Discretization of the PN model for 2D transport of particles with a Trefftz Discontinuous Galerkin method

Christophe Buet, Bruno Després, Guillaume Morel

Discretization of the P_N model for 2D transport of particles with a Trefftz Discontinuous Galerkin method

Christophe Buet[1,5], Bruno Despres[2,3,5], Guillaume Morel[4,5]

[1] CEA, DAM, DIF, F-91297 Arpajon, France

[2] Sorbonne Universités, UPMC Univ Paris 06, UMR 7598, Laboratoire Jacques-Louis Lions, F-75005, Paris, France

[3] Institut Universitaire de France.

[4] Inria Rennes, Campus de Beaulieu, 263 Avenue Général Leclerc, F-35042 Rennes, France.

Abstract

A discretization of the P_N model for 2D transport of particles is presented, together with the derivation of new high order convergence estimates and new numerical results for the P_1 and P_3 models. The discretization method is based on recent advances about the construction of a Trefftz Discontinuous Galerkin (TDG) method for Friedrichs systems coming from linear transport with relaxation are presented in a comprehensive setting. More numerical results in 2 dimensions illustrate the theoretical properties.

1 Introduction

In this work, we consider the P_N approximation in 2D of the transport equation of photons, neutrons or other types of particles in dimension $d = 3$

$$\partial_t I(t,\mathbf{x},\mathbf{\Omega}) + c\mathbf{\Omega}\cdot\nabla I(t,\mathbf{x},\mathbf{\Omega}) = -\sigma_a(\mathbf{x})I(t,\mathbf{x},\mathbf{\Omega}) + \sigma_s(\mathbf{x})\left(\langle I\rangle - I(t,\mathbf{x},\mathbf{\Omega})\right), \tag{1}$$

where I is the distribution function, t the time variable, $\mathbf{x}\in\mathbb{R}^d$ the space variable, $\mathbf{\Omega}$ the direction and $\langle I\rangle = \frac{1}{4\pi}\int_{S^2} I(t,\mathbf{x},\mathbf{\Omega}')d\mathbf{\Omega}'$ is the mean of I. The absorption and scattering coefficients are $\sigma_a(\mathbf{x}) \geq 0$ and $\sigma_s(\mathbf{x}) \geq 0$. We will adopt the common strategy which is to construct the P_N model [1, 2] in the form of a Friedrichs system with relaxation. The numerical method that we use is the Trefftz Discontinuous Galerkin (TDG) method. This method will be comprehensively presented for a class of Friedrichs systems with relaxation which encompasses many physical problems coming from the approximation of transport phenomena.

Given a linear system of partial differential equations, TDG methods are discontinuous Galerkin type schemes that use solutions to the model as basis functions The name comes from the seminal 1926 paper of E. Trefftz which has been recently translated in English [3]. With respect to more traditional numerical approximation methods, Trefftz methods offer in good cases a strong reduction of the number of unknowns, which may seem as a magic property. Trefftz methods have been widely used and studied for time harmonic wave propagation problems, see the review [4] and reference therein, and more recently for time formulation of propagation equations [5, 6, 7, 8, 9, 10, 11]. The TDG method allows to incorporate some information about the solution of the model in the basis functions and, in certain cases, can require fewer degrees of freedom than standard schemes.

TDG method have their pros and cons:

Pros a) Incorporate a priori knowledge in the basis functions, which are therefore well adapted to multiscale problems; b) Often need less degrees of freedom to reach a given accuracy. A typical example for the 2D version of the P_1 model (2.4) in the dominant absorption regime $\sigma_a > 0$ (with $c = \varepsilon = 1$) is illustrated in the table below, where we compare the number of basis functions p needed to achieve a given fractional order. The first line is for our TDG method. One gets $p_{\rm TDG} = 2(\text{order}+1)$ which is a rephrasing of the result of Theorem 6.4 with $N = 1$. The second line is the optimal number of basis function for a general DG method $p_{\rm DG} = \frac{3}{2}(\text{order}+\frac{1}{2})(\text{order}+\frac{3}{2})$.

order	1/2	3/2	5/2	7/2	9/2
$p_{\rm TDG}$	3	5	7	9	11
$p_{\rm DG}$	3	9	18	30	45

[5] E-mail addresses: guillaume.morel@inria.fr, christophe.buet@cea.fr, despres@ann.jussieu.fr

In particular the number of basis functions is the same to get order $= 1/2$. One always gets $p_{\text{TDG}} \leq p_{\text{DG}}$; c) Is easy to incorporate in DG codes, since one only needs to change the basis functions.
Cons d) May suffer ill-conditioning due to poor linear independence of the basis functions [12, 13]. For wave problems, some remedies exist in the literature [14]; e) The practical calculation of the basis functions adds to the computational burden. If one can calculate the basis functions analytically, the computational burden is moderate. If this is not consider, the computational burden is heavier: several options could be considered such as numerically computing the basis functions or relying on the general procedure [15, 16, 17].

In this work we intend to give a review of recent advances of the TDG method for our model problem (4-8), starting from the preliminary works [18, 19, 20]. Assuming that the coefficient σ_a and σ_s are piecewise constant, we will construct families of TDG basis functions adapted to the numerical approximation of the model problem. For first order PDE's the adjoint equations may differ from the direct equations for $R \neq 0$ which is our case, and therefore one can construct two kinds of TDG basis functions: using adjoint solutions or using direct solutions. It turns out that using adjoint solutions is not an efficient method and we will therefore focus on the TDG method with forward solutions. Note this issue does not occur with the already mentioned works [5, 6, 7, 8, 9, 10, 11] because there is no relaxation in their case. Another possibility might be to adopt a Petrov-Galerkin approach choosing test functions as adjoint solutions and trial functions as direct solutions [21, 22]. We have noticed serious stability issues with this method for time dependent problems.

This paper is organized as follows. In Section 2, the Friedrichs system with relaxation is physically motivatedThe as the angular discretization of the kinetic equation (1): in the literature it is called the P_N model and its main properties are given; these properties are directly connected to invariance principles common to many different models. In Section 3 we present the TDG method for Friedrichs systems. Section B is devoted to the numerical analysis the method, in particular a quasi-optimality result and the well-balanced property of the scheme. In Section 4, we determine Trefftz families of exponential and polynomial solutions to our model problem. Next, in Section 5, the stationary solutions (polynomials and exponentials) are explicitly calculated for the P_1 and P_3 models. In section 6 we give a new approximation result showing that with sufficiently exponentials or polynomial solutions we can approach any stationnary 2D solutions of the P_N equations at any order. Thus combining it with the quasi-optimality result we obtain a new high order convergence result. In the final Section 7 we provide numerical examples. First a new numerical example with boundary layers is provided for the P_3 model in Section 7: in terms of accuracy, it shows an important improvement with respect to more standard DG methods. The second test problem illustrates the advantage of the method for a test problem in a diffusion regime. For the last numerical result TDG is compared with standard DG for a time dependent problem.

In this paper all vectors are written in bold. For $\mathbf{v}(\mathbf{x}) \in \mathbb{R}^m$ we use the simplified notation $\mathbf{v} \in L^2(\Omega)$ instead of $\mathbf{v} \in L^2(\Omega)^m$. Moreover we may write $\mathbf{v} = (v_1, ..., v_m)^T$ where ^{T} denotes the transpose and denote $\mathbf{v}^2 = \mathbf{v}^T\mathbf{v}$ to facilitate the distinction with other types of norms or semi-norms.

2 The P_N model

In this section, we start with the kinetic equation (1) and we construct the transport matrices and the relaxation matrix of the P_N approximation in 2D.

2.1 3D configuration

Let $\psi \in [0, 2\pi)$ and $\phi \in [0, \pi)$ be the polar and azimutal angles on the sphere, so that in Cartesian coordinate with usual notations $\mathbf{\Omega} := (\Omega_1, \Omega_2, \Omega_3)^T = (\sin\phi\cos\psi, \sin\phi\sin\psi, \cos\phi)^T \in \mathbb{R}^3$. To be consistent with the standard notation of the spherical harmonics, the uppercase letter $Y_{k,l}$ is used to denoted the real spherical harmonics. We make a slight abuse of notation by not distinguishing between the two forms

$$Y_{k,l}(\mathbf{\Omega}) := Y_{k,l}(\psi, \phi) : \mathbb{R}^2 \to \mathbb{R}, \quad |l| \leq k \leq N, \quad k, l \in \mathbb{N}.$$

The construction and properties of the spherical harmonics are detailed in Appendix A. We introduce some notations and adopt the presentation from [23] but with the spherical harmonics vector arranged as in [2]. In the following, we denote m^{3D} the total number of unknowns, m_e^{3D} the number of even moments (which correspond to k even) and m_o^{3D} the number of odd moments (which correspond to k odd) for the three dimensional P_N model. That is

$$m^{3D} := m_e^{3D} + m_o^{3D} = (N+1)^2, \quad m_e^{3D} := \frac{1}{2}N(N+1), \quad m_o^{3D} := \frac{1}{2}(N+1)(N+2).$$

For any integer $0 \leq k \leq N$ we define $\mathbf{y}_k(\mathbf{\Omega})$ the vectorial function whose components are the $2k+1$ real valued spherical harmonics of order k. Moreover we denote $\mathbf{y}_e(\mathbf{\Omega})$ the vectorial function made of the so-called even moments $\big(\mathbf{y}_{2k}(\mathbf{\Omega})\big)_{0 \leq 2k \leq N}$

and $\mathbf{y}_o(\mathbf{\Omega})$ the vectorial function made of the so-called odd moments $\big(\mathbf{y}_{2k+1}(\mathbf{\Omega})\big)_{0\leq 2k+1\leq N}$. That is

$$\mathbf{y}_k(\mathbf{\Omega}) := \Big(Y_{k,-k}(\mathbf{\Omega}), Y_{k,-k+1}(\mathbf{\Omega}), ..., Y_{k,k-1}(\mathbf{\Omega}), Y_{k,k}(\mathbf{\Omega})\Big)^T \in \mathbb{R}^{2k+1},$$
$$\mathbf{y}_e(\mathbf{\Omega}) := \Big(\mathbf{y}_0^T(\mathbf{\Omega}), \mathbf{y}_2^T(\mathbf{\Omega}), ..., \mathbf{y}_{N-1}^T(\mathbf{\Omega})\Big)^T \in \mathbb{R}^{m_e^{3D}}, \quad \mathbf{y}_o(\mathbf{\Omega}) := \Big(\mathbf{y}_1^T(\mathbf{\Omega}), \mathbf{y}_3^T(\mathbf{\Omega}), ..., \mathbf{y}_N^T(\mathbf{\Omega})\Big)^T \in \mathbb{R}^{m_o^{3D}},$$

Finally, we define $\mathbf{y}(\mathbf{\Omega})$ the vectorial function made of $\mathbf{y}_e(\mathbf{\Omega})$, $\mathbf{y}_o(\mathbf{\Omega})$ and arranged as follow $\mathbf{y}(\mathbf{\Omega}) = \Big(\mathbf{y}_e^T(\mathbf{\Omega}), \mathbf{y}_o^T(\mathbf{\Omega})\Big)^T \in \mathbb{R}^{m^{3D}}$. We generalize this decomposition for any vector $\mathbf{v} \in \mathbb{R}^{m^{3D}}$. We set $\mathbf{v}_k := (v_k^{-k}, v_k^{-k+1}, ..., v_k^{k-1}, v_k^k)^T \in \mathbb{R}^{2k+1}$, $\mathbf{v}_e := (\mathbf{v}_0^T, \mathbf{v}_2^T, ..., \mathbf{v}_{N-1}^T)^T \in \mathbb{R}^{m_e^{3D}}$ and $\mathbf{v}_o := (\mathbf{v}_1^T, \mathbf{v}_3^T, ..., \mathbf{v}_N^T)^T \mathbb{R}^{m_o^{3D}}$, and denote $\mathbf{v}$ as $\mathbf{v} = (\mathbf{v}_e^T, \mathbf{v}_o^T)^T \in \mathbb{R}^{m^{3D}}$. The decomposition of the intensity on the spherical harmonics basis can be written as $I(t, \mathbf{x}, \mathbf{\Omega}) = \sum_{k\geq 0} \sum_{|l|\leq k} Y_{k,l}(\mathbf{\Omega}) u_k^l(t, x)$. The truncation at order N defines the truncated series I_N

$$I_N(t, \mathbf{x}, \mathbf{\Omega}) := \mathbf{y}^T(\mathbf{\Omega})\mathbf{u}(t, \mathbf{x}) = \sum_{k=0}^{N} \mathbf{y}_k^T(\mathbf{\Omega})\mathbf{u}_k(t, x) = \sum_{k=0}^{N} \sum_{|l|\leq k} Y_{k,l}(\mathbf{\Omega}) u_k^l(t, x),$$

where the unknown of the P_N model are $\mathbf{u} \in \mathbb{R}^{m^{3D}}$. With the approximation $I = I_N$ the equation (1) reads

$$\mathbf{y}^T(\mathbf{\Omega})\partial_t \mathbf{u}(t, \mathbf{x}) + c \sum_{i=1}^{3} \Omega_i \mathbf{y}^T(\mathbf{\Omega}) \partial_{x_i} \mathbf{u}(t, \mathbf{x}) = \Big(- \big(\sigma_a + \sigma_s\big)\mathbf{y}^T(\mathbf{\Omega})\mathbf{u}(t, \mathbf{x}) + \sigma_s < \mathbf{y}^T(\mathbf{\Omega}) > \Big)\mathbf{u}(t, \mathbf{x}).$$

Multiplying by $\mathbf{y}(\mathbf{\Omega})$ and integrating over the sphere gives

$$< \mathbf{y}(\mathbf{\Omega})\mathbf{y}^T(\mathbf{\Omega}) > \partial_t \mathbf{u}(t, \mathbf{x}) + c \sum_{i=1}^{3} < \Omega_i \mathbf{y}(\mathbf{\Omega})\mathbf{y}^T(\mathbf{\Omega}) > \partial_{x_i} \mathbf{u}(t, \mathbf{x}) =$$
$$\Big(- \big(\sigma_a + \sigma_s\big) < \mathbf{y}(\mathbf{\Omega})\mathbf{y}^T(\mathbf{\Omega}) > + \sigma_s < \mathbf{y}(\mathbf{\Omega}) >< \mathbf{y}^T(\mathbf{\Omega}) > \Big)\mathbf{u}(t, \mathbf{x}).$$

From the orthogonal properties of the spherical harmonics one has $< \mathbf{y}(\mathbf{\Omega})\mathbf{y}^T(\mathbf{\Omega}) >= I_{m^{3D}}$ and $< \mathbf{y}(\mathbf{\Omega}) >< \mathbf{y}^T(\mathbf{\Omega}) >= \mathbf{e}_1\mathbf{e}_1^T$ with $\mathbf{e}_1 = (1, 0, ..., 0)^T \in \mathbb{R}^{m^{3D}}$. Therefore one gets the system $\partial_t \mathbf{u} + \sum_{i=1}^{3} \mathcal{A}_i \partial_{x_i} \mathbf{u} = -\mathcal{R}\mathbf{u}$, where $\mathbf{u} \in \mathbb{R}^{m^{3D}}$ and $\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3, \mathcal{R} \in \mathbb{R}^{m^{3D}\times m^{3D}}$. The matrices $\mathcal{A}_i = c < \Omega_i \mathbf{y}(\mathbf{\Omega})\mathbf{y}^T(\mathbf{\Omega}) >$ can be computed using the recursion relations (39) to expand $\Omega_i \mathbf{y}(\mathbf{\Omega})$ in terms of spherical harmonics. As pointed out in [2] the matrices $\mathcal{A}_1$, $\mathcal{A}_2$ and $\mathcal{A}_3$ have a block structure $\mathcal{A}_1 = \begin{pmatrix} 0 & \mathcal{A} \\ \mathcal{A}^T & 0 \end{pmatrix}$, $\mathcal{A}_2 = \begin{pmatrix} 0 & \mathcal{B} \\ \mathcal{B}^T & 0 \end{pmatrix}$, $\mathcal{A}_3 = \begin{pmatrix} 0 & \mathcal{C} \\ \mathcal{C}^T & 0 \end{pmatrix}$ where $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C} \in \mathbb{R}^{m_e^{3D}\times m_o^{3D}}$ are rectangular matrices. This block structure is a direct consequence of a decoupling between the even moments and odd moments. The relaxation matrix is diagonal $\mathcal{R} = \text{diag}(\sigma_a, \sigma_a + \sigma_s, ..., \sigma_a + \sigma_s)$. One obtains a system (4) in dimension 3.

2.2 3D rotational invariance

There are some inherent technicalities attached to the description of rotational invariance principles, however what this reveals is extremely valuable for implementations and is why we take the time to describe them.

The matrix representations of the rotation operators in the basis of spherical harmonics, known as D-Wigner matrices [24, 25, 26], are $\mathcal{U}(\alpha, \beta, \gamma) \in \mathbb{R}^{m^{3D}\times m^{3D}}$, where α, β and γ denotes rotation around the axes O_z, O_y and O_z respectively. It is a block matrix

$$\mathcal{U}(\alpha, \beta, \gamma) = \text{diag}\Big(\Delta_0(\alpha, \beta, \gamma), \Delta_2(\alpha, \beta, \gamma), ...\Delta_{m_e}(\alpha, \beta, \gamma), \Delta_1(\alpha, \beta, \gamma), ..., \Delta_{m_o}(\alpha, \beta, \gamma)\Big)$$

where the matrices Δ_k reads [26] $\Delta_k(\alpha, \beta, \gamma) = \mathcal{W}_k(\alpha)\mathcal{D}_k(\beta)\mathcal{W}_k(\gamma) \in \mathbb{R}^{2k+1\times 2k+1}$. Here $\mathcal{D}_k \in \mathbb{R}^{2k+1\times 2k+1}$ is a (reduced) d-Wigner matrix and the matrix $\mathcal{W}_k$ has non-zero elements only on its diagonal and anti-diagonal

$$\mathcal{W}_k(\alpha) = \begin{pmatrix} \cos k\alpha & & & & & & \sin k\alpha \\ & \ddots & & & & \iddots & \\ & & \cos 2\alpha & & \sin 2\alpha & & \\ & & & \cos\alpha & \sin\alpha & & \\ & 0 & & 1 & & 0 & \\ & & & -\sin\alpha & \cos\alpha & & \\ & & -\sin 2\alpha & & & \cos 2\alpha & \\ & \iddots & & & & \ddots & \\ -\sin k\alpha & & & & & & \cos k\alpha \end{pmatrix} \in \mathbb{R}^{2k+1\times 2k+1}.$$

To simplify the matrix $\mathcal{U}$ we may consider a rotation θ in the plane xy only and denote

$$\mathcal{U}_\theta := \mathcal{U}(0,0,\theta) \in \mathbb{R}^{m^{3D}\times m^{3D}}. \tag{2}$$

Using the block rotations, the structure is written as $\mathcal{U}_\theta = \text{diag}\left(\mathcal{W}_0(\theta), \mathcal{W}_2(\theta), ...\mathcal{W}_{m_e}(\theta), \mathcal{W}_1(\theta), ..., \mathcal{W}_{m_o}(\theta)\right)$. The matrix $\mathcal{U}$ represents the orthogonal transformations on $\mathbf{y}(\mathbf{\Omega})$. That is for an orthogonal matrix $Q \in \mathbb{R}^{3\times 3}$ one has $\mathbf{y}(Q\mathbf{\Omega}) = \mathcal{U}(\alpha,\beta,\gamma)\mathbf{y}(\mathbf{\Omega})$ where α, β and γ are the angles of the rotation associated with the matrix Q in $\mathbb{R}^3$.

2.3 2D configuration

We briefly recall the invariance principles which can be found in the references [23, 2, 26].
a) In practice, the P_N model is rarely applied for even values of N (see for example [23, Section 2] for a discussion on the benefits of considering N odd). So we will consider only the case $N = 2n+1$ odd in the following.
b) Provided σ_a and σ_s are invariant with respect to z, then the system is invariant by translation with respect to z. It corresponds to $\partial_z = 0$ (equivalent to setting $\mathcal{A}_3 = 0$) and it has the consequence that the dimension is lowered from $d = 3$ to $d = 2$.
c) We assume that the solution has a mirror symmetry with respect to the plane xy. It is interpreted as pure reflective conditions at the top and bottom boundaries of the 3D domain, as illustrated in Figure 1. This is equivalent to saying that the function $\mathbf{u}$ is an even function of $\cos\phi$. Inspection of the spherical harmonics (37-38) shows that they are odd with respect to $\cos\phi$ if and only if $k+l$ odd, see [2]. Thus one can remove the unknowns u_k^l such that $k+l$ is odd (because

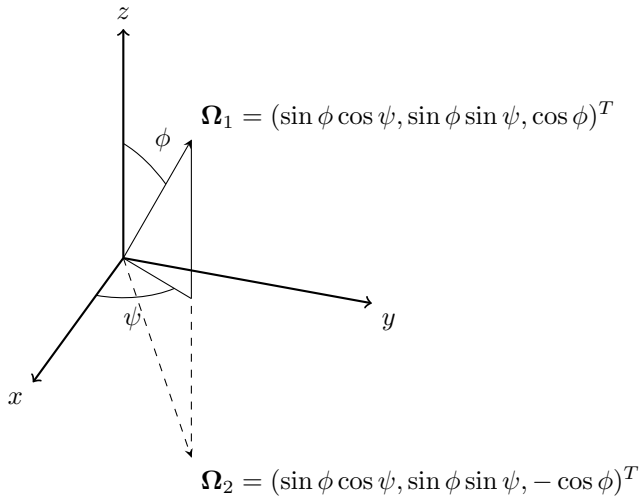

Figure 1: Mirror symmetry. If $\mathbf{u}$ is an even function of $\cos\phi$ then $\mathbf{u}(t,\mathbf{x},\mathbf{\Omega}_1) = \mathbf{u}(t,\mathbf{x},\mathbf{\Omega}_2)$.

they must vanish). This simplifies the matrices $\mathcal{A}_1$, $\mathcal{A}_2$ and $\mathcal{R}$ by removing rows and lines such that $k+l$ is odd. One can check that the size of the remaining part is $m_o^{3D} = (N+1)(N+2)/2$. This procedure defines the matrices

$$A_1 = c\begin{pmatrix} 0 & A \\ A^T & 0 \end{pmatrix} \in \mathbb{R}^{m_o^{3D}\times m_o^{3D}}, \quad A_2 = c\begin{pmatrix} 0 & B \\ B^T & 0 \end{pmatrix} \in \mathbb{R}^{m_o^{3D}\times m_o^{3D}} \text{ and } R \in \mathbb{R}^{m_o^{3D}\times m_o^{3D}} \tag{3}$$

and one gets the P_N model in two dimensions with m_o^{3D} unknowns. These matrices correspond to the ones (5-7) in the introduction and

$$m = m_o^{3D} = \frac{1}{2}(N+1)(N+2).$$

d) Finally, a P_N model has to satisfy some rotational invariance principles. With the previous assumptions this rotational invariance is expressed in the plane xy. The plane xz may also be a possible choice [27, 28], however the rotation matrix associated with the spherical harmonics is more difficult to calculate [24, 29, 26].

After reduction to the plane xy, the rotation matrix $\mathcal{U}_\theta$ becomes $U_\theta \in \mathbb{R}^{m_o^{3D}\times m_o^{3D}}$ with the natural properties $U_\theta U_\mu = U_{\theta+\mu}$ and $(U_\theta)^T = U_{-\theta}$.

2.4 First example: the P_1 model

For the P_1 model, one takes $m = 3$, $m_e = 1$, $m_o = 2$. The matrices A_1, A_2, R and U_θ are

$$A_1 = \frac{c}{\sqrt{3}}\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad A_2 = \frac{c}{\sqrt{3}}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad R = \begin{pmatrix} \sigma_a & 0 & 0 \\ 0 & \sigma_t & 0 \\ 0 & 0 & \sigma_t \end{pmatrix}, \quad U_\theta = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix}.$$

The submatrices are $A = \left(\frac{1}{\sqrt{3}}, 0\right)$ and $B = \left(0, \frac{1}{\sqrt{3}}\right)$.

2.5 Second example: the P_3 model

For the P_3 model one takes $m = 10$, $m_e = 4$, $m_o = 6$. The submatrices read [2]

$$A = \begin{pmatrix} 0 & \frac{1}{\sqrt{3}} & 0 & 0 & 0 & 0 \\ \frac{1}{\sqrt{5}} & 0 & \sqrt{\frac{3}{14}} & -\frac{1}{\sqrt{70}} & 0 & 0 \\ 0 & -\frac{1}{\sqrt{15}} & 0 & 0 & \sqrt{\frac{6}{35}} & 0 \\ 0 & \frac{1}{\sqrt{5}} & 0 & 0 & -\frac{1}{\sqrt{70}} & \sqrt{\frac{3}{14}} \end{pmatrix}, \quad B = \begin{pmatrix} \frac{1}{\sqrt{3}} & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{5}} & 0 & 0 & -\frac{1}{\sqrt{70}} & -\sqrt{\frac{3}{14}} \\ -\frac{1}{\sqrt{15}} & 0 & 0 & \sqrt{\frac{6}{35}} & 0 & 0 \\ -\frac{1}{\sqrt{5}} & 0 & \sqrt{\frac{3}{14}} & \frac{1}{\sqrt{70}} & 0 & 0 \end{pmatrix},$$

$R_1 = \begin{pmatrix} \sigma_a & 0 & 0 & 0 \\ 0 & \sigma_t & 0 & 0 \\ 0 & 0 & \sigma_t & 0 \\ 0 & 0 & 0 & \sigma_t \end{pmatrix}$ and $R_2 = \sigma_t I_{m_o}$ where I_{m_o} is the identity matrix of $\mathbb{R}^{m_o\times m_o}$. The matrices A_1 and A_2 are assembled by a symmetrization of A and B (5). The rotation matrix is constructed in [19][chapter 4]: its structure is $U_\theta = \begin{pmatrix} V_2(\theta) & 0 \\ 0 & V_3(\theta) \end{pmatrix}$ where $V_2(\theta) = \begin{pmatrix} W_0(\theta) & 0 \\ 0 & W_2(\theta) \end{pmatrix}$, $V_3(\theta) = \begin{pmatrix} W_1(\theta) & 0 \\ 0 & W_3(\theta) \end{pmatrix}$ and finally $W_p(\theta)$ is the matrix of rotation for the spherical harmonics of order $p = 1, 2, 3, 4$. These matrices are $W_0(\theta) = 1$,

$$W_1(\theta) = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}, \; W_2(\theta) = \begin{pmatrix} \cos 2\theta & 0 & \sin 2\theta \\ 0 & 1 & 0 \\ -\sin 2\theta & 0 & \cos 2\theta \end{pmatrix} \text{ and } W_3(\theta) = \begin{pmatrix} \cos 3\theta & 0 & 0 & \sin 2\theta \\ 0 & \cos\theta & \sin\theta & 0 \\ 0 & -\sin\theta & \cos\theta & 0 \\ -\sin 3\theta & 0 & 0 & \cos 3\theta \end{pmatrix}.$$

2.6 2D rotational invariance

Rotational invariance is related to the construction of U_θ at the end of Section 2.3 and has important implications for the numerical implementation. It can be checked directly on the P_N model. One introduces the change of frame

$$(x', y') = (x\cos\theta - y\sin\theta, x\sin\theta + y\cos\theta) \Longleftrightarrow (x, y) = (x'\cos\theta + y'\sin\theta, -x'\sin\theta + y'\cos\theta).$$

A possible formulation of the rotational invariance of the P_N model is as follows: if $\mathbf{u}(t, x, y)$ is the solution in the reference frame (x, y), then the function $U_{-\theta}\mathbf{u}(t, x', y')$ is also a solution in the same frame. The minus sign is for compatibility reasons with the notations of [26] and [19, Section 4.1].

Proposition 2.1. *Rotational invariance is equivalent to* $U_\theta R = RU_\theta$, $U_\theta A_1 = (A_1\cos\theta + A_2\sin\theta)\,U_\theta$ *and* $U_\theta A_2 = (-A_1\sin\theta + A_2\cos\theta)\,U_\theta$.

Proof. Set $V(t,x',y') = U_{-\theta}\mathbf{u}(t, x\cos\theta - y\sin\theta, x\sin\theta + y\cos\theta)$. Rotational invariance states that V is the solution of the same model for arbitrary values of θ.

The derivatives are $\partial_t V = U_{-\theta}\partial_t\mathbf{u}$, $\partial_x V = U_{-\theta}\left(\cos\theta\partial_x\mathbf{u} + \sin\theta\partial_y\mathbf{u}\right)$ and $\partial_y V = U_{-\theta}\left(-\sin\theta\partial_x\mathbf{u} + \cos\theta\partial_y\mathbf{u}\right)$. One has that $\partial_t\mathbf{u} = -A_1\partial_x\mathbf{u} - A_2\partial_y\mathbf{u} - R\mathbf{u}$, therefore

$$\partial_t V + A_1\partial_x V + A_2\partial_y V + RV = U_{-\theta}\left(-A_1\partial_x\mathbf{u} - A_2\partial_y\mathbf{u} - R\mathbf{u}\right)$$

$$+A_1U_{-\theta}\left(\cos\theta\partial_x\mathbf{u} + \sin\theta\partial_y\mathbf{u}\right) + A_2U_{-\theta}\left(-\sin\theta\partial_x\mathbf{u} + \cos\theta\partial_y\mathbf{u}\right) + RU_{-\theta}\mathbf{u}$$

that is

$$0 = \left(RU_{-\theta} - U_{-\theta}R\right)\mathbf{u} + \left(A_1U_{-\theta}\cos\theta - A_2U_{-\theta}\sin\theta - U_{-\theta}A_1\right)\partial_x\mathbf{u} + \left(A_1U_{-\theta}\sin\theta + A_2U_{-\theta}\cos\theta - U_{-\theta}A_2\right)\partial_y\mathbf{u}.$$

Considering arbitrary independent values of $\mathbf{u}$, $\partial_x\mathbf{u}$ and $\partial_y\mathbf{u}$ yields

$$0 = RU_{-\theta} - U_{-\theta}R = A_1U_{-\theta}\cos\theta - A_2U_{-\theta}\sin\theta - U_{-\theta}A_1 = A_1U_{-\theta}\sin\theta + A_2U_{-\theta}\cos\theta - U_{-\theta}A_2.$$

Changing θ in $-\theta$, one gets the claim. □

2.7 Other properties

More properties which concern the matrix A (3) can be proved for a general P_N model. The full proofs can be found in [19, Section 4.1]. It is straightforward to verify these properties for the P_1 model and the P_3 model.

Proposition 2.2. *The symmetric matrix AA^T is invertible and all its eigenvalues are strictly positive. The eigenvalues μ_i of$(AA^T)^{-1}R_1$ are strictly positive when $\sigma_a > 0$ and non negative when $\sigma_a = 0$. An important property will be the degeneracy of the lowest eigenvalue as $\sigma_a \to 0$. Assume $\sigma_s > 0$. The lowest eigenvalue of $(AA^T)^{-1}R_1$ is such that $\mu_1 \to 0$ as $\sigma_a \to 0$, and it is non-degenerate (has multiplicity one). Finally, one can count the number of distinct pairs of eigenvalue/eigenvectors of the matrix $(AA^T)^{-1}R_1$ (something we will need for the proof of Theorem 4.1). The eigenvectors of $(AA^T)^{-1}R_1 \in \mathbb{R}^{m_e\times m_e}$ form a basis of $\mathbb{R}^{m_e}$.*

3 Presentation of the Trefftz Discontinuous Galerkin method for Friedrichs systems

The model problem considers both stationary and time dependent problems. Let Ω_S be a bounded polygonal/polyhedral Lipschitz space domain in $\mathbb{R}^d$ and consider a time interval $[0,T]$, $T > 0$. We denote $\Omega = \Omega_S$ for stationary problems and $\Omega = \Omega_S \times [0,T]$ for time dependent problems. Friedrichs systems [30] with linear relaxation can be written as

$$\begin{cases} \sum_{i=0}^{d} A_i\partial_i\mathbf{u} = -R(\mathbf{x})\mathbf{u}, & \text{in } \Omega, \\ M^-\mathbf{u} = M^-\mathbf{g}, & \text{in } \partial\Omega, \end{cases} \tag{4}$$

the dependent variable is $\mathbf{u} \in \mathbb{R}^m$, $\mathbf{x} = (x_1, ..., x_d) \in \mathbb{R}^d$ is the space variable and t is the time variable. The coefficients σ_a and σ_s in (1) are contained in the relaxation matrix R. Recalling that the problem can be stationary or time dependent one may write $\mathbf{u}(t,\mathbf{x})$ or just $\mathbf{u}(\mathbf{x})$ depending on the situation. The matrices $A_i, R(\mathbf{x}) \in \mathbb{R}^{m\times m}$ are symmetric and we assume $R(\mathbf{x}) \in \mathbb{R}^{m\times m}$ is a non negative matrix, i.e. $(R(\mathbf{x})\mathbf{v},\mathbf{v}) \geq 0$ for all $\mathbf{v} \in \mathbb{R}^m, \mathbf{x} \in \mathbb{R}^d$. We use the notation $\partial_0 = \partial_t$, $\partial_i = \partial_{x_i}$ for $i = 1, ..., d$. The outward normal unit vector is $\mathbf{n}(t,\mathbf{x}) = (n_t, n_{x_1}, ..., n_{x_d})$ for $x \in \partial\Omega$ and of course for stationary problems $n_t = 0$ for all $x \in \partial\Omega$. We set $M(\mathbf{n}) = A_0n_t + \sum_{i=1}^{d} A_in_{x_i}$, on $\partial\Omega$. Since M is symmetric one has the standard decomposition $M(\mathbf{n}) = M^+(\mathbf{n}) + M^-(\mathbf{n})$ where M^+ is a non negative matrix and M^- is a non positive matrix. We use the matrix M^- to write the boundary conditions with $\mathbf{g} \in L^2(\partial\Omega)$. Finally we assume the problem (4) admits a unique solution. This family of problems is considered in [5, 6, 7, 8, 9, 10, 11], but without relaxation since they do not model any kind of diffusion. In our case, which stems from the transport of particles or energy together with interactions with the matter in 2D configuration, we set $d = 2$ and adopt the convention coming from [2] that the matrices A_1 and A_2 have the block structure

$$A_1 = c\begin{pmatrix} 0 & A \\ A^T & 0 \end{pmatrix} \in \mathbb{R}^{m\times m}, \quad A_2 = c\begin{pmatrix} 0 & B \\ B^T & 0 \end{pmatrix} \in \mathbb{R}^{m\times m}, \tag{5}$$

where A, $B \in \mathbb{R}^{m_e\times m_o}$ are constant rectangular matrices ($m_e + m_o = m$). The coefficient $c > 0$ is a constant non dimensional wave velocity. For the purposes of mathematical manipulation, the first matrix is

$$A_0 = \varepsilon I_m \in \mathbb{R}^{m\times m} \tag{6}$$

where $I_m \in \mathbb{R}^{m\times m}$ is the identity matrix. The new parameter

$$0 < \varepsilon \leq 1$$

indicates a possible rescaling of the time variable, which is adapted to different physical regimes. In particular $\varepsilon \to 0$ corresponds to the Diffusion Asymptotic regime, where transport is approximated with a diffusion equation [31, 28, 32, 33]. With respect to [7, 5, 11], the originality of our methods is in the non zero relaxation matrix. A natural structure [34] which models relaxation mechanisms is $R + R^t \geq 0$. In our work, we follow closely the convention proposed in [2] by taking a piecewise constant matrix

$$R = \begin{pmatrix} R_1 & 0 \\ 0 & R_2 \end{pmatrix} \in \mathbb{R}^{m\times m}, \tag{7}$$

where R_1 and R_2 are both diagonal matrices $R_1 := \text{diag}(\varepsilon\sigma_a, \sigma_t, ..., \sigma_t) \in \mathbb{R}^{m_e\times m_e}$, $R_2 := \sigma_t I_{m_o} \in \mathbb{R}^{m_o\times m_o}$, with I_{m_o} the identity matrix of $\mathbb{R}^{m_o\times m_o}$. For transfer models [1, 35, 36, 37, 2] the absorption coefficient is $\sigma_a \geq 0$ and the scattering coefficientis $\sigma_s \geq 0$. The weighted sum of the scattering and absorption coefficients will be denoted as

$$\sigma_t := \sigma_t^\varepsilon := \varepsilon\sigma_a + \frac{\sigma_s}{\varepsilon}, \quad \sigma_a, \sigma_s \in \mathbb{R}_+. \tag{8}$$

The matrix $R = R(\mathbf{x})$ will be assumed piecewise constant, because the coefficients $\sigma_a(\mathbf{x})$ and $\sigma_s(\mathbf{x})$ are piecewise constant. The underlying physics is described in more detail in Section 2.

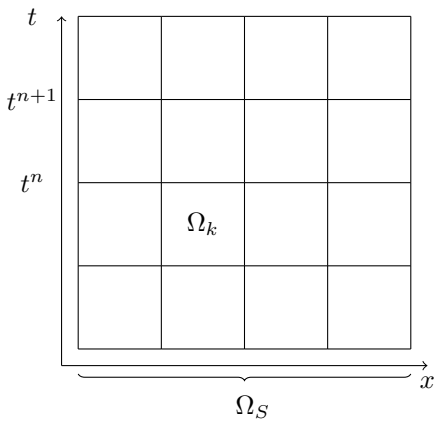

Figure 2: Illustration of the partition $\mathcal{T}_h$ for a time dependent problem.

3.1 Mesh notation and generic discontinuous Galerkin formulation

The partition or mesh of the space domain $\Omega = \Omega_S \subset \mathbb{R}^d$ is denoted as $\mathcal{T}_h$. It is made of polyhedral non overlapping subdomains $\Omega_{S,r}$, that is $\mathcal{T}_h = \cup_r \Omega_{S,r}$. For a space time problem we split the time interval into smaller time intervals (t_n, t_{n+1}) with $0 = t_0 < t_1 < ... < t_N = T$. Making an abuse of notation, the mesh of the space time domain $\Omega = \Omega_S \times [0, T] \subset \mathbb{R}^{d+1}$ is still denoted as $\mathcal{T}_h = \cup_{r,n} \Omega_{S,r} \times (t_n, t_{n+1})$. So $\mathcal{T}_h$ denotes either a purely spatial mesh for stationary models or a space-time mesh for time dependent models. Moreover the cells or subdomains will be referred to with the same notation, that is $\Omega_k = \Omega_{S,r}$ or $\Omega_k = \Omega_{S,r} \times (t_n, t_{n+1})$. The context makes these notations unambiguous. The broken Sobolev space is

$$H^1(\mathcal{T}_h) := \{\mathbf{v} \in L^2(\Omega),\ \mathbf{v}_{|\Omega_k} \in H^1(\Omega_k)\ \forall \Omega_k \in \mathcal{T}_h\}.$$

We assume $\mathbf{u} \in H^1(\mathcal{T}_h)$. We may rewrite (4) under the form $L\mathbf{u} = \mathbf{0}$ where $L(\mathbf{x}) = \sum_i A_i\partial_i + R(\mathbf{x})$. We consider also the adjoint operator $L^*(\mathbf{x}) = -\sum_i A_i\partial_i + R(\mathbf{x})$. All matrices are constant (do not depend either on the time variable or on the space variables). Multiplying (4) by $\mathbf{v} \in H^1(\mathcal{T}_h)$ and integrating on Ω gives $\sum_k \int_{\Omega_k} \mathbf{v}_k^T L(\mathbf{x})\mathbf{u}_k = 0$, where $\mathbf{v}_k = \mathbf{v}_{|\Omega_k}$, $\mathbf{u}_k = \mathbf{u}_{|\Omega_k}$. Integrating by parts one gets $\sum_k \int_{\Omega_k} \left(L^*\mathbf{v}_k\right)^T \mathbf{u}_k + \sum_k \int_{\partial\Omega_k} \mathbf{v}_k^T M_k(\mathbf{x})\mathbf{u}_k = 0$, where $\partial\Omega_k$ is the contour of the element Ω_k with an outward unit normal $\mathbf{n}_k(\mathbf{x}) = (n_t, n_{x_1}, ..., n_{x_d})^T$, $M(\mathbf{x}) = A_0 n_t + \sum_i A_i n_i$ and $M_k(\mathbf{x}) = M(\mathbf{n}_k)$. Since M_k is symmetric it can decomposed under the form $M_k(\mathbf{x}) = M_k^+(\mathbf{x}) + M_k^-(\mathbf{x})$ where M_k^+ is a non negative matrix, M_k^- is a non positive matrix and the matrices annihilate each other $M_k^+ M_k^- = M_k^- M_k^+ = 0$. It is sufficient for this task to compute the eigenvectors $M\mathbf{r} = \lambda\mathbf{r}$, $\|\mathbf{r}\| = 1$, and to set $M^\pm = \sum_{\pm\lambda>0} \lambda\mathbf{r} \otimes \mathbf{r}$. Denoting Σ_{kj} the edge oriented from Ω_k

to Ω_j when $k \neq j$ and Σ_{kk} the edges belonging to $\Omega_k \cap \partial\Omega$ (for simplicity we use the same notation even if there is more than one edge in $\Omega_k \cap \partial\Omega$), one can write

$$\sum_k \int_{\Omega_k} \left(L^*(\mathbf{x})\mathbf{v}_k\right)^T \mathbf{u}_k + \sum_k \sum_{j<k} \int_{\Sigma_{kj}} (\mathbf{v}^T M(\mathbf{x})\mathbf{u})_k + (\mathbf{v}^T M(\mathbf{x})\mathbf{u})_j + \sum_k \int_{\Sigma_{kk}} \mathbf{v}_k^T M_k^+(\mathbf{x})\mathbf{u}_k = -\sum_k \int_{\Sigma_{kk}} \mathbf{v}_k^T M_k^-(\mathbf{x})\mathbf{g}. \quad (9)$$

For $\mathbf{u}$ satisfying the equation (4), the normal flux is continuous on Σ_{kj}: $M_k(\mathbf{x})\mathbf{u}_k(\mathbf{x}) = M_k(\mathbf{x})\mathbf{u}_j(\mathbf{x}) = -M_j(\mathbf{x})\mathbf{u}_j(\mathbf{x})$ for $\mathbf{x} \in \Sigma_{kj}$. This vectorial identity can be projected along the positive and negative eigenvectors of $M_k = -M_j$, leading to similar continuity relations. So, denoting $M_{kj} = M_{k|\Sigma_{kj}} = -M_{j|\Sigma_{jk}} = -M_{jk}$ on Σ_{kj}, one can write also

$$M_k\mathbf{u}_k = M_{kj}\mathbf{u}_j = M_{kj}^+\mathbf{u}_k + M_{kj}^-\mathbf{u}_k = M_{kj}^+\mathbf{u}_k + M_{kj}^-\mathbf{u}_j$$

because the projection of $M_k\mathbf{u}_k = M_{kj}\mathbf{u}_j$ along the eigenvectors yields the continuity $\mathbf{r}_{kj}^T\mathbf{u}_k = \mathbf{r}_{kj}^T\mathbf{u}_j$ for $\lambda \neq 0$. It yields the identity $(\mathbf{v}^T M(\mathbf{x})\mathbf{u})_k + (\mathbf{v}^T M(\mathbf{x})\mathbf{u})_j = (\mathbf{v}_k - \mathbf{v}_j)^T(M_{kj}^+\mathbf{u}_k + M_{kj}^-\mathbf{u}_j)$. So (9) can be recast as

$$\sum_k \int_{\Omega_k} \left(L^*(\mathbf{x})\mathbf{v}_k\right)^T \mathbf{u}_k + \sum_k \sum_{j<k} \int_{\Sigma_{kj}} (\mathbf{v}_k - \mathbf{v}_j)^T(M_{kj}^+(\mathbf{x})\mathbf{u}_k + M_{kj}^-(\mathbf{x})\mathbf{u}_j) + \sum_k \int_{\Sigma_{kk}} \mathbf{v}_k^T M_k^+(\mathbf{x})\mathbf{u}_k = -\sum_k \int_{\Sigma_{kk}} \mathbf{v}_k^T M_k^-(\mathbf{x})\mathbf{g}. \quad (10)$$

We define the bilinear form $a_{DG} : H^1(\mathcal{T}_h) \times H^1(\mathcal{T}_h) \to \mathbb{R}$

$$\begin{aligned} a_{DG}(\mathbf{u},\mathbf{v}) = &\sum_k \int_{\Omega_k} (L^*(\mathbf{x})\mathbf{v}_k)^T \mathbf{u}_k + \sum_k \sum_{j<k} \int_{\Sigma_{kj}} (\mathbf{v}_k - \mathbf{v}_j)^T(M_{kj}^+(\mathbf{x})\mathbf{u}_k + M_{kj}^-(\mathbf{x})\mathbf{u}_j) \\ &+ \sum_k \int_{\Sigma_{kk}} \mathbf{v}_k^T M_k^+(\mathbf{x})\mathbf{u}_k, \quad \mathbf{u},\mathbf{v} \in H^1(\mathcal{T}_h) \end{aligned} \quad (11)$$

and the linear form $l(\mathbf{v}) = -\sum_k \int_{\Sigma_{kk}} \mathbf{v}_k^T M_k^-(\mathbf{x})\mathbf{g}$ for $\mathbf{v} \in H^1(\mathcal{T}_h)$. One can rewrite (10) as $a_{DG}(\mathbf{u},\mathbf{v}) = l(\mathbf{v})$, $\forall \mathbf{v} \in H^1(\mathcal{T}_h)$. We can now define the classic discontinuous Galerkin method for Friedrichs systems with polynomial basis functions [34, 6]. Define $\mathbb{P}_q^d$ the space of polynomials of d variables, of total degree at most q and the broken polynomial space

$$\mathbb{P}_q^d(\mathcal{T}_h) := \{\mathbf{v} \in L^2(\Omega), \mathbf{v}_{|\Omega_k} \in \mathbb{P}_q^d \ \forall \Omega_k \in \mathcal{T}_h\} \subset H^1(\mathcal{T}_h).$$

Definition 3.1. *Assume $P_m(\mathcal{T}_h)$ is a finite subspace of $H^1(\mathcal{T}_h)$, for example $P_m(\mathcal{T}_h) = \mathbb{P}_q^d(\mathcal{T}_h)$. The standard upwind discontinuous Galerkin method for Friedrichs systems is formulated as follows*

$$\begin{cases} \text{find } \mathbf{u}_h \in P_m(\mathcal{T}_h) \text{ such that} \\ a_{DG}(\mathbf{u}_h,\mathbf{v}_h) = l(\mathbf{v}_h), \quad \forall \mathbf{v}_h \in P_m(\mathcal{T}_h). \end{cases} \quad (12)$$

3.2 Trefftz Discontinuous Galerkin formulation

A TDG method takes basis functions which are solutions to (4) in each cell

$$V(\mathcal{T}_h) = \{\mathbf{v} \in H^1(\mathcal{T}_h), L\mathbf{v}_k = \mathbf{0} \quad \forall \Omega_k \in \mathcal{T}_h\} \subset H^1(\mathcal{T}_h).$$

The space $V(\mathcal{T}_h)$ is a genuine subspace of $H^1(\mathcal{T}_h)$ except in the case $L = 0$. Starting from the bilinear form a_{DG}, the volume term can be written as $\int_{\Omega_k} \left(L^*(\mathbf{x})\mathbf{v}_k\right)^T \mathbf{u}_k = 2\int_{\Omega_k} \mathbf{v}_k^T R(\mathbf{x})\mathbf{u}_k$ for all $\mathbf{u},\mathbf{v} \in V(\mathcal{T}_h)$. One can therefore define another bilinear form $a_T : V(\mathcal{T}_h) \times V(\mathcal{T}_h) \to \mathbb{R}$ as

$$\begin{aligned} a_T(\mathbf{u},\mathbf{v}) = &\sum_k 2\int_{\Omega_k} \mathbf{v}_k^T R(\mathbf{x})\mathbf{u}_k + \sum_k \sum_{j<k} \int_{\Sigma_{kj}} (\mathbf{v}_k - \mathbf{v}_j)^T(M_{kj}^+(\mathbf{x})\mathbf{u}_k + M_{kj}^-(\mathbf{x})\mathbf{u}_j) \\ &+ \sum_k \int_{\Sigma_{kk}} \mathbf{v}_k^T M_k^+(\mathbf{x})\mathbf{u}_k, \quad \mathbf{u},\mathbf{v} \in V(\mathcal{T}_h). \end{aligned} \quad (13)$$

Thanks to an integration by part for functions $\mathbf{v} \in V(\mathcal{T}_h)$ which are piecewise homogeneous solutions of the equation, one gets an equivalent formulation of the bilinear form $a_T(\cdot,\cdot)$

$$a_T(\mathbf{u},\mathbf{v}) = -\sum_k \sum_{j<k} \int_{\Sigma_{kj}} (M_{kj}^-(\mathbf{x})\mathbf{v}_k + M_{kj}^+(\mathbf{x})\mathbf{v}_j)^T(\mathbf{u}_k - \mathbf{u}_j) - \sum_k \int_{\Sigma_{kk}} \mathbf{v}_k^T M_k^-(\mathbf{x})\mathbf{u}_k, \quad \mathbf{u},\mathbf{v} \in V(\mathcal{T}_h). \quad (14)$$

The relaxation term R completely disappeared in the formulation (14). It might seem a paradox at first sight but it is not because, for a Trefftz method, some information about R is encoded in the basis functions. Since there is no volume term in the formulation (14) compared to (13) it may be easier to implement. The related bilinear form $l : V(\mathcal{T}_h) \to \mathbb{R}$ is the same as in (11), that is $l(\mathbf{v}) = -\sum_k \int_{\Sigma_{kk}} \mathbf{v}_k^T M_k^-(\mathbf{x})\mathbf{g}$ for all $\mathbf{v} \in V(\mathcal{T}_h)$.

Definition 3.2. *Assume $V_m(\mathcal{T}_h)$ is a finite subspace of $V(\mathcal{T}_h)$. The upwind Trefftz discontinuous Galerkin method for the model problem (4) is formulated as follows*

$$\begin{cases} \text{find } \mathbf{u}_h \in V_m(\mathcal{T}_h) \text{ such that} \\ a_T(\mathbf{u}_h, \mathbf{v}_h) = l(\mathbf{v}_h), \quad \forall \mathbf{v}_h \in V_m(\mathcal{T}_h). \end{cases} \tag{15}$$

In the case of a time dependent problem, even if the classic upwind discontinuous Galerkin formulation (12) and the upwind Trefftz discontinuous Galerkin formulation (15) are posed on the whole space-time domain Ω, they still can be decoupled into a sequence of time-steps. This comes from the fact that the matrix A_0 is definite positive and therefore $M^-(\mathbf{n}) = 0$ if $\mathbf{n} = (1, 0, ..., 0)$. With natural notations with respect to the time step n, we define the time slice bilinear form $a_T^n : V(\mathcal{T}_h) \times V(\mathcal{T}_h) \to \mathbb{R}$

$$a_T^n(\mathbf{u}, \mathbf{v}) = -\sum_k \sum_{j<k} \int_{\Sigma_{k^n j^n}} (M^-_{k^n j^n}\mathbf{v}_k^n + M^+_{k^n j^n}\mathbf{v}_j^n)^T(\mathbf{u}_k^n - \mathbf{u}_j^n) - \sum_k \int_{\partial\Omega_S \cap \partial\Omega_{k^n}} (\mathbf{v}_k^n)^T M^-_{k^n}\mathbf{u}_k^n - \sum_k \int_{\Sigma_{k^n k^{n-1}}} (\mathbf{v}_k^n)^T M^-_{k^n k^{n-1}}\mathbf{u}_k^n$$

and the time slice linear form $l^n(\mathbf{v}) = -\sum_k \int_{\partial\Omega_S \cap \partial\Omega_{k^n}} (\mathbf{v}_k^n)^T \mathbf{g}_S - \sum_k \int_{\Sigma_{k^n k^{n-1}}} (\mathbf{v}_k^n)^T M^-_{k^n k^{n-1}}\mathbf{u}_k^{n-1}$. The convention is $\Sigma_{k^1 k^0} = \partial\Omega_{k^1} \cap (\partial\Omega \times \{0\})$ and $\Sigma_{k^{N+1} k^N} = \partial\Omega_{k^N} \cap (\partial\Omega \times \{T\})$. The space-time formulation (15) is equivalent to the series of space-only problems

$$\begin{cases} \text{find } \mathbf{u}_h^n \in V_m(\mathcal{T}_h), \ n = 1, ..., N, \text{ such that} \\ a_T^n(\mathbf{u}_h^n, \mathbf{v}_h^n) = l^n(\mathbf{v}_h^n), \quad \forall \mathbf{v}_h^n \in V_m(\mathcal{T}_h). \end{cases}$$

4 Construction of local exact solutions

In order to develop a TDG method, one needs to construct basis functions which are exact solutions to the system. It is reasonable to assume that all matrices are piecewise constants. Because it needed to distinguish the case $\sigma_a = 0$ to the case $\sigma_a > 0$, we propose 2 different methods to construct stationary basis functions: the first method constructs exponential functions as solutions of a generalized eigenproblem which comes from [38, 32, 39, 40]; the second one shows how the Birkhoff and Abu-Shumays approach can be used to deduce polynomial solutions of a general P_N model. A third method is specific to the time dependent case. A fourth method is adapted to sources.

4.1 Exponential solutions

Plug the Ansatz $\mathbf{u}(x) = \mathbf{r}e^{\lambda x}$, $\mathbf{r} \neq 0$, into the stationary P_N model. One gets the spectral problem

$$\lambda A_1 \mathbf{r} = -R\mathbf{r}. \tag{16}$$

Here λ is an eigenvalue of the matrix $-R$ in the metric associated to the matrix A_1. Both matrices are real symmetric, however A_1 is degenerated because in general it has positive, negative and vanishing eigenvalues. This is why it is a generalized eigenproblem. Now if the pair $(\mathbf{r}, \lambda)$ is a solution of the generalized eigenproblem, then $\mathbf{u}(x)$ is a one dimensional solution of the stationary P_N model and

$$\mathbf{v}(x, y) = U_\theta \mathbf{u}(x\cos\theta + y\sin\theta) \tag{17}$$

is also a solution of the stationary P_N model, which is linearly independent from the first one provided θ is not a multiple of 2π.

With the decomposition $\mathbf{r} = (\mathbf{w}^T, \boldsymbol{\chi}^T)^T$ with $\mathbf{w} \in \mathbb{R}^{m_e}$ and $\boldsymbol{\chi} \in \mathbb{R}^{m_o}$, the eigenproblem (16) is transformed into

$$-R_1\mathbf{w} = -\lambda A\boldsymbol{\chi} \text{ and } -\sigma_t\boldsymbol{\chi} = -\lambda A^T\mathbf{z}.$$

Using that $\sigma_t > 0$, elimination of $\boldsymbol{\chi}$ yields $R_1\mathbf{w} = \frac{\lambda^2}{\sigma_t} AA^T\mathbf{w}$. In our context, Proposition 2.2 guarantees that AA^T is non-singular, so one gets the reduced eigenproblem $(AA^T)^{-1}R_1\mathbf{w} = \mu\mathbf{w}$ where $\mu = \frac{\lambda^2}{\sigma_t}$. This problem is solvable using Proposition 2.2.

Theorem 4.1. *Let $\sigma_t > 0$ and $\mathbf{w}_1, ..., \mathbf{w}_{m_e} \in \mathbb{R}^{m_e}$ be the eigenvectors of the matrix $(AA^T)^{-1}R_1$ associated with the eigenvalues $\mu_1, ..., \mu_{m_e}$. Note $\boldsymbol{\chi}_i = -\sqrt{\frac{\mu_i}{\sigma_t}}A^T\mathbf{w}_i \in \mathbb{R}^{m_o}$ and $\mathbf{z}_i = (\mathbf{w}_i^T, \boldsymbol{\chi}_i^T)^T \in \mathbb{R}^m$. Let $\mathbf{d} = (\cos\theta, \sin\theta)^T \in \mathbb{R}^2$ be the direction with angle θ and U_θ be the rotation matrix (2). Then the exponential functions*

$$\mathbf{v}_i(\mathbf{x}) = U_\theta \mathbf{z}_i e^{\frac{1}{c}\sqrt{\sigma_t\mu_i}\mathbf{d}^T\mathbf{x}}, \quad i = 1, ..., m_e, \ \theta \in [0, 2\pi), \tag{18}$$

are stationary solutions to the system.

Proof. For $\theta = 0$, the function $\mathbf{v}_i$ is a solution to (16). For $\theta \neq 0$, the rotational invariance (17) yields the claim. □

Examples of exponential solutions are constructed in Section 5 for the P_1 and the P_3 models. This method can be complemented with the study of the secular equation [20] which gives sharp estimates for the eigenvalues and more information about the eigenvectors.

Proposition 4.2. *Take a finite number of directions $\mathbf{d}_k = (\cos\theta_k, \sin\theta_k)^T$ which are different $\theta_k - \theta_l \notin 2\pi\mathbb{Z}$ for $k \neq l$, and assume that $\sigma_a > 0$. Then the functions $(\mathbf{v}_i)_k(\mathbf{x}) = U_{\theta_k}\mathbf{z}_i e^{\frac{1}{c}\sqrt{\sigma_t\mu_i}\mathbf{d_k}^T\mathbf{x}}$ are linearly independent.*

Proof. The condition $\sigma_a > 0$ guarantees that $\mu_i > 0$, refer to Proposition 2.2. Then the scalar functions $e^{\frac{1}{c}\sqrt{\sigma_t\mu_i}\mathbf{d_k}^T\mathbf{x}}$ are all different and are also linearly independent. □

4.2 Harmonic polynomial solutions with Birkhoff and Abu-Shumays work

If $\sigma_a = 0$, then $\mu_1 = 0$ by Proposition 2.2 so the exponential factor degenerates (is equal to 1). It results in linearly dependent functions $(\mathbf{v}_1)_k$ (at least if the number of directions is strictly greater than the size of the system m). This is a critical issue in view of implementation, because linearly dependent basis functions yield singular matrices after discretization. The situation is the same as the one described in [14] for the plane wave basis of the Helmholtz equation when the frequency tends to zero: in the cited reference, the authors show that convenient rescaling of the exponential functions yields special polynomial functions. In our case, we construct polynomial solutions for $\sigma_a = 0$ and $\sigma_s > 0$ with the Birkhoff and Abu-Shumays approach.

The series of all harmonic polynomials is generated as follows: firstly set $q_1(x, y) = 1$, then consider the series for

$$q_{2k}(\mathbf{x}) = \frac{1}{k!}\Re((x - x_0) + i(y - y_0))^k \text{ and } q_{2k+1}(\mathbf{x}) = \frac{1}{k!}\Im((x - x_0) + i(y - y_0))^k \tag{19}$$

All these polynomials are harmonic, that is $\Delta q_k = 0$. Define the function I which depends on a given harmonic polynomial q

$$I(x, y, \boldsymbol{\Omega}) := \sum_{k=0}^{\infty} \left(\frac{-1}{\sigma_s}\right)^k \Big(\boldsymbol{\Omega} \cdot \nabla\Big)^k q(x, y), \quad l \geq 0, \tag{20}$$

where $\boldsymbol{\Omega} := (\sin\phi\cos\psi, \sin\phi\sin\psi, \cos\phi)^T \in \mathbb{R}^3$ with $\psi \in [0, 2\pi)$ and $\phi \in [0, \pi)$. The series is finite and the function I is a polynomial with respect to x and y: its degree is $\deg(q)$ which is the degree of the polynomial q. The terms of the series can be evaluated with the following formula.

Lemma 4.3 (Proof in [20]). *For $k \geq 1$, one has $\Big(\boldsymbol{\Omega} \cdot \nabla\Big)^k q(x, y) = \left(\frac{\sin\phi}{2}\right)^k \Big(e^{-ik\psi}(\partial_x + i\partial_y)^k + e^{ik\psi}(\partial_x - i\partial_y)^k\Big)q(x, y)$.*

Proposition 4.4 ([39]). *For $\sigma_a = 0$ and $\sigma_s > 0$, $I(x, y, \boldsymbol{\Omega})$ is solution to the stationary transport equation (1).*

Proof. Equation (20) yields $I + \frac{1}{\sigma_s}\boldsymbol{\Omega} \cdot \nabla I = q$. But Lemma 4.3 also yields that $< I >= q$. So $\boldsymbol{\Omega} \cdot \nabla I = \sigma_s(< I > - I)$ which is the claim. □

Following Hermeline [2, Appendix A], we redefine the vector $\mathbf{y}(\boldsymbol{\omega})$ as the collection of real valued orthonormal spherical harmonics $Y_{k,l}$ for $k \leq N$ and $k + l$ even. One can check that $\mathbf{y}(\boldsymbol{\Omega}) \in \mathbb{R}^m$ with $m = \frac{1}{2}(N + 1)(N + 2)$. We denote Π_N the L^2 orthogonal projection onto the space of these particular spherical harmonics.

Theorem 4.5 (Proof in [20]). *Take $\sigma_a = 0$, $\sigma_s > 0$. The function $\mathbf{I}_N(x, y) =< \mathbf{y}(\boldsymbol{\Omega})\Pi_N I(x, y, \boldsymbol{\Omega}) > \in \mathbb{R}^m$ is a solution of the P_N model and is a harmonic polynomial with respect to x, y.*

We refer to [19] for additional details. Examples of such solutions are in Section 5 for the P_1 and the P_3 models.

4.3 Time dependent solutions

We give some possible ways to get time dependent solutions to the P_N model which can be used as basis functions for the TDG method when considering a space-time mesh. Other time dependent solutions can be constructed starting from [40]. Once again, we take $\varepsilon = 1$ for simplicity.

A general form is

$$\mathbf{v}(t, \mathbf{x}) = \mathbf{g}(\mathbf{x}) e^{\alpha t}, \tag{21}$$

where $\alpha \in \mathbb{R}$ is arbitrary and $\mathbf{g}$ is a polynomial function of $\mathbf{x}$. One can inject this solution in the P_N model. One gets after removing the exponentials $\Big(A_1 \partial_x + A_2 \partial_y + (R + \alpha I_m)\Big)\mathbf{g}(\mathbf{x}) = \mathbf{0}$ where I_m is the identity matrix of $\mathbb{R}^{m\times m}$. The function $\mathbf{g}(\mathbf{x})$ is very similar to the previous stationary solutions. The matrix R is just replaced by the matrix $\widetilde{R} := R + \alpha I_m$. If α is an eigenvalue of the matrix $-R$, then $\widetilde{R}$ is a non trivial kernel and $\mathbf{g}$ can be taken as a constant-in-space vector (in the kernel of $\widetilde{R}$).

Another second possibility is to start from a one dimensional solution under the form $\mathbf{v}(t, x) = \mathbf{q}(t, x) e^{\lambda x}$, where $\mathbf{q}(t, x) \in \mathbb{R}^m$ is polynomial vector in x and t. A concrete example is given in [18, Proposition 4.2] for the case of the P_1 model. Using rotational invariance, one gets the family $\mathbf{v}(t, \mathbf{x}) = U_\theta \mathbf{q}(t, x\cos\theta + y \sin\theta) e^{\lambda(x\cos\theta + y\sin\theta)}$. Another possibility is to look for more general two dimensional solutions under the form $\mathbf{v}(t, \mathbf{x}) = \mathbf{p}(t, x, y) e^{\lambda(x\cos\theta + y\sin\theta)}$, where $\mathbf{p}(t, x, y) \in \mathbb{R}^m$ is polynomial vector in x, y and t.

4.4 Sources

For non homogeneous problems where there is a non zero source right hand side $\mathbf{f}$, it is valuable to add specific basis functions which are like P_0 Finite Volume basis functions. We refer to Section 2-2.3 of Chapter 2 in [19] for a more detailed presentation. Instead of a general theory, it is sufficient to take an example. Consider the lattice geometry, test case (7.3), and the equation (33), then one may add the basis function $\mathbf{v}_f = R^{-1}\mathbf{f}$ in the central region where R is non singular (where $\mathbf{e}_1 = (1, 0, \dots, 0)$ has only one non zero component)

$$\mathbf{v}_f(\mathbf{x}) = \mathbf{e}_1 \text{ for } \mathbf{x} \in [3, 4]^2 \quad \text{and} \quad \mathbf{v}_f(\mathbf{x}) = \mathbf{0} \text{ everywhere else.} \tag{22}$$

5 Stationary solutions to the P_1 and P_3 models

In this section, we construct explicit stationary solutions to the P_1 and P_3 models. For convenience, we reintroduce the scaling parameter (6) $\varepsilon \in (0, 1]$.

5.1 The P_1 model

We recall that the matrices read $A = \begin{pmatrix} 0 & \frac{1}{\sqrt{3}} \end{pmatrix}$, $B = \begin{pmatrix} \frac{1}{\sqrt{3}} & 0 \end{pmatrix}$, $R_1 = \varepsilon\sigma_a$ and $R_2 = \begin{pmatrix} \sigma_t & 0 \\ 0 & \sigma_t \end{pmatrix}$. We are interested in the stationary solutions to the P_1 model.

Proposition 5.1. *Take $\mathbf{d}_k = (\cos\theta_k, \sin\theta_k)^T \in \mathbb{R}^2$. Let P the permutation matrix $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Assume $\sigma_a > 0$ the following functions are solutions to the P_1 model*

$$\mathbf{v}_k = \begin{pmatrix} \sqrt{\sigma_t} \\ -\sqrt{\varepsilon\sigma_a} P \mathbf{d}_k \end{pmatrix} e^{\frac{1}{c}\sqrt{3\varepsilon\sigma_a\sigma_t}\mathbf{d}_k^T \mathbf{x}}, \quad \sigma_t = \varepsilon\sigma_a + \frac{\sigma_s}{\varepsilon}. \tag{23}$$

Assume $\sigma_a = 0$. The functions below are solutions to the P_1 model

$$\mathbf{v}_k = \begin{pmatrix} \frac{\sigma_s}{\varepsilon} q_k \\ -\frac{c}{\sqrt{3}} P \nabla q_k \end{pmatrix}. \tag{24}$$

The direct verification that harmonic polynomials of any order are solutions of the P_1 equations for $\sigma_a = 0$ is evident. However this is also a consequence of the general Theorem 4.5.

5.2 The P_3 model

The calculations for the proofs can be verified from [20, 19]. First, we give the stationary exponential solutions.

Proposition 5.2. *Take* $\mathbf{d}_k = (\cos\theta_k, \sin\theta_k)^T \in \mathbb{R}^2$. *The following functions are solutions to the* P_3 *model*

$$\mathbf{v}_1(\mathbf{x}) = \begin{pmatrix} 0 \\ -\sqrt{30}\cos 2\theta_k \\ 0 \\ \sqrt{30}\sin 2\theta_k \\ \sqrt{14}\cos\theta_k \\ -\sqrt{14}\sin\theta_k \\ \sqrt{15}\cos 3\theta_k \\ -\cos\theta_k \\ \sin\theta_k \\ -\sqrt{15}\sin 3\theta_k \end{pmatrix} e^{\frac{1}{c}\sqrt{\frac{7}{3}}\sigma_t \mathbf{d}_k^T\mathbf{x}}, \quad \mathbf{v}_2(\mathbf{x}) = \begin{pmatrix} 0 \\ \sqrt{2}\sin 2\theta_k \\ \sqrt{6} \\ \sqrt{2}\cos 2\theta_k \\ 0 \\ 0 \\ -\sqrt{3}\sin 3\theta_k \\ -\sqrt{5}\sin\theta_k \\ -\sqrt{5}\cos\theta_k \\ -\sqrt{3}\cos 3\theta_k \end{pmatrix} e^{\frac{1}{c}\sqrt{7}\sigma_t \mathbf{d}_k^T\mathbf{x}},$$

$$\mathbf{v}_3(\mathbf{x}) = \begin{pmatrix} \frac{\sqrt{\sigma_t}}{14\sqrt{15}}\rho^+ \\ \varepsilon\sqrt{\sigma_t}\sigma_a \sin 2\theta_k \\ -\frac{\varepsilon\sqrt{\sigma_t}\sigma_a}{\sqrt{3}} \\ \varepsilon\sqrt{\sigma_t}\sigma_a \cos 2\theta_k \\ -\frac{1}{630\sqrt{2}}\upsilon^-\tau^+ \sin\theta_k \\ -\frac{1}{630\sqrt{2}}\upsilon^-\tau^+ \cos\theta_k \\ -\frac{\varepsilon}{2\sqrt{21}}\sigma_a\upsilon^- \sin 3\theta_k \\ \frac{\varepsilon}{2\sqrt{35}}\sigma_a\upsilon^- \sin\theta_k \\ \frac{\varepsilon}{2\sqrt{35}}\sigma_a\upsilon^- \cos\theta_k \\ -\frac{\varepsilon}{2\sqrt{21}}\sigma_a\upsilon^- \cos 3\theta_k \end{pmatrix} e^{\frac{1}{c}\upsilon^-\sqrt{\frac{\sigma_t}{18}}\mathbf{d}_k^T\mathbf{x}}, \quad \mathbf{v}_4(\mathbf{x}) = \begin{pmatrix} \frac{\sqrt{\sigma_t}}{14\sqrt{15}}\rho^- \\ \varepsilon\sqrt{\sigma_t}\sigma_a \sin 2\theta_k \\ -\frac{\varepsilon\sqrt{\sigma_t}\sigma_a}{\sqrt{3}} \\ \sqrt{\sigma_t}\sigma_a \cos 2\theta_k \\ -\frac{\sqrt{\varepsilon}}{630\sqrt{2}}\upsilon^+\tau^- \sin\theta_k \\ -\frac{\sqrt{\varepsilon}}{630\sqrt{2}}\upsilon^+\tau^- \cos\theta_k \\ -\frac{\sqrt{\varepsilon}}{2\sqrt{21}}\sigma_a\upsilon^+ \sin 3\theta_k \\ \frac{\sqrt{\varepsilon}}{2\sqrt{35}}\sigma_a\upsilon^+ \sin\theta_k \\ \frac{\sqrt{\varepsilon}}{2\sqrt{35}}\sigma_a\upsilon^+ \cos\theta_k \\ -\frac{\sqrt{\varepsilon}}{2\sqrt{21}}\sigma_a\upsilon^+ \cos 3\theta_k \end{pmatrix} e^{\frac{1}{c}\upsilon^+\sqrt{\frac{\sigma_t}{18}}\mathbf{d}_k^T\mathbf{x}},$$

with $\sigma_t = \varepsilon\sigma_a + \frac{\sigma_s}{\varepsilon}$, $\kappa = \sqrt{605\varepsilon^2\sigma_a^2 + 14\varepsilon\sigma_a\sigma_t + 245\sigma_t^2}$, $\upsilon^\pm = \sqrt{55\varepsilon\sigma_a + 35\sigma_t \pm \sqrt{5}\kappa}$, $\tau^\pm = \sqrt{5}\varepsilon\sigma_a + 35\sqrt{5}\sigma_t \pm 5\kappa$ *and* $\rho^\pm = (\upsilon^\pm)^2 - 110\varepsilon\sigma_a$. *For* $\sigma_a = 0$, *then* $\upsilon^- = \mu_3 = 0$ *and the exponential functions associated with the third family have a degeneracy (as expected from Proposition 2.2).*

Proposition 5.3. *Take* $\sigma_a = 0$. *The polynomial functions below are solutions*

$$\begin{aligned} &\mathbf{v}_1(\mathbf{x}) = \left(1,0,0,0,0,0,0,0,0,0\right), && \mathbf{v}_2(\mathbf{x}) = \left(\sigma_t x,0,0,0,0,-\tfrac{c}{\sqrt{3}},0,0,0,0\right), \\ &\mathbf{v}_3(\mathbf{x}) = \left(\sigma_t y,0,0,0,-\tfrac{c}{\sqrt{3}},0,0,0,0,0\right), && \mathbf{v}_4(\mathbf{x}) = \left(\sigma_t^2 xy,\tfrac{2c^2}{\sqrt{15}},0,0,-\tfrac{\sigma_t c}{\sqrt{3}}x,-\tfrac{\sigma_t c}{\sqrt{3}}y,0,0,0,0\right), \\ &\mathbf{v}_5(\mathbf{x}) = \left(\tfrac{1}{2}\sigma_t^2(x^2-y^2),0,0,\tfrac{2c^2}{\sqrt{15}},\tfrac{\sigma_t c}{\sqrt{3}}y,-\tfrac{\sigma_t c}{\sqrt{3}}x,0,0,0,0\right). \end{aligned} \tag{25}$$

6 High order convergence (stationary case)

The main results of this section are, on the one hand the Theorem 6.4 which establish the h-convergence of the TDG method applied to the P_N model when $\sigma_a > 0$, and, on the other hand, the numerical tests of h-convergence which confirm the theoretical analysis.

We consider a series of meshes $\mathcal{T}_h^n$, $n \in \mathbb{N}$. For a polygonal cell $\Omega_j^n \in \mathcal{T}_h^n$, we define h_j^n the size of its larger edge and ρ_j^n the radius of the largest circle that can be inscribed within Ω_j. The sequence of meshes verifies $h^n := \max_j h_j^n \to 0$ as $n \to \infty$. It is quasi-uniform, that is there exists a constant $C_\tau \in \mathbb{R}^+$ such that $\max_{j,n} \frac{h_j^n}{\rho_j^n} \leq C_\tau$. For the simplicity of notation, the index n is removed in the following. The coefficients σ_a and σ_s are bounded: there exists $C_\sigma \in \mathbb{R}^+$ such that $\sigma_a \leq C_\sigma$ and $\sigma_s \leq C_\sigma$. We also take $\varepsilon = 1$ and $c = 1$. The material below is organized in subsections which are, general bounds, convergence bounds for $\sigma_a > 0$, convergence bounds for $\sigma_a = 0$ and finally numerical tests.

6.1 General bounds

Proposition 6.1. *Let* $\mathbf{u} = (\mathbf{u}_e, \mathbf{u}_o) \in W^{k+1,\infty}(\Omega)$ *be a local solution to the stationary* P_N *model. Let* $\omega \subset \Omega$ *with* $h = diam(\omega)$. *Assume* $\sigma_a > 0$ *and* $\sigma_s > 0$ *and consider the basis functions constructed in Proposition 4.2 for* $2k+1$

different directions

$$0 \leq \theta_1 < \theta_2 < \cdots < \theta_{2k} < \theta_{2k+1} < 2\pi.$$

It yields $(2k+1)m_e$ *solutions* $\mathbf{v}_1, ..., \mathbf{v}_{(2k+1)m_e} \in W^{k+1,\infty}(\omega)$ *decomposed as* $\mathbf{v}_i = (\mathbf{v}_i^e, \mathbf{v}_i^o)$ *for* $1 \leq i \leq (2k+1)m_e$.

Generically, with an additional hypothesis on the linear independance of the coefficients of the Taylor expansion of the $\mathbf{v}_i^e$*, there exists a vector* $\mathbf{a} = (a_i)^T \in \mathbb{R}^{(2k+1)m_e}$ *such that*

$$\left\| \mathbf{u}^e - \sum_{i=1}^{(2k+1)m_e} a_i \mathbf{v}_i^e \right\|_{L^\infty(\omega)} \leq C h^{k+1} \|\mathbf{u}^e\|_{W^{k+2,\infty}(\Omega)}, \tag{26}$$

and $\left\| \mathbf{u}^o - \sum_{i=1}^{(2k+1)m_e} a_i \mathbf{v}_i^o \right\|_{L^\infty(\Omega_j)} \leq C h^k \|\mathbf{u}^e\|_{W^{k+2,\infty}(\Omega_j)}$.

If $\sigma_a = 0$ *and* $\sigma_S > 0$*, the same result holds after replacement of the* $2k+1$ *degenerate exponentials -Proposition 2.2- by* $2k+1$ *harmonic polynomial solutions of degree at most* k *-Theorem 4.5 and Definition (19)-.*

The additional hypothesis on the linear independence holds if one adds $2(N-1)$ *directions. More precisely among the corresponding* $(2(k+N)-1)m_e$ *basis functions, there exist* $(2k+1)m_e$ *basis functions which satisfy the hypothesis.*

Sketch of the proof. It is based on the second order form of the P_N model. For $\sigma_t > 0$ and $\mathbf{u}$ regular enough, one has the decomposition

$$\Big(A\partial_x + B\partial_y\Big)\mathbf{u}_o(\mathbf{x}) = -R_1\mathbf{u}_e(\mathbf{x}) \quad \text{and} \quad \Big(A^T\partial_x + B^T\partial_y\Big)\mathbf{u}_e(\mathbf{x}) = -R_2\mathbf{u}_o(\mathbf{x}). \tag{27}$$

It is equivalent to saying that $\mathbf{u}_e(\mathbf{x})$ is a solution of the second order form of the P_N model

$$\Big(AA^T\partial_{xx} + (AB^T + BA^T)\partial_{xy} + BB^T\partial_{yy}\Big)\mathbf{u}_e(\mathbf{x}) = \sigma_t R_1 \mathbf{u}_e(\mathbf{x}). \tag{28}$$

The functions $\mathbf{v}_i^e$ satisfy the same equation. To obtain (26) it is sufficient to find a vector $\mathbf{a} = (a_i)$ such that all terms of the Taylor expansion of $\sum_{i=1}^{(2k+1)m_e} a_i \mathbf{v}_i^e - \mathbf{u}^e$ vanish up to total degree $k+1$. Since this vectorial function is made of m_e scalar functions, the number of Taylor coefficients that one must annihilate (by conveniently choosing the a_i's) is equal to $N_\# = \frac{(k+1)(k+2)}{2} m_e$. However the equation (28) yields many linear relations between the coefficients of the Taylor expansion of $\mathbf{u}_e$ (and the same linear relations for the coefficients of the Taylor expansion of $\mathbf{v}_i^e$). A counting argument shows that it yields $N_\% = \frac{(k-1)(k)}{2} m_e$ linear relations.

Generically these linear relations are linearly independent: this is not immediate to prove, we refer to [19] for the P_N system, to [12, 41] for the Helmholtz equation. It is proved that, for the Helmholtz equation or the P_1 system, the linear independence of basis functions holds up except on a set of measure zero. Then, the claim (26) is obtained with a total number of basis functions

$$N_{\text{tot}} = N_\# - N_\% = (2k+1)m_e.$$

Eliminating the first expression in (27), a similar result is obtained for the approximation of $\mathbf{u}_o$, but with a loss of one order of approximation. The last part of the proposition is proved in [19], using the Bézout theorem for system of multivariate polynomials. □

Let us denote the space of approximation $V_h := \bigoplus_{\Omega_j \in \mathcal{T}_h} \text{Span}\left\{\mathbf{v}_1, ..., \mathbf{v}_{(2k+1)m_e}\right\}\mathbf{1}_{\Omega_j}$ where $\mathbf{1}_{\Omega_j}$ is the indicatrix function of the cell Ω_j. The norms $\|\dots\|_{DG}$ and $\|\dots\|_{DG^*}$ are defined in the appendix in (40).

Proposition 6.2. *There exists* $\mathbf{v}_h \in V_h$ *such that* $\|\mathbf{u} - \mathbf{v}_h\|_{DG^*} \leq C h^{k-1/2} \|\mathbf{u}\|_{W^{k+1,\infty}(\Omega)}$.

Proof. Combine (B.10) with Proposition 6.1. □

Proposition 6.3. *Let* $\mathbf{u}_h$ *be the solution generated by the TDG method. One has* $\|\mathbf{u} - \mathbf{u}_h\|_{DG} \leq C h^{k-1/2} \|\mathbf{u}\|_{W^{k+1,\infty}(\Omega)}$.

Proof. It comes from the quasi-optimality estimate of Proposition B.6 combined with the previous bound. □

6.2 Convergence estimates for $\sigma_a > 0$

Theorem 6.4 (Convergence of the TDG method for the P_N model)**.** *Assume* $\sigma_a > 0$ *and the previous hypotheses. One has*

$$\|\mathbf{u} - \mathbf{u}_h\|_{L^2(\Omega)} \leq C h^{k-1/2} \|\mathbf{u}\|_{W^{k+1,\infty}(\Omega)}. \tag{29}$$

Proof. Use (B.8) and Proposition 6.3. □

This Theorem shows a remarkable property of the TDG method: the number of additional basis functions to gain one order of convergence from k to $k+1$ does depend linearly on k, . This is not the case for the standard DG method where the number of additional basis functions increases quadratically with k. The table in the introduction summarizes these findings.

6.3 Convergence for $\sigma_a = 0$

In case the absorption coefficient is zero, then the matrix R becomes singular and the previous proof is no longer applicable. In simple cases, it is possible to bypass this technical obstruction. We refer to [18] for such an estimate for $N = 1$, that is for the P_1 model. The rate of convergence is reduced by 1 with respect to the general case (29).

6.4 A numerical test of convergence

Here we display a test of convergence which confirms the theoretical results of the Theorem 6.4 with a number of basis functions equal to $(2k+1)m_e$ with $2k+1$ equidistributed directions. This holds also for all numerical results in this paper.

Consider the stationary P_1 model in two dimensions. Let $\mathbf{x} = (x, y)^T, \Omega = [0, 1]^2$, $\sigma_a = 1/\sqrt{3}, \sigma_s = 1/\sqrt{3}$. The exact solution we consider here is $\mathbf{u}_{ex}(\mathbf{x}) = \left(\cos(y)e^{\sqrt{3}x}, -(\sqrt{3}/2)\cos(y)e^{\sqrt{3}x}, 0.5\sin(y)e^{\sqrt{3}x}\right)^T$. Results obtained with $3, 5$ and 7 basis functions are displayed on the left of Figure 3. As stated in Theorem 6.4 for the particular case $N = 1$, one only needs two additional basis functions to increase the order by 1. Note however that the orders obtained here are slightly better than those predicted in Theorem 6.4: with 3, 5 and 7 basis functions, one gets respectively orders 0.8, 1.5 and 2.5.

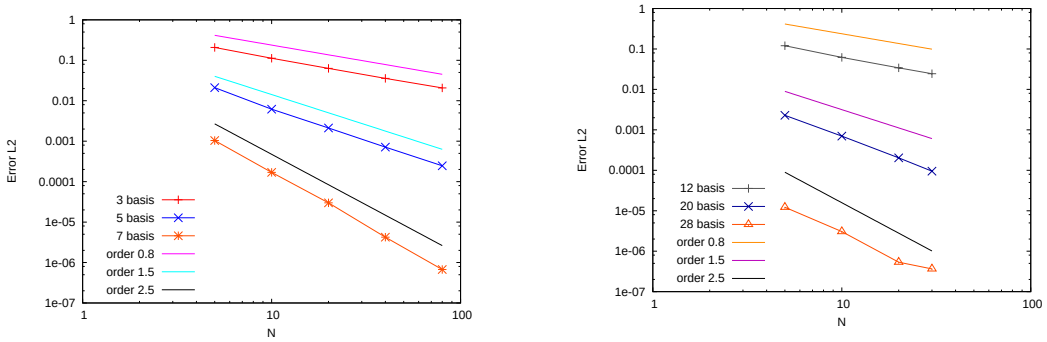

Figure 3: Order depending on the number of basis functions with respect to the number of cells per unit length (denoted as N). On the left P_1 model and on the right P_3 model. L^2 error in logarithmic scale and random meshes.

On the right of Figure 3, we consider the stationary P_3 model in two dimensions. Let $\mathbf{x} = (x, y)^T, \Omega = [0, 1]^2$, $\sigma_a = 0.2, \sigma_s = 0.3$. The exact solution we consider is taken from the solution (18) and has eigenvalue $\sqrt{7}/\sqrt{3}$ with a direction $\mathbf{d} = (\cos\pi/4, \sin\pi/4)^T$ which does not belong to our basis functions. Results obtained with $3, 5$ and 7 directions (for a total of 12, 20 and 28 basis functions) are displayed on the right of Figure 3. The maximal number of degrees of freedom is $N_{\text{tot}} \approx 12^2 \times 28^2 = 112896$, so the test is already expensive on CPU grounds. The order obtained are close to those predicted by Theorem 6.4. Note however that the tests for the P_3 model are displayed on much coarser meshes than for the P_1 model. This comes from the bad condition number of the matrix, which is a well known drawback of the TDG method [12, 13, 19] and occurs when increasing the number of basis functions on fine meshes. Since we do not want the condition number to interfere with the error study we chose not to refine the meshes too much. Still, the bad conditioning of the matrix can probably be seen on the last point of the curve representing 28 basis functions, which is not completely aligned with the other points. Using better preconditioners could solve this issue.

7 Numerical results

Various h-convergence results (theoretical and numerical) are available in the literature for TDG methods for time harmonic equations [12, 11, 14], p-convergence is analyzed in [42]. For the family of Friedrichs systems evaluated in this work, h-convergence can be found in [5, 18, 19].

When the scaling parameter (6) tends to zero ($\varepsilon \to 0$), the model problem admits a diffusion limit [2, 28]. General references which provide accurate numerical methods for the diffusion limit are [31, 28, 32, 33] for asymptotic-preserving methods. In principle, the Trefftz method may be very efficient in the diffusion limit since the exact solutions in the cell have a perfect balance between the transport terms (matrices A_1 and A_2) and the relaxation (matrix R). A simple proof that the Trefftz scheme is indeed Asymptotic Diffusion Preserving can be found in [18].

The relaxation matrix $R(\mathbf{x})$ can be discontinuous in applications. This is typical of the physics of transfer at the interface between two different materials and of neutron propagation: in the application illustrated at the end of this work, the unknown $\mathbf{u}$ comes from an angular discretization of the population of neutrons and the relaxation coefficients model the interaction of neutrons with matter; the issue is that the materials are different on both sides of an interface. Boundary layers may occur when σ_a and/or σ_s vary significantly and the transport equation tends to a diffusion limit when σ_s is high. These two phenomena are challenging for numerical methods and research on devising numerical methods which perform better in these regimes continues to be pursued by the scientific community. The literature is scarce on numerical methods for boundary layers. It has been highlighted in [18] that the TDG method naturally leads to schemes adapted to such problems.

7.1 A test problem with a boundary layer

In this test, taken from [18], a two dimensional test problem with discontinuous coefficients is studied and we focus here on the results obtained with the P_3 model. The structure of the numerical code is classical: assemble matrix and right hand side, invert matrix, display results. It is described in [19]. The domain is $\Omega = [0,1]^2$ and we define Ω_1 (resp. Ω_2) as $\Omega_1 = [0.35, 0.65]^2$ (resp. $\Omega_2 = \Omega \setminus \Omega_1$). We take $\varepsilon = 1$, $c = 1$, $\sigma_a = 2 \times \mathbf{1}_{\Omega_1}(\mathbf{x})$ and $\sigma_s = 2 \times \mathbf{1}_{\Omega_2}(\mathbf{x}) + 10^5 \times \mathbf{1}_{\Omega_1}(\mathbf{x})$. The absorption coefficient has compact support in Ω_1 while the scattering coefficient is discontinuous and takes a high value in Ω_1. These coefficients involve a discontinuous matrix R. When considering random meshes, the interface between Ω_1 and Ω_2 is maintained as a straight line. The geometry and parameters of this test are represented in Figure 4.

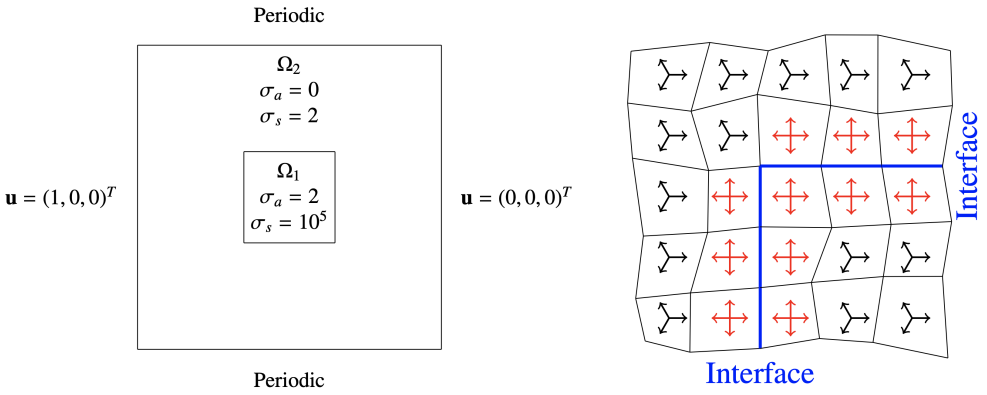

Figure 4: On the left: Domain and boundary condition for the two dimensional boundary layers test. On the right: representation of adaptive directions at the interface. In this example: the 3 equi-distributed directions (30) in each cell except at the interface where the directions are locally adapted into (31).

For the directions, we may consider the following 3-equidistributed directions

$$\mathbf{d}_1 = (1,0),\ \mathbf{d}_2 = (\cos\frac{2\pi}{3}, \sin\frac{2\pi}{3}),\ \mathbf{d}_3 = (\cos\frac{4\pi}{3}, \sin\frac{4\pi}{3}), \tag{30}$$

the 4-equidistributed directions

$$\mathbf{d}_1 = (1,0),\ \mathbf{d}_2 = (0,1),\ \mathbf{d}_3 = (-1,0),\ \mathbf{d}_4 = (0,-1), \tag{31}$$

or the 5-equidistributed directions

$$\mathbf{d}_1 = (1,0),\ \mathbf{d}_2 = (\cos\frac{2\pi}{5}, \sin\frac{2\pi}{5}),\ \mathbf{d}_3 = (\cos\frac{4\pi}{5}, \sin\frac{4\pi}{5}),\ \mathbf{d}_4 = (\cos\frac{6\pi}{5}, \sin\frac{6\pi}{5}),\ \mathbf{d}_5 = (\cos\frac{8\pi}{5}, \sin\frac{8\pi}{5}). \tag{32}$$

As pointed out in [18], the choice of directions at the interface plays an important role in correctly capturing the boundary layers. It is essential to locally get the one dimensional direction perpendicular to the interface associated with the

boundary layer. We make the special choice of directions (31) at the interface. Such directions are well adapted if one considers the one dimensional problem at the interface. A graphical illustration of the adaptive directions at the interface is provided on the right of Figure 4. As stated previously, when $\sigma_a = 0$ the degenerate exponentials are replaced with polynomials. With our parameters, the number of polynomials used in the basis functions is equal to the number of directions.

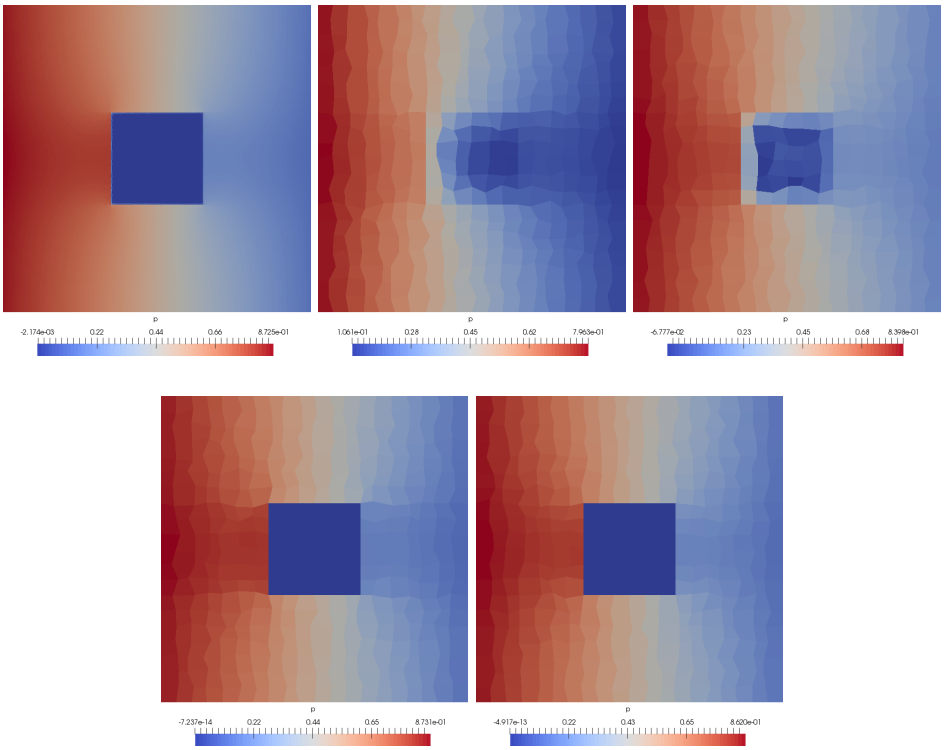

Figure 5: $\boldsymbol{P_3}$ **model**. Representation of the first variable for the P_3 model. Top left: reference solution. Top center: DG scheme with 10 basis functions per cell. Top right: DG scheme with 30 basis functions per cell. Bottom left: TDG scheme with 12 basis functions per cell. Bottom right: TDG scheme with 20 basis functions per cell. For the TDG scheme, the directions at the interface are locally adapted into the 4 directions (31).

The reference solution is calculated on a 200×200 random mesh with the 3 directions (30) and adaptive directions (31) at the interface. The following cases are calculated on a coarse 20×20 mesh

- The DG method with constant basis functions only (= finite volume) for a total of 10 basis functions per cell.
- The DG method with affine basis functions (that is 1, x, y) for a total of 30 basis functions per cell.
- The TDG method with the basis functions of Propositions 5.2 and 5.3 depending on the 3 directions (30) (for a total of 12 basis functions per cell) and on the 4 directions (31) at the interface (for a total of 16 basis functions per cell).
- The TDG method with the basis functions of Propositions 5.2 and 5.3 depending on the 5 directions (32) (for a total of 20 basis functions per cell) and the 4 directions (31) at the interface.

The results are given in Figure 5. One notices a better approximation of the solution for the TDG method with less degrees of freedom compared to the standard DG scheme. If the TDG method gives such good result, it is in fact because the correct exponential solutions (*i.e.* with the right directions) are locally used in the boundary layers. Actually, an enrichment strategy, where the DG basis is locally (in the boundary layers) enriched with some exponential solutions, would give similar result on this numerical test [19, Section 5-4.3.2]. The same kind of idea is used, for example, in the context of the so-called extended finite element method (XFEM) [43, 44].

7.2 A test problem in the regime $\sigma_a = 0$

For the P_3 model we compare the results obtained with the DG and TDG method on a 80×80 mesh. The time step is $\Delta t = T/80$. We consider four different results:

- The limit solution which is the fundamental solution of the 2D heat equation.
- The DG method with constant basis functions only (= finite volume) for a total of 10 basis functions per cell.

• The DG method with affine basis function (that is 1, x and y) for a total of 30 basis functions per cell.
• The TDG method initially with the basis made of three directions for the exponential functions. But the degenerate exponential function (that is $\mathbf{v}_3$ in Proposition 5.2) is systematically removed. Instead it is replaced by the first three polynomial functions $\mathbf{v}_{1,2,3}$ in (25) That is the TDG method applied to the P_3 model uses a combination of exponential and polynomial basis functions. For an implementation with five directions, one should replace the degenerate exponentials with $\mathbf{v}_{1,2,3,4,5}$.

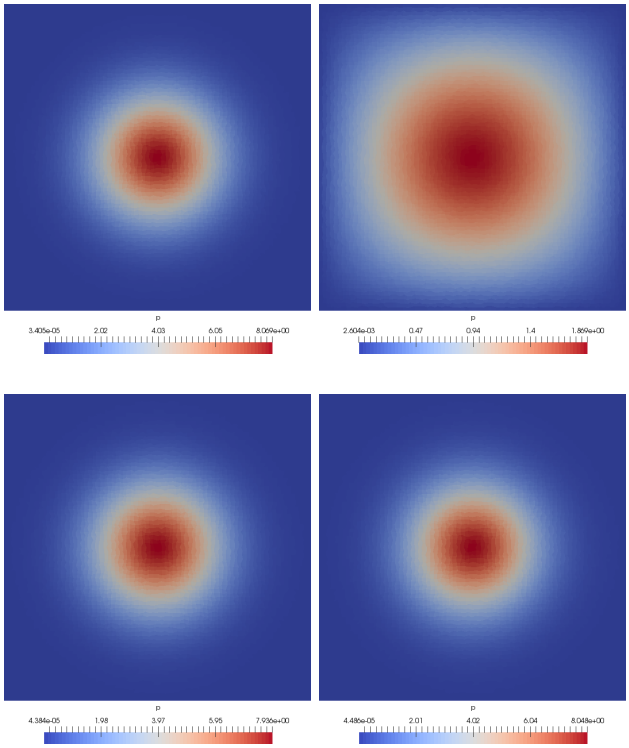

Figure 6: Top left: limit solution. Top right: DG-P_0=FV solution with 10 basis functions per cell. Bottom left: DG-P_1 solution with 30 basis functions per cell. Bottom right: TDG solution with 12 basis functions per cell.

The results presented in Figure 6 illustrate that the DG method with only constant basis functions is too diffusive. On the contrary, one recovers a good approximation with the TDG method. This illustrates the good behavior of TDG approximations for such problems. The DG scheme with affine basis functions is also accurate, but with the disadvantage of using approximately three times more basis functions than the TDG scheme.

7.3 A lattice problem

We consider a lattice problem [28, 35, 2, 45]. The spatial domain $\Omega_S = [0,7] \times [0,7]$ is represented in Figure 7 and we take $T = 3.2$. The white area is a purely scattering region while the striped and black areas are purely absorbing regions. Additionally, the black region contain a source of particles. More precisely, let Ω_c be the union of the eleven striped squares and the black square in Figure 7, then one has

$$\begin{cases} \sigma_a(\mathbf{x}) = 10, & \sigma_s(\mathbf{x}) = 0, \quad \text{if } \mathbf{x} \in \Omega_c, \\ \sigma_a(\mathbf{x}) = 0, & \sigma_s(\mathbf{x}) = 1, \quad \text{else.} \end{cases}$$

Note that for some authors $\sigma_a = 0$, $\sigma_s = 1$, in the central region [35, 2] while other authors take $\sigma_a = 10$, $\sigma_s = 0$ [28, 45]. These two choices give similar numerical results and we consider here the second option. We recall that Friedrichs systems with a source term read

$$\Big(\partial_t + A_1\partial_x + A_2\partial_y\Big)\mathbf{u}(t,\mathbf{x}) = -R\mathbf{u}(t,\mathbf{x}) + \mathbf{f}(\mathbf{x}). \tag{33}$$

In this example, the source $\mathbf{f}(\mathbf{x}) \in \mathbb{R}^m$ is contained in the black region with $\mathbf{f}(\mathbf{x}) = \sigma_a(\mathbf{x}) \times \mathbf{e}_1$ for $\mathbf{x} \in [3,4]^2$ and $\mathbf{f}(\mathbf{x}) = \mathbf{0}$ everywhere else, where $\mathbf{e}_1 = (1,0,...,0)^T \in \mathbb{R}^m$. The boundary conditions are of Dirichlet type, that is $\mathbf{u} = \mathbf{0}$ at the boundaries of the domain and play essentially no role in the problem for the final time considered.

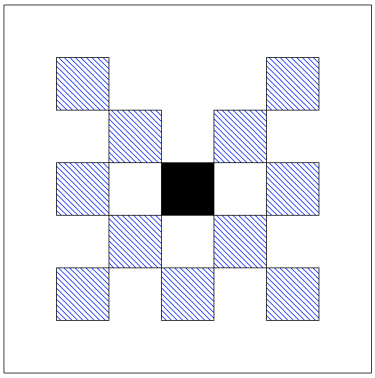

Figure 7: Domain for the lattice problem 7.3.

Following the convention coming from [28, 35, 2, 45], the density of particle (also called the first moment, that is the first component of U) is plotted on a logarithmic color scale. This emphasises the numerical diffusion at the propagation front of the different numerical methods.

7.3.1 The P_1 model.

The numerical results obtained for the P_1 model are displayed in Figure 8. The reference solution is computed with the DG method with affine basis functions for a total of 9 basis functions per cell on a 280×280 random mesh with $dt = 0.01$. We compare the DG and TDG methods on a 140×140 mesh with $dt = 0.02$. We consider the following cases
• The DG method with constant basis functions only for a total of 3 basis functions per cell.
• The DG method with affine basis functions (that is 1, x, y) for a total of 9 basis functions per cell.
• The TDG method with the basis functions (23)-(24) depending on the 5 directions (32), for a total of 5 basis functions per cell (plus one (22) in the black region).
• The TDG method with the basis functions (23)-(24) depending on the 5 directions (32) and the time dependent solutions (21), for a total of 8 basis functions per cell (plus one (22) in the black region).

Figure 8 shows that the DG method with only constant basis functions is too diffusive. However, if one increases the number of basis functions and considers affine basis functions, the DG method recovers a very good accuracy. From Figure 8, one also notices that the TDG method with 5 directions and only stationary basis functions seems too diffusive. Adding the time dependent basis functions (21) to the TDG method allow to recover a good accuracy similar to the affine DG method.

7.3.2 The P_3 model.

The comments are very similar for the P_3 model. Figure 9 represents the numerical results obtained for the P_3 model. The reference solution is computed with the DG method with affine basis functions for a total of 30 basis functions per cell on a 280×280 random mesh with $dt = 0.01$. We compare the DG and TDG methods on a 140×140 mesh with $dt = 0.02$. More precisely, we consider the following cases
• The DG method with constant basis functions only for a total of 10 basis functions per cell.
• The DG method with affine basis functions (that is 1, x, y) for a total of 30 basis functions per cell.
• The TDG method with the basis functions of Propositions 5.2 and 5.3 depending on the 3 directions (30), for a total of 12 basis functions per cell (plus one (22) in the black region).
• The TDG method with the basis functions of Propositions 5.2 and 5.3 depending on the 3 directions (30) and the time dependent solutions (21), for a total of 22 basis functions per cell (plus one (22) in the black region).

As for the P_1 model, Figure 9 illustrates that the DG method recovers a good accuracy when using affine basis functions. For the TDG method, considering only 3 stationary basis functions seems too diffusive. Nevertheless, if one adds the time dependent basis functions (21), the TDG method recovers a good accuracy similar to the affine DG method.

In particular, a benefit of the TDG method compared to the standard DG method is that it uses less basis functions to recover a good approximation of the numerical solution. However, as we will see in the next section, the TDG method may suffer from conditioning issues when considering stationary and time dependent basis functions on fine meshes.

Finally note that, both for the P_1 and P_3 model, the numerical results are similar to those obtained in [35, 28].

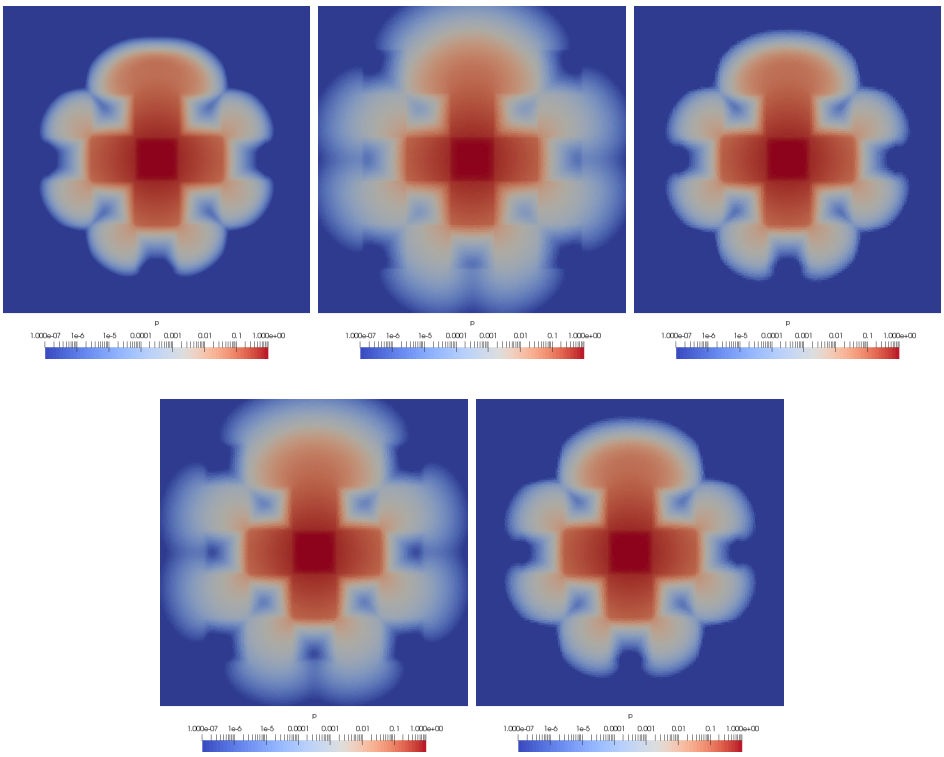

Figure 8: $\boldsymbol{P_1}$ **model**. Representation of the first variable for the test case 7.3, logarithmic scale.Top left: reference solution. Top center: DG scheme with 3 basis functions per cell. Top right: DG scheme with 9 basis functions per cell. Bottom left: TDG scheme with about 5 stationary basis functions per cell. Bottom right: TDG scheme with about 8 basis functions per cell (stationary and time dependent).

7.4 The condition number

It is well known that the Trefftz method can be very sensitive to ill-conditioning. That is why we provide some numerical evidence of this fact.

In Figure 10, we compare an estimation of the condition number of the matrices for the test problem described in Appendix 7.5 for the cases 1 to 4 on random meshes. The estimation is provided using the AztecOO package of the Trilinos library [46]. The Figure illustrates that the conditioning of the mass matrix can deteriorate dramatically depending of the basis functions used in a given calculation. In this case, the temporal exponentials (Case 2) give the best result in term of the condition number. More research is needed to determine if this is a general rule.

In Figure 11, we display the condition number for a stationary P_1 problem depending on a stiff parameter ε. The number of cells is constant. On the left part of the Figure, the condition number goes to 0 when $\varepsilon \to 0$, at the same rate for a DG scheme and for a TDG scheme. This is normal, since $\sigma_a = 0$ in this case, so the basis functions of the TDG are also polynomials. But on the right part of the Figure which corresponds to $\sigma_a > 0$, the condition number of TDG dramatically increases for small ε. Our explanation is that the exponential functions contain a stiff dependence with respect to ε, so it is normal that the conditioning of TDG is more sensitive than that of DG for $\varepsilon \to 0$.

Finally, in Figure 12, we study the behaviour of the condition number as a function of mesh resolution. We compare the condition number of the original matrix with that obtained when the matrix is preconditioned with the usual diagonal scaling technique. One observes a dramatic improvement, since the condition number is even better than the one for the DG method without preconditioning.

7.5 Sensitivity of TDG to the choice of time dependent basis functions

In this section, we show that TDG method (applied to the P_1 model in this section) can be quite sensitive to the choice of the time dependent basis functions. The problem is that we found no theoretical means to determine in advance how the result depends on this choice. It is only by numerical tests that one can evaluate this influence. The results below are also a justification why the approach corresponding to equation (21 (with $\alpha = \sigma_a$ or $\alpha = \sigma_t$) is our favorite one to introduce

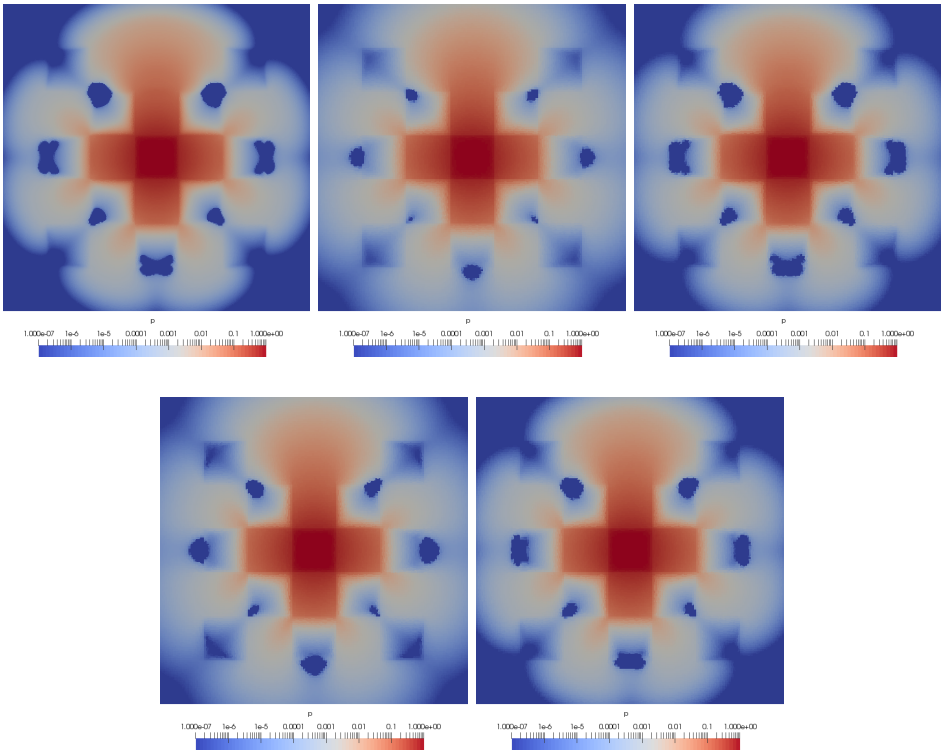

Figure 9: $\boldsymbol{P_3}$ **model**. Representation of the first variable for the test case 7.3. Top left: reference solution. Top center: DG scheme with 10 basis functions per cell. Top right: DG scheme with 30 basis functions per cell. Bottom left: TDG scheme with about 12 stationary basis functions per cell. Bottom right: TDG scheme with about 22 basis functions per cell (stationary and time dependent). Logarithmic scale.

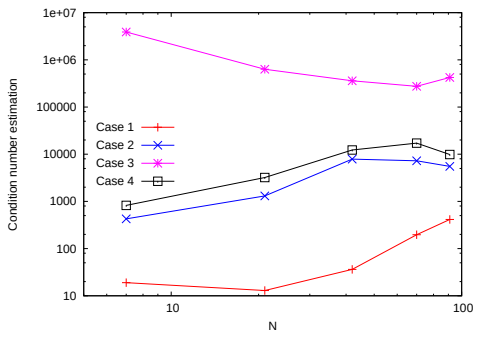

Figure 10: The condition number versus the one-dimensional cell number, for the cases 1 to 4 in the appendix 7.5. Logarithmic scale.

time dependent basis functions.

We consider the following cases

• **Case 1.** The stationary basis functions (23)-(24) only with the 3 directions (30) for a total of about 3 basis functions per cell.

• **Case 2.** The stationary basis functions (23)-(24) with the 3 directions (30) and the time dependent solutions (21) (with $\alpha = \sigma_a$ or $\alpha = \sigma_t$) for a total of about 6 basis functions per cell.

• **Case 3.** The stationary basis functions (23)-(24) and the time dependent solutions (35)-(36) with the 3 directions (30) for a total of about 9 basis functions per cell.

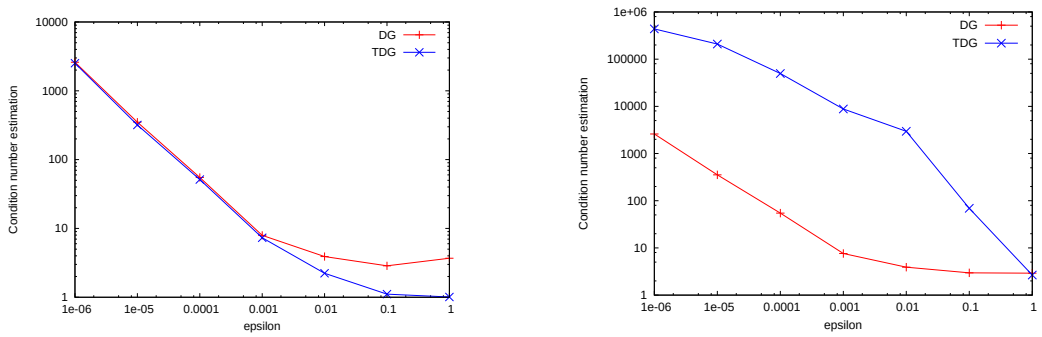

Figure 11: Condition number versus ε for the P_1 model. Logarithmic scale. On the left $\sigma_a = 0$. On the right $\sigma_a = 1$.

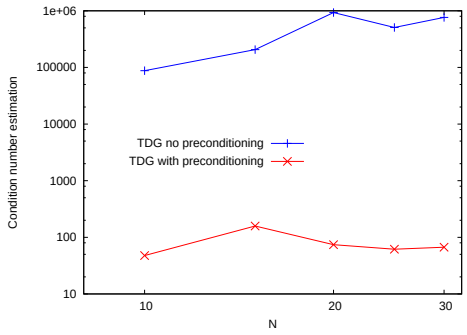

Figure 12: Test case 7 for the P_1 model: Preconditioning/conditioning versus the number of cells. Logarithmic scale.

• **Case 4.** The stationary basis functions (23)-(24) and the time-dependent solutions (34) with the 3 directions (30) for a total of about 6 basis functions per cell.
• **Case 5.** The stationary basis functions (23)-(24) and the time-dependent solutions (34) with the 4 directions (31) for a total of about 8 basis functions per cell.

We give some special time dependent solutions to the P_1 model. In this section, the solutions that we consider are product of time dependent polynomials and stationary exponentials. The proof is by direct calculus, see [19].

Lemma 7.1. *The value $\alpha = \sigma_t$ in the basis functions (21) gives*

$$\mathbf{v}(t,\mathbf{x}) = \begin{pmatrix} \sqrt{\sigma_t(1+\varepsilon)} \\ -\sqrt{\varepsilon(\sigma_a+\sigma_t)}P\mathbf{d} \end{pmatrix} e^{\frac{1}{c}\sqrt{3\varepsilon(\sigma_a+\sigma_t)\sigma_t(1+\varepsilon)}\mathbf{d}^T\mathbf{x}+\sigma_t t}, \tag{34}$$

with $\mathbf{d} = (\cos\theta, \sin\theta)^T \in \mathbb{R}^2$ and P the permutation matrix $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

Lemma 7.2 (Time dependent solutions when $\sigma_a > 0$)**.** *The following functions are solutions to the two dimensional P_1 model ($\mathbf{d}_k = (\cos\theta_k, \sin\theta_k)^T$)*

$$\mathbf{w}_{1,k}(t,\mathbf{x}) = \begin{pmatrix} -2c\varepsilon\sqrt{\sigma_a\sigma_t}\cos\theta_k - \sqrt{3\varepsilon}\sigma_t(\varepsilon\sigma_a+\sigma_t)x - 2c\sqrt{\sigma_a\sigma_t}\sigma_t\cos\theta_k t \\ \varepsilon\sqrt{3\sigma_a\sigma_t}(\varepsilon\sigma_a+\sigma_t)\sin\theta_k x + 2c\sqrt{\varepsilon}\sigma_a\sigma_t\cos\theta_k\sin\theta_k t \\ c\sqrt{\varepsilon}(\varepsilon\sigma_a+\sigma_t) + \varepsilon\sqrt{3\sigma_a\sigma_t}(\varepsilon\sigma_a+\sigma_t)\cos\theta_k x + 2c\sqrt{\varepsilon}\sigma_a\sigma_t\cos^2\theta_k t \end{pmatrix} e^{\frac{1}{c}\sqrt{3\varepsilon\sigma_a\sigma_t}\mathbf{d}_k^T\mathbf{x}}, \tag{35}$$

$$\mathbf{w}_{2,k}(t,\mathbf{x}) = \begin{pmatrix} -2c\varepsilon\sqrt{\sigma_a\sigma_t}\sin\theta_k - \sqrt{3\varepsilon}\sigma_t(\varepsilon\sigma_a+\sigma_t)y - 2c\sqrt{\sigma_a\sigma_t}\sigma_t\sin\theta_k t \\ c\sqrt{\varepsilon}(\varepsilon\sigma_a+\sigma_t) + \varepsilon\sqrt{3\sigma_a\sigma_t}(\varepsilon\sigma_a+\sigma_t)\sin\theta_k y + 2c\sqrt{\varepsilon}\sigma_a\sigma_t\sin^2\theta_k t \\ \varepsilon\sqrt{3\sigma_a\sigma_t}(\varepsilon\sigma_a+\sigma_t)\cos\theta_k y + 2c\sqrt{\varepsilon}\sigma_a\sigma_t\cos\theta_k\sin\theta_k t \end{pmatrix} e^{\frac{1}{c}\sqrt{3\varepsilon\sigma_a\sigma_t}\mathbf{d}_k^T\mathbf{x}}.$$

Lemma 7.3 (Time dependent polynomial solutions when $\sigma_a = 0$)**.** *The following functions are solutions to the two dimensional P_1 model when $\sigma_a = 0$ ($q_k(\mathbf{x})$ is a harmonic polynomial)*

$$\mathbf{v}_{1,k}(\mathbf{x}) = \begin{pmatrix} -\frac{2}{\sqrt{3}}\sqrt{\varepsilon}c^2\partial_x - \sqrt{3\varepsilon}\sigma_t^2 x - \frac{2}{\sqrt{3\varepsilon}}c^2\sigma_t t\partial_x \\ \sqrt{\varepsilon}c\sigma_t x\partial_y + \frac{2}{3\sqrt{\varepsilon}}c^3 t\partial_{xy} \\ \sqrt{\varepsilon}c\sigma_t + \sqrt{\varepsilon}c\sigma_t x\partial_x + \frac{2}{3\sqrt{\varepsilon}}c^3 t\partial_x^2 \end{pmatrix} q_k(\mathbf{x}),\ \mathbf{v}_{2,k}(\mathbf{x}) = \begin{pmatrix} -\frac{2}{\sqrt{3}}\sqrt{\varepsilon}c^2\partial_y - \sqrt{3\varepsilon}\sigma_t^2 y - \frac{2}{\sqrt{3\varepsilon}}c^2\sigma_t t\partial_y \\ \sqrt{\varepsilon}c\sigma_t + \sqrt{\varepsilon}c\sigma_t y\partial_y + \frac{2}{3\sqrt{\varepsilon}}c^3 t\partial_y^2 \\ \sqrt{\varepsilon}c\sigma_t y\partial_x + \frac{2}{3\sqrt{\varepsilon}}c^3 t\partial_{xy} \end{pmatrix} q_k(\mathbf{x}). \quad (36)$$

The results are displayed in Figure 13, where the setup is the lattice problem. We plot the total density of particles (the first component of U) in logscale. This is a severe test because the actual magnitude at the propagation front can be quite small. The random mesh is made of 70×70 cells.

The results with only stationary basis functions is the most diffuse one. One sees that all the time dependent basis functions reduce the diffusion. Compared to Case 2, one notices that the diffusion is lower for cases 3 to 5 but some weird oscillations appear. For the basis functions (34) (Cases 4 and 5), the choice of directions seems important. Indeed, with only the 3 directions (30) (Case 4), the numerical solution is highly asymmetric. Considering the 4 directions (31) (Case 5), fixes this issue. Note that Case 3 also considers the 3 directions (30) without getting the asymmetric result of Case 4.

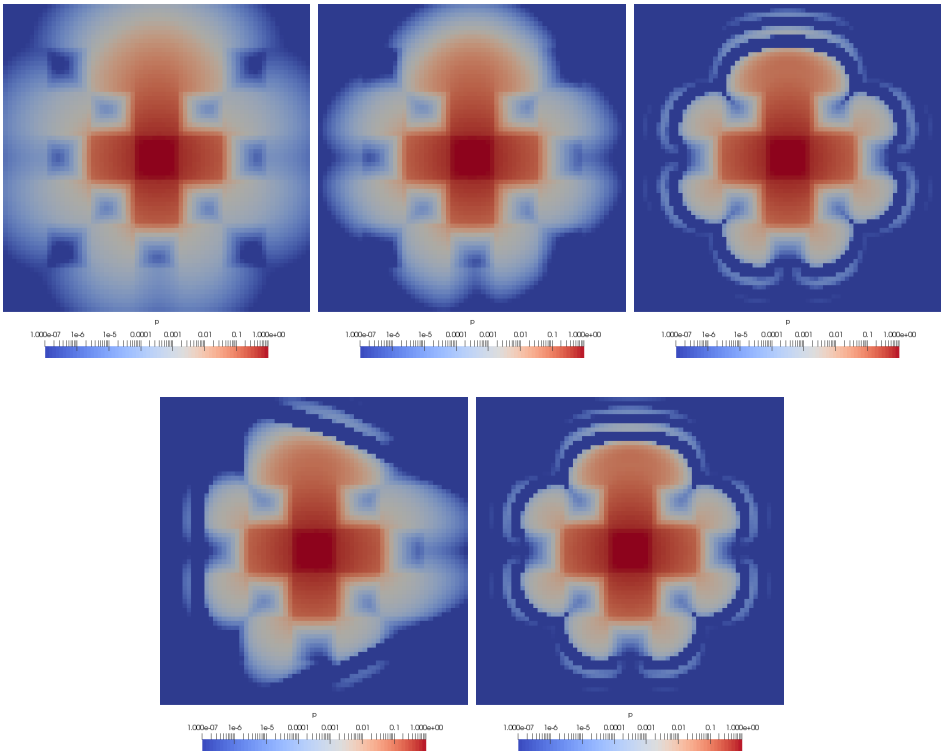

Figure 13: $\boldsymbol{P_1}$ **model**. Representation of the first variable for the test case 7.3. Cases 1 to 5. The cases are numbered from from left to right and top to bottom (top left: Case 1, top center: Case 2...). Logarithmic scale.

8 Conclusions

In this work, the Trefftz discontinuous Galerkin (TDG) method applied to transport models has been studied and analyzed in the general case of the two dimensional P_N model. After recalling the derivation of the P_N model, some of its properties were given. Numerical results for the two dimensional P_1 and P_3 models were provided. It has been shown that the TDG method outperforms the standard DG method for some numerical tests with boundary layers, using less degrees of freedom for a better accuracy. The main drawback of the TDG method is that it may lead to ill-conditioned system when considering too many basis functions per cell or in some asymptotic regimes. The formulation of the TDG method can be easily generalized to the three dimensional case. In 3D for the P_N model, the basis functions can be constructed as in Section 4, starting from a one dimensional solution and then applying a rotation. Note however that the three dimensional rotation is not as simple as in the two dimensional case [24, 29, 26]. Another perspective is to develop good preconditioners

to deal with the ill-conditioning of the linear systems arising from the TDG method. This could be particularly useful when considering, for example, stationary and time dependent basis functions. It could also be interesting to extend the TDG method to the discrete ordinate method (S_N model) which is the other popular approximation of the transport equation. Since the S_N model is naturally written under the form of a Friedrichs system, the general formulation given in Section B can be used. It remains to construct the basis functions.

Appendices

Appendix A Spherical harmonics

We recall some definitions and properties of the spherical harmonics and adopt the presentation given in [2].

A.1 Legendre functions

The spherical harmonics are based on the Legendre functions P_k^l which read

$$P_k^l(\mu) = \frac{1}{2^k k!}(1-\mu^2)^{l/2}\frac{d^{k+l}}{d\mu^{k+l}}((\mu^2-1)^k) \text{ for } l \geq 0 \quad \text{and} \quad P_k^l(\mu) = (-1)^l \frac{(k+l)!}{(k-l)!} P_k^{-l}(\mu) \text{ for } l < 0. \tag{37}$$

The Legendre polynomials satisfy the orthogonality relations $\frac{1}{2}\int_{-1}^{1} P_k^0 d\mu = \delta_k^0$, $\frac{1}{2}\int_{-1}^{1} P_k^l P_m^l d\mu = \frac{1}{(a_k^l)^2}\delta_k^m$, where a_k^l is the normalization factor $a_k^l = \sqrt{(2k+1)\frac{(k-l)!}{(k+l)!}}$. They also satisfy the following recursion relations which are fundamental to deriving the P_N model

$$\begin{cases} \sqrt{1-\mu^2}P_k^m = \frac{1}{2k+1}\Big(P_{k+1}^{m+1} - P_{k-1}^{m+1}\Big), \\ \sqrt{1-\mu^2}P_k^m = \frac{1}{2k+1}\Big(-(k-m+1)(k-m+2)P_{k+1}^{m-1} + (k+m-1)(k+m)P_{k-1}^{m-1}\Big), \\ \mu P_k^m = \frac{1}{2k+1}\Big((k-m+1)P_{k+1}^m + (k+m)P_{k-1}^m\Big). \end{cases}$$

A.2 Spherical harmonics

The complex valued spherical harmonics read $Y_k^l(\psi,\phi) := Y_k^l(\mathbf{\Omega}) := (-1)^l a_k^l P_k^l(\cos\phi)e^{il\psi}$ for $|l| \leq k$. The real valued spherical harmonics $Y_{k,l}$ are

$$. \begin{cases} Y_{k,l}(\mathbf{\Omega}) = Y_k^l(\mathbf{\Omega}) = a_k^l P_k^l(\cos\phi), & l = 0, \\ Y_{k,l}(\mathbf{\Omega}) = \dfrac{(-1)^l}{\sqrt{2}}\big(Y_k^l(\mathbf{\Omega}) + \bar{Y}_k^l(\mathbf{\Omega})\big) = a_k^l\sqrt{2}\cos(l\psi)P_k^l(\cos\phi), & 0 < l \leq k, \\ Y_{k,l}(\mathbf{\Omega}) = \dfrac{i}{\sqrt{2}}\big(Y_k^l(\mathbf{\Omega}) - \bar{Y}_k^l(\mathbf{\Omega})\big) = a_k^{|l|}\sqrt{2}\sin(|l|\psi)P_k^{|l|}(\cos\phi), & -k \leq l < 0. \end{cases} \tag{38}$$

They satisfy the relations $\frac{1}{4\pi}\int_{S^2} Y_{k,l}d\psi d\mu = \delta_k^0\delta_l^0$, $\frac{1}{4\pi}\int_{S^2} Y_{k,l}Y_{m,n}d\psi d\mu = \delta_k^m\delta_l^n$ and the recursion relations

$$\begin{cases} \sin\phi\cos\psi Y_{k,m} &= \varepsilon^m(A_k^m Y_{k+1,m+1} - B_k^m Y_{k-1,m+1}) - \zeta^m(C_k^m Y_{k+1,m-1} - D_k^m Y_{k-1,m-1}), \\ \sin\phi\sin\psi Y_{k,m} &= \eta^m(A_k^m Y_{k+1,-m-1} - B_k^m Y_{k-1,-m-1}) + \phi^m(C_k^m Y_{k+1,-m+1} - D_k^m Y_{k-1,-m+1}), \\ \cos\phi Y_{k,m} &= E_k^m Y_{k+1,m} + F_{k,m}Y_{k-1,m}, \end{cases} \tag{39}$$

where coefficients are given in [2].

Appendix B Numerical analysis of the TDG method

Because of different approximation spaces, TDG methods are not completely standard ones with respect to traditional DG methods. It is therefore valuable to review the basic results [6, 34, 5, 18, 19] which are at the core of the h-convergence analysis of TDG methods. We make the distinction between what we call the classical theory which deals with the case $R = 0$ and the strictly dissipative case $R > 0$.

B.1 Well posedness and quasi-optimality

In this section we show well posedness of (15) and a quasi-optimality bound in mesh-dependent norms. Our analysis follows results of [5] where the special case with $R = 0$ was studied and [18, 19] adapted to the situation where $R > 0$ provides additional L^2 control in the cells. One defines two semi-norms on $H^1(\mathcal{T}_h)$

$$\begin{aligned}\|\mathbf{u}\|_{DG}^2 &= \sum_k \int_{\Omega_k} \mathbf{u}_k^T R \mathbf{u}_k + \sum_k \sum_{j<k} \frac{1}{2} \int_{\Sigma_{kj}} (\mathbf{u}_k - \mathbf{u}_j)^T |M_{kj}| (\mathbf{u}_k - \mathbf{u}_j) + \sum_k \frac{1}{2} \int_{\Sigma_{kk}} \mathbf{u}_k^T |M_k| \mathbf{u}_k, \\ \|\mathbf{u}\|_{DG^*}^2 &= \sum_k \int_{\partial\Omega_k} -\mathbf{u}_k^T M_k^- \mathbf{u}_k,\end{aligned} \tag{40}$$

with $|M_{kj}| = |M_{jk}| = M_{kj}^+ - M_{kj}^-$. First steps are to show that these two semi-norms are in fact norms on the Trefftz space. Further details on the proofs are given in [6, 34, 5, 18, 19].

Lemma B.1. *One has the inequality $\|\mathbf{v}\|_{DG} \leq c\|\mathbf{v}\|_{DG^*}$ for all $\mathbf{v} \in V(\mathcal{T}_h)$, with $c = \sqrt{5/2}$.*

Lemma B.2. *Assume $M \in \mathbb{R}^{n\times n}$ is a symmetric matrix. Then one has $\mathbf{z}^T M^2 \mathbf{z} \leq C\mathbf{z}^T |M| \mathbf{z}$ for all $\mathbf{z} \in \mathbb{R}^n$, where $M = M^+ + M^-$, M^+ is a non negative matrix, M^- is a non positive matrix and $|M| = M^+ - M^-$.*

Proposition B.3. *The semi-norms $\|\cdot\|_{DG}$ and $\|\cdot\|_{DG^*}$ are norms on the Trefftz space $V(\mathcal{T}_h)$.*

Next, we study the coercivity and the continuity of the bilinear form $a(\cdot,\cdot)$ regarding the norms $\|\cdot\|_{DG}$ and $\|\cdot\|_{DG^*}$.

Proposition B.4 (Coercivity)**.** *For $\mathbf{u} \in H^1(\mathcal{T}_h)$ then $a_{DG}(\mathbf{u},\mathbf{u}) = \|\mathbf{u}\|_{DG}^2$. For $\mathbf{u} \in V(\mathcal{T}_h)$ then $a_{DG}(\mathbf{u},\mathbf{u}) = a_T(\mathbf{u},\mathbf{u})$.*

Proposition B.5 (Continuity)**.** *The continuity bound $a_T(\mathbf{u},\mathbf{v}) \leq \sqrt{2}\|\mathbf{u}\|_{DG}\|\mathbf{v}\|_{DG^*}$ holds for all $\mathbf{u},\mathbf{v} \in V(\mathcal{T}_h)$.*

The classical quasi-optimality result is the following.

Proposition B.6 (Quasi-optimality)**.** *For any finite dimensional space $V_m(\mathcal{T}_h) \subset V(\mathcal{T}_h)$, the TDG formulation (15) admits a unique solution $\mathbf{u}_h \in V_m(\mathcal{T}_h)$. The following quasi-optimality bounds holds $\|\mathbf{u} - \mathbf{u}_h\|_{DG} \leq \sqrt{2} \inf_{\mathbf{v}_h \in V_m(\mathcal{T}_h)} \|\mathbf{u} - \mathbf{v}_h\|_{DG^*}$, where $\mathbf{u}$ stands for the exact solution to (4).*

Using the quasi-optimality proposition one has the well-balanced property of the scheme. We refer to [47, 48, 49, 50] for general references for well-balanced methods for similar problems. In one dimension a scheme is well-balanced if it captures all the stationary states of a hyperbolic system. This is possible because, in one dimension, the number of linearly independent stationary solutions is finite. However in two dimensions the space of stationary solutions becomes infinite. It has a huge impact on what is a well-balanced scheme in space dimensions higher than one. One must choose a finite subset of solutions for which the scheme is supposed to be exact. This is a practical definition of a well-balanced scheme and it is immediately deduced from the quasi-optimality result of proposition B.6. Of course a standard DG scheme has the same quasi-optimality result, but it can be well-balanced only for some particular polynomial functions. On the contrary a TDG method can be well-balanced for more general solutions which contain for example exponential factors as in Example 1 in Section 3.2 for which $\sigma_a > 0$.

Proposition B.7 (Well-balanced scheme)**.** *If the solution $\mathbf{u} \in H^1(\Omega)$ of (4) is locally (in each cell) a linear combination of the basis functions (which are by construction exact solutions), then $\mathbf{u}_h = \mathbf{u}$.*

B.2 Estimate in standard norms

In the previous section, the error is bounded in terms of DG-norm. It is of course desirable to have estimates in a more standard norm. In this section we present some elementary L^2 lower bounds of the DG norm which take advantage of the relaxation matrix R and an L^2 upper bound of the DG^* norm. Proofs are in [18, 19].

Proposition B.8. *Assume all $R_k = R(\mathbf{x})_{|\Omega_k} > 0$ are positive. Then $\frac{1}{\sup_{k\in\mathcal{T}_h} \|\sqrt{R_k}^{-1}\|^2} \|\mathbf{w}\|_{L^2(\Omega)} \leq \|\mathbf{w}\|_{DG}$, for all $\mathbf{w} \in H^1(\mathcal{T}_h)$.*

This inequality holds when R is definite positive but degenerates when $R \to 0$. For non stationary problems, one can give an L^2 lower bound at the final time that does not depend on R.

Proposition B.9. *For time dependent problems one has $\|\mathbf{w}\|_{L^2(\Omega_S\times\{T\})} \leq \|\mathbf{w}\|_{DG}$ for all $\mathbf{w} \in H^1(\mathcal{T}_h)$.*

Define the semi-norm $|\mathbf{w}|_{1,\Omega}^2 := \int_\Omega \sum_{i=1}^n \sum_{j=1}^d (\partial_j \mathbf{w}_i)^2$.

Proposition B.10. *One has $\|\mathbf{w}\|_{DG^*}^2 \leq C \sum_j \|\mathbf{w}\|_{L^2(\Omega_j)} \left(\frac{1}{h_j} \|\mathbf{w}\|_{L^2(\Omega_j)} + |\mathbf{w}|_{1,\Omega_j} \right)$ for all $\mathbf{w} \in H^1(\mathcal{T}_h)$, where $h_j = diam(\Omega_j)$ and where the constant C depends on the A_i.*

If the matrices A_i are rescaled with respect to some small parameter, then the constant C must be rescaled as well.

References

[1] S. Chandrasekhar, Radiative transfer, (International Series of Monographs on Physics) Oxford: Clarendon Press; London: Oxford University Press. XIV, 394 p. (1950).

[2] F. Hermeline, A discretization of the multigroup P_N radiative transfer equation on general meshes, J. Comput. Phys. 313 (2016) 549–582.

[3] E. A. Maunder, Trefftz in translation., Comput. Assist. Mech. Eng. Sci. 10 (4) (2003) 545–563.

[4] R. Hiptmair, A. Moiola, I. Perugia, A Survey of Trefftz Methods for the Helmholtz Equation, in: Building Bridges: Connections and Challenges in Modern Approaches to Numerical Partial Differential Equations, Vol. 114, Springer, 2016, pp. 237–278.

[5] F. Kretzschmar, A. Moiola, I. Perugia, S. M. Schnepp, A priori error analysis of space–time Trefftz discontinuous Galerkin methods for wave problems , IMA Journal of Numerical Analysis 36 (4) (2016) 1599.

[6] P. Monk, G. R. Richter, A discontinuous Galerkin method for linear symmetric hyperbolic systems in inhomogeneous media, J. Sci. Comput. 22-23 (2005) 443–477.

[7] J. Gopalakrishnan, P. Monk, P. Sepúlveda, A tent pitching scheme motivated by Friedrichs theory, Comput. Math. Appl. 70 (5) (2015) 1114–1135. doi:10.1016/j.camwa.2015.07.001.
URL https://doi.org/10.1016/j.camwa.2015.07.001

[8] G. Fu, C.-W. Shu, Optimal energy-conserving discontinuous Galerkin methods for linear symmetric hyperbolic systems, J. Comput. Phys. 394 (2019) 329–363. doi:10.1016/j.jcp.2019.05.050.
URL https://doi.org/10.1016/j.jcp.2019.05.050

[9] A. Moiola, I. Perugia, A space-time Trefftz discontinuous Galerkin method for the acoustic wave equation in first-order formulation, Numer. Math. 138 (2) (2018) 389–435. doi:10.1007/s00211-017-0910-x.
URL https://doi.org/10.1007/s00211-017-0910-x

[10] T. Zhang, J. Liu, A space-time discontinuous Galerkin method for first order hyperbolic systems, J. Korean Math. Soc. 51 (4) (2014) 665–678. doi:10.4134/JKMS.2014.51.4.665.
URL https://doi.org/10.4134/JKMS.2014.51.4.665

[11] G. Gabard, Exact integration of polynomial-exponential products with application to wave-based numerical methods, Comm. Numer. Methods Engrg. 25 (3) (2009) 237–246. doi:10.1002/cnm.1123.
URL https://doi.org/10.1002/cnm.1123

[12] O. Cessenat, B. Després, Application of an Ultra Weak Variational Formulation of Elliptic PDEs to the Two-Dimensional Helmholtz Problem, SIAM J. Numer. Anal. 35 (1) (1998) 255–299.

[13] T. Huttunen, P. Monk, J. P. Kaipio, Computational aspects of the ultra-weak variational formulation., J. Comput. Phys. 182 (1) (2002) 27–46.

[14] C. J. Gittelson, R. Hiptmair, I. Perugia, Plane wave discontinuous Galerkin methods: Analysis of the h-version, ESAIM, Math. Model. Numer. Anal. 43 (2) (2009) 297–331.

[15] L.-M. Imbert-Gérard, Well-posedness and generalized plane waves simulations of a 2D mode conversion model, J. Comput. Phys. 303 (2015) 105–124. doi:10.1016/j.jcp.2015.09.033.
URL http://dx.doi.org/10.1016/j.jcp.2015.09.033

[16] L.-M. Imbert-Gérard, Interpolation properties of generalized plane waves, Numer. Math. 131 (4) (2015) 683–711. doi:10.1007/s00211-015-0704-y.
URL http://dx.doi.org/10.1007/s00211-015-0704-y

[17] L.-M. Imbert-Gérard, B. Després, A generalized plane-wave numerical method for smooth nonconstant coefficients, IMA J. Numer. Anal. 34 (3) (2014) 1072–1103. doi:10.1093/imanum/drt030.
URL http://dx.doi.org/10.1093/imanum/drt030

[18] G. Morel, C. Buet, B. Després, Trefftz discontinuous Galerkin method for Friedrichs systems with linear relaxation: application to the P_1 model, Computational Methods in Applied Mathematics (2018) 521–557.

[19] G. Morel, Asymptotic-preserving and well-balanced schemes for transport models using Trefftz discontinuous Galerkin method, Theses, Sorbonne Université (Sep. 2018).
URL https://hal.archives-ouvertes.fr/tel-01911872

[20] C. Buet, B. Després, G. Morel, Trefftz discontinuous Galerkin basis functions for a class of Friedrichs systems coming from linear transport: the P_N model, Advances in Computational Mathematics 46 (41) (2020) 1–27.

[21] G. Gabard, Discontinuous Galerkin methods with plane waves for the displacement-based acoustic equation, International Journal for Numerical Methods in Engineering 66 (2006) 549–569.

[22] G. Gabard, Discontinuous Galerkin methods with plane waves for time-harmonic problems, Journal of Computational Physics 225 (2) (2007) 1961–1984.

[23] C. Garrett, C. D. Hauck, On the eigenstructure of spherical harmonic equations for radiative transport., Comput. Math. Appl. 72 (2) (2016) 264–270. doi:10.1016/j.camwa.2015.05.030.

[24] A. M. Blanco, M. Florez, M. Bermejo, Evaluation of the rotation matrices in the basis of real spherical harmonics, Journal of Molecular Structure-theochem 419 (1997) 19–27.

[25] F. Dai, Y. Xu, Approximation theory and harmonic analysis on spheres and balls, New York, NY: Springer, 2013. doi:10.1007/978-1-4614-6660-4.

[26] D. Pinchon, P. E. Hoggan, Rotation matrices for real spherical harmonics: general rotations of atomic orbitals in space-fixed axes., J. Phys. A, Math. Theor. 40 (7) (2007) 1597–1610. doi:10.1088/1751-8113/40/7/011.

[27] T. A. Brunner, J. P. Holloway, One-dimensional Riemann solvers and the maximum entropy closure, Journal of Quantitative Spectroscopy and Radiative Transfer 69 (5) (2005) 386–399. doi:https://doi.org/10.1016/S0022-4073(00)00099-6.
URL http://www.sciencedirect.com/science/article/pii/S0022407300000996

[28] C. Buet, B. Després, E. Franck, Asymptotic preserving schemes on distorted meshes for Friedrichs systems with stiff relaxation: application to angular models in linear transport, J. Sci. Comput. 62 (2) (2015) 371–398.

[29] J. Ivanic, K. Ruedenberg, Rotation matrices for real spherical harmonics. Direct determination by recursion, Journal of Physical Chemistry 100 (15) (Apr 1996). doi:10.1021/jp953350u.

[30] K. O. Friedrichs, Symmetric positive linear differential equations, Communications on Pure and Applied Mathematics 11 (3) (1958) 333–418.

[31] C. Berthon, R. Turpault, Asymptotic preserving HLL schemes, Numer. Methods Partial Differ. Equations 27 (6) (2011) 1396–1422. doi:10.1002/num.20586.

[32] L. Gosse, Computing qualitatively correct approximations of balance laws. Exponential-fit, well-balanced and asymptotic-preserving, Milano: Springer, 2013. doi:10.1007/978-88-470-2892-0.

[33] S. Jin, C. Levermore, Numerical schemes for hyperbolic conservation laws with stiff relaxation terms, J. Comput. Phys. 126 (2) (1996) 449–467, art. no. 0149.

[34] A. Ern, J. Guermond, Discontinuous galerkin methods for Friedrichs' systems. i. general theory, SIAM J. Numerical Analysis 44 (2) (2006) 753–778. doi:10.1137/050624133.
URL https://doi.org/10.1137/050624133

[35] T. A. Brunner, Form of approximate radiation transport., Sandia report SAND–2002-1778 (2002).

[36] S. Jin, D. Levermore, The discrete-ordinate method in diffusive regimes, Transport Theory and Statistical Physics 20 (1991) 413–439.

[37] Q. Li, J. Lu, W. Sun, Diffusion approximations and domain decomposition method of linear transport equations: asymptotics and numerics., J. Comput. Phys. 292 (2015) 141–167. doi:10.1016/j.jcp.2015.03.014.

[38] K. Case, Elementary solutions of the transport equation and their applications, Ann. Phys. 9 (1960) 1–23. doi:10.1016/0003-4916(60)90060-9.

[39] G. Birkhoff, I. Abu-Shumays, Harmonic solutions of transport equations, J. Math. Anal. Appl. 28 (1969) 211–221. doi:10.1016/0022-247X(69)90123-1.

[40] G. Birkhoff, I. Abu-Shumays, Exact analytic solutions of transport equations, J. Math. Anal. Appl. 32 (1970) 468–481. doi:10.1016/0022-247X(70)90271-4.

[41] A. Moiola, R. Hiptmair, I. Perugia, Vekua theory for the Helmholtz operator, Z. Angew. Math. Phys. 62 (5) (2011) 779–807.

[42] R. Hiptmair, A. Moiola, I. Perugia, Plane wave discontinuous Galerkin methods for the 2D Helmholtz equation: analysis of the p-version, SIAM J. Numer. Anal. 49 (1) (2011) 264–284.

[43] Y. Abdelaziz, A. Hamouine, A survey of the extended finite element, Computers & Structures 86 (11) (2008) 1141 – 1151. doi:https://doi.org/10.1016/j.compstruc.2007.11.001.
URL http://www.sciencedirect.com/science/article/pii/S0045794907002957

[44] T.-P. Fries, T. Belytschko, The extended/generalized finite element method: An overview of the method and its applications, International Journal for Numerical Methods in Engineering 84 (3) (2010) 253–304. doi:10.1002/nme.2914.
URL http://dx.doi.org/10.1002/nme.2914

[45] M. Schäfer, M. Frank, C. Levermore, Diffusive corrections to P_N approximations, Multiscale Model. Simul. 9 (1) (2011) 1–28. doi:10.1137/090764542.

[46] M. Heroux, R. Bartlett, V. H. R. Hoekstra, J. Hu, T. Kolda, R. Lehoucq, K. Long, R. Pawlowski, E. Phipps, A. Salinger, H. Thornquist, R. Tuminaro, J. Willenbring, A. Williams, An Overview of Trilinos, Tech. Rep. SAND2003-2927, Sandia National Laboratories (2003).

[47] B. Després, C. Buet, The structure of well-balanced schemes for Friedrichs systems with linear relaxation, Applied Mathematics and Computation 272 (2016) 440–459.

[48] L. Gosse, G. Toscani, An asymptotic-preserving well-balanced scheme for the hyperbolic heat equations, C. R., Math., Acad. Sci. Paris 334 (4) (2002) 337–342.

[49] J. M. Greenberg, A. Y. Leroux, A well-balanced scheme for the numerical processing of source terms in hyperbolic equations, SIAM Journal on Numerical Analysis 33 (1) (1996) 1–16. arXiv:https://doi.org/10.1137/0733001, doi:10.1137/0733001.
URL https://doi.org/10.1137/0733001

[50] S. Jin, A steady-state capturing method for hyperbolic systems with geometrical source terms, in: N. B. Abdallah, I. M. Gamba, C. Ringhofer, A. Arnold, R. T. Glassey, P. Degond, C. D. Levermore (Eds.), Transport in Transition Regimes, Springer New York, New York, NY, 2004, pp. 177–183.