Lippmann-Schwinger spectrum, composite materials eigenstates and their role in computational homogenization

Cédric Bellis, Hervé Moulinec

Lippmann-Schwinger spectrum, composite materials eigenstates and their role in computational homogenization

Cédric Bellis and Hervé Moulinec

Aix Marseille Univ, CNRS, Centrale Med, LMA UMR7031, Marseille, France

September 2025

Abstract

Focusing on the homogenization of periodic composite materials, this study investigates computational methods based on volume integral equations. Such formulations are revisited from the standpoint of the preconditioning of the original cell problem by the introduction of a comparison material. This allows to recover simple convergence criteria for iterative steepest-descent and fixed-point schemes for composites with general non-linear behaviour. In the case of linear materials, the preconditioned volume integral formulation coincides with the well-known Lippmann-Schwinger equation. The spectral properties of the featured linear integral operator, which is bounded and self-adjoint, is investigated in order to shed light on the behaviour of conventional computational homogenization methods. The so-called Lippmann-Schwinger spectrum is analysed, with its bounds governing the convergence rate of iterative solution methods. The associated eigenvectors, which constitute the eigenstates of the composite material considered, are also described in detail to understand their role in constructing the solution to the cell problem and ultimately in computing the effective properties. Formulated in the continuous setting, this analysis is followed by the investigation of a discrete representation of the integral operator considered. A number of examples on synthetic microstructures are finally considered in the conductivity setting to illustrate the obtained theoretical results and highlight the role of the spectral properties in the operation of computational homogenization methods. This paves the way for the development of reduced models and more efficient computations.

1 Introduction

When considering the homogenization of heterogeneous materials, one has to solve for the local fields in a representative volume element subjected to a uniform loading. This defines a cell problem, which can take the form of local equations or of an equivalent volume integral equation, as recognised in [25, 45]. When the composite material considered is made of constituents with linear behaviours, such an integral equation reduces to the so-called Lippmann-Schwinger equation, which is also widely employed in scattering theory. Establishing such a formulation relies on the introduction of a linear, and in most cases also uniform, comparison medium with its companion Green's function, an approach that traces back to [12]. In the case of periodic composites, this type of formulation has provided the foundation for the fast Fourier transform-based numerical method developed in [35, 36] to compute the effective properties of linear and non-linear composites. This method relies on an iterative scheme that efficiently alternate between the physical space, where the constitutive properties are accounted for, and the Fourier space for applying the Green's operator associated with a homogeneous comparison material suitably chosen. This has opened the door to a breadth of successful developments since, see e.g. the review articles [44, 29].

Prior to, and then in parallel with, the development of these numerical methods, theoretical results were obtained using variational principles and volume integral formulations, in particular to establish bounds on the homogenized properties, starting with the seminal works [18, 19]. In addition, the analytic properties of the effective tensor of the composite as a function of the parameters of its constituents have also been the subject of a number of studies, initiated with [5, 33, 16] for dielectric (or conductive) composites. In the specific case of two-component materials, it is shown that the effective properties can be represented as an integral of a rational function of the material contrast between phases, see [34, Chap. 18]. This formulation turns out to be a Riemann-Stieltjes integral representation associated with the spectral measure of the linear integral operator featured in the Lippmann-Schwinger formulation of the homogenization problem, with the reference medium being defined as one of the two phases. Describing the spectrum of this operator and its associated eigenvectors, also referred to as eigenstates of the composite as in [6], or similarly in [31], is thus key for understanding the expression of the effective properties as functions of the composite micro-geometry and constituent parameters. This approach has been extended to other physics, in particular to elasticity [21], and to the field of inverse homogenization [10, 46] where information on the microstructure is recovered from the knowledge of the effective properties.

To bridge the gap between the generic computational methods on the one hand and the theoretical homogenization results for two-phase composites on the other hand, the perspective adopted in the present study is that the introduction of the comparison medium amounts to a preconditioning of the original cell problem. In the context of solving partial differential equations, preconditioning is a well-known method, see e.g. [39, 13, 30]. It is then crucial to analyse the spectral properties of the preconditioned operator as the performances of iterative solution schemes, such as the conjugate gradient or fixed-point iterations, depend on them, a study that is carried out for example in [15, 41, 40] for second-order differential operators relevant to the problems considered here. This makes it possible to find an optimal preconditioner, i.e., in the present context, an optimal comparison medium. Noticeably, the connection between this type of approach and computational homogenization methods have recently been made in [28, 26, 27]. Spectral properties have also been investigated earlier in [4] in the case of laminated composites.

In the context of computational homogenization methods and to connect with the known theoretical results, the overall objective of the present study is to shed some light on the implications of the preconditioning strategy based on the introduction of the comparison medium and, in the linear case, on the role of the spectral properties of the integral operator of the Lippmann-Schwinger equation in computing the solution to the cell problem and thus the effective properties of a given composite. The results presented hereinafter are intended to be formulated in a framework as general as possible, encompassing in particular the conductivity and elasticity cases, as well as related models, see e.g. [34, Chap. 2]. However, for illustration purposes, some results will be further particularised in the conductivity case, which will then be clearly indicated. Therefore, unless specified otherwise, the presented results are applicable to both the elasticity and conductivity cases.

The homogenization problem is presented in a variational setting in Section 2 for generic composite materials, potentially governed by a non-quadratic energy density. The introduction of the comparison medium is made through the definition of an energetic inner product. This naturally leads to the identification of the gradient of the energy functional as a preconditioned version of the operator featured in the original cell problem governing the local fields. Iterative solutions methods are then studied, starting with steepest descent, which we connect to fixed-point iterations, revisiting the convergence criteria conventionally adopted. A formulation using the unknown potential (or displacement) field is adopted in this study, and the link is also made with the gradient (or strain) based formulations that are widely used in applications. This analysis is then particularised to the case of linear composites, so as to obtain the well-known Lippmann-Schwinger integral equation. The featured linear operator, which is bounded and self-adjoint, is analysed in Section 3 with bounds on its eigenvalues

and a description, in the conductivity setting, of the associated eigenvectors, referred to as eigenstates of the composite. Section 4 focuses on the role of the spectral properties of the operator considered, namely how the solution to the cell problem and the effective properties are expressed in terms of the eigenvalues and eigenvectors in the case of multiple phase composites. These developments are then particularised to two-component materials. This analysis conducted in an infinite-dimensional functional setting is then transposed in the context of numerical methods, with a discretisation scheme presented in Sec. 5. A matrix formulation of the eigenvalue problem is investigated in connection with the conventional tools of FFT-based homogenization methods. Finally, some numerical results are presented and discussed in Section 6 for a set of synthetic microstructures. Some numerical spectra are described together with the associated eigenstates, highlighting in particular the very small portion of the spectrum that actually plays a role in the computation of the solution to the homogenization problem.

2 Homogenization and comparison medium preconditioning

2.1 Variational formulation

Consider a periodic composite material, with representative cell $\Omega \subset \mathbb{R}^d$ and characterised by an energy density $w(\boldsymbol{x}, \boldsymbol{\varepsilon}) : \Omega \times \mathbb{E} \to \mathbb{R}$, with compatible fields taking values in $\mathbb{E}$, e.g. $\mathbb{E} = \mathbb{R}^d$ in conductivity and $\mathbb{R}^{d\times d}_{\text{sym}}$ in elasticity. The density w is assumed to be smooth and convex in its second variable $\boldsymbol{\varepsilon}$. When a loading is applied to the composite, in the form of a uniform term $\bar{\boldsymbol{\varepsilon}} \in \mathbb{E}$, then the objective considered is to compute its effective energy $W_{\text{eff}}(\bar{\boldsymbol{\varepsilon}})$ through the minimization of the macroscopic energy W as:

$$W_{\text{eff}}(\bar{\boldsymbol{\varepsilon}}) = \min_{v^* \in K_{\text{per}}(\Omega)} W(v^*) \qquad \text{with} \qquad W(v^*) = \langle w(\cdot, \bar{\boldsymbol{\varepsilon}} + \boldsymbol{\nabla} v^*) \rangle, \tag{1}$$

where the averaging operator $\langle \cdot \rangle$ is defined as

$$\langle h \rangle = \frac{1}{|\Omega|} \int_\Omega h(\boldsymbol{x}) \, \mathrm{d}V$$

and given the Sobolev space $H^1_{\text{per}}(\Omega)$ of periodic fields in Ω and its subspace:

$$K_{\text{per}}(\Omega) = \{ v^* \in H^1_{\text{per}}(\Omega) \ ; \ \langle v^* \rangle = 0 \}. \tag{2}$$

Periodic functions in the reference cell Ω are considered to be defined on this domain identified as a d-dimensional torus, a compact manifold with no boundary. Doing so naturally ensures periodicity for such functions. In addition, in (2), the Sobolev space $H^1_{\text{per}}(\Omega)$ of functions defined on Ω can be identified as $H^1_{\text{per}}(\Omega) = \{ v^* \in H^1_{\text{loc}}(\mathbb{R}^d) \ ; \ v^* \text{ is } \Omega\text{-periodic} \}$. Note that, in place of (2), we could have considered instead $\mathcal{K}_{\text{per}}(\Omega) = H^1_{\text{per}}(\Omega)/\mathbb{R}$ the quotient space of functions in $H^1_{\text{per}}(\Omega)$ that are equal up to an additive constant or, with a slight abuse of notation, a rigid body translation in elasticity. Yet, working in the subspace $K_{\text{per}}(\Omega)$ defined in (2) is a way to choose a representative element in the equivalence class $\mathcal{K}_{\text{per}}(\Omega)$, see e.g [11, 1]. In addition, the dual spaces $\mathcal{K}_{\text{per}}(\Omega)'$ and $K_{\text{per}}(\Omega)'$ of bounded linear functionals on $\mathcal{K}_{\text{per}}(\Omega)$ and $K_{\text{per}}(\Omega)$, respectively, are such that

$$\mathcal{K}_{\text{per}}(\Omega)' = \{ f \in H^1_{\text{per}}(\Omega)' \ ; \ \langle f, c \rangle = 0 \quad \forall c \in \mathbb{R} \}$$

where $\langle \cdot, \cdot \rangle$ denotes the relevant duality product such that

$$\langle f, [v^*] \rangle_{\mathcal{K}_{\text{per}}(\Omega)', \mathcal{K}_{\text{per}}(\Omega)} = \langle f, v^* \rangle_{K_{\text{per}}(\Omega)', K_{\text{per}}(\Omega)} = \langle f, v^* \rangle_{H^1_{\text{per}}(\Omega)', H^1_{\text{per}}(\Omega)}$$

given $[v^*]$ the class of equivalence of $v^* \in H^1_{\mathrm{per}}(\Omega)$. Therefore, for the sake of brevity, indices in duality products are omitted in what follows.

In this context, the derivative $DW[v^*]$ of W in $K_{\mathrm{per}}(\Omega)$ at v^* and in the direction $\tilde{v}^*$ reads

$$DW[v^*]\tilde{v}^* = \frac{1}{|\Omega|}\int_\Omega \frac{\partial w}{\partial \boldsymbol{\varepsilon}}(\boldsymbol{x}, \bar{\boldsymbol{\varepsilon}} + \boldsymbol{\nabla} v^*) \cdot \boldsymbol{\nabla}\tilde{v}^*(\boldsymbol{x})\,\mathrm{d}V = \frac{1}{|\Omega|}\int_\Omega -\operatorname{div}\left(\frac{\partial w}{\partial \boldsymbol{\varepsilon}}(\boldsymbol{x}, \bar{\boldsymbol{\varepsilon}} + \boldsymbol{\nabla} v^*)\right)\tilde{v}^*(\boldsymbol{x})\,\mathrm{d}V. \quad (3)$$

where the last equality follows by integration by parts. Given $\boldsymbol{h} \in \boldsymbol{L}^2_{\mathrm{per}}(\Omega)$, with $\boldsymbol{L}^2_{\mathrm{per}}(\Omega)$ being the space of square integrable periodic tensor-valued fields in Ω, then $f = -\operatorname{div}\boldsymbol{h}$ can be identified as an element in $K_{\mathrm{per}}(\Omega)'$ according to the relation

$$\langle f, v^*\rangle = \langle -\operatorname{div}\boldsymbol{h}, v^*\rangle = \frac{1}{|\Omega|}\int_\Omega \boldsymbol{h}(\boldsymbol{x})\boldsymbol{\nabla} v^*(\boldsymbol{x})\,\mathrm{d}V \qquad \forall v^* \in K_{\mathrm{per}}(\Omega).$$

Therefore, the identity (3) allows to identify $DW[v^*] \in K_{\mathrm{per}}(\Omega)'$ as:

$$DW[v^*] = -\operatorname{div}\left(\frac{\partial w}{\partial \boldsymbol{\varepsilon}}(\boldsymbol{x}, \bar{\boldsymbol{\varepsilon}} + \boldsymbol{\nabla} v^*)\right). \quad (4)$$

From (4), the optimality condition $DW[u^*] = 0$ for the minimizer $u^* \in K_{\mathrm{per}}(\Omega)$ of W in (1) yields the following local equation:

$$-\operatorname{div}\left(\frac{\partial w}{\partial \boldsymbol{\varepsilon}}(\boldsymbol{x}, \bar{\boldsymbol{\varepsilon}} + \boldsymbol{\nabla} u^*)\right) = \boldsymbol{0} \qquad \text{in } \Omega. \quad (5)$$

In this context, as is common practice in homogenization, a reference comparison medium can be introduced for computational purposes. Here, we adopt the formalism of [39, 13] to show that this amounts to precondition the equation (5) through the choice of an energetic inner product. Consider C_0, symmetric and positive definite, and satisfying the uniform bounds

$$m_0|\boldsymbol{\varepsilon}|^2 \leq C_0\boldsymbol{\varepsilon}\cdot\boldsymbol{\varepsilon} \leq M_0|\boldsymbol{\varepsilon}|^2 \qquad \forall \boldsymbol{\varepsilon} \in \mathbb{E}, \quad (6)$$

with $M_0 \geq m_0 > 0$ and where $|\cdot|$ denotes the Euclidean norm. This allows the following definition to be introduced.

Definition 1. *The space $K_{\mathrm{per}}(\Omega)$ is equipped with the following energetic inner product:*

$$(v_1^*, v_2^*) = \langle \boldsymbol{\nabla} v_1^* \cdot C_0\,\boldsymbol{\nabla} v_2^*\rangle,$$

which also defines a norm on $\mathcal{K}_{\mathrm{per}}(\Omega)$ according to the Poincaré-Wirtinger inequality, see [11].

Remark 1. *Note that, while $C_0 \succ 0$ could be chosen as a non-uniform symmetric field in Ω to define the inner product in the definition above, when it comes to its interpretation as a preconditioner, it is much more advantageous to define it as constant for practical computational reasons, see below. From now on, we will therefore assume that C_0 is uniform.*

Using Definition 1 then, according to the Riesz representation theorem, for any $f \in K_{\mathrm{per}}(\Omega)'$ there exists a unique $r^* \in K_{\mathrm{per}}(\Omega)$ such that

$$\langle f, v^*\rangle = (r^*, v^*) = \frac{1}{|\Omega|}\int_\Omega -\operatorname{div}\left(C_0\boldsymbol{\nabla} r^*\right)v^*\,\mathrm{d}V \qquad \forall v^* \in K_{\mathrm{per}}(\Omega),$$

where the last equality follows from Def. 1 by integration by parts. As a consequence, the field r^* is the weak solution of the problem $-\operatorname{div}\left(C_0\boldsymbol{\nabla} r^*\right) = f$, which is well-posed according to the Lax-Milgram

theorem [11, 1]. The so-called Riesz mapping $R_0 : K_{\text{per}}(\Omega)' \to K_{\text{per}}(\Omega)$ is thus defined by $r^* = R_0 f$ as the linear operator that can formally be expressed as:

$$R_0 f = (-\operatorname{div} C_0 \boldsymbol{\nabla})^{-1} f. \tag{7}$$

Now, considering $C_0 \succ 0$, which defines a uniform reference medium, then the associated gradient $W_0'(v^*) \in K_{\text{per}}(\Omega)$ of W at v^* and associated with Definition 1 can be identified as:

$$DW[v^*]\tilde{v}^* = (W_0'(v^*), \tilde{v}^*),$$

from which, using (4) and (7), we get $W_0'(v^*) = R_0 DW[v^*]$ that we can formally write as:

$$W_0'(v^*) = (-\operatorname{div} C_0 \boldsymbol{\nabla})^{-1} \left(-\operatorname{div}\left(\frac{\partial w}{\partial \boldsymbol{\varepsilon}}(\boldsymbol{x}, \bar{\boldsymbol{\varepsilon}} + \boldsymbol{\nabla} v^*) \right) \right). \tag{8}$$

Considering again the optimality condition by setting $W_0'(u^*) = 0$ yields an equation for u^*, which is now a preconditioned system comparatively to (4–5), where the preconditioner being given by the linear operator A_0 defined as:

$$A_0 : K_{\text{per}}(\Omega) \to K_{\text{per}}(\Omega)' \qquad \text{such that} \qquad A_0 v^* = -\operatorname{div} C_0 \boldsymbol{\nabla} v^*. \tag{9}$$

Note that A_0 is related to the Riesz mapping as $A_0 = R_0^{-1}$, and inverting A_0 to compute (8) is much easier if C_0 is uniform, see Remark 1. The idea of using a change of inner product to preconditioned a system of partial differential equations is extensively discussed in [39, 13] as the so-called *Sobolev gradient* approach, as well as in [30] where it is rather referred to as *operator preconditioning*. These ideas can be traced back to [22, Chap. XV].

2.2 Steepest descent and fixed-point iterations

2.2.1 Iterative scheme

The connection between the iterative schemes commonly employed in computational homogenization, starting from this initially introduced in [35, 36], and gradient-descent algorithms has been established in [20]. This has led to the flourishing development of accelerated methods since, see [44]. For completeness, we revisit here the basic ideas underlying this relationship, with an emphasis on preconditioning.

The solution u^* to (1) can be computed using steepest descent iterations which, given a local gradient $W_0'(u_n^*) \in K_{\text{per}}(\Omega)$ of the energy functional (1), relatively to the inner product in Definition 1 and a weight C_0, which we will specify later on, read as

$$u_{n+1}^* = u_n^* - \alpha_n W_0'(u_n^*), \tag{10}$$

where the step size α_n must be chosen so as to ensure convergence. To discuss this, let consider the operator T_α defined as

$$T_\alpha : K_{\text{per}}(\Omega) \to K_{\text{per}}(\Omega) \qquad \text{with} \qquad T_\alpha = I - \alpha R_0 DW, \tag{11}$$

with $\alpha \in \mathbb{R}^+$ and I being the identity operator on $K_{\text{per}}(\Omega)$, so that (10) can be recast as $u_{n+1}^* = T_\alpha(u_n^*)$. Obviously, the optimal solution u^* is a fixed-point of T_α, i.e. $T_\alpha(u^*) = u^*$ since $DW[u^*] = 0$. With further assumptions on the energy functional and a suitable choice of α we can also show that the operator T_α is a contraction mapping, which is a well-known result, see e.g. [8, 13]. Here, we establish this property under the following assumption.

Hypothesis 1. *The energy density is twice differentiable with* $w(\boldsymbol{x},\cdot) \in C^2(\mathbb{E},\mathbb{R})$*, uniformly in* Ω*, and there exist* $M > m > 0$ *such that the tangent modulus* $\partial^2 w/\partial \boldsymbol{\varepsilon}^2$ *satisfies:*

$$m|\tilde{\boldsymbol{\varepsilon}}|^2 \leq \frac{\partial^2 w}{\partial \boldsymbol{\varepsilon}^2}(\boldsymbol{x},\boldsymbol{\varepsilon})\tilde{\boldsymbol{\varepsilon}}\cdot\tilde{\boldsymbol{\varepsilon}} \leq M|\tilde{\boldsymbol{\varepsilon}}|^2 \qquad \forall \boldsymbol{x}\in\Omega, \forall \boldsymbol{\varepsilon},\tilde{\boldsymbol{\varepsilon}}\in\mathbb{E}. \tag{12}$$

In this setting, the following result applies, the proof of which is given in Appendix A.

Proposition 1. *Introducing the following symmetric linear operator from* $\mathbb{E}$ *into itself for all* $\boldsymbol{x}\in\Omega$*:*

$$P_\alpha(\boldsymbol{x},\boldsymbol{\varepsilon}) = I - \alpha\, C_0^{-1/2}\frac{\partial^2 w}{\partial \boldsymbol{\varepsilon}^2}(\boldsymbol{x},\boldsymbol{\varepsilon})C_0^{-1/2} \tag{13}$$

with $r_\alpha(\boldsymbol{x})$ *being its local spectral radius, i.e.*

$$r_\alpha(\boldsymbol{x}) = \sup_{\boldsymbol{\varepsilon},\tilde{\boldsymbol{\varepsilon}}\in\mathbb{E}} \frac{|P_\alpha(\boldsymbol{x},\boldsymbol{\varepsilon})\tilde{\boldsymbol{\varepsilon}}\cdot\tilde{\boldsymbol{\varepsilon}}|}{|\tilde{\boldsymbol{\varepsilon}}|^2}, \tag{14}$$

then, in the energetic norm $\|\cdot\|$ *associated with Def. 1, it holds*

$$\|T_\alpha(v_2^*) - T_\alpha(v_1^*)\| \leq \sup_{\boldsymbol{x}\in\Omega} r_\alpha(\boldsymbol{x})\|v_2^* - v_1^*\|. \tag{15}$$

As a consequence, the operator T_α *is a contraction mapping as long as* $\sup_\Omega r_\alpha < 1$*. In such a case, the sequence* $\{u_n^*\}$ *generated by (10) will converge to the energy minimizer* u^**, at a convergence rate at most equal to the supremum of the spectral radius of* P_α *defined above.*

2.2.2 Convergence analysis

Let us now examine the conditions on P_α in (13) under which the operator T_α is a contraction. Owing to Hypothesis 1 and the spectral bounds on C_0^{-1}, which can be deduced from (6), then one has the uniform bounds:

$$\frac{m}{M_0}|\tilde{\boldsymbol{\varepsilon}}|^2 \leq C_0^{-1/2}\frac{\partial^2 w}{\partial \boldsymbol{\varepsilon}^2}(\boldsymbol{x},\boldsymbol{\varepsilon})C_0^{-1/2}\tilde{\boldsymbol{\varepsilon}}\cdot\tilde{\boldsymbol{\varepsilon}} \leq \frac{M}{m_0}|\tilde{\boldsymbol{\varepsilon}}|^2 \qquad \forall \boldsymbol{x}\in\Omega, \forall \boldsymbol{\varepsilon},\tilde{\boldsymbol{\varepsilon}}\in\mathbb{E}, \tag{16}$$

so that the spectrum $\sigma(P_\alpha)$ of the operator $P_\alpha(\boldsymbol{x},\boldsymbol{\varepsilon})$ satisfies:

$$\sigma(P_\alpha) \subset \left[1-\alpha\frac{M}{m_0}\,,\; 1-\alpha\frac{m}{M_0}\right] \qquad \forall \boldsymbol{x}\in\Omega, \forall \boldsymbol{\varepsilon}\in\mathbb{E}. \tag{17}$$

Therefore, the spectral radius of the operator P_α defined in (14) is such that

$$\sup_{\boldsymbol{x}\in\Omega} r_\alpha(\boldsymbol{x}) \leq \max\left(\left|1-\alpha\frac{M}{m_0}\right|, \left|1-\alpha\frac{m}{M_0}\right|\right). \tag{18}$$

In addition, for any choice of inner product in Def. 1, the value of the optimal parameter α that minimizes the upper bound in (18) is

$$\alpha = \frac{2}{M/m_0 + m/M_0}, \tag{19}$$

and the associated bound reads:

$$\sup_{\boldsymbol{x}\in\Omega} r_\alpha(\boldsymbol{x}) \leq \left(\frac{M M_0 - m\, m_0}{M M_0 + m\, m_0}\right) < 1. \tag{20}$$

As a consequence, for *any comparison medium* C_0 used to define the inner product in Def. 1, provided that α is chosen as in (19) then the sequence $\{u_n^*\}$ generated by (10) will converge to the energy minimizer u^*, at a convergence rate governed by (20).

Remark 2. *Hypothesis 1 is expressed in terms of the spectral bounds of $\partial^2 w/\partial \boldsymbol{\varepsilon}^2$, with the energy density w being assumed to be sufficiently smooth for the former to be defined. Yet, the above convergence proof can be also established under weaker assumptions, namely Lipschitz continuity and coercivity of w relatively to its second variable.*

Finally, owing to (20), the convergence rate of the fixed-point iterations (10) is minimized when $m_0 = M_0$ i.e., according to the uniform bounds (6), the operator C_0 defining the inner product in Def. 1 is uniform and proportional to the identity, with a single eigenvalue. Constructing a uniform C_0 based on the material bounds of Hypothesis 1 and satisfying

$$C_0 = \frac{m+M}{2} I, \tag{21}$$

where I is the identity operator on $\mathbb{E}$, in elasticity or in conductivity. The choice above is only driven by mathematical arguments and the definition of C_0 need not be characteristic of an existing material. Nevertheless, in the case of elasticity, it would correspond to a material whose first Lamé parameter is zero, which is admissible. Then, the optimal step in (19) and the associated bound reduce to

$$\alpha = 1 \quad \text{and} \quad \sup_{\boldsymbol{x}\in\Omega} r_\alpha(\boldsymbol{x}) \leq \left(\frac{z-1}{z+1}\right) < 1 \quad \text{with} \quad z = \frac{M}{m} > 1. \tag{22}$$

Accordingly, we note $T = T_1$ and the associated iterates generated from (10) correspond to the primal variable formulation of the scheme of [36], for all types of material behavior, whether linear or nonlinear. With $z = M/m$ being interpreted as a contrast characterising the composite material considered, then one recovers the well-known convergence rate of the fixed-point scheme, see [37].

Remark 3. *Here, the convergence proof of the iterative scheme (10), which relies on the uniform upper bound (18) on the spectral radius of P_α, has only been established using the global spectral bounds (6) and (12) on C_0 and $\partial^2 w/\partial \boldsymbol{\varepsilon}^2$, independently. Yet, if we allow C_0 to have multiple eigenvalues, and provided that C_0 and $\partial^2 w/\partial \boldsymbol{\varepsilon}^2$ share the same eigenspaces uniformly in Ω, then we could establish tighter bounds on the spectrum of the composite operator $C_0^{-1/2}(\partial^2 w/\partial \boldsymbol{\varepsilon}^2)C_0^{-1/2}$ in (13), thus improving the upper bound (18). In other words, in the case of an anisotropic composite, rather than defining C_0 as proportional to the identity, the reference medium can be defined as anisotropic as well and optimized so as to improve the convergence rate of the iterative scheme (10). It is also common practice to update C_0 during iterations to improve convergence, see e.g. [28].*

2.2.3 Gradient-based formulation and Green's operator

For computational purposes, see [36], the local equation (5) can advantageously be replaced by an equivalent volume integral equation on the gradient, or strain, field. To do so, consider the periodic gradient Green's operator $\boldsymbol{\Gamma}_0 : \boldsymbol{L}^2_{\text{per}}(\Omega) \to \boldsymbol{L}^2_{\text{per}}(\Omega)$, acting in the space $\boldsymbol{L}^2_{\text{per}}(\Omega)$ of square integrable tensor-valued fields, as:

$$\boldsymbol{\Gamma}_0 : \boldsymbol{\tau} \mapsto \boldsymbol{\Gamma}_0\boldsymbol{\tau} = \boldsymbol{e}^* \text{ such that } \boldsymbol{e}^* = \boldsymbol{\nabla} v^*(\boldsymbol{x}) \text{ with } v^* \in K_{\text{per}}(\Omega) \text{ and } \operatorname{div}(C_0\boldsymbol{e}^* - \boldsymbol{\tau}) = 0 \text{ in } \Omega.$$

The operator $\boldsymbol{\Gamma}_0$ is an integral operator that can formally be written as

$$\boldsymbol{\Gamma}_0 = \boldsymbol{\nabla}(\operatorname{div} C_0 \boldsymbol{\nabla})^{-1} \operatorname{div} = -\boldsymbol{\nabla} A_0^{-1} \operatorname{div},$$

in terms of the preconditioning operator A_0 defined in (9). As a consequence, for all $v^* \in K_{\text{per}}(\Omega)$, it holds:

$$\boldsymbol{\Gamma}_0 C_0 \boldsymbol{\nabla} v^* = -\boldsymbol{\nabla} A_0^{-1} \operatorname{div} C_0 \boldsymbol{\nabla} v^* = \boldsymbol{\nabla} A_0^{-1} A_0 v^* = \boldsymbol{\nabla} v^*. \tag{23}$$

Accordingly, considering the reference medium C_0 chosen as in (21) and (22), then using the associated contraction mapping $T = T_1$ in (11) with (8), and applying the gradient operator to the iterates (10) yield

$$\begin{aligned}\boldsymbol{\nabla} u_{n+1}^* = \boldsymbol{\nabla} T(u_n^*) &= \boldsymbol{\nabla} u_n^* - \boldsymbol{\nabla} W_0'(u_n^*) \\ &= \boldsymbol{\nabla} u_n^* + \boldsymbol{\nabla} A_0^{-1} \operatorname{div} \frac{\partial w}{\partial \boldsymbol{\varepsilon}}(\boldsymbol{x}, \bar{\boldsymbol{\varepsilon}} + \boldsymbol{\nabla} u_n^*),\end{aligned}$$

so that, upon denoting $\boldsymbol{\varepsilon}_n^* = \boldsymbol{\nabla} u_n^*$, we finally obtain

$$\boldsymbol{\varepsilon}_{n+1}^* = \boldsymbol{\varepsilon}_n^* - \boldsymbol{\Gamma}_0 \frac{\partial w}{\partial \boldsymbol{\varepsilon}}(\boldsymbol{x}, \bar{\boldsymbol{\varepsilon}} + \boldsymbol{\varepsilon}_n^*). \tag{24}$$

Based on the discussion in Section 2.2, the iterative scheme above is convergent provided that C_0 is appropriately chosen depending on the properties of the energy density w. This corresponds to the original scheme of [36].

2.3 Linear case

2.3.1 Operators and iterative schemes

In this section, we particularise some of the previous results to the linear case, where the energy density reads:

$$w(\boldsymbol{x}, \boldsymbol{\varepsilon}) = \frac{1}{2} C(\boldsymbol{x}) \boldsymbol{\varepsilon} \cdot \boldsymbol{\varepsilon} \tag{25}$$

in terms of a symmetric, positive definite and bounded operator C on $\mathbb{E}$ into itself. In this context, the derivative of the energy functional reduces to:

$$DW[v^*] = Av^* - b \qquad \text{given} \qquad b = \operatorname{div} C\bar{\boldsymbol{\varepsilon}} \in K_{\text{per}}(\Omega)',$$

using the operator A formally defined as

$$A : K_{\text{per}}(\Omega) \to K_{\text{per}}(\Omega)' \qquad \text{with} \qquad Av^* = -\operatorname{div} C\boldsymbol{\nabla} v^*, \tag{26}$$

i.e. satisfying:

$$\langle Av^*, \tilde{v}^* \rangle = \frac{1}{|\Omega|} \int_\Omega \boldsymbol{\nabla} v^*(\boldsymbol{x}) \cdot C(\boldsymbol{x}) \boldsymbol{\nabla} \tilde{v}^*(\boldsymbol{x}) \, \mathrm{d}V \qquad \forall v^*, \tilde{v}^* \in K_{\text{per}}(\Omega). \tag{27}$$

The equation (5) can therefore be recast as the following problem:

$$\text{Find} \quad u^* \in K_{\text{per}}(\Omega) \quad \text{such that} \quad Au^* = b. \tag{28}$$

The identity (27) defines a bilinear form $a(\cdot, \cdot)$ on $K_{\text{per}}(\Omega) \times K_{\text{per}}(\Omega)$ such that $a(v^*, \tilde{v}^*) = \langle Av^*, \tilde{v}^* \rangle$ for all $v^*, \tilde{v}^*$. Therefore, (28) is equivalent to the weak formulation of the original equation (5) with the solution u^* satisfying

$$a(u^*, v^*) = \langle b, v^* \rangle \qquad \forall v^* \in K_{\text{per}}(\Omega).$$

In the linear case, the objective in homogenization is therefore to compute the effective operator C_{eff} that defines the effective energy W_{eff} through

$$\frac{1}{2} C_{\text{eff}} \, \bar{\boldsymbol{\varepsilon}} \cdot \bar{\boldsymbol{\varepsilon}} = \frac{1}{2} \langle C(\bar{\boldsymbol{\varepsilon}} + \boldsymbol{\nabla} u^*) \cdot (\bar{\boldsymbol{\varepsilon}} + \boldsymbol{\nabla} u^*) \rangle$$

once u^* has been computed from (28).

Using the reference comparison medium defined by C_0 and the associated operator A_0 in (9) then, owing to (8), the linear equation $Au^* = b$ in (28) is modified into the preconditioned form:

$$A_0^{-1} A u^* = A_0^{-1} b. \tag{29}$$

Definition 2. *By introducing the operator* $S = -A_0^{-1}\delta A : K_{\text{per}}(\Omega) \to K_{\text{per}}(\Omega)$ *with* $\delta A = (A - A_0)$ *and defining* $b_0 = A_0^{-1}b$*, then the preconditioned equation* (29) *can be rewritten as*

$$(I - S)u^* = b_0. \tag{30}$$

Considering the operator T_α in (11), then the latter is a contraction mapping if the spectral radius of $P_\alpha(\boldsymbol{x}, \boldsymbol{\varepsilon})$ in (13) is strictly bounded by one, uniformly in $\boldsymbol{\varepsilon} \in \mathbb{E}$, with in the linear setting:

$$P_\alpha(\boldsymbol{x}, \boldsymbol{\varepsilon}) \equiv P_\alpha(\boldsymbol{x}) = I - \alpha C_0^{-1/2} C(\boldsymbol{x}) C_0^{-1/2}. \tag{31}$$

This is satisfied under the conditions (21) and (22), so that in such a case the associated operator $T = T_1$ yields convergent fixed-point iterations $u^*_{n+1} = Tu^*_n$ as in (10). Given that

$$Tu^*_n = u^*_n - A_0^{-1}(Au^*_n - b) = b_0 + Su^*_n. \tag{32}$$

then the fixed-point iterations read

$$u^*_n = T^n u^*_0 = b_0 + Sb_0 + \cdots + S^{n-1}b_0 + S^n u^*_0,$$

which, upon setting $u^*_0 = b_0$, corresponds to the iterates generated using the following convergent Neumann series:

$$u^* = \sum_{n \geq 0} S^n \, b_0 \tag{33}$$

that provides the unique inverse to (30).

2.3.2 Gradient-based formulation

In the linear case (25) where

$$\frac{\partial w}{\partial \boldsymbol{\varepsilon}}(\boldsymbol{x}, \bar{\boldsymbol{\varepsilon}} + \boldsymbol{\varepsilon}^*_n) = C(\boldsymbol{x})(\bar{\boldsymbol{\varepsilon}} + \boldsymbol{\varepsilon}^*_n(\boldsymbol{x})),$$

then (24) reduces to

$$\boldsymbol{\varepsilon}^*_{n+1} = \boldsymbol{\varepsilon}^*_n - \boldsymbol{\Gamma}_0 C(\bar{\boldsymbol{\varepsilon}} + \boldsymbol{\varepsilon}^*_n).$$

Upon introducing $\delta C(\boldsymbol{x}) = (C(\boldsymbol{x}) - C_0)$, using the property (23) and the fact that $\boldsymbol{\Gamma}_0 C_0 \bar{\boldsymbol{\varepsilon}} = \boldsymbol{0}$, then the equation above can be rewritten as

$$\boldsymbol{\varepsilon}_{n+1} = \bar{\boldsymbol{\varepsilon}} - \boldsymbol{\Gamma}_0 \delta C \boldsymbol{\varepsilon}_n$$

in terms of the total field $\boldsymbol{\varepsilon}_n = \bar{\boldsymbol{\varepsilon}} + \boldsymbol{\varepsilon}^*_n$. Again, this corresponds to the iterates generated by the following convergent (when C_0 is appropriately chosen) Neumann series

$$\boldsymbol{\varepsilon} = \sum_{n \geq 0} (-\boldsymbol{\Gamma}_0 \delta C)^n \, \bar{\boldsymbol{\varepsilon}}, \tag{34}$$

which compute the inverse to the well-known gradient-based Lippmann-Schwinger integral equation:

$$\boldsymbol{\varepsilon} + \boldsymbol{\Gamma}_0 \delta C \boldsymbol{\varepsilon} = \bar{\boldsymbol{\varepsilon}}. \tag{35}$$

Remark 4. *The identity* (35) *can be directly obtained from the equation* (30) *on* u^* *by applying to it the gradient operator and using* (23) *as well as the following identities*

$$\begin{cases} \boldsymbol{\nabla} A_0^{-1} A v^* = -\boldsymbol{\nabla} A_0^{-1} \operatorname{div} C \boldsymbol{\nabla} v^* = \boldsymbol{\Gamma}_0 C \boldsymbol{\nabla} v^*, \\ \boldsymbol{\nabla} A_0^{-1} b = \boldsymbol{\nabla} A_0^{-1} \operatorname{div} C \bar{\boldsymbol{\varepsilon}} = -\boldsymbol{\Gamma}_0 C \bar{\boldsymbol{\varepsilon}}, \end{cases}$$

according to which one has in particular $\boldsymbol{\nabla} S v^* = -\boldsymbol{\Gamma}_0 \delta C \boldsymbol{\nabla} v^*$ *in* $K_{\text{per}}(\Omega)$.

3 Lippmann-Schwinger spectrum and composites eigenstates

3.1 Operator properties and eigenvalue problems

In order to investigate the eigenstates associated with the preconditioned linear equation (30), or equivalently of the gradient-based Lippmann-Schwinger integral equation (35), then one focuses on the properties of the featured operators. First, it can be noticed that, considered as operators from $K_{\text{per}}(\Omega)$ into itself, then both A and A_0 are self-adjoint operators for the associated standard inner product, i.e. using Def. 1 with $C_0 = I$, with *compact inverses*, see e.g. [7]. Therefore, there exist orthonormal bases of $K_{\text{per}}(\Omega)$ composed of eigenvectors of A or A_0, respectively, which are associated with sequences of real eigenvalues that tend to infinity. Turning to the operator S in (30) one can establish the following proposition.

Proposition 2. *The operator $S = -A_0^{-1}\delta A : K_{\text{per}}(\Omega) \to K_{\text{per}}(\Omega)$ in Definition 2 is a bounded linear operator, which is self-adjoint for the inner product in Def. 1 defined by C_0.*

Proof. First, recalling the relation (32) between the operator $T = T_1$ and S, namely that $Tv^* = b_0 + Sv^*$ for all $v^* \in K_{\text{per}}(\Omega)$, then the inequality (15) gives

$$\|Sv^*\| \leq \sup_{\boldsymbol{x}\in\Omega} r_\alpha(\boldsymbol{x})\|v^*\|, \tag{36}$$

where $r_\alpha(\boldsymbol{x})$ denotes the spectral radius of the local linear operator $P_\alpha(\boldsymbol{x}) = I - C_0^{-1/2}C(\boldsymbol{x})C_0^{-1/2}$ in (31). Given that C_0 and C belongs to $L^\infty_{\text{per}}(\Omega)$, then (36) establishes that S is a bounded operator.

Second, for $v_1^*, v_2^* \in K_{\text{per}}(\Omega)$ one has

$$\begin{aligned}(-A_0^{-1}\delta A v_1^*, v_2^*) &= \langle -\boldsymbol{\nabla} A_0^{-1}\delta A v_1^* \cdot C_0\, \boldsymbol{\nabla} v_2^*\rangle \\ &= \frac{1}{|\Omega|}\int_\Omega \operatorname{div}\left(C_0 \boldsymbol{\nabla} A_0^{-1}\delta A v_1^*\right) v_2^* \,\mathrm{d}V \\ &= \frac{1}{|\Omega|}\int_\Omega -(\delta A v_1^*) v_2^* \,\mathrm{d}V = -\langle \boldsymbol{\nabla} v_1^* \cdot \delta C\, \boldsymbol{\nabla} v_2^*\rangle,\end{aligned}$$

where we used the definitions (9) and (27) of the linear operators A_0 and A, which establishes that S is self-adjoint as the last identity is symmetric. □

Before examining the implications of Property 2, let us recall some basic results of spectral theory for the convenience of the reader, who is also referred to, e.g., [24, 43] for more details. For a bounded linear operator $\mathcal{L}$ defined in a separable Hilbert space X, its *resolvent set* $\rho(\mathcal{L})$ is defined as:

$$\rho(\mathcal{L}) = \{\lambda \in \mathbb{C}\ ;\ (\mathcal{L} - \lambda I) \text{ is bijective from } X \text{ to } X\}.$$

For $\lambda \in \rho(\mathcal{L})$ then one can define the resolvent operator $R_\lambda(\mathcal{L}) = (\mathcal{L} - \lambda I)^{-1}$, which is also bounded. The *spectrum* of $\mathcal{L}$ is then the complementary set $\sigma(\mathcal{L}) = \mathbb{C}\backslash\rho(\mathcal{L})$. Its subset $\sigma_p(\mathcal{L}) \subset \sigma(\mathcal{L})$ of values λ such that $(\mathcal{L} - \lambda I)$ is not injective is the set of eigenvalues of the operator $\mathcal{L}$, or the so-called *point spectrum*. For $\lambda \in \sigma(\mathcal{L})\backslash\sigma_p(\mathcal{L})$ then $(\mathcal{L} - \lambda I)$ is injective but not surjective. In this case, if its image is dense in X, then λ belongs to the so-called *continuous spectrum* denoted by $\sigma_c(\mathcal{L})$, otherwise λ belongs to the part of the spectrum complementary to $\sigma_p(\mathcal{L}) \cup \sigma_c(\mathcal{L})$ which defines the residual spectrum.

Considering the operator S, then one refers to $\sigma(S)$ as the *Lippmann-Schwinger spectrum*. Given that S is bounded and self-adjoint according to Proposition 2, then one has the following properties:

(i) $\sigma(S) \subset \mathbb{R}$ and $\sigma(S)$ lies in a closed interval $[\lambda^-; \lambda^+]$,

(ii) $\sigma(S) = \sigma_p(S) \cup \sigma_c(S)$,

(iii) eigenvectors corresponding to different eigenvalues in $\sigma_p(S)$ are orthogonal relatively in the sense of Def. 1.

Finally, the operator S admits the following Riemann-Stieltjes integral representation

$$S = \int_{\lambda^-}^{\lambda^+} \lambda \, \mathrm{d}E_\lambda, \tag{37}$$

where $\{E_\lambda\}_{\lambda \in \mathbb{R}}$ is a unique family of orthogonal projection operators associated with S.

The representation above is general in the sense that one cannot *a priori* exclude the existence of a continuous spectrum $\sigma_c(S)$, nor reduce the integral in (37) to a discrete sum and consider that there exists an orthonormal basis of eigenvectors, some issues that have already been discussed in [46]. That acknowledged, given the overall objective of computing numerically the effective properties of a given composite material based on a discretised version of the operator S, we will concentrate solely on its discrete spectrum in the following. Note, however, that we will not be dealing with questions related to the numerical approximation of the spectrum, see for example [9] for a review of such issues and [15] in the context of operator preconditioning.

We now focus on the eigenvalues and eigenvectors of the operator S, i.e. $\lambda_j \in \mathbb{R}$ with associated function $\phi_j^* \in K_{\mathrm{per}}(\Omega)$ satisfying:

$$S\phi_j^* = \lambda_j \phi_j^*, \tag{38}$$

with the fact that λ_j is real-valued being a consequence of Proposition 2 as discussed previously. Note that the eigenvalues are possibly repeated, and the eigenvectors can be normalised in the norm associated with Def. 1. Upon applying the gradient to (38) we immediately arrive at:

$$\boldsymbol{\nabla} S\phi_j^* = -\boldsymbol{\Gamma}_0 \delta C \boldsymbol{\nabla} \phi_j^* = \lambda_j \boldsymbol{\nabla} \phi_j^*. \tag{39}$$

The last identity establishes that the operator $-\boldsymbol{\Gamma}_0 \delta C$ have eigenvectors in the space $\boldsymbol{L}^2_{\mathrm{per}}(\Omega)$ of vector-valued fields that derive from a potential field in $K_{\mathrm{per}}(\Omega)$. As a consequence, solving the eigenvalue problem for the operator S in (30) also provides a solution to (39) for the integral operator $-\boldsymbol{\Gamma}_0 \delta C$, which is featured in the Lippmann-Schwinger integral equation (35) for $\boldsymbol{\varepsilon}$.

For further analysis, (38) can be rewritten in the form of a generalised eigenvalue problem. Indeed, multiplying the former by A_0 leads to:

$$A\phi_j^* = (1 - \lambda_j) A_0 \phi_j^*, \tag{40}$$

that is an equation for the field $\phi_j^* \in K_{\mathrm{per}}(\Omega)$, which we refer to as an *eigenstate* of the composite material considered.

3.2 Eigenvalue bounds

We establish here bounds λ^- and λ^+ on the spectrum of the bounded operator S, as discussed in the previous section. In full, the equation (40) reads:

$$\operatorname{div} C(\boldsymbol{x}) \boldsymbol{\nabla} \phi_j^*(\boldsymbol{x}) = (1 - \lambda_j) \operatorname{div} C_0 \boldsymbol{\nabla} \phi_j^*(\boldsymbol{x}) \qquad \forall \boldsymbol{x} \in \Omega.$$

Therefore, the generalised eigenvalue problem (40) can be recast as this of finding $\lambda_j \in \mathbb{R}$ such that there exist non trivial solutions $\phi_j^* \in K_{\mathrm{per}}(\Omega)$ satisfying in Ω:

$$\operatorname{div} \delta C_j(\boldsymbol{x}) \boldsymbol{\nabla} \phi_j^*(\boldsymbol{x}) = 0 \qquad \text{with} \qquad \delta C_j(\boldsymbol{x}) = C(\boldsymbol{x}) + (\lambda_j - 1) C_0. \tag{41}$$

Following the same argument as in [32] for choosing the reference medium C_0 such that the series (33) and (34) converge, there exists a non-trivial solution to (41) provided that δC_j is neither uniformly positive or negative definite. Given that for all $\boldsymbol{\varepsilon} \in \mathbb{E}$ it holds

$$\delta C_j(\boldsymbol{x})\boldsymbol{\varepsilon}\cdot\boldsymbol{\varepsilon} = \Big(C_0^{-1/2}C(\boldsymbol{x})C_0^{-1/2} + (\lambda_j - 1)I\Big)\tilde{\boldsymbol{\varepsilon}}\cdot\tilde{\boldsymbol{\varepsilon}} \qquad \text{with} \qquad \tilde{\boldsymbol{\varepsilon}} = C_0^{1/2}\boldsymbol{\varepsilon},$$

then upon using the assumptions (6) and (12) as in (16) and taking the infimum or the supremum in the previous relation we have that

$$\text{if } 1-\lambda_j < \frac{m}{M_0} \text{ then } \delta C_j(\boldsymbol{x})\boldsymbol{\varepsilon}\cdot\boldsymbol{\varepsilon} > 0 \text{ and if } 1-\lambda_j > \frac{M}{m_0} \text{ then } \delta C_j(\boldsymbol{x})\boldsymbol{\varepsilon}\cdot\boldsymbol{\varepsilon} < 0 \text{ for all } \boldsymbol{x}\in\Omega.$$

Therefore, in such cases, the tensor δC_j is either positive- or negative-definite so that there is no non-trivial solution to (41). This provides bounds λ^- and λ^+ on the eigenvalues λ_j as in (17), so that the spectrum of the operator S satisfies

$$\sigma(S) \subset \left[1 - \frac{M}{m_0}\,,\, 1 - \frac{m}{M_0}\right] \tag{42}$$

where

$$\frac{m}{M_0} = \inf_{\boldsymbol{x}\in\Omega} \sigma\Big(C_0^{-1/2}C(\boldsymbol{x})C_0^{-1/2}\Big) \qquad \text{and} \qquad \frac{M}{m_0} = \sup_{\boldsymbol{x}\in\Omega} \sigma\Big(C_0^{-1/2}C(\boldsymbol{x})C_0^{-1/2}\Big)\,.$$

As already discussed in Section 2.2.2, the bounds (42) on the spectrum of S govern the convergence of the series (33) and (34), or more precisely their convergence rate is controlled by the spectral radius $\max(|1-M/m_0|, |1-m/M_0|)$.

Remark 5. *Given the inclusion $\sigma(S) \subset [\lambda^-, \lambda^+]$ in (42), the question of whether the reciprocal is true remains open for the configurations considered in this study. In some different settings, with in particular Dirichlet or Neumann boundary conditions instead of periodic conditions, this question has been investigated in [41, 15, 40] where it has been shown that the equality between the spectrum and the interval holds true under appropriate assumptions on the material parameter distributions.*

3.3 Composites eigenstates in the conductivity case

We now turn to the characterisation of the eigenstates, a question that has been investigated in [6, 4] in some specific configurations. Here, we further assume for simplicity that the material considered is piecewise-constant with N_p non-overlapping homogeneous phases Ω_p such that $\Omega = \cup_p \Omega_p$. Placing ourselves in the conductivity setting, we assume that the constituents are linear isotropic, thus defining a piecewise constant conductivity field $C(\boldsymbol{x})$, with $C(\boldsymbol{x}) = C_p > 0$ in each Ω_p, so that it holds

$$C(\boldsymbol{x}) = \sum_{p=1}^{N_p} C_p \chi_p(\boldsymbol{x}), \tag{43}$$

with χ_p being the indicator function of the phase Ω_p. Let $f_p = |\Omega_p|/|\Omega|$ denote the volume fraction of the phase p and $\langle h\rangle_p = \int_{\Omega_p} h(\boldsymbol{x})\,\mathrm{d}V/|\Omega_p|$ the associated averaging operator. Note that the construction of the composite material eigenstates presented in this section in the conductivity case can easily be extended to elasticity using a similar reasoning.

Using the notation $\delta C_j(\boldsymbol{x}) = \big(C(\boldsymbol{x}) + (\lambda_j - 1)C_0\big)$ introduced in (41), then the latter equation for the eigenstates $\phi_j^* \in K_{\mathrm{per}}(\Omega)$ can be rewritten in the present setting as the following local equations

within each phase, with suitable transmission conditions across the interfaces:

$$\begin{cases} (C_p + (\lambda_j - 1)C_0)\Delta\phi_j^*(\boldsymbol{x}) = 0 & \forall p, \forall \boldsymbol{x} \in \Omega_p \\ [\![\phi_j^*(\boldsymbol{x})]\!] = 0 & \forall \boldsymbol{x} \in \partial\Omega_p \\ [\![\delta C_j(\boldsymbol{x})\boldsymbol{\nabla}\phi_j^*(\boldsymbol{x})]\!] \cdot \boldsymbol{n}(\boldsymbol{x}) = 0 & \forall \boldsymbol{x} \in \partial\Omega_p, \end{cases} \tag{44}$$

where $[\![\cdot]\!]$ denotes jumps across the interfaces $\partial\Omega_p$ between phases, and with $\boldsymbol{n}$ being the unit normal on $\partial\Omega_p$, following the use of a suitable orientation convention on each interface. Different classes of solutions to (44) can be constructed, leading to the characterisation of various eigenstates as shown with the two types described below.

Type I. If there exists a phase $q \in \{1, \dots, N_p\}$ such that

$$\lambda_j = 1 - \frac{C_q}{C_0}, \tag{45}$$

then (44) does not constrain the behavior of the associated eigenvectors in the phase Ω_q considered, while in the other phases one must satisfy

$$\Delta\phi_j^*(\boldsymbol{x}) = 0 \qquad \forall p \neq q, \forall \boldsymbol{x} \in \Omega_p,$$

with the continuity conditions of (44). A simple way to find non trivial solutions in $K_{\text{per}}(\Omega)$ to this set of equations is to have ϕ_j^* constant in the different phases Ω_p, with $p \neq q$, so that ϕ_j^* is a trivial harmonic field with $\boldsymbol{\nabla}\phi_j^* = \boldsymbol{0}$ in Ω_p. In addition, to satisfy the continuity equations, one can also expect the gradient of ϕ_j^* to vanish in the neighboring phases near the shared interfaces. In fact, as the sought eigenvectors associated with the eigenvalue (45) are harmonic in each phase Ω_p ($p \neq q$) then, according to the maximum principle, the solutions constructed previously are the only ones possible and constitute a first type of eigenstates.

Remark 6. *Owing to the periodicity conditions, all the solutions to* (44) *satisfy* $\langle \boldsymbol{\nabla}\phi_j^* \rangle = \boldsymbol{0}$. *Therefore, as the eigenstates of the first type just described satisfy* $\boldsymbol{\nabla}\phi_j^* = \boldsymbol{0}$ *in all phases* $p \neq q$, *they also satisfy* $\langle \boldsymbol{\nabla}\phi_j^* \rangle_q = \boldsymbol{0}$.

Remark 7. *Note that if the reference medium* C_0 *is set to be equal to the parameters of one of the phases, then the corresponding eigenvalue* (45) *is* $\lambda_j = 0$, *which characterises the kernel of the operator* S *according to* (38).

Type II. If λ_j within the interval (42) is such that

$$(C_p + (\lambda_j - 1)C_0) \neq 0 \qquad \forall p \in \{1, \dots, N_p\}, \tag{46}$$

then the associated eigenvectors must satisfy

$$\Delta\phi_j^*(\boldsymbol{x}) = 0 \qquad \forall p, \forall \boldsymbol{x} \in \Omega_p,$$

with continuity conditions at the interfaces. Again, according to the maximum principle for harmonic fields, then non-trivial solutions to this set of equations achieve their maxima at the interfaces $\partial\Omega_p$ between phases. Therefore, the second type of eigenvectors, which are associated with the eigenvalues satisfying (46), are characterised by non-trivial behaviors at the interfaces.

4 Eigenstates expansion and effective properties computation

4.1 Multiphase composites

4.1.1 Expression of the solution to the cell problem

In this section, we aim at studying the relation between the eigenvalues λ_j and the effective parameters of the homogenized medium in the linear case, see also similar investigations conducted analytically in [6, 4] and numerically in [14]. To do so, we rely on the assumption that follows, which is related to the previous discussion on the integral representation (37) of the operator S.

Hypothesis 2. *Let assume that the eigenvectors of the operator S constitute a complete set $\{\phi_j^*\}$ of orthonormal functions in $K_{\rm per}(\Omega)$ for the inner product in Definition 1.*

As a consequence, any field $u^* \in K_{\rm per}(\Omega)$ can be expanded in this basis as

$$u^*(\boldsymbol{x}) = \sum_{j\geq 1} p_j\, \phi_j^*(\boldsymbol{x}) \qquad \text{with} \qquad p_j = (u^*, \phi_j^*). \tag{47}$$

Therefore, if u^* is the solution to the preconditioned equation (30), it satisfies

$$\sum_{j\geq 1} p_j (I - S)\phi_j^* = b_0,$$

which, by applying the inner product with a given element of the basis $\{\phi_j^*\}$, gives the projection components as:

$$p_j = \frac{(b_0, \phi_j^*)}{1-\lambda_j}, \tag{48}$$

where, according to (42), one has $1 - \lambda_j \geq \frac{m}{M_0} > 0$ as C is positive definite. The fact that the series (47) with (48) is convergent in $K_{\rm per}(\Omega)$ is a consequence on the requirements on the eigenvalues λ_j for S to be a contraction, i.e. $\|S\| \leq 1$, for exemple using (21), see sections 2.2.2 and 3.2.

It remains to be determined whether b_0 belongs to the range of S. Owing to the definition of the C_0-weighted inner product, we may notice the following identities:

$$\begin{aligned}(b_0, \phi_j^*) = (A_0^{-1} b, \phi_j^*) &= \langle \boldsymbol{\nabla} A_0^{-1} b \cdot C_0\, \boldsymbol{\nabla}\phi_j^* \rangle \\ &= \frac{1}{|\Omega|}\int_\Omega -\operatorname{div}\left(C_0 \boldsymbol{\nabla} A_0^{-1} b\right)\phi_j^*\, \mathrm{d}V = \frac{1}{|\Omega|}\int_\Omega b\, \phi_j^*\, \mathrm{d}V\end{aligned} \tag{49}$$

with the last equality makes use of the definition (9) of the operator A_0. In addition, given that $b = \operatorname{div} C\bar{\boldsymbol{\varepsilon}}$ then one gets

$$\frac{1}{|\Omega|}\int_\Omega b\, \phi_j^*\, \mathrm{d}V = -\frac{1}{|\Omega|}\int_\Omega \bar{\boldsymbol{\varepsilon}} \cdot C\, \boldsymbol{\nabla}\phi_j^*\, \mathrm{d}V = -\bar{\boldsymbol{\varepsilon}} \cdot \boldsymbol{\beta}_j \qquad \text{with} \qquad \boldsymbol{\beta}_j = \langle C \boldsymbol{\nabla}\phi_j^* \rangle, \tag{50}$$

so that the solution to the cell problem (30) reads

$$u^*(\boldsymbol{x}) = \sum_{j\geq 1} \left(\frac{\bar{\boldsymbol{\varepsilon}} \cdot \boldsymbol{\beta}_j}{\lambda_j - 1}\right) \phi_j^*(\boldsymbol{x}). \tag{51}$$

Furthermore, owing to the assumption (43) of a piecewise-constant material, then the terms $\boldsymbol{\beta}_j \in \mathbb{E}$ defined in (50) reduce to:

$$\boldsymbol{\beta}_j = \sum_{p=1}^{N_p} C_p f_p \langle \boldsymbol{\nabla}\phi_j^* \rangle_p.$$

As a consequence, the only coefficients p_j given by (48) that are non-zero in the expansion (51) of the solution are these that involve eigenvectors whose gradients $\boldsymbol{\nabla}\phi_j^*$ do not vanish on average in all phases nor are orthogonal to the uniform loading $\bar{\boldsymbol{\varepsilon}}$ considered. Therefore, the eigenstates of the first type described in Section 3.3 do not contribute to the solution, see Remark 6. In addition, the only *contributing* eigenvectors ϕ_j^* are those satisfying the following formal jump condition at the interfaces $\partial\Omega_p$ between phases:

$$p_j \neq 0 \quad \text{if} \quad \bar{\boldsymbol{\varepsilon}} \cdot \sum_{p=1}^{N_p} C_p \int_{\Omega_p} \boldsymbol{\nabla}\phi_j^* \,\mathrm{d}V = \bar{\boldsymbol{\varepsilon}} \cdot \sum_{\partial\Omega_p} \int_{\partial\Omega_p} [\![C(\boldsymbol{x})\phi_j^*(\boldsymbol{x})]\!] \cdot \boldsymbol{n} \,\mathrm{d}S \neq 0.$$

Hence, only such terms contribute to the expansion (51) of the solution u^* to the cell problem.

4.1.2 Relation with the effective properties

With the solution u^* at hand, one can now find the effective properties C_{eff}. Defining the latter not through the energy but through the average of the fields in (5) as:

$$C_{\text{eff}}\,\bar{\boldsymbol{\varepsilon}} = \langle C(\bar{\boldsymbol{\varepsilon}} + \boldsymbol{\nabla}u^*) \rangle, \tag{52}$$

then, owing to (43) and (51) of which we compute the gradient, we finally identify the effective tensor C_{eff} as the linear operator from $\mathbb{E}$ into itself defined as:

$$C_{\text{eff}} = \langle C \rangle + \sum_{j \geq 1} \left(\frac{\boldsymbol{\beta}_j \otimes \boldsymbol{\beta}_j}{\lambda_j - 1} \right), \tag{53}$$

with $\langle C \rangle = \sum_p C_p f_p$. In (53), the term $\boldsymbol{\beta}_j \otimes \boldsymbol{\beta}_j$ must be interpreted as the linear operator from $\mathbb{E}$ into itself such that $(\boldsymbol{\beta}_j \otimes \boldsymbol{\beta}_j)\bar{\boldsymbol{\varepsilon}} = (\boldsymbol{\beta}_j \cdot \bar{\boldsymbol{\varepsilon}})\boldsymbol{\beta}_j$ for all $\bar{\boldsymbol{\varepsilon}} \in \mathbb{E}$.

Finally, for a given $j \geq 1$, the norm of $\boldsymbol{\beta}_j$ defined in (50) can be interpreted as a *spectral density*, as it weights the contribution of the eigenvalue λ_j to the effective property expansion (53). It has a similar interpretation with respect to the solution u^* of the cell problem, via the expression (51), in which it is related to its projection components p_j onto each eigenstate ϕ_j^*. Finally, it must be reiterated that the series expansion (53) of the effective properties has been established under Hypothesis 2. The latter may not be satisfied for any microstructure, see also Remark 9.

4.2 Two-phase composites

Some known results of the theory of homogenization regarding spectral representations of the effective properties, see [5, 33, 21, 16] and [38, 46], can be recovered from (53) in the case of two-phase conductive composites by introducing some well-chosen variables and defining the comparison material suitably. Considering such a composite material, then (43) reduces to

$$C(\boldsymbol{x}) = C_1\chi_1(\boldsymbol{x}) + C_2\chi_2(\boldsymbol{x}). \tag{54}$$

Parametrising the composite geometry using the indicator function of the phase 1, i.e. defining $\chi \equiv \chi_1$, then the above reads $C(\boldsymbol{x}) = C_1\chi(\boldsymbol{x}) + C_2(1 - \chi(\boldsymbol{x}))$ as $\chi_2 = (1 - \chi_1)$. Introducing the following material parameters

$$z = \frac{C_2}{C_1} \qquad \text{and} \qquad s = \frac{z}{z-1}, \tag{55}$$

then the operator A reads

$$A = -\operatorname{div} C\boldsymbol{\nabla} = -C_2\Delta + \frac{1}{s}C_2 \operatorname{div} \chi\boldsymbol{\nabla}.$$

Given that $A_0 = -\operatorname{div} C_0 \boldsymbol{\nabla} = -C_0\Delta$, then the generalised eigenvalue problem (40) can be recast as:

$$\operatorname{div} \chi \boldsymbol{\nabla} \phi_j^* = \mu_j \Delta \phi_j^* \qquad \text{with} \qquad \mu_j = s\lambda_j, \tag{56}$$

provided that one has set $C_0 = C_2$.

Remark 8. *In the above, the labels* 1 *and* 2 *associated with the two phases of the composite are arbitrary. According to the convergence analysis of Section 2.2.2, the fixed-point scheme for solving the cell problem, which is obtained when $\alpha = 1$, is convergent if one has set the comparison conductivity C_0 to be equal to the largest conductivity value M in the composite. Hence, the choice $C_0 = C_2$ made above is compatible with a convergent numerical scheme when the phase labelled* 2 *is the most conductive phase, i.e. $C_2 = M$ and $C_1 = m$. In such a case, the parameter z in (55) does coincide with the material contrast in Sec. 2.2.2, with $z = M/m > 1$, and the convergence rate of the fixed-point scheme is at most equal to $(z-1)/z$, see Proposition 1 and Eqn. (18).*

As already noticed in Section 3.1, there is an equivalence between the generalised eigenvalue problem (40) involving the operators A and A_0, and the eigenvalue problems (38) for S and (39) for the gradient-based operator $\boldsymbol{\Gamma}_0 \delta C$. Likewise, in the two-phase configuration investigated here, the generalised eigenvalue problem (56) can be rewritten as

$$\Delta^{-1} \operatorname{div} \chi \boldsymbol{\nabla} \phi_j^* = \mu_j \phi_j^* \qquad \text{or} \qquad \boldsymbol{\Gamma} \chi \boldsymbol{\nabla} \phi_j^* = \mu_j \boldsymbol{\nabla} \phi_j^*$$

in terms of the Green's operator formally defined as $\boldsymbol{\Gamma} = \boldsymbol{\nabla} \Delta^{-1} \operatorname{div}$, showing that μ_j can be seen as the eigenvalues of the operator $\boldsymbol{\Gamma}\chi$ with eigenvectors that derive from a potential. This latter formulation is the one employed in the above cited references [5, 33, 21, 16, 46].

One of the main interests of (56) is that, in this formulation, the geometry of the composite, which is fully described by the indicator function χ, and the contrast z between phases, which intervenes through s, are fully decoupled. As a consequence, solving the above generalised eigenvalue problem, i.e. determining the generalised eigenvalues μ_j and eigenvectors ϕ_j^* of the operators $-\operatorname{div} \chi \boldsymbol{\nabla}$ and $-\Delta$, provides the solution u^* and effective property C_{eff} through (51) and (53), respectively and using (56), for any composite having the same microstructure geometry χ but different material properties C_1 and C_2. Accordingly, the effective property C_{eff} can be viewed as a function of the material parameter s defined in (55) and rewritten using the eigenvalues μ_j as:

$$C_{\text{eff}}(s) = \langle C \rangle + \sum_{j \geq 1} \frac{s}{\mu_j - s} (\boldsymbol{\beta}_j \otimes \boldsymbol{\beta}_j). \tag{57}$$

This shows that $C_{\text{eff}}(s)$ has the eigenvalues μ_j as poles, see also [34, Chap. 18]. It should also be noted that the eigenvalue bounds for λ_j established in Section 3.2 give in this setting:

$$\lambda^- = 0 \quad \text{and} \quad \lambda^+ = \frac{1}{s} \quad \text{so that} \quad \mu_j \in [0, 1]. \tag{58}$$

Likewise, according to the developments above, from the component p_j of the solution u^* one can derive a normalised projection $\bar{p}_j$ that is independent of the contrast as:

$$\bar{p}_j = \frac{s(1-\lambda_j)}{C_0} p_j = \bar{\boldsymbol{\varepsilon}} \cdot \langle \chi \boldsymbol{\nabla} \phi_j^* \rangle. \tag{59}$$

Remark 9. *As already emphasised in Section 4.1.2, a series expansion of the effective parameters such as (57) is conditioned to some specific spectral properties (see Hyp. 2). Hence, it does not apply to all two-phase composites. In particular, it is known that the checkerboard and some composites comprised of inclusions with sharp angles (e.g. square inclusions) have effective properties characterised by branch cuts in the complex s-plane, see [34, Chap. 18] and [42]. In such cases, the effective properties cannot uniformly take the form of rational functions of the properties of the constituents.*

Finally, it should be noted that it is possible to make a choice different from $\chi = \chi_2$ and $C_0 = C_1$, for example in order to be consistent with the conventional definition (21) of the reference medium. Indeed, defining

$$\chi(\boldsymbol{x}) = \chi_2(\boldsymbol{x}) - \chi_1(\boldsymbol{x}) = \begin{cases} -1 \text{ in phase 1} \\ +1 \text{ in phase 2,} \end{cases}$$

and using again that $\chi_1 + \chi_2 = 1$, then (54) can be rewritten as

$$C(\boldsymbol{x}) = \frac{C_1 + C_2}{2} + \frac{C_2 - C_1}{2}\chi(\boldsymbol{x}).$$

In this setting, if one chooses $C_0 = (C_1 + C_2)/2$ consistently with (21), then the generalised eigenvalue problem (40) reduces now to

$$\operatorname{div} \chi \boldsymbol{\nabla} \phi_j^* = \mu_j \Delta \phi_j^* \qquad \text{with} \qquad \mu_j = \left(\frac{C_1 + C_2}{C_1 - C_2}\right) \lambda_j. \tag{60}$$

This formulation, which depends only on the microstructure geometry, presents the same advantages than (56) as it allows to deduce the solution and the effective properties for any material parameters C_1 and C_2. The effective properties still has poles, with (57) to be modified according to (60), and the bounds (58) on the eigenvalues being now $\mu_j \in [-1, 1]$.

5 Discrete representation

In the following, the discrete representation of the problem (30) will be considered. In order to distinguish the operators and the fields that appear in the previous sections from their discrete matrices and vectors representations used hereafter, the latter will be denoted using the bracket notation. For example, $[v^*]$ will denote a field v^* represented as a vector, and $[A]$ will denote the matrix representation of the operator A defined in (26).

5.1 Field and operators

5.1.1 Spatial discretisation

A Fourier-based discretisation is considered with the representative volume element $\Omega \subset \mathbb{R}^3$ of size $L_1 \times L_2 \times L_3$ (or $L_1 \times L_2$ in the two-dimensional case) being discretised into a regular grid of $n_1 \times n_2 \times n_3$ voxels (or $n_1 \times n_2$ pixels in the two-dimensional case). To simplify the discrete Fourier transform operations concerning the highest attainable frequencies, the values chosen for n_1, n_2 and n_3 will be odd, see [36]. Thus, a periodic field v in Ω sampled at $N = n_1 n_2 n_3$ voxels is denoted as $v_{\mathrm{S}}(i_1, i_2, i_3)$, which is defined as

$$v_{\mathrm{S}}(i_1, i_2, i_3) = v(x_1, x_2, x_3) \qquad \text{with} \qquad x_d = (i_d - 1)\frac{L_d}{n_d} \text{ and } i_d \in [\![1, n_d]\!] \text{ for } d = 1, 2, 3 \tag{61}$$

where the double bracket notation is used here to denote an interval of integers. As a consequence, the indices $i_d = 1$ designate the origin of coordinates since $x_d(i_d = 1) = 0$.

5.1.2 Vector field representation

As a field $v^* \in K_{\mathrm{per}}(\Omega)$ is such that $\langle v^* \rangle = 0$, its spatial discretisation also satisfies

$$\sum_{i_1, i_2, i_3} v_{\mathrm{S}}^*(i_1, i_2, i_3) = 0$$

so that the total number of degrees of freedom of the latter is $N-1$. Several options are available for representing v^* as an $N-1$ dimensional vector. In this study we chose to use the vector whose elements are the components of the decomposition of the field in a sine-cosine basis as given in (82), and organised in a given order. Details on this choice and on other possible representations are given in Appendix B.

For clarity, discrete vector or matrix representations in the default sine-cosine basis will be denoted as $[\cdot]$ in the following, while $[\cdot]_{\mathrm{F}}$ will designate a representation in the Fourier space and $[\cdot]_{\mathrm{S}}$ for discrete representations sampled in the physical space.

5.1.3 Matrix operator representation

The operator A in (26) can be discretised as a square matrix $[A]$ of dimension $(N-1)\times(N-1)$. It applies on vectors in sine-cosine representation and, considering the definition of operator A, it can be formed as follows

$$[A] = -[P_{\mathrm{T}}^{\mathrm{F}}]^{-1}[\mathrm{div}]_{\mathrm{F}}[F][C]_{\mathrm{S}}[F]^{-1}[\boldsymbol{\nabla}]_{\mathrm{F}}[P_{\mathrm{T}}^{\mathrm{F}}], \tag{62}$$

where we make use of the following matrices

(i) $[P_{\mathrm{T}}^{\mathrm{F}}] \in \mathbb{C}^{(N-1)\times(N-1)}$ transforms a vector from the trigonometric sine-cosine representation into a frequency (Fourier) representation,

(ii) $[\boldsymbol{\nabla}]_{\mathrm{F}} \in \mathrm{i}\mathbb{R}^{3N\times(N-1)}$, a purely imaginary matrix that applies the gradient on a vector in frequency representation and produce a mean-free field by setting the null frequency component to zero,

(iii) $[F] \in \mathbb{C}^{3N\times 3N}$ is the three-dimensional discrete Fourier transform applied on a gradient vector,

(iv) $[C]_{\mathrm{S}} \in \mathbb{R}^{3N\times 3N}$ applies the linear constitutive relation at each element of a gradient vector, corresponding to a given voxel of its spatial discretisation,

(v) $[\mathrm{div}]_{\mathrm{F}} \in \mathrm{i}\mathbb{R}^{(N-1)\times 3N}$, purely imaginary, applies the divergence in Fourier space on a gradient vector, omitting the resulting component at the null frequency, which is equal to zero by construction.

Further details on the construction and on some properties of these matrices are given in Appendix C, according to which (62) can be rewritten using the following identities

$$[P_{\mathrm{T}}^{\mathrm{F}}]^{-1} = 2\overline{[P_{\mathrm{T}}^{\mathrm{F}}]}^{\top}, \qquad [F]^{-1} = \overline{[F]}^{\top} \qquad \text{and} \qquad [\mathrm{div}]_{\mathrm{F}} = -\overline{[\boldsymbol{\nabla}]_{\mathrm{F}}}^{\top}$$

where $\overline{[\cdot]}$ denotes the complex conjugation, as

$$[A] = 2\overline{[P_{\mathrm{T}}^{\mathrm{F}}]}^{\top}\overline{[\boldsymbol{\nabla}]_{\mathrm{F}}}^{\top}[F][C]_{\mathrm{S}}\overline{[F]}^{\top}[\boldsymbol{\nabla}]_{\mathrm{F}}[P_{\mathrm{T}}^{\mathrm{F}}].$$

In this setting, the following result can be shown, see Appendix C.5.

Proposition 3. *The matrix $[A]$ is real and symmetric.*

Similarly, an expression for the matrix representation $[A_0]$ of the operator A_0 in (9) can be obtained. In the case C_0 is a scalar, then in the sine-cosine basis, as well as in Fourier, the matrix $[A_0]$ reads

$$[A_0] = -C_0[\Delta] \tag{63}$$

where $[\Delta] \in \mathbb{R}^{(N-1)\times(N-1)}$ is the discrete Laplace operator in the sine-cosine representation, see Appendix C.5 for more details. Therefore, $[A_0]$ is real, diagonal, and also positive definite since the null frequency has been excluded from its definition. As a consequence, the matrix $[A_0]$ defines a scalar product in $\mathbb{R}^{N-1}$ as $[v_1^*]^{\top}[A_0][v_2^*]$ for $[v_1^*], [v_2^*] \in \mathbb{R}^{N-1}$.

5.2 Generalised eigenvalue problem

Using the notations of Definition 2, the eigenvalues and eigenvectors of the matrix representation $[S]$ of the operator S are investigated, with

$$[S] = -[A_0]^{-1}([A] - [A_0]).$$

From Proposition 3 and the definition (63) of $[A_0]$, which induces an inner product in $\mathbb{R}^{N-1}$ as seen above, we can establish that

$$([S][v_1^*])^\top [A_0][v_2^*] = -[v_1^*]^\top [\delta A][v_2^*] = [v_1^*]^\top [A_0]([S][v_2^*]) \qquad \forall [v_1^*], [v_2^*] \in \mathbb{R}^{N-1},$$

where $[\delta A] = [A] - [A_0]$, which establishes the following discrete counterpart to Proposition 2:

Proposition 4. *The matrix $[S]$ is real and symmetric with respect to the inner product defined by $[A_0]$.*

The discrete version of the eigenvalue problem (38) reads

$$[S][\Phi_j^*] = \Lambda_j [\Phi_j^*] \tag{64}$$

where Λ_j is the j-th eigenvalue of the matrix $[S]$, possibly repeated, and $[\Phi_j^*]$ is an associated eigenvector. Furthermore, the set $\{\Lambda_j\}$ is sorted in ascending order. Owing to Proposition 4 then $\Lambda_j \in \mathbb{R}$ for all j and, by the spectral theorem, there exists an orthonormal basis of eigenvectors $\{[\Phi_j^*]\}$ in $\mathbb{R}^{N-1}$ relatively to the inner product defined by $[A_0]$.

Remark 10. *As already discussed in Section 3.1, the questions related to the numerical approximation of the spectrum of the operator S are beyond the scope of this study and the reader is referred to, e.g., [9, 15] for some accounts on these issues. In particular, in (64) the eigenvalues Λ_j of the matrix $[S]$ cannot be individually compared to the eigenvalues λ_j for the operator S, nor are $[\Phi_j^*]$ direct discretisations of the continuous eigenstates ϕ_j^* of Section 3. Yet, it can be expected, see [9, 23], that any point in the spectrum $\sigma(S)$ of S can be approximated to arbitrary accuracy for N sufficiently large by a point in $\sigma([S])$ as $[S]$ converges, in a suitable sense, to S. Therefore, for a sufficient discretisation, we intend to interpret the numerical results to come in light of the analytical characterisations of the spectrum in the sections 3.2 and 3.3.*

As with (40) formulated in the continuous setting, the discrete equation (64) can be rewritten in the form of a generalised eigenvalue problem as:

$$[A][\Phi_j^*] = (1 - \Lambda_j)[A_0][\Phi_j^*].$$

6 Numerical results

In this section, numerical results are examined for three different microstructures, starting with a multiphase composite, then a two-component version of the latter and finally a polycrystalline model. Additional numerical examples are available at [3], where we treat the case of square inclusions.

6.1 Multiphase microstructure

The periodic microstructure of the system considered first consists of 5 discs with respective volume fractions of 0.05, 0.1, 0.15, 0.20 and 0.25, placed randomly in the domain $\Omega = [0,1) \times [0,1)$. The microstructure is displayed on Figure 1. The cell Ω is discretised as an image with $n_1 = n_2 = 315$, i.e. $N = n_1 n_2 = 99\,225$ pixels. The conductivity of the matrix (phase 1) is set at $C_1 = 1$, that of the 3 smallest discs (phase 2) at $C_2 = 3$, and that of the 2 largest discs (phase 3) at $C_3 = 10$. According to (21), the reference medium is set as $C_0 = \frac{C_1 + C_3}{2} = 5.5$.

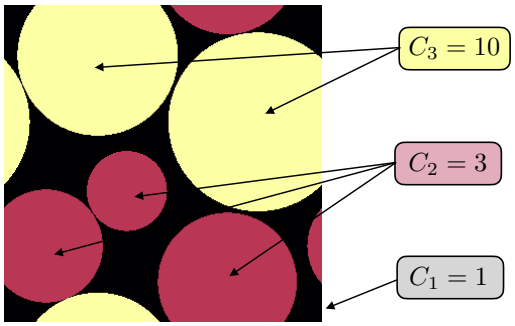


Figure 1: Microstructure comprising 5 discs of different sizes placed randomly. Image discretised into $N = 315^2$ pixels.

6.2 Spectral decomposition

6.2.1 Implementation

The matrix $[A]$ is constructed by successively computing its response to unitary vectors, whose components are all null except one, using the sequence of operations enumerated in Section 5.1.3. The matrix $[A_0]$ is obtained by directly applying the relation (63).

The code that computes the generalised eigensystems of $[A]$ and $[A_0]$ was implemented in Julia, in which the function `eigen` of the `LinearAlgebra` package is used. This function is able to optimize the choice of the algorithm used, according to the properties of the matrices given as argument. In our case, with $[A]$ real and symmetric, $[A_0]$ diagonal, real and positive, it uses the Lapack function `DSYEVR` which is an implementation of the `dqds` algorithm, see [2].

6.2.2 Eigenvalues

The spectrum of the discrete operator $[S]$ is computed for the configuration considered and plotted on Figure 2, with the eigenvalues Λ_j being sorted by increasing values. Three plateaus can be observed on the graphic for eigenvalues approximately equal to -0.818, 0.4545 and 0.818, respectively, which

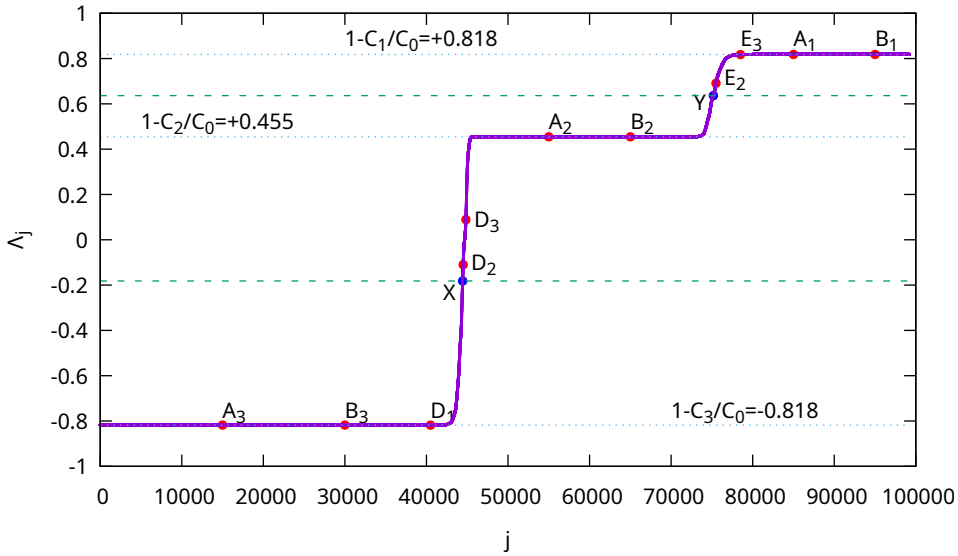


Figure 2: Eigenvalues Λ_j for the microstructure in Fig. 1. The indices j in abscissa vary from 1 to $N-1 = 99224$, the eigenvalues being sorted in ascending order.

correspond to the eigenvalues (45) of the type I and are consistent with the bounds $\pm\frac{C_3-C_1}{C_3+C_1} \approx \pm 0.818$ on the spectral radius of $[S]$, as already highlighted in [32] and further discussed in Section 3.2.

Moreover, the interpolated curve passing through the points crosses the midlines between the plateaus at abscisses that appear to be correlated to the volume fractions of the phases. These intersection are represented by the points X and Y on Figure 2; their respective abscisses, divided by the total number of eigenvalues are $44438/99224 = 0.448$ and $75178/99224 = 0.758$ to be compared to the volume fraction of phase 3 which is 0.45 and the joint volume fraction of the phases 2 and 3 equal to 0.75. Indeed, from the analytical construction of the eigenvalues of type I in Section 3.3, one can expect the size of each plateau to be correlated with the number of degrees of freedom available to describe the associated eigenvectors, i.e. here to the number of pixels in each phase and therefore its volume (or surface) fraction. In this reasoning, the number of points in the transitions between plateaus should be subtracted. Note that the ratio between these quantities depends on the discretisation, as will be seen in Section 6.3.2.

Here, we focus on the spectrum of the discrete operator $[S]$ (there are $N - 1 = 99224$ eigenvalues for the discretisation considered, which have been plotted as dots on the graphic). Noting that this configuration has nearly touching inclusions, and as discussed in the remarks 5, 9 and 10, these numerical results do not allow any conclusions to be drawn about the nature of the spectrum of the operator S. Some complementary material can be found in [3] concerning the study of a composite with square inclusions, using the same discretisation. For such a microstructure with sharp angles, the numerical spectrum shows qualitative differences compared to that of the case studied here, particularly in the transition between plateaus, which reveals possible differences in the properties of the spectra of the operators associated with these two microstructures.

6.2.3 Eigenstates

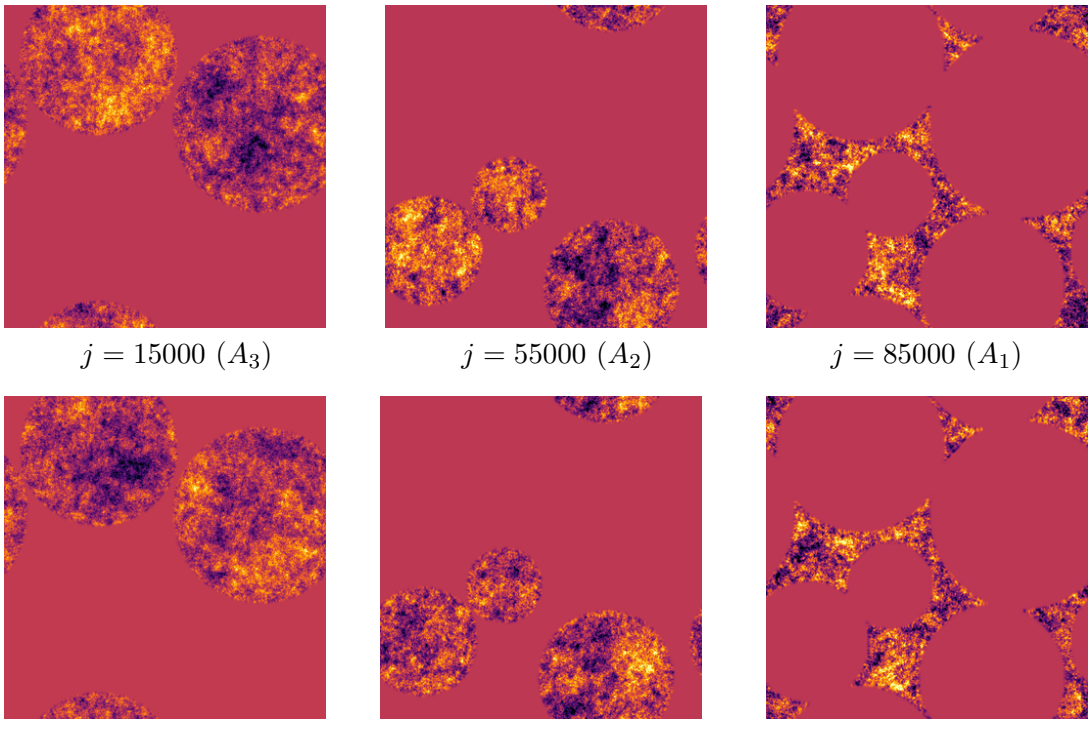
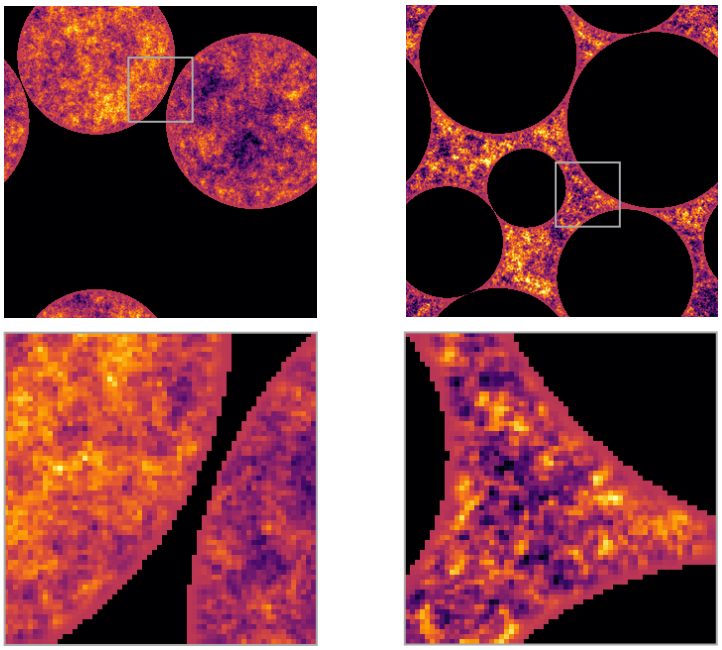


Figure 3: Images of eigenvectors $[\Phi_j^*]$ of type I for the microstructure in Fig. 1, discretised into $N = 315^2$ pixels. The references in parenthesis are those of the corresponding labels on the spectrum in Fig. 2.

The eigenvectors $[\Phi_j^*]$ of the discrete eigenvalue problem (64) are computed and their behaviour in the physical space is now discussed. A selection of eigenvector images are displayed on Figure 3. The two eigenvectors displayed in the left column (eigenvalue indices $j = 15000$ and 30000) are both associated with the eigenvalue $1 - \frac{C_3}{C_0} \approx -0.818$. They must be compared to the solution (45) of the eigenproblem for the operator S, and the associated eigenstates ϕ_j^* of the type I described in Section 3.3, which are allowed to fluctuate in phase 3 while being constant in the matrix (phase 1) and the phase 2. The two eigenvectors displayed in the middle column refer to the eigenvalue equal to $1 - \frac{C_2}{C_0} \approx 0.4545$ and correspond to eigenstates that fluctuate in phase 2 and are constant everywhere else. The two eigenvectors on the right column correspond to the eigenvalue $1 - \frac{C_1}{C_0} \approx +0.818$ and to fields fluctuating in the matrix (phase 1) and constant in all the inclusions (phases 2 and 3).

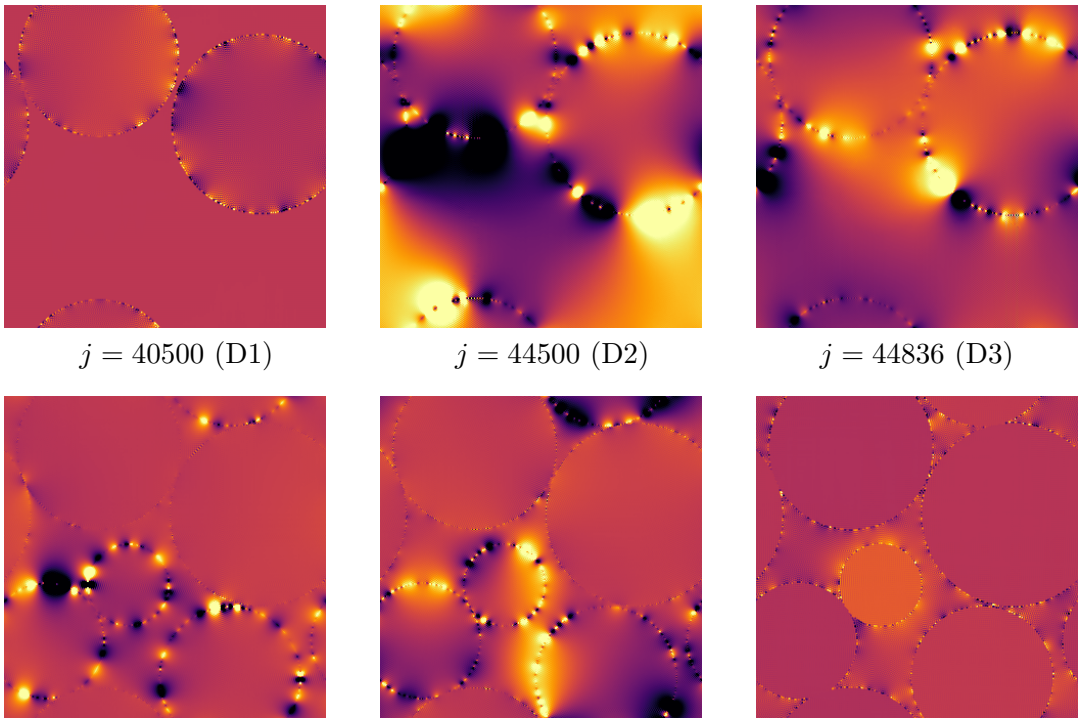
Figure 4 displays close-up images of $[\Phi_j^*]$ for $j = 15000$ where $[\Phi_j^*]$ fluctuates in phase 1 and is constant in the matrix and in the phase 2, and $j = 85000$ where $[\Phi_j^*]$ fluctuates in the matrix and is constant in all the inclusions. The regions where the eigenstates are uniform are masked in black. A thin layer can be observed close to the boundaries between phases, which ensures the continuity conditions for the eigenstates, see (44), from the area where they fluctuate to the other phases where they are uniform, a behaviour that is consistent with the analysis in Section 3.3 for this first type of eigenstates.



$j = 15000$ (A3) $j = 85000$ (A1)

Figure 4: Eigenvectors $[\Phi_j^*]$ of type I for the microstructure in Fig. 1. Left: The pixels in the matrix and in the phase 2 are masked in black. Right: The pixels in the phases 2 and 3 are masked in black. Bottom: close-ups. A boundary layer can be seen at the interfaces between phases where the eigenstates fluctuate (phase 3 for $j = 15000$, matrix for $j = 85000$). This constitutes a buffer zone that allows the continuity conditions to be met.

The images in Figure 5 are associated with eigenvalues of type II, which satisfy (46). They lie between the first and second plateaus at $1 - \frac{C_3}{C_0}$ and $1 - \frac{C_2}{C_0}$ (top row) and between the second and third plateaus at $1 - \frac{C_2}{C_0}$ and $1 - \frac{C_1}{C_0}$ (bottom row). The computed eigenvectors $[\Phi_j^*]$ exhibit localised fluctuations at the boundaries between phases and in the nearby regions, which is again consistent with the analysis of the eigenstates of type II in Section 3.3.



$j = 40500$ (D1) $j = 44500$ (D2) $j = 44836$ (D3)

$j = 75178$ (Y) $j = 75500$ (E2) $j = 78500$ (E3)

Figure 5: Images of eigenvectors $[\Phi_j^*]$ of type II for the microstructure in Fig. 1, discretised into $N = 315^2$ pixels. Between parenthesis: corresponding labels in the spectrum, see Fig. 2.

6.3 Investigation of the solution to the cell problem

6.3.1 Discrete solution and projection onto the eigenvector basis

We now investigate the discrete counterpart of the solution u^* to the preconditioned equation (29), a problem discretised as

$$[A_0]^{-1}[A][u^*] = [A_0]^{-1}[b] \qquad \text{with} \qquad [b] = [P_{\mathrm{T}}^{\mathrm{F}}]^{-1}[\mathrm{div}]_{\mathrm{F}}[F][C]_{\mathrm{S}}[\bar{\boldsymbol{\varepsilon}}], \tag{65}$$

where, with a slight abuse of notation, the vector $[u^*]$ therefore denotes the discrete solution defined by the equation above. Thus, $[u^*]$ does not coincide with the continuous solution u^* directly discretised. Analogously to the eigenstates expansion (47) of the solution in the continuous case, the discrete solution $[u^*]$ can also be expanded onto the orthogonal basis $\{[\Phi_j^*]\}$. Defining the matrix $[\Phi^*] \in \mathbb{R}^{(N-1)\times(N-1)}$ whose columns are the generalised eigenvectors $[\Phi_j^*]$ of $[A]$ and $[A_0]$, i.e. the non-singular matrix defined by blocks as

$$[\Phi^*] = \begin{bmatrix} [\Phi_1^*] & \dots & [\Phi_{N-1}^*] \end{bmatrix}$$

then one can express the discrete solution as

$$[u^*] = [\Phi^*][p] \qquad \text{with} \qquad [p] = [\Phi^*]^\top [A_0][u^*] \in \mathbb{R}^{N-1}, \tag{66}$$

the components of the vector $[p]$ being therefore the projection values of $[u^*]$ onto the vectors $[\Phi_j^*]$, similarly to (47).

Now, defining the diagonal matrix $[\Lambda] \in \mathbb{R}^{(N-1)\times(N-1)}$ as

$$[\Lambda] = \mathrm{diag}(\Lambda_j),$$

and then, by definition of the generalised eigenvalues, one has

$$[A][\Phi^*] = [A_0][\Phi^*]([I] - [\Lambda]),$$

where $[I] \in \mathbb{R}^{(N-1)\times(N-1)}$ denotes the identity matrix. In addition, as it holds

$$[\Phi^*]^\top [A_0][\Phi^*] = [I],$$

since the eigenvectors $\{[\Phi_j^*]\}$ constitute an orthonormal basis relatively to the inner product defined by $[A_0]$, see Section 5.2, then it can be deduced from (65) that $[p]$ defined in (66) reads

$$[p] = ([I] - [\Lambda])^{-1}[\Phi^*]^\top [b], \tag{67}$$

an identity which can be compared to the expression (48) with (49), obtained in the continuous setting.

Therefore, the solution $[u^*]$ to the discrete problem (65) reads

$$[u^*] = [\Phi^*]([I] - [\Lambda])^{-1}[\Phi^*]^\top [b], \tag{68}$$

from which it can be deduced that:

$$[A]^{-1} = [\Phi^*]([I] - [\Lambda])^{-1}[\Phi^*]^\top.$$

6.3.2 Numerical results

The vector $[p]$ in (67) of the projection components onto the eigenvector basis of the discrete solution for the microstructure in Fig. 1 and the loading $\bar{\varepsilon} = (1, 0)$ is plotted on Figure 6, with the associated spectrum being superimposed. As already observed in [14], it clearly appears that a significant portion

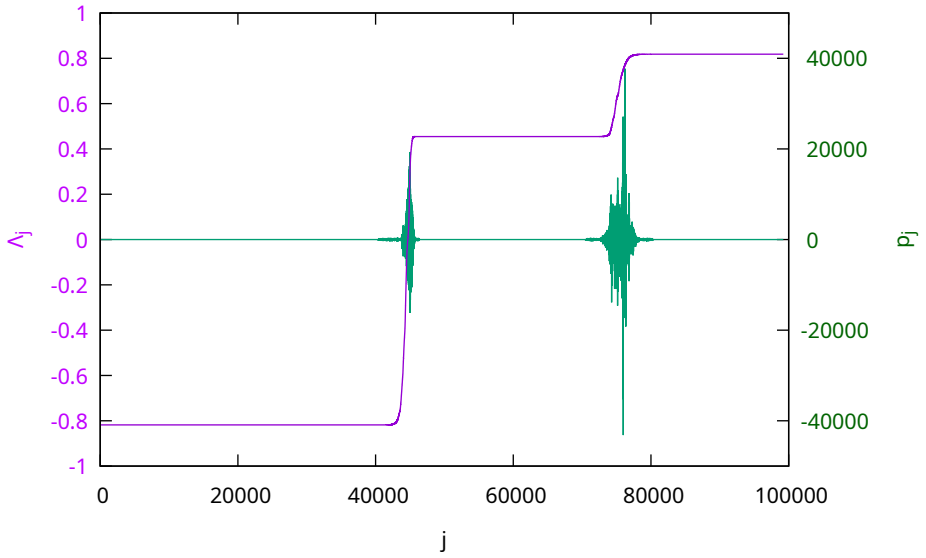


Figure 6: Projection of the solution on the generalised eigenvectors (in green) for the microstructure in Fig. 1, discretised in 315^2 pixels, with the associated spectrum (in magenta).

of the projections has values of zero or close to zero. Consistently with the analysis in Section 4.1.1, this corresponds to the eigenvectors of type I that do not contribute to the solution as the non-zero values of the projection concentrate in the two transition parts between the plateaus, whose interpretation has been discussed previously in Section 6.2.2. Approximately, these transitions extend

between eigenvalue with indices comprised between 41500 and 46200, and between 71000 and 79500. These *useful* eigenvalues and eigenstates therefore only represent approximately 13% of the whole spectrum.

Figure 7 shows the eigenstates giving rise to the largest projection values. They all belong to the second transition part, with eigenvalue indices around 76000. The eigenvectors with highest projection values in the first transition part, with indices around 45000, are depicted in Fig. 8.

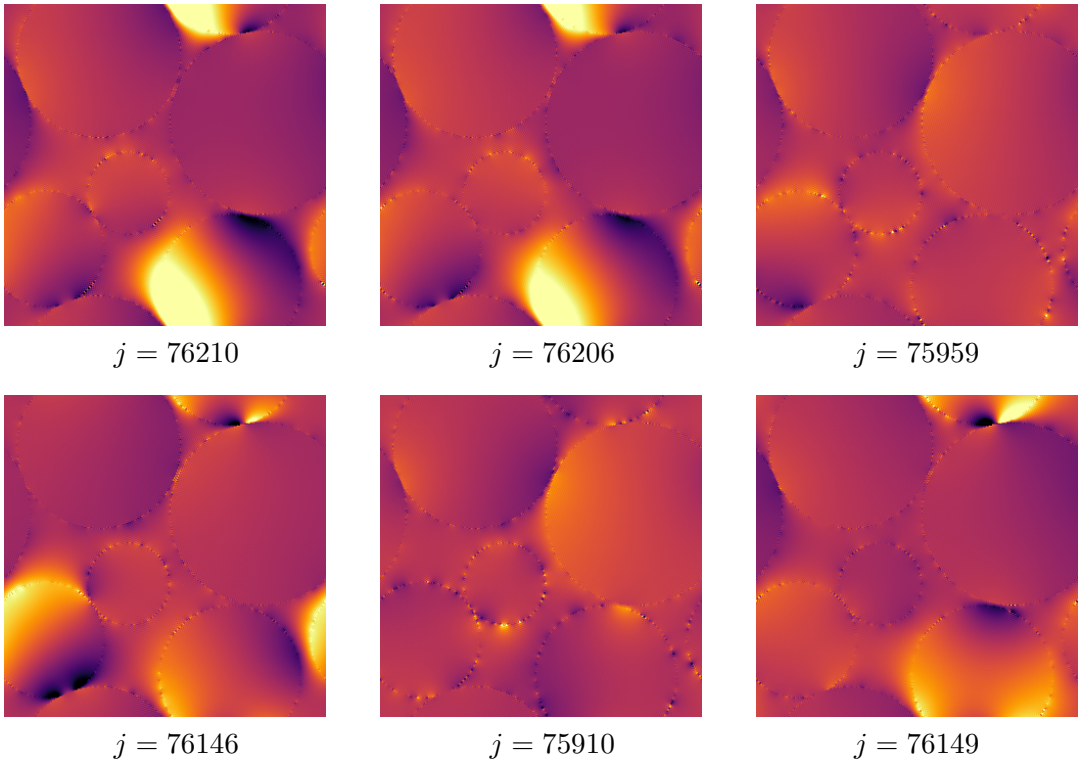


Figure 7: Images of the overall six most contributive eigenstates $[\Phi_j^*]$ for the microstructure in Fig. 1. They are all found in the second transition zone of Fig. 6 and are thus of type II.

Moreover, the extent of these two zones depends largely on the spatial discretisation of the problem. Indeed, as can be seen on Figure 9, the relative width of this zone becomes smaller as the discretisation parameter N increases. For comparison, the eigenvalue indices have been normalised in this figure to fit within the interval $[0, 1]$. To measure this trend, the widths of the different parts of the spectrum, i.e. the three plateaus and the two transitions, are measured for different discretisation N, see figures 6 and 9, by counting the eigenvalues within 10^{-4} of the theoretical values of the plateaus and the eigenvalues beyond that threshold. They are compared with the measurement of a surface with a maximum thickness of two pixels around the interface between phases, as already discussed in Fig. 4. The results are shown in Figure 10, with the blue and green curves representing the number of eigenvalues in the plateaus and transitions, respectively, as functions of the discretisation and measured on the spectra. The red curve represents the total number of pixels in the overall 4-pixel-wide boundary layer between phases. It can be seen that the size of the plateaus (in blue) grows at a rate proportional to N, which is much faster than that of the transitions (in green). The latter correlates with the surface area of the boundary layer between phases, with a growing rate slightly larger than $\sqrt{N}$. Note that, to be fully established, this relation would require measurements with larger discretisations, which exceeds our current computing resources. To conclude, the number of *useful* eigenstates seems to be approximately in a ratio length-to-surface (or surface-to-volume in 3D).

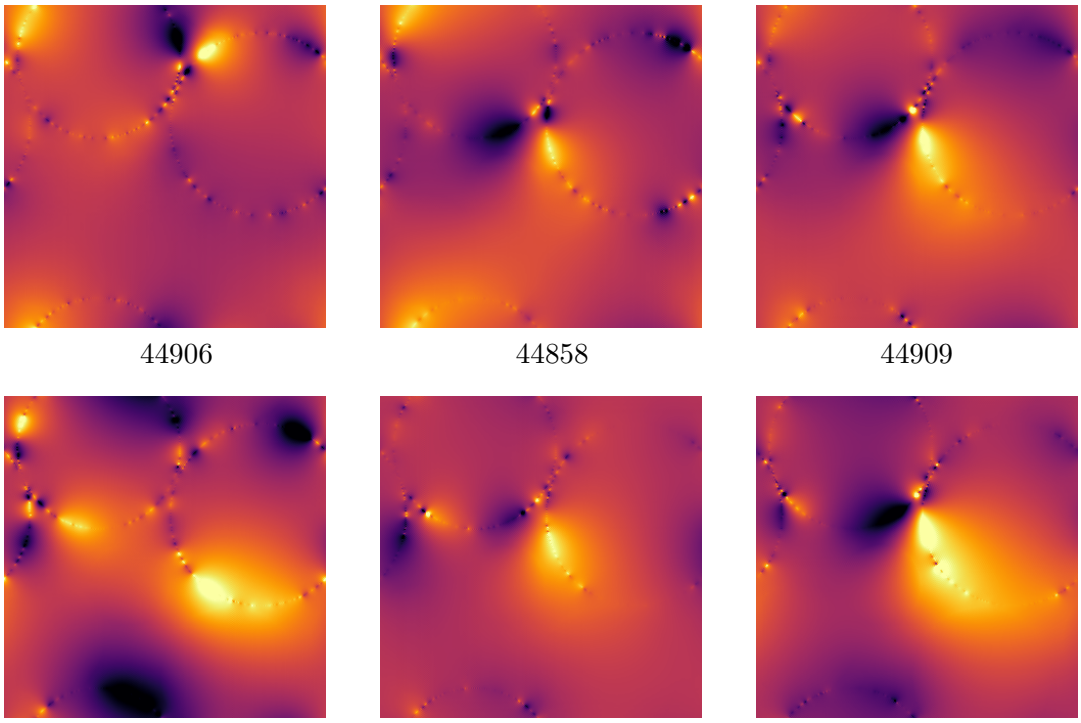


Figure 8: Images of the six most contributive eigenvectors $[\Phi_j^*]$ for the microstructure in Fig. 1 for eigenvalues found in the first transition zone in Fig. 6 around the index 45000. These are Type II eigenvectors.

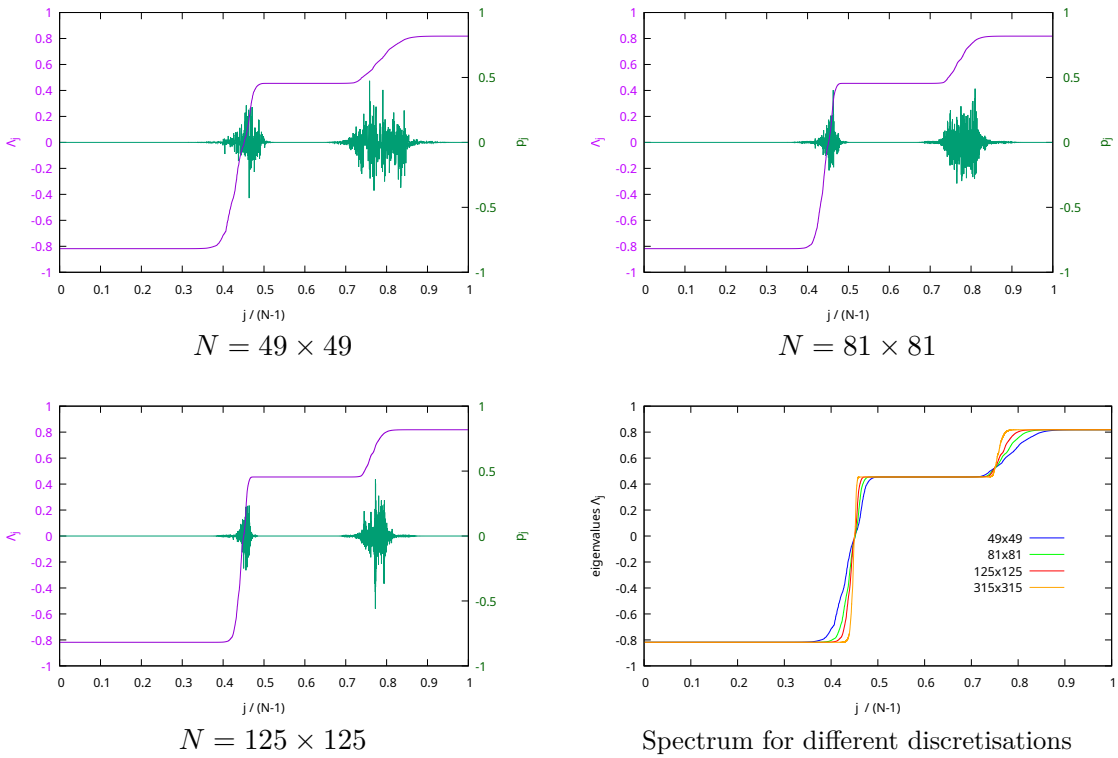


Figure 9: Projection components of the solution and spectra for the microstructure in Fig. 1 for different discretisations N. The eigenvalues indices have been normalised to fit the interval $[0, 1]$.

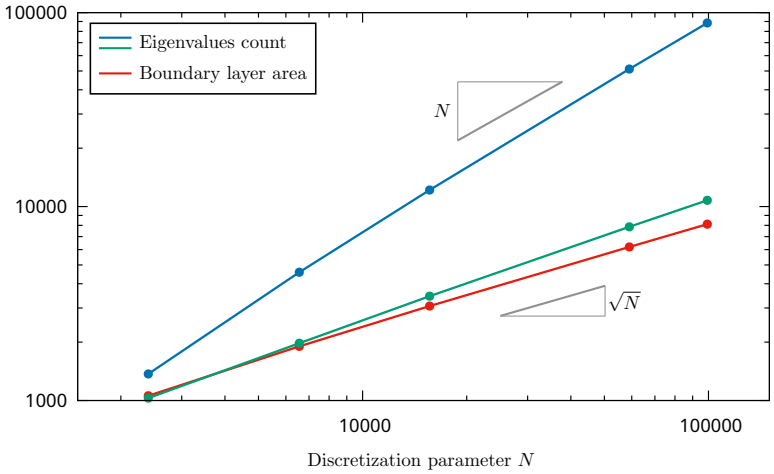


Figure 10: Size of the different parts in the spectra as functions of the discretisation: blue for the three plateaus, green for the transitions in-between. Comparison with the area of the boundary layer, in red, estimated from the image of the microstructure.

6.3.3 Reduced model

Using the relation (68), the exact solution $[u^*]$ of the discrete problem can be obtained. It is also possible to compute the field $[u^*]_{\text{tr}}$ obtained by keeping only the eigenvalues giving rise to the largest projection values. To do so, one can first compute the projection vector $[p]$ using (67), keep only a given number N_{tr} of components with the most significant contributions, the other components being set to zero, and finally, apply the matrix $[\Phi^*]$ on the resulting truncated version of $[p]$. Figure 11 displays the field $[u^*]_{\text{tr}}$ obtained with different numbers of eigenvalues kept. When only a few eigenvalues are retained for the reconstruction of the solution, the resulting field differs qualitatively (image top left) from that obtained with all eigenvalues (bottom right image), but beyond a truncation number around $N_{\text{tr}} = 2500$, the obtained fields are virtually identical.

This strategy can be extended to the computation of the effective conductivity tensor C_{eff}, see (52). Here, we compare the error of a truncated approximation of C_{eff} relatively to its exact discrete version, which is computed using all eigenvalues for $N = 315^2$. For the microstructure considered, the effective medium is anisotropic and the components of C_{eff} are computed by combining the loadings $\bar{\boldsymbol{\varepsilon}}_1 = (1, 0)$ and $\bar{\boldsymbol{\varepsilon}}_2 = (0, 1)$. Quantitative relative error measurements are reported in Figure 12 as functions of the number N_{tr} of predominant eigenvalues used to approximate the effective conductivity. Note that, although C_{eff} is symmetric, the errors on its out-of-diagonal components $(C_{\text{eff}})_{12}$ and $(C_{\text{eff}})_{21}$ as functions of N_{tr} are different as the predominant modes are not the same, hence the associated errors, for the two loadings considered. However, the symmetry does hold for the untruncated discrete solutions. Once again we observe that a reduced number of eigenvalues is sufficient to build a high precision solution: an accuracy of 10^{-4} is reached with less than 10 000 eigenvalues, which represents approximatively only 10% of the total number of eigenvalues of the system for the discretisation considered. Furthermore, if the selected eigenvectors are associated with eigenvalues smaller in absolute value than the spectral radius in Proposition 1, then the convergence of the iterative method will be accelerated. These results advocate for the use of reduced models for the computation of the effective properties of composite materials.

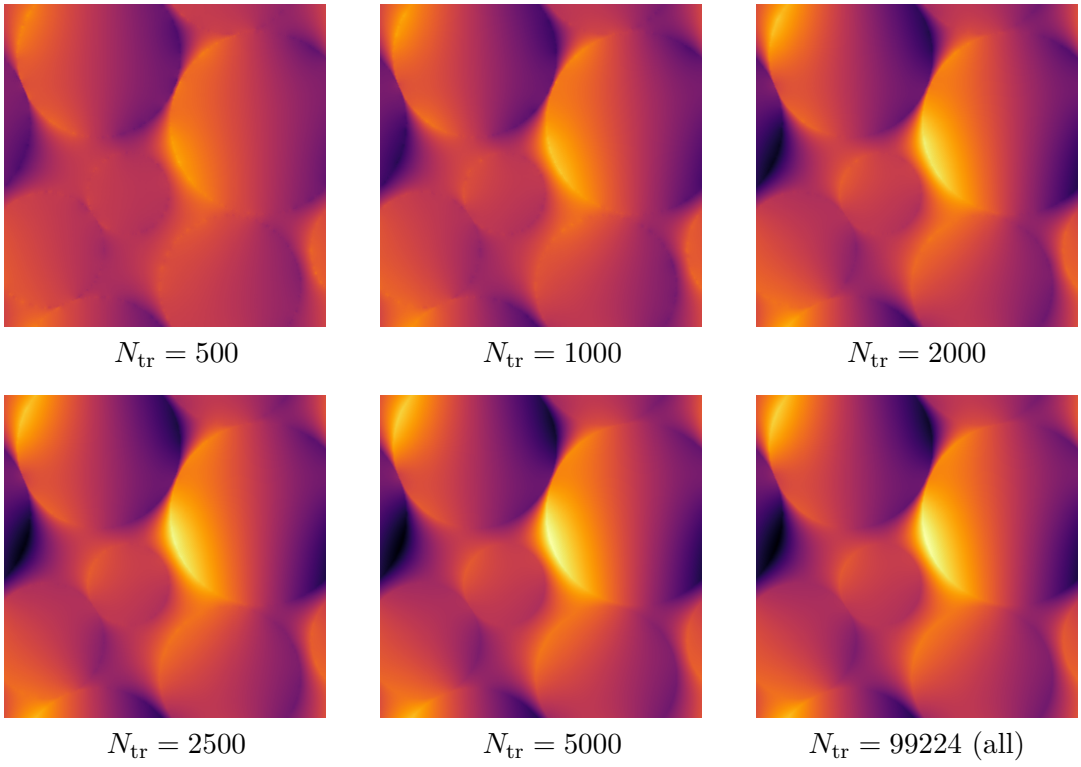


Figure 11: Truncated solution $[u^*]_{\mathrm{tr}}$ constructed using a reduced number N_{tr} of eigenvalues.

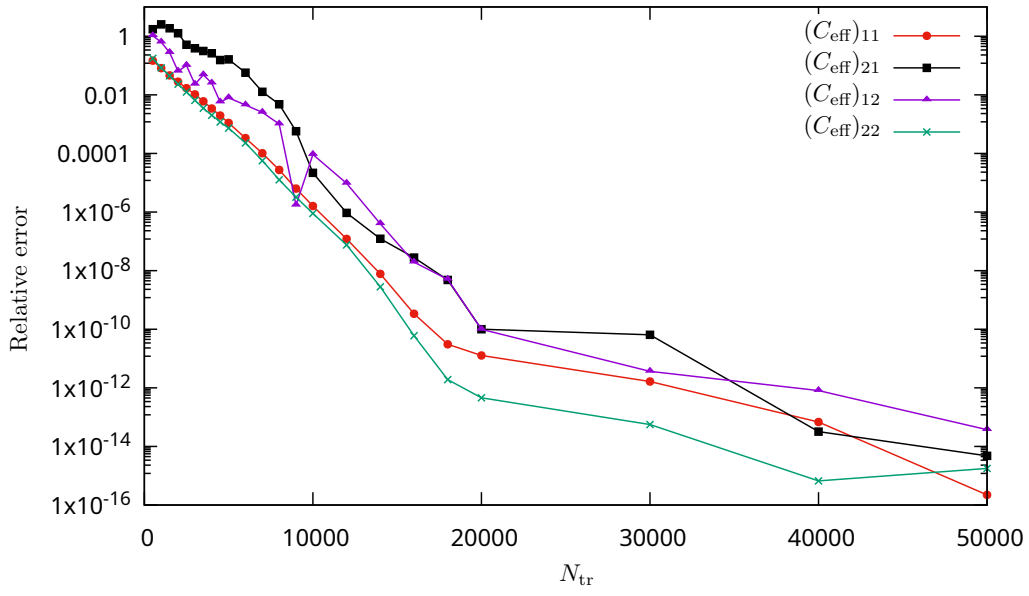


Figure 12: Relative error on the components of the effective conductivity tensor C_{eff} obtained when considering a reduced number N_{tr} of eigenvalues.

6.4 Two-phase configuration

To illustrate numerically the developments of Section 4.2 relative to two-phase conductive composites, the microstructure in Figure 1 investigated previously is now redefined by setting $C_2 = C_3$. Upon defining the indicator function as $\chi \equiv \chi_1$, combined with the choice $C_0 = C_2$ of the reference material, then the discrete version of the generalised eigenvalue problem (56) reads

$$[H][\Phi_j^*] = s\Lambda_j\,[\Delta][\Phi_j^*] \qquad \text{with} \qquad s = \left(\frac{C_2}{C_2 - C_1}\right), \tag{69}$$

with the matrix $[H]$ being expressed in terms of the sampled version $[\chi]_{\rm S}$ of the indicator function of the inclusions as

$$[H] = [P_{\rm T}^{\rm F}]^{-1}[{\rm div}]_{\rm F}[F][\chi]_{\rm S}[F]^{-1}[\boldsymbol{\nabla}]_{\rm F}[P_{\rm T}^{\rm F}].$$

As already highlighted in Section 4.2, the interest of (69) lies in dissociating the geometry of the microstructure from the properties of its constituents. Hence, the value of C_2 need not be specified here, and the following results are valid for the entire range of the ratio $z = C_2/C_1$.

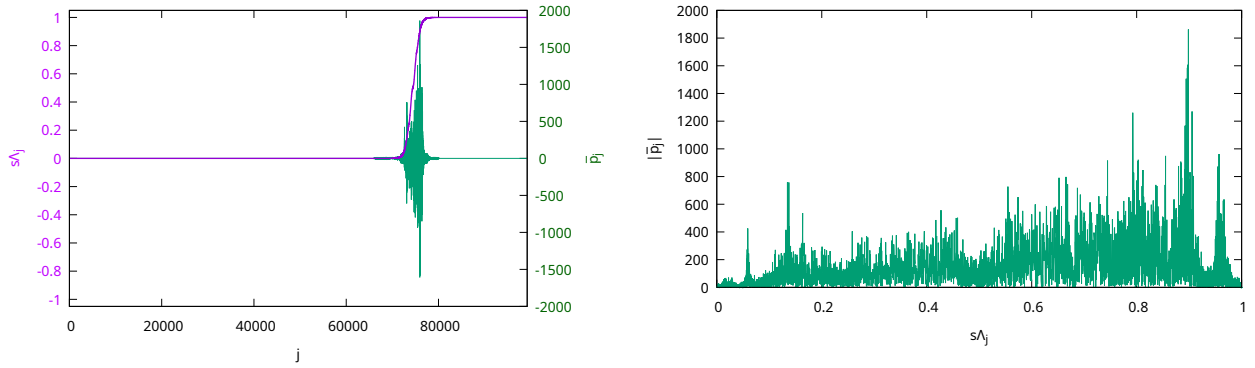


Figure 13: Investigation of the two-phase microstructure derived from Fig. 1. Left: modified eigenvalues $s\Lambda_j$ (69) (in magenta) with the associated normalised projections $\bar{p}_j$ of the solution (in green) on the generalised eigenvectors, see (59). Right: norm of the normalised projection components $\bar{p}_j$ as functions of the modified eigenvalues $s\Lambda_j$.

The spectrum associated with the system (69) is investigated in Figure 13 for the microstructure derived from Fig. 1 when $C_2 = C_3$. To represent only quantities that are independent of the contrast, the left panel shows the normalised spectrum $s\Lambda_j$ together with the associated normalised projections $\bar{p}_j$ of the solution onto the generalised eigenvectors, see (67) and (59), for the loading $\bar{\boldsymbol{\varepsilon}} = (1, 0)$. The results are consistent with the previous analysis, with two plateaus corresponding to eigenvectors of Type I, which do not contribute to the solution, and intermediate eigenvalues associated with contributing eigenvectors of Type II. To better highlight these contributions to the solution, the right panel in Fig. 13 represent the norm of the normalised projection component $\bar{p}_j$ as function of the modified eigenvalues $s\Lambda_j$, see (69). In these figure, the plateaus collapse on the points 0 and 1. Such a figure can be directly compared with representations of the so-called spectral density in, e.g., [38].

As investigated in Section 6.3.3 for the three-phase version of the same microstructure geometry, a reduced model can be constructed using a truncated list of eigenvectors $[\Phi_j^*]$ to provide an accurate approximation of the solution with an accelerated convergence rate. The acceleration provided by using only the first $N_{\rm tr}$ eigenvectors is now quantified. The solution $[u_n^*]$ at the iterate n, with projection components $\{p_j^n\}$ defined as in (66), and the solution $[u^*]$, with projections $\{p_j\}$, obtained at convergence of the iterative scheme are compared. As the eigenvectors $\{[\Phi_j^*]\}$ constitute an orthonormal basis relatively to the inner product defined by the matrix $[A_0]$, the total residual between the

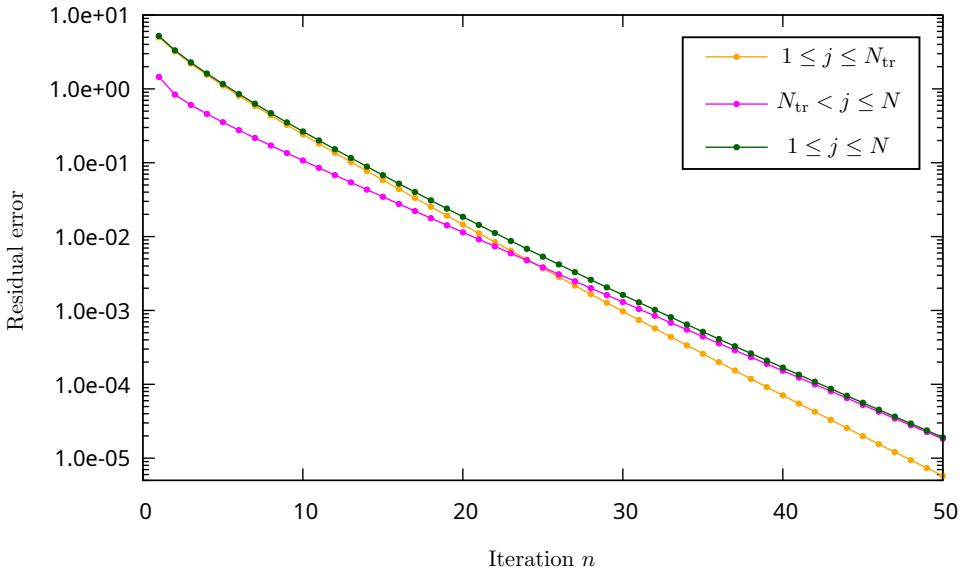


Figure 14: Investigation of the two-phase microstructure derived from Fig. 1. Residual errors on the solution obtained at convergence when considering a subset of eigenvalues with $N_{\rm tr} = 1000$ and as a function of the iteration number, see (70).

two solutions can be decomposed as:

$$\left\| [u_n^*] - [u^*] \right\|_{A_0}^2 = \sum_{j=1}^{N_{\rm tr}} (p_j^n - p_j)^2 \; + \sum_{j=N_{\rm tr}+1}^{N} (p_j^n - p_j)^2. \tag{70}$$

The residual errors for $N_{\rm tr} = 1000$ are plotted in Figure 14. The green curve for $1 \leq j \leq N$ corresponds to the total residual at the left-hand side of (70), while the right-hand side terms are associated with the curves labelled $1 \leq j \leq N_{tr}$ and $N_{tr} < j \leq N$, respectively. This graph highlights that the evolution of the residual exhibits two regimes: a first one governed by the first $N_{\rm tr}$ eigenvectors selected, and a second one with a slower convergence and associated with the remaining $(N - N_{\rm tr})$ eigenvectors. Such a result supports the use of a reduced basis to speed up computations.

6.5 Example of a polycrystal

In a last example, we investigate a microstructure that models a polycrystal as a Voronoi tessellation with 10 random cells, see Figure 15, whose conductivities have been set as $C_j = j$ for $j = 1, \ldots, 10$. The reference medium is defined as $C_0 = (C_1 + C_{10})/2 = 5.5$. Its associated spectrum and the projection components (67) on the eigenvector basis of the discrete solution of (65) for the loading $\bar{\boldsymbol{\varepsilon}} = (1, 0)$ are plotted on Figure 16. Consistently with the previous analyses, the spectrum now exhibits ten plateaus, whose values correspond to the type I eigenvalues, i.e. $1 - \frac{C_j}{C_0}$ for $j = 1, \ldots, 10$. Moreover the projection components reach their highest values in the transition region in-between the plateaus. The eigenvectors $[\Phi_j^*]$ of type I depicted in Fig. 17 are fields that fluctuate in only one given cell, while being uniform in the other ones. The eigenvectors of type II, which yield the highest projection values of the solution, are characterised by not trivial behaviours at the boundaries between cells. Different eigenvectors among the most contributive to the solution $[u^*]$ are displayed on Fig. 18.

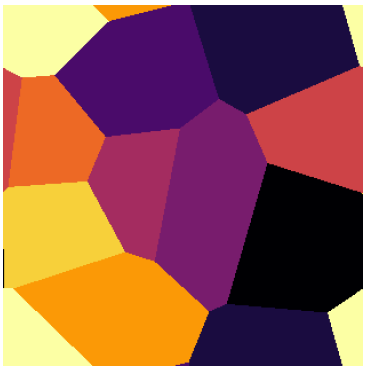


Figure 15: Polycrystal represented as a Voronoi tessellation with 10 cells with conductivities $C_j = j$. Image discretised into $N = 315^2$ pixels.

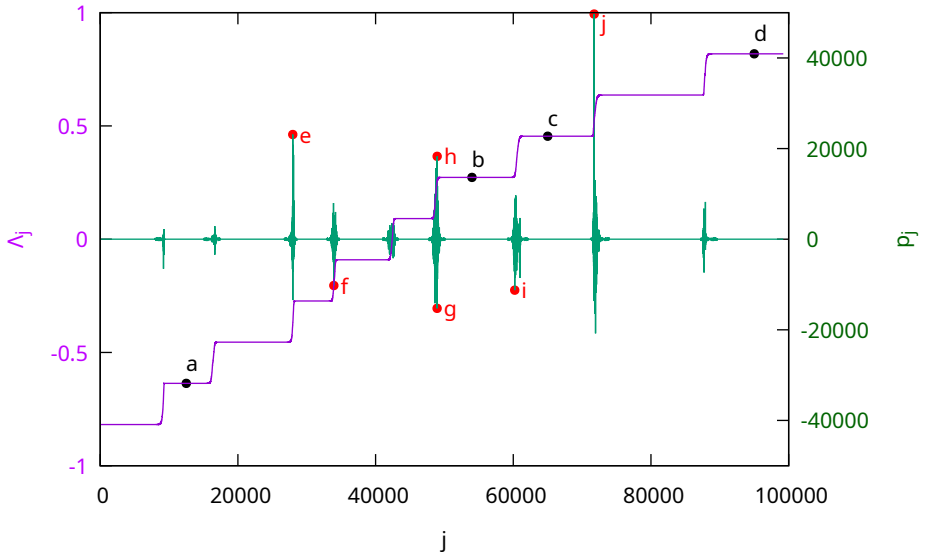
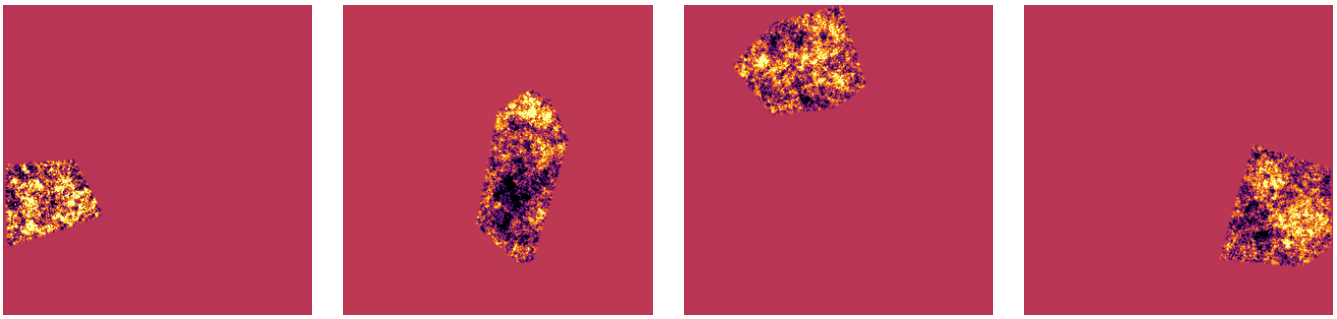
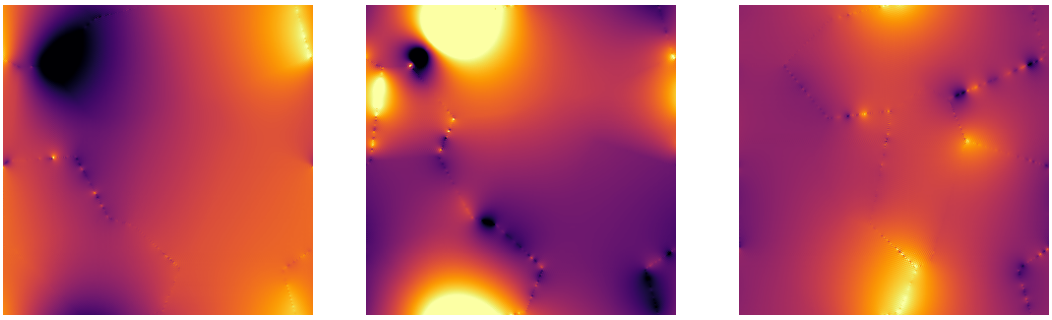


Figure 16: Eigenvalues Λ_j for the microstructure in Fig. 15 (magenta). The indices j in abscissa vary from 1 to $N-1 = 99224$, the eigenvalues being sorted in ascending order. Projection of the solution on the generalised eigenvectors (green).

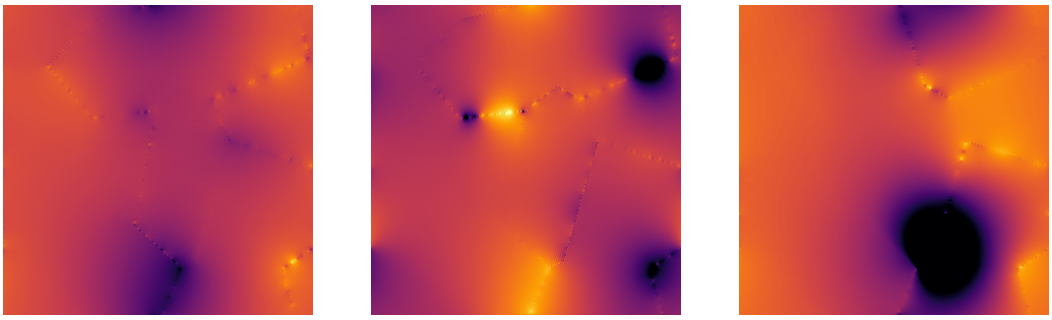


$j = 12500$ (a) $j = 54000$ (b) $j = 65000$ (c) $j = 95000$ (d)

Figure 17: Images of eigenvectors $[\Phi_j^*]$ of type I for the microstructure in Fig. 15, discretised in $N = 315^2$ pixels. Between parenthesis: associated labels on Fig. 16.



$j = 27968$ (e) $j = 33936$ (f) $j = 48935$ (g)



$j = 48945$ (h) $j = 60211$ (i) $j = 71750$ (j)

Figure 18: Images of eigenvectors $[\Phi_j^*]$ of type II for the microstructure in Fig. 15, discretised in $N = 315^2$ pixels. Between parenthesis: corresponding labels on Fig. 16.

7 Conclusion

This study has focused on volume integral formulations of homogenization problems, which provide the foundations for both establishing theoretical results, such as bounds on effective properties in some specific cases, and developing efficient numerical methods for complex microstructures. Connecting these theoretical and numerical approaches requires understanding the effect of introducing a reference comparison material, which amounts in defining a preconditioner for the original cell problem. This results in an integral operator, whose properties govern both the performances of iterative solution methods and the construction of the homogenization solution and thus the effective properties of the composite considered.

The main points of the present study are the following ones:

1. The comparison material is introduced in a variational formulation of the homogenization problem so as to define an inner product, which allows to identify the gradient of the given energy functional. The optimality condition for the gradient thus defined is equivalent to a preconditioned version of the cell problem by a second-order differential operator involving the reference medium.

2. Iterative solution methods such as steepest descent and fixed-point iterations are convergent provided that the comparison materials is adjusted relatively to the spectral bounds of the tangent modulus. If the energy density is not twice differentiable, this condition can be reinterpreted under weaker assumptions.

3. Formulated using the potential (or displacement) field unknown, these developments can be rewritten in an equivalent gradient (or strain) based setting. For linear composites, the preconditioned integral equation reduces to the well-known Lippmann-Schwinger equation. The latter is shown to be expressed in terms of a bounded and self-adjoint linear operator.

4. The spectral properties of the integral operator are analysed. While the existence of a continuous spectrum might not be excluded without further assumptions on the composite constituents, the focus is made on the eigenvalues and the associated eigenstates. Eigenvalues bounds are provided and, in the conductivity setting, eigenstates are constructed and classified into two main types: fields that fluctuate in a given phase while being uniform in the others, and fields that exhibit non trivial behaviours at the interfaces between phases.

5. It is assumed that the solution to the cell problem can be developed on a basis of eigenvectors, an expansion in which the eigenstates of the first type play no role. The effective properties can in turn be expanded as rational fractions of the eigenvalues. This allows to recover some known theoretical results in the case of two-component composites, for a well-chosen set of variables.

6. The linear integral operator is discretised in a setting compatible with conventional FFT-based computational methods. The discrete eigenvalue problem, recast as a generalised eigenvalue problem, is solved. The questions related to the numerical approximation of the continuous operator are beyond the scope of the presented analysis.

7. Some numerical results are presented for three different microstructures, including a two-component composite and a polycrystal model. These results corroborate well the theoretical analysis conducted. In particular, they illustrate on some examples how a very limited portion of the spectrum is actually used to compute the effective properties.

8. Approximate computations of the solution and effective properties are performed using a truncated version of the spectrum. In addition, the typical behaviour of convergence errors in iterative methods is correlated with the eigenvectors involved in the computations.

This study opens the door to future developments for optimizing the choice of the comparison material for non-linear or anisotropic linear composite materials. Reduced models can also be constructed based on the spectral decomposition of the Lippmann-Schwinger integral operator, which can serve to improve the performances of computational methods. In addition, some theoretical developments remain needed, in particular to provide a more accurate description of the useful eigenstates, as well as to provide a clear picture in the case of multiphase materials or in the elasticity case.

A Proof of contraction mapping

Owing to the definition of the Gateaux derivative DT_α of T_α in (11), it holds:

$$T_\alpha(v_2^*) - T_\alpha(v_1^*) = \int_0^1 DT_\alpha[v_1^* + t(v_2^* - v_1^*)](v_2^* - v_1^*)\,\mathrm{d}t, \tag{71}$$

for all $v_1^*, v_2^* \in K_{\mathrm{per}}(\Omega)$, where from (11), one has

$$DT_\alpha[v^*]\tilde{v}^* = \tilde{v}^* - \alpha R_0 D^2W[v^*]\tilde{v}^*. \tag{72}$$

In the identity above, the term $D^2W[v^*]\tilde{v}^*$ belongs to $K_{\mathrm{per}}(\Omega)'$ and reads as:

$$D^2W[v^*]\tilde{v}^* = -\operatorname{div}\left(\frac{\partial^2 w}{\partial \boldsymbol{\varepsilon}^2}(\boldsymbol{x}, \bar{\boldsymbol{\varepsilon}} + \boldsymbol{\nabla} v^*)\boldsymbol{\nabla}\tilde{v}^*\right) \qquad \forall v^*, \tilde{v}^* \in K_{\mathrm{per}}(\Omega), \tag{73}$$

so that its action against a given $\check{v}^* \in K_{\mathrm{per}}(\Omega)$, which we denote as $D^2W[v^*](\tilde{v}^*, \check{v}^*)$, satisfies

$$D^2W[v^*](\tilde{v}^*, \check{v}^*) = \frac{1}{|\Omega|}\int_\Omega \frac{\partial^2 w}{\partial \boldsymbol{\varepsilon}^2}(\boldsymbol{x}, \bar{\boldsymbol{\varepsilon}} + \boldsymbol{\nabla} v^*)\boldsymbol{\nabla}\tilde{v}^* \cdot \boldsymbol{\nabla}\check{v}^*\,\mathrm{d}V.$$

In this context, owing to (71) and (72), then in the energetic norm $\|\cdot\|$ associated with Definition 1 one has:

$$\begin{aligned}\|T_\alpha(v_2^*) - T_\alpha(v_1^*)\| &\leq \int_0^1 \|DT_\alpha[v_1^* + t(v_2^* - v_1^*)](v_2^* - v_1^*)\| \,\mathrm{d}t \\ &\leq \int_0^1 \|DT_\alpha[v_1^* + t(v_2^* - v_1^*)]\| \cdot \|v_2^* - v_1^*\| \,\mathrm{d}t\end{aligned} \tag{74}$$

where the operator norm of the linear operator $DT_\alpha[v^*]$ is defined for all $v^* \in K_{\mathrm{per}}(\Omega)$ as:

$$\|DT_\alpha[v^*]\| = \sup_{v^* \neq 0} \frac{|(DT_\alpha[v^*]\tilde{v}^*, \tilde{v}^*)|}{\|\tilde{v}^*\|^2}. \tag{75}$$

Using (72) we get

$$(DT_\alpha[v^*]\tilde{v}^*, \tilde{v}^*) = (\tilde{v}^* - \alpha R_0 D^2 W[v^*]\tilde{v}^*, \tilde{v}^*) = \langle R_0^{-1}\tilde{v}^* - \alpha D^2 W[v^*]\tilde{v}^*, \tilde{v}^*\rangle \tag{76}$$

where the second equality comes from the definition of the Riesz mapping. Using (73) in the above identity and integrating by parts yields

$$\begin{aligned}\langle R_0^{-1}\tilde{v}^* - \alpha D^2 W[v^*]\tilde{v}^*, \tilde{v}^*\rangle &= \frac{1}{|\Omega|}\int_\Omega \left\{C_0 - \alpha \frac{\partial^2 w}{\partial \boldsymbol{\varepsilon}^2}(\boldsymbol{x}, \bar{\boldsymbol{\varepsilon}} + \boldsymbol{\nabla} v^*)\right\} \boldsymbol{\nabla}\tilde{v}^* \cdot \boldsymbol{\nabla}\tilde{v}^* \,\mathrm{d}V, \\ &= \frac{1}{|\Omega|}\int_\Omega P_\alpha(\boldsymbol{x}, \bar{\boldsymbol{\varepsilon}} + \boldsymbol{\nabla} v^*) C_0 \boldsymbol{\nabla}\tilde{v}^* \cdot \boldsymbol{\nabla}\tilde{v}^* \,\mathrm{d}V.\end{aligned} \tag{77}$$

where one has used the symmetric operator P_α from $\mathbb{E}$ into itself defined in (13). Therefore, taking the supremum in (77) and using (76), one gets

$$|(DT_\alpha[v^*]\tilde{v}^*, \tilde{v}^*)| \leq \sup_{\boldsymbol{x}\in\Omega} r_\alpha(\boldsymbol{x})\|\tilde{v}^*\|^2 \tag{78}$$

where $r_\alpha(\boldsymbol{x})$ is the local spectral radius of P_α in Proposition 1. The inequality (78) provides an upper bound for the operator norm $\|DT_\alpha[v^*]\|$ in (75). Inserting it back into (74) then finally yields the sought inequality (15), i.e.

$$\|T_\alpha(v_2^*) - T_\alpha(v_1^*)\| \leq \sup_{\boldsymbol{x}\in\Omega} r_\alpha(\boldsymbol{x})\|v_2^* - v_1^*\|.$$

B Representations of fields

In this section, details will be given on how the fluctuating part $v^* \in K_{\mathrm{per}}(\Omega)$ of a scalar potential v is discretised and stored as a vector $[v^*]$. Several choices are possible for the representation of v^* as a vector. Here are considered sampled representations in the physical space, denoted as $[v^*]_{\mathrm{S}}$, as well as representations in the Fourier or sine-cosine bases denoted as $[v^*]_{\mathrm{F}}$ or $[v^*]$, respectively.

B.1 Physical space representation

The most obvious way would be to store in a N-dimensional vector $[v]_{\mathrm{S}}$ the values of all the voxels in a given order. For example, the i-th element of vector $[v]_{\mathrm{S}}$ could be

$$\begin{aligned}&[v]_{\mathrm{S}}(i) = v_{\mathrm{S}}(i_1, i_2, i_3) \\ &\text{with } (i_1, i_2, i_3) \in [\![1, n_1]\!] \times [\![1, n_2]\!] \times [\![1, n_3]\!], \\ &\text{and } i = 1 + (i_1 - 1) + (i_2 - 1)n_1 + (i_3 - 1)n_1 n_2,\end{aligned} \tag{79}$$

where v_{S} has been defined in (61). However, as, by definition, a field v^* is of null average, its vector representation $[v^*]_{\mathrm{S}}$ has a dimension $N-1$. This can be achieved by using a representation similar to (79) while excluding one given voxel, for example the voxel of indices $(i_1,i_2,i_3)=(1,1,1)$, to *globally* ensure that $\sum_{i_1,i_2,i_3} v^*_{\mathrm{S}}(i_1,i_2,i_3)=0$, i.e.

$$\begin{aligned}&[v^*]_{\mathrm{S}}(i)=v^*_{\mathrm{S}}(i_1,i_2,i_3)\\&\text{with } (i_1,i_2,i_3)\in[\![1,n_1]\!]\times[\![1,n_2]\!]\times[\![1,n_3]\!]-(1,1,1)\\&\text{and } i=(i_1-1)+(i_2-1)n_1+(i_3-1)n_1n_2,\end{aligned}$$

and assuming that

$$v^*_{\mathrm{S}}(1,1,1)=-\sum_{i=1}^{N-1}[v^*]_{\mathrm{S}}(i)\ .$$

Although the vectors and matrices associated with this representation are real-valued, the choice is rather arbitrary because of the special role given to a voxel in the discretisation of v^*. Finally, as it will be seen next, an alternative representation leading to simple matrix forms for the operators A and A_0 will be adopted.

B.2 Fourier space representation

Considering the decomposition in Fourier series of v:

$$v(x_1,x_2,x_3)=\sum_{(k_1,k_2,k_3)\in\mathcal{K}} c_{k_1,k_2,k_3}e^{+2\pi\mathrm{i}\left(k_1\frac{x_1}{L_1}+k_2\frac{x_2}{L_2}+k_3\frac{x_3}{L_3}\right)}$$

where $\mathcal{K}=[\![-m_1,m_1]\!]\times[\![-m_2,m_2]\!]\times[\![-m_3,m_3]\!]$ with $m_d=\frac{n_d-1}{2}$ for $d=1,2$ or 3, given that n_d is chosen as an odd integer. The vector of N elements $[v]_{\mathrm{F}}$ could be defined, for example, as

$$[v]_{\mathrm{F}}(k)=c_{k_1,k_2,k_3} \tag{80}$$

where $k\in[\![1,N]\!]$, where $k_d\in[\![-m_d,+m_d]\!]$ (for $d=1,2,3$), and where the indices k and (k_1,k_2,k_3) are linked by a one-to-one relationship that can be written as

$$\begin{aligned}&k=k(k_1,k_2,k_3)\\&k_d=k_d(k),\ \text{for } d=1,2,3\ .\end{aligned} \tag{81}$$

The vector of $N-1$ elements $[v^*]_{\mathrm{F}}$, representing the field v^* in the Fourier basis, can be defined by omiting the null frequency, which is generally stored as the first element of $[v]_{\mathrm{F}}$. Accordingly, one has

$$[v^*]_{\mathrm{F}}(k)=[v]_{\mathrm{F}}(k+1),\quad \text{for } k\in[\![1,N-1]\!].$$

This representation of vector $[v^*]_{\mathrm{F}}$ is more convenient than the previous one, as all elements of the vector play an equivalent role and as $[A_0]_{\mathrm{F}}$ is diagonal. Yet, it obliges to form matrix $[A]_{\mathrm{F}}$ as a matrix with complex values, which makes the calculations more cumbersome than the representation described below.

B.3 Sine-cosine decomposition

Finally, we chose a representation of the vector $[v^*]$ using a decomposition of the field v in a trigonometric sine and cosine basis as

$$\begin{aligned}v(x_1,x_2,x_3)=a_{0,0,0}+\sum_{(k_1,k_2,k_3)\in\mathcal{K}^+}&\Big\{a_{k_1,k_2,k_3}\cos\Big(2\pi\Big(k_1\frac{x_1}{L_1}+k_2\frac{x_2}{L_2}+k_3\frac{x_3}{L_3}\Big)\Big)\\&+b_{k_1,k_2,k_3}\sin\Big(2\pi\Big(k_1\frac{x_1}{L_1}+k_2\frac{x_2}{L_2}+k_3\frac{x_3}{L_3}\Big)\Big)\Big\}\end{aligned} \tag{82}$$

where $\mathcal{K}^+$ is defined as the subset of indices $(k_1, k_2, k_3) \in \mathcal{K}$ such that

$$\begin{cases} k_1 > 0 & \text{or} \\ k_1 = 0 \text{ and } k_2 > 0 & \text{or} \\ k_1 = 0 \text{ and } k_2 = 0 \text{ and } k_3 > 0 \ . \end{cases}$$

The $N-1$ terms a_{k_1,k_2,k_3} and b_{k_1,k_2,k_3} for $(k_1, k_2, k_3) \in \mathcal{K}^+$ are stored with a given order as the components of $[v^*]$. It can easily be checked that the frequencies of $\mathcal{K}^+$ and all the frequencies (k_1, k_2, k_3) such that $(-k_1, -k_2, -k_3) \in \mathcal{K}^+$ gather together all the non zero frequencies in $\mathcal{K}$. In the case where the field v is real, the elements a_{k_1,k_2,k_3} and b_{k_1,k_2,k_3} are real.

C Matrix operator representation

This section details the properties of the matrices involved in the construction of $[A]$ and $[A_0]$ introduced in Section 5.1.3

C.1 Matrix $[P_{\mathbf{T}}^{\mathbf{F}}]$: transformation from sine-cosine to Fourier representation

It can easily be shown that the terms c_{k_1,k_2,k_3} of the Fourier decomposition of a scalar field v and the terms of its decomposition in cosine and sine a_{k_1,k_2,k_3} and b_{k_1,k_2,k_3} are linked together by the following relations, for all $(k_1, k_2, k_3) \in \mathcal{K}^+$:

$$\begin{cases} a_{0,0,0} = c_{0,0,0} \\ a_{k_1,k_2,k_3} = (c_{k_1,k_2,k_3} + c_{-k_1,-k_2,-k_3}) \\ b_{k_1,k_2,k_3} = \mathrm{i}(c_{k_1,k_2,k_3} - c_{-k_1,-k_2,-k_3}) \end{cases} \quad \text{and} \quad \begin{cases} c_{0,0,0} = a_{0,0,0} \\ c_{k_1,k_2,k_3} = \dfrac{1}{2}(a_{k_1,k_2,k_3} - \mathrm{i}\, b_{k_1,k_2,k_3}) \\ c_{-k_1,-k_2,-k_3} = \dfrac{1}{2}(a_{k_1,k_2,k_3} + \mathrm{i}\, b_{k_1,k_2,k_3}) \end{cases}$$

When $[v^*]$ is organised such that the elements a_{k_1,k_2,k_3} are placed first, with a given order for $(k_1, k_2, k_3) \in \mathcal{K}^+$, and that the elements b_{k_1,k_2,k_3} are placed next, and when $[v^*]_{\mathrm{F}}$ is organised such that the terms c_{k_1,k_2,k_3} for $(k_1, k_2, k_3) \in \mathcal{K}^+$ ordered in the same manner as in $[v^*]$ are placed first, and that the terms $c_{-k_1,-k_2,-k_3}$ are placed next, the operator $[P_{\mathrm{T}}^{\mathrm{F}}]$ defined as

$$[v^*]_{\mathrm{F}} = [P_{\mathrm{T}}^{\mathrm{F}}][v^*]$$

is a complex matrix that reads by blocks as:

$$[P_{\mathrm{T}}^{\mathrm{F}}] = \frac{1}{2}\begin{bmatrix} [I]_{M\times M} & -\mathrm{i}[I]_{M\times M} \\ [I]_{M\times M} & +\mathrm{i}[I]_{M\times M} \end{bmatrix}$$

where $M = (N-1)/2$ and $[I]_{M\times M}$ is the $M \times M$ identity matrix. Therefore, its inverse $[P_{\mathrm{T}}^{\mathrm{F}}]^{-1}$ is

$$[P_{\mathrm{T}}^{\mathrm{F}}]^{-1} = \begin{bmatrix} [I]_{M\times M} & [I]_{M\times M} \\ +\mathrm{i}[I]_{M\times M} & -\mathrm{i}[I]_{M\times M} \end{bmatrix}$$

so that

$$[P_{\mathrm{T}}^{\mathrm{F}}]^{-1} = 2\overline{[P_{\mathrm{T}}^{\mathrm{F}}]}^{\top}. \tag{83}$$

Note that the factor 2 in (83) can be removed by normalisation but it did not appear necessary to do so here.

Remark 11. *In the case the frequencies (k_1, k_2, k_3) are not organised in the same manner in $[v^*]$ and in $[v^*]_F$, the identity (83) remains valid as matrix $[P_T^F]$ can be written*

$$[P_T^F] = \frac{1}{2}\begin{bmatrix} [I]_{M\times M} & -\mathrm{i}[I]_{M\times M} \\ [I]_{M\times M} & +\mathrm{i}[I]_{M\times M} \end{bmatrix} [P]$$

where $[P]$ is an adequate permutation matrix, thus satisfying $[P]^{-1} = [P]^\top$.

C.2 Discrete Fourier transform matrix

In the one-dimensional case, the discrete Fourier transform matrix $[F_{1D}]$ that applies to a vector of n elements is the complex $n \times n$ matrix that reads

$$[F_{1D}] = \frac{1}{\sqrt{n}}\begin{bmatrix} 1 & 1 & 1 & 1 & \dots & 1 \\ 1 & \omega & \omega^2 & \omega^3 & \dots & \omega^{n-1} \\ 1 & \omega^2 & \omega^4 & \omega^6 & \dots & \omega^{2(n-1)} \\ 1 & \omega^3 & \omega^6 & \omega^9 & \dots & \omega^{3(n-1)} \\ \vdots & \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & \omega^{n-1} & \omega^{2(n-1)} & \omega^{3(n-1)} & \dots & \omega^{(n-1)(n-1)} \end{bmatrix}$$

where $\omega = e^{2\pi\mathrm{i}/n}$, see for example [17]. In the case of a 3-dimensional field v, let $[F_{3D}]$ be the matrix carrying out the discrete Fourier transform of $[v]_S$, which represents a field v sampled in the physical space as a vector of N elements arranged as in (79), and resulting in a vector $[v]_F$ with N elements arranged in the Fourier space as in (80), i.e. such that

$$[v]_F = [F_{3D}][v]_S.$$

It is a $N \times N$ complex matrix whose elements are

$$[F_{3D}](k, n) = \frac{1}{\sqrt{N}}\, \omega_1^{n_1 k_1}\omega_2^{n_2 k_2}\omega_3^{n_3 k_3},$$

with $\omega_d = e^{-2\pi\mathrm{i}/n_d}$ for $d = 1, 2, 3$, and where there are one-to-one relations between k and (k_1, k_2, k_3), and between n and (n_1, n_2, n_3), respectively. The elements of the matrix of the inverse discrete Fourier transform read

$$[F_{3D}]^{-1}(n, k) = \frac{1}{\sqrt{N}}\, \omega_1^{-n_1 k_1}\omega_2^{-n_2 k_2}\omega_3^{-n_3 k_3},$$

so that, one has

$$[F_{3D}]^{-1} = \overline{[F_{3D}]}^\top. \tag{84}$$

The 3D discrete Fourier transform is extended to vector fields by introducing the matrix $[F] \in \mathbb{C}^{3N\times 3N}$ defined by blocks as

$$[F] = \begin{bmatrix} [F_{3D}] & [0] & [0] \\ [0] & [F_{3D}] & [0] \\ [0] & [0] & [F_{3D}] \end{bmatrix}$$

which, according to (84), satisfies $[F]^{-1} = \overline{[F]}^\top$.

C.3 Gradient and divergence matrices

Considering the discretisation in the Fourier space of a given field v as a vector $[v]_{\mathrm{F}}$ of N elements, and the discretisation of the gradient of v in the Fourier space as a vector of $3N$ elements organised as

$$[\boldsymbol{\nabla} v]_{\mathrm{F}} = \begin{bmatrix} [\frac{\partial v}{\partial x_1}]_{\mathrm{F}} \\ [\frac{\partial v}{\partial x_2}]_{\mathrm{F}} \\ [\frac{\partial v}{\partial x_3}]_{\mathrm{F}} \end{bmatrix}. \tag{85}$$

The Fourier-based gradient can be represented by the $3N \times N$ matrix $[\boldsymbol{\nabla}]_{\mathrm{F}}$ such that

$$[\boldsymbol{\nabla} v]_{\mathrm{F}} = [\boldsymbol{\nabla}]_{\mathrm{F}} [v]_{\mathrm{F}}$$

with

$$[\boldsymbol{\nabla}]_{\mathrm{F}} = \begin{bmatrix} [\boldsymbol{\nabla}_1]_{\mathrm{F}} \\ [\boldsymbol{\nabla}_2]_{\mathrm{F}} \\ [\boldsymbol{\nabla}_3]_{\mathrm{F}} \end{bmatrix}$$

where the $N \times N$ block matrices $[\boldsymbol{\nabla}_j]_{\mathrm{F}}$ are given by

$$[\boldsymbol{\nabla}_j]_{\mathrm{F}} = \frac{2\pi \mathrm{i}}{L_j} \operatorname{diag}\Big(k_j(k(1)), \ldots, k_j(k(N))\Big) \tag{86}$$

using the notation of (81) and using the well-known property of the Fourier transform applied on derivatives. The divergence of a vector field $\boldsymbol{a}$ in $\mathbb{R}^3$ is expressed in the Fourier space as

$$\widehat{\operatorname{div}(\boldsymbol{a})} = 2\pi \mathrm{i} \left(\frac{k_1}{L_1} \hat{a}_1 + \frac{k_2}{L_2} \hat{a}_2 + \frac{k_3}{L_3} \hat{a}_3 \right).$$

Using the matrices $[\boldsymbol{\nabla}_j]_{\mathrm{F}}$, the divergence can be represented as the $N \times 3N$ purely imaginary matrix $[\mathrm{div}]_{\mathrm{F}}$ defined as

$$[\mathrm{div}]_{\mathrm{F}} = \begin{bmatrix} [\boldsymbol{\nabla}_1]_{\mathrm{F}}^{\top} & [\boldsymbol{\nabla}_2]_{\mathrm{F}}^{\top} & [\boldsymbol{\nabla}_3]_{\mathrm{F}}^{\top} \end{bmatrix}$$

so that one has

$$[\mathrm{div}]_{\mathrm{F}} = -\overline{[\boldsymbol{\nabla}]}_{\mathrm{F}}^{\top}. \tag{87}$$

Remark 12. *In the construction of the matrix $[A]$, the gradient has to be applied on vector $[v^*]_F$ and the matrix representing the divergence must give a vector of type $[v^*]_F$ as output, i.e. without the null frequency. This leads to slight changes in the expression of the matrices $[\boldsymbol{\nabla}_j]_F$, $[\boldsymbol{\nabla}]_F$ and $[\mathrm{div}]_F$. In the matrices $[\boldsymbol{\nabla}_j]_F$, the null frequency must be omitted among the terms k_j in (86), and a first line of zeros must be added. Despite this, the property (87) remains unchanged.*

C.4 Matrix of constitutive properties

In the conductivity case, when C in (25) is isotropic at all point in Ω, the constitutive equation reads

$$\sigma_j(x_1, x_2, x_3) = C(x_1, x_2, x_3) \frac{\partial v}{\partial x_j}(x_1, x_2, x_3), \quad \forall (x_1, x_2, x_3)$$

and, with a matricial notation, $[C]_{\mathrm{S}}$ is a $3N \times 3N$ diagonal matrix

$$[\sigma]_{\mathrm{S}} = [C]_{\mathrm{S}} [\boldsymbol{\nabla} v]_{\mathrm{S}}$$

with $[\boldsymbol{\nabla} v]_{\mathrm{S}}$ and $[\sigma]_{\mathrm{S}}$ are organised as in (85) with, owing to the isotropy assumption, a block diagonal matrix

$$[C]_{\mathrm{S}} = \begin{bmatrix} [c]_{\mathrm{S}} & [0] & [0] \\ [0] & [c]_{\mathrm{S}} & [0] \\ [0] & [0] & [c]_{\mathrm{S}} \end{bmatrix}$$

and where, as the constitutive relation is local, $[c]_{\mathrm{S}}$ is itself a $N \times N$ real-valued diagonal matrix whose diagonal elements are given by:

$$[c]_{\mathrm{S}}(i,i) = C_{\mathrm{S}}(i_1, i_2, i_3), \tag{88}$$

where C_{S} is defined and indexed consistently with (61) and (79).

C.5 Properties of the matrices $[A]$ and $[A_0]$

Using (83), (87), (84) and (88), the decomposition of $[A]$, already given in (62)

$$[A] = -[P_{\mathrm{T}}^{\mathrm{F}}]^{-1}[\mathrm{div}]_{\mathrm{F}}[F][C]_{\mathrm{S}}[F]^{-1}[\boldsymbol{\nabla}]_{\mathrm{F}}[P_{\mathrm{T}}^{\mathrm{F}}]$$

can be rewritten as follows

$$[A] = 2\overline{[P_{\mathrm{T}}^{\mathrm{F}}]}^{\top}\overline{[\boldsymbol{\nabla}]}_{\mathrm{F}}^{\top}[F][C]_{\mathrm{S}}\overline{[F]}^{\top}[\boldsymbol{\nabla}]_{\mathrm{F}}[P_{\mathrm{T}}^{\mathrm{F}}].$$

Thus $[A]$ is hermitian as

$$[A] = \overline{[A]}^{\top}.$$

Moreover, by construction, $[A]$ is a real-valued matrix, although some of the components of the matrix decomposition (62) are not, thus $[A]$ is symmetrical.

The matrix $[A_0]$ is a special case of $[A]$ matrix where $C(\boldsymbol{x}) = C_0$ for all $\boldsymbol{x} \in \Omega$, with

$$[A_0] = [P_{\mathrm{T}}^{\mathrm{F}}]^{-1}\overline{[\boldsymbol{\nabla}]}_{\mathrm{F}}^{\top}[F][C_0]_{\mathrm{S}}[F]^{-1}[\boldsymbol{\nabla}]_{\mathrm{F}}[P_{\mathrm{T}}^{\mathrm{F}}] \tag{89}$$

The expression of $[A_0]_{\mathrm{F}}$ (in Fourier representation) simply reads

$$[A_0]_{\mathrm{F}} = \overline{[\boldsymbol{\nabla}]}_{\mathrm{F}}^{\top}[F][C_0]_{\mathrm{S}}[F]^{-1}[\boldsymbol{\nabla}]_{\mathrm{F}}.$$

When C_0 is scalar then $[C_0]_{\mathrm{S}} = C_0[I]_{3N\times 3N}$ and $[A_0]_{\mathrm{F}}$ reduces to $[A_0]_{\mathrm{F}} = C_0\overline{[\boldsymbol{\nabla}]}_{\mathrm{F}}^{\top}[\boldsymbol{\nabla}]_{\mathrm{F}}$. Defining the matrix $[D] \in \mathbb{R}^{M\times M}$ as the diagonal matrix with components

$$[D](k,k) = -(2\pi)^2\left(\frac{k_1^2}{L_1^2} + \frac{k_2^2}{L_2^2} + \frac{k_3^2}{L_3^2}\right) \quad \text{for} \quad (k_1,k_2,k_3) \in \mathcal{K}^+,$$

with (k_1, k_2, k_3) being the frequency indices corresponding to k, see (80). Then, one can define the discrete Laplace operator in the Fourier domain as the block matrix $[\Delta]_{\mathrm{F}} \in \mathbb{R}^{(N-1)\times(N-1)}$ given by

$$[\Delta]_{\mathrm{F}} = \begin{bmatrix} [D] & [0] \\ [0] & [D] \end{bmatrix} \tag{90}$$

so that the matrix $[A_0]_{\mathrm{F}}$ considered above reduces to

$$[A_0]_{\mathrm{F}} = -C_0[\Delta]_{\mathrm{F}}.$$

The matrix $[A_0]$ in the sine-cosine representation has a similar expression. Indeed, when C_0 is scalar, (89) becomes

$$[A_0] = -C_0\overline{[P^{\mathrm{F}}_{\mathrm{T}}]}^{-1}[\Delta]_{\mathrm{F}}[P^{\mathrm{F}}_{\mathrm{T}}]$$

and thus

$$[A_0] = -\frac{1}{2}C_0\begin{bmatrix}[I]_{M\times M} & [I]_{M\times M}\\ +\mathrm{i}[I]_{M\times M} & -\mathrm{i}[I]_{M\times M}\end{bmatrix}\begin{bmatrix}[D] & [0]\\ [0] & [D]\end{bmatrix}\begin{bmatrix}[I]_{M\times M} & -\mathrm{i}[I]_{M\times M}\\ [I]_{M\times M} & +\mathrm{i}[I]_{M\times M}\end{bmatrix} = -C_0[\Delta]$$

where the matrix $[\Delta] \in \mathbb{R}^{(N-1)\times(N-1)}$ is defined in the same way than (90), for an ordering of the indices k corresponding to the sine-cosine decomposition.

References

[1] G. Allaire. *Shape Optimization by the Homogenization Method*. Springer, 2002.

[2] E. Anderson, Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, and D. Sorensen. *LAPACK Users' Guide*. Society for Industrial and Applied Mathematics, Philadelphia, PA, third edition, 1999.

[3] C. Bellis and H. Moulinec. Spectrum of the operator of Lippmann-Schwinger equation - a numerical study. In *9th European Congress on Computational Methods in Applied Sciences and Engineering*, Lisbon, Portugal, 2024. https://hal.science/hal-04818271.

[4] C. Bellis, H. Moulinec, and P. Suquet. Eigendecomposition-based convergence analysis of the neumann series for laminated composites and discretization error estimation. *International Journal for Numerical Methods in Engineering*, 121:201–232, 2020.

[5] D. Bergman. The dielectric constant of a composite material – a problem in classical physics. *Phys. Rep. C*, 43(9):377–407, 1978.

[6] D. J. Bergman. The dielectric constant of a simple cubic array of identical spheres. *J. Phys. C: Solid State Phys.*, 12, 1979.

[7] H. Brézis. *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. Springer, New York, 2011.

[8] J. Céa. *Optimisation - Théorie et Algorithmes*. Dunod, 1971.

[9] F. Chatelin. *Spectral Approximation of Linear Operators*. SIAM, 2011.

[10] E. Cherkaev. Inverse homogenization for evaluation of effective properties of a mixture. *Inverse Problems*, 17(4):1203, 2001.

[11] D. Cioranescu and P. Donato. *An Introduction to Homogenization*. Oxford University Press, 1999.

[12] J. D. Eshelby. The Determination of the Elastic Field of an Ellipsoidal Inclusion, and Related Problems. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 241(1226):376–396, 1957.

[13] I. Faragó and J. Karátson. *Numerical solution of nonlinear elliptic problems via preconditioning operators. Theory and applications*, volume 11, Advances in Computation. NOVA Science, 2002.

[14] R. Ferrier and C. Bellis. A posteriori error estimations and convergence criteria in fast fourier transform-based computational homogenization. *International Journal for Numerical Methods in Engineering*, 124(4):834–863, 2023.

[15] T. Gergelits, B. F. Nielsen, and Z. Strakoš. Numerical approximation of the spectrum of self-adjoint operators in operator preconditioning. *Numer. Algor.*, 91:301–325, 2022.

[16] K. Golden and G. Papanicolaou. Bounds for effective parameters of heterogeneous media by analytic continuation. *Commun. Math. Phys.*, 90:473–491, 1983.

[17] G. H. Golub and C. F. Van Loan. *Matrix Computations*. The Johns Hopkins University Press, 4^{th} edition, 2013.

[18] Z. Hashin and S. Shtrikman. On some variational principles in anisotropic and non-homogeneous elasticity. *Journal of the Mechanics and Physics of Solids*, 10:335–342, 1962.

[19] R. Hill. New derivations of some elastic extremum principles. In *Progress in Applied Mechanics: The Prager Anniversary Volume*, pages 99–106. New York: Macmillan Publishing Company, 1963.

[20] M. Kabel, T. Böhlke, and M. Schneider. Efficient fixed point and Newton-Krylov solvers for FFT-based homogenization of elasticity at large deformations. *Comput. Mech.*, 54:1497–1514, 2014.

[21] Y. Kantor and D. Bergman. Elastostatic resonances – a new approach to the calculation of the effective elastic constants of composites. *Journal of Mechanics Physics of Solids*, 30(5):355–376, 1982.

[22] L.V. Kantorovich and G.P. Akilov. *Functional Analysis*. Pergamon, 2^{nd} edition, 1982.

[23] T. Kato. *Perturbation Theory for Linear Operators*. Springer-Verlag, Berlin, Heidelberg, New-York, 1980.

[24] E. Kreyszig. *Introductory Functional Analysis with Applications*. Wiley, 1978.

[25] E. Kröner. *Statistical Continuum Mechanics*. Springer, 1971.

[26] M. Ladecký, R. J. Leute, A. Falsafi, I. Pultarová, L. Pastewka, T. Junge, and J. Zeman. An optimal preconditioned FFT-accelerated finite element solver for homogenization. *Applied Mathematics and Computation*, 446:127835, 2023.

[27] R. J. Leute, M. Ladecký, A. Falsafi, I. Jödicke, I. Pultarová, J. Zeman, T. Junge, and L. Pastewka. Elimination of ringing artifacts by finite-element projection in FFT-based homogenization. *Journal of Computational Physics*, 453:110931, 2022.

[28] Z. Liu, M. A. Bessa, and W. K. Liu. Self-consistent clustering analysis: An efficient multi-scale scheme for inelastic heterogeneous materials. *Computer Methods in Applied Mechanics and Engineering*, 306:319–341, 2016.

[29] S. Lucarini, M. V. Upadhyay, and J. Segurado. FFT based approaches in micromechanics: fundamentals, methods and applications. *Modelling and Simulation in Materials Science and Engineering*, 30(2):023002, 2021.

[30] J. Málek and Z. Strakoš. *Preconditioning and the Conjugate Gradient Method in the Context of Solving PDEs*. SIAM, 2015.

[31] R. C. McPhedran and D. R. McKenzie. Electrostatic and optical resonances of arrays of cylinders. *Applied physics*, 23:223–235, 1980.

[32] J.-C. Michel, H. Moulinec, and P. Suquet. A computational method for linear and nonlinear composites with arbitrary phase contrast. *Int. J. Numer. Meth. Engng*, 52:139–160, 2001.

[33] G. W. Milton. Bounds on the complex dielectric constant of a composite material. *Applied Physics Letters*, 37(3):300–302, 1980.

[34] G. W. Milton. *The Theory of Composites*. Cambridge university press, 2002.

[35] H. Moulinec and P. Suquet. A fast numerical method for computing the linear and nonlinear properties of composites. *C. R. Acad. Sc. Paris, II*, 318:1417–1423, 1994.

[36] H. Moulinec and P. Suquet. A numerical method for computing the overall response of nonlinear composites with complex microstructure. *Comp. Meth. Appl. Mech. Engng.*, 157:69–94, 1998.

[37] H. Moulinec, P. Suquet, and G. W. Milton. Convergence of iterative methods based on Neumann series for composite materials: Theory and practice. *Int. J. Numerical Methods in Engineering*, 2018.

[38] N. B. Murphy, E. Cherkaev, C. Hohenegger, and K. M. Golden. Spectral measure computations for composite materials. *Commun. Math. Sci.*, 13:825–862, 2015.

[39] J. W. Neuberger. *Sobolev Gradients and Differential Equations*. Springer, Berlin, Heidelberg, 2009.

[40] B. F. Nielsen and Z. Strakoš. A simple formula for the generalized spectrum of second order self-adjoint differential operators. *SIAM Review*, 66:125–146, 2024.

[41] B. F. Nielsen, A. Tveito, and W. Hackbusch. Preconditioning by inverting the Laplacian; an analysis of the eigenvalues. *IMA Journal of Numerical Analysis*, 29:24–42, 2009.

[42] Y. V. Obnosov. Periodic heterogeneous structures: New explicit solutions and effective characteristics of refraction of an imposed field. *SIAM Journal on Applied Mathematics*, 59(4):1267–1287, 1999.

[43] M. Reed and B. Simon. *Methods of Modern Mathematical Physics*. Academic Press, 1980.

[44] M. Schneider. A review of nonlinear FFT-based computational homogenization methods. *Acta Mech*, 232:2051–2100, 2021.

[45] J.R. Willis. Bounds and self-consistent estimates for the overall properties of anisotropic composites. *Journal of the Mechanics and Physics of Solids*, 25(3):185 – 202, 1977.

[46] D. Zhang and E. Cherkaev. Reconstruction of spectral function from effective permittivity of a composite material using rational function approximations. *J. Comput. Phys.*, 228:5390–5409, 2009.