

Effective Approximation for the Semiclassical Schrödinger Equation

Philipp Bader · Arieh Iserles ·
Karolina Kropielnicka · Pranav Singh

Received: 16 July 2012 / Revised: 17 August 2013 / Accepted: 4 November 2013 /
Published online: 19 February 2014
© SFoCM 2014

Abstract The computation of the semiclassical Schrödinger equation presents major challenges because of the presence of a small parameter. Assuming periodic boundary conditions, the standard approach consists of semi-discretisation with a spectral method, followed by an exponential splitting. In this paper we sketch an alternative strategy. Our analysis commences with the investigation of the free Lie algebra generated by differentiation and by multiplication with the interaction potential: it turns out that this algebra possesses a structure which renders it amenable to a very effective form of *asymptotic splitting*: exponential splitting where consecutive terms are scaled by increasing powers of the small parameter. This leads to methods which attain high spatial and temporal accuracy and whose cost scales as $\mathcal{O}(M \log M)$, where M is the number of degrees of freedom in the discretisation.

Keywords Semiclassical Schrödinger equation · Exponential splittings · Zassenhaus splitting · Spectral collocation · Krylov subspace methods

Communicated by Elizabeth Mansfield.

P. Bader

Instituto de Matemática Multidisciplinar, Universitat Politècnica de València,
Edificio 8G, piso 2, Camino de Vera s/n, 46022 Valencia, Spain

A. Iserles (✉) · K. Kropielnicka · P. Singh

Department of Applied Mathematics and Mathematical Physics,
University of Cambridge, Wilberforce Road, Cambridge CB3 0WA, UK
e-mail: a.iserles@damtp.cam.ac.uk

K. Kropielnicka

Institute of Mathematics, University of Gdańsk,
Wit Stwosz Str. 57, 90-952 Gdańsk, Poland

Mathematics Subject Classification Primary 65M70 · Secondary 35Q41 · 65L05 · 65F60

1 Introduction

The semiclassical Schrödinger equation plays a central role in a wide range of applications and is the fundamental model of quantum mechanics [7]. Its computation presents numerous enduring challenges [13] which form the centrepiece of this paper.

We consider the standard *linear Schrödinger equation* in a single space variable,

$$i\hbar \frac{\partial u}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 u}{\partial x^2} - \tilde{V}(x)u, \quad t \geq 0, \quad x \in [-(2m)^{-1/2}, (2m)^{-1/2}], \quad (1.1)$$

where $u = u(x, t)$ is given with an initial condition and periodic boundary conditions, the *interaction potential* \tilde{V} is a periodic function and m is the mass of the underlying particle. The parameter \hbar , the *reduced Planck constant*, is truly minute, $\hbar \approx 1.05457168 \times 10^{-34}$ J s, whilst m is a small quantity, although substantially larger than \hbar . However, since physical interest is in fairly small spatial and temporal ‘windows’, it is usual to rescale so that (1.1) is replaced with

$$i\varepsilon \frac{\partial u}{\partial t} = -\varepsilon^2 \frac{\partial^2 u}{\partial x^2} - V(x)u, \quad t \geq 0, \quad x \in [-1, 1], \quad (1.2)$$

where $\varepsilon > 0$ is a small parameter: it is useful to keep in mind the range $10^{-8} \leq \varepsilon \leq 10^{-2}$.

Equation (1.2) is a univariate model for the considerably more important multivariate semiclassical Schrödinger equation with periodic boundary conditions,

$$i\varepsilon \frac{\partial u}{\partial t} = -\varepsilon^2 E \nabla^2 u - V(\mathbf{x})u, \quad t \geq 0, \quad \mathbf{x} \in [-1, 1]^d, \quad (1.3)$$

where $u = u(t, \mathbf{x})$ and E is a diagonal matrix. The methodology of this paper lends itself to straightforward generalisation to (1.3), provided that the dimension d is moderate. Large values of d require combining our approach with other computational techniques, an area under current investigation.

The small size of ε is a source of substantial difficulties in the numerical discretization of (1.2) because, using a naïve approach, rapid oscillations require a resolution of $\mathcal{O}(\varepsilon)$ in both space and time, which is often impractical or, at best, exceedingly expensive. This is the motivation to pursue alternative approaches based in the main on the concept of *exponential splittings* [4, 13–15].

The construction of exponential splitting methods typically commences from space discretisation. Rewriting (1.2) in the form

$$\frac{\partial u}{\partial t} = i\varepsilon \frac{\partial^2 u}{\partial x^2} + i\varepsilon^{-1} V(x)u, \quad t \geq 0, \quad x \in [-1, 1], \quad (1.4)$$

we let the vector $\mathbf{u}(t) \in \mathbb{C}^M$ represent an approximation to the solution at time t : typically, the components of \mathbf{u} are either approximations to the values of u on a

spatial grid or to Fourier coefficients of this function. Replacing the second-derivative operator by a matrix \mathcal{K} (i.e. replacing an infinite-dimensional linear operator by a finite-dimensional one), we obtain the ODE system

$$\mathbf{u}' = i(\varepsilon\mathcal{K} + \varepsilon^{-1}\mathcal{D})\mathbf{u}, \quad t \geq 0, \tag{1.5}$$

where $\mathbf{u}(0)$ is derived from the initial conditions and \mathcal{D} represents a multiplication by the interaction potential V in the finite-dimensional space.

The exact solution of (1.5) is, of course,

$$\mathbf{u}(t) = \exp\left(it(\varepsilon\mathcal{K} + \varepsilon^{-1}\mathcal{D})\right)\mathbf{u}(0),$$

and a natural temptation is to approximate it (using small time steps) by any of many methods to compute the matrix exponential, $\mathbf{u}((n + 1)\Delta t) \approx e^{i\Delta t(\varepsilon\mathcal{K} + \varepsilon^{-1}\mathcal{D})}\mathbf{u}(n\Delta t)$, $n \in \mathbb{Z}_+$. This is generally accepted as a poor idea because the vastly different scales of $\varepsilon\mathcal{K}$ and $\varepsilon^{-1}\mathcal{D}$ require either a very small time step Δt or exceedingly expensive methods to approximate the exponential (e.g. Krylov subspace methods of large dimension) to attain reasonable accuracy. The alternative is to separate scales by means of an exponential splitting. The starting point is usually the *Strang splitting*:

$$e^{it(\varepsilon\mathcal{K} + \varepsilon^{-1}\mathcal{D})} = e^{\frac{1}{2}it\varepsilon\mathcal{K}}e^{it\varepsilon^{-1}\mathcal{D}}e^{\frac{1}{2}it\varepsilon\mathcal{K}} + \mathcal{O}(t^3). \tag{1.6}$$

This has the clear virtue of separating scales. Moreover, usually each individual exponential can be computed very affordably, e.g. once we semidiscretise (1.4) using a spectral method, \mathcal{K} is diagonal and \mathcal{D} a circulant; therefore, $e^{\frac{1}{2}it\varepsilon\mathcal{K}}$ is a diagonal matrix, whilst $e^{it\varepsilon^{-1}\mathcal{D}}$ can be approximated in $\mathcal{O}(M \log M)$ operations with a fast Fourier transform (FFT). Yet the order of approximation is unacceptably low. The standard generalisation of the Strang splitting bears the form

$$e^{i\alpha_1 t \varepsilon \mathcal{K}} e^{i\beta_1 t \varepsilon^{-1} \mathcal{D}} e^{i\alpha_2 t \varepsilon \mathcal{K}} \dots e^{i\alpha_r t \varepsilon \mathcal{K}} e^{i\beta_r t \varepsilon^{-1} \mathcal{D}} e^{i\alpha_r t \varepsilon \mathcal{K}} \dots e^{i\alpha_2 t \varepsilon \mathcal{K}} e^{i\beta_1 t \varepsilon^{-1} \mathcal{D}} e^{i\alpha_1 t \varepsilon \mathcal{K}}.$$

The palindromic form of this splitting (it reads the same from the left and from the right), which is referred to as symmetric splitting in much of the literature, is not accidental since it guarantees higher order. The coefficients α_i and β_i are typically chosen to ensure higher order (because of palindromy, the order is always even), smaller error constants, or both [1, 15].

This approach retains the main virtues of (1.6), namely separation of scales and the ease of computation of individual exponentials. However, an inordinately large number of exponentials is required to attain significant order. The simplest means towards a high-order splitting, the *Yosida method* [15, 20], calls for $r = 3^{p-1}$ (which translates to $2 \times 3^{p-1} + 1$ exponentials) to attain order $2p$. Our aim in this paper is to present splittings which require far fewer exponentials to attain a given order: we wish the number of exponentials to grow linearly, rather than exponentially, with order. Moreover, once the number of exponentials becomes large, ideally we do not want

all of them to fit into the same two scales but wish for them to become increasingly smaller: to have an *asymptotic splitting*.

In this paper we introduce a family of exponential splittings with these favourable features. More specifically, we introduce and analyse exponential splittings of the form

$$e^{i\Delta t(\varepsilon\mathcal{K}+\varepsilon^{-1}\mathcal{D})} = e^{\mathcal{R}_0}e^{\mathcal{R}_1} \dots e^{\mathcal{R}_s}e^{\mathcal{T}_{s+1}}e^{R_s} \dots e^{\mathcal{R}_1}e^{\mathcal{R}_0} + \mathcal{O}\left(\varepsilon^{2s+2}\right), \tag{1.7}$$

where

$$\begin{aligned} \mathcal{R}_0 &= \mathcal{R}_0(\Delta t, \varepsilon, \mathcal{K}, \mathcal{D}) = \mathcal{O}\left(\varepsilon^0\right), \\ \mathcal{R}_k &= \mathcal{R}_k(\Delta t, \varepsilon, \mathcal{K}, \mathcal{D}) = \mathcal{O}\left(\varepsilon^{2k-2}\right), \quad k = 1, \dots, s, \\ \mathcal{T}_{s+1} &= \mathcal{T}_{s+1}(\Delta t, \varepsilon, \mathcal{K}, \mathcal{D}) = \mathcal{O}\left(\varepsilon^{2s}\right), \end{aligned}$$

and variations on this theme. Note a number of critical differences between (1.7) and standard exponential splittings.

Firstly, we quantify the error not in terms of the step-size Δt but of the small parameter ε . There are three small quantities at play: ε , Δt and $1/M$ (where M is the number of degrees of freedom in the semidiscretisation). By letting power laws govern the relationship between ε and the choices of Δt and M , we express the error in the single quantity ε .

Secondly, the number of individual terms in (1.7) is remarkably small and grows *linearly* with s – compare with the exponential growth, as a function of order, in the number of components of familiar splittings. The reason is that the arguments of the exponentials in (1.7) decay increasingly more rapidly in ε .

Thirdly, each of these exponentials can be computed fairly easily. Some of the \mathcal{R}_k s are diagonal matrices, whereby computing the exponential is trivial. Others are circulants and can be computed with FFT. Finally, because of the minute spectral radius of the arguments for sufficiently large k , the remaining exponentials can be evaluated up to the requisite power of ε using a *very* low-dimensional Krylov subspace method. All in all, the cost of these splittings ends up being cubic in the desired order, in contrast with the exponential cost of the Yōšida method.

The asymptotic splitting (1.7) is possible because we have deliberately breached the consensus in the design of exponential splittings: the terms \mathcal{R}_k and \mathcal{T}_{s+1} contain nested commutators. The use of commutators is usually frowned upon because of their cost and because they are believed to increase in norm. However, as we demonstrate in Sect. 2, in the current setting the use of commutators, appropriately handled, is benign. The first idea is to forego the standard steps of first semidiscretising like in (1.5) and then splitting the exponential: we semidiscretise only once the splitting has been done! Thus, the entire narrative takes place within the *free Lie algebra* $\mathfrak{F} = \text{FLA}\{\partial_x^2, V\}$, where $\partial_x = \frac{d}{dx}$ and V is the operation of multiplying with the interaction potential: since we have not yet discretised, both are infinite-dimensional linear operators. We demonstrate in Sect. 2 that \mathfrak{F} can be embedded in a larger Lie algebra \mathfrak{G} , where the commutation has a simple, straightforward interpretation. For all intents and purposes,

commutators are replaced by simple linear combinations of powers of ∂_x . Moreover – and this is what allows this entire procedure to work in a beneficial manner – these are smaller powers of ∂_x than naively expected. Section 2 also describes two Lie-algebraic concepts which are at the heart of our methodology, the symmetric Baker–Campbell–Hausdorff (BCH) formula and the Zassenhaus splitting.

In Sect. 3 we introduce – still working in an infinite-dimensional operatorial setting – our exponential splitting. This requires a recursive procedure, based upon repeated application of the symmetric BCH formula in \mathfrak{G} , working in the Hall basis. Although the underlying algebra is time consuming, it need be done just once, and the outcome is fairly simple.

Section 4 is concerned with semidiscretisation. We consider several alternatives, finally opting for spectral collocation. This allows us to work with nodal function values, attain spectral speed of convergence and calculate matrix exponentials very effectively and affordably.

The computation of matrix exponentials is the theme of Sect. 5. Most exponentials in (1.7) are trivial to calculate because the underlying matrix is either diagonal or a circulant. The one exception are matrices of size $\mathcal{O}(\varepsilon^\alpha)$ for sufficiently large $\alpha > 0$. Owing to the small size of their argument, once they are calculated by Krylov subspace methods, the price tag is very small.

We discuss some variants of the splittings in Sect. 6 and present a number of numerical results concerning these in Sect. 7. Section 8 is devoted to brief conclusions and pointers for future research.

2 A Lie-Algebraic Setting

2.1 An Algebra of Operators

The vector field in the semiclassical Schrödinger equation (1.4) is a linear combination of the action of two operators, ∂_x^2 and multiplication by the interaction potential V . Since the calculation of exponential splittings entails nested commutation, the focus of our interest is on the free Lie algebra

$$\mathfrak{F} = \text{FLA}\{\partial_x^2, V\},$$

i.e. the linear-space closure of all nested commutators generated by ∂_x^2 and V . The elements of \mathfrak{F} are operators acting on sufficiently smooth functions including the initial value of (1.4): for the purpose of this paper and for simplicity’s sake we assume that the initial value, and hence the solution of (1.4) for moderate values of $t \geq 0$, is a periodic function in $C^\infty[-1, 1]$, but our results extend in a straightforward manner to functions of lower smoothness.

To compute commutators, we need in principle to describe their action on functions, e.g.

$$[V, \partial_x^2]u = V(\partial_x^2 u) - \partial_x^2(Vu) = -(\partial_x^2 V)u - 2(\partial_x V)\partial_x u$$

Table 1 Terms of Hall basis of \mathfrak{F} of grade ≤ 4

j	Nested commutator	χ_j	Grade
H_1	∂_x^2	1	1
H_2	V	1	1
H_3	$[V, \partial_x^2]$	0	2
H_4	$[[V, \partial_x^2], \partial_x^2]$	$-\frac{1}{24}$	3
H_5	$[[V, \partial_x^2], V]$	$-\frac{1}{12}$	3
H_6	$[[[V, \partial_x^2], \partial_x^2], \partial_x^2]$	0	4
H_7	$[[[V, \partial_x^2], \partial_x^2], V]$	0	4
H_8	$[[[V, \partial_x^2], V], V]$	0	4

implies that $[V, \partial_x^2] = -(\partial_x^2 V) - 2(\partial_x V)\partial_x$. This algebra necessitates knowing derivatives of the interaction potential, which are assumed for the scope of this paper to be given exactly but in practice can be obtained via differentiation matrices. The higher derivatives of V appearing in our splitting need to be known only to a certain accuracy, and spectral methods or finite difference methods of fairly reasonable orders suffice. It must be noted that these derivatives, if not given exactly, need to be derived only once, and the overhead is bearable.

We list the lowest-order commutators which form a so-called Hall basis [18] of the free Lie algebra \mathfrak{F} in Table 1. In the table, *Grade* refers to the number of “letters” V and ∂_x^2 in the expression, whilst χ_j is the coefficient of this term in the symmetric BCH formula; cf. Sect. 2.2.

Computing the commutators $H_j, j = 3, 4, \dots, 8$, explicitly, we have

$$\begin{aligned}
 H_3 &= -(\partial_x^2 V) - 2(\partial_x V)\partial_x, \\
 H_4 &= (\partial_x^4 V) + 4(\partial_x^3 V)\partial_x + 4(\partial_x^2 V)\partial_x^2, \\
 H_5 &= -2(\partial_x V)^2, \\
 H_6 &= -(\partial_x^6 V) - 6(\partial_x^5 V)\partial_x - 12(\partial_x^4 V)\partial_x^2 - 8(\partial_x^3 V)\partial_x^3, \\
 H_7 &= 4 \left[(\partial_x V)(\partial_x^3 V) + (\partial_x^2 V)^2 \right] + 8(\partial_x V)(\partial_x^2 V)\partial_x, \\
 H_8 &= 0.
 \end{aligned}$$

We note that all the terms belong to the set

$$\mathfrak{G} = \left\{ \sum_{k=0}^n y_k(x) \partial_x^k : n \in \mathbb{Z}_+, y_0, \dots, y_n \in C^\infty[-1, 1] \text{ periodic with period } 2 \right\}.$$

It is trivial to observe that \mathfrak{G} is itself a Lie algebra.

There are numerous cancellations, similar to $H_8 = 0$, because of the special structure induced by the letters ∂_x^2 and $V(x)$; nevertheless, for our exposition it is more appropriate to operate in the larger Lie algebra \mathfrak{G} , where all cancellations will be taken care of by simple computation of the commutators, according to

$$\left[\sum_{i=0}^n f_i(x) \partial_x^i, \sum_{j=0}^m g_j(x) \partial_x^j \right] = \sum_{i=0}^n \sum_{j=0}^m \sum_{\ell=0}^i \binom{i}{\ell} f_i(x) \left(\partial_x^{i-\ell} g_j(x) \right) \partial_x^{\ell+j} - \sum_{j=0}^m \sum_{i=0}^n \sum_{\ell=0}^j \binom{j}{\ell} g_j(x) \left(\partial_x^{j-\ell} f_i(x) \right) \partial_x^{\ell+i}. \tag{2.1}$$

2.2 Symmetric BCH Formula

Let X and Y be two terms in a Lie algebra \mathfrak{g} . The *symmetric Baker–Campbell–Hausdorff formula* (usually known in abbreviated form as the symmetric BCH formula) is

$$e^{\frac{1}{2}X} e^Y e^{\frac{1}{2}X} = e^{\text{sBCH}(X,Y)}, \tag{2.2}$$

where

$$\begin{aligned} \text{sBCH}(tX, tY) &= t(X + Y) - t^3 \left(\frac{1}{24} [[Y, X], X] + \frac{1}{12} [[Y, X], Y] \right) \\ &+ t^5 \left(\frac{7}{5,760} [[[[Y, X], X], X], X] + \frac{7}{1,440} [[[[Y, X], X], X], Y] \right. \\ &+ \frac{1}{180} [[[[Y, X], X], Y], Y] + \frac{1}{720} [[[[Y, X], Y], Y], Y] \\ &+ \left. \frac{1}{480} [[[[Y, X], X], [Y, X]] - \frac{1}{360} [[[[Y, X], Y], [Y, X]]] \right) \\ &+ \mathcal{O}(t^7). \end{aligned} \tag{2.3}$$

Expansion (2.3) can be computed to an arbitrary power of t using an algorithm from [3]. [Because (2.3) is palindromic, only odd powers of t feature in the expansion.] An observant reader would have noticed that the coefficients are the numbers χ_j from Table 1. This is not accidental: once we let $X = \partial_x^2$ and $Y = V$, the table lists the coefficients up to grade four.

2.3 Zassenhaus Splitting

Unless X and Y commute, it is in general not true that $e^{t(X+Y)} = e^{tX} e^{tY}$. The *Zassenhaus splitting* [17]

$$e^{t(X+Y)} = e^{tX} e^{tY} e^{t^2 U_2(X,Y)} e^{t^3 U_3(X,Y)} e^{t^4 U_4(X,Y)} \dots, \tag{2.4}$$

where

$$\begin{aligned}
 U_2(X, Y) &= \frac{1}{2}[Y, X], \\
 U_3(X, Y) &= \frac{1}{3}[[Y, X], Y] + \frac{1}{6}[[Y, X], X], \\
 U_4(X, Y) &= \frac{1}{24}[[[Y, X], X], X] + \frac{1}{8}[[[Y, X], X], Y] + \frac{1}{8}[[[Y, X], Y], Y],
 \end{aligned}$$

quantifies this discrepancy. (More terms can be generated using the – non-symmetric – BCH formula.)

The splitting (2.4) is not well known and seldom used in computation, for the perfectly valid reason that it is not palindromic. The natural temptation is thus to *symmetrise* it and consider a palindromic splitting of the form

$$e^{t(X+Y)} = \dots e^{t^5 Q_5(X,Y)} e^{t^3 Q_3(X,Y)} e^{\frac{1}{2}tX} e^{tY} e^{\frac{1}{2}tX} e^{t^3 Q_3(X,Y)} e^{t^5 Q_5(X,Y)} \dots, \tag{2.5}$$

where we can deduce by inspection of (2.3) that, for example,

$$Q_3(X, Y) = \frac{1}{48}[[Y, X], X] + \frac{1}{24}[[Y, X], Y].$$

Rather than engaging in increasingly tedious calculations to compute Q_5 , we replace (2.5) by a more computation-friendly splitting. We commence from the symmetric BCH formula (2.3),

$$e^{-\frac{1}{2}tX} e^{t(X+Y)} e^{-\frac{1}{2}tX} = e^{\text{sBCH}(-tX, t(X+Y))},$$

which we rewrite in the form

$$e^{t(X+Y)} = e^{\frac{1}{2}tX} e^{\text{sBCH}(-tX, t(X+Y))} e^{\frac{1}{2}tX}. \tag{2.6}$$

It follows from (2.3) that

$$\text{sBCH}(-tX, t(X + Y)) = \mathcal{W}^{[1]} = tY + \mathcal{O}(t^3),$$

and we note that we have extracted the outer term tX from the inner exponent. We iterate (2.6) over the resulting term and continue to symmetrically pull out the lowest-order terms, one by one, until the central exponent reaches the desired high order,

$$\begin{aligned}
 \exp t(X + Y) &= e^{\frac{1}{2}tX} e^{\text{sBCH}(-tX, t(X+Y))} e^{\frac{1}{2}tX} \\
 &= e^{\frac{1}{2}tX} e^{\frac{1}{2}tY} e^{\text{sBCH}(-tY, \text{sBCH}(-tX, t(X+Y)))} e^{\frac{1}{2}tY} e^{\frac{1}{2}tX} = \dots .
 \end{aligned}$$

Notice that by pulling out, we essentially subtract a term and add higher-order corrections. It is important to observe that the order of the exponent given by the sBCH

formula (2.6) is never decreased by this procedure,¹ and thus we can easily control the order of the approximation error when truncating the BCH formula. With the notation

$$\mathcal{W}^{[k+1]} = \text{sBCH}\left(-W^{[k]}, \mathcal{W}^{[k]}\right), \quad \mathcal{W}^{[0]} = t(X + Y), \tag{2.7}$$

the result after s steps can be written as

$$\exp t(X + Y) = e^{\frac{1}{2}W^{[0]}} e^{\frac{1}{2}W^{[1]}} \dots e^{\frac{1}{2}W^{[s]}} e^{\mathcal{W}^{[s+1]}} e^{\frac{1}{2}W^{[s]}} \dots e^{\frac{1}{2}W^{[1]}} e^{\frac{1}{2}W^{[0]}}.$$

We emphasise that, in principle, we can freely choose the elements $W^{[k]}$ which we want to extract. A first idea is to choose the $W^{[k]} = \mathcal{O}(t^{2k-1})$ for $k > 0$ and $W^{[0]} = \mathcal{O}(t)$, which yields a separation of powers, analogous to (2.5), and thus for s stages and approximating $\mathcal{W}^{[s+1]} = W^{[s+1]} + \mathcal{O}(t^{2s+3})$, we obtain a symmetric Zassenhaus splitting of order $2s + 2$.

We have almost established the splitting (1.7) – ‘almost’ because of yet another consideration. In standard splittings, e.g. in the context of a numerical solution of Hamiltonian ordinary differential equations, there is usually a single small parameter, Δt (the time step), and it makes perfect sense to expand in its powers. However, once we contemplate the discretisation of (1.4), we have *three* small parameters to reckon with:

1. The built-in small parameter ε ;
2. The time step Δt ;
3. $1/M$, where M is the number of degrees of freedom in the spatial semidiscretisation.

Although we derive our splitting *before* the infinite-dimensional operator ∂_x^2 has been discretised, we must keep the eventual discretisation at the back of our minds. In other words, sooner or later (more specifically, in Sect. 4) we replace ∂_x^2 with a *differentiation matrix* acting on an appropriate M -dimensional space: M might be the number of nodal values or of Fourier modes. It is elementary that the norm of a differentiation matrix corresponding to ∂_x^n scales as $\mathcal{O}(M^n)$, $n \in \mathbb{N}$. Therefore, we employ in our analysis the shorthand $\partial_x^n = \mathcal{O}(M^n)$.

We propose to deal with three small parameters in unison by converting them into a single currency. More specifically, we assume that our choice of Δt and M is governed by the *scaling laws*

$$M = \mathcal{O}(\varepsilon^{-\rho}), \quad \Delta t = \mathcal{O}(\varepsilon^\sigma), \tag{2.8}$$

where $\rho, \sigma > 0$ are given. As a consequence, each ∂_x^n scales as $\mathcal{O}(\varepsilon^{-n\rho})$.

The choice of these parameters is not entirely arbitrary. *Wentzel–Kramers–Brillouin* (WKB) analysis of the semiclassical Schrödinger equation [13] shows that even with arbitrary well-behaved initial conditions the full solution develops spatial oscillations of order $\mathcal{O}(\varepsilon^{-1})$. A reasonable approximation of the solution therefore necessitates taking $M = \mathcal{O}(\varepsilon^{-1})$ Fourier modes or nodal values at the very least, restricting us to $\rho \geq 1$.

¹ Unless a non-existing term is subtracted and thus newly introduced instead of removed.

It must be noted that, despite equally high oscillations in time, similar considerations do not apply to σ . This is because the solution in time is obtained via exponentials and is exact except for the omission of terms in the splitting. The accuracy in time is tied to the accuracy of the splitting.

The simplest choice of parameters in (2.8), keeping the preceding considerations in mind, is $\rho = \sigma = 1$, and this is what we assume in the next section.

3 An Asymptotic Splitting

3.1 Towards an Asymptotic Splitting

Recalling that $\rho = \sigma = 1$, we commence in this section with the asymptotic splitting (1.7) with $s = 2$, i.e. bearing the error of $\mathcal{O}(\varepsilon^6)$. Given that $\varepsilon > 0$ is very small, this presents a method which is very accurate – arguably, of higher accuracy than is required in standard numerical computations. We will expand the commutators in powers of ε and successively remove them from the core of our expansion, aiming for $W^{[j]} = \mathcal{O}(\varepsilon^{2j-2})$, except for $W^{[0]}$, which will be $\mathcal{O}(\varepsilon^0)$. Our next observation is that Δt is always multiplied by i ; therefore, it is handy to let

$$\tau = i\Delta t = \mathcal{O}(\varepsilon).$$

Note that $\tau\varepsilon\partial_x^2 = \mathcal{O}(\varepsilon^0)$ and $\tau\varepsilon^{-1}V = \mathcal{O}(\varepsilon^0)$, or more generally

$$\tau^\ell\varepsilon^{-m}\partial_x^n = \mathcal{O}(\varepsilon^{\ell-m-n}), \quad \varepsilon \rightarrow 0. \tag{3.1}$$

We can now commence algorithm (2.7), setting

$$\mathcal{W}^{[0]} = \tau\varepsilon^{-1}V + \tau\varepsilon\partial_x^2, \quad W^{[0]} = \tau\varepsilon^{-1}V.$$

With the help of (2.3), we compute the commutators in $\mathcal{W}^{[1]} = \text{sBCH}(-W^{[0]}, \mathcal{W}^{[0]})$ according to (2.1). This task confronts us with long and tedious algebra but can, however be automatized with a computer algebra programme. It is worth pointing out that all simplifications, such as $[[[V, \partial_x^2], V], V] = 0$, are automatically performed once we work in the larger Lie algebra \mathfrak{G} with differential operators and scalar functions. Likewise, there is no need for a tedious representation of expansion elements in, say, the Hall basis because this is done automatically in \mathfrak{G} .

Substituting and aggregating terms of the same order of magnitude, we obtain

$$\begin{aligned} \mathcal{W}^{[1]} = & \underbrace{\tau\varepsilon\partial_x^2}_{\mathcal{O}(\varepsilon^0)} + \underbrace{\frac{1}{12}\tau^3\varepsilon^{-1}(\partial_x V)^2 - \frac{1}{3}\tau^3\varepsilon(\partial_x^2 V)\partial_x^2}_{\mathcal{O}(\varepsilon^2)} - \underbrace{\frac{1}{3}\tau^3\varepsilon(\partial_x^3 V)\partial_x}_{\mathcal{O}(\varepsilon^3)} \\ & + \underbrace{\frac{1}{60}\tau^5\varepsilon^{-1}(\partial_x^2 V)(\partial_x V)^2 - \frac{1}{12}\tau^3\varepsilon(\partial_x^4 V)}_{\mathcal{O}(\varepsilon^4)} \end{aligned}$$

$$\begin{aligned}
 & \overbrace{\left[\tau^5 \varepsilon \left\{ \frac{4}{45} (\partial_x^2 V)^2 - \frac{1}{90} (\partial_x^3 V) (\partial_x V) \right\} \partial_x^2 + \frac{1}{45} \tau^5 \varepsilon^{-3} (\partial_x^4 V) \partial_x^4 \right]}^{\mathcal{O}(\varepsilon^4)} \\
 & + \overbrace{\left[\tau^5 \varepsilon \left\{ \frac{1}{6} (\partial_x^3 V) (\partial_x^2 V) - \frac{1}{90} (\partial_x^4 V) (\partial_x V) \right\} \partial_x \right]}^{\mathcal{O}(\varepsilon^5)} \\
 & + \overbrace{\left[\frac{2}{45} \tau^5 \varepsilon^{-3} (\partial_x^5 V) \partial_x^3 \right]}^{\mathcal{O}(\varepsilon^5)} + \mathcal{O}(\varepsilon^6). \tag{3.2}
 \end{aligned}$$

Unfortunately, (3.2) contains terms of order $\mathcal{O}(\varepsilon^3)$ and $\mathcal{O}(\varepsilon^5)$, both of which are due to the presence of odd powers of ∂_x . This presence is worrisome for an important reason, namely *stability*. Both ∂_x^2 and multiplication by V are Hermitian operators; therefore, $\tau(\varepsilon \partial_x^2 + \varepsilon^{-1} V)$ is a skew-Hermitian operator: its exponential is thus unitary. This survives under eventual discretisation because any reasonable approximation of ∂_x^2 preserves the Hermitian structure. However, ∂_x (and, in general, odd powers of ∂_x) is a skew-Hermitian operator; hence, $i\partial_x$ is Hermitian, as are its reasonable approximations. Therefore, the introduction of odd powers of ∂_x is fraught with loss of unitarity and stability. An extra ingredient is required in our algorithm!

3.2 Intermezzo: Getting Even

Let y be a C^1 function. The starting point for our current construction is the operatorial identity

$$y(x) \partial_x = -\frac{1}{2} \left[\int_{x_0}^x y(\xi) \, d\xi \right] \partial_x^2 - \frac{1}{2} \partial_x y(x) + \frac{1}{2} \partial_x^2 \left[\int_{x_0}^x y(\xi) \, d\xi \cdot \right], \tag{3.3}$$

where x_0 is arbitrary: its direct proof is trivial. Note that, whilst we have ∂_x on the left, the right-hand side features ∂_x^0 and ∂_x^2 , both even powers of the differentiation operator. Since in principle we might be interested in expanding beyond $\mathcal{O}(\varepsilon^{7/2})$ or employing different values of ρ and σ , we wish to cater not just for ∂_x but for all its odd powers. The challenge is thus to generalise (3.3) and express $y(x) \partial_x^{2s+1}$, $s \in \mathbb{Z}_+$, solely by means of even derivatives.

Theorem 1 *Let $s \in \mathbb{Z}_+$, define the real sequence $\{\beta_k\}_{k \geq 0}$ by*

$$\sum_{k=0}^{\infty} \frac{(-1)^k \beta_k}{(2k+1)!} T^k = \frac{1}{T} \left(1 - \frac{T^{1/2}}{\sinh T^{1/2}} \right)$$

and set

$$Q_k(x) = (-1)^{s-k+1} \beta_{s-k} \binom{2s+1}{2k} \partial_x^{2s-2k+1} y(x), \quad k = 0, 1, \dots, s, \quad (3.4)$$

$$Q_{s+1}(x) = \frac{1}{2s+2} \int_{x_0}^x y(\xi) \, d\xi, \quad (3.5)$$

$$P_k(x) = - \sum_{\ell=k}^{s+1} \binom{2\ell}{2k} \partial_x^{2\ell-2k} Q_\ell(x), \quad k = 1, 2, \dots, s+1. \quad (3.6)$$

Then

$$y(x) \partial_x^{2s+1} = \sum_{k=0}^{s+1} P_k(x) \partial_x^{2k} + \sum_{k=0}^{s+1} \partial_x^{2k} [Q_k(x) \cdot]. \quad (3.7)$$

Proof We act on the second sum on the right-hand side of (3.7) with the Leibnitz rule, whereby

$$\begin{aligned} y \partial_x^{2s+1} &= \sum_{k=1}^{s+1} P_k \partial_x^{2k} + \sum_{\ell=0}^{s+1} \sum_{k=0}^{2\ell} \binom{2\ell}{k} \left(\partial_x^{2\ell-k} Q_\ell \right) \partial_x^k \\ &= \sum_{k=1}^{s+1} P_k \partial_x^{2k} + \sum_{k=0}^{s+1} \left[\sum_{\ell=k}^{s+1} \binom{2\ell}{2k} \left(\partial_x^{2(\ell-k)} Q_\ell \right) \right] \partial_x^{2k} \\ &\quad + \sum_{k=0}^s \left[\sum_{\ell=k+1}^{s+1} \binom{2\ell}{2k+1} \left(\partial_x^{2(\ell-k)-1} Q_\ell \right) \right] \partial_x^{2k+1}. \end{aligned}$$

Equating powers of ∂_x on both sides, we obtain (3.5), (3.6) and the equations

$$\sum_{\ell=k+1}^{s+1} \binom{2\ell}{2k+1} \partial_x^{2(\ell-k)-1} Q_\ell = 0, \quad k = s-1, s-2, \dots, 0. \quad (3.8)$$

Our contention is that there exist coefficients $\{\beta_k\}_{k \geq 0}$ such that (3.4) is true. Indeed, substituting (3.4) into (3.8) yields, after simple algebra, the triangular linear system

$$\sum_{\ell=k+1}^s (-1)^{s-\ell} \binom{2s-2k}{2s+1-2\ell} \beta_{s-\ell} = \frac{1}{2s-2k+1}, \quad k = 0, 1, \dots, s-1.$$

We deduce that

$$\sum_{\ell=0}^{k-1} (-1)^\ell \binom{2k}{2\ell+1} \beta_\ell = \frac{1}{2k+1}, \quad k \in \mathbb{N}.$$

Finally, we multiply the last equation by $T^{k-1}/(2k)!$ and sum up for $k \in \mathbb{N}$. On the left-hand side we have

$$\begin{aligned} \sum_{k=1}^{\infty} \frac{1}{(2k)!} \sum_{\ell=0}^{k-1} (-1)^\ell \binom{2k}{2\ell+1} \beta_\ell T^{k-1} &= \sum_{\ell=0}^{\infty} \frac{(-1)^\ell \beta_\ell}{(2\ell+1)!} \sum_{k=\ell+1}^{\infty} \frac{T^{k-1}}{(2k-2\ell-1)!} \\ &= \sum_{\ell=0}^{\infty} \frac{(-1)^\ell \beta_\ell}{(2\ell+1)!} T^\ell \sum_{k=0}^{\infty} \frac{T^k}{(2k+1)!} \\ &= \frac{\sinh T^{1/2}}{T^{1/2}} \sum_{\ell=0}^{\infty} \frac{(-1)^\ell \beta_\ell}{(2\ell+1)!} T^\ell, \end{aligned}$$

whilst on the right we obtain

$$\sum_{k=1}^{\infty} \frac{T^{k-1}}{(2k+1)!} = \frac{1}{T} \left(\frac{\sinh T^{1/2}}{T^{1/2}} - 1 \right).$$

This confirms (3.4) and completes the proof. □

The first few values are $\beta_0 = \frac{1}{6}$, $\beta_1 = \frac{7}{60}$, $\beta_2 = \frac{31}{126}$, $\beta_3 = \frac{127}{120}$, $\beta_4 = \frac{511}{66}$, $\beta_5 = \frac{1414,477}{16,380}$ and $\beta_6 = \frac{8,191}{6}$. Since

$$\frac{te^{xt}}{(e^t - 1)} = \sum_{k=0}^{\infty} \frac{B_k(x)}{k!} t^k,$$

where B_k is the k th Bernoulli polynomial, it is easy to confirm that

$$\beta_k = \frac{(-1)^{k+1} 2^{2k+1} B_{2k+2}(\frac{1}{2})}{k+1}, \quad k \in \mathbb{Z}_+.$$

Practically, just

$$\begin{aligned} y \partial_x &= -\frac{1}{2} \left[\int_0^x y(\xi) d\xi \right] \partial_x^2 - \frac{1}{2} \partial_x y + \frac{1}{2} \partial_x^2 \left[\int_0^x y(\xi) d\xi \cdot \right], \\ y \partial_x^3 &= -(\partial_x y) \partial_x^2 - \frac{1}{4} \left[\int_0^x y(\xi) d\xi \right] \partial_x^4 + \frac{1}{4} \partial_x^3 y - \frac{1}{2} \partial_x^2 [(\partial_x y) \cdot] + \frac{1}{4} \partial_x^4 \left[\int_0^x y(\xi) d\xi \cdot \right], \\ y \partial_x^5 &= \frac{4}{3} (\partial_x^3 y) \partial_x^2 - \frac{5}{3} (\partial_x y) \partial_x^4 - \frac{1}{6} \left[\int_0^x y(\xi) d\xi \right] \partial_x^6 - \frac{1}{2} \partial_x^5 y + \frac{7}{6} \partial_x^2 [(\partial_x^3 y) \cdot] \\ &\quad - \frac{5}{6} \partial_x^4 [(\partial_x y) \cdot] + \frac{1}{6} \partial_x^6 \left[\int_0^x y(\xi) d\xi \cdot \right] \end{aligned}$$

are ever likely to be needed in practical computation.

Table 2 Symmetric Zassenhaus splitting of the first kind in even-order derivatives

Symmetric Zassenhaus algorithm
$s := 0; \quad \mathcal{W}^{[0]} := \tau(\varepsilon\partial_x^2 + \varepsilon^{-1}V); \quad W^{[0]} := \tau\varepsilon^{-1}V$
do
$s := s + 1$
compute $\mathcal{W}^{[s]} := \text{sBCH}(-W^{[s-1]}, \mathcal{W}^{[s-1]})$
rewrite $\mathcal{W}^{[s]}$ in even derivatives, cf. (3.7)
expand result in powers of ε
define $W^{[s]} := \mathcal{O}(\varepsilon^{2s-2})$, s.t. $W^{[s]} - \mathcal{W}^{[s]} = \mathcal{O}(\varepsilon^{2s})$
while $s < \text{desired order } s_{\max}$
Resulting method:
$e^{\mathcal{W}^{[0]}} = e^{W^{[0]}/2} e^{W^{[1]}/2} \dots e^{W^{[s_{\max}]} / 2} \dots e^{W^{[1]}/2} e^{W^{[0]}/2}$ $+ \mathcal{O}(\varepsilon^{2s_{\max}+2})$

3.3 Asymptotic Splitting

All necessary tools are now available, and in this subsection we will illustrate how to compute the splitting (1.7) with the algorithm in Table 2.

Using (3.3) and its generalisations to replace all the occurrences of ∂_x and ∂_x^3 in (3.2), we express $\mathcal{W}^{[1]}$ in the form

$$\begin{aligned}
 \mathcal{W}^{[1]} = & \underbrace{\tau\varepsilon\partial_x^2}_{\mathcal{O}(\varepsilon^0)} + \underbrace{\frac{1}{12}\tau^3\varepsilon^{-1}(\partial_x V)^2 - \frac{1}{6}\tau^3\varepsilon \left\{ (\partial_x^2 V)\partial_x^2 + \partial_x^2 \left[(\partial_x^2 V) \cdot \right] \right\}}_{\mathcal{O}(\varepsilon^2)} \\
 & + \underbrace{\frac{1}{60}\tau^5\varepsilon^{-1}(\partial_x^2 V)(\partial_x V)^2 + \frac{1}{12}\tau^3\varepsilon(\partial_x^4 V)}_{\mathcal{O}(\varepsilon^4)} \\
 & + \underbrace{\frac{1}{180}\tau^5\varepsilon \left\{ 8(\partial_x^2 V)^2\partial_x^2 + 8\partial_x^2 \left[(\partial_x^2 V)^2 \cdot \right] - (\partial_x^3 V)(\partial_x V)\partial_x^2 - \partial_x^2 \left[(\partial_x^3 V)(\partial_x V) \cdot \right] \right\}}_{\mathcal{O}(\varepsilon^4)} \\
 & + \underbrace{\frac{1}{90}\tau^5\varepsilon^{-3} \left\{ (\partial_x^4 V)\partial_x^4 + \partial_x^4 \left[(\partial_x^4 V) \cdot \right] \right\}}_{\mathcal{O}(\varepsilon^4)} + \mathcal{O}(\varepsilon^6).
 \end{aligned}$$

Recall that we started the algorithm with

$$\mathcal{R}_0 = \frac{1}{2}W^{[0]} = \frac{1}{2}\tau\varepsilon^{-1}V,$$

and to progress to the second stage, we choose to eliminate the lowest ε -order term,

$$\mathcal{R}_1 = \frac{1}{2}W^{[1]} = \frac{1}{2}\tau\varepsilon\partial_x^2,$$

from $\mathcal{W}^{[1]}$.

Although the new $W^{[1]}$ and $\mathcal{W}^{[1]}$ are more complicated, the computations are now much simpler. The main reason is that the ε -order behaves under commutation like

$$\left[\tau^{i_1} \varepsilon^{-j_1} f(x) \partial_x^{k_1}, \tau^{i_2} \varepsilon^{-j_2} g(x) \partial_x^{k_2} \right] = \mathcal{O}\left(\tau^{i_1+i_2} \varepsilon^{-(j_1+j_2)} \partial_x^{k_1+k_2-1} \right),$$

and thus, the order increases under very general assumptions. The first commutators then become

$$\left[W^{[1]}, \mathcal{W}^{[1]} \right] = \mathcal{O}\left(\varepsilon^3\right) \quad \text{and} \quad \left[\left[\mathcal{W}^{[1]}, W^{[1]} \right], W^{[1]} \right], \left[\left[\mathcal{W}^{[1]}, W^{[1]} \right], \mathcal{W}^{[1]} \right] = \mathcal{O}\left(\varepsilon^4\right).$$

Continuing the argument we find that all grade four and five commutators scale as $\mathcal{O}(\varepsilon^5)$ and $\mathcal{O}(\varepsilon^6)$ respectively. Subsequent commutators are even smaller, and we obtain

$$\begin{aligned} \mathcal{W}^{[2]} &= \text{sBCH}\left(-W^{[1]}, \mathcal{W}^{[1]}\right) \\ &= -W^{[1]} + \mathcal{W}^{[1]} - \frac{1}{24} \left[\left[\mathcal{W}^{[1]}, W^{[1]} \right], W^{[1]} \right] - \frac{1}{12} \left[\left[\mathcal{W}^{[1]}, W^{[1]} \right], \mathcal{W}^{[1]} \right] + \mathcal{O}\left(\varepsilon^6\right) \\ &= \underbrace{\frac{1}{12} \tau^3 \varepsilon^{-1} (\partial_x V)^2 - \frac{1}{6} \tau^3 \varepsilon \left\{ (\partial_x^2 V) \partial_x^2 + \partial_x^2 \left[(\partial_x^2 V) \cdot \right] \right\}}_{\mathcal{O}(\varepsilon^4)} \\ &\quad + \underbrace{\frac{1}{60} \tau^5 \varepsilon^{-1} (\partial_x^2 V) (\partial_x V)^2 + \frac{1}{12} \tau^3 \varepsilon (\partial_x^4 V)}_{\mathcal{O}(\varepsilon^4)} \\ &\quad + \underbrace{\frac{1}{120} \tau^5 \varepsilon \left\{ 7(\partial_x^2 V)^2 \partial_x^2 + 7\partial_x^2 \left[(\partial_x^2 V)^2 \cdot \right] + (\partial_x^3 V) (\partial_x V) \partial_x^2 + \partial_x^2 \left[(\partial_x^3 V) (\partial_x V) \cdot \right] \right\}}_{\mathcal{O}(\varepsilon^4)} \\ &\quad - \underbrace{\frac{1}{60} \tau^5 \varepsilon^{-3} \left\{ (\partial_x^4 V) \partial_x^4 + \partial_x^4 \left[(\partial_x^4 V) \cdot \right] \right\}}_{\mathcal{O}(\varepsilon^4)} + \mathcal{O}\left(\varepsilon^6\right). \end{aligned}$$

In the next iteration, we pull out the $\mathcal{O}(\varepsilon^2)$ term,

$$2\mathcal{R}_2 = W^{[2]} = \frac{1}{12} \tau^3 \varepsilon^{-1} (\partial_x V)^2 - \frac{1}{6} \tau^3 \varepsilon \left\{ (\partial_x^2 V) \partial_x^2 + \partial_x^2 \left[(\partial_x^2 V) \cdot \right] \right\},$$

and need to compute $\mathcal{W}^{[3]}$. Because of $[W^{[2]}, \mathcal{W}^{[2]}] = \mathcal{O}(\varepsilon^7)$, again, commutators can be disregarded to obtain $\mathcal{T}_3 = \mathcal{W}^{[3]} = \mathcal{O}(\varepsilon^4)$; the asymptotic splitting is therefore

$$S_{(1,1),2}^{[1]} = e^{\mathcal{R}_0} e^{\mathcal{R}_1} e^{\mathcal{R}_2} e^{\mathcal{T}_3} e^{\mathcal{R}_2} e^{\mathcal{R}_1} e^{\mathcal{R}_0}, \tag{3.9}$$

where

$$\mathcal{R}_0 = \frac{1}{2} \tau \varepsilon^{-1} V = \mathcal{O}\left(\varepsilon^0\right), \tag{3.10}$$

$$\begin{aligned} \mathcal{R}_1 &= \frac{1}{2} \tau \varepsilon \partial_x^2 = \mathcal{O}(\varepsilon^0), \\ \mathcal{R}_2 &= \frac{1}{24} \tau^3 \varepsilon^{-1} (\partial_x V)^2 - \frac{1}{12} \tau^3 \varepsilon \left\{ (\partial_x^2 V) \partial_x^2 + \partial_x^2 [(\partial_x^2 V) \cdot] \right\} = \mathcal{O}(\varepsilon^2), \\ \mathcal{T}_3 &= \frac{1}{60} \tau^5 \varepsilon^{-1} (\partial_x^2 V) (\partial_x V)^2 + \frac{1}{12} \tau^3 \varepsilon (\partial_x^4 V) \\ &\quad + \frac{1}{120} \tau^5 \varepsilon \left\{ 7(\partial_x^2 V)^2 \partial_x^2 + 7\partial_x^2 [(\partial_x^2 V)^2 \cdot] \right\} + (\partial_x^3 V) (\partial_x V) \partial_x^2 \\ &\quad + \partial_x^2 [(\partial_x^3 V) (\partial_x V) \cdot] \left. \right\} - \frac{1}{60} \tau^5 \varepsilon^{-3} \left\{ (\partial_x^4 V) \partial_x^4 + \partial_x^4 [(\partial_x^4 V) \cdot] \right\} = \mathcal{O}(\varepsilon^4). \end{aligned}$$

The notation $\mathcal{S}_{(1,1),2}^{[1]}$ is mostly self-explanatory: (1, 1) refers to the values of ρ and σ , whilst $s = 2$. The superscript $^{[1]}$ stands for an asymptotic splitting of the *first kind*: in Sect. 3.5 we consider an alternative splitting (with initial $W^{[0]}$ equalling $\tau \varepsilon \partial_x^2$), which we designate as an asymptotic splitting of the *second kind*.

Once we replace derivatives by differentiation matrices, the evaluation of a single time step $\mathbf{u}^{n+1} = \tilde{\mathcal{S}}_{(1,1),2}^{[1]} \mathbf{u}^n$ requires in principle seven exponentials. However, we note that, once we use nodal values in semidiscretisation, the discretised matrix $\tilde{\mathcal{R}}_0$ is diagonal and the computation of its exponential can be accomplished in $\mathcal{O}(M)$ operations.² The next discretised matrix, $\tilde{\mathcal{R}}_1$, is circulant and its exponentiation involves $\mathcal{O}(M \log M)$ operations. This is an important point because $\tilde{\mathcal{R}}_0$ and $\tilde{\mathcal{R}}_1$ are (spectrally) the largest matrices present. All other matrices are $\mathcal{O}(\varepsilon^2)$ or less, and, as will be clear in Sect. 5, their computation with Krylov subspace methods is very affordable.

3.4 Stability

The convergence of classical methods for initial-value partial differential equations is governed by the Lax equivalence theorem: convergence equals consistency plus stability [11]. Our method is clearly consistent, but the question is whether, once derivatives are replaced by differentiation matrices, the ensuing finite-dimensional operator is stable in the sense of Lax. Within our formalism this is equivalent to

$$\lim_{\varepsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \|(\tilde{\mathcal{S}}_{(1,1),2}^{[1]})^n\| < \infty, \tag{3.11}$$

where $\tilde{\mathcal{S}}_{(1,1),2}^{[1]}$ is the finite-dimensional discretisation of $\mathcal{S}_{(1,1),2}^{[1]}$. Here $\|\cdot\|$ is the standard Euclidean norm.

Condition (3.11) is clearly implied by $\tilde{\mathcal{S}}_{(1,1),2}^{[1]}$ being a unitary matrix for all (sufficiently small) $\varepsilon > 0$, in other words, by the discretisation method’s *unitarity*. This has the added virtue of the discretisation method mimicking the unitarity of the infinite-dimensional operator $\exp(it(\varepsilon \partial_x^2 + \varepsilon^{-1} V))$. (The latter follows because both $i\partial_x^2$ and multiplication by iV are skew-Hermitian.) Consequently, in that case we obtain a *geo-*

² Using a Fourier basis the cost is $\mathcal{O}(M \log M)$.

metric integrator in the sense of [4, 8, 14]. Once we use Krylov methods for approximating the exponentials, unitarity is lost, but the conservation of the ℓ_2 norm guarantees stability nevertheless.

Suppose that $\tilde{\mathcal{R}}_0, \tilde{\mathcal{R}}_1, \tilde{\mathcal{R}}_2$ and $\tilde{\mathcal{T}}_3$ are all skew-Hermitian matrices. Then, by (3.9), $\tilde{\mathcal{S}}_{(1,1),2}^{[1]}$ is unitary. But are they?

The discretisation of ∂_x^2 is the subject of Sect. 4. Here we preempt the discussion by identifying two options. Either we choose the unknowns as nodal values (e.g. by using finite differences or spectral collocation) or as Fourier coefficients (using a spectral method). In the first case, ∂_x^2 is approximated by a *symmetric circulant* \mathcal{K} and multiplication by V by a *diagonal matrix* \mathcal{D} . In the second case, all is reversed: ∂_x^2 is approximated by a diagonal matrix and multiplication by V by a symmetric circulant. In either case, $i\mathcal{K}, i\mathcal{D} \in \mathfrak{su}_M(\mathbb{C})$, the Lie algebra of $M \times M$ complex skew-Hermitian matrices. It follows at once that $\tilde{\mathcal{R}}_0, \tilde{\mathcal{R}}_1 \in \mathfrak{su}_M(\mathbb{C})$; consequently, $e^{\tilde{\mathcal{R}}_0}, e^{\tilde{\mathcal{R}}_1} \in U_M(\mathbb{C})$.³ However,

$$\tilde{\mathcal{R}}_2 = \frac{1}{24} \tau^3 \varepsilon^{-1} \mathcal{D}_{(\partial_x V)^2} - \frac{1}{12} \tau^3 \varepsilon (\mathcal{K} \mathcal{D}_{\partial_x^2} V + \mathcal{D}_{\partial_x^2} V \mathcal{K}),$$

where \mathcal{D}_f is the discretisation of a multiplication by f , may seem problematic: $i\mathcal{K}, i\mathcal{D} \in \mathfrak{su}_M(\mathbb{C})$ need not imply that $i\mathcal{K}\mathcal{D}, i\mathcal{D}\mathcal{K} \in \mathfrak{su}_M(\mathbb{C})$.⁴ Fortunately, it is trivial to verify that $i(\mathcal{K}\mathcal{D} + \mathcal{D}\mathcal{K}) \in \mathfrak{su}_M(\mathbb{C})$, and this proves that $\tilde{\mathcal{R}}_2 \in \mathfrak{su}_M(\mathbb{C})$. Examining carefully (3.10), we observe that so does $\tilde{\mathcal{T}}_3$. We deduce that $e^{\tilde{\mathcal{R}}_2}, e^{\tilde{\mathcal{T}}_3} \in U_M(\mathbb{C})$, and stability (3.11) follows.

The unitarity of $\tilde{\mathcal{S}}_{(1,1),2}^{[1]}$ is not accidental, and we do not need to repeat our analysis on a case-by-case basis for different values of ρ, σ and s or for the asymptotic splittings of the second kind from the next subsection.

Theorem 2 *Supposing that the splitting (1.7) has been derived by the symmetric Zassenhaus algorithm of Table 2, it is true that $W^{[i]} \in \mathfrak{su}_M(\mathbb{C})$ for all $i \geq 0$ and, thus, also $\tilde{\mathcal{R}}_0, \tilde{\mathcal{R}}_1, \dots, \tilde{\mathcal{R}}_s, \tilde{\mathcal{T}}_{s+1} \in \mathfrak{su}_M(\mathbb{C})$.*

Proof The algorithm starts from a skew-Hermitian operator $\mathcal{W}^{[0]}$ and, in each step, pulls out a term $W^{[j]}$ via the symmetric BCH formula (2.3). Assume that $W^{[j]}$ is skew-Hermitian; then so is $\mathcal{W}^{[j+1]}$ because skew-Hermiticity is preserved under commutation. What remains to be shown is that at each step, the lowest-order ε terms in $\mathcal{W}^{[j]}$ after the ‘odd to even’ substitution (3.3), namely $W^{[j]}$, are indeed skew-Hermitian. Recall that, by assumption, $\mathcal{W}^{[j]}$ is skew-Hermitian, and since the substitution is exact, it remains so after it has been applied. For this reason, it is clear that its summands are either skew-Hermitian or feature in skew-Hermitian pairs $i(\mathcal{K}^l \mathcal{D} + \mathcal{D} \mathcal{K}^l)$, where \mathcal{K}^k is a symmetric discretisation of ∂_x^{2k} . The algorithm groups terms with the same scaling, and since $\mathcal{D} \mathcal{K}^k = \mathcal{O}(\varepsilon^{-2k}) = \mathcal{K}^k \mathcal{D}$, the pair will not be split, and thus $W^{[j]} \in \mathfrak{su}_M(\mathbb{C})$. □

³ As before, a tilde denotes a discretisation.

⁴ All powers of τ are odd because of the palindromy of the symmetric BCH formula. Since $\tau = i\Delta t$, this means that they always contribute a multiple of $\pm i$.

$$\begin{aligned}
 & + \frac{1}{60} \tau^5 \varepsilon \left\{ \partial_x^2 \left[(\partial_x^2 V)^2 \cdot \right] + (\partial_x^2 V)^2 \partial_x^2 - 2 \partial_x^2 \left[(\partial_x^3 V)(\partial_x V) \cdot \right] - 2 (\partial_x^3 V)(\partial_x V) \partial_x^2 \right\} \\
 & + \frac{1}{240} \tau^5 \varepsilon^3 \left\{ (\partial_x^4 V) \partial_x^4 + \partial_x^4 \left[(\partial_x^4 V) \cdot \right] \right\}.
 \end{aligned}$$

Finally,

$$\mathcal{R}_2 = \frac{1}{2} \mathcal{W}_1^{[2]}, \quad \mathcal{T}_3 = \mathcal{W}_2^{[2]}.$$

The outcome is the splitting

$$\mathcal{S}_{(1,1),2}^{[2]} = e^{\mathcal{R}_0} e^{\mathcal{R}_1} e^{\mathcal{R}_2} e^{\mathcal{T}_3} e^{\mathcal{R}_2} e^{\mathcal{R}_1} e^{\mathcal{R}_0}, \tag{3.12}$$

where

$$\begin{aligned}
 \mathcal{R}_0 &= \frac{1}{2} \tau \varepsilon \partial_x^2, \\
 \mathcal{R}_1 &= \frac{1}{2} \tau \varepsilon^{-1} V, \\
 \mathcal{R}_2 &= -\frac{1}{12} \tau^3 \varepsilon^{-1} (\partial_x V)^2 + \frac{1}{24} \tau^3 \varepsilon \left\{ \partial_x^2 \left[(\partial_x^2 V) \cdot \right] + (\partial_x^2 V) \partial_x^2 \right\}, \\
 \mathcal{T}_3 &= -\frac{1}{24} \tau^3 \varepsilon (\partial_x^4 V) + \frac{7}{120} \tau^5 \varepsilon^{-1} (\partial_x^2 V)(\partial_x V)^2 \\
 & \quad + \frac{1}{60} \tau^5 \varepsilon \left\{ \partial_x^2 \left[(\partial_x^2 V)^2 \cdot \right] + (\partial_x^2 V)^2 \partial_x^2 - 2 \partial_x^2 \left[(\partial_x^3 V)(\partial_x V) \cdot \right] \right. \\
 & \quad \left. - 2 (\partial_x^3 V)(\partial_x V) \partial_x^2 \right\} + \frac{1}{240} \tau^5 \varepsilon^3 \left\{ (\partial_x^4 V) \partial_x^4 + \partial_x^4 \left[(\partial_x^4 V) \cdot \right] \right\}.
 \end{aligned}$$

As in the case of (3.9), we end up needing to compute seven exponentials. $\tilde{\mathcal{R}}_0$ is either a circulant (once we use nodal values) or a diagonal matrix (if we employ a Fourier expansion), whilst $\tilde{\mathcal{R}}_1$ is then either a diagonal matrix or a circulant – the opposite of $\tilde{\mathcal{R}}_0$. Therefore, the cost of computing $\tilde{\mathcal{R}}_i, i = 0, 1$, is $\mathcal{O}(M \log M)$ operations. This leaves $\tilde{\mathcal{R}}_2 = \mathcal{O}(\varepsilon^2)$ and $\tilde{\mathcal{T}}_3 = \mathcal{O}(\varepsilon^4)$, which need be computed with Krylov subspace methods. The small magnitude of both these matrices means that their exponentials can be computed in a ridiculously small number of Krylov iterations; cf. Sect. 5.

Note that the palindromic property allows us to further reduce the number of exponentials if no output at intermediate steps is required. This so-called First-Same-As-Last (FSAL) property effectively yields a method

$$\tilde{\mathcal{S}}(\alpha)_{(1,1),2}^{[2]} = e^{\tilde{\mathcal{R}}_1} e^{\tilde{\mathcal{R}}_2} e^{\tilde{\mathcal{T}}_3} e^{\tilde{\mathcal{R}}_2} e^{\tilde{\mathcal{R}}_1} e^{\alpha \tilde{\mathcal{R}}_0}, \tag{3.13}$$

where the first step must be calculated with $\alpha = 1$ and further steps with $\alpha = 2$. Whenever output is required, we apply $e^{\tilde{\mathcal{R}}_0}$ and initialise the method by letting $\alpha = 1$ for the next step. All in all, we only need to compute six exponentials in each step,

two of which are diagonal matrices, one is circulant and the remaining three can be approximated cheaply by Krylov methods.

4 Semidiscretisation

The asymptotic splittings (3.9) and (3.12) are expressed in operatorial terms: to render them into proper computational algorithms, we must replace ∂_x^2 with an appropriate differentiation matrix, acting on an M -dimensional space.

It is common in the numerical solution of the Schrödinger equation to use *spectral* discretisation [4, 13]. Thus, the unknowns are the Fourier coefficients of u , \mathcal{K} is a diagonal matrix, $\mathcal{K}_{j,j} = -\pi^2 j^2$, $|j| \leq (M-1)/2$ [note that indeed $\|\mathcal{K}\| = \mathcal{O}(M^2) = \mathcal{O}(\varepsilon^{-2})$] and the operator of a multiplication by f is discretised by a circulant \mathcal{D}_f composed of the Fourier coefficients of f . We deduce that, for any $\mathbf{v} \in \mathbb{C}^M$, the computation of $\mathcal{K}\mathbf{v}$ costs $\mathcal{O}(M)$ operations, whilst the price tag of $\mathcal{D}_f\mathbf{v}$, computed with FFT, is $\mathcal{O}(M \log M)$. The main appeal of spectral methods is that they exhibit *spectral convergence*: for sufficiently large M the error decays faster than $M^{-\alpha} = \mathcal{O}(\varepsilon^\alpha)$ for any $\alpha > 0$. In classical terms, the method is of an infinite order.

Alternative methods of discretisation are based on *nodal values*. In all such methods, multiplication by a function f discretises into a diagonal matrix. Since it is compelling in the presence of period boundary conditions to use equispaced points, the unknowns are thus $u_m \approx u(m/(N + \frac{1}{2}))$, $|m| \leq N$, where $M = 2N + 1$. Two methods that fall into this category are finite differences and spectral collocation.

The idea in spectral collocation [9] is to interpolate the solution at the nodal values using a trigonometric polynomial. Since a trigonometric interpolation can be written as a convolution with the values of the scaled *Dirichlet kernel*

$$D_N(x) = \frac{\sin((N + \frac{1}{2})\pi x)}{(2N + 1) \sin(\frac{1}{2}\pi x)}$$

– in other words $\sum_{\ell=-N}^N D_N(x - \ell/(N + \frac{1}{2}))u_\ell$ is an N th-order trigonometric polynomial which equals u_m at $m/(N + \frac{1}{2})$ – the differentiation matrix given by $\mathcal{K}_{j,\ell} = D_N''((j - \ell)/(N + \frac{1}{2}))$ is a circulant.

Like spectral methods, spectral collocation exhibits spectral convergence. The two are, in fact, equivalent in the context of periodic boundaries and equispaced points – they are just a Fourier transform away from each other. For this reason it is not uncommon in the literature to refer to both by the same name – usually spectral collocation – and the choice between them is mostly a matter of convenience, having little influence on the efficiency. Here we end up favouring the nodal representation of ‘spectral collocation’ over the Fourier representation of ‘spectral methods’.

At this point, a word about finite differences is also in order. At first glance, finite differences might be considered a viable alternative to spectral collocation. An order five central difference method for the second derivative, for instance, incurs an error of $\mathcal{O}((\Delta x)^6) = \mathcal{O}(\varepsilon^6)$, which seems acceptable considering that we already have an error of $\mathcal{O}(\varepsilon^6)$ in the splitting. A closer look at the finite difference error term,

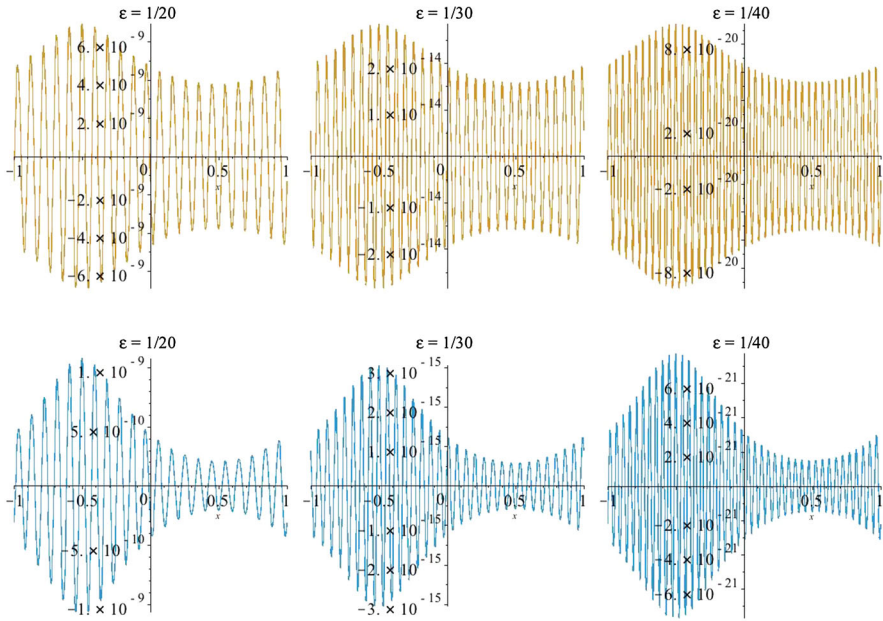


Fig. 1 Error in approximating u'' committed by Fourier method (top row) and spectral collocation (bottom row) for function $u(x) = 1/(2 + \sin \pi x)$ and $M = \epsilon^{-1}$

however, reveals a factor of $u^{(8)}$, which is far from a constant. In fact, it scales as $\mathcal{O}(\epsilon^{-8})$, upsetting the entire balance and bringing the overall error to $\mathcal{O}(\epsilon^{-2})$ – as huge as the norm of the Laplacian! It is possible to coax and cajole finite differences to work with $\rho > 1$, but the $\mathcal{O}(\epsilon^{-1} \log \epsilon^{-1})$ cost of spectral collocation cannot be improved upon in this way, and spectral collocation remains the method of choice.

The error for $1/(2 + \sin \pi x)$ and $e^{\cos \pi x}$ is displayed in Figs. 1 and 2 respectively. Although the spectacular performance of spectrally convergent methods is hardly surprising, it is amazing nonetheless. The reason spectral convergence for $e^{\cos \pi x}$ is so fast – super-exponential, compared to exponential convergence for $1/(2 + \sin \pi x)$ – is because the first function is entire, whereas the second has a polar singularity at $i(\sqrt{3} - 2)$.

5 Computation of Exponentials

Considering splittings (3.9) and (3.12), each step forward in time calls for the computation of

$$\mathbf{u}^{n+1} = e^{\tilde{\mathcal{R}}_0} e^{\tilde{\mathcal{R}}_1} e^{\tilde{\mathcal{R}}_2} e^{\tilde{\mathcal{T}}_3} e^{\tilde{\mathcal{R}}_2} e^{\tilde{\mathcal{R}}_1} e^{\tilde{\mathcal{R}}_0} \mathbf{u}^n, \tag{5.1}$$

where \mathbf{u}^n is the initial value at t_n , say, whilst \mathbf{u}^{n+1} approximates $u(\cdot, t_{n+1})$, where $t_{n+1} = t_n + \Delta t_n$. The matrices $\tilde{\mathcal{R}}_k$ and $\tilde{\mathcal{T}}_3$ depend on Δt_n . We recall that, using finite differences or spectral collocation, \mathbf{u}^n is comprised of equally distributed function values, and in splitting of the second kind, $\tilde{\mathcal{R}}_0$ is a Toeplitz circulant whilst $\tilde{\mathcal{R}}_1$ is

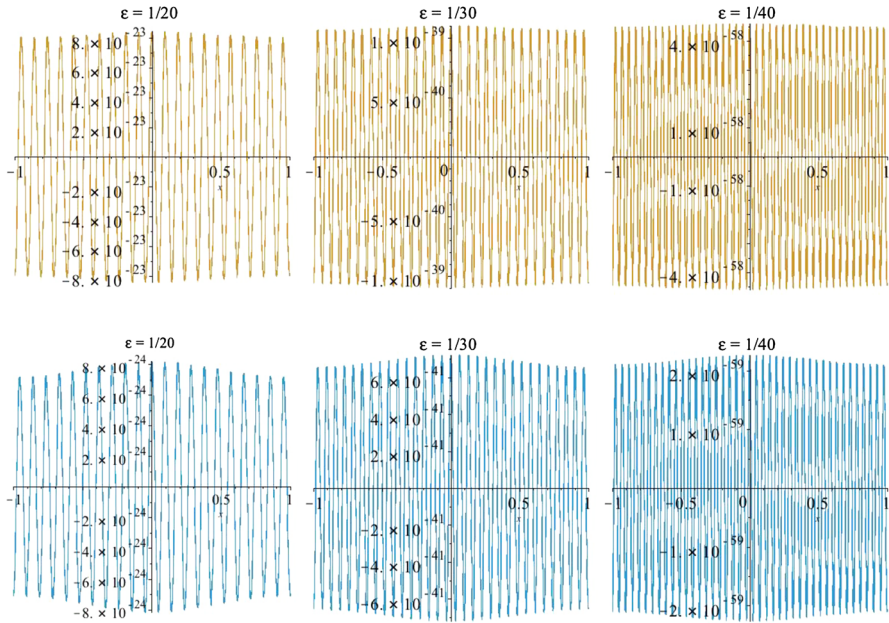


Fig. 2 Error in approximating u'' committed by Fourier method (top row) and spectral collocation (bottom row) for function $u(x) = e^{\cos \pi x}$ and $M = \varepsilon^{-1}$

diagonal. However, once we use a spectral method, the entries of u^n consist of Fourier coefficients, $\tilde{\mathcal{R}}_0$ is diagonal and $\tilde{\mathcal{R}}_1$ is a circulant. The roles are reversed in splitting of the first kind. One way or the other, we need to calculate [or approximate up to $\mathcal{O}(\varepsilon^6)$] the vector $e^{\mathcal{S}} v$ for $v \in \mathbb{C}^M$ and three types of $M \times M$ skew-Hermitian matrices \mathcal{S} : (a) diagonal, (b) Toeplitz circulant and (c) neither, yet small: $\tilde{\mathcal{R}}_2 = \mathcal{O}(\varepsilon^2)$ and $\tilde{\mathcal{T}}_3 = \mathcal{O}(\varepsilon^4)$. Note that we must keep in mind three prerogatives: not just the error of $\mathcal{O}(\varepsilon^6)$ and low cost but also the conservation of the ℓ_2 norm.

Cases (a) and (b) are straightforward. The exponential of a diagonal matrix is itself diagonal and can be computed in $\mathcal{O}(M) = \mathcal{O}(\varepsilon^{-1})$ operations, whilst $e^{\mathcal{S}} v$ for a circulant \mathcal{S} can be calculated by two FFTs, at the cost of $\mathcal{O}(M \log M) = \mathcal{O}(\varepsilon^{-1} \log \varepsilon^{-1})$ operations. Since both calculations are exact (up to machine accuracy), unitarity is maintained. Finally, to deal with case (c), we use a *Krylov subspace method*. Such methods have undergone many enhancements since the pioneering work of [19]: in the current paper we adopt the approach in [5].

Given an $M \times M$ matrix \mathcal{A} and $v \in \mathbb{C}^M$, the m th Krylov subspace is

$$K_m(\mathcal{A}, v) = \text{span} \{v, \mathcal{A}v, \mathcal{A}^2v, \dots, \mathcal{A}^{m-1}v\}, \quad m \in \mathbb{N}.$$

It is well known that $\dim K_{m-1}(\mathcal{A}, v) \leq \dim K_m(\mathcal{A}, v) \leq \min\{m, M\}$, and we refer the reader to [6] for other properties of Krylov subspaces. The main idea is to approximate

$$e^{\mathcal{A}}\mathbf{v} \approx \mathcal{V}_m e^{\mathcal{H}_m} \mathcal{V}_m^* \mathbf{v}, \tag{5.2}$$

where \mathcal{V}_m and \mathcal{H}_m are $M \times m$ and $m \times m$ respectively and $m \ll M$. In addition, the columns of \mathcal{V}_m are orthonormal vectors, which form a basis of $\mathbf{K}_m(\mathcal{A}, \mathbf{v})$, whilst \mathcal{H}_m is upper Hessenberg.

The matrices \mathcal{V}_m and \mathcal{H}_m are generated by the *Arnoldi process* [6,5]. Note that,

The Arnoldi process

```

v1 = v/||v||2
for j = 1, ..., m - 1 do
    t = Avj
    for i = 1, ..., j do
        hi,j = vi*t,    t = t - hi,jvi
    end for
    hj+1,j = ||t||2;    vj+1 = t/hj+1,j
end for
    
```

once $\mathcal{A} \in \mathfrak{su}_M(\mathbb{C})$, it follows that $\mathcal{H}_m \in \mathfrak{su}_m(\mathbb{C})$. Therefore, because the columns of \mathcal{V}_m are orthonormal, the ℓ_2 norm is conserved. Moreover, since $\mathcal{V}_m^* \mathbf{v} = \|\mathbf{v}\|_2 \mathbf{e}_1$, where $\mathbf{e}_1 \in \mathbb{C}^m$ is the first unit vector, it follows that $e^{\mathcal{H}_m} \mathcal{V}_m^* \mathbf{v}$ is merely the first column of $e^{\mathcal{H}_m}$, scaled by $\|\mathbf{v}\|_2$. To compute approximation (5.2), we thus need to evaluate a *small* exponential and calculate a single matrix–vector product.

The computational cost is dominated by the cost of the iterations, each involving a matrix–vector product of the form $\mathcal{A}\mathbf{v}_j$. In any Zassenhaus splitting, matrices which require exponentiation by Krylov methods will involve a multiplication of a diagonal and a circulant. A few FFTs are therefore unavoidable in each iteration, bringing the overall cost to $\mathcal{O}(mM \log M)$ operations.

The question of an appropriate value of m is answered by the inequality

$$\|e^{\mathcal{A}}\mathbf{v} - \mathcal{V}_m e^{\mathcal{H}_m} \mathcal{V}_m^* \mathbf{v}\|_2 \leq 12e^{-\rho^2/(4m)} \left(\frac{e\rho}{2m}\right)^m, \quad m \geq \rho, \tag{5.3}$$

where $\rho = \rho(\mathcal{A})$ is the spectral radius of \mathcal{A} [10]. We know that $\tilde{\mathcal{R}}_2 = \mathcal{O}(\varepsilon^2)$ and assume, with a very minor loss of generality, that $\rho(\tilde{\mathcal{R}}_2) \leq c\varepsilon^2$ for some $c > 0$. We thus deduce from (5.3) that

$$\|e^{\tilde{\mathcal{R}}_2}\mathbf{v} - \mathcal{V}_m e^{\mathcal{H}_m} \mathcal{V}_m^* \mathbf{v}\|_2 \leq 12 \left(\frac{ec}{2m}\right)^m \varepsilon^{2m}, \quad m \geq \rho,$$

and $m = 3$ is sufficient to reduce the error to $\mathcal{O}(\varepsilon^6)$, in line with the error of our symmetric Zassenhaus algorithm. This is true provided that $\rho \leq 3$, i.e. $\varepsilon \leq \sqrt{3}/c$ – since we expect $\varepsilon > 0$ to be very small, this is not much in a way of restriction. Likewise, $\tilde{\mathcal{T}}_3 = \mathcal{O}(\varepsilon^4)$ and the inequality $\rho(\tilde{\mathcal{T}}_3) \leq \tilde{c}\varepsilon^4$ implies that

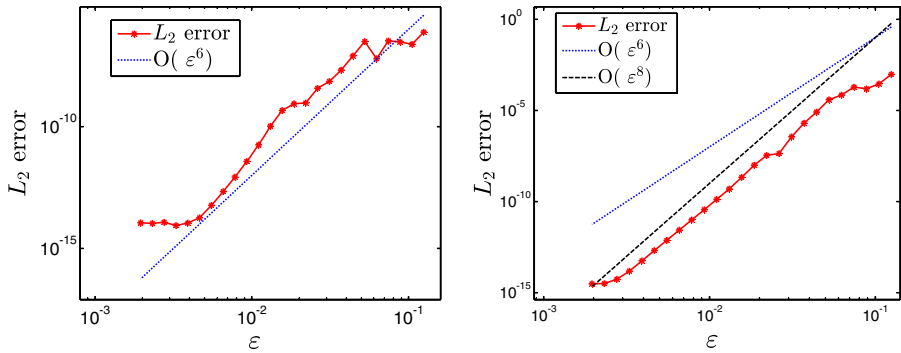


Fig. 3 Error, compared to required order $\mathcal{O}(\varepsilon^6)$, in computing $e^{\tilde{\mathcal{R}}_2} \mathbf{v}$ (left) and $e^{\tilde{\mathcal{T}}_3} \mathbf{v}$ (right)

$$\|e^{\tilde{\mathcal{T}}_3} \mathbf{v} - \mathcal{V}_m e^{\mathcal{H}_m} \mathcal{V}_m^* \mathbf{v}\|_2 \leq 12 \left(\frac{e\tilde{c}}{2m}\right)^m \varepsilon^{4m}, \quad m \geq \rho,$$

and for $\varepsilon \leq (2/\tilde{c})^{1/4}$ we need just $m = 2$. Altogether, we deduce that the computation [consistent with the error of $\mathcal{O}(\varepsilon^6)$] of $e^{\tilde{\mathcal{R}}_2} \mathbf{v}$ (twice) and $e^{\tilde{\mathcal{T}}_3} \mathbf{v}$ in each step (5.1) cost just $\mathcal{O}(M \log M)$ operations.

Figure 3 presents the L_2 error committed in approximating the exponentials $e^{\tilde{\mathcal{R}}_2} \mathbf{v}$ and $e^{\tilde{\mathcal{T}}_3} \mathbf{v}$, where we take $\phi(x) = e^{-20 \sin^2(\pi x/2)}$ as the interaction potential V and $\psi(x) = e^{-4(\sin^2(5\pi x/2) + \sin^2(\pi x/2))}$ as the wave function u , both discretised as nodal values at $M = 2N + 1$ grid points with $N = \lfloor \varepsilon^{-1} \rfloor$. Although we have used just $m = 3$ for $e^{\tilde{\mathcal{R}}_2} \mathbf{v}$ [i.e. approximated the $(2N + 1) \times (2N + 1)$ exponential by an 3×3 one] the error is truly minuscule. Moreover, consistently with our theory (but not with conventional numerical intuition), it decreases with ε . Indeed, the sort of accuracies we obtain for small values of ε are well in excess of what is required in realistic numerical computations. In the case of $e^{\tilde{\mathcal{T}}_3} \mathbf{v}$, we approximate with just a 2×2 exponential! Again, everything is consistent with our analysis.

The slope of the error bound is steeper than ε^6 in the second figure, and this should cause no surprise. The error for $e^{\tilde{\mathcal{T}}_3} \mathbf{v}$ decays as $\mathcal{O}(\varepsilon^8)$, much faster than required.

6 Asymptotic Splittings with Different Values of ρ and σ

The Zassenhaus splitting procedure used for deriving the $S_{(1,1),2}^{[2]}$ splitting (3.12) is hardly tied to the choice of $(\rho, \sigma) = (1, 1)$ and works just as well for any other pair (except that WKB considerations restrict us to $\rho \geq 1$). With small modifications in the working and with little difficulty, for instance, we are able to arrive at variants of this splitting with $\rho = 1$ and $\sigma = \frac{1}{2}$ or $\frac{1}{4}$. The common form of the splittings we present is

$$S_{(1,\sigma),2}^{[2]} = e^{\mathcal{R}_0} e^{\mathcal{R}_1} e^{\mathcal{R}_2} e^{\mathcal{T}_3} e^{\mathcal{R}_2} e^{\mathcal{R}_1} e^{\mathcal{R}_0}, \tag{6.1}$$

with

$$\begin{aligned} \mathcal{R}_0 &= \frac{1}{2} \tau \varepsilon \partial_x^2, \\ \mathcal{R}_1 &= \frac{1}{2} \tau \varepsilon^{-1} V, \\ \mathcal{R}_2 &= \frac{1}{24} \tau^3 \varepsilon \left\{ \partial_x^2 \left[(\partial_x^2 V) \cdot \right] + (\partial_x^2 V) \partial_x^2 \right\} - \frac{1}{12} \tau^3 \varepsilon^{-1} (\partial_x V)^2 \end{aligned}$$

being common to the three cases $\sigma = 1, \frac{1}{2}, \frac{1}{4}$. The components of the final term,

$$\begin{aligned} \mathcal{T}_3 &= \frac{1}{240} \tau^5 \varepsilon^3 \left\{ \partial_x^4 \left[(\partial_x^4 V) \cdot \right] + (\partial_x^4 V) \partial_x^4 \right\} - \frac{1}{24} \tau^3 \varepsilon (\partial_x^4 V) \\ &\quad + \frac{1}{60} \tau^5 \varepsilon \left\{ \partial_x^2 \left[(\partial_x^2 V)^2 \cdot \right] + (\partial_x^2 V)^2 \partial_x^2 \right\} \\ &\quad - \frac{1}{30} \tau^5 \varepsilon \left\{ \partial_x^2 \left[(\partial_x V) (\partial_x^3 V) \cdot \right] + (\partial_x V) (\partial_x^3 V) \partial_x^2 \right\} \\ &\quad + \frac{7}{120} \tau^5 \varepsilon^{-1} (\partial_x V)^2 (\partial_x^2 V), \end{aligned}$$

are also identical, apart from the factor $-\frac{1}{24} \tau^3 \varepsilon (\partial_x^4 V)$, which features in the case of $\sigma = 1$ and is missing for $\sigma = \frac{1}{2}$ and $\sigma = \frac{1}{4}$. The error in these splittings is $\mathcal{O}(\varepsilon^{7\sigma-1})$.

6.1 Implementation Issues

As before, the first two terms $\tilde{\mathcal{R}}_0$ and $\tilde{\mathcal{R}}_1$, which constitute a circulant and a diagonal, are easily exponentiated after discretisation. The remaining terms are $\tilde{\mathcal{R}}_2 = \mathcal{O}(\varepsilon^{3\sigma-1})$ and $\tilde{\mathcal{T}}_3 = \mathcal{O}(\varepsilon^{5\sigma-1})$. Except for cases with $\sigma < \frac{1}{3}$, where the spectral radius of $\tilde{\mathcal{R}}_2$ increases with a decreasing ε , the number of iterations required in the Krylov approximation of the exponentials is small and derived in a straightforward manner from the bound (5.3). The exponentiation of $\tilde{\mathcal{R}}_2$ requires three iterations for $\sigma = 1$ and five iterations for $\sigma = \frac{1}{2}$, whilst $\tilde{\mathcal{T}}_3$ can be exponentiated in two, two and three iterations, respectively, in the cases $\sigma = 1, \frac{1}{2}, \frac{1}{4}$.

For the case $\sigma = \frac{1}{4}$, the spectral radius of $\tilde{\mathcal{R}}_2$ becomes large enough to be of concern with regard to the side condition in (5.3), $m \geq \rho$, where, it should be recalled, m was the number of iterations and ρ the spectral radius. In this case, as noted by [10], the error does not decrease substantially till $m \geq \rho$, decreasing rapidly thereafter. Rewriting (5.3) as

$$\|e^{\mathcal{A}} \mathbf{v} - \mathcal{V}_m e^{\mathcal{H}_m} \mathcal{V}_m^* \mathbf{v}\|_2 \leq 12 \exp\left(\frac{-\rho^2 + 4m^2(1 - \log 2)}{4m}\right) \left(\frac{\rho}{m}\right)^m, \quad m \geq \rho, \tag{6.2}$$

from which, with the choice $m \geq \alpha\rho$ for some $\alpha > 1$, we end up with an estimate

$$\|e^A \mathbf{v} - \mathcal{V}_m e^{\mathcal{H}_m} \mathcal{V}_m^* \mathbf{v}\|_2 \leq 12 \exp\left(\frac{-\rho^2 + 4m^2(1 - \log 2 - \log \alpha)}{4m}\right).$$

Exponential convergence, well in excess of what we require, can be achieved by an appropriate choice of α , whereby $\mathcal{O}(\varepsilon^{-1/4})$ iterations prove adequate. With $\alpha = 2$, for instance, the error term works out to roughly $\exp(-\frac{7}{8}\rho)$. After $\mathcal{O}(\varepsilon^{-1/4})$ iterations we are also left with the task of exponentiating a $\mathcal{O}(\varepsilon^{-1/4}) \times \mathcal{O}(\varepsilon^{-1/4})$ upper Hessenberg matrix, which, although large, can be exponentiated by brute force using MATLAB’s `expm` in $\mathcal{O}(\varepsilon^{-3/4} \log \varepsilon^{-1})$ operations.

The $\mathcal{O}(\varepsilon^{-5/4} \log \varepsilon^{-1})$ cost of the Krylov iterations for $e^{\tilde{\mathcal{R}}_2} \mathbf{v}$ dominates the cost of each time step of the splitting with $\sigma = \frac{1}{4}$, making the overall cost $\mathcal{O}(\varepsilon^{-3/2} \log \varepsilon^{-1})$. This is no less than the cost of the more accurate $\sigma = \frac{1}{2}$ splitting. In the case of $\sigma = 1$ – where, of course, a much smaller error is achieved with the same number of exponentials – the overall cost comes to $\mathcal{O}(\varepsilon^{-2} \log \varepsilon^{-1})$.

There seems little point in considering a σ smaller than $\frac{1}{3}$, where $\tilde{\mathcal{R}}_2$ becomes $\mathcal{O}(1)$ or larger and the number of Krylov iterations required for $e^{\tilde{\mathcal{R}}_2} \mathbf{v}$ starts increasing as $\varepsilon \rightarrow 0$. Even where the spectral radius does decrease with ε , a small σ makes the constraint $m \geq \rho$ in (5.3) a graver concern. With $\sigma = \frac{1}{4}$, for instance, where $\tilde{\mathcal{T}}_3$ scales as $\tilde{c}\varepsilon^{1/4}$, this requires $\varepsilon \leq (3/\tilde{c})^4$. The constant \tilde{c} depends on the interaction potential and circumstances where this constraint can become a serious concern are far from inconceivable.

7 Numerical Results

We present numerical results for two interaction-potential wave-function pairs. The wave functions used in our experiments are $u_1(x) = \frac{1}{100}\phi(x+0.6)e^{i20\pi x}$ and $u_2 = \psi$, where ϕ and ψ were previously introduced as

$$\phi(x) = e^{-20 \sin^2(\pi x/2)}, \quad \psi(x) = e^{-4(\sin^2(5\pi x/2) + \sin^2(\pi x/2))}.$$

The first of these is a moderately oscillating wave train with a periodic Gaussian envelope seen travelling to the right in free space ($V = 0$). In our experiments, these move under the influence of the interaction potentials

$$V_1(x) = b(s(x)) \sin(20\pi x),$$

$$V_2(x) = \frac{1}{5} + \frac{1}{2} b\left(s\left(x + \frac{1}{10}\right)\right) + \frac{3}{10} \left(\sin^4\left(2\pi x - \frac{24}{35}\right) + \sin^2\left(5\pi x - \frac{8}{3}\right)\right),$$

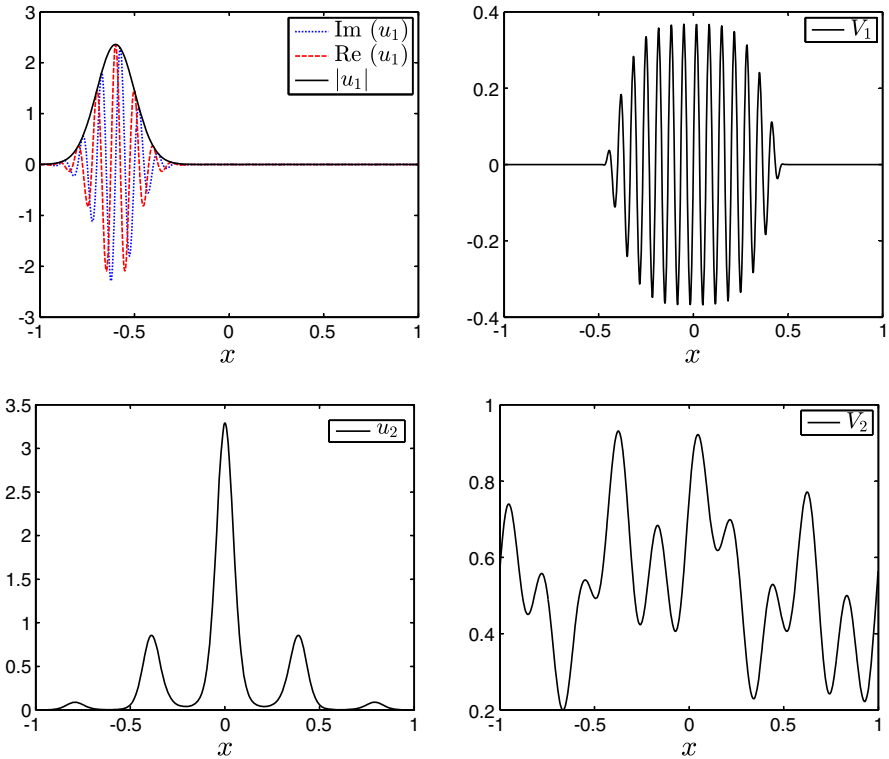


Fig. 4 Interaction-potential wave-function pair u_1, V_1 (top row) and u_2, V_2 (bottom row)

where b is a bump function turned periodic by composition with s ,

$$b(x) = \begin{cases} \exp\left(-\frac{1}{1-x^2}\right) & |x| < 1 \\ 0 & |x| \geq 1 \end{cases},$$

$$s(x) = 1 - \sin(\pi(x + 1/2)).$$

Physically, the first pair shown in Fig. 4 (top row) is an attempt at modelling a wave packet heading towards a periodic lattice. The second pair, Fig 4 (bottom row), has no physical motivation and is chosen for its complexity.

The error estimates are, of course, of an asymptotic nature, and it is hardly surprising that for some cases the true nature of the error estimate does not emerge till very small values of ε (Fig. 5). In the case of V_1 , with $\sigma = \frac{1}{2}$ or $\frac{1}{4}$, for instance, where one of the terms omitted in the splitting, $-\frac{1}{24}\tau^3\varepsilon(\partial_x^4 V)$, is fairly large, we do not see a noticeable decrease till very small values of ε , unless the magnitude of the interaction potential is decreased. In Fig. 6 (top row) the error is seen to approach the asymptotic estimate at an earlier stage in the case of a smaller potential. The asymptotic bounds are very much adhered to, but here we become limited by the inefficiency of the reference method – MATLAB’s `expm` – which does not allow us to go beyond moderate values of $M = \mathcal{O}(\varepsilon^{-1})$.

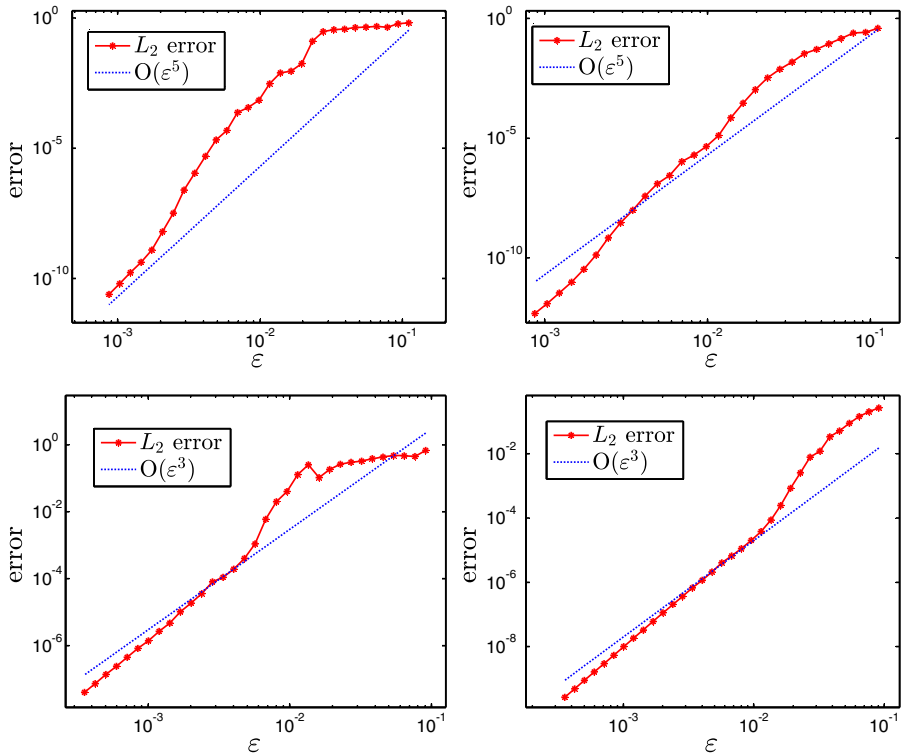


Fig. 5 Error plots for (1, 1)-splittings showing L_2 error for pairs (u_1, V_1) (left column) and (u_2, V_2) (right column): $S_{(1,1),2}^{[2]}$ splitting (top row) with seven exponentials has an error estimate of $\mathcal{O}(\varepsilon^5)$; $S_{(1,1),1}^{[2]}$ splitting (bottom row), which omits \tilde{T}_3 , uses five exponentials and has an error estimate of $\mathcal{O}(\varepsilon^3)$

All estimates and bounds in our analysis were, of course, derived with respect to the L_2 norm. This is approximated by the ℓ_2 norm on the grid $\mathbf{x}_M = (x_1, \dots, x_M)$,

$$\|f\|_{L_2[-1,1]} = \sqrt{\int_{-1}^1 |f(x)|^2 dx} \approx \sqrt{\frac{2}{M} \sum_{i=1}^M |f(x_i)|^2} = \sqrt{\frac{2}{M}} \|f\|_{\ell_2[\mathbf{x}_M]} \tag{7.1}$$

Of more consequence in numerical settings, arguably, is the behaviour of the L_∞ error. The ℓ_∞ norm is a very good approximation for the L_∞ norm, converging rapidly as $M \rightarrow \infty$. Noting the inequality $\|f\|_{\ell_\infty} \leq \|f\|_{\ell_2}$, one should expect the L_∞ error to be worse off than the L_2 error by a factor of $\sqrt{M} = \mathcal{O}(\varepsilon^{-1/2})$, and this is indeed seen to be the case in our experiments (bottom row, Figs. 6, 7).

8 Conclusions

In this paper we presented a methodology for the computation of the semiclassical Schrödinger equation (1.4) with small values of ε . It led to asymptotic exponential

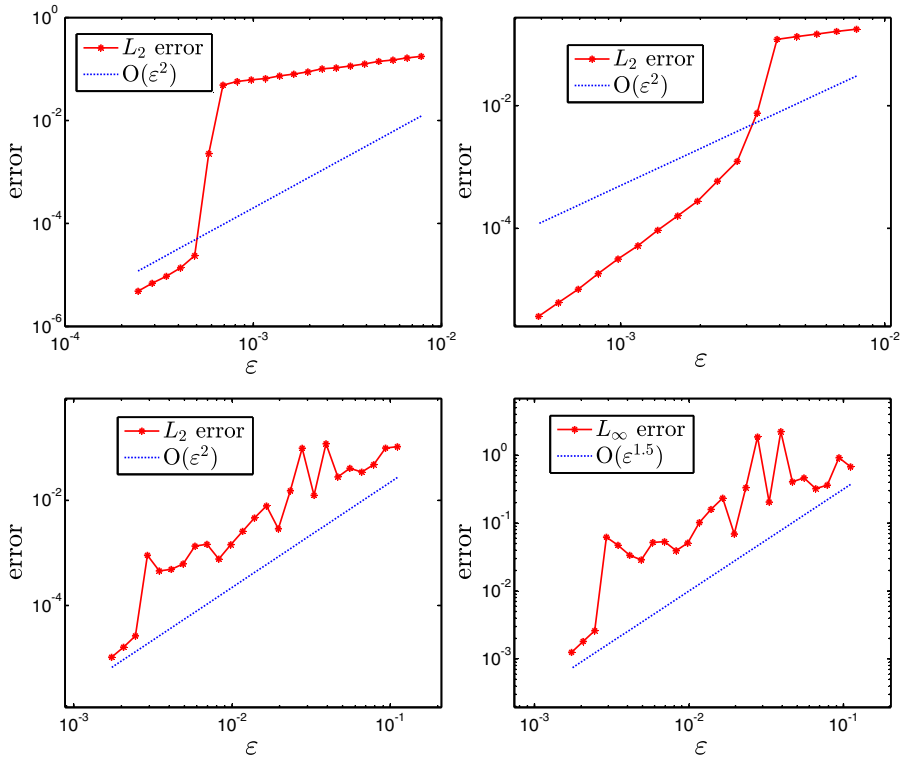


Fig. 6 Error plots for $S_{(1, \frac{1}{2}, 2)}^{[2]}$ splitting: (top) L_2 error for pairs $(u_1, \frac{1}{10} V_1)$ (left) and $(u_1, \frac{1}{100} V_1)$ (right) demonstrates asymptotic nature of error estimate $\mathcal{O}(\varepsilon^2)$; (bottom) relation between L_2 error (left) and L_∞ error (right) for (u_2, V_2) is evident, error estimates being $\mathcal{O}(\varepsilon^2)$ and $\mathcal{O}(\varepsilon^{\frac{3}{2}})$ respectively

splitting *à la* (3.9), (3.12) and (6.1), where each consecutive argument (except perhaps for one) is progressively smaller. Moreover, these arguments are skew-Hermitian (whence stability and unitarity follow) and the underlying exponentials are easy to compute. All this was accomplished by creating a Lie-algebraic framework that uses nested commutators yet avoids their expensive computation, combined with a repeated use of the symmetric BCH formula to form a symmetric Zassenhaus splitting. We also discussed the choice of semidiscretisation and of effective means to approximate matrix exponentials.

We do not view the work of this paper as a finished and complete endeavour: it is more in the nature of an initial foray into a broad and fascinating subject area. There is a wide range of issues that our work raises. Some are already under active investigation, others more speculative:

- (1) *A time-dependent interaction potential* In place of (1.4) we can consider the (non-autonomous) semiclassical Schrödinger equation with time-dependent potential

$$\frac{\partial u}{\partial t} = i\varepsilon \frac{\partial^2 u}{\partial x^2} + i\varepsilon^{-1} V(x, t)u, \quad t \geq 0, \quad x \in [-1, 1],$$

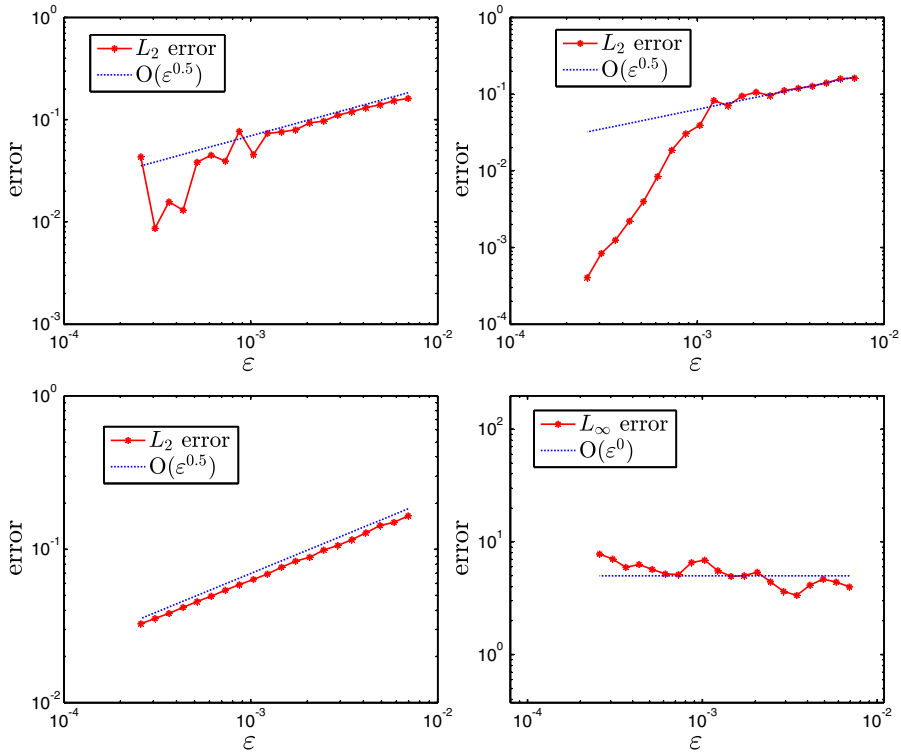


Fig. 7 Error plots for $S_{(1, \frac{1}{4}), 2}^{[2]}$ splitting: (top) L_2 error for pairs $(u_1, \frac{1}{10} V_1)$ (left) and $(u_1, \frac{1}{100} V_1)$ (right) demonstrates asymptotic nature of error estimate $O(\varepsilon^{\frac{1}{2}})$; (bottom) relation between L_2 error (left) and L_∞ error (right) for (u_2, V_2) is evident – error estimates being $O(\varepsilon^{\frac{1}{2}})$ and $O(\varepsilon^0)$ respectively

again with periodic boundary conditions. To this end, we need to combine our methodology – algebra of operators, symmetric Zassenhaus – with Magnus expansions [12]. Preliminary work indicates that, inasmuch as this leads to a considerably more complicated framework, it can fit into our narrative. Specifically, different Magnus terms can be written in a form consistent with the Lie algebra \mathfrak{G} . We expect to report on this work in the near future.

- (2) *A multivariate setting* An effective numerical discretisation of Eq. (1.3), evolving in a torus in \mathbb{C}^d , is the ultimate goal of this work. Insofar as small $d \geq 1$ is concerned, this is a fairly straightforward exercise, but matters are more complicated when d becomes large and the cost of $\mathcal{O}(M^d \log M)$ becomes unsustainable. It is clear that for our methodology to be scaleable to large dimensions, it must be combined with other approaches, e.g. sparse grids [2].
- (3) *The nonlinear Schrödinger equation* A major challenge is to apply our methodology in a nonlinear setting, e.g. to the nonlinear Schrödinger equation

$$i\varepsilon \frac{\partial u}{\partial t} = -\frac{\varepsilon^2}{2m} \frac{\partial^2 u}{\partial x^2} - V(x)u + \lambda |u|^2 u.$$

Preliminary investigation seems to indicate that a naïve generalisation does not work because we are not enjoying the reduction of negative powers of ε after commutation with Lie derivatives corresponding to $|u|^2$.

- (4) *Symmetric Zassenhaus in other settings* Exponential splittings have reached their apogee in the context of symplectic integrators for Hamiltonian ordinary differential equations [8, 15]. Can symmetric Zassenhaus be used in this setting? The idea seems particularly appealing in the context of Hamiltonian functions of the form

$$H(\mathbf{p}, \mathbf{q}) = H_1(\mathbf{p}, \mathbf{q}) + \varepsilon H_2(\mathbf{p}, \mathbf{q}),$$

where $0 < |\varepsilon| \ll 1$. Such systems occur often in celestial mechanics and many-body problems once there exists a large disparity of masses, and it is tempting to use an asymptotic splitting. However, in general we cannot employ in this context the formalism of Sect. 2.1, computing commutators easily. The computation of commutators in this context (in which they become Poisson brackets) is frowned upon because it is expensive. However, for special Hamiltonian functions this approach might be feasible.

Similar reasoning applies to volume-conserving geometric integrators based on splittings [16].

The symmetric Zassenhaus formula might also be relevant within the realm of partial differential equations in the presence of a small parameter, e.g. the Klein–Gordon equation

$$\frac{1}{c^2} \frac{\partial^2 u}{\partial t^2} = \nabla^2 u + \frac{m^2 c^2}{\hbar^2} u.$$

This, again, is a matter for further research.

References

1. Blanes, S., Casas, F. & Murua, A. (2006), Symplectic splitting operator methods tailored for the time-dependent Schrödinger equation, *J. Chem. Phys.* **124**, 234–105.
2. Bungartz, H.-J. & Griebel, M. (2004), Sparse grids, *Acta Numer.* **13**, 147–269.
3. Casas, F. & Murua, A. (2009), An efficient algorithm for computing the Baker–Campbell–Hausdorff series and some of its applications, *J. Math. Phys.* **50**, (electronic).
4. Faou, E. (2012), *Geometric Numerical Integration and Schrödinger Equations*, Zurich Lectures in Advanced Mathematics, The European Mathematical Society, Zürich.
5. Gallopoulos, E. & Saad, Y. (1992), Efficient solution of parabolic equations by Krylov approximation methods, *SIAM J. Sci. Stat. Comput.* **13**, 1236–1264.
6. Golub, G. H. & Van Loan, C. F. (1996), *Matrix Computations*, 3rd edn, Johns Hopkins University Press, Baltimore.
7. Griffiths, D. J. (2004), *Introduction to Quantum Mechanics*, 2nd edn, Prentice Hall, Upper Saddle River, NJ.
8. Hairer, E., Lubich, C. & Wanner, G. (2006), *Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations*, 2nd edn, Springer, Berlin.
9. Hesthaven, J. S., Gottlieb, S. & Gottlieb, D. (2007), *Spectral Methods for Time-Dependent Problems*, Cambridge University Press, Cambridge.

10. Hochbruck, M. & Lubich, C. (1997), On Krylov subspace approximations to the matrix exponential operator, *SIAM J. Numer. Anal.***34**, 1911–1925.
11. Iserles, A. (2008), *A First Course in the Numerical Analysis of Differential Equations*, 2nd edn, Cambridge University Press, Cambridge.
12. Iserles, A., Munthe-Kaas, H. Z., Nørsett, S. P. & Zanna, A. (2000), Lie-group methods, *Acta Numer.***9**, 215–365.
13. Jin, S., Markowich, P. & Sparber, C. (2011), Mathematical and computational methods for semiclassical Schrödinger equations, *Acta Numer.***20**, 121–210.
14. Lubich, C. (2008), *From Quantum to Classical Molecular Dynamics: Reduced Models and Numerical Analysis*, Zurich Lectures in Advanced Mathematics, The European Mathematical Society, Zürich.
15. McLachlan, R. I. & Quispel, G. R. W. (2002), Splitting methods, *Acta Numer.***11**, 341–434.
16. McLachlan, R. I., Munthe-Kaas, H. Z., Quispel, G. R. W. & Zanna, A. (2008), Explicit volume-preserving splitting methods for linear and quadratic divergence-free vector fields, *Found. Comput. Math.***8**, 335–3554.
17. Oteo, J. A. (1991), The Baker–Campbell–Hausdorff formula and nested commutator identities, *J. Math. Phys.***32**, 419–424.
18. Reutenauer, C. (1993), *Free Lie Algebras*, London Mathematical Society Monographs **7**, Oxford University Press, Oxford.
19. Tal Ezer, H. & Kosloff, R. (1984), An accurate and efficient scheme for propagating the time dependent Schrödinger equation, *J. Chem. Phys.***81**, 3967–3976.
20. Yošida, H. (1990), Construction of higher order symplectic integrators, *Phys. Lett.***150**, 262–268.