

Uniform and high-order discretization schemes for Sturm–Liouville problems via Fer streamers

Alberto Gil C. P. Ramos

the date of receipt and acceptance should be inserted later

Abstract The current paper concerns the uniform and high-order discretization of the novel approach to the computation of Sturm–Liouville problems via Fer streamers, put forth in [Ramos and Iserles, Numer. Math., 2015, 131: 541, doi:10.1007/s00211-014-0695-0]. In particular, the discretization schemes are shown to enjoy large step sizes uniform over the entire eigenvalue range and tight error estimates uniform for every eigenvalue. They are made explicit for global orders 4, 7, 10. In addition, the present paper provides total error estimates that quantify the interplay between the truncation and the discretization in the approach by Fer streamers.

Keywords Numerical method · Regular Sturm–Liouville problems · Liouville’s normal form · Continuous and piecewise analytic potential · Self-adjoint separated boundary conditions · Fer expansions · Fer streamers

Mathematics Subject Classification (2000) 65L15 · 65L10 · 65L70 · 34C40

1 Introduction

Regular Sturm–Liouville problems in Liouville’s normal form, with continuous and piecewise analytic potentials

$$\begin{aligned} -y''_{\lambda}(t) + q(t)y_{\lambda}(t) &= \lambda y_{\lambda}(t), \quad t \in [a, b], \quad a, b \in \mathbb{R}, \quad \lambda \in \mathbb{R}, \\ q \in C^0([a, b] \rightarrow [q_{\min}, q_{\max}]) &\text{ is piecewise analytic, } y_{\lambda} \in C^2([a, b] \rightarrow \mathbb{R}), \end{aligned} \quad (1)$$

and self-adjoint separated boundary conditions

$$\begin{aligned} \alpha_1 y_{\lambda}(a) + \alpha_2 y'_{\lambda}(a) &= 0, \quad \alpha_1, \alpha_2 \in \mathbb{R}, \quad \alpha_1^2 + \alpha_2^2 > 0, \\ \beta_1 y_{\lambda}(b) + \beta_2 y'_{\lambda}(b) &= 0, \quad \beta_1, \beta_2 \in \mathbb{R}, \quad \beta_1^2 + \beta_2^2 > 0, \end{aligned} \quad (2)$$

are ubiquitous in applications, and it is of great interest to analyse, discretize and implement numerical methods to compute their eigenvalues and eigenfunctions (λ, y_{λ}) .

Alberto Gil C. P. Ramos
Cambridge Centre for Analysis, University of Cambridge, United Kingdom
E-mail: albertogilramos@cantab.net

The current paper concerns the discretization of the novel approach to the computation of Sturm–Liouville problems via Fer streamers, introduced in [14].

The motivation to discretize the new approach via Fer streamers stems from the local and global error estimates in [14] which guarantee large step sizes uniform over the entire eigenvalue range and tight error estimates uniform for every eigenvalue. The current paper shows how to retain these advantageous features under discretization.

1.1 Uniform versus asymptotic error estimates

It is well known that (1)–(2) possess a unique countable family of solutions $\{\lambda_j, y_{\lambda_j}\}_{j \in \mathbb{Z}_0^+}$ and that its eigenvalues are simple, bounded from below and accumulate only at infinity as well as that its eigenfunctions oscillate as functions of the eigenvalues [13], namely

$$\lambda_j < \lambda_{j+1}, \quad \lim_{j \rightarrow +\infty} (\lambda_j / j^2) = (\pi / (b - a))^2, \quad y_{\lambda_j} \text{ has exactly } j \text{ zeros in } (a, b). \quad (3)$$

Moreover, a representation of the (λ, y_λ) is given by the solution of the initial value problem

$$\begin{aligned} \mathbf{Y}'_\lambda(t) &= \begin{bmatrix} 0 & 1 \\ q(t) - \lambda & 0 \end{bmatrix} \mathbf{Y}_\lambda(t), \quad t \in [a, b], \quad a, b \in \mathbb{R}, \quad \lambda \in \mathbb{R}, \\ q &\in C^0([a, b] \rightarrow [q_{\min}, q_{\max}]) \text{ is piecewise analytic, } \mathbf{Y}_\lambda : [a, b] \rightarrow \mathbb{R}^{2 \times 2}, \end{aligned} \quad (4)$$

with initial condition

$$\mathbf{Y}_\lambda(a) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad (5)$$

which characterises the eigenvalues as roots and the eigenfunctions via the transition matrix

$$\{\lambda_j\}_{j \in \mathbb{Z}_0^+} = \left\{ \lambda \in \mathbb{R} : \det \left(\begin{bmatrix} \alpha_1 & \alpha_2 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ \beta_1 & \beta_2 \end{bmatrix} \mathbf{Y}_\lambda(b) \right) = 0 \right\}, \quad \begin{bmatrix} y_\lambda(c) \\ y'_\lambda(c) \end{bmatrix} = \mathbf{Y}_\lambda(c) \begin{bmatrix} y_\lambda(a) \\ y'_\lambda(a) \end{bmatrix}. \quad (6)$$

To compute (λ, y_λ) one can then for instance set up a mesh with $m \in \mathbb{Z}^+$ intervals

$$c_0 := a < c_1 < \dots < c_m := b, \quad h_k := c_{k+1} - c_k, \quad h_{\min} := \min_k \{h_k\}, \quad h_{\max} := \max_k \{h_k\}, \quad (7)$$

and approximate $(\lambda, c_k) \mapsto \mathbf{Y}_\lambda(c_k)$ by integral series techniques, which require truncation of the series and discretization of the integrals.

The Fer streamers approximation to $(\lambda, t) \mapsto \mathbf{Y}_\lambda(t)$ in the truncation of the integral series in the lead paper [14] and in the discretization of the multivariate integrals in this paper, is virtually unique in the literature because it is based on the uniform regime

$$h_{\max} \rightarrow 0^+, \text{ uniformly w.r.t. } \lambda \in [q_{\max} - h_{\max}^{-2}, +\infty). \quad (8)$$

The only partial exception known to the author is the use of Magnus expansions in [11] which yields a method with global order four based on the bounded uniform regime

$$h \rightarrow 0^+, \text{ uniformly w.r.t. } \lambda \in [-h^{-2}, h^{-2}].$$

The theory based on the uniform regime (8) is very different from the theory based on the asymptotic regimes common throughout the literature

$$\lambda \text{ fixed and } h \rightarrow 0^+, \quad (9)$$

$$h \text{ fixed and } \lambda \rightarrow +\infty, \quad (10)$$

e.g., in the piecewise perturbation methods [7], in the right-correction Magnus series [2] and in the modified Magnus methods [9].

In addition, one can consider another regime. Namely, there are well known asymptotic formulas for λ_j/j^2 in $j \rightarrow +\infty$, of which the leading term is given in (3), which can attain high order for ‘large’ eigenvalues [13].

The main difference of the uniform regime (8) over the asymptotic regimes (9)–(10) is that the former yields approximations which hold uniformly in (λ, h) , while the latter only controls errors in one variable at a time in either h (with fixed λ) or λ (with fixed h).

This has important consequences for the size of the constants in the big \mathcal{O} notation in the approximation estimates as well as their sampling requirements and their ranges of validity. These manifest in four ways: Firstly, (9)–(10) often lead to truncation estimates with ‘large’ constants in the big \mathcal{O} notation for ‘intermediary’ eigenvalues. Secondly, (9)–(10) usually result in more function evaluations for ‘intermediary’ eigenvalues. Thirdly, (9) can lead to quadrature estimates with ‘large’ constants in the big \mathcal{O} notation, which are applicable only to ‘small’ eigenvalues. Fourthly, (10) is usually used with highly oscillatory quadrature, which is applicable only to ‘large’ eigenvalues and non-resonant integrals.

To illustrate the first point above, note the increase in global order from 6 to 8, legitimate under (9), in the truncation of the integral series in [2, Theorem 2] and [9, Theorem 4.2], is built upon Taylor expansions of oscillatory integrals. Thus, the $\mathcal{O}(h^{10})$ terms in [1, p. 35] and [9, p. 759] grow with λ . As a consequence, the constants in the truncation estimates grow with λ and are therefore applicable only to ‘small’ eigenvalues.

As for the second point above, the localized increase in function evaluations outside of (9)–(10) can be traced to the quadrature estimates in [4] for the integral (which appears in approximations for global order greater than or equal to four)

$$h \int_0^1 \mathbf{A}_\lambda(ht) e^{it\varpi_{\lambda,h}} dt, \quad \varpi_{\lambda,h} := 2h \sqrt{\lambda - \frac{\int_c^{c+h} q(\xi) d\xi}{h}},$$

which vary for non-oscillatory $|\varpi_{\lambda,h}| \ll 1$, highly-oscillatory $\varpi_{\lambda,h} \gg 1$ and intermediate $|\varpi_{\lambda,h}| \approx 1$ regimes, where the first two arise from (9)–(10), and the last arises from ‘intermediary’ (λ, h) . With p quadrature points, quadrature estimates read [4]: $\mathcal{O}(h^{2p-1})$ in (9), $\mathcal{O}(h^{p+1}/\varpi_{\lambda,h}^2)$ in (10) and $\mathcal{O}(h^{p+1})$ for ‘intermediary’ (λ, h) . The ratio $(p+1)/(2p-1)$ then quantifies a 50% increase in sampling required for ‘intermediary’ eigenvalues.

Ref. [4] also clarifies the third point above since it is precisely to account for ‘large’ constants in the big \mathcal{O} notation that the aforementioned estimates for $|\varpi_{\lambda,h}| \approx 1$ are worse than the ones for $|\varpi_{\lambda,h}| \ll 1$.

For the fourth point above, we note that quadrature estimates that rely on (10) often arise from asymptotic expansions in $\varpi_{\lambda,h} \gg 1$ [4]. However, since for fixed $\varpi_{\lambda,h}$, smaller h leads to larger λ , these estimates are only valid for ‘large’ eigenvalues. In addition, quadrature estimates under $\varpi_{\lambda,h} \gg 1$ are often valid only in the absence of critical points and subject to a non-resonance condition [6]. In the context of the integral series in [7, 2, 9, 14] and this paper, these considerations are key since the non-resonance condition is not satisfied in the integral (which appears in higher order approximations)

$$h^2 \int_0^1 \int_0^{t_1} \left[\mathbf{A}_\lambda(ht_2), \overline{\mathbf{A}_\lambda(ht_1)} \right] e^{i(t_2-t_1)\varpi_{\lambda,h}} dt_2 dt_1.$$

Working with (8) rather than (9)–(10), Fer streamers bypass such issues and control the truncation and discretization of Sturm–Liouville problems with error bounds which hold with large step sizes equally well for all ‘small’, ‘intermediary’ and ‘large’ eigenvalues.

1.2 Geometric integration

The Lie-group and Lie-algebra

$$\mathrm{SL}(2, \mathbb{R}) := \{\mathbf{U} \in \mathbb{R}^{2 \times 2} : \det(\mathbf{U}) = 1\}, \quad \mathfrak{sl}(2, \mathbb{R}) := \{\mathbf{V} \in \mathbb{R}^{2 \times 2} : \mathrm{tr}(\mathbf{V}) = 0\},$$

play a natural role in the qualitative approximation of the solution of (4)–(5) given that

$$\mathbf{Y}_\lambda([a, b]) \subseteq \mathrm{SL}(2, \mathbb{R}).$$

We note that the Fer streamers approximations and the methods in [2, 9] preserve this geometric property, unlike the methods in [7].

1.3 Computational complexity

We now discuss how the computational complexity of Fer streamers compares with other geometric integration techniques.

1.3.1 Number of steps in each numerical mesh

The truncation in [14] and the discretization in the current paper are based on the following:

Assumption 1 One has that q is continuous in $[a, b]$, $q|_{[c_k^+, c_{k+1}^-]}$ is analytic in $[c_k, c_{k+1}]$ and

$$\lambda \geq q_{\min} \implies h_{\max} \leq 1/\sqrt{q_{\max} - q_{\min}}, \quad (11)$$

$$\lambda < q_{\min} \implies h_{\max} \leq 1/\sqrt{q_{\max} - \lambda}, \quad (12)$$

$$\frac{h_{\max}}{h_{\min}} \leq 2 \text{ (this constant can be increased)}. \quad (13)$$

Assumption (12) prevents ‘large’ constants in the big \mathcal{O} notation in the error estimates. In detail, if $\lambda < q_{\min}$ then the argument of certain hyperbolic cosines and sines is positive. If left unchecked, the argument becomes unbounded and the hyperbolic functions grow exponentially with the size of the argument. Assumption (12) guarantees that the positive argument and the hyperbolic functions are bounded by a small constant. Since [2, p. 423] assumes that $\lambda \gg q_{\max}$, this issue does not arise. Ref. [9] does not assume $\lambda \geq q_{\min}$, and disregards this issue. As discussed in [14], there are Sturm–Liouville problems where (12) is automatically satisfied since there do not exist eigenvalues smaller than q_{\min} .

Assumption (11) enables an unhindered transition of the error estimates between the uniform regime (8). In addition, it quantifies the impact of the magnitude of the potential to the Fer streamers approach to Sturm–Liouville problems. The fact that the scale of the potential influences the step size is noted, but not quantified, in [2, p. 416] and [9, p. 761].

We note that assumptions (11)–(12) use the knowledge of a lower bound to the minimum of the potential and an upper bound to its maximum. In [2] the knowledge of q_{\max} is required since it focuses on $\lambda \gg q_{\max}$. Ref. [9] does not use this information, since it does not control the positive argument of certain hyperbolic functions for $\lambda < q_{\min}$.

Assumption (13) controls the non-uniformity of the mesh, which is intrinsically related to the size of the constants in the big \mathcal{O} notation in the error estimates.

1.3.2 Error estimates and evaluations of the potential

The discretization in this paper with global order 4, 7, 10 with respect to (8) requires 3, 6, 9 (one at the right boundary and the rest in the interior) evaluations of the potential per mesh interval. Given the potential is continuous, this translates to $3m + 1$, $6m + 1$, $9m + 1$ evaluations of the potential for a mesh with m intervals.

The discretization in [2, p. 422–429] with global order 4, 8 with respect to the regime (9) requires 2, 4 (interior) potential evaluations per mesh interval, which corresponds to $2m$, $4m$ potential evaluations, for a mesh with m intervals. It is also shown in [2, Eq. 48] that the global error in that work is bounded in the regime (10).

The work by [9] extends the work by [2] from $\lambda \gg q_{\max}$ to $\lambda \leq q_{\max}$ and suggests different potential evaluations. For global order 4, 8 it uses 3, 5 (one at each boundary and the rest in the interior) potential evaluations per mesh interval. Since the potential is continuous, this results in $2m + 1$, $4m + 1$ potential evaluations, for a mesh with m intervals.

As discussed in Subsection 1.1, the 50% increase in function evaluations from (9)–(10) to (8), is necessary for quadrature over ‘intermediary’ eigenvalues.

1.3.3 Amount of linear algebra

The discretization in this paper boils down to the quadrature of integrals of the form

$$h \int_0^1 \mathbf{Z}_\lambda(ht) dt, \quad (14)$$

$$h^2 \int_0^1 \int_0^{t_1} [\mathbf{Z}_\lambda(ht_2), \mathbf{Z}_\lambda(ht_1)] dt_2 dt_1, \quad (15)$$

$$h^3 \int_0^1 \int_0^{t_1} \int_0^{t_1} [\mathbf{Z}_\lambda(ht_3), [\mathbf{Z}_\lambda(ht_2), \mathbf{Z}_\lambda(ht_1)]] dt_3 dt_2 dt_1, \quad (16)$$

where $t \mapsto \mathbf{Z}_\lambda(ht)$ possesses a plethora of behaviour that varies with (λ, h) . The quadrature schemes in this paper are based on uniform approximations of $\mathbf{Z}_\lambda(ht)$ in (λ, t) by $\tilde{\mathbf{Z}}_{\lambda,h}(t)$ with the property that (14)–(16), with $\mathbf{Z}_\lambda(h \cdot)$ replaced by $\tilde{\mathbf{Z}}_{\lambda,h}(\cdot)$, can be integrated exactly. Thus, the amount of linear algebra in the discretization schemes can be quantified by the number of terms in each integrand in (14)–(16) with $\tilde{\mathbf{Z}}_{\lambda,h}(\cdot)$ instead of $\mathbf{Z}_\lambda(h \cdot)$, which grows exponentially with i) base equal to the product between the number of summands in each representation of $\mathbf{Z}_\lambda(ht)$ times the number of quadrature points, and, ii) exponent equal to the number of commutators in each integrand plus one. Fortunately, the exponential growth of the number of terms in each integrand is heavily attenuated in the quadrature in this paper, since it requires less sampling points for the higher dimensional integrals than for the lower dimensional integrals, which represents a significant saving in linear algebra. In detail, in the sense of (8), global order:

- four requires $\tilde{\mathbf{Z}}_{\lambda,h}(t)$ with 3 sampling points for (14),
- seven requires $\tilde{\mathbf{Z}}_{\lambda,h}(t)$ with 6 sampling points for (14) and 3 for (15),
- ten requires $\tilde{\mathbf{Z}}_{\lambda,h}(t)$ with 9 sampling points for (14), 6 for (15) and 3 for (16).

The discretization in [2, 9] with global order 4, 8 with respect to (9) requires the quadrature, with different $\mathbf{Z}_\lambda(ht)$, of (14) and (14)–(15). In detail, in [2] global order 4, 8 requires $\tilde{\mathbf{Z}}_{\lambda,h}(t)$ with 2, 4 interpolation points for every integral, whereas in [9] global order 4, 8 instead uses 3, 5 interpolation points for every integral.

In particular, [2, 9] do not enjoy the heavy attenuation of the exponential growth of the number of terms in each integrand described above for the quadrature schemes in this paper.

2 Sturm–Liouville problems via Fer streamers and truncation error estimates

The Fer streamers approach to Sturm–Liouville problems introduced in [14] is built around the Fer expansions representation of the solution of (4)–(5) given by the integral series in Theorem 1 below, which relate to the eigensystem of (1)–(2) through (6).

Definition 1 Let $\mathbf{X}, \mathbf{Y} \in \mathfrak{sl}(2, \mathbb{R})$, $l \in \mathbb{Z}^+$ and $t \in [c_k, c_{k+1}]$, and define

$$\begin{aligned} e^{\mathbf{X}} &:= \exp(\mathbf{X}) := \cosh \frac{\rho(\mathbf{X})}{2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \frac{\sinh \frac{\rho(\mathbf{X})}{2}}{\frac{\rho(\mathbf{X})}{2}} \mathbf{X}, & \rho(\mathbf{X}) &:= 2\sqrt{-\det(\mathbf{X})}, \\ \text{ad}_{\mathbf{X}} \mathbf{Y} &:= [\mathbf{X}, \mathbf{Y}] := \mathbf{XY} - \mathbf{YX}, & \text{Ad}_{\exp(\mathbf{X})} \mathbf{Y} &:= e^{\mathbf{X}} \mathbf{Y} e^{-\mathbf{X}}, \\ \mathbf{B}_{\lambda,0}(c_k, t) &:= \begin{bmatrix} 0 & 1 \\ q(t) - \lambda & 0 \end{bmatrix}, & \mathbf{D}_{\lambda,0}(c_k, t) &:= \int_{c_k}^t \mathbf{B}_{\lambda,0}(c_k, \xi) d\xi, \\ \mathbf{B}_{\lambda,l}(c_k, t) &:= \sum_{j=1}^{\infty} \frac{(-1)^j j}{(j+1)!} \text{ad}_{\mathbf{D}_{\lambda,l-1}(c_k, t)}^j \mathbf{B}_{\lambda,l-1}(c_k, t), & \mathbf{D}_{\lambda,l}(c_k, t) &:= \int_{c_k}^t \mathbf{B}_{\lambda,l}(c_k, \xi) d\xi. \end{aligned}$$

Theorem 1 ([3]) *The solution of (4)–(5) is given by the Fer expansions flow*

$$\mathbf{F}_{\lambda}(c_k, c_{k+1}) := e^{\mathbf{D}_{\lambda,0}(c_k, c_{k+1})} e^{\mathbf{D}_{\lambda,1}(c_k, c_{k+1})} \dots, \quad \mathbf{Y}_{\lambda}(c_{k+1}) = \mathbf{F}_{\lambda}(c_k, c_{k+1}) \cdots \mathbf{F}_{\lambda}(c_0, c_1).$$

In practice, integral series such as Magnus, Fer or Neumann expansions, require truncation of the series and discretization of the integrals, which with Sturm–Liouville problems are challenging since the solution of (4)–(5) is often exponentially large or highly oscillatory.

One of the contributions in [14] is to provide a reinterpretation of Fer expansions, which bypasses such issues, and opens the door to approximate $(\lambda, t) \mapsto \mathbf{Y}_{\lambda}(t)$ uniformly and to high order in (8), equally well throughout large, oscillatory and in-between cases.

The first insight in [14] is to use Lie-group/Lie-algebra techniques to rewrite the infinite series in Definition 1, which appear in Fer expansions and are hard to control, by amenable closed-form expressions, named ‘Fer streamers’, given below in Theorem 2.

Definition 2 Let $\mathbf{X} \in \mathfrak{sl}(2, \mathbb{R})$ and $z \in \mathbb{C}$, and define

$$\pi(\mathbf{X}) := \begin{bmatrix} [\mathbf{X}]_{1,1} & [\mathbf{X}]_{1,2} & [\mathbf{X}]_{2,1} \end{bmatrix}^{\top}, \quad \varphi(z) := \frac{\cosh(z) - 1}{z^2} - \frac{\sinh(z)}{z}, \quad \phi(z) := \frac{\cosh(z) - \frac{\sinh(z)}{z}}{z^2}.$$

Theorem 2 ([14]) *If $l \in \mathbb{Z}^+$ and $t \in [c_k, c_{k+1}]$, then the infinite series in Definition 1 for the Fer expansions in Theorem 1 are given in closed-form by the ‘Fer streamers’*

$$\begin{aligned} \mathbf{B}_{\lambda,l}(c_k, t) &= \varphi(\rho(\mathbf{D}_{\lambda,l-1}(c_k, t))) \text{ad}_{\mathbf{D}_{\lambda,l-1}(c_k, t)} \mathbf{B}_{\lambda,l-1}(c_k, t) \\ &\quad + \phi(\rho(\mathbf{D}_{\lambda,l-1}(c_k, t))) \text{ad}_{\mathbf{D}_{\lambda,l-1}(c_k, t)}^2 \mathbf{B}_{\lambda,l-1}(c_k, t). \end{aligned}$$

Moreover, the first Fer streamer takes the form

$$\pi(\mathbf{B}_{\lambda,1}(c_k, t)) = \frac{q(t) - \frac{\int_{c_k}^t q(\xi) d\xi}{t - c_k}}{t - c_k} \begin{bmatrix} \varphi(\rho(\mathbf{D}_{\lambda,0}(c_k, t))) (t - c_k)^2 \\ -2\phi(\rho(\mathbf{D}_{\lambda,0}(c_k, t))) (t - c_k)^3 \\ \frac{1}{2}\phi(\rho(\mathbf{D}_{\lambda,0}(c_k, t))) \rho^2(\mathbf{D}_{\lambda,0}(c_k, t)) (t - c_k) \end{bmatrix}, \quad (17)$$

$$\rho(\mathbf{D}_{\lambda,0}(c_k, t)) = 2(t - c_k) \sqrt{\frac{\int_{c_k}^t q(\xi) d\xi}{t - c_k}} - \lambda. \quad (18)$$

The second insight in [14] is then to partition the λ -axis into

$$\lambda \in [q_{\max} - h_{\max}^{-2}, q_{\min} - 1] \cup [q_{\max} + 1, q_{\max} + h_{\max}^{-2}], \quad (19)$$

$$\lambda \in [q_{\min} - 1, q_{\max} + 1], \quad (20)$$

$$\lambda \in [q_{\max} + h_{\max}^{-2}, +\infty), \quad (21)$$

and use Fer streamers to control the quantities in Theorem 1 via the bounds with respect to (8) in Theorem 3 below, which make possible to derive the truncation estimates with respect to (8) in Theorem 4 below, which forms the main result in [14].

Theorem 3 ([14]) *If Assumption 1 holds true, $l \in \mathbb{Z}^+$ and $t \in [c_k, c_{k+1}]$, then it follows that*

$$e^{\mathbf{D}_{\lambda,0}(c_k, c_{k+1})} \dots e^{\mathbf{D}_{\lambda,0}(c_0, c_1)} = \begin{cases} \begin{bmatrix} \mathcal{O}(1) & \mathcal{O}(h_{\max}) \\ \mathcal{O}(h_{\max}^{-1}) & \mathcal{O}(1) \end{bmatrix}, & \text{in (19) and (20),} \\ \begin{bmatrix} \mathcal{O}(1) & \frac{\mathcal{O}(1)}{\sqrt{\lambda - q_{\max}}} \\ \mathcal{O}(1) \sqrt{\lambda - q_{\max}} & \mathcal{O}(1) \end{bmatrix}, & \text{in (21),} \end{cases}$$

as well as

$$\pi(\mathbf{D}_{\lambda,l}(c_k, t)) = \begin{cases} \|q'\|_{\infty}^{2^{l-1}} h_{\max}^{3 \cdot 2^{l-1} - 1} [\mathcal{O}(h_{\max}) \mathcal{O}(h_{\max}^2) \mathcal{O}(1)]^{\top}, & \text{in (19) and (20),} \\ \|q'\|_{\infty}^{2^{l-1}} h_{\max}^{2^l} (\lambda - q_{\max})^{-\frac{2^{l-1}-1}{2}} \left[\frac{\mathcal{O}(1)}{\sqrt{\lambda - q_{\max}}} \frac{\mathcal{O}(1)}{\lambda - q_{\max}} \mathcal{O}(1) \right]^{\top}, & \text{in (21).} \end{cases}$$

Definition 3 Let $[r] \in \mathbb{Z}^+$, and define the

$$\text{truncated flow: } \tilde{\mathbf{F}}_{\lambda,[r]}(c_k, c_{k+1}) := e^{\mathbf{D}_{\lambda,0}(c_k, c_{k+1})} e^{\mathbf{D}_{\lambda,1}(c_k, c_{k+1})} \dots e^{\mathbf{D}_{\lambda,[r]}(c_k, c_{k+1})},$$

$$\text{truncated solution: } \tilde{\mathbf{Y}}_{\lambda,[r]}(c_{k+1}) := \tilde{\mathbf{F}}_{\lambda,[r]}(c_k, c_{k+1}) \dots \tilde{\mathbf{F}}_{\lambda,[r]}(c_1, c_2) \tilde{\mathbf{F}}_{\lambda,[r]}(c_0, c_1),$$

$$\text{truncation local error: } \mathbf{L}_{\lambda,[r]}^{\text{trun.}}(c_k, c_{k+1}) := \log \left(\mathbf{F}_{\lambda}(c_k, c_{k+1}) \tilde{\mathbf{F}}_{\lambda,[r]}^{-1}(c_k, c_{k+1}) \right),$$

$$\text{truncation global error: } \mathbf{G}_{\lambda,[r]}^{\text{trun.}}(c_{k+1}) := \log \left(\mathbf{Y}_{\lambda}(c_{k+1}) \tilde{\mathbf{Y}}_{\lambda,[r]}^{-1}(c_{k+1}) \right).$$

Theorem 4 ([14]) *If Assumption 1 holds true, and $[r] \in \mathbb{Z}^+$, then*

$$\pi(\mathbf{L}_{\lambda,[r]}^{\text{trun.}}(c_k, c_{k+1})) = \begin{cases} \|q'\|_{\infty}^{2^{[r]}} h_{\max}^{3 \cdot 2^{[r]} - 1} [\mathcal{O}(h_{\max}) \mathcal{O}(h_{\max}^2) \mathcal{O}(1)]^{\top}, & \text{in (19) and (20),} \\ \|q'\|_{\infty}^{2^{[r]}} h_{\max}^{2^{[r]+1}} (\lambda - q_{\max})^{-\frac{2^{[r]}-1}{2}} \left[\frac{\mathcal{O}(1)}{\sqrt{\lambda - q_{\max}}} \frac{\mathcal{O}(1)}{\lambda - q_{\max}} \mathcal{O}(1) \right]^{\top}, & \text{in (21),} \end{cases}$$

$$\pi(\mathbf{G}_{\lambda,[r]}^{\text{trun.}}(c_{k+1})) = \begin{cases} \|q'\|_{\infty}^{2^{[r]}} h_{\max}^{3 \cdot 2^{[r]} - 2} [\mathcal{O}(h_{\max}) \mathcal{O}(h_{\max}^2) \mathcal{O}(1)]^{\top}, & \text{in (19) and (20),} \\ \|q'\|_{\infty}^{2^{[r]}} h_{\max}^{2^{[r]+1}} (\lambda - q_{\max})^{-\frac{2^{[r]}-1}{2}} \left[\frac{\mathcal{O}(1)}{\sqrt{\lambda - q_{\max}}} \frac{\mathcal{O}(1)}{\lambda - q_{\max}} \mathcal{O}(1) \right]^{\top}, & \text{in (21).} \end{cases}$$

Theorem 4 ensures the truncated solution $\tilde{\mathbf{Y}}_{\lambda,[r]}(c_{k+1})$ with $[r] = 1, 2, 3, \dots$ yields an approximation of the exact solution $\mathbf{Y}_{\lambda}(c_{k+1})$ to global order 4, 10, 22, \dots uniformly in (8).

In particular, Theorem 4 reduces the problem of approximating the *infinite* product of exponentials in the exact flow $\mathbf{F}_{\lambda}(c_k, c_{k+1})$ in Theorem 1 to approximating the *finite* product of exponentials in the truncated flow $\tilde{\mathbf{F}}_{\lambda,[r]}(c_k, c_{k+1})$ in Definition 3.

In addition, the fact that $\mathbf{X} \mapsto \exp(\mathbf{X})$ has the simple form in Definition 1, further reduces the problem to approximating the *finite* number of exponents

$$\mathbf{D}_{\lambda,0}(c_k, c_{k+1}), \mathbf{D}_{\lambda,1}(c_k, c_{k+1}), \dots, \mathbf{D}_{\lambda,[r]}(c_k, c_{k+1}). \quad (22)$$

3 Discretisation error estimates

In the previous section, we have seen in Definition 3 and Theorem 4 that $\tilde{\mathbf{F}}_{\lambda, [r]}(c_k, c_{k+1})$ and $\tilde{\mathbf{Y}}_{\lambda, [r]}(c_{k+1})$ approximate $\mathbf{F}_\lambda(c_k, c_{k+1})$ and $\mathbf{Y}_\lambda(c_{k+1})$ to global order 4, 10, 22, ... for $[r] = 1, 2, 3, \dots$, respectively.

In the present section, we will develop specialized quadrature that approximates (22) to high-order uniformly for all eigenvalues. This, in turn, will give rise to the discretized flow $\tilde{\mathbf{F}}_{\lambda, r}(c_k, c_{k+1})$ and discretized solution $\tilde{\mathbf{Y}}_{\lambda, r}(c_{k+1})$ in Definition 10 below, which will themselves be computed exactly. Towards the end, we will prove in Theorem 12 below that these approximate the truncated flow $\tilde{\mathbf{F}}_{\lambda, [r]}(c_k, c_{k+1})$ and truncated solution $\tilde{\mathbf{Y}}_{\lambda, [r]}(c_{k+1})$ to global order 4, 7, 10 for $r = 1, \log(3)/\log(2), 2$, respectively.

To this end, we begin with the observation that each exponent (22) can be approximated by multivariate integrals over polytopes, each of which is non-trivial, apart from the first. Indeed, the first term amounts to the computation of

$$\mathbf{D}_{\lambda, 0}(c_k, c_{k+1}) = (c_{k+1} - c_k) \begin{bmatrix} 0 & 1 \\ \frac{\int_{c_k}^{c_{k+1}} q(t) dt}{c_{k+1} - c_k} - \lambda & 0 \end{bmatrix},$$

which can be carried out without concern, while the second term can be written as

$$\mathbf{D}_{\lambda, 1}(c_k, c_{k+1}) = \int_{c_k}^{c_{k+1}} \mathbf{B}_{\lambda, 1}(c_k, t) dt,$$

and the third term can be controlled by

$$\begin{aligned} \mathbf{D}_{\lambda, 2}(c_k, c_{k+1}) &= \int_{c_k}^{c_{k+1}} \varphi(\rho(\mathbf{D}_{\lambda, 1}(c_k, t))) \text{ad}_{\mathbf{D}_{\lambda, 1}(c_k, t)} \mathbf{B}_{\lambda, 1}(c_k, t) dt \\ &\quad + \int_{c_k}^{c_{k+1}} \phi(\rho(\mathbf{D}_{\lambda, 1}(c_k, t))) \text{ad}_{\mathbf{D}_{\lambda, 1}(c_k, t)}^2 \mathbf{B}_{\lambda, 1}(c_k, t) dt \\ &= -\frac{1}{2} \int_{c_k}^{c_{k+1}} \int_{c_k}^{t_1} [\mathbf{B}_{\lambda, 1}(c_k, t_2), \mathbf{B}_{\lambda, 1}(c_k, t_1)] dt_2 dt_1 \\ &\quad + \frac{1}{3} \int_{c_k}^{c_{k+1}} \int_{c_k}^{t_1} \int_{c_k}^{t_1} [\mathbf{B}_{\lambda, 1}(c_k, t_3), [\mathbf{B}_{\lambda, 1}(c_k, t_2), \mathbf{B}_{\lambda, 1}(c_k, t_1)]] dt_3 dt_2 dt_1 \\ &\quad + \begin{cases} \|q'\|_\infty^4 h_{\max}^{11} \begin{bmatrix} \mathcal{O}(h_{\max}) & \mathcal{O}(h_{\max}^2) \\ \mathcal{O}(1) & -\mathcal{O}(h_{\max}) \end{bmatrix}, & \text{in (19) and (20),} \\ \|q'\|_\infty^4 h_{\max}^8 (\lambda - q_{\max})^{-\frac{3}{2}} \begin{bmatrix} \frac{\mathcal{O}(1)}{\sqrt{\lambda - q_{\max}}} & \frac{\mathcal{O}(1)}{\lambda - q_{\max}} \\ \mathcal{O}(1) & -\frac{\mathcal{O}(1)}{\sqrt{\lambda - q_{\max}}} \end{bmatrix}, & \text{in (21),} \end{cases} \\ &= -\frac{1}{2} \int_{c_k}^{c_{k+1}} \int_{c_k}^{t_1} [\mathbf{B}_{\lambda, 1}(c_k, t_2), \mathbf{B}_{\lambda, 1}(c_k, t_1)] dt_2 dt_1 \\ &\quad + \begin{cases} \|q'\|_\infty^3 h_{\max}^8 \begin{bmatrix} \mathcal{O}(h_{\max}) & \mathcal{O}(h_{\max}^2) \\ \mathcal{O}(1) & -\mathcal{O}(h_{\max}) \end{bmatrix}, & \text{in (19) and (20),} \\ \|q'\|_\infty^3 h_{\max}^6 (\lambda - q_{\max})^{-1} \begin{bmatrix} \frac{\mathcal{O}(1)}{\sqrt{\lambda - q_{\max}}} & \frac{\mathcal{O}(1)}{\lambda - q_{\max}} \\ \mathcal{O}(1) & -\frac{\mathcal{O}(1)}{\sqrt{\lambda - q_{\max}}} \end{bmatrix}, & \text{in (21),} \end{cases} \end{aligned}$$

where the first equality is due to Definition 1 and Theorem 2, and the last two follow from Definition 2 and Theorem 3. Hence, for global order less than or equal to ten, the second and third terms boil down to quadrature of

$$\int_{c_k}^{c_{k+1}} \mathbf{B}_{\lambda,1}(c_k, t) dt, \quad (23)$$

$$\int_{c_k}^{c_{k+1}} \int_{c_k}^{t_1} [\mathbf{B}_{\lambda,1}(c_k, t_2), \mathbf{B}_{\lambda,1}(c_k, t_1)] dt_2 dt_1, \quad (24)$$

$$\int_{c_k}^{c_{k+1}} \int_{c_k}^{t_1} \int_{c_k}^{t_1} [\mathbf{B}_{\lambda,1}(c_k, t_3), [\mathbf{B}_{\lambda,1}(c_k, t_2), \mathbf{B}_{\lambda,1}(c_k, t_1)]] dt_3 dt_2 dt_1. \quad (25)$$

In detail, Fer streamers require:

- for global order 4, the quadrature of (23) to local order 5,
- for global order 7, the quadrature of (23)–(24) to local order 8,
- for global order 10, the quadrature of (23)–(25) to local order 11.

We develop quadrature of (23)–(25) via representations of their integrands that exploit their:

- magnitude to reduce quadrature points, function evaluations and linear algebra, and,
- behaviour to decrease quadrature error without using derivatives of the potential.

These quadrature schemes and representations vary across the three subsets (19)–(21) where $\mathbf{B}_{\lambda,1}(c_k, t)$ exhibits different behaviour and magnitude.

3.1 Exposing the behaviour of the integrands

We start by making the behaviour of $\mathbf{B}_{\lambda,1}(c_k, t)$ explicit in Theorems 5–7 below, which we exploit thereafter.

Definition 4 Let

$$\omega_{\lambda,1}(c_k, t) := 2(t - c_k) \sqrt{\lambda - q(c_k)} \quad \varepsilon_{\lambda,1}(c_k, t) := \omega_{\lambda,1}(c_k, t) (r_{\lambda,1}(c_k, t) - 1), \quad (26)$$

$$r_{\lambda,1}(c_k, t) := \sqrt{\frac{\lambda - \frac{\int_{c_k}^t q(\xi) d\xi}{t - c_k}}{\lambda - q(c_k)}} \quad s_{\lambda,1}(c_k, t) := \omega_{\lambda,1}(c_k, t) \varepsilon_{\lambda,1}(c_k, t). \quad (27)$$

To provide intuition before plunging into technicalities, the reader should be aware that the guiding principle that leads to the representations in Theorems 5–7 below is to rewrite the representation of $\mathbf{B}_{\lambda,1}(c_k, t)$ in (17)–(18) in terms of trigonometric functions with the argument $\omega_{\lambda,1}(c_k, t)$. To this end, recall (17)–(18) and invoke Definitions 2 and 4 to rewrite

$$\rho(\mathbf{D}_{\lambda,0}(c_k, t)) = i \cdot \omega_{\lambda,1}(c_k, t) r_{\lambda,1}(c_k, t) = i \cdot (\varepsilon_{\lambda,1}(c_k, t) + \omega_{\lambda,1}(c_k, t)), \quad (28)$$

$$\begin{aligned} \cosh(\rho(\mathbf{D}_{\lambda,0}(c_k, t))) &= \cos(\varepsilon_{\lambda,1}(c_k, t) + \omega_{\lambda,1}(c_k, t)) \\ &= \cos(\varepsilon_{\lambda,1}(c_k, t)) \cdot \cos(\omega_{\lambda,1}(c_k, t)) \\ &\quad - \sin(\varepsilon_{\lambda,1}(c_k, t)) \cdot \sin(\omega_{\lambda,1}(c_k, t)), \end{aligned} \quad (29)$$

$$\begin{aligned}
\frac{\sinh(\rho(\mathbf{D}_{\lambda,0}(c_k, t)))}{\rho(\mathbf{D}_{\lambda,0}(c_k, t))} &= \frac{\sin(\varepsilon_{\lambda,1}(c_k, t) + \omega_{\lambda,1}(c_k, t))}{\omega_{\lambda,1}(c_k, t) r_{\lambda,1}(c_k, t)} \\
&= \frac{r_{\lambda,1}(c_k, t) - 1}{r_{\lambda,1}(c_k, t)} \frac{\sin(\varepsilon_{\lambda,1}(c_k, t))}{\varepsilon_{\lambda,1}(c_k, t)} \cdot \cos(\omega_{\lambda,1}(c_k, t)) \\
&\quad + \frac{1}{r_{\lambda,1}(c_k, t)} \cos(\varepsilon_{\lambda,1}(c_k, t)) \cdot \frac{\sin(\omega_{\lambda,1}(c_k, t))}{\omega_{\lambda,1}(c_k, t)}. \quad (30)
\end{aligned}$$

Since $r_{\lambda,1}(c_k, t)$ is close to 1 and $\varepsilon_{\lambda,1}(c_k, t)$ is close to 0 uniformly for every eigenvalue, the behaviour of $\mathbf{B}_{\lambda,1}(c_k, t)$ will be encapsulated in terms of trigonometric functions with argument $\omega_{\lambda,1}(c_k, t)$, provided some care is taken to make every singularity removable. As will become clear, this serves to make the behaviour and magnitude of $\mathbf{B}_{\lambda,1}(c_k, t)$ explicit, which, in turn, serves to reduce the function evaluations and linear algebra in the quadrature and to decrease the quadrature error without using derivatives of the potential.

To make this guiding principle precise the next definition introduces the non-oscillatory parts $\mathbf{f}_{\lambda,1}(c_k, t)$, $\mathbf{v}_{\lambda,1}(c_k, t)$ and $\mathbf{g}_{\lambda,1}(c_k, t)$, which appear below in Theorems 5–7. Although important, it is technical in nature and the reader is encouraged to glance over it and return to it as required.

Definition 5

$$\begin{aligned}
\mathbf{f}_{\lambda,1}(c_k, t) &:= \frac{q(t) - \frac{\int_{c_k}^t q(\xi) d\xi}{t - c_k}}{t - c_k} \frac{1}{r_{\lambda,1}^2(c_k, t)} \\
&\quad \times \begin{bmatrix} 1 \\ (r_{\lambda,1}(c_k, t) - 1) \left(r_{\lambda,1}(c_k, t) \phi(i \cdot \varepsilon_{\lambda,1}(c_k, t)) - \frac{1 - \cos(\varepsilon_{\lambda,1}(c_k, t))}{\varepsilon_{\lambda,1}^2(c_k, t)} \right) \\ -r_{\lambda,1}(c_k, t) \left(\cos(\varepsilon_{\lambda,1}(c_k, t)) - \frac{r_{\lambda,1}(c_k, t) - 1}{r_{\lambda,1}(c_k, t)} \frac{\sin(\varepsilon_{\lambda,1}(c_k, t))}{\varepsilon_{\lambda,1}(c_k, t)} \right) \\ -2(r_{\lambda,1}(c_k, t) - 1) \left(\frac{(r_{\lambda,1}(c_k, t) - 1)^2}{r_{\lambda,1}(c_k, t)} \phi(i \cdot \varepsilon_{\lambda,1}(c_k, t)) + \frac{\sin(\varepsilon_{\lambda,1}(c_k, t))}{\varepsilon_{\lambda,1}(c_k, t)} \right) \\ -\frac{2}{r_{\lambda,1}(c_k, t)} \left(\cos(\varepsilon_{\lambda,1}(c_k, t)) + r_{\lambda,1}(c_k, t) s_{\lambda,1}(c_k, t) \frac{\sin(\varepsilon_{\lambda,1}(c_k, t))}{\varepsilon_{\lambda,1}(c_k, t)} \right) \\ \frac{1}{2} r_{\lambda,1}^2(c_k, t) \left(\cos(\varepsilon_{\lambda,1}(c_k, t)) - \frac{r_{\lambda,1}(c_k, t) - 1}{r_{\lambda,1}(c_k, t)} \frac{\sin(\varepsilon_{\lambda,1}(c_k, t))}{\varepsilon_{\lambda,1}(c_k, t)} \right) \\ -\frac{1}{2} r_{\lambda,1}(c_k, t) \left(\cos(\varepsilon_{\lambda,1}(c_k, t)) + r_{\lambda,1}(c_k, t) s_{\lambda,1}(c_k, t) \frac{\sin(\varepsilon_{\lambda,1}(c_k, t))}{\varepsilon_{\lambda,1}(c_k, t)} \right) \end{bmatrix}, \\
\mathbf{f}_{\lambda,1}(c_k, c_k) &:= \frac{q'(c_k^+)}{2} \begin{bmatrix} 1 & 0 & -1 & 0 & -2 & \frac{1}{2} & -\frac{1}{2} \end{bmatrix}^\top, \\
\mathbf{v}_{\lambda,1}(c_k, t) &:= \frac{q(t) - \frac{\int_{c_k}^t q(\xi) d\xi}{t - c_k}}{t - c_k} \begin{bmatrix} \phi(\sqrt{\rho^2(\mathbf{D}_{\lambda,0}(c_k, t))}) \\ -2\phi(\sqrt{\rho^2(\mathbf{D}_{\lambda,0}(c_k, t))}) \\ \frac{1}{2}\phi(\sqrt{\rho^2(\mathbf{D}_{\lambda,0}(c_k, t))}) \rho^2(\mathbf{D}_{\lambda,0}(c_k, t)) \end{bmatrix}, \\
\mathbf{v}_{\lambda,1}(c_k, c_k) &:= \frac{q'(c_k^+)}{2} \begin{bmatrix} -\frac{1}{2} & -\frac{2}{3} & 0 \end{bmatrix}^\top,
\end{aligned}$$

$$\mathbf{g}_{\lambda,1}(c_k, t) := \frac{1}{2} \frac{q(t) - \frac{\int_{c_k}^t q(\xi) d\xi}{t - c_k}}{t - c_k} \frac{1}{r_{\lambda,1}^2(c_k, t)} \times \left[\begin{array}{l} 1 \\ (r_{\lambda,1}(c_k, t) - 1) \frac{1 - \cos(\varepsilon_{\lambda,1}(c_k, t))}{\varepsilon_{\lambda,1}(c_k, t)} - r_{\lambda,1}(c_k, t) \sin(\varepsilon_{\lambda,1}(c_k, t)) \\ - r_{\lambda,1}(c_k, t) \left(\cos(\varepsilon_{\lambda,1}(c_k, t)) - \frac{r_{\lambda,1}(c_k, t) - 1}{r_{\lambda,1}(c_k, t)} \frac{\sin(\varepsilon_{\lambda,1}(c_k, t))}{\varepsilon_{\lambda,1}(c_k, t)} \right) \\ \cos(\varepsilon_{\lambda,1}(c_k, t)) - \frac{r_{\lambda,1}(c_k, t) - 1}{r_{\lambda,1}(c_k, t)} \frac{\sin(\varepsilon_{\lambda,1}(c_k, t))}{\varepsilon_{\lambda,1}(c_k, t)} \\ - \frac{1}{r_{\lambda,1}(c_k, t)} \left(\cos(\varepsilon_{\lambda,1}(c_k, t)) + r_{\lambda,1}(c_k, t) s_{\lambda,1}(c_k, t) \frac{\sin(\varepsilon_{\lambda,1}(c_k, t))}{\varepsilon_{\lambda,1}(c_k, t)} \right) \\ r_{\lambda,1}^2(c_k, t) \left(\cos(\varepsilon_{\lambda,1}(c_k, t)) - \frac{r_{\lambda,1}(c_k, t) - 1}{r_{\lambda,1}(c_k, t)} \frac{\sin(\varepsilon_{\lambda,1}(c_k, t))}{\varepsilon_{\lambda,1}(c_k, t)} \right) \\ - r_{\lambda,1}(c_k, t) \left(\cos(\varepsilon_{\lambda,1}(c_k, t)) + r_{\lambda,1}(c_k, t) s_{\lambda,1}(c_k, t) \frac{\sin(\varepsilon_{\lambda,1}(c_k, t))}{\varepsilon_{\lambda,1}(c_k, t)} \right) \end{array} \right],$$

$$\mathbf{g}_{\lambda,1}(c_k, c_k) := \frac{q'(c_k^+)}{4} [1 \ 0 \ -1 \ 1 \ -1 \ 1 \ -1]^\top.$$

With Definition 5 in hand, we encapsulate the behaviour of $\mathbf{B}_{\lambda,1}(c_k, t)$ in the following three theorems concisely via the Hadamard/entrywise product denoted by \odot .

Theorem 5 *If λ lies in (19) then*

$$\begin{aligned} \pi(\mathbf{B}_{\lambda,1}(c_k, t)) &= \frac{1 - \cos(\omega_{\lambda,1}(c_k, t))}{(\omega_{\lambda,1}(c_k, t))^2} (t - c_k) \begin{bmatrix} t - c_k \\ 0 \\ 0 \end{bmatrix} \odot \begin{bmatrix} [\mathbf{f}_{\lambda,1}(c_k, t)]_{1,1} \\ 0 \\ 0 \end{bmatrix} \\ &+ \cos(\omega_{\lambda,1}(c_k, t)) (t - c_k) \begin{bmatrix} t - c_k \\ (t - c_k)^2 \\ 1 \end{bmatrix} \odot \begin{bmatrix} [\mathbf{f}_{\lambda,1}(c_k, t)]_{2,1} \\ [\mathbf{f}_{\lambda,1}(c_k, t)]_{4,1} \\ [\mathbf{f}_{\lambda,1}(c_k, t)]_{6,1} \end{bmatrix} \\ &+ \frac{\sin(\omega_{\lambda,1}(c_k, t))}{\omega_{\lambda,1}(c_k, t)} (t - c_k) \begin{bmatrix} t - c_k \\ 0 \\ 1 \end{bmatrix} \odot \begin{bmatrix} [\mathbf{f}_{\lambda,1}(c_k, t)]_{3,1} \\ 0 \\ [\mathbf{f}_{\lambda,1}(c_k, t)]_{7,1} \end{bmatrix} \\ &+ \phi(i \cdot \omega_{\lambda,1}(c_k, t)) (t - c_k) \begin{bmatrix} 0 \\ (t - c_k)^2 \\ 0 \end{bmatrix} \odot \begin{bmatrix} 0 \\ [\mathbf{f}_{\lambda,1}(c_k, t)]_{5,1} \\ 0 \end{bmatrix}. \end{aligned}$$

Furthermore, the derivatives $\mathbf{f}_{\lambda,1}^{(j)}(c_k, t)$ can be bounded independently of λ .

Proof Without loss of generality assume $\lambda \in [q_{\max} + 1, q_{\max} + h_{\max}^{-2}]$. The result follows from (17)–(18) and Definitions 2, 5 together with (28)–(30). The terms are arranged to make

$$(t - c_k) [t - c_k \ (t - c_k)^2 \ 1]^\top$$

explicit and to make every singularity removable. Finally, $\mathbf{f}_{\lambda,1}^{(j)}(c_k, t)$ can be bounded independently of λ since the derivatives of (27) can be bounded independently of λ . \square

Theorem 6 If λ lies in (20) then

$$\pi(\mathbf{B}_{\lambda,1}(c_k, t)) = (t - c_k) \begin{bmatrix} t - c_k & (t - c_k)^2 & 1 \end{bmatrix}^\top \odot \mathbf{t}_{\lambda,1}(c_k, t).$$

Furthermore, the derivatives $\mathbf{t}_{\lambda,1}^{(j)}(c_k, \cdot)$ can be bounded independently of λ .

Proof Follows immediately from (17)–(18). \square

Theorem 7 If λ lies in (21) then

$$\begin{aligned} \pi(\mathbf{B}_{\lambda,1}(c_k, t)) &= \frac{1 - \cos(\omega_{\lambda,1}(c_k, t))}{\omega_{\lambda,1}(c_k, t)} (t - c_k) \begin{bmatrix} \frac{1}{\sqrt{\lambda - q(c_k)}} \\ 0 \\ 0 \end{bmatrix} \odot \begin{bmatrix} [\mathbf{g}_{\lambda,1}(c_k, t)]_{1,1} \\ 0 \\ 0 \end{bmatrix} \\ &+ \cos(\omega_{\lambda,1}(c_k, t)) (t - c_k) \begin{bmatrix} \frac{1}{\sqrt{\lambda - q(c_k)}} \\ \frac{1}{\lambda - q(c_k)} \\ 1 \end{bmatrix} \odot \begin{bmatrix} [\mathbf{g}_{\lambda,1}(c_k, t)]_{2,1} \\ [\mathbf{g}_{\lambda,1}(c_k, t)]_{4,1} \\ [\mathbf{g}_{\lambda,1}(c_k, t)]_{6,1} \end{bmatrix} \\ &+ \frac{\sin(\omega_{\lambda,1}(c_k, t))}{\omega_{\lambda,1}(c_k, t)} (t - c_k) \begin{bmatrix} 0 \\ \frac{1}{\lambda - q(c_k)} \\ 1 \end{bmatrix} \odot \begin{bmatrix} 0 \\ [\mathbf{g}_{\lambda,1}(c_k, t)]_{5,1} \\ [\mathbf{g}_{\lambda,1}(c_k, t)]_{7,1} \end{bmatrix} \\ &+ \sin(\omega_{\lambda,1}(c_k, t)) (t - c_k) \begin{bmatrix} \frac{1}{\sqrt{\lambda - q(c_k)}} \\ 0 \\ 0 \end{bmatrix} \odot \begin{bmatrix} [\mathbf{g}_{\lambda,1}(c_k, t)]_{3,1} \\ 0 \\ 0 \end{bmatrix}. \end{aligned}$$

Furthermore, the derivatives $\mathbf{g}_{\lambda,1}^{(j)}(c_k, t)$ can be bounded independently of λ .

Proof The representation follows from (17)–(18) and Definitions 2 and 5 together with (28)–(30). The terms are arranged in order to make

$$(t - c_k) \begin{bmatrix} \frac{1}{\sqrt{\lambda - q(c_k)}} & \frac{1}{\lambda - q(c_k)} & 1 \end{bmatrix}^\top$$

explicit and render every singularity removable. Finally, $\mathbf{g}_{\lambda,1}^{(j)}(c_k, t)$ can be bounded independently of λ because the derivatives of (27) can be bounded independently of λ . \square

3.2 Exposing the magnitude of the integrands

Definition 6 below introduces $\mathbf{B}_{\lambda,1}^{\text{fine}}(c_k, c_k + h_k t)$ as a means to decompose the fine and coarse parts of $\mathbf{B}_{\lambda,1}(c_k, c_k + h_k t)$. This fine and coarse decomposition is made precise with Corollaries 1–3 below which show that $\mathbf{B}_{\lambda,1}^{\text{fine}}(c_k, c_k + h_k t)$ is $\mathcal{O}(1)$ uniformly over the entire eigenvalue range. It is then in the next subsection, in Theorems 8–10, that this fine and coarse decomposition is shown to bear fruit in the form of reduced requirements for quadrature.

Definition 6 Let $\mathbf{B}_{\lambda,1}^{\text{fine}}(c_k, c_k + h_k t)$ be the unique element in $\mathfrak{sl}(2, \mathbb{R})$ such that

$$\begin{aligned} &\pi(\mathbf{B}_{\lambda,1}(c_k, c_k + h_k t)) \\ &=: \pi(\mathbf{B}_{\lambda,1}^{\text{fine}}(c_k, c_k + h_k t)) \odot \begin{cases} h_k \begin{bmatrix} h_k & h_k^2 & 1 \end{bmatrix}^\top, & \text{in (19) and (20),} \\ h_k \begin{bmatrix} \frac{1}{\sqrt{\lambda - q(c_k)}} & \frac{1}{\lambda - q(c_k)} & 1 \end{bmatrix}^\top, & \text{in (21).} \end{cases} \end{aligned}$$

The next three Corollaries follow immediately from Theorems 5–7 and Definition 6.

Corollary 1 *If λ lies in (19) then $\omega_{\lambda,1}(c_k, c_{k+1}) \in [0, 2i]$ or $\omega_{\lambda,1}(c_k, c_{k+1}) \in [0, 2\sqrt{2}]$ and*

$$\begin{aligned} \pi(\mathbf{B}_{\lambda,1}^{fine}(c_k, c_k + h_k t)) &= \frac{1 - \cos(\omega_{\lambda,1}(c_k, c_{k+1})t)}{(\omega_{\lambda,1}(c_k, c_{k+1})t)^2} \begin{bmatrix} t \\ 0 \\ 0 \end{bmatrix} \odot t \begin{bmatrix} [\mathbf{f}_{\lambda,1}(c_k, c_k + h_k t)]_{1,1} \\ 0 \\ 0 \end{bmatrix} \\ &+ \cos(\omega_{\lambda,1}(c_k, c_{k+1})t) \begin{bmatrix} t \\ t^2 \\ 1 \end{bmatrix} \odot t \begin{bmatrix} [\mathbf{f}_{\lambda,1}(c_k, c_k + h_k t)]_{2,1} \\ [\mathbf{f}_{\lambda,1}(c_k, c_k + h_k t)]_{4,1} \\ [\mathbf{f}_{\lambda,1}(c_k, c_k + h_k t)]_{6,1} \end{bmatrix} \\ &+ \frac{\sin(\omega_{\lambda,1}(c_k, c_{k+1})t)}{\omega_{\lambda,1}(c_k, c_{k+1})t} \begin{bmatrix} t \\ 0 \\ 1 \end{bmatrix} \odot t \begin{bmatrix} [\mathbf{f}_{\lambda,1}(c_k, c_k + h_k t)]_{3,1} \\ 0 \\ [\mathbf{f}_{\lambda,1}(c_k, c_k + h_k t)]_{7,1} \end{bmatrix} \\ &+ \phi(i \cdot \omega_{\lambda,1}(c_k, c_{k+1})t) \begin{bmatrix} 0 \\ t^2 \\ 0 \end{bmatrix} \odot t \begin{bmatrix} 0 \\ [\mathbf{f}_{\lambda,1}(c_k, c_k + h_k t)]_{5,1} \\ 0 \end{bmatrix}. \quad (31) \end{aligned}$$

Corollary 2 *If λ lies in (20) then*

$$\pi(\mathbf{B}_{\lambda,1}^{fine}(c_k, c_k + h_k t)) = [t \ t^2 \ 1]^\top \odot t \cdot \mathbf{t}_{\lambda,1}(c_k, c_k + h_k t). \quad (32)$$

Corollary 3 *If λ lies in (21) then $\omega_{\lambda,1}(c_k, c_{k+1}) \in [1, +\infty)$ and*

$$\begin{aligned} \pi(\mathbf{B}_{\lambda,1}^{fine}(c_k, c_k + h_k t)) &= \frac{1 - \cos(\omega_{\lambda,1}(c_k, c_{k+1})t)}{\omega_{\lambda,1}(c_k, c_{k+1})t} t \begin{bmatrix} [\mathbf{g}_{\lambda,1}(c_k, c_k + h_k t)]_{1,1} \\ 0 \\ 0 \end{bmatrix} \\ &+ \cos(\omega_{\lambda,1}(c_k, c_{k+1})t) t \begin{bmatrix} [\mathbf{g}_{\lambda,1}(c_k, c_k + h_k t)]_{2,1} \\ [\mathbf{g}_{\lambda,1}(c_k, c_k + h_k t)]_{4,1} \\ [\mathbf{g}_{\lambda,1}(c_k, c_k + h_k t)]_{6,1} \end{bmatrix} \\ &+ \frac{\sin(\omega_{\lambda,1}(c_k, c_{k+1})t)}{\omega_{\lambda,1}(c_k, c_{k+1})t} t \begin{bmatrix} 0 \\ [\mathbf{g}_{\lambda,1}(c_k, c_k + h_k t)]_{5,1} \\ [\mathbf{g}_{\lambda,1}(c_k, c_k + h_k t)]_{7,1} \end{bmatrix} \\ &+ \sin(\omega_{\lambda,1}(c_k, c_{k+1})t) t \begin{bmatrix} [\mathbf{g}_{\lambda,1}(c_k, c_k + h_k t)]_{3,1} \\ 0 \\ 0 \end{bmatrix}. \quad (33) \end{aligned}$$

3.3 Reduced requirements for quadrature

Theorems 8–10 below highlight the synergy between the Lie bracket and the representations in Theorems 5–7, Definition 6 and Corollaries 1–3: they act together to decrease the magnitude of each multivariate integral, making it smaller than expected! In particular, as discussed in the subsequent text, they illustrate that it is possible to use these representations to develop a quadrature which exploits the magnitude of each integrand in order to reduce the number of function evaluations and amount of linear algebra.

Theorem 8 *If λ belongs to (19), (20) or (21) then*

$$\pi \left(\int_{c_k}^{c_{k+1}} \mathbf{B}_{\lambda,1}(c_k, t) dt \right) = \pi \left(\int_0^1 \mathbf{B}_{\lambda,1}^{\text{fine}}(c_k, c_k + h_k t) dt \right) \odot \begin{cases} h_k^2 [h_k \ h_k^2 \ 1]^\top, & \text{in (19) and (20),} \\ h_k^2 \left[\frac{1}{\sqrt{\lambda - q(c_k)}} \ \frac{1}{\lambda - q(c_k)} \ 1 \right]^\top, & \text{in (21).} \end{cases}$$

Proof Follows by straightforward computation from Definition 6. \square

Theorem 9 *If λ belongs to (19), (20) or (21) then*

$$\begin{aligned} & \pi \left(\int_{c_k}^{c_{k+1}} \int_{c_k}^{t_1} [\mathbf{B}_{\lambda,1}(c_k, t_2), \mathbf{B}_{\lambda,1}(c_k, t_1)] dt_2 dt_1 \right) \\ &= \pi \left(\int_0^1 \int_0^{t_1} [\mathbf{B}_{\lambda,1}^{\text{fine}}(c_k, c_k + h_k t_2), \mathbf{B}_{\lambda,1}^{\text{fine}}(c_k, c_k + h_k t_1)] dt_2 dt_1 \right) \\ & \odot \begin{cases} h_k^5 [h_k \ h_k^2 \ 1]^\top, & \text{in (19) and (20),} \\ h_k^4 (\lambda - q(c_k))^{-\frac{1}{2}} \left[\frac{1}{\sqrt{\lambda - q(c_k)}} \ \frac{1}{\lambda - q(c_k)} \ 1 \right]^\top, & \text{in (21).} \end{cases} \end{aligned}$$

Proof Follows by straightforward computation from Definition 6. \square

Theorem 10 *If λ belongs to (19), (20) or (21) then*

$$\begin{aligned} & \pi \left(\int_{c_k}^{c_{k+1}} \int_{c_k}^{t_1} \int_{c_k}^{t_2} [\mathbf{B}_{\lambda,1}(c_k, t_3), [\mathbf{B}_{\lambda,1}(c_k, t_2), \mathbf{B}_{\lambda,1}(c_k, t_1)]] dt_3 dt_2 dt_1 \right) \\ &= \pi \left(\int_0^1 \int_0^{t_1} \int_0^{t_2} [\mathbf{B}_{\lambda,1}^{\text{fine}}(c_k, c_k + h_k t_3), [\mathbf{B}_{\lambda,1}^{\text{fine}}(c_k, c_k + h_k t_2), \mathbf{B}_{\lambda,1}^{\text{fine}}(c_k, c_k + h_k t_1)]] dt_3 dt_2 dt_1 \right) \\ & \odot \begin{cases} h_k^8 [h_k \ h_k^2 \ 1]^\top, & \text{in (19) and (20),} \\ h_k^6 (\lambda - q(c_k))^{-1} \left[\frac{1}{\sqrt{\lambda - q(c_k)}} \ \frac{1}{\lambda - q(c_k)} \ 1 \right]^\top, & \text{in (21).} \end{cases} \end{aligned}$$

Proof Follows by straightforward computation from Definition 6. \square

As a result of Theorems 8–10, quadrature of (23)–(25) should be replaced with that of

$$\int_0^1 \mathbf{B}_{\lambda,1}^{\text{fine}}(c_k, c_k + h_k t) dt, \quad (34)$$

$$\int_0^1 \int_0^{t_1} [\mathbf{B}_{\lambda,1}^{\text{fine}}(c_k, c_k + h_k t_2), \mathbf{B}_{\lambda,1}^{\text{fine}}(c_k, c_k + h_k t_1)] dt_2 dt_1, \quad (35)$$

$$\int_0^1 \int_0^{t_1} \int_0^{t_2} [\mathbf{B}_{\lambda,1}^{\text{fine}}(c_k, c_k + h_k t_3), [\mathbf{B}_{\lambda,1}^{\text{fine}}(c_k, c_k + h_k t_2), \mathbf{B}_{\lambda,1}^{\text{fine}}(c_k, c_k + h_k t_1)]] dt_3 dt_2 dt_1, \quad (36)$$

since this results in fewer function evaluations and amount of linear algebra, given that:

- for global order 4, as a consequence of Theorem 8, quadrature of (23) to local order 5 is equivalent to quadrature of (34) to local order 3,

- for global order 7, according to Theorem 8, quadrature of (23) to local order 8 is equivalent to quadrature of (34) to local order 6, and, following Theorem 9, quadrature of (24) to local order 8 is equivalent to quadrature of (35) to local order 3,
- for global order 10, Theorem 8 guarantees quadrature of (23) to local order 11 is equivalent to quadrature of (34) to local order 9, Theorem 9 ensures quadrature of (24) to local order 11 is equivalent to quadrature of (35) to local order 6, and, Theorem 10 ensures quadrature of (25) to local order 11 is equivalent to quadrature of (36) to local order 3.

3.4 Interpolation

We develop quadrature for (34)–(36) by polynomial interpolation of the non-oscillatory parts of $\mathbf{B}_{\lambda,1}^{\text{fine}}(c_k, c_k + h_k t)$, given by $\mathbf{f}_{\lambda,1}(c_k, c_k + h_k t)$, $\mathbf{t}_{\lambda,1}(c_k, c_k + h_k t)$ and $\mathbf{g}_{\lambda,1}(c_k, c_k + h_k t)$, as exposed in Corollaries 1–3. This gives rise to the approximations $\tilde{\mathbf{B}}_{\lambda,1,\mathcal{T}_j}^{\text{fine}}(c_k, c_k + h_k t)$ in:

Definition 7 Let $\mathcal{T}_j \subseteq [0, 1]$ be formed by j distinct points and define

$$\tilde{\mathbf{B}}_{\lambda,1,\mathcal{T}_j}^{\text{fine}}(c_k, c_k + h_k t) \in \mathfrak{sl}(2, \mathbb{R})$$

in each of (19), (20) and (21) by the right hand side of, respectively:

- (31) with $t \mapsto [\mathbf{f}_{\lambda,1}(c_k, c_k + h_k t)]_{i,1}$ replaced by polynomial interpolation at \mathcal{T}_j ,
- (32) with $t \mapsto [\mathbf{t}_{\lambda,1}(c_k, c_k + h_k t)]_{i,1}$ replaced by polynomial interpolation at \mathcal{T}_j ,
- (33) with $t \mapsto [\mathbf{g}_{\lambda,1}(c_k, c_k + h_k t)]_{i,1}$ replaced by polynomial interpolation at \mathcal{T}_j .

The question then becomes which (j, \mathcal{T}_j) should be used for quadrature of (34)–(36) to achieve prescribed local order. A question we address in the next three subsubsections:

3.4.1 Smallest number of interpolation points to be consistent with local order

Let $0 \leq \tau_1 < \tau_2 < \dots < \tau_j \leq 1$ be j interpolation points and let $t \mapsto \mathbf{p}_{\mathbf{f}_{\lambda,1}(c_k, c_k + h_k \cdot), j-1}(t)$ be the unique $j-1$ degree interpolation polynomial such that, for every $l \in \{1, \dots, j\}$, $\mathbf{p}_{\mathbf{f}_{\lambda,1}(c_k, c_k + h_k \cdot), j-1}(\tau_l) = \mathbf{f}_{\lambda,1}(c_k, c_k + h_k \tau_l)$. Then [12] for each $t \in [0, 1]$, there exists $\xi \in [0, 1]$:

$$\mathbf{f}_{\lambda,1}(c_k, c_k + h_k t) - \mathbf{p}_{\mathbf{f}_{\lambda,1}(c_k, c_k + h_k \cdot), j-1}(t) = \frac{h_k^j \mathbf{f}_{\lambda,1}^{(j)}(c_k, c_k + h_k \xi)}{j!} \prod_{l=1}^j (t - \tau_l)$$

which yields the pointwise estimate (similarly also for $\mathbf{t}_{\lambda,1}$ and $\mathbf{g}_{\lambda,1}$): for each $t \in [0, 1]$,

$$\begin{aligned} & \left| t \left[\mathbf{f}_{\lambda,1}(c_k, c_k + h_k t) - \mathbf{p}_{\mathbf{f}_{\lambda,1}(c_k, c_k + h_k \cdot), j-1}(t) \right]_{i,1} \right| \\ & \leq \frac{h_k^j}{j!} \max_{\xi \in [c_k, c_k + 1]} \left\{ \left| \left[\mathbf{f}_{\lambda,1}^{(j)}(c_k, \xi) \right]_{i,1} \right| \right\} \max_{\xi \in [0, 1]} \left\{ \left| \xi \prod_{l=1}^j (\xi - \tau_l) \right| \right\}. \end{aligned} \quad (37)$$

Together with the discussion following (34)–(36), the error bound (37) dictates that:

- for global order 4, we need $j = 3$ for (34),
- for global order 7, we need $j = 6$ for (34) and $j = 3$ for (35),
- for global order 10, we need $j = 9$ for (34), $j = 6$ for (35) and $j = 3$ for (36).

In particular, fewer interpolation points are needed for higher dimensional integrals than for lower dimensional integrals, which represents a significant saving in linear algebra.

3.4.2 Interpolation points that reduce quadrature error without derivatives of the potential

As in the analysis in [10, 5, 6] for highly oscillatory Fourier-type integrands, the mildly to highly oscillatory behaviour in (21) made explicit in Corollary 3 also suggests polynomial interpolation at the endpoints since this yields smaller quadrature error. This would lead to the evaluation of $\mathbf{g}_{\lambda,1}(c_k, c_k)$, which depends on $q'(c_k^+)$: something that would be best to avoid since the derivative of the potential might not be available. Fortunately, since in Corollary 3 there is a ‘ t ’ term in front of every $[\mathbf{g}_{\lambda,1}(c_k, c_k + h_k t)]_{i,1}$ term, this is automatically achieved at the left boundary point. Hence, there is no need to interpolate at the left boundary point. As for $t \in (0, 1]$, $\mathbf{g}_{\lambda,1}(c_k, c_k + h_k t)$ does not depend on the derivative of the potential and should be interpolated at the right boundary point. Given this, choose $\tau_1 \neq 0$ and $\tau_j := 1$, in which case, $t \mapsto t \mathbf{p}_{\lambda,1}(c_k, c_k + h_k \cdot)_{j-1}(t)$ is the unique j degree interpolation polynomial which interpolates $t \mapsto t \mathbf{f}_{\lambda,1}(c_k, c_k + h_k t)$ at the $j+1$ points $\{0, \tau_1, \dots, \tau_{j-1}, 1\}$.

3.4.3 Data

As discussed in Subsubsections 3.4.1–3.4.2, it is beneficial to choose $\mathcal{T}_j := \mathcal{S}_{j-1} \cup \{1\}$ where $\mathcal{S}_{j-1} \subseteq (0, 1)$ has $j-1$ distinct points where $j = 3, 6, 9$ vary with (34)–(36) and local order. To reduce potential evaluations, take $\mathcal{S}_2 \subseteq \mathcal{S}_5 \subseteq \mathcal{S}_8$, in which case, one lets

$$\mathcal{S} := \mathcal{S}_{l-2} \text{ for global order } l,$$

in which case, the polynomial interpolation in this section requires the following data

$$\bigcup_{k=0}^{m-1} \left(\{q(c_k)\} \cup \left\{ q(c_k + h_k t), \int_{c_k}^{c_k + h_k t} q(\xi) d\xi : t \in \mathcal{S} \right\} \cup \left\{ \int_{c_k}^{c_{k+1}} q(\xi) d\xi \right\} \right) \cup \{q(b)\}$$

If the antiderivative of the potential is not available, then it is possible to approximate, up to local order, the antiderivative data

$$\left\{ \int_{c_k}^{c_k + h_k t} q(\xi) d\xi : t \in \mathcal{S} \right\} \cup \left\{ \int_{c_k}^{c_{k+1}} q(\xi) d\xi \right\}$$

by the polynomial interpolation of $q(\xi)$ in $\xi \in [c_k, c_{k+1}]$ with the potential data

$$\{q(c_k)\} \cup \{q(c_k + h_k t) : t \in \mathcal{S}\} \cup \{q(c_{k+1})\}$$

and the exact integration of the result.

3.5 Quadrature

Definition 8 below introduces $\tilde{\mathbf{I}}_{\lambda,i,\mathcal{T}_j}^{\text{fine}}(c_k, c_{k+1})$ and $\tilde{\mathbf{I}}_{\lambda,i,\mathcal{T}_j}(c_k, c_{k+1})$, which represent the quadrature developed above for (23)–(25). In particular, $\tilde{\mathbf{I}}_{\lambda,i,\mathcal{T}_j}(c_k, c_{k+1})$ are given by a rescaling of $\tilde{\mathbf{I}}_{\lambda,i,\mathcal{T}_j}^{\text{fine}}(c_k, c_{k+1})$, which, by construction, can be integrated exactly. Specifically:

- for global order 4, (23) is approximated by $\tilde{\mathbf{I}}_{\lambda,1,\mathcal{T}_3}(c_k, c_{k+1})$,
- for global order 7, (23) is approximated by $\tilde{\mathbf{I}}_{\lambda,1,\mathcal{T}_6}(c_k, c_{k+1})$ and (24) by $\tilde{\mathbf{I}}_{\lambda,2,\mathcal{T}_3}(c_k, c_{k+1})$,
- for global order 10, (23) is approximated by $\tilde{\mathbf{I}}_{\lambda,1,\mathcal{T}_9}(c_k, c_{k+1})$, (24) by $\tilde{\mathbf{I}}_{\lambda,2,\mathcal{T}_6}(c_k, c_{k+1})$ and (25) by $\tilde{\mathbf{I}}_{\lambda,3,\mathcal{T}_3}(c_k, c_{k+1})$.

Definition 8 Let

$$\begin{aligned}\tilde{\mathbf{I}}_{\lambda,1,\mathcal{T}_j}^{\text{fine}}(c_k, c_{k+1}) &:= \int_0^1 \tilde{\mathbf{B}}_{\lambda,1,\mathcal{T}_j}^{\text{fine}}(c_k, c_k + h_k t) dt, \\ \tilde{\mathbf{I}}_{\lambda,2,\mathcal{T}_j}^{\text{fine}}(c_k, c_{k+1}) &:= \int_0^1 \int_0^{t_1} [\tilde{\mathbf{B}}_{\lambda,1,\mathcal{T}_j}^{\text{fine}}(c_k, c_k + h_k t_2), \tilde{\mathbf{B}}_{\lambda,1,\mathcal{T}_j}^{\text{fine}}(c_k, c_k + h_k t_1)] dt_2 dt_1, \\ \tilde{\mathbf{I}}_{\lambda,3,\mathcal{T}_j}^{\text{fine}}(c_k, c_{k+1}) &:= \int_0^1 \int_0^{t_1} \int_0^{t_1} [\tilde{\mathbf{B}}_{\lambda,1,\mathcal{T}_j}^{\text{fine}}(c_k, c_k + h_k t_3), \\ &\quad [\tilde{\mathbf{B}}_{\lambda,1,\mathcal{T}_j}^{\text{fine}}(c_k, c_k + h_k t_2), \\ &\quad \tilde{\mathbf{B}}_{\lambda,1,\mathcal{T}_j}^{\text{fine}}(c_k, c_k + h_k t_1)]] dt_3 dt_2 dt_1.\end{aligned}$$

Let also $\tilde{\mathbf{I}}_{\lambda,i,\mathcal{T}_j}(c_k, c_{k+1}) \in \mathfrak{sl}(2, \mathbb{R})$ be the unique elements which satisfy

$$\begin{aligned}\pi(\tilde{\mathbf{I}}_{\lambda,1,\mathcal{T}_j}(c_k, c_{k+1})) &:= \pi(\tilde{\mathbf{I}}_{\lambda,1,\mathcal{T}_j}^{\text{fine}}(c_k, c_{k+1})) \\ &\quad \odot \begin{cases} h_k^2 [h_k \ h_k^2 \ 1]^\top, & \text{in (19) and (20),} \\ h_k^2 \left[\frac{1}{\sqrt{\lambda - q(c_k)}} \ \frac{1}{\lambda - q(c_k)} \ 1 \right]^\top, & \text{in (21),} \end{cases} \\ \pi(\tilde{\mathbf{I}}_{\lambda,2,\mathcal{T}_j}(c_k, c_{k+1})) &:= \pi(\tilde{\mathbf{I}}_{\lambda,2,\mathcal{T}_j}^{\text{fine}}(c_k, c_{k+1})) \\ &\quad \odot \begin{cases} h_k^5 [h_k \ h_k^2 \ 1]^\top, & \text{in (19) and (20),} \\ \frac{h_k^4}{\sqrt{\lambda - q(c_k)}} \left[\frac{1}{\sqrt{\lambda - q(c_k)}} \ \frac{1}{\lambda - q(c_k)} \ 1 \right]^\top, & \text{in (21),} \end{cases} \\ \pi(\tilde{\mathbf{I}}_{\lambda,3,\mathcal{T}_j}(c_k, c_{k+1})) &:= \pi(\tilde{\mathbf{I}}_{\lambda,3,\mathcal{T}_j}^{\text{fine}}(c_k, c_{k+1})) \\ &\quad \odot \begin{cases} h_k^8 [h_k \ h_k^2 \ 1]^\top, & \text{in (19) and (20),} \\ \frac{h_k^6}{\lambda - q(c_k)} \left[\frac{1}{\sqrt{\lambda - q(c_k)}} \ \frac{1}{\lambda - q(c_k)} \ 1 \right]^\top, & \text{in (21).} \end{cases}\end{aligned}$$

3.6 Discretization error estimates

Definition 9 and Theorem 11 below make explicit the quadrature error in Subsection 3.5, while Definition 10 and Theorem 12 below clarify the manner in which the quadrature error, which lives in the Lie algebra $\mathfrak{sl}(2, \mathbb{R})$, affects the quantities in the Lie group $\text{SL}(2, \mathbb{R})$.

Definition 9 Denote the quadrature of the integrals of $\mathbf{D}_{\lambda,1}(c_k, c_{k+1})$ and $\mathbf{D}_{\lambda,2}(c_k, c_{k+1})$ as developed in Subsection 3.5, for $r = 1, \log(3)/\log(2), 2$, i.e., for global order 4, 7, 10, by

$$\tilde{\mathbf{D}}_{\lambda,1,r}(c_k, c_{k+1}) := \begin{cases} \tilde{\mathbf{I}}_{\lambda,1,\mathcal{T}_3}(c_k, c_{k+1}), & \text{for } r = 1, \\ \tilde{\mathbf{I}}_{\lambda,1,\mathcal{T}_6}(c_k, c_{k+1}), & \text{for } r = \log(3)/\log(2), \\ \tilde{\mathbf{I}}_{\lambda,1,\mathcal{T}_9}(c_k, c_{k+1}), & \text{for } r = 2, \end{cases}$$

as well as

$$\tilde{\mathbf{D}}_{\lambda,2,r}(c_k, c_{k+1}) := \begin{cases} \mathbf{0}, & \text{for } r = 1, \\ -\frac{1}{2} \tilde{\mathbf{I}}_{\lambda,2,\mathcal{T}_3}(c_k, c_{k+1}), & \text{for } r = \log(3)/\log(2), \\ -\frac{1}{2} \tilde{\mathbf{I}}_{\lambda,2,\mathcal{T}_6}(c_k, c_{k+1}) + \frac{1}{3} \tilde{\mathbf{I}}_{\lambda,3,\mathcal{T}_3}(c_k, c_{k+1}), & \text{for } r = 2. \end{cases}$$

Theorem 11 *If Assumption 1 holds true: if $r \in \{1, \log(3)/\log(2), 2\}$, then*

$$\pi(\mathbf{D}_{\lambda,i}(c_k, c_{k+1}) - \tilde{\mathbf{D}}_{\lambda,i,r}(c_k, c_{k+1})) = h_{\max}^{3 \cdot 2^r - 1} \begin{cases} [\mathcal{O}(h_{\max}) \mathcal{O}(h_{\max}^2) \mathcal{O}(1)]^\top, & \text{in (19)–(20),} \\ \left[\frac{\mathcal{O}(1)}{\sqrt{\lambda - q_{\max}}} \frac{\mathcal{O}(1)}{\lambda - q_{\max}} \mathcal{O}(1) \right]^\top, & \text{in (21).} \end{cases}$$

Proof Follows from Theorems 8–10 and the discussion in Subsection 3.4. \square

Definition 10 Let $r \in \{1, \log(3)/\log(2), 2\}$, and define the discretized flow, the discretized solution, the discretization local error, and the discretization global error by, respectively:

$$\begin{aligned} \tilde{\mathbf{F}}_{\lambda,1}(c_k, c_{k+1}) &:= e^{\mathbf{D}_{\lambda,0}(c_k, c_{k+1})} e^{\tilde{\mathbf{I}}_{\lambda,1,\mathcal{F}_3}(c_k, c_{k+1})}, \\ \tilde{\mathbf{F}}_{\lambda, \log(3)/\log(2)}(c_k, c_{k+1}) &:= e^{\mathbf{D}_{\lambda,0}(c_k, c_{k+1})} e^{\tilde{\mathbf{I}}_{\lambda,1,\mathcal{F}_6}(c_k, c_{k+1})} e^{-\frac{1}{2}\tilde{\mathbf{I}}_{\lambda,2,\mathcal{F}_3}(c_k, c_{k+1})}, \\ \tilde{\mathbf{F}}_{\lambda,2}(c_k, c_{k+1}) &:= e^{\mathbf{D}_{\lambda,0}(c_k, c_{k+1})} e^{\tilde{\mathbf{I}}_{\lambda,1,\mathcal{F}_9}(c_k, c_{k+1})} e^{-\frac{1}{2}\tilde{\mathbf{I}}_{\lambda,2,\mathcal{F}_6}(c_k, c_{k+1}) + \frac{1}{3}\tilde{\mathbf{I}}_{\lambda,3,\mathcal{F}_3}(c_k, c_{k+1})}, \\ \tilde{\mathbf{Y}}_{\lambda,r}(c_{k+1}) &:= \tilde{\mathbf{F}}_{\lambda,r}(c_k, c_{k+1}) \cdots \tilde{\mathbf{F}}_{\lambda,r}(c_1, c_2) \tilde{\mathbf{F}}_{\lambda,r}(c_0, c_1), \\ \mathbf{L}_{\lambda,r}^{\text{disc.}}(c_k, c_{k+1}) &:= \log \left(\tilde{\mathbf{F}}_{\lambda, \lceil r \rceil}(c_k, c_{k+1}) \tilde{\mathbf{F}}_{\lambda,r}^{-1}(c_k, c_{k+1}) \right), \\ \mathbf{G}_{\lambda,r}^{\text{disc.}}(c_{k+1}) &:= \log \left(\tilde{\mathbf{Y}}_{\lambda, \lceil r \rceil}(c_{k+1}) \tilde{\mathbf{Y}}_{\lambda,r}^{-1}(c_{k+1}) \right). \end{aligned}$$

Theorem 12 *If Assumption 1 holds true, and $r \in \{1, \log(3)/\log(2), 2\}$, then*

$$\begin{aligned} \pi(\mathbf{L}_{\lambda,r}^{\text{disc.}}(c_k, c_{k+1})) &= h_{\max}^{3 \cdot 2^r - 1} \begin{cases} [\mathcal{O}(h_{\max}) \mathcal{O}(h_{\max}^2) \mathcal{O}(1)]^\top, & \text{in (19) and (20),} \\ \left[\frac{\mathcal{O}(1)}{\sqrt{\lambda - q_{\max}}} \frac{\mathcal{O}(1)}{\lambda - q_{\max}} \mathcal{O}(1) \right]^\top, & \text{in (21),} \end{cases} \\ \pi(\mathbf{G}_{\lambda,r}^{\text{disc.}}(c_{k+1})) &= h_{\max}^{3 \cdot 2^r - 2} \begin{cases} [\mathcal{O}(h_{\max}) \mathcal{O}(h_{\max}^2) \mathcal{O}(1)]^\top, & \text{in (19) and (20),} \\ \left[\frac{\mathcal{O}(1)}{\sqrt{\lambda - q_{\max}}} \frac{\mathcal{O}(1)}{\lambda - q_{\max}} \mathcal{O}(1) \right]^\top, & \text{in (21).} \end{cases} \end{aligned}$$

Proof As with the proof of Theorem 4, the main obstacle to estimate the local and global discretization errors in Definition 10, lies in that the lower-left entry of $\exp(\mathbf{D}_{\lambda,0}(c_k, c_{k+1}))$ can be arbitrarily large, as described in Theorem 3. This main obstacle can be circumvented by calling upon three Baker–Campbell–Hausdorff (BCH) type formulas

$$e^{\mathbf{X}} e^{\mathbf{Y}} = e^{\mathbf{X} + \mathbf{Y} + \frac{1}{2}[\mathbf{X}, \mathbf{Y}] + \frac{1}{12}([\mathbf{X}, [\mathbf{X}, \mathbf{Y}]] + [\mathbf{Y}, [\mathbf{Y}, \mathbf{X}]] + \cdots)} \quad (38)$$

$$e^{\mathbf{X}} e^{\mathbf{Y}} e^{-\mathbf{X}} = e^{\mathbf{Y} + [\mathbf{X}, \mathbf{Y}] + \frac{1}{2}[\mathbf{X}, [\mathbf{X}, \mathbf{Y}]] + \frac{1}{6}[\mathbf{X}, [\mathbf{X}, [\mathbf{X}, \mathbf{Y}]] + \cdots} \quad (39)$$

$$= \exp(\text{Ad}_{\exp(\mathbf{X})}(\mathbf{Y})). \quad (40)$$

Regarding local error, for $r = 1$ the local discretization error can be written as

$$\begin{aligned} \mathbf{L}_{\lambda,1}^{\text{disc.}}(c_k, c_{k+1}) &= \log \left(\tilde{\mathbf{F}}_{\lambda, \lceil 1 \rceil}(c_k, c_{k+1}) \tilde{\mathbf{F}}_{\lambda,1}^{-1}(c_k, c_{k+1}) \right) \\ &= \log \left(e^{\mathbf{D}_{\lambda,0}(c_k, c_{k+1})} e^{\mathbf{D}_{\lambda,1}(c_k, c_{k+1})} e^{-\tilde{\mathbf{D}}_{\lambda,1,1}(c_k, c_{k+1})} e^{-\mathbf{D}_{\lambda,0}(c_k, c_{k+1})} \right) \end{aligned}$$

$$\begin{aligned}
&= \log \left(e^{\mathbf{D}_{\lambda,0}(c_k, c_{k+1})} e^{\mathbf{D}_{\lambda,1}(c_k, c_{k+1}) - \tilde{\mathbf{D}}_{\lambda,1,1}(c_k, c_{k+1}) + \text{h.o.t.}} e^{-\mathbf{D}_{\lambda,0}(c_k, c_{k+1})} \right) \\
&= \text{Ad}_{\exp(\mathbf{D}_{\lambda,0}(c_k, c_{k+1}))} \left(\mathbf{D}_{\lambda,1}(c_k, c_{k+1}) - \tilde{\mathbf{D}}_{\lambda,1,1}(c_k, c_{k+1}) + \text{h.o.t.} \right) \\
&= \text{Ad}_{\exp(\mathbf{D}_{\lambda,0}(c_k, c_{k+1}))} \left(\mathbf{D}_{\lambda,1}(c_k, c_{k+1}) - \tilde{\mathbf{D}}_{\lambda,1,1}(c_k, c_{k+1}) \right) + \text{h.o.t.}
\end{aligned}$$

where the first and second equalities are due to Definitions 3 and 10, the third equality is due to (38) and the fourth equality is due to (40). More generally, for arbitrary r , one can show

$$\mathbf{L}_{\lambda,r}^{\text{disc.}}(c_k, c_{k+1}) = \text{Ad}_{\exp(\mathbf{D}_{\lambda,0}(c_k, c_{k+1}))} \left(\mathbf{D}_{\lambda,1}(c_k, c_{k+1}) - \tilde{\mathbf{D}}_{\lambda,1,r}(c_k, c_{k+1}) \right) + \text{h.o.t.}$$

which, with Theorems 3 and 11, yields the desired estimate. Regarding global error, for arbitrary r , the global discretization error obeys the recursion relation with initial condition

$$\mathbf{G}_{\lambda,r}^{\text{disc.}}(c_1) = \mathbf{L}_{\lambda,r}^{\text{disc.}}(c_0, c_1) \quad (41)$$

and general rule

$$\begin{aligned}
\mathbf{G}_{\lambda,r}^{\text{disc.}}(c_{k+1}) &= \log \left(\tilde{\mathbf{Y}}_{\lambda, \lceil r \rceil}(c_{k+1}) \tilde{\mathbf{Y}}_{\lambda,r}^{-1}(c_{k+1}) \right) \\
&= \log \left(\tilde{\mathbf{F}}_{\lambda, \lceil r \rceil}(c_k, c_{k+1}) \tilde{\mathbf{Y}}_{\lambda, \lceil r \rceil}(c_k) \tilde{\mathbf{Y}}_{\lambda,r}^{-1}(c_k) \tilde{\mathbf{F}}_{\lambda,r}^{-1}(c_k, c_{k+1}) \right) \\
&= \log \left(\tilde{\mathbf{F}}_{\lambda, \lceil r \rceil}(c_k, c_{k+1}) e^{\mathbf{G}_{\lambda,r}^{\text{disc.}}(c_k)} \tilde{\mathbf{F}}_{\lambda,r}^{-1}(c_k, c_{k+1}) \right) \\
&= \log \left(e^{\mathbf{L}_{\lambda,r}^{\text{disc.}}(c_k, c_{k+1})} \tilde{\mathbf{F}}_{\lambda,r}(c_k, c_{k+1}) e^{\mathbf{G}_{\lambda,r}^{\text{disc.}}(c_k)} \tilde{\mathbf{F}}_{\lambda,r}^{-1}(c_k, c_{k+1}) \right) \\
&= \log \left(e^{\mathbf{L}_{\lambda,r}^{\text{disc.}}(c_k, c_{k+1})} e^{\mathbf{D}_{\lambda,0}(c_k, c_{k+1})} e^{\mathbf{G}_{\lambda,r}^{\text{disc.}}(c_k) + \text{h.o.t.}} e^{-\mathbf{D}_{\lambda,0}(c_k, c_{k+1})} \right) \\
&= \log \left(e^{\mathbf{L}_{\lambda,r}^{\text{disc.}}(c_k, c_{k+1})} \exp \left(\text{Ad}_{\exp(\mathbf{D}_{\lambda,0}(c_k, c_{k+1}))} \left(\mathbf{G}_{\lambda,r}^{\text{disc.}}(c_k) + \text{h.o.t.} \right) \right) \right) \\
&= \log \left(e^{\mathbf{L}_{\lambda,r}^{\text{disc.}}(c_k, c_{k+1})} \exp \left(\text{Ad}_{\exp(\mathbf{D}_{\lambda,0}(c_k, c_{k+1}))} \left(\mathbf{G}_{\lambda,r}^{\text{disc.}}(c_k) \right) + \text{h.o.t.} \right) \right) \\
&= \mathbf{L}_{\lambda,r}^{\text{disc.}}(c_k, c_{k+1}) + \text{Ad}_{\exp(\mathbf{D}_{\lambda,0}(c_k, c_{k+1}))} \left(\mathbf{G}_{\lambda,r}^{\text{disc.}}(c_k) \right) + \text{h.o.t.} \quad (42)
\end{aligned}$$

where the first four equalities are due to Definitions 3 and 10, the fifth to (39), the sixth to (40), and the last to (38). The global discretization error expressions (41)–(42) then result in

$$\begin{aligned}
\mathbf{G}_{\lambda,r}^{\text{disc.}}(c_{k+1}) &= \text{Ad}_{\exp(\mathbf{D}_{\lambda,0}(c_k, c_{k+1}))} \left(\mathbf{D}_{\lambda,1}(c_k, c_{k+1}) - \tilde{\mathbf{D}}_{\lambda,1,r}(c_k, c_{k+1}) \right) + \dots \\
&\quad + \text{Ad}_{\exp(\mathbf{D}_{\lambda,0}(c_k, c_{k+1})) \dots \exp(\mathbf{D}_{\lambda,0}(c_0, c_1))} \left(\mathbf{D}_{\lambda,1}(c_0, c_1) - \tilde{\mathbf{D}}_{\lambda,1,r}(c_0, c_1) \right) + \text{h.o.t.}
\end{aligned}$$

which, with Assumption 1, Theorem 3 and Theorem 11, result in the desired estimate. \square

4 Total error estimates

Having controlled in Section 2 the truncation error incurred from approximating $\mathbf{F}_\lambda(c_k, c_{k+1})$ and $\mathbf{Y}_\lambda(c_{k+1})$ by $\tilde{\mathbf{F}}_{\lambda, \lceil r \rceil}(c_k, c_{k+1})$ and $\tilde{\mathbf{Y}}_{\lambda, \lceil r \rceil}(c_{k+1})$ for $\lceil r \rceil \in \mathbb{Z}^+$, and controlled in Section 3 the discretization error sustained from approximating $\tilde{\mathbf{F}}_{\lambda, \lceil r \rceil}(c_k, c_{k+1})$ and $\tilde{\mathbf{Y}}_{\lambda, \lceil r \rceil}(c_{k+1})$ by $\tilde{\tilde{\mathbf{F}}}_{\lambda,r}(c_k, c_{k+1})$ and $\tilde{\tilde{\mathbf{Y}}}_{\lambda,r}(c_{k+1})$ for $r \in \{1, \log(3)/\log(2), 2\}$, we now combine these results in Definition 11 and Theorem 13 below which show that the computable quantity $\tilde{\tilde{\mathbf{Y}}}_{\lambda,r}(c_{k+1})$ approximates $\mathbf{Y}_\lambda(c_{k+1})$ with global order 4, 7, 10 for $r = 1, \log(3)/\log(2), 2$, respectively.

The eigensystem of (1)–(2) can then be safely approximated by replacing $\mathbf{Y}_\lambda(c_{k+1})$ by $\tilde{\tilde{\mathbf{Y}}}_{\lambda,r}(c_{k+1})$ in (6) or in any other root-finding or shooting method that uses (4)–(5).

Definition 11 Let $r \in \{1, \log(3)/\log(2), 2\}$, and define the

$$\begin{aligned} \text{total local error:} \quad \mathbf{L}_{\lambda,r}^{\text{total}}(c_k, c_{k+1}) &:= \log \left(\mathbf{F}_{\lambda}(c_k, c_{k+1}) \tilde{\mathbf{F}}_{\lambda,r}^{-1}(c_k, c_{k+1}) \right), \\ \text{total global error:} \quad \mathbf{G}_{\lambda,r}^{\text{total}}(c_{k+1}) &:= \log \left(\mathbf{Y}_{\lambda}(c_{k+1}) \tilde{\mathbf{Y}}_{\lambda,r}^{-1}(c_{k+1}) \right). \end{aligned}$$

Theorem 13 If Assumption 1 holds true, and $r \in \{1, \log(3)/\log(2), 2\}$, then

$$\begin{aligned} \mathbf{L}_{\lambda,r}^{\text{total}}(c_k, c_{k+1}) &= \mathbf{L}_{\lambda, \lceil r \rceil}^{\text{trun.}}(c_k, c_{k+1}) + \mathbf{L}_{\lambda,r}^{\text{disc.}}(c_k, c_{k+1}) + \text{higher order terms}, \quad \text{in (19)–(21),} \\ \mathbf{G}_{\lambda,r}^{\text{total}}(c_{k+1}) &= \mathbf{G}_{\lambda, \lceil r \rceil}^{\text{trun.}}(c_{k+1}) + \mathbf{G}_{\lambda,r}^{\text{disc.}}(c_{k+1}) + \text{higher order terms}, \quad \text{in (19)–(21).} \end{aligned}$$

Proof The first statement follows from

$$\begin{aligned} \mathbf{L}_{\lambda,r}^{\text{total}}(c_k, c_{k+1}) &= \log \left(\mathbf{F}_{\lambda}(c_k, c_{k+1}) \tilde{\mathbf{F}}_{\lambda,r}^{-1}(c_k, c_{k+1}) \right) \\ &= \log \left(\mathbf{F}_{\lambda}(c_k, c_{k+1}) \tilde{\mathbf{F}}_{\lambda, \lceil r \rceil}^{-1}(c_k, c_{k+1}) \tilde{\mathbf{F}}_{\lambda, \lceil r \rceil}(c_k, c_{k+1}) \tilde{\mathbf{F}}_{\lambda,r}^{-1}(c_k, c_{k+1}) \right) \\ &= \log \left(\exp \left(\mathbf{L}_{\lambda, \lceil r \rceil}^{\text{trun.}}(c_k, c_{k+1}) \right) \exp \left(\mathbf{L}_{\lambda,r}^{\text{disc.}}(c_k, c_{k+1}) \right) \right) \\ &= \mathbf{L}_{\lambda, \lceil r \rceil}^{\text{trun.}}(c_k, c_{k+1}) + \mathbf{L}_{\lambda,r}^{\text{disc.}}(c_k, c_{k+1}) + \text{higher order terms}, \end{aligned}$$

where the first three equalities are due to Definitions 3, 10 and 11, and the last to (38). The second statement can be derived similarly by replacing $\mathbf{L}, \mathbf{F}, \tilde{\mathbf{F}}, \tilde{\mathbf{F}}$ by $\mathbf{G}, \mathbf{Y}, \tilde{\mathbf{Y}}, \tilde{\mathbf{Y}}$. \square

The previous theorem links the truncation estimates in [14] with the discretization estimates in this paper in that their sum controls the total error in the Fer streamers approach. As can be seen from Theorems 4 and 12, the bounds on the discretization errors

$$\mathbf{L}_{\lambda,r}^{\text{disc.}}(c_k, c_{k+1}), \quad \mathbf{G}_{\lambda,r}^{\text{disc.}}(c_{k+1}),$$

are larger than the bounds on the truncation errors

$$\mathbf{L}_{\lambda, \lceil r \rceil}^{\text{trun.}}(c_k, c_{k+1}), \quad \mathbf{G}_{\lambda, \lceil r \rceil}^{\text{trun.}}(c_{k+1}).$$

Hence, according to Theorem 13, the total errors obey

$$\mathbf{L}_{\lambda,r}^{\text{total}}(c_k, c_{k+1}) = \mathcal{O} \left(\mathbf{L}_{\lambda,r}^{\text{disc.}}(c_k, c_{k+1}) \right), \quad \mathbf{G}_{\lambda,r}^{\text{total}}(c_{k+1}) = \mathcal{O} \left(\mathbf{G}_{\lambda,r}^{\text{disc.}}(c_{k+1}) \right).$$

5 Conclusions

It has been shown that to preserve the advantageous properties of the Fer streamers approach to Sturm–Liouville problems in [14] under discretization, while simultaneously minimizing the computational complexity by reducing function evaluations and amount of linear algebra in the discretization, quadrature requires not the original representation of Fer streamers, but rather relies on a characterization designed to expose their magnitude and behaviour.

Tight total error estimates, uniform for every eigenvalue, have also been established that quantify the interplay between the truncation and discretization in the approach by Fer streamers which have shown that the discretization errors outweigh the truncation errors.

The principal advantage of the Fer streamers approach to Sturm–Liouville problems is that its truncation and discretization error estimates hold uniformly for all eigenvalues, which guarantee large step sizes over the eigenvalue range. This is especially significant given that the error estimates in alternative techniques apply only to ‘small’ or ‘large’ eigenvalues (c.f. Subsection 1.1). Compared with the alternative geometric integration techniques in the right-correction Magnus series [2] and in the modified Magnus methods [9], which do not possess error estimates uniform over the entire eigenvalue range, the Fer streamers approach presents an interesting trade-off in computational complexity: although it requires a 50% increase in function evaluations for each univariate integral in order to control all ‘small’, ‘intermediary’ and ‘large’ eigenvalues, it also enjoys a significant decrease in linear algebra for each multivariate integral (see Subsection 1.3).

These results open new questions which will be the subject of future work and include:

- Free Lie algebras, and,
- Step size strategies.

5.1 Free Lie algebras

As discussed in this paper, in order to approximate $\mathbf{Y}_\lambda(c_{k+1})$ by $\tilde{\mathbf{Y}}_{\lambda,r}(c_{k+1})$, it is necessary to compute $\tilde{\mathbf{F}}_{\lambda,r}(c_k, c_{k+1})$, which itself relies on the computation of $\tilde{\mathbf{I}}_{\lambda,i,\mathcal{J}_j}^{\text{fine}}(c_k, c_{k+1})$. Given that each integral $\tilde{\mathbf{I}}_{\lambda,i,\mathcal{J}_j}^{\text{fine}}(c_k, c_{k+1})$ can be integrated exactly by first expanding the individual terms of the commutators of the finite sums in $\tilde{\mathbf{B}}_{\lambda,1,\mathcal{J}_j}^{\text{fine}}(c_k, c_k + h_k t)$, then integrating exactly every summand, and finally adding up all the integrated terms, a useful question is how to perform such operations while incurring the smallest possible volume of linear algebra and, by extension, running time. Since this volume of linear algebra relates to the number of commutators that result from integrating the individual terms, there exist three mechanisms that can be used to decrease this volume of linear algebra, which are:

- Firstly, free Lie algebra techniques and Hall basis, which lead to fewer commutators via a systematic use of commutator identities such as: skew symmetry, Jacobi’s identity, etc,
- Secondly, when collected in a Hall basis, certain linear combinations between different integrands are then identically zero, and,
- Thirdly, when collected in a Hall basis, certain linear combinations between different integrands integrate exactly to zero.

Hence, these efficiency mechanisms will be explored to optimise running time.

5.2 Step size strategies

Another question for future work relates to mesh selection, which is known to depend on the interval length and the local variation of the potential.

For illustration purposes, consider a mesh (7) with equal spacing and largest step size that satisfies Assumption 1, namely: $m = \lceil (b-a)\sqrt{q_{\max}-q_{\min}} \rceil$ and $h = h_{\max} = h_{\min} = (b-a)/m$. In addition, take $r = 1, \log(3)/\log(2), 2$ in Theorem 13, i.e., global order 4, 7, 10, and choose Brent’s derivative-free root-finding algorithm to solve (6) with $\mathbf{Y}_\lambda(b)$ replaced by $\tilde{\mathbf{Y}}_{\lambda,r}(b)$. For concreteness, take the Coffey–Evans problem [13, p. 283]

$$\begin{aligned} a &= -\pi/2, \quad b = \pi/2, \quad \beta = 30, \quad q(t) = -2\beta \cos(2t) + \beta^2 \sin(2t)^2, \\ y_\lambda(a) &= y_\lambda(b) = 0, \quad \alpha_1 \neq 0, \quad \beta_1 \neq 0, \quad \alpha_2 = \beta_2 = 0, \end{aligned} \quad (43)$$

and the truncated Gelfand–Levitan problem [13, p. 283]

$$\begin{aligned} a = 0, \quad b = 100, \quad q(t) &= (32 \cos(t)(\cos(t) + (2+t) \sin(t)))/((4+2t + \sin(2t))^2), \\ y_\lambda(a) + y'_\lambda(a) &= 0, \quad y_\lambda(b) = 0, \quad \alpha_1 = \alpha_2 \neq 0, \quad \beta_1 \neq 0, \quad \beta_2 = 0. \end{aligned} \quad (44)$$

For these problems, Figure 1 depicts the absolute and relative errors between reference eigenvalues and eigenvalues computed via Fer streamers with the procedure described above.

Figure 1 highlights two issues related to the mesh strategy described at the start of this subsection. On the one hand, it yields step size $h = 0.03$ for (43) that gives relative error around 10^{-16} for $\lambda \geq 0.5 \times 10^5$ which may be more accuracy than necessary, and thus may consume more resources (e.g., running time or pointwise evaluations of the potential) than necessary. On the other hand, for (44) the same strategy yields relative error around 10^{-8} for every λ with step size $h = 0.57$.

These examples illustrate the fact that the behaviour of the solutions of Sturm–Liouville problems depend on the local behaviour of the potential q . With this in mind, one can:

- Relax Assumption 1 which currently uses global information about q , namely, its minimum q_{\min} and maximum q_{\max} , to instead use local information on each $[c_k, c_{k+1}]$ based on the pointwise evaluations of q at the interpolation points of Subsubsection 3.4.3, and,
- Explore non-equidistant meshes, and investigate automatic mesh selection algorithms, along the lines of [8].

Hence, these mesh algorithms will be explored to optimise running time and to decrease the number of pointwise evaluations of the potential function.

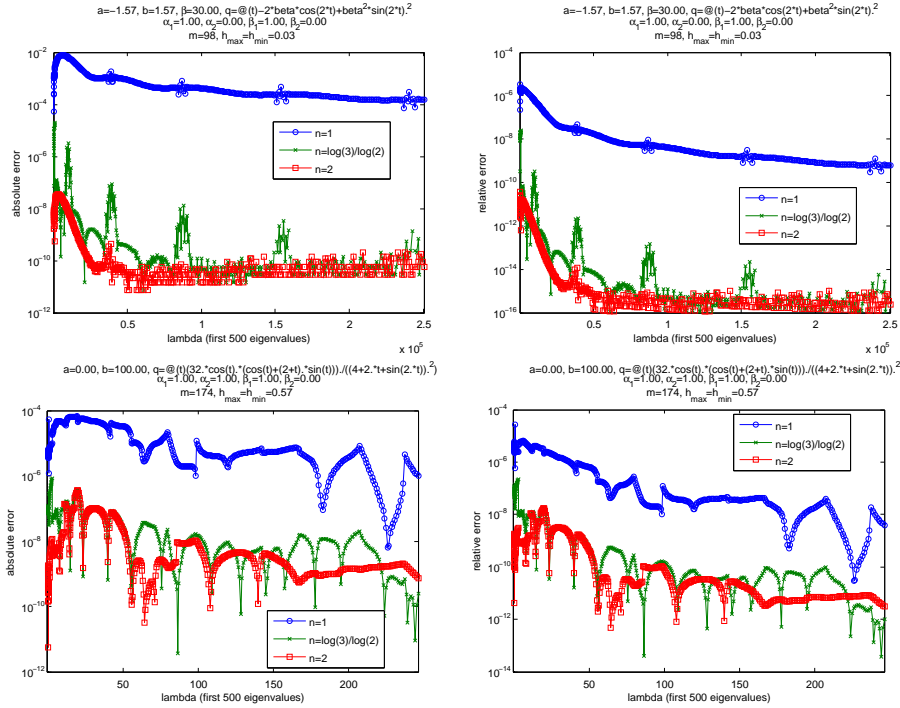


Fig. 1 Absolute (left) and relative (right) errors with Fer streamers with global order 4, 7, 10, i.e., $r = 1, \log(3)/\log(2), 2$, for the Sturm–Liouville problems (43) (top) and (44) (bottom).

Acknowledgments

The author would like to thank the anonymous referees for their comments and suggestions that greatly improved the presentation of the material in this paper, as well as for their thoughts and directions for future work. The author would also like to thank Arieh Iserles and Andreas Asheim for useful discussions. This work was supported by Fundação para a Ciência e a Tecnologia, Portugal through the fellowship SFRH/BD/71692/2010.

Compliance with Ethical Standards

Conflict of Interest: The author declares that he has no conflict of interest.

References

1. Degani, I. [2004], RCMS - Right Correction Magnus Schemes for Oscillatory ODEs, and Cubature Formulae and Commuting Extensions, PhD thesis, Weizmann Institute of Science, Department of Mathematics.
2. Degani, I. and Schiff, J. [2006], ‘RCMS: Right Correction Magnus Series approach for oscillatory ODEs’, *J. Comput. Appl. Math.* **193**(2), 413–436.
3. Fer, F. [1958], ‘Résolution del l’equation matricielle $\dot{U} = pU$ par produit infini d’exponentielles matricielles’, *Bulletin de la Classe des Sciences Académie Royale de Belgique* **44**, 818–829.
4. Iserles, A. [2004], ‘On the numerical quadrature of highly-oscillating integrals I: Fourier transforms’, *IMA Journal of Numerical Analysis* **24**(3), 365–391.
5. Iserles, A. and Nørsett, S. [2005], ‘Efficient quadrature of highly oscillatory integrals using derivatives’, *Proceedings of the Royal Society A* **461**, 1383–1399.
6. Iserles, A. and Nørsett, S. [2006], ‘Quadrature methods for multivariate highly oscillatory integrals using derivatives’, *Mathematics of Computation* **75**(255), 1233–1258.
7. Ixaru, L. [2000], ‘CP methods for the Schrödinger equation’, *Journal of Computational and Applied Mathematics* **125**(1-2), 347–357.
8. Kautský, J. and Nichols, N. K. [1980], ‘Equidistributing meshes with constraints’, *SIAM Journal on Scientific and Statistical Computing* **1**(4), 499–511.
9. Ledoux, V., Daele, M. and van den Berghe, G. [2010], ‘Efficient numerical solution of the 1D Schrödinger eigenvalue problem using Magnus integrators’, *IMA Journal of Numerical Analysis* **30**, 751–776.
10. Levin, D. [1996], ‘Fast integration of rapidly oscillatory functions’, *Journal of Computational and Applied Mathematics* **67**(1), 95–101.
11. Moan, P. C. [1998], Efficient approximation of Sturm–Liouville problems using Lie-group methods, Technical report, Department of Applied Mathematics and Theoretical Physics, University of Cambridge, United Kingdom.
12. Olver, F. W. J., Lozier, D. W., Boisvert, R. F. and Clark, C. W., eds [2010], *NIST Handbook of Mathematical Functions*, Cambridge University Press, New York, NY.
13. Pryce, J. [1993], *Numerical Solution of Sturm–Liouville Problems*, Oxford University Press.
14. Ramos, A. G. C. P. and Iserles, A. [2015], ‘Numerical solution of Sturm–Liouville problems via Fer streamers’, *Numerische Mathematik* **131**(3), 541–565.