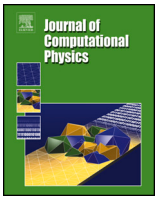
Journal of Computational Physics 533 (2025) 113991

Contents lists available at ScienceDirect

Journal of Computational Physics

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



A quantum lattice algorithm with fourth-order accuracy for the one-dimensional Dirac equation

Paul J. Dellar

Mathematical Institute, University of Oxford, Radcliffe Observatory Quarter, Oxford OX2 6GG, UK

ARTICLE INFO

Keywords:
Lattice algorithms
Splitting algorithms
Composition schemes
Discrete invariants
Resonant interactions

ABSTRACT

The discrete time quantum walk is a quantum cellular automaton whose wavefunction comprises pairs of complex numbers assigned to uniformly spaced points on a line. The wavefunction evolves through the application of an alternating sequence of unitary operators: streaming of wavefunction values to adjacent points, and a Hadamard-type unitary matrix to blend pairs of values at individual points. Each operator generates the exact evolution due to part of the Hamiltonian for the one-dimensional Dirac equation over a finite time step. Composing these operators thus creates a discrete approximation to the Dirac equation. However, the composition of two non-commuting operators creates a global splitting error proportional to the length of the time step. The global error can be reduced from first order to second order in the time step by a unitary pre- and post-processing of the initial conditions and final output. The algorithm then becomes equivalent to a symmetric composition, a Strang splitting, between the two operators. This paper describes a fourth-order accurate composition scheme using nine stages, the fewest possible when the lengths of the time steps employed in the different stages are constrained to be integer multiples of some base time step. Each stage is itself a symmetric composition between two operators. This fourth-order scheme produces quantitatively smaller errors for a typical benchmark problem on spatial lattices with 1024 or more points, and shows the expected fourth-order convergence on sufficiently fine lattices. It has greater accuracy, over sufficiently long times, than three better-known fourth-order composition schemes using fewer stages, but with lengths related by irrational coefficients. The truncation error for plane-wave solutions is due to an operator that separates into a resonant part proportional to the Hamiltonian, and a non-resonant part orthogonal to the Hamiltonian. The resonant part commutes with the exact evolution operator, so its error accumulates to grow linearly with time. The orthogonal part produces oscillations that remain bounded over many time steps. The nine-stage integer scheme has the smallest resonant truncation error of the four schemes, despite being the only scheme that can be implemented using local operations. The other schemes implement streaming by irrational fractions of the lattice spacing using discrete Fourier transforms.

1. Introduction

A good discretisation should preserve structural properties of the system being discretised. This is the goal of the field of geometric numerical integration [1–4]. The structural properties of a quantum system are that states are described by elements of a complex

E-mail address: paul.dellar@maths.ox.ac.uk.

https://doi.org/10.1016/j.jcp.2025.113991
Received 21 April 2024; Received in revised form 18 March 2025; Accepted 4 April 2025

Available online 8 April 2025

Hilbert space, and that states evolve through unitary transformations. The solution of the Schrödinger equation for a quantum system evolving in continuous time can be expressed using a continuous family of unitary transformations generated by the Hamiltonian operator. This ensures that the evolving state always yields a normalised probability distribution for the measured values of observables under the Born rule [5–7].

A quantum system that evolves through discrete time steps must do so through finite unitary transformations of the Hilbert space. It is natural to study systems in which these finite unitary transformations are themselves composed of simpler unitary transformations. The discrete time quantum walk is a widely studied example [8–11]. It employs a lattice of uniformly spaced points on a line. Each point is equipped with two complex degrees of freedom, often called u for "up" and d for "down". One half of an evolutionary step involves translating the values of u and d to adjacent lattice points, upwards for u and downwards for d. This operation corresponds to applying a unitary matrix to a vector of u or d values on a finite lattice with periodic boundary conditions (see Sec. 3 below). The other half of the evolutionary step applies a Hadamard-like unitary matrix to blend u and d at each individual lattice point.

The discrete time quantum walk can be treated as a natural quantum generalisation of a classical random walk [8,11]. It can also be interpreted as one of the simplest quantum cellular automata with a translation-invariant update rule that generates non-trivial dynamics [9,10]. The only quantum cellular automata with translation-invariant update rules for a single degree of freedom per lattice point generate uniform translations and/or global changes of phase [10].

The discrete time quantum walk also turns out to be a discrete approximation to the one-dimensional Dirac equation that describes the relativistic motion of a spin-1/2 quantum particle [5–7,10,12]. One can discretise the one-dimensional Dirac equation by splitting its Hamiltonian operator into differential and algebraic parts. Evolution under each part is exactly soluble. Evolution under the differential part reduces to discrete translations of the wavefunction values on the lattice for a suitable choice of time step Δt. We express this unitary operation using the streaming operator S. The unitary operation C arising from the algebraic part can be computed exactly using matrix exponentiation. Both S and C are defined concretely in Sec. 3 below. Although each part is solved exactly, approximating the evolution under the complete Hamiltonian for a time step of length Δt by the composition SC incurs an $O(\Delta t^2)$ local error because the S and C operators do not commute. Viewed as a numerical algorithm for solving the Dirac equation, the discrete time quantum walk thus has only first order global accuracy. It is equivalent to the "quantum lattice Boltzmann algorithm" as a first-order accurate approximation to the one-dimensional Dirac equation [13]. It is also equivalent to the Feynman checkerboad approximation of the path integral solution of the one-dimensional Dirac equation [14].

Many geometric integration schemes preserve structural properties using compositions of operators with known properties, such as being unitary, symplectic or preserving phase-space volume. The geometric integration literature contains many more elaborate composition schemes that offer higher orders of accuracy [1–4]. The accuracy may be raised to second order by replacing the composition SC with the symmetric composition $\mathsf{C}^{1/2}\mathsf{SC}^{1/2}$, where $\mathsf{C}^{1/2}$ is the operator that evolves the solution for half a time step under the algebraic terms alone. This is known as Strang splitting in the numerical analysis literature [15]. One can rewrite this second-order symmetric composition as the original SC composition applied to a transformed wavefunction (see Sec. 3). The same second-order scheme can also be derived by integrating the one-dimensional Dirac equation along its characteristics for a time step Δt, approximating the integrals of the algebraic terms with the trapezoidal rule, then making a unitary change of variables to create an explicit scheme [16,17]. This is the quantum analog of the derivation of second-order accurate hydrodynamic lattice Boltzmann equations by integrating along characteristics [18]

More recently, the second-order composition scheme for the one-dimensional Dirac equation, and hence the original discrete time quantum walk, have been shown to be equivalent, under further unitary transformations, to the Ablowitz–Kruskal–Ladik scheme for the Klein–Gordon equation [17,19], and to the Du Fort–Frankel scheme for the Schrödinger equation [20–22]. These schemes may be used as building blocks to construct schemes for the two- and three-dimensional Dirac equations [17,23–25]. Streaming along the directions parallel to the different coordinate axes must be implemented separately using unitary transformations to diagonalise the part of the Hamiltonian operator containing derivatives with respect to the corresponding spatial coordinate. These schemes have been used to study the Klein paradox for the oblique incidence of a charge-carrying wavepacket on a straight potential barrier in graphene [25,26]. This problem is very sensitive to the angle of incidence, and hence benefits from highly accurate computations.

The early 1990s saw a renewed interest in constructing composition schemes with third or higher order accuracy in the time step [27–29]. The symmetric composition schemes formed from a symmetric and time-reversible base step have particularly attractive properties. If $\widetilde{\mathsf{U}}_2(\delta t) = \mathsf{C}(\delta t/2)\mathsf{S}(\delta t)\mathsf{C}(\delta t/2)$ denotes the symmetric evolution operator formed by Strang splitting for an arbitrary time interval δt, an m-stage "symmetric symmetric" composition scheme for odd values of m takes the general form

$$\widetilde{\mathsf{U}}(\Delta t) = \widetilde{\mathsf{U}}_2(w_1\Delta t)\widetilde{\mathsf{U}}_2(w_2\Delta t)\ldots\widetilde{\mathsf{U}}_2(w_{(m+1)/2})\ldots\widetilde{\mathsf{U}}_2(w_2\Delta t)\widetilde{\mathsf{U}}_2(w_1\Delta t). \tag{1}$$

This operator is intended to approximate the exact evolution operator over a time step Δt. The coefficients $w_1, \ldots, w_{(m+1)/2}$ determine the relative lengths of the different stages. The symmetry ensures that only odd powers of Δt appear in the series expansion of (1) given by the Baker–Campbell–Hausdorff formula [4,29,30]. This greatly reduces the number of conditions that the coefficients w_i must satisfy, from eight in the most general case down to just two [31]. The scheme (1) is consistent if $2w_1 + 2w_2 + \cdots + w_{(m+1)/2} = 1$, which normalises the length of the time step. It has global fourth-order accuracy if $2w_1^3 + 2w_2^3 + \cdots + w_{(m+1)/2}^3 = 0$. The $O(\Delta t^3)$ errors from the different stages then cancel. The determining equations have the unique solution $w_1 = (2 - 2^{1/3})^{-1}$ and $w_2 = 1 - 2w_1$ for three-stage schemes [27,29]. Another popular choice is the five-stage scheme with coefficients $w_1 = w_2 = (4 - 4^{1/3})^{-1}$ and $w_3 = 1 - 4w_1$ [29,32].

Most attention has been focused on schemes like these with irrational coefficients, and irrational ratios between coefficients. They therefore cannot be efficiently implemented on a lattice with uniformly spaced points. Only integer powers of the base streaming operator S can be implemented efficiently by translating wavefunction values to nearby lattice points. However, a slightly earlier

paper by Blow & Wood [33] described a fourth-order composition scheme that is equivalent to a nine-stage symmetric-symmetric scheme with $w_1 = w_2 = w_3 = w_4 = 1/6$ and $w_5 = -1/3$. This scheme can be implemented on a uniformly spaced lattice if one rescales the base time step to $6\Delta t$. The individual stages then evolve the solution forwards by Δt, and backwards by $2\Delta t$. This scheme has very recently been rediscovered and used to simulate some purely hyperbolic problems such as the inviscid Burgers and shallow water equations [34]. In this paper we shall explore the properties of the Blow & Wood scheme for the one-dimensional Dirac equation.

2. The one-dimensional Dirac equation

The one-dimensional Dirac equation [5–7] for a spin-1/2 particle of mass m moving along the z axis in an external scalar potential $V(z)$ can be written as a first-order hyperbolic system for two complex-valued functions $u(z,t)$ and $d(z,t)$,

$$\partial_t u + \partial_z u = md - iVu, \quad \partial_t d - \partial_z d = -mu - iVd. \tag{2}$$

We use dimensionless natural units in which the speed of light c and the reduced Planck's constant $\hbar$ are both equal to unity. The system (2) may be written as an abstract Schrödinger equation

$$i\partial_t \Psi = \mathcal{H}\Psi, \tag{3}$$

using a two-component vector wavefunction Ψ and a matrix Hamiltonian operator $\mathcal{H}$ whose diagonal elements are themselves differential operators,

$$\Psi = \begin{pmatrix} u \\ d \end{pmatrix}, \quad \mathcal{H} = \begin{pmatrix} V - i\partial_z & im \\ -im & V + i\partial_z \end{pmatrix}. \tag{4}$$

The matrix operator $\mathcal{H}$ is Hermitian (or self-adjoint) with respect to the inner product

$$\langle \Xi, \Phi \rangle = \int \Xi^\dagger \Phi \, dz \tag{5}$$

for pairs of two-component wavefunctions. A superscript dagger denotes a Hermitian conjugate, so $\Psi^\dagger = (u^\star, d^\star)$, where * denotes a complex conjugate. We can thus write the squared norm of a state as

$$\langle \Psi, \Psi \rangle = \int |u|^2 + |d|^2 \, dz. \tag{6}$$

The solution of a linear initial value problem with a time-independent operator is given by the exponential of that operator acting on the initial conditions. The solution of (3) may thus be written as $\Psi(\cdot,t) = \mathsf{U}(t)\Psi(\cdot,0)$ using the evolution operator $\mathsf{U}(t) = \exp(-it\mathcal{H})$. This operator is unitary, so $\langle \Psi, \Psi \rangle$ is independent of time, as required by the Born rule. States that are normalised with $\langle \Psi, \Psi \rangle = 1$ will thus remain normalised. Solutions of (2) also satisfy an energy conservation law. The matrix element of the Hamiltonian operator,

$$\langle \Psi, \mathcal{H}\Psi \rangle = \int V(z)(|u|^2 + |d|^2) + i\left(d^\star \partial_z d - u^\star \partial_z u\right) + im\left(u^\star d - d^\star u\right) \, dz, \tag{7}$$

is constant in time, as shown in Sec. 6. This matrix element is equal to the expectation $\langle \mathcal{H} \rangle$ for normalised states with $\langle \Psi, \Psi \rangle = 1$.

The time-dependent unitary change of variables [6]

$$\begin{pmatrix} u \\ d \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix} \begin{pmatrix} P \\ M \end{pmatrix} e^{-imt} \tag{8}$$

transforms the one-dimensional Dirac equation (2) into

$$\partial_t P + \partial_z M = -iVP, \tag{9a}$$

$$\partial_t M + \partial_z P = 2imM - iVM. \tag{9b}$$

This change of variables is designed firstly to diagonalise the algebraic terms proportional to m in the Hamiltonian (4), and secondly to eliminate the algebraic term proportional to m on the right hand side of (9a). Seeking a slowly varying solution in which $|\partial_t M| \ll 2m|M|$ for a sufficiently weak potential, with $|V(z)| \ll 2m$ for all z, allows us to replace (9b) with the adiabatic approximation $\partial_z P = 2imM$. Substituting this approximation into (9a) gives a closed evolution equation for P,

$$i\partial_t P = -\frac{1}{2m}\partial_{zz} P + VP. \tag{10}$$

This is Schrödinger's equation for a particle moving non-relativistically in a scalar potential, again in natural units with $\hbar = 1$. Schrödinger's equation thus describes the slowly varying solutions of the Dirac equation that are valid in the non-relativistic limit, provided we transform away the $\exp(-imt)$ oscillations from the particle's relativistic rest energy. These oscillations have the Compton frequency $mc^2/\hbar$ in SI units.

3. A second-order accurate scheme for the one-dimensional Dirac equation

We can derive a numerical scheme for the one-dimensional Dirac equation (2) by writing $\mathcal{H} = \mathcal{H}_S + \mathcal{H}_C$ to separate the differential and algebraic parts of the Hamiltonian. This is equivalent to writing (2) as

$$\partial_t \begin{pmatrix} u \\ d \end{pmatrix} = (S + C) \begin{pmatrix} u \\ d \end{pmatrix}, \tag{11}$$

with two anti-Hermitian operators

$$S = \begin{pmatrix} -\partial_z & 0 \\ 0 & \partial_z \end{pmatrix}, \quad C = \begin{pmatrix} -iV & m \\ -m & -iV \end{pmatrix}. \tag{12}$$

These operators $S = -i\mathcal{H}_S$ and $C = -i\mathcal{H}_C$ are anti-Hermitian because we have omitted the imaginary i prefactor from the time derivative in (11). The symbol S denotes streaming. The symbol C denotes either collisions, following the structural analogy between (2) and a discrete Boltzmann equation for interacting classical particles [13], or coin tosses in the context of discrete time quantum walks [8,10].

The exponentials of the C and S operators have the closed-form expressions

$$\mathsf{C} = \exp(\Delta t C) = e^{-iV\Delta t} \begin{pmatrix} \cos m\Delta t & \sin m\Delta t \\ -\sin m\Delta t & \cos m\Delta t \end{pmatrix}, \tag{13}$$

and

$$\mathsf{S} = \exp(\Delta t S) = \begin{pmatrix} \exp(-\Delta t \partial_z) & 0 \\ 0 & \exp(\Delta t \partial_z) \end{pmatrix}. \tag{14}$$

The action of $\exp(\pm\Delta t\partial_z)$ on a smooth function f is given by $(\exp(\pm\Delta t\partial_z)f)(z) = f(z \pm \Delta t)$ using the identity

$$f(z + \Delta z) = \sum_{n=0}^{\infty} \frac{\Delta z^n}{n!} (\partial_z^n f)\Big|_z = \exp(\Delta z \partial_z) f\Big|_z. \tag{15}$$

The $|_z$ notation indicates that f and all its derivatives are evaluated at z. The S operator thus describes the propagation of u along the characteristics $z = z_0 + (t - t_0)$, and of d along the characteristics $z = z_0 - (t - t_0)$, where z_0 and t_0 are constants. The action of S on a vector of u and d values at N uniformly spaced points $z_j = (j-1)\Delta z$ with periodic boundary conditions is given by

$$\mathsf{S}(u_1, \ldots, u_N, d_1, \ldots, d_N)^{\mathsf{T}} = (u_N, u_1, \ldots, u_{N-2}, u_{N-1}, d_2, d_3, \ldots, d_N, d_1)^{\mathsf{T}}. \tag{16}$$

Here $(\cdots)^{\mathsf{T}}$ denotes the transpose of a vector, as distinct from the conjugate transpose $(\cdots)^\dagger$. Similarly, the action of C is given by

$$\mathsf{C}(u_1, \ldots, u_N, d_1, \ldots, d_N)^{\mathsf{T}} = (u'_1, \ldots, u'_N, d'_1, \ldots, d'_N)^{\mathsf{T}}, \tag{17}$$

where, for each j,

$$\begin{pmatrix} u'_j \\ d'_j \end{pmatrix} = \exp\Big(-iV(z_j)\Delta t\Big) \begin{pmatrix} \cos m\Delta t & \sin m\Delta t \\ -\sin m\Delta t & \cos m\Delta t \end{pmatrix} \begin{pmatrix} u_j \\ d_j \end{pmatrix}. \tag{18}$$

Given these two exact solution operators S and C for initial value problems with the two half-systems

$$\partial_t \boldsymbol{\Psi} = S\boldsymbol{\Psi}, \quad \partial_t \boldsymbol{\Psi} = C\boldsymbol{\Psi}, \tag{19}$$

we can compose them to construct a globally first-order accurate scheme for (11) in the Lie–Trotter form [35]

$$\boldsymbol{\Psi}(t + \Delta t) = \mathsf{SC}\boldsymbol{\Psi}(t) + O(\Delta t^2), \tag{20}$$

where $\boldsymbol{\Psi}(t) = (u_1(t), \ldots, u_N(t), d_1(t), \ldots, d_N(t))^{\mathsf{T}}$ is the vector of u and d values on the spatial lattice at time t. This scheme is the discrete time quantum walk, as given explicitly by

$$u_{j+1}^{n+1} = \exp(-iV(z_j)\Delta t)\big(u_j^n \cos m\Delta t + d_j^n \sin m\Delta t\big), \tag{21a}$$

$$d_{j-1}^{n+1} = \exp(-iV(z_j)\Delta t)\big(d_j^n \cos m\Delta t - u_j^n \sin m\Delta t\big). \tag{21b}$$

These equations determine the u and d values on the spatial lattice at time $(n+1)\Delta t$, indicated by the superscript $n+1$ on the left-hand sides, in terms of the u and d values at time $n\Delta t$, indicated by the superscript n on the right-hand sides.

The $O(\Delta t^2)$ local truncation error in (20) arises from the Baker–Campbell–Hausdorff formula for the product of exponentials of noncommuting operators [4,30]

$$\mathsf{SC} = \exp(\Delta t S)\exp(\Delta t C) = \exp(\mathcal{Z}_1), \tag{22}$$

where

$$\mathcal{Z}_1 = \Delta t(\mathcal{S}+\mathcal{C}) + \frac{1}{2}\Delta t^2[\mathcal{S},\mathcal{C}] + \frac{1}{12}\Delta t^3\left([\mathcal{S},[\mathcal{S},\mathcal{C}]] + [\mathcal{C},[\mathcal{C},\mathcal{S}]]\right) - \frac{1}{24}\Delta t^4[\mathcal{C},[\mathcal{S},[\mathcal{S},\mathcal{C}]]] + \cdots. \tag{23}$$

The exponential of the first term gives the exact unsplit evolution operator $\exp(\Delta t(\mathcal{S}+\mathcal{C}))$. However, it is followed by an infinite series of error terms involving sums of nested commutators, as defined by $[\mathcal{A},\mathcal{B}] = \mathcal{A}\mathcal{B} - \mathcal{B}\mathcal{A}$ for any two operators $\mathcal{A}$ and $\mathcal{B}$. Each error term contains a power of Δt one higher than the number of nested commutators.

The symmetric composition scheme, equivalent to Strang splitting [15], replaces (20) with

$$\boldsymbol{\Psi}(t+\Delta t) = \mathsf{C}^{1/2}\mathsf{S}\mathsf{C}^{1/2}\,\boldsymbol{\Psi}(t) \tag{24}$$

to cancel the leading-order truncation error. The symmetric Baker–Campbell–Hausdorff formula gives [4,29,36]

$$\mathsf{C}^{1/2}\mathsf{S}\mathsf{C}^{1/2} = \exp(\tfrac{1}{2}\Delta t\mathcal{C})\exp(\Delta t\mathcal{S})\exp(\tfrac{1}{2}\Delta t\mathcal{C}) = \exp(\mathcal{Z}_2). \tag{25}$$

The series for

$$\mathcal{Z}_2 = \Delta t(\mathcal{S}+\mathcal{C}) - \frac{1}{24}\Delta t^3\left([\mathcal{S},[\mathcal{S},\mathcal{C}]] + 2[\mathcal{C},[\mathcal{C},\mathcal{S}]]\right) + O(\Delta t^5) \tag{26}$$

contains only odd powers of Δt. The symmetric composition eliminates all even powers of Δt, not just the $O(\Delta t^2)$ term, because it is time-reversible:

$$\left(\exp(\tfrac{1}{2}\Delta t\mathcal{C})\exp(\Delta t\mathcal{S})\exp(\tfrac{1}{2}\Delta t\mathcal{C})\right)^{-1} = \exp(-\tfrac{1}{2}\Delta t\mathcal{C})\exp(-\Delta t\mathcal{S})\exp(-\tfrac{1}{2}\Delta t\mathcal{C}), \tag{27}$$

so $\mathcal{Z}_2(\mathcal{S},\mathcal{C}) = -\mathcal{Z}_2(-\mathcal{S},-\mathcal{C})$ must be an odd function of Δt [29,37]. This illustrates the general result that any time-reversible discrete approximation to the exact evolution operator must have an even order of accuracy [30].

The symmetric composition scheme (24) can be rewritten in the previous asymmetric Lie–Trotter form

$$\overline{\boldsymbol{\Psi}}(t+\Delta t) = \mathsf{S}\mathsf{C}\overline{\boldsymbol{\Psi}}(t), \tag{28}$$

by making the unitary change of variables (for all t)

$$\overline{\boldsymbol{\Psi}}(t) = \mathsf{C}^{-1/2}\boldsymbol{\Psi}(t), \tag{29}$$

where $\mathsf{C}^{-1/2}$ is applied to each pair $(u_j, d_j)^{\mathsf{T}}$ as in (18). This is equivalent to coalescing adjacent pairs of $\mathsf{C}^{1/2}$ operations when applying (24) repeatedly to evolve the solution over two or more time steps [38,39]. The Lie–Trotter scheme can thus be upgraded to second-order accuracy merely by pre-processing the initial conditions and post-processing the eventual output [15,40]. An analogous pre- and post-processing has been used to raise the order of accuracy of symplectic integrators for classical Hamiltonian systems [41,42].

4. Fourth-order accurate schemes for the one-dimensional Dirac equation

Having achieved second-order accuracy using the symmetric composition $\mathsf{C}^{1/2}\mathsf{S}\mathsf{C}^{1/2}$ it is natural to seek even higher orders of accuracy by composing longer sequences of exponentials of $\mathcal{S}$ and $\mathcal{C}$ with different coefficients. The symmetric-symmetric schemes use a symmetric operation such as $\mathsf{C}^{1/2}\mathsf{S}\mathsf{C}^{1/2}$ as a building block, and use a symmetric sequence of these blocks to preserve time-reversibility. For example, if we write

$$\widetilde{\mathsf{U}}_2(\delta t) = \exp(\tfrac{1}{2}\delta t\mathcal{C})\exp(\delta t\mathcal{S})\exp(\tfrac{1}{2}\delta t\mathcal{C}) \tag{30}$$

for evolution by the Strang symmetric composition operator over a time step of arbitrary length δt, Suzuki [27] and Yoshida [29] showed that the three-stage composition scheme

$$\widetilde{\mathsf{U}}_4(\Delta t) = \widetilde{\mathsf{U}}_2(w_1\Delta t)\widetilde{\mathsf{U}}_2(w_2\Delta t)\widetilde{\mathsf{U}}_2(w_1\Delta t) \tag{31}$$

is a fourth-order accurate approximation to the exact evolution operator over a time step Δt when the coefficients are

$$w_1 = \frac{1}{2-2^{1/3}} \approx 1.35 \quad w_2 = -\frac{2^{1/3}}{2-2^{1/3}} \approx -1.70. \tag{32}$$

These satisfy the relations $2w_1 + w_2 = 1$, so the overall time step is of length Δt, and $2w_1^3 + w_2^3 = 0$, so the $O(\Delta t^3)$ error terms cancel between the three different substeps. All terms with even powers of Δt automatically vanish, as above.

The w_2 coefficient is negative, so the middle step in (31) is backwards in time. All composition schemes with order three or more must contain a backwards step [4,28,43,44]. This result follows easily for symmetric-symmetric compositions because they must have even order. The determining equation $2w_1^3 + 2w_2^3 + \cdots + w_{(m+1)/2}^3 = 0$ for fourth-order accuracy can only hold if at least one coefficient w_j is negative.

The scheme in (31) has a particularly large negative coefficient w_2, and thus a particularly large coefficient multiplying its $O(\Delta t^5)$ local truncation error. Greater accuracy can be achieved for the same computational work using the five-stage scheme [29,32]

$$\widetilde{\mathsf{U}}_2(w_1\Delta t)\widetilde{\mathsf{U}}_2(w_2\Delta t)\widetilde{\mathsf{U}}_2(w_3)\widetilde{\mathsf{U}}_2(w_2\Delta t)\widetilde{\mathsf{U}}_2(w_1\Delta t) \tag{33}$$

with coefficients

$$w_1 = w_2 = \frac{1}{4 - 4^{1/3}} \approx 0.414, \quad w_3 = 1 - 4w_1 = -\frac{4^{1/3}}{4 - 4^{1/3}} \approx -0.678. \tag{34}$$

The numerical experiments described below confirm that this scheme is indeed more accurate than the three-stage scheme for similar computational work.

The evolution generated by applying these scheme to a classical Hamiltonian system can be interpreted as the exact evolution generated by a modified Hamiltonian. McLachlan [31] showed that the difference between the modified Hamiltonian for a five-stage scheme and the exact Hamiltonian is slightly smaller for the coefficients $w_1 = 0.28$, $w_2 \approx 0.625$, and $w_3 \approx -0.810$. We shall refer to the scheme with these coefficients as the five-stage McLachlan scheme, and to the scheme with the coefficients in (34) as the five-stage Neri scheme.

4.1. A nine-stage scheme with integer coefficients

These schemes all involve coefficients whose ratios are irrational numbers, $w_2 = -2^{1/3} w_1$ for the three-stage scheme, and $w_3 = -4^{1/3} w_1$ for the five-stage scheme. The individual stages $\widetilde{\mathsf{U}}_2(w_i \Delta t)$ therefore cannot be implemented using local operations on a lattice of uniformly spaced points, even after rescaling the base time step Δt. However, the nine-stage scheme used by Blow & Wood [33] takes the general form (1) with

$$\widetilde{\mathsf{U}}_4 = \widetilde{\mathsf{U}}_2(w_1 \Delta t) \widetilde{\mathsf{U}}_2(w_2 \Delta t) \widetilde{\mathsf{U}}_2(w_3 \Delta t) \widetilde{\mathsf{U}}_2(w_4 \Delta t) \widetilde{\mathsf{U}}_2(w_5 \Delta t) \widetilde{\mathsf{U}}_2(w_4 \Delta t) \widetilde{\mathsf{U}}_2(w_3 \Delta t) \widetilde{\mathsf{U}}_2(w_2 \Delta t) \widetilde{\mathsf{U}}_2(w_1 \Delta t). \tag{35}$$

The coefficients are $w_1 = w_2 = w_3 = w_4 = 1$ and $w_5 = -2$. These satisfy $8w_1 + w_5 = 6$ and $8w_1^3 + w_5^3 = 0$, so this scheme advances the solution by $6\Delta t$ at each step with fourth-order accuracy.

This composition can be rewritten as

$$\widetilde{\mathsf{U}}_4 = \left(\widetilde{\mathsf{U}}_2(\Delta t)\right)^4 \widetilde{\mathsf{U}}_2(-2\Delta t) \left(\widetilde{\mathsf{U}}_2(\Delta t)\right)^4 = \left(\mathsf{C}^{1/2}\mathsf{S}\mathsf{C}^{1/2}\right)^4 \mathsf{C}^{-1}\mathsf{S}^{-2}\mathsf{C}^{-1} \left(\mathsf{C}^{1/2}\mathsf{S}\mathsf{C}^{1/2}\right)^4. \tag{36}$$

Coalescing adjacent pairs of collision steps gives

$$\widetilde{\mathsf{U}}_4 = \mathsf{C}^{1/2} (\mathsf{S}\mathsf{C})^4 \mathsf{C}^{-3/2} \mathsf{S}^{-2} \mathsf{C}^{-1/2} (\mathsf{S}\mathsf{C})^4 \mathsf{C}^{-1/2}, \tag{37}$$

which is equivalent to

$$\overline{\mathsf{U}}_4 = (\mathsf{S}\mathsf{C})^4 \mathsf{C}^{-3/2} \mathsf{S}^{-2} \mathsf{C}^{-1/2} (\mathsf{S}\mathsf{C})^4 \tag{38}$$

when applied to the transformed variables $\overline{\boldsymbol{\Psi}} = \mathsf{C}^{-1/2}\boldsymbol{\Psi}$ introduced in (29). We thus require a backwards double streaming step S^{-2}, and two negative fractional powers of the collision step. The latter cause no difficulties as we can implement $\mathsf{C}^{-3/2}$ and $\mathsf{C}^{-1/2}$ exactly using (18) with Δt replaced by $-3\Delta t/2$ and $-\Delta t/2$ respectively.

4.2. A nine-stage scheme with a halved time step

This nine-stage scheme can also be implemented with a halved time step to advance the solution by $3\Delta t$ on the same lattice of uniformly spaced points with an error that is 16 times smaller. This scheme is

$$\widehat{\mathsf{U}}_4 = \left(\mathsf{S}^{1/2}\mathsf{C}^{1/2}\right)^4 \mathsf{C}^{-3/4} \mathsf{S}^{-1} \mathsf{C}^{-1/4} \left(\mathsf{S}^{1/2}\mathsf{C}^{1/2}\right)^4 \tag{39}$$

applied to the transformed variables $\widehat{\boldsymbol{\Psi}} = \mathsf{C}^{-1/4}\boldsymbol{\Psi}$. We can do this because the half streaming steps $\mathsf{S}^{1/2}$ always occur in pairs. Two half streaming steps always result in translations by an integer, as $\pm\frac{1}{2} \pm \frac{1}{2} \in \{-1, 0, 1\}$ for all four choices of the two independent $\pm$ signs. This property does not hold for the more widely used composition schemes using irrational coefficients. The result of applying $\mathsf{S}^{1/2}\mathsf{C}^{1/2}\mathsf{S}^{1/2}\mathsf{C}^{1/2}$ to $\widehat{\boldsymbol{\Psi}} = (u_1, \dots u_N, d_1, \dots d_N)^{\mathsf{T}}$ is $\widehat{\boldsymbol{\Psi}}' = (u'_1, \dots u'_N, d'_1, \dots d'_N)^{\mathsf{T}}$ with

$$u'_j = \exp(-i\Delta t V_{j-1/2}) \left\{ a\left(a u_{j-1} + b d_{j-1}\right) \exp(-i\Delta t V_{j-1}) + b\left(a d_j - b u_j\right) \exp(-i\Delta t V_j) \right\}, \tag{40a}$$

$$d'_j = \exp(-i\Delta t V_{j+1/2}) \left\{ a\left(a d_{j+1} - b u_{j+1}\right) \exp(-i\Delta t V_{j+1}) - b\left(a u_j + b d_j\right) \exp(-i\Delta t V_j) \right\}, \tag{40b}$$

where $a = \cos(m\Delta t/2)$, $b = \sin(m\Delta t/2)$, and $V_j = V(z_j)$ for all j. The half-points $j \pm 1/2$ are now only needed for evaluating the potential V.

5. Numerical experiments

To illustrate the behaviour of these numerical schemes, we apply them to compute the solution of the one-dimensional Dirac equation (2) for a particle with dimensionless mass $m = 16$ in a domain of length $L = 16$ centred at $z = 0$ with periodic boundary conditions. The potential

$$V(z) = 20 \tanh(\tan(\pi z/L))^2 = 20 \left[\left(\frac{\pi z}{L}\right)^2 - \frac{2}{15}\left(\frac{\pi z}{L}\right)^6 + O\left(\left(\frac{z}{L}\right)^8\right) \right] \tag{41}$$

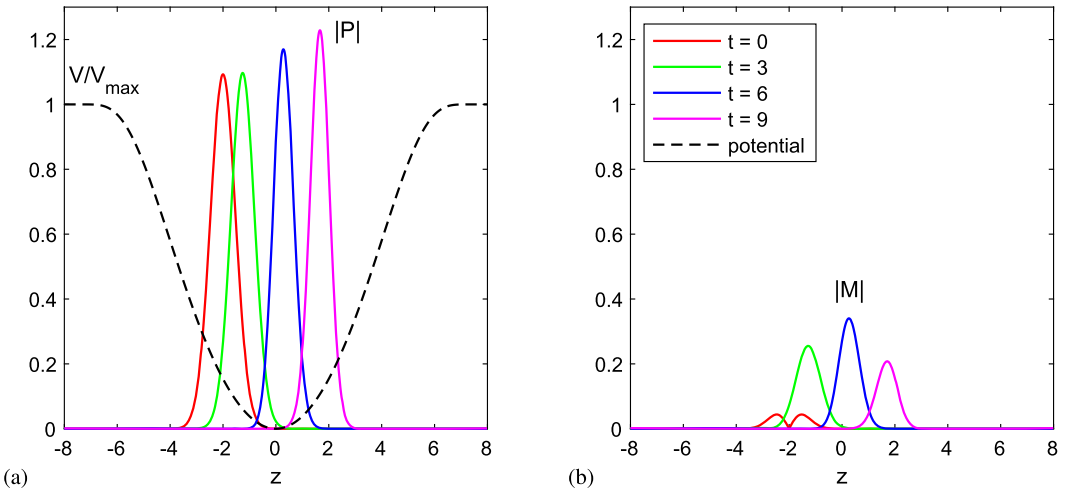


Fig. 1. The moduli of the components (a) $|P|$ and (b) $|M|$ of the wavefunction for the times $t \in \{0, 3, 6, 9\}$ shown on the same scale. The legend applies to both panels. The second component $|M|$ is largest at $t = 6$. The centre of the wavefunction is passing through the minumum of the potential so its momentum is largest. The scaled potential $V(z)/V_{\max}$ with $V_{\max} = V(8)$ is shown by the dashed lines in plot (a).

approximates the parabolic potential of a harmonic oscillator in the region around $z = 0$, as shown by the dashed curve in Fig. 1(a). The function V and all its derivatives are continuous L-periodic functions on the interval $[-L/2, L/2]$.

For initial conditions, we take a Gaussian distribution with standard deviation $\sigma = 1/3$ centred around $z_0 = -2$,

$$P(z,0) = (2\pi)^{-1/4} \left(\frac{16m^2\sigma}{1 + 16m^2\sigma^2} \right)^{1/2} \exp\left(-\frac{(z - z_0)^2}{4\sigma^2} \right). \tag{42}$$

We take $M = -i/(2m)\,\partial_z P$ in accordance with the adiabatic approximation (10) that leads to the Schrödinger equation. The normalisation constant for P in (42) is chosen so that $\int_{-\infty}^{\infty} |P|^2 + |M|^2 \, dz = 1$. These initial conditions for P and M are transformed into initial conditions for u and d using the unitary transformation (8).

Approximating the integral inner product (5) using the trapezoidal rule gives the natural discrete inner product

$$\langle\langle \boldsymbol{\Xi}, \boldsymbol{\Phi} \rangle\rangle = \frac{L}{N} \sum_j \Xi_j^\dagger \Phi_j, \tag{43}$$

between vectors of wavefunction values (indexed by j) on the spatial lattice. The trapezoidal rule integrates trigonometric interpolants exactly (see Sec. 7) so it is exponentially accurate for uniformly spaced points covering a domain with periodic boundary conditions. We can thus use the integral and discrete inner products interchangeably on sufficiently fine lattices. The continuously normalised initial conditions are also discretely normalised to machine precision in the numerical experiments.

Figure 1 shows the P and M components of the wavefunction for the initial conditions, and at the later times $t = 3$, $t = 6$, and $t = 9$ for a fourth-order accurate simulation on a lattice with 4096 points. The P component initially accelerates rightwards due to the gradient of the potential. Figure 2(a) shows time series of the ℓ^2-norm of the difference between solutions computed using this scheme, on lattices with 1024 to 8192 points, and a highly accurate reference solution computed using Fourier differentiation and exponentiation of an unsplit discrete Hamiltonian matrix (see Appendix for details). The differences reduce by a factor of 16 for each doubling of the number of lattice points, consistent with the expected fourth-order accuracy. The differences grow approximately linearly in time for $t \gtrsim 10$ due to the accumulation of phase errors. This is described in detail in Sec. 9.

Figure 2(b) shows the differences between the reference solution at $t = 99$ and the solutions computed using the second-order scheme, the fourth-order integer scheme, and the fourth-order scheme with halved time step, on lattices with $2^9 = 512$ to $2^{17} = 131072$ points. All solutions reach $t = 99$ after an integer number of time steps of length $6\Delta t$. The differences were computed in two different ways. The dots show differences computed using a Fourier interpolation of the reference solution onto the finer lattice used for the composition scheme. The open circles show differences computed by projecting the solution of the composition scheme onto the coarse lattice of 512 Fourier collocation points used to compute the reference solution. The close agreement between these two sets of differences confirms that the true solution is accurately described by the truncated Fourier series used to compute the reference solution.

The data in Fig. 2(b) confirm the expected orders of accuracies of the schemes, and also establish that the fourth-order scheme has a quantitatively smaller error for $N \geq 1024$. For sufficiently fine lattices, the scheme from Sec. 4.2 with a halved time step produces solutions with errors 16 times smaller than the scheme from Sec. 4.1 on the same lattice, as shown in Fig. 2(b). Most experiments were performed using 64-bit IEEE arithmetic. The highest resolution experiment with 131072 points was computed using 128-bit IEEE arithmetic, because the difference from the reference solution was dominated by accumulated round-off error in 64-bit arithmetic.

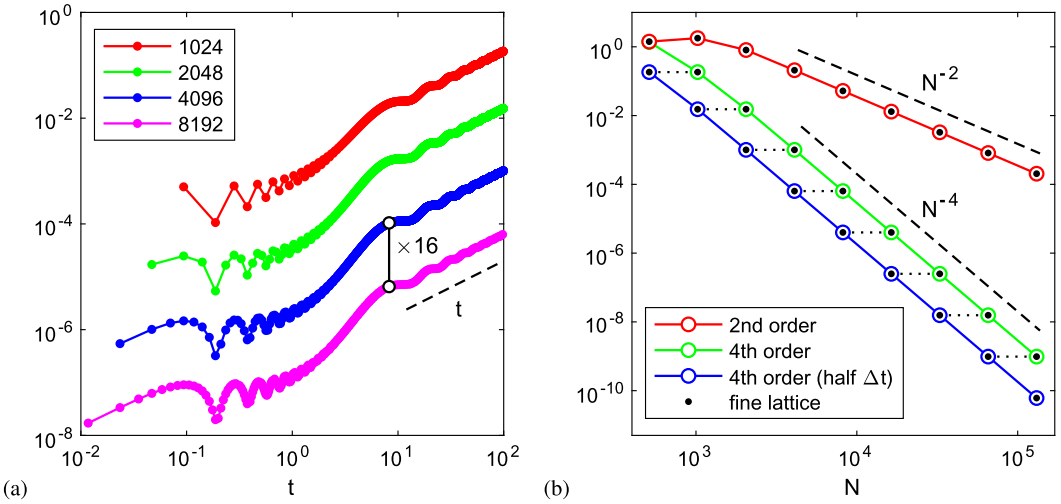


Fig. 2. Plot (a) shows the evolution over time of the ℓ^2 error relative to a highly accurate reference solution for numerical experiments on lattices with $N \in \{1024, 2048, 4096, 8192\}$ points. Plot (b) shows the ℓ^2 error at $t = 99$, and confirms the expected scalings with N of the second and fourth-order schemes. The fourth-order scheme with halved time step from Sec. 4.2 has an error indistinguishable from that of the original fourth-order scheme with twice as many points, as indicated by the horizontal dotted lines. The larger circles show the ℓ^2 error computed by projecting the solutions onto the coarse 512-point lattice of the reference solution (see Appendix). The smaller solid dots show the ℓ^2 error computed using a Fourier interpolation of the reference solution onto the finer N-point lattices.

6. Discrete energy conservation

Solutions of the abstract Schrödinger equation (3) satisfy an energy conservation property when the Hamiltonian operator has no explicit time dependence. We define the expectation of a general operator $\mathcal{A}$, not necessary Hermitian, for a normalised state by $\langle \mathcal{A} \rangle = \langle \Psi, \mathcal{A}\Psi \rangle$ with $\mathcal{A}$ acting on the right. If $\mathcal{A}$ has no explicit time dependence, its expectation evolves according to the generalised Ehrenfest theorem [45,46]

$$\frac{d}{dt} \langle \mathcal{A} \rangle = \langle \partial_t \Psi, \mathcal{A}\Psi \rangle + \langle \Psi, \mathcal{A} \partial_t \Psi \rangle = \langle -i\mathcal{H}\Psi, \mathcal{A}\Psi \rangle + \langle \Psi, \mathcal{A}(-i\mathcal{H})\Psi \rangle = -i \langle [\mathcal{A}, \mathcal{H}] \rangle . \tag{44}$$

This result follows from Schrödinger's equation $i\partial_t \Psi = \mathcal{H}\Psi$, its complex conjugate, and the Hermitian property of $\mathcal{H}$.

The Hamiltonian operator commutes with itself, so $\langle \mathcal{H} \rangle$ is invariant over time. All positive integer powers of $\mathcal{H}$, and hence $\exp(-i\Delta t \mathcal{H})$, also commute with $\mathcal{H}$, so the complex-valued expectation of the unitary evolution operator $\mathsf{U} = \exp(-i\Delta t \mathcal{H})$ is also invariant. Expanding for small Δt gives

$$\langle \exp(-i\Delta t \mathcal{H}) \rangle = 1 - i\Delta t \langle \mathcal{H} \rangle - \frac{1}{2!} \Delta t^2 \langle \mathcal{H}^2 \rangle + \frac{1}{3!} i \Delta t^3 \langle \mathcal{H}^3 \rangle + O(\Delta t^4). \tag{45}$$

The operators $\mathcal{H}^n$ are Hermitian for all positive integers n, so their expectations $\langle \mathcal{H}^n \rangle$ are all real. Taking the imaginary part of (45) thus gives the second-order accurate approximation

$$\langle \mathcal{H} \rangle = -\frac{1}{\Delta t} \operatorname{Im} \Big\{ \langle \exp(-i\Delta t \mathcal{H}) \rangle \Big\} + O(\Delta t^2). \tag{46}$$

This is a discrete analog of the continuous result

$$\mathcal{H} = i \frac{d}{dt} \bigg|_{t=0} \exp(-it\mathcal{H}), \tag{47}$$

with the derivative replaced by a finite difference between $t = \Delta t$ and $t = 0$.

Equation (46) extends straightforwardly to the discrete case, being based on the evolution operator for a finite time step Δt, not the Hamiltonian operator that generates infinitesimal evolution in continuous time. Given a unitary operator $\widetilde{\mathsf{U}}$ that generates evolution in discrete time steps of length Δt, so $\boldsymbol{\Psi}_{n+1} = \widetilde{\mathsf{U}} \boldsymbol{\Psi}_n$ with $\boldsymbol{\Psi}_n = \boldsymbol{\Psi}(n\Delta t)$ for $n \in \mathbb{N}$, the complex discrete expectation of $\widetilde{\mathsf{U}}$ is itself conserved [16]

$$\Big\langle\Big\langle \widetilde{\mathsf{U}} \Big\rangle\Big\rangle_{n+1} = \Big\langle\Big\langle \boldsymbol{\Psi}_{n+1}, \widetilde{\mathsf{U}} \boldsymbol{\Psi}_{n+1} \Big\rangle\Big\rangle = \Big\langle\Big\langle \widetilde{\mathsf{U}} \boldsymbol{\Psi}_n, \widetilde{\mathsf{U}} \widetilde{\mathsf{U}} \boldsymbol{\Psi}_n \Big\rangle\Big\rangle = \Big\langle\Big\langle \boldsymbol{\Psi}_n, \widetilde{\mathsf{U}}^\dagger \widetilde{\mathsf{U}} \widetilde{\mathsf{U}} \boldsymbol{\Psi}_n \Big\rangle\Big\rangle = \Big\langle\Big\langle \boldsymbol{\Psi}_n, \widetilde{\mathsf{U}} \boldsymbol{\Psi}_n \Big\rangle\Big\rangle = \Big\langle\Big\langle \widetilde{\mathsf{U}} \Big\rangle\Big\rangle_n . \tag{48}$$

The product $\widetilde{\mathsf{U}}^\dagger \widetilde{\mathsf{U}}$ is the identity operator because $\widetilde{\mathsf{U}}$ is unitary. Moreover, if $\widetilde{\mathsf{U}}$ approximates $\mathsf{U} = \exp(-i\Delta t \mathcal{H})$ with at least second-order accuracy, then

$$\langle \mathcal{H} \rangle = -\frac{1}{\Delta t} \operatorname{Im} \Big\langle\Big\langle \widetilde{\mathsf{U}} \Big\rangle\Big\rangle + O(\Delta t^2) = -\frac{1}{\Delta t} \operatorname{Im} \big\langle \langle \boldsymbol{\Psi}_n, \boldsymbol{\Psi}_{n+1} \rangle \big\rangle + O(\Delta t^2). \tag{49}$$

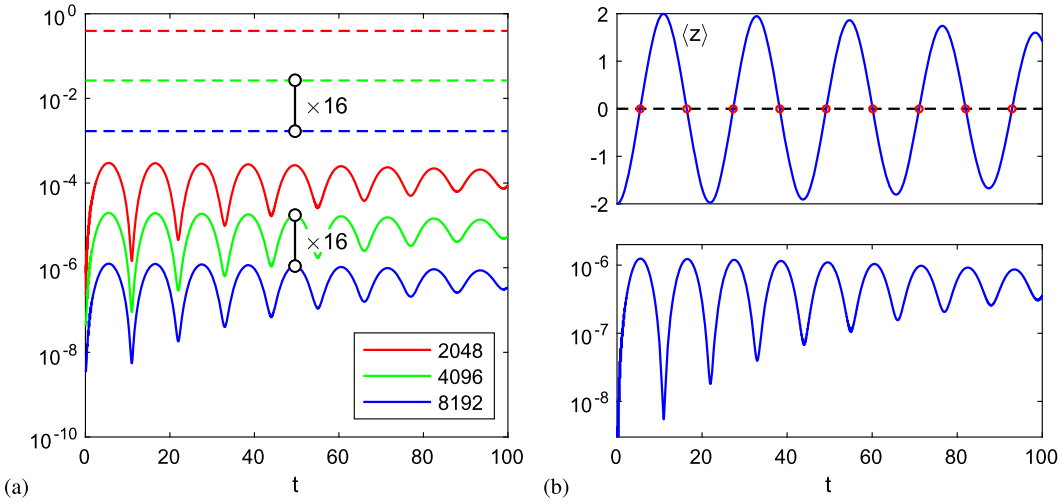


Fig. 3. Plot (a) shows time series of the absolute differences in the discrete invariant (55) (dashed lines) and a spectrally accurate approximation of $\langle \mathcal{H} \rangle$ (solid lines) from the exact value of $\langle \mathcal{H} \rangle$ for lattices with 2048, 4096, and 8192 points. The spectral approximation has a much lower error but oscillates over time. Both quantities exhibit fourth-order convergence towards the exact value with increasing N. Plot (b) shows that the error in the spectral approximation for 8192 points (bottom) is largest when the expected position $\langle z \rangle$ passes through zero (top, dots). The oscillations decrease in amplitude as the wavefunction broadens over time.

The discrete inner product $\langle\langle \cdots \rangle\rangle$ defined in (43) is an exponentially accurate approximation to the integral inner product $\langle \cdots \rangle$ defined in (5). They are thus interchangeable to this level of accuracy.

Equation (49) is a general result for any unitary evolution operator $\widetilde{\mathsf{U}}$. The algorithms in Secs. 3 and 4 achieve second-order and fourth-order accuracies efficiently using the transformed wavefunction $\overline{\boldsymbol{\Psi}} = \mathsf{C}^{-1/2}\boldsymbol{\Psi}$. The transformation from $\boldsymbol{\Psi}$ to $\overline{\boldsymbol{\Psi}}$ is unitary, and hence preserves inner products, so

$$\left\langle\left\langle \boldsymbol{\Psi}_n, \boldsymbol{\Psi}_{n+1} \right\rangle\right\rangle = \left\langle\left\langle \overline{\boldsymbol{\Psi}}_n, \overline{\boldsymbol{\Psi}}_{n+1} \right\rangle\right\rangle. \tag{50}$$

Expressing $\boldsymbol{\Psi}_{n+1}$ in terms of $\boldsymbol{\Psi}_n$, and $\overline{\boldsymbol{\Psi}}_{n+1}$ in terms of $\overline{\boldsymbol{\Psi}}_n$, gives

$$\left\langle\left\langle \boldsymbol{\Psi}_n, \mathsf{C}^{1/2}\mathsf{S}\mathsf{C}^{1/2}\boldsymbol{\Psi}_n \right\rangle\right\rangle = \left\langle\left\langle \overline{\boldsymbol{\Psi}}_n, \mathsf{S}\mathsf{C}\overline{\boldsymbol{\Psi}}_n \right\rangle\right\rangle. \tag{51}$$

The expectation of the evolution operator $\mathsf{C}^{1/2}\mathsf{S}\mathsf{C}^{1/2}$ for the $\boldsymbol{\Psi}_n$ state thus equals the expectation of the SC operator for the $\overline{\boldsymbol{\Psi}}_n$ state. The latter is easier to compute in a concrete implementation for the second-order composition scheme in Sec. 3, giving

$$\langle \mathcal{H} \rangle = -\frac{1}{\Delta t} \operatorname{Im} \left\langle\left\langle \overline{\boldsymbol{\Psi}}_n, \overline{\boldsymbol{\Psi}}_{n+1} \right\rangle\right\rangle + O(\Delta t^2). \tag{52}$$

Adapting the calculation in (48) shows that $\left\langle\left\langle \widetilde{\mathsf{U}}^2 \right\rangle\right\rangle$ is also conserved by discrete unitary evolution under $\widetilde{\mathsf{U}}$. Given a fourth-order accurate unitary approximation $\widetilde{\mathsf{U}}$ to $\exp(-i6\Delta t\mathcal{H})$, such as those constructed in Sec. 4, we can compare the imaginary parts of the expansions of

$$\widetilde{\mathsf{U}} = \exp(-i6\Delta t\mathcal{H}) + O(\Delta t^5) = 1 - 6i\Delta t\mathcal{H} - 18\Delta t^2\mathcal{H}^2 + 36i\Delta t^3\mathcal{H}^3 + 54\Delta t^4\mathcal{H}^4 + O(\Delta t^5), \tag{53}$$

and

$$\widetilde{\mathsf{U}}^2 = \exp(-i12\Delta t\mathcal{H}) + O(\Delta t^5) = 1 - 12i\Delta t\mathcal{H} - 72\Delta t^2\mathcal{H}^2 + 288i\Delta t^3\mathcal{H}^3 + 864\Delta t^4\mathcal{H}^4 + O(\Delta t^5). \tag{54}$$

Taking a linear combination of these two expressions to cancel the $\Delta t^3\mathcal{H}^3$ terms, equivalent to a Richardson extrapolation [47], gives the fourth-order accurate approximation

$$\langle \mathcal{H} \rangle = -\frac{1}{\Delta t} \operatorname{Im} \left\{ \frac{2}{9} \left\langle\left\langle \widetilde{\mathsf{U}} \right\rangle\right\rangle - \frac{1}{36} \left\langle\left\langle \widetilde{\mathsf{U}}^2 \right\rangle\right\rangle \right\} + O(\Delta t^4) = -\frac{1}{\Delta t} \operatorname{Im} \left\{ \frac{2}{9} \left\langle\left\langle \overline{\boldsymbol{\Psi}}_n, \overline{\boldsymbol{\Psi}}_{n+1} \right\rangle\right\rangle - \frac{1}{36} \left\langle\left\langle \overline{\boldsymbol{\Psi}}_n, \overline{\boldsymbol{\Psi}}_{n+2} \right\rangle\right\rangle \right\} + O(\Delta t^4). \tag{55}$$

Figure 3(a) shows a comparison between this invariant and an approximate evaluation of the expression for $\langle \mathcal{H} \rangle$ in (7) using a spectrally accurate approximation for the derivatives $\partial_z u$ and $\partial_z d$ at lattice points (as in the Appendix) and the trapezoidal rule to approximate the integral. Both quantities show fourth-order convergence towards the exact value of $\langle \mathcal{H} \rangle$ calculated for the initial conditions. The spectral approximation has a much smaller error, but it oscillates significantly over time, while the discrete invariant remains constant to machine precision. Figure 3(b) shows that the errors in the spectral approximation are largest when the envelope of the wavefunction passes through the centre of the potential, *i.e.* when the expected position $\langle z \rangle$ equals zero.

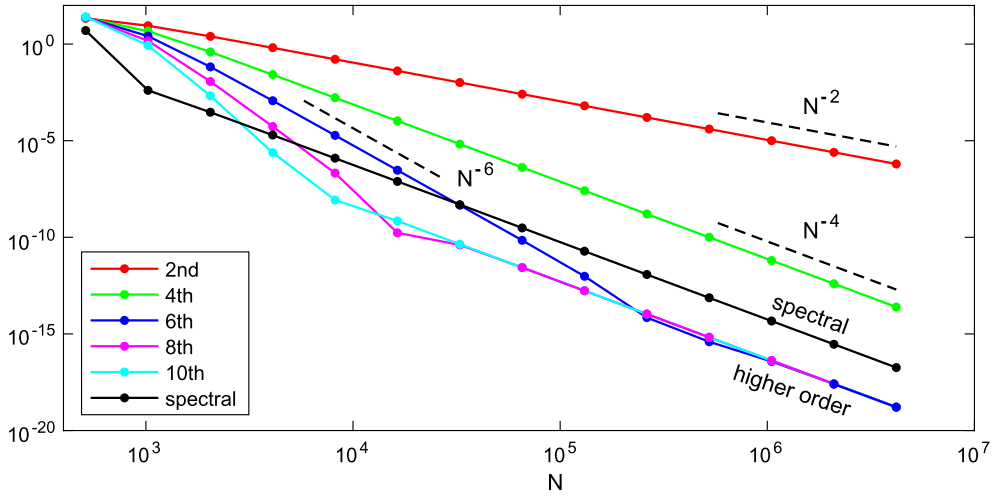


Fig. 4. Convergence of different approximations to $\langle \mathcal{H} \rangle$ based on linear combinations of $\langle \widetilde{\mathsf{U}}^n \rangle$ from (52), (55), (56) and a spectrally accurate approximation to $\langle \mathcal{H} \rangle$ using Fourier differentiation. All data are shown for time $t = 5.53125$. The data for the sixth, eight, and tenth-order invariants (collectively labelled "higher order") overlap for sufficiently large N.

Given a discrete evolution operator that approximates $\exp(-6i\Delta t\mathcal{H})$ to sufficient accuracy, we can construct further invariants with successively higher orders of accuracy. The first few are

$$\langle \mathcal{H} \rangle = -\frac{1}{\Delta t} \operatorname{Im} \left\{ \frac{1}{4} \left\langle\!\left\langle \overline{\boldsymbol{\Psi}}_n, \overline{\boldsymbol{\Psi}}_{n+1} \right\rangle\!\right\rangle - \frac{1}{20} \left\langle\!\left\langle \overline{\boldsymbol{\Psi}}_n, \overline{\boldsymbol{\Psi}}_{n+2} \right\rangle\!\right\rangle + \frac{1}{180} \left\langle\!\left\langle \overline{\boldsymbol{\Psi}}_n, \overline{\boldsymbol{\Psi}}_{n+3} \right\rangle\!\right\rangle \right\} + O(\Delta t^6), \tag{56a}$$

$$\langle \mathcal{H} \rangle = -\frac{1}{\Delta t} \operatorname{Im} \left\{ \frac{4}{15} \left\langle\!\left\langle \overline{\boldsymbol{\Psi}}_n, \overline{\boldsymbol{\Psi}}_{n+1} \right\rangle\!\right\rangle - \frac{1}{15} \left\langle\!\left\langle \overline{\boldsymbol{\Psi}}_n, \overline{\boldsymbol{\Psi}}_{n+2} \right\rangle\!\right\rangle + \frac{4}{315} \left\langle\!\left\langle \overline{\boldsymbol{\Psi}}_n, \overline{\boldsymbol{\Psi}}_{n+3} \right\rangle\!\right\rangle - \frac{1}{840} \left\langle\!\left\langle \overline{\boldsymbol{\Psi}}_n, \overline{\boldsymbol{\Psi}}_{n+4} \right\rangle\!\right\rangle \right\} + O(\Delta t^8), \tag{56b}$$

$$\langle \mathcal{H} \rangle = -\frac{1}{\Delta t} \operatorname{Im} \left\{ \frac{5}{18} \left\langle\!\left\langle \overline{\boldsymbol{\Psi}}_n, \overline{\boldsymbol{\Psi}}_{n+1} \right\rangle\!\right\rangle - \frac{5}{63} \left\langle\!\left\langle \overline{\boldsymbol{\Psi}}_n, \overline{\boldsymbol{\Psi}}_{n+2} \right\rangle\!\right\rangle + \frac{5}{252} \left\langle\!\left\langle \overline{\boldsymbol{\Psi}}_n, \overline{\boldsymbol{\Psi}}_{n+3} \right\rangle\!\right\rangle - \frac{5}{1512} \left\langle\!\left\langle \overline{\boldsymbol{\Psi}}_n, \overline{\boldsymbol{\Psi}}_{n+4} \right\rangle\!\right\rangle \right.$$
$$\left. + \frac{1}{3780} \left\langle\!\left\langle \overline{\boldsymbol{\Psi}}_n, \overline{\boldsymbol{\Psi}}_{n+5} \right\rangle\!\right\rangle \right\} + O(\Delta t^{10}). \tag{56c}$$

Figure 4 shows that these higher-order approximations to $\langle \mathcal{H} \rangle$ are quantitatively more accurate than the fourth-order accurate expression in (55), and even the spectral approximation, for a wide range of lattice sizes. However, they all ultimately reduce to fourth-order accuracy for sufficiently high resolutions because the underlying discrete evolution operator is itself only fourth-order accurate. All approximations were evaluated at the time $t = 5.53125$, an exact multiple of the time step that lies close to the first maximum of the oscillations visible in Fig. 3(a). All these computations were performed using 128-bit arithmetic, which has a precision of 113 bits, or approximately 34 decimal digits. The largest simulation uses $2^{22} = 4194304$ lattice points, and required 1449984 time steps to reach $t = 5.53125$, so over 6×10^{12} lattice point updates in total. By the end of the simulation, the solution was still normalised to within an error of 1.5×10^{-28}. The discrete energy invariants had changed by less than 3×10^{-27} from their initial values of approximately 19. Their relative errors were thus also less than 1.5×10^{-28}. This is consistent with the hypothesis that the accumulated effect of floating-point round-off errors can be modelled by a random walk, and hence grows in proportion to the square root of the number of operations [48].

7. Comparison with conventional fourth-order composition schemes

It is natural to ask how the accuracy of the nine-stage integer scheme compares with the more widely used fourth-order schemes with irrational coefficients. The latter are much more computationally expensive to implement for this problem, because the streaming operation cannot be reduced to a simple translation of wavefunction values between lattice points. A general streaming operation can instead be implemented using the discrete Fourier transform. The resulting scheme is then analogous to the popular split-step Fourier algorithm for the Schrödinger equation [38,49–51].

Given complex numbers $u_1, \ldots, u_N$ we define their discrete Fourier coefficients $\hat{u}_0, \ldots \hat{u}_{N-1}$ by

$$\hat{u}_k = \frac{1}{\sqrt{N}} \sum_{j=1}^{N} u_j \exp\left(-2\pi i \, \frac{k(j-1)}{N} \right), \tag{57}$$

for $k \in \{0, \ldots, N-1\}$. The corresponding inverse transformation is

$$u_j = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} \hat{u}_k \exp\left(2\pi i \, \frac{k(j-1)}{N} \right), \tag{58}$$

for $j \in \{1, \dots, N\}$. This transformation is unitary, which implies Parseval's theorem in the form

$$\sum_{j=1}^{N} |u_j|^2 = \sum_{k=0}^{N-1} |\hat{u}_k|^2. \tag{59}$$

Taking $u_1, \dots, u_N$ to be the values of a complex function u at the discrete points $z_j = (j-1)L/N$ in the interval $[0, L)$, we can use the Fourier coefficients $\hat{u}_k$ to construct the unique minimally oscillatory interpolating trigonometric polynomial [52,53]

$$\tilde{u}(z) = \hat{u}_0 + \sum_{1 \le k < N/2} \left\{ \hat{u}_k \exp\left(2\pi i \frac{kz}{L}\right) + \hat{u}_{N-k} \exp\left(-2\pi i \frac{kz}{L}\right) \right\} + \hat{u}_{N/2} \cos\left(\frac{\pi N z}{L}\right). \tag{60}$$

The last term proportional to $\hat{u}_{N/2}$, the Nyquist or zig-zag mode, is omitted if N is odd. This function $\tilde{u}$ interpolates the data, $\tilde{u}(z_j) = u_j$ for $j \in \{1, \dots, N\}$, and also minimises the oscillation $\int_0^L |\tilde{u}'(z)|^2 dz$. The latter condition removes the ambiguity by which each k in the exponents in (58) can be replaced by $k + Nn$ for any integer n without changing the values of $\tilde{u}(z_j)$, because $\exp(2n\pi i) = 1$. A streaming operation with an arbitrary displacement δz can be performed by computing the $\hat{u}_k$ values using (57), multiplying each $\hat{u}_k$ by $\exp(2\pi i k \delta z/L)$, and each $\hat{u}_{N-k}$ by $\exp(-2\pi i k \delta z/L)$, then applying the inverse transformation (58). This defines a unitary transformation of the vector of values $(u_1, \dots, u_N)^{\mathsf{T}}$ when N is odd.

We must also consider the Nyquist mode when N is even. Evaluating the last term in (60) at $z + \delta z$ involves

$$\cos\left(\frac{\pi N(z+\delta z)}{L}\right) = \cos\left(\frac{\pi N z}{L}\right) \cos\left(\frac{\pi N \delta z}{L}\right) - \sin\left(\frac{\pi N z}{L}\right) \sin\left(\frac{\pi N \delta z}{L}\right). \tag{61}$$

The second term proportional to $\sin(\pi N z/L)$ is not part of the minimally oscillatory trigonometric polynomial (60). It cannot be represented on the spatial lattice as $\sin(\pi N z_j/L) = 0$ for all $j \in \{1, \dots, N\}$. The least bad approach leaves the coefficient of the Nyquist mode unchanged when applying a streaming operation. This is consistent with the standard approach for computing a first derivative by differentiating (60) with respect to z, then omitting the $\sin(\pi N z/L)$ term that vanishes on the spatial lattice. Setting the coefficient of the Nyquist mode in the first derivative to zero ensures that the first derivative of a real even function is a real odd function, and vice versa [53]. Computing a first derivative via (57) and (58) then coincides with the result of applying the Fourier differentiation matrix (A.3) directly to $u_1, \dots, u_N$, as in the Appendix. However, setting the coefficient of the Nyquist mode to zero would make a general streaming operation non-unitary. For the special case when the displacement δz is an integer multiple of $\Delta z = L/N$ one can multiply $\hat{u}_{N/2}$ by $\cos(\pi N \delta z/L) = (-1)^{\delta z/\Delta z}$. The overall operation then exactly coincides with the result of applying the explicit streaming operator S directly to $u_1, \dots, u_N$ as in (16).

To produce a fair comparison, the base time step for each of the fourth-order schemes from Sec. 4 was scaled so that each computation reached $t = 99$ after a similar total number of stages. For example, the three-stage scheme with a base time step of $2\Delta t$ advances the solution by $2\Delta t$ in 3 stages, and hence by $6\Delta t$ in 9 stages. Similarly, the five-stage schemes with a base time step of $3\Delta t$ advance the solution by $3\Delta t$ in 5 stages, and hence by $6\Delta t$ in 10 stages. This comparison deliberately ignores the much greater computational cost of implementing streaming with non-integer displacements via pairs of discrete Fourier transforms. It ignores the minor optimisation available from combining the two adjacent applications of the C operator from the end of one stage and the start of the next stage. It also ignores the possibility of achieving higher accuracy by breaking the link between Δt and Δz that is needed to implement the nine-stage integer scheme using local operations. The spatial discretisation is always exponentially accurate, since the nine-stage integer scheme is equivalent to applying the Fourier streaming algorithm for integer multiples of the lattice spacing. The overall error is thus dominated by the $O(\Delta t^4)$ temporal splitting error.

Figure 5(a) shows the evolution of the ℓ^2 norms of the differences from the reference solutions for the four different fourth-order schemes on a lattice with 8192 points. The nine-stage scheme has the largest initial error, but the smallest error after sufficiently long times, for $t \gtrsim 3$ to be concrete. The five-stage McLachlan scheme [31], despite being optimized for classical Hamiltonian systems via a backwards error analysis, produces larger errors in the wavefunction than the five-stage Neri scheme [32]. Figure 5(b) shows the error in the solution at $t = 99$ for the four different schemes. All four schemes show the expected fourth-order convergence as the number of lattice points increases. As before, the two approaches to computing differences from a spectrally accurate reference solution on a coarse 512-point lattice yield visually indistinguishable results. These plots show convergence data for the nine-stage integer scheme using the same Fourier streaming algorithm as the other schemes, rather than the more efficient S operation used previously. The two different implementations of the nine-stage integer scheme produce visually indistinguishable results when plotted.

8. Analysis of plane wave solutions

To investigate the unexpectedly high accuracy of the nine-stage integer scheme over long times, we consider plane wave solutions proportional to $\exp(ikz)$ for a free particle with no potential. The Hamiltonian in (4) for the one-dimensional Dirac equation, a matrix of differential operators, then reduces to the algebraic 2×2 matrix Hamiltonian

$$\mathsf{H} = \begin{pmatrix} k & im \\ -im & -k \end{pmatrix} = k\boldsymbol{\sigma}_z - m\boldsymbol{\sigma}_y. \tag{62}$$

The matrices $\boldsymbol{\sigma}_y$ and $\boldsymbol{\sigma}_z$ are two of the three Pauli matrices:

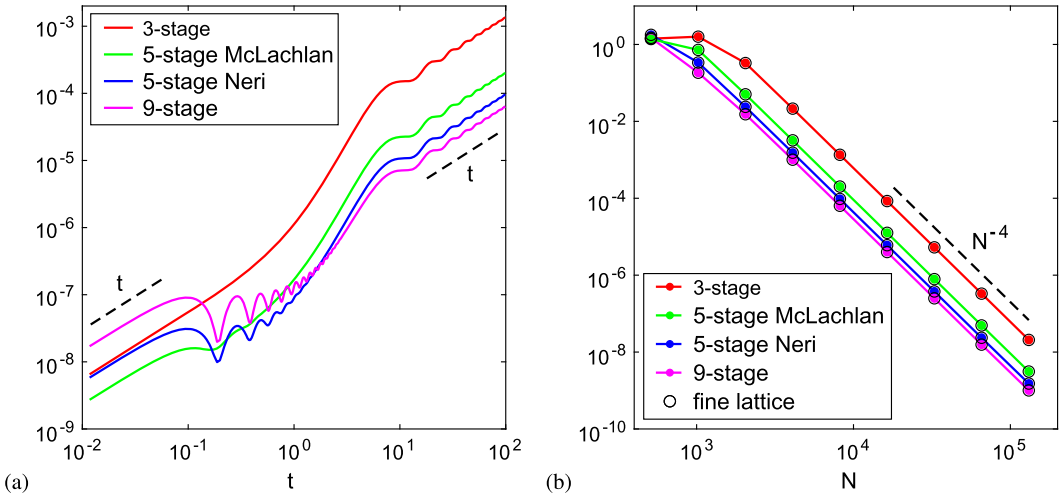


Fig. 5. Accuracy of four different fourth-order composition schemes using 3, 5, and 9 stages. The base time step has been scaled so that all four schemes advance the solution by $6\Delta t$ using 9 or 10 stages in total. Plot (a) shows the evolution of the errors over time for $N = 8192$. The errors grow linearly with time at both early and late times (see Sec. 9). Plot (b) shows the scaling with N of the error at $t = 99$, as computed both by projection onto the coarse 512-point lattice of the reference solution (filled dots) and by Fourier interpolation of the reference solution onto the finer lattice (black open circles).

$$\boldsymbol{\sigma}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \boldsymbol{\sigma}_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \boldsymbol{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{63}$$

Each matrix is both unitary and Hermitian with trace zero. They satisfy the relations of the Pauli algebra [46]

$$\boldsymbol{\sigma}_i \boldsymbol{\sigma}_j = \mathsf{I}\,\delta_{ij} + i\,\epsilon_{ijk}\boldsymbol{\sigma}_k, \tag{64}$$

where I is the 2×2 identity matrix, and ϵ_{ijk} is the alternating Levi-Civita tensor. The Pauli matrices and the identity matrix $\boldsymbol{\sigma}_0 = \mathsf{I}$ form a basis for the space of 2×2 Hermitian matrices, treated as a vector space with real coefficients. This basis is orthogonal with respect to the Frobenius inner product defined by $\langle\langle \mathsf{A}, \mathsf{B} \rangle\rangle = \mathrm{Tr}\,(\mathsf{A}^\dagger \mathsf{B})$ for matrices A and B. The Pauli algebra relations imply $\mathsf{H}^2 = \Omega^2 \mathsf{I}$ with $\Omega = \sqrt{m^2 + k^2}$, so the exponential of the matrix H is

$$\exp(-it\mathsf{H}) = \cos(\Omega t)\,\mathsf{I} - (i/\Omega)\sin(\Omega t)\,\mathsf{H}. \tag{65}$$

The discrete streaming step also becomes an algebraic 2×2 matrix when applied to plane wave solutions proportional to $\exp(ikz)$,

$$\mathsf{S} = \begin{pmatrix} \exp(-ik\Delta t) & 0 \\ 0 & \exp(ik\Delta t) \end{pmatrix}. \tag{66}$$

The second-order scheme thus evolves the wavefunction over a time step Δt by the matrix

$$\widetilde{\mathsf{U}}_2 = \mathsf{C}^{1/2}\mathsf{S}\mathsf{C}^{1/2} = \begin{pmatrix} \cos(m\Delta t)\cos(k\Delta t) - i\sin(k\Delta t) & \sin(m\Delta t)\cos(k\Delta t) \\ -\sin(m\Delta t)\cos(k\Delta t) & \cos(m\Delta t)\cos(k\Delta t) + i\sin(k\Delta t) \end{pmatrix}. \tag{67}$$

As before, the tilde accent on $\widetilde{\mathsf{U}}_2$ indicates a discrete approximation to the exact evolution operator $\mathsf{U} = \exp(-i\Delta t\mathsf{H})$. Comparing (67) with (65) establishes that the local truncation error for the second-order scheme is indeed $O(\Delta t^3)$:

$$\widetilde{\mathsf{U}}_2 - \exp(-i\Delta t\mathsf{H}) = \frac{1}{6}\Delta t^3 \begin{pmatrix} -ikm^2 & -2mk^2 \\ 2mk^2 & ikm^2 \end{pmatrix} + O(\Delta t^4). \tag{68}$$

Replacing Δt with $-2\Delta t$ in (67) gives the double-length reversed-time operator required in the nine-stage integer scheme,

$$\mathsf{C}^{-1}\mathsf{S}^{-2}\mathsf{C}^{-1} = \begin{pmatrix} \cos(2m\Delta t)\cos(2k\Delta t) + i\sin(2k\Delta t) & -\sin(2m\Delta t)\cos(2k\Delta t) \\ \sin(2m\Delta t)\cos(2k\Delta t) & \cos(2m\Delta t)\cos(2k\Delta t) - i\sin(2k\Delta t) \end{pmatrix}. \tag{69}$$

Combining all these expressions with the aid of a symbolic manipulator establishes that the local truncation error of the nine-stage integer scheme $\widetilde{\mathsf{U}}_4$ defined in (36) is indeed $O(\Delta t^5)$:

$$\widetilde{\mathsf{U}}_4 - \exp(-6i\Delta t\mathsf{H}) = \frac{1}{5}\Delta t^5 \begin{pmatrix} ikm^2(48k^2 + 29m^2) & 8mk^2(7k^2 + 4m^2) \\ -8mk^2(7k^2 + 4m^2) & -ikm^2(48k^2 + 29m^2) \end{pmatrix} + O(\Delta t^6). \tag{70}$$

Each component of the leading-order truncation error is a homogeneous polynomial of degree 5 in k and m. More generally, the truncation error for any of the fourth-order composition schemes studied in Sec. 7 can be written using the Pauli matrices as

Table 1
The coefficients in the expressions for the truncation errors of four different fourth-order composition schemes.

scheme	p	q	r	s	λ	μ	$\lvert\mu/\lambda\rvert$
three-stage	-41.06	0.64	-7.20	13.34	-501.86	995	2
five-stage Neri	1.02	2.02	3.43	2.92	-35.71	3248	91
five-stage McLachlan	-4.36	0.97	0.68	2.87	-74.81	1562	21
nine-stage integer	9.60	5.80	11.20	6.40	-23.71	9339	394

$$\widetilde{\mathsf{U}} - \exp(-6i\Delta t\mathsf{H}) = i\Delta t^5\left(\alpha\boldsymbol{\sigma}_z + \beta\boldsymbol{\sigma}_y\right) + O(\Delta t^6), \tag{71}$$

with real coefficients

$$\alpha = km^2(pk^2 + qm^2), \quad \beta = k^2m(rk^2 + sm^2). \tag{72}$$

Table 1 lists the values of the four constant coefficients p, q, r, s for the four different composition schemes. The base time steps for each scheme have been normalised as in Sec. 7 so that each schemes advances the solution by $6\Delta t$ using 9 or 10 applications of the building block $\widetilde{\mathsf{U}}_2$. Each coefficient for the error in the nine-stage scheme is larger than the corresponding coefficients for both five-stage schemes, which does not explain the smaller errors after long times shown in Fig. 5.

9. Long-time behaviour

The explanation of the long-time behaviour requires further exploration of the leading-order truncation errors for plane wave solutions in the different schemes. Each of the composition schemes provides a unitary approximation $\widetilde{\mathsf{U}}$ to the exact evolution operator $\mathsf{U} = \exp(-6\Delta t i\mathsf{H})$. We can thus rewrite the additive expression (71) for the leading-order truncation error as the multiplicative expression

$$\widetilde{\mathsf{U}} = \mathsf{U}\mathsf{E} + O(\Delta t^6), \text{ where } \mathsf{E} = \exp\left\{i\Delta t^5\left(\alpha\boldsymbol{\sigma}_z + \beta\boldsymbol{\sigma}_y\right)\right\}. \tag{73}$$

The real coefficients α and β depend on k and m, and on the four constants p, q, r, s given in Table 1 for the different schemes. The matrix E is unitary because α and β are real numbers, and $\boldsymbol{\sigma}_y$ and $\boldsymbol{\sigma}_z$ are Hermitian matrices. We can introduce a second matrix $\mathsf{H}^\perp = m\boldsymbol{\sigma}_z + k\boldsymbol{\sigma}_y$ that is orthogonal to $\mathsf{H} = k\boldsymbol{\sigma}_z - m\boldsymbol{\sigma}_y$ in the Frobenius inner product. The leading-order error in the evolution over one time step can then be written as

$$\mathsf{E} = \exp\left\{i\Delta t^5\left(\lambda\mathsf{H} + \mu\mathsf{H}^\perp\right)\right\}, \text{ where } \lambda = \frac{\alpha k - \beta m}{k^2 + m^2}, \quad \mu = \frac{\alpha m + \beta k}{k^2 + m^2}. \tag{74}$$

The λ and μ coefficients for the four different fourth-order composition schemes are given in Table 1.

It is useful to introduce the two matrices corresponding to the H and $\mathsf{H}^\perp$ terms separately,

$$\mathsf{E}_\lambda = \exp(i\Delta t^5\lambda\mathsf{H}), \quad \mathsf{E}_\mu = \exp(i\Delta t^5\mu\mathsf{H}^\perp). \tag{75}$$

Suppose first that $\mu = 0$. The remaining matrix E_λ commutes with $\mathsf{U} = \exp(-6\Delta t i\mathsf{H})$, so the discrete evolution operator for n time steps is

$$(\mathsf{U}\mathsf{E}_\lambda)^n = \mathsf{U}^n\mathsf{E}_\lambda^n = \exp\left\{-6in\Delta t\left(1 - (\lambda/6)\Delta t^4\right)\mathsf{H}\right\}. \tag{76}$$

This error thus adjusts the rate at which the solution evolves in time by a factor of $1 - (\lambda/6)\Delta t^4$. The numerical solution evolves a little slower than the exact solution if $\lambda > 0$, and a little faster if $\lambda < 0$. The difference between the numerical and exact solution thus grows linearly with time at first, as shown in Figs. 2(a) and 5(a). This behaviour persists until the number of time steps n becomes comparable to $(2\pi/|\lambda|)\Delta t^{-5}$. The underlying periodicity of the exact and numerical solutions then become apparent. It is not possible to absorb this error by rescaling time, because a different rescaling would be needed for each different wavenumber k.

To explore the general case with $\mu \neq 0$ we introduce the scaled matrices

$$\hat{\mathsf{H}}^\perp = \mathsf{H}^\perp/\Omega, \quad \hat{\mathsf{H}} = \mathsf{H}/\Omega, \tag{77}$$

where $\Omega = \sqrt{k^2 + m^2}$ as before, and write

$$\mathsf{E} = \exp\left\{i\epsilon\left(\hat{\mathsf{H}}\cos\Theta + \hat{\mathsf{H}}^\perp\sin\Theta\right)\right\} \tag{78}$$

with scalar coefficients $\epsilon = \Delta t^5\Omega\sqrt{\lambda^2 + \mu^2}$ and $\Theta = \tan^{-1}(\mu/\lambda)$. The matrices $\mathsf{h}_x = \boldsymbol{\sigma}_x$, $\mathsf{h}_y = \hat{\mathsf{H}}^\perp$, $\mathsf{h}_z = \hat{\mathsf{H}}$ are each unitary and Hermitian with trace zero. They satisfy the relations

$$\mathsf{h}_i\,\mathsf{h}_j = \mathsf{I}\,\delta_{ij} + i\,\epsilon_{ijk}\mathsf{h}_k, \tag{79}$$

so they provide an alternative representation of the Pauli algebra.

The discrete evolution operator for one time step can now be rewritten as

$$\begin{aligned}
\mathsf{UE} &= \exp\Big(-i6\Omega\Delta t\hat{\mathsf{H}}\Big)\exp\Big\{i\epsilon\Big(\hat{\mathsf{H}}\cos(\Theta)+\hat{\mathsf{H}}^{\perp}\sin(\Theta)\Big)\Big\},\\
&= \Big(\mathsf{I}\cos(6\Omega\Delta t)-i\hat{\mathsf{H}}\sin(6\Omega\Delta t)\Big)\Big\{\mathsf{I}\cos(\epsilon)+i\sin(\epsilon)\Big(\hat{\mathsf{H}}\cos(\Theta)+\hat{\mathsf{H}}^{\perp}\sin(\Theta)\Big)\Big\},\\
&= \Big(\cos(6\Omega\Delta t)\cos(\epsilon)+\sin(6\Omega\Delta t)\sin(\epsilon)\cos(\Theta)\Big)\mathsf{I}+i\Big(\cos(6\Omega\Delta t)\sin(\epsilon)\cos(\Theta)-\sin(6\Omega\Delta t)\cos(\epsilon)\Big)\hat{\mathsf{H}}\\
&\qquad\qquad +i\sin(\epsilon)\sin(\Theta)\Big(\hat{\mathsf{H}}^{\perp}\cos(6\Omega\Delta t)-\boldsymbol{\sigma}_x\sin(6\Omega\Delta t)\Big),\\
&= \mathsf{I}\cos(R)-i\sin(R)\Big\{\hat{\mathsf{H}}\cos(\theta)+\sin(\theta)\Big(\hat{\mathsf{H}}^{\perp}\cos(6\Omega\Delta t)-\boldsymbol{\sigma}_x\sin(6\Omega\Delta t)\Big)\Big\},\\
&= \exp\Big\{-iR\Big(\hat{\mathsf{H}}\cos(\theta)+\sin(\theta)\Big(\hat{\mathsf{H}}^{\perp}\cos(6\Omega\Delta t)-\boldsymbol{\sigma}_x\sin(6\Omega\Delta t)\Big)\Big)\Big\}. \qquad (80)
\end{aligned}$$

This calculation is equivalent to using the relations (79) to simplify and sum the series of nested commutators in the Baker–Campbell–Hausdorff formula for a product of exponentials of the non-commuting operators $\hat{\mathsf{H}}$ and $\hat{\mathsf{H}}^{\perp}$.

The coefficients R and θ are determined by

$$\cos(R) = \cos(6\Omega\Delta t)\cos(\epsilon)+\sin(6\Omega\Delta t)\sin(\epsilon)\cos(\Theta), \qquad (81a)$$

$$\tan(\theta) = -\frac{\sin(\epsilon)\sin(\Theta)}{\sin(6\Omega\Delta t)\cos(\epsilon)-\cos(6\Omega\Delta t)\sin(\epsilon)\cos(\Theta)}. \qquad (81b)$$

In the special case when $\mu = 0$, and hence $\Theta = 0$, the solution is $R = 6\Omega\Delta t - \epsilon$ and $\theta = 0$. We thus recover the previous result (76) when U and E commute. Expanding the solutions (81) for $\epsilon \ll 1$ gives

$$R = 6\Omega\Delta t - \epsilon\cos(\Theta)+\frac{1}{2}\epsilon^2\sin^2(\Theta)\cot(6\Omega\Delta t)+O(\epsilon^3), \qquad (82a)$$

$$\theta = -\epsilon\frac{\sin(\Theta)}{\sin(6\Omega\Delta t)}-\epsilon^2\frac{\cos(6\Omega\Delta t)\sin(\Theta)\cos(\Theta)}{\sin(6\Omega\Delta t)^2}+O(\epsilon^3). \qquad (82b)$$

Reverting to the original coefficients λ and μ gives

$$R = 6\Omega\Delta t - \Omega\lambda\Delta t^5+O(\Delta t^{10}), \quad \theta = -\frac{\Omega}{\sin(6\Omega\Delta t)}\mu\Delta t^5+O(\Delta t^{10}). \qquad (83)$$

From (80) we can easily compute the discrete evolution operator for n time steps,

$$\begin{aligned}
\widetilde{\mathsf{U}}^n = (\mathsf{UE})^n &= \exp\Big\{-inR\Big(\hat{\mathsf{H}}\cos(\theta)+\sin(\theta)\Big(\hat{\mathsf{H}}^{\perp}\cos(6\Omega\Delta t)-\boldsymbol{\sigma}_x\sin(6\Omega\Delta t)\Big)\Big)\Big\},\\
&= \cos(nR)\,\mathsf{I}-i\sin(nR)\Big\{\hat{\mathsf{H}}\cos(\theta)+\Big(\hat{\mathsf{H}}^{\perp}\cos(6\Omega\Delta t)-\boldsymbol{\sigma}_x\sin(6\Omega\Delta t)\Big)\sin(\theta)\Big\}, \qquad (84)
\end{aligned}$$

and compare it with the exact evolution operator

$$\mathsf{U}^n = \exp(-6in\Omega\Delta t\hat{\mathsf{H}}) = \cos(6n\Omega\Delta t)\,\mathsf{I}-i\sin(6n\Omega\Delta t)\,\hat{\mathsf{H}}. \qquad (85)$$

This comparison shows that the leading-order truncation error in R only depends upon λ, and the leading-order truncation error in θ only depends upon μ. The error due to λ changes R from $6\Omega\Delta t$, and hence changes the rate at which the solution evolves in time. This is consistent with the earlier calculation that neglected the contribution from μ entirely. The contribution from μ only appears in θ. Moreover, since $\cos\theta = 1 - O(\epsilon^2)$, the coefficient of $\hat{\mathsf{H}}$ in (84) is unchanged by the contribution from μ at leading order. The contribution from μ only appears at leading order through the error term proportional to $\hat{\mathsf{H}}^{\perp}\cos(6\Omega\Delta t)-\boldsymbol{\sigma}_x\sin(6\Omega\Delta t)$.

We can use the operator norm of the difference between $\widetilde{\mathsf{U}}^n$ and U^n to bound the ℓ^2 norm of the error in the solution for any normalised initial conditions. The operator norm of a matrix A induced by the ℓ^2 norm is defined by $||\mathsf{A}||_{\text{op}} = \sup\{||\mathsf{A}\,\mathbf{x}||_2 : ||\mathbf{x}||_2 = 1\}$. This norm is equal to the square root of the largest eigenvalue of the Hermitian matrix $\mathsf{A}^{\dagger}\mathsf{A}$. A unitary matrix thus has unit operator norm. Figure 6 shows the growth of $||\mathsf{U}^n - \widetilde{\mathsf{U}}^n||_{\text{op}}$ over time for a plane wave solution with wavenumber $k = \pi/8$, the longest wave in the previous simulation domain $[-8, 8]$, for a particle of mass $m = 16$ using the nine-stage integer scheme with time step $\Delta t = 1/512$. This corresponds to a lattice with 8192 points. Figure 6 also shows the same information with UE replaced by UE_λ and UE_μ. For these parameter values, $|\mu|$ is much larger than $|\lambda|$, as shown in Table 1. The error after the first time step is thus closely approximated by $||\mathsf{E}_\mu - \mathsf{I}||_{\text{op}} = |\mu|\,\Omega\,\Delta t^5 + O(\Delta t^{10})$. The error grows over the first few time steps, but remains close to the contribution from E_μ alone for the first few hundred time steps. This contribution is oscillatory, and remains bounded at long times. After 2000 time steps, the total error becomes dominated instead by the contribution from $||\mathsf{E}_\lambda - \mathsf{I}||_{\text{op}} = n|\lambda|\,\Omega\,\Delta t^5 + O(\Delta t^{10})$. This is a secularly growing or "resonant" contribution, that arises from the re-appearance of the H operator in the error term E.

This behaviour is reminiscent of seeking a series solution to a weakly nonlinear oscillator. The first correction to the leading order solution grows secularly in time if its evolution involves the same linear operator that governs the leading order solution. The method of multiple scales removes these secular terms via a more flexible series solution that expands the time coordinate as well as the dependent variable [47].

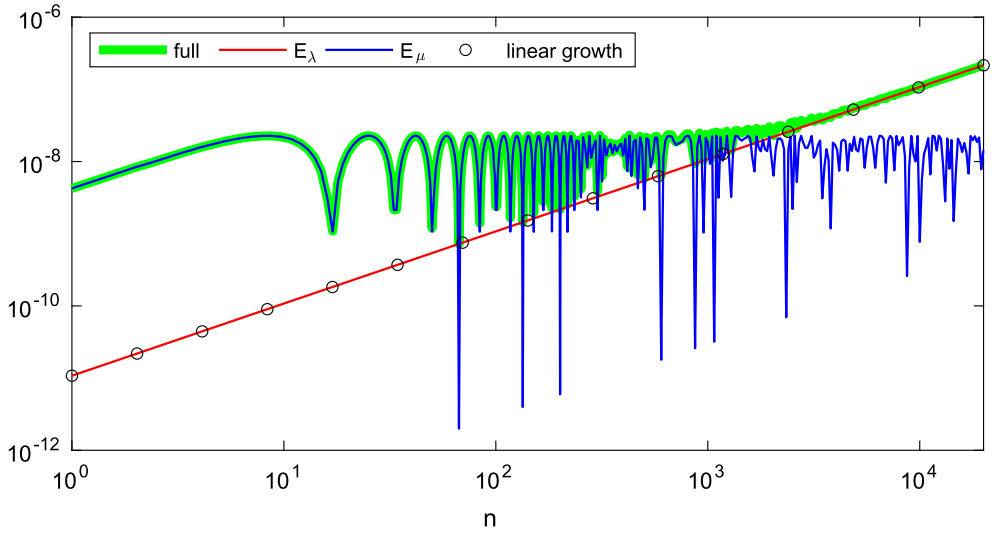


Fig. 6. The behaviour of $||U^n - \widetilde{U}^n||_{op}$ versus n for the nine-stage integer scheme. For small n the norm is dominated by the oscillatory contribution from E_μ. For larger $n \gtrsim 2000$ the norm is dominated by the contribution from E_λ that grows linearly like $n|\lambda|\,\Omega\,\Delta t^5$ (open circles).

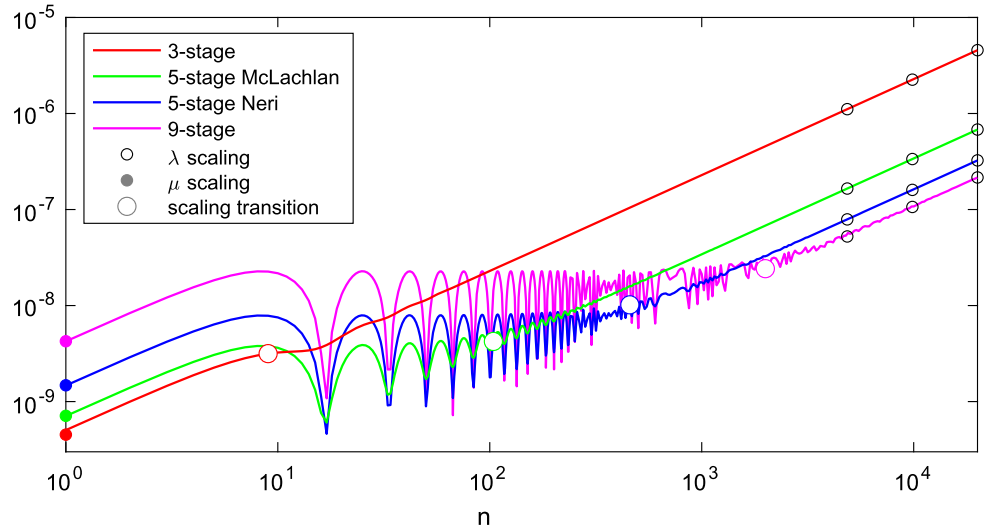


Fig. 7. The behaviour of $||\widetilde{U}^n - U^n||_{op}$ versus n for four different schemes. The norm at $n = 1$ is consistent with $|\mu|\,\Omega\,\Delta t^5$, shown by the filled circles on the left axis. For large n the norm grows like $n|\lambda|\,\Omega\,\Delta t^5$, shown by the small black open circles on the right. The large open circles mark the transition points $n_\star = 5|\mu/\lambda|$ after which the large-n behaviour becomes dominant on this logarithmic plot.

Figure 7 shows that the same qualitative description holds for all four different composition schemes, for which λ and μ values are given in Table 1. In each case, the error after the first time step is close to $|\mu|\,\Omega\,\Delta t^5$, and the error at later times grows linearly in good agreement with $n|\lambda|\,\Omega\,\Delta t^5$. Omitting the common factors of $\Omega\,\Delta t^5$, the two terms in the function $\mu + \lambda n$ have equal magnitude when $n = |\mu/\lambda|$. The two functions $\log(\mu + \lambda n)$ and $\log(\lambda n)$ differ by less than $\log(1 + 1/5) < 0.2$ when $n \geq n_\star = 5|\mu/\lambda|$. The large open circles in Fig. 7 show these $n_\star$ points for the four different schemes. They are all in reasonable agreement with the location of the transition from oscillatory behaviour for small n to monotonically growing behaviour at large n, and also with this transition being at $n_\star \approx 2000$ in Fig. 6.

10. Conclusion

The discrete time quantum walk describes the evolution of a two-component wavefunction on a one-dimensional lattice of uniformly spaces points through a sequence of unitary operations that alternate deterministic translations of the values of wavefunction components to adjacent lattice points ("streaming") with applications of a Hadamard-like operation ("coin tossing") to mix the two components of the wavefunction at each lattice site. This algorithm is equivalent to a quantum cellular automaton. It can be derived as a first-order accurate Lie–Trotter splitting for discretising the one-dimensional Dirac equation. The accuracy can be raised to

second order by pre-processing the initial conditions and post-processing the subsequent output. This is equivalent to replacing the Lie–Trotter splitting SC with the symmetric Strang splitting $\mathsf{C}^{1/2}\mathsf{S}\mathsf{C}^{1/2}$.

We have explored composition schemes that use the symmetric Strang splitting as a building block to achieve fourth-order accuracy. The "symmetric symmetric" composition schemes are time-reversible. This eliminates all even powers of Δt from the local truncation error, leaving just two determining equations for the coefficients in the scheme. One normalises the length of the time step. The other cancels the $O(\Delta t^3)$ truncation errors of the individual stages to achieve fourth-order accuracy. The most popular schemes of this form use three or five stages with lengths related by irrational coefficients. The necessary translations by irrational fractions of a lattice spacing cannot be implemented by local shift operations, but only globally using discrete Fourier transforms, as in the "split-step Fourier" schemes for the Schrödinger equation [49–51]. In contrast, we have described the properties of a nine-stage scheme using integer coefficients implemented using purely local operations. This scheme shows the expected fourth-order convergence on sufficiently fine lattices, and is quantitatively more accurate than the second-order scheme for a typical problem discretised with 1024 points. Due to the restriction to integer coefficients, it has a substantially larger local truncation error than the more widely used fourth-order schemes with irrational coefficients. However, these other schemes all produce larger errors after sufficiently long times. We have shown that the source of the truncation error can be decomposed into a resonant part that grows linearly with time, and a non-resonant part that produces bounded oscillations. The resonant part of the truncation error is smallest for the nine-stage scheme.

Any numerical scheme that generates evolution through a time-independent unitary operation $\widetilde{\mathsf{U}}$ conserves the complex expectation $\langle\widetilde{\mathsf{U}}\rangle$ of the evolution operator [16]. When the scheme is itself second-order accurate, the imaginary part of this discrete invariant is a second-order accurate approximation to the expectation $\langle\mathcal{H}\rangle$ of the continuous Hamiltonian operator. By taking suitable linear combinations of the expectations of $\langle\widetilde{\mathsf{U}}^2\rangle$, $\langle\widetilde{\mathsf{U}}^3\rangle$ and so on, each of which is individually conserved, we have constructed discrete invariants that are fourth, sixth, and up to tenth-order accurate approximations to $\langle\mathcal{H}\rangle$. These are substantially more accurate for certain finite lattice sizes, even when the evolution itself is only fourth-order accurate. For example, Fig. 4 shows that the sixth-order approximation shows sixth-order convergence behaviour until $N = 524288$. The error is then comparable to machine prevision in 64-bit arithmetic.

Finally, the extension to fourth-order accuracy relies on the symmetry and reversibility of the building block $\mathsf{C}^{1/2}\mathsf{S}\mathsf{C}^{1/2}$. Symmetric and reversible compositions of three or more operators, such as $\mathsf{A}^{1/2}\mathsf{B}^{1/2}\mathsf{C}\mathsf{B}^{1/2}\mathsf{A}^{1/2}$, can be used as building blocks to construct fourth-order accurate composition schemes using the approach described in Sec. 4. In principle this applies for any number of operators. The nine-stage integer composition scheme has been confirmed by explicit calculation to have fourth-order accuracy for building blocks formed from up to six non-commuting operators. The expansion in Δt of the composite operator defined by (35) using $\mathsf{A}^{1/2} = \exp(\frac{1}{2}\Delta t\mathcal{A})$, and similarly for the other operators, coincides with the expansion of $\exp(6\Delta t(\mathcal{A}+\mathcal{B}+\mathcal{C}+\mathcal{D}+\mathcal{E}+\mathcal{F}))$ up to and including the $6^4 = 1296$ terms of $O(\Delta t^4)$ arising from the multinomial expansion of $(\mathcal{A}+\mathcal{B}+\mathcal{C}+\mathcal{D}+\mathcal{E}+\mathcal{F})^4$. The Maple Physics package was used to automate the manipulation of series with non-commuting operators as coefficients [54]. The resulting schemes can be used to solve the Dirac equation in two or three spatial dimensions [17,23–25] and to solve the equations of fluid dynamics with sound-wave evolution operators as building blocks [55]. For a composition of three operators, as one might use for the two-dimensional Dirac equation describing charge carriers in graphene, one can coalesce adjacent pairs of $\mathsf{A}^{1/2}$ as in Sec. 4 and apply $\mathsf{B}^{1/2}\mathsf{C}\mathsf{B}^{1/2}\mathsf{A}$ to a transformed wavefunction. One can then use integer streaming operations along the x and y axes, here denoted by A and C, and fractional steps $\mathsf{B}^{1/2}$ for the algebraic terms at lattice points [17,25].

Declaration of competing interest

The author declares that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The computations employed the Advanced Research Computing facilities at the University of Oxford [56].

Appendix A. Reference solutions

The Gaussian wavepacket is an exact solution of the Schrödinger equation for a non-relativistic free particle. There is no equivalent closed-form solution for the Dirac equation. We thus need very accurate numerical reference solutions to evaluate the accuracy of the composition schemes described in the main text.

Representing the wavefunctions u and d by vectors $\mathbf{u} = (u_1, \ldots, d_N)^\mathsf{T}$ and $\mathbf{d} = (d_1, \ldots, d_N)^\mathsf{T}$ of discrete values on a lattice of N points with periodic boundary conditions leads naturally to the abstract Schrödinger equation

$$i\frac{d}{dt}\begin{pmatrix}\mathbf{u}\\ \mathbf{d}\end{pmatrix} = \mathsf{H}_N\begin{pmatrix}\mathbf{u}\\ \mathbf{d}\end{pmatrix} \tag{A.1}$$

with a $2N \times 2N$ block matrix Hamiltonian

$$\mathsf{H}_N = \begin{pmatrix}\mathsf{V} - i\,\mathsf{D} & im\mathsf{I}\\ -im\mathsf{I} & \mathsf{V} + i\,\mathsf{D}\end{pmatrix}. \tag{A.2}$$

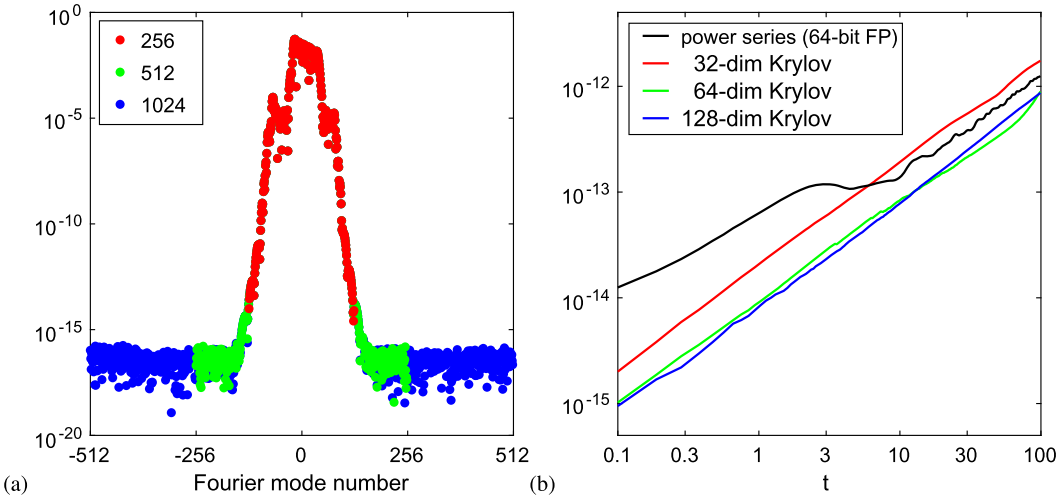


Fig. A.8. (a) Rescaled Fourier coefficients $\sqrt{N}|\hat{P}_k|$ at $t = 99$ computed using a 64-dimensional Krylov space to approximate Hamiltonians with 256, 512, and 1024 Fourier collocation points. (b) Errors in solutions computed with `expokit` using 32, 64, and 128-dimensional Krylov spaces, and the series solution computed in 64-bit arithmetic, relative to the series solution computed in 128-bit arithmetic. All computations for plot (b) used 512 Fourier collocation points.

This is the discrete analogue of the continuous Hamiltonian operator defined in (4). Each block in H_N is an $N \times N$ matrix, I is the identity matrix, and $\mathsf{V} = \mathrm{diag}\,(V(z_1), \ldots, V(z_N))$ is the real diagonal matrix formed by evaluating the scalar potential V at lattice points. The only non-diagonal matrix is the differentiation matrix D that approximates ∂_z on the lattice. The most accurate approximation to ∂_z is provided by the Fourier differentiation matrix, whose elements are given explicitly by [57]

$$D_{ij} = \begin{cases} 0, & \text{for } i = j, \\ (\pi/L)(-1)^{i-j} \cot\left((i-j)\pi/N\right), & \text{for } i \neq j, \end{cases} \tag{A.3}$$

for a domain of length L when N is even. This matrix is anti-Hermitian, so H_N in (A.2) is Hermitian. Matrix-vector products with D, and hence with H_N, can be computed efficiently in $O(N \log N)$ operations using a fast Fourier transform algorithm even though D is a dense matrix.

Solutions of (A.1) with specified initial conditions $\boldsymbol{\Psi}_0 = (\mathbf{u}_0, \mathbf{d}_0)$ can thus be computed efficiently using Arnoldi iteration to construct orthogonal bases for the nested Krylov spaces

$$\mathcal{K}_n = \mathrm{span}\left(\boldsymbol{\Psi}_0, \mathsf{H}_N\boldsymbol{\Psi}_0, \mathsf{H}_N^2\boldsymbol{\Psi}_0, \ldots, \mathsf{H}_N^{n-1}\boldsymbol{\Psi}_0\right),$$

for $n = 1, 2, \ldots$. The action of $\exp(-it\mathsf{H}_N)$ on $\boldsymbol{\Psi}_0$ can be approximated within these spaces by computing the exponential of a much smaller $n \times n$ matrix [58,59]. Most of the comparisons shown in the figures used reference solutions computed using the Krylov space approach implemented in the `ZGEXPV` subroutine from the `expokit` package [60], and the fast Fourier transform library `FFTW3` [61]. The `expokit` package required a minor modification to preserve floating point precision when forming complex numbers from pairs of 64-bit real floating point numbers. The simpler modification replaces the Fortran `CMPLX` intrinsic function with the widely supported extension `DCMPLX`, a simple string substitution in the source code. The same functionality is available in standard modern Fortran by supplying an optional third argument to `CMPLX` to specify the Fortran `kind` (floating point data type) of the complex number.

Figure A.8(a) shows the moduli of the rescaled discrete Fourier coefficients $\sqrt{N}\hat{P}_k$ for the solution at $t = 99$ of the initial value problem in Sec. 5 with 256, 512, and 1024 Fourier collocation points. The wavefunction $P = (u + id)/\sqrt{2}$ is the combination of u and d that appears in the Schrödinger equation (10) describing the non-relativistic regime. The discrete Fourier coefficients $\hat{P}_k$ are defined in (57). The solution was evolved from $t = 0$ to $t = 99$ over steps of length $6\Delta t$ using `ZGEXPV` with a 64-dimensional Krylov space and a prescribed tolerance of 10^{-14}. The rescaled Fourier coefficients $\sqrt{N}\hat{P}_k$ tend to a limiting spectrum that is independent of N when the initial conditions are normalised as in Sec. 5. This is confirmed by the three overlapping sets of data shown in Fig. A.8(a). For $N \geq 512$, the magnitudes of the highest retained Fourier coefficients $\sqrt{N}\hat{P}_k$ are comparable to the machine epsilon in 64-bit arithmetic.

Figure A.8(a) also shows that 512 Fourier collocation points are sufficient when using an unsplit Hamiltonian and a dense Fourier differentiation matrix. Most of the numerical experiments with the composition schemes uses finer lattices with more points. The simplest way to compute the difference between two solutions on different lattices is to project the solution from the finer lattice onto the coarse 512-point lattice of the reference solution. As an alternative, the Fourier series form of the reference solution can be interpolated onto the finer lattice by extending the Fourier coefficients with zeros before computing an inverse Fourier transform on the finer lattice. Parseval's theorem (59) implies that the discrete ℓ^2 norms of the differences computed by projection and interpolation

are comparable when the coefficients of the additional Fourier modes supported by the finer lattice are negligibly small. Figures 2(b) and 5(b) show that the ℓ^2 norms of the differences computed by projection and interpolation are indeed visually indistinguishable.

The solutions computed using `ZGEXPV` turn out to have absolute errors of around 10^{-12} at $t = 99$, as shown in Fig. A.8(b), so still more accurate reference solutions were needed for comparison with the highest resolution computations in the main text. These were computed by forming the matrix H_N explicitly using the concrete form for D in (A.3), then truncating the series expansion of the evolution operator

$$\exp(-3i\Delta t \mathsf{H}_N) = \sum_{r=0}^{\infty} \frac{1}{r!} \left(-3i\Delta t \mathsf{H}_N \right)^r \tag{A.4}$$

so that the Frobenius norm of the last retained term was below 10^{-20}. The Frobenius norm, defined by $||\mathsf{A}||_F^2 = \mathrm{Tr}\,(\mathsf{A}^\dagger \mathsf{A}) = \sum_{i,j} |a_{ij}|^2$, is quick to compute and invariant under unitary transformations. It provides an upper bound for the operator norm $||\mathsf{A}||_{\mathrm{op}}$ defined in Sec. 9. The time interval $3\Delta t$ in (A.4) matches the time step for the fourth-order composition scheme with a halved time step in Sec. 4.2. The solutions were evolved by repeated multiplication using the matrix defined by truncating (A.4). All computations were performed in 128-bit arithmetic using Kahan summation to preserve precision while summing terms of widely differing magnitudes [48,62].

Figure A.8(b) shows a comparison of the ℓ^2 norm of the difference between the series solution and solutions computed using 32, 64, and 128-dimensional Krylov spaces, all using $N = 512$ Fourier collocation points. The discrete ℓ^2 norm of the difference $\delta\boldsymbol{\Psi}$ between two solutions is defined by

$$||\delta\boldsymbol{\Psi}||_2 = \left(\frac{L}{N} \sum_{j=1}^{N} ||\delta\Psi_j||^2 \right)^{1/2}, \tag{A.5}$$

where $||\delta\Psi_j||^2 = |\delta u_j|^2 + |\delta d_j|^2$ is the ℓ^2 norm of the 2-component wavefunction $\delta\Psi_j = (\delta u_j, \delta d_j)^{\mathsf{T}}$ at lattice point j. The 64- and 128-dimensional Krylov space solutions have similar accuracies. Both differ from the series solution by around 10^{-12} at time $t = 99$, comparable to the errors shown in Figs. 2(b) and 5(b) for the composition schemes at the highest resolutions used. The 32-dimensional Krylov space solution has a substantially larger error. The series solution computed in 64-bit arithmetic with Kahan summation has a larger error than the 64- and 128-dimensional Krylov space solutions computed in 64-bit arithmetic.

Neither the Krylov space solutions nor the series solutions exploit the Hermitian property of H_N to guarantee unitary evolution. The extent to which the computed solutions preserve the discrete normalisation $\langle\langle \boldsymbol{\Psi}, \boldsymbol{\Psi} \rangle\rangle$, and the expectation of the discrete Hamiltonian $\langle\langle \boldsymbol{\Psi}, \mathsf{H}_N \boldsymbol{\Psi} \rangle\rangle$, can thus be used to estimate their accuracy [38]. Suppose the numerical solution is $\widetilde{\boldsymbol{\Psi}} = \boldsymbol{\Psi} + \delta\boldsymbol{\Psi}$, where $\boldsymbol{\Psi}$ is the discretely normalised exact solution with $\langle\langle \boldsymbol{\Psi}, \boldsymbol{\Psi} \rangle\rangle = 1$, and $\delta\boldsymbol{\Psi}$ is the error. The squared discrete norm of the numerical solution is

$$\left\langle\left\langle \widetilde{\boldsymbol{\Psi}}, \widetilde{\boldsymbol{\Psi}} \right\rangle\right\rangle = \langle\langle \boldsymbol{\Psi}, \boldsymbol{\Psi} \rangle\rangle + 2\,\langle\langle \boldsymbol{\Psi}, \delta\boldsymbol{\Psi} \rangle\rangle + \langle\langle \delta\boldsymbol{\Psi}, \delta\boldsymbol{\Psi} \rangle\rangle\,, \tag{A.6}$$

so

$$\langle\langle \boldsymbol{\Psi}, \delta\boldsymbol{\Psi} \rangle\rangle = \frac{1}{2} \left(\left\langle\left\langle \widetilde{\boldsymbol{\Psi}}, \widetilde{\boldsymbol{\Psi}} \right\rangle\right\rangle - 1 \right) + O(||\delta\boldsymbol{\Psi}||_2^2), \tag{A.7}$$

where $||\cdot||_2 = \langle\langle \cdot, \cdot \rangle\rangle^{1/2}$ is the ℓ^2 norm derived from the discrete inner product. The discrete and continuous inner products can be used interchangeably, because the trapezoidal rule with uniformly spaced points is exponentially accurate for approximating integrals on domains with periodic boundary conditions. Neglecting the $O(||\delta\boldsymbol{\Psi}||_2^2)$ term and using the Cauchy–Schwartz inequality,

$$\langle\langle \boldsymbol{\Psi}, \delta\boldsymbol{\Psi} \rangle\rangle^2 \leq \langle\langle \boldsymbol{\Psi}, \boldsymbol{\Psi} \rangle\rangle\ \langle\langle \delta\boldsymbol{\Psi}, \delta\boldsymbol{\Psi} \rangle\rangle\,, \tag{A.8}$$

and the normalisation of the exact solution, gives

$$||\delta\boldsymbol{\Psi}||_2 \geq \frac{1}{2} \left| \left\langle\left\langle \widetilde{\boldsymbol{\Psi}}, \widetilde{\boldsymbol{\Psi}} \right\rangle\right\rangle - 1 \right| + O(||\delta\boldsymbol{\Psi}||_2^2). \tag{A.9}$$

Keeping the $O(||\delta\boldsymbol{\Psi}||_2^2)$ term throughout leads to the rigorous, but less illuminating, bound

$$||\delta\boldsymbol{\Psi}||_2 \geq \frac{1}{2} \left(\sqrt{2 \left| \left\langle\left\langle \widetilde{\boldsymbol{\Psi}}, \widetilde{\boldsymbol{\Psi}} \right\rangle\right\rangle - 1 \right| + 1} - 1 \right). \tag{A.10}$$

The error in the normalisation thus gives a lower bound for the ℓ^2 norm of the error in the solution.

Similarly, the expectation of the discrete Hamiltonian operator in (A.2) is

$$\left\langle\left\langle \widetilde{\boldsymbol{\Psi}}, \mathsf{H}_N \widetilde{\boldsymbol{\Psi}} \right\rangle\right\rangle = \langle\langle \boldsymbol{\Psi}, \mathsf{H}_N \boldsymbol{\Psi} \rangle\rangle + 2\,\langle\langle \mathsf{H}_N \boldsymbol{\Psi}, \delta\boldsymbol{\Psi} \rangle\rangle + \langle\langle \delta\boldsymbol{\Psi}, \mathsf{H}_N \delta\boldsymbol{\Psi} \rangle\rangle\,, \tag{A.11}$$

using the Hermitian property of H_N to combine the two cross terms. Neglecting the $O(||\delta\boldsymbol{\Psi}||_2^2)$ term as above, and using the Cauchy–Schwartz inequality for $\delta\boldsymbol{\Psi}$ and $\mathsf{H}_N \boldsymbol{\Psi}$, gives

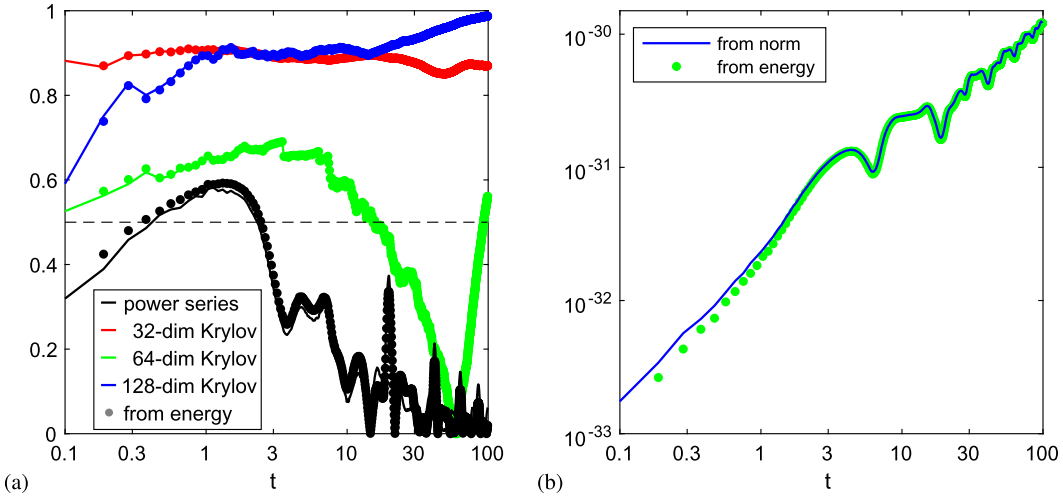


Fig. A.9. (a) Ratios of the lower bounds (A.9) from the norm (lines), and (A.12) from the energy (dots), to the true error $||\delta\mathbf{\Psi}||_2$ for 32, 64, and 128-dimensional Krylov space solutions, and for the series solution computed in 64-bit arithmetic. The horizontal dashed line marks a ratio of 0.5. (b) The two lower bounds (A.9) and (A.12) for the error in the series solution computed in 128-bit arithmetic.

$$||\delta\mathbf{\Psi}||_2 \geq \frac{\left|\left\langle\left\langle \widetilde{\mathbf{\Psi}}, \mathsf{H}_N \widetilde{\mathbf{\Psi}} \right\rangle\right\rangle - E_0\right|}{2||\mathsf{H}_N \mathbf{\Psi}||_2} + O(||\delta\mathbf{\Psi}||_2^2, \tag{A.12}$$

where $E_0 = \langle\langle \mathbf{\Psi}_0, \mathsf{H}_N \mathbf{\Psi}_0 \rangle\rangle$ is the expectation of H_N for the initial conditions. This gives a second lower bound for the ℓ^2 norm of the error in the solution.

Figure A.9(a) shows the ratio of these two bounds for $||\delta\mathbf{\Psi}||_2$ to its actual value, as computed by the difference from the series solution in 128-bit arithmetic, for 32-, 64- and 128-dimensional Krylov space solutions, and for the series solution in 64-bit arithmetic. The two bounds are very close to each other, and for the 32- and 128-dimensional Krylov space solutions they are within a factor of 2 of the actual error relative to the solution in 128-bit arithmetic. For $t \lesssim 10$ they are within a factor of 2 of the actual error for the 64-dimensional Krylov space solution. Figure A.9(b) shows the two lower bounds for $||\delta\mathbf{\Psi}||_2$ calculated for the series solution in 128-bit arithmetic. Although these are only lower bounds, and no more accurate solution is available for calculating $||\delta\mathbf{\Psi}||_2$ directly, the observation that both lower bounds remain below 10^{-29} provides some grounds for confidence in the accuracy of the series solution.

Data availability

The data used to create the figures can be obtained from the University of Oxford Research Archive.

References

[1] R.I. McLachlan, G.R. Quispel, Six lectures on the geometric integration of ODEs, in: R. Devore, A. Iserles, E. Süli (Eds.), Foundations of Computational Mathematics, Cambridge University Press, 2001, pp. 155–210.
[2] E. Hairer, C. Lubich, G. Wanner, Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations, 2nd ed., Springer, 2006.
[3] A. Iserles, G.R.W. Quispel, Why geometric numerical integration?, in: K. Ebrahimi-Fard, M. Barbero Liñán (Eds.), Discrete Mechanics, Geometric Integration and Lie–Butcher Series, Springer, Cham, Switzerland, 2018, pp. 1–28.
[4] S. Blanes, F. Casas, A Concise Introduction to Geometric Numerical Integration, CRC Press, Boca Raton, FL, 2016.
[5] J.D. Bjorken, S.D. Drell, Relativistic Quantum Mechanics, McGraw–Hill, New York, London, 1964.
[6] W. Pauli, General Principles of Quantum Mechanics, Springer, Berlin, Heidelberg, New York, 1980. Translated by P. Achuthan & K. Venkatesan from articles in Handbuch der Physik (1933 and 1958).
[7] F. Schwabl, Advanced Quantum Mechanics, 3rd ed., Springer, Berlin, London, 2005.
[8] Y. Aharonov, L. Davidovich, N. Zagury, Quantum random walks, Phys. Rev. A 48 (1993) 1687–1690.
[9] I. Bialynicki-Birula, Weyl, Dirac, and Maxwell equations on a lattice as unitary cellular automata, Phys. Rev. D 49 (1994) 6920–6927.
[10] D.A. Meyer, From quantum cellular automata to quantum lattice gases, J. Stat. Phys. 85 (1996) 551–574.
[11] J. Kempe, Quantum random walks: an introductory overview, Contemp. Phys. 44 (2003) 307–327.
[12] F.W. Strauch, Relativistic effects and rigorous limits for discrete- and continuous-time quantum walks, J. Math. Phys. 48 (2007) 082102.
[13] S. Succi, R. Benzi, Lattice Boltzmann equation for quantum mechanics, Physica D 69 (1993) 327–332.
[14] R.P. Feynman, A.R. Hibbs, Quantum Mechanics and Path Integrals, McGraw–Hill, New York, 1965, pp. 35–36.
[15] G. Strang, On the construction and comparison of difference schemes, SIAM J. Numer. Anal. 5 (1968) 506–517.
[16] P.J. Dellar, An exact energy conservation property of the quantum lattice Boltzmann algorithm, Phys. Lett. A 376 (2011) 6–13.
[17] P.J. Dellar, Quantum lattice algorithms: similarities and connections to some classic finite difference algorithms, ESAIM Proc. 52 (2015) 76–104.
[18] X. He, S. Chen, G.D. Doolen, A novel thermal model of the lattice Boltzmann method in incompressible limit, J. Comput. Phys. 146 (1998) 282–300.
[19] M. Ablowitz, M. Kruskal, J. Ladik, Solitary wave collisions, SIAM J. Appl. Math. 36 (1979) 428–437.
[20] E.C. Du Fort, S.P. Frankel, Stability conditions in the numerical treatment of parabolic differential equations, Math. Tables Other Aids Comput. 7 (1953) 135–152.

[21] L. Wu, Dufort–Frankel-type methods for linear and nonlinear Schrödinger equations, SIAM J. Numer. Anal. 33 (1996) 1526–1533.
[22] P.J. Dellar, Relativistic properties and invariants of the Du Fort–Frankel scheme for the one-dimensional Schrödinger equation, J. Comput. Phys. X 2 (2019) 100004.
[23] J. Yepez, Quantum Algorithms for Computational Physics: Volume 3 of Lattice Gas Dynamics, Tech. Rep. AFRL-VS-HA-TR-2006-1143, Air Force Research Laboratory, Hanscom, MA, 2007 Available from http://www.dtic.mil/docs/citations/ADA474659.
[24] P.J. Dellar, D. Lapitski, S. Palpacelli, S. Succi, Isotropy of three-dimensional quantum lattice Boltzmann schemes, Phys. Rev. E 83 (2011) 046706.
[25] D. Lapitski, Development of the Quantum Lattice Boltzmann method for simulation of quantum electrodynamics with applications to graphene, DPhil thesis, University of Oxford, 2014, Available from http://ora.ox.ac.uk/objects/uuid:e89cd11b-da2c-4c34-be9f-7b3d711e2e64.
[26] M.I. Katsnelson, K.S. Novoselov, A.K. Geim, Chiral tunnelling and the Klein paradox in graphene, Nat. Phys. 2 (2006) 620–625.
[27] M. Suzuki, Fractal decomposition of exponential operators with applications to many-body theories and Monte Carlo simulations, Phys. Lett. A 146 (1990) 319–323.
[28] M. Suzuki, General theory of fractal path integrals with applications to many-body theories and statistical physics, J. Math. Phys. 32 (1991) 400–407.
[29] H. Yoshida, Construction of higher order symplectic integrators, Phys. Lett. A 150 (1990) 262–268.
[30] R.I. McLachlan, G.R.W. Quispel, Splitting methods, Acta Numer. 11 (2002) 341–434.
[31] R.I. McLachlan, On the numerical integration of ordinary differential equations by symmetric composition methods, SIAM J. Sci. Comput. 16 (1995) 151–168.
[32] F. Neri, Lie Algebras and Canonical Integration, Preprint, Department of Physics, University of Maryland, 1988.
[33] K.J. Blow, D. Wood, Theoretical description of transient stimulated Raman scattering in optical fibers, IEEE J. Quantum Electron. 25 (1989) 2665–2673.
[34] T. Bellotti, P. Helluy, L. Navoret, Fourth-order entropy-stable lattice Boltzmann schemes for hyperbolic systems, SIAM J. Sci. Comput. 47 (2025) A586–A611.
[35] H.F. Trotter, On the product of semi-groups of operators, Proc. Am. Math. Soc. 10 (1959) 545–551.
[36] A. Van-Brunt, M. Visser, Simplifying the Reinsch algorithm for the Baker–Campbell–Hausdorff series, J. Math. Phys. 57 (2016) 023507.
[37] M. Suzuki, Decomposition formulas of exponential operators and Lie exponentials with some applications to quantum mechanics and statistical physics, J. Math. Phys. 26 (1985) 601–612.
[38] C. Leforestier, R.H. Bisseling, C. Cerjan, M.D. Feit, R. Friesner, A. Guldberg, et al., A comparison of different propagation schemes for the time dependent Schrödinger equation, J. Comput. Phys. 94 (1991) 59–80.
[39] P.J. Dellar, An interpretation and derivation of the lattice Boltzmann method using Strang splitting, Comput. Math. Appl. 65 (2013) 129–141.
[40] J.C. Butcher, The effective order of Runge–Kutta methods, Lect. Notes Math. 109 (1969) 133–139.
[41] J. Wisdom, M. Holman, J. Touma, Symplectic correctors, in: J.E. Marsden, G.W. Patrick, W.F. Shadwick (Eds.), Integration Algorithms and Classical Mechanics, in: Fields Institute Communications, vol. 10, Amer. Math. Soc., 1996, pp. 217–244.
[42] M.A. López-Marcos, J.M. Sanz-Serna, R.D. Skeel, Cheap enhancement of symplectic integrators, in: D.F. Griffiths, G.A. Watson (Eds.), Numerical Analysis 1995, Longman, Harlow, 1996, pp. 107–122.
[43] Q. Sheng, Solving linear partial differential equations by exponential splitting, IMA J. Numer. Anal. 9 (1989) 199–212.
[44] S. Blanes, F. Casas, On the necessity of negative coefficients for operator splitting schemes of order higher than two, Appl. Numer. Math. 54 (2005) 23–37.
[45] P. Ehrenfest, Bemerkung über die angenäherte Gültigkeit der klassischen Mechanik innerhalb der Quantenmechanik, Z. Phys. 45 (1927) 455–457.
[46] L.D. Landau, E. Lifshitz, Quantum Mechanics: Non-Relativistic Theory, 3 ed., Butterworth–Heinemann, 1977.
[47] C.M. Bender, S.A. Orszag, Advanced Mathematical Methods for Scientists and Engineers, Springer, New York, 1999.
[48] N. Higham, Accuracy and Stability of Numerical Algorithms, second ed., Society for Industrial and Applied Mathematics, Philadelphia, 2002.
[49] R.H. Hardin, F.D. Tappert, Applications of the split-step Fourier method to the numerical solution of nonlinear and variable coefficient wave equations, SIAM Rev. 15 (1973) 423.
[50] M.D. Feit, J.A. Fleck, A. Steiger, Solution of the Schrödinger equation by a spectral method, J. Comput. Phys. 47 (1982) 412–433.
[51] J. Weideman, B. Herbst, Split-step methods for the solution of the nonlinear Schrödinger equation, SIAM J. Numer. Anal. 23 (1986) 485–507.
[52] P. Henrici, Applied and Computational Complex Analysis, vol. 3, Wiley, New York, 1993.
[53] S.G. Johnson, Notes on FFT-based differentiation, Available from https://math.mit.edu/~stevenj/fft-deriv.pdf, 2011.
[54] E.S. Cheb-Terrab, Symbolic computing with anti-commutative and non-commutative variables, MapleTech 5 (1998) 16–22.
[55] R. Salmon, An ocean circulation model based on operator-splitting, Hamiltonian brackets, and the inclusion of sound waves, J. Phys. Oceanogr. 39 (2009) 1615–1633.
[56] A. Richards, University of Oxford Advanced Research Computing, Technical Note, 2015.
[57] D. Gottlieb, M.Y. Hussaini, S.A. Orszag, Theory and applications of spectral methods, in: R. Voigt, D. Gottlieb, M. Hussaini (Eds.), Spectral Methods for Partial Differential Equation, Society of Industrial and Applied Mathematics, Philadelphia, 1984, pp. 1–54.
[58] Y. Saad, Analysis of some Krylov subspace approximations to the matrix exponential operator, SIAM J. Numer. Anal. 29 (1992) 209–228.
[59] M. Hochbruck, C. Lubich, On Krylov subspace approximations to the matrix exponential operator, SIAM J. Numer. Anal. 34 (1997) 1911–1925.
[60] B.R. Sidje, Expokit: a software package for computing matrix exponentials, ACM Trans. Math. Softw. 24 (1998) 130–156.
[61] M. Frigo, S.G. Johnson, The design and implementation of FFTW3, Proc. IEEE 93 (2005) 216–231.
[62] W. Kahan, Further remarks on reducing truncation errors, Commun. ACM 8 (1965) 40.