



HAL
open science

Comparison of three accelerated FFT-based schemes for computing the mechanical response of composite materials

Hervé Moulinec, Fabrice Silva

► **To cite this version:**

Hervé Moulinec, Fabrice Silva. Comparison of three accelerated FFT-based schemes for computing the mechanical response of composite materials. *International Journal for Numerical Methods in Engineering*, 2014, 97 (13), pp.960-985. 10.1002/nme.4614 . hal-00787089

HAL Id: hal-00787089

<https://hal.science/hal-00787089>

Submitted on 11 Feb 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Comparison of three accelerated FFT-based schemes for computing the mechanical response of composite materials

H. Moulinec, F. Silva

February 8, 2013

L.M.A./ C.N.R.S., 31 Chemin Joseph Aiguier, 13402 Marseille cedex 20, France

Abstract

Since Moulinec & Suquet ([1], [2]) introduced an iterative method based on Fourier transforms to compute the mechanical properties of heterogeneous materials, improved algorithms have been proposed to increase the convergence rate of the scheme. This paper is devoted to the comparison of the accelerated schemes proposed by Eyre & Milton ([8]), by Michel et al ([10]) and by Monchiet & Bonnet ([11]). It shows that the algorithms by Eyre-Milton and by Michel et al are particular cases of Monchiet-Bonnet algorithm, corresponding to particular choices of parameters of the method. An upper bound of the spectral radius of the schemes is determined, which enables to propose sufficient conditions of convergence of the schemes. Conditions are found for minimizing this upper bound. This study shows that the scheme which minimizes this upper bound is the scheme of Eyre-Milton. The paper discusses the choice of the convergence test used in the schemes.

1 Introduction

An iterative method has been proposed by Moulinec & Suquet [1, 2] to compute the overall and local response of a heterogeneous material submitted to a mechanical solicitation. This method is based on the iterative resolution of the Lippmann-Schwinger equation which involves the Green's operator associated to a reference linear elastic material. This scheme, hereafter called "the basic scheme", is briefly recalled in section 2.2.

Since then, several authors proposed variants or improvements of the initial method (among others: Müller [3], Lebensohn [4], Vinogradov & Milton [5], Zeman et al [6], Brisard & Dormieux [7], ...). In particular, Eyre & Milton [8, 9], Michel et al [10] and Monchiet & Bonnet [11] proposed accelerated algorithms to overcome the low convergence rate of the basic scheme for highly-contrasted materials. These three schemes are presented in section 3; their respective iterative relations are re-written in order to make appear their similarity. The section 4 is devoted to the mathematical study of the convergence of the schemes. The choice of the convergence test used in the iterations of the algorithms is discussed in section 5. An efficient implementation of the algorithm is proposed in section 6. Finally, numerical simulations are presented in section 7 to illustrate and confirm results of 4 and 5.

2 General considerations and basic scheme

2.1 Local problem

The problem under consideration consists in finding the response of a volume element V of a heterogeneous linear elastic material submitted to a prescribed macroscopic strain \mathbf{E} .

The microstructure is assumed to be periodic and is therefore fully described by a unit cell V . The stiffness tensor is known at each point \mathbf{x} of V and denoted by $\mathbf{c}(\mathbf{x})$. The constituents are supposed to be perfectly bonded at interfaces. The deformations are assumed to be infinitesimal. The local problem to be solved can be summarized as:

$$\begin{cases} \boldsymbol{\sigma}(\mathbf{x}) = \mathbf{c}(\mathbf{x}) : \boldsymbol{\varepsilon}(\mathbf{x}) , \forall \mathbf{x} \in V & \text{(constitutive relations)} \\ \operatorname{div} \boldsymbol{\sigma} = 0 & \text{(equilibrium)} \\ \boldsymbol{\varepsilon}(\mathbf{x}) = \frac{1}{2}(\nabla \mathbf{u}(\mathbf{x}) + \nabla^T \mathbf{u}(\mathbf{x})) & \text{(compatibility)} \\ \langle \boldsymbol{\varepsilon} \rangle = \mathbf{E} + \text{periodicity conditions} , \end{cases} \quad (1)$$

where $\boldsymbol{\varepsilon}$, $\boldsymbol{\sigma}$ and \mathbf{u} respectively denote the local strain, the local stress, the local displacement in V , and where $\langle . \rangle$ denotes the average of a given field in volume V .

2.2 “Basic” iterative scheme

Using the idea of a reference medium (Eshelby [12]), Moulinec and Suquet ([1], [2]) have introduced an iterative scheme to solve problem (1) which can be summarized by the following relation where the strain field $\boldsymbol{\varepsilon}^{i+1}$ at iterate $i+1$ is updated from the previous iterate i :

$$\boldsymbol{\varepsilon}^{i+1}(\mathbf{x}) = -\mathbf{\Gamma}^0 * ((\mathbf{c}(\mathbf{x}) - \mathbf{c}^0) : \boldsymbol{\varepsilon}^i(\mathbf{x})) + \mathbf{E} , \quad (2)$$

where \mathbf{c}^0 is a homogeneous linear elastic material, where $\mathbf{\Gamma}^0$ is the Green’s operator associated to \mathbf{c}^0 (an analytical expression of $\mathbf{\Gamma}^0$ is given in appendix (A.1) in the usual case when the reference medium \mathbf{c}^0 is isotropic), and where $*$ denotes the convolution operator.

Relation (2) can profitably be applied in the Fourier space as the convolution operator in real space becomes a simple product in Fourier space:

$$\forall \boldsymbol{\xi} \neq \mathbf{0} \quad \hat{\boldsymbol{\varepsilon}}^{i+1}(\boldsymbol{\xi}) = -\hat{\mathbf{\Gamma}}^0(\boldsymbol{\xi}) : (\widehat{\mathbf{c} - \mathbf{c}^0}) : \hat{\boldsymbol{\varepsilon}}^i(\boldsymbol{\xi}) , \quad \hat{\boldsymbol{\varepsilon}}^{i+1}(\mathbf{0}) = \mathbf{E} , \quad (3)$$

where a ‘hat’ over a quantity means the Fourier transform of that quantity and where $\boldsymbol{\xi}$ is the spatial frequency in Fourier space.

As proved in [13], when the material is isotropic, a sufficient condition of convergence of the scheme (2) is that:

$$k^0 > \frac{k(\mathbf{x})}{2} \quad \text{and} \quad \mu^0 > \frac{\mu(\mathbf{x})}{2} \quad \forall \mathbf{x} \in V, \quad (4)$$

where $k(\mathbf{x})$ and $\mu(\mathbf{x})$ are respectively the bulk modulus and the shear modulus of $\mathbf{c}(\mathbf{x})$, and where k^0 and μ^0 are respectively the bulk modulus and the shear modulus of \mathbf{c}^0 (which is set to be isotropic too). An upper bound of the spectral radius of the scheme has been found; this bound is minimized when:

$$k^0 = \frac{1}{2}(k_{min} + k_{max}) , \quad \mu^0 = \frac{1}{2}(\mu_{min} + \mu_{max}) , \quad (5)$$

with k_{min} and k_{max} (respectively μ_{min} and μ_{max}) are the minimum and maximum values of $k(\mathbf{x})$ (respectively $\mu(\mathbf{x})$) in V :

$$\begin{aligned} k_{min} &= \min_{\mathbf{x} \in V}(k(\mathbf{x})) , & k_{max} &= \max_{\mathbf{x} \in V}(k(\mathbf{x})) , \\ \mu_{min} &= \min_{\mathbf{x} \in V}(\mu(\mathbf{x})) , & \mu_{max} &= \max_{\mathbf{x} \in V}(\mu(\mathbf{x})) . \end{aligned}$$

This lowest upper bound is equal to:

$$\frac{\mathcal{C} - 1}{\mathcal{C} + 1} \quad (6)$$

where \mathcal{C} is the contrast between the mechanical properties in the material and is defined as:

$$\mathcal{C} = \max(k_{max}/k_{min}, \mu_{max}/\mu_{min}) \quad (7)$$

3 Presentation of the three schemes

3.1 Scheme of Eyre and Milton

Eyre and Milton [8] have introduced an accelerated scheme which has been reformulated in [13] as:

$$e^{i+1} = e^i - 2(\delta c)^{-1} : c^0 : (\Gamma^0 * (c^0 : e^i) - e^i + E + \Gamma^0 * (c : e^i)) .$$

To avoid negative stiffness for the reference medium as initially proposed by the authors, it is preferable to change the sign of c^0 and thus to write the iterative relation as:

$$e^{i+1} = e^i + 2(c + c^0)^{-1} : c^0 : (\Gamma^0 * (c^0 : e^i) - e^i + E - \Gamma^0 * (c : e^i)) ,$$

which can be re-written as:

$$(c + c^0) : e^{i+1} = (c + c^0) : e^i - 2c^0 : \Gamma^0 * (c : e^i) - 2\Delta^0 * e^i - 2c^0 : (\langle e^i \rangle - E) , \quad (8)$$

where Δ^0 is the so-called ‘‘stress Green’s tensor’’ which has been introduced by Bhattacharya & Suquet in [14] and which is related to $\hat{\Gamma}^0(\xi)$ in Fourier space by:

$$\hat{\Delta}^0(\xi) = c^0 - c^0 : \hat{\Gamma}^0(\xi) : c^0 , \quad \forall \xi \neq \mathbf{0} \quad \text{and} \quad \Delta^0(\xi) = \mathbf{0} , \quad \text{when} \quad \xi = \mathbf{0} . \quad (9)$$

Alternatively, the operator Δ^0 can be defined in real space as:

$$\Delta^0 * t(x) = c^0 : (t(x) - \langle t \rangle - \Gamma^0 * c^0 : t(x)) , \quad \text{for any 2d order tensor field } t(x). \quad (10)$$

The following remarks are in order:

- the tensor field denoted here by e^i (rather than by ε^i) converges to the strain tensor field ε solution of (1), but, before convergence has been reached, e^i can violate the compatibility conditions,
- similarly, the condition of prescribed macroscopic strain is met at convergence, but generally not before. Indeed, using properties of Γ^0 and Δ^0 detailed in (A.2), averaging relation (8) gives:

$$\langle (c + c^0) : e^{i+1} \rangle = \langle (c + c^0) : e^i \rangle - 2c^0 : (\langle e^i \rangle - E) ,$$

which implies at convergence that:

$$\langle e^{i=\infty} \rangle - E = \mathbf{0} .$$

3.2 Augmented Lagrangian scheme

A scheme based on the augmented Lagrangian method has been proposed by Michel, Moulinec, Suquet in [10]. The three unknowns of the problem are $e(x)$, $\varepsilon(x)$, $\lambda(x)$, where $e(x)$ satisfies the constitutive relations, but not necessarily the compatibility condition, $\varepsilon(x)$ is a compatible strain field and $\lambda(x)$ is the Lagrange multiplier for the constraint $e(x) = \varepsilon(x)$. Using Uzawa’s algorithm, the proposed scheme can be shortly written as:

Iterate $i + 1$: given e^i and λ^i ,

$$(a) \quad \tau^i(x) = \lambda^i(x) - c^0 : e^i(x) ,$$

$$(b) \quad \hat{\tau}^i = \mathcal{F}(\tau^i) ,$$

$$(c) \quad \hat{\varepsilon}^{i+1}(\xi) = -\hat{\Gamma}^0 : \hat{\tau}^i(\xi) \quad \forall \xi \neq \mathbf{0} , \quad \hat{\varepsilon}^{i+1}(\mathbf{0}) = E ,$$

$$(d) \quad e^{i+1} = \mathcal{F}^{-1}(\hat{\varepsilon}^{i+1}) ,$$

$$(e) \quad \text{Solve for } e^{i+1}(x) :$$

$$\frac{\partial w}{\partial e}(x, e^{i+1}) + c^0 : e^{i+1}(x) = c^0 : \varepsilon^{i+1}(x) + \lambda^i(x) ,$$

$$(f) \quad \lambda^{i+1}(x) = \lambda^i(x) + d^0 : (\varepsilon^{i+1}(x) - e^{i+1}(x)) ,$$

$$(g) \quad \text{convergence test} ,$$

where $\overline{\mathcal{F}}$ and $\overline{\mathcal{F}}^{-1}$ respectively denote the Fourier transform and the inverse Fourier transform.

When \mathbf{d}^0 is chosen to be equal to \mathbf{c}^0 , if λ^i is replaced by its expression derived from (e), step (f) simplifies into:

$$\lambda^{i+1}(\mathbf{x}) = \frac{\partial w}{\partial \mathbf{e}}(\mathbf{x}, \mathbf{e}^{i+1}).$$

The Lagrangian multiplier on the compatibility constraint appears to be the stress related to \mathbf{e}^i . Eliminating $\lambda^{i+1}(\mathbf{x})$ and $\lambda^i(\mathbf{x})$ in the other steps, the algorithm can thus be summarized in a sole relation which reads:

$$\frac{\partial w}{\partial \mathbf{e}}(\mathbf{x}, \mathbf{e}^{i+1}) + \mathbf{c}^0 : \mathbf{e}^{i+1}(\mathbf{x}) = -\mathbf{c}^0 : \Gamma^0 * \left(\frac{\partial w}{\partial \mathbf{e}}(\mathbf{x}, \mathbf{e}^i) - \mathbf{c}^0 : \mathbf{e}^i(\mathbf{x}) \right) + \mathbf{c}^0 : \mathbf{E} + \frac{\partial w}{\partial \mathbf{e}}(\mathbf{x}, \mathbf{e}^i).$$

In the linear elastic case, $\frac{\partial w}{\partial \mathbf{e}}(\mathbf{x}, \mathbf{e})$ can be replaced by $\mathbf{c}(\mathbf{x}) : \mathbf{e}(\mathbf{x})$ and the previous relation becomes:

$$(\mathbf{c}(\mathbf{x}) + \mathbf{c}^0) : \mathbf{e}^{i+1}(\mathbf{x}) = -\mathbf{c}^0 : \Gamma^0 * (\mathbf{c}(\mathbf{x}) : \mathbf{e}^i(\mathbf{x}) - \mathbf{c}^0 : \mathbf{e}^i(\mathbf{x})) + \mathbf{c}^0 : \mathbf{E} + \mathbf{c}(\mathbf{x}) : \mathbf{e}^i(\mathbf{x}),$$

which in turn, can be written as:

$$(\mathbf{c} + \mathbf{c}^0) : \mathbf{e}^{i+1} = (\mathbf{c} + \mathbf{c}^0) : \mathbf{e}^i - \mathbf{c}^0 : \Gamma^0 * \mathbf{c} : \mathbf{e}^i - \Delta^0 * \mathbf{e}^i - \mathbf{c}^0 : (\langle \mathbf{e}^i \rangle - \mathbf{E}). \quad (11)$$

3.3 Polarization-based scheme

Monchiet and Bonnet ([11]) have recently introduced a scheme based on a polarization field defined as:

$$\boldsymbol{\tau} = (\mathbf{c} - \mathbf{c}^0) : \mathbf{e} \quad ,$$

to solve the problem of heterogeneous linear elastic material submitted to a prescribed macroscopic polarization \mathbf{T} . The associated iterative relation reads as:

$$\boldsymbol{\tau}^{i+1} = \boldsymbol{\tau}^i - \alpha \mathbf{c}^0 : \Gamma^0 * \mathbf{c} : \mathbf{e}^i - \beta \Delta^0 * \mathbf{e}^i - \langle \boldsymbol{\tau}^i \rangle + \mathbf{T} \quad .$$

If one chooses to replace \mathbf{c}^0 by $-\mathbf{c}^0$ and to change β into $-\beta$, the polarization-based scheme becomes:

$$(\mathbf{c} + \mathbf{c}^0) : \mathbf{e}^{i+1} = (\mathbf{c} + \mathbf{c}^0) : \mathbf{e}^i - \alpha \mathbf{c}^0 : \Gamma^0 * \mathbf{c} : \mathbf{e}^i - \beta \Delta^0 * \mathbf{e}^i - \langle (\mathbf{c} + \mathbf{c}^0) : \mathbf{e}^i \rangle + \mathbf{T} \quad . \quad (12)$$

In order to compare the polarization-based scheme with the schemes presented above in 3.1 and 3.2, we propose to modify it slightly to deal with the prescribed macroscopic strain (which is more commonly used in practice than prescribed macroscopic polarization):

$$(\mathbf{c} + \mathbf{c}^0) : \mathbf{e}^{i+1} = (\mathbf{c} + \mathbf{c}^0) : \mathbf{e}^i - \alpha \mathbf{c}^0 : \Gamma^0 * \mathbf{c} : \mathbf{e}^i - \beta \Delta^0 * \mathbf{e}^i - \beta \mathbf{c}^0 : (\langle \mathbf{e}^i \rangle - \mathbf{E}) \quad . \quad (13)$$

This algorithm ensures that the average of the strain at convergence is equal to the prescribed macroscopic strain \mathbf{E} . Indeed, one has the following iterative relation on the average of the strain and stress fields:

$$\langle (\mathbf{c} + \mathbf{c}^0) : \mathbf{e}^{i+1} \rangle = \langle (\mathbf{c} + \mathbf{c}^0) : \mathbf{e}^i \rangle - \beta \mathbf{c}^0 : (\langle \mathbf{e}^i \rangle - \mathbf{E}),$$

which implies at convergence that: $\langle \mathbf{e}^{i=\infty} \rangle = \mathbf{E}$.

3.4 Comparison of the 3 schemes

The simple comparison of the 3 schemes as presented in (8), (11) and (13) reveals their strong similarity. The scheme of Eyre-Milton and the augmented Lagrangian scheme appear to be particular cases of polarization-based scheme with $\alpha = \beta = 2$ (as already noted in [11]) and $\alpha = \beta = 1$ respectively.

This similarity is very remarkable because the underlying ideas from which the schemes were derived are different:

- the scheme of Eyre-Milton is based on a re-conditioning of the operator involved in the iterative relation of the basic scheme to improve its convergence rate,
- the scheme in (11) relies on an augmented Lagrangian to deal with non-linear problems,
- the polarization-based scheme has been devised to take advantage, in terms of convergence rate, of both the basic scheme (2) based on an iterative relation on strains, and of the dual stress-based scheme.

The three schemes update the so-called ‘‘polarization’’ term $((\mathbf{c} + \mathbf{c}^0) : \mathbf{e}^i)$ at each iteration with a combination of three quantities:

- $\mathbf{\Gamma}^0 * \mathbf{c} : \mathbf{e}^i$ which is a measure of the error on the divergence of the stress,
- $\mathbf{\Delta}^0 * \mathbf{e}^i$ which is a measure of the error on compatibility of the strain,
- $\langle \mathbf{e}^i \rangle - \mathbf{E}$ which is a measure of the deviation from the prescribed macroscopic strain.

4 Rate of convergence of the iterative scheme

4.1 Conditions for convergence

The error at iterate i of the schemes (8), (11) or (13) can be defined as the difference between the estimated strain and the strain of the exact solution:

$$\epsilon_i = \|\mathbf{e}^i - \mathbf{e}^\infty\| \quad (14)$$

which can be bounded by:

$$\epsilon_i \leq \mathcal{R}^i \cdot \epsilon_0 \quad (15)$$

where \mathcal{R} is the spectral radius of the operator involved in the iterations, i.e. the supremum of the absolute value of its eigenvalues. The scheme converges if and only if the spectral radius is smaller than 1, i.e.:

$$\mathcal{R} < 1. \quad (16)$$

In the following paragraphs, conditions on the parameters of the scheme (13) are searched for to ensure that condition (16) is satisfied.

An eigenvalue Φ of the operator used in the iterations, and an associated eigenvector \mathbf{e}_Φ , satisfy the homogeneous equation:

$$\Phi(\mathbf{c} + \mathbf{c}^0) : \mathbf{e}_\Phi = (\mathbf{c} + \mathbf{c}^0) : \mathbf{e}_\Phi - \alpha \mathbf{c}^0 : \mathbf{\Gamma}^0 * \mathbf{c} : \mathbf{e}_\Phi - \beta \mathbf{\Delta}^0 * \mathbf{e}_\Phi - \beta \mathbf{c}^0 : \langle \mathbf{e}_\Phi \rangle. \quad (17)$$

Applying $\mathbf{c}^0 : \mathbf{\Gamma}^0$ to relation (17), and using properties of $\mathbf{\Gamma}^0$ and $\mathbf{\Delta}^0$ detailed in appendix A, one easily obtains that:

$$\mathbf{c}^0 : \mathbf{\Gamma}^0 * ((\Phi - 1)(\mathbf{c} + \mathbf{c}^0) : \mathbf{e}_\Phi + \alpha \mathbf{c} : \mathbf{e}_\Phi) = \mathbf{0}. \quad (18)$$

Now, applying $\mathbf{\Delta}^0 * (\mathbf{c}^0)^{-1}$ to (17), one obtains that:

$$\mathbf{\Delta}^0 * ((\Phi - 1)(\mathbf{c}^0)^{-1} : (\mathbf{c} + \mathbf{c}^0) : \mathbf{e}_\Phi + \beta \mathbf{e}_\Phi) = \mathbf{0}. \quad (19)$$

Finally, averaging (17) on the volume V , one obtains that:

$$\langle (\Phi - 1)(\mathbf{c}^0)^{-1} : (\mathbf{c} + \mathbf{c}^0) : \mathbf{e}_\Phi + \beta \mathbf{e}_\Phi \rangle = \mathbf{0}. \quad (20)$$

Following [11], if one introduces:

$$\begin{aligned} \boldsymbol{\sigma}_\Phi &= (\Phi - 1) \times (\mathbf{c} + \mathbf{c}^0) : \mathbf{e}_\Phi + \alpha \mathbf{c} : \mathbf{e}_\Phi \\ \boldsymbol{\varepsilon}_\Phi &= (\Phi - 1) \times \mathbf{c}^{0^{-1}} : (\mathbf{c} + \mathbf{c}^0) : \mathbf{e}_\Phi + \beta \mathbf{e}_\Phi, \end{aligned} \quad (21)$$

the relations (18), (19) and (20) can be re-formulated as:

$$\begin{aligned}\mathbf{\Gamma}^0 * \boldsymbol{\sigma}_\Phi &= 0 \\ \mathbf{\Delta}^0 * \boldsymbol{\varepsilon}_\Phi &= 0 \\ \langle \boldsymbol{\varepsilon}_\Phi \rangle &= 0,\end{aligned}\tag{22}$$

i.e. $\boldsymbol{\sigma}_\Phi$ must be equilibrated, $\boldsymbol{\varepsilon}_\Phi$ must be compatible, and the average of $\boldsymbol{\varepsilon}_\Phi$ is null.

On the other hand, because of (21), $\boldsymbol{\sigma}_\Phi$ and $\boldsymbol{\varepsilon}_\Phi$ are related by:

$$\boldsymbol{\sigma}_\Phi = \mathbf{c}'(\mathbf{x}) : \boldsymbol{\varepsilon}_\Phi\tag{23}$$

with:

$$\mathbf{c}'(\mathbf{x}) = [(\Phi - 1) \cdot (\mathbf{c}(\mathbf{x}) + \mathbf{c}^0) + \alpha \mathbf{c}(\mathbf{x})] : [(\Phi - 1) \cdot (\mathbf{c}(\mathbf{x}) + \mathbf{c}^0) + \beta \mathbf{c}^0]^{-1} : \mathbf{c}^0\tag{24}$$

In conclusion, solving the eigenvalue problem is equivalent to finding when the heterogeneous linear elastic problem, with elastic stiffness \mathbf{c}' given above and with prescribed macroscopic strain: $\mathbf{E} = \mathbf{0}$ has a solution other than the trivial one: $\boldsymbol{\varepsilon}_\Phi(\mathbf{x}) = \mathbf{0}$, $\forall x$.

Thus, for a given heterogeneous elastic problem $\mathbf{c}(\mathbf{x})$, a *sufficient* condition of convergence for the iterative scheme consists in finding α , β and \mathbf{c}^0 which guarantee that \mathbf{c}' is either positive definite or negative definite whenever $|\Phi| \geq 1$, implying that all eigenvalues lie in the range $] - 1, 1[$.

When $\mathbf{c}(\mathbf{x})$ is isotropic, $\mathbf{c}'(\mathbf{x})$ is isotropic too and its bulk and shear moduli are given by:

$$\begin{aligned}k'(\mathbf{x}) &= \frac{(\Phi-1) \cdot (k(\mathbf{x})+k^0) + \alpha k(\mathbf{x})}{(\Phi-1) \cdot (k(\mathbf{x})+k^0) + \beta k^0} \cdot k^0 \\ \mu'(\mathbf{x}) &= \frac{(\Phi-1) \cdot (\mu(\mathbf{x})+\mu^0) + \alpha \mu(\mathbf{x})}{(\Phi-1) \cdot (\mu(\mathbf{x})+\mu^0) + \beta \mu^0} \cdot \mu^0.\end{aligned}\tag{25}$$

For a given set of α , β , k^0 , μ^0 , $k(\mathbf{x})$ and $\mu(\mathbf{x})$, k' changes its sign when:

$$\begin{aligned}\Phi &= 1 - \alpha \frac{k(\mathbf{x})}{k(\mathbf{x})+k^0} \\ \Phi &= 1 - \beta \frac{k^0}{k(\mathbf{x})+k^0},\end{aligned}\tag{26}$$

and μ' changes its sign when:

$$\begin{aligned}\Phi &= 1 - \alpha \frac{\mu(\mathbf{x})}{\mu(\mathbf{x})+\mu^0} \\ \Phi &= 1 - \beta \frac{\mu^0}{\mu(\mathbf{x})+\mu^0}.\end{aligned}\tag{27}$$

Therefore, if the following condition is satisfied:

$$\forall \mathbf{x} : \begin{cases} \Phi > 1 - \alpha \frac{k(\mathbf{x})}{k(\mathbf{x})+k^0} \\ \Phi > 1 - \beta \frac{k^0}{k(\mathbf{x})+k^0} \\ \Phi > 1 - \alpha \frac{\mu(\mathbf{x})}{\mu(\mathbf{x})+\mu^0} \\ \Phi > 1 - \beta \frac{\mu^0}{\mu(\mathbf{x})+\mu^0}, \end{cases}$$

assuming that k^0 and μ^0 are positive, one has: $k'(x) > 0$ and $\mu'(x) > 0 \forall x$, i.e. \mathbf{c}' is positive definite and thus Φ cannot be an eigenvalue.

On the other hand, when:

$$\forall \mathbf{x} : \begin{cases} \Phi < 1 - \alpha \frac{k(\mathbf{x})}{k(\mathbf{x})+k^0} \\ \Phi < 1 - \beta \frac{k^0}{k(\mathbf{x})+k^0} \\ \Phi < 1 - \alpha \frac{\mu(\mathbf{x})}{\mu(\mathbf{x})+\mu^0} \\ \Phi < 1 - \beta \frac{\mu^0}{\mu(\mathbf{x})+\mu^0}, \end{cases}$$

\mathbf{c}' is negative definite and thus Φ cannot either be an eigenvalue.

In conclusion, an eigenvalue Φ of the operator used in the iterations verifies:

$$\left\{ \begin{array}{l} \Phi \geq \min_x \left(\min \left(1 - \alpha \frac{k(\mathbf{x})}{k(\mathbf{x}) + k^0}, 1 - \beta \frac{k^0}{k(\mathbf{x}) + k^0}, 1 - \alpha \frac{\mu(\mathbf{x})}{\mu(\mathbf{x}) + \mu^0}, 1 - \beta \frac{\mu^0}{\mu(\mathbf{x}) + \mu^0} \right) \right) \\ \text{and} \\ \Phi \leq \max_x \left(\max \left(1 - \alpha \frac{k(\mathbf{x})}{k(\mathbf{x}) + k^0}, 1 - \beta \frac{k^0}{k(\mathbf{x}) + k^0}, 1 - \alpha \frac{\mu(\mathbf{x})}{\mu(\mathbf{x}) + \mu^0}, 1 - \beta \frac{\mu^0}{\mu(\mathbf{x}) + \mu^0} \right) \right) . \end{array} \right. \quad (28)$$

Therefore, the spectral radius \mathcal{R} of the scheme can be bounded by:

$$\mathcal{R} \leq \mathcal{R}_u \quad (29)$$

with the upper bound \mathcal{R}_u of the spectral radius being defined as:

$$\mathcal{R}_u = \max_x \left(\max \left(\left| 1 - \alpha \frac{k(\mathbf{x})}{k(\mathbf{x}) + k^0} \right|, \left| 1 - \beta \frac{k^0}{k(\mathbf{x}) + k^0} \right|, \left| 1 - \alpha \frac{\mu(\mathbf{x})}{\mu(\mathbf{x}) + \mu^0} \right|, \left| 1 - \beta \frac{\mu^0}{\mu(\mathbf{x}) + \mu^0} \right| \right) \right) . \quad (30)$$

Hence, a *sufficient* condition for the scheme to converge, is to set $\alpha, \beta, k^0, \mu^0$, such that:

$$\mathcal{R}_u < 1 \quad (31)$$

or in other words, such that:

$$\forall x \left\{ \begin{array}{l} \left| 1 - \alpha \frac{k(\mathbf{x})}{k(\mathbf{x}) + k^0} \right| < 1 \\ \left| 1 - \beta \frac{k^0}{k(\mathbf{x}) + k^0} \right| < 1 \\ \left| 1 - \alpha \frac{\mu(\mathbf{x})}{\mu(\mathbf{x}) + \mu^0} \right| < 1 \\ \left| 1 - \beta \frac{\mu^0}{\mu(\mathbf{x}) + \mu^0} \right| < 1 \end{array} \right. \quad (32)$$

Monchiet & Bonnet propose the following inequalities as sufficient conditions for the convergence of their iterative scheme:

$$0 \leq \alpha < 2, \quad 0 \leq \beta < 2, \quad k^0 > 0, \quad \mu^0 > 0$$

(according to Monchiet-Bonnet the cases $\alpha = 0$ or $\beta = 0$ are possible, as they consider $\Phi = 1$ as an acceptable eigenvalue with respect to convergence). However, less strict conditions may be achieved, as shown, for example, by Eyre & Milton who proved convergence for $\alpha = \beta = 2$.

4.2 Optimizing convergence conditions

Combining (29) and (15) one obtains:

$$\epsilon_i \leq \mathcal{R}^i \cdot \epsilon_0 \leq \mathcal{R}_u^i \cdot \epsilon_0 . \quad (33)$$

In order to improve the convergence rate of the scheme, one can endeavor to minimize the upper bound \mathcal{R}_u of its spectral radius. But some remarks are in order:

- when minimizing \mathcal{R}_u , one does not take into account any geometrical information about the microstructure, thus this does not guarantee the optimal convergence rate for a particular microstructure,
- relation (15) only shows how the error ϵ_i *asymptotically* depends on the iterate i . When i is large enough, and thus when the convergence is governed by the largest eigenvalue of the operator, one has:

$$\log \epsilon_i \simeq i \times \log \mathcal{R} . \quad (34)$$

But for small or moderate values of i , smaller eigenvalues can still play an important role, and the corresponding part of the curves ϵ_i vs i may greatly vary from one case to the other.

This will be later illustrated in section (7.2).

Although not compulsory, a reasonable choice is to take: $k^0 \geq 0$ and $\mu^0 \geq 0$ which implies that α and β must be strictly positive to fulfil the condition in (32). This assumption will be made in the remainder of this section.

Let us denote the particular values of Φ :

$$\begin{aligned}
\Phi_{\alpha, k_{min}} &= 1 - \alpha \frac{k_{min}}{k_{min} + k^0} \\
\Phi_{\alpha, k_{max}} &= 1 - \alpha \frac{k_{max}}{k_{max} + k^0} \\
\Phi_{\beta, k_{min}} &= 1 - \beta \frac{k^0}{k_{min} + k^0} \\
\Phi_{\beta, k_{max}} &= 1 - \beta \frac{k^0}{k_{max} + k^0} \\
\Phi_{\alpha, \mu_{min}} &= 1 - \alpha \frac{\mu_{min}}{\mu_{min} + \mu^0} \\
\Phi_{\alpha, \mu_{max}} &= 1 - \alpha \frac{\mu_{max}}{\mu_{max} + \mu^0} \\
\Phi_{\beta, \mu_{min}} &= 1 - \beta \frac{\mu^0}{\mu_{min} + \mu^0} \\
\Phi_{\beta, \mu_{max}} &= 1 - \beta \frac{\mu^0}{\mu_{max} + \mu^0} .
\end{aligned} \tag{35}$$

Relation (28) can be simplified into:

$$\left\{ \begin{array}{l} \Phi \geq \min \left(1 - \alpha \frac{k_{max}}{k_{max} + k^0}, 1 - \beta \frac{k^0}{k_{min} + k^0}, 1 - \alpha \frac{\mu_{max}}{\mu_{max} + \mu^0}, 1 - \beta \frac{\mu^0}{\mu_{min} + \mu^0} \right) \\ \text{and} \\ \Phi \leq \max \left(1 - \alpha \frac{k_{min}}{k_{min} + k^0}, 1 - \beta \frac{k^0}{k_{max} + k^0}, 1 - \alpha \frac{\mu_{min}}{\mu_{min} + \mu^0}, 1 - \beta \frac{\mu^0}{\mu_{max} + \mu^0} \right) . \end{array} \right. \tag{36}$$

or:

$$\left\{ \begin{array}{l} \Phi \geq \min (\Phi_{\alpha, k_{max}}, \Phi_{\beta, k_{min}}, \Phi_{\alpha, \mu_{max}}, \Phi_{\beta, \mu_{min}}) \\ \text{and} \\ \Phi \leq \max (\Phi_{\alpha, k_{min}}, \Phi_{\beta, k_{max}}, \Phi_{\alpha, \mu_{min}}, \Phi_{\beta, \mu_{max}}) . \end{array} \right. \tag{37}$$

and (30) can be written as:

$$\mathcal{R}_u = \max(|\Phi_{\alpha, k_{min}}|, |\Phi_{\alpha, k_{max}}|, |\Phi_{\beta, k_{min}}|, |\Phi_{\beta, k_{max}}|, |\Phi_{\alpha, \mu_{min}}|, |\Phi_{\alpha, \mu_{max}}|, |\Phi_{\beta, \mu_{min}}|, |\Phi_{\beta, \mu_{max}}|) \tag{38}$$

The values $\Phi_{\alpha, k_{min}}$, $\Phi_{\beta, k_{min}}$, $\Phi_{\alpha, k_{max}}$, $\Phi_{\beta, k_{max}}$ are plotted in figure 1 (respectively figure 2) as functions of k_0 in the case where $k_{min} = 1$, $k_{max} = 2$ with $\alpha = \beta = 1$ (respectively, $\alpha = \beta = 2$).

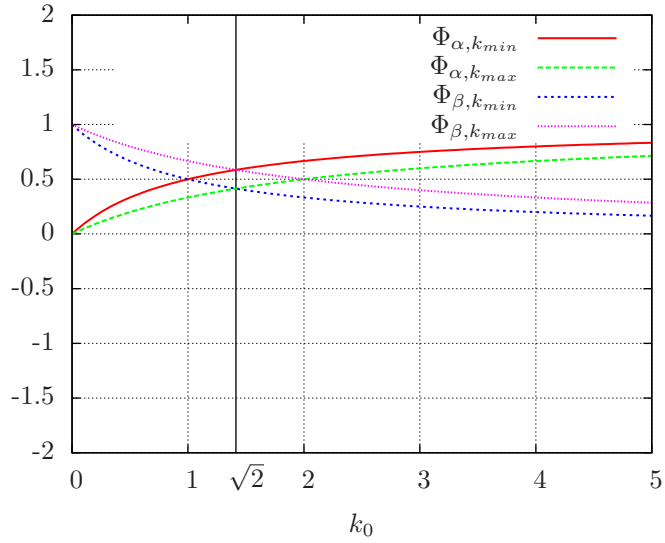


Figure 1: Bounds of the eigenvalues vs k_0 for $k_{min} = 1$, $k_{max} = 2$ and with $\alpha = \beta = 1$ (augmented Lagrangian scheme). The upper bound \mathcal{R}_u of the spectral radius is deduced from the envelope of the curves on the diagram.

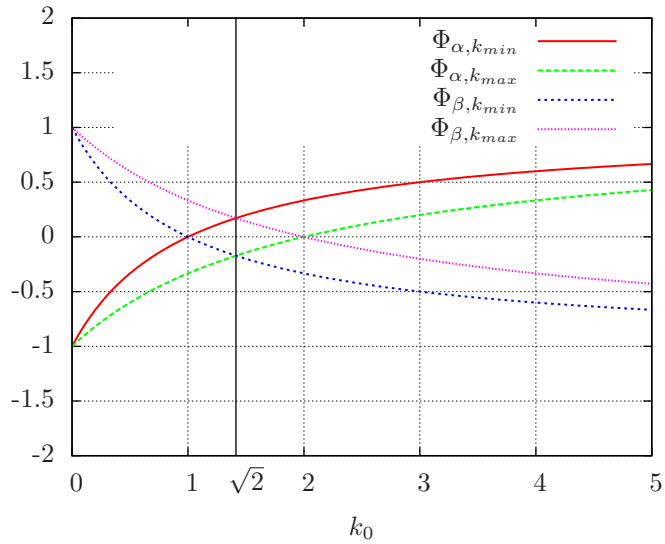


Figure 2: Bounds of the eigenvalues vs k_0 for $k_{min} = 1$, $k_{max} = 2$ and with $\alpha = \beta = 2$ (scheme of Eyre-Milton)

When $\alpha = \beta$, the minimum of the upper envelope of this set of curves, corresponding to the case where $\Phi_{\alpha, k_{min}} = \Phi_{\beta, k_{max}}$ and the maximum of the lower envelope, corresponding to the case where $\Phi_{\beta, k_{min}} = \Phi_{\alpha, k_{max}}$ take place at the same value of k_0 , which can easily be determined as:

$$k^0 = \sqrt{k_{min} \cdot k_{max}}$$

Similarly, the set of curves giving $\Phi_{\alpha, \mu_{min}}$, $\Phi_{\beta, \mu_{min}}$, $\Phi_{\alpha, \mu_{max}}$, $\Phi_{\beta, \mu_{max}}$ as a function of μ_0 has its upper envelope minimized when $\Phi_{\alpha, \mu_{min}} = \Phi_{\beta, \mu_{max}}$ and its lower envelope maximized

when $\Phi_{\beta, \mu_{min}} = \Phi_{\alpha, \mu_{max}}$. When $\alpha = \beta$, these two cases appear when:

$$\mu^0 = \sqrt{\mu_{min} \cdot \mu_{max}}$$

Thus, when $\alpha = \beta$, \mathcal{R}_u is minimized when:

$$\begin{cases} k^0 = \sqrt{k_{min} \cdot k_{max}} \\ \mu^0 = \sqrt{\mu_{min} \cdot \mu_{max}} \end{cases} \quad (39)$$

This result generalizes what had already been proved in [8] and [13] for the special case when $\alpha = \beta = 2$ (scheme of Eyre-Milton).

With this specific choice of k^0 and μ^0 , the eigenvalues Φ are bounded by:

$$\begin{cases} \Phi \geq \min \left(1 - \alpha \frac{1}{1 + \sqrt{k_{min}/k_{max}}}, 1 - \alpha \frac{1}{1 + \sqrt{\mu_{min}/\mu_{max}}} \right) \\ \text{and} \\ \Phi \leq \max \left(1 - \alpha \frac{1}{1 + \sqrt{k_{max}/k_{min}}}, 1 - \alpha \frac{1}{1 + \sqrt{\mu_{max}/\mu_{min}}} \right) \end{cases} \quad (40)$$

Defining the contrast \mathcal{C} between the phases as:

$$\mathcal{C} = \max(k_{max}/k_{min}, \mu_{max}/\mu_{min}) \quad (41)$$

(40) becomes:

$$1 - \alpha \frac{1}{1 + 1/\sqrt{\mathcal{C}}} \leq \Phi \leq 1 - \alpha \frac{1}{1 + \sqrt{\mathcal{C}}}.$$

In other words, the upper bound \mathcal{R}_u of the spectral radius verifies:

$$\mathcal{R}_u = \max \left(\left| 1 - \alpha \frac{1}{1 + \sqrt{\mathcal{C}}} \right|, \left| 1 - \alpha \frac{1}{1 + 1/\sqrt{\mathcal{C}}} \right| \right), \quad (42)$$

which can even be simplified into:

$$\begin{cases} \mathcal{R}_u = 1 - \alpha \frac{1}{1 + \sqrt{\mathcal{C}}}, & \text{if } 0 < \alpha \leq 2 \\ \mathcal{R}_u = -1 + \alpha \frac{1}{1 + 1/\sqrt{\mathcal{C}}}, & \text{if } \alpha > 2 \end{cases}, \quad (43)$$

taking into account that $\mathcal{C} \geq 1$.

Thus, to guarantee that $\mathcal{R}_u < 1$ for *all finite contrasts*, α must be lower than or equal to 2.

Concerning infinite contrasts (material with void or rigid phases), the upper bound \mathcal{R}_u exhibited here tends to 1, the convergence obtained numerically as shown in [10] still has to be proven.

In the case of the augmented Lagrangian scheme, $\alpha = \beta = 1$ and one can calculate the optimal bound for the spectral radius as:

$$\mathcal{R}_u = \frac{\sqrt{\mathcal{C}}}{\sqrt{\mathcal{C}} + 1}$$

In the case of the scheme of Eyre-Milton, $\alpha = \beta = 2$ one has:

$$\mathcal{R}_u = \frac{\sqrt{\mathcal{C}} - 1}{\sqrt{\mathcal{C}} + 1}$$

(which corroborates [8], [13] and [5]).

4.3 “Best” scheme

Keeping in mind the preliminary remarks in section (4.2), let us find the values for α and β which minimize \mathcal{R}_u . For simplicity, we restrict ourselves to the case when $\alpha = \beta$ although a more general proof is presented in appendix B.

To minimize \mathcal{R}_u , relation (43) directly leads to the choice:

$$\alpha = \beta = 2 ,$$

which corresponds to the scheme of Eyre-Milton.

4.4 Rate of convergence for high contrast between the phases

From (33) it comes that the number N of iterations necessary to reach a given accuracy \mathcal{E} , satisfies:

$$N \leq \log(\mathcal{E}/\epsilon_0)/\log(\mathcal{R}_u) , \quad (44)$$

when $\alpha \leq 2$ and when the contrast is high:

$$N \leq \log(\epsilon_0/\mathcal{E}) \frac{1}{\alpha} \sqrt{\mathcal{C}} . \quad (45)$$

This result implies that the scheme of Eyre-Milton should be, more or less, two times faster than the augmented Lagrangian scheme, when the contrast between the phases is large.

For comparison, the basic scheme would reach the error tolerance \mathcal{E} in N iterations with:

$$N \leq \frac{1}{2} \log(\epsilon_0/\mathcal{E}) \mathcal{C} . \quad (46)$$

5 Convergence test

The convergence test at each iteration i consists in comparing the deviations from equilibrium, from compatibility and from the prescribed loading conditions with a prescribed tolerance.

The convergence tests which have been chosen are the following:

- criterion on the equilibrium of the stress

The deviation from equilibrium at iterate i can be evaluated using the L2-norm of the divergence of the stress defined as:

$$\|\text{div}(\boldsymbol{\sigma}^i)\|_2 = \sqrt{\frac{1}{V} \int_V |\text{div}(\boldsymbol{\sigma}^i)|^2 dx} ,$$

where $|\cdot|$ denotes the euclidan norm of a vector.

Invoking Parseval’s theorem, it can be easily evaluated in the Fourier space as:

$$\|\text{div}(\boldsymbol{\sigma}^i)\|_2 = \sqrt{\sum_{\boldsymbol{\xi}} |\boldsymbol{\xi} \cdot \widehat{\boldsymbol{\sigma}}^i(\boldsymbol{\xi})|^2}$$

and it may be normalized by the macroscopic stress to make it insensitive to a linear factor on the prescribed strain:

$$\epsilon_{equilibrium} = \frac{\|\text{div}(\boldsymbol{\sigma}^i)\|_2}{\|\langle \boldsymbol{\sigma}^i \rangle\|} = \frac{\sqrt{\sum_{\boldsymbol{\xi}} |\boldsymbol{\xi} \cdot \widehat{\boldsymbol{\sigma}}^i(\boldsymbol{\xi})|^2}}{\|\widehat{\boldsymbol{\sigma}}^i(\mathbf{0})\|} , \quad (47)$$

where $\|\cdot\|$ denotes the Frobenius norm of the second order tensor, i.e. the square root of the sum of the squares of its components: $\|\boldsymbol{\sigma}\| = \sqrt{\sum_{i,j} |\sigma_{ij}|^2}$.

- criterion on the compatibility equations

There are six compatibility relations to be satisfied:

$$\begin{aligned}
\frac{\partial^2 e_{11}}{\partial x_2^2} + \frac{\partial^2 e_{22}}{\partial x_1^2} - 2 \frac{\partial^2 e_{12}}{\partial x_1 \partial x_2} &= 0 \\
\frac{\partial^2 e_{22}}{\partial x_3^2} + \frac{\partial^2 e_{33}}{\partial x_2^2} - 2 \frac{\partial^2 e_{23}}{\partial x_2 \partial x_3} &= 0 \\
\frac{\partial^2 e_{33}}{\partial x_1^2} + \frac{\partial^2 e_{11}}{\partial x_3^2} - 2 \frac{\partial^2 e_{13}}{\partial x_3 \partial x_1} &= 0 \\
\frac{\partial^2 e_{11}}{\partial x_2 \partial x_3} - \frac{\partial^2 e_{13}}{\partial x_1 \partial x_2} - \frac{\partial^2 e_{12}}{\partial x_1 \partial x_3} + \frac{\partial^2 e_{23}}{\partial x_1 \partial x_1} &= 0 \\
\frac{\partial^2 e_{22}}{\partial x_3 \partial x_1} - \frac{\partial^2 e_{12}}{\partial x_2 \partial x_3} - \frac{\partial^2 e_{23}}{\partial x_2 \partial x_1} + \frac{\partial^2 e_{13}}{\partial x_2 \partial x_2} &= 0 \\
\frac{\partial^2 e_{33}}{\partial x_1 \partial x_2} - \frac{\partial^2 e_{23}}{\partial x_3 \partial x_1} - \frac{\partial^2 e_{13}}{\partial x_3 \partial x_2} + \frac{\partial^2 e_{12}}{\partial x_3 \partial x_3} &= 0 .
\end{aligned}$$

The deviation from compatibility can be easily evaluated in Fourier space by:

$$\epsilon_{compatibility} = \frac{\max_{\boldsymbol{\xi}} (\max_{j=1, \dots, 6} (|\widehat{c}_j(\boldsymbol{\xi})|))}{\sqrt{\sum_{\boldsymbol{\xi}} \widehat{e}_{ij}(\boldsymbol{\xi}) : \widehat{e}_{ij}^*(\boldsymbol{\xi})}} \quad (48)$$

with:

$$\begin{aligned}
\widehat{c}_1(\boldsymbol{\xi}) &= -\xi_2 \xi_2 \widehat{e}_{11}(\boldsymbol{\xi}) - \xi_1 \xi_1 \widehat{e}_{22}(\boldsymbol{\xi}) + 2 \xi_1 \xi_2 \widehat{e}_{12}(\boldsymbol{\xi}) \\
\widehat{c}_2(\boldsymbol{\xi}) &= -\xi_3 \xi_3 \widehat{e}_{22}(\boldsymbol{\xi}) - \xi_2 \xi_2 \widehat{e}_{33}(\boldsymbol{\xi}) + 2 \xi_2 \xi_3 \widehat{e}_{23}(\boldsymbol{\xi}) \\
\widehat{c}_3(\boldsymbol{\xi}) &= -\xi_1 \xi_1 \widehat{e}_{33}(\boldsymbol{\xi}) - \xi_3 \xi_3 \widehat{e}_{11}(\boldsymbol{\xi}) + 2 \xi_3 \xi_1 \widehat{e}_{13}(\boldsymbol{\xi}) \\
\widehat{c}_4(\boldsymbol{\xi}) &= -\xi_2 \xi_3 \widehat{e}_{11}(\boldsymbol{\xi}) + \xi_1 \xi_2 \widehat{e}_{13}(\boldsymbol{\xi}) + \xi_1 \xi_3 \widehat{e}_{12}(\boldsymbol{\xi}) - \xi_1 \xi_1 \widehat{e}_{23}(\boldsymbol{\xi}) \\
\widehat{c}_5(\boldsymbol{\xi}) &= -\xi_3 \xi_1 \widehat{e}_{22}(\boldsymbol{\xi}) + \xi_2 \xi_3 \widehat{e}_{12}(\boldsymbol{\xi}) + \xi_2 \xi_1 \widehat{e}_{23}(\boldsymbol{\xi}) - \xi_2 \xi_2 \widehat{e}_{13}(\boldsymbol{\xi}) \\
\widehat{c}_6(\boldsymbol{\xi}) &= -\xi_1 \xi_2 \widehat{e}_{33}(\boldsymbol{\xi}) + \xi_3 \xi_1 \widehat{e}_{23}(\boldsymbol{\xi}) + \xi_3 \xi_2 \widehat{e}_{13}(\boldsymbol{\xi}) - \xi_3 \xi_3 \widehat{e}_{12}(\boldsymbol{\xi}) .
\end{aligned}$$

- criterion on the loading conditions:

In the case of prescribed macroscopic strain, the convergence on loading conditions can be tested using:

$$\epsilon_{loading} = \frac{\| \langle \mathbf{e} \rangle - \mathbf{E} \|}{\| \mathbf{E} \|} = \frac{\sqrt{\langle e_{ij} \rangle - E_{ij} : \langle e_{ij} \rangle - E_{ij}}}{\sqrt{E_{ij} : E_{ij}}} . \quad (49)$$

Remark: similar tests can be obtained with other loading conditions.

The equilibrium test proposed by Monchiet and Bonned in [11] involves an error estimate defined in the Fourier space by:

$$\epsilon_{MB}^i = \frac{\| \mathbb{P} * \boldsymbol{\sigma}^i \|_2}{\| \boldsymbol{\sigma}^i \|_2} = \frac{\sqrt{\sum_{\boldsymbol{\xi}} \| \mathbb{P}(\boldsymbol{\xi}) : \widehat{\boldsymbol{\sigma}}^i(\boldsymbol{\xi}) \|^2}}{\sqrt{\sum_{\boldsymbol{\xi}} \| \widehat{\boldsymbol{\sigma}}^i(\boldsymbol{\xi}) \|^2}} \quad (50)$$

where \mathbb{P} is a fourth-order tensor, which is equal to the Green operator Γ_0 for the choice of the reference medium $c_0 = \mathbb{I}$ (in other words when the Lamé coefficients are $\lambda_0 = 0$ and $\mu_0 = 0.5$), i.e.:

$$\widehat{\mathbb{P}}_{ijkl}(\boldsymbol{\xi}) = \frac{1}{2|\boldsymbol{\xi}|^2} \{ \delta_{ik} \xi_j \xi_l + \delta_{il} \xi_j \xi_k + \delta_{jk} \xi_i \xi_l + \delta_{jl} \xi_i \xi_k \} - \frac{\xi_i \xi_j \xi_k \xi_l}{|\boldsymbol{\xi}|^4} . \quad (51)$$

It is worth noting the relation between this criterion and the previous ones:

$$\forall \boldsymbol{\xi}, \quad \left\| \widehat{\mathbb{P}}(\boldsymbol{\xi}) : \widehat{\boldsymbol{\sigma}}^i(\boldsymbol{\xi}) \right\| = \frac{1}{|\boldsymbol{\xi}|} \sqrt{2 \left| \widehat{\text{div}} \boldsymbol{\sigma}^i(\boldsymbol{\xi}) \right|^2 - \left| \frac{\boldsymbol{\xi}}{|\boldsymbol{\xi}|} \cdot \widehat{\text{div}} \boldsymbol{\sigma}^i(\boldsymbol{\xi}) \right|^2} . \quad (52)$$

Using the decomposition into subspaces $\widehat{\text{div}} \boldsymbol{\sigma}^i(\boldsymbol{\xi}) = \alpha(\boldsymbol{\xi})\boldsymbol{\xi} + B(\boldsymbol{\xi})$ with α scalar and $\boldsymbol{\xi} \cdot B(\boldsymbol{\xi}) = 0$, it comes that

$$\forall \boldsymbol{\xi}, \quad \left\| \widehat{\mathbb{P}}(\boldsymbol{\xi}) : \widehat{\boldsymbol{\sigma}}^i(\boldsymbol{\xi}) \right\| = \frac{1}{2\pi} \sqrt{2 \frac{|B(\boldsymbol{\xi})|^2}{|\boldsymbol{\xi}|^2} + |\alpha|^2} \quad (53)$$

and the error ϵ_{MB} vanishes if and only if both components are zero for all frequencies, i.e. the divergence of the local stress vanishes too. However, $\epsilon_{\text{equilibrium}}$ defined in (47) and ϵ_{MB} defined in (50) do not behave in the same manner in the vicinity of 0. Roughly speaking $\epsilon_{\text{equilibrium}}$ behaves as $|\widehat{\text{div}} \boldsymbol{\sigma}(\boldsymbol{\xi})|$ and ϵ_{MB} behaves as $\frac{|\widehat{\text{div}} \boldsymbol{\sigma}(\boldsymbol{\xi})|}{|\boldsymbol{\xi}|}$, and thus the latter is less sensitive to high frequency errors than $\epsilon_{\text{equilibrium}}$. This can lead to great consequences concerning the accuracy of a computation as is illustrated in section 7.5.

6 Implementation of the algorithms

The following algorithm has been used in this study to apply the polarization-based scheme in the case of prescribed macroscopic strain \mathbf{E} :

$$\begin{aligned} \text{Iterate } i+1 : & \text{ given } \mathbf{e}^i \text{ and } \boldsymbol{\sigma}^i, \\ (a) \quad & \mathbf{s}_a^i(\mathbf{x}) = \boldsymbol{\sigma}^i(\mathbf{x}) + (1-\beta) \cdot \mathbf{c}^0 : \mathbf{e}^i(\mathbf{x}) \\ & \mathbf{s}_b^i(\mathbf{x}) = \alpha \cdot \boldsymbol{\sigma}^i(\mathbf{x}) - \beta \cdot \mathbf{c}^0 : \mathbf{e}^i(\mathbf{x}) \\ (b) \quad & \widehat{\mathbf{s}}_b^i = \mathcal{F}(\mathbf{s}_b^i) \\ (c) \quad & \widehat{\mathbf{e}}_b^i(\boldsymbol{\xi}) = -\widehat{\Gamma}^0(\boldsymbol{\xi}) : \widehat{\mathbf{s}}_b^i(\boldsymbol{\xi}) \quad \forall \boldsymbol{\xi} \neq \mathbf{0}, \quad \widehat{\mathbf{e}}_b^i(\mathbf{0}) = \beta \cdot \mathbf{E} \\ (d) \quad & \mathbf{e}_b^i(\mathbf{x}) = \mathcal{F}(\widehat{\mathbf{e}}_b^i(\boldsymbol{\xi})) \\ (e) \quad & \mathbf{e}^{i+1}(\mathbf{x}) = (\mathbf{c}(\mathbf{x}) + \mathbf{c}^0)^{-1} : (\mathbf{s}_a^i(\mathbf{x}) + \mathbf{c}^0 : \mathbf{e}_b^i(\mathbf{x})) \\ (f) \quad & \boldsymbol{\sigma}^{i+1}(\mathbf{x}) = \mathbf{c}(\mathbf{x}) : \mathbf{e}^{i+1}(\mathbf{x}) \\ (g) \quad & \text{convergence tests.} \end{aligned} \quad (54)$$

(55)

One can take advantage of this algorithm to save memory space, since only two arrays are necessary to store \mathbf{e}^i , \mathbf{s}_b^i , $\widehat{\mathbf{s}}_b^i$, $\widehat{\mathbf{e}}_b^i$ and \mathbf{e}_b^i in one hand, $\boldsymbol{\sigma}^i$ and \mathbf{s}_a^i in the other hand.

Similar algorithms can be constructed for prescribed macroscopic stress (for example for creep loading) or prescribed direction of macroscopic stress (for example for simple tension).

All the computations reported in this article have been performed with the FFT-based code CraFT, which is freely available under the terms of the Cecill B license agreement at <http://craft.lma.cnrs-mrs.fr/>.

7 numerical simulations

7.1 Composite material

Numerical simulations have been carried out to check the theoretical results presented above. The material which has been chosen is composed of two phases with equal bulk and shear moduli:

$$k_{\text{inclusion}} = \mu_{\text{inclusion}} \text{ and } k_{\text{matrix}} = \mu_{\text{matrix}}$$

in other words, the Poisson coefficient of both phases satisfies:

$$\nu = 0.125 .$$

So the contributions of the bulk moduli and of the shear moduli in the upper bound \mathcal{R}_u of the spectral radius are equal (see (30)).

3 different microstructures were studied:

- a microstructure consisting of a matrix reinforced by long fibers arranged in a square array with circular section and with a volume fraction of 50.26%. This microstructure has been discretized into a 2D image composed of 243×243 pixels (see (3a))
- a microstructure consisting of a matrix reinforced by spherical inclusions arranged in a cubic array with a volume fraction of 26.85%, discretized into a 3D image composed of $45 \times 45 \times 45$ voxels and containing 1 inclusion (see (3b)).
- a microstructure consisting of a matrix reinforced by impenetrable spherical inclusions with a volume fraction of 26.85%, discretized into a 3D image composed of $81 \times 81 \times 81$ voxels and containing 6 inclusions (see (3c)).

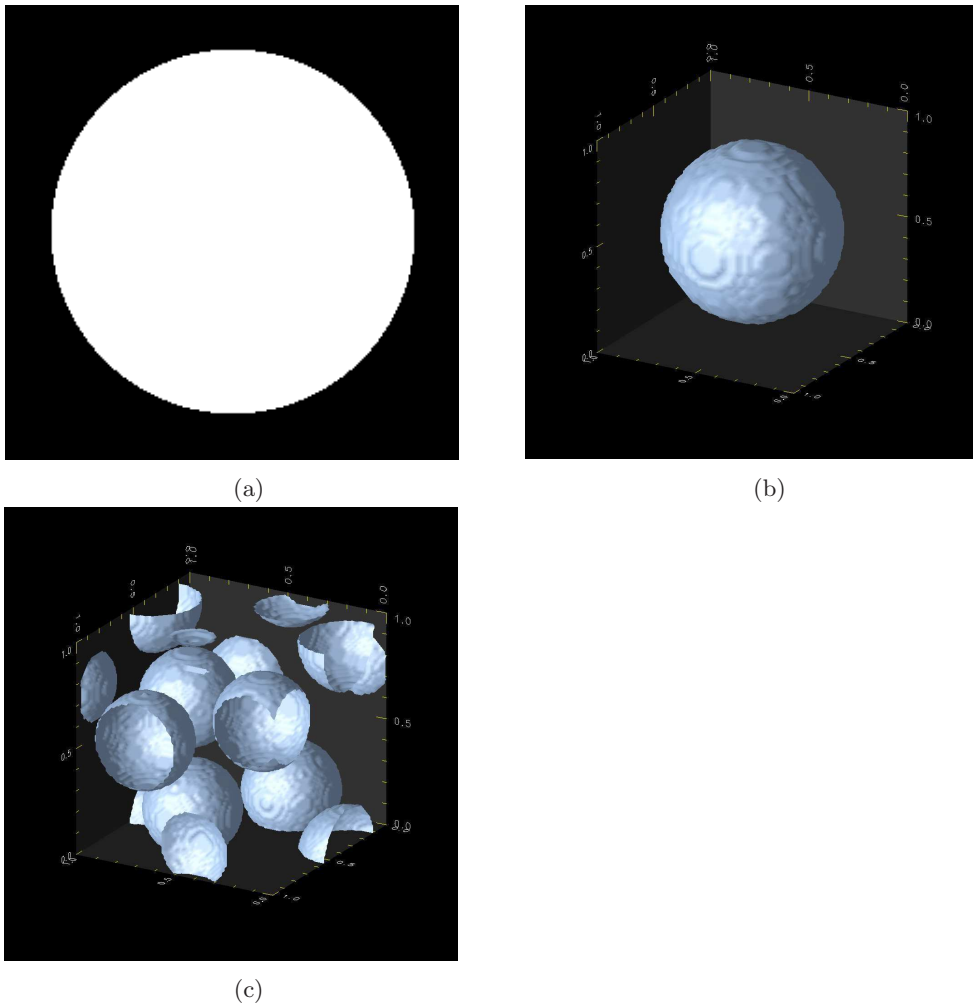


Figure 3: (a) : 2D image of a periodic long fiber reinforced material with circular sections (b) : 3D image of a periodic material with 1 spherical inclusion (c) : 3D image of a periodic material with 6 spherical inclusions

For each configuration, the following macroscopic strain was prescribed: $\mathbf{E}_{11} = 1$, $\mathbf{E}_{ij \neq 11} = 0$

All the computations were performed with a required precision of 10^{-10} for the equilibrium and the compatibility conditions (i.e. the algorithms had to iterate until the error on equilibrium and compatibility reached values lower than 10^{-10}).

The same accuracy was required for the error on loading conditions.

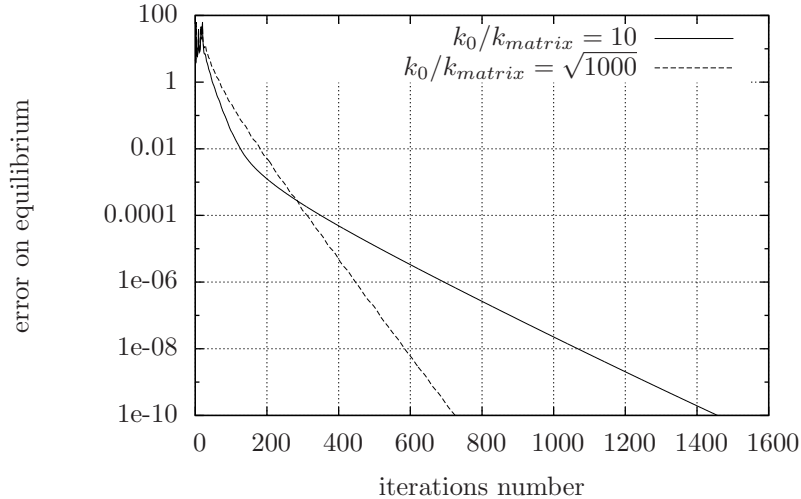


Figure 4: Error on equilibrium at each iteration of a computation on the microstructure (3b) (spherical inclusions) using the augmented Lagrangian scheme. The contrast on mechanical properties of the two phase is 1000. Two different values for the reference material were chosen: $k_0/k_{matrix} = \mu_0/\mu_{matrix} = 10$ and $k_0/k_{matrix} = \mu_0/\mu_{matrix} = \sqrt{1000}$

7.2 Error as function of number of iterations

Typical curves of error as a function of the number of iterates are presented in figure (4). They were obtained using the augmented Lagrangian scheme applied to microstructure (3b) (1 spherical inclusion) with a contrast of 1000 between the mechanical properties of the two phases:

$$\mathcal{C} = k_{inclusion}/k_{matrix} = \mu_{inclusion}/\mu_{matrix} = 1000 .$$

The 2 curves differ by the reference material which has been chosen. After a transient phase during the first iterations, the error is decreasing and its logarithm is asymptotically linear, with a slope expected to be the logarithm of the spectral radius: $\log(\mathcal{R})$ as predicted by (34).

Comparing the two curves in figure (4), it is worth noting that the error is lower in the computation with the reference medium $k_0/k_{matrix} = 10$ than in the computation with $k_0/k_{matrix} = \sqrt{1000}$ during the first iterations (before iterate $i \simeq 290$ when the error reaches about $2.8 \cdot 10^{-4}$). Then, the error decreases much more rapidly when $k_0/k_{matrix} = \sqrt{1000}$, as expected according to (39). This illustrates the fact that the comparison between the different schemes and between the choices of reference medium made in section (4), makes sense when the tolerance is small enough (i.e. when the number of iterations is large enough) so that the part of the error due to the spectral radius (i.e. the largest eigenvalue of the iteration operator) prevails.

This is why all the computations presented in this section have been performed with a high precision (10^{-10}), which is much more than necessary for common computations.

Conversely, if one has to carry out a computation with a moderate precision, the reference medium found as optimum in section (4) may be not the fastest one. For example, to reach a precision of 10^{-2} in the cases presented in figure (4), one needs 183 iterations with “optimal” reference material such that $k_0/k_{matrix} = \sqrt{1000}$, but only 129 iterations with reference medium $k_0/k_{matrix} = 10$.

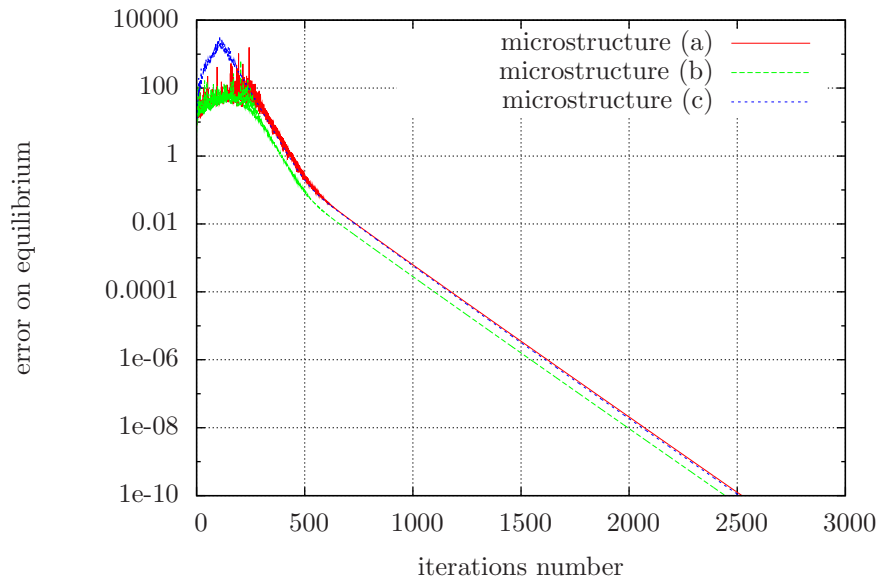


Figure 5: Error on equilibrium at each iteration of a computation on the microstructures (3a), (3b) and (3c) using the scheme of Eyre-Milton. The contrast on mechanical properties of the two phase is 10^4 . The reference material: $k_0/k_{matrix} = \mu_0/\mu_{matrix} = 50$

In figure 5, the error curves are plotted for computations made using the scheme of Eyre-Milton, with a contrast of 10^4 between the mechanical properties of the constituents, for the 3 different microstructures (3a), (3b) and (3c). As soon as the iterate is large enough (larger than 700 in this example), the 3 curves are linear and parallel, illustrating that the rate of convergence is governed by the spectral radius and does not seem to depend much on the microstructure.

In the attempt to make the figures clearer, they will be presented in the following sections for one microstructure only.

7.3 Optimal convergence

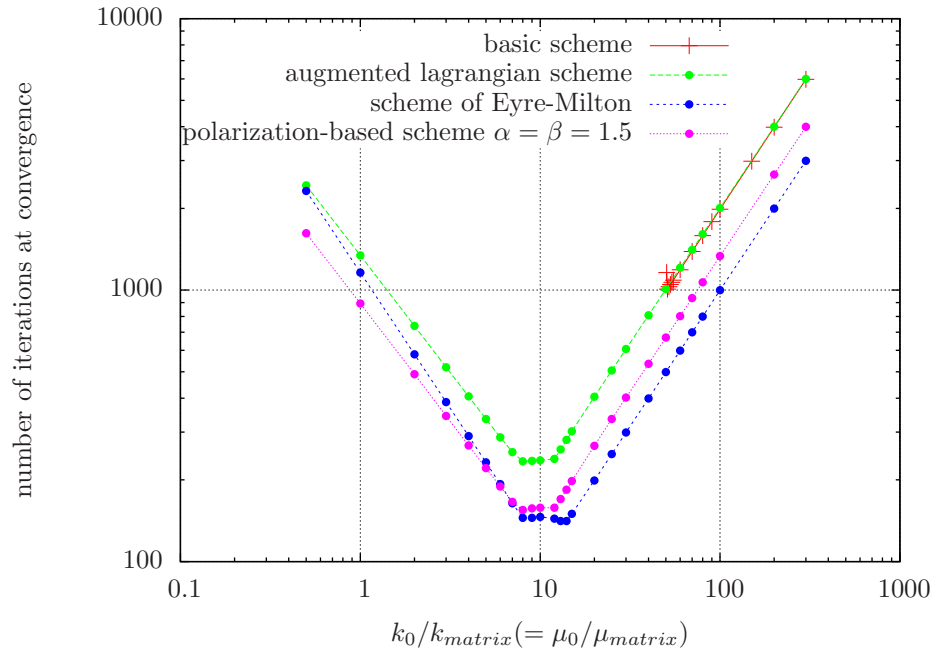


Figure 6: Number of iterations at convergence (tolerance = 10^{-10}) for different choices of reference material \mathbf{c}_0 , for a contrast of 100 between the mechanical properties of inclusion and matrix. Microstructure (3c)

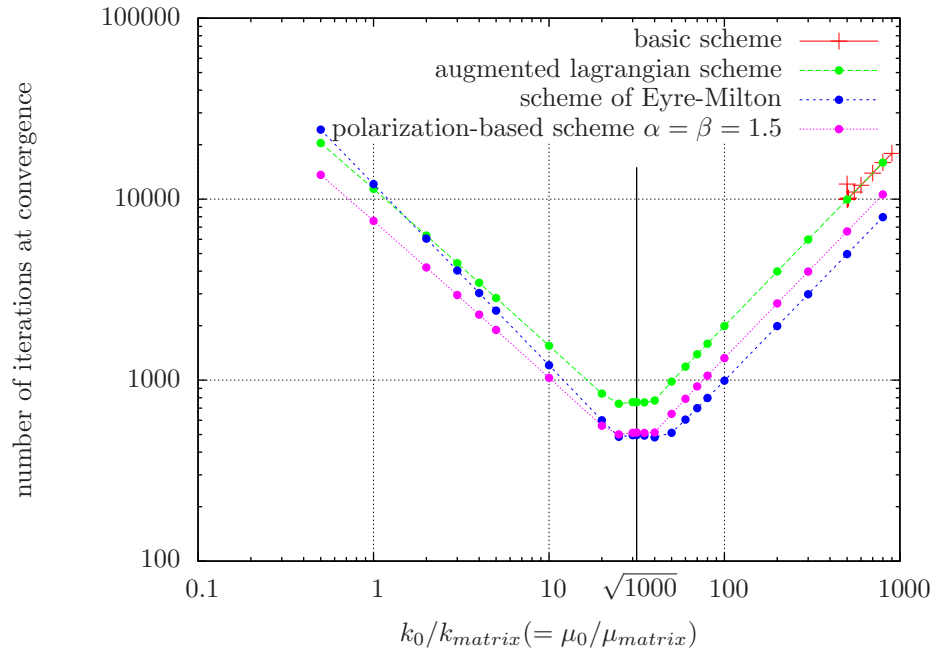


Figure 7: Number of iterations at convergence (tolerance = 10^{-10}) for different choices of reference material \mathbf{c}_0 , for a contrast of 1000 between the mechanical properties of inclusion and matrix. Microstructure (3c)

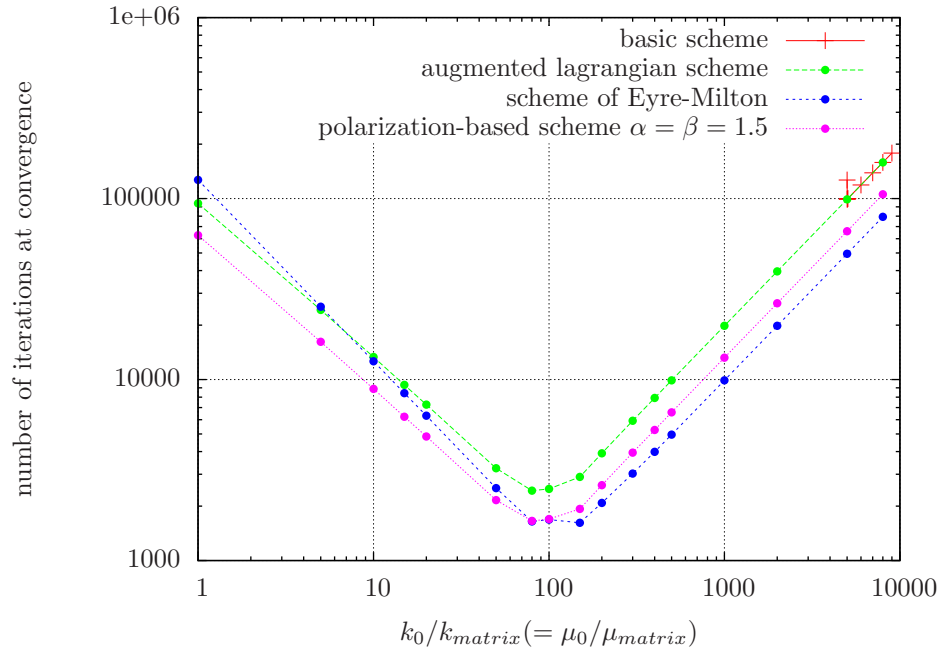


Figure 8: Number of iterations at convergence (tolerance = 10^{-10}) for different choices of reference material \mathbf{c}_0 , for a contrast of 10000 between the mechanical properties of inclusion and matrix. Microstructure (3c)

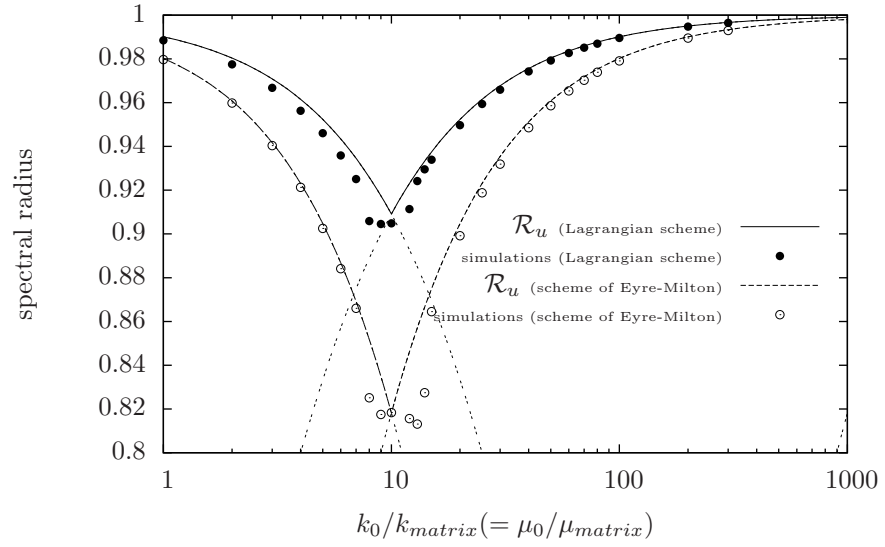


Figure 9: Comparison between the upper bound of the spectral radius \mathcal{R}_u given by (30) and the spectral radius estimated from numerical simulations. Microstructure (3c), contrast 100.

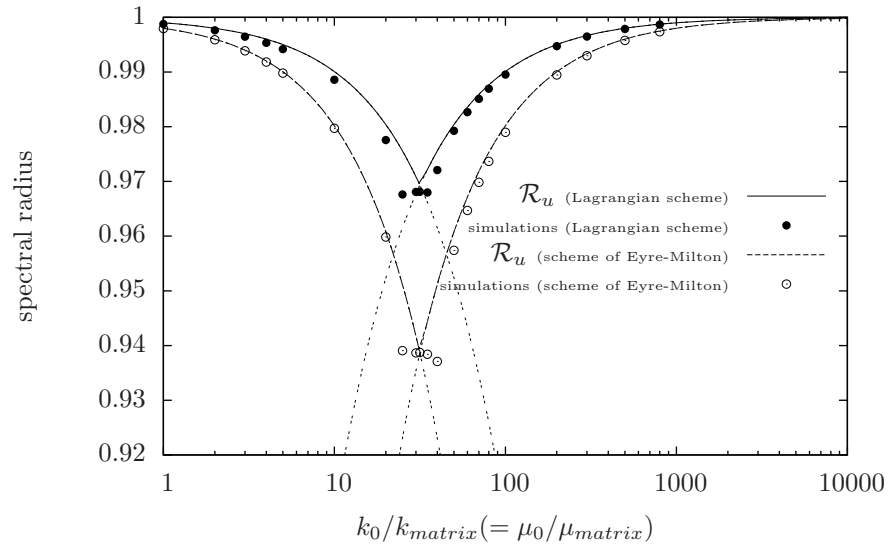


Figure 10: Comparison between the upper bound of the spectral radius \mathcal{R}_u given by (30) and the spectral radius estimated from numerical simulations. Microstructure (3c), contrast 1000.

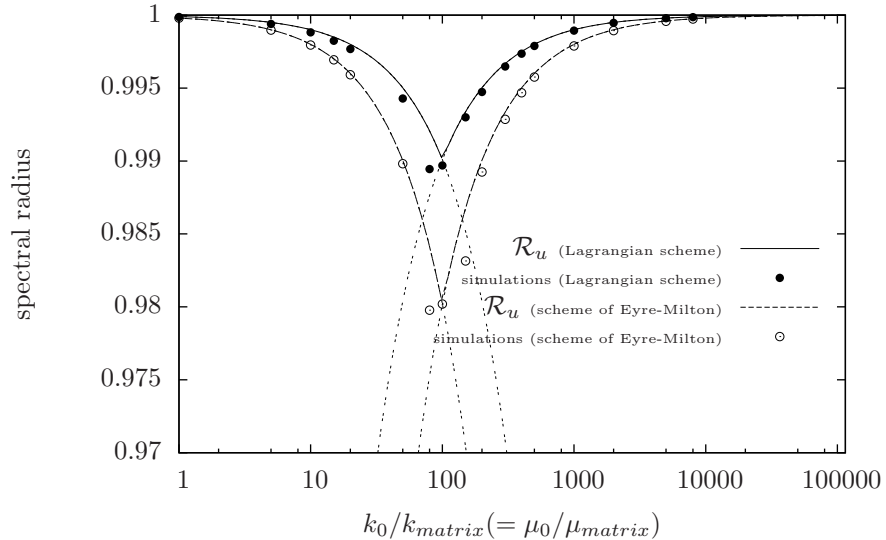


Figure 11: Comparison between the upper bound of the spectral radius \mathcal{R}_u given by (30) and the spectral radius estimated from numerical simulations. Microstructure (3c), contrast 10000.

Figure (6), (7) and (8) show the number of iterations necessary to reach the required precision, for various values of the reference medium, when the contrast between the mechanical properties of the 2 phases is, respectively, 100, 1000 and 10000.

Relation (39) appears to give a good prediction of the actual optimal reference media for all accelerated schemes (augmented Lagrangian scheme, scheme of Eyre-Milton, polarization-based scheme with parameters $\alpha = \beta = 1.5$). Furthermore, when the reference medium follows (39), the scheme of Eyre-Milton exhibits the fastest rate of convergence.

For each computation, the spectral radius \mathcal{R} of the iterative operator is estimated using equation (34). When the logarithm of the measured error as a function of the iterates number is sufficiently close to its linear asymptot (in practical, when the error is lower than 10^{-8}), it is fitted to a line by linear regression. The values of \mathcal{R} are plotted in figures (9), (10) and (11) for the different contrasts studied and compared to the curves giving \mathcal{R}_u obtained by relation (30).

As predicted, the \mathcal{R}_u curves provide an upper estimate of the actual rate of convergence of the different schemes and yet they approximate it very closely in all cases investigated, even without any supplementary information about the microstructure.

7.4 Rate of convergence vs contrast

Computations have been carried out to study the rate of convergence of the different schemes. The reference medium has been chosen following (5) for the basic scheme and (39) for the augmented Lagrangian scheme, for the scheme of Eyre-Milton and for the polarization-based scheme.

In this study, the contrast of the mechanical properties between the phases was defined as the ratio between inclusion and matrix:

$$\mathcal{C}' = k_{inclusion}/k_{matrix} (= \mu_{inclusion}/\mu_{matrix}), \quad (56)$$

which is simply related to the contrast \mathcal{C} defined in (41) by:

$$\mathcal{C} = \max(\mathcal{C}', \frac{1}{\mathcal{C}'}).$$

This distinguishes the cases where the inclusion is weaker or stronger than the matrix.

The number of iterations when the tolerance $\mathcal{E} = 10^{-10}$ has been reached, is plotted for different contrasts \mathcal{C}' on figure 12 with a log-log scale.

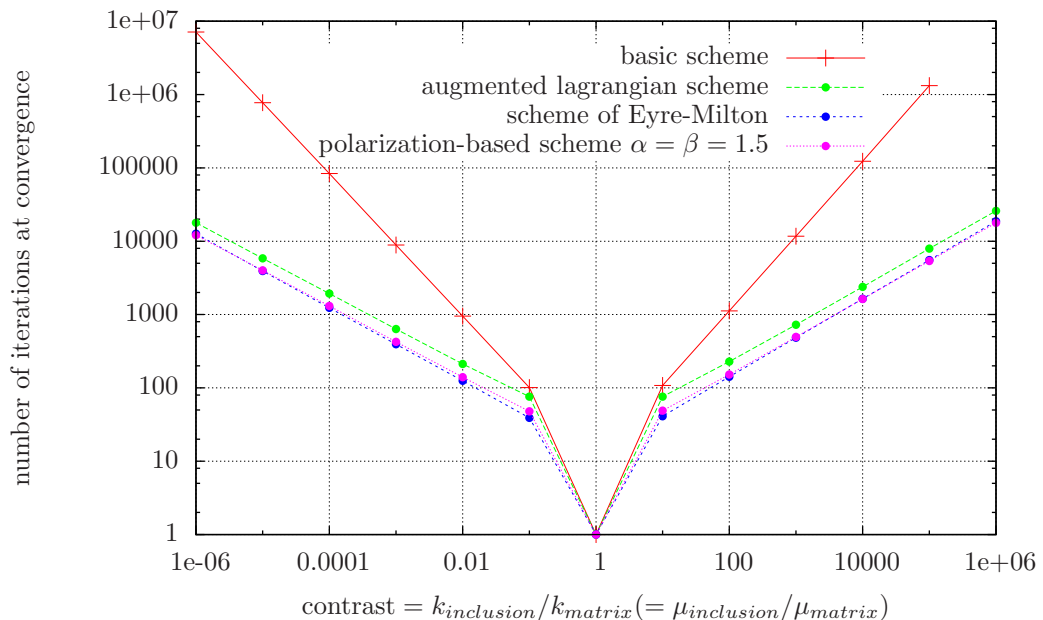


Figure 12: Number of iterations at convergence (tolerance = 10^{-10}) for different contrasts between the phases. Reference medium chosen according to (39) for the augmented Lagrangian scheme, for the scheme of Eyre-Milton and for the polarization-based scheme; reference medium chosen according to (5) for the basic scheme. Microstructure (3b)

As expected, the accelerated schemes are much faster than the basic scheme, even for small contrasts. The augmented Lagrangian scheme is slower than the polarization-based scheme with $\alpha = \beta = 1.5$ and than the scheme of Eyre-Milton. The Eyre-Milton is the fastest at least for moderate contrast ($10^{-3} \leq \mathcal{C}' \leq 10^{+3}$), but converges with almost the same rate than the polarization-based scheme with $\alpha = \beta = 1.5$ for higher contrasts.

Furthermore, the right terms of the inequalities (45) and (46) appears to predict fairly the actual rates of convergence. Indeed, fitting the logarithm of the data in figure 12 by linear regression, one obtains two parameters denoted by a and b in the relation:

$$N = a \cdot \mathcal{C}'^b \quad (57)$$

for the different schemes, summarized in table 1.

The exponent b for the basic scheme is close to 1 of relation (45), and the exponents of the accelerated schemes are close to 1/2 of relation (45).

On the other hand, the shape of the curves in figure 12 is very different from the shape of those obtained by Monchiet and Bonnet in [11] for similar simulations. Indeed, in our simulations, the logarithm of the number of iterations at convergence linearly depends on the logarithm of the contrast, whereas in the simulations by Monchiet and Bonnet the number of iterations seems to reach an asymptote for high contrasts. In our opinion, this great dissimilarity is due to the difference between the convergence tests used in this article, and the ones used in [11]. This will be discussed in the next section.

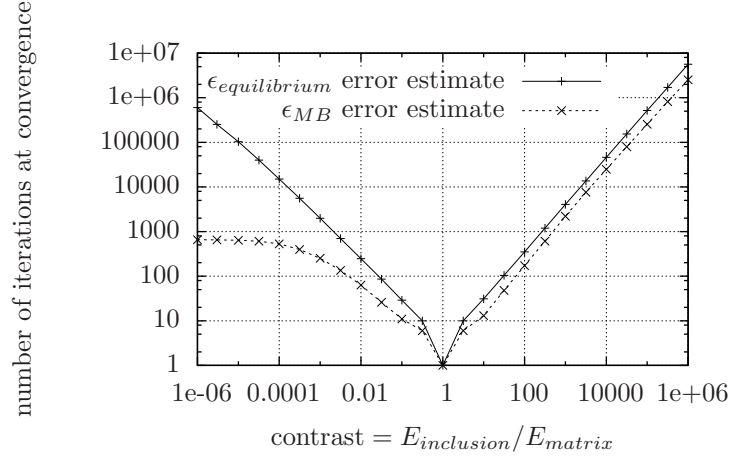
scheme	contrasts	a	b
basic	$\mathcal{C}' < 1$	10.92	0.970
basic	$\mathcal{C}' > 1$	10.18	1.022
augmented Lagrangian	$\mathcal{C}' < 1$	24.33	0.476
augmented Lagrangian	$\mathcal{C}' > 1$	22.46	0.509
polarization-based ($\alpha = \beta = 1.5$)	$\mathcal{C}' < 1$	15.48	0.482
polarization-based ($\alpha = \beta = 1.5$)	$\mathcal{C}' > 1$	14.57	0.513
Eyre-Milton	$\mathcal{C}' < 1$	12.33	0.501
Eyre-Milton	$\mathcal{C}' > 1$	12.22	0.531

Table 1: Parameters a and b of the relation between the number of iterations and the contrast: $N = a.C^b$, obtained by linear regression on the logarithms of data of figure 12.

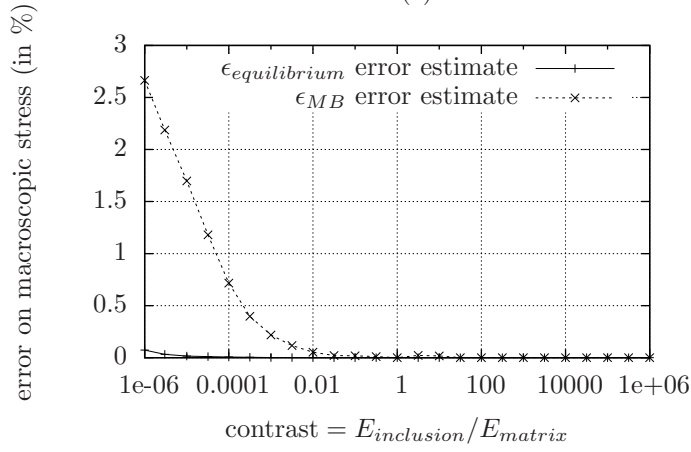
7.5 Error test

Numerical simulations using the “basic-scheme” have been performed for a prescribed macroscopic strain $E_{11} = 1$, $E_{ij \neq 11} = 0$, on a 2D microstructure similar to the one plotted in (3a) but with a different resolution: 128×128 pixels (instead of 243×243). Several contrasts between the mechanical properties of the two phases were tested: both phases had the same Poisson coefficient $\nu = 0.3$, the Young’s modulus of the matrix was fixed to 1 and the Young’s modulus of the inclusion varied from 10^{-6} to 10^{+6} . The required accuracy was $\mathcal{E} = 0.001$ with the error estimated either by criterion (47) or by criterion (50).

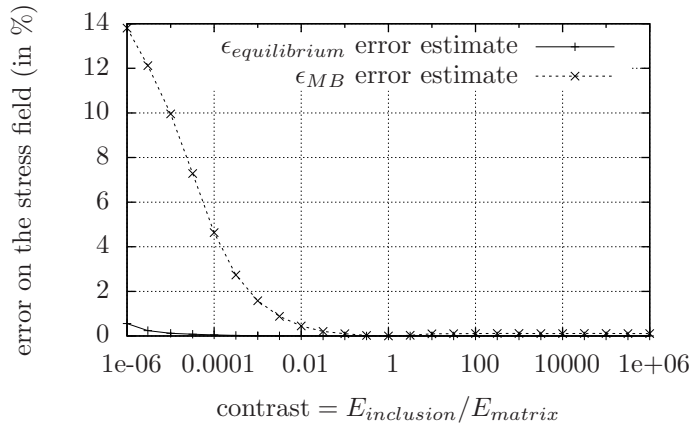
In the upper figure of Fig. 13, the number of iterations necessary to reach the required accuracy are plotted for several mechanical contrasts on a log-log scale. The two curves obtained with error criteria (47) or (50) are similar for ratios $E_{inclusion}/E_{matrix} > 1$ but greatly differ when $E_{inclusion}/E_{matrix} < 1$, the difference increasing when the ratio $E_{inclusion}/E_{matrix}$ decreases. As the curve using (47) is more or less symmetric with respect to the vertical axis corresponding to $E_{inclusion}/E_{matrix} = 1$, the curve using (50) has a visibly asymmetric shape: as it is almost linear (in log-log scale) when $E_{inclusion}/E_{matrix} > 1$, it seems to reach an asymptot when $\log(E_{inclusion}/E_{matrix})$ decreases. This is surprising because these curves must be perfectly symmetric when the two phases are interchangeable in the microstructure (as can be checked on figure (14) which presents results obtained on a checkerboard-like microstructure), and, thus, they are expected to be more or less symmetric when the phases are not interchangeable.



(a)



(b)



(c)

Figure 13: Computations made with the basic iterative scheme for a prescribed macroscopic strain $E_{11} = 1$, $E_{ij \neq 11} = 0$, for several ratios between the Young moduli of the inclusion and of the matrix in the microstructure (3a) (with 128×128 pixels). Poisson coefficients of both phases: $\nu = 0.3$. The error is estimated either with (47) (solid line) or with (50) (dashed line). (a): number of iterations to reach error tolerance 0.001

(b): comparison between the macroscopic stress $\langle \sigma_{11} \rangle$ obtained at the end of each computation and the macroscopic stress of a reference (high precision) computation

(c): comparison between the stress field obtained at the end of each computation and the stress field of the reference computation $\frac{\sqrt{\langle \|\sigma(\mathbf{x}) - \sigma_{\text{ref}}(\mathbf{x})\|^2 \rangle}}{\sqrt{23 \langle \|\sigma_{\text{ref}}(\mathbf{x})\|^2 \rangle}}$

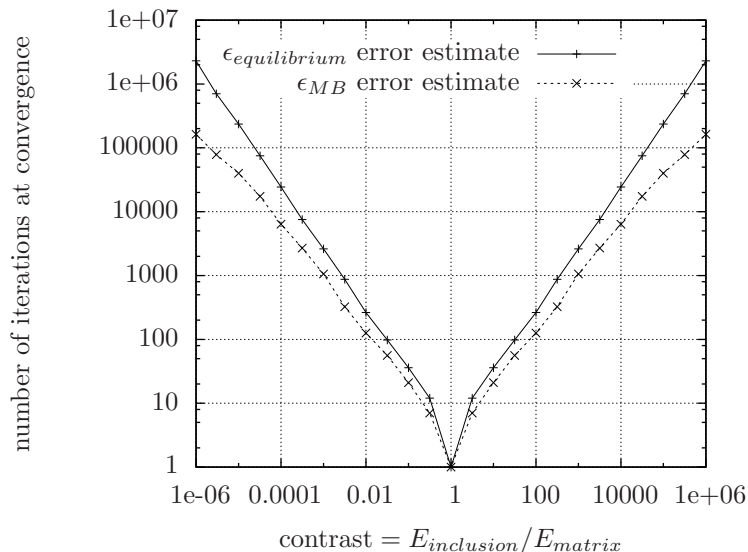


Figure 14: Same computations as presented in figure 13, but on a **checkerboard**-like microstructure.

The two phases of the microstructure are interchangeable and thus the curves plotted here are perfectly symmetric, whatever is the error estimate.

A simple test of the efficiency of both error estimates consists in comparing the results of each computation with reference results. The relative differences between the macroscopic stress obtained for each computation and the macroscopic stress σ_{ref} of a reference high precision computation (accelerated scheme of Eyre-Milton, error criterion (47), error tolerance $\mathcal{E} = 10^{-9}$), i.e.:

$$\frac{\langle \sigma_{11} \rangle - \langle \sigma_{11}^{ref} \rangle}{\langle \sigma_{11}^{ref} \rangle}$$

were evaluated and plotted on the middle diagram of Fig. 13. The root mean square difference between the stress field of each computation and the stress field of the reference computation, normalized by the root mean square of the reference stress field, i.e.:

$$\frac{\|\sigma - \sigma_{ref}\|_2}{\|\sigma_{ref}\|_2} = \frac{\sqrt{\langle \|\sigma(\mathbf{x}) - \sigma_{ref}(\mathbf{x})\|^2 \rangle}}{\sqrt{\langle \|\sigma_{ref}(\mathbf{x})\|^2 \rangle}}$$

is plotted on the lower diagram of figure 13. As can be seen, when the error estimate (50) is used, this difference increases significantly as the ratio $E_{inclusion}/E_{matrix}$ decreases, whereas it remains small when the error estimate (47) is used.

In conclusion, we suspect the error estimate (50) proposed by Monchiet and Bonnet of being too permissive and of being responsible of too early stop of the iterative process when the ratio $E_{inclusion}/E_{matrix}$ is small.

Acknowledgements

Mihail Garajeu is gratefully acknowledged for his help in the mathematical proof of (B).

References

- [1] Moulinec H., Suquet P. 1994. *A fast numerical method for computing the linear and nonlinear properties of composites. C. R. Acad. Sc. Paris II*, **318**, 1417–1423.
- [2] Moulinec H., P. Suquet. 1998. *A numerical method for computing the overall response of nonlinear composites with complex microstructure. Comp. Meth. Appl. Mech. Engng.*, **157**, 69–94.
- [3] Müller W.H. 1996. *Mathematical versus experimental stress analysis of inhomogeneities in solids. J. Physique IV*, **6**:C1.139C1148
- [4] Lebensohn R.A. 2001. *N-site modelling of a 3D viscoplastic polycrystal using Fast Fourier Transform. Acta Materialia*, **49**, 2723–2737.
- [5] Vinogradov V., Milton G.W. 2008. *An accelerated FFT algorithm for thermoelastic and non linear composites. Int. J. Numer. Meth. Engng.*, **76**: 1678–1695. doi: 10.1002/nme.2375
- [6] Zeman J., Vondrej J., Novak J. and Marek I. 2010. *Accelerating a FFT-based solver for numerical homogenization of periodic media by conjugate gradients. Journal of Computational Physics*, **229**, 8065–8071
- [7] Brisard S., Dormieux L. 2010. *FFT-based methods dor the mechanics of composites: A general variational framework. Computational Materials Science*, **49**, 663–671
- [8] Eyre D.J., Milton G.W. 1999. *A fast numerical scheme for computing the response of composites using grid refinement. J. Physique III*, **6**, 41–47.
- [9] Milton G.W. 2002. *The Theory of Composites'* (1st edition). Cambridge University Press.
- [10] Michel J.C, Moulinec H, Suquet P. 2000. *A computational method based on augmented Lagrangians and Fast Fourier Transforms for composites with high contrast Comput. Modelling Engng. Sc.*, **1** (2), 79–88.
- [11] Monchiet V. and Bonnet G. (2012): *A polarization-based FFT iterative scheme for computing the effective properties of elastic composites with arbitrary contrast. Int. J. Numer. Meth. Engng.*, **89**: 1419-1436. doi: 10.1002/nme.3295
- [12] Eshelby J.D. 1957. *The Determination of the Elastic Field of an Ellipsoidal Inclusion, and Related Problems Proc. R. Soc. Lond. A*, **241**, 376-396
- [13] Michel J.C, Moulinec H, Suquet P. 2001. *A computational method for linear and nonlinear composites with arbitrary phase contrast INT. J. Numer. Meth. Engng* **52**, 139-160.
- [14] Bhattacharya K., Suquet P.M. 2005. *A Model Problem concerning Recoverable Strains of Shape-Memory Polycrystals . Proc. R. Soc. Lond. A*, **461**; 2797–2816

A Properties of the operators Γ^0 and Δ^0

A.1 Analytical expression of Γ^0

When \mathbf{c}^0 is isotropic, the associated Green's operator $\mathbf{\Gamma}^0$ has a simple expression in Fourier space which reads:

$$\widehat{\Gamma}_{ijkh}^0(\boldsymbol{\xi}) = \frac{1}{4\mu^0|\boldsymbol{\xi}|^2}(\delta_{ki}\xi_h\xi_j + \delta_{hi}\xi_k\xi_j + \delta_{kj}\xi_h\xi_i + \delta_{hj}\xi_k\xi_i) - \frac{\lambda^0 + \mu^0}{(\lambda^0 + 2\mu^0)\mu^0} \frac{\xi_i\xi_j\xi_k\xi_h}{|\boldsymbol{\xi}|^4}. \quad (58)$$

A.2 Properties of Γ^0 and Δ^0

For any second-order tensor field $\mathbf{t}(\mathbf{x})$:

$$\Gamma^0 * \mathbf{t} \text{ is compatible and: } \langle \Gamma^0 * \mathbf{t} \rangle = 0 \quad (59)$$

$$\Gamma^0 * \mathbf{t} = \mathbf{0} \Leftrightarrow \mathbf{t} \text{ is equilibrated (i.e. } \operatorname{div}(\mathbf{t}) = \mathbf{0} \text{)} \quad (60)$$

$$\Delta^0 \mathbf{t} = \mathbf{0} \Leftrightarrow \mathbf{t} \text{ is compatible} \quad (61)$$

and:

$$\Gamma^0 * \mathbf{c}^0 : \Gamma^0 = \Gamma^0 \quad (62)$$

$$\Delta^0 * \mathbf{c}^{0^{-1}} : \Delta^0 = \Delta^0 \quad (63)$$

$$\Gamma^0 * \Delta^0 = \mathbf{0} \quad (64)$$

$$\Delta^0 \Gamma^0 = \mathbf{0} \quad (65)$$

B “Best” scheme

Let us find the values of α , β , k_0 and μ_0 , which minimize \mathcal{R}_u in the general case, i.e. without the hypothesis that $\alpha = \beta$. Thus the problem consists in finding:

$$\min_{\alpha, \beta, k_0, \mu_0} \mathcal{R}_u$$

or, using relation (38),

$$\min_{\alpha, \beta, k_0, \mu_0} (\max(|\Phi_{\alpha, k_{min}}|, |\Phi_{\alpha, k_{max}}|, |\Phi_{\beta, k_{min}}|, |\Phi_{\beta, k_{max}}|, |\Phi_{\alpha, \mu_{min}}|, |\Phi_{\alpha, \mu_{max}}|, |\Phi_{\beta, \mu_{min}}|, |\Phi_{\beta, \mu_{max}}|)) \quad (66)$$

This implies that:

$$\text{and} \begin{cases} \min_{\alpha, \beta, k_0, \mu_0} \mathcal{R}_u \geq \min_{\alpha, \beta, k_0} (\max(|\Phi_{\alpha, k_{min}}|, |\Phi_{\alpha, k_{max}}|, |\Phi_{\beta, k_{min}}|, |\Phi_{\beta, k_{max}}|)) \\ \min_{\alpha, \beta, k_0, \mu_0} \mathcal{R}_u \geq \min_{\alpha, \beta, \mu_0} (\max(|\Phi_{\alpha, \mu_{min}}|, |\Phi_{\alpha, \mu_{max}}|, |\Phi_{\beta, \mu_{min}}|, |\Phi_{\beta, \mu_{max}}|)) \end{cases}$$

In the first step of the demonstration, we will look for:

$$\mathcal{A} = \min_{\alpha, \beta, k_0} (\max(|\Phi_{\alpha, k_{min}}|, |\Phi_{\alpha, k_{max}}|, |\Phi_{\beta, k_{min}}|, |\Phi_{\beta, k_{max}}|)) \quad (67)$$

As:

$$\max(|\Phi_{\alpha, k_{min}}|, |\Phi_{\alpha, k_{max}}|, |\Phi_{\beta, k_{min}}|, |\Phi_{\beta, k_{max}}|) \geq \max(|\Phi_{\alpha, k_{min}}|, |\Phi_{\alpha, k_{max}}|)$$

and:

$$\max(|\Phi_{\alpha, k_{min}}|, |\Phi_{\alpha, k_{max}}|, |\Phi_{\beta, k_{min}}|, |\Phi_{\beta, k_{max}}|) \geq \max(|\Phi_{\beta, k_{min}}|, |\Phi_{\beta, k_{max}}|)$$

one has:

$$\mathcal{A} \geq \min_{k_0} \left(\max \left(\min_{\alpha} (\max(|\Phi_{\alpha, k_{min}}|, |\Phi_{\alpha, k_{max}}|)), \min_{\beta} (\max(|\Phi_{\beta, k_{min}}|, |\Phi_{\beta, k_{max}}|)) \right) \right) \quad (68)$$

For k_0 fixed, $\max(|\Phi_{\alpha, k_{min}}|, |\Phi_{\alpha, k_{max}}|)$ is minimized when:

$$\Phi_{\alpha, k_{min}} = -\Phi_{\alpha, k_{max}}$$

i.e. when:

$$1 - \alpha \frac{k_{min}}{k_0 + k_{min}} = -1 + \alpha \frac{k_{max}}{k_0 + k_{max}}$$

thus, optimal value for α is:

$$\alpha_{opt} = \frac{2}{\frac{k_{min}}{k_0 + k_{min}} + \frac{k_{max}}{k_0 + k_{max}}} \quad (69)$$

and, when $\alpha = \alpha_{opt}$, one has:

$$\max(|\Phi_{\alpha, k_{min}}|, |\Phi_{\alpha, k_{max}}|) = \frac{k_0(k_{max} - k_{min})}{k_0(k_{max} + k_{min}) + 2k_{max}k_{min}} \quad (70)$$

For k_0 , and μ_0 supposed to be fixed, $\max(|\Phi_{\beta, k_{min}}|, |\Phi_{\beta, k_{max}}|)$ is minimized when:

$$\Phi_{\beta, k_{min}} = -\Phi_{\beta, k_{max}}$$

i.e. when:

$$1 - \beta \frac{k_0}{k_0 + k_{min}} = -1 + \beta \frac{k_0}{k_0 + k_{max}}$$

thus, optimal value for β is:

$$\beta_{opt} = \frac{2}{\frac{k_0}{k_0 + k_{min}} + \frac{k_0}{k_0 + k_{max}}} \quad (71)$$

and, when $\beta = \beta_{opt}$, one has:

$$\max(|\Phi_{\beta, k_{min}}|, |\Phi_{\beta, k_{max}}|) = \frac{k_{max} - k_{min}}{2k_0 + k_{max} + k_{min}} \quad (72)$$

Expression (70) is an increasing function of k_0 and expression (72) is a decreasing function of k_0 and, thus: $\max(\min_{\alpha}(\max(|\Phi_{\alpha, k_{min}}|, |\Phi_{\alpha, k_{max}}|)), \min_{\beta}(\max(|\Phi_{\beta, k_{min}}|, |\Phi_{\beta, k_{max}}|)))$ is minimized when:

$$\frac{k_0(k_{max} - k_{min})}{k_0(k_{max} + k_{min}) + 2k_{max}k_{min}} = \frac{k_{max} - k_{min}}{2k_0 + k_{max} + k_{min}}$$

i.e. when:

$$k_0 = k_{0opt} = \sqrt{k_{min} \times k_{max}} \quad (73)$$

for which:

$$\alpha_{opt} = \beta_{opt} = 2 \quad (74)$$

and expression (70) and (72) are equal to:

$$\frac{\sqrt{k_{max}} - \sqrt{k_{min}}}{\sqrt{k_{max}} + \sqrt{k_{min}}} \quad (75)$$

And thus, expression (68) simplifies into:

$$\mathcal{A} \geq \frac{\sqrt{k_{max}} - \sqrt{k_{min}}}{\sqrt{k_{max}} + \sqrt{k_{min}}} \quad (76)$$

It is simple to verify that, when $\alpha = \beta = 2$ and $k_0 = \sqrt{k_{min} \times k_{max}}$, one has:

$$\Phi_{\alpha, k_{min}} = \Phi_{\alpha, k_{max}} = \Phi_{\beta, k_{min}} = \Phi_{\beta, k_{max}} = \frac{\sqrt{k_{max}} - \sqrt{k_{min}}}{\sqrt{k_{max}} + \sqrt{k_{min}}}$$

and finally that:

$\min_{\alpha,\beta,k_0}(\max(|\Phi_{\alpha,k_{min}}|, |\Phi_{\alpha,k_{max}}|, |\Phi_{\beta,k_{min}}|, |\Phi_{\beta,k_{max}}|))$ is reached when $\alpha = \beta = 2$ and $k_0 = \sqrt{k_{min} \times k_{max}}$ and equal to $\frac{\sqrt{k_{max}} - \sqrt{k_{min}}}{\sqrt{k_{max}} + \sqrt{k_{min}}}$

In a very similar manner, it can be proven that $\min_{\alpha,\beta,\mu_0}(\max(|\Phi_{\alpha,\mu_{min}}|, |\Phi_{\alpha,\mu_{max}}|, |\Phi_{\beta,\mu_{min}}|, |\Phi_{\beta,\mu_{max}}|))$ is reached when $\alpha = \beta = 2$ and $\mu_0 = \sqrt{\mu_{min} \times \mu_{max}}$ and equal to $\frac{\sqrt{\mu_{max}} - \sqrt{\mu_{min}}}{\sqrt{\mu_{max}} + \sqrt{\mu_{min}}}$

Finally: $\min_{\alpha,\beta,k_0,\mu_0} \mathcal{R}_u$ is reached when $\alpha = 2, \beta = 2, k_0 = \sqrt{k_{min} \times k_{max}}$ and $\mu_0 = \sqrt{\mu_{min} \times \mu_{max}}$ and is equal to:

$$\min_{\alpha,\beta,k_0,\mu_0} \mathcal{R}_u = \max\left(\frac{\sqrt{k_{max}} - \sqrt{k_{min}}}{\sqrt{k_{max}} + \sqrt{k_{min}}}, \frac{\sqrt{\mu_{max}} - \sqrt{\mu_{min}}}{\sqrt{\mu_{max}} + \sqrt{\mu_{min}}}\right) \quad (77)$$