HOMOGENIZATION THEORY FOR PARTIAL DIFFERENTIAL EQUATIONS

G.A. Pavliotis
Department of Mathematics
South Kensington Campus
Imperial College London
London SW7 2AZ
UK

October 14, 2004
# Contents

1 Introduction
   1.1 Introduction ......................................................... 1
   1.2 An Example: Steady Heat Conduction in a Composite Material ....... 3
   1.3 Some Remarks on the Literature on Homogenization ................. 7

2 The Method of Multiple Scales: Formal Asymptotics .................. 9
   2.1 Introduction–Setting of The Problem .................................. 9
   2.2 The Multiple Scales Expansion ......................................... 10
   2.3 Some Exactly Solvable Cases ........................................... 17
   2.4 Properties of the Homogenized Equation ............................. 20
   2.5 Remarks on the Method of Multiple Scales–Extensions ............ 23
      2.5.1 Higher Order Correctors .......................................... 23
      2.5.2 Different Boundary Conditions .................................... 24
      2.5.3 Locally Periodic Coefficients .................................... 25
      2.5.4 Time Dependent Problems ........................................ 26

3 Elliptic Boundary Value Problems ....................................... 31
   3.1 Introduction .......................................................... 31
   3.2 Background Material .................................................. 32
   3.3 Dirichlet Boundary Conditions ........................................ 37
   3.4 Periodic Boundary Conditions ........................................ 41
   3.5 The Fredholm Alternative for Second Order Uniformly Elliptic Operators in Divergence Form ................................. 44

4 The Homogenization Theorem .............................................. 55
   4.1 Introduction .......................................................... 55
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.2</td>
<td>The Homogenization Theorem I: Tartar’s Method of Oscillating Test Functions</td>
<td>57</td>
</tr>
<tr>
<td>4.3</td>
<td>The Homogenization Theorem II: The Method of Multiple Scales</td>
<td>60</td>
</tr>
<tr>
<td>5</td>
<td>Two–Scale Convergence</td>
<td>67</td>
</tr>
<tr>
<td>5.1</td>
<td>Introduction</td>
<td>67</td>
</tr>
<tr>
<td>5.2</td>
<td>Two–Scale Convergence</td>
<td>68</td>
</tr>
<tr>
<td>5.3</td>
<td>The Homogenization Theorem through Two–Scale Convergence</td>
<td>76</td>
</tr>
<tr>
<td>5.4</td>
<td>Two Examples of Non Linear Problems</td>
<td>83</td>
</tr>
<tr>
<td>5.4.1</td>
<td>Introduction</td>
<td>83</td>
</tr>
<tr>
<td>5.4.2</td>
<td>Homogenization for Convex Energy Functionals</td>
<td>83</td>
</tr>
<tr>
<td>5.4.3</td>
<td>Homogenization for Monotone Operators</td>
<td>83</td>
</tr>
<tr>
<td>6</td>
<td>Homogenization for Parabolic PDE</td>
<td>89</td>
</tr>
<tr>
<td>6.1</td>
<td>Introduction</td>
<td>89</td>
</tr>
<tr>
<td>6.2</td>
<td>Homogenization for Linear Transport Equations</td>
<td>91</td>
</tr>
<tr>
<td>6.3</td>
<td>Homogenization for Advection–Diffusion Equations</td>
<td>91</td>
</tr>
<tr>
<td>6.4</td>
<td>Homogenization for Parabolic Equations with Time Dependent Coefficients</td>
<td>91</td>
</tr>
<tr>
<td>7</td>
<td>Periodic Homogenization for Parabolic Equations: A Probabilistic Approach</td>
<td>93</td>
</tr>
<tr>
<td>7.1</td>
<td>Introduction</td>
<td>93</td>
</tr>
<tr>
<td>7.2</td>
<td>Background Material on Diffusion Processes</td>
<td>95</td>
</tr>
<tr>
<td>7.3</td>
<td>The Martingale Central Limit Theorem</td>
<td>95</td>
</tr>
<tr>
<td>7.3.1</td>
<td>The Central Limit Theorem for Martingales</td>
<td>96</td>
</tr>
<tr>
<td>7.4</td>
<td>The Homogenization Theorem for Advection–Diffusion Equations Revisited</td>
<td>100</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

1.1 Introduction

Homogenization theory is concerned with the analysis of Partial Differential Equations (PDE) with rapidly oscillating coefficients. The basic problem is this: we are given a one parameter family of partial differential operators $A^\varepsilon$ and we consider the equation

$$A^\varepsilon u^\varepsilon = f,$$

(1.1)

together with the appropriate boundary and/or initial conditions. Here $\varepsilon$ is a small parameter, $\varepsilon \ll 1$, associated with the oscillations—indeed, it measures the period of oscillations. We are interested in studying the solution of (1.1) in the limit as $\varepsilon \to 0$. In particular, we would like to understand the following issues:

- **Convergence to a limit.** Is there a limit $u$ of $u^\varepsilon$, as $\varepsilon \to 0$? In which sense should we understand the convergence (i.e., in which norm, which topology etc.)? What is the convergence rate?

- **Characterization of the limiting process.** What kind of equation does the limit $u$ satisfy? Suppose that the limiting equation is of the following form:

$$Au = f$$

(1.2)

Is equation (1.2) of the same type as the original equation (1.1)? In other words, is the operator $A$ of the same type as $A^\varepsilon$, i.e. is it a partial differential operator? We will call the coefficients of the homogenized operator $A$ the effective coefficients or the effective parameters.
Explicit analytical construction of $A$. How can we compute the effective coefficients?

Properties of the limiting equation. How do the properties of the solution $u$ of the limiting equation compare with those of $u^\varepsilon$? How do the effective coefficients depend on the coefficients of $A^\varepsilon$?

We remark that from an applied point of view another issue of fundamental importance is the efficient computation of the effective coefficients. Indeed, the reason why the method of homogenization is of practical interest is because the calculation of the effective coefficients, together with the solution of the limiting equation is (computationally) much easier than the solution of the original problem (1.1).

In these notes we will try to address the above issues—in particular the existence and characterization of the limit—for various types of partial differential equations under appropriate assumptions on the type of oscillations.

There are various problems in physics and engineering that can be adequately placed within the framework that we will develop in these notes. As examples we mention composite materials [5], flow in porous media [27], atmospheric turbulence [13]. A common feature of all these problems is that phenomena occur at various length and time scales. This results in the PDE which describe the physical phenomenon under investigation to be very complicated and hard to analyze. Let us for simplicity consider the case when there are only two characteristic length scales, a microscopic one $\ell$ and a macroscopic one $L$. It quite often the case that there is a clear scale separation between the microscale and the macroscale. Mathematically this implies that we can introduce a small parameter $\varepsilon := \frac{\ell}{L} \ll 1$ which controls the scale separation. Now, the PDE which model the aforementioned problems are of the form (1.1), the parameter $\varepsilon$ controlling the scale separation. In this context, the idea is that a homogenized equation of the form (1.2) can adequately describe the physical phenomenon that we are interested in analyzing when viewed from the macroscale, without having to resolve explicitly the microscale. Thus, the limit as $\varepsilon \to 0$ corresponds to the case of infinite scale separation. This is indeed the basic goal of homogenization from a physical point of view: determine the macroscopic behavior of a physical system induced by a given microstructure.

In this course we will be mostly concerned with problems with a periodic structure (for example, with periodic composite materials). The periodicity assumption implies that the coefficients

---

1. The method of homogenization has found recently applications in finance, see for instance the book [21]
2. Throughout these notes we will be dealing with equations that have already been non-dimensionalized: the characteristic length and times are in fact non-dimensional quantities.
1.2. AN EXAMPLE: STEADY HEAT CONDUCTION IN A COMPOSITE MATERIAL

There are various reasons why we will make this assumption: first, the periodicity assumption is quite realistic in various cases, such as composite materials or diffusion in porous media. Moreover, as we will see later, this assumption will enable us to compute the homogenized coefficients explicitly. Finally, the periodicity assumption makes the mathematical analysis of the problem much simpler than that of the random or deterministic non–periodic problem.

Let us now assume, like we did before, that the problem we are studying has only two characteristic length scales and no temporal dependence. Let us fix the macroscopic length scale $L$ to be an $O(1)$ quantity. Now, we can associate the microscopic scale $\ell$ with the period of oscillations $\epsilon$. Thus, the coefficients of the PDE under investigation are $\epsilon$–periodic functions of the spatial variable $x$. If $a^\epsilon(x)$ denotes a generic coefficient of our PDE, then we can introduce the periodicity into account by writing

$$a^\epsilon(x) = a \left( \frac{x}{\epsilon} \right).$$

Here $a(y)$ is a given $1$–periodic function. In this manner, we have introduced a reference cell $Y$ with period 1. This the unit cell where the reference heterogeneities are given and is of fundamental importance in the theory of periodic homogenization for PDE. To understand better the issues that arise in periodic homogenization and the type of results that we obtain let us discuss an example in detail.

1.2 An Example: Steady Heat Conduction in a Composite Material

Now we wish to discuss an example in some detail in order to illustrate the basic ideas and techniques of periodic homogenization as well as the mathematical tools needed. To this end, let us

---

3To be more precise, the computation of the effective coefficients will require the solution of a partial differential equation. However, the point is that periodic homogenization is completely constructive.

4In fact, random homogenization is also constructive and it leads to expressions for the effective coefficients which are very similar to the ones that we get in periodic homogenization. The analysis, however, is much more difficult. See however the remarks in Chapter (7).

5The order symbols $O$ and $o$: Let $f(x)$, $g(x)$ be smooth functions and let $\epsilon \in \mathbb{R}$. We will say that $f(\epsilon)$ is of order $g(\epsilon)$ as $\epsilon \to 0$ and we will write $f(\epsilon) = O(g(\epsilon))$ if $\lim_{\epsilon \to 0} \frac{f(\epsilon)}{g(\epsilon)} = L$, $|L| < \infty$. Moreover, the notation $f(\epsilon) = o(g(\epsilon))$ is used when $\lim_{\epsilon \to 0} \frac{f(\epsilon)}{g(\epsilon)} = 0$. We refer e.g. to [26] for details.

6We will use the term $T$–periodic for functions $f(x)$ which are periodic with period $T$: $f(x) = f(x + T)$.

7In these notes we will always assume that we have non–dimensionalized the PDE under investigation in such a way that the reference cell is of period 1. Since this cell can be identified with the unit torus, we will also use the notation $\mathbb{T}^d$. We have chosen to work with the unit cell as opposed to a cell of the form $[0, L_1] \times [0, L_2] \cdots \times [0, L_d] \in \mathbb{R}^d$, for notational simplicity.
consider the problem of steady heat conduction in a periodic composite material. If \( \Omega \in \mathbb{R}^3 \) denotes the domain occupied by the material, then the size of the domain defines the macroscopic length scale \( L \). On the other hand, the period of heterogeneities defines the microscopic length scale \( \ell \) of the problem. Assuming as before that \( L = \mathcal{O}(1) \) and that the size of heterogeneities is small \( \ell = \epsilon \ll 1 \), then the phenomenon of steady heat conduction is described by the following elliptic boundary value problem:

\[
-\frac{\partial}{\partial x_i} \left( a_{ij} \left( \frac{x}{\epsilon} \right) \frac{\partial u^\epsilon}{\partial x_j} \right) = f, \quad \text{for } x \in \Omega \tag{1.4a}
\]

\[
u^\epsilon(x) = 0, \quad \text{for } x \in \partial\Omega. \tag{1.4b}
\]

In writing the above equation we have used the summation convention, i.e. repeated indices imply summation. The matrix \( A(y) = \{a_{ij}(y)\}_{i,j=1}^3 \) is the \textit{thermal conductivity tensor} and \( u^\epsilon(x) \) denotes the temperature field. Notice that equation (1.4) is of the form (1.1) with \( A^\epsilon := -\frac{\partial}{\partial x_i} \left( a_{ij} \left( \frac{x}{\epsilon} \right) \frac{\partial}{\partial x_j} \right) \), together with the Dirichlet boundary conditions. The purpose of homogenization theory is to study the limit of this equation as \( \epsilon \to 0 \).

In later chapters we will show that, under appropriate assumptions on the coefficients \( \{a_{ij}\}_{i,j=1}^3 \), the function \( f(x) \) and the domain \( \Omega \) the homogenized equation is

\[
-\overline{a}_{ij} \frac{\partial u}{\partial x_i \partial x_j} = f, \quad \text{for } x \in \Omega \tag{1.5a}
\]

\[
u(x) = 0, \quad \text{for } x \in \partial\Omega. \tag{1.5b}
\]

The constant homogenized coefficients \( \{\overline{a}_{ij}\}_{i,j=1}^3 \) are given by the formula:

\[
\overline{a}_{ij} = \int_Y \left( a_{ij}(y) - a_{ik} \frac{\partial x^j}{\partial y_k} \right) dy, \quad i,j = 1, \ldots, d. \tag{1.6}
\]

The \textit{(first order) corrector} \( \{\chi^k(y)\}_{k=1}^d \) solves the \textit{cell problem}

\[
-\frac{\partial}{\partial y_i} \left( a_{ij}(y) \frac{\partial \chi^k(y)}{\partial y_j} \right) = -\frac{\partial a_{ik}(y)}{\partial y_i}, \quad \chi^k(y) \text{ is } 1\text{-periodic, } k = 1, \ldots, d. \tag{1.7}
\]

Thus, the calculation of the effective coefficients involves the solution of a partial differential equation posed on the unit cell (i.e. with periodic boundary conditions), together with the computation of the integrals in (1.6). We will prove that the solution \( u^\epsilon(x) \) of (1.4) converges to the solution

\[\text{from a physical point of view, this limit corresponds to the case where the heterogeneities become smaller and smaller. In other words, we want to replace the original, highly heterogeneous material characterized by the coefficients } A \left( \frac{x}{\epsilon} \right) \text{ with an effective, homogeneous which is characterized by the constant coefficients } \overline{A}. \text{ Hence the name homogenization}\]
$u(x)$ of the homogenized equation (1.5) as $\epsilon \to 0$ in the appropriate function space$^9$ and in the appropriate topology. Moreover, we will be able to estimate the difference between $u^\epsilon$ and $u$: we will prove that $\|u^\epsilon - u\| \leq C \epsilon^\alpha$ where $\| \cdot \|$ denotes the norm of the appropriate (Hilbert) space, $C$ is a constant independent of $\epsilon$ and the exponent $\alpha > 0$ will be determined from the homogenization procedure.

In these lecture notes we will study various methods which enable us to obtain homogenization results of this form for various types of (mostly linear) PDE with (primarily) periodic coefficients. The most widely used method for studying homogenization problems in mechanics and physics is that of the multiple scales. The idea behind this method is look for solutions of equation (1.4), or more generally (1.1), in the form

$$u^\epsilon(x) = u_0 \left( x, \frac{x}{\epsilon} \right) + \epsilon u_1 \left( x, \frac{x}{\epsilon} \right) + \epsilon^2 u_2 \left( x, \frac{x}{\epsilon} \right) + \ldots,$$

where $u_j = u_j(x, y), j = 1, 2, \ldots$ are $1$–periodic in $y$. That is, we make the–physically reasonable–assumption that the solutions depends explicitly upon both scales which appear in the problem. Upon substituting (1.8) into (1.1), equating power–like terms in $\epsilon$ and applying repeatedly the Fredholm alternative$^{10}$, we obtain the homogenized equation for $u_0(x)$ together with the cell problem. We can also compute higher order terms, if we wish. The validity of the ansatz (1.8) can be justified a posteriori by computing the difference $R^\epsilon$ between $u^\epsilon$ and $u$ and then using energy estimates or the maximum principle to prove that it is small.

The method of multiple scales is very useful in providing us with the right answer, it is not very appropriate however when trying to prove homogenization theorems$^{11}$. The main problem is that its rigorous justification requires a lot of smoothness from the solution of our PDE (1.4). For most problems that arise in applications the coefficients of the PDE that we wish to analyze are not smooth functions and consequently the solution $u^\epsilon$ is not smooth, either. Therefore, in order to develop a rigorous basis for the method of homogenization for PDE with non–smooth coefficients, we first need to define an appropriate concept of non–smooth solution, that of a weak solution. This is accomplished by studying our PDE in an averaged sense, after integrating against a smooth function. Our original PDE (1.4) is transformed then into an equation of the form

$$a[u^\epsilon, \phi] = f(\phi), \quad \text{for all smooth } \phi,$$

$^9$We will be dealing almost exclusively with Hilbert spaces in these notes.

$^{10}$The Fredholm alternative will be analyzed in section 3.5.

$^{11}$See, however the analysis presented in section 4.3
where \( a[\cdot, \cdot] \) is an appropriate bilinear form and \( f(\cdot) \) is an appropriate linear functional. The function \( v \) are called test functions. Once our PDE has been recast in the abstract form (1.9) and once an appropriate existence and uniqueness theory of solutions has been developed, a variety of tool from functional analysis and modern PDE theory are at disposal. The general setting for proving homogenization theorems is this: first, we obtain energy estimates within the appropriate function spaces by replacing \( v \) by appropriate functionals of \( u^\varepsilon \); this enables us to conclude that \( u^\varepsilon \) is uniformly bounded in some appropriate function space which in turn implies that it converges to some function \( u \). second, we pass to the limit as \( \varepsilon \) tends to 0 in (1.9) and characterize \( u \) by making appropriate choices of test functions. In this setting, the problem of proving the homogenization theorem reduces to the construction of the right test functions. Let us now review some of the methods for constructing appropriate test functions within the framework of periodic homogenization.

A very powerful method is L. Tartar’s method of oscillating test functions. The idea of this method is to construct appropriate test functions by using the solution of the cell problem, or rather, of the adjoint cell problem. This method has been developed precisely within the framework of periodic homogenization and will be presented in section 4.2.

Another approach to this problem is through the perturbed test function method. The idea here is to use test functions in the form of a multiple scales expansion:

\[
\phi^\varepsilon \approx \phi_0 + \varepsilon \phi_1 + \varepsilon^2 \phi_2 + \ldots
\]

This method is in fact quite general and can be used in a variety of contexts, [31, 18, 17, 44, 43]. A variant of this approach which is tailored to periodic problems is that of two–scale convergence [3, 2, 38, 39]. In addition to considering test function of the form (1.10), an additional ingredient of this method is based on the introduction of a new type of convergence, the two–scale convergence. This method will be introduced and analyzed in Chapter 5, and then used to study the homogenization problem for a variety of PDE. We mention that a similar methodology is employed in [18, 17] in order to study homogenization for nonlinear PDE within the framework of viscosity solutions [19, ch. 10].

Various PDE, in particular linear elliptic and parabolic PDE, admit a probabilistic interpretation: their solutions can be expressed as expectation values of functions of continuous time
stochastic processes [22] [40, ch. 8]. For this type of PDE homogenization theorems can be obtained through analysis of the stochastic differential equations associated with the PDE [40]. From this point of view, the proof of the homogenization theorem is essentially a form of the central limit theorem from probability theory. For details on this approach we refer to e.g. [7, ch. 3], [46, 6, 8]. The probabilistic approach to homogenization is also related to averaging problems for stochastic differential equations [31].

Most of the methods mentioned so far are tailored to the analysis of periodic homogenization, since in one way or another they make explicit use of the cell problem and its solution. These methods break down when dealing with non–periodic homogenization and more abstract methods have to be developed, since explicit formulas for the homogenized coefficients of the form (1.5) cannot be obtained. Various notions of convergence have been developed which are more suitable for the non–periodic case. Concepts such as that of $G$–convergence (for symmetric non–periodic problems), and of $H$–convergence (for non–symmetric non–periodic problems) have been introduced. Moreover, various types of PDE–in particular elliptic equations–admit a variational characterization [19, ch. 8]. A general mathematical framework for studying convergence problems for this type of PDE has been developed under the name of $\Gamma$–convergence [36]. This type of convergence is concerned with the convergence of minimizers of (energy) functionals.

As we have already mentioned, no constructive way of computing the effective coefficients is available in non–periodic homogenization. For these problems the best one can hope for is the derivation of (hopefully optimal) bounds on the homogenized coefficients in terms of the coefficients of the original PDE. This is a very important problem from an applied perspective and the literature on this issue is vast.

1.3 Some Remarks on the Literature on Homogenization

Questions related to homogenization have been investigated since the early nineteenth century\textsuperscript{15}. However, the mathematical theory of homogenization was initiated from the pioneering works of Spagnolo on $G$–convergence in the late sixties. Since then, the literature of homogenization has been enormous. Nowadays, there are various excellent monographs, textbooks and lecture notes which cover material related to homogenization. An excellent monograph, which covers a great

\textsuperscript{14}There are various extensions of all of the methods mentioned above which are appropriate for the study of random homogenization, i.e. for the homogenization of PDE with random coefficients. We will not be concerned with problems of this form in these lecture notes. We refer to e.g. [45, 11, 9] for various approaches to random homogenization.

\textsuperscript{15}For example, in connection to the study of effective coefficients of composite materials.
deal of material on periodic homogenization is the book by Bensoussan, Lions and Papanicolaou [7]. We recommend this book for additional material, alternative approaches and the analysis of various problems that will not be covered in these notes (for example, high frequency wave propagation). An excellent recent textbook, which is also strongly recommended, is that of Cioranescu and Donato [12]. We also mention the encyclopedic monograph by Jikov, Kozlov and Oleinik [28]. These are the main references upon which these lecture notes are based. Another classic monograph on homogenization is the book by Sachez–Palenci [50]. Our presentation of two–scale convergence is also influenced by [3] and [25]. A very readable presentation of non–periodic homogenization is contained in [14]. For additional material on applications of periodic homogenization to continuum mechanics (composite materials, flow in porous media etc.) we refer to the books [5, 27]. We also mention the lecture notes [41] which deal with various aspects of homogenization for elliptic and parabolic equations.

The mathematical tools that we will use in these lecture notes are, mostly, linear functional analysis and linear PDE theory. There are various excellent textbooks devoted to this subjects. The material in [19, 48] is more than adequate for our purposes. Chapters 1–4 in [12] contain useful background material. Other useful references, in particular in connection to Sobolev spaces of periodic functions, are [28, ch. 1] and [51, ch. 2].
Chapter 2

The Method of Multiple Scales: Formal Asymptotics

2.1 Introduction–Setting of The Problem

In this chapter we will use the method of multiple scales to study the problem of homogenization
for the Dirichlet problem (1.4) which describes steady state heat conduction in a composite material

\[-\frac{\partial}{\partial x_i} \left( a_{ij} \left( \frac{x}{\epsilon} \right) \frac{\partial u^\epsilon}{\partial x_j} \right) = f, \quad \text{for } x \in \Omega \tag{2.1a} \]

\[u^\epsilon(x) = 0, \quad \text{for } x \in \partial \Omega. \tag{2.1b} \]

Our goal is to derive the homogenized equation (1.5), together with the cell problem (1.7). We also want to study various properties of the homogenized coefficients. We take \( \Omega \in \mathbb{R}^d \), open, bounded with smooth boundary. We will also assume that the coefficients \( A(y) = \{a_{ij}(y)\}_{i,j=1}^d \) are smooth, 1-periodic and uniformly elliptic. Furthermore, will take the function \( f(x) \) to be smooth

\[a_{ij}(y), f(x) \in C^\infty(\mathbb{R}^d), \quad i, j = 1, \ldots, d. \tag{2.2a} \]

\[a_{ij}(y + \hat{e}_k) = a_{ij}(y), \quad i, j, k = 1, \ldots, d. \tag{2.2b} \]

\[a_{ij}(y)\xi_i\xi_j \geq \alpha |\xi|^2, \quad \alpha > 0, \quad \forall y \in Y \forall \xi \in \mathbb{R}^d. \tag{2.2c} \]

In the above \( \{\hat{e}_k\}_{k=1}^d \) denotes the standard basis in \( \mathbb{R}^d \). Moreover, \( Y = [0, 1]^d \) denotes the (reference) unit cell\(^1\). We will also assume that the function \( f \) is smooth and independent of \( \epsilon \). Notice that we do not assume that the matrix \( A \) is symmetric. Let us also remark that the regularity

\(^1\)The unit cell is actually the \( d \)-dimensional unit torus \( T^d \).
assumptions are too stringent and we make them at this point in order to carry out the formal calculations that follow. We will see in later chapters that we actually need much less regularity. This is an important issue, in particular from in connection to the applications of the method in e.g. continuum mechanics. For example, in the case of an isotropic composite with two constituents with different thermal conductivities\(^2\), it is clear that the coefficient \(a(y)\) will have jumps when passing from one phase to the other and hence the assumption of smoothness is not realistic. We will see later on that all we really need is \(a_{ij}(y) \in L^\infty(\mathbb{R}^d)\).

### 2.2 The Multiple Scales Expansion

Let us now consider smooth functions \(\phi(x, y)\), \(x \in \Omega, y \in \mathbb{R}^d\) which are \(1\)-periodic in \(y\). With functions \(\phi(x, y)\) of this form we associate functions \(\phi(x, \frac{x}{\varepsilon})\). The idea behind the method of multiple scales is to assume that the solution \(u^\varepsilon(x)\) of (2.1) depends explicitly on \(x\) as well as \(\frac{x}{\varepsilon}\). Thus, we assume the following ansatz for \(u^\varepsilon\):

\[
 u^\varepsilon(x) = u_0 \left( x, \frac{x}{\varepsilon} \right) + \varepsilon u_1 \left( x, \frac{x}{\varepsilon} \right) + \varepsilon^2 u_2 \left( x, \frac{x}{\varepsilon} \right) + \ldots, \tag{2.3}
\]

Now, from a physical point of view it is reasonable to expect that the solution of (2.1) is of this form, since there are two different length scales in our problem and the above expansion takes this fact explicitly into account. However, at this point, the above expansion is simply an intelligent guess: its validity will be justified later, using either the maximum principle, energy estimates or the method of two–scale convergence. Notice also that no information concerning the boundary conditions is incorporated into this expansion. We would have used the same expansion for Neumann or mixed boundary conditions. This is a very important observation and we will come back to it later.

The variables \(x\) and \(y = \frac{x}{\varepsilon}\) represent the ”slow” (macroscopic) and ”fast” (microscopic) scales of the problem, respectively\(^3\). For \(\varepsilon \ll 1\) the variable \(y\) changes much more rapidly than \(x\) and we can think of \(x\) as being a constant, when looking at the problem at the microscopic scale. This is where the assumption of scale separation enters: we treat \(x\) and \(y\) as independent variables. Justifying the validity of this assumption as \(\varepsilon \to 0\) is one of the main issues of the mathematical theory of homogenization.

\(^2\)The isotropy implies that we only have one component, as opposed to a matrix; indeed, the coefficient matrix is of the form \(A(y) = a(y)\mathbb{I}\), where \(\mathbb{I}\) is the identity matrix in \(\mathbb{R}^d\).

\(^3\)The terms fast and slow scales would be more appropriate for time dependent problems and in fact this terminology has been introduced for these problems, [26]. However, we will retain this terminology even when referring to spatial and not temporal scales.
The fact that \( y = \frac{z}{\varepsilon} \) implies that the partial derivatives with respect to \( x_j \) become:

\[
\frac{\partial}{\partial x_j} \rightarrow \frac{\partial}{\partial x_j} + \frac{1}{\varepsilon} \frac{\partial}{\partial y_j}, \quad j = 1, \ldots, d.
\]

Using this we can write the differential operator \( \mathcal{A}^\varepsilon := -\frac{\partial}{\partial x_i} \left( a_{ij} \left( \frac{z}{\varepsilon} \right) \frac{\partial}{\partial x_j} \right) \) in the form

\[
\mathcal{A}^\varepsilon = \frac{1}{\varepsilon^2} \mathcal{A}_0 + \frac{1}{\varepsilon} \mathcal{A}_1 + \mathcal{A}_2, \quad (2.5)
\]

where

\[
\mathcal{A}_0 := -\frac{\partial}{\partial y_i} \left( a_{ij}(y) \frac{\partial}{\partial y_j} \right), \quad (2.6a)
\]

\[
\mathcal{A}_1 := -\frac{\partial}{\partial y_i} \left( a_{ij}(y) \frac{\partial}{\partial x_j} \right) - \frac{\partial}{\partial x_i} \left( a_{ij}(y) \frac{\partial}{\partial y_j} \right), \quad (2.6b)
\]

\[
\mathcal{A}_2 := -\frac{\partial}{\partial x_i} \left( a_{ij}(y) \frac{\partial}{\partial x_j} \right), \quad (2.6c)
\]

Now, equation (2.1), on account of (2.5), becomes:

\[
\left( \frac{1}{\varepsilon^2} \mathcal{A}_0 + \frac{1}{\varepsilon} \mathcal{A}_1 + \mathcal{A}_2 \right) u^\varepsilon = f, \quad \text{for } x \in \Omega, \quad (2.7a)
\]

\[
u^\varepsilon(x) = 0, \quad \text{for } x \in \partial \Omega. \quad (2.7b)
\]

We substitute now (2.3) into (2.7) to deduce:

\[
\frac{1}{\varepsilon^2} \mathcal{A}_0 u_0 + \frac{1}{\varepsilon} (\mathcal{A}_0 u_1 + \mathcal{A}_1 u_0) + (\mathcal{A}_0 u_2 + \mathcal{A}_1 u_1 + \mathcal{A}_2 u_0) + \varepsilon (\mathcal{A}_1 u_2 + \mathcal{A}_2 u_1) + \varepsilon^2 \mathcal{A}_2 u_2 + \cdots = f. \quad (2.8)
\]

We equate equal powers of \( \varepsilon \) in the above equation and disregard all terms of order higher than 1 to obtain the following sequence of problems:

\[
\mathcal{O} \left( \frac{1}{\varepsilon^2} \right): \quad \mathcal{A}_0 u_0 = 0, \quad u_0 \text{ is } 1\text{-periodic}, \quad (2.9a)
\]

\[
\mathcal{O} \left( \frac{1}{\varepsilon} \right): \quad \mathcal{A}_0 u_1 = -\mathcal{A}_1 u_0, \quad u_1 \text{ is } 1\text{-periodic}, \quad (2.9b)
\]

\[
\mathcal{O}(1): \quad \mathcal{A}_0 u_2 = -\mathcal{A}_1 u_1 - \mathcal{A}_2 u_0 + f, \quad u_2 \text{ is } 1\text{-periodic}. \quad (2.9c)
\]

Notice that the first to equations in the above sequence, equations (2.9a) and (2.9b), have to be satisfied so that no singularities appear; indeed, the first two terms on the right hand side of equation (2.8) diverge as \( \varepsilon \rightarrow 0 \), unless equations (2.9a) and (2.9b) are satisfied. From this point of
view, the problem of homogenization may be thought of a typical problem in the theory of singular
perturbations for differential equations. We refer to, e.g., [26, 29] for a presentation of singular
perturbation theory.

In order to proceed with the derivation of the homogenized equation we have to study equations
(2.9). These equations are of the form

$$\mathcal{A}_0 u = h, \quad u \text{ is } 1\text{--periodic}, \quad (2.10)$$

with $u = u(x, y)$ and similarly $h = h(x, y)$. Let us make some remarks and derive some simple
properties of the boundary value problem (2.10). A more detailed, rigorous treatment will be given
in the next chapter, sections

refsec:per and (3.5). Good references for elliptic PDE with periodic boundary conditions are [37, 
ch. 5] and [49, ch. 6].

We start our study of (2.10) by noticing that, although both $u(x, y)$ and $h(x, y)$ are functions
of $x$ as well as $y$, the variable $x$ enters merely as a parameter; indeed, $\mathcal{A}_0$ is a partial differential
operator with respect to $y$. This, of course, is related to our assumption of separation of length
scales: equation (2.10) is posed on the unit cell $Y$ and the operator $\mathcal{A}_0$ “sees” only the small scale
structures which are expressed in our problem through the variable $y$.

Now, we need to find necessary and sufficient conditions for problem 2.10 to be well posed.
For this we will need a calculation which is presented in the next lemma.

**LEMMA 2.1.** Let $F(y)$ be a smooth 1--periodic function. Then

$$\int_Y \frac{\partial F(y)}{\partial y_i} \, dy = 0, \quad i = 1, \ldots, d.$$

**Proof.** This is a simple consequence of the fact that $F(y)$ is periodic in $y$, together with the fundamental theorem of calculus:

$$\int_Y \frac{\partial F(y)}{\partial y_i} \, dy = \int_{y_1} \cdots \int_{y_i} \cdots \int_{y_d} F(y_1, \ldots, y_i, \ldots, y_d) \, dy$$

$$= \int_{y_1} \cdots \int_{y_{i-1}} \int_{y_{i+1}} \cdots \int_{y_d} (F(y_1, \ldots, 1, \ldots, y_d) - F(y_1, \ldots, 0, \ldots, y_d)) \prod_{j=1, j \neq i}^d dy_j$$

$$= 0, \quad i = 1, \ldots, d. \quad \Box$$

\(^4\)Throughout these notes we will be using the notation $dy = \prod_{j=1}^d dy_j$. 

---

1. (Continued)
More generally, let $F(y)$ and $G(y)$ be smooth, 1-periodic functions. Then, the following integration by parts formula holds.

$$
\int_Y \frac{\partial F}{\partial y_i} G \, dy = - \int_Y F \frac{\partial G}{\partial y_i} \, dy, \quad i = 1, \ldots, d. \tag{2.11}
$$

Lemma 2.1 enables to show that, in order for equation (2.10) to be well posed\(^5\), it is necessary for the right hand side of the equation to average to 0 over the unit cell. We start by proving that it is a necessary condition for a solution to exist.

**Lemma 2.2.** A necessary condition for the existence of a solution to (2.10) is

$$
\int_Y h(y) \, dy = 0. \tag{2.12}
$$

*Proof.* Let $u$ be a solution of (2.10). We integrate the left hand side of the equation and use Lemma 2.1 to obtain:

$$
\int_Y \mathcal{A}_0 u = - \int_Y \frac{\partial}{\partial y_k} \left( a_{ij} (y) \frac{\partial u}{\partial y_j} \right) \, dy = 0
$$

which makes sense only if condition (2.12) holds. \(\Box\)

We will call (2.12) the *solvability condition*. It will play a very important role in our analysis.

What happens now if the right hand side of (2.10) is identically equal to 0? This situation is analyzed in the following proposition.

**Proposition 2.3.** The only solutions of the homogeneous equation

$$
\mathcal{A}_0 u = 0, \tag{2.13}
$$

are constants in $y$.

*Proof.* Let $u$ be a solution of (2.13). We multiply the equation by $u$, integrate over $Y$, integrate by

\(^5\)By the term well-posed we mean the existence and uniqueness of a solution which depends continuously on the data of the problem.
parts using (2.11) and use the uniform ellipticity assumption on \( A(y) \) to obtain

\[
0 = \int_Y u A_0 u \, dy
= -\int u \frac{\partial}{\partial y_k} \left( a_{ij}(y) \frac{\partial u}{\partial y_j} \right) \, dy
= \int a_{ij}(y) \frac{\partial u}{\partial y_j} \frac{\partial y_i}{\partial y_j} \, dy
\geq \alpha \int_Y |\nabla_y u|^2 \, dy,
\]

where the notation

\[
|\nabla_y u|^2 = \sum_{i=1}^d \left| \frac{\partial u}{\partial y_i} \right|^2
\]

has been used. We have thus obtained the inequality

\[
\int_Y |\nabla_y u|^2 \, dy \leq 0.
\]

This inequality can be satisfied only if the gradient of \( u \) with respect to \( y \) vanishes and, hence, only if \( u \) is constant in \( y \). \( \square \)

An immediate corollary of the above proposition is that if solutions to equation (2.10) exist, then they are unique up to constants in \( y \).

**COROLLARY 2.4.** All solutions of (2.10) differ by a constant in \( y \).

*Proof.* Let \( u_1 \) and \( u_2 \) be two solutions of (2.10) and let \( u = u_1 - u_2 \). We use the linearity of the operator \( A_0 \) to obtain an equation for \( u \):

\[
A_0 u = 0.
\]

Hence, Proposition 2.3 applies and we conclude that the functions \( u_1 \) and \( u_2 \) differ by a constant. \( \square \)

Notice carefully that we haven’t proved the existence of a solution to (2.10). In the next chapter we will prove the existence and uniqueness of solutions—up to constants in \( y \) of (2.10), and we will see that this is a consequence of the **Fredholm alternative**. However, the above results will be sufficient for our purposes in this chapter. Let us also remark that, among all solutions of (2.10), we will choose the one whose integral over \( Y \) vanishes:

\[
A_0 u = h, \quad u \text{ is } 1\text{-periodic, } \int_Y u \, dy = 0.
\]
Now we are ready to analyze equations (2.9). We start with (2.9b). The Fredholm alternative, in particular Proposition 2.3 implies that $u_0$ is constant in $y$:

$$u_0 = u(x). \quad (2.14)$$

This means that the first term in the multiple scales expansion is independent of the fast scales which are represented by $y$. Consequently we can hope to derive a homogenized equation for $u(x)$ which is independent of the microscopic scales.

Let us proceed now with (2.9b), which in view of equation (2.14) becomes:

$$\mathcal{A}_0 u_1 = \frac{\partial a_{ij}}{\partial y_i} \frac{\partial u}{\partial x_j}. \quad (2.15)$$

We need to check that the solvability condition is satisfied. We use Lemma 2.1, together with the fact that $u(x)$ is independent of $y$ to deduce:

$$\int_Y \frac{\partial a_{ij}}{\partial y_i} \frac{\partial u}{\partial x_j} \, dy = \frac{\partial u}{\partial x_j} \int_Y \frac{\partial a_{ij}}{\partial y_i} \, dy = 0.$$

Thus, the solvability condition is satisfied and equation (2.15) is well posed: it admits a unique, up to constants in $y$, solution. We can solve (2.15) using separation of variables. To this end we look for a solution which has the following form:

$$u_1(x, y) = \chi^j(y) \frac{\partial u}{\partial x_j} + \hat{u}_1(x). \quad (2.16)$$

Upon substituting (2.16) into (2.15) we obtain the cell problem:

$$\mathcal{A}_0 \chi^j = \frac{\partial a_{ij}}{\partial y_i}, \quad \chi^j(y) \text{ is 1–periodic, } j = 1, \ldots, d. \quad (2.17)$$

This is precisely equation (1.7). As we have already mentioned $\chi^j(y)$ is called the first order corrector field. Notice that the periodicity of the coefficients implies that the right hand side of equation (2.17) averages to zero over the unit cell and consequently the cell problem is well posed. We ensure the uniqueness of solutions to (2.17) by requiring the corrector field to have zero average. We also remark that the function $\hat{u}_1(x)$ is undetermined at this point. It will become clear from the subsequent analysis, however, that it is not needed for the derivation of the homogenized equation.

Now we consider equation (2.9a). In order for this equation to be well posed it is necessary and sufficient for the right hand side of this equation to average to zero. Since we have assumed that the function $f(x)$ is independent of $y$ the solvability condition implies:

$$\int_Y (A_1 u_1 + A_2 u_0) \, dy = f. \quad (2.18)$$
We start with the first term on the left hand side of the above equation:
\[
\int_Y \mathcal{A}_2 u_0 \, dy = \int_Y -\frac{\partial}{\partial x_i} \left( a_{ij}(y) \frac{\partial u(x)}{\partial x_j} \right) \, dy \\
= -\frac{\partial}{\partial x_i} \left[ \left( \int_Y a_{ij}(y) \, dy \right) \frac{\partial u(x)}{\partial x_j} \right].
\] (2.19)

Moreover:
\[
\int_Y \mathcal{A}_1 u_2 \, dy = \int_Y \left[ -\frac{\partial}{\partial y_i} \left( a_{ij}(y) \frac{\partial u_1}{\partial x_j} \right) - \frac{\partial}{\partial x_i} \left( a_{ij}(y) \frac{\partial u_1}{\partial y_j} \right) \right] \, dy \\
= I_1 + I_2.
\]

The first term on the right hand side of the above equation can be dealt with upon using Lemma 2.1:
\[
I_1 = \int_Y -\frac{\partial}{\partial y_i} \left( a_{ij}(y) \frac{\partial u_1}{\partial x_j} \right) \, dy \\
= 0.
\]

now we consider to \( I_2 \):
\[
I_2 = \int_Y -\frac{\partial}{\partial x_i} \left( a_{ij}(y) \frac{\partial u_1}{\partial y_j} \right) \, dy \\
= \int_Y a_{ij}(y) \frac{\partial}{\partial x_i} \left( \frac{\partial}{\partial y_j} \left( \chi^k(y) \frac{\partial}{\partial x_k} u(x) \right) \right) \, dy \\
= \left[ \int_Y a_{ij}(y) \frac{\partial}{\partial x_j} \chi^k(y) \right] \frac{\partial^2 u}{\partial x_i \partial x_k}.
\] (2.20)

We substitute (2.20) and (2.19) in (2.18) to obtain the homogenized equation
\[
-\overline{\alpha}_{ij} \frac{\partial^2 u}{\partial x_i \partial x_j} = f, \quad \text{for } x \in \Omega,
\] (2.21a)
\[
u(x) = 0, \quad \text{for } x \in \partial \Omega,
\] (2.21b)

where the homogenized coefficients \( \{\overline{\alpha}_{ij}\}_{i,j=1}^d \) are given by the formula:
\[
\overline{\alpha}_{ij} = \int_Y \left( a_{ij}(y) + a_{ik}(y) \frac{\partial \chi^i(y)}{\partial y_k} (y) \right) \, dy, \quad i,j = 1, \ldots, d.
\] (2.22)

Let us now summarize the homogenization procedure. The first step is to assume the ansatz (2.3) for the solution \( u^\epsilon \) of problem (2.1). Upon substituting this expansion into this equation and equating equal powers in \( \epsilon \) we obtain equations (2.9). Application of Fredholm alternative to equation
(2.9a) gives that the first term in the expansion is independent of the microscopic variable \( y \). Solving equation (2.9b) through separation of variables provides us with the cell problem (2.17). Finally, we obtain the homogenized equation (2.21) by imposing the solvability condition to (2.9a). We can see that this method is quite general and it can be applied to a variety of problems, both linear and non–linear. The basic tool is a solvability condition of the form (2.12). In the case of the boundary value problem that we analyzed in this section the solvability condition was a result of assumptions (2.2). We refer to [17, 18] where similar arguments—though more involved!—are applied to nonlinear PDE. However, it is quite often the case that the solvability condition is not straightforward even for linear problems. We will study such an example in section 6.2

Now, the solution of the homogenized equation (2.21) involves the following steps: First we need to solve the cell problem (2.17) which is a uniformly elliptic PDE with periodic boundary conditions. Then, we need to calculate the integrals in (2.22). Finally, we solve the homogenized equation (2.21). The numerical solution of the cell problem is quite standard by means of a spectral method\(^6\). Similarly, it is quite straightforward to compute the effective coefficients \( \{ a_{ij} \}_{i,j=1}^{d} \). Finally, depending of course on the domain \( \Omega \), it is not hard to solve the homogenized equation, which is a PDE with constant coefficients. The above discussion shows that it is very advantageous to solve the homogenized equation (2.21), as opposed to solving the original equation (2.1); this equation leads to an ill–posed numerical problem when \( \epsilon \ll 1 \).

### 2.3 Some Exactly Solvable Cases

Generally, neither the cell problem (2.17), can be solved in closed form, nor is it possible to compute the integrals (2.22) analytically. There are some cases, however, where this is possible. The prime example when this is the case is when considering problem (2.1a) in one dimension, \( d = 1 \). In one dimension our PDE become ordinary differential equations which can be solved by quadratures.

Let then \( d = 1 \) and let \( \Omega = [0, L] \). Now the Dirichlet problem (2.1a) reduces to a two–point boundary value problem:

\[
- \frac{d}{dx} \left( a \left( \frac{x}{\epsilon} \right) \frac{du^\epsilon}{dx} \right) = f, \quad x \in (0, L),
\]

\[(2.23a)\]

\^6\ Which for the case of equation (2.17) amounts to expanding the corrector field in a truncated Fourier series and computing Fourier coefficients by solving the resulting linear system of equations. We refer e.g. to [34, 35] for some details.
Similarly, the cell problem becomes an ordinary differential equation:

\[-\frac{d}{dy} \left( a(y) \frac{d\chi}{dy} \right) = \frac{d}{dy}(a(y)), \quad y \in (0, 1),\]

\[\chi(0) = \chi(1), \quad \int_0^1 \chi \, dy = 0.\]

In the one dimensional case we only have one effective coefficient which is given by the one dimensional version of (2.22)

\[\bar{a} = \int_0^1 \left( a(y) + a(y) \frac{d\chi(y)}{dy} \right) \, dy\]

We assume, as before, that \(a(y)\) is smooth, periodic with period 1. The uniform ellipticity assumption implies that \(a(y)\) is bounded away from 0. We combine this with the fact that it is uniformly bounded to write

\[\alpha \leq a(y) \leq \beta, \quad y \in [0, 1],\]

for some positive constants \(\alpha \leq \beta\). We also assume that \(f\) is a smooth function.

Now, equation (2.24) can be solved exactly after two integrations. To this end, we integrate the equation once to obtain:

\[a(y) \frac{d\chi}{dy} = -a(y) + c_1.\]

The constant \(c_1\) is undetermined at this point. The left part of inequality (2.26), allows us to divide (2.27) by \(a(y)\). We then integrate to deduce:

\[\chi(y) = -y + c_1 \int_y^1 \frac{1}{a(y)} \, dy + c_2.\]

In order to determine the constant \(c_1\) we use the fact that \(\chi(y)\) is a periodic function:

\[\chi(0) = \chi(1) \Rightarrow 0 = 1 - c_1 \int_0^1 \frac{1}{a(y)} \, dy\]

\[\Rightarrow c_1 = \frac{1}{\int_0^1 \frac{1}{a(y)} \, dy}.\]

\(^7\)Notice that in order to ensure uniqueness of solutions for the cell problem we need to impose an additional condition on \(\chi(y)\) than we did on \(u^\varepsilon\) which solves (2.23). Namely, we need to specify its average over the unit cell. This is a manifestation of the difference between Dirichlet and periodic boundary value problems.
We could use the fact that the average of $\chi(y)$ over the unit cell vanishes in order to determine the second coefficient $c_2$. However, only the derivative of the corrector field enters in the formula for the effective diffusivity, equation (2.25), which does not require the calculation of $c_2$:

$$\frac{d\chi}{dy} = 1 - \frac{1}{\int_0^1 a(y)^{-1} dy} \int_0^y \frac{1}{a(y)} dy.$$ 

We substitute this expression in equations (2.25) to obtain:

$$\bar{\alpha} = \frac{1}{\int_0^1 a(y)^{-1} dy}. \tag{2.29}$$

This is the formula which gives the homogenized, effective, coefficients in one dimension. Notice that even this formula involves the integral of the inverse of the original coefficients $a(y)$. The computation of this integral will involve, in general, numerical quadrature.

Even if cannot compute the integral in (2.29), we can obtain some qualitative information on $\bar{\alpha}$. This is the content of the following lemma.

**Lemma 2.5.** The homogenized coefficient $\bar{\alpha}$ has the same upper and lower bound as $a(y)$. Moreover, it is bounded from above by the average of $a(y)$:

$$\alpha \leq \bar{\alpha} \leq \beta,$$

and

$$\bar{\alpha} \leq \int_0^1 a(y) \, dy.$$

**Proof.** 1. We invert inequality (2.26) to obtain

$$\beta^{-1} \leq a(y)^{-1} \leq \alpha^{-1}.$$

We integrate the above inequality over $[0, 1]$ to derive:

$$\beta^{-1} \leq \int_0^1 a(y)^{-1} \, dy \leq \alpha^{-1},$$

from which the first claim in the lemma follows, upon inversion.

2. We use the Cauchy–Schwarz inequality to deduce:

$$1 = \int_0^1 1 \, dy = \int_0^1 \sqrt{a} \sqrt{a^{-1}} \, dy \leq \sqrt{\int_0^1 a \, dy} \sqrt{\int_0^1 a^{-1} \, dy},$$

from which the second claim in the lemma follows, upon squaring and inverting. \qed
2.4 Properties of the Homogenized Equation

In section 2.2 we derived formally the homogenized equation (2.21) using the method of multiple scales and the Fredholm alternative. In this section we study some of the properties of this equation or, rather, of the homogenized coefficients. In particular, we show that the effective coefficients matrix \( \overline{A} = \{ \overline{a}_{ij} \}_{i,j=1}^{d} \) is positive definite, which means that the homogenized differential operator is uniformly elliptic. Moreover, we prove that if the coefficients matrix \( A = \{ a_{ij}(y) \}_{i,j=1}^{d} \) is symmetric, then the homogenized coefficients matrix \( \overline{A} \) is also symmetric. Finally, we show that the homogenization process can create anisotropies: even if the matrix \( A(y) \) is a diagonal, the matrix of homogenized coefficients \( \overline{A} \) need not be.

In order to analyze the properties of the matrix of homogenized coefficients it will be useful to find an alternative representation for \( \overline{A} \). To this end, we introduce the bilinear form

\[
a_1(\phi, \psi) = \int_{Y} a_{ij} \frac{\partial \phi}{\partial x_j} \frac{\partial \psi}{\partial x_i} \, dy,
\]

which is defined for all smooth periodic functions \( \phi, \psi \in C^\infty_{\text{per}}(Y) \). We start with obtaining an alternative, equivalent formulation for the cell problem.

**Lemma 2.6.** The cell problem (2.17) can be written in the form

\[
a_1(\chi^\ell - y_k, \phi) = 0, \quad \forall \phi \in C^\infty_{\text{per}}(Y).
\]

**Proof.** We multiply the cell problem by a test function \( \phi \in C^\infty_{\text{per}}(Y) \) and integrate over the unit cell:

\[
- \int_{Y} A_{0}u \phi \, dy = - \int_{Y} \sum_{i,j} \frac{\partial a_{ij}}{\partial y_i} \phi \, dy.
\]

We use the integration by parts formula (2.11) to rewrite the left hand side of (2.32) in the form

\[
- \int_{Y} A_{0}u \phi \, dy = \int_{Y} a_{ij}(y) \frac{\partial \chi^\ell}{\partial y_j} \frac{\partial \phi}{\partial y_i} \, dy.
\]

Consider now the right hand side of (2.32). We can rewrite it as follows:

\[
\int_{Y} a_{ij}(y) \frac{\partial \chi^\ell}{\partial y_j} \frac{\partial \phi}{\partial y_i} \, dy = \int_{Y} a_{ij}(y) \frac{\partial \phi}{\partial y_i} \, dy = \int_{Y} a_{ij}(y) \frac{\partial \phi}{\partial y_i} \delta_{ij} \, dy = \int_{Y} a_{ij}(y) \frac{\partial \phi}{\partial y_i} \frac{\partial y_j}{\partial y_i} \, dy.
\]

\(^{8}\text{Actually, we need much less smoothness. See Chapter 3}\)
In the above $\delta_{ij}$ denotes the Kronecker delta. Consequently, the cell problem can be written as

$$a_1(\chi^\ell, \phi) = a_1(y_\ell, \phi), \quad \forall \phi \in C^\infty_{per}(Y), \quad \ell = 1, \ldots, d,$$  
(2.34)

which is equivalent to (2.31).

Now we are ready to give an alternative representation formula for the homogenized coefficients.

**Lemma 2.7.** The effective coefficients $\overline{a}_{ik}$ are given by the following formula:

$$\overline{a}_{ij} = a_1(\chi^j - y_j, \chi^i - y_i), \quad i, j = 1, \ldots, d.$$  
(2.35)

**Proof.** We have:

$$\overline{a}_{ij} = \int_Y \left( a_{ij} - a_{ik} \frac{\partial \chi^j}{\partial y_k} \right) dy = \int_Y \left( a_{mn} \frac{\partial y_j}{\partial y_n} \frac{\partial y_k}{\partial y_m} - a_{nk} \frac{\partial \chi^j}{\partial y_k} \frac{\partial y_i}{\partial y_n} \right) dy = a_1(y_j, y_i) - a_1(\chi^j, y_i) = a_1(y_j - \chi^j, y_i) = a_1(y_j - \chi^j, y_i - \chi^j) = a_1(y_j - \chi^j, \chi^i - y_i),$$  
(2.36)

for all $i, j = 1, \ldots, d$. The lemma is proved.

Now we can prove that the matrix $A$ is positive definite. We have the following.

**Lemma 2.8.** The homogenized differential operator

$$A = \overline{a}_{ij} \frac{\partial^2}{\partial x_i \partial x_j}$$

is uniformly elliptic.

**Proof.** We need to show that

$$\overline{a}_{ij} \xi_i \xi_j \geq \alpha |\xi|^2, \quad \alpha > 0.$$  

This is the weak formulation of the cell problem.
We use the representation formula (2.35) to deduce:

\[
\overline{a}_{ij} \xi_i \xi_j = a_1 (\chi^j - y_j, \chi^i - y_i) \xi_i \xi_j \\
= a_1 ((\chi^j - y_j) \xi_j, (\chi^i - y_i) \xi_i) \\
:= a_1 (w, w),
\]

with \( w = \xi_i (\chi - y_k) \). We use now the positive definiteness of the \( A(y) \) to obtain:

\[
\overline{a}_{ij} \xi_i \xi_j = a_1 (w, w) \\
= \int_Y a_{ij} \frac{\partial w}{\partial y_i} \frac{\partial w}{\partial y_j} dy \\
\geq \alpha \int_Y |\nabla_y w|^2 dy,
\]

with \( \nabla_y w = \frac{\partial w}{\partial y_k} e_k \). The above computation shows that \( \overline{A} \) is nonnegative. To show that it is positive definite we argue as follows. Let \( \overline{a}_{ij} \xi_i \xi_j = 0 \). This implies that \( \frac{\partial w}{\partial y_k} = 0, \ k = 1, \ldots, d \) which means that \( w = c \) and consequently that \( \xi_i y_k = \xi_i \chi^i (y) - c \). The right hand side of this equation is 1–periodic in \( y \) and consequently the left hand side should also be. The only way this can happen is if \( \xi_i = 0, \ i = 1, \ldots, d \). Consequently, \( \overline{a}_{ij} \xi_i \xi_j = 0 \) if and only if \( \xi_i = 0, \ i = 1, \ldots, d \). This completes the proof of the lemma. \( \square \)

This above lemma shows that uniform ellipticity is a property that is preserved under the homogenization procedure. In particular, this implies that the homogenized equation is well posed, since it it a uniformly elliptic PDE with constant coefficients. Another property that is preserved is that of the symmetry of the diffusion tensor \( A(y) \). We have the following lemma.

**Lemma 2.9.** Assume that the coefficients matrix \( A(y) \) is symmetric. Then the homogenized matrix \( \overline{A} \) is also symmetric.

**Proof.** The symmetry of \( A \) implies the symmetry of the bilinear form \( a_1 (\phi, \psi) \):

\[
a_1 (\phi, \psi) = a_1 (\psi, \phi) \ \forall \ \psi, \phi \in C^\infty (Y).
\]

Consequently:

\[
\overline{a}_{ij} = a_1 (\chi^j - y_j, \chi^i - y_i) \\
= a_1 (\chi^i - y_i, \chi^j - y_j) \\
= a_1 (w, w),
\]
and thus $\overline{A} = \overline{A}^T$, where $\overline{A}^T$ denotes the transpose of $A$.

On the contrary, homogenization does not preserve isotropy. In particular, even if the diffusion
text

matrix $A$ has only diagonal non–zero elements, the homogenized diffusion matrix will in general
have non–zero off–diagonal elements. To see this, let us assume that $a_{ij} = 0$, $i \neq j$. Then, the
off–diagonal elements of the homogenized diffusion matrix are given by the formula

$$
\overline{a}_{ij} = -\int_Y a_{ij} \frac{\partial \chi^j}{\partial y_i} dy, \quad i \neq j \quad \text{no summation},
$$

This expression is not identically equal to zero. This leads to the surprising result that an isotropic

composite material behaves, at the limit where the microstructure becomes finer and finer, like an

anisotropic homogeneous material.

### 2.5 Remarks on the Method of Multiple Scales–Extensions

#### 2.5.1 Higher Order Correctors

In section (2.1) we studied homogenization for the Dirichlet boundary value problem (2.1) using
the method of multiple scales. We derived the homogenized equation (2.21) and the cell problem
(2.17). We also computed the first order correction $u_1(\xi, \eta)$, up to an unknown function $\hat{u}_1(x)$:

$$
u_1(\xi, \eta) = -\chi^i(\eta) \frac{\partial u(\xi)}{\partial x_i} + \hat{u}_1(x),
$$

with $y = \frac{\xi}{\epsilon}$. Now we can proceed with solving equation (2.9c) and computing the second correc-
tion $u_2(x, y)$. We substitute (2.40) into (2.9a) and use the homogenized equation (2.21) to obtain:

$$
A_0 u_2 = f - A_1 u_1 - A_2 u_0 \\
= -\overline{a}_{ij} \frac{\partial^2 u}{\partial x_i \partial x_j} + \frac{\partial}{\partial x_i} \left( a_{ij}(\eta) \frac{\partial u_1}{\partial y_j} \right) + \frac{\partial}{\partial y_j} \left( a_{ij}(\eta) \frac{\partial u_1}{\partial x_j} \right) + \frac{\partial}{\partial x_i} \left( a_{ij}(\eta) \frac{\partial u}{\partial x_j} \right) \\
= -\overline{a}_{ij} \frac{\partial^2 u}{\partial x_i \partial x_j} + \frac{\partial}{\partial x_i} \left( a_{ij}(\eta) \frac{\partial \chi^k(\eta)}{\partial y_j} \frac{\partial u}{\partial x_k} \right) - \frac{\partial}{\partial y_j} \left( a_{ij}(\eta) \chi^k(\eta) \frac{\partial^2 u}{\partial x_j \partial x_k} \right) + a_{ij}(\eta) \frac{\partial^2 u}{\partial y_i \partial x_j} \\
= -\overline{a}_{ij} \frac{\partial^2 u}{\partial x_i \partial x_j} + a_{ij}(\eta) \frac{\partial^2 u}{\partial x_i \partial x_j} - \left( a_{ij}(\eta) \frac{\partial \chi^k(\eta)}{\partial y_j} \right) \frac{\partial^2 u}{\partial x_i \partial x_k} - \frac{\partial}{\partial y_j} \left( a_{ij}(\eta) \chi^k(\eta) \right) \frac{\partial^2 u}{\partial x_j \partial x_k} \\
:= b_{ij}(\eta) \frac{\partial^2 u}{\partial x_i \partial x_j},
$$

with

$$
b_{ij}(\eta) = -a_{ij} + a_{ij}(\eta) - a_{ik}(\eta) \frac{\partial \chi^j}{\partial y_k} - \frac{\partial (a_{kj}(\eta) \chi^j)}{\partial y_k}, \quad i, j = 1, \ldots, d.
$$
In writing (2.42) we have renamed various dummy indices which appear in (2.41). Equation (2.41) is well posed since the right hand side averages to zero over the unit cell\(^{10}\). Now we can solve this equation using separation of variables:

\[
 u_2(x, y) = -\theta^{ij}(y) \frac{\partial^2 u}{\partial x_i \partial x_j} + \hat{u}_2(x), \tag{2.43}
\]

where the *second order corrector field* \( \{ \theta^{ij}(y) \}^d_{i,j=1} \) satisfies the following cell problem:

\[
 \mathcal{A}_0 \theta^{ij} = -b_{ij}, \quad i, j = 1, \ldots, d. \tag{2.44}
\]

where \( b_{ij}(y) \) is given by (2.42). The second order corrector \( u_2(x, y) \) given by (2.43) will be needed in the proof of the convergence theorem in Chapter 3.

Of course, it is also possible to obtain higher order corrections by solving higher order equations iteratively using separation of variables and introducing additional cell problems. All higher order equations are of the form

\[
 \mathcal{A}_0 u_2 = -\mathcal{A}_1 u_{k-1} - \mathcal{A}_2 u_{k-2}, \quad u_k \text{ is 1–periodic, } k = 3, \ldots. \tag{2.45}
\]

We remark that the \( k \)th order corrector \( u_k(x, y) \) will be proportional to the \( k \)th order partial derivative of the solution \( u(x) \) to the homogenized equation. After having computed the first \( k \) correctors we can approximate \( u^\epsilon(x) \) as follows:

\[
 u^\epsilon(x) \approx \sum_{i=0}^{k} u_i(x, \frac{x}{\epsilon}).
\]

It should be intuitively clear that by adding higher order terms in the expansion we are getting closer to the solution of the Dirichlet problem (2.1); that is, the distance between the solution of (2.1) and the expansion defined above becomes smaller as we add terms. We will make this intuition precise in the next chapter.

**EXERCISE 2.10.** Compute all higher order terms in the expansion and obtain the corresponding cell problems.

### 2.5.2 Different Boundary Conditions

The elliptic boundary value problem (2.1) that we considered in the previous section was a Dirichlet problem. However, an inspection of the analysis presented in section 2.2 reveals that the boundary

\(^{10}\)We ensured that by imposing the solvability condition which led to the homogenized equation.
conditions did not play any role in the derivation of the homogenized equation. In particular, the two-scale expansion (2.3) that we used in order to derive the homogenized equation did not contain any information concerning the boundary conditions of the problem under investigation. Indeed, the boundary conditions become somewhat irrelevant in the homogenization procedure: exactly the same calculations would enable us to obtain the homogenized equation for Neumann or mixed boundary conditions.

The boundary conditions become very important when trying to prove the homogenization theorem. The fact that the two-scale expansion (2) does not satisfy the boundary conditions of our PDE exactly but only up to $O(\epsilon)$ introduces boundary layers [26, ch. 3]\(^{11}\). Boundary layers affect the convergence rate, i.e. the rate with which $u^\epsilon(x)$ converges to $u(x)$ as $\epsilon \to 0$. We can solve this problem by modifying the two-scale expansion (2.3), adding additional terms which take care of the boundary layer and vanish exponentially fast as we move away from the boundary so that they do not affect the solution in the interior. We refer to [5] for details.

### 2.5.3 Locally Periodic Coefficients

In the Dirichlet problem that we analyzed in section 2.2 we assumed that the coefficients $\{a_{ij}^\epsilon(x)\}_{i,j}^d$ depend only on microscale:

$$a_{ij}^\epsilon(x) = a_{ij}(\frac{x}{\epsilon}), \quad i, j = 1, \ldots, d,$$

with $\{a_{ij}(y)\}_{i,j}^d$ being 1-periodic functions. However, the method of multiple scales would also be applicable to the case where the coefficients depend explicitly on the macroscale as well as the microscale, i.e. when they are *locally periodic*\(^{12}\):

$$a_{ij}^\epsilon(x) = a_{ij}(x, \frac{x}{\epsilon}), \quad i, j = 1, \ldots, d,$$

with $\{a_{ij}(x, y)\}_{i,j}^d$ being 1-periodic in $y$ and smooth in $x$. An analysis similar to the one presented in section 2.2 enables us to obtain the homogenized equation for the Dirichlet problem

$$-\frac{\partial}{\partial x_i} \left( a_{ij} \left( x, \frac{x}{\epsilon} \right) \frac{\partial u^\epsilon}{\partial x_j} \right) = f, \quad \text{for } x \in \Omega \quad \text{(2.49a)}$$

$$u^\epsilon(x) = 0, \quad \text{for } x \in \partial \Omega. \quad \text{(2.49b)}$$

\(^{11}\)The presence of boundary and initial layers is a common feature in all problems of singular perturbations. See e.g. [26] and [29] for further details.

\(^{12}\)The term nonuniformly periodic coefficients is also used.
Now the homogenized coefficients are functions of $x$:

\[
- \frac{\partial}{\partial x_i} \left( \overline{a}_{ij}(x) \frac{\partial u}{\partial x_j} \right) = f, \quad \text{for } x \in \Omega \tag{2.50a}
\]

\[
u(x) = 0, \quad \text{for } x \in \partial \Omega, \tag{2.50b}
\]

and the cell problem reads:

\[
- \frac{\partial}{\partial y_i} \left( a_{ij}(x, y) \frac{\partial \chi^k(x, y)}{\partial y_j} \right) = - \frac{\partial a_{ik}(x, y)}{\partial y_i}, \quad k = 1, \ldots, d. \tag{2.51}
\]

The homogenized coefficients are given by the formula:

\[
\overline{a}_{ij}(x) = \int_Y \left( a_{ij}(x, y) - a_{ik}(x, y) \frac{\partial \chi^j(x, y)}{\partial y_k} \right) dy, \quad i, j = 1, \ldots, d. \tag{2.52}
\]

We emphasize the fact that the "macroscopic variable" $x$ enters in the above two equations merely as a parameter: the operator $\mathcal{A}_0 := - \frac{\partial}{\partial y_i} \left( a_{ij}(x, y) \frac{\partial}{\partial y_j} \right)$ which appears in the cell problem (2.51) is a partial differential operator in $y$ for every $x \in \Omega$ and the integrals in (2.52) are taken over the unit cell $Y$ with respect to $y$, for every $x \in \Omega$. Consequently, in order to compute the effective coefficients $\{\overline{a}_{ij}(x)\}_{i,j=1}^d$ we need to solve the cell problem (2.51) and evaluate the integrals in (2.52) at all points $x \in \Omega$.

**EXERCISE 2.11.** Consider the boundary value problem (2.49). Use the method of multiple scales to obtain the homogenized equation (2.50), the cell problem (2.51) and the formula for the homogenized coefficients (2.52). Verify that the properties of these coefficients presented in section (2.4) are still valid.

### 2.5.4 Time Dependent Problems

So far in this chapter we have used the method of multiple scales to study the homogenization of elliptic, i.e. time independent problems. However, the method is also applicable to evolution PDE with rapidly oscillating coefficients. Let us consider the initial/boundary value problem for the following parabolic(diffusion) PDE:

\[
\frac{\partial u^\epsilon}{\partial t} - \frac{1}{\epsilon^2} \frac{\partial}{\partial x_i} \left( a_{ij}(\frac{x}{\epsilon}) \frac{\partial u^\epsilon}{\partial x_j} \right) = f(x, t) \quad \text{in } \Omega \times (0, T), \tag{2.53a}
\]

\[
u^\epsilon(x, t) = 0 \quad \text{on } \partial \Omega \times (0, T) \tag{2.53b}
\]

\[
u^\epsilon(x, 0) = u_{in}(x) \quad \text{in } \Omega. \tag{2.53c}
\]
We assume that the coefficients \( \{a_{ij}(y)\}_{i,j=1}^{d} \) satisfy conditions (2.2). Despite the fact that (2.53) is an evolution problem, the fact that the coefficients do not depend on time implies that we still have oscillations only in the spatial variable \( x \); in other words, we do not have a temporal microscale. Consequently, an expansion of the form (2.3) is still valid, where the time variable \( t \) enters merely as a parameter:

\[
 u^\varepsilon(x, t) = u_0(x, x \varepsilon, t) + \varepsilon u_1(x, x \varepsilon, t) + \varepsilon^2 u_2(x, x \varepsilon, t) + \ldots \tag{2.54}
\]

We can carry out an analysis similar to the one presented in section 2.2 for the Dirichlet problem. In fact, the \( O(\varepsilon^2) \) and \( O(\varepsilon) \) equations are exactly the same as before, equation (2.9c), (2.9b). The \( O(1) \) equation is modified by the addition of the partial derivative with respect to time of \( u(x, t) \) on the right hand side of (2.9c):

\[
 A_0 u_2 = -A_1 u_1 - A_2 u_2 - \frac{\partial u}{\partial t} + f.
\]

All the operators in the previous equation are the ones defined in section 2.2. Now, an application of the Fredholm alternative enables us to obtain the homogenized problem:\(^{13}\)

\[
 \frac{\partial u}{\partial t} - \frac{\partial}{\partial x_i} \left( a_{ij} \frac{\partial u}{\partial x_j} \right) = f(x, t) \quad \text{in} \quad \Omega \times (0, T), \tag{2.56a}
\]

\[
 u(x, t) = 0 \quad \text{on} \quad \partial \Omega \times (0, T) \tag{2.56b}
\]

\[
 u(x, 0) = u_{in}(x) \quad \text{in} \quad \Omega. \tag{2.56c}
\]

The homogenized coefficients \( \{\overline{a}_{ij}\}_{i,j=1}^{d} \) are given by formula (2.22) and the cell problem is also the same as before and given by equation (2.17).

**EXERCISE 2.12.** Derive the homogenized equation (2.56), together with the cell problem and the formula for the homogenized coefficients using the method of multiple scales with the expansion (2.54).

The situation becomes somewhat different when the coefficients of our evolution PDE oscillate also in time. Consider the following parabolic PDE

\[
 \frac{\partial u^\varepsilon}{\partial t} - \frac{\partial}{\partial x_i} \left( a_{ij} \left( \frac{x_i}{\varepsilon}, \frac{t}{\varepsilon^2} \right) \frac{\partial u^\varepsilon}{\partial x_j} \right) = f(x, t) \quad \text{in} \quad \Omega \times (0, T), \tag{2.57a}
\]

\(^{13}\)To be more precise, Fredholm’s alternative provides us only with the homogenized equation, not the initial and boundary conditions. This is related to the discussion in subsection 2.5.2: initial and boundary conditions become somewhat irrelevant in the homogenization procedure.
28
CHAPTER 2. THE METHOD OF MULTIPLE SCALES: FORMAL ASYMPTOTICS

\[ u^\varepsilon(x,t) = 0 \text{ on } \partial \Omega \times (0,T), \]
\[ u^\varepsilon(x,0) = u_{in}(x) \text{ in } \Omega. \]  

Now we take the coefficients \( \{a_{ij}(y,\tau)\}_{i,j=1}^d \) to be 1–periodic in both \( y \) and \( \tau \). The fact that we have fast oscillations in both space and time means that we have to introduce two fast variables: \( y = \frac{x}{\varepsilon} \) and \( \tau = \frac{t}{\varepsilon^2} \). Now we have to use a two–scale expansion of the following form:

\[ u^\varepsilon(x,t) = u_0(x,\frac{x}{\varepsilon}, t, \frac{t}{\varepsilon^2}) + \epsilon u_1(x,\frac{x}{\varepsilon}, t, \frac{t}{\varepsilon^2}) + \epsilon^2 u_2(x,\frac{x}{\varepsilon}, t, \frac{t}{\varepsilon^2}) + \ldots, \]  

with \( u_j(x,y,t,\tau), j = 1, \ldots, d \) being 1–periodic in both \( y \) and \( \tau \). By treating now \( x, y \) and \( t, \tau \) as independent variables, we can write:

\[ R^\varepsilon := \frac{\partial u^\varepsilon}{\partial t} - \frac{\partial}{\partial x_i} \left( a_{ij} \left( \frac{x}{\varepsilon}, \frac{t}{\varepsilon^2} \right) \frac{\partial u^\varepsilon}{\partial x_j} \right) \]
\[ = \frac{1}{\varepsilon^2} R_0 + \frac{1}{\varepsilon} R_1 + \frac{1}{\varepsilon} R_2, \]  

where

\[ R_0 := \frac{\partial}{\partial \tau} - \frac{\partial}{\partial y_i} \left( a_{ij}(y,\tau) \frac{\partial}{\partial y_j} \right), \]
\[ R_1 := -\frac{\partial}{\partial y_i} \left( a_{ij}(y,\tau) \frac{\partial}{\partial x_j} \right) - \frac{\partial}{\partial x_i} \left( a_{ij}(y,\tau) \frac{\partial}{\partial y_j} \right), \]
\[ R_2 := \frac{\partial}{\partial t} - \frac{\partial}{\partial x_i} \left( a_{ij}(y) \frac{\partial}{\partial x_j} \right). \]  

EXERCISE 2.13. Discuss about the solvability condition (i.e. Fredholm alternative) for equation (2.61)\(^{14}\).

The analysis of the sequence of equations that we have obtained using the Fredholm alternative enables us to obtain the homogenized equation (2.56). The homogenized coefficients for this problem are given by the formula

\[ \bar{a}_{ij} = \int_Y \left( a_{ij}(y,\tau) - a_{ik}(y,\tau) \frac{\partial \theta^j(y,\tau)}{\partial y_k} \right) dy, \quad i, j = 1, \ldots, d. \]  

\(^{14}\)See [7, sec 3.10.2] for a rigorous analysis of the solvability condition for this problem.
The corrector field \( \{ \theta^j(y, \tau) \} \) satisfies the following cell problem:

\[
\mathcal{R}_0 \theta^j(y, \tau) = -\frac{\partial a_{ij}(y, \tau)}{\partial y_i} \tag{2.63a}
\]

\( \theta^j(y, \tau) \) is \( 1 \)-periodic in \( y, \tau \). \( \tag{2.63b} \)

We emphasize the fact that the cell problem (2.63) is not an evolution problem: we have periodic boundary conditions in both space and time and \( \tau \) plays the role of a space–like variable.

**EXERCISE 2.14.** Use the expansion (2.58) in order to obtain the homogenized equation (2.56) and the cell problem (2.63).

**EXERCISE 2.15.** Why do you think we have set the period of oscillations in time to be \( \epsilon^2 \), whereas the period of oscillations in space is \( \epsilon \)? (Hint: what is the order of the highest derivative in time? what is the order of the highest derivative in space?). Carry out the homogenization analysis based on the method of multiple scales for the cases where the coefficients are of the form \( \{ a_{ij}(\xi, \frac{1}{\epsilon}) \} \) and \( \{ a_{ij}(\xi, \frac{\tau}{\epsilon}) \} \) \( \epsilon \)-periodic in \( \xi, \tau \).

Similarly, one can also study the problem of homogenization for hyperbolic (wave) equations:

\[
\frac{\partial^2 u^\epsilon}{\partial t^2} - \frac{\partial}{\partial x_i} \left( a_{ij} \left( \frac{x}{\epsilon} \right) \frac{\partial u^\epsilon}{\partial x_j} \right) = f(x, t) \text{ in } \Omega \times (0, T), \tag{2.64a}
\]

\[
u(x, t) = 0 \text{ on } \partial \Omega \times (0, T) \tag{2.64b}
\]

\[
u(x, 0) = u_{\text{in}}(x) \text{ in } \Omega. \tag{2.64c}
\]

\[
\frac{\partial u^\epsilon}{\partial t}(x, 0) = v_{\text{in}}(x) \text{ in } \Omega. \tag{2.64d}
\]

In this case the homogenized equation is

\[
\frac{\partial^2 u}{\partial t^2} - \frac{\partial}{\partial x_i} \left( a_{ij} \frac{\partial u}{\partial x_j} \right) = f(x, t) \text{ in } \Omega \times (0, T), \tag{2.65a}
\]

\[
u(x, t) = 0 \text{ on } \partial \Omega \times (0, T) \tag{2.65b}
\]

\[
u(x, 0) = u_{\text{in}}(x) \text{ in } \Omega. \tag{2.65c}
\]

\[
\frac{\partial u}{\partial t}(x, 0) = v_{\text{in}}(x) \text{ in } \Omega. \tag{2.65d}
\]

The homogenized coefficients as well as the cell problem as the same as in the case of the Dirichlet problem, formulas (2.22) and (2.17) respectively.

\footnote{See [7, ch.3] for further details on the derivation of the homogenized equations using the method of multiple scales. In chapter 6 we will revisit this problem using the method of two–scale convergence.}
EXERCISE 2.16. Use the method of multiple scales to derive the homogenized equation (2.65).

EXERCISE 2.17. Consider problem (2.64) with space–time dependent coefficients \( \{a_{ij}(z, \frac{1}{\epsilon})\}_{i,j=1}^d \). What do you think the "natural" exponent \( p \) is? Derive the homogenized equation for this exponent using an appropriate multiple scales expansion.
Chapter 3

Elliptic Boundary Value Problems

3.1 Introduction

In chapter 2 we used the methods of multiple scales in order to derive the homogenized equation, together with the cell problem and the formula for the homogenized coefficients for the the Dirichlet problem (2.1). Further, we discussed about the applicability of this method to various other PDE, such as elliptic PDE with different boundary conditions and evolution equations. The asymptotic technique that we introduced led naturally to the use of the solvability condition (Fredholm alternative) for second order, uniformly elliptic partial differential operators in divergence form with periodic boundary conditions.

All the calculations of the previous chapter were formal in character and were based on the assumption that the coefficients $A = \{a_{ij}\}_{i,j=1}^d$, the forcing term $f(x)$ and the solution of the PDE $u^\epsilon(x)$ are smooth functions. As we have already discussed, the assumption of smoothness for the data of the problem $A$ and $f$, which leads to the smoothness of the solution, is not a realistic one and has to be removed. It is necessary, therefore, to develop and existence and uniqueness theory for elliptic partial differential equations with non smooth coefficients, and, furthermore, to prove that the partial differential operators of under investigation satisfy the Fredholm alternative.

Furthermore, no justification of the validity of the method of multiple scales has been presented. The principal hypothesis of this method is that the solution of partial differential equations with rapidly oscillating coefficients depend explicitly on the fast as well and the slow scales which are present in the problem. Despite the fact that this assumption is physically reasonable, there is no reason a priori why it should hold true. Consequently, we have to prove that that the solution of the original boundary value problem converges in some appropriate sense, as $\epsilon$ tends to 0, to the solution of the homogenized PDE.
In order to address the aforementioned issues we will need to use various tools from linear functional analysis. It will prove necessary for the subsequent analysis to recast our PDE as an abstract equation in the appropriate Hilbert space. This will enable us to develop rather painlessly the appropriate existence and uniqueness theory and to prove the Fredholm alternative. In the following chapters energy estimates and appropriate choices of test functions will enable us to prove the homogenization theorem and various extensions.

To be more specific, in this and chapter we will try to answer the following questions.

(i) How do we define the solution of the Dirichlet problem when the coefficients $A$ are not smooth but merely bounded and uniformly elliptic?

(ii) How do we define the solution of elliptic PDE with periodic boundary conditions of the form $A_0u = f$ with $A_0$ being defined in (2.6a)? What conditions should $f$ satisfy so that this equation is well posed? How can we develop a similar solvability theory for more general elliptic partial differential operators, i.e. with lower order terms?

The material that we will develop in this chapter will also provide us with background material which will enable us to study in later chapters the following issues.

(i) How do we compare between solutions of the Dirichlet problem (2.1) and of the homogenized equation (2.21)? In particular, how do we prove that, as $\epsilon \to 0$, we have that $u^\epsilon \to u$? In what sense do we understand this convergence?

(ii) It is intuitively clear that if we as we add more terms in the multiple scales expansion we get a better approximation to the solution $u^\epsilon$ of (2.1). How do we make this intuition precise?

Our analysis will be based upon some standard tools from linear functional analysis. Section 3.2 provided a very brief overview all necessary tools. These techniques are then used to study elliptic PDE with Dirichlet and periodic boundary conditions in sections 3.3 and 3.4, respectively. The Fredholm alternative for general divergence form uniformly elliptic operators with periodic boundary conditions is stated and proved in section 3.5.

### 3.2 Background Material

In this section we put together some definitions and theorems, without proofs, from linear functional analysis and the theory of Sobolev spaces. Only results that will be needed later on in these
notes are included. Our presentation is influenced by [19, ch. 5, ch.6] and [49, ch. 5]. A comprehensive treatment of the theory of Sobolev spaces can be found in [1]. The reader is assumed to be familiar with the basic properties of Hilbert and Banach spaces. Various facts about the space of square integrable functions will also be used. reader is already familiar with basic Hilbert space theory and with basic properties of $L^p$ The reader is referred to e.g. [33, 20, 10, 54] for material on these issues.

In the following $\Omega$ denotes a bounded open subset of $\mathbb{R}^d$. We do not make any assumptions on the regularity of the boundary $\partial \Omega$ at this point. We will use the standard notation for $L^p$ spaces. We will denote by $C_0^\infty(\Omega)$ the space of infinitely differentiable functions from $\Omega$ to $\mathbb{R}$ with compact support.

The rigorous study of homogenization leads naturally to partial differential equations with non-smooth coefficients. It should be intuitively clear that that the classical definition of a solution, i.e. of a $C^2(\Omega) \cap C(\overline{\Omega})$ function which solves (2.1a) and vanishes on the boundary, is too strong and has to be weakened. In particular, the classical definition of differentiability is not appropriate for our problem. The analysis of PDE with non-smooth coefficients requires the introduction of the concept of a \textit{weak derivative}. This will lead naturally to the definition of a \textit{weak solution}. Weak solutions for boundary value problems of the form (2.1) are elements of functions spaces of weakly differentiable functions, which turn to have a Banach or, for the problems that we will consider, a Hilbert space structure.

We start with the definition of the weak derivative.

**DEFINITION 3.1.** Let $u, v \in L^1_{\text{loc}}(\Omega)$. We say that $v$ is the \textit{first weak derivative} of $u$ with respect to $x_i$ if

$$
\int_{\Omega} u \frac{\partial \phi}{\partial x_i} \, dx = - \int_{\Omega} v \phi \, dx \quad \forall \phi \in C_0^\infty(\Omega).
$$

**REMARK 3.2.** In the above definition of the weak derivative no differentiation of the function is involved: the trick is to base the definition of differentiability on the integration by parts formula—against a smooth function—rather than on the limit of difference quotients.

We will use the notation $\frac{\partial u}{\partial x_i}$ to denote the weak derivative with respect to $x_i$. We will also use $\nabla u = \frac{\partial u}{\partial x_i} \hat{e}_i$ where $\{\hat{e}_i\}_{i=1}^d$ denotes the standard basis in $\mathbb{R}^d$. Of course, we can define weak derivatives of higher orders and we can verify that they satisfy all the rules of standard differentiation. Now we are ready to define the basic function space that we will need for our analysis.
DEFINITION 3.3. The Sobolev space $H^1(\Omega)$ consists of all square integrable functions from $\Omega$ to $\mathbb{R}$ whose first order weak derivatives exist and they are square integrable:

$$H^1(\Omega) = \left\{ u \big| u \in L^2(\Omega), \nabla u \in L^2(\Omega)^d \right\}.$$ 

Now we list without proof some properties of the space $H^1(\Omega)$.

(i) ($H^1(\Omega)$ as a function space) $H^1(\Omega)$ is a separable Hilbert space with norm

$$\|u\|_{H^1(\Omega)} = \left( \|u\|_{L^2(\Omega)}^2 + \|\nabla u\|_{L^2(\Omega)}^2 \right)^{\frac{1}{2}}$$

and inner product

$$(u, v)_{H^1(\Omega)} = (u, v)_{L^2(\Omega)} + (\nabla u, \nabla v)_{L^2(\Omega)}.$$ 

In writing the above equations we used the notation $(\nabla u, \nabla v)_{L^2(\Omega)} = \int_{\Omega} \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_j} \, dx$ with $\|\nabla u\|_{L^2(\Omega)} = \sqrt{(\nabla u, \nabla u)_{L^2(\Omega)}}$.

(ii) (Rellich compactness Theorem). The embedding of $H^1(\Omega)$ into $L^2(\Omega)$ is compact. This implies, in particular, that from every bounded sequence in $H^1(\Omega)$ we can extract a subsequence which is strongly convergent in $L^2(\Omega)$.

(iii) The space of infinitely differentiable functions $C^\infty(\Omega)$ is dense in $H^1(\Omega)$. In fact, we can define $H^1(\Omega)$ as the closure of $C^\infty(\Omega)$ under the $H^1$–norm [33, sec. 7.5].

In the next two sections we will define weak solutions of the PDE under investigation which are elements of appropriate subsets of $H^1(\Omega)$. Of course, since we are dealing with boundary value problems the solutions that we define will have to satisfy the given boundary conditions in some appropriate sense. Since functions in $H^1(\Omega)$ are defined up to sets of measure zero and the boundary of the domain $\Omega$ has measure zero in $\mathbb{R}^d$, the definition of boundary conditions for functions in $H^1(\Omega)$ requires some thought\(^1\). This is accomplished through the definition of the trace operator:

THEOREM 3.4. (The trace theorem). Assume $\Omega$ is bounded and $\partial \Omega$ is Lipschitz continuous. Then there exists a bounded linear operator $T : H^1(\Omega) \to L^2(\partial \Omega)$ such that:

(i) $Tu = u\mid_{\partial \Omega}$ $\forall u \in H^1(\Omega) \cap C(\overline{\Omega}).$

(ii) $\|Tu\|_{L^2(\partial \Omega)} \leq C\|u\|_{H^1(\Omega)}$ $\forall u \in H^1(\Omega).$

\(^1\) Apart from the case $d = 1$ since then $H^1(\Omega) \subset C^0(\overline{\Omega})$—see e.g. [49, Thm 5.3.1]—and we can define the value on the boundary in the classical sense.
We remark that, if \( \partial \Omega \) is Lipchitz continuous, we can define a surface measure on \( \partial \Omega \) and thus the space \( L^2(\partial \Omega) \) is well defined.

Now, the trace operator \( T \) is not onto \( L^2(\partial \Omega) \): there exist functions in \( L^2(\partial \Omega) \) which are not traces of elements of \( H^1(\Omega) \). In particular, we have the following definition:

**DEFINITION 3.5.** Let \( \partial \Omega \) be Lipschitz continuous. We define the space \( H^{1/2}(\partial \Omega) \) to be the range of the trace operator \( T 
\)
\[
H^{1/2}(\partial \Omega) = \{ f \in L^2(\Omega) | Tu = f, \ u \in H^1(\Omega) \}.
\]

The space \( H^{1/2}(\partial \Omega) \) is a Banach space for the norm
\[
\| u \|_{H^{1/2}(\partial \Omega)} = \int_{\partial \Omega} |u(x)|^2 \, ds_x + \int_{\partial \Omega} \int_{\partial \Omega} \frac{|u(x) - u(y)|^2}{|x - y|^{d+1}} \, ds_x \, ds_y.
\]

Moreover, when \( \partial \Omega \) is Lipschitz continuous, we can control the \( H^{1/2}(\partial \Omega) \)-norm of the trace of an element of \( H^1(\Omega) \) in terms of its \( H^1(\Omega) \)-norm; that is, there exists a constant \( C \) such that
\[
\| Tu \|_{H^{1/2}(\partial \Omega)} \leq C \| u \|_{H^1(\Omega)}, \quad \forall \ u \in H^1(\Omega).
\]

We are mostly interested in studying Dirichlet problems with homogeneous boundary conditions and, naturally, we would like to consider elements of Sobolev spaces which vanish on the boundary.

For the analysis of these boundary value problems we will need the following subset of \( H^1(\Omega) \):

**DEFINITION 3.6.** The Sobolev space \( H^1_0(\Omega) \) is defined as the closure of \( C_0^\infty(\Omega) \) with respect to the \( H^1 \)-norm.

The space \( H^1_0(\Omega) \) consists of elements of \( H^1(\Omega) \) whose trace vanishes on the boundary:

**THEOREM 3.7.** Let \( \partial \Omega \) be Lipschitz continuous. Then
\[
H^1_0(\Omega) = \{ u \in H^1(\Omega) | Tu = 0 \}.
\]

A very important property of \( H^1_0(\Omega) \) is the fact that we can control the \( L^2 \) norm of its elements in terms of the \( L^2 \)-norm of their gradient. Below we present this result, together with its analogue for elements of \( H^1(\Omega) \).

**THEOREM 3.8.** (i) (Poincaré inequality) Let \( \Omega \) be a bounded open set in \( \mathbb{R}^d \). Then there is a constant \( C_\Omega \) such that
\[
\| u \|_{L^2(\Omega)} \leq C_\Omega \| \nabla u \|_{L^2(\Omega)}.
\] (3.8)

for every \( u \in H^1_0(\Omega) \).
(ii) (Poincare–Wirtinger inequality) Let $\Omega$ be a bounded open set in $\mathbb{R}^d$ with $C^1$ boundary $\partial \Omega$. Then there exists a constant $C_{\Omega}$ such that

$$
\left\| u - \int_{\Omega} u(y) \, dy \right\|_{L^2(\Omega)} \leq C_{\Omega} \| \nabla u \|_{L^2(\Omega)}
$$

(3.9)

for every $u \in H^1(\Omega)$.

We refer to e.g. [19, sec. 5.8] for a proof of this theorem.

An immediate corollary of the first part of the above theorem is that $\| \nabla \cdot \|_{L^2(\Omega)}$ defines an equivalent norm in $H^1_0(\Omega)$:

**COROLLARY 3.9.** The quantity

$$
\| u \|_{H^1_0(\Omega)} = \| \nabla u \|_{L^2(\Omega)}
$$

(3.10)

defines a norm in $H^1_0(\Omega)$ which is equivalent to the $H^1$–norm.

This is the norm that we will use when studying Dirichlet problems with homogeneous boundary conditions.

Another space that we will need in our studies is the dual of $H^1_0(\Omega)$, that is, the space of bounded linear functionals over $H^1_0(\Omega)$. We will use the notation $H^{-1}(\Omega)$ for $(H^1_0(\Omega))^*$. Further, we will denote by $\langle \cdot, \cdot \rangle$ the pairing between $H^{-1}(\Omega)$ and $H^1(\Omega)^2$. We remark that $H^{-1}(\Omega)$ is a Banach space equipped with the norm

$$
\| f \|_{H^{-1}(\Omega)} = \sup_{\| v \|_{H^1(\Omega)}} | \langle f, v \rangle |.
$$

We can give an explicit characterization of $H^{-1}(\Omega)$ in terms of functions in $L^2(\Omega)$:

**THEOREM 3.10.** Let $f \in H^{-1}(\Omega)$. Then there exist functions $\{ f_j \}_{j=0}^d \in L^2(\Omega)$ such that

$$
\langle f, v \rangle = (f_0, v)_{L^2(\Omega)} + \sum_{j=1}^d \left( f_k, \frac{\partial v}{\partial x_j} \right)_{L^2(\Omega)} \quad \forall v \in H^1_0(\Omega),
$$

(3.12)

Moreover, we have:

$$
\| f \|_{H^{-1}(\Omega)}^2 = \inf_{\| f_j \|_{L^2(\Omega)}} \sum_{j=0}^d \| f_j \|_{L^2(\Omega)}^2.
$$

(3.13)

The infimum is taken over all functions $\{ f_j \}_{j=0}^d \in L^2(\Omega)$ for which (3.12) holds.

---

$^2$In other words, the action of $f \in H^{-1}(\Omega)$ on $v \in H^1(\Omega)$ will be denoted by $\langle f, v \rangle$. 
In the remaining of this chapter we will see that the space \( H^1(\Omega) \) and its subsets (for example the space \( H^1_0(\Omega) \) or spaces of periodic functions) are the appropriate spaces in which to look for weak solutions of boundary value problems for second order elliptic PDE. Once we recast our PDE problems into the appropriate Hilbert space framework, then existence and uniqueness of solutions follows from the following abstract existence and uniqueness theorem.

**Theorem 3.11.** (Lax–Milgram). Let \( H \) be a Hilbert space with norm \( \| \cdot \| \) and inner product \( \langle \cdot, \cdot \rangle \). Further, let \( \langle \cdot, \cdot \rangle \) denote the pairing between \( H^* \) and \( H \). Let \( a : H \times H \to \mathbb{R} \) be a bilinear mapping which satisfies the following properties:

1. **(Coercivity)** There exists a constant \( \alpha > 0 \) such that
   \[
   a[u, u] \geq \alpha\|u\|^2 \quad \forall u \in H.
   \]
2. **(Continuity)** There exists a constant \( \beta > 0 \) such that
   \[
   a[u, v] \leq \beta\|u\|\|v\| \quad \forall u, v \in H.
   \]

Let now \( f : H \to \mathbb{R} \) be a bounded linear functional on \( H \). Then there exists a unique element \( u \in H \) such that

\[
a[u, v] = \langle f, v \rangle
\]

for all \( v \in H \).

### 3.3 Dirichlet Boundary Conditions

Now we are ready to prove existence and uniqueness of weak solution for the Dirichlet problem with homogeneous boundary conditions. We first need to define the class of coefficients \( A = \{a_{ij}(x)\}_{i,j=1}^d \) of the partial differential operators that we will be concerned with.

**Definition 3.12.** Let \( \alpha, \beta \in \mathbb{R} \), such that \( 0 < \alpha < \beta \). We define \( M(\alpha, \beta, \Omega) \) to be the set of \( d \times d \) matrices \( A = \{a_{ij}\}_{i,j=1}^d \in (L^\infty(\Omega))^{d \times d} \) such that

1. \( a_{ij} \xi_i \xi_j \geq \alpha|\xi|^2 \),
2. \( |a_{ij} \xi_j| \leq \beta|\xi| \).

Further, we define \( M_{\text{per}}(\alpha, \beta, Y) \) to be the set of matrices in \( M(\alpha, \beta, \Omega) \) with \( Y \)-periodic coefficients.
Since we are interested in studying PDE whose coefficients are not differentiable, we will need to formulate the equations in a way that involves no derivatives of these coefficients and only weak derivatives of the solution. The idea behind the weak formulation of the Dirichlet problem and the concept of a weak solution is to multiply our equation by a test function and then integrate by parts.

**DEFINITION 3.13.** Consider the Dirichlet problem

\[
- \frac{\partial}{\partial x_i} \left( a_{ij}(x) \frac{\partial u}{\partial x_j} \right) = f, \quad \text{for } x \in \Omega \tag{3.17a}
\]

\[
u(x) = 0, \quad \text{for } x \in \partial \Omega, \tag{3.17b}
\]

with \( A = \{a_{ij}\}_{i,j=1}^d \in M(\alpha, \beta, \Omega), f \in H^{-1}(\Omega) \).

(i) The bilinear form associated with the Dirichlet problem (3.17) is

\[
a(\phi, \psi) = \int_{\Omega} a_{ij}(x) \frac{\partial \phi}{\partial x_j} \frac{\partial \psi}{\partial x_i} \, dx, \tag{3.18}
\]

for \( \phi, \psi \in H^1_0(\Omega) \).

(ii) We say that \( u \) is a weak solution of the Dirichlet problem (3.17) if

\[
a[u, \phi] = \langle f, \phi \rangle, \tag{3.19}
\]

for all \( \phi \in H^1_0(\Omega) \), where \( \langle \cdot, \cdot \rangle \) denotes the pairing between \( H^{-1}(\Omega) \) and \( H^1_0(\Omega) \).

Notice that the above definition of a weak solution for the Dirichlet problem involves only first order weak derivatives of our solution and no derivatives of the coefficients \( A \). Moreover, the right hand side of the equation enters in a weak, ”average” sense, after having been integrated against a test function. It looks like (3.19) is the right formulation for the Dirichlet problem (3.17). Of course, in order for this to be the case we need to prove existence and uniqueness of solutions of (3.19). This is the context of the next theorem.

**THEOREM 3.14.** The Dirichlet problem (3.17) with \( A \in M(\alpha, \beta, \Omega) \) and \( f \in H^{-1}(\Omega) \) has a unique weak solution \( u \in H^1_0(\Omega) \). Moreover, the following estimate holds:

\[
\|u\|_{H^1_0(\Omega)} \leq \frac{1}{\alpha} \|f\|_{H^{-1}(\Omega)}. \tag{3.20}
\]
3.3. Dirichlet Boundary Conditions

Proof. We have to verify the conditions of the Lax–Milgram Lemma. We start with coercivity. We use the positive definiteness of the matrix $A$ to obtain:

$$a[u, u] = \int_{\Omega} a_{ij}(x) \frac{\partial u}{\partial x_i} \frac{\partial u}{\partial x_j} \, dx \geq \alpha \int_{\Omega} |\nabla u|^2 \, dx = \alpha \|u\|^2_{H^1_0(\Omega)}.$$ 

In the above we used the fact that $\|\nabla \cdot \|_{L^2(\Omega)}$ defines an equivalent norm in $H^1_0(\Omega)$.

Now we proceed with continuity. We use the $L^\infty$ bound on the coefficients $\{a_{ij}(x)\}_{i,j=1}^d$, together with the Cauchy–Schwarz inequality to estimate:

$$a[u, v] = \int_{\Omega} a_{ij}(x) \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_j} \, dx \leq \beta \int_{\Omega} |\nabla u||\nabla v| \, dx \leq \beta \|\nabla u\|_{L^2(\Omega)} \|\nabla v\|_{L^2(\Omega)} = \beta \|u\|_{H^1_0(\Omega)} \|v\|_{H^1_0(\Omega)}.$$ 

Let us now prove estimate (3.20). We have:

$$\alpha \|u\|^2_{H^1_0(\Omega)} \leq a[u, v] = \langle f, u \rangle \leq \|f\|_{H^{-1}(\Omega)} \|u\|_{H^1_0(\Omega)},$$

from which the estimate follows. \hfill \Box

**Exercise 3.15.** Consider the problem (3.17) with $A \in M(\alpha, \beta, \Omega)$ and $f \in L^2(\Omega)$. Write down the weak formulation of the problem. Prove that in this case estimate (3.20) becomes

$$\|u\|_{H^1_0(\Omega)} \leq \frac{C_\Omega}{\alpha} \|f\|_{L^2(\Omega)},$$

where $C_\Omega$ is the Poincaré constant for the domain $\Omega$ defined in Theorem 3.8.

**Remark 3.16.** In the case where the data of the problem is regular enough so that the Dirichlet problem (3.17) admits a classical solution (i.e. a function $u \in C^2(\Omega) \cap C^0(\overline{\Omega})$ satisfying (3.17)), then the weak and classical solutions coincide. See e.g. [23].

**Remark 3.17.** Let $\{A^\varepsilon\}_{\varepsilon > 0}$ be a sequence of matrices parametrized by $\varepsilon$ such that $A^\varepsilon \in M(\alpha, \beta, \Omega)$ for every $\varepsilon > 0$. Consider the Dirichlet problem

$$-\frac{\partial}{\partial x_i} \left( a_{ij}^\varepsilon(x) \frac{\partial u^\varepsilon}{\partial x_j} \right) = f, \text{ for } x \in \Omega \quad (3.22a)$$
\[ u^\varepsilon(x) = 0, \text{ for } x \in \partial \Omega, \]  
(3.22b)

with \( f(x) \in H^{-1}(\Omega) \). Notice that we do not assume that the coefficients are of the form \( a^\varepsilon(x) = a_{ij}(\xi) \) with \( a_{ij}(y) \) being periodic. Estimate (3.20) enables us to conclude that

\[ \|u^\varepsilon\|_{H^1_0(\Omega)} \leq C, \]

the constant \( C \) being independent of \( \varepsilon \). This means that the sequence \( \{u^\varepsilon\}_{\varepsilon > 0} \) is uniformly bounded in \( H^1_0(\Omega) \) which implies that we can extract a subsequence \( \{u'^\varepsilon\}_{\varepsilon > 0} \) such that

\[ u'^\varepsilon \rightharpoonup u \text{ weakly in } H^1_0(\Omega), \]

for some \( u \in H^1_0(\Omega) \). From the Rellich compactness theorem we deduce that

\[ u'^\varepsilon \to u \text{ strongly in } L^2(\Omega). \]

Thus, estimate (3.20) applied to the solution of Dirichlet problem (3.22) enables us to conclude the existence of a (homogenized) limit \( u \). \(^3\) The above argument however does not provide us with any information concerning this limit, other than it is an element of \( H^1_0(\Omega) \). Characterizing this limit, that is, obtaining the homogenized equation, is a more subtle and difficult problem.

So far in this section we have been concerned with the homogeneous Dirichlet problem (3.17). For the proof of the homogenization theorem using the method of multiple scales we will need estimates on the solution of the non–homogeneous Dirichlet problem:

\[ -\frac{\partial}{\partial x_i} \left( a_{ij}(x) \frac{\partial u}{\partial x_j} \right) = f, \text{ for } x \in \Omega \]
(3.26a)

\[ u(x) = g, \text{ for } x \in \partial \Omega, \]
(3.26b)

The following theorem, which we state without proof, will be sufficient for our purposes:

**THEOREM 3.18.** Consider the non–homogeneous Dirichlet problem (3.26) with \( A \in M(\alpha, \beta, \Omega) \), \( f \in H^{-1}(\Omega) \) and \( g \in H^{1/2}(\partial \Omega) \) where \( \partial \Omega \) is Lipchitz continuous. Then the problem has a unique weak solution \( u \in H^1(\Omega) \). Moreover, the following estimate holds:

\[ \|u\|_{H^1(\Omega)} \leq C \left( \|f\|_{H^{-1}(\Omega)} + \|g\|_{H^{1/2}(\partial \Omega)} \right). \]
(3.27)

The constant \( C \) in the above estimate depends on the domain \( \Omega \) and the constants \( \alpha, \beta. \)

\(^3\)Notice, however, that the above arguments does not imply the uniqueness of the limit.
### 3.4 Periodic Boundary Conditions

The analysis of periodic homogenization for the Dirichlet problem (2.1) using the method of multiple scales is based on the study of elliptic PDE with periodic boundary conditions:

\[- \frac{\partial}{\partial y_i} \left( a_{ij}(y) \frac{\partial u}{\partial y_j} \right) = f(y), \quad u(y) \text{ is } 1\text{–periodic.} \tag{3.28}\]

Here the coefficients \( \{a_{ij}(y)\} \) and the function \( f(y) \) are 1–periodic. In order to study this problem within the framework of weak solutions of elliptic PDE we will need some basic properties of Sobolev spaces of periodic functions.

**DEFINITION 3.19.** Let \( C^\infty_{\text{per}}(Y) \) consist of all elements of \( C^\infty(\mathbb{R}^d) \) which are 1–periodic. \( H^1_{\text{per}}(Y) \) is defined to be the closure of \( C^\infty_{\text{per}}(Y) \) with respect to the \( H^1 \)-norm.

A similar definition holds for \( L^2_{\text{per}}(Y) \). Naturally, a function \( u \in H^1_{\text{per}}(Y) \) has the same trace \( Tu \) on opposite faces of \( Y \) (see Theorem 3.4).

Now we would like to define an appropriate concept of solution for elliptic PDE with periodic boundary conditions, problem (3.28). As we have already discussed in the previous chapter, in order for (3.28) to have a solution we need to assume that the right hand side of this equation averages to zero over the unit cell. That is, we need to restrict the set of functions \( f \) which appear in (3.28). Even if we do so, the solutions of this equation can be determined only up to a constant in \( y \) and uniqueness within the space \( H^1_{\text{per}}(Y) \) cannot be ensured. In fact, since Poincare inequality does not hold in \( H^1_{\text{per}}(Y) \), the bilinear form corresponding to (3.28) is not coercive and the Lax–Milgram lemma does not apply. It will be shown in Section 3.5 that this is due to the fact that the null space of the adjoint of the partial differential operator

\[ \mathcal{A} := - \frac{\partial}{\partial y_i} \left( a_{ij}(y) \frac{\partial}{\partial y_j} \right), \]

is not empty, but, rather, it consists of constant functions in \( y \).

The above considerations imply that in order to be able to prove existence and uniqueness of weak solutions of (3.28) we need to be able to identify between elements of \( H^1_{\text{per}}(Y) \) which differ by a constant. We have the following definition.

---

\(^4\)Since we are dealing with a second order linear elliptic PDE we expect that the weak formulation of this problem will only involve the first weak derivative of the solution. Thus, it is natural to expect that the appropriate function space in which to look for weak solutions should be (a subset of) \( H^1_{\text{per}}(Y) \).
DEFINITION 3.20. The quotient space

\[ W = H^1_{\text{per}}(Y) / \mathbb{R}, \]

is defined as the space of equivalence classes with respect to the relation

\[ u \simeq v \iff u - v \text{ is constant in } y \ \forall \ u, v \in H^1_{\text{per}}(Y). \]

We will denote by \( \bar{u} \in W \) the equivalence class represented by \( u \in H^1_{\text{per}}(Y) \). The above definition means that if \( u, v \in \bar{u} \in W \) then \( u, v \in H^1_{\text{per}}(Y) \) with \( u - v = c \) for some constant \( c \).

Now, the Poincaré inequality does hold for elements in \( W \). This means that we can use \( \| \nabla \cdot \| \mathcal{L}^2(Y) \) as a norm in \( W \), equivalent to the one induced by the \( H^1 \)-norm:

\[ \| \bar{u} \|_W = \| \nabla u \|_{\mathcal{L}^2(Y)} \ \forall \ u \in \bar{u} \in W. \tag{3.31} \]

Moreover, we can characterize the dual space \( W^* \) in terms of elements of \( (H^1_{\text{per}}(Y))^* \):

\[ W^* = \{ f \in (H^1_{\text{per}}(Y))^* | \langle f, 1 \rangle = 0 \} \tag{3.32} \]

with

\[ \langle F, \bar{u} \rangle_{W^*, W} = \langle F, u \rangle_{(H^1_{\text{per}}(Y))^*, (H^1_{\text{per}}(Y))} =: \langle F, u \rangle, \ \forall \ u \in \bar{u}, \forall \bar{u} \in W. \tag{3.33} \]

Now we are ready to present the weak formulation of equation (3.28).

DEFINITION 3.21. We will say that \( \bar{u} \in W \) is a weak solution of the boundary value problem (3.28) if

\[ a_1(\bar{u}, \bar{v}) = \langle f, \bar{v} \rangle_{W^*, W} \ \forall \ \bar{u} \in W, \forall \bar{v} \in W \tag{3.34} \]

with \( f \in W^* \). The bilinear form \( a_1(\cdot, \cdot) \) is defined as:

\[ a_1(\bar{u}, \bar{v}) = \int_Y a_{ij}(y) \frac{\partial u}{\partial y_i} \frac{\partial v}{\partial y_j} \, dy \ \forall \ u \in \bar{u}, \forall \bar{v} \in \bar{v}. \tag{3.35} \]

Existence and uniqueness of weak solutions to (3.28) holds within the space \( W \). Indeed, we have the following theorem.

THEOREM 3.22. The problem (3.28) with \( A \in M_{\text{per}}(\alpha, \beta, Y) \) and \( f \in W^* \) has a unique weak solution \( \bar{u} \in W \). Moreover, the following estimate holds:

\[ \| \bar{u} \|_W \leq \frac{1}{\alpha} \| f \|_{W^*}. \tag{3.36} \]
EXERCISE 3.23. Use the Lax–Milgram lemma to prove the above theorem.

REMARK 3.24. From the above theorem we have existence and uniqueness of weak solutions of (3.28) in $W$. Let $\tilde{u} \in W$ denote this unique solution. Then, every $u \in \tilde{u}$ solves equation (3.34). This means that we have infinitely many weak solutions in $H^1_{\text{per}}(Y)$ which differ by a constant in $y$. Indeed, the bilinear form (3.35) is not coercive in $H^1_{\text{per}}(Y)$ since the Poincaré inequality does not hold in this space; this is the reason for lack of uniqueness.

REMARK 3.25. Notice that the solvability condition that we used in the previous chapter is contained in the characterization of $W^*$, equation (3.32).

The above existence and uniqueness theorem might seem to be rather unsatisfactory since the uniqueness of solutions is proved up to a class of equivalent functions. We can alleviate this problem by fixing a representative from each equivalence class. A natural choice for this representative is the one which averages to zero over the unit cell. To this end, we define the space

$$H = \left\{ u \in H^1_{\text{per}}(Y) \mid \int_Y u \, dy = 0 \right\}.$$  \hfill (3.37)

Since we have fixed the average of elements of $H$ to be equal to zero, the Poincaré–Wirtinger inequality which holds in $H^1_{\text{per}}(Y)$ reduces to the Poincaré inequality and we can use

$$\|u\|_H = \|\nabla u\|_{L^2(Y)}, \quad \forall u \in H,$$  \hfill (3.38)

as the norm in $H$.

Now we can express the existence and uniqueness results of Theorem 3.22 within the space $H$. As before, we use the notation $\langle \cdot, \cdot \rangle$ for the pairing between $(H^1_{\text{per}}(Y))^*$ and $H^1_{\text{per}}(Y)$.

THEOREM 3.26. Let $f \in (H^1_{\text{per}}(Y))^*$. Then the problem

$$a_1(u, v) = \langle f, v \rangle, \quad \forall v \in H^1_{\text{per}}(Y)$$  \hfill (3.39)

has a unique solution $u \in H$ if and only if

$$\langle f, 1 \rangle = 0.$$  \hfill (3.40)

Proof. We set $v = 1 \in H^1_{\text{per}}(Y)$ to deduce

$$a_1(u, 1) = 0 = \langle f, 1 \rangle,$$
which implies that condition (3.40) is necessary. Now, calculations similar to the ones presented in the proof of Theorem 3.14 enable us to conclude that $a_1(\cdot, \cdot)$ defines a continuous, coercive, bilinear form on $W$. This implies, by Lax–Milgram lemma, that there exists a unique solution $u \in W$ of the problem

$$a_1(u, v) = \langle f, v \rangle \quad \forall v \in H.$$

Now we need to show that this is in fact true for all $v \in H^{1}_{\text{per}}(Y)$. Let $v \in H^{1}_{\text{per}}(Y)$. Then $v - \int_Y v(y) \, dy = \hat{v} \in H$ and we compute:

$$a_1(u, v) = a_1(u, \hat{v}) = \langle f, \hat{v} \rangle$$

$$= \langle f, v \rangle - \left( \int_Y v(y) \, dy \right) \langle f, 1 \rangle$$

$$= \langle f, v \rangle \quad \forall v \in H^{1}_{\text{per}}(Y).$$

Thus, the problem $a_1(u, v) = \langle f, v \rangle$ admits a unique solution in $H$ for every $v \in H^{1}_{\text{per}}(Y)$. 

In this section we gave two slightly different weak formulations of problem (3.28) and we proved the corresponding existence and uniqueness theorems. In the next chapter we will have the occasion to use both formulations. The important point, in both formulations, is the solvability condition (3.40). In the next section we will see that the results of this section are a special case of the Fredholm alternative for second order, divergence form, uniformly elliptic partial differential operators.

### 3.5 The Fredholm Alternative for Second Order Uniformly Elliptic Operators in Divergence Form

In this and the previous chapter we studied uniformly elliptic operators of the form

$$\mathcal{A}_0 = -\frac{\partial}{\partial x_i} \left( a_{ij}(x) \frac{\partial}{\partial x_j} \right),$$

with either Dirichlet or periodic boundary conditions. We saw that it is relatively straightforward to characterize the null space of $\mathcal{A}_0$ and to develop a solvability theory. The reason for this is the absence of lower order terms.

It is quite often the case in applications, however, that more general elliptic operators of the form

$$\mathcal{A} = -\frac{\partial}{\partial x_i} \left( a_{ij}(x) \frac{\partial}{\partial x_j} \right) + b_i(x) \frac{\partial}{\partial x_i} + c(x),$$

(3.45)
appears. The problem of periodic homogenization for elliptic PDE of the form

\[ -\mathcal{A}u = f, \]

together with boundary conditions relies on the solvability theory for Poisson equations of the form

\[ -\mathcal{A}u = f, \quad u(x) \text{ is } 1\text{-periodic.} \quad (3.46) \]

Clearly, the solvability theory for (3.46) is related to the properties of the null space of the operator \( \mathcal{A} \), as well as of its adjoint \( \mathcal{A}^* \).

Generally speaking, the first order term \( b_i(x) \frac{\partial}{\partial x_i} \) in \( \mathcal{A} \) represents advection and the zeroth order term \( c(x) \) represents (linear) chemical reaction. As examples of operators which fall into the general class described by (3.45) we mention the Schrödinger operator from quantum mechanics

\[ \mathcal{H} = -\Delta + V(x), \]

where \( V(x) \) is the potential\(^6\) and the advection–diffusion operator

\[ \mathcal{A} = -\Delta + v(x) \cdot \nabla, \]

where \( v(x) \) is the fluid velocity field which is divergence–free: \( \nabla \cdot v(x) = 0 \). As we will see in Chapter 7, second order elliptic operators appear as the generators of continuous time Markov processes\(^7\).

The goal of this section is to develop a solvability theory for partial differential operators of the form (3.45) with periodic boundary conditions\(^8\). Throughout this section we will make the following assumptions on the coefficients of \( \mathcal{A} \):

\begin{align*}
A(x) & = \{a_{ij}(x)\}_{i,j=1}^d \in M_{\text{per}}(\alpha, \beta, Y), \quad (3.47a) \\
b(x) & = \{b_i(x)\}_{i=1}^d \in (C^1_{\text{per}}(Y))^d, \quad (3.47b) \\
c(x) & \in L^\infty_{\text{per}}(Y). \quad (3.47c)
\end{align*}

\(^5\)Let \( \mathcal{A} : D(\mathcal{A}) \subset X \to Y \), where \( X \) and \( Y \) are Banach spaces. The null space \( \mathcal{N}(\mathcal{A}) \) is

\[ \mathcal{N}(\mathcal{A}) = \{u \in D(\mathcal{A}) : \mathcal{A}u = 0\}. \]

\(^6\)In this case the zeroth order term is not related to chemical reactions.

\(^7\)Though not in divergence form.

\(^8\)Analysis very similar to the one that we will use in this section gives solvability for Dirichlet boundary conditions.
Notice that we do need to assume that the vector $b(x)$ is continuously differentiable. The reason for this is that we will need to consider the $L^2$-adjoint operator of $A$, which involves derivatives of $b(x)$.

Of course, our analysis will be based on Fredholm theorem which we recall here for convenience. We refer the reader to e.g. [19, pp. 641–643] for a proof and further details.

**THEOREM 3.27. (Fredholm Alternative.)** Let $H$ be a Hilbert space and let $K : H \to H$ be a compact operator\(^9\). Then the following alternative holds.

1. Either the equations

   \[
   (I - K)u = f \tag{3.48a}
   \]
   \[
   (I - K^*)U = F \tag{3.48b}
   \]

   have unique solutions for every $f, F \in H$ or

2. the homogeneous equations

   \[
   (I - K)u_0 = 0, \quad (I - K^*)U_0 = 0 \tag{3.49}
   \]

   have the same number of non trivial solutions:

   \[
   \text{dim } (\mathcal{N}(I - K)) = \text{dim } (\mathcal{N}(I - K^*)),
   \]

   with

   \[
   \text{dim } (\mathcal{N}(I - K)) < \infty.
   \]

   In this case equations (3.48a) and (3.48b) have a solution if and only of

   \[(f, U_0) = 0 \quad \forall U_0 \in \mathcal{N}(I - K^*)\]

   and

   \[(F, u_0) = 0 \quad \forall u_0 \in \mathcal{N}(I - K),\]

   respectively.

**REMARK 3.28.** The above theorem holds in the case where $H$ is a normed space, see for instance [30, sec. 8.7]. However, the Fredholm alternative as stated above will be sufficient for our purposes.

---

\(^9\)A bounded operator $K : H \to H$ is called compact if it maps bounded sets into precompact ones. Equivalently, $K$ is compact if and only if for every bounded sequence $\{u_n\}_{n=1}^\infty \in H$, the sequence $\{Ku_n\}_{n=1}^\infty$ has a strongly convergent subsequence in $H$. 
Needless to say, the differential operator (3.45) is not even bounded, let alone compact and consequently Fredholm's theory does not directly apply to our problem. In order to apply Theorem 3.27 to the study of operator (3.45) we need to introduce an appropriate integral operator, the resolvent operator

$$R_\lambda(A) = (A + \lambda I)^{-1}. \quad (3.50)$$

Here $I$ stands for the identity operator on $L^2_{\text{per}}(Y)$ and $\lambda > 0$. The strategy of the proof that the operator $A$ defined in (3.45) satisfies the Fredholm alternative is to first use appropriate energy estimates to prove that $R_\lambda(A)$ is a bounded operator from $H^1_{\text{per}}(Y)$ to $L^2_{\text{per}}(Y)$, for $\lambda$ sufficiently large. The Rellich compactness theorem will enable us then to deduce that $R_\lambda(A)$ is a compact operator from $L^2_{\text{per}}(Y)$ to $L^2_{\text{per}}(Y)$.

Let us now carry out this program in detail. We start by defining the formal $L^2$–adjoint of $A$:

$$A^* = -\frac{\partial}{\partial x_j} \left( a_{ij}(x) \frac{\partial}{\partial x_i} \right) - b_i(x) \frac{\partial}{\partial x_i} + \left( c(x) - \frac{\partial b_i(x)}{\partial x_i} \right), \quad (3.51)$$

We need to define the bilinear forms associated to the operators $A$ and $A^*$ as well as the concept of a weak solution for the problems $A u = f$ and $A^* U = F$ with $f, F \in L^2_{\text{per}}(Y)$ and periodic boundary conditions.

As in the case of Definition 3.13, the bilinear form $a[\cdot, \cdot] : H^1_{\text{per}}(Y) \times H^1_{\text{per}}(Y) \to \mathbb{R}$ associated with the divergence form, uniformly elliptic operator $A$ is defined with integrating $A u$, $u \in H^1_{\text{per}}(Y)$ against a function $v \in H^1_{\text{per}}(Y)$ and integrating by parts:

$$a(u, v) = \int_Y a \frac{\partial u}{\partial x_j} \frac{\partial v}{\partial x_i} dx + \int_Y b_i \frac{\partial u}{\partial x_i} v dx + \int_Y c u v dx,$$

for $u, v \in H^1_{\text{per}}(Y)$. Similarly, the adjoint bilinear form $a^*[\cdot, \cdot] : H \times H \to \mathbb{R}$ associated with $A^*$ is

$$a^*(u, v) = \int_Y a_{ij} \frac{\partial u}{\partial x_j} \frac{\partial v}{\partial x_i} dx - \int_Y b_i \frac{\partial v}{\partial x_i} u dx + \int_Y \left( c - \frac{\partial b_i}{\partial x_i} \right) uv dx$$

$$= a[v, u],$$

for $u, v \in H^1_{\text{per}}(Y)$. The last equality follows after an integration by parts.

Let now $f, F \in L^2_{\text{per}}(Y)$. We say that $u$ and $U$ are weak solutions of the boundary value problems

$$A u = f, \quad u(x) \text{ is } 1\text{–periodic} \quad (3.52)$$
and
\[ A^* U = F, \quad U(x) \text{ is 1–periodic} \]  \hspace{1cm} (3.53)

if and only if
\[ a[u, v] = (f, v) \quad \forall v \in H^1_{\text{per}}(Y) \]  \hspace{1cm} (3.54)

and
\[ a^*[u, v] = (F, v) \quad \forall v \in H^1_{\text{per}}(Y), \]  \hspace{1cm} (3.55)

respectively. Now we are ready to state the main result of this section.

**THEOREM 3.29. ( Fredholm Alternative for divergence form uniformly elliptic operators with periodic boundary conditions. )**

1. Assume conditions (3.47). Then precisely one of the following statements holds: either
   
   a. For every \( f \in L^2_{\text{per}}(Y) \) there exists a unique solution of (3.52) or else
   
   b. There exist a weak non trivial solution of the homogeneous problem
   \[ A u = 0, \quad u(x) \text{ is 1–periodic}. \]  \hspace{1cm} (3.56)

2. Furthermore, should assertion b hold, the dimension of the subspace \( N \in H^1_{\text{per}}(Y) \) of weak solutions of (3.56) is finite and equals the dimension of the subspace \( N^* \in H^1_{\text{per}}(Y) \) of
   \[ A^* U = 0, \quad U(x) \text{ is 1–periodic}. \]

3. Finally, the BVP (3.52) has a weak solution if and only if
   \[ (f, U) = 0, \quad \forall U \in N^*. \]

For the proof of the theorem we will need two lemmas. In the first lemma we obtain some energy estimates which are necessary for the proof of Theorem 3.29. In the second we study some properties of the (inverse of the) resolvent operator defined in (3.50). In order to simplify the notation we will denote by \( \| \cdot \|_0 \) and \( \| \cdot \|_1 \) the norms in \( L^2_{\text{per}}(Y) \) and \( H^1_{\text{per}}(Y) \), respectively. The \( L^2 \)–inner product is denoted by \( (\cdot, \cdot) \) and the pairing between \( H^1_{\text{per}}(Y) \) and its dual by \( \langle \cdot, \cdot \rangle \).

Notice that since we look for solutions in the whole space \( H^1_{\text{per}}(Y) \), and not in its subspace \( H \) defined in (3.37), we cannot use Poincaré inequality.
**Lemma 3.30.** Assume conditions (3.47). Then there exists constants \( \nu, \mu > 0 \) such that

\[
|a[u, v]| \leq \nu \|u\|_1 \|v\|_1,
\]

and

\[
\frac{\alpha}{2} \|u\|_1^2 \leq a[u, u] + \mu \|u\|_0^2,
\]

for all \( u, v \in H^1_{\text{per}}(Y) \).

**Proof.**

1. We use the \( L^\infty \) bounds on the coefficients \( A, b \) and \( c \), together with the Cauchy–Schwarz inequality to deduce:

\[
|a(u, v)| \leq \left| \int_Y A \nabla u \nabla v \, dx + \int_Y b \nabla u v \, dx + \int_Y c uv \, dx \right|
\]

\[
\leq \|A\|_{L^\infty} \int_Y |\nabla u| |\nabla v| \, dx + \|b\|_{L^\infty} \int_Y |\nabla u| |v| \, dx + \|c\|_{L^\infty} \int_Y |u| |v| \, dx
\]

\[
\leq C (\|\nabla u\|_0 \|\nabla v\|_0 + \|\nabla u\|_0 \|v\|_0 + \|u\|_0 \|v\|_0 + \|c\|_{L^\infty} \|u\|_0^2)
\]

\[
\leq C \|u\|_1 \|v\|_1.
\]

2. We use now the uniform ellipticity of \( A \) to compute:

\[
\alpha \|\nabla u\|_0^2 \leq \int_Y (\nabla u)^T A \nabla u \, dx
\]

\[
= a[u, u] - \int_Y \nabla u u \, dx + \int_Y c u^2 \, dx
\]

\[
\leq a[u, u] + \|b\|_{L^\infty} \int_Y |\nabla u| |u| \, dx + \|c\|_{L^\infty} \|u\|_0^2.
\]

Now we make use of the Cauchy inequality with \( \delta \)

\[
a b \leq \delta a^2 + \frac{1}{4\delta} b^2;
\]

to the second term on the right hand side of (3.58) to obtain

\[
\int_Y |\nabla u| |u| \, dx \leq \delta \|\nabla u\|_0^2 + \frac{1}{4\delta} \|u\|_0^2.
\]

We chose \( \delta \) so that

\[
\alpha - \|b\|_{L^\infty} \delta = \frac{\alpha}{2}.
\]

---

\(^{10} \)Proof:

\[
(\sqrt{2\delta}a - \frac{1}{\sqrt{2\delta}}b)^2 \geq 0.
\]
We use inequality (3.58) with \( \delta \) chosen as above in (3.58) to obtain
\[
\frac{\alpha}{2} \| \nabla u \|^2_0 \leq a[u,u] + \left( \frac{1}{4\delta} \| b \|_{L^\infty} + \| c \|_{L^\infty} \right) \| u \|^2_0.
\]
We add now \( \frac{\alpha}{2} \| u \|^2_0 \) on both sides of the above inequality to obtain
\[
\frac{\alpha}{2} \| u \|^2_0 \leq a[u,u] + \mu \| u \|^2_0,
\]
with
\[
\mu = \frac{1}{2\alpha} \| b \|_{L^\infty} + \| c \|_{L^\infty} + \frac{\alpha}{2}.
\]

**Lemma 3.31.** Assume conditions (3.47). Take \( \mu \) from Lemma 3.30. Then for every \( \lambda \geq \mu \) and each function \( f \in L^2_{\text{per}}(Y) \) there exists a unique weak solution \( u \in H^1_{\text{per}}(Y) \) of the problem
\[
(\mathcal{A} + \lambda I)u = f, \quad u(x) \text{ is } 1\text{-periodic.} \tag{3.59}
\]

**Proof.** Let \( \lambda \geq \mu \). Define the operator
\[
\mathcal{A}_\lambda := \mathcal{A} + \lambda I. \tag{3.60}
\]
The bilinear form associated to \( \mathcal{A}_\lambda \) is
\[
a_\lambda[u,v] = a[u,v] + \lambda(u,v) \quad \forall u, v \in H^1_{\text{per}}(Y). \tag{3.61}
\]
Now, Lemma 3.30 together with our assumption that \( \lambda \geq \mu \) imply that the bilinear form \( a_\lambda[u,v] \) is continuous and coercive on \( H^1_{\text{per}}(Y) \). Hence the Lax–Milgram theorem applies\(^1\) and we deduce the existence and uniqueness of a solution \( u \in H^1_{\text{per}}(Y) \) of the equation
\[
a_\lambda[u,v] = (f,v), \quad \forall v \in H^1_{\text{per}}(Y). \tag{3.62}
\]
This is precisely the weak formulation of the boundary value problem (3.59). \( \square \)

**Proof of Theorem 3.29.** 1. By Lemma 3.31 there exists, for every \( g \in L^2_{\text{per}}(Y) \), a unique solution \( u \in H^1_{\text{per}}(Y) \) of
\[
a_\mu[u,v] = (g,v), \quad \forall v \in H^1_{\text{per}}(Y). \tag{3.63}
\]
We use the resolvent operator defined in (3.50) to write the solution of (3.63) in the following form:
\[
u = R_A(\mu)g. \tag{3.64}
\]
\(^1\)We have that \( \langle f,v \rangle = (f,v) \) defines a bounded linear functional on \( H^1_{\text{per}}(Y) \).
Consider now equation (3.52). We add the term $\mu u$ on both sides of this equation to obtain

$$\mathcal{A}_\mu u = \mu u + f,$$

where $\mathcal{A}_\mu$ is defined in (3.60). The weak formulation of this equation is

$$a_\lambda[u, v] = (\lambda u + f, v), \quad \forall v \in H^1_{\text{per}}(Y). \quad (3.65)$$

We can rewrite this as an integral equation (see (3.64))\(^{12}\)

$$u = R_A(\lambda)(\mu u + f),$$

or, equivalently,

$$(I - K)u = h,$$

where

$$K := \mu R_A(\lambda), \quad h = R_A(\lambda)f.$$  

2. Now we claim that the operator $K : L^2_{\text{per}}(Y) \to L^2_{\text{per}}(Y)$ is compact. Indeed, let $u$ be the solution of (3.63) which is given by (3.64). We use the second estimate in Lemma 3.30, the definition of the bilinear form (3.61) and the Cauchy–Schwarz inequality in (3.63) to obtain

$$\frac{\alpha}{2} \|u\|^2 \leq a_\mu[u, u] = (g, u) \leq \|g\|_0 \|\mu u\|_1.$$  

Consequently,

$$\|u\|_1 \leq C\|g\|_0.$$  

We use now (3.64), the definition of $K$ and the above estimate to deduce that

$$\mu \|u\|_1 = \|Kg\|_1 \leq C\mu \|g\|_0. \quad (3.66)$$

By the Rellich compactness theorem $H^1_{\text{per}}(Y)$ is compactly embedded in $L^2_{\text{per}}(Y)$ and consequently estimate (3.66) implies that $K$ maps bounded sets in $L^2_{\text{per}}(Y)$ into compact ones in $L^2_{\text{per}}(Y)$. Hence, it is a compact operator.

3. We apply now the Fredholm alternative (Theorem 3.27) to the operator $K$: either

\(^{12}\)Since the resolvent is the inverse of a differential operator, it is intuitively clear that it is an integral operator. The kernel of this operator is the Green function associated to the problem at hand, see e.g. [24, ch. 7] for further information.
a. there exists a unique \( u \in L^2_{\text{per}}(Y) \) such that

\[
(I - K)u = h,
\]  
(3.67)

or

b. there exists a non trivial solution \( u \in L^2_{\text{per}}(Y) \) of the homogenous equation

\[
(I - K)u = 0.
\]  
(3.68)

Let us assume that (3.67) holds. From the preceding analysis we deduce that there exists a unique weak solution \( u \in H^1_{\text{per}}(Y) \) of (3.52). Assume now that (3.68) holds. Let \( N \) and \( N^* \) denote the dimensions of null spaces of \( I - K \) and \( I - K^* \), respectively. From Theorem 3.27 we know that \( N = N^* \). Moreover, it is straightforward to prove (check!) that \( u \in \mathcal{N}(I - K) = 0 \iff a[u, \phi] = 0, \ \forall u \in H^1_{\text{per}}(Y) \) and \( v \in \mathcal{N}(I - K^*) = 0 \iff a^*[v, \phi] = 0, \ \forall \phi \in H^1_{\text{per}}(Y) \), Thus, the Fredholm alternative for \( K \) implies the Fredholm alternative for \( \mathcal{A}^* \).  

4. Now we prove the third part of the theorem. Let \( v \in \mathcal{N}(I - K^*) \). By Theorem 3.27 we know that (3.68) has a solution if and only if

\[
(h, v) = 0 \ \forall v \in \mathcal{N}(I - K^*).
\]

We compute

\[
(h, v) = (R_\mathcal{A}(\mu)f, v) = \frac{1}{\mu} (Kf, v) = \frac{1}{\mu} (f, K^*v) = \frac{1}{\mu} (f, v).
\]

Hence, problem (3.52) has a weak solution if and only if

\[
(f, v) = 0 \ \forall v \in \mathcal{N}(\mathcal{A}^*).
\]

This completes the proof of the theorem. \( \square \)

In order to apply Theorem 3.29 to a specific homogenization problem we will need to characterize the null space of the adjoint operator \( \mathcal{A}^* \). In particular, we will need to prove that it is one–dimensional. We will see in Chapter 7 that this is intimately related to the ergodic theory of Markov processes.

\[^{13}\text{Of course, within the context of weak solutions.} \]
EXERCISE 3.32. Use Theorem 3.29 to derive the results of the previous section. (Hint: What is the adjoint $A_0^*$ of operator $A_0$ defined in (3.44)? What is the null space of $A_0$ and $A_0^*$? ).

EXERCISE 3.33. Prove the Fredholm alternative for operator $A$ defined in (3.45) under assumptions (3.47) for Dirichlet boundary conditions.
Chapter 4

The Homogenization Theorem

4.1 Introduction

In this chapter we prove the homogenization theorem for the Dirichlet problem with homogeneous boundary conditions and periodic coefficients:

\begin{align}
- \frac{\partial}{\partial x_i} \left( a_{ij}^\varepsilon(x) \frac{\partial u^\varepsilon}{\partial x_j} \right) &= f, \quad \text{for } x \in \Omega \quad (4.1a) \\
u^\varepsilon(x) &= 0, \quad \text{for } x \in \partial \Omega. \quad (4.1b)
\end{align}

In chapter 2 we derived the homogenized equation

\begin{align}
- \overline{a}_{ij} \frac{\partial u}{\partial x_i \partial x_j} &= f, \quad \text{for } x \in \Omega \quad (4.2a) \\
u(x) &= 0, \quad \text{for } x \in \partial \Omega, \quad (4.2b)
\end{align}

together with the formula for the homogenized coefficients:

\[ \overline{a}_{ij} = \int_Y \left( a_{ij}(y) - a_{ik}(y) \frac{\partial \chi^k(y)}{\partial y_k} \right) dy, \quad i, j = 1, \ldots, d. \quad (4.3) \]

The correction fields \( \chi^k(y), k = 1, \ldots d \) satisfy the cell problem

\[ \frac{\partial}{\partial y_i} \left( a_{ij}(y) \frac{\partial \chi^k}{\partial y_j} \right) = - \frac{\partial a_{ik}}{\partial y_i}, \quad \chi^k(y) \text{ is 1–periodic, } k = 1, \ldots, d. \quad (4.4) \]

Moreover, we obtained the following two–scale expansion for \( u^\varepsilon(x) \):

\[ u^\varepsilon(x) \approx u(x) - \varepsilon \chi^i \left( \frac{x}{\varepsilon} \right) \frac{\partial u}{\partial x_i} - \varepsilon^2 \theta^{ij} \left( \frac{x}{\varepsilon} \right) \frac{\partial^2 u}{\partial x_i \partial x_j} + \ldots \quad (4.5) \]

where the fields \( \theta^{k\ell}(y), k, \ell = 1, \ldots d \) satisfy the higher order cell problem

\[ \frac{\partial}{\partial y_i} \left( a_{ij}(y) \frac{\partial \theta^{k\ell}}{\partial y_j} \right) = -b_{k\ell}, \quad \theta^{k\ell}(y) \text{ is 1–periodic, } k, \ell = 1, \ldots, d, \quad (4.6) \]
with
\[ b_{ij}(y) = -\bar{a}_{ij} + a_{ij}(y) - a_{ik}(y) \frac{\partial \chi^j(y)}{\partial y_k} - \frac{\partial (a_{kj}(y) \chi^{(d)}(y))}{\partial y_k}, \quad i, j = 1, \ldots, d. \]  

Our goal in this chapter is to prove that \( u^\varepsilon \) which solves (4.1) converges to \( u \) which solves (4.2), weakly in \( H^1_0(\Omega) \) and to justify the expansion (4.5). We will prove the convergence of \( u^\varepsilon \) to \( u \) using Tartar’s method of oscillating test functions. The proof is based on the construction of appropriate test functions which enable us to replace products of weakly convergent sequences with products of a weakly and a strongly convergent sequence. Thus, Tartar’s method enables us to pass to the limit using a ”compensated compactness” argument. The statement and proof of the basic convergence theorem will be presented in section 4.2.

On the other hand, the justification of the two–scale expansion (4.5) consists of two parts: first we prove that the first three terms in the expansion are well defined and hence the expression in (4.5) makes sense. The proof of this results is quite is based on our analysis of elliptic boundary value problems in the previous chapter, together with standard results from elliptic regularity theory. Then, we estimate the difference between \( u^\varepsilon(x) \) and the first three terms in the expansion. This is accomplished by a bootstrapping argument [43, 42]. The idea is to prove that the remainder term satisfies an inhomogeneous Dirichlet boundary value problem with ”small” data and to then use estimate (3.27) to bound the remainder.

In our proofs we will need the weak convergence properties of sequences of rapidly oscillating periodic functions. In particular, we will make use of the following result, whose proof can be found in e.g. [12, ch. 2].

**Theorem 4.1.** Let \( 1 \leq p \leq +\infty \) and \( f \) be a \( Y \)–periodic function in \( L^p(Y) \). Set
\[ f_\varepsilon(x) = f \left( \frac{x}{\varepsilon} \right) \quad \text{a.e. on } \mathbb{R}^d. \]

Then, if \( p < +\infty \), as \( \varepsilon \to 0 \)
\[ f_\varepsilon \rightharpoonup \int_Y f(y) \, dy \text{ weakly in } L^p(\Omega), \]
for any bounded open subset \( \Omega \) of \( \mathbb{R}^d \).

If \( p = +\infty \) we have
\[ f_\varepsilon \rightharpoonup^* \int_Y f(y) \, dy \text{ weakly--* in } L^\infty(\mathbb{R}^d). \]
4.2 The Homogenization Theorem I: Tartar’s Method of Oscillating Test Functions

In this section we will prove the following theorem:

**THEOREM 4.2.** Let $u^\varepsilon$ be the solution of (4.1) with $f \in H^{-1}(\Omega)$ and $A^\varepsilon(x) = A \left( \frac{x}{\varepsilon} \right)$, $A(y) \in \mathcal{M}_{\text{per}}(\alpha, \beta, \gamma)$. Further, let $u$ be the solution of the homogenized problem (4.2) with $\overline{A}$ given by (4.3) and $\chi^k(y)$, $k = 1, \ldots, d$ satisfy (4.4). Then

(i) $u^\varepsilon \rightharpoonup u$ weakly in $H^1_0(\Omega)$.

(ii) $A^\varepsilon \nabla u^\varepsilon \rightharpoonup \overline{A} \nabla u$ weakly in $(L^2(\Omega))^d$.

**Proof.** 1. In the previous chapter we proved the following estimate for the solution of the Dirichlet problem (4.1):

$$\|u^\varepsilon\|_{H^1_0(\Omega)} \leq C.$$  

We combine this with the $L^\infty$-bound on the coefficients of the matrix $A^\varepsilon$ to obtain:

$$\|\xi^\varepsilon\|_{(L^2(\Omega))^d} := \|A^\varepsilon \nabla u^\varepsilon\|_{(L^2(\Omega))^d} \leq \beta\|u^\varepsilon\|_{H^1_0(\Omega)} \leq C.$$  

The above two estimates imply that we can extract subsequences, still denoted by $u^\varepsilon$, $\xi^\varepsilon$ such that:

$$u^\varepsilon \rightharpoonup u \quad \text{weakly in } H^1_0(\Omega) \quad \text{(4.8a)}$$

$$\xi^\varepsilon \rightharpoonup \xi \quad \text{weakly in } (L^2(\Omega))^d \quad \text{(4.8b)}$$

Using the notation $\xi^\varepsilon := A^\varepsilon \nabla u^\varepsilon$ we can write the weak form of the Dirichlet problem (4.1) as follows

$$\int_{\Omega} \xi^\varepsilon_i \frac{\partial v}{\partial x_i} \, dx = \langle f, v \rangle, \quad \forall v \in H^1_0(\Omega). \quad \text{(4.9)}$$

We can pass to the limit in (4.9) using (4.8b) to obtain:

$$\int_{\Omega} \xi_i \frac{\partial v}{\partial x_i} \, dx = \langle f, v \rangle, \quad \forall v \in H^1_0(\Omega). \quad \text{(4.10)}$$

Thus, to complete the proof of the theorem we need to show that $\xi = \overline{A} \nabla u$. In order to do this we need to construct appropriate test functions. We will accomplish this by using the *adjoint cell problem*. 
2. The formal $L^2$–adjoint operator of $A_0$ is:

$$A_0^* := - \frac{\partial}{\partial y_i} \left( a_{ji}(y) \frac{\partial}{\partial y_j} \right).$$

Let now $\lambda \in \mathbb{R}^d$ and let $\chi^\lambda(y) \in H^1_{\text{per}}(\Gamma)$ be the unique, mean zero solution of the PDE

$$A_0^* \chi^\lambda = - \frac{\partial a_{ji}(y)}{\partial y_i} \lambda_j.$$

(4.12)

Define now

$$w^\epsilon_\lambda(x) = \lambda \cdot x - \epsilon \chi^\lambda_\epsilon(x),$$

with $\chi^\lambda_\epsilon(x) = \chi^{\lambda \left( \frac{x}{\epsilon} \right)}$. We clearly have that $\|w^\epsilon_\lambda(x)\|_{H^1(\Omega)} \leq C$. We use this estimate, together with Theorem 4.1 to deduce that

$$w^\epsilon_\lambda \rightharpoonup \lambda \cdot x \quad \text{weakly in } H^1_0(\Omega),$$

(4.13)

which, on account of the Rellich compactness theorem, implies

$$w^\epsilon_\lambda \to \lambda \cdot x \quad \text{strongly in } L^2(\Omega).$$

(4.14)

Moreover, it is easy to check that $w^\epsilon_\lambda$ satisfies the following PDE:

$$(A^\epsilon)^* w^\epsilon_\lambda = 0,$$

(4.15)

with

$$(A^\epsilon)^* := - \frac{\partial}{\partial x_i} \left( a_{ji}^\epsilon(x) \frac{\partial}{\partial x_j} \right).$$

Consider now the weak formulation of the Dirichlet problem (4.1) with a test function of the form $v = w^\epsilon_\lambda \phi$ with $\phi \in C_0^\infty(\Omega)$:

$$\int_{\Omega} a_{ij}^\epsilon \frac{\partial u^\epsilon}{\partial x_i} \frac{\partial w^\epsilon_\lambda}{\partial x_j} \phi \, dx + \int_{\Omega} \frac{\partial u^\epsilon}{\partial x_i} \frac{\partial \phi}{\partial x_i} w^\epsilon_\lambda \, dx = \langle f, w^\epsilon_\lambda \phi \rangle.$$

(4.17)

Now we multiply (4.15) by $\phi u^\epsilon$, $\phi \in C_0^\infty(\Omega)$, integrate over $\Omega$, integrate by parts and substract the result from (4.17) to obtain:

$$\int_{\Omega} a_{ij}^\epsilon \frac{\partial u^\epsilon}{\partial x_j} \frac{\partial \phi}{\partial x_i} w^\epsilon_\lambda \, dx - \int_{\Omega} a_{ij}^\epsilon \frac{\partial u^\epsilon}{\partial x_j} \frac{\partial \phi}{\partial x_i} w^\epsilon_\lambda \, dx = \langle f, w^\epsilon_\lambda \phi \rangle.$$

(4.18)

Notice that none of the terms in (4.18) involves products of two weakly convergent sequences in $L^2(\Omega)$. This is exactly how Tartar’s method works: by choosing our test functions appropriately and by subtracting the appropriate expression from the weak formulation of the Dirichlet problem
(4.17) we are able to obtain an equation which does not involve any "problematic" terms. Thus we can actually pass to the limit in (4.18) and obtain the homogenized equation.

3. The integrand in the first term on the left hand side of (4.18) is the product of a strongly and a weakly convergent sequence in $L^2(\Omega)$. Thus, we can use (4.8b) and (4.14) to pass to the limit:

$$
\int_{\Omega} a_{ij}^\epsilon \frac{\partial w^\epsilon}{\partial x_j} \frac{\partial \phi}{\partial x_i} \ dx \to \int_{\Omega} \xi_i \frac{\partial \phi}{\partial x_i} \lambda \cdot x \ dx, \quad \forall \phi \in C_0^\infty(\Omega).
$$

Let us consider now the second term on the left hand side of (4.18). We notice that

$$
a_{ij}^\epsilon \frac{\partial u^\epsilon}{\partial x_j} = a_{ij}^\epsilon \lambda_j - \frac{\partial \chi^\epsilon_j}{\partial x_j} \to \int_{\Omega} \left( a_{ij}(y) - \frac{\partial \chi_j^*(y)}{\partial y_k} \right) \lambda_j \ dy \text{ weakly in } L^2(\Omega)
$$

(4.20)

We have used the notation $\chi^*_j$ to denote the solutions of the adjoint cell problem. We also used Theorem 4.1. Now, since the integrand of the second term on the left hand side of equation is a product of as strongly and a weakly convergence sequence, we can pass to the limit as $\epsilon \to 0$ to obtain:

$$
\int_{\Omega} a_{ij}^\epsilon u^\epsilon \frac{\partial w^\epsilon}{\partial x_j} \frac{\partial \phi}{\partial x_i} \ dx \to \int_{\Omega} \bar{a}_{ij} \lambda_j \frac{\partial \phi}{\partial x_i} \ dx, \quad \forall \phi \in C_0^\infty(\Omega).
$$

Moreover:

$$
\langle f, w^\epsilon \phi \rangle \to \langle f, \lambda \cdot x \phi \rangle, \quad \forall \phi \in C_0^\infty(\Omega).
$$

We combine the above results to obtain the limiting equation:

$$
\int_{\Omega} \xi_i \frac{\partial \phi}{\partial x_i} \lambda \cdot x \ dx - \int_{\Omega} \bar{a}_{ij} \lambda_j \frac{\partial \phi}{\partial x_i} \ dx = \langle f, \lambda \cdot x \phi \rangle. \quad (4.23)
$$

Consider now equation (4.10) with $v = \lambda \cdot x \phi, \phi \in C_0^\infty(\Omega)$:

$$
\int_{\Omega} \xi_i \frac{\partial \phi}{\partial x_i} \lambda \cdot x \ dx + \int_{\Omega} \xi_j \lambda_j \phi \ dx = \langle f, \lambda \cdot x \phi \rangle. \quad (4.24)
$$

We solve for the first term on the left hand side of the above equation and substitute into (4.23) to obtain:

$$
\int_{\Omega} \xi_j \lambda_j \phi \ dx = - \int_{\Omega} \bar{a}_{ij} \lambda_j \phi \ dx \\
= \int_{\Omega} \bar{a}_{ij} \lambda_j \frac{\partial u}{\partial x_i} \phi \ dx.
$$

Now, since the above equation holds for every $\phi \in C_0^\infty(\Omega)$, we can conclude that

$$
\xi = \bar{A} \nabla u.
$$

This completes the proof of the theorem.
REMARK 4.3. The above theorem is optimal in the sense that the assumption on the data of the problems are limited to what is necessary in order to have existence and uniqueness of solutions of the equation for $u^c(x)$, the homogenized equation and the cell problem; hence, none of the assumptions of the theorem can be relaxed. Notice moreover that the presence of oscillations in the problem implies that one cannot hope to get strong convergence in $H^1_0(\Omega)$.

REMARK 4.4. For the case of periodic homogenization part (ii) of the above theorem follows from part (i), on account of Theorem 4.1. However, this is not true when the coefficients $A^c$ are not periodic and in this case part (ii) has to be proved separately.

REMARK 4.5. The reason why we had to consider the adjoint cell problem in the proof of Theorem 4.2 was because the we did not take the matrix $A$ to be symmetric. In the case where $A = A^T$ the above proof is considerably simplified. The difference between the symmetric and the non–periodic case will become more profound in the non–periodic setting.

4.3 The Homogenization Theorem II: The Method of Multiple Scales

Now we turn our attention to the rigorous justification of the tw–scale expansion (4.5). We start by observing that the existence and uniqueness results for elliptic PDE with periodic boundary conditions, together with the assumptions on $A(y)$ and $f(x)$ enable us to conclude that the cell problem (4.4) admits a unique solution $\chi^k(y) \in \hat{W}$, $k = 1, \ldots, d$, where the space $\hat{W}$ is defined in (3.37). Similarly, there exist unique functions $\theta^{k\ell}(Y) \in \hat{W}, k, \ell = 1, \ldots d$ that solve (4.6).

Continuing in the same manner, we can prove existence and uniqueness of all higher order cell problems and compute the corresponding terms in the two–scale expansion:

$$u^c(x) \approx \sum_{k=0}^{n} \epsilon^k u_k \left(x, \frac{x}{\epsilon}\right).$$

(4.25)

Now, in order for the two–scale expansion (4.25) to be well defined we need to ensure that all higher order derivatives of the solution $u$ of the homogenized equation which appear in the expansion exist in $L^2(\Omega)$. In particular, since the $k$th term of the two–scale expansion involves the $k - 1$ order derivatives of $u$, the solution of the homogenized equation should belong at least to $H^{k-1}_0(\Omega)^1$. Now, elliptic regularity theory [23, Thm. 8.13] ensures that the solution of the homog-
enized equation will have the desired regularity provided that \( f \in H^{k-3}(\Omega) \), \( k \geq 2 \). In particular, for expansion (4.5) to be well defined in the \( L^2 \)-sense we need to assume that \( f \in L^2(\Omega) \). Notice that the above considerations do not involve any regularity assumptions on the coefficients \( A^\varepsilon(x) \).

The above considerations enable us to rigorously justify the fact that the solution \( u^\varepsilon(x) \) of the Dirichlet problem (4.1) admits a two–scale expansion of the form (4.25) provided that the data of the problem are sufficiently regular. With a little more work we can also estimate the difference between \( u^\varepsilon(x) \) and the two–scale expansion in \( H^1(\Omega) \). In this section we will prove an estimate of this form for the expansion (4.5). We have the following theorem.

**Theorem 4.6.** Let \( u^\varepsilon \) be the solution of (4.1) with \( f \in L^2(\Omega) \) and \( A^\varepsilon(x) = A\left(\frac{x}{\varepsilon}\right), A(y) \in M_{\text{per}}(\alpha, \beta, Y) \). Further, let \( u \) be the solution of the homogenized problem (4.2) with \( \overline{A} \) given by (4.3) and \( \chi^k(y), \theta^k(y) k, \ell = 1, \ldots, d \) satisfy (4.4) and (4.6). Then \( u^\varepsilon \) admits the asymptotic expansion (4.5). Moreover, if \( f \in C^\infty(\Omega), \partial \Omega \) is of class \( C^\infty \) and \( \chi^k(y), \theta^k(y) \in W^{1,\infty}(Y) k, \ell = 1, \ldots, d, \) then we have the following estimate:

\[
\left\| u^\varepsilon(x) - \left( u(x) - \varepsilon \chi^j \left( \frac{x}{\varepsilon} \right) \frac{\partial u}{\partial x_j} - \varepsilon^2 \theta^j \left( \frac{x}{\varepsilon} \right) \frac{\partial^2 u}{\partial x_i \partial x_j} \right) \right\|_{H^1(\Omega)} \leq C \varepsilon^{\frac{1}{2}}.
\]

(4.26)

**Proof.** The proof that the \( u^\varepsilon(x) \) admits the asymptotic expansion (4.5) is based on the discussion preceding the statement of the theorem. Its details are left as an exercise. Let us now prove estimate (4.26). Let \( R^\varepsilon \) denote the difference between \( u^\varepsilon \) and \( u + u_1 + \varepsilon^2 u_2 \):

\[
R^\varepsilon(x) = u^\varepsilon(x) - \left( u(x) + \varepsilon u_1 \left( x, \frac{x}{\varepsilon} \right) + \varepsilon^2 u_2 \left( x, \frac{x}{\varepsilon} \right) \right).
\]

The terms \( u_1 \) and \( u_2 \) are defined in (4.5). We apply the operator

\[
A^\varepsilon = -\frac{\partial}{\partial x_i} \left( a_{ij} \left( \frac{x}{\varepsilon} \frac{\partial}{\partial x_j} \right) \right),
\]

to both sides of the above equation and use the decomposition (2.5) with \( A_0, A_1, A_2 \) being given by (2.6) to obtain:

\[
A^\varepsilon R^\varepsilon = A^\varepsilon u^\varepsilon - A^\varepsilon \left( u + \varepsilon u_1 + \varepsilon^2 u_2 \right)
= f - \left[ \frac{1}{\varepsilon^2} A_0 u_0 + \frac{1}{\varepsilon} (A_0 u_1 + A_1 u_0) + (A_0 u_2 + A_1 u_1 + A_2 u_0) + \varepsilon (A_1 u_2 + A_2 u_1) + \varepsilon^2 A_2 u_2 \right]
= -\frac{1}{\varepsilon^2} A_0 u_0 - \frac{1}{\varepsilon} (A_0 u_1 + A_1 u_0) - (A_0 u_2 + A_1 u_1 + A_2 u_0 - f) - \varepsilon (A_1 u_2 + A_2 u_1) - \varepsilon^2 A_2 u_2
= -\varepsilon (A_1 u_2 + A_2 u_1) - \varepsilon^2 A_2 u_2
:= f^\varepsilon(x).
\]

(4.27)
In the above calculation we used equations (2.9). Moreover, we have:

\[ R^\varepsilon(x)|_{x \in \partial \Omega} = u^\varepsilon(x)|_{x \in \partial \Omega} - \left( u(x) + \varepsilon u_1 \left( x, \frac{x}{\varepsilon} \right) + \varepsilon^2 u_2 \left( x, \frac{x}{\varepsilon} \right) \right)|_{x \in \partial \Omega} \]
\[ = \left( \varepsilon \chi^j \left( \frac{x}{\varepsilon} \right) \frac{\partial u}{\partial x_j} + \varepsilon^2 \theta^jr \left( x, \frac{x}{\varepsilon} \right) \frac{\partial^2 u}{\partial x_i \partial x_j} \right)|_{x \in \partial \Omega} \]
\[ := \varepsilon g^\varepsilon|_{x \in \partial \Omega}. \quad (4.28) \]

From the above considerations we conclude that the remainder \( R^\varepsilon \) satisfies the following inhomogeneous Dirichlet boundary value problem:

\[ -\frac{\partial}{\partial x_i} \left( a^i_{ij}(x) \frac{\partial R^\varepsilon}{\partial x_j} \right) = f^\varepsilon, \quad \text{for } x \in \Omega \quad (4.29a) \]
\[ R^\varepsilon(x) = \varepsilon g^\varepsilon, \quad \text{for } x \in \partial \Omega. \quad (4.29b) \]

Now, from Lemmas 4.7 and 4.8 we have that there exists a constant \( C \) independent of \( \varepsilon \) such that

\[ \| f^\varepsilon \|_{H^{-1}(\Omega)} \leq C \quad (4.30) \]

and

\[ \| g^\varepsilon \|_{H^\frac{1}{2}(\partial \Omega)} \leq C \varepsilon^{\frac{1}{2}}. \quad (4.31) \]

Thus, estimate (3.27) applies and we have:

\[ \| R^\varepsilon(x) \|_{H^1(\Omega)} \leq C \left( \varepsilon \| f^\varepsilon \|_{H^{-1}(\Omega)} + \varepsilon \| g^\varepsilon \|_{H^\frac{1}{2}(\partial \Omega)} \right) \]
\[ \leq C_1 \varepsilon + C_2 \varepsilon^{\frac{1}{2}}, \quad (4.32) \]

from which estimate (4.26) follows. \( \square \)

In order to conclude the proof of Theorem 4.6 we need to prove estimates (4.30) and (4.31).

**Lemma 4.7.** Let \( f \in C^\infty(\Omega) \) and \( A(y) \in M_{\text{per}}(\alpha, \beta, Y) \). Consider \( f^\varepsilon \) defined in (4.27). We have that \( f^\varepsilon \in H^{-1}(\Omega) \) and moreover there exists a constant \( C \) independent of \( \varepsilon \) such that \( \| f^\varepsilon \|_{H^{-1}(\Omega)} \leq C \).

**Proof.** 1. We compute:

\[ \mathcal{A}_2 u_1(x, y) = a_{ij}(y)\chi^k(y) \frac{\partial^3 u}{\partial x_i \partial x_j \partial x_k}. \]

Moreover:

\[ \mathcal{A}_2 u_2(x, y) = a_{ij}(y)\chi^{kt}(y) \frac{\partial^4 u}{\partial x_i \partial x_j \partial x_k \partial x_l}. \]
Furthermore:

$$A_1 u_2(x, y) = a_{ij}(y) \frac{\partial \theta^k(y)}{\partial y_j} \frac{\partial^3 u}{\partial x_i \partial x_k \partial x_\ell} + \frac{\partial}{\partial y_j} \left( a_{ji}(y) \frac{\partial \theta^k(y)}{\partial y_j} \right) \frac{\partial^3 u}{\partial x_i \partial x_k \partial x_\ell}.$$ 

2. Let $$\phi^\varepsilon(x) = \phi\left(\frac{x}{\varepsilon}\right)$$. From the chain rule we have:

$$\frac{\partial}{\partial y_j} \phi^\varepsilon(y) = \varepsilon \frac{\partial \phi\left(\frac{x}{\varepsilon}\right)}{\partial x_j} := \varepsilon \frac{\partial \phi^\varepsilon(x)}{\partial x_j}.$$

Now we have:

$$\left. \frac{\partial}{\partial y_j} \left( a_{ji}(y) \frac{\partial \theta^k(y)}{\partial y_j} \right) \right|_{y = \frac{x}{\varepsilon}} \frac{\partial^3 u}{\partial x_i \partial x_k \partial x_\ell} = \varepsilon \frac{\partial}{\partial x_j} \left( a_{ji} \left( \frac{x}{\varepsilon} \right) \theta^k \left( \frac{x}{\varepsilon} \right) \right) \frac{\partial^3 u}{\partial x_i \partial x_k \partial x_\ell}$$

$$= \varepsilon \frac{\partial}{\partial x_j} \left( a_{ji} \left( \frac{x}{\varepsilon} \right) \theta^k \left( \frac{x}{\varepsilon} \right) \frac{\partial^3 u}{\partial x_i \partial x_k \partial x_\ell} \right)$$

$$- \varepsilon a_{ji} \left( \frac{x}{\varepsilon} \right) \frac{\partial \theta^k \left( \frac{x}{\varepsilon} \right)}{\partial x_j} \frac{\partial^3 u}{\partial x_i \partial x_k \partial x_\ell}$$

$$= \varepsilon \frac{\partial}{\partial x_j} \left( a_{ji} \left( \frac{x}{\varepsilon} \right) \theta^k \left( \frac{x}{\varepsilon} \right) \frac{\partial^3 u}{\partial x_i \partial x_k \partial x_\ell} \right) - \varepsilon A_2 u_2 \left( x, \frac{x}{\varepsilon} \right).$$

We substitute the above expressions in the definition of $$f^\varepsilon$$ to obtain:

$$f^\varepsilon = -A_1 u_2 \left( x, \frac{x}{\varepsilon} \right) - A_2 u_1 \left( x, \frac{x}{\varepsilon} \right) - \varepsilon A_2 u_2 \left( x, \frac{x}{\varepsilon} \right)$$

$$= - \left( a_{ik} \left( \frac{x}{\varepsilon} \right) \chi^i \left( \frac{x}{\varepsilon} \right) + a_{ij} \left( \frac{x}{\varepsilon} \right) \frac{\partial \theta^k \left( \frac{x}{\varepsilon} \right)}{\partial y_j} \right) \frac{\partial^3 u}{\partial x_i \partial x_k \partial x_\ell}$$

$$- \varepsilon a_{ij} \left( \frac{x}{\varepsilon} \right) \frac{\partial \theta^k \left( \frac{x}{\varepsilon} \right)}{\partial x_j} \frac{\partial^3 u}{\partial x_i \partial x_k \partial x_\ell}$$

$$:= f_0 + \frac{\partial f_i}{\partial x_i},$$

with

$$f_0 = \left( a_{ik} \left( \frac{x}{\varepsilon} \right) \chi^i \left( \frac{x}{\varepsilon} \right) + a_{ij} \left( \frac{x}{\varepsilon} \right) \frac{\partial \theta^k \left( \frac{x}{\varepsilon} \right)}{\partial y_j} \right) \frac{\partial^3 u}{\partial x_i \partial x_k \partial x_\ell}$$

and

$$f_i = a_{ij} \left( \frac{x}{\varepsilon} \right) \theta^k \left( \frac{x}{\varepsilon} \right) \frac{\partial^3 u}{\partial x_j \partial x_k \partial x_\ell}, \quad i = 1, \ldots, d.$$
Moreover, we have:

\[ \| f_0 \|_{L^2(\Omega)} = \left\| \left( a_{ik} \left( \frac{x}{\epsilon} \right) x^k + a_{ij} \left( \frac{x}{\epsilon} \right) \frac{\partial \theta^{k\ell}(y)}{\partial y_j} \right) \frac{\partial^3 u}{\partial x_i \partial x_k \partial x_\ell} \right\|_{L^2(\Omega)} \]

\[ \leq \left\| \frac{\partial^3 u}{\partial x_i \partial x_k \partial x_\ell} \right\|_{L^\infty(\Omega)} \left\| a_{ik} \left( \frac{x}{\epsilon} \right) x^k + a_{ij} \left( \frac{x}{\epsilon} \right) \frac{\partial \theta^{k\ell}(y)}{\partial y_j} \right\|_{L^2(\Omega)} \]

\[ \leq C. \]

The uniform bound follows from the fact that, from Theorem 4.1, the sequence \( a_{ik} \left( \frac{x}{\epsilon} \right) x^k + a_{ij} \left( \frac{x}{\epsilon} \right) \frac{\partial \theta^{k\ell}(y)}{\partial y_j} \bigg|_{y=\frac{x}{\epsilon}} \) converges weakly in \( L^2(\Omega) \) and is consequently uniformly bounded in \( L^2(\Omega) \) [54, sec. V.1, Thm 1]. A similar argument shows that

\[ \| f_i \|_{L^2(\Omega)} \leq C, \quad i = 1, \ldots, d. \]

We use now definition (3.13) of the \( H^{-1} \)-norm to deduce:

\[ \| f^\epsilon \|^2_{H^{-1}(\Omega)} \leq \sum_{i=0}^{d} \| f_i \|^2_{L^2(\Omega)} \leq C. \]

This completes the proof of the lemma. \( \square \)

**Lemma 4.8.** Let \( f \in C^\infty(\Omega), \partial \Omega \) and \( \chi^k(y), \theta^{k\ell}(y) \in W^{1,\infty}(\Omega) \) \( k, \ell = 1, \ldots, d. \) Consider \( g^\epsilon \) defined in (4.28). We have that \( g^\epsilon \in H^{1/2}(\partial \Omega) \) and moreover there exists a constant \( C \) independent of \( \epsilon \) such that \( \| g^\epsilon \|^2_{H^{1/2}(\partial \Omega)} \leq C \epsilon^{-\frac{1}{2}}. \)

We refer to [12, pp. 136–137] for a proof of this lemma.

**Remark 4.9.** The regularity on the solutions of the cell problems (4.4) and (4.6) is only needed for the proof of Lemma 4.8. We also remark that an argument similar to the one used in the proof of Lemma 4.7 which is based in the smoothness of the solution of the homogenized equation would not be sufficient in this case.

**Remark 4.10.** The justification of the error estimate (4.26) requires to assume much more smoothness upon the data of the problem, in comparison to what was needed for the proof of Theorem 4.2. From this point of view Tartar’s method is provides us with a better way of proving the homogenization theorem. On the other hand, the analysis based on the two–scale expansion provides us with more informations. In particular, we can compute higher order corrections to the solution of the homogenized equation and to also obtain error estimates. These are very important issues in problems in applied science where the parameter \( \epsilon \) is small but finite.
REMARK 4.11. The proof of the error estimate (4.26) is based on the derivation of a PDE for the remainder $R^\varepsilon$ which is of the same form as the original PDE for $u^\varepsilon$ but inhomogeneous with "small" data. Energy estimates are then used to bound it. This is a quite general method and it actually works for all PDE for which energy estimates of the form (3.27) are available. From this point of view, the construction of the two–scale expansion furnishes us also with the proof of the homogenization theorem, provided that we are willing to assume enough regularity on the data of the problem.
Chapter 5

Two–Scale Convergence

5.1 Introduction

In the previous chapter we proved the homogenization theorem for the Dirichlet problem using two different methods. In section 4.2 we used Tartar’s method to construct appropriate test functions which enabled us to pass to the limit of products of weakly convergent sequences. In section 4.3 we used the two–scale expansion that we had obtained previously in order to estimate the difference between the solution of our PDE and the solution of the homogenized equation.

In this chapter we will combine these two approaches to develop an alternative homogenization procedure. The basic idea will be to consider test functions in the form of a two–scale expansion. In this way we will be able to obtain the homogenized equation and prove the convergence theorem in one step. For this we will need to define a new concept, that of two–scale convergence. This concept was introduced by Nguetseng [38, 39] and later popularized and developed further by Allaire [2, 3]. Our presentation follows closely that of [3]. Corollary 5.18 is taken from [47].

In our study of two–scale convergence we will make extensive use of functions of two arguments $x$ and $y$, say, of the form $L^p(\Omega; X)$, where $X$ is a Banach space and $\Omega$ is an open subset of $\mathbb{R}^d$, not necessarily bounded. This is also a Banach space with norm

$$
\|u\|_{L^p(\Omega; X)}^p = \int_{\Omega} \|u(x, y)\|_X^p \, dx.
$$

An important function space for our subsequent considerations is $L^2(\Omega; L^2(Y)) := L^2(\Omega \times Y)$. This is also a Banach space with norm

$$
\|u\|_{L^2(\Omega \times Y)} = \left( \int_{\Omega} \int_Y \|u(x, y)\|^2 \, dy \, dx \right)^{1/2}.
$$

We will also have the occasion to use the space $L^2(\Omega; C_{\text{per}}(Y))$, which is the set of all measurable functions $u : x \in \Omega \rightarrow u(x) \in C_{\text{per}}(Y)$ such that $\|u(x)\| \in L^2(\Omega \times Y)$. 

...
According to (5.1), the norm of this space is

\[ \|u\|^2_{L^2(\Omega; C_{\text{per}}(Y))} = \int_{\Omega} \left( \sup_{y \in Y} |u(x, y)| \right)^2 \, dy. \]

This is a separable Banach space which is dense in \( L^2(\Omega \times Y) \) [12, Thm. 3.61]. It enjoys various properties which we will need.

**Theorem 5.1.** Let \( u \in L^2(\Omega; C_{\text{per}}(Y)) \). Then

(i) \( u \left( x, \frac{\cdot}{\epsilon} \right) \in L^2(\Omega) \) with

\[ \left\| u \left( x, \frac{x}{\epsilon} \right) \right\|_{L^2(\Omega)} \leq \|u(x, y)\|_{L^2(\Omega; C_{\text{per}}(Y))}. \]

(ii) \( u \left( x, \frac{\cdot}{\epsilon} \right) \) converges to \( \int_Y u(x, y) \, dy \), weakly in \( L^2(\Omega) \).

(iii) We have

\[ \left\| u \left( x, \frac{x}{\epsilon} \right) \right\|_{L^2(\Omega)} \to \|u(x, y)\|_{L^2(\Omega \times Y)}. \]

**5.2 Two–Scale Convergence**

In this section we define two–scale convergence and study some of its basic properties. In order to do so we will need to consider appropriate test functions, the *admissible test functions*.

**Definition 5.2.** A function \( \phi(x, y) \in L^2(\Omega \times Y) \) is called an admissible test function if it satisfies

\[ \lim_{\epsilon \to 0} \int_{\Omega} \left| \phi \left( x, \frac{x}{\epsilon} \right) \right|^2 \, dx = \int_{\Omega} \int_{Y} |\phi (x, y)|^2 \, dy \, dx. \] (5.6)

**Remark 5.3.** A certain amount of regularity in either \( x \) or \( y \) is needed in order for a test function to be admissible. In particular, not all elements of \( L^2(\Omega \times Y) \) satisfy condition (5.6).\(^1\) From Theorem 5.1 we know that \( L^2(\Omega; C_{\text{per}}(Y)) \) is a set of admissible test functions. Moreover, it is straightforward to prove that functions of the form \( \phi(x, y) = \phi_1(y)\phi_2(x, y) \) with \( \phi_1(y) \in L^\infty(Y) \) and \( \phi_2(x, y) \in L^2[\Omega; C_{\text{per}}(Y)] \) are also admissible test functions.

\(^1\)A counterexample is provided in [3, Prop. 5.8]
DEFINITION 5.4. Let $u^\epsilon$ be a sequence in $L^2(\Omega)$. We will say that $u^\epsilon$ two–scale converges to $u_0(x, y) \in L^2(\Omega \times Y)$ and write $u^\epsilon \overset{\omega}{\rightharpoonup} u_0$ if for every admissible test function $\phi$ we have
\[
\lim_{\epsilon \to 0} \int_{\Omega} u^\epsilon(x) \phi \left( x, \frac{x}{\epsilon} \right) \, dx = \int_{\Omega} \int_{Y} u_0(x, y) \phi(x, y) \, dy \, dx.
\] (5.7)

Two scale convergence implies weak convergence in $L^2(\Omega)$. In particular, we have the following lemma.

LEMMA 5.5. Let $u^\epsilon$ be a sequence in $L^2(\Omega)$ which two–scale converges to $u_0(x, y) \in L^2(\Omega \times Y)$. Then
\[
u^\epsilon \rightharpoonup \nu_0(x) := \int_{Y} u_0(x, y) \, dy, \quad \text{weakly in } L^2(\Omega).
\]

Proof. Choose a test function $\phi(x) \in L^2(\Omega)$, independent of $\epsilon$. This is clearly an admissible test function and we can use it in (5.7) to deduce:
\[
\lim_{\epsilon \to 0} \int_{\Omega} u^\epsilon(x) \phi (x) \, dx = \int_{\Omega} \int_{Y} u_0(x, y) \phi(x) \, dy \, dx = \int_{\Omega} \left( \int_{Y} u_0(x, y) \, dy \right) \phi(x) \, dx = (\nu_0, \phi)_{L^2(\Omega)}.
\]
The above holds for every $\phi(x) \in L^2(\Omega)$ and, hence, $u^\epsilon \rightharpoonup \nu_0$ weakly in $L^2(\Omega)$.

An immediate consequence of the above lemma is the following.

COROLLARY 5.6. Let $u^\epsilon$ be a sequence in $L^2(\Omega)$ which two–scale converges to $u_0(x) \in L^2(\Omega)$, i.e. the two–scale limit is independent of $y$. Then the weak $L^2$–limit and the two–scale limit coincide.

In this chapter we will see that two–scale convergence is the appropriate tool for studying periodic homogenization. In particular, it enables us to rigorously justify the two–scale expansion that we have been obtaining when applying the method of multiple scales, as the next result indicates.

LEMMA 5.7. Consider a function $u^\epsilon(x) \in L^2(\Omega)$ which admits the following two–scale expansion
\[
u^\epsilon(x) = u_0 \left( x, \frac{x}{\epsilon} \right) + \epsilon u_1 \left( x, \frac{x}{\epsilon} \right) + \ldots
\]
where $u_j(x, y) \in L^2(\Omega; C_{\text{per}}(Y))$, $j = 0, 1, \ldots, N$, $\Omega$ being a bounded domain in $\mathbb{R}^d$. Then $u^\epsilon \overset{\omega}{\rightharpoonup} u_0$. 
CHAPTER 5. TWO-SCALE CONVERGENCE

Proof. It is enough to consider the case where $N = 1$. Let $\psi(x, y) \in L^2(\Omega; C_{\text{per}}(Y))$ and define $f_j(x, y) = u_j(x, y)\psi(x, y)$, $j = 0, 1$. We will use the notation $f^\epsilon(x) = f(x, x^\epsilon)$. We clearly have

$$
\int_\Omega u^\epsilon(x) \phi \left(x, \frac{x}{\epsilon}\right) \, dx = \int_\Omega f_0^\epsilon(x) \, dx + \epsilon \int_\Omega f_1^\epsilon(x) \, dx.
$$

(5.9)

Now, $f_j^\epsilon(x) \in L^2(\Omega; C_{\text{per}}(Y))$ for $j = 0, 1$. This implies, by Theorem 5.1, that $f_0^\epsilon$ converges to its average over $Y \int_Y f_0(x, y) \, dy$, weakly in $L^2(\Omega)$. We choose the test function $\phi = 1 \in L^2(\Omega)$, since $\Omega$ was assumed to be bounded:

$$
\int_\Omega f_0^\epsilon(x) \, dx \rightarrow \int_\Omega \int_Y f_0(x, y) \, dy \, dx.
$$

Let us consider now the second integral on the right hand side of (5.9). Since the sequence $f_1^\epsilon$ is weakly convergent in $L^2(\Omega)$, it is bounded [54, V.1, Thm. 1]. Thus, and using again the boundedness of $\Omega$, together with Cauchy–Schwartz inequality, we obtain:

$$
\epsilon \int_\Omega f_1^\epsilon(x) \, dx \leq \epsilon C \|f_1^\epsilon\|_{L^2(\Omega)} \leq \epsilon C \rightarrow 0.
$$

We use the above two calculations in (5.9) to obtain:

$$
\int_\Omega u^\epsilon(x) \phi \left(x, \frac{x}{\epsilon}\right) \, dx \rightarrow \int_\Omega \int_Y u_0(x, y) \phi(x, y) \, dy \, dx.
$$

Hence, $u^\epsilon$ two-scale converges to $u_0$.

The above result justifies, in some sense, the formal calculations presented in Chapter 2: the first term in the expansion $u_0(x, y)$ is the two-scale limit of the sequence $u^\epsilon$ and further, by Lemma 5.5, the average of $u_0(x, y)$ over the unit cell is the weak $L^2$–limit of the sequence.

Now, we would like to find criteria which enable to conclude that a given sequence in $L^2(\Omega)$ is two-scale convergent. The following compactness result provides us with such a criterion.

**THEOREM 5.8.** Let $u^\epsilon$ be a bounded sequence in $L^2(\Omega)$. Then there exists a subsequence, still denoted by $u^\epsilon$, and function $u_0(x, y) \in L^2(\Omega \times Y)$ such that $u^\epsilon$ two-scale converges to $u_0(x, y)$.

Proof. 1. To ease the notation we will denote by $X$ the space $L^2(\Omega; C_{\text{per}}(Y))$. Let now $\phi \in X$. From Theorem 5.1 we have

$$
\|\phi \left(x, \frac{x}{\epsilon}\right)\|_{L^2(\Omega)} \leq \|\phi(x, y)\|_X.
$$

\[\square\]
Consequently:

\[
\left| \int_{\Omega} u^\varepsilon(x) \phi \left( x, \frac{x}{\varepsilon} \right) \, dx \right| \leq \| u^\varepsilon(x) \|_{L^2(\Omega)} \| \phi \left( x, \frac{x}{\varepsilon} \right) \|_{L^2(\Omega)} \\
\leq C \| \phi(x, y) \|_X \leq C,
\]

since, by assumption, \( u^\varepsilon \) is bounded in \( L^2(\Omega) \).

2. From (5.11) we deduce that \( \int u^\varepsilon \cdot dx \) defines a bounded linear functional over \( X \). That is, we can define a \( U^\varepsilon \in X^* \) such that

\[
\langle U^\varepsilon, \phi \rangle_{X, X^*} = \int_{\Omega} u^\varepsilon(x) \phi \left( x, \frac{x}{\varepsilon} \right) \, dx
\]

for every \( \phi \in X \). From (5.11) we have that

\[
|\langle U^\varepsilon, \phi \rangle_{X, X^*}| \leq \| \phi \|_X, \quad \forall \phi \in X
\]

and consequently \( \| U^\varepsilon \|_{X^*} \leq C \). Since \( X \) is a separable Banach space, we can extract a weak--* convergent subsequence [32, Thm.12 Ch. 10], still denoted by \( U^\varepsilon \), such that

\[
U^\varepsilon \rightharpoonup U_0, \quad \text{weak--* in } X^*
\]

for some \( U_0 \in X^* \). Consequently

\[
\int_{\Omega} u^\varepsilon(x) \phi \left( x, \frac{x}{\varepsilon} \right) \, dx = \langle U^\varepsilon, \phi \rangle_{X, X^*} \rightarrow \langle U_0, \phi \rangle_{X, X^*}, \quad \forall \phi \in X.
\]

3. Equation (5.11) implies that

\[
\lim_{\varepsilon \to 0} \left| \int_{\Omega} u^\varepsilon(x) \phi \left( x, \frac{x}{\varepsilon} \right) \, dx \right| \leq C \| \phi(x, y) \|_X.
\]

We combine the above equation with (5.12) to deduce that

\[
|\langle U_0, \phi \rangle_{X, X^*}| \leq C \| \phi(x, y) \|_X, \quad \forall \phi \in X.
\]

Now, \( X \) is dense in \( L^2(\Omega \times Y) \) which implies that (5.13) actually holds for every \( \phi \in L^2(\Omega \times Y) \). Hence, \( \langle U_0, \cdot \rangle_{X, X^*} \) can be extended to become a bounded linear functional on \( L^2(\Omega \times Y) \). This is a Hilbert space and thus the Riesz representation theorem applies [54, III.6], which enables us to identify the limiting bounded linear functional by a unique element \( u_0(x, y) \) of \( L^2(\Omega \times Y) \):

\[
\langle U_0, \phi \rangle_{X, X^*} = \int_{\Omega} \int_Y u_0(x, y) \, dy \, dx, \quad \forall \phi \in X.
\]

We combine (5.14) with (5.12) to obtain

\[
\int_{\Omega} u^\varepsilon(x) \phi \left( x, \frac{x}{\varepsilon} \right) \, dx \rightarrow \int_{\Omega} \int_Y u_0(x, y) \, dy \, dx, \quad \forall \phi \in X.
\]

The theorem is proved.
**REMARK 5.9.** As has already been mentioned, the space $X = L^2(\Omega; C_{\text{per}}(Y))$ is not the only possible choice for the space of test functions that we consider. Any set of admissible test functions would do, provided that it forms a separable Banach space $X$ and that it is dense in $L^2(\Omega \times Y)$. Separability is necessary so that we can extract a weak–* convergent subsequence from a bounded sequence in $X^*$.

The two–scale convergence defined in Definition 5.4 is still a weak type of convergence, since it is defined in terms of the product of a sequence $u^\varepsilon$ with an appropriate test function. We can also define a notion of strong two–scale convergence.

**DEFINITION 5.10.** Let $u^\varepsilon$ be a sequence in $L^2(\Omega)$. We will say that $u^\varepsilon$ two–scale converges strongly to $u_0(x, y) \in L^2(\Omega \times Y)$ and write $u^\varepsilon \overset{2}{\to} u_0$ if

$$
\lim_{\varepsilon \to 0} \int_\Omega |u^\varepsilon(x)|^2 \, dx = \int_\Omega \int_Y |u_0(x, y)|^2 \, dy \, dx.
$$

(5.16)

**REMARK 5.11.** Although every strongly two–scale convergent sequence is also two–scale convergent, the converse is not true: not all two–scale convergent sequences are strongly two–scale convergent.

**REMARK 5.12.** In view of the above definition, we can define the class of admissible test functions to be the subset of elements of $L^2(\Omega \times Y)$ which are periodic in $y$ and strongly two–scale convergent.

As it is always the case with weak convergence, the limit of the product of two two–scale convergent sequences is not in general the product of the limits. However, we can pass to the limit when we one of the two sequences is strongly two–scale convergent. Moreover, a strongly two–scale convergent sequence converges to its two–scale limit strongly in $L^2(\Omega)$, provided that the limit is regular enough. Notice however that the two–scale limit will not in general possess any further regularity. In fact, every function $u_0(x, y)$ in $L^2(\Omega \times Y)$ is attained as a two–scale limit of some sequence in $L^2(\Omega)$ [3, Lem. 1.13].

**THEOREM 5.13.** (i) Let $u^\varepsilon, v^\varepsilon$ be sequences in $L^2(\Omega)$ such that $u^\varepsilon \overset{2}{\to} u_0$ and $v^\varepsilon \overset{2}{\to} v_0$. Then $u^\varepsilon v^\varepsilon \overset{2}{\to} u_0 v_0$.

(ii) Assume further that $u_0(x, y) \in L^2(\Omega; C_{\text{per}}(Y))$. Then

$$
\left\| u^\varepsilon(x) - u_0\left(x, \frac{x}{\varepsilon}\right) \right\|_{L^2(\Omega)} \to 0.
$$
We refer to [3] for a proof of this theorem.

So far we have only considered bounded sequences in $L^2(\Omega)$ whose two–scale limit is an element of $L^2(\Omega \times Y)$ and depends explicitly on $y$. It is now natural to ask whether more information on the two–scale limit can be obtained when our sequence is bounded in a stronger norm. In our study of homogenization for the Dirichlet problem, whose solution $u^\epsilon$ is bounded in $H^1_0(\Omega)$, we saw that the first term in the expansion, which by Lemma 5.7 is the two–scale limit of the solution of our PDE, is independent of $y$. This is a general result which, together with additional information, is the content of the next theorem.

THEOREM 5.14. (i) Let $u^\epsilon$ be a bounded sequence in $H^1(\Omega)$. Then $u^\epsilon$ two-scale converges to its weak–$H^1$ limit $u(x)$. Further, there exists a function $u_1(x, y) \in L^2(\Omega; H^1_{\text{per}}(Y)/\mathbb{R})$ such that, up to a subsequence, $\nabla u^\epsilon$ two–scale converges to $\nabla_x u(x) + \nabla_y u_1(x, y)$.

(ii) Let $u^\epsilon$ and $\epsilon \nabla u^\epsilon$ be uniformly bounded sequences in $L^2(\Omega)$. Then there exists a function $u_0(x, y) \in L^2(\Omega; H^1_{\text{per}}(Y))$ such that, up to a subsequence, $u^\epsilon$ and $\epsilon \nabla u^\epsilon$ two–scale converge to $u_0(x, y)$ and to $\nabla_y u_0(x, y)$, respectively.

(iii) Let $v^\epsilon$ be a divergence–free field which is bounded in $[L^2(\Omega)]^d$. Then the two–scale limit satisfies $\nabla_y \cdot v_0(x, y) = 0$ and $\int_Y \nabla_x \cdot v_0(x, y) \, dy = 0$

REMARK 5.15. Since we only know that $v_0(x, y) \in [L^2(\Omega \times Y)]^d$, we have to interpret the divergence of $v_0(x, y)$ with respect to $x$ and $y$ in the appropriate sense, namely in the $H^{-1}$ sense. In all the instances where we will have the occasion to use part (iii) of the above theorem we will consider vector fields which are smooth enough so that we will not have to worry about this issue.

In order to prove this theorem we will need the following lemma, which states that divergence free vectors are orthogonal to gradients of scalars, orthogonality being interpreted in the $L^2$–sense.

LEMMA 5.16. Let $F \in (L^2_{\text{per}}(Y))^d$ be such that $(F, G)_{L^2(Y)} := \sum_{i=1}^d (F_i, G_i)_{L^2(Y)} = 0$ for all $G \in (C^\infty_{\text{per}}(\Omega))^d$ with $\nabla \cdot G := \sum_{i=1}^d \frac{\partial G_i}{\partial x_i} = 0$. Then there exists a unique function $p \in H^1_{\text{per}}(Y)/\mathbb{R}$ such that $F = -\nabla p$.

This lemma holds in any open set $\Omega$ of $\mathbb{R}^d$. For a proof we refer to e.g. [52, 53]. The proof of this lemma for a general domain $\Omega$ is quite complicated. It is rather straightforward to prove it, however, in the periodic setting using Fourier series.
REMARK 5.17. Let us introduce the notation

\[ L^2_{\text{div}} = \{ f \in (L^2(\Omega))^d; \nabla \cdot f = 0 \} \]

and

\[ L^2_{\text{grad}} = \{ f \in (L^2(\Omega))^d; f = -\nabla p, p \in H^1_{\text{per}}(Y)/\mathbb{R} \}. \]

Then a consequence of the above theorem is that the space of square integrable periodic vectors admits the following orthogonal decomposition:

\[ [L^2(\Omega)]^d = L^2_{\text{div}} \oplus L^2_{\text{grad}}. \]  \hfill (5.17)

Proof of Theorem 5.14. 1. The assumption \[ \|u^\epsilon\|_H^1(\Omega) \leq C \] implies that there exist functions \( u_0(x, y) \in L^2(\Omega \times Y) \) and \( V \in [L^2(\Omega \times Y)]^d \) such that, up to a subsequence,

\[ u^\epsilon \rightharpoonup u_0(x, y) \quad \text{and} \quad \nabla u^\epsilon \rightharpoonup V(x, y). \] \hfill (5.18)

Let now \( \Phi \in [L^2(\Omega; C^1_{\text{per}}(Y))]^d \). We have

\[
\epsilon \int_\Omega \nabla u^\epsilon(x) \cdot \Phi \left( x, \frac{x}{\epsilon} \right) \, dx = - \int_\Omega \int_\Omega u^\epsilon \left[ \epsilon \nabla_x \Phi \left( x, \frac{x}{\epsilon} \right) + \nabla_y \Phi \left( x, \frac{x}{\epsilon} \right) \right] \, dx \\
\quad \rightarrow - \int_\Omega \int_\Omega u_0(x, y) \nabla_y \Phi(x, y) \, dy \, dx.
\]

On the other hand, \( \nabla u^\epsilon(x) \cdot \Phi \left( x, \frac{x}{\epsilon} \right) \) is bounded in \( L^2(\Omega) \) which implies that

\[ \epsilon \int_\Omega \nabla u^\epsilon(x) \cdot \Phi \left( x, \frac{x}{\epsilon} \right) \, dx \rightarrow 0 \]

and consequently

\[ \int_\Omega \int_Y u_0(x, y) \nabla_y \Phi(x, y) \, dy \, dx = 0 \quad \forall \Phi \in [L^2(\Omega; C^1_{\text{per}}(Y))]^d. \]

Consequently, the two–scale limit is independent of \( y, u_0 = u_0(x) \). Moreover, by Lemma 5.5 we conclude that \( u_0 \) is actually the (in this case strong) \( L^2 \)–limit which is the weak \( H^1 \)–limit.

2. Choose now an admissible test function \( \Phi(x, y) \) with \( \nabla_y \Phi(x, y) = 0 \). We compute

\[
\int_\Omega \nabla u^\epsilon(x) \cdot \Phi \left( x, \frac{x}{\epsilon} \right) \, dx = - \int_\Omega \int_\Omega u^\epsilon \left[ \nabla_x \Phi \left( x, \frac{x}{\epsilon} \right) + \frac{1}{\epsilon} \nabla_y \Phi \left( x, \frac{x}{\epsilon} \right) \right] \, dx \\
\quad = - \int_\Omega \int_\Omega u^\epsilon \nabla_x \Phi \left( x, \frac{x}{\epsilon} \right) + \frac{1}{\epsilon} \, dx \\
\quad \rightarrow - \int_\Omega \int_Y u_0(x) \nabla_x \Phi(x, y) \, dy \, dx.
\]
The two–scale convergence of \( \nabla u^\varepsilon \) (5.18) implies that
\[
\int_\Omega \nabla u^\varepsilon(x) \cdot \Phi \left( x, \frac{x}{\varepsilon} \right) \, dx = - \int_\Omega \int_\Omega u^\varepsilon \left[ \nabla_x \Phi \left( x, \frac{x}{\varepsilon} \right) + \frac{1}{\varepsilon} \nabla_y \Phi \left( x, \frac{x}{\varepsilon} \right) \right] \, dx \\
\to \int_\Omega \int_Y V(x, y) \cdot \Phi(x, y) \, dydx.
\]
We combine the above two equations to obtain
\[
- \int_\Omega \int_Y u_0(x) \nabla_x \Phi(x, y) \, dydx = \int_\Omega \int_Y V(x, y) \cdot \Phi(x, y) \, dydx,
\]
for all divergence free admissible test functions. An integration by parts yields
\[
\int_\Omega \int_Y (\nabla_x u_0 - V(x, y)) \cdot \Phi(x, y) \, dydx = 0.
\]
Since the test functions are taken to be divergence free, Lemma 5.16 applies and consequently there exists a function \( u_1(x, y) \in L^2(\Omega; H^1_{\text{per}}(Y)/\mathbb{R}) \) such that
\[
V(x, y) = \nabla_x u_0 + \nabla_y u_1(x, y).
\]

The proofs of parts (ii) and (iii) of the theorem are similar and they are left as an exercise. \( \square \)

Similar techniques used in the proof of Theorem 5.14 enable us to obtain information on the two–scale limit in the case where the control we have over the \( L^2 \)–norm of the gradient of \( u^\varepsilon \) is in between that assumed in parts (i) and (ii) of the theorem.

**COROLLARY 5.18.** Let \( u^\varepsilon \) and \( \varepsilon^\gamma \nabla u^\varepsilon \) be bounded sequences in \( L^2(\Omega) \) with \( \gamma \in (0, 1) \). Then the two–scale limit \( u_0 \) of \( u^\varepsilon \) is independent of \( y \) and is the weak–\( L^2 \) limit of \( u^\varepsilon \). Moreover, there exists a function \( u_1(x, y) \) in \( L^2(\Omega; H^1_{\text{per}}(Y)/\mathbb{R}) \) such that \( \varepsilon^\gamma \nabla u^\varepsilon \) two–scale converges to \( \nabla_y u_1(x, y) \).

**Proof.** From part (ii) of Theorem 5.14 we have that \( \varepsilon u^\varepsilon \rightharpoonup \nabla_y u_0(x, y) \). On the other hand, the bound on \( \varepsilon^\gamma \nabla u^\varepsilon , \gamma \in (0, 1) \) implies that \( \varepsilon \nabla u^\varepsilon \rightharpoonup 0 \). Consequently, \( \nabla_y u_0(x, y) = 0 \) (in the weak sense), and \( u_0 = u_0(x) \). The fact that \( u_0 \) is the weak \( L^2 \) limit of \( u^\varepsilon \) follows from Theorem 5.14.

Now, the bound on \( \varepsilon^\gamma \nabla u^\varepsilon \) implies that, for every admissible test function \( \Phi \) we have:
\[
\lim_{\varepsilon \to 0} \varepsilon^\gamma \int_\Omega \nabla u^\varepsilon(x) \cdot \Phi \left( x, \frac{x}{\varepsilon} \right) \, dx = \int_\Omega \int_Y V(x, y)\Phi(x, y) \, dydx.
\]
We now consider divergence free–in \( y \)– admissible test functions \( \Phi \). We integrate by parts to obtain:
\[
\lim_{\varepsilon \to 0} \varepsilon^\gamma \int_\Omega \nabla u^\varepsilon(x) \cdot \Phi \left( x, \frac{x}{\varepsilon} \right) \, dx = \lim_{\varepsilon \to 0} \left( -\varepsilon^\gamma \int_\Omega u^\varepsilon(x) \nabla \cdot \Phi \left( x, \frac{x}{\varepsilon} \right) \, dx \right) \\
= \lim_{\varepsilon \to 0} \left( -\varepsilon^\gamma \int_\Omega u^\varepsilon(x) \nabla_x \cdot \Phi \left( x, \frac{x}{\varepsilon} \right) \, dx \right) \\
= 0,
\]
(5.21)
on account of the bound on $u^\epsilon$. Combining now the above two equations we deduce:

$$\int_{\Omega} \int_{\gamma} V(x, y) \cdot \Phi(x, y) \, dy \, dx = 0$$

(5.22)

for all test functions with $\nabla_y \cdot \Phi = 0$. We use now Lemma 5.16 to conclude that there exists a function $u_1(x, y) \in L^2(\Omega; H^1_{\text{per}}(\gamma)/\mathbb{R})$ such that $\epsilon^\gamma \nabla u^\epsilon \overset{\ast}{\rightharpoonup} \nabla_y u_1(x, y)$ and the proof is complete.

\[\square\]

**REMARK 5.19.** The same conclusion is valid for $\gamma > 1$. However, in this case the two-scale limit of $u^\epsilon$ will, in general, depend of $y$.

**REMARK 5.20.** Consider a two– scale expansion of the form

$$u^\epsilon \approx u_0(x) + \epsilon^{1-\gamma} u_1(x, \frac{x}{\epsilon})$$

with $\gamma \in (0, 1)$. We have:

$$\epsilon^\gamma \nabla u^\epsilon \approx \epsilon^\gamma \nabla_x u_0(x) + \left( \nabla_x + \frac{1}{\epsilon} \nabla_y \right) \epsilon u_1 \left( x, \frac{x}{\epsilon} \right)$$

$$\approx \epsilon^\gamma \nabla_x u_0 + \epsilon \nabla_x u_1 + \nabla_y u_1$$

$$\rightarrow \nabla_y u_1.$$

(5.23)

Thus, in this setting, the function $u_1(x, y)$ is exactly the higher order term in the expansion, only its order differs from the order of the first term in the expansion by a fractional exponent of $\epsilon$.

### 5.3 The Homogenization Theorem through Two–Scale Convergence

In this section we prove the homogenization theorem for the Dirichlet problem, Theorem 4.2, together with a corrector result, as in Theorem 4.6, using the method of two–scale convergence. Before stating the precise results of this section let us make some remarks on our approach. The first step of in our analysis is to use the energy estimates from Chapter 3 to deduce that $u^\epsilon$ as well as $\nabla u^\epsilon$ have two–scale convergent subsequences. The second step is to use a test function of the form

$$\phi^\epsilon(x) = \phi_0(x) + \epsilon \phi_1 \left( x, \frac{x}{\epsilon} \right)$$

(5.24)

in order to pass to the two scale limit. In this way we obtain a coupled system of equations for the first two terms in the expansion $\{u_0, u_1\}$, the two–scale system, see equation (5.29) below. Well
The posedness of this system is proved using the Lax–Milgram lemma. The final step is to decouple this system of equations using separation of variables, obtaining thus the homogenized equation for \( u_0 \).

Let us now state the homogenization theorem that we will prove in this section.

**THEOREM 5.21.** Let \( u^\varepsilon \) be the solution of

\[
- \frac{\partial}{\partial x_i} \left( a^\varepsilon_{ij}(x) \frac{\partial u^\varepsilon}{\partial x_j} \right) = f, \quad \text{for } x \in \Omega
\]  (5.25a)

\[
u^\varepsilon(x) = 0, \quad \text{for } x \in \partial \Omega.
\]  (5.25b)

with \( f \in H^{-1}(\Omega) \) and \( A^\varepsilon(x) = A \left( \frac{x}{\varepsilon} \right) \), \( A(y) \in M_{\text{per}}(\alpha, \beta, Y) \). Further, let \( u \) be the solution of the homogenized problem

\[
- \overline{a}_{ij} \frac{\partial^2 u}{\partial x_i \partial x_j} = f, \quad \text{for } x \in \Omega
\]  (5.26a)

\[
u(x) = 0, \quad \text{for } x \in \partial \Omega,
\]  (5.26b)

with \( \overline{A} \) given by

\[
\overline{a}_{ij} = \int_Y \left( a_{ij}(y) - a_{ik}(y) \frac{\partial \chi^i(y)}{\partial y_k(y)} \right) dy, \quad i, j = 1, \ldots, d.
\]  (5.27)

and \( \chi^k(y) \), \( k = 1, \ldots, d \) satisfy

\[
- \frac{\partial}{\partial y_i} \left( a_{ij}(y) \frac{\partial \chi^k}{\partial y_j} \right) = - \frac{\partial a_{ik}}{\partial y_i}, \quad \chi^k(y) \text{ is 1-periodic, } k = 1, \ldots, d.
\]  (5.28)

Then \( u^\varepsilon \rightharpoonup u \) weakly in \( H^1_0(\Omega) \).

**REMARK 5.22.** The compactness theorems of the previous section enable us to extract a two–scale convergent subsequence but they do not allow us to conclude that the whole sequence converges to a limit. In the case of Theorem 5.21 the limit \( u \) is unique, since the homogenized equation (5.26) has a unique solution. This implies that the whole sequence converges, not just a subsequence. In the calculations that follow we will use this result without any further reference.

We will break the proof of this theorem into three parts, as advertised. The following lemma provides us with the first two terms of the two–scale expansion, together with the coupled system of equations that they satisfy.
**Lemma 5.23.** Let \( u^\epsilon(x) \) be the solution of (5.25) with the assumptions of Theorem 5.21. Then there exist functions \( u(x) \in H^1_0(\Omega) \), \( u_1(x, y) \in L^2(\Omega; H^1_{\text{per}}(Y)/\mathbb{R}) \) such that \( u^\epsilon \) and \( \nabla u^\epsilon \) two-scale converge to \( u(x) \) and \( \nabla_x u + \nabla_y u_1 \). Furthermore, \( \{u, u_1\} \) satisfy the two-scale system

\[
- \nabla_y [A(y)(\nabla_x u + \nabla_y u_1)] = 0 \quad \text{in } \Omega \times Y, \tag{5.29a}
\]

\[
- \nabla_x \left[ \int_Y A(y)(\nabla_x u + \nabla_y u_1) \, dy \right] = f \quad \text{in } \Omega, \tag{5.29b}
\]

\[
u(x) = 0 \quad \text{on } \partial \Omega, \quad u_1(x, y) \text{ is periodic in } y. \tag{5.29c}
\]

**Proof.**

1. We have that \( \|u^\epsilon\|_{H^1_0(\Omega)} \leq C \) which implies the first part of the lemma: there exist functions \( u(x) \in H^1_0(\Omega) \), \( u_1(x, y) \in L^2(\Omega; H^1_{\text{per}}(Y)/\mathbb{R}) \) such that

\[
u^\epsilon(x) \rightharpoonup u(x), \tag{5.30a}
\]

\[
\nabla \nu^\epsilon(x) \rightharpoonup \nabla_x u(x) + \nabla_y u_1(x, y). \tag{5.30b}
\]

Further, the two-scale limit of \( u^\epsilon(x) \) is also the weak \( H^1_0(\Omega) \)–limit of this sequence.

2. The weak formulation of (5.25) is

\[
\int \Omega a_{ij}(x) \frac{\partial \nu^\epsilon}{\partial x_j} \frac{\partial \phi^\epsilon}{\partial x_i} \, dx = \langle f, \phi^\epsilon \rangle \quad \forall \phi^\epsilon \in H^1_0(\Omega). \tag{5.31}
\]

Now, we expect the solution of (5.25) to be of the form \( u^\epsilon \approx u(x) + \epsilon u_1 \left( x, \frac{x}{\epsilon} \right) + \ldots \). This suggests to use a test function of the form (5.24)

\[
\phi^\epsilon(x) = \phi_0(x) + \epsilon \phi_1 \left( x, \frac{x}{\epsilon} \right), \quad \phi_0 \in C^\infty_0(\Omega), \phi_1 \in C^\infty_0[\Omega; C^\infty_{\text{per}}(Y)].
\]

We clearly have that \( \phi^\epsilon \in H^1_0(\Omega) \). Upon using this test function in (5.31) and rearranging terms a bit we obtain:

\[
I_1 + \epsilon I_2 := \int \Omega \frac{\partial \nu^\epsilon}{\partial x_j} \left[ a_{ij}(x) \left( \frac{\partial \phi_0}{\partial x_i}(x) + \frac{\partial \phi_1}{\partial y_i} \left( x, \frac{x}{\epsilon} \right) \right) \right] \, dx + \epsilon \int \Omega \frac{\partial \nu^\epsilon}{\partial x_j} \left[ a_{ij}(x) \frac{\partial \phi_1}{\partial x_i} \left( x, \frac{x}{\epsilon} \right) \right] \, dx = \langle f, \phi_0 + \epsilon \phi_1 \rangle.
\]

Now, the function \( a_{ij}(x) \left( \frac{\partial \phi_0}{\partial x_i}(x) + \frac{\partial \phi_1}{\partial y_i} \left( x, \frac{x}{\epsilon} \right) \right) \) is of the form \( \phi_1(y) \phi_2(x, y) \) with \( \phi_1(y) \in L^\infty(Y) \) and \( \phi_2(x, y) \in L^2[\Omega; C^\infty_{\text{per}}(Y)] \) and, in view of Remark 5.3, it is an admissible test function. We can thus pass to the two scale limit to obtain:

\[
I_1 \to \int \Omega \int_Y a_{ij}(y) \left( \frac{\partial u}{\partial x_j} + \frac{\partial u_1}{\partial y_j} \right) \left( \frac{\partial \phi_0}{\partial x_i} + \frac{\partial \phi_1}{\partial y_i} \right) \, dy \, dx.
\]
The function $a_{ij}^\varepsilon(x) \frac{\partial \phi_1}{\partial x_i} (x, \frac{x}{\varepsilon})$ is also an admissible test function. Passing to the limit in $I_2$ we obtain:

$$I_2 \to 0.$$  

Moreover, we have that $\phi_0 + \varepsilon \phi_1 \rightharpoonup \phi_0$ weakly in $H^1_0(\Omega)$ which implies that

$$\langle f, \phi_0 + \varepsilon \phi_1 \rangle \to \langle f, \phi_0 \rangle.$$  

Putting the above considerations together we obtain the limiting equation

$$\int_\Omega \int_Y a_{ij}(y) \left( \frac{\partial u}{\partial x_j} + \frac{\partial u_1}{\partial y_j} \right) \left( \frac{\partial \phi_0}{\partial x_i} + \frac{\partial \phi_1}{\partial y_i} \right) dy dx = \langle f, \phi_0 \rangle. \tag{5.36}$$  

In deriving (5.36) we assumed that the test functions $\phi_0, \phi_1$ are smooth, a density argument however enables us to conclude that (5.36) holds for every $\phi_0 \in L^2(\Omega), \phi_1 \in L^2[\Omega; H^1(\Gamma)/\mathbb{R}]$.

3. Now, (5.36) is the weak formulation of the two–scale system (5.29). To see this, we set $\phi_0 = 0$ to obtain:

$$\int_\Omega \int_Y a_{ij}(y) \frac{\partial u}{\partial x_j} \frac{\partial \phi_1}{\partial y_i} dy dx = 0,$$

which is precisely the weak formulation of (5.29a). Setting now $\phi_1 = 0$ in (5.36) we get

$$\int_\Omega \int_Y a_{ij}(y) \left( \frac{\partial u}{\partial x_j} + \frac{\partial u_1}{\partial y_j} \right) \frac{\partial \phi_0}{\partial x_i} dy dx = \langle f, \phi_0 \rangle,$$

which is the weak formulation of (5.29b). The boundary conditions (5.29c) follow from the fact that $u(x) \in H^1_0(\Omega)$ and $u_1(x, y) \in L^2(\Omega; H^1(\Gamma)/\mathbb{R})$. ◼

Now we need to prove that the two–scale system is well posed. This is the content of the next lemma.

**LEMMA 5.24.** The two–scale system (5.29) has a unique solution $(u(x), u_1(x, y)) \in H^1_0(\Omega) \times L^2(\Omega; H^1_1(\Gamma)/\mathbb{R})$, under the assumptions of Theorem 5.21.

**Proof.** We will use the Lax–Milgram Lemma. The weak formulation of the two–scale system is given by equation (5.36) for every $(\phi_0, \phi_1) \in H^1_0(\Omega) \times L^2(\Omega; H^1_1(\Gamma)/\mathbb{R})$. We will denote this product space by $X$. Notice that it is a Hilbert space with norm

$$\|U\|^2_X = \|\nabla u\|_{L^2(\Omega)} + \|\nabla_y u_1\|_{L^2(\Omega \times \Gamma)},$$

where $U = (u, u_1)$. The inner product in $X$ is defined similarly. Let us define now the bilinear form

$$a[U, \Phi] = \int_\Omega \int_Y a_{ij}(y) \left( \frac{\partial u}{\partial x_j} + \frac{\partial u_1}{\partial y_j} \right) \left( \frac{\partial \phi_0}{\partial x_i} + \frac{\partial \phi_1}{\partial y_i} \right) dy dx.$$
Again, we use the notation $\Phi = (\phi_0, \phi_1)$. We have to check that this bilinear form is continuous and coercive. Let us start with continuity. We use the $L^\infty$ bound on $A$, together with the Cauchy–Schwartz inequality to obtain:

$$
a[U, \Phi] = \int_{\Omega} \int_Y a_{ij}(y) \left( \frac{\partial u}{\partial x_i} + \frac{\partial u_1}{\partial y_j} \right) \left( \frac{\partial \phi_0}{\partial x_i} + \frac{\partial \phi_1}{\partial y_j} \right) \, dy \, dx
\leq \beta \int_{\Omega} \int_Y (\nabla_x u + \nabla_y u_1) \cdot (\nabla_x \phi_0 + \nabla_y \phi_1) \, dy \, dx
\leq C \|U\|_X \|\Phi\|_X.
$$

We proceed with coercivity. We use the fact that the integral of the derivative of a periodic function over the unit cell vanishes to obtain:

$$
a[U, U] = \int_{\Omega} \int_Y a_{ij}(y) \left( \frac{\partial u}{\partial x_i} + \frac{\partial u_1}{\partial y_j} \right) \left( \frac{\partial u}{\partial x_i} + \frac{\partial u_1}{\partial y_j} \right) \, dy \, dx
\geq \alpha \int_{\Omega} \int_Y \left| \nabla_x u + \nabla_y u_1 \right|^2 \, dy \, dx
= \alpha \left( \int_{\Omega} \int_Y \left| \nabla_x u \right|^2 \, dy \, dx + 2 \int_{\Omega} \int_Y \nabla_x u \cdot \nabla_y u_1 \, dy \, dx + \int_{\Omega} \int_Y \left| \nabla_y u_1 \right|^2 \, dy \, dx \right)
= \alpha \left( \|u\|^2_X + \|\nabla_y u_1\|^2_{L^2(\Omega \times Y)} \right).
$$

Hence, the bilinear form $a[U, \Phi]$ is continuous and coercive and the Lax–Milgram Lemma applies. This proves existence and uniqueness of solutions of the two–scale system in $X$.

Now we can conclude the proof of Theorem 5.21.

**LEMMA 5.25.** Consider the unique solution $(u, u_1) \in H^1_0(\Omega) \times L^2(\Omega; H^1_{\text{per}}(Y)/\mathbb{R})$ of the two–scale system (5.29). Then $u$ is the unique solution of the homogenized equation (5.26) and $u_1(x, y)$ is of the form

$$u_1(x, y) = -\chi^j(y) \frac{\partial u}{\partial x_j}, \quad (5.41)$$

where $\{\chi^j(y)\}_{j=1}^d$ is the solution of the cell problem (5.28).

**Proof.** We substitute (5.41) into (5.29a) to obtain

$$- \frac{\partial}{\partial y_i} \left( a_{ij}(y) \frac{\partial \chi^k}{\partial x_j} \right) \frac{\partial u}{\partial x_k} = - \frac{\partial a_{ij}}{\partial y_i} \frac{\partial u}{\partial x_j}.
$$

This equation is satisfied provided that $\{\chi^j(y)\}_{j=1}^d \in H^1_{\text{per}}(Y)/\mathbb{R}$ is the unique solution of the cell problem. Now equation (5.29b) becomes

$$- \frac{\partial}{\partial x_i} \left( \int_Y a_{ij}(y) \left( \frac{\partial u}{\partial x_j} - \frac{\partial \chi^k}{\partial y_j} \frac{\partial u}{\partial x_k} \right) \, dy \right) = f.
$$

This is precisely the homogenized equation with the homogenized coefficients given by (5.27).
REMARK 5.26. The fact that the choice (5.41) for \( u_1 \) enables us to solve the two-scale system, provided that \( u_0 \) satisfies the homogenized equation, implies that this the only possible set of functions \((u, u_1)\) which solves the two-scale system, since we have already proved uniqueness of solutions.

The above three lemmas provide us with the proof of Theorem 5.21.

REMARK 5.27. The proof of the homogenization theorem using the method of two-scale convergence might seem quite complicated, since in order to prove the theorem we first had to obtain the two-scale system (5.29) and to prove that it is well posed. However, this method of proof has several advantages.

(i) The choice of the test functions (5.24) which enables us to pass to the limit is ”global”, in the sense that it does not depend on the specific PDE under investigation but only on the fact that the solution admits a two-scale expansion. This is very important, since finding the right test functions for our problem can be quite tricky. We will see later that the same type of test functions will enable us to prove the homogenization theorem for parabolic problems with time dependent coefficients rather easily. The proof based on Tartar’s method of oscillating test functions is much more complicated.

(ii) The well posedness of the two-scale system is not very hard to prove, using the Lax–Milgram Lemma. Now, provided that we can actually decouple the homogenized equation from the cell problem, we can proceed to do so using separation of variables. Notice however that it is not a priori clear that such a decoupling is possible. In particular, in order to be able to obtain a homogenized equation which is independent of the microscale we need good control on the oscillations of the solution to our PDE, in particular we need an \( H^1 \)–estimate. This enables us to conclude that the two-scale limit is independent of the microscale and consequently a homogenized equation actually exists. This is not always the case. We will see examples later on where only \( L^2 \)–estimates can be had, which implies that the two-scale limit does depend on the microstructure.

Now we proceed with the proof of a strong convergence result. In section 4.3 we used the two-scale expansion that we had constructed previously in order to prove that

\[
\lim_{\epsilon \to 0} \| u^\epsilon (x) - \left( u(x) + \epsilon u_1 \left( x, \frac{x}{\epsilon} \right) + \epsilon^2 u_2 \left( x, \frac{x}{\epsilon} \right) \right) \|_{H^1(\Omega)} = 0.
\]
We have also able to compute the convergence rate. Results of this type are called corrector results. We can prove a result of this type using the method of two–scale convergence. We have the following.

**THEOREM 5.28.** Consider \( u^\varepsilon(x) \) and \( u(x) \) as in Theorem 5.21. Further, let \( u_1(x,y) \) be given by (5.41). Then

\[
\lim_{\varepsilon \to 0} \left\| u^\varepsilon(x) - \left( u(x) + \varepsilon u_1 \left( x, \frac{x}{\varepsilon} \right) \right) \right\|_{H^1(\Omega)} = 0.
\]

**Proof.** Since we already know that \( u^\varepsilon \) converges to \( u(x) \) strongly in \( L^2(\Omega) \), in order to prove the theorem it is enough to prove that

\[
\lim_{\varepsilon \to 0} \left\| \nabla u^\varepsilon(x) - \nabla \left( u(x) + \varepsilon u_1 \left( x, \frac{x}{\varepsilon} \right) \right) \right\|_{(L^2(\Omega))^d} = 0.
\]

Equivalently:

\[
\lim_{\varepsilon \to 0} \left\| \nabla u^\varepsilon(x) - \left( \nabla u(x) + \varepsilon \nabla_x u_1 \left( x, \frac{x}{\varepsilon} \right) + \nabla_y u_1 \left( x, \frac{x}{\varepsilon} \right) \right) \right\|_{(L^2(\Omega))^d} = 0.
\]

Assuming now enough regularity on \( u_1 \) so that it can be considered to be an admissible test function we have that \( \| \varepsilon \nabla_x u_1 \left( x, \frac{x}{\varepsilon} \right) \|_{(L^2(\Omega))^d} \rightarrow 0 \). Hence, it is enough to prove that

\[
\lim_{\varepsilon \to 0} \left\| \nabla u^\varepsilon(x) - \left( \nabla u(x) + \nabla_y u_1 \left( x, \frac{x}{\varepsilon} \right) \right) \right\|_{(L^2(\Omega))^d} = 0.
\]

The uniform ellipticity of \( A \) now implies:

\[
\begin{aligned}
\alpha \| \nabla u^\varepsilon(x) - \nabla u(x) - \nabla_y u_1 \left( x, \frac{x}{\varepsilon} \right) \|_{(L^2(\Omega))^d}^2 &= \alpha \int_{\Omega} \left| \nabla u^\varepsilon(x) - \left( \nabla u(x) + \nabla_y u_1 \left( x, \frac{x}{\varepsilon} \right) \right) \right|^2 dx \\
&\leq \int_{\Omega} a_{ij} \left( \frac{x}{\varepsilon} \right) \left( \frac{\partial u^\varepsilon}{\partial x_j}(x) - \frac{\partial u}{\partial x_j}(x) - \frac{\partial u_1}{\partial y_j} \left( x, \frac{x}{\varepsilon} \right) \right) \left( \frac{\partial u^\varepsilon}{\partial x_i}(x) - \frac{\partial u}{\partial x_i}(x) - \frac{\partial u_1}{\partial y_i} \left( x, \frac{x}{\varepsilon} \right) \right) dx \\
&\leq \int_{\Omega} a_{ij} \left( \frac{x}{\varepsilon} \right) \frac{\partial u^\varepsilon}{\partial x_j}(x) \frac{\partial u^\varepsilon}{\partial x_i}(x) dx \\
&\quad + \int_{\Omega} a_{ij} \left( \frac{x}{\varepsilon} \right) \frac{\partial u}{\partial x_j}(x) \frac{\partial u_1}{\partial y_j} \left( x, \frac{x}{\varepsilon} \right) \frac{\partial u^\varepsilon}{\partial x_i}(x) dx \\
&\quad - \int_{\Omega} \left( a_{ij} \left( \frac{x}{\varepsilon} \right) + a_{ji} \left( \frac{x}{\varepsilon} \right) \right) \frac{\partial u^\varepsilon}{\partial x_j}(x) \left( \frac{\partial u}{\partial x_i}(x) + \frac{\partial u_1}{\partial y_i} \left( x, \frac{x}{\varepsilon} \right) \right) dx \\
&\quad \rightarrow \langle f, u \rangle + \int_{\Omega} \int_{Y} a_{ij}(y) \left( \frac{\partial u}{\partial x_i}(x) + \frac{\partial u_1}{\partial y_i}(x, y) \right) \left( \frac{\partial u}{\partial x_j}(x) + \frac{\partial u_1}{\partial y_j}(x, y) \right) dy dx \\
&\quad + \int_{\Omega} \int_{Y} a(y) \left( \frac{\partial u}{\partial x_i}(x) + \frac{\partial u_1}{\partial y_i}(x, y) \right) \left( \frac{\partial u}{\partial x_j}(x) + \frac{\partial u_1}{\partial y_j}(x, y) \right) dy dx \\
\end{aligned}
\]

Consequently \( \lim_{\varepsilon \to 0} \| \nabla u^\varepsilon(x) - \nabla u(x) - \nabla_y u_1 \left( x, \frac{x}{\varepsilon} \right) \|_{(L^2(\Omega))^d} = 0 \). and the theorem is proved \( \square \).
**REMARK 5.29.** The fact that the proof of this theorem is much simpler than the proof of Theorem 4.6 is due to the fact that we didn’t have to study the behavior of the various terms in the two–scale expansion on the boundary of the domain. The price we have to pay for this is that we do not get a convergence rate. Notice however that, apart for the simplicity of this proof, the regularity assumptions are rather minimal.

### 5.4 Two Examples of Non Linear Problems

#### 5.4.1 Introduction

So far we have been exclusively concerned with homogenization for linear problems. In this section we will study the problem of homogenization for a non linear PDE by means of the method of two–scale convergence, namely a PDE governed by a monotone operator. Before doing so we will make some general remarks concerning non linear problems.

The first issue one is confronted with when attempting to study the properties of solutions to non linear PDE is that of the lack of smoothness. Thus, an appropriate concept of weak solutions has to be defined. For various types of non linear PDE, however, the concept of weak solution introduced and studied in Chapter 3 is still adequate. In this section we will study a problem of this form.

#### 5.4.2 Homogenization for Convex Energy Functionals

Consider the following integral functional

#### 5.4.3 Homogenization for Monotone Operators

In this section we consider the problem of homogenization for the following non linear elliptic boundary value problem

\[
-\nabla \left( a \left( \frac{x}{\varepsilon}, \nabla u^\varepsilon \right) \right) = f, \quad \text{for } x \in \Omega \tag{5.46a}
\]

\[
u^\varepsilon(x) = 0, \quad \text{for } x \in \partial \Omega. \tag{5.46b}
\]

We take \( f \in L^2(\Omega) \). For the function \( a(y, \lambda) : Y \times \mathbb{R}^d \rightarrow \mathbb{R} \) we make the following assumptions

1. The map \( \lambda \rightarrow a(y, \lambda) \) is measurable and \( y \)-periodic for every \( \lambda \).

2. The map \( y \rightarrow a(y, \lambda) \) is continuous a.e in \( y \in Y \).
3. There exists a $c > 0$ such that
\[ c|\lambda|^2 \leq a(y, \lambda) \cdot \lambda, \quad \forall y \in Y, \forall \lambda \in \mathbb{R}^d. \tag{5.47} \]

4. There exists a $c > 0$ such that
\[ |a(y, \lambda)| \leq (1 + |\lambda|), \quad \forall y \in Y, \forall \lambda \in \mathbb{R}^d. \tag{5.48} \]

5. $a(y, \lambda)$ is strongly monotone:
\[ [a(y, \lambda) - a(y, \mu)] \geq c|\lambda - \mu|^2, \quad \forall y \in Y, \forall \lambda, \mu \in \mathbb{R}^d. \tag{5.49} \]

**REMARK 5.30.** If we define the nonlinear operator $A : H^1_0(\Omega) \rightarrow H^{-1}(\Omega)$ by
\[ Au = a(x, \nabla u), \]
then the strict monotonicity of $a(y, \lambda)$ immediately yields
\[ (Au - Av, u - v) \geq c\|u - v\|^2 \quad \forall u, v \in H^1_0(\Omega). \]

A weak solution of equation (5.46) is a function $u^\varepsilon \in H^1_0(\Omega)$ such that
\[ \int_\Omega a(x, \varepsilon \nabla u^\varepsilon) \nabla v \, dx = \int_\Omega f v \, dx \quad \forall v \in H^1_0(\Omega) \tag{5.50} \]

A non-linear variant of the Lax–Milgram Lemma, valid for monotone operators which satisfy the assumptions stated above ensures the existence and uniqueness of weak solutions for boundary value problems of the form (5.46); we refer the interested reader to [55, sec. 2.14] for details.\footnote{Alternatively, one can use variational techniques to prove existence and uniqueness of solutions for (5.50): the strict monotonicity of $a(y, \lambda)$ ensures that the corresponding stored energy functional is strictly convex and the direct method in the calculus of variations trivially applies.}

Without getting into any of the details we just state that under the assumptions stated above and for every $\varepsilon > 0$ there exists a unique solution $u^\varepsilon \in H^1_0(\omega)$ of equation (5.50). Now we are ready to state and prove the homogenization theorem for monotone equation (5.46).

**THEOREM 5.31.** The sequence of solutions of (5.50) converges to a function $u(x)$ weakly in $H^1_0(\Omega)$ and the sequence $\nabla u^\varepsilon$ 2-scale converges to $\nabla_x u + \nabla_y u_1(x, y)$ where the vector $(u, u_1)$ is the unique solution in $H^1_0(\Omega) \times L^2[\Omega; H^1_{per}(Y)/\mathbb{R}]$ of the homogenized problem
\[ -\nabla_x \left[ \int_Y a(y, \nabla u(x) + \nabla_y u_1(x, y)] \, dy \right] = f \quad \text{in } \Omega \tag{5.51a} \]
5.4. TWO EXAMPLES OF NON LINEAR PROBLEMS

\[-\nabla_y a[y, \nabla u(x) + \nabla_y u_1(x, y)] = 0 \quad \text{in } Y \tag{5.52}\]
\[u = 0 \quad \text{on } \partial \Omega \tag{5.53}\]
\[y \to u_1(x, y) \text{ is } y-\text{periodic.} \tag{5.54}\]

**Proof.** As it is always the case, the proof of the above homogenization theorem will consist of the two parts. In the first part we derive the necessary energy estimates which enable us to conclude that a limit exists. In the second part we make appropriate choices of test functions in order to pass to the limit.

1. We set \( v = u^\epsilon \) in (5.50) and use assumption (5.48) to obtain

\[
\int_\Omega |\nabla u^\epsilon|^2 \, dx \leq \int_\Omega a \left( \frac{x}{\epsilon}, \nabla u^\epsilon \right) \nabla u^\epsilon \, dx
\]
\[
= \int_\Omega f u^\epsilon \, dx \leq \|f\|_{L^2(\Omega)} \|u^\epsilon\|_{L^2(\Omega)},
\]

from which, together with Poincaré inequality, we deduce that

\[
\|u^\epsilon\|_{H_0^1(\Omega)} \leq C.
\]

Furthermore, we use assumption (5.47) to compute

\[
\int_\Omega a \left( \frac{x}{\epsilon}, \nabla u^\epsilon \right) \nabla u^\epsilon \, dx \leq \int_\Omega (1 + |\nabla u^\epsilon|^2) \nabla u^\epsilon \, dx
\]
\[
\leq \|u^\epsilon\|^2_{H_0^1(\Omega)} \leq C.
\]

Consequently

\[
\left\| a \left( \frac{x}{\epsilon}, \nabla u^\epsilon \right) \right\|_{L^2(\Omega)} \leq C.
\]

We introduce the notation

\[
h^\epsilon := a \left( \frac{x}{\epsilon}, \nabla u^\epsilon \right). \tag{5.55}
\]

The above estimates imply that there exist functions \( u(x) \in H_0^1(\Omega) \), \( h_0(x, y) \in L^2(\Omega; H_{\text{per}}^1(Y)/\mathbb{R}) \) such that

\[
u^\epsilon \rightharpoonup u \quad \text{w–}H_0^1(\Omega),
\]
\[
\nabla u^\epsilon \rightharpoonup \nabla u(x) + \nabla_y u_1(x, y),
\]
\[
h^\epsilon \rightharpoonup h_0(x, y).
\]
We use now the notation (5.55) to write equation (5.46) in the form
\[ f + \nabla \cdot h^\epsilon = 0. \]

We use now the decomposition (5.17) to pass to the limit as \( \epsilon \) tends to 0 to obtain
\[ \nabla_y \cdot h_0(x, y) = 0 \]  
(5.56)

and
\[ f(x) + \nabla_x \left[ \int_Y h_0(x, y) \, dy \right] = 0. \]  
(5.57)

Our goal now is to express \( h_0(x, y) \) in terms of \( a(x, \nabla u), u(x) \) and \( u_1(x, y) \) and, thus, to characterize the limit.

2. We let \( t > 0 \) and consider the following one parameter family of test functions
\[ \mu^\epsilon(x) = \nabla \left[ u(x) + \epsilon \phi_1 \left( x, \frac{x}{\epsilon} \right) \right] + t \epsilon \phi \left( x, \frac{x}{\epsilon} \right), \]
where \( \phi, \phi_1 \in C^\infty_0[\Omega; C^\infty_{\text{per}}(Y)] \). We clearly have that
\[ \mu^\epsilon \to \nabla u(x) + \nabla_y \phi_1(x, y) + t \phi(x, y) =: \mu_0. \]

We use now the strict monotonicity of \( a(y, \lambda) \) to obtain
\[ \int_{\Omega} \left[ a \left( \frac{x}{\epsilon}, \nabla u^\epsilon \right) - a \left( \frac{x}{\epsilon}, \mu^\epsilon \right) \right] \cdot (\nabla u^\epsilon - \mu^\epsilon) \, dx \geq \int_{\Omega} |\nabla u^\epsilon - \mu^\epsilon|^2 \, dx. \]

We rearrange terms a bit and integrate by parts to rewrite the above inequality in the form
\[ \int_{\Omega} \left[ -\nabla a \left( \frac{x}{\epsilon}, \nabla u^\epsilon \right) u^\epsilon - a \left( \frac{x}{\epsilon}, \nabla u^\epsilon \right) \mu^\epsilon \
- a \left( \frac{x}{\epsilon}, \mu^\epsilon \right) \nabla u^\epsilon + a \left( \frac{x}{\epsilon}, \mu^\epsilon \right) \mu^\epsilon \right] \, dx \geq \int_{\Omega} |\nabla u^\epsilon - \mu^\epsilon|^2 \, dx. \]

We pass to the limit as \( \epsilon \to 0 \) in the above inequality to deduce
\[ \int_{\Omega} \int_Y [f u - a(y, \mu_0) \cdot (\nabla u + \nabla_y u_1) - h_0 \cdot \mu_0 + a(y, \mu_0) \cdot \mu_0] \, dy \, dx \geq \int_{\Omega} \int_Y |\nabla_x u + \nabla_y u_1 - \mu_0|^2 \, dy \, dx. \]  
(5.58)

Let now \( \phi_1(x, y) \) be such that
\[ \phi_1^\epsilon \to u_1(x, y), \quad s \in L^2(\Omega; H^1_{\text{per}}(Y)). \]
In this case we have that
\[ \mu_0 = \nabla_x u(x) + \nabla_y u_1(x, y) + t\phi \]

We use this into (5.58) to deduce
\[
\int_\Omega \int_Y [f u + a(y, \mu_0) t\phi - h_0 \cdot \mu_0] \, dy \, dx \geq t^2 \int_\Omega \int_Y |\phi|^2 \, dy \, dx. \tag{5.59}
\]

Upon using equations (5.56) and (5.57), (5.59) becomes
\[
\int_\Omega \int_Y [a(y, \mu_0) - h_0] t\phi \, dy \, dx \geq t^2 \int_\Omega \int_Y |\phi|^2 \, dy \, dx.
\]

We divide through by \( t \) and pass to the limit as \( t \to 0 \) to derive
\[
\int_\Omega \int_Y [a(y, \nabla_x u + \nabla_y u_1) - h_0(x, y)] \phi \, dy \, dx \geq 0.
\]

The above inequality implies
\[ h_0(x, y) = a(y, \nabla_x u + \nabla_y u_1). \tag{5.60} \]

We combine now (5.60) with (5.56) and (5.57) to obtain the two–scale system (5.31). The theorem is proved.

**REMARK 5.32.** Results of corrector type can be obtained for the case of monotone equations. We refer to ?? for details.
Chapter 6

Homogenization for Parabolic PDE

6.1 Introduction

In this chapter we will apply the method of two–scale convergence in order to study the problem of homogenization for the following initial–value (or Cauchy) problem.

\[
\frac{\partial \rho(x, t)}{\partial t} + \nabla \cdot (v(x)\rho(x, t)) = \kappa \Delta \rho(x, t) \quad \text{for} \quad (x, t) \in \mathbb{R}^d \times \mathbb{R}^+, \quad (6.1a)
\]

\[
\theta(x, t = 0) = \theta_0(x) \quad \text{for} \quad x \in \mathbb{R}^d, \quad (6.1b)
\]

with \( \kappa \geq 0 \) and \( \nabla \cdot u(x) = 0 \). When \( \kappa > 0 \) equation (6.1a) is called the advection–diffusion equation. It governs the evolution of a passive tracer which is immersed in a fluid whose velocity field is given by \( v(x) \). Examples of physical entities which adequately be modelled as passive tracers are the ozon in the atmosphere or pollutants in the ocean. The term advection–diffusion stems from the two physical mechanisms which are present in equation (6.1a): advection due to the fluid velocity field \( v(x) \) and molecular diffusion which is controlled by the molecular diffusivity \( \kappa \).

In addition to the incompressibility assumption the velocity field will be taken to be 1–periodic and sufficiently smooth. Notice that the incompressibility assumption implies that the advection term in (6.1a) has the form \( v(x) \cdot \theta \). We also remark that the velocity field is taken to be divergence–free because this is a physically realistic assumption; there are analogues of the homogenization theorems that will be proved in the next chapter for general, compressible, velocity fields.\(^1\)

\(^1\)Alternatively, (6.1a) for \( \kappa > 0 \) can also be considered as the Fokker–Planck equation from the probabilistic theory of diffusion processes. See Chapter 7 or details. Similarly, in the case \( \kappa = 0 \) and when \( d \) is even, equation (6.1a) reduces to the Liouville equation from Hamiltonian mechanics, with \( v(x) = J \nabla H(x) \), where \( H(x) \) is the Hamiltonian and \( J \) is the standard symplectic matrix. We refer to e.g. [4] for details. In the Fokker–Planck or Liouville equation approach, equation (6.1a) governs the evolution of a probability density.
Naturally, a dominant role in the subsequent analysis will be played by the operator

\[ L = \kappa \nabla_y - a(y) \cdot \nabla_y, \quad (6.2) \]

which is defined on the unit torus with periodic boundary conditions. It turns out that the operator \( L \) has completely different properties for \( \kappa = 0 \) and \( \kappa > 0 \). In particular, it is rather straightforward to prove that, when \( \kappa > 0 \), the null space of \( L \) is consists of constants in \( y \) and that it has compact resolvent and, hence, Fredholm theory applies. The above statements are not true in general, for the case \( \kappa = 0 \): the null space of \( L \) depends crucially on the (ergodic) properties of the velocity field \( v(y) \). This is a first instance that the concept of ergodicity plays an important role on the problem of homogenization of various PDE. We will have the occasion to explore this issue further in Chapter 7.

A consequence of the above is that the problem of homogenization for (6.1a) becomes much harder to solve when \( \kappa = 0 \). In fact, we will only be able to study this problem fully in the case when the problem is posed on \( \mathbb{R}^2 \).

Before we proceed with our analysis we need to modify the definition of two–scale convergence, Definition 5.4 to take into account the time dependence of problem (6.1a). We have the following two definitions.

**DEFINITION 6.1.** A function \( \phi(x, y) \in L^2(\Omega \times Y) \) is called an admissible test function if it satisfies

\[ \lim_{\epsilon \to 0} \int_{\Omega} \left| \phi \left( x, \frac{x}{\epsilon} \right) \right|^2 \, dx = \int_{\Omega} \int_{Y} \left| \phi \left( x, y \right) \right|^2 \, dy \, dx. \quad (6.3) \]

**DEFINITION 6.2.** Let \( u^\epsilon \) be a sequence in \( L^2(\Omega) \). We will say that \( u^\epsilon \) two–scale converges to \( u_0(x, y) \in L^2(\Omega \times Y) \) and write \( u^\epsilon \xrightarrow{\epsilon \to 0} u_0 \) if for every admissible test function \( \phi \) we have

\[ \lim_{\epsilon \to 0} \int_{\Omega} u^\epsilon(x) \phi \left( x, \frac{x}{\epsilon} \right) \, dx = \int_{\Omega} \int_{Y} u_0(x, y) \phi(x, y) \, dy \, dx. \quad (6.4) \]

**THEOREM 6.3.** Let \( u^\epsilon \) be a bounded sequence in \( L^2(\Omega) \). Then there exists a subsequence, still denoted by \( u^\epsilon \), and function \( u_0(x, y) \in L^2(\Omega \times Y) \) such that \( u^\epsilon \) two–scale converges to \( u_0(x, y) \).

**THEOREM 6.4.** (i) Let \( u^\epsilon \) be a bounded sequence in \( H^1(\Omega) \). Then \( u^\epsilon \) two-scale converges to its weak–\( H^1 \) limit \( u(x) \). Further, there exists a function \( u_1(x, y) \in L^2(\Omega; H^1_{\text{per}}(Y)/\mathbb{R}) \) such that, up to a subsequence, \( \nabla u^\epsilon \) two–scale converges to \( \nabla_x u(x) + \nabla_y u_1(x, y) \).

\[ ^2 \text{This is to be expected: for } \kappa > 0 \text{ the operator } L \text{ is a second order differential operator, whereas for } \kappa = 0 \text{ it is a first order differential operator. This is the prime example of a singularly perturbed problem.} \]
(ii) Let $u^\epsilon$ and $\epsilon \nabla u^\epsilon$ be uniformly bounded sequences in $L^2(\Omega)$. Then there exists a function $u_0(x, y) \in L^2(\Omega; H^1_{\text{per}}(Y))$ such that, up to a subsequence, $u^\epsilon$ and $\epsilon \nabla u^\epsilon$ two-scale converge to $u_0(x, y)$ and to $\nabla_y u_0(x, y)$, respectively.

(iii) Let $v^\epsilon$ be a divergence-free field which is bounded in $[L^2(\Omega)]^d$. Then the two-scale limit satisfies $\nabla_y \cdot v_0(x, y) = 0$ and $\int_Y \nabla_x \cdot v_0(x, y) \, dy = 0$

6.2 Homogenization for Linear Transport Equations

6.3 Homogenization for Advection–Diffusion Equations

6.4 Homogenization for Parabolic Equations with Time Dependent Coefficients
Chapter 7

Periodic Homogenization for Parabolic Equations: A Probabilistic Approach

7.1 Introduction

In the previous chapter we introduced the advection–diffusion equation as as an equation which governs the evolution of the concentration of a passive tracer, in the presence of molecular diffusion. Alternatively, the advection–diffusion equation can be thought of as an equation for the transition probability density of a particle which moves in a fluid, subject to molecular diffusion. In this chapter we will use this deep connection between solutions of parabolic equations—such as the advection–diffusion equation—and diffusion processes in order to provide a different proof of the homogenization theorem discussed in the previous chapter which is purely probabilistic in nature.

Let us consider the following initial value problem

\[
\frac{\partial T(x, t)}{\partial t} = v(x) \cdot \nabla T(x, t) + \kappa \Delta T(x, t), \tag{7.1a}
\]

\[
T(x, t = 0) = f(x). \tag{7.1b}
\]

where the initial condition \( f(x) \) is taken to be smooth and bounded. We also assume that the velocity field is divergence free, \( \nabla \cdot v(x) = 0 \). The solution of (7.1) can be expressed as a path integral

\[
T(x, t) = \mathbb{E}(f(X_{x,t}^i)), \tag{7.2}
\]

\[\textsuperscript{1}\text{The incompressibility of the velocity field is not necessary for the discussion that follows, we make this assumption only for consistency with the previous chapter.}\]
where $X_{x,t}^\varepsilon$ denotes the particle position which satisfies the *stochastic differential equation* of motion (SDE for short)

$$dX_{x,t}^\varepsilon = v(X_{x,t}^\varepsilon)dt + \sqrt{2\kappa}dW$$

(7.3a)

$$X_{x,0}^\varepsilon = x.$$  

(7.3b)

Here $W(t)$ denotes the standard Brownian motion in $\mathbb{R}^d$ and $\mathbb{E}$ stands for the expectation with respect to the probability measure associated with the process $X_{x,t}^\varepsilon$ (precise definitions will be given in the next section). In this setting, equation (7.1) is called the *backward Kolmogorov equation* associated to the SDE (7.3). The $L^2$–adjoint equation of (7.1)

$$\frac{\partial u(x,t)}{\partial t} + v(x) \cdot \nabla u(x,t) = \kappa \Delta u(x,t),$$

(7.4a)

$$u(x,t = 0) = u_0(x).$$

(7.4b)

is called the *forward Kolmogorov* or Fokker–Planck equation. It governs the evolution of the transition probability density of the stochastic process $X_{t}^\varepsilon$.

Assume now that the velocity field is 1–periodic and consider the rescaled advection–diffusion equation

$$\frac{\partial u^\varepsilon(x,t)}{\partial t} + \frac{1}{\varepsilon} v\left(\frac{x}{\varepsilon}\right) \cdot \nabla u^\varepsilon(x,t) = \kappa \Delta u^\varepsilon(x,t),$$

(7.5a)

$$u(x,t = 0) = u_0(x).$$

(7.5b)

according to the results of the previous chapter the solution of the initial value problem (7.5) $u^\varepsilon$ converges as $\varepsilon \to 0$ to the solution of

$$\frac{\partial \overline{u}(x,t)}{\partial t} = \mathcal{K}D^2 \overline{u}(x,t),$$

(7.6a)

$$u(x,t = 0) = u_0(x).$$

(7.6b)

where the effective diffusivity is given by ?? . Equation (7.6) can be solved explicitly to yield

$$\overline{u}(x,t) = \int_{\mathbb{R}^d} G(x - y,t)u_0(y) \, dy$$

where the *heat kernel* is

$$G(x,t) = \frac{1}{(4\pi t)^{\frac{d}{2}}} \frac{1}{(\text{det} \mathcal{K})^{\frac{1}{2}}} e^{-\frac{(x-\kappa^{-1}x,\varepsilon)^2}{4t}}$$

\[2\text{In deriving (7.4) we have used the incompressibility of the velocity field.}\]
Let try to reinterpret the above homogenization result in terms of the stochastic process (7.3). The SDE corresponding to the rescaled equation (7.5) reads

\[
\frac{dX_t^{c,x}}{\epsilon} = \frac{v}{\epsilon} \left( \frac{X_t^{c,x}}{\epsilon} \right) dt + \sqrt{2\kappa} dW
\]

\[X_{t=0}^{c,x} = x.\] (7.7a)

On the other hand, the SDE corresponding to the homogenized equation (7.6) is

\[
d\bar{X}_t^x = \sqrt{2A}dW
\]

\[\bar{X}_{t=0}^x = x.\] (7.8a)

where the matrix \(A\) is such that \(AA^T = \mathcal{K}\). In other words, the homogenization theorem implies that the process \(X_t^{c,x}\) defined through (7.7) converges to the Gaussian process \(\bar{X}_t^x\) with mean and variance

\[\mathbb{E}\bar{X}_t^x = 0 \quad \text{and} \quad \mathbb{E}\bar{X}_t^x \bar{X}_s^x = \mathcal{K}\min(s, t),\]

respectively. Thus, in terms of the underlying stochastic process, the homogenization theorem for the advection diffusion equation (7.4) tells us that—for fixed time \(t\)—the sequence of random variables \(X_t^{c,x}\) converges to the mean zero Gaussian random variable whose variance can be computed from the cell problem ???. But this is a manifestation of the central limit theorem from probability theory! This observation suggests that it might be fruitful to use ideas from the theory of stochastic processes in order to prove homogenization theorems for partial differential equations, at least for those which admit a probabilistic interpretation. This is precisely the purpose of this chapter: to develop probabilistic techniques for the study of homogenization.

The rest of this chapter is organized as follows. In section (7.2) we will recall some basic facts from the theory of diffusion processes. In section (7.3) we will present an abstract central limit theorem for martingales. In section (7.4) we will use this central limit theorem in order to provide an alternative proof of the homogenization theorem for advection–diffusion equations.

### 7.2 Background Material on Diffusion Processes

### 7.3 The Martingale Central Limit Theorem

In this section we will present without proof the martingale central limit theorem (MCLT for short). The MCLT says, roughly speaking, that the long time behavior of a martingale is governed by an
effective Brownian motion, provided that appropriate compactness and ergodicity assumptions are satisfied. Next, we will use the MCLT in order to prove an abstract homogenization theorem. This theorem says—again without being very precise— that a homogenization theorem holds whenever the effective diffusivity is well defined. In other words, we will prove that the well posedness of the cell problem is a necessary and sufficient condition for the homogenization theorem to hold. In the next section we will use this theorem to study the problem of homogenization for the advection diffusion equation (7.1).

7.3.1 The Central Limit Theorem for Martingales

The basic technical tool which is needed for the analysis in this chapter is the following.

THEOREM 7.1. (MCLT) Let \( \{M_t : t \geq 0\} \) be a right continuous square integrable martingale on a probability space \((\Omega, \{\mathcal{F}_t\}, \mathbb{P})\) with respect to a given increasing filtration \(\{\mathcal{F}_t : t \geq 0\}\). Further, let \(\langle M_t, M_t \rangle\) denote its quadratic variation and assume that

1. \( M_0 = 0 \).
2. The increments of \( M_t \) are stationary.
3. Its quadratic variation converges in \( L^1(\mathbb{P}) \) to some positive constant \( \sigma^2 \):

\[
\lim_{t \to \infty} \mathbb{E} \left[ \frac{\langle M_t, M_t \rangle - t}{t} - \sigma^2 \right] = 0. \tag{7.9}
\]

Then \( M_t / \sqrt{t} \) converges in distribution to a mean zero Gaussian law with variance \( \sigma^2 \).

We refer to e.g. [15, ch. 7] for a proof of this theorem.

The only martingales that we will need in order to prove the homogenization theorem for advection–diffusion equations are stochastic integrals. An immediate corollary of the above theorem, in conjunction with the martingale representation theorem, that we will need later on is the following.

COROLLARY 7.2. Let \( I^\varepsilon_t = \int_0^t f(x(s)) \, d\beta(s) \) where \( x(s) \) is a stationary Markov process. Assume that

\[
\lim_{t \to \infty} \mathbb{E} \left[ \epsilon \int_0^{t/\varepsilon} f^2(s) \, ds - \sigma^2 \right] = 0.
\]

Then \( \epsilon I^\varepsilon_t / \varepsilon^2 \) converges in distribution to a Brownian motion with variance \( \sigma^2 \).

\(^3\)The theorem that we will prove is in fact less general than that; the proof of the sharpest possible result requires more work. See however the remarks at the end of this and the following section.
Now we can use Theorem (7.1) to prove a central limit theorem for additive functionals of Markov processes. In the next section we will see that the homogenization theorem for advection–diffusion equations follows as a corollary from this functional central limit theorem. In order to state our theorem we need to introduce some notation. Let \( X_t \) be a Markov process taking values on a complete separable metric space \( E \) endowed with its Borel \( \sigma \)-algebra \( \mathcal{E} \). We assume that there exists a stationary state \( \pi \). Let \( L^2_\pi \) be the space of square integrable functions with respect to the stationary measure \( \pi \) with norm \( \| \cdot \|_\pi \) and inner product \( \langle \cdot , \cdot \rangle_\pi \). Let \( L \) denote the generator of the process in \( L^2_\pi \) with domain \( D(L) \). Let \( L^* \) be the adjoint of \( L \) in \( L^2_\pi \). Let \( \mathbb{P}_\pi \) be the measure on path space \( D(\mathbb{R}_+, E) \) induced by the Markov process \( X_t \) starting from \( \pi \) and by the \( \mathbb{E}_\pi \) the expectation with respect to \( \mathbb{P}_\pi \).

Let now \( V \in L^2_\pi \) with \( \mathbb{E}_\pi V = 0 \). We want to find conditions on \( V \) and the generator \( L \) which are sufficient for the existence of a central limit theorem for the process

\[
Z^\epsilon_t = \epsilon \int_0^{t/\epsilon^2} V(X_s) \, ds. \tag{7.10}
\]

We will express these conditions in terms of the properties of solutions of the Poisson equation \(^4\)

\[
-Lf = V. \tag{7.11}
\]

Now we can state the functional central limit theorem.

**Theorem 7.3. (Central Limit Theorem for Additive Functionals of Markov Processes.)** Let \( X_t \) be a Markov process with generator \( L \) which is ergodic with invariant measure \( \pi \). Further, let \( V(x) \) be a function such that \( V \in L^2_\pi \) with \( \mathbb{E}_\pi V = 0 \). Assume that there exists a unique solution \( f \) of (7.11) with \( f, f^2 \in D(L) \). Assume finally that the initial condition of \( X_t \) is distributed according to the invariant measure \( \pi \). Then \( Z^\epsilon_t \) converges in distribution to a Gaussian process with variance

\[
\sigma^2(V) = 2\langle f, (-L)f \rangle_\pi. \tag{7.12}
\]

**Proof.** Since \( f \in D(L) \) the process

\[
M_t = f(X_t) - f(X_0) - \int_0^t (Lf)(X_s) \, ds \tag{7.13}
\]

is well defined and, further, it is a martingale. Now, the function \( f \) solves the Poisson equation

\(^4\)This is precisely the cell problem.
(7.11). We use this in (7.13) and solve for $Z_t^\varepsilon$ to obtain:

$$Z_t^\varepsilon = \varepsilon \int_0^{t/\varepsilon^2} V(X_s) \, ds$$

$$= -\varepsilon \int_0^{t/\varepsilon^2} Lf(X_s) \, ds$$

$$= \varepsilon M_{t/\varepsilon^2} + \varepsilon (f(X_0) - f(X_{t/\varepsilon^2}))$$

$$=: \varepsilon M_{t/\varepsilon^2} + I_t^\varepsilon.$$  \hspace{1cm} (7.14)

Since by assumption $f \in L_\pi^2$ and the measure $\pi$ is stationary we have that the second term in the above equation converges to 0 as $t \to \infty$ in $L_\pi^2$:

$$\lim_{\varepsilon \to 0} \mathbb{E}_\pi |I_t^\varepsilon|^2 = 0.$$

In order to complete the proof of the theorem we need to check that that the martingale $\varepsilon M_{t/\varepsilon^2}$ satisfies the conditions of Theorem 7.1. First, since $X_t$ under $P_\pi$ is a stationary Markov process, $\varepsilon M_{t/\varepsilon^2}$ has stationary increments. Further, $M_0 = 0$. Since $f^2 \in D(L)$ the quadratic variation of $M_t$ is well defined and thus

$$\varepsilon \langle M_{t/\varepsilon}, M_{t/\varepsilon} \rangle = \varepsilon \int_0^{t/\varepsilon} \left( Lf^2(X_s) - 2f(X_s)Lf(X_s) \right) \, ds.$$  \hspace{1cm} (7.15)

Now, since the process $X_t$ is ergodic we can apply Birkhoff’s (individual) ergodic theorem cite petersen to deduce

$$\lim_{\varepsilon \to 0} \varepsilon \langle M_{t/\varepsilon}, M_{t/\varepsilon} \rangle = 2 \langle f, (-L)f \rangle_\pi = 2\sigma^2(V) \quad \text{in } L^1(\pi).$$

In the above we used the fact that $\langle Lf^2 \rangle_\pi = 0$. Hence, Theorem 7.1 applies and we conclude that $\varepsilon M_{t/\varepsilon^2}$, and consequently $Z_t^\varepsilon$ converges in distribution to a Gaussian process with variance $\sigma^2(V)$ given by (7.12).

\[ \square \]

**REMARK 7.4.** If we do not assume that $f^2 \in D(L)$ then the formula for the quadratic variation of $M_t$ (7.15) is not valid. However, even in this case $\langle M, M \rangle_t$ is an increasing additive functional and, using an approximation argument, we can prove that the expectation of $\langle M, M \rangle_t$ is still given by $2t\langle f, (-L)f \rangle_\pi$.

**REMARK 7.5.** Connection with Fredholm alternative, connection with Kurtz’s theorem.

**REMARK 7.6.** Remark that not necessary to assume stationary initial conditions.
REMARK 7.7. From the formula for the variance, or effective diffusivity, we immediately get that it is nonnegative definite. Hence, this result follows from the theorem and we do not need to prove separately.

Now we want to weaken the conditions of Theorem 7.3. We first need to introduce some notation. Assume that there exists a common core $C \in D(L) \cap D(L^*)$. We denote by $\mathcal{H}_1 \subset C$ the Hilbert space with norm
$$
\|f\|_1^2 = \langle f, (I - L)f \rangle_\pi
$$
and inner product endowed with polarization
$$
\langle f, (-L)f \rangle_\pi = \frac{1}{4} (\|f + g\|_1^2 - \|f - g\|_1^2)
$$
Notice that only the symmetric part $S = \frac{1}{2}(L - L^*)$ of $L$ enters into the definition of $\mathcal{H}_1$. We denote by $\mathcal{H}_1^*$ the dual space of $\mathcal{H}_1$. The norm in this space is
$$
\|f\|_{-1}^2 = \langle f, (I - L)^{-1}f \rangle_\pi.
$$
The inner product is also obtained by polarization. We have that, for every $f \in D(L)$, $gL^2_\pi \cap \mathcal{H}_{-1}$
$$
|\langle f, g \rangle_\pi| \leq \|f\|_1 \|g\|_{-1}.
$$
Moreover, a function $f \in D(L)$ is an element of $\mathcal{H}_{-1}$ if and only if there exists a constant $C$ such that
$$
|\langle f, g \rangle_\pi| \leq C \|g\|_1
$$
for every $g \in D(L)$. In this case we have $\|f\|_{-1} \leq C$. Assume not that $V \in L^2_\pi \cap \mathcal{H}_{-1}$ and let $f_\lambda$ denote the solution of the resolvent equation
$$
\lambda f_\lambda - Lf_\lambda = V.
$$
We have the following theorem:

THEOREM 7.8. Suppose that
$$
\sup_{0 < |\lambda| \leq 1} \|\lambda f_\lambda\|_{-1} < \infty. \quad (7.16)
$$
Then the $\mathbb{P}_\pi$–law of the process $I_t$ defined in (7.10) converges to a mean zero Gaussian distribution with variance
$$
\sigma^2(V) = 2 \lim_{\lambda \to 0} \|f_\lambda\|_1^2. \quad (7.17)
$$
We emphasize that the theorem above is sharp, since we merely assume that the effective diffusivity can be defined. compare with Kubo theory.
7.4 The Homogenization Theorem for Advection–Diffusion Equations Revisited
Bibliography


