and $\mathbf{n}$ is the unit normal from triangle 1 to triangle 2. This formula is right thanks to the property that the gradient of a continuous function across the edge has its tangential derivative along the edge continuous. Namely $(\mathbf{q}_2 - \mathbf{q}_1).\mathbf{t} = 0$ where $\mathbf{t}$ is the unit tangent vector.

The proof relies also on the fact that $(\mathbf{a}, \mathbf{b}) \mapsto (\mathbf{n}.\mathbf{a})(\mathbf{n}.\mathbf{b})$ is a bilinear form which quadratic form is semidefinite positive, whatever $\mathbf{n}$.

Through a proof very similar to the one of Theorem 4 in [9], one may prove a wider Theorem:

Theorem 2 *Let Ω be an open subset of $\mathbb{R}^2$ and $\mathcal{T}_h$ a triangulation of Ω. If there exists an open subset $\Omega' \subset \Omega$ such that the following property is satisfied in Ω' and for the mesh $\mathcal{T}_h$:*

$$(PM) \begin{cases} \exists (\mathbf{a}, \mathbf{b}) \text{ two independent unit vectors such that } (\mathbf{n}.\mathbf{a})(\mathbf{n}.\mathbf{b}) \geq 0 \\ \text{for all } \mathbf{n} \text{ unit normal to an edge of } \mathcal{T}_h \bigcap \Omega' \end{cases} \tag{7}$$

then any function u_h convex and P_1 will satisfy the following equation in the sense of distributions on Ω':

$$\frac{\partial^2 u_h}{\partial \mathbf{a} \partial \mathbf{b}} \geq 0. \tag{8}$$

The proof relies on the fact that if a function u_h is convex (in a dual definition with $\mathcal{C}_0^\infty$ test functions), then by Lemma 1, the gradient's jumps are non-negative. As a consequence, in Formula (6), if u_h is convex, the scalar coefficients $(\mathbf{q}_2 - \mathbf{q}_1).\mathbf{n}$ will be all non-negative. But separately, if condition (PM) is satisfied, $(\mathbf{n}.\mathbf{a})(\mathbf{n}.\mathbf{b}) \geq 0$ and so even the non-diagonal terms of the hessian $(\partial^2 u_h/(\partial \mathbf{a} \partial \mathbf{b}))$ will be sign-constrained !

Notice first that property (PM) only depends on the geometry of the mesh and not on any function. Notice then that the property (PM) is at given h but, if it remains for a sequence of $h \to 0$ in $\Omega' \subset \Omega$, then the associated sequence of function u_h given by Theorem 2 will satisfy (8) even at the limit in $\Omega' \subset \Omega$. Yet such a property (8) is contradictory with the property of convexity. This is the key of the obstruction to the convergence stated in [9]. The next subsection is devoted to discussing the generality of (PM).

2.2   Is property (PM) frequent ?

In this subsection, we start from some elementary observations on particular meshes, then we argue on whether such meshes are frequent and state a Theorem. Four types of meshes may be considered that are depicted in Figure (1).

Concerning mesh 1, three different unit normals (up to a multiplicative factor -1) exist in *all* the domain: $\{(0, 1); (1, 0); (-1/\sqrt{2}, 1/\sqrt{2})\}$. If we choose $\mathbf{a} = (-1, 0)$ and $\mathbf{b} = (0, 1)$, the values taken by $(\mathbf{n}.\mathbf{a})(\mathbf{n}.\mathbf{b})$ are $\{0; 0; (1/\sqrt{2})(1/\sqrt{2})\}$. They are all nonnegative.
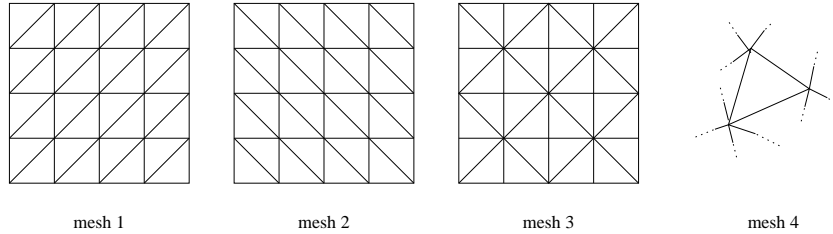
Figure 1: Four meshes.

Concerning mesh 2, three different unit normals (up to a multiplicative factor -1) exist in *all* the domain: $\{(0,1);(1,0);(-1/\sqrt{2},-1/\sqrt{2})\}$. It suffices to choose $\mathbf{a} = (1,0)$ and $\mathbf{b} = (0,1)$ to have the three values taken by $(\mathbf{n.a})(\mathbf{n.b})$ among $\{0;0;(-1/\sqrt{2})(-1/\sqrt{2})\}$. They are all nonnegative.

Concerning mesh 3, four types of unit normals can be found in *all* the domain: $\{(0,1); (1,0); (-1/\sqrt{2},1/\sqrt{2}); (1/\sqrt{2},1/\sqrt{2})\}$. If we choose $\mathbf{a} = (1,0)$ and $\mathbf{b} = (1/\sqrt{2},-1/\sqrt{2})$, the values taken by $(\mathbf{n.a})(\mathbf{n.b})$ are $\{1/\sqrt{2};0;(-1/\sqrt{2})(-1);0\}$. They are all nonnegative.

Any structured mesh like mesh 1, 2 or 3, if it is refined while keeping the same structure, will satisfy (PM) with the very same $\mathbf{a}$ and $\mathbf{b}$ for any h. But what can be stated about a more general mesh like mesh 4 ?

We are going to prove the following Theorem.

Theorem 3 *Let Ω be an open set of $\mathbb{R}^2$, $\Omega' \subset \Omega$ an open subset and $\mathcal{T}_h$ a triangulation of Ω.*
For any given h and Ω', there exist $(\mathbf{a}, \mathbf{b})$ such that (PM) is true in Ω' and for $\mathcal{T}_h$. Moreover, if the refinement process does not enrich the edges' directions of $\mathcal{T}_{h_n}$ in Ω', then the property (PM) will be true for any $\mathcal{T}_{h_n}$ in Ω'.

Proof of Theorem 3
First, one must notice that the condition $(\mathbf{n.a})(\mathbf{n.b}) \geq 0$ is invariant through changing $\mathbf{n}$ into $-\mathbf{n}$. So, it does not depend on the choice of the normal to the hyperplane.

Let us take a very general mesh like mesh 4 for which we may isolate a single triangle. Then, for each of the three unit normals (up to the multiplicative coefficient -1), having $(\mathbf{n.a})(\mathbf{n.b}) \geq 0$ is equivalent to having $\mathbf{a}$ and $\mathbf{b}$ in the *same* closed half plane whose normal is $\mathbf{n}$. Equivalently, we can take either $\mathbf{a}$ and $\mathbf{b}$ on the same side as $\mathbf{n}$ or on the opposite side. Gathering the conditions associated to each of the three edges, we are led to choosing $\mathbf{a}$ and $\mathbf{b}$ in the same half cone that contains none of the three hyperplane. Such a configuration is depicted on Figure (2). The three unit normals are drawn and called $n1, n2, n3$ and their associated hyperplanes are called $P1, P2, P3$ and are bold. Then one may choose the pair $(\mathbf{a}, \mathbf{b})$ both in any of the six half cones. So, one may easily find such a couple $(\mathbf{a}, \mathbf{b})$ for any given triangle.

Even better, whatever might be the finite number of edges for a larger Ω' and a given $\mathcal{T}_h$, one sees obviously that such a choice of $\mathbf{a}$ and $\mathbf{b}$ is easy since there will always be a finite number of hyperplanes. As a consequence, $\mathbf{a}$ and $\mathbf{b}$ can be choosen to be *both* in any of the cones which partition the whole space. So the property (PM) is true even on any *given* and potentially unstructured mesh of a subdomain of Ω.

Could a refinement process let the property (PM) be true for a given couple $(\mathbf{a}, \mathbf{b})$, a given Ω' and any $h \to 0$?

In case of any general given mesh $\mathcal{T}_h$ we have just proved that one may find $\mathbf{a}$ and $\mathbf{b}$, should they suit the property only in Ω'. Then, for instance, if the refinement process is
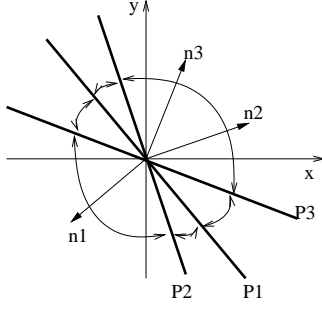
5

Figure 2: Unit normals and the associated hyperplanes in case of a triangle.

such that every triangle in Ω' is refined into four homothetic subtriangles, it is obvious that no more direction of edges will be provided. As a consequence, the very same $\mathbf{a}$ and $\mathbf{b}$ will then suit property (PM) in Ω' for *any* refined mesh and *any* h. As stated in the Theorem, if the refinement process does not enrich the edges, the $(\mathbf{a}, \mathbf{b})$ will suit and the proof is complete.

$\square$

In a sense to be defined, the set of refinement processes that enable (PM) is open and non-empty. For those interested in refinement processes, we conjecture a sufficient condition although we do not know whether there could exist a refinement process satisfying this property. It means that any segment has a normal unit vector which can be approximated by normals of successive edges of the mesh's triangles (for space step sufficiently small).

Conjecture 4 *Let $\mathcal{T}_{h_n}$ and its refinement process of the open set Ω in $\mathbb{R}^2$ be such that*

$\forall [b_0, b_1] \subset \Omega' \subset \Omega$ *and its unit normal* $\mathbf{n}, \forall V_0 \in \mathcal{V}(b_0), \forall V_1 \in \mathcal{V}(b_1), \forall V \in \mathcal{V}([b_0, b_1])$,
$\forall \varepsilon > 0, \exists N \geq 0, \forall n \geq N, \exists \gamma_n,$ *a continuous union of edges from the mesh* $\mathcal{T}_{h_n}$ *such that*
$\gamma_n(0) \in V_0, \gamma_n(1) \in V_1, \gamma_n(t) \in V$ *for all* t *and* $\int_0^1 |< \mathbf{n}, \mathbf{n}(t) >| \, dt \geq 1 - \varepsilon,$

where $\mathbf{n}(t)$ is a unit normal to the path γ_n.
Then there exists no $(\mathbf{a}, \mathbf{b})$ such that property (PM) be true for any h_n.

2.3   More precision in non-convergence

In [9], the authors state that the "conformal method may not converge for some limit function" because the second derivative of the limit is forced to satisfy an unnatural condition. In this subsection, we give a more precise result in L^2 and still use our Lemma 1 that enables to identify the convexity of u_h and the non-negativity of its gradients' jumps (for P_1 FE). This enables us to state our main Theorem:

Theorem 5 *Let Ω be an open domain in $\mathbb{R}^2$, Ω' an open subdomain of Ω and a family of meshes $(\mathcal{T}_h)_{h \to 0^+}$. If the property (PM) is satisfied in Ω' and for the meshes $(\mathcal{T}_h)_{h \to 0}$, then there exists $\varepsilon > 0$ and a $\mathcal{C}^\infty$ convex function u_{exact} such that*

$$\min_{u_h \in P_1 \ and \ u_h \ convex} | u_h - u_{exact} |_{L^2} \geq \varepsilon.$$

This Theorem uses that if Property (PM) is satisfied in a domain Ω for a family of meshes $(\mathcal{T}_h)_{h \to 0^+}$, should it be only locally in Ω', there will be two unit vectors $\mathbf{a}, \mathbf{b}$ such that (8) holds. Given those $\mathbf{a}, \mathbf{b}$, there exists a function u_{exact} that may not be the limit in L^2 of *any* sequence of functions both P_1 and convex (in a strong definition) on *this* family of meshes.

We need some more preliminary results and definitions before proving our result concerning P_1 FE.

Definition 1 *Let Ω be an open subset of $\mathbb{R}^2$, and $u \in \mathcal{C}^0(\Omega)$. For any $\mathbf{x} = (x, y) \in \Omega$ and $(\mathbf{a}, \mathbf{b})$ two independent unit vectors, there exists $(\alpha_0, \beta_0) \in \mathbb{R}^{+*2}$ such that*

$$\forall \alpha, \ 0 \le \alpha \le \alpha_0, \forall \beta, \ 0 \le \beta \le \beta_0,$$
$$\mathbf{x} + \alpha \mathbf{a} + \beta \mathbf{b} = (x + \alpha a_1 + \beta b_1, y + \alpha a_2 + \beta b_2) \in \Omega.$$

For such $(\alpha_0, \beta_0, \mathbf{a}, \mathbf{b})$, one defines:

$$\phi_{(\alpha_0, \beta_0, \mathbf{a}, \mathbf{b})}(u) = (u(\mathbf{x} + \alpha_0 \mathbf{a} + \beta_0 \mathbf{b}) - u(\mathbf{x} + \alpha_0 \mathbf{a}) - u(\mathbf{x} + \beta_0 \mathbf{b}) + u(\mathbf{x})) / (\alpha_0 \beta_0).$$

If $\alpha_0 = \beta_0 \to 0^+$, then $\phi_{(\alpha_0, \beta_0, \mathbf{a}, \mathbf{b})}(u) \to \partial^2 u / (\partial \mathbf{a} \partial \mathbf{b})$. Then $\phi_{(\alpha_0, \beta_0, \mathbf{a}, \mathbf{b})}(u)$ may be considered as a kind of double integral of $\partial^2 u / (\partial \mathbf{a} \partial \mathbf{b})$. We are going to prove that such a quantity $\phi_{(\alpha_0, \beta_0, \mathbf{a}, \mathbf{b})}(u)$ will be overconstrained by the mere discretization and the limit u will not satisfy the right sign of the second derivative.

We want now to define an explicit solution depending on some parameters. A well-chosen combination of these parameters will trigger a non-approximable function. Let $\Omega =]1, 2[^2$ and the problem (2, 3) for a positive symmetric definite matrix $\mathbf{C}$ be such that:

$$\mathbf{C} = \begin{pmatrix} \mu_2 & \rho \\ \rho & \mu_1 \end{pmatrix}. \tag{9}$$

Then there exists an exact solution thanks to (4):

$$u_{exact} = \frac{1}{\mu_1 \mu_2 - \rho^2} \left(\mu_1 (x^2 - 1)/2 + \mu_2 (y^2 - 1)/2 - \rho(xy - 1) \right) \text{ in } \Omega. \tag{10}$$

The function u_{exact} is such that it is zero at the corner $(1, 1)$ of the domain Ω chosen but it can easily be generalized to other domains. Simple computations prove the following formula:

$$\phi_{(\alpha_0, \beta_0, \mathbf{a}, \mathbf{b})}(u_{exact}) = \frac{1}{\mu_1 \mu_2 - \rho^2} \left(\mu_1 a_1 b_1 + \mu_2 a_2 b_2 - \rho(a_1 b_2 + a_2 b_1) \right) = \mathbf{a}' \mathbf{C}^{-1} \mathbf{b}, \tag{11}$$

where a_1, a_2, b_1, b_2 are the components of $\mathbf{a}, \mathbf{b}$. It is then useful to state the following Lemma.

Lemma 6 *Let $\mathbf{a}, \mathbf{b}$ be two given independent unit vectors in $\mathbb{R}^2$. Let $\Omega \subset \mathbb{R}^2$ and $\eta > 0$ given. Then, there exists (μ_1, μ_2, ρ) such that $\mu_1 \ge 0, \mu_2 \ge 0, \mu_1 \mu_2 - \rho^2 \ge 0$ and*

$$\phi_{(\alpha_0, \beta_0, \mathbf{a}, \mathbf{b})}(u_{exact}) = \mathbf{a}' \mathbf{C}^{-1} \mathbf{b} \le -\eta \text{ in } \Omega, \tag{12}$$

provided $(\alpha_0, \beta_0) \in \mathbb{R}^{+*2}$ are such that there exists $\phi_{(\alpha_0, \beta_0, \mathbf{a}, \mathbf{b})}(u_{exact})$.

7

Roughly speaking, Lemma 6 claims that for any independent $\mathbf{a}, \mathbf{b}$, one may find a positive definite matrix $\mathbf{C}^{-1}$ such that the associated u_{exact} satisfies

$$\phi_{(\alpha_0, \beta_0, \mathbf{a}, \mathbf{b})}(u_{exact}) = \mathbf{a}' \mathbf{C}^{-1} \mathbf{b} < 0.$$

Proof of Lemma 6
From the formula (11), one sees that it is sufficient to find a positive symmetric bilinear form $(\mathbf{C}^{-1})$ such that, for independent $\mathbf{a}, \mathbf{b}$ given, $\mathbf{a}' \mathbf{C}^{-1} \mathbf{b} < 0$. We will build up $\mathbf{C}^{-1}$ from its eigenvalues and eigenvectors.

Let $\mathbf{e}_1$ be a unit vector between $\mathbf{a}$ and $\mathbf{b}$ (normalized mean for instance). Then let $\mathbf{e}_2$ be a unit vector normal to $\mathbf{e}_1$. With the same notations for $\mathbf{a}$ and $\mathbf{b}$ in this new basis as above, $a_1 b_1 > 0$ and $a_2 b_2 < 0$.

Let now $\mathbf{C}^{-1}$ be a matrix in the canonical basis with the eigenvectors $\mathbf{e}_1, \mathbf{e}_2$ and the associated positive eigenvalues λ_1, λ_2. Obviously, $\mathbf{C}$ is positive semidefinite. Then $\mathbf{a}' \mathbf{C}^{-1} \mathbf{b} = \lambda_1 a_1 b_1 + \lambda_2 a_2 b_2$ which may be less than $-\eta$ given, for an appropriate choice of λ_1, λ_2. Then the (μ_1, μ_2, ρ) are the coefficients of the matrix $\mathbf{C}$ in the canonical basis given by (9).

$\square$

We have now proved all what is needed to start the following proof.

Proof of Theorem 5.
Since we assume (PM) is satisfied for $(\mathcal{T}_h)_{h \to 0}$ and $(\mathbf{a}, \mathbf{b})$ in Ω', Theorem 2 enables us to claim that any P_1 function u_h satisfies (8) in the sense of distributions in Ω':

$$0 \ \le\ <\frac{\partial^2 u_h}{\partial \mathbf{a} \partial \mathbf{b}}, \Psi>_\Omega = <u_h, \frac{\partial^2 \Psi}{\partial \mathbf{a} \partial \mathbf{b}}>_\Omega, \tag{13}$$

for any nonnegative $\Psi \in \mathcal{C}_0^\infty(\Omega')$. This property obviously remains for open subdomains of Ω'.

Should it be needed, one could decrease $\alpha_0, \beta_0 > 0$ and find a subdomain $\Omega'' \subset \Omega'$ such that for any nonnegative $\varphi \in \mathcal{C}_0^\infty(\Omega'')$, the function

$$\Psi : \ \mathbf{x} \mapsto \int_0^1 \int_0^1 \varphi(\mathbf{x} - t\alpha_0 \mathbf{a} - t'\beta_0 \mathbf{b}) \, dt \, dt', \tag{14}$$

be nonnegative and in $\mathcal{C}_0^\infty(\Omega')$. So the function Ψ is eligible for equation (13). The second derivative of Ψ may be computed:

$$\begin{aligned}\frac{\partial^2 \Psi}{\partial \mathbf{a} \partial \mathbf{b}} &= (\varphi(\mathbf{x} - \alpha_0 \mathbf{a} - \beta_0 \mathbf{b}) - \varphi(\mathbf{x} - \beta_0 \mathbf{b}) - \varphi(\mathbf{x} - \alpha_0 \mathbf{a}) + \varphi(\mathbf{x})) / (\alpha_0 \beta_0) \\ &= \phi_{(\alpha_0, \beta_0, -\mathbf{a}, -\mathbf{b})}(\varphi).\end{aligned}$$

So we have, for any nonnegative $\varphi \in \mathcal{C}_0^\infty(\Omega'')$:

$$\begin{aligned}0 \le\ <\frac{\partial^2 u_h}{\partial \mathbf{a} \partial \mathbf{b}}, \Psi>_\Omega \ \ &= <u_h, \frac{\partial^2 \Psi}{\partial \mathbf{a} \partial \mathbf{b}}>_\Omega \\ &= <u_h, \phi_{(\alpha_0, \beta_0, -\mathbf{a}, -\mathbf{b})}(\varphi)>_{\Omega''} \\ &= <\phi_{(\alpha_0, \beta_0, \mathbf{a}, \mathbf{b})}(u_h), \varphi>_{\Omega''} .\end{aligned} \tag{15}$$

Separately, given $\mathbf{a}, \mathbf{b}, \eta > 0$, Lemma 6 enables us to claim there exists a convex quadratic function u_{exact} such that (12) is true for $\eta > 0$ given. Moreover, because

of (15), since the open set $\Omega'' \subset \Omega'$ is of non-zero measure, and $\phi_{(\alpha_0,\beta_0,\mathbf{a},\mathbf{b})}(u_{exact})$ is a constant known by (11), we have

$$< \phi_{(\alpha_0,\beta_0,\mathbf{a},\mathbf{b})}(u_h - u_{exact}), \varphi >_{\Omega'} \geq -\phi_{(\alpha_0,\beta_0,\mathbf{a},\mathbf{b})}(u_{exact}) \int_\Omega \varphi.$$

As a consequence, for convenient Ω'', we are led to

$$< \phi_{(\alpha_0,\beta_0,\mathbf{a},\mathbf{b})}(u_h - u_{exact}), \varphi >_{\Omega''} \geq \eta \int_\Omega \varphi, \tag{16}$$

for any *nonnegative* $\varphi \in \mathcal{C}_0^\infty(\Omega'')$ and u_h convex and P_1. This last inequality contradicts any possible convergence in $L^2(\Omega'')$ of a sequence u_h of convex functions P_1 in $(\mathcal{T}_h)_h$ to u_{exact} given by Lemma 6.

□

The previous Theorem adds one more argument to the need for convenient numerical tools to discretize the constraint of convexity.

3   The P_1 FEM

First we study a strong discretization of convexity then a weak one.

3.1   Use of the gradients' jumps for convexity

The non-convergence of the convexity problem (2, 3) discretized by conformal P_1 Finite Elements is proved and numerically illustrated in [9] and is made more precise in the subsection 2.3. We give here a different argument.
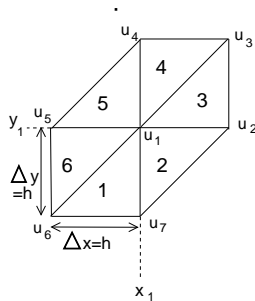

Figure 3: Local shape of the mesh close to (x_1, y_1).

Given a function u, one may interpolate it in P_1 FE as $u_h = \sum_{i=1}^N u_{hi} \phi_i(x)$. If the mesh is structured like in Figure 3, should it be local, one may compute the gradients' jumps accross the edges as functions of the values at the various involved nodes. Then one may compute the series expansion of the sampled values u_i from the exact initial function u. The jumps between triangles 1 and 2, 2 and 3 and last 3 and 4 are respectively:

$$\begin{array}{rcl}
\text{Jump}(1/2) = \frac{-u_6+u_7+u_1-u_2}{h} & = & (u_{xx}+u_{xy})h + O(h^2), \\
\text{Jump}(2/3) = \frac{-u_1+u_7+u_3-u_2}{h} & = & (u_{xy}+u_{yy})h + O(h^2), \\
\text{Jump}(3/4) = \frac{-u_1+u_2+u_4-u_3}{h} & = & -\sqrt{2}*(u_{xy})h + O(h^2).
\end{array} \tag{17}$$

When $h \to 0$, instead of forcing the solution of a problem to be convex, we force it to some mesh-dependent combination of its second order derivatives to have a sign or another. In

addition, this combination is meaningless since for the convex limit function, it may have whatever sign. We will say that such a discretization is not consistent.
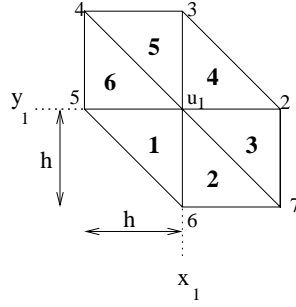

Figure 4: Local shape of the second mesh close to (x_1, y_1) and global numbering.

One could also study the use of gradients' jumps on the very different mesh like in Figure 4. With similar computations available on the web site of the author, one may prove the following Proposition.

Proposition 7 *The gradients' jumps on a mesh given in Figure 4 between triangles 1 and 2, 2 and 3 and last 3 and 4 are respectively:*

$$
\begin{array}{rcl}
Jump(1/2) = \frac{-u_1 + u_5 - u_6 + u_7}{h} & = & (u_{xx} - u_{xy})_{x=x_1, y=y_1} h + O(h^2), \\
Jump(2/3) = \frac{(u_6 - u_7 - u_1 + u_2)\sqrt{2}}{h} & = & u_{xy}\sqrt{2}h + O(h^2), \\
Jump(3/4) = \frac{u_7 - u_2 - u_1 + u_3}{h} & = & (u_{yy} - u_{xy})_{x=x_1, y=y_1} h + O(h^2).
\end{array}
\tag{18}
$$

Such a discretization of convexity is not consistent.

Below, we investigate the consistency of various other discretizations.

3.2 Use of a weak version

We use here a weak P_1 definition of convexity which is identical to the one of [2] which is fully discussed in Section 5.

3.2.1 The subharmonicity constraint

A weak definition of subharmonicity ($\Delta u \geq 0$) is, for any test function ϕ_i in the discrete basis:

$$
< \Delta u_h, \phi_i >= \mathrm{Tr} \ < \mathbf{D}^2 u_h, \phi_i >= -\mathrm{Tr} \ < \nabla u_h \otimes \nabla \phi_i >\geq 0.
\tag{19}
$$

One may then prove the following Proposition:

Proposition 8 *The weak P_1 discretization of the subharmonicity constraint on a mesh like in Figure 3 is consistent:*

$$
< \Delta u_h, \phi_i >= -4u_1 + u_2 + u_5 + u_7 + u_4 = (u_{xx} + u_{yy})_{(x=x_1, y=y_1)}h^2 + O(h^3),
\tag{20}
$$

where i is the node at the center of the cell's group (x_1, y_1).

10

The proof is very easy and left to the reader. Indeed, it is only the discretization of the Laplacian which is known to be consistent and even convergent !

In order to test this discretization, we used the Matlab package `optim` to minimize the functional $\int_\Omega |\nabla u|^2/2 + fu$, where $f = \Delta u_{exact}$, over the set of subharmonic functions. The exact solution of this H_0^1 projection is u_{exact}. The convergence with the procedure `quadprog` of quadratic programming can be seen on Figure 5. It is quite satisfactory.
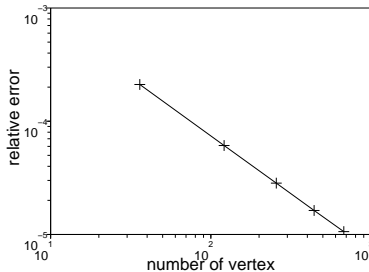

Figure 5: convergence of P_1 FE in case of a subharmonicity constraint.

3.2.2 The convexity constraint

One may give a weak definition of convexity:

$$\text{Tr } < \mathbf{D}^2 u_h, \phi_i > \geq 0 \text{ and } \det < \mathbf{D}^2 u_h, \phi_i > \geq 0, \tag{21}$$

for any ϕ_i, basis function of the FEM at the node i. One may then prove the following Proposition:

Proposition 9 *The weak P_1 discretization of the convexity constraint (21) on a mesh like in Figure 3 is consistent as can be seen from (20) and*

$$\begin{aligned}
\det < \mathbf{D}^2 u_h, \phi_i > &= \det < \nabla u_h \otimes \nabla \phi_i > = \\
&= (-2u_1 + u_2 + u_5)(u_7 + u_4 - 2u_1) - (-u_7 + u_3 - u_4 + u_6 - u_5 - u_2 + 2u_1)^2/4 \\
&= (u_{xx}u_{yy} - u_{xy}u_{yx})h^4 + O(h^5).
\end{aligned} \tag{22}$$

where i is the node at the center of the cell's group.

This discretization (21) for the convexity contraint appears to be consistent although the numerical treatment of the linear and nonlinear constraints should raise inaccuracies since they are of very different orders of magnitude.

4 The P_2 FEM

In the present section, we investigate successive discretizations through P_2 FEM of two second order contraints: subharmonicity and convexity.

First, we interpolate a continuous function u to a P_2 function $u_h = \sum_i u_i \phi_i(x)$ in a domain Ω meshed with triangles. Here, the index i denotes both vertices and edge midpoints. Then we compute the discretized version of the second order term constrained to be nonnegative (various versions are treated) and compute its series expansion. If

forcing it to be nonnegative amounts to forcing the continuous limit function to the correct constraint, then we claim the discretization is consistent. Otherwise it is inconsistent.

For the whole section, we assume the mesh is (at least locally) structured around the point of coordinates (x_1, y_1). The local numbering of triangles is depicted in Figure 6. The node (x_1, y_1) is localy numbered 1 in every triangle and the local numbering of nodes is depicted in triangle 3 of Figure 6.
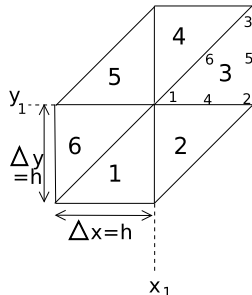

Figure 6: Local shape of the mesh close to (x_1, y_1). Local numbering in triangle 3.

4.1 Gradients jumps for the convexity constraint

We use here a strong definition of convexity. In a way similar to the P_1 case, one may prove for P_2 functions u_h:

$$< \frac{\partial^2 u_h}{\partial \mathbf{a} \partial \mathbf{b}}, \varphi > = \sum_e \int_e (\mathbf{q}_2(s) - \mathbf{q}_1(s)).\mathbf{n}\, \varphi(s) \mathrm{d}s\, (\mathbf{n}.\mathbf{a})(\mathbf{n}.\mathbf{b}) + \sum_K \frac{\partial^2 u_h}{\partial \mathbf{a} \partial \mathbf{b}}|_K \int_K \varphi,$$

for any e interior edge of the mesh, $\mathbf{a}$ and $\mathbf{b}$ are two unit normal vectors, φ is a C^∞ function with compact support in Ω. By taking φ localized along the edge e, one may state that this strong convexity (with $\mathcal{C}_0^\infty$ test functions) implies the non-negativity of the gradients' jumps at least in an integral form. This strong definition implies the weak P_2 definition where one takes the P_2 basis functions associated to the edge midpoints as test functions (as do [2]).

In order to test the consistence of such a discretization, we computed the gradients' jumps accross the edges common to triangles 1 and 2, 2 and 3 and last 3 and 4. Of course, they are not constant as they are P_1 FE. After an exact computation and a series expansion (details may be found on a MAPLE worksheet available on the web page of the author), one may state the following Proposition

Proposition 10 *The discretization of the convexity constraint with the jump of the gradients between triangles 1 and 2, 2 and 3, 3 and 4 on a mesh like in Figure 6 gives terms*

$$\begin{aligned} Jump(1/2) &= (u_{xxy} + u_{xyy})(x_1, y_1)(y - y_1)h/2 + O(h^2); \ y \in [y_1 - h, y_1] \\ Jump(2/3) &= (u_{xxy} + u_{xyy})(x_1, y_1)(x - x_1)h/2 + O(h^2); \ x \in [x_1, x_1 + h] \\ Jump(3/4) &= -(u_{xxy} + u_{xyy})(x_1, y_1)(x - x_1)\sqrt{2}h/2 + O(h^2); \ x \in [x_1, x_1 + h]. \end{aligned} \qquad (23)$$

Such a discretization is non-consistent.

One deduces from (23) that forcing the non-negativity of the gradient's jumps forces the limit function to satisfy the non-natural equality condition $u_{xxy} + u_{xyy} = 0$! So the P_2 discretization of the strong definition of convexity ($\mathcal{C}_0^\infty$ test functions) implies the non-negativity of gradient's jumps which is non-consistent and so must be rejected. This does not enable to reject the weak P_2 definition of convexity with the basis functions associated to the edge midpoints as test functions because it is only implied by the strong one. We will check below that the weak definition is consistent (subsection 4.3).

4.2   Weak version of the second derivative at a vertex

We define the weak (with P_2 test functions like in [2]) version of the Hessian at vertex i as:

$$< \mathbf{D}^2 u_h, \phi_i > = - \begin{pmatrix} \int_\Omega \frac{\partial u_h}{\partial x} \frac{\partial \phi_i}{\partial x} & \int_\Omega \frac{\partial u_h}{\partial x} \frac{\partial \phi_i}{\partial y} \\ \int_\Omega \frac{\partial u_h}{\partial y} \frac{\partial \phi_i}{\partial x} & \int_\Omega \frac{\partial u_h}{\partial y} \frac{\partial \phi_i}{\partial y} \end{pmatrix}. \tag{24}$$

Strictly speaking, this may not provide a correct weak version of non-negativity since ϕ_i changes sign. In addition, using only the vertices functions as test functions would provide a (too ?) small number of constraints. Anyway, one may state the following Proposition which proof is left to the reader:

Proposition 11 *The discretization of the linear part of the convexity constraint according to (24) on a mesh like in Figure 6 ($h = \Delta x = \Delta y$) gives:*

$$Tr\left(< \mathbf{D}^2 u_h, \phi_i >\right) = -(u_{xxxx} + u_{yyyy})(x_1, y_1) * h^4/48 + O(h^5). \tag{25}$$

Such a discretization is non-consistent.

It appears that such a discretization is not even consistent for the linear part of the constraint. More precisely, while one could believe one forces the solution to be subharmonic, indeed, one forces it to be such that $u_{xxxx} + u_{yyyy} \leq 0$! The full nonlinear convexity constraint on the same vertices may only work worse. This proves that the proof of convergence of FE methods by [2], which fails if the test functions change sign, maynot be improved on that point.

One must notice that the fourth order of the expansion in (25) is meaningful. Indeed, on the one hand the three midpoints quadrature is exact for P_2 functions in a triangle. On the other hand the basis function for a vertex vanishes on these midpoints. As a result, the order two term is identically zero. So the first non-zero term is the fourth one and the constraint is of fourth order too.

4.3   Weak version of the second derivative at an edge midpoint

We define the weak version of the second derivative at an edge midpoint (denoted by index j) in a way very similar to (24). Indeed, the function ϕ_i in (24) is replaced by the basis function ϕ_j associated to an edge midpoint indexed by j.

Unlike in subsection 4.2, the basis functions (of the edge midpoints) are non-negative. So it is *a priori* an admissible weak formulation of semidefinite positiveness.

4.3.1 The subharmonicity constraint

Let us assume we discretize the constraint $\Delta u \geq 0$ by

$$\int_\Omega \nabla u_h . \nabla \phi_j \leq 0, \tag{26}$$

for all j index of an edge midpoint interior to Ω. One may then state a Proposition (which proof is left to the reader):

Proposition 12 *The discretization of the linear part of the convexity constraint according to (26) on a mesh like in Figure 6 ($h = \Delta x = \Delta y$) gives the same series expansion, whether the edge is vertical, horizontal or diagonal:*

$$Tr \left(< \mathbf{D}^2 u_h, \phi_j > \right) = \Delta u(x_1, y_1) h^2/3 + O(h^3), \tag{27}$$

for any j index of an edge midpoint interior to Ω. Such a discretization is consistent.

One question remains: if we discretize the subharmonicity (or the convexity) constraint only at edge midpoints, is it enough constraints or not ?

4.3.2 The convexity constraint

Like in the subharmonic case, we take a weak version of the continuous nonlinear constraint $\det D^2 u \geq 0$ with (nonnegative) test functions associated to every edge midpoint j in the interior. One may then prove easily the following Proposition:

Proposition 13 *The discretization of the nonlinear part of the convexity constraint on a mesh like in Figure 6 ($h = \Delta x = \Delta y$) gives:*

$$\det(< \mathbf{D}^2 u_h, \phi_j >) = \left(\frac{\partial^2 u}{\partial x^2} \frac{\partial^2 u}{\partial y^2} - \frac{\partial^2 u}{\partial x \partial y} \frac{\partial^2 u}{\partial y \partial x} \right)(x_1, y_1) h^4/9 + O(h^5), \tag{28}$$

for any j index of an edge midpoint interior to Ω. Such a discretization is consistent.

So the linear constraint is of order two while the nonlinear is of order four for convergence. Such a discrepancy should be managed in numerical simulation.

5 Discussion on the Aguilera and Morin's articles

Let us notice that [9] seems not to have been known of the authors of [1, 2].

Roughly speaking, the first article [1] proves convergence of the FD discretization of problems like ours. The second [2] proves convergence of the FE discretization of the same problems if it is at least P_2 (and some more assumptions that will prove to be contradictory). So, the proof for FE does not include P_1 FE as a by-product, and so, in a sense, nor the FD. But this is not sufficient to claim the proof of FD is wrong.

5.1 Finite Differences

In [1], Aguilera and Morin deal with FD and "prove convergence under very general conditions, even when the continuous solution is not smooth". But since FD discretizations trigger the very same matrices as the P_1 FE for a mesh like our mesh 1 on Figure 1 (except boundary conditions irrelevant here), there seems to be a contradiction with [9]. This forces to study further.

We could not find any error in their proof, which relies on approximation theory. So we looked over the numerical experiments. While they claim the error in the L^∞ norm on the monopolist problem "is smaller or approximately equal to h", their table of convergence is reproduced in a loglog scale in the left part of Figure 7. It is not convincing.

Moreover, they provide also the error table for the 3D monopolist problem which can be seen on the right part of Figure 7. As they notice, "the L^∞ error is not converging to zero with order $O(h)$" and even does not converge at all. Moreover the execution time grows faster than polynomialy.

The next article [2] gives one example of P_1 discretization and a weak definition of convexity (with P_1 test functions) that does not converge neither (3.7 p. 3150).
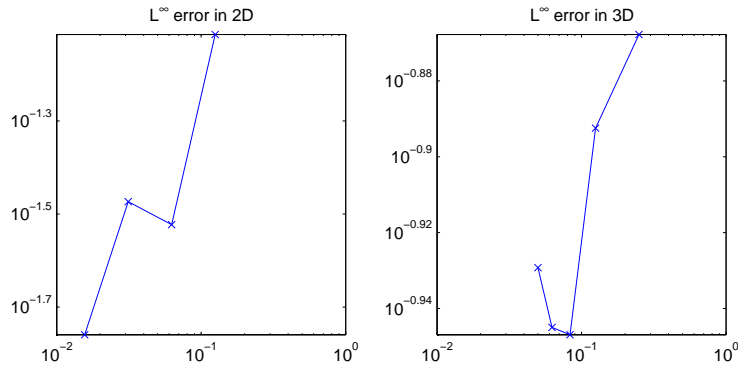

Figure 7: L^∞ error versus space step h

5.2 Finite Elements

In [2], Aguilera and Morin prove convergence of the discretization of the full problem: not only the approximation of the Hessian, but also of the functional together, but under some (contradictory) conditions.

Especially, they define u_h (interpolated on a given mesh) to be FE-convex (with respect to the test-functions basis $\{\phi_s^h\}_s$) if and only if the discrete Hessian satisfies

$$H_s^h u_h = - \left(< \partial_i u_h, \partial_j \phi_s^h > \right)_{i,j} \succeq 0,$$

for all ϕ_s^h in the test-functions basis, where they denote $\mathbf{A} \succeq 0$ if $\mathbf{A}$ is semidefinite positive (SDP). Indeed, in their proof ((3.1) or the equation before (3.2) for instance), they extend this definition to a u (and not a u_h) that only needs to be H^1 and not in the discrete space V_h. It amounts to assuming that the continuous operator $u \mapsto H_v u = - < \nabla u, \nabla v >$ for v in H_0^1 (or $\mathcal{C}_0^\infty$ as defined in their (2.3, 2.4)) and the discrete operator (on a given mesh !) $u_h \in P_1 \mapsto - < \nabla u_h, \nabla \phi_s^h >$ are identical. Of course there is no difference on the u or u_h since they are H^1. But the test functions are different. This difference will be

highlighted in the next subsection 5.3. Then their main result states convergence of the discretization once one assumes at least (p. 3147):

C.2 There exists a linear operator $\mathcal{I}^h$ with values in the discretization space V_h (the interpolant), an integer $m \geq 2$, and a constant C independent of u and h such that

$$\| u - \mathcal{I}^h u \|_{H^1(\Omega)} \leq C h^{m-1} \| u \|_{H^m(\Omega)},$$

C.3 The basis test-functions are such that

$$\phi_s^h(x) \geq 0 \text{ for all } x \in \Omega, \text{ all } h > 0, \text{ and all } s.$$

In their article, they split the proof of convergence of the constraint from the proof of convergence of the functional. An analogy could be done with the incompressibility condition (div $u = 0$) on Navier-Stokes equations. Separately, the Navier-Stokes and the incompressibility equations can be discretized and converge. Yet, there are compatibility conditions (Babushka-Brezzi condition) for the discretization of all the fields together, namely velocity and pressure. We do believe one could find such a compatibility condition when minimizing a functional under convex constraint.

In addition to C.2 and C.3, the proof requires $m > 2$ (even if one improves the proof by a $L^2 - L^2$ instead of $L^\infty - L^1$ duality in their (3.3)). Then, condition C.2 forces to use at least P_2 FE (and higher). But then, since some test functions of the associated basis change sign (as seen by the authors p. 3150), assumption C.3 forces to use *at most* P_1 FE ... So the theorem is right but unusable since the assumptions never may occur together.

This is seen by the authors who say they "need to elaborate" on the condition $m > 2$. On their p. 3150, they notice the FE discretized Hessian expands into the FD scheme (for one specific mesh !) at first order and so they refer to their FD article to conclude that their Theorem "also holds for $m = 2$" (or P_1 FE).

The very next example they provide is P_1 and they state "since u is convex, the projection u_h should converge to u as $h \to 0$, but this is not the case in this example.". They summarize this result by saying "although there is some sort of super-convergence for some meshes, for general meshes[...] FE-convex piecewise *linear* function may not suffice". We do not share this analysis and consider this example as a counter-example. This non-convergence weakens also [1].

All this is even complexified when the authors cope with numerical experiments of P_2 FE discretization. They notice the FE basis functions are not all nonnegative. So they "considered the usual piecewise *linear* nodal basis for the vertices" (p. 3152). In other words, they replace some P_2 basis functions with P_1 basis functions with space step divided by two. But the P_1 FE basis being not proved to converge, and even numerically proved not to converge in some cases, the proof does not apply to these experiments.

Their tables of convergence are transformed into loglog graphics in Figure 8. One may not conclude the discretization converges.
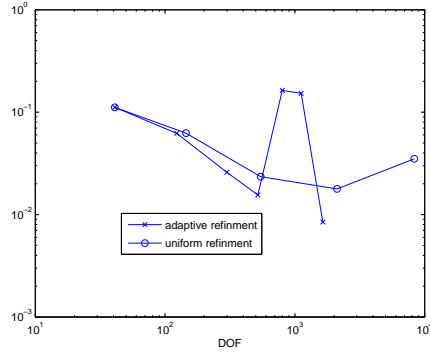
16

Figure 8: L^∞ error versus Degrees of Freedom

5.3 Are [1, 2] contradictory with [9] ?

The answer is no. But it deserves to be explained.

Both [2] and [9] use a dual definition of convexity and have conclusions that could seem to be contradictory. The only difference is that [2] uses the discretized basis as test functions:

$$- < \nabla u_h \otimes \nabla \phi_s^h > \succeq 0,$$

while [9] uses C_0^∞ as test functions:

$$- < \nabla u_h \otimes \nabla \phi > \succeq 0.$$

The difference appears when [9], and us in the present article, take $\phi \in C_0^\infty$ localized on an edge e of the mesh, use (6) for P_1 FE and then, forcing convexity amounts to forcing the matrix

$$(\mathbf{q}_2 - \mathbf{q}_1).\mathbf{n} \begin{pmatrix} n_1^2 & n_1 n_2 \\ n_1 n_2 & n_2^2 \end{pmatrix},$$

to be SDP, where $\mathbf{n} = (n_1, n_2)$ is the unit normal to the edge e from triangle 1 to triangle 2 and $\mathbf{q}_1$ (resp. $\mathbf{q}_2$) is the (constant) gradient of u_h in triangle 1 (resp. 2). But since the $\mathbf{n}$ components matrix is SDP whatever the normal, forcing the semidefiniteness of $- < \nabla u_h \otimes \nabla \phi >$ forces the *scalar* $(\mathbf{q}_2 - \mathbf{q}_1).\mathbf{n}$ to be nonnegative on this localized edge e (in P_1). Then, [9] and the present article prove that this strong convexity will prohibit convergence for P_1 FE discretization of some given u on *some* given meshes (not all !) since u_h (the P_1 interpolation of the sampled u) is overconstrained.

By using only a test function in the finite basis, Aguilera and Morin may not localize as Choné and Le Meur [9] and so do not fall on their obstruction. In the formula (6), a $\phi = \phi_h \in P_1$ associated to any vertex has various edges in its support. So the informations on various edges are melted if the test functions are P_1 (weak convexity) while they are isolated if the test functions are $\mathcal{C}_0^\infty$ (strong convexity).

We proved in subsection 4.2 that some weak P_2 definition of convexity is not even consistent if the test functions do no satisfy condition C3. So the result of [2] maynot be improved on that assumption. Anyway, the proof of convergence for FE maynot be used since the assumptions never may occur together. We are still annoyed at the numerical results of [1] for FD discretizations that do not match the claim of convergence although the proof is not contradictory with [9].

6 Conclusion

In this article we prove, from a point of view complementary to [9], that the P_1 discretization of a function that satisfies a strong definition of convexity (with $\mathcal{C}_0^\infty$ test functions), which is equivalent to the gradients' jumps positivity (for P_1 FE), leads to an additional constraint on the limit function. The error of such a discretization of the constraint does not vanish with the space step. We justify it is localized where the additional constraint on the limit function is not satisfied. The condition for existence of such a counter-example requires information both from the mesh and its refinement.

In addition, such a P_1 discretization of strong convexity is not even consistent. The definition of consistency is very similar to the one of partial differential equations and it has been used here to discriminate likely discretizations and unlikely ones. But this does not guarantee good numerical results even when the discretization is consistant. We also prove the P_1 discretization of weak convexity to be consistent.

We also test the gradients' jumps in P_2 and various weak P_2 discretizations. Some of them are consistent and not others. We also discuss thoroughly the literature.

The author wants to thank G. Carlier for drawing his attention to the problem of subharmonicity approximation.

References

[1] Néstor E. Aguilera and Pedro Morin.
    Approximating optimization problems over convex functions.
    Numer. Math., 111(1):1–34, 2008.

[2] Néstor E. Aguilera and Pedro Morin.
    On convex functions and the finite element method.
    SIAM J. Numer. Anal., 47(4):3139–3157, 2009.

[3] Giuseppe Buttazzo, Vincenzo Ferone, and Bernhard Kawohl.
    Minimum problems over sets of concave functions and related questions.
    Math. Nachr., 173:71–89, 1995.

[4] Guillaume Carlier.
    On a theorem of Alexandrov.
    J. Nonlinear Convex Anal., 5(1):49–58, 2004.

[5] Guillaume Carlier, Myriam Comte, and Gabriel Peyré.
    Approximation of maximal Cheeger sets by projection.
    M2AN Math. Model. Numer. Anal., 43(1):139–150, 2009.

[6] Guillaume Carlier and Thomas Lachand-Robert.
    Regularity of solutions for some variational problems subject to a convexity constraint.
    Comm. Pure Appl. Math., 54(5):583–594, 2001.

[7] Guillaume Carlier, Thomas Lachand-Robert, and Bertrand Maury.
    H^1-projection into the set of convex functions: a saddle-point formulation.
    In *CEMRACS 1999 (Orsay)*, volume 10 of *ESAIM Proc.*, pages 277–289 (electronic).
    Soc. Math. Appl. Indust., Paris, 1999.

[8] Guillaume Carlier, Thomas Lachand-Robert, and Bertrand Maury.

A numerical approach to variational problems subject to convexity constraint.
Numer. Math., 88(2):299–318, 2001.

[9] Philippe Choné and Hervé V. J. Le Meur.
Non-convergence result for conformal approximation of variational problems subject to a convexity constraint.
Numer. Funct. Anal. Optim., 22(5-6):529–547, 2001.

[10] Ivar Ekeland and Santiago Moreno-Bromberg.
An algorithm for computing solutions of variational problems with global convexity constraints.
Numer. Math., 115(1):45–69, 2010.

[11] Herman H. Goldstine.
A history of the calculus of variations from the 17th through the 19th century, volume 5 of *Studies in the History of Mathematics and Physical Sciences*.
Springer-Verlag, New York, 1980.

[12] Thomas Lachand-Robert and Édouard Oudet.
Minimizing within convex bodies using a convex hull method.
SIAM J. Optim., 16(2):368–379 (electronic), 2005.

[13] Thomas Lachand-Robert and Mark A. Peletier.
An example of non-convex minimization and an application to Newton's problem of the body of least resistance.
Ann. Inst. H. Poincaré Anal. Non Linéaire, 18(2):179–198, 2001.

[14] Nim K. Leung and Robert J. Renka.
C^1 convexity-preserving interpolation of scattered data.
SIAM J. Sci. Comput., 20(5):1732–1752 (electronic), 1999.

[15] Pierre-Louis Lions.
Identification du cône dual des fonctions convexes et applications.
C. R. Acad. Sci. Paris Sér. I Math., 326(12):1385–1390, 1998.

[16] Bertrand Maury.
Version continue de l'algorithme d'Uzawa.
C. R. Math. Acad. Sci. Paris, 337(1):31–36, 2003.

[17] Isaac Newton.
Philosophiae Naturalis Principia Mathematica.
1686.

[18] Jean-Claude Rochet and Philippe Choné.
Ironing, sweeping and multidimensionnal screening.
Econometrica, 66:783–826, 1998.