Computers and Mathematics with Applications 181 (2025) 84–110

Contents lists available at ScienceDirect

Computers and Mathematics with Applications

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

Analysis of difference schemes for the Fokker–Planck angular diffusion operator

Óscar López Pouso [a,b,*], Javier Segura [c]

[a] Department of Applied Mathematics, Faculty of Mathematics, University of Santiago de Compostela, Santiago de Compostela (A Coruña), Spain
[b] Galician Centre for Mathematical Research and Technology (CITMAga), Spain
[c] Departamento de Matemáticas, Estadística y Computación, Universidad de Cantabria, Santander, Spain

ARTICLE INFO

MSC:
primary 65D25
secondary 35K65, 35Q84, 65Z05, 78A35

Keywords:
Fokker–Planck angular diffusion operator
Numerical differentiation
Discrete ordinates method
Charged particles
Light propagation

ABSTRACT

This paper is dedicated to the mathematical analysis of difference schemes for discretizing the angular diffusion operator present in the azimuth–independent Fokker–Planck equation. The study establishes sets of sufficient conditions to ensure that the schemes achieve convergence of order 2, and provides insights into the rationale behind certain widely recognized discrete ordinates methods. In the process, interesting properties regarding Gaussian nodes and weights, which until now have remained unnoticed by mathematicians, naturally emerge.

1. Introduction

This paper focuses on analyzing *numerical differentiation formulas based on finite differences* that approximate the Fokker-Planck (FP) angular diffusion operator in the azimuth-independent case

$$\Delta_{\mathrm{FP}} f(\mu) = (\mathscr{D}(\mu) f'(\mu))', \quad \mu \in [-1, 1], \tag{1}$$

where $\mathscr{D}(\mu) = 1 - \mu^2$. We will use the term *difference schemes* or simply *schemes* for brevity to refer to these formulas; any alternative meaning of the word 'scheme' will be evident from the context when it arises.

Operator (1) is important because it is a fundamental part of the Fokker-Planck equation (FPE). In turn, the FPE is a forward-backward parabolic partial differential equation (PDE), highly significant in the field of nuclear engineering, in which f represents the angular flux of particles, while μ, which is the cosine of the polar angle, determines the direction of particle propagation. Interested readers can consult various references, including [2], [9], or [10], to delve deeper into this topic. The FP angular diffusion operator is also known by other names such as *continuous scattering operator*, *FP Laplacian*, *Laplacian on the unit sphere*, *spherical Laplacian*, or *Laplace-Beltrami operator*.

A commonly employed technique for solving the FPE is the use of a discrete ordinates method (DOM), which discretizes the operator (1) by utilizing a suitably selected set of nodes. Although various choices are possible, a frequently adopted approach is to use the Gauss-Legendre (GL) nodes. In this paper, DOM discretizations that use GL nodes will be referred to as *GL schemes*.

* Corresponding author.
E-mail addresses: oscar.lopez@usc.es (Ó. López Pouso), javier.segura@unican.es (J. Segura).

https://doi.org/10.1016/j.camwa.2025.01.005
Received 8 April 2024; Received in revised form 17 September 2024; Accepted 1 January 2025

Available online 9 January 2025

This work originated with the primary intention of carrying out a mathematical analysis of the GL scheme proposed by Morel in [14]. To conduct this analysis, it has been valuable for us to define two categories of schemes which we have named *type I* and *type II*. Morel's scheme belongs to the type II category, while type I schemes encompass another well-known DOM, which is once again a GL scheme: the one employed by Haldy and Ligou in [5].

The main objective is to establish the convergence of these schemes with second-order accuracy. While addressing this problem is relatively straightforward when considering uniform meshes, it becomes significantly more challenging when the nodes are not equally spaced, such as in the case of GL schemes. Type II schemes present an additional difficulty as they deviate from typical numerical differentiation formulas, using convenient approximations instead of exact values of $\mathscr{D}$.

We notice that the computing power of current PCs, together with recent research that allows the calculation of nodes and weights of GL formulas with millions of nodes in a few seconds of laptop time (see [3] and references therein), makes it possible to program GL schemes without too much cost even when the number of nodes is large.

For the purposes of this study, the term *diffusivity* will be used to refer to $\mathscr{D}$, recognizing that this decision entails some linguistic flexibility, given that $\mathscr{D}$ originates from the mathematical expression of the spherical Laplacian and does not directly represent any physical property of the medium.

Many of the ideas presented herein can also be used if $\mathscr{D}(\mu)$ is different from $1-\mu^2$, as long as it satisfies some natural conditions.

After the elementary remainder that the reader will find in Section 2, this paper is structured as follows:

- Section 3 focuses on defining the specific type of meshes considered in the paper and on setting the properties they must satisfy.
- In Section 4, we review established properties of GL nodes and weights, while also presenting novel properties discovered during the study of the schemes in this article. These additional properties play a crucial role in proving that some important schemes converge with order 2.
- Section 5 comprises two lemmas that serve as the foundation for proving the main results in subsequent sections.
- Section 6 explains the concepts of convergence of order p, full and half range mode, and preservation of moments.
- Sections 7 and 8 form the core of the paper, providing a detailed description and analysis of type I and type II schemes, and also including numerical results.
- Section 9 contains a discussion about the potential usefulness of non-convergent schemes for solving the FPE.
- Section 10 finishes the paper by summarizing the findings and drawing overall conclusions.

2. An elementary reminder

Let μ be in $(-1,1)$, and let us understand that, for a general function G and small $h>0$, $\overline{G}_s = G(\mu+sh)$.

It will be useful to keep in mind that the classical formula

$$\Delta_{\mathrm{FP}} f(\mu) \approx \frac{\overline{\mathscr{D}}_{-1/2}\overline{f}_{-1} - (\overline{\mathscr{D}}_{-1/2}+\overline{\mathscr{D}}_{1/2})\overline{f}_0 + \overline{\mathscr{D}}_{1/2}\overline{f}_1}{h^2} \tag{2}$$

can be interpreted as the outcome of repeatedly applying, with step-size $h/2$, the centered formula for the first derivative:

$$\varphi'(\mu) = \frac{\overline{\varphi}_1 - \overline{\varphi}_{-1}}{2h} + E(h). \tag{3}$$

Indeed, (2) follows from

$$\Delta_{\mathrm{FP}} f(\mu) \approx \frac{\overline{\mathscr{D}}_{1/2}\overline{f'}_{1/2} - \overline{\mathscr{D}}_{-1/2}\overline{f'}_{-1/2}}{h} \approx \frac{\overline{\mathscr{D}}_{1/2}\frac{\overline{f}_1-\overline{f}_0}{h} - \overline{\mathscr{D}}_{-1/2}\frac{\overline{f}_0-\overline{f}_{-1}}{h}}{h}. \tag{4}$$

If $\varphi \in \mathrm{C}^3([-1,1])$, the formula (3) achieves order 2, i.e., $E(h)=O(h^2)$. However, one cannot infer from this property that the formula (2) also possesses second-order accuracy. This is because the presence of h in the denominator of the last fraction in Equation (4) could make the order decay down to 1. Fortunately, this undesired effect does not occur, and the following theorem holds. The proof, which relies on Taylor expansions, is omitted here since this is a well-established result.

Theorem 1. *If $f \in \mathrm{C}^4([-1,1])$, then the numerical differentiation formula (2) has order 2, i.e.,*

$$\Delta_{\mathrm{FP}} f(\mu) = \frac{\overline{\mathscr{D}}_{-1/2}\overline{f}_{-1} - (\overline{\mathscr{D}}_{-1/2}+\overline{\mathscr{D}}_{1/2})\overline{f}_0 + \overline{\mathscr{D}}_{1/2}\overline{f}_1}{h^2} + O(h^2). \tag{5}$$

Remark 1. Theorem 1 still holds if $\mathscr{D}$ is replaced by any other diffusivity, as long as it belongs to $\mathrm{C}^3([-1,1])$.

The formula (2) can be applied at the interior points of a uniform mesh of $[-1,1]$ in a quite obvious way. Since, as said above, the GL nodes are not equally spaced, a broader framework is needed, and this will be the focus of the next sections.

3. The mesh

Considering the influence of the schemes utilized in nuclear engineering that served as a motivation for this work, we will focus exclusively on meshes comprising interior nodes. While, as exemplified in [9], it is feasible to devise schemes that incorporate -1 and 1 as nodes, the study of such cases will be deferred for future research.

Specifically, we will consider several instances of the following situation: for every natural N, we want to approximate the operator (1) on a mesh of N nodes $\mu_1^N, \dots, \mu_N^N$, located in the open interval $(-1,1)$ and not necessarily equally spaced, with the aid of an auxiliary set of $N+1$ points $\mu_{1/2}^N, \dots, \mu_{N+1/2}^N$, also not necessarily equally spaced. Note the difference in meaning between 'node' and 'point.'

The sets of nodes and points are supposed to be interlaced conforming to the following pattern:

$$-1 = \mu_{1/2}^N < \mu_1^N < \mu_{1+1/2}^N < \cdots < \mu_{N-1/2}^N < \mu_N^N < \mu_{N+1/2}^N = 1. \tag{6}$$

Definition 1. M_N and $\widetilde{M}_N$ are the numbers defined by

$$M_N = \max_{1\le n\le N-1} \{\mu_{n+1}^N - \mu_n^N\}, \tag{7}$$

$$\widetilde{M}_N = \max\{\mu_1^N + 1, M_N, 1 - \mu_N^N\}. \tag{8}$$

The minimum requirement for $[\{\mu_n^N\}_{n=1}^N,\ N \in \mathbb{N}]$ to be considered a collection of meshes of $[-1,1]$ is that

$$\lim_{N\to\infty} \widetilde{M}_N = 0, \tag{9}$$

but here a stronger assumption is needed, namely that

$$\widetilde{M}_N = O(N^{-1}), \tag{10}$$

as it happens for uniform meshes.

Remark 2. Since $\mu_1^N + 1 + \sum_{n=1}^{N-1}(\mu_{n+1}^N - \mu_n^N) + 1 - \mu_N^N = 2$ by (6), it is sure that $2 \le (N+1)\widetilde{M}_N$, which in turn implies that it is impossible to have $\widetilde{M}_N = O(N^{-p})$ with $p > 1$. However, $\widetilde{M}_N$ could potentially be a $O(N^{-p})$ with $p \in (0,1)$ if one assumes only (6) and (9).

Remark 3. According to Remark 2, with (10) we are supposing that the elements of $\{-1, \text{nodes}, 1\}$ are as close together as they can be, but this does not prevent the order 1 from being exceeded locally; for example, GL nodes satisfy (10) and accumulate quadratically at the end-points of $(-1,1)$; other examples can be furnished by applying appropriate functions to the nodes of a uniform mesh.

It is clear that (10) implies that

$$M_N = O(N^{-1}). \tag{11}$$

The scheme (2) can be easily adapted to this more general situation, and, naturally, we would like to get conditions which make the new scheme to have order 2. Recalling Section 2, one can correctly intuit in this regard that the hypotheses (6) and (10) will not be enough, because μ_n^N and $\mu_{n+1/2}^N$ are not necessarily located at the center of the cells $[\mu_{n-1/2}^N, \mu_{n+1/2}^N]$ and $[\mu_n^N, \mu_{n+1}^N]$. What may be less apparent is that these hypotheses not only fail to guarantee second-order convergence, but they are also insufficient to ensure mere convergence. Later we will prove that everything unfolds smoothly if μ_n^N and $\mu_{n+1/2}^N$ are *sufficiently close* to the mentioned central points as long as several appropriate assumptions are added to the picture.

Accordingly, we proceed by introducing a set of new conditions that build upon the existing hypotheses (6) and (10), bringing us closer to the desired objective.

Definition 2. $M_N^* = \max_{1\le n\le N}\{\mu_{n+1/2}^N - \mu_{n-1/2}^N\}$.

Since the elements of $\{-1, \text{nodes}, 1\}$ are supposed to be as close together as they can be, the hypothesis (6) implies that the same will happen to the points, that is,

$$M_N^* = O(N^{-1}). \tag{12}$$

More precisely, the following lemma holds.

Lemma 1. *Under the hypothesis (6), conditions (10) and (12) are equivalent.*

Proof. Simply notice that (10) implies (12) because $M_N^* \le 2\widetilde{M}_N$ and (12) implies (10) because $\widetilde{M}_N \le 2M_N^*$. Both inequalities are readily delivered from hypothesis (6). □

Definition 3. The set of secondary nodes $\{\hat{\mu}_n^N\}_{n=1}^N$ is defined as follows:

$$\hat{\mu}_n^N = (\mu_{n-1/2}^N + \mu_{n+1/2}^N)/2, \tag{13}$$

i.e., $\hat{\mu}_n^N$ is the midpoint of the cell $[\mu_{n-1/2}^N, \mu_{n+1/2}^N]$.

The set of secondary points $\{\hat{\mu}_{n+1/2}^N\}_{n=1}^{N-1}$ is defined as follows:

$$\hat{\mu}_{n+1/2}^N = (\mu_n^N + \mu_{n+1}^N)/2, \tag{14}$$

i.e., $\hat{\mu}_{n+1/2}^N$ is the midpoint of the cell $[\mu_n^N, \mu_{n+1}^N]$.

Definition 4. $D_N^\star = \max_{1\le n\le N} |\hat{\mu}_n^N - \mu_n^N|$.

Definition 5. $D_N = \max_{1\le n\le N-1} |\hat{\mu}_{n+1/2}^N - \mu_{n+1/2}^N|$.

The following result holds.

Lemma 2. *Under the hypotheses (6) and (10), there exist $q \ge 1$ and $r \ge 1$ such that*

$$D_N^\star = O(N^{-q}), \tag{15}$$

$$D_N = O(N^{-r}). \tag{16}$$

Proof. Due to (6), it is sure that $\hat{\mu}_n^N, \mu_n^N \in (\mu_{n-1/2}^N, \mu_{n+1/2}^N)$ for $n = 1, \dots, N$. So,

$$\max_{1\le n\le N} |\hat{\mu}_n^N - \mu_n^N| \le M_N^*. \tag{17}$$

Now (15) is implied by (12).

Analogously, (16) is implied by the inequality

$$\max_{1\le n\le N-1} |\hat{\mu}_{n+1/2}^N - \mu_{n+1/2}^N| \le M_N \tag{18}$$

and (11). □

However, $D_N^\star = O(N^{-1})$ and $D_N = O(N^{-1})$ are not enough for ensuring quadratic convergence. To achieve this goal, we will make the assumption that both q and r in (15) and (16) are not less than 2:

$$D_N^\star = O(N^{-q}) \text{ with } q \ge 2, \tag{19}$$

$$D_N = O(N^{-r}) \text{ with } r \ge 2. \tag{20}$$

Definition 6. $m_N^* = \min_{1\le n\le N} \{\mu_{n+1/2}^N - \mu_{n-1/2}^N\}$.

The hypotheses that we have enunciated so far are necessary to have convergence of order 2. On the contrary, there are signs that the one that comes now could be weakened if f were regular enough. It is not very restrictive though, and simplifies the proofs that will come later. Specifically, it will be assumed that

$$\frac{1}{m_N^*} = O(N^s) \text{ with } 1 \le s \le 4m-2, \text{ where } m = \min\{q, r\}. \tag{21}$$

Remark 4. Notice that $s < 1$ is impossible because the trivial equality $\sum_{n=1}^N (\mu_{n+1/2}^N - \mu_{n-1/2}^N) = 2$ implies that $1/m_N^* \ge N/2$. The upper bound $4m-2$ prevents m_N^* from decreasing too fast, but the rate of decrease could still be considerably high, since $4m-2 \ge 6$. This is why we say above that this hypothesis is not very restrictive.

The following lemma will be useful. Its proof is simple from Definition 3 and is omitted.

Lemma 3. *For $n = 2, \dots, N-1$,*

$$\frac{\mu_{n-1}^N + \mu_{n+1}^N}{2} - \mu_n^N = (\hat{\mu}_{n-1/2}^N - \mu_{n-1/2}^N) + (\hat{\mu}_{n+1/2}^N - \mu_{n+1/2}^N) + 2(\hat{\mu}_n^N - \mu_n^N). \tag{22}$$

Hence, under the hypotheses (19) and (20),

$$\max_{2\le n\le N-1}\left|\frac{\mu_{n-1}^N+\mu_{n+1}^N}{2}-\mu_n^N\right| = O(N^{-m}),\ \textit{with}\ m=\min\{q,r\}\ge 2. \tag{23}$$

4. Properties of GL nodes and weights

Here we collect a brief list of facts about GL quadrature on $(-1,1)$ that will be needed later. Extending the results from this section to a general interval (a,b) is straightforward, which is relevant in this paper as similar considerations will be necessary for the intervals $(-1,0)$ and $(0,1)$.

Symmetry of weights and antisymmetry of nodes with respect to 0 are assumed to be known. Whenever GL nodes are mentioned, it must be understood that they are arranged in increasing order.

The following result expresses in a formal way what was said about GL nodes in Remark 3.

Proposition 1. *If $\{\mu_n^N\}_{n=1}^N$ are the GL nodes, then the following assertions, where the exponents 1 and 2 are optimal, hold:*

(A) $\widetilde{M}_N = O(N^{-1})$, *that is, hypothesis (10) is met.*
(B) *For any fixed natural k,*

$$0<\mu_1^N+1<\mu_2^N-\mu_1^N<\cdots<\mu_k^N-\mu_{k-1}^N \quad \textit{if} \quad N\ge 2k \tag{24}$$

and

$$\mu_k^N+1 = 1-\mu_{N-k+1}^N = O(N^{-2}). \tag{25}$$

Remark 5 *(Meaning of 'optimal exponent').* An equivalent way of saying that the exponent 1 is optimal in the expression $\widetilde{M}_N = O(N^{-1})$ is to say that $\widetilde{M}_N=\Theta(N^{-1})$ ('Big Theta' of N^{-1}). Similarly, $\mu_k^N+1=\Theta(N^{-2})$.

Proposition 2. *Let $\{w_n^N\}_{n=1}^N$ be the set of GL weights and let k be any fixed natural number. Then*

$$0<w_1^N<w_2^N<\cdots<w_k^N \quad \textit{if} \quad N\ge 2k \tag{26}$$

and

$$w_k^N = O(N^{-2}), \tag{27}$$

being the exponent 2 optimal.

Therefore,

$$\frac{1}{\min_{1\le n\le N} w_n^N} = \max_{1\le n\le N}\frac{1}{w_n^N} = \frac{1}{w_1^N} = \frac{1}{w_N^N} = O(N^2). \tag{28}$$

Other properties of GL nodes and weights. The statements in Propositions 1 and 2 are established facts (proofs can be derived from results in [18]), but, as far as we know, the properties that follow are new. We have become aware of them since they are inherent to schemes used in nuclear engineering such as Haldy-Ligou's or Morel's, to be described later. Although a complete mathematical proof is unavailable, the properties' validity is strongly supported by the theoretical results in reference [11]. Furthermore, they have been corroborated by numerical experiments.

New properties. Let us suppose that $\{\mu_n^N\}_{n=1}^N$ and $\{w_n^N\}_{n=1}^N$ are, respectively, the GL nodes and weights, and that the points $\{\mu_{n+1/2}^N\}_{n=0}^N$ are defined by

$$\mu_{1/2}^N = -1, \tag{29}$$
$$\mu_{n+1/2}^N = \mu_{n-1/2}^N + w_n^N \ \text{for}\ n=1,\dots,N. \tag{30}$$

Then

- The hypothesis (6) holds, and
- The hypotheses (19) and (20) are met with $q=r=2$, that is, $D_N^*=O(N^{-2})$ and $D_N=O(N^{-2})$. Accordingly, by Lemma 3,

$$\max_{2\le n\le N-1}\left|\frac{\mu_{n-1}^N+\mu_{n+1}^N}{2}-\mu_n^N\right| = O(N^{-2}). \tag{31}$$

For the sake of ease, the superscript N will be omitted in what follows.

5. The underlying formulas

Definition 7. $\mathbb{P}_k$, with $k \in \mathbb{N}$, will be the real vector space of all polynomials with real coefficients having degree less than or equal to k.

Definition 8 *(quantities of interest related to cell* $[\mu_n, \mu_{n+1}]$*).* For $n = 1, \dots, N-1$:

$$h_n = (\mu_{n+1} - \mu_n)/2, \tag{32}$$

$$d_n = \hat{\mu}_{n+1/2} - \mu_{n+1/2}, \tag{33}$$

$$h_{n-} = \mu_{n+1/2} - \mu_n, \tag{34}$$

$$h_{n+} = \mu_{n+1} - \mu_{n+1/2}. \tag{35}$$

Definition 9 *(quantities of interest related to cell* $[\mu_{n-1/2}, \mu_{n+1/2}]$*).* For $n = 1, \dots, N$:

$$h_n^* = (\mu_{n+1/2} - \mu_{n-1/2})/2, \tag{36}$$

$$d_n^* = \hat{\mu}_n - \mu_n, \tag{37}$$

$$h_{n-}^* = \mu_n - \mu_{n-1/2}, \tag{38}$$

$$h_{n+}^* = \mu_{n+1/2} - \mu_n. \tag{39}$$

Remark 6. It is obvious that $h_{n-}^* = h_{(n-1)+}$ if $n \in \{2, \dots, N\}$, and that $h_{n+}^* = h_{n-}$ if $n \in \{1, \dots, N-1\}$.

Remark 7. Due to (6), h_n, h_{n-}, h_{n+}, h_n^*, h_{n-}^* and h_{n+}^* are always positive. On the other hand, d_n and d_n^* can be positive, negative, or zero.

The above Definitions 8 and 9 imply that, for $n = 1, \dots, N-1$,

$$h_{n-} = h_n - d_n, \tag{40}$$

$$h_{n+} = h_n + d_n, \tag{41}$$

$$h_{n-} + h_{n+} = 2h_n = \mu_{n+1} - \mu_n, \tag{42}$$

$$h_{n+} - h_{n-} = 2d_n, \tag{43}$$

$$d_n + d_n^* = h_n - h_n^*, \tag{44}$$

for $n = 2, \dots, N$,

$$d_{n-1} + d_n^* = h_n^* - h_{n-1}, \tag{45}$$

for $n = 2, \dots, N-1$,

$$d_{n-1} - d_n = 2h_n^* - (h_{n-1} + h_n), \tag{46}$$

and, for $n = 1, \dots, N$,

$$h_{n-}^* = h_n^* - d_n^*, \tag{47}$$

$$h_{n+}^* = h_n^* + d_n^*, \tag{48}$$

$$h_{n-}^* + h_{n+}^* = 2h_n^* = \mu_{n+1/2} - \mu_{n-1/2}, \tag{49}$$

$$h_{n+}^* - h_{n-}^* = 2d_n^*. \tag{50}$$

Besides,

$$M_N = 2 \max_{1 \le n \le N-1} h_n, \tag{51}$$

$$M_N^* = 2 \max_{1 \le n \le N} h_n^*, \tag{52}$$

$$D_N = \max_{1 \le n \le N-1} |d_n|, \tag{53}$$

$$D_N^* = \max_{1\le n\le N} |d_n^*|, \tag{54}$$

$$m_N^* = 2 \min_{1\le n\le N} h_n^*. \tag{55}$$

In light of Section 2, we will exploit the following two lemmas. We will use the notation $\|\psi\|_\infty = \max_{\mu\in[-1,1]} |\psi(\mu)|$, understanding that $\psi \in \mathrm{C}([-1,1])$. Also, the notations ξ_{n-}^*, ξ_{n+}^*, ξ_{n-}, ξ_{n+} will stand for intermediate values appearing in the Lagrange form of the Taylor remainder.

Lemma 4. *Assume that the hypothesis (6) holds.*

The approximation

$$\varphi'(\mu_n) \approx \frac{\varphi(\mu_{n+1/2}) - \varphi(\mu_{n-1/2})}{\mu_{n+1/2} - \mu_{n-1/2}}, \ n = 1,\dots,N, \tag{56}$$

converges with order 2 *if, and only if, the hypotheses (10) and (19) are met. More precisely, if $E_n^*(\varphi)$ is defined by*

$$E_n^*(\varphi) = \varphi'(\mu_n) - \frac{\varphi(\mu_{n+1/2}) - \varphi(\mu_{n-1/2})}{\mu_{n+1/2} - \mu_{n-1/2}}, \ n = 1,\dots,N, \tag{57}$$

then

$$\max_{1\le n\le N} |E_n^*(\varphi)| = O(N^{-2}) \text{ for all } \varphi \in \mathrm{C}^3([-1,1]) \tag{58}$$

if, and only if, the hypotheses (10) and (19) are met.

The maximal possible order is 2.

Moreover, the formula (56) is exact if $\varphi \in \mathbb{P}_1$ or if $[D_N^* = 0$ and $\varphi \in \mathbb{P}_2]$.

Proof. That the formula (56) is exact on $\mathbb{P}_1$ is a triviality, although this fact will also be deduced, along with the rest of the conclusions, from the reasoning that follows.

We will write E_n^* instead of $E_n^*(\varphi)$. Recall that, under (6), conditions (10) and (12) are equivalent by Lemma 1.

Take $\varphi \in \mathrm{C}^3([-1,1])$ and $n \in \{1,\dots,N\}$, and consider the Taylor expansions

$$\varphi(\mu_{n+1/2}) = \varphi(\mu_n) + h_{n+}^*\varphi'(\mu_n) + \frac{(h_{n+}^*)^2}{2}\varphi''(\mu_n) + \frac{(h_{n+}^*)^3}{6}\varphi'''(\xi_{n+}^*), \tag{59}$$

$$\varphi(\mu_{n-1/2}) = \varphi(\mu_n) - h_{n-}^*\varphi'(\mu_n) + \frac{(h_{n-}^*)^2}{2}\varphi''(\mu_n) - \frac{(h_{n-}^*)^3}{6}\varphi'''(\xi_{n-}^*). \tag{60}$$

Subtracting (59) and (60) and dividing the result by $\mu_{n+1/2} - \mu_{n-1/2} = h_{n-}^* + h_{n+}^*$, we have

$$-E_n^* = \frac{\varphi(\mu_{n+1/2}) - \varphi(\mu_{n-1/2})}{\mu_{n+1/2} - \mu_{n-1/2}} - \varphi'(\mu_n) = \frac{(h_{n+}^*)^2 - (h_{n-}^*)^2}{2(h_{n-}^* + h_{n+}^*)}\varphi''(\mu_n) + \frac{(h_{n+}^*)^3\varphi'''(\xi_{n+}^*) + (h_{n-}^*)^3\varphi'''(\xi_{n-}^*)}{6(h_{n-}^* + h_{n+}^*)}, \tag{61}$$

or, taking account of

$$\frac{(h_{n+}^*)^2 - (h_{n-}^*)^2}{2(h_{n-}^* + h_{n+}^*)} = \frac{h_{n+}^* - h_{n-}^*}{2} = d_n^*, \tag{62}$$

$$E_n^* = -d_n^*\varphi''(\mu_n) - \frac{(h_{n+}^*)^3\varphi'''(\xi_{n+}^*) + (h_{n-}^*)^3\varphi'''(\xi_{n-}^*)}{6(h_{n-}^* + h_{n+}^*)}. \tag{63}$$

Now, since h_{n-}^* and h_{n+}^* are positive due to (6) and

$$\begin{aligned}
\frac{(h_{n+}^*)^3 + (h_{n-}^*)^3}{6(h_{n-}^* + h_{n+}^*)} &= \frac{(h_{n+}^*)^2 - h_{n+}^*h_{n-}^* + (h_{n-}^*)^2}{6} \\
&= \frac{(h_n^* + d_n^*)^2 - (h_n^* + d_n^*)(h_n^* - d_n^*) + (h_n^* - d_n^*)^2}{6} \\
&= \frac{(h_n^*)^2 + 3(d_n^*)^2}{6} \le \frac{(M_N^*/2)^2 + 3(D_N^*)^2}{6} = \frac{(M_N^*)^2 + 12(D_N^*)^2}{24},
\end{aligned} \tag{64}$$

we get from Equation (63) the following inequality:

$$\max_{1\le n\le N} |E_n^*| \le D_N^*\|\varphi''\|_\infty + \frac{(M_N^*)^2 + 12(D_N^*)^2}{24}\|\varphi'''\|_\infty. \tag{65}$$

The ‘if part’ is a consequence of (65), (12), and (19). Equation (65) also implies that the formula (56) is exact if $\varphi \in \mathbb{P}_1$ or if $[D_N^* = 0$ and $\varphi \in \mathbb{P}_2]$.

The 'only if part' can be proved in two steps:

Step 1 If the hypothesis (19) does not hold, that is, if $D^*_N \neq O(N^{-2})$, then $\max_{1\le n\le N}|E^*_n| \neq O(N^{-2})$ for certain $\varphi \in C^3([-1,1])$. Indeed, if one takes $\varphi(\mu)=\mu^2$, then $E^*_n = -2d^*_n$ by (63), and hence $\max_{1\le n\le N}|E^*_n| = 2D^*_N \neq O(N^{-2})$.

Step 2 If the hypothesis (19) holds but the hypothesis (10) does not hold, then $\max_{1\le n\le N}|E^*_n| \neq O(N^{-2})$ for certain $\varphi\in C^3([-1,1])$. To see this, let us take $\varphi(\mu)=\mu^3$. Then, by (63) and calculations within (64), $E^*_n = -(h^*_n)^2 - 3(d^*_n)^2 - 6\mu_n d^*_n$. Now we will prove that $\max_{1\le n\le N}|E^*_n| \neq O(N^{-2})$. Notice that $(h^*_n)^2 + 3(d^*_n)^2 - 6|\mu_n d^*_n| \le |E^*_n|$, and hence, for $n=1,\dots,N$,

$$(h^*_n)^2 \le |E^*_n| + 6|\mu_n d^*_n| - 3(d^*_n)^2 \le |E^*_n| + 6|\mu_n d^*_n| \le \max_{1\le n\le N}|E^*_n| + 6D^*_N, \tag{66}$$

from where

$$(M^*_N)^2 \le 4\max_{1\le n\le N}|E^*_n| + 24D^*_N. \tag{67}$$

So, M^*_N would be a $O(N^{-1})$, i.e., the hypothesis (10) would be satisfied, if $\max_{1\le n\le N}|E^*_n|$ were a $O(N^{-2})$. This ends the proof of Step 2.

The examples above are also useful to demonstrate that, whether hypotheses (10) and (19) are met or not, the order 2 cannot be improved:

- If $D^*_N \neq O(N^{-q})$ for all $q>2$, then the example given by $\varphi(\mu)=\mu^2$ shows that $\max_{1\le n\le N}|E^*_n| = 2D^*_N$ is of the same order than D^*_N, so less than or equal to 2.
- If $D^*_N = O(N^{-q})$ for some $q>2$, then the example given by $\varphi(\mu)=\mu^3$ shows, using (67), that $\max_{1\le n\le N}|E^*_n| \ge (M^*_N/2)^2 - 6D^*_N \ge N^{-2} - 6D^*_N$, and so $\max_{1\le n\le N}|E^*_n|$ is again at most of order 2. The inequality $M^*_N/2 \ge 1/N$ follows from $2=\sum_{n=1}^N(\mu_{n+1/2} - \mu_{n-1/2}) \le N M^*_N$.

This ends the proof of Lemma 4. □

The following result is analogous to Lemma 4, but contains a finer expression of the error term that will be needed later.

Lemma 5. *Assume that the hypothesis (6) holds.*

(A) *The approximation*

$$\varphi'(\mu_{n+1/2}) \approx \frac{\varphi(\mu_{n+1})-\varphi(\mu_n)}{\mu_{n+1}-\mu_n},\quad n=1,\dots,N-1, \tag{68}$$

converges with order 2 *if, and only if, the hypotheses (10) and (20) are met. More precisely, if* $E_n(\varphi)$ *is defined by*

$$E_n(\varphi) = \varphi'(\mu_{n+1/2}) - \frac{\varphi(\mu_{n+1})-\varphi(\mu_n)}{\mu_{n+1}-\mu_n},\quad n=1,\dots,N-1, \tag{69}$$

then

$$\max_{1\le n\le N-1}|E_n(\varphi)| = O(N^{-2}) \text{ for all } \varphi\in C^3([-1,1]) \tag{70}$$

if, and only if, the hypotheses (10) and (20) are met.
The maximal possible order is 2.
Moreover, the formula (68) is exact if $\varphi\in\mathbb{P}_1$ *or if [*$D_N=0$ *and* $\varphi\in\mathbb{P}_2$*].*

(B) *If* $\varphi\in C^5([-1,1])$*, then, for* $n=1,\dots,N-1$*, there exist certain* $\xi_{n-}\in(\mu_n,\mu_{n+1/2})$ *and* $\xi_{n+}\in(\mu_{n+1/2},\mu_{n+1})$ *such that*

$$E_n(\varphi) = -d_n\varphi''(\mu_{n+1/2}) - \frac{h_n^2+3d_n^2}{6}\varphi'''(\mu_{n+1/2}) - \frac{h_n^2 d_n + d_n^3}{6}\varphi^{4)}(\mu_{n+1/2}) - \frac{h_{n+}^5\varphi^{5)}(\xi_{n+}) + h_{n-}^5\varphi^{5)}(\xi_{n-})}{120(h_{n-}+h_{n+})}. \tag{71}$$

Proof. The proof of (A) is like that of Lemma 4. Let us prove (B).
Subtracting the Taylor expansions

$$\varphi(\mu_{n+1}) = \varphi(\mu_{n+1/2}) + h_{n+}\varphi'(\mu_{n+1/2}) + \frac{h_{n+}^2}{2}\varphi''(\mu_{n+1/2}) + \frac{h_{n+}^3}{6}\varphi'''(\mu_{n+1/2}) + \frac{h_{n+}^4}{24}\varphi^{4)}(\mu_{n+1/2}) + \frac{h_{n+}^5}{120}\varphi^{5)}(\xi_{n+}), \tag{72}$$

$$\varphi(\mu_n) = \varphi(\mu_{n+1/2}) - h_{n-}\varphi'(\mu_{n+1/2}) + \frac{h_{n-}^2}{2}\varphi''(\mu_{n+1/2}) - \frac{h_{n-}^3}{6}\varphi'''(\mu_{n+1/2}) + \frac{h_{n-}^4}{24}\varphi^{4)}(\mu_{n+1/2}) - \frac{h_{n-}^5}{120}\varphi^{5)}(\xi_{n-}), \tag{73}$$

and then dividing the result by $\mu_{n+1}-\mu_n = h_{n-}+h_{n+}$, one gets

$$-E_n = \frac{\varphi(\mu_{n+1}) - \varphi(\mu_n)}{\mu_{n+1} - \mu_n} - \varphi'(\mu_{n+1/2})$$
$$= \frac{h_{n+}^2 - h_{n-}^2}{2(h_{n-} + h_{n+})}\varphi''(\mu_{n+1/2}) + \frac{h_{n+}^3 + h_{n-}^3}{6(h_{n-} + h_{n+})}\varphi'''(\mu_{n+1/2})$$
$$+ \frac{h_{n+}^4 - h_{n-}^4}{24(h_{n-} + h_{n+})}\varphi^{4)}(\mu_{n+1/2}) + \frac{h_{n+}^5\varphi^{5)}(\xi_{n+}) + h_{n-}^5\varphi^{5)}(\xi_{n-})}{120(h_{n-} + h_{n+})}. \tag{74}$$

Finally, the error representation (71) results from (74) and the following equalities:

$$\frac{h_{n+}^2 - h_{n-}^2}{2(h_{n-} + h_{n+})} = \frac{h_{n+} - h_{n-}}{2} = d_n, \tag{75}$$
$$\frac{h_{n+}^3 + h_{n-}^3}{6(h_{n-} + h_{n+})} = \frac{1}{6}(h_{n+}^2 - h_{n+}h_{n-} + h_{n-}^2) = \frac{h_n^2 + 3d_n^2}{6}, \tag{76}$$
$$\frac{h_{n+}^4 - h_{n-}^4}{24(h_{n-} + h_{n+})} = \frac{(h_{n+}^2 + h_{n-}^2)(h_{n+} - h_{n-})}{24} = \frac{h_n^2 d_n + d_n^3}{6}, \tag{77}$$

where we have used the identities $h_{n-} = h_n - d_n$ and $h_{n+} = h_n + d_n$. □

6. Some general comments

We will describe in the following sections difference schemes for approximating the FP angular diffusion operator $\Delta_{\rm FP} f$ defined by Equation (1). In what follows, $\Delta_{{\rm FP},N} f(\mu_n)$ will stand for an approximation of $\Delta_{\rm FP} f(\mu_n)$ obtained on a mesh of N nodes.

Definition 10. For each $n = 1, \ldots, N$, we define the truncation error $R_n(f)$ as

$$R_n(f) = \Delta_{\rm FP} f(\mu_n) - \Delta_{{\rm FP},N} f(\mu_n). \tag{78}$$

Definition 11. A numerical scheme for computing $\Delta_{{\rm FP},N} f(\mu_n)$

1. Converges for the function f if

$$\lim_{N\to\infty} \max_{1\le n\le N} |R_n(f)| = 0. \tag{79}$$

2. Converges with (at least) order p for the function f if

$$\max_{1\le n\le N} |R_n(f)| = O(N^{-p}) \tag{80}$$

for certain positive real number p.
3. Converges with order p if converges with order p for all f regular enough, which in this paper will mean that there exists $k \in \mathbb{N}$ such that converges with order p for all $f \in {\rm C}^k([-1,1])$.

As anticipated in the introduction, a particular case of DOM schemes will have a special relevance in this paper: the GL schemes, the definition of which is formalized as follows.

Definition 12. A GL scheme is any formula of numerical differentiation for approximating the operator (1) that takes as nodes

- Either the set $\{\mu_n\}_{n=1}^N$ of GL nodes on $(-1,1)$, in which case we will say that the GL scheme is being operated in full range (FR) mode,
- Or, being N even and $M = N/2$, the set

$$\{\mu_n\}_{n=1}^N = \{\mu_n\}_{n=1}^M \cup \{\mu_n\}_{n=M+1}^{2M}, \tag{81}$$

where $\{\mu_n\}_{n=1}^M$ are the GL nodes on $(-1,0)$ and $\{\mu_n\}_{n=M+1}^{2M}$ are the GL nodes on $(0,1)$, in which case we will say that the GL scheme is being operated in half range (HR) mode.

Unless specified otherwise, whenever we say from now on that $\{\mu_n\}_{n=1}^N$ are the GL nodes, we mean that they are either the GL nodes on $(-1,1)$ or the nodes defined by Equation (81). The same applies to the GL weights $\{w_n\}_{n=1}^N$.

Definition 12 contemplates two typical modes of operation when solving the FPE: the FR and the HR mode (see for instance [2]). We notice that, automatically:

1. The FR mode refines the mesh in the vicinity of -1 and 1. The FPE degenerates at $\mu = 0$, and so the node 0 is typically avoided by taking N even, but the parity of N is not at all relevant when studying the convergence of the schemes that discretize the angular diffusion operator in isolation. We think that the ideas contained in this paper can be used to design a DOM scheme for the FPE which can use the FR mode with N odd while maintaining good properties as order of convergence and (see Subsection 6.1) discrete moments preservation, but this will be part of future research.
2. The HR mode avoids the node 0 and refines the mesh in the vicinity of -1, 0, and 1.

6.1. The zeroth and first moment properties

Associated with the FP Laplacian, there are two properties of interest, namely the zeroth and the first moment properties:

$$\int_{-1}^{1} \Delta_{\mathrm{FP}} f(\mu)\, d\mu = 0, \tag{82}$$

$$\int_{-1}^{1} \mu \Delta_{\mathrm{FP}} f(\mu)\, d\mu = -2 \int_{-1}^{1} \mu f(\mu)\, d\mu, \tag{83}$$

both of which are easy to verify. The reader can think about how these properties should be written for diffusivities other than $\mathscr{D}(\mu) = 1 - \mu^2$.

According to [14], it is desirable for a scheme to satisfy discrete versions of these two properties.

Definition 13. We say that a GL scheme

- Satisfies the discrete zeroth moment property (or preserves the zeroth moment) if, for all function $f : [-1, 1] \to \mathbb{R}$ and for all $N \geq 2$,

$$\sum_{n=1}^{N} w_n \Delta_{\mathrm{FP},N} f(\mu_n) = 0. \tag{84}$$

- Satisfies the discrete first moment property (or preserves the first moment) if, for all function $f : [-1, 1] \to \mathbb{R}$ and for all $N \geq 2$,

$$\sum_{n=1}^{N} w_n \mu_n \Delta_{\mathrm{FP},N} f(\mu_n) = -2 \sum_{n=1}^{N} w_n \mu_n f(\mu_n), \tag{85}$$

where $\{w_n\}_{n=1}^{N}$ are the GL weights.

Remark 8. A GL scheme may or may not preserve moments, yet, if it is convergent, it invariably preserves moments asymptotically, i.e., as N tends to infinity, when f is regular enough. This fact is a consequence of the convergence of the GL approximations in Equations (84) and (85) to the exact value of the integrals in Equations (82) and (83).

Definition 13 relies on GL quadrature, which is natural for GL schemes, but, when dealing with a non-GL scheme, an analogous definition can be written based on some other appropriate quadrature rule.

7. Schemes of type I

After (4), and noticing that $\mathscr{D}(\mu_{1/2}) = \mathscr{D}(\mu_{N+1/2}) = 0$, let us consider the following scheme:

$$\Delta_{\mathrm{FP},N} f(\mu_1) = \frac{\mathscr{D}(\mu_{1+1/2}) \frac{f(\mu_2) - f(\mu_1)}{\mu_2 - \mu_1}}{\mu_{1+1/2} + 1}, \tag{86}$$

$$\Delta_{\mathrm{FP},N} f(\mu_n) = \frac{\mathscr{D}(\mu_{n+1/2}) \frac{f(\mu_{n+1}) - f(\mu_n)}{\mu_{n+1} - \mu_n} - \mathscr{D}(\mu_{n-1/2}) \frac{f(\mu_n) - f(\mu_{n-1})}{\mu_n - \mu_{n-1}}}{\mu_{n+1/2} - \mu_{n-1/2}}$$

$$\text{for } n = 2, \dots, N-1, \tag{87}$$

$$\Delta_{\mathrm{FP},N} f(\mu_N) = \frac{-\mathscr{D}(\mu_{N-1/2}) \frac{f(\mu_N) - f(\mu_{N-1})}{\mu_N - \mu_{N-1}}}{1 - \mu_{N-1/2}}. \tag{88}$$

The scheme (86)–(88) can be written simply as

$$\Delta_{\mathrm{FP},N} f(\mu_n) = \frac{\mathscr{D}(\mu_{n+1/2})\frac{f(\mu_{n+1})-f(\mu_n)}{\mu_{n+1}-\mu_n} - \mathscr{D}(\mu_{n-1/2})\frac{f(\mu_n)-f(\mu_{n-1})}{\mu_n-\mu_{n-1}}}{\mu_{n+1/2}-\mu_{n-1/2}} \quad \text{for } n=1,\ldots,N, \tag{89}$$

understanding that the terms containing the undefined nodes μ_0 and μ_{N+1} must be ignored as they are multiplied by zero.

This is really a family of schemes depending upon the choice of the nodes μ_n and the points $\mu_{n+1/2}$. We shall refer to the members of this family as schemes of type I.

7.1. First example: midpoint GL scheme

We will call *midpoint GL scheme* the scheme of type I obtained when $\{\mu_n\}_{n=1}^N$ are the GL nodes and the points $\{\mu_{n+1/2}\}_{n=0}^N$ are defined by $\mu_{1/2}=-1$, $\mu_{n+1/2}=(\mu_n+\mu_{n+1})/2$ for $n=1,\ldots,N-1$, $\mu_{N+1/2}=1$. The use of midpoints in this scheme brings to mind the discretizations employed by Antal and Lee in [1] and by Mehlhorn and Duderstadt in [13].

It does not preserve either the zeroth or the first moment. Later, we will show that it converges with order 2 regardless of whether it is operated in FR or HR mode.

7.2. Second example: Haldy-Ligou's scheme

If $\{\mu_n\}_{n=1}^N$ are the GL nodes, and the points $\{\mu_{n+1/2}\}_{n=0}^N$ are defined by $\mu_{1/2}=-1$, $\mu_{n+1/2}=\mu_{n-1/2}+w_n$ for $n=1,\ldots,N$, being $\{w_n\}_{n=1}^N$ the GL weights, one recovers, according to [14, Eqns. (10), (10a), (10b)], the scheme used by Haldy and Ligou in [5, Eq. (5a)].

So, Haldy-Ligou's scheme reads as follows:

$$\Delta_{\mathrm{FP},N} f(\mu_n) = \frac{\mathscr{D}(\mu_{n+1/2})\frac{f(\mu_{n+1})-f(\mu_n)}{\mu_{n+1}-\mu_n} - \mathscr{D}(\mu_{n-1/2})\frac{f(\mu_n)-f(\mu_{n-1})}{\mu_n-\mu_{n-1}}}{w_n} \quad \text{for } n=1,\ldots,N. \tag{90}$$

It was designed so that the discrete zeroth moment property is satisfied, but it does not preserve the first moment. Also observe that the subtraction $\mu_{n+1/2}-\mu_{n-1/2}$ in Equation (89) has been replaced with w_n, which is better from the viewpoint of rounding.

Since $w_n = w_{N-n+1}$ for $n=1,\ldots,N$ and

$$\sum_{n=1}^{N} w_n = 2, \tag{91}$$

points $\mu_{n+1/2}$ are antisymmetric with respect to 0:

$$\mu_{n+1/2} = -\mu_{N-n+1/2} \text{ for } n=0,\ldots,N. \tag{92}$$

When programming this scheme, it is convenient to take advantage of Equation (92) by calculating only those points $\mu_{n+1/2}$ that belong to $[-1,0]$, and then determining the ones in $(0,1]$ by means of the antisymmetry. In this way, roundoff errors are reduced. In HR mode, computing the points $\mu_{n+1/2}$ within $[-1,-0.5]$ is sufficient for completing the list.

It will be seen later that this scheme converges with order 2 when used in FR mode and, somewhat surprisingly, is not convergent when used in HR mode (see Remark 9).

7.3. Analysis of convergence

Results in this subsection hold for generic diffusivities and not only for $\mathscr{D}(\mu)=1-\mu^2$.

We start with a result on the error representation.

Proposition 3 *(Error representation for schemes of type I). Assume that the hypothesis (6) holds. Let $\mathscr{D}$ be a function of class $\mathrm{C}^1([-1,1])$ such that $\mathscr{D}(-1)=\mathscr{D}(1)=0$. Suppose that $f\in\mathrm{C}^2([-1,1])$ and that $\Delta_{\mathrm{FP},N} f(\mu_n)$ is defined by Equation (89). Then, for $n=1,\ldots,N$,*

$$\Delta_{\mathrm{FP}} f(\mu_n) = \Delta_{\mathrm{FP},N} f(\mu_n) + R_n(f), \tag{93}$$

with

$$R_n(f) = \varepsilon_n(f) + E_n^{\star}(\mathscr{D} f'), \tag{94}$$

being

$$\varepsilon_1(f) = \frac{\mathscr{D}(\mu_{1+1/2})E_1(f)}{\mu_{1+1/2}+1}, \tag{95}$$

$$\varepsilon_n(f) = \frac{\mathscr{D}(\mu_{n+1/2})E_n(f) - \mathscr{D}(\mu_{n-1/2})E_{n-1}(f)}{\mu_{n+1/2}-\mu_{n-1/2}} \text{ for } n = 2,\dots,N-1, \tag{96}$$

$$\varepsilon_N(f) = -\frac{\mathscr{D}(\mu_{N-1/2})E_{N-1}(f)}{1-\mu_{N-1/2}}. \tag{97}$$

In the expressions above, $E_n^\star(\mathscr{D} f')$ and $E_n(f)$ are those defined by Equations (57) and (69), respectively.

Proof. $\Delta_{\rm FP} f$ is well defined in the classical sense because $\mathscr{D} \in {\rm C}^1([-1,1])$ and $f \in {\rm C}^2([-1,1])$.

Using (57), the equality $\mathscr{D}(\mu_{1/2}) = 0$, and (69),

$$\begin{aligned}\Delta_{\rm FP} f(\mu_1) &= \frac{\mathscr{D}(\mu_{1+1/2})f'(\mu_{1+1/2}) - \mathscr{D}(\mu_{1/2})f'(\mu_{1/2})}{\mu_{1+1/2}+1} + E_1^\star(\mathscr{D} f')\\ &= \frac{\mathscr{D}(\mu_{1+1/2})f'(\mu_{1+1/2})}{\mu_{1+1/2}+1} + E_1^\star(\mathscr{D} f')\\ &= \frac{\mathscr{D}(\mu_{1+1/2})}{\mu_{1+1/2}+1}\left\{\frac{f(\mu_2)-f(\mu_1)}{\mu_2-\mu_1} + E_1(f)\right\} + E_1^\star(\mathscr{D} f')\\ &= \Delta_{{\rm FP},N} f(\mu_1) + \varepsilon_1(f) + E_1^\star(\mathscr{D} f'),\end{aligned} \tag{98}$$

with $\varepsilon_1(f)$ given by (95). The missing proofs can be done analogously. □

Our goal is to fix certain conditions on the set of nodes and points so that the scheme converges with order 2. Thanks to Equation (94) and Lemma 4, the point is to establish conditions for $\max_{1\le n\le N} |\varepsilon_n(f)|$ to be a $O(N^{-2})$ when f is regular enough.

As anticipated by (95) and (97), the determination of bounds for $|\varepsilon_1(f)|$ and $|\varepsilon_N(f)|$ is special because $\mathscr{D}(\mu_{1/2}) = \mathscr{D}(\mu_{N+1/2}) = 0$. It turns out to be a very easy task.

Proposition 4 *(Bound for $\max\{|\varepsilon_1(f)|, |\varepsilon_N(f)|\}$). Assume that the hypothesis (6) holds. Suppose that $\mathscr{D}$ is a function of class ${\rm C}^1([-1,1])$ such that $\mathscr{D}(-1) = \mathscr{D}(1) = 0$. Let f be a function of class ${\rm C}^2([-1,1])$ and let $\varepsilon_1(f)$, $\varepsilon_N(f)$ be the quantities defined by Equations (95) and (97), respectively. Then,*

$$\max\{|\varepsilon_1(f)|, |\varepsilon_N(f)|\} \le \|\mathscr{D}'\|_\infty \left(\max_{1\le n\le N-1} |E_n(f)|\right). \tag{99}$$

Proof. Notice that $\mu_{1+1/2} + 1 = 2h_1^*$. Then, Equation (95) and the equality

$$\mathscr{D}(\mu_{1+1/2}) = \mathscr{D}(\mu_{1/2}) + 2h_1^*\mathscr{D}'(c_1) = 2h_1^*\mathscr{D}'(c_1), \tag{100}$$

obtained by means of Taylor's theorem, imply

$$|\varepsilon_1(f)| \le \|\mathscr{D}'\|_\infty \left(\max_{1\le n\le N-1} |E_n(f)|\right). \tag{101}$$

Proceeding in a similar way, one sees that the same upper bound is valid for $|\varepsilon_N(f)|$. □

Obtaining an appropriate bound for $\max_{2\le n\le N-1} |\varepsilon_n(f)|$ is much more difficult. We need to introduce some new definitions and, as will be seen in the proof of Proposition 5 below, use the second part of Lemma 5 and break the problem into several simpler ones.

Definition 14. For $n = 1,\dots,N-1$, we define $a_n = d_n + d_n^*$ and

$$A_N = \max_{1\le n\le N-1} |a_n|. \tag{102}$$

Definition 15. For $n = 2,\dots,N$, we define $b_n = d_{n-1} + d_n^*$ and

$$B_N = \max_{2\le n\le N} |b_n|. \tag{103}$$

Definition 16. $C_N = D_N + D_N^*$.

Notice that C_N can be used to bound both A_N and B_N.

Definition 17. $\beta_N(\mathscr{D})$ is the number defined by

$$\beta_N(\mathscr{D}) = \max_{2\le n\le N-1} \left| \frac{(d_{n-1}-d_n)\mathscr{D}(\mu_{n+1/2})}{\mu_{n+1/2}-\mu_{n-1/2}} \right| \tag{104}$$

or, equivalently,

$$\beta_N(\mathscr{D}) = \max_{2\le n\le N-1} \left| \frac{(d_{n-1}-d_n)\mathscr{D}(\mu_{n+1/2})}{2h_n^\star} \right|. \tag{105}$$

Proposition 5 *(Bound for* $\max_{2\le n\le N-1}|\varepsilon_n(f)|$*). Assume that the hypothesis (6) holds. Suppose that* $\mathscr{D}$ *is a function of class* $C^1([-1,1])$*. Let us understand that* $\beta_N=\beta_N(\mathscr{D})$ *and let* f *be a function of class* $C^5([-1,1])$*. Fix* $n\in\{2,\dots,N-1\}$ *and let* $\varepsilon_n(f)$ *be the quantity defined by Equation (96). Then,*

$$\varepsilon_n(f) = \varepsilon_n^{(1)}(f)+\varepsilon_n^{(2)}(f)+\varepsilon_n^{(3)}(f)+\varepsilon_n^{(4)}(f), \tag{106}$$

with

$$|\varepsilon_n^{(1)}(f)| \le \beta_N\|f''\|_\infty + D_N\|(\mathscr{D}f'')'\|_\infty, \tag{107}$$

$$|\varepsilon_n^{(2)}(f)| \le \frac{\beta_N(3D_N+C_N)}{3}\|f'''\|_\infty + \frac{C_N}{3}\|\mathscr{D}f'''\|_\infty + \frac{M_N^2+12D_N^2}{24}\|(\mathscr{D}f''')'\|_\infty, \tag{108}$$

$$|\varepsilon_n^{(3)}(f)| \le \frac{\beta_N\{(M_N^*)^2+8C_ND_N+4C_N^2+12D_N^2\}}{24}\|f^{4)}\|_\infty + \frac{C_ND_N}{3}\|\mathscr{D}f^{4)}\|_\infty + \frac{M_N^2D_N+4D_N^3}{24}\|(\mathscr{D}f^{4)})'\|_\infty, \tag{109}$$

$$|\varepsilon_n^{(4)}(f)| \le \frac{Z_N}{960}\|\mathscr{D}\|_\infty\|f^{5)}\|_\infty, \tag{110}$$

where

$$Z_N = (M_N^*)^3 + 8(M_N^*)^2C_N + 24M_N^*C_N^2 + 40M_N^*D_N^2 + 32C_N^3 + 160C_ND_N^2 + \frac{16}{m_N^\star}(C_N^4+10C_N^2D_N^2+5D_N^4). \tag{111}$$

Proof. According to Equations (71) and (96),

$$\begin{aligned}\varepsilon_n(f) = \frac{1}{2h_n^*}\Bigg\{ &\mathscr{D}(\mu_{n-1/2})\Bigg[d_{n-1}f''(\mu_{n-1/2}) + \frac{h_{n-1}^2+3d_{n-1}^2}{6}f'''(\mu_{n-1/2}) + \frac{h_{n-1}^2d_{n-1}+d_{n-1}^3}{6}f^{4)}(\mu_{n-1/2})\\ &+ \frac{h_{(n-1)+}^5 f^{5)}(\xi_{(n-1)+}) + h_{(n-1)-}^5 f^{5)}(\xi_{(n-1)-})}{120(h_{(n-1)-}+h_{(n-1)+})}\Bigg]\\ &-\mathscr{D}(\mu_{n+1/2})\Bigg[d_nf''(\mu_{n+1/2}) + \frac{h_n^2+3d_n^2}{6}f'''(\mu_{n+1/2})\\ &+\frac{h_n^2d_n+d_n^3}{6}f^{4)}(\mu_{n+1/2}) + \frac{h_{n+}^5f^{5)}(\xi_{n+})+h_{n-}^5f^{5)}(\xi_{n-})}{120(h_{n-}+h_{n+})}\Bigg]\Bigg\},\end{aligned} \tag{112}$$

which gives (106) with

$$\varepsilon_n^{(1)}(f) = \frac{1}{2h_n^*}\{d_{n-1}\mathscr{D}(\mu_{n-1/2})f''(\mu_{n-1/2}) - d_n\mathscr{D}(\mu_{n+1/2})f''(\mu_{n+1/2})\}, \tag{113}$$

$$\varepsilon_n^{(2)}(f) = \frac{1}{2h_n^*}\left\{\frac{h_{n-1}^2+3d_{n-1}^2}{6}\mathscr{D}(\mu_{n-1/2})f'''(\mu_{n-1/2}) - \frac{h_n^2+3d_n^2}{6}\mathscr{D}(\mu_{n+1/2})f'''(\mu_{n+1/2})\right\}, \tag{114}$$

$$\varepsilon_n^{(3)}(f) = \frac{1}{2h_n^*}\left\{\frac{h_{n-1}^2d_{n-1}+d_{n-1}^3}{6}\mathscr{D}(\mu_{n-1/2})f^{4)}(\mu_{n-1/2}) - \frac{h_n^2d_n+d_n^3}{6}\mathscr{D}(\mu_{n+1/2})f^{4)}(\mu_{n+1/2})\right\}, \tag{115}$$

$$\varepsilon_n^{(4)}(f) = \frac{1}{2h_n^*}\left\{\mathscr{D}(\mu_{n-1/2})\frac{h_{(n-1)+}^5f^{5)}(\xi_{(n-1)+})+h_{(n-1)-}^5f^{5)}(\xi_{(n-1)-})}{120(h_{(n-1)-}+h_{(n-1)+})} - \mathscr{D}(\mu_{n+1/2})\frac{h_{n+}^5f^{5)}(\xi_{n+})+h_{n-}^5f^{5)}(\xi_{n-})}{120(h_{n-}+h_{n+})}\right\}. \tag{116}$$

We will prove (107)–(109) firstly and leave the proof of (110) for later.

Notice that, for $r=1,2,3$, we have by Taylor that

$$\mathscr{D}(\mu_{n-1/2})f^{r+1)}(\mu_{n-1/2}) = \mathscr{D}(\mu_{n+1/2})f^{r+1)}(\mu_{n+1/2}) - 2h_n^*(\mathscr{D}f^{r+1)})'(c_n^{(r)}), \tag{117}$$

and hence the expressions (113)–(115) can be rewritten as follows:

$$\varepsilon_n^{(1)}(f) = \frac{(d_{n-1}-d_n)\mathscr{D}(\mu_{n+1/2})}{2h_n^*} f''(\mu_{n+1/2}) - d_{n-1}(\mathscr{D} f'')'(c_n^{(1)}), \tag{118}$$

$$\varepsilon_n^{(2)}(f) = \frac{(h_{n-1}^2+3d_{n-1}^2)-(h_n^2+3d_n^2)}{12h_n^*}\mathscr{D}(\mu_{n+1/2})f'''(\mu_{n+1/2}) - \frac{(h_{n-1}^2+3d_{n-1}^2)}{6}(\mathscr{D} f''')'(c_n^{(2)}), \tag{119}$$

$$\varepsilon_n^{(3)}(f) = \frac{(h_{n-1}^2 d_{n-1}+d_{n-1}^3)-(h_n^2 d_n+d_n^3)}{12h_n^*}\mathscr{D}(\mu_{n+1/2})f^{4)}(\mu_{n+1/2}) - \frac{(h_{n-1}^2 d_{n-1}+d_{n-1}^3)}{6}(\mathscr{D} f^{4)})'(c_n^{(3)}). \tag{120}$$

Now we proceed to bound each of these three terms separately.

- Bound for $|\varepsilon_n^{(1)}(f)|$: the bound (107) follows immediately from (118).
- Bound for $|\varepsilon_n^{(2)}(f)|$: thanks to Equation (46) we have

$$h_{n-1}^2 - h_n^2 = (h_{n-1}+h_n)(h_{n-1}-h_n) = \{2h_n^* - (d_{n-1}-d_n)\}(h_{n-1}-h_n). \tag{121}$$

Thus, noticing that $h_{n-1}-h_n = -a_n - b_n$, which holds in virtue of Definitions 14, 15 and Equations (44), (45),[1] we arrive at

$$\begin{aligned}\frac{(h_{n-1}^2+3d_{n-1}^2)-(h_n^2+3d_n^2)}{12h_n^*} &= \frac{\{2h_n^*-(d_{n-1}-d_n)\}(-a_n-b_n)+3(d_{n-1}+d_n)(d_{n-1}-d_n)}{12h_n^*}\\ &= -\frac{a_n+b_n}{6} + \frac{\{3(d_{n-1}+d_n)+a_n+b_n\}(d_{n-1}-d_n)}{12h_n^*},\end{aligned} \tag{122}$$

and then Equation (119) can be rewritten as

$$\begin{aligned}\varepsilon_n^{(2)}(f) = &-\frac{a_n+b_n}{6}\mathscr{D}(\mu_{n+1/2})f'''(\mu_{n+1/2}) + \frac{\{3(d_{n-1}+d_n)+a_n+b_n\}}{6}\frac{(d_{n-1}-d_n)\mathscr{D}(\mu_{n+1/2})}{2h_n^*}f'''(\mu_{n+1/2})\\ &- \frac{(h_{n-1}^2+3d_{n-1}^2)}{6}(\mathscr{D} f''')'(c_n^{(2)}),\end{aligned} \tag{123}$$

which implies

$$|\varepsilon_n^{(2)}(f)| \le \frac{C_N}{3}\|\mathscr{D} f'''\|_\infty + \frac{3D_N+C_N}{3}\beta_N\|f'''\|_\infty + \frac{1}{6}\{(M_N/2)^2+3D_N^2\}\|(\mathscr{D} f''')'\|_\infty. \tag{124}$$

Finally, observe that (124) is equivalent to (108).
- Bound for $|\varepsilon_n^{(3)}(f)|$: keeping Equation (120) in mind, we will begin by obtaining expressions for $h_{n-1}^2 d_{n-1} - h_n^2 d_n$ and for $d_{n-1}^3 - d_n^3$ that allow us to bound $|\varepsilon_n^{(3)}(f)|$ in an optimal way. Taking into account the previous bounds, we realize that it is convenient to bring up the $d_{n-1}-d_n$ factor as many times as possible.
The easiest part is $d_{n-1}^3 - d_n^3$:

$$d_{n-1}^3 - d_n^3 = (d_{n-1}^2 + d_{n-1}d_n + d_n^2)(d_{n-1}-d_n). \tag{125}$$

Let us proceed now with $h_{n-1}^2 d_{n-1} - h_n^2 d_n$. It is known, from Equations (44) and (45), that $h_n = h_n^* + a_n$ and $h_{n-1} = h_n^* - b_n$. Hence,

$$\begin{aligned}h_{n-1}^2 d_{n-1} - h_n^2 d_n &= (h_n^*-b_n)^2 d_{n-1} - (h_n^*+a_n)^2 d_n\\ &= (h_n^*)^2(d_{n-1}-d_n) - 2h_n^*(b_n d_{n-1} + a_n d_n) + b_n^2 d_{n-1} - a_n^2 d_n.\end{aligned} \tag{126}$$

Next step is to prove that $b_n^2 d_{n-1} - a_n^2 d_n$ is a multiple of $d_{n-1}-d_n$. Note that, in virtue of Definitions 14 and 15, $b_n - a_n = d_{n-1} - d_n$. So,

$$\begin{aligned}b_n^2 d_{n-1} - a_n^2 d_n &= b_n^2(d_{n-1}-d_n) + (b_n^2-a_n^2)d_n\\ &= b_n^2(d_{n-1}-d_n) + (b_n+a_n)(b_n-a_n)d_n\\ &= \{b_n^2 + (b_n+a_n)d_n\}(d_{n-1}-d_n),\end{aligned} \tag{127}$$

and Equation (126) becomes

$$h_{n-1}^2 d_{n-1} - h_n^2 d_n = \{(h_n^*)^2 + b_n^2 + (a_n+b_n)d_n\}(d_{n-1}-d_n) - 2h_n^*(b_n d_{n-1} + a_n d_n). \tag{128}$$

In summary,

$$(h_{n-1}^2 d_{n-1} + d_{n-1}^3) - (h_n^2 d_n + d_n^3) = \{(h_n^*)^2 + b_n^2 + (a_n+b_n)d_n + d_{n-1}^2 + d_{n-1}d_n + d_n^2\}(d_{n-1}-d_n) - 2h_n^*(b_n d_{n-1} + a_n d_n). \tag{129}$$

[1] The identity $h_n - h_{n-1} = a_n + b_n$ is also a rewriting of Equation (22) in Lemma 3.

If we define now

$$x_n = (h_n^*)^2 + b_n^2 + (a_n + b_n)d_n + d_{n-1}^2 + d_{n-1}d_n + d_n^2, \tag{130}$$

Equation (120) can be rewritten as follows:

$$\varepsilon_n^{(3)}(f) = \frac{x_n}{6}\frac{(d_{n-1} - d_n)\mathscr{D}(\mu_{n+1/2})}{2h_n^*} f^{4)}(\mu_{n+1/2}) - \frac{b_n d_{n-1} + a_n d_n}{6}\mathscr{D}(\mu_{n+1/2}) f^{4)}(\mu_{n+1/2}) - \frac{(h_{n-1}^2 d_{n-1} + d_{n-1}^3)}{6}(\mathscr{D} f^{4)})'(c_n^{(3)}). \tag{131}$$

Finally, taking account of

$$|x_n| \le (M_N^*/2)^2 + C_N^2 + 2C_N D_N + 3D_N^2, \tag{132}$$

$$|b_n d_{n-1} + a_n d_n| \le 2C_N D_N, \tag{133}$$

the bound (109) is deduced from Equation (131).

We now proceed with the proof of (110).
Firstly note that, for $k \in \{n-1, n\}$,

$$\frac{h_{k+}^5 + h_{k-}^5}{h_{k-} + h_{k+}} = h_{k+}^4 - h_{k+}^3 h_{k-} + h_{k+}^2 h_{k-}^2 - h_{k+} h_{k-}^3 + h_{k-}^4 = h_k^4 + 10h_k^2 d_k^2 + 5d_k^4. \tag{134}$$

The equalities $h_{k-} = h_k - d_k$ and $h_{k+} = h_k + d_k$ have been used in the last step.
Then, in virtue of Equation (116) and the positivity of h_{k-} and h_{k+},

$$|\varepsilon_n^{(4)}(f)| \le \frac{\|\mathscr{D}\|_\infty \|f^{5)}\|_\infty}{240}\frac{1}{h_n^*}(h_{n-1}^4 + 10h_{n-1}^2 d_{n-1}^2 + 5d_{n-1}^4 + h_n^4 + 10h_n^2 d_n^2 + 5d_n^4). \tag{135}$$

This bound is sharper than it seems, as we are about to see below. Since $h_{n-1} = h_n^* - b_n$ and $h_n = h_n^* + a_n$, we have

$$h_{n-1}^4 + 10h_{n-1}^2 d_{n-1}^2 = (h_n^*)^4 - 4(h_n^*)^3 b_n + 6(h_n^*)^2 b_n^2 + 10(h_n^*)^2 d_{n-1}^2 - 4h_n^* b_n^3 - 20h_n^* b_n d_{n-1}^2 + b_n^4 + 10b_n^2 d_{n-1}^2 \tag{136}$$

and

$$h_n^4 + 10h_n^2 d_n^2 = (h_n^*)^4 + 4(h_n^*)^3 a_n + 6(h_n^*)^2 a_n^2 + 10(h_n^*)^2 d_n^2 + 4h_n^* a_n^3 + 20h_n^* a_n d_n^2 + a_n^4 + 10a_n^2 d_n^2, \tag{137}$$

which allows rewriting the inequality (135) as

$$\begin{aligned}|\varepsilon_n^{(4)}(f)| \le \frac{\|\mathscr{D}\|_\infty \|f^{5)}\|_\infty}{240}\{&(h_n^*)^3 - 4(h_n^*)^2 b_n + 6h_n^* b_n^2 + 10h_n^* d_{n-1}^2 - 4b_n^3 - 20b_n d_{n-1}^2 + (h_n^*)^3 + 4(h_n^*)^2 a_n + 6h_n^* a_n^2 \\ &+ 10h_n^* d_n^2 + 4a_n^3 + 20a_n d_n^2 + \frac{1}{h_n^*}(b_n^4 + 10b_n^2 d_{n-1}^2 + 5d_{n-1}^4 + a_n^4 + 10a_n^2 d_n^2 + 5d_n^4)\}.\end{aligned} \tag{138}$$

The proof ends by using the bounds $h_n^* \le M_N^*/2$, $|b_n| \le C_N$, $|a_n| \le C_N$, $|d_n| \le D_N$, and $1/h_n^* \le 2/m_N^*$. □

We can now state the main result of this subsection.

Theorem 2 *(Order 2 of convergence for schemes of type I). Let $\mathscr{D}$ be a function of class $C^3([-1,1])$ such that $\mathscr{D}(-1) = \mathscr{D}(1) = 0$. Suppose that the sets of nodes and points satisfy the conditions stated in Section 3 and that $\beta_N(\mathscr{D})$ goes to zero at least with order 2. That is to say, suppose that*

$$-1 = \mu_{1/2} < \mu_1 < \mu_{1+1/2} < \cdots < \mu_{N-1/2} < \mu_N < \mu_{N+1/2} = 1, \tag{139}$$

$$\widetilde{M}_N = O(N^{-1}), \tag{140}$$

$$D_N^\star = O(N^{-q}) \text{ with } q \ge 2, \tag{141}$$

$$D_N = O(N^{-r}) \text{ with } r \ge 2, \tag{142}$$

$$\frac{1}{m_N^*} = O(N^s) \text{ with } 1 \le s \le 4m - 2, \text{ where } m = \min\{q, r\}, \tag{143}$$

$$\beta_N(\mathscr{D}) = O(N^{-t}) \text{ with } t \ge 2. \tag{144}$$

Then, the scheme (89) converges with order 2 for any function f of class $C^5([-1,1])$, and the same is true if [$D_N = D_N^* = 0$ and the hypotheses (139) and (140) hold] or if [$d_1 = \cdots = d_{N-1}$ and the hypotheses (139)–(143) hold].

Proof. Let f be a function of class $C^5([-1,1])$.

Thanks to Proposition 4 and Lemma 5 we know that

$$\max\{|\varepsilon_1(f)|, |\varepsilon_N(f)|\} = O(N^{-2}). \tag{145}$$

Moreover, since $C_N = D_N + D_N^* = O(N^{-m})$, Proposition 5 implies that

$$\max_{2\le n\le N-1} |\varepsilon_n^{(1)}(f)| = O(N^{-\min\{r,t\}}), \tag{146}$$

$$\max_{2\le n\le N-1} |\varepsilon_n^{(2)}(f)| = O(N^{-2}), \tag{147}$$

$$\max_{2\le n\le N-1} |\varepsilon_n^{(3)}(f)| = O(N^{-\min\{r+2,t+2\}}), \tag{148}$$

$$\max_{2\le n\le N-1} |\varepsilon_n^{(4)}(f)| = O(N^{-\min\{3,4m-s\}}). \tag{149}$$

In summary,

$$\max_{1\le n\le N} |\varepsilon_n(f)| = O(N^{-2}). \tag{150}$$

On the other hand,

$$\max_{1\le n\le N} |E_n^\star(\mathscr{D} f')| = O(N^{-2}) \tag{151}$$

by Lemma 4, and so, in virtue of Proposition 3,

$$\max_{1\le n\le N} |R_n(f)| = O(N^{-2}), \tag{152}$$

which proves convergence of order 2.

The validity of the two assertions yet to be demonstrated follows without difficulty from a simple observation of the bounds in Proposition 5, as well as taking into account Definition 17. □

Since $\mathscr{D}$ is bounded, the hypothesis (144) is automatically satisfied if

$$\max_{2\le n\le N-1} \left| \frac{d_{n-1} - d_n}{\mu_{n+1/2} - \mu_{n-1/2}} \right| = O(N^{-p}) \text{ with } p \ge 2, \tag{153}$$

so the reader might wonder why we have not used this assumption in the previous theorem. After all, that way the set of hypotheses would be independent of $\mathscr{D}$. The reason is that, for $\mathscr{D}(\mu) = 1 - \mu^2$ and the choices of nodes and points made by Haldy and Ligou (Subsection 7.2), condition (153) is not satisfied, while (144) holds with $t = 2$.

To finish this subsection, let us comment that in Proposition 5, and hence in Theorem 2, we can change $\beta_N(\mathscr{D})$ for

$$\tilde{\beta}_N(\mathscr{D}) = \max_{2\le n\le N-1} \left| \frac{(d_{n-1} - d_n)\mathscr{D}(\mu_{n-1/2})}{\mu_{n+1/2} - \mu_{n-1/2}} \right|. \tag{154}$$

Indeed, we could have used

$$\mathscr{D}(\mu_{n+1/2}) f^{r+1)}(\mu_{n+1/2}) = \mathscr{D}(\mu_{n-1/2}) f^{r+1)}(\mu_{n-1/2}) + 2h_n^*(\mathscr{D} f^{r+1)})'(c_n^{(r)}) \tag{155}$$

instead of Equation (117) in the proof of Proposition 5, and thus eliminate the evaluations at $\mu_{n+1/2}$ to be left with the evaluations at $\mu_{n-1/2}$.

7.4. Application of the theory to some examples. Numerical results

In the tables below, E will denote the maximum of the absolute values of the errors in the complete set of nodes, i.e.,

$$E = \max_{1\le n\le N} |\Delta_{\mathrm{FP}} f(\mu_n) - \Delta_{\mathrm{FP},N} f(\mu_n)|. \tag{156}$$

Midpoint GL scheme For this scheme, operated in FR mode,

- Hypothesis (139) is obviously satisfied given the definition of the points $\{\mu_{n+1/2}\}_{n=0}^N$.
- Hypothesis (140) is satisfied in virtue of Proposition 1.
- Hypothesis (141), with $q = 2$, is supported by the results in Section 4, since

$$|\hat{\mu}_1 - \mu_1| = |\hat{\mu}_N - \mu_N| < \mu_2 + 1 \quad \text{and} \tag{157}$$

Table 1
Numerical results for the midpoint GL scheme operated in FR mode. $f(\mu) = e^{\mu}$, $\mathscr{D}(\mu) = 1 - \mu^2$.

N	E	order	q	s
50	1.54×10^{-2}			
100	3.96×10^{-3}	1.96	1.98	1.98
500	1.61×10^{-4}	1.99	1.99	1.99
1000	4.02×10^{-5}	2.00	2.00	2.00
5000	1.61×10^{-6}	2.00	2.00	2.00
10000	4.10×10^{-7}	1.97	2.00	2.00
20000	1.42×10^{-7}	1.53	2.00	2.00

$$\hat{\mu}_n - \mu_n = \frac{1}{2}\left(\frac{\mu_{n-1} + \mu_{n+1}}{2} - \mu_n\right) \text{ for } n \in \{2, \dots, N-1\}. \tag{158}$$

- Hypothesis (142) is obviously satisfied because $D_N = 0$.
- Hypothesis (143) holds with $s = 2$ in virtue of Proposition 1, because

$$\frac{1}{m_N^*} < \frac{1}{\mu_1 + 1}. \tag{159}$$

  In fact, since it is also true that $1/(\mu_2 + 1) < 1/m_N^*$, the value $s = 2$ is optimal.
- Hypothesis (144) is obviously satisfied because $\beta_N(\mathscr{D}) = 0$.

According to Theorem 2, the midpoint GL scheme in FR mode is expected to converge with order 2. Table 1 shows the numerical results got for the FP Laplacian of $f(\mu) = e^{\mu}$. These results are in agreement with the theoretical prediction. Roundoff errors start spoiling the computations in the last row, where the order decays down to 1.53.

In HR mode, the midpoint GL scheme behaves similarly, that is, converges with order 2, but roundoff errors appear earlier, due to the extreme proximity of the nodes in the neighborhood of 0.

As $D_N = 0$, $\beta_N(\mathscr{D})$ is equal to 0 regardless of $\mathscr{D}$, which in turn implies that the scheme is still convergent if other diffusivities are considered. Specifically, the midpoint GL scheme converges with order 2 if $\mathscr{D}$ is any function of class $C^3([-1, 1])$ such that $\mathscr{D}(-1) = \mathscr{D}(1) = 0$.

Haldy-Ligou's scheme For this scheme, operated in FR mode,

- Hypothesis (139) is supported by the results in Section 4.
- Hypothesis (140) is satisfied in virtue of Proposition 1.
- Hypotheses (141) and (142) hold with $q = r = 2$, which is again supported by the results in Section 4.
- Hypothesis (143) holds with $s = 2$ in virtue of Proposition 2, because

$$\frac{1}{m_N^*} = \frac{1}{w_1}. \tag{160}$$

- If $\mathscr{D}(\mu) = 1 - \mu^2$, hypothesis (144) is satisfied with $t = 2$. This assertion is supported by some asymptotic analysis of the same type as that considered in [11]. We observe that the number $\beta_N(\mathscr{D})$ can alternatively be written as

$$\beta_N(\mathscr{D}) = \max_{2 \le n \le N-1} \left| \left(\frac{\mu_{n+1} - \mu_{n-1}}{2w_n} - 1 \right) \mathscr{D}(\mu_{n+1/2}) \right|,$$

  and that it is known (see [11, Theorem 1]) that

$$\Delta_k = \left| \left(\frac{\mu_{k+1} - \mu_{k-1}}{2w_k} - 1 \right) \right| = O(N^{-2})$$

  if $\mu_{k\pm1}$ are in a fixed interval $[a, b] \subset (-1, 1)$. Contrarily, when k is fixed, it is known that $\Delta_k = O(1)$ (with a small error constant), but in that case we have $\mathscr{D}(\mu_{k+1/2}) = O(N^{-2})$.

The comments made in the previous example are valid for this one. Table 2 shows the numerical results, which corroborate that Haldy-Ligou's scheme in FR mode converges with order 2.

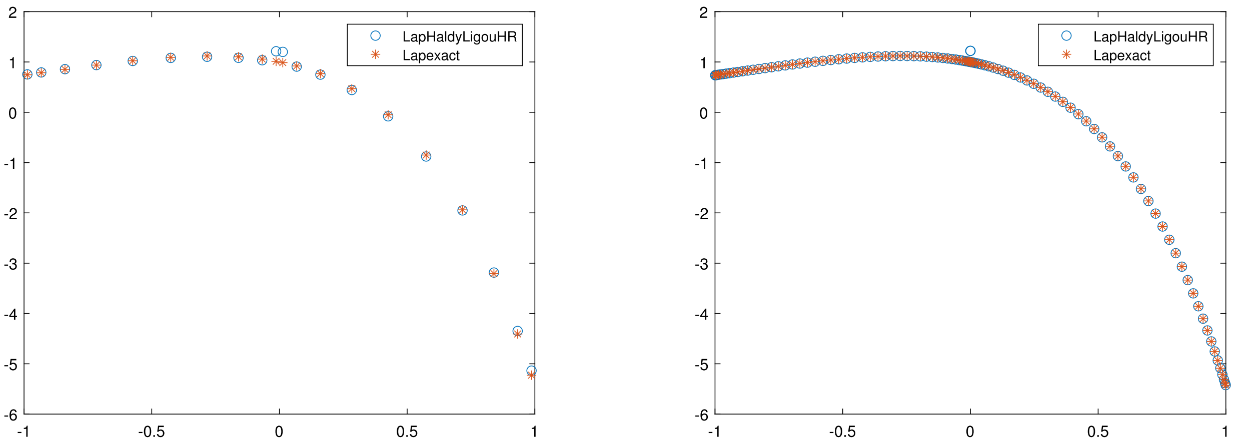
In HR mode, however, Haldy-Ligou's scheme is not convergent. This is shown in Table 3. The 't' column tells us that the problem is that the hypothesis (144) is no longer satisfied. We have not included the 'order' column since, in the absence of convergence, this value loses interest. On the other hand, Fig. 1 shows that it is at nodes close to 0 where the scheme fails, which the reader can connect with the definition of $\beta_N(\mathscr{D})$ and the fact that $\mathscr{D}$ is not zero at 0, while points $\mu_{n+1/2}$ are accumulating quadratically on both sides of 0.

Table 2
Numerical results for Haldy-Ligou's scheme operated in FR mode. $f(\mu) = e^{\mu}$, $\mathscr{D}(\mu) = 1 - \mu^2$.

N	E	order	q	r	s	t
50	8.68×10^{-3}					
100	2.20×10^{-3}	1.98	1.98	1.98	1.99	1.99
500	8.92×10^{-5}	1.99	1.99	1.99	1.99	2.00
1000	2.23×10^{-5}	2.00	2.00	2.00	2.00	2.00
5000	8.95×10^{-7}	2.00	2.00	2.00	2.00	2.00
10000	2.31×10^{-7}	1.95	2.00	2.00	2.00	2.00
20000	9.75×10^{-8}	1.24	2.00	2.00	2.00	2.00

Table 3
Numerical results showing that Haldy-Ligou's scheme operated in HR mode is not convergent. $f(\mu) = e^{\mu}$, $\mathscr{D}(\mu) = 1 - \mu^2$.

$N = 2M$	E	q	r	s	t
50	2.20×10^{-1}				
100	2.21×10^{-1}	1.97	1.96	1.97	-1.14×10^{-2}
500	2.21×10^{-1}	1.99	1.99	1.99	-1.61×10^{-3}
1000	2.21×10^{-1}	2.00	2.00	2.00	-1.19×10^{-4}
5000	2.22×10^{-1}	2.00	2.00	2.00	-1.64×10^{-5}
10000	2.28×10^{-1}	2.00	2.00	2.00	-1.08×10^{-6}
20000	3.64×10^{-1}	2.00	2.00	2.00	-4.54×10^{-9}



(a) $N = 2M = 20$. (b) $N = 2M = 100$.

Fig. 1. Haldy-Ligou's scheme in HR mode cannot compute good approximations of the FP Laplacian in the vicinity of 0. $f(\mu) = e^{\mu}$, $\mathscr{D}(\mu) = 1 - \mu^2$.

The above mentioned accumulation of points $\mu_{n+1/2}$ around 0 does not exist in FR mode, and the quadratic accumulation towards -1 and 1 is not a problem, since there $\mathscr{D}(\mu_{n+1/2})$ tends to zero at a rate that compensates for this accumulation and is enough for $\beta_N(\mathscr{D})$ to be a $O(N^{-2})$.

Remark 9. As it will be explained in Section 9, non-convergence of Haldy-Ligou's scheme in HR mode does not necessarily render it ineffective for solving the FPE.

Uniform mesh (a non-GL scheme of type I and order 2) Let us take $h = 2/N$ and define

- $\mu_1 = -1 + h/2$, $\mu_{n+1} = \mu_n + h$ for $n = 1, \ldots, N-1$,
- $\mu_{1/2} = -1$, $\mu_{n+1/2} = (\mu_n + \mu_{n+1})/2$ for $n = 1, \ldots, N-1$, and $\mu_{N+1/2} = 1$.

Then, by Theorem 2, the corresponding scheme of type I converges with order 2, because $D_N^* = D_N = 0$ and the hypotheses (139) and (140) are trivially met. Results are shown in Table 4.

Table 4
Numerical results for the scheme of type I and of order 2 on uniform mesh. $f(\mu)=\mathrm{e}^{\mu}$, $\mathscr{D}(\mu)=1-\mu^2$.

N	E	order
50	2.44×10^{-3}	
100	6.23×10^{-4}	1.97
500	2.53×10^{-5}	1.99
1000	6.33×10^{-6}	2.00
5000	2.54×10^{-7}	2.00
10000	6.34×10^{-8}	2.00
20000	5.12×10^{-8}	3.09×10^{-1}

Table 5
Numerical results for an instance of scheme of type I on uniform mesh having order 1. $f(\mu)=\mathrm{e}^{\mu}$, $\mathscr{D}(\mu)=1-\mu^2$.

N	E	order	q
50	7.38×10^{-3}		
100	3.68×10^{-3}	1.00	1.01
500	7.36×10^{-4}	1.00	1.00
1000	3.68×10^{-4}	1.00	1.00
5000	7.36×10^{-5}	1.00	1.00
10000	3.68×10^{-5}	1.00	1.00
20000	1.84×10^{-5}	1.00	1.00

This scheme preserves both the zeroth and the first moments, as long as the discrete zeroth and first moment properties of Definition 13 are reinterpreted using, instead GL, the quadrature formula

$$\int_{-1}^{1} G(\mu)\,d\mu = \int_{-1}^{\mu_1} G(\mu)\,d\mu + \int_{\mu_1}^{\mu_N} G(\mu)\,d\mu + \int_{\mu_N}^{1} G(\mu)\,d\mu$$

$$\approx (h/2)G_1 + (\text{trapezoidal approximation}) + (h/2)G_N = h\sum_{n=1}^{N} G_n, \tag{161}$$

where it is understood that G_n is an approximation of $G(\mu_n)$.

If the mesh $\{\mu_n\}_{n=1}^N$ is uniform, but the distance between μ_1 and -1 or between μ_N and 1 is different from $h/2$, then the scheme can easily stop being of order 2. Table 5 shows that the order reduces to 1 if $\mu_1=-1+2/N$, $\mu_N=1-1/N$ and $\{\mu_n\}_{n=2}^{N-1}$ are placed so that $\{\mu_n\}_{n=1}^N$ is uniform. The reason for the order drop is that now D_N^* is only a $O(N^{-1})$ (i.e., $q=1$ in Table 5).

8. Schemes of type II

We will call schemes of type II those schemes obtained by substituting in Equations (86)–(88) the values $\{\mathscr{D}(\mu_{n+1/2})\}_{n=0}^N$ by $\{\alpha_{n+1/2}\}_{n=0}^N$, being $\alpha_{1/2}=\mathscr{D}(\mu_{1/2})=0$, and $\alpha_{n+1/2}$ a certain approximation of $\mathscr{D}(\mu_{n+1/2})$ for $n=1,\dots,N$. We notice that $\alpha_{N+1/2}$ is not necessarily equal to 0.

So, these schemes are defined as follows:

$$\Delta_{\mathrm{FP},N} f(\mu_1) = \frac{\alpha_{1+1/2}\frac{f(\mu_2)-f(\mu_1)}{\mu_2-\mu_1}}{\mu_{1+1/2}+1}, \tag{162}$$

$$\Delta_{\mathrm{FP},N} f(\mu_n) = \frac{\alpha_{n+1/2}\frac{f(\mu_{n+1})-f(\mu_n)}{\mu_{n+1}-\mu_n} - \alpha_{n-1/2}\frac{f(\mu_n)-f(\mu_{n-1})}{\mu_n-\mu_{n-1}}}{\mu_{n+1/2}-\mu_{n-1/2}} \quad \text{for } n=2,\dots,N-1, \tag{163}$$

$$\Delta_{\mathrm{FP},N} f(\mu_N) = \frac{-\alpha_{N-1/2}\frac{f(\mu_N)-f(\mu_{N-1})}{\mu_N-\mu_{N-1}}}{1-\mu_{N-1/2}}. \tag{164}$$

Obviously, the family of schemes of type I is strictly contained in the family of schemes of type II.

After (89), when $\alpha_{N+1/2}=0$ a scheme of type II can be written as

$$\Delta_{\mathrm{FP},N} f(\mu_n)=\frac{\alpha_{n+1/2}\frac{f(\mu_{n+1})-f(\mu_n)}{\mu_{n+1}-\mu_n}-\alpha_{n-1/2}\frac{f(\mu_n)-f(\mu_{n-1})}{\mu_n-\mu_{n-1}}}{\mu_{n+1/2}-\mu_{n-1/2}}$$

$$\text{for } n=1,\dots,N. \tag{165}$$

The values of $\alpha_{n+1/2}$ are computed from those of the nodes μ_n and the points $\mu_{n+1/2}$. Let us explain how this can be done. Notice that, having fixed $\alpha_{1/2}=0$, there is only one way of choosing $\{\alpha_{n+1/2}\}_{n=1}^{N}$ that makes the scheme exact on $\mathbb{P}_1$. Indeed, since it is obviously exact when f is constant and $\Delta_{\mathrm{FP}} f(\mu)=\mathscr{D}'(\mu)$ when $f(\mu)=\mu$, we conclude that a scheme of type II is exact on $\mathbb{P}_1$ if, and only if,

$$\frac{\alpha_{n+1/2}-\alpha_{n-1/2}}{\mu_{n+1/2}-\mu_{n-1/2}}=\mathscr{D}'(\mu_n) \text{ for } n=1,\dots,N. \tag{166}$$

When $\mathscr{D}(\mu)=1-\mu^2$, Equation (166) is telling that $\{\alpha_{n+1/2}\}_{n=0}^{N}$ must be defined by

$$\alpha_{1/2}=\mathscr{D}(\mu_{1/2})=0, \tag{167}$$

$$\alpha_{n+1/2}=\alpha_{n-1/2}-2\mu_n(\mu_{n+1/2}-\mu_{n-1/2}) \text{ for } n=1,\dots,N \tag{168}$$

if we want exactness on $\mathbb{P}_1$.

Definition 18. For $n=0,\dots,N$,

$$\lambda_n=\mathscr{D}(\mu_{n+1/2})-\alpha_{n+1/2}, \tag{169}$$

where it is understood that $\mathscr{D}(\mu)=1-\mu^2$ and that $\{\alpha_{n+1/2}\}_{n=0}^{N}$ is the set defined by (167)–(168).

Definition 19. $\Lambda_N=\max_{0\le n\le N}|\lambda_n|$.

For simplicity, we have decided to use the notations $\alpha_{n+1/2}$, λ_n and Λ_N, and not $\alpha_{n+1/2}(\mathscr{D}),\dots$, even when all three depend on $\mathscr{D}$. Later on, it will be useful to remember this fact.

The following result states precisely what we mean by saying that $\alpha_{n+1/2}$ is an approximation of $\mathscr{D}(\mu_{n+1/2})$.

Theorem 3. *Assume that $\mathscr{D}(\mu)=1-\mu^2$ and that the hypotheses (6), (10), and (19) hold. Then*

$$\Lambda_N=O(N^{-2}). \tag{170}$$

Proof. The initial value problem (IVP)

$$\begin{cases} y'=f(x), \ -1<x<1, \\ y(-1)=\eta\in\mathbb{R}, \end{cases} \tag{171}$$

with $f\in C^2([-1,1])$, has got a unique solution $y\in C^3([-1,1])$.

Since $\mu_{n+1/2}-\mu_{n-1/2}=2h_n^*$, the hypotheses (6) and (10) guarantee that this IVP is solved with order 2 of convergence by the numerical scheme

$$\begin{cases} y_0=\eta, \\ y_n=y_{n-1}+2h_n^* f(\mu_{n-1/2}+h_n^*), \ n=1,\dots,N, \end{cases} \tag{172}$$

where y_n represents an approximation of $y(\mu_{n+1/2})$.

The order 2 of convergence is kept if we replace $f(\mu_{n-1/2}+h_n^*)$ by $f(\mu_{n-1/2}+h_{n-}^*)=f(\mu_n)$, because

$$\max_{1\le n\le N}|h_n^*-h_{n-}^*|=\max_{1\le n\le N}|d_n^*|=D_N^*=O(N^{-q}) \text{ with } q\ge 2 \tag{173}$$

in virtue of the hypothesis (19).

If we solve with the adapted scheme

$$\begin{cases} y_0=\eta, \\ y_n=y_{n-1}+2h_n^* f(\mu_n), \ n=1,\dots,N, \end{cases} \tag{174}$$

the IVP determined by the data $f(x)=-2x$ and $\eta=0$, the solution of which is $y(x)=\mathscr{D}(x)=1-x^2$, we find that $y_n=\alpha_{n+1/2}$ for all $n=0,\dots,N$, and the proof is done. Details are given in [11]. □

8.1. Example: Morel's scheme

Morel's scheme is the scheme of type II obtained when:

- Nodes and points are the same than in the Haldy-Ligou's scheme, that is, $\{\mu_n\}_{n=1}^N$ are the GL nodes, and the points $\{\mu_{n+1/2}\}_{n=0}^N$ are those defined by $\mu_{1/2} = -1$, $\mu_{n+1/2} = \mu_{n-1/2} + w_n$ for $n = 1, \dots, N$, being $\{w_n\}_{n=1}^N$ the GL weights.
- $\{\alpha_{n+1/2}\}_{n=0}^N$ are the values defined by Equations (167)–(168), that is, $\alpha_{1/2} = 0$, $\alpha_{n+1/2} = \alpha_{n-1/2} - 2\mu_n w_n$ for $n = 1, \dots, N$. Since

$$\sum_{n=1}^{N} w_n \mu_n = \int_{-1}^{1} \mu \, d\mu = 0, \tag{175}$$

it turns out that these values satisfy the following symmetry relation:

$$\alpha_{n+1/2} = \alpha_{N-n+1/2} \text{ for } n = 0, \dots, N. \tag{176}$$

In particular, $\alpha_{N+1/2} = 0$.

So, this scheme reads as follows:

$$\Delta_{\mathrm{FP},N} f(\mu_n) = \frac{\alpha_{n+1/2} \frac{f(\mu_{n+1}) - f(\mu_n)}{\mu_{n+1} - \mu_n} - \alpha_{n-1/2} \frac{f(\mu_n) - f(\mu_{n-1})}{\mu_n - \mu_{n-1}}}{w_n}$$

$$\text{for } n = 1, \dots, N. \tag{177}$$

Morel introduced it in [14] expressly so that the discrete zeroth and first moment properties were fulfilled. References [2], [6], [15], [16], and [19] provide examples of its application. Preservation of the zeroth and first moments holds true in both FR and HR modes.

Morel's scheme converges with order 2 when it is operated in FR mode. However, it is not convergent when applied in HR mode, just like Haldy-Ligou's scheme (see Section 9).

As it is clear from Equation (177) and the definition of $\alpha_{n+1/2}$, the points $\mu_{n+1/2}$ are not needed for describing this scheme, and in fact Morel did not mention them at all in [14]. We had to consider these points, though, in order to carry out the forthcoming convergence analysis.

Remark 10 *(Other choices of $\alpha_{n+1/2}$ can be made).* When compared to having order 2, having exactness on $\mathbb{P}_1$ is not that important (notice that schemes of type I are not exact on $\mathbb{P}_1$ unless μ_n be the midpoint of the cell $[\mu_{n-1/2}, \mu_{n+1/2}]$), but presenting the problem of calculating $\alpha_{n+1/2}$ from the exactness on $\mathbb{P}_1$ offers a very simple way to arrive at the same choice that Morel made using a different method. Moreover, our approach can be applied to get good values of $\alpha_{n+1/2}$ via Equation (166) even when a diffusivity other than $\mathscr{D}$ is used.

Having said that, fixing $\mathscr{D}(\mu) = 1 - \mu^2$ and observing the proof of Theorem 3, we could modify the values of $\alpha_{n+1/2}$ simply by using a different numerical method from the one used in this proof. To have an instance, let us suppose that $\{\mu_n\}_{n=1}^N$ and $\{\mu_{n+1/2}\}_{n=0}^N$ are those of Morel's scheme. Then, the choice

$$\alpha_{1/2} = 0, \tag{178}$$

$$\alpha_{n+1/2} = \alpha_{n-1/2} - w_n(2\mu_{n-1/2} + w_n) \text{ for } n = 1, \dots, N, \tag{179}$$

which results from solving the IVP in the proof of Theorem 3 with the classic Runge-Kutta method of fourth order, provides us with a scheme of type II which, at least experimentally, has order 2 of convergence when it is used in FR mode (again, is not convergent in HR mode). The trade-off for using these new values of $\alpha_{n+1/2}$ is that the discrete first moment property ceases to be met. Incidentally, condition (176) is still satisfied.

8.2. Analysis of convergence

We are going to analyze schemes of type II only for $\mathscr{D}(\mu) = 1 - \mu^2$ and restricting ourselves to the case in which the numbers $\{\alpha_{n+1/2}\}_{n=0}^N$ are given by (167)–(168). Therefore, we can use Theorem 3. Thanks to the fact that we have already analyzed the convergence of schemes of type I, the task ahead will not be so complicated.

Let us start with a useful lemma.

Lemma 6. *If $\{\lambda_n\}_{n=0}^N$ is given by Definition 18, then, for $n = 1, \dots, N$,*

$$\lambda_n = \lambda_{n-1} - 2d_n^*(\mu_{n+1/2} - \mu_{n-1/2}). \tag{180}$$

Proof. By the definitions of d_n^* and $\hat{\mu}_n$,

$$2d_n^* = 2\hat{\mu}_n - 2\mu_n = \mu_{n-1/2} + \mu_{n+1/2} - 2\mu_n. \tag{181}$$

So,

$$\lambda_n = \lambda_{n-1} - 2d_n^*(\mu_{n+1/2} - \mu_{n-1/2}) \Leftrightarrow \lambda_n = \lambda_{n-1} + \{2\mu_n - (\mu_{n-1/2} + \mu_{n+1/2})\}(\mu_{n+1/2} - \mu_{n-1/2})$$
$$\Leftrightarrow \lambda_n = \lambda_{n-1} + 2\mu_n(\mu_{n+1/2} - \mu_{n-1/2}) - \{(\mu_{n+1/2})^2 - (\mu_{n-1/2})^2\}. \tag{182}$$

Noticing now that $(\mu_{n+1/2})^2 - (\mu_{n-1/2})^2 = \mathscr{D}(\mu_{n-1/2}) - \mathscr{D}(\mu_{n+1/2})$, one has

$$\lambda_n = \lambda_{n-1} - 2d_n^*(\mu_{n+1/2} - \mu_{n-1/2}) \Leftrightarrow \mathscr{D}(\mu_{n+1/2}) - \alpha_{n+1/2} = \mathscr{D}(\mu_{n-1/2}) - \alpha_{n-1/2} + 2\mu_n(\mu_{n+1/2} - \mu_{n-1/2})$$
$$- \mathscr{D}(\mu_{n-1/2}) + \mathscr{D}(\mu_{n+1/2}) \Leftrightarrow \alpha_{n+1/2} = \alpha_{n-1/2} - 2\mu_n(\mu_{n+1/2} - \mu_{n-1/2}), \tag{183}$$

which ends the proof, as the last equality is known to be true. □

The basic idea in this section is to use Lemma 6 to recast the scheme as a perturbation of a scheme of type I. Having done that, Theorem 2 solves much of the problem.

Proposition 6 *(Error representation for schemes of type II). Suppose that $f \in C^2([-1,1])$ and that $\Delta_{\mathrm{FP},N} f(\mu_n)$ is defined by Equations (162)–(164), with $\{\alpha_{n+1/2}\}_{n=0}^N$ given by Equations (167)–(168). Then, for $n = 1,\dots,N$,*

$$\Delta_{\mathrm{FP},N} f(\mu_n) = \Delta_{\mathrm{FP}} f(\mu_n) - \{R_n(f) + R_n^*(f)\}, \tag{184}$$

with $R_n(f)$ defined by Equation (93) and

$$R_1^*(f) = 2d_1^*\{E_1(f) - f'(\mu_{1+1/2})\}, \tag{185}$$

$$R_n^*(f) = \frac{\lambda_n}{2h_n^*}\{E_{n-1}(f) - E_n(f) + f'(\mu_{n+1/2}) - f'(\mu_{n-1/2})\} + 2d_n^*\{E_{n-1}(f) - f'(\mu_{n-1/2})\} \text{ for } n = 2,\dots,N-1, \tag{186}$$

$$R_N^*(f) = \left(2d_N^* - \frac{\alpha_{N+1/2}}{1-\mu_{N-1/2}}\right)\{E_{N-1}(f) - f'(\mu_{N-1/2})\}. \tag{187}$$

In the expressions above, $E_n(f)$ is that defined by Equation (69).

Proof. Let us distinguish the three possible cases.

- Case $n = 1$: use Definition 18 and Lemma 6 to see that

$$\alpha_{1+1/2} = \mathscr{D}(\mu_{1+1/2}) - \lambda_1 = \mathscr{D}(\mu_{1+1/2}) + 2d_1^*(\mu_{1+1/2} + 1). \tag{188}$$

  Then,

$$\Delta_{\mathrm{FP},N} f(\mu_1) = \frac{\alpha_{1+1/2}\frac{f(\mu_2)-f(\mu_1)}{\mu_2-\mu_1}}{\mu_{1+1/2}+1}$$
$$= \frac{\mathscr{D}(\mu_{1+1/2})\frac{f(\mu_2)-f(\mu_1)}{\mu_2-\mu_1}}{\mu_{1+1/2}+1} + 2d_1^*\frac{f(\mu_2)-f(\mu_1)}{\mu_2-\mu_1}$$
$$= \frac{\mathscr{D}(\mu_{1+1/2})\frac{f(\mu_2)-f(\mu_1)}{\mu_2-\mu_1}}{\mu_{1+1/2}+1} + 2d_1^*\{f'(\mu_{1+1/2}) - E_1(f)\}. \tag{189}$$

  Lemma 5 has been used in the last step above.
  In other words,

$$\Delta_{\mathrm{FP},N} f(\mu_1) = \Delta_{\mathrm{FP}} f(\mu_1) - \{R_1(f) + R_1^*(f)\}, \tag{190}$$

  with $R_1^*(f)$ defined by (185).
- Case $n \in \{2,\dots,N-1\}$: since, by Definition 18,

$$\alpha_{n+1/2} = \mathscr{D}(\mu_{n+1/2}) - \lambda_n \text{ and } \alpha_{n-1/2} = \mathscr{D}(\mu_{n-1/2}) - \lambda_{n-1}, \tag{191}$$

  we have

$$\Delta_{\mathrm{FP},N} f(\mu_n) = \frac{\alpha_{n+1/2}\frac{f(\mu_{n+1})-f(\mu_n)}{\mu_{n+1}-\mu_n} - \alpha_{n-1/2}\frac{f(\mu_n)-f(\mu_{n-1})}{\mu_n-\mu_{n-1}}}{\mu_{n+1/2}-\mu_{n-1/2}}$$
$$= \frac{\mathscr{D}(\mu_{n+1/2})\frac{f(\mu_{n+1})-f(\mu_n)}{\mu_{n+1}-\mu_n} - \mathscr{D}(\mu_{n-1/2})\frac{f(\mu_n)-f(\mu_{n-1})}{\mu_n-\mu_{n-1}}}{\mu_{n+1/2}-\mu_{n-1/2}} - \frac{\lambda_n\frac{f(\mu_{n+1})-f(\mu_n)}{\mu_{n+1}-\mu_n} - \lambda_{n-1}\frac{f(\mu_n)-f(\mu_{n-1})}{\mu_n-\mu_{n-1}}}{\mu_{n+1/2}-\mu_{n-1/2}}$$
$$= \Delta_{\mathrm{FP}} f(\mu_n) - R_n(f) - \frac{\lambda_n\frac{f(\mu_{n+1})-f(\mu_n)}{\mu_{n+1}-\mu_n} - \lambda_{n-1}\frac{f(\mu_n)-f(\mu_{n-1})}{\mu_n-\mu_{n-1}}}{\mu_{n+1/2}-\mu_{n-1/2}}. \tag{192}$$

Now use $\lambda_{n-1} = \lambda_n + 2d_n^*(\mu_{n+1/2} - \mu_{n-1/2})$, which holds by Lemma 6, $\mu_{n+1/2} - \mu_{n-1/2} = 2h_n^*$, and finally Lemma 5 to get

$$\Delta_{\mathrm{FP},N} f(\mu_n) = \Delta_{\mathrm{FP}} f(\mu_1) - \{R_n(f) + R_n^*(f)\}, \tag{193}$$

with $R_n^*(f)$ defined by (186).

- Case $n = N$: Equation (93) implies

$$\frac{-\mathscr{D}(\mu_{N-1/2})\frac{f(\mu_N)-f(\mu_{N-1})}{\mu_N-\mu_{N-1}}}{1-\mu_{N-1/2}} = \Delta_{\mathrm{FP}} f(\mu_N) - R_N(f), \tag{194}$$

which, used in combination with the three identities

$$\Delta_{\mathrm{FP},N} f(\mu_N) = \frac{-\alpha_{N-1/2}\frac{f(\mu_N)-f(\mu_{N-1})}{\mu_N-\mu_{N-1}}}{1-\mu_{N-1/2}}, \tag{195}$$

$$\alpha_{N-1/2} = \mathscr{D}(\mu_{N-1/2}) - \lambda_{N-1} = \mathscr{D}(\mu_{N-1/2}) - \{\lambda_N + 2d_N^*(1-\mu_{N-1/2})\}, \tag{196}$$

and

$$\lambda_N = \mathscr{D}(\mu_{N+1/2}) - \alpha_{N+1/2} = -\alpha_{N+1/2}, \tag{197}$$

ends the proof of this case. □

To properly understand the notation used in the following definition, recall that λ_n depends on $\mathscr{D}$.

Definition 20. $\beta_N^*(\mathscr{D})$ is the number defined by

$$\beta_N^*(\mathscr{D}) = \max_{2\le n\le N-1}\left|\frac{(d_{n-1}-d_n)\lambda_n}{\mu_{n+1/2}-\mu_{n-1/2}}\right| \tag{198}$$

or, equivalently,

$$\beta_N^*(\mathscr{D}) = \max_{2\le n\le N-1}\left|\frac{(d_{n-1}-d_n)\lambda_n}{2h_n^*}\right|. \tag{199}$$

Definition 21. $X_N = |\alpha_{N+1/2}/(1-\mu_{N-1/2})|$.

We are now in a position to prove the main result of this subsection.

Theorem 4 *(Order 2 of convergence for schemes of type II). Suppose that*

$$-1 = \mu_{1/2} < \mu_1 < \mu_{1+1/2} < \cdots < \mu_{N-1/2} < \mu_N < \mu_{N+1/2} = 1, \tag{200}$$

$$\widetilde{M}_N = O(N^{-1}), \tag{201}$$

$$D_N^\star = O(N^{-q}) \text{ with } q \ge 2, \tag{202}$$

$$D_N = O(N^{-r}) \text{ with } r \ge 2, \tag{203}$$

$$\frac{1}{m_N^*} = O(N^s) \text{ with } 1 \le s \le 4m-2, \text{ where } m = \min\{q,r\}, \tag{204}$$

$$\beta_N(\mathscr{D}) = O(N^{-t}) \text{ with } t \ge 2, \tag{205}$$

$$\beta_N^*(\mathscr{D}) = O(N^{-u}) \text{ with } u \ge 2, \tag{206}$$

$$X_N = O(N^{-v}) \text{ with } v \ge 2. \tag{207}$$

Then, the scheme (162)–(164), with $\{\alpha_{n+1/2}\}_{n=0}^{N}$ given by (167)–(168), converges with order 2 for any function f of class $C^5([-1,1])$, and the same is true if $[D_N = D_N^* = 0$ and the hypotheses (200), (201) and (207) hold] or if $[d_1 = \cdots = d_{N-1}$ and the hypotheses (200)–(204) and (207) hold].

Furthermore, if $s \le r$, where r and s are those in (203) and (204), the hypothesis (206) can be ignored, since it will be automatically fulfilled as a consequence of the others.

Proof. Let us start by noticing that $\Lambda_N = \max_{0\le n\le N} |\lambda_n| = O(N^{-2})$, in virtue of Theorem 3. Then,

$$\beta_N^*(\mathscr{D}) \le \frac{2D_N\Lambda_N}{m_N^*} = O(N^{s-r-2}), \tag{208}$$

and so the hypothesis (206) will indeed be automatically fulfilled if $s \le r$.

Now we will prove the main part of the theorem. Thanks to Proposition 6, we only need to prove that $\max_{1\le n\le N} |R_n^*(f)| = O(N^{-2})$, because we already know that $\max_{1\le n\le N} |R_n(f)| = O(N^{-2})$ by Theorem 2.

Notice that

$$\max_{1\le n\le N-1} |E_n(f)| = O(N^{-2}) \tag{209}$$

by Lemma 5.

- Bound for $|R_1^*(f)|$:

$$\begin{aligned} |R_1^*(f)| &= |2d_1^*\{E_1(f) - f'(\mu_{1+1/2})\}| \\ &\le 2D_N^*\left\{\left(\max_{1\le n\le N-1} |E_n(f)|\right) + \|f'\|_\infty\right\} = O(N^{-q}). \end{aligned} \tag{210}$$

- Bound for $\max_{2\le n\le N-1} |R_n^*(f)|$: let us fix $n \in \{2,\ldots,N-1\}$ and understand that $\beta_N^* = \beta_N^*(\mathscr{D})$. We know from Equation (186) that

$$R_n^*(f) = \frac{\lambda_n}{2h_n^*}\{E_{n-1}(f) - E_n(f) + f'(\mu_{n+1/2}) - f'(\mu_{n-1/2})\} + 2d_n^*\{E_{n-1}(f) - f'(\mu_{n-1/2})\} \text{ for } n = 2,\ldots,N-1. \tag{211}$$

Two parts of the expression above can be easily bounded:

$$|2d_n^*\{E_{n-1}(f) - f'(\mu_{n-1/2})\}| \le 2D_N^*\left\{\left(\max_{1\le n\le N-1} |E_n(f)|\right) + \|f'\|_\infty\right\} = O(N^{-q}) \tag{212}$$

and

$$\begin{aligned} \left|\frac{\lambda_n}{2h_n^*}\{f'(\mu_{n+1/2}) - f'(\mu_{n-1/2})\}\right| &= \left|\frac{\lambda_n}{2h_n^*}\{f'(\mu_{n-1/2}) + 2h_n^* f''(c_n) - f'(\mu_{n-1/2})\}\right| \\ &= |\lambda_n f''(c_n)| \le \Lambda_N \|f''\|_\infty = O(N^{-2}). \end{aligned} \tag{213}$$

Finding a bound for

$$\left|\frac{\lambda_n}{2h_n^*}\{E_{n-1}(f) - E_n(f)\}\right| \tag{214}$$

is in principle more difficult, but, introducing the definition

$$\tilde{\varepsilon}_n(f) = \frac{\lambda_n}{2h_n^*}\{E_{n-1}(f) - E_n(f)\}, \tag{215}$$

noting the resemblance of $\tilde{\varepsilon}_n(f)$ to $\varepsilon_n(f)$ in Equation (96), and using the same ideas than those in the proof of Proposition 5 (with $\mathscr{D} \equiv 1$), one gets

$$\tilde{\varepsilon}_n(f) = \tilde{\varepsilon}_n^{(1)}(f) + \tilde{\varepsilon}_n^{(2)}(f) + \tilde{\varepsilon}_n^{(3)}(f) + \tilde{\varepsilon}_n^{(4)}(f), \tag{216}$$

with

$$|\tilde{\varepsilon}_n^{(1)}(f)| \le \beta_N^* \|f''\|_\infty + D_N\Lambda_N \|f'''\|_\infty, \tag{217}$$

$$|\tilde{\varepsilon}_n^{(2)}(f)| \le \frac{\beta_N^*(3D_N + C_N)}{3}\|f'''\|_\infty + \frac{C_N\Lambda_N}{3}\|f'''\|_\infty + \frac{(M_N^2 + 12D_N^2)\Lambda_N}{24}\|f^{4)}\|_\infty, \tag{218}$$

$$|\tilde{\varepsilon}_n^{(3)}(f)| \le \frac{\beta_N^*\{(M_N^*)^2 + 8C_ND_N + 4C_N^2 + 12D_N^2\}}{24}\|f^{4)}\|_\infty + \frac{C_ND_N\Lambda_N}{3}\|f^{4)}\|_\infty + \frac{(M_N^2D_N + 4D_N^3)\Lambda_N}{24}\|f^{5)}\|_\infty, \tag{219}$$

$$|\tilde{\varepsilon}_n^{(4)}(f)| \le \frac{Z_N\Lambda_N}{960}\|f^{5)}\|_\infty, \tag{220}$$

Table 6
Numerical results for Morel's scheme operated in FR mode. $f(\mu) = e^{\mu}$, $\mathscr{D}(\mu) = 1 - \mu^2$.

N	E	order	q	r	s	t	u
50	6.94×10^{-3}						
100	1.76×10^{-3}	1.98	1.98	1.98	1.99	1.99	3.97
500	7.14×10^{-5}	1.99	1.99	1.99	1.99	2.00	3.99
1000	1.79×10^{-5}	2.00	2.00	2.00	2.00	2.00	4.00
5000	7.16×10^{-7}	2.00	2.00	2.00	2.00	2.00	4.00
10000	1.86×10^{-7}	1.94	2.00	2.00	2.00	2.00	4.00
20000	8.64×10^{-8}	1.11	2.00	2.00	2.00	2.00	4.00

where Z_N is given by Equation (111).
Now, recalling that $C_N = D_N + D_N^* = O(N^{-m})$, it is clear that

$$\max_{2\le n\le N-1} |R_n^*(f)| = O(N^{-2}). \tag{221}$$

- Bound for $|R_N^*(f)|$:

$$\begin{aligned} |R_N^*(f)| &= \left| \left(2d_N^* - \frac{\alpha_{N+1/2}}{1-\mu_{N-1/2}} \right) \{E_{N-1}(f) - f'(\mu_{N-1/2})\} \right| \\ &\le \left(2D_N^* + X_N\right) \left\{ \left(\max_{1\le n\le N-1} |E_n(f)| \right) + \|f'\|_\infty \right\} \\ &= O(N^{-\min\{q,v\}}). \end{aligned} \tag{222}$$

In summary, if $f \in C^5([-1,1])$,

$$\max_{1\le n\le N} |R_n(f) + R_n^*(f)| = O(N^{-2}), \tag{223}$$

and so the scheme converges with order 2. The statements that remain to be proved follow easily. □

8.3. Application of the theory to some examples. Numerical results

Recalling Equation (156), E will denote the maximum of the absolute values of the errors in the complete set of nodes.

Morel's scheme For this scheme, operated in FR mode, the hypotheses (200)–(205) are met; the justifications given for the Haldy-Ligou's scheme are also valid for this one. Moreover, as $r = s$, the hypothesis (206) is automatically satisfied, while the last hypothesis (207) also holds because $\alpha_{N+1/2} = 0$.

Thus, according to Theorem 4, Morel's scheme in FR mode is expected to converge with order 2. Numerical results in agreement with the theoretical prediction are displayed in Table 6, the rows of which stop at the moment where roundoff errors start to spoil the approximation.

Like Haldy-Ligou's scheme, Morel's does not converge when used in HR mode, and the reason is the same: the hypothesis (205) is not fulfilled. The numerical and graphical results are very similar to those of Haldy-Ligou's and are therefore omitted.

And again as it happened for the Haldy-Ligou's scheme, the non-convergent nature of Morel's scheme in HR mode does not necessarily preclude its utility for solving the FPE. This will be stressed in Section 9.

Remark 11. We know from Theorem 3 that $\Lambda_N = O(N^{-2})$ under hypotheses (200)–(202). In [11], theoretical support for the same identity $\Lambda_N = O(N^{-2})$ is provided in a direct manner, without relying on these hypotheses, as long as the $\alpha_{n+1/2}$ are those of Morel's scheme.

Uniform mesh (a non-GL scheme of type II and order 2) Let us take, as $\{\mu_n\}_{n=1}^N$ and $\{\mu_{n+1/2}\}_{n=0}^N$, the uniformly spaced sets that we took when defining the scheme of type I and of order 2 in Subsection 7.4 (uniform mesh). Then, the corresponding scheme of type II satisfies $D_N = D_N^* = \alpha_{N+1/2} = 0$, which implies convergence of order 2 according to Theorem 4. In fact, it is easily verified that in this case we have $\alpha_{n+1/2} = \mathscr{D}(\mu_{n+1/2})$ for all $n = 0, \dots, N$, so this scheme turns out to be precisely the scheme of type I just mentioned.

9. On the applicability of non-convergent schemes

Ganapol and López Pouso noted that Morel's scheme performs better in FR mode than in HR mode when using the RM/DOM method introduced in [2].[2] So, one might be tempted to assert that the aforementioned non-convergent schemes should always be avoided, but this conclusion would be misleading.

Given that this paper focuses on the convergence of numerical differentiation formulas for the FP Laplacian, our study becomes a part of the *consistency* analysis of schemes that fully discretize the FPE if they perform the μ-discretization by means of one of these formulas. Since consistency is not necessary for convergence, one can wonder whether, let us say, Haldy-Ligou's or Morel's discretizations can still be used in HR mode when solving the FPE. The answer is affirmative: if the scheme described in [9] is adapted by altering the discretization with respect to μ for either Haldy-Ligou's or Morel's scheme, then the numerical experiments show a convergent behavior regardless of whether FR or HR mode is implemented.

The existence of non-consistent but convergent schemes for solving PDEs, as well as schemes having a consistency order lower than their convergence order, has been known for a long time. Interested readers may consult references [4], [7], [12], and [17].

The order of convergence in the variable μ is expected to be 2 when using a second-order numerical differentiation formula, such as Morel's, to discretize the FP Laplacian. However, it is important to note that previous studies (see graphs in references [2], [6], and [9]) suggest that the solution to the FPE can be non-differentiable. In such cases, using second-order approximations for the differential operators, such as Morel's scheme, does not necessarily provide an advantage over lower-order approximations, as the lack of regularity in the solution prevents achieving second-order accuracy.

10. Conclusions

Widely recognized difference schemes for discretizing the FP angular diffusion operator have been incorporated into a comprehensive framework, which has undergone thorough analysis. This analysis has allowed us to derive sets of sufficient conditions that guarantee the convergence with second-order accuracy for the schemes falling into the two categories defined in this work: type I and type II schemes.

By applying these general results, the study provides theoretical evidence supporting second-order convergence of Haldy-Ligou's and Morel's schemes when they are operated in FR mode. Moreover, the study highlights that Haldy-Ligou's and Morel's schemes do not exhibit convergence when operated in HR mode. The importance of this observation when solving the FPE depends on the method employed: numerical experiments show that HR mode should be avoided when using the RM/DOM method described in [2], but it works as well as FR mode when adapting the method in [9] to use GL nodes instead of a uniform mesh. A theoretical analysis of this phenomenon would be interesting.

Regarding the experiments conducted with the method described in [9], up to this point, we have limited ourselves to the even scheme, which excludes the $\mu = 0$ node. An idea for future research is the development of an odd scheme that utilizes GL nodes.

Another idea is to find a method for accurately computing the gaps between GL nodes appearing in the schemes (referring to the differences $\mu_{n+1} - \mu_n$ and $\mu_n - \mu_{n-1}$ in Equations (90) and (177)), as this might help reduce roundoff errors. It is noteworthy to mention that this necessity was already observed in another context by Dirk Laurie in [8, Subsection 6.3].

Lastly, this research has uncovered new properties of GL nodes and weights. The analysis of these properties, which necessitates the use of specialized techniques, is conducted in [11]. It would be interesting to see if analogous results can be obtained for Gauss-Lobatto and Gauss-Radau quadratures, as full range Gauss-Lobatto sets and half range Gauss-Radau sets are also widely used in the field of nuclear engineering. For both, the endpoints -1 and 1 become nodes, which can also be achieved with the schemes described in reference [9].

Funding

OLP acknowledges support from Ministerio de Ciencia e Innovación, project PID2021-122625OB-I00 with funds from

MCIN/AEI/10.13039/501100011033/ ERDF, EU,

and from the Xunta de Galicia (2021 GRC Gl-1563 - ED431C 2021/15).

JS acknowledges support from Ministerio de Ciencia e Innovación, project PID2021-127252NB-I00 with funds from

MCIN/AEI/10.13039/501100011033/ ERDF, EU.

Acknowledgements

The authors are grateful to Prof. Barry Ganapol from the Aerospace and Mechanical Department at the University of Arizona for his interest in this work and helpful advice after carefully reviewing parts of the paper. Furthermore, we appreciate his efforts in conducting new numerical experiments specifically for us, which corroborate that the RM/DOM method described in [2] is less efficient when implemented in HR mode compared to FR mode.

[2] Excerpted from [2]: 'After experimentation, our choice will be FRLGQ. Apparently, HRLGQ, while appropriate for the neutron transport, performs poorly for the FPE'.

Appendix A. List of acronyms

The following acronyms are used in this paper:

- DOM: discrete ordinates method.
- FP, FPE: Fokker-Planck, Fokker-Planck equation.
- FR: full range (meaning that the GL nodes in $(-1, 1)$ are used).
- GL: Gauss-Legendre.
- HR: half range (meaning that the union of the GL nodes in $(-1, 0)$ and in $(0, 1)$ are used).
- PDE: partial differential equation.

Data availability

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

References

[1] Michael J. Antal, Clarence E. Lee, Charged particle mass and energy transport in a thermonuclear plasma, J. Comput. Phys. 20 (3) (1976) 298–312, https://doi.org/10.1016/0021-9991(76)90083-8.
[2] Barry Ganapol, Óscar López Pouso, Response matrix/discrete ordinates solution of the 1D Fokker-Planck equation, Nucl. Sci. Eng. 197 (9) (2023) 2327–2342, https://doi.org/10.1080/00295639.2023.2194228.
[3] Amparo Gil, Javier Segura, Nico M. Temme, Fast and reliable high-accuracy computation of Gauss-Jacobi quadrature, Numer. Algorithms 87 (4) (2021) 1391–1419, https://doi.org/10.1007/s11075-020-01012-6.
[4] Rolf Dieter Grigorieff, Some stability inequalities for compact finite difference schemes, Math. Nachr. 135 (1) (1988) 93–101, https://doi.org/10.1002/mana.19881350110.
[5] Pierre-André Haldy, Jacques Ligou, A multigroup formalism to solve the Fokker-Planck equation characterizing charged particle transport, Nucl. Sci. Eng. 74 (3) (1980) 178–184, https://doi.org/10.13182/NSE80-A20117.
[6] Arnold D. Kim, Paul Tranquilli, Numerical solution of the Fokker-Planck equation with variable coefficients, J. Quant. Spectrosc. Radiat. Transf. 109 (5) (2008) 727–740, https://doi.org/10.1016/j.jqsrt.2007.09.011.
[7] Heinz-Otto Kreiss, Thomas A. Manteuffel, Blair Swartz, Burton Wendroff, Andrew B. White Jr., Supra-convergent schemes on irregular grids, Math. Comput. 47 (176) (1986) 537–554, https://doi.org/10.1090/S0025-5718-1986-0856701-5.
[8] Dirk P. Laurie, Computation of Gauss-type quadrature formulas, in: Walter Gautschi, Francisco Marcellán, Lothar Reichel (Eds.), Special Issue on Numerical Analysis 2000, vol. V, Quadrature and Orthogonal Polynomials, J. Comput. Appl. Math. 127 (1–2) (2001) 201–217, https://doi.org/10.1016/S0377-0427(00)00506-9.
[9] Óscar López Pouso, Nizomjon Jumaniyazov, Numerical experiments with the Fokker-Planck equation in 1D slab geometry, J. Comput. Theor. Transp. 45 (3) (2016) 184–201, https://doi.org/10.1080/23324309.2016.1150856.
[10] Óscar López Pouso, Nizomjon Jumaniyazov, Numerical solution of the azimuth-dependent Fokker-Planck equation in 1D slab geometry, J. Comput. Theor. Transp. 50 (2) (2021) 102–133, https://doi.org/10.1080/23324309.2021.1896554.
[11] Óscar López Pouso, Javier Segura, Uniform relations between the Gauss-Legendre nodes and weights, Preprint https://doi.org/10.48550/arXiv.2305.19128, submitted for publication.
[12] Thomas A. Manteuffel, Andrew B. White Jr., The numerical solution of second-order boundary value problems on nonuniform meshes, Math. Comput. 47 (176) (1986) 511–535, https://doi.org/10.1090/S0025-5718-1986-0856700-3.
[13] Thomas A. Mehlhorn, James J. Duderstadt, A discrete ordinates solution of the Fokker-Planck equation characterizing charged particle transport, J. Comput. Phys. 38 (1) (1980) 86–106, https://doi.org/10.1016/0021-9991(80)90013-3.
[14] Jim E. Morel, An improved Fokker-Planck angular differencing scheme, Nucl. Sci. Eng. 89 (2) (1985) 131–136, https://doi.org/10.13182/NSE85-A18187.
[15] Edgar Olbrant, Martin Frank, Generalized Fokker-Planck theory for electron and photon transport in biological tissues: application to radiotherapy, Comput. Math. Methods Med. 11 (4) (2010) 313–339, https://doi.org/10.1080/1748670X.2010.491828.
[16] Japan K. Patel, James S. Warsa, Anil Kant Prinja, Accelerating the solution of the S_N equations with highly anisotropic scattering using the Fokker-Planck approximation, Ann. Nucl. Eng. 147 (2020) 107665, https://doi.org/10.1016/j.anucene.2020.107665.
[17] Jesús María Sanz Serna, Stability and convergence in numerical analysis I: linear problems, a simple, comprehensive account, in: Jack Kenneth Hale, Pedro Martínez Amores (Eds.), Nonlinear Differential Equations and Applications, Pitman, London, 1985, pp. 64–113.
[18] Gabor Szegö, Orthogonal Polynomials, 4th edition, American Mathematical Society, Providence, Rhode Island, 1975, 1st edition published in 1939.
[19] James S. Warsa, Anil Kant Prinja, A moment-preserving S_N discretization for the one-dimensional Fokker-Planck equation, Trans. Am. Nucl. Soc. 106 (1) (2012) 362–365.