An ADER discontinuous Galerkin method on moving meshes for Liouville's equation of geometrical optics

Journal of Computational Physics 488 (2023) 112208

Contents lists available at ScienceDirect

Journal of Computational Physics

journal homepage: www.elsevier.com/locate/jcp

An ADER discontinuous Galerkin method on moving meshes for Liouville's equation of geometrical optics

Robert A.M. van Gestel [a,*], Martijn J.H. Anthonissen [a], Jan H.M. ten Thije Boonkkamp [a], Wilbert L. IJzerman [a,b]

[a] Eindhoven University of Technology, PO Box 513, 5600 MB Eindhoven, the Netherlands
[b] Signify, High Tech Campus 7, 5656 AE, Eindhoven, the Netherlands

ARTICLE INFO

Article history:
Received 19 July 2022
Accepted 1 May 2023
Available online 8 May 2023

Keywords:
Liouville's equation
Geometrical optics
Discontinuous Galerkin
ADER
Arbitrary-Lagrangian Eulerian

ABSTRACT

Liouville's equation describes light propagation through an optical system. It governs the evolution of an energy distribution on phase space. This energy distribution is discontinuous across optical interfaces. Curved optical interfaces manifest themselves as moving boundaries on phase space. In this paper, an ADER discontinuous Galerkin (DG) method on a moving mesh is applied to solve Liouville's equation. In the ADER approach a temporal Taylor series is computed by replacing temporal derivatives with spatial derivatives using the Cauchy-Kovalewski procedure. The result is a fully discrete explicit scheme of arbitrary high order of accuracy. A moving mesh is not sufficient to be able to solve Liouville's equation numerically for the optical systems considered in this article. To that end, we combine the scheme with a new method we refer to as sub-cell interface method. When dealing with optical interfaces in phase space, non-local boundary conditions arise. These are incorporated in the DG method in an energy-preserving manner. Numerical experiments validate energy-preservation up to machine precision and show the high order of accuracy. Furthermore, the DG method is compared to quasi-Monte Carlo ray tracing for two examples showing that the DG method yields better accuracy in the same amount of computational time.

1. Introduction

Illumination optics deals with the design of optical systems for various applications, for instance, street lighting [1] and automotive headlamps [2,3]. (Quasi-)Monte Carlo ray tracing is typically employed to compute relevant photometric quantities, such as illuminance or intensity, on a target. (Quasi-)Monte Carlo ray tracing can be expensive in computing the photometric quantities to high accuracy, due to its rather slow convergence. Furthermore, low accuracy results make the task of performing numerical optimisation on the optical system very challenging.

A different approach to ray tracing is based on a phase space description of light propagation [4–6]. Here, phase space is defined as the collection of all positions and direction coordinates of light rays. The evolution of these coordinates, of a single light ray, is described by a Hamiltonian system, whenever the refractive index field is smooth. When a light ray hits an optical interface, that is a discontinuity in the refractive index field, the laws of refraction or reflection have to be

* Corresponding author.
E-mail address: r.a.m.v.gestel@tue.nl (R.A.M. van Gestel).

https://doi.org/10.1016/j.jcp.2023.112208

applied. On phase space one can define a quantity describing an energy density known as the basic luminance [7]. From the basic luminance the illuminance and intensity can be computed by integration. The evolution of the basic luminance is governed by Liouville's equation for geometrical optics. At an optical interface Snell's law of refraction and the law of specular reflection have to be applied, which describe a discontinuous change in the direction coordinate for a light ray. This results in non-local boundary conditions for the basic luminance at an optical interface.

In this work we consider a discontinuous Galerkin (DG) method to solve Liouville's equation. DG methods are particularly suitable for solving Liouville's equation due to their compact stencil, their ability to deal with complex geometries [8] and their potential for high parallelisation efficiency [9]. Curved optical interfaces manifest themselves as moving boundaries in phase space. To accommodate the moving boundaries we employ an Arbitrary Lagrangian-Eulerian (ALE) formulation [10,11]. In the ALE formulation Liouville's equation is transformed from a moving domain to a static domain with an appropriate transformation. This allows us to align the mesh with the optical interfaces.

The DG approach allows for arbitrary high order of accuracy in space. By combining DG with an Arbitrary Derivative (ADER) approach one can also achieve arbitrary high order of accuracy in the evolution coordinate. The ADER methodology was first developed for finite volume methods by Titarev and Toro in [12–14]. Later it was extended to DG schemes by Qiu et al. [15] and Dumbser et al. [16]. In those works, an element-local temporal Taylor expansion is computed where temporal derivatives are replaced with spatial derivatives using the Cauchy-Kovalewski or Lax-Wendroff procedure. This procedure becomes rather cumbersome for non-linear partial differential equations since it is problem dependent. To allow for a more general treatment a local space-time Galerkin predictor method based on a space-time weak formulation was developed by Dumbser et al. [17,18]. For recent applications of the latter approach see for example [19–21]. For a comparison between different ADER approaches, we refer the reader to [22]. The ADER-DG schemes yield a fully-discrete explicit scheme as opposed to a semi-discrete multi-stage scheme when DG is combined with an explicit Runge-Kutta method.

In [23,24], ADER-DG methods in the ALE formulation have been used with a local space-time Galerkin predictor. Furthermore, in [25], the authors employ an explicit Taylor series on a moving mesh for a spatially one-dimensional setting, but only for up to second order in time. Higher order in time predictors is achieved using continuous explicit Runge-Kutta schemes [26]. In this work, we employ the Cauchy-Kovalewski procedure to derive an element-local Taylor expansion on a moving mesh in a spatially two-dimensional setting, up to arbitrary order of accuracy, and we will exploit specifics of our problem to simplify the procedure. That is, a particular choice of mesh and mesh movement yields simpler computations for the considered optical systems.

The moving mesh method alone is not sufficient to solve Liouville's equation for geometrical optics numerically, as for certain optical systems it can lead to an increasingly smaller mesh spacing which via a CFL stability condition leads to an ever decreasing stepsize. Therefore, we introduce a new method we refer to as the sub-cell interface method, to resolve this issue. Similar to the ADER-DG method, the method is also based on a weak formulation over a phase space element, where now an optical interface is allowed to cut the element into two pieces during a single step.

As mentioned before, at an optical interface there are non-local boundary conditions which have to be incorporated into the DG scheme as numerical fluxes across the optical interface. Snell's law of refraction and the law of specular reflection describe how phase space is connected at an optical interface. Moreover, both laws depend on the unit surface normal of the optical interface which can change for curved interfaces. This makes dealing with optical interfaces rather complicated. In our previous work [7] we presented a method that incorporated the non-local boundary conditions into the numerical fluxes of a DG method in an energy-preserving manner, for a simple flat optical interface, by satisfying an energy balance discretely. There light rays propagated only in the forward direction even after reflection and/or refraction, where the forward direction means that the position of a light ray is always increasing along one position axis. For arbitrary curved interfaces light rays can change direction. Here, we extend our previous work by formally incorporating this change in direction into the discretisation at optical interfaces. Moreover, we formulate and prove energy balances that should hold discretely for curved optical interfaces. Consequently, we are able to deal with arbitrary curved optical interfaces in an energy-preserving manner.

The structure of this paper is as follows: in Section 2 we describe the setting for Liouville's equation and the relevant laws of optics. In Section 3.1 we present the discretisation of Liouville's equation on a moving mesh using DG and in Section 3.2 we develop the necessary temporal Taylor expansions used in the ADER approach. Next, we present the discretisation using the sub-cell interface method and how to deal with optical interfaces in Sections 3.3 and 3.4. As the mesh movement can cause large deformations, we briefly discuss mesh refinement in Section 3.5. Numerical experiments and comparisons with quasi-Monte Carlo ray tracing are carried out in Section 4. Finally we present our conclusions in Section 5.

2. Liouville's equation

A light ray can be described by its position vector $(\boldsymbol{q}, z) \in \mathbb{R}^{d+1}$ and momentum vector $(\boldsymbol{p}, p_z) \in \mathbb{R}^{d+1}$ $(d = 1, 2)$, where the momentum vector is defined as the unit direction vector multiplied by the refractive index n. The momentum vector has a fixed length $|(\boldsymbol{p}, p_z)| = n$, hence, one can write $p_z = \sigma\sqrt{n^2 - |\boldsymbol{p}|^2}$ with $\sigma \in \{-1, 1\}$. Rather than taking time or the arc-length to describe the evolution of a light ray, we take its z-coordinate as the evolution coordinate [6]. The ray coordinates $\boldsymbol{q}(z)$ and $\boldsymbol{p}(z)$ evolve according to Hamilton's equations

$$\frac{\mathrm{d}\boldsymbol{q}}{\mathrm{d}z} = \frac{\partial H}{\partial \boldsymbol{p}}, \tag{1a}$$

$$\frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}z} = -\frac{\partial H}{\partial \boldsymbol{q}}, \tag{1b}$$

with the Hamiltonian H given by

$$H(z, \boldsymbol{q}, \boldsymbol{p}) = -\sigma\sqrt{n(z, \boldsymbol{q})^2 - |\boldsymbol{p}|^2}, \tag{1c}$$

where $n(z, \boldsymbol{q})$ describes the refractive index field as a function of the position coordinates. In the Hamiltonian system (1) and the definition of p_z the term σ describes the direction of light rays with respect to the z-axis, i.e., $\sigma = -1$ describes backward rays and $\sigma = 1$ describes forward rays.

The collection of all momentum vectors with fixed length n lies on Descartes' sphere with radius n [6]. We denote the d-dimensional unit sphere as $S^d \subset \mathbb{R}^{d+1}$ and the sphere with radius n as $S^d(n)$. Due to $(\boldsymbol{p}, p_z) \in S^d(n)$, the momentum $\boldsymbol{p}$ is restricted by $|\boldsymbol{p}| \leq n$, and the collection of momenta $\boldsymbol{p}$ lies on one of the two discs describing either forward ($\sigma = 1$) or backward rays ($\sigma = -1$). At a plane $z = \text{const}$, the collection of all positions $\boldsymbol{q}$ and momenta on either disk combine to the $2d$-dimensional phase space domain $\mathcal{P}_\sigma$ for either forward rays ($\sigma = 1$) or backward rays ($\sigma = -1$). Furthermore, $d = 1$ for two-dimensional optics and $d = 2$ for three-dimensional optics. The flow generated by Hamilton's equations describes symplectic transformations which means that a volume element of phase space $\mathrm{d}\mathcal{U} = \mathrm{d}\boldsymbol{q}\mathrm{d}\boldsymbol{p}$ remains constant [27]. A volume in phase space is known in optics as étendue [28].

The energy (per time) of a beam of light remains constant when there are no losses due to, e.g., scattering, Fresnel reflections or absorption. In illumination optics this energy is known as the luminous flux, denoted by Φ. An infinitesimal element of luminous flux can be related to a quantity ρ_σ defined on phase space $\mathcal{P}_\sigma$ by $\mathrm{d}\Phi = \rho_\sigma\,\mathrm{d}\mathcal{U}$. The subscript σ for ρ_σ is used to distinguish between forward and backward propagating light. Now, if the beam of light is propagated along the z-axis over some distance, then in the absence of losses $\mathrm{d}\Phi$ remains constant. Since also the phase space volume element $\mathrm{d}\mathcal{U}$ remains constant, ρ_σ must remain invariant. Whenever ρ_σ is sufficiently smooth one can write this invariance of ρ_σ as follows

$$\frac{\mathrm{d}}{\mathrm{d}z}\rho_\sigma(z, \boldsymbol{q}(z), \boldsymbol{p}(z)) = 0. \tag{2}$$

In illumination optics the quantity ρ_σ is referred to as the basic luminance [28]. Finally, we remark that if the basic luminance is known at a plane $z = \text{const}$, other quantities such as the illuminance and luminous intensity can be computed. For instance, the illuminance can be computed by integrating ρ_σ over all the momenta $\boldsymbol{p}$, and the luminous intensity can be computed by integrating $\rho_\sigma\, p_z n$ over all the positions $\boldsymbol{q}$; see [7] for more detail.

Assuming sufficient smoothness, one can derive Liouville's equation by taking the total $\frac{\mathrm{d}}{\mathrm{d}z}$-derivative in (2), yielding

$$\frac{\partial \rho_\sigma}{\partial z} + \frac{\partial H}{\partial \boldsymbol{p}} \cdot \frac{\partial \rho_\sigma}{\partial \boldsymbol{q}} - \frac{\partial H}{\partial \boldsymbol{q}} \cdot \frac{\partial \rho_\sigma}{\partial \boldsymbol{p}} = 0, \tag{3}$$

where we have made use of Hamilton's equations (1). The advective form of Liouville's equation can be transformed to a conservative form, i.e.,

$$\frac{\partial \rho_\sigma}{\partial z} + \nabla \cdot (\rho_\sigma \boldsymbol{u}) = 0 \tag{4a}$$

with the velocity field $\boldsymbol{u}$ defined by

$$\boldsymbol{u} = \begin{pmatrix} \frac{\partial H}{\partial \boldsymbol{p}} \\ -\frac{\partial H}{\partial \boldsymbol{q}} \end{pmatrix} = \frac{1}{\sigma\sqrt{n^2 - |\boldsymbol{p}|^2}} \begin{pmatrix} \boldsymbol{p} \\ n\frac{\partial n}{\partial \boldsymbol{q}} \end{pmatrix}, \tag{4b}$$

where we have used that the velocity field $\boldsymbol{u}$ is divergence-free and $\nabla = \left(\frac{\partial}{\partial \boldsymbol{q}}, \frac{\partial}{\partial \boldsymbol{p}}\right)$.

At an optical interface the Hamiltonian H is discontinuous and therefore (2) is not valid at an optical interface. However, at an optical interface the basic luminance does remain invariant, in the absence of losses, even when light is reflected or refracted [28,29]. Consequently, at an optical interface we enforce invariance of the basic luminance together with Snell's law of refraction or the law of specular reflection. This is expressed in the following jump condition

$$\rho_{\sigma(z^+)}(z^+, \boldsymbol{q}(z^+), \boldsymbol{p}(z^+)) = \rho_{\sigma(z^-)}(z^-, \boldsymbol{q}(z^-), \boldsymbol{p}(z^-)), \tag{5a}$$

where we explicitly denote σ as $\sigma(z^\pm)$ since it might change, and the superscript $\pm$ denotes one-sided limits towards the optical interface that correspond to incident and outgoing light for $-$ and $+$, respectively. We compute the full momentum vector $(\boldsymbol{p}, p_z)(z^+)$ as

$$(\boldsymbol{p}, p_z)(z^+) = \mathcal{S}\Big((\boldsymbol{p}, p_z)(z^-); n_0, n_1, \hat{\boldsymbol{\nu}}\Big) \quad \text{and} \quad \operatorname{sgn} p_z(z^+) = \sigma(z^+). \tag{5b}$$

In (5a)-(5b) we explicitly denote the sign of p_z with $\sigma(z^{\pm})$. The change in momentum (5b) is described by vectorial versions of the law of specular reflection and Snell's law of refraction, which depend on the refractive indices of the incident and transmitted media denoted by n_0 and n_1, respectively, and the surface unit normal $\hat{\boldsymbol{\nu}} \in \mathbb{R}^{d+1}$ at the point $(\boldsymbol{q}(z^-), z^-)$. To be explicit, in equation (5b) the function $\mathcal{S}$ can either describe refraction or specular reflection, relating the incident momentum $\boldsymbol{i} = (\boldsymbol{p}, p_z) \in \mathbb{R}^{d+1}$ to an outgoing momentum as follows

$$\mathcal{S}(\boldsymbol{i}; n_0, n_1, \hat{\boldsymbol{\nu}}) = \begin{cases} \mathcal{S}_{\mathrm{R}} = \boldsymbol{i} - 2\psi\hat{\boldsymbol{\nu}} & \text{if } \delta \leq 0, \\ \mathcal{S}_{\mathrm{T}} = \boldsymbol{i} - (\psi + \sqrt{\delta})\hat{\boldsymbol{\nu}} & \text{if } \delta > 0, \end{cases} \tag{6a}$$

with

$$\psi = \boldsymbol{i} \cdot \hat{\boldsymbol{\nu}} \text{ and } \delta = n_1^2 - n_0^2 + \psi^2. \tag{6b}$$

The sign of the normal should be taken such that $\psi \leq 0$, i.e., $\hat{\boldsymbol{\nu}}$ points towards the medium of the incident ray.

In the definition of $\mathcal{S}$ we have also defined the functions for reflection and refraction denoted by $\mathcal{S}_{\mathrm{R}}$ and $\mathcal{S}_{\mathrm{T}}$, respectively. Reflection transforms quantities on a Descartes' sphere, that is $\mathcal{S}_{\mathrm{R}} : \mathrm{S}^d(n_0) \to \mathrm{S}^d(n_0)$. For refraction $\mathcal{S}_{\mathrm{T}}$ it can happen that $\delta < 0$ so that the result yields complex numbers, whereas for $\delta \geq 0$ it takes a momentum from $\mathrm{S}^d(n_0)$ and returns a momentum on $\mathrm{S}^d(n_1)$.

In (6) we either use reflection $\mathcal{S}_{\mathrm{R}}$ or refraction $\mathcal{S}_{\mathrm{T}}$ depending on the sign of δ, where for reflection the case is usually referred to as total internal reflection. We will often abuse the notation to write $\mathcal{S}(\boldsymbol{p}) = \mathcal{S}(\boldsymbol{p}; n_0, n_1, \hat{\boldsymbol{\nu}})$, and similarly for $\mathcal{S}_{\mathrm{R}}$ and $\mathcal{S}_{\mathrm{T}}$, where the refractive indices and normal should be clear from the context. Moreover, we will need reflection and refraction in reversed directions, which we denote by $\mathcal{S}^{-1}$, i.e., for reflection the expression reads

$$\mathcal{S}_{\mathrm{R}}^{-1}(\boldsymbol{p}; n_0, n_1, \hat{\boldsymbol{\nu}}) = -\mathcal{S}_{\mathrm{R}}(-\boldsymbol{p}; n_0, n_1, \hat{\boldsymbol{\nu}}), \tag{7a}$$

and for refraction [30]

$$\mathcal{S}_{\mathrm{T}}^{-1}(\boldsymbol{p}; n_0, n_1, \hat{\boldsymbol{\nu}}) = -\mathcal{S}_{\mathrm{T}}(-\boldsymbol{p}; n_1, n_0, -\hat{\boldsymbol{\nu}}). \tag{7b}$$

For completeness, we remark that the relations (5) and (6) state that light can only be either fully reflected or fully refracted, that is, there are no partial reflections, so Fresnel reflections are not taken into account.

In geometrical optics, light rays evolve according to Hamilton's equations (1) and describe the location of the characteristics of Liouville's equation (4a), along which the basic luminance remains constant. Furthermore, at an optical interface characteristics change discontinuously according to (6) whereby (5a) the basic luminance remains constant [30]. Solving Hamilton's equation and applying (5b)-(6) at an optical interface, is commonly referred to as ray tracing [28]. For simple optical systems it can be manageable to trace a light ray back from a certain target, e.g., $z = Z$, to the source at $z = 0$, such that we can determine the exact solution to Liouville's equation at $z = Z$. In particular, we will apply this technique in the method described in Section 3.3 and in the first example shown in Section 4.

3. Numerical method

In what follows we will only consider two-dimensional optics ($d = 1$). Therefore, we will omit the bold-faced notation for d-dimensional vectors and instead write q for the position and p for the momentum coordinate on phase space. Moreover, we will consider only forward propagating light rays unless stated otherwise. Thus we take $\sigma = 1$ and omit the σ in Liouville's equation (4a), so that we can write equation (4a) as

$$\frac{\partial \rho}{\partial z} + \nabla \cdot (\rho \boldsymbol{u}) = 0, \tag{8}$$

where now $\nabla = (\frac{\partial}{\partial q}, \frac{\partial}{\partial p})$. Furthermore, we will only consider piecewise constant refractive index fields. A curved optical interface given by $q = Q(z)$ manifests itself as a moving boundary in phase space. The phase space domain, hence, is z-dependent, so we denote the phase space domain as $\mathcal{P}(z)$. In Sections 3.1-3.5 we will describe the discretisation of Liouville's equation on the moving phase space domain $\mathcal{P}(z)$.

3.1. DG on moving mesh

We employ an Arbitrary Lagrangian-Eulerian discontinuous Galerkin (ALE-DG) method, where we can prescribe a velocity to move the mesh such that it remains aligned with optical interfaces. In other words, we consider the DG method on a moving mesh. The phase space domain $\mathcal{P}(z)$ is partitioned into Cartesian elements, where each element is a Cartesian product of one-dimensional intervals. Let $\Omega(z) = [q_0(z), q_1(z)] \times [p_0, p_1]$ denote one such a Cartesian element.

First, we transform Liouville's equation (8) to a static reference domain by considering the following transformation from the reference square $\chi = [0, 1]^2$ to the element $\Omega(z)$, which reads

$$\boldsymbol{x}(\tau,\boldsymbol{\xi}) = \begin{pmatrix} q(\tau,\xi) \\ p(\eta) \end{pmatrix} = \begin{pmatrix} (1-\xi)q_0(\tau) + \xi q_1(\tau) \\ (1-\eta)p_0 + \eta p_1 \end{pmatrix} = \begin{pmatrix} q_0(\tau) + \xi\,\Delta q(\tau) \\ p_0 + \eta \Delta p \end{pmatrix}, \tag{9}$$

where $\Delta q(\tau) = q_1(\tau) - q_0(\tau)$, $\Delta p = p_1 - p_0$, $z = \tau$ and $\boldsymbol{\xi} = (\xi,\eta)$. Let us introduce $\rho^*(\tau,\boldsymbol{\xi}) = \rho(\tau,\boldsymbol{x}(\tau,\boldsymbol{\xi}))$, then the τ-derivative of ρ^* reads

$$\frac{\partial \rho^*}{\partial \tau} = \frac{\partial \rho}{\partial z}\frac{\mathrm{d}z}{\mathrm{d}\tau} + \boldsymbol{v}\cdot\nabla\rho, \tag{10}$$

where $\boldsymbol{v} = \frac{\partial \boldsymbol{x}}{\partial \tau}$ denotes the mesh velocity, i.e., the velocity at which we move the mesh. Subsequently, we insert Liouville's equation (8), use $z = \tau$ and apply the product rule on the last term of (10) which leads after some rewriting to

$$\frac{\partial \rho^*}{\partial \tau} = -\nabla\cdot\left(\rho^*(\boldsymbol{u}-\boldsymbol{v})\right) - \rho^*\nabla\cdot\boldsymbol{v}. \tag{11}$$

The transformation (9) of the spatial domain to the reference domain, transforms the divergence of a general function $\boldsymbol{g} = (g_1, g_2)$ as

$$\nabla\cdot\boldsymbol{g} = \frac{1}{\mathcal{J}}\nabla_{\boldsymbol{\xi}}\cdot\tilde{\boldsymbol{g}} \text{ with } \tilde{\boldsymbol{g}} = \begin{pmatrix} g_1\,\Delta p \\ g_2\,\Delta q \end{pmatrix}, \tag{12}$$

where $\mathcal{J}(\tau) = \Delta q(\tau)\Delta p$ denotes the Jacobian determinant of the transformation (9), $\nabla_{\boldsymbol{\xi}} = (\frac{\partial}{\partial\xi}, \frac{\partial}{\partial\eta})$ denotes the gradient on the reference domain and the quantities with a tilde denote their transformed counterparts. Consequently, transforming the divergences of (11) to the reference domain yields

$$\frac{\partial \rho^*}{\partial \tau} = -\frac{1}{\mathcal{J}}\nabla_{\boldsymbol{\xi}}\cdot\left(\rho^*(\tilde{\boldsymbol{u}}-\tilde{\boldsymbol{v}})\right) - \frac{\rho^*}{\mathcal{J}}\nabla_{\boldsymbol{\xi}}\cdot\tilde{\boldsymbol{v}}. \tag{13}$$

Next, we multiply (13) by $\mathcal{J}$ and make use of the so-called geometric conservation law [10]

$$\frac{\mathrm{d}\mathcal{J}}{\mathrm{d}\tau} = \frac{\mathrm{d}}{\mathrm{d}\tau}\left(\Delta q(\tau)\Delta p\right) = \nabla_{\boldsymbol{\xi}}\cdot\tilde{\boldsymbol{v}}, \tag{14}$$

such that we obtain

$$\frac{\partial\left(\rho^*\mathcal{J}\right)}{\partial\tau} + \nabla_{\boldsymbol{\xi}}\cdot\left(\rho^*(\tilde{\boldsymbol{u}}-\tilde{\boldsymbol{v}})\right) = 0. \tag{15}$$

The geometric conservation law states that mesh motion does not disturb a uniform solution [10], i.e., if ρ^* is uniform then (15) reduces to the geometric conservation law (14). Finally, we omit the $*$ in (15) and introduce the transformed flux $\tilde{\boldsymbol{f}}$, such that the conservation law on the reference domain can be written as

$$\frac{\partial\left(\rho\mathcal{J}\right)}{\partial\tau} + \nabla_{\boldsymbol{\xi}}\cdot\tilde{\boldsymbol{f}} = 0 \text{ with } \tilde{\boldsymbol{f}} = \rho(\tilde{\boldsymbol{u}}-\tilde{\boldsymbol{v}}). \tag{16}$$

Typically, the numerical solution of ρ at $z = z^t$ is known and we want to evolve the numerical solution to $z = z^{t+1}$, where t denotes the step index. The DG method is based on the weak formulation. The weak formulation of equation (16) with test function $\phi_k = \phi_k(\boldsymbol{\xi})$ is written as

$$\int_{z^t}^{z^{t+1}}\int_{\chi}\phi_k\left(\frac{\partial\left(\rho\mathcal{J}\right)}{\partial\tau} + \nabla_{\boldsymbol{\xi}}\cdot\tilde{\boldsymbol{f}}\right)\mathrm{d}\boldsymbol{\xi}\,\mathrm{d}\tau = 0. \tag{17}$$

We apply the product rule and Gauss's theorem to the last term on the left-hand side of equation (17), which yields

$$\int_{\chi}\left(\rho\mathcal{J}\right)^{t+1}\phi_k\,\mathrm{d}\boldsymbol{\xi} - \int_{\chi}\left(\rho\mathcal{J}\right)^{t}\phi_k\,\mathrm{d}\boldsymbol{\xi} = \int_{z^t}^{z^{t+1}}\left(\int_{\chi}\left(\nabla_{\boldsymbol{\xi}}\phi_k\right)\cdot\tilde{\boldsymbol{f}}\,\mathrm{d}\boldsymbol{\xi} - \int_{\partial\chi}\phi_k\tilde{\boldsymbol{F}}\cdot\hat{\boldsymbol{N}}\,\mathrm{d}\sigma\right)\mathrm{d}\tau, \tag{18}$$

where we have replaced the flux in the boundary integral with a numerical flux $\tilde{\boldsymbol{F}}$ and $\hat{\boldsymbol{N}}$ denotes the outward unit normal on the reference domain χ. The numerical flux $\tilde{\boldsymbol{F}}$ depends on the left and right states at the interface denoted by ρ^- and ρ^+, respectively. For the numerical flux, we employ the upwind flux, i.e.,

$$\tilde{\boldsymbol{F}}(\rho^-,\rho^+)\cdot\hat{\boldsymbol{N}} = \left(\tilde{\boldsymbol{u}}-\tilde{\boldsymbol{v}}\right)\cdot\hat{\boldsymbol{N}}\begin{cases}\rho^- & \text{if } \left(\tilde{\boldsymbol{u}}-\tilde{\boldsymbol{v}}\right)\cdot\hat{\boldsymbol{N}} \geq 0, \\ \rho^+ & \text{if } \left(\tilde{\boldsymbol{u}}-\tilde{\boldsymbol{v}}\right)\cdot\hat{\boldsymbol{N}} < 0.\end{cases} \tag{19}$$

The numerical solution on each element is represented by an expansion into basis functions. As basis functions we employ a tensor-product of one-dimensional Lagrange polynomials ℓ_i of degree N, defined on Gauss-Legendre quadrature nodes $\{\xi_i\}_{i=0}^N$ over the interval $[0,1]$. The Lagrange polynomials $\ell_i(\xi)$ have the property

$$\ell_i(\xi_j) = \delta_{ij} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{otherwise,} \end{cases}$$

with δ_{ij} denoting the Kronecker delta. Hence, the polynomial basis functions form a nodal basis. Moreover, the Lagrange polynomials are orthogonal with respect to the L_2-inner product, consequently, applying $(N+1)$-point Gauss-Legendre quadrature with nodes $\{\xi_i\}_{i=0}^N$ and weights $\{w_i\}_{i=0}^N$ one obtains

$$\int_0^1 \ell_i(\xi)\ell_j(\xi)\,\mathrm{d}\xi = \sum_{n=0}^N w_n \ell_i(\xi_n)\ell_j(\xi_n) = w_i\delta_{ij}. \tag{20}$$

On the reference domain χ the basis functions are denoted as ϕ_l, which formally are given by

$$\phi_l(\boldsymbol{\xi}) = \ell_i(\xi)\ell_j(\eta) \text{ with } l = (N+1)j + i + 1, \tag{21}$$

and $i = 0, 1, \ldots, N$ and $j = 0, 1, \ldots, N$. The expansion of ρ in terms of these basis functions reads

$$\rho_{\mathrm{h}}(z^t, \boldsymbol{\xi}) = \sum_{l=1}^{N_d} \rho_l^t \phi_l(\boldsymbol{\xi}), \tag{22}$$

where $N_d = (N+1)^2$ denotes the number of degrees of freedom. Inserting the expansion (22) into the left-hand side of equation (18), yields

$$\sum_{l=1}^{N_d} \left(\int_\chi \phi_l\phi_k\,\mathrm{d}\boldsymbol{\xi} \right) \left((\rho_l\mathcal{J})^{t+1} - (\rho_l\mathcal{J})^t \right) = \int_{z^t}^{z^{t+1}} \left(\int_\chi \left(\nabla_{\boldsymbol{\xi}}\phi_k\right)\cdot\tilde{\boldsymbol{f}}\,\mathrm{d}\boldsymbol{\xi} - \int_{\partial\chi} \phi_k \tilde{\boldsymbol{F}}\cdot\hat{\boldsymbol{N}}\,\mathrm{d}\sigma \right) \mathrm{d}\tau, \tag{23}$$

with $(\rho_l\mathcal{J})^t = \rho_l^t\mathcal{J}^t$ where $\mathcal{J}^t$ denotes the numerical approximation of $\mathcal{J}(z^t)$. By letting $k = 1, 2, \ldots, N_d$ we arrive at the N_d equations for the expansion coefficients ρ_l^{t+1}. All the integrals in equation (23) are evaluated with $(N+1)$-point Gauss-Legendre quadrature. For the right-hand side of equation (23) we require the solution ρ at intermediate levels of $[z^t, z^{t+1}]$. In the next section, we will describe how these values are computed using the ADER approach.

In addition to solving equation (23), we also solve the trajectory equation for the vertices $q_i(\tau)$ with $i = 0, 1$, e.g.,

$$\frac{\mathrm{d}q_i}{\mathrm{d}\tau} = a_i(\tau), \tag{24}$$

with $a_i(\tau)$ the mesh velocity at the vertex q_i. We solve (24) in the same way as equation (23), i.e., we compute

$$q_i^{t+1} = q_i^t + \int_{z^t}^{z^{t+1}} a_i(\tau)\,\mathrm{d}\tau, \tag{25}$$

where q_i^t denotes the numerical approximation of $q_i(z^t)$, etc., by applying $(N+1)$-point Gauss-Legendre quadrature. After computing the new vertex locations q_0^{t+1} and q_1^{t+1}, the Jacobian $\mathcal{J}$ is updated using

$$\mathcal{J}^{t+1} = (q_1^{t+1} - q_0^{t+1})\Delta p, \tag{26}$$

in agreement with solving

$$\mathcal{J}^{t+1} = \mathcal{J}^t + \Delta p \int_{z^t}^{z^{t+1}} (a_1(\tau) - a_0(\tau))\,\mathrm{d}\tau, \tag{27}$$

which is the integration of the geometric conservation law (14).

3.2. z-integration using local ADER predictor

3.2.1. Moving element

To compute the right-hand side of equation (23) we generally require the solution ρ at intermediate levels of $[z^t, z^{t+1}]$. In the ADER approach one computes a predictor approximating the z-evolution locally on each element without considering neighbouring elements. In particular, we employ a Taylor expansion about the old level and subsequently apply the Cauchy-Kovalewski procedure [16,22] where we repeatedly replace τ-derivatives with spatial derivatives using the governing equation. The Taylor expansion up to degree M about the old level $\tau = z^t$, where the solution is known, on the reference domain reads

$$\rho(z^t + \tau, \boldsymbol{\xi}) \approx \sum_{k=0}^{M} \frac{1}{k!} \tau^k \frac{\partial^k \rho}{\partial \tau^k}(z^t, \boldsymbol{\xi}). \tag{28}$$

We require the governing equation of ρ on the reference domain. Therefore, we first rewrite equation (11) in an advective form, by using $\nabla \cdot \boldsymbol{u} = 0$, yielding

$$\frac{\partial \rho}{\partial \tau} = -(\boldsymbol{u} - \boldsymbol{v}) \cdot \nabla \rho,$$

and subsequently we transform the gradient to the reference domain using (12) which leads to

$$\frac{\partial \rho}{\partial \tau} = -\frac{1}{\mathcal{J}} \left(\tilde{\boldsymbol{u}} - \tilde{\boldsymbol{v}}\right) \cdot \nabla_{\boldsymbol{\xi}} \rho. \tag{29}$$

Recall, that we consider only a piecewise constant refractive index field and thus by relation (4b) the last component of $\boldsymbol{u}$ is zero. Moreover, we consider only mesh movement w.r.t. the q-axis so that we can write $\boldsymbol{u} = (u_0, 0)$ and $\boldsymbol{v} = (v_0, 0)$. This allows us to rewrite relation (29) as

$$\frac{\partial \rho}{\partial \tau} = -\frac{1}{\Delta q} (u_0 - v_0) \frac{\partial \rho}{\partial \xi}, \tag{30}$$

where we have used (12) and $\mathcal{J} = \Delta q \Delta p$. To simplify the notation it is convenient to write (30) as

$$\frac{\partial \rho}{\partial \tau} = c(\tau, \boldsymbol{\xi}) \frac{\partial \rho}{\partial \xi} \text{ with } c(\tau, \xi) = a(\tau) + \xi b(\tau), \tag{31a}$$

where we use that the velocity field v_0 is linear in ξ, cf. (9), such that a and b read

$$a(\tau) = \frac{1}{\Delta q} \left(\frac{\partial q_0}{\partial \tau} - u_0 \right) \text{ and } b(\tau) = \frac{1}{\Delta q} \frac{\partial \Delta q}{\partial \tau}, \tag{31b}$$

where we omit u_0's dependence on η.

From equation (31a) we can generate higher order τ-derivatives solely in terms of spatial derivatives as follows. First, we take the τ-derivative and the ξ-derivative of equation (31a) such that we obtain

$$\frac{\partial^2 \rho}{\partial \tau^2} = \frac{\partial c}{\partial \tau} \frac{\partial \rho}{\partial \xi} + c \frac{\partial^2 \rho}{\partial \xi \partial \tau},$$
$$\frac{\partial^2 \rho}{\partial \xi \partial \tau} = b \frac{\partial \rho}{\partial \xi} + c \frac{\partial^2 \rho}{\partial \xi^2},$$

where combining both relations leads to

$$\frac{\partial^2 \rho}{\partial \tau^2} = \left(\frac{\partial c}{\partial \tau} + bc \right) \frac{\partial \rho}{\partial \xi} + c^2 \frac{\partial^2 \rho}{\partial \xi^2}. \tag{32}$$

Similarly, expressions for higher order τ-derivatives can be found. For higher derivatives the expressions can become rather large, however, they can still be found with the aid of a computer algebra program. For example, the third order derivative reads

$$\frac{\partial^3 \rho}{\partial \tau^3} = \left(\frac{\partial^2 c}{\partial \tau^2} + 2b \frac{\partial c}{\partial \tau} + \frac{\partial b}{\partial \tau} c + b^2 c \right) \frac{\partial \rho}{\partial \xi} + 3c \left(\frac{\partial c}{\partial \tau} + bc \right) \frac{\partial^2 \rho}{\partial \xi^2} + c^3 \frac{\partial^3 \rho}{\partial \xi^3}. \tag{33}$$

Finally, we can insert the relations for the τ-derivatives into the Taylor expansion (28).

From an implementation point of view it is more efficient to rearrange the Taylor expansion (28) by expressing it in terms of ξ-derivatives, and thus reducing the number of ξ-derivative evaluations. If we consider all the τ-derivatives from order 0 to M, then we rewrite the Taylor expansion (28) as follows

$$\rho(z^t+\tau,\boldsymbol{\xi}) \approx \sum_{k=0}^{M} \mathcal{C}_k(\tau,\boldsymbol{\xi}) \frac{\partial^k \rho}{\partial \xi^k}(z^t,\boldsymbol{\xi}), \tag{34}$$

where $\mathcal{C}_k(\tau,\boldsymbol{\xi})$ describes the coefficient of the k-th ξ-derivative. For example, for $M=3$ the coefficients $\mathcal{C}_k(\tau,\boldsymbol{\xi})$ read

$$\begin{aligned}
\mathcal{C}_0(\tau,\boldsymbol{\xi}) &= 1,\\
\mathcal{C}_1(\tau,\boldsymbol{\xi}) &= \tau c + \frac{\tau^2}{2}\left(\frac{\partial c}{\partial \tau} + bc\right) + \frac{\tau^3}{3!}\left(\frac{\partial^2 c}{\partial \tau^2} + 2b\frac{\partial c}{\partial \tau} + \frac{\partial b}{\partial \tau}c + b^2 c\right),\\
\mathcal{C}_2(\tau,\boldsymbol{\xi}) &= \frac{\tau^2}{2}c^2 + \frac{\tau^3}{3!}3c\left(\frac{\partial c}{\partial \tau} + bc\right),\\
\mathcal{C}_3(\tau,\boldsymbol{\xi}) &= \frac{\tau^3}{3!}c^3,
\end{aligned} \tag{35}$$

where b and c are evaluated at $(z^t,\boldsymbol{\xi})$.

Finally, we combine the Taylor expansion (34) with the expansion for ρ_{h} (22) to compute the spatial derivatives, completing the local ADER predictor for moving elements. Consequently, we can compute ρ at the required Gauss-Legendre quadrature points to compute the right-hand side of equation (23).

The complete scheme obeys a property called constant state preservation, whenever the refractive index field is constant. Constant state preservation means that a uniform solution must remains uniform, hence, obeying (14) and (15). Numerically, this means that the discretisation will exactly (up to machine precision) preserve a constant state, which needs to be independent of the mesh motion. This property is proven in Appendix A.

3.2.2. Static element

In the special case where an element does not move, i.e., $v_0(\tau,\xi)=0$, the advection equation (30) reduces to

$$\frac{\partial \rho}{\partial \tau} = -\frac{u_0}{\Delta q}\frac{\partial \rho}{\partial \xi}, \tag{36}$$

and hence, the higher order τ-derivatives can easily be found to be

$$\frac{\partial^k \rho}{\partial \tau^k} = \left(-\frac{u_0}{\Delta q}\right)^k \frac{\partial^k \rho}{\partial \xi^k}. \tag{37}$$

Hence, the Taylor expansion on a static element reads

$$\rho(z^t+\tau,\boldsymbol{\xi}) \approx \sum_{k=0}^{M} \frac{1}{k!}\left(-\frac{u_0\tau}{\Delta q}\right)^k \frac{\partial^k \rho}{\partial \xi^k}(z^t,\boldsymbol{\xi}). \tag{38}$$

Another consequence of $v_0(\tau,\xi)=0$ is that the flux $\tilde{\boldsymbol{f}}$ reduces to $\tilde{\boldsymbol{f}}=\rho\tilde{\boldsymbol{u}}$. The transformed velocity $\tilde{\boldsymbol{u}}$ does not depend on τ, since we are considering piecewise constant refractive index fields. Therefore, inserting the Taylor expansion (38) into the flux $\tilde{\boldsymbol{f}}$ results in a flux that is a polynomial in τ. Similarly, the numerical flux (19) with the expansion (38) is a polynomial in τ.

Once again, we insert the Taylor expansion (38) into the right-hand side of equation (23). The τ-integral can now easily be computed without quadrature since the integrand is just a polynomial in τ. Hence, for static elements we compute the τ-integral analytically.

Since no quadrature rule for the τ-integral is necessary, a static element is cheaper to update than a moving element. Moreover, the coefficients in the Taylor expansion of a static element only depend on η via u_0 and are less complicated than those in a moving element; one can see this by comparing expansion (38) to (34)-(35).

3.2.3. CFL condition

The ALE-ADER-DG method described thus far is an explicit one-step high order DG method. This explicit method must also obey a CFL stability condition, which imposes a condition on the stepsize Δz and reads for explicit DG schemes as [19,20]

$$\Delta z \leq \frac{1}{2d}\frac{\mathrm{CFL}}{2N+1}\min_e \frac{h_e}{w_{\max,e}}, \tag{39}$$

where the minimum runs over all elements, and h_e denotes a characteristic element size for the element e and $w_{\max,e}$ denotes the maximum velocity on the element e. We refer the reader to [18] for a von Neumann stability analysis of ADER-DG schemes for a linear scalar advection equation in 1D.

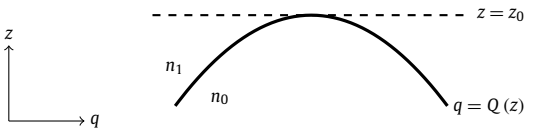


Fig. 1. Optical interface given by $q = Q(z)$ with $\left|\frac{\mathrm{d}Q}{\mathrm{d}z}\right| \to \infty$ at z_0.

(a) Element in (q, z)-space.

(b) Element in phase space.

Fig. 2. Two-dimensional cross sections of element T.

Chalmers and Krivodonova [31] have shown that the CFL condition depends on the width of a cell along the characteristic direction of flow. In our case, we have $\boldsymbol{u} - \boldsymbol{v} = (u_0 - v_0, 0)$ meaning the flow direction is along the q-axis. This means that the CFL condition (39) effectively reduces to

$$\Delta z \leq \frac{1}{2d}\frac{\mathrm{CFL}}{2N+1}\min_e \frac{\Delta q_e}{|u_0 - v_0|_{\max,e}}. \tag{40}$$

The CFL condition (40) thus states that the mesh spacing Δp has no impact on the CFL condition, something that was shown in [31]. In our numerical experiments we employ (40) to compute the maximum stable stepsize Δz.

3.3. Sub-cell interface method

The moving mesh approach described in Sections 3.1 and 3.2 has its limitations regarding the optical interfaces we can solve for. In particular, by the CFL condition (40) the maximum stable stepsize Δz_{stable} scales with $1/\left|u_{0,\max} - v_{0,\max}\right|$. If we want to align the mesh with an optical interface given by $q = Q(z)$, then at the optical interface $v_0 = \frac{\mathrm{d}Q}{\mathrm{d}z}$. Consequently, if $\left|\frac{\mathrm{d}Q}{\mathrm{d}z}\right| \to \infty$ at some z_0-value, as for example in Fig. 1, then as we approach z_0 we have to take rapidly decreasing stepsizes Δz by CFL condition (40). To resolve this problem we propose a new method, we refer to as the sub-cell interface method.

The starting point of the sub-cell interface method is to define a three-dimensional control volume $T = [z^t, z^{t+1}] \times [q_0, q_1] \times [p_0, p_1]$, where an optical interface separates T into two parts, see Fig. 2. Let the optical interface be defined by $q = Q(z)$, $q_0 \leq Q(z) \leq q_1$, then we require the optical interface to either satisfy $q_0 = Q(z^t)$ and $q_1 = Q(z^{t+1})$, or $q_1 = Q(z^t)$ and $q_0 = Q(z^{t+1})$, such that the optical interface connects diagonally opposite corners. This ensures that there is no discontinuity in the refractive index field on the interval $[q_0, q_1]$ at $z = z^t$ and $z = z^{t+1}$. Other configurations are also possible, however, they are not considered in this paper.

The phase space domain of T is cut into two pieces along the optical interface where we denote the pieces with $\Omega_0(z) = [Q(z), q_1] \times [p_0, p_1]$ and $\Omega_1(z) = [q_0, Q(z)] \times [p_0, p_1]$, see Fig. 2b. The two parts combined are denoted as $\Omega = \Omega_0(z) \cup \Omega_1(z) = [q_0, q_1] \times [p_0, p_1]$.

We start the derivation of the weak formulation by introducing the test function $\hat{\phi}_k = \hat{\phi}_k(\boldsymbol{x})$ defined on the domain Ω. The test function is related to the test function $\phi_k(\boldsymbol{\xi})$ by a transformation of Ω to the reference domain that is given by

$$\boldsymbol{x}(\boldsymbol{\xi}) = \begin{pmatrix} q(\xi) \\ p(\eta) \end{pmatrix} = \begin{pmatrix} q_0 + \xi\Delta q \\ p_0 + \eta\Delta p \end{pmatrix}, \tag{41}$$

which is just the transformation (9) for a static element, where as before $\Delta q = q_1 - q_0$ and $\Delta p = p_1 - p_0$. Hence, the test function $\hat{\phi}_k$ is formally given by $\hat{\phi}_k(\boldsymbol{x}) = \phi_k\big(\boldsymbol{\xi}(\boldsymbol{x})\big)$, where $\boldsymbol{\xi}(\boldsymbol{x})$ denotes the inverse of the transformation (41).

Next, we multiply Liouville's equation (8) with the test function $\hat{\phi}_k$ and integrate over the element Ω to obtain

$$\int_\Omega \left(\frac{\partial \rho}{\partial z} + \nabla \cdot (\rho \boldsymbol{u})\right)\hat{\phi}_k \,\mathrm{d}\mathcal{U} = 0. \tag{42}$$

Consider now the first term on the left-hand side of (42). Since ρ is discontinuous across an optical interface, we first split the domain as $\Omega = \Omega_0(z) \cup \Omega_1(z)$ and then apply Reynolds' transport theorem to each part, yielding

$$\int_\Omega \frac{\partial \rho}{\partial z}\hat{\phi}_k \,\mathrm{d}\mathcal{U} = \sum_{i=0}^{1} \int_{\Omega_i(z)} \frac{\partial \rho}{\partial z}\hat{\phi}_k \,\mathrm{d}\mathcal{U} = \frac{\mathrm{d}}{\mathrm{d}z}\int_\Omega \rho\hat{\phi}_k \,\mathrm{d}\mathcal{U} - \sum_{i=0}^{1} \int_{\partial\Omega_i(z)} \hat{\phi}_k \rho \boldsymbol{v}\cdot\hat{\boldsymbol{N}}\,\mathrm{d}\sigma \tag{43a}$$

with $\hat{\boldsymbol{N}}$ denoting once again the outward unit normal. The velocity $\boldsymbol{v}$ is defined as

$$\boldsymbol{v} = \begin{cases} \left(\frac{\mathrm{d}Q}{\mathrm{d}z}, 0\right) & \text{if } q = Q(z), \\ \boldsymbol{0} & \text{otherwise,} \end{cases} \tag{43b}$$

i.e., the velocity $\boldsymbol{v}$ is only non-zero at the optical interface.

For the second term in (42) we use once again $\Omega = \Omega_0(z) \cup \Omega_1(z)$ and apply the product rule and Gauss's theorem to both parts such that we obtain

$$\sum_{i=0}^{1} \int_{\Omega_i(z)} \nabla\cdot(\rho\boldsymbol{u})\,\hat{\phi}_k \,\mathrm{d}\mathcal{U} = \sum_{i=0}^{1} \int_{\partial\Omega_i(z)} \hat{\phi}_k \rho\boldsymbol{u}\cdot\hat{\boldsymbol{N}}\,\mathrm{d}\sigma - \int_{\Omega_i(z)} \left(\nabla\hat{\phi}_k\right)\cdot(\rho\boldsymbol{u})\,\mathrm{d}\mathcal{U}. \tag{44}$$

Substituting (43a) and (44) into equation (42) yields

$$\frac{\mathrm{d}}{\mathrm{d}z}\int_\Omega \rho\hat{\phi}_k \,\mathrm{d}\mathcal{U} = \sum_{i=0}^{1} \int_{\Omega_i(z)} \left(\nabla\hat{\phi}_k\right)\cdot(\rho\boldsymbol{u})\,\mathrm{d}\mathcal{U} - \int_{\partial\Omega_i(z)} \hat{\phi}_k \rho\,(\boldsymbol{u}-\boldsymbol{v})\cdot\hat{\boldsymbol{N}}\,\mathrm{d}\sigma. \tag{45}$$

To complete the weak formulation we integrate over the interval $[z^t, z^{t+1}]$ yielding

$$\int_\Omega \rho^{t+1}\,\hat{\phi}_k \,\mathrm{d}\mathcal{U} - \int_\Omega \rho^t\,\hat{\phi}_k \,\mathrm{d}\mathcal{U} = \mathcal{R}_\mathrm{V} - \mathcal{R}_\mathrm{S} \tag{46a}$$

with the volume term $\mathcal{R}_\mathrm{V}$ and surface term $\mathcal{R}_\mathrm{S}$ defined as

$$\mathcal{R}_\mathrm{V} = \int_{z^t}^{z^{t+1}} \sum_{i=0}^{1} \int_{\Omega_i(z)} \left(\nabla\hat{\phi}_k\right)\cdot(\rho\boldsymbol{u})\,\mathrm{d}\mathcal{U}\,\mathrm{d}z, \tag{46b}$$

$$\mathcal{R}_\mathrm{S} = \int_{z^t}^{z^{t+1}} \sum_{i=0}^{1} \int_{\partial\Omega_i(z)} \hat{\phi}_k \rho\,(\boldsymbol{u}-\boldsymbol{v})\cdot\hat{\boldsymbol{N}}\,\mathrm{d}\sigma\,\mathrm{d}z. \tag{46c}$$

Both $\boldsymbol{u}$ and $\boldsymbol{v}$ have only one non-zero component because the refractive index field is piecewise constant and because of (43b). Hence, we can write $\boldsymbol{u} = (u_0, 0)$ and $\boldsymbol{v} = (v_0, 0)$. The surface term $\mathcal{R}_\mathrm{S}$ can, therefore, be written as

$$\begin{aligned} \mathcal{R}_\mathrm{S} &= \int_{z^t}^{z^{t+1}} \left[\int_{p_0}^{p_1} \hat{\phi}_k \rho(u_0 - v_0)\mathrm{d}p \Big|_{q=Q(z)^+}^{q_1} + \int_{p_0}^{p_1} \hat{\phi}_k \rho(u_0 - v_0)\mathrm{d}p \Big|_{q=q_0}^{Q(z)^-} \right] \mathrm{d}z \\ &= \int_{z^t}^{z^{t+1}} \left[\int_{p_0}^{p_1} \hat{\phi}_k \rho u_0 \mathrm{d}p \Big|_{q=q_0}^{q_1} + \int_{p_0}^{p_1} \hat{\phi}_k \rho \left(u_0 - \frac{\mathrm{d}Q}{\mathrm{d}z}\right) \mathrm{d}p \Big|_{q=Q(z)^+}^{Q(z)^-} \right] \mathrm{d}z, \end{aligned} \tag{47}$$

where $Q(z)^\pm$ denote one-sided limits towards $Q(z)$. The simplifications also allow us to write the volume term as

$$\mathcal{R}_\mathrm{V} = \int_{z^t}^{z^{t+1}} \left[\int_{p_0}^{p_1} \int_{q_0}^{Q(z)} \frac{\partial\hat{\phi}_k}{\partial q} \rho u_0 \,\mathrm{d}q\mathrm{d}p + \int_{p_0}^{p_1} \int_{Q(z)}^{q_1} \frac{\partial\hat{\phi}_k}{\partial q} \rho u_0 \,\mathrm{d}q\mathrm{d}p \right] \mathrm{d}z. \tag{48}$$

An alternative formulation for the expressions (47) and (48) is desired, since $\left|\frac{\mathrm{d}Q}{\mathrm{d}z}\right| \to \infty$ for the optical interface shown in Fig. 1. Therefore, we make use of the (local) inverse of $Q(z)$ on the interval $[z^t, z^{t+1}]$, which we denote by $Z(q)$, i.e., $z = Q^{-1}(q) = Z(q)$. Applying a change of variables to the integrals in (47) yields

$$\mathcal{R}_S = \int_{z^t}^{z^{t+1}} \int_{p_0}^{p_1} \hat{\phi}_k \rho u_0 \mathrm{d}p \Big|_{q=q_0}^{q_1} \mathrm{d}z + \int_{q_0}^{q_1} \int_{p_0}^{p_1} \hat{\phi}_k \rho \left(u_0 \frac{\mathrm{d}Z}{\mathrm{d}q} - 1 \right) \mathrm{d}p \Big|_{z=Z(q)^-}^{Z(q)^+} \mathrm{d}q, \tag{49}$$

where $Z(q)^{\pm}$ denote one-sided limits towards $Z(q)$. In the volume term (48) we change the order of the integrals such that we obtain

$$\mathcal{R}_V = \int_{q_0}^{q_1} \left[\int_{p_0}^{p_1} \int_{z^t}^{Z(q)} \frac{\partial \hat{\phi}_k}{\partial q} \rho u_0 \, \mathrm{d}z \mathrm{d}p + \int_{p_0}^{p_1} \int_{Z(q)}^{z^{t+1}} \frac{\partial \hat{\phi}_k}{\partial q} \rho u_0 \, \mathrm{d}z \mathrm{d}p \right] \mathrm{d}q. \tag{50}$$

We refer to expressions (47)-(48) as the $Q(z)$-formulation and to expressions (49)-(50) as the $Z(q)$-formulation. Either formulation has its benefit, depending on the shape of the optical interface. Note that in the surface term $\mathcal{R}_S$ the fluxes are replaced with an upwind numerical flux, equivalent to (19), except at the optical interface. At the optical interface we compute the numerical fluxes using the methods which we will present in Section 3.4. In the results section we will describe which formulation we choose.

Making yet again use of the expansion of ρ described by (22) and inserting it into the left-hand side of equation (46a), and subsequently transforming Ω to the static reference domain χ via (41) leads to

$$\int_{\Omega} \rho^{t+1} \hat{\phi}_k \, \mathrm{d}\mathcal{U} - \int_{\Omega} \rho^t \hat{\phi}_k \, \mathrm{d}\mathcal{U} = \sum_{l=1}^{N_d} \left(\int_{\chi} \phi_l \phi_k \, \mathrm{d}\boldsymbol{\xi} \right) \mathcal{J} \left(\rho_l^{t+1} - \rho_l^t \right), \tag{51}$$

where $\mathcal{J} = \Delta q \Delta p$ is the Jacobian determinant of the transformation (41), which is a constant for this type of element. The final formulation then reads

$$\sum_{l=1}^{N_d} \left(\int_{\chi} \phi_l \phi_k \, \mathrm{d}\boldsymbol{\xi} \right) \mathcal{J} \left(\rho_l^{t+1} - \rho_l^t \right) = \mathcal{R}_V - \mathcal{R}_S, \tag{52}$$

with the two formulations for $\mathcal{R}_V$ and $\mathcal{R}_S$. All the integrals that appear in (52) are approximated by $(N+1)$-point Gauss-Legendre quadrature. To complete the scheme we need to describe how we determine ρ at intermediate levels of $[z^t, z^{t+1}]$.

The optical interface cutting the cell in two pieces makes it rather complicated to construct a Taylor expansion, since at the optical interface we have to apply the jump condition (5). Instead, we use that ρ remains invariant along a light ray, i.e., we apply (2) combined with jump condition (5). In particular, at $z = z^t$ we know the solution, hence, at a given (final) three-dimensional point $(z_f, \boldsymbol{x}_f)$ we apply the following relation

$$\rho(z_f, \boldsymbol{x}_f) = \rho(z^t, \boldsymbol{x}^t(z_f, \boldsymbol{x}_f)), \tag{53}$$

i.e., we apply the method of characteristics to determine ρ at the final point $(z_f, \boldsymbol{x}_f)$. Here, $\boldsymbol{x}^t$ describes the start of a characteristic at $z = z^t$ that ends at the point $(z_f, \boldsymbol{x}_f)$. To use (53) we need to determine the phase space point $\boldsymbol{x}^t$, by solving Hamilton's equations (1)

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}z} = \boldsymbol{u}(z, \boldsymbol{x}), \tag{54a}$$

with final condition

$$\boldsymbol{x}(z_f) = \boldsymbol{x}_f, \tag{54b}$$

with the velocity field $\boldsymbol{u}$ given by (4b). At an optical interface we have to apply the law of specular reflection or Snell's law of refraction. As mentioned in Section 2, solving Hamilton's equations and applying the laws of optics at optical interfaces is known as ray tracing [28], where in this particular case it is local ray tracing. Note that with a piecewise constant refractive index field the light rays are piecewise straight lines in the (q, z)-plane. Therefore, solving (54) reduces to computing the intersections of light rays with optical interfaces and the plane $z = z^t$.

After computing $\boldsymbol{x}^t(z_f, \boldsymbol{x}_f)$ we perform a search over the elements of the mesh to find from which element the characteristic originated. Once the correct element has been identified, the point $\boldsymbol{x}^t$ can be transformed to reference coordinates $\boldsymbol{\xi}$ on that particular element such that we can compute the value using expression (22). This process of determining the value at a new level and tracing back the characteristic, is also known as a semi-Lagrangian step. The semi-Lagrangian step is the key component in semi-Lagrangian methods such as the semi-Lagrangian discontinuous Galerkin methods presented in [32–34]. We employ the semi-Lagrangian step at every quadrature node required to evaluate the integrals in $\mathcal{R}_V$ and $\mathcal{R}_S$, which then allows us to update ρ_h using equation (52).

3.4. Optical interfaces

At an optical interface the momentum of a light ray changes discontinuously according to the law of specular reflection or Snell's law of refraction, which combined with the invariance of the basic luminance lead to the jump condition (5). The jump condition describes how phase space is connected at an optical interface. In terms of the mesh used in the DG method described thus far, an optical interface at a fixed point (z, q) is described by a collection of momentum intervals, on either side of the interface, describing the momentum domain. In the ALE-ADER-DG method we align the mesh with the optical interface and, therefore, each momentum interval corresponds to a face of an element. In the sub-cell interface method the optical interface cuts an element in two, where at the optical interface each momentum interval can be interpreted as a face. On both sides of the optical interface, we know a piecewise polynomial solution ρ at $z = z^t$ and we can compute, with either the Taylor expansions (34) or (38), or the semi-Lagrangian step (53), values on the interval $[z^t, z^{t+1}]$. Since the piecewise polynomial solution ρ is in general a discontinuous function of the momentum p, we have to be careful in how we compute the numerical flux. For example, it can happen that the flux leaving one face of the optical interface is determined by the fluxes striking multiple faces, whilst the total flux should remain the same by conservation of energy. In other words, the laws of optics cause the elements to be connected in a highly non-trivial way at the optical interface, moreover, a change in the normal of the optical interface causes a change in the connectivity of the elements.

In our previous work [7] we presented a method that incorporates the jump condition (5) in an energy-preserving manner into a DG spectral element method, for a fixed optical interface described by $q = \text{const}$. In this section, we will summarise the key components of the method and the extension to arbitrary curved optical interfaces. In particular, for arbitrary curved optical interfaces we have to separate the contributions at an optical interface into forward and backward propagating light.

3.4.1. Partitioning of momentum intervals

To facilitate the usage of the vectorial laws of reflection and refraction described by (6), we first introduce some notation to transfer from the full momentum vector (p, p_z) to p and back. We define $C_\sigma(n) = \{(p, p_z) \in \mathrm{S}^1(n) \mid \operatorname{sgn} p_z = \sigma\}$ with sgn the sign function defined as $\operatorname{sgn}(x) = 1$ if $x \geq 0$ and $\operatorname{sgn}(x) = -1$ if $x < 0$, so that $C_{\sigma=1}(n) \cup C_{\sigma=-1}(n) = \mathrm{S}^1(n)$. Given the momentum vector $(p, p_z) \in C_\sigma(n_0)$ we compute the momentum p with the mapping $P_\sigma : C_\sigma(n_0) \to [-n_0, n_0]$ as $(p, p_z) \mapsto p$. Furthermore, given a momentum p we compute its full momentum vector with the (inverse) mapping $P_\sigma^{-1} : [-n_0, n_0] \to C_\sigma(n_0)$ as $p \mapsto (p, \sigma\sqrt{n_0^2 - p^2})$.

We aim towards defining a formal definition for finding the incident light, given that we know the momenta values after reflection or refraction. For exposition purposes we will only consider reflection. Let R now be some momentum interval $R = [p_0, p_1] \subset [-n_0, n_0]$ describing light after reflection. Applying now P_σ^{-1} to R yields

$$P_\sigma^{-1}(R) = \left\{ (p, p_z) \in C_\sigma(n_0) \mid p \in R,\, p_z = \sigma\sqrt{n_0^2 - p^2} \right\}, \tag{55a}$$

and similarly for $U = P_\sigma^{-1}(R)$ we can transfer back down to R with the mapping P_σ, i.e.,

$$P_\sigma(U) = \{p \mid (p, p_z) \in U\}. \tag{55b}$$

For a physical interpretation see Fig. 3. Let now $U = P_\sigma^{-1}(R)$, so that $U \subset C_\sigma(n_0)$ describes the momentum vectors of the reflected light. Then the momentum vectors of the incident light can be found by applying $\mathcal{S}_\mathrm{R}^{-1}$ to $U = P_\sigma^{-1}(R)$, i.e.,

$$\mathcal{S}_\mathrm{R}^{-1}(P_\sigma^{-1}(R)) = \left\{ \mathcal{S}_\mathrm{R}^{-1}\big((p, p_z)\big) \mid (p, p_z) \in P_\sigma^{-1}(R) \right\}. \tag{56}$$

We can restrict $\mathcal{S}_\mathrm{R}^{-1}(P_\sigma^{-1}(R))$ to $C_\sigma(n_0)$ by computing its intersection with either $C_{\sigma=1}(n_0)$ or $C_{\sigma=-1}(n_0)$. Subsequently applying P_σ yields the momentum values on $[-n_0, n_0]$ corresponding to the incident light. This process is illustrated in Fig. 3.

Finally, the actions are combined such that we can formally find the forward ($\sigma = 1$) or backward ($\sigma = -1$) incident light by applying $\mathcal{I}$ to R, where $\mathcal{I}(R; \sigma^-, \sigma^+)$ is defined as

$$\mathcal{I}(R; \sigma^-, \sigma^+) = P_{\sigma-}\left(\mathcal{S}_\mathrm{R}^{-1}(P_{\sigma+}^{-1}(R)) \cap C_{\sigma-}(n_0) \right), \tag{57}$$

here the $-$ and $+$ are used to distinguish incident and reflected light, respectively. The result of (57) is shown in Fig. 3. Similar to (57) we define the operation also for transmission with $\mathcal{S}_\mathrm{R}^{-1}$ replaced by $\mathcal{S}_\mathrm{T}^{-1}$ with appropriate changes to the refractive indices.

With the formal definition of the incident light, we can relate the total flux for incident and outgoing light at the optical interface for the interval R with $\sigma^+ = 1$. The total flux leaving R is equal to the flux striking the intervals $\mathcal{I}(R; -1, 1)$ and $\mathcal{I}(R; 1, 1)$, describing the intervals corresponding to incident light with $\sigma^- = -1$ and $\sigma^- = 1$, respectively. This is expressed in the following energy balance

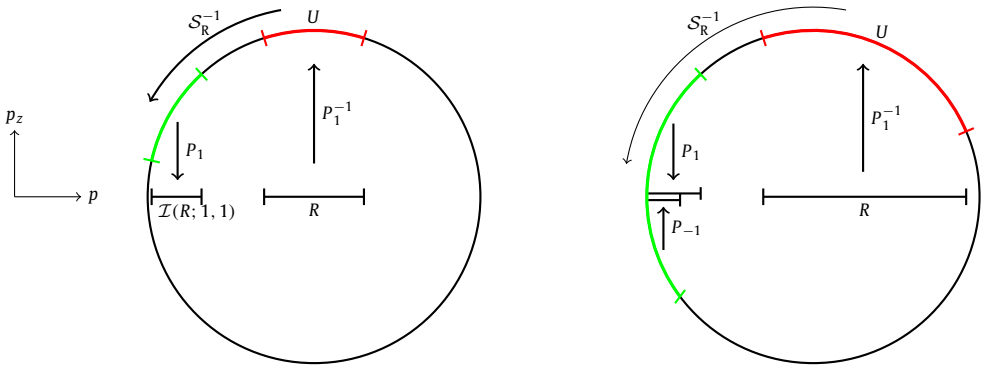


Fig. 3. Two examples where we apply the mappings $P_{\sigma-}$, $P_{\sigma+}^{-1}$, and apply $\mathcal{S}_{\mathrm{R}}^{-1}$ on the sphere $\mathrm{S}^1(n_0)$. On the left the incident light only has $\sigma = 1$, whereas on the right both forward and backward incident light contribute.

$$\int_R \rho_{\sigma=1}\left(u_0 - \frac{\mathrm{d}Q}{\mathrm{d}z}\right)\Big|_+ \mathrm{d}p = \int_{\mathcal{I}(R;-1,1)} \rho_{\sigma=-1}\left(u_0 - \frac{\mathrm{d}Q}{\mathrm{d}z}\right)\Big|_- \mathrm{d}p + \int_{\mathcal{I}(R;1,1)} \rho_{\sigma=1}\left(u_0 - \frac{\mathrm{d}Q}{\mathrm{d}z}\right)\Big|_- \mathrm{d}p, \tag{58}$$

where $\cdot|_\pm$ denotes, once more, one-sided limits towards the optical interface. A proof for the energy balance (58) is given in Appendix B. In general, forward and backward propagating light can contribute to an interval R. We separate the contributions by partitioning the interval as $R = R_0 \cup R_1$ so that $\mathcal{I}(R_0;-1,1) = \emptyset$ and $\mathcal{I}(R_1;1,1) = \emptyset$. Hence with the partitioning of R the energy balance (58) leads to the following two balances

$$\int_{R_0} \rho_{\sigma=1}\left(u_0 - \frac{\mathrm{d}Q}{\mathrm{d}z}\right)\Big|_+ \mathrm{d}p = \int_{\mathcal{I}(R_0;1,1)} \rho_{\sigma=1}\left(u_0 - \frac{\mathrm{d}Q}{\mathrm{d}z}\right)\Big|_- \mathrm{d}p, \tag{59a}$$

$$\int_{R_1} \rho_{\sigma=1}\left(u_0 - \frac{\mathrm{d}Q}{\mathrm{d}z}\right)\Big|_+ \mathrm{d}p = \int_{\mathcal{I}(R_1;-1,1)} \rho_{\sigma=-1}\left(u_0 - \frac{\mathrm{d}Q}{\mathrm{d}z}\right)\Big|_- \mathrm{d}p. \tag{59b}$$

We remark that for transmission only $\mathcal{S}_{\mathrm{R}}^{-1}$ needs to be replaced with $\mathcal{S}_{\mathrm{T}}^{-1}$ in (57).

Recall, that we only consider forward propagating light, i.e., we solve Liouville's equation (8) where $\sigma = 1$. This means that we in general do not know $\rho_{\sigma=-1}$. Depending on the optical system and initial/boundary conditions it is not necessary to solve for backward propagating light. In particular, for the examples presented in Section 4, backward propagating light does not play a role. Hence, we simply take $\rho_{\sigma=-1} = 0$ when computing (59).

3.4.2. Energy-preserving fluxes

Consider now an optical interface $q = Q(z)$ with slope $\frac{\mathrm{d}Q}{\mathrm{d}z}$ separating the media with refractive indices n_0 and n_1. An example of a geometry of the elements is sketched in Fig. 4a. The figure shows a number of faces on both sides of the optical interface. Let L_0 and L_1 be the faces where light strikes the interface, and R_0 and R_1 the faces where light leaves the interface, in other words, at L_0 and L_1 the velocity field is directed towards the optical interface while at R_0 and R_1 the velocity field is directed away from the optical interface.

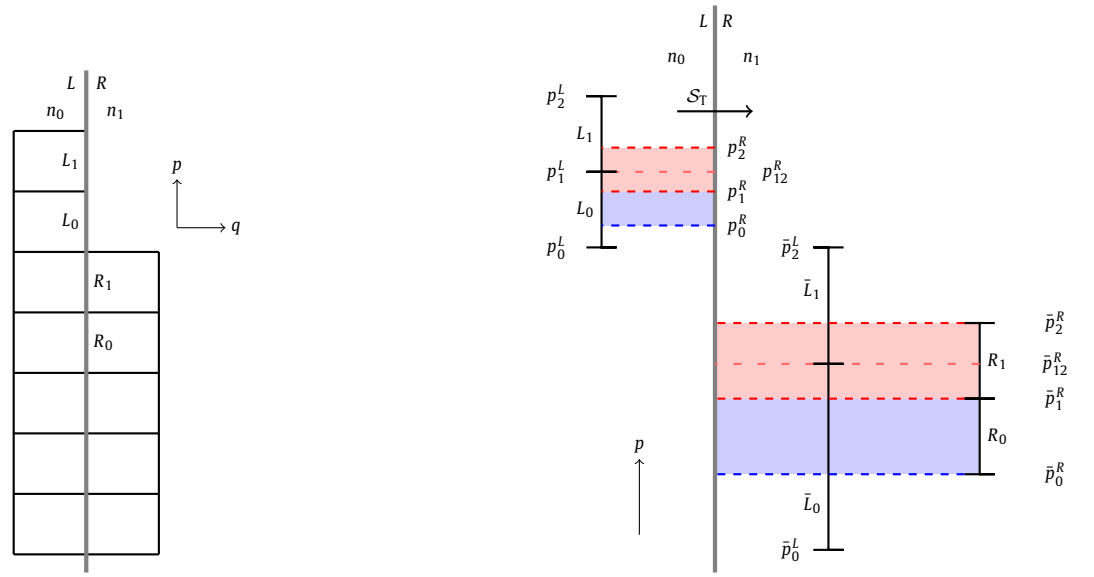
Let us now consider a face R_i $(i = 0, 1)$ and for sake of simplicity assume that all its corresponding incident light is forward propagating, see Fig. 4b. With this assumption we have $\sigma^- = 1$ and $\sigma^+ = 1$, and the energy balance for the face reads

$$\int_{R_i} \rho\left(u_0 - \frac{\mathrm{d}Q}{\mathrm{d}z}\right)\Big|_+ \mathrm{d}p = \int_{\mathcal{I}(R_i;1,1)} \rho\left(u_0 - \frac{\mathrm{d}Q}{\mathrm{d}z}\right)\Big|_- \mathrm{d}p, \tag{60}$$

where we omit the σ subscript for ρ. The energy balance (60) will be important in ensuring energy conservation in the ALE-ADER-DG scheme. The energy balance for an interface described by $z = Z(q)$, which is a (local) inverse of $q = Q(z)$, can be found by multiplying the balance (60) with $\frac{\mathrm{d}Z}{\mathrm{d}q}$ and applying $\frac{\mathrm{d}Q}{\mathrm{d}z}\frac{\mathrm{d}Z}{\mathrm{d}q} = 1$, resulting in the energy balance

$$\int_{R_i} \rho\left(u_0 \frac{\mathrm{d}Z}{\mathrm{d}q} - 1\right)\Big|_+ \mathrm{d}p = \int_{\mathcal{I}(R_i;1,1)} \rho\left(u_0 \frac{\mathrm{d}Z}{\mathrm{d}q} - 1\right)\Big|_- \mathrm{d}p. \tag{61}$$

Due to the partitioning of the momentum intervals described in the previous section we know σ for the incident and the reflected/refracted light and, therefore, introduce the shorthand notation $S(p) = S(p; n_0, n_1, \hat{\mathbf{v}})$. Here $S(p)$ simply takes



(a) Geometry in phase space

(b) Connectivity of faces

Fig. 4. Sketch of the geometry at an optical interface.

the first component of $\mathcal{S}$ (6). In particular, in the case considered here, we have $\sigma^- = 1$ and $\sigma^+ = 1$, so that the function $S(p)$ reads

$$S(p) = \begin{cases} S_R = p - 2\psi\nu_q & \text{if } \delta \le 0, \\ S_T = p - (\psi + \sqrt{\delta})\nu_q & \text{if } \delta > 0, \end{cases} \tag{62a}$$

with

$$\psi = \left(\frac{p}{\sqrt{n_0^2 - p^2}}\right) \cdot \begin{pmatrix} \nu_q \\ \nu_z \end{pmatrix} \quad \text{and} \quad \delta = n_1^2 - n_0^2 + \psi^2. \tag{62b}$$

Moreover, we will use the notation S_T to describe the q-component of $\mathcal{S}_T$ and similarly we will use S_T^{-1} to denote the q-component for refraction in reverse.

We proceed by determining for each face R_i the contributing faces. Therefore, we first transform the faces L_0 and L_1 to the other side of the optical interface by applying Snell's law of refraction $\mathcal{S}_T$ resulting in the virtual faces $\bar{L}_0$ and $\bar{L}_1$ with $\bar{L}_i = S_T(L_i)$. The virtual faces can now be related to the face R_i, see Fig. 4b.

As mentioned before, we have to be careful in how we compute the numerical flux in order to ensure we obey the energy balance (60) discretely. At a fixed point (z, q) on the optical interface we can write the solution on each face as a polynomial of the momentum p. We denote the polynomial on a face L_i by $\rho^{L_i}(p) \in \mathbb{P}_N$. Application of the jump condition (5) allows us to relate ρ on the face L_i to its counterpart on the virtual face $\bar{L}_i$ by

$$\rho^{\bar{L}_i}(\bar{p}) = \rho^{L_i}\left(S_T^{-1}(\bar{p})\right) = \rho^{L_i}(p), \quad \text{with } \bar{p} = S_T(p). \tag{63}$$

Combining relation (63) with the geometric connectivity of the faces from Fig. 4a allows us to describe how we compute the polynomial $\rho^{R_i}(p) \in \mathbb{P}_N$ for each face R_i. For example, the polynomial on face R_1 depends on ρ at the faces L_0 and L_1. The polynomial $\rho^{R_1} \in \mathbb{P}_N$ must thus be computed from a piecewise polynomial ρ^L with the additional constraint of the energy balance (60), therefore, we pose the problem as a constrained least-squares problem that reads

$$\min_{\rho^{R_1} \in \mathbb{P}_N} \int_{\bar{p}_1^R}^{\bar{p}_2^R} \left(\rho^{R_1}(\bar{p}) - \rho^L(S_T^{-1}(\bar{p}))\right)^2 \mathrm{d}\bar{p}, \tag{64a}$$

$$\text{subject to} \int_{\bar{p}_1^R}^{\bar{p}_2^R} F^{R_1}(\bar{p})\,\mathrm{d}\bar{p} = \int_{p_1^R}^{p_2^R} F^L(p)\,\mathrm{d}p, \tag{64b}$$

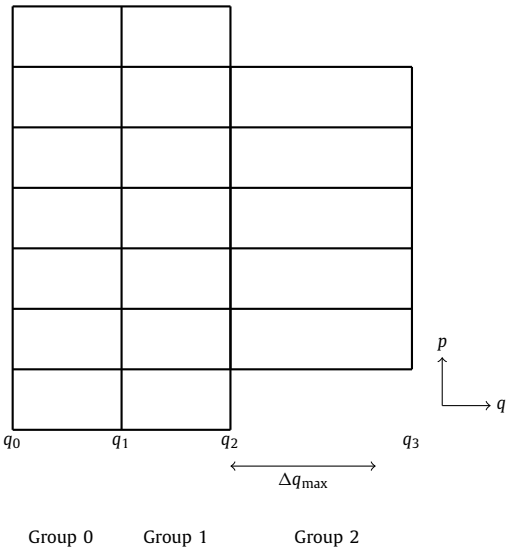


Fig. 5. Sketch of a mesh. Group 2 is a candidate for mesh refinement and will be split into two smaller groups with the mesh refinement algorithm.

where ρ^L and F^L denote piecewise polynomials given by

$$\rho^L(p) = \begin{cases} \rho^{L_0}(p) & \text{if } p \in L_0, \\ \rho^{L_1}(p) & \text{if } p \in L_1, \end{cases} \text{ and } F^L(p) = \begin{cases} F^{L_0}(p) & \text{if } p \in L_0, \\ F^{L_1}(p) & \text{if } p \in L_1. \end{cases}$$

Here, the numerical flux $F^{L_i}(p)$ is written in a basis of Lagrange polynomials, i.e.,

$$F^{L_i}(p) = \sum_{j=0}^{N} \rho_j^{L_i} a_j \ell_j(\eta(p)) \text{ with } a_j = u_0 - \frac{\mathrm{d}Q}{\mathrm{d}z}\bigg|_{(-,p_j)}, \tag{65}$$

where $\eta(p)$ denotes a transformation from the face L_i to the reference interval $[0,1]$, and $\{p_j\}_{j=0}^N$ denote the Gauss-Legendre quadrature points on the interval L_i. The numerical flux F^{R_1} is similarly written in the form (65), where the coefficients $\rho_j^{R_1}$ are the expansion coefficients of the polynomial ρ^{R_1} that are to be determined from solving equation (64).

The constrained least-squares problem (64) is now solved as described in [7]. In short, we start by writing the problem in terms of a Lagrangian with a Lagrangian multiplier and subsequently impose the requirements for a stationary point and apply $(N+1)$-point Gauss-Legendre quadrature on each (part of a) face resulting in a linear system for the $N+1$ coefficients $\rho_j^{R_1}$ and a Lagrange multiplier. The linear system is solved analytically, see [7], to obtain the $N+1$ coefficients $\rho_j^{R_1}$ on the face R_1. Finally, from these coefficients we can compute the numerical flux that is used in the ALE-ADER-DG method or the sub-cell interface method.

In a similar manner, the method can be applied when considering total internal reflection where in the equations (60), (63) and (64) refraction should be replaced by reflection, i.e., replacing S_T by S_R. Moreover, the method works for arbitrary configurations of faces at an optical interface.

3.5. Mesh refinement

In the moving mesh method small or large elements can appear. Moreover, we might want to prepare the mesh such that we can apply the sub-cell interface method and still have control over the stepsize Δz. In the description of the ALE-ADER-DG scheme we have made use of Cartesian elements. Furthermore, the mesh is only allowed to move in the q-direction and the velocity field's p-component is zero, hence, for mesh refinement we only need to consider the geometry in the q-direction. Thus, we do not consider mesh refinement along the p-direction, and interpret the mesh as a collection of q-intervals with multiple Cartesian elements per q-interval. Then, at every q-interval we collect all the elements that share the same q-interval, see Fig. 5. We refer to the collection of elements that share a q-interval as a group. The groups are sorted based on their q-values. In addition to simplifying the mesh refinement, this specific structure also simplifies the search for the correct element in the semi-Lagrangian step (53).

The mesh refinement algorithm requires a minimum mesh spacing $\Delta q_{\min}$ and a maximum mesh spacing $\Delta q_{\max} = \alpha \Delta q_{\min}$, with $\alpha > 1$. In the mesh refinement algorithm we loop over all the groups. If either of the current group and the next group have a mesh spacing Δq smaller than $\Delta q_{\min}$ and have a cumulative mesh spacing smaller than $\Delta q_{\max}$, then we combine these two groups and as a result coarsen the mesh. Otherwise, if the current group's mesh spacing is bigger than

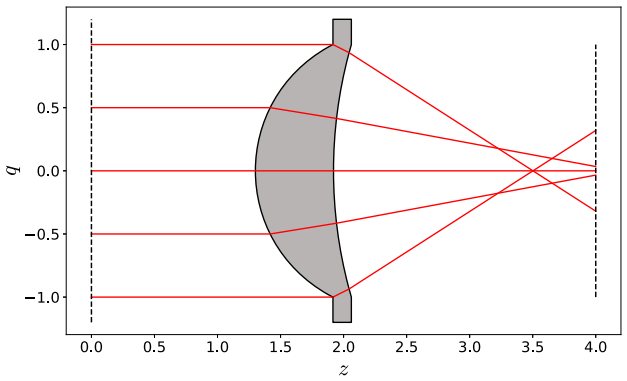


Fig. 6. Meniscus lens with a couple of light rays. The refractive index of the lens is $n_1 = 1.5$ and for the background medium $n_0 = 1$.

$\Delta q_{\max}$, then we split the current group at its midpoint into two smaller groups. Since ρ is discontinuous across an optical interface, we only allow the coarsening of groups that share the same refractive index. We iterate the mesh refinement procedure until no groups have been modified in an iteration.

The coarsening of the groups is performed by means of an L_2-projection. For example, if we consider two adjacent elements with known piecewise polynomial $w(\boldsymbol{x})$ defined on $\Omega = [q_0, q_1] \times [p_0, p_1]$, then we compute the polynomial

$$\rho(\boldsymbol{x}) = \sum_{l=1}^{N_d} \rho_l \hat{\phi}_l(\boldsymbol{x})$$

by solving

$$\int_{\Omega} (\rho(\boldsymbol{x}) - w(\boldsymbol{x}))\, \hat{\phi}_k(\boldsymbol{x})\, \mathrm{d}\mathcal{U} = 0 \quad \text{for } k = 1, 2, \ldots, N_d, \tag{66}$$

where the basis functions are once again defined by a transformation to the reference domain, such that the expansion is equivalent to (22). By solving (66) we compute the polynomial ρ that minimises the L_2-norm of $\rho - w$. The refining, or splitting, of a group is also performed by means of an L_2-projection, which in this case is equivalent to interpolating the given polynomial. We remark that solving (66) exactly will be energy-preserving, since the constant function 1 is contained in the span of the basis functions.

4. Results

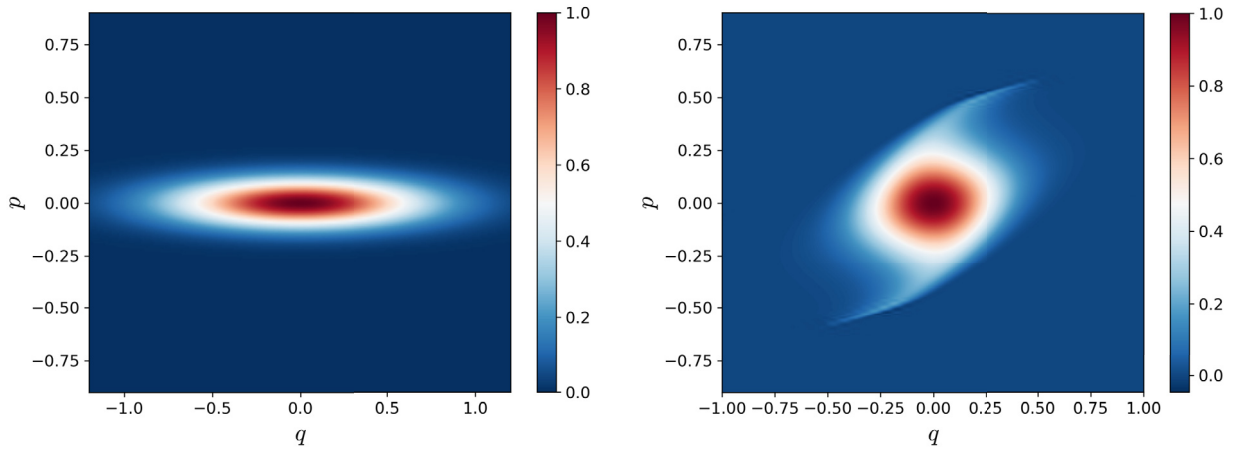
In the following we will discuss two examples, a meniscus lens and a dielectric total internal reflection concentrator. To solve Liouville's equation, we apply a few fixed settings for these problems. Namely, we take $M = N$ in the Taylor expansions (34) and (38), such that the ALE-ADER-DG scheme has a formal $(N+1)$th order accuracy in space and z. Moreover, we use the CFL condition (40) with $\mathrm{CFL} = 0.9$ fixed and we take $\alpha = 2.25$ in the mesh refinement procedure, as described in Section 3.5. In the sub-cell interface method we choose the $Z(q)$-formulation given by equations (49)-(50). The ALE-ADER-DG scheme was implemented in C++ and all the computations were performed using a single core of a laptop that has an Intel Core i7-7700HQ CPU @ 2.80 GHz.

4.1. Meniscus lens

As a first example we consider the meniscus lens, that features two spherically curved surfaces. The geometry that we consider is shown in Fig. 6. The spherical surfaces of the meniscus lens satisfy

$$q^2 + (z - z_c)^2 = R^2. \tag{67}$$

For the left surface we take $z_c = 2.42$ and $R = 1.12$, whereas for the right surface we take $z_c = 5.52$ and $R = 3.6$. For this example, the q-domain is given by the interval $[-1.2, 1.2]$ for $z \leq z_2 = 5.52 - \sqrt{3.6^2 - 1} \approx 2.06$ and $[-1, 1]$ for $z > z_2$. One can imagine the meniscus lens being fixed onto some physical system such that at $z = z_2$ the light striking $q < -1$ and $q > 1$ is fully absorbed. To numerically solve for the meniscus lens we apply the sub-cell interface method only for one single step at $z = z_c - R$ for each surface, and for the remaining curved part of the lens we apply the moving mesh method to align the mesh with the optical interface. The remaining parts of the system do not require a moving mesh, hence, we simply use a static mesh in those regions. In the moving mesh method we prescribe the mesh velocity at the optical interface by



(a) $z = 0$ (b) $z = 4$

Fig. 7. Distributions of ρ for meniscus lens with Gaussian initial condition computed with the $N = 7$ ALE-ADER-DG scheme.

writing (67) as $q = Q(z)$, such that the mesh velocity is given by $\frac{\mathrm{d}Q}{\mathrm{d}z}$ at the interface. In the sub-cell interface method the intersections of a light ray with the surface described by equation (67) are computed analytically.

To show the effects of the lens we compute a numerical solution. At $z = 0$ we start with a Gaussian distribution, given by

$$\rho_0(q, p) = \exp\left(-\frac{q^2}{2\sigma_q^2}\right)\exp\left(-\frac{p^2}{2\sigma_p^2}\right), \tag{68}$$

where we take $\sigma_q = 0.5$ and $\sigma_p = 0.08$. For this particular problem we limit the maximum momentum, since the velocity $\boldsymbol{u}$ (4b) blows up as $|p|$ approaches n, therefore, we limit the maximum momentum to $0.9n(z, q)$. Furthermore, we choose mesh spacings $\Delta q_{\max} = 0.2$ and $\Delta p = 0.09$, and use $N = 7$. Initially the mesh has 520 elements and at the end at $z = 4$ the mesh contains 320 elements. The initial condition and the numerical solution are shown in Fig. 7. From the figure, we observe that the initial condition has been compressed in the q-direction and expanded in the p-direction. Moreover, one can see values below 0 on the target distribution which is due to a cut-off of the initial distribution. The cut-off generates a discontinuity in the distribution, which appears as an oscillation resulting in undershoot in the numerical solution.

The optical interface discretisation as described in Section 3.4 should be energy-preserving for this example. Therefore, the luminous flux inside the domain plus the luminous flux leaving the domain through the physical boundaries of the system (excluding optical interfaces) should remain constant. The former is computed by integrating ρ over the phase space domain, whereas the latter is computed by adding the numerical fluxes that leave the system. We compute the absolute relative deviation from energy preservation at every step and find that the maximum deviation from energy preservation is $2.66 \cdot 10^{-15}$ and, thus, we observe energy preservation up to machine precision.

Next, we compare two strategies to apply the moving mesh method. As explained in Section 3.2, the update of a moving element is more expensive than a static element. Since we only require the moving mesh method to align optical interfaces with the mesh, we can use the freedom in the mesh velocity to optimise for better performance by using fewer moving elements. In the first strategy, which we will refer to as the global strategy, we let the mesh velocity at a certain z-position be a piecewise linear interpolant between points where we prescribe the mesh velocity. These points are the boundary of the domain where the mesh velocity is 0, and the optical interfaces where the mesh velocity is computed according to the shape of the optical interface. In the second strategy, which we call the local strategy, we only move elements adjacent to an optical interface, while the other elements remain fixed. An example of the mesh velocity for both strategies is shown in Fig. 8.

To compare both strategies we measure the computation time in the moving mesh region of the example, where for the global strategy we denote the computation time for this region with t_{global} and for the local strategy we denote the computation time with t_{local}. The speed up $t_{\text{global}}/t_{\text{local}}$ is plotted in Fig. 9 as a function of the polynomial degree N for various refinement levels r. A refinement level r indicates

$$\Delta q_{r,\max} = 2^{-r}\Delta q_{0,\max} \text{ and } \Delta p_r = 2^{-r}\Delta p_0, \tag{69}$$

where we choose $\Delta q_{0,\max} = 0.4$ and $\Delta p_0 = 0.2$ for $r = 0$. From Fig. 9 we observe that only moving the mesh elements close to the optical interfaces is significantly more efficient, especially for larger N values. Therefore, in the rest of the paper we will only use the local strategy where only elements next to the optical interface are moving.

Next, we study the convergence of the scheme. To that end, we use an initial condition so that the solution at $z = 4$ is sufficiently smooth. In particular, we take

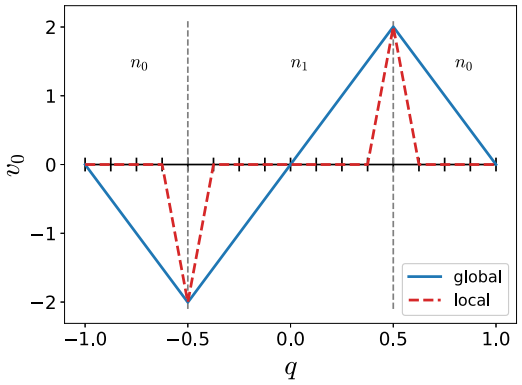


Fig. 8. Global versus local strategy of defining the mesh velocity v_0 for the meniscus lens at some z. Black intervals denote the q-intervals of elements and grey dashed lines denote optical interfaces.

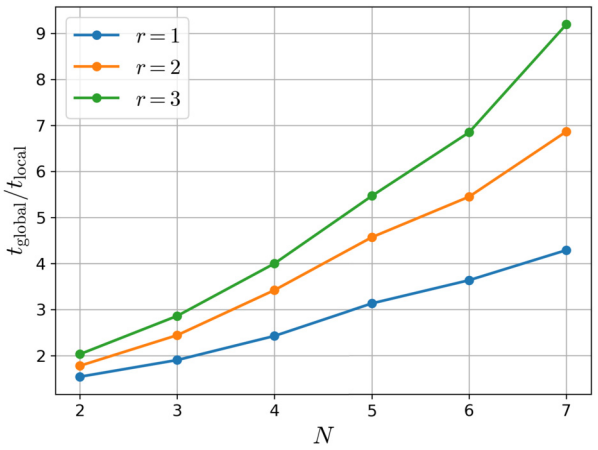


Fig. 9. Speed up $t_{\text{global}}/t_{\text{local}}$ of the moving mesh portion of the meniscus lens.

$$\rho_0(q,p) = \varphi_{m,k}\left(\frac{q}{\lambda_q}\right)\varphi_{m,k}\left(\frac{p}{\lambda_p}\right), \tag{70}$$

with parameters $\lambda_q = 0.5$ and $\lambda_p = 0.25$. Here, $\varphi_{m,k}$, with both m and k positive integers, is the function defined by

$$\varphi_{m,k}(x) = \begin{cases} \cos^{m+1}\left(\frac{\pi}{2}x^k\right) & \text{if } |x| < 1, \\ 0 & \text{otherwise,} \end{cases} \tag{71}$$

which is a C_0^m-function, meaning its first m derivatives are continuous and it has compact support. We choose $m = 10$ and $k = 2$. With the chosen initial condition, the exact solution at $z = 4$ can be obtained by tracing light rays backwards through the spherical surfaces of the lens. The convergence results for the L_2 and L_∞ norms are listed in Table 1, where the convergence rate is measured as $\log_2(e_{r-1}/e_r)$ with e_r the error for refinement level r. The computed orders of convergence are in good agreement with the expected $N+1$ order of convergence.

As a final study for this example, we compare solving Liouville's equation with the ALE-ADER-DG method to quasi-Monte Carlo ray tracing. Quasi-Monte Carlo ray tracing is typically used to determine the illuminance or luminous intensity directly, rather than to compute the basic luminance. Therefore, we will compute the illuminance using both methods. The illuminance E at a certain z-value can be computed by integrating out the momentum coordinate, i.e.,

$$E(z,q) = \int_{-n(z,q)}^{n(z,q)} \rho(z,q,p)\,\mathrm{d}p. \tag{72}$$

For the ALE-ADER-DG method we thus first have to solve Liouville's equation followed by evaluation of (72). The evaluation of (72) is performed exactly with respect to the polynomial basis of the numerical solution.

In the quasi-Monte Carlo ray tracing method we fix a number of bins B and employ a uniform grid on the target interval $q \in [-1,1]$. The jth bin is then $[Q_j, Q_{j+1}]$ with Q_j given by

Table 1
Convergence data for the meniscus lens example with the ALE-ADER-DG scheme.

r	L_2	$\mathcal{O}(L_2)$	L_∞	$\mathcal{O}(L_\infty)$	L_2	$\mathcal{O}(L_2)$	L_∞	$\mathcal{O}(L_\infty)$
		$N=2$				$N=3$		
0	6.41e-02		3.69e-01		3.86e-02		2.80e-01	
1	1.22e-02	2.40	1.10e-01	1.74	4.41e-03	3.13	4.93e-02	2.51
2	1.87e-03	2.70	2.40e-02	2.20	3.60e-04	3.61	6.24e-03	2.98
3	2.46e-04	2.93	3.83e-03	2.64	2.44e-05	3.88	4.40e-04	3.83
4	3.13e-05	2.98	5.17e-04	2.89	1.56e-06	3.97	2.91e-05	3.92
		$N=4$				$N=5$		
0	2.42e-02		1.89e-01		1.47e-02		1.45e-01	
1	1.48e-03	4.03	1.89e-02	3.32	5.08e-04	4.85	9.70e-03	3.90
2	6.77e-05	4.45	1.25e-03	3.92	1.25e-05	5.34	2.79e-04	5.12
3	2.37e-06	4.84	4.66e-05	4.75	2.27e-07	5.78	5.45e-06	5.68
4	7.60e-08	4.96	1.58e-06	4.88	3.70e-09	5.94	9.26e-08	5.88
		$N=6$				$N=7$		
0	8.81e-03		9.49e-02		5.35e-03		6.07e-02	
1	1.76e-04	5.64	3.61e-03	4.72	6.06e-05	6.47	1.38e-03	5.46
2	2.33e-06	6.24	5.91e-05	5.93	4.36e-07	7.12	1.28e-05	6.76
3	2.17e-08	6.74	6.14e-07	6.59	2.10e-09	7.70	6.28e-08	7.67
4	1.80e-10	6.92	5.08e-09	6.92	8.98e-12	7.87	2.73e-10	7.85

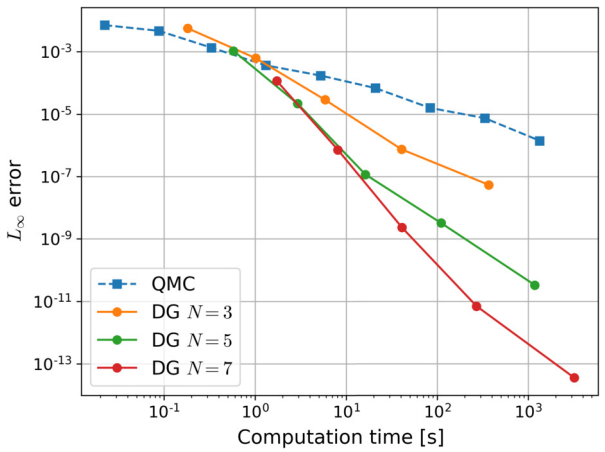


Fig. 10. Comparison between quasi-Monte Carlo (QMC) ray tracing and ALE-ADER-DG scheme (DG) for the meniscus lens.

$$Q_j = j\Delta q - 1, \quad j = 0, 1, \ldots, B, \tag{73}$$

with $\Delta q = 2/B$. The midpoint of the jth bin is given by $\frac{1}{2}(Q_j + Q_{j+1})$. The global error for quasi-Monte Carlo integration using a 2D Sobol sequence behaves as $\mathcal{O}(\log(N_{\text{RT}})^2/N_{\text{RT}})$ with N_{RT} the number of rays used [35]. The 2D Sobol sequence generates initial phase coordinates $(q_i, p_i) \in \mathcal{P}$ at $z = 0$. For more details on the Sobol sequence and quasi-Monte Carlo integration we refer the reader to [36]. For this particular example, ray tracing can compute exact intersections for each part of the meniscus lens, avoiding the need for a root-finder.

To compare the performance of both methods, we want to compute the error as the L_∞-norm of the illuminance. The quasi-Monte Carlo method computes an average illuminance on each bin, therefore, we also compute the average illuminance for the ALE-ADER-DG scheme when computing the error. Once again, we take the initial condition (70) such that we can use the exact solution to Liouville's equation to compute the exact illuminance.

The comparison of both methods is plotted in Fig. 10, where the error is plotted as a function of the computation time. For quasi-Monte Carlo ray tracing we choose $B = 200$ and vary the number of rays used. To be specific, for the first data point we use $N_{\text{RT}} = 31250$ rays and quadruple the number of rays for each subsequent point, such that at the last point we are using $N_{\text{RT}} = 2.048 \cdot 10^9$ rays. For the ALE-ADER-DG method we choose finer mesh spacings for subsequent points, see (69).

From Fig. 10 it can be seen that for a 10 second computation time the ALE-ADER-DG scheme with $N = 3$ achieves roughly 1 order of magnitude lower error compared to quasi-Monte Carlo ray tracing, and for $N = 5$ and $N = 7$ the difference in error has increased to roughly 2 orders of magnitude. Moreover, the ALE-ADER-DG scheme converges much faster than the quasi-Monte Carlo ray tracing method, in other words the ALE-ADER-DG scheme is more efficient to compute high accuracy solutions.

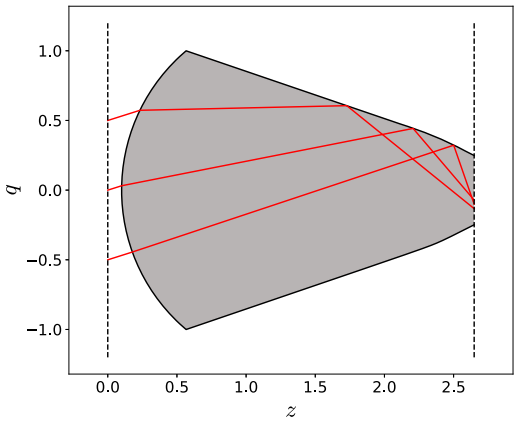


Fig. 11. A DTIRC and a couple of light rays. The gray colour represents a refractive index $n_1 = 1.5$ and the white colour the background medium with $n_0 = 1$.

Table 2
Parameters for the DTIRC.

Parameter	Value
z_c	1.405407
R	1.305407
a_1	−0.423579
b_0	−0.194090
b_1	−0.875464
b_2	0.191880
Z_{target}	2.648668

4.2. Dielectric TIR concentrator

As a second example we consider the dielectric TIR concentrator (DTIRC). The geometry that we consider is shown in Fig. 11. The optical system concentrates light that is emitted within a certain acceptance angle, from $z = 0$ towards the target in the (dielectric) medium with $n_1 = 1.5$. The rays shown in Fig. 11 are first refracted at a spherical surface, followed by reflection at one of the side walls. These side walls are designed such that the light rays satisfy the condition for total internal reflection. Details about the design process of such a system can be found in [28,37]. The spherical surface of the device is given by (67), whereas the top side wall satisfies $q = Q_{\text{top}}(z)$ with $q > 0$ and the bottom side wall is given by $q = -Q_{\text{top}}(z)$. Here $Q_{\text{top}}(z)$ reads

$$Q_{\text{top}}(z) = a_0 + a_1 z + b_0\sqrt{1 + b_1 z + b_2 z^2}, \tag{74}$$

and the target is placed at $z = Z_{\text{target}}$. The parameters for the DTIRC are listed in Table 2. The parameter a_0 is fixed by requiring that the spherical surface connects to the top side wall at $q = 1$, yielding the value $a_0 = 1.3519991422999297$.

As initial condition we use

$$\rho_0(q, p) = \varphi_{m,k}\left(\frac{q}{\lambda_q}\right)\varphi_{m,k}\left(\frac{p}{\lambda_p}\right), \tag{75}$$

with parameters $m = 10$, $k = 4$, $\lambda_q = 0.8$ and $\lambda_p = \sin\left(20\frac{\pi}{180}\right)$. Furthermore, we limit the maximum momentum to $\sin\left(85\frac{\pi}{180}\right)n(z, q)$. Then, with mesh spacings $\Delta q_{\max} = 0.1$, $\Delta p = 0.1$, and taking $N = 6$, we compute with the ALE-ADER-DG scheme the numerical solutions. The resulting distributions are shown in Fig. 12, where the initial condition and the numerical solutions at $z = \frac{1}{2}Z_{\text{target}}$ and $z = Z_{\text{target}}$ are shown. Only in the solution at $z = Z_{\text{target}}$ one can see that some of the light has been reflected. Furthermore, the light remains contained within the dielectric medium n_1 as expected.

As was done in the meniscus lens example, we will compare quasi-Monte Carlo ray tracing and the ALE-ADER-DG scheme for computing the illuminance. For this example, we modify the quasi-Monte Carlo ray tracing grid to ensure no bin cuts the side walls given by $q = \pm Q_{\text{top}}(Z_{\text{target}}) \approx \pm 0.248562$. Specifically, we modify the grid to be piecewise uniform, such that the grid spacing is uniform on the q-intervals $[-1.2, -Q_{\text{top}}(Z_{\text{target}})]$, $[-Q_{\text{top}}(Z_{\text{target}}), Q_{\text{top}}(Z_{\text{target}})]$ and $[Q_{\text{top}}(Z_{\text{target}}), 1.2]$. In quasi-Monte Carlo ray tracing we compute exact intersections with the spherical surface, whereas for intersections with the side wall we employ a Newton's method that resorts to bisection when necessary.

The resulting illuminance at $z = Z_{\text{target}}$ for both methods is shown in Fig. 13, where for QMC we use $B = 400$ bins and $N_{\text{RT}} = 8 \cdot 10^6$ rays and for the ALE-ADER-DG scheme we integrate the solution shown in Fig. 12. From the figure we observe that the solutions for both methods are almost indistinguishable by eye.

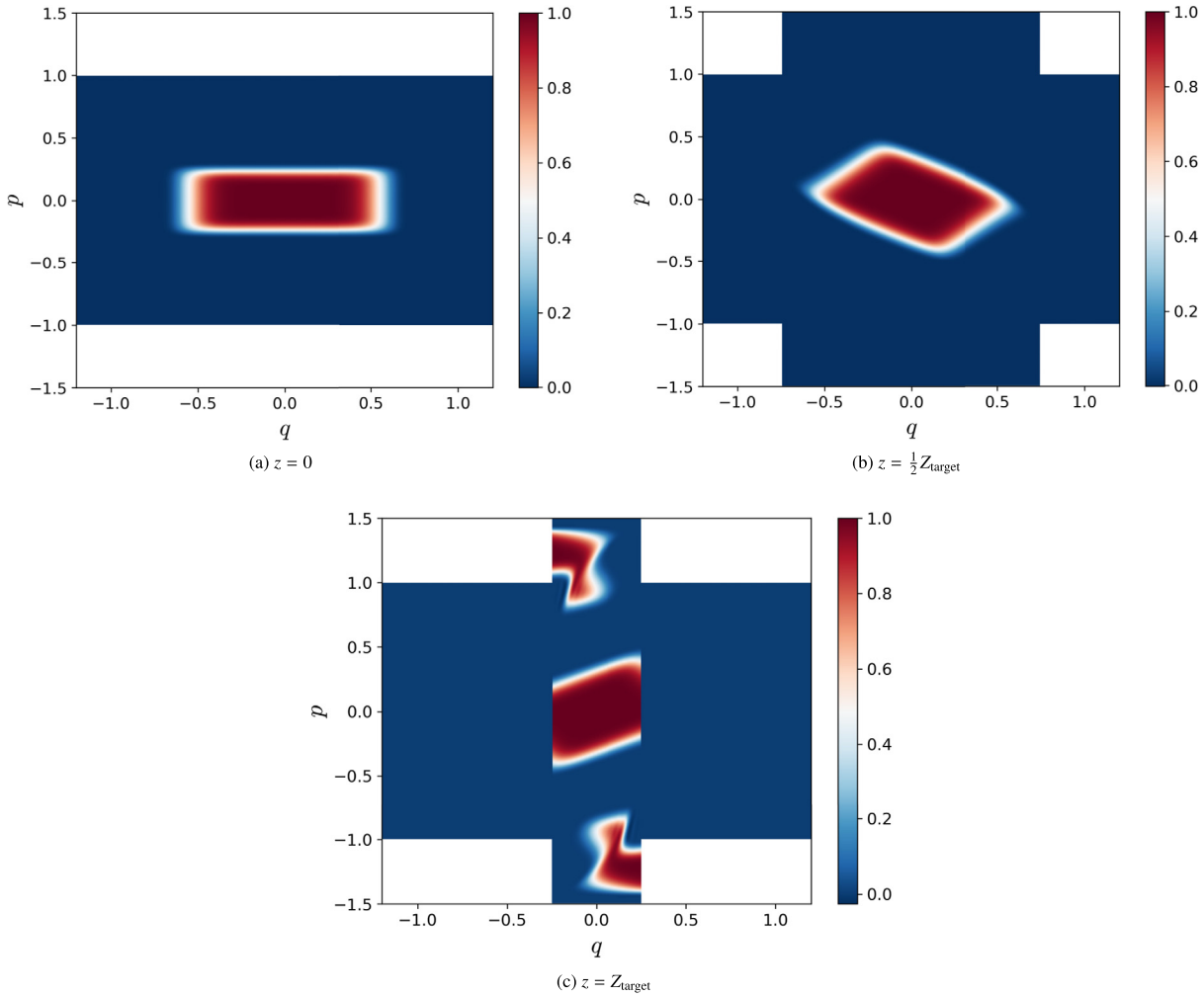


Fig. 12. Distributions of ρ for the DTIRC computed with the $N = 6$ ALE-ADER-DG scheme.

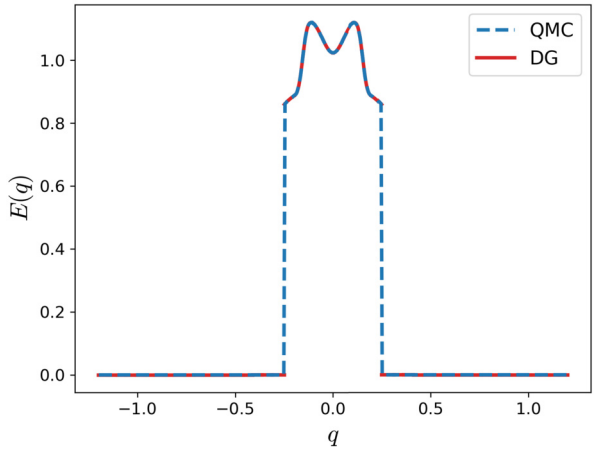


Fig. 13. Illuminance at $z = Z_{\text{target}}$ for the DTIRC computed with quasi-Monte Carlo ray tracing (QMC) on $B = 400$ bins and the $N = 6$ ALE-ADER-DG (DG) scheme.

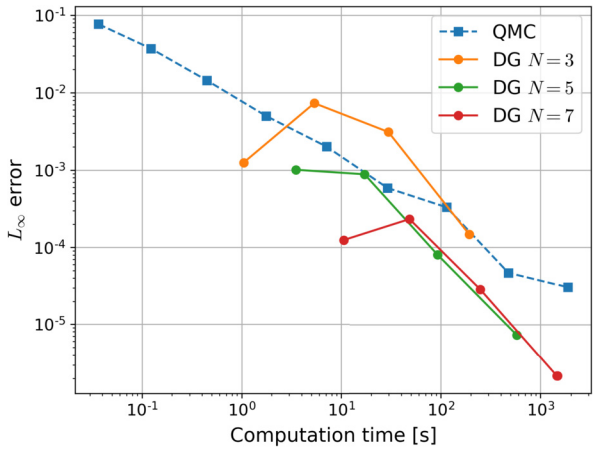


Fig. 14. Comparison between quasi-Monte Carlo (QMC) ray tracing and ALE-ADER-DG scheme (DG) for the dielectric TIR concentrator.

Next, we compare the performance of both methods where we once again compute the error as the L_∞-norm of the average illuminance. To compute the error, we use a reference solution computed with the ALE-ADER-DG scheme with $N = 7$, $\Delta q_{\max} = 0.025$ and $\Delta p = 0.0125$ ($r = 4$ in (69)). For quasi-Monte Carlo ray tracing we fix the number of bins to $B = 400$. The comparison between quasi-Monte Carlo ray tracing and the ALE-ADER-DG scheme is plotted in Fig. 14. For QMC the first data point corresponds to $N_{\text{RT}} = 31250$, whereas the last data point corresponds to $N_{\text{RT}} = 2.048 \cdot 10^9$ rays. For the ALE-ADER-DG method we choose the mesh spacings (69) with $r = 0, 1, 2, 3$. One can observe from Fig. 14, that the ALE-ADER-DG scheme with $N = 3$ for $r = 1, 2$ is beaten by quasi-Monte Carlo ray tracing in terms of accuracy whereas for $r = 3$ both methods perform similar. For $N = 5$ and $N = 7$ the ALE-ADER-DG scheme achieves better accuracy. We remark that this particular example is computationally expensive for the ALE-ADER-DG scheme since a significant portion of the mesh does not cover the solution. For example, the solution is only active in $[-Q_{\text{top}}(Z_{\text{target}}), Q_{\text{top}}(Z_{\text{target}})]$ as can be seen in Fig. 12. To reduce the computational cost in the inactive region, one could use an adaptive mesh refinement approach with a refinement criterion based on the solution, see for example [38].

5. Conclusions

We have solved Liouville's equation for two-dimensional optical systems on a moving mesh using the ALE-ADER-DG scheme. The moving mesh method was used to align the mesh with optical interfaces. The moving mesh method alone is not sufficient to solve for the considered geometries and, therefore, we introduced the sub-cell interface method. The non-local boundary conditions at optical interfaces are dealt with in an energy-preserving manner. We numerically verified that the optical interface discretisation is energy-preserving up to machine precision.

In the ALE-ADER-DG scheme an arbitrary order of accuracy for smooth solutions can be chosen both in space and the evolution coordinate z. The expected order of convergence was verified in a numerical example. Moreover, in the ADER approach we made a distinction between moving and static elements. Combined with a choice in the mesh velocity we found that letting only elements that are adjacent to an optical interface move, leads to a more efficient scheme as static elements are cheaper to update than moving elements.

The performance of the ALE-ADER-DG scheme was compared to quasi-Monte Carlo ray tracing for computing the illuminance. The numerical experiments show that the ALE-ADER-DG scheme is much more efficient than QMC ray tracing for computing high accuracy solutions. Moreover, in the first example the ALE-ADER-DG scheme achieves two orders of magnitude lower error compared to QMC ray tracing within only 10 seconds of computation time. In the second example, the ALE-ADER-DG scheme still outperforms QMC ray tracing although less pronounced. For this particular example, the ALE-ADER-DG scheme is computationally expensive since the solution is only active within a small region. Hence, an adaptive mesh refinement approach will reduce the computational cost in the inactive region.

So far we have assumed that light can only be fully reflected or fully refracted at optical interfaces. An alternative would be to incorporate Fresnel reflections [39], which would require a modification to the jump condition (5) and to the way we compute the energy-preserving fluxes across optical interfaces. This will be a topic of future research. Furthermore, we plan to extend the solver to three-dimensional optics. For three-dimensional optics we require a four-dimensional phase space, making it a computationally challenging problem. Nevertheless, the ALE-ADER-DG schemes still have the benefit of high order of convergence over QMC ray tracing.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Robert van Gestel reports financial support was provided by Dutch Research Council.

Data availability

No data was used for the research described in the article.

Acknowledgements

This work is part of the research programme NWO-TTW Perspectief with project number P15-36, which is (partly) financed by the Netherlands Organisation for Scientific Research (NWO).

Appendix A. Constant state preservation

In this section we will prove that the ALE-ADER-DG scheme described by the approximation of equation (23) together with the Taylor expansion (34) is constant state preserving when the refractive index field is constant. Since the integrals in equation (23) are computed with $(N+1)$-point Gauss-Legendre quadrature, we introduce notation to denote the usage of quadrature as follows. The application of $(N+1)$-point Gauss-Legendre quadrature to approximate the integral $\int_I g(x)\,\mathrm{d}x$ with interval $I=[a,b]$ is denoted as $\int_{I,N} g(x)\,\mathrm{d}x$ with the meaning

$$\int_{I,N} g(x)\,\mathrm{d}x = |I| \sum_{n=0}^{N} w_n g(a+|I|\,\xi_n), \tag{A.1}$$

with $|I|=b-a$, and $\{w_n\}_{n=0}^N$ and $\{\xi_n\}_{n=0}^N$ denoting the quadrature weights and points on the interval $[0,1]$. Similarly, the notation is used for multidimensional integrals where the multidimensional integral is evaluated as an iterated integral.

With the above notation equation (23) with integrals approximated by quadrature is written as

$$\sum_{l=1}^{N_d} \left(\int_{\chi,N} \phi_l \phi_k \,\mathrm{d}\boldsymbol{\xi} \right) \left((\rho_l \mathcal{J})^{t+1} - (\rho_l \mathcal{J})^t \right) = \int_{Z,N} \left(\int_{\chi,N} \left(\nabla_{\boldsymbol{\xi}} \phi_k\right) \cdot \tilde{\boldsymbol{f}} \,\mathrm{d}\boldsymbol{\xi} - \int_{\partial\chi,N} \phi_k \tilde{\boldsymbol{F}} \cdot \hat{\boldsymbol{N}} \,\mathrm{d}\sigma \right) \mathrm{d}\tau, \tag{A.2}$$

with $Z=[z^t, z^{t+1}]$. Assume now that ρ is a constant at z^t, i.e., $\rho(z^t,\boldsymbol{\xi})=c=\text{const}$, then the Taylor expansion (34) is equal to $\rho(z^t+\tau,\boldsymbol{\xi})=c$ since all spatial derivatives are zero. The numerical flux (19) is consistent, i.e., for $\rho^-=c$ and $\rho^+=c$ the numerical flux (19) yields

$$\tilde{\boldsymbol{F}}(c,c)\cdot\hat{\boldsymbol{N}} = c\left(\tilde{\boldsymbol{u}}-\tilde{\boldsymbol{v}}\right)\cdot\hat{\boldsymbol{N}}. \tag{A.3}$$

Inserting $\rho(z^t,\boldsymbol{\xi})=c$ into the left-hand side and $\rho(z^t+\tau,\boldsymbol{\xi})=c$ into the right-hand side of equation (A.2), and using (A.3) leads to

$$\sum_{l=1}^{N_d} \left(\int_{\chi,N} \phi_l \phi_k \,\mathrm{d}\boldsymbol{\xi} \right) \left((\rho_l \mathcal{J})^{t+1} - c\mathcal{J}^t \right) = c \int_{Z,N} \left(\int_{\chi,N} \left(\nabla_{\boldsymbol{\xi}} \phi_k\right) \cdot \left(\tilde{\boldsymbol{u}}-\tilde{\boldsymbol{v}}\right) \,\mathrm{d}\boldsymbol{\xi} - \int_{\partial\chi,N} \phi_k \left(\tilde{\boldsymbol{u}}-\tilde{\boldsymbol{v}}\right)\cdot\hat{\boldsymbol{N}} \,\mathrm{d}\sigma \right) \mathrm{d}\tau. \tag{A.4}$$

Next, we will prove that

$$\int_{\chi,N} \left(\nabla_{\boldsymbol{\xi}} \phi_k\right) \cdot \tilde{\boldsymbol{u}} \,\mathrm{d}\boldsymbol{\xi} - \int_{\partial\chi,N} \phi_k \tilde{\boldsymbol{u}} \cdot \hat{\boldsymbol{N}} \,\mathrm{d}\sigma = 0, \tag{A.5}$$

for a constant refractive index field $n=n_0$. Using (12) we can write $\tilde{\boldsymbol{u}}=(u_0\,\Delta p, u_1\,\Delta q)$, where for a constant refractive index field we have $u_1=0$ so that $\tilde{\boldsymbol{u}}=(u_0\,\Delta p, 0)$. Combining this with the transformation of χ to $\Omega(z)$ as described by (9), leads to

$$\tilde{\boldsymbol{u}}(\xi,\eta) = \begin{pmatrix} u_0\big(p(\eta)\big)\,\Delta p \\ 0 \end{pmatrix}. \tag{A.6}$$

Let $I=[0,1]$ denote the unit interval and recall that $\phi_k(\boldsymbol{\xi})=\ell_i(\xi)\ell_j(\eta)$, see (21), then we can write the left-hand side of (A.5) as

$$\begin{aligned} &\int_{\chi,N} \left(\nabla_{\boldsymbol{\xi}} \phi_k\right) \cdot \tilde{\boldsymbol{u}} \,\mathrm{d}\boldsymbol{\xi} - \int_{\partial\chi,N} \phi_k \tilde{\boldsymbol{u}} \cdot \hat{\boldsymbol{N}} \,\mathrm{d}\sigma \\ &\qquad = \Delta p \int_{I,N}\int_{I,N} \frac{\mathrm{d}\ell_i(\xi)}{\mathrm{d}\xi} \ell_j(\eta)\, u_0(p(\eta))\,\mathrm{d}\xi\,\mathrm{d}\eta - \Delta p\Big(\ell_i(1)-\ell_i(0)\Big) \int_{I,N} \ell_j(\eta)\, u_0(p(\eta))\,\mathrm{d}\eta \end{aligned}$$

$$= \Delta p \int_{I,N} \ell_j(\eta)\, u_0(p(\eta))\, \mathrm{d}\eta \left[\int_{I,N} \frac{\mathrm{d}\ell_i(\xi)}{\mathrm{d}\xi} \mathrm{d}\xi - (\ell_i(1) - \ell_i(0)) \right]. \tag{A.7}$$

The integral in the square brackets is evaluated exactly since $\frac{\mathrm{d}\ell_i(\xi)}{\mathrm{d}\xi}$ is a polynomial of degree $N-1$, hence,

$$\int_{I,N} \frac{\mathrm{d}\ell_i(\xi)}{\mathrm{d}\xi} \mathrm{d}\xi = \int_0^1 \frac{\mathrm{d}\ell_i(\xi)}{\mathrm{d}\xi} \mathrm{d}\xi = \ell_i(1) - \ell_i(0). \tag{A.8}$$

Therefore, we obtain

$$\int_{\chi,N} \left(\nabla_{\boldsymbol{\xi}} \phi_k\right) \cdot \tilde{\boldsymbol{u}}\, \mathrm{d}\boldsymbol{\xi} - \int_{\partial\chi,N} \phi_k \tilde{\boldsymbol{u}} \cdot \hat{\boldsymbol{N}}\, \mathrm{d}\sigma = 0, \tag{A.9}$$

proving relation (A.5).

Using relation (A.5) in (A.4) leads to

$$\sum_{l=1}^{N_d} \left(\int_{\chi,N} \phi_l \phi_k \, \mathrm{d}\boldsymbol{\xi} \right) \left((\rho_l \mathcal{J})^{t+1} - c\mathcal{J}^t \right) = c \int_{Z,N} \left(\int_{\chi,N} \left(\nabla_{\boldsymbol{\xi}} \phi_k\right) \cdot \left(-\tilde{\boldsymbol{v}}\right) \mathrm{d}\boldsymbol{\xi} - \int_{\partial\chi,N} \phi_k \left(-\tilde{\boldsymbol{v}}\right) \cdot \hat{\boldsymbol{N}}\, \mathrm{d}\sigma \right) \mathrm{d}\tau. \tag{A.10}$$

The velocity field $\boldsymbol{v} = \frac{\partial \boldsymbol{x}}{\partial \tau}$ is linear in $\boldsymbol{\xi}$, cf. (9), and since the test functions are polynomials of degree N in ξ and η the integrals over χ and $\partial\chi$ are evaluated exactly. Thus, we can replace $\int_{\chi,N}$ with $\int_{\chi}$ and subsequently apply Gauss's theorem and the product rule, such that we obtain

$$\begin{aligned} \sum_{l=1}^{N_d} \left(\int_{\chi} \phi_l \phi_k \, \mathrm{d}\boldsymbol{\xi} \right) \left((\rho_l \mathcal{J})^{t+1} - c\mathcal{J}^t \right) &= c \int_{Z,N} \left(\int_{\chi} \left(\nabla_{\boldsymbol{\xi}} \phi_k\right) \cdot \left(-\tilde{\boldsymbol{v}}\right) \mathrm{d}\boldsymbol{\xi} - \int_{\partial\chi} \phi_k \left(-\tilde{\boldsymbol{v}}\right) \cdot \hat{\boldsymbol{N}}\, \mathrm{d}\sigma \right) \mathrm{d}\tau. \\ &= c \int_{Z,N} \int_{\chi} \phi_k \nabla_{\boldsymbol{\xi}} \cdot \tilde{\boldsymbol{v}}\, \mathrm{d}\boldsymbol{\xi}\, \mathrm{d}\tau. \end{aligned} \tag{A.11}$$

Applying the geometric conservation law (14) to replace $\nabla_{\boldsymbol{\xi}} \cdot \tilde{\boldsymbol{v}}$ with $\frac{\mathrm{d}\mathcal{J}}{\mathrm{d}\tau}$ leads to

$$\sum_{l=1}^{N_d} \left(\int_{\chi} \phi_l \phi_k \, \mathrm{d}\boldsymbol{\xi} \right) \left((\rho_l \mathcal{J})^{t+1} - c\mathcal{J}^t \right) = c \left(\int_{Z,N} \frac{\mathrm{d}\mathcal{J}}{\mathrm{d}\tau} \mathrm{d}\tau \right) \int_{\chi} \phi_k \, \mathrm{d}\boldsymbol{\xi}. \tag{A.12}$$

A quick computation by applying the $(N+1)$-point Gauss-Legendre quadrature shows that

$$\begin{aligned} \sum_{l=1}^{N_d} \rho_l \int_{\chi} \phi_l \phi_k \, \mathrm{d}\boldsymbol{\xi} &= \sum_{n,m=0}^{N} \rho_{n,m} \int_{\chi} \ell_n(\xi) \ell_m(\eta) \ell_i(\xi) \ell_j(\eta)\, \mathrm{d}\xi \mathrm{d}\eta = w_i w_j \rho_{ij}, \\ \int_{\chi} \phi_k \, \mathrm{d}\boldsymbol{\xi} &= \int_{\chi} \ell_i(\xi) \ell_j(\eta)\, \mathrm{d}\xi \mathrm{d}\eta = w_i w_j, \end{aligned}$$

where in the former relation one could have alternatively applied the orthogonality of the Lagrange polynomials (20). Combining this with the fact that we integrate $\frac{\mathrm{d}\mathcal{J}}{\mathrm{d}\tau}$ numerically, see (27), as opposed to exact integration, leads to

$$w_i w_j \left(\rho_{ij}^{t+1} \mathcal{J}^{t+1} - c\mathcal{J}^t \right) = c\, w_i w_j \left(\mathcal{J}^{t+1} - \mathcal{J}^t \right), \tag{A.13}$$

from which we can directly obtain $\rho_{ij}^{t+1} = c$ showing that the scheme is constant state preserving.

Appendix B. Energy balance

In Section 3.4 we formulated the energy balance (58) that relates the fluxes across an optical interface. The total flux leaving a momentum interval R is related to the flux striking the intervals $\mathcal{I}(R;-1,\sigma)$ and $\mathcal{I}(R;1,\sigma)$ by

$$\int_R \rho_\sigma \left(u_0 - \frac{\mathrm{d}Q}{\mathrm{d}z}\right)\Big|_+ \mathrm{d}p = \int_{\mathcal{I}(R;-1,\sigma)} \rho_{\sigma=-1}\left(u_0 - \frac{\mathrm{d}Q}{\mathrm{d}z}\right)\Big|_- \mathrm{d}p + \int_{\mathcal{I}(R;1,\sigma)} \rho_{\sigma=1}\left(u_0 - \frac{\mathrm{d}Q}{\mathrm{d}z}\right)\Big|_- \mathrm{d}p, \tag{B.1}$$

with $\mathcal{I}$ defined in (57). As described in Section 3.4.1 we partition the interval R as $R = R_0 \cup R_1$, so that $\mathcal{I}(R_0;-1,\sigma) = \emptyset$ and $\mathcal{I}(R_1;1,\sigma) = \emptyset$. Henceforth, we will assume $\mathcal{I}(R;-1,\sigma) = \emptyset$, so that the balance (B.1) reduces to

$$\int_R \rho \left(u_0 - \frac{\mathrm{d}Q}{\mathrm{d}z}\right)\Big|_+ \mathrm{d}p = \int_{\mathcal{I}(R;1,\sigma)} \rho\left(u_0 - \frac{\mathrm{d}Q}{\mathrm{d}z}\right)\Big|_- \mathrm{d}p, \tag{B.2}$$

where we omit the subscript σ for ρ. Here, we will prove the balance (B.2) for reflection and refraction. Let the optical interface be given by $q = Q(z)$ and let prime $'$ denote differentiation with respect to z, i.e., $' = \frac{\mathrm{d}}{\mathrm{d}z}$, then we have the following unit normal

$$\hat{\boldsymbol{\nu}} = \begin{pmatrix} \nu_q \\ \nu_z \end{pmatrix} = \frac{\pm 1}{\sqrt{1 + Q'(z)^2}} \begin{pmatrix} -1 \\ Q'(z) \end{pmatrix}. \tag{B.3}$$

In what follows the sign of the normal is not important.

Furthermore, we will require the jump condition (5) which we will shorten to $\rho^-(p^-) = \rho^+(p^+)$ where we omit the position coordinates as they remain constant at the optical interface and omit σ for sake of simplicity, and use the shorthand notation $p^+ = S(p^-)$ where S was introduced in Section 3.4. Recall that the $\pm$ superscript denotes one-sided limits towards the optical interface, where the $-$ refers to the incident side while the $+$ denotes the outgoing side, i.e., after reflection or refraction.

We start the derivation by transforming the left-hand side of relation (B.2), by applying the jump condition and subsequently making a coordinate transformation by using $p = S(i_q)$ where i_q denotes the q-component of the incident momentum vector, i.e.,

$$\begin{aligned}\int_R \rho^+(p)\left(u_0^+(p) - Q'(z)\right)\mathrm{d}p &= \int_R \rho^-\left(S^{-1}(p)\right)\left(u_0^+(p) - Q'(z)\right)\mathrm{d}p \\ &= \int_{\mathcal{I}(R;1,\sigma)} \rho^-\left(i_q\right)\left(u_0^+\left(S(i_q)\right) - Q'(z)\right)\frac{\mathrm{d}S(i_q)}{\mathrm{d}i_q}\mathrm{d}i_q.\end{aligned} \tag{B.4}$$

From (B.2) it then follows that the following relation should hold

$$\int_{\mathcal{I}(R;1,\sigma)} \rho^-\left(i_q\right)\left(u_0^+\left(S(i_q)\right) - Q'(z)\right)\frac{\mathrm{d}S(i_q)}{\mathrm{d}i_q}\mathrm{d}i_q = \int_{\mathcal{I}(R;1,\sigma)} \rho^-\left(i_q\right)\left(u_0^-(i_q) - Q'(z)\right)\mathrm{d}i_q, \tag{B.5}$$

where upon subtracting the right-hand side from the left-hand side, we see that relation (B.5) holds for arbitrary ρ^- if

$$\left(u_0^+\left(S(i_q)\right) - Q'(z)\right)\frac{\mathrm{d}S(i_q)}{\mathrm{d}i_q} = u_0^-(i_q) - Q'(z). \tag{B.6}$$

Note that relation (B.6) must hold independently of whether we assume $\mathcal{I}(R;-1,\sigma) = \emptyset$ or $\mathcal{I}(R;1,\sigma) = \emptyset$. In other words, if we would assume $\mathcal{I}(R;1,\sigma) = \emptyset$ instead, we would still end up again with relation (B.6). We will first prove relation (B.6) for reflection and then for refraction.

Recall that the law of reflection transforms an incident momentum $\boldsymbol{i} = (i_q, i_z)$ to the reflected momentum $\boldsymbol{r} = (r_q, r_z)$, by

$$\boldsymbol{r} = \boldsymbol{i} - 2\psi\hat{\boldsymbol{\nu}} \quad \text{with } \boldsymbol{i} = \begin{pmatrix} i_q \\ \sigma\sqrt{n^2 - i_q^2} \end{pmatrix} \text{ and } \psi = \boldsymbol{i}\cdot\hat{\boldsymbol{\nu}},$$

where σ denotes the direction of the light ray and n is the refractive index of the incident and reflected light ray. The q-component of the law of reflection is denoted as $S_{\mathrm{R}}(i_q)$, see (62), and its derivative reads

$$\frac{\mathrm{d}S_{\mathrm{R}}(i_q)}{\mathrm{d}i_q} = 1 - 2\nu_q\left(\nu_q - \frac{i_q}{i_z}\nu_z\right).$$

Note that the velocity field $u_0(i_q) = i_q/i_z$, see relation (4b). Hence, we write the velocities $u_0^{\pm}$ in terms of the vectors $\boldsymbol{i}$ and $\boldsymbol{r}$, i.e.,

$$u_0^-(i_q) = \frac{i_q}{i_z} \text{ and } u_0^+(S_R(i_q)) = \frac{r_q}{r_z}.$$

Moreover, $Q'(z)$ can be written as $Q'(z) = -\nu_z/\nu_q$, cf. (B.3). Therefore, relation (B.6) can be written as

$$\left(\frac{r_q}{r_z} + \frac{\nu_z}{\nu_q}\right)\left(1 - 2\nu_q\left(\nu_q - \frac{i_q}{i_z}\nu_z\right)\right) = \frac{i_q}{i_z} + \frac{\nu_z}{\nu_q}. \tag{B.7}$$

We proceed by subtracting the right-hand side from the left-hand side so that we obtain

$$0 = \frac{r_q}{r_z} - \frac{i_q}{i_z} - 2\nu_q\left(\frac{r_q}{r_z} + \frac{\nu_z}{\nu_q}\right)\left(\nu_q - \frac{i_q}{i_z}\nu_z\right).$$

Next, we rewrite the terms to a common denominator of $r_z i_z$ as follows

$$\begin{aligned} 0 &= \frac{r_q i_z}{r_z i_z} - \frac{i_q r_z}{r_z i_z} - 2\nu_q \frac{1}{r_z}\left(r_q + r_z\frac{\nu_z}{\nu_q}\right)\frac{1}{i_z}\left(i_z\nu_q - i_q\nu_z\right) \\ &= \frac{1}{r_z i_z}\Big[r_q i_z - i_q r_z - 2\left(r_q\nu_q + r_z\nu_z\right)\left(i_z\nu_q - i_q\nu_z\right)\Big]. \end{aligned} \tag{B.8}$$

The term $r_q\nu_q + r_z\nu_z = \boldsymbol{r}\cdot\hat{\boldsymbol{\nu}} = -\psi$, which can be derived by multiplying the law of reflection with $\hat{\boldsymbol{\nu}}$. The expression (B.8) is rewritten using cross products, i.e., let $\vec{i} = (i_q, i_z, 0)$ denote the 3-vector of $\boldsymbol{i}$ etc. Moreover, let $\hat{e}_3 = (0,0,1)$, then from expression (B.8) we find

$$\begin{aligned} \frac{1}{r_z i_z}\Big[\left(\vec{r}\times\vec{i}\right)\cdot\hat{e}_3 + 2\psi\left(\vec{\nu}\times\vec{i}\right)\cdot\hat{e}_3\Big] &= \frac{1}{r_z i_z}\Big[\left(\vec{i} - 2\psi\vec{\nu}\right)\times\vec{i} + 2\psi\vec{\nu}\times\vec{i}\Big]\cdot\hat{e}_3 \\ &= \frac{1}{r_z i_z}\Big[-2\psi\vec{\nu}\times\vec{i} + 2\psi\vec{\nu}\times\vec{i}\Big]\cdot\hat{e}_3 \\ &= 0, \end{aligned}$$

completing the proof for reflection.

Now for refraction, recall that Snell's law of refraction transforms an incident momentum $\boldsymbol{i} = (i_q, i_z)$ to the transmitted/refracted momentum $\boldsymbol{t} = (t_q, t_z)$. Snell's law of refraction reads

$$\boldsymbol{t} = \boldsymbol{i} - \left(\psi + \sqrt{\delta}\right)\hat{\boldsymbol{\nu}} \quad \text{with } \delta = n_1^2 - n_0^2 + \psi^2,$$

with again $\psi = \boldsymbol{i}\cdot\hat{\boldsymbol{\nu}}$ and where n_0 and n_1 are the incident and transmitted media, respectively. The q-component of Snell's law is written as $S_T(i_q)$, see (62), and its derivative reads

$$\frac{\mathrm{d}S_T(i_q)}{\mathrm{d}i_q} = 1 - \nu_q\left(1 + \frac{\psi}{\sqrt{\delta}}\right)\left(\nu_q - \frac{i_q}{i_z}\nu_z\right).$$

As before, we can write the velocities $u_0^{\pm}$ in terms of the vectors $\boldsymbol{i}$ and $\boldsymbol{t}$, i.e.,

$$u_0^-(i_q) = \frac{i_q}{i_z} \text{ and } u_0^+(S_T(i_q)) = \frac{t_q}{t_z}.$$

Therefore, relation (B.6) can be written as

$$\left(\frac{t_q}{t_z} + \frac{\nu_z}{\nu_q}\right)\left(1 - \nu_q\left(1 + \frac{\psi}{\sqrt{\delta}}\right)\left(\nu_q - \frac{i_q}{i_z}\nu_z\right)\right) = \frac{i_q}{i_z} + \frac{\nu_z}{\nu_q}. \tag{B.9}$$

We proceed by subtracting the right-hand side from the left-hand side such that we obtain

$$0 = \frac{t_q}{t_z} - \frac{i_q}{i_z} - \nu_q\left(1 + \frac{\psi}{\sqrt{\delta}}\right)\left(\frac{t_q}{t_z} + \frac{\nu_z}{\nu_q}\right)\left(\nu_q - \frac{i_q}{i_z}\nu_z\right).$$

Next, we rewrite terms to a common denominator of $t_z i_z$ as follows

$$\begin{aligned} 0 &= \frac{t_q i_z}{t_z i_z} - \frac{i_q t_z}{t_z i_z} - \nu_q\left(1 + \frac{\psi}{\sqrt{\delta}}\right)\frac{1}{t_z}\left(t_q + t_z\frac{\nu_z}{\nu_q}\right)\frac{1}{i_z}\left(i_z\nu_q - i_q\nu_z\right) \\ &= \frac{1}{t_z i_z}\Big[t_q i_z - i_q t_z - \left(1 + \frac{\psi}{\sqrt{\delta}}\right)\left(t_q\nu_q + t_z\nu_z\right)\left(i_z\nu_q - i_q\nu_z\right)\Big]. \end{aligned} \tag{B.10}$$

The term $t_q \nu_q + t_z \nu_z = \boldsymbol{t} \cdot \hat{\boldsymbol{\nu}} = -\sqrt{\delta}$, which can be derived by multiplying Snell's law of refraction with $\hat{\boldsymbol{\nu}}$. Once again, we employ 3-vectors to write expression (B.10) as

$$\begin{aligned}\frac{1}{t_z i_z}\Big[\left(\vec{t}\times\vec{i}\right)\cdot\hat{e}_3 + \left(\sqrt{\delta}+\psi\right)\left(\vec{\nu}\times\vec{i}\right)\cdot\hat{e}_3\Big] &= \frac{1}{t_z i_z}\Big[\left(\vec{i}-\left(\psi+\sqrt{\delta}\right)\vec{\nu}\right)\times\vec{i} + \left(\psi+\sqrt{\delta}\right)\vec{\nu}\times\vec{i}\Big]\cdot\hat{e}_3 \\ &= \frac{1}{t_z i_z}\Big[-\left(\psi+\sqrt{\delta}\right)\vec{\nu}\times\vec{i} + \left(\psi+\sqrt{\delta}\right)\vec{\nu}\times\vec{i}\Big]\cdot\hat{e}_3 \\ &= 0,\end{aligned}$$

completing the proof for refraction.

Finally, we remark that for an interface given by $z = Z(q)$ relation (B.6) needs to multiplied by ν_q/ν_z to find the same result, where the normal now reads

$$\hat{\boldsymbol{\nu}} = \frac{\pm 1}{\sqrt{1+\frac{\mathrm{d}Z}{\mathrm{d}q}^2}}\begin{pmatrix}\frac{\mathrm{d}Z}{\mathrm{d}q}\\ -1\end{pmatrix}.$$

References

[1] X.-H. Lee, I. Moreno, C.-C. Sun, High-performance LED street lighting using microlens arrays, Opt. Express 21 (2013) 10612–10621.
[2] X. Zhu, Q. Zhu, H. Wu, C. Chen, Optical design of LED-based automotive headlamps, Opt. Laser Technol. 45 (2013) 262–266.
[3] A. Cvetkovic, O. Dross, J. Chaves, P. Benitez, J.C. Miñano, R. Mohedano, Etendue-preserving mixing and projection optics for high-luminance LEDs, applied to automotive headlamps, Opt. Express 14 (2006) 13014.
[4] A.M. Herkommer, Phase space optics: an alternate approach to freeform optical systems, Opt. Eng. 53 (2013) 031304.
[5] D. Rausch, M. Rommel, A.M. Herkommer, T. Talpur, Illumination design for extended sources based on phase space mapping, Opt. Eng. 56 (2017) 065103.
[6] K.B. Wolf, Geometric Optics on Phase Space, Springer Science & Business Media, 2004.
[7] R.A.M. van Gestel, M.J.H. Anthonissen, J.H.M. ten Thije Boonkkamp, W.L. IJzerman, An energy conservative hp-method for Liouville's equation of geometrical optics, J. Sci. Comput. 89 (2021) 1–35.
[8] F.D. Witherden, A.M. Farrington, P.E. Vincent, PyFR: an open source framework for solving advection–diffusion type problems on streaming architectures using the flux reconstruction approach, Comput. Phys. Commun. 185 (2014) 3028–3040.
[9] M. Dumbser, F. Fambri, M. Tavelli, M. Bader, T. Weinzierl, Efficient implementation of ADER discontinuous Galerkin schemes for a scalable hyperbolic PDE engine, Axioms 7 (2018) 63.
[10] C.A.A. Minoli, D.A. Kopriva, Discontinuous Galerkin spectral element approximations on moving meshes, J. Comput. Phys. 230 (2011) 1876–1902.
[11] D.A. Kopriva, A.R. Winters, M. Bohm, G.J. Gassner, A provably stable discontinuous Galerkin spectral element approximation for moving hexahedral meshes, Comput. Fluids 139 (2016) 148–160.
[12] V.A. Titarev, E.F. Toro, ADER: arbitrary high order Godunov approach, J. Sci. Comput. 17 (2002) 609–618.
[13] E.F. Toro, V.A. Titarev, ADER schemes for scalar non-linear hyperbolic conservation laws with source terms in three-space dimensions, J. Comput. Phys. 202 (2005) 196–215.
[14] E.F. Toro, V.A. Titarev, Derivative Riemann solvers for systems of conservation laws and ADER methods, J. Comput. Phys. 212 (2006) 150–165.
[15] J. Qiu, M. Dumbser, C.-W. Shu, The discontinuous Galerkin method with Lax–Wendroff type time discretizations, Comput. Methods Appl. Mech. Eng. 194 (2005) 4528–4543.
[16] M. Dumbser, C.-D. Munz, Building blocks for arbitrary high order discontinuous Galerkin schemes, J. Sci. Comput. 27 (2006) 215–230.
[17] M. Dumbser, C. Enaux, E.F. Toro, Finite volume schemes of very high order of accuracy for stiff hyperbolic balance laws, J. Comput. Phys. 227 (2008) 3971–4001.
[18] M. Dumbser, D.S. Balsara, E.F. Toro, C.-D. Munz, A unified framework for the construction of one-step finite volume and discontinuous Galerkin schemes on unstructured meshes, J. Comput. Phys. 227 (2008) 8209–8253.
[19] O. Zanotti, F. Fambri, M. Dumbser, Solving the relativistic magnetohydrodynamics equations with ADER discontinuous Galerkin methods, a posteriori subcell limiting and adaptive mesh refinement, Mon. Not. R. Astron. Soc. 452 (2015) 3010–3029.
[20] F. Fambri, M. Dumbser, S. Köppel, L. Rezzolla, O. Zanotti, ADER discontinuous Galerkin schemes for general-relativistic ideal magnetohydrodynamics, Mon. Not. R. Astron. Soc. 477 (2018) 4543–4564.
[21] O. Zanotti, F. Fambri, M. Dumbser, A. Hidalgo, Space–time adaptive ADER discontinuous Galerkin finite element schemes with a posteriori sub-cell finite volume limiting, Comput. Fluids 118 (2015) 204–224.
[22] G. Gassner, M. Dumbser, F. Hindenlang, C.-D. Munz, Explicit one-step time discretizations for discontinuous Galerkin and finite volume schemes based on local predictors, J. Comput. Phys. 230 (2011) 4232–4247.
[23] W. Boscheri, M. Dumbser, Arbitrary-Lagrangian–Eulerian discontinuous Galerkin schemes with a posteriori subcell finite volume limiting on moving unstructured meshes, J. Comput. Phys. 346 (2017) 449–479.
[24] E. Gaburro, W. Boscheri, S. Chiocchetti, C. Klingenberg, V. Springel, M. Dumbser, High order direct arbitrary-Lagrangian-Eulerian schemes on moving Voronoi meshes with topology changes, J. Comput. Phys. 407 (2020) 109167.
[25] J. Badwaik, P. Chandrashekar, C. Klingenberg, Single-step arbitrary Lagrangian–Eulerian discontinuous Galerkin method for 1-D Euler equations, Commun. Appl. Math. Comput. Sci. 2 (2020) 541–579.
[26] B. Owren, M. Zennaro, Derivation of efficient, continuous, explicit Runge–Kutta methods, SIAM J. Sci. Stat. Comput. 13 (1992) 1488–1501.
[27] V.I. Arnold, Mathematical Methods of Classical Mechanics, vol. 60, Springer Science & Business Media, 2013.
[28] J. Chaves, Introduction to Nonimaging Optics, CRC Press, 2017.
[29] F.E. Nicodemus Radiance, Am. J. Phys. 31 (1963) 368–377.
[30] B.S. van Lith, J.H.M. ten Thije Boonkkamp, W.L. IJzerman, T.W. Tukker, A novel scheme for Liouville's equation with a discontinuous Hamiltonian and applications to geometrical optics, J. Sci. Comput. 68 (2016) 739–771.
[31] N. Chalmers, L. Krivodonova, A robust CFL condition for the discontinuous Galerkin method on triangular meshes, J. Comput. Phys. 403 (2020) 109095.
[32] J.-M. Qiu, C.-W. Shu, Positivity preserving semi-Lagrangian discontinuous Galerkin formulation: theoretical analysis and application to the Vlasov–Poisson system, J. Comput. Phys. 230 (2011) 8386–8409.

[33] M. Restelli, L. Bonaventura, R. Sacco, A semi-Lagrangian discontinuous Galerkin method for scalar advection by incompressible flows, J. Comput. Phys. 216 (2006) 195–215.
[34] L. Einkemmer, A performance comparison of semi-Lagrangian discontinuous Galerkin and spline based Vlasov solvers in four dimensions, J. Comput. Phys. 376 (2019) 937–951.
[35] C. Filosa, Phase Space Ray Tracing for Illumination Optics, Ph.D. thesis, Eindhoven University of Technology, 2018.
[36] G. Leobacher, F. Pillichshammer, Introduction to Quasi-Monte Carlo Integration and Applications, Springer, 2014.
[37] X. Ning, R. Winston, J. O'Gallagher, Dielectric totally internally reflecting concentrators, Appl. Opt. 26 (1987) 300–305.
[38] M. Dumbser, O. Zanotti, A. Hidalgo, D.S. Balsara, ADER-WENO finite volume schemes with space–time adaptive mesh refinement, J. Comput. Phys. 248 (2013) 257–286.
[39] D.J. Griffiths, Introduction to Electrodynamics, American Association of Physics Teachers, 2005.