Stable Boundary Conditions and Discretization for P_N Equations

Jonas Bünger, Neeraj Sarna and Manuel Torrilhon

Center for Computational Engineering & Department of Mathematics
RWTH Aachen University, Germany[*]
$\big($buenger@mathcces.rwth-aachen.de,sarna@mathcces.rwth-aachen.de,mt@mathcces.rwth-aachen.de$\big)$

October 15, 2019

Abstract

A solution to the linear Boltzmann equation satisfies an energy bound, which reflects a natural fact: The number of particles in a finite volume is bounded in time by the number of particles initially occupying the volume augmented by the total number of particles that entered the domain over time. In this paper, we present boundary conditions (BCs) for the spherical harmonic (P_N) approximation, which ensure that this fundamental energy bound is satisfied by the P_N approximation. Our BCs are compatible with the characteristic waves of P_N equations and determine the incoming waves uniquely. Both, energy bound and compatibility, are shown based on abstract formulations of P_N equations and BCs to isolate the necessary structures and properties. The BCs are derived from a Marshak type formulation of BC and base on a non-classical even/odd-classification of spherical harmonic functions and a stabilization step, which is similar to the truncation of the series expansion in the P_N method. We show that summation by parts (SBP) finite differences on staggered grids in space and the method of simultaneous approximation terms (SAT) allows to maintain the energy bound also on the semi-discrete level.

1 Linear Boltzmann and Radiative Transfer Equation

The transport of particles in a background medium is governed by the *linear Boltzmann equation*[1]. In the absence of sources and absorption it takes the form

$$\frac{1}{|v(\epsilon)|}\partial_t\psi(t,\epsilon,x,\Omega) + \Omega\cdot\nabla_x\psi(t,\epsilon,x,\Omega) = \mathcal{Q}(t,\epsilon,x)\left[\psi(t,\epsilon,x,\Omega)\right] \tag{1}$$

where the unknown $\psi(t,\epsilon,x,\Omega)$ is the number density of particles, with respect to the measure $d\epsilon d\Omega dx$, located at x and moving in direction $\Omega\in\mathbb{S}^2$ with energy ϵ at time t. $|v(\epsilon)|$ is the absolute velocity of a particle with energy ϵ. The scattering operator $\mathcal{Q}$ describes the change in time due to angular deflections and energy-loss as particles interact with the background medium through collisions.

Two situations that are of particular practical relevance allow to consider models of the same structure but reduced phase space.

---

[*]Mathematics (CCES), Schinkelstr. 2, 52062 Aachen, Germany

1

1. When the energy of particles is fixed we can omit the energy variable and obtain the *radiative transfer equation*[2] (RT)

$$\partial_t \psi(t,x,\Omega) + \Omega \cdot \nabla_x \psi(t,x,\Omega) = \underbrace{\int_{S^2} \sigma_s(x,\Omega' \cdot \Omega)\psi(t,x,\Omega')\,d\Omega' - \sigma_t\psi(t,x,\Omega)}_{:=\mathcal{Q}^{RT}(x)\psi(t,x,\Omega)} \qquad (2)$$

where σ_s is a non-negative scattering cross-section consisting of the density of scattering centers $N_V(x)$ in the background medium such that $\sigma_s(x,\cdot)/N_V(x)$ is the probability density of the angular deflection; σ_t is the respective total cross-section. Note that σ_s, and thereby also σ_t, can depend on time.

2. Many situations allow to neglect the time dependency and it can be assumed that particles loose a significant amount of energy only by a sequence of collisions with each collision changing the energy only slightly. In these situations one usually describes the particle system in terms of the particle fluence $\hat{\psi}(x,\epsilon,\Omega) := |v(\epsilon)|\psi(x,\epsilon,\Omega)$ and employs an evolution equation in energy space called *Boltzmann equation in continuous slowing down approximation*[3] (BCSD)

$$-\partial_\epsilon\left(S(\epsilon,x)\hat{\psi}(\epsilon,x,\Omega)\right) + \Omega \cdot \nabla_x\hat{\psi}(\epsilon,x,\Omega) = \underbrace{\int_{S^2} \tilde{\sigma}_s(\epsilon,x,\Omega'\cdot\Omega)\,\hat{\psi}(\epsilon,x,\Omega')\,d\Omega' - \tilde{\sigma}_t(\epsilon,x)\,\hat{\psi}(\epsilon,x,\Omega)}_{:=\mathcal{Q}^{CSD}(\epsilon,x)\hat{\psi}(\epsilon,x,\Omega)},$$

$$(3)$$

where the stopping power S describes the average energy loss per distance traveled, $\tilde{\sigma}_s$ denotes a scattering cross-section involving elastic and inelastic collisions and $\tilde{\sigma}_t$ is the joint total scattering cross-section of elastic and inelastic collisions.

Note that the BCSD can be transformed into a pseudo equation of radiation transport if the stopping power S is constant with respect to energy ϵ using the variable transformation $\epsilon(t) = \epsilon_{\max} - t$.

From the viewpoint of this work, all three equations (eqs. (1) to (3)) are of the same structure. We will focus on the equation for radiative transfer in the following analysis.

The **boundary conditions** for radiative transfer are of Dirichlet type and prescribe the incoming half of the particle distribution at the boundary:

$$\psi(t,x,\Omega) \overset{!}{=} \psi_{in}(t,x,\Omega) \qquad \forall\Omega \in \mathbb{S}^2 : n \cdot \Omega < 0, \qquad (4)$$

where n denotes the outward-pointing normal vector at $x \in \partial G$ and, with $(\partial G \times \mathbb{S}^2)_- := \{(x,\Omega) \in \partial G \times \mathbb{S}^2 : n \cdot \Omega < 0\}$, $\psi_{in} : \mathbb{R}_+ \times (\partial G \times \mathbb{S}^2)_- \to \mathbb{R}_+$ is a given distribution of incoming particles. It is common sense, that the number of particles in a bounded domain $G \subset \mathbb{R}^3$ at time T is bounded from above by the number of particles that initially ($t = 0$) occupied the domain augmented by the number of particles that entered through the boundary. This natural fact is reflected by the following *energy bound* for a solution ψ of of radiative transfer equation (2), see also [4, 5]

$$\|\psi(T,\,\cdot\,)\|^2_{L^2(G\times\mathbb{S}^2)} \le \|\psi_0\|^2_{L^2(G\times\mathbb{S}^2)} + \|\psi_{\mathrm{in}}\|^2_{L^2((0,T)\times(\partial G\times\mathbb{S}^2)_-)}. \qquad (5)$$

Note that this energy bound, due to linearity of eq. (1) in ψ, implies continuous dependency of solutions on the initial and boundary data given by ψ_0 and ψ_{in}, and assures uniqueness of solutions. For the

particular case that no particles enter the domain $\psi_{\text{in}}(t, \cdot) \equiv 0$ at any time t, we can deduce that the number of particles cannot increase, i.e. the energy bound becomes

$$\partial_t \|\psi(t, \cdot)\|^2_{L^2(G \times \mathbb{S}^2)} \leq 0. \tag{6}$$

The energy bound (5) reflects a fundamental property of the physics described by the radiative transfer equation and is relevant for the analysis of well posedness in terms of uniqueness and continuous dependency on data of solutions. Therefore it is desirable to assure an analogous energy bound also for reduced models aiming to approximate the linear Boltzmann equation, such as the P_{N} approximation considered in this work.

In this paper we present boundary conditions for the P_{N} approximation of radiative transfer, which assure an energy bound analogous to (5) for the approximate solution and are compatible with the P_{N} equations, i.e. incoming characteristic waves are uniquely determined by outgoing characteristic waves and boundary data. We introduce the boundary conditions in two steps. In the first step (section 2), we investigate a generic first order system of linear transport equations with relaxation, an abstraction of P_{N} equations, and show stability and compatibility for boundary conditions which have a particular structure. In the second step (section 3), we derive the boundary conditions for P_{N} equations and show that these are of the desired structure. In section 4 we describe a discretization of P_{N} equations in space using summation by parts finite difference on staggered grid for spatial derivatives and simultaneous approximation terms for the boundary conditions, show that a semi discrete solution satisfies an energy bound analogous to (5) and present first numerical results in section 5.

2 Boundary conditions for Onsager Compatible Systems

In this section we investigate boundary conditions of a specific structure for an abstract form of transport equations.

Problem 1. *Let $G \subset \mathbb{R}^d$. Find solution $u : \mathbb{R}_+ \times G \to \mathbb{R}^m$ of the first-order system*

$$\partial_t u + \sum_{i=1}^{d} A^{(i)} \partial_{x_i} u = Pu \qquad \forall x \in G, \ \forall t > 0 \tag{7}$$

with transport matrices $A^{(i)} \in \mathbb{R}^{m \times m}$ and negativ semi definit relaxation matrix $P \in \mathbb{R}^{m \times m}$, for given initial datum $u_0 : G \to \mathbb{R}^m$

$$u(0, \cdot) = u_0 \qquad \forall x \in G. \tag{8}$$

We augment Problem 1 with the following assumptions.

Assumption 2 (Rotational Invariance)**.** *Equation (7) is rotational invariant, i.e. for a orthogonal matrix $\mathcal{R} \in \mathbb{R}^{d \times d}$ mapping a cartesian coordinate system C onto $\hat{C}$ there is an orthogonal rotation matrix $Q(\mathcal{R}) \in \mathbb{R}^{m \times m}$ that commutes with the spatial operator $\mathcal{L}u := \sum_{i=1}^{d} A^{(i)} \partial_{x_i} u + Pu$ in the sense that*

$$\hat{\mathcal{L}} \circ Q(\mathcal{R}) = Q(\mathcal{R}) \circ \mathcal{L}, \tag{9}$$

where $\hat{\mathcal{L}}u := \sum_{i=1}^{d} A^{(i)} \partial_{\hat{x}_i} u + Pu$ with $\partial_{\hat{x}_i} \cdot = \sum_{j=1}^{d} \mathcal{R}_{ji} \partial_{x_j} \cdot$.

The rotational invariance allows to express the directional transport matrix $A^{(n)} := \sum_{i=1}^{d} n_i A^{(i)}$ with a unit direction $n \in \mathbb{S}^{d-1}$ in terms of a transport matrix $A^{(j)}$ through rotation

$$A^{(n)} = Q(R_{n \to e_j}) A^{(j)} \left(Q(R_{n \to e_j}) \right)^T \tag{10}$$

where $R_{n \to e_j}$ denotes a rotation that maps direction n onto unit vector e_j. Vice versa, we assume $A^{(n)}$ to be given, use eq. (10) to express $A^{(i)}$ in terms of $A^{(n)}$ and obtain

$$\sum_{i=1}^{d} A^{(i)} \partial_{x_i}(\cdot) = \sum_{i=1}^{d} \left(Q(R_{e_i \to n}) A^{(n)} \left(Q(R_{e_i \to n}) \right)^T \right) \partial_{x_i}(\cdot). \tag{11}$$

Consequently, it suffices to characterize the transport operator by essentially one transport matrix $A^{(n)}$.

Assumption 3 (Block Structure Symmetry/Onsager Compatibility). For any given direction n the variables in u can be ordered such that the transport matrix $A^{(n)}$ takes the form

$$A^{(n)} = \begin{pmatrix} 0 & \hat{A}^{(n)} \\ \left(\hat{A}^{(n)} \right)^T & 0 \end{pmatrix} \tag{12}$$

where $\hat{A}^{(n)} \in \mathbb{R}^{r \times s}$ has full row rank, i.e. $r \leq s$ and $\text{rank}(\hat{A}^{(n)}) = r$ independent of n.

We introduce compact notation for a splitting in u with respect to a direction $n \in \mathbb{S}^{d-1}$:

$$u^{n+} := (Q^{n+})^T u \qquad u^{n-} := (Q^{n-})^T u, \tag{13}$$

where $Q^{n+} \in \mathbb{R}^{m \times r}$ and $Q^{n-} \in \mathbb{R}^{m \times (m-r)}$ such that $Q(R_{n \to e_j}) = [Q^{n+}, Q^{n-}]$ for some rotation $R_{n \to e_j}$. The splitting is induced by the block structure of $A^{(n)}$ in the sense that

$$u^T A^{(n)} u = (u^{n+})^T \hat{A} u^{n-} + (u^{n-})^T \hat{A}^T u^{n+}. \tag{14}$$

Later we will refer to the variable sets u^{n+} and u^{n-} as odd and even. Note that the symmetry of the transport matrix $A^{(n)}$ also implies hyperbolicity of eq. (7).

We define the class of boundary conditions investigated in this work, similar to the work done in [5, 6, 7].

Definition 4 (Onsager Boundary Condition). Let $n \in \mathbb{S}^{d-1}$ denotes the outward-pointing boundary normal at the respective boundary point $x \in \partial G$. We call a boundary condition *Onsager boundary condition*, if it is of the form

$$u^{n+}(x) = L \hat{A} u^{n-}(x) + g(x) \qquad \forall x \in \partial G, \tag{15}$$

with

(i) $L \in \mathbb{R}^{r \times r}$ symmetric semi positive definite and

(ii) boundary source $g(x) \in \text{Im}(L)$.

In the remainder of this section we show that OBCs imply an energy bound for a solution u in terms of its data u_0 and g, and verify their compatibility with the hyperbolic characteristics of the field equation (7). We start with proofing the energy bound, where we use $\|v\|_{L^2(D)}$ to denote the energy-norm of a vector valued function $v : D \to \mathbb{R}^s$ with $v_i \in L^2(D)$

$$\|v\|_{L^2(D)} := \int_D (v(x))^T v(x)\, dx. \tag{16}$$

Theorem 5 (Energy Bound). *A solution u to problem 1 with Onsager boundary condition is bounded by the initial and boundary data*

$$\|u(T, \cdot)\|^2_{L^2(G)} \le \|u_0\|^2_{L^2(G)} + \|L^+\|^2_2 \|g\|^2_{L^2((0,T) \times \partial G)}, \tag{17}$$

where $\|L^+\|_2$ denotes the operator norm of the pseudo inverse L^+ of L, i.e. $\|L^+\|_2 = \max\{\lambda^{-1} : \lambda \in \sigma(L) \setminus \{0\}\}$. For the particular case that boundary source vanishes $g(T, \cdot) \equiv 0$ at all times t we get that energy cannot increase

$$\partial_t \|u\|_{L^2(G)} \le 0. \tag{18}$$

Proof. We multiply equation (7) with u^T from the left, use the product rule of differentiation and integrate over the domain G to obtain

$$\partial_t u + \sum_{i=1}^d A_i \partial_{x_i} u + Pu = 0 \quad \Rightarrow \quad \partial_t \|u\|^2_{L^2(G)} = -\int_G \sum_{i=1}^d \partial_{x_i}(u^T A_i u)\, dx + 2 \int_G u^T Pu\, dx.$$

Applying the Gauss-theorem and using eq. (14) we obtain

$$\partial_t \|u\|^2_{L^2(G)} = -\int_{\partial G} u^T \underbrace{\left(\sum_{i=1}^d n_i A_i \right)}_{A^{(n)}} u\, dS + 2 \int_G u^T Pu\, dx = -2 \int_{\partial G} (u^{n+})^T \hat{A} u^{n-}\, dS + 2 \int_G u^T Pu\, dx.$$

Now we use the Onsager boundary condition to express u^{n+} in terms of u^{n-} and g to obtain

$$\partial_t \|u\|^2_{L^2(G)} = -2 \int_{\partial G} (L\hat{A} u^{n-} + g)^T \hat{A} u^{n-}\, dS + 2 \int_G u^T Pu\, dx$$

$$= \int_{\partial G} \underbrace{-2g^T \hat{A} u^{n-} - (\hat{A} u^{n-})^T L(\hat{A} u^{n-})}_{(v:=-\hat{A} u^{n-})\, 2g^T v - v^T Lv}\, dS + \int_{\partial G} \underbrace{-(\hat{A} u^{n-})^T L(\hat{A} u^{n-})}_{\le 0}\, dS + 2 \int_G \underbrace{u^T Pu}_{\le 0}\, dx.$$

We now bound $2g^T v - v^T Lv$ from above by g and L: As g is in the image of L, we can always find a $v_g \in \mathbb{R}^c$ with $g = Lv_g$ and with L symmetric positive-definite we obtain $\forall v$

$$2g^T v - v^T Lv = v_g{}^T Lv_g - (v - v_g)^T L(v - v_g) \le |v_g{}^T Lv_g| = |g^T v_g|$$

$$\le \|g\|_2 \|v_g\|_2 = \|g\|_2 \|L^+ g\|_2 \le \|L^+\|_2 \|g\|_2^2 = \|L^+\|_2\, g^T g.$$

With the above inequality

$$\partial_t \|u\|^2_{L^2(G)} \le \|L^+\|_2 \int_{\partial G} g^T g\, dS = \|L^+\|_2 \|g(t, \cdot)\|^2_{L^2(\partial G)}$$

and we obtain the energy estimate

$$\|u(T,\cdot)\|^2_{L^2(G)} \le \|u_0\|^2_{L^2(G)} + \|L^+\|_2 \|\|g\|^2_{L^2((0,T)\times\partial G)}.$$

$\square$

It remains to show that OBCs are compatible with the characteristics of the field equations (7), which we show using the following result on the eigenstructure of Onsager compatible matrices.

Theorem 6 (Eigenstructure of Onsager Compatible Matrices). *Let $A \in \mathbb{R}^{m\times m}$ be an Onsager compatible matrix such that*

$$A = \begin{pmatrix} 0 & \hat{A} \\ \hat{A}^T & 0 \end{pmatrix}.$$

$\hat{A} \in \mathbb{R}^{r\times(m-r)}$. *Then A has an eigendecomposition $A = X\Lambda X^T$ with*

$$X = \frac{1}{\sqrt{2}} \begin{pmatrix} \hat{X} & 0 & \hat{X} \\ \tilde{X} & \sqrt{2}X^k & -\tilde{X} \end{pmatrix} \in \mathbb{R}^{m\times m} \qquad \Lambda = \begin{pmatrix} \Lambda_p & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -\Lambda_p \end{pmatrix} \in \mathbb{R}^{m\times m} \tag{19}$$

where

(i) $\Lambda_p \in \mathbb{R}^{r\times r}$ *diagonal positive definit,*

(ii) $\hat{X} \in \mathbb{R}^{r\times r}$ *orthogonal,*

(iii) $\tilde{X} \in \mathbb{R}^{(m-r)\times r}$ *has orthonormal columns and*

(iv) *the columns of $X^k \in \mathbb{R}^{(m-r)\times(m-2r)}$ form an orthonormal basis of $\ker(\hat{A})$ and are orthogonal to the columns of $\tilde{X}$.*

Proof. We can always write $\hat{A}$ by its singular value decomposition, say $\hat{A} = U\Sigma V^T$ where $U \in \mathbb{R}^{r\times r}, V \in \mathbb{R}^{(m-r)\times(m-r)}$ orthogonal and $\Sigma \in \mathbb{R}^{r\times(m-r)}$ with $\Sigma_{ij} = 0$ for $i \ne j$. We identify $V = [V_1, V_2]$ with $V_1 \in \mathbb{R}^{(m-r)\times r}$ and $V_2 \in \mathbb{R}^{(m-r)\times((m-2r)}$, and $\Sigma = [\Sigma_1, 0]$ with $\Sigma_1 \in \mathbb{R}^{r\times r}$, such that $U\Sigma_1 V_1^T = \hat{A}$. Choosing $U = \hat{X}, \tilde{X} = V_1, X^k = V_2, \Lambda_p = \Sigma_1$ yields the result. $\square$

The structure of the diagonal matrix Λ in (19) tells us that the spectrum of the transport matrix $A^{(n)}$, is symmetric about the origin. We can deduce that at a boundary point we have r incoming, r outgoing and $(m-2r)$ standing characteristic waves. Existence and uniqueness of solutions to problem 1 in the context of hyperbolic equations require that the boundary conditions uniquely determine the incoming, and only the incoming, characteristic waves. In order to allow an energy bound of the form in (5) the incoming waves must be determined by the boundary data and outgoing waves, and be independent of incoming waves. From the following theorem we can immediately deduce that Onsager boundary conditions satisfy the aforementioned requirements.

Theorem 7 (Characteristic Compatibility). *Onsager boundary condition (15) is equivalent to the boundary condition*

$$(L\hat{X}\Lambda_p + \hat{X})w_- = (L\hat{X}\Lambda_p - \hat{X})w_+ + \sqrt{2}g \tag{20}$$

for the outgoing and incoming characteristic variables w_+ and w_-, given by

$$\begin{pmatrix} w_- \\ w_+ \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \hat{X}^T & \tilde{X}^T \\ \hat{X}^T & -\tilde{X}^T \end{pmatrix} \begin{pmatrix} u^{n+} \\ u^{n-} \end{pmatrix}. \tag{21}$$

Matrix $(L\hat{X}\Lambda_p + \hat{X})$ is invertible and therefore Onsager boundary conditions uniquely determine the incoming characteristic waves by the outgoing waves and the boundary data.

6

Proof. To prove the equivalence of the Onsager boundary condition (15) and the characteristic boundary condition (20) we first express u^{n+} and u^{n-} by the characteristic variables

$$\begin{pmatrix} u^{n+} \\ u^{n-} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \hat{X} & 0 & \hat{X} \\ \tilde{X} & \sqrt{2}X^k & -\tilde{X} \end{pmatrix} \begin{pmatrix} w_+ \\ w_0 \\ w_- \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \hat{X}(w_+ + w_-) \\ \tilde{X}(w_+ - w_-) + \sqrt{2}X^k w_0 \end{pmatrix}$$

where w_+ denotes the outgoing, w_- the incoming characteristics and w_0 the characteristics with zero velocity in normal direction. We replace u^{n+} and u^{n-} in the boundary condition by their expression in characteristic variables, factor out the characteristic variables and with $\hat{A} = \hat{X}\Lambda_p\tilde{X}^T$ we obtain

$$u^{n+} = L\hat{A}u^{n-} + g$$
$$\Leftrightarrow \quad \hat{X}(w_+ + w_-) = L(\hat{X}\Lambda_p\tilde{X}^T)(\tilde{X}(w_+ - w_-) + \sqrt{2}X^k w_0) + \sqrt{2}g$$
$$\Leftrightarrow \quad (L\hat{X}\Lambda_p + \hat{X})w_- = (L\hat{X}\Lambda_p - \hat{X})w_+ + \sqrt{2}g.$$

The matrix $(L\hat{X}\Lambda_p + \hat{X})$ is invertible as we can write it as a product of invertible matrices

$$(L\hat{X}\Lambda_p + \hat{X}) = \hat{X} \underbrace{(\hat{X}^T L\hat{X} + \Lambda_p^{-1})}_{\text{positiv definit}} \Lambda_p.$$

$\square$

3 Stable Boundary Conditions for P_{N} equations

In this section we present boundary conditions, which ensure that the P_{N} approximation maintains the energy estimate (5). We will exploit the previous section 2: First, in section 3.1, we embed P_{N} equations into the context of section 2, i.e. we show that P_{N} equations are instances of the abstract field equations (7) in problem 1. Secondly, in section 3.2, we discuss the formulation of boundary conditions for P_{N} equations and present Onsager boundary conditions.

3.1 P_{N} Approximation

3.1.1 Motivation

Due to the high dimension of the number density ψ, in general ψ is a function of six variables, solving radiation transport equation by a direct discretization in phase space is usually computationally too expensive. Besides the curse of dimensionality, the evolution of ψ is non-local as the collision operator introduces a coupling in energy and direction space. Therefore it is usually necessary to use reduced models, one of which are so-called P_{N} models, which we consider in this work.

A P_{N} model follows from a spectral Galerkin discretization in direction space using spherical harmonic functions as basis [8]. The number density ψ is approximated by an expansion in spherical harmonic functions $Y_l^k : \mathbb{S}^2 \to \mathbb{R}$ up to order $N \in \mathbb{N}$

$$\psi(t,x,\Omega) \approx \psi_{P_{\mathrm{N}}}(t,x,\Omega) = \sum_{l \leq N, |k| \leq l} u_l^k(t,x)Y_l^k(\Omega), \tag{22}$$

where the expansion coefficients are the spherical harmonic moments of the angular distribution. Evolution equations for the coefficients u_l^k, the P_{N} equations, are obtained by testing the equation of

n	Y_l^k even	Y_l^k odd
e_x	$(k < 0 \wedge k \text{ odd})$ $\vee (k \geq 0 \wedge k \text{ even})$	$(k < 0 \wedge k \text{ even})$ $\vee (k \geq 0 \wedge k \text{ odd})$
e_y	$k \geq 0$	$k < 0$
e_z	$(l + k)$ even	$(l + k)$ odd

Table 1: Even/Odd classification formulas for spherical harmonic functions with respect to the standard basis vectors e_1, e_2 and e_3 of the Cartesian coordinate system.

radiation transport with spherical harmonic functions up to order N

$$\int_{S^2} (\partial_t \psi_{P_N} + \Omega \cdot \nabla_x \psi_{P_N}) Y_{l'}^{k'} \, d\Omega = \int_{S^2} (\mathcal{Q}^{RT} \psi_{P_N}) Y_{l'}^{k'} \, d\Omega \qquad \forall l' \leq N, \forall |k'| \leq l'. \tag{23}$$

The phase space gets reduced by two dimensions at the cost of replacing a scalar equation by a system of equations. More importantly, the spherical harmonic functions are eigenfunctions of the scattering operator whenever the cross sections are isotropic, i.e. depend only on the deflection angle $(\Omega \cdot \Omega')$, which reduces the computational cost of the scattering operator in the Galerkin method, as the expansion coefficients decouple in direction space.

3.1.2 Spherical Harmonic Functions and Notation

Spherical harmonic functions form a *complete orthonormal set* of the space $L^2(\mathbb{S}^2; \mathbb{R})$. The real spherical harmonic $Y_l^k : \mathbb{S}^2 \to \mathbb{R}$ of degree $l \in \mathbb{N}_0$ and order k ($|k| \leq l$) is given by

$$Y_l^k(\Omega(\mu, \varphi)) C_{l,|k|} P_l^{|k|}(\mu) \begin{cases} \cos(|k|\varphi) & k > 0 \\ 1/\sqrt{2} & k = 0 \\ \sin(|k|\varphi) & k < 0, \end{cases} \qquad C_{l,|k|} = (-1)^{|k|} \sqrt{\frac{2l+1}{2\pi} \frac{(l-|k|)!}{(l+|k|)!}} \tag{24}$$

where $P_l^{|k|}$ denotes the associated Legendre function of respective degree and order, and μ and φ are the polar and azimuthal angle of direction Ω. These functions are well-studied, e.g. [9, 10]. We recapitulate properties essential to this work in a form that is customized to an observation we will use to classify spherical harmonics: A spherical harmonic Y_l^k is either even or odd in the direction of a cartesian axes, where we call a function $f : S^2 \to \mathbb{R}^n$ *even/odd in direction* $n \in \mathbb{S}^2$, if

$$f(\Omega - 2(n \cdot \Omega)n) = +/- f(\Omega) \qquad \forall \Omega \in \mathbb{S}^2. \tag{25}$$

Table 1 gives simple formulas for the even/odd classification based on the degree l and order k.

When considering all three cartesian axes simultaneously the set of all spherical harmonics is separated into $2^3 = 8$ non-trivial sets, which can easily be verified using the formulas. Note, that our classification is different from the usual even/odd classification of spherical harmonics based on the *parity* referring to inversion about the origin, which solely depends on the even-odd property of the integer degree l, that is, $Y_l^k(-\Omega) = (-1)^l Y_l^k(\Omega)$.

For the sake of compact notation we introduce abbreviations for vector functions of spherical harmonics

$$\bar{\Upsilon}_l(\Omega) := (Y_l^{-l}(\Omega), Y_l^{-l+1}(\Omega), \ldots, Y_l^l(\Omega))^T, \qquad \bar{\Upsilon}_l : \mathbb{S}^2 \to \mathbb{R}^{2l+1} \tag{26}$$

$$\Upsilon_N(\Omega) := (\bar{\Upsilon}_0(\Omega)^T, \bar{\Upsilon}_1(\Omega)^T, \ldots, \bar{\Upsilon}_N(\Omega)^T)^T. \qquad \Upsilon_N : \mathbb{S}^2 \to \mathbb{R}^{(N+1)^2} \tag{27}$$

We use $\bar{\Upsilon}_l^{e,i}(\Omega) \in \mathbb{R}^{l+1}$ to denote a vector function containing the spherical harmonic functions of degree l, that are even with respect to the direction e_i ($i \in \{1,2,3\}$) and denote the odd counterpart by $\bar{\Upsilon}_l^{o,i}(\Omega) \in \mathbb{R}^l$. As above we also designate their collections up to degree N

$$\Upsilon_{\mathrm{N}}^{e,i}(\Omega) := (\bar{\Upsilon}_0^{e,i}(\Omega)^T, \ldots, \bar{\Upsilon}_{\mathrm{N}}^{e,i}(\Omega)^T)^T, \qquad \Upsilon_{\mathrm{N}}^{e,i} : S^2 \to \mathbb{R}^{\frac{(N+2)(N+1)}{2}} \tag{28}$$

$$\Upsilon_{\mathrm{N}}^{o,i}(\Omega) := (\bar{\Upsilon}_0^{o,i}(\Omega)^T, \ldots, \bar{\Upsilon}_{\mathrm{N}}^{o,i}(\Omega)^T)^T, \qquad \Upsilon_{\mathrm{N}}^{o,i} : S^2 \to \mathbb{R}^{\frac{N(N+1)}{2}}. \tag{29}$$

As scalar producst for given $v \in L^2(\mathbb{S}^2; \mathbb{R}^r)$ and $w \in L^2(\mathbb{S}^2; \mathbb{R}^c)$ we use $\langle v, w^T \rangle$ and $\langle v, w^T \rangle_{n_\pm}$ to denote the $(r \times c)$-matrices defined by

$$\left(\langle v, w^T \rangle \right)_{ij} := \int_{S^2} v_i\, w_j\, d\Omega \qquad \left(\langle v, w^T \rangle_{n_\pm} \right)_{ij} := \int_{\Omega \cdot n \gtrless 0} v_i\, w_j\, d\Omega.$$

We now recapitulate two properties which, alongside orthogonality, ensure that P_N equations are instances of the abstract field equations eq. (7). First, the function space spanned by spherical harmonics of degree l is *rotational invariant* in the sense that for a rotation $\mathcal{R}$ about the origin mapping a unit vector Ω onto $\Omega' = \mathcal{R}\Omega$ the following equality holds

$$Y_l^k(\mathcal{R}\Omega) = Y_l^k(\Omega') = \sum_{k'=-l}^{l} D_{k'}^l(\mathcal{R}) Y_l^{k'}(\Omega) \quad \text{with} \quad D_{k'}^l(\mathcal{R}) = \langle Y_l^k(\mathcal{R} \cdot), Y_l^{k'}(\cdot) \rangle. \tag{30}$$

We can immediately deduce that for the vector of spherical harmonics up to order N

$$\Upsilon_{\mathrm{N}}(\mathcal{R}\Omega) = D(\mathcal{R}) \Upsilon_{\mathrm{N}}(\Omega) \quad \text{with} \quad D(\mathcal{R}) := \langle \Upsilon_{\mathrm{N}}(\mathcal{R} \cdot), \Upsilon_{\mathrm{N}}(\cdot)^T \rangle. \tag{31}$$

Secondly, spherical harmonics fulfill *recursion relations*, which, in order to emphasize the coupling between even and odd functions, we write as

$$\Omega_i \bar{\Upsilon}_l o, i(\Omega) = \hat{A}_{l,-}^{(i)} \bar{\Upsilon}_{l\text{-}1} e, i(\Omega) + \hat{A}_{l,+}^{(i)} \bar{\Upsilon}_{l+1} e, i(\Omega) \tag{32a}$$

$$\Omega_i \bar{\Upsilon}_l e, i(\Omega) = (\hat{A}_{l-1,+}^{(i)})^T \bar{\Upsilon}_{l\text{-}1} o, i(\Omega) + (\hat{A}_{l+1,-}^{(i)})^T \bar{\Upsilon}_{l+1} o, i(\Omega) \tag{32b}$$

with

$$\hat{A}_{l,\pm}^{(i)} = \langle \Omega_i \bar{\Upsilon}_l o, i, (\bar{\Upsilon}_{l\pm 1} e, i)^T \rangle \in \mathbb{R}^{l \times (l\pm 1+1)}. \tag{33}$$

In order to point out the specific structure of the recursion relation more clearly, we schematically write out the recursion relation for all spherical harmonics. Gathering all even and odd spherical harmonics, we can write the recursion relation as

$$\Omega_i \begin{pmatrix} \Upsilon_\infty^{o,i}(\Omega) \\ \Upsilon_\infty^{e,i}(\Omega) \end{pmatrix} = \begin{pmatrix} 0 & \hat{A}^{(i)} \\ \left(\hat{A}^{(i)} \right)^T & 0 \end{pmatrix} \begin{pmatrix} \Upsilon_\infty^{o,i}(\Omega) \\ \Upsilon_\infty^{e,i}(\Omega) \end{pmatrix} \tag{34}$$

with the infinite matrix

$$\hat{A}^{(i)} = \langle \Omega_i \Upsilon_\infty^{o,i}, \left(\Upsilon_\infty^{e,i} \right)^T \rangle = \begin{pmatrix} \hat{A}_{1,-}^{(i)} & 0 & \hat{A}_{1,+}^{(i)} & 0 & \cdots \\ 0 & \hat{A}_{2,-}^{(i)} & 0 & \hat{A}_{2,+}^{(i)} & \ddots \\ \vdots & \ddots & \ddots & \ddots & \ddots \end{pmatrix}. \tag{35}$$

3.1.3 P_{N} Equations

Following (22) we replace ψ in the equations of radiation transport (2) by an expansion in spherical harmonics up to order N

$$\psi_{P_N}(t,x,\Omega) = u^T(t,x)\Upsilon_{\mathrm{N}}(\Omega),$$

test the equation (2) by multiplication with spherical harmonics up to order N, and integrate over the unit sphere

$$\langle \partial_t(u^T\Upsilon_{\mathrm{N}}), \Upsilon_{\mathrm{N}} \rangle + \langle \Omega \cdot \nabla_x(u^T\Upsilon_{\mathrm{N}}), \Upsilon_{\mathrm{N}} \rangle = \langle \mathcal{Q}_{RT}(u^T\Upsilon_{\mathrm{N}}), \Upsilon_{\mathrm{N}} \rangle. \tag{36}$$

Equation (36) is equivalent to

$$\underbrace{\langle \Upsilon_{\mathrm{N}}, \Upsilon_{\mathrm{N}}^T \rangle}_{=I_N} \partial_t u + \sum_{i=1}^{3} \underbrace{\langle \Omega_i \Upsilon_{\mathrm{N}}, \Upsilon_{\mathrm{N}}^T \rangle}_{=:A_N^{(i)}} \partial_{x_i} u = \underbrace{\langle \mathcal{Q}_{RT}\Upsilon_{\mathrm{N}}, \Upsilon_{\mathrm{N}}^T \rangle}_{=:Q_N^{RT}} u, \tag{37}$$

with $(\mathcal{Q}_{RT}\Upsilon_{\mathrm{N}})_i = \mathcal{Q}_{RT}(\Upsilon_{\mathrm{N}})_i$. Using the orthonormality and recursion relation of spherical harmonics, as indicated by underbraces, we can confirm that P_{N} equations can be written in the abstract form of eq. (7).

   Rotational invariance of the P_{N} equations follows from the rotational invariance of spherical harmonics, eq. (30). It is obvious that the scattering matrix Q_N and the advection matrices $A_N^{(i)}$ do not depend on the orientation of the coordinate system. Now let $\mathcal{R} \in \mathbb{R}^3$ be the matrix rotating a cartesian coordinate system and $D_N(\mathcal{R})$ the respective rotation matrix as defined in eq. (31). It is easy to verify that

$$Q_N D_N(\mathcal{R}) = D_N(\mathcal{R})Q_N \tag{38}$$

using the invariance of integrals over the unit sphere under rotation, orthogonality of the rotation matrices $\mathcal{R}$ and $D_N(\mathcal{R})$ and that scattering cross sections depend on the pre- and post-collision direction solely via the deflection angle. Likewise the transport operator commutes with the rotation $D_N(\mathcal{R})$

$$\sum_{i=1}^{d} A_N^{(i)} \partial_{\tilde{x}_i}(D(\mathcal{R}) \cdot) = D(\mathcal{R}) \sum_{i=1}^{d} A_N^{(i)} \partial_{x_i}(\cdot), \tag{39}$$

which can be verified using the same arguments and $\partial_{\tilde{x}_i}(\cdot) = \sum_{s=1}^{d} \mathcal{R}_{is}\partial_{x_s}(\cdot)$.

   Block Structure Symmetry / Onsager compatibility of the P_{N} equations immediately follows from the structure of the recursion relations eqs. (32a) and (32b) regarding the even/odd classification of spherical harmonics as in eq. (25). As indicated by eq. (34) it is clear that any permutation gathering first odd and then even components of Υ_{N} allows to transform $A_N^{(i)}$ into the form of (12). The full row rank requirement is satisfied, as the elements of $\Omega_i\Upsilon_{\mathrm{N}}^{\mathrm{o,i}}(\Omega)$ are linearly independent.

3.2 Boundary Conditions

While we know how to formulate boundary conditions for the linear kinetic equation, the right formulation of boundary conditions for P_{N} equations is not clear. We cannot expect $\psi_{P_{\mathrm{N}}}$ to fulfill the original kinetic boundary conditions eq. (4) for general ψ_{in}, that is $\psi_{P_{\mathrm{N}}}$ can not be of the form

$$\psi_{P_{\mathrm{N}}}(\Omega) = \begin{cases} \psi_{P_{\mathrm{N}}}(\Omega) & n \cdot \Omega \geq 0 \\ \psi_{in}(\Omega) & n \cdot \Omega < 0, \end{cases} \tag{40}$$

at the boundary, because the right hand side is discontinuous in general, while ψ_{P_N} is continuous. From theorem 6 we know that the boundary conditions have to determine $N(N+1)/2$ incoming characteristic waves and hence, the right number of independent boundary conditions is $N(N+1)/2$.

Two different strategies to transfer the kinetic boundary condition (4) into boundary conditions for P_N equations are being used in the literature: Enforcing that the P_N approximation ψ_{P_N} matches ψ_{in} on a set of inward-pointing directions Ω_i ($\Omega_i \cdot n < 0$)

$$\psi_{P_N}(\Omega_i) = \psi_{\text{in}}(\Omega_i) \qquad i = 1, \ldots, N(N+1)/2,$$

and secondly, equating ingoing half moments of the known distribution ψ_{in} and the P_N approximation ψ_{P_N} for a certain set of test function ϕ_i

$$\int\limits_{\Omega \cdot n < 0} \psi_{P_N} \, \phi_i \, d\Omega = \int\limits_{\Omega \cdot n < 0} \psi_{\text{in}} \, \phi_i \, d\Omega \qquad i = 1, \ldots, N(N+1)/2.$$

The two types of boundary conditions are referred to as Mark[11, 12] and Marshak[13] type boundary conditions. In this work we will consider only boundary conditions of Marshak type with test functions $\phi_i \in \text{span}\{Y_l^k : |k| \leq l \leq N\}$. Rather than the strategy, it is the set of test functions or directions which is crucial, as the set determines the boundary conditions for the expansion coefficients.

From our previous observations regarding well-posedness, it is clear that the choice of test functions is crucial from a mathematical perspective. However, it is important to not forget, that the primary objective of P_N equations and the accompanying boundary conditions is to serve as an accurate physical model and hence, the choice of test functions should also be reasonable from a modeling point of view. This is however difficult to check, as, apart from the macroscopic density u_0^0 and fluxes u_1^i, we lack a macroscopic interpretation to a general harmonic moment u_l^k. We briefly discuss classical choices of test functions.

3.2.1 Classical Marshak Boundary Conditions

The common choice of test functions, which results in the right number of boundary conditions, is to test with the spherical harmonics of even (odd) parity, if the approximation order N is odd (even). These choices are attractive for arbitrarily oriented boundary normal n, as the parity is invariant under rotation and the test space thereby independent of n. This approach of choosing the test functions is nevertheless questionable. For one, completely exchanging the test space when in- or decreasing the order N by one is somehow unnatural, considering that P_N equations follow from a truncated series expansion and therefore from a hierarchical family of models. And secondly, questions arise from the viewpoint of a physical model. The total incoming particle flux is probably the most fundamental quantity we want to capture exactly within a boundary model. This is not assured, when testing with a finite number of even parity harmonics. When testing with spherical harmonics of odd parity, the correct total incoming particle flux is enforced indirectly by the test functions Y_1^{-1}, Y_1^0 and Y_1^1. In [14] this observation for the case of vacuum boundary condition is used to justify testing with odd parity spherical harmonics, i.e. $\phi_i \in \{Y_l^k : l \text{ odd}\}$. This is not a valid argument for choosing all spherical harmonics of odd parity, as the correct total incoming particle flux can be imposed by a single equation (choose $\phi_i = \Omega \cdot n$).

To our knowledge, the issue of stability is taken into account only in the works [15] and [16], where stable variational formulations for P_N equations are derived. These works however do not provide an explicit form of boundary conditions for the strong formulation of P_N equations and rely on the parity splitting of spherical harmonics.

3.2.2 Onsager Boundary Conditions

In the following we present a modification of Marshak type boundary conditions, which can be seen as a truncation of the recursion relation, consistent with the truncation of the series expansion, and show that the modification leads to Onsager boundary condition. We first motivate our choice of test functions in the Marshak conditions and translate the integral formulation into explicit conditions for the expansion coefficients.

We consider boundary conditions that are based on enforcing the original Boltzmann boundary condition (40) weakly on the subspace of functions that are odd with respect to the boundary normal n, which we write as

$$\langle \psi_{P_N}, \Upsilon_N^{o,n} \rangle = \langle \psi_{P_N}, \Upsilon_N^{o,n} \rangle_{n_+} + \langle \psi_{in}, \Upsilon_N^{o,n} \rangle_{n_-}. \tag{41}$$

Note that Equation (41) is equivalent to the classical formulation of Marshak type boundary conditions, as the case $n \cdot \Omega \geq 0$ in eq. (40) is superfluous and hence the integral equation is governed by integrals over the incoming half sphere. This test space is reasonable from various points of views, and compared to the previously described classical choice more natural:

- the number of test functions is correct for all approximation orders N, in particular independent on whether N is even or odd,

- the boundary conditions are hierarchical in N, that is, they form a cascade analogously to the hierarchy of P_N-equations.

- in the limit $N \to \infty$ the test space forms a complete set on the ingoing half sphere $\{\Omega \in \mathbb{S}^2 : \Omega \cdot n < 0\}$, hence the boundary condition is satisfied weakly in the limit,

- the correct total incoming particle flux, as most fundamental model property, is incorporated explicitly by the first equation and

- they allow to obtain Onsager boundary conditions and hence the desired consistency with the P_N equations and an energy bound for solutions.

First, we rewrite eq. (41) as an equation for the expansion coefficients by splitting ψ_{P_N} into even and odd part $\psi_{P_N} = \Upsilon_N^{o,n\,T} u^o + \Upsilon_N^{e,n\,T} u^e$, using the orthogonality property and replacing half sphere integrals of even functions by integrals over the full sphere

$$\langle \Upsilon_N^{o,n\,T} u^o + \Upsilon_N^{e,n\,T} u^e, \Upsilon_N^{o,n} \rangle = \langle \Upsilon_N^{o,n\,T} u^o + \Upsilon_N^{e,n\,T} u^e, \Upsilon_N^{o,n} \rangle_{n_+} + \langle \psi_{in}, \Upsilon_N^{o,n} \rangle_{n_-} \tag{42}$$

$$\Leftrightarrow \quad \langle \Upsilon_N^{o,n}, \Upsilon_N^{o,n\,T} \rangle u^o = \frac{1}{2} \langle \Upsilon_N^{o,n}, \Upsilon_N^{o,n\,T} \rangle u^o + \langle \Upsilon_N^{o,n}, \Upsilon_N^{e,n\,T} \rangle_{n_+} u^e + \langle \psi_{in}, \Upsilon_N^{o,n} \rangle_{n_-} \tag{43}$$

$$\Leftrightarrow \quad \underbrace{\langle \Upsilon_N^{o,n}, \Upsilon_N^{o,n\,T} \rangle}_{I} u^o = \underbrace{2\langle \Upsilon_N^{o,n}, \Upsilon_N^{e,n\,T} \rangle_{n_+}}_{=:\tilde{M}} u^e + \underbrace{2\langle \psi_{in}, \Upsilon_N^{o,n} \rangle_{n_-}}_{=:g} \tag{44}$$

$$\Leftrightarrow \quad u^o = \tilde{M} u^e + g. \tag{45}$$

We will point out a specific structure in the matrix $\tilde{M}$, which allows us to apply a natural modification to $\tilde{M}$ leading to Onsager boundary conditions. To keep the notation simple we consider the case $n = e_i$, i.e. the boundary normal n coincides with the cartesian basis vectors e_i. The general case follows from rotation. We first insert an artificial $1 = \frac{\Omega \cdot n}{\Omega \cdot n} = \frac{\Omega_i}{\Omega_i}$ for fixed i of which we assign the nominator to the odd part and the denominator to the even

$$\tilde{M} = 2 \int\limits_{\Omega_i \geq 0} \Upsilon_N^{o,i}(\Omega) \left(\Upsilon_N^{e,i}(\Omega) \right)^T d\Omega = 2 \int\limits_{\Omega_i \geq 0} \frac{1}{\Omega_i} \Upsilon_N^{o,i}(\Omega) \left(\Omega_i \Upsilon_N^{e,i}(\Omega) \right)^T d\Omega. \tag{46}$$

Note that the term $\frac{1}{\Omega_i}\Upsilon_N^{o,i}(\Omega)$ is finite for $\Omega_i \to 0$ as $\Upsilon_N^{o,i}(\Omega)$ is at least first order in Ω_i. The term $\Omega_i\Upsilon_N^{e,i}(\Omega)$ is odd and can be written in terms of $\Upsilon_{N+1}^{o,i}$ using the recursion relation (32b)

$$\Omega_i\Upsilon_N^{e,i}(\Omega) = \langle\Omega_i\Upsilon_N^{e,i}(\Omega), \left(\Upsilon_{N+1}^{o,i}(\Omega)\right)^T\rangle\Upsilon_{N+1}^{o,i}(\Omega) = \left(\hat{A}^{(i)}\right)^T\Upsilon_N^{o,i}(\Omega) + R_{l=N+1}\bar{\Upsilon}_{l=N+1}^{o,i}(\Omega) \qquad (47)$$

The last term describes the contribution of higher order moments ($l > N$) in the recursion relation. Similar to the truncation of the series expansion, we truncate the recursion relation by neglecting the contribution of higher order moments to obtain an approximation of the recursion relation involving only terms up to order N

$$\Omega_i\Upsilon_N^{e,i}(\Omega) \approx \left(\hat{A}^{(i)}\right)^T\Upsilon_N^{o,i}(\Omega). \qquad (48)$$

Inserting this approximation into eq. (46) we obtain

$$\tilde{M} \approx M = 2\int_{\Omega_i \geq 0} \frac{1}{\Omega_i}\Upsilon_N^{o,i}(\Omega)\left(\left(\hat{A}_N^{(i)}\right)^T\Upsilon_N^{o,i}(\Omega)\right)^T d\Omega = 2\underbrace{\int_{\Omega_i \geq 0} \frac{1}{\Omega_i}\Upsilon_N^{o,i}(\Omega)\left(\Upsilon_N^{o,i}(\Omega)\right)^T d\Omega}_{=:L_N^{(i)}} \hat{A}_N^{(i)}, \qquad (49)$$

where we observe that the resulting matrix has the form of Onsager boundary condition (15), i.e. the matrix is written as a symmetric positive-definite matrix L_N multiplied by the flux matrix $\hat{A}_N^{(i)}$ and hence we obtain from (45) the desired form of boundary conditions

$$u^{o,i} = L_N^{(i)}\hat{A}_N^{(i)}u^{e,i} + g. \qquad (50)$$

Note that $\tilde{M}$ in eq. (45) differs from M only in the columns that correspond to even expansion coefficients of highest order ($l = N$). The effect of replacing $\tilde{M}$ by M is therefore minor whenever the highest order moments are small, i.e. when the P_N approximation, being based on a truncated series expansion, is a reasonable approximation in the first place. Note also, that $L_N^{(i)}$ is invertible and hence the condition $g(x) \in \text{Im}\left(L_N^{(i)}\right)$ is always satisfied. Therefore our boundary condition ensures the energy bound (17) for general boundary data.

4 Stable Discretization of P_N Equation

In this section we describe the summation by parts (SBP)[17, 18, 19, 20] finite difference discretization (FD) in space for P_N equations with simultaneous approximation terms (SAT)[21] for Onsager boundary conditions. Similar to the discretization of two dimensional P_N equations with periodic or constant extrapolation boundaries conditions presented in [22], we exploit the coupling structure to discretize moments on staggered grids. The discretization allows to assure an discrete analogue to the energy bound in theorem 5 for general Onsager boundary conditions. We first describe the discretization in space using the summation by parts (SBP) finite difference method on staggered grids, then describe the inclusion of the boundary condition using simultaneous approximation terms (SAT) in a manner that ensures stability of the semi-discrete system. We will neglect the relaxation term as its discretization is straight forward and, analogous to the continuous system, the conditions assuring energy stability of the semi-discrete system are independent of the relaxation term.
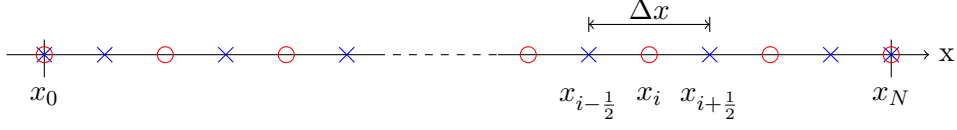
Figure 1: Illustration of one-dimensional staggered grid: Red circles accommodate even variables (even grid) and blue crosses accommodate odd variabes (odd grid). At the boundary nodes both variables are present.

4.1 Energy Stable Discretization in 1D

We first describe the discretization in space and implementation of boundary conditions for the following one-dimensional problem on the an interval $I = (x_L, x_R) \subset \mathbb{R}$.

Problem 8. *Find $u^o : \mathbb{R}_+ \times I \to \mathbb{R}^{n_o}$ and $u^e : \mathbb{R}_+ \times I \to \mathbb{R}^{n_e}$ such that*

$$\partial_t \begin{pmatrix} u^o \\ u^e \end{pmatrix} + \begin{pmatrix} 0 & \hat{A} \\ \hat{A}^T & 0 \end{pmatrix} \partial_x \begin{pmatrix} u^o \\ u^e \end{pmatrix} = 0 \qquad \forall x \in I, \ \forall t > 0 \tag{51}$$

with Onsager boundary conditions

$$u^o = -L\hat{A}\,u^e + g \qquad x = x_L \tag{52}$$

$$u^o = L\hat{A}\,u^e + g \qquad x = x_R \tag{53}$$

and a compatible initial condition $u^o(0, \cdot) = u_0^o, \ u^e(0, \cdot) = u_0^e$.

We discretize the interval I by two grids, which are staggered to one another in the interior, as shown in Figure 1. We refer to the two grids as even and odd grid; identify the vectors of odd and even grid nodes by $x^o = (x_0, x_1, \ldots, x_N)^T \in \mathbb{R}^{N+1}$ and $x^e = (x_0, x_{\frac{1}{2}}, x_{1+\frac{1}{2}}, \ldots, x_{N-\frac{1}{2}}, x_N)^T \in \mathbb{R}^{N+2}$; and denote the sets of grid nodes by $I_h^o \subset I$ and $I_h^e \subset I$. We use $f_{h_\star}$ to denote the grid function approximating the restricting of $f : I \to \mathbb{R}^n$ to $I_h^\star$, identify function values by subscripts $f_{h_\star,i} := f_{h_\star}(x_i)$ and abuse $f_{h_\star}$ to also denote the vector obtained by concatenating the function values of all nodes, e.g.

$$f_{h_\star} := \left(f_{h_\star,0}^o{}^T, \ldots, f_{h_\star,N}{}^T \right)^T. \tag{54}$$

The common notation

$$\langle f_h, g_h \rangle = \sum_i (f_{h,i})^T g_{h,i} \tag{55}$$

is used for the scalar product.

For the sake of a compact notation we introduce notation for operations on grid functions that can be interpreted as the product of a matrix with the vector of node values: Given two grids consisting of the nodes I_h^1 and I_h^2, with $|I_h^1| = N_1 + 1$ and $|I_h^2| = N_2 + 1$, and a matrix $M \in \mathbb{R}^{(N_2+1) \times (N_1+1)}$ we use $\mathcal{M}$ to denote the mapping of grid functions on I_h^1 to grid functions on I_h^2 given by

$$(\mathcal{M}f_{h_1})_i := \sum_{j=0}^{N_1} M_{ij} f_{h_1,j} \qquad i = 0, \ldots, N_2 \tag{56}$$

Furthermore we use $\mathcal{MN}$, $\mathcal{M} \pm \mathcal{N}$, $\mathcal{M}^{-1}$ and $\mathcal{M}^T$ to denote the mappings defined by the matrices MN, $M + N$, M^{-1} and M^T whenever these are well defined.

14

The finite difference (FD) operators on staggered grids, which we consider here, have the character of a projection, in the sense that they approximates the first order spatial derivative of an odd/even grid function on the even/odd grid. In particular, we consider the special class of summation by parts (SBP) FD operators, which are constructed such that integration by parts can be imitated on the discrete level [23].

Definition 9 (Staggered SBP-FD)**.** We call $\mathcal{D}^{e \to o}$ and $\mathcal{D}^{o \to e}$ SBP FD operators on the staggered grids x^o and x^e, if they can be identified by matrices of the form

$$D^o := D^{e \to o} = (P^o)^{-1} Q^o, \qquad D^e := D^{o \to e} = (P^e)^{-1} Q^e,$$

with $P^o \in \mathbb{R}^{(N+1) \times (N+1)}$, $P^e \in \mathbb{R}^{(N+2) \times (N+2)}$, $Q^o \in \mathbb{R}^{(N+1) \times (N+2)}$ and $Q^o \in \mathbb{R}^{(N+2) \times (N+1)}$, such that

(i) the order of *accuracy* is $s \in \mathbb{N}$ for boundary nodes and $2s$ for interior nodes, e.g.

$$(\mathcal{D}^{e \to o}(x^n)_{h_e})_i = (nx^{n-1})_{h_o,i} \qquad \begin{matrix} n = 0, \dots, s & \text{if } i \text{ is boundary node} \\ n = 0, \dots, 2s & \text{if } i \text{ is internal node,} \end{matrix}$$

(ii) the so-called *SBP-property*

$$Q^o + (Q^e)^T = B^o = (B^e)^T = B \qquad \text{with} \qquad \langle f_{h_o}, \mathcal{B}g_{h_e} \rangle = f_{h_o,N}{}^T g_{h_e,N} - f_{h_o,0}{}^T g_{h_e,0}$$

is fullfilled and

(iii) the matrices P^o and P^e are diagonal positive definite and thereby introduce *discrete norms*

$$\|f_{h_o}\|_{h_o}^2 := \langle f_{h_o}, \mathcal{P}^o f_{h_o} \rangle, \qquad \|f_{h_e}\|_{h_e}^2 := \langle f_{h_e}, \mathcal{P}^e f_{h_e} \rangle.$$

For a detailed description of the construction of the matrices P^o, P^e, Q^o and Q^e we refer the reader to [23].

We now discretize system (51) in space by restricting u^o to I_h^o and u^e to I_h^e and replacing the spatial derivatives of $(\hat{A}u^e)$ and $(\hat{A}^T u^o)$ by SBP-FD operators

$$\begin{aligned} \partial_t u_{h_o}^o + \mathcal{D}^{e \to o}(\hat{A}u^e)_{h_e} &= 0 \\ \partial_t u_{h_e}^e + \mathcal{D}^{o \to e}(\hat{A}^T u^o)_{h_o} &= 0. \end{aligned} \tag{57}$$

SBP-FD operators allow to mimic integration by parts of the continuous system, as

$$\frac{1}{2} \partial_t \|u_{h_o}^o\|_{h_o}^2 = -\langle \mathcal{D}^{e \to o}(\hat{A}u^e)_{h_e}, \mathcal{P}^o u_{h_o}^o \rangle = -\langle (\mathcal{P}^o)^{-1} \mathcal{Q}^o(\hat{A}u^e)_{h_e}, \mathcal{P}^o u_{h_o}^o \rangle = -\langle \mathcal{Q}^o(\hat{A}u^e)_{h_e}, u_{h_o}^o \rangle,$$

analogously $\frac{1}{2} \partial_t \|u_{h_e}^e\|_{h_e}^2 = -\langle \mathcal{Q}^e(\hat{A}^T u^o)_{h_o}, u_{h_e}^e \rangle$ and with the SBP-property

$$\begin{aligned} \frac{1}{2} \partial_t \left(\|u_h^o\|_{h_o}^2 + \|u_h^e\|_{h_e}^2 \right) &= -\langle \left(\mathcal{Q}^{o\,T} + \mathcal{Q}^e \right) u_{h_o}^o, (\hat{A}u^e)_{h_e} \rangle \\ &= \left(u_{h_o,0}^o \right)^T \hat{A} u_{h_e,0}^e - \left(u_{h_o,N}^o \right)^T \hat{A} u_{h_e,N}^e. \end{aligned} \tag{58}$$

The SAT method consists of augmenting the equations of the boundary nodes by a penalty term that imposes the boundary condition in a weak sense ($a \in \{o, e\}$)

$$\partial_t u_{h_a,0}^a + \dots = (P_{0,0}^a)^{-1} \tau_0^a \left(u_{h_o,0}^o - \left(-L\hat{A}u_{h_e,0}^e + g_{h,0} \right) \right) \tag{59}$$

$$\partial_t u_{h_a,N}^a + \dots = (P_{N,N}^a)^{-1} \tau_N^a \left(u_{h_o,N}^o - \left(L\hat{A}u_{h_e,N}^e + g_{h,N} \right) \right), \tag{60}$$

with penalization matrices $\tau_0^o, \tau_N^o \in \mathbb{R}^{n_o \times n_0}$ and $\tau_0^e, \tau_N^e \in \mathbb{R}^{n_e \times n_o}$. With theses terms eq. (58) becomes

$$
\begin{aligned}
\frac{1}{2}\partial_t \left(\|u_{h_o}^o\|_{h_o}^2 + \|u_{h_e}^e\|_{h_e}^2 \right) =& \begin{pmatrix} u_{h_o,0}^o \\ u_{h_e,0}^e \end{pmatrix}^T \begin{pmatrix} \tau_0^o & (\tau_0^e)^T + \hat{A} + \tau_0^o L \hat{A} \\ 0 & \tau_0^e L \hat{A} \end{pmatrix} \begin{pmatrix} u_{h_o,0}^o \\ u_{h_e,0}^e \end{pmatrix} - \begin{pmatrix} u_{h_o,0}^o \\ u_{h_e,0}^e \end{pmatrix}^T \begin{pmatrix} \tau_0^o \\ \tau_0^e \end{pmatrix} g_{h,0} \\
&+ \begin{pmatrix} u_{h_o,N}^o \\ u_{h_e,N}^e \end{pmatrix}^T \begin{pmatrix} \tau_N^o & (\tau_N^e)^T - \hat{A} - \tau_N^o L \hat{A} \\ 0 & -\tau_N^e L \hat{A} \end{pmatrix} \begin{pmatrix} u_{h_o,N}^o \\ u_{h_e,N}^e \end{pmatrix} - \begin{pmatrix} u_{h_o,N}^o \\ u_{h_e,N}^e \end{pmatrix}^T \begin{pmatrix} \tau_N^o \\ \tau_N^e \end{pmatrix} g_{h,N}.
\end{aligned}
\tag{61}
$$

In order to obtain an energy bound on the semi-discrete level, we need to choose the penalization matrices τ_*^o, τ_*^e appropriately. We first reduce the complexity by relating τ_*^e to τ_*^o such that the bilinear terms in eq. (61) vanish, i.e.

$$
\tau_0^e = - \left(\hat{A} + \tau_0^o L \hat{A} \right)^T \quad \text{and} \quad \tau_N^e = \left(\hat{A} + \tau_N^o L \hat{A} \right)^T.
\tag{62}
$$

With eq. (62) we obtain necessary and sufficient conditions for τ_0^o, τ_N^o to assure an energy bound analogous to the continuous version (5) for the semi-discrete approximation.

Theorem 10 (Discrete Energy Bound (1D)). *A solution to problem 8 discretized in space using SBP-FD and SAT as described above satisfies the energy bound*

$$
\|u_{h_o}^o(T)\|_{h_o}^2 + \|u_{h_e}^e(T)\|_{h_e}^2 \leq \|(u_0^o)_{h_o}\|_{h_o}^2 + \|(u_0^o)_{h_e}\|_{h_e}^2 + C \left(\|g_L\|_{L^2(0,T)}^2 + \|g_R\|_{L^2(0,T)}^2 \right)
\tag{63}
$$

with $C = \max \bigcup\limits_{\star \in \{0,N\}} \{\|\tau_\star^o\|_2, \|L^+ + \tau_\star^{oT}\|_2\}$,
iff the penalization matrices $\tau_\star^o$ ($\star \in \{0,N\}$) are semi-negative definite and $\forall x \in \mathbb{R}^{n_o}$

$$
x^T L \left(-\tau_\star^o\right)^T L x \leq x^T L x.
\tag{64}
$$

Proof. Assume the energy bound (63) holds. For the energy bound to hold for all $u_{h_o,\star}^o$ and $u_{h_e,\star}^e$ at all times T for the particular case of vanishing boundary source ($g = 0$), we can deduce from eq. (61) that the matrices $\tau_\star^o$, $(\tau_0^e L \hat{A}) = -((\hat{A} + \tau_0^o L \hat{A})^T L \hat{A}$ and $-(\tau_N^e L \hat{A}) = -(\hat{A} + \tau_N^o L \hat{A})^T L \hat{A}$ have to be semi-negative definite. Condition (64) now follows from

$$
(\hat{A} + \tau_\star^o L \hat{A})^T L \hat{A} = \hat{A}^T \left(L + L (\tau_\star^o)^T L \right) \hat{A} \overset{L \text{ inv.}}{=} \hat{A}^T L \left(L^{-1} + (\tau_\star^o)^T \right) L \hat{A},
\tag{65}
$$

rank $\hat{A} = n_o$ and L semi-positive definite.
Now assume that $\tau_\star^o$ is semi-negative definite and (64) holds. We derive a bound of the right hand side of eq. (61) by bounding terms involving $u_{h_o,\star}^o$ and $u_{h_e,\star}^e$ separately. We will use

$$
x^T M x - 2 x^T M p = (x - p)^T M (x - p) - p^T M p \leq |p^T M p| \leq \|M\|_2 \|p\|_2^2 = \|M\|_2 \, p^T p
\tag{66}
$$

for semi-negative definite $M \in \mathbb{R}^{n \times n}$. With eq. (66) can conclude the following bound for the terms in $u_{h_o,\star}^o$

$$
(u_{h_o,\star}^o)^T \tau_\star^o u_{h_o,\star}^o - (u_{h_o,\star}^o)^T \tau_\star^o g_\star = (u_{h_o,\star}^o)^T \tau_\star^o u_{h_o,\star}^o - 2(u_{h_o,\star}^o)^T \tau_\star^o \frac{g_\star}{2} \leq \frac{1}{4} \|\tau_\star^o\|_2 \, g_\star^T g_\star.
\tag{67}
$$

As $-(L + L\,(\tau_\star^o)^T\,L)$ is semi-negative definite we can bound the terms in $u_{h_e,\star}^e$ similarly

$$(u_{h_e,\star}^e)^T(-(\hat{A}^T L\hat{A} + \hat{A}^T L\,(\tau_\star^o)^T\,L\hat{A}))u_{h_e,\star}^e \pm (u_{h_e,\star}^e)^T(\hat{A}^T + \hat{A}^T L(\tau_\star^o)^T)g_\star \tag{68}$$

$$=(L\hat{A}u_{h_e,\star}^e)^T(-(L^+ + (\tau_\star^o)^T))(L\hat{A}u_{h_e,\star}^e) \pm (L\hat{A}u_{h_e,\star}^e)^T(L^+ + (\tau_\star^o)^T)g_\star \tag{69}$$

$$\leq \frac{1}{4}\|L^+ + \tau_\star^{oT}\|_2\, g_\star^T g_\star. \tag{70}$$

The energy bound now follows from

$$\frac{1}{2}\partial_t\left(\|u_{h_o}^o\|_{h_o}^2 + \|u_{h_e}^e\|_{h_e}^2\right) \leq \sum_{\star\in\{0,N\}}\frac{1}{4}\left(\|\tau_\star^o\|_2 + \|L^+ + \tau_\star^{oT}\|_2\right)g_\star^T g_\star \leq \frac{C}{2}\sum_{\star\in\{0,N\}}g_\star^T g_\star \tag{71}$$

with $C = \max\bigcup_{\star\in\{0,N\}}\{\|\tau_\star^o\|_2, \|L^+ + \tau_\star^{oT}\|_2\}$. $\qquad\qquad\square$

The most apparent choice for the penalization matrix $\tau_\star^o$ to assure the discrete energy bound is $\tau_\star^o = 0$. For the case that L is invertible, the condition (64) simplifies to

$$0 \leq x^T\,(-\tau_\star^o)^T\,x \leq x^T L^{-1}x$$

and an obvious family of penalization matrices assuring energy stability would be $\tau_\star^o = -\alpha L^{-1}$ with $\alpha \in [0,1]$.

4.2 Extension to 3D

Before explaining the extension to the three-dimensional case, we realize that the even/odd classification with respect to the three cartesian coordinate axes divides the spherical harmonic functions into $8\,(=2^3)$ families with two families coupling within the recursion relation of direction i, iff their type is distinct only in direction i. We classify the respective expansion coefficients accordingly. This insight into the coupling allows us to apply the staggered SBP-FD discretization to the three-dimensional P_N equations. We discretize the eight families of expansion coefficients on an rectangular domain $G = [x_{1,L}, x_{1,R}] \times [x_{2,L}, x_{2,R}] \times [x_{3,L}, x_{3,R}] \subset \mathbb{R}^3$ using 8 staggered grids, such that the coupling structure within the equations is respected, see Figure 2. When representing grid functions by vectors of node values with x_3 as the contiguous direction, following x_2 and then the x_1 direction, the matrices identifying the staggered SBP difference operators in three dimensions can be written using the Kronecker product

$$D_1^{\mathbf{a}} = D_{x_1}^{a_1} \otimes I_{x_2}^{a_2} \otimes I_{x_3}^{a_3}, \quad D_2^{\mathbf{a}} = I_{x_1}^{a_1} \otimes D_{x_2}^{a_2} \otimes I_{x_3}^{a_3}, \quad D_3^{\mathbf{a}} = I_{x_1}^{a_1} \otimes I_{x_2}^{a_3} \otimes D_{x_3}^{a_3}. \tag{72}$$

The triplet $\mathbf{a} = (a_1, a_2, a_3)$ with $a_i \in \{o, e\}$ identifies grids $h_{\mathbf{a}}$ and variables $u_{h_{\mathbf{a}}}^{\mathbf{a}}$ with respect to the odd-even classification in three dimensional space and $I_{x_1}^o$, etc. denote identity matrices of the same dimension as the number of grid nodes in the direction indicated by the subscript. The same approach can be used to construct scalar products assigned to the different grids

$$\langle f_{h_{\mathbf{a}}}, g_{h_{\mathbf{a}}}\rangle_{h_{\mathbf{a}}} := \langle f_{h_{\mathbf{a}}}, (\mathcal{P}_{x_1}^{a_1} \otimes \mathcal{P}_{x_2}^{a_2} \otimes \mathcal{P}_{x_3}^{a_3})\,g_{h_{\mathbf{a}}}\rangle \qquad \|f_{h_{\mathbf{a}}}\|_{h_{\mathbf{a}}}^2 = \langle f_{h_{\mathbf{a}}}, f_{h_{\mathbf{a}}}\rangle_{h_{\mathbf{a}}}. \tag{73}$$

and for the boundary grids $\partial h_{\mathbf{a}} = h_{\mathbf{a}} \cap \overline{G}$

$$\langle f_{\partial h_{\mathbf{a}}}, g_{\partial h_{\mathbf{a}}}\rangle_{\partial h_{\mathbf{a}}} = \sum_{i\in\partial h_{\mathbf{a}}} w_i^{\mathbf{a}}\,(f_{\partial h_{\mathbf{a}}})_i^T\,(g_{\partial h_{\mathbf{a}}})_i \qquad \|f_{\partial h_{\mathbf{a}}}\|_{\partial h_{\mathbf{a}}}^2 = \langle f_{\partial h_{\mathbf{a}}}, g_{\partial h_{\mathbf{a}}}\rangle_{\partial h_{\mathbf{a}}} \tag{74}$$
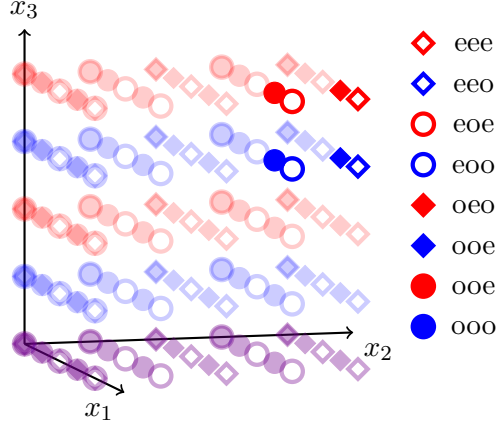
Figure 2: Illustration of staggered grids in three-dimensions. The grid boundary is assumed to be on the surfaces spanned by the axes. Each of the eight symbols stores a certain odd/even class of variables. In total there are eight classes. In the interior there is the following pattern: In x_1-direction filled and open variable classes interchange, in x_2-direction diamonds and circles interchange, in x_3-direction blue and red variable classes interchange. Analogously to the 1d setting in Fig. 1 the respective variable classes are jointly present on the boundary surfaces.

with $w_i^a = \sum_{i=1}^{d} \sum_{\star \in \{0,N_d\}} \delta_{i_d,\star} \prod_{j \neq d} P_{x_j}^{a_j}$. Analogous to the one-dimensional case we discretize in space by replacing the spatial derivatives by the difference operators (72) and augmenting the equations of the boundary nodes with SAT terms. Using $c_d(\mathbf{a})$ to denote the complement of odd-even class $\mathbf{a}$ in direction d (e.g. $c_2(ooo) = oeo$), and $o_d(\mathbf{a})$ and $(e_d(\mathbf{a}))$ the triplet obtained by replacing the d-th component of $\mathbf{a}$ by o and e respectively (e.g $e_3(\mathbf{a}) = (a_1 a_2 e)$) we can write the semi-discrete system by

$$\partial_t u_{h_\mathbf{a},i}^\mathbf{a} + \sum_{d=1}^{3} A_d^\mathbf{a} \left(\mathcal{D}_d^\mathbf{a} u_{h_{c_d(\mathbf{a})}}^{c_d(\mathbf{a})} \right)_i = \sum_{\substack{d=1,\dots,3 \\ \star \in \{0,N_d\}}} \frac{\delta_{i_d,\star}}{(P_{x_d}^{a_d})_{\star,\star}} \tau_\star^{d,\mathbf{a}} \left(u_{h_{o_d(\mathbf{a})},i}^{o_d(\mathbf{a})} - \left(\pm L_d^{e_d(\mathbf{a})} \hat{A}_d^{e_d(\mathbf{a})} u_{h_{e_d(\mathbf{a})},i}^{e_d(\mathbf{a})} + (g_d^a)_i \right) \right),$$

$$(75)$$

where i is a triplet identifying a node of grid $\mathbf{a}$. We used $\pm$ to merge the formulas for $i_d = 0$ with a minus sign and $i_d = N_d$ with a plus sign. Although blurred by the indices and superscripts the discretization (75) is the three dimensional counterpart of eqs. (59) and (60) and we can use the same arguments as used in the proof of theorem 10 to obtain an analogous energy bound, given that

$$\tau_0^{d,e_d(\mathbf{a})} = -\left(\hat{A}_d^{e_d(\mathbf{a})} + \tau_0^{d,o_d(\mathbf{a})} L_d^{e_d(\mathbf{a})} \hat{A}_d^{e_d(\mathbf{a})} \right)^T \quad \text{and} \quad \tau_{N_d}^{d,e_d(\mathbf{a})} = \left(\hat{A}_d^{e_d(\mathbf{a})} + \tau_{N_d}^{d,o_d(\mathbf{a})} L_d^{e_d(\mathbf{a})} \hat{A}_d^{e_d(\mathbf{a})} \right)^T. \quad (76)$$

Theorem 11 (Discrete Energy Bound (3D)). *A solution to problem 8 discretized in space using SBP-FD and SAT as described above satisfies the energy bound*

$$\sum_{\mathbf{a} \in \{e,o\}^3} \|u_{h_\mathbf{a}}^a(T)\|_{h_\mathbf{a}}^2 \leq \sum_{\mathbf{a} \in \{e,o\}^3} \|(u_0^a)_{h_\mathbf{a}}\|_{h_\mathbf{a}}^2 + C\|g_d^a\|_{\partial h_\mathbf{a}(0,T)}^2 \quad (77)$$

with $C = \max \bigcup_{\substack{d=1,\dots,3 \\ \mathbf{a} \in \{e,o\}^3}} \{\|\tau^{d,o_d(\mathbf{a})}\|_2, \|L_d^+ + \left(\tau^{d,o_d(\mathbf{a})}\right)^T\|_2\}$ *and* $\| \cdot \|_{\partial h_\mathbf{a}(0,T)}^2 = \int_0^T \| \cdot \|_{\partial h_\mathbf{a}}^2 dt$,

iff the penalization matrices $\tau^{d,\mathbf{a}}$ are semi-negative definite and $\forall x \in \mathbb{R}^{n_{\mathbf{a}}}$

$$x^T L_d^{e_d(\mathbf{a})} \left(-\tau^{d,\mathbf{a}} \right)^T L_d^{e_d(\mathbf{a})} x \leq x^T L_d^{e_d(\mathbf{a})} x. \tag{78}$$

5   Numerical Results

In this section we present numerical results of four test cases with different objectives. The first test case confirms and visualizes the decay of energy over time for vacuum boundary conditions. The second demonstrates the effect of incoming waves depending on standing waves and the interpretation as artificial boundary sources. Test cases three and four are motivated by electron probe microscopy (EPMA)[24] and aim to address the accuracy of the boundary conditions in an application of engineering relevance. We consider the two boundary scenarios relevant in EPMA: Electrons leaving a solid material into vacuum and beam electrons entering a solid material. For both scenarios we compare different approximation orders N and address the accuracy of P_N solutions through comparison to Monte Carlo (MC) results obtained by taking averages over a set of 10^7 electron trajectories simulated with the software package DTSA-II[25, 26]. It is generally accepted that the odd order P_N approximation is superior to the next even order approximation P_{N+1} and hence, excluding the second test case, we follow the common practice of choosing N odd.[27]

Unless stated otherwise, the presented results were obtained from discretization of P_N equations as explained in section 4 with second order SBP operators based on `github.com/ooreilly/sbp` [23] and with a second order Strang splitting for numerical integration in time.

5.1   Test Case 1: Energy Stability

In the first test case we consider particles moving in vacuum, i.e. particles are in free-flight and do not interact with a background medium. The initial particle distribution is constant in the directions x_2 and x_3, follows a normal distribution in x_1 direction centered at $x_1 = 0$ and is independent of the direction Ω. We restrict our view onto the spatial domain $(-1, 1) \times \mathbb{R} \times \mathbb{R}$ and assume, that the initial particle density is negligible outside, so particles do not enter the domain ($\psi_{in} = 0$). The P_N model for this case reduces to a pseudo one dimensional problem and, identifying x_1 with x, we can write the initial condition for the P_N approximation by

$$u_l^k(0, x) = \begin{cases} e^{-\left(\frac{x - \mu_x}{\sqrt{2}\sigma_x} \right)^2} & l = k = 0 \\ 0 & \text{else.} \end{cases}$$

Figure 3a shows the particle density u_0^0 at different times t^* for the initial distribution with $\mu_x = 0$ and $\sigma_x = 0.2$ obtained with approximation order $N = 13$. The density profile smooths out symmetrically towards both boundaries and approaches zero, which coincides with the picture of the underlying particle system, where particles leave the domain and no new particles enter over the boundary. The solution also illustrates two artifacts of the P_N method: The density is not assured to be positive and the set of velocities at which information is transported is finite.

In Figure 3b the discrete SBP energy norm of the solution is plotted over time. The graph confirms the decrease of the energy in time for vacuum boundary conditions. As the set of transport velocities is finite, the initial data centered around $x = 0$ reach the boundaries in a package manner, which leads to the terraced shape of the graph.

(a) Snapshots of a particle density u_0^0.
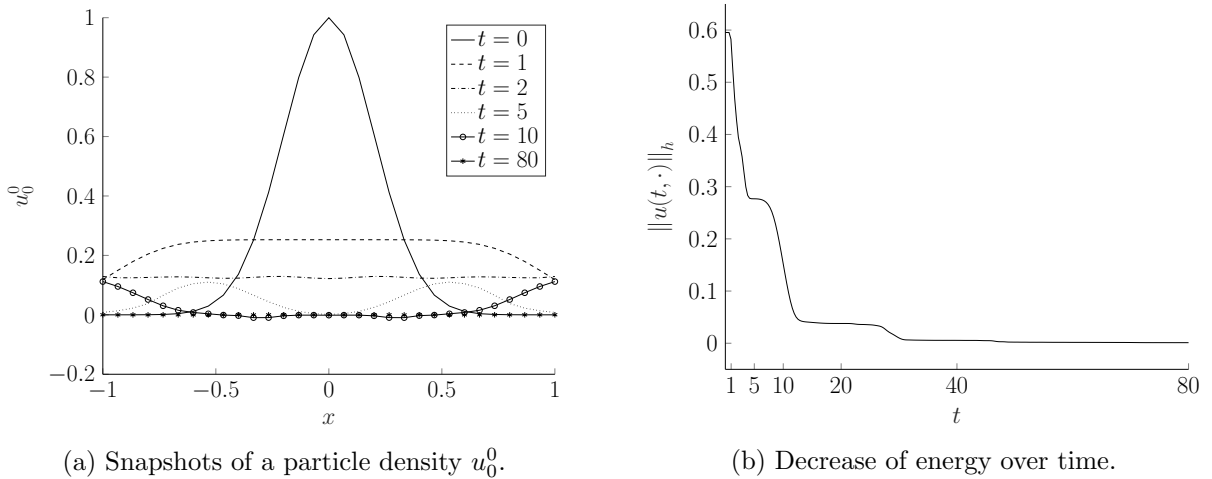
(b) Decrease of energy over time.

Figure 3: Results obtained with P_{13} for an initial bulk of particles with constant distribution in direction space in the absence of collisions and physical boundaries.

5.2 Test Case 2: Artificial Boundary Source

The purpose of this case is to demonstrate the necessity of stable boundary conditions, to assure a strict energy bound, i.e. the energy does not increase in the absence of boundary sources ($g = 0$). As for the first test case, we neglect scattering, consider pseudo one dimensional P_{N} equations ($N = 2$) on the x_1 axis and identify x_1, x_2 and x_3 by x, y and z. We choose the domain to be $G = \mathbb{R}_+$, and construct an initial condition that reduces the number of non-trivial coefficients to only four and allows to consider solely the domain close to the boundary at $x = 0$. We set all coefficients u_l^k to zero, except for the four coefficients that are even in both y and z direction, namely $u^e = (u_0^0, u_2^0, u_2^2)^T$ and $u^o = u_1^1$. The pseudo one dimensional P_2 equations are

$$\partial_t \begin{pmatrix} u^o \\ u^e \end{pmatrix} + \begin{pmatrix} 0 & \hat{A} \\ \hat{A}^T & 0 \end{pmatrix} \partial_x \begin{pmatrix} u^o \\ u^e \end{pmatrix} = 0, \qquad \text{with} \quad \hat{A} = \begin{pmatrix} 0.5773\dots & -0.2581\dots & 0.4472\dots \end{pmatrix}.$$

For the demonstration we consider the boundary condition with zero source

$$u^o = \begin{pmatrix} 0.8660\dots & -0.2420\dots & 0.4192\dots \end{pmatrix} u^e + \overset{0}{\cancel{g}}, \tag{79}$$

obtained from the submatrix of $\tilde{M}$ in eq. (45) that couples u^o to u^e. This boundary condition prescribes the incoming waves dependent of standing waves such that no strict stability is given, which can be avoided by neglecting the higher order coefficients of the recursion relation appearing in matrix $\tilde{M}$ (eq. (46)), as described in section 3.2.2. The initial condition is constructed to serve the purpose: We choose u^e and compute u^o from the boundary condition

$$u^e|_{t=0} \overset{!}{=} \begin{pmatrix} 1 \\ 2.5 \\ -1 \end{pmatrix} e^{\left(\frac{x}{0.1}\right)^2} \quad \overset{BC}{\longrightarrow} \quad u^o|_{t=0} = -0.1583\dots\, e^{\left(\frac{x}{0.1}\right)^2}.$$

The red graph in Figure 4 shows the change of the energy over time for the unstable boundary condition (79). The change of the energy results from two opposing processes: The energy-loss caused by the outgoing waves, and the energy-gain caused by incoming waves prescribed by the boundary condition at $x = 0$. While the energy loss is initially dominant, both processes balance out around $t = 0.1$, and

as time increases the outgoing waves vanish and we can observe the constant increase in energy caused by the standing waves acting as artificial sources in the boundary condition.

The black graph in Figure 4 visualizes the energy evolution for the stabilized version of the test case, i.e. replacing boundary condition (79) by the stabilized version as described in section 3.2.2 and adapting the initial condition of u^o. For the stable boundary conditions the energy decays and, as incoming waves are independent of standing waves, approaches steady state when outgoing waves become negligible.
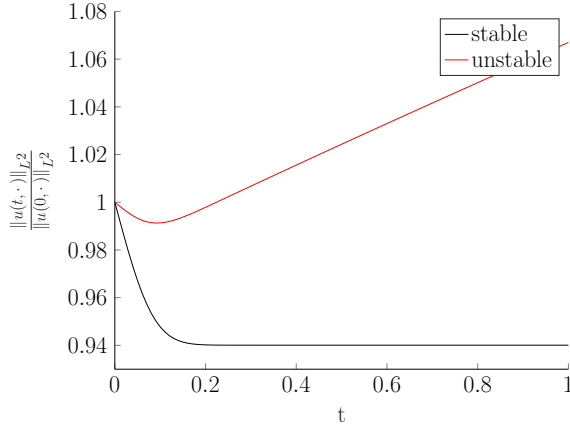
Figure 4: Evolution of energy for test case 2 with stable and unstable boundary conditions; Increase of energy in the absence of boundary source terms caused by standing waves acting as artificial source terms in unstable boundary condition.

5.3 Boundary Conditions for EPMA

We now consider two practically relevant test cases motivated by EPMA, a method to quantify the micro structure of materials[24]. The method is based on exciting a small polished material sample by a focused electron beam within a vacuum chamber. In the context of EPMA the electron transport is usually described in terms of the electron fluence modeled by BCSD eq. (3). The electron fluence density, given by the zeroth order moment u_0^0 in the series expansion, is the quantity of interest for EPMA and will be used for comparison and evaluation. For the sake of demonstration we consider only two dimensions and assume solutions to be constant in the third direction.

In both test cases we consider a sample of pure copper of density 8960 kg/m^3. Our P_N code and the MC code DTSA-II implement different models of stopping power and scattering cross-sections. For comparability of the MC and P_N results we set the stopping power to a constant $S = -2 \cdot 10^{-13}$ (J/m)/(kg/m^3) in both codes. The scattering cross-sections used in the P_N simulations were taken from the database of the ICRU report [28] which was generated using the ELSEPA code[29] and the MC simulations employed scattering cross sections from a NIST database[30] that was generated by an earlier version of the ELSEPA code.

5.3.1 Test Case 3: Vacuum

First we look at the situation of beam electrons leaving a material sample into vacuum. Figure 5 shows a P_{13} result of the electron fluence density u_0^0 for a bulk of electrons inside copper, which is
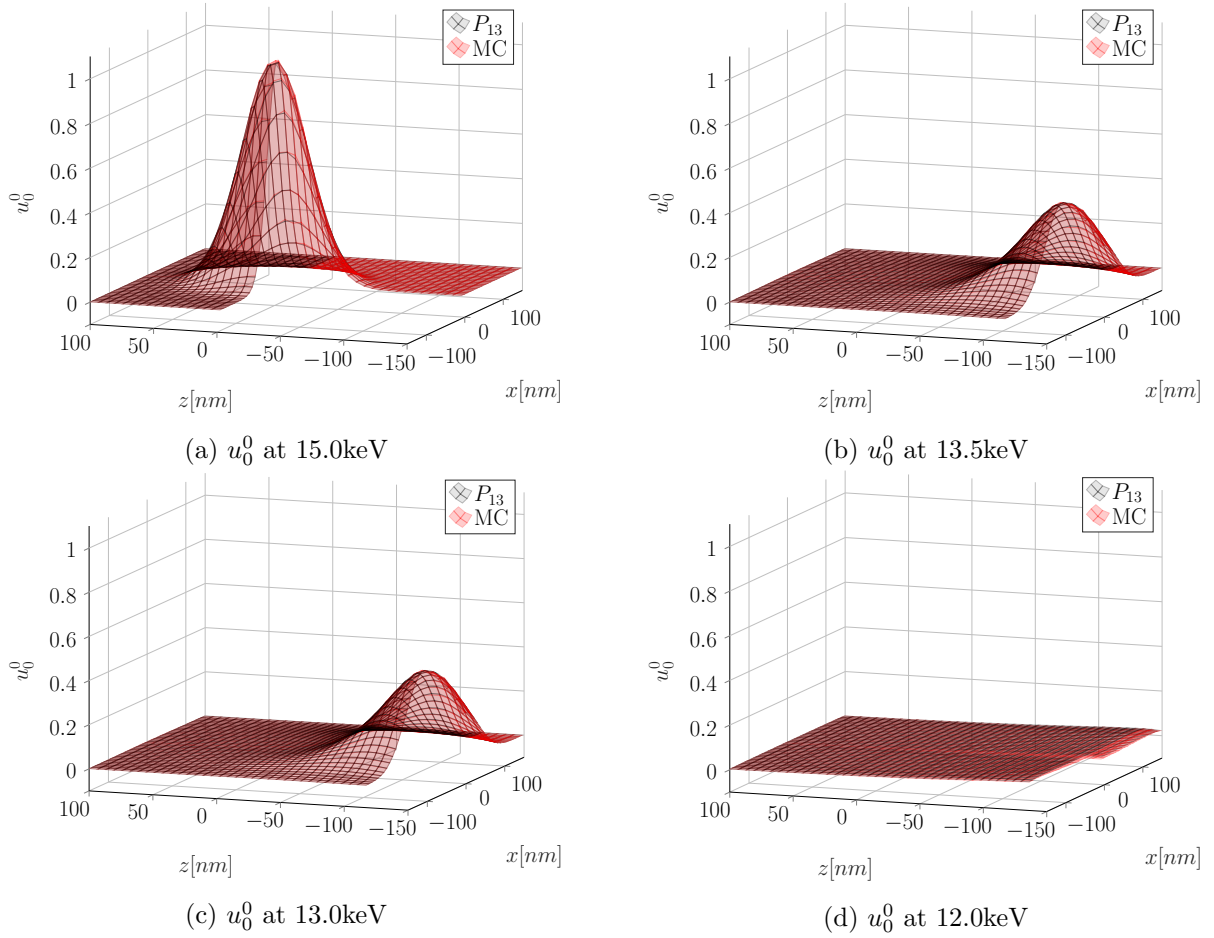
(a) u_0^0 at 15.0keV

(b) u_0^0 at 13.5keV

(c) u_0^0 at 13.0keV

(d) u_0^0 at 12.0keV

Figure 5: Fluence density u_0^0 of a bulk of electrons escaping a material sample of pure copper into vacuum over a polished surface at z=−150nm computed with P_{13} and MC.

initially ($\epsilon = 15$keV) centered around $x = z = 0$nm and moving in negative z direction

$$\psi_{15\text{keV}}(x, z, \Omega) = e^{-\left(\frac{x^2 + z^2}{2\sigma^2}\right)} \delta(\Omega_z + 1) \qquad \text{with } \sigma = 25\text{nm},$$

escaping into vacuum over a polished surface at $z = -150$nm. To assess the quality of the P_{13} result Figure 5 also shows MC results. For all four snapshots we can observe a very good agreement between P_{13} and MC, in fact, it is hard to see deviations between the results of the two methods. The good agreement is more visible in Figure 6a, which shows snapshots of u_0^0 sliced at $x = 0$. Figure 6b shows the same slices obtained with approximation orders $N = 3, 7, 13$. Taking the P_{13} as reference, we observe that the P_7 result is already very accurate whereas the P_3 result deviates clearly. Comparing the deviation at the boundary with the deviations in the interior, it seems that the deviation is caused by the low order rather then a poor physical model at the boundary.

5.3.2 Test Case 4: Gaussian Beam

In this test case we look at the other important boundary scenario in the context of EPMA: Beam electrons entering a material probe normal to the polished material surface. As for the previous test case we assume solutions to be constant in the y ($\hat{=}x_2$) direction and look at a two dimensional scenario
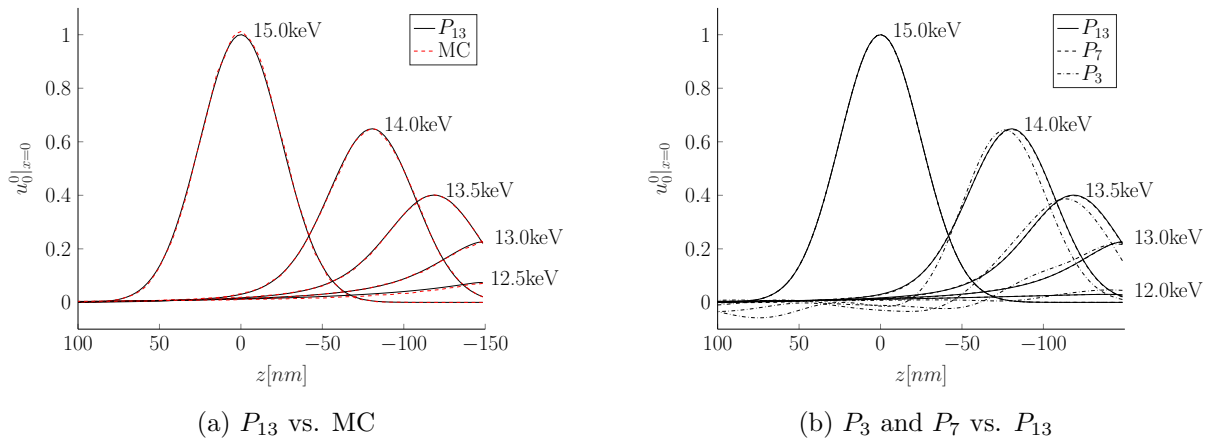
(a) P_{13} vs. MC

(b) P_3 and P_7 vs. P_{13}

Figure 6: Fluence density u_0^0 at x=0nm of a bulk of electrons inside a material sample of pure copper, initially (15keV) moving in negative z-direction, escaping into vacuum over a polished surface at $z = -150$nm computed with P_N and MC.

in x $(\hat{=}x_1)$ and z $(\hat{=}x_3)$ direction.

We consider a 14keV Gaussian electron beam that is aligned with the z-axis and shoots electrons onto a polished surface located at $z = 0$, which we can model via the incoming electron fluence

$$\psi_{in}(\epsilon, x, \Omega) = e^{-\left(\frac{\epsilon - 14\text{keV}}{\sqrt{2}\sigma_\epsilon}\right)^2} e^{-\left(\frac{x}{\sqrt{2}\sigma_x}\right)^2} e^{-\left(\frac{\Omega_z + 1}{\sqrt{2}\sigma_\Omega}\right)^2}.$$

In Figure 7 we compare the electron fluence density u_0^0 obtained with P_{13} to a MC result for $\sigma_\epsilon = \frac{14\text{keV}}{100}$, $\sigma_x = 25$nm and $\sigma_{\Omega_z} = 0.1$. As before, P_{13} reproduces MC very well and no significant deviations between the results can be observed. The good agreement is more apparent in Figure 8a where slices of u_0^0 at different energies can be compared. In Figure 8b we compare different approximation orders N and make the same observation as for the test case on vacuum boundary conditions: The P_7 result shows only minor deviations from the P_{13} result, whereas the P_3 result deviates significantly. Again, it seems that the strong deviation of P_3 from P_{13} is related to the low approximation order rather then introduced by a poor translation of the Gaussian electron beam into boundary conditions for the P_N equations.

6 Acknowledgements

(a) u_0^0 at 14.1keV

(b) u_0^0 at 14.0keV
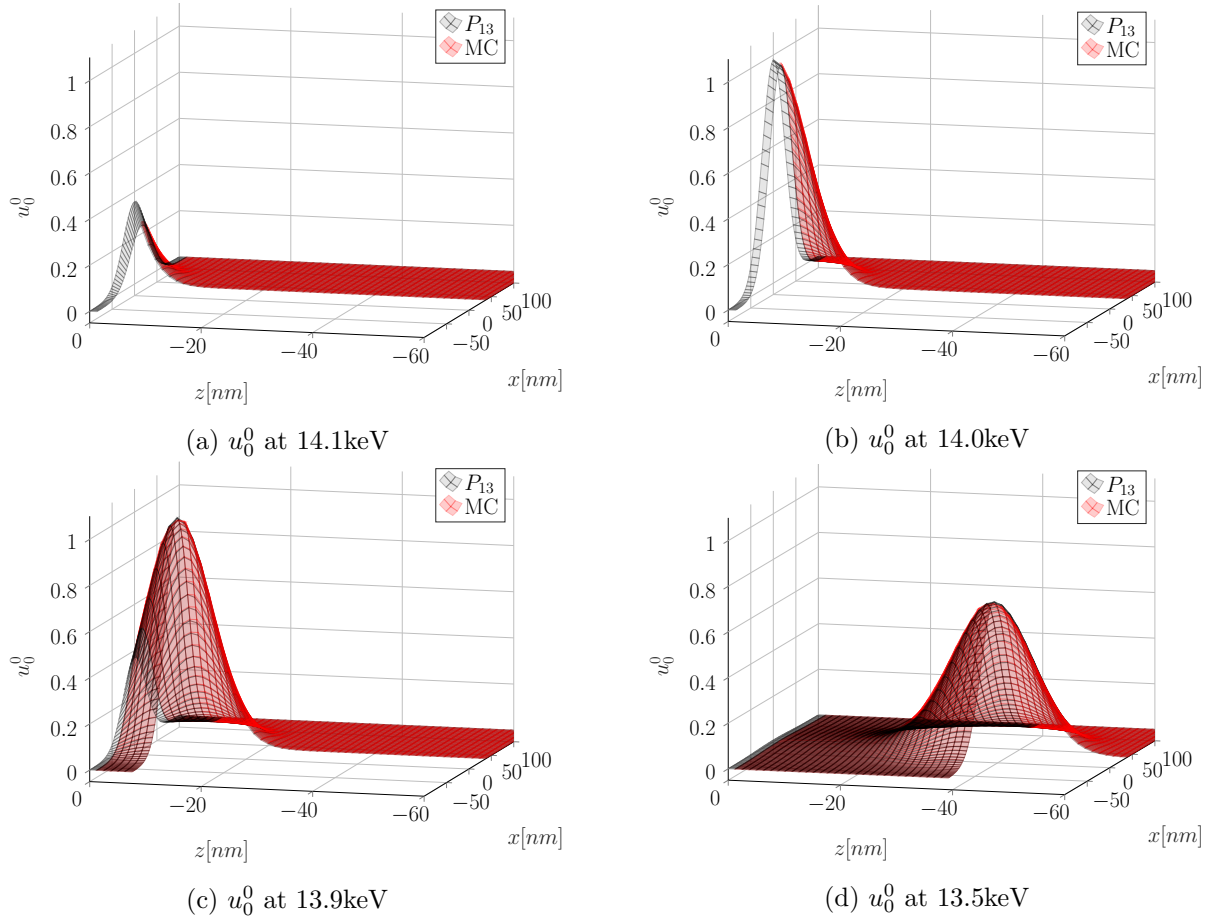
(c) u_0^0 at 13.9keV

(d) u_0^0 at 13.5keV

Figure 7: Fluence density u_0^0 of beam electrons, shot from a 14keV Gaussian electron beam, penetrating a material sample of pure copper over a polished surface at $z=0$nm computed with P_{13} and MC.
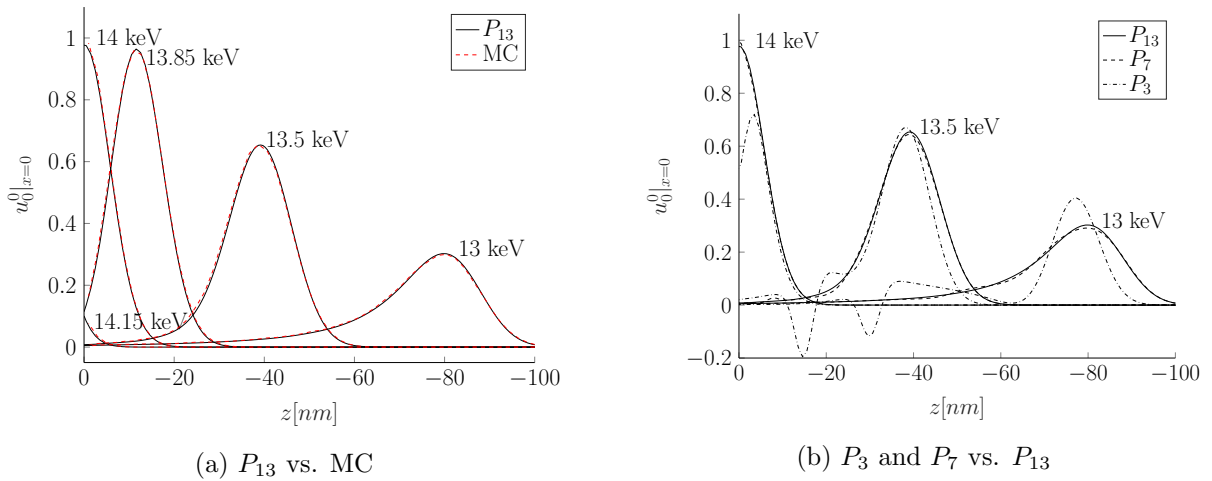

(a) P_{13} vs. MC

(b) P_3 and P_7 vs. P_{13}

Figure 8: Fluence density u_0^0 at $x=0$nm of beam electrons, shot from a 14keV Gaussian electron beam, penetrating a material sample of pure copper over a polished surface at $z=0$nm computed with P_N and MC.

References

[1] C. Cercignani, *The Boltzmann Equation and Its Applications*. Springer-Verlag New York, 1988.

[2] K. M. Case and P. F. Zweifel, *Linear transport theory*. Reading, Mass. Addison-Wesley Pub. Co., 1967.

[3] E. W. Larsen, M. M. Miften, B. A. Fraass, and I. A. D. Bruinvis, "Electron dose calculations using the method of moments," *Medical Physics*, vol. 24, no. 1, pp. 111–125, 1997.

[4] C. Ringhofer, C. Schmeiser, and A. Zwirchmayer, "Moment methods for the semiconductor boltzmann equation in bounded position domains," *SIAM Journal on Numerical Analysis*, vol. 39, 03 2001.

[5] N. Sarna and M. Torrilhon, "On stable wall boundary conditions for the hermite discretization of the linearised boltzmann equation," *Journal of Statistical Physics*, vol. 170, pp. 101–126, Jan 2018.

[6] A. S. Rana and H. Struchtrup, "Thermodynamically admissible boundary conditions for the regularized 13 moment equations," *Physics of Fluids*, vol. 28, no. 2, p. 027105, 2016.

[7] H. Struchtrup and M. Torrilhon, "H theorem, regularization, and boundary conditions for linearized 13 moment equations," *Physical review letters*, vol. 99, p. 014502, 08 2007.

[8] B. Davison, *Neutron transport theory*. International series of monographs on physics, Clarendon Press, 1958.

[9] K. Atkinson and W. Han, *Spherical Harmonics and Approximations on the Unit Sphere: An Introduction*, vol. 2014. 01 2012.

[10] C. Müller, *Spherical harmonics / Claus Müller*. Springer-Verlag Berlin ; New York, 1966.

[11] J. C. Mark, "The spherical harmonics method, part I," Tech. Report MT 92, National Research Council of Canada, 1944.

[12] J. C. Mark, "The spherical harmonics method, part II," Tech. Report MT 97, National Research Council of Canada, 1945.

[13] R. E. Marshak, "Note on the spherical harmonic method as applied to the milne problem for a sphere," *Phys. Rev.*, vol. 71, pp. 443–446, Apr 1947.

[14] M. Williams, *Mathematical methods in particle transport theory*. 01 1971.

[15] H. Egger and M. Schlottbom, "A mixed variational framework for the radiative transfer equation," *Mathematical Models and Methods in Applied Sciences*, vol. 22, 03 2012.

[16] M. Schlottbom and H. Egger, "A class of galerkin schemes for time-dependent radiative transfer," *SIAM Journal on Numerical Analysis*, vol. 54, pp. 3577–3599, 12 2015.

[17] H.-O. KREISS and G. SCHERER, "Finite element and finite difference methods for hyperbolic partial differential equations," in *Mathematical Aspects of Finite Elements in Partial Differential Equations* (C. de Boor, ed.), pp. 195 – 212, Academic Press, 1974.

[18] B. Strand, "Summation by parts for finite difference approximations for d/dx," *Journal of Computational Physics*, vol. 110, no. 1, pp. 47 – 67, 1994.

[19] P. Olsson, "Summation by parts, projections, and stability. i," *Math. Comput.*, vol. 64, pp. 1035–1065, July 1995.

[20] D. C. D. R. Fernndez, J. E. Hicken, and D. W. Zingg, "Review of summation-by-parts operators with simultaneous approximation terms for the numerical solution of partial differential equations," *Computers & Fluids*, vol. 95, pp. 171 – 196, 2014.

[21] M. H. Carpenter, D. Gottlieb, and S. Abarbanel, "Time-stable boundary conditions for finite-difference schemes solving hyperbolic systems: Methodology and application to high-order compact schemes," *Journal of Computational Physics*, vol. 111, no. 2, pp. 220 – 236, 1994.

[22] B. Seibold and M. Frank, "Starmap—a second order staggered grid method for spherical harmonics moment equations of radiative transfer," *ACM Trans. Math. Softw.*, vol. 41, pp. 4:1–4:28, Oct. 2014.

[23] O. O'Reilly, T. Lundquist, E. M. Dunham, and J. Nordstrm, "Energy stable and high-order-accurate finite difference methods on staggered grids," *Journal of Computational Physics*, vol. 346, pp. 572 – 589, 2017.

[24] L. Reimer, *Scanning electron microscopy.* Berlin: Springer, 2nd ed., 1998.

[25] N. W. M. Ritchie, "A new monte carlo application for complex sample geometries," *Surf. Interface Anal.*, vol. 37, no. 11, pp. 1006–1011, 2005.

[26] N. W. M. Ritchie, "Spectrum simulation in DTSA-II," *Microscopy and Microanalysis*, vol. 15, no. 5, pp. 454 – 522, 2009.

[27] J. Howlet, "Neutron transport theory. by b. davison and j. b. sykes. 75s. 1957. (oxford university press)," *The Mathematical Gazette*, vol. 43, no. 345, p. 221222, 1959.

[28] I. C. on Radiation Units and Measurements, *Elastic Scattering of Electrons and Positrons.* ICRU report, Oxford University Press, 2007.

[29] F. Salvat, A. Jablonski, and C. J. Powell, "elsepadirac partial-wave calculation of elastic scattering of electrons and positrons by atoms, positive ions and molecules," *Computer Physics Communications*, vol. 165, no. 2, pp. 157 – 190, 2005.

[30] A. Jablonski, F. Salvat, and C. Powell, *NIST Electron Elastic-Scattering Cross-Section DatabaseVersion 3.1.* National Institute of Standards and Technology, 2003.