Metric Based Upscaling for Partial Differential Equations with a Continuum of Scales

Thesis by

Lei Zhang

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2007

(Submitted May 16, 2007)

For my parents.

Acknowledgments

I am deeply grateful to my advisor, Professor Houman Owhadi, for his invaluable guidance. His brilliance, insight and enthusiasm to the unknown world has been a constant inspiration for me. I also wish to thank him for his kindness and encouragements.

I would like to thank Professor Jerrold E. Marsden, Professor Thomas Y. Hou, and Professor Emmanuel J. Candes for taking their precious time to serve on my thesis committee.

Studying at Caltech is a greatly enriching experience, I would like to thank all the professors who taught me at Caltech, especially Jerrold Marsden, Kip Thorne, and Houman.

I would like to thank Professor Niles Pierce for his help and guidance when I was in his group. I would like to thank Wuan, Nawaf, Lexing, Theo, Xinwei, Jay, Ruo Li, Roger, Stephan, Zhiyi, Lei Zhen et al. for many stimulating and interesting conversations academically or non-academically. Many thanks to Sheila, Chad, and Sydney for their remarkable work.

Thanks to my parents, although oceans apart, your constant support is my motivation to proceed. Finally, thanks to my wife, Hong Wang, who has been accompanying me and supporting me in these years, experiencing everything no matter good or bad together and making our life more colorful.

Abstract

Numerical upscaling of problems with multiple scale structures have attracted increasing attention in recent years. In particular, problems with non-separable scales pose a great challenge to mathematical analysis and simulation. Most existing methods are either based on the assumption of scale separation or heuristic arguments.

In this thesis, we present rigorous results on homogenization of partial differential equations with L^∞ coefficients which allow for a continuum of spatial and temporal scales. We propose a new type of compensation phenomena for elliptic, parabolic, and hyperbolic equations. The main idea is the use of the so-called "harmonic coordinates" ("caloric coordinates" in the parabolic case). Under these coordinates, the solutions of these differential equations have one more degree of differentiability. It has been deduced from this compensation phenomenon that numerical homogenization methods formulated as oscillating finite elements can converge in the presence of a continuum of scales, if one uses global caloric coordinates to obtain the test functions instead of using solutions of a local cell problem.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

In this work, we focus on upscaling problems with non-separable scales, which is both important for applications and far from understood from a mathematical point of view. The upscaling method lies on a new type of compensation phenomena for partial differential equations with L^∞ coefficients [108, 106, 107]. We could design numerical homogenization methods through the use of a coordinate transformation which brings in an extra degree of regularization. The results are presented in a rigorous mathematical framework.

1.1 Overview

Problems with many scales are ubiquitous in nature. To make them more accessible to analysis, it is often preferable to make the assumption of scale separation (a small parameter $\varepsilon \to 0$) and periodicity (or quasi-periodicity, ergodicity). Essential progress has been achieved in the study of such problems, to give a few examples out of a vast literature, let us refer to [33] (Bensoussan, Lions, and Papanicolaou) and [81] (Jikov, Kozlov, and Oleinik). However, although an infinite perfectly periodic crystal lattice is elegant and more amenable to mathematical analysis, nature is often more nasty and disordered. Even the purest material cannot escape the effect of defects, fractures, and phase boundaries [117]. High Reynolds turbulence flow, plasma instability, and earthquakes give us some outstanding examples where strong scale coupling is present and scale separation no longer works.

Among all the multi-scale problems, the following divergence form elliptic equation

with highly oscillatory coefficients $a(x)$ is perhaps the most intensively studied one. It is also one of the main objects of this thesis.

$$\begin{cases} -\nabla a(x)\nabla u(x) = g & \text{in} \quad \Omega \\ u = 0 & \text{on} \quad \partial\Omega \end{cases} \tag{1.1.1}$$

Many methods for other multi-scale problems can find their roots in the elliptic counterpart. Therefore, in the following, we will use (1.1.1) as a benchmark problem.

If a solution at the fine scale is sought, we often use multilevel or multigrid type methods [1] which go from fine to coarse scale and back from coarse to fine iteratively. However, optimal convergence for multigrid method cannot be easily achieved if the coefficients are non-smooth or highly oscillatory. Diffusion problems with small scale oscillations can be handled by the so-called "matrix dependent" prolongation/restriction operators [132]. In recent years, robust and efficient multilevel methods such as algebraic multigrid method (AMG) [113] and smoothed aggregation (SA) [126] were proposed for general elliptic problems by mostly heuristic strategies. Domain decomposition methods are used to provide good preconditioners and facilitate parallel processing.

However, even with modern state-of-the-art supercomputers and algorithms, a direct simulation of the highly heterogeneous media, which involves a wide range of spatial scales and time scales, is still difficult, if not impossible. That is why we will pursue multi-scale methods to solve (1.1.1) on the coarse scale. More precisely, we want to know how to transfer information from fine scales to coarse scales and how to use the information obtained to solve the problem on the coarse scale with much fewer degrees of freedom. We often refer this procedure as *numerical homogenization* or *numerical upscaling*.

Homogenization theory (Γ-, G- and H- convergence) answers the question with the assumption of scale separation. The idea is to average heterogeneous media on the fine scale in order to derive effective properties. The most general theory in homogenization is that of H-convergence which was introduced by Spagnolo [121] and further generalized by Tartar and Murat[125, 98]. With the powerful oscillating test functions method or compen-

[1] See [39] on systematic upscaling, which is a multi-scale computational methodology developed from multigrid method.

sated compactness method, the H-convergence result for elliptic equations can be proved independent from ergodicity or scale separation assumptions, although the homogenized problem is not known a priori unless the media is periodic.

The Multi-scale Finite Element Method (MsFEM) of Hou and Wu [78] has been a large source of inspiration in numerical applications (particularly for reservoir modeling in geophysics), we refer to [130], [88], [1], and [131] for recent developments. It leads to a coarse scale operator while keeping the fine scale structures of the solutions. The construction of the base functions is decoupled from element to element, leading to a scheme adapted to parallel computers. A proof of the convergence of the method is given in periodic settings when the size of the heterogeneities is smaller than the grid size and an "oversampling technique" is proposed to remove the so-called "cell resonance" error [79] (when the size of the heterogeneities is comparable to the grid size).

In fact, The issue of numerical homogenization of partial differential equations with heterogeneous coefficients has received a great deal of attention and many methods have been proposed. Let us mention a few of them:

- Multi-scale finite element methods [56], [103], [78], [74], [65], [70], [8]
- Multi-scale finite volume methods [89]
- Heterogeneous multi-scale methods [128]
- Wavelet based homogenization [68], [55], [51], [37], [18], [41]
- Residual free bubbles methods [42]
- Discontinuous enrichment methods [61], [60]
- Partition of unity methods [66]
- Energy minimizing multigrid methods [129].

Most multi-scale methods are based on solving local cell problems. Some approaches use the cell problem to calculate effective media properties, then solve an effective equation on the coarse scale. A detailed review of this kind of upscaling methods can be found in

[62]. Other methods, like MsFEM, incorporate the fine scale features of the problems into basis elements. The coupling of small scales with coarse scales is then performed through a numerical formulation of the global problem using these multi-scale basis. These methods can often be justified in dimension one, in the case of periodic or ergodic media with scale separation, or in the case of partial differential equations with sufficiently smooth coefficients. However, the separation of scales is not always possible. For example, in the subsurface modeling, the reservoirs often contain rocks of very different types, and the permeability usually covers several orders of magnitude, from impermeable barriers to highly permeable fast channels. It is difficult to make methods based on local cell problems succeed in solving problems with non-separable scales.

Another perspective to approach multi-scale problems is the compression of operators, for example, the homogenized equation can be seen as a compressed version of the original equation in the case of scale separation. This question has received an answer within the context of the fast multiplication of vectors with fully populated special matrices arising in various applications [64, 50]. Let us recall the fast multi-pole method and the hierarchical multi-pole method designed by L. Greengard and V. Rokhlin [71], which are based on the singular value decomposition of Green's function. Wavelet based methods for the reduction of integral and differential operators have been designed by G. Beylkin, R. Coifman and V. Rokhlin [9, 36, 35]. The concept of Hierarchical matrices has been developed by W. Hackbusch et al. [73, 31, 30, 29, 27, 28] and is based on approximating a matrix to a degenerate sum using a hierarchical partitioning procedure. For divergence form elliptic equations with L^∞ coefficients, it has been proven that the inverse of the stiffness matrix can be represented by H-Matrix, provided the discretization is stable. It has been shown that the complexity for solving (1.1.1) is $O\big(N(\ln N)^{n+3}\big)$ operations (N is the degree of freedom of the dicretization, n is the dimension of the space).

Composition Rule Allaire and Brizzi [8] have introduced the composition rule in the multi-scale finite element formulation, and have observed that a multi-scale finite element method with higher order Lagrange polynomials has a higher accuracy. In fact, I. Babuška et al. introduced the so called "change of variable" technique in the general setting of par-

tition of unity method (PUM) with p-version of finite elements. Through the change of variable, the original problem is mapped into a problem which can be better approximated. In [20], special class of second order elliptic problems with essentially one-dimensional rough coefficients $a(x,y) = a(x)$ was considered, using change of variables, the divergence form equation can be converted to a nondivergence form equation, and by Bernstein theorem [34] the approximation property of mapped polynomials can be obtained. For elliptic problems in 2-d with corners or interfaces, conformal mapping was used to map the rough solution to a smoother function in [104, 105],

Global Information Numerous efforts have been made to deal with problems with non-separable scales or global features. There exist some approaches which incorporate large scale effects in the setting of upscaling. For example, iterating between coarse and fine scales [48, 49] or solving a minimization problem [75, 101]. The idea to use global fine scale information to homogenize transport equations for reservoir modeling in geophysics is currently implemented in the industry and has been shown to be more accurate than local methods ([131] and [130]). It is applied in practice because the porosity of the medium is time independent and one can solve an elliptic equation only at $t = 0$ to upscale the transport equations. Some recent results using global information by Efendiev et al. [2, 80] are formulated in a partition of unity finite element framework.

The multi-scale elliptic problems are in some sense the "easiest" multi-scale problems. The methods mentioned here can be seen as building blocks of a much larger quest aimed at capturing high dimensional problems with a few coarse parameters [119], [14], [99], [72]. Many extensions to diverse physical situations have to be developed and justified. Paraphrasing the outcome of a recent DOE workshop [54], we may understand the physics of multi-scale structures at each individual scale nevertheless *"without the capability to 'bridge the scales', a significant number of important scientific and engineering problems will remain out of reach"*.

1.2 Summary of the Thesis

The thesis is divided into three parts. Chapter 1 is the introduction. Chapter 2, 3 and 4 discuss elliptic equations, parabolic equations and hyperbolic equations respectively. In each chapter, we will first introduce the corresponding compensation result and formulate the numerical homogenization method, then give the detailed proof of all the results in the chapter, and finally show the results of numerical experiments. In Chapter 5, we will make some concluding remarks.

To make the thesis self-contained, in the Appendices, we will state some results on regularity of partial differential equations with L^∞ coefficients, a-harmonic mapping, and C^1 finite element using B-splines which are needed in the main body.

1.2.1 Metric Based Upscaling for Elliptic Equation

In Chapter 2 we consider the numerical homogenization of divergence form elliptic equations:

$$\begin{cases} -\nabla a(x)\nabla u(x) = g & \text{in} \quad \Omega \\ u = 0 & \text{on} \quad \partial\Omega \end{cases} \tag{1.2.1}$$

where $a(x)$ is a symmetric $n \times n$ matrix with entries in $L^\infty(\Omega)$. We assume $a(x)$ is uniformly bounded and coercive. $p > 2$ is some constant depending on a and Ω, g is a function in $L^p(\Omega)$.

The harmonic coordinates $F(x) = \big(F_1(x), \ldots, F_n(x)\big)$ associated to (1.2.1) satisfy the following equations,

$$\begin{cases} \operatorname{div} a\nabla F_i = 0 & \text{in} \quad \Omega \\ F_i(x) = x_i & \text{on} \quad \partial\Omega. \end{cases} \tag{1.2.2}$$

It can be shown that F is an automorphism over Ω [6], we refer to Appendix B and references therein. Recall that the natural distance associated to the Laplace operator on a fractal space is called resistance metric [83, 122, 26]. It is thus natural to find that a similar (not equivalent) notion of distance allows the numerical homogenization of PDEs with arbitrary coefficients. More precisely the analogue of the resistance metric here is the

harmonic mappings F. The analysis of these mappings allows bypassing of boundary layer effects in homogenization in periodic media [8].

We discover the following compensation phenomena: Though in Euclidean coordinates solutions u of (1.2.1) are only $W^{1,p}$ (by Meyers Theorem) or Hölder continuous (by De Giorgi-Nash-Moser theory), with respect to the harmonic coordinates they are $W^{2,p}$ ($C^{1,\alpha}$ in 2d). Namely, they have one more degree of differentiability. Indeed, $u \circ F^{-1}$ satisfies a non-divergence elliptic equation which is known to have $W^{2,p}$ estimate under a Cordes type condition [93].

More precisely, write $\sigma := {}^t\nabla F a \nabla F$, then there exists $p > 2$, if σ satisfies the following Cordes type condition

$$\beta_\sigma := \operatorname{esssup}_{(x,t)\in\Omega_T}\left(n - \frac{\left(\operatorname{Trace}[\sigma]\right)^2}{\operatorname{Trace}[{}^t\sigma\sigma]}\right) < 1 \tag{1.2.3}$$

and $\left\| \operatorname{Trace}(\sigma)\right)^{\frac{n}{2p}-1}\|_{L^\infty(\Omega)} < \infty$, we have the following result [2]:

$$\|u \circ F^{-1}\|_{W^{2,p}(\Omega)} \leq C\|g\|_{L^p(\Omega)}, \tag{1.2.4}$$

and in 2d, there exists $\alpha > 0$, the derivative of u with respect to F is Hölder continuous

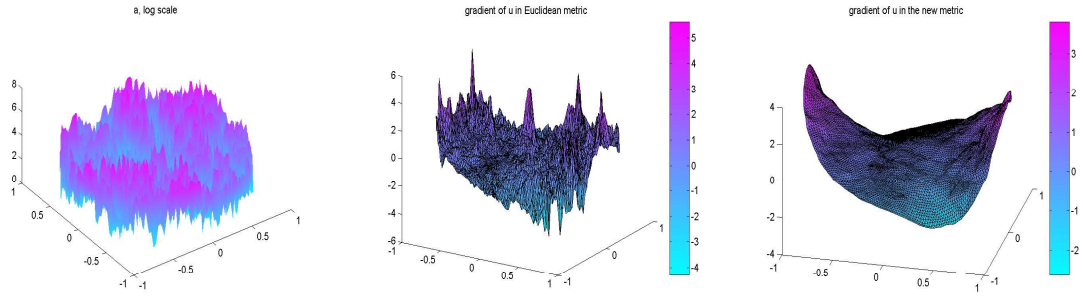
$$\|(\nabla F)^{-1}\nabla u\|_{C^\alpha(\Omega)} \leq C\|g\|_{L^p(\Omega)}. \tag{1.2.5}$$

(1.2.4) also holds for $p = 2$.

This phenomenon can be observed numerically. In figure 1.1(a) a is given by a product of random functions oscillating over a continuum of scales. The entries of the matrix ∇F and ∇u are in L^p, while $(\nabla F)^{-1}\nabla u$ is Hölder continuous.

It can be deduced from this compensation phenomena that numerical homogenization methods based on oscillating finite elements can converge in the presence of a continuum of scales, if one uses global harmonic coordinates to obtain the test functions instead of solving a local cell problem. Compared with methods which perturb the test functions with

[2] The results presented in the thesis are slightly different from but essentially the same as those in [108].



(a) a ***in log scale*** **(b)** ∇u **(c)** $(\nabla F)^{-1}\nabla u$

Figure 1.1: Gradient of u in Euclidean coordinates and harmonic coordinates

the solution of a local cell problem, the global change of coordinates allows avoidance of the cell resonance problem and the means to obtain a scheme converging uniformly in h.

We can roughly explain the numerical homogenization method using the following analogue: When we solve the equation $Ax = y$, we do not compute A^{-1} directly. Instead, we solve the equation with n different right hand sides $Ax_i = y_i$. With the information from x_i, we can construct A_c with much fewer degrees of freedom and compute an approximate solution using A_c by a prescribed accuracy.

For example, we can use the composition rule to construct the numerical homogenization method, write V^h the finite dimensional subspace of $H_0^1(\Omega)$,

$$V^h := \{\varphi \circ F \,:\, \varphi \in X^h\} \tag{1.2.6}$$

where X^h is a usual C^0 or C^1 finite element space.

We have the following error estimate:

$$\|u - u_h\|_{H_0^1(\Omega)} \leq Ch\|g\|_{L^2(\Omega)}. \tag{1.2.7}$$

Once one understand that the key idea for the homogenization of (1.2.1) lies in its higher regularity properties with respect to harmonic coordinates, one can homogenize (1.2.1) through a different formulation.

Instead of using the finite element $\psi = \varphi \circ F$, which has a deformed support, we can

construct a finite element space ξ with regular support on the original quasi-uniform mesh, but the price to pay is the discontinuity of the elements which results in a nonconforming finite element method. We will prove the convergence of this method. The approximation error on the coarse mesh would depend on the aspect ratio of the triangles of the coarse mesh in the metric induced by F, and the approximation error of F by a piecewise linear map.

Both ψ and ξ contain the whole fine scale structure of F. In fact, it is possible to compress the elliptic operator by a effective operator associated to the coarse mesh. In fact, the resulting numerical method can be formulated as a Petrov-Galerkin method, with test functions piecewise linear on the coarse mesh and trial functions the nonconforming finite element ξ. To define the effective operator, we only need compressed information, the bulk quantities $\langle a\nabla F\rangle$ and the non-averaged quantities $F(b)-F(a)$, where a,b are nodes of the coarse mesh triangles K.

The elliptic operator appearing in (2.0.1) can be seen as the generator of a stochastic differential equation. This stochastic differential equation can reflect the transport process of a pollutant in a highly heterogeneous medium such as soil. The following operator $\Delta - \nabla V \nabla$ whose numerical homogenization is similar to that of (2.0.1) can represent a physical system evolving in a highly irregular energy landscape V. The simple fact that this evolution taking place in a continuous domain can be captured by a Markov chain evolving on a graph is far from being obvious [123]. We propose to accurately simulate a Markov chain living on a fine graph by an 'up-scaled' Markov chain living on a coarse graph using the information from F. The main question is how to choose the jump rate γ_{ij} of the random walk between the nodes of the coarse graph. The answer to that question is conveyed by a multi-scale finite volume method.

We have seen that if σ is stable then $u \circ F^{-1}$ belongs to $W^{2,p}(\Omega)$ with some $p > 2$. It is thus natural to expect a better accuracy by C^1 finite elements (described in Appendix C and references therein) rather than piecewise linear elements. This increase of accuracy has already been observed by Allaire and Brizzi in [8] when F is approximated as the solution of a local cell problem. When the harmonic coordinates are computed globally, we observe a sharp increase of the accuracy for the finite elements $\psi = \varphi \circ F$ by using splines as the

elements φ.

1.2.2 Numerical Homogenization for Parabolic Equations with a Continuum of Space-Time Scales

In Chapter 3 we address the issue of numerical homogenization of linear divergence form parabolic equations with a continuum of space-time scales, which describe many important problems such as nuclear waste storage in deep geological formations [12]. Consider the following equation:

$$\begin{cases} \partial_t u = \nabla\big(a(x,t)\nabla u(x,t)\big) + g \quad \text{in} \quad \Omega \times (0,T) \\ u(x,t) = 0 \quad \text{for} \quad (x,t) \in \big(\partial\Omega \times (0,T)\big) \cup \big(\Omega \times \{t=0\}\big) \end{cases} \tag{1.2.8}$$

where Ω is a bounded and convex domain of class C^2 of $\mathbb{R}^n$, $T > 0$ and $\Omega_T := \Omega \times (0,T)$. g is a function in $L^2(\Omega_T)$, and $a(x,t)$ is a symmetric positive definite matrix with entries in $L^\infty(\Omega_T)$ and uniformly elliptic on the closure of Ω_T.

Under the assumption of scale separation and time independent coefficients a, numerical homogenization methods have been proposed and analyzed in [3, 45] in the framework of MsFEM or HMM method.

If the medium is time independent or the dependence on time is smooth, it is sufficient to solve an associated elliptic equation n times to use the time independent harmonic coordinates F. Otherwise, we need to solve for the time dependent caloric coordinates $F := (F_1, \ldots, F_n)$ satisfying

$$\begin{cases} \partial_t F_i = \nabla\big(a(x,t)\nabla F_i(x,t)\big) \quad \text{in} \quad \Omega_T \\ F_i(x,t) = x_i \quad \text{for} \quad (x,t) \in \big(\partial\Omega \times (0,T)\big) \\ \nabla\big(a(x,0)\nabla F_i(x,0)\big) = 0 \quad \text{in} \quad \Omega \times \{t=0\} \end{cases} \tag{1.2.9}$$

which is essentially different from the elliptic case.

The compensation phenomena can now be read off from the following estimate, under

a parabolic Cordes condition we have:

$$\|u\circ F^{-1}\|_{L^2(0,T,W^{2,2}(\Omega))}+\|\partial_t(u\circ F^{-1})\|_{L^2(\Omega_T)}\leq C\|g\|_{L^2(\Omega_T)} \tag{1.2.10}$$

which implies although $u\in L^2(0,T,H_0^1(\Omega))$ and $\partial_t u\in L^2(0,T,H^{-1}(\Omega))$, $u\circ F^{-1}\in L^2(0,T,W^{2,2}(\Omega))$ and $\partial_t(u\circ F^{-1})\in L^2(\Omega_T)$.

For $t\in[0,T]$, let us define the time-space finite element space

$$V^h(t):=\left\{\varphi\circ F(x,t)\,:\,\varphi\in X^h\right\}. \tag{1.2.11}$$

Define Y_T^h the subspace of $L^2\big(0,T;H_0^1(\Omega)\big)$ as

$$Y_T^h:=\{v\in L^2\big(0,T;H_0^1(\Omega)\big)\,:\,v(x,t)\in V^h(t)\}. \tag{1.2.12}$$

Write u_h the solution in Y_T^h of the following system of ordinary differential equations:

$$\begin{cases}(\psi,\partial_t u_h)_{L^2(\Omega)}+a[\psi,u_h](t)=(\psi,g)_{L^2(\Omega)} & \text{for all } t\in(0,T) \text{ and } \psi\in V^h(t)\\ u_h(x,0)=0.\end{cases} \tag{1.2.13}$$

We have the following error estimate,

$$\big\|(u-u_h)(.,T)\big\|_{L^2(\Omega)}+\big\|u-u_h\big\|_{L^2(0,T;H_0^1(\Omega))}\leq Ch\|g\|_{L^2(\Omega_T)}. \tag{1.2.14}$$

Furthermore, we can introduce the time discrete numerical homogenization method. Suppose $(t_n=n\frac{T}{M})_{0\leq n\leq M}$ is a discretization of $[0,T]$ with $M\in\mathbb{N}$. Let (φ_i) be a basis of X^h. Write Z_T^h as the subspace of Y_T^h,

$$Z_T^h:=\{w\in Y_T^h:w(x,t)=\sum_i c_i(t)\varphi_i\big(F(x,t)\big),\ c_i(t) \text{ are constants on } (t_n,t_{n+1}]\} \tag{1.2.15}$$

U_T^h is the subspace of Y_T^h defined as,

$$U_T^h = \{\psi \in Y_T^h : \psi(x,t) = \sum_i d_i \varphi_i\big(F(x,t)\big),\ d_i \text{ are constants (on } [0,T]).\} \tag{1.2.16}$$

define $w_n \in U_T^h$ by

$$w_n(x,t) := \sum_i c_i(t_n)\varphi_i(F(x,t)). \tag{1.2.17}$$

Let v be the solution in Z_T^h which satisfies the following implicit weak formulation, (suppose that $v(x,0) \equiv 0$): for $n \in \{0,\dots,M-1\}$ and $\forall \psi \in U_T^h$,

$$\begin{aligned}\big(\psi(t_{n+1}), v_{n+1}(t_{n+1})\big)_{L^2(\Omega)} =& \big(\psi(t_n), v_n(t_n)\big)_{L^2(\Omega)} + \int_{t_n}^{t_{n+1}} \Big(\big(\partial_t \psi(t), v_{n+1}(t)\big)_{L^2(\Omega)} \\ & - a\big[\psi, v_{n+1}\big](t) + \big(\psi(t), g(t)\big)_{L^2(\Omega)}\Big)\, dt.\end{aligned} \tag{1.2.18}$$

The following theorem gives an error bound for the time discretization scheme

$$\big\|(u_h - v)(T)\big\|_{L^2(\Omega)} + \|u_h - v\|_{L^2(0,T,H_0^1(\Omega))} \le C\frac{\Delta t}{h}\|g\|_{L^2(\Omega_T)} \tag{1.2.19}$$

When a is independent of t, the error bound can be improved to $C\Delta t$.

1.2.3 Numerical Homogenization for the Acoustic Wave Equation with a Continuum of Space Scales

Based on the upscaling techniques for elliptic equations and extended to parabolic equations with a continuum of scales, we can numerically homogenize acoustic wave equations with a continuum of scales.

Waves in heterogeneous media is a field of great mathematical interest and applicable to many real problems in geophysics, seismology, and electromagnetics [124, 25, 127, 19]. We refer to [124] for a review of the acoustic wave equation in relation to seismic imaging. For an extensive work on the wave equation in complex or random media we refer to [111], [24], [23], [87], [110], [109], [47], [86], [116], [84], and [85].

We consider the homogenization of wave equation in heterogeneous media where the bulk modulus $K(x)$ and the density of the medium $\rho(x)$ are only bounded,

$$K(x)^{-1}\partial_{tt}u = \nabla\rho(x)^{-1}\nabla u + g(x,t). \tag{1.2.20}$$

For example, in geophysical and seismic prospecting, K stands for the bulk modulus, ρ the density and u the unknown pressure. The velocity c and acoustic impedance σ are given by

$$c = \sqrt{K/\rho} \quad \text{and} \quad \sigma = \sqrt{K\rho}. \tag{1.2.21}$$

The main difference with parabolic equations lies in the fact that with hyperbolic equations, energy is conserved and after homogenization there is no hope of recovering the energy (or information) lying in the highest frequencies. However when the medium is highly heterogeneous the eigenfunctions associated to the highest frequencies are localized, thus energy is mainly transported by the lowest frequencies. That is why, when one is only interested in the large scale transport of energy, it is natural to approximate the solutions of (1.2.20) by the solutions of a homogenized operator. For localization of waves in heterogeneous media, we refer to [118, 11, 84, 85, 86].

Different numerical schemes have been developed to solve that equation (with different assumption on the regularity of the coefficients), we refer to [25], [19], [127], and [22] for an incomplete list.

We show that under a Cordes type condition, as well as some mild assumptions for forcing term $g(x,t)$ and initial data $u(x,0)$ (for example, assume that $\partial_t g \in L^2(\Omega_T)$, $g \in L^\infty(0,T,L^2(\Omega))$, $\partial_t u(x,0) \in H^1(\Omega)$, and $\nabla a(x)\nabla u(x,0) \in L^2(\Omega)$), the second order derivatives of the solution with respect to harmonic coordinates F are in L^2 (instead of H^{-1} with respect to Euclidean coordinates) and the solution itself is in $L^\infty(0,T,H^2(\Omega))$ (instead of $L^\infty(0,T,H^1(\Omega))$ with respect to Euclidean coordinates). Therefore, using the composition rule, we can construct numerical homogenization methods to solve the acoustic wave equation.

Chapter 2

Metric Based Upscaling for Elliptic Equations

Suppose $\Omega \subset \mathbb{R}^n$ to be a bounded and convex domain of class C^2. We consider the following benchmark PDE

$$\begin{cases} -\operatorname{div}\big(a(x)\nabla u(x)\big) = g \quad \text{in} \quad \Omega \\ u = 0 \quad \text{in} \quad \partial\Omega \end{cases} \tag{2.0.1}$$

where g is a function in $L^\infty(\Omega)$ (depending on the context, we can make different assumption $g \in L^p(\Omega)$ with $p \geq 2$). $a(x)$ is symmetric, uniformly elliptic with entries in $L^\infty(\Omega)$.

2.1 Compensation Phenomena

We introduce the so called a-harmonic coordinates associated to (2.0.1), i.e., the weak solution of the following boundary value problem

$$\begin{cases} \operatorname{div} a\nabla F = 0 \quad \text{in} \quad \Omega \\ F(x) = x \quad \text{on} \quad \partial\Omega. \end{cases} \tag{2.1.1}$$

By (2.1.1) we mean that F is a n-dimensional vector field $F(x) = \big(F_1(x), \ldots, F_n(x)\big)$

such that each of its entries satisfies

$$\begin{cases} \operatorname{div} a\nabla F_i = 0 \quad \text{in} \quad \Omega \\ F_i(x) = x_i \quad \text{on} \quad \partial\Omega. \end{cases} \tag{2.1.2}$$

Define σ by,

$$\sigma := {}^t\nabla F a \nabla F. \tag{2.1.3}$$

Define the anisotropic distortion of σ by

$$\mu_\sigma := \operatorname{esssup}_{x\in\Omega}\Big(\frac{\lambda_{\max}\big(\sigma(x)\big)}{\lambda_{\min}\big(\sigma(x)\big)}\Big). \tag{2.1.4}$$

where $\lambda_{\max}(M(x))$ ($\lambda_{\min}(M(x))$) denote the maximal (minimal) eigenvalue of matrix $M(x)$, we also use the notation $\lambda_{\max}(M) := \operatorname{esssup}_{x\in\Omega}\sup_{|\xi|=1} {}^t\xi a\xi$ and $\lambda_{\min}(M) := \operatorname{essinf}_{x\in\Omega}\inf_{|\xi|=1} {}^t\xi a\xi$ for the supremum of $\lambda_{\max}(M(x))$ and infimum of $\lambda_{\min}(M(x))$ over Ω.

In dimension $n = 2$, we say that σ is *stable* if and only if $\mu_\sigma < \infty$ and $\big(\operatorname{Trace}(\sigma)\big)^{-1} \in L^\infty(\Omega)$. According to [10], in dimension two if a is smooth then σ is stable. According to [6], F is always an homeomorphism in dimension two even with $a_{i,j} \in L^\infty(\Omega)$. Also see Appendix B and references therein.

We will use the notation $\nabla_F u := (\nabla F)^{-1}\nabla u$. In dimension two, it is known ([10], [15], [6]) that the determinant of ∇F is strictly positive almost everywhere and the object $\nabla_F u$ is well defined. In dimension three and higher $\nabla_F u$ is well defined when σ is stable and F is an automorphism.

Theorem 2.1.1. *Assume that σ is stable and $n = 2$, then there exists constants $p > 2$, $\alpha > 0$, and $C > 0$ such that $(\nabla F)^{-1}\nabla u \in C^\alpha(\Omega)$ and*

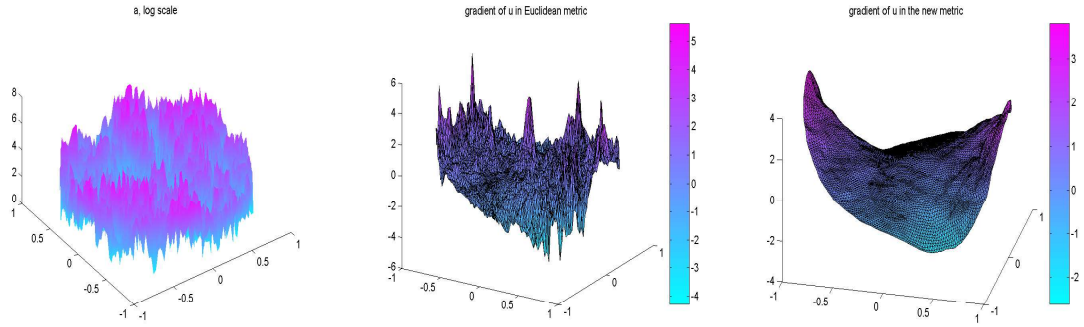
$$\big\|(\nabla F)^{-1}\nabla u\big\|_{C^\alpha(\Omega)} \le C\|g\|_{L^p(\Omega)}. \tag{2.1.5}$$

The constant α depends on $\Omega, \lambda_{\max}(a)/\lambda_{\min}(a)$, and μ_σ. The constant C depends on the constants above and $\big\|\operatorname{Trace}(\sigma)\big)^{\frac{n}{2p}-1}\big\|_{L^\infty(\Omega)}$.

Remark.

If one considers a sequence $a_\varepsilon = a(x/\varepsilon)$ *such that* μ_{σ_ε} *and* $\left\| \mathrm{Trace}(\sigma_\varepsilon)\right)^{\frac{n}{2p}-1}\Big\|_{L^\infty(\Omega)}$ *are uniformly bounded away from* ∞, $\lambda_{\min}(a_\varepsilon)$ *and* $\lambda_{\max}(a_\varepsilon)$ *are uniformly bounded away from* 0 *and* ∞, *then (2.1.5) is uniformly true.*

This compensation phenomenon can be observed numerically. In figures 2.1(a) and 2.2(a) a is given by a product of random functions oscillating over a continuum of scales. The entries of the matrix ∇F are in L^p by Meyers Theorem (figure 2.2(b)), the entries of the gradient of u in the Euclidean metric are in L^p (figures 2.1(b) and 2.2(c)), yet $(\nabla F)^{-1}\nabla u$ is Hölder continuous (figures 2.1(c) and 2.2(d)).



(a) a ***in log scale*** **(b)** ∇u **(c)** $(\nabla F)^{-1}\nabla u$

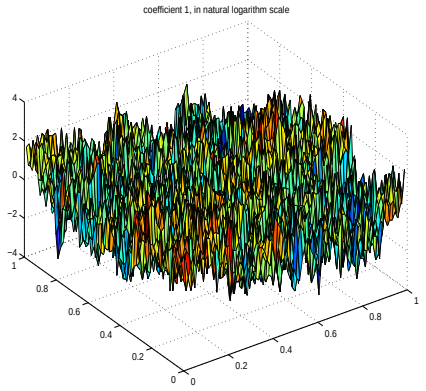
Figure 2.1: Change of metric on the disk.

Moreover, we can introduce the compensation phenomenon in dimension $n \geq 3$. As in Appendix A.1.2, the Cordes parameter β_σ associated to σ is defined by

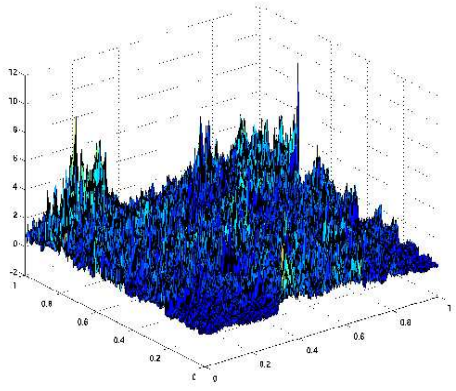
$$\begin{aligned}\beta_\sigma :&= \mathrm{esssup}_{x\in\Omega}\Big(n - \frac{\big(\mathrm{Trace}[\sigma(x)]\big)^2}{\mathrm{Trace}[{}^t\sigma(x)\sigma(x)]}\Big)\\ &= \mathrm{esssup}_{x\in\Omega}\Big(n - \frac{\big(\sum_{i=1}^n \lambda_i(\sigma(x))\big)^2}{\sum_{i=1}^n \lambda_i(\sigma(x))^2}\Big).\end{aligned} \tag{2.1.6}$$

where $\lambda_i(M)$ denotes the ith eigenvalue of M.

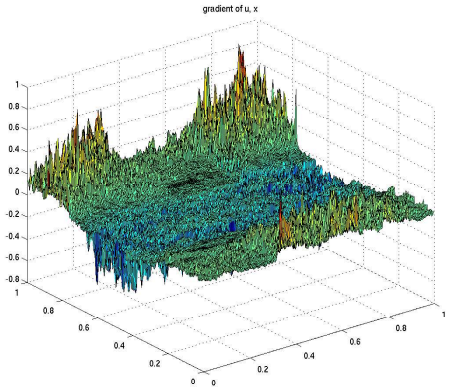
In dimension $n \geq 3$, we say that σ is *stable* if and only if, $\beta_\sigma < 1$ and exists $p \geq 2$ such



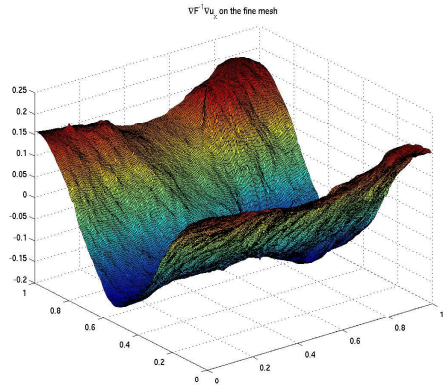
(a) a ***in log scale***



(b) ***one of the entries*** ∇F



(c) ∇u



(d) $(\nabla F)^{-1}\nabla u$

Figure 2.2: Change of metric on the torus

that $\| \big(\mathrm{Trace}(\sigma)\big)^{\frac{n}{2p}-1} \|_{L^\infty(\Omega)} < \infty$. In fact, in dimension 2, we have

$$\frac{1}{1-\beta_\sigma} \leq \frac{1}{2}(\mu_\sigma + \frac{1}{\mu_\sigma}) \tag{2.1.7}$$

therefore $\mu_\sigma < \infty \Rightarrow \beta_\sigma < 1$.

Remark. *According to [10] and [44] in dimension three and higher σ can be unstable even if a is smooth. We refer to figure 2.6 for an explicit example.*

Let us write

$$\|v\|_{W_0^{2,p}(\Omega)} := \Big(\int_\Omega \big(\sum_{i,j=1}^n |\partial_i \partial_j v|^2 \big)^{\frac{p}{2}} dx \Big)^{\frac{1}{p}}. \tag{2.1.8}$$

Theorem 2.1.2. *For $n \geq 3$, assume that σ is stable and F is an automorphism on Ω, then there exist constants $p > 2$ and $C > 0$ such that $u \circ F^{-1} \in W_0^{2,p}(\Omega)$ and*

$$\big\| u \circ F^{-1} \big\|_{W_0^{2,p}(\Omega)} \leq C \|g\|_{L^p(\Omega)}. \tag{2.1.9}$$

The constant p depends on n, Ω, $\lambda_{\max}(a)$, $\lambda_{\min}(a)$, and β_σ. The constant C depends on the constants above and $\| \big(\mathrm{Trace}(\sigma)\big)^{\frac{n}{2p}-1} \|_{L^\infty(\Omega)}$.

In the following theorem we do not need to assume Ω to be convex.

Theorem 2.1.3. *Assume $n \geq 2$ and $(\mathrm{Trace}(\sigma))^{-1} \in L^\infty(\Omega)$. Let $p > 2$. There exist a constant $C^* = C^*(n, \partial\Omega) > 0$ such that if $\beta_\sigma < C^*$ then there exists a real number $\gamma > 0$ depending only on n, Ω, and p such that*

$$\big\| (\nabla F)^{-1} \nabla u \big\|^2_{C^\gamma(\Omega)} \leq C \|g\|^2_{L^p(\Omega)}. \tag{2.1.10}$$

The constant C in (2.1.10) depends on n, γ, Ω, C^*, $\lambda_{\min}(a)$, $\lambda_{\max}(a)$, $\|a\|_{L^\infty(\Omega)}$, μ_σ and $\| \big(\mathrm{Trace}(\sigma)\big)^{\frac{n}{2p}-1} \|_{L^\infty(\Omega)}$

2.2 Dimensionality Reduction

2.2.1 Finite Element Using Composition Rule

According to Theorems 2.1.1 and 2.1.2, whatever the choice of g at small scales, solutions to (2.0.1) live in the neighborhood of a functional space correlated to F of dimension n. We will propose a rigorous justification of a variation of the multi-scale finite element method introduced by Hou and Wu [78] in the refined form by Allaire and Brizzi [8] in situations where the medium is not assumed to be periodic or ergodic (these methods are already rigorously justified when the medium is periodic [78], [8]).

Let $\mathscr{T}_h$ be a conformal simplicial coarse mesh on Ω composed of n-simplices (triangles in dimension two and tetrahedra in dimension three). Here h is the resolution of the mesh defined as the maximal length of the edges of the tessellation. By 'coarse' mesh we assume h is much greater than the scale of oscillations of the problem. By 'conformal' mesh, we mean: call $\gamma(\mathscr{T}_h)$ the maximum ratio of the n-simplices K over $\mathscr{T}_h$ of the ratio between the radius of the smallest ball containing K and the largest ball inscribed in K. Assume $\gamma(\mathscr{T}_h)$ to be uniformly bounded in h.

Write the coarse mesh finite element space $X^h \subset H_0^1(\Omega)$ the set of piecewise linear functions on the coarse mesh vanishing at the boundary of the tessellation. $\mathscr{N}_h$ is the set of interior nodes of the tessellation and φ_i ($i \in \mathscr{N}_h$) is the usual nodal basis function of X^h satisfying

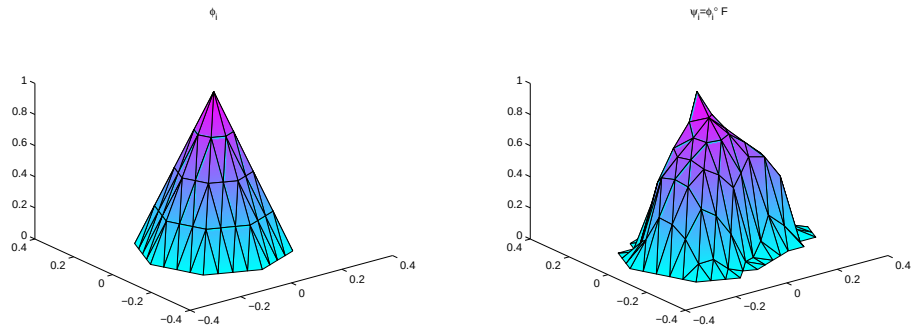
$$\varphi_i(y_j) = \delta_{ij}. \tag{2.2.1}$$

The finite elements $(\psi_i)_{i\in\mathscr{N}_h}$ are defined by

$$\psi_i := \varphi_i \circ F(x). \tag{2.2.2}$$

Let V^h be the space spanned by ψ_i. Write $u_h \in V^h$ the solution of the Galerkin scheme associated to (2.1.1) based on the shape functions $(\psi_i)_{i\in\mathscr{N}_h}$.

$$a[\psi_i, u_h] = (\psi_i, g)_{L^2(\Omega)}. \tag{2.2.3}$$



(a) φ_i **(b)** $\psi_i = \varphi_i \circ F$

Figure 2.3: The Galerkin elements

where $a[\cdot,\cdot]$ is the bilinear form defined on $H_0^1(\Omega)$,

$$a[v,w] := \int_\Omega \nabla v a \nabla w. \tag{2.2.4}$$

We have the following Theorem 2.2.1 which implies that solutions to (2.0.1) live in the H^1-norm neighborhood of a low dimensional space.

Theorem 2.2.1. *Assume that σ is stable and $n = 2$, there exist constants $\alpha, C > 0$ such that*

$$\|u - u_h\|_{H^1} \leq C h^\alpha \|g\|_{L^p(\Omega)}. \tag{2.2.5}$$

The constant α depends only on n, Ω, $\lambda_{\min}(a)$, $\lambda_{\max}(a)$, and μ_σ. The constant C depends on the objects mentioned above plus $\gamma(\mathscr{T}_h)$ and $\left\| \left(\mathrm{Trace}(\sigma) \right)^{-1} \right\|_{L^\infty(\Omega)}$.

Remark. *Theorem 2.2.1 is also valid with $\alpha = 1$ as in Theorem 2.2.2. The only difference between these two theorems lies in the constant C. In the proof of Theorem 2.2.1 we use the property $u \circ F^{-1} \in C^{1,\alpha}(\Omega)$ and in the proof of Theorem 2.2.2 we use the property $u \circ F^{-1} \in W^{2,2}(\Omega)$.*

Remark. *The proof of Theorems 2.2.1 and 2.2.2 is done for the exact function ψ_i and not for its discrete version. In the implementation, we use piecewise linear function on the fine mesh to approximate ψ_i. If a is regular at a given small scale h_0 then it is easy to check*

that Theorem 2.2.1 remains valid as long as the edges of the fine mesh are smaller than h_0. A more intriguing case is when a is discrete and discontinuous on a fine mesh. Numerical experiments show that Theorems such as 2.2.1 and 2.1.1 remain valid.

Remark. *We keep the composition rule used in [8]. The only difference between the elements (2.2.2) and the ones proposed by Hou, Wu, Allaire, and Brizzi lies in the fact that we use the global solution to (2.1.1) and not a local one computed on each triangle of the coarse mesh through an over-sampling technique.*

For dimension $n \geq 2$ we have the following estimate:

Theorem 2.2.2. *Assume that σ is stable, $n \geq 2$, $\left\| \left(\mathrm{Trace}(\sigma) \right)^{-1} \right\|_{L^\infty(\Omega)} < \infty$ and $\left\| \mathrm{Trace}(\sigma) \right\|_{L^\infty(\Omega)} < \infty$. Then there exist constants $p > 2$, $C > 0$ such that*

$$\|u - u_h\|_{H^1(\Omega)} \leq Ch \|g\|_{L^p(\Omega)}. \tag{2.2.6}$$

Furthermore we have the L^2 estimate,

$$\|u - u_h\|_{L^2(\Omega)} \leq Ch^2 \|g\|_{L^p(\Omega)}. \tag{2.2.7}$$

The constant C depends on n, $\gamma(\mathscr{T}_h)$, Ω, β_σ, $\lambda_{\max}(a)$, $\lambda_{\min}(a)$, $\left\| \left(\mathrm{Trace}(\sigma) \right)^{-1} \right\|_{L^\infty(\Omega)}$ and $\left\| \mathrm{Trace}(\sigma) \right\|_{L^\infty(\Omega)}$.

Remark. *Compared with the numerical results in section 2.4, we believe that the theoretical estimates are not optimal. We conjecture that the resulted equation is better conditioned after the coordinate transformation, although the comparison of $\mu(\sigma)$ and the aspect ratio of a does not directly quantifies this effect. See the discussion at p.54.*

2.2.2 Localized Nonconforming Finite Element Method

For clarity, we will restrict to dimension two from now on, the generalization of the statements to higher dimensions is conditioned on the stability of σ (and the application of Theorem 2.1.3).

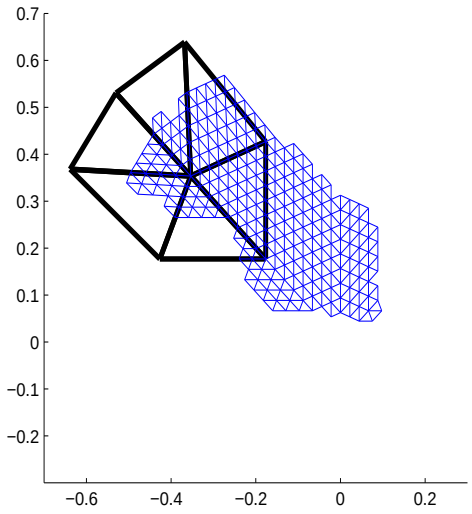


Figure 2.4: Support of the elements φ_i and ψ_i

The elements ψ (2.2.2) can be supported on highly distorted and non-local domain (figure 2.4) since

$$\text{support}(\psi_i) := F^{-1}\big(\text{support}(\varphi_i)\big). \tag{2.2.8}$$

Is it possible to avoid that difficulty by solving (2.0.1) on a coarse mesh with localized elements supported on a regular domain? The answer is yes, but the price to pay will be the discontinuity of the elements, which results in a nonconforming finite element method.

Recall that when the coefficients $a(x)$ of the PDE (2.0.1) are L^∞, F is Hölder continuous by Theorem A.1.2 in Appendix A.1.1, also see [120, 67]). It is meaningful to look at the point value of F. Now let v be a function defined on the nodes a, b, c of the triangle $K \in \mathscr{T}_h$. It is natural to look at the so-called *'coarse gradient'* of v evaluated at the nodes of the triangle K, i.e., the vector defined by

$$\nabla v(K) := \begin{pmatrix} b-a \\ c-a \end{pmatrix}^{-1} \begin{pmatrix} v(b)-v(a) \\ v(c)-v(a) \end{pmatrix}. \tag{2.2.9}$$

Define $\eta^F_{\min}(K)$ the weak aspect ratio induced by F

$$\eta^F_{\min}(K) = \frac{1}{\sin\theta}. \tag{2.2.10}$$

where θ is the interior angle of the triangle $K_F = (F(a), F(b), F(c))$ which is closest to $\pi/2$. $\eta^F_{\min}(K)$ is large if the triangle K_F is flat. We define

$$\eta^*_{\min} = \sup_{K\in\mathscr{T}_h} \eta^F_{\min}(K) \tag{2.2.11}$$

If the weak aspect ratio of the triangle K $\eta^F_{\min}(K) < \infty$ then the following object called the *'coarse gradient'* of v with respect to the metric induced by F is well defined.

$$\nabla_F v(K) := \begin{pmatrix} F(b)-F(a) \\ F(c)-F(a) \end{pmatrix}^{-1} \begin{pmatrix} v(b)-v(a) \\ v(c)-v(a) \end{pmatrix}. \tag{2.2.12}$$

Now consider the nodal elements $(\xi_i)_{i\in\mathscr{N}_h}$, defined by

$$\begin{cases} \xi_i(x_j) = \delta_{ij} \\ \nabla_F \xi(x) = \text{constant within each} \quad K \in \mathscr{T}_h. \end{cases} \tag{2.2.13}$$

If the mesh is not unadapted to F then the elements (figure 2.5) (2.2.13) are well defined and given by

$$\begin{cases} \xi_i(x) = 1 + \big(F(x) - F(x_i)\big)\nabla_F \varphi_i(K) \quad \text{if} \quad i \sim K \quad \text{and} \quad x \in K \\ \xi_i(0) = 0 \quad \text{in other cases} \end{cases} \tag{2.2.14}$$

where the notation $i \sim K$ means that i is a node of K. Observe that the elements ξ_i are discontinuous at the boundaries of the triangles of the coarse mesh, but continuous at the vertices, therefore we obtain a nonconforming finite element method. ξ_i are easier to implement since they are localized in these triangles. Write Z^h the vector space spanned by the functions ξ_i.

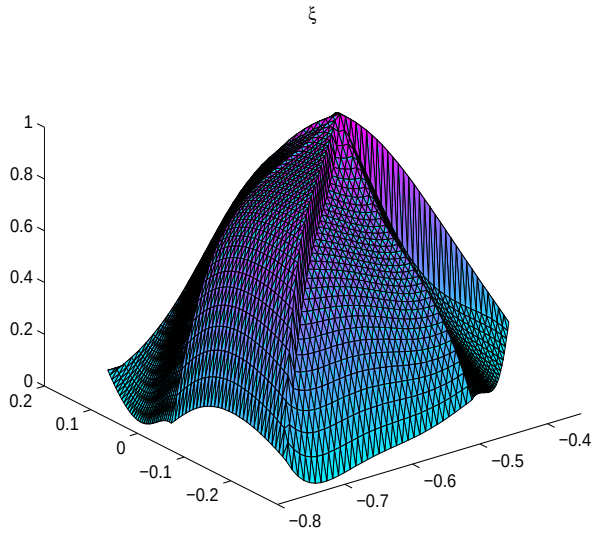


Figure 2.5: Localized Galerkin elements ξ_i

For $K \in \mathscr{T}_h$ we write a_K the bilinear form on $H^1(K)$ defined by

$$a_K[v,w] := \int_K {}^t\nabla v a \nabla w \, dx. \tag{2.2.15}$$

Write $H^1(\mathscr{T}_h)$ the space of functions $v \in L^2(\Omega)$ such that the restriction of v to each triangle K belongs to $H^1(K)$. For $v, w \in H^1(\mathscr{T}_h)$, define

$$a^*[v,w] := \sum_{K \in \mathscr{T}_h} a_K[v,w]. \tag{2.2.16}$$

The nonconforming finite element method can be formulated in the following way: look for $u^f \in \mathscr{Z}^h$ such that for all $i \in \mathscr{N}_h$,

$$a^*[\xi_i, u^f] = (\xi_i, g)_{L^2(\Omega)}. \tag{2.2.17}$$

Let $\|v\|_{h,K} = (\int_K |\nabla u|^2 dx)^{1/2}$, $\|v\|_h = \sum_K \|v\|_{h,K}$, then $\|v\|_h$ is a norm of $H^1(\mathscr{T}_h)$. Write $\mathscr{Z}^h u$ the interpolation of u over space Z^h:

$$\mathscr{Z}^h u(x) = \sum_{i \in \mathscr{N}_h} u(x_i)\xi_i(x). \tag{2.2.18}$$

It is well known that the numerical error of nonconforming finite element method is composed of interpolation error and nonconforming error (see e.g. [40, Lemma 8.1.7]),

$$\|u-u^f\|_h \le C\big(\|u-\mathscr{Z}^h u\|_h + \sup_{w_h\in Z^h\setminus\{0\}} \frac{|(w_h,g)-a^*(w_h,u)|}{\|w_h\|_h}\big). \tag{2.2.19}$$

The nonconforming error is

$$e_n(Z^h) := \sup_{w_h\in Z^h\setminus\{0\}} \frac{|(w_h,g)-a^*(w_h,u)|}{\|w_h\|_h}. \tag{2.2.20}$$

We will show that the interpolation error $\|u-\mathscr{Z}^h u\|_h$ depends on $\eta^*_{\min}$, which is the weak aspect ratio induced by F, and the nonconforming error $e_n(Z^h)$ depends on $\chi^*(h)$, which is the error of piecewise linear approximation to F.

$\chi^*(h)$ is defined in the following: Let $A\ominus B$ be the symmetric set difference of A and B. For $K\in\mathscr{T}_h$, we can quantify the approximation error of piecewise linear approximation to F by

$$\chi^F(K) := \frac{\text{area}(F(K)\ominus K^F)}{\text{area}(K^F)}, \tag{2.2.21}$$

and,

$$\chi^*(h) := \sup_{K\in\mathscr{T}_h} \chi^F(K). \tag{2.2.22}$$

The accuracy of this approximation itself is an interesting problem, see [95, P. 46-51, 'PL approximations of homeomorphisms'] and [96]. We make the following assumption:

Assumption 2.2.1. *$\chi^*(h)\to 0$ as $h\to 0$. Furthermore, since F is a C^α homeomorphism, there exists some $\beta>0$, such that $\chi^*(h)\le Ch^\beta$.*

We say that the tessellation $\mathscr{T}_h$ is *not unadapted* to F if and only if the determinant of $\nabla F(K)$ is strictly positive for all $K\in\mathscr{T}_h$ and Assumption 2.2.1 holds. Observe that if the tessellation $\mathscr{T}_h$ is not unadapted to F then $\eta^*_{\min}(K)<\infty$, the requirement $\det(\nabla F(K))>0$ contains additional condition that there is no inversion in the images of the triangles of $\mathscr{T}_h$ by F. Now we have the following theorem,

Theorem 2.2.3. *Assume that σ is stable and that the mesh is not unadapted to F. Then*

there exists a constant $\alpha > 0$ such that

$$\left(a^*[u-u^f]\right)^{\frac{1}{2}} \leq Ch^{\alpha}\|g\|_{L^p(\Omega)}. \tag{2.2.23}$$

The constant α depends only on n, Ω, ε, μ_σ, $\lambda_{\max}(a)$, $\lambda_{\min}(a)$, and χ^*. The constant C depends on the objects mentioned above plus $\eta^*_{\min}$ and $\left\|\left(\mathrm{Trace}(\sigma)\right)^{-1}\right\|_{L^\infty(\Omega)}$.

Remark. *The bilinear operator $a^*[\cdot,\cdot]$ on Z^h is characterized by a constant matrix within each triangle $K \in \mathscr{T}_h$ equal to*

$$^t\left(\nabla F(K)\right)^{-1} \langle {}^t\nabla F a \nabla F\rangle_K \left(\nabla F(K)\right) \tag{2.2.24}$$

where $\langle v\rangle_K$ means the average of v over K with respect to the Lebesgue measure

$$\langle v\rangle_K := \frac{1}{\mathrm{Vol}(K)}\int_K v(x)\,dx \tag{2.2.25}$$

and $\mathrm{Vol}(K)$ is the Lebesgue measure of volume of K.

Observe that u^f is discontinuous at the boundaries of the triangles of the coarse mesh; we have to find an accurate way to interpolate u^f in the whole space using its values at the nodes of the coarse mesh. Let us write $F(\mathscr{N}_h)$ the image of the nodes of $\mathscr{T}_h$ by F, $\mathscr{T}^F$ the triangulation of $F(\mathscr{N}_h)$. Suppose φ_i^F is the standard piecewise linear nodal basis of $\mathscr{T}^F$. Write $\mathscr{J}_h$ the interpolation operator from the space of functions defined on the nodes of $\mathscr{T}_h$ into $H^1(\Omega)$ defined by

$$\mathscr{J}_h v(x) := \sum_{i\in\mathscr{N}_h} v(x_i)\varphi_i^F \circ F(x). \tag{2.2.26}$$

Observe that for $i \in \mathscr{N}_h$, $v(x_i) = \mathscr{J}_h v(x_i)$. We have the following estimate

Theorem 2.2.4. *Assume that σ is stable and that the mesh is not unadapted to F. Then there exist constants $\alpha, C_f > 0$ such that*

$$\|u - \mathscr{J}_h u^f\|_{H^1(\Omega)} \leq C_f h^{\alpha}\|g\|_{L^p(\Omega)}. \tag{2.2.27}$$

The constant α is the same as in Theorem 2.2.3. The constant C_f depends on the same objects as in Theorem 2.2.3 plus ν^*, which is defined by

$$\nu^* := \sup_{K\in\mathscr{T}_h} \frac{\mathrm{Vol}(K^F)}{\mathrm{Vol}(F(K))} \tag{2.2.28}$$

where K^F is the triangle whose nodes are the images of the nodes of K by F.

2.2.3 Numerical Homogenization from the Information Point of View: Effective Operator on Coarse Scale

The Galerkin schemes described in subsections 2.2.1 and 2.2.2 are based on elements containing the complete fine scale structure of F, which represents too much information. We can wonder: What minimal information should be kept from the fine scales in order to up-scale (2.0.1)? We would like to keep an accurate version of (2.0.1) with minimal computer memory. Otherwise stated, we are considering compression from the point of view of numerical homogenization. We view the operator (2.0.1) as a bilinear form on $H_0^1(\Omega)$ and we will use X^h as space of test functions to *zoom at* the operator associated to a at a given arbitrary resolution.

$$a : \begin{cases} H_0^1(\Omega)\times H_0^1(\Omega) & \to \mathbb{R} \\ (v,w) & \to \int_\Omega {}^t\nabla v a \nabla w. \end{cases} \tag{2.2.29}$$

The up-scaled or compressed operator, written $\mathscr{U}_h a$, will naturally be a bilinear form on the space of piecewise linear functions on the coarse mesh with Dirichlet boundary condition.

$$\mathscr{U}_h a : \begin{cases} X^h \times X^h & \to \mathbb{R} \\ (v,w) & \to \mathscr{U}_h a[v,w]. \end{cases} \tag{2.2.30}$$

The question is how to choose $\mathscr{U}_h a$? To answer that question we can integrate (2.0.1)

against a test function φ in X^h, then we obtain that

$$\int_\Omega \nabla\varphi a \nabla u \, dx = \int_\Omega \varphi g \, dx. \tag{2.2.31}$$

We will use the test function φ to 'look at' the operator (2.0.1) at the given resolution h. We can decompose the first term in the integral above as a sum of integrals over the triangles of the coarse mesh to obtain (we assume that σ is stable),

$$\int_\Omega \nabla\varphi a \nabla u \, dx = \sum_{K\in\mathscr{T}_h} \int_K \nabla\varphi(x) a(x) \nabla F(x) \big(\nabla F(x)\big)^{-1} \nabla u(x) \, dx. \tag{2.2.32}$$

Now $\nabla\varphi$ is constant within each triangle $K \in \mathscr{T}_h$. $\big(\nabla F(x)\big)^{-1}\nabla u(x)$ is Hölder continuous, thus we can approximate it by a constant within each triangle K and equal to the coarse gradient of u induced by F, i.e.,

$$\nabla_F u(K) := \begin{pmatrix} F(b)-F(a) \\ F(c)-F(a) \end{pmatrix}^{-1} \begin{pmatrix} u(b)-u(a) \\ u(c)-u(a) \end{pmatrix}. \tag{2.2.33}$$

where a, b, c are the nodes of the triangle K. It follows that the tensor $a\nabla F$ can be averaged over each triangle of the coarse mesh and we will write $\langle a\nabla F\rangle_K$ its average. In conclusion a good candidate for the up-scaled operator $\mathscr{U}_h a$ is the bilinear form given by the following formula: for $v, w \in X^h$

$$\mathscr{U}_h a[v,w] := \sum_{K\in\mathscr{T}_h} \int_K {}^t\nabla v \big\langle a\nabla F\big\rangle_K \big(\nabla F(K)\big)^{-1} \nabla w. \tag{2.2.34}$$

Observe that the only information kept from the small scales in the compressed operator (2.2.34) is the bulk quantities $\langle a\nabla F\rangle_K$ and the non averaged quantities $F(b)-F(a)$, where a and b are nodes of the triangles of the coarse mesh. The latter quantity can be interpreted as a deformation of the coarse mesh induced by the small scales (or a new distance defining coarse gradient). In particular, when $a = M(\frac{x}{\varepsilon})$ and M is ergodic, then as $\varepsilon \downarrow 0$ $\langle a\nabla F\rangle_K$ converges to the usual effective conductivity obtained from homogenization theory and $\nabla F(K)$ converges to the identity matrix. It follows that the object (2.2.34) recovers the formulae

obtained from homogenization theory when the medium is ergodic and characterized by scale separation.

To estimate the accuracy of compression we have to use the up-scaled operator $\mathscr{U}_h a$ to obtain an approximation of the linear interpolation of u on the coarse mesh. We look for $u^m \in X^h$ such that for all $i \in \mathscr{N}_h$,

$$\mathscr{U}_h a[\varphi_i, u^m] = (\varphi_i, g)_{L^2(\Omega)}. \tag{2.2.35}$$

The price to pay for the loss of information on the small scales is the loss of ellipticity. This loss can be caused by two correlated factors:

- The new metric can generate flat triangles.
- The up-scaled operator can become singular.

The first factor is due to the localization of the scheme. The second factor does not appear with Galerkin schemes. It is not observed in dimension two but it can not be avoided in dimension greater or equal to three in the sense that the up-scaled operator has no reason to remain elliptic and local. Indeed, consider a box of dimension three, and set in that box tubes of high conductivity as shown in figure 2.6. Set the left side of the box to temperature $0^0 c$ and the right side to temperature $100^0 c$. Then an inversion in the temperature profile is produced around the critical points shown in figure 2.6 (see [10] and [44], instead of increasing from left to right in these regions temperature decreases). Now as the operator is up-scaled, the information on the geometry of the tubes is lost but the inversion phenomenon remains in the loss of ellipticity and locality of the operator.

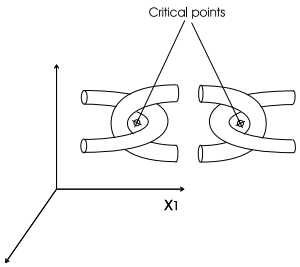


Figure 2.6: a in dimension three

Nevertheless it is possible to prove that once stability is achieved then the method is accurate (if σ is stable). Notice that we can rewrite (2.2.34) into

$$\begin{aligned}\mathscr{U}_h a[v,w] :=& \sum_{K\in\mathscr{T}_h}\int_K {}^t\nabla v\big\langle a\nabla F\big\rangle_K\big(\nabla F(K)\big)^{-1}\nabla w\,dx\\ =& \sum_{K\in\mathscr{T}_h}\int_K {}^t\nabla v a\nabla\mathscr{Z}^h w\,dx\\ =& a^*[v,\mathscr{Z}^h w].\end{aligned} \tag{2.2.36}$$

Thus we can rewrite (2.2.35) as

$$a^*[v,u^z] = (v,g)_{L^2(\Omega)}. \tag{2.2.37}$$

$\forall v\in X^h$ and $u^z\in\mathscr{Z}^h$. For a nodal function v, let us define the homogeneous Dirichlet form on the graph induced by $\mathscr{T}_h$:

$$\mathscr{E}_h[v] := \sum_{i\sim j}|v_i - v_j|^2 \tag{2.2.38}$$

where $v_i = v(x_i)$.

Remark. *Let us recall that for $v\in X^h$, $\mathscr{E}_h[v]$ can be bounded from below and above by the L^2-norm of the gradient of v. More precisely*

$$\frac{1}{4\eta_{\max}}\mathscr{E}_h[v] \le \|\nabla v\|^2_{L^2(\Omega)} \le \eta_{\max}\mathscr{E}_h[v]. \tag{2.2.39}$$

which means that $\mathscr{E}_h[v]$ is equivalent to $\|v\|_{H^1_0(\Omega)}$. $\eta_{\max} = 1/\sin(\theta)$ where θ is the closest interior angle of the triangles of $\mathscr{T}_h$ to 0 or π.

By (2.2.37), the test function and trial function are in different spaces, this category of finite element methods is called Petrov-Galerkin method [16, Section 8.2]. We write $i\sim j$ when those nodes share an edge on the coarse mesh. Let us define the following stability parameter

$$\mathscr{S}^m := \inf_{w\in Z^h\setminus\{0\}}\sup_{v\in X^h\setminus\{0\}}\frac{\mathscr{U}_h a[v,w]}{\big(\mathscr{E}_h[v]\big)^{\frac{1}{2}}\big(\mathscr{E}_h[w]\big)^{\frac{1}{2}}}. \tag{2.2.40}$$

Remember the inf-sup condition (or Banach-Nečas-Babuška Theorem) [21, p.112][57,

p.84][16, Section 8.2], if and only if $\mathscr{S}^m > 0$, the scheme (2.2.37) is well defined in the sense that a unique solution $u^z \in \mathscr{Z}^h$ exists. In [79], Hou et al. proved the inf-sup condition for the scale separation periodic case and proposed a Petrov-Galerkin MsFEM formulation with nonconforming multi-scale trial functions and linear test functions in order to eliminate the cell resonance error. Observe that $\mathscr{S}^m$ depends only on the up-scaled parameters so we have a control on the stability.

Let us write $\mathscr{I}_h u$ the linear interpolation of u over $\mathscr{T}_h$:

$$\mathscr{I}_h u := \sum_{i \in \mathscr{N}_h} u(x_i)\varphi_i(x). \tag{2.2.41}$$

We have the following estimate,

Theorem 2.2.5. *Assume that σ and the scheme are stable and that the mesh is not unadapted to F. Then there exist constants $\alpha > 0, C_m > 0$ such that*

$$\|\mathscr{I}_h u - u^m\|_{H^1(\Omega)} \leq C_m h^\alpha \|g\|_{L^\infty(\Omega)}. \tag{2.2.42}$$

The constant α depends only on n, Ω, $\lambda_{\min}(a)$, $\lambda_{\max}(a)$, and μ_σ. The constant C_m can be written

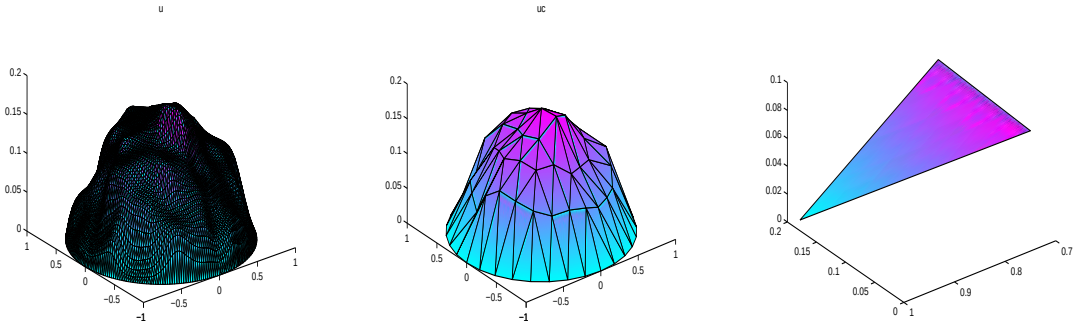
$$C_m := C\frac{\eta^*_{\min}\eta_{\max}}{\mathscr{S}^m}. \tag{2.2.43}$$

where C depends on the objects mentioned above plus and $\left\|\left(\mathrm{Trace}(\sigma)\right)^{-1}\right\|_{L^\infty(\Omega)}$.

The compressed operator allows us to capture the solution of (2.0.1) on a coarse mesh (figure 2.7). We can add information to the compressed operator in order to obtain fine resolution approximation of u (figure 2.8). Indeed let $\mathscr{J}_h$ be the interpolation operator introduced in (2.2.26), we then have the following estimate:

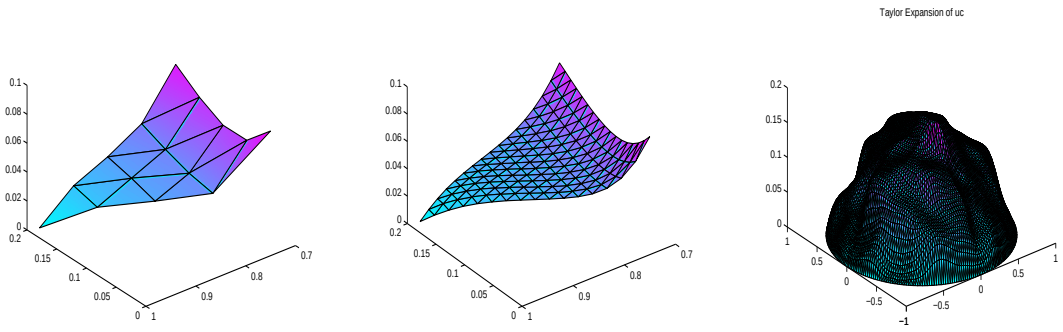
Theorem 2.2.6. *Assume that σ and the scheme 2.2.37 are stable and the mesh is not unadapted to F. Then there exist constants $\alpha > 0, C_m > 0$ such that*

$$\|u - \mathscr{J}_h u^m\|_{H^1(\Omega)} \leq C_m h^\alpha \|g\|_{L^\infty(\Omega)}. \tag{2.2.44}$$



(a) u **(b)** u^m **(c)** ***Triangle of the coarse mesh***

Figure 2.7: u estimated with the up-scaled operator



(a) ***Refined*** F**.** **(b)** ***Refined*** F **(c)** $\mathcal{J}_h u$

Figure 2.8: Taylor expansion with respect to the new metric

The constant α depends only on n, Ω, $\lambda_{\min}(a)$, $\lambda_{\max}(a)$, and μ_σ. The constant C_m can be written as,

$$C_m := C\eta_{\max}\eta^*_{\min}\left(\frac{\mu_\sigma \nu^*}{\mathscr{S}m}\right)^{\frac{1}{2}} \tag{2.2.45}$$

where C depends on the objects mentioned above plus $\left\|\left(\mathrm{Trace}(\sigma)\right)^{-1}\right\|_{L^\infty(\Omega)}$.

2.2.4 Numerical Homogenization from the Transport Point of View: a Multi-scale Finite Volume Method

Let us write $\mathscr{T}_h^*$ the dual mesh associated to $\mathscr{T}_h$. $\mathscr{T}_h^*$ can be obtained by drawing segments from the midpoints of the edges of the triangles of $\mathscr{T}_h$ to an interior point in these triangles: We can choose the interior point to be the circumcenter to obtain a Voronoï tessellation but one can also choose the barycenter [59].

Let us write V_i the control volume associated to the node i of the primal mesh and χ_i the characteristic function of V_i. The finite volume method can be expressed in the following way: Look for $u^\nu \in \mathscr{Z}^h$ ($\mathscr{Z}^h$ being the space spanned by the elements ξ_i introduced in (2.2.13)) such that for all $i \in \mathscr{N}_h$,

$$a^*[\chi_i, u^\nu] = (\chi_i, g)_{L^2}. \tag{2.2.46}$$

Again, it follows from equation (2.2.46) that the only information kept from fine scales are the usual bulk quantities (effective conductivities at the edges of the dual mesh) plus the metric information $F(b) - F(a)$ where a and b are nodes of the triangles of the primal mesh. According to (2.2.46) the good choice for the jump rates of the random walk should be

$$\gamma_{ij} = a^*[\chi_i, \xi_j] \quad \text{if} \quad i \sim j \quad \text{and} \quad i \neq j. \tag{2.2.47}$$

To properly describe the transport process one should look at a parabolic operator instead of the elliptic one. We will restrict ourselves to the elliptic case characterizing the equilibrium properties of the random walk.

Notice that the test function and trial function in (2.2.46) are in different spaces. Similar to (2.2.40), write $\mathscr{S}^\nu$ the stability parameter of the up-scaled finite volume operator, which

is defined by

$$\mathscr{S}^{\nu} := \inf_{w\in\mathscr{Z}^h\setminus\{0\}} \sup_{v\in\mathscr{Y}_h\setminus\{0\}} \frac{a^*[v,w]}{\left(\mathscr{E}_h[v]\right)^{\frac{1}{2}}\left(\mathscr{E}_h[w]\right)^{\frac{1}{2}}}. \tag{2.2.48}$$

Theorem 2.2.7. *Assume that σ and the scheme are stable ($\mathscr{S}^{\nu} > 0$) and that the mesh is not unadapted to F. Then there exist constants $\alpha > 0, C_{\nu} > 0$ such that*

$$\|\mathscr{I}_h u - \mathscr{I}_h u^{\nu}\|_{H^1(\Omega)} \leq C_{\nu,1} h^{\alpha} \|g\|_{L^p(\Omega)} \tag{2.2.49}$$

and

$$\|u - \mathscr{J}_h u^{\nu}\|_{H^1(\Omega)} \leq C_{\nu,2} h^{\alpha} \|g\|_{L^p(\Omega)}. \tag{2.2.50}$$

The constant α depends only on n, Ω, $\lambda_{\min}(a)$, $\lambda_{\max}(a)$, and μ_{σ}. The constant $C_{\nu,1}$ can be written as

$$C_{\nu,1} := C\frac{\eta^*_{\min}\eta_{\max}}{\mathscr{S}^{\nu}}. \tag{2.2.51}$$

The constant $C_{\nu,2}$ can be written

$$C_{\nu,2} := C\left(\frac{\eta^*_{\min}\eta_{\max}\mu_{\sigma}\nu^*}{\mathscr{S}^{\nu}}\right)^{\frac{1}{2}}. \tag{2.2.52}$$

C depends on the objects mentioned above plus and $\left\|\left(\mathrm{Trace}(\sigma)\right)^{-1}\right\|_{L^{\infty}(\Omega)}$.

Remark. *Numerical experiments show that although the finite volume method keeps very little information from small scales it is more stable and accurate than the method presented in 2.2.3 (it is also more stable and almost as accurate as Galerkin method in which the whole fine scale structure of F is used). That is why we believe that the constants (2.2.51) and (2.2.52) are not optimal.*

2.3 Proofs

2.3.1 Compensation Phenomena

2.3.1.1 Proof of Theorem 2.1.1–2.1.3

To prove Theorem 2.1.1, we need a variation of Campanato's result [46] on non-divergence form elliptic operators. The $W^{2,p}$ solvability Theorem A.1.3 is included in the appendix A.1.2.

Assume that σ is stable. Write F^{-1} the inverse of F (which is well defined if σ is stable). Write Q the symmetric positive matrix given by the following equation

$$Q(y) := \Big(\frac{\big({}^t\nabla F a \nabla F\big)}{|\det\big(\nabla F\big)|}\Big) \circ F^{-1}(y). \tag{2.3.1}$$

w is the strong solution of the following non-divergence form elliptic equation:

$$\sum_{i,j=1}^{n} Q_{ij}\partial_i\partial_j w = -\frac{\tilde{g}}{|\det\big(\nabla F\big)| \circ F^{-1}}. \tag{2.3.2}$$

Let us now prove the following theorem,

Theorem 2.3.1. *Assume that σ is stable and that Ω is convex. Then there exists constants $p > 2, C > 0$ such that the solution of (2.3.2) belongs to $W_0^{2,p}(\Omega)$ and satisfies*

$$\|w\|_{W_0^{2,p}(\Omega)} \leq \frac{C}{1-\beta_\sigma^{\frac{1}{2}}}\|g\|_{L^p(\Omega)}. \tag{2.3.3}$$

C depends on $\lambda_{\min}(a)$, n, p, Ω, and $\|\big(\mathrm{Trace}(\sigma)\big)^{\frac{n}{2p}-1}\|_{L^\infty(\Omega)}$ The theorem also holds for p=2.

Proof. Since

$$\nu_Q = \frac{\sum_{i=1}^n \lambda_{i,Q}}{\sum_{i=1}^n \lambda_{i,Q}^2} = \frac{\mathrm{Trace}(Q)}{\mathrm{Trace}(Q^2)} \circ F^{-1} \tag{2.3.4}$$

we have

$$\frac{\nu_Q}{\det(\nabla F)\circ F^{-1}} = \frac{\mathrm{Trace}(\sigma)}{\mathrm{Trace}(\sigma^2)} \circ F^{-1}. \tag{2.3.5}$$

Using the change of variables $y = F(x)$, we obtain that

$$\|\frac{\nu_Q}{\det(\nabla F)\circ F^{-1}}\tilde{g}\|_{L^p(\Omega)} \leq \|g\|_{L^p(\Omega)}\|\frac{\mathrm{Trace}(\sigma)}{\mathrm{Trace}(\sigma^2)}(\det(\nabla F))^{\frac{1}{p}}\|_{L^\infty(\Omega)}. \tag{2.3.6}$$

Since $\sigma = {}^t\nabla F a \nabla F$, we have $\det(\sigma) = (\det(\nabla F))^2 \det(a)$, use inequalities

$$|\det(M)| \leq C(\mathrm{Trace}(M))^n, \tag{2.3.7}$$

and

$$\mathrm{Trace}(M^2) \leq C(\mathrm{Trace}(M))^2. \tag{2.3.8}$$

It is easy to check that

$$\|\frac{\mathrm{Trace}(\sigma)}{\mathrm{Trace}(\sigma^2)}(\det(\nabla F))^{\frac{1}{p}}\|^p_{L^\infty(\Omega)} \leq \frac{C_{qp,n}}{(\lambda_{\min}(a))^{\frac{n}{2}}}\|\big(\mathrm{Trace}(\sigma)\big)^{\frac{n}{2p}-1}\|^p_{L^\infty(\Omega)}. \tag{2.3.9}$$

Observe that $\beta_Q = \beta_\sigma$, a direct application of Theorem A.1.3 and estimate (2.3.5) to equation (2.3.2) implies the theorem. The proof also applies to the case $p = 2$. □

Using the well known De Giorgi-Moser-Nash theory ([69], [97], [100]) for divergence form elliptic operators with discontinuous coefficients (more precisely we refer to [120] for the Global Hölder regularity), there exists $C, \alpha' > 0$ depending on Ω and $\lambda_{\max}(a)/\lambda_{\min}(a)$ such that F is α' Hölder continuous and

$$\|F\|_{C^{\alpha'}} \leq C. \tag{2.3.10}$$

which is stated as Theorem A.1.2 in Appendix A.1.1.

The following lemma is the key observation,

Lemma 2.3.1. *For $\varphi \in C_0^\infty(\Omega)$, We have*

$$a[\varphi, u] = Q[\tilde{\varphi}, \tilde{u}] \tag{2.3.11}$$

and $\mathrm{div}_y Q(y) = 0$ in the weak sense.

Proof. Since $F \in C^{\alpha'}$, $\tilde{\varphi} \in H_0^1(\Omega)$,

$$\begin{aligned}\int_\Omega \nabla\varphi a \nabla u\, dx &= \int_\Omega \nabla\varphi \nabla F^{-1} \nabla F a \nabla F \nabla F^{-1} \nabla u dx \\ &= \int_\Omega \nabla\tilde{\varphi} Q \nabla \tilde{u} dy.\end{aligned}$$

To prove $Q(y)$ is divergence free, it is enough to prove that for all $\tilde{\varphi} \in C_0^\infty(\Omega)$,

$$\int_\Omega Q(y) \nabla\tilde{\varphi} dy = 0. \tag{2.3.12}$$

In fact,

$$\begin{aligned}\int_\Omega \sigma \circ F^{-1} \nabla\tilde{\varphi} \frac{dy}{\det(\nabla F)\circ F^{-1}} &= \int_\Omega \nabla F a \nabla F \nabla F^{-1} \nabla\varphi\, dx \\ &= \int_\Omega \nabla F a \nabla\varphi\, dx \\ &= -\int_\Omega \varphi \nabla a \nabla F\, dx \\ &= 0.\end{aligned}$$

□

Let $\varphi \in C_0^\infty(\Omega)$. Write $\tilde{\varphi} := \varphi \circ F^{-1}$. Using Theorem 2.3.1 we obtain that

$$\Big(\tilde{\varphi}, \sum_{i,j=1}^n Q_{ij} \partial_i \partial_j w\Big)_{L^2(\Omega)} = -\Big(\tilde{\varphi}, \frac{\tilde{g}}{|\det(\nabla F)| \circ F^{-1}}\Big)_{L^2(\Omega)}. \tag{2.3.13}$$

By Lemma 2.3.1,

$$a[\varphi, w \circ F] = (\varphi, g)_{L^2(\Omega)}. \tag{2.3.14}$$

It follows from the uniqueness of the solution of the Dirichlet problem (2.3.14) that $w \circ F = u$. Theorem 2.1.2 is then a straightforward consequence of Theorem 2.3.1 and the equality $u \circ F^{-1} = w$.

Theorem 2.1.1 is a straightforward consequence of the Sobolev embedding inequality (A.1.13), Theorem 2.3.1, Lemma A.1.3, and the fact that $\nabla_F u = \nabla\tilde{u} \circ F$.

2.3.1.2 Hölder Continuity for Nonconvex Domain

In this subsection we will not assume Ω to be convex. Let $N^{p,\lambda}(\Omega)$ $(1 < p < \infty, 0 < \lambda < n)$ be the weighted Morrey space formed by functions $v : \Omega \to \mathbb{R}$ such that $\|v\|_{N^{p,\lambda}(\Omega)} < \infty$ with

$$\|v\|_{N^{p,\lambda}(\Omega)} = \sup_{x_0 \in \Omega} \left(\int_\Omega |x - x_0|^{-\lambda} |v(x)|^p \right)^{\frac{1}{p}}. \tag{2.3.15}$$

To obtain the Hölder continuity of $u \circ F^{-1}$ in dimension $n \geq 3$ we will use corollary 4.1 of [90]. We will give the result of S. Leonardi below in a form adapted to our context. Consider the nondivergence elliptic Dirichlet problem (A.1.3). Write $W^{2,p,\lambda}(\Omega)$ the functions in $W^{2,p}(\Omega)$ such that their second order derivatives are in $N^{p,\lambda}(\omega)$.

Theorem 2.3.2. *There exist a constant $C^* = C^*(n, p, \lambda, \partial\Omega) > 0$ such that if $\beta_M < C^*$ and $f \in N^{p,\lambda}(\Omega)$ then the Dirichlet problem (A.1.3) has a unique solution in $W^{2,p,\lambda} \cap W_0^{1,p}(\Omega)$. Moreover, if $0 < \lambda < n < p$ then $\nabla v \in C^\alpha(\Omega)$ with $\alpha = 1 - n/p$ and*

$$\|\nabla v\|_{C^\alpha(\Omega)} \leq \frac{C}{\lambda_{\min}(M)} \|f\|_{N^{p,\lambda}(\Omega)} \tag{2.3.16}$$

where $C = C(n, p, \lambda, \partial\Omega)$.

Theorem 2.1.3 is then a straightforward application of Theorem 2.3.2.

2.3.2 Finite Element Using Composition Rule

2.3.2.1 Proof of Theorem 2.2.1–2.2.2

Let us prove Theorem 2.2.1. We write V^h the linear space spanned by the elements ψ_i. The solution of the Galerkin scheme satisfies $a[u - u_h, v] = 0$ for all $v \in V^h$. Thus

$$a[u - u_h] = a[u - u_h, u - v]. \tag{2.3.17}$$

It follows by Cauchy-Schwartz inequality that

$$a[u - u_h] \leq \inf_{v \in V^h} a[u - v]. \tag{2.3.18}$$

Let $\tilde{v} := v \circ F^{-1}$ and using the change of variable $y = F(x)$ we obtain

$$a[u-v] = Q[\tilde{u}-\tilde{v}]. \tag{2.3.19}$$

In dimension $n = 2$ it follows that

$$\|u-u_h\|^2_{H^1} \leq \frac{D}{\lambda_{\min}(a)} \inf_{w \in X^h} \|\nabla \tilde{u} - \nabla w\|^2_{L^\infty(\Omega)}. \tag{2.3.20}$$

with

$$D := \text{Trace}\Big[\int_\Omega {}^t\nabla F a \nabla F \, dx\Big]. \tag{2.3.21}$$

Using the following standard approximation properties of the elements φ_i (see for instance [57]),

$$\inf_{w \in X^h} \|\nabla \tilde{u} - \nabla w\|_{L^2(\Omega)} \leq C\gamma(\mathscr{T}_h) h^\alpha \|\tilde{u}\|_{C^{1,\alpha}_0(\Omega)}. \tag{2.3.22}$$

we obtain that

$$\|u-u_h\|_{H^1} \leq C\gamma(\mathscr{T}_h)\Big(\frac{D}{\lambda_{\min}(a)}\Big)^{\frac{1}{2}} \|\nabla \tilde{u}\|_{C^\alpha} h^\alpha. \tag{2.3.23}$$

We conclude by observing that for $l \in \mathbb{R}^n$

$$\int_\Omega {}^t l {}^t\nabla F a \nabla F l = \inf_{f \in C^\infty_0(\Omega)} \int_\Omega {}^t(l+\nabla f) a (l+\nabla f). \tag{2.3.24}$$

which means that D is indeed bounded. Theorem 2.2.1 becomes a direct consequence of (2.3.20) and Theorem 2.3.1.

In dimension $n \geq 3$ we obtain from (2.3.19) that

$$\|u-u_h\|^2_{H^1} \leq \frac{\lambda_{\max}(Q)}{\lambda_{\min}(a)} \inf_{w \in X^h} \|\nabla \tilde{u} - \nabla w\|^2_{L^2(\Omega)}. \tag{2.3.25}$$

It is easy to obtain that

$$\lambda_{\max}(Q) \leq \big(\det(a)\big)^{\frac{1}{2}} \mu_{\sigma}^{\frac{n}{2}} \big(\mathrm{Trace}(\sigma)\big)^{1-\frac{n}{2}}. \tag{2.3.26}$$

We conclude by observing that $\mu_{\sigma} < C(\beta_{\sigma})$ and using the following standard approximation properties of the elements φ_i (see for instance [57]).

$$\inf_{w \in X^h} \|\nabla \tilde{u} - \nabla w\|_{L^2(\Omega)} \leq C\gamma(\mathscr{T}_h) h \|\tilde{u}\|_{W^{2,2}(\Omega)}. \tag{2.3.27}$$

We can use the Aubin-Nitsche technique [17, 102] for L^2 estimate, the proof follows from standard duality techniques (see for instance Theorem 5.7.6 of [40]). Choose $v \in H_0^1(\Omega)$ to be the solution of the following linear problem: for all $w \in H_0^1(\Omega)$),

$$a[w, v] = (w, u - u_h)_{L^2(\Omega)}. \tag{2.3.28}$$

Choosing $w = u - u_h$ in equation (2.3.28), for $\forall \varphi_h \in X^h$, we have

$$\|u - u_h\|^2_{L^2(\Omega)} = a[u - u_h, v - \varphi_h]. \tag{2.3.29}$$

Using Cauchy Schwartz inequality, since φ_h is arbitrary, we deduce that

$$\begin{aligned}
\|u - u_h\|^2_{L^2(\Omega_T)} &\leq \big(a[u - u_h]\big)^{\frac{1}{2}} \inf_{\varphi_h \in X^h} \big(a[v - \varphi_h]\big)^{\frac{1}{2}} && (2.3.30)\\
&\leq Ch\|\tilde{u}\|_{W^{2,2}(\Omega)} Ch\|\tilde{v}\|_{W^{2,2}(\Omega)} && (2.3.31)\\
&\leq Ch^2 \|\tilde{u}\|_{W^{2,2}(\Omega)} \|u - u_h\|_{L^2(\Omega)}. && (2.3.32)
\end{aligned}$$

Therefore,

$$\|u - u_h\|_{L^2(\Omega)} \leq Ch^2 \|\tilde{u}\|_{W^{2,2}(\Omega)}. \tag{2.3.33}$$

2.3.3 Localized Nonconforming Finite Element Method

2.3.3.1 Proof of Theorem 2.2.3

Let us prove Theorem 2.2.3. We assume that the coarse mesh is not unadapted to F. Let K be a triangle of $\mathscr{T}_h$ and let a be a node of K, such that $\eta^F_{\min}(K)$ is the weak aspect ratio induced by F over triangle K. $\eta^F_{\min}(K) = \frac{1}{\sin\theta}$, where θ is the interior angle between $(F(a),F(b))$ and $(F(a),F(c))$, b and c are the other nodes of K. We can prove the following lemma,

Lemma 2.3.2.

$$|\nabla_F u(K) - \nabla_F u(a)| \leq 3\eta^F_{\min}(K)\|\nabla\tilde{u}\|_{C^\alpha}\|F\|^\alpha_{C^{\alpha'}} h^{\alpha\alpha'}. \tag{2.3.34}$$

Proof. It is easy to check that

$$u(b) - u(a) = \big(F(b) - F(a)\big)\nabla\tilde{u}\circ F(a) + (F(b) - F(a))\cdot q_{ba}. \tag{2.3.35}$$

where the vector q_{ba} is defined by

$$q_{ba} := \int_0^1 \Big[\nabla\tilde{u}\big[F(a) + s(F(b) - F(a))\big] - \nabla\tilde{u}\big[F(a)\big]\Big]\, ds. \tag{2.3.36}$$

Use the notation $f_{ba} := (F(b) - F(a))/|F(b) - F(a)|$ and write $f^\perp_{ba}$ the unit vector obtained by a 90° rotation of f_{ba} towards f_{ca}. Defining q_{ca} as in (2.3.36) we obtain that

$$\nabla_F u(K) = \nabla_F u(a) + k, \tag{2.3.37}$$

with

$$k = q_{ba} - \lambda\, f^\perp_{ba}, \tag{2.3.38}$$

and

$$\lambda := \frac{f_{ca}.(q_{ba} - q_{ca})}{f_{ca}.f_{ba^\perp}}. \tag{2.3.39}$$

which leads us to

$$|\nabla_F u(K) - \nabla_F u(a)| \leq \frac{3}{f_{ca} \cdot f_{ba^\perp}} \|\nabla \tilde{u}\|_{C^\alpha} \|F\|^\alpha_{C^{\alpha'}} h^{\alpha\alpha'}. \tag{2.3.40}$$

□

The following lemma is a direct consequence of Lemma 2.3.2,

Lemma 2.3.3. *Let $K \in \mathscr{T}_h$ and let $x \in \Omega$ then*

$$|\nabla_F u(K) - \nabla_F u(x)| \leq 3\eta^*_{\min} \|\nabla \tilde{u}\|_{C^\alpha} (1 + \|F\|^\alpha_{C^{\alpha'}}) \big(h + \operatorname{dist}(x,K)\big)^{\alpha\alpha'}. \tag{2.3.41}$$

Lemma 2.3.4. *We have*

$$a^*[u - \mathscr{Z}^h u] \leq C\eta^*_{\min} \|\nabla \tilde{u}\|_{C^\alpha} \|F\|^\alpha_{C^{\alpha'}} h^{\alpha\alpha'} D. \tag{2.3.42}$$

Proof. Since

$$a_K[u - \mathscr{Z}^h u] = \int_K {}^t\big(\nabla_F u(x) - \nabla_F u(K)\big) \sigma(x) \big(\nabla_F u(x) - \nabla_F u(K)\big)\, dx. \tag{2.3.43}$$

with $\sigma(x) := {}^t\nabla F a \nabla F$. Using the change of variables $F(x) = y$ we obtain that

$$a_K[u - \mathscr{Z}^h u] = \int_{F(K)} {}^t\big(\nabla \tilde{u}(y) - \nabla_F u(K)\big) Q(y) \big(\nabla \tilde{u}(y) - \nabla_F u(K)\big)\, dy. \tag{2.3.44}$$

from which we deduce that

$$a_K[u - \mathscr{Z}^h u] \leq \big(3\eta^*_{\min} \|\nabla \tilde{u}\|_{C^\alpha} \|F\|^\alpha_{C^{\alpha'}} h^{\alpha\alpha'}\big)^2 \int_{F(K)} \sup_{|e|=1} {}^t e Q e\, dy. \tag{2.3.45}$$

Thus

$$a^*[u - \mathscr{Z}^h u] \leq C\big(\eta^*_{\min} \|\nabla \tilde{u}\|_{C^\alpha} \|F\|^\alpha_{C^{\alpha'}} h^{\alpha\alpha'}\big)^2 D. \tag{2.3.46}$$

where D has been defined by (2.3.21). □

We need the following lemma to control the nonconforming error:

Lemma 2.3.5.

$$\sup_{w_h \in \mathscr{Z}^h \setminus \{0\}} \frac{|(w_h, g) - a^*(w_h, u)|}{\|w_h\|_h} \leq Ch^\alpha. \tag{2.3.47}$$

Proof. Let $\mathscr{E}_h$ be the set of all the edges in $\mathscr{T}_h$, $\mathscr{E}_h^F$ be the set of all the edges in $\mathscr{T}_h^F$. Observe that

$$\begin{aligned}
a^*[u, w_h] - (w_h, g) &= \sum_{K \in \mathscr{T}_h} \int_K (\nabla u a \nabla w_h - w_h g)\, dx \\
&= \sum_{K \in \mathscr{T}_h} \int_{F(K)} (\nabla \tilde{u} Q(y) \nabla \tilde{u} - \tilde{w}_h Q \nabla \nabla \tilde{u})\, dy \\
&= \sum_{K \in \mathscr{T}_h} \int_{F(K)} \nabla (Q \nabla \tilde{u} \cdot \tilde{w}_h)\, dy \\
&= \sum_{e \in \mathscr{E}_h} \int_{F(e)} Q \frac{\partial \tilde{u}}{\partial n} [\tilde{w}_h]\, ds \\
&= \sum_{e \in \mathscr{E}_h} \int_{F(e)} Q \frac{\partial \tilde{u}}{\partial n} [\tilde{w}_h] ds - \sum_{e \in \mathscr{E}_h^F} \int_e Q \frac{\partial \tilde{u}}{\partial n} [\tilde{w}_h]\, ds \\
&= \sum_{K \in \mathscr{T}_h} \int_{F(K) \ominus K^F} \nabla (Q \nabla \tilde{u} \cdot [\tilde{w}_h])\, dy \\
&= \sum_{K \in \mathscr{T}_h} \int_{F(K) \ominus K^F} (\nabla \tilde{u} Q \nabla [\tilde{w}_h] - [\tilde{w}_h] Q \nabla \nabla \tilde{u})\, dy.
\end{aligned} \tag{2.3.48}$$

where $[\cdot]$ is the jump across $\partial F(K)$. Suppose that K_1 and K_2 are adjacent triangles in $\mathscr{T}_h$. x_1, x_2, x_3 are vertices of K_1. x_1, x_2, x_4 are vertices of K_2. $W_1 = \tilde{w}_h|_{F(K_1)}$, $W_2 = \tilde{w}_h|_{F(K_2)}$, W_1 and W_2 are piecewise linear on $F(K_1)$ and $F(K_2)$ respectively. Since $\tilde{w}_h$ is continuous on vertices of $F(K)$, let $w_i = \tilde{w}_h(F(x_i))$. Let $F(\overline{x_1 x_2})$ be the image of segment $\overline{x_1 x_2}$ by F, for $y \in F(\overline{x_1 x_2})$, $i = 1, 2$, $j = 1, 2$,

$$W_i(y) - w_j = \nabla W_i (y - y_j). \tag{2.3.49}$$

therefore,

$$W_1(y) - W_2(y) = (\nabla W_1 - \nabla W_2)(y - \lambda y_1 - (1 - \lambda) y_2). \tag{2.3.50}$$

Choose λ such that $|y - \lambda y_1 - (1 - \lambda) y_2| \leq Ch^{\alpha'}$. By (2.3.48) and Assumption (2.2.1),

it results

$$
\begin{aligned}
|a^*[u,w_h]-(w_h,g)| &\le \sum_{K\in\mathscr{T}_h}\Big|\int_{F(K)\ominus K^F}(\nabla\tilde{u}Q\nabla[\tilde{w}_h]-[\tilde{w}_h]Q\nabla\nabla\tilde{u})\Big|\,dy\\
&\le \sum_{K\in\mathscr{T}_h}\int_{F(K)\ominus K^F}|\nabla\tilde{u}Q\nabla[\tilde{w}_h]|+|[\tilde{w}_h]Q\nabla\nabla\tilde{u}|\,dy\\
&\le \sum_{K\in\mathscr{T}_h}\int_{F(K)\ominus K^F}|\nabla\tilde{u}Q\nabla[\tilde{w}_h]|dy\\
&\quad +Ch^{\alpha'}\sum_{K\in\mathscr{T}_h}\int_{F(K)\ominus K^F}|[\nabla\tilde{w}_h]|\cdot|Q\nabla\nabla\tilde{u}|\,dy \qquad (2.3.51)\\
&\le Ch^{\beta}\Big(\underbrace{\sum_{K\in\mathscr{T}_h}\int_{F(K)}|\nabla\tilde{u}Q\nabla\tilde{w}_h|dy}_{I_1}+\\
&\quad \underbrace{Ch^{\alpha'}\sum_{K\in\mathscr{T}_h}\int_{F(K)}|\nabla\tilde{w}_h|\cdot\frac{\tilde{g}}{\det(\nabla F)\circ F^{-1}}dy}_{I_2}\Big).
\end{aligned}
$$

Using Cauchy-Schwarz inequality, we have the following estimate for I_1 and I_2,

$$
\begin{aligned}
I_1=\sum_{K\in\mathscr{T}_h}\int_{F(K)}|^t\nabla\tilde{u}Q(y)\nabla\tilde{w}_h|dy &\le \sum_{K\in\mathscr{T}_h}\int_K|^t\nabla w_h a(x)\nabla u|dx\\
&\le (a^\star[w_h])^{1/2}(a[u])^{1/2} \qquad (2.3.52)\\
&\le C\lambda_{\max}(a)\|w_h\|_h\|g\|_{L^2(\Omega)}.
\end{aligned}
$$

$$
\begin{aligned}
I_2&=\sum_{K\in\mathscr{T}_h}\int_{F(K)}|\nabla\tilde{w}_h|\cdot\frac{\tilde{g}}{\det(\nabla F)\circ F^{-1}}dy\\
&\le \sum_{K\in\mathscr{T}_h}\int_K\nabla w_h(\nabla F\nabla F)^{-1}\nabla w_h g dx\\
&\le \frac{1}{(\lambda_{\min}(a))^{1/2}}\sum_{K\in\mathscr{T}_h}\Big(\int_K\nabla w_h\sigma\nabla w_h\Big)^{1/2}\Big(\int_K g^2dx\Big)^{1/2} \qquad (2.3.53)\\
&\le \frac{1}{(\lambda_{\min}(a))^{1/2}}\Big(\sum_{K\in\mathscr{T}_h}\int_K\nabla w_h\sigma\nabla w_h\Big)^{1/2}\Big(\int_\Omega g^2dx\Big)^{1/2}\\
&\le \Big(\frac{\mu_\sigma}{\lambda_{\min}(a)\|(\mathrm{Trace}\,\sigma)^{-1}\|_{L^\infty(\Omega)}}\Big)^{1/2}\|w_h\|_h\|g\|_{L^2(\Omega)}.
\end{aligned}
$$

□

Theorem 2.2.3 is implied by Theorem 2.3.1, Theorem A.1.2, Lemma 2.3.4, Lemma 2.3.5, (2.3.51), (2.3.52), (2.3.53), and the inequality (2.2.19) for nonconforming finite elements.

2.3.3.2 Proof of Theorem 2.2.4

Let us now prove Theorem 2.2.4. We have

$$a\big[u-\mathscr{J}_h u^f\big] \le 2a\big[u-\mathscr{J}_h u\big]+2a\big[\mathscr{J}_h u-\mathscr{J}_h u^f\big]. \tag{2.3.54}$$

Write $\tilde{\mathscr{J}}_h u := (\mathscr{J}_h u)\circ F^{-1}$. $\tilde{\mathscr{J}}_h u$ is a linear interpolation of $\tilde{u}$ on the tessellation $\mathscr{T}^F$. Using the identity

$$a[u-\mathscr{J}_h u] = Q[\tilde{u}-\tilde{\mathscr{J}}_h u], \tag{2.3.55}$$

we obtain that

$$a[u-\mathscr{J}_h u] \le \|\tilde{u}\|_{C^\alpha}\tilde{h}^\alpha D. \tag{2.3.56}$$

where $\tilde{h}$ is the maximal length of the edges of $\mathscr{T}^F$. Observe that $\tilde{h} \le h^{\alpha'}\|F\|_{C^{\alpha'}}$.

Lemma 2.3.6. *We have*

$$a\big[\mathscr{J}_h u-\mathscr{J}_h u^f\big] \le \mu_\sigma \nu^* a^*\big[\mathscr{Z}^h u-u^f\big]. \tag{2.3.57}$$

Proof. Let $w := \mathscr{J}_h u-\mathscr{J}_h u^f$, to obtain (2.3.57), we observe that

$$a[w] = a^*[\mathscr{J}_h w] = Q[\tilde{w}]. \tag{2.3.58}$$

Thus

$$a[w] = \sum_{K\in\mathscr{T}_h}\int_{K^F} {}^t\nabla\tilde{w}(K^F)Q(y)\nabla\tilde{w}(K^F)\,dy. \tag{2.3.59}$$

It follows that

$$a[w] \le \nu^* \sum_{K\in\mathscr{T}_h}\int_{F(K)} \frac{\lambda_{\max}(Q)}{\lambda_{\min}(Q)}{}^t\nabla\tilde{w}(K^F)Q(y)\nabla\tilde{w}(K^F)\,dy. \tag{2.3.60}$$

from equation

$$Q \circ F = \frac{\sigma}{(\det(\sigma))^{\frac{1}{2}}} (\det(a))^{\frac{1}{2}}, \tag{2.3.61}$$

we obtain that

$$\frac{\lambda_{\max}(Q)}{\lambda_{\min}(Q)} \leq \mu_{\sigma}. \tag{2.3.62}$$

Next, observing that

$$\sum_{K \in \mathscr{T}_h} \int_{F(K)} {}^t\nabla \tilde{w}(K^F) Q(y) \nabla \tilde{w}(K^F)\, dy = a^*[\mathscr{Z}^h w]. \tag{2.3.63}$$

we obtain (2.3.57). □

Theorem 2.2.4 is a consequence of inequalities (2.3.54), (2.3.56), Lemmas 2.3.6 and 2.3.4, Theorem 2.3.1, and the following inequality

$$a^*\left[\mathscr{Z}^h u - u^f\right] \leq a^*\left[u - u^f\right] + a^*\left[u - \mathscr{Z}^h u\right]. \tag{2.3.64}$$

2.3.4 Numerical Homogenization from the Information Point of View

2.3.4.1 Proof of Theorem 2.2.5

In this subsection we will prove Theorem 2.2.5. The method introduced in subsection 2.2.3 can be formulated in the following way: Look for $u^m \in X^h$ such that for all $v \in X^h$,

$$a^*[v, \mathscr{Z}^h u^m] = (v, g)_{L^2\Omega}. \tag{2.3.65}$$

which implies the following finite element orthogonality property for all $v \in X^h$,

$$a^*[v, \mathscr{Z}^h u^m - u] = 0. \tag{2.3.66}$$

Let us write $w = u - \mathscr{Z}^h u^m$. By the inf-sup condition (2.2.40) we obtain that

$$\left(\mathscr{E}_h[w]\right)^{\frac{1}{2}} \leq \frac{1}{\mathscr{S}m} \sup_{v \in X^h} \frac{a^*[v, \mathscr{Z}^h w]}{\left(\mathscr{E}_h[v]\right)^{\frac{1}{2}}}. \tag{2.3.67}$$

By the orthogonality property (2.3.66) we have

$$a^*[v, \mathscr{Z}^h w] = a^*[v, \mathscr{Z}^h u - u]. \tag{2.3.68}$$

Thus

$$a^*[v, \mathscr{Z}^h w] \leq \big(\lambda_{\max}(a)\big)^{\frac{1}{2}} \|\nabla v\|_{L^2(\Omega)} \big(a^*[\mathscr{Z}^h u - u]\big)^{\frac{1}{2}}. \tag{2.3.69}$$

using the inequality (2.2.39)

$$\|\nabla v\|^2_{L^2(\Omega)} \leq \eta_{\max} \mathscr{E}_h[v], \tag{2.3.70}$$

combine (2.3.67), (2.3.69), we deduce that

$$\big(\mathscr{E}_h[w]\big)^{\frac{1}{2}} \leq \frac{1}{\mathscr{S}m} \big(\lambda_{\max}(a)\eta_{\max}\big)^{\frac{1}{2}} \big(a^*[\mathscr{Z}^h u - u]\big)^{\frac{1}{2}}. \tag{2.3.71}$$

It follows from (2.2.39) that

$$\|\nabla \mathscr{I}_h u - \nabla u^m\|_{L^2(\Omega)} \leq \frac{\eta_{\max}}{\mathscr{S}m} \big(\lambda_{\max}(a)\big)^{\frac{1}{2}} \big(a^*[\mathscr{Z}^h u - u]\big)^{\frac{1}{2}}. \tag{2.3.72}$$

And we deduce from Poincaré inequality that

$$\|\mathscr{I}_h u - u^m\|_{H^1_0(\Omega)} \leq C_\Omega \frac{\eta_{\max}}{\mathscr{S}m} \big(\lambda_{\max}(a)\big)^{\frac{1}{2}} \big(a^*[\mathscr{Z}^h u - u]\big)^{\frac{1}{2}}. \tag{2.3.73}$$

We obtain Theorem 2.2.5 from equations (2.3.72), (2.3.73), Lemma 2.3.4, and Theorem 2.3.1.

2.3.4.2 Proof of Theorem 2.2.6

Let us now prove Theorem 2.2.6. Using triangle inequality we obtain

$$a\big[u - \mathscr{J}_h u^m\big] \leq a\big[u - \mathscr{J}_h u\big] + a\big[\mathscr{J}_h u - \mathscr{J}_h u^m\big]. \tag{2.3.74}$$

The object $a\big[u - \mathscr{J}_h u\big]$ has already been bounded from above by (2.3.56). Writing

$w := \mathscr{J}_h u - \mathscr{J}_h u^m$ we have

$$a[w] = \frac{a[\mathscr{J}_h w]}{\mathscr{E}_h[w]} \mathscr{E}_h[w]. \tag{2.3.75}$$

$\mathscr{E}_h[w]$ has already been estimated in equation (2.3.71). It remains to notice that

$$a[\mathscr{J}_h w] = Q[\tilde{w}]. \tag{2.3.76}$$

From this point the arguments are similar to the ones employed in Lemma 2.3.6, indeed

$$Q[\tilde{w}] \leq \mu_\sigma \nu^* \lambda_{\max}(a) \|\nabla w\|^2_{L^2(\Omega)} \leq \mu_\sigma \nu^* \lambda_{\max}(a) \eta_{\max} \mathscr{E}_h[w]. \tag{2.3.77}$$

2.3.5 Numerical Homogenization from a Transport Point of View

2.3.5.1 Proof of Theorem 2.2.7

We assume the mesh to be regular in the following sense: The nodes of the Voronoï diagram of $\mathscr{T}_h$ belong to elements of the primal mesh. In dimension 2 this means that each triangle $K \in \mathscr{T}_h$ is acute. Let us write Y_h the vector space spanned by the functions χ_i. For $v \in Z^h$ we define $\mathscr{Y}_h v$ by

$$\mathscr{Y}_h v := \sum_{i \in \mathscr{N}_h} v_i \chi_i. \tag{2.3.78}$$

The metric based numerical homogenization method can be formulated in the following way: look for $u^v \in Z^h$ (the space spanned by the elements ξ_i p.23) such that for all $i \in \mathscr{N}_h$,

$$a^*[\chi_i, u^v] = (\chi_i, g). \tag{2.3.79}$$

Notice that the finite volume solution u is given by

$$a[\chi_i, u] = (\chi_i, g), \tag{2.3.80}$$

which implies the following finite volume orthogonality property for all $i \in \mathscr{N}_h$,

$$a^*[\chi_i, u^v - u] = 0. \tag{2.3.81}$$

Equation (2.3.79) can be written as

$$\sum_{j\sim i} u_j^v \int_{\partial V_i} n.a.\nabla \xi_i = \int_{V_i} g. \tag{2.3.82}$$

Write $w := \mathscr{Z}^h u - u^v$, by the inf-sup condition (2.2.48) we get

$$(\mathscr{E}_h[w])^{\frac{1}{2}} \leq \frac{1}{\mathscr{S}^v} \sup_{v\in\mathscr{Y}_h\setminus\{0\}} \frac{a^*[v,w]}{(\mathscr{E}_h[v])^{\frac{1}{2}}}. \tag{2.3.83}$$

Using the orthogonality property of the finite volume method we obtain that for $v \in \mathscr{Y}_h$,

$$a^*[v,w] = -\sum_{i\in\mathscr{N}_h} v_i \int_{\partial V_i} n.a\big(\nabla \mathscr{Z}^h u - \nabla u\big). \tag{2.3.84}$$

Let $\mathscr{E}_h^*$ be the edges of the dual tessellation (edges of the control volumes), we obtain that

$$a^*[v,w] = \sum_{e_{ij}\in\mathscr{E}_h^*} (v_j - v_i) \int_{e_{ij}} n_{ij}.a\big(\nabla \mathscr{Z}^h u - \nabla u\big). \tag{2.3.85}$$

where e_{ij} is the edge separating the control volume V_i from the control volume V_j and n_{ij} is the unit vector orthogonal to e_{ij} pointing outside of V_i. It follows that

$$a^*[v,w] \leq \big(\mathscr{E}_h[v]\big)^{\frac{1}{2}} \|\nabla_F \mathscr{Z}^h u - \nabla_F u\|_{L^\infty(\mathscr{E}_h^*)} \Big(\sum_{e_{ij}\in\mathscr{E}_h^*} |e_{ij}|^2 \lambda_{\max}(a)\lambda_{\max}(\sigma)\Big)^{\frac{1}{2}}. \tag{2.3.86}$$

Note that

$$\lambda_{\max}(\sigma) \leq \frac{\mu_\sigma}{2\|(\mathrm{Trace}(\sigma))^{-1}\|_{L^\infty}}, \tag{2.3.87}$$

and

$$\sum_{e_{ij}\in\mathscr{E}_h^*} |e_{ij}|^2 \leq 4\eta_{\max}\,\mathrm{Vol}(\Omega). \tag{2.3.88}$$

It follows from equation (2.3.83),

$$\big(\mathscr{E}_h[w]\big)^{\frac{1}{2}} \leq \frac{4}{\mathscr{S}^v} \|\nabla_F \mathscr{Z}^h u - \nabla_F u\|_{L^\infty(\mathscr{E}_h^*)} \eta_{\max}\,\mathrm{Vol}(\Omega)\big(\lambda_{\max}(a)\lambda_{\max}(\sigma)\big)^{\frac{1}{2}}. \tag{2.3.89}$$

Equation (2.2.49) of Theorem 2.2.7 is then a straightforward consequence of (2.2.39)

and Lemma 2.3.3.

Let us now proceed to prove (2.2.50) of Theorem 2.2.7. By triangle inequality

$$a\big[u-\mathscr{J}_h u^\nu\big] \le a\big[u-\mathscr{J}_h u\big] + a\big[\mathscr{J}_h u-\mathscr{J}_h u^\nu\big]. \tag{2.3.90}$$

$a\big[u-\mathscr{J}_h u\big]$ has already been estimated in equation (2.3.56). Writing $w := \mathscr{J}_h u - \mathscr{J}_h u^\nu$ we have

$$a[w] = \frac{a^*[\mathscr{J}_h w]}{\mathscr{E}_h[w]}\mathscr{E}_h[w]. \tag{2.3.91}$$

But $\mathscr{E}_h[w]$ has already been estimated in equation (2.3.89). It remains to estimate $\frac{a^*[\mathscr{J}_h w]}{\mathscr{E}_h[w]}$. Similar to the argument in Lemma 2.3.6, we have

$$a^*[\mathscr{J}_h w] \le \mu_\sigma \nu^* \lambda_{\max}(a)\eta_{\max}\mathscr{E}_h[w]. \tag{2.3.92}$$

which concludes the proof of Theorem 2.2.7.

2.4 Numerical Experiments

Let us now illustrate the implementation of the numerical method. The computational domain is the unit disk in dimension two. Equation (2.0.1) is solved on a fine tessellation characterized by 66049 nodes and 131072 triangles. The coarse tessellation has 289 nodes and 512 triangles (figure 2.9). Since our methods involve the computation of global harmonic coordinates, the memory and CPU time requirements are not improved if one needs to solve (2.0.1) only one time, whereas localized methods such as the one of Hou and Wu or E. and Engquist do improve the memory requirement or the CPU time. We will show, as a trade-off, the accuracy of the numerical homogenization will be sharply improved.

The elliptic operator associated to equation (2.0.1) has been up-scaled to an operator defined on the coarse mesh (compression by a factor of ~ 300) using 5 different methods:

- FEM_ψ: The Galerkin scheme described in subsection 2.2.1 using the multi-scale finite element shape function ψ_i.

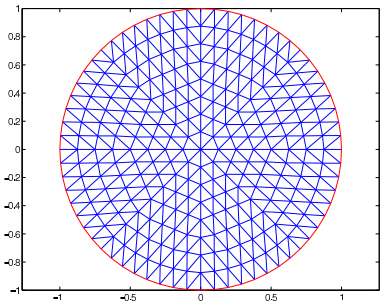
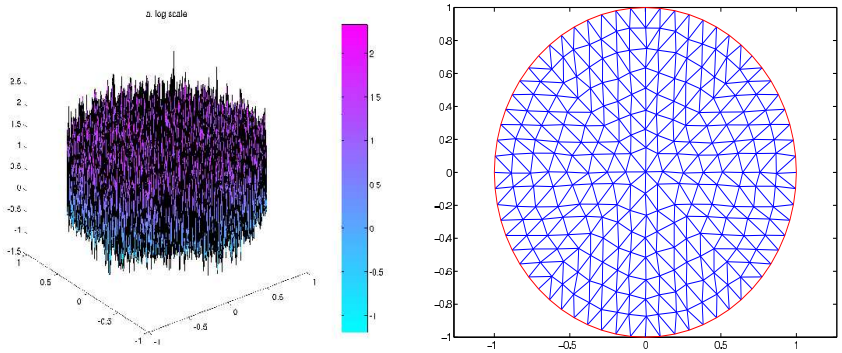


Figure 2.9: Coarse grid

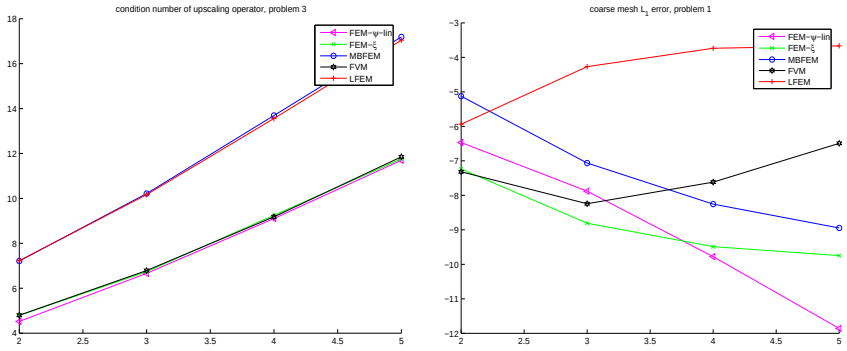
- FEM_ξ: The Galerkin scheme described in subsection 2.2.2 using the nonconforming elements ξ_i.
- MBFEM: The metric based compression scheme described in subsection 2.2.3.
- FVM: The finite volume method described in subsection 2.2.4.
- LFEM: A multi-scale finite element where F is computed locally [1] on each triangle K of the coarse mesh as the solution of a cell problem with boundary condition $F(x) = x$ on ∂K. This method has been implemented in order to understand the effect of the removal of global information in the structure of the metric induced by F.



(a) a **(b)** $\mathscr{T}^F$

Figure 2.10: Example 2.4.1, Trigonometric multi-scale

[1] instead of globally



(a) *Condition number* **(b) *Coarse mesh* L^1 *error***

Figure 2.11: Example 2.4.1, Trigonometric multi-scale

Example 2.4.1. *Trigonometric multi-scale*

The following example is extracted from [94] as a problem without scale separation:

$$a(x) = \frac{1}{6}\Big(\frac{1.1+\sin(2\pi x/\varepsilon_1)}{1.1+\sin(2\pi y/\varepsilon_1)} + \frac{1.1+\sin(2\pi y/\varepsilon_2)}{1.1+\cos(2\pi x/\varepsilon_2)} + \frac{1.1+\cos(2\pi x/\varepsilon_3)}{1.1+\sin(2\pi y/\varepsilon_3)} + \frac{1.1+\sin(2\pi y/\varepsilon_4)}{1.1+\cos(2\pi x/\varepsilon_4)} + \frac{1.1+\cos(2\pi x/\varepsilon_5)}{1.1+\sin(2\pi y/\varepsilon_5)} + \sin(4x^2y^2) + 1\Big) \tag{2.4.1}$$

where $\varepsilon_1 = \frac{1}{5}, \varepsilon_2 = \frac{1}{13}, \varepsilon_3 = \frac{1}{17}, \varepsilon_4 = \frac{1}{31}, \varepsilon_5 = \frac{1}{65}$.

Figure 2.10 demonstrates the deformation of $\mathscr{T}^F$ (figure 2.9) induced by F. The deformation is small since the medium is quasi-periodic and F is close to x. The weak aspect ratio induced by F is $\eta^*_{\text{min}} = 1.1252$. Table 2.1 gives the relative error estimated on the nodes of the coarse mesh between the solution u of the PDE (2.0.1) and the solutions obtained using the up-scaled operators. Table 2.2 gives the relative error estimated on the nodes of the fine mesh between u and the fine mesh approximation of the coarse mesh solutions. Figure 2.11(a) gives the condition number of the stiffness matrix associated to the up-scaled operator versus $-\log_2 h$ (logarithm of the resolution). Figure 2.11(b) gives the relative L_1-distance between u and its approximation on the coarse mesh in log scale versus $-\log_2 h$. Obviously FEM_ψ has the best performance. For the method LFEM, numerical error increases with the resolution, this is an effect of the cell resonance observed in [79] and [8]. This cell resonance does not occur with the methods proposed in this paper. The

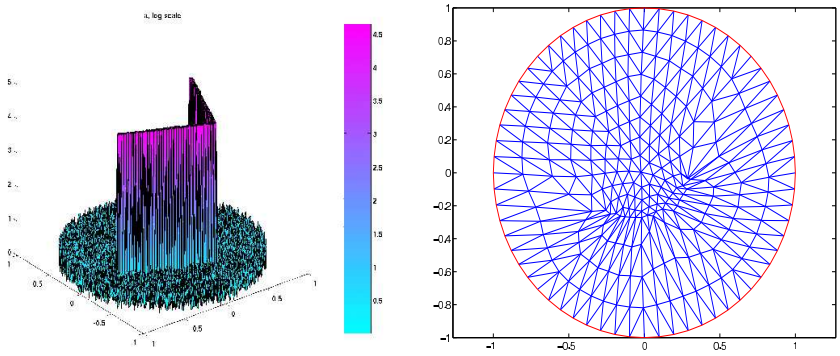
Coarse Mesh Error	FEM_ψ	FEM_ξ	MBFEM	FVM	LFEM
L^1	0.0042	0.0022	0.0075	0.0032	0.0411
L^2	0.0039	0.0024	0.0074	0.0040	0.0441
L^∞	0.0059	0.0090	0.0154	0.0117	0.0496
H^1	0.0060	0.0262	0.0568	0.0203	0.0763

Table 2.1: Example 2.4.1, Trigonometric multi-scale

Fine mesh Error	FEM_ψ	FEM_ξ	MBFEM	FVM	LFEM
L^1	0.0042	0.0085	0.0053	0.0080	0.0593
L^2	0.0043	0.0082	0.0061	0.0078	0.0591
L^∞	0.0063	0.0112	0.0154	0.0141	0.0597
H^1	0.0581	0.0540	0.0778	0.0601	0.0943

Table 2.2: Example 2.4.1, Trigonometric multi-scale

finite volume method is characterized by the the best stability and one of the best accuracies at a coarse resolution. The increase in the error observed for FVM as the resolution is increased is a numerical artifact created by the fine mesh: one has to divide the coarse tessellation into coarse control volumes. These coarse control volumes are unions of the control volumes defined on a fine mesh, and when the ratio between the coarse resolution and the fine mesh resolution is small and the triangulation is irregular, it is not possible to divide the coarse tessellation into control volumes intersecting the edges of the primal mesh at positions close to the midpoints of those edges.



(a) a. **(b)** $\mathscr{T}^F$

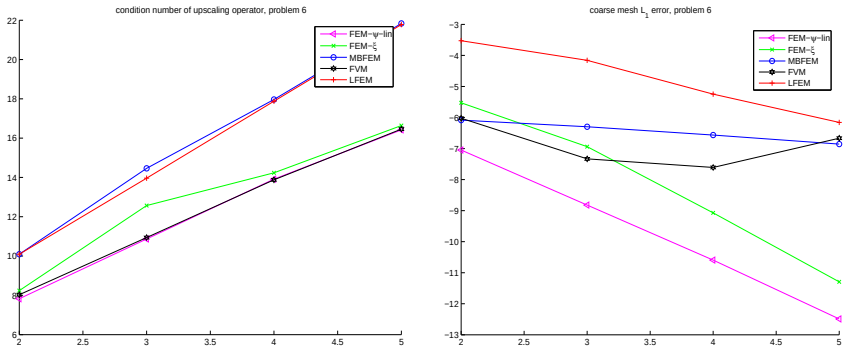
Figure 2.12: Example 2.4.2, High conductivity channel

Example 2.4.2. *High conductivity channel*

In this example a is characterized by a fine and long-ranged high conductivity channel. We choose $a(x) = 100$, if x is in the channel, and $a(x) = O(1)$, if x is not in the channel. The weak aspect ratio induced by F is $\eta^{\star}_{\min} = 2.2630$. μ_σ is 1.03×10^6, and the aspect ratio of a is 103. The condition number of the linear system associated with linear elements on the fine mesh is 3.29×10^8. The condition number of the linear system associated with LFEM (*dof* 225) is 1.48×10^4, but the condition number of the linear system associated with FEM_ψ is merely 621. We can see that the equation is better conditioned in the harmonic coordinates, and the theoretical error estimate is in fact not optimal.

Table 2.3 gives the relative error estimated on the nodes of the coarse mesh between the solution u of the PDE (2.0.1) and an approximation obtained from the up-scaled operators. Table 2.4 gives the relative error estimated on the nodes of the fine mesh between u and the fine mesh interpolation of the coarse mesh solutions. Figure 2.13(a) gives the condition number of the stiffness matrix associated to the up-scaled operator versus $-\log_2 h$ (logarithm of the resolution). Figure 2.13(b) gives the relative L_1-distance between u and its approximation on the coarse mesh in log scale versus $-\log_2 h$.

Observe in figure 2.12 that the effect of the new metric on the mesh is to bring close together nodes linked by a path of low electrical resistance.



(a) *Condition number* **(b) *Coarse mesh* L^1 *error***

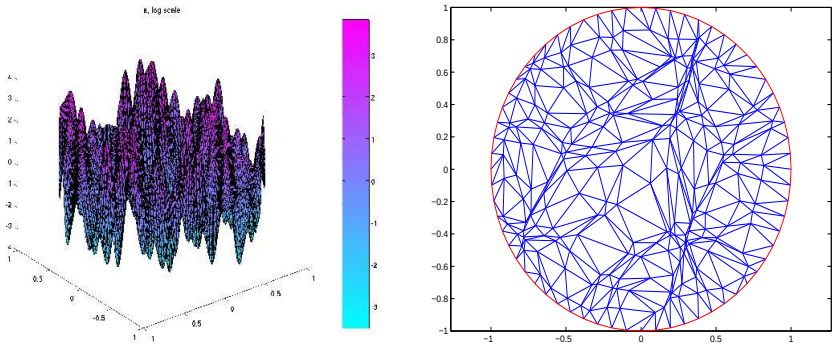
Figure 2.13: Example 2.4.2, High conductivity channel

Coarse Mesh Error	FEM_ψ	FEM_ξ	MBFEM	FVM	LFEM
L^1	0.0022	0.0081	0.0127	0.0062	0.0519
L^2	0.0025	0.0096	0.0179	0.0081	0.0606
L^∞	0.0120	0.0227	0.0549	0.0174	0.1223
H^1	0.0120	0.0384	0.0919	0.0265	0.1514

Table 2.3: Example 2.4.2, High conductivity channel

Fine mesh Error	FEM_ψ	FEM_ξ	MBFEM	FVM	LFEM
L^1	0.0070	0.0155	0.0164	0.0121	0.0612
L^2	0.0069	0.0153	0.0202	0.0123	0.0743
L^∞	0.0133	0.0227	0.0573	0.0214	0.1226
H^1	0.0760	0.1032	0.1838	0.0820	0.2142

Table 2.4: Example 2.4.2, High conductivity channel



(a) a. **(b)** $\mathscr{T}^F$

Figure 2.14: Example 2.4.3, Random Fourier modes

Example 2.4.3. *Random Fourier modes.*

In this case, $a(x) = e^{h(x)}$, with

$$h(x) = \sum_{|k|\leq R} (a_k \sin(2\pi k \cdot x) + b_k \cos(2\pi k \cdot x)) \tag{2.4.2}$$

where a_k and b_k are independent identically distributed random variables on $[-0.3, 0.3]$ and $R = 6$. This is another example where scales are not separated. The weak aspect ratio induced by F is $\eta^{\star}_{\min} = 3.4997$. The deformation induced by F is given in figure 2.14. Errors between u and the fine mesh interpolation of the coarse mesh solutions are larger (tables 2.5 and 2.6), which is due to the fact that those errors depend on the aspect ratio $\eta^{*}_{\max}$. Of course one could improve the compression by adapting the mesh to the new metric but this has not been our point of view here. We have preferred to show raw data obtained with a given coarse mesh. The figures 2.16 and 2.17 give the log plot of L^1, L^2, L^∞ and H^1 relative error. The x-axis corresponds to the refinement of coarse mesh, the y-axis is the error. The tables 2.7 and 2.8 show the convergence rate in different norms (the parameter α in the error of the order of h^α).

Coarse Mesh Error	FEM_ψ	FEM_ξ	MBFEM	FVM	LFEM
L^1	0.0027	0.0075	0.0117	0.0106	0.1197
L^2	0.0028	0.0087	0.0130	0.0125	0.1169
L^∞	0.0066	0.0278	0.0320	0.0376	0.1358
H^1	0.0133	0.0648	0.0805	0.0597	0.1514

Table 2.5: Example 2.4.3, Random Fourier Modes

Fine mesh Error	FEM_ψ	FEM_ξ	MBFEM	FVM	LFEM
L^1	0.0112	0.0148	0.0148	0.0188	0.1304
L^2	0.0177	0.0223	0.0184	0.0202	0.1265
L^∞	0.0773	0.0824	0.0614	0.0680	0.1669
H^1	0.0972	0.1152	0.1307	0.1659	0.1725

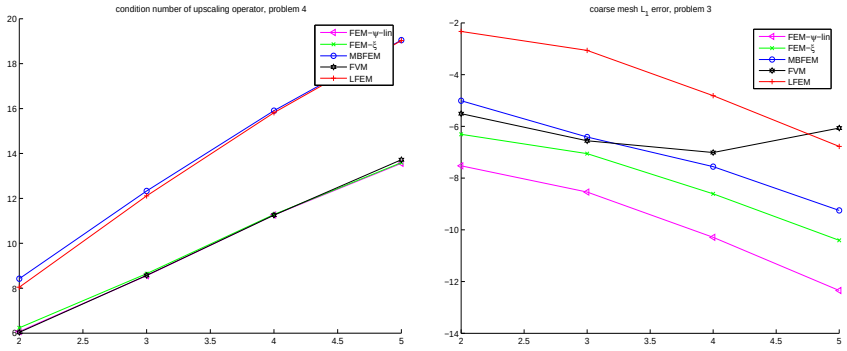
Table 2.6: Example 2.4.3, Random Fourier Modes

Method	L^1	L^2	L^∞	H^1
FEM_ψ	1.62	1.66	1.56	1.44
FEM_ξ	1.38	1.27	1.23	1.18
MBFEM	1.38	1.40	1.27	1.08
FVM	0.53	1.14	1.26	1.03
LFEM	1.51	1.53	1.62	1.46

Table 2.7: Coarse mesh approximation convergence rate

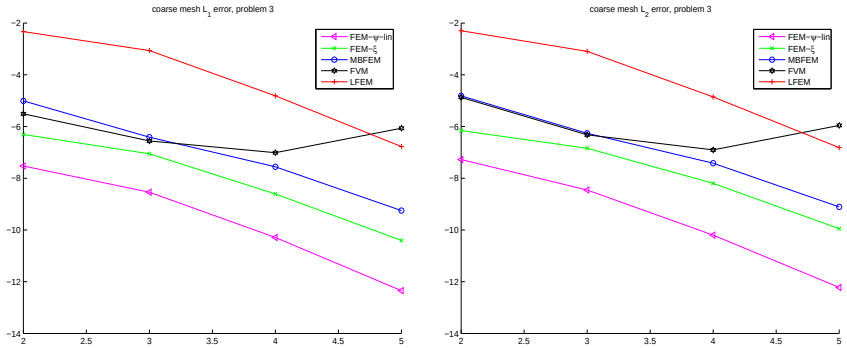
Method	L^1	L^2	L^∞	H^1
FEM_ψ	1.74	1.61	1.23	0.89
FEM_ξ	1.57	1.47	1.23	0.91
MBFEM	1.54	1.52	1.21	0.96
FVM	0.75	1.16	1.22	0.58
LFEM	1.52	1.54	1.42	1.10

Table 2.8: Fine mesh approximation convergence rate

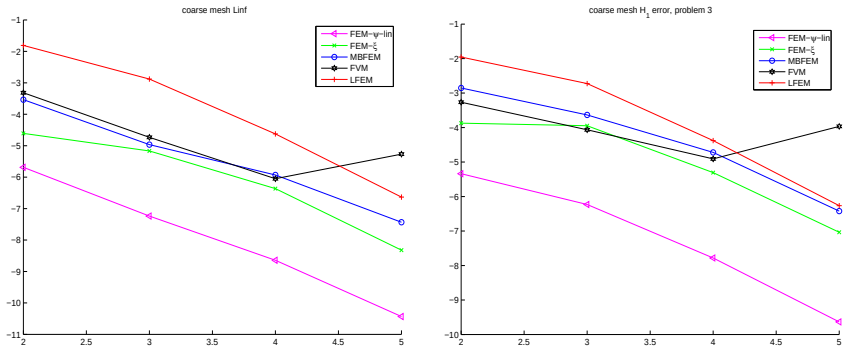


(a) *Condition number* **(b) *Coarse mesh* L^1 *error***

Figure 2.15: Example 2.4.3, Random Fourier modes

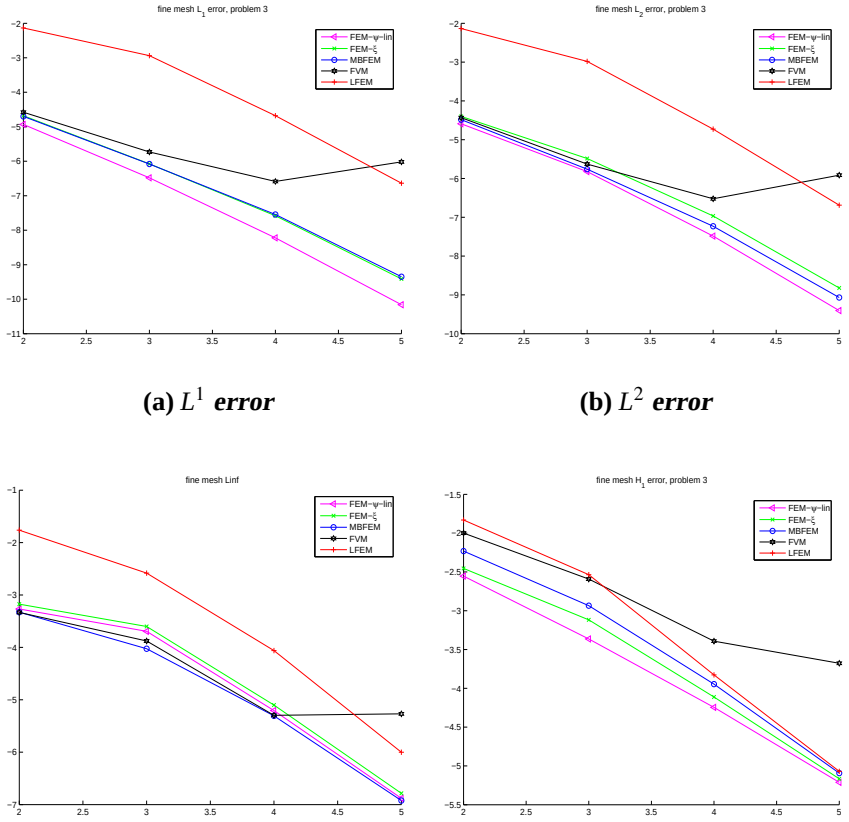


(a) L^1 ***error*** **(b)** L^2 ***error***



(c) L^∞ ***error*** **(d)** H^1 ***error***

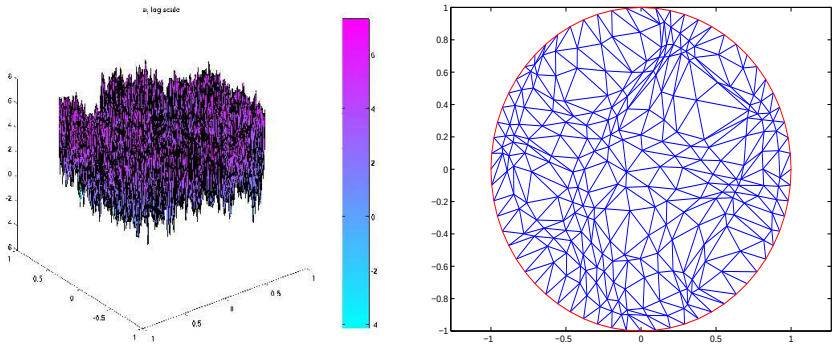
Figure 2.16: Coarse mesh error ($\log_2$) L^1, L^2, L^∞ and H^1 errors vs. coarse mesh refinement, Example 2.4.3, Random Fourier modes



(a) L^1 ***error*** **(b)** L^2 ***error***

(c) L^∞ ***error*** **(d)** H^1 ***error***

Figure 2.17: Fine mesh approximation error ($\log_2$) L^1, L^2, L^∞ and H^1 errors vs. coarse mesh refinement, Example 2.4.3, Random Fourier modes



(a) a **(b)** $\mathscr{T}^F$

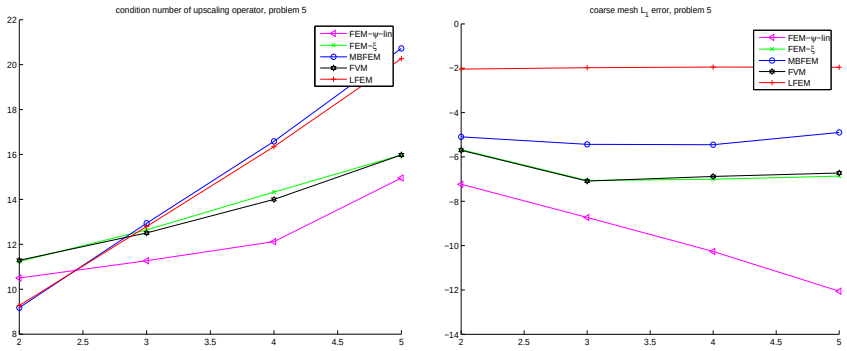
Figure 2.18: Example 2.4.4, Random fractal

Coarse Mesh Error	FEM_ψ	FEM_ξ	MBFEM	FVM	LFEM
L^1	0.0024	0.0075	0.0231	0.0073	0.0519
L^2	0.0025	0.0085	0.0241	0.0100	0.0606
L^∞	0.0094	0.0399	0.0920	0.0398	0.1694
H^1	0.0161	0.0718	0.1553	0.0493	0.3107

Table 2.9: Example 2.4.4, Random fractal

Example 2.4.4. *Random fractal*

In this case, a is given by a product of discontinuous functions oscillating randomly at different scales, $a(x) = a_1(x)a_2(x)\cdots a_n(x)$, and $a_i(x) = c_{pq}$ for $x \in [\frac{p}{2^i}, \frac{p+1}{2^i}) \times [\frac{q}{2^i}, \frac{q+1}{2^i})$, c_{pq} is uniformly random in $[\frac{1}{\gamma}, \gamma]$, $n = 5$ and $\gamma = 2$. The weak aspect ratio induced by F is $\eta^\star_{\min} = 2.4796$. Table 2.9 gives the relative error estimated on the nodes of the coarse mesh between the solution u of the PDE (2.0.1) and an approximation obtained from the up-scaled operator. Table 2.10 gives the relative error of the fine mesh approximation of the coarse mesh solutions. Figure 2.19(a) illustrates the condition number of the stiffness matrix associated to the up-scaled operator versus $-\log_2 h$ (logarithm of the resolution). Figure 2.19(b) gives the relative L_1-distance between u and its approximation on the coarse mesh in log scale versus $-\log_2 h$.

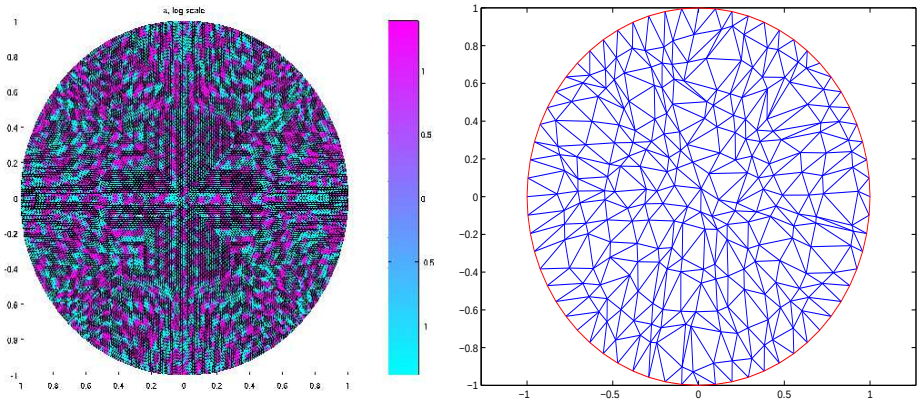


(a) *Condition number* **(b) *Coarse mesh* L^1 *error***

Figure 2.19: Example 2.4.4, Random fractal

Fine mesh Error	FEM_ψ	FEM_ξ	MBFEM	FVM	LFEM
L^1	0.0108	0.0147	0.0245	0.0142	0.0765
L^2	0.0155	0.0198	0.0280	0.0173	0.0812
L^∞	0.0662	0.0802	0.0919	0.0720	0.1694
H^1	0.1015	0.1231	0.1838	0.1433	0.2642

Table 2.10: Example 2.4.4, Random fractal



(a) a **(b)** $\mathscr{T}^F$

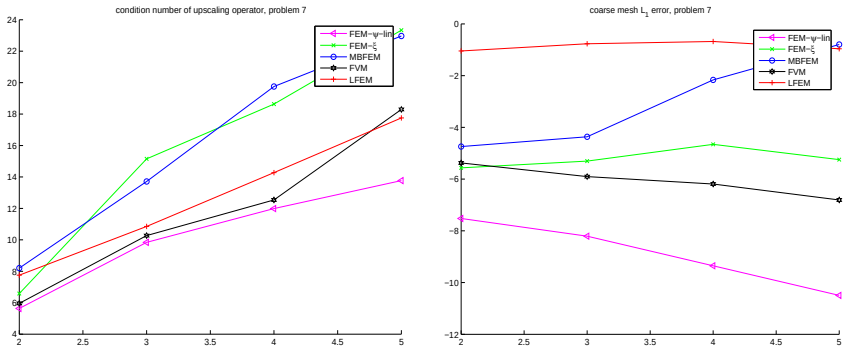
Figure 2.20: Example 2.4.5, Percolation.

Coarse Mesh Error	FEM_ψ	FEM_ξ	MBFEM	FVM	LFEM
L^1	0.0034	0.0253	0.0485	0.0167	0.2848
L^2	0.0041	0.0265	0.0523	0.0189	0.2851
L^∞	0.0163	0.0813	0.0643	0.0499	0.3018
H^1	0.0343	0.0843	0.1070	0.0713	0.3740

Table 2.11: Example 2.4.5, Percolation.

Example 2.4.5. *Site percolation*

In this case, the conductivity of each site is equal to γ or $1/\gamma$ with probability $1/2$. We have chosen $\gamma = 4$ in this example. Observe that some errors are larger for this challenging case because a percolating medium generates flat triangles in the new metric — indeed $\eta^\star_{\min} = 22.3395$ which is much larger than previous examples. Table 2.11 gives the relative error of the coarse mesh approximations of the PDE (2.0.1). Table 2.12 gives the relative error of the fine mesh $\mathscr{J}_h$ interpolation of the coarse mesh solutions. Figure 2.21(a) demonstrates the condition number of the stiffness matrix associated to the up-scaled operator versus $-\log_2 h$. Figure 2.21(b) gives the relative L_1-error of the coarse mesh solution in log scale versus $-\log_2 h$.



(a) *Condition number* **(b) *Coarse mesh* L^1 *error***

Figure 2.21: Example 2.4.5, Percolation

Fine mesh Error	FEM_ψ	FEM_ξ	MBFEM	FVM	LFEM
L^1	0.0115	0.0265	0.0585	0.0216	0.3024
L^2	0.0152	0.0268	0.0628	0.0229	0.3015
L^∞	0.0500	0.0527	0.0940	0.0497	0.3135
H^1	0.1000	0.1712	0.1954	0.1343	0.3964

Table 2.12: Example 2.4.5, Percolation

2.4.1 Numerical Experiments with Splines

We have seen that if σ is stable then $u \circ F^{-1}$ belongs to $W^{2,p}(\Omega)$ with some $p > 2$. It is thus natural to expect a better accuracy by using C^1 elements as φ in the method $\psi = \varphi \circ F$ instead of C^0 elements. In [8], the use of higher order Lagrangian polynomials as coarse mesh shape functions was suggested and the increase of accuracy has been observed when F is obtained as the solution of a local cell problem. In our case, when the harmonic coordinates are computed globally, we would alter the finite element method $\psi = \varphi \circ F$ by using C^1 elements for the φ.

We refer to Appendix C and references therein for C^1 finite element methods. One possibility is to use weighted extended B-splines (WEB) method developed by K. Höllig in [76, 77], these elements are in general C^1-continuous. They are obtained from tensor products of one dimensional elements. The Dirichlet boundary condition is satisfied using a smooth weight function ω, such that $\omega = 0$ at the boundary. The condition number of the stiffness matrix is bounded from above by $O(h^{-2})$ (we have the same optimal bound on a Galerkin system with piecewise linear elements).

We have considered two challenging multi-scale medium for our numerical experiments: random Fourier modes and percolation. For the simplicity of the implementation a square domain has been considered, and weighted spline basis are used instead of the WEB spline basis. For a square domain $[-1,1] \times [-1,1]$, the weight function is chosen as $\omega = (1-x^2)(1-y^2)$. Two methods have been compared,

- FEM_ψ_{lin}: The Galerkin scheme using the finite elements $\psi_i = \varphi_i \circ F$, where φ_i are the piecewise linear nodal basis elements.

- FEM_ψ_{sp}: The Galerkin scheme using the finite element $\psi_i = \varphi_i \circ F$, where φ_i are weighted cubic B-spline elements.

As shown in table 2.13–2.16 and figures 2.22–2.23, a sharp increase in accuracy is observed for the method FEM_ψ_{sp}.

Coarse Mesh Error	FEM_ψ_{lin}	FEM_ψ_{sp}
L^1	0.0437	0.0046
L^2	0.0426	0.0052
L^∞	0.0614	0.0096
H^1	0.0746	0.0227

Table 2.13: Example 2.4.3, Random Fourier modes

Fine mesh Error	FEM_ψ_{lin}	FEM_ψ_{sp}
L^1	0.0546	0.0077
L^2	0.0529	0.0096
L^∞	0.0920	0.0289
H^1	0.2109	0.0547

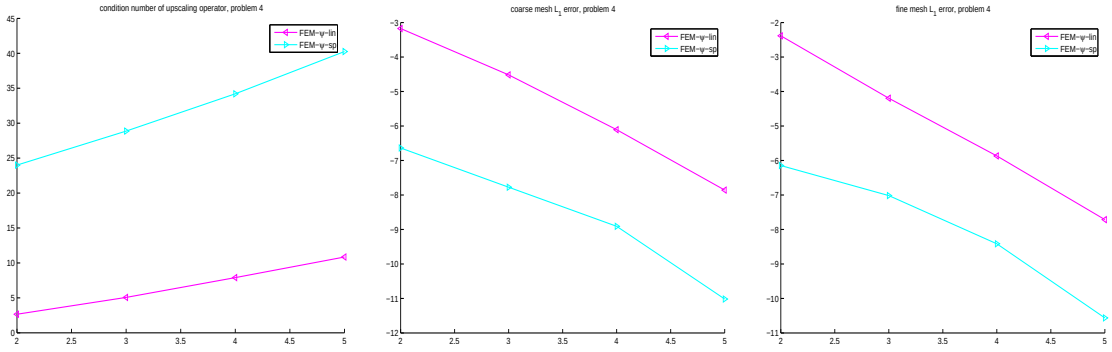
Table 2.14: Example 2.4.3, Random Fourier modes

Coarse Mesh Error	FEM_ψ_{lin}	FEM_ψ_{sp}
L^1	0.0393	0.0080
L^2	0.0379	0.0098
L^∞	0.0622	0.0309
H^1	0.0731	0.0404

Table 2.15: Example 2.4.5, Percolation

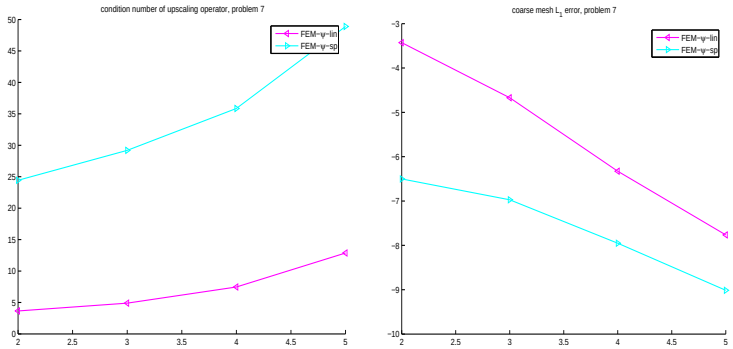
Fine mesh Error	FEM_ψ_{lin}	FEM_ψ_{sp}
L^1	0.0470	0.0099
L^2	0.0464	0.0130
L^∞	0.1174	0.0554
H^1	0.2030	0.0838

Table 2.16: Example 2.4.5, Percolation

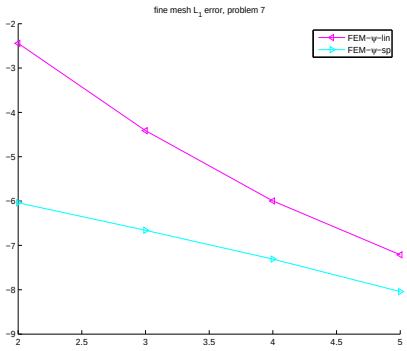


(a) *Condition number* **(b) *Coarse mesh* L^1 *error*** **(c) *Fine mesh* L^1 *error***

Figure 2.22: Example 2.4.3. Random Fourier modes.



(a) *Condition number.* **(b) *Coarse mesh* L^1 *error***



(c) *Fine mesh* L^1 *error*

Figure 2.23: Example 2.4.5. Percolation

Chapter 3

Numerical Homogenization for Parabolic Equations with Continuum Time and Space Scales

In this chapter, we address the issue of the homogenization of divergence form parabolic equations in space and time in situations where scale separation and ergodicity at small scales are not available. In section 3.1 we give the main results, i.e., the compensation result and error estimate of numerical homogenization scheme. In subsection 3.1.1 we will establish the increase of regularity of solutions of (3.1.1), show that under a (parabolic) Cordes type condition the first order time derivatives and second order space derivatives of the solution with respect to caloric coordinates are in L^2 instead of H^{-1} in Euclidean coordinates. In subsection 3.1.2 we will formulate the corresponding semidiscrete numerical homogenization method. In section 3.1.3 we will state the formulation and error estimate of the corresponding full discrete numerical homogenization method. In section 3.2 we give the detailed proof of the regularity result and error estimate. In section 3.3 we show the results of numerical experiments.

3.1 Main Results

Let Ω be a bounded and convex domain of class C^2 of $\mathbb{R}^n$. Let $T > 0$. Consider the following parabolic PDE

$$\begin{cases} \partial_t u = \operatorname{div}\big(a(x,t)\nabla u(x,t)\big) + g \quad \text{in} \quad \Omega \times (0,T) \\ u(x,t) = 0 \quad \text{for} \quad (x,t) \in \big(\partial\Omega \times (0,T)\big) \cup \big(\Omega \cup \{t=0\}\big). \end{cases} \tag{3.1.1}$$

Write $\Omega_T := \Omega \times (0,T)$. g is a function in $L^2(\Omega_T)$. $a(x,t)$ is a symmetric positive definite $n \times n$ matrix with entries in $L^\infty(\Omega_T)$. We assume a to be uniformly elliptic on the closure of Ω_T.

3.1.1 Compensation Phenomena

Let F be the solution of the following parabolic equation

$$\begin{cases} \partial_t F = \operatorname{div}\big(a(x,t)\nabla F(x,t)\big) \quad \text{in} \quad \Omega_T \\ F(x,t) = x \quad \text{for} \quad (x,t) \in \big(\partial\Omega \times (0,T)\big) \\ \operatorname{div}\big(a(x,0)\nabla F(x,0)\big) = 0 \quad \text{in} \quad \Omega. \end{cases} \tag{3.1.2}$$

By (3.1.2) we mean that $F := (F_1, \ldots, F_n)$ is a n-dimensional vector field such that each of its entries satisfies

$$\begin{cases} \partial_t F_i = \operatorname{div}\big(a(x,t)\nabla F_i(x,t)\big) \quad \text{in} \quad \Omega_T \\ F_i(x,t) = x_i \quad \text{for} \quad (x,t) \in \big(\partial\Omega \times (0,T)\big) \\ \operatorname{div}\big(a(x,0)\nabla F_i(x,0)\big) = 0 \quad \text{in} \quad \Omega. \end{cases} \tag{3.1.3}$$

We call F the caloric coordinates associated with a. In the case of time independent medium, F is the harmonic coordinates defined in (2.1.1). Similar as in Chapter 2, we write

$$\sigma := {}^t\nabla F a \nabla F. \tag{3.1.4}$$

3.1.1.1 Time independent medium

In this subsection we assume that a does not depend on time t. Recall that β_σ, the Cordes parameter associated to σ is defined by

$$\beta_\sigma := \operatorname{esssup}_{x\in\Omega}\Big(n - \frac{\big(\operatorname{Trace}[\sigma]\big)^2}{\operatorname{Trace}[{}^t\sigma\sigma]}\Big), \tag{3.1.5}$$

and

$$\mu_\sigma := \operatorname{esssup}_{\Omega_T}\frac{\lambda_{\max}(\sigma)}{\lambda_{\min}(\sigma)}. \tag{3.1.6}$$

Write for $p \geq 2$, $W_D^{2,p}$ (D for Dirichlet boundary condition) the Banach space $W^{2,p}(\Omega)\cap W_0^{1,p}(\Omega)$. Equip $W_D^{2,p}(\Omega)$ with the norm

$$\|v\|^p_{W_D^{2,p}(\Omega)} := \int_\Omega \big(\sum_{i,j}(\partial_i\partial_j v)^2\big)^{\frac{p}{2}}\,dx. \tag{3.1.7}$$

Equip the space $L^p(0,T,W_D^{2,p}(\Omega))$ with the norm

$$\|v\|^p_{L^p(0,T,W_D^{2,p}(\Omega))} = \int_0^T\int_\Omega \big(\sum_{i,j}(\partial_i\partial_j v)^2\big)^{\frac{p}{2}}\,dx\,dt. \tag{3.1.8}$$

Theorem 3.1.1. *Assume that $\partial_t a \equiv 0$, $g \in L^2(\Omega_T)$, Ω is convex, $\beta_\sigma < 1$, and $(\operatorname{Trace}[\sigma])^{\frac{n}{4}-1} \in L^\infty(\Omega)$ then $u\circ F^{-1} \in L^2(0,T,W_D^{2,2}(\Omega))$ and*

$$\|u\circ F^{-1}\|_{L^2(0,T,W_D^{2,2}(\Omega))} \leq \frac{C}{1-\beta_\sigma^{\frac{1}{2}}}\|g\|_{L^2(\Omega_T)}. \tag{3.1.9}$$

The constant C can be written as

$$C = \frac{C_n}{(\lambda_{\min}(a))^{\frac{n}{4}}}\big\|(\operatorname{Trace}[\sigma])^{\frac{n}{4}-1}\big\|_{L^\infty(\Omega)}. \tag{3.1.10}$$

Remark. *According to theorem 3.1.1, although the second order derivatives of u with respect to Euclidean coordinates are only in $L^2(0,T,H^{-1}(\Omega))$, they are in $L^2(\Omega_T)$ with respect to harmonic coordinates.*

The compensation phenomena presented in this subsection can be observed numerically. In figure 3.1, the value of a is set to be equal to 1 or 100 with probability $1/2$ on each triangle of a fine mesh with 16641 nodes and 32768 triangles.

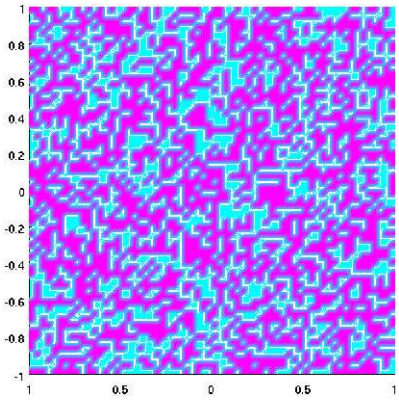
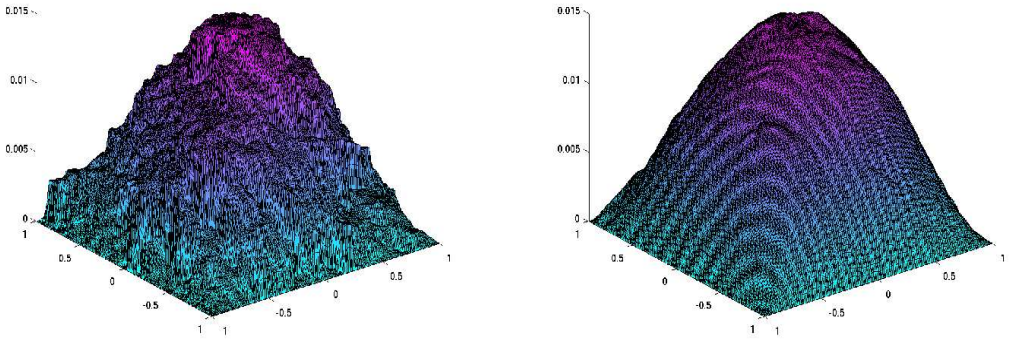
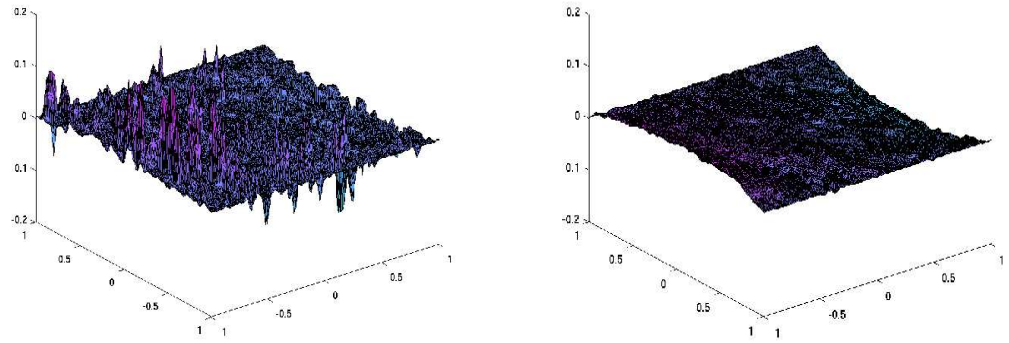


Figure 3.1: Site percolation

(3.1.1) has been solved numerically on that mesh with $g \equiv 1$. u, $u \circ F^{-1}$, $\partial_x u$ and $\partial x(u \circ F^{-1})$ have been plotted at time $t = 1$ in figure 3.2.



(a) u **(b)** $u \circ F^{-1}$



(c) $\partial_x u$ **(d)** $\partial x(u \circ F^{-1})$

Figure 3.2: u, $u \circ F^{-1}$, $\partial_x u$, and $\partial x(u \circ F^{-1})$ at time $t = 1$ for the time independent site percolating medium

In situations where $g \in L^\infty(0,T,L^2(\Omega))$, $\partial_t g \in L^2(0,T,H^{-1}(\Omega))$, or $g \in L^p(\Omega_T)$ with $p > 2$, one can obtain a higher regularity for $u \circ F^{-1}$. This is the object of the following theorems.

Theorem 3.1.2. *Assume that Ω is convex, $g \in L^\infty(0,T,L^2(\Omega))$, $\partial_t g \in L^2(0,T,H^{-1}(\Omega))$, $\partial_t a \equiv 0$, $\beta_\sigma < 1$ and $(\mathrm{Trace}[\sigma])^{\frac{n}{4}-1} \in L^\infty(\Omega_T)$, then for all $t \in [0,T]$, $u \circ F^{-1}(.,t) \in W_D^{2,2}(\Omega)$ and*

$$\|u \circ F^{-1}(.,t)\|_{W_D^{2,2}(\Omega)} \leq \frac{C}{1-\beta_\sigma^{\frac{1}{2}}} \Big(\|g\|_{L^\infty(0,T,L^2(\Omega))} + \|\partial_t g\|_{L^2(0,T,H^{-1}(\Omega))} \Big). \tag{3.1.11}$$

Theorem 3.1.3. *Assume that Ω is convex, $g(.,0) \in L^2(\Omega)$, $\partial_t g \in L^2(0,T,H^{-1}(\Omega))$, $g \in L^p(\Omega_T)$, $\partial_t a \equiv 0$, $\beta_\sigma < 1$, and $(\mathrm{Trace}[\sigma])^{\frac{n}{4}-1} \in L^\infty(\Omega_T)$, then there exists a real number $p_0 > 2$ depending only on n, Ω and β_σ such that for each p, $2 \leq p < p_0$, one has*

$$\begin{aligned} \|u \circ F^{-1}\|_{L^p(0,T,W_D^{2,p}(\Omega))} \leq & \frac{C}{1-\beta_\sigma^{\frac{1}{2}}} \big(\|g\|_{L^p(\Omega_T)} \\ & + \|g(.,0)\|_{L^2(\Omega)} + \|\partial_t g\|_{L^2(0,T,H^{-1}(\Omega))} \big). \end{aligned} \tag{3.1.12}$$

The constant C in Theorem 3.1.2 and 3.1.3 can be written as

$$C = \frac{C_{n,\Omega,p}}{(\lambda_{\min}(a))^{\frac{n}{4}}} \big\| (\mathrm{Trace}[\sigma])^{\frac{n}{4}-1} \big\|_{L^\infty(\Omega_T)} \big(1 + \frac{1}{\lambda_{\min}(a)}\big)^{\frac{1}{2}}.$$

Write

$$\|v\|_{C^\gamma(\Omega)} := \sup_{x,y \in \Omega, x \neq y} \frac{|v(x) - v(y)|}{|x-y|^\gamma} \tag{3.1.13}$$

Theorem 3.1.4. *Assume that $n \leq 2$, Ω is convex, $g(.,0) \in L^2(\Omega)$, $\partial_t g \in L^2(0,T,H^{-1}(\Omega))$, $g \in L^p(\Omega_T)$, $\partial_t a \equiv 0$, $\beta_\sigma < 1$, $(\mathrm{Trace}[\sigma])^{-1} \in L^\infty(\Omega_T)$, and $g \in L^2[0,T;L^{p^*}(\Omega)]$ with $2 < p^*$. Then there exists $p \in (2,p^*]$ and $\gamma(p) > 0$ such that*

$$\begin{aligned} \Big(\int_0^T \big\| \nabla (u \circ F^{-1})(.,t) \big\|_{C^\gamma(\Omega)}^2 dt \Big)^{\frac{1}{2}} \leq & \frac{C}{1-\beta_\sigma^{\frac{1}{2}}} \big(\|g\|_{L^p(\Omega_T)} \\ & + \|g(.,0)\|_{L^2(\Omega)} + \|\partial_t g\|_{L^2(0,T,H^{-1}(\Omega))} \big). \end{aligned} \tag{3.1.14}$$

The constant C in (3.1.14) depends on n, p, Ω, $\lambda_{\min}(a)$, $\lambda_{\max}(a)$ and $\left\|(\mathrm{Trace}(\sigma))^{-1}\right\|_{L^\infty(\Omega_T)}$.

Remark. *It easy to check that if $n=1$ then the theorem is valid with $\gamma = 1/2$.*

3.1.1.2 Medium with a continuum of time scales

In this subsection the entries of a are merely in $L^\infty(\Omega_T)$. We need to introduce the following parabolic Cordes type condition.

Condition 3.1.1 (Cordes Type Condition). *We say that condition 3.1.1 is satisfied if and only if there exists $\delta \in (0,\infty)$ and $\varepsilon \in (0,1)$ such that*

$$\mathrm{esssup}_{\Omega_T} \frac{\delta^2 \,\mathrm{Trace}[{}^t\sigma\sigma] + 1}{\left(\delta\,\mathrm{Trace}[\sigma] + 1\right)^2} \leq \frac{1}{n+\varepsilon}. \tag{3.1.15}$$

Write

$$z_\sigma := \mathrm{esssup}_{\Omega_T} n \frac{\mathrm{Trace}[{}^t\sigma\sigma]}{(\mathrm{Trace}[\sigma])^2}. \tag{3.1.16}$$

Observe that z_σ is a measure of anisotropy of σ, in particular $1 \leq z_\sigma \leq n$ and $z_\sigma = 1$ if σ is isotropic.

Write

$$y_\sigma := \|\,\mathrm{Trace}[\sigma]\|_{L^\infty(\Omega_T)} \left\|(\mathrm{Trace}[\sigma])^{-1}\right\|_{L^\infty(\Omega_T)}. \tag{3.1.17}$$

Proposition 3.1.1. *If $\|\,\mathrm{Trace}[\sigma]\|_{L^\infty(\Omega_T)} < \infty$ and $\left\|(\mathrm{Trace}[\sigma])^{-1}\right\|_{L^\infty(\Omega_T)} < \infty$ then condition 3.1.1 is satisfied with*

$$\delta := n \left\|(\mathrm{Trace}[\sigma])^{-1}\right\|_{L^\infty(\Omega_T)}. \tag{3.1.18}$$

and with $\varepsilon := (2y_\sigma - 1)/(2y_\sigma^2)$ provided that $z_\sigma \leq 1 + \varepsilon/n$.

Remark. *Notice that in dimension one $z_\sigma = 1$, thus for $n=1$ condition 3.1.1 is satisfied if $\mathrm{Trace}[\sigma] \in L^\infty(\Omega_T)$ and $(\mathrm{Trace}[\sigma])^{-1} \in L^\infty(\Omega_T)$.*

We have the following theorems:

Theorem 3.1.5. *Assume that Ω is convex, and condition 3.1.1 is satisfied then $u\circ F^{-1}\in L^2\big(0,T,W_D^{2,2}(\Omega)\big)$, $\partial_t(u\circ F^{-1})\in L^2(\Omega_T)$, and*

$$\|u\circ F^{-1}\|_{L^2(0,T,W_D^{2,2}(\Omega))}+\|\partial_t(u\circ F^{-1})\|_{L^2(\Omega_T)}\leq C\|g\|_{L^2(\Omega_T)} \tag{3.1.19}$$

where C depends on Ω, n, δ, and ε.

Remark. *According to theorem 3.1.5 although the second order space derivatives and first order time derivatives of u with respect to Euclidean coordinates are only in $L^2(0,T,H^{-1}(\Omega))$, they are in $L^2(\Omega_T)$ with respect to caloric coordinates.*

Similarly we can obtain the following theorems in situations where $g\in L^p(\Omega_T)$ with some $p>2$.

Theorem 3.1.6. *Assume that Ω is convex, and condition 3.1.1 is satisfied then there exists a number $p_0>2$ depending on n,Ω,ε such that for $p\in(2,p_0)$, $u\circ F^{-1}\in L^p\big(0,T,W_D^{2,p}(\Omega)\big)$, $\partial_t(u\circ F^{-1})\in L^p(\Omega_T)$ and*

$$\|u\circ F^{-1}\|_{L^p(0,T,W_D^{2,p}(\Omega))}+\|\partial_t(u\circ F^{-1})\|_{L^p(\Omega_T)}\leq C\|g\|_{L^p(\Omega_T)} \tag{3.1.20}$$

where C depends on Ω, n, δ, and ε.

Theorem 3.1.7. *Assume that Ω is convex, and condition 3.1.1 is satisfied, then there exists a number $\alpha_0>2$ depending on n,Ω,ε such that for $\alpha\in(0,\alpha_0)$, $\nabla(u\circ F^{-1})\in L^2(0,T,C^\alpha(\Omega))$ and*

$$\|\nabla(u\circ F^{-1})(.,t)\|_{L^2(0,T,C^\alpha(\Omega))}\leq C\|g\|_{L^p(\Omega_T)} \tag{3.1.21}$$

where C depends on Ω, δ, n, and ε.

These compensation phenomena can also be observed numerically. Choose

$$a(x,y,t)=\frac{1}{6}\Big(\sum_{i=1}^{5}\frac{1.1+\sin(2\pi x'/\varepsilon_i)}{1.1+\sin(2\pi y'/\varepsilon_i)}+\sin(4x'^2y'^2)+1\Big) \tag{3.1.22}$$

with $x'=x+\sqrt{2}t$, $y'=y-\sqrt{2}t$, $\varepsilon_1=\frac{1}{5}$, $\varepsilon_2=\frac{1}{13}$, $\varepsilon_3=\frac{1}{17}$, $\varepsilon_4=\frac{1}{31}$, and $\varepsilon_5=\frac{1}{65}$. This medium has been shown in figure 3.3 at time 0 (note that $\lambda_{\max}(a)/\lambda_{\min}(a)\sim 100$).

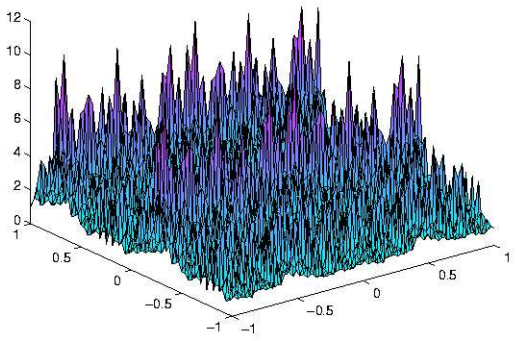
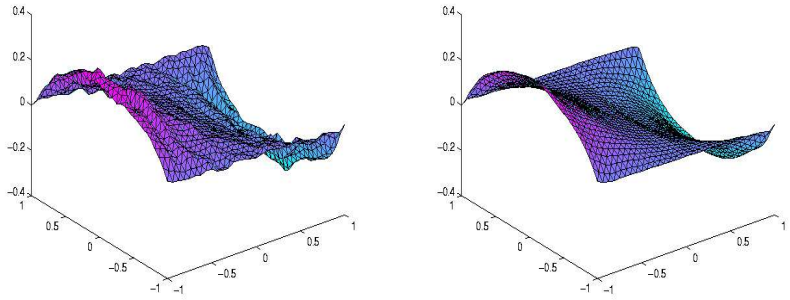


Figure 3.3: a at time 0

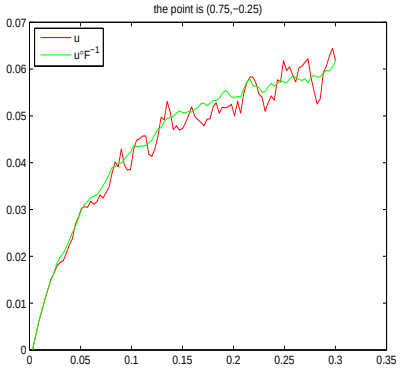
(3.1.1) has been solved numerically with $g \equiv 1$ on the fine mesh with 16641 nodes and 32768 triangles. Figure 3.4 illustrates $\partial_x u$ and $\partial_x (u \circ F^{-1})$ at time 0.3.



(a) $\partial_x u$ **(b)** $\partial_x (u \circ F^{-1})$

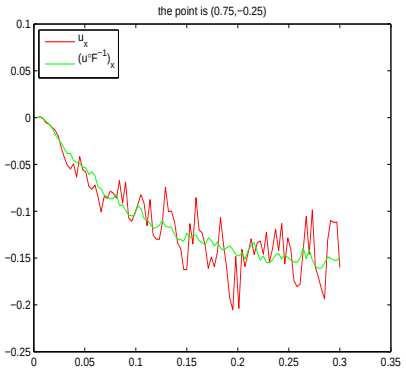
Figure 3.4: $\partial_x u$ and $\partial_x (u \circ F^{-1})$ at time $t = 0.3$ for the multi-scale trigonometric time dependent medium.

In figure 3.5 and 3.6, $x_0 = (0.75, -0.25)$ and the curves $t \to u(x_0,t), u \circ F^{-1}(x_0,t), \nabla u(x_0,t), \nabla u \circ F^{-1}(x_0,t)$ are plotted from $t = 0$ to $t = 0.3$.

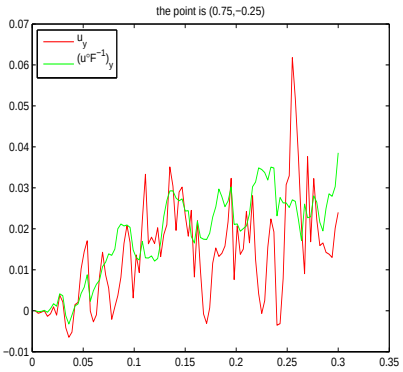


(a) u ***and*** $u \circ F^{-1}$

Figure 3.5: $t \to u(x_0,t), u \circ F^{-1}(x_0,t)$ from $t = 0$ to $t = 0.3$ with $x_0 = (0.75, -0.25)$



(a) $\partial_x u$ ***and*** $\partial_x(u \circ F^{-1})$



(b) $\partial_y u$ ***and*** $\partial_y(u \circ F^{-1})$

Figure 3.6: $t \to \nabla u(x_0,t), \nabla u \circ F^{-1}(x_0,t)$ from $t = 0$ to $t = 0.3$ with $x_0 = (0.75, -0.25)$

3.1.2 Numerical Homogenization in Space

Suppose we have a quasi-uniform coarse mesh with its aspect ratio uniformly bounded. Let X_h be a finite dimensional subspace of $H_0^1(\Omega) \cap W^{1,\infty}(\Omega)$ [1] defined on the coarse mesh with the following approximation property: There exists a constant C_X such that for all $f \in W_D^{2,2}(\Omega)$

$$\inf_{v \in X_h} \|f - v\|_{H_0^1(\Omega)} \leq C_X h \|f\|_{W_D^{2,2}(\Omega)}. \tag{3.1.23}$$

It is known that the set of piecewise linear functions on a triangulation of Ω satisfies condition (3.1.23) provided that the length of the edges of the triangles are bounded by h (C_X in (3.1.23) is given by the aspect ratio of the triangles).

For media characterized by a continuum of time scales we will use C^1 differentiable elements which satisfy the following inverse inequalities (see [57, Section 1.7]): for $v \in X_h$,

$$\|v\|_{W_D^{2,2}(\Omega)} \leq C_X h^{-1} \|v\|_{H_0^1(\Omega)}. \tag{3.1.24}$$

and

$$\|v\|_{H_0^1(\Omega)} \leq C_X h^{-1} \|v\|_{L^2(\Omega)}. \tag{3.1.25}$$

As in Chapter 2, we use WEB spline based finite element to ensure that the conditions (3.1.24) and (3.1.25) are satisfied.

For $t \in [0,T]$ let us define the time-space finite element space

$$V^h(t) := \{\varphi \circ F(x,t) \,:\, \varphi \in X_h\}. \tag{3.1.26}$$

Write

$$a[v,w](t) := \int_\Omega {}^t\nabla v(x,t) a(x,t) \nabla w(x,t)\, dx. \tag{3.1.27}$$

Define Y_T^h the subspace of $L^2\big(0,T;H_0^1(\Omega)\big)$ as

$$Y_T^h := \{v \in L^2\big(0,T;H_0^1(\Omega)\big) : v(x,t) \in V^h(t)\}. \tag{3.1.28}$$

[1] $W^{1,\infty}$ is the usual space of uniformly Lipschitz continuous functions.

Write u_h the solution in Y_T^h of the following system of ordinary differential equations:

$$\begin{cases} (\psi, \partial_t u_h)_{L^2(\Omega)} + a[\psi, u_h](t) = (\psi, g)_{L^2(\Omega)} \quad \text{for all } t \in (0,T) \text{ and } \psi \in V^h(t) \\ u_h(x,0) = 0. \end{cases} \tag{3.1.29}$$

We have the following theorem for time independent medium:

Theorem 3.1.8. *Assume that $\partial_t a \equiv 0$, Ω is convex, $\beta_\sigma < 1$ and $(\mathrm{Trace}[\sigma])^{-1} \in L^\infty(\Omega_T)$, then*

$$\big\|(u-u_h)(.,T)\big\|_{L^2(\Omega)} + \big\|u-u_h\big\|_{L^2(0,T;H_0^1(\Omega))} \leq Ch\|g\|_{L^2(\Omega_T)}. \tag{3.1.30}$$

The constant C depends on C_X, n, Ω, $\lambda_{\min}(a)$, $\lambda_{\max}(a)$, and $\big\|(\mathrm{Trace}[\sigma])^{-1}\big\|_{L^\infty(\Omega_T)}$. If $n \geq 3$ it also depends on $\big\|\mathrm{Trace}[\sigma]\big\|_{L^\infty(\Omega_T)}$.

We have the following theorem for medium with a continuum of time scales:

Theorem 3.1.9. *Assume that Ω is convex, and condition 3.1.1 is satisfied then*

$$\big\|(u-u_h)(.,T)\big\|_{L^2(\Omega)} + \big\|u-u_h\big\|_{L^2(0,T,H_0^1(\Omega))} \leq Ch\|g\|_{L^2(\Omega_T)}. \tag{3.1.31}$$

The constant C depends on C_X, n, Ω, δ, ε, $\lambda_{\min}(a)$, and $\lambda_{\max}(a)$.

3.1.3 Numerical Homogenization in Space and Time

The ordinary differential equations system (3.1.29) is still characterized by a continuum of time scales in situations where the entries of a merely belong to $L^\infty(\Omega_T)$. They need to be discretized (homogenized) in time in order to be solved numerically. This will be the object of the this subsection. Loosely speaking, although the parabolic operator (3.1.1) is associated to a fine tessellation and fine time steps, it is possible to approximate it on a coarse tessellation with coarse time steps.

Let $M \in \mathbb{N}$, $(t_n = n\frac{T}{M})_{0\leq n\leq M}$ is a discretization of $[0,T]$. (φ_i) is a basis of X_h. Write Z_T^h the subspace of Y_T^h such that

$$Z_T^h = \{w \in Y_T^h : w(x,t) = \sum_i c_i(t)\varphi_i\big(F(x,t)\big),\ c_i(t) \text{ are constants on } (t_n, t_{n+1}]\}. \tag{3.1.32}$$

Write U_T^h the subspace of Y_T^h such that

$$U_T^h = \{\psi \in Y_T^h : \psi(x,t) = \sum_i d_i \varphi_i\big(F(x,t)\big),\ d_i \text{ are constants (on } [0,T]).\}. \tag{3.1.33}$$

For $w \in Y_T^h$, define $w_n \in U_T^h$ by

$$w_n(x,t) := \sum_i c_i(t_n)\varphi_i(F(x,t)). \tag{3.1.34}$$

Let $v \in Z_T^h$ be the solution of the following implicit weak formulation (suppose that $v(x,0) \equiv 0$): for $n \in \{0,\dots,M-1\}$ and $\forall \psi \in U_T^h$,

$$\begin{aligned}\big(\psi(t_{n+1}), v_{n+1}(t_{n+1})\big)_{L^2(\Omega)} =& \big(\psi(t_n), v_n(t_n)\big)_{L^2(\Omega)} + \int_{t_n}^{t_{n+1}} \Big(\big(\partial_t \psi(t), v_{n+1}(t)\big)_{L^2(\Omega)} \\ & - a\big[\psi, v_{n+1}\big](t) + \big(\psi(t), g(t)\big)_{L^2(\Omega)}\Big)\, dt.\end{aligned} \tag{3.1.35}$$

The following Theorem 3.1.10 shows the stability of the implicit scheme (3.1.35):

Theorem 3.1.10. *Let $v \in Z_T^h$ be the solution of (3.1.35). We have*

$$\big\|v(T)\big\|_{L^2(\Omega)} + \|v\|_{L^2(0,T,H_0^1(\Omega))} \le C\|g\|_{L^2(\Omega_T)}. \tag{3.1.36}$$

The constant C depends on n, Ω, and $\lambda_{\min}(a)$.

The following Theorem 3.1.11 gives an error bound for the time discretization scheme (3.1.35) when a does not depend on time:

Theorem 3.1.11. *Let $v \in Z_T^h$ be the solution of (3.1.35) and u_h be the solution of (3.1.29). Assume that $\partial_t a \equiv 0$. We have*

$$\big\|(u_h - v)(T)\big\|_{L^2(\Omega)} + \|u_h - v\|_{L^2(0,T,H_0^1(\Omega))} \le C\Delta t \Big(\|\partial_t g\|_{L^2(0,T,H^{-1}(\Omega))} + \big\|g(.,0)\big\|_{L^2(\Omega)} \Big). \tag{3.1.37}$$

The constant C depends on n, Ω, $\lambda_{\min}(a)$, and $\lambda_{\max}(a)$.

The following Theorem 3.1.12 gives an error bound for the time discretization scheme (3.1.35) when a is time dependent:

Theorem 3.1.12. *Assume that Ω is convex, and condition 3.1.1 is satisfied. Let $v \in Z_T^h$ be the solution of (3.1.35) and u_h be the solution of (3.1.29), we have*

$$\left\|(u_h - v)(T)\right\|_{L^2(\Omega)} + \|u_h - v\|_{L^2(0,T,H_0^1(\Omega))} \leq C\frac{|\Delta t|}{h}\|g\|_{L^2(\Omega_T)}. \tag{3.1.38}$$

where C depends on Ω, n, δ, ε, $\lambda_{\min}(a)$, and $\lambda_{\max}(a)$.

Remark. *Observe that the accuracy of the time discretization scheme (3.1.35) requires that $|\Delta t| << h$ when a has no bounded time derivatives.*

3.2 Proofs

3.2.1 Compensation Phenomena

3.2.1.1 Time independent medium: Proof of Theorem 3.1.1–3.1.4

Suppose that $\partial_t a \equiv 0$. We will need the following lemmas. Let $\mathscr{A}_T$ be the bilinear form on $L^2\left(0,T;H_0^1(\Omega)\right)$ defined by

$$\mathscr{A}_T[v,w] := \int_0^T a[v,w](t)\,dt. \tag{3.2.1}$$

where

$$a[v,w](t) := \int_\Omega {}^t\nabla v(x,t)a(x,t)\nabla w(x,t)\,dx. \tag{3.2.2}$$

Let $\mathscr{A}_T[u] := \mathscr{A}_T[u,u]$.

The following Lemmas 3.2.1–3.2.3 are the standard energy estimates. We show them by assuming a smooth. When a is nonsmooth, we can use Galerkin approximations of u and then pass to the limit. We refer to [58, Section 7.1.2] for a reminder.

Lemma 3.2.1. *We have*

$$\left\|u(.,T)\right\|^2_{L^2(\Omega)} + \mathscr{A}_T[u] \leq \frac{C_{n,\Omega}}{\lambda_{\min}(a)} \|g\|_{L^2(\Omega_T)}. \tag{3.2.3}$$

Proof. Multiplying (3.1.1) by u and integrating over Ω_T we obtain that

$$\frac{1}{2}\left\|u(.,T)\right\|^2_{L^2(\Omega)} + \mathscr{A}_T[u] = (u,g)_{L^2(\Omega_T)}. \tag{3.2.4}$$

(3.2.3) is established by using Poincaré and Cauchy-Schwarz inequalities. □

Lemma 3.2.2. *Assume $\partial_t a \equiv 0$. We have*

$$\left\|\partial_t u\right\|^2_{L^2(\Omega_T)} + a\left[u(.,T)\right] \leq \left\|g\right\|^2_{L^2(\Omega_T)}. \tag{3.2.5}$$

Proof. Multiplying (3.1.1) by $\partial_t u$ and integrating by parts we obtain that

$$\left\|\partial_t u(.,t)\right\|^2_{L^2(\Omega)} + a[\partial_t u, u] = (\partial_t u, g)_{L^2(\Omega)}. \tag{3.2.6}$$

Noticing that

$$a[\partial_t u, u] = \frac{1}{2}\partial_t\big(a[u]\big). \tag{3.2.7}$$

we conclude by integration with respect to time and using Cauchy-Schwarz inequality. □

Lemma 3.2.3. *Assume $\partial_t a \equiv 0$. We have*

$$\left\|\partial_t u(.,T)\right\|^2_{L^2(\Omega)} + \mathscr{A}_T[\partial_t u] \leq \frac{1}{\lambda_{\min}(a)} \|\partial_t g\|^2_{L^2(0,T,H^{-1}(\Omega))} + \left\|g(.,0)\right\|^2_{L^2(\Omega)}. \tag{3.2.8}$$

Proof. We obtain from (3.1.1) that

$$\partial_t^2 u = \operatorname{div}\big(a(x)\nabla\partial_t u(x,t)\big) + \partial_t g. \tag{3.2.9}$$

Multiplying (3.2.9) by $\partial_t u$ and integrating over Ω_T we obtain that

$$\frac{1}{2}\left\|\partial_t u(.,T)\right\|^2_{L^2(\Omega)} + \mathscr{A}_T[\partial_t u] = \int_0^T (\partial_t u, \partial_t g)_{L^2(\Omega)}\, dt + \frac{1}{2}\left\|\partial_t u(.,0)\right\|^2_{L^2(\Omega)}. \tag{3.2.10}$$

We conclude by the H^{-1}-duality inequality and Cauchy-Schwarz inequality. □

Let us now prove the compensation theorems. Choose

$$M := \frac{\sigma}{|\det(\nabla F)|^{\frac{1}{2}}} \circ F^{-1}. \tag{3.2.11}$$

It easy to see that $\beta_M = \beta_\sigma$ [2]. Observe that

$$\|\nu_M\|^2_{L^\infty(\Omega_T)} \leq \frac{C_n}{(\lambda_{\min}(a))^{\frac{n}{2}}} \left\|(\mathrm{Trace}[\sigma])^{\frac{n}{4}-1}\right\|^2_{L^\infty(\Omega_T)}. \tag{3.2.12}$$

Fix $t \in [0,T]$, choose

$$f := \frac{(\partial_t u - g)}{|\det(\nabla F)|^{\frac{1}{2}}} \circ F^{-1}. \tag{3.2.13}$$

Observe that by the change of variable $y = F(x)$ one obtains that if $\partial_t a \equiv 0$ (which implies that F is time independent), $\partial_t u \in L^2(\Omega)$, and $g(.,t) \in L^2(\Omega)$ then $f \in L^2(\Omega)$ and

$$\|f\|_{L^2(\Omega)} \leq \|\partial_t u\|_{L^2(\Omega)} + \|g\|_{L^2(\Omega)}. \tag{3.2.14}$$

It follows from Theorem A.1.3 that there exists a unique $v \in W^{2,2}_D(\Omega)$ satisfying

$$\|v\|^2_{W^{2,2}_D(\Omega)} \leq \frac{C\|\nu_M\|^2_{L^\infty(\Omega_T)}}{(1-\beta_\sigma^{\frac{1}{2}})^2} \left(\|\partial_t u\|^2_{L^2(\Omega)} + \|g\|^2_{L^2(\Omega)}\right). \tag{3.2.15}$$

and the following equation

$$\partial_t \tilde{u}(y,t) = \sum_{i,j} \left(\sigma(F^{-1}(y,t),t)\right)_{i,j} \partial_i \partial_j v(y,t) + \tilde{g}(y,t). \tag{3.2.16}$$

We use the notation $\tilde{g} := g \circ F^{-1}$ and $\tilde{u} := u \circ F^{-1}$. Using the change of variable $y = F(x)$

[2] see p.139 for the definition of Cordes parameter β_M

and using the property $\operatorname{div} a\nabla F = 0$ when $\partial_t a \equiv 0$ we obtain that (3.2.16) can be written

$$\partial_t u = \operatorname{div}\big(a\nabla(v\circ F)\big) + g. \tag{3.2.17}$$

If $\partial_t u \in L^2(\Omega)$ and $g(.,t) \in L^2(\Omega)$ we can use the uniqueness property of the solution of the divergence form elliptic Dirichlet problem

$$\operatorname{div}\big(a\nabla w\big) = \partial_t u - g \tag{3.2.18}$$

to obtain that $v\circ F = u$. Thus using Lemma 3.2.3 we have proven Theorem 3.1.2. Moreover assume that $g \in L^2(\Omega_T)$ and $\partial_t u \in L^2(\Omega_T)$. It follows that for $t \in [0,T]-B$, $g(.,t) \in L^2(\Omega)$ and $\partial_t u(.,t) \in L^2(\Omega)$ where B is a subset of $[0,T]$ of 0-Lebesgue measure. It results from the previous arguments that on $[0,T]-B$, $u\circ F^{-1}(.,t) \in W_D^{2,2}(\Omega)$ and

$$\|u\circ F^{-1}(.,t)\|^2_{W_D^{2,2}(\Omega)} \leq \frac{C\|\nu_M\|^2_{L^\infty(\Omega_T)}}{(1-\beta_\sigma^{\frac{1}{2}})^2}\big(\|\partial_t u(.,t)\|^2_{L^2(\Omega)} + \|g(.,t)\|^2_{L^2(\Omega)}\big). \tag{3.2.19}$$

Integrating (3.2.19) with respect to time we obtain that $u\circ F^{-1} \in L^2(0,T,W_D^{2,2}(\Omega))$ and

$$\|u\circ F^{-1}\|^2_{L^2(0,T,W_D^{2,2}(\Omega))} \leq \frac{C\|\nu_M\|^2_{L^\infty(\Omega_T)}}{(1-\beta_\sigma^{\frac{1}{2}})^2}\big(\|\partial_t u\|^2_{L^2(\Omega_T)} + \|g\|^2_{L^2(\Omega_T)}\big). \tag{3.2.20}$$

thus by Lemma 3.2.2 we obtain Theorem 3.1.1.

Let us now prove Theorem 3.1.3. Assume that there exists $q_0 > 2$ such that for $2 \leq p < q_0$, $\partial_t u \in L^p(\Omega_T)$, and $g \in L^p(\Omega_T)$. Let us now apply Theorem A.1.3 with $p < \min(p_0, q_0)$, M given by (3.2.11), and f given by (3.2.13). It follows that for $t \in [0,T]-B$ (where B is a subset of $[0,T]$ of 0-Lebesgue measure), $g(.,t) \in L^p(\Omega)$ and $\partial_t u(.,t) \in L^p(\Omega)$. We deduce from Theorem A.1.3 and the argumentation related to equation (3.2.18) that on $[0,T]-B$, $u\circ F^{-1}(.,t) \in W_D^{2,p}(\Omega)$ and

$$\|u\circ F^{-1}(.,t)\|^p_{W_D^{2,p}(\Omega)} \leq \frac{C_{n,p,\Omega}\|\nu_M\|^p_{L^\infty(\Omega_T)}}{(1-\beta_\sigma^{\frac{1}{2}})^p}\big(\|\partial_t u(.,t)\|^p_{L^p(\Omega)} + \|g(.,t)\|^p_{L^p(\Omega)}\big). \tag{3.2.21}$$

Integrating (3.2.21) with respect to time we obtain that $u\circ F^{-1}\in L^p(0,T,W^{2,p}_D(\Omega))$ and

$$\|u\circ F^{-1}\|_{L^p(0,T,W^{2,p}_D(\Omega))}\leq \frac{C_{n,p,\Omega}\|\nu_M\|_{L^\infty(\Omega_T)}}{1-\beta_\sigma^{\frac{1}{2}}}\big(\|\partial_t u\|_{L^p(\Omega_T)}+\|g\|_{L^p(\Omega_T)}\big). \tag{3.2.22}$$

It remains to show that under the assumptions of theorem 3.1.3, $\partial_t u\in L^p(\Omega_T)$.

In order to bound $\big\|\partial_t u(.,t)\big\|_{L^p(\Omega)}$ we use general Sobolev inequalities [58, Section 5.6].

- If $n\geq 3$, write $p^*=2n/(n-2)$. We have for $2<p\leq p^*$,

$$\Big(\int_\Omega(\partial_t u)^p\,dx\Big)^{\frac{2}{p}}\leq C_{n,\Omega}\Big(\int_\Omega(\partial_t u)^{p^*}\,dx\Big)^{\frac{2}{p^*}} \tag{3.2.23}$$

  thus, using Gagliardo-Nirenberg-Sobolev inequality

$$\Big(\int_\Omega(\partial_t u)^p\,dx\Big)^{\frac{2}{p}}\leq C_{n,p,\Omega}\frac{1}{\lambda_{\min}(a)}a[\partial_t u]. \tag{3.2.24}$$

- If $n=2$, we write for $(x_1,x_2,x_3)\in\Omega\times(0,1)$, $v(x_1,x_2,x_3):=\partial_t u(x_1,x_2)$. Using Gagliardo-Nirenberg-Sobolev inequality in dimension three we obtain that for $2<p\leq 6$

$$\Big(\int_\Omega(\partial_t u)^p\,dx\Big)^{\frac{2}{p}}\leq C_{n,p,\Omega}\int_\Omega(\nabla\partial_t u)^2\,dx. \tag{3.2.25}$$

  which leads us to (3.2.24).

- If $n=1$ then using Morrey's inequality we obtain that with $\gamma:=1/2$,

$$\|\partial_t u\|^2_{C^{0,\gamma}(\Omega)}\leq C_\Omega\frac{1}{\lambda_{\min}(a)}a[\partial_t u]. \tag{3.2.26}$$

We conclude the proof of Theorem 3.1.3 using Lemma 3.2.3.

Theorem 3.1.4 follows from Morrey's inequality and theorem 3.1.3.

3.2.1.2 Medium with a continuum of time scales: Proof of Theorem 3.1.5–3.1.7

We will need the adapted version of Theorems 1.6.2 and 1.6.3 of [93], which are adapted as Theorems A.2.1 and A.2.2 in the Appendix. These theorems show that under a parabolic

Cordes type condition, nondivergence form parabolic equation has a unique solution in $L^p\big(0,T,W_D^{2,p}(\Omega)\big)$. Consider the following parabolic problem:

$$\partial_t v = \sum_{i,j=1}^{n} M_{ij}(x)\partial_i\partial_j v + f \tag{3.2.27}$$

assume M to be symmetric, uniformly bounded, and elliptic, $v = 0$ at $t = 0$ and on the boundary $\partial\Omega$. Write

$$\eta_M := \operatorname{esssup}_{x\in\Omega_T} \frac{\operatorname{Trace}[{}^tMM]+1}{\big(\operatorname{Trace}[M]+1\big)^2}, \tag{3.2.28}$$

and

$$\alpha_M := \operatorname{esssup}_{x\in\Omega_T} \frac{\operatorname{Trace}[M]+1}{\operatorname{Trace}[{}^tMM]+1}. \tag{3.2.29}$$

Write for $p \geq 2$

$$S_p(\Omega_T) := \Big\{ v \in L^p\big(0,T,W_D^{2,p}(\Omega)\big); \partial_t v \in L^p(\Omega_T); v(.,0) \equiv 0 \Big\}, \tag{3.2.30}$$

and

$$\|v\|^p_{S_p(\Omega_T)} := \int_{\Omega_T} \big(\sum_{i,j} (\partial_i\partial_j v)^2 + (\partial_t v)^2\big)^{\frac{p}{2}}\, dy\, dt. \tag{3.2.31}$$

Parabolic Cordes condition can be given by: there exists $\varepsilon \in (0,1)$, such that

$$\eta_M \leq \frac{1}{n+\varepsilon}. \tag{3.2.32}$$

Let $\delta > 0$. Let us now apply Theorem A.2.1 on $[0, T/\delta]$ with

$$M := \delta\sigma \circ F^{-1}(y, \delta t), \tag{3.2.33}$$

and

$$f := \delta(g \circ F^{-1})(y, \delta t). \tag{3.2.34}$$

Notice that $\eta_M < \infty$ and $\alpha_M < \infty$ since

$$\operatorname{esssup}_{\Omega_{\frac{T}{\delta}}} \frac{\operatorname{Trace}[{}^tMM]+1}{\left(\operatorname{Trace}[M]+1\right)^2} = \operatorname{esssup}_{\Omega_T} \frac{\delta^2 \operatorname{Trace}[{}^t\sigma\sigma]+1}{\left(\delta \operatorname{Trace}[\sigma]+1\right)^2}. \tag{3.2.35}$$

Let us now prove proposition 3.1.1. Recall that $x = \operatorname{Trace}[\sigma]$ and $z = n\frac{\operatorname{Trace}[{}^t\sigma\sigma]}{(\operatorname{Trace}[\sigma])^2}$ (observe that $1 \leq z \leq n$). It is easy to check that condition 3.1.1 can be written as,

$$-\delta^2 x^2 (\frac{\varepsilon+n}{n} z - 1) + 2x\delta - (n+\varepsilon-1) \geq 0. \tag{3.2.36}$$

Choose $\delta = n\left\|(\operatorname{Trace}[\sigma])^{-1}\right\|_{L^\infty(\Omega_T)}$. Observing that $\delta x \geq n$ and $\delta x = n y_\sigma$ it is easy to conclude the proof of proposition 3.1.1. Similarly we obtain the following lemma by straightforward computation from equation (3.2.36):

Lemma 3.2.4. *Assume that condition 3.1.1 is satisfied then*

$$\left\|(\operatorname{Trace}[\sigma])^{-1}\right\|_{L^\infty(\Omega_T)} \leq C(n,\varepsilon,\delta), \tag{3.2.37}$$

and

$$\left\|\operatorname{Trace}[\sigma]\right\|_{L^\infty(\Omega_T)} \leq C(n,\varepsilon,\delta). \tag{3.2.38}$$

If condition 3.1.1 is satisfied and F is an homeomorphism, It follows that the following equation admits a unique solution in $S_2(\Omega_{\frac{T}{\delta}})$ by Theorem A.2.1.

$$\partial_t w(y,t) = \sum_{i,j} M_{i,j}(y,t)\partial_i\partial_j w(y,t) + k(y,t) \tag{3.2.39}$$

with $k(y,t) = \delta\tilde{g}(y,\delta t)$. And we have

$$\int_0^{\frac{T}{\delta}} \int_\Omega \left((\partial_t w)^2 + \sum_{i,j}(\partial_i\partial_j w)^2\right) dy\,dt \leq \frac{C}{(1-\sqrt{1-\varepsilon})^2} \|f\|_{L^2(\Omega_{\frac{T}{\delta}})}. \tag{3.2.40}$$

Using the change of variables $t \to \delta t$ and writing

$$w(y,t) := v(y,\delta t). \tag{3.2.41}$$

we obtain that v satisfies the following equation on Ω_T

$$\partial_t v(y,t) = \sum_{i,j} \big(\sigma(F^{-1}(y,t),t)\big)_{i,j} \partial_i \partial_j v(y,t) + \tilde{g}(y,t). \tag{3.2.42}$$

Using the change of variable $y = F(x)$ and observing that $\partial_t F = \operatorname{div} a\nabla F$ we obtain that $v \circ F$ satisfies

$$\partial_t (v \circ F) = \operatorname{div}\big(a\nabla(v \circ F)\big) + g. \tag{3.2.43}$$

It follows from the uniqueness of the solution of (3.2.43) that $u = v \circ F$. In resume we have obtained Theorem 3.1.5 (use Lemma 3.2.4 to control the constants). The proof of Theorem 3.1.6 is similar and based on Theorem A.2.2. The proof of Theorem 3.1.7 follows from Theorem 3.1.6 and Morrey's inequality.

3.2.2 Analysis of Numerical Homogenization

Write $\mathscr{R}_h$ the projection operator mapping $L^2\big(0,T;H_0^1(\Omega)\big)$ onto Y_T^h, $\mathscr{R}_h$ is defined by: $\forall v \in Y_T^h$

$$\mathscr{A}_T[v, u - \mathscr{R}_h u] = 0. \tag{3.2.44}$$

Write $\rho := u - \mathscr{R}_h u$ and $\theta := \mathscr{R}_h u - u_h$.

Lemma 3.2.5.

$$\frac{1}{2}\big\|(u-u_h)(T)\big\|^2_{L^2(\Omega)} + \mathscr{A}_T[u-u_h] = \int_{\Omega_T} \rho\,\partial_t(u-u_h)\,dxdt + \mathscr{A}_T[\rho, u-u_h]. \tag{3.2.45}$$

Proof. Multiply (3.1.1) using test function $\psi \in V^h(t)$ and integrate over Ω, by subtracting (3.1.29) we obtain that

$$\big(\psi, \partial_t(u-u_h)\big) + a[\psi, u-u_h] = 0 \quad \text{for all } \psi \in V^h(t). \tag{3.2.46}$$

Integrating by parts with respect to time leads us to,

$$\big(\psi,(u-u_h)(.,T)\big)_{L^2(\Omega)}+\mathscr{A}_T[\psi,u-u_h]=\int_{\Omega_T}\partial_t\psi(u-u_h)\,dxdt. \tag{3.2.47}$$

Taking $\psi=\theta$ in (3.2.47) we deduce that

$$\begin{aligned}\big\|(u-u_h)(.,T)\big\|^2_{L^2(\Omega)}+\mathscr{A}_T[u-u_h]=&\int_{\Omega_T}\partial_t\theta(u-u_h)\,dxdt+\big(\rho,u-u_h\big)_{L^2(\Omega)}(T)\\&+\mathscr{A}_T[\rho,u-u_h].\end{aligned} \tag{3.2.48}$$

Integrating by parts again, it results

$$\begin{aligned}\int_{\Omega_T}\partial_t\theta(u-u_h)\,dxdt+\big(\rho,u-u_h\big)_{L^2(\Omega)}(T)=&\frac{1}{2}\big\|(u-u_h)(.,T)\big\|^2_{L^2(\Omega)}\\&+\int_{\Omega_T}\rho\partial_t(u-u_h))\,dxdt\end{aligned} \tag{3.2.49}$$

which completes the proof of the lemma. □

3.2.2.1 Time Independent Medium. Proof of Theorem 3.1.8

Lemma 3.2.6.

$$\big\|(u-u_h)(T)\big\|^2_{L^2(\Omega)}+\mathscr{A}_T[u-u_h]\le 2\|\rho\|_{L^2(\Omega_T)}\|\partial_t(u-u_h)\|_{L^2(\Omega_T)}+\mathscr{A}_T[\rho]. \tag{3.2.50}$$

Proof. Lemma 3.2.6 is a straightforward consequence of Lemma 3.2.5 and Cauchy-Schwartz inequality. □

Similar to Lemmas 3.2.1 and 3.2.2, we have the energy estimates for u_h in the following Lemmas 3.2.7 and 3.2.8.

Lemma 3.2.7. *We have*

$$\big\|u_h(.,T)\big\|^2_{L^2(\Omega)}+\mathscr{A}_T[u_h]\le\frac{C_{n,\Omega}}{\lambda_{\min}(a)}\|g\|_{L^2(\Omega_T)}. \tag{3.2.51}$$

Proof. Taking $\psi = u_h$ in (3.1.29) and integrating over Ω_T we obtain that

$$\frac{1}{2}\|u_h(.,T)\|^2_{L^2(\Omega)} + \mathscr{A}_T[u_h] = (u_h, g)_{L^2(\Omega_T)}. \tag{3.2.52}$$

Using Poincaré and Minkowski inequalities leads us to (3.2.51). □

Lemma 3.2.8. *Assume $\partial_t a \equiv 0$. We have*

$$\|\partial_t u_h\|^2_{L^2(\Omega_T)} + a[u_h(.,T)] \leq \|g\|^2_{L^2(\Omega_T)}. \tag{3.2.53}$$

Proof. The proof is similar to Lemma 3.2.2. We need to take $\psi = \partial_t u_h$ in (3.1.29). □

Let $t \in [0,T]$ and $v \in H^1_0(\Omega)$, we will write $\mathscr{R}_{h,t}v(.,t)$ the solution of:

$$\int_\Omega {}^t\nabla\psi a(x,t)(\psi, v - \mathscr{R}_{h,t}v)\,dx = 0 \quad \forall \psi \in V^h(t). \tag{3.2.54}$$

We need the following lemma:

Lemma 3.2.9. *Assume the mapping $x \to F(x,t)$ to be invertible for fixed t, then for $v \in H^1_0(\Omega)$ we have*

- *For $n = 1$,*

$$\left(a[v - \mathscr{R}_{h,t}v]\right)^{\frac{1}{2}} \leq C_X h \|v \circ F^{-1}(.,t)\|_{W^{2,2}_D} \|a\nabla F\|^{\frac{1}{2}}_{L^\infty(\Omega_T)}, \tag{3.2.55}$$

- *For $n \geq 2$,*

$$\begin{aligned}\left(a[v - \mathscr{R}_{h,t}v]\right)^{\frac{1}{2}} \leq & C_X h \|v \circ F^{-1}(.,t)\|_{W^{2,2}_D} \\ & \times C_n \mu_\sigma^{\frac{n-1}{4}} \left\|(\mathrm{Trace}[\sigma])^{-1}\right\|^{\frac{n-2}{4}}_{L^\infty(\Omega_T)}.\end{aligned} \tag{3.2.56}$$

Proof. Using the change of coordinates $y = F(x,t)$ we obtain that (write $\tilde{v} := v \circ F^{-1}$)

$$a[v] = Q[\tilde{v}] \tag{3.2.57}$$

with

$$\mathscr{Q}[w] := \int_\Omega {}^t\nabla w(y,t) Q(y,t) \nabla w(y,t)\, dy \tag{3.2.58}$$

and

$$Q(y,t) := \frac{\sigma}{\det(\nabla F)} \circ F^{-1}. \tag{3.2.59}$$

Using the definition of $\mathscr{R}_{h,t} v$ we derive that

$$\mathscr{Q}[\tilde{v} - \widehat{\mathscr{R}_{h,t} v}] = \inf_{\varphi \in X_h} \mathscr{Q}[\tilde{v} - \varphi]. \tag{3.2.60}$$

Using property (3.1.23) it follows,

$$\mathscr{Q}[\tilde{v} - \widehat{\mathscr{R}_{h,t} v}] \leq \lambda_{\max}(Q) C_X^2 h^2 \|\tilde{v}\|^2_{W_D^{2,2}(\Omega)}. \tag{3.2.61}$$

It is easy to obtain that

- For $n = 1$,

$$\lambda_{\max}(Q) \leq \|a \nabla F\|_{L^\infty(\Omega_T)}, \tag{3.2.62}$$

- For $n \geq 2$,

$$\lambda_{\max}(Q) \leq C_n \mu_\sigma^{\frac{n-1}{2}} \left\| (\mathrm{Trace}[\sigma])^{-1} \right\|^{\frac{n}{2}-1}_{L^\infty(\Omega_T)}. \tag{3.2.63}$$

□

Lemma 3.2.10. *Assume that $\partial_t a \equiv 0$, Ω is convex, $\beta_\sigma < 1$, and $(\mathrm{Trace}[\sigma])^{-1} \in L^\infty(\Omega_T)$ then*

$$\mathscr{A}_T[\rho] \leq C h^2 \|g\|^2_{L^2(\Omega_T)}. \tag{3.2.64}$$

The constant C depends on C_X, n, Ω, $\lambda_{\min}(a)$, $\lambda_{\max}(a)$, and $\left\| (\mathrm{Trace}[\sigma])^{-1} \right\|_{L^\infty(\Omega_T)}$. If $n \geq 3$ it may also depend on $\left\| \mathrm{Trace}[\sigma] \right\|_{L^\infty(\Omega_T)}$.

Proof. The proof is a straightforward application of Lemma 3.2.9 and theorem 3.1.1. Observe that in dimension one $a \nabla F = (\int_\Omega a^{-1})^{-1}$. □

Lemma 3.2.11. *Assume that $\partial_t a \equiv 0$, Ω is convex, $\beta_\sigma < 1$, and $(\mathrm{Trace}[\sigma])^{-1} \in L^\infty(\Omega_T)$ then*

$$\|\rho\|_{L^2(\Omega_T)} \leq Ch^2 \|g\|_{L^2(\Omega_T)} \tag{3.2.65}$$

The constant C depends on C_X, n, Ω, $\lambda_{\min}(a)$, $\lambda_{\max}(a)$, and $\left\|(\mathrm{Trace}[\sigma])^{-1}\right\|_{L^\infty(\Omega_T)}$. If $n \geq 3$ it also depends on $\left\|\mathrm{Trace}[\sigma]\right\|_{L^\infty(\Omega_T)}$.

Proof. The proof follows from standard duality techniques (see for instance [40, Theorem 5.7.6]). We choose $v \in L^2(0,T,H_0^1(\Omega))$ to be the solution of the following linear problem: for all $w \in L^2(0,T,H_0^1(\Omega))$

$$\mathscr{A}_T[w,v] = (w,\rho)_{L^2(\Omega_T)}. \tag{3.2.66}$$

Choosing $w = \rho$ in equation (3.2.66) we deduce that

$$\|\rho\|^2_{L^2(\Omega_T))} = \mathscr{A}_T[\rho, v - \mathscr{R}_h v]. \tag{3.2.67}$$

By Cauchy Schwartz inequality,

$$\|\rho\|^2_{L^2(\Omega_T)} \leq \left(\mathscr{A}_T[\rho]\right)^{\frac{1}{2}} \left(\mathscr{A}_T[v - \mathscr{R}_h v]\right)^{\frac{1}{2}}. \tag{3.2.68}$$

By Theorem 3.1.1 we have

$$\|\tilde{v}\|_{L^2(0,T,W_D^{2,2}(\Omega))} \leq C\|\rho\|_{L^2(\Omega_T)}. \tag{3.2.69}$$

Using Lemma 3.2.9 it results

$$\left(\mathscr{A}_T[v - \mathscr{R}_h v]\right)^{\frac{1}{2}} \leq Ch\|\rho\|_{L^2(\Omega_T)}. \tag{3.2.70}$$

It follows that

$$\|\rho\|_{L^2(\Omega_T)} \leq Ch\left(\mathscr{A}_T[\rho]\right)^{\frac{1}{2}}. \tag{3.2.71}$$

We deduce the lemma by applying Lemma 3.2.10 to bound $\mathscr{A}_T[\rho]$. □

Theorem 3.1.8 is a straightforward application of Lemmas 3.2.2, 3.2.6, 3.2.8, 3.2.10,

and 3.2.11.

3.2.2.2 Medium with a Continuum of Time Scales: Proof of Theorem 3.1.9

In this subsection we will assume that the finite elements space X_h satisfies the inverse inequality (3.1.24).

Lemma 3.2.12.

$$\begin{aligned}\frac{1}{2}\big\|(u-u_h)(T)\big\|^2_{L^2(\Omega)} + \mathscr{A}_T[u-u_h] = \int_{\Omega_T} \frac{\tilde{\rho}}{|\det \nabla F| \circ F^{-1}} \\ \big(\tilde{g} + \sum_{i,j=1}^{n} \sigma_{i,j} \circ F^{-1} \partial_i \partial_j \tilde{u}_h - \partial_t \tilde{u}_h\big)\, dx\, dt\end{aligned} \tag{3.2.72}$$

Proof. Using change of variable $x = F^{-1}(y)$ in equation (3.2.45), we have

$$\begin{aligned}\int_{\Omega_T} \rho \partial_t (u-u_h)\, dx\, dt = \int_{\Omega_T} \frac{\tilde{\rho}}{|\det \nabla F| \circ F^{-1}} \partial_t (\tilde{u} - \tilde{u}_h)\, dy\, dt \\ + \int_{\Omega_T} \rho \partial_t F (\nabla F)^{-1} \nabla (u-u_h)\, dx\, dt.\end{aligned} \tag{3.2.73}$$

Using equation (3.1.2) we obtain that

$$\int_{\Omega_T} \rho \partial_t F (\nabla F)^{-1} \nabla (u-u_h)\, dx\, dt = -\mathscr{A}_T[\rho, u-u_h] - \sum_{i,j=1}^{n} \int_{\Omega_T} \tilde{\rho} Q_{i,j} \partial_i \partial_j (\tilde{u} - \tilde{u}_h)\, dy\, dt. \tag{3.2.74}$$

□

Lemma 3.2.13.

$$\Big\| \frac{\partial_t \tilde{u}_h}{|\det(\nabla F)|^{\frac{1}{2}} \circ F^{-1}} \Big\|^2_{L^2(\Omega_T)} \leq 2\|g\|^2_{L^2(\Omega_T)} + C\|\tilde{u}_h\|^2_{L^2(0,T,W^{2,2}_D(\Omega))} \tag{3.2.75}$$

where the constant C depends on n, $\lambda_{\max}(a)$, $\|\operatorname{Trace}[\sigma]\|_{L^\infty(\Omega_T)}$, and μ_σ.

Proof. Using the change of variable $y = F(x,t)$ in (3.1.29) we obtain that for all $\varphi \in X_h$

$$\begin{cases} (\varphi, \frac{\partial_t \tilde{u}_h}{|\det(\nabla F)|\circ F^{-1}})_{L^2(\Omega)} = & \sum_{i,j=1}^n \int_\Omega (\varphi, Q_{i,j}\partial_i\partial_j \tilde{u}_h)_{L^2(\Omega)} \\ & +(\varphi, \frac{\tilde{g}}{|\det(\nabla F)|\circ F^{-1}})_{L^2(\Omega)} \\ \tilde{u}_h(x,0) = 0. \end{cases} \tag{3.2.76}$$

Recall that Q is given by (3.2.59). We choose $\varphi = \partial_t \tilde{u}_h$ and observe that

$$\frac{\sigma}{|\det \nabla F|^{\frac{1}{2}}} = \frac{\sigma}{|\det \sigma|^{\frac{1}{4}}} |\det a|^{\frac{1}{4}} \tag{3.2.77}$$

thus

$$\Big\| \frac{\sigma}{|\det \nabla F|^{\frac{1}{2}}} \Big\|_{L^\infty(\Omega_T)} \leq C\big(n, \lambda_{\max}(a), \| \operatorname{Trace}[\sigma]\|_{L^\infty(\Omega_T)}, \mu_\sigma\big). \tag{3.2.78}$$

We deduce the lemma by Minkowski inequality. □

Combining Lemma 3.2.12 and Lemma 3.2.13 we obtain the following lemma:

Lemma 3.2.14.

$$\begin{aligned} \frac{1}{2}\big\|(u-u_h)(T)\big\|^2_{L^2(\Omega)} + \mathscr{A}_T[u-u_h] \leq & C\|\rho\|_{L^2(\Omega_T)}\Big(\|g\|_{L^2(\Omega_T)} \\ & + \|\tilde{u}_h\|_{L^2(0,T,W^{2,2}_D(\Omega))}\Big) \end{aligned} \tag{3.2.79}$$

where the constant C depends on n, $\lambda_{\max}(a)$, $\| \operatorname{Trace}[\sigma]\|_{L^\infty(\Omega_T)}$, μ_σ.

Lemma 3.2.15. *Assume that Ω is convex, and condition 3.1.1 is satisfied, then*

$$\|\rho\|_{L^2(\Omega_T)} \leq Ch^2 \|g\|_{L^2(\Omega_T)}. \tag{3.2.80}$$

The constant C depends on C_X, n, Ω, δ, ε, $\lambda_{\min}(a)$, and $\lambda_{\max}(a)$.

Proof. The proof is similar to that for Lemma 3.2.11. □

Lemma 3.2.16. *Assume Ω is convex and* $\mathrm{Trace}[\sigma] \in L^\infty(\Omega_T)$*, we have*

$$\|\tilde{u}_h\|_{L^2(0,T,W_D^{2,2}(\Omega))} \leq \frac{C}{h}\|g\|_{L^2(\Omega_T)}. \tag{3.2.81}$$

The constant C depends on C_X, n, Ω, $\lambda_{\min}(a)$, $\lambda_{\max}(a)$, and $\|\mathrm{Trace}[\sigma]\|_{L^\infty(\Omega_T)}$.

Proof. Using the inverse inequality (3.1.24) we obtain that

$$\|\tilde{u}_h\|_{L^2(0,T,W_D^{2,2}(\Omega))} \leq \frac{C_X}{h}\|\nabla\tilde{u}_h\|_{L^2(0,T,W_D^{2,2}(\Omega))}. \tag{3.2.82}$$

Using the change of variables $y = F(x)$ it follows that,

$$\|\nabla\tilde{u}_h\|^2_{L^2(0,T,W_D^{2,2}(\Omega))} \leq C\mathscr{A}_T[u_h] \tag{3.2.83}$$

where C depends on n, $\lambda_{\min}(a)$ and $\|\mathrm{Trace}[\sigma]\|_{L^\infty(\Omega_T)}$. We deduce the lemma by using Lemma 3.2.7. □

Now Theorem 3.1.9 is a straightforward application of Lemma 3.2.14, Lemma 3.2.15 and Lemma 3.2.16.

3.2.2.3 Numerical homogenization in time and space: Proof of Theorem 3.1.10–3.1.12

We use the notation of subsection 3.1.3. First let us show that the numerical scheme (3.1.35) is stable. Indeed, choosing $\psi = v_{n+1}(t)$ one gets

$$\begin{aligned}\big|v_{n+1}(t_{n+1})\big|^2_{L^2(\Omega)} =& \big(v_{n+1}(t_n), v_n(t_n)\big)_{L^2(\Omega)} + \frac{1}{2}\Big(\big|v_{n+1}(t_{n+1})\big|^2_{L^2(\Omega)} - \big|v_{n+1}(t_n)\big|^2_{L^2(\Omega)}\Big)\\ &- \int_{t_n}^{t_{n+1}} \Big(a\big[v_{n+1}(t)\big] + \big(v_{n+1}(t), g(t)\big)_{L^2(\Omega)}\Big)\,dt.\end{aligned} \tag{3.2.84}$$

It follows by Cauchy-Schwartz inequality that

$$\begin{aligned}\frac{1}{2}\left|v_{n+1}(t_{n+1})\right|^2_{L^2(\Omega)} \leq & \frac{1}{2}\left|v_n(t_n)\right|^2_{L^2(\Omega)} - \int_{t_n}^{t_{n+1}} \Big(a\big[v_{n+1}(t)\big] \\ & + \big(v_{n+1}(t), g(t)\big)_{L^2(\Omega)} \Big)\, dt.\end{aligned} \tag{3.2.85}$$

Hence using Poincaré and Minkowski inequalities one obtains that

$$\begin{aligned}\left|v_{n+1}(t_{n+1})\right|^2_{L^2(\Omega)} + \int_{t_n}^{t_{n+1}} a\big[v_{n+1}(t)\big]\, dt \leq & \left|v_n(t_n)\right|^2_{L^2(\Omega)} \\ & + \frac{C_{n,\Omega}}{\lambda_{\min}(a)} \int_{t_n}^{t_{n+1}} \left|g(t)\right|^2_{L^2(\Omega)} dt\end{aligned} \tag{3.2.86}$$

which implies Theorem 3.1.10 and the stability of the scheme.

Integrating (3.1.29) with respect to time over $[t_n, t_{n+1}]$ we obtain that for $\psi \in U_T^h$,

$$\begin{aligned}\big(\psi(t_{n+1}), u_h(t_{n+1})\big)_{L^2(\Omega)} = & \big(\psi(t_n), u_h(t_n)\big)_{L^2(\Omega)} + \int_{t_n}^{t_{n+1}} \Big(\big(\partial_t \psi(t), u_h(t)\big)_{L^2(\Omega)} \\ & - a\big[\psi(t), u_h(t)\big] + \big(\psi(t), g(t)\big)_{L^2(\Omega)} \Big)\, dt.\end{aligned} \tag{3.2.87}$$

Let us write (u^i) the coordinates of u_h associated to the basis $\varphi_i \circ F$, i.e.,

$$u_h(x,t) := \sum_i u^i(t)\varphi_i(F(x,t)). \tag{3.2.88}$$

Define

$$u_n(x,t) := \sum_i u^i(t_n)\varphi_i(F(x,t)). \tag{3.2.89}$$

Subtracting (3.2.87) and (3.1.35) we obtain that for $\psi \in Z_T^h$,

$$\begin{aligned}\big(\psi(t_{n+1}), (u_{n+1} - v_{n+1})(t_{n+1})\big)_{L^2(\Omega)} = & \big(\psi(t_n), (u_n - v_n)(t_n)\big)_{L^2(\Omega)} \\ & + \int_{t_n}^{t_{n+1}} \Big(\big(\partial_t \psi(t), (u_h - v_{n+1})(t)\big)_{L^2(\Omega)} \\ & - a\big[\psi(t), (u_h - v_{n+1})(t)\big] \Big)\, dt.\end{aligned} \tag{3.2.90}$$

Choosing $\psi = u_{n+1} - v_{n+1}$, by Cauchy-Schwartz inequality we deduce

$$\begin{aligned}&\frac{1}{2}\big\|(u_{n+1}-v_{n+1})(t_{n+1})\big\|^2_{L^2(\Omega)} + \int_{t_n}^{t_{n+1}} a\big[(u_{n+1}-v_{n+1})(t)\big]\,dt \leq \\ &\quad \frac{1}{2}\big\|(u_n-v_n)(t_n)\big\|^2_{L^2(\Omega)} + \int_{t_n}^{t_{n+1}} \Big(\big(\partial_t(u_{n+1}-v_{n+1})(t), (u_h-u_{n+1})(t)\big)_{L^2(\Omega)} \\ &\quad - a\big[(u_{n+1}-v_{n+1})(t), (u_h-u_{n+1})(t)\big]\Big)\,dt.\end{aligned} \tag{3.2.91}$$

In the following, we will discuss two different situations where a is time independent or time dependent.

Time independent medium. If the medium is time independent then (3.2.91) can be written as

$$\begin{aligned}&\frac{1}{2}\big\|(u_{n+1}-v_{n+1})(t_{n+1})\big\|^2_{L^2(\Omega)} + \int_{t_n}^{t_{n+1}} a\big[(u_{n+1}-v_{n+1})(t)\big]\,dt \leq \\ &\quad \frac{1}{2}\big\|(u_n-v_n)(t_n)\big\|^2_{L^2(\Omega)} - \int_{t_n}^{t_{n+1}} a\big[(u_{n+1}-v_{n+1})(t), (u_h-u_{n+1})(t)\big]\,dt\end{aligned} \tag{3.2.92}$$

which leads us to

$$\begin{aligned}&\frac{1}{2}\big\|(u_{n+1}-v_{n+1})(t_{n+1})\big\|^2_{L^2(\Omega)} + \int_{t_n}^{t_{n+1}} a\big[(u_{n+1}-v_{n+1})(t)\big]\,dt \leq \\ &\quad \frac{1}{2}\big\|(u_n-v_n)(t_n)\big\|^2_{L^2(\Omega)} + \int_{t_n}^{t_{n+1}} \int_{t_n}^{t_{n+1}} 1(t<s)\, a\big[(u_{n+1}-v_{n+1})(t), \partial_s u_h(s)\big]\,ds dt.\end{aligned} \tag{3.2.93}$$

Write $\Delta t := t_{n+1} - t_n$. Using Minkowski inequality we obtain that

$$\begin{aligned}a\big[(u_{n+1}-v_{n+1})(t), \partial_s u_h(s)\big] \leq &\frac{1}{2\Delta t} a\big[(u_{n+1}-v_{n+1})(t)\big] \\ &+ \frac{\Delta t}{2} a\big[\partial_s u_h(s)\big].\end{aligned} \tag{3.2.94}$$

It follows from (3.2.93) that

$$\begin{aligned}&\big|(u_{n+1}-v_{n+1})(t_{n+1})\big|^2_{L^2(\Omega)} + \int_{t_n}^{t_{n+1}} a\big[(u_{n+1}-v_{n+1})(t)\big]\,dt \leq \\ &\big|(u_n-v_n)(t_n)\big|^2_{L^2(\Omega)} + |\Delta t|^2 \int_{t_n}^{t_{n+1}} a\big[\partial_s u_h(s)\big]\,ds.\end{aligned} \tag{3.2.95}$$

Observing that

$$\int_{t_n}^{t_{n+1}} a\big[(u_{n+1}-v_{n+1})(t)\big]\,dt \geq 0.5\int_{t_n}^{t_{n+1}} a\big[(u_h-v_{n+1})(t)\big]\,dt - \int_{t_n}^{t_{n+1}} a\big[(u_h-u_{n+1})(t)\big]\,dt, \tag{3.2.96}$$

and

$$\int_{t_n}^{t_{n+1}} a\big[(u_h-u_{n+1})(t)\big]\,dt \leq \frac{\Delta t^2}{2}\int_{t_n}^{t_{n+1}} a\big[\partial_s u_h(s)\big]\,ds. \tag{3.2.97}$$

We obtain that

$$\big\|(u_{n+1}-v_{n+1})(t_{n+1})\big\|^2_{L^2(\Omega)} + \frac{1}{2}\int_{t_n}^{t_{n+1}} a\big[(u_h-v_{n+1})(t)\big]\,dt \leq \big\|(u_n-v_n)(t_n)\big\|^2_{L^2(\Omega)} + \frac{3}{2}\Delta t^2\int_{t_n}^{t_{n+1}} a\big[\partial_s u_h(s)\big]\,ds. \tag{3.2.98}$$

In conclusion we have obtained the following lemma:

Lemma 3.2.17. *Let $v \in Z_T^h$ be the solution of (3.1.35). We have*

$$\big\|(u_h-v)(T)\big\|^2_{L^2(\Omega)} + \frac{1}{2}\int_0^T a\big[(u_h-v)(t)\big]\,dt \leq \frac{3}{2}\Delta t^2\int_0^T a\big[\partial_s u_h(s)\big]\,ds. \tag{3.2.99}$$

Combining Lemma 3.2.3 with Lemma 3.2.17 we obtain Theorem 3.1.11.

Time dependent medium. Observe that

$$\partial_t(u_{n+1}-v_{n+1}) = \partial_t F(\nabla F)^{-1}\nabla(u_{n+1}-v_{n+1}). \tag{3.2.100}$$

It follows after writing $\partial_t F = \operatorname{div} a\nabla F$, integration by parts, and using the change of

variables $y = F(x,t)$ in (3.2.91) that

$$\begin{aligned}
&\frac{1}{2}\big\|(u_{n+1}-v_{n+1})(t_{n+1})\big\|^2_{L^2(\Omega)} + \int_{t_n}^{t_{n+1}} a\big[(u_{n+1}-v_{n+1})(t)\big]\,dt \leq \\
&\frac{1}{2}\big\|(u_n-v_n)(t_n)\big\|^2_{L^2(\Omega)} - 2\int_{t_n}^{t_{n+1}} a\big[(u_{n+1}-v_{n+1})(t),(u_h-u_{n+1})(t)\big]\,dt \\
&-\sum_{i,j}\int_{t_n}^{t_{n+1}}\int_{\Omega}(\tilde{u}_h-\tilde{u}_{n+1})Q_{i,j}\partial_i\partial_j(\tilde{u}_{n+1}-\tilde{v}_{n+1})\,dt\,dy.
\end{aligned} \tag{3.2.101}$$

Hence using Minkowski inequality we obtain that

$$\begin{aligned}
&\big\|(u_{n+1}-v_{n+1})(t_{n+1})\big\|^2_{L^2(\Omega)} + \int_{t_n}^{t_{n+1}} a\big[(u_{n+1}-v_{n+1})(t)\big]\,dt \leq \\
&\big\|(u_n-v_n)(t_n)\big\|^2_{L^2(\Omega)} + 4\int_{t_n}^{t_{n+1}} a\big[(u_h-u_{n+1})(t)\big]\,dt \\
&-2\sum_{i,j}\int_{t_n}^{t_{n+1}}\int_{\Omega}(\tilde{u}_h-\tilde{u}_{n+1})Q_{i,j}\partial_i\partial_j(\tilde{u}_{n+1}-\tilde{v}_{n+1})\,dt\,dy.
\end{aligned} \tag{3.2.102}$$

Again using Minkowski inequality it follows that

$$\begin{aligned}
&\Big|\sum_{i,j}\int_{t_n}^{t_{n+1}}\int_{\Omega}(\tilde{u}_h-\tilde{u}_{n+1})Q_{i,j}\partial_i\partial_j(\tilde{u}_{n+1}-\tilde{v}_{n+1})\,dt\,dy\Big| \leq \\
&\frac{1}{2}C_A N^2\int_{t_n}^{t_{n+1}}\int_{\Omega}|\tilde{u}_h-\tilde{u}_{n+1}|^2\,dt\,dy + \frac{\lambda_{\max}(Q)}{2C_A}\int_{t_n}^{t_{n+1}}\sum_{i,j}\int_{\Omega}|\partial_i\partial_j(\tilde{u}_{n+1}-\tilde{v}_{n+1})|^2\,dt\,dy.
\end{aligned} \tag{3.2.103}$$

Using the inverse inequality (3.1.24) and the change of variable $y = F(x,t)$ we obtain that

$$\int_{t_n}^{t_{n+1}}\sum_{i,j}\int_{\Omega}|\partial_i\partial_j(\tilde{u}_{n+1}-\tilde{v}_{n+1})|^2\,dt\,dy \leq \frac{C_X}{h^2\lambda_{\min}(Q)}\int_{t_n}^{t_{n+1}} a\big[(u_{n+1}-v_{n+1})(t)\big]\,dt. \tag{3.2.104}$$

Choosing $C_A = \frac{C_X \lambda_{\max}(Q)}{h^2 \lambda_{\min}(Q)}$, we have

$$\begin{aligned}
&\big|(u_{n+1}-v_{n+1})(t_{n+1})\big|^2_{L^2(\Omega)} + \frac{1}{2}\int_{t_n}^{t_{n+1}} a\big[(u_{n+1}-v_{n+1})(t)\big]\,dt \leq \\
&\big|(u_n-v_n)(t_n)\big|^2_{L^2(\Omega)} + 4\int_{t_n}^{t_{n+1}} a\big[(u_h-u_{n+1})(t)\big]\,dt \\
&+ \frac{C_X \lambda_{\max}(Q)}{2h^2\lambda_{\min}(Q)} N^2 \int_{t_n}^{t_{n+1}}\int_\Omega |\tilde{u}_h - \tilde{u}_{n+1}|^2\,dt\,dy.
\end{aligned} \tag{3.2.105}$$

Using (3.2.96) gives us

$$\begin{aligned}
&\big|(u_{n+1}-v_{n+1})(t_{n+1})\big|^2_{L^2(\Omega)} + \frac{1}{4}\int_{t_n}^{t_{n+1}} a\big[(u_h-v_{n+1})(t)\big]\,dt \leq \\
&\big|(u_n-v_n)(t_n)\big|^2_{L^2(\Omega)} + \frac{9}{2}\int_{t_n}^{t_{n+1}} a\big[(u_h-u_{n+1})(t)\big]\,dt \\
&+ \frac{C_X \lambda_{\max}(Q)}{2h^2\lambda_{\min}(Q)} N^2 \int_{t_n}^{t_{n+1}}\int_\Omega |\tilde{u}_h - \tilde{u}_{n+1}|^2\,dt\,dy.
\end{aligned} \tag{3.2.106}$$

Moreover using the change of variables $F(x,t) = y$ and the inverse inequality (3.1.25) we deduce that

$$\int_{t_n}^{t_{n+1}} a\big[(u_h-u_{n+1})(t)\big]\,dt \leq \frac{C_X \lambda_{\max}(Q)}{h^2}\int_{t_n}^{t_{n+1}}\int_\Omega |\tilde{u}_h - \tilde{u}_{n+1}|^2\,dt\,dy. \tag{3.2.107}$$

Let us observe that

$$\int_{t_n}^{t_{n+1}}\int_\Omega |\tilde{u}_h - \tilde{u}_{n+1}|^2\,dt\,dy \leq \frac{1}{2}\Delta t^2 \int_{t_n}^{t_{n+1}}\int_\Omega |\partial_t \tilde{u}_h|^2\,dt\,dy. \tag{3.2.108}$$

It follows

$$\begin{aligned}
&\big|(u_{n+1}-v_{n+1})(t_{n+1})\big|^2_{L^2(\Omega)} + \frac{1}{4}\int_{t_n}^{t_{n+1}} a\big[(u_h-v_{n+1})(t)\big]\,dt \leq \\
&\big|(u_n-v_n)(t_n)\big|^2_{L^2(\Omega)} + C_B\frac{|\Delta t|^2}{h^2}\int_{t_n}^{t_{n+1}}\int_\Omega |\partial_t \tilde{u}_h|^2\,dy\,dt
\end{aligned} \tag{3.2.109}$$

with

$$C_B = C_X \lambda_{\max}(Q)\Big(\frac{9}{2} + \frac{N^2}{4\lambda_{\min}(Q)}\Big). \tag{3.2.110}$$

We deduce,

$$\left\|(u_h - v)(T)\right\|^2_{L^2(\Omega)} + \frac{1}{4}\int_0^T a\left[(u_h - v)(t)\right] dt \leq C_B \frac{|\Delta t|^2}{h^2} \int_0^T \int_\Omega |\partial_t \tilde{u}_h|^2 \, dt \, dy. \quad (3.2.111)$$

Using Lemma 3.2.4 to control C_B and combining (3.2.111) with Theorem 3.1.5 we obtain Theorem 3.1.12.

3.3 Numerical Experiments

The purpose of this section is to give several illustrations of the implementation and performance of the numerical homogenization method. The computational domain is the unit square in dimension two. Equation (3.1.1) is solved on a fine tessellation characterized by 16129 interior nodes (degrees of freedom).

In the time independent case, for instance, a is a constant matrix in each triangle of a given fine mesh of Ω. In the time dependent case, on each fine mesh triangle a is step function from $[0,T]$ into the set of positive definite symmetric matrices. The caloric coordinates F have been computed on the same fine triangulation of Ω associated to a through a standard finite element method.

Three different coarse tessellations with 9, 49, and 225 degrees of freedom (*dof*) are considered. The parabolic operator associated to equation (3.1.1) has been homogenized over these coarse meshes using the method introduced in subsection 3.1.3.. We have chosen splines to be the space X^h introduced in subsection 3.1.2.

3.3.1 Time Independent Medium.

Example 3.3.1. *Time independent site percolation.*

In this example we consider the site percolating medium associated to figure 3.1. (3.1.1) has been solved with $g = 1$ and $g = \sin(2.4x - 1.8y + 2\pi t)$. The fine mesh and coarse mesh errors are given in tables 3.1, 3.2, 3.3, and 3.4.

Table 3.1: Coarse mesh error: Time independent site percolation with $g \equiv 1$

dof	L^1	L^∞	L^2	H^1
9	0.0142	0.0389	0.0168	0.0366
49	0.0077	0.0450	0.0101	0.0482
225	0.0035	0.0228	0.0060	0.0293

Table 3.2: Fine mesh error: Time independent site percolation with $g \equiv 1$

dof	L^1	L^∞	L^2	H^1
9	0.0196	0.0843	0.0251	0.1193
49	0.0136	0.0698	0.0184	0.1028
225	0.0040	0.0243	0.0070	0.0485

Table 3.3: Coarse mesh error: Time independent site percolation with $g = \sin(2.4x - 1.8y + 2\pi t)$

dof	L^1	L^∞	L^2	H^1
9	0.0278	0.0400	0.0274	0.0377
49	0.0336	0.0612	0.0324	0.0619
225	0.0321	0.0492	0.0289	0.0710

Table 3.4: Fine mesh error: Time independent site percolation with $g = \sin(2.4x - 1.8y + 2\pi t)$.

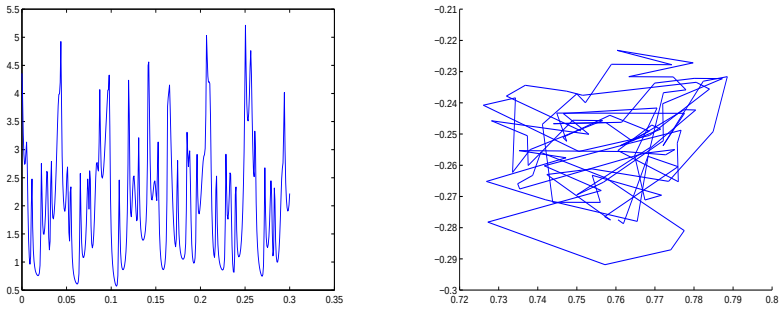
dof	L^1	L^∞	L^2	H^1
9	0.0418	0.1099	0.0428	0.1287
49	0.0390	0.0907	0.0383	0.1177
225	0.0174	0.0544	0.0318	0.0977

3.3.2 Time Dependent Medium

In the following examples we consider media characterized by a continuum of time scales. The solution is obtained by the numerical homogenization method described in subsection 3.1.3.

Example 3.3.2. *Time dependent multi-scale trigonometric*

In this example a is given by equation (3.1.22). Although the number of fine time steps to solve (3.1.1) is 2663, only 134 coarse time steps have been used to solve the homogenized formulation. If one also takes into account homogenization in space, the compression factor is of the order of 35000 for the coarse mesh with 9 interior nodes. Figure 3.7 shows the curves of $t \longrightarrow a(x_0,t)$ and $t \longrightarrow F(x_0,t)$ for a given $x_0 \in \Omega$.



(a) $t \longrightarrow a(x_0,t)$ **(b)** ***Top view of*** $t \longrightarrow F(x_0,t)$

Figure 3.7: Multi-scale time dependent trigonometric medium

The coarse and fine mesh relative L^1, L^2, L^∞, and H^1 errors with respect to time have been plotted in figures 3.8, 3.9, 3.10, and 3.11. The initial boost of the relative error is due to the initial value $u(x,0) \equiv 0$.

The coarse and fine mesh errors at $t = 0.1$ are given in tables 3.5 and 3.6 for $g \equiv 1$, for $g = \sin(2.4x - 1.8y + 2\pi t)$ the errors are given in tables 3.7 and 3.8.

Table 3.5: Coarse mesh error: Multi-scale trigonometric time dependent medium, $g = 1$

dof	L^1	L^∞	L^2	H^1
9	0.0018	0.0045	0.0019	0.0039
49	0.0012	0.0054	0.0015	0.0060

Table 3.6: Fine mesh error: Multi-scale trigonometric time dependent medium, $g = 1$

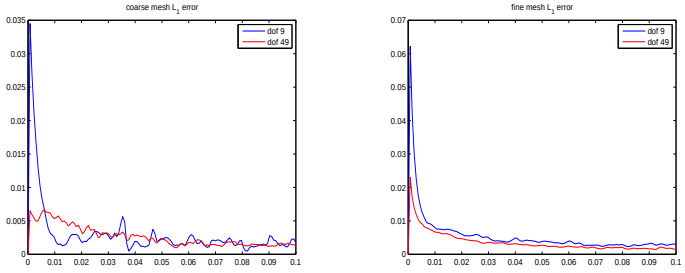
dof	L^1	L^∞	L^2	H^1
9	0.0031	0.0096	0.0034	0.0242
49	0.0014	0.0059	0.0016	0.0166

Table 3.7: Coarse mesh error: Multi-scale trigonometric time dependent medium, $g = \sin(2.4x - 1.8y + 2\pi t)$

dof	L^1	L^∞	L^2	H^1
9	0.0043	0.0087	0.0044	0.0085
49	0.0033	0.0079	0.0035	0.0084

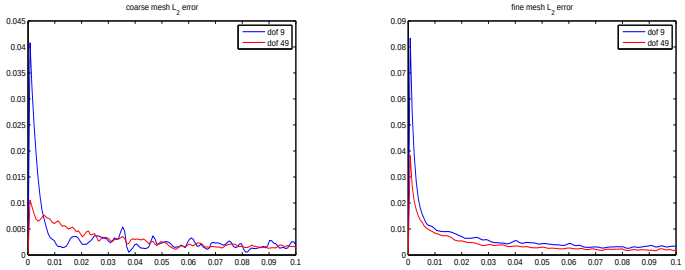
Table 3.8: Fine mesh error: Multi-scale trigonometric time dependent medium, $g = \sin(2.4x - 1.8y + 2\pi t)$

dof	L^1	L^∞	L^2	H^1
9	0.0082	0.0199	0.0087	0.0379
49	0.0038	0.0104	0.0040	0.0244



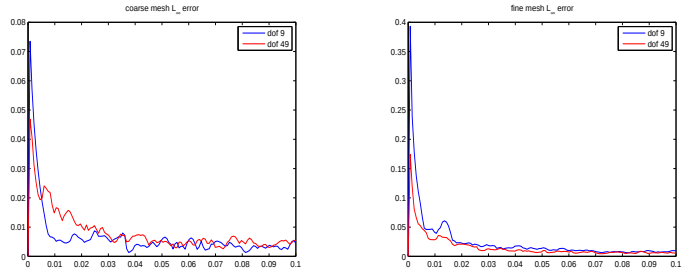
(a) *Coarse mesh* L^1 *error* **(b) *Fine mesh* L^1 *error***

Figure 3.8: L^1 error: Multi-scale trigonometric time dependent medium



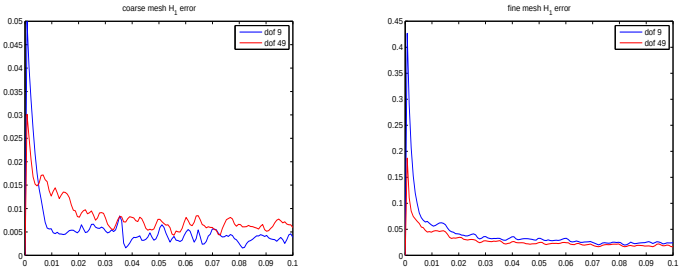
(a) *Coarse mesh* L^2 *error* **(b) *Fine Mesh* L^2 *error***

Figure 3.9: L^2 error: Multi-scale trigonometric time dependent medium



(a) *Coarse mesh* L^∞ *error* **(b) *Fine mesh* L^∞ *error***

Figure 3.10: L^∞ error: Multi-scale trigonometric time dependent medium



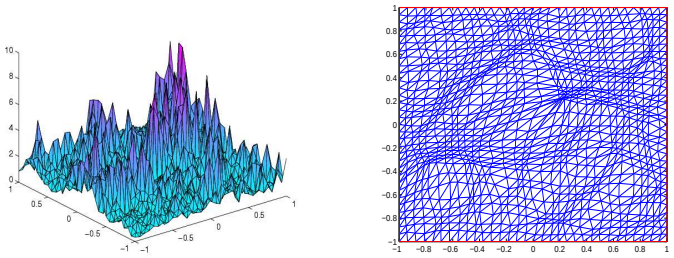
(a) *Coarse Mesh* H^1 *error* **(b) *Fine Mesh* H^1 *error***

Figure 3.11: H^1 error: Multi-scale trigonometric time dependent medium

Example 3.3.3. *Time dependent random fractal*

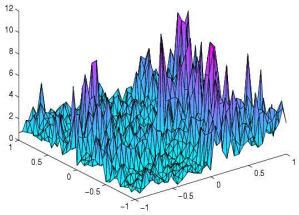
In this case, a is given by a product of discontinuous functions oscillating randomly at different scales, $a(x,t) = a_1(x,t)a_2(x,t)\cdots a_n(x,t)$, and $a_i(x,t) = c_{pq}$ for $x \in [\frac{p}{2^i}, \frac{p+1}{2^i}) \times [\frac{q}{2^i}, \frac{q+1}{2^i})$ in the time interval $0.1 \times [\frac{k}{4^i}, \frac{k+1}{4^i})$. c_{pq} is uniformly random in $[\frac{1}{\gamma}, \gamma]$, $n = 6$ and $\gamma = 0.7$. In this example, we have $\frac{\lambda_{\max}(a)}{\lambda_{\min}(a)} = 160.3295$. The number of fine time steps are 3482, and the number of coarse time steps are 175.

The time-dependent media and (F_1, F_2) is drawn in the following figure 3.12.

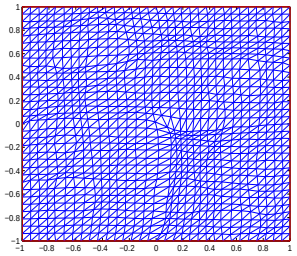


(a) a ***at*** $t = 0$

(b) (F_1, F_2) ***at*** $t = 0$



(c) a ***at*** $t = 0.1$



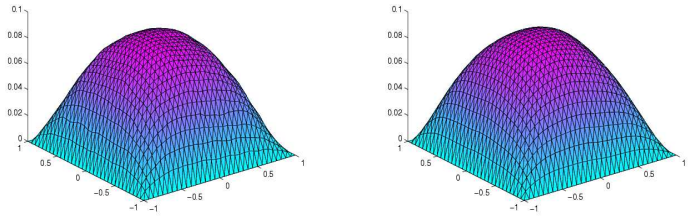
(d) (F_1, F_2) ***at*** $t = 0.1$

Figure 3.12: a and (F_1, F_2) at time $t = 0$, $t = 0.1$ for time dependent random fractal medium

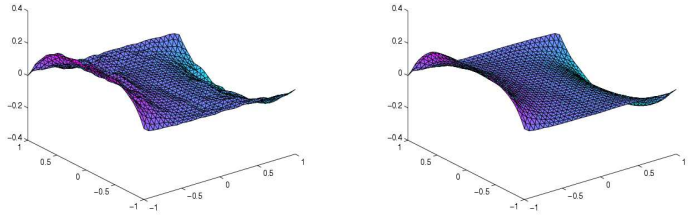
Figures 3.13 is the plot of u, $\tilde{u}$, u_x and $\tilde{u}_x$ at time $t = 0.1$.

We illustrate L^1, L^2, L^∞, and H^1 errors with respect to time in figure 3.14-3.17.

In the following tables, we show the coarse mesh error (table 3.9) and fine mesh error (table 3.10) at time $t = 0.1$.



(a) u **(b)** $\tilde{u}$



(c) u_x **(d)** $\tilde{u}_x$

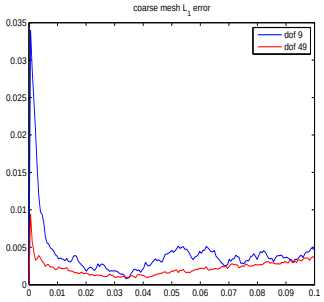
Figure 3.13: u, $\tilde{u}$, u_x, and $\tilde{u}_x$ at time $t = 0.1$ for time dependent random fractal medium

Table 3.9: Coarse mesh error of random fractal case with spline elements

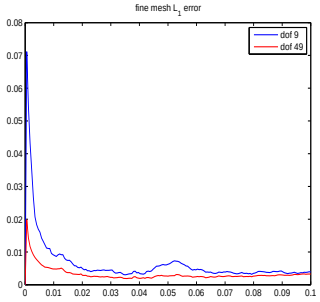
dof	L^1	L^∞	L^2	H^1
9	0.0046	0.0074	0.0052	0.0065
49	0.0036	0.0046	0.0036	0.0059

Table 3.10: Fine mesh error of random fractal case with spline elements

dof	L^1	L^∞	L^2	H^1
9	0.0039	0.0082	0.0043	0.0222
49	0.0033	0.0054	0.0034	0.0168

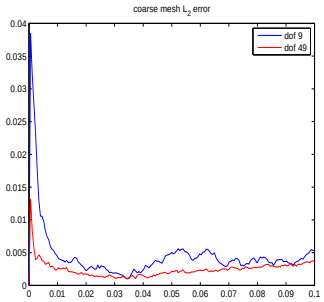


(a) *coarse mesh* L^1 *error*

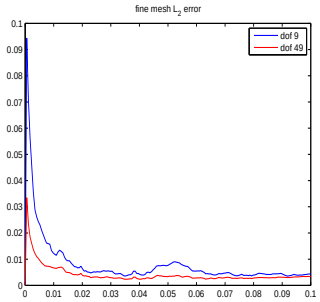


(b) *fine mesh* L^1 *error*

Figure 3.14: L^1 error for time dependent random fractal medium at $t = .1$

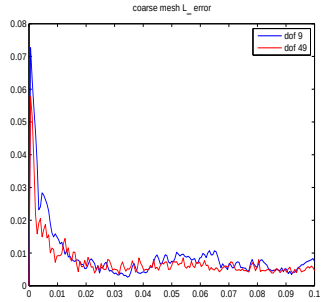


(a) *coarse mesh* L^2 *error*

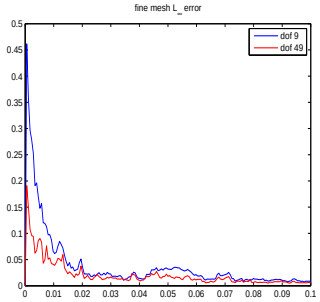


(b) *Fine Mesh* L^2 *error*

Figure 3.15: L^2 error for time dependent random fractal medium $t = .1$

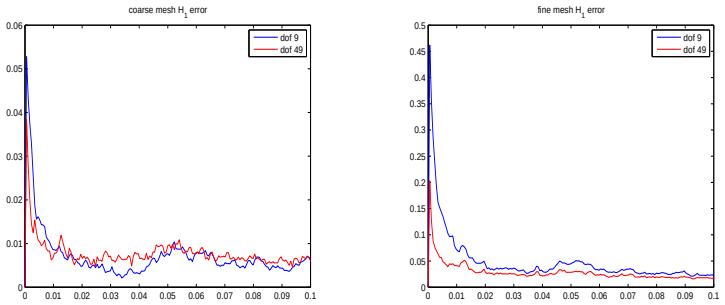


(a) *coarse mesh* L^∞ *error*



(b) *fine mesh* L^∞ *error*

Figure 3.16: L^∞ error for time dependent random fractal medium $t = .1$.



(a) *coarse Mesh* H^1 *error* **(b) *fine Mesh* H^1 *error***

Figure 3.17: H^1 error for time dependent random fractal medium $t = .1$

Chapter 4

Numerical Homogenization for Acoustic Wave Equation with Continuum Space Scales

In this chapter, we apply the metric based upscaling method to acoustic wave equation in dimension n in situations where the bulk modulus and the density of the medium are only bounded. In section 4.1, we show that under a Cordes type condition the second order derivatives of the solution with respect to harmonic coordinates are in L^2 (instead of H^{-1} with respect to Euclidean coordinates). It follows that it is possible to homogenize the wave equation numerically without assumptions of scale separation or ergodicity by pre-computing n solutions of the associated elliptic equation. In section 4.2, we give proofs of compensation phenomena and convergence of the numerical upscaling method. In section 4.3, we show the numerical experiment results.

4.1 Main Results

Let $\Omega \subset \mathbb{R}^n$ be a bounded and convex domain of class C^2. Let $T > 0$. Consider the following hyperbolic partial differential equation

$$\begin{cases} K^{-1}(x)\partial_t^2 u = \operatorname{div}(\rho^{-1}(x)\nabla u(x,t)) + g \quad \text{in} \quad \Omega \times (0,T), \\ u(x,t) = 0 \quad \text{for} \quad (x,t) \in \partial\Omega \times (0,T), \\ u(x,t) = u(x,0) \quad \text{for} \quad (x,t) \in \Omega \times \{t=0\}, \\ \partial_t u(x,t) = u_t(x,0) \quad \text{for} \quad (x,t) \in \Omega \times \{t=0\}. \end{cases} \tag{4.1.1}$$

Write $\Omega_T := \Omega \times (0,T)$ and $a := \rho^{-1}$. We assume that a is a $n \times n$ symmetric, uniformly elliptic matrix on the closure of Ω whose entries are in $L^\infty(\Omega)$. K is assumed to be scalar such that K and K^{-1} belong to $L^\infty(\Omega)$. We assume that g belongs to $L^2(\Omega_T)$.

4.1.1 Compensation phenomena

Since the media is time independent, we use the same notation as in Chapter 2 and focus on space dimension $n = 2$. The extension to higher dimension is straightforward conditioned on the stability of σ. Let $F := (F_1, F_2)$ be the harmonic coordinates satisfying

$$\begin{cases} \operatorname{div} a\nabla F = 0 \quad \text{in} \quad \Omega, \\ F(x) = x \quad \text{on} \quad \partial\Omega. \end{cases} \tag{4.1.2}$$

Let $\sigma := {}^t\nabla F a \nabla F$ and

$$\mu_\sigma := \operatorname{esssup}_{x\in\Omega}\Big(\frac{\lambda_{\max}\big(\sigma(x)\big)}{\lambda_{\min}\big(\sigma(x)\big)}\Big). \tag{4.1.3}$$

We assume that σ satisfies Cordes type condition $\mu_\sigma < \infty$ and $(\operatorname{Trace}[\sigma])^{-1} \in L^\infty(\Omega)$. We also write

$$\|v\|_{L^\infty(0,T,H^2(\Omega))} = \operatorname{esssup}_{0\le t\le T}\Big(\int_\Omega \sum_{i,j}\big(\partial_i\partial_j v(x,t)\big)^2\,dx\Big)^{\frac{1}{2}}. \tag{4.1.4}$$

Throughout the chapter, we made the following assumptions

Assumption A. Assume that the g satisfies $\partial_t g \in L^2(\Omega_T)$, $g \in L^\infty(0,T,L^2(\Omega))$, initial data $u(x,0)$ and $\partial_t u(x,0)$ satisfy $\partial_t u(x,0) \in H^1(\Omega)$ and $\nabla a(x)\nabla u(x,0) \in L^2(\Omega)$ or equivalently $\partial_t^2 u(x,0) \in L^2(\Omega)$.

Theorem 4.1.1. *Suppose that **Assumption A** holds, then $u \circ F^{-1} \in L^\infty(0,T,H^2(\Omega))$ and*

$$\begin{aligned}\|u\circ F^{-1}\|_{L^\infty(0,T,H^2(\Omega))} \leq & C\big(\|g\|_{L^\infty(0,T,L^2(\Omega))} + \|\partial_t g\|_{L^2(\Omega_T)} + \|\partial_t u(x,0)\|_{H^1(\Omega)} \\ & + \|\partial_t^2 u(x,0)\|_{L^2(\Omega)}\big).\end{aligned} \tag{4.1.5}$$

The constant C can be written

$$C = \frac{C(n,\Omega,K_{\min},K_{\max},\lambda_{\min}(a),\lambda_{\max}(a))}{(1-\beta_\sigma^{\frac{1}{2}})}\big\|(\mathrm{Trace}[\sigma])^{\frac{n}{4}-1}\big\|_{L^\infty(\Omega_T)}. \tag{4.1.6}$$

We use the notation

$$K_{\max} := \|K\|_{L^\infty(\Omega)} \quad \text{and} \quad K_{\min} := (\|K^{-1}\|_{L^\infty(\Omega)})^{-1}. \tag{4.1.7}$$

Remark. *The condition $g \in L^2(\Omega_T)$ is sufficient to obtain Theorem 4.1.1 and the following theorems. For the sake of clarity we have preferred to restrict ourselves to $g \in L^\infty(0,T,L^2(\Omega))$.*

4.1.2 Numerical homogenization in space

We will largely use the notation similar to parabolic equation in Section 3.1.2. Suppose we have a quasi-uniform mesh. Let X^h be a finite dimensional subspace of $H_0^1(\Omega)\cap W^{1,\infty}(\Omega)$ with the following approximation properties: There exists a constant C_X such that

- Interpolation property, i.e., for all $f \in H^2(\Omega)\cap H_0^1(\Omega)$

$$\inf_{v\in X^h}\|f-v\|_{H_0^1(\Omega)} \leq C_X h\|f\|_{H^2(\Omega)}. \tag{4.1.8}$$

- Inverse Sobolev inequality, i.e., for all $v \in X^h$,

$$\|v\|_{H^2(\Omega)} \leq C_X h^{-1} \|v\|_{H_0^1(\Omega)}, \tag{4.1.9}$$

and

$$\|v\|_{H_0^1(\Omega)} \leq C_X h^{-1} \|v\|_{L^2(\Omega)}. \tag{4.1.10}$$

These properties are known to be satisfied when X^h is a space of WEB (Weighted Extended B-splines) finite element of resolution h.

Write

$$V^h := \left\{ \varphi \circ F(x) \,:\, \varphi \in X^h \right\}. \tag{4.1.11}$$

For $v \in H_0^1(\Omega)$ write $\mathscr{R}_h v$ the projection of v on V_h with respect to the bilinear operator $a[\cdot,\cdot]$, i.e., the unique element of V_h such that for all $w \in V_h$,

$$a[w, v - \mathscr{R}_h v] = 0. \tag{4.1.12}$$

Define Y_T^h the subspace of $L^2\left(0,T;H_0^1(\Omega)\right)$ as

$$Y_T^h := \{ v \in L^2\left(0,T;H_0^1(\Omega)\right) : v(x,t) \in V_h, \forall t \in [0,T] \}. \tag{4.1.13}$$

We use the following notation

$$a[v,w] := \int_\Omega {}^t\nabla v(x,t) a(x) \nabla w(x,t)\, dx. \tag{4.1.14}$$

Write u_h the solution in Y_T^h of the following system of ordinary differential equations:

$$\begin{cases} (K^{-1}\psi, \partial_t^2 u_h)_{L^2(\Omega)} + a[\psi, u_h] = (\psi, g)_{L^2(\Omega)} \quad \text{for all } t \in (0,T) \text{ and } \psi \in V^h, \\ u_h(x,0) = \mathscr{R}_h u(x,0), \\ \partial_t u_h(x,0) = \mathscr{R}_h \partial_t u(x,0). \end{cases} \tag{4.1.15}$$

The following theorem shows the accuracy of the semidiscrete solution. We need an improved assumption on the regularity of the forcing term g and the initial data, which guarantees $O(h)$ convergence of the scheme (4.1.15). On the other hand, we can see that even g and all the initial data are smooth, under the assumption $a(x) \in L^\infty$, we can at best expect $u \in L^\infty(0,T,H^1(\Omega))$ instead of the improved regularity $L^\infty(0,T,H^2(\Omega))$ in the harmonic coordinate.

Assumption B. Assume that the forcing term g satisfies $\partial_t^2 g \in L^2(\Omega_T)$, $\partial_t g \in L^\infty(0,T,L^2(\Omega))$, initial value $u(x,0)$ and $\partial_t u(x,0)$ satisfy $\partial_t^2 u(x,0) \in H^1(\Omega)$ and $\nabla a(x)\nabla \partial_t u(x,0) \in L^2(\Omega)$ or equivalently $\partial_t^3 u(x,0) \in L^2(\Omega)$.

From now on we will always suppose without explicitly mentioning that **Assumption B** is satisfied in the discussion of numerical homogenization method.

Theorem 4.1.2.

$$\begin{aligned}\left\|\partial_t(u-u_h)(.,T)\right\|_{L^2(\Omega)} + \left\|(u-u_h)(.,T)\right\|_{H_0^1(\Omega)} \leq Ch\big(&\|\partial_t g\|_{L^\infty(0,T,L^2(\Omega))} + \|\partial_t^2 g\|_{L^2(\Omega_T)} \\ &+\|\partial_t^2 u(x,0)\|_{H^1(\Omega)} + \|\partial_t^3 u(x,0)\|_{L^2(\Omega)}\big).\end{aligned} \tag{4.1.16}$$

The constant C depends on C_X, n, Ω, β_σ, $K_{\min}$, $K_{\max}$, $\lambda_{\min}(a)$, $\lambda_{\max}(a)$, and $\left\|(\mathrm{Trace}[\sigma])^{-1}\right\|_{L^\infty(\Omega_T)}$.

4.1.3 Numerical homogenization in time and space

Let $M \in \mathbb{N}$, $(t_n = n\frac{T}{M})_{0\leq n\leq M}$ is a discretization of $[0,T]$. (φ_i) is a basis of X_h. Write trial space Z_T^h the subspace of Y_T^h such that

$$Z_T^h = \{w \in Y_T^h : w(x,t) = \sum_i c_i(t)\varphi_i\big(F(x,t)\big),\ c_i(t) \text{ are linear on } (t_n,t_{n+1}]\}. \tag{4.1.17}$$

Write test space U_T^h the subspace of Y_T^h such that

$$U_T^h = \{\psi \in Y_T^h : \psi(x,t) = \sum_i d_i\varphi_i\big(F(x,t)\big),\ d_i \text{ are linear (on } [0,T]).\}. \tag{4.1.18}$$

Write v_h the solution in Z_T^h of the following system of implicit weak formulation: for $n \in \{0,\dots,M-1\}$ and $\psi \in U_T^h$,

$$\begin{aligned}(K^{-1}\psi, \partial_t v_h)(t_{n+1}) - (K^{-1}\psi, \partial_t v_h)(t_n) = \int_{t_n}^{t_{n+1}} (K^{-1}\partial_t \psi, \partial_t v_h) dt \\ - \int_{t_n}^{t_{n+1}} a[\psi, v_h] dt + \int_{t_n}^{t_{n+1}} (\psi, g) dt.\end{aligned} \tag{4.1.19}$$

In equation (4.1.19), $\partial_t v_h(t)$ stands for $\lim_{\varepsilon \downarrow 0}(v_h(t) - v_h(t-\varepsilon))/\varepsilon$. The unknowns are $\partial_t v_h(t)$, once we know the values of them at t_n, we can use the following relation to obtain $v_h(t_{n+1})$,

$$\partial_t v_h(t_{n+1}) = \frac{v_h(t_{n+1}) - v_h(t_n)}{t_{n+1} - t_n}. \tag{4.1.20}$$

The trial function space Z_T^h and test function space U_T^h introduced in this section are of different degrees of freedom, therefore we are solving a least square problem at each time step.

Denote mass matrix M by $M_{ij} = \big(\varphi_i(F(x)), \varphi_j(F(x))\big)$ and stiffness matrix K by $K_{ij} = a[\varphi_i(F(x)), \varphi_j(F(x))]$, we can show that (4.1.19) is equivalent to a linear equation

$$Ce = f, \tag{4.1.21}$$

with the matrix C

$$C = \begin{bmatrix} M + \frac{\Delta t^2}{2} K \\ t_n M + \frac{\Delta t^2 (2t_{n+1} + t_n)}{6} K \end{bmatrix}. \tag{4.1.22}$$

Since M and K are positive definite, the least square problem has a unique solution, which also gives the existence and uniqueness of v_h. The computational cost of solving the least square problem only depends the degrees of freedom of the coarse mesh.

The following Theorem 4.1.3 shows the stability of the implicit scheme (4.1.19):

Theorem 4.1.3.

$$\begin{aligned}\|\partial_t v_h(.,T)\|_{L^2(\Omega)} + \|v_h(.,T)\|_{H_0^1(\Omega)} \le C\big(\|g\|_{L^2(\Omega_T)} + \|\partial_t u(x,0)\|_{L^2(\Omega)} \\ + \|u(x,0)\|_{H^1(\Omega)}\big) + Ch\big(\|\partial_t g\|_{L^\infty(0,T,L^2(\Omega))} + \|\partial_t^2 g\|_{L^2(\Omega_T)} \\ + \|\partial_t^2 u(x,0)\|_{H^1(\Omega)} + \|\partial_t^3 u(x,0)\|_{L^2(\Omega)}\big).\end{aligned} \tag{4.1.23}$$

The constant C depends on $\lambda_{min}(a)$, $\lambda_{max}(a)$, $K_{\min}$, $K_{\max}$, and T.

The following Theorems 4.1.4 and 4.1.5 give us the accuracy of the implicit scheme (4.1.19).

Theorem 4.1.4.

$$\begin{aligned}&\big\|(\partial_t u_h-\partial_t v_h)(.,T)\|_{L^2(\Omega)}+\big\|(u_h-v_h)(.,T)\big\|_{H^1(\Omega)}\leq C\Delta t(1+h^{-1})\\&\big(\|\partial_t g\|_{L^\infty(0,T,L^2(\Omega))}+\|\partial_t^2 g\|_{L^2(\Omega_T)}+\|\partial_t^2 u(x,0)\|_{H^1(\Omega)}+\|\partial_t^3 u(x,0)\|_{L^2(\Omega)}\big).\end{aligned}\tag{4.1.24}$$

The constant C depends on C_X, T, $\lambda_{\min}(a)$, $\lambda_{\max}(a)$, $K_{\min}$, $K_{\max}$, β_σ, $\big\|(\mathrm{Trace}[\sigma])^{-1}\big\|_{L^\infty(\Omega_T)}$, and $\big\|\mathrm{Trace}[\sigma]\big\|_{L^\infty(\Omega_T)}$.

Theorem 4.1.5.

$$\begin{aligned}&\big\|(u_h-v_h)(.,T)\big\|_{L^2(\Omega)}+\big\|\int_0^T(u_h-v_h)(.,t)\,dt\big\|_{H_0^1(\Omega)}\leq C\frac{\Delta t}{h}(\Delta t)^{\frac{1}{2}}\\&\big(\|\partial_t g\|_{L^\infty(0,T,L^2(\Omega))}+\|\partial_t^2 g\|_{L^2(\Omega_T)}+\|\partial_t^2 u(x,0)\|_{H^1(\Omega)}+\|\partial_t^3 u(x,0)\|_{L^2(\Omega)}\big).\end{aligned}\tag{4.1.25}$$

The constant C depends on C_X, T, $\lambda_{\min}(a)$, $\lambda_{\max}(a)$, $K_{\min}$, $K_{\max}$, β_σ, $\big\|(\mathrm{Trace}[\sigma])^{-1}\big\|_{L^\infty(\Omega_T)}$, $\big\|\mathrm{Trace}[\sigma]\big\|_{L^\infty(\Omega_T)}$, and $\|\partial_t v_h\|_{L^\infty(0,T,L^2(\Omega))}$ (which is bounded by Theorem 4.1.3) .

4.2 Proofs

The proofs are organized into three subsections corresponding to the three subsections of section 4.1.

4.2.1 Compensation phenomena: Proof of Theorem 4.1.1

Lemma 4.2.1. *We have*

$$\begin{aligned}\|\partial_t^2 u\|^2_{L^2(\Omega)}(T)+a[\partial_t u](T)\leq&C(T,\frac{K_{\max}}{K_{\min}},K_{\max})\Big(a[\partial_t u](0)\\&+\|\partial_t^2 u(x,0)\|^2_{L^2(\Omega)}+\|\partial_t g\|^2_{L^2(\Omega_T)}\Big).\end{aligned}\tag{4.2.1}$$

Proof. In case a is smooth, differentiating (4.1.1) with respect to t, we have

$$K^{-1}\partial_t^3 u - \operatorname{div} a\nabla\partial_t u = \partial_t g. \tag{4.2.2}$$

multiplying by $\partial_t^2 u$, and integrating over Ω, we obtain that

$$\frac{1}{2}\frac{d}{dt}\|K^{-\frac{1}{2}}\partial_t^2 u\|^2_{L^2(\Omega)} + \frac{1}{2}\frac{d}{dt}a[\partial_t u] = (\partial_t g, \partial_t^2 u)_{L^2(\Omega)}. \tag{4.2.3}$$

Integrating the latter equation with respect to t and using Cauchy-Schwartz inequality we obtain that

$$\begin{aligned}\|K^{-\frac{1}{2}}\partial_t^2 u\|^2_{L^2(\Omega)}(T) + a[\partial_t u](T) \leq &\|K^{-\frac{1}{2}}\partial_t^2 u\|^2_{L^2(\Omega)}(0) + a[\partial_t u](0)\\ &+ \|\partial_t g\|_{L^2(\Omega_T)}\|\partial_t^2 u\|_{L^2(\Omega_T)}.\end{aligned} \tag{4.2.4}$$

Consider the following differential inequality,

$$X(t) \leq A(t) + B(t)\Big(\int_0^t X(s)\,ds\Big)^{\frac{1}{2}}. \tag{4.2.5}$$

Write $Y(t) = \sup_{s\in[0,t]} X(s)$, one has

$$X(t) \leq A(t) + B(t)t^{\frac{1}{2}}\big(Y(t)\big)^{\frac{1}{2}}. \tag{4.2.6}$$

Let t be a time such that $Y(t) = X(t)$, then

$$Y(t) \leq 2A(t) + t\big(B(t)\big)^2. \tag{4.2.7}$$

It follows that

$$\begin{aligned}\|\partial_t^2 u\|^2_{L^2(\Omega)}(T) + a[\partial_t u](T) \leq &C(T, \frac{K_{\max}}{K_{\min}}, K_{\max})\Big(a[\partial_t u](0)\\ &+ \|\partial_t^2 u\|^2_{L^2(\Omega)}(0) + \|\partial_t g\|^2_{L^2(\Omega_T)}\Big).\end{aligned} \tag{4.2.8}$$

In the case where a is nonsmooth we use Galerkin approximations of u in (4.1.1) and then pass to limit. This technique is standard and we refer to [58, Section 7.3.2.c] for a

reminder. □

Lemma 4.2.2.

$$\begin{aligned}\|\partial_t u\|^2_{L^2(\Omega)}(T)+a[u](T) \leq & C(T,\frac{K_{\max}}{K_{\min}},K_{\max})\Big(a[u](0)\\ &+\|\partial_t u\|^2_{L^2(\Omega)}(0)+\|g\|^2_{L^2(\Omega_T)}\Big).\end{aligned} \tag{4.2.9}$$

Proof. Multiplying 4.1.1 by $\partial_t u$, and integrating over Ω, we obtain that

$$\frac{1}{2}\frac{d}{dt}\|K^{-\frac{1}{2}}\partial_t u\|^2_{L^2(\Omega)}+\frac{1}{2}\frac{d}{dt}a[u]=(g,\partial_t u)_{L^2(\Omega)}. \tag{4.2.10}$$

The remaining part of the proof is similar to the proof of Lemma 4.2.1. □

Let us now prove the compensation theorem Theorem 4.1.1, this is similar to the proof of Theorem 3.1.1 in Chapter 3. Choose

$$M:=\frac{\sigma}{|\det(\nabla F)|^{\frac{1}{2}}}\circ F^{-1}. \tag{4.2.11}$$

It is easy to check that (4.2.11) is well defined. Moreover observe that $\beta_M=\beta_\sigma$ and

$$\|\nu_M\|^2_{L^\infty(\Omega_T)}\leq\frac{C_n}{(\lambda_{\min}(a))^{\frac{n}{2}}}\left\|(\mathrm{Trace}[\sigma])^{\frac{n}{4}-1}\right\|^2_{L^\infty(\Omega_T)}. \tag{4.2.12}$$

Fix $t\in[0,T]$. Choose

$$f:=\frac{(K^{-1}\partial_t^2 u-g)}{|\det(\nabla F)|^{\frac{1}{2}}}\circ F^{-1}. \tag{4.2.13}$$

Observe that by the change of variable $y=F(x)$ one obtains that

$$\|f\|^2_{L^2(\Omega)}\leq K_{\min}^{-1}\|\partial_t^2 u\|_{L^2(\Omega)}+\|g\|^2_{L^2(\Omega)}. \tag{4.2.14}$$

It follows from Theorem A.1.3 that there exists a unique $v\in W^{2,2}(\Omega)$ satisfying the

equation

$$\sum_{i,j}\left(\sigma(F^{-1}(y))\right)_{i,j}\partial_i\partial_j v(y,t)=\tilde{K}^{-1}(y)\partial_t^2\tilde{u}(y,t)-\tilde{g}(y,t), \tag{4.2.15}$$

and

$$\|v\|^2_{H^2(\Omega)}\leq\frac{C\|v_M\|^2_{L^\infty(\Omega_T)}}{(1-\beta_\sigma^{\frac{1}{2}})^2}\left(K_{\min}^{-1}\|\partial_t^2 u\|_{L^2(\Omega)}+\|g\|^2_{L^2(\Omega)}\right). \tag{4.2.16}$$

Use the notation $\tilde{K}:=K\circ F^{-1}$, $\tilde{g}:=g\circ F^{-1}$, and $\tilde{u}:=u\circ F^{-1}$. By change of variable $y=F(x)$ and the identity $\operatorname{div} a\nabla F=0$ we deduce that (4.2.15) can be written as

$$\operatorname{div}\left(a\nabla(v\circ F)\right)=K^{-1}\partial_t^2 u-g. \tag{4.2.17}$$

If $\partial_t^2 u\in L^2(\Omega)$ and $g(.,t)\in L^2(\Omega)$ we can use the uniqueness property of the solution of the divergence form elliptic equation with homogeneous Dirichlet boundary condition.

$$\operatorname{div}\left(a\nabla w\right)=K^{-1}\partial_t^2 u-g. \tag{4.2.18}$$

to obtain that $v\circ F=u$. Thus we have proven Theorem 4.1.1. In the following sections we will prove the convergence of semidiscrete and fully discrete numerical homogenization formulation (4.1.15) and (4.1.19).

4.2.2 Numerical homogenization in space: Proof of Theorem 4.1.2.

Lemma 4.2.3. *We have*

$$\begin{aligned}\|\partial_t^2 u_h\|^2_{L^2(\Omega)}(T)+a[\partial_t u_h](T)\leq& C(T,\frac{K_{\max}}{K_{\min}},K_{\max})\Big(a[\partial_t u_h](0)\\&+\|\partial_t^2 u_h(x,0)\|^2_{L^2(\Omega)}+\|\partial_t g\|^2_{L^2(\Omega_T)}\Big).\end{aligned} \tag{4.2.19}$$

Proof. The proof is similar to the proof of Lemma 4.2.1. □

Lemma 4.2.4.

$$\begin{aligned}\|\partial_t u_h\|^2_{L^2(\Omega)}(T)+a[u_h](T) \leq & C(T,\frac{K_{\max}}{K_{\min}},K_{\max})\Big(a[u_h](0)\\ &+\|\partial_t u_h\|^2_{L^2(\Omega)}(0)+\|g\|^2_{L^2(\Omega)_T}\Big).\end{aligned} \tag{4.2.20}$$

Proof. The proof is similar to the proof of Lemma 4.2.2. □

Write $\mathscr{R}_h$ the projection operator mapping $L^2(0,T;H^1_0(\Omega))$ onto Y^h_T, for all $v \in Y^h_T$:

$$\mathscr{A}_T[v, u-\mathscr{R}_h u] = 0 \tag{4.2.21}$$

let $\rho := u-\mathscr{R}_h u$ and $\theta := \mathscr{R}_h u - u_h$.

For $t \in [0,T]$ and $v \in H^1_0(\Omega)$, we write $\mathscr{R}_{h,t}v(.,t)$ the solution of:

$$\int_\Omega {}^t\nabla\psi a(x)\nabla(v-\mathscr{R}_{h,t}v(x,t))\,dx = 0 \quad \text{for all } \psi \in V^h \tag{4.2.22}$$

It is obvious that $\mathscr{R}_h u(.,t) = \mathscr{R}_{h,t}u(.,t)$. For example, we can choose a series of test functions which is separable in space and time, $v(x,t) = T(t)X(x)$, $T(t)$ is smooth in t and has $\delta(t)$ function as its weak limit.

We will need the Lemma 3.2.9 of Chapter 3 which gives the bound

$$a[v-\mathscr{R}_{h,t}v]\big)^{\frac{1}{2}} \leq Ch \tag{4.2.23}$$

We will use the Lemmas 3.2.9, 4.2.5, 4.2.6, 4.2.7 and 4.2.8 to obtain the approximation property of the projection operator $\mathscr{R}_h$.

Lemma 4.2.5. *$\partial_t(u\circ F^{-1}) \in L^\infty(0,T,H^2(\Omega))$ and*

$$\begin{aligned}\|\partial_t(u\circ F^{-1})\|_{L^\infty(0,T,H^2(\Omega))} \leq & C\big(\|\partial_t g\|_{L^\infty(0,T,L^2(\Omega))}+\|\partial_t^2 g\|_{L^2(\Omega_T)}\\ &+\|\partial_t^3 u(x,0)\|_{L^2(\Omega)}+\|\partial_t^2 u(x,0)\|_{H^1(\Omega)}\big).\end{aligned} \tag{4.2.24}$$

Remark. *The constant C is the one given in Theorem 4.1.1.*

Proof. The proof follows from the differentiation of (4.1.1) and is similar to the proof of Theorem 4.1.1. □

Lemma 4.2.6. *We have*

$$\begin{aligned}\big(\mathscr{A}_T[\partial_t\rho]\big)^{\frac{1}{2}} \leq & Ch\big(\|\partial_t g\|_{L^\infty(0,T,L^2(\Omega))} + \|\partial_t^2 g\|_{L^2(\Omega_T)} \\ & + \|\partial_t^3 u(x,0)\|_{L^2(\Omega)} + \|\partial_t^2 u(x,0)\|_{H^1(\Omega)}\big).\end{aligned} \tag{4.2.25}$$

Proof. The proof is a straightforward application of Lemma 4.2.5 and Lemma 3.2.9. □

Lemma 4.2.7. *We have*

$$\begin{aligned}\|\partial_t\rho\|_{L^2(\Omega_T)} \leq & Ch^2\big(\|\partial_t g\|_{L^\infty(0,T,L^2(\Omega))} + \|\partial_t^2 g\|_{L^2(\Omega_T)} \\ & + \|\partial_t^3 u(x,0)\|_{L^2(\Omega)} + \|\partial_t^2 u(x,0)\|_{H^1(\Omega)}\big).\end{aligned} \tag{4.2.26}$$

The constant C in Lemma 4.2.6 and 4.2.7 depends on C_X, n, Ω, β_σ, $\lambda_{\min}(a)$, $\lambda_{\max}(a)$, $K_{\min}$, $K_{\max}$, and $\big\|(\mathrm{Trace}[\sigma])^{-1}\big\|_{L^\infty(\Omega_T)}$. If $n \geq 3$ it also depends on $\big\|\mathrm{Trace}[\sigma]\big\|_{L^\infty(\Omega_T)}$.

Proof. The proof follows from standard duality techniques and similar to that of Lemma 3.2.11. We can deduce that,

$$\big(\mathscr{A}_T[v - \mathscr{R}_{h,t}v]\big)^{\frac{1}{2}} \leq Ch\|\partial_t\rho\|_{L^2(\Omega_T)}. \tag{4.2.27}$$

Since

$$\|\partial_t\rho\|_{L^2(\Omega_T)} \leq Ch\big(\mathscr{A}_T[\partial_t\rho]\big)^{\frac{1}{2}}. \tag{4.2.28}$$

We deduce the lemma by applying Lemma 4.2.6 to bound $A_T[\partial_t\rho]$. □

Lemma 4.2.8. *We have the following estimates for initial data,*

$$\begin{aligned}\|\partial_t(u_h(x,0) - u(x,0))\|_{L^2(\Omega)} &\leq Ch^2\big(\|\partial_t g(x,0)\|_{L^2(\Omega)} + \|\partial_t^2 u(x,0)\|_{H^1(\Omega)} + \|\partial_t^3 u(x,0)\|_{L^2(\Omega)}\big) \\ \|u_h(x,0) - u(x,0)\|_{H_0^1(\Omega)} &\leq Ch\big(\|\partial_t g(x,0)\|_{L^2(\Omega)} + \|\partial_t^2 u(x,0)\|_{H^1(\Omega)} + \|\partial_t^3 u(x,0)\|_{L^2(\Omega)}\big)\end{aligned} \tag{4.2.29}$$

Proof. We can estimate $\|\partial_t\rho\|_{L^2(\Omega)}$ similar to Lemma 4.2.7 and apply Lemma 3.2.9 to get the second inequality.

□

Lemma 4.2.9. *We have*

$$\|\partial_t(u-u_h)\|^2_{L^2(\Omega)}(T)+a[u-u_h](T)\leq C(K_{\min},K_{\max},T)\Big(\|\partial_t(u-u_h)\|^2_{L^2(\Omega)}(0) \\ +a[u-u_h](0)+\|\partial_t\rho\|_{L^2(\Omega_T)}\|\partial_t^2(u-u_h)\|_{L^2(\Omega_T)}+\mathscr{A}_T[\partial_t\rho]\Big). \tag{4.2.30}$$

Proof. We have

$$(K^{-1}\psi,\partial_t^2(u-u_h))+a[\psi,u-u_h]=0. \tag{4.2.31}$$

Let $\psi=\partial_t\theta=\partial_t(u-u_h)-\partial_t\rho$, it results

$$\frac{1}{2}\frac{d}{dt}\|K^{-\frac{1}{2}}\partial_t(u-u_h)\|^2_{L^2(\Omega)}+\frac{1}{2}\frac{d}{dt}a[u-u_h]=(K^{-1}\partial_t\rho,\partial_t^2(u-u_h))+a[\partial_t\rho,u-u_h]. \tag{4.2.32}$$

Integrate with respect to t, using Cauchy-Schwartz inequality, we have

$$\frac{1}{2}\|K^{-\frac{1}{2}}\partial_t(u-u_h)\|^2_{L^2(\Omega)}(T)-\frac{1}{2}\|K^{-\frac{1}{2}}\partial_t(u-u_h)\|^2_{L^2(\Omega)}(0)+\frac{1}{2}a[u-u_h](T) \\ -\frac{1}{2}a[u-u_h](0)\leq\int_0^T K_{\min}^{-1}\|\partial_t\rho\|_{L^2(\Omega)}\|\partial_t^2(u-u_h)\|_{L^2(\Omega)}dt+\big(\mathscr{A}_T[\partial_t\rho]\mathscr{A}_T[u-u_h]\big)^{\frac{1}{2}}. \tag{4.2.33}$$

The remaining part of the proof is similar to the proof of Lemma 4.2.1. □

Theorem 4.1.2 is a straightforward application of Lemma 4.2.1, Lemma 4.2.3, Lemma 4.2.6, Lemma 4.2.7, Lemma 4.2.8, and Lemma 4.2.9.

4.2.3 Numerical homogenization in space and time: Proof of Theorem 4.1.3–4.1.5

Stability Choose $\psi\in U_T^h$ in equation (4.1.19) such that $\psi(x,t)=\partial_t v_h(x,t)$ for $t\in(t_n,t_{n+1}]$. We obtain that

$$\|K^{-\frac{1}{2}}\partial_t v_h\|^2_{L^2(\Omega)}(t_{n+1})-(K^{-1}\partial_t v_h(t_{n+1}),\partial_t v_h(t_n))_{L^2(\Omega)}=-\int_{t_n}^{t_{n+1}}a[\partial_t v_h,v_h]dt \\ +\int_{t_n}^{t_{n+1}}(\partial_t v_h,g)_{L^2(\Omega)}dt. \tag{4.2.34}$$

Observing that

$$\int_{t_n}^{t_{n+1}} a[\partial_t v_h, v_h]dt = \frac{1}{2}a[v_h](t_{n+1}) - \frac{1}{2}a[v_h](t_n). \tag{4.2.35}$$

using Cauchy-Schwartz inequality we obtain

$$\begin{aligned}\|K^{-\frac{1}{2}}\partial_t v_h\|^2(t_{n+1}) + a[v_h](t_{n+1}) \leq \|K^{-\frac{1}{2}}\partial_t v_h\|^2(t_n) + a[v_h](t_n)\\ +2\int_{t_n}^{t_{n+1}} (\partial_t v_h, g)_{L^2(\Omega)}(t)\,dt.\end{aligned} \tag{4.2.36}$$

Summing from 0 to M,

$$\|K^{-\frac{1}{2}}\partial_t v_h\|^2(T) + a[v_h](T) \leq \|K^{-\frac{1}{2}}\partial_t v_h\|^2(0) + a[v_h](0) + 2\int_0^T (\partial_t v_h, g)_{L^2(\Omega)}\,dt. \tag{4.2.37}$$

We conclude the proof of Theorem 4.1.3 in a way similar to the proof of Lemma 4.2.1.

H^1 Error Estimate We derive from equations (4.1.19) and (4.1.15) that

$$\begin{aligned}(K^{-1}\psi, \partial_t u_h - \partial_t v_h)(t_{n+1}) - (K^{-1}\psi, \partial_t u_h - \partial_t v_h)(t_n)\\ -\int_{t_n}^{t_{n+1}} (K^{-1}\partial_t\psi, \partial_t u_h - \partial_t v_h)dt + \int_{t_n}^{t_{n+1}} a[\psi, u_h - v_h]dt = 0.\end{aligned} \tag{4.2.38}$$

Let $\psi = \partial_t \hat{u}_h - \partial_t v_h$ where $\hat{u}_h$ is the linear interpolation of u_h over Z_T^h. Write $y_h = u_h - v_h$ and $w_h = \hat{u}_h - u_h$, it follows that

$$\begin{aligned}(K^{-1}\partial_t y_h, \partial_t y_h)(t_{n+1}) + (K^{-1}\partial_t w_h, \partial_t y_h)(t_{n+1}) - (K^{-1}\partial_t y_h, \partial_t y_h)(t_n)\\ -(K^{-1}\partial_t w_h, \partial_t y_h)(t_n) + \int_{t_n}^{t_{n+1}} a[\partial_t y_h, y_h]dt + \int_{t_n}^{t_{n+1}} a[\partial_t w_h, y_h]dt = 0.\end{aligned} \tag{4.2.39}$$

Observing $\int_{t_n}^{t_{n+1}} \partial_t w_h(x,t)\,dt = 0$ we need the following lemma:

Lemma 4.2.10. *If $\int_{t_n}^{t_{n+1}} u(s)\,ds = 0$, then*

$$u^2 \leq \frac{1}{4}\Delta t \int_{t_n}^{t_{n+1}} u'(s)^2\,ds. \tag{4.2.40}$$

Proof. We have

$$\begin{aligned} u(t)^2 &= (u(t) - \frac{1}{\Delta t}\int_{t_n}^{t_{n+1}} u(s)ds)^2 \\ &= \frac{1}{\Delta t^2}(\int_{t_n}^{t_{n+1}}(\int_s^t u'(\xi)d\xi)ds)^2 \\ &\leq \frac{1}{\Delta t^2}(\int_{t_n}^{t_{n+1}} 1^2 ds)(\int_{t_n}^{t_{n+1}}(\int_s^t (u'(\xi))d\xi)^2 ds) \\ &= \frac{1}{\Delta t}\int_{t_n}^{t_{n+1}}(\int_s^t (u'(\xi))d\xi)^2 ds \\ &\leq \frac{1}{\Delta t}\int_{t_n}^{t_{n+1}} |\int_s^t 1^2 d\xi| \cdot |\int_s^t (u'(\xi))^2 d\xi| ds \\ &= \frac{\Delta t}{4}\int_{t_n}^{t_{n+1}} (u'(s))^2 ds. \end{aligned}$$

□

Since $\partial_t^2 w_h(x,t) = -\partial_t^2 u_h(x,t)$ in $(t_n, t_{n+1}]$, by Lemma 4.2.10 we have

$$\int_\Omega |\partial_t w_h(x,t)|^2 dxdt \leq \frac{1}{4}\Delta t \int_{t_n}^{t_{n+1}}\int_\Omega |\partial_t^2 u_h(x,t)|^2 dxdt, \tag{4.2.41}$$

and

$$\int_{t_n}^{t_{n+1}}\int_\Omega |\partial_t w_h(x,t)|^2 dxdt \leq \frac{1}{4}\Delta t^2 \int_{t_n}^{t_{n+1}}\int_\Omega |\partial_t^2 u_h(x,t)|^2 dxdt. \tag{4.2.42}$$

Using the inverse Sobolev inequality (4.1.10) we obtain from equation (4.2.42) that

$$\int_{t_n}^{t_{n+1}}\int_\Omega a[\partial_t w_h]\, dxdt \leq C\frac{\Delta t^2}{h^2}\int_{t_n}^{t_{n+1}}\int_\Omega |\partial_t^2 u_h(x,t)|^2 dxdt. \tag{4.2.43}$$

Summing (4.2.39) over n, notice $y_h(0) = 0$, $\partial_t y_h(0) = 0$ we obtain that

$$(K^{-1}\partial_t y_h, \partial_t y_h)_{L^2(\Omega)}(T) + \frac{1}{2}a[y_h(.,T)] = -\int_0^T a[\partial_t w_h, y_h]dt - (K^{-1}(\partial_t w_h, \partial_t y_h)_{L^2(\Omega)}(T). \tag{4.2.44}$$

Theorem 4.1.4 is a straightforward consequence of (4.2.44), the estimates (4.2.41), (4.2.43), Lemma 4.2.3 and Lemma 4.2.8.

L^2 Error Estimate Following [22], write $y(t) := \int_t^T (u_h - v_h)ds$ and $\psi(t)$ the linear interpolation of $y(t)$ on Z_T^h. Write $z(t) = y(t) - \psi(t)$. Using the test function ψ in (4.2.38) we obtain that

$$\begin{aligned}&(K^{-1}(y-z), \partial_t(u_h - v_h))(t_{n+1}) - (K^{-1}(y-z), \partial_t(u_h - v_h))(t_n)\\ &+ \int_{t_n}^{t_{n+1}} (K^{-1}(u_h - v_h + \partial_t z), \partial_t(u_h - v_h))dt + \int_{t_n}^{t_{n+1}} a[y-z, u_h - v_h)]dt = 0\end{aligned} \tag{4.2.45}$$

Observe that

$$\int_{t_n}^{t_{n+1}} (K^{-1}(u_h - v_h), \partial_t(u_h - v_h))dt = \frac{1}{2}\|K^{-\frac{1}{2}}(u_h - v_h)\|^2(t_{n+1}) - \frac{1}{2}\|K^{-\frac{1}{2}}(u_h - v_h)\|^2(t_n). \tag{4.2.46}$$

Moreover

$$\begin{aligned}\int_{t_n}^{t_{n+1}} a[y, u_h - v_h]dt &= -\int_{t_n}^{t_{n+1}} \frac{1}{2}\frac{d}{dt}a[y]dt\\ &= -\frac{1}{2}a[y](t_{n+1}) + \frac{1}{2}a[y](t_n).\end{aligned} \tag{4.2.47}$$

Write

$$I_{n+1} := (K^{-1}(y-z), \partial_t(u_h - v_h))(t_{n+1}) + \frac{1}{2}\|K^{-\frac{1}{2}}(u_h - v_h)\|^2(t_{n+1}) - \frac{1}{2}a[\int_{t_{n+1}}^T (u_h - v_h)ds]. \tag{4.2.48}$$

We have

$$I_{n+1} \;=\; I_n - \int_{t_n}^{t_{n+1}} (K^{-1}\partial_t z, \partial_t(u_h - v_h))dt + \int_{t_n}^{t_{n+1}} a[z, u_h - v_h)]dt.$$

It follows by induction that

$$I_{n+1} = I_0 - \underbrace{\int_0^{t_{n+1}} (K^{-1}\partial_t z, \partial_t(u_h - v_h))dt}_{J_1} + \underbrace{\int_0^{t_{n+1}} a[z, u_h - v_h)]dt}_{J_2}. \tag{4.2.49}$$

Choose $T = t_{n+1}$. Observe that $v_h(0) = u_h(0)$, $\partial_t v_h(0) = \partial_t u_h(0)$, $I_0 = -\frac{1}{2}a[\int_0^T (u_h - v_h)ds]$, and $I_{n+1} = \frac{1}{2}\|K^{-\frac{1}{2}}(u_h - v_h)\|^2(T)$. Moreover using Theorems 4.1.3, 4.1.4 and Lemma 4.2.4, it follows that $\|\partial_t z\|_{L^2(\Omega)}$ is bounded by $\Delta t^2(1 + h^{-1})$ and $\|\partial_t(u_h - v_h)\|_{L^2(\Omega)}$

is bounded by $\Delta t(1+h^{-1})$, hence we can deduce that $J_1 \leq C\Delta t^3/h^2$.

Similarly we obtain that

$$\begin{aligned} J_2 &\leq \int_0^T a[z]^{\frac{1}{2}} a[u_h - v_h]^{\frac{1}{2}} dt \\ &\leq \frac{C\Delta t}{h} \sup_{t\in[0,T]} \|u_h - v_h\|_{H^1(\Omega)} \int_0^T \|u_h - v_h\|_{H^1(\Omega)} dt \\ &\leq \frac{C}{h^2}\Delta t^3. \end{aligned} \tag{4.2.50}$$

This concludes the proof of Theorem 4.1.5.

4.3 Numerical Experiments

In this section, we will first give the numerical algorithm. Several examples will be presented. The computational domain is the unit square in dimension two. Equation (4.1.1) is solved on a fine tessellation characterized by 16129 interior nodes (degrees of freedom). Three different coarse tessellations with 9, 49, and 225 degrees of freedom (*dof*) are considered.

The hyperbolic equation (4.1.1) has been homogenized on these coarse meshes. We use web extended B-spline based finite element [76] to be the space X^h introduced in subsection 4.1.2. For all the numerical examples, we compute the solutions up to time $T = 1$. The initial condition is $u(x,0) = 0$ and $u_t(x,0) = 0$.

The fine mesh solver for the wave equation is *Matlab* routine *hyperbolic*, which uses linear finite element basis in space and adaptive integrator in time. The fine mesh solver for F is *Matlab* routine *assempde*.

Algorithm 4.3.1 (Algorithm for Homogenization).

1. *Compute F on fine mesh.*
2. *Construct multi-scale finite element basis $\psi = \varphi \circ F$, compute stiffness matrix K and mass matrix M (ψ is piecewise linear on the fine mesh).*
3. *Time march (4.1.19) and (4.1.20) on coarse mesh.*

Table 4.1: Coarse mesh error (dof 49): Time independent site percolation with $g = 1$ with different methods

Method	L^1	L^∞	L^2	H^1
LFEM	0.1446	0.2159	0.1615	0.3431
FEM_ψ_{lin}	0.0157	0.0690	0.0443	0.1504
FEM_ψ_{sp}	0.0064	0.0233	0.0070	0.0522

Example 4.3.1. *Time independent site percolation*

In this example we consider the site percolating medium associated to figure 4.1. Equation (4.1.1) has been solved with $g = 1$, $g = \sin(2.4x - 1.8y + 2\pi t)$.

and a Gaussian source function given by

$$g(x,y) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\Big(-\frac{x^2 + (y-0.15)^2}{2\sigma^2}\Big) \tag{4.3.1}$$

with $\sigma = 0.05$. Notice that as $\sigma \to 0$, the source function will become more singular. Figure 4.2 shows u computed on 16129 interior nodes and u_h computed on 9 interior nodes in the case $g = 1$ at time 1. Numerical errors are given in tables 4.1 –4.6.

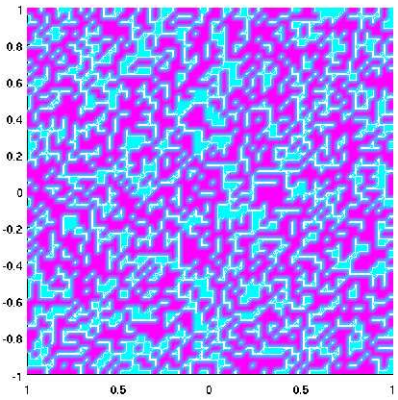
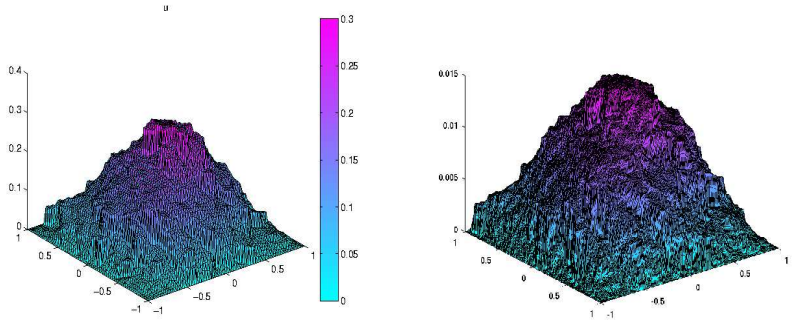


Figure 4.1: Site percolation

Tables 4.1 and 4.2 show the comparison between different numerical homogenization methods; here we use the notation in Chapter 2. LFEM is a multi-scale finite element method where F is computed locally (instead of globally) on each triangle K of the coarse



(a) u. **(b)** u_h.

Figure 4.2: u computed on 16129 interior nodes and u_h computed on 9 interior nodes at time 1

Table 4.2: Fine mesh error (dof 49): Time independent site percolation with $g = 1$ with different methods

Method	L^1	L^∞	L^2	H^1
LFEM	0.1582	0.2557	0.2231	0.3304
FEM_ψ_{lin}	0.0439	0.0518	0.0791	0.1236
FEM_ψ_{sp}	0.0097	0.0493	0.0126	0.0767

mesh as the solution of a cell problem with boundary condition $F(x) = x$ on ∂K. FEM_ψ_{lin} is the Galerkin scheme using the finite elements $\psi_i = \varphi_i \circ F$, where φ_i are the piecewise linear nodal basis elements. FEM_ψ_{sp} is the Galerkin scheme using the finite elements $\psi_i = \varphi_i \circ F$, where φ_i are the weighted cubic B-spline basis elements. We observe that the methods using global F have much better performance, and FEM_ψ_{sp} is better than FEM_ψ_{lin}.

Tables 4.3, 4.4, 4.5, and 4.6 show that we have reasonable error with different forcing term using FEM_ψ_{sp}.

Table 4.3: Coarse mesh error: Time independent site percolation with $g = \sin(2.4x - 1.8y + 2\pi t)$

dof	L^1	L^∞	L^2	H^1
9	0.0696	0.0920	0.0701	0.1014
49	0.0337	0.0431	0.0305	0.0648
225	0.0318	0.0653	0.0292	0.0921

Table 4.4: Fine mesh error: Time independent site percolation with $g = \sin(2.4x - 1.8y + 2\pi t)$

dof	L^1	L^∞	L^2	H^1
9	0.0998	0.1232	0.0887	0.2428
49	0.0592	0.1150	0.0536	0.1778
225	0.0404	0.1031	0.0380	0.1398

Table 4.5: Coarse mesh error: Time independent site percolation with Gaussian source

dof	L^1	L^∞	L^2	H^1
9	0.0748	0.1235	0.0799	0.3767
49	0.0546	0.1092	0.0580	0.2602
225	0.0368	0.0601	0.0406	0.0974

Table 4.6: Fine mesh error: Time independent site percolation with Gaussian source

dof	L^1	L^∞	L^2	H^1
9	0.0977	0.4595	0.1192	0.4857
49	0.0927	0.4144	0.1102	0.3857
225	0.0866	0.2030	0.1098	0.3802

Table 4.7: Coarse mesh error (dof 49): High conductivity channel with different strength

A	L^1	L^∞	L^2	H^1
10	0.0041	0.0197	0.0083	0.0208
100	0.0080	0.0459	0.0126	0.0492
1000	0.0349	0.0934	0.0484	0.1051

Table 4.8: Fine mesh error (dof 49): High conductivity channel with different strength

A	L^1	L^∞	L^2	H^1
10	0.0053	0.0246	0.0120	0.0375
100	0.0117	0.0501	0.0179	0.0958
1000	0.0454	0.01253	0.0611	0.1491

Example 4.3.2. *Time independent high conductivity channel*

In this example a is characterized by a narrow and long ranged high conductivity channel. We choose $a(x) = A \gg 1$, if x is in the channel, and $a(x) = O(1)$ and random, if x is not in the channel. The media is illustrated in figure 4.3.

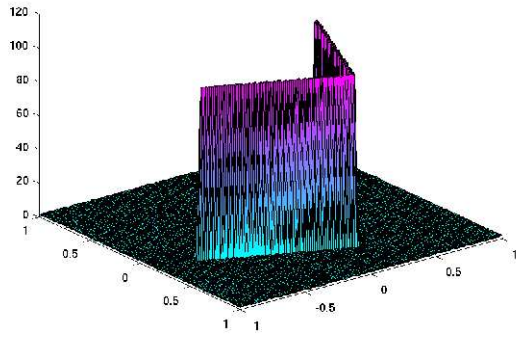


Figure 4.3: High conductivity channel superposed on a random medium

Tables 4.7 and 4.8 give the coarse and fine meshes errors for $g = 1$ with fixed coarse mesh (*dof* 49) and $A = 10, 100, 1000$ respectively. From the table we can see that the errors grow with A increasing, but the growth is moderate. The fine mesh and coarse mesh errors for $g = \sin(2.4x - 1.8y + 2\pi t)$ are given in tables 4.9 and 4.10.

Table 4.9: Coarse mesh error: High conductivity channel case with $g = \sin(2.4x - 1.8y + 2\pi t)$

dof	L^1	L^∞	L^2	H^1
9	0.0364	0.0338	0.0335	0.0541
49	0.0193	0.0282	0.0196	0.0447
225	0.0081	0.0092	0.0078	0.0204

Table 4.10: Fine mesh error: High conductivity channel case with $g = \sin(2.4x - 1.8y + 2\pi t)$

dof	L^1	L^∞	L^2	H^1
9	0.0748	0.0790	0.0729	0.1514
49	0.0295	0.0339	0.0291	0.0760
225	0.0095	0.0119	0.0091	0.0315

Table 4.11: Coarse mesh error: Neumann boundary condition, fine mesh *dof* 16129

dof	L^1	L^∞	L^2	H^1
9	0.0468	0.893	0.0506	0.1320
49	0.0138	0.0713	0.0166	0.1353
225	0.0094	0.0436	0.0115	0.1778

Example 4.3.3. *Time dependent source with Neumann boundary condition*

In this example we consider the site percolating medium. (4.1.1) has been solved with Neumann boundary condition. The source term is given by $g(x,t) = T(t)X(x,y)$, $X(x,y)$ is the Gaussian source function described by

$$X(x,y) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\big(-\frac{x^2+y^2}{2\sigma^2}\big), \tag{4.3.2}$$

with $\sigma = 0.05$, $T(t) = T_1(t)T_2(t)$

$$T_1(t) = \sum_1^{10} 2\frac{1-(-1)^k}{k\pi} \sin(2k\pi t), \tag{4.3.3}$$

and $T_2(t) = \text{erfc}(8(t-0.5))$, erfc is the complementary error function. See figure 4.4 for $T(t)$ in $(0,1)$.

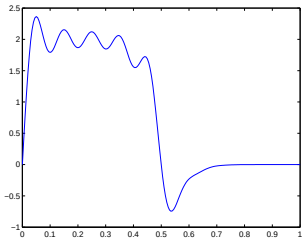
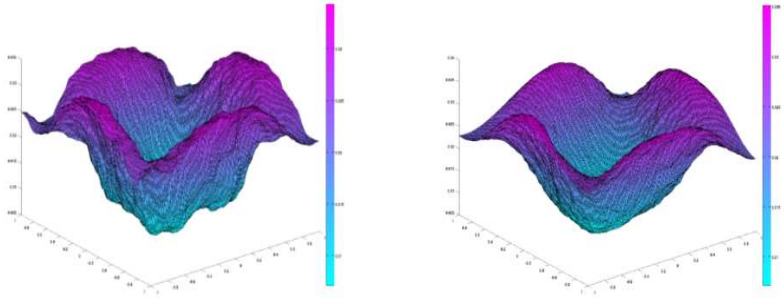


Figure 4.4: $t \to g(0,t)$

Errors are given in tables 4.11 and 4.12 for fine mesh with *dof* 16129, in tables 4.13 and 4.14 for fine mesh with *dof* 65025. Figure 4.5 shows u computed on 16129 interior nodes and u_h computed on 9 interior nodes at time 1.



(a) u **(b)** u_h

Figure 4.5: u computed on 16129 interior nodes and u_h computed on 9 interior nodes at time 1

Table 4.12: Fine mesh error: Neumann boundary condition, fine mesh *dof* 16129

dof	L^1	L^∞	L^2	H^1
9	0.0484	0.1240	0.0571	0.4334
49	0.0261	0.0803	0.0316	0.3025
225	0.0183	0.0520	0.0216	0.2575

Table 4.13: Coarse mesh error: Neumann boundary condition, fine mesh *dof* 65025

dof	L^1	L^∞	L^2	H^1
9	0.0477	0.1217	0.0593	0.1381
49	0.0140	0.1184	0.0178	0.1761

Table 4.14: Fine mesh error: Neumann boundary condition, fine mesh *dof* 65025

dof	L^1	L^∞	L^2	H^1
9	0.0523	0.1209	0.0550	0.3914
49	0.0263	0.1176	0.0314	0.2444

Chapter 5

Conclusion and Discussions

Global information As we have stated in the previous chapters, if one has solved the initial equation at least n times and those solutions are (locally) linearly independent it is also possible to use them as new coordinates for numerical homogenization. The idea to use global fine scale information to homogenize transport equations for reservoir modeling in geophysics is currently implemented in the industry and has been shown to be more accurate than local methods ([131] and [130]). It is applied in practice because the porosity of the medium is time independent and one has to solve an elliptic equation only at $t = 0$ to upscale the transport equations. We notice that some recent results using global information [2, 80] are formulated in a partition of unity framework. In this case, $\{1, F_1, \cdots, F_n\}$ can be used to construct the local approximation space.

Higher dimension In dimension three and higher it has been known since the work of Fenchenko and Khruslov [63], [82] that the homogenization of divergence form elliptic operators $-\operatorname{div} a_\varepsilon \nabla u_\varepsilon = g$ can lead to a non-local homogenized operator if the sequence of matrices a_ε is uniformly elliptic but with entries uniformly bounded only in $L^1(\Omega)$. From a numerical point of view this non-local effect implies that a non-local numerical homogenization method cannot be avoided to obtain accuracy. Recently Briane has shown [43] that this non-local effect is absent in dimension two in the H-convergence setting. In Chapter 2, it is shown that the accuracy of local methods depend on the aspect ratio of the triangles of the tessellation with respect to caloric coordinates and the invertibility of F on both continuous and discrete level (which is not the case if one uses non-local finite

elements). In fact, observe that in dimensions higher than three the harmonic coordinates are not always invertible, an idea to bypass this difficulty could be either to choose the change of coordinates locally and adaptively or to enrich the coordinates by writing down the initial equations as degenerate equations in a space of higher dimension. These points have not been explored.

Optimality condition Moreover, it could be conjectured that one could use any set of n linearly independent solutions of (4.1.1) instead of the harmonic coordinates. The key property allowing the homogenization of (4.1.1) lies in the fact that if g has enough integrability then the space of solutions is at small scales close in H^1 norm to a space of dimension n. Once one has observed at least n linearly independent solutions of (4.1.1), one has seen all of them at small scales. Let us further explain this in the following sense:

Write $L := -\nabla a \nabla$. L^{-1} maps $H^{-1}(\Omega)$ into $H^1_0(\Omega)$, it also maps $L^2(\Omega)$ into V a sub-vector space of $H^1_0(\Omega)$. The elements of V is close in H^1 norm to a space of dimension n (n is the dimension of the physical space Ω). Introduce the Kolmogorov n-width optimality condition [112],

$$d(n, \|\cdot\|, S) = \inf_{E_n} \sup_{f \in S} \inf_{g \in E_n} \|f - g\| \tag{5.0.1}$$

which is the error per degree of freedom for a whole class of functions. E_n denotes an n-dimensional space, and S is the class of functions that we wish to approximate, S is chosen as the unit ball of some appropriate Banach space. A minimizing space E_n is called an optimal space. Therefore, Let $\mathscr{T}_h$ be a triangulation of $\Omega \subset \mathbb{R}^n$ of resolution h (where $0 < h < \texttt{diam}(\Omega)$), in terms of H^1 norm or C^α norm, $\{F_i\}$ is the optimal space for V uniformly in h.

We have the following conjecture, let Λ be set of mappings from $\mathscr{T}_h$ into the unit sphere of $\mathbb{R}^{n+1}$ (if $\lambda \in \Lambda$ then λ is constant on each triangle $K \in \mathscr{T}_h$ and $\|\lambda(K)\| = 1$), then

$$\sup_{v_1, v_2, \ldots, v_{n+1} \in V} \inf_{\lambda \in \Lambda} \frac{\|\sum_{i=1}^{n+1} \lambda_i v_i\|_{H^1_0(\Omega)}}{\sum_{i=1}^{n+1} \|\nabla a \nabla v_i\|_{L^2(\Omega)}} \leq Ch. \tag{5.0.2}$$

Equation (5.0.2) is saying that any $n+1$ elements of V are (at an h approximation in

H^1 norm) linearly dependent. Recall that $n+1$ vectors are linearly dependent in a linear combination (with non-zero coefficients) of these vectors in the null vector. In (5.0.2) the linear combination of the $n+1$ vectors is at relative distance of order h (resolution of the triangulation) from 0.

Further plans Some possible extensions of this work may include:

- Homogenization of partial differential equations
  - Homogenization of linear elasticity equation: Numerical experiments suggest that similar compensation phenomenon still exists. Although the generalization from the scalar case to vectorial case is highly nontrivial, one could still find a change of coordinates in the space of elastic deformations. The difficulty is the injectivity of the mapping.
  - Homogenization of convection-diffusion-reaction equation and Navier-Stokes equation: The idea is to rewrite the equation into a parabolic equation or the perturbation of a parabolic equation, and apply the upscaling method for parabolic equations. Petrov-Galerkin like methods (for example ELLAM) will be investigated. It is of its own theoretical interest to investigate whether or in what assumption we can find similar compensation phenomena for nonlinear equations (such as Navier-Stokes).
- Numerical Analysis and Scientific Computing
  - Implementing and developing a fast, scalable method to solve the fine scale problem. It is interesting to compare numerical methods such as AMG and H-Matrix method, which are able to solve problems with non-separable scales, to find the connection between algebraic approaches and upscaling approaches, and to benefit from such a connection [38]. This would also include attempts to compare several categories of methods: locally adaptive finite element method, multi-scale finite element method with overlapping or non-overlapping local

problems, and our method [1] The goal is to obtain an optimal trade-off which keeps the essential fine scale information and reduces the cost of precomputation.

- Applications
  - It is possible to apply our method to a wide variety of practical problems: composite material, reservoir modeling, inverse problem, and seismic imaging, just to name a few. For instance, in a joint project with Professor Haibin Su of Nanyang Technological University of Singapore, we intend to optimize the charge distribution of composite ferroelectric material in order to lower the voltage needed to actuate the composite. In this case, the equation has to be solved with multiple right hand sides, which is an ideal case on which to apply our method.

[1] In terms of a global mesh generation method, my method is closely related to the harmonic map based moving mesh method [91].

Appendix A

Regularity Theory of PDEs

A.1 Elliptic Equations

This section introduces regularity results for divergence form and nondivergence form elliptic equations.

A.1.1 Divergence Form

Consider the following divergence form elliptic equation,

$$\begin{cases} -\operatorname{div}\left(a(x)\nabla u(x)\right) = g \quad \text{in} \quad \Omega \\ u = 0 \quad \text{in} \quad \partial\Omega. \end{cases} \tag{A.1.1}$$

where $a(x)$ is symmetric, uniformly bounded, and elliptic.

Theorem A.1.1 (Meyers Theorem [7]). *There exists a number $p > 2$ and a positive constant $C > 0$, which both depend only on α, β, and Ω, such that, if $g \in W^{-1,p}(\Omega)$, then the solution u of (A.1.1) belongs to $W_0^{1,p}(\Omega)$ and satisfies*

$$\|u\|_{W_0^{1,p}(\Omega)} \leq C\|g\|_{W^{-1,p}(\Omega)}. \tag{A.1.2}$$

De Giorgi-Moser-Nash theory ([69], [97], [100]) gives Hölder estimate of divergence form elliptic operators with discontinuous coefficients; more precisely we refer to [120] for the global Hölder regularity. For example, we have the following theorem [67, Theorem

8.29]:

Theorem A.1.2 (Hölder Continuity). *If u is a $W^{1,2}$ solution of equation (A.1.1), it follows that u is locally Hölder continuous in Ω. A uniform Hölder estimate may also be obtained if the domain is further restricted, for example, if Ω satisfies a uniform exterior cone condition.*

A.1.2 Nondivergence Form

In this case, we consider strong solutions of the linear second order elliptic Dirichlet problem

$$\begin{cases} \sum_{i,j=1}^{n} M^{ij}(x) D_{ij} v = f \quad \text{in} \quad \Omega, \\ u = 0 \quad \text{in} \quad \partial\Omega. \end{cases} \tag{A.1.3}$$

We assume M to be symmetric, uniformly elliptic and bounded. Strong solution of (A.1.3) is a twice weakly differentiable function satisfying the equation (A.1.3) a.e. in Ω and assuming boundary values on $\partial\Omega$ in classical or in generalized sense.

The strong solution of (A.1.3) may not be unique. In fact, we need the following Cordes condition to ensure the unique solvability of (A.1.3).

Let $\lambda_i(\cdot)$ be the ith eigenvalue. Write for the symmetric matrix M,

$$\begin{aligned} \nu_M :&= \frac{\sum_{i=1}^{n} \lambda_i(M(x))}{\sum_{i=1}^{n} \lambda_i({}^tM(x)M(x))} \\ &= \frac{\text{Trace}[M(x)]}{\text{Trace}[{}^tM(x)M(x)]}, \end{aligned} \tag{A.1.4}$$

and the Cordes parameter β_M associated to M

$$\begin{aligned} \beta_M :&= \text{esssup}_{x\in\Omega} \Big(n - \frac{\big(\sum_{i=1}^{n} \lambda_i(M(x))\big)^2}{\sum_{i=1}^{n} \lambda_i({}^tM(x)M(x))}\Big) \\ &= \text{esssup}_{x\in\Omega} \Big(n - \frac{\big(\text{Trace}[M(x)]\big)^2}{\text{Trace}[{}^tM(x)M(x)]}\Big). \end{aligned} \tag{A.1.5}$$

Theorem A.1.3. *Assume that $\beta_M < 1$. If Ω is convex, then there exists a real number $p_0 > 2$ depending only on n, Ω, and β_M such that for $2 \le p < p_0$, if $f \in L^p(\Omega)$ the Dirichlet*

problem (A.1.3) has a unique solution satisfying

$$\|v\|_{W_0^{2,p}(\Omega)} \leq \frac{C}{1-\beta_M^{\frac{1}{2}}} \|\nu_M f\|_{L^p(\Omega)}. \tag{A.1.6}$$

The constant C only depends on Ω, M, and p.

Theorem A.1.3 is a slight modification of theorem 1.2.3 of [93], for the sake of completeness we will give the main ideas leading to estimate (A.1.6). Let us recall the Miranda-Talenti estimate [93, Section 1.4].

Lemma A.1.1. *Let $\Omega \subset \mathbb{R}^n$ be a bounded and convex domain of class C^2. Then for each $v \in W_0^{2,2}(\Omega)$ it results*

$$\int_\Omega \sum_{i,j=1}^{n} (\partial_i \partial_j v)^2 \, dx \leq \int_\Omega (\Delta v)^2 \, dx. \tag{A.1.7}$$

Remark. *The only place where we use the convexity of Ω is for the validity of Lemma A.1.1 (we refer to [93]).*

The Laplacian $\Delta : W_0^{2,p}(\Omega) \to L^p(\Omega)$ is an isomorphism for each $p > 1$. Let $\Delta^{-1}(p)$ be the inverse operator $\Delta^{-1} : L^p(\Omega) \to W_0^{2,p}$. It is clear from (A.1.7) that $\|\Delta^{-1}(2)\| \leq 1$. We also know that $\|\Delta^{-1}(p)\|$ is monotone increasing. Fix a number $r \in (2, \infty)$, by interpolation inequality we have,

$$\|\Delta^{-1}(p)\| \leq C(p) = \|\Delta^{-1}(r)\|^{\frac{r(p-2)}{p(r-2)}}. \tag{A.1.8}$$

Let v be a solution of (A.1.3) (the existence of v can be obtained from a fix point theorem by [93, p. 21]), we have

$$\|v\|_{W_0^{2,p}(\Omega)} \leq \|\Delta^{-1}(p)\| \|\Delta v\|_{L^p(\Omega)}. \tag{A.1.9}$$

Observing that $\Delta v = \nu_M f + \Delta v - \nu_M L_M v$, one can obtain

$$\|\Delta v\|_{L^p(\Omega)} \leq \|\nu_M f\|_{L^p(\Omega)} + \|\Delta v - \nu_M L_M v\|_{L^p(\Omega)}. \tag{A.1.10}$$

Then following the proof of Theorem 1.2.1 and Theorem 1.2.3 of [93] we have

$$\|\Delta v - v_M L_M v\|_{L^p(\Omega)}^p \leq \int_\Omega \beta_M^{p/2} \Big(\sum_{i,j=1}^n (\partial_i \partial_j v)^p \Big) \, dx. \tag{A.1.11}$$

Since $C(p) \geq 1$, there exists some $p > 2$ such that $1 - C(p)\beta_M^{1/2} \geq (1 - \beta_M^{1/2})/2$. Combining (A.1.9), (A.1.10), and (A.1.11) we obtain that

$$\Big(\int_\Omega \Big(\sum_{i,j=1}^n (\partial_i \partial_j v)^p \Big) \, dx \Big)^{\frac{1}{p}} \leq \frac{2C(p)}{1 - \beta_M^{1/2}} \|v_M f\|_{L^p(\Omega)}. \tag{A.1.12}$$

which leads to estimate (A.1.6).

By Sobolev embedding inequality

$$\|\nabla v\|_{C^{1-\frac{n}{p}}(\bar{\Omega})} \leq C \|v\|_{W_0^{2,p}(\Omega)}. \tag{A.1.13}$$

Theorem A.1.3 implies the Hölder continuity of v in dimension $n = 2$.

A.2 Parabolic Equations

A.2.1 Divergence Form

Consider the following parabolic equation,

$$\begin{cases} \partial_t u = \operatorname{div}\big(a(x,t)\nabla u(x,t)\big) + g \quad \text{in} \quad \Omega \times (0,T) \\ u(x,t) = 0 \quad \text{for} \quad (x,t) \in \big(\partial\Omega \times (0,T)\big) \cup \big(\Omega \cup \{t = 0\}\big). \end{cases} \tag{A.2.1}$$

By energy estimate and Galerkin approximation, there exists unique weak solution of divergence form parabolic equation $u \in L^2(0,T;H_0^1(\Omega))$, with $u' \in L^2(0,T;H^{-1}(\Omega))$. For example, see [58, Section 7.1].

Similar to divergence form elliptic equation, we have Hölder continuity of u [92, Chapter IV].

A.2.2 Nondivergence Form

Consider the following non-divergence form parabolic problem:

$$\partial_t v = \sum_{i,j=1}^{n} M^{ij}(x,t)\partial_i\partial_j v + f. \tag{A.2.2}$$

We assume M to be symmetric, uniformly bounded, and elliptic. $v = 0$ at $t = 0$ and on the boundary $\partial\Omega$. Write

$$\eta_M := \operatorname{esssup}_{x\in\Omega_T} \frac{\operatorname{Trace}[{}^tMM] + 1}{\big(\operatorname{Trace}[M] + 1\big)^2}, \tag{A.2.3}$$

and

$$\alpha_M := \operatorname{esssup}_{x\in\Omega_T} \frac{\operatorname{Trace}[M] + 1}{\operatorname{Trace}[{}^tMM] + 1}. \tag{A.2.4}$$

Write for $p \geq 2$

$$S_p(\Omega_T) := \Big\{ v \in L^p\big(0,T,W_D^{2,p}(\Omega)\big); \partial_t v \in L^p(\Omega_T); v(.,0) \equiv 0 \Big\}. \tag{A.2.5}$$

and

$$\|v\|^p_{S_p(\Omega_T)} := \int_{\Omega_T} \big(\sum_{i,j} (\partial_i\partial_j v)^2 + (\partial_t v)^2\big)^{\frac{p}{2}}\, dy\, dt \tag{A.2.6}$$

We will restate Theorems 1.6.2 and 1.6.3 of [93] below in a version adapted to our framework:

Theorem A.2.1. *Assume Ω to be convex and that there exists $\varepsilon \in (0,1)$ such that $\eta_M \leq 1/(n+\varepsilon)$, then for each $f \in L^2(\Omega_T)$ the Cauchy-Dirichlet problem (A.2.2) admits a unique solution in $S_2(\Omega_T)$ which satisfies the bound*

$$\|v\|_{S_2(\Omega_T)} \leq \frac{\alpha_M}{1-\sqrt{1-\varepsilon}} \|f\|_{L^2(\Omega_T)}. \tag{A.2.7}$$

Theorem A.2.2. *Assume Ω to be convex and that there exists $\varepsilon \in (0,1)$ such that $\eta_M \leq 1/(n+\varepsilon)$, then there exists a number $p_0 > 2$ depending on Ω, n, ε such that for each $f \in L^p(\Omega_T)$ the Cauchy-Dirichlet problem (A.2.2) admits a unique solution in $S_p(\Omega_T)$ which*

satisfies the bound

$$\|v\|_{S_p(\Omega_T)} \leq C_p \frac{\alpha_M}{1-\sqrt{1-\varepsilon}} \|f\|_{L^p(\Omega_T)}. \tag{A.2.8}$$

Remark. *In fact theorem 1.6.3 of [93] is written with* $1 - C(p)\sqrt{1-\varepsilon}$ *in the denominator of (A.2.8) but it is easy to modify it to obtain (A.2.8) by lowering the value of* p_0 *and changing the value of* C_p.

Appendix B

a-Harmonic Mapping

Consider the weak solutions to the divergence form elliptic equation

$$\operatorname{div}(a\nabla F) = 0 \quad \text{in } \Omega \tag{B.0.1}$$

where a is a two by two symmetric matrix and satisfies the uniform ellipticity condition with entries in $L^\infty(\Omega)$. Let Ω be an open set in $\mathbb{R}^2$. F will be said to be a-harmonic if it satisfies (B.0.1).

B.1 Periodic Media

Let $\Omega = \mathbb{R}^2$. Suppose a is 1-periodic with respect to each of its variables x_1 and x_2. Let

$$\begin{aligned} W^{1,2}_{\sharp}(\mathbb{R}^2,\mathbb{R}^2) \equiv \{U \in W^{1,2}_{\text{loc}}(\mathbb{R}^2,\mathbb{R}^2) | U(x_1+m,x_2+n) = U(x_1,x_2), \\ \text{for a.e. } (x_1,x_2) \in \mathbb{R}^2, \forall m,n \in \mathscr{Z}\}, \end{aligned} \tag{B.1.1}$$

and

$$W^{1,2}_{\sharp,A}(\mathbb{R}^2,\mathbb{R}^2) \equiv \{U \in W^{1,2}_{\text{loc}}(\mathbb{R}^2,\mathbb{R}^2) | U - Ax \in W^{1,2}_{\sharp}(\mathbb{R}^2,\mathbb{R}^2)\}. \tag{B.1.2}$$

The following is the a-harmonic result in periodic setting (Theorem 2.1 of [4]).

Theorem B.1.1. *Let $a \in L^\infty_\sharp(R^2, M^s_K)$ and let A be strict positive definite matrix. If $U^A \in W^{1,2}_{\sharp,A}(\mathbb{R}^2,\mathbb{R}^2)$ is a $a-$harmonic mapping, then U^A is a homeomorphism of $\mathbb{R}^2$ onto itself*

and

$$\det DU^A > 0 \quad \textit{a.e. in } \mathbb{R}^2. \tag{B.1.3}$$

B.2 For Bounded Domain

Let $\Omega \subset \mathbb{R}^2$ be a bounded simply connected open set, whose boundary is a simple closed curve. According to [10], in dimension two if a is smooth then F is a homeomorphism. According to [6], F is always a homeomorphism in dimension two even with $a_{i,j} \in L^\infty(\Omega)$. For example, we have the following result (Theorem 4 of [5]):

Theorem B.2.1. *Let $\Omega \subset \mathbb{R}^2$ be a bounded simply connected open set, whose boundary $\partial\Omega$ is a simple closed curve. Let $\Phi = (\phi_1, \phi_2)$ be a homeomorphism of $\partial\Omega$ onto a convex closed curve Γ and let D be the bounded convex domain bounded by Γ. Let $U \in W^{1,2}_{loc}(\Omega, \mathbb{R}^2) \cap C(\overline{\Omega}, \mathbb{R}^2)$ be the a-harmonic mapping whose components are the solutions of the Dirichlet problems*

$$\operatorname{div}(a\nabla u_i) = 0 \quad \textit{in } \Omega, \tag{B.2.1}$$

$$u_i = \phi_i \quad \textit{on } \partial\Omega \quad i = 1, 2. \tag{B.2.2}$$

Then U is a homeomorphism of $\overline{\Omega}$ onto $\overline{D}$.

Appendix C

C^1 Finite Element Method

As the shape functions on the coarse mesh, we find that C^1 finite elements sharply increase the accuracy compared with C^0 finite elements. C^1 finite elements were developed to satisfy the regularity requirement of higher order partial differential equations, for example, in linear elasticity. The construction of C^1 element is more cumbersome compared to C^0 ones. Traditional C^1 methods use higher order polynomials, for example, Argyris element etc. [13, 32, 53]. However they are generally hard to implement and the solutions using these shape functions may have spurious oscillations. Recently, subdivision schemes using B-splines were introduced to construct C^1 finite elements [77, 52]. In our numerical experiments, we use weighted extended B-splines (WEB) method developed by K. Höllig in [76, 77].

C.1 Multivariate B-splines

In one dimension, the uniform B-spline b^n of degree n is defined by the recursion relation,

$$b^n(x) = \int_{x-1}^{x} b^{n-1}(s)ds. \tag{C.1.1}$$

b^0 is defined as the characteristic function of the unit interval $[0,1)$. $b^n(0) = 0$.

Though there is no unique way to construct multivariate B-splines, the simplest one is to use tensor products of uniform B-splines. For a bounded domain Ω, let $b^n_{k,h}$ be m variate

nth order B-splines with grid width h and index $k = (k_1, \cdots, k_m)$. We define,

$$\mathscr{B}_h^n(\Omega) = \text{span}_{k \in K} b_{k,h}^n. \tag{C.1.2}$$

which the set K of relevant indices contains all k such that $b_{k,h}^n \neq 0$ for some $x \in \Omega$.

C.2 Weight Functions

To make B-splines based finite element approximations comply with a specific boundary condition, we can use a weight function w [114, 115]. For example, for Dirichlet boundary condition, we can choose w such that it is positive and vanishes at the boundary.

Therefore, we define the weighted spline space by

$$w\mathscr{B}_h^n(\Omega) = \text{span}_{k \in K} w b_k. \tag{C.2.1}$$

where K is the set of relevant indices.

C.3 Web-Splines

Although the spaces $\mathscr{B}$ and $w\mathscr{B}$ provide optimal approximation order, the B-spline basis is not uniformly stable with respect to the grid width h. Since some B-splines close to the boundary of the computational domain, Ω may have very little support. Thus, the Galerkin system may be ill-conditioned. The remedy of this problem is to use the Lagrangian interpolation of the 'inside' B-splines to approximate those 'boundary' B-splines. We introduce the following notation: Inner and Outer B-splines. Partition the grid cells $Q = lh + [0,1]^m h$ into interior, boundary, and exterior cells, depending on whether $Q \subseteq \bar{\Omega}$, the interior of Q intersects $\partial\Omega$, or $Q \cap \Omega = \emptyset$. Among the relevant B-splines b_k, $k \in K$, we distinguish between inner B-splines

$$b_i, i \in I, \tag{C.3.1}$$

which have at least one interior cell Q_i in their support, and outer B-splines

$$b_j, j \in J = K \backslash I, \tag{C.3.2}$$

for which suppb_j consists entirely of boundary and exterior cells.

For an outer index $j \in J$, let $I(i) = l + \{0, ..., n\}^m \subset I$ be an m-dimensional array of inner indices closed to j, assuming that h is small enough so that such an array exists. Moreover, denote by

$$e_{i,j} = \prod_{\nu=1}^{m} \prod_{\mu=0, l_\nu+\mu \neq i_\nu}^{n} \frac{j_\nu - l_\nu - \mu}{i_\nu - l_\nu - \mu} \tag{C.3.3}$$

the values of the Lagrange polynomials associated with $I(j)$ and by $J(i)$ the set of all j with $i \in I(j)$. Then the web-splines

$$B_i = \frac{\omega}{\omega(x_i)} [b_i + \sum_{j \in J(i)} e_{i,j} b_j], \quad i \in I \tag{C.3.4}$$

form a basis for the web-space $\omega^e \mathscr{B}_h^n(\Omega)$.

Bibliography

[1] J. E. Aarnes. On the use of a mixed multiscale finite element method for greater flexibility and increased speed or improved accuracy in reservoir simulation. *Multiscale Model. Simul.*, 2(3):421–439, 2004. 1.1

[2] J. E. Aarnes, Y. Efendiev, and L. Jiang. Analysis of multiscale finite element methods using global information for two-phase flow simulations. submitted, 2007. 1.1, 5

[3] A. Abdulle and W. E. Finite difference heterogeneous multi-scale method for homogenization problems. *J. Comput. Phys.*, 191:18–39, 2003. 1.2.2

[4] G. Alessandrini and V. Nesi. Univalence of σ-harmonic mappings and applications, 2000. Newton Institute preprint NI00006-SMM. B.1

[5] G. Alessandrini and V. Nesi. Univalent σ-harmonic mappings. *Arch. Rational Mech. Anal.*, 158:155–171, 2001. B.2

[6] G. Alessandrini and V. Nesi. Univalent σ-harmonic mappings: connections with quasiconformal mappings. *J. Anal. Math.*, 90:197–215, 2003. 1.2.1, 2.1, B.2

[7] G. Allaire. *Shape Optimization by the Homogenization Method*. Springer, 2001. A.1.1

[8] G. Allaire and R. Brizzi. A multiscale finite element method for numerical homogenization. *Multiscale Modeling & Simulation*, 4(3):790–812, 2005. (Preliminary version: internal report, n. 545, CMAP, Ecole Polytechnique, July 2004). 1.1, 1.1, 1.2.1, 1.2.1, 2.2.1, 2.2.1, 2.4, 2.4.1

[9] B. Alpert, G. Beylkin, R. Coifman, and V. Rokhlin. Wavelet-like bases for the fast solution of second-kind integral equations. *SIAM J. Sci. Comput.*, 14(1):159–184, 1993. 1.1

[10] A. Ancona. Some results and examples about the behavior of harmonic functions and Green's functions with respect to second order elliptic operators. *Nagoya Math. J.*, 165:123–158, 2002. 2.1, 2.1, 2.2.3, B.2

[11] P. W. Anderson. Absence of diffusion in certain random lattices. *Phys. Rev.*, 109:1492–1505, 1958. 1.2.3

[12] ANDRA. Dossier 2005 argile, tome *Evolution phénoménologique du stockage géologique*, Juin 2005. 1.2.2

[13] J. H. Argyris and D. W. Scharpf. The sheba family of shell elements for the matrix displacement method. *Aeron. J. Roy. Aeron. Soc.*, 71:873–883, 1968. C

[14] A. Armaou, C. I. Siettos, and I. G. Kevrekidis. Time-steppers and 'coarse' control of distributed microscopic processes. *Internat. J. Robust Nonlinear Control*, 14(2):89–111, 2004. 1.1

[15] K. Astala and V. Nesi. Composites and quasiconformal mappings: new optimal bounds in two dimensions. *Calc. Var. Partial Differential Equations*, 18(4):335–355, 2003. 2.1

[16] K. E. Atkinson and W. Han. *Theoretical Numerical Analysis: A Functional Analysis Framework*, volume 39 of *Texts in Applied Mathematics*. Springer, New York, 2001. 2.2.3, 2.2.3

[17] J. P. Aubin. Behavior of the error of the approximate solutions of boundary value problems for linear elliptic operators by galerkin's and finite difference methods. *Ann. Scuola Norm. Sup. Pisa*, 21:599–637, 1967. 2.3.2.1

[18] A. Averbuch, G. Beylkin, R. Coifman, and M. Israeli. Multiscale inversion of elliptic operators. In *Signal and image representation in combined spaces*, volume 7 of *Wavelet Anal. Appl.*, pages 341–359. Academic Press, San Diego, CA, 1998. 1.1

[19] M. Azaiez, B. F. Belgacem, H. El Fekih, and M. Ismail. Numerical simulation of the wave equation with discontinuous coefficients by nonconforming finite elements. *Numer. Methods Partial Differential Equations*, 15(6):637–656, 1999. 1.2.3, 1.2.3

[20] I . Babuška, G . Caloz, and J .E. Osborn. Special finite element methods for a class of second order elliptic problems with rough coefficients. *SIAM J. Numer. Anal.*, 31(4):945–981, 1994. 1.1

[21] I. Babuška and A. Aziz. Survey lectures on the mathematical foundation of the finite elment method. In *The Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations*, pages 1–359. Academic Press, New York, 1972. 2.2.3

[22] G. A. Baker. Error estimates for finite element methods for second order hyperbolic equations. *Siam J. Numer. Anal.*, 13(4):564–576, 1976. 1.2.3, 4.2.3

[23] G. Bal, J. B. Keller, G. Papanicolaou, and L. Ryzhik. Transport theory for acoustic waves with reflection and transmission at interfaces. *Wave Motion*, 30(4):303–327, 1999. 1.2.3

[24] G. Bal, G. Papanicolaou, and L. Ryzhik. Self-averaging in time reversal for the parabolic wave equation. *Stoch. Dyn.*, 2(4):507–531, 2002. 1.2.3

[25] A. Bamberger, R. Glowinski, and Q. H. Tran. A domain decomposition method for the acoustic wave equation with discontinuous coefficients and grid change. *SIAM J. Numer. Anal*, 34(2):603–639, 1997. 1.2.3, 1.2.3

[26] M.T. Barlow, T. Coulhon, and T. Kumagai. Characterization of sub-gaussian heat kernel estimates on strongly recurrent graphs. *Comm. Pure Appl. Math.*, 58(12):1642–1677, 2005. 1.2.1

[27] M. Bebendorf. Approximate inverse preconditioning of fe systems for elliptic operators with non-smooth coefficients. *Preprint 7/2004, Max-Planck-Institute MiS, Leipzig*, 2005. 1.1

[28] M. Bebendorf. Efficient inversion of galerkin matrices of general second-order elliptic differential operators with nonsmooth coefficients. *Math. Comp.*, 74:1179–1199, 2005. 1.1

[29] M. Bebendorf. Why approximate lu decompositions of finite element discretizations of elliptic operators can be computed with almost linear complexity. *Preprint 8/2005, Max-Planck-Institute MiS, Leipzig*, 2005. 1.1

[30] M. Bebendorf and Y. Chen. Efficient solution of nonlinear elliptic problems using hierarchical matrices with broyden updates. *preprint 51/2005, Max-Planck-Institut MiS, Leipzig*, 2005. 1.1

[31] M. Bebendorf and W. Hackbusch. Existence of $\mathscr{H}$-matrix approximants to the inverse FE-matrix of elliptic operators with L^∞-coefficients. *Numer. Math.*, 95(1):1–28, 2003. 1.1

[32] K. Bell. A refined triangular plate bending finite element. *Internat. J. Numer. Methods Engrg.*, 1:101–122, 1969. C

[33] A. Bensoussan, J. L. Lions, and G. Papanicolaou. *Asymptotic analysis for periodic structure*. North Holland, Amsterdam, 1978. 1.1

[34] S. Bernstein. Sur la generalisation du probleme de dirichlet. *Math. Ann.*, 69:82–136, 1910. 1.1

[35] G. Beylkin, R. Coifman, and V. Rokhlin. Fast wavelet transforms and numerical algorithms. I. *Comm. Pure Appl. Math.*, 44(2):141–183, 1991. 1.1

[36] G. Beylkin, R. Coifman, and V. Rokhlin. Fast wavelet transforms and numerical algorithms. I. In *Wavelets and applications (Marseille, 1989)*, volume 20 of *RMA Res. Notes Appl. Math.*, pages 368–393. Masson, Paris, 1992. 1.1

[37] G. Beylkin and N. Coult. A multiresolution strategy for reduction of elliptic PDEs and eigenvalue problems. *Appl. Comput. Harmon. Anal.*, 5(2):129–155, 1998. 1.1

[38] A. Brandt, R. Ewing, and O. Iliev. Mini-workshop: Numerical upscaling: Theory and applications. *Oberwolfach Reports*, 2:1127–1176, 2005. 5

[39] A. E. Brandt. Methods of systematic upscaling, 2006. *Technical Report MCS06-05. Computer Science and Applied Mathematics, Weizmann Institute of Science.* 1

[40] S. C. Brenner and L. R. Scott. *The Mathematical Theory of Finite Element Methods*. Springer, 2002. second edition. 2.2.2, 2.3.2.1, 3.2.2.1

[41] M. E. Brewster and G. Beylkin. A multiresolution strategy for numerical homogenization. *Appl. Comput. Harmon. Anal.*, 2(4):327–349, 1995. 1.1

[42] F. Brezzi and D. Marini. Subgrid phenomena and numerical schemes. In *Frontiers in numerical analysis (Durham, 2002)*, Universitext, pages 1–16. Springer, Berlin, 2003. 1.1

[43] M. Briane. About nonlocal effects in two-dimension conductivity. *preprint*, 2005. 5

[44] M. Briane, G. W. Milton, and V. Nesi. Change of sign of the corrector's determinant for homogenization in three-dimensional conductivity. *Arch. Ration. Mech. Anal.*, 173(1):133–150, 2004. 2.1, 2.2.3

[45] R. Brizzi and G. Allaire. Report on the numerical approximation of parabolic problems with highly oscillating coefficients using a multiscale finite element method, 2006. *CMAP preprint 598*. 1.2.2

[46] S. Campanato. Un risultato relativo ad equazioni ellittiche del secondo ordine di tipo non variazionale. *Ann. Scuola Norm. Sup. Pisa (3)*, 21:701–707, 1967. 2.3.1.1

[47] G. Chavent, G. Papanicolaou, P. Sacks, and W. W. Symes, editors. *Inverse problems in wave propagation*, volume 90 of *The IMA Volumes in Mathematics and its Appli-*

cations, New York, 1997. Springer-Verlag. (*Papers from the workshop held at the University of Minnesota, Minneapolis, MN, March 6–17, 1995*). 1.2.3

[48] Y. Chen and L. J. Durlofsky. A coupled local-global upscaling approach for simulating flow in highly heterogeneous formations. *Advances in Water Resources*, 26:1041–1060, 2003. 1.1

[49] Y. Chen and L. J. Durlofsky. Adaptive local-global upscaling for general flow scenarios in heterogeneous formations. *Advances in Water Resources*, 26:157–185, 2006. 1.1

[50] H. Cheng, Z. Gimbutas, P. G. Martinsson, and V. Rokhlin. On the compression of low rank matrices. *SIAM. J. Sci. Comput.*, 26(4):1389–1404, 2005. 1.1

[51] A. Chertock and D. Levy. On wavelet-based numerical homogenization. *Multiscale Modeling and Simulation*, 3:65–88, 2004. 1.1

[52] F. Cirak, M. Ortiz, and P. Schröder. Subdivision surfaces: A new paradigm for thin-shell finite-element analysis. *Internat. J. Numer. Methods Eng.*, 47(12):2039–2072, 2000. C

[53] R. W. Clough and J. L. Tocher. Finite element stiffness matrices for analysis of plates in bending. In *Proceedings of the Conference on Matrix Methods in Structural Mechanics*. Wright Patterson A.F.B., Ohio, 1965. C

[54] J. Dolbow, M. A. Khaleel, and J. Mitchell. Multiscale mathematics initiative: A roadmap. Technical report, DOE, 2004. (Prepared for the U.S. Department of Energy under Contract DE-AC05-76RL01830). 1.1

[55] M. Dorobantu and B. Engquist. Wavelet-based numerical homogenization. *SIAM J. Numer. Anal.*, 35(2):540–559 (electronic), 1998. 1.1

[56] Y. R. Efendiev, T. Y. Hou, and X. H. Wu. Convergence of a nonconforming multiscale finite element method. *SIAM J. Numer. Anal.*, 37(3):888–910 (electronic), 2000. 1.1

[57] A. Ern and J.-L. Guermond. *Theory and practice of finite element methods.*, volume 159 of *Applied Mathematical Sciences*. Springer, 2004. 2.2.3, 2.3.2.1, 2.3.2.1, 3.1.2

[58] L. Evans. *Partial Differential Equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, 1997. 3.2.1.1, 3.2.1.1, 4.2.1, A.2.1

[59] R. Eymard, T. Gallouët, and R. Herbin. Finite volume method. In P. G. Ciarlet and J. L. Lions, editors, *Solution of Equations in $\mathbb{R}^n$ (Part 3), Techniques of Scientific Computing (Part 3)*, volume 7 of *Handbook of Numerical Analysis*, pages 713–1020. North-Holland, Amsterdam, 2000. 2.2.4

[60] C. Farhat, I. Harari, and L. P. Franca. The discontinuous enrichment method. *Comput. Methods Appl. Mech. Engrg.*, 190(48):6455–6479, 2001. 1.1

[61] C. Farhat, I. Harari, and U. Hetmaniuk. The discontinuous enrichment method for multiscale analysis. *Comput. Methods Appl. Mech. Engrg.*, 192(28-30):3195–3209, 2003. 1.1

[62] C. L. Farmer. Upscaling: A review. *Numerical Methods in Fluids*, 40:63–78, 2002. 1.1

[63] V. N. Fenchenko and Ē. Ya. Khruslov. Asymptotic behavior of solutions of differential equations with a strongly oscillating coefficient matrix that does not satisfy a uniform boundedness condition. *Dokl. Akad. Nauk Ukrain. SSR Ser. A*, 4:24–27, 95, 1981. 5

[64] M. Fenn and G. Steidl. Fmm and $\mathscr{H}$-matrices: a short introduction to the basic idea. Technical report, Department for Mathematics and Computer Science, University of Mannheim ; TR-2002-008, 2004. 1.1

[65] J. Fish and A. Wagiman. Multiscale finite element method for a locally nonperiodic heterogeneous medium. *Comput. Mech.*, 12(3):164–180, 1993. 1.1

[66] J. Fish and Y. Zheng. Multi-scale enrichment based on partition of unity. *Int. J. Num. Meth. Engng*, 2005. 1.1

[67] D. Gilbarg and N.S. Trudinger. *Elliptic Partial Differential Equations of Second Order*. Springer-Verlag, 1983. 2nd ed. 2.2.2, A.1.1

[68] A. C. Gilbert. A comparison of multiresolution and classical one-dimensional homogenization schemes. *Appl. Comput. Harmon. Anal.*, 5(1):1–35, 1998. 1.1

[69] E. De Giorgi. Sulla differenziabilità e l'analiticità delle estremali degli integrali multipli regolari. *Mem. Accad. Sci. Torino. Cl. Sci. Fis. Mat. Nat. (3)*, 3:25–43, 1957. 2.3.1.1, A.1.1

[70] O. Goubet. Séparation des variables dans le problème de stokes. application à son approximation multiéchelles éléments finis. *C. R. Acad. Sci. Paris Sér. I Math.*, 315(12):1315–1318, 1992. 1.1

[71] L. Greengard and V. Rokhlin. A fast algorithm for particle simulations. *J. Comput. Phys.*, 73(2):325–348, 1987. 1.1

[72] W. Hackbusch. On the multigrid method applied to difference equations. *Computing*, 20(4):291–306, 1978. 1.1

[73] W. Hackbusch, L. Grasedyck, and S. Börm. An introduction to hierarchical matrices. *Math. Bohem.*, 127(2):229–241, 2002. 1.1

[74] V. H. Hoang and C. Schwab. High-dimensional finite elements for elliptic problems with multiple scales. *Multiscale Model. Simul.*, 3(1):168–194 (electronic), 2004/05. 1.1

[75] L. Holden and B. Nielsen. Global upscaling of permeability in heterogeneous reservoirs: the output least squares (ols) method. *Trans. Porous Media*, 40(2):115–143, 2000. 1.1

[76] K. Höllig. *Finite element methods with B-splines*, volume 26 of *Frontiers in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2003. 2.4.1, 4.3, C

[77] K. Höllig, U. Reif, and J. Wipper. Weighted extended B-spline approximation of Dirichlet problems. *SIAM J. Numer. Anal.*, 39(2):442–462, 2001. 2.4.1, C

[78] T. Y. Hou and X. H. Wu. A multiscale finite element method for elliptic problems in composite materials and porous media. *J. Comput. Phys.*, 134(1):169–189, 1997. 1.1, 2.2.1

[79] T. Y. Hou, X. H. Wu, and Y. Zhang. Removing the cell resonance error in the multiscale finite element method via a Petrov-Galerkin formulation. *Commun. Math. Sci.*, 2(2):185–205, 2004. 1.1, 2.2.3, 2.4

[80] L. Jiang, Y. Efendiev, and V. Ginting. Multiscale methods for parabolic equations with continuum spatial scales, 2007. (submitted to Discrete and Continuous Dynamical Systems, Series BDCDS-B). 1.1, 5

[81] V. V. Jikov, S. M. Kozlov, and O. A. Oleinik. *Homogenization of Differential Operators and Integral Functionals*. Springer-Verlag, 1991. 1.1

[82] Ē. Ya. Khruslov. Homogenized models of composite media. In *Composite media and homogenization theory (Trieste, 1990)*, volume 5 of *Progr. Nonlinear Differential Equations Appl.*, pages 159–182. Birkhäuser Boston, Boston, MA, 1991. 5

[83] J. Kigami. Harmonic calculus on limits of networks and its application to dendrites. *J. Funct. Anal.*, 128(1):48–86, 1995. 1.2.1

[84] W. Kohler, G. Papanicolaou, and B. White. Localization and mode conversion for elastic waves in randomly layered media. i. *Wave Motion*, 23(1):1–22, 1996. 1.2.3, 1.2.3

[85] W. Kohler, G. Papanicolaou, and B. White. Localization and mode conversion for elastic waves in randomly layered media. ii. *Wave Motion*, 23(2):181–201, 1996. 1.2.3, 1.2.3

[86] W. Kohler, G. Papanicolaou, and B. White. Localization of low frequency elastic waves. In *Stochastic models in geosystems (Minneapolis, MN, 1994)*, volume 85 of *IMA Vol. Math. Appl.*, pages 209–217. Springer, New York, 1997. 1.2.3, 1.2.3

[87] W. Kohler, G. Papanicolaou, and B. White. Reflection and transmission of acoustic waves by a locally-layered random slab. In *Diffuse waves in complex media (Les Houches, 1998)*, volume 531 of *NATO Sci. Ser. C Math. Phys. Sci.*, pages 347–381. Kluwer, Dordrecht, 1999. 1.2.3

[88] S. H. Lee, P. Jenny, and H. A. Tchelepi. A finite-volume method with hexahedral multiblock grids for modeling flow in porous media. *Comput. Geosci.*, 6(3-4):353–379, 2002. 1.1

[89] S. H. Lee, P. Jenny, and H. A. Tchelepi. Multi-scale finite-volume method for elliptic problems in subsurface flow simulation. *Journal of Computational Physics*, 187:47–67, 2003. 1.1

[90] S. Leonardi. Weighted Miranda-Talenti inequality and applications to equations with discontinuous coefficients. *Comment. Math. Univ. Carolin.*, 43(1):43–59, 2002. 2.3.1.2

[91] R. Li, T. Tang, and P. Zhang. Moving Mesh Methods in Multiple Dimensions Based on Harmonic Maps. *Journal of Computational Physics*, 170:562–588, July 2001. 1

[92] G. M. Lieberman. *Second order parabolic differential equations*. World Scientific, Singapore, 1996. A.2.1

[93] A. Maugeri, D. K. Palagachev, and L. G. Softova. *Elliptic and Parabolic Equations with Discontinuous Coefficients*, volume 109 of *Mathematical Research*. Wiley-VCH, 2000. 1.2.1, 3.2.1.2, A.1.2, A.1.2, A.1.2, A.1.2, A.2.2, A.2.2

[94] P. Ming and X. Yue. Numerical methods for multiscale elliptic problems. *J. of Comput. Phys.*, 214:421–445, 2006. 2.4

[95] E. E. Moise. *Geometric Topology in Dimension 2 and 3*, volume 47 of *Graduate Texts in Mathematics*. Springer Verlag, New York, 1977. 2.2.2

[96] C. Mora-Corral and J. C. Bellido. Approximation of hölder homeomorphisms by piecewise affine homeomorphisms, 2006. preprint. 2.2.2

[97] J. Moser. On Harnack's theorem for elliptic differential equations. *Comm. Pure Appl. Math.*, 14:577–591, 1961. 2.3.1.1, A.1.1

[98] F. Murat and L. Tartar. H-convergence. In *Topics in the mathematical modelling of composite materials*, volume 31 of *Progr. Nonlinear Differential Equations Appl.*, pages 21–43. Birkhäuser Boston, Boston, MA, 1997. 1.1

[99] B. Nadler, S. Lafon, R. R. Coifman, and I. G. Kevrekidis. Diffusion maps, spectral clustering and reaction coordinates of dynamical systems. *Arxiv math.NA/0503445*, 2005. 1.1

[100] J. Nash. Continuity of solutions of parabolic and elliptic equations. *Amer. J. Math.*, 80:931–954, 1958. 2.3.1.1, A.1.1

[101] B .F. Nielsen and A. Tveito. An upscaling method for one-phase flow in heterogeneous reservoirs: a weighted output least squares (wols) approach. *Comput. Geosci.*, 2:92–123, 1998. 1.1

[102] J. A. Nitsche. Ein kriterium für die quasi-optimalität des ritzschen verfahrens. *Numer. Math.*, 11:346–348, 1968. 2.3.2.1

[103] A. A. Oberai and P. M. Pinsky. A multiscale finite element method for the Helmholtz equation. *Comput. Methods Appl. Mech. Engrg.*, 154(3-4):281–297, 1998. 1.1

[104] H. Oh and I. Babuška. The p-version of the finite element method for the elliptic boundary value problems with interfaces. *Comput. Meths. Appl. Mech. Engrg.*, 97:211–231, 1992. 1.1

[105] H. Oh and I. Babuška. The method of auxiliary mapping for the finite element solutions of elasticity problems containing singularities. *J. Comput. Phys.*, 121(2):193–212, 1995. 1.1

[106] H. Owhadi and L. Zhang. Homogenization of parabolic equations with a continuum of space and time scales, 2005. (*in review by SINUM. Preprint available at http://www.acm.caltech.edu/∼owhadi/ and Arxiv math.AP/0512504*). 1

[107] H. Owhadi and L. Zhang. Homogenization of the acoustic wave equation with a continuum of scales., 2006. (*submitted. Preprint available at http://www.acm.caltech.edu/∼owhadi/ and Arxiv math.NA/0604380*). 1

[108] H. Owhadi and L. Zhang. Metric based upscaling. *Communications on Pure and Applied Mathematics*, 60(5):675–723, 2007. (*Preprint available at Arxiv, math.NA/0505223.*). 1, 2

[109] G. Papanicolaou. Mathematical problems in geophysical wave propagation. In *Proceedings of the International Congress of Mathematicians, Vol. I (Berlin, 1998)*, pages 403–427. Deutscher Mathematiker-Vereinigung, 1998. 1.2.3

[110] G. Papanicolaou and L. Ryzhik. Waves and transport. In *Hyperbolic equations and frequency interactions (Park City, UT, 1995)*, volume 5 of *IAS/Park City Math. Ser.*, pages 305–382. Amer. Math. Soc., Providence, RI, 1999. 1.2.3

[111] G. Papanicolaou, L. Ryzhik, and K. Sølna. The parabolic wave approximation and time reversal. *Mat. Contemp.*, 23:139–159, 2002. (*Seventh Workshop on Partial Differential Equations, Part II (Rio de Janeiro, 2001*). 1.2.3

[112] A. Pinkus. *n-Widths in approximation theory*. Springer-Verlag, Berlin, 1985. 5

[113] J. Ruge and K. Stüben. Algebraic multigrid. In Stephen F. McCormick, editor, *Multigrid Methods*, pages 73–130. SIAM, 1987. 1.1

[114] V. L. Rvachev and T. I. Sheiko. R-functions in boundary value problems in mechanics. *Appl. Mech. Rev.*, 48:151–188, 1995. C.2

[115] V. L. Rvachev, T. I. Sheiko, V. Shapiro, and I. Tsukanov. On completeness of rfm solution structures. *Comp. Mech.*, pages 305–316, 2000. C.2

[116] L. Ryzhik, G. Papanicolaou, and J. B. Keller. Transport equations for elastic and other waves in random media. *Wave Motion*, 24(4):327–370, 1996. 1.2.3

[117] M. Sahimi. *Heterogeneous Materials I.* Interdisciplinary Applied Mathematics. Springer, 2003. 1.1

[118] P. Sheng. *Introduction to wave scattering, localization and mesoscopic phenomena*. Academic, San Diego, 1995. 1.2.3

[119] V. B. Shenoy, R. Miller, E. B. Tadmor, D. Rodney, R. Phillips, and M. Ortiz. An adaptive finite element approach to atomic-scale mechanics—the quasicontinuum method. *J. Mech. Phys. Solids*, 47(3):611–642, 1999. 1.1

[120] L. Simon. Global estimates of Hölder continuity for a class of divergence-form elliptic equations. *Arch. Rational Mech. Anal.*, 56:253–272, 1974. 2.2.2, 2.3.1.1, A.1.1

[121] S. Spagnolo. Sulla convergenza di soluzioni di equazioni paraboliche ed ellittiche. *Ann. Scuola Norm. Sup. Pisa (3) 22 (1968), 571-597; errata, ibid. (3)*, 22:673, 1968. 1.1

[122] R. S. Strichartz. Function spaces on fractals. *J. Funct. Anal.*, 198(1):43–83, 2003. 1.2.1

[123] D. W. Stroock and W. A. Zheng. Markov chain approximations to symmetric diffusions. *Ann. Inst. H. Poincaré Probab. Statist.*, 33(5):619–649, 1997. 1.2.1

[124] W. W. Symes. Mathematics of reflection seismolgy, 1998. (*Lecture notes, available at http://www.trip.caam.rice.edu/txt/tripinfo/other_list.html*). 1.2.3

[125] L. Tartar. Homogénéisation et compacité par compensation. In *Séminaire Goulaouic-Schwartz (1978/1979)*, pages Exp. No. 9, 9. École Polytech., Palaiseau, 1979. 1.1

[126] P. Vaněk, J. Mandel, and M. Brezina. Algebraic multigrid based on smoothed aggregation for second and fourth order problems. *Computing*, 56(3):179–196, 1996. 1.1

[127] T. Vdovina, S. E. Minkoff, and O. Korostyshevskaya. Operator upscaling for the acoustic wave equation. *Multiscale Model. Simul.*, 4(4):1305–1338, 2005. 1.2.3, 1.2.3

[128] B. Engquist W. E, X. Li, W. Ren, and E. Vanden-Eijnden. The heterogeneous multiscale method: A review., 2004. (*Preprint, http://www.math.princeton.edu/multiscale/review.pdf*). 1.1

[129] W. L. Wan, Tony F. Chan, and Barry Smith. An energy-minimizing interpolation for robust multigrid methods. *SIAM J. Sci. Comput.*, 21(4):1632–1649 (electronic), 1999/00. 1.1

[130] X. H. Wu, Y. Efendiev, and T. Y. Hou. Analysis of upscaling absolute permeability. *Discrete Contin. Dyn. Syst. Ser. B*, 2(2):185–204, 2002. 1.1, 1.1, 5

[131] X. H. Wu, Y. Efendiev, and T. Y. Hou. Accurate multiscale finite element methods for two-phase flow simulations. *J. Comput. Phys.*, 220(1):155–174, 2006. 1.1, 1.1, 5

[132] P .M. De Zeeuw. Matrix-dependent prolongation and restrictions in a blackbox multigrid solver. *J. Comput. Applied Math.*, 33(1), 1990. 1.1