THE DIFFUSION APPROXIMATION
FOR THE LINEAR BOLTZMANN EQUATION
WITH VANISHING SCATTERING COEFFICIENT[*]

CLAUDE BARDOS[†], ETIENNE BERNARD[‡], FRANÇOIS GOLSE[§], AND RÉMI SENTIS[¶]

Pour George Papanicolaou, en témoignage de gratitude et d'amitié.

Abstract. The present paper discusses the diffusion approximation of the linear Boltzmann equation in cases where the collision frequency is not uniformly large in the spatial domain. Our results apply for instance to the case of radiative transfer in a composite medium with optically thin inclusions in an optically thick background medium. The equation governing the evolution of the approximate particle density coincides with the limit of the diffusion equation with infinite diffusion coefficient in the optically thin inclusions.

Key words. Linear Boltzmann equation, diffusion approximation, neutron transport equation, radiative transfer equation.

AMS subject classifications. 45M05, 82A70, 82C70, 85A25.

1. Introduction

The linear Boltzmann equation is a kinetic model used in many different contexts. It appeared (perhaps for the first time) in a paper by Lorentz [23] on the motion of electrons in metals, and has been since then used in various branches of mathematical physics such as radiative transfer [24, 26] and neutron transport theory [28]. A typical model linear Boltzmann equation is

$$(\partial_t + \omega \cdot \nabla_x) f(t,x,\omega) = \sigma \left(\frac{1}{4\pi} \int_{\mathbf{S}^2} f(t,x,\omega') ds(\omega') - f(t,x,\omega) \right) \tag{1.1}$$

where $t \geq 0$, $x \in \mathbf{R}^3$, and $\omega \in \mathbf{S}^2$ are respectively the time, position, and direction, while ds is the surface element on $\mathbf{S}^2$. The function $f \equiv f(t,x,\omega)$ is the distribution function of a population of monokinetic particles (such as photons) moving at speed 1 in a background medium (such as a planetary or stellar atmosphere). The coefficient $\sigma > 0$ is the scattering rate in the medium. The simple model above assumes that the scattering mechanism is isotropic and that the absorption and scattering rate are equal. (The model above is essentially equation (4.216) in [26].)

A classical approximation of solutions f of the linear Boltzmann equation above is the so-called "P1 approximation", where f is replaced by its spherical harmonics expansion in the angle variable ω, truncated at order 1 (see chapter IX in [28]). This approximation is also known as the "Eddington approximation" in radiative transfer (see for instance chapter III.2 in [26]). The P1 approximation is

$$f(t,x,\omega) \simeq \rho(t,x) + 3j(t,x) \cdot \omega,$$

so that

$$\rho(t,x) \simeq \tfrac{1}{4\pi} \int_{\mathbf{S}^2} f(t,x,\omega')ds(\omega') \quad \text{and} \quad j(t,x) \simeq \tfrac{1}{4\pi} \int_{\mathbf{S}^2} \omega' f(t,x,\omega')ds(\omega').$$

This approximation is used in regimes where the scattering coefficient $\sigma \gg 1$. In other words, the diffusion approximation is justified when the mean free path of the particles between scattering events is much smaller than the typical length scale of the spatial domain. In this case $f \simeq \rho$ to leading order, so that $j = -\tfrac{1}{3\sigma}\nabla_x\rho$. Thus, the P1 approximation becomes

$$f(t,x,\omega) \simeq \rho(t,x) - \tfrac{1}{\sigma}\omega \cdot \nabla_x\rho(t,x). \tag{1.2}$$

Averaging both sides of the linear Boltzmann equation in the variable ω leads to the local conservation law of mass (or particle number)

$$\partial_t \int_{\mathbf{S}^2} f(t,x,\omega')\tfrac{ds(\omega')}{4\pi} + \text{div}_x \int_{\mathbf{S}^2} \omega' f(t,x,\omega')\tfrac{ds(\omega')}{4\pi} = 0.$$

Substituting the P1 approximate expression for f in the left hand side of this equality results in the diffusion equation

$$\partial_t\rho - \text{div}_x\left(\tfrac{1}{3\sigma}\nabla_x\rho\right) = 0. \tag{1.3}$$

While the P1 or diffusion approximation of the linear Boltzmann equation has been used for a long time, its mathematical justification is more recent. A proof based on Hilbert's expansion [20], a formal expansion of the solution f of the linear Boltzmann equation in powers of $1/\sigma$, can be found in [5]; see also chapter XXI.5 in [11].

Other proofs are based on the representation of the solution f of the linear Boltzmann equation in terms of a stochastic process: see for instance [25] and the references therein. For the original contributions to the subject, see [6] and [18].

The diffusion approximation (1.3) of the linear Boltzmann equation (1.1) is also used for inhomogeneous media where the scattering rate $\sigma \equiv \sigma(x)$ varies smoothly with the position variable. Yet the Cauchy problems for the linear Boltzmann equation (1.1) and for the diffusion equation (1.3) are well posed, in some weak sense to be discussed below, under the only assumption that both σ and $1/\sigma$ belong to $L^\infty(\mathbf{R}^3)$. However both the proof of the diffusion approximation based on the Hilbert expansion [5] and the proof based on stochastic processes [25] require that the scattering rate σ satisfy rather stringent smoothness assumptions. In particular, these smoothness assumptions exclude scattering rates that are discontinuous functions of the position variable. This is unfortunate, since discontinuities in the scattering rate $\sigma(x)$ appear in the case of inhomogeneous or composite materials.

Another related issue is the order of magnitude of the scattering rate. As recalled above, the diffusion approximation (1.3) of the linear Boltzmann equation (1.1) is justified if the scattering rate $\sigma(x)$ depends smoothly on the space variable x and is large uniformly in x. In some applications involving strongly inhomogeneous media, the order of magnitude of the scattering rate may vary considerably in the spatial domain. For instance, in the context of neutron transport and nuclear reactor design, the scattering cross-section for neutron collisions of non-fission type in the uranium oxide is about 100 times higher than in water. Another example is the case of nondiffusive objects embedded in a diffusive medium in the context of medical imaging: see [2], especially the examples given on p. 1678.

The present work addresses the following problem.

PROBLEM: to extend the validity of the diffusion approximation to cases where the scattering coefficient $\sigma \equiv \sigma(x)$ is neither continuous nor large uniformly as x runs through the spatial domain.

The formulation of this problem obviously includes cases where there is no separation of scale in the size of the scattering coefficient in the diffusive and nondiffusive regions. Asymptotic expansions "à la Hilbert" cannot be used in such generality.

Our approach is based instead on an energy method similar to the one used earlier in the derivation of the Rosseland equation from the radiative transfer equation [4], or of the drift-diffusion equation from the Boltzmann equation with Fermi-Dirac statistics [15], or even in the derivation of incompressible fluid dynamics from the Boltzmann equation of the kinetic theory of gases [3].

The main results in this paper were presented by one of us (C.B.) to the conference "Recent developments in applied mathematics" in honor of G. Papanicolaou. Our own work on the various asymptotic theories of kinetic models, including the diffusion approximation of the linear Boltzmann equation, the Rosseland approximation of the radiative transfer equation or the hydrodynamic limits of the Boltzmann equation in the kinetic theory of gases, was strongly influenced by G. Papanicolaou's remarkable contributions [7, 25] to this subject. We are happy to dedicate this paper to our friend and colleague G. Papanicolaou on the occasion of his 70th birthday.

2. Presentation of the problem and main result

2.1. The linear Boltzmann equation and the scaling assumptions.
Consider the linear Boltzmann equation

$$(\partial_t + v \cdot \nabla_x) f(t,x,v) + \mathcal{L}_x f(t,x,v) = 0 \tag{2.1}$$

for the unknown $f \equiv f(t,x,v)$ that is the distribution function for a system of identical point particles interacting with some background material. In other words, $f(t,x,v)$ is the number density of particles located at the position $x \in \Omega$, with velocity $v \in \mathbf{R}^N$ at time $t \geq 0$. Henceforth, we assume that Ω is a bounded domain of $\mathbf{R}^N$ with C^1 boundary $\partial\Omega$, and that Ω is locally on one side of $\partial\Omega$.

The notation $\mathcal{L}_x$ designates a linear integral operator acting on the v variable in f, i.e.

$$\mathcal{L}_x f(t,x,v) = \int_{\mathbf{R}^N} k(x,v,w)(f(t,x,v) - f(t,x,w)) d\mu(w) \tag{2.2}$$

where μ is a Borel probability measure on $\mathbf{R}^N$, while k is a nonnegative function defined $\mu \otimes \mu$-a.e. on $\mathbf{R}^N \times \mathbf{R}^N$. We assume that k satisfies the semi-detailed balance condition

$$\int_{\mathbf{R}^N} k(x,v,w) d\mu(w) = \int_{\mathbf{R}^N} k(x,w,v) d\mu(w) \tag{2.3}$$

and introduce the notation

$$a(x,v) := \int_{\mathbf{R}^N} k(x,v,w) d\mu(w) \tag{2.4}$$

for the scattering rate, so that the kernel $k(x,v,w)$, up to the multiplicative coefficient $1/a$, measures the probability of a transition from velocity w to velocity v for particles located at the position x. Eventually one has:

$$\mathcal{L}_x f(t,x,v) = a(x,v) f(t,x,v) - \mathcal{K}_x f(t,x,v)$$

where $\mathcal{K}_x$ designates the integral operator

$$\mathcal{K}_x f(t,x,v) := \int_{\mathbf{R}^N} k(x,v,w) f(t,x,w) d\mu(w). \tag{2.5}$$

The semi-detailed balance assumption appears for instance in [19] — see formula (2.9) in §2. The assumptions on the transition kernel other than (2.3) used in our discussion are introduced later.

Denoting by n_x the unit outward normal field at $x \in \partial\Omega$, we henceforth consider the outgoing, characteristic, and incoming components of $\partial\Omega \times \mathbf{R}^N$:

$$\Gamma_+ := \{(x,v) \in \partial\Omega \times \mathbf{R}^N \,|\, v \cdot n_x > 0\},$$
$$\Gamma_0 := \{(x,v) \in \partial\Omega \times \mathbf{R}^N \,|\, v \cdot n_x = 0\},$$
$$\Gamma_- := \{(x,v) \in \partial\Omega \times \mathbf{R}^N \,|\, v \cdot n_x < 0\}.$$

The linear Boltzmann equation is supplemented with the absorption boundary condition

$$f(t,x,v) = 0, \quad (x,v) \in \Gamma_-, \quad t > 0. \tag{2.6}$$

(In other words, it is assumed that there are no particles entering the domain Ω.) This choice is made for the sake of simplicity; other boundary conditions will be discussed later.

We next introduce the scaling assumption pertaining to the diffusion approximation of the linear Boltzmann equation (2.1). Set L to be a length scale that measures the size of Ω while V is the average particle speed; consider the time scale $T := L/V$. The diffusion limit of (2.1) is based on the assumption that the dimensionless quantity $Ta(x,v)$ is large. We introduce a scaling parameter $0 < \epsilon \ll 1$ and set

$$\hat{k}_\epsilon(x,v,w) := \epsilon k(x,v,w)$$

so that $\hat{k}_\epsilon(x,v,w)$ is of order unity. Accordingly, we define

$$\hat{a}_\epsilon(x,v) := \epsilon a(x,v), \quad \hat{\mathcal{L}}_x = \epsilon \mathcal{L}_x, \quad \text{and} \quad \hat{\mathcal{K}}_x = \epsilon \mathcal{K}_x.$$

(For notational simplicity, we do not mention explicitly the dependence of $\mathcal{L}_x$ and $\mathcal{K}_x$ in ϵ.) Assume further that

$$\int_{\mathbf{R}^N} v \, d\mu(v) = 0 \tag{2.7}$$

and that variations of order unity of the boundary data driving the solution of (2.1) do not occur on time scales shorter than T/ϵ. This is obvious for the boundary condition (2.6); however, this assumption is crucial and needs to be satisfied for some more general boundary conditions. In that case, the solution f of (2.1) is sought in the form

$$f(t,x,v) = \hat{f}_\epsilon(\epsilon t, x, v),$$

with the notation $\hat{t} = \epsilon t$ for the rescaled time variable. Thus (2.1) takes the form

$$\epsilon \partial_{\hat{t}} \hat{f}_\epsilon(\hat{t},x,v) + v \cdot \nabla_x \hat{f}_\epsilon(\hat{t},x,v) + \frac{1}{\epsilon} \hat{\mathcal{L}}_x \hat{f}_\epsilon(\hat{t},x,v) = 0.$$

Henceforth we drop hats on rescaled variables and consider the initial-boundary value problem for the scaled linear Boltzmann equation

$$\begin{cases} (\epsilon\partial_t + v\cdot\nabla_x)f_\epsilon(t,x,v) + \dfrac{1}{\epsilon}\mathcal{L}_x f_\epsilon(t,x,v) = 0, & x\in\Omega,\, v\in\mathbf{R}^N,\, t>0, \\ f_\epsilon(t,x,v) = 0, & (x,v)\in\Gamma_-,\, t>0, \\ f_\epsilon(0,x,v) = f^{in}(x,v), & x\in\Omega,\, v\in\mathbf{R}^N, \end{cases} \quad (2.8)$$

with

$$\mathcal{L}_x f(t,x,v) = a_\epsilon(x,v)f(t,x,v) - \int k_\epsilon(x,v,w)f(t,x,w)d\mu(w), \quad (2.9)$$

in the limit as $\epsilon\to 0$.

2.2. Statement of the diffusion approximation.   From now on, we assume that the probability measure satisfies (2.7) and

$$0 < \int_{\mathbf{R}^N} |v\cdot\xi|^2 d\mu(v) < \infty \qquad \text{for all } \xi\in\mathbf{R}^N\setminus\{0\}. \quad (2.10)$$

Decompose the spatial domain as follows: $\Omega = A\cup B$ with $A\cap B = \varnothing$. Assume that A is open and B is closed in $\mathbf{R}^N$ (i.e. $B\cap\partial\Omega = \varnothing$), with finitely many connected components denoted B_l, for $l = 1,\ldots,m$. We further assume that B_l has C^1 boundary, and that B_l is locally on one side of its boundary ∂B_l. Finally, we denote by n_x the unit normal field at $x\in\partial A$, oriented towards the exterior of A.

The characteristic part of $\partial A\times\mathbf{R}^N$, i.e.

$$(\partial A\times\mathbf{R}^N)_0 := \{(x,v)\in\partial A\times\mathbf{R}^N \text{ s.t. } v\cdot n_x = 0\}, \quad (2.11)$$

is assumed to be negligible in the following sense:

$$d\sigma(x)d\mu(v) - \text{meas}((\partial A\times\mathbf{R}^N)_0) = 0, \quad (2.12)$$

where $d\sigma$ is the surface element on ∂A.

We further assume that the scattering kernel k_ϵ in the linear Boltzmann equation is a $dxd\mu(v)d\mu(w)$-a.e. nonnegative measurable function on $\Omega\times\mathbf{R}^N\times\mathbf{R}^N$ satisfying the following assumptions, in addition to (2.3):

(H1) the absorption rate a_ϵ is uniformly small on $B\times\mathbf{R}^N$ as $\epsilon\to 0$, i.e.

$$\|a_\epsilon\|_{L^\infty(B\times\mathbf{R}^N,dxd\mu)} \to 0 \quad \text{as } \epsilon\to 0; \quad (2.13)$$

(H2) the restriction of k_ϵ to $A\times\mathbf{R}^N\times\mathbf{R}^N$ is assumed to be independent of ϵ and denoted $k_A\equiv k_A(x,v,w)$; it satisfies

$$C_K := \underset{(x,v)\in A\times\mathbf{R}^N}{\text{esssup}} \int_{\mathbf{R}^N} \left(k_A(x,v,w) + \frac{1}{k_A(x,v,w)} + \frac{1}{k_A(x,w,v)}\right) d\mu(w) < \infty. \quad (2.14)$$

We henceforth denote

$$a_A(x,v) := \int_{\mathbf{R}^N} k_A(x,v,w)d\mu(w), \quad \text{for } dxd\mu(v) - \text{a.e. } (x,v)\in A\times\mathbf{R}^N. \quad (2.15)$$

The diffusion approximation requires one additional assumption on the set B where the scattering rate vanishes as $\epsilon \to 0$. A first possibility is to postulate a lower bound on the scattering rate on B which is compatible with assumption (H1) as $\epsilon \to 0$:

(H3) the restriction of k_ϵ to $B \times \mathbf{R}^N \times \mathbf{R}^N$ is assumed to satisfy the bound

$$\operatorname*{esssup}_{(x,v) \in B \times \mathbf{R}^N} \int_{\mathbf{R}^N} \frac{d\mu(w)}{k_\epsilon(x,v,w)} = o(1/\epsilon^2) \quad \text{as } \epsilon \to 0. \tag{2.16}$$

However, the diffusion approximation can be proven even for scattering rates of order $O(\epsilon^2)$, including the case $a_\epsilon = 0$ corresponding to vacuum, at the expense of an ergodicity assumption on the free transport operator in each connected component B_l of B. For each $l = 1, \ldots, m$, denote by $\tau_l \equiv \tau_l(x,v)$ the forward exit time from B_l starting from the position $x \in \mathring{B}_l$ with the velocity v; in other words

$$\tau_l(x,v) := \inf\{t > 0 \text{ s.t. } x + tv \in \partial B_l\}. \tag{2.17}$$

We extend $\tau_l(x,v)$ to the noncharacteristic part of $\partial B_l \times \mathbf{R}^N$ as follows: first we set

$$\tau_l(x,v) = 0 \quad \text{if } x \in \partial B_l \text{ and } v \cdot n_x < 0.$$

On the other hand, if $x \in \partial B_l$ and $v \cdot n_x > 0$, we define $\tau_l(x,v)$ so that

$$\tau_l(x + tv, v) \to \tau_l(x,v) \quad \text{as } t \to 0^+.$$

Instead of condition (H3), one can assume that

(H4) the Borel probability measure μ satisfies $\mu(\{0\}) = 0$ and, for each $l = 1, \ldots, m$ and each $g \in L^2(\partial B_l)$,

$$g(x + \tau_l(x,v)v) = g(x) \text{ for } d\sigma(x)d\mu(v) - \text{a.e. } (x,v) \in \partial B_l \times \mathbf{R}^N$$
$$\Rightarrow g(x) = \frac{1}{|\partial B_l|} \int_{\partial B_l} g(y)d\sigma(y) \text{ for a.e. } x \in \partial B_l, \tag{2.18}$$

where $d\sigma$ is the surface element on ∂B_l.

For instance, the assumption (H4) is satisfied if the measure μ is spherically symmetric and if B_l is convex for each $l = 1, \ldots, m$.

The main result in this paper is summarized in the following theorem.

THEOREM 2.1. *Assume that the Borel probability measure μ satisfies (2.7)-(2.10), while the scattering kernel k_ϵ satisfies (2.3), together with the conditions (H1)-(H2) and at least one of the conditions (H3) or (H4). Let $f^{in} \in L^2(\Omega \times \mathbf{R}^N; dxd\mu)$.*

(a) For each $\epsilon > 0$, the Cauchy problem (2.8) has a unique weak solution f_ϵ belonging to $C_b(\mathbf{R}_+; L^2(\Omega \times \mathbf{R}^N; dxd\mu(v)))$.

(b) For a.e. $x \in A$, there exists a unique $\mathbf{R}^N$-valued vector field $b^*(x, \cdot) \in L^2(\mathbf{R}^N, d\mu)^N$ such that

$$\mathcal{L}_x^* b^*(x,v) = v \quad \text{and} \quad \int_{\mathbf{R}^N} b^*(x,v)d\mu(v) = 0.$$

(c) The $M_N(\mathbf{R})$-valued matrix field M defined by

$$M_{ij}(x) := \int_{\mathbf{R}^N} b_i^*(x,v)v_j d\mu(v) \quad \text{for a.e. } x \in A \text{ and all } i, j = 1, \ldots, N$$

satisfies

$$|M_{ij}(x)| \leq 2C_K \|v_i\|_{L^2(\mathbf{R}^N, d\mu)} \|v_j\|_{L^2(\mathbf{R}^N, d\mu)} \quad \text{for a.e. } x \in A \text{ and all } i, j = 1, \ldots, N$$

and

$$\sum_{i,j=1}^{N} M_{ij}(x) \xi_i \xi_j \geq \frac{\beta}{2C_K} |\xi|^2 \text{ for all } \xi \in \mathbf{R}^N, \quad \text{for a.e. } x \in A,$$

where $\beta > 0$ is the smallest eigenvalue of the real symmetric matrix S defined by

$$S_{ij} := \int_{\mathbf{R}^N} v_i v_j d\mu(v). \tag{2.19}$$

(d) In the limit as $\epsilon \to 0$,

$$f_\epsilon \to \rho \text{ weakly-* in } L^\infty(\mathbf{R}_+; L^2(\Omega \times \mathbf{R}^N; dx d\mu))$$

where ρ is the unique weak solution of

$$\begin{cases} \partial_t \rho(t, x) = \text{div}_x(M(x) \nabla_x \rho(t, x)), & x \in A, \, t > 0, \\ \rho(t, x) = 0, & x \in \partial\Omega, \, t > 0, \\ \rho(t, x) = \rho_l(t), & x \in B_l, \, l = 1, \ldots, m, \, t > 0, \\ \dot{\rho}_l(t) = \dfrac{1}{|B_l|} \displaystyle\int_{\partial B_l} \dfrac{\partial \rho}{\partial n_M}(t, x) d\sigma(x), & l = 1, \ldots, m, \, t > 0, \\ \rho(0, x) = \rho^{in}(x), & x \in \Omega, \end{cases} \tag{2.20}$$

and where ρ^{in} is given by the following formula:

$$\rho^{in}(x) = \begin{cases} \displaystyle\int_{\mathbf{R}^N} f^{in}(x, v) d\mu(v) & \text{for a.e. } x \in A \\ \dfrac{1}{|B_l|} \displaystyle\iint_{B_l \times \mathbf{R}^N} f^{in}(y, v) dy d\mu(v) & \text{for a.e. } x \in B_l \end{cases}$$

In (2.20), we have used the standard notation

$$\frac{\partial \rho}{\partial n_M}(t, x) := \sum_{i,j=1}^{N} M_{ij}(x) n_{x,i} \partial_{x_j} \rho(t, x).$$

The diffusion approximation stated in Theorem 2.1 (d) can be strengthened as follows, provided that the initial condition f^{in} is independent of v and constant in each one of the connected components B_l of B.

THEOREM 2.2. *Assume that the Borel probability measure μ satisfies (2.7)-(2.10), while the scattering kernel k_ϵ satisfies (2.3), together with the conditions (H1)-(H2), and at least one of the conditions (H3) or (H4). Let $\rho^{in} \in L^2(\Omega)$ satisfy the condition*

$$\rho^{in}(x) = \frac{1}{|B_l|} \int_{B_l} \rho^{in}(y) dy \quad \text{for a.e. } x \in B_l, \quad l = 1, \ldots, m.$$

Assume further that

$$\mathcal{L}_x b^*(x,v) = v \quad \text{for a.e. } (x,v) \in A \times \mathbf{R}^N,$$

where b^* is the vector field defined in Theorem 2.1 (b). Then

(a) for a.e. $x \in A$, the matrix $M(x)$ defined in Theorem 2.1 is symmetric;

(b) the solution f_ϵ of the Cauchy problem for the linear Boltzmann equation (2.8) with initial data $f^{in}(x,v) = \rho^{in}(x)$ satisfies

$$f_\epsilon(t,\cdot,\cdot) \to \rho(t,\cdot) \text{ strongly in } L^2(\Omega \times \mathbf{R}^N; dxd\mu) \text{ for all } t \geq 0$$

and

$$\frac{1}{\epsilon}\left(f_\epsilon - \int_{\mathbf{R}^N} f_\epsilon d\mu(v)\right) \to -b^* \cdot \nabla_x \rho \text{ strongly in } L^2(\mathbf{R}_+ \times A \times \mathbf{R}^N; dxd\mu)$$

as $\epsilon \to 0$, where ρ is the unique weak solution of the diffusion problem (2.20).

Notice that the strong diffusion limit theorem (Theorem 2.2) involves the condition $\mathcal{L}_x b^*(x,v) = v$, implying that the diffusion matrix $M(x)$ is symmetric. On the contrary, the weak diffusion limit theorem (Theorem 2.1) applies to situations where $M(x)$ may fail to be symmetric. At the time of this writing, we do not know whether strong convergence in the diffusion limit can be obtained with this level of generality in cases where the diffusion $M(x)$ is not symmetric for a.e. $x \in A$.

2.3. Remarks on Theorem 2.1. The class of collision integrals $\mathcal{L}_x$ considered here is obviously more general than in [5]. In [5], it is assumed that the measure μ is the uniform probability measure on the set V of admissible velocities, that can be a ball, or a sphere, or a spherical annulus centered at the origin in $\mathbf{R}^N$. The scattering kernel $k_\epsilon(x,v,w)$ is independent of x and ϵ, and is of the form

$$k_\epsilon(x,v,w) = \sigma \kappa(v,w),$$

where $\sigma > 0$,

$$0 \leq \frac{1}{C} \leq \kappa(v,w) = \kappa(w,v) \leq C \text{ for a.e. } (v,w) \in V \times V,$$

and

$$\int_V \kappa(v,w)dw = 1 \quad \text{for a.e. } v \in V,$$

for some positive constant C. Furthermore, the main result in [5] assumes that

$$\kappa(Rv,Rv') = \kappa(v,v') \text{ for a.e. } (v,v') \in V \times V, \text{ for all } R \in O_N(\mathbf{R}).$$

Under this assumption, the vector field b^* is of the form $b^*(x,v) = \beta^*(|v|)v$ for some real-valued measurable function β defined a.e. on $\mathbf{R}_+$ (see Lemma 4.2.4 in [1] or [12] for an analogous result in a more complex situation), and the diffusion matrix field is of the form $M(x) = mI$, where $m > 0$ is a positive constant (see the formulas (43)-(44) in [5]).

Assumption (H1) is obviously satisfied if $k_\epsilon(x,v,w) = 0$ for $dxd\mu(v)d\mu(w)$-a.e. $(x,v,w) \in B \times \mathbf{R}^N \times \mathbf{R}^N$, or if $k_\epsilon(x,v,w) = O(\epsilon)$ on B. However the assumption used

in the present paper is obviously much more general. For instance, it is satisfied if one has $k_\epsilon(x,v,w) = O(|\ln\epsilon|^{-\gamma_l})$ on B_l with $\gamma_l > 0$ for each $l = 1, \ldots, m$. The Hilbert expansion method used in [5] does not apply to this situation, and therefore cannot be used on the problem considered here in its fullest generality. The treatment of nondiffusive embedded objects in [2] assumes that $k_\epsilon = O(\epsilon)$ in B (see [2] on p. 1683), a situation much less general than the one considered here, which can be treated with the Hilbert expansion method and leads to a diffusion system analogous to (2.20).

Even in the nondegenerate case[1] where $B = \varnothing$, observe that our assumptions on the transition kernel k_ϵ do not imply that the vector field b^* in Theorem 2.1 (b) depends smoothly on x. This again excludes the possibility of using the Hilbert expansion as in [5] to establish the validity of the diffusion limit. Accordingly, the diffusion matrix field M defined in Theorem 2.1 (c) is in general not even continuous. The classical interpretation of the diffusion equation with diffusion matrix M in terms of the associated stochastic differential equation fails in such a case (see for instance Section 5.1 and Remark 5.1.6 in [27]).

Yet, even though the Hilbert expansion method cannot be used on the scaled linear Boltzmann equation (2.8) with the level of generality implied by assumptions (H1)-(H4), notice that the second convergence statement in Theorem 2.2 provides information analogous to the knowledge of the next to leading order term in Hilbert's expansion. (This is easily seen for instance on formula (1.2) in the special case of the radiative transfer equation with isotropic scattering.) At variance with the usual diffusion approximation theory, this information is available in the diffusion region only (i.e. for $x \in A$).

The special case where the diffusion matrix has a type I discontinuity across some smooth surface is equivalent to a transmission problem for two diffusion equations on each side of the discontinuity surface, with continuity of the solution and of the normal component of the current across the discontinuity surface. See for instance [21] on p. 107 or Lemma 1.1 in [13] for a discussion of this well known issue.

If B_l is convex for $l = 1, \ldots, m$ and μ is of the form $d\mu(v) = r(|v|)dv$ or μ is the uniform probability measure on a sphere included in $\mathbf{R}^N$ centered at the origin, the condition (H4) is obviously satisfied. Indeed, for each $x \neq y \in \partial B_l$, the segment $[x,y]$ is included in B_l. Set $\omega_{xy} := (y-x)/|y-x|$; then $\omega_{xy} \cdot n_x > 0$ and $\tau_l(x, \omega_{xy}) = |x-y|$, so that

$$x + \tau(x, c\omega_{xy})c\omega_{xy} = y \quad \text{for all } c > 0.$$

Thus, if $g \in L^2(\partial B_l)$ satisfies

$$g(x + \tau_l(x,v)v) = g(x) \quad \text{for } d\sigma(x)d\mu(v)\text{-a.e. } (x,v) \in \partial B_l \times \mathbf{R}^N,$$

one has

$$g(y) = g(x + \tau(x, c\omega_{xy})c\omega_{xy}) = g(x)$$

for a.e. $(x,y,c) \in \partial B_l \times \partial B_l \times (0,\infty)$. Since the left- and right hand side of this equality are independent of c, one concludes that $g(x) = g(y)$ for a.e. $(x,y) \in (\partial B_l)^2$. Averaging in $y \in \partial B_l$, one concludes that

$$g(x) = \frac{1}{|\partial B_l|} \int_{\partial B_l} g(y) d\sigma(y), \quad \text{for a.e. } x \in \partial B_l.$$

---

[1]We call this case "nondegenerate" because the scattering kernel is controlled from above and below by quantities of order $O(1/\epsilon)$ on the whole spatial domain Ω, in the sense that A fills Ω and (H2) holds with $A = \Omega$. This is precisely the scaling assumption used in the diffusion approximation: see for instance the classical references [5, 20, 25].

But even when B_l is convex, the condition (H4) may fail to be satisfied for some measures μ. For instance, assume that $N = 2$, and take $B_l = \{x \in \mathbf{R}^2 \text{ s.t. } |x| \le 1\}$. Denote by (e_1, e_2) the canonical basis of $\mathbf{R}^2$, and let

$$\mu = \tfrac{1}{4}(\delta_{e_1} + \delta_{-e_1} + \delta_{e_2} + \delta_{-e_2}).$$

For $x = (x_1, x_2) \in \partial B_l$ and $\omega_i \in \{\pm e_1, \pm e_2\}$, one has

$$\tau_l(x, \omega_i) = 2|x_i| \quad \text{if } x \cdot \omega_i < 0.$$

Therefore

$$(-x_1, x_2) + \tau_l(x, e_1)e_1 = (x_1, x_2), \quad (x_1, x_2) - \tau_l(x, -e_1)e_1 = (-x_1, x_2), \quad x_1 > 0,$$
$$(x_1, -x_2) + \tau_l(x, e_2)e_2 = (x_1, x_2), \quad (x_1, x_2) - \tau_l(x, -e_2)e_2 = (x_1, -x_2), \quad x_2 > 0.$$

Thus $g(x) = |x_1|$ or $g(x) = |x_2|$ are not a.e. constant on ∂B_l and yet satisfy the condition

$$g(x + \tau_l(x, v)v) = g(x) \quad \text{for } dx \, d\mu(v) - \text{a.e. } (x, v) \in \partial B_l \times \mathbf{R}^2.$$

Comparing theorems 2.1 and 2.2 with the result in [14] is a more delicate issue. We recall that the problem considered in [14] involves the juxtaposition of a medium where the collision cross-section is of order 1 and a highly collisional medium, where the collision cross section is of order $1/\epsilon$. The setting is one dimensional, but extensions to higher dimensions are possible and discussed in [14]. The main result in [14] is a proof of the validity of a domain decomposition strategy where the highly collisional medium is treated by the diffusion equation, with a boundary layer term that is the solution of a Milne problem (see [5]) to accurately describe the interface. The interested reader is referred to [14] for a more accurate description of this domain decomposition algorithm.

At first sight, the situation considered in the present paper is of the same type, as the case of a transition kernel k_ϵ such that $k_\epsilon(x, v, w) = O(\epsilon)$ for a.e. $x \in B$ is covered by our assumptions. Yet the result in theorems 2.1 and 2.2 obviously does not involve any sophisticated treatment of the interface between A and B that would require solving a Milne problem. The difference between both results comes from the type of boundary data considered in [14] and here. In the situation considered in theorems 2.1 and 2.2, the distribution function of particles on the boundary of each connected component of B, i.e. of the region where the collision cross-section is of order 1, is independent of the variable v in the vanishing ϵ limit. Indeed, for each $l = 1, \ldots, m$, we have assumed that $B_l \cap \partial\Omega = \varnothing$, so that the domain B_l is surrounded by the diffusive region A where the distribution function is independent of v in the vanishing ϵ limit (see (5.11) below). For such boundary data, one easily verifies that the boundary layer matching the kinetic and the diffusion domains in [14] is trivial to leading order.

Finally, notice that the time scale on which the evolution of the linear Boltzmann equation (2.8) is observed is the same in the diffusive as well as in the nondiffusive parts of Ω. This may be slightly surprising at first sight, since the diffusion approximation involves a near equilibrium regime, and therefore needs to be observed on a time scale much longer than the original time scale for the linear Boltzmann equation. One of the difficulties in matching diffusive and nondiffusive regions in the theory of the linear Boltzmann equation is that diffusion and transport are phenomena evolving on different time scales. However, in the situation considered here, the boundary of each one of the connected components B_l of B does not touch $\partial\Omega$. Therefore, the transport process in each B_l is driven by the surrounding diffusive region. This explains why our asymptotic theory in Theorems 2.1 and 2.2 involves the same time scale in the diffusive and in the nondiffusive regions.

3. The collision integral and the diffusion matrix

3.1. Properties of the integral operator $\mathcal{L}_x$.   Henceforth we denote

$$\langle\phi\rangle := \int_{\mathbf{R}^N}\phi(v)d\mu(v), \quad \text{for all } \phi\in L^1(\mathbf{R}^N,d\mu).$$

LEMMA 3.1. *Assume that μ is a Borel probability measure on $\mathbf{R}^N$, and that k_ϵ is a nonnegative $dxd(\mu\otimes\mu)$-measurable function defined $dxd(\mu\otimes\mu)$-a.e. on $\Omega\times\mathbf{R}^N\times\mathbf{R}^N$, satisfying (2.3). Assume further that the function $a_\epsilon\equiv a_\epsilon(x,v)$ defined in (2.4) satisfies the condition*

$$a_\epsilon\in L^\infty(\Omega\times\mathbf{R}^N,dxd\mu). \tag{3.1}$$

(a) The integral operators $\mathcal{L}_x$ and $\mathcal{K}_x$ are bounded on $L^2(\mathbf{R}^N,\mu)$ for a.e. $x\in\Omega$, with

$$\|\mathcal{K}_x\|_{\mathcal{L}(L^2(\mathbf{R}^N,\mu))}\leq \|a_\epsilon(x,\cdot)\|_{L^\infty(\mathbf{R}^N,\mu)}.$$

(b) The adjoints of $\mathcal{K}_x$ and $\mathcal{L}_x$ are given by the formulas

$$\begin{cases} \mathcal{K}_x^*\phi(v) = \displaystyle\int_{\mathbf{R}^N}k_\epsilon(x,w,v)\phi(w)d\mu(w) \text{ and} \\[2mm] \mathcal{L}_x^*\phi(v) = \displaystyle\int_{\mathbf{R}^N}k_\epsilon(x,w,v)(\phi(v)-\phi(w))d\mu(w) \end{cases}$$

for a.e. $x\in\Omega$.

(c) For each $\phi\in L^2(\mathbf{R}^N;d\mu)$, and for a.e. $x\in\Omega$,

$$\langle\phi\mathcal{L}_x\phi\rangle = \tfrac{1}{2}\iint_{\mathbf{R}^N\times\mathbf{R}^N}k_\epsilon(x,v,w)(\phi(v)-\phi(w))^2d\mu(v)d\mu(w).$$

(d) For a.e. $x\in\Omega$,

$$\{ \text{ functions a.e. constant on } \mathbf{R}^N\} = \mathbf{R}\subset\mathrm{Ker}(\mathcal{L}_x)\cap\mathrm{Ker}(\mathcal{L}_x^*);$$

if in addition $k_\epsilon(x,v,w)>0$ for $d\mu(v)d\mu(w)$-a.e. $(v,w)\in\mathbf{R}^N\times\mathbf{R}^N$, then

$$\mathrm{Ker}(\mathcal{L}_x) = \mathrm{Ker}(\mathcal{L}_x^*) = \{ \text{ functions a.e. constant on } \mathbf{R}^N\} = \mathbf{R}.$$

Proof.   Statement (a) follows from Schur's lemma (Lemma 18.1.12 in [17] or Lemma 1 in §2 of chapter XXI in [11]). The formula for $\mathcal{K}_x^*$ in statement (b) and the inclusion in statement (d) are obvious.

As for the formula for $\mathcal{L}_x^*$ in statement (b), observe that

$$\mathcal{L}_x\phi(v)+\mathcal{K}_x\phi(v) = \int_{\mathbf{R}^N}k_\epsilon(x,v,w)\phi(v)d\mu(w)$$

$$= \int_{\mathbf{R}^N}k_\epsilon(x,w,v)\phi(v)d\mu(w) = a_\epsilon(x,v)\phi(v)$$

for $dxd\mu$-a.e. $(x,v)\in\Omega\times\mathbf{R}^N$ by the semi-detailed balance assumption (2.3).

For each $\phi \in L^2(\mathbf{R}^N; d\mu)$ and a.e. $x \in \Omega$

$$
\begin{aligned}
\langle \phi \mathcal{L}_x \phi \rangle &= \iint_{\mathbf{R}^N \times \mathbf{R}^N} k_\epsilon(x,v,w)(\phi(v)^2 - \phi(v)\phi(w)) d\mu(v) d\mu(w) \\
&= \int_{\mathbf{R}^N} a_\epsilon(x,v)\phi(v)^2 d\mu(v) - \iint_{\mathbf{R}^N \times \mathbf{R}^N} k_\epsilon(x,v,w)\phi(v)\phi(w) d\mu(v) d\mu(w) \\
&= \tfrac{1}{2} \int_{\mathbf{R}^N} a_\epsilon(x,v)\phi(v)^2 d\mu(v) + \tfrac{1}{2} \int_{\mathbf{R}^N} a_\epsilon(x,w)\phi(w)^2 d\mu(v) \\
&\quad - \iint_{\mathbf{R}^N \times \mathbf{R}^N} k_\epsilon(x,v,w)\phi(v)\phi(w) d\mu(v) d\mu(w) \\
&= \iint_{\mathbf{R}^N \times \mathbf{R}^N} k_\epsilon(x,v,w)\tfrac{1}{2}(\phi(v)^2 + \phi(w)^2) d\mu(v) d\mu(w) \\
&\quad - \iint_{\mathbf{R}^N \times \mathbf{R}^N} k_\epsilon(x,v,w)\phi(v)\phi(w) d\mu(v) d\mu(w) \\
&= \tfrac{1}{2} \iint_{\mathbf{R}^N \times \mathbf{R}^N} k_\epsilon(x,v,w)(\phi(v) - \phi(w))^2 d\mu(v) d\mu(w)
\end{aligned}
$$

by Fubini's theorem and the semi-detailed balance assumption (2.3). This proves statement (c).

By statement (c), if $\phi \in L^2(\mathbf{R}^N; d\mu)$ satisfies $\mathcal{L}_x \phi = 0$, then

$$
0 = \langle \phi \mathcal{L}_x \phi \rangle = \tfrac{1}{2} \iint_{\mathbf{R}^N \times \mathbf{R}^N} k_\epsilon(x,v,w)(\phi(v) - \phi(w))^2 d\mu(v) d\mu(w).
$$

Therefore

$$
k_\epsilon(x,v,w)(\phi(v) - \phi(w)) = 0 \quad \text{for } d\mu(v) d\mu(w)\text{-a.e. } (v,w) \in \mathbf{R}^N \times \mathbf{R}^N
$$

so that

$$
\phi(v) - \phi(w) = 0 \quad \text{for } d\mu(v) d\mu(w)\text{-a.e. } (v,w) \in \mathbf{R}^N \times \mathbf{R}^N.
$$

Since $k_\epsilon(x,v,w) > 0$ for $d\mu(v) d\mu(w)$-a.e. $(v,w) \in \mathbf{R}^N$, averaging in w shows that

$$
\phi(v) = \langle \phi \rangle \quad \text{for } d\mu(v)\text{-a.e. } v \in \mathbf{R}^N,
$$

so that

$$
\mathrm{Ker}(\mathcal{L}_x) \subset \{ \text{ functions a.e. constant on } \mathbf{R}^N \} = \mathbf{R}.
$$

Since the function $(v,w) \mapsto k_\epsilon(x,w,v)$ satisfies the same properties as the function $(v,w) \mapsto k_\epsilon(x,v,w)$,

$$
\mathrm{Ker}(\mathcal{L}_x^*) \subset \{ \text{ functions a.e. constant on } \mathbf{R}^N \} = \mathbf{R},
$$

and the proof is complete. □

3.2. The Fredholm alternative: proof of Theorem 2.1 (b)-(c).

PROPOSITION 3.2. *Assume that the Borel probability measure μ satisfies (2.7)-(2.10), and that k_ϵ is a nonnegative $dx d(\mu \otimes \mu)$-measurable function, defined $dx d(\mu \otimes \mu)$-a.e. on $\Omega \times \mathbf{R}^N \times \mathbf{R}^N$, satisfying (2.3) and (H2).*

(a) *For a.e.* $x \in A$ *and each* $\phi \in L^2(\mathbf{R}^N, d\mu)$

$$\|\phi - \langle\phi\rangle\|_{L^2(\mathbf{R}^N;d\mu)} \leq 2C_K \|\mathcal{L}_x\phi\|_{L^2(\mathbf{R}^N,d\mu)},$$
$$\|\phi - \langle\phi\rangle\|_{L^2(\mathbf{R}^N;d\mu)} \leq 2C_K \|\mathcal{L}_x^*\phi\|_{L^2(\mathbf{R}^N,d\mu)}.$$

(b) *For a.e.* $x \in A$, *the operators* $\mathcal{L}_x$ *and* $\mathcal{L}_x^*$ *are bounded Fredholm operators on* $L^2(\mathbf{R}^N; d\mu)$, *with*

$$\mathrm{Ran}(\mathcal{L}_x) = \mathbf{R}^\perp \quad and \quad \mathrm{Ran}(\mathcal{L}_x^*) = \mathbf{R}^\perp.$$

(c) *For a.e.* $x \in A$, *there exists unique* $\mathbf{R}^N$-*valued vector fields* $b(x, \cdot)$ *and* $b^*(x, \cdot)$ *in* $L^2(\mathbf{R}^N, d\mu)^N$ *such that*

$$\begin{cases} \mathcal{L}_x b(x,v) = v & and \quad \langle b(x, \cdot) \rangle = 0 \\ \mathcal{L}_x^* b^*(x,v) = v & and \quad \langle b^*(x, \cdot) \rangle = 0. \end{cases}$$

(d) *For a.e.* $x \in A$ *and all* $i, j = 1, \ldots, N$, *one has*

$$\int_{\mathbf{R}^N} b_i^*(x,v) v_j \, d\mu(v) = \int_{\mathbf{R}^N} v_i b_j(x,v) \, d\mu(v).$$

Notice that statement (b) in Theorem 2.1 is exactly the part of statement (c) in Proposition 3.2 concerning the vector field b^*.

Proof.   Set $\mathcal{L}_x \phi = \psi$; by statement (c) in Lemma 3.1

$$\langle \phi\psi \rangle = \langle \phi\mathcal{L}_x\phi \rangle = \tfrac{1}{2} \iint_{\mathbf{R}^N \times \mathbf{R}^N} k_\epsilon(x,v,w)(\phi(v) - \phi(w))^2 d\mu(v) d\mu(w) \geq 0.$$

By the Cauchy-Schwarz inequality, for a.e. $x \in A$,

$$|\phi(v) - \langle\phi\rangle|^2 = \left(\int_{\mathbf{R}^N} (\phi(v) - \phi(w)) d\mu(w) \right)^2$$
$$\leq \int_{\mathbf{R}^N} \frac{d\mu(w)}{k_A(x,v,w)} \int_{\mathbf{R}^N} k_A(x,v,w)(\phi(v) - \phi(w))^2 d\mu(w),$$

so that

$$\|\phi - \langle\phi\rangle\|_{L^2(\mathbf{R}^N;d\mu)}^2 \leq C_K \iint_{\mathbf{R}^N \times \mathbf{R}^N} k_A(x,v,w)(\phi(v) - \phi(w))^2 d\mu(v) d\mu(w) = 2C_K\langle\phi\psi\rangle.$$

Next,

$$\langle\psi\rangle = \langle\mathcal{L}_x\phi\rangle = \langle(\mathcal{L}_x^*1)\phi\rangle = 0,$$

since $\mathcal{L}_x^* 1 = 0$ by Lemma 3.1 (d), so that

$$\langle\phi\psi\rangle = \langle(\phi - \langle\phi\rangle)\psi\rangle \leq \|\psi\|_{L^2(\mathbf{R}^N,d\mu)} \|\phi - \langle\phi\rangle\|_{L^2(\mathbf{R}^N,d\mu)}$$

by the Cauchy-Schwarz inequality. Putting together the last two inequalities, we obtain the bound

$$\|\phi - \langle\phi\rangle\|_{L^2(\mathbf{R}^N;d\mu)} \leq 2C_K \|\psi\|_{L^2(\mathbf{R}^N,d\mu)}.$$

Since the function $k_A(x,w,v)$ satisfies the same assumptions as k_A, the analogous inequality for the adjoint operator $\mathcal{L}_x^*$ immediately follows, which proves statement (a).

Pick $x \in A$ such that $a_A(x,\cdot) \in L^\infty(\mathbf{R}^N; d\mu)$, and the inequalities in statement (a) are verified. By Lemma 3.1 (a)-(b), the operators $\mathcal{L}_x$ and $\mathcal{L}_x^*$ are bounded on $L^2(\mathbf{R}^N; d\mu)$. Besides, $\mathrm{Ran}(\mathcal{L}_x)$ is closed, as shown by the following argument. Let $\psi \in L^2(\mathbf{R}^N, d\mu)$ and ϕ_n be a sequence in $L^2(\mathbf{R}^N, d\mu)$ such that

$$\mathcal{L}_x\phi_n \to \psi \quad \text{in } L^2(\mathbf{R}^N, d\mu) \text{ as } n \to \infty.$$

In particular $\mathcal{L}_x\phi_n$ is a Cauchy sequence; since

$$\|(\phi_n - \langle\phi_n\rangle) - (\phi_m - \langle\phi_m\rangle)\|_{L^2(\mathbf{R}^N, d\mu)} \leq 2C_K\|\mathcal{L}_x\phi_n - \mathcal{L}_x\phi_m\|_{L^2(\mathbf{R}^N, d\mu)},$$

we conclude that $\phi_n - \langle\phi_n\rangle$ is a Cauchy sequence in $L^2(\mathbf{R}^N, d\mu)$. Therefore,

$$\phi_n - \langle\phi_n\rangle \to \tilde{\phi} \quad \text{in } L^2(\mathbf{R}^N, d\mu) \text{ as } n \to \infty,$$

so that

$$\mathcal{L}_x(\phi_n - \langle\phi_n\rangle) = \mathcal{L}_x\phi_n \to \mathcal{L}_x\tilde{\phi} \quad \text{in } L^2(\mathbf{R}^N, d\mu) \text{ as } n \to \infty.$$

Hence $\psi = \mathcal{L}_x\tilde{\phi} \in \mathrm{Ran}(\mathcal{L}_x)$. By the same token, $\mathrm{Ran}(\mathcal{L}_x^*)$ is closed. Applying Corollary II.17 in [9] and Lemma 3.1 (d) shows that

$$\mathrm{Ran}(\mathcal{L}_x)^\perp = \mathrm{Ker}(\mathcal{L}_x^*) = \mathbf{R} \quad \text{and} \quad \mathrm{Ran}(\mathcal{L}_x^*)^\perp = \mathrm{Ker}(\mathcal{L}_x) = \mathbf{R}.$$

By Proposition II.12 in [9], we conclude that $\mathcal{L}_x$ and $\mathcal{L}_x^*$ are Fredholm operators with

$$\mathrm{Ran}(\mathcal{L}_x) = \mathrm{Ker}(\mathcal{L}_x^*)^\perp = \mathbf{R}^\perp \quad \text{and} \quad \mathrm{Ran}(\mathcal{L}_x^*) = \mathrm{Ker}(\mathcal{L}_x)^\perp = \mathbf{R}^\perp.$$

The existence and uniqueness of the vector fields b and b^* in statement (c) follow from the orthogonality condition (2.7) and the Fredholm alternative in statement (b).

Finally, for a.e. $x \in A$ and all $i,j = 1,\ldots,N$, one has

$$\int_{\mathbf{R}^N} b_i^*(x,v)v_j d\mu(v) = \int_{\mathbf{R}^N} b_i^*(x,v)\mathcal{L}_x b_j(x,v)d\mu(v)$$
$$= \int_{\mathbf{R}^N} \mathcal{L}_x^* b_i^*(x,v)b_j(x,v)d\mu(v) = \int_{\mathbf{R}^N} v_i b_j(x,v)d\mu(v),$$

which proves statement (d).   □

It remains to prove statement (c) in Theorem 2.1.

Proof of Theorem 2.1 (c). The inequality in Proposition 3.2 (a) implies that

$$\|b_i^*(x,\cdot)\|_{L^2(\mathbf{R}^N, d\mu)} \leq 2C_K\|\mathcal{L}^* b_i^*(x,\cdot)\|_{L^2(\mathbf{R}^N, d\mu)} = 2C_K\|v_i\|_{L^2(\mathbf{R}^N, d\mu)}$$

for all $i = 1,\ldots,N$ and a.e. $x \in A$. This bound implies the first inequality in statement (c) of Theorem 2.1 as a consequence of the definition of the $M_{ij}(x)$ and of the Cauchy-Schwarz inequality.

As for the second inequality in statement (c), for a.e. $x \in A$ and all $\xi \in \mathbf{R}^N$,

$$(\xi \cdot v)^2 = (\mathcal{L}_x(\xi \cdot b(x,v)))^2$$
$$= \left(\int_{\mathbf{R}^N} k_A(x,v,w)(\xi \cdot b(x,v) - \xi \cdot b(x,w))d\mu(w)\right)^2$$
$$\leq a_A(x,v)\int_{\mathbf{R}^N} k_A(x,v,w)(\xi \cdot b(x,v) - \xi \cdot b(x,w))^2 d\mu(w)$$

by the Cauchy-Schwarz inequality. On the other hand, by Lemma 3.1 (c) and Proposition 3.2 (d),

$$\sum_{i,j=1}^{N} M_{ij}(x)\xi_i\xi_j = \int_{\mathbf{R}^N} (\xi \cdot b(x,v))\mathcal{L}_x(\xi \cdot b(x,v))d\mu(v)$$

$$= \tfrac{1}{2} \iint_{\mathbf{R}^N \times \mathbf{R}^N} k_A(x,v,w)(\xi \cdot b(x,v) - \xi \cdot b(x,w))^2 d\mu(v)d\mu(w).$$

Hence

$$\sum_{i,j=1}^{N} S_{ij}\xi_i\xi_j = \int_{\mathbf{R}^N} (\xi \cdot v)^2 d\mu(v)$$

$$\leq C_K \iint_{\mathbf{R}^N \times \mathbf{R}^N} k_A(x,v,w)(\xi \cdot b(x,v) - \xi \cdot b(x,w))^2 d\mu(v)d\mu(w)$$

$$= 2C_K \sum_{i,j=1}^{N} M_{ij}(x)\xi_i\xi_j.$$

The conclusion follows from the definition of $\beta > 0$ which implies the inequality

$$\sum_{i,j=1}^{N} S_{ij}\xi_i\xi_j \geq \beta|\xi|^2.$$

$\square$

4. Existence and uniqueness theory for the linear Boltzmann equation and for the diffusion problem

4.1. Proof of Theorem 2.1 (a). Henceforth we denote

$$\langle\!\langle \Phi \rangle\!\rangle := \iint_{\mathbf{R}^N \times \mathbf{R}^N} \Phi(v,w)d\mu(v)d\mu(w), \quad \text{for all } \Phi \in L^1(\mathbf{R}^N \times \mathbf{R}^N, d(\mu \otimes \mu)).$$

PROPOSITION 4.1. *Assume that k_ϵ is a nonnegative measurable function defined $dxd(\mu \otimes \mu)$-a.e. on $\Omega \times \mathbf{R}^N \times \mathbf{R}^N$ satisfying (2.3) and (3.1) with a_ϵ defined by (2.4). For each $\epsilon > 0$ and each $f^{in} \in L^2(\Omega \times \mathbf{R}^N; dxd\mu)$, there exists a unique weak solution of the initial-boundary value problem (2.8) in the space $C_b(\mathbf{R}_+; L^2(\Omega \times \mathbf{R}^N; dxd\mu))$. This solution satisfies*
(a) the continuity equation in the sense of distributions on $\mathbf{R}_+^* \times \Omega$:

$$\partial_t \langle f_\epsilon \rangle + \operatorname{div}_x \frac{1}{\epsilon} \langle v f_\epsilon \rangle = 0;$$

(b) the "entropy inequality"

$$\int_\Omega \langle f_\epsilon(t,x,\cdot)^2 \rangle dx + \int_0^t \int_\Omega \langle\!\langle k_\epsilon(x,\cdot,\cdot)q_\epsilon(s,x,\cdot,\cdot)^2 \rangle\!\rangle dxds \leq \int_\Omega \langle f^{in}(x,\cdot)^2 \rangle dx$$

for each $\epsilon > 0$ and each $t \geq 0$, where

$$q_\epsilon(t,x,v,w) := \frac{1}{\epsilon}(f_\epsilon(t,x,v) - f_\epsilon(t,x,w)).$$

Proof. The operator $\phi(x,v) \mapsto \mathcal{L}_x\phi(x,v)$ is a bounded perturbation of the advection operator $-v\cdot\nabla_x$ with absorbing boundary condition (2.6) that is the generator of a strongly continuous contraction semigroup on $L^2(\Omega\times\mathbf{R}^N;dxd\mu)$. This implies the existence and uniqueness of the weak solution f_ϵ of the initial-boundary value problem (2.8) in the functional space $C_b(\mathbf{R}_+;L^2(\Omega\times\mathbf{R}^N;dxd\mu))$. Statement (a) follows from the inclusion $\mathbf{R}\subset\mathrm{Ker}(\mathcal{L}_x^*)$ in Lemma 3.1 (d), while statement (b) follows from Lemma 3.1 (c) and the usual energy estimate for the transport equation with source.

<div align="right">□</div>

Statement (a) in Theorem 2.1 follows from Proposition 4.1.

4.2. Existence and uniqueness theory for the diffusion problem (2.20).
Define

$$\mathcal{H} := \left\{ u\in L^2(\Omega) \text{ s.t. } u(x) = \frac{1}{|B_l|}\int_{B_l} u(y)dy \text{ for a.e. } x\in B_l,\, l=1,\dots,n \right\},$$

and

$$\mathcal{V} := \mathcal{H}\cap H_0^1(\Omega) = \{u\in H_0^1(\Omega) \text{ s.t. } \nabla u(x) = 0 \text{ for a.e. } x\in B_l,\, l=1,\dots,n\}.$$

These spaces are equipped with the norms

$$\|u\|_{\mathcal{H}} := \|u\|_{L^2(\Omega)} \quad \text{for each } u\in L^2(\Omega),$$
$$\|u\|_{\mathcal{V}} := (\|u\|_{\mathcal{H}}^2 + \|\nabla u\|_{L^2(A)}^2)^{1/2}.$$

For each $\rho^{in}\in\mathcal{H}$, consider the following variational problem

$$\begin{cases} \rho\in C_b(\mathbf{R}_+;\mathcal{H})\cap L^2(\mathbf{R}_+;\mathcal{V}), \quad \partial_t\rho\in L^2(\mathbf{R}_+;\mathcal{V}'), \quad\quad \text{and } \rho\big|_{t=0} = \rho^{in}, \\ \dfrac{d}{dt}\int_\Omega \rho(t,x)w(x)dx + \int_A \nabla w(x)\cdot M(x)\nabla_x\rho(t,x)dx = 0, \quad \text{for a.e. } t\geq 0, \\ \hspace{9cm} \text{for all } w\in\mathcal{V}, \end{cases} \quad (4.1)$$

where $x\mapsto M(x)$ is an $M_N(\mathbf{R})$-valued measurable matrix field such that $M_{ij}\in L^\infty(A)$ for all $i,j=1,\dots,N$.

LEMMA 4.2. *Assume that $x\mapsto M(x)$ is an $M_N(\mathbf{R})$-valued measurable matrix field on A such that $M_{ij}\in L^\infty(A)$ for all $i,j=1,\dots,N$. Let $\rho\in C([0,T];\mathcal{H})\cap L^2([0,T];\mathcal{V})$ with $\partial_t\rho\in L^2([0,T];\mathcal{V}')$. Then ρ is a solution of the variational problem (4.1) if and only if*

$$\begin{cases} \partial_t\rho - \mathrm{div}_x(M\nabla_x\rho) = 0 & \text{in } \mathcal{D}'((0,T)\times A), \\ \rho(t,\cdot)\big|_{\partial\Omega} = 0 & \text{in } L^2([0,T];H^{1/2}(\partial\Omega)), \\ \dot{\rho}_l = \dfrac{1}{|B_l|}\left\langle \dfrac{\partial\rho}{\partial n_M},1 \right\rangle_{H^{-1/2}(\partial B_l),H^{1/2}(\partial B_l)} & \text{in } H^{-1}((0,T)),\, l=1,\dots,m \\ \rho\big|_{t=0} = \rho^{in}. \end{cases}$$

Proof. Specializing (4.1) to the case where $w\in C_c^\infty(A)$ is equivalent to

$$\partial_t\rho - \mathrm{div}_x(M\nabla_x\rho) = 0 \quad \text{in } \mathcal{D}'((0,T)\times A).$$

In particular, the (a.e. defined) vector field

$$(0,T)\times A \ni (t,x) \mapsto (\rho(t,x), -M(x)\nabla_x\rho(t,x))$$

is divergence free in $(0,T) \times A$. Applying statement (b) in Lemma A.3 shows that

$$
\begin{aligned}
0 &= \frac{d}{dt} \int_\Omega \rho(t,x) w(x) dx + \int_A \nabla w(x) \cdot M(x) \nabla_x \rho(t,x) dx \\
&= \frac{d}{dt} \int_A \rho(t,x) w(x) dx + \sum_{l=1}^m |B_l| w_l \dot\rho_l(t) + \int_A \nabla w(x) \cdot M(x) \nabla_x \rho(t,x) dx \\
&= \sum_{l=1}^m w_l \left(|B_l| \dot\rho_l(t) - \left\langle \frac{\partial \rho}{\partial n_M}(t,\cdot) \Big|_{\partial B_l}, 1 \right\rangle_{H^{-1/2}(\partial B_l), H^{1/2}(\partial B_l)} \right)
\end{aligned}
$$

for each $w \in \mathcal{V}$, where

$$
w_l := \frac{1}{|B_l|} \int_{B_l} w(y) dy, \quad l = 1,\dots,m.
$$

Since this is true for all $w \in \mathcal{V}$, and therefore for all $(w_1,\dots,w_m) \in \mathbf{R}^m$, one concludes that

$$
|B_l| \dot\rho_i - \left\langle \frac{\partial \rho}{\partial n_M} \Big|_{\partial B_l}, 1 \right\rangle_{H^{-1/2}(\partial B_l), H^{1/2}(\partial B_l)} = 0
$$

in $H^{-1}((0,T))$ for all $l = 1,\dots,m$, which is precisely the transmission condition on ∂B_l. Finally, the Dirichlet condition on $\partial\Omega$ comes from the condition $\rho \in L^2(\mathbf{R}_+; \mathcal{V})$ since $\mathcal{V} \subset H_0^1(\Omega)$.

Conversely, if $\rho \in C_b([0,T],\mathcal{H}) \cap L^2([0,T],\mathcal{V})$ s.t. $\partial_t \rho \in L^2([0,T],\mathcal{V}')$ satisfies the initial condition and the diffusion equation in (2.20) in the sense of distributions on $(0,T) \times A$, together with the transmission condition on ∂B_l for each $l = 1,\dots,m$, it follows from the identity above that ρ must satisfy (4.1). □

This lemma justifies the following definition.

DEFINITION 4.3.   *For $\rho^{in} \in \mathcal{H}$, a weak solution of the problem (2.20) is a function $\rho \equiv \rho(t,x)$ such that $\rho \in C_b(\mathbf{R}_+; \mathcal{H}) \cap L^2(\mathbf{R}_+; \mathcal{V})$ and $\partial_t \rho \in L^2(\mathbf{R}_+; \mathcal{V}')$, which satisfies the variational formulation and the initial condition in (4.1).*

The existence and uniqueness theory for the limiting diffusion problem is summarized in the following proposition.

PROPOSITION 4.4.   *Assume that $x \mapsto M(x)$ is an $M_N(\mathbf{R})$-valued measurable matrix field on A satisfying*

$$
M_{ij} \in L^\infty(A) \text{ for all } i,j = 1,\dots,N, \text{ and there exists } \alpha > 0 \text{ s.t.}
$$
$$
\xi \cdot M(x)\xi \geq \alpha |\xi|^2 \text{ for a.e. } x \in A \text{ and all } \xi \in \mathbf{R}^N.
$$

For each $\rho^{in} \in \mathcal{H}$, the diffusion problem (2.20) has a unique weak solution. This solution satisfies the "energy" identity for each $t \geq 0$:

$$
\tfrac{1}{2} \int_\Omega \rho(t,x)^2 dx + \int_0^t \int_A \nabla_x \rho(s,x) \cdot M(x) \nabla_x \rho(s,x) dx ds = \tfrac{1}{2} \int_\Omega \rho^{in}(x)^2 dx.
$$

Proof.   The existence and uniqueness of the solution of the variational problem (4.1) is a straightforward consequence of the Lions-Magenes Theorem X.9 in [9], with the bilinear form

$$a(u,w) := \int_A \nabla u(x) \cdot M(x) \nabla w(x) dx, \quad u, w \in \mathcal{V}.$$

Indeed, this bilinear form satisfies the assumptions of the Lions-Magenes theorem, since the first inequality in Theorem 2.1 (c) (already established in Section 3.2) implies that

$$|a(u,w)| \le 2C_K \langle |v|^2 \rangle \|\nabla u\|_{L^2(A)} \|\nabla w\|_{L^2(A)} \le 2C_K \langle |v|^2 \rangle \|u\|_{\mathcal{V}} \|w\|_{\mathcal{V}},$$

while the second inequality there implies that

$$a(u,u) \ge \frac{\beta}{2C_K} \|\nabla u\|_{L^2(A)}^2 = \frac{\beta}{2C_K} \|\nabla u\|_{L^2(\Omega)}^2 = \frac{\beta}{2C_K} (\|u\|_{\mathcal{V}}^2 - \|u\|_{\mathcal{H}}^2)$$

for each $u, w \in \mathcal{V}$.

Consider next the linear functional

$$L(t): \mathcal{V} \ni w \mapsto \langle \partial_t \rho, w \rangle_{\mathcal{V}', \mathcal{V}} + a(\rho(t, \cdot), w)$$

defined for a.e. $t \ge 0$. By Lemma A.2, one has

$$L(t) = 0 \text{ in } \mathcal{V}' \text{ for a.e. } t \in \mathbf{R}_+.$$

In particular, for a.e. $s \ge 0$, one has

$$0 = \langle L(s), \rho(s, \cdot) \rangle_{\mathcal{V}', \mathcal{V}} = \langle \partial_t \rho(s, \cdot), \rho(s, \cdot) \rangle_{\mathcal{V}', \mathcal{V}} + \int_A \nabla_x \rho(s, x) \cdot M(x) \nabla_x \rho(s, x) dx,$$

and one concludes by integrating in $s \in [0, t]$ and applying Lemma A.1 (b).   □

5. Diffusion approximation: proof of Theorem 2.1 (d)

The proof is split in several steps.

Step 1: uniform bounds and weak compactness.

By the entropy inequality (statement (b) in Proposition 4.1), one has the bounds

$$\begin{cases} \|f_\epsilon(t, \cdot, \cdot)\|_{L^2(\Omega \times \mathbf{R}^N; dxd\mu)} \le \|\rho^{in}\|_{L^2(\Omega)} & \text{and} \\ \|\sqrt{k_\epsilon} q_\epsilon\|_{L^2(\mathbf{R}_+ \times \Omega \times \mathbf{R}^N \times \mathbf{R}^N; dtdxd(\mu \otimes \mu))} \le \|\rho^{in}\|_{L^2(\Omega)}. \end{cases} \tag{5.1}$$

By the Banach-Alaoglu theorem, the families f_ϵ and $\sqrt{k_\epsilon} q_\epsilon$ are relatively compact in $L^\infty(\mathbf{R}_+; L^2(\Omega \times \mathbf{R}^N; dxd\mu))$ weak-* and $L^2(\mathbf{R}_+ \times \Omega \times \mathbf{R}^N \times \mathbf{R}^N; dtdxd(\mu \otimes \mu))$ weak respectively. Extracting subsequences if needed, one has

$$f_\epsilon \to f \text{ in } L^\infty(\mathbf{R}_+; L^2(\Omega \times \mathbf{R}^N; dxd\mu)) \text{ weak-*} \tag{5.2}$$

while

$$\sqrt{k_\epsilon} q_\epsilon \to r \text{ in } L^2(\mathbf{R}_+ \times \Omega \times \mathbf{R}^N \times \mathbf{R}^N; dtdxd(\mu \otimes \mu)) \text{ weak.} \tag{5.3}$$

In particular

$$q_\epsilon \to q \text{ in } L^2(\mathbf{R}_+ \times A \times \mathbf{R}^N \times \mathbf{R}^N; k_A(x, v, w) dtdxd(\mu \otimes \mu)) \text{ weak,} \tag{5.4}$$

where

$$q(t,x,v,w) := r(t,x,v,w)/\sqrt{k_A(x,v,w)}, \qquad (5.5)$$

for $dtdxd\mu(v)d\mu(w)$-a.e. $(t,x,v,w) \in \mathbf{R}_+ \times A \times \mathbf{R}^N \times \mathbf{R}^N$.

Step 2: asymptotic form of the linear Boltzmann equation.

One has

$$\frac{1}{\epsilon}\mathcal{L}_x f_\epsilon(t,x,v) = \int_{\mathbf{R}^N} \mathbf{1}_A(x)k_A(x,v,w)q_\epsilon(t,x,v,w)d\mu(w)$$
$$+ \int_{\mathbf{R}^N} \mathbf{1}_B(x)k_\epsilon(x,v,w)q_\epsilon(t,x,v,w)d\mu(w).$$

Since $(x,v,w) \mapsto \mathbf{1}_A(x)$ belongs to $L^2(A \times \mathbf{R}^N \times \mathbf{R}^N; k_A(x,v,w)dxd(\mu \otimes \mu))$ by (2.14),

$$\int_{\mathbf{R}^N} \mathbf{1}_A(x)k_A(x,v,w)q_\epsilon(t,x,v,w)d\mu(w) \to \int_{\mathbf{R}^N} \mathbf{1}_A(x)k_A(x,v,w)q(t,x,v,w)d\mu(w)$$

in the weak topology of $L^2(\mathbf{R}_+ \times A \times \mathbf{R}^N; dtdxd\mu)$ as $\epsilon \to 0$. On the other hand, the Cauchy-Schwarz inequality and (2.14) imply that

$$\left\| \int_{\mathbf{R}^N} k_\epsilon(\cdot,\cdot,w)q_\epsilon(\cdot,\cdot,\cdot,w)d\mu(w) \right\|_{L^2(\mathbf{R}_+ \times B \times \mathbf{R}^N; dtdxd\mu)}^2$$
$$\leq \|a_\epsilon\|_{L^\infty(B \times \mathbf{R}^N)} \iint_{\mathbf{R}_+ \times \Omega} \langle\langle k_\epsilon(x,\cdot,\cdot)q_\epsilon(t,x,\cdot,\cdot)^2 \rangle\rangle dtdx$$
$$\leq \|a_\epsilon\|_{L^\infty(B \times \mathbf{R}^N)} \|\rho^{in}\|_{L^2(\Omega)}^2 \to 0$$

as $\epsilon \to 0$, by (2.13) and the entropy inequality in Proposition 4.1. Thus

$$\frac{1}{\epsilon}\mathcal{L}_x f_\epsilon(t,x,v) \to \int_{\mathbf{R}^N} \mathbf{1}_A(x)k_A(x,v,w)q(t,x,v,w)d\mu(w) \qquad (5.6)$$

in the weak topology of $L^2(\mathbf{R}_+ \times \Omega \times \mathbf{R}^N; dxd\mu)$ as $\epsilon \to 0$. Passing to the limit in the scaled Boltzmann equation (2.8), we see that

$$v \cdot \nabla_x f \in L^2(\mathbf{R}_+ \times \Omega \times \mathbf{R}^N, dtdxd\mu) \quad \text{and}$$
$$\int_{\mathbf{R}^N} k_A(\cdot,\cdot,w)q(\cdot,\cdot,\cdot,w)d\mu(w) \in L^2(\mathbf{R}_+ \times A \times \mathbf{R}^N, dtdxd\mu), \qquad (5.7)$$

while

$$v \cdot \nabla_x f(t,x,v) + \mathbf{1}_A(x)\int_{\mathbf{R}^N} k_A(x,v,w)q(t,x,v,w)d\mu(w) = 0, \qquad (5.8)$$

for $dtdxd\mu$-a.e. $(t,x,v) \in \mathbf{R}_+ \times \Omega \times \mathbf{R}^N$.

Step 3: asymptotic form of f_ϵ.

Multiplying both sides of the scaled linear Boltzmann equation (2.8) by ϵ and passing to the limit in the sense of distributions as $\epsilon \to 0$, one finds that

$$\mathcal{L}_x f(t,x,v) = 0 \quad \text{for a.e. } (t,x,v) \in \mathbf{R}_+^* \times A \times \mathbf{R}^N.$$

By Lemma 3.1 (d), this implies that $f(t,x,v)$ is independent of v for a.e. $x \in A$, i.e. is of the form

$$f(t,x,v) = \rho(t,x) \quad \text{for a.e. } (t,x,v) \in \mathbf{R}_+^* \times A \times \mathbf{R}^N. \tag{5.9}$$

By (5.2) and (5.7),

$$\rho \in L^\infty(\mathbf{R}_+; L^2(A)) \quad \text{and} \quad \nabla_x \rho \in L^2(\mathbf{R}_+ \times A), \tag{5.10}$$

since

$$(v \cdot \nabla_x f)v = (v \otimes v) \cdot \nabla_x \rho \in L^2(\mathbf{R}_+ \times A; L^1(\mathbf{R}^N, d\mu)),$$

so that

$$S \cdot \nabla_x \rho = \langle v \otimes v \rangle \cdot \nabla_x \rho \in L^2(\mathbf{R}_+ \times A).$$

Since $\det(S) \neq 0$ by assumption (2.10), one concludes that $\nabla_x \rho \in L^2(\mathbf{R}_+ \times A)$.

In particular,

$$\rho\big|_{\partial B_i} \in L^2([0,T]; H^{1/2}(\partial B_i))$$

for each $T > 0$ and each $i = 1, \ldots, n$.

In particular, the first condition in (5.7) implies that $s \mapsto f(t, x+sv, v)$ is continuous in s for $dt dx d\mu$-a.e. $(t,x,v) \in \mathbf{R}_+ \times \Omega \times \mathbf{R}^N$. Therefore, we deduce from (5.8) and (5.9) that, for each $l = 1, \ldots, m$,

$$\begin{cases} v \cdot \nabla_x f(t,x,v) = 0, & x \in B_l, \, v \in \mathbf{R}^N, \, t > 0, \\ f(t,x,v) = \rho(t,x), & x \in \partial B_l, \, v \in \mathbf{R}^N, \, t > 0. \end{cases} \tag{5.11}$$

At this point, things are different according to whether (H3) or (H4) holds.

Under assumption (H3), applying the Cauchy-Schwarz inequality and the entropy inequality in Proposition 4.1, one finds that,

$$\int_0^\infty \int_B \|f_\epsilon(t,x,\cdot) - \langle f_\epsilon \rangle(t,x)\|_{L^2(\mathbf{R}^N, d\mu)}^2 \, dx dt$$

$$= \int_0^\infty \int_B \int_{\mathbf{R}^N} \left(\int_{\mathbf{R}^N} (f_\epsilon(t,x,v) - f_\epsilon(t,x,w)) d\mu(w) \right)^2 d\mu(v) dx dt$$

$$\leq \operatorname*{esssup}_{(x,v) \in B \times \mathbf{R}^N} \int_{\mathbf{R}^N} \frac{d\mu(w)}{k_\epsilon(x,v,w)}$$

$$\times \int_0^\infty \int_B \iint_{\mathbf{R}^N \times \mathbf{R}^N} k_\epsilon(x,v,w)(f_\epsilon(t,x,v) - f_\epsilon(t,x,w))^2 d\mu(v) d\mu(w) dx dt$$

$$\leq \epsilon^2 \operatorname*{esssup}_{(x,v) \in B \times \mathbf{R}^N} \int_{\mathbf{R}^N} \frac{d\mu(w)}{k_\epsilon(x,v,w)} \|f^{in}\|_{L^2(\Omega \times \mathbf{R}^N)} \to 0$$

as $\epsilon \to 0$. The condition (H3) is used precisely at this point, in order to bound the right hand side of the last inequality above. Therefore

$$f(t,x,v) = \langle f \rangle(t,x) =: \rho(t,x) \quad \text{for a.e. } (t,x,v) \in \mathbf{R}_+ \times B_l \times \mathbf{R}^N, \quad l = 1, \ldots, m.$$

This implies that $v \cdot \nabla_x \rho(t,x) = 0$ for a.e $(t,x,v) \in \mathbf{R}_+ \times B_l \times \mathbf{R}^N$, and we conclude that $S \nabla_x \rho(t,x) = 0$, by the first equality in (5.11). This implies in turn that $\nabla_x \rho(t,x) = 0$

for a.e. $(t,x) \in \mathbf{R}_+ \times B_l$ since the matrix $S = \langle v \otimes v \rangle$ is invertible by (2.10). Since B_l is connected, we conclude that, for $l = 1, \ldots, m$,

$$f(t,x,v) = \rho_l(t) := \frac{1}{|B_l|} \iint_{B_l \times \mathbf{R}^N} f(t,x,v) dx d\mu(v), \tag{5.12}$$

for $dt dx d\mu$-a.e. $(t,x,v) \in \mathbf{R}_+ \times B_l \times \mathbf{R}^N$.

Under assumption (H4), observe that the first equality in (5.11) implies that

$$\frac{d}{ds} f(t, x+sv, v) = 0 \quad \text{for all } s \text{ s.t. } x+sv \in B_l$$

for $dt dx d\mu(v)$-a.e. $(t,x,v) \in \mathbf{R}_+ \times B_l \times \mathbf{R}^N$. Thus, in view of the condition on μ in (H4), one concludes that

$$\rho(t, x+\tau_l(x,v)v) = \rho(t,x) \text{ for } d\sigma(x) d\mu(v)\text{-a.e. } (x,v) \in \partial B_l \times \mathbf{R}^N$$

by solving the boundary value problem above by the method of characteristics. By assumption (2.18),

$$\rho(t,x) = \frac{1}{|\partial B_l|} \int_{\partial B_l} \rho(t,y) d\sigma(y) =: \rho_l(t) \quad \text{for a.e. } x \in \partial B_l,$$

for a.e. $t \geq 0$. In other words, $\rho(t, \cdot)$ is a.e. equal to a constant on ∂B_l. Solving again for f along characteristics, we conclude that (5.12) holds under assumption (H4) even if (H3) is not verified.

Summarizing, we have proved that

$$\begin{aligned} f(t,x,v) &= \rho(t,x) \text{ for } dt dx d\mu\text{-a.e. } (t,x,v) \in \Omega \\ &\text{with } \rho \in L^\infty(\mathbf{R}_+; \mathcal{H}) \text{ and } \nabla_x \rho \in L^2(\mathbf{R}_+ \times \Omega). \end{aligned} \tag{5.13}$$

Step 4: Fourier's law and continuity equation.

Observe that the flux satisfies

$$\begin{aligned} \frac{1}{\epsilon} \langle v f_\epsilon(t,x,\cdot) \rangle &= \frac{1}{\epsilon} \langle (\mathcal{L}_x^* b^*(x,\cdot)) f_\epsilon(t,x,\cdot) \rangle = \left\langle b^*(x,\cdot) \frac{1}{\epsilon} \mathcal{L}_x f_\epsilon(t,x,\cdot) \right\rangle \\ &= \iint_{\mathbf{R}^N \times \mathbf{R}^N} b^*(x,v) k_A(x,v,w) q_\epsilon(t,x,v,w) d\mu(v) d\mu(w) \end{aligned} \tag{5.14}$$

for a.e. $(t,x) \in \mathbf{R}_+ \times A$ and for all $\epsilon > 0$. Since $b^* \in L^\infty(A; L^2(\mathbf{R}^N; d\mu))$ by Proposition 3.2 (c), one concludes that the function $(x,v,w) \mapsto \sqrt{k_A(x,v,w)} b^*(x,v)$ belongs to the space $L^\infty(A; L^2(\mathbf{R}^N \times \mathbf{R}^N; d\mu(v)\mu(w)))$. Thus

$$\begin{aligned} \frac{1}{\epsilon} \langle v f_\epsilon(t,x,\cdot) \rangle &= \iint_{\mathbf{R}^N \times \mathbf{R}^N} b^*(x,v) k_A(x,v,w) q_\epsilon(t,x,v,w) d\mu(v) d\mu(w) \\ &\to \iint_{\mathbf{R}^N \times \mathbf{R}^N} b^*(x,v) k_A(x,v,w) q(t,x,v,w) d\mu(v) d\mu(w) \\ &= -\langle b^*(x,\cdot) v \cdot \nabla_x \rho(t,x) \rangle = -M(x) \nabla_x \rho(t,x) \end{aligned} \tag{5.15}$$

in the weak topology of $L^2(\mathbf{R}_+ \times A)$ as $\epsilon \to 0$, on account of (5.8).

Therefore, for each $w \in \mathcal{V}$, one has

$$\frac{d}{dt} \int_\Omega \langle f_\epsilon(t,x,\cdot) \rangle w(x) dx - \int_A \frac{1}{\epsilon} \langle v f_\epsilon(t,x,\cdot) \rangle \cdot \nabla w(x) dx = 0,$$

(since $\nabla w = 0$ on $\mathring{B}$) and passing to the limit in each side of this identity as $\epsilon \to 0$ shows that

$$\frac{d}{dt}\int_\Omega \rho(t,x)w(x)dx + \int_A \nabla w(x)\cdot M(x)\nabla_x\rho(t,x)dx = 0 \qquad (5.16)$$

in the sense of distributions on $\mathbf{R}_+^*$.

Step 5: limiting initial condition

By (5.14) and the Cauchy-Schwarz inequality

$$\left\|\frac{1}{\epsilon}\langle vf_\epsilon\rangle\right\|^2_{L^2([0,T]\times A)} \leq \int_{\mathbf{R}_+}\int_A \langle\!\langle k_A(x,\cdot,\cdot)q_\epsilon(t,x,\cdot,\cdot)^2\rangle\!\rangle dxdt$$

$$\times \operatorname*{esssup}_{x\in A}\iint_{\mathbf{R}^N\times\mathbf{R}^N} k_A(x,v,w)|b^*(x,v)|^2 d\mu(v)d\mu(w)$$

$$\leq 8C_K^3\langle|v|^2\rangle\|\rho^{in}\|^2_{L^2(\Omega)},$$

using the entropy inequality in Proposition 4.1 and Proposition 3.2 (a)-(c). Since

$$\frac{d}{dt}\int_\Omega \langle f_\epsilon(t,x,\cdot)\rangle w(x)dx = \int_A \frac{1}{\epsilon}\langle vf_\epsilon(t,x,\cdot)\rangle\cdot\nabla w(x)dx$$

for each $w\in\mathcal{V}$, one has

$$\left\|\frac{d}{dt}\int_\Omega \langle f_\epsilon(\cdot,x,\cdot)\rangle w(x)dx\right\|_{L^2([0,T])} \leq (2C_K)^{3/2}\langle|v|^2\rangle^{1/2}\|\rho^{in}\|_{L^2(\Omega)}\|\nabla w\|_{L^2(\Omega)}. \qquad (5.17)$$

Applying the Ascoli-Arzela theorem shows that, for each $w\in\mathcal{V}$,

$$\int_\Omega (\langle f_\epsilon(t,x,\cdot)\rangle - \rho(t,x))w(x)dx \to 0 \text{ uniformly in } t\in[0,T] \qquad (5.18)$$

for all $T>0$. In particular

$$\iint_\Omega f^{in}(x,v)w(x)dxd\mu(v) = \int_\Omega \langle f_\epsilon(0,x,\cdot)\rangle w(x)dx \to \int_\Omega \rho(0,x)w(x)dx$$

as $\epsilon\to 0$. Since the test function w is constant on B_l for each $l=1,\ldots,m$,

$$\begin{aligned}\int_\Omega \rho(0,x)w(x)dx &= \iint_\Omega f^{in}(x,v)w(x)dxd\mu(v)\\ &= \int_\Omega \rho^{in}(x)w(x)dx \quad\text{for each } w\in\mathcal{V},\end{aligned} \qquad (5.19)$$

with ρ^{in} defined by the formula in Theorem 2.1.

Returning to (5.17), we have proved that $\partial_t\langle f_\epsilon\rangle$ is bounded in $L^2(\mathbf{R}_+,\mathcal{V}')$ for each $T>0$, so that

$$\partial_t\rho \in L^2(\mathbf{R}_+;\mathcal{V}'). \qquad (5.20)$$

Since $\rho\in L^\infty(\mathbf{R}_+;\mathcal{H})\cap L^2([0,T];\mathcal{V})$ for each $T>0$ by (5.13), we conclude from (5.20) that

$$\rho\in C_b(\mathbf{R}_+;\mathcal{H})$$

so that (5.19) implies that ρ satisfies the initial condition in (2.20).

Step 6: Dirichlet condition.

Next we establish the Dirichlet condition on $\partial\Omega$ for the diffusion equation. The scaled linear Boltzmann equation implies that, for each $\chi\in C_c^1(\mathbf{R}_+^*)$,

$$v\cdot\nabla_x\int_0^\infty\chi(t)f_\epsilon(t,x,v)dt=-\int_0^\infty\chi(t)\frac{1}{\epsilon}\mathcal{L}_xf_\epsilon(t,x,v)dt+\epsilon\int_0^\infty\chi'(t)f_\epsilon(t,x,v)dt$$

is bounded in $L^2(\Omega\times\mathbf{R}^N;dxd\mu)$ by (5.6), the uniform boundedness principle (the Banach-Steinhaus theorem), and the entropy inequality in Proposition 4.1, while

$$\int_0^\infty\chi(t)f_\epsilon(t,x,v)dt$$

is bounded in $L^2(\Omega\times\mathbf{R}^N;dxd\mu)$ by the same entropy inequality. Hence

$$0=\int_0^\infty\chi(t)f_\epsilon(t,\cdot,\cdot)dt\bigg|_{\Gamma^-}\ \to\ \int_0^\infty\chi(t)\rho(t,\cdot)dt\bigg|_{\Gamma^-}$$

in $L^2(\Gamma^-;|v\cdot n_x|\tau(x,v)\wedge 1d\sigma(x)dv)$ by Cessenat's trace theorem [10], where the notation $\tau(x,v)$ designates the forward exit time from Ω starting from x with velocity v, i.e.

$$\tau(x,v):=\inf\{t>0\text{ s.t. }x+tv\in\partial\Omega\},\quad x\in\Omega,\,v\in\mathbf{R}^N.$$

In particular

$$\int_0^\infty\chi(t)\rho(t,\cdot)dt\bigg|_{\partial\Omega}=0.$$

By (5.13), we already know that the limiting density $\rho\in L^2([0,T];H^1(\Omega))$. Therefore

$$\rho(t,\cdot)\big|_{\partial\Omega}=0\quad\text{in }L^2([0,T];H^{1/2}(\partial\Omega))\tag{5.21}$$

for each $T>0$.

Step 7: convergence to the diffusion equation.

Summarizing, we have proved that

$$f_\epsilon\text{ is relatively compact in }L^\infty(\mathbf{R}_+;L^2(\Omega\times\mathbf{R}^N,dxd\mu(v)))\text{ weak-*}$$

and that, if f is a limit point of f_ϵ as $\epsilon\to 0$, it is of the form

$$f(t,x,v)=\rho(t,x)\quad dtdxd\mu(v)\text{-a.e. in }(t,x,v)\in\mathbf{R}_+\times\Omega\times\mathbf{R}^N,$$

where

$$\begin{aligned}&\rho\in L^\infty(\mathbf{R}_+;\mathcal{H})\cap L^2(\mathbf{R}_+;H_0^1(\Omega))=L^\infty(\mathbf{R}_+;\mathcal{H})\cap L^2(\mathbf{R}_+;\mathcal{V})\\&\text{and }\partial_t\rho\in L^2(\mathbf{R}_+;\mathcal{V}'),\end{aligned}\tag{5.22}$$

since ρ satisfies the Dirichlet boundary condition (5.21) and $\nabla_x\rho\in L^2(\mathbf{R}_+\times\Omega)$ by (5.13), together with (5.20). In particular, this implies that

$$\rho\in C_b(\mathbf{R}_+;\mathcal{H}).\tag{5.23}$$

Besides ρ satisfies (5.16) for each test function $w \in \mathcal{V}$, together with the initial condition (5.19). Therefore ρ is the unique solution of the Dirichlet problem for the diffusion equation with diffusion matrix $M(x)$ defined in Theorem 2.1 (c) with infinite diffusivity in B, with initial data ρ^{in}. By compactness and uniqueness of the limit point, we conclude that

$$f_\epsilon \to \rho \quad \text{in } L^\infty(\mathbf{R}_+; L^2(\Omega \times \mathbf{R}^N, dxd\mu)) \text{ weak-*}$$

as $\epsilon \to 0$.

5.1. Proof of Theorem 2.2. Since $\mathcal{L}_x b^*(x,v) = v$ for $dxd\mu$-a.e. $(x,v) \in A \times \mathbf{R}^N$, we conclude from the uniqueness of the vector field b in statement (c) of Proposition 3.2 that $b(x,v) = b^*(x,v)$ for $dxd\mu$-a.e. $(x,v) \in A \times \mathbf{R}^N$. The identity in Proposition 3.2 (d) shows that

$$M_{ij}(x) = \int_{\mathbf{R}^N} b_i(x,v) v_j d\mu(v) = \int_{\mathbf{R}^N} v_i b_j(x,v) d\mu(v) = M_{ji}(x)$$

for a.e. $x \in A$ and all $i,j = 1,\ldots,N$. This proves statement (a) in Theorem 2.2.

It remains to prove the strong convergence in statement (b). The proof is based on the weak convergence already established in Theorem 2.1 and on a squeezing argument based on the entropy inequality for (2.8) and on the energy identity in (2.20).

Step 1: limiting entropy production.

By definition of q_ϵ, one has

$$q_\epsilon(t,x,v,w) = -q_\epsilon(t,x,w,v)$$

for $dtdxd\mu(v)d\mu(w)$-a.e. $(t,x,v,w) \in \mathbf{R}_+ \times A \times \mathbf{R}^N \times \mathbf{R}^N$ and each $\epsilon > 0$; by passing to the limit as $\epsilon \to 0$,

$$q(t,x,v,w) = -q(t,x,w,v)$$

for $dtdxd\mu(v)d\mu(w)$-a.e. $(t,x,v,w) \in \mathbf{R}_+ \times A \times \mathbf{R}^N \times \mathbf{R}^N$. Defining

$$k_A^s(x,v,w) = \tfrac{1}{2}(k_A(x,v,w) + k_A(x,w,v)),$$

one has

$$\left\langle\!\!\left\langle k_A(x,\cdot,\cdot) q(t,x,\cdot,\cdot)^2 \right\rangle\!\!\right\rangle = \left\langle\!\!\left\langle k_A^s(x,\cdot,\cdot) q(t,x,\cdot,\cdot)^2 \right\rangle\!\!\right\rangle$$

for a.e. $(t,x) \in \mathbf{R}_+ \times A$. Likewise,

$$\iint_{\mathbf{R}^N \times \mathbf{R}^N} k_A(x,v,w)(\phi(v) - \phi(w))^2 d\mu(v) d\mu(w)$$
$$= \iint_{\mathbf{R}^N \times \mathbf{R}^N} k_A^s(x,v,w)(\phi(v) - \phi(w))^2 d\mu(v) d\mu(w)$$

and

$$\iint_{\mathbf{R}^N \times \mathbf{R}^N} k_A(x,v,w)(\phi(v) - \phi(w)) q(t,x,v,w) d\mu(v) d\mu(w)$$
$$= \iint_{\mathbf{R}^N \times \mathbf{R}^N} k_A^s(x,v,w)(\phi(v) - \phi(w)) q(t,x,v,w) d\mu(v) d\mu(w),$$

for a.e. $(t,x) \in \mathbf{R}_+ \times A$. With $\phi(v) = \xi \cdot b(x,v)$ for some $\xi \in \mathbf{R}^N$ to be chosen later, and applying the Cauchy-Schwarz inequality, one finds that

$$\left(\iint_{\mathbf{R}^N \times \mathbf{R}^N} k_A^s(x,v,w) \xi \cdot (b(x,v) - b(x,w)) q(t,x,v,w) d\mu(v) d\mu(w) \right)^2$$
$$\leq \iint_{\mathbf{R}^N \times \mathbf{R}^N} k_A(x,v,w)(\xi \cdot (b(x,v) - b(x,w)))^2 d\mu(v) d\mu(w) \langle\!\langle k_A(x,\cdot,\cdot) q(t,x,\cdot,\cdot)^2 \rangle\!\rangle.$$
(5.24)

On the other hand, by definition of k_A^s

$$\iint_{\mathbf{R}^N \times \mathbf{R}^N} k_A^s(x,v,w) \xi \cdot (b(x,v) - b(x,w)) q_\epsilon(t,x,v,w) d\mu(v) d\mu(w)$$
$$= \frac{2}{\epsilon} \iint_{\mathbf{R}^N \times \mathbf{R}^N} k_A^s(x,v,w) \xi \cdot (b(x,v) - b(x,w)) f_\epsilon(t,x,v) d\mu(v) d\mu(w)$$
$$= \frac{1}{\epsilon} \langle f_\epsilon(t,x,\cdot)(\mathcal{L}_x + \mathcal{L}_x^*) \xi \cdot b(x,\cdot) \rangle = \frac{2}{\epsilon} \langle \xi \cdot v f_\epsilon(t,x,\cdot) \rangle$$

for a.e. $(t,x) \in \mathbf{R}_+ \times A$, where the last equality follows from the definitions of the vector fields b and b^* in Proposition 3.2 (c). Passing to the limit as $\epsilon \to 0$, one finds that

$$\iint_{\mathbf{R}^N \times \mathbf{R}^N} k_A^s(x,v,w) \xi \cdot (b(x,v) - b(x,w)) q(t,x,v,w) d\mu(v) d\mu(w)$$
$$= -2 \xi \cdot M(x) \nabla_x \rho(t,x)$$
(5.25)

for a.e. $(t,x) \in \mathbf{R}_+ \times A$, in view of (5.15) . On the other hand,

$$\iint_{\mathbf{R}^N \times \mathbf{R}^N} k_A(x,v,w)(\xi \cdot (b(x,v) - b(x,w)))^2 d\mu(v) d\mu(w)$$
$$= 2 \langle \xi \cdot b(x,\cdot) \mathcal{L}_x(\xi \cdot b(x,\cdot)) \rangle = 2 \langle \xi \cdot b(x,\cdot) \xi \cdot v \rangle = 2 \xi \cdot M(x) \xi$$
(5.26)

for a.e. $x \in A$, by Lemma 3.1 (c). Applying the Cauchy-Schwarz inequality in (5.25) and using (5.26) implies that

$$2(\xi \cdot M(x) \nabla_x \rho(t,x))^2 \leq \xi \cdot M(x) \xi \langle\!\langle k_A(x,\cdot,\cdot) q(t,x,\cdot,\cdot)^2 \rangle\!\rangle.$$

With the choice $\xi = \nabla_x \rho(t,x)$, we conclude that

$$2 \nabla_x \rho(t,x) \cdot M(x) \nabla_x \rho(t,x) \leq \langle\!\langle k_A(x,\cdot,\cdot) q(t,x,\cdot,\cdot)^2 \rangle\!\rangle$$
(5.27)

for a.e. $(t,x) \in \mathbf{R}_+ \times A$.

Step 2: strong convergence.

By Proposition 4.1 (b), for each $t \geq 0$ and each $\epsilon > 0$, one has

$$\int_\Omega \langle f_\epsilon(t,x,\cdot)^2 \rangle dx + \int_0^t \int_\Omega \langle\!\langle k_\epsilon(x,\cdot,\cdot) q_\epsilon(s,x,\cdot,\cdot)^2 \rangle\!\rangle dx ds \leq \int_\Omega \rho^{in}(x)^2 dx.$$

By Jensen's inequality,

$$\int_\Omega \langle f_\epsilon(t,x,\cdot)^2 \rangle dx \geq \int_\Omega \langle f_\epsilon(t,x,\cdot) \rangle^2 dx,$$

while, by convexity and weak convergence,

$$\varliminf_{\epsilon \to 0} \int_\Omega \langle f_\epsilon(t,x,\cdot)\rangle^2 dx \geq \int_\Omega \rho(t,x)^2 dx$$

for each $t \geq 0$ by (5.18). By the same token, for each $T \in [0,\infty]$,

$$\varliminf_{\epsilon \to 0} \int_0^T \int_A \langle\!\langle k_A(x,\cdot,\cdot)q_\epsilon(t,x,\cdot,\cdot)^2\rangle\!\rangle dxdt \geq \int_0^T \int_A \langle\!\langle k_A(x,\cdot,\cdot)q(t,x,\cdot,\cdot)^2\rangle\!\rangle dxdt. \quad (5.28)$$

Since the weak solution ρ of the diffusion problem (2.20) satisfies

$$\int_\Omega \rho(t,x)^2 dx + 2\int_0^t \int_A \nabla_x \rho(s,x) \cdot M(x)\nabla_x \rho(s,x)dxds = \int_\Omega \rho^{in}(x)^2 dx$$

for each $t > 0$ by Proposition 4.4, we conclude that

$$\int_\Omega \langle f_\epsilon(t,x,\cdot)^2\rangle dx \to \int_\Omega \rho(t,x)^2 dx \text{ for each } t \geq 0,$$

while

$$\int_0^\infty \int_A \langle\!\langle k_A(x,\cdot,\cdot)q_\epsilon(t,x,\cdot,\cdot)^2\rangle\!\rangle dxdt \to \int_0^\infty \int_A \langle\!\langle k_A(x,\cdot,\cdot)q(t,x,\cdot,\cdot)^2\rangle\!\rangle dxdt$$

$$= 2\int_0^\infty \int_A \nabla_x \rho(t,x) \cdot M(x)\nabla_x \rho(t,x)dxdt$$

as $\epsilon \to 0$ by (5.27).

Therefore,

$$f_\epsilon(t,\cdot,\cdot) \to \rho \text{ strongly in } L^2(\Omega \times \mathbf{R}^N, dxd\mu) \text{ for all } t \geq 0,$$

and

$$q_\epsilon \to q \text{ strongly in } L^2(\mathbf{R}_+ \times A \times \mathbf{R}^N \times \mathbf{R}^N, dtdxd(\mu \otimes \mu))$$

as $\epsilon \to 0$, so (5.27) is an equality. In other words, for a.e. $(t,x) \in \mathbf{R}_+ \times A$, one has

$$2\nabla_x \rho(t,x) \cdot M(x)\nabla_x \rho(t,x) = \langle\!\langle k_A(x,\cdot,\cdot)q(t,x,\cdot,\cdot)^2\rangle\!\rangle.$$

Therefore, q satisfies the equality in the Cauchy-Schwarz inequality (5.24) (with $\xi = \nabla_x \rho(t,x)$). This implies that q is of the form

$$q(t,x,v,w) = \lambda(t,x)(b(x,v) - b(x,w)) \cdot \nabla_x \rho(t,x),$$

for some measurable function λ defined a.e. on $\mathbf{R}_+ \times A$. Inserting this expression for q in (5.8), we find that

$$v \cdot \nabla_x \rho(t,x) = -\int_{\mathbf{R}^N} k_A(x,v,w)q(t,x,v,w)d\mu(w)$$

$$= -\lambda(t,x)\nabla_x \rho(t,x) \cdot \mathcal{L}_x b(x,v) = -\lambda(t,x)v \cdot \nabla_x \rho(t,x),$$

for a.e. $(t,x,v) \in \mathbf{R}_+ \times A \times \mathbf{R}^N$. Therefore $\lambda(t,x) = -1$ for a.e. $(t,x) \in \mathbf{R}_+ \times A$ such that $\nabla_x \rho(t,x) \neq 0$, so that

$$q(t,x,v,w) = -(b(x,v) - b(x,w)) \cdot \nabla_x \rho(t,x),$$

for a.e. $(t,x,v,w) \in \mathbf{R}_+ \times A \times \mathbf{R}^N \times \mathbf{R}^N$. Averaging in w, one finds that

$$\frac{1}{\epsilon}\left(f_\epsilon(t,x,v) - \int_{\mathbf{R}^N} f_\epsilon(t,x,w)d\mu(w)\right) = \int_{\mathbf{R}^N} q_\epsilon(t,x,v,w)d\mu(w)$$

$$\to \int_{\mathbf{R}^N} q(t,x,v,w)d\mu(w) = -b(x,v)\cdot\nabla_x\rho(t,x)$$

in the strong topology of $L^2(\mathbf{R}_+ \times A \times \mathbf{R}^N, dtdxd\mu)$ as $\epsilon \to 0$. By the uniqueness property in Proposition 3.2 (c), one has $b = b^*$ since $\mathcal{L}_x b^*(x,v) = v$ for $dxd\mu$-a.e. $(x,v) \in A \times \mathbf{R}^N$, and the conclusion of Theorem 2.2 follows.

6. Conclusions

The main results presented above (Theorems 2.1 and 2.2) can be generalized in several directions.

First, our method obviously applies to a scaled linear Boltzmann equation of the form

$$(\epsilon\partial_t + v\cdot\nabla_x)f_\epsilon(t,x,v) + \frac{1}{\epsilon}\mathcal{L}_x f_\epsilon(t,x,v) + \epsilon\mathcal{B}f_\epsilon(t,x,v) = \epsilon S(t,x,v),$$

where $\mathcal{B}$ is a bounded operator on $L^2(\Omega \times \mathbf{R}^N; dxd\mu)$, while the source term S is chosen so that $S \in L^1(\mathbf{R}_+; L^2(\Omega \times \mathbf{R}^N, dxd\mu))$. For instance $\mathcal{B}$ could be the multiplication by an amplifying or damping coefficient, i.e. $\mathcal{B}f_\epsilon(t,x,v) = \gamma(x)f_\epsilon(t,x,v)$ as in [5]. In other words, problems where the collision process is nearly (but not exactly) conservative can be treated exactly as above.

For some applications, for instance in the context of neutron transport theory, it would be important to extend the validity of the results presented in this paper to the case of scattering kernels which fail to satisfy the semi-detailed balance condition (2.3).

More general boundary conditions than the absorbing condition on $\partial\Omega$ can also be considered. For instance, imposing a specular or diffuse reflection condition at the boundary, or a convex combination thereof, i.e. assuming that

$$f_\epsilon(t,x,v) = (1-\theta(x))f_\epsilon(t,x,v-2v\cdot n_x n_x) + \frac{\theta(x)}{\langle(w\cdot n_x)_+\rangle}\int_{\mathbf{R}^N} f_\epsilon(t,x,w)(w\cdot n_x)_+d\mu(w)$$

with $\theta \in C(\partial\Omega)$ satisfying $0 \le \theta(x) \le 1$ for all $x \in \partial\Omega$ and with a measure μ invariant under all transformations of the form $v \mapsto Qv$ for $Q \in O_N(\mathbf{R})$ leads to the same result as in Theorems 2.1-2.2, except that the homogeneous Dirichlet condition on $\partial\Omega$ should be replaced with the homogeneous Neumann boundary condition.

Another possibility is to impose the density of incoming particles on $\partial\Omega$, in other words, to replace the absorbing boundary condition on $\partial\Omega$ with the condition

$$f_\epsilon(t,x,v) = f_b(t,x,v), \quad v\cdot n_x < 0, \, x \in \partial\Omega,$$

where f_b is given.

If f_b depends on the velocity variable v, one needs a boundary layer term of thickness $O(\epsilon)$ on $\partial\Omega$ in order to match the interior diffusion solution in A. This boundary layer term is, to leading order, governed by a half-space problem for the linear Boltzmann equation in space dimension 1: see for instance [5, 7, 14] for a detailed account of this theory. These boundary layer terms are used in general to correct the standard Hilbert type expansions of the solution of the linear Boltzmann equation in powers of the small parameter ϵ. As mentioned above, the Hilbert expansion method cannot be used on the problem considered in the present paper, at least not in its full generality.

Even if f_b is independent of the velocity variable v but depends on (t,x), the "entropy inequality" (statement (b) in Proposition 4.1) involves the flux of f_b^2 on $\partial\Omega$ in the right hand side of the inequality, and this term is of order $1/\epsilon$, which destroys the second uniform estimate in (5.1). In the nondegenerate case $B=\varnothing$, this difficulty can be circumvented by measuring the deviation of the solution f_ϵ to some appropriate profile $U(t,x)$ such that $U\big|_{\partial\Omega}=f_b$ instead of f_ϵ itself in the entropy inequality: see [16] for a precise description of the method. Whether this method can be used in the degenerate case where $B\neq\varnothing$ remains to be checked. One can probably combine the boundary layer theory above with the method of [16]; this again remains to be checked in detail.

Finally, since the compactness method used in the proof of Theorems 2.1-2.2, finds its origin in [4], we expect that the methods presented in this paper should also apply to some nonlinear problems, such as the radiative transfer equations.

Appendix A. Auxiliary lemmas on evolution equations. Let $\mathcal{V}$ and $\mathcal{H}$ be two separable Hilbert spaces such that $\mathcal{V}\subset\mathcal{H}$ with continuous inclusion and $\mathcal{V}$ is dense in $\mathcal{H}$. The Hilbert space $\mathcal{H}$ is identified with its dual and the map

$$\mathcal{H}\ni u\mapsto L_u\in\mathcal{V}',$$

where L_u is the linear functional

$$L_u:\mathcal{V}\ni v\mapsto(u|v)_\mathcal{H}\in\mathbf{R},$$

identifies $\mathcal{H}$ with a dense subspace of $\mathcal{V}'$.

LEMMA A.1. *Assume that*

$$v\in L^2(0,T;\mathcal{V})\quad\text{and}\quad\frac{dL_v}{dt}\in L^2(0,T;\mathcal{V}').$$

Then
(a) the function v is a.e. equal to a unique element of $C([0,T],\mathcal{H})$ still denoted v;
(b) this function $v\in C([0,T],\mathcal{H})$ satisfies

$$\tfrac{1}{2}|v(t_2)|_\mathcal{H}^2-\tfrac{1}{2}|v(t_1)|_\mathcal{H}^2=\int_{t_1}^{t_2}\left\langle\frac{dL_v}{dt}(t),v(t)\right\rangle_{\mathcal{V}',\mathcal{V}}dt$$

for all $t_1,t_2\in[0,T]$.

Statement (a) follows from Proposition 2.1 and Theorem 3.1 in chapter 1 of [22], and statement (b) from Theorem II.5.12 of [8].

LEMMA A.2. *Let $L\in L^2(0,T;\mathcal{V}')$ satisfy*

$$\langle L(t),w\rangle_{\mathcal{V}',\mathcal{V}}=0\ \text{for a.e.}\ t\in[0,T]$$

for all $w\in\mathcal{V}$. Then

$$L(t)=0\ \text{for a.e.}\ t\in[0,T].$$

Proof. Pick $\mathcal{N}_w\subset[0,T]$ negligible such that L is defined on $[0,T]\setminus\mathcal{N}_w$ and

$$\langle L(t),w\rangle_{\mathcal{V}',\mathcal{V}}=0\ \text{for all}\ t\in[0,T]\setminus\mathcal{N}_w.$$

Let $\mathcal{D}$ be a dense countable subset of $\mathcal{V}$ and let

$$\bar{\mathcal{N}} := \bigcup_{w \in \mathcal{D}} \mathcal{N}_w.$$

For all $t \in [0,T] \setminus \bar{\mathcal{N}}$, one has

$$\langle L(t), w \rangle_{\mathcal{V}', \mathcal{V}} = 0 \text{ for all } w \in \mathcal{D} \quad \text{so that } L(t) = 0$$

because $L(t)$ is a continuous linear functional on $\mathcal{V}$ and $\mathcal{D}$ is dense in $\mathcal{V}$. $\qquad\square$

The next lemma recalls the functional background for Green's formula in the context of evolution equations.

LEMMA A.3. *Let Ω be an open subset of $\mathbf{R}^N$ with smooth boundary, and let $T > 0$. Denote by n the unit outward normal field on $\partial\Omega$. Let $\rho \in C([0,T]; L^2(\Omega))$, and let $m \in L^2((0,T) \times \Omega, \mathbf{R}^N)$. Assume that*

$$\partial_t \rho + \mathrm{div}_x\, m = 0 \quad \text{in the sense of distributions in } (0,T) \times \Omega.$$

Then
(a) the vector field m has a normal trace[2] $m \cdot n\big|_{(0,T) \times \partial\Omega} \in H_{00}^{1/2}((0,T) \times \partial\Omega)'$;
(b) for each $\psi \in H^1(\Omega)$,

$$\frac{d}{dt} \int_\Omega \rho(\cdot, x) \psi(x) dx - \int_\Omega m(\cdot, x) \cdot \nabla_x \psi(x) dx = -\langle m \cdot n\big|_{\partial\Omega}, \psi\big|_{\partial\Omega}\rangle_{H^{-1/2}(\partial\Omega), H^{1/2}(\partial\Omega)}$$

in $H^{-1}(0,T)$.

Proof. Let $\chi \in C_c^\infty(\mathbf{R})$ be such that

$$\chi(t) = 1 \text{ for } t \in [-1, T+1] \quad \text{and} \quad \mathrm{supp}(\chi) \subset [-2, T+2].$$

Define

$$\bar{\rho}(t,x) := \begin{cases} \rho(t,x) & \text{if } 0 \le t \le T \\ \chi(t)\rho(0,x) & \text{if } t < 0 \\ \chi(t)\rho(T,x) & \text{if } t > T \end{cases}$$

and

$$\bar{m}(t,x) := \begin{cases} m(t,x) & \text{if } 0 \le t \le T \\ 0 & \text{if } t \notin [0,T], \end{cases}$$

so that the vector field $X := (\bar{\rho}, \bar{m})$ is an extension of (ρ, m) to $\mathbf{R} \times \Omega$ satisfying

$$X \in L^2(\mathbf{R} \times \Omega; \mathbf{R}^{N+1}).$$

Besides,

$$(\partial_t \bar{\rho} + \mathrm{div}_x\, \bar{m})(t,x) = \chi'(t)(\mathbf{1}_{t<0}\rho(0,x) + \mathbf{1}_{t>T}\rho(T,x)) =: S(t,x)$$

---

[2]We recall that $H_{00}^{1/2}((0,T) \times \partial\Omega)$ is the Lions-Magenes subspace of functions in $H^{1/2}((0,T) \times \partial\Omega)$ whose extension by 0 to $\mathbf{R} \times \partial\Omega$ defines an element of $H^{1/2}(\mathbf{R} \times \partial\Omega)$; the notation $H_{00}^{1/2}((0,T) \times \partial\Omega)'$ designates the dual of $H_{00}^{1/2}((0,T) \times \partial\Omega)$.

with $S \in L^2(\mathbf{R} \times \Omega)$, so that

$$\operatorname{div}_{t,x} X = S \in L^2(\mathbf{R} \times \Omega).$$

Therefore, X has a normal trace on the boundary $\partial(\mathbf{R} \times \Omega) = \mathbf{R} \times \partial\Omega$, which is denoted $X \cdot n\big|_{\mathbf{R} \times \partial\Omega} \in H^{-1/2}(\mathbf{R} \times \partial\Omega)$.

Let $\phi \in H_{00}^{1/2}((0,T) \times \partial\Omega)$; denote by $\bar{\phi}$ its extension by 0 to $\mathbf{R} \times \partial\Omega$. Therefore $\bar{\phi} \in H^{1/2}(\mathbf{R} \times \partial\Omega)$, and there exists $\bar{\Phi} \in H^1(\mathbf{R} \times \Omega)$ such that $\bar{\phi} = \bar{\Phi}\big|_{\mathbf{R} \times \partial\Omega}$. The normal trace of m is then defined as follows: by Green's formula,

$$\langle m \cdot n\big|_{\mathbf{R} \times \partial\Omega}, \phi \rangle_{H_{00}^{1/2}((0,T) \times \partial\Omega)', H_{00}^{1/2}((0,T) \times \partial\Omega)}$$

$$:= \langle X \cdot n\big|_{\mathbf{R} \times \partial\Omega}, \bar{\phi} \rangle_{H^{1/2}(\mathbf{R} \times \partial\Omega)', H^{1/2}(\mathbf{R} \times \partial\Omega)}$$

$$= \iint_{\mathbf{R} \times \Omega} (\bar{\rho} \partial_t \bar{\Phi} + \bar{m} \cdot \nabla_x \bar{\Phi} + S\bar{\Phi})(t,x) \, dx \, dt.$$

Applying Green's formula on $(0,T) \times \Omega$ shows that two different extensions of the vector field (ρ, m) define the same distribution $m \cdot n\big|_{(0,T) \times \partial\Omega}$ on $(0,T) \times \partial\Omega$. This completes the proof of statement (a).

As for statement (b), let $\kappa \in H_0^1(0,T)$ and $\psi \in H^1(\Omega)$, define $\Phi(t,x) := \kappa(t)\psi(x)$, and let $\bar{\Phi}$ be the extension of Φ by 0 to $\mathbf{R} \times \Omega$, so that $\bar{\Phi} \in H^1(\mathbf{R} \times \Omega)$. Thus $\phi = \Phi\big|_{(0,T) \times \partial\Omega}$ belongs to $H_{00}^{1/2}((0,T) \times \partial\Omega)$, and

$$\langle \langle m \cdot n\big|_{\partial\Omega}, \psi\big|_{\partial\Omega} \rangle_{H^{-1/2}(\partial\Omega), H^{1/2}(\partial\Omega)}, \kappa \rangle_{H^{-1}(0,T), H_0^1(0,T)}$$

$$:= \langle m \cdot n\big|_{(0,T) \times \partial\Omega}, \phi \rangle_{H_{00}^{1/2}((0,T) \times \partial\Omega)', H_{00}^{1/2}((0,T) \times \partial\Omega)}$$

$$= \iint_{\mathbf{R} \times \Omega} (\bar{\rho} \partial_t \bar{\Phi} + \bar{m} \cdot \nabla_x \bar{\Phi} + S\bar{\Phi})(t,x) \, dx \, dt$$

$$= \int_0^T \int_\Omega (\rho(t,x)\kappa'(t)\psi(x) + m(t,x) \cdot \nabla\psi(x)\kappa(t)) \, dx \, dt$$

$$= -\left\langle \frac{d}{dt} \int_\Omega \rho(t,x)\psi(x) \, dx, \kappa \right\rangle_{H^{-1}(0,T), H_0^1(0,T)} + \int_0^T \int_\Omega m(t,x) \cdot \nabla\psi(x)\kappa(t)) \, dx \, dt,$$

which is precisely the identity in statement (b). $\qquad\square$

REFERENCES

[1] G. Allaire and F. Golse, *Transport et diffusion*, Lecture notes, Ecole Polytechnique, 2013.

[2] G. Bal, *Transport through diffusive and nondiffusive regions, embedded objects, and clear layers*, SIAM J. Appl. Math., 62, 1677–1697, 2002.

[3] C. Bardos, F. Golse, and C.D. Levermore, *Fluid dynamic limits of kinetic equations. II. Convergence proofs for the Boltzmann equation*, Commun. Pure Appl. Math., 46, 667–753, 1993.

[4] C. Bardos, F. Golse, B. Perthame, and R. Sentis, *The nonaccretive radiative transfer equations: existence of solutions and Rosseland approximation*, J. Funct. Anal., 77, 434–460, 1988.

[5] C. Bardos, R. Santos, and R. Sentis, *Diffusion approximation and computation of the critical size*, Trans. Amer. Math. Soc., 284, 617–649, 1984.

[6] P. Benoist, *Théorie du coefficient de diffusion dans un réseau comportant des cavités*, Thèse de Doctorat d'Etat, Université de Paris, 21 Janvier 1964; Rapport CEA-R-2278, 1964.

[7] A. Bensoussan, J.-L. Lions, and G.C. Papanicolaou, *Boundary layers and homogenization of transport processes*, Publ. Res. Inst. Math. Sci. Kyoto., 15, 53–157, 1979.

[8] F. Boyer and P. Fabrie, *Eléments d'analyse pour l'étude de quelques modèles d'écoulements de fluides visqueux incompressibles*, Mathématiques et Applications, Springer-Verlag, Berlin, Heidelberg, 52, 2006.

[9] H. Brezis, *Analyse Fonctionnelle. Théorie et Applications*, Masson, Paris 1987. English translation "Functional Analysis, Sobolev Spaces and Partial Differential Equations", Springer, New York, 2011.

[10] M. Cessenat, *Théorèmes de trace L^p pour des espaces de fonctions de la neutronique*, C. R. Acad. Sci. Paris Sér. I Math., 299, 831–834, 1984.

[11] R. Dautray and J.-L. Lions, *Analyse mathématique et calcul numérique pour les sciences et les techniques*, Masson, Paris 1985. English translation: "Mathematical analysis and numerical methods for science and technology. Evolution problems. II", Springer-Verlag, Berlin, 6, 1993.

[12] L. Desvillettes and F. Golse, *A remark concerning the Chapman-Enskog asymptotics*. in Advances in kinetic theory and computing, B; Perthame (ed.), Ser. Adv. Math. Appl. Sci., World Sci. Publ., River Edge, NJ, 22, 191–203, 1994.

[13] L. Desvillettes, F. Golse, and V. Ricci, *Homogenization of the heat equation with infinitely conducting inclusions*, ESAIM Math. Model. Numer. Anal. 48, 1583–1613, 2014.

[14] F. Golse, S. Jin, and C.D. Levermore, *A domain decomposition analysis for a two-scale linear transport problem*, M2AN Math. Model. Numer. Anal., 37, 869–892, 2003.

[15] F. Golse and F. Poupaud, *Limite fluide des équations de Boltzmann des semi-conducteurs pour une statistique de Fermi-Dirac*, Asymptotic Anal., 6, 135–160, 1992.

[16] F. Golse and F. Salvarani, *The nonlinear diffusion limit for generalized Carleman models: the initial-boundary value problem*, Nonlinearity, 20, 927–942, 2007.

[17] L. Hörmander, *The Analysis of Linear Partial Differential Operators III*, Springer-Verlag, Berlin, Heidelberg, 1985.

[18] A.M. Il'in and R.Z. Has'minskiĭ, *On the equations of Brownian motion* (Russian), Teor. Verojatnost. i Primenen., 9, 466–491, 1964.

[19] L. Landau and E. Lifshitz, *Physical Kinetics*, Pergamon Press, Oxford, 1981.

[20] E.W. Larsen and J.B. Keller, *Asymptotic solution of neutron transport problems for small mean free paths*, J. Math. Phys., 15, 75–81, 1974.

[21] J.-L. Lions, *Problèmes aux limites et équations aux dérivées partielles*, Séminaire de Mathématiques Supérieures, (été 1962), Les Presses de l'Université de Montréal, Montréal, 1, 1965.

[22] J.-L. Lions and E. Magenes, *Problèmes aux limites non homogènes et applications*, Dunod, Paris, 1968. English translation "Non-homogeneous boundary value problems and applications. Vol. I", Springer-Verlag, New York-Heidelberg, 1972.

[23] H.A. Lorentz, *Le mouvement des électrons dans les métaux*, Arch. Neerl., 10, 336–371, 1905.

[24] B.W. Mihalas and D. Mihalas, *Foundations of Radiation Hydrodynamics*, Oxford University Press, 1984.

[25] G.C. Papanicolaou, *Asymptotic analysis of transport processes*, Bull. Amer. Math. Soc., 81, 330–392, 1975.

[26] G. Pomraning, *The Equations of Radiation Hydrodynamics*, Pergamon Press, 1973.

[27] D. Stroock and S.R.S. Varadhan, *Multidimensional Diffusion Processes*, Springer-Verlag, Berlin, Heidelberg, 2006.

[28] A.M. Weinberg and E.P. Wigner, *The Physical Theory of Neutron Chain Reactors*, The University of Chicago Press, 1958.