

DUALITY-BASED ASYMPTOTIC-PRESERVING METHOD FOR HIGHLY ANISOTROPIC DIFFUSION EQUATIONS*

PIERRE DEGOND[†], FABRICE DELUZET[‡], ALEXEI LOZINSKI[§],
JACEK NARSKI[¶], AND CLAUDIA NEGULESCU^{||}

To Dave Levermore on his 60th birthday with friendship and appreciation

Abstract. The present paper introduces an efficient and accurate numerical scheme for the solution of a highly anisotropic elliptic equation, the anisotropy direction being given by a variable vector field. This scheme is based on an asymptotic preserving reformulation of the original system, permitting an accurate resolution independently of the anisotropy strength and without the need of a mesh adapted to this anisotropy. The counterpart of this original procedure is the larger system size, enlarged by adding auxiliary variables and Lagrange multipliers. This Asymptotic-Preserving method generalizes the method investigated in a previous paper [P. Degond, F. Deluzet, and C. Negulescu, *Multiscale Model. Simul.*, 8(2), 645–666, 2009/10] to the case of an arbitrary anisotropy direction field.

Key words. Anisotropic diffusion, asymptotic preserving scheme, finite element method.

AMS subject classifications. 65N30.

1. Introduction

Anisotropic problems are common in the mathematical modeling of physical problems. They occur in various fields of applications such as flows in porous media [3, 23], semiconductor modeling [35], quasi-neutral plasma simulations [11], image processing [44, 45], atmospheric or oceanic flows [43] and so on, the list being not exhaustive. The initial motivation for this work is closely related to magnetized plasma simulations such as atmospheric plasma [27, 29], internal fusion plasma [4, 13] or plasma thrusters [1]. In this context, the medium is structured by the magnetic field, which may be strong in some regions and weak in others. Indeed, the gyration of the charged particles around magnetic field lines dominates the motion in the plane perpendicular to magnetic field. This explains the large number of collisions in the perpendicular plane while the motion along the field lines is rather undisturbed. As a consequence the mobility of particles in different directions differs by many orders of magnitude; this ratio can be as huge as 10^{10} . On the other hand, when the magnetic field is weak the anisotropy is much smaller. As the regions with weak and strong magnetic field can coexist in the same computational domain, one needs a numerical scheme which gives accurate results for a large range of anisotropy strengths. The relevant boundary conditions in many fields of application are periodic (for instance in simulations

*Received: August 19, 2010; accepted (in revised version): January 3, 2011.

[†]Université de Toulouse, UPS, INSA, UT1, UTM, Institut de Mathématiques de Toulouse, F-31062 Toulouse, France. CNRS, Institut de Mathématiques de Toulouse UMR 5219, F-31062 Toulouse, France (pierre.degond@math.univ-toulouse.fr).

[‡]Université de Toulouse, UPS, INSA, UT1, UTM, Institut de Mathématiques de Toulouse, F-31062 Toulouse, France. CNRS, Institut de Mathématiques de Toulouse UMR 5219, F-31062 Toulouse, France (fabrice.deluzet@math.univ-toulouse.fr).

[§]Université de Toulouse, UPS, INSA, UT1, UTM, Institut de Mathématiques de Toulouse, F-31062 Toulouse, France (alexei.lozinski@math.univ-toulouse.fr).

[¶]Université de Toulouse, UPS, INSA, UT1, UTM, Institut de Mathématiques de Toulouse, F-31062 Toulouse, France (jacek.narski@math.univ-toulouse.fr).

^{||}CMI/LATP, Université de Provence, 39 rue Frédéric Joliot-Curie 13453 Marseille cedex 13, France (claudia.negulescu@cmi.univ-mrs.fr).

of the tokamak plasmas on a torus) or Neumann boundary conditions (atmospheric plasma; see for example [5]). For these reasons we propose a strongly anisotropic model problem for which we wish to introduce an efficient and accurate numerical scheme. This model problem reads

$$\begin{cases} -\nabla \cdot \mathbb{A} \nabla \phi^\varepsilon = f & \text{in } \Omega, \\ n \cdot \mathbb{A} \nabla \phi^\varepsilon = 0 & \text{on } \partial\Omega_N, \\ \phi^\varepsilon = 0 & \text{on } \partial\Omega_D, \end{cases} \quad (1.1)$$

where $\Omega \subset \mathbb{R}^2$ or $\Omega \subset \mathbb{R}^3$ is a bounded domain with boundary $\partial\Omega = \partial\Omega_D \cup \partial\Omega_N$ and outward normal n . The direction of the anisotropy is defined by a vector field B , where we suppose $\operatorname{div} B = 0$ and $B \neq 0$. The direction of B is given by a vector field $b = B/|B|$. The anisotropy matrix is then defined as

$$\mathbb{A} = \frac{1}{\varepsilon} A_{\parallel} b \otimes b + (Id - b \otimes b) A_{\perp} (Id - b \otimes b), \quad (1.2)$$

and $\partial\Omega_D = \{x \in \partial\Omega \mid b(x) \cdot n = 0\}$. The scalar field $A_{\parallel} > 0$ and the symmetric positive definite matrix field A_{\perp} are of order one while the parameter $0 < \varepsilon < 1$ can be very small, thus provoking the high anisotropy of the problem. This work extends the results of [12], where the special case of a vector field b , aligned with the z -axis, was studied. An extension of this approach is proposed in [6] to handle more realistic anisotropy topologies. It relies on the introduction of a curvilinear coordinate system with one coordinate aligned with the anisotropy direction. Adapted coordinates are widely used in the framework of plasma simulation (see for instance [4, 15, 37]), these systems being either developed to fit particular magnetic field geometries or plasma equilibria (Euler potentials [42], toroidal and poloidal [21, 26], quasiballooning [16], Hamada [22] and Boozer [7] coordinates). Note that the study of certain plasma regions in a tokamak have motivated the use of non-orthogonal coordinates systems [24]. In contrast with all these methods, we propose here a numerical scheme that uses coordinates and meshes independent of the anisotropy direction, like in [39]. This feature offers the capability to easily treat time evolving anisotropy directions. This is very important in the context of tokamak plasma simulation, where the anisotropy is driven by a time dependent magnetic field.

One of the difficulties associated with the numerical solution of problem (1.1) lies in the fact that this problem becomes very ill-conditioned for small $0 < \varepsilon \ll 1$. Indeed, replacing ε by zero yields an ill-posed problem as it has an infinite number of solutions (any function constant along the b field solves the problem with $\varepsilon = 0$). In the discrete case the problem translates into a linear system which is ill-conditioned, as it mixes the terms of different orders of magnitude for $\varepsilon \ll 1$. As a consequence the numerical algorithm for solving this linear system gives unacceptable errors (in the case of direct solvers) or fails to converge in a reasonable time (in the case of iterative methods).

This difficulty arises when the boundary conditions supplied to the dominant $O(1/\varepsilon)$ operator lead to an ill-posed problem. This is the case for Neumann boundary conditions imposed on the part of the boundary with $b \cdot n \neq 0$ as well as for periodic boundary conditions. If instead, the boundary conditions are such that the dominant operator gives a well-posed problem, the numerical difficulty vanishes. One can resort to standard methods, as the dominant operator is sufficient to determine the limit solution. This is the case for Dirichlet and Robin boundary conditions. The problem addressed in this paper arises therefore only with specific boundary conditions. It has

however a considerable impact in numerous physical problems concerning plasmas, geophysical flows, plates and shells (for example). In this paper, we will focus on Neumann boundary condition since they represent a larger range of physical applications. The periodic boundary conditions can be addressed in a similar way.

Numerical methods for anisotropic problems have been extensively studied in the literature. Distinct methods have been developed. For example, finite volumes schemes with carefully designed approximations to the normal fluxes on the cell edges were proposed in [32, 17]. The mimetic finite difference discretization [33] is a closely related but sometimes more efficient approach. The symmetric and asymmetric difference schemes were investigated in [20, 41]. In the context of finite element methods, the use of slope limiters is advocated in [31]. Domain decomposition techniques using multiple coarse grid corrections are adapted to the anisotropic equations in [19, 30]. Multigrid methods have been studied in [18, 38]. For anisotropy aligned with one or two directions, point or plane smoothers are shown to be very efficient [34]. The *hp*-finite element method is also known to give good results for singular perturbation problems [36]. All of these methods try to discretize the anisotropic PDE as it is written and then apply purely numerical tricks to circumvent the problems related to lack of accuracy of the discrete solution or to the slow convergence of iterative algorithms. This leads to methods which are sometimes difficult to implement.

The approach that we pursue in this paper is entirely different: we reformulate first the original PDE in such a way that the resulting problem can be efficiently and accurately discretized by straight-forward and easily implementable numerical methods for any anisotropy strength. Our scheme is related to the *Asymptotic Preserving* method introduced in [25]. These techniques are designed to give a precise solution in the various regimes with no restrictions on the computational meshes and with the additional property of converging to the limit solution when $\varepsilon \rightarrow 0$. The derivation of the Asymptotic Preserving method requires identification of the limit model. In the case of Singular Perturbation problems, the original problem is reformulated in such a way that the obtained set of equations contain both the limit model and the original problem with a continuous transition between them, according to the values of ε . This reformulated system of equation sets the foundation of the AP-scheme. These Asymptotic Preserving techniques have been explored in previous studies, for instance quasi-neutral or gyro-fluid limits [10, 13]. A similar approach was also pursued in [9] where a reformulation of magnetohydrodynamics equations was used in conjunction with the Jacobian-free Newton-Krylov technology and multigrid methods.

In this paper, we present a new algorithm which extends the results of [12]. The originality of this algorithm consists in the fact that it is applicable for variable anisotropy directions b without additional work. The discretization mesh need not be adapted to the field direction b , but is simply a Cartesian grid whose mesh-size is governed by the desired accuracy, independently on the anisotropy strength ε . All this is possible in a well-adapted mathematical framework (optimally chosen spaces, introduction of Lagrange multipliers). The key idea, as in [12], is to decompose the solution ϕ into two parts: a mean part p which is constant along the field lines and the fluctuation part q consisting of a correction to the mean part needed to recover the full solution. Both parts p and q are solutions to well-posed problems for any $\varepsilon > 0$. In the limit of $\varepsilon \rightarrow 0$ the AP-reformulation reduces to the so called *Limit* model (L-model), whose solution is an acceptable approximation of the P-model solution for $\varepsilon \ll 1$ (see Theorem 2.2). In [12] the Asymptotic Preserving reformulation of the original problem was obtained in two steps. First, the original problem was

integrated along the field lines (z -axis) leading to an ε -independent elliptic problem for the mean part p . Second, the mean equation was subtracted from the original problem and the ε -dependent elliptic problem for the fluctuating part q was obtained. This approach however is not applicable if the field b is arbitrary. In this paper we present a new approach. Instead of integrating the original problem along the arbitrary field lines, we choose to force the mean part p to lie in the Hilbert space of functions constant along the field lines and the fluctuating part q to be orthogonal (in L^2 sense) to this space. This is done by a Lagrange multiplier technique and requires introduction of additional variables, thus enlarging the linear system to be solved. This method allows to treat the arbitrary b field case, regardless of the field topology, and thus eliminates the limitations of the algorithm presented in [12]. We note that an alternative method, bypassing the need in Lagrange multipliers, is proposed in [8]; it is based on a reformulation of the original problem as a fourth order equation.

The outline of this paper is the following. Section 2 introduces the original anisotropic elliptic problem. The original problem will be referred to as the *Singular-Perturbation* model (P-model). The mathematical framework is introduced and the *Asymptotic Preserving* reformulation (AP-model) is then derived. Section 3 is devoted to the numerical implementation of the AP-formulation. Numerical results are presented for 2D and 3D test cases, for constant and variable fields b . Three methods are compared (AP-formulation, P-model and L-model) according to their precision for different values of ε . The rigorous numerical analysis of this new algorithm will be the subject of a forthcoming publication.

2. Problem definition

We consider a two or three dimensional anisotropic problem, given on a sufficiently smooth, bounded domain $\Omega \subset \mathbb{R}^d$, $d=2,3$ with boundary $\partial\Omega$. The direction of the anisotropy is defined by the vector field $b \in (C^\infty(\Omega))^d$, satisfying $|b(x)|=1$ for all $x \in \Omega$.

Given this vector field b , one can decompose vectors $v \in \mathbb{R}^d$, gradients $\nabla\phi$, with $\phi(x)$ a scalar function, and divergences $\nabla \cdot v$, with $v(x)$ a vector field, into a part parallel to the anisotropy direction and a part perpendicular to it. These parts are defined as follows:

$$\begin{aligned} v_{\parallel} &:= (v \cdot b)b, & v_{\perp} &:= (Id - b \otimes b)v, & \text{such that } v &= v_{\parallel} + v_{\perp}, \\ \nabla_{\parallel}\phi &:= (b \cdot \nabla\phi)b, & \nabla_{\perp}\phi &:= (Id - b \otimes b)\nabla\phi, & \text{such that } \nabla\phi &= \nabla_{\parallel}\phi + \nabla_{\perp}\phi, \\ \nabla_{\parallel} \cdot v &:= \nabla \cdot v_{\parallel}, & \nabla_{\perp} \cdot v &:= \nabla \cdot v_{\perp}, & \text{such that } \nabla \cdot v &= \nabla_{\parallel} \cdot v + \nabla_{\perp} \cdot v, \end{aligned} \quad (2.1)$$

where we denote by \otimes the vector tensor product. With these notations we can now introduce, the so-called Singular Perturbation problem, whose numerical solution is the main concern of this paper.

2.1. The singular Perturbation problem (P-model). We consider the following Singular Perturbation problem:

$$(P) \begin{cases} -\frac{1}{\varepsilon} \nabla_{\parallel} \cdot (A_{\parallel} \nabla_{\parallel} \phi^{\varepsilon}) - \nabla_{\perp} \cdot (A_{\perp} \nabla_{\perp} \phi^{\varepsilon}) = f & \text{in } \Omega, \\ \frac{1}{\varepsilon} n_{\parallel} \cdot (A_{\parallel} \nabla_{\parallel} \phi^{\varepsilon}) + n_{\perp} \cdot (A_{\perp} \nabla_{\perp} \phi^{\varepsilon}) = 0 & \text{on } \partial\Omega_N, \\ \phi^{\varepsilon} = 0 & \text{on } \partial\Omega_D, \end{cases} \quad (2.2)$$

where n is the outward normal to Ω and the boundaries are defined by

$$\partial\Omega_D = \{x \in \partial\Omega \mid b(x) \cdot n = 0\}, \quad \partial\Omega_N = \partial\Omega \setminus \partial\Omega_D. \quad (2.3)$$

The parameter $0 < \varepsilon < 1$ can be very small and is responsible for the high anisotropy of the problem. The aim is to introduce a numerical scheme whose computational costs (simulation time and memory), for fixed precision, are independent of ε .

We shall assume in the rest of this paper the following hypothesis on the diffusion coefficients and the source terms.

Hypothesis A. *Let $f \in L^2(\Omega)$ and $\partial\Omega_D \neq \emptyset$. The diffusion coefficients $A_{\parallel} \in L^\infty(\Omega)$ and $A_{\perp} \in \mathbb{M}_{d \times d}(L^\infty(\Omega))$ are supposed to satisfy*

$$0 < A_0 \leq A_{\parallel}(x) \leq A_1, \quad f.a.a. \ x \in \Omega, \quad (2.4)$$

$$A_0 \|v\|^2 \leq v^t A_{\perp}(x) v \leq A_1 \|v\|^2, \quad \forall v \in \mathbb{R}^d \text{ and } f.a.a. \ x \in \Omega. \quad (2.5)$$

As we intend to use the finite element method for the numerical solution of the P-problem, let us put (2.2) under variational form. For this let \mathcal{V} be the Hilbert space

$$\mathcal{V} := \{\phi \in H^1(\Omega) / \phi|_{\partial\Omega_D} = 0\}, \quad (\phi, \psi)_{\mathcal{V}} := (\nabla_{\parallel}\phi, \nabla_{\parallel}\psi)_{L^2} + \varepsilon(\nabla_{\perp}\phi, \nabla_{\perp}\psi)_{L^2}.$$

Thus, we are seeking $\phi^\varepsilon \in \mathcal{V}$, the solution of

$$a_{\parallel}(\phi^\varepsilon, \psi) + \varepsilon a_{\perp}(\phi^\varepsilon, \psi) = \varepsilon(f, \psi), \quad \forall \psi \in \mathcal{V}, \quad (2.6)$$

where (\cdot, \cdot) stands for the standard L^2 inner product and the continuous bilinear forms $a_{\parallel} : \mathcal{V} \times \mathcal{V} \rightarrow \mathbb{R}$ and $a_{\perp} : \mathcal{V} \times \mathcal{V} \rightarrow \mathbb{R}$ are given by

$$a_{\parallel}(\phi, \psi) := \int_{\Omega} A_{\parallel} \nabla_{\parallel}\phi \cdot \nabla_{\parallel}\psi \, dx, \quad a_{\perp}(\phi, \psi) := \int_{\Omega} (A_{\perp} \nabla_{\perp}\phi) \cdot \nabla_{\perp}\psi \, dx. \quad (2.7)$$

Thanks to Hypothesis A and the Lax-Milgram theorem, problem (2.2) admits a unique solution $\phi^\varepsilon \in \mathcal{V}$ for all fixed $\varepsilon > 0$.

2.2. The Limit problem (L-model). The direct numerical solution of (2.2) may be very inaccurate for $\varepsilon \ll 1$. Indeed, when ε tends to zero, the system reduces to

$$\begin{cases} -\nabla_{\parallel} \cdot (A_{\parallel} \nabla_{\parallel}\phi) = 0 & \text{in } \Omega, \\ n_{\parallel} \cdot (A_{\parallel} \nabla_{\parallel}\phi) = 0 & \text{on } \partial\Omega_N, \\ \phi = 0 & \text{on } \partial\Omega_D. \end{cases} \quad (2.8)$$

This is an ill-posed problem as it has an infinite number of solutions $\phi \in \mathcal{G}$, where

$$\mathcal{G} = \{\phi \in \mathcal{V} \mid \nabla_{\parallel}\phi = 0\} \quad (2.9)$$

is the Hilbert space of functions which are constant along the field lines of b . This shows that the condition number of the system obtained by discretizing (2.2) tends to ∞ as $\varepsilon \rightarrow 0$ so that its solution will suffer from round-off errors.

For this reason, we should approximate (2.2) in the limit $\varepsilon \rightarrow 0$ differently. Supposing that $\phi^\varepsilon \rightarrow \phi^0$ as $\varepsilon \rightarrow 0$ we identify (at first formally) the problem satisfied by ϕ^0 . From the above arguments we know that $\phi^0 \in \mathcal{G}$. Taking now test functions $\psi \in \mathcal{G}$ in (2.6), we obtain

$$\int_{\Omega} A_{\perp} \nabla_{\perp}\phi^\varepsilon \cdot \nabla_{\perp}\psi \, dx = \int_{\Omega} f \psi \, dx. \quad (2.10)$$

Passing to the limit $\varepsilon \rightarrow 0$ yields the variational formulation of the problem satisfied by ϕ^0 : (Limit problem): find $\phi^0 \in \mathcal{G}$, the solution of

$$(L) \int_{\Omega} A_{\perp} \nabla_{\perp} \phi^0 \cdot \nabla_{\perp} \psi \, dx = \int_{\Omega} f \psi \, dx, \quad \forall \psi \in \mathcal{G}, \quad (2.11)$$

which is a well posed problem. Indeed, the space $\mathcal{G} \subset \mathcal{V}$ is a Hilbert space associated with the inner product

$$(\phi, \psi)_{\mathcal{G}} := (\nabla_{\perp} \phi, \nabla_{\perp} \psi)_{L^2}, \quad \forall \phi, \psi \in \mathcal{G}, \quad (2.12)$$

and the norm $\|\cdot\|_{\mathcal{G}}$ is equivalent to the H^1 norm. This is due to the Poincaré inequality, as

$$\|\phi\|_{L^2}^2 \leq C \|\nabla \phi\|_{L^2}^2 = C \|\nabla_{\parallel} \phi\|_{L^2}^2 + C \|\nabla_{\perp} \phi\|_{L^2}^2 = C \|\nabla_{\perp} \phi\|_{L^2}^2, \quad \forall \phi \in \mathcal{G}.$$

Hypothesis A and the Lax-Milgram lemma imply the existence and uniqueness of a solution $\phi^0 \in \mathcal{G}$ of the Limit problem (2.11).

REMARK 2.1. For the moment let us restrict ourselves to the simple special case (considered in a previous paper [12]) of the two dimensional domain $\Omega = (0, L_x) \times (0, L_z)$ in the (x, z) plane with a constant b -field aligned with the Z -axis:

$$b = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (2.13)$$

The functions in the space \mathcal{G} are independent of z , so that \mathcal{G} can be identified with $H_0^1(0, L_x)$. The limit problem (2.11) now reads: Find ϕ^0 in $H_0^1(0, L_x)$ satisfying

$$\int_0^{L_x} \bar{A}_{\perp}(x) \partial_x \phi^0(x) \partial_x \psi(x) \, dx = \int_0^{L_x} \bar{f}(x) \psi(x) \, dx, \quad \forall \psi \in H_0^1(0, L_x),$$

where $\bar{A}_{\perp}(x) = (1/L_z) \int_0^{L_z} A_{\perp, 11}(x, z) \, dz$ and $\bar{f}(x) = (1/L_z) \int_0^{L_z} f(x, z) \, dz$ are the mean values of A_{\perp} and f along the field lines. The limit solution ϕ^0 thus satisfies a one-dimensional elliptic equation whose coefficients are integrated along the anisotropy direction:

$$\begin{aligned} -\partial_x \left(\bar{A}_{\perp}(x) \partial_x \phi^0(x) \right) &= \bar{f}(x) \text{ on } (0, L_x), \\ \phi^0(0) &= \phi^0(L_x) = 0. \end{aligned} \quad (2.14)$$

We see now that $\phi^0(x)$ is a solution to the one dimensional elliptic problem, so that it belongs to $H^2(0, L_x)$ provided $f \in L^2(\Omega)$. Since ϕ^0 as a function of (x, z) does not depend on z , we have also $\phi^0 \in H^2(\Omega)$. This conclusion ($\phi^0 \in H^2(\Omega)$) remains valid in the case of a cylindrical three dimensional domain $\Omega = \Omega_{xy} \times (0, L_z)$ in the (x, y, z) space with any sufficiently smooth Ω_{xy} in the (x, y) plane and the field b aligned with the Z -axis, $b = (0, 0, 1)^t$. Indeed, it is easy to see that in this case $\phi^0 = \phi^0(x, y)$ solves an elliptic two dimensional problem in Ω_{xy} similar to (2.14), so that we can apply the standard regularity results for elliptic problems. These examples show that it is reasonable to suppose $\phi^0 \in H^2(\Omega)$ in more general geometries of Ω and b . This can be indeed proved under the hypotheses in Appendix A by specifying the $(d-1)$ dimensional elliptic problem for ϕ^0 . The proof, being rather lengthy and technical, is postponed to a forthcoming work [14].

2.3. The Asymptotic Preserving approach (AP-model). In this section we introduce the AP-formulation, which is a reformulation of the Singular Perturbation problem (2.2), and permits a “continuous” transition from the (P)-problem (2.2) to the (L)-problem (2.11) as $\varepsilon \rightarrow 0$. For this purpose, each function is decomposed into its mean part along the anisotropy direction (lying in the subspace \mathcal{G} of \mathcal{V}) and a fluctuating part (cf. [12]) lying in the L^2 -orthogonal complement \mathcal{A} of \mathcal{G} in \mathcal{V} , defined by

$$\mathcal{A} := \{\phi \in \mathcal{V} \mid (\phi, \psi) = 0, \forall \psi \in \mathcal{G}\}. \quad (2.15)$$

Note that (\cdot, \cdot) denotes here and elsewhere the inner product of $L^2(\Omega)$.

In what follows, we need the following

Hypothesis B. *The Hilbert-space \mathcal{V} admits the decomposition*

$$\mathcal{V} = \mathcal{G} \oplus^\perp \mathcal{A}, \quad (2.16)$$

with \mathcal{G} given by (2.9) and \mathcal{A} given by (2.15), and where the orthogonality of the direct sum is taken with respect to the L^2 -norm. Denoting by P the orthogonal projection on \mathcal{G} with respect to the L^2 inner product:

$$P: \mathcal{V} \rightarrow \mathcal{G} \text{ such that } (P\phi, \psi) = (\phi, \psi) \quad \forall \phi \in \mathcal{V}, \psi \in \mathcal{G}, \quad (2.17)$$

we shall suppose that this mapping is continuous and that we have the Poincaré-Wirtinger inequality

$$\|\phi - P\phi\|_{L^2(\Omega)} \leq C \|\nabla_{\parallel} \phi\|_{L^2(\Omega)}, \quad \forall \phi \in \mathcal{V}. \quad (2.18)$$

Applying the projection P to a function ϕ is nothing but a weighted average of ϕ along the anisotropy field lines of b . The space \mathcal{G} is the space of averaged functions (the parallel Gradient of these averaged functions being equal to zero), whereas the space \mathcal{A} is the space of the fluctuations (the Average of the fluctuations being equal to zero). Note that the decomposition (2.16) is not self evident (the orthogonality is assumed here in the L^2 sense, not that of H^1 !) and it may in fact fail on some “pathological” domains Ω , cf. Example A.1. Indeed, although one can always define an L^2 -orthogonal projection $\tilde{P}\phi$ on the space of functions constant along each field line, for any ϕ with square-integrable $\nabla_{\parallel} \phi$, one cannot assure in general that $\tilde{P}\phi$ belongs to \mathcal{V} for $\phi \in \mathcal{V}$ since one may lose control of the perpendicular part of the gradient of $\tilde{P}\phi$. Fortunately however, Hypothesis B is typically satisfied for the domains of practical interest. The interested reader is referred to Appendix A for an example of a set of assumptions on Ω and b which entail Hypothesis B and which essentially reduces to the requirement that the field b to intersect $\partial\Omega_N$ in a uniformly non-tangential manner and for the boundary components $\partial\Omega_N$ and $\partial\Omega_D$ to be sufficiently smooth.

Let us also define the operator

$$Q: \mathcal{V} \rightarrow \mathcal{A}, \quad Q = I - P. \quad (2.19)$$

Each function $\phi \in \mathcal{V}$ can be decomposed uniquely as $\phi = p + q$, where $p = P\phi \in \mathcal{G}$ and $q = Q\phi \in \mathcal{A}$. Using this decomposition, we reformulate the Singular-Perturbation problem (2.2). Indeed, replacing $\phi^\varepsilon := p^\varepsilon + q^\varepsilon$ in problem (2.2) and taking test functions

$\eta \in \mathcal{G}$ and $\xi \in \mathcal{A}$ leads to an asymptotic preserving formulation of the original problem: Find $(p^\varepsilon, q^\varepsilon) \in \mathcal{G} \times \mathcal{A}$ such that

$$\begin{cases} a_\perp(p^\varepsilon, \eta) + a_\perp(q^\varepsilon, \eta) = (f, \eta), & \forall \eta \in \mathcal{G}, \\ a_{||}(q^\varepsilon, \xi) + \varepsilon a_\perp(q^\varepsilon, \xi) + \varepsilon a_\perp(p^\varepsilon, \xi) = \varepsilon(f, \xi), & \forall \xi \in \mathcal{A}. \end{cases} \quad (2.20)$$

Contrary to the Singular Perturbation problem (2.2), formally setting $\varepsilon = 0$ in (2.20) yields the system

$$\begin{cases} a_\perp(p^0, \eta) + a_\perp(q^0, \eta) = (f, \eta), & \forall \eta \in \mathcal{G}, \\ a_{||}(q^0, \xi) = 0, & \forall \xi \in \mathcal{A}, \end{cases} \quad (2.21)$$

which has a unique solution $(p^0, q^0) \in \mathcal{G} \times \mathcal{A}$, where p^0 is the unique solution of the L-problem (2.11) and $q^0 \equiv 0$. Indeed, taking $\xi = q^0$ as test function in the second equation of (2.21) yields $\nabla_{||} q^0 = 0$, which means $q^0 \in \mathcal{G}$. But at the same time, $q^0 \in \mathcal{A}$, so that $q^0 \in \mathcal{G} \cap \mathcal{A} = \{0\}$. Setting $q^0 \equiv 0$ in the first equation of (2.21), shows that p^0 is the unique solution of the L-problem.

THEOREM 2.2. *For every $\varepsilon > 0$ the Asymptotic Preserving formulation (2.20), under Hypotheses A and B, admits a unique solution $(p^\varepsilon, q^\varepsilon) \in \mathcal{G} \times \mathcal{A}$, where $\phi^\varepsilon := p^\varepsilon + q^\varepsilon$ is the unique solution in \mathcal{V} of the Singular Perturbation model (2.2).*

These solutions satisfy the bounds

$$\|\phi^\varepsilon\|_{H^1(\Omega)} \leq C\|f\|_{L^2(\Omega)}, \quad \|q^\varepsilon\|_{H^1(\Omega)} \leq C\|f\|_{L^2(\Omega)}, \quad \|p^\varepsilon\|_{H^1(\Omega)} \leq C\|f\|_{L^2(\Omega)}, \quad (2.22)$$

with an ε -independent constant $C > 0$. Moreover, we have

$$\phi^\varepsilon \rightarrow \phi^0, \quad p^\varepsilon \rightarrow \phi^0, \quad \text{and} \quad q^\varepsilon \rightarrow 0 \quad \text{in} \quad H^1(\Omega) \quad \text{as} \quad \varepsilon \rightarrow 0, \quad (2.23)$$

where $\phi^0 \in \mathcal{G}$ is the unique solution of the Limit model (2.11).

Proof. The existence and uniqueness of a solution for the P-problem as well as L-problem are consequences of the Lax-Milgram theorem. The existence and uniqueness of a solution of (2.20) is then immediate by construction, after noting that the decomposition $\phi^\varepsilon = p^\varepsilon + q^\varepsilon$ is unique.

The bound $\|\phi^\varepsilon\|_{H^1(\Omega)} \leq C\|f\|_{L^2(\Omega)}$ is obtained by a standard elliptic argument. Furthermore $p^\varepsilon = P\phi^\varepsilon$, where P is the L^2 -orthogonal projector on \mathcal{G} , which is a bounded operator in \mathcal{V} by (A.4). This implies the estimates for p^ε and q^ε in (2.22). Since $p^\varepsilon \in \mathcal{G}$ and $q^\varepsilon \in \mathcal{A}$ are bounded, there exist subsequences p^{ε_n} and q^{ε_n} that weakly converge for $\varepsilon_n \rightarrow 0$ to some $p^0 \in \mathcal{G}$ and $q^0 \in \mathcal{A}$. Taking $\varepsilon = \varepsilon_n$ in (2.20) and passing to the limit $\varepsilon_n \rightarrow 0$ we identify (p^0, q^0) with the unique solution of (2.21), i.e. $p^0 = \phi^0$ is the unique solution of (2.11) and $q^0 \equiv 0$. Since the limit does not depend on the choice of the subsequence, we have the weak convergence as $\varepsilon \rightarrow 0$, i.e.

$$p^\varepsilon \rightharpoonup p^0 \quad \text{in} \quad H^1(\Omega), \quad q^\varepsilon \rightharpoonup 0 \quad \text{in} \quad H^1(\Omega).$$

We shall prove now that these convergences are actually strong. Introducing $e^\varepsilon = p^\varepsilon - p^0$, we have

$$a_\perp(e^\varepsilon, \eta) + a_\perp(q^\varepsilon, \eta) = 0, \quad \forall \eta \in \mathcal{G}.$$

Taking $\eta = e^\varepsilon$ in this relation and adding it to the second equation in (2.20), where we put $\xi = q^\varepsilon/\varepsilon$, yields

$$\frac{1}{\varepsilon}a_{\parallel}(q^\varepsilon, q^\varepsilon) + a_{\perp}(q^\varepsilon + e^\varepsilon, q^\varepsilon + e^\varepsilon) = (f, q^\varepsilon) - a_{\perp}(p^0, q^\varepsilon). \quad (2.24)$$

Due to the Poincaré-Wirtinger Equation (2.18), there exist a constant $C > 0$ such that

$$\|q\|_{L^2(\Omega)} \leq C a_{\parallel}(q, q)^{1/2}, \quad \forall q \in \mathcal{A}. \quad (2.25)$$

In combination with a Young inequality this gives $(f, q^\varepsilon) \leq \|f\|_{L^2} \|q^\varepsilon\|_{L^2} \leq \varepsilon \frac{C^2}{2} \|f\|_{L^2}^2 + \frac{1}{2\varepsilon} a_{\parallel}(q^\varepsilon, q^\varepsilon)$. Using this in the right hand side of (2.24), we arrive at

$$\frac{1}{2\varepsilon} a_{\parallel}(q^\varepsilon, q^\varepsilon) + a_{\perp}(q^\varepsilon + e^\varepsilon, q^\varepsilon + e^\varepsilon) \leq \varepsilon \frac{C^2}{2} \|f\|_{L^2}^2 - a_{\perp}(p^0, q^\varepsilon).$$

Noting that $q^\varepsilon + e^\varepsilon = \phi^\varepsilon - p^0$ and $\nabla_{\parallel} e^\varepsilon = 0$ we can rewrite this last inequality as

$$\frac{1}{2\varepsilon} a_{\parallel}(\phi^\varepsilon - p^0, \phi^\varepsilon - p^0) + a_{\perp}(\phi^\varepsilon - p^0, \phi^\varepsilon - p^0) \leq \varepsilon \frac{C^2}{2} \|f\|_{L^2}^2 - a_{\perp}(p^0, q^\varepsilon).$$

Since $a_{\perp}(p^0, q^\varepsilon) \rightarrow 0$ as $\varepsilon \rightarrow 0$ (thanks to the weak convergence $q^\varepsilon \rightharpoonup 0$) we observe that $\phi^\varepsilon \rightarrow p^0$ strongly in $H^1(\Omega)$. Reminding again that $p^\varepsilon = P\phi^\varepsilon$ and P is bounded in the norm of $H^1(\Omega)$, we obtain also $p^\varepsilon \rightarrow Pp^0 = p^0$, which entails $q^\varepsilon \rightarrow 0$. \square

REMARK 2.3. Let us return to the simple special case discussed in Remark 2.1, i.e. $\Omega = (0, L_x) \times (0, L_z)$ and the b -field given by (2.13). Recall that the space \mathcal{G} can be identified in this case with the space of functions constant along the Z -axis, which means $\mathcal{G} := \{\phi \in \mathcal{V} / \partial_z \phi = 0\}$. The space \mathcal{A} is orthogonal (with respect to the L^2 -norm) to \mathcal{G} and thus contains the functions that have zero mean value along the Z -axis, i.e. $\mathcal{A} := \{\phi \in \mathcal{V} / \int_0^{L_z} \phi(x, z) dz = 0\}$. Therefore, for $\phi^\varepsilon = p^\varepsilon + q^\varepsilon \in \mathcal{V}$, the function p^ε is the mean value of ϕ^ε in the direction of the field b :

$$p^\varepsilon = \frac{1}{L_z} \int_0^{L_z} \phi^\varepsilon dz, \quad (2.26)$$

and q^ε is the fluctuating part with zero mean value:

$$q^\varepsilon = \phi^\varepsilon - \frac{1}{L_z} \int_0^{L_z} \phi^\varepsilon dz. \quad (2.27)$$

Hypothesis B is thus easily verified. The results obtained in this special case were presented in a previous paper [12]. In the case of an arbitrary b -field, formula (2.26) is generalized as (A.2) in Appendix A, where the length element along the b -field line is weighted by the infinitesimal cross-sectional area of the field tube around the considered b -field-line. This formula can be thus interpreted as a consequence of the co-area formula. Note that in the special case of a uniform anisotropy direction, the limit problem can easily be formulated as an elliptic problem depending only on the transverse coordinates (see Equation (2.14)). The size of the problem is thus significantly smaller than that of the initial one. This feature still occurs for non-uniform b -fields as long as adapted coordinates and meshes are used. In our case, aligned and transverse coordinates are not at our disposal and the solution of the limit problem must be searched as a function of the whole set of coordinates.

2.4. Lagrange multiplier space. The objective of this work is the numerical solution of system (2.20) and the comparison of the obtained results with those obtained by directly solving the original problem (2.2). In a general case, when the field b is not necessarily constant, the discretization of the subspaces \mathcal{G} and \mathcal{A} is not straightforward as in the simpler case [12]. In order to overcome this difficulty a Lagrange multiplier technique will be used.

2.4.1. The \mathcal{A} space. To avoid the use of the constrained space \mathcal{A} , we can remark that \mathcal{A} can be characterized as being the orthogonal complement (in the L^2 sense) of the \mathcal{G} -space. Thus, instead of (2.20), the slightly changed system will be solved: find $(p^\varepsilon, q^\varepsilon, l^\varepsilon) \in \mathcal{G} \times \mathcal{V} \times \mathcal{G}$ such that

$$\begin{cases} a_\perp(p^\varepsilon, \eta) + a_\perp(q^\varepsilon, \eta) = (f, \eta) & \forall \eta \in \mathcal{G}, \\ a_\parallel(q^\varepsilon, \xi) + \varepsilon a_\perp(q^\varepsilon, \xi) + \varepsilon a_\perp(p^\varepsilon, \xi) + (l^\varepsilon, \xi) = \varepsilon(f, \xi) & \forall \xi \in \mathcal{V}, \\ (q^\varepsilon, \chi) = 0 & \forall \chi \in \mathcal{G}. \end{cases} \quad (2.28)$$

The constraint $(q^\varepsilon, \chi) = 0, \forall \chi \in \mathcal{G}$ is forcing the solution q^ε to belong to \mathcal{A} , and this property is carried over to the limit $\varepsilon \rightarrow 0$. We have thus circumvented the difficulty of discretizing \mathcal{A} by introducing a new variable and enlarging the linear system.

PROPOSITION 2.4. *Problems (2.20) and (2.28) are equivalent. Indeed, $(p^\varepsilon, q^\varepsilon) \in \mathcal{G} \times \mathcal{A}$ is the unique solution of (2.20) if and only if $(p^\varepsilon, q^\varepsilon, l^\varepsilon) \in \mathcal{G} \times \mathcal{V} \times \mathcal{G}$ with $l^\varepsilon \equiv 0$ is the unique solution of (2.28).*

Proof. Let $(p^\varepsilon, q^\varepsilon) \in \mathcal{G} \times \mathcal{A}$ be the unique solution of (2.20). Then it is trivial to show that $(p^\varepsilon, q^\varepsilon, 0)$ solves (2.28). Let now $(p^\varepsilon, q^\varepsilon, l^\varepsilon) \in \mathcal{G} \times \mathcal{V} \times \mathcal{G}$ be a solution of (2.28). Then the last equation of (2.28) implies that $q^\varepsilon \in \mathcal{A}$. Choosing in the second equation a test function $\xi \in \mathcal{G}$, one gets

$$\varepsilon a_\perp(q^\varepsilon, \xi) + \varepsilon a_\perp(p^\varepsilon, \xi) + (l^\varepsilon, \xi) = \varepsilon(f, \xi), \quad \forall \xi \in \mathcal{G},$$

which, because of the first equation in (2.28), yields $(l^\varepsilon, \xi) = 0$ for all $\xi \in \mathcal{G}$. Thus $l^\varepsilon \equiv 0$. \square

2.4.2. The \mathcal{G} space. In order to eliminate the problems that arise when dealing with the discretization of \mathcal{G} , the Lagrange multiplier method will again be used. First note that

$$p \in \mathcal{G} \Leftrightarrow \begin{cases} \nabla_\parallel p = 0, \\ p \in \mathcal{V} \end{cases} \Leftrightarrow \begin{cases} \int_\Omega A_\parallel \nabla_\parallel p \cdot \nabla_\parallel \lambda dx = a_\parallel(p, \lambda) = 0, \forall \lambda \in \mathcal{L}, \\ p \in \mathcal{V}, \end{cases} \quad (2.29)$$

where \mathcal{L} is a functional space that should be chosen large enough so that one could find for any $p \in \mathcal{V}$ a $\lambda \in \mathcal{L}$ with $\nabla_\parallel \lambda = \nabla_\parallel p$. On the other hand, the space \mathcal{L} should be not too large in order to ensure the uniqueness of the Lagrange multipliers in the unconstrained system. A space that satisfies these two requirements, under some quite general assumptions to be detailed later, can be defined as

$$\mathcal{L} := \{\lambda \in L^2(\Omega) / \nabla_\parallel \lambda \in L^2(\Omega), \lambda|_{\partial\Omega_{in}} = 0\}, \quad \text{with} \quad \partial\Omega_{in} := \{x \in \partial\Omega / b(x) \cdot n < 0\}. \quad (2.30)$$

Using the characterization (2.29) of the constrained space \mathcal{G} , we shall now reformulate the system (2.28) as follows: Find $(p^\varepsilon, \lambda^\varepsilon, q^\varepsilon, l^\varepsilon, \mu^\varepsilon) \in \mathcal{V} \times \mathcal{L} \times \mathcal{V} \times \mathcal{V} \times \mathcal{L}$ such that

$$(AP) \begin{cases} a_\perp(p^\varepsilon, \eta) + a_\perp(q^\varepsilon, \eta) + a_\parallel(\eta, \lambda^\varepsilon) = (f, \eta), & \forall \eta \in \mathcal{V}, \\ a_\parallel(p^\varepsilon, \kappa) = 0, & \forall \kappa \in \mathcal{L}, \\ a_\parallel(q^\varepsilon, \xi) + \varepsilon a_\perp(q^\varepsilon, \xi) + \varepsilon a_\perp(p^\varepsilon, \xi) + (l^\varepsilon, \xi) = \varepsilon(f, \xi), & \forall \xi \in \mathcal{V}, \\ (q^\varepsilon, \chi) + a_\parallel(\chi, \mu^\varepsilon) = 0, & \forall \chi \in \mathcal{V}, \\ a_\parallel(l^\varepsilon, \tau) = 0, & \forall \tau \in \mathcal{L}. \end{cases} \quad (2.31)$$

The advantage of the above formulation, as compared to (2.20), is that we only have to discretize the spaces \mathcal{V} and \mathcal{L} (at the price of the introduction of three additional variables), which is much easier than the discretization of the constrained spaces \mathcal{G} and \mathcal{A} . More importantly, the dual formulation (2.31) does not require any change of coordinates to express the fact that p^ε is constant along the b -field lines and that q^ε averages to zero along these lines. Therefore this formulation is particularly well adapted to time-dependent b -fields, as it does not require any operation which would have to be reinitiated as b evolves. The system (2.31) will be called the Asymptotic-Preserving formulation in the sequel.

To analyze this Asymptotic-Preserving formulation, we need the following

Hypothesis B' *The trace $\lambda|_{\partial\Omega_{in}}$ is well defined for any $\lambda \in \tilde{\mathcal{V}}$ as an element of $L^2(\partial\Omega_{in})$, with continuous dependence of the trace norm in $L^2(\partial\Omega_{in})$ on $\|\lambda\|_{\tilde{\mathcal{V}}}$. Moreover, the Hilbert space*

$$\tilde{\mathcal{V}} = \{\phi \in L^2(\Omega) / \nabla_\parallel \phi \in L^2(\Omega)\}, \quad (\phi, \psi)_{\tilde{\mathcal{V}}} := (\phi, \psi) + (\nabla_\parallel \phi, \nabla_\parallel \psi), \quad (2.32)$$

admits the decomposition

$$\tilde{\mathcal{V}} = \tilde{\mathcal{G}} \oplus \mathcal{L}, \quad (2.33)$$

where $\tilde{\mathcal{G}}$ is given by

$$\tilde{\mathcal{G}} := \{\phi \in \tilde{\mathcal{V}} / \nabla_\parallel \phi = 0\}, \quad (2.34)$$

and \mathcal{L} is given by (2.30). The spaces $\tilde{\mathcal{G}}$ and $G = \tilde{\mathcal{G}} \cap \mathcal{V}$ are related in the following way: if $g \in \tilde{\mathcal{G}}$ is such that $\int_{\partial\Omega_{in}} \eta g d\sigma = 0$ for all $\eta \in \mathcal{G}$, then $g = 0$.

The decomposition (2.33) is quite natural. It simply says that any function ϕ can be decomposed on each field line as a sum of a function that vanishes at one given point on this line and a constant (which is therefore the value of ϕ at this point). Hypothesis B' will be thus normally satisfied in cases of practical interest. For example, we prove in Appendix A that the set of assumptions on the domain Ω and the b -field, which can be used to verify Hypothesis B, is also sufficient (but far from necessary) for Hypothesis B'. We are now able to show the relation between systems (2.28) and (2.31).

PROPOSITION 2.5. *Assuming Hypotheses A, B, and B', problem (2.31) admits a unique solution*

$(p^\varepsilon, \lambda^\varepsilon, q^\varepsilon, l^\varepsilon, \mu^\varepsilon) \in \mathcal{V} \times \mathcal{L} \times \mathcal{V} \times \mathcal{V} \times \mathcal{L}$, where $(p^\varepsilon, q^\varepsilon, l^\varepsilon) \in \mathcal{G} \times \mathcal{V} \times \mathcal{G}$ is the unique solution of (2.28).

The proof of Proposition 2.5 is based on the following two lemmas.

LEMMA 2.6. *Assume Hypothesis B' and let $p \in \tilde{\mathcal{V}}$ be such that $a_{||}(p, \lambda) = 0 \quad \forall \lambda \in \mathcal{L}$. Then $p \in \tilde{\mathcal{G}}$.*

Proof. Take any $\eta \in \tilde{\mathcal{V}}$ and write $\eta = \lambda + g$ with $\lambda \in \mathcal{L}$ and $g \in \tilde{\mathcal{G}}$. We have $a_{||}(p, g) = 0$, and hence $a_{||}(p, \eta) = 0$ for all $\eta \in \tilde{\mathcal{V}}$. This entails $\nabla_{||} p = 0$, hence $p \in \tilde{\mathcal{G}}$. \square

LEMMA 2.7. *Assume Hypothesis B' and let $F \in \tilde{\mathcal{V}}^*$ be such that $F(\eta) = 0$ for all $\eta \in \mathcal{G}$. Then the problem of finding $\lambda \in \mathcal{L}$ such that*

$$a_{||}(\eta, \lambda) = F(\eta), \quad \forall \eta \in \tilde{\mathcal{V}}, \quad (2.35)$$

has a unique solution.

Proof. Consider the bilinear form b on $\tilde{\mathcal{V}} \times \tilde{\mathcal{V}}$

$$b(u, v) = a_{||}(u, v) + \int_{\partial\Omega_{in}} uv d\sigma.$$

By Hypothesis B', this is an inner product on $\tilde{\mathcal{V}}$. Indeed, if $b(u, u) = 0$ then $u \in \tilde{\mathcal{G}} \cap \mathcal{L}$ so that $u = 0$. The Riesz representation theorem implies that the problem of finding $\mu \in \tilde{\mathcal{V}}$ such that

$$b(\eta, \mu) = F(\eta), \quad \forall \eta \in \tilde{\mathcal{V}},$$

has a unique solution. We can now decompose $\mu = \lambda + g$ with $\lambda \in \mathcal{L}$ and $g \in \tilde{\mathcal{G}}$. This yields

$$a_{||}(\eta, \lambda) + \int_{\partial\Omega_{in}} \eta g d\sigma = F(\eta), \quad \forall \eta \in \tilde{\mathcal{V}},$$

so that, in particular, $\int_{\partial\Omega_{in}} \eta g d\sigma = 0$ for all $\eta \in \mathcal{G}$, which implies $g = 0$. We see now that λ is a solution to (2.35). The uniqueness follows easily. \square

Let us now prove Proposition 2.5.

Proof of existence in Proposition 2.5.

Proof. Take $(p^\varepsilon, q^\varepsilon, l^\varepsilon) \in \mathcal{G} \times \mathcal{V} \times \mathcal{G}$ as the unique solution of (2.28). Then equations 2,3,5 in (2.31) are immediately satisfied. It remains to properly choose the Lagrange multipliers $\lambda^\varepsilon, \mu^\varepsilon \in \mathcal{L}$ to satisfy equations 1,4 in (2.31). For this, let us define $F_1, F_2 \in \tilde{\mathcal{V}}^*$ by

$$F_1(\eta) := \frac{1}{\varepsilon} a_{||}(q^\varepsilon, \eta), \quad F_2(\eta) := -(q^\varepsilon, \eta), \quad \forall \eta \in \tilde{\mathcal{V}}. \quad (2.36)$$

These functionals are indeed continuous in the norm of $\tilde{\mathcal{V}}$ since their definitions do not contain the derivatives in directions perpendicular to b . Since $F_1(\eta) = F_2(\eta) = 0$ for all $\eta \in \mathcal{G}$, Lemma 2.7 implies the existence of $\lambda^\varepsilon \in \mathcal{L}$ and $\mu^\varepsilon \in \mathcal{L}$ such that

$$a_{||}(\eta, \lambda^\varepsilon) = F_1(\eta), \quad a_{||}(\chi, \mu^\varepsilon) = F_2(\chi), \quad \forall \eta, \chi \in \tilde{\mathcal{V}}. \quad (2.37)$$

Taking $\eta, \chi \in \mathcal{V} \subset \tilde{\mathcal{V}}$ we observe (cf. the second line in (2.28) where $l^\varepsilon = 0$)

$$\begin{aligned} a_{||}(\eta, \lambda^\varepsilon) &= \frac{1}{\varepsilon} a_{||}(q^\varepsilon, \eta) = (f, \eta) - a_\perp(p^\varepsilon, \eta) - a_\perp(q^\varepsilon, \eta), \quad \forall \eta \in \mathcal{V}, \\ a_{||}(\chi, \mu^\varepsilon) &= -(q^\varepsilon, \chi), \quad \forall \chi \in \mathcal{V}, \end{aligned}$$

which coincides with equations 1,4 in (2.31). \square

Proof of uniqueness in Proposition 2.5.

Proof. Consider the solution to system (2.31) with $f = 0$. Lemma 2.6 implies that $p^\varepsilon, l^\varepsilon \in \tilde{\mathcal{G}} \cap \mathcal{V} = \mathcal{G}$ and $(p^\varepsilon, q^\varepsilon, l^\varepsilon) \in \mathcal{G} \times \mathcal{V} \times \mathcal{G}$ satisfies (2.28) with $f = 0$, so that $p^\varepsilon = q^\varepsilon = l^\varepsilon = 0$ by Proposition 2.4. Equations 1,4 in (2.31) now tell us that $\lambda^\varepsilon, \mu^\varepsilon \in \tilde{\mathcal{G}}$. But $\tilde{\mathcal{G}} \cap \mathcal{L} = \{0\}$, so $\lambda^\varepsilon = \mu^\varepsilon = 0$. \square

The presence of $1/\varepsilon$ in the formulas (2.36), (2.37) defining λ^ε indicates at a first sight that λ^ε may tend to ∞ as $\varepsilon \rightarrow 0$, which would be disastrous for an AP numerical method based on (2.31) at very small ε . Fortunately λ^ε remains uniformly bounded in ε in the cases of practical interest. To ensure that this holds, it suffices to suppose that the limit solution ϕ^0 is in $H^2(\Omega)$, which is a reasonable assumption as discussed in Remark 2.1.

PROPOSITION 2.8. *Assume Hypotheses A, B, B', and $\phi^0 \in H^2(\Omega)$ where ϕ^0 is the solution to (2.11). Then λ^ε introduced in (2.31) satisfies*

$$\|\nabla_{||} \lambda^\varepsilon\|_{L^2} \leq C \max(\|f\|_{L^2}, \|\phi^0\|_{H^2}) \quad (2.38)$$

with a constant C independent of ε .

Proof. We will denote all the ε -independent constants by C in this proof. We start from relation (2.24) in the proof of Theorem 2.2. Dropping the positive term $a_\perp(q^\varepsilon + e^\varepsilon, q^\varepsilon + e^\varepsilon)$, it can be rewritten as

$$\frac{1}{\varepsilon} a_{||}(q^\varepsilon, q^\varepsilon) \leq (f, q^\varepsilon) - a_\perp(\phi^0, q^\varepsilon).$$

Since $\phi^0 \in H^2(\Omega)$ we can integrate by parts in the integral defining $a_\perp(\phi^0, q^\varepsilon)$:

$$\begin{aligned} -a_\perp(\phi^0, q^\varepsilon) &= - \int_{\Omega} A_\perp \nabla_\perp \phi^0 \cdot \nabla_\perp q^\varepsilon dx \\ &= - \int_{\partial\Omega_N} (Id - bb^t) A_\perp \nabla_\perp \phi^0 \cdot n q^\varepsilon d\sigma + \int_{\Omega} (\nabla_\perp \cdot A_\perp \nabla_\perp \phi^0) q^\varepsilon dx \\ &\leq C \|\phi^0\|_{H^2} (\|q^\varepsilon\|_{L^2(\partial\Omega_N)} + \|q^\varepsilon\|_{L^2(\Omega)}) \end{aligned}$$

since $\nabla \phi^0$ has a trace on $\partial\Omega$ and its norm in $L^2(\partial\Omega_N)$ is bounded by $C \|\phi^0\|_{H^2}$. Thus,

$$\frac{1}{\varepsilon} \|\nabla_{||} q^\varepsilon\|_{L^2}^2 \leq \frac{C}{\varepsilon} a_{||}(q^\varepsilon, q^\varepsilon) \leq C \|f\|_{L^2} \|q^\varepsilon\|_{L^2(\Omega)} + C \|\phi^0\|_{H^2} (\|q^\varepsilon\|_{L^2(\partial\Omega_N)} + \|q^\varepsilon\|_{L^2(\Omega)}).$$

By the Poincaré-Wirtinger inequality (2.18) (note that $Pq^\varepsilon = 0$) and by Hypothesis B' we have

$$\max(\|q^\varepsilon\|_{L^2(\Omega)}, \|q^\varepsilon\|_{L^2(\partial\Omega_N)}) \leq C \|\nabla_{||} q^\varepsilon\|_{L^2},$$

so that

$$\frac{1}{\varepsilon} \|\nabla_{||} q^\varepsilon\|_{L^2} \leq C \max(\|f\|_{L^2}, \|\phi^0\|_{H^2}).$$

This is the same as (2.38) since, according to (2.36) and (2.37), $\nabla_{||}\lambda^\varepsilon = \frac{1}{\varepsilon}\nabla_{||}q^\varepsilon$. \square

REMARK 2.9. The Limit model (2.11), reformulated using the Lagrange multiplier technique, now reads: Find $(\phi^0, \lambda^0) \in \mathcal{V} \times \mathcal{L}$ such that

$$(L') \begin{cases} \int_{\Omega} A_{\perp} \nabla_{\perp} \phi^0 \cdot \nabla_{\perp} \psi \, dx + \int_{\Omega} A_{||} \nabla_{||} \psi \cdot \nabla_{||} \lambda^0 \, dx = \int_{\Omega} f \psi \, dx, & \forall \psi \in \mathcal{V}, \\ \int_{\Omega} A_{||} \nabla_{||} \phi^0 \cdot \nabla_{||} \kappa \, dx = 0, & \forall \kappa \in \mathcal{L}. \end{cases} \quad (2.39)$$

Problem (2.39) is also well posed assuming Hypotheses A, B, B', and $\phi^0 \in H^2(\Omega)$. Indeed, the uniqueness of the solution to (2.39) can be proved in exactly the same manner as in the proof of Proposition 2.5 above. To prove the existence of a solution, it suffices to take the limit $\varepsilon \rightarrow 0$ in the first two lines of (2.31). Indeed, we know by Theorem 2.2 that $p^\varepsilon \rightarrow \phi^0$, the solution to (2.11), and $q^\varepsilon \rightarrow 0$ in $H^1(\Omega)$. Moreover, the family $\{\nabla_{||}\lambda^\varepsilon\}$ is bounded in the norm of $L^2(\Omega)$ by Proposition 2.8. We can take therefore a weakly convergence subsequence $\{\nabla_{||}\lambda^{\varepsilon_n}\}$ and identify its limit with $\{\nabla_{||}\lambda^0\}$ with some $\lambda^0 \in \mathcal{L}$ (cf. Lemma 2.7) to see that $(\phi^0, \lambda^0) \in \mathcal{V} \times \mathcal{L}$ solves (2.39).

3. Numerical method

This section concerns the discretization of the Asymptotic Preserving formulation (2.31), based on a finite element method, and the detailed study of the obtained numerical results. The numerical analysis of the present scheme is investigated in a forthcoming work [14]; in particular we are interested in the convergence of the scheme, independently of the parameter $\varepsilon > 0$.

Let us denote by $\mathcal{V}_h \subset \mathcal{V}$ and $\mathcal{L}_h \subset \mathcal{L}$ the finite dimensional approximation spaces, constructed by means of appropriate numerical discretizations (see Section 3.1 and Appendix B). We are therefore looking for a discrete solution $(p_h^\varepsilon, \lambda_h^\varepsilon, q_h^\varepsilon, l_h^\varepsilon, \mu_h^\varepsilon) \in \mathcal{V}_h \times \mathcal{L}_h \times \mathcal{V}_h \times \mathcal{V}_h \times \mathcal{L}_h$ of the following system:

$$\begin{cases} a_{\perp}(p_h^\varepsilon, \eta) + a_{\perp}(q_h^\varepsilon, \eta) + a_{||}(\eta, \lambda_h^\varepsilon) = (f, \eta), & \forall \eta \in \mathcal{V}_h, \\ a_{||}(p_h^\varepsilon, \kappa) = 0, & \forall \kappa \in \mathcal{L}_h, \\ a_{||}(q_h^\varepsilon, \xi) + \varepsilon a_{\perp}(q_h^\varepsilon, \xi) + \varepsilon a_{\perp}(p_h^\varepsilon, \xi) + (l_h^\varepsilon, \xi) = \varepsilon (f, \xi), & \forall \xi \in \mathcal{V}_h, \\ (q_h^\varepsilon, \chi) + a_{||}(\chi, \mu_h^\varepsilon) = 0, & \forall \chi \in \mathcal{V}_h, \\ a_{||}(l_h^\varepsilon, \tau) = 0, & \forall \tau \in \mathcal{L}_h. \end{cases} \quad (3.1)$$

Our numerical experiments indicate that the spaces \mathcal{V}_h and \mathcal{L}_h can be always taken of the same type and on the same mesh. The only difference between these two finite element spaces lies thus in the incorporation of boundary conditions. In general, let X_h denote the complete finite element space (without any restrictions on the boundary) which should be H^1 conforming but otherwise arbitrarily chosen. We define then

$$\mathcal{V}_h = \{v_h \in X_h / v_h|_{\partial\Omega_D} = 0\}, \quad (3.2)$$

$$\mathcal{L}_h = \{\lambda_h \in X_h / \lambda_h|_{\partial\Omega_{in} \cup \partial\Omega_D} = 0\}. \quad (3.3)$$

While this choice of \mathcal{V}_h is straightforward, the boundary conditions in \mathcal{L}_h require special attention. Indeed, nothing in the definition (2.30) of space \mathcal{L} , on the continuous

level, indicates that its elements should vanish on $\partial\Omega_D$. However, this liberty on $\partial\Omega_D$ is somewhat counter-intuitive. Indeed, the Lagrange multiplier $\lambda^\varepsilon \in \mathcal{L}$ serves to impose $\nabla_{||} p^\varepsilon = 0$ for some function p^ε taken from the space \mathcal{V} . But, for $p \in \mathcal{V}$ the trace on $\partial\Omega_D$ is zero so that $\nabla_{||} p^\varepsilon = 0$ there without the help of a Lagrange multiplier. Of course, this argument is not valid on the continuous level since the trace of functions in \mathcal{L} does not even necessarily exist. However this may become very important on the finite element level. Indeed, we provide in Appendix B an example of a finite element setting without incorporating $\lambda_h|_{\partial\Omega_D} = 0$ into the definition of \mathcal{L}_h , which leads to an ill-posed system (3.1). To avoid this difficulty, we choose \mathcal{L}_h as in (3.3) in all our experiments, thus obtaining well-posed problems.

3.1. Discretization of the (AP) problem. Let us present the discretization in a 2D case, the 3D case being a simple generalization. The here considered computational domain Ω is a square $\Omega = [0, 1] \times [0, 1]$. All simulations are performed on structured meshes. Let us introduce the Cartesian, homogeneous grid

$$x_i = i/N_x, \quad 0 \leq i \leq N_x, \quad y_j = j/N_y, \quad 0 \leq j \leq N_y, \quad (3.4)$$

where N_x and N_y are positive even constants corresponding to the number of discretization intervals in the x - resp. y -direction. The corresponding mesh-sizes are denoted by $h_x > 0$ resp. $h_y > 0$. Choosing a \mathbb{Q}_2 finite element method (\mathbb{Q}_2 -FEM), based on the quadratic base functions

$$\theta_{x_i} = \begin{cases} \frac{(x-x_{i-2})(x-x_{i-1})}{2h_x^2} & x \in [x_{i-2}, x_i], \\ \frac{(x_{i+2}-x)(x_{i+1}-x)}{2h_x^2} & x \in [x_i, x_{i+2}], \\ 0 & \text{else} \end{cases}, \quad \theta_{y_j} = \begin{cases} \frac{(y-y_{j-2})(y-y_{j-1})}{2h_y^2} & y \in [y_{j-2}, y_j], \\ \frac{(y_{j+2}-y)(y_{j+1}-y)}{2h_y^2} & y \in [y_j, y_{j+2}], \\ 0 & \text{else} \end{cases} \quad (3.5)$$

for even i, j and

$$\theta_{x_i} = \begin{cases} \frac{(x_{i+1}-x)(x-x_{i-1})}{h_x^2} & x \in [x_{i-1}, x_{i+1}], \\ 0 & \text{else} \end{cases}, \quad \theta_{y_j} = \begin{cases} \frac{(y_{j+1}-y)(y-y_{j-1})}{h_y^2} & y \in [y_{j-1}, y_{j+1}], \\ 0 & \text{else} \end{cases} \quad (3.6)$$

for odd i, j , we define

$$X_h := \{v_h = \sum_{i,j} v_{ij} \theta_{x_i}(x) \theta_{y_j}(y)\}.$$

We then search for discrete solutions $(p_h^\varepsilon, q_h^\varepsilon, l_h^\varepsilon) \in \mathcal{V}_h \times \mathcal{V}_h \times \mathcal{V}_h$ and $(\lambda_h^\varepsilon, \mu_h^\varepsilon) \in \mathcal{L}_h \times \mathcal{L}_h$ with \mathcal{V}_h and \mathcal{L}_h defined by (3.2) and (3.3). This leads to the inversion of a linear system, the corresponding matrix being non-symmetric and given by

$$A = \begin{pmatrix} A_1 & \tilde{A}_0 & A_1 & 0 & 0 \\ \tilde{A}_0^T & 0 & 0 & 0 & 0 \\ \varepsilon A_1 & 0 & A_0 + \varepsilon A_1 & C & 0 \\ 0 & 0 & C & 0 & \tilde{A}_0 \\ 0 & 0 & 0 & \tilde{A}_0^T & 0 \end{pmatrix}. \quad (3.7)$$

The matrices A_0, A_1 resp. C correspond to the bilinear forms $a_{||}(\cdot, \cdot), a_{\perp}(\cdot, \cdot)$ resp. (\cdot, \cdot) on $\mathcal{V}_h \times \mathcal{V}_h$, used in equations (2.31) and belong to $\mathbb{R}^{N \times N}$ where $N = \dim \mathcal{V}_h$, so that

$N = (N_x + 1)(N_y - 1)$ if, as in our numerical examples, $\partial\Omega_D$ consists of the bottom ($y=0$) and top ($y=1$) parts of the boundary. The matrix $\tilde{A}_0 \in \mathbb{R}^{N \times \tilde{N}}$ with $\tilde{N} = \dim \mathcal{L}_h$ is the sub-matrix of A_0 obtained by crossing out the columns corresponding to the mesh nodes on $\partial\Omega_{in}$. In the following numerical examples, we take $\partial\Omega_{in}$ as the left part of the boundary ($x=0$), so that $\tilde{N} = N_x(N_y - 1)$.

The matrix elements are computed using the 2D Gauss quadrature formula, with 3 points in the x and y direction:

$$\int_{-1}^1 \int_{-1}^1 f(x, y) = \sum_{i, j=-1}^1 \omega_i \omega_j f(x_i, y_j), \quad (3.8)$$

where $x_0 = y_0 = 0$, $x_{\pm 1} = y_{\pm 1} = \pm \sqrt{\frac{3}{5}}$, $\omega_0 = 8/9$ and $\omega_{\pm 1} = 5/9$, which is exact for polynomials of degree 5.

3.2. Discretization of (P) and (L) problems. To make comparisons, we have also discretized (P) and (L) problems using the same finite element spaces. Thus, the discretized (P) problem reads: Find $\phi_h^\varepsilon \in \mathcal{V}_h$ such that

$$a_{||}(\phi_h^\varepsilon, \psi) + \varepsilon a_{\perp}(\phi_h^\varepsilon, \psi) = \varepsilon(f, \psi), \quad \forall \psi \in \mathcal{V}_h.$$

This leads to the inversion of a linear system with the matrix $A_0 + \varepsilon A_1$.

The discretized (L) problem reads (cf Remark 2.9): find $(\phi_h^0, \lambda_h^0) \in \mathcal{V}_h \times \mathcal{L}_h$ such that

$$\begin{cases} a_{\perp}(\phi_h^0, \psi) + a_{||}(\psi, \lambda_h^0) = \int_{\Omega} f \psi dx, & \forall \psi \in \mathcal{V}_h, \\ a_{||}(\phi_h^0, \kappa) = 0, & \forall \kappa \in \mathcal{L}_h. \end{cases}$$

This leads to the inversion of a linear system with the matrix

$$\begin{pmatrix} A_1 & \tilde{A}_0 \\ \tilde{A}_0^T & 0 \end{pmatrix}.$$

In the following, we compare the three discretized methods outlined above on several test cases with manufactured solutions. The linear systems in all the discretized problems are solved directly, using the LU decomposition implemented in the library MUMPS [2].

3.3. Numerical Results.

3.3.1. 2D test case, uniform and aligned b -field. In this section we compare the numerical results obtained via the \mathbb{Q}_2 -FEM, by discretizing the Singular Perturbation model (2.2), the Limit model (2.11) and the Asymptotic Preserving reformulation (2.31). In all numerical tests we set $A_{\perp} = Id$ and $A_{||} = 1$. We start with a simple test case, where the analytical solution is known. Let the source term f be given by

$$f = (4 + \varepsilon) \pi^2 \cos(2\pi x) \sin(\pi y) + \pi^2 \sin(\pi y) \quad (3.9)$$

and the b field be aligned with the x -axis. Hence, the solution ϕ^ε of (2.2) and its decomposition $\phi^\varepsilon = p^\varepsilon + q^\varepsilon$ write

$$\phi^\varepsilon = \sin(\pi y) + \varepsilon \cos(2\pi x) \sin(\pi y), \quad (3.10)$$

$$p^\varepsilon = \sin(\pi y), \quad q^\varepsilon = \varepsilon \cos(2\pi x) \sin(\pi y). \quad (3.11)$$

We denote by ϕ_P , ϕ_L , ϕ_A the numerical solution of the Singular Perturbation model (2.2), the Limit model (2.11) and the Asymptotic Preserving reformulation (2.31) respectively. The comparison will be done in the L^2 -norm as well as the H^1 -norm. The linear systems obtained after discretization of the three methods with \mathbb{Q}_2 -FEM are solved using the same numerical algorithm — a direct solver, namely a LU decomposition implemented in a solver MUMPS[2]. In Figure 3.1 we plotted the

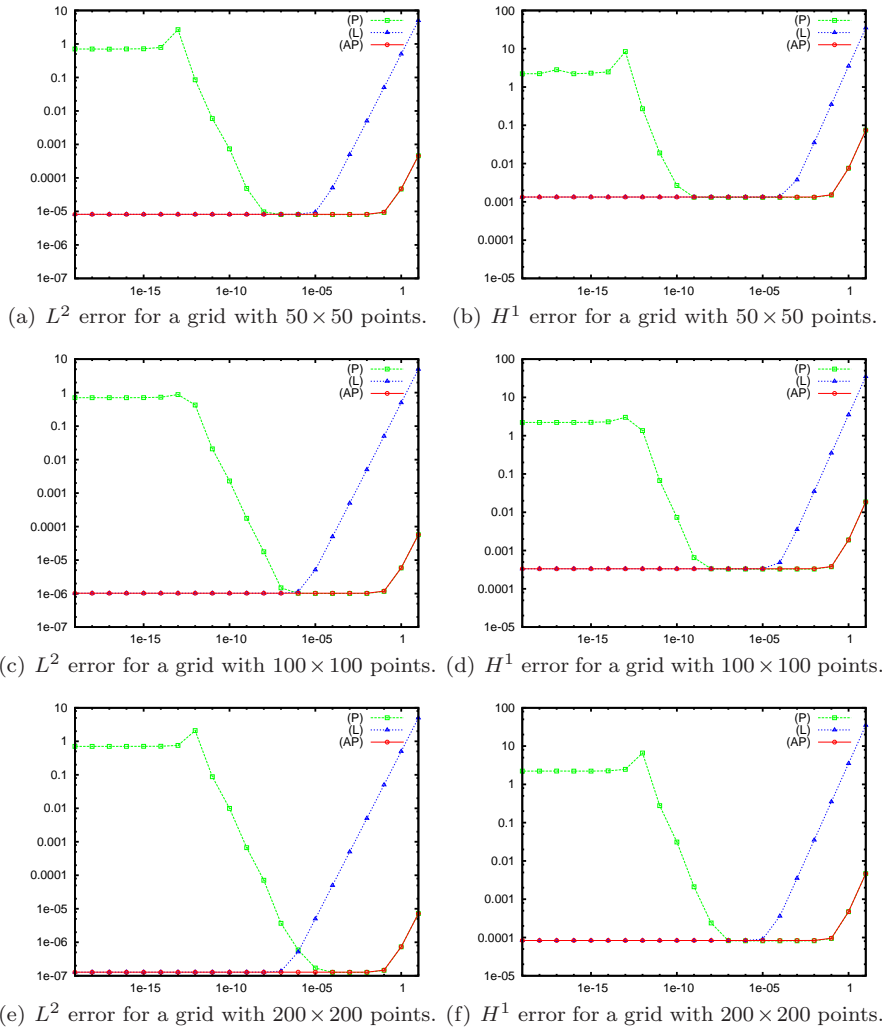


FIGURE 3.1. Absolute L^2 (left column) and H^1 (right column) errors between the exact solution ϕ^ε and the computed numerical solution ϕ_A (AP), ϕ_L (L), ϕ_P (P) for the test case with constant b . The error is plotted as a function of the parameter ε and for three different mesh-sizes.

absolute errors (in the L^2 resp. H^1 -norms) between the numerical solutions obtained with one of the three methods and the exact solution, and this as a function of the parameter ε and for several mesh-sizes. In Table 3.1 we specified the error values for one fixed grid and several ε -values. One observes that the Singular Perturbation finite element approximation is accurate only for ε bigger than some critical value ε_P ,

ε	AP scheme		Limit model		Singular Perturbation scheme	
	L^2 error	H^1 error	L^2 error	H^1 error	L^2 error	H^1 error
10	7.2×10^{-6}	4.7×10^{-3}	5.0×10^0	3.51×10^1	7.2×10^{-6}	4.7×10^{-3}
1	7.3×10^{-7}	4.7×10^{-4}	5.0×10^{-1}	3.51×10^0	7.3×10^{-7}	4.7×10^{-4}
10^{-1}	1.47×10^{-7}	9.6×10^{-5}	5.0×10^{-2}	3.51×10^{-1}	1.45×10^{-7}	9.4×10^{-5}
10^{-4}	1.28×10^{-7}	8.3×10^{-5}	5.0×10^{-5}	3.61×10^{-4}	1.26×10^{-7}	8.2×10^{-5}
10^{-6}	1.28×10^{-7}	8.3×10^{-5}	5.2×10^{-7}	8.4×10^{-5}	5.9×10^{-7}	8.2×10^{-5}
10^{-10}	1.28×10^{-7}	8.3×10^{-5}	1.28×10^{-7}	8.3×10^{-5}	9.9×10^{-3}	3.12×10^{-2}
10^{-15}	1.28×10^{-7}	8.3×10^{-5}	1.28×10^{-7}	8.3×10^{-5}	7.1×10^{-1}	2.23×10^0

TABLE 3.1. Comparison between the Asymptotic Preserving scheme, the Limit model and the Singular Perturbation model for $h=0.005$ (200 mesh points in each direction) and constant b : absolute L^2 -error and H^1 -error, for different ε -values.

method	# rows	# non zero	time	L^2 -error	H^1 -error
AP	50×10^3	1563×10^3	13.212 s	1.02×10^{-6}	3.34×10^{-4}
L	20×10^3	469×10^3	5.227 s	1.14×10^{-6}	3.34×10^{-4}
P	10×10^3	157×10^3	3.707 s	1.02×10^{-6}	3.27×10^{-4}

TABLE 3.2. Comparison between the Asymptotic Preserving scheme (AP), the Limit model (L), and the Singular Perturbation model (P) for $h=0.01$ (100 mesh points in each direction) and fixed $\varepsilon=10^{-6}$: matrix size, number of nonzero elements, average computational time and error in L^2 and H^1 norms.

the Limit model gives reliable results for ε smaller than ε_L , whereas the AP-scheme is accurate independently of ε . The order of convergence for all three methods is three in the L^2 -norm and two in the H^1 -norm, which is an optimal result for \mathbb{Q}_2 finite elements. When designing a robust numerical method one has therefore two options. The first one is to use an Asymptotic Preserving scheme, which is accurate independently of ε , but requires the solution of a bigger linear system. The second one is to design a coupling strategy that involves the solution of the Singular Perturbation formulation and the Limit problem in their respective validity domains. This is however a very delicate problem, since we observe that the critical values ε_P and ε_L are mesh dependent, namely ε_P is inversely proportional to h and ε_L is proportional to h . Therefore for small meshes there may exist a range of ε -values where neither the Singular Perturbation nor the Limit model finite element approximations give accurate results. For our test case, this is even the case for meshes as big as 200×200 points, if one regards the L^2 -norm. This mesh-size is generally insufficient in the case of real physical applications. Another interesting aspect with respect to which the three methods must be compared is the computational time and the size of the matrices involved in the linear systems. Table 3.2 shows that the Asymptotic Preserving scheme is expensive in computational time and memory requirements, as compared to the other methods. Indeed, the computational time required to solve the problem is almost four times bigger than that of the Singular Perturbation scheme. Moreover, the Asymptotic Preserving method involves matrices that have five times more rows and ten times more nonzero elements than the matrices obtained with the Singular Perturbation approximation. It is however the only scheme that provides the h -convergence regardless of ε . In order to reduce the computational costs, a coupling strategy for problems with variable ε will be proposed in a forthcoming paper.

In sub-domains where $\varepsilon > \varepsilon_P$ the Singular Perturbation problem will be solved, in sub-domains where $\varepsilon < \varepsilon_L$ the Limit problem will be solved and only in the remaining part, where neither the Limit nor the Singular Perturbation model are valid, the Asymptotic Preserving formulation will be solved.

3.3.2. 2D test case, non-uniform and non-aligned b -field. We now focus our attention on the original feature of the here introduced numerical method, namely its ability to treat nonuniform b fields. In this section we present numerical simulations performed for a variable field b .

First, let us construct a numerical test case. Finding an analytical solution for an arbitrary b presents a considerable difficulty. We have therefore chosen a different approach. First, we choose a limit solution

$$\phi^0 = \sin(\pi y + \alpha(y^2 - y)\cos(\pi x)), \quad (3.12)$$

where α is a numerical constant aimed at controlling the variations of b . For $\alpha=0$, the limit solution of the previous section is obtained. The limit solution for $\alpha=2$ is shown in Figure 3.2. We set $\alpha=2$ in what follows. Since ϕ^0 is a limit solution, it

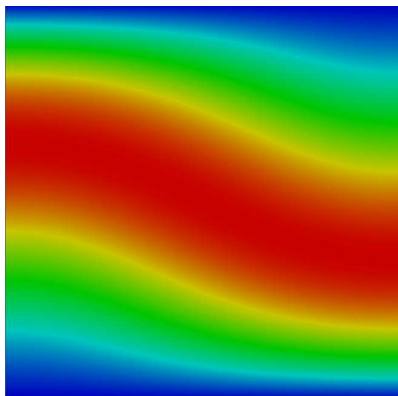


FIGURE 3.2. *The limit solution for the test case with variable b .*

is constant along the b field lines. Therefore we can determine the b field using the following implication:

$$\nabla_{\parallel} \phi^0 = 0 \quad \Rightarrow \quad b_x \frac{\partial \phi^0}{\partial x} + b_y \frac{\partial \phi^0}{\partial y} = 0, \quad (3.13)$$

which yields for example

$$b = \frac{B}{|B|}, \quad B = \begin{pmatrix} \alpha(2y-1)\cos(\pi x) + \pi \\ \pi\alpha(y^2-y)\sin(\pi x) \end{pmatrix}. \quad (3.14)$$

Note that the field B , constructed in this way, satisfies $\text{div}B=0$ — an important property in the framework of plasma simulation. Furthermore, we have $B \neq 0$ in the computational domain. Now, we choose ϕ^ε to be a function that converges, as $\varepsilon \rightarrow 0$, to the limit solution ϕ^0 :

$$\phi^\varepsilon = \sin(\pi y + \alpha(y^2 - y)\cos(\pi x)) + \varepsilon \cos(2\pi x)\sin(\pi y). \quad (3.15)$$

ε	AP scheme		Limit model		Singular Perturbation scheme	
	L^2 error	H^1 error	L^2 error	H^1 error	L^2 error	H^1 error
10	7.2×10^{-6}	4.6×10^{-3}	5.0×10^0	3.50×10^1	7.2×10^{-6}	4.6×10^{-3}
1	7.1×10^{-7}	4.6×10^{-4}	5.0×10^{-1}	3.50×10^0	7.1×10^{-7}	4.6×10^{-4}
10^{-2}	2.05×10^{-7}	1.33×10^{-4}	5.0×10^{-3}	3.50×10^{-2}	2.05×10^{-7}	1.33×10^{-4}
10^{-4}	2.12×10^{-7}	1.38×10^{-4}	5.0×10^{-5}	3.77×10^{-4}	1.74×10^{-6}	1.43×10^{-4}
10^{-7}	2.17×10^{-7}	1.41×10^{-4}	2.22×10^{-7}	1.41×10^{-4}	1.68×10^{-3}	1.26×10^{-2}
10^{-10}	2.17×10^{-7}	1.41×10^{-4}	2.17×10^{-7}	1.41×10^{-4}	3.93×10^{-1}	1.35×10^0
10^{-15}	2.17×10^{-7}	1.41×10^{-4}	2.17×10^{-7}	1.41×10^{-4}	6.7×10^{-1}	2.32×10^0

TABLE 3.3. Comparison between the Asymptotic preserving scheme, the Limit model and the Singular Perturbation model for $h=0.005$ (200 mesh points in each direction) and variable b : absolute L^2 -error and H^1 -error.

Finally, the force term is calculated using the equation, i.e.

$$f = -\nabla_{\perp} \cdot (A_{\perp} \nabla_{\perp} \phi^{\varepsilon}) - \frac{1}{\varepsilon} \nabla_{\parallel} \cdot (A_{\parallel} \nabla_{\parallel} \phi^{\varepsilon}).$$

As in the previous section, we compare the numerical solution of the Singular Perturbation model (2.2), the Limit model (2.11), and the Asymptotic Preserving reformulation (2.31), i.e. ϕ_P , ϕ_L , ϕ_A with the exact solution (3.15). The L^2 and H^1 -errors are reported on Figure 3.3 and Table 3.3. Once again the Asymptotic Preserving scheme proves to be valid for all values of ε , contrary to the other schemes. There is however a difference compared to the constant- b case. For a variable b , the threshold value ε_P seems to be independent of the mesh size and is much larger than that of the uniform b test case. This observation limits further the possible choice of coupling strategies, since even for coarse meshes there exists a range of ε -values where neither the Singular Perturbation nor the Limit model are valid. The coupling strategy, involving all three models, remains interesting to investigate.

In the next test case we investigate the influence of the variations of the b field on the accuracy of the solution. We would like to answer the following question: What is the minimal number of points per characteristic length of b variations required to obtain an acceptable solution? For this, let us modify the previous test case. Let $b = B/|B|$, with

$$B = \begin{pmatrix} \alpha(2y-1)\cos(m\pi x) + \pi \\ m\pi\alpha(y^2-y)\sin(m\pi x) \end{pmatrix}, \quad (3.16)$$

m being an integer. The limit solution and ϕ^{ε} are chosen to be

$$\phi^0 = \sin(\pi y + \alpha(y^2 - y)\cos(m\pi x)), \quad (3.17)$$

$$\phi^{\varepsilon} = \sin(\pi y + \alpha(y^2 - y)\cos(m\pi x)) + \varepsilon \cos(2\pi x)\sin(\pi y). \quad (3.18)$$

We perform two tests: first, we fix the mesh size and vary m to find the minimal period of b for which the Asymptotic Preserving method yields still acceptable results. We define a result to be acceptable when the relative error is less than 0.01. In the second test m remains fixed and the convergence of the scheme is studied. The results are presented on Figures 3.4 and 3.5.

For $\varepsilon=1$ and 400 mesh points in each direction ($h=0.0025$) the relative error in the L^2 -norm, defined as $\frac{\|\phi^{\varepsilon} - \phi_A\|_{L^2(\Omega)}}{\|\phi_A\|_{L^2(\Omega)}}$, is below 0.01 for all tested values of $1 \leq m \leq 50$.

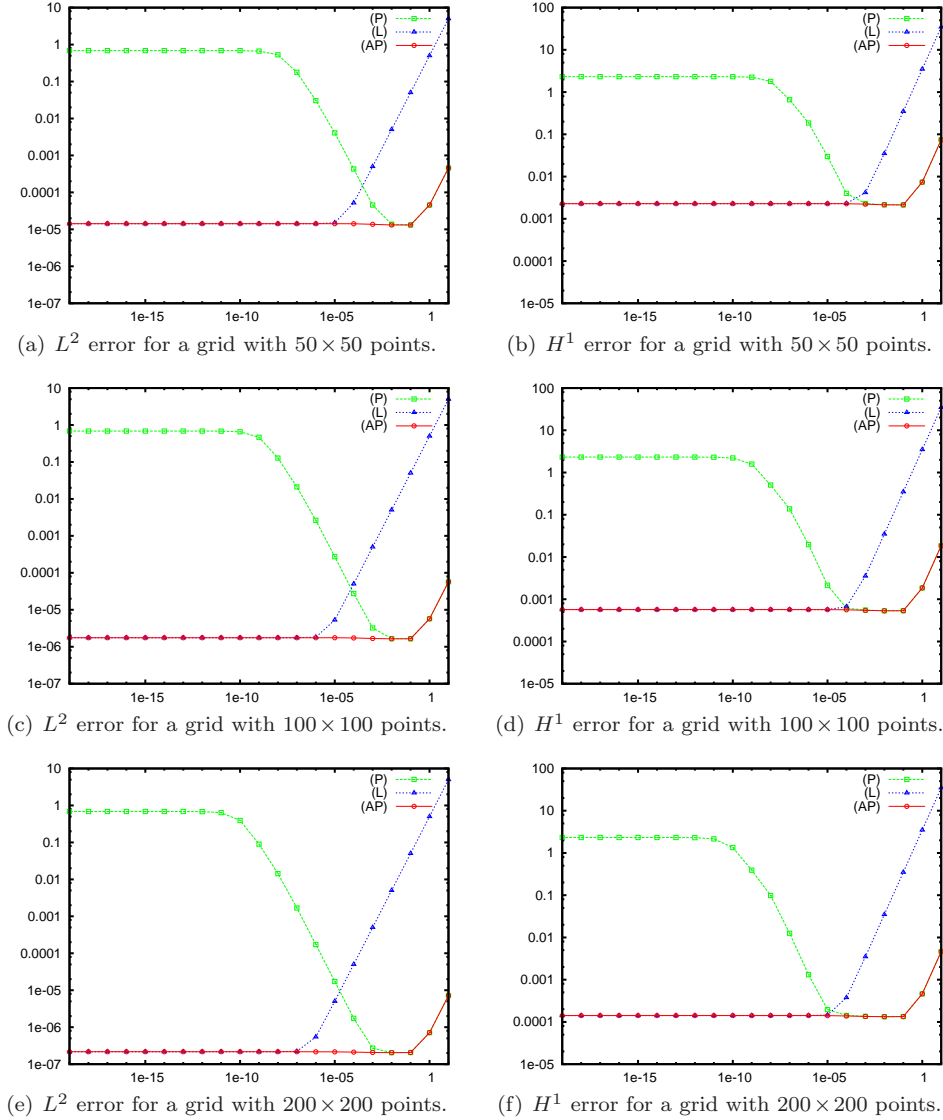


FIGURE 3.3. Absolute L^2 (left column) and H^1 (right column) errors between the exact solution ϕ^ε and the computed solution ϕ_A (AP), ϕ_L (L), ϕ_P (P) for the test case with variable b . Plotted are the errors as a function of the small parameter ε , for three different meshes.

The relative H^1 -error $\frac{\|\phi^\varepsilon - \phi_A\|_{H^1(\Omega)}}{\|\phi_A\|_{H^1(\Omega)}}$ exceeds the critical value for $m > 25$. For $\varepsilon = 10^{-20}$ the maximal m for which the error is acceptable in both norms is 20. The minimal number of mesh points per period of b variations is 40 in the worst case, in order to obtain a 1% relative error.

Figure 3.5 and Table 3.4 show the convergence of the Asymptotic Preserving scheme with respect to h for $m = 10$ and $\varepsilon = 10^{-10}$. We observe that for big values of h the error does not diminish with h . For $h < 0.025$ the scheme converges at a better rate than 2 for H^1 -error and 3 for L^2 -error. For $h < 0.00625$ (160 points) the optimal

convergence rate in the H^1 -norm is obtained (which is 32 mesh points per period of b). The method is super-convergent in the whole tested range for the L^2 -error.

These results are reassuring, as they prove that the Asymptotic Preserving scheme is precise even for strongly varying fields for relatively small mesh sizes, which was not evident. Indeed, the optimal convergence rate in the H^1 -norm is obtained for 32 mesh points per b period, and an 1% relative error for 40 points. It shows that accurate results can be obtained in more complex simulations, such as tokamak plasma. The application of the method to bigger scale problems is the subject of an ongoing work.

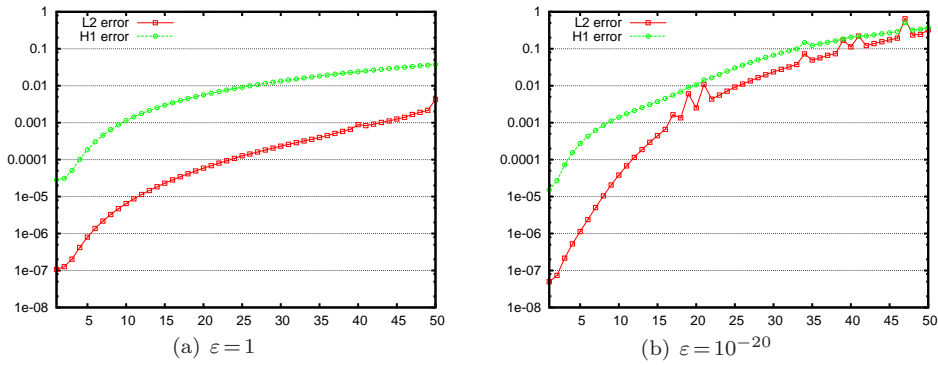


FIGURE 3.4. Relative L^2 and H^1 errors between the exact solution ϕ^ε and the computed solution ϕ_A (AP) for $h=0.0025$ (400 points in each direction) as a function of m and for $\varepsilon=1$ respectively 10^{-20} .

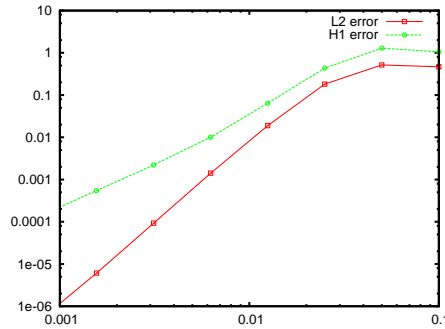


FIGURE 3.5. Relative L^2 and H^1 errors between the exact solution ϕ^ε and the computed solution ϕ_A (AP) for $m=10$ and $\varepsilon=10^{-10}$ as a function of h .

3.3.3. 3D test case, uniform and aligned b -field. Finally, we test our method on a simple 3D case. Let the field b be aligned with the X -axis:

$$b = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}. \quad (3.19)$$

h	# points per period	L^2 -error	H^1 -error
0.1	2	4.7×10^{-1}	1.05
0.05	4	5.2×10^{-1}	1.29
0.025	8	1.82×10^{-1}	4.3×10^{-1}
0.0125	16	1.89×10^{-2}	6.4×10^{-2}
0.00625	32	1.41×10^{-3}	1.00×10^{-2}
0.0003125	64	9.3×10^{-5}	2.21×10^{-3}
0.0015625	128	6.1×10^{-6}	5.5×10^{-4}
0.00078125	256	4.6×10^{-7}	1.36×10^{-4}

TABLE 3.4. Relative L^2 and H^1 errors between the exact solution ϕ^ε and the computed solution ϕ_A (AP) for $m=10$ and $\varepsilon=10^{-10}$ as a function of h .

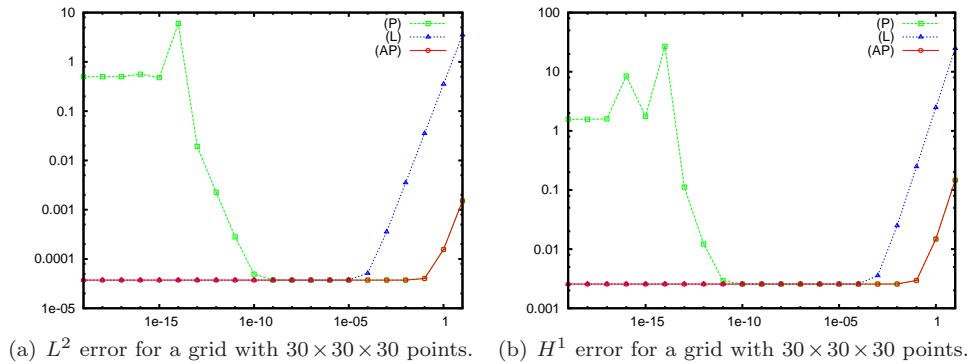


FIGURE 3.6. Absolute L^2 (left column) and H^1 (right column) errors between the exact solution ϕ^ε and the computed solution ϕ_A (AP), ϕ_L (L), ϕ_P (P) for the 3D test case. The errors are plotted as a function of the anisotropy ratio ε .

Let $\Omega = [0, 1] \times [0, 1] \times [0, 1]$, and let the source term f be such that the solution is given by

$$\phi^\varepsilon = \sin(\pi y) \sin(\pi z) + \varepsilon \cos(2\pi x) \sin(\pi y) \sin(\pi z), \quad (3.20)$$

$$p^\varepsilon = \sin(\pi y) \sin(\pi z), \quad q^\varepsilon = \varepsilon \cos(2\pi x) \sin(\pi y) \sin(\pi z). \quad (3.21)$$

Numerical simulations were performed on a $30 \times 30 \times 30$ grid. Once again all three methods are compared. The L^2 and H^1 -errors are given on Figure 3.6. The numerical results are equivalent with those obtained in the 2D test with constant b . Note that it is difficult to perform 3D simulations with more refined grids due to memory requirements on standard desktop equipment. Every row in the matrix constructed for the Singular Perturbation model can contain up to 125 non-zero entries (for \mathbb{Q}_2 finite elements), while matrices associated with the Asymptotic Preserving reformulation have rows with up to 375 non-zero entries. Furthermore the dimension of the latter is five times bigger. The memory requirements of the direct solver used in our simulations grow rapidly. The remedy could be to use an iterative solver with suitable preconditioner. Finding the most efficient method to invert these matrices is however

beyond the scope of this paper. In future work we will address this problem as well as a parallelization of this method.

4. Conclusions

The Asymptotic Preserving method presented in this paper is shown to be very efficient for the solution of highly anisotropic elliptic equations, where the anisotropy direction is given by an arbitrary but smooth vector field b with non-adapted coordinates and meshes. The results presented here generalize the procedure used in [12] and have the important advantage of permitting the use of Cartesian grids which are independently of the shape of the anisotropy. Moreover, the scheme is equally accurate, independently of the anisotropy strength, avoiding thus the use of coupling methods. The numerical study of this AP-scheme shall be investigated in a forthcoming paper, in particular the ε -independent convergence results shall be stated.

Another important related work consists in extending our methods to the case of anisotropy ratios ε , which are variable in Ω from moderate to very small values. This is important, for example, in plasma physics simulations as already noted in the introduction. An alternative strategy to the Asymptotic Preserving schemes would be to couple a standard discretization in subregions with moderate ε with a limit ($\varepsilon \rightarrow 0$) model in subregions with small ε as suggested, for example, in [5, 28]. However, the limit model is only valid for $\varepsilon \ll 1$ and cannot be applied for weak anisotropies. Thus, the coupling strategy requires existence of a range of anisotropy strength where both methods are valid. This is rather undesirable since this range may not exist at all, as illustrated by our results in Figure 3.1.

Appendix A. Decompositions $\mathcal{V} = \mathcal{G} \oplus^\perp \mathcal{A}$, $\tilde{\mathcal{V}} = \tilde{\mathcal{G}} \oplus \mathcal{L}$ and related estimates. We shall show in this Appendix that all the statements in Hypotheses B and B' can be rigorously derived under some assumptions on the domain boundary $\partial\Omega$ and on the manner in which it is intersected by the field b . We assume essentially that b is tangential to $\partial\Omega$ on $\partial\Omega_D$ and that b penetrates the remaining part of the boundary $\partial\Omega_N$ at an angle that stays away from 0 on $\partial\Omega_N$. This assumption is indeed essential as shows the following counter-example:

EXAMPLE A.1. Consider a two dimensional domain Ω of the form

$$\Omega = \{(x, y) : 0 < x < 1, 0 < y < L(x)\},$$

where $L(x) = 1 + \sqrt[4]{x(1-x)}$, and take $b = (0, 1)$ so that $\partial\Omega_D$ is the part of the boundary consisting of $\{x = 0, 0 < y < 1\}$ and $\{x = 1, 0 < y < 1\}$. Note that the angle of penetration of b in Ω is arbitrarily close to 0 on $\partial\Omega_N$, near its junction with $\partial\Omega_D$. Consider the function $u \in \mathcal{V}$ given by

$$u(x, y) = \begin{cases} (y-1), & \text{if } y > 1, \\ 0, & \text{otherwise.} \end{cases}$$

We have

$$\|\nabla u\|_{L^2}^2 = \left\| \frac{\partial u}{\partial y} \right\|_{L^2}^2 = \int_0^1 (L(x) - 1) dx < \infty,$$

hence $u \in \mathcal{V}$. Let $p = Pu$ be the orthogonal (with respect to the L^2 scalar product) projection of u on \mathcal{G} :

$$p = p(x) = \frac{(L(x) - 1)^2}{2L(x)}.$$

Computing the norm of its derivatives gives

$$\|\nabla p\|_{L^2}^2 = \left\| \frac{\partial p}{\partial x} \right\|_{L^2}^2 = \int_0^1 \left(\frac{d}{dx} \left(\frac{(L(x)-1)^2}{2L(x)} \right) \right)^2 L(x) dx = \infty$$

since the expression inside the last integral is equal to $\frac{1}{16x} + O(1)$ as $x \rightarrow 0$. We have thus a function $u \in \mathcal{V}$ such that $Pu \notin \mathcal{V}$. Hence, the property $\mathcal{V} = \mathcal{G} \oplus^\perp \mathcal{A}$ does not hold for such a domain.

To prove $\mathcal{V} = \mathcal{G} \oplus^\perp \mathcal{A}$, we assume also that $\partial\Omega_N$ consists of two disjoint components for which there exist global and smooth parametrizations. This last assumption can be weakened (existence of an atlas of local smooth parametrizations should be sufficient) at the expense of lengthening the proofs. The precise set of our assumptions is assembled in Hypothesis C below. We then proceed by proving that this hypothesis implies $\mathcal{V} = \mathcal{G} \oplus^\perp \mathcal{A}$ and also $\tilde{\mathcal{V}} = \tilde{\mathcal{G}} \oplus \mathcal{L}$.

Hypothesis C. The boundary of Ω is the union of three components: $\partial\Omega_D$ where $b \cdot n = 0$, $\partial\Omega_{in}$ where $b \cdot n \leq -\alpha$ and $\partial\Omega_{out}$ where $b \cdot n \geq \alpha$ with some constant $\alpha > 0$. Moreover, there is a smooth system of coordinates ξ_1, \dots, ξ_{d-1} on $\partial\Omega_{in}$, meaning that there is a bounded domain $\Gamma_{in} \in \mathbb{R}^{d-1}$ and a one-to-one map $h_{in} : \Gamma_{in} \rightarrow \mathbb{R}^d$ such that $h_{in} \in C^2(\bar{\Gamma}_{in})$ and $\partial\Omega_{in}$ is the image of $h_{in}(\xi_1, \dots, \xi_{d-1})$ as $(\xi_1, \dots, \xi_{d-1})$ goes over Γ_{in} . The matrix formed by the vectors $(\partial h_{in} / \partial \xi_1, \dots, \partial h_{in} / \partial \xi_{d-1}, n)$ is invertible for all $(\xi_1, \dots, \xi_{d-1}) \in \bar{\Gamma}_{in}$. Similar assumptions hold also for $\partial\Omega_{out}$ (changing Γ_{in} to Γ_{out} and h_{in} to h_{out}).

Using this hypothesis we can introduce a system of coordinates in Ω such that the field lines of b coincide with the coordinate lines. To do this consider the initial value problem for a parametrized ordinary differential equation (ODE):

$$\frac{\partial X}{\partial \xi_d}(\xi', \xi_d) = b(X(\xi', \xi_d)), \quad X(\xi', 0) = h_{in}(\xi'). \quad (\text{A.1})$$

Here $X(\xi', \xi_d)$ is \mathbb{R}^d -valued and ξ' stands for $(\xi_1, \dots, \xi_{d-1})$. For any fixed $\xi' \in \Gamma_{in}$, equation (A.1) should be understood as an ODE for a function of ξ_d . Its solution $X(\xi', \xi_d)$ goes then over the field line of b starting (as $\xi_d = 0$) at the point on the inflow boundary $\partial\Omega_{in}$, parametrized by ξ' . This field line hits the outflow boundary $\partial\Omega_{out}$ somewhere. In other words, for any $\xi' \in \Gamma_{in}$ there exists $L(\xi') > 0$ such that $X(\xi', L(\xi')) \in \partial\Omega_{out}$. The domain of definition of X is thus

$$D = \{(\xi', \xi_d) \in \mathbb{R}^d / \xi' \in \Gamma_{in} \text{ and } 0 < \xi_d < L(\xi')\}.$$

Gathering the results on parametrized ODEs, from for instance [40], we conclude that $X(\xi', \xi_d) = X(\xi_1, \dots, \xi_d)$ is a smooth function of all its d parameters, or more precisely $X \in C^2(\bar{D})$. Evidently, the map X is one-to-one from \bar{D} to $\bar{\Omega}$ and thus ξ_1, \dots, ξ_d provides a system of coordinates for $\bar{\Omega}$. Moreover this system is not degenerate in the sense that the vectors $\partial X / \partial \xi_1, \dots, \partial X / \partial \xi_d$ are linearly independent at each point of $\bar{\Omega}$. Indeed, if this was not the case, then there would exist a non trivial linear combination $\lambda_1 \partial X / \partial \xi_1 + \dots + \lambda_d \partial X / \partial \xi_d$ that would vanish at some point in $\bar{\Omega}$. But, ODE (A.1) implies

$$\frac{\partial}{\partial \xi_d} \sum_{i=1}^d \lambda_i \frac{\partial X}{\partial \xi_i}(\xi', \xi_d) = \nabla b(X(\xi', \xi_d)) \cdot \sum_{i=1}^d \lambda_i \frac{\partial X}{\partial \xi_i}(\xi', \xi_d),$$

so that the unique solution of this ODE, i.e. the linear combination $\sum_{i=1}^d \lambda_i \frac{\partial X}{\partial \xi_i}$, would vanish on the whole field line, in particular on the inflow. But this is impossible since $\frac{\partial X}{\partial \xi_i} = \frac{\partial h_{in}}{\partial \xi_i}$, $i=1, \dots, d-1$ on the inflow, while $\frac{\partial X}{\partial \xi_d} = b$ and the vectors $\left(\frac{\partial h_{in}}{\partial \xi_1}, \dots, \frac{\partial h_{in}}{\partial \xi_{d-1}}, b\right)$ are linearly independent for all $(\xi_1, \dots, \xi_{d-1}) \in \bar{\Gamma}_{in}$. We see thus that the Jacobian $J = \det(\partial X_j / \partial \xi_i)$ does not vanish on $\bar{\Omega}$, and we can assume that $m < J < M$ everywhere on $\bar{\Omega}$ with some positive constants m and M (assuming that J is positive does not harm the generality since if J is negative in Ω than one can replace ξ_1 by $-\xi_1$). Since $X \in C^2(\bar{\Omega})$, we have also that $J \in C^1(\bar{\Omega})$.

One also sees easily that the top of D is given by a smooth function $L(\xi')$. Indeed, $L(\xi')$ is determined for each $\xi' \in \Gamma_{in}$ from the equation $X(\xi', L(\xi')) = h_{out}(\eta)$ with some $\eta = (\eta_1, \dots, \eta_{d-1}) \in \Gamma_{out}$. We know already that this equation is solvable for $\xi_d = L(\xi')$, $\eta_1, \dots, \eta_{d-1}$ for any $\xi' \in \Gamma_{in}$. To conclude that the solution depends smoothly on ξ' we can apply the implicit function theorem to the equation

$$F(\xi'; \xi_d, \eta_1, \dots, \eta_{d-1}) = X(\xi', \xi_d) - h_{out}(\eta_1, \dots, \eta_{d-1}) = 0.$$

Indeed, the R^d -valued function F is smooth and the matrix of its partial derivatives with respect to $\xi_d, \eta_1, \dots, \eta_{d-1}$ is invertible, since $\partial F / \partial \xi_d = b$ and $\partial F / \partial \eta_i = -\partial h_{out} / \partial \eta_i$ at some point at the outflow and the vectors $\partial h_{out} / \partial \eta_i$ lie in the tangent plane to $\partial \Omega_{out}$ while b is nowhere in this plane. We have moreover that $L \in C^1(\bar{\Gamma}_{in})$. Indeed, we can prove that all the derivatives of L are bounded. In order to do it, let us remark that the differential of $X(\xi', L(\xi'))$ represents a vector in the tangent plane at some point on $\partial \Omega_{out}$ so that it is perpendicular to the outward normal n . We have thus for any $i=1, \dots, d-1$

$$0 = n \cdot \left(\frac{\partial X}{\partial \xi_i}(\xi', L(\xi')) + \frac{\partial X}{\partial \xi_d}(\xi', L(\xi')) \frac{\partial L}{\partial \xi_i}(\xi') \right) = n \cdot \left(\frac{\partial X}{\partial \xi_i}(\xi', L(\xi')) + b(\xi', L(\xi')) \frac{\partial L}{\partial \xi_i}(\xi') \right)$$

so that

$$\frac{\partial L}{\partial \xi_i}(\xi') = - \frac{n \cdot \frac{\partial X}{\partial \xi_i}(\xi', L(\xi'))}{n \cdot b(\xi', L(\xi'))},$$

which is bounded since X has bounded partial derivatives and $n \cdot b \geq \alpha$ by hypothesis. Note also that L is strictly positive.

- We can now establish the decomposition $\mathcal{V} = \mathcal{G} \oplus^\perp \mathcal{A}$. Take any $\phi \in \mathcal{V} \cap C^1(\bar{\Omega})$ and introduce $p \in L^2(\Omega)$ by

$$p(x) = p(\xi', \xi_d) = p(\xi') = \frac{\int_0^{L(\xi')} \phi(\xi', t) J(\xi', t) dt}{\int_0^{L(\xi')} J(\xi', t) dt}. \quad (\text{A.2})$$

From now on we switch back and forth between the Cartesian coordinates $x = (x_1, \dots, x_d)$ and the new ones $(\xi', \xi_d) = (\xi_1, \dots, \xi_d)$. Evidently, p is constant along each field line. Moreover, p is the L^2 -orthogonal projection of ϕ on the space of such functions. Indeed, if $\psi = \psi(\xi') \in L^2(\Omega)$ is any function constant along each field line, then

$$\begin{aligned} \int_{\Omega} p \psi dx &= \int_D p \psi J d\xi = \int_{\Gamma_{in}} p(\xi') \psi(\xi') \int_0^{L(\xi')} J(\xi', \xi_d) d\xi_d d\xi' \\ &= \int_{\Gamma_{in}} \int_0^{L(\xi')} \phi(\xi', \xi_d) \psi(\xi') J(\xi', \xi_d) d\xi_d d\xi' = \int_{\Omega} \phi \psi dx. \end{aligned}$$

Let us prove that $p \in \mathcal{V}$, i.e. that its derivatives are square integrable. The change of variable $t = L(\xi')s$ yields the function

$$p(\xi') = \frac{\int_0^1 \phi(\xi', L(\xi')s) J(\xi', L(\xi')s) ds}{\int_0^1 J(\xi', L(\xi')s) ds}.$$

Now we have $\partial p / \partial \xi_d = 0$ and for all $\partial p / \partial \xi_{i, i=1, \dots, d-1}$, denoting $a = a(\xi') = (\int_0^1 J(\xi', L(\xi')s) ds)^{-1}$, $\phi = \phi(\xi', L(\xi')s)$, and the same for J , we obtain

$$\begin{aligned} \frac{\partial p}{\partial \xi_i} &= \frac{\partial a}{\partial \xi_i} \int_0^1 \phi J ds + a \int_0^1 \frac{\partial \phi}{\partial \xi_i} J ds + a \int_0^1 \frac{\partial \phi}{\partial \xi_d} \frac{\partial L}{\partial \xi_i} s J ds \\ &\quad + a \int_0^1 \phi \frac{\partial J}{\partial \xi_i} ds + a \int_0^1 \phi \frac{\partial J}{\partial \xi_d} \frac{\partial L}{\partial \xi_i} s ds. \end{aligned} \quad (\text{A.3})$$

Using all the previous bounds on the functions L and J and skipping the details of somewhat tedious calculations, we arrive at

$$\begin{aligned} \int_{\Omega} \left(\frac{\partial p}{\partial \xi_i} \right)^2 dx &= \int_{\Gamma_{in}} \int_0^{L(\xi')} \left(\frac{\partial p}{\partial \xi_i} \right)^2 J d\xi_d d\xi' \\ &\leq C \int_{\Gamma_{in}} \int_0^{L(\xi')} \left(\phi^2 + \left(\frac{\partial \phi}{\partial \xi_i} \right)^2 + \left(\frac{\partial \phi}{\partial \xi_d} \right)^2 \right) J d\xi_d d\xi', \end{aligned}$$

implying

$$\left\| \frac{\partial p}{\partial \xi_i} \right\|_{L^2(\Omega)}^2 \leq C \left(\|\phi\|_{L^2(\Omega)}^2 + \left\| \frac{\partial \phi}{\partial \xi_i} \right\|_{L^2(\Omega)}^2 + \left\| \frac{\partial \phi}{\partial \xi_d} \right\|_{L^2(\Omega)}^2 \right) \leq C \|\phi\|_{H^1(\Omega)}^2.$$

Thus $p \in H^1(\Omega)$, hence $p \in \mathcal{G}$ and $q = \phi - p \in \mathcal{A}$. Since the dependence of p on ϕ is continuous in the norm of $H^1(\Omega)$, a density argument shows that the decomposition $\phi = p + q$ with $p \in \mathcal{G}$ and $q \in \mathcal{A}$ exists for any $\phi \in \mathcal{V}$.

- Let us now introduce the operator P as the L^2 -orthogonal projector on \mathcal{G} , i.e.

$$P: \mathcal{V} \rightarrow \mathcal{G}, \quad \phi \in \mathcal{V} \mapsto P\phi \in \mathcal{G} \quad \text{given by} \quad (\text{A.2}).$$

The estimates in the preceding paragraph show that the operator P is continuous in the norm of $H^1(\Omega)$:

$$\|\nabla_{\perp}(P\phi)\|_{L^2(\Omega)} \leq C \|\nabla \phi\|_{L^2(\Omega)}, \quad \forall \phi \in \mathcal{V}. \quad (\text{A.4})$$

- We have also the following Poincaré-Wirtinger inequality:

$$\|\phi - P\phi\|_{L^2(\Omega)} \leq C \|\nabla_{\parallel} \phi\|_{L^2(\Omega)}, \quad \forall \phi \in \mathcal{V}. \quad (\text{A.5})$$

To prove this, it is sufficient to establish that $\|q\|_{L^2(\Omega)} \leq C \|\nabla_{\parallel} q\|_{L^2(\Omega)}$ for all $q \in \mathcal{A}$. We observe that

$$\|q\|_{L^2(\Omega)}^2 = \int_{\Gamma_{in}} \int_0^{L(\xi')} q^2(\xi', \xi_d) J(\xi', \xi_d) d\xi_d d\xi'$$

and

$$\|\nabla\|\phi\|_{L^2(\Omega)}^2 = \int_{\Gamma_{in}} \int_0^{L(\xi')} \left(\frac{\partial q}{\partial \xi_d} \right)^2 (\xi', \xi_d) J(\xi', \xi_d) d\xi_d d\xi'.$$

The requirement $q \in \mathcal{A}$ is equivalent to

$$\int_0^{L(\xi')} q(\xi', \xi_d) J(\xi', \xi_d) d\xi_d = 0 \quad \text{f.a.a. } \xi' \in \Gamma_{in}. \quad (\text{A.6})$$

We have thus to prove, for every ξ' ,

$$\int_0^{L(\xi')} q^2(\xi', \xi_d) J(\xi', \xi_d) d\xi_d \leq C^2 \int_0^{L(\xi')} \left(\frac{\partial q}{\partial \xi_d} \right)^2 (\xi', \xi_d) J(\xi', \xi_d) d\xi_d$$

provided (A.6). Fixing any ξ' , making the change of integration variable $\xi_d = L(\xi')t$ and introducing the functions $u(t) = q(\xi', L(\xi')t) J(\xi', L(\xi')t)$ and $J(t) = J(\xi', L(\xi')t)$, we rewrite the last inequality as

$$\int_0^1 \frac{u^2(t)}{J(t)} dt \leq \frac{C^2}{L^2(\xi')} \int_0^1 \left(\frac{u'(t)}{J(t)} - \frac{u(t)}{J^2(t)} J'(t) \right)^2 J(t) dt. \quad (\text{A.7})$$

Since $\int_0^1 u(t) dt = 0$, we have by the standard Poincaré inequality

$$\int_0^1 u^2(t) dt \leq C_P^2 \int_0^1 (u'(t))^2 dt. \quad (\text{A.8})$$

- Let us turn to the verification of Hypothesis B'. Take any $u \in \tilde{\mathcal{V}}$. We want to prove that one can decompose $u = p + q$ with $p \in \tilde{\mathcal{G}}$ and $q \in \mathcal{L}$ and the trace of u on $\partial\Omega_{in}$ (denoted g) is in $L^2(\partial\Omega_{in})$. In the ξ -coordinates we can write a surface element of $\partial\Omega_{in}$ as $d\sigma = S(\xi') d\xi'$ with a function S smoothly depending on ξ' . We see now that for u sufficiently smooth,

$$\begin{aligned} \|g\|_{L^2(\partial\Omega_{in})}^2 &= \int_{\Gamma_{in}} g^2(\xi') S(\xi') d\xi' \\ &\leq C \int_0^1 \int_{\Gamma_{in}} \left[u^2(\xi', L(\xi')s) + \frac{1}{L(\xi')} \left(\frac{\partial u}{\partial \xi_d} \right)^2 (\xi', L(\xi')s) \right] S(\xi') d\xi' ds \\ &\quad (\text{by a one-dimensional trace inequality}) \\ &\leq C \|u\|_{\tilde{\mathcal{V}}}^2. \end{aligned}$$

By density, the trace g is thus defined for any $u \in \tilde{\mathcal{V}}$ with $\|g\|_{L^2(\partial\Omega_{in})} \leq C \|u\|_{\tilde{\mathcal{V}}}$. Taking $p = p(\xi') = g(\xi')$, we observe by a similar calculation that $\|p\|_{L^2(\Omega)} \leq C \|u\|_{\tilde{\mathcal{V}}}$, so that $p \in \tilde{\mathcal{G}}$. By definition $q = u - p \in \mathcal{L}$.

Appendix B. On the choice of the finite element space \mathcal{L}_h . Let Ω be the rectangle $(0, L_x) \times (0, L_y)$ and the anisotropy direction be constant and aligned with the y -axis: $b = (0, 1)$. Let us use the \mathbb{Q}_k finite elements on a Cartesian grid, i.e. take some basis function $\theta_{x_i}(x)$, $i = 0, \dots, N_x$ and $\theta_{y_j}(y)$, $j = 0, \dots, N_y$ and define the complete finite element space X_h (without any restrictions on the boundary) as

$\text{span}\{\theta_{x_i}(x)\theta_{y_j}(y) \mid 0 \leq i \leq N_x, 0 \leq j \leq N_y\}$. The following subspace is then used for the approximation of the unknowns $p, q, l \in \mathcal{V}$:

$$\mathcal{V}_h = \{v_h \in X_h / v_h|_{\partial\Omega_D} = 0\}.$$

We want to prove that taking for the approximation of $\lambda, \mu \in \mathcal{L}$, the space \mathcal{L}_h under the form

$$\mathcal{L}_h = \{\lambda_h \in X_h / \lambda_h|_{\partial\Omega_{in}} = 0\}, \quad (\text{A.1})$$

leads to an ill posed problem (3.1).

Claim. There exist $\lambda_h \in \mathcal{L}_h$, $\lambda_h \neq 0$ such that $a_{||}(\lambda_h, p_h) = 0$ for all $p_h \in \mathcal{V}_h$. In fact there are exactly $2N_y$ linearly independent functions having this property.

REMARK B.1. In the continuous case, the equation

$$a_{||}(p, \lambda) = 0, \quad \forall p \in \mathcal{V},$$

implies $\lambda = 0$ by density arguments. These density arguments are lost when discretizing the spaces \mathcal{V} and \mathcal{L} .

Proof of the Claim. We can suppose that the basis functions $\theta_{ij}(x, y) := \theta_{x_i}(x)\theta_{y_j}(y)$ are enumerated, so that $\theta_{ij}(0, y) = 0$ for all $i \geq 1$ and $\theta_{0j}(0, y) \neq 0$. Hence for all $p_h = \sum p_{ij}\theta_{ij} \in \mathcal{V}_h$ the coefficients satisfy $p_{0j} = 0$, since the part of the boundary $\{x = 0\}$ is in $\partial\Omega_D$. Let $M = (m_{ik})_{0 \leq i, k \leq N_x}$ be the mass matrix in the x -direction: $m_{ik} = \int \theta_{x_i}(x)\theta_{x_k}(x)dx$. This matrix is invertible, hence there is a vector $a \in \mathbb{R}^{N_x+1}$ that solves $Ma = e$ with $e \in \mathbb{R}^{N_x+1}$, $e = (1, 0, \dots, 0)^t$. Take any fixed integer j , $1 \leq j \leq N_y$, and define $\lambda_h \in \mathcal{L}_h$ as $\lambda_h = \sum a_i \theta_{ij}$. Then for all $p_h = \sum p_{kl}\theta_{kl} \in \mathcal{V}_h$ we have

$$\begin{aligned} a_{||}(\lambda_h, p_h) &= \sum_{i,k,l} a_i p_{kl} \int_{\Omega} \frac{\partial \theta_{ij}}{\partial y} \frac{\partial \theta_{kl}}{\partial y} dx dy \\ &= \sum_{i,k,l} a_i p_{kl} \int_0^{L_x} \theta_{x_i}(x)\theta_{x_k}(x) dx \int_0^{L_y} \theta'_{y_j}(y)\theta'_{y_l}(y) dy \\ &= \sum_{k,l} \delta_{k0} p_{kl} \int_0^{L_y} \theta'_{y_j}(y)\theta'_{y_l}(y) dy = 0. \end{aligned}$$

As we can do this for all (i, j) , $i = 0, 1 \leq j \leq N_y$, and in the same manner for all (i, j) , $i = N_x, 1 \leq j \leq N_y$, there are $2N_y$ linearly independent functions with the property $a_{||}(\lambda_h, p_h) = 0$ for all $p_h \in \mathcal{V}_h$.

We see now that the system (3.1) with zero right hand side $f = 0$ possesses non-zero solutions $(p_h^\varepsilon, \lambda_h^\varepsilon, q_h^\varepsilon, l_h^\varepsilon, \mu_h^\varepsilon) = (0, \lambda_h^j, 0, 0, 0)$, where λ_h^j is any of the functions constructed in the preceding paragraph. It means that (3.1) is ill-posed, i.e. the corresponding matrix is singular.

Acknowledgment. This work has been supported by the Marie Curie Actions of the European Commission under the contract DEASE (MEST-CT-2005-021122), by the French 'Commissariat à l'Énergie Atomique (CEA)' under contracts ELMAG (CEA-Cesta 4600156289) and GYRO-AP (Euratom-CEA V 3629.001), by the Agence Nationale de la Recherche (ANR) under contract IODISEE (ANR-09-COSI-007-02), by the 'Fondation Sciences et Technologies pour l'Aéronautique et l'Espace (STAE)'

under contract PLASMAX (RTRA-STAE/2007/PF/002), and by the Scientific Council of the Université Paul Sabatier, under contract MOSITER. Support from the French magnetic fusion programme 'fédération de recherche sur la fusion par confinement magnétique' is also acknowledged. The authors would like to express their gratitude to G. Gallice and C. Tessieras from CEA-Cesta for bringing their attention to this problem, and to G. Falchetto, X. Garbet, and M. Ottaviani from CEA-Cadarache, for their constant support of this research programme.

REFERENCES

- [1] J. Adam, J. Boeuf, N. Dubuit, M. Dudeck, L. Garrigues, D. Gresillon, A. Heron, G. Hagelaar, V. Kulaev, and N. Lemoine, et al., *Physics, simulation and diagnostics of Hall effect thrusters*, Plasma Physics and Controlled Fusion, 50, 124041, 2008.
- [2] P.R. Amestoy, I.S. Duff, J.Y. L'Excellent, and J. Koster, *A fully asynchronous multifrontal solver using distributed dynamic scheduling*, SIAM J. Matrix Anal. Appl., 23(1), 15–41, 2001.
- [3] S.F. Ashby, W.J. Bosl, R.D. Falgout, S.G. Smith, A.F. Tompson, and T.J. Williams, *A numerical simulation of groundwater flow and contaminant transport on the CRAY T3D and C90 supercomputers*, Int. J. High Performance Computing Applications, 13(1), 80–93, 1999.
- [4] M. Beer, S. Cowley, and G. Hammett, *Field-aligned coordinates for nonlinear simulations of tokamak turbulence*, Physics of Plasmas, 2(7), 2687, 1995.
- [5] C. Besse, P. Degond, F. Deluzet, J. Claudel, G. Gallice, and C. Tessieras, *A model hierarchy for ionospheric plasma modeling*, Math. Models Methods Appl. Sci., 14(3), 393–415, 2004.
- [6] C. Besse, F. Deluzet, C. Negulescu, and C. Yang, *Efficient numerical methods for strongly anisotropic elliptic equations*, submitted to SIAM Multiscale Modeling and Simulation.
- [7] A.H. Boozer, *Establishment of magnetic coordinates for a given magnetic field*, Physics of Fluids, 25(3), 520–521, 1982.
- [8] S. Brull, P. Degond, and F. Deluzet, *Degenerate anisotropic elliptic problems and magnetized plasma simulations*, Comm. Comp. Phys. (CICP), to appear. <http://arxiv.org/abs/1010.5968>.
- [9] L. Chacón, *An optimal, parallel, fully implicit Newton–Krylov solver for three-dimensional viscoresistive magnetohydrodynamics*, Physics of Plasmas, 15, 056103, 2008.
- [10] P. Crispel, P. Degond, and M.H. Vignal, *An asymptotic preserving scheme for the two-fluid Euler–Poisson model in the quasineutral limit*, J. Comput. Phys., 223(1), 208–234, 2007.
- [11] P. Degond, F. Deluzet, L. Navoret, A.B. Sun, and M.H. Vignal, *Asymptotic-preserving particle-in-cell method for the Vlasov–Poisson system near quasineutrality*, J. Comput. Phys., 229(16), 5630–5652, 2010.
- [12] P. Degond, F. Deluzet, and C. Negulescu, *An asymptotic preserving scheme for strongly anisotropic elliptic problems*, Multiscale Model. Simul., 8(2), 645–666, 2009/10.
- [13] P. Degond, F. Deluzet, A. Sangam, and M.H. Vignal, *An asymptotic preserving scheme for the Euler equations in a strong magnetic field*, J. Comput. Phys., 228(10), 3540–3558, 2009.
- [14] F. Deluzet, A. Lozinski, J. Narski, and C. Negulescu, *Error estimates for an asymptotic-preserving method for highly anisotropic diffusion equations*, in preparation.
- [15] W.D. D'haeseleer, W.N.G. Hitchon, J.D. Callen, and J.L. Shohet, *Flux Coordinates and Magnetic Field Structure. A Guide to a Fundamental Tool of Plasma Theory*, Springer Series in Computational Physics, Springer-Verlag, Berlin, 1991.
- [16] A.M. Dimits, *Fluid simulations of tokamak turbulence in quasiballooning coordinates*, Phys. Rev. E, 48(5), 4070–4079, Nov. 1993.
- [17] R. Eymard, T. Gallouët, and R. Herbin, *A new finite volume scheme for anisotropic diffusion problems on general grids: Convergence analysis*, C.R. Math. Acad. Sci. Paris, 344(6), 403–406, 2007.
- [18] M.W. Gee, J.J. Hu, and R.S. Tuminaro, *A new smoothed aggregation multigrid method for anisotropic problems*, Numer. Linear Algebra Appl., 16(1), 19–37, 2009.
- [19] L. Giraud and R.S. Tuminaro, *Schur complement preconditioners for anisotropic problems*, IMA J. Numer. Anal., 19(1), 1–18, 1999.
- [20] S. Gnani, Q. Yu, J. Kruger, and K. Lackner, *Modelling of heat transport in magnetised plasmas using non-aligned coordinates*, J. Comput. Phys., 209(1), 354–370, 2005.
- [21] V. Grandgirard, Y. Sarazin, X. Garbet, G. Dif-Pradalier, P. Ghendrih, N. Crouseilles, G. Latu, E. Sonnendrücker, N. Besse, and P. Bertrand, *Computing ITG turbulence with a full-f*

- semi-lagrangian code*, Comm. Nonlin. Sci. Num. Sim., 13(1), 81–87, 2008.
- [22] S. Hamada, *Hydromagnetic equilibria and their proper coordinates*, Nucl. Fusion, 2, 23–37, 1962.
- [23] T.Y. Hou and X.H. Wu, *A multiscale finite element method for elliptic problems in composite materials and porous media*, J. Comput. Phys., 134(1), 169–189, 1997.
- [24] Y. Igitkhanov, M. Mikhailov, and W. Feneberg, *Fluid description of edge plasma transport in a non-orthogonal coordinate system*, Contributions to Plasma Physics, 34(2-3), 398–403, 1994.
- [25] S. Jin, *Efficient asymptotic-preserving (AP) schemes for some multiscale kinetic equations*, SIAM J. Sci. Comput., 21(2), 441–454, 1999.
- [26] E. Kaveeva and V. Rozhansky, *Poloidal and toroidal flows in tokamak plasma near magnetic islands*, Technical Physics Letters, 30(7), 538–540, 2004.
- [27] M. Kelley, W. Swartz, and J. Makela, *Mid-latitude ionospheric fluctuation spectra due to secondary $E\bar{U}B$ instabilities*, Journal of Atmospheric and Solar-Terrestrial Physics, 66(17), 1559–1565, 2004.
- [28] M. Keskinen, *Nonlinear theory of the $E \times B$ instability with an inhomogeneous electric field*, Journal of Geophysical Research, 89(A6), 3913–3920, 1984.
- [29] M. Keskinen, S. Ossakow, and B. Fejer, *Three-dimensional nonlinear evolution of equatorial ionospheric spread-F bubbles*, Geophys. Res. Lett., 30(16), 1–4, 2003.
- [30] B.N. Khoromskij and G. Wittum, *Robust Schur complement method for strongly anisotropic elliptic equations*, Numer. Linear Algebra Appl., 6(8), 621–653, 1999.
- [31] D. Kuzmin, M.J. Shashkov, and D. Svyatskiy, *A constrained finite element method satisfying the discrete maximum principle for anisotropic diffusion problems*, J. Comput. Phys., 228(9), 3448–3463, 2009.
- [32] C. Le Potier, *Schéma volumes finis pour des opérateurs de diffusion fortement anisotropes sur des maillages non structurés*, C.R. Math. Acad. Sci. Paris, 340(12), 921–926, 2005.
- [33] K. Lipnikov, M. Shashkov, and D. Svyatskiy, *The mimetic finite difference discretization of diffusion problem on unstructured polyhedral meshes*, J. Comput. Phys., 211(2), 473–491, 2006.
- [34] I. Llorente and N. Melson, *Robust multigrid smoothers for three dimensional elliptic equations with strong anisotropies*, ICASE Technical Report: TR-98-37, 1998.
- [35] T. Manku and A. Nathan, *Electrical properties of silicon under nonuniform stress*, J. Appl. Phys., 74(3), 1832–1837, 1993.
- [36] J.M. Melenk, *hp-finite Element Methods for Singular Perturbations*, Lecture Notes in Mathematics, Springer-Verlag, Berlin, 1796, 2002.
- [37] K. Miyamoto, *Controlled Fusion and Plasma Physics*, Chapman & Hall, 2007.
- [38] Y. Notay, *An Aggregation-based Algebraic Multigrid Method*, Report GANMN 08-02, Université Libre de Bruxelles, Brussels, Belgium, 2008.
- [39] M.A. Ottaviani, *An alternative approach to field-aligned coordinates for plasma turbulence simulations*, Physics Letters A, 375(15), 1677–1685, 2011.
- [40] L.S. Pontryagin, *Ordinary Differential Equations*, Addison-Wesley Publishing Co., Inc., Reading, Mass.-Palo Alto, Calif.-London, 1962.
- [41] P. Sharma and G.W. Hammett, *Preserving monotonicity in anisotropic diffusion*, J. Comput. Phys., 227, 123–142, November 2007.
- [42] D. Stern, *Geomagnetic Euler potentials*, J. Geophys. Res., 72(15), 3995–4005, 1967.
- [43] A.M. Tréguier, *Modélisation numérique pour l’océanographie physique*, Ann. Math. Blaise Pascal, 9(2), 345–361, 2002.
- [44] W.W. Wang and X.C. Feng, *Anisotropic diffusion with nonlinear structure tensor*, Multiscale Model. Simul., 7(2), 963–977, 2008.
- [45] J. Weickert, *Anisotropic Diffusion in Image Processing*, European Consortium for Mathematics in Industry. B. G. Teubner, Stuttgart, 1998.