# Numerical analysis of the neutron multigroup $SP\_N$ equations

Erell Jamelot, François Madiot

# Numerical analysis of the neutron multigroup $SP_N$ equations

Erell Jamelot *[1] and François Madiot †[2]

[1] DEN-Service de Thermohydraulique et de Mécanique des Fluides-STMF, LMSF, CEA, Université Paris-Saclay, F-91191, Gif-sur-Yvette, France.
[2] DEN-Service d'Etudes des Réacteurs et de Mathématiques Appliquées (SERMA), CEA,Université Paris-Saclay, F-91191, Gif-sur-Yvette, France.

July 20, 2020

**Abstract**

The multigroup neutron $SP_N$ equations, which are an approximation of the neutron transport equation, are used to model nuclear reactor cores. In their steady state, these equations can be written as a source problem or an eigenvalue problem. We study the resolution of those two problems with an $H^1$-conforming finite element method and a Discontinuous Galerkin method, namely the Symmetric Interior Penalty Galerkin method.

## 1 Introduction

The neutron transport equation describes the neutron flux density in a reactor core. It depends on 7 variables: 3 for the space, 2 for the motion direction, 1 for the energy (or the speed), and 1 for the time.
The energy variable is discretized using the multigroup theory [9]. In this method, the entire range of neutron energies is divided into $G$ intervals, called energy groups. In each energy group, the neutron flux density is lumped and all parameters are averaged. We denote by $\mathcal{I}_G := \{1, \cdots, G\}$, the set of energy group indices.
Concerning the motion direction, the $P_N$ transport equations are obtained by developing the neutron flux on the spherical harmonics from order 0 to order

*erell.jamelot@cea.fr
†francois.madiot@cea.fr

$N$. This approach is very time-consuming. The simplified $P_N$ ($SP_N$) transport theory [11] was developed to address this issue. The two fundamental hypotheses to obtain the $SP_N$ equations are that locally, the angular flux has a planar symmetry; and that the axis system evolves slowly. The neutron flux and the scattering cross sections are then developed on the Legendre polynomials. From a mathematical point of view, $SP_N$ equations correspond to tensorized $1D$ $P_N$ transport equations, so that some couplings are missing. Consequently, the $SP_N$ equations do not converge to transport equations. Nevertheless, they are commonly used by physicists since their resolution is cheap in terms of computational cost. The order $N$ is odd, and the number of $SP_N$ odd (resp. even) moments is $\widehat{N} := \frac{N+1}{2}$. We will denote by $\mathcal{I}_e$ (resp. $\mathcal{I}_o$) the subset of even (resp. odd) integers of the integer set $\{0, \cdots, N\}$.
Finally, the (motion direction and energy) discretization of the neutron flux is such that there are $\widehat{N} \times G$ even and odd moments of the neutron flux.
We will denote by $\phi = (\,(\phi_m^g)_{m\in\mathcal{I}_e}\,)_{g\in\mathcal{I}_G} \in \mathbb{R}^{\widehat{N}\times G}$ the set of functions containing, for all energy group $g$, the even moments of the neutron flux.
Likewise, we will denote by $\mathbf{p} = \big(\,(\,(p_{x,m}^g)_{m\in\mathcal{I}_o}\,)_{g\in\mathcal{I}_G}\big)_{x=1}^d \in \left(\mathbb{R}^{\widehat{N}\times G}\right)^d$ the set of functions containing the odd moments of the neutron flux.
Note that while modelling the core of a pressurized water reactor, the number of groups if such that $2 \leq G \lesssim 30$, physicists usually choose $N = 1$ or 3, more rarely $N = 5$.

## 2 Setting of the model

The reactor core is modelled by a bounded, connected and open subset $\mathcal{R}$ of $\mathbb{R}^d$, $d = 1, 2, 3$, having a Lipschitz boundary which is piecewise regular. The coefficients are piecewise regular, so that we split $\mathcal{R}$ into $\widetilde{N}$ open disjoint parts $(\mathcal{R}_i)_{i=1}^{\widetilde{N}}$ with Lipschitz, piecewise regular boundaries: $\overline{\mathcal{R}} = \cup_{i=1}^{\widetilde{N}} \overline{\mathcal{R}_i}$. For this reason, we will use the following space of piecewise regular functions:

$$\mathcal{P}W^{1,\infty}(\mathcal{R}) = \left\{ D \in L^\infty(\mathcal{R}) \mid \vec{\nabla}\, D|_{\mathcal{R}_i} \in (L^\infty(\mathcal{R}_i))^d, i = 1, \cdots, \widetilde{N} \right\}.$$

For a set of functions $\psi = (\phi_m^g)_{m,g} \in \mathbb{R}^{\widehat{N}\times G}$, we make the following abuse of notation: $\vec{\nabla}\, \psi = (\,(\partial_x \psi_m^g)_{m,g}\,)_{x=1}^d \in \left(\mathbb{R}^{\widehat{N}\times G}\right)^d$.

For a set of vector valued functions $\mathbf{q} = \left(\,\left((q_{x,m}^g)_{m,g}\right)_{x=1}^d\right) \in \left(\mathbb{R}^{\widehat{N}\times G}\right)^d$, we make the following abuse of notation:

$$\operatorname{div} \mathbf{q} = \big(\operatorname{div}(\,(q_{x,m}^g)_{x=1}^d\,)\big)_{m,g}, \quad \mathbf{q} \cdot \mathbf{p} = \left( \sum_{x=1}^d q_{x,m}^g\, p_{x,m}^g \right)_{m,g} \in \mathbb{R}^{\widehat{N}\times G}.$$

Let us use these notations: for $E \subset \mathbb{R}^d$, $L(E) = L^2(E)$; $L := L^2(\mathcal{R})$; $V := H_0^1(\mathcal{R})$; $V' := H^{-1}(\mathcal{R})$ its dual and $Q := H(\operatorname{div}, \mathcal{R})$. For $W = L(E)$, $L$, $V$ or

$Q$ we define the product space $\underline{W} := W^{\widehat{N}\times G}$ endowed with the following scalar product and associated norm:

$$(\mathtt{u},\mathtt{v})_{\underline{W}} = \sum_{g\in\mathcal{I}_G}\sum_{m\in\mathcal{I}e,o}(\mathtt{u}^g_m,\mathtt{v}^g_m)_W, \quad \|\mathtt{u}\|^2_{\underline{W}} = \sum_{g\in\mathcal{I}_G}\sum_{m\in\mathcal{I}e,o}\|\mathtt{u}^g_m\|^2_W. \tag{1}$$

We also set $\underline{V}' := (V')^{\widehat{N}\times G}$, $\underline{\mathbf{L}}(E) = (\underline{L}(E))^d$ and $\underline{L}^p(\cdot) = (L^p(\cdot))^{\widehat{N}\times G}$.
Let $\mathbf{q} \in \left(\mathbb{R}^{\widehat{N}\times G}\right)^d$ and $\mathbb{M} \in \left(\mathbb{R}^{\widehat{N}\times\widehat{N}}\right)^{G\times G}$. We set $\mathbf{q}_x = (q^g_{x,m})_{m,g}$ and we use the notation $\mathbb{M}\,\mathbf{q} = (\mathbb{M}\mathbf{q}_x)^d_{x=1}$.

Given a source term $S_f \in \underline{L}$, the multigroup $SP_N$ equations with zero-flux boundary conditions[1] read as coupled diffusion-like equations set in a mixed formulation:

$$\text{Solve in } (\phi,\mathbf{p}) \in \underline{V}\times\underline{Q}\,|\left\{\begin{array}{rcl} \mathbb{T}_o\,\mathbf{p} + \vec{\nabla}\,(\mathbb{H}\,\phi) & = & 0, \\ {}^t\mathbb{H}\,\mathrm{div}\,\mathbf{p} + \mathbb{T}_e\,\phi & = & S_f. \end{array}\right. \tag{2}$$

When $S_f$ depends on $\phi$, the steady state multigroup $SP_N$ equations read as the following generalized eigenproblem:

$$\text{Solve in } (\lambda,\phi,\mathbf{p}) \in \mathbb{R}^*\times\underline{V}\times\underline{Q}\,|\left\{\begin{array}{rcl} \mathbb{T}_o\,\mathbf{p} + \vec{\nabla}\,(\mathbb{H}\,\phi) & = & 0, \\ {}^t\mathbb{H}\,\mathrm{div}\,\mathbf{p} + \mathbb{T}_e\,\phi & = & \lambda^{-1}\mathbb{M}_f\phi. \end{array}\right. \tag{3}$$

The physical solution to Problem (3) corresponds to the eigenfunction associated to the smallest eigenvalue, which in addition is simple [7]. In neutronics, the *multiplication factor* $k_{eff} = \max_\lambda \lambda$ characterizes the physical state of the core reactor: if $k_{eff} = 1$: the nuclear chain reaction is self-sustaining; if $k_{eff} > 1$: the chain reaction is diverging; if $k_{eff} < 1$: the chain reaction vanishes.
The matrices $\mathbb{H}$, $\mathbb{T}_e$, $\mathbb{T}_o$, $\mathbb{M}_f \in \left(\mathbb{R}^{\widehat{N}\times\widehat{N}}\right)^{G\times G}$ are such that $\forall(g,g') \in \mathcal{I}_G\times\mathcal{I}_G$ ($\delta_{\cdot,\cdot}$ is the Kronecker symbol):

- $(\mathbb{H})_{g,g'} = \delta_{g,g'}\widehat{\mathbb{H}} \in \mathbb{R}^{\widehat{N}\times\widehat{N}}$, with $\forall(i,j) \in \{1,\cdots,\widehat{N}\}^2$, $\widehat{\mathbb{H}}_{i,j} = \delta_{i,j} + \delta_{i,j-1}$.
- $(\mathbb{T}_e)_{g,g} := \mathbb{T}^g_e \in \mathbb{R}^{\widehat{N}\times\widehat{N}}$ denotes the even removal matrix, such that:

  $$\mathbb{T}^g_e = diag\left(t_0\sigma^g_{r,0}, t_2\sigma^g_{r,2}, ...\right),$$

  $(\mathbb{T}_o)_{g,g} := \mathbb{T}^g_o \in \mathbb{R}^{\widehat{N}\times\widehat{N}}$ denotes the odd removal matrix, such that:

  $$\mathbb{T}^g_o = diag\left(t_1\sigma^g_{r,1}, t_3\sigma^g_{r,3}, ...\right),$$

  where $\forall m \in \mathcal{I}_{e,o}$, $\sigma^g_{r,m} := \sigma^g_t - \sigma^{g\to g}_{s,m}$, and $\forall m > 0$, $t_m > 0$.

  The coefficient $\sigma^g_t$ is the macroscopic total cross section of energy group $g$, and the coefficients $\sigma^{g\to g}_{s,m}$ denote the $P_N$ moments of the macroscopic self scattering cross sections from energy group $g$ to itself.

[1] ie: for $1 \le g \le G$, $m \in \mathcal{I}_e$, $(\phi^g_m)|_{\partial\mathcal{R}} = 0$.

- For $g' \neq g$:

  $(\mathbb{T}_e)_{g,g'} := -\mathbb{S}_e^{g' \to g} \in \mathbb{R}^{\widehat{N} \times \widehat{N}}$ denotes the even scattering matrix, such that:

  $$\mathbb{S}_e^{g' \to g} = diag\left(t_0 \sigma_{s,0}^{g' \to g}, t_2 \sigma_{s,2}^{g' \to g}, ...\right),$$

  $(\mathbb{T}_o)_{g,g'} := -\mathbb{S}_o^{g' \to g} \in \mathbb{R}^{\widehat{N} \times \widehat{N}}$ denotes the odd scattering matrix, such that:

  $$\mathbb{S}_o^{g' \to g} = diag\left(t_1 \sigma_{s,1}^{g' \to g}, t_3 \sigma_{s,3}^{g' \to g}, ...\right),$$

  where $\sigma_{s,m}^{g' \to g}$ are the $P_N$ moments of the macroscopic scattering cross sections from energy group $g'$ to energy group $g$.

- $(\mathbb{M}_f)_{g,g'} := \chi^g \mathbb{M}_f^{g'} \in \mathbb{R}^{\widehat{N} \times \widehat{N}}$ is such that $\mathbb{M}_f^{g'} \phi^{g'} = {}^t(\underline{\nu \sigma}_f^{g'} \phi_0^{g'}, 0, ...)$ where the coefficient $\underline{\nu \sigma}_f^{g'}$ is the product of the number of neutrons emitted per fission times the macroscopic fission cross section; and the coefficient $\chi_g$ is the fission spectrum of energy group $g$.

The coefficients of the matrices $\mathbb{T}_{e,o}$, $\mathbb{M}_f$ are supposed to be such that:

$$\left\{\begin{array}{ll} (0) & \forall\, g,\, g' \in \mathcal{I}_G, \forall\, m \in \mathcal{I}_{e,o} : \\ & (\sigma_{r,m}^g, \sigma_{s,m}^{g' \to g}, \underline{\nu \sigma}_f^g) \in \mathcal{P}W^{1,\infty}(\mathcal{R}) \times L^\infty(\mathcal{R}) \times L^\infty(\mathcal{R}). \\ (i) & \exists\, (\sigma_{r,(e,o)})_*,\, (\sigma^{r,(e,o)})^* > 0 \,|\, \forall\, g \in \mathcal{I}_G,\, \forall\, m \in \mathcal{I}_{e,o} : \\ & (\sigma_{r,(e,o)})_* \leq t_m \sigma_{r,m}^g \leq (\sigma^{r,(e,o)})^* \text{ a.e. in } \mathcal{R}. \\ (ii) & \exists\, (\underline{\nu \sigma}_f)^* > 0 \,|\, \forall\, g \in \mathcal{I}_G,\, 0 \leq \underline{\nu \sigma}_f^g \leq (\underline{\nu \sigma}_f)^* \text{ a.e. in } \mathcal{R} \text{ and } \exists\, g' \,|\, \underline{\nu \sigma}_f^{g'} \neq 0. \\ (iii) & \exists\, 0 < \varepsilon < \dfrac{1}{G-1} \,|\, \forall\, m \in \mathcal{I}_{e,o},\, \forall\, g, g' \in \mathcal{I}_G, g' \neq g, \\ & |\sigma_{s,m}^{g \to g'}| \leq \varepsilon \sigma_{r,m}^g \text{ a.e. in } \mathcal{R}. \end{array}\right. \tag{4}$$

It happens that the coefficient $\underline{\nu \sigma}_f^g$ vanishes in some regions.
Hypothesis 4$-(iii)$ is valid while modelling the core of a pressurized water reactor: the scattering cross-sections are weaker than the removal cross-sections of an order $0 < \varepsilon << 1$. Thus, the matrices ${}^t\mathbb{T}_{e,o}$ are strictly diagonally dominant matrices: they are invertible.
Let us set $\mathbb{D} = {}^t\mathbb{H}\, \mathbb{T}_o^{-1}\, \mathbb{H}$.
Problem 2 can be written as a set of coupled primal diffusion-like equations with single unknown $\phi \in \underline{V}$:

$$\text{Solve in } \phi \in \underline{V} \,|\, -\text{div}\left(\mathbb{D}\, \vec{\nabla}\, \phi\right) + \mathbb{T}_e\, \phi = S_f. \tag{5}$$

The variational formulation of (5) writes:

$$\text{Solve in } \phi \in \underline{V} \,|\, \forall \psi \in \underline{V} :\; c(\phi, \psi) = \ell(\psi), \tag{6}$$

where: $\left\{\begin{array}{lcl} c : \underline{V} \times \underline{V} & \to & \mathbb{R} \\ c(\phi, \psi) & = & (\mathbb{D}\, \vec{\nabla}\, \phi, \vec{\nabla}\, \psi)_{\underline{\mathbf{L}}} + (\mathbb{T}_e \phi, \psi)_{\underline{L}} \end{array}\right.$, and $\left\{\begin{array}{lcl} \ell : \underline{V} & \to & \mathbb{R} \\ \ell(\psi) & = & (S_f, \psi)_{\underline{L}} \end{array}\right.$.

**Theorem 1** *Suppose that $\mathbb{D}$ is positive definite. For a given source term $S_f \in \underline{L}$, it exists a unique $\phi \in \underline{V}$ that solves Problem 6. In addition, it holds: $\|\phi\|_{\underline{V}} \lesssim \|S_f\|_{\underline{L}}$.*

**Proof:** The bilinear form $c$ and the linear form $\ell$ are continuous and under the hypothesis on $\mathbb{D}$, the bilinear form $c$ is coercive: we can apply Lax-Milgram theorem to conclude. □

In the same way, Problem 3 can be written as:

$$\text{Solve in } (\lambda, \phi) \in \mathbb{R}^* \times \underline{V}\backslash\{0\} \mid -\operatorname{div}\left(\mathbb{D}\,\vec{\nabla}\,\phi\right) + \mathbb{T}_e\,\phi = \lambda^{-1}\mathbb{M}_f\phi. \tag{7}$$

The variational formulation of (7) writes:

$$\text{Solve in } (\lambda, \phi) \in \mathbb{R}^* \times \underline{V}\backslash\{0\} \,|\, \forall \psi \in \underline{V} :\, c(\phi,\psi) = \lambda^{-1}\ell_f(\phi,\psi), \tag{8}$$

where: $\left\{ \begin{array}{rcl} \ell_f : \underline{L} \times \underline{L} & \rightarrow & \mathbb{R} \\ \ell_f(\phi,\psi) & = & (\mathbb{M}_f\phi,\psi)_{\underline{L}} \end{array} \right.$.

**Theorem 2** *Suppose that $\mathbb{D}$ is positive definite. There exists a unique compact operator $T_f : \underline{L} \rightarrow \underline{L}$ such that $\forall(\phi,\psi) \in \underline{L} \times \underline{V}$: $c(T_f\phi,\psi) = \ell_f(\phi,\psi)$.*

**Proof:** The bilinear form $c$ is a continuous and under the hypothesis on $\mathbb{D}$, it is coercive onto $\underline{V} \times \underline{V}$. The bilinear form $\ell_f$ is a continuous onto $\underline{L} \times \underline{V}$. Finally, $\underline{V}$ is a subset of $\underline{L}$ with a compact embedding. We can then apply the work of Babuška and Osborn in [2]. □

Thus, the couple $(\phi, \lambda^{-1})$ is a solution to Problem 8 iff the couple $(\phi, \lambda)$ is an eigenpair of operator $T_f$. Moreover, Problem 8 admits a countable number of eigenvalues.

We propose first to derive conditions on the macroscopic cross sections so that Problems 5 and 7 are well-posed. Then we obtain a priori error estimates for a discretization performed with some $H^1$-conforming FEM and a Discontinuous Galerkin method, namely the Symmetric Interior Penalty Galerkin method (SIPG) [8, Chapter 4]. The outline is as follows: in Section 3, we exhibit some conditions so that the matrix $\mathbb{T}_o^{-1}$ and $\mathbb{T}_e$ are positive definite. Then we study the discretization of the source problem (5) in Section 5, and the discretization of the eigenproblem in Section 6. Finally, we perform in Section 7 a numerical study of convergence on a benchmark representative of a nuclear core.

# 3 Properties of $\mathbb{T}_e$ and $\mathbb{T}_o^{-1}$

Consider the diagonal matrix containing the even (resp. odd) removal macroscopic cross sections: $\mathbb{T}_{r,(e,o)} = diag(\mathbb{T}^1_{e,o}, \cdots, \mathbb{T}^G_{e,o})$. We split $\mathbb{T}_{e,o}$ so that: $\mathbb{T}_{e,o} = \mathbb{T}_{r,(e,o)}(\mathbb{I} - \varepsilon\mathbb{U}_{e,o})$, where $\mathbb{I} \in \left(\mathbb{R}^{\widehat{N}\times\widehat{N}}\right)^{G\times G}$ is the identity matrix, and:

$$\begin{array}{llcl} \forall g,\, g' \in \mathcal{I}_G,\, g' \neq g, & (\mathbb{U}_{e,o})_{g,g'} & = & diag\left(\left(\dfrac{\sigma^{g'\rightarrow g}_{s,m}}{\varepsilon\sigma^{g}_{r,m}}\right)_{m\in\mathcal{I}_{e,o}}\right) \in \mathbb{R}^{\widehat{N}\times\widehat{N}}; \\ \forall g \in \mathcal{I}_G, & (\mathbb{U}_{e,o})_{g,g} & = & 0 \in \mathbb{R}^{\widehat{N}\times\widehat{N}}. \end{array}$$

We have then: $\|\mathbb{U}_{e,o}\|_2 \lesssim \dfrac{\alpha_{s,(e,o)}}{\varepsilon}$ where: $\alpha_{s,(e,o)} := (G{-}1) \max\limits_{m\in\mathcal{I}_{e,o}} \max\limits_{g\neq g'\in\mathcal{I}_G} \sup\limits_{\vec{x}\in\mathcal{R}} \dfrac{|\sigma_{s,m}^{g'\to g}(\vec{x})|}{\sigma_{r,m}^{g}(\vec{x})}$.

Let us set $\alpha_{r,(e,o)} = \dfrac{(\sigma_{r,(e,o)})^*}{(\sigma_{r,(e,o)})_*} > 1$. We have the following properties.

**Property 3** *Suppose that* $\alpha_{s,e} < \dfrac{1}{\alpha_{r,e}}$. *The matrix* $\mathbb{T}_e$ *is such that:*

$$\forall X \in \mathbb{R}^{\widehat{N}\times G} \quad (\mathbb{T}_e X|X) \geq \tau_e \|X\|_2^2 \text{ where } \tau_e = (\sigma_{r,e})_* \left(1 - \alpha_{r,e}\alpha_{s,e}\right). \tag{9}$$

**Proof:** We have: $\forall X \in \mathbb{R}^{\widehat{N}\times G}$, $(\mathbb{T}_e X|X) = (\mathbb{T}_{r,e}X|X) - \varepsilon(\mathbb{U}_e X|\mathbb{T}_{r,e}X)$, so that: $(\,(\mathbb{T}_e X|X) \geq (\sigma_{r,e})_* - \varepsilon\|\mathbb{U}_e\|_2\,\|\mathbb{T}_{r,e}\|_2\,)\ \|X\|_2$, where $\|\mathbb{T}_{r,e}\|_2 \leq (\sigma_{r,e})^*$. $\square$

**Property 4** *Suppose that* $\alpha_{s,o} < \dfrac{1}{\alpha_{r,o}+1}$, *the matrix* $\mathbb{T}_o^{-1}$ *is such that:*

$$\forall X \in \mathbb{R}^{\widehat{N}\times G} \quad (\mathbb{T}_o^{-1}X|X) \geq \tau_o \|X\|_2^2 \text{ where } \tau_o = \frac{1}{(\sigma_{r,o})^*}\left(1 - \frac{\alpha_{r,o}\alpha_{s,o}}{1-\alpha_{s,o}}\right). \tag{10}$$

**Proof:** The Taylor expansion of $\mathbb{T}_o^{-1}$ writes: $\mathbb{T}_o^{-1} = \left(\mathbb{I} + \sum\limits_{l>0} \varepsilon^l \mathbb{U}_o^l\right)\mathbb{T}_{r,o}^{-1}$.

We get that $\forall X \in \mathbb{R}^{\widehat{N}\times G}$:

$$\begin{aligned}
\left(\mathbb{T}_o^{-1}X \,|\, X\right) &= \left(\mathbb{T}_{r,o}^{-1}X \,|\, X\right) + \sum_{l>0} \varepsilon^l \left(\mathbb{U}_o^l \mathbb{T}_{r,o}^{-1}X \,|\, X\right) \\
&\geq \frac{1}{(\sigma_{r,o})^*}\left(1 - \alpha_{r,o}\sum_{l>0}\varepsilon^l \|\mathbb{U}_o\|_2^l\right)\|X\|_2^2, \\
&\geq \frac{1}{(\sigma_{r,o})^*}\left(1 - \alpha_{r,o}\frac{\varepsilon\,\|\mathbb{U}_o\|_2}{1-\varepsilon\,\|\mathbb{U}_o\|_2}\right)\|X\|_2^2, \\
&\geq \frac{1}{(\sigma_{r,o})^*}\left(1 - \frac{\alpha_{r,o}\alpha_{s,o}}{1-\alpha_{s,o}}\right)\|X\|_2^2.
\end{aligned}$$

$\square$

Under assumptions of Properties 3 and 4 the matrices $\mathbb{T}_e$ and $\mathbb{T}_o^{-1}$ are positive definite. Moreover, one can show that $\|\mathbb{H}\vec{\nabla}\,\phi\|_{\underline{\mathbf{L}}} \gtrsim \|\vec{\nabla}\,\phi\|_{\underline{\mathbf{L}}}$ [12]. We infer that the matrix $\mathbb{D}$ is positive definite and that there exists a constant $C_{\mathbb{D}} > 0$ such that for all $\xi \in \mathbb{R}^{\widehat{N}\times G}$,

$$(\mathbb{D}\,\xi|\mathbb{D}\,\xi) \leq C_{\mathbb{D}}\|\xi\|_2^2. \tag{11}$$

From now on, we suppose that this property holds.

## 4 Discretizations

Let $\mathcal{T}_h$ be a shape-regular mesh of $\mathcal{R}$, with mesh size $h$. We denote by $K$ its elements and $F$ its facets. To simplify the presentation, we assume that the meshes are such that in every element, the cross-sections are regular. We define by $\mathcal{F}_h^i$ the set of interior faces of $\mathcal{T}_h$, $\mathcal{F}_h^b$ the set of boundary facets and $\mathcal{F}_h = \mathcal{F}_h^i \cup \mathcal{F}_h^b$. We denote by $N_\partial$ the maximum number of mesh faces composing the boundary of mesh elements

$$N_\partial := \max_{K\in\mathcal{T}_h} \mathrm{Card}\{F \in \mathcal{F}_h, F \subset \partial K\}.$$

We will first consider an $H^1$-conforming finite element method (FEM). For $k \in \mathbb{N}^*$, $V_h^k \subset V$ and $\underline{V}_h^k \subset \underline{V}$ are the finite dimension spaces defined by:

$$V_h^k = \{v_h \in V, \forall K \in \mathcal{T}_h, v_h|_K \in \mathbb{P}_k\}, \quad \underline{V}_h^k := (V_h^k)^{\widehat{N}\times G}.$$

The discrete variational formulation associated to Problem (6) writes:

$$\text{Solve in } \phi_h \in \underline{V}_h^k \,|\, \forall \psi_h \in \underline{V}_h^k : \, c(\phi_h, \psi_h) = \ell(\psi_h), \tag{12}$$

Similarly, the discrete variational formulation associated to Problem (7) writes:

$$\text{Solve in } (\lambda_h, \phi_h) \in \mathbb{R}^* \times \underline{V}_h^k\backslash\{0\} \,|\, \forall \psi \in \underline{V}_h^k : \, c(\phi_h, \psi_h) = \lambda_h^{-1} \ell_f(\phi_h, \psi_h). \tag{13}$$

Then, we will consider a non-conforming FEM. We define the broken spaces:

$$V_{\mathrm{NC}} = \left\{v \in L^2(\mathcal{R}) \,|\, \forall K \in \mathcal{T}_h, \, v \in H^1(K)\right\}, \quad \underline{V}_{\mathrm{NC}} = (V_{\mathrm{NC}})^{\widehat{N}\times G}.$$

For $(\phi, \psi) \in \underline{V}_{\mathrm{NC}} \times \underline{V}_{\mathrm{NC}}$, and $\mathbb{T} \in \mathbb{R}^{\widehat{N}\times G}$, we set:

$$\left(\mathbb{D}\,\vec{\nabla}_h \phi, \vec{\nabla}_h \psi\right)_{\mathcal{T}_h} = \sum_{K\in\mathcal{T}_h} \left(\mathbb{D}\,\vec{\nabla}\,\phi, \vec{\nabla}\,\psi\right)_{\underline{\mathbf{L}}(K)}, \quad \text{and} \quad \left\|\vec{\nabla}_h \psi\right\|_{\mathcal{T}_h} = \left(\vec{\nabla}_h \psi, \vec{\nabla}_h \psi\right)_{\mathcal{T}_h}^{1/2}.$$

For $F \in \mathcal{F}_h^i$ such that $F = \partial K_1 \cap \partial K_2$, we define the average $\{\mathbb{D}\,\vec{\nabla}_h \psi\}$ and the jump $[\![\psi]\!]$ as:

$$\begin{aligned}
\{\mathbb{D}\,\vec{\nabla}_h \psi\}|_F &= \frac{1}{2}\left((\mathbb{D}_1\,\vec{\nabla}\,\psi_1)|_F + (\mathbb{D}_2\,\vec{\nabla}\,\psi_2)|_F\right) \in \left(\mathbb{R}^{\widehat{N}\times G}\right)^d, \\
[\![\psi]\!]|_F &= \psi_1|_F\,\mathbf{n}_1 + \psi_2|_F\,\mathbf{n}_2 \in \left(\mathbb{R}^{\widehat{N}\times G}\right)^d.
\end{aligned}$$

where $\mathbf{n}_i$ is is the unit outward normal to $K_i$ at face $F$ and $\mathbb{D}_i = \mathbb{D}|_{K_i}$, $\psi_i = \psi|_{K_i}$.

For $F \in \mathcal{F}_h^b$ such that $F \in K$, we set $\{\mathbb{D}\,\vec{\nabla}_h \psi\}|_F = \mathbb{D}|_K\,\vec{\nabla}\,\psi|_K$ and $[\![\psi]\!]|_F = (\psi_K)|_F\,\mathbf{n}$, where $\psi_K = \psi|_K$ and $\mathbf{n}$ is the unit outward normal to $K$ at face $F$.

For $k \in \mathbb{N}^*$, $V_{h,\mathrm{NC}}^k \subset H^1(\mathcal{T}_h)$ and $\underline{V}_{h,\mathrm{NC}}^k$ are the finite dimension spaces defined by:

$$V_{h,\mathrm{NC}}^k = \{v_h \in L^1(\mathcal{R}); \forall K \in \mathcal{T}_h, v_h|_K \in \mathbb{P}_k\}, \quad \underline{V}_{h,\mathrm{NC}}^k := \left(V_{h,\mathrm{NC}}^k\right)^{\widehat{N}\times G}.$$

For $\phi_h,\ \psi_h \in \underline{V}^k_{h,\mathrm{NC}}$, we set: $\left(\{\mathbb{D}\,\vec{\nabla}_h\phi_h\}, [\![\psi_h]\!]\right)_{\mathcal{F}^i_h} = \sum_{F\in\mathcal{F}^i_h} \left(\{\mathbb{D}\,\vec{\nabla}_h\phi_h\}, [\![\psi_h]\!]\right)_{\underline{\mathbf{L}}(F)}$.

Let us set

$$c_h(\phi_h,\psi_h) = c_{\mathcal{T}_h}(\phi_h,\psi_h) + c_{\mathcal{F}_h}(\phi_h,\psi_h), \tag{14}$$

with

$$\begin{aligned} c_{\mathcal{T}_h}(\phi_h,\psi_h) &= \left(\mathbb{D}\,\vec{\nabla}_h\phi_h, \vec{\nabla}_h\psi_h\right)_{\mathcal{T}_h} + (\mathbb{T}_e\phi_h,\psi_h)_{\underline{L}}\,, \\ c_{\mathcal{F}_h}(\phi_h,\psi_h) &= \sum_{F\in\mathcal{F}_h} \frac{\alpha}{h_F}\left([\![\phi_h]\!],[\![\psi_h]\!]\right)_{\underline{\mathbf{L}}(F)} - \left(\{\mathbb{D}\,\vec{\nabla}_h\psi_h\},[\![\phi_h]\!]\right)_{\mathcal{F}^i_h} - \left(\{\mathbb{D}\,\vec{\nabla}_h\phi_h\},[\![\psi_h]\!]\right)_{\mathcal{F}^i_h}, \end{aligned}$$

where $\alpha$ is a stabilization parameter.

The Symmetric Interior Penalty Galerkin method (SIPG) associated to Problem (6) writes:

$$\text{Solve in } \phi_h \in \underline{V}^k_{h,\mathrm{NC}} \,|\, \forall\psi_h \in \underline{V}^k_{h,\mathrm{NC}} :\ c_h(\phi_h,\psi_h) = \ell(\psi_h). \tag{15}$$

Similarly, the SIPG method associated to Problem (8) writes:

$$\begin{array}{c} \text{Solve in } (\lambda_h,\phi_h) \in \mathbb{R}^* \times \underline{V}^k_{h,\mathrm{NC}}\backslash\{0\} \,|\, \forall\psi_h \in \underline{V}^k_{h,\mathrm{NC}} : \\ c_h(\phi_h,\psi_h) = \lambda_h^{-1}\ell_f(\phi_h,\psi_h). \end{array} \tag{16}$$

# 5 The source problem

## 5.1 Conforming discretization

**Theorem 5** *Suppose that there exists $r_{\max}$ in $[0,1]$ such that $\forall r \in [0, r_{\max}[$, $\phi \in \left(H^{1+r}(\mathcal{R})\right)^{\widehat{N}\times G}$ ([6], Proposition 1). Let us set $\mu = \min(r_{\max}, k)$. The solution of (12), $\phi_h$ is such that: $\|\phi-\phi_h\|_{\underline{V}} \lesssim h^\mu \|S_f\|_{\underline{L}}$ and $\|\phi-\phi_h\|_{\underline{L}} \lesssim h^{2\mu}\|S_f\|_{\underline{L}}$.*

**Proof:** From Céa's lemma and Aubin-Nitsche lemma as detailed in ([10], §2.3).□

## 5.2 SIPG discretization

**Assumption 5.1 (Regularity of exact solution and space $V^\star$)** *Let us denote by $W^{2,p}(\mathcal{T}_h)$ the broken Sobolev space spanned by those functions $v$ such that for all $K \in \mathcal{T}_h$, $v|_K \in W^{2,p}(K)$. We set $\underline{W}^{2,p}(\mathcal{T}_h) = \left(W^{2,p}(\mathcal{T}_h)\right)^{\widehat{N}\times G}$. We assume that $d \geq 2$ and that there is $2d/(d+2) < p \leq 2$ such that, for the exact solution $\phi \in \underline{V}^\star := \underline{V} \cap \underline{W}^{2,p}(\mathcal{T}_h)$. This holds for our assumptions on the coefficients, which are piecewise constant with respect to the triangulation [15].*

This assumption requires $p > 1$ for $d = 2$ and $p > 6/5$ for $d = 3$. In particular, we observe that, in two space dimensions, $\phi \in \underline{W}^{2,p}(\mathcal{T}_h)$ in polygonal domains. Moreover, using Sobolev embeddings [4, Sect. IX.3], this implies

$$\phi \in \left(H^{1+\alpha_p}(\mathcal{R})\right)^{\widehat{N}\times G}, \quad \alpha_p = \frac{d+2}{2} - \frac{d}{p} > 0.$$

We state the following lemma [8, Lemma 1.46, p.27].

**Lemma 6** *Suppose that $(\mathcal{T}_h)_h$ is a shape- and contact-regular mesh sequence. Then, we have for all $h > 0$:*

$$\forall \psi_h \in \underline{V}^k_{h,\mathrm{NC}}, \forall K \in \mathcal{T}_h, \forall F \in \partial K, \quad h_K^{1/2} \|\psi_h\|_{\underline{L}^2(F)} \le C_{tr} \|\psi_h\|_{\underline{L}^2(K)}, \tag{17}$$

*where $h_K$ is the diameter of element $K$.*

We aim at asserting the discrete coercivity using the following norm:

$$\forall \psi_h \in \underline{V}^k_{h,\mathrm{NC}}, \quad |||\psi_h|||^2_{sip} := c_{\mathcal{T}_h}(\psi_h, \psi_h) + \|\psi_h\|^2_J$$

with the jump semi-norm

$$\|\psi_h\|^2_J := \sum_{F \in \mathcal{F}_h} \frac{1}{h_F} \|[\![\psi_h]\!]\|^2_{\underline{\mathbf{L}}(F)} .$$

Under assumption (4), there exists $\beta > 0$ we have for all $\psi_h \in \underline{V}^k_{h,\mathrm{NC}}$

$$c_{\mathcal{T}_h}(\psi_h, \psi_h) \ge \beta \left( \left\| \vec{\nabla}_h \psi_h \right\|^2_{\mathcal{T}_h} + \|\psi_h\|^2_{\underline{L}} \right), \tag{18}$$

so that

$$|||\psi_h|||^2_{sip} \ge \beta \left( \left\| \vec{\nabla}_h \psi_h \right\|^2_{\mathcal{T}_h} + \|\psi_h\|^2_{\underline{L}} + \|\psi_h\|^2_J \right).$$

**Lemma 7 (Discrete coercivity)** *Let $\underline{\alpha} := C^2_{tr} N_\partial \frac{C_\mathbb{D}}{\beta}$ where*

- *$C_{tr}$ results from the discrete trace inequality* (17)*,*
- *$N_\partial$ is defined in Section 4,*
- *$C_\mathbb{D}$ is defined in* (11)*.*

*For all $\alpha \ge \underline{\alpha}$, the SIP bilinear form defined by* (14) *is coercive on $\underline{V}^k_{h,\mathrm{NC}}$ with respect to the $|||\cdot|||_{sip}$-norm, i.e.,*

$$c_h(\psi_h, \psi_h) \ge C_\alpha |||\psi_h|||^2_{sip},$$

*with* $C_\alpha := \left( \alpha - C^2_{tr} N_\partial \frac{C_\mathbb{D}}{\beta} \right) \min \left\{ \frac{1}{2}, \beta \left( \alpha + C^2_{tr} N_\partial \frac{C_\mathbb{D}}{\beta} \right)^{-1} \right\}.$

**Proof:** We follow the proof of [8, Lemma 4.12]. For all $\psi_h \in \underline{V}^k_{h,\mathrm{NC}}$,

$$\begin{aligned}
&c_h(\psi_h, \psi_h) \\
&= c_{\mathcal{T}_h}(\psi_h, \psi_h) + c_{\mathcal{F}_h}(\psi_h, \psi_h) \\
&= c_{\mathcal{T}_h}(\psi_h, \psi_h) + \sum_{F \in \mathcal{F}_h} \frac{\alpha}{h_F} \|[\![\psi_h]\!]\|^2_{\underline{\mathbf{L}}(F)} - 2 \left( \{\mathbb{D} \vec{\nabla}_h \psi_h\}, [\![\psi_h]\!] \right)_{\mathcal{F}^i_h} \\
&\ge c_{\mathcal{T}_h}(\psi_h, \psi_h) + \alpha \|\psi_h\|^2_J - 2 C_{tr} (N_\partial)^{1/2} \left\| \mathbb{D} \vec{\nabla}_h \psi_h \right\|_{\mathcal{T}_h} \|\psi_h\|_J
\end{aligned}$$

where we used Cauchy-Schwarz and Lemma 6 in the last line. Using the inequality $2ab \leq \varepsilon a + \varepsilon^{-1} b$ for any $\varepsilon > 0$, we obtain

$$\begin{aligned} 2C_{tr}(N_\partial)^{1/2} \left\| \mathbb{D}\vec{\nabla}_h \psi_h \right\|_{\mathcal{T}_h} \|\psi_h\|_J &\leq \varepsilon C_{tr}^2 N_\partial \left\| \mathbb{D}\vec{\nabla}_h \psi_h \right\|_{\mathcal{T}_h}^2 + \varepsilon^{-1} \|\psi_h\|_J^2 \\ &\leq \varepsilon C_{tr}^2 N_\partial C_{\mathbb{D}} \left\| \vec{\nabla}_h \psi_h \right\|_{\mathcal{T}_h}^2 + \varepsilon^{-1} \|\psi_h\|_J^2. \end{aligned}$$

Using (18), we obtain that there exists a constant $\beta > 0$ such that

$$\begin{aligned} & c_h(\psi_h, \psi_h) \\ & \geq \beta(1 - \varepsilon\underline{\alpha}) \left\| \vec{\nabla}_h \psi_h \right\|_{\mathcal{T}_h}^2 + \beta \|\psi_h\|_{\underline{L}}^2 + (\alpha - \varepsilon^{-1}) \|\psi_h\|_J^2. \end{aligned}$$

Choosing $\varepsilon = 2\left(\alpha + \underline{\alpha}\right)^{-1}$ yields the assertion. □

Thus, it only remains to prove boundedness. To this purpose, we need to define $\underline{V}^{\star,h} = \underline{V}^{\star} + \underline{V}_{h,\mathrm{NC}}^{k}$ and the following norm

$$|||\psi|||_{sip,\star} := \left( |||\psi|||_{sip}^p + \sum_{K \in \mathcal{T}_h} h_K^{1+\gamma_p} \left\| \vec{\nabla}\psi|_K \cdot \mathbf{n}_K \right\|_{\underline{L}^p(\partial K)} \right)^{1/p},$$

where $\gamma_p = \dfrac{d(p-2)}{2}$ and $\mathbf{n}_K$ is the unit outward normal to $K$. Following [8, Section 4.2], we obtain the following results.

**Lemma 8 (Boundedness)** *There is $C_{bnd}$ , independent of $h$, such that for all $(\phi, \psi_h) \in \underline{V}^{\star,h} \times \underline{V}_h$*

$$c_h(\phi, \psi_h) \leq C_{bnd} |||\phi|||_{sip,\star} |||\psi_h|||_{sip}$$

**Theorem 9 (Convergence)** *Suppose that there exists $r_{\max}$ in $(0, 1]$ such that $\forall r \in [0, r_{\max}]$, $\phi \in \left(H^{1+r}(\mathcal{R})\right)^{\widehat{N} \times G}$ ([6], Proposition 1). Then the solution of (15), $\phi_h$ is such that:*

$$|||\phi - \phi_h|||_{sip} \lesssim C \inf_{\psi_h \in \underline{V}_{h,\mathrm{NC}}} |||\phi - \psi_h|||_{sip,\star},$$

*where $C$ is a constant independent of $h$. Moreover, under Assumption 5.1, there holds*

$$|||\phi - \phi_h|||_{sip} \leq C |\phi|_{\underline{W}^{2,p}(\mathcal{T}_h)} h^\mu,$$

*where $\mu = r_{\max}$, $C$ is a constant independent of $h$ and $p$ is such that $\mu = \dfrac{d+2}{2} - \dfrac{d}{p}$.*

**Theorem 10 ($L^2$-norm estimate)** *Suppose that there exists $r_{\max}$ in $(0, 1]$ such that $\forall r \in [0, r_{\max}]$, $\phi_m^g \in H^{1+r}(\mathcal{R})$ ([6], Proposition 1). Under Assumption 5.1, the solution of (15), $\phi_h$ is such that: $\|\phi - \phi_h\|_{\underline{L}} \lesssim h^{2\mu} \|S_f\|_{\underline{L}}$, where $\mu = r_{\max}$.*

**Proof:** We apply the Aubin-Nitsche similarly as in [8, Theorem 4.25]. □

# 6 The eigenproblem

## 6.1 Conforming discretization

**Theorem 11** *Let $\mu$ be the regularity of the eigenfunction $\varphi$ associated to $\lambda$, and $\omega = \min(\mu, k)$. Let $\lambda_h$ be the discrete eigenvalue associated to Problem* (13). *The following a priori error estimate holds: $|\lambda - \lambda_h| \lesssim h^{2\omega}$.*

**Proof:** As in the continuous case (Theorem 2), since the discretization is conforming, there exists a unique compact operator $T_h : \underline{V}_h^k \to \underline{V}_h^k$ such that $\forall(\phi_h, \psi_h) \in \underline{V}_h^k \times \underline{V}_h^k$: $c(T_h\phi_h, \psi_h) = \ell_f(\phi_h, \psi_h)$. According to Thm. 5, the sequence of the operators $(T_h)_h$ is pointwise converging towards $T$. As $T_h$ and $T$ are a compact operators, the sequence of the operators $(T_h)_h$ is then converging in $\mathscr{L}(\underline{V})$ towards $T$: $\|T_h - T\|_{\mathscr{L}(\underline{V})} \to 0$. The norm convergence guarantees that there is no spectral pollution (see [17]). Morevover, we can apply Theorem 8.3 in [2] to state the error estimate on the eigenvalue. We remark that $(\mathbb{M}_f\phi, \phi)_{\underline{L}}$ is a norm over $\underline{V}_\lambda := \{\phi \in \underline{V} \,|\, \forall\psi \in \underline{V},\, c(\phi, \psi) = \lambda\ell_f(\phi, \psi)\}$ [12, Section 5.2.2 p. 78]. $\square$

## 6.2 SIPG discretization

We recall that, in this section, we work under the assumption 5.1.

**Theorem 12** *Let $\mu$ be the regularity of the eigenfunction $\varphi$ associated to $\lambda$, and $\omega = \min(\mu, k)$. Let $\lambda_h$ be the discrete eigenvalue associated to Problem* (16). *The following a priori error estimate holds: $|\lambda - \lambda_h| \lesssim h^{2\omega}$.*

**Proof:** We apply the theory developed in [1]. The proof is decomposed as follows. We first show that there is no spectral pollution. Then, we derive the error estimate.

Let $E : \underline{V} + \underline{V}_{h,\mathrm{NC}}^k \to \underline{V} + \underline{V}_{h,\mathrm{NC}}^k$ be the continuous spectral projector relative to $\lambda$ defined by

$$E = \frac{1}{2\pi i} \int_\Gamma \left(z - T|_{\underline{V} + \underline{V}_{h,\mathrm{NC}}^k}\right)^{-1} dz,$$

where $\Gamma$ is a circle in the complex plane centred at $\lambda$ which lies in $\rho(T|_{\underline{V} + \underline{V}_{h,\mathrm{NC}}^k})$ and encloses no other points of $\sigma(T|_{\underline{V} + \underline{V}_{h,\mathrm{NC}}^k})$. The absence of spectral pollution relies on two properties. First, using interpolation results [8, Assumption 4.31] we have for all $\phi \in E(\underline{V} + \underline{V}_{h,\mathrm{NC}}^k)$,

$$\inf_{\psi_h \in \underline{V}_{h,\mathrm{NC}}^k} |||\phi - \psi_h|||_{sip} \leq Ch^\mu,$$

where $C$ is a constant independent of $h$. Second, we have for all $\phi_h \in \underline{V}^k_{h,\mathrm{NC}}$,

$$\begin{aligned} |||(T - T_h)\phi_h|||_{sip} &\le Ch^\mu |T\phi_h|_{\underline{W}^{2,p}(\mathcal{T}_h)} \\ &\le Ch^\mu \|T\phi_h\|_{\left(H^{1+\alpha_p}(\mathcal{R})\right)^{\widehat{N}\times G}} \\ &\le Ch^\mu \|\phi_h\|_{\underline{L}} \\ &\le Ch^\mu |||\phi_h|||_{sip}, \end{aligned}$$

where we used Theorem 9 in the second line and regularity results [15] in the third line. Applying [1, Theorem 3.7], we obtain that there is no spectral pollution.

Moreover, we apply [1, Theorem 3.14] to state the error estimate on the eigenvalue,

$$|\lambda - \lambda_h| \le C\delta_h \delta_{*,h},$$

where

$$\begin{aligned} \delta_h &= \gamma_h + \left|\left|\left|(T - T_h)|_{E(\underline{V}+\underline{V}^k_{h,\mathrm{NC}})}\right|\right|\right|_{sip} \\ \delta_{*,h} &= \gamma_{*,h} + \left|\left|\left|(T_* - T_{*,h})|_{E(\underline{V}+\underline{V}^k_{h,\mathrm{NC}})}\right|\right|\right|_{sip}, \end{aligned}$$

with

$$\begin{aligned} \gamma_h &= \delta(E(V + \underline{V}^k_{h,\mathrm{NC}}), \underline{V}^k_{h,\mathrm{NC}}), \\ \gamma_{*,h} &= \delta(E_*(V + \underline{V}^k_{h,\mathrm{NC}}), \underline{V}^k_{h,\mathrm{NC}}), \end{aligned}$$

where

$$\delta(Y, Z) = \sup_{y\in Y, |||y|||_{sip}=1} \left( \inf_{z\in Z} |||y - z|||_{sip} \right),$$

and $E_* : \underline{V} + \underline{V}^k_{h,\mathrm{NC}} \to \underline{V} + \underline{V}^k_{h,\mathrm{NC}}$ is the continuous spectral projector of the adjoint operator $T_*|_{\underline{V}+\underline{V}^k_{h,\mathrm{NC}}}$ relative to $\bar{\lambda}$.

Using again elliptic regularity results [15] and Theorem 9, we obtain

$$\begin{aligned} \left|\left|\left|(T - T_h)|_{E(\underline{V}+\underline{V}^k_{h,\mathrm{NC}})}\right|\right|\right|_{sip} &\le Ch^\mu \\ \left|\left|\left|(T_* - T_{*,h})|_{E(\underline{V}+\underline{V}^k_{h,\mathrm{NC}})}\right|\right|\right|_{sip} &\le Ch^\mu. \end{aligned}$$

Using elliptic regularity results, we get

$$\|\varphi\|_{\left(H^{1+\alpha_p}(\mathcal{R})\right)^{\widehat{N}\times G}} \le C\|\varphi\|_{\underline{L}} \le C\|\varphi\|_{\underline{V}}.$$

Applying Theorem 9, we infer that

$$\gamma_h \leq Ch^{\mu}$$
$$\gamma_{*,h} \leq Ch^{\mu}.$$

This concludes the proof. □

# 7 Numerical Results

We consider the test case Model 2, case 1 from the benchmark of Takeda and Ikeda [19]. The geometry of the core is three-dimensional and the domain is $\{(x, y, z) \in \mathbb{R}^3, 0 \leq x \leq 140\ cm; 0 \leq y \leq 140\ cm; 0 \leq z \leq 150\ cm\}$. This test is defined with 4 energy groups, isotropic scattering and vacuum boundary conditions. Figure 1 represents the cross-sectional geometry on the plane $z = 75\ cm$.

Since the scattering is isotropic, the $SP_3$ formulation can easily be reformulated as a multigroup diffusion problem with 8 energy groups and an isotropic albedo boundary condition [3]. We then made the computations with the PRIAM solver from the code CRONOS2 [14] for the conforming case and with the MINARET solver [13] from the APOLLO3® code [18] for the SIPG discretization.

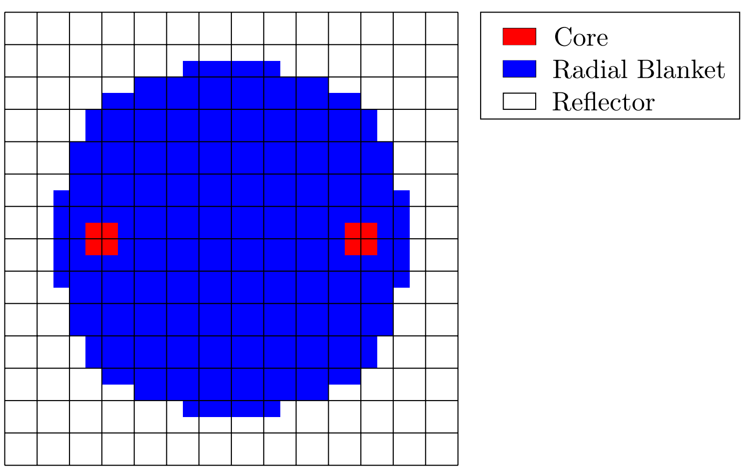


Figure 1: Cross-sectional view of the core ($z = 75\ cm$).

In Figure 2, we consider the convergence of the fundamental mode where we used the $SP_3$ formulation with $Q^1$ finite elements and a regular cartesian mesh of size $h$. The approximated order of convergence is 2.22.

In Figure 3, we consider the convergence of the fundamental mode for different the $SP_N$ formulations with discontinuous $P^1$ finite elements and a prismatic mesh of size $h$. The approximated orders of convergence are given in Table 1.

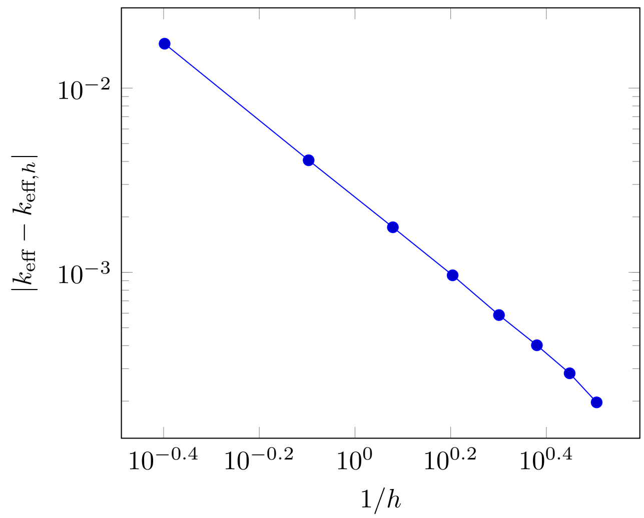


Figure 2: Error on the discrete eigenvalue for the $SP_3$ formulation with $Q^1$ finite elements

| $SP_3$ | $SP_5$ | $SP_7$ |
|---|---|---|
| 1.88 | 1.96 | 1.92 |

Table 1: Approximated order of convergence associated to Figure 3

# 8 Conclusion

We did the numerical analysis of the approximation with an $H^1$-conforming finite element method of the neutron multigroup $SP_N$ equations. We also studied the numerical analysis of the approximation with the Symmetric Interior Penalty Galerkin method of the neutron multigroup $SP_N$ equations. We then illustrated numerically the convergence results on a benchmark representative of a nuclear core. Those results can be extended to a mixed finite element method, see [5] for the diffusion case with an $H^1$-conforming finite element method.

# Acknowledgements

The authors gratefully acknowledge P. Ciarlet for fruitful discussions.

# References

[1] A. Alonso and A. D. Russo. Spectral approximation of variationally-posed eigenvalue problems by nonconforming methods. *Journal of Computational*

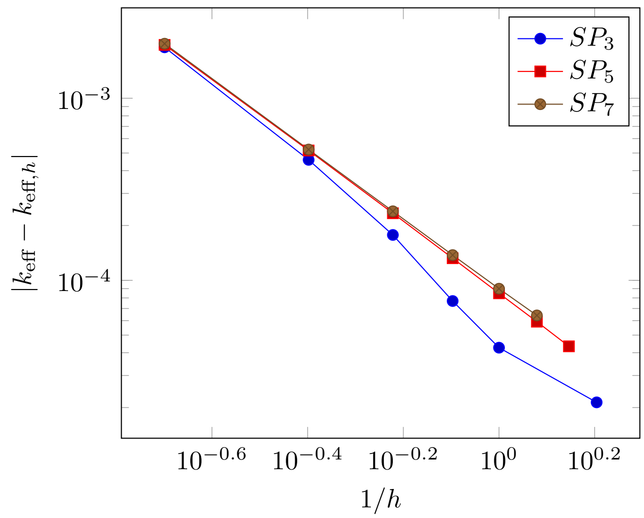


Figure 3: Error on the discrete eigenvalue for the $SP_3$ formulation with discontinuous linear finite elements

*and Applied Mathematics*, 223(1):177–197, 2009.

[2] I. Babuska and J. Osborn. Eigenvalue problems, in handbook of numerical analysis, vol. II. *PG Ciarlet and JL Lions Editions, Elsevier science Publishers BV (North-Holland)*, pages 645–785, 1991.

[3] A.-M. Baudron and J.-J. Lautard. Simplified $\mathrm{P_N}$ transport core calculations in the Apollo3 system. *International Conference on Mathematics and Computational Methods Applied to Nuclear Science and Engineering (M&C 2011)*, Rio de Janeiro, RJ, Brazil, May 8-12, 2011.

[4] H. Brézis. Analyse fonctionnelle, théorie et applications, 1983.

[5] P. Ciarlet, L. Giret, E. Jamelot, and F.D. Kpadonou. Numerical analysis of the mixed finite element method for the neutron diffusion eigenproblem with heterogeneous coefficients. *ESAIM: Mathematical Modelling and Numerical Analysis*, 52(5):2003–2035, 2018.

[6] P. Ciarlet Jr, E. Jamelot, and F.D. Kpadonou. Domain decomposition methods for the diffusion equation with low-regularity solution. *Computers & Mathematics with Applications*, 74(10):2369–2384, 2017.

[7] R. Dautray and J.-L. Lions. *Analyse mathématique et calcul numérique pour les sciences et les techniques*. Masson, 1985.

[8] D. A. Di Pietro and A. Ern. *Mathematical aspects of discontinuous Galerkin methods*, volume 69. Springer Science & Business Media, 2011.

[9] J. J. Duderstadt and L. J. Hamilton. *Nuclear reactor analysis*, volume 650. John Wiley & Sons, Inc., 1976.

[10] A. Ern and J.-L. Guermond. *Theory and practice of finite elements*, volume 159. Springer Science & Business Media, 2013.

[11] E. M. Gelbard. Application of spherical harmonics method to reactor problems. *Bettis Atomic Power Laboratory, West Mifflin, PA, Technical Report No. WAPD-BT-20*, 1960.

[12] L. Giret. *Non-Conforming Domain Decomposition for the Multigroup Neutron SPN Equation*. PhD thesis, Paris Saclay, 2018.

[13] J.-J. Lautard and J.-Y. Moller. Minaret, a deterministic neutron transport solver for nuclear core calculations. In M&C, Rio de Janeiro, 2011.

[14] J.J. Lautard, S. Loubière, and C. Fedon-Magnaud. Cronos: a modular computational system for neutronic core calculations. 1992.

[15] S. Nicaise and A.-M. Sändig. General interface problems. I, II. *Mathematical Methods in the Applied Sciences*, 17(6):395–429, 431–450, 1994.

[16] ie: for $1 \leq g \leq G$, $m \in \mathcal{I}_e$, $(\phi_m^g)|_{\partial\mathcal{R}} = 0$.

[17] J. E. Osborn. Spectral approximation for compact operators. *Mathematics of computation*, 29(131):712–725, 1975.

[18] D Schneider et al. APOLLO3® : CEA/DEN deterministic multi-purpose code for reactor physics analysis. PHYSOR-2016, May 1-5 2016, Sun Valley, Idaho, USA.

[19] T. Takeda and H. Ikeda. 3-d neutron transport benchmarks. *Journal of Nuclear Science and Technology*, 28(7):656–669, 1991.