M2AN. MATHEMATICAL MODELLING AND NUMERICAL ANALYSIS - MODÉLISATION MATHÉMATIQUE ET ANALYSE NUMÉRIQUE

M. AVELLANEDA
TH. Y. HOU
G. C. PAPANICOLAOU

Finite difference approximations for partial differential equations with rapidly oscillating coefficients

M2AN. Mathematical modelling and numerical analysis - Modélisation mathématique et analyse numérique, tome 25, nº 6 (1991), p. 693-710

<http://www.numdam.org/item?id=M2AN_1991__25_6_693_0>

MATHEMATICAL MODELLING AND NUMERICAL ANALYSIS
MODÉLISATION MATHÉMATIQUE ET ANALYSE NUMÉRIQUE
(Vol. 25, n° 6, 1991, p. 693 à 710)

FINITE DIFFERENCE APPROXIMATIONS FOR PARTIAL DIFFERENTIAL EQUATIONS WITH RAPIDLY OSCILLATING COEFFICIENTS (*)

M. AVELLANEDA (1), Th. Y. HOU (1) and G. C. PAPANICOLAOU (1)

Communicated by R. TEMAM

Abstract. — *It is possible to solve numerically problems with rapidly oscillating coefficients (homogenization) without resolving the rapid oscillations? In some cases, in one dimensional problems and in some multidimensional hyperbolic problems, this can be done by using grids that are irregularly spaced relative to the rapid oscillations. In this paper we show that this simple idea does not generalize to multidimensional elliptic problems except when the coefficients are periodic.*

Résumé. — *Peut-on résoudre numériquement des équations avec des coefficients à variation rapide (homogénéisation) sans un échantillonnage détaillé des échelles les plus fines? Pour certains problèmes monodimensionnels, ainsi que pour certaines équations hyperboliques, cela peut se faire en utilisant des maillages placés de façon irrégulière par rapport aux échelles d'oscillation rapide. Dans cette note, nous montrons que cette méthode simple ne peut se généraliser aux problèmes elliptiques multidimensionnels, à l'exception du cas où les coefficients sont périodiques.*

1. INTRODUCTION

The asymptotic analysis or homogenization of solutions of partial differential equations with rapidly oscillating coefficients has been studied in many different settings [1-12], for elliptic, parabolic or wave equations, for equations with periodic, almost periodic, random coefficients, etc. When the analysis is successful the limit problem is a differential equation, usually but not always of the same form as the original, with coefficients that do not oscillate. The homogenized or effective coefficients are obtained by solving a canonical, local or cell problem. In the periodic case, for example, one has

(*) Received June 1990.
(1) Courant institute, New York University, 251 Mercer Street, New York, NY 10012, U.S.A.

to solve a problem over one period cell to determine the effective coefficients.

How is homogenization to be implemented numerically? The obvious answer is to calculate the effective coefficients by solving a cell problem and then to solve the homogenized equations. Standard numerical methods can be used in both steps since no rapidly oscillating coefficients are involved. However, this simple recipe works well only when the coefficients are periodic, and even in that case it can lead to an expensive computation when the effective coefficients are not constant since, typically, several cell problems have to be solved. In the almost periodic or random case the cell problem is difficult to solve numerically. We need a numerical method that « captures » efficiently the homogenized problem without resorting explicitly to a cell problem. Efficiency means here that, if $\varepsilon > 0$ denotes the parameter involved in the homogenization (ratio of microscopic to macroscopic length scales) and $h > 0$ denotes the size of a finite difference mesh for a discrete version of the problem, then h need not be much smaller than ε in order to get a good numerical solution.

In fact we can take $\varepsilon \sim h$ and still get good approximate solutions if the numerical grid samples well the oscillating coefficients. This idea is due to Engquist [13] and was subsequently developed by Engquist and Hou in [14, 15]. They considered systems of semilinear hyperbolic equations with rapidly oscillating initial data. The propagation of oscillations and the homogenization of such problems was considered earlier by Tartar [16] and by McLaughlin, Papanicolaou and Tartar [17]. The main result in [13-15] is that if the oscillations in the data are periodic of period ε and the numerical grid has size h then we get convergence of the numerical solution to the homogenized solution as $\varepsilon \to 0$, $h \to 0$ without requiring h/ε to be small, provided that the numerical grid spacings are irrational multiples of the period.

The purpose of this paper is to formulate this idea of capturing homogenization numerically by sampling for second order elliptic equations and to show that when homogenization is formulated in a suitable abstract framework [6] which is natural, then the sampling scheme of Engquist [13] and Engquist-Hou [14, 15] can be analyzed through an application of the ergodic theorem. In the special case of one space dimension both the usual homogenization [4] and the numerical homogenization of [13-15] can be obtained by elementary, essentially explicit computations. We review this one-dimensional case in section 2.

In several space dimensions we find that only in the case of rapidly oscillating *periodic* coefficients do the results of Engquist-Hou [13, 15] generalize, in a weaker form. In the case of almost periodic or random coefficients in several space dimensions we show, both theoretically and with a simple counterexample, that numerical homogenization by sampling does not work efficienty, at least in the manner in which the sampling idea is implemented here. There may, of course, be other ways of doing the discretization and the sampling that do capture homogenization effectively. In section 3 we present our main results for the multidimensional problem after introducing the relevant ideas from [6] and their finite difference version [18, 19].

2. THE ONE-DIMENSIONAL CASE

Consider the one-dimensional elliptic boundary value problem

$$-\frac{d}{dx}\left[a\left(x, \frac{x}{\varepsilon}\right)\frac{d}{dx}u^{\varepsilon}(x)\right] = f(x), \qquad 0 < x < 1, \tag{2.1}$$

$$u^{\varepsilon}(0) = b, \qquad u^{\varepsilon}(1) = c. \tag{2.2}$$

Here $a(x, y)$ is a smooth function, unit periodic in y, positive and bounded

$$0 < a_1 \leq a(x, y) \leq a_2 < \infty .$$

It is easily seen that

$$\sup_{0 \leq x \leq 1} |u^{\varepsilon}(x) - \bar{u}(x)| \to 0 \tag{2.3}$$

as $\varepsilon \to 0$ where $\bar{u}(x)$ is the solution of the homogenized problem

$$-\frac{d}{dx}\left[a^*(x)\frac{d}{dx}\bar{u}(x)\right] = f(x), \qquad 0 < x < 1,$$

$$\bar{u}(0) = b, \qquad \bar{u}(1) = c \tag{2.4}$$

and

$$a^*(x) = \left[\int_0^1 \frac{1}{a(x, y)}\, dy\right]^{-1}. \tag{2.5}$$

One can get this result using asymptotic expansions with two scales [4] or directly from the explicit solution of (2.1) and the fact that if $g(x, y)$ is a bounded continuous function of x and periodic in y of period one, then

$$\lim_{\varepsilon \to 0} \int_0^1 g\left(x, \frac{x}{\varepsilon}\right) dx = \int_0^1 \int_0^1 g(x, y)\, dy\, dx . \tag{2.6}$$

The simplest finite difference approximation to (2.1) has the form

$$-\nabla^{-h}\left[a\left(x_i, \frac{x_i}{\varepsilon}\right)\nabla^{+h} u_h^{\varepsilon}(x_i)\right] = f(x_i), \qquad 1 \leq i \leq n-1, \tag{2.7}$$

$$u_h^{\varepsilon}(x_0) = b, \qquad u_h^{\varepsilon}(x_n) = c \tag{2.8}$$

in which

$$x_i = ih\,, \qquad nh = 1 \tag{2.9}$$

and

$$\nabla^{\pm h} g(x_i) = \frac{g(x_i \pm h) - g(x_i)}{h}\,. \tag{2.10}$$

We want to know how well u_h^ε approximates $\bar{u}$. Clearly if h is much smaller than ε, u_h^ε is close to u^ε. A more interesting situation is contained in the following theorem.

THEOREM 1 : *For any bounded continuous function $f(x)$, $0 \leqslant x \leqslant 1$*

$$\lim_{\varepsilon, h \to 0} \sup_{0 \leqslant i \leqslant n} \left| u_h^\varepsilon(x_i) - \bar{u}(x_i) \right| = 0 \tag{2.11}$$

where in the limit the ratio

$$\frac{\varepsilon}{h} = r \tag{2.12}$$

is held fixed and r is any irrational number.

In this one-dimensional case we can actually prove results that contain more information. Specifically, we have the following version of Theorem 1.

THEOREM 1a : *For any bounded continuous function $f(x)$, $0 \leqslant x \leqslant 1$ and any $\tau > 0$, there is an $h_0 > 0$ and a set $S(\varepsilon, h_0) \subset [0, h_0]$ with Lebesgue measure $|S(\varepsilon, h_0)| \geqslant (1 - \tau)\, h_0$, such that*

$$\sup_{0 \leqslant i \leqslant n} \left| u_h^\varepsilon(x_i) - \bar{u}(x_i) \right| \leqslant \tau \tag{2.13}$$

for all $h \in S(\varepsilon, h_0)$ and any $0 < \varepsilon \leqslant 1$.

The type of convergence given in this theorem is due to Engquist [13], who calls it *convergence essentially independent of* ε.

The proof of both Theorem 1 and 1a follows easily from the explicit solution of (2.7), (2.8). Let

$$V_i^{\varepsilon, h} = a(x_{i+1/2}, x_{i+1/2}/\varepsilon)\, \nabla^{+h} u_h^\varepsilon(x_i)\,. \tag{2.14}$$

Then

$$-\frac{V_i^{\varepsilon, h} - V_{i-1}^{\varepsilon, h}}{h} = f(x_i) \tag{2.15}$$

and hence

$$V_i^{\varepsilon, h} = V_0^{\varepsilon, h} - \sum_{k=1}^{i} f(x_k)\, h \tag{2.16}$$

or

$$u^{\varepsilon, h}(x_i) = b + h \sum_{j=1}^{i} \frac{1}{a(x_{j+1/2}, x_{j+1/2}/\varepsilon)} \left(V_0^{\varepsilon, h} - \sum_{k=1}^{j} f(x_k)\, h \right). \tag{2.17}$$

The constant $V_0^{\varepsilon, h}$ in (2.17) is determined from the boundary condition $u^{\varepsilon, h}(x_n) = c$.

To study the limit of $u^{\varepsilon, h}$ from (2.17) we need a lemma. To simplify the argument in this lemma, which is an elementary ergodic theorem [20], we make liberal assumptions about smoothness.

LEMMA 1 : *Suppose $g(x, y) \in C^3([a, b] \times [0, 1])$ and is periodic in y of period one. Let $x_k = kh$ and $r = h/\varepsilon$.* If *r is an irrational number, then we have*

$$\left| \sum_{k=1}^{n} g\left(x_k, \frac{x_k}{\varepsilon} \right) h - \int_a^b \left(\int_0^1 g(x, y)\, dy \right) dx \right| \leqslant C(r)\, h \tag{2.18}$$

where $C(r)$ is a constant that depends on r. If moreover $h \in S(\varepsilon, h_0)$ where

$$S(\varepsilon, h_0) = \left\{ 0 < h \leqslant h_0 \,\middle|\, \frac{kh}{\varepsilon} \in \left(i - \frac{\tau}{|k|^{3/2}}, i + \frac{\tau}{|k|^{3/2}} \right) \right.$$

$$\left. \text{for} \quad i = 1, 2, \dots, \left[\frac{kh_0}{\varepsilon} \right] + 1, \quad 0 \neq k \in Z, \quad 0 < \varepsilon \leqslant 1 \right\} \tag{2.19}$$

then

$$C(r) \leqslant \frac{C_0}{\tau} \tag{2.20}$$

with C_0 a constant and

$$|S(\varepsilon, h_0)| \geqslant h_0 \left(1 - \tau \sum_{k=1}^{\infty} \frac{1}{k^{3/2}} \right) \geqslant h(1 - 3\tau). \tag{2.21}$$

Remark :

If we take $\tau = \sqrt{h}$ then for all $0 < \varepsilon \leqslant 1$ there is an $h_0 > 0$ independent of ε such that for all $h \in S(\varepsilon, h_0)$ we have

$$\left| \sum_{k=1}^{n} g\left(x_k, \frac{x_k}{\varepsilon} \right) h - \int_a^b \int_0^1 g(x, y)\, dy\, dx \right| \leqslant c_0 \sqrt{h} \tag{2.22}$$

and

$$|S(\varepsilon, h_0)| \geq h_0(1 - 3\sqrt{h}). \tag{2.23}$$

Proof of Lemma 1 :

Since g has continuous derivatives we make at most an error of order h if we replace on the left side of (2.18) the integral over x by a sum. We shall show that

$$\left| \sum_{k=1}^{n} g\left(x_k, \frac{x_k}{\varepsilon}\right) h - \sum_{k=1}^{n} \left[\int_0^1 g(x_k, y)\, dy \right] h \right| \leq \frac{C_0 h}{\tau} \tag{2.24}$$

when $h \in S(\varepsilon, h_0)$ as defined by (2.19). Define

$$\tilde{g}(x, y) = g(x, y) - \int_0^1 g(x, y)\, dy .$$

Since $\tilde{g}$ is periodic in y with mean zero it can be expanded in a Fourier series

$$\tilde{g}(x, y) = \sum_{m \neq 0} a_m(x)\, e^{2\pi i m y} \tag{2.25}$$

and since $\tilde{g}(x, y) \in C^3$ we have

$$|a_m(x)| \leq \frac{C}{|m|^3} .$$

This implies that

$$\left| \sum_{k=1}^{n} \tilde{g}\left(x_k, \frac{x_k}{\varepsilon}\right) h \right| = \left| \sum_{k=1}^{n} \sum_{m \neq 0} a_m(x_k)\, e^{2\pi i m x_k/\varepsilon} h \right|$$

$$= \left| \sum_{m \neq 0} \sum_{k=1}^{n} a_m(x_k)\, e^{2\pi i m x_k/\varepsilon} h \right| .$$

Summation by parts yields

$$\sum_{k=1}^{n} a_m(x_k)\, e^{2\pi i m x_k/\varepsilon} = a_m(x_n) \sum_{k=1}^{n} e^{2\pi i m x_k/\varepsilon} +$$

$$+ \sum_{k=1}^{n-1} \left(\sum_{j=1}^{k} e^{2\pi m x_j/\varepsilon} \right) (a_m(x_k) - a_m(x_{k+1})) .$$

Therefore, we obtain

$$\left| \sum_{k=1}^{n} \tilde{g}\left(x_k, \frac{x_k}{\varepsilon}\right) h \right| \leq \sum_{m \neq 0} \frac{ch}{|m|^3\, |1 - e^{2\pi i m h/\varepsilon}|} .$$

But for $h \in S(\varepsilon, h_0)$ we have

$$|1 - e^{2\pi imh/\varepsilon}| = 2|\sin(\pi mh/\varepsilon)| \geqslant \frac{4\tau}{m^{3/2}} m \neq 0 .$$

Hence for $h \in S(\varepsilon, h_0)$

$$\left| \sum_{k=1}^{n} \tilde{g}\left(x_k, \frac{x_k}{\varepsilon}\right) h \right| \leqslant \frac{ch}{4\tau} \sum_{m \neq 0} \frac{1}{m^{3/2}} \leqslant \frac{C_0 h}{\tau} . \tag{2.26}$$

The Lebesgue measure of the set $S(\varepsilon, h_0)$ is easily calculated from (2.19) so that (2.21) holds. The estimate (2.26) and the remarks above (2.24) complete the proof of the lemma.

3. THE MULTI-DIMENSIONAL CASE

The direct analysis of the previous section does not go over in the multi-dimensional case because we do not have an explicit representation such as (2.17). Moreover, we will see that although a result similar to Theorem 1 can be obtained there are essential differences in the multi-dimensional case. It will turn out however that once the problem is suitably formulated, a generalization of Theorem 1 suggests itself in the periodic case immediately and its proof is quite simple. But no useful generalization is possible in the almost periodic or random case. In the next section we introduce the abstract framework and summarize the relevant results from [6] that we need. In section 3.2 we review the finite difference version of the results in [6], as carried out in detail in [18] and in [19]. Our main theorems are stated in section 3.3 along with their proof.

3.1. Abstract framework for homogenization

Let (Ω, F, P) be a probability space and let $(a_{ij}(y, \omega))$, $i, j = 1, 2, \dots, d$ be a strictly stationary matrix-valued random field with $y \in \mathbf{R}^d$. We assume throughout that

$$a_0 |\xi|^2 \leqslant \sum_{i,j=1}^{d} a_{ij}(y, \omega)\, \xi_i\, \xi_j \leqslant \frac{1}{a_0} |\xi|^2 , \tag{3.1.1}$$

for all $y \in \mathbf{R}^d$, $\omega \in \Omega$ and $\xi = (\xi_1, \dots, \xi_d)$. We want to consider the asymptotic behavior of the elliptic boundary value problem.

$$-\sum_{i,j=1}^{d} \frac{\partial}{\partial x_i} \left[a_{ij}\left(\frac{x}{\varepsilon}, \omega\right) \frac{\partial u^\varepsilon(x, \omega)}{\partial x_j} \right] = f(x) , \qquad x \in O$$

$$u^\varepsilon(x, \omega) = 0 , \qquad x \in \partial O \tag{3.1.2}$$

where O is a bounded open subset of $\mathbf{R}^d$ and $f(x) \in L^2(O)$. The asymptotic analysis of $u^\varepsilon(x, \omega)$ as $\varepsilon \downarrow 0$ is the random homogenization problem studied in [6]. It contains the case of periodic and almost periodic coefficient as a special case as we will explain. The main result in [6] is that if the stationary stochastic coefficient are ergodic with respect to space translations, then there exist constants a_{ij}^*, $i, j = 1, 2, \ldots, d$, the effective coefficients, such that if $\bar{u}(x)$ satisfies

$$- \sum_{i,j=1}^{d} a_{ij}^* \frac{\partial^2 \bar{u}(x)}{\partial x_i \, \partial x_j} = f(x) , \qquad x \in O ,$$

$$\bar{u}(x) = 0 , \qquad x \in \partial O \tag{3.1.3}$$

we have

$$\lim_{\varepsilon \downarrow 0} \int_O \int_\Omega |u^\varepsilon(x, \omega) - \bar{u}(x)|^2 \, P(d\omega) \, dx = 0 . \tag{3.1.4}$$

A similar result holds for a finite difference version of (3.1.2) and that is discussed in the next section. To explain role of ergodicity in the convergence of u^ε to $\bar{u}$ and to give the characterization of the effective coefficients a_{ij}^* we need the abstract framework of [6].

We may take the probability space (Ω, F, P) as follows. The set Ω is the set of Lebesgue measurable $d \times d$ matrix-valued function on $\mathbf{R}^d$ satisfying (3.1.1). The value of $\omega \in \Omega$ at $y \in \mathbf{R}^d$ is defined almost everywhere and is denoted by $\omega_{ij}(y, \omega)$. Thus Ω is the set of all coefficients for (3.1.2). We take F to be the σ-algebra generated by cylinder sets and arrange that it be countably generated (stochastic continuity suffices for this). The probability measure P defined on (Ω, F) is invariant with respect to the group of translations $\tau_x : \Omega \to \Omega$

$$(\tau_x \omega)(y) = \omega(y - x), x, y \in \mathbf{R}^d . \tag{3.1.5}$$

We assume that the action of τ_x is ergodic. That is the only sets $A \in F$ that are invariant ($\tau_x A \subset A$) have $P(A) = 0$ or 1.

Let $H = L^2(\Omega, F, P ; \mathbf{R})$ be the Hilbert space of square integrable random variables on Ω with inner product

$$(\tilde{f}, \tilde{g}) = \int_\Omega \tilde{f}(\omega) \, \tilde{g}(\omega) \, P(d\omega), \tilde{f}, \tilde{g} \in H .$$

If $\tilde{f} \in H$, define for P-almost all ω the operators

$$(T_x \tilde{f})(\omega) = \tilde{f}(\tau_{-x} \omega), x \in \mathbf{R}^d . \tag{3.1.6}$$

These operators form a unitary group on H which is strongly continuous (this follows from stochastic continuity of $\omega_{ij}(y, \omega)$ which is the same as $a_{ij}(y, \omega)$ in (3.1.2)). With $\tilde{f}$ in H we associate a stationary process

$$f(x, \omega) = (T_x \tilde{f})(\omega) = \tilde{f}(\tau_{-x} \omega) . \tag{3.1.7}$$

If for example we define

$$\tilde{a}_{ij}(\omega) = \omega_{ij}(0, \omega) \quad i, j = 1, 2, \dots, d \tag{3.1.8}$$

for P-almost all ω, then

$$a_{ij}(x, \omega) = (T_x \tilde{a}_{ij})(\omega) = \tilde{a}_{ij}(\tau_{-x} \omega) = \omega_{ij}(x, \omega) \tag{3.1.9}$$

as noted already.

Let D_i, $i = 1, 2, \dots, d$ denote the infinitesimal generators of the strongly continuous, unitary group T_x on H, $x \in \mathbf{R}^d$. They defined on dense subsets D_i of H by

$$D_i = \frac{\partial}{\partial x_i} T_x|_{x=0} \quad i = 1, 2, \dots, d . \tag{3.1.10}$$

Following [21] we can now define the effective coefficients a_{ij}^* as follows. Let $\bar{E}_j$ $j = 1, 2, \dots, d$ be given constants. Construct square integrable random variables $\tilde{E}_j(\omega)$ and $\tilde{F}_j(\omega)$, $j = 1, 2, \dots, d$ such that

$$\tilde{F}_i(\omega) = \sum_{j=1}^{d} \tilde{a}_{ij}(\omega) \tilde{E}_j(\omega) \tag{3.1.11}$$

$$D_i \tilde{E}_j = D_j \tilde{E}_i , \quad \text{weakly} , \quad i, j = 1, 2, \dots, d \tag{3.1.12}$$

$$\sum_{i=1}^{d} D_i \tilde{F}_i = 0 , \quad \text{weakly} \tag{3.1.13}$$

$$\int_{\Omega} \tilde{E}_j(\omega) P(d\omega) = \bar{E}_j, j = 1, 2, \dots, d \tag{3.1.14}$$

Then

$$\sum_{J=1}^{d} a_{ij}^* \bar{E}_j = \int_{\Omega} \tilde{F}_i(\omega) P(d\omega) . \tag{3.1.15}$$

Problem (3.1.11)-(3.1.15) is, of course, the abstract analog of the cell problem in the periodic case. In fact the periodic case is realized in this framework as follows. Let $\tilde{a}_{ij}(x)$ be the given periodic coefficients which are bounded measurable functions satisfying (3.1.1). Let T^d be the d-dimensional unit torus. We take

$$\Omega = \{\tilde{a}_{ij}(\cdot + \omega), \omega \in T^d\}$$

so that Ω is identical with T^d and is in this case a much smaller set than what it is in general. Let F be the σ-algebra of Lebesgue measurable sets and P be Lebesgue measure on T^d. It is invariant under translation $\tau_x \omega = \omega - x$ (mod 1). The action of τ_x on Ω is ergodic and the infinitesimal generators of the group T_x are the usual partial derivatives

$$D_i = \frac{\partial}{\partial \omega_i} \quad (\omega) = (\omega_i, \dots \omega_d) .$$

The coefficients a_{ij} are given by $a_{ij}(x, \omega) = \tilde{a}_{ij}(x + \omega)$.

The periodic case can be put into the abstract framework by the essentially trivial process of letting the center of the period cell be a random variable that is uniformly distributed over the unit torus in d-dimension. The way the case of the almost periodic coefficients fits into the abstract framework is described in [6].

The ergodicity of the random coefficients is essential in proving the convergence (3.1.4) with the limit problem having coefficients a_{ij}^* that are constant, independent of ω. Without ergodicity the effective coefficients will be functions of ω, $a_{ij}^*(\omega)$, that are τ_x invariant.

3.2. The discrete case

A finite difference version of (3.1.2) is as follows. Let $\mathbf{h} = (h_1, h_2, h_d)$ be a vector with positive components and denote by $\Lambda_{\mathbf{h}}$ the subgroup of $\mathbf{R}^d$

$$\Lambda_{\mathbf{h}} = \{y \in \mathbf{R}^d \mid y = (z_1 h_1, z_2 h_2, \dots, z_d h_d), z_j \in \mathbf{Z}, 1 \leq j \leq d\} . \quad (3.2.1)$$

Let e_j, $j = 1, \dots, d$ be unit vectors in the coordinate directions. Define

$$\nabla_i^{\pm \mathbf{h}} g(x) = \frac{1}{\pm h_i} [g(x \pm h_i e_i) - g(x)] , \quad (3.2.2)$$

with $i = 1, 2, \dots, d$, and let $\Lambda_{\mathbf{h}}^O$ be the intersection of O and $\Lambda_{\mathbf{h}}$. The discrete version of (3.1.2) is

$$- \sum_{i,j=1}^{d} \nabla_i^{-\mathbf{h}} \left[a_{ij} \left(\frac{x}{\varepsilon}, \omega \right) \nabla_j^{+\mathbf{h}} u_{\mathbf{h}}^{\varepsilon}(x, \omega) \right] = f(x), x \in \Lambda_{\mathbf{h}}^O$$

$$u_{\mathbf{h}}^{\varepsilon}(x, \omega) = 0 , \quad x \in \partial \Lambda_{\mathbf{h}}^O . \quad (3.2.3)$$

The boundary of the discrete set $\Lambda_{\mathbf{h}}^{O}$ is defined as the set of points in the complement of $\Lambda_{\mathbf{h}}^{O}$ with at least one nearest neighbor in $\Lambda_{\mathbf{h}}^{O}$.

Let

$$\mathbf{h} = |\mathbf{h}|(r_1, r_2, \dots, r_d) = h\mathbf{r} . \tag{3.2.4}$$

With $\varepsilon > 0$ and $r_j > 0$ $j = 1, \dots, d$ fixed, the solution $u_{\mathbf{h}}^{\varepsilon}(x, \omega)$ of (3.2.3) converges to $u^{\varepsilon}(x, \omega)$ of (3.1.2)

$$\lim_{h \downarrow 0} \sum_{x \in \Lambda_{h\mathbf{r}}^{O}} h^d \int_{\Omega} |u_{\mathbf{h}}^{\varepsilon}(x, \omega) - u^{\varepsilon}(x, \omega)|^2 P(d\omega) = 0 . \tag{3.2.5}$$

The interesting case to consider is when $h = \varepsilon$, that is $\mathbf{h} = \varepsilon\mathbf{r}$ with $\mathbf{r} = (r_1, r_2, \dots, r_d)$ fixed and with positive components. Then we have the discrete homogenization problem studied by Kuhnemann [18] and in [19]. The main convergence result is this : if the restriction of the translation group τ_x to $x \in \Lambda_{\mathbf{r}}$ is an ergodic subgroup for P then there exist constants $a_{ij}^*(\mathbf{r})$, depending on the lattice scale factors $\mathbf{r}$, such that if $\bar{u}_{\mathbf{r}}(x)$ is the solution of

$$-\sum_{i,j=1}^{d} a_{ij}^*(\mathbf{r}) \frac{\partial^2 \bar{u}_{\mathbf{r}}(x)}{\partial x_i \, \partial x_j} = f(x), x \in O$$

$$\bar{u}(x) = 0, \quad x \in \partial O \tag{3.2.6}$$

$$\lim_{\varepsilon \to 0} \sum_{x \in \Lambda_{\varepsilon\mathbf{r}}^{O}} (\varepsilon |\mathbf{r}|)^d \int_{\Omega} |u_{\varepsilon\mathbf{r}}^{\varepsilon}(x, \omega) - \bar{u}_{\mathbf{r}}(x)| \, P(d\omega) = 0 . \tag{3.2.7}$$

We note again that even if the subgroup τ_x, $x \in \Lambda_{\mathbf{r}}$ is ergodic for P, the effective coefficients $a_{ij}^*(\mathbf{r})$ are constants that depend on the lattice scales $\mathbf{r}$. This does not happen in one space dimension. Before explaining this let us first write the discrete analog of (3.1.11)-(3.1.15). This discrete analog of the operators D_i defined by (3.1.10) and corresponding to (3.2.2) is

$$D_i^{\pm\mathbf{h}} \tilde{g}(\omega) = \frac{1}{\pm h_i} [(T_{\pm h_i e_i} \tilde{g})(\omega) - \tilde{g}(\omega)] , \tag{3.2.8}$$

Then the problem that determines $a_{ij}^*(\mathbf{r})$ is: given constants $\bar{E}_j$, $j = 1, 2, \dots, d$, find square integrable random variables $\tilde{E}_j(\omega)$, $\tilde{F}_j(\omega)$, $j = 1, 2, \dots, d$ such that

$$\tilde{F}_i(\omega) = \sum_{j=1}^{d} \tilde{a}_{ij}(\omega) \tilde{E}_j(\omega) \tag{3.2.9}$$

$$D_i^{+\mathbf{r}} \tilde{E}_j = D_j^{+\mathbf{r}} \tilde{E}_i \,, \quad \text{weakly}\,, \tag{3.2.10}$$

$$\sum_{j=1}^{d} D_j^{-\mathbf{r}} \tilde{F}_j = 0\,, \quad \text{weakly}\,, \tag{3.2.11}$$

$$\int_\Omega \tilde{E}_j(\omega)\, P(d\omega) = \bar{E}_j \tag{3.2.12}$$

and then

$$\sum_{j=1}^{d} a_{ij}^*(\mathbf{r})\, \bar{E}_j = \int_\Omega \tilde{F}_i(\omega)\, P(d\omega)\,. \tag{3.2.13}$$

In one space dimension the effective coefficient is given by

$$a^* = \left[\int_\Omega \frac{1}{\tilde{a}(\omega)}\, P(d\omega) \right]^{-1}. \tag{3.2.14}$$

As we saw in section 2 we can also derive (3.2.14) by solving (3.1.11)-(3.1.15). The main step in this calculation is to note that (3.1.13) in one dimension implies that $\tilde{F}$ is a constant, in the ergodic case. We have the same conclusion from (3.2.11) in the discrete situation when we have ergodicity. Thus $a^*(\mathbf{r})$ is again given by formula (3.2.14).

So in the one-dimensional case the only thing that matters is the ergodicity of the translation subgroup τ_x, $x \in \Lambda_\mathbf{r}$ relative to P. We thus recover immediately Theorem 1 of section 2 in the periodic case because if r is irrational and τ_x are translations on T^1, the subgroup τ_{nr}, $n \in \mathbf{Z}$, is ergodic for Lebesgue measure on the one-dimensional unit torus T^1.

3.3. Consistency of the discrete approximation

We have seen in the previous section that in order to be able to capture the homogenized solution $\bar{u}$ of (3.1.3) by the discrete approximation $u_\mathbf{h}^\varepsilon$ of (3.2.3) with $\varepsilon/|\mathbf{h}| =$ fixed as $\varepsilon \to 0$, it is not enough that the discrete translation group τ_x, $x \in \Lambda_\mathbf{r}$ be ergodic for P. However we have the following result.

THEOREM 2: *Given $\delta > 0$ there exist $\varepsilon_0 > 0$ and $r_0 > 0$ such that if $\varepsilon \leq \varepsilon_0$ and $|\mathbf{r}| \leq r_0$ and $\mathbf{r}$ is such that τ_x, $x \in \Lambda_\mathbf{r}$ is ergodic for P then the solution $u_{\varepsilon\mathbf{r}}^\varepsilon(x, \omega)$ of the finite difference problem* (3.2.3) *and the solution $\bar{u}(x)$ of the homogenized problem* (3.1.3) *satisfy.*

$$\sum_{x \in \Lambda_{\varepsilon\mathbf{r}}^O} (\varepsilon |\mathbf{r}|)^d \int_\Omega |u_{\varepsilon\mathbf{r}}^\varepsilon(x, \omega) - \bar{u}(x)|^2\, P(d\omega) < \delta\,. \tag{3.3.1}$$

Proof :

Given the convergence result (3.2.7), the only thing that has to be shown is that $a_{ij}^*(\mathbf{r})$ defined by (3.2.9)-(3.2.13) converges to a_{ij}^* of (3.1.11)-(3.1.15) as $|\mathbf{r}| \to 0$. This however follows immediately from the fact that if $\tilde{g}$ is in the domain of the operator D_i defined by (3.1.10), then the abstract difference operators defined by (3.2.8) converge to the generators

$$D_i^{\pm \mathbf{h}} \tilde{g} \to D_i \tilde{g}$$

strongly as $|\mathbf{h}| \to 0$. That is, (3.2.9)-(3.2.13) is a consistent, strong approximation of (3.1.11)-(3.1.15) as $|\mathbf{r}| \to 0$. This completes the proof.

Let us compare Theorem 2 with the one-dimensional Theorem 1. The main difference is that $\mathbf{r}$, which is now a scalar r, need not be small in one dimension because consistency is automatic for any r. In the multidimensional case the fact that $|\mathbf{r}|$ must be small is a definite but unavoidable restriction in general. We show below by an explicit calculation why $|\mathbf{r}|$ small is unavoidable. In the multidimensional *periodic* case, however, the restriction to $|\mathbf{r}|$ small is superficial because if we write

$$\mathbf{r} = [\mathbf{r}] + \boldsymbol{\rho} \tag{3.3.2}$$

where $[\mathbf{r}]$ is the integer part of $\mathbf{r}$ and $\boldsymbol{\rho}$ the residual, then it is enough that $|\boldsymbol{\rho}| \to 0$ through a sequence that has positive, irrational components. Therefore in the multidimensional periodic case we recover Theorem 1 as follows.

THEOREM 3 : *Assume that the coefficients* $\tilde{a}_{ij}(x)$ *are periodic so that* $a_{ij}(x, \omega) = \tilde{a}_{ij}(x + \omega)$ *with* ω *in* T^d, $d \geq 1$. *Then given* $\delta > 0$, *there exists an* $\varepsilon_0 > 0$ *and a* $\rho_0 > 0$ *such that if* $\varepsilon \leq \varepsilon_0$ *and* $|\boldsymbol{\rho}| \leq \rho_0$ *with* $\boldsymbol{\rho}$ *having irrational components, then the solution of* $u_{\varepsilon\mathbf{r}}^{\varepsilon}$ *of the finite difference problem* (3.2.3) *with* $\mathbf{r} = \mathbf{n} + \boldsymbol{\rho}$, $\mathbf{n}$ *a fixed vector with positive integer components, and the solution* $\bar{u}(x)$ *of the homogenized problem* (3.1.3) *satisfy*

$$\sum_{x \in \Lambda_{\varepsilon\mathbf{r}}^{O}} (\varepsilon |\mathbf{r}|)^d \int_{T^d} |u_{\varepsilon\mathbf{r}}^{\varepsilon}(x, \omega) - \bar{u}(x)|^2 \, d\omega < \delta \,. \tag{3.3.2}$$

The extension of Theorem 1*a* to several dimensions is much more difficult because it involves estimates of the allowable set of scale factors. We do not have any such results at present.

We will now show that the restriction to $|\mathbf{r}|$ small in Theorem 2 is necessary in general. We have to produce an example where the discrete homogenized coefficients $a_{ij}^*(\mathbf{r})$, defined by (3.2.9)-(3.2.13), differ from the continuous homogenized coefficients a_{ij}^* defined by (3.1.11)-(3.1.15) when $|\mathbf{r}| > 0$.

Assume that the coefficients $a_{ij}(x, \omega) = \tilde{a}_{ij}(\tau_{-x}\omega)$ have the form

$$\tilde{a}_{ij}(\omega) = \delta_{ij}[1 + \zeta\tilde{b}(\omega)] \tag{3.3.3}$$

where $\tilde{b}(\omega)$ is bounded and has mean zero and ζ is a small real parameter. The homogenized coefficients a_{ij}^* will now depend on ζ and it is easy to see [21] that they are analytic functions of ζ near the origin. Similarly, the discrete homogenized coefficients $a_{ij}^*(\mathbf{r})$ depend on ζ and they are also analytic near the origin. Their Taylor expansions can be computed easily. We will compute the expansion for the discrete case since the continuous case is similar and is also given in [21].

Let the discrete, abstract gradient $D^{\mathbf{h}}$ be defined by (3.2.8) and let

$$\Delta^{\mathbf{h}} = \sum_{j=1}^{d} D_j^{-\mathbf{h}} D_j^{+\mathbf{h}} \tag{3.3.4}$$

be the discrete, abstract Laplace operator. We also define the discrete projection operators

$$\Gamma_{pq}^{\mathbf{h}} = D_p^{+\mathbf{h}}(-\Delta^{\mathbf{h}})^{-1} D_q^{-\mathbf{h}}. \tag{3.3.5}$$

We can now rewrite problem (3.2.9)-(3.2.12) that determines the discrete homogenized coefficients in the following form

$$\tilde{E}_i = \bar{E}_i + \zeta \sum_{j=1}^{d} \Gamma_{ij}^{\mathbf{h}}(\tilde{b}\tilde{E}_j) \tag{3.3.6}$$

which is an integral equation version of (3.2.9)-(3.2.12).

To see the equivalence of (3.3.6) with (3.2.9)-(3.2.12) note first that (3.2.12) is immediate because the operator Γ projects to functions with zero mean. The abstract zero-curl condition (3.2.10) is also immediate because the field $\bar{\mathbf{E}}$ is constant and the operator Γ has a gradient as its last action, as can be seen from (3.3.5), and the curl of a gradient is zero. To verify the divergence-zero condition (3.2.11) we write (3.3.6) in vector form

$$\tilde{\mathbf{E}} = \bar{\mathbf{E}} + \Gamma(\zeta\tilde{\mathbf{b}}\tilde{\mathbf{E}})$$

multiply both sides by $1 + \zeta\tilde{\mathbf{b}}$ and take divergence

$$D^{\mathbf{h}}.\ [(1 + \zeta\tilde{b})\,\tilde{\mathbf{E}}] = D^{\mathbf{h}}.\ [(1 + \zeta\tilde{b})\,\bar{\mathbf{E}}] + D^{\mathbf{h}}.\ [(1 + \zeta\tilde{b})\,\Gamma(\zeta\tilde{b}\tilde{\mathbf{E}})].$$

— From (3.3.5) it follows that $D^{\mathbf{h}}.\Gamma = -D^{\mathbf{h}}$ and since $\bar{\mathbf{E}}$ is constant we have

$$D^{\mathbf{h}}.\ [(1 + \zeta\tilde{b})\,\tilde{\mathbf{E}}] = D^{\mathbf{h}}.\ [\zeta\tilde{b}(\bar{\mathbf{E}} - \tilde{\mathbf{E}} + \Gamma(\zeta\tilde{b}\tilde{\mathbf{E}}))].$$

The right side is now zero by the definition of $\tilde{\mathbf{E}}$, so the divergence-zero property (3.2.11) is proved.

— From (3.3.6) we get the Taylor expansion of $\tilde{\mathbf{E}}$ which we write in vector form

$$\tilde{\mathbf{E}} = \bar{\mathbf{E}} + \zeta\Gamma(\tilde{b}\bar{\mathbf{E}}) + \zeta^2\,\Gamma(\tilde{b}\Gamma(\tilde{b}\bar{\mathbf{E}})) + \cdots \tag{3.3.7}$$

Inserting this expression into the definition (3.2.13) of the discrete homogenized coefficients we have

$$a^*(\mathbf{r})\bar{\mathbf{E}} = \int_\Omega (1+\zeta\tilde{b})\,[\bar{\mathbf{E}} + \zeta\Gamma(\tilde{b}\bar{\mathbf{E}}) + \zeta^2\,\Gamma(\tilde{b}\Gamma(\tilde{b}\bar{\mathbf{E}})) + \cdots\,]\,P(d\omega)$$

$$= \bar{\mathbf{E}} + \zeta\int_\Omega \tilde{b}P(d\omega)\,\bar{\mathbf{E}} + \zeta^2\int_\Omega \tilde{b}\Gamma(\tilde{b}\bar{\mathbf{E}})\,P(d\omega) + \cdots$$

Since $\tilde{b}$ has mean zero and $\bar{\mathbf{E}}$ is arbitrary we get the following expansion for the discrete homogenized coefficients

$$a^*_{ij}(\mathbf{r}) = \delta_{ij} + \zeta^2\int_\Omega \tilde{b}(\omega)\,\Gamma^{\mathbf{r}}_{ij}\,\tilde{b}(\omega)\,P(d\omega) + \cdots \tag{3.3.8}$$

The translation operators T_x, $x \in \mathbf{R}^d$, which are defined by (3.1.6), form a unitary group and have therefore the spectral representation

$$T_x = \int_{\mathbf{R}^d} e^{ik.x}\,G(dk) \tag{3.3.9}$$

where $G(A)$, $A \subset \mathbf{R}^d$ is a projection-valued measure on H. From the definition (3.2.8) of the discrete, abstract gradient operators and (3.3.4), (3.3.5) we get the spectral representation of $\Gamma^{\mathbf{h}}_{pq}$

$$\Gamma^{\mathbf{h}}_{pq} = -\int_{\mathbf{R}^d} \frac{(e^{-ih_p k_p}-1)\,(e^{ih_q k_q}-1)}{\sum_{j=1}^{d} |e^{ih_j k_j}-1|^2}\,G(dk) \tag{3.3.10}$$

where $h = (h_1, ..., h_d)$ and $k = (k_1, ..., k_d)$. Thus the expansion (3.3.8) has the form

$$a^*_{pq}(\mathbf{r}) = \delta_{pq} - \zeta^2\int_{\mathbf{R}^d} \frac{(e^{-ir_p k_p}-1)\,(e^{ir_q k_q}-1)}{\sum_{j=1}^{d} |e^{ir_j k_j}-1|^2}\,\hat{R}_b(dk) + \cdots \tag{3.3.11}$$

where $\hat{R}_b(dk)$ is the spectral measure of the stationary process $b(x, \omega) = \tilde{b}(\tau_{-x}\omega)$. That is, if

$$\int_{\Omega} b(x, \omega)\, b(0, \omega)\, P(d\omega) = R_b(x) \tag{3.3.12}$$

is the covariance of $b(x, \omega)$, then

$$R_b(x) = \int_{\mathbf{R}^d} e^{ik \cdot x}\, \hat{R}_b(dk)\,. \tag{3.3.13}$$

By a calculation similar to the above it is easy to see that the continuous homogenized coefficients have the Taylor expansion

$$a_{pq}^* = \delta_{pq} - \zeta^2 \int_{\mathbf{R}^d} \frac{k_p\, k_q}{|k|^2}\, \hat{R}_b(dk) + \cdots \tag{3.3.14}$$

which is consistent with (3.3.11) in the limit $|\mathbf{r}| \to 0$, as should be.

Suppose the spectral measure of b ig genuinely multidimensional, i.e. that $\hat{R}_b(dk)$ is concentrated on more than one direction in k space. For example, suppose that $\hat{R}_b$ has a continuous density with respect to Lebesgue measure. Then the coefficients of ζ^2 in (3.3.11) and (3.3.14) are different for any $\mathbf{r}$ with $|\mathbf{r}| > 0$. *This is the difference between the one-dimensional and the multidimensional cases.* In the multidimensional case, if there was a set of mesh widths $\mathbf{r}$ with $|\mathbf{r}|$ finite, but not necessarily close to zero, for which $a_{pq}^*(\mathbf{r}) \sim a_{pq}^*$, then we would have a successful capture of the homogenized problem by the numerical scheme without resolving the oscillations. But when we consider $a_{pq}^*(\mathbf{r})$ and a_{pq}^* for small ζ, we see from (3.3.11) and (3.3.14) that there is no set of mesh widths $\mathbf{r}$ that will do unless the spectral measure $\hat{R}_b(dk)$ is concentrated on a periodic lattice, the periodic case of Theorem 3, and $\mathbf{r}$ has irrational residual components. In one space dimension the expansions for the discrete and continuous case are the same and they agree with the expansion of (2.5) when the form (3.3.3) of the coefficients is used.

ACKNOWLEDGEMENT

We thank B. Engquist for several discussions on numerical homogenization. The work of M. Avellaneda was supported by grants : NSF-DMS 8802739 and ARO-DAAL 03 89 K 0039. The work of T. Hou was supported by the grant AFOSR 86-0352 and the work of G. Papanicolaou by grants : AFOSR 86-0352 and NSF DMS 8701895.

REFERENCES

[1] E. DEGIORGI and S. SPAGNOLO, Sulla convergenza degli integrali dell'energia per operatori ellitici del secundo ordine, Boll. I.M.I., **8**, 391-411, 1973.

[2] I. BABUSKA, Solution of interface problems by homogenization I, II, III, SIAM J. Math. Analysis, **7**, 603-345, 635-645, and **8**, 923-937, 1976.

[3] L. TARTAR, Cours Peccot, Paris 1977.

[4] A. BENSOUSSAN, J. L. LIONS and G. PAPANICOLAOU, Asymptotic Analysis for Periodic Structures, North Holland, Amsterdam, 1978.

[5] E. SANCHEZ-PALENCIA, Nonhomogeneous Media and Vibration Theory, Springer, 1980.

[6] G. PAPANICOLAOU and S.R.S. VARADHAN, Boundary value problems with rapidly oscillating random coefficients, at Proceedings of Conference on Random Fields, Esztergom, Hungary, 1979, published in Seria Colloquia Mathematica Societatis Janos Bolyai, **27**, 835-873, North Holland, 1981.

[7] V. V. ZHIKOV, S. M. KOZLOV, O. OLEINIK and KHA T'EN NGOAN, AVERAGING and G-convergence of differential operators, Uspekhi Mat. Nauk. **34**, 65-133, 1979.

[8] N. BAHVALOV and G. PANACHENKO, Average of processes in periodic media, Nauka, Moscow, 1984.

[9] S. KOZLOV, The method of averaging and walks in inhomogeneous environments, Uspekhi Mat. Nauk **40**, 61-120, 1985.

[10] G. PAPANICOLAOU, Macroscopic properties of composites, bubbly fluids, suspensions and related problems, in Les méthodes de l'homogénéization : Théorie et applications en physique, CEA-EDF-INRIA séries **57**, 229-318, Eyrolles, Paris, 1985.

[11] J. L. LIONS, Remarques sur les problèmes d'homogénéization dans les milieux à structure périodique et sur quelques problèmes raides, in Les méthodes de l'homogénéization : Théorie et applications en physique, CEA-EDF-INRIA séries **57**, 129-228, Eyrolles, Paris, 1985.

[12] F. MURAT and L. TARTAR, Calcul des variations et homogénéization, in Les méthodes de l'homogénéization : Théorie et applications en physique, CEA-EDF-INRIA séries **57**, 319-370, Eyrolles, Paris, 1985.

[13] B. ENGQUIST, Computation of oscillatory solutions to hyperbolic differential equations, to appear. Also in Lecture Notes in Math. **1270**, 10-22, Springer, 1987.

[14] B. ENGQUIST and T. HOU, Particle method approximation of oscillatory solutions to hyperbolic differential equations, SIAM J. Numer. Anal. **26**, 289-319, 1989.

[15] B. ENGQUIST and T. HOU, Computation of oscillatory solutions to hyperbolic differential equations using particle methods, Lecture Notes in Math., **1360**, 68-82, 1988.

[16] L. TARTAR, Solutions oscillantes des équations de Carleman, Séminaire Goulaouic-Meyer-Schwartz, 1981.

[17] D. MCLAUGHLIN, G. PAPANICOLAOU and L. TARTAR, Weak limits of semilinear hyperbolic equations with oscillating data, Springer Lecture Notes in Phys., **230**, 277-289, 1985.

[18] R. KUHNEMANN, The diffusion limit for reversible jump processes on $\mathbf{Z}^d$ with ergodic random bond conductivities, Comm. Math. Phys., **90**, 27-68, 1983.

[19] R. FIGARI, E. ORLANDI and G. PAPANICOLAOU, Diffusive behavior of a random walk in a random medium, in « Stochastic Analysis » by K. Ito, 105-120, North Holland, 1984.

[20] V. ARNOLD and A. AVEZ, Ergodic Problems of Classical Mechanics, Benjamin, New York, 1968.

[21] K. GOLDEN and G. PAPANICOLAOU, Bounds for effective parameters of heterogeneous media by analytic continuation, Comm. Math. Phys., **90**, 473-491, 1983.