Homogenization via formal multiscale asymptotics and volume averaging: How do the two techniques compare?

Yohan Davit [a,b,*], Christopher G. Bell [c], Helen M. Byrne [c,d], Lloyd A.C. Chapman [c,d], Laura S. Kimpton [c], Georgina E. Lang [c], Katherine H.L. Leonard [d], James M. Oliver [c], Natalie C. Pearson [c], Rebecca J. Shipley [e], Sarah L. Waters [c], Jonathan P. Whiteley [d], Brian D. Wood [f], Michel Quintard [a,b]

[a] Université de Toulouse, INPT, UPS, IMFT (Institut de Mécanique des Fluides de Toulouse), Allée Camille Soula, F-31400 Toulouse, France
[b] CNRS, IMFT, F-31400 Toulouse, France
[c] Mathematical Institute, University of Oxford, 24–29 St Giles', Oxford OX1 3LB, UK
[d] Department of Computer Science, University of Oxford, Parks Road, Oxford OX1 3QD, UK
[e] Department of Mechanical Engineering, University College London, Torrington Place, London WC1 E7JE, UK
[f] School of Chemical, Biological, and Environmental Engineering, Oregon State University, Corvallis, OR 97331, United States

Keywords:
Homogenization
Upscaling
Volume averaging
Multiscale asymptotics
Porous media

A B S T R A C T

A wide variety of techniques have been developed to homogenize transport equations in multiscale and multiphase systems. This has yielded a rich and diverse field, but has also resulted in the emergence of isolated scientific communities and disconnected bodies of literature. Here, our goal is to bridge the gap between formal multiscale asymptotics and the volume averaging theory. We illustrate the methodologies via a simple example application describing a parabolic transport problem and, in so doing, compare their respective advantages/disadvantages from a *practical point of view*. This paper is also intended as a pedagogical guide and may be viewed as *a tutorial for graduate students* as we provide historical context, detail subtle points with great care, and reference many fundamental works.

1. Introduction

The effective behavior of multiscale, multiphase materials has been of interest to researchers from the 19th Century. The earliest examples include Maxwell's work on the conductivity of dilute suspensions [1] and Einstein's analysis of the viscosity of a dilute suspension of neutrally buoyant hard spheres [2]. Several precursory ideas were presented in these studies, in particular the concepts of *effective* conductivity and viscosity. The continued use of these early results as limit cases or approximate correlations serves to illustrate how fundamental and remarkable they were. Nowadays, effective theories have applications as diverse as composite materials [3], biological tissues [4], biofilms [5], networks of large-scale bodies such as buildings [6], the mechanics of masonry structures [7], reservoirs with large faults [8] or transport in vascular networks [9].

A typical multiscale problem is illustrated in Fig. 1 for a porous medium. Pore-scale properties, such as the indicator field describing the phase geometry, vary *rapidly* with the spatial coordinates relative to the scales of the macroscopic domain. In Fig. 1, this is to say that the characteristic lengthscales μ and ℓ are much smaller than a characteristic large-scale length, L,

$$\mu,\ \ell \ll L. \tag{1.1}$$

Discretization of domains that satisfy Eq. (1.1) necessarily yields a substantial amount of mesh cells, making it challenging to compute solutions of partial differential equations in such multiscale systems. A solution to this numerical problem is to adopt a macroscopic viewpoint and use models in which high frequency fluctuations have been filtered out (see Fig. 2).

In most cases, such effective medium approaches were first formulated from an empirical point of view, e.g., Darcy's law [10], the dispersion equation [11–13], and the generalized Darcy's laws for multiphase flows [14]. Later on, the hypothesis that these descriptions could be obtained theoretically by averaging microscale equations found its way into the scientific community, before the sixties, and with a rapid pace thereafter. One of the first fundamental analyses related to porous media was devised in the fifties by Taylor [15] and Aris [16]. It was concerned with solute transport in a Poiseuille flow and deriving an asymptotic equation that would describe the transport of the average cross-section concentration in a tube. Taylor and Aris showed that this average satisfies a one-dimensional advection–dispersion equation and that the

* Corresponding author. Tel.: +33 534322882.
E-mail address: yohan.davit@imft.fr (Y. Davit).

http://dx.doi.org/10.1016/j.advwatres.2013.09.006

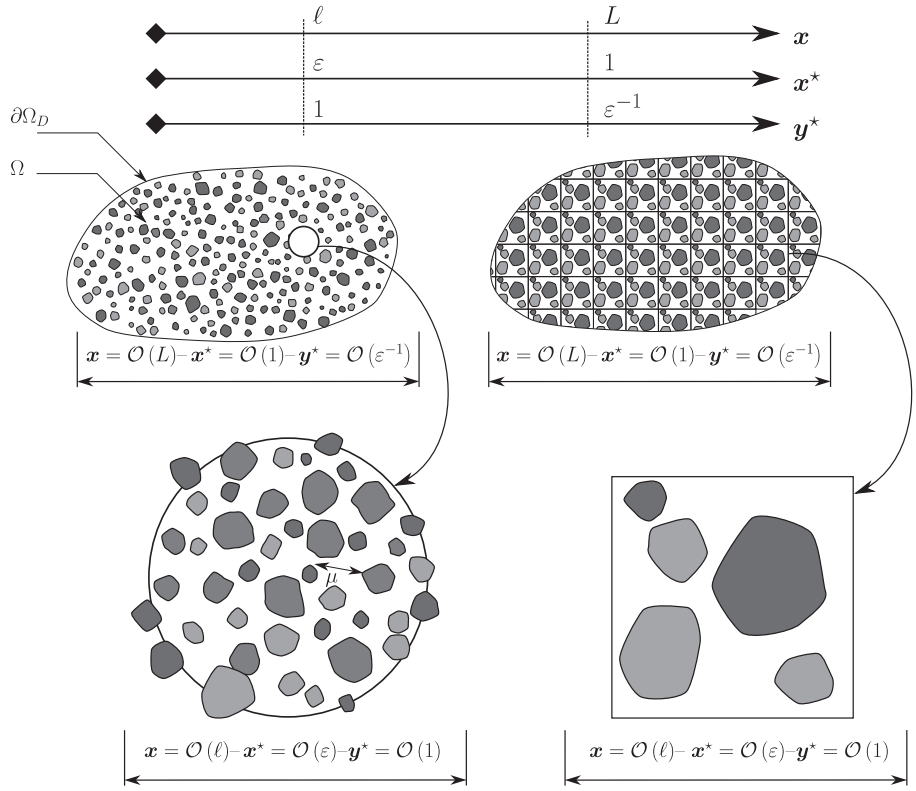


Fig. 1. Illustration of the different coordinate systems and the hierarchy of scales for *a priori* non-periodic (left-hand side) and periodic (right-hand side) media. The dimensionalized system corresponds to the spatial variable $\boldsymbol{x}$ where μ is a pore-scale characteristic length, ℓ is the size of the averaging volume and L is a macroscale characteristic length. Further, we have illustrated two additional coordinate systems that correspond to the spatial variables $\boldsymbol{x}^\star$ (macroscale) and $\boldsymbol{y}^\star$ (microscale), nondimensionalized with L and ℓ respectively.

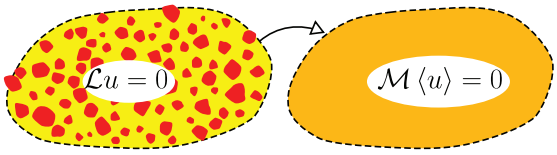


Fig. 2. Schematic diagram illustrating micro- and macroscale descriptions. The microscale differential operator, $\mathcal{L}$, applying to u, is transformed into a macroscale operator, $\mathcal{M}$, that involves effective parameters and applies to the average value $\langle u\rangle$. The microscale geometry exhibits high-frequency fluctuations that have been filtered out in the macroscale geometry.

dispersion coefficient is proportional to the square of the Péclet number. This result is valid only asymptotically (in the long-time limit), the relevant timescale being the time for a molecule of solute to travel the entire width of the tube. Hence, the analysis is particularly useful when the width of the tube is much smaller than the total length. More generally, this notion of separation of scales, Eq. (1.1), is central to the development of macroscale theories (see [17] or [18] for broad historical perspectives on mechanics).

Not only have averaging approaches led to thousands of contributions, but also a proliferation of theoretical frameworks (see [19] for a review). Generic homogenization techniques include deterministic methods such as volume averaging, multiscale asymptotics, mixture theories and the generalized method of moments (or Taylor–Aris–Brenner method, see [20]); and stochastic approaches based on ensemble averaging, i.e., where macroscale quantities are sought as mathematical expectations [21–24]. These frameworks led to significant advances in the field and to the development of new application areas such as optimal design [25] or shape optimization [26]. However, this enormous volume of works (with little connection between them) has also resulted in a lot of confusion. Indeed, how many times have we heard arguments about the relationship between the different theories, even from the most prominent contributors themselves? Surprisingly, there has been little effort to clarify these questions. One of the authors remembers hours of discussion at UC Davis with Stephen Whitaker during, or after, the visits of known contributors to porous media theories: Bourgeat, Cushman, Dagan, Gray, to cite a few. These discussions raised interesting and fundamental questions. However, on only one occasion did this lead to a public contribution: a short note (in French) comparing asymptotic homogenization and the method of volume averaging [27]. The goal that the authors outlined in this short paper remains largely unachieved and the purpose of this contribution is to advance further in this direction.

2. Historical background: volume averaging and multiscale asymptotics

2.1. Volume averaging

The idea underlying volume averaging is that macroscale variables can be defined through the use of spatial averaging. Early works in the sixties include [28–31]. For example, Marle [28,29] tried to justify Darcy's law using irreversible thermodynamics and out of equilibrium fundamental relationships (Onsager reciprocal relations). The idea that macroscale models should be compatible with thermodynamical principles was not new (see [12,32,33]), but the introduction of the volume averaging framework initiated a highly productive methodology that was

subsequently used by several authors (for instance by Hassanizadeh and Gray [34] and Bennethum and Cushman [35]). This methodology bears connections with developments belonging to the *theory of mixtures* (see [36]). Averaging balance equations typically yields additional terms that can be interpreted heuristically via physical notions such as tortuosity, dispersion, permeability or interphase exchanges (see [31,37]). Such empirical approaches have been used to interpret microscale numerical simulations and to propose new forms of macroscale equations (e.g., in [38,39]). Alternatively, macroscale equations may be derived via the method of volume averaging with *closure*, as developed by Whitaker (see [40]), which relies on a perturbation analysis and scaling approximations.

2.2. Multiscale asymptotics

In mathematics, asymptotic methods are often used to study differential operators involving rapidly oscillating coefficients. It is far beyond the scope of this paper to provide an exhaustive review of asymptotic methods, so we will focus on developments in the field of *partial differential equations* that appeared in the seventies (for reference, example works on ordinary differential equations include [41,42]). The term *homogenization* was first coined in a computational context by Babuška [43], where he presents multiscale asymptotic developments for an application to periodic domains in nuclear reactors. Early works also include contributions by Sanchez-Palencia for applications to a variety of problems (e.g., [44–46]). The methodology of Sanchez-Palencia, which is the core of the formal multiscale asymptotics that will be presented in this paper, consisted of a perturbative decomposition (a two-scale expansion) of variables as a function of a (very) small parameter ε, defined as the ratio of micro- and macro-lengthscales, $\varepsilon = \frac{\ell}{L}$ (see Eq. (1.1)). One of the most important references in the field, that discusses formal approaches along with convergence proofs, is the book by Bensoussan et al. [47].

These early results focused on periodic multiscale asymptotics. Non-periodic results, which were developed concomitantly in the seventies, are based on more general and powerful convergence principles such as $\Gamma-, G-$ or $H-$ convergence (see e.g., [48,49] for a broad perspective regarding such developments). We remark that a specific advantage of periodic homogenization, which is particularly interesting for practical engineering applications, is that it may enable effective parameters to be determined as functions of tensor fields that are calculated locally over a unit-cell. For further references on multiscale asymptotics, we refer also to the book [50], the paper by Auriault [51] and the many contributions of Allaire, Arbogast, Auriault, Ene, Hornung, Lions, Mei, Mikelić, Papanicolaou, Sanchez-Palencia and Tartar.

3. Objectives and organization

Several branches of volume averaging and multiscale asymptotics were developed in parallel. Although they share similar goals, their nature is significantly different. In this work, we will study the particular type of *volume averaging* presented in Whitaker's book [40]. We will compare this technique to *formal periodic asymptotics*, as developed in the early works of E. Sanchez-Palencia, in [47] or in [52]. We will not consider "mathematical homogenization" and convergence aspects, although we will discuss some results in Section 9.

In addition to serving as a comparison of multiscale asymptotics and volume averaging, this paper may be viewed as a tutorial for graduate students and researchers. We have taken particular care to detail nontrivial points that are rarely explained in the literature. To further clarify presentation, we use "*Technical Notes*" throughout the paper, in order to distinguish essential material from more advanced points. We recommend that on first reading, these notes are ignored. We remark that learning homogenization theories takes time and effort, but it is a worthwhile investment. The process of averaging can provide additional insight into many different problems in physics and mathematics.

The remainder of this paper is organized as follows. We will first illustrate formal asymptotics and volume averaging via an example application to a transport problem in a hierarchical multiscale system. This is done in three steps through Sections 4–6. In Section 4, we define the physical framework and the example transport problem. In Section 5, we explain how the method of volume averaging may be used to homogenize this problem. In Section 6, we present results obtained via formal multiscale asymptotics. In Section 7, we discuss fundamental properties of the homogenized problem and show how they can be used to check for errors and missing terms. Section 8 is dedicated to comparing both methods in terms of final results, assumptions and algorithms. *Again, this comparison is done in the context of a practical application*. Finally, in Section 9, we discuss more advanced aspects of the methods, including current open problems and highlight important contributions to the field.

4. Definitions and problem formulation

4.1. Notations and algebra

Algebraic notations often differ from one paper to another, making it difficult for researchers to fight their way through the jungle of symbols. The variety of alternatives used to represent similar objects (e.g., whether to use matrices, dyadics, vectors or tensors; index or Gibbs notation; upper or lower indices; nesting or other conventions; tensor or Kronecker symbols for outer products) can be bewildering to say the least. To further complicate matters, conventions and notations are often only implicitly defined. In this paper, we will work with tensors and tensor fields as presented below (for detailed introductions to this topic, see [53,54]).

For the Gibbs (symbolic compact) notations, we will use the following notations:

- zero-order tensors (scalar) or zero-order tensor fields (scalar fields) with Roman characters, e.g., $A \in \mathbb{R}$;
- first-order tensors (vector) or first-order tensor fields (vector fields) with bold Roman characters, e.g., $\boldsymbol{A} \in \mathbb{R}^n$ (for our purposes we will focus on $n \leq 3$). Roman subscripts will identify the components of a vector, e.g., A_i;
- second-order tensors (dyadic or matrix) or second-order tensor fields (dyadic fields) with bold sans serif characters, e.g., $\mathbf{A} \in \mathbb{R}^n \times \mathbb{R}^n$. We will refer to each component of this tensor as A_{ij}.

For the index notation, we will use Einstein's summation convention via repeated indices:

$$A_i x_i \equiv \sum_i A_i x_i. \tag{4.1}$$

We will not perform changes of bases (except for rescaling), so that we will not worry about covariance and contravariance.

Apart from tensor addition and multiplication by a scalar, two additional operations are needed in this work. First, we will use the outer product of two vectors defined by

$$\boldsymbol{A}\boldsymbol{B} \equiv \mathbf{C}, \tag{4.2}$$

with $\mathsf{C}_{ij} = A_i B_j$.

Second, we will use the dot-product (inner product) that will follow the nesting convention, i.e., *the last index of the first tensor is contracted with the first index of the last tensor*. To be explicit, we have

- Vector · Vector = Scalar: $\boldsymbol{A}\cdot\boldsymbol{B}\equiv A_iB_i$.
- Vector · Dyadic = Vector: $\boldsymbol{A}\cdot\mathbf{B}\equiv\boldsymbol{C}$ with $C_i=A_j\mathsf{B}_{ji}$. We remark that this operation is illegal in matrix algebra and may yield confusion. This is, however, a convenient convention for partial differential equations as it unambiguously defines the operation $\boldsymbol{\nabla}\cdot\mathbf{A}$ in which nabla is seen a vector with components $\boldsymbol{\nabla}_i=\partial_i$.
- Dyadic · Vector = Vector: $\mathbf{B}\cdot\boldsymbol{A}\equiv\boldsymbol{C}$ with $C_i=\mathsf{B}_{ij}A_j$.
- Dyadic · Dyadic = Dyadic: $\mathbf{A}\cdot\mathbf{B}\equiv\mathbf{C}$ with $\mathsf{C}_{ij}=\mathsf{A}_{ik}\mathsf{B}_{kj}$. This may also be written with vectors as $(\boldsymbol{AB})\cdot(\boldsymbol{CD})=\boldsymbol{A}(\boldsymbol{B}\cdot\boldsymbol{C})\boldsymbol{D}$.
- Dyadic : Dyadic = Scalar: $\mathbf{A}:\mathbf{B}\equiv C=\mathsf{A}_{ij}\mathsf{B}_{ji}$.

For clarity, we will use both the Gibbs (symbolic representations of tensors) and index notations (component representation of tensors) when necessary.

4.2. Parabolic transport problems

In this subsection, we present example transport problems that are described by pure diffusive equations, before proposing a prototypical mathematical formulation in the next subsection. One such problem is heat conduction in composite systems or porous media (see e.g., [55]). Energy conservation within each region is often written as

$$\partial_t T=\boldsymbol{\nabla}\cdot\left(\frac{k}{\rho c_p}\boldsymbol{\nabla}T\right),\tag{4.3}$$

where T [K] is the temperature, k [W m^{-1} K^{-1}] the thermal conductivity, ρ [kg m^{-3}] the mass density and c_p [J kg^{-1} K^{-1}] the specific heat capacity.

Another example problem is the diffusion of a solute in multiphase or multi-region systems. Mass conservation yields (assuming no reactions or sources/sinks of solute)

$$\partial_t c=\boldsymbol{\nabla}\cdot(\mathbf{D}(\boldsymbol{x})\cdot\boldsymbol{\nabla}c),\tag{4.4}$$

with c [mol m^3] the solute concentration and $\mathbf{D}$ [m^2 s^{-1}] the diffusion tensor.

For weakly compressible isothermal fluids (see a discussion of these points in [56]), the deviation of mass density ρ [kg m^3] from its reference value ρ_0 may be approximated by $\rho\approx\rho_0[1+\gamma(P-P_0)]$ with γ [N^{-1} m^2] the fluid compressibility, P [N m^{-2}] the pressure and P_0 [N m^{-2}] is the reference pressure corresponding to ρ_0. For porous media flows with negligible gravitational effects, mass conservation combined with Darcy's law may yield

$$\partial_t P=\boldsymbol{\nabla}\cdot\left(\frac{\mathbf{K}(\boldsymbol{x})}{\epsilon\gamma\mu}\cdot\boldsymbol{\nabla}P\right),\tag{4.5}$$

where $\mathbf{K}$ [m^2] is the permeability tensor, ϵ is the porosity, and μ [N m^{-2} s^{-1}] is the dynamic viscosity.

All these examples lead to a similar diffusion equation and we will use this equation as a generic example throughout this paper. We remark that terms such as ρc_p or $\epsilon\gamma\mu$ are considered constant through time and space, otherwise their introduction in the divergence term would not be possible. A generalization to non-constant coefficients would lead to additional complexity which is beyond the scope of this work.

4.3. Prototypical initial boundary value problem

In Sections 5 and 6 of this paper, we will study homogenization of the following partial differential equation (a more mathematical formalization is described in Technical note 1),

$$\underbrace{\partial_t u}_{\text{rate of change}}=\underbrace{\boldsymbol{\nabla}\cdot(\mathbf{A}(\boldsymbol{x})\cdot\boldsymbol{\nabla}u)}_{\text{diffusion}}=\partial_{x_i}\left(\mathsf{A}_{ij}\partial_{x_j}u\right),\tag{4.6}$$

in which u is a scalar field that may correspond to a variety of physical parameters (e.g., pressure, temperature or concentration). We will assume that the system is subject to the following Dirichlet boundary condition,

$$u=f(\boldsymbol{x},t)\quad\text{on }\partial\Omega_D,\tag{4.7}$$

and the initial condition

$$u(\boldsymbol{x},0)=0.\tag{4.8}$$

This problem can be generalized by incorporating source terms into Eq. (4.6) and considering nontrivial initial conditions in place of Eq. (4.8). Additionally, a spatially varying weight could be included in front of the rate of change term, $\alpha(\boldsymbol{x})\partial_t u$, on the left-hand side of Eq. (4.6). In order to highlight the differences between volume averaging and multiscale asymptotics, we will focus on the simplest problem defined by Eqs. (4.6) to (4.8).

In Eq. (4.6), we have assumed that $\mathbf{A}$ varies with position $\boldsymbol{x}$. This will enable us to analyze materials with smoothly varying coefficients without explicitly needing to incorporate different phases. This approach is only valid when there are no jumps in the variables and no sources or sinks on the boundaries between phases. In practice, weak derivatives (and partial differential equations defined in the sense of distributions) may be used to relax such restrictions.

Technical note 1: The same problem may be written more rigorously as follows. Let Ω be an open bounded subset of $\mathbb{R}^n$ with boundary $\partial\Omega_D$. We consider the following parabolic problem

$$\partial_t u=\boldsymbol{\nabla}\cdot(\mathbf{A}(\boldsymbol{x})\cdot\boldsymbol{\nabla}u)\quad\text{in }\Omega\times\mathbb{R}^+,$$

along with the boundary condition

$$u=f(\boldsymbol{x},t)\quad\text{on }\partial\Omega_D\times\mathbb{R}^+,$$

and the initial condition $u(\boldsymbol{x},0)=0$ in Ω, wherein $\mathbf{A}(\boldsymbol{x})$ is a dyadic field defined on $\overline{\Omega}=\Omega\cup\partial\Omega_D$. We may further complete this problem by defining functional spaces and properties of the problem, e.g., $u\in L^2(\Omega)$, $\mathbf{A}$ is positive-definite, symmetric, or there exist $\alpha,\beta\in\mathbb{R}^{+*}$ with $\alpha\leqslant\beta$ such that $\alpha|\boldsymbol{\xi}|^2\leqslant\mathsf{A}_{ij}\xi_i\xi_j\leqslant\beta|\boldsymbol{\xi}|^2$. In the remainder of this paper, however, attention focuses on formal asymptotic developments, with little concern for functional spaces, topologies and convergence (mathematical approaches that can be used to address such issues are discussed in Section 9).

4.4. Nondimensionalization

We consider a hierarchical system, as illustrated in Fig. 1, for which natural lengthscales, ℓ and L, characterize the micro- and macroscales, respectively. Using these lengthscales, we define the dimensionless spatial variables $\boldsymbol{x}^\star\equiv\frac{\boldsymbol{x}}{L}$ and $\boldsymbol{y}^\star\equiv\frac{\boldsymbol{x}}{\ell}$ (see Fig. 1) where $\boldsymbol{x}$ is the dimensional variable. The dimensionless lengths are often referred to as the micro- ($\boldsymbol{y}^\star$ is $\mathcal{O}(1)$ at the microscale) and macro- ($\boldsymbol{x}^\star$ is $\mathcal{O}(1)$ at the macroscale) variables (see also Fig. 1), with the

relationship $\boldsymbol{y}^\star = \varepsilon^{-1}\boldsymbol{x}^\star$ with $\varepsilon \equiv \frac{\ell}{L} \ll 1$, see Eq. (1.1). We scale time $t = Tt^\star$ with $T = \frac{L^2}{\mathsf{A}}$ (where A is the norm of the arithmetic mean of $\mathbf{A}(\boldsymbol{x})$), $\mathbf{A}^\star(\boldsymbol{x}^\star) = \frac{T\mathbf{A}(\boldsymbol{x})}{L^2}$, $u^\star(\boldsymbol{x}^\star) = \frac{u(\boldsymbol{x})}{u_0}$ and $f^\star(\boldsymbol{x}^\star, t^\star) = \frac{f(\boldsymbol{x},t)}{u_0}$ where u_0 is a reference value, e.g., $u_0 = \max|f|$. We rescale Eqs. (4.6)–(4.8) to give

$$\partial_{t^\star} u^\star = \nabla_{\boldsymbol{x}^\star} \cdot \left(\mathbf{A}^\star(\boldsymbol{x}^\star) \cdot \nabla_{\boldsymbol{x}^\star} u^\star\right), \tag{4.9}$$

$$u^\star = f^\star, \tag{4.10}$$

$$u^\star(\boldsymbol{x}^\star, 0) = 0. \tag{4.11}$$

Henceforth, we remove the stars on the dimensionless variables for simplicity. Further, we will assume that $\mathbf{A}$ exhibits only high (spatial) frequency fluctuations, i.e., varies with a characteristic lengthscale which is $\mathcal{O}(\varepsilon)$. In the next two sections, we will present homogenization of this initial boundary value problem.

5. Homogenization via volume averaging

The volume averaging approach relies on direct spatial averaging of the partial differential equations. To do this, we proceed schematically as follows:

1. Define the averaging process.
2. State theorems that allow the interchange of spatial averaging and differential operators.
3. Average equations and use an "average-plus-perturbation" decomposition.
4. Make assumptions, primarily concerning the timescales and lengthscales of the processes, to simplify the problem and obtain an approximate form of the perturbations.
5. Derive a closed form of the macroscopic equations.

A more detailed description of the theory's algorithm is given in Section 8 (see Fig. 10).

5.1. Definitions of averaging volume and moving average

In Section 4.3, the initial boundary value problem which describes u was defined for $u \in \Omega \subset \mathbb{R}^n$ with n a positive integer. In practice, we will focus on cases for which $n \leq 3$ with a particular emphasis on $n = 3$. In the *volume* averaging literature, the notation used often reflects this specific choice, with sets measured as "volumes". To avoid confusion and simplify comparison with previous works, we will follow this terminology and denote by $\mathcal{V}(\boldsymbol{x}) \subset \Omega$ the averaging (closed) set at point $\boldsymbol{x}$, and by V its volume ($V \equiv \int_{\mathcal{V}(\boldsymbol{x})} dV$) (see Fig. 3).

For any tensor $\boldsymbol{\psi}$ (including scalars, vectors and dyadics), we define the moving volume average (a more general definition is given in Technical note 2) as

$$\langle\boldsymbol{\psi}\rangle(\boldsymbol{x}, t) \equiv \frac{1}{V}\int_{\xi\in\mathcal{V}(\boldsymbol{x})} \boldsymbol{\psi}(\xi, t) dV_\xi, \tag{5.1}$$

or equivalently,

$$\langle\boldsymbol{\psi}\rangle(\boldsymbol{x}, t) = \frac{1}{V}\int_{\zeta\in\mathcal{V}(\mathbf{0})} \boldsymbol{\psi}(\boldsymbol{x}+\zeta, t) dV_\zeta. \tag{5.2}$$

We will assume that $V \equiv$ constant and suppose that $\mathcal{V}(\boldsymbol{x})$ is the closed ball $\mathcal{B}_r(\boldsymbol{x})$ with radius r centered at point $\boldsymbol{x}$. In multiphase systems, we also often define the phase and intrinsic averages. For phase α (the closed set $\mathcal{V}_\alpha$ within the averaging volume $\mathcal{V}$, see Fig. 3), these correspond to

$$\text{Phase average:} \quad \langle\boldsymbol{\psi}_\alpha\rangle = \frac{1}{V}\int_{\xi\in\mathcal{V}_\alpha(\boldsymbol{x})} \boldsymbol{\psi}_\alpha(\xi, t) dV_\xi \tag{5.3}$$

and

$$\text{Intrinsic average:} \quad \langle\boldsymbol{\psi}_\alpha\rangle^\alpha = \frac{1}{V_\alpha}\int_{\xi\in\mathcal{V}_\alpha(\boldsymbol{x})} \boldsymbol{\psi}_\alpha(\xi, t) dV_\xi, \tag{5.4}$$

where $\boldsymbol{\psi}_\alpha$ is the restriction of $\boldsymbol{\psi}$ to the subset $\mathcal{V}_\alpha$ and V_α is the volume associated with $\mathcal{V}_\alpha$. With these notations, we have the simple relationship

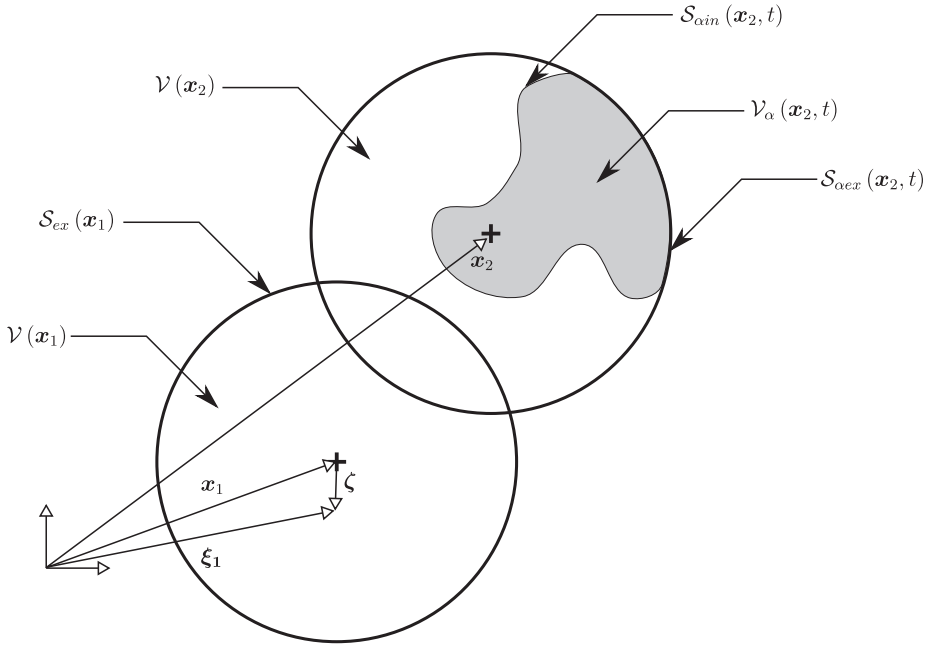


Fig. 3. A schematic representation of the averaging sets and interfaces, their positions and the coordinate system notations. In this figure, the centroid of the averaging volume $\mathcal{V}(\boldsymbol{x})$ is represented using the variable $\boldsymbol{x}$ and the integration vector using the vectors ζ or ξ. The surface delineating between $\mathcal{V}(\boldsymbol{x})$ and the rest of the medium is $\mathcal{S}_{ex}(\boldsymbol{x})$. If the averaging contains multiple phases, each portion of phase α contained in $\mathcal{V}(\boldsymbol{x})$ is denoted $\mathcal{V}_\alpha(\boldsymbol{x})$ with internal and external boundaries $\mathcal{S}_{\alpha in}(\boldsymbol{x})$ and $\mathcal{S}_{\alpha ex}(\boldsymbol{x})$, respectively.

$$\langle \boldsymbol{\psi}_\alpha \rangle = \frac{V_\alpha}{V} \langle \boldsymbol{\psi}_\alpha \rangle^\alpha. \tag{5.5}$$

Technical note 2: Moving averages are more generally defined by (see [57–60])

$$\langle \boldsymbol{\psi} \rangle_m \equiv \int_{\mathbb{R}^n} m(\xi) \bar{\boldsymbol{\psi}}(\boldsymbol{x} - \xi, t) dV_\xi \equiv m \bigstar \bar{\boldsymbol{\psi}}, \tag{5.6}$$

where $\bigstar$ denotes the spatial convolution, $m : \mathbb{R}^n \to \mathbb{R}$ is a smoothing function that has compact support in $\mathbb{R}^n$ and is normalized so that $\int_{\mathbb{R}^n} m(\xi) d\xi \equiv 1$. Further, $\bar{\boldsymbol{\psi}}$ denotes an extension of $\boldsymbol{\psi}$ outside Ω,

$$\bar{\boldsymbol{\psi}} = \begin{cases} \boldsymbol{\psi} \text{ in } \bar{\Omega}, \\ \check{\boldsymbol{\psi}} \text{ in } \mathbb{R}^n \setminus \bar{\Omega}, \end{cases}$$

where $\bar{\Omega}$ is the closure of Ω and the function $\check{\boldsymbol{\psi}}$ remains to be defined. This extension is required to define *averages close to domain boundaries*. In most cases, since the boundary layer on which the difficulty exists is $\mathcal{O}(\varepsilon)$ and since we are interested in the limit $\varepsilon \to 0$, we proceed "as if" the domain was unbounded. However, this treatment is not always satisfactory, e.g., the error may propagate further than $\mathcal{O}(\varepsilon)$ or this may yield difficulties in the treatment of the macroscale boundary conditions. The reader is referred to Prat [61] for a discussion of this issue. A similar problem for multiphase systems and solutions via jump boundary conditions are discussed in, e.g., [62,63]. An advantage of using the convolution definition is that further mathematics can be simplified by choosing $m \in C^\infty$, i.e., is continuous and has continuous derivatives of all orders. An example of one such possible choice for $m(\xi)$ is:

$$m(\xi) = \begin{cases} C \exp\left(-(r^2 - |\xi|^2)^{-1}\right) & \text{if } \|\xi\| < r, \\ 0 & \text{if } \|\xi\| \geqslant r. \end{cases} \tag{5.7}$$

We may also obtain $\langle \boldsymbol{\psi} \rangle_m = \langle \boldsymbol{\psi} \rangle$ by considering the discontinuous functional

$$m(\xi) = \begin{cases} 1 & \text{if } \|\xi\| \leqslant r, \\ 0 & \text{if } \|\xi\| > r. \end{cases}$$

5.2. Theorems

Application of the volume averaging methodology relies heavily upon the following three theorems (proofs are available in, among others, [58,59,64–66]) for sufficiently smooth (see Technical note 3) tensor fields T_α, $\boldsymbol{T}_\alpha$ and $\mathbf{T}_\alpha$ defined in phase α.

Theorem 1.

$$\langle \nabla \cdot \boldsymbol{T}_\alpha \rangle = \nabla \cdot \langle \boldsymbol{T}_\alpha \rangle + \frac{1}{V} \int_{\mathcal{S}_{\alpha in}(\boldsymbol{x}, t)} \boldsymbol{n}_\alpha \cdot \boldsymbol{T}_\alpha \, dS$$

and

$$\langle \nabla \cdot \mathbf{T}_\alpha \rangle = \nabla \cdot \langle \mathbf{T}_\alpha \rangle + \frac{1}{V} \int_{\mathcal{S}_{\alpha in}(\boldsymbol{x}, t)} \boldsymbol{n}_\alpha \cdot \mathbf{T}_\alpha \, dS.$$

The equivalent index notations are $\langle \partial_i T_{\alpha i} \rangle = \partial_i \langle T_{\alpha i} \rangle + \frac{1}{V} \int_{\mathcal{S}_{\alpha in}(\boldsymbol{x}, t)} n_{\alpha i} T_{\alpha i} \, dS$ *and* $\langle \partial_i \mathsf{T}_{\alpha ij} \rangle = \partial_i \langle \mathsf{T}_{\alpha ij} \rangle + \frac{1}{V} \int_{\mathcal{S}_{\alpha in}(\boldsymbol{x}, t)} n_{\alpha i} \mathsf{T}_{\alpha ij} \, dS$.

Theorem 2.

$$\langle \nabla T_\alpha \rangle = \nabla \langle T_\alpha \rangle + \frac{1}{V} \int_{\mathcal{S}_{\alpha in}(\boldsymbol{x}, t)} \boldsymbol{n}_\alpha T_\alpha \, dS$$

and

$$\langle \nabla \boldsymbol{T}_\alpha \rangle = \nabla \langle \boldsymbol{T}_\alpha \rangle + \frac{1}{V} \int_{\mathcal{S}_{\alpha in}(\boldsymbol{x}, t)} \boldsymbol{n}_\alpha \boldsymbol{T}_\alpha \, dS.$$

The equivalent index notations are $\langle \partial_i T_\alpha \rangle = \partial_i \langle T_\alpha \rangle + \frac{1}{V} \int_{\mathcal{S}_{\alpha in}(\boldsymbol{x}, t)} n_{\alpha i} T_\alpha \, dS$ *and* $\langle \partial_i T_{\alpha j} \rangle = \partial_i \langle T_{\alpha j} \rangle + \frac{1}{V} \int_{\mathcal{S}_{\alpha in}(\boldsymbol{x}, t)} n_{\alpha i} T_{\alpha j} \, dS$.

Theorem 3. $\langle \partial_t T_\alpha \rangle = \partial_t \langle T_\alpha \rangle - \frac{1}{V} \int_{\mathcal{S}_{\alpha in}(\boldsymbol{x}, t)} (\boldsymbol{n}_\alpha \cdot \boldsymbol{w}_\alpha) T_\alpha \, dS$. *The equivalent index notation is* $\langle \partial_t T_\alpha \rangle = \partial_t \langle T_\alpha \rangle - \frac{1}{V} \int_{\mathcal{S}_{\alpha in}(\boldsymbol{x}, t)} n_{\alpha i} w_{\alpha i} T_\alpha \, dS$.

In these equations, $\mathcal{S}_{\alpha in}(\boldsymbol{x}, t)$ represents internal boundaries of the phase α with centroid at $\boldsymbol{x}$; $\boldsymbol{n}_\alpha$ denotes the outward unit normal vector to the boundary of phase α; and $\boldsymbol{w}_\alpha$ is velocity of the corresponding boundary. In Theorems 1 and 2, the average of a derivative is systematically expressed as the sum of two terms. Loosely speaking, this can be interpreted as a decomposition of the surface terms that appear in the divergence theorem. For instance, we may write

$$\langle \nabla \cdot \boldsymbol{T}_\alpha \rangle = \frac{1}{V} \int_{\mathcal{S}_{\alpha ex}(\boldsymbol{x}, t)} \boldsymbol{n}_\alpha \cdot \boldsymbol{T}_\alpha \, dS + \frac{1}{V} \int_{\mathcal{S}_{\alpha in}(\boldsymbol{x}, t)} \boldsymbol{n}_\alpha \cdot \boldsymbol{T}_\alpha \, dS, \tag{5.8}$$

where $\mathcal{S}_{\alpha ex}(\boldsymbol{x}, t)$ is the intersection of $\mathcal{S}_{ex}(\boldsymbol{x})$, the surface external to the averaging volume centered at $\boldsymbol{x}$ (see Fig. 3), with phase α. We can demonstrate that $\frac{1}{V} \int_{\mathcal{S}_{\alpha ex}(\boldsymbol{x}, t)} \boldsymbol{n}_\alpha \cdot \boldsymbol{T}_\alpha \, dS = \nabla \cdot \langle \boldsymbol{T}_\alpha \rangle$ (see [66]), which provides an explicit representation in terms of average quantities. The second term, however, corresponds to the integral over the internal surface $\mathcal{S}_{\alpha in}(\boldsymbol{x}, t)$ which characterizes microscale internal heterogeneities.

To facilitate solution, in Section 4, we defined a problem that has no explicit internal phase boundaries. This means that we may average via the operator $\frac{1}{V} \int_{\mathcal{V}} \bullet dV$, as opposed to averaging via $\frac{1}{V} \int_{\mathcal{V}_\alpha} \bullet dV$, so that we can treat the medium "as a single phase" with (smoothly) spatially varying coefficients. We remark, however, that this approach is only valid when there are no jumps in the variables that are differentiated and no sources or sinks on the boundaries between phases (no singularities, see Technical note 3). If these assumptions hold, we can write $\mathcal{S}_{in}(\boldsymbol{x}, t) = \emptyset$ and all the second terms in Theorems 1 to 3 vanish so that averages and derivatives can be readily interchanged. Therefore, Theorems 1 to 3 simplify to.

Theorem 4. $\langle \nabla \cdot \boldsymbol{T} \rangle = \nabla \cdot \langle \boldsymbol{T} \rangle$ *and* $\langle \nabla \cdot \mathbf{T} \rangle = \nabla \cdot \langle \mathbf{T} \rangle$.

Theorem 5. $\langle \nabla T \rangle = \nabla \langle T \rangle$ *and* $\langle \nabla \boldsymbol{T} \rangle = \nabla \langle \boldsymbol{T} \rangle$.

Theorem 6. $\langle \partial_t T \rangle = \partial_t \langle T \rangle$.

In many cases, this simplification does not apply. For example, if we consider porous media then we must deal with perforated domains, internal solid interfaces and, possibly, source terms on boundaries. Further, jumps and discontinuities in the concentrations or fluxes across phase boundaries may need to be accounted for via surface integrals (see Technical note 3).

Technical note 3: A condition underlying Theorems 1 and 2 is that T_α, $T_{\alpha i}$ and $T_{\alpha ij}$ are continuous and have continuous derivatives, i.e., are $\mathcal{C}^1$ within phase α. Theorems 4 and 5 further require that T, T_i or T_{ij} are $\mathcal{C}^1$ over the averaging volume, i.e., that there are no jumps on the boundaries of internal phases. To relax such assumptions and accommodate discontinuities, we must use weak derivatives and partial differential equations defined in the sense of distributions. In particular, for functions that are piecewise $\mathcal{C}^1$ ($\mathcal{C}^n$ functions almost everywhere are relatively common in mathematical representations of multiphase systems, whereby fields are smooth within phases and exhibit singularities on boundaries), we can use Schwartz's jump formula (see [67]). For example, for $n=3$ and a surface discontinuity, $\nabla f = (\nabla f)^u + (f_+ - f_-)\boldsymbol{n}_{-+}\sigma$ in which $(\nabla f)^u$ is the usual gradient within $\mathcal{C}^1$ parts; f_+ and f_- are continuous prolongations of f on the $+$ and $-$ sides, respectively; $\boldsymbol{n}_{-+}$ is the unit vector normal to the discontinuous surface $\mathcal{S}_{\text{jump}}$ oriented from $-$ to $+$; and σ is the surface "Dirac distribution" such that $\frac{1}{V}\int_{\mathcal{V}} \sigma f dV = \frac{1}{V}\int_{\mathcal{S}_{\text{jump}}} f dS$. This representation is particularly interesting when used in conjunction with the convolution representation of the moving average, as it yields a more concise formalism for which averaging theorems are more rigorously defined and demonstrated (see [58,59]).

5.3. Average-plus-perturbation decomposition

We introduce the following decomposition of $u(\boldsymbol{x}, t)$ into an average component, $\langle u\rangle(\boldsymbol{x}, t)$, and a perturbation, $\tilde{u}(\boldsymbol{x}, t)$ (for another definition of the perturbation, see Technical note 4):

$$u(\boldsymbol{x}, t) = \langle u\rangle(\boldsymbol{x}, t) + \tilde{u}(\boldsymbol{x}, t). \tag{5.9}$$

In many branches of mathematics, e.g., linear stability analysis, a perturbation analysis is introduced in order to study the influence of *small* deviations from a reference or equilibrium state. This often leads to confusion with the different purposes of multiscale perturbation decompositions. To further complicate the matter, it often happens that the multiscale perturbations are small (i.e., $\langle u\rangle = \mathcal{O}(1)$ and $\tilde{u} = \mathcal{O}(\varepsilon)$ where $0 < \varepsilon \ll 1$) but this is not necessarily the case. In Eq. (5.9), *we do not require $\tilde{u}$ to be small*. At this stage, we do not impose any constraint on the perturbations. In Section 5, we will however use this decomposition to impose constraints on the spatial and temporal variations of $\tilde{u}$ and $\langle u\rangle$. More specifically, spatial localization will be obtained by imposing that spatial frequencies of $\langle u\rangle$ are $\mathcal{O}(1)$, while $\tilde{u}$ contains larger frequencies of order $\mathcal{O}(\varepsilon^{-1})$ (see Fig. 4). Similarly, we will impose scalings of the temporal frequencies for $\langle u\rangle$ and $\tilde{u}$.

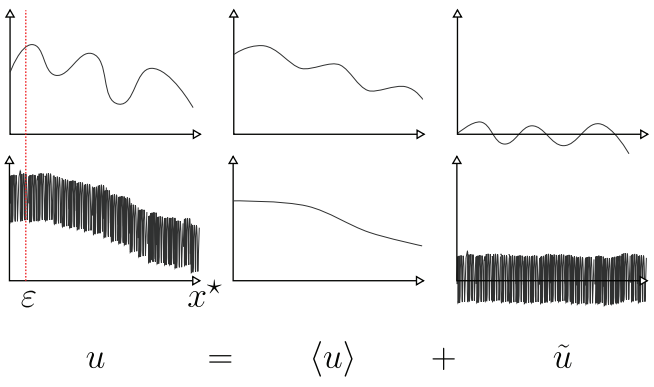


Fig. 4. Illustration of the average plus perturbation decomposition for: (top line) the case where the size of the averaging volume is similar to that of the signal wavelength; and (bottom line) the case where the size of the averaging volume is much larger the signal wavelength.

Technical note 4: In the literature, for example in Bear's work (see e.g., [68]), the perturbation has also been defined in the following way. Consider the unique field $u(\boldsymbol{x}, t)$ and the average value, $\langle u\rangle(\boldsymbol{\chi}, t)$, at point $\boldsymbol{\chi}$ such that we may have $\boldsymbol{\chi} \neq \boldsymbol{x}$. We can define the perturbation decomposition as

$$u(\boldsymbol{x}, t) = \langle u\rangle(\boldsymbol{\chi}, t) + \mathring{u}(\boldsymbol{x}, t; \boldsymbol{\chi}).$$

This equation defines the perturbation at any given point $\boldsymbol{x}$ relative to the average value at any point $\boldsymbol{\chi}$ in the system. In most cases, we will have $\mathring{u}(\boldsymbol{x}, t; \boldsymbol{\chi}_1) \neq \mathring{u}(\boldsymbol{x}, t; \boldsymbol{\chi}_2)$ if $\boldsymbol{\chi}_1 \neq \boldsymbol{\chi}_2$ and $\mathring{u}(\boldsymbol{x}, t; \boldsymbol{\chi}_1) = \mathring{u}(\boldsymbol{x}, t; \boldsymbol{\chi}_2)$ if $\boldsymbol{\chi}_1 = \boldsymbol{\chi}_2$, even though $\mathring{u}(\boldsymbol{x}, t; \boldsymbol{\chi}_1)$ and $\mathring{u}(\boldsymbol{x}, t; \boldsymbol{\chi}_2)$ always refer to the same $(\boldsymbol{x}, t)$.

5.4. Averaging

We first average both sides of Eq. (4.9) to obtain

$$\langle \partial_t u\rangle = \langle \nabla \cdot (\mathbf{A}(\boldsymbol{x}) \cdot \nabla u)\rangle. \tag{5.10}$$

Since there are no internal boundaries, we apply Theorems 4 to 6 which yield

$$\partial_t \langle u\rangle = \nabla \cdot \langle \mathbf{A}(\boldsymbol{x}) \cdot \nabla u\rangle. \tag{5.11}$$

Again, we emphasize that we are manipulating a special case here: in general, boundary conditions should be introduced carefully when averaging.

We now introduce the "average-plus-perturbation" decomposition (Eq. (5.9)), and obtain

$$\partial_t \langle u\rangle = \nabla \cdot \langle \mathbf{A}(\boldsymbol{x}) \cdot (\nabla\langle u\rangle + \nabla\tilde{u})\rangle. \tag{5.12}$$

At this stage of the development, we have obtained an average version of Eq. (4.9) with only few constraints upon the fields (e.g., no separation of scales). This equation involves $\nabla\tilde{u}$, so that its resolution requires knowledge of $\tilde{u}$. We often say that the macroscopic equation is not in a "closed form" since it does not involve only $\langle u\rangle$. In order to obtain a closed form, we seek an approximate expression for $\tilde{u}$ that will simplify the coupling. To this end, we must derive the initial boundary value problem which is satisfied by $\tilde{u}$ and then obtain a convenient form of the solution.

5.5. Perturbation

Subtraction of Eq. (5.12) from Eq. (4.9) (consider that $\tilde{u} = u - \langle u\rangle$, so that it becomes an intuitive operation) yields

$$\begin{aligned}\partial_t \tilde{u} = {} & \nabla \cdot (\mathbf{A}(\boldsymbol{x}) \cdot \nabla\tilde{u}) - \nabla \cdot \langle \mathbf{A}(\boldsymbol{x}) \cdot \nabla\tilde{u}\rangle + \nabla \cdot (\mathbf{A}(\boldsymbol{x}) \cdot \nabla\langle u\rangle) \\ & - \nabla \cdot \langle \mathbf{A}(\boldsymbol{x}) \cdot \nabla\langle u\rangle\rangle,\end{aligned} \tag{5.13}$$

with boundary and initial conditions

$$\tilde{u} = f - \langle u \rangle \quad \text{on } \partial\Omega_D, \tag{5.14}$$

$$\tilde{u}(\boldsymbol{x}, 0) = 0. \tag{5.15}$$

Although we have obtained an initial boundary value problem that is solved by $\tilde{u}$, this problem involves source terms $\nabla \cdot (\mathbf{A}(\boldsymbol{x}) \cdot \nabla\langle u \rangle) - \nabla \cdot \langle \mathbf{A}(\boldsymbol{x}) \cdot \nabla\langle u \rangle \rangle$ and $\langle u \rangle$ that act on the macroscale. If possible, this micro/macro coupling should be eliminated: otherwise we will have to solve for $\tilde{u}$, coupled with $\langle u \rangle$, over the entire domain, gaining no practical advantage as compared to a direct resolution of the microscopic problem (or we could use Green's functions as discussed in Technical note 5).

Technical note 5: Without further assumptions, we may exploit the nature of the differential operator to obtain a solution in terms of the Green's function $G(\boldsymbol{x}, \boldsymbol{x}', t, t')$ (see [69–71]). $\tilde{u}$ can then be expressed as an integral form that involves G or $\widetilde{G} = G - \langle G \rangle$, along with the source terms that appear in Eq. (5.13) ($\nabla \cdot (\mathbf{A}(\boldsymbol{x}) \cdot \nabla\langle u \rangle) - \nabla \cdot \langle \mathbf{A}(\boldsymbol{x}) \cdot \nabla\langle u \rangle \rangle$) and the boundary data $f - \langle u \rangle$. In this case, G or $\widetilde{G}$ needs to be solved microscopically over Ω and we have gained no advantage as compared to a direct resolution of the initial problem. Further, the mathematical expressions are rather tedious and the process leads to a complex integrodifferential form of Eq. (5.12). Such integral descriptions are generally termed nonlocal (see discussion in Section 5.7). While these problems are undoubtedly complex, they are of interest because of the mathematical and physical insight that they can provide.

5.6. Assumptions

To facilitate solution and decoupling of the micro- and macroscales, we formulate the following concepts (which are explained in further detail in the next subsection):

5.6.1. The averaging volume as a representative volume element (RVE)

(**A1**): The system is hierarchical and we can define a representative volume element, i.e., a relatively small portion of Ω over which parameter fields (not variables such as u) are spatially quasi-stationary. Further, we choose a RVE and fix the averaging volume $\mathcal{V}$ as this RVE, so that ε corresponds to the size of the averaging volume/RVE.

5.6.2. Localization

For our test problem, we aim to construct a homogenized model that is *local* in space and time, i.e., parameter fields and variables behave locally. To this end, we will constrain the variations of fields relative to the value of ε that was fixed in (**A1**). Here, we assume that:

(**A2**): Nondimensionalized spatial frequencies of an average quantity, $\langle \boldsymbol{\psi} \rangle$ (e.g., $\langle u \rangle$), are $\mathcal{O}(1)$ whereas $\tilde{\boldsymbol{\psi}}$ contains larger frequencies of order $\mathcal{O}(\varepsilon^{-1})$ with $\varepsilon \ll 1$, i.e., $\langle \boldsymbol{\psi} \rangle$ varies very slowly in space compared to $\tilde{\boldsymbol{\psi}}$.

(**A3**): Nondimensionalized temporal frequencies of $\langle u \rangle$ are $\mathcal{O}(1)$ and those of $\tilde{u}$ are $\mathcal{O}(\delta^{-1})$ with $\delta \ll 1$, i.e., temporal variations in $\langle u \rangle$ are much slower than in $\tilde{u}$. Further, we assume that δ is "small enough" to justify temporal quasi-stationarity of the perturbation problem.

5.6.3. Unit-cell and periodicity

(**A4**): We can express $\tilde{u}(\boldsymbol{x})$ in terms of a macroscale variable, $\boldsymbol{x}$, and a microscale variable, $\boldsymbol{y}$, so that $\tilde{u}(\boldsymbol{x}) = \tilde{w}(\boldsymbol{x}, \boldsymbol{y})$. We will further assume that $\tilde{w}(\boldsymbol{x}, \boldsymbol{y})$ is periodic in $\boldsymbol{y}$ with a corresponding unit-cell Y of size $\mathcal{O}(\varepsilon)$.

5.7. Discussion of assumptions

5.7.1. Discussion of the representative volume element (RVE)

(**A1**) The notion of RVE (sometimes also termed representative elementary volume (REV), see Technical note 6) is of particular importance to the volume averaging approach. In the definition, the term "representative" implies that only a relatively small part of Ω is representative of the entire domain, i.e., can be used to determine equivalent parameters associated with the macroscale equations. This notion is closely related to the quasi-stationarity of the parameter fields (see [72,73]), which can be defined intuitively in the following way. Consider the porosity of a porous medium consisting of a solid structure and a saturated fluid phase, γ. We can define the porosity in the averaging volume as the ratio of the fluid phase and the volume of the averaging set, $\frac{V_\gamma}{V}$. Further, we position the center of the averaging ball within the solid phase, so that for r sufficiently small, the porosity is equal to zero. The dependence of the porosity upon the radius of the averaging ball may exhibit various features, some of which are represented in Fig. 5.

On the left-hand side in Fig. 5, we observe the following behavior: below the minimum critical value $r < r_{\min}$ (dimensional coordinate system), the porosity fluctuates with spatial heterogeneities. In the range $r_{\min} \leqslant r \leqslant r_{\max}$, new heterogeneities are introduced within the averaging set but its volume is sufficiently large to smooth out fluctuations and the porosity becomes quasi-stationary. Beyond the maximum critical value $r > r_{\max}$, the system of interest exhibits larger scale heterogeneities that will induce a deviation from the quasi-stationary value. For such hierarchical systems, the RVE corresponds to choosing the radius, $r_0 = \frac{\ell}{2}$, of the averaging ball to lie in the range $r_{\min} \leqslant r_0 \leqslant r_{\max}$ and ℓ is the diameter of the RVE. In the volume averaging literature, this constraint for disordered media is often expressed by imposing $\mu \ll \ell$. We remark, however, that this constraint does not automatically apply to classes of ordered media, as illustrated in the periodic domain represented on the left-hand side of Fig. 6, for which $\ell = \mathcal{O}(\mu)$.

In the literature, the averaging volume is often defined *ab initio* as a RVE, so that simplifications can be made initially. If it exists, a RVE is a convenient choice of averaging volume because it simplifies mathematical developments and splits the problem into a hierarchy of scales by allowing a decomposition of the parameter fields into high and low spatial frequency components. However, RVEs are only a small subset of possible averaging volumes. *A priori*, we can average over any volume that suits our purposes, for instance via megascale averaging (see e.g., in [66]), even if a RVE does not exist for the system of interest.

There are situations for which the system of interest is not hierarchical, parameter fields are not quasi-stationary and RVEs cannot be defined. This is the case for the medium represented on the right-hand side of Fig. 5 for which characteristic length-scales evolve continuously. This is also the case for fractal porous media, which have received an increasing amount of

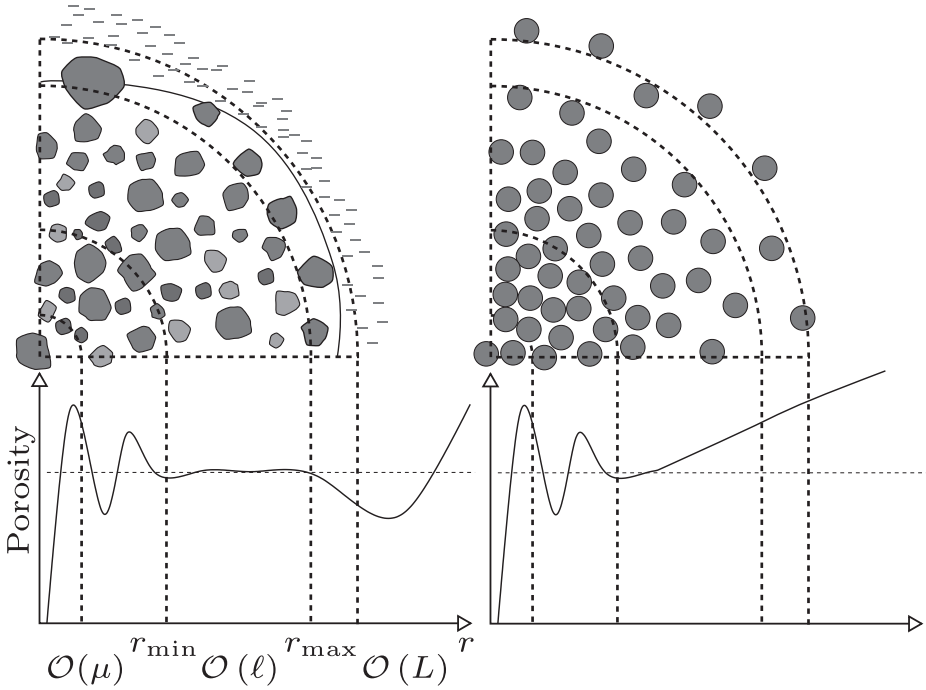
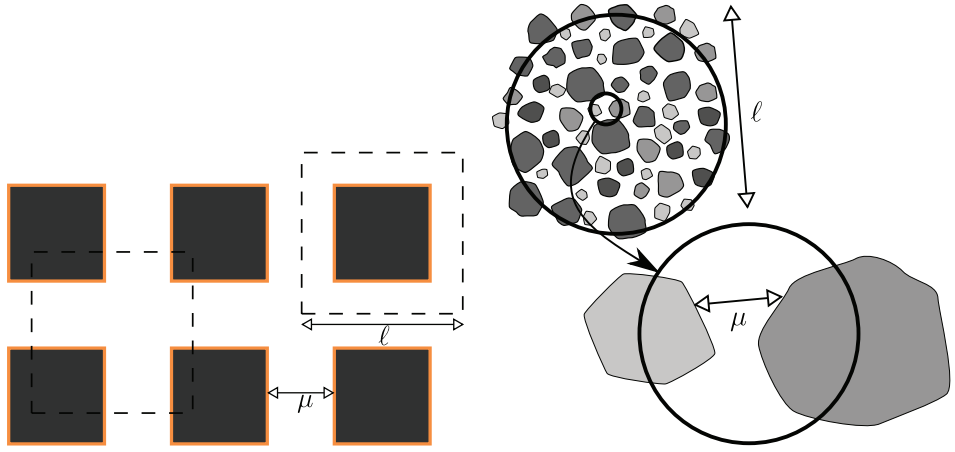


Fig. 5. Evolution of the porosity as a function of the radius, expressed in the dimensional coordinate system, of the averaging volume for hierarchical (left-hand side) and non-hierarchical (right-hand side) porous media. In the latter case, a RVE cannot be defined as characteristic lengthscales evolve continuously. For the hierarchical medium, $r_{\min}$ and $r_{\max}$ are defined as the bounds of the interval over which the porosity is quasi-stationary. A RVE is an averaging volume of radius r_0, such that $r_{\min} \leqslant r_0 \leqslant r_{\max}$. We will further denote, $\ell = 2r_0$, the diameter of this RVE.



(a) Case for which $\ell = \mathcal{O}(\mu)$. (b) Case for which $\ell \gg \mu$.

Fig. 6. Schematics illustrating the difference between the pore-scale characteristic length, μ, and the RVE diameter, ℓ, for a case where $\ell = \mathcal{O}(\mu)$ (left-hand side) and $\ell \gg \mu$ (right-hand side) in the dimensional coordinate system.

attention in the last couple of decades. Although the details are beyond the scope of this paper, it is important to mention that continuum average descriptions of such systems can be used in many instances. For example, we may exploit self-similarity or the appearance of cutoff lengths (e.g., in microvascular systems in [9]) in order to describe the non-hierarchical nature of the system via parameter gradients, e.g., volume fraction (see an example for dendritic solidification in [74] and mushy zones in [75]), permeability and dispersion coefficients, or nonlocal formulations involving fractional derivatives. We may also generate random fields of effective parameters to approximate spatial heterogeneities (see [76] for an example application to the volume averaging framework in the case of heterogeneous permeabilities, and [77] in the case of dispersion coefficients).

Technical note 6: RVE vs. REV In the literature, the RVE is often replaced by the notion of ''representative elementary volume'' (REV), for instance in most of Whitaker's work. The REV terminology introduces the concept of ''elementary'' volume which further constrains the range of radii allowed. ''Elementary'' implies that this is the smallest representative volume that can be defined. In Fig. 5, it means that the RVE can be anywhere inbetween $r_{\min}$ and $r_{\max}$, while the radius of the REV is equal to $r_{\min}$. In ordered porous media, for instance in the purely periodic case represented in Fig. 1, the REV can be uniquely and unambiguously defined. In practice, however, the REV may be difficult to determine as we often deal with disordered porous media for which the ''elementary'' size will depend on the process of interest. In such cases, the size of the REV will be different for different classes of effective parameters, say, permeability or volume fractions. A variety of tools may be used to characterize such real systems and determine the size of the REV directly from three-dimensional images, obtained for example using X-ray tomography. A non-exhaustive list includes correlation lengths, fractal dimensions, wavelet or Fourier analysis of a phase indicator function or numerical resolution of partial differential equations. An excellent generic review of methods that can be used to analyze X-ray tomography to quantify pore-scale structures can be found in [78].

5.7.2. Discussion of localization

A spatially *local* macroscale operator, $\mathcal{M}_l$, is one for which, for all $\boldsymbol{x} \in \Omega$, we can verify that $\mathcal{M}_l\langle u\rangle(\boldsymbol{x}, t) = 0$ in any arbitrarily small neighborhood of $\boldsymbol{x}$. A spatially *nonlocal* macroscale operator, $\mathcal{M}_n$, is one for which we need information at a large distance from $\boldsymbol{x}$ to verify $\mathcal{M}_n\langle u\rangle(\boldsymbol{x}, t) = 0$. An example case that illustrates "action at a distance" is the homogenization of random walks, including Lévy flights, yielding nonlocal diffusion at the macroscale, that may be described via fractional derivatives (see [79]). Another example for time nonlocality (see Technical note 7) involves history and temporal delays.

Technical note 7: *Nonlocal* models generally take the form of integrodifferential equations. For instance, consider the nonlocal time-dependent diffusion model described by

$$\begin{aligned}\partial_t\langle u\rangle &= \boldsymbol{\nabla}\cdot\left[\int_0^t \partial_\tau A(\tau)\boldsymbol{\nabla}\langle u\rangle(t-\tau)d\tau\right]\\ &= \boldsymbol{\nabla}\cdot(\partial_t A \bigstar_t \boldsymbol{\nabla}\langle u\rangle),\end{aligned} \tag{5.16}$$

where $\bigstar_t$ is a temporal convolution. We remark that the integration $\int_0^t$ is equivalent to the integration $\int_{-\infty}^{+\infty}$, when we weight functions in the integrand of the convolution by the unit-step function. This equation may be used to describe high frequency fluctuations for transport phenomena in porous media. The expression on the right-hand side involves a time convolution that accounts for history dependence. In other words, the rate of change term $\partial_t\langle u\rangle$ does not only depend on the divergence of the flux at time t but also at all other previous times ($\int_0^t(\cdot)d\tau$). Spatial nonlocal effects may be described in a similar way, using, for example, fractional derivatives (note that the fractional derivative is simply a nonlocal model with a specific type of kernel). Such models are generally not desirable because their resolution is computationally intensive. In the above example, the numerical resolution would require that all time steps are readily accessible in memory. For spatial nonlocality, this often means that long distance interactions occur, resulting in spatial discretization schemes involving dense matrices that are numerically intensive to invert. Other models, that are also termed nonlocal (although, strictly speaking, they are local), are higher order gradient theories. Here, the ''action at a distance'' is partly captured by higher (but finite) spatial order derivatives (e.g., consider the matrices with larger bands produced by spatial discretization of high order derivatives).

Localization assumptions are central to most upscaling methodologies and, in general, represent the strongest form of constraint upon the fields. Paradoxically, they may be relaxed to obtain nonlocal formulations. This paradox stems from the fact that nonlocal homogenized models have a broader domain of validity, but are more complex to derive and more intensive to compute. We must trade-off information for a smaller computational cost, but the extent to which that has to be done is a delicate choice which is both user- and problem-specific.

(**A2**) In the volume averaging literature, this assumption is often written in terms of the geometrical constraint, $\ell \ll L$ in dimensional form, or $\varepsilon \ll 1$ in nondimensional form. Loosely speaking, we often think about this constraint as: average quantities are approximately constant over the lengthscale of the RVE (the difficulty in this assertion being to clearly define what we mean by approximately). The idea is that there are primarily two lengthscales characterizing the spatial variations of the variables and therefore we use the average plus perturbation decomposition to separate high- and low-frequencies.

We emphasize that this constraint is only conceptual and does not reflect the process-dependence of the lengthscales. In other words, the lengthscales that matter are not those associated with the geometry of the system but with the spatial evolution of the variable fields. In many instances, the constraint $\varepsilon \ll 1$ is deemed sufficient since the characteristic length for microscale processes is $\mathcal{O}(\ell)$ while the characteristic length for macroscale processes is $\mathcal{O}(L)$, *but it is not always so simple*. A basic example is the propagation of a pulse through a porous medium for which the initial width of the pulse is close to zero and there is no obvious separation of scales (although, geometrically we still have $\varepsilon \ll 1$). Therefore, in (**A2**), we have defined the separation of scales via constraints applied to spatial frequencies that characterize the sig-

Technical note 8: The assumption (**A2**) has various consequences, including the important relationships $\langle\langle\boldsymbol{\psi}\rangle\rangle \simeq \langle\boldsymbol{\psi}\rangle$ and $\langle\tilde{\boldsymbol{\psi}}\rangle \simeq 0$. This can be readily demonstrated by considering the average of the perturbation decomposition equation (5.12)

$$\langle\boldsymbol{\psi}\rangle = \langle\langle\boldsymbol{\psi}\rangle\rangle + \langle\tilde{\boldsymbol{\psi}}\rangle \tag{5.17}$$

The average of the average, $\langle\langle\boldsymbol{\psi}\rangle\rangle$ is given by

$$\langle\langle\boldsymbol{\psi}\rangle\rangle = \frac{1}{V}\int_{\mathcal{V}}\langle\boldsymbol{\psi}\rangle(\boldsymbol{x}+\boldsymbol{\zeta}, t)\,dV_\zeta. \tag{5.18}$$

In this equation, we consider dimensionless variables $\boldsymbol{x}$ and $\boldsymbol{\zeta}$ so that $\|\boldsymbol{\zeta}\| = \mathcal{O}(\varepsilon)$. Therefore, we introduce the following scalings $\boldsymbol{\zeta} = \varepsilon\hat{\boldsymbol{\zeta}}$ and $V = \varepsilon^3\widehat{V}$ with $\left\|\hat{\boldsymbol{\zeta}}\right\| = \mathcal{O}(1)$ and $\widehat{V} = \mathcal{O}(1)$. With these notations, Eq. (5.18) becomes:

$$\langle\langle\boldsymbol{\psi}\rangle\rangle = \frac{1}{\widehat{V}}\int_{\mathcal{V}}\langle\boldsymbol{\psi}\rangle(\boldsymbol{x}+\varepsilon\hat{\boldsymbol{\zeta}}, t)\,dV_\zeta. \tag{5.19}$$

In (**A2**), we have assumed that $\varepsilon \ll 1$ so that we can Taylor expand about $\boldsymbol{x}$. We find that

$$\langle\langle\boldsymbol{\psi}\rangle\rangle = \langle\boldsymbol{\psi}\rangle + \varepsilon\boldsymbol{\nabla}_{\boldsymbol{x}}\langle\boldsymbol{\psi}\rangle\cdot\frac{1}{\widehat{V}}\int_{\mathcal{V}}\hat{\boldsymbol{\zeta}}\,dV_\zeta + O(\varepsilon^2). \tag{5.20}$$

Since the origin of the RVE is at its centroid and the medium is not perforated, we have $\int_{\hat{\mathcal{V}}} \hat{\boldsymbol{\xi}} dV_{\xi} = \mathbf{0}$, so that we obtain the result:

$$\langle\langle \boldsymbol{\psi} \rangle\rangle(\boldsymbol{x}, t) = \langle \boldsymbol{\psi} \rangle(\boldsymbol{x}, t) + O(\varepsilon^2). \tag{5.21}$$

We remark that substitution of this result back into Eq. (5.17) yields

$$\langle \tilde{\boldsymbol{\psi}} \rangle(\boldsymbol{x}, t) = O(\varepsilon^2). \tag{5.22}$$

Perforated domains may yield further difficulties as is discussed in [59].

nals $\langle u \rangle$ and $\tilde{u}$, not to the geometry. An important consequence of (**A2**) is discussed in Technical note 8.

We remark that (**A1**) and (**A2**) are similar in nature. The essential difference between these two assumptions is that (**A1**) applies to parameter fields (e.g., **A**), while (**A2**) applies to all tensor fields of interest (including **A** and the variable u). In practice, however, it is necessary to treat them separately, so that we can first determine the size of the RVE by studying *known* parameter fields, e.g., **A** or the porosity, and then use this RVE to impose constraints (**A2**) and (**A3**) on u.

5.7.3. Discussion of temporal quasi-stationarity

(**A3**) In the volume averaging literature, this assumption is known as the quasi-stationarity of the perturbation problem, Eq. (5.13). Its physical basis is that microscale processes generally relax much faster than macroscale ones. Hence, the perturbation problem can often be considered steady while evolution still occurs at the macroscale. In our problem, we may assume that microscale and macroscale timescales are $\mathcal{O}\left(\frac{l^2}{\mathsf{A}}\right)$ and $\mathcal{O}\left(\frac{L^2}{\mathsf{A}}\right)$, respectively wherein the notation A represents a norm of an average value of the tensor **A**. In nondimensional form, with $T = \frac{L^2}{\mathsf{A}}$, this means that micro- and macro-timescales are $\mathcal{O}(\varepsilon^2)$ and $\mathcal{O}(1)$, respectively. Hence, we may write $\delta = \mathcal{O}(\varepsilon^2)$, which is often termed a diffusive scaling. In Technical note 9, we discuss how the assumption of quasi-stationarity may be relaxed.

Technical note 9: The assumption (**A3**) is not mandatory, i.e., the perturbation problem may be treated in a completely transient form. However, relaxing (**A3**) yields a structure of the perturbation that ''mixes'' the micro- and macro- timescales via temporal convolutions, as discussed in Technical note 7. This may be interpreted as a Green's function solution of the perturbation problem. The resulting macroscale equation will therefore involve temporal nonlocal effects (see discussions in [80–82]). In this context, the quasi-stationarity assumption may be formalized by studying the correlation structure of the kernels (see [71]). Loosely speaking, this corresponds to a kernel that relaxes to equilibrium with a timescale that is much faster than the macro-timescale.

5.7.4. Discussion of unit-cell and periodicity

(**A4**) In addition to the RVE, we introduce here the (different) concept of the unit-cell. The unit-cell can be extracted directly from images of real porous media and, in most cases, is of size $\mathcal{O}(\varepsilon)$. However, it does not necessarily have to be a subset of Ω. Rather, it represents a conceptualized medium which is utilized to capture the most important topological features of the real system. For instance, it may be composed artificially from statistical properties of the porous medium.

Further, we treat $\tilde{u}(\boldsymbol{x})$ as a function of the micro- ($\boldsymbol{y}$) and macroscale ($\boldsymbol{x}$) variables, $\tilde{u}(\boldsymbol{x}) = \tilde{w}(\boldsymbol{x}, \boldsymbol{y})$ and suppose that $\tilde{w}$ is periodic in $\boldsymbol{y}$ (see Technical note 10). To build a periodic geometry from a non-periodic one, we may use symmetry operations, i.e., reflect copies of the original unit cell, although this may lead to a critical loss of information such as anisotropic features. We may also treat $\tilde{w}$ as periodic in $\boldsymbol{y}$ with a non-periodic geometry and parameter jumps on the interfaces, although this may be erroneous if the system is close to the percolation threshold (see Fig. 7 for an illustration of a fractured porous medium). Other boundary conditions can be used, such as linear gradients (see discussions and references in [83,84]).

Finally, we will consider that, for quantities that are defined in the periodic unit-cell, the spatial averaging operator $\langle \bullet \rangle$ corresponds to averaging over the unit-cell, independent of the initial averaging support and the geometry of the unit-cell.

Technical note 10: The periodicity condition in (**A4**) has strong limitations. If we consider Technical note 5, this means that the macroscale boundary conditions have a negligible impact upon the Green's function corresponding to Eqs. (4.6)–(4.8). The idea that the perturbation is only weakly coupled to the macroscale boundary conditions is closely linked with the separation of scales and the fact that we consider $\varepsilon \ll 1$. We remark, however, that this is not always correct, a typical example being an advection–diffusion problem for which the microscale Péclet number scales as ε^{-1}.

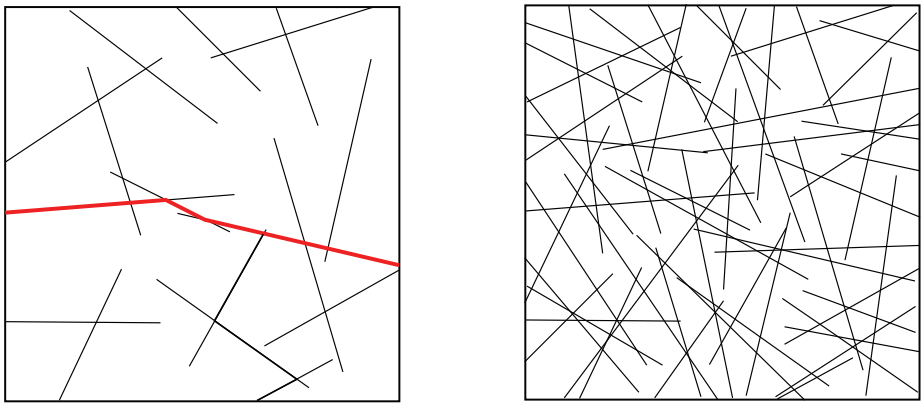


Fig. 7. A conceptual representation of a fractured porous medium with low density of fractures on the left-hand side and a larger density of fractures on the right-hand side. We remark that for the case on the left-hand side, the intersections of fractures with boundaries do not match between the left/right and top/bottom boundaries, whereas some do for the case on the right-hand side. Therefore, application of periodic conditions for flow in such systems will eliminate percolation (red thicker line, single preferential flow) on the left-hand side while it may preserve the correct percolation properties on the right-hand side (much larger percolation). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

5.8. Simplifications to Eq. (5.13)

We recall that the perturbation problem Eqs. (5.13) to (5.15) is

$$\partial_t \tilde{u} = \nabla \cdot (\mathbf{A} \cdot \nabla \tilde{u}) - \nabla \cdot \langle \mathbf{A} \cdot \nabla \tilde{u} \rangle + \nabla \cdot (\mathbf{A} \cdot \nabla \langle u \rangle) - \nabla \cdot \langle \mathbf{A} \cdot \nabla \langle u \rangle \rangle, \tag{5.23}$$

with conditions

$$\tilde{u} = f - \langle u \rangle, \tag{5.24}$$

$$\tilde{u}(\boldsymbol{x}, 0) = 0. \tag{5.25}$$

Further, recall that our goal is to simplify further Eqs. (5.23) to (5.25), so that we may obtain an approximate solution for the perturbation. A rigorous derivation of this solution uses order of magnitudes estimate of each term as powers of ε, and is developed in Appendix A. For simplicity, we will only present a simplified derivation using the approximation symbol, $\simeq$, to "hide" the tedious analysis of higher order terms. Assumptions (**A1**) to (**A4**) are then used as follows:

- Application of (**A1**): We first decompose $\mathbf{A} = \langle \mathbf{A} \rangle + \tilde{\mathbf{A}}$. In (**A1**), we assume that our system is hierarchical so that we can use a RVE as the averaging volume. This means that $\langle \mathbf{A} \rangle$ varies with a characteristic lengthscale that is $\mathcal{O}(1)$ while $\tilde{\mathbf{A}}$ captures higher frequencies. To further simplify the example, we have imposed in Section 4.4 that $\langle \mathbf{A} \rangle$ does not vary at all, i.e., $\langle \mathbf{A} \rangle(\boldsymbol{x}) \equiv$ constant $\equiv \langle \mathbf{A} \rangle$.
- Application of (**A2**): With the size of the averaging volume fixed by (**A1**), we now apply (**A2**) that yields $\nabla_{\boldsymbol{y}} \langle u \rangle \simeq 0$, $\nabla \cdot \langle \mathbf{A}(\boldsymbol{x}) \cdot \nabla \langle u \rangle \rangle \simeq \nabla \cdot (\langle \mathbf{A} \rangle \cdot \nabla \langle u \rangle)$ and $\langle \tilde{u} \rangle \simeq 0$ (see Technical note 8, loosely speaking we may say that average quantities do not vary *over the lengthscale of the RVE* and can be treated as constant). Therefore, we obtain

$$\nabla \cdot (\mathbf{A} \cdot \nabla \langle u \rangle) - \nabla \cdot \langle \mathbf{A} \cdot \nabla \langle u \rangle \rangle \simeq \nabla \cdot \big(\tilde{\mathbf{A}} \cdot \nabla \langle u \rangle\big). \tag{5.26}$$

We can further simplify this expression by neglecting higher-order terms, e.g., $\nabla \nabla \langle u \rangle$ (see Appendix A for a more rigorous approach). This yields

$$\begin{aligned} \nabla \cdot \big(\tilde{\mathbf{A}} \cdot \nabla \langle u \rangle\big) &= \tilde{\mathbf{A}} : \nabla \nabla \langle u \rangle + \big(\nabla \cdot \tilde{\mathbf{A}}\big) \cdot \nabla \langle u \rangle \\ &\simeq \big(\nabla \cdot \tilde{\mathbf{A}}\big) \cdot \nabla \langle u \rangle. \end{aligned} \tag{5.27}$$

Therefore, Eq. (5.23) simplifies to

$$\partial_t \tilde{u} = \nabla \cdot (\mathbf{A} \cdot \nabla \tilde{u}) - \nabla \cdot \langle \mathbf{A} \cdot \nabla \tilde{u} \rangle + \big(\nabla \cdot \tilde{\mathbf{A}}\big) \cdot \nabla \langle u \rangle. \tag{5.28}$$

- Application of (**A3**): On imposing quasi-stationarity, we can write Eq. (5.28) as

$$-\big(\nabla \cdot \tilde{\mathbf{A}}\big) \cdot \nabla \langle u \rangle = \nabla \cdot (\mathbf{A} \cdot \nabla \tilde{u}) - \nabla \cdot \langle \mathbf{A} \cdot \nabla \tilde{u} \rangle. \tag{5.29}$$

Finally, we write this equation in the microscale coordinate system:

$$-\nabla_{\boldsymbol{y}} \cdot \big(\tilde{\mathbf{A}}\big) \cdot \varepsilon \nabla \langle u \rangle = \nabla_{\boldsymbol{y}} \cdot \big(\mathbf{A} \cdot \nabla_{\boldsymbol{y}} \tilde{u}\big) - \nabla_{\boldsymbol{y}} \cdot \langle \mathbf{A} \cdot \nabla_{\boldsymbol{y}} \tilde{u} \rangle. \tag{5.30}$$

5.9. Structure of the solution for the perturbation (closure)

In what follows, it will be convenient to assume that there exist $\boldsymbol{b}(\boldsymbol{x})$ and $\boldsymbol{c}(\boldsymbol{y})$ such that

$$\tilde{u} = \boldsymbol{b}(\boldsymbol{x}) \cdot \nabla \langle u \rangle(\boldsymbol{x}),$$

and

$$\tilde{u} = \varepsilon \boldsymbol{c}(\boldsymbol{y}) \cdot \nabla \langle u \rangle(\boldsymbol{x}). \tag{5.31}$$

We have not formulated these solutions arbitrarily, but rather are proposing a solution that exploits the linearity of the problem to separate contributions at the micro- and macroscale (reminiscent of a standard separation of variables). In other words, we have expressed the perturbations as $\tilde{u} = \boldsymbol{\pi}(\boldsymbol{y}) \cdot \boldsymbol{g}(\boldsymbol{x})$. Introducing Eq. (5.31) into Eq. (5.30), we obtain the following equation for $\boldsymbol{c}(\boldsymbol{y})$

$$-\nabla_{\boldsymbol{y}} \cdot \tilde{\mathbf{A}} = \nabla_{\boldsymbol{y}} \cdot \big(\mathbf{A} \cdot \nabla_{\boldsymbol{y}} \boldsymbol{c}\big) - \nabla_{\boldsymbol{y}} \cdot \big\langle \mathbf{A} \cdot \nabla_{\boldsymbol{y}} \boldsymbol{c} \big\rangle. \tag{5.32}$$

- Application of (**A4**): We now assume that Eq. (5.32) can be solved over a unit-cell in which we replace Eq. (5.24) by periodic boundary conditions. The operator $\langle \bullet \rangle$ now refers to the average over the periodic unit-cell. This yields $\nabla_{\boldsymbol{y}} \cdot \big\langle \mathbf{A} \cdot \nabla_{\boldsymbol{y}} \boldsymbol{c} \big\rangle = 0$ (periodic boundary conditions) and

$$-\nabla_{\boldsymbol{y}} \cdot \tilde{\mathbf{A}} = \nabla_{\boldsymbol{y}} \cdot \big(\mathbf{A} \cdot \nabla_{\boldsymbol{y}} \boldsymbol{c}\big), \tag{5.33}$$

with the condition,

$$\text{Periodicity in } \boldsymbol{y}. \tag{5.34}$$

Finally, we remark that the solution of this problem is unique up to a constant. We recall that $\langle \tilde{u} \rangle \simeq 0$ (see Technical note 8), so that we fix the constant by imposing

$$\langle \boldsymbol{c} \rangle = 0. \tag{5.35}$$

5.9.1. Macroscale equation

Recall that the unclosed Eq. (5.12) has the following form

$$\partial_t \langle u \rangle = \nabla \cdot \langle \mathbf{A} \cdot (\nabla \langle u \rangle + \nabla \tilde{u}) \rangle. \tag{5.36}$$

Applying Eq. (5.31) supplies

$$\nabla \tilde{u} = \varepsilon^{-1} \nabla_{\boldsymbol{y}} \tilde{u} \simeq \nabla_{\boldsymbol{y}} \boldsymbol{c} \cdot \nabla \langle u \rangle, \tag{5.37}$$

and we obtain the following closed form of the macroscale equation:

$$\partial_t \langle u \rangle = \nabla \cdot (\mathbf{A}_e \cdot \nabla \langle u \rangle), \tag{5.38}$$

with

$$\mathbf{A}_e = \big\langle \mathbf{A} \cdot \big(\mathbf{I} + \nabla_{\boldsymbol{y}} \boldsymbol{c}\big) \big\rangle. \tag{5.39}$$

6. Homogenization via multiscale asymptotics

We now focus on homogenization of the same parabolic problem via multiscale asymptotics.

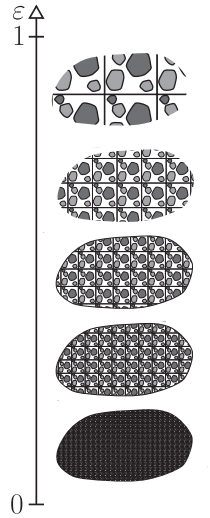


Fig. 8. Illustration of a sequence of geometries associated to a variation of ε for a fixed value of the macroscopic lengthscale, L.

6.1. Assumptions

When using an asymptotic approach, we start by making a series of assumptions that enable us to cast the problem in a more convenient form. These can be summarized as follows, with subscripts ε to distinguish them from the volume averaging assumptions:

6.1.1. One physical problem as a fictitious sequence of problems

$(\mathbf{A1})_\varepsilon$: We homogenize by considering the limit $\varepsilon \to 0$ of a sequence of problems, $(u_\varepsilon)_{0<\varepsilon<1}$, rather than the problem defined by Eqs. (4.9) to (4.11) which implicitly contains a fixed, finite (but small) value of the scales ratio, ε_0. We use a similar notation for the corresponding sequence of parameter fields.

6.1.2. Scalings

$(\mathbf{A2})_\varepsilon$: In the current nondimensional setting, $\mathsf{A}_{\varepsilon ij}$ scales as $\mathcal{O}(1)$ and we can write $\mathbf{A}_\varepsilon = \mathbf{A}_0$ with $\mathsf{A}_{0ij} = \mathcal{O}(1)$. In addition, we recall that in SubSection 4.4 we assumed that $\mathbf{A}_0$ exhibits only high-frequency fluctuations, i.e., $\mathbf{A}_0 = \mathbf{A}_0\left(\frac{\boldsymbol{x}}{\varepsilon}\right)$.

6.1.3. Formal two-scale expansion

$(\mathbf{A3})_\varepsilon$: We assume that we can use the series approximation $u_\varepsilon(\boldsymbol{x}) = u_\varepsilon^\dagger\left(\boldsymbol{x}, \frac{\boldsymbol{x}}{\varepsilon}\right) = \sum u_i\left(\boldsymbol{x}, \frac{\boldsymbol{x}}{\varepsilon}\right)\varepsilon^i$, with $u_\varepsilon : \mathbb{R}^n \to \mathbb{R}$ and $u_\varepsilon^\dagger : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$. It is implicit (but fundamental) in this decomposition that $u_i = \mathcal{O}(1)$ for $i = 0, 1, 2, \ldots$.

6.1.4. Unit-cell and periodicity

$(\mathbf{A4})_\varepsilon$: We consider that our physical problem is periodic or that it can be approximated by a conceptual periodic setting. The fundamental periodic unit is referred to as the unit-cell. Further, we assume that $u_i(\boldsymbol{x}, \boldsymbol{y})$ is periodic in $\boldsymbol{y}$ for $i = 0, 1, 2, \ldots$.

6.2. Discussion of assumptions

$(\mathbf{A1})_\varepsilon$ Eqs. (4.9) to (4.11) were written for a fixed geometry and therefore a fixed value of ε. In multiscale asymptotics, we are interested in the sequence $(u_\varepsilon)_{0<\varepsilon<1}$ in the limit $\varepsilon \to 0$. This sequence corresponds to varying geometrical properties of the system, as illustrated in Fig. 8 for a fixed value of L. We remark further that our physical problem of interest corresponds to a fixed finite (small) value of ε, so that by considering the limit $\varepsilon \to 0$, we will obtain only an asymptotic approximation.

$(\mathbf{A2})_\varepsilon$ In the current setting, we require that $\mathsf{A}_{\varepsilon ij} = \mathcal{O}(1)$. The system was nondimensionalized with the macroscopic lengthscale, L, and $\mathsf{A} = \mathcal{O}(1)$ means that we use the timescale $T = L^2/\mathsf{A}$ in Section 4.4 for global diffusion (a more general approach is discussed in Technical note 11). This is a natural choice for a diffusive system for which we seek the global behavior. More generally, the choice of nondimensionalization and scalings are both subtle and fundamentally important for the multiscale asymptotic method. The scaling of all dimensionless parameters as series of ε must be decided carefully: alternative choices lead to different macroscale equations. Although in the current setting, our attention is restricted to diffusive transport, the challenges associated with scaling can be well illustrated for an advection–diffusion problem (see Technical note 12). In practice, each scaling decision is equivalent to choosing the timescale of interest (e.g. that of diffusion or advection) on the macroscopic lengthscale. Whilst the multiscale asymptotics method cannot, therefore, be applied to a system with arbitrary physical properties without defining their relative importance, it lends itself naturally to approaches in applied modeling and engineering.

Technical note 11: There are two natural timescales in our problem: the timescale for transport at the macro- and micro-scales (there may be more, see [47]). We expect there to be orders of magnitude separation between these characteristic times since they depend on the micro- and macro-lengthscales and we have $\varepsilon \ll 1$. Therefore, we may define a second small parameter as ε^θ with $\theta \in \mathbb{N}$ and, as for the spatial operators, we decompose time so that $t \to (t, \tau = \varepsilon^{-\theta} t)$. Further, we may use $u_\varepsilon(\boldsymbol{x}, t) = u_\varepsilon^\dagger(\boldsymbol{x}, t, \tau)$, so that $\partial_t u_\varepsilon(\boldsymbol{x}, t) = \partial_t u_\varepsilon^\dagger(\boldsymbol{x}, t, \tau) + \varepsilon^{-\theta}\partial_\tau u_\varepsilon^\dagger(\boldsymbol{x}, t, \tau)$. Performing such developments for temporal derivatives is consistent with the spatial analysis (for a detailed discussion, see [47]).

Technical note 12: Consider a situation where transport occurs by advection and diffusion (coefficient D constant) in a perforated (porous) medium (for a detailed discussion of advection–diffusion problems, see [52]). We often characterize transport by defining local ($Pe_\ell = U\ell/D$) and global ($Pe_L = UL/D$) Péclet numbers with $Pe_L = \varepsilon^{-1} Pe_\ell$ (the Péclet number describes the ratio of diffusive to advective timescales and takes the form Ud/D where U is a typical flow velocity, d the lengthscale of interest and D a diffusion coefficient). We consider the following cases:

- If diffusion is much stronger than advection locally ($Pe_\ell = \mathcal{O}(\varepsilon^2)$) then it also dominates at the macroscale ($Pe_L = \mathcal{O}(\varepsilon)$). This situation yields at the macroscale a dispersion tensor with only tortuosity effects.
- If diffusion dominates advection locally ($Pe_\ell = \mathcal{O}(\varepsilon)$) then diffusion and advection are equally important at the macroscale ($Pe_L = \mathcal{O}(1)$). This situation yields at the macroscale an advection–dispersion equation with a dispersion tensor exhibiting only tortuosity effects.
- If diffusion and advection are in balance at the microscale ($Pe_\ell = \mathcal{O}(1)$) then advection is dominant at the macroscale ($Pe_L = \mathcal{O}(\varepsilon^{-1})$). In this case, we obtain an advection–dispersion at the macroscale and the dispersion tensor exhibits tortuosity and hydrodynamic dispersion effects.

$(\mathbf{A3})_\varepsilon$ There are two aspects to this approximation. Firstly, we transform the functional $u_\varepsilon : \mathbb{R}^n \to \mathbb{R}$, into a new functional $u_\varepsilon^\dagger : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ (see Technical note 13 for a simple example). The idea is that u_ε exhibits variations at two distinct lengthscales, characterized by $\boldsymbol{x}$ and $\varepsilon^{-1}\boldsymbol{x}$ respectively so that we can separate both contributions by creating a new function $u_\varepsilon^\dagger$ that depends on these two variables. This is a key step in the method. In particular, the chain rule yields

$$\boldsymbol{\nabla} = \boldsymbol{\nabla}_{\boldsymbol{x}} + \varepsilon^{-1}\boldsymbol{\nabla}_{\boldsymbol{y}}, \tag{6.1}$$

$$\boldsymbol{\nabla}\cdot = \boldsymbol{\nabla}_{\boldsymbol{x}}\cdot + \varepsilon^{-1}\boldsymbol{\nabla}_{\boldsymbol{y}}\cdot, \tag{6.2}$$

$$\boldsymbol{\nabla}\cdot\boldsymbol{\nabla} = \boldsymbol{\nabla}_{\boldsymbol{x}}\cdot\boldsymbol{\nabla}_{\boldsymbol{x}} + \varepsilon^{-1}\left(\boldsymbol{\nabla}_{\boldsymbol{y}}\cdot\boldsymbol{\nabla}_{\boldsymbol{x}} + \boldsymbol{\nabla}_{\boldsymbol{x}}\cdot\boldsymbol{\nabla}_{\boldsymbol{y}}\right) + \varepsilon^{-2}\boldsymbol{\nabla}_{\boldsymbol{y}}\cdot\boldsymbol{\nabla}_{\boldsymbol{y}}. \tag{6.3}$$

Technical note 13: Consider the following function

$$u_\varepsilon : \mathbb{R} \to \mathbb{R},$$
$$x \mapsto \alpha x + \varepsilon \sin\left(\varepsilon^{-1} x\right). \tag{6.4}$$

We define $u_\varepsilon^\dagger$ by

$$u_\varepsilon^\ddagger : \mathbb{R}\times\mathbb{R}\to\mathbb{R},\quad x,y(x)\mapsto \alpha x+\varepsilon\sin(y(x)). \tag{6.5}$$

So that

$$u_\varepsilon(x)=u_\varepsilon^\ddagger(x,y(x)=\varepsilon^{-1}x). \tag{6.6}$$

Now, it follows that

$$\frac{du_\varepsilon}{dx}(x)=\frac{\partial u_\varepsilon^\ddagger}{\partial x}(x,\varepsilon^{-1}x)+\varepsilon^{-1}\frac{\partial u_\varepsilon^\ddagger}{\partial y}(x,\varepsilon^{-1}x) =\alpha+\cos(\varepsilon^{-1}x). \tag{6.7}$$

We further remark that we have not assumed that x and y are independent.

The second step of this approximation is to assume the two-scale expansion

$$u_\varepsilon^\ddagger\Big(\boldsymbol{x},\frac{\boldsymbol{x}}{\varepsilon}\Big)=\sum_i u_i\Big(\boldsymbol{x},\frac{\boldsymbol{x}}{\varepsilon}\Big)\varepsilon^i. \tag{6.8}$$

The decomposition is similar to that used for the decomposition in terms of $\boldsymbol{x}$ and $\frac{\boldsymbol{x}}{\varepsilon}$, but is useful when considering partial derivatives. Loosely speaking, the order of magnitude separation in the spatial frequencies can be captured via an order of magnitude separation in the amplitude of the derivatives.

$(\mathbf{A4})_\varepsilon$ We limit our analysis to periodic domains, i.e., assume that the system is periodic or that the periodic setting constitutes a reasonable approximation (see also Technical note 10). We will consider spatial averages as

$$\langle\psi\rangle_\varepsilon(\boldsymbol{x},t)=\frac{1}{V_\varepsilon}\int_{\boldsymbol{y}\in\mathcal{V}_\varepsilon(\boldsymbol{x})}\psi_\varepsilon(\boldsymbol{x},\boldsymbol{y},t)dV_{\boldsymbol{y}}, \tag{6.9}$$

where $\mathcal{V}_\varepsilon(\boldsymbol{x})$ is a sequence of sets with origin at its centroid $\boldsymbol{x}$, and V_ε is the corresponding sequence of volumes. Further, since we have defined the periodic domain *a priori*, the size of the averaging volume corresponds to the size of the unit-cell *and the variable of integration is* $\boldsymbol{y}$, *not* $\boldsymbol{x}$. We complete the concept of periodicity of a function $h(\boldsymbol{x},\boldsymbol{y})$ in $\boldsymbol{y}$ by imposing continuity (no jump) in the value of the function or its flux, $\boldsymbol{n}\cdot(\mathbf{A}_\varepsilon\cdot\nabla_{\boldsymbol{y}}h)$, where $\boldsymbol{n}$ is the unit outward facing normal to the periodic boundary across the periodic unit. As a result, we have

$$\big\langle\nabla_{\boldsymbol{y}}\cdot(\mathbf{A}_\varepsilon\cdot\nabla_{\boldsymbol{y}}h)\big\rangle_\varepsilon=0. \tag{6.10}$$

6.3. Derivation

Our goal here is to derive an initial boundary value problem for u_0, the first term in the asymptotic expansion for u.

- Application of $(\mathbf{A1})_\varepsilon$: We rewrite Eqs. (4.9)–(4.11) as

$$\partial_t u_\varepsilon=\nabla\cdot(\mathbf{A}_\varepsilon\cdot\nabla u_\varepsilon), \tag{6.11}$$

$$u_\varepsilon=f_\varepsilon, \tag{6.12}$$

$$u_\varepsilon(\boldsymbol{x},0)=0. \tag{6.13}$$

- Application of $(\mathbf{A2})_\varepsilon$: Assuming that $\mathbf{A}_\varepsilon=\mathbf{A}_0=\mathcal{O}(\mathbf{1})$ yields

$$\partial_t u_\varepsilon=\nabla\cdot(\mathbf{A}_0\cdot\nabla u_\varepsilon). \tag{6.14}$$

- Application of $(\mathbf{A3})_\varepsilon$: Formally applying Eqs. (6.1)–(6.3) to Eq. (6.11) yields

$$\partial_t u_\varepsilon=\varepsilon^{-2}\nabla_{\boldsymbol{y}}\cdot(\mathbf{A}_0\cdot\nabla_{\boldsymbol{y}}u_\varepsilon) +\varepsilon^{-1}\big[\nabla_{\boldsymbol{x}}\cdot(\mathbf{A}_0\cdot\nabla_{\boldsymbol{y}}u_\varepsilon)+\nabla_{\boldsymbol{y}}\cdot(\mathbf{A}_0\cdot\nabla_{\boldsymbol{x}}u_\varepsilon)\big] +\nabla_{\boldsymbol{x}}\cdot(\mathbf{A}_0\cdot\nabla_{\boldsymbol{x}}u_\varepsilon). \tag{6.15}$$

We then use the two-scale expansion up to second order, i.e., $u_\varepsilon(\boldsymbol{x})=u_0(\boldsymbol{x},\frac{\boldsymbol{x}}{\varepsilon})+\varepsilon u_1(\boldsymbol{x},\frac{\boldsymbol{x}}{\varepsilon})+\varepsilon^2u_2(\boldsymbol{x},\frac{\boldsymbol{x}}{\varepsilon})+h.o.t.$. Substituting this expression into Eq. (6.15), we obtain a coupled problem for the components u_i. *Now that we have explicitly expressed the complete problem in terms of powers of* ε, *we can consider the limit* $\varepsilon\to0$ *and, in so doing, decompose the unique problem into a sequence of nested sub-problems*. Equating coefficients of $\mathcal{O}(\varepsilon^{-2})$ in Eq. (6.15) yields

$$\nabla_{\boldsymbol{y}}\cdot(\mathbf{A}_0\cdot\nabla_{\boldsymbol{y}}u_0)=0. \tag{6.16}$$

Equating coefficients of $\mathcal{O}(\varepsilon^{-1})$ in Eq. (6.15) yields the unit-cell problem

$$\nabla_{\boldsymbol{y}}\cdot(\mathbf{A}_0\cdot\nabla_{\boldsymbol{y}}u_1)+\nabla_{\boldsymbol{x}}\cdot(\mathbf{A}_0\cdot\nabla_{\boldsymbol{y}}u_0)+\nabla_{\boldsymbol{y}}\cdot(\mathbf{A}_0\cdot\nabla_{\boldsymbol{x}}u_0)=0. \tag{6.17}$$

Equating coefficients of $\mathcal{O}(\varepsilon^0)$ in Eq. (6.15) yields the equation

$$\partial_t u_0=\nabla_{\boldsymbol{x}}\cdot(\mathbf{A}_0\cdot\nabla_{\boldsymbol{x}}u_0)+\nabla_{\boldsymbol{y}}\cdot(\mathbf{A}_0\cdot\nabla_{\boldsymbol{y}}u_2) +\nabla_{\boldsymbol{x}}\cdot(\mathbf{A}_0\cdot\nabla_{\boldsymbol{y}}u_1)+\nabla_{\boldsymbol{y}}\cdot(\mathbf{A}_0\cdot\nabla_{\boldsymbol{x}}u_1). \tag{6.18}$$

Henceforth, we will consider that $\boldsymbol{x}$ and $\boldsymbol{y}$ can be treated as independent variables.

- Application of $(\mathbf{A4})_\varepsilon$: By periodicity, the problem equation (6.16) only admits solutions of the form, $u_0=\bar{u}(\boldsymbol{x})$, i.e., u_0 only varies on the macro-lengthscale $\boldsymbol{x}$ and does not depend on the micro-lengthscale $\mathbf{y}$ (for a proof, see Technical note 14).

Technical note 14: To simplify presentation, we will limit our analysis to the case $\mathbf{A}_0=\mathrm{A}_0\mathbf{I}$ where $\mathbf{I}$ is the identity. From Eq. (6.16) we deduce

$$\big\langle u_0\nabla_{\boldsymbol{y}}\cdot(\mathrm{A}_0\nabla_{\boldsymbol{y}}u_0)\big\rangle_\varepsilon=0, \tag{6.19}$$

which may be rewritten as

$$\big\langle\nabla_{\boldsymbol{y}}\cdot(u_0\mathrm{A}_0\nabla_{\boldsymbol{y}}u_0)\big\rangle_\varepsilon-\Big\langle\mathrm{A}_0\big\|\nabla_{\boldsymbol{y}}u_0\big\|^2\Big\rangle_\varepsilon=0. \tag{6.20}$$

The periodicity assumption in $(\mathbf{A4})_\varepsilon$ guarantees that the first term in Eq. (6.20) is zero. Accordingly, we see that satisfying Eq. (6.20) is equivalent to requiring $\|\nabla_{\boldsymbol{y}}u_0\|=0$, and so $u_0=\bar{u}(\boldsymbol{x})$.

With $u_0=\bar{u}(\boldsymbol{x})$, Eq. (6.17) simplifies to

$$\nabla_{\boldsymbol{y}}\cdot(\mathbf{A}_0\cdot\nabla_{\boldsymbol{y}}u_1)+\nabla_{\boldsymbol{y}}\cdot(\mathbf{A}_0\cdot\nabla_{\boldsymbol{x}}\bar{u})=0, \tag{6.21}$$

together with periodicity of u_1. We can also rewrite Eq. (6.18) as follows

$$\partial_t\bar{u}=\nabla_{\boldsymbol{x}}\cdot(\mathbf{A}_0\cdot\nabla_{\boldsymbol{x}}\bar{u})+\nabla_{\boldsymbol{y}}\cdot(\mathbf{A}_0\cdot\nabla_{\boldsymbol{y}}u_2)+\nabla_{\boldsymbol{x}}\cdot(\mathbf{A}_0\cdot\nabla_{\boldsymbol{y}}u_1) +\nabla_{\boldsymbol{y}}\cdot(\mathbf{A}_0\cdot\nabla_{\boldsymbol{x}}u_1). \tag{6.22}$$

Technical note 15: Solvability conditions via the Fredholm Alternative. To formally seek solutions to the systems of equations that arise when we equate coefficients of powers of ε, we could introduce an invariant distribution $\rho(\boldsymbol{x},\boldsymbol{y})$ that solves the homogeneous self-adjoint counterpart to Eq. (6.16). For the simplified system considered in this paper, the local operator is $L\bullet:=\nabla_{\boldsymbol{y}}\cdot(\mathbf{A}_0\cdot\nabla_{\boldsymbol{y}}\bullet)$, equipped with periodic boundary conditions. This is a self–adjoint system, and so the invariant distribution satisfies $L^*\rho=\nabla_{\boldsymbol{y}}\cdot(\mathbf{A}_0\cdot\nabla_{\boldsymbol{y}}\rho)=0$, with periodic boundary conditions. As for the $\mathcal{O}(\varepsilon^{-2})$ system defined by

Eq. (6.16), the only solution for the invariant distribution is that $\rho(\boldsymbol{x})$ is locally constant.

Finally, we remark that subsequent systems of equations that arise when we equate coefficients of higher powers of ε (for example, Eqs. 6.17,6.18) may be written as non-homogeneous systems involving the same operator in the form

$$Lu_i = \nabla_{\boldsymbol{y}} \cdot (\mathbf{A}_0 \cdot \nabla_{\boldsymbol{y}} u_i) = g_i, \tag{6.23}$$

(where the index i indicates the system of equations that arise when we equate coefficients of powers of ε^{i-2}). By the Fredholm alternative, the solvability condition for these systems takes the form

$$\langle \rho L u_i \rangle_\varepsilon - \langle u_i L^* \rho \rangle_\varepsilon = \langle \rho g_i \rangle_\varepsilon, \tag{6.24}$$

or, equivalently, for the current example

$$\left\langle \rho \nabla_{\boldsymbol{y}} \cdot (\mathbf{A}_0 \cdot \nabla_{\boldsymbol{y}} u_i) \right\rangle_\varepsilon - \left\langle u_i \nabla_{\boldsymbol{y}} \cdot (\mathbf{A}_0 \cdot \nabla_{\boldsymbol{y}} \rho) \right\rangle_\varepsilon = \langle \rho g_i \rangle_\varepsilon. \tag{6.25}$$

We note that for the example problem under consideration, the second term on the left-hand side of Eq. (6.25) is zero as the invariant distribution ρ is locally constant. Therefore, applying the solvability condition is equivalent to averaging (as defined by Eq. (6.9)) the non-homogeneous equations (6.23) that arise when we equate coefficients of each power of ε. For more general systems of partial differentiation equations, the relationship between averaging and solvability is defined by Eq. (6.24).

The macroscale equation is derived by averaging Eq. (6.22) over the unit-cell, which in this scenario is equivalent to introducing an invariant distribution and applying the Fredholm alternative (see Technical note 15). This yields

$$\begin{aligned}\langle \partial_t \bar{u} \rangle_\varepsilon = {} & \langle \nabla_{\boldsymbol{x}} \cdot (\mathbf{A}_0 \cdot \nabla_{\boldsymbol{x}} \bar{u}) \rangle_\varepsilon + \left\langle \nabla_{\boldsymbol{y}} \cdot (\mathbf{A}_0 \cdot \nabla_{\boldsymbol{y}} u_2) \right\rangle_\varepsilon \\ & + \left\langle \nabla_{\boldsymbol{x}} \cdot (\mathbf{A}_0 \cdot \nabla_{\boldsymbol{y}} u_1) \right\rangle_\varepsilon + \left\langle \nabla_{\boldsymbol{y}} \cdot (\mathbf{A}_0 \cdot \nabla_{\boldsymbol{x}} u_1) \right\rangle_\varepsilon.\end{aligned} \tag{6.26}$$

We consider the terms in order. On the left-hand side, we have

$$\langle \partial_t \bar{u} \rangle_\varepsilon = \partial_t \langle \bar{u} \rangle_\varepsilon = \partial_t \bar{u}, \tag{6.27}$$

since $\bar{u}$ depends only on the macroscopic variable $\boldsymbol{x}$. For the first term on the right-hand side, we recall that averaging applies to the variable $\boldsymbol{y}$ and that $\boldsymbol{x}$ and $\boldsymbol{y}$ are treated as independent in the limit $\varepsilon \to 0$, so that we have

$$\langle \nabla_{\boldsymbol{x}} \cdot (\mathbf{A}_0 \cdot \nabla_{\boldsymbol{x}} \bar{u}) \rangle_\varepsilon = \nabla_{\boldsymbol{x}} \cdot \langle \mathbf{A}_0 \cdot \nabla_{\boldsymbol{x}} \bar{u} \rangle_\varepsilon. \tag{6.28}$$

Using the periodicity of u_2 and of the corresponding flux, the second term on the right-hand side yields

$$\left\langle \nabla_{\boldsymbol{y}} \cdot (\mathbf{A}_0 \cdot \nabla_{\boldsymbol{y}} u_2) \right\rangle_\varepsilon = 0. \tag{6.29}$$

We also have

$$\left\langle \nabla_{\boldsymbol{y}} \cdot (\mathbf{A}_0 \cdot \nabla_{\boldsymbol{x}} u_1) \right\rangle_\varepsilon = 0. \tag{6.30}$$

Therefore, Eq. (6.26) simplifies to

$$\partial_t \bar{u} = \nabla_{\boldsymbol{x}} \cdot \left\langle \mathbf{A}_0 \cdot (\nabla_{\boldsymbol{x}} \bar{u} + \nabla_{\boldsymbol{y}} u_1) \right\rangle_\varepsilon. \tag{6.31}$$

6.4. Structure of the perturbation solution

To solve for u_1, we exploit the linearity of Eq. (6.21) and write u_1 in the form

$$u_1 = \boldsymbol{\chi}(\boldsymbol{y}) \cdot \nabla_{\boldsymbol{x}} \bar{u} \tag{6.32}$$

which is motivated by separating terms that depend on $\boldsymbol{x}$ and $\boldsymbol{y}$. Here, $\boldsymbol{\chi}$ depends only on the microscale coordinate $\boldsymbol{y}$ (see similar discussion for volume averaging in SubSection 5.9). Substitution of (6.32) into (6.21) yields a local "cell" problem for $\boldsymbol{\chi}$ given by

$$-\nabla_{\boldsymbol{y}} \cdot (\mathbf{A}_0) = \nabla_{\boldsymbol{y}} \cdot (\mathbf{A}_0 \cdot \nabla_{\boldsymbol{y}} \boldsymbol{\chi}), \tag{6.33}$$

with the condition

$$\text{Periodicity in } \boldsymbol{y}. \tag{6.34}$$

We note that $\boldsymbol{\chi}$ may be determined once the periodic geometry and $\mathbf{A}_0$ are prescribed. Finally, $\boldsymbol{\chi}$ is only unique up to additive constants in $\boldsymbol{y}$, so to determine it uniquely we impose

$$\langle \boldsymbol{\chi} \rangle_\varepsilon = \mathbf{0}. \tag{6.35}$$

6.5. Macroscale equation

We can introduce Eq. (6.32) into Eq. (6.31) to obtain

$$\partial_t \bar{u} = \nabla_{\boldsymbol{x}} \cdot (\mathbf{A}_e \cdot \nabla_{\boldsymbol{x}} \bar{u}), \tag{6.36}$$

with

$$\mathbf{A}_e = \left\langle \mathbf{A}_0 \cdot (\mathbf{I} + \nabla_{\boldsymbol{y}} \boldsymbol{\chi}) \right\rangle_\varepsilon. \tag{6.37}$$

A discussion of the dependence of variables $\boldsymbol{x}$ and $\boldsymbol{y}$ is given in Technical note 16.

Technical note 16: Should $\boldsymbol{x}$ and $\boldsymbol{y}$ be viewed as dependent or independent variables? This is a delicate point that may yield significant confusion (in the formal version of multiscale asymptotics at least). At the start, we transform the sequence of functions u_ε that are $\mathbb{R}^n \to \mathbb{R}$ into a new sequence of functions $u_\varepsilon^\ddagger$ that are $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$. Although this new function depends on two spatial variables $(\boldsymbol{x}, \boldsymbol{y})$, we must consider $u_\varepsilon^\ddagger(\boldsymbol{x}, \varepsilon^{-1}\boldsymbol{x}) = u_\varepsilon(\boldsymbol{x})$. Therefore, $\boldsymbol{x}$ and $\boldsymbol{y}$ are, in some sense, related. This becomes clear when we wish to determine derivatives and realize that we must apply the chain rule. We remark further that in practical terms, $\boldsymbol{x}$ and $\boldsymbol{y}$ can be considered independent when solving Eqs. (6.33) to (6.37) (e.g., we solve the unit-cell and macroscale equations separately). This may be understood as follows. In Eq. (6.15), we view ε as a parameter (we may even think about it as being not particularly small), use two-scale expansions and scale the complete problem according to ε. Only when each term in these equations has been explicitly expressed as powers of ε, we can consider the limit $\varepsilon \to 0$ and decompose Eq. (6.15) into a sequence of nested sub-problems. In so doing, we treat $\boldsymbol{x}$ and $\boldsymbol{y}$ as independent, only in the asymptotic limit $\varepsilon \to 0$, a step that corresponds to the intuitive idea of separation of scales.

7. Looking for errors

Given that the above case is one of the simplest that may be studied, the process of homogenization, either via multiscale asymptotics or volume averaging, is often rather involved. Errors can occur at different steps in the developments, e.g., during the initial theoretical derivation, the numerical resolution of the unit-cell problem or the calculation of effective parameters. In addition to careful calculations, it is important to develop techniques for systematically checking results *a posteriori*. This can be achieved, in part, by verifying that fundamental properties of effective tensors and unit-cell problems are satisfied.

7.1. Properties of the effective tensor

Effective tensors must often satisfy properties such as boundedness, positive definiteness or symmetry. For instance, in the parabolic case, we can show that: $\mathbf{A}_e$ must be positive definite; $\mathbf{A}$ symmetric implies $\mathbf{A}_e$ symmetric; and $\mathbf{A}$ diagonal does not imply $\mathbf{A}_e$ diagonal. Further, several bounds exist for $\mathbf{A}_e$ that

correspond to different configurations. Such bounds have received a considerable amount of attention from physical and mathematical points of view, since these can be used to generate approximate values of effective conductivities, elastic moduli for composite materials, or equivalent permeabilities for porous media, without the need to compute the unit-cell problem over a specific geometry.

The fundamental bounds (or Reuss–Voigt–Wiener bounds, [85]) on $\mathbf{A}_e$ are the arithmetic and harmonic means. However, these are usually sub-optimal and tighter bounds have been developed for different cases. Consider for example, the famous Hashin and Shtrikman bounds that apply to isotropic binary mixtures (see [86]). A rigorous proof of these bounds can be derived via *H*-convergence for non-symmetric $\mathbf{A}(\boldsymbol{x})$ (see discussions in [49]). There is an incredible amount of work on the topic so here we only reference the works of Cardwell and Parsons [87], Matheron [88], Rubinstein and Torquato [89], Ené [90]. A good review of these problems in the context of equivalent permeabilities is given by Renard and De Marsily in [83].

7.2. A quick trick for volume averaging

For volume averaging, we usually check that averaging the perturbation and unit-cell problems yields consistent results. On imposing the condition that $\langle\langle\cdot\rangle\rangle = \langle\cdot\rangle$ (spatial localization, see Technical note 8), the averages of the left- and right-hand sides of Eqs. (5.13) to (5.15) should both vanish. This procedure is often useful for detecting missing terms or inconsistent assumptions that will generate contradictions.

To be more explicit, consider a linear spatial differential operator $\mathcal{L}(\cdot)$, an averaging operator $\langle\cdot\rangle$ and suppose that $\mathcal{L}(u) = \text{Source}$. The perturbation problem should have the following structure

$$\mathcal{L}(\tilde{u}) - \langle\mathcal{L}(\tilde{u})\rangle = \text{Source} - \langle\text{Source}\rangle - \mathcal{L}(\langle u\rangle) + \langle\mathcal{L}(\langle u\rangle)\rangle, \tag{7.1}$$

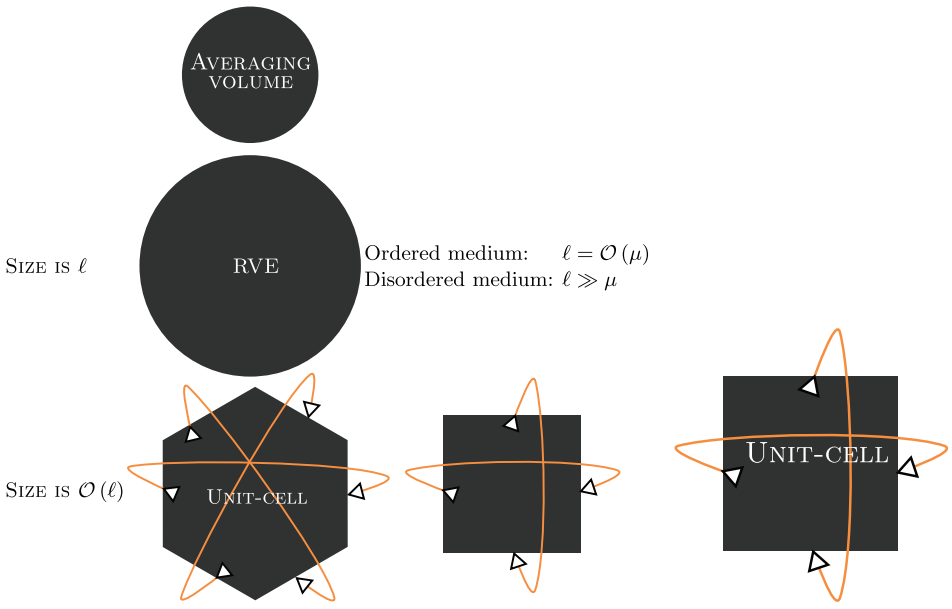
so that averaging this problem with the localization assumption yields $0 = 0$. If terms are missing on the left- or the right-hand side, this will usually yield source terms in these equations that do not disappear upon averaging.

7.3. Validation and convergence

For a variety of operators and boundary conditions, it is possible to rigorously prove convergence (in a sense that is discussed in more detail in Section 9) of the sequence of microscale solutions u_ε towards u_0 (or $\langle u_0\rangle$) or of the sequence of tensors $\mathbf{A}_\varepsilon$ towards $\mathbf{A}_e$ (see e.g., [47,49]). Such demonstrations are, however, based on highly technical functional analysis. Further, there are many cases for which formal homogenization provides a useful description of the physical system but convergence has not been proven. In such cases, a more physical approach to studying convergence involves either comparing the solution of the homogenized model to experiments or to numerical solutions of the microscale problem. In the literature, the latter choice is often referred to as validation of the homogenized model against direct numerical simulations (DNSs, see [91,92]). A DNS is a numerical procedure that aims to solve the microscale problem over Ω. This requires significant computational effort, so we often simplify the system by reducing dimensionality or exploiting symmetries. This approach can also be used to assess the domains of validity of the homogenized model by exploring a broad range of scalings/values for dimensionless parameters, a crucial step in homogenization (see Sections 8 and 9).

8. Comparison of volume averaging and formal asymptotics

In this section, we compare both techniques in terms of the definitions of the different terms, the final results, the algorithms and



(a) Averaging sets for volume averaging.

(b) Averaging set for multiscale asymptotics.

Fig. 9. Schematic representation of the different averaging sets considered in (a) volume averaging and (b) multiscale asymptotics. We show that three different concepts are used in the volume averaging theory: the averaging volume is the moving set over which variables are averaged; the RVE corresponds to a particular size of averaging volume, ℓ, *specific to hierarchical media* which is large enough to capture the microscale features (i.e., is representative) and small enough so that properties can vary at the macroscale; and the unit-cell is a *fictitious* entity with, in most cases, periodic boundary conditions that is used to calculate effective properties of spatially local homogenized models. In the multiscale asymptotic framework, the medium is periodized *a priori* so that there is only one averaging set: the unit-cell.

the approximations. We also give a personal viewpoint of the relative advantages of the two methods.

8.1. Difference in the definitions of the averages

Asymptotics are often associated with L being fixed and $\ell \to 0$, whereas averaging is usually associated with ℓ fixed and $L \to \infty$. In most cases, however, this is simply a matter of definition since the relevant parameter is ε, not ℓ or L. We remark that there are special cases for which fields do not scale with ε (e.g., gravitational terms in momentum transport equations) and for which particular care is needed.

A more important difference concerns the nature of the sets that are used for averaging. The first step of volume averaging consists in applying an averaging operator to the boundary value problems of interest. This is done via a spatial convolution with a smoothing function that has compact support. This support may be a ball, a polyhedron or any set that suits our purposes. In many cases, we will average over a closed ball since, *a priori*, there is no reason to favor one direction (e.g., the diagonals of a cuboid). If the medium is hierarchical, the second step consists in determining an averaging volume that will facilitate mathematical developments. This particular averaging volume is the RVE. In the final step, we often assume that closure fields and effective parameters can be solved over periodic unit-cells. These notions are different and are schematized in Fig. 9. With asymptotics, the averages are defined only over unit-cell cuboids, in a way that is compatible with periodicity.

Henceforth, we will focus on more fundamental differences by assuming that averaging over the periodic cuboids is equivalent to averaging over the RVE/unit-cell combination. Therefore, we will simply use the notation $\langle \bullet \rangle$ to denote any form of spatial averaging. We remark that this assumption is rigorously true only if the choice of the RVE matches the unit-cell used for volume averaging and multiscale asymptotics.

8.2. Differences between u_0 and $\langle u \rangle$

In the example parabolic problem, we have the relationship $u_0 = \langle u \rangle$. To understand this equality, consider the following average of the two-scale expansion up to the first order

$$\langle u \rangle = \langle u_0 \rangle + \varepsilon \langle u_1 \rangle, \tag{8.1}$$

with the uniqueness condition $\langle u_1 \rangle = 0$ (combine Eqs. 6.35 and 6.32). Further, we have shown that u_0 does not depend on $\boldsymbol{y}$, so that $\langle u_0 \rangle = u_0$ and $\langle u \rangle = u_0$.

More generally, there is a fundamental difference between u_0 and $\langle u \rangle$: u_0 may depend on the micro- and macroscales, $\boldsymbol{x}$ and $\boldsymbol{y}$, whereas the assumption (**A2**) imposes that $\langle u \rangle$ varies only at the macroscale. Quantities that are equivalent in both frameworks are in fact $\langle u \rangle$ and $\langle u_0 \rangle$. Loosely speaking, we may average the generic two-scale expansion

$$\langle u_\varepsilon \rangle = \langle u_0 \rangle + \sum_i \langle u_i \rangle \varepsilon^i, \tag{8.2}$$

to obtain $\langle u \rangle = \langle u_0 \rangle$ (for further detail, see Technical note 17).

Technical note 17: In the most general case, there is a *fundamental* difference between u_0 and $\langle u \rangle$. In the volume averaging theory, assumption (**A2**) means that $\langle u \rangle$ varies with frequencies of $\mathcal{O}(1)$ (i.e., it depends only on $\boldsymbol{x}$). In the multiscale asymptotics framework, however, u_0 may depend on $\boldsymbol{x}$ and $\boldsymbol{y}$: *a priori*, there is no reason why u_0 should depend on $\boldsymbol{x}$ only. In the methodology exposed above, it follows from the $\mathcal{O}(\varepsilon^{-2})$ boundary value problem and the two-scale expansion.

Finally, we mention an important result of two-scale convergence (see definition in Section 9): if $u_\varepsilon \in L^2(\Omega)$ two-scale converges to $u_0(\boldsymbol{x},\boldsymbol{y}) \in L^2(\Omega \times Y)$ then u_ε converges to $\int_Y u_0(\boldsymbol{x},\boldsymbol{y}) dV_{\boldsymbol{y}}$ weakly in $L^2(\Omega)$ (see [95]), where Y is the unit cube. This suggests that $\langle u_0 \rangle$ and $\langle u \rangle$ are equivalent, rather than u_0 and $\langle u \rangle$.

8.3. Comparison of macroscale and unit-cell problems

Recall for volume averaging that the approximate representation of u (that we shall denote u_V for clarity) is obtained by combining Eq. (5.9) with Eq. (5.31) (see Section 5), so that

$$u_V(\boldsymbol{x},\boldsymbol{y}) = \langle u \rangle(\boldsymbol{x}) + \varepsilon \boldsymbol{c}(\boldsymbol{y}) \cdot \boldsymbol{\nabla} \langle u \rangle(\boldsymbol{x}), \tag{8.3}$$

where $\boldsymbol{c}$ solves

$$-\boldsymbol{\nabla}_{\boldsymbol{y}} \cdot \tilde{\mathbf{A}} = \boldsymbol{\nabla}_{\boldsymbol{y}} \cdot \left(\mathbf{A} \cdot \boldsymbol{\nabla}_{\boldsymbol{y}} \boldsymbol{c} \right), \tag{8.4}$$

with conditions

$$\text{Periodicity in } \boldsymbol{y}, \tag{8.5}$$

$$\langle \boldsymbol{c} \rangle = 0. \tag{8.6}$$

Similarly, the approximate asymptotic representation of u (that we shall denote u_A) is obtained by combining Eqs. (6.8) and (6.32) (see Section 6) to give

$$u_A(\boldsymbol{x},\boldsymbol{y}) = \langle u \rangle(\boldsymbol{x}) + \varepsilon \boldsymbol{\chi}(\boldsymbol{y}) \cdot \boldsymbol{\nabla}_{\boldsymbol{x}} \langle u \rangle, \tag{8.7}$$

where $\boldsymbol{\chi}$ solves

$$-\boldsymbol{\nabla}_{\boldsymbol{y}} \cdot \mathbf{A}_0 = \boldsymbol{\nabla}_{\boldsymbol{y}} \cdot \left(\mathbf{A}_0 \cdot \boldsymbol{\nabla}_{\boldsymbol{y}} \boldsymbol{\chi} \right), \tag{8.8}$$

with conditions

$$\text{Periodicity in } \boldsymbol{y}, \tag{8.9}$$

$$\langle \boldsymbol{\chi} \rangle = \mathbf{0}.$$

While there appear to be several discrepancies between these results, these differences are only superficial. First, consider that $\mathbf{A} = \mathbf{A}_0$ by definition ($\mathbf{A}_0(\boldsymbol{y}) \equiv \mathbf{A}(\boldsymbol{x})$). Next, consider that $\boldsymbol{\nabla}_{\boldsymbol{y}} \cdot \tilde{\mathbf{A}} \simeq \boldsymbol{\nabla}_{\boldsymbol{y}} \cdot \mathbf{A}$ since $\boldsymbol{\nabla}_{\boldsymbol{y}} \cdot \langle \mathbf{A} \rangle \simeq 0$. Given the uniqueness of solutions of these problems, it follows that $\boldsymbol{c} = \boldsymbol{\chi}$ and $u_A = u_V$. Therefore, the macroscale equations and effective parameters obtained via volume averaging, Eqs. (5.38) and (5.39), are identical to those obtained via multiscale asymptotics, Eqs. (6.36) and (6.37). This result was also obtained by Bourgeat et al. [27].

8.4. Comparison of algorithms

In Figs. 10 and 11, algorithms for volume averaging and multiscale asymptotics respectively are presented as flowcharts. Both flowcharts share the same starting and ending actions: "Identify microscale initial boundary value problems (IBVPs)"; "Go to SOLVE" if the procedure has been successful and we can go to the SOLVE flowchart, Fig. 13, corresponding to the resolution of the homogenized problem; and "Use another method" if the procedure has not yielded a model which domain of validity matches our needs.

The volume averaging method can be summarized as follows. The first steps involve averaging the microscale IBVP, using the perturbative decomposition and then obtaining the perturbation IBVP. From this point, we *make several assumptions* (these are detailed in Section 5 for the parabolic transport problem) in order

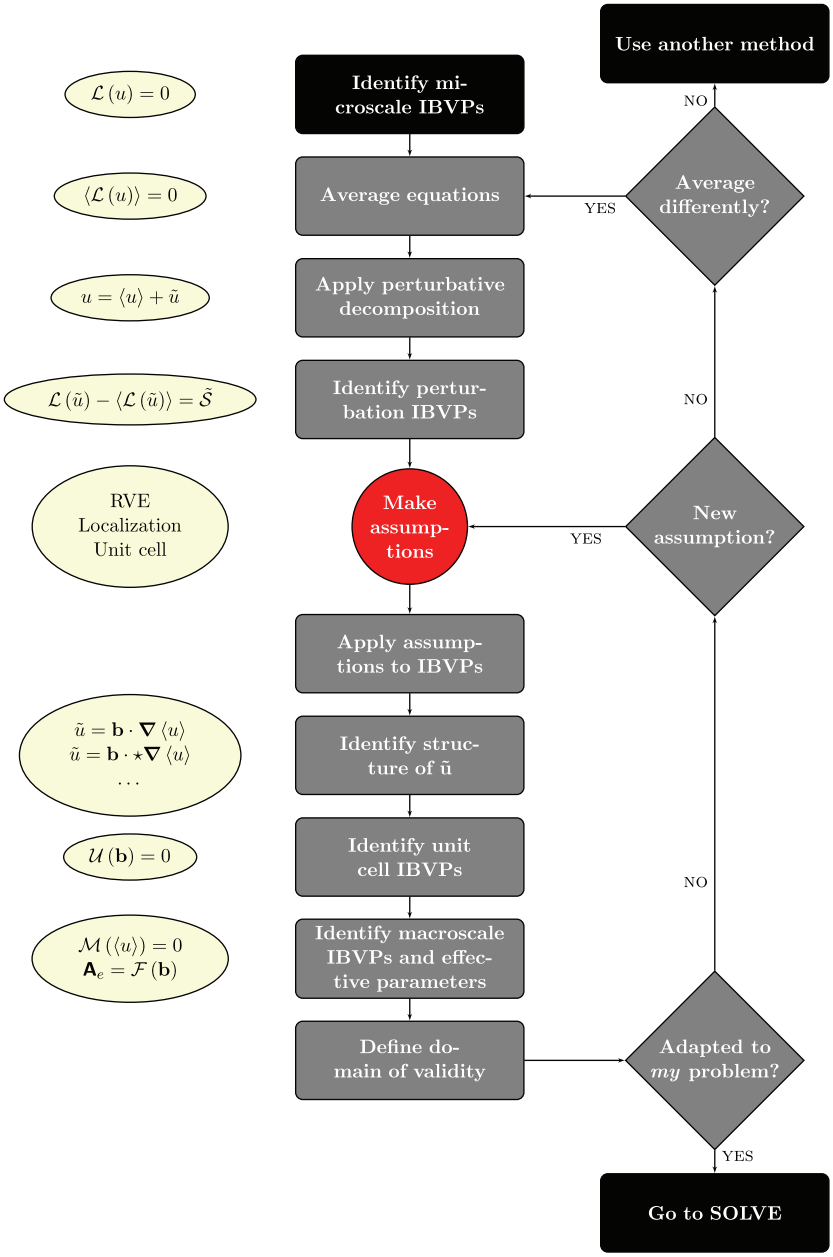


Fig. 10. Flowchart illustrating the algorithm used for volume averaging (see Section 5 for an example application to a parabolic problem).

to obtain an approximate solution to the perturbation IBVP. This approximate solution is then introduced into the unclosed form of the average equation and may yield a local or nonlocal homogenized problem. If the problem is local in space (that was the case of the example parabolic problem), effective parameters can be expressed as function of closure variables that may be calculated separately over the unit-cell.

For the formal asymptotics method, the process *starts with assumptions*, so that the initial problem can be casted into a form that is better suited to homogenization. These assumptions imply that we can explicitly express the dependence of each term upon the scale ratio, ε, so that we may consider the limit $\varepsilon \to 0$. In so doing, we obtain a coupled series of IBVPs from which we can extract the homogenized model. In the previous parabolic case, we can simplify the coupling between these IBVPs and obtain a macroscale equation that involves effective parameters that are expressed as functions of variables that may be calculated separately over the unit-cell. We remark that, in many cases, it may not be possible to simplify coupling and we may obtain nonlocal or mixed scale formulations. As for volume averaging, we recommend decoupling the micro- and macroscale models as much as possible.

Finally, for each method, we must specify the domains of validity, i.e., translate assumptions into clear physical restrictions. In the example problem presented in Section 4, the assumptions associated with each method and their physical interpretation were thoroughly discussed in Sections 5 and 6. Example representations of more complex domains of validity, in particular involving a broader set of dimensionless numbers, can be found in [96]. Other examples include timescale limitations as discussed in [97]. If the domain of validity is compatible with the physical system and satisfies criteria fixed by the user (e.g., "the model should be accurate only in the time-asymptotic regime" or "the model should be local in space"), we can proceed

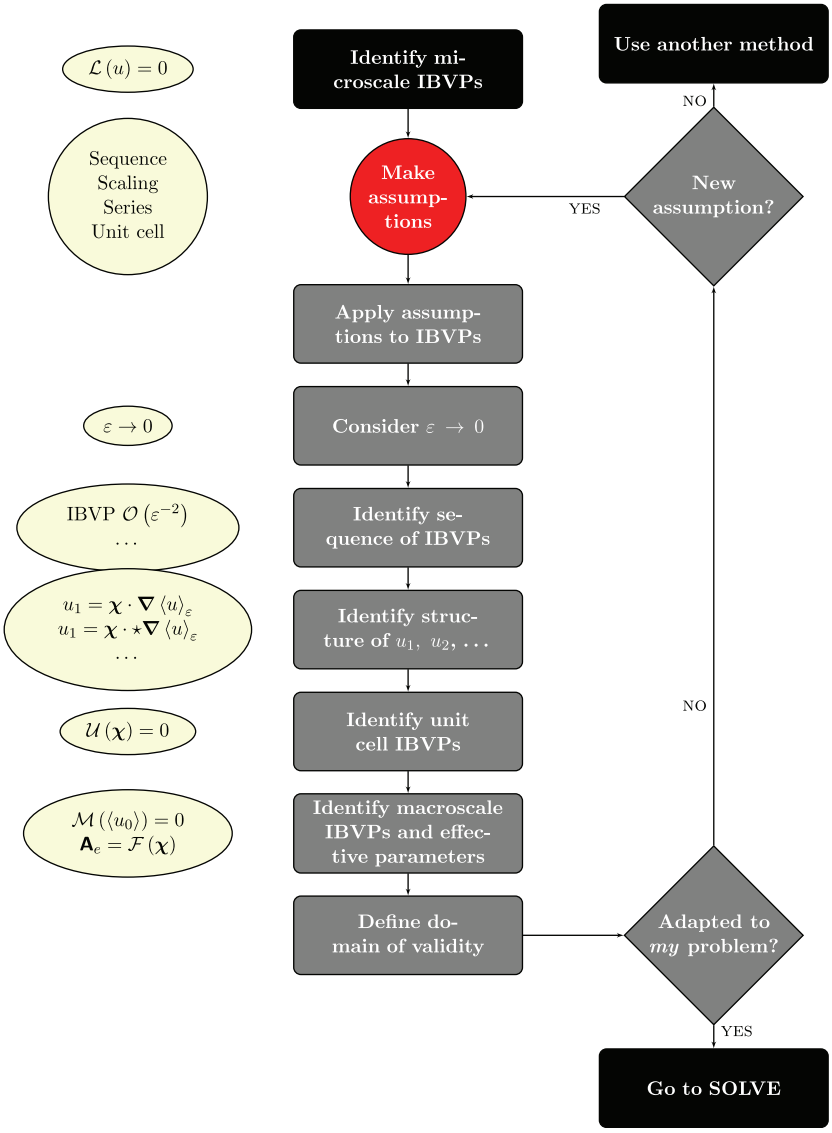


Fig. 11. Flowchart illustrating the algorithm used for formal asymptotics (see Section 6 for an example application to a parabolic problem).

to the resolution of the problems (see Section 9.3). If not, we may consider different assumptions (e.g., relax the quasi-stationarity assumption and develop a macroscale model that is nonlocal in time). We may also try to average the problem differently, for instance by splitting the domain into different regions and deriving an explicit representation of the average flux. This will not change the problem, rather it sheds new light on the equations, which may in turn make novel approximations visible. This is the case with two-equation models (see [93] and [94]) which may be viewed as higher order approaches.

An important difference between volume averaging and formal multiscale asymptotics, is, therefore, the chronology of actions in the algorithms. In particular, the "Make assumptions" action, which is the most important step in both techniques, occurs at different stages of the developments (compare Figs. 10 and 11). For the formal asymptotics, the assumptions are made *a priori*, i.e., we transform our equations into a new sequence of problems that are further away from the physical problem but more convenient for homogenization. For volume averaging, we do not make strong assumptions *ab initio*; instead, we start by using averaging results that apply to most systems, even those that violate $\varepsilon \ll 1$.

8.5. Comparison of approximations

Another important difference between the techniques is the nature of the assumptions that are made. Since the final results are identical, there are strong correspondences between these assumptions. In this subsection, we aim to highlight these relationships. A summary is available in Table 1.

(*RVE and then periodic unit-cell*) vs. (*Periodic unit-cell*). With multiscale asymptotics, we assume that the medium is periodic *a priori*, whereas volume averaging uses a periodic unit-cell *a posteriori* only to calculate effective parameters. As a result, the averaging volume in the multiscale asymptotics framework always corresponds to the unit-cell. The volume averaging approach does not require this approximation *a priori* via the notion of RVE.

In essence, the significance of the periodicity assumption is the same in both cases: either the medium is periodic, or we assume that effective properties can be approximated by considering a periodic case. Further, periodicity also implies that the unit-cell problem and effective parameters do not depend upon macroscale boundary conditions (a strong assumption that allows the formulation of uncoupled micro- and macroscale descriptions).

Table 1
Summary highlighting the principal differences between volume averaging and multiscale asymptotics.

		Volume averaging	Multiscale asymptotics
Definitions	Averaging operator	$\langle\bullet\rangle$	$\langle\bullet\rangle_\varepsilon$
	Expansions	$u = \langle u\rangle(\boldsymbol{x}) + \tilde{u}(\boldsymbol{x}, \frac{\boldsymbol{x}}{\varepsilon})$	$u_\varepsilon(\boldsymbol{x}) = \sum u_i(\boldsymbol{x}, \frac{\boldsymbol{x}}{\varepsilon})\varepsilon^i$
	First term	$\langle u\rangle(\boldsymbol{x})$	$u_0(\boldsymbol{x}, \boldsymbol{y})$
	Equivalent averages	$\langle u\rangle$	$\langle u_0\rangle$
Results	Final models	One model for all scalings	One model for one scaling
Algorithms	Chronology	*A posteriori* assumptions	*A priori* assumptions
Approximations	Set up	One problem	Sequence of problems
	Averaging volumes	RVE and then unit-cell	Unit-cell
	Periodicity	*A posteriori* periodicity	*A priori* periodicity
	Spatial localization	Neglecting higher order terms	Two-scale expansion, scaling of **A** and limit $\varepsilon \to 0$
	Temporal localization	Quasi-stationarity of the perturbations	Two-scale expansion, scaling of **A** and limit $\varepsilon \to 0$.

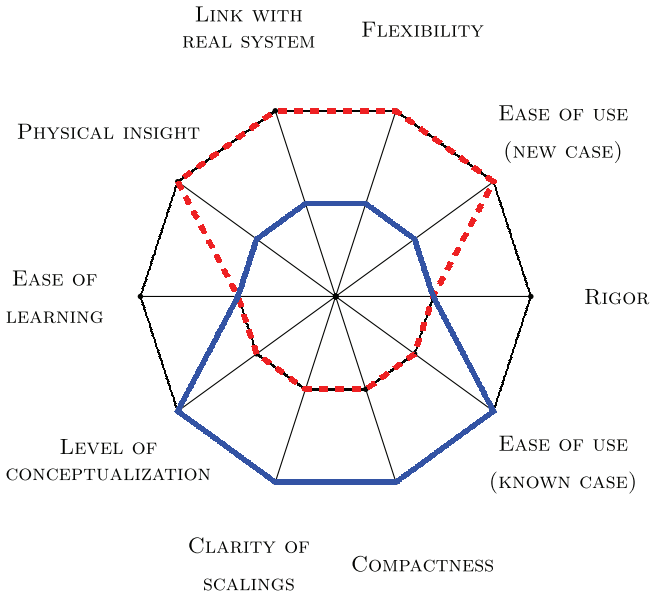


Fig. 12. Spiderweb diagram showing a personal view of the advantages associated with volume averaging (dashed red) and formal multiscale asymptotics (blue). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

(*Average-plus-perturbation*) vs. (*Two-scale expansion*). Both methods rely on a perturbative decomposition of the initial fields, but the initial decompositions are rather different. With volume averaging, we decompose the microscale fields into a spatially averaged term and a perturbation, without imposing a more specific structure of the perturbation. With formal asymptotics, we assume that the dependent variables can be expanded as power series of ε (the two-scale expansion *ansatz*), and that each u_i is a function of $\boldsymbol{x}$ and $\boldsymbol{y}$, so that the perturbation structure is constrained from the start. Further, the first terms in both decompositions, u_0 and $\langle u\rangle$, are not identical, as is shown in SubSection 8.2.

However, we have proved in SubSection 8.3 that the final expressions are identical for both decompositions. This results from the additional constraints that are imposed on the perturbation in the volume averaging methodology (in the "Make assumptions"). In particular, we assume in SubSection 5.6 that the dimensionless spatial frequencies of any average quantity, $\langle\cdot\rangle$, are $\mathcal{O}(1)$ and $\tilde{\psi}$ contains larger frequencies of order $\mathcal{O}(\varepsilon^{-1})$. This yields results similar to using a two-scale expansion with terms that depend on the slow, $\boldsymbol{x}$, and rapid, $\boldsymbol{y}$, spatial variables.

(*Neglecting higher order spatial terms*) vs. (*Two-scale series, scaling of* **A** *and limit* $\varepsilon \to 0$). To obtain an approximate structure of the perturbation, both techniques rely on the fact that there is a separation of spatial scales, $\varepsilon \ll 1$, so that we can approximate the system by considering the asymptotic limit $\varepsilon \to 0$. In multiscale asymptotics, we explicitly detail all ε dependencies via the two-scale expansion and the scaling of **A**. In the volume averaging framework, the scaling of **A** is arbitrary and we adopt a more physical approach whereby we neglect higher-order terms *a posteriori*. While this does not impact the results in the purely diffusive case, it plays an important role in advection–diffusion problems for which different scalings of the Péclet number yield different results. In the volume averaging approach, these results are typically contained in the same equations, whereas in the formal asymptotic approach these different scalings are analyzed one by one. As discussed in Technical note 18, an advantage of multiscale asymptotics is that it enables the identification of scalings that result in non-homogenizablesystems.

Technical note 18: In the most general case, the relative scaling of dimensionless numbers will influence spatial localization and some scalings may yield problems for which a local solution does not exist. Such cases are sometimes referred to as ''non-homogenizable systems''. They arise when there is a strong coupling between the micro- and macroscales. For instance, this is the case if the Péclet number scales as ε^{-1} (see [52] and Technical note 12). This does not appear in our previous analysis as we have used only one scaling and the problem exhibits only one dimensionless number.

(*Temporal quasi-stationarity of the perturbations*) vs. (*Two-scale series, scaling of* **A** *and limit* $\varepsilon \to 0$). With the volume averaging technique, we suppose that the unit-cell problem is quasi-steady. As detailed in Section 5, this means that the timescale for relaxation of the transient unit-cell problem is extremely rapid compared to characteristic timescales for the macroscale problem. With formal asymptotics, we obtain a steady problem on u_1 as a result of the diffusive scaling choosen for **A**. This diffusive scaling, however, represents the ratio of characteristic times associated with the micro- and macroscale processes, so that both assumptions have the same physical significance. We remark further that these assumptions may be relaxed to obtain a nonlocal form of the macroscale problem.

8.6. A more personal comparison

The two communities that led the development of volume averaging and multiscale asymptotics become self-evident when we consider the advantages and disadvantages of the two methods (see Fig. 12). The volume averaging approach was proposed by

physicists and engineers, whereas asymptotics was developed by mathematicians. This has led to fundamental differences in the way in which assumptions are expressed and in the algorithms that are used, even though the final results are the same. The volume averaging method provides a more physical framework that brings both advantages and disadvantages. One of its principal strengths is that the whole process remains more faithful to the physical systems. Further, we do not have to make approximations at the outset (for instance, we do not have to explicitly scale parameters as functions of ε): the first step involves only averaging in space. This yields a physical interpretation of the averaging procedure that facilitates physical intuition, especially in multiphase systems (see Technical note 19) where interfacial terms are more easily dealt with.

The multiscale asymptotics method, on the other hand, is more upfront, introduces a clearer scaling of parameters in terms of ε and provides a more systematic framework with a higher level of conceptualization. Once the model is defined in a mathematical form, the method is straightforward to apply and just requires calculations to be performed, without a strong need to "think" about the physics.

Technical note 19: Consider the diffusive transport of a solute through two regions (γ and ω). We denote the solute concentrations in each phase by c_γ and c_ω, and by D_γ and D_ω the corresponding diffusion coefficients. Transport may be described by the following transmission problem

$$\partial_t c_\gamma = \nabla \cdot (D_\gamma \nabla c_\gamma),$$
$$\partial_t c_\omega = \nabla \cdot (D_\omega \nabla c_\omega),$$

with continuity of concentrations and their fluxes across phase boundaries. On averaging the two transport equations within each phase, we obtain

$$\partial_t \langle c_\gamma \rangle = \nabla \cdot (D_\gamma \langle \nabla c_\gamma \rangle) + \frac{1}{V} \int_{\mathcal{S}_\gamma} \mathbf{n}_{\gamma\omega} \cdot \nabla c_\gamma d\mathcal{S},$$
$$\partial_t \langle c_\omega \rangle = \nabla \cdot (D_\omega \langle \nabla c_\omega \rangle) + \frac{1}{V} \int_{\mathcal{S}_\omega} \mathbf{n}_{\omega\gamma} \cdot \nabla c_\omega d\mathcal{S}.$$

The structure of this averaged system provides insight into the physics before homogenization assumptions are made. In particular, we remark that $\frac{1}{V}\int_{\mathcal{S}_\gamma} \mathbf{n}_{\gamma\omega} \cdot \nabla c_\gamma d\mathcal{S} = -\frac{1}{V}\int_{\mathcal{S}_\omega} \mathbf{n}_{\omega\gamma} \cdot \nabla c_\omega d\mathcal{S}$ is the average interfacial flux. This suggests that it is also possible to derive an alternative two-equation macroscale model which includes solute exchange between the two phases.

If our purpose is to tackle a real problem, then neither technique is more rigorous than the other. While multiscale asymptotics may seem to be more rigorous, our viewpoint is that this is a false impression created by the strong simplifications of the problem made at the start. Once this has been done, the later stages are straightforward but the relationships with the original physical problem and the validity of the two-scale expansions are less clear. In any case, the reader should bear in mind that, from a mathematical point of view, neither technique is exact: function spaces and metrics are not defined, convergence is not proven, the two-scale decomposition is an *ansatz*, the volume averaging method uses many order of magnitude estimates. Rigorous analyses exist for some cases (see Section 9).

8.7. Which method should you use?

Ideally, we recommend that you learn and use both methods as this will give you a unique perspective upon multiscale problems. From our point of view, dogmatism and sectarianism of the scientific communities are part of the reasons that both methods developed separately for so many years. We hope that this paper will motivate new generations of researchers to go beyond the confines of one school of thought.

9. One step further

This section is a succinct introduction to supplementary material that may assist the interested reader to take one step further and enter contemporary research areas. *We acknowledge that the content is by no means exhaustive and reflects choices of the authors.*

9.1. Corrector results

Corrector results refer to reconstructions of the microscale fields from the solutions of the unit-cell and macroscale problems. These apply to volume averaging and multiscale asymptotics in exactly the same way.

To be explicit, let us consider the case of the previous parabolic problem. The resolution of the macroscale problem yields the value of $\langle u \rangle$ over Ω, which, by definition, filters high frequency spatial components of the microscale variable u. From the solution of $\langle u \rangle$, we can also determine the vector field $\nabla\langle u \rangle$ pointwise and use the solution of the unit-cell problem, $\mathbf{b} = \varepsilon\mathbf{c} = \varepsilon\boldsymbol{\chi}$, to calculate the approximate value of the fluctuations, $\tilde{u} \simeq \mathbf{b} \cdot \nabla\langle u \rangle$. From these results, we can construct the corrector result as

$$u_{cor} = \langle u \rangle + \mathbf{b} \cdot \nabla\langle u \rangle.$$

The corrector results are important in many instances. First, they allow us to access the approximated value of the microscale field and may provide useful information for specific problems. For example, it can be convenient when making comparisons with experiments (see [98]) or DNS (see in [99]). Second, because it captures microscale fluctuations, it modifies convergence properties and, without going into (important) details, yields a "strong" convergence of u to u_{cor} when $\varepsilon \to 0$, as opposed to a "weaker" convergence of u to $\langle u \rangle$(or to u_0). In other words, it gives us a much better approximation of u, without extensive computational efforts.

9.2. SOLVE algorithm

In Fig. 13, we present the algorithm used for the resolution of a homogenized problem. We include nonlocal descriptions, so that micro- and macroscale equations may still be coupled. If that is the case the numerical resolution of the problem will be much more intensive than the resolution of a purely local system. If these are decoupled, then we can solve the unit-cell problem separately, use it to compute the effective parameters and then solve the macroscale equations.

9.3. Examples of unit-cell problems and homogenized models

The goal of this subsection is to explain how the homogenized model may be applied to particular cases of interest, and to provide examples of unit-cell solutions and macroscale models. We will first study the case of a two-dimensional stratified medium and then solve the unit-cell problem for a complex three-dimensional geometry.

9.3.1. Analytical solution for a stratified porous medium

We consider the two-dimensional (spatial variables in the $\boldsymbol{y}$ coordinate system are y_1 and y_2) cell geometry given in Fig. 14, where the parameter A is now a scalar that is piecewise constant, given by

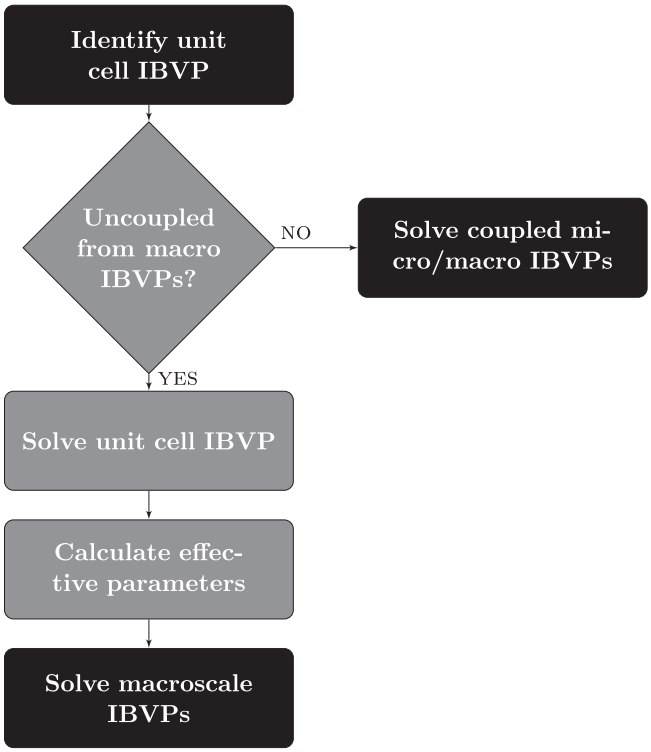


Fig. 13. Flowchart illustrating the resolution algorithm (referred to as SOLVE in Figs. 10 and 11).

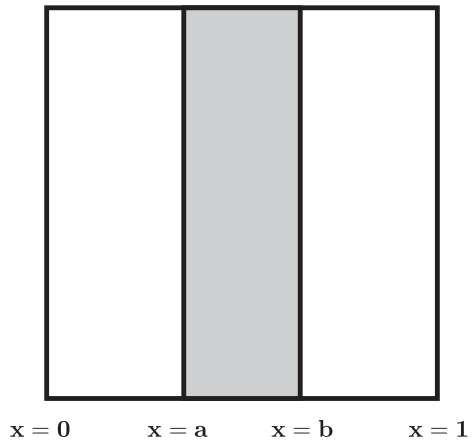


Fig. 14. The stratified cell problem geometry.

$$A = \begin{cases} A_\sigma & \text{for } a \leqslant y_2 \leqslant b, \\ A_\omega & \text{for } 0 < y_2 < a \text{ and } b < y_2 < 1, \end{cases}$$

where A_ω and A_σ are both constants.

The unit-cell problem is

$$-\nabla \cdot \left(\widetilde{A}\mathbf{I}\right) = \nabla \cdot (A\nabla \boldsymbol{c}),$$

with periodicity and the zero-average constraint. For piecewise constant A, we can use Schwartz's derivatives so that

$$\boldsymbol{n}_{\omega\sigma}\sigma(A_\sigma - A_\omega) = \nabla \cdot (A\nabla \boldsymbol{c}),$$

where σ is the interfacial Dirac distribution. Equivalently, we can split this problem as $\Delta \boldsymbol{c}_\alpha = 0$ within each phase $\alpha = \omega$ or σ (open sets) with boundary conditions,

$$\boldsymbol{n}_{\omega\sigma} \cdot (A_\omega \nabla \boldsymbol{c_\omega} - A_\sigma \nabla \boldsymbol{c_\sigma}) = \boldsymbol{n}_{\omega\sigma}(A_\sigma - A_\omega),$$
$$\boldsymbol{c_\omega} = \boldsymbol{c_\sigma}.$$

Resolution of this problem and integration yields (see Appendix B)

$$\mathbf{A}_e = \begin{bmatrix} A_a & 0 \\ 0 & A_h \end{bmatrix},$$

where the components are the weighted arithmetic mean, $A_a = \sum \phi_i A_i$, with ϕ_i the volume fraction of phase i, and the weighted harmonic mean, $A_h = 1/\left(\sum \frac{\phi_i}{A_i}\right)$. For the case where A_ω and A_σ are generic second order tensors, equivalent relationships have been derived by Quintard and Whitaker [100]. Such unit-cell problems that can be solved analytically, either exactly or approximately, are rare (another example consists of a sphere embedded in a matrix see [1,4,101,102]).

9.3.2. Validation against analytical solutions and comparison with direct numerical simulations

We validate the homogenization of the diffusion operator by comparing the solution of the microscale problem with analytical and numerical solutions of the homogenized equations for a collection of test problems. In all cases, we consider N strips of length 10 units in the y_1-direction, and 1 unit in the y_2-direction, stacked on top of each other in the y_2-direction. The domain therefore occupies $0 < y_1 < 10$, $0 < y_2 < N$. Inside each individual strip, diffusion is homogeneous and the diffusion coefficient alternatively takes the value of A_ω and A_σ: we therefore have diffusion coefficient A given by

$$A = \begin{cases} A_\sigma, & n < y_2 < n+1 \text{ if } n \text{ is an odd integer}, \\ A_\omega, & n < y_2 < n+1 \text{ if } n \text{ is an even integer}. \end{cases} \tag{9.1}$$

Noting that $\phi_\omega = \phi_\sigma = 0.5$, this gives rise to a homogenized diffusion tensor $\mathbf{A}_e$ given by

$$\mathbf{A}_e = \begin{pmatrix} 0.5(A_\omega + A_\sigma) & 0 \\ 0 & \left(\frac{1}{2A_\omega} + \frac{1}{2A_\sigma}\right)^{-1} \end{pmatrix}. \tag{9.2}$$

In all simulations below we fix $A_\omega = 1$ and $A_\sigma = 0.1$.

Our microscale problem is, for $0 < y_1 < 10, 0 < y_2 < 100$, find u such that

$$\nabla \cdot (A\nabla u) = S, \tag{9.3}$$

where both u and $A\partial u/\partial \boldsymbol{n}$ are continuous across the interface between strips, and A is given by Eq. (9.1). The corresponding homogenized problem is, for $0 < y_1 < 10, 0 < y_2 < 100$, find $\langle u \rangle$ such that

$$\nabla \cdot (\mathbf{A}_e \cdot \nabla \langle u \rangle) = S, \tag{9.4}$$

where $\mathbf{A}_e$ is given by Eq. (9.2).

We will now investigate the choice of different boundary conditions and the source/sink term, S. The derivation of the homogenized diffusion tensor, as presented above with multiscale asymptotics and volume averaging, did not take account of Neumann boundary conditions and the source term, S. We therefore consider two different sets of model problems: first those without Neumann boundary conditions and with $S = 0$ (where the analysis is more likely to hold), and those with Neumann boundary conditions and $S \neq 0$ (where the analysis may not hold). In all cases, the numerical solutions were calculated using the finite element method.

Model Problems without Neumann boundary conditions and with $S = 0$. Our first model problem in this section uses $S = 0$ and boundary conditions

$$u = \langle u \rangle = 1, \quad y_1 = 0,$$
$$u = \langle u \rangle = 0, \quad y_1 = 10,$$
$$u, \langle u \rangle, \quad \text{periodic on } y_2 = 0, 100.$$

This problem has solution

$$u = \langle u \rangle = \frac{10 - y_1}{10},$$

for any choice of A_ω and A_σ, so that the micro- and macroscale solutions are equal.

Our second model problem in this section has boundary conditions

$$\begin{aligned} &u = \langle u \rangle = 1, \quad y_2 = 0, \\ &u = \langle u \rangle = 0, \quad y_2 = 100, \\ &u, \langle u \rangle, \quad \text{periodic on } y_1 = 0, 10. \end{aligned} \tag{9.5}$$

This problem has solution

$$u = u(y_2),$$

$$\langle u \rangle = \langle u \rangle(y_2) = 1 - \frac{y_2}{100}.$$

The solution for u is shown in Fig. 15(a), and the solution for $\langle u \rangle$ (solid line), and the corrector (broken line) are shown in Fig. 15(b). We see that the correction works exactly in this case. We remark that we would need to adjust boundary conditions to obtain $\langle \tilde{u} \rangle = 0$.

Model Problems with Neumann boundary conditions and $S \neq 0$. For our first model problem in this section, we take $N = 100$, the sink $S = -0.01$, and apply the following boundary conditions:

$$\begin{aligned} &u = \langle u \rangle = 1, \quad y_1 = 0, \\ &\frac{\partial u}{\partial \boldsymbol{n}} = \frac{\partial \langle u \rangle}{\partial \boldsymbol{n}} = 0, \quad \text{on other boundaries.} \end{aligned} \tag{9.6}$$

Contour plots of the solution for u and $\langle u \rangle$ are presented in Fig. 16(a) and (b). We see that the solution for u exhibits fluctuations with the same frequency as the microstructure due to the combination of the diffusion coefficient for the microscale problem varying between strips and the sink term. As expected this microscale feature is not observed for $\langle u \rangle$, the solution to the homogenized problem, and would not be captured by a corrector result either. Nevertheless, away from the boundaries $y_2 = 0, 100$ the solution to the homogenized problem does an excellent job of approximating the macroscale features. At the boundaries $y_2 = 0, 100$ the microscale and homogenized models diverge in order to satisfy the respective asymmetric boundary conditions (top and bottom). To verify that no flux boundary conditions are the cause of this boundary effect we solve a second model problem with $N = 100$ and boundary conditions given by

$$\begin{aligned} &u = \langle u \rangle = 1, \quad y_1 = 0, \\ &\frac{\partial u}{\partial y_1} = \frac{\partial \langle u \rangle}{\partial y_1} = 0, \quad y_1 = 10, \\ &u, \langle u \rangle, \quad \text{periodic on } y_2 = 0, 100. \end{aligned} \tag{9.7}$$

A contour plot of the solution for u is shown in Fig. 16(c). We see that in this case, boundary effects are no longer present, as expected.

To demonstrate the inhomogeneities in the homogenized diffusion tensor that arise from the microscale problem, we consider a further model problem with boundary conditions given by

$$\begin{aligned} &u = \langle u \rangle = 1, \quad y_1, y_2 = 0, \\ &\frac{\partial u}{\partial \boldsymbol{n}} = \frac{\partial \langle u \rangle}{\partial \boldsymbol{n}} = 0, \quad \text{on other boundaries.} \end{aligned} \tag{9.8}$$

Contour plots of the solutions for u and $\langle u \rangle$ are presented in Fig. 17(a) and (b). As in Fig. 16(a) and (b), we see that the microscale fluctuations of u caused by the microstructure and the no flux boundary conditions, are not captured by the homogenized solution, $\langle u \rangle$. Nevertheless, the macroscale features of u are evident in $\langle u \rangle$.

9.3.3. Solution of the unit-cell problem on a complex geometry

In most cases, analytical solutions of the generic unit-cell problem are not available and it must instead be solved numerically. In Fig. 18(a), we present the three dimensional geometry used to compute $\mathbf{c}(\boldsymbol{y}) = \boldsymbol{\chi}(\boldsymbol{y})$ as per Fig. 18(b). The geometry consists of nine (blue) cuboids (see Fig. 18(a)) embedded within a cubic unit-cell so that:

$$A = \begin{cases} 0.1 & \text{for the blue cuboids (volume fraction 0.570).} \\ 1.0 & \text{for the transparent matrix (volume fraction 0.430).} \end{cases}$$

The numerical procedure is based on a finite volume formulation detailed by Quintard [103]. We remark that the results have the same reflection symmetries as the geometry.

Note that, throughout this subsection, we present approximate values rounded to the third decimal. The effective tensor corresponding to this unit-cell is calculated to be

$$\mathbf{A}_e = 0.387\mathbf{I},$$

where $\mathbf{I}$ is the identity tensor. Harmonic and arithmetic bounds are 0.163 and 0.487, respectively. Hashin–Shtrikman bounds (that apply only to the isotropic case) are approximately 0.243 and 0.411 (see Section 7.1 for more details).

9.4. Homogenization of a variety of differential operators

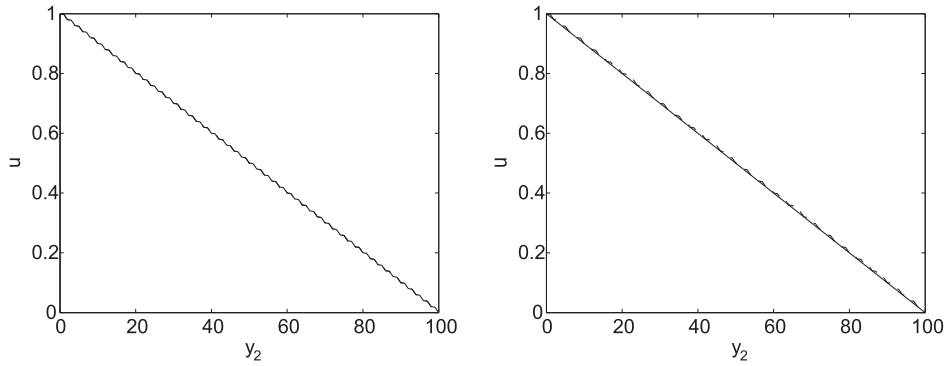
Application of the two formal methods presented in this paper are not restricted to the parabolic problem presented in Section 4. They can be applied to numerous differential operators that relate to different applications, provided that a notion of separation of scales or a small parameter can be defined and that the nature of the initial operator remains deterministic and linear. Volume averaging and formal periodic asymptotics can be used to deal with elliptic, parabolic or hyperbolic equations for numerous applications including: Newtonian and non-Newtonian Stokes and Navier–Stokes, advection–diffusion-reaction, acoustic wave propagation, Schrödinger's, solid mechanics, Maxwell's, Reynolds' and poromechanics problems.

Some results exist for nonlinear differential operators (see an example discussion in [95]), but these are typically of mathematical, rather than practical, interest. Further, such results are primarily nonlocal (i.e., there is still a strong coupling between micro- and macroscale equations). This considerably reduces their practical value, as it increases the computational cost associated with a real model resolution. From our viewpoint, the development of practical models in homogenization of nonlinear differential operators remains a considerable challenge.

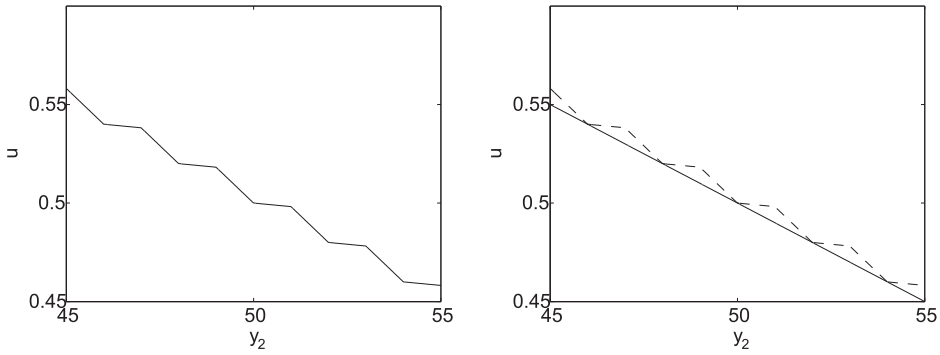
There is also no particular scale restriction that applies to these methods. They have been used to study problems ranging from the cellular level within biological tissues (see [104]) to large-scale averaging of dispersion equations in aquifers [105], through uptake in plant roots [106] or transport in bioreactors [107]. Further, reiterated or successive homogenization can be performed to obtain macroscale equations (see, for instance [108] or [47] for nonlinear monotone operators). One example of a reiterated averaging leading to a practical model is the successive upscaling of Stokes equations to Darcy's law and then from Darcy's law to a two-pressure model (see [56,109]).

9.5. Non-uniqueness of macroscale models

In general, there are several macroscale models associated with a given (unique) microscopic problem. Therefore, the development of the homogenized model cannot be fully decoupled from the physical situation of interest and the corresponding assumptions that are made during upscaling. Consider, for example, the different models that are used to describe momentum transport in por-

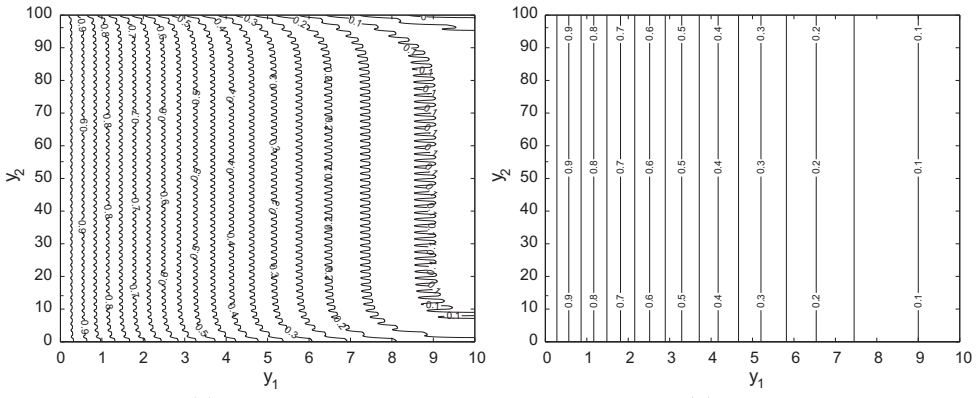


(a) Microscale. (b) Average and corrector.

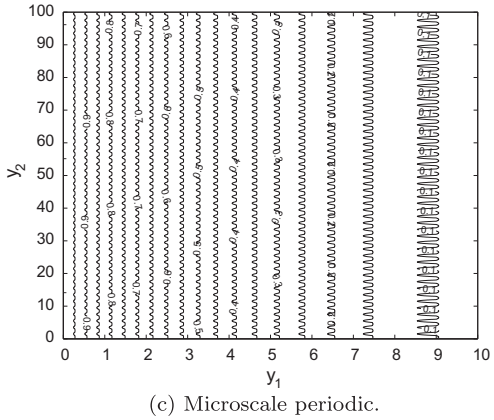


(c) Microscale (zoomed in). (d) Average and corrector (zoomed in).

Fig. 15. (a) The solution for u satisfying boundary conditions given by Eq. (9.5); (b) the solution for $\langle u \rangle$ (solid line), and the corrected $\langle u \rangle$ (broken line), satisfying boundary conditions given by Eq. (9.5). Figures (c) and (d) are corresponding zoomed in representations.

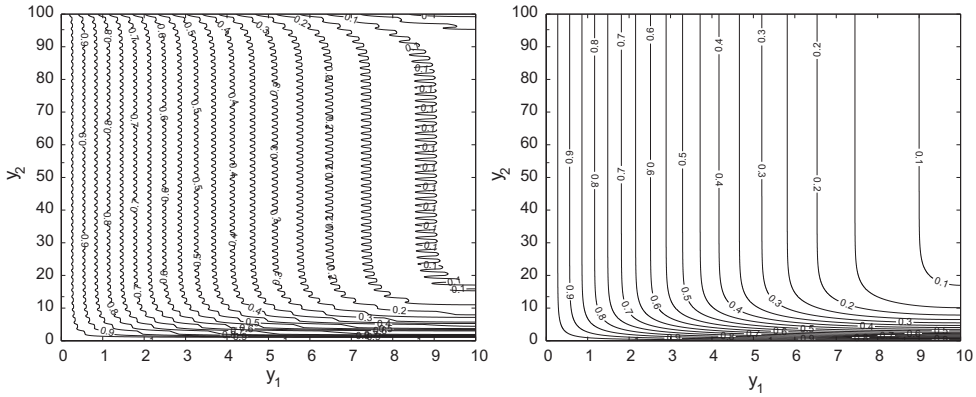


(a) Microscale. (b) Average.



(c) Microscale periodic.

Fig. 16. (a) The solution for u satisfying boundary conditions given by Eq. (9.6), (b) the solution for $\langle u \rangle$ satisfying boundary conditions given by Eq. (9.6) and (c) the solution for u satisfying boundary conditions given by Eq. (9.7).



(a) Microscale. (b) Average.

Fig. 17. (a) The solution for u satisfying boundary conditions given by Eq. (9.8), and (b) the solution for $\langle u \rangle$ satisfying boundary conditions given by Eq. (9.8).

ous media, i.e., Darcy, Darcy with memory effects (a time nonlocal formulation [110]), Darcy–Brinkman (including the viscous term in the macroscale equation [111]) or Darcy–Forcheimer (including higher order velocity expansions in the macroscale equation to deal with inertial effects, see discussions in [112]).

An obvious modification that generates different models is the choice of scaling that is applied during the asymptotic developments. Again, this choice should be closely related to the physics of the system of interest. An important example of this corresponds to the scaling of the Péclet number during homogenization of advection–diffusion equations (see Technical note 12). The form of the dispersion coefficients will strongly depend upon scalings. Other macroscale problems can be obtained by modifying the spatial averages (see [59]), or via mixed scales formulations (e.g., one phase is described at the pore-scale while the other is described at Darcy's scale and both are coupled, see [96]).

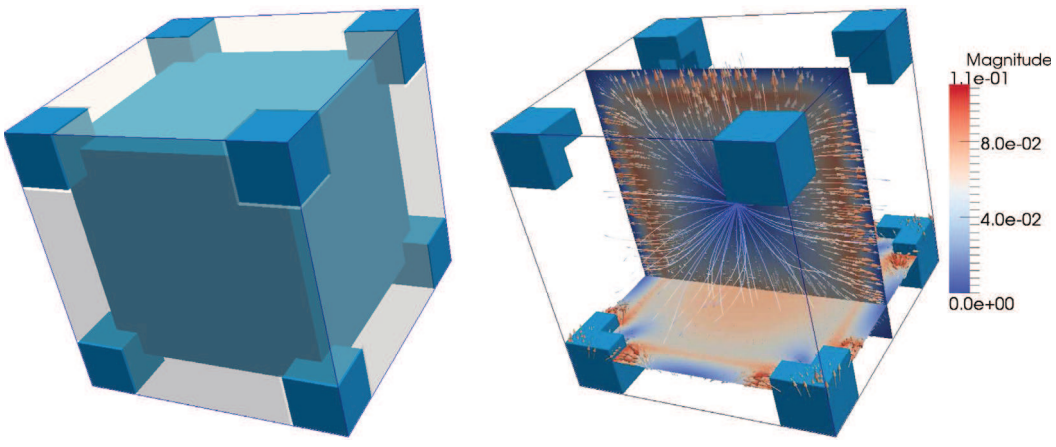
In multiphase systems, assumptions relative to spatial gradients of the microscale fields may yield different macroscale models. For instance, if the variations of the microscale field over the length-scale of the unit-cell are relatively small, we may assume that average values of the field within each phase are approximately the same, an assumption termed "local equilibrium" (see Fig. 19 and [113]). If this is not the case, then asymptotic models or higher order homogenization theories may be used (see e.g., [97]). In the volume averaging theory, higher order approaches often take the form of two- (or more) equation models in which the macroscale flux between phases is explicit. There are also many semi-heuristic models in hydrology that account for exchange dynamics between multiple regions, including formulations such as multi-rate transport models [114]. There are considerable challenges associated with understanding these models from a homogenization viewpoint.

9.6. Domains of validity

Determining the domains of validity of macroscale models is an essential step of homogenization. Constraints associated with homogenized models are often restrictive, and it is fundamental for practical applications to express them clearly. Further, we must understand how different macroscale models relate to each other (see [115] for a link between one- and two-equation models) and how they apply to specific situations. Open questions include: can we develop consistent higher order theories? Is there a systematic way of choosing models depending on the position in the space of dimensionless parameters? Can we combine different descriptions and create hybrid models that increase the domains of validity?

9.7. Weighted averages

In general, we expect average quantities to behave smoothly, so that they can be described by partial differential equations at the



(a) Geometry. (b) Solution of the vector field, $\mathbf{c}(\boldsymbol{y}) = \boldsymbol{\chi}(\boldsymbol{y})$.

Fig. 18. Resolution of the unit-cell problem in the three dimensional geometry (a). In (b), we plot the contour lines (center of the cube), the amplitude of $\mathbf{c}(\boldsymbol{y}) = \boldsymbol{\chi}(\boldsymbol{y})$ within two orthogonal cutting planes and the vector fields as arrows (glyphs) with origins within these same planes.

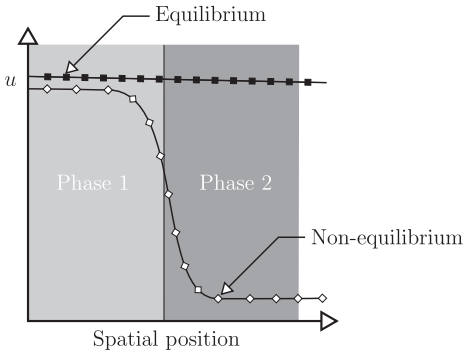


Fig. 19. Schematics illustrating the concept of local equilibrium between two phases.

macroscale. In the volume averaging framework, weighted average, as described in Technical note 2, may be used to regularize spatial fields, i.e., we can often choose the weight function so that it eliminates singularities. Further, weighted averages are also used to simplify macroscale expressions. For example, Quintard and Whitaker [58,59] show that, in classes of*ordered porous media,* Darcy's law is obtained for the hydrostatic problem only for a careful choice of weighted average.

9.8. Mathematical homogenization

From a purely mathematical point of view, the methods presented in this paper are formal but not rigorous. For example, the *ansatz* that u_ε can be expressed as a two-scale expansion is generally not correct at all orders. Consequently, an entire branch of homogenization has been developed, that we will term "mathematical homogenization" for the remainder of this discussion. Mathematical homogenization is concerned with constructing the function spaces, topologies and notions of convergence that allow a rigorous study of the problem.

In this paper, we focused on periodic homogenization since it lends itself very well for engineering applications. We emphasize that there are homogenization methods that apply to nonperiodic cases. These are based on convergence concepts such as Γ-convergence, *G*-convergence or *H*-convergence (see e.g., [49]). These methods are extremely powerful and can provide great mathematical insight into the problems, but do not generally yield a unit-cell problem and explicit formulae.

More recently, a particularly popular mathematical homogenization technique was introduced for *periodic* media. Initial propositions were made by Ngutseng [116] and the theory was further developed by Allaire (see [95]). It relies on the notion of two-scale convergence, a special type of weak convergence that is defined in the Technical note 20, and its properties, in particular regarding the *existence and uniqueness* of two-scale limits. Its popularity stems from the fact that it is a self-contained rigorous method that does not require two separate steps (formal asymptotics and then convergence), as was the case for early convergence results (see e.g., [47]).

Technical note 20: Let u_ε be a bounded sequence in $L^2(\Omega)$. There exists a subsequence, still denoted u_ε, and a function $u_0(\boldsymbol{x},\boldsymbol{y}) \in L^2(\Omega \times Y)$ (Y is the unit cube) such that

$$\lim_{\varepsilon\to 0} \int_\Omega u_\varepsilon(\boldsymbol{x})\psi(\boldsymbol{x},\boldsymbol{y})d\boldsymbol{x} = \int_\Omega \int_Y u_0(\boldsymbol{x},\boldsymbol{y})\psi(\boldsymbol{x},\boldsymbol{y})d\boldsymbol{x}d\boldsymbol{y},$$

for any smooth function $\psi(\boldsymbol{x},\boldsymbol{y})$, which is Y-periodic in $\boldsymbol{y}$. This sequence u_ε is said to two-scale converge to $u_0(\boldsymbol{x},\boldsymbol{y})$.

10. Conclusions

This paper was initiated on the basis that very few works have studied relationships between different homogenization methods. As a result, there is little connection and exchange of knowledge between larger bodies of the literature. In this paper, we have provided a detailed comparison of volume averaging with closure and formal periodic asymptotics based on multiscale expansions. We have shown that, in a simple parabolic case, the main differences between both methods are not the results themselves, but the methodologies, the insight that they provide, the algorithms and the formalisms of assumptions. We have also detailed many points of the developments that are rarely discussed in the literature, which we hope will help graduate students and researchers entering this field.

We anticipate that this work will inspire researchers to fight dogmatic conflicts between different communities and focus on understanding relationships between the many upscaling methods that are available. These include, for instance, the generalized Taylor-Aris-Brenner method of moments, a variety of stochastic approaches or mixture theories. Comparison of volume averaging and multiscale asymptotics should also be performed in more complex cases, e.g., the case with sources and/or drift (see e.g., [117]) or the use of one- and two-equation models (see e.g., [118]). From our viewpoint, such knowledge would be an important step towards facilitating access to many important papers in the literature and could possibly shed a new light on multiscale problems.

Acknowledgements

This work was supported in part by Award No. KUK-C1–013-04 made by King Abdullah University of Science and Technology (KAUST). Dr. Wood was supported in part by the U. S. Department of Energy, Office of Science (Subsurface Biogeochemistry Research program through the PNNL Subsurface Science Focus Area), and by NSF Mathematics under Grant 1122699.

Appendix A. Rigorous derivation for volume averaging

Here, our goal is to provide a brief overview of a more rigorous manner to perform the order of magnitude analysis in the volume averaging methodology. Similar derivations for other approximations using the symbol $\simeq$ in Section 5 are straightforward, so that we will only perform calculations up to the point where we obtained the simplified form of the perturbation problem.

Recall that the quasi-stationary perturbation problem is

$$0 = \nabla\cdot(\mathbf{A}\cdot\nabla\tilde{u}) - \nabla\cdot\langle\mathbf{A}\cdot\nabla\tilde{u}\rangle + \nabla\cdot(\mathbf{A}\cdot\nabla\langle u\rangle) - \nabla \cdot\langle\mathbf{A}\cdot\nabla\langle u\rangle\rangle. \tag{A.1}$$

The last term on the right-hand side can be written

$$\nabla\cdot\langle\mathbf{A}\cdot\nabla\langle u\rangle\rangle = \nabla\cdot(\langle\mathbf{A}\rangle\cdot\nabla\langle u\rangle) + h.o.t.. \tag{A.2}$$

This stems from the Taylor expansion about $\boldsymbol{x}$ within the "average of the average" expression, i.e., (for a similar discussion, see Technical note 8)

$$\langle\mathbf{A}\cdot\nabla\langle u\rangle\rangle = \langle\mathbf{A}\rangle\cdot\nabla\langle u\rangle + \langle\mathbf{A}\cdot\boldsymbol{\zeta}\rangle\cdot\nabla\nabla\langle u\rangle + h.o.t., \tag{A.3}$$

so that we have

$$\nabla\cdot(\mathbf{A}\cdot\nabla\langle u\rangle) - \nabla\cdot\langle\mathbf{A}\cdot\nabla\langle u\rangle\rangle = \nabla\cdot\left(\tilde{\mathbf{A}}\cdot\nabla\langle u\rangle\right) - \nabla\cdot\langle\mathbf{A}\cdot\boldsymbol{\zeta}\rangle \cdot\nabla\nabla\langle u\rangle + h.o.t.. \tag{A.4}$$

Further, we can decompose the first term on the right-hand side as

$$\nabla \cdot \left(\tilde{\mathbf{A}} \cdot \nabla \langle u \rangle\right) = \nabla \cdot \left(\tilde{\mathbf{A}}\right) \cdot \nabla \langle u \rangle + \tilde{\mathbf{A}} : \nabla\nabla \langle u \rangle, \tag{A.5}$$

so that

$$0 = \nabla \cdot (\mathbf{A} \cdot \nabla \tilde{u}) - \nabla \cdot \langle \mathbf{A} \cdot \nabla \tilde{u} \rangle + \nabla \cdot \left(\tilde{\mathbf{A}}\right) \cdot \nabla \langle u \rangle + \tilde{\mathbf{A}} : \nabla\nabla \langle u \rangle - \nabla \cdot \langle \mathbf{A} \cdot \zeta \rangle \cdot \nabla\nabla \langle u \rangle + h.o.t.. \tag{A.6}$$

We rewrite derivatives of terms that vary at the microscale in the $\boldsymbol{y}$ coordinate system, so that

$$0 = \varepsilon^{-2} \nabla_{\boldsymbol{y}} \cdot (\mathbf{A}_{\boldsymbol{y}} \cdot \nabla \tilde{u}) - \varepsilon^{-2} \nabla_{\boldsymbol{y}} \cdot \langle \mathbf{A} \cdot \nabla_{\boldsymbol{y}} \tilde{u} \rangle + \varepsilon^{-1} \nabla_{\boldsymbol{y}} \cdot \left(\tilde{\mathbf{A}}\right) \cdot \nabla \langle u \rangle + \tilde{\mathbf{A}} : \nabla\nabla \langle u \rangle - \nabla_{\boldsymbol{y}} \cdot \langle \mathbf{A} \cdot \zeta^{\star}(\boldsymbol{y}) \rangle \cdot \nabla\nabla \langle u \rangle + h.o.t., \tag{A.7}$$

where $\zeta^{\star}(\boldsymbol{y}) = \varepsilon^{-1}\zeta(\boldsymbol{x})$. Multiplying by ε^2 on both sides yields

$$0 = \nabla_{\boldsymbol{y}} \cdot (\mathbf{A}_{\boldsymbol{y}} \cdot \nabla \tilde{u}) - \nabla_{\boldsymbol{y}} \cdot \langle \mathbf{A} \cdot \nabla_{\boldsymbol{y}} \tilde{u} \rangle + \varepsilon \nabla_{\boldsymbol{y}} \cdot \left(\tilde{\mathbf{A}}\right) \cdot \nabla \langle u \rangle + \mathcal{O}(\varepsilon^2). \tag{A.8}$$

Appendix B. Analytical resolution of the unit-cell problem for the stratified case

The orientation of strata is such that the unit normal vector on the interface pointing from ω to σ is $\boldsymbol{n}_{\omega\sigma} = [0, 1]^{\mathrm{T}}$ where T denotes transpose. Symmetries yield a one dimensional problem with variations along the y_2-axis:

$$\frac{d^2 c_{\alpha x}}{dy_2^2} = 0 \text{ and } \frac{d^2 c_{\alpha y}}{dy_2^2} = 0.$$

The flux boundary conditions read $A_\omega \frac{dc_{\omega y_2}}{dy_2} - A_\sigma \frac{dc_{\sigma y_2}}{dy_2} = A_\sigma - A_\omega$ and $A_\omega \frac{dc_{\omega y_1}}{dy_1} - A_\sigma \frac{dc_{\sigma y_1}}{dy_1} = 0$. We will denote by ω_1 the phase corresponding to $0 < y_2 < a, \sigma$ the phase corresponding to $a < y_2 < b$ and ω_2 the phase corresponding to $b < y_2 < 1$. The continuity of the vectors yields $c_{\omega_1 y_2}(0) = c_{\omega_2 y_2}(1)$, $c_{\omega_1 y_2}(a) = c_{\sigma y_2}(a)$ and $c_{\omega_2 y_2}(b) = c_{\sigma y_2}(b)$. It is straightforward to show that $c_{\alpha y_1} = 0$. Further, we obtain

$$c_{\omega_1 y_2} = B\frac{y_2}{\phi_\omega}, \quad c_{\omega_2 y_2} = B\left[\frac{y_2 - b}{\phi_\omega} - \frac{1}{2}\right], \quad c_{\sigma y_2} = -B\left[\frac{y_2 - a}{\phi_\sigma} - \frac{1}{2}\right],$$

with $B = \frac{\phi_\omega \phi_\sigma (A_\sigma - A_\omega)}{\phi_\sigma A_\omega + \phi_\omega A_\sigma}$ and $\phi_\sigma = b - a$, $\phi_\omega = 1 - b + a = 2a$ (the unit-cell is symmetric so that $1 - b = a$). The effective tensor reads

$$\mathbf{A}_e = \langle A(y_2)(\mathbf{I} + \nabla \mathbf{c}) \rangle = \left\langle A(y_2) \begin{bmatrix} 1 & 0 \\ 0 & 1 + \frac{dc_{y_2}}{dy_2} \end{bmatrix} \right\rangle.$$

So that

$$\mathrm{A}_{ey_2 y_2} = \phi_\omega A_\omega \left(1 + \frac{B}{\phi_\omega}\right) + \phi_\sigma A_\sigma \left(1 - \frac{B}{\phi_\sigma}\right) = \frac{A_\omega A_\sigma}{\phi_\sigma A_\omega + \phi_\omega A_\sigma} = \frac{1}{\sum \frac{\phi_i}{A_i}},$$

which is the weighted harmonic mean. Further,

$$\mathrm{A}_{ey_1 y_1} = \phi_\omega A_\omega + \phi_\sigma A_\sigma,$$

which is the weighted arithmetic mean.

References

[1] Maxwell J. A treatise on electricity and magnetism. MacMillan & Co., Oxford at the Clarendon Press, London, 1873.
[2] Einstein A. Eine neue Bestimmung der Moleküldimensionen. Ann Phys 1906;324(2):289–306.
[3] Milton G. The theory of composites. Cambridge.: University Press; 2002.
[4] Wood B, Whitaker S. Multi-species diffusion and reaction in biofilms and cellular media. Chem Eng Sci 2000;55:3397–418.
[5] Davit Y, Byrne H, Osborne J, Pitt-Francis J, Gavaghan D, Quintard M. Hydrodynamic dispersion within porous biofilms. Phys Rev E 2013;87:012718.
[6] Valdés-Parada F, Varela J, Alvarez-Ramirez J. Upscaling pollutant dispersion in the Mexico City Metropolitan Area. Physica A 2012;391:606–15.
[7] Luciano R. Homogenization technique and damage model for old masonry material. Int J Solids Struct 1997;34:3191–208.
[8] Henn N, Quintard M, Bourbiaux B, Sakthikumar S. Modelling of Conductive Faults with a Multiscale Approach. Oil Gas Sci Technol 2004;59(2):197–214.
[9] Lorthois S, Cassot F. Fractal analysis of vascular networks: Insights from morphogenesis. J Theoret Biol 2010;262(4):614–33.
[10] Darcy H. Les Fontaines Publiques de la Ville de Dijon. Paris: Dalmont; 1856.
[11] De Josselin De Jong G. Longitudinal and transverse diffusion in granular deposits. Trans Amer Geophys Union 1958;39(1):67–74.
[12] Scheidegger AE. General theory of dispersion in porous media. J Geophys Res 1961;66(10):3273–8.
[13] Bear J. On the tensor form of dispersion. J Geophys Res 1961;66(4):1185–97.
[14] Muskat M. Physical principles of oil production. New York: McGraw-Hill; 1949.
[15] Taylor G. Dispersion of soluble matter in solvent flowing slowly through a tube. Proc Roy Soc A 1953;219:186–203.
[16] Aris R. On the dispersion of a solute in a fluid flowing through a tube. Adv Water Resour 1956;235:67–77.
[17] Scheidegger A. The physics of flow through porous media. University of Toronto Press, 1957.
[18] Markov K. Elementary Micromechanics of Heterogeneous Media, in: Heterogeneous Media: Modeling and Simulation in Science. Birkhäuser Boston: Engineering and Technology; 2000.
[19] Cushman J, Bennethum L, Hu B. A primer on upscaling tools for porous media. Adv Water Resour 2002;25:1043–67.
[20] Brenner H. Dispersion resulting from flow through spatially periodic porous media. Phil Trans Roy Soc Lond A 1980;297:81.
[21] Matheron G. Genèse et signification énergétique de la loi de Darcy. Rev Inst Français du Pétrole 1966;21(11):1697–706.
[22] Gelhar LW, Gutjahr AL, Naff RL. Stochastic Analysis of Macrodispersion in a Stratified Aquifer. Water Resour Res 1979;15(6):1387–97.
[23] Dagan G. The generalization of Darcy's law for nonuniform flows. Water Resour Res 1979;15(1):1–7.
[24] Dagan G. Flow and Transport in Porous Formations. Heidelberg, Berlin, New York: Springer-Verlag; 1989.
[25] Bendsoe M, Kikuchi N. Generating optimal topologies in structural design using a homogenization method. Comput Meth Appl Mech Eng 1988;71(2):197–224.
[26] Allaire G. Shape optimization by the homogenization method. New York: Springer-Verlag; 2002.
[27] Bourgeat A, Quintard M, Whitaker S. Eléments de comparaison entre la méthode d'homogénéisation et la méthode de prise de moyenne avec fermeture, C RAcad Sci Paris 306, sér II 1988. p. 463–6.
[28] Marle CM. Application de la méthode de la thermodynamique des processus irréversible á l'écoulement d'un fluide à travers un milieux poreux. Bull RILEM 1965;29:1066–71.
[29] Marle CM. Ecoulements monophasiques en milieu poreux. Rev Inst Français du Pétrole 1967;22(10):1471–509.
[30] Whitaker S. Diffusion and dispersion in porous media. AICHE J 1967;13:420–7.
[31] Slattery JC. Flow of Viscoelastic Fluids Through Porous Media. AIChE J 1967;13:1066–71.
[32] Mokadam RG. Thermodyn Anal Darcy Law 1961;28(2):208–12.
[33] Biot MA. Mechanics of Deformation and Acoustic Propagation in Porous Media. J Appl Phys 1962;33(4):1482–98.
[34] Hassanizadeh M, Gray WG. General Conservation Equations for Multi-Phase Systems: 1. Averaging procedure. Adv Water Resour 1979;2:131–44.
[35] Bennethum L, Cushman J. Multiscale, hybrid mixture theory for swelling systems. I. Balance laws. Int J Eng Sci 1996;34(2):125–45.
[36] Bowen RM. Continuum physics, chap. theory of mixtures. New York: Academic Press; 1976.
[37] Bear J, Bachmat Y. Introduction to Modeling of Transport Phenomena in Porous Media, Theory and Applications of Transport in Porous Media. Springer; 1990.
[38] Hsu C. A Closure Model for Transient Heat Conduction in Porous Media. J Heat Transfer 1999;121:733–9.
[39] Nakayama A, Kuwahara F, Sugiyama M, Xu G. A Two-Energy Equation Model for Conduction and Convection in Porous Media. Int J Heat Mass Transfer 2001;44(22):4375–9.
[40] Whitaker S. The Method of Volume Averaging. Dordrecht, The Netherlands: Kluwer Academic Publishers; 1999.
[41] Bogoliubov N, Mitropolsky Y. Asymptotic Methods in the Theory of Nonlinear Oscillations. New York: Gordon and Breach; 1961.
[42] Cole J. Uniformly valid approximations for certain nonlinear differential equations. In: La Salle JP, Lefschetz S, editors. Proceedings of the 1961 international symposium on nonlinear differential equations and nonlinear mechanics. New York: Academic Press; 1963. p. 113–20.
[43] Babuška I. Homogenization approach in engineering. Lecture note in Econom. and Math. Systems (Second Internat. Sympos., Versailles, 1975), vol. 134 (Part 1), 1976. p. 137–53.

[44] Sanchez-Palencia E. Solutions périodiques par rapport aux variables d'espaces et applications. C R Acad Sci Paris, Sér A–B 1970;271:A1129–32.
[45] Sanchez-Palencia E. Equations aux dérivées partielles dans un type de milieux hétérogènes. C R Acad Sci Paris, Sér A–B 1971;272:A1410–3.
[46] Ene H, Sanchez-Palencia E. Equations et phénomènes de surface pour l'écoulement dans un modèle de milieu poreux. J Mecanique 1975;14:73–108.
[47] Bensoussan A, Lions J-L, Papanicolaou G. Asymptotic analysis for periodic structures, Studies in Mathematics and its Applications 5. Amsterdam: North-Holland Publishing Company; 1978.
[48] Spagnolo S. Sulla convergenza di soluzioni di equazioni paraboliche ed ellittiche. Ann Sc Norm Super Pisa, Sci Fis Mat, III 1968;22:571–97.
[49] Tartar L. The general theory of homogenization: a personalized introduction, Lecture notes of the Unione Matematica Italiana. Springer; 2009.
[50] Hornung U, editor. Homogenization and porous media. Heidelberg, Berlin, New York: Springer-Verlag; 1997.
[51] Auriault J-L. Heterogeneous medium. Is an equivalent macroscopic description possible? Int J Eng Sci 1991;29(7):785–95.
[52] Auriault J-L, Adler P. Taylor dispersion in porous media: Analysis by multiple scale expansions. Adv Water Resour 1995;18(4):217–26.
[53] Aris R. Vectors, tensors, and the basic equations of fluid mechanics. Dover; 1989.
[54] Simmonds J. A brief on tensor analysis. Springer-Verlag; 1994.
[55] Bird R, Stewart W, Lightfoot E. Transport phenomena. New York: John Wiley and Sons; 1960.
[56] Quintard M, Whitaker S. Transport in chemically and mechanically heterogeneous porous media I: Theoretical development of region averaged equations for slightly compressible single-phase flow. Adv Water Resour 1996;19:29–47.
[57] Baveye P, Sposito G. The operational significance of the continuum hypothesis in the theory of water movement through soils and aquifers. Water Resour Res 1984;20:521–30.
[58] Quintard M, Whitaker S. Transport in ordered and disordered porous media I: The cellular average and the use of weighting functions. Trans Porous Med 1994;14(2):163–77.
[59] Quintard M, Whitaker S. Transport in ordered and disordered porous media. II: Generalized volume averaging. Trans Porous Med 1994;14(2): 179–206.
[60] Pride S, Berryman J. Connecting theory to experiment in poroelasticity. J Mech Phys Solids 1998;46(4):719–47.
[61] Prat M. On the boundary conditions at the macroscopic level. Trans Porous Med 1989;4(3):259–80.
[62] Beavers G, Joseph D. Boundary conditions at a naturally permeable wall. J Fluid Mech 1967;30:197–207.
[63] Valdés-Parada F, Goyeau B, Ochoa-Tapia J. Diffusive mass transfer between a microporous medium and an homogeneous fluid: Jump boundary conditions. Chem Eng Sci 2006;61:1692–704.
[64] Gray W, Lee P. On the theorems for local volume averaging of multiphase systems. Int J Multiphase Flow 1977;3:333–42.
[65] Cushman J. Proofs of the volume averaging theorems for multiphase flow. Adv Water Resour 1982;5:248–53.
[66] Gray W. Mathematical tools for changing spatial scales in the analysis of physical systems. Boca Raton, FL: CRC Press; 1993.
[67] Schwartz L. La théorie des distributions. Hermann Paris, 1978.
[68] Bachmat Y, Bear J. Macroscopic modelling of transport phenomena in porous media. 1: The continuum approach. Trans Porous Med 1986;1:213–40.
[69] Koch D, Brady J. A non-local description of advection-diffusion with application to dispersion in porous media. J Fluid Mech Digit Arch 1987;180:387–403.
[70] Koch J, Brady J. Anomalous diffusion in heterogeneous porous media. Phys Fluids 1988;31:965–73.
[71] Wood B, Valdés-Parada F. Volume averaging: local and nonlocal closures using a Green's function approach. Adv Water Resour 2012;51:139–67.
[72] Christakos G. Random field models in earth sciences. San Diego: Academic Press; 1992.
[73] Wood B. Technical note: Revisiting the geometric theorems for volume averaging and the meaning of disorder in porous media. Adv Water Resour, SI Whitaker.
[74] Goyeau B, Bousquet-Melou P, Gobin D, Quintard M, Fichot F. Macroscopic modeling of columnar dendritic solidification. Comput Appl Math 2004;23:381–400.
[75] Bousquet-Melou P, Goyeau B, Quintard M, Fichot F, Gobin D. Average momentum equation for interdendritic flow in a solidifying columnar mushy zone. Int J Heat Mass Transfer 2002;45:3651–65.
[76] Ahmadi A, Quintard M. Large-scale properties for two-phase flow in random porous media. J Hydrol 1996;183:69–99.
[77] Wood B, Cherblanc F, Quintard M, Whitaker S. Volume averaging for determining the effective dispersion tensor: closure using periodic unit cells and comparison with ensemble averaging. Water Resour Res 2003;39(6):1154–73.
[78] Wildenschild D, Sheppard A. Diffusion and dispersion in porous media. Adv Water Resour 2013;51:217–46.
[79] Metzler R, Klafter J. The random walk's guide to anomalous diffusion: a fractional dynamics approach. Phys Rep 2000;339:1–77.
[80] Valdés-Parada F, Alvarez-Ramirez J. Frequency-dependent dispersion in porous media. Phys Rev E 2011;84:031201.
[81] Davit Y, Quintard M. Comment on frequency-dependent dispersion in porous media. Phys Rev E 2012;86:013201.
[82] Valdés-Parada F, Alvarez-Ramirez J. Reply to comment on 'Frequency-dependent dispersion in porous media'. Phys Rev E 2012;86:013202.
[83] Renard P, de Marsily G. Calculating equivalent permeability: a review. Adv Water Resour 1997;26:253–78.
[84] Lunati I, Bernard D, Giudici M, Parravicini G, Ponzini G. A numerical comparison between two upscaling techniques: non-local inverse based scaling and simplified renormalization. Adv Water Resour 2001;24:913–29.
[85] Wiener O. Abhandlungen der Mathema sch. Phys.. Klasse Königlichen Sächsischen Ges. 1967;32:309.
[86] Hashin Z, Shtrikman S. A variational approach to the elastic behavior of multiphase minerals. J Mech Phys Solids 1963;11:127–40.
[87] Cardwell W, Parsons R. Average permeabilities of heterogeneous oil sands. Trans Am Inst Min Metall Pet Eng 1945;160:34–42.
[88] Matheron G. Eléments pour une théorie des milieux poreux. Masson et Cie Ed., Paris, 1967.
[89] Rubinstein J, Torquato S. Flow in random porous media: mathematical formulation, variational principles, and rigorous bounds. J Fluid Mech 1989;206:25–46.
[90] Ene H. Estimations du tenseur de perméabilité. C R Acad Sci II Paris 1991;312:1269–72.
[91] Davit Y, Debenest G, Wood BD, Quintard M. Modeling non-equilibrium mass transport in biologically reactive porous media. Adv Water Resour 2010;33:1075–93.
[92] Veran S, Aspa Y, Quintard M. Effective boundary conditions for rough reactive walls in laminar boundary layers. Int J Heat Mass Transfer 2009;52:3712–25.
[93] Ahmadi A, Quintard M, Whitaker S. Transport in chemically and mechanically heterogeneous porous media, V, two-equation model for solute transport with adsorption. Adv Water Resour 1998;22:59–86.
[94] Zanotti F, Carbonell R. Development of transport equations for multiphase system – 1: General Development for two phase system. Chem Eng Sci 1984;39(2):263–78.
[95] Allaire G. Homogenization and two-scale convergence. SIAM J Math Anal 1992;6:1482–518.
[96] Golfier F, Quintard M, Cherblanc F, Harvey C, Zinn B, Wood B. Comparison of theory and experiment for solute transport in highly heterogeneous porous medium. Adv Water Resour 2007;30:2235–61.
[97] Davit Y, Quintard M, Debenest G. Equivalence between volume averaging and moments matching techniques for mass transport models in porous media. Int J Heat Mass Transfer 2010;53:4985–93.
[98] Laribi S, Bertin H, Quintard M. Ecoulements polyphasiques en milieu poreux stratifié: résultats expérimentaux et interprétation. C R Acad Sci Paris 1990;311:271–6.
[99] Quintard M, Whitaker S. Two-phase flow in heterogeneous porous media I: the influence of large spatial and temporal gradients. Trans Porous Med 1990;5:341–79.
[100] Quintard M, Whitaker S. Ecoulement monophasique en milieu poreux: effet des hétérogénéités locales. J Meca Theor Appl 1987;6(5):691–726.
[101] Chang H-C. Effective diffusion and conduction in two-phase media: a unified approach. AIChE J 1983;29:846–53.
[102] Ochoa-Tapia J, Stroeve P, Whitaker S. Diffusive transport in two-phase media: spatially periodic models and Maxwell's theory for isotropic and anisotropic systems. Chem Eng Sci 1994;49(5):709–26.
[103] Quintard M. Diffusion in isotropic and anisotropic porous systems: three-dimensional calculations. Trans Porous Med 1993;11:187–99.
[104] Wood B, Whitaker S. Diffusion and reaction in biofilms. Chem Eng Sci 1998;53:397–425.
[105] Quintard M, Cherblanc F, Whitaker S. Dispersion in heterogeneous porous media: one-equation non-equilibrium model. Trans Porous Med 2001;44(1): 181–203.
[106] Zygalakis K, Roose T. A mathematical model for investigating the effect of cluster roots on plant nutrient uptake. Eur Phys J Spec Top 2012;204:103–18.
[107] Shipley R, Jones G, Dyson R, Sengers B, Bailey C, Catt C, Please C, Malda J. Design criteria for a printed tissue engineering construct: a mathematical homogenization approach. J Theoret Biol 2009;259(3):489–502.
[108] Lions J, Lukkassen D, Persson L, Wall P. Reiterated homogenization of nonlinear monotone operators. Chin. Annal Math. 22.
[109] Barenblatt G, Zheltov I, Kochina I. Basic concepts in the theory of homogeneous liquids in fissured rocks. J Appl Math (USSR) 1960;24(5):1286–303.
[110] Lions J. Some methods in the mathematical analysis of systems and their control. New York: Gordon and Breach; 1981.
[111] Brinkman HC. A calculation of the viscous force exerted by a flowing fluid on a dense swarm of particles. Appl Scient Res 1949;1:27–34.
[112] Lasseux D, Abbasian Arani A, Ahmadi A. On the stationary macroscopic inertial effects for one phase flow in ordered and disordered porous media. Phys Fluids 2011;23:073103.
[113] Golfier F, Wood B, Orgogozo L, Quintard M, Buès M. Biofilms in porous media: development of macroscopic transport equations via volume averaging with closure for local mass equilibrium conditions. Adv Water Resour 2009;32(3):463–85.
[114] Haggerty R, Gorelick S. Multiple-rate mass transfer for modeling diffusion and surface reactions in media with pore-scale heterogeneity. Water Resour Res 1995;31:2383–400.
[115] Davit Y, Wood B, Debenest G, Quintard M. Correspondence between one- and two-equation models for solute transport in two-region heterogeneous porous media. Trans Porous Med 2012;95(1):213–38.

[116] Nguetseng G. A general convergence result for a functional related to the theory of homogenization. SIAM J Math Anal 1989;20:608–23.
[117] Allaire G, Raphael A-L. Homogenization of a convection–diffusion model with reaction in a porous medium. C R Acad Sci Paris, Ser I 2007;344(8):523–8.
[118] Auriault J-L, Ene H. Macroscopic modelling of heat transfer in composites with interfacial thermal barrier. Int J Heat Mass Transfer 1994;37(18): 2885–92.