

ASYMPTOTIC ANALYSIS OF UPWIND DISCONTINUOUS GALERKIN APPROXIMATION OF THE RADIATIVE TRANSPORT EQUATION IN THE DIFFUSIVE LIMIT*

JEAN-LUC GUERMOND[†] AND GUIDO KANSCHAT[‡]

Abstract. We revisit some results from M. L. Adams [*Nucl. Sci. Engrg.*, 137 (2001), pp. 298–333]. Using functional analytic tools we prove that a necessary and sufficient condition for the standard upwind discontinuous Galerkin approximation to converge to the correct limit solution in the diffusive regime is that the approximation space contains a linear space of continuous functions, and the restrictions of the functions of this space to each mesh cell contain the linear polynomials. Furthermore, the discrete diffusion limit converges in the Sobolev space H^1 to the continuous one if the boundary data is isotropic. With anisotropic boundary data, a boundary layer occurs, and convergence holds in the broken Sobolev space H^s with $s < \frac{1}{2}$ only.

Key words. finite elements, discontinuous Galerkin, neutron transport, diffusion limit

AMS subject classifications. 65N35, 65N22, 65F05, 35J05

DOI. 10.1137/090746938

1. Introduction. The purpose of the present paper is to analyze the upwind discontinuous Galerkin (DG) approximation of the radiative transport equation in the diffusive limit. We focus our attention on the distinguished limit $\lim_{h \rightarrow 0} \lim_{\varepsilon \rightarrow 0}$, where h is the mesh size and ε is the nondimensional mean free path length. This limit corresponds to approximating the radiative transport equation on meshes that are such that $h \gg \varepsilon$; i.e., the approximate solutions are underresolved with respect to the mean free path length. This problem has been addressed in the groundbreaking paper [1] using formal asymptotic analysis. We revisit [1] using a functional analytic point of view and thereby rigorously justify some of the conclusions in [1]. More precisely we prove that the upwind DG method has the following properties when $h \gg \varepsilon$:

1. In the limit of vanishing mean free path length, the upwind DG method yields a mixed discretization of the diffusion equation with a continuous primal variable (see Theorem 4.4 and Corollary 4.5).
2. If the incoming flux is isotropic, the upwind DG scheme approximates the diffusion solution well if and only if the approximation space contains a space of continuous functions that are at least linear on each mesh cell. The upwind DG solution converges to the diffusion solution in H^1 (there is no boundary layer), as stated in Theorem 5.3.
3. If the incoming flux is not isotropic, a boundary layer effect occurs. Nevertheless, provided the approximation space contains a subspace of continuous

*Received by the editors January 20, 2009; accepted for publication (in revised form) December 21, 2009; published electronically April 2, 2010. This material is based upon work supported by the Department of Homeland Security under agreement 2008-DN-077-ARI018-02, National Science Foundation grants DMS-0713829, DMS-0810387, and CBET-0736202, and is partially supported by award KUS-C1-016-04, made by King Abdullah University of Science and Technology (KAUST).
<http://www.siam.org/journals/sinum/48-1/74693.html>

[†]Department of Mathematics, Texas A&M University, College Station, TX 77843 (guermond@math.tamu.edu). This work was completed while the author was on leave from LIMSI, UPRR 3251 CNRS, BP 133, 91403 Orsay Cedex, France.

[‡]Department of Mathematics, Texas A&M University, College Station, TX 77843 (kanschat@tamu.edu).

functions that are at least linear on each mesh cell, the interior solution approximates the correct limit. See the H^s -estimate in Lemma 5.2, $s \in [0, \frac{1}{2})$, and the L^2 -convergence result in Theorem 5.4.

4. Piecewise constant approximation is not appropriate for radiative transport in optically thick materials, since the approximate solution becomes globally constant; i.e., the upwind DG approximation locks in this case (see Corollary 4.5 and Remark 5.2).

These statements, minus the convergence estimates in Sobolev spaces, have already been made in [1] using heuristic asymptotic arguments.

The paper is organized as follows: The continuous problem and the discrete DG settings together with notation are introduced in section 2. The formal asymptotic analysis of the DG approximation is performed in section 3 under the assumption that the nondimensional mean free path goes to zero. This section reproduces more or less the heuristic arguments from [1, 10]. The material in section 3 is not new, but it makes the paper self-contained and introduces intuitive concepts that are useful to understand the functional analytic arguments of section 4. We perform the asymptotic analysis of the upwind DG approximation in section 4 using functional analysis. The limit discrete problem obtained formally in section 3 and rigorously in section 4 is analyzed in section 5, and its limit as $h \rightarrow 0$ is investigated; various convergence issues are sorted out therein. In particular we investigate a boundary layer effect mentioned in [1, 9]; we characterize this effect by proving new convergence estimates in the fractional Sobolev spaces H^s with $s \in [0, \frac{1}{2})$.

The main results of the paper are Theorem 4.4, Corollary 4.5, Lemma 5.2, and Theorems 5.3 and 5.4. To the best of our knowledge, the results stated in Lemma 5.2 and Theorems 5.3 and 5.4 are original.

2. Setting of the problem.

2.1. The transport equation. Let D be a bounded, open, Lipschitz domain in \mathbb{R}^3 , and let S^2 be the unit sphere in \mathbb{R}^3 . The boundary of D is denoted by ∂D , and the outer unit normal on ∂D is denoted by n . The influx boundary is defined by

$$\Gamma_-(D) := \{(\Omega, x) \in S^2 \times \partial D \mid \Omega \cdot n(x) < 0\}.$$

We consider a scaled version of the transport equation:

$$(2.1) \quad \begin{aligned} \Omega \cdot \nabla \psi(\Omega, x) + \frac{\sigma(x)}{\varepsilon} \psi(\Omega, x) - \left(\frac{\sigma(x)}{\varepsilon} - \varepsilon \sigma_a(x) \right) \bar{\psi}(x) &= \varepsilon q(x) && \text{in } S^2 \times D, \\ \psi(\Omega, x) &= \alpha(\Omega, x) && \text{on } \Gamma_-(D), \end{aligned}$$

where the independent variables (Ω, x) span $S^2 \times D$. The dependent variable $\psi(\Omega, x)$ is referred to as the angular intensity. The averaging operator with respect to the angular variable is defined by

$$(2.2) \quad \bar{\psi}(x) := \frac{1}{4\pi} \int_{S^2} \psi(x, \Omega) \, d\Omega.$$

The data are the source term $q(x)$ and the boundary value $\alpha(\Omega, x)$. We refer to $\sigma(x)$ and $\sigma_a(x)$ as the total cross section and the absorption cross section, respectively. To avoid degeneracy we assume that the total cross section and the absorption cross

section are uniformly positive:

$$(2.3) \quad \inf_{x \in D} \sigma(x) > 0,$$

$$(2.4) \quad \inf_{x \in D} \sigma_a(x) > 0.$$

The parameter ε measures the extent to which the problem is diffusive. The limit $\varepsilon \rightarrow 0$ corresponds to the ratio of the mean free path of the particles to the diameter of D going to zero. In this regime the media becomes optically thick and is dominated by scattering.

The objective of this paper is to study the behavior of the approximation of (2.1) using a standard upwind DG approximation as ε goes to zero. More specifically, we investigate what happens when the nondimensional characteristic mesh size is significantly larger than ε .

2.2. The discrete setting. The approximation setting is obtained by tensoring an approximation space for the space domain D and an approximation space for the unit sphere S^2 .

2.2.1. Space discretization. Let \mathcal{T}_h be a subdivision of D into disjoint (open) cells K such that the closure of D is equal to $\cup_{K \in \mathcal{T}_h} \overline{K}$. The meshes are assumed to be affine to avoid unnecessary technicalities; i.e., D is assumed to be a polyhedron. The diameter of $K \in \mathcal{T}_h$ is denoted by h_K , and we set $h = \max_{K \in \mathcal{T}_h} h_K$. We suppose that we have at hand a family of meshes $\{\mathcal{T}_h\}_{h>0}$ and that this family is uniformly shape-regular. We also assume that the mesh is quasi-uniform; i.e., there is $c > 0$ so that

$$(2.5) \quad ch \leq h_K \leq h \quad \forall K \in \mathcal{T}_h.$$

This hypothesis is used when invoking inverse inequalities. It could be avoided by localizing the inverse estimate arguments, but we shall refrain from doing so to steer clear of unnecessary technicalities.

We denote by \mathcal{F}_h^i the set of interior faces (also called interfaces). For $F \in \mathcal{F}_h^i$, we denote by $K_1(F)$ and $K_2(F)$ the two cells so that $F = \overline{K_1(F)} \cap \overline{K_2(F)}$. For any point x on F we denote by $n_1(x)$ and $n_2(x)$ the unit normal vectors on F that point toward $K_2(F)$ and $K_1(F)$, respectively. Although the ordering of $K_1(F)$ and $K_2(F)$ is arbitrary, nothing that is said hereafter depends on the choice that is made.

The set of faces on ∂D is denoted \mathcal{F}_h^∂ . For any $F \in \mathcal{F}_h^\partial$, the cell $K \in \mathcal{T}_h$ which is such that $F = \overline{K} \cap \partial \Omega$ is denoted $K_1(F)$. For any point x on F , the unit normal vector at x that points outward is denoted $n_1(x)$. Finally, we set $\mathcal{F}_h = \mathcal{F}_h^i \cup \mathcal{F}_h^\partial$.

We define a discontinuous approximation space based on the mesh \mathcal{T}_h as follows:

$$(2.6) \quad \mathcal{D}_h = \{v_h \in L^2(\Omega); \forall K \in \mathcal{T}_h, v_h|_K \in P_K\},$$

where P_K is a finite-dimensional function space on the cell K . We assume henceforth that the derivatives of functions in P_K stay in P_K ; i.e.,

$$(2.7) \quad \partial_i p \in P_K, \quad \forall i = 1, 2, 3, \quad \forall p \in P_K, \quad \forall K \in \mathcal{T}_h.$$

Note that (2.7) holds when P_K is a polynomial space.

We define the subspace of \mathcal{D}_h composed of the functions that are continuous:

$$(2.8) \quad \mathcal{C}_h = \mathcal{D}_h \cap \mathcal{C}^0(\overline{D}).$$

For every continuous function β on ∂D , we set

$$(2.9) \quad \mathcal{C}_{h,\beta} = \{\varphi \in \mathcal{C}_h; \varphi|_{\partial D} = \beta\}.$$

$\mathcal{C}_{h,\beta}$ is an affine space in general and a linear subspace of \mathcal{C}_h when β is zero.

Let $\mathcal{D}(h) = H^1(D) + \mathcal{D}_h$. Since every function v in $\mathcal{D}(h)$ has a two-valued trace on $F \in \mathcal{F}_h^i$, we set

$$(2.10) \quad v_1(x) = \lim_{y \rightarrow x}^{y \in K_1(F)} v(y), \quad v_2(x) = \lim_{y \rightarrow x}^{y \in K_2(F)} v(y), \quad \text{for a.e. } x \in F,$$

$$(2.11) \quad \llbracket v \rrbracket = v_1 - v_2, \quad \{v\} = \frac{1}{2}(v_1 + v_2), \quad \text{a.e. on } F.$$

2.2.2. Angular discretization. The angular discretization is done using finite elements in angle and the Galerkin method. We define a mesh on the unit sphere S^2 , and we denote by \mathcal{S}_h the finite-dimensional space composed of scalar-valued functions that are piecewise polynomial on this mesh. Henceforth, all integrals over S^2 are assumed to be evaluated exactly.

By this choice, energy conservation holds automatically. The only requirement we make on the discrete space \mathcal{S}_h is that it contains the constant function of the angle variable:

$$(2.12) \quad 1 \in \mathcal{S}_h \quad \forall h > 0.$$

As an example, consider the discretization used in [8]: we partition S^2 into a triangular mesh obtained by projecting an icosahedron (or refinements of an icosahedron) onto the sphere. On each spherical triangle, we choose constant shape functions. Since the area of each spherical triangle is analytically computable, we can evaluate all integrals exactly. By taking the same constant on each triangle, condition (2.12) holds trivially.

2.2.3. Discretization of $S^2 \times D$. We construct an approximation space for $L^2(S^2 \times D)$ by tensoring \mathcal{S}_h and \mathcal{D}_h :

$$(2.13) \quad \mathcal{W}_h := \mathcal{S}_h \otimes \mathcal{D}_h \quad \forall h > 0.$$

2.3. The upwind DG approximation. In this section, we construct an approximation of the solution to (2.1) by using the standard upwind DG method [11, 14].

We start by defining the bilinear form

$$(2.14) \quad \mathcal{L}(v, w) = \sum_{K \in \mathcal{T}_h} \int_{S^2 \times K} \left(-v \Omega \cdot \nabla w + \frac{\sigma}{\varepsilon} v w + \left(\varepsilon \sigma_a - \frac{\sigma}{\varepsilon} \right) \bar{v} w \right) d\Omega dx \\ + \sum_{F \in \mathcal{F}_h^i} \int_{S^2 \times F} \Omega \cdot n_1 v^\uparrow \llbracket w \rrbracket d\Omega dx + \sum_{F \in \mathcal{F}_h^\partial} \int_{\{\Omega \cdot n \geq 0\} \times F} \Omega \cdot n v w d\Omega dx,$$

where the quantity v^\uparrow is the so-called upwind flux

$$(2.15) \quad v^\uparrow = \begin{cases} v_1(\Omega, x) & \text{if } \Omega \cdot n_1(x) > 0, \\ v_2(\Omega, x) & \text{if } \Omega \cdot n_1(x) < 0. \end{cases}$$

After integrating by parts, the bilinear form \mathcal{L} can be rewritten in the following form, which will also prove useful:

$$(2.16) \quad \mathcal{L}(v, w) = \sum_{K \in \mathcal{T}_h} \int_{S^2 \times K} \left(w \Omega \cdot \nabla v + \frac{\sigma}{\varepsilon} v w + \left(\varepsilon \sigma_a - \frac{\sigma}{\varepsilon} \right) \bar{v} w \right) d\Omega dx \\ + \sum_{F \in \mathcal{F}_h^i} \int_{S^2 \times F} -\Omega \cdot n_1 \llbracket v \rrbracket w^\downarrow d\Omega dx + \sum_{F \in \mathcal{F}_h^o} \int_{\{\Omega \cdot n \leq 0\} \times F} |\Omega \cdot n| v w d\Omega dx,$$

where the quantity v^\downarrow is the downwind flux

$$(2.17) \quad v^\downarrow = \begin{cases} v_2(\Omega, x) & \text{if } \Omega \cdot n_1(x) \geq 0, \\ v_1(\Omega, x) & \text{if } \Omega \cdot n_1(x) < 0. \end{cases}$$

We also define the linear form

$$(2.18) \quad \ell(w) = \varepsilon 4\pi \int_D q \bar{w} dx + \sum_{F \in \mathcal{F}_h^o} \int_{\{\Omega \cdot n \leq 0\} \times F} |\Omega \cdot n| \alpha w d\Omega dx.$$

The discrete DG version of (2.1) that we investigate in the present paper is as follows: Seek $\psi_h \in \mathcal{W}_h$ so that the following holds:

$$(2.19) \quad \mathcal{L}(\psi_h, w) = \ell(w) \quad \forall w \in \mathcal{W}_h.$$

At variance with [1], we do not use quadratures to approximate integrals over $S^2 \times D$. We assume that all the integrals are evaluated exactly. This choice simplifies the presentation without affecting the conclusions in any dramatic way.

3. Formal asymptotic analysis, $\varepsilon \rightarrow 0$. In this section we perform the formal asymptotic analysis of the problem (2.19) under the assumption $\varepsilon \rightarrow 0$. Rigorous convergence statements are reported in section 4. The reader who is familiar with the formal results from [1, 9] or wants to see the rigorous argument can skip this section and go directly to section 4.

3.1. The continuous problem. Let us focus first on the continuous problem (2.1). The presence of the small parameter ε suggests the following expansion for ψ (see, e.g., [1, 9] or [5, Chap. XXI, sect. 5]):

$$(3.1) \quad \psi = \psi^{(0)} + \varepsilon \psi^{(1)} + \varepsilon^2 \psi^{(2)} + o(\varepsilon^2).$$

The leading-order angular intensity can be shown to be isotropic (here and below, a function is said to be isotropic if it is constant with respect to Ω and therefore depends on space only) and to satisfy the following diffusion equation:

$$(3.2a) \quad -\nabla \cdot \left(\frac{1}{3\sigma} \nabla \psi^{(0)} \right) + \sigma_a \psi^{(0)} = q \quad \text{in } D,$$

$$(3.2b) \quad \psi^{(0)}(x) = \frac{1}{2\pi} \int_{\Omega \cdot n(x) < 0} W(|\Omega \cdot n(x)|) \alpha(\Omega, x) d\Omega \quad \text{on } \partial D,$$

where $W(\mu) = \frac{\sqrt{3}}{2} \mu H(\mu)$ is defined in terms of Chandrasekhar's H -function for isotropic scattering in a conservative medium (see [13] for the asymptotic analysis and [2] for details on the H -function). It is shown in [13] that $\lim_{\varepsilon \rightarrow 0} \psi = \psi^{(0)}$, and

the convergence is not uniform unless the incident flux is isotropic. The purpose of the present paper is to investigate whether we recover $\psi^{(0)}$ from (2.1) under the process $\lim_{h \rightarrow 0} \lim_{\varepsilon \rightarrow 0} \psi_h$.

Observe that if α is isotropic, we have

$$\begin{aligned} \psi^{(0)}(x) &= \frac{1}{2\pi} \alpha(x) \int_{\Omega \cdot n(x) < 0} W(|\Omega \cdot n(x)|) \, d\Omega \\ &= \frac{1}{2\pi} \alpha(x) 2\pi \int_0^{\pi/2} W(\cos(\theta)) \sin(\theta) \, d\theta = \alpha(x) \int_0^1 W(\mu) \, d\mu \\ &= \alpha(x), \end{aligned}$$

where we used $\int_0^1 W(\mu) \, d\mu = 1$; cf. [2, p. 109].

For further reference we define the following scalar-valued and vector-valued functions:

$$(3.3a) \quad m(x) = \frac{1}{\pi} \int_{\Omega \cdot n(x) < 0} \alpha(\Omega, x) |\Omega \cdot n(x)| \, d\Omega,$$

$$(3.3b) \quad M(x) = \frac{1}{4\pi} \int_{\Omega \cdot n(x) < 0} \alpha(\Omega, x) |\Omega \cdot n(x)| \Omega \, d\Omega.$$

Observe that

$$(3.4) \quad m = \alpha \quad \text{and} \quad M = -\frac{1}{6} \alpha n \quad \text{if } \alpha \text{ is isotropic.}$$

Remark 3.1. We will assume that $m(x)$ is the trace of a function in \mathcal{C}_h ; i.e., $m(x)$ is continuous and piecewise polynomial on the boundary. This assumption is not restrictive. If $m(x)$ is not the trace of a function in \mathcal{C}_h , then we construct an approximation of m , say, m_h , and the rest of the paper is unchanged but for perturbations involving the difference $m - m_h$. The extra technical consistency terms involving $m - m_h$ are a distraction we want to avoid.

In the rest of the paper we determine conditions on \mathcal{W}_h and on α that are sufficient to guarantee that the leading-order of the solution to the discrete problem (2.19) converges to the solution of (3.2).

3.2. Formal asymptotic on the discrete problem. Let ψ_h be the solution to (2.19). Let us assume the following formal expansion for the discrete angular intensity:

$$(3.5) \quad \psi_h = \psi_h^{(0)} + \varepsilon \psi_h^{(1)} + \varepsilon^2 \psi_h^{(2)} + o(\varepsilon^2).$$

We now derive the problem solved by the leading term $\psi_h^{(0)}$.

We proceed as in [1]. After inserting (3.5) into (2.19) and using the representation of \mathcal{L} that involves the upwind flux (2.14), the zeroth order term gives

$$(3.6) \quad \int_{S^2 \times D} \left(\psi_h^{(0)} - \overline{\psi_h^{(0)}} \right) w(\Omega, x) \sigma(x) \, d\Omega \, dx = 0 \quad \forall w \in \mathcal{W}_h,$$

the first-order term gives

$$\begin{aligned}
 (3.7) \quad & \sum_{K \in \mathcal{T}_h} \int_{S^2 \times K} \left(-\psi_h^{(0)} \Omega \cdot \nabla w + \left(\psi_h^{(1)} - \bar{\psi}_h^{(1)} \right) \sigma w \right) d\Omega dx \\
 & + \sum_{F \in \mathcal{F}_h^i} \int_{S^2 \times F} \Omega \cdot n_1 \psi_h^{(0)\uparrow} \llbracket w \rrbracket d\Omega dx + \sum_{F \in \mathcal{F}_h^\partial} \int_{\{\Omega \cdot n \geq 0\} \times F} \Omega \cdot n \psi_h^{(0)} w d\Omega dx \\
 & = \sum_{F \in \mathcal{F}_h^\partial} \int_{\{\Omega \cdot n \leq 0\} \times F} |\Omega \cdot n| \alpha w d\Omega dx \quad \forall w \in \mathcal{W}_h,
 \end{aligned}$$

and the second-order term gives

$$\begin{aligned}
 (3.8) \quad & \sum_{K \in \mathcal{T}_h} \int_{S^2 \times K} \left(-\psi_h^{(1)} \Omega \cdot \nabla w + \left(\psi_h^{(2)} - \bar{\psi}_h^{(2)} \right) \sigma w + \sigma_a \bar{\psi}_h^{(0)} w \right) d\Omega dx \\
 & + \sum_{F \in \mathcal{F}_h^i} \int_{S^2 \times F} \Omega \cdot n_1 \psi_h^{(1)\uparrow} \llbracket w \rrbracket d\Omega dx + \sum_{F \in \mathcal{F}_h^\partial} \int_{\{\Omega \cdot n \geq 0\} \times F} \Omega \cdot n \psi_h^{(1)} w d\Omega dx \\
 & = \int_D 4\pi q \bar{w} dx \quad \forall w \in \mathcal{W}_h.
 \end{aligned}$$

In the rest of section 3 we assume that (3.6)–(3.8) define $\psi_h^{(0)}$, $\psi_h^{(1)}$, $\psi_h^{(2)}$. Our goal is to investigate the properties of $\psi_h^{(0)}$, $\psi_h^{(1)}$, $\psi_h^{(2)}$ and see how $\psi_h^{(0)}$ relates to $\psi^{(0)}$ as defined by (3.2) when the mesh size h goes to zero.

A first consequence of (3.6) is the following lemma.

LEMMA 3.1 (isotropy). $\psi_h^{(0)}$ is isotropic, provided (2.3) and (2.12) hold.

Proof. Owing to (2.12), $\bar{\psi}_h^{(0)}$ is a member of \mathcal{W}_h . As a result, $\psi_h^{(0)} - \bar{\psi}_h^{(0)}$ is a member of \mathcal{W}_h . The above observation implies that we can test (3.6) with $\psi_h^{(0)} - \bar{\psi}_h^{(0)}$, giving $\|\sigma^{\frac{1}{2}}(\psi_h^{(0)} - \bar{\psi}_h^{(0)})\|_{L^2(S^2 \times D)} = 0$. Observing that $\sigma(x) dx$ is a strictly positive measure owing to (2.3), we infer that $\psi_h^{(0)} = \bar{\psi}_h^{(0)} \in \mathcal{D}_h$, which proves that $\psi_h^{(0)}$ is isotropic. \square

The following holds as a consequence of (3.7).

LEMMA 3.2 (continuity). $\psi_h^{(0)}$ satisfies the following identity, provided (2.3) and (2.12) hold:

$$(3.9) \quad \sum_{F \in \mathcal{F}_h^i} \int_F \llbracket \psi_h^{(0)} \rrbracket \llbracket \varphi \rrbracket dx + \sum_{F \in \mathcal{F}_h^\partial} \int_F (\psi_h^{(0)} - m(x)) \varphi dx = 0 \quad \forall \varphi \in \mathcal{D}_h,$$

where $m(x)$ is defined in (3.3a).

Proof. Let φ be a member of \mathcal{D}_h ; then φ is a member of \mathcal{W}_h owing to (2.12). We then test (3.7) with φ . Owing to the isotropy of $\psi_h^{(0)}$ (see Lemma 3.1), we have

$$\sum_{K \in \mathcal{T}_h} \int_{S^2 \times K} -\psi_h^{(0)} \Omega \cdot \nabla \varphi d\Omega dx = \sum_{K \in \mathcal{T}_h} \int_K -\psi_h^{(0)}(x) \left(\int_{S^2} \Omega d\Omega \right) \cdot \nabla \varphi(x) dx = 0.$$

Moreover, by the definition of the average intensity $\bar{\psi}_h^{(1)}$, we also have

$$\sum_{K \in \mathcal{T}_h} \int_{S^2 \times K} \left(\psi_h^{(1)} - \bar{\psi}_h^{(1)} \right) \sigma \varphi d\Omega dx = \sum_{K \in \mathcal{T}_h} \int_K \left(\int_{S^2} \psi_h^{(1)} d\Omega - 4\pi \bar{\psi}_h^{(1)} \right) \sigma \varphi dx = 0.$$

We treat the flux terms in the left-hand side of (3.7) as follows:

$$\begin{aligned}
& \sum_{F \in \mathcal{F}_h^i} \int_{S^2 \times F} \Omega \cdot n_1 \psi_h^{(0)\uparrow} \llbracket \varphi \rrbracket d\Omega dx + \sum_{F \in \mathcal{F}_h^\partial} \int_{\{\Omega \cdot n \geq 0\} \times F} \Omega \cdot n \psi_h^{(0)} \varphi d\Omega dx \\
&= \sum_{F \in \mathcal{F}_h^i} \int_F \left(\int_{\Omega \cdot n_1 \geq 0} |\Omega \cdot n_1| d\Omega \right) (\psi_{h1}^{(0)} - \psi_{h2}^{(0)}) \llbracket \varphi \rrbracket dx \\
&\quad + \sum_{F \in \mathcal{F}_h^\partial} \int_F \left(\int_{\Omega \cdot n \geq 0} |\Omega \cdot n| d\Omega d\Omega \right) \psi_h^{(0)} \varphi dx \\
&= \sum_{F \in \mathcal{F}_h^i} \pi \int_F \llbracket \psi_h^{(0)} \rrbracket \llbracket \varphi \rrbracket dx + \sum_{F \in \mathcal{F}_h^\partial} \pi \int_F \psi_h^{(0)} \varphi dx.
\end{aligned}$$

Using the definition of m in (3.3a), the right-hand side in (3.7) gives

$$\sum_{F \in \mathcal{F}_h^\partial} \int_{\{\Omega \cdot n \leq 0\} \times F} |\Omega \cdot n| \alpha \varphi d\Omega dx = \sum_{F \in \mathcal{F}_h^\partial} \pi \int_F m \varphi dx.$$

Combining the above computations gives the desired result. \square

COROLLARY 3.3. *Assume that (2.3) and (2.12) hold. If m is the restriction over ∂D of a function in \mathcal{C}_h , then $\psi_h^{(0)} \in \mathcal{C}_h$ with boundary values equal to m .*

Proof. Since we assume m is a trace of a function in \mathcal{C}_h , we can choose a lifting $l_h(m) \in \mathcal{C}_h$ such that $l_h(m) = m$ on ∂D . Then (3.9) implies

$$\sum_{F \in \mathcal{F}_h^i} \int_F \llbracket \psi_h^{(0)} - l_h(m) \rrbracket \llbracket \varphi \rrbracket dx + \sum_{F \in \mathcal{F}_h^\partial} \int_F (\psi_h^{(0)} - l_h(m)) \varphi dx = 0 \quad \forall \varphi \in \mathcal{D}_h.$$

Taking $\varphi = \psi_h^{(0)} - l_h(m)$ in the above equation yields

$$\begin{aligned}
\int_F \llbracket \psi_h^{(0)} - l_h(m) \rrbracket^2 dx &= \int_F \llbracket \psi_h^{(0)} \rrbracket^2 dx = 0 & \forall F \in \mathcal{F}_h^i, \\
\int_F (\psi_h^{(0)} - l_h(m))^2 dx &= 0 & \forall F \in \mathcal{F}_h^\partial,
\end{aligned}$$

which proves the corollary. \square

Remark 3.2. When $m = 0$, we can take $l_h(m) = 0$, and Corollary 3.3 applies by assuming only that (2.3) and (2.12) hold. If m is not the trace of a function in \mathcal{C}_h , then the boundary value of $\psi_h^{(0)}$ can be shown to be close to a suitable approximation of m ; see Remark 3.1.

3.3. The limit problem. We now construct the problem solved by $\psi_h^{(0)}$. We start with a standard technical result.

LEMMA 3.4. *The following holds for all vectors a and b in \mathbb{R}^3 :*

$$(3.10) \quad \int_{S^2} (\Omega \cdot a)(\Omega \cdot b) d\Omega = \frac{4\pi}{3} a \cdot b.$$

Proof. By observing that on the unit sphere we have $\Omega = x$ and $\Omega = n$, where x is the position vector and n the unit normal pointing outward, we derive the following:

$$\int_{S^2} (\Omega \cdot a)(\Omega \cdot b) d\Omega = \int_{S^2} (x \cdot a)(n \cdot b) d\Omega = \int_{B^3} \nabla \cdot ((x \cdot a)b) dx = \frac{4\pi}{3} a \cdot b,$$

where B^3 denotes the unit ball in \mathbb{R}^3 . \square

Let us define the vector field

$$(3.11) \quad J_h^{(0)}(x) = \int_{S^2} \psi_h^{(1)}(\Omega, x) \Omega \, d\Omega.$$

Due to the tensor product structure of the discretization, each component of the vector field $J_h^{(0)}$ is a member of D_h ; i.e., $J_h^{(0)} \in \mathcal{D}_h^3$.

We make the following assumption on the discrete setting \mathcal{S}_h :

$$(3.12) \quad \text{the map } S^2 \ni \Omega \longmapsto \Omega \in \mathbb{R}^3 \text{ belongs to } \mathcal{S}_h.$$

Remark 3.3. Assumption (3.12) is quite restrictive. For instance, it does not hold if \mathcal{S}_h is composed of piecewise constant functions, which is a commonly used approximation in practice. We explain in Appendix A how this hypothesis can be removed, and we show that the conclusions of this and the following sections hold up to minor nonessential modifications. We retain assumption (3.12) in the rest of the paper, since it significantly simplifies the notation and keeps us focused on the essential.

LEMMA 3.5. *Assume that (2.3), (2.12), and (3.12) hold. The fields $\psi_h^{(0)} \in \mathcal{C}_{h,m}$ and $J_h^{(0)} \in \mathcal{D}_h^3$ solve the following equations for all $\varphi \in \mathcal{C}_{h,0}$ and for all $\varpi \in \mathcal{D}_h^3$:*

$$(3.13a) \quad \sum_{K \in \mathcal{T}_h} \int_K \left(-J_h^{(0)} \cdot \nabla \varphi + 4\pi \sigma_a \psi_h^{(0)} \varphi \right) dx = \sum_{K \in \mathcal{T}_h} \int_K 4\pi q \varphi \, dx,$$

$$(3.13b) \quad \sum_{K \in \mathcal{T}_h} \int_K \left(\frac{4\pi}{3} \nabla \psi_h^{(0)} + \sigma J_h^{(0)} \right) \cdot \varpi \, dx = \sum_{F \in \mathcal{F}_h^p} \int_F 4\pi \left(\frac{1}{6} mn + M \right) \cdot \varpi \, dx.$$

Proof. We proceed in three steps.

Step 1. Due to Corollary 3.3, the hypotheses (2.3) and (2.12) imply that $\psi_h^{(0)}$ is in $\mathcal{C}_{h,m}$. That $J_h^{(0)}$ belongs to \mathcal{D}_h^3 has already been established above.

Step 2. Let us test (3.8) using an isotropic test function $\varphi \in \mathcal{D}_h$. The first term inside the summation over the mesh elements gives

$$\int_{S^2 \times K} -\psi_h^{(1)} \Omega \cdot \nabla \varphi \, d\Omega \, dx = \int_K - \left(\int_{S^2} \psi_h^{(1)} \Omega \, d\Omega \right) \cdot \nabla \varphi \, dx = \int_K -J_h^{(0)} \cdot \nabla \varphi \, dx.$$

The two terms involving $\psi_h^{(2)}$ and $\overline{\psi}_h^{(2)}$ cancel each other, owing to the definition of the average intensity $\overline{\psi}_h^{(2)}$. The fourth term involving $\overline{\psi}_h^{(0)}$ gives

$$\int_{S^2 \times K} \sigma_a \overline{\psi}_h^{(0)} \varphi \, d\Omega \, dx = \int_K 4\pi \sigma_a \overline{\psi}_h^{(0)} \varphi \, dx.$$

We now restrict the test functions to those that are continuous and zero at the boundary of D , i.e., we pick φ in $\mathcal{C}_{h,0}$ (see definition (2.9)); as a result the boundary integrals in (3.8) are zero. The right-hand-side in (3.8) gives

$$\int_D 4\pi \overline{\varphi} \, dx = 4\pi \int_D q \varphi \, dx.$$

By combining the above results, we finally infer that when (3.8) is tested with functions φ in $\mathcal{C}_{h,0}$, the following holds:

$$\sum_{K \in \mathcal{T}_h} \int_K \left(-J_h^{(0)} \cdot \nabla \varphi + 4\pi \sigma_a \overline{\psi}_h^{(0)} \varphi \right) dx = \sum_{K \in \mathcal{T}_h} \int_K 4\pi q \varphi \, dx,$$

which is (3.13a).

Step 3. Let us now test (3.7) using $w(\Omega, x) = \varphi(x)\Omega \cdot e_i$, with $\varphi(x)$ being an arbitrary function in \mathcal{D}_h . Note that w is a legitimate test function, owing to the hypothesis (3.12). Let us observe first that after integration by parts, (3.7) can be rewritten as follows (see also (2.16)):

$$\begin{aligned} & \sum_{K \in \mathcal{T}_h} \int_{S^2 \times K} \left(\Omega \cdot \nabla \psi_h^{(0)} + \psi_h^{(1)} - \bar{\psi}_h^{(1)} \right) w \, d\Omega \, dx + \sum_{F \in \mathcal{F}_h^\partial} \int_{\{\Omega \cdot n \leq 0\} \times F} |\Omega \cdot n| \psi_h^{(0)} w \, d\Omega \, dx \\ & + \sum_{F \in \mathcal{F}_h^i} \int_{S^2 \times F} -\Omega \cdot n_1 \llbracket \psi_h^{(0)} \rrbracket w^\downarrow \, d\Omega \, dx = \sum_{F \in \mathcal{F}_h^\partial} \int_{\{\Omega \cdot n \leq 0\} \times F} |\Omega \cdot n| \alpha w \, d\Omega \, dx. \end{aligned}$$

First, using the Einstein summation convention together with (3.10) and $e_i \cdot e_j = \delta_{ij}$, we infer

$$\int_{S^2 \times K} w \Omega \cdot \nabla \psi_h^{(0)} \, d\Omega \, dx = \int_K \varphi \partial_j \psi_h^{(0)} \, dx \int_{S^2} (\Omega \cdot e_j)(\Omega \cdot e_i) \, d\Omega = \frac{4\pi}{3} \int_K \varphi \partial_i \psi_h^{(0)} \, dx.$$

Second, we compute the contribution of $\psi_h^{(1)}$:

$$\int_{S^2 \times K} \left(\psi_h^{(1)} - \bar{\psi}_h^{(1)} \right) w \sigma \, d\Omega \, dx = \int_{S^2 \times K} \psi_h^{(1)} (\Omega \cdot e_i) \varphi \sigma \, d\Omega \, dx = \int_K (J_h^{(0)} \cdot e_i) \varphi \sigma \, dx.$$

Third, we take care of the flux terms. Since $\psi_h^{(0)}$ is a member of $\mathcal{C}_{h,m}$, the flux term in the left-hand side gives

$$\sum_{F \in \mathcal{F}_h^\partial} \int_F \left(\int_{\Omega \cdot n \leq 0} |\Omega \cdot n| (\Omega \cdot e_i) \, d\Omega \right) m \varphi \, dx = -\frac{2\pi}{3} \sum_{F \in \mathcal{F}_h^\partial} \int_F (n \cdot e_i) m \varphi \, dx.$$

Similarly, for the flux term in the right-hand side, we have

$$\sum_{F \in \mathcal{F}_h^\partial} \int_F \left(\int_{\Omega \cdot n < 0} |\Omega \cdot n| (\Omega \cdot e_i) \alpha(\Omega, x) \, d\Omega \right) \varphi \, dx = \sum_{F \in \mathcal{F}_h^\partial} \int_F 4\pi (M \cdot e_i) \varphi \, dx.$$

We now combine all the above results, and we obtain

$$\sum_{K \in \mathcal{T}_h} \int_K \left(\frac{4\pi}{3} \nabla \psi_h^{(0)} + \sigma J_h^{(0)} \right) \cdot e_i \varphi \, dx = \sum_{F \in \mathcal{F}_h^\partial} \int_F 4\pi \left(\frac{1}{6} mn + M \right) \cdot e_i \varphi \, dx,$$

which is (3.13b). \square

To some extent (3.13) looks like an local discontinuous Galerkin (LDG) approximation of (3.2a); see, e.g., [4]. The flux $J_h^{(0)}$ is approximated by discontinuous vector-valued functions in \mathcal{D}_h^3 , but contrary to most LDG techniques, the primal variable ψ_h^0 is approximated by continuous functions.

To be able to better analyze the well-posedness of (3.13) we are going to eliminate $J_h^{(0)}$. For this purpose we make the following simplifying assumption:

$$(3.14) \quad \sigma \text{ is constant on each cell of each mesh of the family } \{\mathcal{T}_h\}_{h>0}.$$

This means that the manifold across which σ might be discontinuous does not cross any mesh element; in other words, the mesh family matches the possible discontinuities of the scalar field σ .

We now derive the main result of this section.

PROPOSITION 3.6. *Assume that (2.3), (2.12), (3.12), and (3.14) hold. Then $\psi_h^{(0)} \in \mathcal{C}_{h,m}$ solves the following problem: for all $\varphi \in \mathcal{C}_{h,0}$,*

$$(3.15) \quad \int_D \left(\frac{1}{3\sigma} \nabla \psi_h^{(0)} \cdot \nabla \varphi + \sigma_a \psi_h^{(0)} \varphi \right) dx = \int_{\partial D} \frac{1}{\sigma} \left(\frac{m}{6} + M \cdot n \right) \cdot \partial_n \varphi dx + \int_D q \varphi dx.$$

Proof. Let φ be a member of $\mathcal{C}_{h,0}$, and set $\varpi = \nabla \varphi$. Owing to (2.7), ϖ is a member of \mathcal{D}_h^3 ; moreover, using the fact that σ is piecewise constant over \mathcal{T}_h , $\frac{1}{\sigma} \varpi$ is again a member of \mathcal{D}_h^3 . Using $\frac{1}{\sigma} \varpi$ as a test function in (3.13b) and making use of (3.13a), we obtain

$$\begin{aligned} 0 &= \sum_{K \in \mathcal{T}_h} \int_K \left(\frac{4\pi}{3\sigma} \nabla \psi_h^{(0)} + J_h^{(0)} \right) \cdot \nabla \varphi dx - \sum_{F \in \mathcal{F}_h^\partial} \int_F \frac{4\pi}{\sigma} \left(\frac{1}{6} mn + M \right) \cdot \nabla \varphi dx \\ &= \sum_{K \in \mathcal{T}_h} \int_K \left(\frac{4\pi}{3\sigma} \nabla \psi_h^{(0)} \cdot \nabla \varphi + 4\pi (\sigma_a \psi_h^{(0)} - q) \varphi \right) dx \\ &\quad - \sum_{F \in \mathcal{F}_h^\partial} \int_F \frac{4\pi}{\sigma} \left(\frac{1}{6} mn + M \right) \cdot \nabla \varphi dx. \end{aligned}$$

Now we observe that for any function φ in $\mathcal{C}_{h,0}$, the tangent component of $\nabla \varphi$ at the boundary of the domain D is zero; this means that $\nabla \varphi = n \partial_n \varphi$. Inserting this information in the above equality yields the desired result. \square

4. Rigorous derivation of the limit problem. The purpose of this section is to rederive rigorously the limit problem (3.15) without invoking the formal expansion (3.1). We now ignore the definitions of $\psi_h^{(0)}$, $\psi_h^{(1)}$, and $\psi_h^{(2)}$ from the previous section. We are going to show in Lemma 4.3 that ψ_h has an isotropic limit in an appropriate Sobolev space as $\varepsilon \rightarrow 0$ that we denote by $\psi_h^{(0)}$, i.e.,

$$(4.1) \quad \psi_h^{(0)} := \lim_{\varepsilon \rightarrow 0} \psi_h.$$

We show at the end of this section that this limit solves the same limit equation as that already derived in section 3, which justifies our using the same symbol.

We assume throughout this section that (2.3), (2.12), (3.12), and (3.14) hold and

$$(4.2) \quad \alpha \in L^2(\Gamma_-(D)).$$

4.1. The a priori estimates. Let us introduce the following discrete semi-norms

$$(4.3) \quad \|v\|_{J^i}^2 = \sum_{F \in \mathcal{F}_h^i} \int_{S^2 \times F} v^2 |\Omega \cdot n_1| d\Omega dx, \quad \|v\|_{J^\partial}^2 = \sum_{F \in \mathcal{F}_h^\partial} \int_{S^2 \times F} v^2 |\Omega \cdot n| d\Omega dx,$$

and let us set $\|v\|_J^2 := \|v\|_{J^i}^2 + \|v\|_{J^\partial}^2$.

We recall the following coercivity property of the bilinear form \mathcal{L} .

LEMMA 4.1 (*L*²-coercivity). *The following identity holds for all $v \in \mathcal{W}_h$:*

$$(4.4) \quad \mathcal{L}(v, v) = \frac{1}{\varepsilon} \|\sigma^{\frac{1}{2}}(v - \bar{v})\|_{L^2(S^2 \times D)}^2 + \varepsilon 4\pi \|\sigma_a^{\frac{1}{2}} \bar{v}\|_{L^2(D)}^2 + \frac{1}{2} \|[v]\|_J^2.$$

Proof. Let us use definition (2.16). The main technicality consists of handling the advection and interface terms, which are dealt with by integration by parts [11]:

$$\begin{aligned}
& \sum_{K \in \mathcal{T}_h} \int_{S^2 \times K} \Omega \cdot \nabla \frac{1}{2} v^2 \, d\Omega \, dx + \sum_{F \in \mathcal{F}_h^i} \int_{S^2 \times F} -\Omega \cdot n_1 \llbracket v \rrbracket v^\downarrow \, d\Omega \, dx \\
& \quad + \sum_{F \in \mathcal{F}_h^\partial} \int_{\{\Omega \cdot n \leq 0\} \times F} |\Omega \cdot n| v^2 \, d\Omega \, dx \\
& = \sum_{F \in \mathcal{F}_h^i} \int_{S^2 \times F} \Omega \cdot n_1 (\{v\} - v^\downarrow) \llbracket v \rrbracket \, d\Omega \, dx \\
& \quad + \sum_{F \in \mathcal{F}_h^\partial} \int_{S^2 \times F} |\Omega \cdot n| \frac{1}{2} v^2 \, d\Omega \, dx = \frac{1}{2} \|\llbracket v \rrbracket\|_J^2.
\end{aligned}$$

Then, after realizing that

$$\int_D \sigma \|v - \bar{v}\|_{L^2(S^2)}^2 \, dx = \int_D \sigma (\|v\|_{L^2(S^2)}^2 - 4\pi \bar{v}^2) \, dx,$$

the rest of the proof follows easily. \square

We now deduce a priori estimates that are uniform with respect to ε but possibly nonuniform with respect to h ; uniformity with respect to h is dealt with in section 5.

LEMMA 4.2. *Assume that (2.3)–(2.4) and (4.2) hold. There is $c(\alpha, q, h)$, uniform with respect to ε , so that*

$$(4.5) \quad \frac{1}{\varepsilon} \|\psi_h - \bar{\psi}_h\|_{L^2(S^2 \times D)}^2 + \varepsilon \|\bar{\psi}_h\|_{L^2(D)}^2 + \|\llbracket \psi_h \rrbracket\|_{J_i}^2 + \|\psi_h - m\|_{J^\partial}^2 \leq c(\alpha, q, h)\varepsilon,$$

where $c(\alpha, q, h) = ch^{-1} \|\alpha\|_{L^2(\{\Omega \cdot n \leq 0\} \times D)}^2 + c' \|q\|_{L^2(D)}^2$.

Proof. The idea is to use the L^2 -coercivity of \mathcal{L} .

Step 1. We define $\phi_m \in \mathcal{C}_{h,m}$ so that

$$\|\phi_m\|_{H^1(D)} \leq c \|m\|_{H^{\frac{1}{2}}(\partial D)} \leq c' \|h^{-\frac{1}{2}} m\|_{L^2(D)} \leq c'' \|h^{-\frac{1}{2}} \alpha\|_{L^2(\{\Omega \cdot n \leq 0\} \times D)}.$$

One can construct ϕ_m by taking the Clément [3] or Scott–Zhang [15] interpolant of any H^1 -lifting of m (recall that we assumed that m is the trace of function in \mathcal{C}_h). Then, using the fact that ϕ_m is continuous and isotropic and using the representation (2.16) for the bilinear form \mathcal{L} , we rewrite (2.19) as follows:

$$\begin{aligned}
(4.6) \quad \mathcal{L}(\psi_h - \phi_m, w) & = \ell(w) - \mathcal{L}(\phi_m, w) \\
& = \varepsilon \int_D 4\pi q \bar{w} \, dx + \sum_{F \in \mathcal{F}_h^\partial} \int_{\{\Omega \cdot n \leq 0\} \times F} |\Omega \cdot n| (\alpha - \phi_m) w \, d\Omega \, dx \\
& \quad - \int_{S^2 \times D} (\Omega \cdot \nabla \phi_m + \varepsilon \sigma_a \phi_m) w \, d\Omega \, dx := R_1(w) + R_2(w) + R_3(w),
\end{aligned}$$

with obvious notation for the three terms R_1, R_2, R_3 .

Step 2. Now we use $w = \psi_h - \phi_m$ to test the above equation. Note that $\psi_h - \phi_m$ is an admissible test function, since $\psi_h - \phi_m \in \mathcal{C}_{h,0}$. Using (4.4), we infer

$$\begin{aligned}
(4.7) \quad \frac{1}{\varepsilon} \|\sigma^{\frac{1}{2}}(\psi_h - \bar{\psi}_h)\|_{L^2(S^2 \times D)}^2 + \varepsilon 4\pi \|\sigma_a^{\frac{1}{2}}(\bar{\psi}_h - \phi_m)\|_{L^2(D)}^2 + \frac{1}{2} \|\llbracket \psi_h - \phi_m \rrbracket\|_J^2 \\
= R_1(\psi_h - \phi_m) + R_2(\psi_h - \phi_m) + R_3(\psi_h - \phi_m).
\end{aligned}$$

The rest of the proof consists of bounding each of the three terms $|R_1(\psi_h - \phi_m)|$, $|R_2(\psi_h - \phi_m)|$, and $|R_3(\psi_h - \phi_m)|$ in the right-hand side.

Step 3. For R_1 we have

$$\begin{aligned} |R_1(\psi_h - \phi_m)| &\leq \varepsilon 4\pi \int_D |q(\bar{\psi}_h - \phi_m)| \, dx \leq c\varepsilon \|q\|_{L^2(D)} \|\sigma_a^{\frac{1}{2}}(\bar{\psi}_h - \phi_m)\|_{L^2(D)} \\ &\leq c\varepsilon \|q\|_{L^2(D)}^2 + \varepsilon\pi \|\sigma_a^{\frac{1}{2}}(\bar{\psi}_h - \phi_m)\|_{L^2(D)}^2. \end{aligned}$$

Owing to the definition of m , the term R_2 can be handled as follows:

$$\begin{aligned} R_2(\psi_h - \phi_m) &:= \sum_{F \in \mathcal{F}_h^\partial} \int_{\{\Omega \cdot n \leq 0\} \times F} |\Omega \cdot n| (\alpha - m)(\psi_h - \phi_m) \, d\Omega \, dx \\ &= \sum_{F \in \mathcal{F}_h^\partial} \int_{\{\Omega \cdot n \leq 0\} \times F} |\Omega \cdot n| (\alpha - m)(\psi_h - \bar{\psi}_h) \, d\Omega \, dx. \end{aligned}$$

We use an inverse inequality and proceed as follows:

$$\begin{aligned} |R_2(\psi_h - \phi_m)| &\leq \|\alpha - m\|_{L^2(\{\Omega \cdot n \leq 0\} \times \partial D)} \|\psi_h - \bar{\psi}_h\|_{L^2(S^2 \times \partial D)} \\ &\leq ch^{-\frac{1}{2}} \|\alpha - m\|_{L^2(\{\Omega \cdot n \leq 0\} \times \partial D)} \|\sigma^{\frac{1}{2}}(\psi_h - \bar{\psi}_h)\|_{L^2(S^2 \times D)} \\ &\leq ch^{-1}\varepsilon \|\alpha\|_{L^2(\{\Omega \cdot n \leq 0\} \times \partial D)}^2 + \frac{1}{4\varepsilon} \|\sigma^{\frac{1}{2}}(\psi_h - \bar{\psi}_h)\|_{L^2(S^2 \times D)}^2. \end{aligned}$$

We handle the third term as follows:

$$\begin{aligned} R_3(\psi_h - \phi_m) &:= - \int_{S^2 \times D} (\Omega \cdot \nabla \phi_m + \varepsilon \sigma_a \phi_m) (\psi_h - \phi_m) \, d\Omega \, dx \\ &= - \int_{S^2 \times D} (\Omega \cdot \nabla \phi_m) (\psi_h - \bar{\psi}_h) \, dx - \int_{S^2 \times D} \varepsilon \sigma_a \phi_m (\bar{\psi}_h - \phi_m) \, dx. \end{aligned}$$

This gives

$$\begin{aligned} |R_3(\psi_h - \phi_m)| &\leq c_1 \|\nabla \phi_m\|_{L^2(D)} \|\psi_h - \bar{\psi}_h\|_{L^2(S^2 \times D)} + c_2 \varepsilon \|\phi_m\|_{L^2(D)} \|\bar{\psi}_h - \phi_m\|_{L^2(D)} \\ &\leq c_1 \varepsilon \|\phi_m\|_{H^1(D)}^2 + \varepsilon\pi \|\sigma^{\frac{1}{2}}(\bar{\psi}_h - \phi_m)\|_{L^2(D)}^2 + \frac{1}{4\varepsilon} \|\sigma^{\frac{1}{2}}\psi_h - \bar{\psi}_h\|_{L^2(S^2 \times D)}^2. \end{aligned}$$

Step 4. By inserting all the bounds that have been derived above into (4.7) we finally infer

$$\frac{1}{\varepsilon} \|\sigma^{\frac{1}{2}}(\psi_h - \bar{\psi}_h)\|_{L^2(S^2 \times D)}^2 + \varepsilon \|\sigma_a^{\frac{1}{2}}(\bar{\psi}_h - \phi_m)\|_{L^2(D)} + \|\llbracket \psi_h - \phi_m \rrbracket\|_J^2 \leq c(\alpha, q, h)\varepsilon,$$

where $c(\alpha, q, h) = ch^{-1} \|\alpha\|_{L^2(\{\Omega \cdot n \leq 0\} \times D)}^2 + c' \|q\|_{L^2(D)}^2$, and (4.5) follows easily. \square

Let us define the discrete vector field

$$(4.8) \quad J_h := \int_{S^2} \varepsilon^{-1} \psi_h \Omega \, d\Omega \in \mathcal{D}_h^3.$$

We can now derive the counterparts of Lemma 3.1 and Corollary 3.3.

LEMMA 4.3. *Assume that (2.3)–(2.4) and (4.2) hold. One can extract from the sequence $(\psi_h)_{\varepsilon>0}$ a subsequence that converges in every norm to a function $\psi_h^{(0)}$ in*

$\mathcal{C}_{h,m}$ when $\varepsilon \rightarrow 0$. One can also extract from the sequence $(J_h)_{\varepsilon>0}$ a subsequence that converges in every norm to a function $J_h^{(0)}$ in \mathcal{D}_h^3 when $\varepsilon \rightarrow 0$.

Proof. From the a priori estimate (4.5), we deduce that $\|\overline{\psi}_h\|_{L^2(D)} \leq c(\alpha, q, h)^{\frac{1}{2}}$ and

$$\|\psi_h\|_{L^2(S^2 \times D)} \leq \|\psi_h - \overline{\psi}_h\|_{L^2(S^2 \times D)} + \|\overline{\psi}_h\|_{L^2(D)} \leq c(\alpha, q, h)^{\frac{1}{2}}(1 + \varepsilon).$$

Then by compactness of the unit ball in \mathcal{W}_h (recall that \mathcal{W}_h is finite-dimensional), we can extract a subsequence (still denoted by $(\psi_h)_{\varepsilon>0}$) so that $\lim_{\varepsilon \rightarrow 0} \psi_h = \psi_h^{(0)}$, where $\psi_h^{(0)}$ is a member of \mathcal{W}_h and the convergence holds in every norm (recall that the mesh size is fixed for the time being). We also deduce from (4.5) that $\|\psi_h - \overline{\psi}_h\|_{L^2(D)} \leq c(\alpha, q, h)^{\frac{1}{2}}\varepsilon$, which means

$$\psi_h^{(0)} = \lim_{\varepsilon \rightarrow 0} \psi_h = \lim_{\varepsilon \rightarrow 0} \overline{\psi}_h = \overline{\psi_h^{(0)}}.$$

As a result $\psi_h^{(0)}$ is isotropic, i.e., $\psi_h^{(0)} \in \mathcal{D}_h$. The estimate (4.5) also implies

$$\|[\psi_h^{(0)}]\|_{J^i}^2 + \|\psi_h^{(0)} - m\|_{J^\partial}^2 = \lim_{\varepsilon \rightarrow 0} \|[\psi_h]\|_{J^i}^2 + \|\psi_h - m\|_{J^\partial}^2 \leq \lim_{\varepsilon \rightarrow 0} c(\alpha, q, h)\varepsilon = 0.$$

This means that $\psi_h^{(0)}$ is a member of $\mathcal{C}_{h,m}$.

For the field J_h we observe that

$$J_h(x) = \int_{S^2} \frac{1}{\varepsilon} (\psi_h(\Omega, x) - \overline{\psi}_h(x)) \Omega \, d\Omega \quad \forall x \in D.$$

Then, using the bound (4.5) we infer

$$\|J_h\|_{L^2(D)} \leq \frac{(4\pi)^{\frac{1}{2}}}{\varepsilon} \|\psi_h - \overline{\psi}_h\|_{L^2(S^2 \times D)} \leq (4\pi c(\alpha, q, h))^{\frac{1}{2}}.$$

We conclude from the compactness of the unit ball in \mathcal{D}_h (recall that \mathcal{D}_h is finite-dimensional) that we can extract a subsequence (still denoted by $(J_h)_{\varepsilon>0}$) so that $\lim_{\varepsilon \rightarrow 0} J_h = J_h^{(0)}$, where $J_h^{(0)}$ is a member of \mathcal{D}_h and the convergence holds in every norm (recall that the mesh size is fixed for the time being). This concludes the proof. \square

We are now ready to characterize the limit function $\psi_h^{(0)}$. The following theorem is the main result of this section.

THEOREM 4.4. *Assume that (2.3)–(2.4), (2.12), (3.12), and (4.2) hold. Then $\psi_h^{(0)} \in \mathcal{C}_{h,m}$ and $J_j^{(0)} \in \mathcal{D}_h^3$ solve the following mixed problem:*

$$(4.9a) \quad \sum_{K \in \mathcal{T}_h} \int_K \left(-J_h^{(0)} \cdot \nabla \varphi + 4\pi \sigma_\alpha \psi_h^{(0)} \varphi \right) dx = 4\pi \int_D q \varphi dx,$$

$$(4.9b) \quad \sum_{K \in \mathcal{T}_h} \int_K \left(\frac{4\pi}{3} \nabla \psi_h^{(0)} + \sigma J_h^{(0)} \right) \cdot \varpi dx = 4\pi \sum_{F \in \mathcal{F}_h^\partial} \int_F \left(\frac{1}{6} mn + M \right) \cdot \varpi dx$$

$\forall \varphi \in \mathcal{C}_{h,0}$ and $\forall \varpi \in \mathcal{D}_h$, respectively.

Proof. We proceed as in the proof of Lemma 3.5.

Step 1. Owing to (2.12), we can take $w = \varepsilon^{-1}\varphi \in \mathcal{C}_{h,0}$ to test (2.19). Using the representation (2.14) for the bilinear form \mathcal{L} , we obtain

$$\sum_{K \in \mathcal{T}_h} \int_K (-J_h \cdot \nabla \varphi + 4\pi \sigma_a \bar{\psi}_h \varphi) \, dx = 4\pi \int_D q \varphi \, dx.$$

We can now pass to the limit as $\varepsilon \rightarrow 0$ in this equation by using the fact that $J_h \rightarrow J_h^{(0)}$ and $\psi_h \rightarrow \psi_h^{(0)}$ in very norm in \mathcal{D}_h^3 and \mathcal{C}_h , respectively (see Lemma 4.3).

Step 2. Now we take $w(\Omega, x) = \varphi(x)\Omega \cdot e_i$, $i \in \overline{1,3}$, with φ in \mathcal{D}_h , to test (2.19) where we use the representation (2.16) for \mathcal{L} . (See Step 3 in the proof of Lemma 3.5.) For the first term, using the Einstein summation convention together with (3.10) and $e_i \cdot e_j = \delta_{ij}$, we infer

$$\int_{S^2 \times K} (\Omega \cdot e_i) \varphi \Omega \cdot \nabla \psi_h \, d\Omega \, dx = \int_K \varphi \int_{S^2} (\Omega \cdot e_j) (\Omega \cdot e_i) \partial_j \psi_h \, d\Omega \, dx \xrightarrow{\varepsilon \rightarrow 0} \frac{4\pi}{3} \int_K \varphi \partial_i \psi_h^{(0)} \, dx.$$

Second, we compute the contribution of the scattering term

$$\begin{aligned} \int_{S^2 \times K} \left((\psi_h - \bar{\psi}_h) \frac{\sigma}{\varepsilon} + \varepsilon \sigma_a \bar{\psi}_h \right) w \, d\Omega \, dx &= \int_{S^2 \times K} (\psi_h \Omega \cdot e_i) \varphi \frac{\sigma}{\varepsilon} \, d\Omega \, dx \\ &= \int_K (J_h \cdot e_i) \varphi \sigma \, dx \xrightarrow{\varepsilon \rightarrow 0} \int_K (J_h^{(0)} \cdot e_i) \varphi \sigma \, dx. \end{aligned}$$

Third, we rewrite the flux term in (2.16) as follows:

$$\begin{aligned} \sum_{F \in \mathcal{F}_h^i} \int_{S^2 \times F} -\Omega \cdot n_1 [\psi_h] w^\downarrow \, d\Omega \, dx + \sum_{F \in \mathcal{F}_h^\partial} \int_{\{\Omega \cdot n \leq 0\} \times F} |\Omega \cdot n| \psi_h w \, d\Omega \, dx \\ \xrightarrow{\varepsilon \rightarrow 0} \sum_{F \in \mathcal{F}_h^\partial} \int_F m \varphi \int_{\{\Omega \cdot n \leq 0\}} |\Omega \cdot n| (\Omega \cdot e_i) \, d\Omega \, dx = \sum_{F \in \mathcal{F}_h^\partial} -\frac{2\pi}{3} \int_F m \varphi (n \cdot e_i) \, dx. \end{aligned}$$

For the right-hand side, we have

$$\ell(w) \xrightarrow{\varepsilon \rightarrow 0} \sum_{F \in \mathcal{F}_h^\partial} \varphi \int_{\Omega \cdot n \leq 0} |\Omega \cdot n| (\Omega \cdot e_i) \alpha \, d\Omega \, dx = 4\pi \sum_{F \in \mathcal{F}_h^\partial} \int_F \varphi M \cdot e_i \, dx.$$

Putting everything together, we obtain

$$\sum_{K \in \mathcal{T}_h} \int_K \left(\frac{4\pi}{3} \partial_i \psi_h^{(0)} + \sigma J_h^{(0)} \cdot e_i \right) \varphi \, dx = 4\pi \sum_{F \in \mathcal{F}_h^\partial} \int_F \left(\frac{1}{6} mn + M \right) \cdot e_i \varphi \, dx.$$

This concludes the proof. \square

We summarize the results of this section in

COROLLARY 4.5. *Assume that (2.3)–(2.4), (2.12), (3.12), (3.14), and (4.2) hold.*

Then

(i) *the whole sequences $(\psi_h)_{\varepsilon > 0}$ and $(J_h)_{\varepsilon > 0}$ converge to $\psi_h^{(0)}$ and $J_h^{(0)}$, respectively, as $\varepsilon \rightarrow 0$;*

(ii) *the limits $\psi_h^{(0)}$ and $J_h^{(0)}$ obtained in this section are equal to those formally derived in section 3;*

(iii) *$\psi_h^{(0)} \in \mathcal{C}_{h,m}$ solves (3.15).*

Proof. The problem (4.9) has a unique solution (see Proposition 5.1 below); therefore, $\psi_h^{(0)}$ and $J_h^{(0)}$ are uniquely defined; as a result, the entire sequence converges to the same limit. Part (ii) is obvious, since (3.13) and (4.9) are the same. Finally, the proof of (iii) is identical to that of Proposition 3.6. \square

5. Analysis of the limit problem (3.15). We have proved in section 4 that the sequence $(\psi_h)_{\varepsilon>0}$ converges to the function $\psi_h^{(0)} \in \mathcal{C}_{h,m}$, solving (3.15) as ε goes to zero. The purpose of this section is to study the limit problem (3.15): we discuss well-posedness, convergence properties as the mesh size h goes zero, and possible incompatibilities on the boundary conditions.

5.1. (Nonuniform) H^1 -well-posedness. To reformulate (3.15) in a more convenient way we introduce the bilinear form

$$(5.1) \quad b(\phi, \varphi) = \int_D \left(\frac{1}{3\sigma} \nabla \phi \cdot \nabla \varphi + \sigma_a \phi \varphi \right) dx$$

and the linear form

$$(5.2) \quad r(\varphi) = \int_{\partial D} \frac{1}{\sigma} \left(\frac{m}{6} + M \cdot n \right) \partial_n \varphi dx + \int_D q \varphi dx.$$

Then (3.15) consists of finding $\psi_h^{(0)}$ in $\mathcal{C}_{h,m}$ satisfying

$$(5.3) \quad b(\psi_h^{(0)}, \varphi) = r(\varphi), \quad \varphi \in \mathcal{C}_{h,0}.$$

PROPOSITION 5.1. *The discrete problem (3.15) has a unique solution $\psi_h^{(0)}$, and the following bound holds:*

$$(5.4) \quad \|\psi_h^{(0)}\|_{H^1(D)} \leq \begin{cases} c \|q\|_{H^{-1}(D)} & \text{if } \alpha \text{ is isotropic,} \\ c(h^{-\frac{1}{2}} \|\frac{1}{6}m + M \cdot n\|_{L^2(\partial D)} + \|q\|_{H^{-1}(D)}) & \text{otherwise.} \end{cases}$$

Proof. We first observe that, owing to (2.3), the bilinear form b is coercive and bounded on $H_0^1(D) \times H_0^1(D)$ uniformly with respect to the mesh size. Moreover, the linear form r satisfies the following bound:

$$(5.5) \quad \sup_{0 \neq \varphi \in \mathcal{C}_{h,0}} \frac{\int_D r(\varphi) dx}{\|\varphi\|_{H^1(D)}} \leq (ch^{-\frac{1}{2}} \|\frac{1}{6}m + M \cdot n\|_{L^2(\partial D)} + \|q\|_{H^{-1}(D)}).$$

Note that, owing to (3.4), $\frac{1}{6}m + M \cdot n = 0$ if α is isotropic. The Lax–Milgram lemma implies the desired result. \square

5.2. Uniform H^s -well-posedness. Proposition 5.1 shows that if the incoming flux is not isotropic, the H^1 -norm of $\psi_h^{(0)}$ is not bounded uniformly as $h \rightarrow 0$. Loss of H^1 -boundedness is symptomatic of a boundary condition incompatibility which manifests itself by a numerical boundary layer. This issue will be explored in more detail in section 5.5. For the time being we want to determine whether uniform stability can be achieved for the solution of (3.15) in a norm which is weaker than that of $H^1(D)$.

Taking inspiration from [7], we now show that uniform stability holds in $H^s(D)$ with $s \in [0, \frac{1}{2})$ under some simplifying assumptions. In particular we now assume that the mesh family $\{\mathcal{T}_h\}_{h>0}$ is such that there is $c > 0$, uniform with respect to h , so that

$$(5.6) \quad \inf_{v_h \in \mathcal{C}_{h,0}} \|\phi - v_h\|_{H^p(D)} \leq ch^{l-p} \|\phi\|_{H^l(D)}, \quad \forall \phi \in H^l(D), \forall p \in [0, 1], \forall l \in [r, 2],$$

$$(5.7) \quad \inf_{v_h \in \mathcal{C}_{h,0}} (\|\phi - v_h\|_{L^2(\partial D)} + h \|\partial_n(\phi - v_h)\|_{L^2(\partial D)}) \leq ch^{l-\frac{1}{2}} \|\phi\|_{H^l(D)} \quad \forall l \in [1, 2].$$

Note that these are standard interpolation estimates in Sobolev spaces with non-integer derivative order.

LEMMA 5.2. *Assume that (2.3)–(2.4), (2.12), (3.12), (3.14), and (4.2) hold. Assume that σ and σ_a are constants and that Ω is a convex polyhedron. Assume that there is $\mu \in (0, \frac{1}{2})$ so that $m \in H^\mu(\partial D)$. Assume that (5.6)–(5.7) hold. Let $\psi_h^{(0)}$ be the solution to (3.15). Then, the following holds for all $s \in [0, 1/2)$:*

$$(5.8) \quad \|\psi_h^{(0)}\|_{H^s(D)} \leq c(\|\alpha\|_{L^2(\{\Omega \cdot n < 0\} \times D)} + \|m\|_{H^\mu(\partial D)} + \|q\|_{L^2(D)}).$$

Proof. Since there is $\mu \in (0, \frac{1}{2})$ so that $m \in H^\mu(\partial D)$, we construct $\phi_m \in \mathcal{C}_h$ so that $\phi_m|_{\partial D} = m$ and $\|\phi_m\|_{H^{\frac{1}{2}+\mu}(D)} \leq c\|m\|_{H^\mu(\partial D)}$, with c uniform with respect to h . This can be done by taking the Clément [3] or Scott–Zhang [15] interpolant of any $H^{\frac{1}{2}+\mu}$ -lifting of m . Let us define $\psi_{h,0}^{(0)} = \psi_h^{(0)} - \phi_m$; by construction $\psi_{h,0}^{(0)}$ is a member of $\mathcal{C}_{h,0}$. We then recast the problem (5.3) as follows:

$$(5.9) \quad b(\psi_{h,0}^{(0)}, \varphi) = r(\varphi) - b(\phi_m, \varphi) \quad \forall \varphi \in \mathcal{C}_{h,0}.$$

Let us define the operator $A_h : \mathcal{C}_{h,0} \rightarrow \mathcal{C}_{h,0}$ as follows:

$$\int_D \varphi A_h \phi \, dx = \int_D \nabla \phi \cdot \nabla \varphi \, dx \quad \forall (\phi, \varphi) \in \mathcal{C}_{h,0} \times \mathcal{C}_{h,0}.$$

Since A_h is self-adjoint and positive definite, A_h^p can be defined for all $p \in \mathbb{R}$. The following result is proved in [7]: There are $c_1(p) \in (0, +\infty)$, $c_2(p) \in (0, +\infty)$, uniform with respect to h so that for all $p \in (-\frac{3}{2}, \frac{3}{2})$

$$(5.10) \quad c_1(p)\|\varphi\|_{H^p}^2 \leq \int_D \varphi A_h^p \varphi \, dx \leq c_2(p)\|\varphi\|_{H^p}^2 \quad \forall \varphi \in \mathcal{C}_{h,0}.$$

Let $s \in [0, 1/2)$, and set $\delta := \frac{1}{2} - s$. Then taking $\phi_h := A_h^{s-1} \psi_{h,0}^{(0)}$ as a test function in (5.9) and using (5.10), we obtain

$$(5.11) \quad \begin{aligned} \frac{c_1}{3\sigma} \|\psi_{h,0}^{(0)}\|_{H^s}^2 + c_1 \sigma_a \|\psi_{h,0}^{(0)}\|_{H^{s-1}}^2 &\leq \frac{1}{3\sigma} \int_D A_h \psi_{h,0}^{(0)} A_h^{s-1} \psi_{h,0}^{(0)} \, dx + \sigma_a \int_D \psi_{h,0}^{(0)} A_h^{s-1} \psi_{h,0}^{(0)} \, dx \\ &= b(\psi_{h,0}^{(0)}, \phi_h) = r(\phi_h) - b(\phi_m, \phi_h). \end{aligned}$$

The rest of the argument consists of bounding $|r(\phi_h)| + |b(\phi_m, \phi_h)|$ from above.

We start by estimating $|b(\phi_m, \phi_h)|$. Using the fact that $H^\mu(D) = H_0^\mu(D)$, $\forall \mu \in (0, \frac{1}{2})$ (see, e.g., [12, Thm. 11.1] or [6, Cor. 1.4.4.5]), we infer

$$\begin{aligned} |b(\phi_m, \phi_h)| &\leq c \left(\|\nabla \phi_m\|_{H^{-\frac{1}{2}+\mu}(D)} \|\nabla A_h^{s-1} \psi_{h,0}^{(0)}\|_{H^{\frac{1}{2}-\mu}(D)} + \|\phi_m\|_{L^2(D)} \|\psi_{h,0}^{(0)}\|_{L^2(D)} \right) \\ &\leq c \left(\|\phi_m\|_{H^{\frac{1}{2}+\mu}(D)} \|A_h^{s-1} \psi_{h,0}^{(0)}\|_{H^{\frac{3}{2}-\mu}(D)} + \|\phi_m\|_{H^{\frac{1}{2}+\mu}(D)} \|\psi_{h,0}^{(0)}\|_{H^s(D)} \right). \end{aligned}$$

Using (5.10) (since $\frac{3}{2} - \mu < \frac{3}{2}$) and the $H^{\frac{1}{2}+\mu}$ -boundedness of ϕ_m , we deduce

$$\begin{aligned} |b(\phi_m, \phi_h)| &\leq c \|m\|_{H^\mu(\partial D)} \left(\|\psi_{h,0}^{(0)}\|_{H^{-\frac{1}{2}+2s-\mu}(D)} + \|\psi_{h,0}^{(0)}\|_{H^s(D)} \right) \\ &\leq c \|m\|_{H^\mu(\partial D)} \|\psi_{h,0}^{(0)}\|_{H^s(D)}, \quad \text{since } -\frac{1}{2} + s \leq 0 \leq \mu. \end{aligned}$$

Now we estimate $r(\phi_h)$; this is done by estimating $\|\partial_n \phi_h\|_{L^2(\partial D)}$. For this purpose we introduce the function $\phi = H_0^1(D)$, solving

$$-\Delta \phi = A_h^s \psi_{h,0}^{(0)}, \quad \phi|_{\partial D} = 0.$$

The rest of the argument consists of estimating $\|\partial_n \phi_h\|_{L^2(\partial D)}$ in terms of $\|\partial_n \phi\|_{L^2(\partial D)}$ (observe that $A_h \phi_h = A_h^s \psi_{h,0}^{(0)}$). Since Ω is a convex polyhedron, we have $\|\phi\|_{H^{3/2+\delta}} \leq c \|A_h^s \psi_{h,0}^{(0)}\|_{H^{-1/2+\delta}}$, (see Grisvard [6, Thm. 3.2.1.2]) which together with (5.10) implies the estimate

$$\|\partial_n \phi\|_{L^2(\partial D)} \leq c \|\phi\|_{H^{3/2+\delta}} \leq c' \|A_h^s \psi_{h,0}^{(0)}\|_{H^{-1/2+\delta}} \leq c'' \|\psi_{h,0}^{(0)}\|_{H^{2s-1/2+\delta}} = c'' \|\psi_{h,0}^{(0)}\|_{H^s}.$$

The connection between ϕ and ϕ_h is elucidated by observing that for all $\varphi \in \mathcal{C}_{h,0}$

$$\int_D \nabla \phi \cdot \nabla \varphi \, dx = \int_D \varphi A_h^s \psi_{h,0}^{(0)} \, dx = \int_D A_h^{s-1} \psi_{h,0}^{(0)} A_h \varphi \, dx = \int_D \nabla \phi_h \cdot \nabla \varphi \, dx.$$

This means that ϕ_h is the Galerkin approximation of ϕ in $\mathcal{C}_{h,0}$. This together with the approximability hypothesis (5.6)–(5.7) immediately implies that the following error estimate holds:

$$\|\phi - \phi_h\|_{L^2(\partial D)} \leq ch^{1+\delta} \|\phi\|_{H^{\frac{3}{2}+\delta}} \leq ch^{1+\delta} \|\psi_{h,0}^{(0)}\|_{H^s}.$$

Using again the approximability hypothesis (5.7), we now derive a bound on $\|\partial_n \phi_h\|_{L^2(\partial D)}$ as follows:

$$\begin{aligned} \|\partial_n \phi_h\|_{L^2(\partial D)} &\leq \inf_{v_h \in \mathcal{C}_{h,0}} (\|\partial_n(\phi_h - v_h)\|_{L^2(\partial D)} + \|\partial_n(v_h - \phi)\|_{L^2(\partial D)} + \|\partial_n \phi\|_{L^2(\partial D)}) \\ &\leq \inf_{v_h \in \mathcal{C}_{h,0}} (ch^{-1} \|\phi_h - v_h\|_{L^2(\partial D)} + \|\partial_n(v_h - \phi)\|_{L^2(\partial D)}) + \|\phi\|_{H^{\frac{3}{2}+\delta}(D)} \\ &\leq \inf_{v_h \in \mathcal{C}_{h,0}} (ch^{-1} \|\phi - v_h\|_{L^2(\partial D)} + \|\partial_n(v_h - \phi)\|_{L^2(\partial D)}) \\ &\quad + ch^{-1} \|\phi_h - \phi\|_{L^2(\partial D)} + \|\phi\|_{H^{\frac{3}{2}+\delta}(D)} \\ &\leq c(1 + h^{1+\delta-1}) \|\phi\|_{H^{\frac{3}{2}+\delta}(D)} \leq c' \|\psi_{h,0}^{(0)}\|_{H^s}. \end{aligned}$$

Using this estimate plus the obvious bound $\|\phi_h\|_{L^2(D)} \leq c \|\psi_{h,0}^{(0)}\|_{H^s}$, we finally derive a bound from above on $|r(\phi_h)|$ by proceeding as follows:

$$\begin{aligned} |r(\phi_h)| &\leq c(\|\frac{1}{6}m + M \cdot n\|_{L^2(\partial D)} \|\partial_n \phi_h\|_{L^2(\partial D)} + \|q\|_{L^2(D)} \|\phi_h\|_{L^2(D)}) \\ &\leq c(\|\alpha\|_{L^2(\{\Omega \cdot n < 0\} \times D)} + \|q\|_{L^2(D)}) \|\psi_{h,0}^{(0)}\|_{H^s}. \end{aligned}$$

By combining the above bounds we obtain

$$|r(\phi_h) - b(\phi_m, \psi_{h,0}^{(0)})| \leq c(\|m\|_{H^\mu(\partial D)} + \|\alpha\|_{L^2(\{\Omega \cdot n < 0\} \times D)} + \|q\|_{L^2(D)}) \|\psi_{h,0}^{(0)}\|_{H^s}.$$

The conclusion follows readily by inserting this estimate into (5.11) and using the triangle inequality. \square

Remark 5.1. The hypothesis $m \in H^\mu(\partial D)$ is not restrictive, since we just require $\mu > 0$. This amounts to assuming that m is slightly more regular than an L^2 -function over ∂D . Any discontinuous but piecewise smooth function over ∂D belongs to every $H^\mu(\partial D)$, $\mu \in (0, \frac{1}{2})$.

In the rest of the paper we characterize the limit solution of (3.15) when $h \rightarrow 0$; that is, we compute $\psi_{\text{lim}} := \lim_{h \rightarrow 0} \lim_{\varepsilon \rightarrow 0} \psi_h$.

5.3. Limit solution of (3.15) for isotropic incoming flux. If the incoming flux is isotropic, then $\frac{1}{6}m + M \cdot n = 0$ (see (3.4)). This condition turns out to be sufficient to establish the optimal convergence result stated in this subsection.

THEOREM 5.3. *Assume that α is such that $\frac{1}{6}m + M \cdot n = 0$. Assume that (2.3)–(2.4), (2.12), (3.12), (3.14), and (4.2) hold. Assume also (5.6) and $m \in H^{\frac{1}{2}}(\partial D)$; then $\psi_h^{(0)}$ converges in $H^1(D)$ to $\psi_{\text{lim}} = \psi_0$, solution of (3.2a)–(3.2b), and the following error estimate holds:*

$$(5.12) \quad \|\psi_0 - \psi_h^{(0)}\|_{H^1(D)} \leq c \inf_{v_h \in \mathcal{C}_{h,0}} \|\psi_0 - v_h\|_{H^1(D)}.$$

Proof. Construct a H^1 -lifting of $m \in H^{\frac{1}{2}}(\partial D)$, say, $\phi_m \in H^1(D)$, so that $\psi_h^{(0)} - \phi_m \in H_0^1(D)$. Then reproduce the arguments of the proof of Proposition 5.1 and conclude using Cea’s lemma. \square

The critical assumption here is (5.6), which requires the spaces \mathcal{C}_h to be rich enough so as to have reasonable approximation properties. This is a condition on the mesh family $\{\mathcal{T}_h\}_{h>0}$ and the associated discrete space family $\{\mathcal{D}_h\}_{h>0}$. More precisely (5.6) holds if the following two conditions are satisfied:

(i) The meshes are conforming; i.e., each face of a cell is either the face of a neighboring cell or at the boundary. This condition can be weakened to accommodate for local refinement, and in this case each face of any cell may be a subset of a face of its neighbor.

(ii) The polynomial spaces on each cell must allow continuity across interfaces of neighboring cells without losing approximation properties. This is usually achieved by using multidimensional polynomial spaces \mathbb{P}_k of total order $k \geq 1$ for triangles and tetrahedra or mapped tensor product spaces \mathbb{Q}_k of order $k \geq 1$ in each coordinate direction on quadrilaterals and hexahedra.

Remark 5.2. For instance, condition (ii) is violated if piecewise constant elements are used. While piecewise constant approximation is admissible for solving the transport problem (2.1), the continuity condition (3.9) forces the diffusion limit solution to be globally constant, i.e., $\psi_h^{(0)}$ does not converge to $\psi^{(0)}$, unless $\psi^{(0)}$ is constant.

Remark 5.3. Conditions (i)–(ii) have been identified in [1] and termed “locality” and “surface-matching” properties. We think though that the condition (5.6) gives a complementary rationale to that given in [1]. Lists of admissible and nonadmissible finite elements are given in Tables I and II in [1].

5.4. Limit solution of (3.15) for general incoming flux. If the incoming flux is not isotropic and more generally if $\frac{1}{6}m + M \cdot n \neq 0$, the convergence analysis is more complicated, since a boundary layer occurs. Let us consider ψ_{lim} to be the solution to the following boundary value problem:

$$(5.13a) \quad -\nabla \cdot \left(\frac{1}{3\sigma} \nabla \psi_{\text{lim}} \right) + \sigma_a \psi_{\text{lim}} = q,$$

$$(5.13b) \quad \psi_{\text{lim}}|_{\partial D} = \frac{1}{2}m - 3M \cdot n.$$

Note that ψ_{lim} solves the same PDE as $\psi^{(0)}$ (see (3.2a)), but the boundary value is slightly different; see (3.2b). The boundary condition for ψ_{lim} can be rewritten as follows:

$$\psi_{\text{lim}}|_{\partial D} = \frac{1}{2\pi} \int_{\Omega \cdot n \leq 0} \overline{W}(|\Omega \cdot n|) \alpha(\Omega, x) \, d\Omega$$

with $\overline{W}(\mu) = \mu + \frac{3}{2}\mu^2$.

The main result of this section is the following theorem.

THEOREM 5.4. *Assume that (2.3)–(2.4), (2.12), (3.12), (3.14), (4.2), and (5.6)–(5.7) hold. Assume in addition that $M \cdot n$ is the trace of a function in \mathcal{C}_h and that there is $\mu > 0$ so that $m \in H^\mu(\partial D)$, $M \cdot n \in H^\mu(\partial D)$. Then $\psi_h^{(0)}$ converges to ψ_{lim} in $H^s(D)$ for all $s \in [0, \frac{1}{2})$ and denoting by d the space dimension, the following error estimate holds:*

$$(5.14) \quad \|\psi_{\text{lim}} - \psi_h^{(0)}\|_{L^2(D)} \leq c' h^{\frac{s}{d}} \quad \forall s \in \left[0, \frac{1}{2}\right).$$

Proof. Owing to the regularity assumptions on m and $M \cdot n$, it can be shown that $\psi_{\text{lim}} \in H^{\frac{1}{2}+\mu}(D)$. Let $\phi_h \in \mathcal{C}_h$ be the Scott–Zhang interpolant of ψ_{lim} ; then $M \cdot n$ being the trace of a function in \mathcal{C}_h by hypothesis, we have

$$\|\phi_h - \psi_{\text{lim}}\|_{L^2(D)} \leq ch^{\frac{1}{2}+\mu} \|\psi_{\text{lim}}\|_{H^{\frac{1}{2}+\mu}(D)} \quad \text{and} \quad \phi_h|_{\partial D} = \frac{1}{2}m - 3M \cdot n.$$

We now try to correct the boundary value of $\psi_h^{(0)}$. Let \mathcal{B}_h be the set of elements in the mesh that touch ∂D (either by a face, an edge, or a vertex). We denote by $\mathcal{F}_h^{\mathcal{B}}$ the set of the interfaces of the elements in \mathcal{B}_h . We now define $\tilde{\psi}_h^{(0)}$ so that

$$\tilde{\psi}_h^{(0)}|_{\partial D} = \phi_h|_{\partial D} \quad \text{and} \quad \tilde{\psi}_h^{(0)}|_{D \setminus \mathcal{B}_h} = \psi_h^{(0)}|_{D \setminus \mathcal{B}_h}.$$

This function is well defined on each cell in \mathcal{B}_h by simply interpolating between the boundary data on ∂D and on $D \setminus \mathcal{B}_h$. By using standard estimates, it follows that

$$\begin{aligned} \sum_{K \in \mathcal{B}_h} \|\tilde{\psi}_h^{(0)}\|_{L^2(K)}^2 &\leq c \sum_{K \in \mathcal{B}_h} \left(\|\phi_h\|_{L^2(K)}^2 + \|\psi_h^{(0)}\|_{L^2(K)}^2 \right), \\ \sum_{F \in \mathcal{F}_h^{\mathcal{B}}} \|\tilde{\psi}_h^{(0)}\|_{L^2(F)}^2 &\leq c \sum_{F \in \mathcal{F}_h^{\mathcal{B}}} \left(\|\phi_h\|_{L^2(F)}^2 + \|\psi_h^{(0)}\|_{L^2(F)}^2 \right). \end{aligned}$$

Let us denote $e := \phi_h - \tilde{\psi}_h^{(0)}$. Then the following holds:

$$\begin{aligned} b(e, \varphi_h) &= b(\phi_h - \psi_{\text{lim}}, \varphi_h) + b(\psi_{\text{lim}}, \varphi_h) + b(\psi_h^{(0)} - \tilde{\psi}_h^{(0)}, \varphi_h) - b(\psi_h^{(0)}, \varphi_h) \\ &= b(\phi_h - \psi_{\text{lim}}, \varphi_h) + b(\psi_h^{(0)} - \tilde{\psi}_h^{(0)}, \varphi_h) - \int_{\partial D} \frac{1}{3\sigma} \left(\frac{1}{2}m + 3M \cdot n \right) \partial_n \varphi_h \, dx \end{aligned}$$

$\forall \varphi_h$ in $\mathcal{C}_{h,0}$. The term $b(\psi_h^{(0)} - \tilde{\psi}_h^{(0)}, \varphi_h)$ is expanded as follows:

$$\begin{aligned} b(\psi_h^{(0)} - \tilde{\psi}_h^{(0)}, \varphi_h) &= \sum_{K \in \mathcal{B}_h} \int_K \left(\frac{1}{3\sigma} \nabla(\psi_h^{(0)} - \tilde{\psi}_h^{(0)}) \cdot \nabla \varphi_h + \sigma_a(\psi_h^{(0)} - \tilde{\psi}_h^{(0)}) \varphi_h \right) \, dx \\ &= \sum_{K \in \mathcal{B}_h} \int_K (\psi_h^{(0)} - \tilde{\psi}_h^{(0)}) \left(\sigma_a \varphi_h - \frac{1}{3\sigma} \Delta \varphi_h \right) \, dx \\ &\quad + \int_{\partial D} \frac{1}{3\sigma} \left(\frac{1}{2}m + 3M \cdot n \right) \partial_n \varphi_h \, dx \\ &\quad + \sum_{F \in \mathcal{F}_h^{\mathcal{B}}} \int_F \frac{1}{3\sigma} (\psi_h^{(0)} - \tilde{\psi}_h^{(0)}) n_1 \cdot [\nabla \varphi_h]. \end{aligned}$$

We finally have

$$(5.15) \quad b(e, \varphi_h) = R_1(\varphi_h) + R_2(\varphi_h) + R_3(\varphi_h),$$

with $R_1(\varphi_h) := b(\phi_h - \psi_{\text{lim}}, \varphi_h)$ and

$$\begin{aligned} R_2(\varphi_h) &= \sum_{K \in \mathcal{B}_h} \int_K (\psi_h^{(0)} - \tilde{\psi}_h^{(0)}) \left(\sigma_a \varphi_h - \frac{1}{3\sigma} \Delta \varphi_h \right) dx, \\ R_3(\varphi_h) &= \sum_{F \in \mathcal{F}_h^{\text{E}}} \int_F \frac{1}{3\sigma} (\psi_h^{(0)} - \tilde{\psi}_h^{(0)}) n_1 \cdot [\nabla \varphi_h]. \end{aligned}$$

We now choose $\varphi_h = A_h^{-1}e$ to test (5.15) (observe that this is legitimate, since e is a member of $\mathcal{C}_{h,0}$, which was not the case of $\phi_h - \psi_h^{(0)}$), and, owing to (5.10), we obtain

$$\frac{c}{3\sigma} \|e_h\|_{L^2(D)}^2 + c' \sigma_a \|e_h\|_{H^{-1}(D)}^2 \leq R_1(\varphi_h) + R_2(\varphi_h) + R_3(\varphi_h).$$

We define $\varphi \in H_0^1(D)$ to be the solution of the following problem:

$$-\Delta \varphi = A_h \varphi_h = e, \quad \varphi|_{\partial D} = 0.$$

Since D is a convex polyhedron, we have $\|\varphi\|_{H^2(D)} \leq c\|e\|_{L^2(D)}$. The definition of φ implies that φ_h is the Galerkin approximation of φ , and the following estimates hold:

$$\begin{aligned} \sum_{F \in \mathcal{F}_h^i} \|\partial_n(\varphi_h - \varphi)\|_{L^2(F)}^2 &\leq ch\|e\|_{L^2(D)}^2, \\ \sum_{K \in \mathcal{T}_h} \left(\|\nabla(\varphi - \varphi_h)\|_{L^2(K)}^2 + h^2 \|\Delta \varphi_h\|_{L^2(K)}^2 \right) &\leq ch^2 \|e\|_{L^2(D)}^2. \end{aligned}$$

We handle $R_1(\varphi_h)$ as follows:

$$\begin{aligned} R_1(\varphi_h) &= \sum_{K \in \mathcal{T}_h} \int_K (\phi_h - \psi_{\text{lim}}) \left(\sigma_s \varphi_h - \frac{1}{3\sigma} \Delta \varphi_h \right) dx \\ &\quad + \sum_{F \in \mathcal{F}_h^i} \int_F \frac{1}{3\sigma} (\phi_h - \psi_{\text{lim}}) n_1 \cdot [\nabla \varphi_h] := R_{11}(\varphi_h) + R_{12}(\varphi_h). \end{aligned}$$

The term $R_{11}(\varphi_h)$ is controlled as follows:

$$\begin{aligned} R_{11}(\varphi_h)^2 &\leq c \sum_{K \in \mathcal{T}_h} \left(\|\phi_h - \psi_{\text{lim}}\|_{L^2(K)}^2 \right) \sum_{K \in \mathcal{T}_h} \left(\|\varphi_h\|_{L^2(K)}^2 + \|\Delta \varphi_h\|_{L^2(K)}^2 \right) \\ &\leq ch^{2(\frac{1}{2}+\mu)} \|\psi_{\text{lim}}\|_{H^{\frac{1}{2}+\mu}(D)}^2 \|e\|_{L^2(D)}^2. \end{aligned}$$

Similarly the term $R_{12}(\varphi_h)$ is controlled as follows:

$$\begin{aligned} R_{12}(\varphi_h)^2 &\leq c \sum_{F \in \mathcal{F}_h^i} \left(\|\phi_h - \psi_{\text{lim}}\|_{L^2(F)}^2 \right) \sum_{F \in \mathcal{F}_h^{\text{E}}} \|\llbracket \nabla(\varphi_h - \varphi) \rrbracket\|_{L^2(F)}^2 \\ &\leq ch^{2\mu} \|\psi_{\text{lim}}\|_{H^{\frac{1}{2}+\mu}(D)}^2 h \|e\|_{L^2(D)}^2 \leq ch^{2(\frac{1}{2}+\mu)} \|\psi_{\text{lim}}\|_{H^{\frac{1}{2}+\mu}(D)}^2 \|e\|_{L^2(D)}^2. \end{aligned}$$

We handle $R_2(\varphi_h)$ as follows:

$$\begin{aligned} R_2(\varphi_h)^2 &\leq c \sum_{K \in \mathcal{B}_h} \left(\|\psi_h^{(0)}\|_{L^2(K)}^2 + \|\tilde{\psi}_h^{(0)}\|_{L^2(K)}^2 \right) \sum_{K \in \mathcal{B}_h} \left(\|\varphi_h\|_{L^2(K)}^2 + \|\Delta\varphi_h\|_{L^2(K)}^2 \right) \\ &\leq c \|e\|_{L^2(D)}^2 \sum_{K \in \mathcal{B}_h} \left(\|\psi_h^{(0)}\|_{L^2(K)}^2 + \|\phi_h\|_{L^2(K)}^2 \right) \\ &\leq c \|e\|_{L^2(D)}^2 \left(\sum_{K \in \mathcal{B}_h} |K| \right)^{\frac{p-2}{p}} \left(\|\psi_h^{(0)}\|_{L^p(D)}^2 + \|\phi_h\|_{L^p(D)}^2 \right), \end{aligned}$$

where p is chosen so that the continuous embedding $H^s(D) \subset L^p(D)$ holds, where s is a number in $(0, \frac{1}{2})$. In d space dimensions, $p = \frac{2d}{d-2s}$. For instance, in three space dimensions, $p = \frac{6}{3-2s} \in (2, 3)$, and in two space dimensions $p = \frac{2}{1-s} \in (2, 4)$. In conclusion

$$|R_2(\varphi_h)| \leq ch^{\frac{p-2}{2p}} \|e\|_{L^2(D)} \left(\|\psi_h^{(0)}\|_{H^s(D)} + \|\phi_h\|_{H^s(D)} \right).$$

Then owing to Lemma 5.2 (see estimate (5.8)) and using $\|\phi_h\|_{H^s(D)} \leq c\|\phi_{\text{lim}}\|_{H^s(D)}$, we infer

$$|R_2(\varphi_h)| \leq ch^{\frac{s}{d}} \|e\|_{L^2(D)}.$$

We handle $R_3(\varphi_h)$ as follows:

$$\begin{aligned} R_3(\varphi_h)^2 &\leq c \sum_{F \in \mathcal{F}_h^{\mathcal{E}}} \left(\|\psi_h^{(0)}\|_{L^2(F)}^2 + \|\tilde{\psi}_h^{(0)}\|_{L^2(F)}^2 \right) \sum_{F \in \mathcal{F}_h^{\mathcal{E}}} \|\llbracket \nabla(\varphi_h - \varphi) \rrbracket\|_{L^2(F)}^2 \\ &\leq ch^{-1} \sum_{K \in \mathcal{B}_h} \left(\|\psi_h^{(0)}\|_{L^2(K)}^2 + \|\tilde{\psi}_h^{(0)}\|_{L^2(K)}^2 \right) h \|e\|_{L^2(D)}^2. \end{aligned}$$

Then using the same arguments as those for $R_2(\varphi_h)$, we obtain

$$|R_3(\varphi_h)| \leq ch^{\frac{p-2}{2p}} \|e\|_{L^2(D)} = ch^{\frac{s}{d}} \|e\|_{L^2(D)}.$$

In conclusion we have

$$\|e\|_{L^2(D)} \leq c(h^{\frac{s}{d}} + h^{\frac{1}{2}+\mu}) \leq c'h^{\frac{s}{d}}.$$

Then, using the triangle inequality

$$\begin{aligned} \|\psi_{\text{lim}} - \psi_h^{(0)}\|_{L^2(D)} &\leq \|\psi_{\text{lim}} - \phi_h\|_{L^2(D)} + \|\phi_h - \tilde{\psi}_h^{(0)}\|_{L^2(D)} + \|\tilde{\psi}_h^{(0)} - \psi_h^{(0)}\|_{L^2(D)} \\ &\leq c(h^{\frac{1}{2}+\mu} \|\psi_{\text{lim}}\|_{H^{\frac{1}{2}+\mu}(D)} + \|e\|_{L^2(D)} + h^{\frac{s}{d}}) \leq c'h^{\frac{s}{d}}. \end{aligned}$$

This concludes the proof of (5.14) and shows that $\psi_h^{(0)} \rightarrow \psi_{\text{lim}}$ strongly in $L^2(D)$ as $\varepsilon \rightarrow 0$. That $\psi_h^{(0)} \rightarrow \psi_{\text{lim}}$ in $H^s(D)$ strong, for all $s \in [0, \frac{1}{2})$, follows from the boundedness in $H^{s'}(D)$ and the compact injection $H^{s'}(D) \subset H^s(D)$ for all s' such that $0 \leq s < s' < \frac{1}{2}$; see (5.8). \square

We finish this section by observing that (5.13) can be formulated in a very weak form as follows: Find $\psi_{\text{lim}} \in L^2(D)$ so that for all $\varphi \in H := \{v \in H_0^1(D), \nabla \cdot \frac{1}{3\sigma} \nabla v \in L^2(D)\}$,

$$(5.16) \quad \int_D \psi_{\text{lim}} \left(\sigma_a \varphi - \nabla \cdot \frac{1}{3\sigma} \nabla \varphi \right) dx = \int_D q \varphi dx - \int_{\partial D} \frac{1}{3\sigma} \left(\frac{1}{2} m - 3M \cdot n \right) \partial_n \varphi dx.$$

The fact that the above problem is well-posed is a consequence of $-\Delta : H \rightarrow L^2(D)$ being an isomorphism.

5.5. Boundary condition incompatibilities. The convergence rate given in Theorem 5.4 is very slow; it is $\mathcal{O}(h^{\frac{1}{4}})$ and $\mathcal{O}(h^{\frac{1}{6}})$ in two and three space dimensions, respectively. Whether the exponent in the estimate (5.14) is sharp is unclear to us at the moment. The fact that we have convergence in a norm weaker than that of $H^{\frac{1}{2}}(\partial D)$ is due to incompatible boundary conditions. Observe that for every h , $\psi_h^{(0)}|_{\partial D} = m$, but $\psi_{\text{lim}}|_{\partial D} = \frac{1}{2}m - 3M \cdot n$; i.e., the boundary conditions of $\psi_h^{(0)}$ and ψ_{lim} are incompatible. This observation also implies that $\psi_h^{(0)}$ cannot converge to ψ_{lim} in any norm stronger than that of $H^{\frac{1}{2}}(\partial D)$, thus showing that the a priori estimate (5.8) is the best that can be obtained for incoming fluxes such that $\frac{1}{6}m + M \cdot n \neq 0$.

Finally, note that ψ_{lim} and $\psi^{(0)}$ solve the same PDE but satisfy two different boundary conditions, implying that

$$(5.17) \quad \lim_{h \rightarrow 0} \lim_{\varepsilon \rightarrow 0} \psi_h := \psi_{\text{lim}} \neq \psi^{(0)} \quad \text{unless } \frac{1}{6}m + M \cdot n = 0.$$

For any practical purpose the above result says that unless $\frac{1}{6}m + M \cdot n = 0$, the approximate solution of (2.19), ψ_h , may not be close to $\psi^{(0)}$ if the mesh size is significantly larger than the mean free path, even if the mesh size is small.

As observed in [1] though, the above negative conclusion is moderated by the fact that the boundary values of ψ_{lim} and $\psi^{(0)}$ are very close for all practical purposes. Actually, it can be shown that $W(\mu) \approx \mu + \frac{3}{2}\mu^2, \forall \mu \in [0, 1]$, and the difference $|W(\mu) - (\mu + \frac{3}{2}\mu^2)|$ is a few percents in the maximum norm over $[0, 1]$ (see [1]); as a result, the following approximate identity holds:

$$(5.18) \quad \psi_{\text{lim}}|_{\partial D} = \frac{1}{2\pi} \int_{\Omega \cdot n \leq n} \left(|\Omega \cdot n| + \frac{3}{2} |\Omega \cdot n|^2 \right) \alpha(\Omega, x) d\Omega \approx \psi^{(0)}|_{\partial D}.$$

In other words ψ_{lim} and $\psi^{(0)}$ are different but differ by a few percents only.

In conclusion, if the incoming flux is not isotropic, a numerical boundary layer occurs when the mesh size is significantly larger than the mean free path, but the interior approximation is not too far from the correct limit.

6. Conclusions. By using functional analytic tools, we confirmed the results in [1], namely, that in the limit of vanishing mean free path length, the upwind DG approximation of the radiative transfer problem yields a mixed discretization of the diffusion equation. A key feature of this limit is that the discrete primal variable is continuous. This property implies that grid convergence can be achieved in optically thick regions only if the DG approximation space contains a linear space of piecewise linear continuous functions with optimal interpolation properties. Under this condition, the solution of the discrete diffusion problem converges in $H^1(D)$ to the diffusion solution if the incoming flux is isotropic. If the incoming flux is not isotropic,

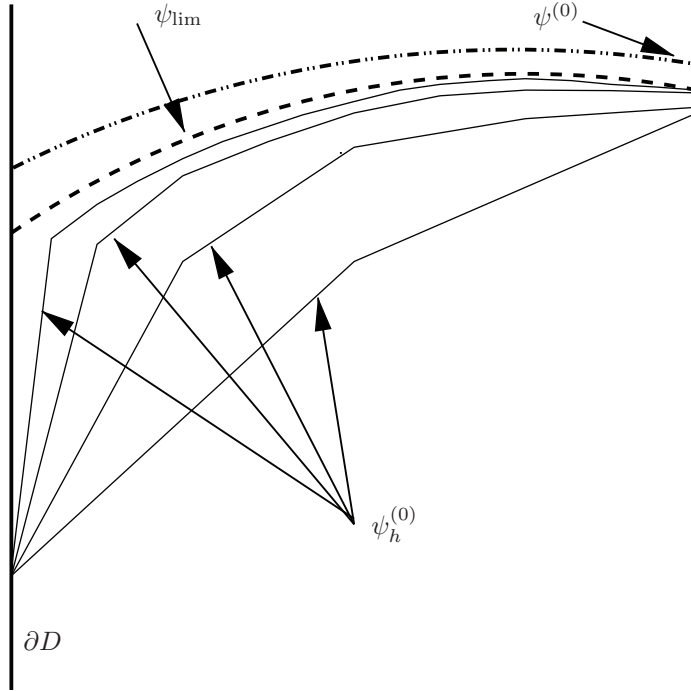


FIG. 6.1. Illustration of the conclusions of this paper when $\frac{1}{6}m + M \cdot n \neq 0$. Schematic representation of the graphs of $\psi^{(0)}$, ψ_{lim} , and $\psi_h^{(0)}$ for several h in a cross section of D . $\psi^{(0)}$ is the diffusion limit solving (3.2), $\psi_h^{(0)}$ is the discrete diffusion limit solving (3.15), and $\psi_{\text{lim}} = \lim_{h \rightarrow 0} \psi_h^{(0)}$ solves (5.13). ψ_{lim} and $\psi^{(0)}$ solve the same PDE, but the boundary conditions are different: $\psi^{(0)}|_{\partial D} = \int W(|\Omega \cdot n|) \frac{\alpha}{2\pi} d\Omega$, $\psi_{\text{lim}}|_{\partial D} = \frac{1}{2}m - 3M \cdot n$, and $\psi_h^{(0)}|_{\partial D} = m$ for any h . In general, $\psi_{\text{lim}} \neq \psi^{(0)}$ unless $\frac{1}{6}m + M \cdot n = 0$. $\psi_h^{(0)}$ converges to ψ_{lim} as $h \rightarrow 0$ except in a discrete boundary layer. There is no boundary layer when $\frac{1}{6}m + M \cdot n = 0$, and, in this case, $\psi_{\text{lim}} = \psi^{(0)}$.

the discrete solution converges in $H^s(D)$, with $s < \frac{1}{2}$, to a function which is close to the diffusion limit. A boundary layer effect occurs, and the convergence holds only in the interior of the domain.

A schematic representation of the situation is shown in Figure 6.1.

Appendix A. Assumption (3.12) may seem to be quite restrictive. For instance, it does not hold if \mathcal{S}_h is composed of piecewise constant functions, which is a commonly used approximation. As an alternative to (3.12), we assume that the L^2 -projection operator $p_h : L^2(S^2; \mathbb{R}^3) \rightarrow (\mathcal{S}_h)^3$ is such that

$$(A.1) \quad \int_{S^2} p_h(\Omega) d\Omega = 0,$$

$$(A.2) \quad \int_{\{\Omega \cdot a < 0\}} p_h(\Omega) \otimes p_h(\Omega) d\Omega = \int_{\{\Omega \cdot a > 0\}} p_h(\Omega) \otimes p_h(\Omega) d\Omega \quad \forall a \in \mathbb{R}^3.$$

These two conditions hold, provided the mesh defining \mathcal{S}_h is symmetric with respect to the origin. We are now going to show that everything which is said in sections 3.3 and 4 holds up to minor modifications.

Lemma 3.5 must be reworked a little. Equation (3.13a) is unchanged, since the proof of this statement does not depend on assumption (3.12). Equation (3.13b), however, must be changed. In Step 3 of the proof of Lemma 3.5 we must use $w(\omega, x) = p_h(\Omega) \cdot e_i$ to test (3.7) instead of $\Omega \cdot e_i$, since $\Omega \cdot e_i$ may not be in \mathcal{S}_h and the integral identity (3.10) does not hold. Nevertheless, we can define a symmetric 3×3 matrix I_h by

$$(A.3) \quad \frac{4\pi}{3} I_h := \int_{S^2} p_h(\Omega) \otimes p_h(\Omega) \, d\Omega = \int_{S^2} \Omega \otimes p_h(\Omega) \, d\Omega.$$

If we assume that the angular discretization is sufficiently fine, I_h approximates the identity. In particular, it is positive definite.

When we test (3.7) with $p_h(\Omega) \cdot e_i$, the above definition implies that the first term becomes

$$\begin{aligned} \int_{S^2 \times K} w \Omega \cdot \nabla \psi_h^{(0)} \, d\Omega \, dx &= \int_K \varphi \partial_j \psi_h^{(0)} \, dx \int_{S^2} (\Omega \cdot e_j) (p_h(\Omega) \cdot e_i) \, d\Omega \\ &= \frac{4\pi}{3} \int_K \varphi e_i^T I_h \nabla \psi_h^{(0)} \, dx. \end{aligned}$$

Second, we compute the contribution of $\psi_h^{(1)}$,

$$\int_{S^2 \times K} \left(\psi_h^{(1)} - \bar{\psi}_h^{(1)} \right) w \sigma \, d\Omega \, dx = \int_{S^2 \times K} \psi_h^{(1)} (p_h(\Omega) \cdot e_i) \varphi \sigma \, d\Omega \, dx = \int_K (J_h^{(0)} \cdot e_i) \varphi \sigma \, dx,$$

where we used $\int_{S^2} \psi_h^{(1)} p_h(\Omega) \, d\Omega = \int_{S^2} \psi_h^{(1)} \Omega \, d\Omega$, since p_h is the L^2 -projection and $\mathcal{W}_h = \mathcal{S}_h \otimes \mathcal{D}_h$ is a tensor product space. Third, we take care of the flux terms. Since $\psi_h^{(0)}$ is a member of $\mathcal{C}_{h,m}$, the flux term in the left-hand side gives

$$\sum_{F \in \mathcal{F}_h^\partial} \int_F \left(\int_{\Omega \cdot n \leq 0} |\Omega \cdot n| (p_h(\Omega) \cdot e_i) \, d\Omega \right) m \varphi \, dx = -\frac{2\pi}{3} \sum_{F \in \mathcal{F}_h^\partial} \int_F (n^T I_h e_i) m \varphi \, dx,$$

where we used (A.2). Similarly, for the flux term in the right-hand side, we obtain

$$\sum_{F \in \mathcal{F}_h^\partial} \int_F \left(\int_{\Omega \cdot n < 0} |\Omega \cdot n| (\Omega \cdot e_i) \alpha(\Omega, x) \, d\Omega \right) \varphi \, dx = \sum_{F \in \mathcal{F}_h^\partial} \int_F 4\pi (M_h \cdot e_i) \varphi \, dx,$$

where we have defined the following approximation of M :

$$(A.4) \quad M_h(x) := \frac{1}{4\pi} \int_{\Omega \cdot n(x) < 0} \alpha(\Omega, x) |\Omega \cdot n(x)| p_h(\Omega) \, d\Omega.$$

We now combine all the above results, and we obtain the new equation replacing (3.13b):

$$\sum_{K \in \mathcal{T}_h} \int_K \left(\frac{4\pi}{3} I_h \nabla \psi_h^{(0)} + \sigma J_h^{(0)} \right) \cdot e_i \varphi \, dx = \sum_{F \in \mathcal{F}_h^\partial} \int_F 4\pi \left(\frac{1}{6} m(I_h n) + M_h \right) \cdot e_i \varphi \, dx.$$

We are now in position to state the counterpart of Proposition 3.6.

PROPOSITION A.1. *Assume that (2.3), (2.12), (3.14), and (A.1)–(A.2) hold. Then $\psi_h^{(0)} \in \mathcal{C}_{h,m}$ solves the following problem: For all $\varphi \in \mathcal{C}_{h,0}$,*

$$(A.5) \quad \int_D \left(\frac{1}{3\sigma} \left(\nabla \psi_h^{(0)} \right)^T I_h \nabla \varphi + \sigma_a \psi_h^{(0)} \varphi \right) dx \\ = \int_{\partial D} \frac{1}{\sigma} \left(\frac{m}{6} (n^T I_h n) + M_h \cdot n \right) \cdot \partial_n \varphi dx + \int_D q \varphi dx.$$

The statements (3.15) and (A.5) differ only by the presence of the approximate identity I_h and the approximation of M_h . Since these differences are nonessential, we retain assumption (3.12) in sections 4 and 5.

Acknowledgments. We thank Marvin L. Adams and J. Morel for bringing to our attention the subject under investigation in this paper. Helpful discussions with Jean Ragusa are also acknowledged.

REFERENCES

- [1] M.L. ADAMS, *Discontinuous finite element methods in thick diffusive problems*, Nucl. Sci. Engrg., 137 (2001), pp. 298–333.
- [2] S. CHANDRASEKHAR, *Radiative Transfer*, Oxford University Press, London, 1950.
- [3] P. CLÉMENT, *Approximation by finite element functions using local regularization*, RAIRO, Anal. Numer., 9 (1975), pp. 77–84.
- [4] B. COCKBURN AND C.W. SHU, *The local discontinuous Galerkin method for time-dependent convection-diffusion systems*, SIAM J. Numer. Anal., 35 (1998), pp. 2440–2463.
- [5] R. DAUTRAY AND J.-L. LIONS, *Mathematical Analysis and Numerical Methods for Science and Technology*, Vol. 5, Springer, 2000.
- [6] P. GRISVARD, *Elliptic Problems in Nonsmooth Domains*, Monogr. Stud. Math. 24, Pitman, Boston, MA, 1985.
- [7] J.-L. GUERMOND AND J. PASCIAK, *Stability of discrete Stokes operators in fractional Sobolev spaces*, J. Math. Fluid Mech., 10 (2008), pp. 588–610.
- [8] G. KANSCHAT, *Parallel and Adaptive Galerkin Methods for Radiative Transfer Problems*, Dissertation, Universität Heidelberg, Heidelberg, 1996; available online at <http://www.ub.uni-heidelberg.de/archiv/6331>.
- [9] E.W. LARSEN AND J.B. KELLER, *Asymptotic solution of neutron transport problems for small mean free paths*, J. Math. Phys., 15 (1974), pp. 75–81.
- [10] E.W. LARSEN, J.E. MOREL, AND W.F. MILLER, JR., *Asymptotic solutions of numerical transport problems in optically thick, diffusive regimes*, J. Comput. Phys., 69 (1987), pp. 283–324.
- [11] P. LESAINTE AND P.-A. RAVIART, *On a finite element method for solving the neutron transport equation*, in *Mathematical Aspects of Finite Elements in Partial Differential Equations*, C. de Boors, ed., Academic Press, New York, 1974, pp. 89–123.
- [12] J.-L. LIONS AND E. MAGENES, *Problèmes aux Limites non Homogènes et Applications*, Vol. 1, Dunod, Paris, 1968.
- [13] F. MALVAGI AND G.C. POMRANING, *Initial and boundary conditions for diffusive linear transport problems*, J. Math. Phys., 32 (1991), pp. 805–820.
- [14] W. REED AND T. HILL, *Triangular Mesh Methods for the Neutron Transport Equation*, Technical report LA-UR-73-479, Los Alamos Scientific Laboratory, Los Alamos, NM, 1973.
- [15] R.L. SCOTT AND S. ZHANG, *Finite element interpolation of nonsmooth functions satisfying boundary conditions*, Math. Comp., 54 (1990), pp. 483–493.