Optimization of Eigenvalues and Eigenmodes by using the Adjoint Method

Anca-Maria Toader and Cristian Barbarosie *

Abstract

Application of the Adjoint Method has proven successful in Shape Optimization and Topology Optimization. In the present paper the Adjoint Method is applied to the optimization of eigenvalues and eigenmodes (eigenvectors). The general case of an arbitrary cost function depending on the first n eigenvalues and eigenmodes is detailed. The direct problem does not involve a bilinear form and a linear form as usual in other applications. However, it is possible to follow the spirit of the method and deduce n adjoint problems and obtain n adjoint states, where n is the number of eigenmodes taken into account for optimization. An optimization algorithm based on the derivative of the cost function is developed. This derivative depends on the derivatives of the eigenmodes and the Adjoint Method allows one to express it in terms of the the adjoint states and of the solutions of the direct eigenvalue problem. The formulas hold for the case when the eigenvalues are simple. A section is dedicated to discussions on the case when there are multiple eigenvalues. The same procedures are applied to optimization of microstructures, modeled by Bloch waves. The results obtained hold for general functionals depending on the eigenvalues and on the eigenmodes of the microstructure. However, the wave vector $\vec{k}$ is a more delicate case of optimization parameter. The derivative of a general functional with respect to $\vec{k}$ is obtained, which has interesting implications in band-gap maximization problems.

1 Introduction

The Adjoint Method in the Calculus of Variations is being used in a rigorous framework in optimization problems governed by partial differential equations since the eighties, with the papers of J. Céa, F. Murat and L. Tartar.

F. Murat and L. Tartar in [MuratTartar1985] applied the Adjoint Method in the Theory of Homogenization and pointed out its origins in the works of Pontryagin, namely in Pontryagin's Principle. J. Céa in [Cea1986] uses the Adjoint Method from the perspective of the Lagrange multipliers in a practical way.

*Centro de Matemática e Aplicações Fundamentais, Faculdade de Ciências, Universidade de Lisboa, 1749-016 Lisboa, Portugal. atoader@fc.ul.pt, cabarbarosie@fc.ul.pt

Applications of the Adjoint Method to control problems, shape optimization and topology optimization were considered in a large number of scientific papers. Among them we mention [AllaireJouveT2004], [AllaireJouve2008], [AllaireJouveVG2011], [AllaireDaponyF14] in the domain of structural optimization of solid bodies. However the optimization of functionals depending only on the eigenvalues, which involves a self-adjoint problem, was studied earlier in [KikuchiChengM1995].

A brief description of the Adjoint Method is given in the sequel. The main problem in the Calculus of Variations can be stated in a simplistic way like : Given two physical quantities u and θ that are related $u = u(\theta)$, determine the relation between their variations. More precisely, given a variation of the quantity θ, denoted by $\delta\theta$, it produces a variation of the quantity u, denoted by δu; one seeks for the relation between their variations δu and $\delta\theta$. When the variations $\delta\theta$ and δu are infinitesimal, this leads naturally to the notion of derivative of u with respect to θ, having in this case a linear dependency between δu and $\delta\theta$.

In the framework of applications, the relation of u and θ is given by a state equation, which may be a partial differential equation having the variational formulation

$$A_\theta(u_\theta, v) = l_\theta(v), \forall v \in H, \tag{P_θ}$$

where A_θ is a bilinear, continuous and coercive application depending on θ, l_θ is a linear, continuous application depending on θ, H is a Hilbert space. The physical quantity θ is called control.

The state u_θ depends on θ in an implicit way, through the problem (P_θ) and one seeks to describe the variation of u_θ in terms of the variation of θ, under convenient differentiability hypothesis on the families of operators A_θ and l_θ with respect to θ.

In optimal control problems, given a functional to minimize/maximize, the goal is to find a control θ producing the state u_θ that minimizes/maximizes the functional. Namely, given a functional

$$J(\theta) = \mathcal{J}(\theta, u_\theta),$$

that depends on θ, in a direct mode and through the solution u_θ do problem (P_θ), one is looking for a θ that minimizes/maximizes J. Since a necessary condition for extremum is the vanishing of the total derivative

$$\frac{dJ}{d\theta} = 0,$$

one of the most important ingredients is the analytic calculus of the above referred derivative. Note that the total derivative of J involves the derivative of u_θ with respect to θ :

$$\frac{dJ}{d\theta}(\theta)\,\tau = \frac{\partial\mathcal{J}}{\partial\theta}(\theta, u_\theta)\,\tau + \frac{\partial\mathcal{J}}{\partial u}(\theta, u_\theta)\frac{du_\theta}{d\theta}(\theta)\,\tau. \tag{dJ}$$

The difficulty lays in the fact that the term $\frac{\partial \mathcal{J}}{\partial u}(\theta, u_\theta)\frac{du_\theta}{d\theta}(\theta)\,\tau$ involves $\frac{du_\theta}{d\theta}(\theta)\,\tau$; note that u_θ depends on θ implicitly through problem (P_θ). It is the *Adjoint Method* that allows one to transform this implicit dependency in an explicit one with respect to τ.

Consider the adjoint problem in the form

$$A_\theta(p_\theta, w) = \frac{\partial \mathcal{J}}{\partial u}(\theta, u_\theta)w, \forall w \in H, \qquad (PA_\theta)$$

the solution $p_\theta \in H$ is called the *adjoint state*. Then under symmetry hypothesis on the operators A_θ it is possible to prove that the implicit term in the expression of (dJ) has the form :

$$\frac{\partial \mathcal{J}}{\partial u}(\theta, u_\theta)\frac{du_\theta}{d\theta}(\theta)\,\tau = A_\theta(p_\theta, \frac{du_\theta}{d\theta}(\theta)\,\tau) = A_\theta(\frac{du_\theta}{d\theta}(\theta)\,\tau, p_\theta) =$$

$$= \left(-\frac{dA_\theta}{d\theta}(u_\theta, p_\theta) + \frac{dl_\theta}{d\theta}(p_\theta)\right)\tau.$$

Therefore, the total derivative of J writes as the following expression, where the dependency on τ is explicit :

$$\frac{dJ}{d\theta}(\theta)\,\tau = \left(\frac{\partial \mathcal{J}}{\partial \theta}(\theta, u_\theta) - \frac{dA_\theta}{d\theta}(u_\theta, p_\theta) + \frac{dl_\theta}{d\theta}(p_\theta)\right)\,\tau.$$

The Calculus of Variations provides the derivative of the functional J in an arbitrary point θ. However in the shape and topology optimization methods only the derivative in $\theta = 0$ is employed since these methods use the current configuration of the control in order to determine the configuration of the next iteration.

The problem of optimization of the eigenvalues and eigenmodes of an elastic structure or microstructure makes the goal of the present work. The context is essentially different from the above exposed one since the state problem does not involve any more a bilinear form and a linear form.

In section 2 is given the setting of the problem and the objective functionals are discussed. The derivatives, with respect to an optimization parameter, of the eigenvalues and of the eigenmodes of vibration are deduced in section 3. In section 4 the Adjoint Method is employed and the derivative of a general objective functional is obtained. Some comments on the case of multiple eigenvalues are presented in section 5. Section 6 is dedicated to the optimization of the eigenvalues and of the eigenmodes of a microstructure. The framework of Bloch waves is chosen and the Adjoint Method applies in order to obtain the derivative of a general functional with respect to a parameter. When the parameter is the wave vector itself, the derivative may be used in band-gap maximization problems.

2 Setting of the problem and the objective functionals

Consider a body occupying a domain $\Omega \subset \mathbb{R}^3$; the boundary $\partial\Omega$ is splitted in two disjoint parts: the part Γ_D, where the Dirichlet condition is applied, and the part Γ_N, where the Neumann condition is applied.

The symbol ϵ denotes the strain tensor and ρ the specific mass (density) of the material. The symbol $\mathbf{C} = \mathbf{C}(s)$ denotes the fourth-order elasticity tensor and depends on the material parameters s that characterize the chosen model. Both ρ and $\mathbf{C}$ may vary in different parts of the domain.

The eigenvalue problem

$$\begin{cases} -div(\mathbf{C}\epsilon(u)) &= \lambda\rho u \ \text{ in } \Omega, \\ u &= 0 \ \text{ on } \Gamma_D, \\ \mathbf{C}\epsilon(u) \cdot n &= 0 \ \text{ on } \Gamma_N. \end{cases} \tag{1}$$

arises from the system of dynamic equations by searching its solutions of the form $\mathbf{u}(t,x) = \phi(t)u(x)$, where ϕ is a sinusoidal function. One concludes that the function $u = u(x)$ must satisfy the eigenvalue problem (1).

The natural frequencies $\omega_i(s)$ of the model are calculated as $\dfrac{\sqrt{\lambda_i}}{2\pi}$, where λ_i are the eigenvalues of the eigenvalue problem (1). The vibration modes u_i are the eigenvectors of problem (1). Since $\mathbf{C} = \mathbf{C}(s)$, the solutions $(\lambda_i(s), u_i(s))_{i\geq 1}$ of the eigenvalue problem (1) depend on s as well. The material parameters are denoted by s, which in this case represent an m-tuple in $\mathbb{R}^m$ (that is, the model has m material parameters $s_1, \ldots, s_m$).

The norm $\|\cdot\|_{L^2_\rho}$ on the space $L^2(\Omega)^3$ is induced by the following inner product associated to the function $\rho \in L^\infty(\Omega)$ representing the specific mass ($\rho(x) \geq \rho_0 > 0$ almost everywhere in Ω):

$$(u, v) \mapsto \int_\Omega \rho u \cdot v \, dx \, .$$

Consider a general objective functional $J(\lambda_i(s), u_i(s))$ depending of the first n eigenvalues and vibration modes. The dependency of the eigenvalues $\lambda_i(s)$ and of the vibration modes $u_i(s)$ on the parameters s is characterized by the problem (1) itself and is highly non-linear. Therefore J is non-linear itself.

3 The derivatives of the eigenvalues and eigenmodes of vibration

In order to obtain the derivatives of the eigenvalues and modes of vibration, the variational formulation of the eigenvalue problem (1) is used. For a fixed parameter s, the (weak) solutions $(\lambda(s),\ u(s))$ of the eigenvalue problem (1) are

the solutions of the following variational problem:

$$\left\{ \begin{array}{c} \text{find} \quad \lambda(s) \in \mathbb{R} \quad \text{and} \quad u(s) \in V \text{ such that} \\ \displaystyle\int_\Omega \mathbf{C}(s)\,\epsilon(u(s)) \cdot \epsilon(v)\,dx = \lambda \int_\Omega \rho u(s) \cdot v\,dx, \quad \forall v \in V, \end{array} \right. \tag{2}$$

where the space V of test functions is defined as

$$V = \{v \in H^1(\Omega)^3 : v_{|\Gamma_D} = 0\}.$$

A classical result states the existence of a countably infinite set of solutions of problem (2), with the eigenvalues $(\lambda_i)_{i\geq 1}$ forming an increasing sequence of real numbers, the eigenvectors $(u_i)_{i\geq 1}$ forming an hilbertian basis of L^2_ρ, and such that the pair (λ_i, u_i) verifies problem (1) in the almost everywhere sense, for all $i \geq 1$.

For simplicity of the exposition, we will assume that the material model in consideration for the body Ω has only one material parameter $s \in \mathbb{R}$. However the following results hold in the vectorial case.

Theorem 1 *Provided differentiability properties of the elasticity tensor* $\mathbf{C} = \mathbf{C}(s)$ *with respect to a general material parameter* s *and assuming that the eigenvalues of problem (1) are simple, then the eigenvalues and the eigenvectors are differentiable with respect to* s.

The derivative of the eigenvalue $\lambda_i = \lambda_i(s)$ *is*

$$\frac{d\lambda_i}{ds}(s) = \int_\Omega \frac{d\mathbf{C}}{ds}(s)\epsilon(u_i) \cdot \epsilon(u_i)\,dx, \tag{3}$$

where the corresponding eigenvector u_i *is normalized in* L^2_ρ*:* $\|u_i\|_{L^2_\rho} = 1$.

The derivative $\dfrac{du_i}{ds}$ *of the eigenvector* $u_i = u_i(s)$ *is the solution of the problem below:*

$$\left\{ \begin{array}{c} \textit{find} \quad \dfrac{du_i}{ds} \in < u_i >^\perp \quad \textit{such that} \\ \displaystyle\int_\Omega \mathbf{C}(s)\epsilon\Big(\frac{du_i}{ds}\Big) \cdot \epsilon(v)\,dx - \lambda_i \int_\Omega \rho \frac{du_i}{ds} \cdot v\,dx = \\ \displaystyle\frac{d\lambda_i}{ds}(s) \int_\Omega \rho\, u_i \cdot v\,dx - \int_\Omega \frac{d\mathbf{C}}{ds}(s)\epsilon(u_i) \cdot \epsilon(v)\,dx, \\ \forall v \in V. \end{array} \right. \tag{4}$$

where $< u_i >^\perp$ *denotes the orthogonal complement of the eigenspace generated by* u_i*, with respect to the inner product in* L^2_ρ.

Proof : In order to simplify the notations, the index i will be omitted. Consider the mappings $s \mapsto \lambda(s)$ and $u \mapsto u(s)$, where $u(s)$ is normalized in

L^2_ρ . The differentiability of the above mappings is proved with detail in [RousseletChenais1990].

Since $\lambda(s)$ is an eigenvalue and the associated eigenvector $u(s)$ is in the space V, equality (2) holds with $v = u(s)$:

$$\int_\Omega \mathbf{C}(s)\epsilon(u(s)) \cdot \epsilon(u(s))dx = \lambda(s) \int_\Omega \rho u(s) \cdot u(s)dx = \lambda(s).$$

Thus:

$$\begin{aligned}
\frac{d\lambda}{ds}(s) &= \frac{d}{ds}\int_\Omega \mathbf{C}(s)\epsilon(u(s)) \cdot \epsilon(u(s))dx \\
&= \int_\Omega \frac{d}{ds}\Big[\mathbf{C}(s)\epsilon(u(s))\Big] \cdot \epsilon(u(s))\, dx + \int_\Omega \mathbf{C}(s)\epsilon(u(s)) \cdot \frac{d}{ds}\epsilon(u(s))\, dx \\
&= \int_\Omega \frac{d\mathbf{C}}{ds}(s)\epsilon(u(s)) \cdot \epsilon(u(s))\, dx + 2\int_\Omega \mathbf{C}(s)\epsilon(u(s)) \cdot \frac{d}{ds}\epsilon(u(s))\, dx.
\end{aligned}$$

Taking $v = \dfrac{du}{ds}(s)$ in (2), we conclude that the second term in the last expression is null:

$$\begin{aligned}
\int_\Omega \mathbf{C}(s)\epsilon(u(s)) \cdot \epsilon\Big(\frac{du}{ds}(s)\Big)dx &= \lambda(s)\int_\Omega \rho u(s) \cdot \frac{du}{ds}(s)dx \\
&= \frac{1}{2}\,\lambda(s)\,\frac{d}{ds}\int_\Omega \rho\, u(s) \cdot u(s)dx \\
&= \frac{1}{2}\,\lambda(s)\,\frac{d}{ds}\|u(s)\|^2_{L^2_\rho} \\
&= 0,
\end{aligned}$$

Therefore,

$$\frac{d\lambda}{ds}(s) = \int_\Omega \frac{d\mathbf{C}}{ds}(s)\epsilon(u(s)) \cdot \epsilon(u(s))dx. \tag{5}$$

On the other hand, from (2), by derivation with respect to the parameter s, it turns out that $\dfrac{du}{ds}(s)$ verifies the following problem:

$$\left\{\begin{array}{l}
\displaystyle\int_\Omega \mathbf{C}(s)\epsilon\big(\frac{du}{ds}(s)\big) \cdot \epsilon(v)\, dx - \lambda(s)\int_\Omega \rho\frac{du}{ds}(s) \cdot v\, dx = \\
\qquad \displaystyle\frac{d\lambda}{ds}(s)\int_\Omega \rho\, u(s) \cdot v\, dx - \int_\Omega \frac{d\mathbf{C}}{ds}(s)\epsilon(u(s)) \cdot \epsilon(v)\, dx, \\
\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \forall v \in V.
\end{array}\right.$$

According to Fredholm alternative, the above problem has a unique solution in $< u >^\perp$ if the compatibility condition below holds:

$$\frac{d\lambda}{ds}(s)\int_\Omega \rho\, u(s) \cdot u(s)\, dx - \int_\Omega \frac{d\mathbf{C}}{ds}(s)\epsilon(u(s)) \cdot \epsilon(u(s))\, dx = 0. \tag{6}$$

Since u is normalized in L^2_ρ and since the condition (6) holds, it turns out that the compatibility condition above is verified and thus the derivative $\dfrac{du}{ds}(s)$ is the unique solution of the above problem.

Remark 1 *Formula (3) has already been obtained in [RousseletChenais1990]. A similar formula has been used in the framework of shape optimization, see [AllaireJouve2005].*

Remark 2 *For a fixed index i, the above mapping $s \mapsto (\lambda_i(s), u_i(s))$ with $\|u_i\|_{L^2_\rho} = 1$, is not well defined (there are two possibilities, u_i and $-u_i$). However, for the calculation of $\dfrac{du_i}{ds}(s)$ the sign of the eigenvector is important.*

Remark 3 *Since the natural frequency ν_i writes in terms of the eigenvalues λ_i as $\nu_i = \dfrac{\sqrt{\lambda_i}}{2\pi}$, the derivative of ν_i has consequently the form*

$$\frac{d\nu_i}{ds} = \frac{1}{4\pi\sqrt{\lambda_i}}\frac{d\lambda_i}{ds} = \frac{1}{4\pi\sqrt{\lambda_i}}\int_\Omega \frac{d\mathbf{C}}{ds}\epsilon(u_i)\cdot\epsilon(u_i)\,dx. \tag{7}$$

Remark 4 *The case when the eigenvalues are multiple is discussed in section 5.*

4 The derivative of the objective functional by the adjoint method

Given a generic functional $J(\lambda_i(s), u_i(s))$ depending on the first n eigenvalues and on the corresponding n eigenvectors, a straightforward computation gives its derivative as

$$\frac{dJ}{ds}(s) = \sum_{i=1}^{n}\frac{\partial J}{\partial \lambda_i}\,\frac{d\lambda_i}{ds}(s) + \sum_{i=1}^{n}\frac{\partial J}{\partial u_i}\,\frac{du_i}{ds}(s).$$

and from (3) it turns out that

$$\frac{dJ}{ds}(s) = \sum_{i=1}^{n}\frac{\partial J}{\partial \lambda_i}\int_\Omega \frac{d\mathbf{C}}{ds}\epsilon(u_i(s))\cdot\epsilon(u_i(s))\,dx + \sum_{i=1}^{n}\frac{\partial J}{\partial u_i}\,\frac{du_i}{ds}(s). \tag{8}$$

The derivative of J with respect to the parameter s involves the derivatives of the eigenvectors $\dfrac{du_i}{ds}(s)$. In order to turn explicit the last term in the above expression of the derivative of J, the adjoint method will be employed.

Consider n adjoint problems with the form below for $1 \le i \le n$:

$$\begin{cases} \text{find } p_i \in V, \text{ with } \displaystyle\int_\Omega \rho p_i \cdot u_i(s)dx = 0 \text{ such that} \\ \displaystyle\int_\Omega \mathbf{C}(s)\epsilon(p_i)\cdot\epsilon(w)dx - \lambda_i(s)\int_\Omega \rho p_i\cdot w dx = \frac{\partial J}{\partial u_i}w, \forall w \in < u_i >^\perp . \end{cases} \tag{9}$$

Each one of the adjoint problems (9) will allow, by considering a particular test function $w = \frac{du_i}{ds}(s)$, to express the terms $\frac{\partial J}{\partial u_i}\,\frac{du_i}{ds}(s)$:

$$\int_\Omega \mathbf{C}(s)\epsilon(p_i)\cdot\epsilon(\frac{du_i}{ds})dx - \lambda_i(s)\int_\Omega \rho p_i\cdot\frac{du_i}{ds}dx = \frac{\partial J}{\partial u_i}\frac{du_i}{ds}.$$

Using problem (4) that defines $\frac{du_i}{ds}(s)$ with the test function equal to the corresponding adjoint state $v = p_i$, the above terms $\frac{\partial J}{\partial u_i}\,\frac{du_i}{ds}(s)$ become:

$$\frac{\partial J}{\partial u_i}\,\frac{du_i}{ds}(s) = \frac{d\lambda_i}{ds}(s)\int_\Omega \rho\, u_i\cdot p_i\, dx - \int_\Omega \frac{d\mathbf{C}}{ds}(s)\epsilon(u_i)\cdot\epsilon(p_i)\, dx.$$

Since p_i belongs to $< u_i(s) >^\perp$,

$$\frac{\partial J}{\partial u_i}\,\frac{du_i}{ds}(s) = -\int_\Omega \frac{d\mathbf{C}}{ds}(s)\epsilon(u_i)\cdot\epsilon(p_i)\, dx.$$

Consequently, the derivative of the functional J is obtained in the following explicit form:

$$\begin{aligned}\frac{dJ}{ds}(s) = &\sum_{i=1}^{n}\frac{\partial J}{\partial\lambda_i}\int_\Omega \frac{d\mathbf{C}}{ds}(s)\epsilon(u_i(s))\cdot\epsilon(u_i(s))\, dx\\ &-\sum_{i=1}^{n}\int_\Omega \frac{d\mathbf{C}}{ds}(s)\epsilon(u_i(s))\cdot\epsilon(p_i)\, dx,\end{aligned} \tag{10}$$

where p_i, for each $1 \le i \le n$, is the solution of the adjoint problem (9) and $(\lambda_i(s), u_i(s))$ is the solution of the variational eigenvalue problem (2).

An optimization algorithm based on the above derivative was developed and applied in [OliveiraToaderV2012] for identification purposes in material damage for the case of a dam.

5 Multiple eigenvalues

This section is devoted to the case of multiple eigenvalues. This case is considered difficult and is excluded by many theoretical studies, or, at the best, some rough, partial results are stated. Also, many numerical approaches ignore this case altogether.

Let us begin with a very simple example, the Laplace operator in a circular domain Ω :

$$\begin{cases} -\Delta u = \lambda u & \text{in} \quad \Omega \\ u = 0 & \text{on} \quad \partial\Omega \end{cases} \tag{11}$$

No structural parameters for now, no optimization, just analysis of a fixed domain. The first eigenvalue of problem (11) is simple; the second and third eigenvalues coincide.

Remark 5 *It can be argued that the above situation is "pathological", unstable and thus of no practical interest. This argument lies on the fact that in nature there are no perfect circles. Any object or structure in real world has imperfections. For instance, it may present a slight excentricity or other tiny imperfections. For a real world object, the probability of getting* $\lambda_2 = \lambda_3$ *is negligible. Thus, this situation has no practical interest; it represents only an abstract toy case. This argument applies even to numerical approaches to the problem : most softwares approximate the domain* Ω *by a mesh; thus,* Ω *will be computationally not a perfect circle. Even if meshless methods are employed, round-off errors will almost always prevent* $\lambda_2 = \lambda_3$ *from happening.*

The above argument is quite strong; however, when considering a varying parameter (like in structural optimization), equality between two eigenvalues becomes an unavoidable fact, as the following example illustrates. Consider the same problem (11) in a different domain

$$\Omega = \left\{(x, y) \in \mathbb{R}^2 : \frac{x^2}{s^2} + y^2 \leq 1\right\}$$

This is an ellipse whose excentricity depends on the parameter $s > 0$. If we plot λ_1, λ_2 and λ_3 against s, for a discrete family of values of s between 0.8 and 1.2, we get a picture like in Figure 1.

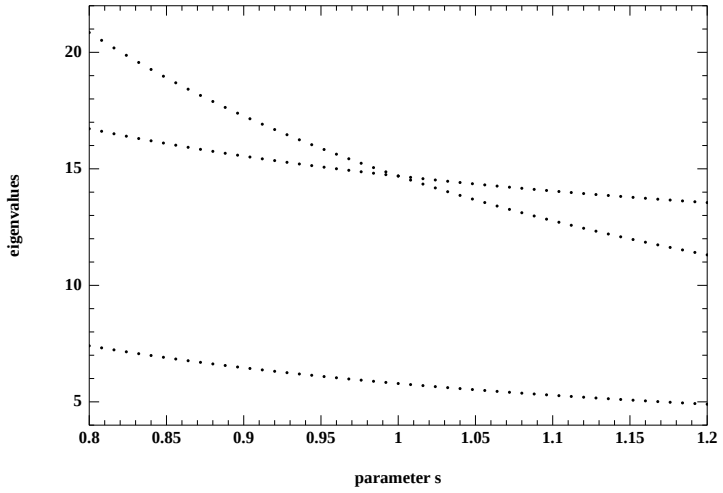

Figure 1: First three eigenvalues for an ellipse

The crossing occuring (at $s = 1$) between λ_2 and λ_3, is "stable" in the sense that small imperfections in the physical object (or numerical errors) will not prevent it from happening.

Remark 6 *With regard to Figure 1, several interesting questions arise. For instance, why should one call it "a crossing" between eigenvalues rather than*

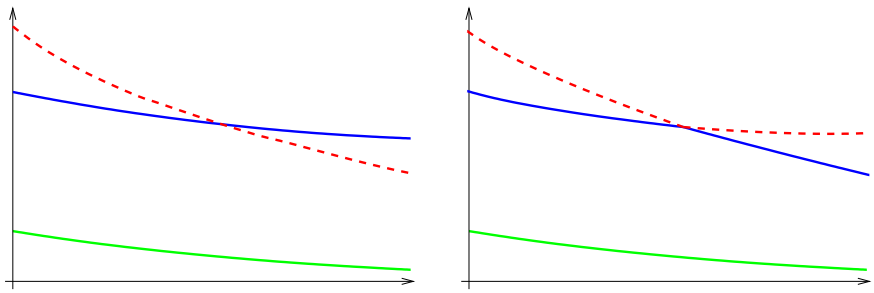

Figure 2: Crossing or just brief touch ?

considering that λ_2 and λ_3 merely touch each other at $s = 1$ and then go apart again ? In other words, can we robustly distinguish between the two situations depicted in Figure 2 ? Yet another way of putting it : how do we label eigenvalues ? Should we always stick stubbornly to the increasing order ? If we do so, it makes no sense to call it "crossing"; eigenvalues will allways just briefly touch each other. On the other hand, this strategy will produce points of non-differentiability. But if we "relax" the labelling procedure by giving up the "increasing" requirement, we may get a smooth dependence of both λ_1 and λ_2 with respect to the structural parameter s. Smoothness of course represents a huge advantage, but it comes at the cost of new difficulties, new callenges. If the labelling of eigenvalues is no longer pre-established, we need a new, reliable labelling strategy.

Remark 7 *People interested in optimization may argue that non-differentiability of the objective functional at only one point is no obstacle. The algorithm builds a sequence of discrete points s_k which converge to a minimum point, and the probability of one of these points hitting at the non-differentiability point is negligible. This argument is fallacious. If one tries to use the usual (steepest-descent-like) approach to this optimization problem, the algorithm will become unstable as soon as it reaches the neighborhood of the multiple eigenvalue, precisely due to jumps in the derivative of the objective functional. To make things worse, many objective functional will attain their minimum at a point corresponding to a multiple eigenvalue.*

Remark 8 *There are optimization algorithms specially designed for non-smooth objective functionals (like the bundle method or derivative-free methods). At first sight, these algorithms transform the question of labelling (or sorting) eigenvalues into an irrelevant detail. Recall, however, that derivative-based methods are usually faster than the ones which do not use derivatives. Thus, it is still interesting to try to obtain smooth curves whenever possible.*

In [BarbarosieRochaT], an algorithm is described which "colours" a set of data points in order to obtain a number of curves as smooth as possible. That is, the algorithm receives an input like in the left hand side of Figure 3 and produces

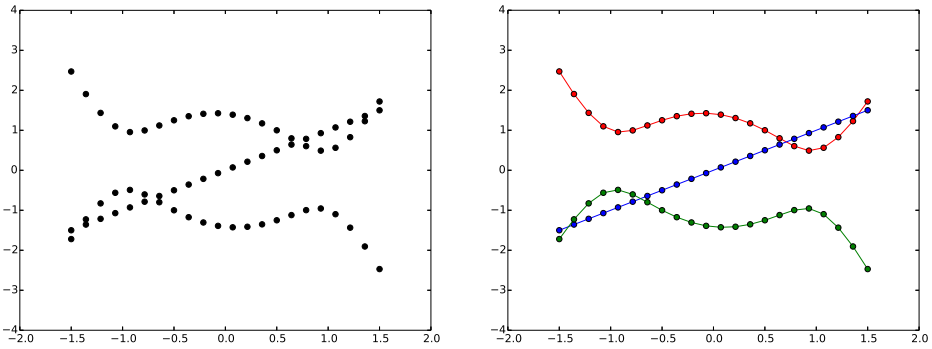

Figure 3: Colouring algorithm

an output like in the right hand side of the same Figure. This algorithm has no knowledge of where the data points come from; in particular, it has no knowledge of the eigenvectors associated to them. However, an interesting phenomenon has been observed : if we draw not only the eigenvalues but also the eigenvectors (like in Figure 4), it can be seen that this "colouring" procedure turns the dependence of the eigenvectors continuous or even smooth. Of course, the eigenvectors are not uniquely determined at the crossing point (there is an increased degree of arbitrariness due to the multiplicity of the eigenvalue). But, when looking immediately before and immediately after the crossing point, one can easily see the continuity (and perhaps even differentiability) of the eigenvectors.

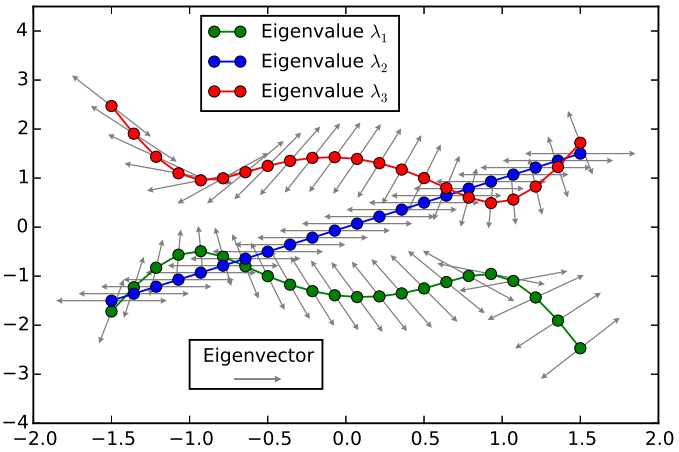

Figure 4: Eigenvalues and eigenvectors

Note that the eigenvectors usually belong to a space of high dimension and cannot be easily represented graphically. For instance, Figure 4 shows the eigen-

values and eigenvectors of the matrix

$$\begin{bmatrix} 0 & 1-s^2 & -s \\ 1-s^2 & s/2 & 1-s^2 \\ -s & 1-s^2 & 0 \end{bmatrix}$$

with s varying between -1.5 and 1.5. The grey arrows represent a projection of the three-dimensional eigenvectors on a certain two-dimensional subspace.

Remark 9 *For more than one structural parameter, things are more complicated; see, for instance, Exercise 2.4 in [ConcaPlanchardV1995, Chapter III]. This is object of ongoing work.*

Results in this section were obtained using the softwares FreeFem++ [Hecht2012], Matplotlib [Hunter2007], xgraphic [Jouve] and xfig [xfig].

6 The adjoint method in the framework of Bloch waves

The above considerations may be adapted to the framework of Bloch waves. See [AllaireConca1998] and [ConcaPlanchardV1995, Chapter III] for a more detailed presentation on Bloch wave theory.

Consider a body in $\mathbb{R}^N$ ($N = 2$ or 3) made of a periodic material, that is, a material whose inhomogeneities are periodically distributed. Suppose that the period is small when compared to the overall size of the body. A general notion of periodicity is needed, according to which a function is periodic if it is invariant to a group of translations. Note that any translation can be identified with a vector in $\mathbb{R}^N$ and the composition of translations corresponds to the sum of the respective vectors, thus any notion of periodicity can be described by an additive subgroup $\mathcal{G}$ of $\mathbb{R}^N$ with N linearly independent generators $\vec{g}_1, \vec{g}_2, \ldots, \vec{g}_N$ (such a subgroup is called a lattice). A complex function $\varphi : \mathbb{R}^N \to \mathbb{C}$ is said to be $\mathcal{G}$-periodic if

$$\varphi(x + \vec{g}) = \varphi(x), \ \forall x \in \mathbb{R}^N \ \forall \vec{g} \in \mathcal{G}.$$

An elastic Bloch wave is the superposition of a plane wave of the form $e^{i\langle \vec{k}, x\rangle}$ and a perturbation φ which is a $\mathcal{G}$-periodic function. Thus, a Bloch wave writes :

$$u(x) = e^{i\langle \vec{k}, x\rangle}\varphi(x).$$

Note that this is equivalent to the following conditions on u :

$$u(x + \vec{g}_j) = e^{i\langle \vec{k}, \vec{g}_j\rangle}u(x), \ \forall x \in \mathbb{R}^N, \ \forall j = 1, \ldots N.$$

The vector $\vec{k}$ is called "wave vector".

It is useful to define the dual group $\mathcal{G}^*$ as

$$\mathcal{G}^* = \left\{\vec{h} \in \mathbb{R}^N \text{ such that } \langle \vec{h}, \vec{g}\rangle \in 2\pi\mathbb{Z}, \ \forall \vec{g} \in \mathcal{G}\right\}$$

When an elastic Bloch wave propagates through a body made of a $\mathcal{G}$-periodic material, and supposing that the wavelength is comparable to the size of the periodicity cell, then the following problem, called Bloch cellular problem, characterizes the propagation phenomenon :

$$\left\{\begin{array}{rcl} -div(\mathbf{C}\epsilon(u)) & = & \lambda\rho u \ \text{ in } \mathbb{R}^N_{\text{perf}}(T), \\ \mathbf{C}\epsilon(u)\cdot n & = & 0 \quad \text{on } \partial T, \\ u(x+\vec{g}_j) & = & e^{i\langle\vec{k},\vec{g}_j\rangle}u(x),\ \forall x\in\mathbb{R}^N_{\text{perf}}(T),\ \forall j=1,\dots N. \end{array}\right. \tag{12}$$

In the above, T is a compact set representing a model hole in the periodicity cell Y and $\mathbb{R}^N_{\text{perf}}(T)$ is the perforated space defined by

$$\mathbb{R}^N_{\text{perf}}(T) = \mathbb{R}^N \setminus \bigcup_{m\in\mathbb{Z}^N} (T + m_1\vec{g}_1 + \dots m_N\vec{g}_N).$$

Denote by $H^1_{\#}(\mathbb{R}^N_{\text{perf}}(T),\mathbb{C}^N)$ the completion in the norm of $H^1(Y\setminus T,\mathbb{C}^N)$ of the space of functions in $C^\infty(\mathbb{R}^N_{\text{perf}}(T),\mathbb{C}^N)$ which are $\mathcal{G}$-periodic.

For $\vec{k}\in\mathbb{C}$ arbitrarily fixed, denote by $W_{\vec{k}}(\mathbb{R}^N_{\text{perf}}(T))$ the set of Bloch waves having the plane wave in the direction $\vec{k}$:

$$W_{\vec{k}}(\mathbb{R}^N_{\text{perf}}(T)) = \{u : \mathbb{R}^N_{\text{perf}}(T)\to\mathbb{C}^N \mid u(x) = e^{i\langle\vec{k},x\rangle}\varphi(x), \varphi\in H^1_{\#}(\mathbb{R}^N_{\text{perf}}(T),\mathbb{C}^N)\}.$$

Thus the last equation in (12) is equivalent to $u\in W_{\vec{k}}(\mathbb{R}^N_{\text{perf}}(T))$. Note that $W_0(\mathbb{R}^N_{\text{perf}}(T)) = H^1_{\#}(\mathbb{R}^N_{\text{perf}}(T),\mathbb{C}^N)$ and $W_{\vec{k}}(\mathbb{R}^N_{\text{perf}}(T))$ is a Hilbert space, obtained from $H^1_{\#}(\mathbb{R}^N_{\text{perf}}(T),\mathbb{C}^N)$ by multiplication with the fixed function $e^{i\langle\vec{k},x\rangle}$.

The natural norm $\|\cdot\|_{L^2_\rho}$ on the space $L^2(\mathbb{R}^N_{\text{perf}}(T),\mathbb{C}^N)$ is induced by the following inner product associated to the function $\rho\in L^\infty(\mathbb{R}^N_{\text{perf}})$ representing the specific mass ($\rho(x)\geq\rho_0>0$ almost everywhere in $\mathbb{R}^N_{\text{perf}}$):

$$(u,v)\mapsto \int_{Y\setminus T}\rho u\cdot\bar{v}\,dx\,. \tag{13}$$

In this section, the bar denotes the complex conjugate.

The variational formulation of the Bloch cellular problem (12) is

$$\left\{\begin{array}{c} \text{find } \ \lambda\in\mathbb{R} \ \text{ and } \ u\in W_{\vec{k}}(\mathbb{R}^N_{\text{perf}}(T)) \text{ such that} \\ \displaystyle\int_{Y\setminus T}\mathbf{C}(s)\,\epsilon(u)\cdot\epsilon(\bar{v})\,dx = \lambda\int_{Y\setminus T}\rho u\cdot\bar{v}\,dx, \ \ \forall v\in W_{\vec{k}}(\mathbb{R}^N_{\text{perf}}(T)). \end{array}\right. \tag{14}$$

The following theorem, similar to Theorem 1, holds.

Theorem 2 *Provided differentiability properties of the elasticity tensor* $\mathbf{C} = \mathbf{C}(s)$ *with respect to a general material parameter* s *and assuming that the eigenvalues of problem (1) are simple, then the eigenvalues and the eigenvectors are differentiable with respect to* s*.*

The derivative of the eigenvalue $\lambda_l = \lambda_l(s)$ *is*

$$\frac{d\lambda_l}{ds}(s) = \int_{Y\setminus T} \frac{d\mathbf{C}}{ds}(s)\,\epsilon(u_l)\cdot\epsilon(\bar{u}_l)\,dx, \tag{15}$$

where the corresponding eigenvector u_l *is normalized in the* L^2_ρ *norm of* $L^2(\mathbb{R}^N_{perf}(T))$ *:* $\|u_l\|_{L^2_\rho} = 1$.

The derivative $\dfrac{du_l}{ds}$ *of the eigenvector* $u_l = u_l(s)$ *is the solution of the problem below:*

$$\left\{\begin{array}{l} \text{find}\quad \dfrac{du_l}{ds} \text{ in } W_{\vec{k}}(\mathbb{R}^N_{perf}(T)),\ \text{such that}\ \displaystyle\int_{Y\setminus T} \rho\frac{du_l}{ds}\bar{u}_l\,dx = 0\ \text{and} \\ \\ \displaystyle\int_{Y\setminus T} \mathbf{C}(s)\,\epsilon\Big(\frac{du_l}{ds}\Big)\cdot\epsilon(\bar{v})\,dx - \lambda_l\int_{Y\setminus T}\rho\frac{du_l}{ds}\cdot\bar{v}\,dx = \\ \dfrac{d\lambda_l}{ds}(s)\displaystyle\int_{Y\setminus T}\rho\,u_l\cdot\bar{v}\,dx - \int_{Y\setminus T}\frac{d\mathbf{C}}{ds}(s)\,\epsilon(u_l)\cdot\epsilon(\bar{v})\,dx, \quad \forall v\in W_{\vec{k}}(\mathbb{R}^N_{perf}(T)). \end{array}\right. \tag{16}$$

where $< u_l >^\perp$ *denotes the orthogonal complement of the eigenspace generated by* u_l*, with respect to the inner product (13).*

The derivative of a functional $F(\lambda_l(s), u_l(s))$ depending on the first n eigenvalues and on the corresponding n eigenvectors may be written in terms of n adjoint states p_l, as :

$$\begin{aligned} \frac{dF}{ds}(s) = &\sum_{l=1}^{n}\frac{\partial F}{\partial\lambda_l}\int_{Y\setminus T}\frac{d\mathbf{C}}{ds}(s)\,\epsilon(u_l(s))\cdot\epsilon(\bar{u}_l(s))\,dx \\ &-\sum_{l=1}^{n}\int_{Y\setminus T}\frac{d\mathbf{C}}{ds}(s)\,\epsilon(u_l(s))\cdot\epsilon(\bar{p}_l)\,dx. \end{aligned} \tag{17}$$

Each adjoint state p_l is the solution of the following adjoint problem, for $1\le l\le n$:

$$\left\{\begin{array}{l} \text{find } p_l\in W_{\vec{k}}(\mathbb{R}^N_{\text{perf}}(T)),\ \text{with}\ \displaystyle\int_{Y\setminus T}\rho p_l\cdot\bar{u}_l(s)dx = 0 \text{ such that} \\ \displaystyle\int_{Y\setminus T}\mathbf{C}(s)\,\epsilon(p_l)\cdot\epsilon(\bar{w})\,dx - \lambda_l(s)\int_{Y\setminus T}\rho\,p_l\cdot\bar{w}\,dx = \frac{\partial F}{\partial u_l}\,w, \\ \qquad\qquad \forall w\in W_{\vec{k}}(\mathbb{R}^N_{\text{perf}}(T)). \end{array}\right. \tag{18}$$

When trying to design acoustic filters or optical waveguides, as described in [SigmundJensen2003] or [BarbarosieNeves2004], one is interested in varying the wave vector. Typically, one is looking for a microstructure having a bandgap as large as possible. For instance, one may want to maximize the gap between

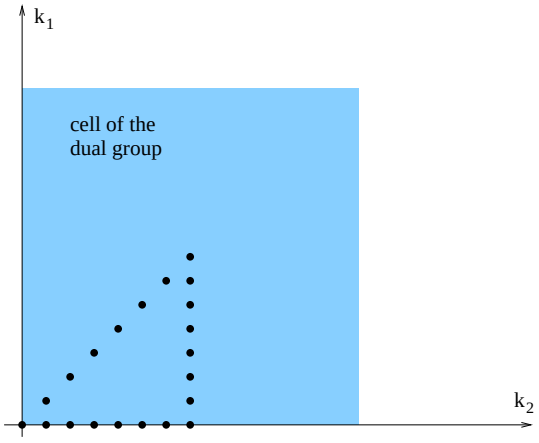

Figure 5: Brillouin triangle in the space of wave vectors

λ_2 and λ_3 (assuming that the eigenvalues are numbered in increasing order). In this case, the problem can be written as

$$\max_s \min_{\vec{k}} \big(\lambda_3(s,\vec{k}) - \lambda_2(s,\vec{k})\big) \tag{19}$$

Previous numerical approaches to this problem consider a fixed grid of wave vectors $\vec{k}_1, \vec{k}_2, \ldots$ (often placed on the boundary of the triangle shown in Figure 5) and seek to maximize the "worst case" :

$$\max_s \min_i \big(\lambda_3(s,\vec{k}_i) - \lambda_2(s,\vec{k}_i)\big) \tag{20}$$

This is equivalent to the maximization of the minimum between a finite family of functionals depending on the parameter s. However, it would be advantageous to be able to actively look for the worst case (the wave vector where λ_2 and λ_3 are closest), by using a steepest descent algorithm, for instance. Thus, it is interesting to try to differentiate the eigenvalues with respect to $\vec{k}$.

In the sequel, we shall focus on a more general formulation

$$\max_s \min_{\vec{k}} F\big(\lambda_l(s,\vec{k}), u_l(s,\vec{k})\big) \tag{21}$$

The derivative of the eigenvalues λ_l and of the eigenmodes u_l with respect to $\vec{k}$ will be computed and, by making use of the adjoint method, the derivative of the functional $F\big(\lambda_l(s,\vec{k}), u_l(s,\vec{k})\big)$ with respect to $\vec{k}$ will be explicitly obtained.

In order to achieve this, it is preferable to write $u(x)$ in the form of Bloch wave $u(x) = e^{i\langle\vec{k},x\rangle}\varphi(x)$. Thus, the cellular problem may be written as :

$$\left\{\begin{array}{c} \text{find} \quad \lambda \in \mathbb{R} \quad \text{and} \quad \varphi \in H^1_{\#}(\mathbb{R}^N_{\text{perf}}(T), \mathbb{C}^N) \text{ such that} \\ \displaystyle\int_{Y\setminus T} \mathbf{C}_{\alpha\beta\gamma\delta}\,(e^{i\langle\vec{k},x\rangle}\varphi_\alpha)_{,\beta}\,(e^{-i\langle\vec{k},x\rangle}\psi_\gamma)_{,\delta}\,dx = \lambda \int_{Y\setminus T} \rho\,\varphi_\alpha\psi_\alpha\,dx, \\ \forall \psi \in H^1_{\#}(\mathbb{R}^N_{\text{perf}}(T), \mathbb{C}^N). \end{array}\right. \tag{22}$$

The derivative of λ_l is obtained as

$$\frac{\partial \lambda_l}{\partial k_j} = 2Re \int_{Y\setminus T} i\, \mathbf{C}_{\alpha j\gamma\delta}\, u_{l\alpha}\, \bar{u}_{l\gamma,\delta}; \tag{23}$$

while the problem that defines the derivative of φ_l, $\dfrac{\partial \varphi_l}{\partial k_j}$, writes

$$\left\{ \begin{array}{c} \text{find} \quad \dfrac{\partial \varphi_l}{\partial k_j} \in H^1_{\#}(\mathbb{R}^N_{\text{perf}}(T), \mathbb{C}^N), \text{ such that } \displaystyle\int_{Y\setminus T} \rho\, \frac{\partial \varphi_{l\alpha}}{\partial k_j}\, \bar{\varphi}_{l\alpha}\, dx = 0 \text{ and} \\ \displaystyle\int_{Y\setminus T} \mathbf{C}_{\alpha\beta\gamma\delta} \left(e^{i\langle \vec{k},x\rangle} \frac{\partial \varphi_{l\alpha}}{\partial k_j}\right)_{,\beta} (e^{-i\langle \vec{k},x\rangle}\psi_\gamma)_{,\delta}\, dx - \lambda_l \int_{Y\setminus T} \rho\, \frac{\partial \varphi_{l\alpha}}{\partial k_j}\, \psi_\alpha\, dx = \\ \dfrac{\partial \lambda_l}{\partial k_j} \displaystyle\int_{Y\setminus T} \rho\, \varphi_{l\alpha}\, \psi_\alpha\, dx - \\ - \displaystyle\int_{Y\setminus T} i\, \mathbf{C}_{\alpha\beta\gamma\delta} \big(\delta_{\beta j}\, \varphi_{l\alpha}\, e^{i\langle \vec{k},x\rangle} (e^{-i\langle \vec{k},x\rangle}\psi_\gamma)_{,\delta} - \delta_{\delta j} (e^{i\langle \vec{k},x\rangle}\varphi_{l\alpha})_{,\beta}\, e^{-i\langle \vec{k},x\rangle}\psi_\gamma\big)\, dx, \\ \forall \psi \in H^1_{\#}(\mathbb{R}^N_{\text{perf}}(T), \mathbb{C}^N). \end{array} \right. \tag{24}$$

Consider the adjoint problem in the form

$$\left\{ \begin{array}{c} \text{find} \quad p_l \in H^1_{\#}(\mathbb{R}^N_{\text{perf}}(T), \mathbb{C}^N) \text{ with } \displaystyle\int_{Y\setminus T} \rho\, p_l\, \bar{\varphi}_l\, dx = 0 \text{ such that} \\ \displaystyle\int_{Y\setminus T} \mathbf{C}_{\alpha\beta\gamma\delta} \left(e^{i\langle \vec{k},x\rangle} p_{l\alpha}\right)_{,\beta} (e^{-i\langle \vec{k},x\rangle}\bar{\psi}_\gamma)_{,\delta}\, dx - \lambda \int_{Y\setminus T} \rho\, p_{l\alpha}\, \bar{\psi}_\alpha\, dx = \frac{\partial F}{\partial \varphi_\alpha}\, \psi_\alpha, \\ \forall \psi \in H^1_{\#}(\mathbb{R}^N_{\text{perf}}(T), \mathbb{C}^N). \end{array} \right. \tag{25}$$

Then, the derivative of the functional F may be expressed as

$$\begin{array}{c} \dfrac{dF}{dk_j}(\vec{k}) = \displaystyle\sum_{l=1}^{n} \frac{\partial F}{\partial \lambda_l} 2Re \int_{Y\setminus T} i\, \mathbf{C}_{\alpha j\gamma\delta}\, u_{l\alpha}\, \bar{u}_{l\gamma,\delta}\, dx - \\ - \displaystyle\sum_{l=1}^{n} \int_{Y\setminus T} i\, \mathbf{C}_{\alpha\beta\gamma\delta} \left(\delta_{\beta j}\, \varphi_{l\alpha}\, e^{i\langle \vec{k},x\rangle} (e^{-i\langle \vec{k},x\rangle}\, \bar{p}_{l\gamma})_{,\delta} - \right. \\ \left. - \delta_{\delta j} (e^{i\langle \vec{k},x\rangle}\varphi_{l\alpha})_{,\beta}\, e^{-i\langle \vec{k},x\rangle}\, \bar{p}_{l\gamma} \right) dx. \end{array} \tag{26}$$

This formula opens the possibility of solving (19) through a gradient based algorithm instead of considering a discrete set of wavevectors like in (20).

Acknowledgements We thank to Pedro Antunes for useful discussions. This work was supported by Fundação para a Ciência e a Tecnologia, PEst-OE/MAT/UI0209/2013.

References

[AllaireConca1998] G. Allaire, C. Conca, Bloch wave homogenization and spectral asymptotic analysis *J. Math. Pures et Appli.* 77:153-208, 1998.

[AllaireJouveT2004] G. Allaire, F. Jouve, A.-M. Toader, Structural optimization using sensitivity analysis and a level-set method *Journal of Computational Physics* 194(1):363-393, 2004.

[AllaireJouve2005] G. Allaire, F. Jouve A level-set method for vibration and multiple loads structural optimization *Computer Methods in Applied Mechanics and Engineering*, 194(30-33):3269-3290, 2005.

[AllaireJouve2008] G. Allaire, F. Jouve, Minimum stress optimal design with the level set method *Engineering Analysis with Boundary Elements* 32:909-918, 2008.

[AllaireJouveVG2011] G. Allaire, F. Jouve, N. Van Goethem, Damage and fracture evolution in brittle materials by shape optimization methods *Journal of Computational Physics* 230:5010-5044, 2011.

[AllaireDaponyF14] G. Allaire, Ch. Dapogny, P. Frey, Shape optimization with a level set based mesh evolution method *CMAME* 282:22-53, 2014.

[BarbarosieNeves2004] C. Barbarosie, M.M. Neves, Periodic structures for frequency filtering: analysis and optimization *Computers & Structures* 82:1399-1403, 2004.

[BarbarosieRochaT] C. Barbarosie, P. Rocha, A.-M. Toader, Smoothness of eigenvalues of structures depending on parameters, in preparation

[Cea1986] J. Céa, Conception optimale ou identification de formes. Calcul rapide de la dérivée directionnelle de la fonction coût *RAIRO, Modélisation Math. Anal.* 20:371-402, 1986.

[ConcaPlanchardV1995] C. Conca, J. Planchard, M. Vanninathan, Fluids and periodic structures, Wiley, 1995

[Hecht2012] F. Hecht, New development in FreeFem++, *J. Numer. Math.* 20:251-265, 2012.

[Hunter2007] J.D. Hunter, Matplotlib : A 2D graphics environement, *Computing in Science & Engineering* 9:90-95, 2007.

[Jouve] F. Jouve, `http://www.cmap.polytechnique.fr/~jouve/xd3d/`, web page

[KikuchiChengM1995] N. Kikuchi, H.-C. Cheng, Z.-D. Ma, Optimal Shape and Topology Design of Vibrating Structures *Advances in Structural Optimization. Solid Mechanics and Its Applications* 25:189-222, 1995.

[MuratTartar1985] F. Murat, L. Tartar, Calculus of variations and homogenization (translation from french Calcul des variations et homogénéisation 1985). *Topics in the mathematical modelling of composite materials, Progr. Nonlinear Differential Equations Appl.*, 31:139-173, Birkhäuser Boston, Boston, 1997.

[RousseletChenais1990] B. Rousselet, D. Chenais, Continuité et Différentiabilité d'Éléments Propres: Application à l'Optimisation de Structures. *Applied Mathematics and Optimization*, 22(1):27-59, 1990.

[OliveiraToaderV2012] S. Oliveira, A.-M. Toader, P. Vieira, Damage identification in a concrete dam by fitting measured modal parameters *Nonlinear Analysis: Real World Applications*, 13(6):2888-2899, 2012.

[SigmundJensen2003] O. Sigmund, J. Jensen, Systematic design of phononic band gap materials and structures by topology optimization *Philos. Trans. R. Soc. Lond. A* 361:1001-1019, 2003.

[xfig] `http://en.wikipedia.org/wiki/Xfig`, web page