MULTISCALE METHODS AND ANALYSIS FOR THE DIRAC/NONLINEAR DIRAC EQUATION

YIN JIA

NATIONAL UNIVERSITY OF SINGAPORE

2019

MULTISCALE METHODS AND ANALYSIS FOR THE DIRAC/NONLINEAR DIRAC EQUATION

by

YIN JIA

(B.Sc., Tsinghua University)

A THESIS SUBMITTED FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

NUS GRADUATE SCHOOL FOR INTEGRATIVE SCIENCES AND ENGINEERING

NATIONAL UNIVERSITY OF SINGAPORE

2019

Supervisor:

Professor Bao Weizhu

Examiners:

Professor Ren Weiqing
Professor Thalhammer Mechthild, University of Innsbruck
Associate Professor Wang Li-Lian, Nanyang Technological University

Declaration

I hereby declare that this thesis is my original work and it has been written by me in its entirety. I have duly acknowledged all the sources of information which have been used in the thesis.

This thesis has also not been submitted for any degree in any university previously.

Yin Jia

July 30, 2019

Acknowledgments

It is my great honor here to thank those who make this thesis possible.

First and foremost, I would like to express my most heartfelt gratitude to my supervisor, Prof. Weizhu Bao, who led me to this exciting research field and patiently guided me throughout my PhD candidature, for his generous support, valuable suggestions, as well as continuous care and encouragement. His broad knowledge in many areas and profound insight always surprised me. I learnt quite a lot from him through profession to life.

Secondly, I would like to thank my TAC members, Prof. Yuan Ping Feng and Prof. Jiangbin Gong for providing much helpful advice on my project, especially from the viewpoint of physics. I sincerely appreciate their attendance to all the TAC meetings in spite of their busy schedule.

My sincere thanks also go to my collaborator, Prof. Yongyong Cai, for his generous help and contribution to the work. This project can not be done without the fruitful discussion with him.

Besides, I am grateful to the colleagues in my group, especially Prof. Yan Wang, Prof. Wenfan Yi, Prof. Qinglin Tang, Prof. Hanquan Wang, Prof. Huihui Kong, Dr. Chunmei Su, Dr. Weijie Huang, Dr. Quan Zhao, Dr. Zhaoqiang Liu, Dr. Ying Ma, Miss Yue Feng, Mr. Teng Zhang and Mr. Bo Lin for their help, encouragement and suggestions. I would like to thank my friends Tong Teng, Han Kou, Menghua Sun, my seniors Yangjing Zhang, Yu Guan, Tongyao Pang, Bin Wu, and Liuquan Wang for their emotional support and sharing of life.

I also greatly appreciate NUS and NGS for awarding me the research scholarship during my PhD candidature, and the financial support for the conference leaves.

Last but not least, thank my dearest family, my mother Wenfeng Jin, father Dawei Yin, and husband Xianzhe Chen for their unconditional love, trust, support, and consideration.

Jia YIN

July 2019

Contents

Summary

The Dirac equation is a relativistic wave equation describing the motion of spin-1/2 massive particles such as electrons and quarks. It can be seen as the relativistic version of the Schrödinger equation, and thus plays an important role in quantum mechanics. There are three interesting scaling parameters in the non-dimensionalized Dirac equation. Taking different limits of these parameters results in the nonrelativistic limit regime, the semiclassical regime and the massless regime of the Dirac equation. The dynamics of the Dirac equation behaves differently under these regimes, and it is worthwhile to study the performance of various numerical methods respectively.

The aim of this thesis is to propose efficient numerical methods to solve the Dirac equation in different limit regimes, and study their properties. Rigorous proof are presented and numerical results are reported to verify the error bounds and compare the performance of the methods. The thesis mainly consists of three parts:

In the first part, a new fourth-order compact time-splitting (S_{4c}) Fourier pseudospectral method is put forward for the Dirac equation. The method splits the Dirac equation into two parts and uses a double commutator between them to integrate the Dirac equation at each time interval. It is explicit, fourth-order in time and spectral order in space, and it is called a compact time-splitting method since, at each time step, the number of sub-steps in S_{4c} is much less than those of the standard fourth-order splitting method and the fourth-order partitioned Runge-Kutta splitting method. Another advantage of S_{4c} is that it avoids using negative time steps in integrating sub-problems at each time interval. Comparison between S_{4c} and many other existing time-splitting methods for the Dirac equation are carried out in terms of accuracy and efficiency as well as long time behavior. Numerical results demonstrate the advantage in terms of efficiency and accuracy of the proposed S_{4c}. The spatial/temporal resolutions of S_{4c} for the Dirac equation in different parameter regimes including the nonrelativistic limit regime and the semiclassical limit regime are reported through numerical examples.

The second part deals with super-resolution of the time-splitting methods, especially the Lie-Trotter splitting (S_1) and the Strang splitting (S_2) for the Dirac and nonlinear Dirac equation without external magnetic potentials in the nonrelativistic limit regime, with a small parameter $0 < \varepsilon \leq 1$ inversely proportional to the speed of light. In this limit regime, the

solution highly oscillates in time with wavelength at $O(\varepsilon^2)$ in time. The splitting methods surprisingly show super-resolution, in the sense of breaking the resolution constraint under the Shannon's sampling theorem, i.e. the methods can capture the solution accurately even if the time step size τ is much larger than the sampled wavelength at $O(\varepsilon^2)$. In both the Dirac equation and the nonlinear Dirac equation cases, S_1 shows $1/2$ order convergence uniformly with respect to ε, as there are two independent error bounds $\tau+\varepsilon$ and $\tau+\tau/\varepsilon$. Moreover, if τ is non-resonant, i.e. τ is away from certain region determined by ε, S_1 would yield an improved uniform first order $O(\tau)$ error bound. In addition, S_2 is uniformly convergent for the Dirac/nonlinear Dirac equation with 1/2 order rate for general time step size τ and uniformly convergent with $3/2$ order rate for non-resonant time step size. Numerical results are reported to confirm these rigorous results. Furthermore, it is noted that super-resolution is still valid for higher order splitting methods.

The third part is devoted to studying rigorously the error bounds of four frequently-used finite difference time domain (FDTD) methods for the Dirac equation in the semiclassical regime, involving a small dimensionless parameter $0<\delta\leq 1$ representing the scaled Planck constant. In this regime, there are highly oscillatory propogating waves with wavelength $O(\delta)$ in both time and space of the solution. The leap-frog, two semi-implicit, and the Crank-Nicolson finite difference methods are applied to numerically solve the Dirac equation in the semiclassical regime, and their error estimates are rigorously established respectively. It is proved that these methods share the same error bounds, which are explicitly related to time step size τ, mesh size h, as well as the small parameter δ. Furthermore, the dependence of the observables, i.e. the total probability density and the current density on the parameters τ, h and δ are found out. Based on the error bounds, in the semiclassical regime, i.e. $0<\delta\ll 1$, to obtain 'correct' numerical solutions and related observables, the δ-scalabilities $\tau=O(\delta^{3/2})$ and $h=O(\delta^{3/2})$ are required for all these FDTD methods. Numerical tests are carried out to support the error estimates.

List of Figures

List of Tables

List of Symbols and Abbreviations

$\hbar$	Planck constant
c	speed of light
t	time variable
$\mathbb{R}^d$	d-dimensional Euclidean space
$\mathbb{C}^d$	d-dimensional complex space
$\mathbf{x} = (x_1, ..., x_d)^T$	spatial variable in $\mathbb{R}^d$
τ	time step size
h	mesh size
ε	a dimensionless parameter in $(0, 1]$, inversely proportional to the light speed
δ	scaled Planck constant
ν	the dimensionless mass parameter
$\Psi := \Psi(t, \mathbf{x}) \in \mathbb{C}^4$	4-component complex wave function
$\Phi := \Phi(t, \mathbf{x}) \in \mathbb{C}^2$	2-component complex wave function
σ_j $(j = 1, 2, 3)$	Pauli matrices
∂_t	the partial derivative of t
∂_j $(j = 1, 2, 3)$	the partial derivative of x_j
$A \lesssim B$	$\lvert A\rvert \leq C \cdot B$ for some generic constant $C > 0$
$A \lesssim_\delta B$	$\lvert A\rvert \leq C_\delta \cdot B$ for some generic constant $C_\delta > 0$ related to δ
$\mathrm{Re}(f)$	the real part of f
$\mathrm{Im}(f)$	the imaginary part of f
$\overline{f}$	the conjugate of f
A^T	the transpose of matrix A
A^*	the conjugate transpose of matrix A

1D	one dimension
2D	two dimension
3D	three dimension
LFFD	leap-frog finite difference
SIFD	semi-implicit finite difference
CNFD	Crank-Nicolson finite difference
EWI	exponential wave integrator
EWI-FP	exponential wave integrator Fourier pseudospectral
TSFP	time-splitting Fourier pseudospectral
S_1	first-order splitting (Lie-Trotter splitting)
S_2	second-order splitting (Strang splitting)
S_4	fourth-order splitting
$S_{4\mathrm{RK}}$	fourth-order Runge-Kutta splitting
$S_{4\mathrm{c}}$	fourth order compact splitting

Chapter 1

Introduction

This chapter serves as an introduction of the thesis. A brief overview of different regimes of the Dirac equation is presented, and the relation among Dirac, Weyl and Majorana equations is discussed. In the third section, we introduce the nonlinear Dirac equation. The last two sections summarize the problems studied and show the structure and scope of the thesis.

1.1 The Dirac equation and its different regimes

The Schrödinger equation plays an important role in quantum mechanics, just as Newton's second law does in classical physics. It is a scalar equation describing the evolution of a quantum system [103]. However, the Schrödinger equation would no longer be valid when the velocity of the particle is very high so that special relativity should be taken into accout. In this case, the Klein-Gordon equation was first proposed in 1926 [49]. It solved many problems, but the most severe drawback is that its probability density may be negative. To solve this problem, Paul Dirac derived the Dirac equation in 1928 , which could be seen as the square root of the Klein-Gordon equation, and has only the first order time derivative.

The standard expression of the Dirac equation under external electromagnetic potentials is given as [27, 50, 51, 52, 121]

$$i\hbar\partial_t\Psi = \left(-ic\hbar\sum_{j=1}^{3}\alpha_j\partial_j + mc^2\beta\right)\Psi + e\left(V(t,\mathbf{x})I_4 - \sum_{j=1}^{3}A_j(t,\mathbf{x})\alpha_j\right)\Psi, \quad \mathbf{x}\in\mathbb{R}^3. \tag{1.1.1}$$

In the equation, $\Psi := \Psi(t,\mathbf{x}) = (\psi_1(t,\mathbf{x}),\ \psi_2(t,\mathbf{x}), \psi_3(t,\mathbf{x}), \psi_4(t,\mathbf{x}))^T \in \mathbb{C}^4$ is the complex-valued spinor wave function, with t representing the time, and $\mathbf{x} = (x_1, x_2, x_3)^T$ representing

the spatial coordinate. ∂_j means ∂_{x_j} for $j = 1,2,3$. $V(t) := V(t,\mathbf{x})$, and $\mathbf{A}(t) := \mathbf{A}(t,\mathbf{x}) = (A_1(t,\mathbf{x}),A_2(t,\mathbf{x}),A_3(t,\mathbf{x}))^T$ respectively stand for the external electric and magnetic potentials, which are all real-valued given functions. There are also many constants including $i = \sqrt{-1}$, $\hbar$ the Planck constant, m the mass, c the speed of light and e the unit charge. Finally, β and α_j ($j = 1,2,3$) are 4×4 Dirac representation matrices of the four-dimensional Clifford algebra:

$$\beta = \begin{pmatrix} I_2 & \mathbf{0} \\ \mathbf{0} & -I_2 \end{pmatrix}, \quad \alpha_j = \begin{pmatrix} \mathbf{0} & \sigma_j \\ \sigma_j & \mathbf{0} \end{pmatrix}, \quad j = 1,2,3, \tag{1.1.2}$$

where I_n is the $n\times n$ identity matrix and σ_j ($j = 1,2,3$) are the Pauli matrices defined as:

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{1.1.3}$$

The Dirac equation is widely applied in relativistic quantum mechanics. It describes the motion of relativistic spin-1/2 massive particles, such as electrons and quarks. It fully explained the hydrogen spectrum and predicted the existence of antimatter. Recently, the Dirac equation has been extensively adopted to investigate theoretically the structures and/or dynamical properties of graphene and graphite as well as other two-dimensional (2D) materials [1, 59, 96, 95], and to study the relativistic effects in molecules in super intense lasers, e.g., attosecond lasers [31, 63].

The Dirac equation (1.1.1) could be nondimensionalized using

$$\tilde{\mathbf{x}} = \frac{\mathbf{x}}{x_s}, \quad \tilde{t} = \frac{t}{t_s}, \quad \tilde{m} = \frac{m}{m_s}, \quad \tilde{V} = \frac{V}{A_s}, \quad \tilde{\mathbf{A}} = \frac{\mathbf{A}}{A_s}, \quad \tilde{\Psi}(\tilde{t},\tilde{\mathbf{x}}) = \frac{\Psi(t,\mathbf{x})}{\psi_s}, \tag{1.1.4}$$

where x_s, t_s and m_s are respectively length unit, time unit and mass unit. Plugging (1.1.4) into (1.1.1) and taking $\psi_s = x_s^{-3/2}$ and $A_s = \frac{m_s x_s^2}{e t_s^2}$, after some simplification and then removing all $\tilde{}$, we obtain the dimensionless Dirac equation in 3D

$$i\delta\partial_t\Psi = \left(-i\frac{\delta}{\varepsilon}\sum_{j=1}^{3}\alpha_j\partial_j + \frac{\nu}{\varepsilon^2}\beta\right)\Psi + \left(V(t,\mathbf{x})I_4 - \sum_{j=1}^{3}A_j(t,\mathbf{x})\alpha_j\right)\Psi, \quad \mathbf{x}\in\mathbb{R}^3, \tag{1.1.5}$$

where the three dimensionless parameters $0 < \varepsilon,\delta,\nu \le 1$ are given as

$$\varepsilon = \frac{x_s}{t_s c} = \frac{v_s}{c}, \quad \delta = \frac{\hbar t_s}{m_s x_s^2}, \quad \nu = \tilde{m} = \frac{m}{m_s}, \tag{1.1.6}$$

with $v_s = x_s/t_s$ defined as the velocity unit. Indeed, here ε indicates the ratio between the wave velocity and the speed of light, δ stands for the scaled Planck constant and ν is the ratio between the mass of the particle and the mass unit.

As discussed in [15], under certain assumptions on the electromagnetic potentials $V(t,\mathbf{x})$ and $\mathbf{A}(t,\mathbf{x})$, the Dirac equation (1.1.5) in 3D could be reduced to equations in two dimensions (2D) and one dimension (1D). Specifically, the Dirac equation in 2D has been widely applied in modeling the electron structure and dynamics of graphene and other 2D materials as they share the same dispersion relation on certain points in the phase space which are called Dirac or conical points [59, 61, 62, 95]. Actually, the Dirac equation (1.1.5) in 3D and its dimension reduction to 2D and 1D can be expressed in a unified way as

$$i\delta\partial_t\Psi = \left(-i\frac{\delta}{\varepsilon}\sum_{j=1}^{d}\alpha_j\partial_j + \frac{\nu}{\varepsilon^2}\beta\right)\Psi + \left(V(t,\mathbf{x})I_4 - \sum_{j=1}^{d}A_j(t,\mathbf{x})\alpha_j\right)\Psi, \quad \mathbf{x}\in\mathbb{R}^d, \quad (1.1.7)$$

where $d = 1,2,3$ indicates the dimension, $\mathbf{x}$ is set to be $(x_1,x_2)^T$ in 2D and x_1 in 1D. To study the dynamics of the Dirac equation (1.1.7), the initial condition is usually taken as

$$\Psi(t=0,\mathbf{x}) = \Psi_0(\mathbf{x}), \quad \mathbf{x}\in\mathbb{R}^d. \quad (1.1.8)$$

The Dirac equation (1.1.7) with (1.1.8) is dispersive, time-symmetric, and it conserves the total *probability* [15]

$$\|\Psi(t,\cdot)\|^2 := \int_{\mathbb{R}^d}|\Psi(t,\mathbf{x})|^2 d\mathbf{x} = \int_{\mathbb{R}^d}\sum_{j=1}^{4}|\psi_j(t,\mathbf{x})|^2 d\mathbf{x} \equiv \|\Psi(0,\cdot)\|^2 = \|\Psi_0\|^2, \ t\geq 0, \quad (1.1.9)$$

and the *energy* [15]

$$\begin{aligned} E(\Psi(t,\cdot)) &:= \int_{\mathbb{R}^d}\left(-i\frac{\delta}{\varepsilon}\sum_{j=1}^{d}\Psi^*\alpha_j\partial_j\Psi + \frac{\nu}{\varepsilon^2}\Psi^*\beta\Psi + V(t,\mathbf{x})|\Psi|^2 - \sum_{j=1}^{d}A_j(t,\mathbf{x})\Psi^*\alpha_j\Psi\right)d\mathbf{x} \\ &\equiv E(\Psi_0), \qquad t\geq 0, \end{aligned} \quad (1.1.10)$$

where $\Psi^* = \overline{\Psi}^T$ with $\overline{f}$ denoting the complex conjugate of f.

Introduce the total probability density $\rho := \rho(t,\mathbf{x})$ as

$$\rho(t,\mathbf{x}) = \sum_{j=1}^{4}\rho_j(t,\mathbf{x}) = \Psi(t,\mathbf{x})^*\Psi(t,\mathbf{x}), \quad \mathbf{x}\in\mathbb{R}^d, \quad (1.1.11)$$

where the probability density $\rho_j := \rho_j(t,\mathbf{x})$ of the j-th ($j = 1,2,3,4$) component is defined as

$$\rho_j(t,\mathbf{x}) = |\psi_j(t,\mathbf{x})|^2, \quad \mathbf{x} \in \mathbb{R}^d, \tag{1.1.12}$$

and the current density $\mathbf{J}(t,\mathbf{x}) = (J_1(t,\mathbf{x}),\ldots,J_d(t,\mathbf{x})))^T$ is defined as

$$J_l(t,\mathbf{x}) = \frac{1}{\varepsilon}\Psi(t,\mathbf{x})^*\alpha_l\Psi(t,\mathbf{x}), \quad l = 1,\ldots,d, \tag{1.1.13}$$

then we could derive the following conservation law from the Dirac equation (1.1.7) [15]

$$\partial_t\rho(t,\mathbf{x}) + \nabla\cdot\mathbf{J}(t,\mathbf{x}) = 0, \quad \mathbf{x} \in \mathbb{R}^d, \quad t \geq 0. \tag{1.1.14}$$

Moreover, if the electric potential V is perturbed by a real constant V^0, i.e., $V \to V + V^0$, then the solution $\Psi(t,\mathbf{x}) \to e^{-i\frac{V^0 t}{\delta}}\Psi(t,\mathbf{x})$, which implies that the probability density of each component $\rho_j (j = 1,2,3,4)$ and the total probability density ρ are all unchanged. In addition, when $d = 1$, if the magnetic potential A_1 is perturbed by a real constant A_1^0, i.e., $A_1 \to A_1 + A_1^0$, then the solution $\Psi(t,\mathbf{x}) \to e^{i\frac{A_1^0 t}{\delta}\alpha_1}\Psi(t,\mathbf{x})$, which implies that only the total probability density ρ is unchanged; however, this property is unfortunately not valid in 2D and 3D. Furthermore, if the external electromagnetic potentials are all real constants, i.e. $V(t,\mathbf{x}) \equiv V^0$ and $A_j(t,\mathbf{x}) \equiv A_j^0$ ($j = 1,\ldots,d$) with $\mathbf{A}^0 = (A_1^0,...,A_d^0)^T$, the Dirac equation (1.1.7) admits the plane wave solution $\Psi(t,\mathbf{x}) = \mathbf{B}e^{i(\mathbf{k}\cdot\mathbf{x} - \frac{\omega}{\delta}t)}$ with ω the time frequency, $\mathbf{B} \in \mathbb{R}^4$ the amplitude vector and $\mathbf{k} = (k_1,...,k_d)^T \in \mathbb{R}^d$ the spatial wave number, which satisfies the following eigenvalue problem

$$\omega\mathbf{B} = \left(\sum_{j=1}^{d}\left(\frac{\delta k_j}{\varepsilon} - A_j^0\right)\alpha_j + \frac{\nu}{\varepsilon^2}\beta + V^0 I_4\right)\mathbf{B}. \tag{1.1.15}$$

Solving the above equation, we can get the *dispersion relation* of the Dirac equation (1.1.7)

$$\omega := \omega(\mathbf{k}) = V^0 \pm \frac{1}{\varepsilon^2}\sqrt{\nu^2 + \varepsilon^2|\delta\mathbf{k} - \varepsilon\mathbf{A}^0|^2}, \quad \mathbf{k} \in \mathbb{R}^d. \tag{1.1.16}$$

In 2D and 1D, i.e. $d = 2$ or 1 in (1.1.7), similar to the process in [14], the Dirac equation (1.1.7) can be decoupled into simplified PDEs with two-component wave function $\Phi := \Phi(t,\mathbf{x}) = (\phi_1(t,\mathbf{x}),\phi_2(t,\mathbf{x}))^T \in \mathbb{C}^2$ satisfying

$$i\delta\partial_t\Phi = \left(-i\frac{\delta}{\varepsilon}\sum_{j=1}^{d}\sigma_j\partial_j + \frac{\nu}{\varepsilon^2}\sigma_3\right)\Phi + \left(V(t,\mathbf{x})I_2 - \sum_{j=1}^{d}A_j(t,\mathbf{x})\sigma_j\right)\Phi, \quad \mathbf{x} \in \mathbb{R}^d, \tag{1.1.17}$$

where $\Phi = (\psi_1, \psi_4)^T$ (or $\Phi = (\psi_2, \psi_3)^T$) gives the relation between Φ and Ψ. Again, to study the dynamics of the Dirac equation (1.1.17), we usually take the initial condition as

$$\Phi(t=0,\mathbf{x}) = \Phi_0(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^d. \tag{1.1.18}$$

The Dirac equation (1.1.17) with (1.1.18) share similar properties with (1.1.7). It is dispersive, time-symmetric, and it conserves the total *probability* [15]

$$\begin{aligned} \|\Phi(t,\cdot)\|^2 &:= \int_{\mathbb{R}^d} |\Phi(t,\mathbf{x})|^2 d\mathbf{x} = \int_{\mathbb{R}^d} \sum_{j=1}^{2} |\phi_j(t,\mathbf{x})|^2 d\mathbf{x} \\ &\equiv \|\Phi(0,\cdot)\|^2 = \|\Phi_0\|^2, \quad t \geq 0, \end{aligned} \tag{1.1.19}$$

and the *energy* [15]

$$\begin{aligned} E(\Phi(t,\cdot)) &:= \int_{\mathbb{R}^d} \left(-i\frac{\delta}{\varepsilon} \sum_{j=1}^{d} \Phi^* \sigma_j \partial_j \Phi + \frac{\nu}{\varepsilon^2} \Phi^* \sigma_3 \Phi + V(t,\mathbf{x}) |\Phi|^2 - \sum_{j=1}^{d} A_j(t,\mathbf{x}) \Phi^* \sigma_j \Phi \right) d\mathbf{x} \\ &\equiv E(\Phi_0), \qquad t \geq 0. \end{aligned} \tag{1.1.20}$$

By taking proper definitions of the total probability density $\rho := \rho(t,\mathbf{x})$ and the current density $\mathbf{J}(t,\mathbf{x}) = (J_1(t,\mathbf{x}), \ldots, J_d(t,\mathbf{x})))^T$, we could get the same conservation law (1.1.14) from the Dirac equation (1.1.17) [15].

Furthermore, the *dispersion relation* of (1.1.17) is

$$\omega := \omega(\mathbf{k}) = V^0 \pm \frac{1}{\varepsilon^2} \sqrt{\nu^2 + \varepsilon^2 |\delta \mathbf{k} - \varepsilon \mathbf{A}^0|^2}, \quad \mathbf{k} \in \mathbb{R}^d, \tag{1.1.21}$$

which is derived from a similar process as for (1.1.7).

If mass unit is chosen as $m_s = m$, length unit $x_s = \frac{\hbar}{mc}$, and time unit $t_s = \frac{x_s}{c} = \frac{\hbar}{mc^2}$, then we would have $\varepsilon = \delta = \nu = 1$, which corresponds to the classical (or standard) scaling. This choice of x_s, m_s and t_s is appropriate when the wave speed is at the same order of the light velocity. However, a different choice of x_s, m_s and t_s is more appropriate in other cases. We remark that the choice of x_s, m_s and t_s determines the observation scale of time evolution of the system and decides which phenomena can be resolved by discretization on specified spatial/temporal grids and which phenomena is 'visible' by asymptotic analysis.

More specifically, different parameter regimes could be considered for the Dirac equation (1.1.7) (or (1.1.17)), which are displayed in Fig. 1.1.1:

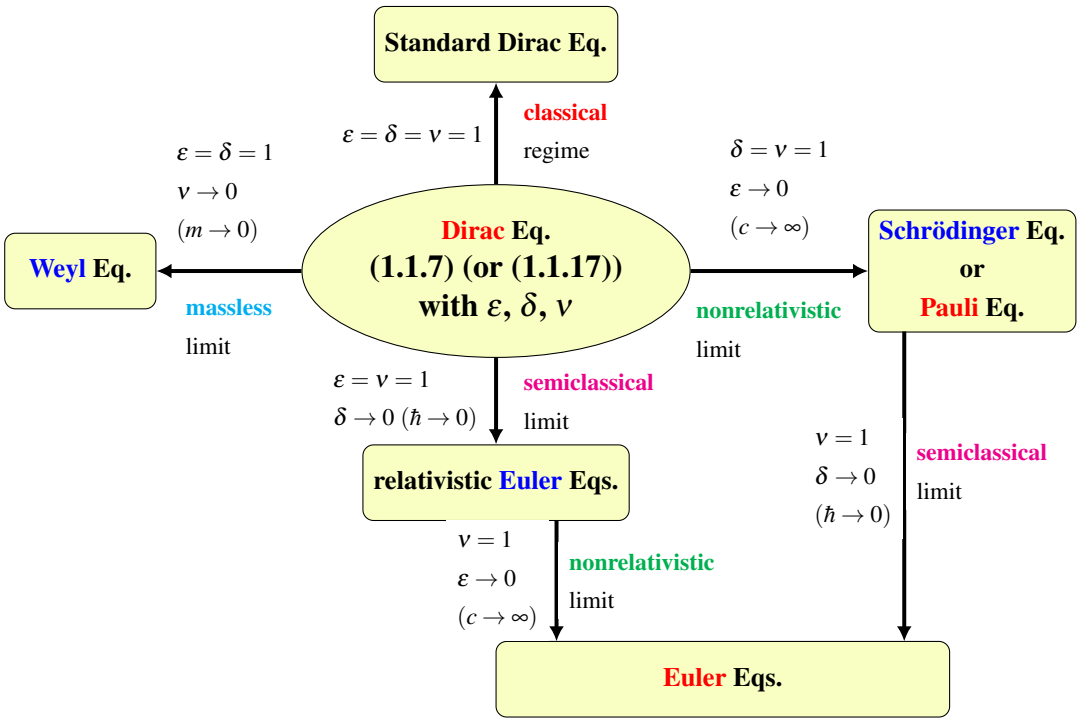

Figure 1.1.1: Diagram of different parameter regimes and limits of the Dirac equation (1.1.7) (or (1.1.17)).

- Standard (or classical) regime, i.e. $\varepsilon = \delta = \nu = 1$ ($\Longleftrightarrow m_s = m$, $x_s = \frac{\hbar}{mc}$, and $t_s = \frac{\hbar}{mc^2}$). In this regime, the wave velocity is at the order of the speed of light. The dispersion relation (1.1.16) (or (1.1.21)) suggests that in this case $\omega(\mathbf{k}) = O(1)$ when $|\mathbf{k}| = O(1)$, and thus the solution propagates waves with wavelength at $O(1)$ in space and time. In addition, if the initial data $\Psi_0 = O(1)$ in (1.1.8) (or $\Phi_0 = O(1)$ in (1.1.18)), then the solution $\Psi = O(1)$ of (1.1.7) with (1.1.8) (or $\Phi = O(1)$ of (1.1.17) with (1.1.18)), which implies that the probability density $\rho = O(1)$, current density $\mathbf{J} = O(1)$, and the energy $E(\Psi(t,\cdot)) = O(1)$. For the classical regime of the Dirac equation, there have been extensive analytical and numerical studies in the literatures. In the analytical aspect, we refer to [47, 48, 56, 71, 72, 101] and references therein for the existence and multiplicity of bound states and/or standing wave solutions. In the numerical part, there are many efficient and accurate numerical methods [4], including the finite difference time domain (FDTD) methods [5, 36, 97], time-splitting Fourier pseudospectral (TSFP) method [15, 80], exponential wave integrator Fourier pseudospectral (EWI-FP) method

[15], the Gaussian beam method [126], etc.

- Massless regime, i.e. $\varepsilon = \delta = 1$ and $0 < \nu \ll 1$ ($\Longleftrightarrow x_s = \frac{\hbar}{m_s c}$ and $t_s = \frac{\hbar}{m_s c^2}$). In this regime, the mass of the particle is much less than the mass unit. When $\nu \to 0$, the Dirac equation (1.1.7) (or (1.1.17)) converges to the Weyl equation [98, 129] with linear convergence rate in terms of ν. Any numerical methods for the Dirac equation (1.1.7) (or (1.1.17)) in the standard regime can be extended to apply in this parameter regime.

- Nonrelativistic regime, i.e. $\delta = \nu = 1$ and $0 < \varepsilon \ll 1$ ($\Longleftrightarrow m_s = m$ and $t_s = \frac{m x_s^2}{\hbar}$). In this regime, the wave speed is much less than the speed of light. From the dispersion relation (1.1.16) (or (1.1.21)), this regime suggests $\omega(\mathbf{k}) = \varepsilon^{-2} + O(1)$ when $|\mathbf{k}| = O(1)$, and thus the solution propagates waves with wavelength at $O(\varepsilon^2)$ and $O(1)$ in time and space, respectively, when $0 < \varepsilon \ll 1$. In addition, if the initial data $\Psi_0 = O(1)$ in (1.1.8) (or $\Phi_0 = O(1)$ in (1.1.18)), then the solution $\Psi = O(1)$ of (1.1.7) with (1.1.8) (or $\Phi = O(1)$ of (1.1.17) with (1.1.18)), which implies that the probability density $\rho = O(1)$, current density $\mathbf{J} = O(\varepsilon^{-1})$ and the energy $E(\Psi(t,\cdot)) = O(\varepsilon^{-2})$. The highly oscillatory nature of the solution in time and the unboundedness of the energy bring significant difficulty in mathematical analysis and numerical simulation of the Dirac equation in the nonrelativistic regime, i.e. $0 < \varepsilon \ll 1$. It is proved that the Dirac equation (1.1.7) (or (1.1.17)) converges – 'singularly' – to the Pauli equation [29, 81] and/or the Schrödinger equation [7, 29] when $\varepsilon \to 0^+$ through diagonalizing the Dirac operator and using proper ansatz. Rigorous error estimates have been established for the FDTD, TSFP and EWI-FP methods in this parameter regime [15]. The error bounds depend explicitly on the mesh size h, time step τ and the small parameter ε. Recently, a uniformly accurate multiscale time integrator pseudospectral method was proposed and analyzed for the Dirac equation in the nonrelativistic regime, which converges uniformly with respect to $\varepsilon \in (0,1]$ [14, 90], making the time step sizes independent of the small parameter ε.

- Semiclassical regime, i.e. $\varepsilon = \nu = 1$ and $0 < \delta \ll 1$ ($\Longleftrightarrow m_s = m$ and $t_s = \frac{x_s}{c}$), where the quantum effect is neglected. In this regime, the solution propagates waves with wavelength at $O(\delta)$ in space and time [32] when $0 < \delta \ll 1$. In addition, if the initial

data $\Psi_0 = O(1)$ in (1.1.8) (or $\Phi_0 = O(1)$ in (1.1.18)), then the solution $\Psi = O(1)$ of (1.1.7) with (1.1.8) (or $\Phi = O(1)$ of (1.1.17) with (1.1.18)), which implies that the probability density $\rho = O(1)$, current density $\mathbf{J} = O(1)$ and the energy $E(\Psi(t,\cdot)) = O(1)$. Similar to the nonrelativistic regime, the severe oscillation of the solution in time and space makes it difficult to carry out the mathematical analysis and numerical simulation of the Dirac equation in the semiclassical regime. When $\delta \to 0$, the Dirac equation (1.1.7) (or (1.1.17)) converges – 'singularly' – to the relativistic Euler equations [6, 70, 112]. The convergence could be proved by using the Wigner transformation method. It is an interesting question to establish rigorous error estimates of different numerical methods for the Dirac equation in the semiclassical regime, just as the case for the Schrödinger equation [3, 11, 21, 22, 39, 40, 86]. Specifically, it is meaningful to find out the dependence of the mesh size h and time step τ on the small parameter $\delta \in (0,1]$.

- Simultaneously nonrelativistic and massless regime, i.e. $\delta = 1$, $\nu \sim \varepsilon$ and $0 < \varepsilon \ll 1$ ($\Longleftrightarrow t_s = \frac{m_s x_s^2}{\hbar}$). In this regime, the wave speed is much less than the speed of light and the mass of the particle is much less than the mass unit. Here we assume $\nu = \nu_0 \varepsilon$ with $\nu_0 > 0$ a constant independent of $\varepsilon \in (0,1]$. In this case, the Dirac equation (1.1.7) can be re-written as ($d = 1,2,3$)

$$i\partial_t \Psi = \left(-i\frac{1}{\varepsilon} \sum_{j=1}^{d} \alpha_j \partial_j + \frac{\nu_0}{\varepsilon} \beta \right) \Psi + \left(V(t,\mathbf{x}) I_4 - \sum_{j=1}^{d} A_j(t,\mathbf{x}) \alpha_j \right) \Psi, \quad \mathbf{x} \in \mathbb{R}^d, \tag{1.1.22}$$

and similarly, the Dirac equation (1.1.17)) can be re-written as ($d = 1,2$)

$$i\partial_t \Phi = \left(-i\frac{1}{\varepsilon} \sum_{j=1}^{d} \sigma_j \partial_j + \frac{\nu_0}{\varepsilon} \sigma_3 \right) \Phi + \left(V(t,\mathbf{x}) I_2 - \sum_{j=1}^{d} A_j(t,\mathbf{x}) \sigma_j \right) \Phi, \quad \mathbf{x} \in \mathbb{R}^d. \tag{1.1.23}$$

In this parameter regime, formally the dispersion relation (1.1.16) (or (1.1.21)) suggests $\omega(\mathbf{k}) = O(\varepsilon^{-1})$ when $|\mathbf{k}| = O(1)$ and thus the solution propagates waves with wavelength at $O(\varepsilon)$ and $O(1)$ in time and space respectively when $0 < \varepsilon \ll 1$. In addition, if the initial data $\Psi_0 = O(1)$ in (1.1.8) (or $\Phi_0 = O(1)$ in (1.1.18)), then the solution $\Psi = O(1)$ of (1.1.22) with (1.1.8) (or $\Phi = O(1)$ of (1.1.23) with (1.1.18)), which implies that the probability density $\rho = O(1)$, current density $\mathbf{J} = O(\varepsilon^{-1})$ and the energy

$E(\Psi(t,\cdot)) = O(\varepsilon^{-1})$. Again, because of the difficulty in analysis and numerical simulation of the Dirac equation in this regime, which is caused by the highly oscillatory nature of the solution in time and the unboundedness of the energy, it is worthwhile to study the singular limit of the Dirac equation (1.1.22) (or (1.1.23)) when $\varepsilon \to 0^+$ and establish rigorous error estimates of various numerical methods for it. Specifically, we could try to find out the explicit dependence of the mesh size h and time step size τ on the small parameter $\varepsilon \in (0,1]$.

1.2 Relation to the Weyl and Majorana equations

In the representation of the Dirac equation (1.1.1), if there is no external electromagnetic potentials, then by taking natural units ($\hbar = c = 1$), the equation can be expressed in a compact form

$$\left(i \sum_{k=0}^{3} \gamma^k \partial_k - m \right) \Psi = 0, \quad \mathbf{x} \in \mathbb{R}^3. \tag{1.2.1}$$

In this expression, ∂_0 is used to represent ∂_t, and ∂_1, ∂_2, ∂_3 have the same meaning as in (1.1.1). The γ matrices are

$$\gamma^0 = \begin{pmatrix} I_2 & 0 \\ 0 & -I_2 \end{pmatrix}, \quad \gamma^j = \begin{pmatrix} 0 & \sigma_j \\ -\sigma_j & 0 \end{pmatrix}, \quad j = 1,2,3, \tag{1.2.2}$$

where $\sigma_j \, (j = 1,2,3)$ are still the Pauli matrices defined in (1.1.3).

Actually, to describe a fermion field, the γ-matrices could also be taken in other forms as long as they satisfy the two requirements [99]

$$\begin{cases} [\gamma^j, \gamma^k]_+ = 2 g_{jk} I_4 \\ \gamma_0 \gamma_k \gamma_0 = (\gamma_k)^* \end{cases} \tag{1.2.3}$$

which are obtained from the energy-momentum relation, and the requirement that the Hamiltonian should be Hermitian, respectively. In (1.2.3), the anticommutator of two elements A, B is defined as $[A,B]_+ = AB + BA$; g indicates the Minkowski metric with signature $(+---)$, which means it could be seen as a 4×4 matrix with diagonal line $(1,-1,-1,-1)$; g_{jk} is the element on the j-th row and k-th column of g in the matrix form; $\gamma_k = g_{kk}\gamma^k$, which gives

$$\gamma_0 = \gamma^0, \quad , \gamma_j = -\gamma^j, \quad j = 1,2,3. \tag{1.2.4}$$

The γ-matrices in (1.2.2) is one possible choice satisfying (1.2.3). It is called *the Dirac representation* of the γ-matrices, as it results in the Dirac equation (1.1.1). Besides this choice, there are also other two meaningful representations which are named *the Weyl representation* and *the Majorana representation* respectively, because from these two set of γ-matrices, Weyl equation and Majorana equation could be derived. In this sense, the Dirac equation, Weyl equation and Majorana equation are closely related through the unified equation (1.2.1).

- **The Weyl representation**

  The choice of the γ-matrices

  $$\gamma^0 = \begin{pmatrix} 0 & I_2 \\ I_2 & 0 \end{pmatrix}, \quad \gamma^j = \begin{pmatrix} 0 & \sigma_j \\ -\sigma_j & 0 \end{pmatrix}, \quad j = 1,2,3, \tag{1.2.5}$$

  is called *the Weyl representation* or *the chiral representation* [99]. Plugging (1.2.5) into (1.2.1), and taking $m = 0$ for massless particles, we could get two decoupled equations

  $$\partial_t \Psi_+(t,\mathbf{x}) + \sum_{j=1}^{3} \sigma_3 \partial_j \Psi_+(t,\mathbf{x}) = 0, \quad \mathbf{x} \in \mathbb{R}^3, \tag{1.2.6}$$

  $$\partial_t \Psi_-(t,\mathbf{x}) - \sum_{j=1}^{3} \sigma_3 \partial_j \Psi_-(t,\mathbf{x}) = 0, \quad \mathbf{x} \in \mathbb{R}^3, \tag{1.2.7}$$

  where $\Psi_+(t,\mathbf{x})$, $\Psi_-(t,\mathbf{x})$ respectively correspond to the upper two and lower two components of $\Psi(t,\mathbf{x})$, i.e. $(\Psi_+^T(t,\mathbf{x}), \Psi_-^T(t,\mathbf{x}))^T = \Psi(t,\mathbf{x})$. Equations (1.2.6) and (1.2.7) are both Weyl equations, (1.2.6) describes right-handed Weyl spinors, while (1.2.7) decribes left-handed Weyl spinors.

- **The Majorana representation**

  It could be noticed that if all the non-zero elements in the γ-matrices are purely imaginary, then with real initial conditions, the solution to (1.2.1) would always be real. In fact, there is such a choice of γ-matrices:

  $$\begin{aligned} &\gamma^0 = \begin{pmatrix} 0 & \sigma_2 \\ \sigma_2 & 0 \end{pmatrix}, \quad \gamma^1 = \begin{pmatrix} i\sigma_1 & 0 \\ 0 & i\sigma_1 \end{pmatrix}, \\ &\gamma^2 = \begin{pmatrix} 0 & \sigma_2 \\ -\sigma_2 & 0 \end{pmatrix}, \quad \gamma^3 = \begin{pmatrix} i\sigma_3 & 0 \\ 0 & i\sigma_3 \end{pmatrix}. \end{aligned} \tag{1.2.8}$$

By taking this set of γ-matrices in (1.2.1), we would get an equation equivalent to the Majorana equation [99], and this choice of γ-matrices is called *the Majorana representation*. The Majorana equation depicts Majorana fermions, which are quantum particles which are their own antiparticles. As a result, their wave function should always remain real with real initial conditions, which corresponds with the property of (1.2.1) with (1.2.8).

The three choices of γ-matrices mentioned (1.2.2), (1.2.5) and (1.2.8) are just frequently used ones out of infinitely many choices satisfying (1.2.3). There is definitely some connection among different sets of the matrices. Actually, there is a theorem stating that any two of the choices are related by a similarity transformation with a unitary matrix [83]. More detailedly, suppose $\{\gamma_k | k = 0,1,2,3\}$ and $\{\tilde{\gamma}_k | k = 0,1,2,3\}$ are two representations of the γ-matrices satisfying (1.2.3), then there exists a unitary matrix U, such that

$$\gamma^k = U\tilde{\gamma}^k U^*, \quad k = 0,1,2,3, \tag{1.2.9}$$

and the respective solutions Ψ and $\tilde{\Psi}$ are related by

$$\Psi(t,\mathbf{x}) = U\tilde{\Psi}(t,\mathbf{x}), \quad t > 0, \quad \mathbf{x} \in \mathbb{R}^3, \tag{1.2.10}$$

which can easily be checked from (1.2.1).

1.3 The nonlinear Dirac equation

To meet the need of simulating self-interacting Dirac fermions [64, 67, 111, 122], the nonlinear Dirac equation (NLDE) was introduced in 1938 [82], which has the form [50, 64, 67, 73, 74, 111, 122]

$$i\hbar\partial_t\Psi = \Big[-ic\hbar\sum_{j=1}^{3}\alpha_j\partial_j + mc^2\beta\Big]\Psi + e\Big[V(t,\mathbf{x})I_4 - \sum_{j=1}^{3}A_j(t,\mathbf{x})\alpha_j\Big]\Psi + \mathbf{F}(\Psi)\Psi, \quad \mathbf{x} \in \mathbb{R}^3. \tag{1.3.1}$$

The nonlinear Dirac equation (1.3.1) is similar to the Dirac equation (1.1.1) except for the nonlinear term $\mathbf{F}(\Psi)$. The nonlinearity is introduced for self-interaction, and in the resulting field equations, it is cubic with respect to the wave function, which is only significant

at extremely high densities.There have been different cubic nonlinearities generated from different applications [64, 67, 73, 74, 111, 122, 128]. Here we take $\mathbf{F}(\Psi) = g_1(\Psi^*\beta\Psi)\beta + g_2|\Psi|^2 I_4$ with $g_1, g_2 \in \mathbb{R}$ two constants and $\Psi^* = \overline{\Psi}^T$, while $\overline{f}$ denotes the complex conjugate of f. The first term, i.e. $g_2 = 0$ and $g_1 \neq 0$ is motivated from the Soler model in quantum field theory [64, 67, 111, 122], and the second term i.e. $g_1 = 0$ and $g_2 \neq 0$ is generated from BECs with a chiral confinement and/or spin-orbit coupling [42, 73, 74]. A remark is given here that our numerical methods and their error estimates in this thesis can be easily extended to the NLDE with other nonlinearities [102, 105, 122].

In fact, the NLDE has also been proposed in the Einstein-Cartan-Sciama-Kibble theory of gravity in order to extend general relativity to matter with intrinsic angular momentum (spin) [78]. And recently, the NLDE has been adapted as a mean field model for Bose-Einstein condensates (BECs) [42, 73, 74] and/or cosmology [102]. Moreover, the experimental advances in BECs, graphene and other 2D materials have also stimulated the research interests on the mathematical analysis and numerical simulations of the Dirac equation and/or the NLDE without/with electromagnetic potentials, especially the honeycomb lattice potential [2, 59, 61].

Similar to the process for Dirac equation [15], through a proper nondimensionalization (with the choice of x_s, $t_s = \frac{mx_s^2}{\hbar}$, $A_s = \frac{mv^2}{e}$ and $\psi_s = x_s^{-3/2}$ as the dimensionless length unit, time unit, potential unit and spinor field unit, respectively) and dimension reduction [15], we can obtain the dimensionless NLDE in d-dimensions ($d = 3, 2, 1$)

$$i\partial_t\Psi = \Big[-\frac{i}{\varepsilon}\sum_{j=1}^{d}\alpha_j\partial_j + \frac{1}{\varepsilon^2}\beta\Big]\Psi + \Big[V(t,\mathbf{x})I_4 - \sum_{j=1}^{d}A_j(t,\mathbf{x})\alpha_j\Big]\Psi + \mathbf{F}(\Psi)\Psi, \quad \mathbf{x}\in\mathbb{R}^d, \tag{1.3.2}$$

where ε is a dimensionless parameter inversely proportional to the light speed given by

$$0 < \varepsilon := \frac{x_s}{t_s c} = \frac{v}{c} \leq 1, \tag{1.3.3}$$

with $v = \frac{x_s}{t_s}$ the wave speed, and

$$\mathbf{F}(\Psi) = \lambda_1(\Psi^*\beta\Psi)\beta + \lambda_2|\Psi|^2 I_4, \qquad \Psi \in \mathbb{C}^4, \tag{1.3.4}$$

where $\lambda_1 = \frac{g_1}{mv^2x_s^3} \in \mathbb{R}$ and $\lambda_2 = \frac{g_2}{mv^2x_s^3} \in \mathbb{R}$ are two dimensionless constants for the interaction strength.

To study the dynamics, we give the initial condition

$$\Psi(t=0,\mathbf{x})=\Psi_0(\mathbf{x}),\qquad \mathbf{x}\in\mathbb{R}^d.$$

The NLDE (1.3.2) is dispersive and time symmetric [127]. Similar to the case for the Dirac equation, after introducing proper total probability density ρ as well as the current density $\mathbf{J}(t,\mathbf{x})=(J_1(t,\mathbf{x}),J_2(t,\mathbf{x})$, we could get the conservation law (1.1.14). Moreover, the NLDE (1.3.2) conserves the total mass. The energy is conserved if the electromagnetic potentials are time-independent, i.e. if $V(t,\mathbf{x})=V(\mathbf{x})$ and $A_j(t,\mathbf{x})=A_j(\mathbf{x})$ for $j=1,2,3$, then

$$\begin{aligned} E(t) &:= \int_{\mathbb{R}^d}\left[-\frac{i}{\varepsilon}\sum_{j=1}^{d}\Psi^*\alpha_j\partial_j\Psi+\frac{1}{\varepsilon^2}\Psi^*\beta\Psi+V(\mathbf{x})|\Psi|^2+G(\Psi)-\sum_{j=1}^{d}A_j(\mathbf{x})\Psi^*\alpha_j\Psi\right]d\mathbf{x}\\ &\equiv E(0),\qquad t\geq 0, \end{aligned}\tag{1.3.5}$$

where

$$G(\Psi)=\frac{\lambda_1}{2}\left(\Psi^*\beta\Psi\right)^2+\frac{\lambda_2}{2}|\Psi|^4,\qquad \Psi\in\mathbb{C}^4.\tag{1.3.6}$$

In (1.3.2), if the external electromagnetic potentials are taken to be constants, i.e. $V(t,\mathbf{x})\equiv V^0$ and $A_j(t,\mathbf{x})\equiv A_j^0$ for $j=1,2,3$, then the NLDE (1.3.2) admits the plane wave solution as $\Psi(t,\mathbf{x})=\mathbf{B}e^{i(\mathbf{k}\cdot\mathbf{x}-\omega t)}$, where the time frequency ω, amplitude vector $\mathbf{B}\in\mathbb{R}^4$ and spatial wave number $\mathbf{k}=(k_1,\ldots,k_d)^T\in\mathbb{R}^d$ satisfy

$$\omega\mathbf{B}=\left[\sum_{j=1}^{d}\left(\frac{k_j}{\varepsilon}-A_j^0\right)\alpha_j+\frac{1}{\varepsilon^2}\beta+V^0I_4+\lambda_1\left(\mathbf{B}^*\beta\mathbf{B}\right)\beta+\lambda_2|\mathbf{B}|^2I_4\right]\mathbf{B},\tag{1.3.7}$$

which immediately gives the *dispersion relation* of the NLDE (1.3.2) as

$$\omega:=\omega(\mathbf{k},\mathbf{B})=V^0+\lambda_2|\mathbf{B}|^2\pm\frac{1}{\varepsilon^2}\sqrt{\left[1+\varepsilon^2\lambda_1\left(\mathbf{B}^*\beta\mathbf{B}\right)\right]^2+\varepsilon^2\left|\mathbf{k}-\varepsilon\mathbf{A}^0\right|^2},\qquad \mathbf{k}\in\mathbb{R}^d.\tag{1.3.8}$$

Again, similar to the Dirac equation [15], for one dimension (1D) and two dimensions (2D), the NLDE (1.3.2) can be simplified to the following one [64, 67, 111]

$$i\partial_t\Phi=\Big[-\frac{i}{\varepsilon}\sum_{j=1}^{d}\sigma_j\partial_j+\frac{1}{\varepsilon^2}\sigma_3\Big]\Phi+\Big[V(t,\mathbf{x})I_2-\sum_{j=1}^{d}A_j(t,\mathbf{x})\sigma_j\Big]\Phi+\mathbf{F}(\Phi)\Phi,\quad \mathbf{x}\in\mathbb{R}^d,\tag{1.3.9}$$

where

$$\mathbf{F}(\Phi) = \lambda_1 \left(\Phi^* \sigma_3 \Phi\right) \sigma_3 + \lambda_2 |\Phi|^2 I_2, \qquad \Phi \in \mathbb{C}^2, \tag{1.3.10}$$

with λ_1 and λ_2 both real numbers. In (1.3.9), the two-component wave function Φ is defined as $\Phi := \Phi(t,\mathbf{x}) = (\phi_1(t,\mathbf{x}), \phi_2(t,\mathbf{x}))^T \in \mathbb{C}^2$. The initial condition for dynamics is given as

$$\Phi(t=0,\mathbf{x}) = \Phi_0(\mathbf{x}), \qquad \mathbf{x} \in \mathbb{R}^d. \tag{1.3.11}$$

The NLDE (1.3.9) has similar properties to its four-component version (1.3.2). It is dispersive and time symmetric, satisfies the conservation law (1.1.14) [34], conserves total mass, and also conserves energy

$$\begin{aligned} E(t) &:= \int_{\mathbb{R}^d} \left(-\frac{i}{\varepsilon} \sum_{j=1}^d \Phi^* \sigma_j \partial_j \Phi + \frac{1}{\varepsilon^2} \Phi^* \sigma_3 \Phi + V(\mathbf{x}) |\Phi|^2 - \sum_{j=1}^d A_j(\mathbf{x}) \Phi^* \sigma_j \Phi + G(\Phi) \right) d\mathbf{x} \\ &\equiv E(0), \qquad t \geq 0, \end{aligned} \tag{1.3.12}$$

where

$$G(\Phi) = \frac{\lambda_1}{2} \left(\Phi^* \sigma_3 \Phi\right)^2 + \frac{\lambda_2}{2} |\Phi|^4, \qquad \Phi \in \mathbb{C}^2. \tag{1.3.13}$$

if the electromagnetic potentials are time-independent.

Under constant external electromagnetic potentials, i.e. $V(t,\mathbf{x}) \equiv V^0$ and $A_j(t,\mathbf{x}) \equiv A_j^0$ for $j = 1,2$, the NLDE (1.3.9) admits the plane wave solution as $\Phi(t,\mathbf{x}) = \mathbf{B}\, e^{i(\mathbf{k}\cdot\mathbf{x} - \omega t)}$, with the time frequency ω, amplitude vector $\mathbf{B} \in \mathbb{R}^2$ and spatial wave number $\mathbf{k} = (k_1, \ldots, k_d)^T \in \mathbb{R}^d$ satisfy

$$\omega \mathbf{B} = \Big[\sum_{j=1}^d \left(\frac{k_j}{\varepsilon} - A_j^0 \right) \sigma_j + \frac{1}{\varepsilon^2} \sigma_3 + V^0 I_2 + \lambda_1 \left(\mathbf{B}^* \sigma_3 \mathbf{B}\right) \sigma_3 + \lambda_2 |\mathbf{B}|^2 I_2 \Big] \mathbf{B}. \tag{1.3.14}$$

which implies the *dispersion relation* of the NLDE (1.3.9) directly as

$$\omega := \omega(\mathbf{k}, \mathbf{B}) = V^0 + \lambda_2 |\mathbf{B}|^2 \pm \frac{1}{\varepsilon^2} \sqrt{\left[1 + \varepsilon^2 \lambda_1 \left(\mathbf{B}^* \sigma_3 \mathbf{B}\right)\right]^2 + \varepsilon^2 \left|\mathbf{k} - \varepsilon \mathbf{A}^0\right|^2}, \qquad \mathbf{k} \in \mathbb{R}^d. \tag{1.3.15}$$

The NLDE (1.3.2) (or (1.3.9)) has different regimes with different choices of the dimensionless parameter ε. When $\varepsilon = 1$, which corresponds to the classical regime, extensive analytical and numerical results have been obtained in the literature. We refer to [8, 9, 28, 41, 53, 54, 55, 88] and references therein for the existence and multiplicity of bound states

and/or standing wave solutions, Specifically, when $\varepsilon = 1$ with $d = 1$, $V(t,x) \equiv A_1(t,x) \equiv 0$ in (1.3.9), and in the nonlinearity (1.3.10), $\lambda_1 = -1$ and $\lambda_2 = 0$ is taken , the NLDE (1.3.9) gives soliton solutions with explicit form derived in [46, 67, 75, 89, 93, 100, 114, 119]. On the other hand, we refer to[23, 33, 34, 77, 79, 80, 97, 106, 108] and references therein for the numerical methods and comparison. The numerical methods include the finite difference time domain (FDTD) methods [34, 77, 97], time-splitting Fourier spectral (TSFP) methods [23, 33, 66, 80], and Runge-Kutta discontinuous Galerkin methods [107, 125, 127].

When $0 < \varepsilon \ll 1$, this is the nonrelativistic regime. In this case, there has not been much work on the analysis and computation of the NLDE (1.3.2) (or (1.3.9)). This is because as indicated by the dispersion relation (1.1.16) (or (1.3.14)), the solution of the NLDE propogates waves with wavelength $O(\varepsilon^2)$ and $O(1)$ respectively in time and space, i.e. the solution is highly oscillatory in time. Furthermore, the corresponding energy functionals (1.3.5) and (1.3.12) are indefinite [29, 55] and would become unbounded when $\varepsilon \to 0$. Recently, several numerical methods were applied to the NLDE and the error estimates were carried out [16]. The methods include the finite difference time domain (FDTD) methods, the exponential wave integrator Fourier pseudospectral (EWI-FP) method, and the time-splitting Fourier pseudospectral (TSFP) method. To overcome the strict dependency of the time step size on ε, uniform accurate (UA) schemes with two-scale formulation approach [90] or multiscale time integrator pseudospectral method [35] were also designed for NLDE in the nonrelativistic regime.

1.4 Problems to study

As is pointed out in the previous sections, although there has been much effort devoted to the study of the Dirac equation, there still lacks thorough understanding of it in different regimes. This motivates us to design new efficient and accurate numerical methods to solve the Dirac equation in different regimes, and establish the error estimates. Specifically, the purposes of the thesis are:

- Find out a proper splitting of the Dirac operator so that a new fourth order compact splitting method (S_{4c}) could be designed and applied to the Dirac equation. Compare

the performance of S_{4c} with other fourth-order splitting methods in efficiency and accuracy. Test numerically the error bounds of S_{4c} for the Dirac equation in different regimes. Moreover, extend the method to the case of time-dependent electromagnetic potentials.

- Give rigorous proof for the super-resolution property, which is observed through extensive numerical tests, of the splitting methods in solving the Dirac and nonlinear Dirac equation in the nonrelativistic regime without external magnetic potential.
- Apply several finite difference time domain methods to the Dirac equation in the semiclassical regime. Prove rigorously the error estimates and validate them through numerical examples.

1.5 Structure and scope of the thesis

The thesis is organized as follows.

Chapter 2 proposes a new fourth-order compact time-splitting (S_{4c}) Fourier pseudospectral method for the Dirac equation. It is applied through splitting the Dirac equation into two parts and introducing a double commutator between them to help reduce computational cost. This method successfully cuts down the number of sub-steps in S_{4c}, compared to the standard fourth-order splitting and the fourth-order partitioned Runge-Kutta splitting. Comparison in accuracy, efficiency as well as long time behavior among S_{4c} and many other existing time-splitting methods for the Dirac equation are carried out. The error bounds and the spatial/temporal resolutions of S_{4c} are also inferred for the Dirac equation in different parameter regimes including the nonrelativistic regime and the semiclassical regime. Furthermore, extension to time-dependent potentials is also considered by using the time-ordering operator.

In Chapter 3, super-resolution of the time-splitting methods for the Dirac equation in the absence of external magnetic potentials in the nonrelativistic regime is studied. Specifically, the first-order splitting S_1 and the second-order splitting S_2 are examined carefully. These methods surprisingly break the resolution constraint under the Shannon's sampling theorem, as they can capture the solution accurately with much larger time step sizs τ than the sampled

wavelength. Rigorous error estimates and proof for all time step sizes and non-resonant time step sizes are established respectively, and they are verified through numerical examples. The error bounds in the full-discretization form are also given and proved.

Chapter 4 deals with the super-resolution of splitting methods for the nonlinear Dirac equation in the nonrelativistic regime, still without magnetic potentials. The results are similar to those in the linear case, but the proof is quite different because of the nonlinearity. Furthermore, it is noticed that super-resolution also holds true for higher-order splitting methods. Numerical results are presented to give an intuitive understanding.

Chapter 5 is devoted to studying rigorously the Dirac equation in the semiclassical regime, a small dimensionless parameter $0 < \delta \leq 1$ representing the scaled Planck constant. In this regime, there are highly oscillatory propogating waves with wavelength $O(\delta)$ in both time and space of the solution. Four frequently-used finite difference time domain (FDTD) methods are applied, and their error estimates are rigorously proved. Numerical tests are carried out to support the error estimates.

Finally, the conclusions are drawn in Chapter 6, and some possible future work is also put forward.

Research in this thesis may give new insights into the regimes not well studied for Dirac equation, including the nonrelativistic regime, the semiclassical regime, and the simultaneous nonrelativistic and massless regime. It would also improve the computational efficiency for Dirac equation with small parameters, especially in the absence of external magnetic potential.

This thesis mainly deals with Dirac equation in 1D. Extension to 2D and 3D is briefly mentioned, but the details are omitted for concision. Furthermore, during computation, a bounded domain with periodic boundary conditions is always assumed, which is an acceptable approximation of the real domain for highly centered wave functions.

Chapter 2

A Fourth-Order Compact Time-splitting Method

To solve the Dirac equation (1.1.7) (or (1.1.17)), first- and second-order time-splitting spectral methods have been applied and analyzed [15]. The splitting methods could be straightforwardly extended to higher order, e.g. fourth-order methods [24, 94, 116], such as the standard fourth-order splitting (S_4) [65, 118, 130] and the fourth-order partitioned Runge-Kutta ($S_{4\mathrm{RK}}$) splitting [30, 69]. However, as has been observed in the literature [94], S_4 has to use negative time step in at least one of the sub-problems at each time interval [65, 118, 130], which causes some drawbacks in practical computation, and the number of sub-problems in $S_{4\mathrm{RK}}$ at each time interval is much bigger than that of S_4 [30], which increases the computational cost at each time step a lot.

In this chapter, we introduce a fourth-order compact time-splitting method ($S_{4\mathrm{c}}$) for (1.1.7) (and (1.1.17)), in order to overcome the above mentioned problems caused by S_4 and $S_{4\mathrm{RK}}$. We first give a brief review of the time-splitting methods, and then show the detailed computation of the double commutator for (1.1.7) (and/or (1.1.17)), which is the key point in applying $S_{4\mathrm{c}}$. The full discretization and properties of $S_{4\mathrm{c}}$ are discussed, and comparison among it and other splitting methods is carried out through numerical examples. Finally, we show the extension to the equation with time-dependent potentials.

2.1 Review of different time-splitting schemes

The splitting technique introduced by Trotter in 1959 [123] has been widely applied in analysis and numerical simulation [3, 21, 22, 39, 40, 94], especially in computational quantum physics. For details, we refer to [115, 116, 117] and references therein. In the Hamiltonian system and general ordinary differential equations (ODEs), the splitting approach has been shown to preserve the structural/geometric properties [76, 124] and is superior in many applications. Developments of splitting type methods in solving partial differential equations (PDEs) include utilization in Schrödinger/nonlinear Schrödinger equations [3, 21, 22, 39, 40, 92, 120], Dirac/nonlinear Dirac equations [15, 16, 27, 91], Maxwell-Dirac system [23, 80], Zakharov system [24, 25, 26, 68, 85, 87], Stokes equation [38], and Enrenfest dynamics [57], etc.

To review the frequently used time-splitting schemes for integrating differential equations, we introduce a model equation ($d = 1, 2, 3$)

$$\partial_t u(t, \mathbf{x}) = (T + W)u(t, \mathbf{x}), \quad t > 0, \quad \mathbf{x} \in \mathbb{R}^d, \tag{2.1.1}$$

with the initial data

$$u(0, \mathbf{x}) = u_0(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^d, \tag{2.1.2}$$

where T and W are two time-independent operators. For any time step $\tau > 0$, the solution of (2.1.1) with (2.1.2) can be formally represented as

$$u(\tau, \mathbf{x}) = e^{\tau(T+W)} u_0(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^d. \tag{2.1.3}$$

A splitting (or split-step or time-splitting) scheme can be designed by approximating the operator $e^{\tau(T+W)}$ by a sequential product of $e^{\tau T}$ and $e^{\tau W}$ [118, 130], i.e.

$$e^{\tau(T+W)} \approx \Pi_{j=1}^{n} e^{a_j \tau T} e^{b_j \tau W}, \tag{2.1.4}$$

where $n \geq 1$, a_j, $b_j \in \mathbb{R}$ ($j = 1, \ldots, n$) are to be determined so that the approximation has certain order of accuracy in terms of τ [118, 130]. Without loss of generality, here we suppose that the computation of $e^{\tau W}$ is easier and/or more efficient than of $e^{\tau T}$.

2.1.1 First- and second-order time-splitting schemes

Taking $n = 1$ and $a_1 = b_1 = 1$ in (2.1.4), one can obtain the **first-order Lie-Trotter splitting** (S_1) as [123] ($d = 1, 2, 3$)

$$u(\tau, \mathbf{x}) \approx S_1(\tau)u_0(\mathbf{x}) := e^{\tau T} e^{\tau W} u_0(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^d. \tag{2.1.5}$$

In this method, one needs to integrate each of the operators T and W once. By using Taylor expansion, one can formally show the local truncation error as [113]

$$\|u(\tau, \mathbf{x}) - S_1(\tau)u_0(\mathbf{x})\| \leq C_1 \tau^2, \tag{2.1.6}$$

where $C_1 > 0$ is a constant independent of τ and $\|\cdot\|$ is a norm depending on the problem. As a result, S_1 is formally a first-order integrator [94].

Similarly, taking $n = 2$, $a_1 = 0$, $b_1 = \frac{1}{2}$, $a_2 = 1$ and $b_2 = \frac{1}{2}$, one can obtain the **second-order Strang splitting** (S_2) method as [113] ($d = 1, 2, 3$)

$$u(\tau, \mathbf{x}) \approx S_2(\tau)u_0(\mathbf{x}) := e^{\frac{\tau}{2} W} e^{\tau T} e^{\frac{\tau}{2} W} u_0(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^d. \tag{2.1.7}$$

In this method, one needs to integrate the operator T once and W twice. Again, by using Taylor expansion, one can formally show the local truncation error as [113]

$$\|u(\tau, \mathbf{x}) - S_2(\tau)u_0(\mathbf{x})\| \leq C_2 \tau^3, \tag{2.1.8}$$

where $C_2 > 0$ is a constant independent of τ. As a result, S_2 is formally a second-order integrator [94].

2.1.2 Fourth-order time-splitting schemes

Besides first- and second-order, high order, especially fourth-order, splitting methods for (2.1.1) with (2.1.2) via the construction (2.1.4) have been extensively studied in the literature [43, 44].

For brevity, here we only mention a popular **fourth-order Forest-Ruth (or Yoshida) splitting** (S_4) method [65, 118, 130] as ($d = 1, 2, 3$)

$$u(\tau, \mathbf{x}) \approx S_4(\tau)u_0(\mathbf{x}) := S_2(w_1\tau)S_2(w_2\tau)S_2(w_1\tau)u_0(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^d, \tag{2.1.9}$$

where

$$w_1 = \frac{1}{2-2^{1/3}}, \quad w_2 = -\frac{2^{1/3}}{2-2^{1/3}}. \tag{2.1.10}$$

In this method, the operators T and W need to be integrated three times and four times, respectively. Still by using Taylor expansion, one can formally show the local truncation error as [65]

$$\|u(\tau,\mathbf{x}) - S_4(\tau)u_0(\mathbf{x})\| \le C_4\tau^5, \tag{2.1.11}$$

where $C_4 > 0$ is a constant independent of τ. As a result, S_4 is formally a fourth-order integrator [94]. As mentioned before, due to the fact that negative time steps, e.g. $w_2 < 0$, are used in the method, in general, it cannot be applied to solve dissipative differential equations. In addition, as noticed in the literature [94], some drawbacks of the S_4 method were reported, such as the constant C_4 is usually much larger than C_1 and C_2, and the fourth-order accuracy could be observed only when τ is very small [94, 117].

To overcome the drawbacks of the S_4 method, the **fourth-order partitioned Runge-Kutta splitting** ($S_{4\mathrm{RK}}$) was proposed [30, 69] for $\mathbf{x} \in \mathbb{R}^d$ ($d = 1,2,3$) as

$$\begin{aligned} u(\tau,\mathbf{x}) &\approx S_{4\mathrm{RK}}(\tau)u_0(\mathbf{x}) \\ &:= e^{a_1\tau W}e^{b_1\tau T}e^{a_2\tau W}e^{b_2\tau T}e^{a_3\tau W}e^{b_3\tau T}e^{a_4\tau W}e^{b_3\tau T}e^{a_3\tau W}e^{b_2\tau T}e^{a_2\tau W}e^{b_1\tau T}e^{a_1\tau W}u_0(\mathbf{x}), \end{aligned} \tag{2.1.12}$$

where

$$\begin{aligned} &a_1 = 0.0792036964311957, \quad a_2 = 0.353172906049774, \\ &a_3 = -0.0420650803577195, \quad a_4 = 1-2(a_1+a_2+a_3), \\ &b_1 = 0.209515106613362, \quad b_2 = -0.143851773179818, \quad b_3 = \frac{1}{2} - (b_1+b_2). \end{aligned}$$

This method requires much more repetitions of the operators T and W. It can be easily observed that six integration of T and seven integration of W are required for each time step. Again, by using Taylor expansion, one can formally show the local truncation error as [30]

$$\|u(\tau,\mathbf{x}) - S_{4\mathrm{RK}}(\tau)u_0(\mathbf{x})\| \le \widetilde{C}_4\tau^5, \tag{2.1.13}$$

where $\widetilde{C}_4 > 0$ is a constant independent of τ. As a result, $S_{4\mathrm{RK}}$ is also formally a fourth-order integrator [94]. Although some problems caused by S_4 are solved, it is easy to see that the computational cost of $S_{4\mathrm{RK}}$ is about twice that of S_4. Meanwhile, in this method, negative time steps, e.g. $a_3 < 0$, are still not totally prevented.

	S_1	S_2	S_4	$S_{4\mathrm{RK}}$	$S_{4\mathrm{c}}$
T	1	1	3	6	2
W	1	2	4	7	3

Table 2.1.1: The numbers of operators T and W to be implemented in different time-splitting methods.

2.1.3 Fourth-order compact time-splitting schemes

In order to avoid the negative time steps, a **fourth-order gradient symplectic integrator** was proposed by S. A. Chin, motivated by the numerical integration of the Schrödinger equation [43, 44, 45] as ($d = 1,2,3$)

$$u(\tau,\mathbf{x}) \approx S_{4\mathrm{c}}(\tau)u_0(\mathbf{x}) := e^{\frac{1}{6}\tau W} e^{\frac{1}{2}\tau T} e^{\frac{2}{3}\tau \widehat{W}} e^{\frac{1}{2}\tau T} e^{\frac{1}{6}\tau W} u_0(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^d, \tag{2.1.14}$$

where

$$\widehat{W} := W + \frac{1}{48}\tau^2 [W,[T,W]], \tag{2.1.15}$$

with $[T,W] := TW - WT$ the commutator of the two operators T and W and $[W,[T,W]]$ a double commutator. Again, Taylor expansion formally gives the local truncation error as [43, 44]

$$\|u(\tau,\mathbf{x}) - S_{4\mathrm{c}}(\tau)u_0(\mathbf{x})\| \le \widehat{C}_4 \tau^5, \tag{2.1.16}$$

where $\widehat{C}_4 > 0$ is a constant independent of τ. As a result, $S_{4\mathrm{c}}$ is also a fourth-order integrator [94]. In this method, the operator T only needs to be integrated twice while the operator W needs to be integrated three times in one time step, under the assumption that the computation of $\widehat{W}$ is equivalent to that of W, which means that $S_{4\mathrm{c}}$ is much more efficient than S_4 and $S_{4\mathrm{RK}}$. In this sense, it is appropriate to name it a **fourth-order compact splitting** ($S_{4\mathrm{c}}$) since, at each time step, the number of sub-steps in it is much less than those in S_4 and $S_{4\mathrm{RK}}$. We could also observe that there is no negative time step in $S_{4\mathrm{c}}$, which serves as its advantage as well.

For comparison, Table 2.1.1 lists the numbers of integration for T and W required by different splitting methods in each time step. From the table, under the assumptions that the computation for $e^{\tau W}$ is easier and/or more efficient than that for $e^{\tau T}$ and the computation of $e^{\tau \widehat{W}}$ is similar to that for $e^{\tau W}$, we could draw the following conclusions: (i) the computational time for S_2 is almost the same as that for S_1; (ii) the computational time for $S_{4\mathrm{c}}$ is about

twice that for S_2 (or S_1); (iii) among the three fourth-order splitting methods, S_{4c} is the most efficient and S_{4RK} is the most time-consuming.

2.2 Derivation of double commutators and full discretization

Motivated by S_{4c} introduced above, a new fourth-order compact time-splitting Fourier pseudospectral method could be proposed for the Dirac equation.

In this section, we first show that the double commutator $[W,[T,W]]$ is linear in T and then compute it for the Dirac equations (1.1.17) for $d = 1,2$ and (1.1.7) for $d = 1,2,3$ with time-independent electromagnetic potentials. After that, we introduce the full discretization of S_{4c} for the Dirac equation (1.1.17) in 1D as a simple illustration.

Lemma 2.1. *Let T and W be two operators, then we have*

$$[W,[T,W]] = 2WTW - WWT - TWW. \tag{2.2.1}$$

Thus the double commutator $[W,[T,W]]$ is linear in T, i.e. for any two operators T_1 and T_2, we have

$$[W,[a_1T_1 + a_2T_2,W]] = a_1[W,[T_1,W]] + a_2[W,[T_2,W]], \qquad a_1,a_2 \in \mathbb{R}. \tag{2.2.2}$$

Proof. Noticing $[T,W] := TW - WT$, we have

$$\begin{aligned} [W,[T,W]] &= [W,(TW-WT)] = W(TW-WT) - (TW-WT)W \\ &= WTW - WWT - TWW + WTW \\ &= 2WTW - WWT - TWW. \end{aligned} \tag{2.2.3}$$

From (2.2.3), it is easy to see that the double commutator $[W,[T,W]]$ is linear in T, i.e. (2.2.2) is valid. □

2.2.1 The double commutator for 1D

Lemma 2.2. *For the Dirac equation* (1.1.17) *in 1D, i.e. $d = 1$, with time-independent potentials $V(x)$, $A_1(x)$, define*

$$T = -\frac{1}{\varepsilon}\sigma_1\partial_1 - \frac{i\nu}{\delta\varepsilon^2}\sigma_3, \qquad W = -\frac{i}{\delta}\Big(V(x)I_2 - A_1(x)\sigma_1\Big), \tag{2.2.4}$$

we have

$$[W,[T,W]] = -\frac{4i\nu}{\delta^3\varepsilon^2}A_1^2(x)\sigma_3. \tag{2.2.5}$$

Proof. Combining (2.2.4) and (2.2.2), we obtain

$$[W,[T,W]] = -\frac{1}{\varepsilon}[W,[\sigma_1\partial_1,W]] - \frac{i\nu}{\delta\varepsilon^2}[W,[\sigma_3,W]]. \tag{2.2.6}$$

Noticing (2.2.1) and (2.2.4), we have

$$\begin{aligned}
[W,[\sigma_1\partial_1,W]] &= 2\left(-\frac{i}{\delta}\big(V(x)I_2 - A_1(x)\sigma_1\big)\right)(\sigma_1\partial_1)\left(-\frac{i}{\delta}\big(V(x)I_2 - A_1(x)\sigma_1\big)\right) \\
&\quad -\left(-\frac{i}{\delta}\big(V(x)I_2 - A_1(x)\sigma_1\big)\right)^2(\sigma_1\partial_1) - (\sigma_1\partial_1)\left(-\frac{i}{\delta}\big(V(x)I_2 - A_1(x)\sigma_1\big)\right)^2 \\
&= -\frac{2}{\delta^2}\big(V(x)I_2 - A_1(x)\sigma_1\big)\sigma_1\partial_1\big(V(x)I_2 - A_1(x)\sigma_1\big) \\
&\quad +\frac{1}{\delta^2}\big(V(x)I_2 - A_1(x)\sigma_1\big)^2\sigma_1\partial_1 + \frac{1}{\delta^2}\sigma_1\partial_1\big(V(x)I_2 - A_1(x)\sigma_1\big)^2 \\
&= -\frac{2}{\delta^2}\sigma_1\big(V(x)I_2 - A_1(x)\sigma_1\big)\partial_1\big(V(x)I_2 - A_1(x)\sigma_1\big) \\
&\quad -\frac{2}{\delta^2}\sigma_1\big(V(x)I_2 - A_1(x)\sigma_1\big)^2\partial_1 + \frac{2}{\delta^2}\sigma_1\big(V(x)I_2 - A_1(x)\sigma_1\big)^2\partial_1 \\
&\quad +\frac{2}{\delta^2}\sigma_1\big(V(x)I_2 - A_1(x)\sigma_1\big)\partial_1\big(V(x)I_2 - A_1(x)\sigma_1\big) \\
&= 0.
\end{aligned} \tag{2.2.7}$$

$$\begin{aligned}
[W,[\sigma_3,W]] &= 2\left(-\frac{i}{\delta}\big(V(x)I_2 - A_1(x)\sigma_1\big)\right)\sigma_3\left(-\frac{i}{\delta}\big(V(x)I_2 - A_1(x)\sigma_1\big)\right) \\
&\quad -\left(-\frac{i}{\delta}\big(V(x)I_2 - A_1(x)\sigma_1\big)\right)^2\sigma_3 - \sigma_3\left(-\frac{i}{\delta}\big(V(x)I_2 - A_1(x)\sigma_1\big)\right)^2 \\
&= -\frac{2}{\delta^2}\big(V(x)I_2 - A_1(x)\sigma_1\big)\big(V(x)I_2 + A_1(x)\sigma_1\big)\sigma_3 + \frac{1}{\delta^2}\big(V(x)I_2 - A_1(x)\sigma_1\big)^2\sigma_3 \\
&\quad +\frac{1}{\delta^2}\big(V(x)I_2 + A_1(x)\sigma_1\big)^2\sigma_3 \\
&= -\frac{1}{\delta^2}\Big(2V^2(x)I_2 - 2A_1^2(x)I_2 - \big(V^2(x)I_2 + A_1^2(x)I_2 - 2A_1(x)V(x)\sigma_1\big) \\
&\quad -\big(V^2(x)I_2 + A_1^2(x)I_2 + 2A_1(x)V(x)\sigma_1\big)\Big)\sigma_3 \\
&= -\frac{1}{\delta^2}\big(-4A_1^2(x)I_2\big)\sigma_3 = \frac{4}{\delta^2}A_1^2(x)\sigma_3.
\end{aligned} \tag{2.2.8}$$

Plugging (2.2.7) and (2.2.8) into (2.2.6), we can obtain (2.2.5) immediately. □

Combining (2.2.5), (2.2.4) and (2.1.15), we have

$$\widehat{W} = W + \frac{1}{48}\tau^2[W,[T,W]] = -\frac{i}{\delta}\Big(V(x)I_2 - A_1(x)\sigma_1\Big) - \frac{i\nu\tau^2}{12\delta^3\varepsilon^2}A_1^2(x)\sigma_3, \tag{2.2.9}$$

which immediately implies that the computation of $e^{\tau\widehat{W}}$ is similar (or at almost the same computational cost) to that for $e^{\tau W}$ in this case.

Corollary 2.1. *For the Dirac equation* (1.1.7) *in 1D, i.e.* $d = 1$*, define*

$$T = -\frac{1}{\varepsilon}\alpha_1\partial_1 - \frac{i\nu}{\delta\varepsilon^2}\beta, \qquad W = -\frac{i}{\delta}\Big(V(x)I_4 - A_1(x)\alpha_1\Big), \tag{2.2.10}$$

we have

$$[W,[T,W]] = -\frac{4i\nu}{\delta^3\varepsilon^2}A_1^2(x)\beta. \tag{2.2.11}$$

2.2.2 The double commutators for 2D and 3D

Similar to the 1D case, we have

Lemma 2.3. *For the Dirac equation* (1.1.17) *in 2D, i.e.* $d = 2$*, with time-independent potentials, define*

$$T = -\frac{1}{\varepsilon}\sigma_1\partial_1 - \frac{1}{\varepsilon}\sigma_2\partial_2 - \frac{i\nu}{\delta\varepsilon^2}\sigma_3, \qquad W = -\frac{i}{\delta}\Big(V(\mathbf{x})I_2 - A_1(\mathbf{x})\sigma_1 - A_2(\mathbf{x})\sigma_2\Big), \tag{2.2.12}$$

we have

$$[W,[T,W]] = F_3(\mathbf{x}) + F_1(\mathbf{x})\partial_1 + F_2(\mathbf{x})\partial_2, \tag{2.2.13}$$

where

$$\begin{aligned}
F_1(\mathbf{x}) &= \frac{4}{\delta^2\varepsilon}\Big(-A_2^2(\mathbf{x})\sigma_1 + A_1(\mathbf{x})A_2(\mathbf{x})\sigma_2\Big), \quad F_2(\mathbf{x}) = \frac{4}{\delta^2\varepsilon}\Big(A_1(\mathbf{x})A_2(\mathbf{x})\sigma_1 - A_1^2(\mathbf{x})\sigma_2\Big),\\
F_3(\mathbf{x}) &= \frac{4}{\delta^2\varepsilon}\Big(A_1(\mathbf{x})\partial_2 A_2(\mathbf{x}) - A_2(\mathbf{x})\partial_1 A_2(\mathbf{x})\Big)\sigma_1 + \frac{4}{\delta^2\varepsilon}\Big(A_2(\mathbf{x})\partial_1 A_1(\mathbf{x}) - A_1(\mathbf{x})\partial_2 A_1(\mathbf{x})\Big)\sigma_2\\
&\quad + \frac{4i}{\delta^2\varepsilon}\left(A_2(\mathbf{x})\partial_1 V(\mathbf{x}) - A_1(\mathbf{x})\partial_2 V(\mathbf{x}) - \frac{\nu}{\delta\varepsilon}\big(A_1^2(\mathbf{x}) + A_2^2(\mathbf{x})\big)\right)\sigma_3.
\end{aligned}$$

Proof. Combining (2.2.12) and (2.2.2), we obtain

$$[W,[T,W]] = -\frac{1}{\varepsilon}[W,[\sigma_1\partial_1,W]] - \frac{1}{\varepsilon}[W,[\sigma_2\partial_2,W]] - \frac{i\nu}{\delta\varepsilon^2}[W,[\sigma_3,W]]. \tag{2.2.14}$$

From (1.1.3), we have

$$\begin{aligned} &\sigma_j^2 = I_2, \quad \sigma_j\sigma_l = -\sigma_l\sigma_j, \qquad 1 \le j \ne l \le 3, \\ &\sigma_1\sigma_2 = i\sigma_3, \quad \sigma_2\sigma_3 = i\sigma_1, \quad \sigma_3\sigma_1 = i\sigma_2. \end{aligned} \tag{2.2.15}$$

Noticing (2.2.12), (2.2.1) and (2.2.15), we get

$$\begin{aligned} &[W,[\sigma_1\partial_1,W]] \\ &= -\frac{1}{\delta^2}\Big(2\big(V(\mathbf{x})I_2 - A_1(\mathbf{x})\sigma_1 - A_2(\mathbf{x})\sigma_2\big)(\sigma_1\partial_1)\big(V(\mathbf{x})I_2 - A_1(\mathbf{x})\sigma_1 - A_2(\mathbf{x})\sigma_2\big) \\ &\quad -\big(V(\mathbf{x})I_2 - A_1(\mathbf{x})\sigma_1 - A_2(\mathbf{x})\sigma_2\big)^2(\sigma_1\partial_1) - (\sigma_1\partial_1)\big(V(\mathbf{x})I_2 - A_1(\mathbf{x})\sigma_1 - A_2(\mathbf{x})\sigma_2\big)^2\Big) \\ &= -\frac{2}{\delta^2}\sigma_1A_2(\mathbf{x})\sigma_2\big(\partial_1V(\mathbf{x})I_2 - \partial_1A_1(\mathbf{x})\sigma_1 - \partial_1A_2(\mathbf{x})\sigma_2\big) \\ &\quad -\frac{2}{\delta^2}\sigma_1\big(V(\mathbf{x})I_2 - A_1(\mathbf{x})\sigma_1 + A_2(\mathbf{x})\sigma_2\big)\big(V(\mathbf{x})I_2 - A_1(\mathbf{x})\sigma_1 - A_2(\mathbf{x})\sigma_2\big)\partial_1 \\ &\quad +\frac{1}{\delta^2}\sigma_1\big(V(\mathbf{x})I_2 - A_1(\mathbf{x})\sigma_1 + A_2(\mathbf{x})\sigma_2\big)^2\partial_1 + \frac{1}{\delta^2}\sigma_1\big(V(\mathbf{x})I_2 - A_1(\mathbf{x})\sigma_1 - A_2(\mathbf{x})\sigma_2\big)^2\partial_1 \\ &\quad -\frac{2}{\delta^2}\sigma_1A_2(\mathbf{x})\sigma_2\big(\partial_1V(\mathbf{x})I_2 - \partial_1A_1(\mathbf{x})\sigma_1 - \partial_1A_2(\mathbf{x})\sigma_2\big) \\ &= -\frac{4}{\delta^2}A_2(\mathbf{x})\big(\partial_1V(\mathbf{x})\sigma_1\sigma_2 + \partial_1A_1(\mathbf{x})\sigma_2 - \partial_1A_2(\mathbf{x})\sigma_1\big) + \frac{4}{\delta^2}A_2^2(\mathbf{x})\sigma_1\partial_1 \\ &\quad -\frac{4}{\delta^2}A_1(\mathbf{x})A_2(\mathbf{x})\sigma_2\partial_1 \\ &= \frac{4}{\delta^2}\big(A_2^2(\mathbf{x})\sigma_1 - A_1(\mathbf{x})A_2(\mathbf{x})\sigma_2\big)\partial_1 + \frac{4}{\delta^2}A_2(\mathbf{x})\big(\partial_1A_2(\mathbf{x})\sigma_1 - \partial_1A_1(\mathbf{x})\sigma_2\big) \\ &\quad -\frac{4i}{\delta^2}A_2(\mathbf{x})\partial_1V(\mathbf{x})\sigma_3. \end{aligned} \tag{2.2.16}$$

$$\begin{aligned} [W,[\sigma_3,W]] &= -\frac{1}{\delta^2}\Big(2\big(V(\mathbf{x})I_2 - A_1(\mathbf{x})\sigma_1 - A_2(\mathbf{x})\sigma_2\big)\sigma_3\big(V(\mathbf{x})I_2 - A_1(\mathbf{x})\sigma_1 - A_2(\mathbf{x})\sigma_2\big) \\ &\quad -\big(V(\mathbf{x})I_2 - A_1(\mathbf{x})\sigma_1 - A_2(\mathbf{x})\sigma_2\big)^2\sigma_3 - \sigma_3\big(V(\mathbf{x})I_2 - A_1(\mathbf{x})\sigma_1 - A_2(\mathbf{x})\sigma_2\big)^2\Big) \\ &= \frac{2}{\delta^2}\sigma_3\big(V(\mathbf{x})I_2 + A_1(\mathbf{x})\sigma_1 + A_2(\mathbf{x})\sigma_2\big)\big(A_1(\mathbf{x})\sigma_1 + A_2(\mathbf{x})\sigma_2\big) \\ &\quad -\frac{2}{\delta^2}\sigma_3\big(A_1(\mathbf{x})\sigma_1 + A_2(\mathbf{x})\sigma_2\big)\big(V(\mathbf{x})I_2 - A_1(\mathbf{x})\sigma_1 - A_2(\mathbf{x})\sigma_2\big) \\ &= \frac{4}{\delta^2}\big(A_1^2(\mathbf{x}) + A_2^2(\mathbf{x})\big)\sigma_3. \end{aligned} \tag{2.2.17}$$

$$\begin{aligned} [W,[\sigma_2\partial_2,W]] &= -\frac{4}{\delta^2}\big(A_1(\mathbf{x})A_2(\mathbf{x})\sigma_1 - A_1^2(\mathbf{x})\sigma_2\big)\partial_2 - \frac{4}{\delta^2}A_1(\mathbf{x})\big(\partial_2A_2(\mathbf{x})\sigma_1 - \partial_2A_1(\mathbf{x})\sigma_2\big) \\ &\quad +\frac{4i}{\delta^2}A_1(\mathbf{x})\partial_2V(\mathbf{x})\sigma_3. \end{aligned} \tag{2.2.18}$$

Plugging (2.2.16), (2.2.18) and (2.2.17) into (2.2.14), after some computation, we can get (2.2.13). □

Corollary 2.2. *For the Dirac equation* (1.1.7) *in 2D, i.e.* $d = 2$*, with time-independent potentials, define*

$$T = -\frac{1}{\varepsilon}\alpha_1\partial_1 - \frac{1}{\varepsilon}\alpha_2\partial_2 - \frac{i\nu}{\delta\varepsilon^2}\beta, \quad W = -\frac{i}{\delta}\Big(V(\mathbf{x})I_2 - A_1(\mathbf{x})\alpha_1 - A_2(\mathbf{x})\alpha_2\Big), \tag{2.2.19}$$

we have

$$[W,[T,W]] = F_3(\mathbf{x}) + F_1(\mathbf{x})\partial_1 + F_2(\mathbf{x})\partial_2, \tag{2.2.20}$$

where

$$\begin{aligned} F_1(\mathbf{x}) &= \frac{4}{\delta^2\varepsilon}\Big(-A_2^2(\mathbf{x})\alpha_1 + A_1(\mathbf{x})A_2(\mathbf{x})\alpha_2\Big), \quad F_2(\mathbf{x}) = \frac{4}{\delta^2\varepsilon}\Big(A_1(\mathbf{x})A_2(\mathbf{x})\alpha_1 - A_1^2(\mathbf{x})\alpha_2\Big), \\ F_3(\mathbf{x}) &= \frac{4}{\delta^2\varepsilon}\Big(A_1(\mathbf{x})\partial_2 A_2(\mathbf{x}) - A_2(\mathbf{x})\partial_1 A_2(\mathbf{x})\Big)\alpha_1 + \frac{4}{\delta^2\varepsilon}\Big(A_2(\mathbf{x})\partial_1 A_1(\mathbf{x}) - A_1(\mathbf{x})\partial_2 A_1(\mathbf{x})\Big)\alpha_2 \\ &\quad + \frac{4i}{\delta^2\varepsilon}\Big(A_2(\mathbf{x})\partial_1 V(\mathbf{x}) - A_1(\mathbf{x})\partial_2 V(\mathbf{x})\Big)\gamma\alpha_3 - \frac{4i\nu}{\delta^3\varepsilon^2}\big(A_1^2(\mathbf{x}) + A_2^2(\mathbf{x})\big)\beta, \end{aligned}$$

where

$$\gamma = \begin{pmatrix} \mathbf{0} & I_2 \\ I_2 & \mathbf{0} \end{pmatrix}. \tag{2.2.21}$$

For the Dirac equation (1.1.7) in 3D, i.e. $d = 3$, we have

Lemma 2.4. *For the Dirac equation* (1.1.7) *in 3D, i.e.* $d = 3$*, with time-independent potentials, define*

$$T = -\frac{1}{\varepsilon}\sum_{j=1}^{3}\alpha_j\partial_j - \frac{i\nu}{\delta\varepsilon^2}\beta, \quad W = -\frac{i}{\delta}\Big(V(\mathbf{x})I_4 - \sum_{j=1}^{3}A_j(\mathbf{x})\alpha_j\Big), \tag{2.2.22}$$

we have

$$[W,[T,W]] = F_4(\mathbf{x}) + F_1(\mathbf{x})\partial_1 + F_2(\mathbf{x})\partial_2 + F_3(\mathbf{x})\partial_3, \tag{2.2.23}$$

where

$$
\begin{aligned}
F_1(\mathbf{x}) &= \frac{4}{\delta^2\varepsilon}\Big(-\big(A_2^2(\mathbf{x})+A_3^2(\mathbf{x})\big)\alpha_1+A_1(\mathbf{x})A_2(\mathbf{x})\alpha_2+A_1(\mathbf{x})A_3(\mathbf{x})\alpha_3\Big),\\
F_2(\mathbf{x}) &= \frac{4}{\delta^2\varepsilon}\Big(A_2(\mathbf{x})A_1(\mathbf{x})\alpha_1-\big(A_1^2(\mathbf{x})+A_3^2(\mathbf{x})\big)\alpha_2+A_2(\mathbf{x})A_3(\mathbf{x})\alpha_3\Big),\\
F_3(\mathbf{x}) &= \frac{4}{\delta^2\varepsilon}\Big(A_3(\mathbf{x})A_1(\mathbf{x})\alpha_1+A_3(\mathbf{x})A_2(\mathbf{x})\alpha_2-\big(A_1^2(\mathbf{x})+A_2^2(\mathbf{x})\big)\alpha_3\Big),\\
F_4(\mathbf{x}) &= \frac{4}{\delta^2\varepsilon}\Big(A_1(\mathbf{x})\big(\partial_2A_2(\mathbf{x})+\partial_3A_3(\mathbf{x})\big)-A_2(\mathbf{x})\partial_1A_2(\mathbf{x})-A_3(\mathbf{x})\partial_1A_3(\mathbf{x})\Big)\alpha_1\\
&\quad+\frac{4}{\delta^2\varepsilon}\Big(A_2(\mathbf{x})\big(\partial_1A_1(\mathbf{x})+\partial_3A_3(\mathbf{x})\big)-A_1(\mathbf{x})\partial_2A_1(\mathbf{x})-A_3(\mathbf{x})\partial_2A_3(\mathbf{x})\Big)\alpha_2\\
&\quad+\frac{4}{\delta^2\varepsilon}\Big(A_3(\mathbf{x})\big(\partial_1A_1(\mathbf{x})+\partial_2A_2(\mathbf{x})\big)-A_1(\mathbf{x})\partial_3A_1(\mathbf{x})-A_2(\mathbf{x})\partial_3A_2(\mathbf{x})\Big)\alpha_3\\
&\quad+\frac{4i}{\delta^2\varepsilon}\Big(A_1(\mathbf{x})\big(\partial_2A_3(\mathbf{x})-\partial_3A_2(\mathbf{x})\big)+A_2(\mathbf{x})\big(\partial_3A_1(\mathbf{x})-\partial_1A_3(\mathbf{x})\big)\\
&\quad+A_3(\mathbf{x})\big(\partial_1A_2(\mathbf{x})-\partial_2A_1(\mathbf{x})\big)\Big)\gamma+\frac{4i}{\delta^2\varepsilon}\Big(A_3(\mathbf{x})\partial_2V(\mathbf{x})-A_2(\mathbf{x})\partial_3V(\mathbf{x})\Big)\gamma\alpha_1\\
&\quad+\frac{4i}{\delta^2\varepsilon}\Big(A_1(\mathbf{x})\partial_3V(\mathbf{x})-A_3(\mathbf{x})\partial_1V(\mathbf{x})\Big)\gamma\alpha_2\\
&\quad+\frac{4i}{\delta^2\varepsilon}\Big(A_2(\mathbf{x})\partial_1V(\mathbf{x})-A_1(\mathbf{x})\partial_2V(\mathbf{x})\Big)\gamma\alpha_3-\frac{4i\nu}{\delta^3\varepsilon^2}\Big(A_1^2(\mathbf{x})+A_2^2(\mathbf{x})+A_3^2(\mathbf{x})\Big)\beta.
\end{aligned}
$$

Proof. By combining (2.2.22) and (2.2.2), we obtain

$$
[W,[T,W]]=-\frac{1}{\varepsilon}[W,[\alpha_1\partial_1,W]]-\frac{1}{\varepsilon}[W,[\alpha_2\partial_2,W]]-\frac{1}{\varepsilon}[W,[\alpha_3\partial_3,W]]-\frac{i\nu}{\delta\varepsilon^2}[W,[\beta,W]]. \tag{2.2.24}
$$

From (1.1.2) and (2.2.21), we have

$$
\begin{aligned}
&\beta^2=I_4,\quad \alpha_j^2=I_4,\quad \alpha_j\alpha_l=-\alpha_l\alpha_j,\\
&\beta\alpha_j=-\alpha_j\beta,\quad \gamma\alpha_j=\alpha_j\gamma,\qquad 1\le j\ne l\le 3,\\
&\alpha_1\alpha_2=i\gamma\alpha_3,\quad \alpha_2\alpha_3=i\gamma\alpha_1,\quad \alpha_3\alpha_1=i\gamma\alpha_2.
\end{aligned} \tag{2.2.25}
$$

Noticing (2.2.22), (2.2.1) and (2.2.25), we get

$$\begin{aligned}
[W,[\beta,W]] &= -\frac{1}{\delta^2}\Big(2\Big(V(\mathbf{x})I_4-\sum_{j=1}^{3}A_j(\mathbf{x})\alpha_j\Big)\beta\Big(V(\mathbf{x})I_4-\sum_{j=1}^{3}A_j(\mathbf{x})\alpha_j\Big)\\
&\quad -\Big(V(\mathbf{x})I_4-\sum_{j=1}^{3}A_j(\mathbf{x})\alpha_j\Big)^2\beta-\beta\Big(V(\mathbf{x})I_4-\sum_{j=1}^{3}A_j(\mathbf{x})\alpha_j\Big)^2\Big)\\
&= -\frac{2}{\delta^2}\beta\Big(V(\mathbf{x})I_4+\sum_{j=1}^{3}A_j(\mathbf{x})\alpha_j\Big)\Big(V(\mathbf{x})I_4-\sum_{j=1}^{3}A_j(\mathbf{x})\alpha_j\Big)\\
&\quad +\frac{1}{\delta^2}\beta\Big(V(\mathbf{x})I_4+\sum_{j=1}^{3}A_j(\mathbf{x})\alpha_j\Big)^2+\frac{1}{\delta^2}\beta\Big(V(\mathbf{x})I_4-\sum_{j=1}^{3}A_j(\mathbf{x})\alpha_j\Big)^2\\
&= \frac{4}{\delta^2}\big(A_1^2(\mathbf{x})+A_2^2(\mathbf{x})+A_3^2(\mathbf{x})\big)\beta. \qquad (2.2.26)
\end{aligned}$$

$$\begin{aligned}
&[W,[\alpha_1\partial_1,W]]\\
&= -\frac{1}{\delta^2}\Big(2\Big(V(\mathbf{x})I_4-\sum_{j=1}^{3}A_j(\mathbf{x})\alpha_j\Big)(\alpha_1\partial_1)\Big(V(\mathbf{x})I_4-\sum_{j=1}^{3}A_j(\mathbf{x})\alpha_j\Big)\\
&\quad -\Big(V(\mathbf{x})I_4-\sum_{j=1}^{3}A_j(\mathbf{x})\alpha_j\Big)^2(\alpha_1\partial_1)-(\alpha_1\partial_1)\Big(V(\mathbf{x})I_4-\sum_{j=1}^{3}A_j(\mathbf{x})\alpha_j\Big)^2\Big)\\
&= -\frac{4}{\delta^2}\alpha_1\big(A_2(\mathbf{x})\alpha_2+A_3(\mathbf{x})\alpha_3\big)\big(\partial_1V(\mathbf{x})I_4-\partial_1A_1(\mathbf{x})\alpha_1-\partial_1A_2(\mathbf{x})\alpha_2-\partial_1A_3(\mathbf{x})\alpha_3\big)\\
&\quad +\frac{1}{\delta^2}\alpha_1\Big(\Big(V(\mathbf{x})I_4-A_1(\mathbf{x})\alpha_1+A_2(\mathbf{x})\alpha_2+A_3(\mathbf{x})\alpha_3\Big)^2+\Big(V(\mathbf{x})I_4-\sum_{j=1}^{3}A_j(\mathbf{x})\alpha_j\Big)^2\\
&\quad -2\Big(V(\mathbf{x})I_4-A_1(\mathbf{x})\alpha_1+A_2(\mathbf{x})\alpha_2+A_3(\mathbf{x})\alpha_3\Big)\Big(V(\mathbf{x})I_4-\sum_{j=1}^{3}A_j(\mathbf{x})\alpha_j\Big)\Big)\partial_1,\\
&= \frac{4}{\delta^2}\big(A_2(\mathbf{x})\alpha_2+A_3(\mathbf{x})\alpha_3\big)\alpha_1\big(\partial_1V(\mathbf{x})I_4-\partial_1A_1(\mathbf{x})\alpha_1-\partial_1A_2(\mathbf{x})\alpha_2-\partial_1A_3(\mathbf{x})\alpha_3\big)\\
&\quad +\frac{4}{\delta^2}\Big(\big(A_2^2(\mathbf{x})+A_3^2(\mathbf{x})\big)\alpha_1-A_1(\mathbf{x})A_2(\mathbf{x})\alpha_2-A_1(\mathbf{x})A_3(\mathbf{x})\alpha_3\Big)\partial_1\\
&= \frac{4}{\delta^2}\Big(\big(A_2(\mathbf{x})\partial_1A_2(\mathbf{x})+A_3(\mathbf{x})\partial_1A_3(\mathbf{x})\big)\alpha_1-A_2(\mathbf{x})\partial_1A_1(\mathbf{x})\alpha_2-A_3(\mathbf{x})\partial_1A_1(\mathbf{x})\alpha_3\\
&\quad +\big(iA_2(\mathbf{x})\partial_1A_3(\mathbf{x})-iA_3(\mathbf{x})\partial_1A_2(\mathbf{x})\big)\gamma+iA_3(\mathbf{x})\partial_1V(\mathbf{x})\gamma\alpha_2-iA_2(\mathbf{x})\partial_1V(\mathbf{x})\gamma\alpha_3\Big)\\
&\quad +\frac{4}{\delta^2}\Big(\big(A_2^2(\mathbf{x})+A_3^2(\mathbf{x})\big)\alpha_1-A_1(\mathbf{x})A_2(\mathbf{x})\alpha_2-A_1(\mathbf{x})A_3(\mathbf{x})\alpha_3\Big)\partial_1. \qquad (2.2.27)
\end{aligned}$$

$$
\begin{aligned}
&[W,[\alpha_2\partial_2,W]]\\
&=\frac{4}{\delta^2}\Big(-A_1(\mathbf{x})\partial_2A_2(\mathbf{x})\alpha_1+\big(A_1(\mathbf{x})\partial_2A_1(\mathbf{x})+A_3(\mathbf{x})\partial_2A_3(\mathbf{x})\big)\alpha_2-A_3(\mathbf{x})\partial_2A_2(\mathbf{x})\alpha_3\\
&\quad+\big(iA_3(\mathbf{x})\partial_2A_1(\mathbf{x})-iA_1(\mathbf{x})\partial_2A_3(\mathbf{x})\big)\gamma-iA_3(\mathbf{x})\partial_2V(\mathbf{x})\gamma\alpha_1+iA_1(\mathbf{x})\partial_2V(\mathbf{x})\gamma\alpha_3\Big)\\
&\quad+\frac{4}{\delta^2}\Big(\big(A_1^2(\mathbf{x})+A_3^2(\mathbf{x})\big)\alpha_2-A_2(\mathbf{x})A_1(\mathbf{x})\alpha_1-A_2(\mathbf{x})A_3(\mathbf{x})\alpha_3\Big)\partial_2. \qquad (2.2.28)
\end{aligned}
$$

$$
\begin{aligned}
&[W,[\alpha_3\partial_3,W]]\\
&=\frac{4}{\delta^2}\Big(-A_1(\mathbf{x})\partial_3A_3(\mathbf{x})\alpha_1-A_2(\mathbf{x})\partial_3A_3(\mathbf{x})\alpha_2+\big(A_1(\mathbf{x})\partial_3A_1(\mathbf{x})+A_2(\mathbf{x})\partial_3A_2(\mathbf{x})\big)\alpha_3\\
&\quad+\big(iA_1(\mathbf{x})\partial_3A_2(\mathbf{x})-iA_2(\mathbf{x})\partial_3A_1(\mathbf{x})\big)\gamma+iA_2(\mathbf{x})\partial_3V(\mathbf{x})\gamma\alpha_1-iA_1(\mathbf{x})\partial_3V(\mathbf{x})\gamma\alpha_2\Big)\\
&\quad+\frac{4}{\delta^2}\Big(\big(A_1^2(\mathbf{x})+A_2^2(\mathbf{x})\big)\alpha_3-A_3(\mathbf{x})A_1(\mathbf{x})\alpha_1-A_3(\mathbf{x})A_2(\mathbf{x})\alpha_2\Big)\partial_3. \qquad (2.2.29)
\end{aligned}
$$

Plugging (2.2.27), (2.2.28), (2.2.29) and (2.2.26) into (2.2.24), after some computation, we obtain (2.2.23). □

From Lemmas 2.2, 2.3 and 2.4 and Corollaries 2.1 and 2.2, it is easy to observe that the double commutator will vanish when the Dirac equation (1.1.17) (or (1.1.7)) has no magnetic potentials, as is stated in the following lemma.

Lemma 2.5. *For the Dirac equation* (1.1.17) *in 1D and 2D, and for the Dirac equation* (1.1.7) *in 1D, 2D and 3D, when there is no magnetic potential, i.e., when* $A_1(\mathbf{x})=A_2(\mathbf{x})=A_3(\mathbf{x})\equiv 0$*, we have*

$$[W,[T,W]]=0. \qquad (2.2.30)$$

2.2.3 Full discretization in 1D

In this section, we present the fourth-order compact time-splitting Fourier pseudospectral method for the Dirac equation (1.1.7) (or (1.1.17)) by applying S_{4c} (2.1.14) for time integration and the Fourier pseudospectral discretization in space. For simplicity of notations, we present the numerical method for (1.1.17) in 1D first. Similar to most works in the literature for the analysis and computation of the Dirac equation (cf. [14, 15, 16, 23] and references therein), in practical computation, we truncate the whole space problem onto an interval $\Omega=(a,b)$ with periodic boundary conditions. The truncated interval is large enough such

that the truncation error is negligible. In 1D, the Dirac equation (1.1.17) ($V(t,\mathbf{x}) \equiv V(\mathbf{x})$, $A_j(t,\mathbf{x}) \equiv A_j(\mathbf{x})$, $j = 1,2$) with periodic boundary conditions collapses to

$$\begin{aligned} &i\delta\partial_t\Phi = \left(-i\frac{\delta}{\varepsilon}\sigma_1\partial_x + \frac{\nu}{\varepsilon^2}\sigma_3\right)\Phi + \Big(V(x)I_2 - A_1(x)\sigma_1\Big)\Phi, \quad x \in \Omega, \quad t > 0, \\ &\Phi(t,a) = \Phi(t,b), \quad \partial_x\Phi(t,a) = \partial_x\Phi(t,b), \quad t \geq 0; \\ &\Phi(0,x) = \Phi_0(x), \quad a \leq x \leq b; \end{aligned} \tag{2.2.31}$$

where $\Phi := \Phi(t,x)$, $\Phi_0(a) = \Phi_0(b)$ and $\Phi_0'(a) = \Phi_0'(b)$.

Choose a time step $\tau > 0$, denote $t_n = n\tau$ for $n \geq 0$ and let $\Phi^n(x)$ be an approximation of $\Phi(t_n,x)$. Re-write the Dirac equation (2.2.31) as

$$\partial_t\Phi = \left(-\frac{1}{\varepsilon}\sigma_1\partial_x - \frac{i\nu}{\delta\varepsilon^2}\sigma_3\right)\Phi - \frac{i}{\delta}\Big(V(x)I_2 - A_1(x)\sigma_1\Big)\Phi := (T+W)\Phi, \tag{2.2.32}$$

then we can apply the S_{4c} method (2.1.14) for time integration over the time interval $[t_n, t_{n+1}]$ as

$$\Phi^{n+1}(x) = S_{4c}(\tau)\Phi^n(x) := e^{\frac{1}{6}\tau W}e^{\frac{1}{2}\tau T}e^{\frac{2}{3}\tau\widehat{W}}e^{\frac{1}{2}\tau T}e^{\frac{1}{6}\tau W}\Phi^n(x), \quad a \leq x \leq b, \quad n \geq 0, \tag{2.2.33}$$

where the two operators T and W are given in (2.2.4) and the operator $\widehat{W}$ is given in (2.2.9). In order to calculate $e^{\frac{1}{2}\tau T}$, we can discretize it in space via Fourier spectral method and then integrate (in phase space or Fourier space) in time **exactly** [15, 23]. Since W is diagonalizable [15], $e^{\frac{1}{6}\tau W}$ can be evaluated very efficiently [15]. For $e^{\frac{2}{3}\tau\widehat{W}}$, by plugging (1.1.3) into (2.2.9), we can diagonalize it as

$$\widehat{W} = -\frac{i}{\delta}\Big(V(x)I_2 - A_1(x)\sigma_1\Big) - \frac{i\nu\tau^2}{12\delta^3\varepsilon^2}A_1^2(x)\sigma_3 = -iP_2(x)\,\Lambda_2(x)\,P_2(x)^* := \widehat{W}(x), \tag{2.2.34}$$

where $\Lambda_2(x) = \mathrm{diag}(\lambda_+^{(2)}(x), \lambda_-^{(2)}(x))$ with $\lambda_\pm^{(2)}(x) = \frac{V(x)}{\delta} \pm \frac{A_1(x)}{12\delta^3\varepsilon^2}\sqrt{144\delta^4\varepsilon^4 + \nu^2\tau^4A_1^2(x)}$, and

$$P_2(x) = \frac{1}{\sqrt{2\beta_1(x)}}\begin{pmatrix} \sqrt{\beta_1(x)+\beta_2(x)} & \sqrt{\beta_1(x)-\beta_2(x)} \\ -\sqrt{\beta_1(x)-\beta_2(x)} & \sqrt{\beta_1(x)+\beta_2(x)} \end{pmatrix}, \quad a \leq x \leq b, \tag{2.2.35}$$

with

$$\beta_1(x) = \sqrt{144\delta^4\varepsilon^4 + \nu^2\tau^4A_1^2(x)}, \quad \beta_2(x) = \nu\tau^2A_1(x), \quad a \leq x \leq b. \tag{2.2.36}$$

Thus we have

$$e^{\frac{2}{3}\tau\widehat{W}} = e^{-\frac{2i}{3}\tau P_2(x)\Lambda_2(x)P_2(x)^*} = P_2(x)\,e^{-\frac{2i}{3}\tau\Lambda_2(x)}\,P_2(x)^*, \quad a \le x \le b. \tag{2.2.37}$$

Choose a mesh size $h := \Delta x = \frac{b-a}{M}$ with M being an even positive integer and denote the grid points as $x_j := a + jh$, for $j = 0, 1, \ldots, M$. Denote $X_M = \{U = (U_0, U_1, ..., U_M)^T \mid U_j \in \mathbb{C}^2, j = 0, 1, \ldots, M,\ U_0 = U_M\}$. For any $U \in X_M$, its Fourier representation is given as

$$U_j = \sum_{l=-M/2}^{M/2-1} \widetilde{U}_l\, e^{i\mu_l(x_j-a)} = \sum_{l=-M/2}^{M/2-1} \widetilde{U}_l\, e^{2ijl\pi/M}, \qquad j = 0, 1, \ldots, M, \tag{2.2.38}$$

where μ_l and $\widetilde{U}_l \in \mathbb{C}^2$ are defined as

$$\mu_l = \frac{2l\pi}{b-a}, \qquad \widetilde{U}_l = \frac{1}{M}\sum_{j=0}^{M-1} U_j\, e^{-2ijl\pi/M}, \qquad l = -\frac{M}{2}, \ldots, \frac{M}{2}-1. \tag{2.2.39}$$

For $U \in X_M$ and $u(x) \in L^2(\Omega)$, their l^2-norms are defined as

$$\|U\|_{l^2}^2 := h\sum_{j=0}^{M-1} |U_j|^2, \qquad \|u\|_{l^2}^2 := h\sum_{j=0}^{M-1} |u(x_j)|^2. \tag{2.2.40}$$

Let Φ_j^n be the numerical approximation of $\Phi(t_n, x_j)$ and denote $\Phi^n = \left(\Phi_0^n, \Phi_1^n, \ldots, \Phi_M^n\right)^T \in X_M$ as the solution vector at $t = t_n$. Take $\Phi_j^0 = \Phi_0(x_j)$ for $j = 0, \ldots, M$, then a **fourth-order compact time-splitting Fourier pseudospectral** (S_{4c}) discretization for the Dirac equation (2.2.31) is given as

$$\begin{aligned}
\Phi_j^{(1)} &= e^{\frac{\tau}{6}W(x_j)}\Phi_j^n = P_1\, e^{-\frac{i\tau}{6}\Lambda_1(x_j)}P_1^*\,\Phi_j^n,\\
\Phi_j^{(2)} &= \sum_{l=-M/2}^{M/2-1} e^{\tau\Gamma_l}\left(\widetilde{\Phi^{(1)}}\right)_l e^{i\mu_l(x_j-a)} = \sum_{l=-M/2}^{M/2-1} Q_l\, e^{-i\tau D_l}\, Q_l^*\left(\widetilde{\Phi^{(1)}}\right)_l e^{2ijl\pi/M},\\
\Phi_j^{(3)} &= e^{\frac{2\tau}{3}\widehat{W}(x_j)}\Phi_j^{(2)} = P_2(x_j)\, e^{-\frac{2i\tau}{3}\Lambda_2(x_j)}\, P_2(x_j)^*\,\Phi_j^{(2)}, \qquad j = 0, 1, \ldots, M,\\
\Phi_j^{(4)} &= \sum_{l=-M/2}^{M/2-1} e^{\tau\Gamma_l}\left(\widetilde{\Phi^{(3)}}\right)_l e^{i\mu_l(x_j-a)} = \sum_{l=-M/2}^{M/2-1} Q_l\, e^{-i\tau D_l}\, Q_l^*\left(\widetilde{\Phi^{(3)}}\right)_l e^{2ijl\pi/M},\\
\Phi_j^{n+1} &= e^{\frac{\tau}{6}W(x_j)}\Phi_j^{(4)} = P_1\, e^{-\frac{i\tau}{6}\Lambda_1(x_j)}\, P_1^*\,\Phi_j^{(4)},
\end{aligned} \tag{2.2.41}$$

where

$$\begin{aligned}
W(x_j) &:= -\frac{i}{\delta}\Big(V(x_j)I_2 - A_1(x_j)\sigma_1\Big) = -iP_1\,\Lambda_1(x_j)\,P_1^*, \quad j = 0, 1, \ldots, M,\\
\Gamma_l &= -\frac{i\mu_l}{\varepsilon}\sigma_1 - \frac{i\nu}{\delta\varepsilon^2}\sigma_3 = -iQ_l\,D_l\,Q_l^*, \quad l = -\frac{M}{2}, \ldots, \frac{M}{2}-1,
\end{aligned} \tag{2.2.42}$$

with $D_l = \mathrm{diag}\left(\frac{1}{\delta\varepsilon^2}\sqrt{\nu^2+\delta^2\varepsilon^2\mu_l^2}, -\frac{1}{\delta\varepsilon^2}\sqrt{\nu^2+\delta^2\varepsilon^2\mu_l^2}\right)$, $\Lambda_1(x) = \mathrm{diag}\left(\lambda_+^{(1)}(x), \lambda_-^{(1)}(x)\right)$ with $\lambda_\pm^{(1)}(x) = \frac{1}{\delta}\big(V(x) \pm A_1(x)\big)$, $\eta_l = \sqrt{\nu^2+\delta^2\varepsilon^2\mu_l^2}$, and

$$P_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}, \quad Q_l = \frac{1}{\sqrt{2\eta_l(\eta_l+\nu)}} \begin{pmatrix} \eta_l+\nu & -\delta\varepsilon\mu_l \\ \delta\varepsilon\mu_l & \eta_l+\nu \end{pmatrix}, \quad l = -\frac{M}{2}, \ldots, \frac{M}{2}-1. \tag{2.2.43}$$

We remark here that full discretization by other time-splitting methods together with Fourier pseudospectral method for spatial discretization can be implemented similarly [15] and the details are omitted here for brevity.

The S_{4c} (2.2.41) is explicit, its memory cost is $O(M)$ and its computational cost per time step is $O(M \ln M)$. It obtains fourth-order accuracy in time and spectral accuracy in space. In addition, it conserves the total probability in the discretized level, as shown in the following lemma.

Lemma 2.6. *For any $\tau > 0$, S_{4c} (2.2.41) conserves the mass in the discretized level, i.e.*

$$\left\|\Phi^{n+1}\right\|_{l^2}^2 := h\sum_{j=0}^{M-1}\left|\Phi_j^{n+1}\right|^2 \equiv h\sum_{j=0}^{M-1}\left|\Phi_j^0\right|^2 = h\sum_{j=0}^{M-1}\left|\Phi_0(x_j)\right|^2 = \left\|\Phi_0\right\|_{l^2}^2, \quad n \geq 0. \tag{2.2.44}$$

Proof. Noticing $W(x_j)^* = -W(x_j)$ and thus $\left(e^{\frac{\tau}{6}W(x_j)}\right)^* e^{\frac{\tau}{6}W(x_j)} = I_2$, from (2.2.41) and summing for $j = 0, 1, \ldots, M-1$, we get

$$\begin{aligned} \left\|\Phi^{n+1}\right\|_{l^2}^2 &= h\sum_{j=0}^{M-1}\left|\Phi_j^{n+1}\right|^2 = h\sum_{j=0}^{M-1}\left|e^{\frac{\tau}{6}W(x_j)}\Phi_j^{(4)}\right|^2 = h\sum_{j=0}^{M-1}(\Phi_j^{(4)})^*\left(e^{\frac{\tau}{6}W(x_j)}\right)^* e^{\frac{\tau}{6}W(x_j)}\Phi_j^{(4)} \\ &= h\sum_{j=0}^{M-1}(\Phi_j^{(4)})^* I_2\,\Phi_j^{(4)} = h\sum_{j=0}^{M-1}\left|\Phi_j^{(4)}\right|^2 = \left\|\Phi^{(4)}\right\|_{l^2}^2, \qquad n \geq 0. \end{aligned} \tag{2.2.45}$$

Similarly, we have

$$\left\|\Phi^{(3)}\right\|_{l^2}^2 = \left\|\Phi^{(2)}\right\|_{l^2}^2, \qquad \left\|\Phi^{(1)}\right\|_{l^2}^2 = \left\|\Phi^{n}\right\|_{l^2}^2, \qquad n \geq 0. \tag{2.2.46}$$

Similarly, using the Parsval's identity and noticing $\Gamma_l^* = -\Gamma_l$ and thus $\left(e^{\tau\Gamma_l}\right)^* e^{\tau\Gamma_l} = I_2$, we get

$$\left\|\Phi^{(4)}\right\|_{l^2}^2 = \left\|\Phi^{(3)}\right\|_{l^2}^2, \qquad \left\|\Phi^{(2)}\right\|_{l^2}^2 = \left\|\Phi^{(1)}\right\|_{l^2}^2. \tag{2.2.47}$$

Combining (2.2.45), (2.2.46) and (2.2.47), we obtain

$$\left\|\Phi^{n+1}\right\|_{l^2}^2 = \left\|\Phi^{(4)}\right\|_{l^2}^2 = \left\|\Phi^{(3)}\right\|_{l^2}^2 = \left\|\Phi^{(2)}\right\|_{l^2}^2 = \left\|\Phi^{(1)}\right\|_{l^2}^2 = \left\|\Phi^{n}\right\|_{l^2}^2, \quad n \geq 0. \tag{2.2.48}$$

Using the mathematical induction, we get the mass conservation (2.2.44). □

2.2.4 Discussion on extension to 2D and 3D

When there is no magnetic potential, i.e., when $A_1(\mathbf{x}) = A_2(\mathbf{x}) = A_3(\mathbf{x}) \equiv 0$ in the Dirac equation (1.1.17) in 2D and (1.1.7) in 2D and 3D, from Lemma 2.5, we know that the double commutator $[W,[T,W]] = 0$. In this case, noticing (2.1.15), we have

$$\widehat{W} = W + \frac{1}{48}\tau^2[W,[T,W]] = W, \tag{2.2.49}$$

and then the S_{4c} method (2.1.14) collapses to

$$u(\tau,\mathbf{x}) \approx S_{4c}(\tau)u_0(\mathbf{x}) := e^{\frac{1}{6}\tau W}e^{\frac{1}{2}\tau T}e^{\frac{2}{3}\tau W}e^{\frac{1}{2}\tau T}e^{\frac{1}{6}\tau W}u_0(\mathbf{x}). \tag{2.2.50}$$

Applying the S_{4c} method (2.2.50) to integrate the Dirac equation (1.1.17) in 2D over the time interval $[t_n, t_{n+1}]$ with $\Phi(t_n,\mathbf{x}) = \Phi^n(\mathbf{x})$ given, we obtain

$$\Phi^{n+1}(\mathbf{x}) = S_{4c}(\tau)\Phi^n(\mathbf{x}) = e^{\frac{1}{6}\tau W}e^{\frac{1}{2}\tau T}e^{\frac{2}{3}\tau W}e^{\frac{1}{2}\tau T}e^{\frac{1}{6}\tau W}\Phi^n(\mathbf{x}), \quad \mathbf{x}\in\Omega, \quad n \geq 0, \tag{2.2.51}$$

where T and W are given in (2.2.12). Similarly, applying the S_{4c} method (2.2.50) to integrate the Dirac equation (1.1.7) in 2D and 3D over the time interval $[t_n, t_{n+1}]$ with $\Psi(t_n,\mathbf{x}) = \Psi^n(\mathbf{x})$ given, we obtain

$$\Psi^{n+1}(\mathbf{x}) = S_{4c}(\tau)\Psi^n(\mathbf{x}) = e^{\frac{1}{6}\tau W}e^{\frac{1}{2}\tau T}e^{\frac{2}{3}\tau W}e^{\frac{1}{2}\tau T}e^{\frac{1}{6}\tau W}\Psi^n(\mathbf{x}), \quad \mathbf{x}\in\Omega, \quad n \geq 0, \tag{2.2.52}$$

where T and W are given in (2.2.19) and (2.2.22) for 2D and 3D, respectively. In practical computation, the operators $e^{\frac{1}{6}\tau W}$ and $e^{\frac{2}{3}\tau W}$ in (2.2.51) and (2.2.52) can be evaluated in physical space directly and easily [15]. For the operator $e^{\frac{1}{2}\tau T}$, it can be discretized in space via Fourier spectral method and then integrate (in phase space or Fourier space) in time **exactly**. For details, we refer to [15, 23] and references therein. In fact, the implementation of the S_{4c} method in this case is much simpler than that of the S_4 and S_{4RK} methods.

On the other hand, when the magnetic potential is nonzero in the Dirac equation (1.1.17) in 2D and (1.1.7) in 2D and 3D, one has to adapt the formulation (2.2.50) for S_{4c} method. In this case, the main difficulty is how to efficiently and accurately evaluate the operator $e^{\frac{2}{3}\tau\hat{W}}$. This can be done by using the method of characteristics and the nonuniform fast Fourier transform (NUFFT), which has been developed for the magnetic Schrödinger equation. For details, we refer to [37, 84] and references therein. Of course, in this situation, it is a little more tedious in practical implementation for S_{4c} than that for S_4 and S_{4RK}.

2.3 Numerical results

In this section, we compare the accuracy and efficiency as well as long time behavior of the fourth-order compact time-splitting Fourier pseudospectral S_{4c} method (2.2.41) with other time-splitting methods including the first-order time-splitting (S_1) method, the second-order time-splitting (S_2) method, the fourth-order time-splitting (S_4) method and the fourth-order partitioned Runge-Kutta time-splitting (S_{4RK}) method for the Dirac equation in the classical regime. We also report the spatial/temporal resolution of the S_{4c} method for the Dirac equation in different parameter regimes.

2.3.1 Comparison with other time-splitting methods in the classical regime

For simplicity, we first consider an example in 1D. In the Dirac equation (1.1.17), we take $d = 1$, $\varepsilon = \delta = \nu = 1$ and

$$V(x) = \frac{1-x}{1+x^2}, \quad A_1(x) = \frac{(x+1)^2}{1+x^2}, \quad x \in \mathbb{R}. \tag{2.3.1}$$

The initial data in (1.1.18) is taken as:

$$\phi_1(0,x) = e^{-x^2/2}, \quad \phi_2(0,x) = e^{-(x-1)^2/2}, \quad x \in \mathbb{R}. \tag{2.3.2}$$

The problem is solved numerically on a bounded domain $\Omega = (-32, 32)$, i.e. $a = -32$ and $b = 32$.

Due to the fact that the exact solution is unavailable, we obtain a numerical 'exact' solution by utilizing the S_{4c} method with a fine mesh size $h_e = \frac{1}{16}$ and a small time step $\tau_e = 10^{-5}$.

	$h_0 = 1$	$h_0/2$	$h_0/2^2$	$h_0/2^3$
S_1	1.01	5.16E-2	7.07E-5	–
S_2	1.01	5.16E-2	6.96E-5	1.92E-10
S_4	1.01	5.16E-2	6.96E-5	3.52E-10
S_{4c}	1.01	5.16E-2	6.96E-5	3.06E-10
S_{4RK}	1.01	5.16E-2	6.96E-5	5.15E-10

Table 2.3.1: Spatial errors $e_\Phi(t=6)$ of different time-splitting methods under different mesh size h for the Dirac equation (1.1.17) in 1D.

Let Φ^n be the numerical solution obtained by a numerical method with mesh size h and time step τ, then the error is quantified as

$$e_\Phi(t_n) = \|\Phi^n - \Phi(t_n,\cdot)\|_{l^2} = \sqrt{h\sum_{j=0}^{M-1}|\Phi(t_n,x_j)-\Phi_j^n|^2}. \tag{2.3.3}$$

In order to compare the spatial errors, we take time step $\tau = \tau_e = 10^{-5}$ such that the temporal discretization error could be negligible. Table 2.3.1 lists numerical errors $e_\Phi(t=6)$ for different time-splitting methods under different mesh sizes h. We remark here that, for the S_1 method, in order to observe the spatial error when the mesh size $h = h_0/2^3$, one has to choose time step $\tau \le 10^{-10}$ which is too small and thus the error is not shown in the table for this case. From Table 2.3.1, we could see that all the numerical methods are spectral order accurate in space (cf. each row in Table 2.3.1).

On the other hand, in order to compare the temporal errors, we take the mesh size $h = h_e = \frac{1}{16}$ such that the spatial discretization error is negligible. Table 2.3.2 lists numerical errors $e_\Phi(t=6)$ for different time-splitting methods under different time step τ. In the table, we use second (s) as the unit for CPU time. For comparison, Figure 2.3.1 plots $e_\Phi(t=6)$ and $e_\Phi(t=6)/\tau^\alpha$ with α taken as the order of accuracy of a certain numerical method (in order to show the constants C_1 in (2.1.6), C_2 in (2.1.8), C_4 in (2.1.11), $\widetilde{C}_4$ in (2.1.13) and $\widehat{C}_4$ in (2.1.16)) for different time-splitting methods under different time step τ.

From Table 2.3.2 and Figure 2.3.1, we can draw the following conclusions: (i) S_1 is first-order in time, S_2 is second-order in time, and S_4, S_{4c} and S_{4RK} are all fourth-order in time (cf. Table 2.3.2 and Figure 2.3.1 left). (ii) For any fixed mesh h and time τ, the computational

		$\tau_0 = 1/2$	$\tau_0/2$	$\tau_0/2^2$	$\tau_0/2^3$	$\tau_0/2^4$	$\tau_0/2^5$	$\tau_0/2^6$
S_1	$e_\Phi(t=6)$	1.17	4.71E-1	2.09E-1	9.90E-2	4.82E-2	2.38E-2	1.18E-2
	rate	–	1.31	1.17	1.08	1.04	1.02	1.01
	CPU Time	0.02	0.05	0.11	0.16	0.37	0.62	**1.31**
S_2	$e_\Phi(t=6)$	7.49E-1	1.87E-1	4.66E-2	1.16E-2	2.91E-3	7.27E-4	1.82E-4
	rate	–	2.00	2.00	2.00	2.00	2.00	2.00
	CPU Time	0.04	0.06	0.11	0.21	0.37	0.75	**1.42**
S_4	$e_\Phi(t=6)$	3.30E-1	3.73E-2	3.05E-3	2.07E-4	1.32E-5	8.29E-7	5.20E-8
	rate	–	3.15	3.61	3.89	3.97	3.99	4.00
	CPU Time	0.10	0.16	0.38	0.58	1.09	2.23	**4.41**
S_{4c}	$e_\Phi(t=6)$	1.66E-2	9.54E-4	5.90E-5	3.68E-6	2.30E-7	1.43E-8	8.12E-10
	rate	–	4.12	4.01	4.00	4.00	4.01	4.13
	CPU Time	0.06	0.09	0.18	0.35	0.68	1.36	**2.68**
S_{4RK}	$e_\Phi(t=6)$	2.87E-3	1.78E-4	1.11E-5	6.97E-7	4.34E-8	2.58E-9	1.66E-10
	rate	–	4.01	3.99	4.00	4.00	4.07	3.96
	CPU Time	0.15	0.28	0.57	1.24	2.66	3.94	**7.79**

Table 2.3.2: Temporal errors $e_\Phi(t=6)$ of different time-splitting methods under different time step τ for the Dirac equation (1.1.17) in 1D. Here we also list convergence rates and computational time (CPU time in seconds) for comparison.

times for S_1 and S_2 are quite similar, the computational times of S_{4c}, S_4 and S_{4RK} are about two times, three times and six times of the S_2 method, respectively (cf. Table 2.3.2). (iii) Among the three fourth-order time-splitting methods, S_{4c} and S_{4RK} are quite similar in terms of numerical errors for any fixed τ and they are much smaller than that of the S_4 method, especially when the τ is not so small (cf. Table 2.3.2 and Figure 2.3.1 left). (iv) For the constants in front of the convergence rates of different methods, $C_4 \gg C_1 \sim C_2 \gg \widehat{C_4} \sim \widetilde{C}_4$ (cf. Figure 2.3.1 right). (v) S_4 suffers from convergence rate reduction when the time step is not small and there is a very large constant in front of the convergence rate. As a result, this method is, in general, to be avoided in practical computation, as has been observed when it is

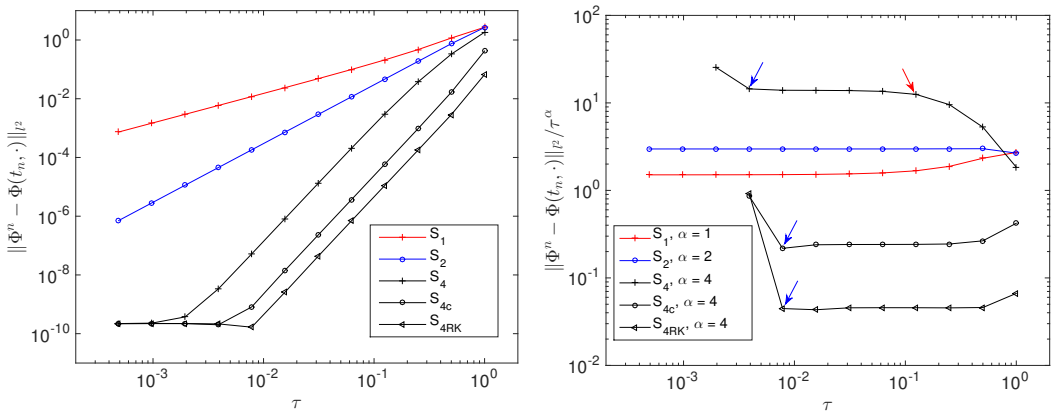

Figure 2.3.1: Temporal errors $e_\Phi(t=6)$ (left) and $e_\Phi(t=6)/\tau^\alpha$ with α taken as the order of accuracy of a certain numerical method (right) of different time-splitting methods under different time step τ for the Dirac equation (1.1.17) in 1D.

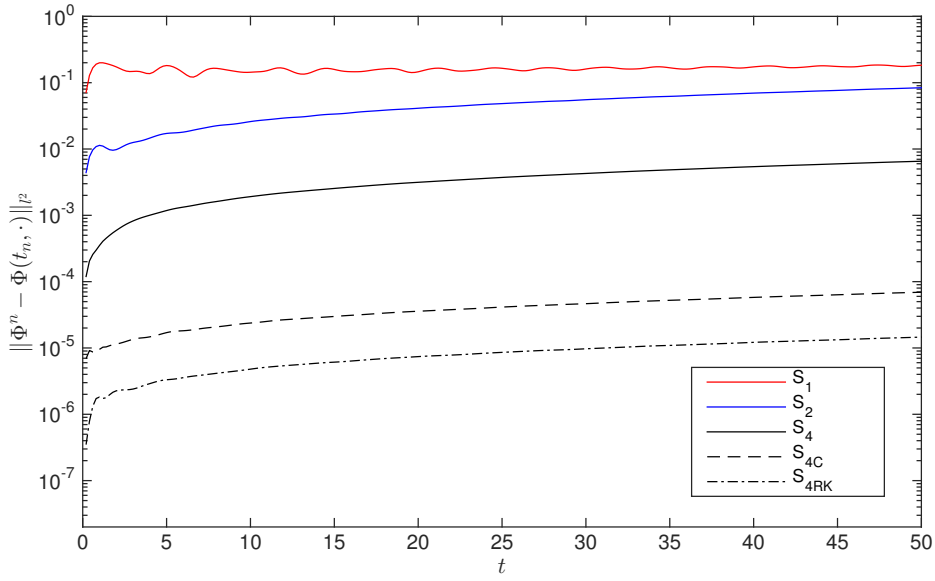

Figure 2.3.2: Time evolution of the errors $e_\Phi(t)$ under $h = \frac{1}{16}$ and $\tau = 0.1$ over long time of different time-splitting methods for the Dirac equation (1.1.17) in 1D.

applied for the nonlinear Schrödinger equation as well [120].

To compare the long time behavior of different time-splitting methods, Figure 2.3.2 depicts $e_\Phi(t)$ under mesh size $h = \frac{1}{16}$ and time step $\tau = 0.1$ for $0 \le t \le T := 50$.

From Figure 2.3.2, we can observe: (i) The errors increase very fast when t is small, e.g. $0 \le t \le O(1)$, and they almost don't change when $t \gg 1$, thus they are suitable for long time

simulation, especially the fourth-order methods. (ii) When t is not large, the error of S_4 is about 10 times bigger than that of S_{4c}; however, when $t \gg 1$, it becomes about 100 times larger. (iii) The error of S_{4RK} is always the smallest among all the time-splitting methods.

Based on the efficiency and accuracy as well as long time behavior, in conclusion, among the three fourth-order time-splitting methods, S_{4c} is more accurate than S_4 and it is more efficient than S_{4RK}. Thus S_{4c} is highly recommended for studying the dynamics of the Dirac equation, especially in 1D.

Next, we consider an example in 2D. For simplicity, here we only compare the three fourth-order integrators, i.e., S_{4c}, S_4 and S_{4RK}. In order to do so, in the Dirac equation (1.1.17), we take $d = 2$, $\varepsilon = \delta = \nu = 1$ and take the potential in honey-comb form

$$\begin{aligned} V(\mathbf{x}) &= \cos\left(\frac{4\pi}{\sqrt{3}}\mathbf{e_1}\cdot\mathbf{x}\right) + \cos\left(\frac{4\pi}{\sqrt{3}}\mathbf{e}_2\cdot\mathbf{x}\right) + \cos\left(\frac{4\pi}{\sqrt{3}}\mathbf{e}_3\cdot\mathbf{x}\right), \\ A_1(\mathbf{x}) &= A_2(\mathbf{x}) = 0, \qquad \mathbf{x}\in\mathbb{R}^2, \end{aligned} \tag{2.3.4}$$

with

$$\mathbf{e}_1 = (-1,0)^T, \quad \mathbf{e}_2 = (1/2,\sqrt{3}/2)^T, \quad \mathbf{e}_3 = (1/2,-\sqrt{3}/2)^T. \tag{2.3.5}$$

The initial data in (1.1.18) is taken as:

$$\phi_1(0,\mathbf{x}) = e^{-\frac{x_1^2+x_2^2}{2}}, \quad \phi_2(0,\mathbf{x}) = e^{-\frac{(x_1-1)^2+x_2^2}{2}}, \qquad \mathbf{x} = (x_1,x_2)^T \in \mathbb{R}^2. \tag{2.3.6}$$

The problem is solved numerically on a bounded domain $\Omega = (-10,10)\times(-10,10)$.

Similar to the 1D example, we obtain a numerical 'exact' solution by using S_{4c} with a fine mesh size $h_e = \frac{1}{32}$ and a small time step $\tau_e = 10^{-4}$. The error for the numerical solution Φ^n with mesh size h and time step τ is quantified as

$$e_\Phi(t_n) = \|\Phi^n - \Phi(t_n,\cdot)\|_{l^2} = h\sqrt{\sum_{j=0}^{M-1}\sum_{l=0}^{M-1}|\Phi(t_n,x_{1j},x_{2l}) - \Phi_{jl}^n|^2}. \tag{2.3.7}$$

Similar to the 1D case, in order to compare the spatial errors, we take time step $\tau = \tau_e = 10^{-4}$ such that the temporal discretization error could be negligible. Table 2.3.3 lists numerical errors $e_\Phi(t=2)$ for different time-splitting methods under different mesh size h. On the other hand, in order to compare the temporal errors, we take mesh size $h = h_e = \frac{1}{32}$

	$h_0 = 1/2$	$h_0/2$	$h_0/2^2$	$h_0/2^3$
S_4	1.10	1.01E-1	3.83E-4	7.33E-10
S_{4c}	1.10	1.01E-1	3.83E-4	7.33E-10
S_{4RK}	1.10	1.01E-1	3.83E-4	7.34E-10

Table 2.3.3: Spatial errors $e_\Phi(t = 2)$ of different time-splitting methods under different mesh size h for the Dirac equation (1.1.17) in 2D.

		$\tau_0 = 1/2$	$\tau_0/2$	$\tau_0/2^2$	$\tau_0/2^3$	$\tau_0/2^4$	$\tau_0/2^5$	$\tau_0/2^6$
S_4	Error	4.33E-1	2.57E-2	3.53E-3	2.83E-4	1.88E-5	1.20E-6	7.51E-8
	Order	–	4.07	2.87	3.64	3.91	3.98	3.99
	CPU Time	0.20	0.26	0.45	1.04	1.63	3.37	**6.54**
S_{4c}	Error	6.75E-2	3.18E-3	7.91E-5	4.70E-6	2.91E-7	1.81E-8	1.13E-9
	Order	–	4.41	5.33	4.07	4.01	4.00	4.00
	CPU Time	0.12	0.28	0.31	0.55	1.11	2.09	**4.14**
S_{4RK}	Error	8.32E-3	3.56E-4	7.42E-6	4.43E-7	2.75E-8	1.71E-9	1.07E-10
	Order	–	4.55	5.59	4.07	4.01	4.00	4.00
	CPU Time	0.26	0.43	0.87	1.52	2.92	6.20	**11.74**

Table 2.3.4: Temporal errors $e_\Phi(t = 2)$ of different fourth order time-splitting methods under different time step τ for the Dirac equation (1.1.17) in 2D. Here we also list convergence rates and computational time (CPU time in seconds) for comparison.

such that the spatial discretization error could be negligible. Table 2.3.4 lists numerical errors $e_\Phi(t = 2)$ for different time-splitting methods under different time step τ.

From Table 2.3.3&Table 2.3.4, we can draw the following conclusions: (i) All the three methods are spectrally accurate in space and fourth-order in time. (ii) For any fixed mesh size h and time step τ, the computational times of S_4 and S_{4RK} are approximately 1.5 times and 3 times of S_{4c}, respectively. (iii) S_{4c} and S_{4RK} are quite similar in terms of numerical errors for any fixed τ and the errors are much smaller than that of S_4, especially when τ is not so small. (iv) Again, order reduction in time is observed in S_4 when τ is not small, however, there is almost no order reduction in time for S_{4c} and S_{4RK}.

Based on the efficiency and accuracy for the Dirac equation in high dimensions, in conclusion, among the three fourth-order time-splitting methods, S_{4c} is more accurate than S_4 and it is more efficient than S_{4RK}. Thus S_{4c} is highly recommended for studying the dynamics of the Dirac equation in high dimensions, especially when there is no magnetic potential.

2.3.2 Application and performance in different regimes

In this subsection, we study numerically temproal/spatial resolution of the fourth-order compact time-splitting Fourier pseudospectral S_{4c} method (2.2.41) for the Dirac equation in different parameter regimes. We take $d = 1$ and the electromagnetic potentials as (2.3.1) in Dirac equation (1.1.17). To quantify the numerical error, we adapt the relative errors of the wave function Φ, the total probability density ρ and the current $\mathbf{J}$ as

$$e^r_\Phi(t_n) = \frac{\|\Phi^n - \Phi(t_n,\cdot)\|_{l^2}}{\|\Phi(t_n,\cdot)\|_{l^2}}, \quad e^r_\rho(t_n) = \frac{\|\rho^n - \rho(t_n,\cdot)\|_{l^2}}{\|\rho(t_n,\cdot)\|_{l^2}}, \quad e^r_{\mathbf{J}}(t_n) = \frac{\|\mathbf{J}^n - \mathbf{J}(t_n,\cdot)\|_{l^2}}{\|\mathbf{J}(t_n,\cdot)\|_{l^2}}, \tag{2.3.8}$$

where ρ^n and $\mathbf{J}^n$ are obtained from the wave function Φ^n via

$$\rho(t,\mathbf{x}) = \sum_{j=1}^{2} \rho_j(t,\mathbf{x}) = \Phi(t,\mathbf{x})^* \Phi(t,\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^d, \tag{2.3.9}$$

and

$$J_l(t,\mathbf{x}) = \frac{1}{\varepsilon} \Phi(t,\mathbf{x})^* \sigma_l \Phi(t,\mathbf{x}), \quad l = 1,\ldots,d, \tag{2.3.10}$$

with $d = 1$, respectively. Again, the numerical 'exact' solution is obtained by using the S_{4c} method with a very fine mesh $h = h_e$ and a very small time step $\tau = \tau_e$.

- **In the nonrelativistic regime**

  Here we take $\delta = \nu = 1$, $\varepsilon \in (0,1]$ and the initial data in (1.1.18) is taken as (2.3.2). In this parameter regime, the solution propagates waves with wavelength at $O(1)$ and $O(\varepsilon^2)$ in space and time, respectively. The problem is solved numerically on a bounded domain $\Omega = (-32,32)$, i.e. $a = -32$ and $b = 32$. Similar to the second-order time-splitting Fourier pseudospectral method [15], the S_{4c} method converges uniformly with respect to $\varepsilon \in (0,1]$ at spectral order in space. Detailed numerical results are omitted here for brevity. Here we only present temporal errors by taking $h = h_e = \frac{1}{16}$ so that the spatial discretization error could be negligible. Table 2.3.5 shows the temporal errors

	$\tau_0 = 1$	$\tau_0/2^2$	$\tau_0/2^4$	$\tau_0/2^6$	$\tau_0/2^8$	$\tau_0/2^{10}$
$\varepsilon_0 = 1$	2.24E-1	**5.07E-4**	1.95E-6	7.63E-9	<1E-10	<1E-10
order	–	**4.39**	4.01	4.00	–	–
$\varepsilon_0/2$	1.18	1.05E-2	**3.61E-5**	1.40E-7	5.67E-10	<1E-10
order	–	3.41	**4.09**	4.00	3.97	–
$\varepsilon_0/2^2$	1.46	2.07E-1	1.69E-3	**6.09E-6**	2.37E-8	<1E-10
order	–	1.41	3.47	**4.06**	4.00	–
$\varepsilon_0/2^3$	1.41	1.50	5.88E-2	3.84E-4	**1.39E-6**	5.40E-9
order	–	-0.04	2.33	3.63	**4.06**	4.00
$\varepsilon_0/2^4$	1.43	1.47	6.80E-1	1.46E-2	9.33E-5	**3.38E-7**
order	–	-0.02	0.56	2.77	3.65	**4.05**

Table 2.3.5: Temporal errors $e^r_\Phi(t=6)$ of S_{4c} under different τ and ε for the Dirac equation (1.1.17) in 1D in the nonrelativistic regime.

	$\tau_0 = 1$	$\tau_0/2^2$	$\tau_0/2^4$	$\tau_0/2^6$	$\tau_0/2^8$	$\tau_0/2^{10}$
$\varepsilon_0 = 1$	1.71E-1	**3.73E-4**	1.44E-6	5.62E-9	<1E-10	<1E-10
order	–	**4.42**	4.01	4.00	–	–
$\varepsilon_0/2$	1.31	7.17E-3	**2.45E-5**	9.50E-8	3.94E-10	<1E-10
order	–	3.76	**4.10**	4.01	3.96	–
$\varepsilon_0/2^2$	8.19E-1	2.20E-1	8.16E-4	**2.92E-6**	1.13E-8	<1E-10
order	–	0.95	4.04	**4.06**	4.00	–
$\varepsilon_0/2^3$	8.75E-1	4.77E-1	5.76E-2	1.65E-4	**5.89E-7**	2.29E-9
order	–	0.44	1.52	4.22	**4.07**	4.00
$\varepsilon_0/2^4$	1.00	1.12	2.04E-1	1.49E-2	4.03E-5	**1.43E-7**
order	–	-0.08	1.23	1.88	4.27	**4.07**

Table 2.3.6: Temporal errors $e^r_\rho(t=6)$ of S_{4c} under different τ and ε for the Dirac equation (1.1.17) in 1D in the nonrelativistic regime.

$e^r_\Phi(t=6)$ for the wave function under different τ and $\varepsilon \in (0,1]$. Similarly, Table 2.3.6 and Table 2.3.7 depict the temporal errors $e^r_\rho(t=6)$ and $e^r_{\mathbf{J}}(t=6)$ for the probability and current, respectively.

From Table 2.3.5-Table 2.3.7, when $\tau \lesssim \varepsilon^2$, fourth-order convergence is observed for the S_{4c} method in the relative error for the wave function, probability and current.

	$\tau_0=1$	$\tau_0/2^2$	$\tau_0/2^4$	$\tau_0/2^6$	$\tau_0/2^8$	$\tau_0/2^{10}$
$\varepsilon_0=1$	2.92E-1	**6.76E-4**	2.61E-6	1.02E-8	<1E-10	<1E-10
order	–	**4.38**	4.01	4.00	–	–
$\varepsilon_0/2$	1.30	1.98E-2	**6.88E-5**	2.67E-7	1.06E-9	<1E-10
order	–	3.02	**4.09**	4.00	3.99	–
$\varepsilon_0/2^2$	1.29	2.98E-1	3.40E-3	**1.23E-5**	4.76E-8	<1E-10
order	–	1.06	3.23	**4.06**	4.00	–
$\varepsilon_0/2^3$	1.21	1.29	8.82E-2	7.85E-4	**2.85E-6**	1.11E-8
order	–	-0.05	1.94	3.41	**4.05**	4.00
$\varepsilon_0/2^4$	1.52	1.44	1.30	2.41E-2	1.92E-4	**6.98E-7**
order	–	0.04	0.07	2.88	3.48	**4.05**

Table 2.3.7: Temporal errors $e^r_{\mathbf{J}}(t=6)$ of S_{4c} under different τ and ε for the Dirac equation (1.1.17) in 1D in the nonrelativistic regime.

This suggests that the ε-scalability for the S_{4c} method in the nonrelativistic regime is: $h=O(1)$ and $\tau=O(\varepsilon^2)$. In addition, noticing $\Phi=O(1)$, $\rho=O(1)$ and $\mathbf{J}=O(\varepsilon^{-1})$ when $0\le\varepsilon\ll 1$, we can formally observe the following error bounds for $0<\varepsilon\le 1$, $\tau\lesssim\varepsilon^2$ and $0\le n\le\frac{T}{\tau}$

$$\begin{aligned}&\|\Phi^n-\Phi(t_n,\cdot)\|_{l^2}\lesssim h^{m_0}+\frac{\tau^4}{\varepsilon^6},\quad \|\rho^n-\rho(t_n,\cdot)\|_{l^2}\lesssim h^{m_0}+\frac{\tau^4}{\varepsilon^6},\\ &\|\mathbf{J}^n-\mathbf{J}(t_n,\cdot)\|_{l^2}\lesssim\frac{1}{\varepsilon}\left(h^{m_0}+\frac{\tau^4}{\varepsilon^6}\right).\end{aligned}\tag{2.3.11}$$

where $m_0\ge 2$ depends on the regularity of the solution. Rigorous mathematical justification is still on-going.

- **In the semiclassical regime** Here we take $\varepsilon=\nu=1$, $\delta\in(0,1]$. The initial data in (1.1.18) is taken as

$$\begin{aligned}\phi_1(0,x)&=\frac{1}{2}e^{-4x^2}e^{iS_0(x)/\delta}\left(1+\sqrt{1+S_0'(x)^2}\right),\\ \phi_2(0,x)&=\frac{1}{2}e^{-4x^2}e^{iS_0(x)/\delta}S_0'(x),\quad x\in\mathbb{R},\end{aligned}\tag{2.3.12}$$

with

$$S_0(x)=\frac{1}{40}\big(1+\cos(2\pi x)\big),\qquad x\in\mathbb{R}.\tag{2.3.13}$$

	$h_0 = 1$	$h_0/2$	$h_0/2^2$	$h_0/2^3$	$h_0/2^4$	$h_0/2^5$	$h_0/2^6$
$\delta_0 = 1$	**8.25E-1**	2.00E-1	9.52E-3	6.66E-6	3.78E-10	<1E-10	<1E-10
$\delta_0/2$	1.20	**7.40E-1**	5.31E-2	8.87E-5	3.43E-10	<1E-10	<1E-10
$\delta_0/2^2$	1.41	9.89E-1	**5.12E-1**	3.81E-3	9.24E-10	<1E-10	<1E-10
$\delta_0/2^3$	1.76	1.21	7.30E-1	**2.76E-1**	1.91E-5	4.17E-10	<1E-10
$\delta_0/2^4$	1.37	1.36	1.36	5.31E-1	**1.54E-1**	5.31E-10	<1E-10
$\delta_0/2^5$	2.44	1.92	1.36	1.36	4.36E-1	**5.49E-2**	2.90E-10

Table 2.3.8: Spatial errors $e^r_\Phi(t=2)$ of S_{4c} under different h and δ for the Dirac equation (1.1.17) in 1D in the semiclassical regime.

	$h_0 = 1$	$h_0/2$	$h_0/2^2$	$h_0/2^3$	$h_0/2^4$	$h_0/2^5$	$h_0/2^6$
$\delta_0 = 1$	**5.83E-1**	1.39E-1	8.27E-3	4.36E-6	4.92E-10	<1E-10	<1E-10
$\delta_0/2$	1.29	**5.22E-1**	3.71E-2	5.56E-5	2.79E-10	<1E-10	<1E-10
$\delta_0/2^2$	9.22E-1	7.44E-1	**2.41E-1**	1.54E-3	6.75E-10	<1E-10	<1E-10
$\delta_0/2^3$	1.63	9.39E-1	6.11E-1	**6.33E-2**	4.78E-6	8.19E-10	<1E-10
$\delta_0/2^4$	2.04	1.40	1.00	3.57E-1	**1.97E-2**	6.76E-10	<1E-10
$\delta_0/2^5$	5.81	3.65	1.07	1.01	1.86E-1	**3.35E-3**	5.67E-10

Table 2.3.9: Spatial errors $e^r_\rho(t=2)$ of S_{4c} under different h and δ for the Dirac equation (1.1.17) in 1D in the semiclassical regime.

In this parameter regime, the solution propagates waves with wavelength at $O(\delta)$ in both space and time. The problem is solved numerically on a bounded domain $\Omega = (-16, 16)$, i.e. $a = -16$ and $b = 16$.

Table 2.3.8 shows the spatial errors $e^r_\Phi(t=2)$ for the wave function under different h and $\delta \in (0,1]$ with $\tau = \tau_e = 10^{-4}$ such that the temporal discretization error could be negligible. Table 2.3.9 and Table 2.3.10 depict the spatial errors $e^r_\rho(t=2)$ and $e^r_{\mathbf{J}}(t=2)$ for the probability and current, respectively. Similarly, Table 2.3.11 shows the temporal errors $e^r_\Phi(t=2)$ for the wave function under different τ and $\delta \in (0,1]$ with $h = h_e = \frac{1}{128}$ so that the spatial discretization error could be negligible. Table 2.3.12 and Table 2.3.13 depict the temporal errors $e^r_\rho(t=2)$ and $e^r_{\mathbf{J}}(t=2)$ for the probability and current, respectively.

From Table 2.3.8-Table 2.3.10, when $h \lesssim \delta$, spectral convergence (in space) is observed

	$h_0=1$	$h_0/2$	$h_0/2^2$	$h_0/2^3$	$h_0/2^4$	$h_0/2^5$	$h_0/2^6$
$\delta_0=1$	**8.07E-1**	1.67E-1	1.05E-2	5.69E-6	5.10E-10	<1E-10	<1E-10
$\delta_0/2$	1.45	**6.89E-1**	4.28E-2	6.46E-5	3.06E-10	<1E-10	<1E-10
$\delta_0/2^2$	1.94	1.05	**3.52E-1**	2.13E-3	7.96E-10	<1E-10	<1E-10
$\delta_0/2^3$	2.52	1.03	7.07E-1	**1.24E-1**	7.75E-6	8.16E-10	<1E-10
$\delta_0/2^4$	2.85	1.77	1.10	5.84E-1	**4.72E-2**	6.75E-10	<1E-10
$\delta_0/2^5$	3.88	4.06	1.11	1.07	3.81E-1	**1.22E-2**	5.63E-10

Table 2.3.10: Spatial errors $e^r_{\mathbf{J}}(t=2)$ of S_{4c} under different h and δ for the Dirac equation (1.1.17) in 1D in the semiclassical regime.

	$\tau_0=1$	$\tau_0/2$	$\tau_0/2^2$	$\tau_0/2^3$	$\tau_0/2^4$	$\tau_0/2^5$	$\tau_0/2^6$
$\delta_0=1$	1.60E-1	**1.58E-2**	5.09E-4	2.08E-5	1.27E-6	7.89E-8	4.94E-9
order	–	**3.34**	4.96	4.61	4.04	4.01	4.00
$\delta_0/2$	8.66E-1	1.48E-1	**7.17E-3**	3.90E-4	2.41E-5	1.50E-6	9.39E-8
order	–	2.55	**4.36**	4.20	4.02	4.00	4.00
$\delta_0/2^2$	1.26	9.52E-1	1.38E-1	**7.38E-3**	4.50E-4	2.80E-5	1.75E-6
order	–	0.40	2.78	**4.23**	4.03	4.01	4.00
$\delta_0/2^3$	1.45	1.20	9.94E-1	1.62E-1	**9.11E-3**	5.57E-4	3.46E-5
order	–	0.27	0.27	2.62	**4.15**	4.03	4.01
$\delta_0/2^4$	1.40	1.44	1.12	9.46E-1	2.62E-1	**1.50E-2**	9.15E-4
order	–	-0.04	0.36	0.25	1.85	**4.13**	4.03
$\delta_0/2^5$	1.44	1.44	1.42	1.22	1.07	4.43E-1	**2.83E-2**
order	–	-0.01	0.03	0.22	0.19	1.27	**3.97**

Table 2.3.11: Temporal errors $e^r_{\Phi}(t=2)$ of S_{4c} under different τ and δ for the Dirac equation (1.1.17) in 1D in the semiclassical regime.

for the S_{4c} method in the relative error for the wave function, probability and current. Similarly, from Table 2.3.11-Table 2.3.13, when $\tau \lesssim \delta$, fourth-order convergence (in time) is observed for the S_{4c} method in the relative error for the wave function, probability and current. These suggest that the δ-scalability for the S_{4c} method in the semiclassical regime is: $h = O(\delta)$ and $\tau = O(\delta)$. In addition, noticing $\Phi = O(1)$, $\rho = O(1)$ and $\mathbf{J} = O(1)$ when $0 \le \delta \ll 1$, we can formally observe the following error

	$\tau_0 = 1$	$\tau_0/2$	$\tau_0/2^2$	$\tau_0/2^3$	$\tau_0/2^4$	$\tau_0/2^5$	$\tau_0/2^6$
$\delta_0 = 1$	1.15E-1	**1.23E-2**	4.11E-4	1.70E-5	1.03E-6	6.40E-8	4.11E-9
order	–	**3.23**	4.90	4.59	4.05	4.01	3.96
$\delta_0/2$	5.05E-1	9.20E-2	**4.93E-3**	2.36E-4	1.44E-5	8.98E-7	5.62E-8
order	–	2.45	**4.22**	4.39	4.03	4.01	4.00
$\delta_0/2^2$	7.69E-1	4.22E-1	4.32E-2	**2.85E-3**	1.73E-4	1.08E-5	6.72E-7
order	–	0.86	3.29	**3.92**	4.04	4.01	4.00
$\delta_0/2^3$	1.28	9.03E-1	5.67E-1	3.77E-2	**2.03E-3**	1.23E-4	7.66E-6
order	–	0.51	0.67	3.91	**4.21**	4.04	4.01
$\delta_0/2^4$	8.80E-1	1.25	9.86E-1	7.53E-1	2.58E-2	**1.35E-3**	8.15E-5
order	–	-0.50	0.34	0.39	4.87	**4.26**	4.05
$\delta_0/2^5$	9.60E-1	9.90E-1	1.09	1.08	8.82E-1	2.59E-2	**1.16E-3**
order	–	-0.04	-0.14	0.02	0.29	5.09	**4.48**

Table 2.3.12: Temporal errors $e_\rho^r(t=2)$ of S_{4c} under different τ and δ for the Dirac equation (1.1.17) in 1D in the semiclassical regime.

	$\tau_0 = 1$	$\tau_0/2$	$\tau_0/2^2$	$\tau_0/2^3$	$\tau_0/2^4$	$\tau_0/2^5$	$\tau_0/2^6$
$\delta_0 = 1$	1.98E-1	**2.21E-2**	6.42E-4	2.34E-5	1.42E-6	8.84E-8	5.55E-9
order	–	**3.16**	5.11	4.78	4.04	4.01	3.99
$\delta_0/2$	6.61E-1	1.93E-1	**8.72E-3**	4.34E-4	2.67E-5	1.66E-6	1.04E-7
order	–	1.78	**4.47**	4.33	4.02	4.01	4.00
$\delta_0/2^2$	1.25	6.66E-1	1.46E-1	**8.44E-3**	5.16E-4	3.21E-5	2.00E-6
order	–	0.91	2.19	**4.12**	4.03	4.01	4.00
$\delta_0/2^3$	1.57	1.19	7.29E-1	1.23E-1	**7.10E-3**	4.35E-4	2.71E-5
order	–	0.39	0.71	2.57	**4.11**	4.03	4.01
$\delta_0/2^4$	1.04	1.47	1.15	8.24E-1	9.50E-2	**5.86E-3**	3.60E-4
order	–	-0.50	0.35	0.48	3.12	**4.02**	4.02
$\delta_0/2^5$	1.02	1.14	1.19	1.19	9.39E-1	7.34E-2	**5.22E-3**
order	–	-0.16	-0.06	0.01	0.34	3.68	**3.81**

Table 2.3.13: Temporal errors $e_{\mathbf{J}}^r(t=2)$ of S_{4c} under different τ and δ for the Dirac equation (1.1.17) in 1D in the semiclassical regime.

bounds for $0 < \delta \leq 1$, $\tau \lesssim \delta$, $h \lesssim \delta$ and $0 \leq n \leq \frac{T}{\tau}$

$$\begin{aligned} &\|\Phi^n - \Phi(t_n,\cdot)\|_{l^2} \lesssim \frac{h^{m_0}}{\delta^{m_0}} + \frac{\tau^4}{\delta^4}, \quad \|\rho^n - \rho(t_n,\cdot)\|_{l^2} \lesssim \frac{h^{m_0}}{\delta^{m_0}} + \frac{\tau^4}{\delta^4}, \\ &\|\mathbf{J}^n - \mathbf{J}(t_n,\cdot)\|_{l^2} \lesssim \frac{h^{m_0}}{\delta^{m_0}} + \frac{\tau^4}{\delta^4}. \end{aligned} \tag{2.3.14}$$

	$\tau_0 = 1$	$\tau_0/2$	$\tau_0/2^2$	$\tau_0/2^3$	$\tau_0/2^4$	$\tau_0/2^5$	$\tau_0/2^6$
$\varepsilon_0 = 1$	**1.12E-1**	4.20E-3	2.18E-4	1.33E-5	8.30E-7	5.18E-8	3.24E-9
order	–	4.74	4.27	4.03	4.01	4.00	4.00
$\varepsilon_0/2$	4.72E-1	**3.66E-2**	1.17E-3	6.64E-5	4.09E-6	2.55E-7	1.59E-8
order	–	**3.69**	4.97	4.14	4.02	4.01	4.00
$\varepsilon_0/2^2$	1.14	2.72E-1	**1.27E-2**	3.64E-4	2.10E-5	1.30E-6	8.08E-8
order	–	2.07	**4.42**	5.12	4.11	4.02	4.00
$\varepsilon_0/2^3$	1.29	5.84E-1	1.60E-1	**5.19E-3**	1.41E-4	8.22E-6	5.07E-7
order	–	1.14	1.87	**4.94**	5.20	4.10	4.02
$\varepsilon_0/2^4$	1.40	7.31E-1	3.40E-1	9.81E-2	**2.46E-3**	6.16E-5	3.58E-6
order	–	0.94	1.10	1.79	**5.32**	5.32	4.10
$\varepsilon_0/2^5$	1.39	1.06	3.90E-1	2.09E-1	6.32E-2	**1.27E-3**	2.84E-5
order	–	0.40	1.44	0.90	1.72	**5.64**	5.48
$\varepsilon_0/2^6$	1.48	1.48	5.90E-1	2.19E-1	1.32E-1	4.21E-2	**7.04E-4**
order	–	0.00	1.32	1.43	0.72	1.65	**5.90**

Table 2.3.14: Temporal errors $e^r_\Phi(t=2)$ of S_{4c} under different τ and ε for the Dirac equation (1.1.17) in 1D in the simultaneously nonrelativistic and massless regime.

where $m_0 \geq 2$ depends on the regularity of the solution. Rigorous mathematical justification is still on-going.

- **In the simultaneously nonrelativistic and massless regime** We take $d = 1$, $\delta = 1$ and $\nu = \varepsilon$ in (1.1.17) with $\varepsilon \in (0,1]$. The initial data in (1.1.18) is taken as (2.3.2). In this parameter regime, the solution propagates waves with wavelength at $O(1)$ and $O(\varepsilon)$ in space and time, respectively. The problem is solved numerically on a bounded domain $\Omega = (-128, 128)$, i.e. $a = -128$ and $b = 128$ by S_{4c}. Similar to the nonrelativistic regime, the S_{4c} method converges uniformly with respect to $\varepsilon \in (0,1]$ at spectral order in space. Detailed numerical results are omitted here for brevity. Here we only present temporal errors by taking $h = h_e = \frac{1}{16}$ so that the spatial discretization error could be negligible. Table 2.3.14 shows the temporal errors $e^r_\Phi(t=2)$ for the wave function under different τ and $\varepsilon \in (0,1]$. Similarly, Table 2.3.15 and Table 2.3.16 depict the temporal errors $e^r_\rho(t=2)$ and $e^r_{\mathbf{J}}(t=2)$ for the probability and current, respectively.

  From Table 2.3.14-2.3.16, when $\tau \lesssim \varepsilon$, fourth-order convergence is observed for the

	$\tau_0 = 1$	$\tau_0/2$	$\tau_0/2^2$	$\tau_0/2^3$	$\tau_0/2^4$	$\tau_0/2^5$	$\tau_0/2^6$
$\varepsilon_0 = 1$	**8.62E-2**	3.48E-3	1.91E-4	1.17E-5	7.28E-7	4.54E-8	2.82E-9
order	–	4.63	4.19	4.03	4.01	4.00	4.01
$\varepsilon_0/2$	3.56E-1	**2.97E-2**	7.90E-4	4.56E-5	2.82E-6	1.76E-7	1.10E-8 0
order	–	**3.59**	5.23	4.12	4.01	4.00	4.00
$\varepsilon_0/2^2$	9.98E-1	2.83E-1	**1.22E-2**	2.54E-4	1.45E-5	8.95E-7	5.57E-8
order	–	1.82	**4.53**	5.59	4.13	4.02	4.01
$\varepsilon_0/2^3$	8.15E-1	5.58E-1	1.60E-1	**4.18E-3**	9.00E-5	5.29E-6	3.27E-7
order	–	0.55	1.80	**5.26**	5.54	4.09	4.02
$\varepsilon_0/2^4$	9.32E-1	7.05E-1	3.32E-1	1.02E-1	**1.69E-3**	3.69E-5	2.19E-6
order	–	0.40	1.09	1.70	**5.92**	5.52	4.08
$\varepsilon_0/2^5$	1.05	6.88E-1	3.28E-1	2.07E-1	6.70E-2	**8.68E-4**	1.63E-5
order	–	0.61	1.07	0.67	1.63	**6.27**	5.73
$\varepsilon_0/2^6$	8.39E-1	8.04E-1	4.76E-1	1.72E-1	1.27E-1	4.33E-2	**5.49E-4**
order	–	0.06	0.76	1.47	0.44	1.55	**6.30**

Table 2.3.15: Temporal errors $e_\rho^r(t=2)$ of S_{4c} under different τ and ε for the Dirac equation (1.1.17) in 1D in the simultaneously nonrelativistic and massless regime.

S_{4c} method in the relative error for the wave function, probability and current. This suggests that the ε-scalability for the S_{4c} method in the simultaneously nonrelativistic and massless regime is: $h = O(1)$ and $\tau = O(\varepsilon)$. In addition, noticing $\Phi = O(1)$, $\rho = O(1)$ and $\mathbf{J} = O(\varepsilon^{-1})$ when $0 \le \varepsilon \ll 1$, we can formally observe the following error bounds for $0 < \varepsilon \le 1$, $\tau \lesssim \varepsilon$ and $0 \le n \le \frac{T}{\tau}$

$$\begin{aligned} &\|\Phi^n - \Phi(t_n,\cdot)\|_{l^2} \lesssim h^{m_0} + \frac{\tau^4}{\varepsilon^3}, \quad \|\rho^n - \rho(t_n,\cdot)\|_{l^2} \lesssim h^{m_0} + \frac{\tau^4}{\varepsilon^3}, \\ &\|\mathbf{J}^n - \mathbf{J}(t_n,\cdot)\|_{l^2} \lesssim \frac{1}{\varepsilon}\left(h^{m_0} + \frac{\tau^4}{\varepsilon^3}\right). \end{aligned} \tag{2.3.15}$$

where $m_0 \ge 2$ depends on the regularity of the solution. Rigorous mathematical justification is still on-going.

Based on the discussion in the introduction chapter and numerical comparison results in this section, Table 2.3.17 summarizes spatial/temporal wavelengths of the Dirac

	$\tau_0 = 1$	$\tau_0/2$	$\tau_0/2^2$	$\tau_0/2^3$	$\tau_0/2^4$	$\tau_0/2^5$	$\tau_0/2^6$
$\varepsilon_0 = 1$	**2.03E-1**	7.11E-3	4.03E-4	2.47E-5	1.54E-6	9.61E-8	5.98E-9
order	–	4.84	4.14	4.03	4.01	4.00	4.01
$\varepsilon_0/2$	7.37E-1	**5.58E-2**	1.89E-3	1.11E-4	6.84E-6	4.26E-7	2.66E-8
order	–	**3.72**	4.88	4.09	4.02	4.00	4.00
$\varepsilon_0/2^2$	1.34	4.30E-1	**1.81E-2**	5.59E-4	3.31E-5	2.05E-6	1.28E-7
order	–	1.64	**4.57**	5.01	4.08	4.02	4.00
$\varepsilon_0/2^3$	1.20	7.03E-1	2.30E-1	**6.14E-3**	1.89E-4	1.13E-5	7.00E-7
order	–	0.77	1.61	**5.23**	5.02	4.06	4.01
$\varepsilon_0/2^4$	1.36	1.04	4.15E-1	1.31E-1	**2.52E-3**	7.59E-5	4.57E-6
order	–	0.39	1.32	1.66	**5.71**	5.05	4.05
$\varepsilon_0/2^5$	1.63	1.32	5.79E-1	2.47E-1	8.28E-2	**1.27E-3**	3.26E-5
order	–	0.30	1.19	1.23	1.58	**6.03**	5.28
$\varepsilon_0/2^6$	1.38	1.47	8.97E-1	3.04E-1	1.52E-1	5.54E-2	**7.52E-4**
order	–	-0.09	0.71	1.56	1.00	1.45	**6.20**

Table 2.3.16: Temporal errors $e^r_{\mathbf{J}}(t=2)$ of S_{4c} under different τ and ε for the Dirac equation (1.1.17) in 1D in the simultaneously nonrelativistic and massless regime.

equation under different parameter regimes and the corresponding spatial/temporal resolution of the S_{4c} method.

2.4 Application to the dynamics of graphene

In this section, we show the simulation of the dynamics of graphene through a 2D numerical example applying the S_{4c} method. In the example, we choose $\Omega = (-18,18) \times (-18,18)$, and the initial state ($\mathbf{x} = (x_1, x_2)^T$):

$$\phi_1(0,\mathbf{x}) = e^{-\frac{x_1^2+x_2^2}{2}}, \quad \phi_2(0,\mathbf{x}) = e^{-\frac{(x_1-1)^2+x_2^2}{2}}, \tag{2.4.1}$$

the density of which is illustrated in Figure 2.4.1.

During the computation, the mesh size is set to be $h = \frac{1}{16}$, and the time step size is fixed at $\tau = 0.01$. There is no magnetic potential ($\mathbf{x} = (x_1, x_2)^T$), i.e.

$$A_j(\mathbf{x}) \equiv 0, \quad j = 1,2, \quad \mathbf{x} \in \Omega.$$

	Spatial wavelength	Temporal wavelength	Spatial accuracy	Temporal accuracy	Spatial resolution	Temporal resolution
Standard regime	$O(1)$	$O(1)$	spectral	$O(\tau^4)$	$O(1)$	$O(1)$
Nonrelativistic regime	$O(1)$	$O(\varepsilon^2)$	spectral	$O(\frac{\tau^4}{\varepsilon^6})$	$O(1)$	$O(\varepsilon^2)$
Semiclassical regime	$O(\delta)$	$O(\delta)$	spectral	$O(\frac{\tau^4}{\delta^4})$	$O(\delta)$	$O(\delta)$
Nonrelativistic &massless regime	$O(1)$	$O(\varepsilon)$	spectral	$O(\frac{\tau^4}{\varepsilon^3})$	$O(1)$	$O(\varepsilon)$
Massless regime	$O(1)$	$O(1)$	spectral	$O(\tau^4)$	$O(1)$	$O(1)$

Table 2.3.17: Spatial/temporal wavelengths of the Dirac equation under different parameter regimes and the corresponding spatial/temporal resolution of the S_{4c} method.

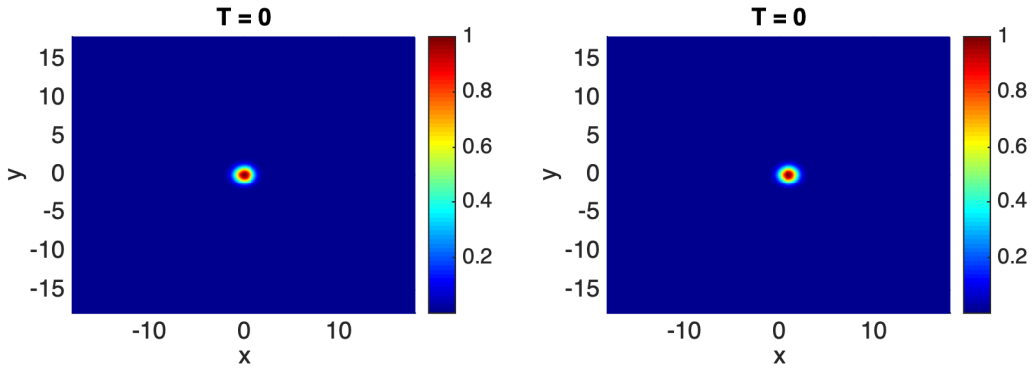

Figure 2.4.1: The initial density of the example, the left figure is for $\rho_1(0,\mathbf{x})$, and the right figure is for $\rho_2(0,\mathbf{x})$.

From $t=0$ to $t=6$, we consider the nonrelativistic regime of the Dirac equation, and choose $\varepsilon = 10^{-3}$. The electric potential $V(\mathbf{x})$ is chosen to be the honeycomb lattice potential ($\mathbf{x}=(x_1,x_2)^T$):

$$V(\mathbf{x}) = \cos(\frac{4\pi}{\sqrt{3}}\mathbf{e_1}\cdot\mathbf{x}) + \cos(\frac{4\pi}{\sqrt{3}}\mathbf{e}_2\cdot\mathbf{x}) + \cos(\frac{4\pi}{\sqrt{3}}\mathbf{e}_3\cdot\mathbf{x}), \quad t\in[0,5], \quad \mathbf{x}\in\Omega, \tag{2.4.2}$$

with

$$\mathbf{e}_1 = (-1,0)^T, \quad \mathbf{e}_2 = (1/2,\sqrt{3}/2)^T, \quad \mathbf{e}_3 = (1/2,-\sqrt{3}/2)^T. \tag{2.4.3}$$

From $t = 6$ to $t = 12$, we consider the Dirac equation with $\varepsilon = 1$, and discard the electric potential, which means, we set ($\mathbf{x} = (x_1, x_2)^T$)

$$V(\mathbf{x}) \equiv 0, \quad \mathbf{x} \in \Omega. \tag{2.4.4}$$

Figure 2.4.2 and Figure 2.4.3 depict the densities $\rho_j(t, \mathbf{x}) = |\phi_j(t, \mathbf{x})|^2$ (j = 1, 2) from $t = 0$ to $t = 12$.

From the two figures, we find out that the dynamics of the density depends heavily on ε, which stands for the regime the equation is in. When $\varepsilon = 10^{-3}$, it is in the nonrelativistic regime, and under the honeycomb lattice potential, it will generate a honeycomb-like density, which is similar to the electron density in graphene. After that, by taking $\varepsilon = 1$, we set it in the classical regime to simulate the dynamics of graphene without external electromagnetic potentials. The figures show that the density will fluctuate in the Zitterbewegung form, as has been demonstrated through experiments. Remarkably, here the time step size $\tau = 0.01$ is suitable for both two regimes, i.e., it is irrelevant to the dimensionless parameter ε.

2.5 Extension to the case of time-dependent potentials

In this section, we aim to extend the S_{4c} introduced in this chapter earlier to the case of time-dependent potentials $V(t, x)$ and $A_1(t, x)$. For simplicity, we only consider the 1D dimensionless Dirac equation (1.1.17) with the initial condition (1.1.18) ($d = 1$). Extension to higher dimensions is straightforward.

2.5.1 The method

For illustration, we first consider a model equation

$$\partial_t \psi(t) = H(t)\psi(t), \quad t > 0, \quad \text{with } \psi(t_0) = \psi_0, \tag{2.5.1}$$

where $H(t)$ is a time-dependent operator. Suppose the exact solution $\psi(t)$ propagates with the operator $U(t, t_0)$, i.e.,

$$\psi(t) = U(t, t_0)\psi(t_0), \tag{2.5.2}$$

then plugging into (2.5.1), we can get the differential equation

$$\partial_t U(t, t_0) = H(t)U(t, t_0), \quad t > 0, \tag{2.5.3}$$

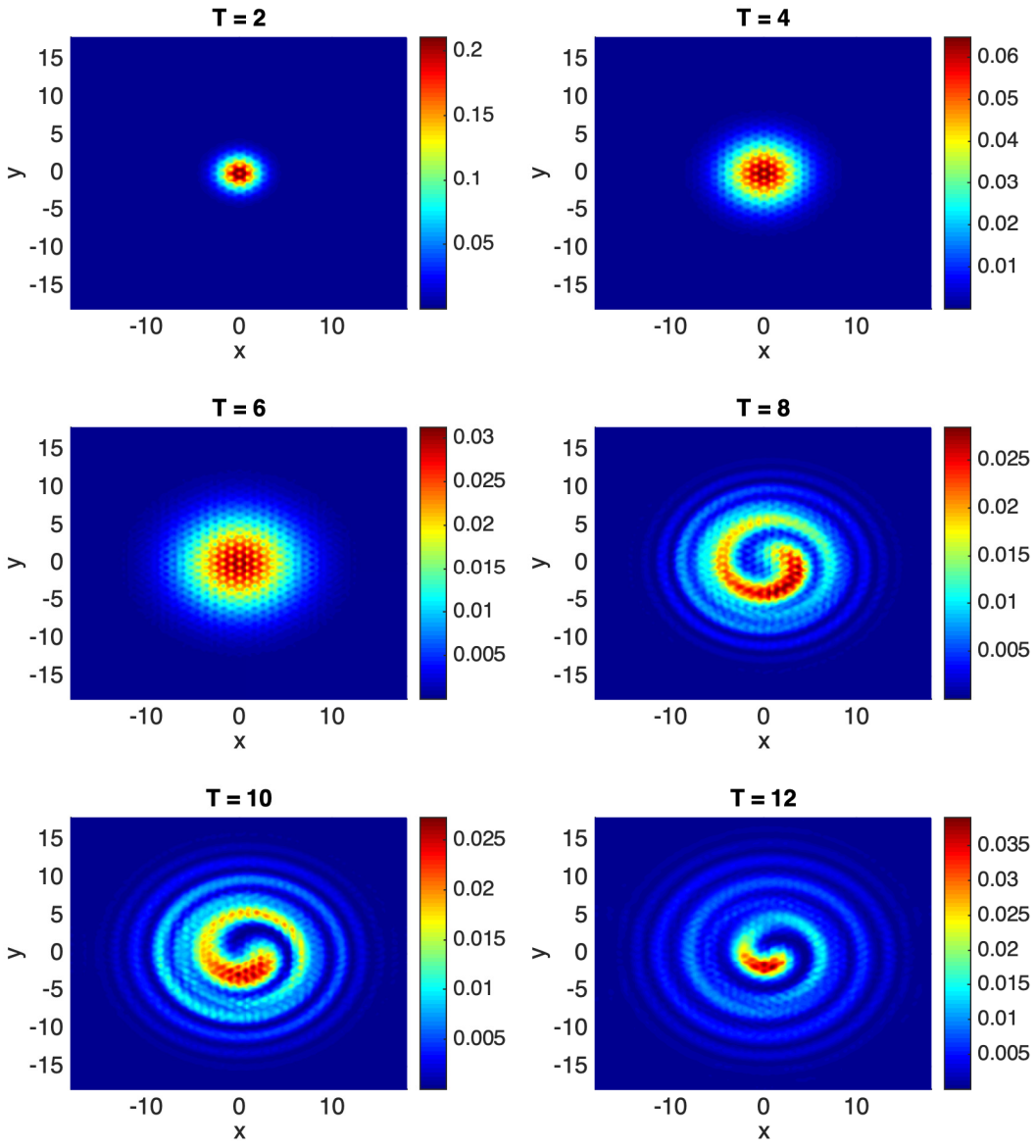

Figure 2.4.2: Dynamics of the density $\rho_1(t,\mathbf{x})$ to $T = 12$.

with $U(t_0,t_0) = Id$ the identity operator. Take $\triangle t > 0$, then by Taylor expansion,

$$\begin{aligned} U(t_0+\triangle t,t_0) &= U(t_0,t_0)+\triangle t\partial_t U(t_0,t_0)+O((\triangle t)^2) \\ &= (1+\triangle t H(t_0))+O((\triangle t)^2) \\ &= e^{\triangle t H(t_0)}+O((\triangle t)^2). \end{aligned} \tag{2.5.4}$$

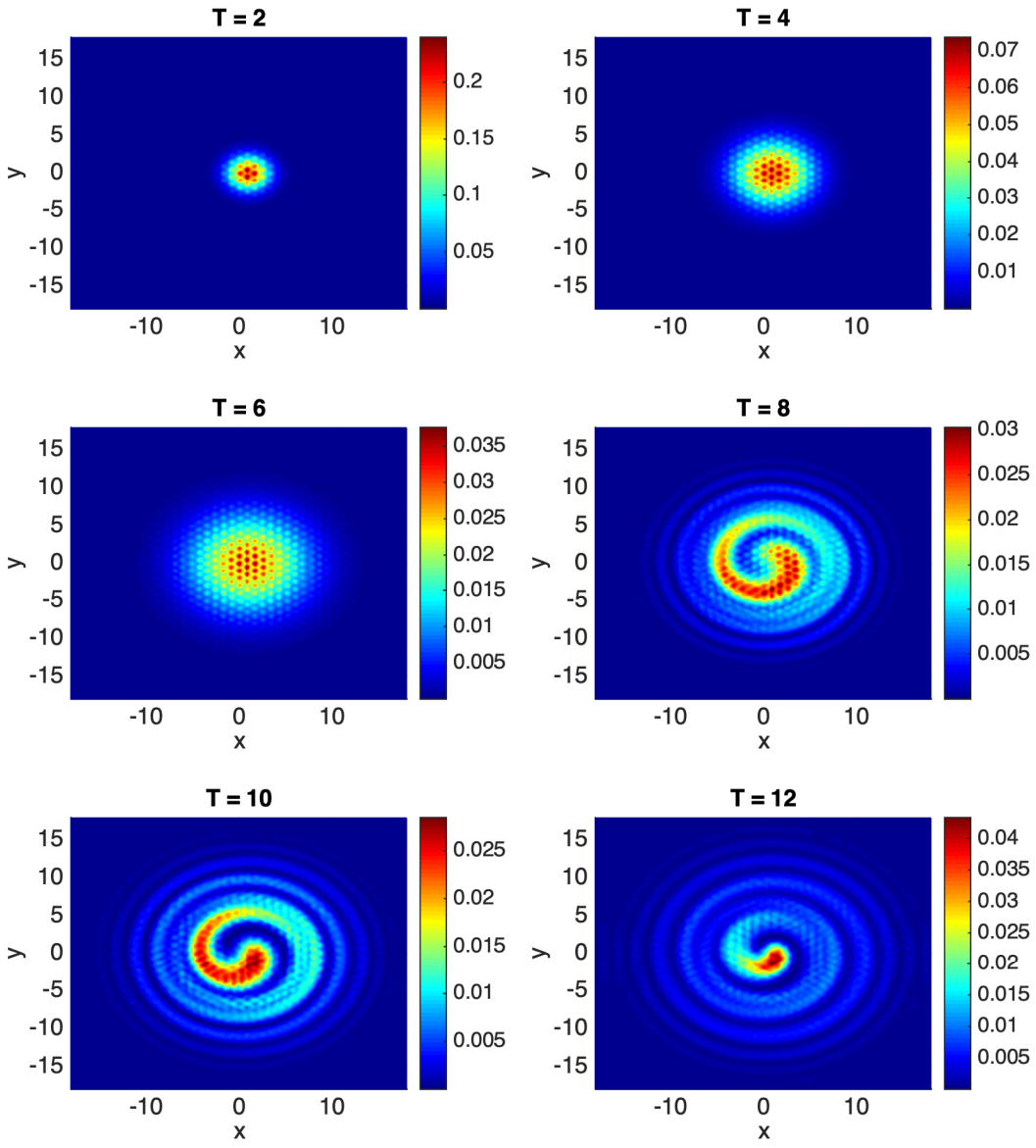

Figure 2.4.3: Dynamics of the density $\rho_2(t,\mathbf{x})$ to $T = 12$.

As a result, because of the fact that

$$U(t+\triangle t,t) = \Pi_{k=1}^{n} U(t+\frac{k}{n}\triangle t, t+\frac{k-1}{n}\triangle t) \tag{2.5.5}$$

holds for any positive integer n, we have

$$U(t+\triangle t,t) = \lim_{n\to\infty} e^{\frac{\triangle t}{n}H(t+\frac{k-1}{n}\triangle t)}...e^{\frac{\triangle t}{n}H(t+\frac{\triangle t}{n})}e^{\frac{\triangle t}{n}H(t)} \tag{2.5.6}$$

On the other hand, from (2.5.3) with initial condition $U(t_0,t_0) = Id$, we have

$$\begin{aligned} U(t,t_0) &= Id + \int_{t_0}^{t} H(s)U(s,t_0)ds \\ &= Id + \int_{t_0}^{t} H(s_1)ds_1 + \int_{t_0}^{t} H(s_1)\int_{t_0}^{s} H(s_2)U(s_2,t_0)ds_2 ds_1 \\ &= Id + \sum_{n=1}^{\infty}\int_{t_0}^{t}\int_{t_0}^{s_1}\ldots\int_{t_0}^{s_{n-1}} ds_n\ldots ds_1 H(s_1)\ldots H(s_n) \qquad (2.5.7) \\ &=: \mathscr{T}(e^{\int_{t_0}^{t} H(s)ds}), \qquad (2.5.8) \end{aligned}$$

where $\mathscr{T}(\cdot)$ is defined as the time-ordering operator, with the expression given in (2.5.7).

From the above discussion, we get

$$\begin{aligned} U(t+\triangle t,t) &= \mathscr{T}(e^{\int_{t}^{t+\triangle t} H(s)ds}) \\ &= \lim_{n\to\infty} e^{\frac{\triangle t}{n}H(t+\frac{k-1}{n}\triangle t)}\ldots e^{\frac{\triangle t}{n}H(t+\frac{\triangle t}{n})}e^{\frac{\triangle t}{n}H(t)}. \end{aligned}$$

Define a forward time derivative operator [45] $\mathscr{D} := \frac{\overleftarrow{\partial}}{\partial t}$ with

$$F(t)e^{\triangle t\mathscr{D}}G(t) = F(t+\triangle t)G(t), \qquad (2.5.9)$$

where $F(\cdot)$ and $G(\cdot)$ are any two time-dependent operators. Then we have the following lemma.

Lemma 2.7. *The following equality holds true for any time-dependent operator $H(t)$.*

$$\mathscr{T}(e^{\int_{t}^{t+\triangle t} H(s)ds}) = \exp[\triangle t(H(t)+\mathscr{D})]. \qquad (2.5.10)$$

Proof. We start from the right-hand-side of (2.5.10).

$$\begin{aligned} \exp[\triangle t(H(t)+\mathscr{D})] &= \lim_{n\to\infty}\left(e^{\frac{\triangle t}{n}\mathscr{D}}e^{\frac{\triangle t}{n}H(t)}\right)^n \\ &= \lim_{n\to\infty} e^{\frac{\triangle t}{n}\mathscr{D}}e^{\frac{\triangle t}{n}H(t)}\ldots e^{\frac{\triangle t}{n}\mathscr{D}}e^{\frac{\triangle t}{n}H(t)}e^{\frac{\triangle t}{n}\mathscr{D}}e^{\frac{\triangle t}{n}H(t)} \\ &= \lim_{n\to\infty} e^{\frac{\triangle t}{n}H(t+\frac{k-1}{n}\triangle t)}\ldots e^{\frac{\triangle t}{n}H(t+\frac{\triangle t}{n})}e^{\frac{\triangle t}{n}H(t)} \\ &= \mathscr{T}(e^{\int_{t}^{t+\triangle t} H(s)ds}), \qquad (2.5.11) \end{aligned}$$

where the first equality comes from the fact that $e^{x(A+B)} = \lim_{n\to\infty}\left(e^{\frac{x}{n}A}e^{\frac{x}{n}B}\right)^n$. □

Under this lemma, for the model problem (2.5.1), suppose $H(t) = T + W(t)$, define $\widetilde{T} = T + \mathscr{D}$, then the solution can be expressed as

$$\begin{aligned} \psi(t+\triangle t) &= U(t+\triangle t, t)\psi(t) = \exp[\triangle t(T + W(t) + \mathscr{D})]\psi(t) \\ &= \exp[\triangle t(\widetilde{T} + W(t))]\psi(t). \end{aligned} \tag{2.5.12}$$

We could apply the ordinary splitting methods for (2.5.12), for example, Strang splitting gives

$$\begin{aligned} \psi(t+\triangle t) \approx \widetilde{S}_2(\triangle t)\psi(t) &:= e^{\frac{\triangle t}{2}\widetilde{T}} e^{\triangle t W(t)} e^{\frac{\triangle t}{2}\widetilde{T}} \psi(t) \\ &= e^{\frac{\triangle t}{2}T} e^{\frac{\triangle t}{2}\mathscr{D}} e^{\triangle t W(t)} e^{\frac{\triangle t}{2}T} e^{\frac{\triangle t}{2}\mathscr{D}} \\ &= e^{\frac{\triangle t}{2}T} e^{\triangle t W(t+\frac{\triangle t}{2})} e^{\frac{\triangle t}{2}T}, \end{aligned} \tag{2.5.13}$$

where we use $e^{\frac{\triangle t}{2}\widetilde{T}} = e^{\frac{\triangle t}{2}T} e^{\frac{\triangle t}{2}\mathscr{D}}$, as it is straightforward that $[T, \mathscr{D}] = 0$.

Similarly, we can apply the fourth-order compact splitting to (2.5.12), and get the expression

$$\begin{aligned} \psi(t+\triangle t) \approx \widetilde{S_{4c}}(\triangle t)\psi(t) &:= e^{\frac{\triangle t}{6}W(t)} e^{\frac{\triangle t}{2}\widetilde{T}} e^{\frac{2\triangle t}{3}\overline{W}(t)} e^{\frac{\triangle t}{2}\widetilde{T}} e^{\frac{\triangle t}{6}W(t)} \psi(t) \\ &= e^{\frac{\triangle t}{6}W(t+\triangle t)} e^{\frac{\triangle t}{2}T} e^{\frac{2\triangle t}{3}\widehat{W}(t+\frac{\triangle t}{2})} e^{\frac{\triangle t}{2}T} e^{\frac{\triangle t}{6}W(t)} \psi(t), \end{aligned} \tag{2.5.14}$$

where

$$\begin{aligned} \widehat{W}(t) &= W(t) + \frac{1}{48}(\triangle t)^2 [W(t), [\widetilde{T}, W(t)]] \\ &= W(t) + \frac{1}{48}(\triangle t)^2 [W(t), [T, W(t)]] + \frac{1}{48}(\triangle t)^2 [W(t), [\mathscr{D}, W(t)]]. \end{aligned} \tag{2.5.15}$$

Through simple computation, we have

$$\begin{aligned} [W(t), [\mathscr{D}, W(t)]] &= [W(t), [\mathscr{D}W(t) - W(t)\mathscr{D}]] = [W(t), [\mathscr{D}W(t) - (W'(t) + \mathscr{D}W(t))]] \\ &= [W(t), -W'(t)] = 0, \end{aligned} \tag{2.5.16}$$

As a result,

$$\widehat{W}(t) = W(t) + \frac{1}{48}(\triangle t)^2 [W(t), [T, W(t)]]. \tag{2.5.17}$$

Now we can move on to (1.1.17). For simplicity, we only consider the 1D case. Extension to 2D and (1.1.7) for $d = 1, 2, 3$ is straightforward. Actually, it is similar to the case with time-independent electromagnetic potentials. Define

$$T = -\frac{1}{\varepsilon}\sigma_1 \partial_x - \frac{i\nu}{\delta\varepsilon^2}\sigma_2\sigma_3, \quad W(t) = -\frac{i}{\delta}\left(V(t,x)I_2 - A_1(t,x)\sigma_1\right), \tag{2.5.18}$$

then we could derive

$$[W(t),[T,W(t)]] = -\frac{4i\nu}{\delta^3\varepsilon^2}A_1^2(t,x)\sigma_3. \tag{2.5.19}$$

As a result, the semi-discretized fourth-order compact time-splitting method (S_{4c}) could be defined as:

$$\Phi^{n+1}(x) = e^{\frac{1}{6}\tau W(t_{n+1})}e^{\frac{1}{2}\tau T}e^{\frac{2}{3}\tau\widehat{W}(t_n+\tau/2)}e^{\frac{1}{2}\tau T}e^{\frac{1}{6}\tau W(t_n)}\Phi^n(x), \quad n=0,1,..., \tag{2.5.20}$$

where $\widehat{W}(t)$ is defined above. Here $\Phi^n(x)$ is the semi-discretized approximation of $\Phi(t,x)$ at $t=t_n:=n\tau$, with τ the time step size. The initial value $\Phi^0(x):=\Phi_0(x)$ is given.

2.5.2 Numerical results

This sections give the results of applying S_{4c} for the Dirac equation with time-dependent electromagnetic potentials in the nonrelativistic regime and the semiclassical regime. In the numerical examples, the time-dependent potentials are taken as

$$V(t,x) = \frac{1-tx}{1+t^2x^2}, \quad A_1(t,x) = \frac{(tx+1)^2}{1+t^2x^2}, \quad t>0, \quad x\in\mathbb{R}. \tag{2.5.21}$$

And the initial condition is set to be

$$\phi_1(0,x) = e^{-x^2/2}, \quad \phi_2(0,x) = e^{-(x-1)^2/2}, \quad x\in\mathbb{R}, \tag{2.5.22}$$

The relative error is quantified as

$$e_\Phi^r(t_n) = \frac{\|\Phi^n - \Phi(t_n,\cdot)\|_{l^2}}{\|\Phi(t_n,\cdot)\|_{l^2}}. \tag{2.5.23}$$

In computation, the problem is solved on a bounded domain Ω with periodic boundary conditions.

I. In the nonrelativistic regime

Take $\delta=\nu=1$. During the computation, the domain is set to be $\Omega=(-32,32)$. To obtain the 'exact' solution, fine mesh size $h_e=1/16$ and time step size $\tau_e=10^{-5}$ are used. Results are shown in Table 2.5.1.

From the table, it is clearly observed that when $\tau\lesssim\varepsilon^2$, there is fourth-order convergence for the S_{4c} method. This suggests that the ε-scalability for the S_{4c} method in the nonrelativistic

$e^r_\Phi(t=6)$	$\tau_0=1$	$\tau_0/2^2$	$\tau_0/2^4$	$\tau_0/2^6$	$\tau_0/2^8$	$\tau_0/2^{10}$
$\varepsilon_0=1$	1.79E-1	**1.67E-3**	1.50E-6	7.02E-9	<1E-10	<1E-10
order	–	**3.37**	5.06	3.87	–	–
$\varepsilon_0/2$	9.81E-1	2.12E-2	**3.01E-5**	1.06E-7	5.12E-10	<1E-10
order	–	2.77	**4.73**	4.08	3.84	–
$\varepsilon_0/2^2$	1.61	1.87E-1	1.87E-3	**4.72E-6**	1.83E-8	<1E-10
order	–	1.55	3.33	**4.31**	4.00	–
$\varepsilon_0/2^3$	1.37	1.55	4.49E-2	3.00E-4	**1.08E-6**	4.21E-9
order	–	-0.09	2.55	3.61	**4.06**	4.00
$\varepsilon_0/2^4$	1.42	1.62	5.39E-1	1.10E-2	7.25E-5	**2.63E-7**
order	–	-0.10	0.80	2.81	3.62	**4.05**

Table 2.5.1: Temporal errors $e^r_\Phi(t=6)$ of S_{4c} under different τ and ε for the Dirac equation (1.1.17) in 1D in the nonrelativistic regime.

regime is: $h = O(1)$ and $\tau = O(\varepsilon^2)$, which is the same as the case with time-independent potentials.

II. In the semiclassical regime

Take $\varepsilon = \nu = 1$. During the computation, the domain is set to be $\Omega = (-16, 16)$. To obtain the 'exact' solution, fine mesh size $h_e = 1/16$ and time step size $\tau_e = 10^{-4}$ are used. Results are shown in Table 2.5.2.

From the table, when $\tau \lesssim \delta$, fourth-order convergence in time is observed for the S_{4c} method, which suggests that the δ-scalability for the S_{4c} method in the semiclassical regime is: $h = O(\delta)$ and $\tau = O(\delta)$. The result is also the same as the case with time-independent potentials.

$e^r_\Phi(t=2)$	$\tau_0=1$	$\tau_0/2$	$\tau_0/2^2$	$\tau_0/2^3$	$\tau_0/2^4$	$\tau_0/2^5$	$\tau_0/2^6$
$\delta_0=1$	1.38E-1	**1.24E-2**	3.43E-4	1.48E-5	9.10E-7	5.66E-8	3.54E-9
order	–	**3.48**	5.17	4.53	4.03	4.01	4.00
$\delta_0/2$	8.81E-1	1.06E-1	**4.60E-3**	2.71E-4	1.69E-5	1.05E-6	6.57E-8
order	–	3.06	**4.52**	4.08	4.01	4.00	4.00
$\delta_0/2^2$	1.41E	9.90E-1	8.55E-2	**4.79E-3**	2.94E-4	1.83E-5	1.14E-6
order	–	0.52	3.53	**4.16**	4.03	4.01	4.00
$\delta_0/2^3$	1.44	1.49	9.43E-1	1.03E-1	**5.72E-3**	3.49E-4	2.16E-5
order	–	-0.05	0.66	3.20	**4.17**	4.04	4.01
$\delta_0/2^4$	1.49	1.38	1.43	1.09	2.01E-1	**1.26E-2**	1.52E-3
order	–	0.11	-0.05	0.39	2.44	**4.00**	3.05
$\delta_0/2^5$	1.39	1.44	1.43	1.47	1.34	2.31E-1	**1.09E-2**
order	–	-0.04	0.01	-0.04	0.13	2.54	**4.40**

Table 2.5.2: Temporal errors $e^r_\Phi(t=2)$ of S_{4c} under different τ and δ for the Dirac equation (1.1.17) in 1D in the semiclassical regime.

Chapter 3

Super-Resolution of Time-splitting Methods for the Dirac Equation

In this chapter, a superior property called super-resolution of time-splitting methods for the Dirac equation without magnetic potential in the nonrelativisitic regime is exhibited and studied rigorously. This advantageous property makes the time-splitting methods perform much better in this case, as the time step size no longer needs to be dependent on the small parameter ε.

3.1 Introduction

Time-splitting methods have been utilized to study the dynamics of the Dirac equation, and it is found to be efficient in the nonrelativistic regime [15]. In fact, when dealing with oscillatory problems, the splitting methods usually perform much better than traditional numerical methods [21, 76]. For instance, in order to obtain "correct" observables of the Schrödinger equation in the semiclassical regime, the time-splitting spectral method requires much weaker constraints on time step size and mesh size than the finite difference methods [21]. Similar properties have been observed for the nonlinear Schrödinger equation (NLSE)/Gross-Pitaevskii equation (GPE) in the semiclassical regime [3] and the Enrenfest dynamics [57]. However, in general, splitting methods still suffer from the mesh size/time step constraints related to the high frequencies in the aforementioned problems, i.e. they need to obey the resolution constraint determined by the Shannon's sampling theorem [104] – in order to resolve a wave one needs to use a few grid points per wavelength. For Dirac equation

in the nonrelativistic regime, from the analysis in [15], the error bound for second order Strang splitting TSFP (S_2) depends on the small parameter ε as τ^2/ε^4, which corresponds to such constraints.

In this chapter, we report a surprising finding that the splitting methods are uniformly accurate (w.r.t. the rapid oscillations), when applied to the Dirac equation in the nonrelativistic regime in the absence of external magnetic field. This fact reveals that there is no mesh size/time step restriction for splitting methods in this situation, e.g. the splitting methods have **super-resolution**, which is highly nontrivial. Specifically, through our extensive numerical experiments, we find out that if the magnetic potentials $A_j \equiv 0$ for $j = 1, \ldots, d$ in (1.1.7) with $\delta = \nu = 1$, the errors of TSFP are then independent of ε and uniform w.r.t. ε, i.e., S_2 for Dirac equation (1.1.7) with $\delta = \nu = 1$ without magnetic potentials A_j has super-resolution w.r.t. ε. In such case, (1.1.7) reduces to ($d = 1, 2, 3$, $\delta = \nu = 1$)

$$i\partial_t \Psi(t, \mathbf{x}) = \left(-\frac{i}{\varepsilon} \sum_{j=1}^{d} \alpha_j \partial_j + \frac{1}{\varepsilon^2} \beta + V(t, \mathbf{x}) I_4 \right) \Psi(t, \mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^d, \ t > 0, \tag{3.1.1}$$

with the initial value given in (1.1.8). In lower dimensions ($d = 1, 2$), the four component Dirac equation (3.1.1) can be reduced to the following two-component form for $\Phi(t, \mathbf{x}) = (\phi_1(t, \mathbf{x}), \phi_2(t, \mathbf{x}))^T \in \mathbb{C}^2$ ($d = 1, 2$) [15]:

$$i\partial_t \Phi(t, \mathbf{x}) = \left(-\frac{i}{\varepsilon} \sum_{j=1}^{d} \sigma_j \partial_j + \frac{1}{\varepsilon^2} \sigma_3 + V(t, \mathbf{x}) I_2 \right) \Phi(t, \mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^d, \ t > 0, \tag{3.1.2}$$

with initial value

$$\Phi(t = 0, \mathbf{x}) = \Phi_0(\mathbf{x}), \qquad \mathbf{x} \in \mathbb{R}^d. \tag{3.1.3}$$

The two-component form (3.1.2) is widely used in lower dimensions $d = 1, 2$ due to its simplicity compared to the four component form (3.1.1).

Our extensive numerical studies and theoretical analysis show that for first-order, second-order, and even higher order time-splitting Fourier pseudospectral methods, there are always uniform error bounds w.r.t. $\varepsilon \in (0, 1]$. In other words, the splitting methods can capture the solutions accurately even if the time step size τ is much larger than the sampled wavelength at $O(\varepsilon^2)$, i.e. they exhibit **super-resolution** in the sense of breaking the resolution constraint under the Shannon's sampling theorem [104]. This super-resolution property of the splitting

methods makes them more efficient and reliable for solving the Dirac equation without magnetic potentials in the nonrelativisitc regime, compared to other numerical approaches in the literature. In the sequel, we will study rigorously the super-resolution phenomenon for first-order (S_1) and second-order (S_2) time-splitting methods, and present numerical results to validate the conclusions.

3.2 Semi-discretization

In this section, we recall the first- and second-order time-splitting methods applied to the Dirac equation. For simplicity of presentation, we only carry out the splitting methods and corresponding analysis for (3.1.2) in 1D ($d = 1$). Generalization to (3.1.1) and/or higher dimensions is straightforward and results remain valid without modifications.

Denote the Hermitian operator

$$\mathscr{T}^{\varepsilon} = -i\varepsilon\sigma_1\partial_x + \sigma_3, \quad x \in \mathbb{R}, \tag{3.2.1}$$

then the Dirac equation (3.1.2) in 1D can be written as

$$i\partial_t\Phi(t,x) = \frac{1}{\varepsilon^2}\mathscr{T}^{\varepsilon}\Phi(t,x) + V(t,x)\Phi(t,x), \quad x \in \mathbb{R}, \tag{3.2.2}$$

with initial value

$$\Phi(0,x) = \Phi_0(x), \quad x \in \mathbb{R}. \tag{3.2.3}$$

Choose $\tau > 0$ to be the time step size and $t_n = n\tau$ for $n = 0, 1, \ldots$ as the time steps. Denote $\Phi^n(x)$ as the numerical approximation of $\Phi(t_n, x)$, where $\Phi(t,x)$ is the exact solution to (3.2.2) with (3.2.3), then the semi-discretization of the first- and second-order time-splitting methods can be expressed as follows.

First-order splitting (Lie-Trotter splitting). The discrete-in-time first-order splitting (S_1) is written as [123]

$$\Phi^{n+1}(x) = e^{-\frac{i\tau}{\varepsilon^2}\mathscr{T}^{\varepsilon}} e^{-i\int_{t_n}^{t_{n+1}} V(s,x)\,ds}\Phi^n(x), \quad x \in \mathbb{R}, \tag{3.2.4}$$

with $\Phi^0(x) = \Phi_0(x)$.

Second-order splitting (Strang splitting). The discrete-in-time second-order splitting (S_2) is written as [113]

$$\Phi^{n+1}(x) = e^{-\frac{i\tau}{2\varepsilon^2}\mathscr{T}^\varepsilon} e^{-i\int_{t_n}^{t_{n+1}} V(s,x)\,ds} e^{-\frac{i\tau}{2\varepsilon^2}\mathscr{T}^\varepsilon} \Phi^n(x), \quad x \in \mathbb{R}. \tag{3.2.5}$$

with $\Phi^0(x) = \Phi_0(x)$.

3.3 Uniform error bounds

For any $T > 0$, we consider smooth enough solutions, i.e. we assume the electric potential satisfies

$$(A) \qquad V(t,x) \in W^{m,\infty}([0,T];L^\infty(\mathbb{R})) \cap L^\infty([0,T];W^{2m+m_*,\infty}(\mathbb{R})),$$

with $m \in \mathbb{N}^*$, $m_* \in \{0,1\}$. In addition, we assume the exact solution $\Phi(t,x)$ satisfies

$$(B) \qquad \Phi(t,x) \in L^\infty([0,T],(H^{2m+m_*}(\mathbb{R}))^2), \quad m \in \mathbb{N}^*, \quad m_* \in \{0,1\}.$$

We remark here that if the initial value $\Phi_0(\mathbf{x}) \in (H^{2m+m_*}(\mathbb{R}))^2$, then condition (B) is implied by condition (A).

For the numerical approximation $\Phi^n(x)$ obtained from S_1 (3.2.4) or S_2 (3.2.5), we introduce the error function

$$\mathbf{e}^n(x) = \Phi(t_n,x) - \Phi^n(x), \quad 0 \le n \le \frac{T}{\tau}, \tag{3.3.1}$$

then the following error estimates hold.

Theorem 3.1. *Let $\Phi^n(x)$ be the numerical approximation obtained from S_1 (3.2.4), then under the assumptions (A) and (B) with $m = 1$ and $m_* = 0$, we have the following error estimates*

$$\|\mathbf{e}^n(x)\|_{L^2} \lesssim \tau + \varepsilon, \quad \|\mathbf{e}^n(x)\|_{L^2} \lesssim \tau + \tau/\varepsilon, \quad 0 \le n \le \frac{T}{\tau}. \tag{3.3.2}$$

As a result, there is a uniform error bound for S_1

$$\|\mathbf{e}^n(x)\|_{L^2} \lesssim \tau + \max_{0<\varepsilon\le 1} \min\{\varepsilon, \tau/\varepsilon\} \lesssim \sqrt{\tau}, \quad 0 \le n \le \frac{T}{\tau}. \tag{3.3.3}$$

Theorem 3.2. *Let $\Phi^n(x)$ be the numerical approximation obtained from S_2 (3.2.5), then under the assumptions (A) and (B) with $m = 2$ and $m_* = 0$, we have the following error estimates*

$$\|\mathbf{e}^n(x)\|_{L^2} \lesssim \tau^2 + \varepsilon, \quad \|\mathbf{e}^n(x)\|_{L^2} \lesssim \tau^2 + \tau^2/\varepsilon^3, \quad 0 \le n \le \frac{T}{\tau}. \tag{3.3.4}$$

As a result, there is a uniform error bound for S_2

$$\|\mathbf{e}^n(x)\|_{L^2} \lesssim \tau^2 + \max_{0<\varepsilon\le 1} \min\{\varepsilon, \tau^2/\varepsilon^3\} \lesssim \sqrt{\tau}, \quad 0 \le n \le \frac{T}{\tau}. \tag{3.3.5}$$

Remark 3.1. *The error bounds in Theorem 3.1 can be expressed as*

$$\|\mathbf{e}^n(x)\|_{L^2} \le (C_1 + C_2 T)\|\Phi(t,x)\|_{L^\infty([0,T];(H^2)^2)} \left(\tau + \max_{0<\varepsilon\le 1} \min\{\varepsilon, \tau/\varepsilon\}\right), \quad 0 \le n \le \frac{T}{\tau},$$

and the error estimates in Theorem 3.2 can be restated as

$$\|\mathbf{e}^n(x)\|_{L^2} \le (C_3 + C_4 T)\|\Phi(t,x)\|_{L^\infty([0,T];(H^4)^2)} \left(\tau^2 + \max_{0<\varepsilon\le 1} \min\{\varepsilon, \tau^2/\varepsilon^3\}\right), \quad 0 \le n \le \frac{T}{\tau},$$

where C_j, $j = 1,2,3,4$ are constants depending only on $V(t,x)$.

This remark could be easily derived by examining the proofs of Theorem 3.1 and Theorem 3.2, and the details will be skipped. We notice that the constants before the error bounds have linear relations with T, instead of usual exponential ones.

We also remark that higher order time-splitting methods also share the super-resolution property, but for simplicity, we only focus on S_1 and S_2 here.

In the following, we derive the proof for Theorem 3.1 and Theorem 3.2, i.e. the uniform error bounds for the splitting methods S_1 and S_2. As $\mathscr{T}^\varepsilon$ is diagonalizable in the phase space (Fourier domain), it can be decomposed as [14, 15, 29]

$$\mathscr{T}^\varepsilon = \sqrt{Id - \varepsilon^2\Delta}\,\Pi_+^\varepsilon - \sqrt{Id - \varepsilon^2\Delta}\,\Pi_-^\varepsilon, \tag{3.3.6}$$

where $\Delta = \partial_{xx}$ is the Laplace operator in 1D and Id is the identity operator. Π_+^ε and Π_-^ε are projectors defined as

$$\Pi_+^\varepsilon = \frac{1}{2}\left[I_2 + \left(Id - \varepsilon^2\Delta\right)^{-1/2}\mathscr{T}^\varepsilon\right], \quad \Pi_-^\varepsilon = \frac{1}{2}\left[I_2 - \left(Id - \varepsilon^2\Delta\right)^{-1/2}\mathscr{T}^\varepsilon\right]. \tag{3.3.7}$$

It is straightforward to see that $\Pi_+^\varepsilon + \Pi_-^\varepsilon = I_2$, and $\Pi_+^\varepsilon \Pi_-^\varepsilon = \Pi_-^\varepsilon \Pi_+^\varepsilon = \mathbf{0}$, $(\Pi_\pm^\varepsilon)^2 = \Pi_\pm^\varepsilon$. Furthermore, through Taylor expansion, we have [29]

$$\Pi_+^\varepsilon = \Pi_+^0 + \varepsilon \mathscr{R}_1 = \Pi_+^0 - i\frac{\varepsilon}{2}\sigma_1 \partial_x + \varepsilon^2 \mathscr{R}_2, \quad \Pi_+^0 = \mathrm{diag}(1,0), \tag{3.3.8}$$

$$\Pi_-^\varepsilon = \Pi_-^0 - \varepsilon \mathscr{R}_1 = \Pi_-^0 + i\frac{\varepsilon}{2}\sigma_1 \partial_x - \varepsilon^2 \mathscr{R}_2, \quad \Pi_-^0 = \mathrm{diag}(0,1), \tag{3.3.9}$$

where $\mathscr{R}_1 : (H^m(\mathbb{R}))^2 \to (H^{m-1}(\mathbb{R}))^2$ for $m \geq 1$, $m \in \mathbb{N}^*$, and $\mathscr{R}_2 : (H^m(\mathbb{R}))^2 \to (H^{m-2}(\mathbb{R}))^2$ for $m \geq 2$, $m \in \mathbb{N}^*$ are uniformly bounded operators with respect to ε.

To help capture the features of solutions, denote

$$\mathscr{D}^\varepsilon = \frac{1}{\varepsilon^2}(\sqrt{Id - \varepsilon^2\Delta} - Id) = -(\sqrt{Id - \varepsilon^2\Delta} + Id)^{-1}\Delta, \tag{3.3.10}$$

which is a uniformly bounded operator with respect to ε from $(H^m(\mathbb{R}))^2$ to $(H^{m-2}(\mathbb{R}))^2$ for $m \geq 2$, then we have the decomposition for the unitary evolution operator $e^{\frac{it}{\varepsilon^2}\mathscr{T}^\varepsilon}$ as

$$e^{\frac{it}{\varepsilon^2}\mathscr{T}^\varepsilon} = e^{\frac{it}{\varepsilon^2}(\sqrt{Id-\varepsilon^2\Delta}\,\Pi_+^\varepsilon - \sqrt{Id-\varepsilon^2\Delta}\,\Pi_-^\varepsilon)} = e^{it/\varepsilon^2} e^{it\mathscr{D}^\varepsilon}\,\Pi_+^\varepsilon + e^{-it/\varepsilon^2} e^{-it\mathscr{D}^\varepsilon}\,\Pi_-^\varepsilon. \tag{3.3.11}$$

For the ease of the proof, we first introduce the following two lemmas for the Lie-Trotter splitting S_1 (3.2.4) and the Strang splitting S_2 (3.2.5), respectively. For simplicity, we denote $V(t) := V(t,x)$, and $\Phi(t) := \Phi(t,x)$ in short.

Lemma 3.1. *Let $\Phi^n(x)$ be the numerical approximation obtained from the Lie-Trotter splitting S_1 (3.2.4), then under the assumptions (A) and (B) with $m = 1$ and $m_* = 0$, we have*

$$\mathbf{e}^{n+1}(x) = e^{-\frac{i\tau}{\varepsilon^2}\mathscr{T}^\varepsilon} e^{-i\int_{t_n}^{t_{n+1}} V(s,x)ds}\mathbf{e}^n(x) + \eta_1^n(x) + \eta_2^n(x), \quad 0 \leq n \leq \frac{T}{\tau} - 1, \tag{3.3.12}$$

with $\|\eta_1^n(x)\|_{L^2} \lesssim \tau^2$, $\eta_2^n(x) = -ie^{-\frac{i\tau}{\varepsilon^2}\mathscr{T}^\varepsilon}\left(\int_0^\tau f_2^n(s)ds - \tau f_2^n(0)\right)$, where

$$\begin{aligned} f_2^n(s) =& e^{i2s/\varepsilon^2} e^{is\mathscr{D}^\varepsilon}\Pi_+^\varepsilon\left(V(t_n)\Pi_-^\varepsilon e^{is\mathscr{D}^\varepsilon}\Phi(t_n)\right) \\ &+ e^{-i2s/\varepsilon^2} e^{-is\mathscr{D}^\varepsilon}\Pi_-^\varepsilon\left(V(t_n)\Pi_+^\varepsilon e^{-is\mathscr{D}^\varepsilon}\Phi(t_n)\right). \end{aligned} \tag{3.3.13}$$

Proof. From the definition of $\mathbf{e}^n(x)$, noticing the Lie-Trotter splitting formula (3.2.4), we have

$$\mathbf{e}^{n+1}(x) = e^{-\frac{i\tau}{\varepsilon^2}\mathscr{T}^\varepsilon} e^{-i\int_{t_n}^{t_{n+1}} V(s,x)ds}\mathbf{e}^n(x) + \eta^n(x), \quad 0 \leq n \leq \frac{T}{\tau} - 1, \quad x \in \mathbb{R}, \tag{3.3.14}$$

where $\eta^n(x)$ is the local truncation error defined as

$$\eta^n(x) = \Phi(t_{n+1},x) - e^{-\frac{i\tau}{\varepsilon^2}\mathscr{T}^\varepsilon} e^{-i\int_{t_n}^{t_{n+1}} V(s,x)\,ds}\Phi(t_n,x), \quad x\in\mathbb{R}. \tag{3.3.15}$$

Noticing (3.2.2), applying Duhamel's principle, we derive

$$\Phi(t_{n+1},x) = e^{-\frac{i\tau}{\varepsilon^2}\mathscr{T}^\varepsilon}\Phi(t_n,x) - i\int_0^\tau e^{-\frac{i(\tau-s)}{\varepsilon^2}\mathscr{T}^\varepsilon} V(t_n+s,x)\Phi(t_n+s,x)ds, \tag{3.3.16}$$

while Taylor expansion gives

$$\begin{aligned} &e^{-\frac{i\tau}{\varepsilon^2}\mathscr{T}^\varepsilon} e^{-i\int_{t_n}^{t_{n+1}} V(s,x)\,ds}\Phi(t_n,x) \\ &= e^{-\frac{i\tau}{\varepsilon^2}\mathscr{T}^\varepsilon}\left(1 - i\int_{t_n}^{t_{n+1}} V(s,x)\,ds + O(\tau^2)\right)\Phi(t_n,x). \end{aligned} \tag{3.3.17}$$

Combining (3.3.16), (3.3.17) and (3.3.15), we get

$$\begin{aligned} \eta^n(x) =& \tau i e^{-\frac{i\tau}{\varepsilon^2}\mathscr{T}^\varepsilon} V(t_n,x)\Phi(t_n,x) - i\int_0^\tau e^{-\frac{i(\tau-s)}{\varepsilon^2}\mathscr{T}^\varepsilon}\left(V(t_n,x)e^{-\frac{is}{\varepsilon^2}\mathscr{T}^\varepsilon}\Phi(t_n,x)\right)ds \\ &+ \sum_{j=1}^2 R_j^n(x), \end{aligned} \tag{3.3.18}$$

where

$$\begin{aligned} R_1^n(x) &= e^{-\frac{i\tau}{\varepsilon^2}\mathscr{T}^\varepsilon}\left(\lambda_1^n(x) + \lambda_2^n(x)\right)\Phi(t_n,x), \\ R_2^n(x) &= -i\int_0^\tau e^{-\frac{i(\tau-s)}{\varepsilon^2}\mathscr{T}^\varepsilon}\left(V(t_n)\lambda_4^n(s,x) + \lambda_3^n(s,x)\Phi(t_n+s,x)\right)ds, \end{aligned}$$

with

$$\lambda_1^n(x) = e^{-i\int_{t_n}^{t_{n+1}} V(s,x)ds} - \left(1 - i\int_{t_n}^{t_{n+1}} V(s,x)ds\right), \tag{3.3.19}$$

$$\lambda_2^n(x) = -i\int_{t_n}^{t_{n+1}} V(u,x)\,du + i\tau V(t_n,x), \tag{3.3.20}$$

$$\lambda_3^n(s,x) = V(t_n+s,x) - V(t_n,x), \quad 0\le s\le\tau, \tag{3.3.21}$$

$$\lambda_4^n(s,x) = -i\int_0^s e^{-\frac{i(s-w)}{\varepsilon^2}\mathscr{T}^\varepsilon}\left(V(t_n+w,x)\Phi(t_n+w,x)\right)dw, \quad 0\le s\le\tau. \tag{3.3.22}$$

It is easy to see that for $0\le n\le\frac{T}{\tau}-1$,

$$\begin{aligned} &\|\lambda_1^n(x)\|_{L^\infty} \lesssim \tau^2\|V(t,x)\|_{L^\infty(L^\infty)}^2, \quad \|\lambda_2^n(x)\|_{L^\infty} \lesssim \tau^2\|\partial_t V(t,x)\|_{L^\infty(L^\infty)}, \\ &\|\lambda_3^n(s,x)\|_{L^\infty([0,\tau];L^\infty)} \lesssim \tau\|\partial_t V(t,x)\|_{L^\infty(L^\infty)}, \\ &\|\lambda_4^n(s,x)\|_{L^\infty([0,\tau];(L^2)^2)} \lesssim \tau\|V(t,x)\|_{L^\infty(L^\infty)}\|\Phi(t,x)\|_{L^\infty\left((L^2)^2\right)}, \end{aligned}$$

As a consequence, we obtain the following bounds for $0 \le n \le \frac{T}{\tau} - 1$,

$$\|R_1^n(x)\|_{L^2} \lesssim (\|\lambda_1^n(x)\|_{L^\infty} + \|\lambda_2^n(x)\|_{L^\infty}) \|\Phi(t_n)\|_{L^2} \lesssim \tau^2, \tag{3.3.23}$$

$$\begin{aligned}\|R_2^n(x)\|_{L^2} \lesssim \tau \Big(& \|V(t_n)\|_{L^\infty} \|\lambda_4^n(s,x)\|_{L^\infty([0,\tau];(L^2)^2)} \\ & + \|\lambda_3^n(s,x)\|_{L^\infty([0,\tau];L^\infty)} \|\Phi\|_{L^\infty((L^2)^2)} \Big) \lesssim \tau^2.\end{aligned} \tag{3.3.24}$$

Recalling $\eta_2^n(x)$ given in Lemma 3.1, we introduce for $0 \le s \le \tau$

$$f^n(s) := f^n(s,x) = e^{\frac{is}{\varepsilon^2}\mathscr{T}^\varepsilon} \left(V(t_n,x) e^{-\frac{is}{\varepsilon^2}\mathscr{T}^\varepsilon} \Phi(t_n,x) \right) = f_1^n(s) + f_2^n(s), \tag{3.3.25}$$

with f_2^n given in (3.3.13) and f_1^n from the decomposition (3.3.11) as

$$f_1^n(s) = e^{is\mathscr{D}^\varepsilon} \Pi_+^\varepsilon \left(V(t_n) e^{-is\mathscr{D}^\varepsilon} \Pi_+^\varepsilon \Phi(t_n) \right) + e^{-is\mathscr{D}^\varepsilon} \Pi_-^\varepsilon \left(V(t_n) e^{is\mathscr{D}^\varepsilon} \Pi_-^\varepsilon \Phi(t_n) \right),$$

and then $\eta^n(x)$ (3.3.18) can be written as

$$\begin{aligned}\eta^n(x) = & -ie^{-\frac{i\tau}{\varepsilon^2}\mathscr{T}^\varepsilon} \left(\int_0^\tau (f_1^n(s) + f_2^n(s))ds - \tau(f_1^n(0) + f_2^n(0)) \right) \\ & + R_1^n(x) + R_2^n(x).\end{aligned} \tag{3.3.26}$$

Now, it is easy to verify that $\eta^n(x) = \eta_1^n(x) + \eta_2^n(x)$ with $\eta_2^n(x)$ given in Lemma 3.1 if we let

$$\eta_1^n(x) = -ie^{-\frac{i\tau}{\varepsilon^2}\mathscr{T}^\varepsilon} \left(\int_0^\tau f_1^n(s)ds - \tau f_1^n(0) \right) + R_1^n(x) + R_2^n(x). \tag{3.3.27}$$

Noticing that

$$\begin{aligned}& \left\| e^{-\frac{i\tau}{\varepsilon^2}\mathscr{T}^\varepsilon} \left(\int_0^\tau f_1^n(s)ds - \tau f_1^n(0) \right) \right\|_{L^2} \\ & \lesssim \tau^2 \|\partial_s f_1^n(\cdot)\|_{L^\infty([0,\tau];(L^2)^2)} \lesssim \tau^2 \|V(t_n)\|_{W^{2,\infty}} \|\Phi(t_n)\|_{H^2},\end{aligned}$$

recalling the regularity assumptions (A) and (B), combining (3.3.23) and (3.3.24) , we can get

$$\begin{aligned}\|\eta_1^n(x)\|_{L^2} & \le \|R_1^n(x)\|_{L^2} + \|R_2^n(x)\|_{L^2} + \left\| e^{-\frac{i\tau}{\varepsilon^2}\mathscr{T}^\varepsilon} \left(\int_0^\tau f_1^n(s)ds - \tau f_1^n(0) \right) \right\|_{L^2} \\ & \lesssim \tau^2,\end{aligned}$$

which completes the proof of Lemma 3.1. □

Lemma 3.2. *Let $\Phi^n(x)$ be the numerical approximation obtained from the Strang splitting S_2 (3.2.5), then under the assumptions (A) and (B) with $m = 2$ and $m_* = 0$, we have for $0 \le n \le \frac{T}{\tau} - 1$,*

$$\mathbf{e}^{n+1}(x) = e^{-\frac{i\tau}{2\varepsilon^2}\mathscr{T}^\varepsilon} e^{-i\int_{t_n}^{t_{n+1}} V(s,x)ds} e^{-\frac{i\tau}{2\varepsilon^2}\mathscr{T}^\varepsilon} \mathbf{e}^n(x) + \eta_1^n(x) + \eta_2^n(x) + \eta_3^n(x), \tag{3.3.28}$$

with

$$\|\eta_1^n(x)\|_{L^2} \lesssim \tau^3, \quad \eta_2^n(x) = -ie^{-\frac{i\tau}{\varepsilon^2}\mathscr{T}^\varepsilon}\left(\int_0^\tau f_2^n(s)ds - \tau f_2^n(\tau/2)\right), \tag{3.3.29}$$

$$\eta_3^n(x) = -e^{-\frac{i\tau}{\varepsilon^2}\mathscr{T}^\varepsilon}\left(\int_0^\tau \int_0^s \sum_{j=2}^4 g_j^n(s,w)dwds - \frac{\tau^2}{2}\sum_{j=2}^4 g_j^n(\tau/2,\tau/2)\right), \tag{3.3.30}$$

where

$$\begin{aligned}
f_2^n(s) =& e^{\frac{i2s}{\varepsilon^2}} e^{is\mathscr{D}^\varepsilon}\, \Pi_+^\varepsilon (V(t_n+s)e^{is\mathscr{D}^\varepsilon}\Pi_-^\varepsilon \Phi(t_n)) \\
&+ e^{\frac{-i2s}{\varepsilon^2}} e^{-is\mathscr{D}^\varepsilon}\, \Pi_-^\varepsilon (V(t_n+s)e^{-is\mathscr{D}^\varepsilon}\Pi_+^\varepsilon \Phi(t_n)), \qquad (3.3.31)\\
g_2^n(s,w) =& e^{i2w/\varepsilon^2} e^{is\mathscr{D}^\varepsilon}\Pi_+^\varepsilon \left(V(t_n)e^{-i(s-w)\mathscr{D}^\varepsilon}\Pi_+^\varepsilon \left(V(t_n)e^{iw\mathscr{D}^\varepsilon}\Pi_-^\varepsilon \Phi(t_n)\right)\right) \\
&+ e^{-i2w/\varepsilon^2} e^{-is\mathscr{D}^\varepsilon}\Pi_-^\varepsilon \left(V(t_n)e^{i(s-w)\mathscr{D}^\varepsilon}\Pi_-^\varepsilon \left(V(t_n)e^{-iw\mathscr{D}^\varepsilon}\Pi_+^\varepsilon \Phi(t_n)\right)\right), \qquad (3.3.32)\\
g_3^n(s,w) =& e^{\frac{i2(s-w)}{\varepsilon^2}} e^{is\mathscr{D}^\varepsilon}\Pi_+^\varepsilon \left(V(t_n)e^{i(s-w)\mathscr{D}^\varepsilon}\Pi_-^\varepsilon \left(V(t_n)e^{-iw\mathscr{D}^\varepsilon}\Pi_+^\varepsilon \Phi(t_n)\right)\right) \\
&+ e^{-\frac{i2(s-w)}{\varepsilon^2}} e^{-is\mathscr{D}^\varepsilon}\Pi_-^\varepsilon \left(V(t_n)e^{-i(s-w)\mathscr{D}^\varepsilon}\Pi_+^\varepsilon \left(V(t_n)e^{iw\mathscr{D}^\varepsilon}\Pi_-^\varepsilon \Phi(t_n)\right)\right), \qquad (3.3.33)\\
g_4^n(s,w) =& e^{i2s/\varepsilon^2} e^{is\mathscr{D}^\varepsilon}\Pi_+^\varepsilon \left(V(t_n)e^{i(s-w)\mathscr{D}^\varepsilon}\Pi_-^\varepsilon \left(V(t_n)e^{iw\mathscr{D}^\varepsilon}\Pi_-^\varepsilon \Phi(t_n)\right)\right) \\
&+ e^{-i2s/\varepsilon^2} e^{-is\mathscr{D}^\varepsilon}\Pi_-^\varepsilon \left(V(t_n)e^{-i(s-w)\mathscr{D}^\varepsilon}\Pi_+^\varepsilon \left(V(t_n)e^{-iw\mathscr{D}^\varepsilon}\Pi_+^\varepsilon \Phi(t_n)\right)\right). \qquad (3.3.34)
\end{aligned}$$

Proof. From the definition of $\mathbf{e}^n(x)$, noticing the Strang splitting formula (3.2.5), we have

$$\mathbf{e}^{n+1}(x) = e^{-\frac{i\tau}{2\varepsilon^2}\mathscr{T}^\varepsilon} e^{-i\int_{t_n}^{t_{n+1}} V(s,x)\,ds} e^{-\frac{i\tau}{2\varepsilon^2}\mathscr{T}^\varepsilon} \mathbf{e}^n(x) + \eta^n(x), \quad x \in \mathbb{R}, \tag{3.3.35}$$

where $\eta^n(x)$ is the local truncation error defined as

$$\eta^n(x) = \Phi(t_{n+1},x) - e^{-\frac{i\tau}{2\varepsilon^2}\mathscr{T}^\varepsilon} e^{-i\int_{t_n}^{t_{n+1}} V(s,x)\,ds} e^{-\frac{i\tau}{2\varepsilon^2}\mathscr{T}^\varepsilon} \Phi(t_n,x), \quad x \in \mathbb{R}. \tag{3.3.36}$$

Similar to the S_1 case, repeatedly using Duhamel's principle and Taylor expansion, we can obtain

$$\begin{aligned}
&\Phi(t_{n+1}) \\
&= e^{-\frac{i\tau}{\varepsilon^2}\mathscr{T}^\varepsilon}\Phi(t_n) - i\int_0^\tau e^{-\frac{i(\tau-s)}{\varepsilon^2}\mathscr{T}^\varepsilon}\left(V(t_n+s)e^{-\frac{is}{\varepsilon^2}\mathscr{T}^\varepsilon}\Phi(t_n)\right)ds \\
&\quad - \int_0^\tau\int_0^s e^{-\frac{i(\tau-s)}{\varepsilon^2}\mathscr{T}^\varepsilon}\left(V(t_n,x)e^{-\frac{i(s-w)}{\varepsilon^2}\mathscr{T}^\varepsilon}\left(V(t_n+w)\Phi(t_n+w)\right)\right)dwds, \qquad (3.3.37)
\end{aligned}$$

$$\begin{aligned}
&e^{-i\int_{t_n}^{t_{n+1}}V(s)\,ds}e^{-\frac{i\tau}{2\varepsilon^2}\mathscr{T}^\varepsilon}\Phi(t_n) \\
&= e^{-\frac{i\tau}{2\varepsilon^2}\mathscr{T}^\varepsilon}\left(1 - i\int_0^\tau V(t_n+s)\,ds - \frac{1}{2}\left(\int_0^\tau V(t_n+s)\,ds\right)^2\right)e^{-\frac{i\tau}{2\varepsilon^2}\mathscr{T}^\varepsilon}\Phi(t_n) \\
&\quad + e^{-\frac{i\tau}{2\varepsilon^2}\mathscr{T}^\varepsilon}\left(O(\tau^3)\right)e^{-\frac{i\tau}{2\varepsilon^2}\mathscr{T}^\varepsilon}\Phi(t_n). \qquad (3.3.38)
\end{aligned}$$

Denoting

$$f^n(s) = e^{\frac{is}{\varepsilon^2}\mathscr{T}^\varepsilon}\left(V(t_n+s,x)e^{-\frac{is}{\varepsilon^2}\mathscr{T}^\varepsilon}\Phi(t_n,x)\right), \qquad (3.3.39)$$

for $0 \le s \le \tau$, and

$$g^n(s,w) = e^{\frac{is}{\varepsilon^2}\mathscr{T}^\varepsilon}\left(V(t_n,x)e^{-\frac{i(s-w)}{\varepsilon^2}\mathscr{T}^\varepsilon}\left(V(t_n,x)e^{-\frac{iw}{\varepsilon^2}\mathscr{T}^\varepsilon}\Phi(t_n,x)\right)\right), \qquad (3.3.40)$$

for $0 \le s,w \le \tau$, in view of (3.3.37) and (3.3.38), $\eta^n(x)$ (3.3.36) can be written as

$$\begin{aligned}
\eta^n(x) = -e^{-\frac{i\tau}{\varepsilon^2}\mathscr{T}^\varepsilon}\Bigg[i\int_0^\tau f^n(s)\,ds - i\tau f^n\left(\frac{\tau}{2}\right) + \int_0^\tau\int_0^s g^n(s,w)\,dwds \\
- \frac{\tau^2}{2}g^n\left(\frac{\tau}{2},\frac{\tau}{2}\right)\Bigg] + \sum_{j=1}^2 R_j^n(x), \qquad (3.3.41)
\end{aligned}$$

where

$$\begin{aligned}
&R_1^n(x) = -e^{-\frac{i\tau}{2\varepsilon^2}\mathscr{T}^\varepsilon}(\lambda_1^n(x)+\lambda_2^n(x))e^{-\frac{i\tau}{2\varepsilon^2}\mathscr{T}^\varepsilon}\Phi(t_n,x), \\
&R_2^n(x) \\
&= -\int_0^\tau\int_0^s e^{-\frac{i(\tau-s)}{\varepsilon^2}\mathscr{T}^\varepsilon}\left(V(t_n+s,x)e^{-\frac{i(s-w)}{\varepsilon^2}\mathscr{T}^\varepsilon}\left(V(t_n+w,x)\lambda_3^n(w,x)\right)\right)dwds,
\end{aligned}$$

with

$$\begin{aligned}
\lambda_1^n(x) &= -i\left(\int_0^\tau V(t_n+s,x)\,ds - \tau V(t_n+\frac{\tau}{2},x)\right) - \frac{1}{2}\left(\int_0^\tau V(t_n+s,x)ds\right)^2 \\
&\quad + \frac{1}{2}\tau^2 V^2(t_n,x), \\
\lambda_2^n(x) &= e^{-i\int_0^\tau V(t_n+s,x)ds} - 1 + i\int_0^\tau V(t_n+s,x)ds + \frac{1}{2}\left(\int_0^\tau V(t_n+s,x)ds\right)^2, \\
\lambda_3^n(w,x) &= -i\int_0^w e^{-\frac{i(w-u)}{\varepsilon^2}\mathscr{T}^\varepsilon}\left(V(t_n+u,x)\Phi(t_n+u,x)\right)du.
\end{aligned}$$

It is easy to check that $\|\lambda_2^n(x)\|_{L^\infty} \lesssim \tau^3\|V(t,x)\|^3_{L^\infty(L^\infty)}$ and

$$\begin{aligned}
&\|\lambda_1^n(x)\|_{L^\infty} \lesssim \tau^3\|\partial_{tt}V(t,x)\|_{L^\infty(L^\infty)} + \tau^3\|\partial_t V(t,x)\|_{L^\infty(L^\infty)}\|V(t,x)\|_{L^\infty(L^\infty)}, \\
&\|\lambda_3^n(w,x)\|_{L^\infty([0,\tau];(L^2)^2)} \lesssim \tau\|V(t,x)\|_{L^\infty(L^\infty)}\|\Phi\|_{L^\infty((L^2)^2)},
\end{aligned}$$

which immediately implies that

$$\|R_1^n(x)\|_{L^2} \lesssim \left(\|\lambda_1^n(x)\|_{L^\infty} + \|\lambda_2^n(x)\|_{L^\infty}\right)\|\Phi(t_n)\|_{L^2} \lesssim \tau^3, \tag{3.3.42}$$

$$\|R_2^n(x)\|_{L^2} \lesssim \tau^2\|V(t,x)\|^2_{L^\infty(L^\infty)}\|\lambda_3^n(w,x)\|_{L^\infty([0,\tau];L^2)} \lesssim \tau^3. \tag{3.3.43}$$

In view of (3.3.11), recalling the definitions of $f_2^n(s)$ and $g_j^n(s,w)$ $(j=2,3,4)$ given in Lemma 3.2, we introduce $f_1^n(s)$ and $g_1^n(s,w)$ such that

$$f^n(s) = f_1^n(s) + f_2^n(s), \quad g^n(s,w) = \sum_{j=1}^4 g_j^n(s,w) \tag{3.3.44}$$

where

$$\begin{aligned}
f_1^n(s) &= e^{is\mathscr{D}^\varepsilon}\,\Pi_+^\varepsilon\left(V(t_n+s)e^{-is\mathscr{D}^\varepsilon}\Pi_+^\varepsilon\Phi(t_n)\right) \\
&\quad + e^{-is\mathscr{D}^\varepsilon}\,\Pi_-^\varepsilon\left(V(t_n+s)e^{is\mathscr{D}^\varepsilon}\Pi_-^\varepsilon\Phi(t_n)\right), \\
g_1^n(s,w) &= e^{is\mathscr{D}^\varepsilon}\Pi_+^\varepsilon\left(V(t_n)e^{-i(s-w)\mathscr{D}^\varepsilon}\Pi_+^\varepsilon\left(V(t_n)e^{-iw\mathscr{D}^\varepsilon}\Pi_+^\varepsilon\Phi(t_n)\right)\right) \\
&\quad + e^{-is\mathscr{D}^\varepsilon}\Pi_-^\varepsilon\left(V(t_n)e^{i(s-w)\mathscr{D}^\varepsilon}\Pi_-^\varepsilon\left(V(t_n)e^{iw\mathscr{D}^\varepsilon}\Pi_-^\varepsilon\Phi(t_n)\right)\right).
\end{aligned}$$

Denote

$$\begin{aligned}
\zeta_1^n(x) &= -ie^{-\frac{i\tau}{\varepsilon^2}\mathscr{T}^\varepsilon}\left(\int_0^\tau f_1^n(s)\,ds - \tau f_1^n(\tau/2)\right), \\
\zeta_2^n(x) &= -e^{-\frac{i\tau}{\varepsilon^2}\mathscr{T}^\varepsilon}\left(\int_0^\tau\int_0^s g_1^n(s,w)\,dwds - \frac{\tau^2}{2}g_1^n(\tau/2,\tau/2)\right),
\end{aligned}$$

then it is easy to show that for $J = [0,\tau]^2$,

$$\|\zeta_1^n(x)\|_{L^2} \lesssim \tau^3 \|\partial_{ss} f_1(s)\|_{L^\infty([0,\tau];(L^2)^2)} \lesssim \tau^3, \tag{3.3.45}$$

$$\|\zeta_2^n(x)\|_{L^2} \lesssim \tau^3 (\|\partial_s g_1(s,w)\|_{L^\infty(J;(L^2)^2)} + \|\partial_w g_1(s,w)\|_{L^2(J;(L^2)^2)}) \lesssim \tau^3, \tag{3.3.46}$$

by noticing that $V \in L^\infty(W^{2m,\infty})$ and $\Phi(t,x) \in L^\infty((H^{2m})^2)$ with $m = 2$ as well as the fact that $\mathscr{D}^\varepsilon : (H^l)^2 \to (H^{l-2})^2$ ($l \geq 2$) is uniformly bounded w.r.t. ε. Recalling (3.3.39), (3.3.40), (3.3.41), (3.3.44) and η_j^n ($j = 2,3$) (3.3.29)-(3.3.30) given in Lemma 3.2 , we have

$$\eta^n(x) = \eta_1^n(x) + \eta_2^n(x) + \eta_3^n(x), \tag{3.3.47}$$

where $\eta_2^n(x)$ and $\eta_3^n(x)$ are given in Lemma 3.2, and

$$\eta_1^n(x) = R_1^n(x) + R_2^n(x) + \zeta_1^n(x) + \zeta_2^n(x).$$

Combining (3.3.42), (3.3.43), (3.3.45) and (3.3.46), we can get

$$\|\eta_1^n(x)\|_{L^2} \leq \|R_1^n(x)\|_{L^2} + \|R_2^n(x)\|_{L^2} + \|\zeta_1^n(x)\|_{L^2} + \|\zeta_2^n(x)\|_{L^2} \lesssim \tau^3, \tag{3.3.48}$$

which completes the proof. □

Utilizing these lemmas, we now proceed to prove Theorem 3.1 and Theorem 3.2.

Proof of Theorem 3.1

Proof. From Lemma 3.1, it is straightforward that

$$\|\mathbf{e}^{n+1}(x)\|_{L^2} \leq \|\mathbf{e}^n(x)\|_{L^2} + \|\eta_1^n(x)\|_{L^2} + \|\eta_2^n(x)\|_{L^2}, \quad 0 \leq n \leq \frac{T}{\tau} - 1, \tag{3.3.49}$$

with $\mathbf{e}^0(x) = 0$, $\|\eta_1^n(x)\|_{L^2} \lesssim \tau^2$ and $\eta_2^n(x) = -ie^{-i\tau\mathscr{T}^\varepsilon/\varepsilon^2}\left(\int_0^\tau f_2^n(s)ds - \tau f_2^n(0)\right)$, where $f_2^n(s)$ is defined in (3.3.13).

To analyze $f_2^n(s)$, using (3.3.8) and (3.3.9), we expand $\Pi_+^\varepsilon V(t_n)\Pi_-^\varepsilon$ and $\Pi_-^\varepsilon V(t_n)\Pi_+^\varepsilon$ to get

$$\begin{aligned}
\Pi_+^\varepsilon V(t_n)\Pi_-^\varepsilon =& -\varepsilon\Pi_+^0 V(t_n)\mathscr{R}_1 + \varepsilon\mathscr{R}_1 V(t_n)\Pi_-^\varepsilon,\\
\Pi_-^\varepsilon V(t_n)\Pi_+^\varepsilon =& \varepsilon\Pi_-^0 V(t_n)\mathscr{R}_1 - \varepsilon\mathscr{R}_1 V(t_n)\Pi_+^\varepsilon.
\end{aligned}$$

As $\mathscr{R}_1 : (H^m)^2 \to (H^{m-1})^2$ is uniformly bounded with respect to $\varepsilon \in (0,1]$, we have

$$\left\| \Pi_+^\varepsilon \left(V(t_n) \Pi_-^\varepsilon e^{is\mathscr{D}^\varepsilon} \Phi(t_n) \right) \right\|_{L^2} \lesssim \varepsilon \|V(t_n)\|_{W^{1,\infty}} \|\Phi(t_n)\|_{H^1}, \tag{3.3.50}$$

$$\left\| \Pi_-^\varepsilon \left(V(t_n) \Pi_+^\varepsilon e^{is\mathscr{D}^\varepsilon} \Phi(t_n) \right) \right\|_{L^2} \lesssim \varepsilon \|V(t_n)\|_{W^{1,\infty}} \|\Phi(t_n)\|_{H^1}. \tag{3.3.51}$$

Noticing the assumptions (A) and (B) with $m = 1$ and $m_* = 0$, we obtain from (3.3.13) ($0 \le s \le \tau$)

$$\|f_2^n(s)\|_{L^\infty([0,\tau];(L^2)^2)} \lesssim \varepsilon, \quad \|\partial_s(f_2^n)(\cdot)\|_{L^\infty([0,\tau];(L^2)^2)} \lesssim \varepsilon/\varepsilon^2 = 1/\varepsilon. \tag{3.3.52}$$

As a result, from the first inequality, we get

$$\left\| \int_0^\tau f_2^n(s)\, ds - \tau f_2^n(0) \right\|_{L^2} \lesssim \tau\varepsilon. \tag{3.3.53}$$

On the other hand, noticing Taylor expansion and the second inequality in (3.3.52), we have

$$\left\| \int_0^\tau f_2^n(s)\, ds - \tau f_2^n(0) \right\|_{L^2} \le \frac{\tau^2}{2} \|\partial_s f_2^n(\cdot)\|_{L^\infty([0,\tau];(L^2)^2)} \lesssim \tau^2/\varepsilon. \tag{3.3.54}$$

Combining (3.3.53) and (3.3.54), we arrive at

$$\|\eta_2^n(x)\|_{L^2} \lesssim \min\{\tau\varepsilon, \tau^2/\varepsilon\}. \tag{3.3.55}$$

Then from (3.3.49) and $\mathbf{e}^0 = 0$, we get

$$\begin{aligned} \|\mathbf{e}^{n+1}(x)\|_{L^2} \le & \|\mathbf{e}^0(x)\|_{L^2} + \sum_{k=0}^n \|\eta_1^k(x)\|_{L^2} + \sum_{k=0}^n \|\eta_2^k(x)\|_{L^2} \\ \lesssim & n\tau^2 + n\min\{\tau\varepsilon, \tau^2/\varepsilon\} \lesssim \tau + \min\{\varepsilon, \tau/\varepsilon\}, \quad 0 \le n \le \frac{T}{\tau} - 1, \end{aligned}$$

which gives the desired results. □

Proof of Theorem 3.2

Proof. From Lemma 3.2, it is easy to get that

$$\|\mathbf{e}^{n+1}(x)\|_{L^2} \le \|\mathbf{e}^n(x)\|_{L^2} + \|\eta_1^n(x)\|_{L^2} + \|\eta_2^n(x)\|_{L^2} + \|\eta_3^n(x)\|_{L^2}, \tag{3.3.56}$$

with $\mathbf{e}^0(x) = 0$ and $\|\eta_1^n(x)\|_{L^2} \lesssim \tau^3$.

Through similar computations in the S_1 case, under the hypothesis of Theorem 3.2, we can show that for $0 \le s, w, \le \tau$,

$$\|f_2^n(s)\|_{L^2} \lesssim \varepsilon, \quad \|\partial_s f_2^n(s)\|_{L^2} \lesssim \varepsilon/\varepsilon^2 = 1/\varepsilon, \quad \|\partial_{ss} f_2^n(s)\|_{L^2} \lesssim 1/\varepsilon^3;$$
$$\|g_j^n(s,w)\|_{L^2} \lesssim \varepsilon, \quad \|\partial_s g_j^n(s,w)\|_{L^2} \lesssim 1/\varepsilon, \quad \|\partial_w g_j^n(s,w)\|_{L^2} \lesssim 1/\varepsilon, \quad j = 2,3,4.$$

As a result, for $j = 2,3,4$, we have

$$\left\| \int_0^\tau f_2^n(s)\, ds - \tau f_2^n(\frac{\tau}{2}) \right\|_{L^2} \lesssim \tau\varepsilon, \ \left\| \int_0^\tau \int_0^s g_j^n(s,w)\, dwds - \frac{\tau^2}{2} g_j^n(\frac{\tau}{2}, \frac{\tau}{2}) \right\|_{L^2} \lesssim \tau^2\varepsilon.$$

On the other hand, for $j = 2,3,4$,Taylor expansion will lead to

$$\left\| \int_0^\tau f_2^n(s)\, ds - \tau f_2^n(\frac{\tau}{2}) \right\|_{L^2} \lesssim \frac{\tau^3}{\varepsilon^3}, \ \left\| \int_0^\tau \int_0^s g_j^n(s,w)\, dwds - \frac{\tau^2}{2} g_j^n(\frac{\tau}{2}, \frac{\tau}{2}) \right\|_{L^2} \lesssim \frac{\tau^3}{\varepsilon}.$$

The two estimates above together with (3.3.29) and (3.3.30) imply

$$\|\eta_2^n(x)\|_{L^2} + \|\eta_3^n(x)\|_{L^2} \lesssim \min\{\tau\varepsilon, \tau^3/\varepsilon^3\}. \tag{3.3.57}$$

Recalling (3.3.56), we can get

$$\begin{aligned}\|\mathbf{e}^{n+1}(x)\|_{L^2} \le & \|\mathbf{e}^0(x)\|_{L^2} + \sum_{k=0}^n \|\eta_1^k(x)\|_{L^2} + \sum_{k=0}^n \|\eta_2^k(x)\|_{L^2} + \sum_{k=0}^n \|\eta_3^k(x)\|_{L^2} \\ \lesssim & n\tau^3 + n\min\{\tau\varepsilon, \tau^3/\varepsilon^3\} \lesssim \tau^2 + \min\{\varepsilon, \tau^2/\varepsilon^3\}, \quad 0 \le n \le \frac{T}{\tau} - 1,\end{aligned}$$

which gives the desired result. □

3.4 Improved uniform error bounds for non-resonant time steps

In the Dirac equation (3.1.2) or (3.1.1), the leading term is $\frac{1}{\varepsilon^2}\sigma_3\Phi$ or $\frac{1}{\varepsilon^2}\beta\Psi$, which suggests that the solution propagates almost periodically in time with periods $2k\pi\varepsilon^2$ ($k \in \mathbb{N}^*$, which are the periods of $e^{-i\sigma_3/\varepsilon^2}$ and $e^{-i\beta/\varepsilon^2}$). From numerical results, we find out that the errors perform much better than the error bounds in Theorem 3.1& Theorem 3.2, when 2τ

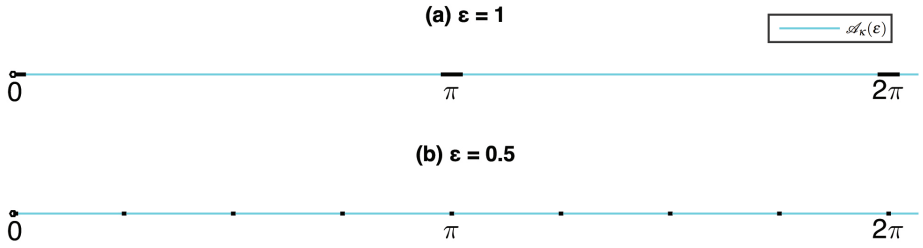

Figure 3.4.1: Illustration of non-resonant time steps $\mathscr{A}_\kappa(\varepsilon)$ with $\kappa = 0.15$ for (a) $\varepsilon = 1$ and (b) $\varepsilon = 0.5$.

is away from the leading temporal oscillation periods $2k\pi\varepsilon^2$. In fact, for given $0 < \kappa \leq 1$, define

$$\mathscr{A}_\kappa(\varepsilon) := \bigcup_{k=0}^{\infty} \left[\varepsilon^2 k\pi + \varepsilon^2 \arcsin\kappa, \varepsilon^2(k+1)\pi - \varepsilon^2 \arcsin\kappa\right], \quad 0 < \varepsilon \leq 1, \tag{3.4.1}$$

and the errors of S_1 and S_2 can be improved compared to the previous section when $\tau \in \mathscr{A}_\kappa(\varepsilon)$. To illustrate $\mathscr{A}_\kappa(\varepsilon)$, we show in Figure 3.4.1 for $\varepsilon = 1$ and $\varepsilon = 0.5$ with fixed $\kappa = 0.15$.

For $\tau \in \mathscr{A}_\kappa(\varepsilon)$, we can derive improved uniform error bounds for S_1 and S_2 as shown in Theorem 3.3 and Theorem 3.4.

Theorem 3.3. *Let $\Phi^n(x)$ be the numerical approximation obtained from S_1 (3.2.4). If the time step size τ is non-resonant, i.e. there exists $0 < \kappa \leq 1$, such that $\tau \in \mathscr{A}_\kappa(\varepsilon)$, under the assumptions (A) and (B) with $m = 1$ and $m_* = 1$, we have an improved uniform error bound*

$$\|\mathbf{e}^n(x)\|_{L^2} \lesssim_\kappa \tau, \quad 0 \leq n \leq \frac{T}{\tau}. \tag{3.4.2}$$

Proof. We divide the proof into three steps.

Step 1 (Explicit representation of the error). From Lemma 3.1, we have

$$\mathbf{e}^{n+1}(x) = e^{-\frac{i\tau}{\varepsilon^2}\mathscr{T}^\varepsilon} e^{-i\int_{t_n}^{t_{n+1}} V(s)ds}\mathbf{e}^n(x) + \eta_1^n(x) + \eta_2^n(x), \quad 0 \leq n \leq \frac{T}{\tau} - 1, \tag{3.4.3}$$

with $\|\eta_1^n(x)\|_{L^2} \lesssim \tau^2$, $\mathbf{e}^0 = 0$ and

$$\eta_2^n(x) = -ie^{-i\tau\mathscr{T}^\varepsilon/\varepsilon^2}\left(\int_0^\tau f_2^n(s)ds - \tau f_2^n(0)\right). \tag{3.4.4}$$

where

$$f_2^n(s) = e^{i2s/\varepsilon^2} e^{is\mathscr{D}^\varepsilon} \Pi_+^\varepsilon \left(V(t_n) \Pi_-^\varepsilon e^{is\mathscr{D}^\varepsilon} \Phi(t_n) \right) + e^{-i2s/\varepsilon^2} e^{-is\mathscr{D}^\varepsilon} \Pi_-^\varepsilon \left(V(t_n) \Pi_+^\varepsilon e^{-is\mathscr{D}^\varepsilon} \Phi(t_n) \right).$$

Denote the numerical solution propagator $S_{n,\tau} := e^{-\frac{i\tau}{\varepsilon^2}\mathscr{T}^\varepsilon} e^{-i\int_{t_n}^{t_{n+1}} V(s,x)ds}$ for $n \geq 0$, then $\forall \widetilde{\Phi} \in \mathbb{C}^2$, for $m \geq 1$,

$$\left\| S_{n,\tau} \widetilde{\Phi} \right\|_{L^2} = \left\| \widetilde{\Phi} \right\|_{L^2}, \ \left\| S_{n,\tau} \widetilde{\Phi} \right\|_{H^m} \leq e^{C\tau \|V(t,x)\|_{L^\infty([0,T];W^{m,\infty})}} \|\widetilde{\Phi}\|_{H^m}, \tag{3.4.5}$$

with some generic constant $C > 0$ and

$$\begin{aligned} \mathbf{e}^{n+1}(x) &= S_{n,\tau} \mathbf{e}^n(x) + (\eta_1^n(x) + \eta_2^n(x)) \\ &= S_{n,\tau}(S_{n-1,\tau} \mathbf{e}^{n-1}(x)) + S_{n,\tau} \left(\eta_1^{n-1}(x) + \eta_2^{n-1}(x) \right) + (\eta_1^n(x) + \eta_2^n(x)) \\ &= ... \\ &= S_{n,\tau} S_{n-1,\tau} ... S_{0,\tau} \mathbf{e}^0(x) + \sum_{k=0}^{n} S_{n,\tau} ... S_{k+2,\tau} S_{k+1,\tau} \left(\eta_1^k(x) + \eta_2^k(x) \right), \end{aligned} \tag{3.4.6}$$

where for $k = n$, we take $S_{n,\tau} ... S_{k+2,\tau} S_{k+1,\tau} = Id$. Since $S_{n,\tau}$ preserves the L^2 norm, noticing $\|\eta_1^k(x)\|_{L^2} \lesssim \tau^2$, $k = 0, 1, ..., n$, we have

$$\left\| \sum_{k=0}^{n} S_{n,\tau} ... S_{k+1,\tau} \eta_1^k(x) \right\|_{L^2} \lesssim \sum_{k=0}^{n} \tau^2 \lesssim \tau,$$

which leads to

$$\|\mathbf{e}^{n+1}(x)\|_{L^2} \lesssim \tau + \left\| \sum_{k=0}^{n} S_{n,\tau} ... S_{k+1,\tau} \eta_2^k(x) \right\|_{L^2}. \tag{3.4.7}$$

The improved estimates rely on the refined analysis of the terms involving η_2^k in (3.4.7). To this aim, we introduce the following approximation of η_2^k to focus on the most relevant terms,

$$\tilde{\eta}_2^k(x) = \int_0^\tau \tilde{f}_2^k(s) ds - \tau \tilde{f}_2^k(0), \quad k = 0, 1, ..., n, \tag{3.4.8}$$

with

$$\tilde{f}_2^k(s) = -ie^{i(2s-\tau)/\varepsilon^2} \Pi_+^\varepsilon \left(V(t_k) \Pi_-^\varepsilon \Phi(t_k) \right) - ie^{i(\tau-2s)/\varepsilon^2} \Pi_-^\varepsilon \left(V(t_k) \Pi_+^\varepsilon \Phi(t_k) \right), \tag{3.4.9}$$

then it is easy to verify that (using Taylor expansion $e^{i\tau\mathscr{D}^\varepsilon} = Id + O(\tau\mathscr{D}^\varepsilon)$)

$$\|\eta_2^k(x) - \tilde{\eta}_2^k(x)\|_{L^2} \lesssim \tau^2 \|V(t_k)\|_{H^2} \|\Phi(t_k)\|_{H^2} \lesssim \tau^2. \tag{3.4.10}$$

As a result, from (3.4.7), we have

$$\begin{aligned}
\|\mathbf{e}^{n+1}(x)\|_{L^2} \lesssim & \tau + \left\| \sum_{k=0}^{n} S_{n,\tau} \ldots S_{k+1,\tau} (\eta_2^k(x) - \tilde{\eta}_2^k(x)) \right\|_{L^2} + \left\| \sum_{k=0}^{n} S_{n,\tau} \ldots S_{k+1,\tau} \tilde{\eta}_2^k(x) \right\|_{L^2} \\
\leq & \tau + \sum_{k=0}^{n} \|\eta_2^k(x) - \tilde{\eta}_2^k(x)\|_{L^2} + \left\| \sum_{k=0}^{n} S_{n,\tau} \ldots S_{k+1,\tau} \tilde{\eta}_2^k(x) \right\|_{L^2} \\
\lesssim & \tau + \left\| \sum_{k=0}^{n} S_{n,\tau} \ldots S_{k+1,\tau} \tilde{\eta}_2^k(x) \right\|_{L^2}.
\end{aligned}$$

Step 2 (Representation of the error using the exact solution flow). Denote $S_e(t;t_k)$ ($k = 0,1,\ldots,n$) to be the exact solution operator of the Dirac equation, acting on some $\tilde{\Phi}(x) = (\tilde{\phi}_1(x), \tilde{\phi}_2(x))^T \in \mathbb{C}^2$ so that $S_e(t;t_k)\tilde{\Phi}(x)$ is the exact solution $\Psi(t,x)$ at time t of

$$\begin{cases} i\partial_t \Psi(t,x) = \dfrac{\mathscr{T}^\varepsilon}{\varepsilon^2}\Psi(t,x) + V(t,x)\Psi(t,x), \\ \Psi(t_k,x) = \tilde{\Phi}(x). \end{cases} \tag{3.4.11}$$

and the following properties hold true for $t \geq t_k$, $m \geq 1$ and some generic constant $C > 0$

$$\left\| S_e(t;t_k)\widetilde{\Phi} \right\|_{L^2} = \left\| \widetilde{\Phi} \right\|_{L^2}, \ \left\| S_e(t;t_k)\widetilde{\Phi} \right\|_{H^m} \leq e^{C(t-t_k)\|V(t,x)\|_{L^\infty([0,T];W^{m,\infty})}} \|\widetilde{\Phi}\|_{H^m}. \tag{3.4.12}$$

It is convenient to write $\tilde{\eta}_2^k(x)$ (3.4.8) as

$$\tilde{\eta}_2^k(x) = p_+(\tau)\Pi_+^\varepsilon \left(V(t_k)\Pi_-^\varepsilon \Phi(t_k) \right) + p_-(\tau)\Pi_-^\varepsilon \left(V(t_k)\Pi_+^\varepsilon \Phi(t_k) \right), \tag{3.4.13}$$

with $p_\pm(\tau) = -ie^{\mp i\tau/\varepsilon^2} \left(\int_0^\tau e^{\pm i2s/\varepsilon^2} ds - \tau \right)$ and by the inequality $\left| \int_0^\tau e^{i2s/\varepsilon^2} ds - \tau \right| + \left| \int_0^\tau e^{-i2s/\varepsilon^2} ds - \tau \right| \leq 4\tau$ and similar computations in (3.3.50)-(3.3.51), it follows that

$$\|\tilde{\eta}_2^k\|_{H^2} \lesssim \tau\varepsilon \|V(t_k)\|_{W^{3,\infty}} \|\Phi(t_k)\|_{H^3} \lesssim \varepsilon\tau. \tag{3.4.14}$$

Recalling the error bounds in Theorem 3.1 and Remark 3.1, we have

$$\begin{aligned}
\|(S_{n,\tau} \ldots S_{k+1,\tau} - S_e(t_{n+1};t_{k+1}))\tilde{\eta}_2^k(x)\|_{L^2} &\lesssim \left(\tau + \frac{\tau}{\varepsilon} \right) \|\tilde{\eta}_2^k\|_{H^2} \\
&\lesssim \left(\tau + \frac{\tau}{\varepsilon} \right) \varepsilon\tau \lesssim \tau^2,
\end{aligned} \tag{3.4.15}$$

and

$$\begin{aligned}\|\mathbf{e}^{n+1}(x)\|_{L^2} \lesssim \tau + \sum_{k=0}^{n} \left\| (S_{n,\tau}...S_{k+1,\tau} - S_e(t_{n+1};t_{k+1}))\tilde{\eta}_2^k(x) \right\|_{L^2} \\ + \left\| \sum_{k=0}^{n} S_e(t_{n+1};t_{k+1})\tilde{\eta}_2^k(x) \right\|_{L^2} \\ \lesssim \tau + \left\| \sum_{k=0}^{n} S_e(t_{n+1};t_{k+1})\tilde{\eta}_2^k(x) \right\|_{L^2}. \end{aligned} \tag{3.4.16}$$

Noticing (3.4.13), we have

$$\begin{aligned} S_e(t_{n+1};t_{k+1})\tilde{\eta}_2^k(x) = p_+(\tau)S_e(t_{n+1};t_{k+1})\Pi_+^\varepsilon V(t_k)\Pi_-^\varepsilon S_e(t_k;t_0)\Phi(0) \\ + p_-(\tau)S_e(t_{n+1};t_{k+1})\Pi_-^\varepsilon V(t_k)\Pi_+^\varepsilon S_e(t_k;t_0)\Phi(0), \end{aligned} \tag{3.4.17}$$

and it remains to estimate S_e part in (3.4.16).

Step 3 (Improved error bounds for non-resonant time steps). From [14], we know that the exact solution of Dirac equation is structured as follows

$$S_e(t_n;t_k)\tilde{\Phi}(x) = e^{-i(t_n-t_k)/\varepsilon^2}\Psi_+(t,x) + e^{i(t_n-t_k)/\varepsilon^2}\Psi_-(t,x) + R_k^n\tilde{\Phi}(x), \tag{3.4.18}$$

where $R_k^n : (L^2)^2 \to (L^2)^2$ is the residue operator and $\|R_k^n\tilde{\Phi}(x)\|_{L^2} \lesssim \varepsilon^2\|\tilde{\Phi}(x)\|_{H^2}$ $(0 \le k \le n)$, and

$$\begin{cases} i\partial_t\Psi_\pm(t,x) = \pm\mathscr{D}^\varepsilon\Psi_\pm(t,x) + \Pi_\pm^\varepsilon(V(t)\Psi_\pm(t,x)), \\ \Psi_\pm(t_k,x) = \Pi_\pm^\varepsilon\tilde{\Phi}(x). \end{cases} \tag{3.4.19}$$

Denote $S_e^+(t;t_k)\tilde{\Phi}(x) = \Psi_+(t,x)$, $S_e^-(t;t_k)\tilde{\Phi}(x) = \Psi_-(t,x)$ to be the solution propagator of the above equation for $\Psi_+(t,x)$, $\Psi_-(t,x)$, respectively, and $S_e^\pm$ share the same properties in

(3.4.12). Plugging (3.4.18) into (3.4.17), we derive

$$
\begin{aligned}
&\sum_{k=0}^{n} S_e(t_{n+1};t_{k+1})\tilde{\eta}_2^k(x)\\
=&\sum_{k=0}^{n}\sum_{\sigma=\pm}\left(e^{-i\frac{t_{n+1}-t_{k+1}}{\varepsilon^2}}S_e^+(t_{n+1};t_{k+1})+e^{i\frac{t_{n+1}-t_{k+1}}{\varepsilon^2}}S_e^-(t_{n+1};t_{k+1})+R_{k+1}^{n+1}\right)\\
&\quad\Pi_\sigma^\varepsilon V(t_k)\Pi_{\sigma^*}^\varepsilon\left(e^{-i\frac{t_k-t_0}{\varepsilon^2}}S_e^+(t_k;t_0)+e^{i\frac{t_k-t_0}{\varepsilon^2}}S_e^-(t_k;t_0)+R_0^k\right)\Phi(0)p_\sigma(\tau)\\
=&\underbrace{\sum_{k=0}^{n}\sum_{\sigma=\pm}e^{-i\sigma\frac{t_{n+1}-t_{k+1}}{\varepsilon^2}}S_e^\sigma(t_{n+1};t_{k+1})\Pi_\sigma^\varepsilon V(t_k)\Pi_{\sigma^*}^\varepsilon e^{i\sigma\frac{t_k-t_0}{\varepsilon^2}}S_e^{\sigma^*}(t_k;t_0)\Phi(0)p_\sigma(\tau)}_{I_1^n(x)}\\
&+\underbrace{\sum_{k=0}^{n}\sum_{\sigma=\pm}\left(R_{k+1}^{n+1}\Pi_\sigma^\varepsilon V(t_k)\Pi_{\sigma*}^\varepsilon\Phi(t_k)+S_e(t_{n+1};t_{k+1})\Pi_\sigma^\varepsilon V(t_k)\Pi_{\sigma*}^\varepsilon R_0^k\Phi(0)\right)p_\sigma(\tau)}_{I_2^n(x)}\\
=&I_1^n(x)+I_2^n(x),
\end{aligned}
$$

where $\sigma^*=+$ if $\sigma=-$ and $\sigma^*=-$ if $\sigma=+$. As $|p_\pm(\tau)|=\left|\int_0^\tau e^{\pm 2is/\varepsilon^2}ds-\tau\right|\lesssim\tau^2/\varepsilon^2$ by Taylor expansion, we have

$$
\|I_2^n(x)\|_{L^2}\lesssim\frac{\tau^2}{\varepsilon^2}\sum_{k=0}^{n}\left(\varepsilon^2\|V(t_k)\|_{W^{2,\infty}}\|\Phi(t_k)\|_{H^2}+\varepsilon^2\|V(t_k)\|_{L^\infty}\|\Phi(t_0)\|_{H^2}\right)\lesssim\tau^2.
$$

We can rewrite $I_1^n(x)$ as

$$
\begin{aligned}
I_1^n(x)=&\sum_{k=0}^{n}\sum_{\sigma=\pm}e^{-i\sigma\frac{t_{n+1}-2t_k-\tau}{\varepsilon^2}}S_e^\sigma(t_{n+1};t_{k+1})\Pi_\sigma^\varepsilon V(t_k)\Pi_{\sigma^*}^\varepsilon S_e^{\sigma^*}(t_k;t_0)\Phi(0)p_\sigma(\tau),\\
=&p_+(\tau)\sum_{k=0}^{n}(\theta_k-\theta_{k-1})S_e^+(t_{n+1};t_{k+1})\Pi_+^\varepsilon V(t_k)\Pi_-^\varepsilon S_e^-(t_k;t_0)\Phi(0)\\
&+p_-(\tau)\sum_{k=0}^{n}\overline{(\theta_k-\theta_{k-1})}S_e^-(t_{n+1};t_{k+1})\Pi_-^\varepsilon V(t_k)\Pi_+^\varepsilon S_e^+(t_k;t_0)\Phi(0)\\
=&\gamma_1^n(x)+\gamma_2^n(x),
\end{aligned}
$$

where $\bar{\theta}$ is the complex conjugate of θ and for $\quad 0 \le k \le n$,

$$\theta_k = \sum_{l=0}^{k} e^{-i(t_{n+1}-2t_l-\tau)/\varepsilon^2} = \frac{e^{-in\tau/\varepsilon^2} - e^{-i(n-2k-2)\tau/\varepsilon^2}}{1-e^{2i\tau/\varepsilon^2}}, \quad \theta_{-1} = 0, \tag{3.4.20}$$

$$\gamma_1^n(x) = p_+(\tau) \sum_{k=0}^{n} (\theta_k - \theta_{k-1}) S_e^+(t_{n+1};t_{k+1}) \Pi_+^\varepsilon V(t_k) \Pi_-^\varepsilon S_e^-(t_k;t_0)\Phi(0), \tag{3.4.21}$$

$$\gamma_2^n(x) = p_-(\tau) \sum_{k=0}^{n} \overline{(\theta_k - \theta_{k-1})} S_e^-(t_{n+1};t_{k+1}) \Pi_-^\varepsilon V(t_k) \Pi_+^\varepsilon S_e^+(t_k;t_0)\Phi(0). \tag{3.4.22}$$

It is easy to check that if $\tau \in \mathscr{A}_\kappa(\varepsilon)$, it satisfies $|1-e^{2i\tau/\varepsilon^2}| = 2|\sin(\tau/\varepsilon^2)| \ge 2\kappa > 0$, then we have

$$|\theta_k| \le \frac{1}{\kappa}, \quad k = 0,1,...,n.$$

As a result, noticing $|p_\pm(\tau)| \le 2\tau$, we can get

$$\begin{aligned}
&\|\gamma_1^n(x)\|_{L^2} \\
&\le 2\tau \Bigg\| \sum_{k=0}^{n-1} \theta_k \big[S_e^+(t_{n+1};t_{k+1}) \Pi_+^\varepsilon V(t_k) \Pi_-^\varepsilon S_e^-(t_k;t_0) - S_e^+(t_{n+1};t_{k+2}) \Pi_+^\varepsilon V(t_{k+1}) \\
&\qquad \Pi_-^\varepsilon S_e^-(t_{k+1};t_0) \big] \Phi(0) \Bigg\|_{L^2} + \tau \|\theta_n S_e^+(t_{n+1};t_{n+1}) \Pi_+^\varepsilon V(t_n) \Pi_-^\varepsilon S_e^-(t_n;t_0)\Phi(0)\|_{L^2} \\
&\lesssim \tau \sum_{k=0}^{n-1} \tau/\kappa + \tau/\kappa \lesssim_\kappa \tau,
\end{aligned}$$

where we have used the triangle inequality and properties of the solution flows $S_e^\pm$ to deduce that (omitted for brevity as they are standard)

$$\begin{aligned}
&\Big\| \big[S_e^+(t_{n+1};t_{k+1}) \Pi_+^\varepsilon V(t_k) \Pi_-^\varepsilon S_e^-(t_k;t_0) - S_e^+(t_{n+1};t_{k+2}) \Pi_+^\varepsilon V(t_{k+1}) \Pi_-^\varepsilon S_e^-(t_{k+1};t_0) \big] \Phi(0) \Big\|_{L^2} \\
&\le \big\| S_e^+(t_{n+1};t_{k+1}) \Pi_+^\varepsilon \left((V(t_k) - V(t_{k+1})) \Pi_-^\varepsilon S_e^-(t_k;t_0) \right) \Phi(0) \big\|_{L^2} \\
&\quad + \big\| S_e^+(t_{n+1};t_{k+1}) \Pi_+^\varepsilon \left(V(t_{k+1}) \Pi_-^\varepsilon (S_e^-(t_k;t_0) - S_e^-(t_{k+1};t_0)) \right) \Phi(0) \big\|_{L^2} \\
&\quad + \big\| \left(S_e^+(t_{n+1};t_{k+1}) - S_e^+(t_{n+1};t_{k+2}) \right) \Pi_+^\varepsilon V(t_{k+1}) \Pi_-^\varepsilon S_e^-(t_{k+1};t_0) \Phi(0) \big\|_{L^2} \\
&\lesssim \tau \|\partial_t V\|_{L^\infty(L^\infty)} \|\Phi(0)\|_{L^2} + \tau \big\| \partial_t S_e^-(t;t_0)\Phi(0) \big\|_{L^\infty([0,T];(L^2)^2)} \\
&\qquad + \tau \big\| \partial_t \left(S_e^+(t_{n+1};t) \Pi_+^\varepsilon V(t_{k+1}) \Pi_-^\varepsilon S_e^-(t_{k+1};t_0)\Phi(0) \right) \big\|_{L^\infty([t_{k+1},t_{n+1}];(L^2)^2)} \\
&\lesssim \tau + \tau \|\Phi(0)\|_{H^2} + \tau \|V(t_{k+1})\|_{W^{2,\infty}} \|\Phi(0)\|_{H^2} \lesssim \tau.
\end{aligned}$$

Similarly, we could get $\|\gamma_2^n(x)\|_{L^2} \lesssim_\kappa \tau$ and hence $\|I_1^n(x)\|_{L^2} \lesssim_\kappa \tau$. In summary, we have

$$\|\mathbf{e}^{n+1}(x)\|_{L^2} \lesssim \tau + \|I_1^n(x)\|_{L^2} + \|I_2^n(x)\|_{L^2} \lesssim_\kappa \tau,$$

which gives the desired results. □

Theorem 3.4. *Let $\Phi^n(x)$ be the numerical approximation obtained from S_2 (3.2.5). If the time step size τ is non-resonant, i.e. there exists $0<\kappa\leq 1$, such that $\tau\in\mathscr{A}_\kappa(\varepsilon)$, under the assumptions (A) and (B) with $m=2$ and $m_*=1$, we assume an extra regulariry $V(t,x)\in W^{1,\infty}([0,T];H^3(\mathbb{R}))$ and then the following two error estimates hold*

$$\|\mathbf{e}^n(x)\|_{L^2}\lesssim_\kappa \tau^2+\tau\varepsilon,\quad \|\mathbf{e}^n(x)\|_{L^2}\lesssim_\kappa \tau^2+\tau^2/\varepsilon,\quad 0\leq n\leq\frac{T}{\tau}. \tag{3.4.23}$$

As a result, there is an improved uniform error bound for S_2

$$\|\mathbf{e}^n(x)\|_{L^2}\lesssim_\kappa \tau^2+\max_{0<\varepsilon\leq 1}\min\{\tau\varepsilon,\tau^2/\varepsilon\}\lesssim_\kappa \tau^{3/2},\quad 0\leq n\leq\frac{T}{\tau}. \tag{3.4.24}$$

Proof. We divide the proof into two steps.

Step 1 (Representation of the error using the exact solution flow). From Lemma 3.2, we have for $0\leq n\leq\frac{T}{\tau}-1$,

$$\mathbf{e}^{n+1}(x)=e^{-\frac{i\tau}{2\varepsilon^2}\mathscr{T}^\varepsilon}e^{-i\int_{t_n}^{t_{n+1}}V(s,x)ds}e^{-\frac{i\tau}{2\varepsilon^2}\mathscr{T}^\varepsilon}\mathbf{e}^n(x)+\eta_1^n(x)+\eta_2^n(x)+\eta_3^n(x), \tag{3.4.25}$$

with η_j^n $(j=1,2,3)$ stated in Lemma 3.2 as

$$\|\eta_1^n(x)\|_{L^2}\lesssim\tau^3,\quad \eta_2^n(x)=-ie^{-\frac{i\tau}{\varepsilon^2}\mathscr{T}^\varepsilon}\left(\int_0^\tau f_2^n(s)ds-\tau f_2^n(\tau/2)\right), \tag{3.4.26}$$

$$\eta_3^n(x)=-e^{-\frac{i\tau}{\varepsilon^2}\mathscr{T}^\varepsilon}\left(\int_0^\tau\int_0^s\sum_{j=2}^4 g_j^n(s,w)dwds-\frac{\tau^2}{2}\sum_{j=2}^4 g_j^n(\tau/2,\tau/2)\right), \tag{3.4.27}$$

where f_2^n and g_j^n $(j=2,3,4)$ are given in (3.3.31)-(3.3.34).

Denote the second order splitting integrator $S_{n,\tau}=e^{-\frac{i\tau}{2\varepsilon^2}\mathscr{T}^\varepsilon}e^{-i\int_{t_n}^{t_{n+1}}V(s)ds}e^{-\frac{i\tau}{2\varepsilon^2}\mathscr{T}^\varepsilon}$ for $n\geq 0$, and $S_e(t;t_k)$ to be the exact solution flow (3.4.11) for the Dirac equation (3.2.2), then $S_{n,\tau}$ enjoys the similar properties as those in the first order Lie-Trotter splitting case (3.4.5) and we can get

$$\begin{aligned}\mathbf{e}^{n+1}(x)=&S_e(t_{n+1};t_n)\mathbf{e}^n(x)+\eta_1^n(x)+\eta_2^n(x)+\eta_3^n(x)+(S_{n,\tau}-S_e(t_{n+1};t_n))\,\mathbf{e}^n(x)\\ =&\ldots\\ =&S_e(t_{n+1};t_0)\mathbf{e}^0(x)+\sum_{k=0}^n S_e(t_{n+1};t_{k+1})(\eta_1^k(x)+\eta_2^k(x)+\eta_3^k(x))\\ &+\sum_{k=0}^n S_e(t_{n+1};t_{k+1})(S_{k,\tau}-S_e(t_{k+1};t_k))\mathbf{e}^k(x).\end{aligned} \tag{3.4.28}$$

By Duhamel's principle, it is straightforward to compute

$$\begin{aligned}\left(S_{k,\tau}-S_e(t_{k+1};t_k)\right)\tilde{\Phi}(x) &= e^{-\frac{i\tau}{2\varepsilon^2}\mathscr{T}^\varepsilon}(e^{-i\int_{t_k}^{t_{k+1}}V(s,x)ds}-1)e^{-\frac{i\tau}{2\varepsilon^2}\mathscr{T}^\varepsilon}\\ &\quad -i\int_0^\tau e^{-\frac{i(\tau-s)\mathscr{T}^\varepsilon}{\varepsilon^2}}V(t_k+s,x)S_e(t_k+s;t_k)\tilde{\Phi}(x)\,ds. \qquad (3.4.29)\end{aligned}$$

Recalling $\|e^{-i\int_{t_k}^{t_{k+1}}V(s,x)ds}-1\|_{L^\infty}\le\tau\|V(t,x)\|_{L^\infty([t_k,t_{k+1}];L^\infty)}$ and the properties of $S_e(t;t_k)$ (3.4.12), we obtain from (3.4.29)

$$\begin{aligned}&\left\|\left(S_{k,\tau}-S_e(t_{k+1};t_k)\right)\tilde{\Phi}(x)\right\|_{L^2}\\ &\le\tau\|V(t,x)\|_{L^\infty([t_k,t_{k+1}];L^\infty)}\|\tilde{\Phi}\|_{L^2}+\tau\|V(t,x)\|_{L^\infty([t_k,t_{k+1}];L^\infty)}\|\tilde{\Phi}\|_{L^2}\lesssim\tau\|\tilde{\Phi}\|_{L^2},\end{aligned}$$

and

$$\|S_e(t_{n+1};t_{k+1})(S_{k,\tau}-S_e(t_{k+1};t_k))\mathbf{e}^k(x)\|_{L^2}\lesssim\tau\|\mathbf{e}^k(x)\|_{L^2},\quad k=0,...,n. \qquad (3.4.30)$$

Noticing $\|\mathbf{e}^0(x)\|_{L^2}=0$, combining (3.4.30) and (3.4.28), recalling $\|\eta_1^n(x)\|_{L^2}\lesssim\tau^3$, we can control

$$\begin{aligned}&\|\mathbf{e}^{n+1}(x)\|_{L^2}\\ &\le\sum_{k=0}^n\|S_e(t_{n+1};t_{k+1})(S_{k,\tau}-S_e(t_{k+1};t_k))\mathbf{e}^k\|_{L^2}+\sum_{j=1}^3\left\|\sum_{k=0}^n S_e(t_{n+1};t_{k+1})\eta_j^k(x)\right\|_{L^2}\\ &\lesssim\tau^2+\sum_{k=0}^n\tau\|\mathbf{e}^k(x)\|_{L^2}+\sum_{j=2}^3\left\|\sum_{k=0}^n S_e(t_{n+1};t_{k+1})\eta_j^k(x)\right\|_{L^2}. \qquad (3.4.31)\end{aligned}$$

Similar to the Lie-Trotter splitting S_1, the key to establish the improved error bounds for non-resonant τ is to derive refined estimates for the terms involving η_j^k $(j=2,3)$ in (3.4.31). To this purpose, we introduce the approximations $\tilde{\eta}_l^k(x)$ of $\eta_l^k(x)$ $(l=2,3,\,k=0,1,\ldots,n)$ as

$$\tilde{\eta}_2^k(x)=\int_0^\tau\tilde{f}_2^k(s)ds-\tau\tilde{f}_2^k(\frac{\tau}{2}),\ \tilde{\eta}_3^k(x)=\int_0^\tau\int_0^s\sum_{j=2}^4\tilde{g}_j^k(s,w)dwds-\frac{\tau^2}{2}\sum_{j=2}^4\tilde{g}_j^k(\frac{\tau}{2},\frac{\tau}{2}), \qquad (3.4.32)$$

where we expand $V(t_k+s,x)=V(t_k,x)+s\partial_t V(t_k,x)+O(s^2)$ and $e^{is\mathscr{D}^\varepsilon}=Id+is\mathscr{D}^\varepsilon+O(s^2)$ up to the linear term in $f_2^k(s)$ (3.3.31) and the zeroth order term in $g_j^k(s,w)$ $(j=2,3,4)$

(3.3.32)-(3.3.34), respectively,

$$\begin{aligned}
\tilde{f}_2^k(s) &= e^{\frac{i(2s-\tau)}{\varepsilon^2}}\left((s-\tau)\mathscr{D}^\varepsilon\Pi_+^\varepsilon(V(t_k)\Pi_-^\varepsilon\Phi(t_k))+s\Pi_+^\varepsilon(V(t_k)\mathscr{D}^\varepsilon\Pi_-^\varepsilon\Phi(t_k))\right)\\
&\quad - e^{\frac{i(\tau-2s)}{\varepsilon^2}}\left((s-\tau)\mathscr{D}^\varepsilon\Pi_-^\varepsilon(V(t_k)\Pi_+^\varepsilon\Phi(t_k))+s\Pi_-^\varepsilon(V(t_k)\mathscr{D}^\varepsilon\Pi_+^\varepsilon\Phi(t_k))\right)\\
&\quad - ise^{\frac{i(2s-\tau)}{\varepsilon^2}}\Pi_+^\varepsilon(\partial_t V(t_k)\Pi_-^\varepsilon\Phi(t_k)) - ise^{\frac{i(\tau-2s)}{\varepsilon^2}}\Pi_-^\varepsilon(\partial_t V(t_k)\Pi_+^\varepsilon\Phi(t_k))\\
&\quad - ie^{\frac{i(2s-\tau)}{\varepsilon^2}}\Pi_+^\varepsilon(V(t_k)\Pi_-^\varepsilon\Phi(t_k)) - ie^{\frac{i(\tau-2s)}{\varepsilon^2}}\Pi_-^\varepsilon(V(t_k)\Pi_+^\varepsilon\Phi(t_k)),\\
\tilde{g}_2^k(s,w) &= -ie^{\frac{i(2w-\tau)}{\varepsilon^2}}\Pi_+^\varepsilon\left(V(t_k)\Pi_+^\varepsilon\left(V(t_k)\Pi_-^\varepsilon\Phi(t_k)\right)\right)\\
&\quad - ie^{\frac{i(\tau-2w)}{\varepsilon^2}}\Pi_-^\varepsilon\left(V(t_k)\Pi_-^\varepsilon\left(V(t_k)\Pi_+^\varepsilon\Phi(t_k)\right)\right),\\
\tilde{g}_3^k(s,w) &= -ie^{\frac{i(2(s-w)-\tau)}{\varepsilon^2}}\Pi_+^\varepsilon\left(V(t_k)\Pi_-^\varepsilon\left(V(t_k)\Pi_+^\varepsilon\Phi(t_k)\right)\right)\\
&\quad - ie^{\frac{i(\tau-2(s-w))}{\varepsilon^2}}\Pi_-^\varepsilon\left(V(t_k)\Pi_+^\varepsilon\left(V(t_k)\Pi_-^\varepsilon\Phi(t_k)\right)\right),\\
\tilde{g}_4^k(s,w) &= -ie^{\frac{i(2s-\tau)}{\varepsilon^2}}\Pi_+^\varepsilon\left(V(t_k)\Pi_-^\varepsilon\left(V(t_k)\Pi_-^\varepsilon\Phi(t_k)\right)\right)\\
&\quad - ie^{\frac{i(\tau-2s)}{\varepsilon^2}}\Pi_-^\varepsilon\left(V(t_k)\Pi_+^\varepsilon\left(V(t_k)\Pi_+^\varepsilon\Phi(t_k)\right)\right).
\end{aligned}$$

Using Taylor expansion in $f_2^k(s)$ (3.3.31) and $g_j^k(s,w)$ $(j=2,3,4)$ (3.3.32)-(3.3.34) as well as properties of $\mathscr{D}^\varepsilon$, it is not difficult to check that

$$\begin{aligned}
&\|\eta_2^k(x)-\tilde{\eta}_2^k(x)\|_{L^2}\\
&\lesssim \tau^3\Bigg(\|V(t,x)\|_{W^{2,\infty}([0,T];L^\infty)}\|\Phi(t_k)\|_{L^2}+\|\partial_t V(t,x)\|_{W^{1,\infty}([0,T];H^2)}\|\|\Phi(t_k)\|_{H^2}\\
&\qquad +\|V(t,x)\|_{L^\infty([0,T];H^4)}\|\Phi(t_k)\|_{H^4}\Bigg)\lesssim\tau^3,\\
&\|\eta_3^k(x)-\tilde{\eta}_3^k(x)\|_{L^2}\lesssim\tau^3\|V(t_n,x)\|_{W^{2,\infty}}^2\|\Phi(t_k)\|_{H^2}\lesssim\tau^3,
\end{aligned}$$

which would yield for $k\le n\le\frac{T}{\tau}-1$,

$$\left\|S_e(t_{n+1};t_{k+1})\eta_2^k(x)-S_e(t_{n+1};t_{k+1})\tilde{\eta}_2^k(x)\right\|_{L^2}\lesssim\|\eta_2^k(x)-\tilde{\eta}_2^k(x)\|_{L^2}\lesssim\tau^3, \tag{3.4.33}$$

$$\left\|S_e(t_{n+1};t_{k+1})\eta_3^n(x)-S_e(t_{n+1};t_{k+1})\tilde{\eta}_3^k(x)\right\|_{L^2}\lesssim\|\eta_3^k(x)-\tilde{\eta}_3^k(x)\|_{L^2}\lesssim\tau^3. \tag{3.4.34}$$

Plugging the above inequalities (3.4.33)-(3.4.34) into (3.4.31), we derive

$$\begin{aligned}\|\mathbf{e}^{n+1}(x)\|_{L^2} &\lesssim \tau^2 + \sum_{k=0}^{n}\tau^3 + \sum_{j=2}^{3}\left\|\sum_{k=0}^{n} S_e(t_{n+1};t_{k+1})\tilde{\eta}_j^k(x)\right\|_{L^2} + \sum_{k=0}^{n}\tau\|\mathbf{e}^k(x)\|_{L^2}\\ &\lesssim \tau^2 + \left\|\sum_{k=0}^{n} S_e(t_{n+1};t_{k+1})\tilde{\eta}_2^k(x)\right\|_{L^2} + \left\|\sum_{k=0}^{n} S_e(t_{n+1};t_{k+1})\tilde{\eta}_3^k(x)\right\|_{L^2}\\ &\quad + \sum_{k=0}^{n}\tau\|\mathbf{e}^k(x)\|_{L^2}. \end{aligned} \tag{3.4.35}$$

Step 2 (Improved estimates for non-resonant time steps). It remains to show the estimates on the terms related to $\tilde{\eta}_2^k$ and $\tilde{\eta}_3^k$. The arguments will be similar to those in the proof of the Lie-Trotter splitting case Theorem 3.3, so we only sketch the proof below. Taking $\tilde{\eta}_2^k$ for example, we write

$$\tilde{\eta}_2^k(s) = \tilde{\eta}_{2+}^k(s) + \tilde{\eta}_{2-}^k(s), \quad \tilde{\eta}_{2\pm}^k(x) = \int_0^\tau \tilde{f}_{2\pm}^k(s)ds - \tau\tilde{f}_{2\pm}^k(\tau/2), \quad k = 0,1,...,n, \tag{3.4.36}$$

with

$$\begin{aligned}\tilde{f}_{2\pm}^k(s) =& e^{\pm i(2s-\tau)/\varepsilon^2}\left(\pm(s-\tau)\mathscr{D}^\varepsilon\Pi_\pm^\varepsilon(V(t_k)\Pi_\mp^\varepsilon\Phi(t_k)) \pm s\Pi_\pm^\varepsilon(V(t_k)\mathscr{D}^\varepsilon\Pi_\mp^\varepsilon\Phi(t_k))\right)\\ &- ise^{\pm i(2s-\tau)/\varepsilon^2}\,\Pi_\pm^\varepsilon(\partial_t V(t_k)\Pi_\mp^\varepsilon\Phi(t_k)) - ie^{\pm i(2s-\tau)/\varepsilon^2}\Pi_\pm^\varepsilon(V(t_k)\Pi_\mp^\varepsilon\Phi(t_k))\end{aligned}$$

and $\tilde{f}_2^k(s) = \tilde{f}_{2+}^n(s) + \tilde{f}_{2-}^n(s)$.

Recalling the structure of the exact solution to the Dirac equation in (3.4.18), we have for $0 \le k \le n$

$$S_e(t_n;t_k)\tilde{\Phi}(x) = e^{-i(t_n-t_k)/\varepsilon^2}S_e^+(t_n;t_k)\tilde{\Phi}(x) + e^{i(t_n-t_k)/\varepsilon^2}S_e^-(t_n;t_k)\tilde{\Phi}(x) + R_k^n\tilde{\Phi}(x),$$

where the propagators $S_e^\pm$ and the residue operator $R_k^n : (L^2)^2 \to (L^2)^2$ are defined in (3.4.18). Therefore, we can get

$$\sum_{k=0}^{n} S_e(t_{n+1};t_{k+1})\tilde{\eta}_{2+}^k(x) = \sum_{j=1}^{4}\tilde{I}_j^n(x),$$

with

$$\begin{aligned}
\tilde{I}_1^n(x) &= \tilde{p}_1(\tau)\sum_{k=0}^{n} e^{-\frac{i(t_{n+1}-2t_k-\tau)}{\varepsilon^2}} S_e^+(t_{n+1};t_{k+1})\Pi_+^\varepsilon V(t_k)\Pi_-^\varepsilon S_e^-(t_k;t_0)\Phi(0),\\
\tilde{I}_2^n(x) &= \tilde{p}_1(\tau)\sum_{k=0}^{n} (R_{k+1}^{n+1}\Pi_+^\varepsilon V(t_k)\Pi_-^\varepsilon \Phi(t_k) + S_e(t_{n+1};t_{k+1})\Pi_+^\varepsilon V(t_k)\Pi_-^\varepsilon R_0^k\Phi(0)),\\
\tilde{I}_3^n(x) &= \sum_{k=0}^{n} e^{-\frac{i(t_{n+1}-2t_k-\tau)}{\varepsilon^2}} S_e^+(t_{n+1};t_{k+1})\Big(\tilde{p}_2(\tau)\mathscr{D}^\varepsilon\Pi_+^\varepsilon V(t_k)\\
&\qquad + \tilde{p}_3(\tau)\Pi_+^\varepsilon V(t_k)\mathscr{D}^\varepsilon - i\tilde{p}_3(\tau)\Pi_+^\varepsilon\partial_t V(t_k)\Big)\Pi_-^\varepsilon S_e^-(t_k;t_0)\Phi(0),\\
\tilde{I}_4^n(x) &= \sum_{k=0}^{n}\Big(R_{k+1}^{n+1}\left(\tilde{p}_2(\tau)\mathscr{D}^\varepsilon\Pi_+^\varepsilon V(t_k) + \tilde{p}_3(\tau)\left(\Pi_+^\varepsilon V(t_k)\mathscr{D}^\varepsilon - i\Pi_+^\varepsilon\partial_t V(t_k)\right)\right)\Pi_-^\varepsilon\Phi(t_k)\\
&\qquad + S_e(t_{n+1};t_{k+1})\left(\tilde{p}_2(\tau)\mathscr{D}^\varepsilon\Pi_+^\varepsilon V(t_k) + \tilde{p}_3(\tau)\left(\Pi_+^\varepsilon V(t_k)\mathscr{D}^\varepsilon - i\Pi_+^\varepsilon\partial_t V(t_k)\right)\right)\\
&\qquad \Pi_-^\varepsilon R_0^k\Phi(0)\Big),
\end{aligned}$$

where

$$\begin{aligned}
&\tilde{p}_1(\tau) = -i\left(\int_0^\tau e^{i(2s-\tau)/\varepsilon^2}ds - \tau\right), \quad \tilde{p}_2(\tau) = \left(\int_0^\tau (s-\tau)e^{i(2s-\tau)/\varepsilon^2}ds + \frac{\tau^2}{2}\right),\\
&\tilde{p}_3(\tau) = \left(\int_0^\tau s e^{i(2s-\tau)/\varepsilon^2}ds - \frac{\tau^2}{2}\right).
\end{aligned}$$

The residue terms $\tilde{I}_2^n$ and $\tilde{I}_4^n$ will be estimated first. Using the properties of R_k^n and S_e, noticing (3.3.50)-(3.3.51), we have

$$\begin{aligned}
&\|R_{k+1}^{n+1}\Pi_+^\varepsilon V(t_k)\Pi_-^\varepsilon\Phi(t_k) + S_e(t_{n+1};t_{k+1})\Pi_+^\varepsilon V(t_k)\Pi_-^\varepsilon R_0^k\Phi(0)\|_{L^2}\\
&\lesssim \varepsilon^3\|V(t_k)\|_{W^{3,\infty}}\left(\|\Phi(t_k)\|_{H^3} + \|\Phi(0)\|_{H^3}\right),\\
&\|R_{k+1}^{n+1}\mathscr{D}^\varepsilon\Pi_+^\varepsilon V(t_k)\Pi_-^\varepsilon\Phi(t_k)\|_{L^2} + \|R_{k+1}^{n+1}(\Pi_+^\varepsilon V(t_k)\mathscr{D}^\varepsilon - i\Pi_+^\varepsilon\partial_t V(t_k))\Pi_-^\varepsilon\Phi(t_k)\|_{L^2}\\
&\lesssim \varepsilon^3\|V(t,x)\|_{W^{1,\infty}([0,T];W^{5,\infty})}\|\Phi(t_k)\|_{H^5},\\
&\|S_e(t_{n+1};t_{k+1})\mathscr{D}^\varepsilon\Pi_+^\varepsilon V(t_k)\Pi_-^\varepsilon R_0^k\Phi(0)\|_{L^2} \lesssim \varepsilon^3\|V(t,x)\|_{W^{1,\infty}([0,T];W^{3,\infty})}\|\Phi(0)\|_{H^5},\\
&\|S_e(t_{n+1};t_{k+1})(\Pi_+^\varepsilon V(t_k)\mathscr{D}^\varepsilon - i\Pi_+^\varepsilon\partial_t V(t_k))\Pi_-^\varepsilon R_0^k\Phi(0)\|_{L^2}\\
&\lesssim \varepsilon^3\|V(t,x)\|_{W^{1,\infty}([0,T];W^{3,\infty})}\|\Phi(0)\|_{H^5},
\end{aligned}$$

which will lead to the following conclusions in view of the fact that $|\tilde{p}_1(\tau)| = |\int_0^\tau e^{i(2s-\tau)/\varepsilon^2}ds - \tau| \lesssim \min\{\tau^2/\varepsilon^2, \tau^3/\varepsilon^4\}$ and $|\tilde{p}_2(\tau)|, |\tilde{p}_3(\tau)| \lesssim \min\{\tau^2/\varepsilon^2, \tau^3/\varepsilon^4\}$ (Taylor expansion up to

the linear or the quadratic term),

$$\|\tilde{I}_2^n(x)\|_{L^2} \lesssim \min\{\tau\varepsilon, \tau^2/\varepsilon\}, \quad \|\tilde{I}_4^n(x)\|_{L^2} \lesssim \min\{\tau\varepsilon, \tau^2/\varepsilon\}. \tag{3.4.37}$$

Now, we proceed to treat $\tilde{I}_1^n$ and $\tilde{I}_3^n$. For $\tilde{I}_1^n(x)$, it is similar to (3.4.21) which has been analyzed in the S_1 case. Using the same idea (details omitted for brevity here), and the fact that $|\tilde{p}_1(\tau)| = \left|\int_0^\tau e^{i(2s-\tau)/\varepsilon^2} ds - \tau\right| \lesssim \min\{\tau, \tau^2/\varepsilon^2\}$ as well as $\Pi_\pm^\varepsilon V(t_k)\Pi_\mp^\varepsilon = O(\varepsilon)$, under the regularity assumptions, we can get for $\tau \in \mathscr{A}_\kappa(\varepsilon)$,

$$\|\tilde{I}_1^n(x)\|_{L^2} \lesssim \min\{\tau, \tau^2/\varepsilon^2\}(\sum_{k=0}^{n-1} \tau\varepsilon/\kappa + \varepsilon/\kappa) \lesssim_\kappa \min\{\tau\varepsilon, \tau^2/\varepsilon\}. \tag{3.4.38}$$

Similarly, noticing $|\tilde{p}_2(\tau)|, |\tilde{p}_3(\tau)| \leq \tau^2$, we can get

$$\|\tilde{I}_3^n(x)\|_{L^2} \lesssim \tau^2(\sum_{k=0}^{n-1} \tau\varepsilon/\kappa + \varepsilon/\kappa) \lesssim_\kappa \tau^2\varepsilon. \tag{3.4.39}$$

Combing the estimates for $\tilde{I}_j^n$ $(j = 1, 2, 3, 4)$, we have

$$\left\|\sum_{k=0}^{n} S_e(t_{n+1}; t_{k+1})\tilde{\eta}_{2+}^k(x)\right\|_{L^2} \leq \sum_{j=1}^{4} \|\tilde{I}_j^n(x)\|_{L^2} \lesssim_\kappa \min\{\tau\varepsilon, \tau^2/\varepsilon\}. \tag{3.4.40}$$

For $\sum_{k=0}^n S_e(t_{n+1}; t_{k+1})\tilde{\eta}_{2-}^k(x)$, we can have the same results as

$$\left\|\sum_{k=0}^{n} S_e(t_{n+1}; t_{k+1})\tilde{\eta}_{2-}^k(x)\right\|_{L^2} \lesssim_\kappa \min\{\tau\varepsilon, \tau^2/\varepsilon\}, \tag{3.4.41}$$

which yield the following results in view of (3.4.40) and (3.4.36)

$$\left\|\sum_{k=0}^{n} S_e(t_{n+1}; t_{k+1})\tilde{\eta}_{2}^k(x)\right\|_{L^2} \lesssim_\kappa \min\{\tau\varepsilon, \tau^2/\varepsilon\}. \tag{3.4.42}$$

The same technique works for $S_e(t_{n+1}; t_{k+1})\tilde{\eta}_3^k(x)$ and we can get

$$\left\|\sum_{k=0}^{n} S_e(t_{n+1}; t_{k+1})\tilde{\eta}_{3}^k(x)\right\|_{L^2} \lesssim_\kappa \min\{\tau\varepsilon, \tau^2/\varepsilon\}. \tag{3.4.43}$$

Plugging these results into (3.4.35), we have

$$\|\mathbf{e}^{n+1}(x)\|_{L^2} \lesssim_\kappa \tau^2 + \sum_{k=0}^{n} \tau\|\mathbf{e}^k(x)\|_{L^2} + \min\{\tau\varepsilon, \tau^2/\varepsilon\}. \tag{3.4.44}$$

Gronwall's inequality then implies for τ satisfying $\tau \in \mathscr{A}_\kappa(\varepsilon)$,

$$\|\mathbf{e}^{n+1}(x)\|_{L^2} \lesssim_\kappa \tau^2 + \min\{\tau\varepsilon, \tau^2/\varepsilon\}, \quad 0 \le n \le \frac{T}{\tau} - 1. \tag{3.4.45}$$

This completes the proof for Theorem 3.4. □

Remark 3.2. *In Theorem 3.3 and Theorem 3.4, the constants in the error estimates depend on κ and the proof suggests that the constants are bounded from above by $\frac{T}{\tau}C$ and $\frac{2}{\kappa}C$ with some common factor C independent of κ and τ. The optimality of the uniform error bounds in Theorem 3.3 and Theorem 3.4 will be verified by numerical examples later.*

3.5 Numerical results

In this section, we report two numerical examples to verify our theorems. For spatial discretization, we use the Fourier pseudospectral method.

In both examples, we choose the electric potential in (3.2.2) as

$$V(t,x) = \frac{1-x}{1+x^2}, \quad x \in \mathbb{R}, \quad t \ge 0, \tag{3.5.1}$$

and the initial data in (3.2.3) as

$$\phi_1(0,x) = e^{-\frac{x^2}{2}}, \quad \phi_2(0,x) = e^{-\frac{(x-1)^2}{2}}, \quad x \in \mathbb{R}. \tag{3.5.2}$$

In the numerical simulations, as a common practice, we truncate the whole space onto a sufficiently large bounded domain $\Omega = (a,b)$, and assume periodic boundary conditions. The mesh size is chosen as $h := \triangle x = \frac{b-a}{M}$ with M being an even positive integer. The grid points can be denoted as $x_j := a + jh$, for $j = 0, 1, ..., M$.

To show the numerical results, we introduce the discrete l^2 errors of the numerical solution. Let $\Phi^n = (\Phi_0^n, \Phi_1^n, ..., \Phi_{M-1}^n, \Phi_M^n)^T$ be the numerical solution obtained by a numerical method with time step τ and ε as well as a very fine mesh size h at time $t = t_n$, and $\Phi(t,x)$ be the exact solution, then the discrete l^2 error is quantified as

$$e^{\varepsilon,\tau}(t_n) = \|\Phi^n - \Phi(t_n,\cdot)\|_{l^2} = \sqrt{h\sum_{j=0}^{M-1} |\Phi(t_n,x_j) - \Phi_j^n|^2}, \tag{3.5.3}$$

and $e(t_n)$ should be close to the L^2 errors in Theorem 3.1, Theorem 3.2, Theorem 3.3 & Theorem 3.4 for fine spatial mesh sizes h.

Example 1 We first test the uniform error bounds for the splitting methods. In this example, we choose resonant time step size, that is, for small enough chosen ε, there is a positive k_0, such that $\tau = k_0\varepsilon\pi$.

The bounded computational domain is set as $\Omega = (-32, 32)$. Because we mainly care about the temporal errors to verify super-resolution, during the computation, the spatial mesh size is always set to be $h = \frac{1}{16}$ so that the spatial error is negligible. As there is no exact solution available, for comparison, we use a numerical 'exact' solution generated by the S_2 method with a very fine time step size $\tau_e = 2\pi \times 10^{-6}$.

Table 3.5.1 & Table 3.5.2 show the numerical errors $e^{\varepsilon,\tau}(t = 2\pi)$ with different ε and time step size τ for S_1 and S_2, respectively.

In Table 3.5.1 & Table 3.5.2, the last two rows show the largest error of each column for fixed τ. They both give $1/2$ order of convergence, which coincides well with Theorem 3.1 & Theorem 3.2. More specifically, in Table 3.5.1, we can see when $\tau \gtrsim \varepsilon$ (below the lower bolded line), there is first order convergence, which agrees with the error bound $\|\Phi(t_n, x) - \Phi^n(x)\|_{L^2} \lesssim \tau + \varepsilon$. When $\tau \lesssim \varepsilon^2$ (above the upper bolded line), we observe first order convergence, which matches the other error bound $\|\Phi(t_n, x) - \Phi^n(x)\|_{L^2} \lesssim \tau + \tau/\varepsilon$. Similarly, in Table 3.5.2, the second order convergence can be clearly observed when $\tau \lesssim \varepsilon^2$ (above the upper bolded line) or when $\tau \gtrsim \sqrt{\varepsilon}$ (below the lower bolded line), which fits well with the two error bounds $\|\Phi(t_n, x) - \Phi^n(x)\|_{L^2} \lesssim \tau^2 + \tau^2/\varepsilon^3$ and $\|\Phi(t_n, x) - \Phi^n(x)\|_{L^2} \lesssim \tau^2 + \varepsilon$.

Through the results of this example, we successfully validate the uniform error bounds for the splitting methods in Theorem 3.1 & Theorem 3.2.

Example 2 In this example, we test the improved uniform error bounds for non-resonant time step size. Here we choose $\tau \in \mathscr{A}_\kappa(\varepsilon)$ for some given ε and $0 < \kappa \leq 1$.

The bounded computational domain is set as $\Omega = (-16, 16)$. The numerical 'exact' solution is computed by S_2 with a very small time step $\tau_e = 8 \times 10^{-6}$. Spatial mesh size is

Table 3.5.1: Discrete l^2 temporal errors $e^{\varepsilon,\tau}(t=2\pi)$ for the wave function with resonant time step size, S_1 method.

$e^{\varepsilon,\tau}(t=2\pi)$	$\tau_0=\pi/4$	$\tau_0/4$	$\tau_0/4^2$	$\tau_0/4^3$	$\tau_0/4^4$	$\tau_0/4^5$
$\varepsilon_0=1$	4.84E-1	**1.27E-1**	3.20E-2	8.03E-3	2.01E-3	5.02E-4
order	–	**0.97**	0.99	1.00	1.00	1.00
$\varepsilon_0/2$	6.79E-1	1.21E-1	**3.10E-2**	7.78E-3	1.95E-3	4.87E-4
order	–	1.24	**0.98**	1.00	1.00	1.00
$\varepsilon_0/2^2$	5.78E-1	2.71E-1	3.07E-2	**7.76E-3**	1.95E-3	4.87E-4
order	–	0.55	1.57	**0.99**	1.00	1.00
$\varepsilon_0/2^3$	**5.33E-1**	1.85E-1	1.21E-1	7.75E-3	**1.95E-3**	4.87E-4
order	**–**	0.76	0.30	1.98	**1.00**	1.00
$\varepsilon_0/2^4$	5.13E-1	1.48E-1	7.02E-2	5.76E-2	1.95E-3	**4.88E-4**
order	–	0.90	0.54	0.14	2.44	**1.00**
$\varepsilon_0/2^5$	5.04E-1	**1.34E-1**	4.70E-2	3.07E-2	2.82E-2	4.88E-4
order	–	**0.96**	0.75	0.31	0.06	2.93
$\varepsilon_0/2^7$	4.98E-1	1.25E-1	**3.37E-2**	1.18E-2	7.68E-3	7.05E-3
order	–	1.00	**0.95**	0.76	0.31	0.06
$\varepsilon_0/2^9$	4.97E-1	1.24E-1	3.17E-2	**8.46E-3**	2.95E-3	1.92E-3
order	–	1.00	0.98	**0.95**	0.76	0.31
$\varepsilon_0/2^{11}$	4.96E-1	1.23E-1	3.13E-2	7.94E-3	**2.12E-3**	7.37E-4
order	–	1.00	0.99	0.99	**0.95**	0.76
$\max_{0<\varepsilon\le 1} e^{\varepsilon,\tau}(t=2\pi)$	6.79E-1	2.71E-1	1.21E-1	5.76E-2	2.82E-2	1.39E-2
order	–	0.66	0.58	0.54	0.52	0.51

fixed as $h=1/16$ for all the numerical simulations.

Table 3.5.3 & Table 3.5.4 show the numerical errors $e^{\varepsilon,\tau}(t=4)$ with different ε and time step size τ for S_1 and S_2, respectively.

In Table 3.5.3, we could see that overall, for fixed time step size τ, the error $e^{\varepsilon,\tau}(t=4)$ does not change with different ε. This verifies the uniform first order convergence in time for S_1 with non-resonant time step size, as stated in Theorem 3.3. In Table 3.5.4, the last two rows show the largest error of each column for fixed τ, which gives $3/2$ order of convergence, and it is consistent with Theorem 3.4. More specifically, in Table 3.5.4, we can observe the second order convergence when $\tau \gtrsim \varepsilon$ (below the lower bolded line) or when $\tau \lesssim \varepsilon^2$ (above the upper bolded line). The lower bolded diagonal line agrees with the error bound

Table 3.5.2: Discrete l^2 temporal errors $e^{\varepsilon,\tau}(t=2\pi)$ for the wave function with resonant time step size, S_2 method.

$e^{\varepsilon,\tau}(t=2\pi)$	$\tau_0=\pi/4$	$\tau_0/4^2$	$\tau_0/4^3$	$\tau_0/4^4$	$\tau_0/4^5$	$\tau_0/4^6$
$\varepsilon_0=1$	8.08E-2	**4.44E-3**	2.76E-4	1.73E-5	1.08E-6	6.74E-8
order	–	**2.09**	2.00	2.00	2.00	2.00
$\varepsilon_0/2$	4.13E-1	9.66E-3	**5.73E-4**	3.57E-5	2.23E-6	1.39E-7
order	–	2.71	**2.04**	2.00	2.00	2.00
$\varepsilon_0/2^2$	2.63E-1	2.15E-1	1.21E-3	**7.22E-5**	4.50E-6	2.81E-7
order	–	0.15	3.74	**2.03**	2.00	2.00
$\varepsilon_0/2^3$	**2.08E-1**	1.10E-1	1.10E-1	1.51E-4	**9.05E-6**	5.64E-7
order	**–**	0.46	0.00	4.75	**2.03**	2.00
$\varepsilon_0/2^4$	1.92E-1	5.56E-2	5.51E-2	5.51E-2	1.89E-5	**1.13E-6**
order	–	0.89	0.01	0.00	5.76	**2.03**
$\varepsilon_0/2^5$	1.88E-1	2.85E-2	2.76E-2	2.76E-2	2.76E-2	2.36E-6
order	–	1.36	0.02	0.00	0.00	6.76
$\varepsilon_0/2^6$	1.87E-1	1.55E-2	1.38E-2	1.38E-2	1.38E-2	1.38E-2
order	–	1.79	0.08	0.00	0.00	0.00
$\varepsilon_0/2^7$	1.87E-1	**9.86E-3**	6.92E-3	6.90E-3	6.90E-3	6.90E-3
order	–	**2.12**	0.26	0.00	0.00	0.00
$\varepsilon_0/2^{11}$	1.87E-1	6.97E-3	**5.93E-4**	4.32E-4	4.31E-4	4.31E-4
order	–	2.37	**1.78**	0.23	0.00	0.00
$\varepsilon_0/2^{15}$	1.87E-1	6.95E-3	4.03E-4	**3.75E-5**	2.71E-5	2.70E-5
order	–	2.37	2.05	**1.71**	0.23	0.00
$\max_{0<\varepsilon\le 1} e^{\varepsilon,\tau}(t=2\pi)$	4.13E-1	2.15E-1	1.10E-1	5.51E-2	2.76E-2	1.38E-2
order	–	0.47	0.49	0.50	0.50	0.50

$\|\Phi(t_n,x)-\Phi^n(x)\|_{L^2} \lesssim \tau^2+\tau\varepsilon$, and the upper bolded diagonal line matches the other error bound $\|\Phi(t_n,x)-\Phi^n(x)\|_{L^2} \lesssim \tau^2+\tau^2/\varepsilon$.

Through the results of this example, we successfully validate the improved uniform error bounds for the splitting methods in Theorem 3.3 and Theorem 3.4, with non-resonant time step size.

Table 3.5.3: Discrete l^2 temporal errors $e^{\varepsilon,\tau}(t=4)$ for the wave function with non-resonant time step size, S_1 method.

$e^{\varepsilon,\tau}(t=4)$	$\tau_0=1/2$	$\tau_0/2$	$\tau_0/2^2$	$\tau_0/2^3$	$\tau_0/2^4$	$\tau_0/2^5$	$\tau_0/2^6$
$\varepsilon_0=1$	3.51E-1	1.78E-1	8.96E-2	4.50E-2	2.25E-2	1.13E-2	5.64E-3
order	–	0.98	0.99	0.99	1.00	1.00	1.00
$\varepsilon_0/2$	3.52E-1	1.65E-1	8.34E-2	4.20E-2	2.11E-2	1.05E-2	5.28E-3
order	–	1.10	0.98	0.99	1.00	1.00	1.00
$\varepsilon_0/2^2$	3.25E-1	1.64E-1	8.04E-2	4.07E-2	2.05E-2	1.03E-2	5.15E-3
order	–	0.99	1.03	0.98	0.99	1.00	1.00
$\varepsilon_0/2^3$	3.24E-1	1.69E-1	8.10E-2	4.13E-2	2.02E-2	1.02E-2	5.13E-3
order	–	0.94	1.06	0.97	1.03	0.99	0.99
$\varepsilon_0/2^4$	3.12E-1	1.61E-1	8.24E-2	4.22E-2	2.05E-2	1.03E-2	5.10E-3
order	–	0.95	0.97	0.97	1.04	0.99	1.02
$\varepsilon_0/2^5$	3.25E-1	1.61E-1	8.10E-2	4.10E-2	2.07E-2	1.04E-2	5.13E-3
order	–	1.02	0.99	0.98	0.99	0.98	1.02
$\varepsilon_0/2^6$	3.19E-1	1.63E-1	8.43E-2	4.09E-2	2.05E-2	1.03E-2	5.16E-3
order	–	0.97	0.95	1.04	1.00	0.99	0.99
$\varepsilon_0/2^7$	3.18E-1	1.60E-1	8.10E-2	4.06E-2	2.05E-2	1.03E-2	5.13E-3
order	–	0.99	0.99	0.99	0.99	0.99	1.00
$\max\limits_{0<\varepsilon\le 1} e^{\varepsilon,\tau}$	3.52E-1	1.78E-1	8.96E-2	4.50E-2	2.25E-2	1.13E-2	5.64E-3
order	–	0.98	0.99	0.99	1.00	1.00	1.00

3.6 Extension to full-discretization

Theorem 3.1 to Theorem 3.4 in the above sections only deal with semi-discretization. In this section, we extend these error estimates to full-discretization.

Consider (3.2.2) with the initial condition (3.2.3) on a bounded domain $\Omega=[a,b]$ with periodic boundary conditions. Choose mesh size $h=\frac{b-a}{M}$ with M being an even positive integer, time step size $\tau:=\kappa t>0$ and denote the grid points and time steps as:

$$x_j:=a+jh,\quad j=0,1,\ldots,M;\qquad t_n:=n\tau,\quad n=0,1,2,\ldots.$$

Denote $X_M=\{U=(U_0,U_1,...,U_M)^T \mid U_j\in\mathbb{C}^2, j=0,1,\ldots,M,\ U_0=U_M\}$ and we always use $U_{-1}=U_{M-1}$ and $U_{M+1}=U_1$ if they are involved.

Table 3.5.4: Discrete l^2 temporal errors $e^{\varepsilon,\tau}(t=4)$ for the wave function with non-resonant time step size, S_2 method.

$e^{\varepsilon,\tau}(t=4)$	$\tau_0=1/2$	$\tau_0/4$	$\tau_0/4^2$	$\tau_0/4^3$	$\tau_0/4^4$	$\tau_0/4^5$
$\varepsilon_0=1/2$	1.69E-1	3.85E-3	**2.36E-4**	1.47E-5	9.20E-7	5.75E-8
order	–	2.73	**2.01**	2.00	2.00	2.00
$\varepsilon_0/2$	9.79E-2	1.16E-2	4.61E-4	**2.83E-5**	1.77E-6	1.10E-7
order	–	1.54	2.33	**2.01**	2.00	2.00
$\varepsilon_0/2^2$	6.76E-2	3.93E-3	1.32E-3	5.76E-5	**3.54E-6**	2.21E-7
order	–	2.05	0.78	2.26	**2.01**	2.00
$\varepsilon_0/2^3$	7.86E-2	4.49E-3	2.63E-4	1.72E-4	7.59E-6	**4.67E-7**
order	–	2.06	2.05	0.31	2.25	**2.01**
$\varepsilon_0/2^4$	7.55E-2	5.04E-3	5.33E-4	2.64E-5	2.14E-5	9.43E-7
order	–	1.95	1.62	2.17	0.15	2.25
$\varepsilon_0/2^5$	**7.01E-2**	1.94E-2	2.38E-4	6.50E-5	3.02E-6	2.61E-6
order	**–**	0.93	3.18	0.94	2.22	0.10
$\varepsilon_0/2^7$	6.84E-2	**2.67E-3**	2.77E-4	2.31E-4	2.76E-6	1.04E-6
order	–	**2.34**	1.64	0.13	3.19	0.70
$\varepsilon_0/2^9$	6.84E-2	2.67E-3	**1.65E-4**	1.03E-5	2.08E-6	2.10E-6
order	–	2.34	**2.01**	2.00	1.15	-0.00
$\varepsilon_0/2^{11}$	6.84E-2	2.67E-3	1.66E-4	**1.03E-5**	6.53E-7	4.53E-8
order	–	2.34	2.00	**2.00**	1.99	1.92
$\varepsilon_0/2^{13}$	6.84E-2	2.67E-3	1.64E-4	1.04E-5	**7.51E-7**	1.51E-7
order	–	2.34	2.01	1.99	**1.89**	1.16
$\max_{0<\varepsilon\le 1} e^{\varepsilon,\tau}(t=4)$	1.69E-1	1.94E-2	4.11E-3	2.31E-4	2.14E-5	2.61E-6
order	–	1.56	1.12	2.08	1.72	1.52

Denote

$$Y_M = Z_M \times Z_M,\ Z_M = \text{span}\left\{\phi_l(x) = e^{i\mu_l(x-a)},\ l = -\frac{M}{2},\ldots,\frac{M}{2}-1\right\},$$

where $\mu_l = \frac{2l\pi}{b-a}$ with $l = -\frac{M}{2},...,\frac{M}{2}-1$. Let $[C_p(\overline{\Omega})]^2$ be the function space consisting of all periodic vector function $U(x):\ \overline{\Omega}=[a,b]\to\mathbb{C}^2$. For any $U(x)\in[C_p(\overline{\Omega})]^2$ and $U\in X_M$, define $P_M:\ [L^2(\Omega)]^2\to Y_M$ as the standard projection operator [109], $I_M:\ [C_p(\overline{\Omega})]^2\to Y_M$ and $I_M: X_M\to Y_M$ as the standard interpolation operator [109], i.e. for $a\le x\le b$

$$(P_M U)(x) = \sum_{l=-M/2}^{M/2-1} \widehat{U}_l\, e^{i\mu_l(x-a)}, \quad (I_M U)(x) = \sum_{l=-M/2}^{M/2-1} \widetilde{U}_l\, e^{i\mu_l(x-a)}, \tag{3.6.1}$$

with

$$\widehat{U}_l = \frac{1}{b-a}\int_a^b U(x)\, e^{-i\mu_l(x-a)}\, dx, \quad \widetilde{U}_l = \frac{1}{M}\sum_{j=0}^{M-1} U_j\, e^{-2ijl\pi/M}, \tag{3.6.2}$$

where $U_j = U(x_j)$ when U is a function.

We first consider the Lie-Trotter splitting S_1. Denote $\Phi^{[n]}(x)$ to be the semi-discretized numerical solution from S_1 (3.2.4), and Φ^n to be the full-discretized numerical solution with Fourier spectral discretization in space, i.e. we have for $n = 0, 1, ..., \frac{T}{\tau} - 1$

$$\Phi^{[n+1]}(x) = e^{-i\int_{t_n}^{t_{n+1}} V(s,x)ds} e^{-\frac{i\tau}{\varepsilon^2}\mathscr{T}^\varepsilon} \Phi^{[n]}(x), \quad x \in [a,b], \tag{3.6.3}$$

with

$$\Phi^{[0]}(x) = \Phi(0,x), \quad x \in [a,b], \tag{3.6.4}$$

and

$$\Phi_j^{n+1} = e^{-i\int_{t_n}^{t_{n+1}} V(s,x_j)ds} \big(e^{-\frac{i\tau}{\varepsilon^2}\mathscr{T}^\varepsilon} I_M(\Phi^n)\big)(x_j), \quad j = 0, 1, ..., M-1, \tag{3.6.5}$$

with

$$\Phi_j^0 = \Phi(0,x_j), \quad j = 0, 1, ...M-1. \tag{3.6.6}$$

Moreover, we introduce the full-discretized error

$$\mathbf{e}_f^n(x) := P_M(\Phi(t_n,x)) - I_M(\Phi^n), \tag{3.6.7}$$

then the uniform and improved uniform error bounds for S_1 in Theorem 3.1 and Theorem 3.3 can be extended to full-discretization as follows

Theorem 3.5. *(i) Under the assumptions (A) and (B) with $2m+m_* \geq 2$, we have the following full-discretized error estimate for S_1*

$$\|\mathbf{e}_f^n(x)\|_{L^2} \lesssim \sqrt{\tau} + h^{2m+m_*}, \quad 0 \leq n \leq \frac{T}{\tau}. \tag{3.6.8}$$

(ii) If the time step size τ is non-resonant, i.e. there exists $0 < \kappa \leq 1$, such that $\tau \in \mathscr{A}_\kappa(\varepsilon)$, then under the assumptions (A) and (B) with $2m+m_* \geq 3$, we have an improved uniform error bound for S_1

$$\|\mathbf{e}_f^n(x)\|_{L^2} \lesssim_\kappa \tau + h^{2m+m_*}, \quad 0 \leq n \leq \frac{T}{\tau}. \tag{3.6.9}$$

Proof. (i) It is obvious that

$$\begin{aligned}\|\mathbf{e}_f^n(x)\|_{L^2} \quad \leq \quad & \|P_M(\Phi(t_n,x)) - \Phi(t_n,x)\|_{L^2} + \|\Phi(t_n,x) - \Phi^{[n]}(x)\|_{L^2} \\ & + \|\Phi^{[n]}(x) - I_M(\Phi^{[n]})(x)\|_{L^2} + \|I_M(\Phi^{[n]})(x) - I_M(\Phi^n)\|_{L^2}. \quad (3.6.10)\end{aligned}$$

From the regularity conditions, we have

$$\|P_M(\Phi(t_n,x)) - \Phi(t_n,x)\|_{L^2} \lesssim h^{2m+m_*}, \quad \|\Phi^{[n]}(x) - I_M(\Phi^{[n]})(x)\|_{L^2} \lesssim h^{2m+m_*}. \quad (3.6.11)$$

Moreover, Theorem 3.1 suggests

$$\|\Phi(t_n,x) - \Phi^{[n]}(x)\|_{L^2} \lesssim \sqrt{\tau}. \quad (3.6.12)$$

As a result, we only need to focus on the term $\|I_M(\Phi^{[n]})(x) - I_M(\Phi^n)\|_{L^2}$.

Notice that

$$\|\Phi^{[n]}(x) - I_M(\Phi^{[n]})(x)\|_{L^2} \lesssim h^{2m+m_*}, \quad (3.6.13)$$

then from (3.6.3), we have

$$\Phi^{[n+1]}(x) = e^{-i\int_{t_n}^{t_{n+1}} V(t,x)dt} e^{-\frac{i\tau}{\varepsilon^2}\mathscr{T}^\varepsilon} I_M(\Phi^{[n]})(x) + O(h^{2m+m_*}). \quad (3.6.14)$$

Subtracting from (3.6.5), and taking interpolation, we get

$$I_M(\Phi^{[n+1]})(x) - I_M(\Phi^{n+1}) = I_M\left(e^{-i\int_{t_n}^{t_{n+1}} V(t,x)dt} e^{-\frac{i\tau}{\varepsilon^2}\mathscr{T}^\varepsilon}\left(I_M(\Phi^{[n]})(x) - I_M(\Phi^n)\right)\right) + O(h^{2m+m_*}).$$

As $I_M(\Phi^{[0]})(x) - I_M(\Phi^0) = 0$, through recursion, we have

$$\|I_M(\Phi^{[n+1]})(x) - I_M(\Phi^{n+1})\|_{L^2} \lesssim h^{2m+m_*}. \quad (3.6.15)$$

Plugging into (3.6.10), taking into account (3.6.11) and (3.6.12), we finally get

$$\|\mathbf{e}_f^n(x)\|_{L^2} \lesssim \sqrt{\tau} + h^{2m+m_*}, \quad (3.6.16)$$

which completes the proof.

(ii) The proof for non-resonant time step is similar to the proof for (i). The details are omitted here for brevity. □

Next we consider the Strang splitting S_2. Similarly, denote $\Phi^{[n]}(x)$ to be the semi-discretized numerical solution from S_2 (3.2.5), and Φ^n to be the full-discretized numerical solution with Fourier spectral discretization in space, i.e. we have for $n = 0, 1, ..., \frac{T}{\tau} - 1$

$$\Phi^{[n+1]}(x) = e^{-\frac{i\tau}{2\varepsilon^2}\mathscr{T}^\varepsilon} e^{-i\int_{t_n}^{t_{n+1}} V(s,x)\,ds} e^{-\frac{i\tau}{2\varepsilon^2}\mathscr{T}^\varepsilon} \Phi^{[n]}(x), \quad x \in [a,b], \tag{3.6.17}$$

with

$$\Phi^{[0]}(x) = \Phi(0,x), \quad x \in [a,b], \tag{3.6.18}$$

and

$$\Phi_j^{n+1} = e^{-\frac{i\tau}{2\varepsilon^2}\mathscr{T}^\varepsilon} I_M\big(e^{-i\int_{t_n}^{t_{n+1}} V(s,x)\,ds} e^{-\frac{i\tau}{2\varepsilon^2}\mathscr{T}^\varepsilon} I_M(\Phi^n)\big)(x_j), \quad j = 0, 1, ..., M-1, \tag{3.6.19}$$

with

$$\Phi_j^0 = \Phi(0,x_j), \quad j = 0, 1, ...M-1. \tag{3.6.20}$$

The full-discretized error is still defined as (3.6.7), and then the uniform and improved uniform error bounds for S_2 in Theorem 3.2 and Theorem 3.4 can be extended to full-discretization as follows

Theorem 3.6. *(i) Under the assumptions (A) and (B) with $2m+m_* \geq 4$, we have the following full-discretized error estimate for S_2*

$$\|\mathbf{e}_f^n(x)\|_{L^2} \lesssim \sqrt{\tau} + h^{2m+m_*}, \quad 0 \leq n \leq \frac{T}{\tau}. \tag{3.6.21}$$

(ii) If the time step size τ is non-resonant, i.e. there exists $0 < \kappa \leq 1$, such that $\tau \in \mathscr{A}_\kappa(\varepsilon)$, under the assumptions (A) and (B) with $2m + m_* \geq 5$, with an extra regularity $V(t,x) \in W^{1,\infty}([0,T];H^3(\mathbb{R}))$ and then the following improved uniform error estimate for S_2 holds

$$\|\mathbf{e}_f^n(x)\|_{L^2} \lesssim_\kappa \tau^{3/2} + h^{2m+m_*}, \quad 0 \leq n \leq \frac{T}{\tau}. \tag{3.6.22}$$

Proof. The proof for Theorem 3.6 is similar to the proof for Theorem 3.5. We have the same inequality (3.6.10) and (3.6.11). The right hand side of (3.6.12) should still be $\sqrt{\tau}$ for all time step sizes, and would become $\tau^{3/2}$ for non-resonant time step sizes. The main task remains analyzing $\|I_M(\Phi^{[n]})(x) - I_M(\Phi^n)\|_{L^2}$.

Noticing the fact that

$$\|\Phi^{[n]}(x) - I_M(\Phi^{[n]})(x)\|_{L^2} \lesssim h^{2m+m_*}, \tag{3.6.23}$$

and furthermore

$$\|I_M\Big(e^{-i\int_{t_n}^{t_{n+1}}V(s,x)\,ds}e^{-\frac{i\tau}{2\varepsilon^2}\mathscr{T}^\varepsilon}I_M(\Phi^{[n]})\Big)(x)-e^{-i\int_{t_n}^{t_{n+1}}V(s,x)\,ds}e^{-\frac{i\tau}{2\varepsilon^2}\mathscr{T}^\varepsilon}\Phi^{[n]}(x)\|_{L^2}\lesssim h^{2m+m_*},$$

we have

$$\begin{aligned}I_M(\Phi^{[n+1]})(x)-I_M(\Phi^{n+1}) =& I_M\Bigg(e^{-\frac{i\tau}{2\varepsilon^2}\mathscr{T}^\varepsilon}I_M\Big(e^{-i\int_{t_n}^{t_{n+1}}V(s,x)\,ds}e^{-\frac{i\tau}{2\varepsilon^2}\mathscr{T}^\varepsilon}\big(I_M(\Phi^{[n]})(x)-I_M(\Phi^n)\big)\Big)\Bigg)\\ &+O(h^{2m+m_*})\end{aligned}$$

from (3.6.17) and (3.6.19). As $I_M(\Phi^{[0]})(x)-I_M(\Phi^0)=0$, through recursion, we have

$$\|I_M(\Phi^{[n+1]})(x)-I_M(\Phi^{n+1})\|_{L^2}\lesssim h^{2m+m_*}. \tag{3.6.24}$$

As a result, plugging into (3.6.10), we can get

$$\|\mathbf{e}_f^n(x)\|_{L^2}\lesssim \sqrt{\tau}+h^{2m+m_*}, \tag{3.6.25}$$

for all time step sizes and

$$\|\mathbf{e}_f^n(x)\|_{L^2}\lesssim_\kappa \tau^{3/2}+h^{2m+m_*}, \tag{3.6.26}$$

for non-resonant time step sizes, which completes the proof. □

Chapter 4

Uniform Error Bounds of Time-Splitting Methods for Nonlinear Dirac Equation

This chapter extends the super-resolution of time-splitting methods discussed in the previous chapter to the nonlinear Dirac equation. We still consider the equation in the absence of external magnetic potential in the nonrelativistic regime. Our numerical studies show similar results to the linear case, but the proofs are established in a different way because of the nonlinearity [18].

4.1 Introduction

In this chapter, we consider the splitting methods applied to the nonlinear Dirac equation [42, 50, 51, 58, 59, 60, 61, 64, 67, 73, 74, 78, 102] in the nonrelativistic regime in the absence of magnetic potential. In one or two dimension (1D or 2D), the equation with time-independent electric potential can be represented in the two-component form with wave function $\Phi := \Phi(t,\mathbf{x}) = (\phi_1(t,\mathbf{x}),\phi_2(t,\mathbf{x}))^T \in \mathbb{C}^2$ by taking $A_j \equiv 0$ ($j = 1,...,d$) in (1.3.9) as:

$$i\partial_t\Phi = \left(-\frac{i}{\varepsilon}\sum_{j=1}^{d}\sigma_j\partial_j + \frac{1}{\varepsilon^2}\sigma_3\right)\Phi + V(\mathbf{x})\Phi + \mathbf{F}(\Phi)\Phi, \quad \mathbf{x}\in\mathbb{R}^d, \quad d = 1,2, \quad t > 0, \tag{4.1.1}$$

where the initial value is given as

$$\Phi(t=0,x) = \Phi_0(\mathbf{x}), \quad \mathbf{x}\in\mathbb{R}^d, \quad d = 1,2. \tag{4.1.2}$$

and the nonlinearity $\mathbf{F}(\Phi)$ is chosen to be (λ_1, $\lambda_2 \in \mathbb{R}$)

$$\mathbf{F}(\Phi) = \lambda_1(\Phi^*\sigma_3\Phi)\sigma_3 + \lambda_2|\Phi|^2 I_2. \tag{4.1.3}$$

Though the TSFP method (also called S_2) has a τ^2/ε^4 dependence on the small parameter ε [16], under the specific case where there is a lack of magnetic potential, as in (4.1.1), we find out through our recent extensive numerical experiments that the errors of S_2 will be independent of ε and uniform w.r.t. ε. In other words, S_2 for NLDE (4.1.1) in the absence of magnetic potentials displays **super-resolution** w.r.t ε.

The super-resolution property for the time-splitting methods makes them superior in solving the NLDE in the absence of magnetic potentials in the nonrelativistic regime as they are more efficient and reliable compared to other numerical methods in the literature. In this chapter, the super-resolution for first-order (S_1) and second-order (S_2) time-splitting methods will be rigorously analyzed, and numerical results will be presented to validate the conclusions. We remark that similar results have been analyzed for the linear Dirac equation [17], where the linearity enables us to explicitly track the error exactly and make estimation at the target time step without the use of Gronwall type arguments. However, in the nonlinear case, it is impossible to follow the error propagation exactly and estimations have to be done at each time step. As a result, Gronwall arguments will be involved together with the mathematical induction to control the nonlinearity. In particular, instead of the previously adopted Lie calculus approach [92], Taylor expansion and Duhamel principle are employed to study the local error of the splitting methods, which can identify how temporal oscillations propagate numerically.

4.2 Semi-discretization

For simplicity of analysis, here we only consider (4.1.1) in 1D ($d = 1$). Extension to (4.1.1) in 2D and/or the four-component form with $d = 1, 2, 3$ is straightforward.

Denote the free Dirac Hermitian operator

$$\mathscr{T}^{\varepsilon} = -i\varepsilon\sigma_1\partial_x + \sigma_3, \quad x \in \mathbb{R}, \tag{4.2.1}$$

then the NLDE (4.1.1) in 1D can be written as

$$i\partial_t\Phi(t,x) = \frac{1}{\varepsilon^2}\mathscr{T}^{\varepsilon}\Phi(t,x) + V(x)\Phi(t,x) + \mathbf{F}(\Phi(t,x))\Phi(t,x), \quad x \in \mathbb{R}, \tag{4.2.2}$$

with nonlinearity (4.1.3) and the initial condition (4.1.2).

Choose $\tau > 0$ as the time step size and $t_n = n\tau$ for $n = 0, 1, ...$ as the time steps. Denote $\Phi^n(x)$ to be the numerical approximation of $\Phi(t_n, x)$, where $\Phi(t,x)$ is the exact solution to (4.2.2) with (4.1.3) and (4.1.2), then through applying the discrete-in-time first-order splitting (Lie-Trotter splitting) [123], S_1 can be represented as:

$$\Phi^{n+1}(x) = e^{-\frac{i\tau}{\varepsilon^2}\mathscr{T}^\varepsilon} e^{-i\tau[V(x)+\mathbf{F}(\Phi^n(x))]}\Phi^n(x), \quad x \in \mathbb{R}, \tag{4.2.3}$$

with $\Phi^0(x) = \Phi_0(x)$. For simplicity, we also write $\Phi^{n+1}(x) := S^{\text{Lie}}_{n,\tau}(\Phi^n)$, where $S^{\text{Lie}}_{n,\tau}$ denotes the numerical propagator of Lie-Trotter splitting.

Similarly, applying the discrete-in-time second-order splitting (Strang splitting, S_2) to (4.2.2), we have the numerical method as [113]

$$\Phi^{n+1}(x) = e^{-\frac{i\tau}{2\varepsilon^2}\mathscr{T}^\varepsilon} e^{-i\tau\left[V(x)+\mathbf{F}\left(e^{-\frac{i\tau}{2\varepsilon^2}\mathscr{T}^\varepsilon}\Phi^n(x)\right)\right]} e^{-\frac{i\tau}{2\varepsilon^2}\mathscr{T}^\varepsilon}\Phi^n(x). \tag{4.2.4}$$

with $\Phi^0(x) = \Phi_0(x)$. We write the numerical propagator for S_2 as $\Phi^{n+1}(x) := S^{\text{Str}}_{n,\tau}(\Phi^n)$.

4.3 Uniform error bounds

For any $0 < T < T^*$, where T^* denotes the maximal existence time of the solution for (4.1.1) with (4.1.2), we are going to consider smooth enough solutions, i.e. we assume the electric potential satisfies

$$(C) \qquad V(x) \in W^{2m+m_*,\infty}(\mathbb{R}), \quad m \in \mathbb{N}^*, m_* = 0, 1.$$

In addition, we assume the exact solution $\Phi(t,x)$ satisfies

$$(D) \qquad \Phi(t,x) \in L^\infty([0,T]; (H^{2m+m*}(\mathbb{R}))^2), \quad m \in \mathbb{N}^*, m_* = 0, 1.$$

For the numerical approximation $\Phi^n(x)$ obtained from S_1 (4.2.3), we introduce the error function

$$\mathbf{e}^n(x) = \Phi(t_n, x) - \Phi^n(x), \quad 0 \le n \le \frac{T}{\tau}, \tag{4.3.1}$$

then the following uniform error bound can be established.

Theorem 4.1. *Let $\Phi^n(x)$ be the numerical approximation obtained from S_1 (4.2.3), then under assumptions (C) and (D) with $m = m_* = 1$, there exists $0 < \tau_0 \leq 1$ independent of ε such that the following two error estimates hold for $0 < \tau < \tau_0$*

$$\|\mathbf{e}^n(x)\|_{H^1} \lesssim \tau + \varepsilon, \quad \|\mathbf{e}^n(x)\|_{H^1} \lesssim \tau + \tau/\varepsilon, \quad 0 \leq n \leq \frac{T}{\tau}. \tag{4.3.2}$$

Consequently, there is a uniform error bound for S_1 when $0 < \tau < \tau_0$

$$\|\mathbf{e}^n(x)\|_{H^1} \lesssim \tau + \max_{0<\varepsilon\leq 1} \min\{\varepsilon, \tau/\varepsilon\} \lesssim \sqrt{\tau}, \quad 0 \leq n \leq \frac{T}{\tau}. \tag{4.3.3}$$

For simplicity of the presentation, in the proof for this theorem and other theorems later for NLDE, we will take $V(x) \equiv 0$. Extension to the case where $V(x) \neq 0$ or time-dependent is straightforward [17]. Compared to the linear case [17], the nonlinear term is much more complicated to analyze. A key issue of the error analysis for NLDE is to control the nonlinear term of numerical solution Φ^n, and for which we require the following stability lemma [92].

Lemma 4.1. *Suppose $V(t_n,x) \in W^{1,\infty}(\mathbb{R})$, and $\Phi(x), \Psi(x) \in (H^1(\mathbb{R}))^2$ satisfy $\|\Phi\|_{H^1}, \|\Psi\|_{H^1} \leq M$, we have*

$$\|S^{\text{Lie}}_{n,\tau}(\Phi) - S^{\text{Lie}}_{n,\tau}(\Psi)\|_{H^1} \leq e^{c_1\tau}\|\Phi - \Psi\|_{H^1}, \tag{4.3.4}$$

where c_1 depends on M and $\|V(x)\|_{W^{1,\infty}}$.

Proof. The proof is quite similar to the nonlinear Schrödinger equation case in [92] and we omit it here for brevity. □

Under the assumption (D) ($m \geq 1$), for $\varepsilon \in (0,1]$, we denote $M_1 > 0$ as

$$M_1 = \sup_{\varepsilon\in(0,1]} \|\Phi(t,x)\|_{L^\infty([0,T];(H^1(\mathbb{R}))^2)}. \tag{4.3.5}$$

Based on (4.3.5) and Lemma 4.1, one can control the the nonlinear term once the hypothesis of the lemma is fulfilled. Making use of the fact that S_1 is explicit, together with the uniform error estimates in Theorem 4.1, we can use mathematical induction to complete the proof.

The following properties of $\mathscr{T}^\varepsilon$ will be frequently used in the analysis. $\mathscr{T}^\varepsilon$ is diagonalizable in the phase space (Fourier domain) and can be decomposed as

$$\mathscr{T}^\varepsilon = \sqrt{Id - \varepsilon^2\Delta}\Pi^\varepsilon_+ - \sqrt{Id - \varepsilon^2\Delta}\Pi^\varepsilon_-, \tag{4.3.6}$$

where $\Delta = \partial_{xx}$ is the Laplace operator in 1D, Id is the identity operator, and Π_+^ε, Π_-^ε are projectors defined as

$$\Pi_+^\varepsilon = \frac{1}{2}\left[Id + (Id-\varepsilon^2\Delta)^{-1/2}\mathscr{T}^\varepsilon\right], \quad \Pi_-^\varepsilon = \frac{1}{2}\left[Id - (Id-\varepsilon^2\Delta)^{-1/2}\mathscr{T}^\varepsilon\right]. \tag{4.3.7}$$

It is straightforward to verify that $\Pi_+^\varepsilon + \Pi_-^\varepsilon = Id$, $\Pi_+^\varepsilon\Pi_-^\varepsilon = \Pi_-^\varepsilon\Pi_+^\varepsilon = 0$, $(\Pi_\pm)^2 = \Pi_\pm$, and through Taylor expansion, we have [29]

$$\Pi_\pm^\varepsilon = \Pi_\pm^0 \pm \varepsilon\mathscr{R}_1 = \Pi_\pm^0 + \mp i\frac{\varepsilon}{2}\sigma_1\partial_x \pm \varepsilon^2\mathscr{R}_2, \Pi_+^0 = \mathrm{diag}(1,0), \Pi_-^0 = \mathrm{diag}(0,1) \tag{4.3.8}$$

with $\mathscr{R}_1 : (H^m(\mathbb{R}))^2 \to (H^{m-1}(\mathbb{R}))^2$ for $m \geq 1$, $\mathscr{R}_2 : (H^m(\mathbb{R}))^2 \to (H^{m-2}(R))^2$ for $m \geq 2$ being uniformly bounded operators w.r.t. ε.

In order to characterize the oscillatory features of the solution, denote

$$\mathscr{D}^\varepsilon = \frac{1}{\varepsilon^2}(\sqrt{Id-\varepsilon^2\Delta} - Id) = -(\sqrt{Id-\varepsilon^2\Delta} + Id)^{-1}\Delta, \tag{4.3.9}$$

which is a uniformly bounded operator w.r.t ε from $(H^m(\mathbb{R}))^2 \to (H^{m-2}(\mathbb{R}))^2$ for $m \geq 2$, then the evolution operator $e^{\frac{it}{\varepsilon^2}\mathscr{T}^\varepsilon}$ can be expressed as

$$e^{\frac{it}{\varepsilon^2}\mathscr{T}^\varepsilon} = e^{\frac{it}{\varepsilon^2}(\sqrt{Id-\varepsilon^2\Delta}\Pi_+^\varepsilon - \sqrt{Id-\varepsilon^2\Delta}\Pi_-^\varepsilon)} = e^{\frac{it}{\varepsilon^2}}e^{it\mathscr{D}^\varepsilon}\Pi_+^\varepsilon + e^{-\frac{it}{\varepsilon^2}}e^{-it\mathscr{D}^\varepsilon}\Pi_-^\varepsilon. \tag{4.3.10}$$

For simplicity, here we use $\Phi(t) := \Phi(t,x)$, $\Phi^n := \Phi^n(x)$ in short.

Now we are ready to introduce the following lemma for proving Theorem 4.1.

Lemma 4.2. *Let $\Phi^n(x)$ ($0 \leq n \leq \frac{T}{\tau} - 1$) be obtained from S_1 (4.2.3) satisfying $\|\Phi^n(x)\|_{H^1} \leq M_1 + 1$, under the assumptions of Theorem 4.1, we have*

$$\mathbf{e}^{n+1}(x) = e^{-\frac{i\tau}{\varepsilon^2}\mathscr{T}^\varepsilon}e^{-i\tau\mathbf{F}(\Phi^n)}\mathbf{e}^n(x) + \eta_1^n(x) + e^{-\frac{i\tau}{\varepsilon^2}\mathscr{T}^\varepsilon}\eta_2^n(x), \tag{4.3.11}$$

with $\|\eta_1^n(x)\|_{H^1} \leq c_1\tau^2 + c_2\tau\|\mathbf{e}^n(x)\|_{H^1}$, $\eta_2^n(x) = \int_0^\tau f_2^n(s)ds - \tau f_2^n(0)$, where c_1 depends on M_1, λ_1, λ_2 and $\|\Phi(t,x)\|_{L^\infty([0,T];(H^3)^2)}$; c_2 depends on M_1, λ_1, and λ_2;

$$\begin{aligned} f_2^n(s) = &- ie^{\frac{-4is}{\varepsilon^2}}\Pi_-^\varepsilon\left(\mathbf{g}_1^n(x)\Pi_+^\varepsilon\Phi(t_n)\right) - ie^{\frac{4is}{\varepsilon^2}}\Pi_+^\varepsilon\left(\overline{\mathbf{g}_1^n}(x)\Pi_-^\varepsilon\Phi(t_n)\right) \\ &- ie^{\frac{-i2s}{\varepsilon^2}}\left[\Pi_+^\varepsilon\left(\mathbf{g}_1^n(x)\Pi_+^\varepsilon\Phi(t_n)\right) + \Pi_-^\varepsilon\left(\mathbf{g}_2^n(x)\Pi_+^\varepsilon\Phi(t_n) + \mathbf{g}_1^n(x)\Pi_-^\varepsilon\Phi(t_n)\right)\right] \\ &- ie^{\frac{2is}{\varepsilon^2}}\left[\Pi_-^\varepsilon\left(\overline{\mathbf{g}_1^n}(x)\Pi_-^\varepsilon\Phi(t_n)\right) + \Pi_+^\varepsilon\left(\mathbf{g}_2^n(x)\Pi_-^\varepsilon\Phi(t_n) + \overline{\mathbf{g}_1^n}(x)\Pi_+^\varepsilon\Phi(t_n)\right)\right], \end{aligned} \tag{4.3.12}$$

where $\mathbf{g}_j^n(x) = \mathbf{g}_j(\Phi_+(t_n), \Phi_-(t_n))$ *with* $\Phi_\pm(t_n) = \Pi_\pm^\varepsilon \Phi(t_n)$ *and*

$$\mathbf{g}_1(\Phi_+(t_n), \Phi_-(t_n)) = \lambda_1 \left((\Phi_-(t_n))^* \sigma_3 \Phi_+(t_n)\right) \sigma_3 + \lambda_2 \left((\Phi_-(t_n))^* \Phi_+(t_n)\right) I_2, \tag{4.3.13}$$

$$\mathbf{g}_2(\Phi_+(t_n), \Phi_-(t_n)) = \sum_{\sigma=\pm} \left[\lambda_1 ((\Phi_\sigma(t_n))^* \sigma_3 \Phi_\sigma(t_n)) \sigma_3 + \lambda_2 |\Phi_\sigma(t_n)|^2 I_2\right] \tag{4.3.14}$$

Proof. Through the definition of $\mathbf{e}^n(x)$ (4.3.1), noticing the formula (4.2.3), we have

$$\mathbf{e}^{n+1}(x) = e^{-\frac{i\tau}{\varepsilon^2}\mathscr{T}^\varepsilon} e^{-i\tau \mathbf{F}(\Phi^n)} \mathbf{e}^n(x) + \eta^n(x), \quad 0 \le n \le \frac{T}{\tau} - 1, \quad x \in \mathbb{R}, \tag{4.3.15}$$

where $\eta^n(x)$ is the "local truncation error" (notice that this is not the usual local truncation error, compared with $\Phi(t_{n+1}, x) - S^{\mathrm{Lie}}_{n,\tau}\Phi(t_n, x)$),

$$\eta^n(x) = \Phi(t_{n+1}, x) - e^{-\frac{i\tau}{\varepsilon^2}\mathscr{T}^\varepsilon} e^{-i\tau \mathbf{F}(\Phi^n)} \Phi(t_n, x), \quad x \in \mathbb{R}. \tag{4.3.16}$$

By Duhamel's principle, the solution $\Phi(t, x)$ to (4.2.2) satisfies

$$\Phi(t_n + s, x) = e^{-\frac{is}{\varepsilon^2}\mathscr{T}^\varepsilon} \Phi(t_n, x) - i \int_0^s e^{-\frac{i(s-w)}{\varepsilon^2}\mathscr{T}^\varepsilon} \mathbf{F}(\Phi(t_n + w, x)) \Phi(t_n + w, x) dw, \quad 0 \le s \le \tau, \tag{4.3.17}$$

which implies that $\|\Phi(t_n + s, x) - e^{-\frac{is}{\varepsilon^2}\mathscr{T}^\varepsilon} \Phi(t_n, x)\|_{H^1} \lesssim \tau$ ($s \in [0, \tau]$). Setting $s = \tau$ in (4.3.17), we have from (4.3.16),

$$\eta^n(x) = e^{-\frac{i\tau}{\varepsilon^2}\mathscr{T}^\varepsilon} \left(\int_0^\tau f^n(s) ds - \tau f^n(0) \right) + R_1^n(x) + R_2^n(x), \tag{4.3.18}$$

where

$$f^n(s) = -ie^{\frac{is}{\varepsilon^2}\mathscr{T}^\varepsilon} \left(\mathbf{F}(e^{-\frac{is}{\varepsilon^2}\mathscr{T}^\varepsilon} \Phi(t_n)) e^{-\frac{is}{\varepsilon^2}\mathscr{T}^\varepsilon} \Phi(t_n, x) \right), \quad R_1^n(x) = e^{-\frac{i\tau}{\varepsilon^2}\mathscr{T}^\varepsilon} \left(\Lambda_1^n(x) + \Lambda_2^n(x) \right) \tag{4.3.19}$$

$$R_2^n(x) = -i \int_0^\tau e^{-\frac{i(\tau-s)}{\varepsilon^2}\mathscr{T}^\varepsilon} \left[\mathbf{F}(\Phi(t_n + s)) \Phi(t_n + s) - \mathbf{F}(\Phi(e^{-\frac{is}{\varepsilon^2}\mathscr{T}^\varepsilon} \Phi(t_n))) e^{-\frac{is}{\varepsilon^2}\mathscr{T}^\varepsilon} \Phi(t_n) \right] ds, \tag{4.3.20}$$

with

$$\Lambda_1^n(x) = -\left(e^{-i\tau \mathbf{F}(\Phi^n)} - (I_2 - i\tau \mathbf{F}(\Phi^n)) \right) \Phi(t_n), \quad \Lambda_2^n(x) = \left(i\tau \left(\mathbf{F}(\Phi^n) - \mathbf{F}(\Phi(t_n)) \right) \right) \Phi(t_n). \tag{4.3.21}$$

Noticing (4.3.17), the assumption that $\|\Phi^n\|_{H^1} \leq M_1+1$, and the fact that $e^{-is\mathscr{T}^\varepsilon/\varepsilon^2}$ preserves H^k norm, it is not difficult to find

$$\|R_2^n(x)\|_{H^1} \lesssim (M_1+1)^2 \int_0^\tau \|\Phi(t_n+s,x) - e^{-\frac{is}{\varepsilon^2}\mathscr{T}^\varepsilon}\Phi(t_n,x)\|_{H^1}\, ds \lesssim \tau^2 \tag{4.3.22}$$

On the other hand, using Taylor expansion in $\Lambda_1^n(x)$ and the local H^1 Lipschitz property of $\mathbf{F}$, we get

$$\|R_1^n(x)\|_{H^1} \lesssim \tau^2\|\Phi^n\|_{H^1}^2\|\Phi(t_n)\|_{H^1} + \tau(M_1+1)^2\|\Phi^n-\Phi(t_n)\|_{H^1} \lesssim \tau^2 + \tau\|\mathbf{e}^n(x)\|_{H^1}. \tag{4.3.23}$$

It remains to estimate the $f^n(s)$ part. Using the decomposition (4.3.10) and the Taylor exapnsion $e^{i\tau\mathscr{D}^\varepsilon} = Id + O(\tau\mathscr{D}^\varepsilon)$, we have $e^{\frac{-is\mathscr{T}^\varepsilon}{\varepsilon^2}}\Phi(t_n) = e^{\frac{-is}{\varepsilon^2}}\Phi_+(t_n) + e^{\frac{is}{\varepsilon^2}}\Phi_-(t_n) + O(\tau)$ ($\Phi_\pm(t_n) = \Pi_\pm^\varepsilon\Phi(t_n)$),

$$f^n(s) = -i\sum_{\sigma=\pm} e^{\frac{\sigma is}{\varepsilon^2}}\Pi_\sigma^\varepsilon\left\{\mathbf{F}\left(e^{\frac{-is}{\varepsilon^2}}\Phi_+(t_n)+e^{\frac{is}{\varepsilon^2}}\Phi_-(t_n)\right)\left(e^{\frac{-is}{\varepsilon^2}}\Phi_+(t_n)+e^{\frac{is}{\varepsilon^2}}\Phi_-(t_n)\right)\right\} + f_1^n(s), \tag{4.3.24}$$

where for $s\in[0,\tau]$,

$$\|f_1^n(s)\|_{H^1} \lesssim \tau\|\Phi(t_n)\|_{H^3}^3 \lesssim \tau. \tag{4.3.25}$$

Since $\mathbf{F}$ is of polynomial type, by direct computation, we can further simplify (4.3.24) to get

$$f^n(s) = f_1^n(s) + f_2^n(s) + \tilde{f}^n(s), \quad 0\leq s\leq\tau, \tag{4.3.26}$$

where $f_2^n(s)$ is given in (4.3.12) and $\tilde{f}^n(s)$ is independent of s as

$$\tilde{f}^n(s) \equiv -i\left[\Pi_+^\varepsilon\left(\mathbf{g}_2^n(x)\Pi_+^\varepsilon\Phi(t_n)+\mathbf{g}_1^n(x)\Pi_-^\varepsilon\Phi(t_n)\right)+\Pi_-^\varepsilon\left(\mathbf{g}_2^n(x)\Pi_-^\varepsilon\Phi(t_n)+\overline{\mathbf{g}_1^n}(x)\Pi_+^\varepsilon\Phi(t_n)\right)\right] \tag{4.3.27}$$

with $\mathbf{g}_{1,2}^n$ defined in (4.3.13)-(4.3.14).

Now, it is easy to verify that $\eta^n(x) = \eta_1^n(x)+\eta_2^n(x)$ with $\eta_2^n(x)$ given in Lemma 4.2 by choosing

$$\eta_1^n(x) = e^{-\frac{i\tau}{\varepsilon^2}\mathscr{T}^\varepsilon}\left(\int_0^\tau (f_1^n(s)+\tilde{f}^n(s))ds - \tau(f_1^n(0)+\tilde{f}^n(0))\right) + R_1^n(x) + R_2^n(x) \tag{4.3.28}$$

Noticing that $\tilde{f}^n(s)$ is independent of s and $\|f_1^n(s)\|_{H^1}\lesssim\tau$, combining (4.3.22) and (4.3.23), we can get

$$\|\eta_1^n(x)\|_{H^1} \leq \sum_{j=1}^2\|R_j^n(x)\|_{H^1} + \left\|\int_0^\tau f_1^n(s)ds - \tau f_1^n(0)\right\|_{H^1} \lesssim \tau\|\mathbf{e}^n(x)\|_{H^1} + \tau^2,$$

which completes the proof of Lemma 4.2. □

Now, we proceed to prove Theorem 4.1.

Proof. We will prove by induction that the estimates (4.3.2)-(4.3.3) hold for all time steps $n \le \frac{T}{\tau}$ together with

$$\|\Phi^n\|_{H^1} \le M_1 + 1. \tag{4.3.29}$$

Since initially $\Phi^0 = \Phi_0(x)$, $n = 0$ case is obvious. Assume (4.3.2)-(4.3.3) and (4.3.29) hold true for all $0 \le n \le m \le \frac{T}{\tau} - 1$, then we are going to prove the case $n = m+1$.

From Lemma 4.2, we have

$$\mathbf{e}^{n+1}(x) = e^{-\frac{i\tau}{\varepsilon^2}\mathscr{T}^\varepsilon} e^{-i\tau\mathbf{F}(\Phi^n)}\mathbf{e}^n(x) + \eta_1^n(x) + e^{-\frac{i\tau}{\varepsilon^2}\mathscr{T}^\varepsilon}\eta_2^n(x), \quad 0 \le n \le m, \tag{4.3.30}$$

with $\|\eta_1^n(x)\|_{H^1} \lesssim \tau^2 + \tau\|\mathbf{e}^n(x)\|_{H^1}$, $\mathbf{e}^0 = 0$ and $\eta_2^n(x)$ given in Lemma 4.2.

Denote $\mathscr{L}_n = e^{-\frac{i\tau}{\varepsilon^2}\mathscr{T}^\varepsilon}\left(e^{-i\tau\mathbf{F}(\Phi^n)} - I_2\right)$ $(0 \le n \le m \le \frac{T}{\tau} - 1)$, and it is straightforward to calculate

$$\|\mathscr{L}_n\Psi(x)\|_{H^1} \le C_{M_1}\tau\|\Psi\|_{H^1}, \quad \forall \Psi \in (H^1(\mathbb{R}))^2, \tag{4.3.31}$$

with C_{M_1} only depending on M_1. Thus we can obtain from (4.3.30) that for $0 \le n \le m$,

$$\begin{aligned}
\mathbf{e}^{n+1}(x) &= e^{-\frac{i\tau}{\varepsilon^2}\mathscr{T}^\varepsilon}\mathbf{e}^n(x) + \eta_1^n(x) + e^{-\frac{i\tau}{\varepsilon^2}\mathscr{T}^\varepsilon}\eta_2^n(x) + \mathscr{L}_n\mathbf{e}^n(x) \\
&= e^{-\frac{2i\tau}{\varepsilon^2}\mathscr{T}^\varepsilon}\mathbf{e}^{n-1}(x) + e^{-\frac{i\tau}{\varepsilon^2}\mathscr{T}^\varepsilon}\left(\eta_1^{n-1}(x) + e^{-\frac{i\tau}{\varepsilon^2}\mathscr{T}^\varepsilon}\eta_2^{n-1}(x) + \mathscr{L}_{n-1}\mathbf{e}^{n-1}\right) \\
&\qquad + \left(\eta_1^n(x) + e^{-\frac{i\tau}{\varepsilon^2}\mathscr{T}^\varepsilon}\eta_2^n(x) + \mathscr{L}_n\mathbf{e}^n\right) \\
&= \ldots \\
&= e^{-i(n+1)\tau\mathscr{T}^\varepsilon/\varepsilon^2}\mathbf{e}^0(x) + \sum_{k=0}^{n} e^{-\frac{i(n-k)\tau}{\varepsilon^2}\mathscr{T}^\varepsilon}\left(\eta_1^k(x) + e^{-\frac{i\tau}{\varepsilon^2}\mathscr{T}^\varepsilon}\eta_2^k(x) + \mathscr{L}_k\mathbf{e}^k(x)\right).
\end{aligned} \tag{4.3.32}$$

Since $\|\eta_1^k(x)\|_{H^1} \lesssim \tau^2 + \tau\|\mathbf{e}^n(x)\|_{H^1}$, $k = 0, 1, \ldots, n$, and $e^{-is/\varepsilon^2\mathscr{T}^\varepsilon}$ $(s \in \mathbb{R})$ preserves H^1 norm, we have from (4.3.31)

$$\begin{aligned}
\left\|\sum_{k=0}^{n} e^{-\frac{i(n-k)\tau}{\varepsilon^2}\mathscr{T}^\varepsilon}\left(\eta_1^k(x) + \mathscr{L}_k\mathbf{e}^k\right)\right\|_{H^1} &\lesssim \sum_{k=0}^{n}\tau^2 + \sum_{k=0}^{n}\tau\|\mathbf{e}^k(x)\|_{H^1} \\
&\lesssim \tau + \tau\sum_{k=0}^{n}\|\mathbf{e}^k(x)\|_{H^1},
\end{aligned} \tag{4.3.33}$$

which leads to

$$\|\mathbf{e}^{n+1}(x)\|_{H^1} \lesssim \tau + \tau \sum_{k=0}^{n} \|\mathbf{e}^{k}(x)\|_{H^1} + \left\| \sum_{k=0}^{n} e^{-\frac{i(n-k+1)\tau}{\varepsilon^2}\mathscr{T}^{\varepsilon}} \eta_2^k(x) \right\|_{H^1}, \quad n \le m. \tag{4.3.34}$$

To analyze $\eta_2^n(x) = \int_0^\tau f_2^n(s)ds - \tau f_2^n(0)$, using (4.3.8), we can find $f_2^n(s) = O(\varepsilon)$, e.g.

$$(\Pi_+^{\varepsilon}\Phi(t_n))^*\sigma_3(\Pi_-^{\varepsilon}\Phi(t_n)) = -\varepsilon(\Pi_+^{\varepsilon}\Phi(t_n))^*\sigma_3(\mathscr{R}_1\Phi(t_n)) + \varepsilon(\mathscr{R}_1\Phi(t_n))^*\sigma_3(\Pi_-^{\varepsilon}\Phi(t_n)),$$

and the other terms in $f_2^n(s)$ can be estimated similarly. As $\mathscr{R}_1 : (H^m)^2 \to (H^{m-1})^2$ is uniformly bounded with respect to $\varepsilon \in (0,1]$, we have (with detailed computations omitted)

$$\|f_2^n(\cdot)\|_{L^\infty([0,\tau];(H^1)^2)} \lesssim \varepsilon \|\Phi(t_n)\|_{H^2}^3 \lesssim \varepsilon. \tag{4.3.35}$$

Noticing the assumptions of Theorem 4.1, we obtain from (4.3.12)

$$\|f_2^n(\cdot)\|_{L^\infty([0,\tau];(H^1)^2)} \lesssim \varepsilon, \quad \|\partial_s(f_2^n)(\cdot)\|_{L^\infty([0,\tau];(H^1)^2)} \lesssim \varepsilon/\varepsilon^2 = 1/\varepsilon, \tag{4.3.36}$$

which leads to

$$\left\| \int_0^\tau f_2^n(s)\,ds - \tau f_2^n(0) \right\|_{H^1} \lesssim \tau\varepsilon. \tag{4.3.37}$$

On the other hand, using Taylor expansion and the second inequality in (4.3.36), we have

$$\left\| \int_0^\tau f_2^n(s)\,ds - \tau f_2^n(0) \right\|_{H^1} \le \frac{\tau^2}{2} \|\partial_s f_2^n(\cdot)\|_{L^\infty([0,\tau];(H^1)^2)} \lesssim \tau^2/\varepsilon. \tag{4.3.38}$$

Combining (4.3.37) and (4.3.38), we arrive at

$$\|\eta_2^n(x)\|_{H^1} \lesssim \min\{\tau\varepsilon, \tau^2/\varepsilon\}. \tag{4.3.39}$$

Then from (4.3.34), we get for $n \le m$

$$\|\mathbf{e}^{n+1}(x)\|_{H^1} \lesssim n\tau^2 + n\min\{\tau\varepsilon, \tau^2/\varepsilon\} + \tau \sum_{k=0}^{n} \|\mathbf{e}^{n}(x)\|_{H^1}. \tag{4.3.40}$$

Using discrete Gronwall's inequality, we have

$$\|\mathbf{e}^{n+1}(x)\|_{H^1} \lesssim \tau + \min\{\varepsilon, \tau/\varepsilon\}, \quad n \le m, \tag{4.3.41}$$

which shows that (4.3.2)-(4.3.3) hold for $n = m+1$. It can be checked that all the constants appearing in the estimates depend only on $M_1, \lambda_1, \lambda_2, T$ and $\|\Phi(t,x)\|_{L^\infty([0,T];(H^3)^2)}$, and

$$\|\Phi^{m+1}\|_{H^1} \le \|\Phi(t_{m+1})\|_{H^1} + \|\mathbf{e}^{m+1}\|_{H^1} \le M_1 + C\sqrt{\tau} \tag{4.3.42}$$

for some $C = C(M_1, \lambda_1, \lambda_2, T, \|\Phi(t,x)\|_{L^\infty([0,T];(H^3)^2)})$. Choosing $\tau \le \frac{1}{C^2}$ will justify (4.3.29) at $n = m+1$, which finishes the induction process, and the proof for Theorem 4.1 is completed.

□

Remark 4.1. *In Theorem 4.1 and the other results in this chapter for the 1D case, we prove the H^1 error bounds for $\mathbf{e}^n(x)$ due to the fact that $H^1(\mathbb{R})$ is an algebra, and the corresponding estimates should be in H^2 norm for 2D and 3D cases with of course higher regularity assumptions.*

For the numerical approximation $\Phi^n(x)$ obtained from S_2 (4.2.4), we introduce the error function as in (4.3.1), and the following uniform error bounds hold.

Theorem 4.2. *Let $\Phi^n(x)$ be the numerical approximation obtained from S_2* (4.2.4)*, then under the assumptions (C) and (D) with $m = 2$, $m_* = 1$, there exists $0 < \tau_0 \le 1$ independent of ε such that the following error estimates hold for $0 < \tau < \tau_0$,*

$$\|\mathbf{e}^n(x)\|_{H^1} \lesssim \tau^2 + \varepsilon, \quad \|\mathbf{e}^n(x)\|_{H^1} \lesssim \tau^2 + \tau^2/\varepsilon^3, \quad 0 \le n \le \frac{T}{\tau}. \tag{4.3.43}$$

As a result, there is a uniform error bound for S_2 for $\tau > 0$ small enough

$$\|\mathbf{e}^n(x)\|_{H^1} \lesssim \tau^2 + \max_{0<\varepsilon\le 1} \min\{\varepsilon, \tau^2/\varepsilon^3\} \lesssim \sqrt{\tau}, \quad 0 \le n \le \frac{T}{\tau}. \tag{4.3.44}$$

Proof. As the proof of the theorem is not difficult to establish through combining the techniques used in proving Theorem 4.1 and the ideas in the proof of the uniform error bounds for S_2 in the linear case [17], we only give the outline of the proof here. For simplicity, we denote $\Phi(t) := \Phi(t,x)$, $\Phi^n := \Phi^n(x)$ in short. Similar to the S_1 case, the H^1 bound of the numerical solution Φ^n is needed and can be done by using mathematical induction. For simplicity, we will assume the H^1 bound of Φ^n as in (4.3.29).

Step 1. Use Taylor expansion and Duhamel's principle repeatedly to represent the 'local truncation error' $\eta^n(x) = \Phi(t_{n+1}) - e^{-\frac{i\tau}{2\varepsilon^2}\mathscr{T}^\varepsilon} e^{-i\tau\mathbf{F}(e^{-\frac{i\tau}{2\varepsilon^2}\mathscr{T}^\varepsilon}\Phi^n)} e^{-\frac{i\tau}{2\varepsilon^2}\mathscr{T}^\varepsilon}\Phi(t_n)$ [16, 92] as

$$\eta^n(x) = e^{-\frac{i\tau}{\varepsilon^2}\mathscr{T}^\varepsilon}\left[\int_0^\tau (f^n(s)+h^n(s))\,ds - \tau f^n\left(\frac{\tau}{2}\right) - \int_0^\tau\int_0^s g^n(s,w)\,dwds + \frac{\tau^2}{2}g^n\left(\frac{\tau}{2},\frac{\tau}{2}\right)\right] + R^n(x),$$

where $\|R^n(x)\|_{H^1} \lesssim \tau^3 + \tau\|\mathbf{e}^n(x)\|_{H^1}$, $f^n(s)$ is the same as that in Lie splitting S_1 case (4.3.19) and

$$h^n(s) = -ie^{\frac{is}{\varepsilon^2}\mathscr{T}^\varepsilon}\left[\left(\mathbf{F}\left(\Phi(t_n+s)\right) - \mathbf{F}\left(e^{-\frac{is}{\varepsilon^2}\mathscr{T}^\varepsilon}\Phi(t_n)\right)\right)e^{-\frac{is}{\varepsilon^2}\mathscr{T}^\varepsilon}\Phi(t_n)\right], \quad 0 \le s \le \tau, \tag{4.3.45}$$

$$g^n(s,w) = e^{\frac{is}{\varepsilon^2}\mathscr{T}^\varepsilon}\left(\mathbf{F}(e^{-\frac{is}{\varepsilon^2}\mathscr{T}^\varepsilon}\Phi(t_n))e^{-\frac{i(s-w)}{\varepsilon^2}\mathscr{T}^\varepsilon}\left(\mathbf{F}(e^{-\frac{is}{\varepsilon^2}\mathscr{T}^\varepsilon}\Phi(t_n))e^{-\frac{iw}{\varepsilon^2}\mathscr{T}^\varepsilon}\Phi(t_n)\right)\right), \ 0 \le s,w \le \tau. \tag{4.3.46}$$

Step 2. For $h^n(s)$, using Duhamel's principle to get

$$\begin{aligned}\Phi(t_n+s) =& e^{-\frac{is}{\varepsilon^2}\mathscr{T}^\varepsilon}\Phi(t_n) - ie^{-\frac{is}{\varepsilon^2}\mathscr{T}^\varepsilon}\int_0^s f^n(w)\,dw + O(s^2) \\ =& \phi^n(s) - is\mathbf{F}(\phi^n(s))\phi^n(s) - \hat{f}^n(s) + O(s^2),\end{aligned} \tag{4.3.47}$$

where $\phi^n(s) = e^{-\frac{is}{\varepsilon^2}\mathscr{T}^\varepsilon}\Phi(t_n)$, $\hat{f}^n(s) = ie^{-\frac{is}{\varepsilon^2}\mathscr{T}^\varepsilon}\int_0^s (f^n(w) - f^n(s))\,dw$, and we could find

$$\mathbf{F}\left(\Phi(t_n+s)\right) - \mathbf{F}\left(e^{-\frac{is}{\varepsilon^2}\mathscr{T}^\varepsilon}\Phi(t_n)\right) = -2\lambda_1\mathrm{Re}\left((\phi^n(s))^*\sigma_3\hat{f}^n(s)\right)\sigma_3 - 2\lambda_2\mathrm{Re}\left((\phi^n(s))^*\hat{f}^n(s)\right)I_2 + O(s^2).$$

Recalling $\hat{f}^n(s) = O(s)$ and (4.3.24), we get $f^n(s) - f^n(w) = f_2^n(s) - f_2^n(w) + O(s)$ with $f_2^n(s)$ given in (4.3.12). Finally, under the assumption of Theorem 4.2, expanding $e^{-\frac{is\mathscr{T}^\varepsilon}{\varepsilon^2}}\Phi(t_n) = e^{-\frac{is}{\varepsilon^2}}\Phi_+(t_n) + e^{\frac{is}{\varepsilon^2}}\Phi_-(t_n) + O(s)$ ($\Phi_\pm(t_n) = \Pi_\pm^\varepsilon\Phi(t_n)$), we can write the $h^n(s)$ term as

$$\int_0^\tau h^n(s)\,ds = \zeta_1^n(x) + \kappa_1^n(x), \quad \|\kappa_1^n(x)\|_{H^1} \lesssim \tau^3, \tag{4.3.48}$$

with $\zeta_1^n(x)$ as the simplification of

$$2i\int_0^\tau e^{\frac{is}{\varepsilon^2}\mathscr{T}^\varepsilon}\left[\left(\lambda_1\mathrm{Re}\left((\phi^n(s))^*\sigma_3\hat{f}^n(s)\right)\sigma_3 + \lambda_2\mathrm{Re}\left((\phi^n(s))^*\hat{f}^n(s)\right)I_2\right)\phi^n(s)\right]ds \tag{4.3.49}$$

by taking $e^{-\frac{is\mathscr{T}^\varepsilon}{\varepsilon^2}} \approx e^{-\frac{is}{\varepsilon^2}}\Pi_+ + e^{\frac{is}{\varepsilon^2}}\Pi_-$, and it can be proved that $\|\zeta_1^n(x)\|_{H^1} \lesssim \min\{\tau^2\varepsilon, \frac{\tau^3}{\varepsilon}\}$.

Similarly, $g^n(s,w)$ can be written as

$$g^n(s,w) = \mathscr{G}_1^n(s,w) + \mathscr{G}_2^n(s,w) + \mathscr{G}_3^n(s,w), \tag{4.3.50}$$

where $\|\mathscr{G}_3^n(s,w)\|_{H^1} \lesssim \tau$, the oscillatory term (in time) $\mathscr{G}_1^n(s,w)$ simplifies $g^n(s,w)$ by using $e^{-\frac{is\mathscr{T}^\varepsilon}{\varepsilon^2}} \approx e^{-\frac{is}{\varepsilon^2}}\Pi_+ + e^{\frac{is}{\varepsilon^2}}\Pi_-$ and removing the non-oscillatory terms as in (4.3.27), $\mathscr{G}_2^n(s,w) =$

$\mathscr{G}_2^n(0,0)$ is the non-oscillatory term (s,w independent) similar to (4.3.27), $\|\mathscr{G}_1^n(s,w)\|_{H^1} \lesssim \varepsilon$. We can prove $\|\partial_s \mathscr{G}_1^n(s,w)\|_{H^1} \lesssim 1/\varepsilon$, $\|\partial_w \mathscr{G}_1^n(s,w)\|_{H^1} \lesssim 1/\varepsilon$.

Lastly, $f^n(s)$ can be decomposed as

$$f^n(s) = \mathscr{F}_1^n(s) + \mathscr{F}_2^n(s) + \mathscr{F}_3^n(s), \tag{4.3.51}$$

where $\|\mathscr{F}_3^n(s)\|_{H^1} \lesssim \tau^2$, the oscillatory term (in time) $\mathscr{F}_1^n(s)$ simplifies $f^n(s)$ by using $e^{-\frac{is\mathscr{T}^\varepsilon}{\varepsilon^2}} = e^{-\frac{is}{\varepsilon^2}}(I_2 - is\mathscr{D}^\varepsilon)\Pi_+ + e^{\frac{is}{\varepsilon^2}}(I_2 + is\mathscr{D}^\varepsilon)\Pi_- + O(s^2)$ and removing the non-oscillatory terms as in (4.3.27), $\mathscr{F}_2^n(s) = \mathscr{F}_2^n(0)$ is the non-oscillatory term (s independent) similar to (4.3.27). We can prove $\|\mathscr{F}_1^n(s)\|_{H^1} \lesssim \varepsilon$, $\|\partial_s \mathscr{F}_1^n(s)\|_{H^1} \lesssim 1/\varepsilon$, $\|\partial_{ss} \mathscr{F}_1^n(s)\|_{H^1} \lesssim 1/\varepsilon^3$.

Denote

$$\zeta_2^n(x) = \left(\int_0^\tau \mathscr{F}_1^n(s)\,ds - \tau \mathscr{F}_1^n(\tau/2)\right), \quad \zeta_3^n(x) = \left(\int_0^\tau \int_0^s \mathscr{G}_1^n(s,w)\,dwds - \frac{\tau^2}{2}\mathscr{G}_1^n(\tau/2,\tau/2)\right), \tag{4.3.52}$$

and we have

$$\eta^n(x) = e^{-\frac{i\tau}{\varepsilon^2}\mathscr{T}^\varepsilon}\left[\zeta_1^n(x) + \zeta_2^n(x) - \zeta_3^n(x)\right] + \kappa^n(x). \tag{4.3.53}$$

where

$$\begin{aligned}\kappa^n(x) = R^n(x) \quad &+ \quad e^{-\frac{i\tau}{\varepsilon^2}\mathscr{T}^\varepsilon}\left(\kappa_1^n(x) + \int_0^\tau \mathscr{F}_3^n(s)\,ds - \tau\mathscr{F}_3^n\left(\frac{\tau}{2}\right) - \int_0^\tau\int_0^s \mathscr{G}_3^n(s,w)\,dwds\right.\\ &+ \quad \left.\frac{\tau^2}{2}\mathscr{G}_3^n\left(\frac{\tau}{2},\frac{\tau}{2}\right)\right),\end{aligned} \tag{4.3.54}$$

and $\|\kappa^n(x)\|_{H^1} \lesssim \tau^3 + \tau\|\mathbf{e}^n(x)\|_{H^1}$.

Following the idea in S_1 case (4.3.32), we have the error equation for S_2 with $0 \le n \le \frac{T}{\tau} - 1$

$$\mathbf{e}^{n+1}(x) = e^{-\frac{i\tau}{\varepsilon^2}\mathscr{T}^\varepsilon}\mathbf{e}^n(x) + \zeta_1^n(x) + \zeta_2^n(x) - \zeta_3^n(x) + \kappa^n(x) + \widetilde{L}_n(\mathbf{e}^n(x)), \tag{4.3.55}$$

where $\widetilde{L}_n(\mathbf{e}^n(x)) = e^{-\frac{i\tau}{2\varepsilon^2}\mathscr{T}^\varepsilon}\left(e^{-i\tau\mathbf{F}\left(e^{-\frac{i\tau}{2\varepsilon^2}\mathscr{T}^\varepsilon}\Phi^n\right)} - I_2\right)e^{-\frac{i\tau}{2\varepsilon^2}\mathscr{T}^\varepsilon}$, and $\|\widetilde{L}_n\mathbf{e}^n(x)\|_{H^1} \le e^{c_{M_1}\tau}\|\mathbf{e}^n(x)\|_{H^1}$ (c_{M_1} depends on M_1) . For $0 \le n \le \frac{T}{\tau} - 1$, we would have

$$\|\mathbf{e}^{n+1}(x)\|_{H^1} \lesssim \tau^2 + \tau\sum_{k=0}^n \|\mathbf{e}^k(x)\|_{H^1} + \sum_{j=1,2,3}\left\|\sum_{k=0}^n e^{-\frac{i(n-k+1)\tau}{\varepsilon^2}\mathscr{T}^\varepsilon}\zeta_j^k(x)\right\|_{H^1}. \tag{4.3.56}$$

Under the hypothesis of Theorem 4.2, we have

$$\|\mathscr{F}_1^n(s)\|_{H^1} \lesssim \varepsilon, \quad \|\partial_s \mathscr{F}_1^n(s)\|_{H^1} \lesssim \varepsilon/\varepsilon^2 = 1/\varepsilon, \quad \|\partial_{ss}\mathscr{F}_1^n(s)\|_{H^1} \lesssim 1/\varepsilon^3, \quad 0 \le s \le \tau;$$
$$\|\mathscr{G}_1^n(s,w)\|_{H^1} \lesssim \varepsilon, \quad \|\partial_s \mathscr{G}_1^n(s,w)\|_{H^1} \lesssim 1/\varepsilon, \quad \|\partial_w \mathscr{G}_1^n(s,w)\|_{H^1} \lesssim 1/\varepsilon, \quad 0 \le s,w \le \tau.$$

which together with (4.3.52) gives $\|\zeta_2^n(x)\|_{H^1} \lesssim \min\{\varepsilon\tau, \tau^3/\varepsilon^3\}$ and $\|\zeta_3^n(x)\|_{H^1} \lesssim \min\{\varepsilon\tau^2, \tau^3/\varepsilon\}$. Since $\|\zeta_1^n(x)\|_{H^1} \lesssim \min\{\tau^2\varepsilon, \frac{\tau^3}{\varepsilon}\}$, we derive from (4.3.56) that

$$\begin{aligned}\|\mathbf{e}^{n+1}(x)\|_{H^1} \lesssim & \tau^2 + \tau\sum_{k=0}^{n}\|\mathbf{e}^k(x)\|_{H^1} + n\min\{\varepsilon\tau^2, \tau^3/\varepsilon\} + \left\|\sum_{k=0}^{n} e^{-\frac{i(n-k+1)\tau}{\varepsilon^2}\mathscr{T}^\varepsilon}\zeta_2^k(x)\right\|_{H^1} \\ \lesssim & \tau^2 + n\min\{\varepsilon\tau, \tau^3/\varepsilon^3\} + \tau\sum_{k=0}^{n}\|\mathbf{e}^k(x)\|_{H^1}, \quad 0 \le n \le \frac{T}{\tau} - 1.\end{aligned} \tag{4.3.57}$$

The discrete Gronwall's inequality will give the desired results in Theorem 4.2 with the help of mathematical induction. □

4.4 Improved uniform error bounds for non-resonant time steps

The leading term in NLDE (4.2.2) is $\frac{1}{\varepsilon^2}\sigma_3\Phi$, suggesting that the solution behaves almost periodically in time with periods $2k\pi\varepsilon^2$ ($k \in \mathbb{N}^*$, the periods of $e^{-i\sigma_3/\varepsilon^2}$). From numerical results, we observe that S_1 behave much better than the results in Theorem 4.1 when 4τ is not close to the leading temporal oscillation periods $2k\pi\varepsilon^2$. In fact, for given $0 < \kappa \le 1$, define

$$\mathscr{A}_\kappa(\varepsilon) := \bigcup_{k=0}^{\infty}\left[0.5\varepsilon^2 k\pi + 0.5\varepsilon^2\arcsin\kappa, 0.5\varepsilon^2(k+1)\pi - 0.5\varepsilon^2\arcsin\kappa\right], \quad 0 < \varepsilon \le 1, \tag{4.4.1}$$

then when $\tau \in \mathscr{A}_\kappa(\varepsilon)$, i.e., when non-resonant time step sizes are chosen, the errors of S_1 can be improved. To illustrate $\mathscr{A}_\kappa(\varepsilon)$ (compared to the linear case [17], the resonant steps $\mathscr{A}_\kappa^c(\varepsilon)$ for fixed ε double due to the cubic nonlinearity), we show in Figure 4.4.1 for $\varepsilon = 1$ and $\varepsilon = 0.5$ with fixed $\kappa = 0.15$.

For $\tau \in \mathscr{A}_\kappa(\varepsilon)$, we can derive improved uniform error bounds for S_1 as follows.

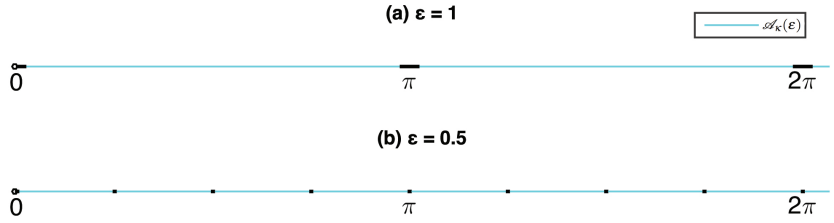

Figure 4.4.1: Illustration of the non-resonant time step $\mathscr{A}_\kappa(\varepsilon)$ with $\kappa = 0.15$ for (a) $\varepsilon = 1$ and (b) $\varepsilon = 0.5$.

Theorem 4.3. *Let $\Phi^n(x)$ be the numerical approximation obtained from S_1 (4.2.3). If the time step size τ is non-resonant, i.e. there exists $0 < \kappa \leq 1$, such that $\tau \in \mathscr{A}_\kappa(\varepsilon)$, then under the assumptions (C) and (D) with $m = m_* = 1$, we have an improved uniform error bound for small enough $\tau > 0$*

$$\|\mathbf{e}^n(x)\|_{H^1} \lesssim_\kappa \tau, \quad 0 \leq n \leq \frac{T}{\tau}. \tag{4.4.2}$$

Proof. First of all, the assumptions of Theorem 4.1 are satisfied in Theorem 4.3, so we can directly use the results of Theorem 4.1. In particular, the numerical solution Φ^n are bounded in H^1 as $\|\Phi^n\|_{H^1} \leq M_1 + 1$ (4.3.29) and Lemma 4.2 for local truncation error holds.

We start from (4.3.34). The improved estimates rely on the cancellation phenomenon for the η_2^k term in (4.3.34). From Lemma 4.2, (4.3.12), (4.3.13) and (4.3.14), we can write $\eta_2^k(x)$ with $\Phi_\pm(s) := \Pi_\pm^\varepsilon \Phi(s,x)$ as

$$\begin{aligned}\eta_2^k(x) :=& p_1(\tau)\mathscr{R}_{4,-}(\Phi_+(t_k),\Phi_-(t_k)) - \overline{p_1(\tau)}\mathscr{R}_{4,+}(\Phi_+(t_k),\Phi_-(t_k)) \\ &+ p_2(\tau)\mathscr{R}_{2,-}(\Phi_+(t_k),\Phi_-(t_k)) - \overline{p_2(\tau)}\mathscr{R}_{2,+}(\Phi_+(t_k),\Phi_-(t_k)),\end{aligned} \tag{4.4.3}$$

where $\mathscr{R}_{j,\pm}(\Phi_+,\Phi_-)$ ($j = 2,4, \Phi_+,\Phi_- : \mathbb{R} \to \mathbb{C}^2$) are as follows

$$\begin{aligned}&\mathscr{R}_{4,-}(\Phi_+,\Phi_-) = \Pi_-^\varepsilon\left(\mathbf{g}_1(\Phi_+,\Phi_-)\Phi_+\right), \quad \mathscr{R}_{4,+}(\Phi_+,\Phi_-) = \Pi_-^\varepsilon\left(\overline{\mathbf{g}_1(\Phi_+,\Phi_-)}\Phi_-\right), \\ &\mathscr{R}_{2,-}(\Phi_+,\Phi_-) = \Pi_+^\varepsilon\left(\mathbf{g}_1(\Phi_+,\Phi_-)\Phi_+\right) + \Pi_-^\varepsilon\left(\mathbf{g}_2(\Phi_+,\Phi_-)\Phi_+ + \mathbf{g}_1(\Phi_+,\Phi_-)\Phi_-\right), \\ &\mathscr{R}_{2,+}(\Phi_+,\Phi_-) = \Pi_+^\varepsilon\left(\overline{\mathbf{g}_1(\Phi_+,\Phi_-)}\Phi_+\right) + \Pi_-^\varepsilon\left(\mathbf{g}_2(\Phi_+,\Phi_-)\Phi_+ + \overline{\mathbf{g}_1(\Phi_+,\Phi_-)}\Phi_-\right),\end{aligned} \tag{4.4.4}$$

with $\mathbf{g}_1, \mathbf{g}_2$ given in (4.3.13)-(4.3.14) (Lemma 4.2), and

$$p_1(\tau) = -i\left(\int_0^\tau e^{-\frac{4si}{\varepsilon^2}}\,ds - \tau\right), \quad p_2(\tau) = -i\left(\int_0^\tau e^{-\frac{2si}{\varepsilon^2}}\,ds - \tau\right). \tag{4.4.5}$$

It is obvious that $|p_1(\tau)|, |p_2(\tau)| \leq 2\tau$ and (4.3.34) implies that

$$\begin{aligned}\|\mathbf{e}^{n+1}(x)\|_{H^1} \lesssim & \tau \sum_{\sigma=\pm, j=2,4} \left\| \sum_{k=0}^{n} e^{-\frac{i(n-k+1)\tau}{\varepsilon^2}\mathscr{T}^{\varepsilon}} \mathscr{R}_{j,\sigma}(\Phi_+(t_k), \Phi_-(t_k)) \right\|_{H^1} \\ & + \tau + \tau \sum_{k=0}^{n} \|\mathbf{e}^k(x)\|_{H^1}. \end{aligned} \tag{4.4.6}$$

To proceed, we introduce $\widetilde{\Phi}_{\pm}(t)$ as

$$\widetilde{\Phi}_{\pm}(t) := \tilde{\Phi}_{\pm}(t,x) = e^{\pm\frac{it}{\varepsilon^2}}\Phi_{\pm}(t,x) = e^{\pm\frac{it}{\varepsilon^2}}\Pi_{\pm}^{\varepsilon}\Phi(t,x), \quad 0 \leq t \leq T. \tag{4.4.7}$$

Since $\Phi(t,x)$ solves NLDE (4.1.1) (or (4.2.2)), noticing the properties of $\mathscr{T}^{\varepsilon}$ as in (4.3.6) and (4.3.9) and the L^2 orthogonal projections $\Pi_{\pm}^{\varepsilon}$, it is straightforward to compute that

$$i\partial_t\widetilde{\Phi}_{\pm}(t) = \mathscr{D}^{\varepsilon}\widetilde{\Phi}_{\pm}(t) + \Pi_{\pm}\left(e^{\mp\frac{it}{\varepsilon^2}}\mathbf{F}(\Phi(t))\Phi(t)\right), \tag{4.4.8}$$

and the assumptions of Theorem 4.1 would yield

$$\|\tilde{\Phi}_{\pm}(\cdot)\|_{L^\infty([0,T];(H^3)^2)} \lesssim 1, \quad \|\partial_t\tilde{\Phi}_{\pm}(\cdot)\|_{L^\infty([0,T];(H^1)^2)} \lesssim 1. \tag{4.4.9}$$

Now, we can deal with the terms involving $\mathscr{R}_{j,\pm}$ $(j = 2,4)$ in (4.4.4).

For $\mathscr{R}_{4,-}$. By direct computation, we get $\mathscr{R}_{4,-}(\Phi_+(t_k), \Phi_-(t_k)) = e^{-\frac{3it_k}{\varepsilon^2}}\mathscr{R}_{4,-}(\widetilde{\Phi}_+(t_k), \widetilde{\Phi}_-(t_k))$. In view of (4.3.10) and (4.4.4), we have for $0 \leq k \leq n \leq \frac{T}{\tau} - 1$,

$$e^{-\frac{i(n-k+1)\tau}{\varepsilon^2}\mathscr{T}^{\varepsilon}}\mathscr{R}_{4,-}(\Phi_+(t_k), \Phi_-(t_k)) = e^{\frac{i(n+1-4k)\tau}{\varepsilon^2}}e^{i(t_{n+1}-t_k)\mathscr{D}^{\varepsilon}}\mathscr{R}_{4,-}(\widetilde{\Phi}_+(t_k), \widetilde{\Phi}_-(t_k)). \tag{4.4.10}$$

Denoting

$$A(t) := A(t,x) = e^{-it\mathscr{D}^{\varepsilon}}\mathscr{R}_{j,\sigma}(\widetilde{\Phi}_+(t), \widetilde{\Phi}_-(t)), \quad 0 \leq t \leq T, \tag{4.4.11}$$

and noticing that $\partial_t A(t) = -ie^{-it\mathscr{D}^{\varepsilon}}\mathscr{D}^{\varepsilon}\mathscr{R}_{j,\sigma}(\widetilde{\Phi}_+(t), \widetilde{\Phi}_-(t)) + e^{-it\mathscr{D}^{\varepsilon}}\partial_t\mathscr{R}_{j,\sigma}(\widetilde{\Phi}_+(t), \widetilde{\Phi}_-(t))$, we can derive from (4.4.9) and the fact that $\mathscr{D}^{\varepsilon} : (H^m)^2 \to (H^{m-2})^2$ is uniformly bounded w.r.t ε,

$$\begin{aligned}\|A(t_k) - A(t_{k-1})\|_{H^1} \lesssim & \tau\left[\|\mathscr{R}_{j,\sigma}(\widetilde{\Phi}_+(t_k), \widetilde{\Phi}_-(t_k))\|_{H^3} + \|\partial_t\mathscr{R}_{j,\sigma}(\widetilde{\Phi}_+(t), \widetilde{\Phi}_-(t))\|_{L^\infty([0,T];(H^1)^2)}\right] \\ \lesssim & \tau, \quad 1 \leq k \leq \frac{T}{\tau}. \end{aligned} \tag{4.4.12}$$

Using (4.4.12), (4.4.10), $\|A(t)\|_{L^\infty([0,T];(H^1)^2)} \lesssim 1$, the property that $e^{it\mathscr{D}^\varepsilon}$ preserves H^1 norm, summation by parts formula and triangle inequality, we have

$$\begin{aligned}\left\|\sum_{k=0}^{n} e^{-\frac{i(n-k+1)\tau}{\varepsilon^2}\mathscr{T}^\varepsilon}\mathscr{R}_{4,-}(\Phi_+(t_k),\Phi_-(t_k))\right\|_{H^1} &= \left\|\sum_{k=0}^{n} e^{-\frac{i4k\tau}{\varepsilon^2}}A(t_k)\right\|_{H^1} \\ \le \left\|\sum_{k=0}^{n-1}\theta_k(A(t_k)-A(t_{k+1})\right\|_{H^1} &+ \|\theta_n A(t_n)\|_{H^1} \lesssim \tau\left|\sum_{k=0}^{n-1}\theta_k\right|+1\end{aligned} \tag{4.4.13}$$

with

$$\theta_k = \sum_{j=0}^{k} e^{-\frac{i4j\tau}{\varepsilon^2}} = \frac{1-e^{-\frac{i4(k+1)\tau}{\varepsilon^2}}}{1-e^{-\frac{i4\tau}{\varepsilon^2}}}, \quad k\ge 0, \quad \theta_{-1}=0. \tag{4.4.14}$$

For $\tau\in\mathscr{A}_\kappa(\varepsilon)$ (4.4.1), we have $|1-e^{-\frac{i4\tau}{\varepsilon^2}}| = |2\sin(2\tau/\varepsilon^2)| \ge 2\kappa$ and $|\theta_k|\le \frac{2}{2\kappa} = 1/\kappa$, and (4.4.13) leads to

$$\left\|\sum_{k=0}^{n} e^{-\frac{i(n-k+1)\tau}{\varepsilon^2}\mathscr{T}^\varepsilon}\mathscr{R}_{4,-}(\Phi_+(t_k),\Phi_-(t_k))\right\|_{H^1} \lesssim \frac{n\tau+1}{\kappa} \lesssim \frac{1}{\kappa}. \tag{4.4.15}$$

For $\mathscr{R}_{2,-}$. Similar to the case $\mathscr{R}_{4,-}$ (slightly different), it is straightforward to show that

$$e^{-\frac{i(n-k+1)\tau}{\varepsilon^2}\mathscr{T}^\varepsilon}\mathscr{R}_{2,-}(\Phi_+(t_k),\Phi_-(t_k)) = e^{\frac{i(n+1-2k)\tau}{\varepsilon^2}}\left[e^{-it_{n+1}\mathscr{D}^\varepsilon}B(t_k)+e^{it_{n+1}\mathscr{D}^\varepsilon}C(t_k)\right], \tag{4.4.16}$$

where

$$B(t) = e^{it\mathscr{D}^\varepsilon}\Pi_+^\varepsilon\left(\mathbf{g}_1(\widetilde{\Phi}_+(t),\widetilde{\Phi}_-(t))\widetilde{\Phi}_+(t)\right), \tag{4.4.17}$$

$$C(t) = e^{-it\mathscr{D}^\varepsilon}\Pi_-^\varepsilon\left(\mathbf{g}_2(\widetilde{\Phi}_+(t),\widetilde{\Phi}_-(t))\widetilde{\Phi}_+(t)+\mathbf{g}_1(\widetilde{\Phi}_+(t),\widetilde{\Phi}_-(t))\Phi_-(t)\right). \tag{4.4.18}$$

$B(t)$ and $C(t)$ satisfies the same estimates as $A(t)$ (4.4.12). Therefore, similar procedure will give

$$\begin{aligned}\left\|\sum_{k=0}^{n} e^{-\frac{i(n-k+1)\tau}{\varepsilon^2}\mathscr{T}^\varepsilon}\mathscr{R}_{2,-}(\Phi_+(t_k),\Phi_-(t_k))\right\|_{H^1} &\le \left\|\sum_{k=0}^{n} e^{-\frac{i2k\tau}{\varepsilon^2}}B(t_k)\right\|_{H^1} + \left\|\sum_{k=0}^{n} e^{-\frac{i2k\tau}{\varepsilon^2}}C(t_k)\right\|_{H^1} \\ \lesssim \tau\left|\sum_{k=0}^{n-1}\widetilde{\theta}^k\right|+1,\end{aligned} \tag{4.4.19}$$

with $\widetilde{\theta}_k = \sum_{j=0}^{k} e^{-\frac{i2j\tau}{\varepsilon^2}} = \frac{1-e^{-\frac{i2(k+1)\tau}{\varepsilon^2}}}{1-e^{\frac{-i2\tau}{\varepsilon^2}}}$, $k\ge 0$, $\widetilde{\theta}_{-1}=0$. For $\tau\in\mathscr{A}_\kappa(\varepsilon)$ (4.4.1), we know $|1-e^{\frac{-i2\tau}{\varepsilon^2}}| = |2\sin(\tau/\varepsilon^2)| \ge |4\sin(2\tau/\varepsilon^2)| \ge 4\kappa$ and $|\widetilde{\theta}_k| \le \frac{2}{4\kappa} = 2/\kappa$, which shows

$$\left\|\sum_{k=0}^{n} e^{-\frac{i(n-k+1)\tau}{\varepsilon^2}\mathscr{T}^\varepsilon}\mathscr{R}_{2,-}(\Phi_+(t_k),\Phi_-(t_k))\right\|_{H^1} \lesssim \tau\left|\sum_{k=0}^{n-1}\widetilde{\theta}^k\right|+1 \lesssim \frac{1}{\kappa}. \tag{4.4.20}$$

For $\mathscr{R}_{4,+}$ and $\mathscr{R}_{2,+}$. It is easy to see that the $\mathscr{R}_{4,+}$ and $\mathscr{R}_{2,+}$ terms in (4.4.6) can be bounded exactly the same as the $\mathscr{R}_{4,-}$ and $\mathscr{R}_{2,-}$ terms, respectively.

Finally, combining (4.4.6), (4.4.15), (4.4.20) and above observations, we have for $\tau \in \mathscr{A}_\kappa(\varepsilon)$,

$$\|\mathbf{e}^{n+1}(x)\|_{H^1} \lesssim \frac{\tau}{\kappa} + \tau \sum_{k=0}^{n} \|\mathbf{e}^k(x)\|_{H^1}, \quad 0 \le n \le \frac{T}{\tau} - 1, \tag{4.4.21}$$

and discrete Gronwall inequality will yield $\|\mathbf{e}^{n+1}(x)\|_{H^1} \lesssim \frac{\tau}{\kappa}$ $(0 \le n \le \frac{T}{\tau} - 1)$ for small enough $\tau \in \mathscr{A}_\kappa(\varepsilon)$. The proof is complete. □

Similar to the S_1 case, for non-resonant time steps, i.e., $\tau \in \mathscr{A}_\kappa(\varepsilon)$, we can derive improved uniform error bounds for S_2 as shown in the following theorem.

Theorem 4.4. *Let $\Phi^n(x)$ be the numerical approximation obtained from S_2 (4.2.4). If the time step size τ is non-resonant, i.e. there exists $0 < \kappa \le 1$, such that $\tau \in \mathscr{A}_\kappa(\varepsilon)$, then under the assumptions (C) and (D) with $m = 2$, $m_* = 1$, the following two error estimates hold for small enough $\tau > 0$*

$$\|\mathbf{e}^n(x)\|_{H^1} \lesssim_\kappa \tau^2 + \tau\varepsilon, \quad \|\mathbf{e}^n(x)\|_{H^1} \lesssim_\kappa \tau^2 + \tau^2/\varepsilon, \quad 0 \le n \le \frac{T}{\tau}. \tag{4.4.22}$$

As a result, there is an improved uniform error bound for S_2 when $\tau > 0$ is small enough

$$\|\mathbf{e}^n(x)\|_{H^1} \lesssim_\kappa \tau^2 + \max_{0<\varepsilon\le 1} \min\{\tau\varepsilon, \tau^2/\varepsilon\} \lesssim_\kappa \tau^{3/2}, \quad 0 \le n \le \frac{T}{\tau}. \tag{4.4.23}$$

Proof. As the proof is extended from the techniques used for S_1 and the proof for improved uniform error bounds for S_2 in the linear case [17], here we just show the outline of the proof for brevity.

We start from (4.3.57). Following the strategy in the S_1 case, the key idea is to extract the leading terms from $\Phi(t,x)$ as (4.4.7) for estimating $\zeta_2^n(x)$, and the computations are more or less the same. Recalling (4.3.52) , noticing $\mathscr{F}_1^n(s)$ is similar to $f_2^n(s)$ (3.3.13) and $\|\zeta_2^n(x)\|_{H^1} \lesssim \min\{\varepsilon\tau, \tau^2/\varepsilon\}$, following the computations in the proof of Theorem 4.3, we would get for $0 \le n \le \frac{T}{\tau} - 1$ and $\tau \in A_\kappa(\varepsilon)$,

$$\left\| \sum_{k=0}^{n} e^{-\frac{i(n-k+1)\tau}{\varepsilon^2}\mathscr{T}^\varepsilon} \zeta_2^k(x) \right\|_{H^1} \lesssim \sum_{k=0}^{n} \frac{1}{\kappa}\tau \min\{\varepsilon\tau, \tau^2/\varepsilon\} \lesssim \frac{1}{\kappa}\min\{\varepsilon\tau, \tau^2/\varepsilon\}, \tag{4.4.24}$$

and the conclusions of Theorem 4.4 hold by applying discrete Gronwall inequality to (4.3.57).

□

4.5 Numerical results

To verify our error bounds in Theorem 4.1 to Theorem 4.4, we show two numerical examples here. In the examples, we always use Fourier pseudospectral method for spatial discretization.

As a common practice when applying the Fourier pseudospectral method, in our numerical simulations, we truncate the whole space onto a sufficiently large bounded domain $\Omega = (a,b)$, and assume periodic boundary conditions. The mesh size is chosen as $h := \triangle x = \frac{b-a}{M}$ with M being an even positive integer. Then the grid points can be denoted as $x_j := a + jh$, for $j = 0,1,...,M$.

In this example and the examples later, we always choose the electric potential $V(x) \equiv 0$. For the nonlinearity (4.1.3), we take $\lambda_1 = 1$, $\lambda_2 = 0$, i.e.

$$\mathbf{F}(\Phi) = (\Phi^* \sigma_3 \Phi)\sigma_3, \tag{4.5.1}$$

and the initial data $\Phi_0 = (\phi_1, \phi_2)$ in (5.1.15) is given as

$$\phi_1(0,x) = e^{-\frac{x^2}{2}}, \quad \phi_2(0,x) = e^{-\frac{(x-1)^2}{2}}, \quad x \in \mathbb{R}. \tag{4.5.2}$$

As only the temporal errors are concerned in this paper, during the computation, the spatial mesh size is always set to be $h = \frac{1}{16}$ so that the spatial errors are negligible.

We first take resonant time steps, that is, for small enough chosen ε, there is a positive k_0, such that $\tau = k_0 \varepsilon \pi$, to check the error bounds in Theorem 4.1 and Theorem 4.2. The bounded computational domain is taken as $\Omega = (-32, 32)$. Because of the lack of available exact solution, for comparison, we use a numerical 'exact' solution generated by the second-order time-splitting method (S_2), which will be introduced later, with a very fine time step size $\tau_e = 2\pi \times 10^{-6}$.

To display the numerical results, we introduce the discrete H^1 errors of the numerical solution. Let $\Phi^n = (\Phi_0^n, \Phi_1^n, ..., \Phi_{M-1}^n, \Phi_M^n)^T$ be the numerical solution obtained by a numerical method with given ε, time step size τ as well as the fine mesh size h at time $t = t_n$, and $\Phi(t,x)$ be the exact solution, then the discrete H^1 error is defined as

$$e^{\varepsilon,\tau}(t_n) = \|\Phi^n - \Phi(t_n,\cdot)\|_{H^1} = \sqrt{h\sum_{j=0}^{M-1} |\Phi(t_n,x_j) - \Phi_j^n|^2 + h\sum_{j=0}^{M-1} |\Phi'(t_n,x_j) - (\Phi')_j^n|^2}, \tag{4.5.3}$$

where

$$(\Phi')_j^n = i\sum_{l=-M/2}^{M/2-1} \mu_l \widehat{\Phi}_l^n e^{i\mu_l(x_j-a)}, \quad j=0,1,...,M-1, \tag{4.5.4}$$

with μ_l, $\widehat{\Phi}_l^n \in \mathbb{C}^2$ defined as

$$\mu_l = \frac{2l\pi}{b-a}, \quad \widehat{\Phi}_l^n = \frac{1}{M}\sum_{j=0}^{M-1} \Phi_j^n e^{-i\mu_l(x_j-a)}, \quad l=-\frac{M}{2},...,\frac{M}{2}-1, \tag{4.5.5}$$

and $\Phi'(t_n,x_j)$ is defined similarly. Then $e^{\varepsilon,\tau}(t_n)$ should be close to the H^1 errors in Theorem 4.1 for fine spatial mesh sizes h.

Table 4.5.1 and Table 4.5.2 show the numerical temporal errors $e^{\varepsilon,\tau}(t=2\pi)$ with different ε and time step size τ for S_1 and S_2 respectively, up to the time $t=2\pi$.

The last two rows of Table 4.5.1 show the largest error of each column for fixed τ. The errors exhibit $1/2$ order convergence, which coincides well with Theorem 4.1. More specifically, we can observe when $\tau \gtrsim \varepsilon$ (below the lower bolded line), there is first order convergence, which agrees with the error bound $\|\Phi(t_n,x)-\Phi^n(x)\|_{H^1} \lesssim \tau+\varepsilon$. When $\tau \lesssim \varepsilon^2$ (above the upper bolded line), there is also first order convergence, which matches the other error bound $\|\Phi(t_n,x)-\Phi^n(x)\|_{H^1} \lesssim \tau+\tau/\varepsilon$.

In Table 4.5.2, the last two rows show the largest error of each column for fixed τ. We could clearly observe that there is $1/2$ order convergence, which agrees well with Theorem 4.2. More specifically, in Table 4.5.2, we can see when $\tau \gtrsim \sqrt{\varepsilon}$ (below the lower bolded line), there is second order convergence, which coincides with the error bound $\|\Phi(t_n,x)-\Phi^n(x)\|_{H^1} \lesssim \tau^2+\varepsilon$; when $\tau \lesssim \varepsilon^2$ (above the upper bolded line), we also observe second order convergence, which matches the other error bound $\|\Phi(t_n,x)-\Phi^n(x)\|_{H^1} \lesssim \tau^2+\tau^2/\varepsilon^3$.

The results from the example successfully validate the uniform error bounds for S_1 and S_2 in Theorem 4.1 and Theorem 4.2.

Moreover, to support the improved uniform error bound in Theorem 4.3 and Theorem 4.4, we further test the errors using non-resonant time steps here, i.e., we choose $\tau \in \mathscr{A}_\kappa(\varepsilon)$ for some given ε and fixed $0<\kappa\leq 1$. In this case, the bounded computational domain is set as $\Omega=(-16,16)$.

Table 4.5.1: Discrete H^1 temporal errors $e^{\varepsilon,\tau}(t=2\pi)$ for the wave function with resonant time step size, S_1 method.

$e^{\varepsilon,\tau}(t=2\pi)$	$\tau_0=\pi/4$	$\tau_0/4$	$\tau_0/4^2$	$\tau_0/4^3$	$\tau_0/4^4$	$\tau_0/4^5$
$\varepsilon_0=1$	4.18	**7.09E-1**	1.69E-1	4.17E-2	1.04E-2	2.59E-3
order	–	**1.28**	1.04	1.01	1.00	1.00
$\varepsilon_0/2$	2.54	6.37E-1	**1.44E-1**	3.55E-2	8.84E-3	2.21E-3
order	–	1.00	**1.07**	1.01	1.00	1.00
$\varepsilon_0/2^2$	2.25	1.15	1.47E-1	**3.53E-2**	8.73E-3	2.18E-3
order	–	0.49	1.48	**1.03**	1.01	1.00
$\varepsilon_0/2^3$	2.29	6.69E-1	6.56E-1	3.62E-2	**8.84E-3**	2.20E-3
order	–	0.89	0.01	2.09	**1.02**	1.00
$\varepsilon_0/2^4$	2.32	5.33E-1	3.24E-1	3.49E-1	8.98E-3	**2.22E-3**
order	–	1.06	0.36	-0.05	2.64	**1.01**
$\varepsilon_0/2^5$	2.34	**5.29E-1**	1.76E-1	1.70E-1	1.79E-1	2.24E-3
order	–	**1.07**	0.79	0.03	-0.04	3.16
$\varepsilon_0/2^7$	2.35	5.57E-1	**1.30E-1**	4.46E-2	4.28E-2	4.49E-2
order	–	1.04	**1.05**	0.77	0.03	-0.03
$\varepsilon_0/2^9$	2.35	5.68E-1	1.38E-1	**3.26E-2**	1.12E-2	1.07E-2
order	–	1.02	1.02	**1.04**	0.77	0.03
$\varepsilon_0/2^{11}$	2.35	5.71E-1	1.41E-1	3.45E-2	**8.14E-3**	2.80E-3
order	–	1.02	1.01	1.02	**1.04**	0.77
$\varepsilon_0/2^{13}$	2.35	5.72E-1	1.42E-1	3.53E-2	8.64E-3	**2.04E-3**
order	–	1.02	1.00	1.00	1.02	**1.04**
$\max_{0<\varepsilon\le 1} e^{\varepsilon,\tau}(t=2\pi)$	4.18	1.15	6.56E-1	3.49E-1	1.79E-1	9.07E-2
order	–	0.93	0.40	0.45	0.48	0.49

And for comparison, the numerical 'exact' solution is computed by the second-order time-splitting method (S_2) with a very small time step size $\tau_e = 8\times 10^{-6}$. Spatial mesh size is fixed as $h = 1/16$ for all the numerical simulations.

Table 4.5.3 and Table 4.5.4 respectively show the numerical errors $e^{\varepsilon,\tau}(t=4)$ with different ε and time step size τ for S_1 and S_2 up to the time $t=4$.

From Table 4.5.3, we could see that overall, for fixed time step size τ, i.e., for each column, the error $e^{\varepsilon,\tau}(t=4)$ does not change much with different ε. This verifies the temporal uniform first order convergence for S_1 with non-resonant time step size, as stated in Theorem 4.3.

The last two rows in Table 4.5.4 show the largest error of each column for fixed τ, which

Table 4.5.2: Discrete H^1 temporal errors $e^{\varepsilon,\tau}(t=2\pi)$ for the wave function of the NLDE (4.2.2) with resonant time step size, S_2 method.

$e^{\varepsilon,\tau}(t=2\pi)$	$\tau_0=\pi/4$	$\tau_0/4$	$\tau_0/4^2$	$\tau_0/4^3$	$\tau_0/4^4$	$\tau_0/4^5$
$\varepsilon_0=1$	4.51	**8.81E-2**	5.31E-3	3.31E-4	2.07E-5	1.29E-6
order	–	**2.84**	2.03	2.00	2.00	2.00
$\varepsilon_0/2$	3.81	1.57E-1	**4.70E-3**	2.90E-4	1.81E-5	1.13E-6
order	–	2.30	**2.53**	2.01	2.00	2.00
$\varepsilon_0/2^2$	1.78	1.56	7.98E-3	**4.41E-4**	2.73E-5	1.71E-6
order	–	0.09	3.81	**2.09**	2.00	2.00
$\varepsilon_0/2^3$	1.35	7.18E-1	7.74E-1	8.98E-4	**5.14E-5**	3.20E-6
order	–	0.46	-0.05	4.88	**2.06**	2.00
$\varepsilon_0/2^4$	1.26	3.69E-1	3.65E-1	3.80E-1	1.11E-4	**6.41E-6**
order	–	0.88	0.01	-0.03	5.87	**2.05**
$\varepsilon_0/2^5$	**1.25**	1.93E-1	1.83E-1	1.83E-1	1.87E-1	1.39E-5
order	**–**	1.35	0.04	0.00	-0.01	6.86
$\varepsilon_0/2^9$	1.25	**5.24E-2**	1.20E-2	1.15E-2	1.15E-2	1.15E-2
order	–	**2.29**	1.06	0.03	0.00	0.00
$\varepsilon_0/2^{13}$	1.25	5.01E-2	**2.66E-3**	7.53E-4	7.18E-4	7.17E-4
order	–	2.32	**2.12**	0.91	0.03	0.00
$\varepsilon_0/2^{17}$	1.25	5.00E-2	2.47E-3	**1.80E-4**	7.96E-5	7.78E-5
order	–	2.32	2.17	**1.89**	0.59	0.02
$\max_{0<\varepsilon\le 1} e^{\varepsilon,\tau}(t=2\pi)$	4.51	1.56	7.74E-1	3.80E-1	1.87E-1	9.26E-2
order	–	0.76	0.51	0.51	0.51	0.51

gives $3/2$ order of convergence, and it is consistent with Theorem 4.4. More specifically, in Table 4.5.4, we can observe the second order convergence when $\tau \gtrsim \varepsilon$ (below the lower bolded line) or when $\tau \lesssim \varepsilon^2$ (above the upper bolded line), agreeing with the error bound $\|\Phi(t_n,x)-\Phi^n(x)\|_{H^1} \lesssim \tau^2+\tau\varepsilon$ and the other error bound $\|\Phi(t_n,x)-\Phi^n(x)\|_{H^1} \lesssim \tau^2+\tau^2/\varepsilon$, respectively.

The results from the example successfully validate the improved uniform error bounds for S_1 and S_2 in Theorem 4.3 and Theorem 4.4.

Table 4.5.3: Discrete H^1 temporal errors $e^{\varepsilon,\tau}(t=4)$ for the wave function with non-resonant time step size, S_1 method.

$e^{\varepsilon,\tau}(t=4)$	$\tau_0=1/2$	$\tau_0/2$	$\tau_0/2^2$	$\tau_0/2^3$	$\tau_0/2^4$	$\tau_0/2^5$	$\tau_0/2^6$
$\varepsilon_0=1$	2.25	9.50E-1	4.55E-1	2.23E-1	1.10E-1	5.47E-2	2.73E-2
order	–	1.25	1.06	1.03	1.02	1.01	1.00
$\varepsilon_0/2$	3.32	1.03	3.81E-1	1.85E-1	9.14E-2	4.54E-2	2.27E-2
order	–	1.69	1.43	1.04	1.02	1.01	1.00
$\varepsilon_0/2^2$	2.08	7.67E-1	5.35E-1	1.90E-1	9.17E-2	4.51E-2	2.24E-2
order	–	1.44	0.52	1.49	1.05	1.02	1.01
$\varepsilon_0/2^3$	1.50	6.42E-1	3.99E-1	1.72E-1	1.01E-1	4.67E-2	2.29E-2
order	–	1.23	0.69	1.22	0.76	1.12	1.03
$\varepsilon_0/2^4$	1.56	7.49E-1	3.50E-1	1.68E-1	9.25E-2	4.39E-2	2.40E-2
order	–	1.06	1.10	1.06	0.86	1.08	0.87
$\varepsilon_0/2^5$	1.48	7.51E-1	3.99E-1	1.80E-1	8.75E-2	4.20E-2	2.29E-2
order	–	0.97	0.91	1.15	1.04	1.06	0.88
$\varepsilon_0/2^6$	1.50	7.12E-1	3.46E-1	1.81E-1	9.17E-2	4.49E-2	2.21E-2
order	–	1.08	1.04	0.94	0.98	1.03	1.02
$\varepsilon_0/2^7$	1.52	7.43E-1	3.76E-1	1.99E-1	1.16E-1	4.53E-2	2.28E-2
order	–	1.04	0.98	0.92	0.78	1.36	0.99
$\max_{0<\varepsilon\le 1} e^{\varepsilon,\tau}(t=4)$	3.32	1.03	5.35E-1	2.23E-1	1.16E-1	5.47E-2	2.73E-2
order	–	1.69	0.95	1.26	0.94	1.08	1.00

4.6 Extension to full-discretization

Similar to the case of the Dirac equation, the error estimates in Theorem 4.1 to Theorem 4.4 can be extended to full-discretization.

For this purpose, consider (4.2.2) with the initial condition (4.1.2) on a bounded domain $\Omega=[a,b]$ with periodic boundary conditions. Choose mesh size $h=\frac{b-a}{M}$ with M being an even positive integer, time step $\tau>0$, denote the grid points x_j, $j=0,1,\ldots,M$ and time steps t_n, $n=0,1,2,\ldots$ as before. Moreover, X_M, Y_M, Z_M, $[C_p(\overline{\Omega})]^2$, the projection and the interpolation operator are all defined to be the same as in the linear case.

We first consider the Lie-Trotter splitting S_1. Denote $\Phi^{[n]}(x)$ to be the semi-discretized numerical solution from S_1 (4.2.3), and Φ^n to be the full-discretized numerical solution with

Table 4.5.4: Discrete H^1 temporal errors $e^{\varepsilon,\tau}(t=4)$ for the wave function with non-resonant time step size, S_2 method.

$e^{\varepsilon,\tau}(t=4)$	$\tau_0=1/4$	$\tau_0/4$	$\tau_0/4^2$	$\tau_0/4^3$	$\tau_0/4^4$	$\tau_0/4^5$
$\varepsilon_0=1$	1.62E-1	**9.55E-3**	5.95E-4	3.72E-5	2.32E-6	1.45E-7
order	–	**2.04**	2.00	2.00	2.00	2.00
$\varepsilon_0/2$	6.31E-1	7.67E-3	**4.71E-4**	2.94E-5	1.84E-6	1.15E-7
order	–	3.18	**2.01**	2.00	2.00	2.00
$\varepsilon_0/2^2$	4.33E-1	1.49E-2	7.16E-4	**4.43E-5**	2.77E-6	1.73E-7
order	–	2.43	2.19	**2.01**	2.00	2.00
$\varepsilon_0/2^3$	3.88E-1	4.33E-2	1.52E-3	8.20E-5	**5.08E-6**	3.17E-7
order	–	1.58	2.42	2.11	**2.01**	2.00
$\varepsilon_0/2^4$	2.02E-1	4.29E-2	5.97E-3	1.86E-4	1.02E-5	**6.34E-7**
order	–	1.12	1.42	2.50	2.09	2.01
$\varepsilon_0/2^6$	**1.36E-1**	6.15E-3	1.10E-3	8.67E-4	1.00E-4	2.99E-6
order	**–**	2.23	1.24	0.17	1.56	2.53
$\varepsilon_0/2^8$	9.73E-2	**7.82E-3**	6.80E-3	6.59E-5	1.70E-5	1.40E-5
order	–	**1.82**	0.10	3.34	0.98	0.14
$\varepsilon_0/2^{10}$	9.65E-2	4.18E-3	**2.73E-4**	3.18E-5	2.56E-5	1.03E-6
order	–	2.27	**1.97**	1.55	0.15	2.32
$\varepsilon_0/2^{12}$	9.69E-2	4.00E-3	2.93E-4	**1.64E-5**	2.05E-6	4.31E-7
order	–	2.30	1.89	**2.08**	1.50	1.12
$\max_{0<\varepsilon\leq 1} e^{\varepsilon,\tau}(t=4)$	6.31E-1	5.88E-2	6.80E-3	8.67E-4	1.11E-4	1.40E-5
order	–	1.71	1.56	1.49	1.49	1.49

Fourier spectral discretization in space, i.e. we have for $n=0,1,...,\frac{T}{\tau}-1$

$$\begin{aligned}\Phi^{<1>}(x) &= e^{-i\tau\left[V(x)+\mathbf{F}(\Phi^{[n]}(x))\right]}\Phi^{[n]}(x),\\ \Phi^{[n+1]}(x) &= e^{-\frac{i\tau}{\varepsilon^2}\mathscr{T}^{\varepsilon}}\Phi^{<1>}(x), \quad x\in[a,b],\end{aligned} \tag{4.6.1}$$

with

$$\Phi^{[0]}(x)=\Phi(0,x), \quad x\in[a,b], \tag{4.6.2}$$

and

$$\begin{aligned}\Phi_j^{(1)} &= e^{-i\tau\left[V(x_j)+\mathbf{F}(\Phi_j^n))\right]}\Phi_j^n,\\ \Phi_j^{n+1} &= e^{-\frac{i\tau}{\varepsilon^2}\mathscr{T}^{\varepsilon}}I_M(\Phi^{(1)})(x_j) \quad j=0,1,...,M-1,\end{aligned} \tag{4.6.3}$$

with

$$\Phi_j^0=\Phi(0,x_j), \quad j=0,1,...M-1. \tag{4.6.4}$$

Moreover, the full-discretized error introduced in (3.6.7) is used. Then the uniform and improved uniform error bounds for S_1 in Theorem 4.1 and Theorem 4.3 can be extended to full-discretization as follows

Theorem 4.5. *(i) Under the assumptions (C) and (D) with $2m+m_* \geq 3$, we have the following full-discretized error estimate for S_1*

$$\|\mathbf{e}_f^n(x)\|_{H^1} \lesssim \sqrt{\tau} + h^{2m+m_*-1}, \quad 0 \leq n \leq \frac{T}{\tau}. \tag{4.6.5}$$

(ii) If the time step size τ is non-resonant, i.e. there exists $0 < \kappa \leq 1$, such that $\tau \in \mathscr{A}_\kappa(\varepsilon)$, then under the assumptions (C) and (D) with $2m+m_* \geq 3$, we have an improved uniform error bound for S_1

$$\|\mathbf{e}_f^n(x)\|_{H^1} \lesssim_\kappa \tau + h^{2m+m_*-1}, \quad 0 \leq n \leq \frac{T}{\tau}. \tag{4.6.6}$$

Proof. (i) Similar to the proof in the linear case, it is obvious that

$$\begin{aligned} \|\mathbf{e}_f^n(x)\|_{H^1} \quad \leq \quad & \|P_M(\Phi(t_n,x)) - \Phi(t_n,x)\|_{H^1} + \|\Phi(t_n,x) - \Phi^{[n]}(x)\|_{H^1} \\ & + \|\Phi^{[n]}(x) - P_M(\Phi^{[n]})(x)\|_{H^1} + \|P_M(\Phi^{[n]})(x) - I_M(\Phi^n)\|_{H^1}. \end{aligned} \tag{4.6.7}$$

From the regularity conditions, we have

$$\begin{aligned} &\|P_M(\Phi(t_n,x)) - \Phi(t_n,x)\|_{H^1} \lesssim h^{2m+m_*-1}, \\ &\|\Phi^{[n]}(x) - P_M(\Phi^{[n]})(x)\|_{H^1} \lesssim h^{2m+m_*-1}. \end{aligned} \tag{4.6.8}$$

Moreover, Theorem 4.1 suggests

$$\|\Phi(t_n,x) - \Phi^{[n]}(x)\|_{H^1} \lesssim \sqrt{\tau}. \tag{4.6.9}$$

As a result, we only need to focus on the term $\|P_M(\Phi^{[n]})(x) - I_M(\Phi^n)\|_{H^1}$.

Define the difference to be

$$\mathbf{e}_t^n(x) = P_M(\Phi^{[n]})(x) - I_M(\Phi^n)(x), \quad 0 \leq n \leq T/\tau, \tag{4.6.10}$$

then the result can be proved by mathematical induction.

It is easy to check that when $n = 0$, we have $\|\mathbf{e}_t^0(x)\|_{H^1} \lesssim h^{2m+m_*-1}$, so the error estimate holds.

Assume that for $0 \le n \le m \le \frac{T}{\tau} - 1$, the error estimate (4.6.5) holds. Then take $n = m+1$, we have for $\Phi^{[m+1]}$ and Φ^{m+1}

$$P_M(\Phi^{<1>}) = P_M(e^{-i\tau\left[V+\mathbf{F}\left(\Phi^{[m]}(x)\right)\right]}\Phi^{[m]}), \quad P_M(\Phi^{[m+1]}) = e^{-\frac{i\tau}{\varepsilon^2}\mathscr{T}^\varepsilon} P_M(\Phi^{<1>}),$$
$$I_M(\Phi^{(1)}) = I_M(e^{-i\tau\left[V(x_j)+\mathbf{F}(\Phi_j^n))\right]}(\Phi^m)), \quad I_M(\Phi^{m+1}) = e^{-\frac{i\tau}{\varepsilon^2}\mathscr{T}^\varepsilon} I_M(\Phi^{(1)}).$$

As $e^{-\frac{i\tau}{\varepsilon^2}\mathscr{T}^\varepsilon}$ preserves H^1 norm, we get

$$\|\mathbf{e}_t^{m+1}(\cdot)\|_{H^1} = \|P_M(\Phi^{<1>}) - I_M(\Phi^{(1)})\|_{H^1}.$$

On the other hand, we have

$$\begin{aligned} &P_M(\Phi^{<1>}) - I_M(\Phi^{(1)}) \\ &= P_M(e^{-i\tau\left[V+\mathbf{F}\left(\Phi^{[m]}(x)\right)\right]}\Phi^{[m]}) - I_M(e^{-i\tau\left[V(x_j)+\tau\mathbf{F}(\Phi_j^n))\right]}(\Phi^m)), \end{aligned} \tag{4.6.11}$$

which together with $\Phi^{<1>} \in H_p^{2m+m_*}$ implies

$$\|P_M(\Phi^{<1>}) - I_M(\Phi^{(1)})\|_{H^1} \lesssim h^{2m+m_*-1} + \|W(x)\|_{H^1}, \tag{4.6.12}$$

where

$$W(x) := I_M(e^{-i\tau\left[V+\mathbf{F}\left(\Phi^{[m]}(x)\right)\right]}\Phi^{[m]}) - I_M(e^{-i\tau\left[V(x_j)+\mathbf{F}(\Phi_j^n))\right]}(\Phi^m).$$

As shown in [11, 13, 19], $W(x)$ can be estimated through finite difference approximation as

$$\begin{aligned} \|W(x)\|_{H^1} &\lesssim \tau\left(\left\|\Phi_j^m - \Phi^{[m]}(x_j)\right\|_{l^2} + \left\|\kappa_x^+(\Phi_j^m - \Phi^{[m]}(x_j))\right\|_{l^2}\right) \\ &\lesssim \tau\left(\|\mathbf{e}_t^m(\cdot)\|_{H^1} + h^{2m+m_*-1}\right), \end{aligned}$$

where $\kappa_x^+\Phi_j = \frac{\Phi_{j+1}-\Phi_j}{h}$ is the forward finite difference operator. The key point is that $\|\partial_x(I_M\Psi_j)\|_{L^2} \sim \|\kappa_x^+\Psi_j\|_{l^2}$. Thus, we have

$$\|\mathbf{e}_t^{m+1}(\cdot)\|_{H^1} \lesssim \tau\left(\|\mathbf{e}_t^m(\cdot)\|_{H^s} + h^{2m+m_*-1}\right). \tag{4.6.13}$$

Indeed, it is true for all $n \le m$,

$$\|\mathbf{e}_t^{n+1}(\cdot)\|_{H^1} \lesssim \tau\left(\|\mathbf{e}_t^n(\cdot)\|_{H^s} + h^{2m+m_*-1}\right). \tag{4.6.14}$$

Using discrete Gronwal inequality, we get

$$\|\mathbf{e}_t^{n+1}(\cdot)\|_{H^1} \lesssim h^{2m+m_*-1}, \quad n \le m \le \frac{T}{\tau} - 1. \tag{4.6.15}$$

Thus (4.6.5) holds true for $n = m+1$ by using the discrete Sobolev inequality with sufficiently small h and τ, together with (4.6.8) and (4.6.9). This completes the induction and the proof.

(ii) The proof for non-resonant time step is similar to the proof for (i). The details are omitted here for brevity. □

Next we consider the Strang splitting S_2. Similarly, denote $\Phi^{[n]}(x)$ to be the semi-discretized numerical solution from S_2 (4.2.4), and Φ^n to be the full-discretized numerical solution with Fourier spectral discretization in space, i.e. we have for $n = 0, 1, ..., \frac{T}{\tau} - 1$

$$\begin{aligned}
&\Phi^{<1>}(x) = e^{-\frac{i\tau}{2\varepsilon^2}\mathscr{T}^\varepsilon}(\Phi^{[n]})(x),\\
&\Phi^{<2>}(x) = e^{-i\tau\left[V(x)+\mathbf{F}\left(\Phi^{<1>}(x)\right)\right]}\Phi^{<1>}(x), \quad j = 0, 1, ..., M-1,\\
&\Phi^{[n+1]}(x) = e^{-\frac{i\tau}{2\varepsilon^2}\mathscr{T}^\varepsilon}\Phi^{<2>}(x),
\end{aligned} \tag{4.6.16}$$

with

$$\Phi^{[0]}(x) = \Phi(0,x), \quad x \in [a,b], \tag{4.6.17}$$

and

$$\begin{aligned}
&\Phi_j^{(1)} = e^{-\frac{i\tau}{2\varepsilon^2}\mathscr{T}^\varepsilon} I_M(\Phi^n)(x_j),\\
&\Phi_j^{(2)} = e^{-i\tau\left[V(x_j)+\mathbf{F}\left(\Phi_j^{(1)}\right)\right]}\Phi_j^{(1)}, \quad j = 0, 1, ..., M-1,\\
&\Phi_j^{n+1} = e^{-\frac{i\tau}{2\varepsilon^2}\mathscr{T}^\varepsilon} I_M(\Phi^{(2)})(x_j),
\end{aligned} \tag{4.6.18}$$

with

$$\Phi_j^0 = \Phi(0,x_j), \quad j = 0, 1, ...M-1. \tag{4.6.19}$$

The full-discretized error is still defined as (3.6.7), and then the uniform and improved uniform error bounds for S_2 in Theorem 4.2 and Theorem 4.4 can be extended to full-discretization as follows

Theorem 4.6. *(i) Under the assumptions (C) and (D) with $2m+m_* \geq 5$, we have the following full-discretized error estimate for S_2*

$$\|\mathbf{e}_f^n(x)\|_{H^1} \lesssim \sqrt{\tau} + h^{2m+m_*-1}, \quad 0 \leq n \leq \frac{T}{\tau}. \tag{4.6.20}$$

(ii) If the time step size τ is non-resonant, i.e. there exists $0 < \kappa \leq 1$, such that $\tau \in \mathscr{A}_\kappa(\varepsilon)$, under the assumptions (C) and (D) with $2m+m_* \geq 5$, then the following improved uniform error estimate for S_2 holds

$$\|\mathbf{e}_f^n(x)\|_{H^1} \lesssim_\kappa \tau^{3/2} + h^{2m+m_*-1}, \quad 0 \leq n \leq \frac{T}{\tau}. \tag{4.6.21}$$

Proof. (i) The process of the proof is similar to the S_1 case in Theorem 4.5. The inequality (4.6.7), together with the estimates (4.6.8) and (4.6.9) still hold. As a result, we only need to focus on the term $\|P_M(\Phi^{[n]})(x) - I_M(\Phi^n)\|_{H^1}$.

With the definition for $\mathbf{e}_t^n(x)$ in (4.6.10), the result can be proved by mathematical induction.

It is easy to check that when $n = 0$, we have $\|\mathbf{e}_t^0(x)\|_{H^1} \lesssim h^{2m+m_*-1}$, so the error estimate holds.

Assume that for $0 \le n \le m \le \frac{T}{\tau} - 1$, the error estimate (4.6.20) holds. Then take $n = m+1$, we have for $\Phi^{[m+1]}$ and Φ^{m+1}

$$
\begin{aligned}
&P_M(\Phi^{<1>}) = e^{-\frac{i\tau}{2\varepsilon^2}\mathscr{T}^\varepsilon} P_M(\Phi^{[m]}), \quad P_M(\Phi^{[m+1]}) = e^{-\frac{i\tau}{2\varepsilon^2}\mathscr{T}^\varepsilon} P_M(\Phi^{<2>}),\\
&P_M(\Phi^{<2>}) = P_M(e^{-i\tau\left[V(x)+\mathbf{F}\left(\Phi^{<1>}(x)\right)\right]}\Phi^{<1>}),\\
&I_M(\Phi^{(1)}) = e^{-\frac{i\tau}{2\varepsilon^2}\mathscr{T}^\varepsilon} I_M(\Phi^m), \quad I_M(\Phi^{m+1}) = e^{-\frac{i\tau}{2\varepsilon^2}\mathscr{T}^\varepsilon} I_M(\Phi^{(2)}),\\
&I_M(\Phi^{(2)}) = I_M(e^{-i\tau\left[V+\mathbf{F}\left(\Phi^{(1)}\right)\right]}\Phi^{(1)}).
\end{aligned}
$$

As $e^{-\frac{i\tau}{2\varepsilon^2}\mathscr{T}^\varepsilon}$ preserves H^1 norm, we get

$$
\|\mathbf{e}_t^m(\cdot)\|_{H^1} = \|P_M(\Phi^{<1>} - I_M(\Phi^{(1)}))\|_{H^1}, \ \ \|\mathbf{e}_t^{m+1}(\cdot)\|_{H^1} = \|P_M(\Phi^{<2>}) - I_M(\Phi^{(2)})\|_{H^1}.
$$

On the other hand, we have

$$
\begin{aligned}
&P_M(\Phi^{<2>}) - I_M(\Phi^{(2)})\\
&= P_M(e^{-i\tau\left[V(x)+\mathbf{F}\left(\Phi^{<1>}(x)\right)\right]}\Phi^{<1>}) - I_M(e^{-i\tau\left[V+\mathbf{F}\left(\Phi^{(1)}\right)\right]}\Phi^{(1)}),
\end{aligned}
\tag{4.6.22}
$$

which together with $\Phi^{<2>} \in H_p^{2m+m_*}$ implies

$$
\|I_M(\Phi^{(2)}) - P_M(\Phi^{<2>})\|_{H^1} \lesssim h^{2m+m_*-1} + \|W(x)\|_{H^1}, \tag{4.6.23}
$$

where

$$
W(x) := I_M(e^{-i\tau\left[V+\mathbf{F}\left(\Phi^{(1)}\right)\right]}\Phi^{(1)}) - I_M(e^{-i\tau\left[V(x)+\mathbf{F}\left(\Phi^{<1>}(x)\right)\right]}\Phi^{<1>}).
$$

Similar to the S_1 case, $W(x)$ can be estimated through finite difference approximation as

$$
\begin{aligned}
\|W(x)\|_{H^1} &\lesssim \tau\left(\left\|\Phi_j^{(1)} - \Phi^{<1>}(x_j)\right\|_{l^2} + \left\|\kappa_x^+(\Phi_j^{(1)} - \Phi^{<1>}(x_j))\right\|_{l^2}\right)\\
&\lesssim \tau\left(\|\mathbf{e}_t^m(\cdot)\|_{H^1} + h^{2m+m_*-1}\right).
\end{aligned}
$$

Thus, we have

$$\|\mathbf{e}_t^{m+1}(\cdot)\|_{H^1} \lesssim \tau \left(\|\mathbf{e}_t^{m}(\cdot)\|_{H^s} + h^{2m+m_*-1} \right). \tag{4.6.24}$$

Indeed, it is true for all $n \leq m$,

$$\|\mathbf{e}_t^{n+1}(\cdot)\|_{H^1} \lesssim \tau \left(\|\mathbf{e}_t^{n}(\cdot)\|_{H^s} + h^{2m+m_*-1} \right). \tag{4.6.25}$$

Using discrete Gronwal inequality, we get

$$\|\mathbf{e}_t^{n+1}(\cdot)\|_{H^1} \lesssim h^{2m+m_*-1}, \quad n \leq m \leq \frac{T}{\tau} - 1. \tag{4.6.26}$$

Thus (4.6.20) holds true for $n = m+1$ and use the discrete Sobolev inequality with sufficiently small h and τ, together with (4.6.8) and (4.6.9). This completes the induction and the proof.

(ii) The proof for non-resonant time step is similar to the proof for (i). The details are omitted here for brevity. □

4.7 Extension to fourth-order splitting methods

From our numerical experiments, we observe that super-resolution does not only hold for first-order (S_1) and second-order (S_2) time-splitting methods. Indeed, higher order splitting methods also have this property. As an illustration, here we apply two fourth-order time-splitting methods for the Dirac and nonlinear Dirac equation respectively, to show that super-resolution also takes place for higher order splitting methods.

4.7.1 The methods

As has been extensively studied in Chapter 2, in the linear case (1.1.17), where there is no magnetic potential, and the electric potential is time-independent, i.e., $A(t,\mathbf{x}) \equiv 0$, and $V(\mathbf{x})$ does not depend on time t, the fourth-order compact time-splitting method (S_{4c}) is superior in accuracy and efficiency among fourth-order methods. The discrete-in-time S_{4c} method can be represented as:

$$\Phi^{n+1}(x) = e^{-\frac{i\tau}{6}V(x)} e^{-\frac{i\tau}{2\varepsilon^2}\mathscr{T}^\varepsilon} e^{-\frac{2i\tau}{3}\widehat{V}(x)} e^{-\frac{i\tau}{2\varepsilon^2}\mathscr{T}^\varepsilon} e^{-\frac{i\tau}{6}V(x)} \Phi^n(x), \tag{4.7.1}$$

with the initial value $\Phi^0(x) = \Phi_0(x)$, where

$$\widehat{V}(x) := V(x) - \frac{\tau^2}{48\varepsilon^2}[V(x),[\mathscr{T}^\varepsilon, V(x)]], \tag{4.7.2}$$

with $[A,B] := AB - BA$ the commutator. According to Chapter 2, the double commutator $[V(x),[\mathscr{T}^\varepsilon,V(x)] \equiv 0$, so that $\widehat{V}(x) \equiv V(x)$, and S_{4c} can be simplified to

$$\Phi^{n+1}(x) = e^{-\frac{i\tau}{6}V(x)} e^{-\frac{i\tau}{2\varepsilon^2}\mathscr{T}^\varepsilon} e^{-\frac{2i\tau}{3}V(x)} e^{-\frac{i\tau}{2\varepsilon^2}\mathscr{T}^\varepsilon} e^{-\frac{i\tau}{6}V(x)} \Phi^n(x), \quad \Phi^0(x) = \Phi_0(x). \tag{4.7.3}$$

On the other hand, as S_{4c} can not be easily extended to the nonlinear case, for (4.2.2), we use the fourth-order partitioned Runge-Kutta splitting method (S_{4RK}) [27, 30, 69], which is also a highly accurate fourth-order splitting method, as an alternative here.

Denote the nonlinear propagator $e^{i\tau\mathbf{W}^n(x)}$ as

$$e^{i\tau\mathbf{W}^n(x)}\Psi(x) := e^{i\tau V(x)+i\tau\mathbf{F}(\Psi(x))}\Psi(x), \tag{4.7.4}$$

then the discrete-in-time S_{4RK} method can be represented as:

$$\begin{aligned}\Phi^{n+1}(x) =& e^{-\frac{ia_1\tau}{\varepsilon^2}\mathscr{T}^\varepsilon} e^{-ib_1\mathbf{W}^n(x)} e^{-\frac{ia_2\tau}{\varepsilon^2}\mathscr{T}^\varepsilon} e^{-ib_2\mathbf{W}^n(x)} e^{-\frac{ia_3\tau}{\varepsilon^2}\mathscr{T}^\varepsilon} e^{-ib_3\mathbf{W}^n(x)} e^{-\frac{ia_4\tau}{\varepsilon^2}\mathscr{T}^\varepsilon} \\ & e^{-ib_3\mathbf{W}^n(x)} e^{-\frac{ia_3\tau}{\varepsilon^2}\mathscr{T}^\varepsilon} e^{-ib_2\mathbf{W}^n(x)} e^{-\frac{ia_2\tau}{\varepsilon^2}\mathscr{T}^\varepsilon} e^{-ib_1\mathbf{W}^n(x)} e^{-\frac{ia_1\tau}{\varepsilon^2}\mathscr{T}^\varepsilon} \Phi^n(x),\end{aligned} \tag{4.7.5}$$

with the initial value $\Phi^0(x) = \Phi_0(x)$, and the constants

$$a_1 = 0.0792036964311957, \quad a_2 = 0.353172906049774, \tag{4.7.6}$$

$$a_3 = -0.0420650803577195, \quad a_4 = 1 - 2(a_1 + a_2 + a_3), \tag{4.7.7}$$

$$b_1 = 0.209515106613362, \quad b_2 = -0.143851773179818, \quad b_3 = \frac{1}{2} - (b_1 + b_2). \tag{4.7.8}$$

From our numerical experiments, we find out that both methods present super-resolution in time, and the error bounds can be inferred as

Theorem 4.7. *Let $\Phi^n(x)$ be the numerical approximation obtained from S_{4c} (4.7.3) for (3.2.2) or S_{4RK} (4.7.5) for (4.2.2), then under the assumptions (A) and (B) with $m = 4$ for the Dirac equation, or (C) and (D) with $m = 4$, $m_* = 1$ for the NLDE, we have the following error estimates for small enough time step size $\tau > 0$ ($s = 0$ for the Dirac equation, and $s = 1$ for the NLDE)*

$$\|\mathbf{e}^n(x)\|_{H^s} \lesssim \tau^4 + \varepsilon, \quad \|\mathbf{e}^n(x)\|_{H^s} \lesssim \tau^4 + \tau^4/\varepsilon^7, \quad 0 \le n \le \frac{T}{\tau}. \tag{4.7.9}$$

As a result, there is a uniform error bound for $\tau > 0$ small enough

$$\|\mathbf{e}^n(x)\|_{H^s} \lesssim \tau^4 + \max_{0<\varepsilon\le 1} \min\{\varepsilon, \tau^4/\varepsilon^7\} \lesssim \sqrt{\tau}, \quad 0 \le n \le \frac{T}{\tau}. \tag{4.7.10}$$

Furthermore, for non-resonant time-steps, similar to S_1 and S_2, we have improved uniform error bounds for the fourth-order splitting methods.

Theorem 4.8. *Let $\Phi^n(x)$ be the numerical approximation obtained from S_{4c} (4.7.3) for (3.2.2) or S_{4RK} (4.7.5) for (4.2.2). If the time step size τ is non-resonant, i.e. there exists $0 < \kappa \leq 1$, such that $\tau \in \mathscr{A}_\kappa(\varepsilon)$, then under the assumptions (A) and (B) with $m = 4$ for the Dirac equation, or (C) and (D) with $m = 4$, $m_* = 1$ for the NLDE, then following two error estimates hold for small enough $\tau > 0$ ($s = 0$ for the Dirac equation, and $s = 1$ for the NLDE)*

$$\|\mathbf{e}^n(x)\|_{H^s} \lesssim_\kappa \tau^4 + \tau\varepsilon, \quad \|\mathbf{e}^n(x)\|_{H^s} \lesssim_\kappa \tau^4 + \tau^4/\varepsilon^5, \quad 0 \leq n \leq \frac{T}{\tau}. \tag{4.7.11}$$

As a result, there is an improved uniform error bound when $\tau > 0$ is small enough

$$\|\mathbf{e}^n(x)\|_{H^s} \lesssim_\kappa \tau^4 + \max_{0<\varepsilon\leq 1} \min\{\tau\varepsilon, \tau^4/\varepsilon^5\} \lesssim_\kappa \tau^{3/2}, \quad 0 \leq n \leq \frac{T}{\tau}. \tag{4.7.12}$$

Proof of the theorems can be extended from the proof for S_1 and S_2, and is omitted here for brevity.

4.7.2 Numerical results

In this subsection, numerical results are exhibited to validate the uniform error bounds in Theorem 4.7 and Theorem 4.8.

In all the examples, we still choose the nonlinearity and the initial values as (4.5.1) and (4.5.2), respectively. For the linear case, we take the electric potential in (3.2.2) to be time-independent as

$$V(x) = \frac{1-x}{1+x^2}, \quad x \in \Omega, \tag{4.7.13}$$

while for the nonliear case, we always take $V(x) \equiv 0$.

We first look at the case for resonant time steps. The bounded computational domain is taken as $\Omega = (-32, 32)$, the spatial mesh size is always set to be $h = \frac{1}{16}$ so that the spatial errors are negligible. The numerical 'exact' solution is generated by S_2 with a very fine time step size $\tau_e = 2\pi \times 10^{-6}$.

The discrete l^2 (for the Dirac equation) or H^1 (for the NLDE) error $e^{\varepsilon,\tau}(t_n)$ used to show the results is defined in (3.5.3) or (4.5.3) respectively. It should be close to the errors in Theorem 4.7.

Table 4.7.1 and Table 4.7.2 show the numerical temporal errors $e^{\varepsilon,\tau}(t=4\pi)$ for the fourth-order methods with different ε and time step size τ, up to time $t=4\pi$, in linear and nonlinear cases respectively.

Table 4.7.1: Discrete l^2 temporal errors $e^{\varepsilon,\tau}(t=4\pi)$ for the wave function of the Dirac equation (3.2.2) with resonant time step size, S_{4c} method.

$e^{\varepsilon,\tau}(t=4\pi)$	$\tau_0=\pi/4$	$\tau_0/2$	$\tau_0/2^2$	$\tau_0/2^3$	$\tau_0/2^4$	$\tau_0/2^5$
$\varepsilon_0=1$	1.29E-1	1.11E-3	**5.98E-5**	3.64E-6	2.26E-7	1.42E-8
order	–	6.86	**4.22**	4.04	4.01	3.99
$\varepsilon_0/2^{1/2}$	5.60E-1	1.17E-2	2.38E-4	**1.39E-5**	8.56E-7	5.34E-8
order	–	5.58	5.62	**4.09**	4.02	4.00
$\varepsilon_0/2$	8.00E-1	1.66E-1	1.16E-3	5.94E-5	**3.57E-6**	2.21E-7
order	–	2.27	7.15	4.29	**4.06**	4.01
$\varepsilon_0/2^{3/2}$	1.04	3.33E-1	2.21E-2	3.26E-4	4.38E-5	**1.11E-6**
order	–	1.65	3.91	6.08	2.90	**5.30**
$\varepsilon_0/2^2$	9.09E-1	6.41E-1	2.08E-1	3.64E-3	9.07E-4	3.39E-5
order	**–**	0.50	1.62	5.84	2.00	4.74
$\varepsilon_0/2^6$	1.06	2.26E-1	4.74E-2	3.70E-2	3.66E-2	3.65E-2
order	–	2.22	2.25	0.36	0.01	0.00
$\varepsilon_0/2^{10}$	1.06	2.16E-1	2.29E-2	3.60E-3	2.67E-3	2.44E-3
order	–	2.29	3.24	2.67	0.43	0.13
$\varepsilon_0/2^{14}$	1.06	2.16E-1	2.27E-2	2.51E-3	1.26E-3	8.00E-4
order	–	2.29	3.25	3.18	0.99	0.65
$\varepsilon_0/2^{18}$	1.06	2.16E-1	2.28E-2	2.58E-3	1.31E-3	8.48E-4
order	–	2.29	3.25	3.14	0.98	0.63
$\max_{0<\varepsilon\le 1} e^{\varepsilon,\tau}(t=4\pi)$	1.07	6.41E-1	4.39E-1	3.04E-1	2.12E-1	1.49E-1
order	–	0.74	0.55	0.53	0.52	0.51

In Table 4.7.1 and Table 4.7.2, we take the last two rows to show the maximum discrete l^2 and H^1 error respectively of each column for fixed τ, and their convergence order as before. We could clearly observe that there is a $1/2$ order uniform convergence in both cases. Moreover, although the lower bold lines as in the previous examples are hard to examine in the table, the upper bold lines are evident, i.e. when $\tau \lesssim \varepsilon^2$ (above the upper bold lines), there is always fourth order convergence. We infer that on the other side, when $\tau \gtrsim \sqrt[4]{\varepsilon}$, there should

Table 4.7.2: Discrete H^1 temporal errors $e^{\varepsilon,\tau}(t=4\pi)$ for the wave function of the NLDE (4.2.2) with resonant time step size, $S_{4\mathrm{RK}}$ method.

$e^{\varepsilon,\tau}(t=4\pi)$	$\tau_0=\pi/8$	$\tau_0/2$	$\tau_0/2^2$	$\tau_0/2^3$	$\tau_0/2^4$	$\tau_0/2^5$
$\varepsilon_0=1/2^{3/2}$	1.89E-1	**2.06E-2**	2.45E-4	7.28E-6	4.27E-7	2.60E-8
order	–	**3.19**	6.40	5.07	4.09	4.03
$\varepsilon_0/2^{1/2}$	4.42E-1	8.42E-2	**4.91E-3**	3.54E-5	1.72E-6	1.02E-7
order	–	2.39	**4.10**	7.12	4.36	4.08
$\varepsilon_0/2$	1.05	2.57E-1	5.40E-2	**2.35E-3**	8.95E-6	4.60E-7
order	–	2.04	2.25	**4.52**	8.04	4.28
$\varepsilon_0/2^{3/2}$	5.65E-1	7.37E-1	1.74E-1	3.55E-2	**1.17E-3**	2.49E-6
order	–	-0.38	2.08	2.29	**4.92**	8.88
$\varepsilon_0/2^2$	1.79E-1	3.30E-1	5.20E-1	1.19E-1	2.40E-2	**5.85E-4**
order	–	-0.88	-0.66	2.13	2.31	**5.36**
$\varepsilon_0/2^4$	1.23E-1	1.74E-1	1.83E-1	1.94E-1	3.52E-2	8.33E-2
order	–	-0.50	-0.07	-0.08	2.46	-1.24
$\varepsilon_0/2^8$	1.82E-2	4.76E-3	8.85E-4	7.19E-3	6.95E-3	1.25E-3
order	–	1.93	2.43	-3.02	0.05	2.47
$\varepsilon_0/2^{12}$	1.53E-2	1.59E-3	9.75E-5	1.55E-4	7.57E-5	5.79E-4
order	–	3.27	4.03	-0.67	1.03	-2.94
$\varepsilon_0/2^{16}$	1.53E-2	1.59E-3	3.50E-4	3.60E-4	3.49E-4	3.56E-4
order	–	3.27	2.18	-0.04	0.05	-0.03
$\max\limits_{0<\varepsilon\le 1} e^{\varepsilon,\tau}(t=4\pi)$	1.05	7.37E-1	5.20E-1	3.66E-1	2.57E-1	1.80E-1
order	–	0.52	0.50	0.51	0.51	0.51

also be fourth order convergence. However, as it may require the ε to be extremely small in order to observe this relation, we do not validate it here. These two diagonal lines indicate the two error bounds $\tau^4+\varepsilon$, and $\tau^4+\tau^4/\varepsilon^7$ in this case, which corresponds to Theorem 4.7.

To justify the improved uniform error bounds in Theorem 4.8, we further test the errors using non-resonant time step sizes, i.e., we choose $\tau\in\mathscr{A}_\kappa(\varepsilon)$ for some given ε and fixed $0<\kappa\le 1$. The bounded computational domain is set as $\Omega=(-16,16)$.

For comparison, the numerical 'exact' solution is computed by the second-order time-splitting method (S_2) with a very small time step size $\tau_e=8\times 10^{-6}$. Spatial mesh size is fixed as $h=1/16$ for all the numerical simulations.

The discrete l^2 or H^1 error $e^{\varepsilon,\tau}(t_n)$ used to show the results is defined in (3.5.3) or (4.5.3) respectively. It should be close to the errors in Theorem 4.8.

Tables 4.7.3 and 4.7.4 show the numerical temporal errors $e^{\varepsilon,\tau}(t=4)$ with different ε and time step size τ for S_{4c} and S_{4RK} respectively, up to time $t=4$.

Table 4.7.3: Discrete l^2 temporal errors $e^{\varepsilon,\tau}(t=4)$ for the wave function of the Dirac equation (3.2.2) with non-resonant time step size, S_{4c} method.

$e^{\varepsilon,\tau}(t=4)$	$\tau_0=1/2$	$\tau_0/2$	$\tau_0/2^2$	$\tau_0/2^3$	$\tau_0/2^4$	$\tau_0/2^5$
$\varepsilon_0=1$	5.40E-3	**1.70E-4**	1.02E-5	6.28E-7	3.92E-8	2.50E-9
order	–	**4.99**	4.07	4.02	4.00	3.97
$\varepsilon_0/2^{1/2}$	6.52E-2	6.79E-4	**3.78E-5**	2.31E-6	1.43E-7	8.99E-9
order	–	6.59	**4.17**	4.04	4.01	4.00
$\varepsilon_0/2$	4.20E-1	5.47E-3	1.62E-4	**9.47E-6**	5.83E-7	3.64E-8
order	–	6.26	5.08	**4.09**	4.02	4.00
$\varepsilon_0/2^{3/2}$	4.73E-1	1.34E-1	8.70E-4	4.45E-5	**2.67E-6**	1.65E-7
order	–	1.82	7.26	4.29	**4.06**	4.01
$\varepsilon_0/2^2$	3.14E-1	1.07E-1	2.15E-2	2.48E-4	1.35E-5	**8.18E-7**
order	–	1.55	2.32	6.43	4.20	**4.05**
$\varepsilon_0/2^5$	3.46E-1	6.51E-2	1.49E-2	3.13E-3	2.95E-3	1.12E-3
order	–	2.41	2.13	2.25	0.09	1.40
$\varepsilon_0/2^8$	3.39E-1	5.23E-2	4.64E-3	2.37E-3	2.35E-3	2.35E-3
order	–	2.70	3.50	0.97	0.01	0.00
$\varepsilon_0/2^{11}$	3.40E-1	5.22E-2	4.13E-3	4.56E-4	1.69E-4	7.92E-5
order	–	2.70	3.66	3.18	1.43	1.09
$\varepsilon_0/2^{14}$	3.40E-1	5.22E-2	4.06E-3	4.79E-4	1.71E-4	7.80E-5
order	–	2.70	3.68	3.08	1.48	1.14
$\varepsilon_0/2^{17}$	3.40E-1	5.22E-2	4.07E-3	4.90E-4	1.79E-4	8.47E-5
order	–	2.70	3.68	3.05	1.45	1.08
$\max_{0<\varepsilon\le 1} e^{\varepsilon,\tau}(t=4)$	4.73E-1	1.34E-1	5.00E-2	3.00E-2	1.11E-2	3.70E-3
order	–	1.82	1.42	0.74	1.44	1.58

The last two rows of Table 4.7.3 and Table 4.7.4 show the maximum values of each column and the corresponding convergence rate. From them, we could clearly observe approximately $3/2$ order uniform convergence for S_{4c} to Dirac equation (3.2.2) and S_{4RK} to NLDE (4.2.2) under non-resonant time step sizes. Meanwhile, when $\tau \lesssim \varepsilon^2$ (above the upper

Table 4.7.4: Discrete H^1 temporal errors $e^{\varepsilon,\tau}(t=4)$ for the wave function of the NLDE (4.2.2) with non-resonant time step size, $S_{4\mathrm{RK}}$ method.

$e^{\varepsilon,\tau}(t=4)$	$\tau_0=1/4$	$\tau_0/2$	$\tau_0/2^2$	$\tau_0/2^3$	$\tau_0/2^4$	$\tau_0/2^5$
$\varepsilon_0=1/4$	1.23E-1	**1.32E-2**	1.52E-4	4.44E-6	2.55E-7	1.56E-8
order	–	**3.23**	6.43	5.10	4.12	4.04
$\varepsilon_0/2^{1/2}$	3.65E-1	1.61E-2	**1.00E-3**	3.13E-5	1.18E-6	6.89E-8
order	–	4.50	**4.01**	5.00	4.73	4.10
$\varepsilon_0/2$	1.63E-1	6.54E-2	3.70E-3	**1.79E-4**	7.29E-6	3.39E-7
order	–	1.32	4.14	**4.37**	4.62	4.43
$\varepsilon_0/2^{3/2}$	1.21E-1	5.45E-2	4.47E-3	1.32E-3	**5.52E-5**	2.16E-6
order	–	1.15	3.61	1.76	**4.58**	4.68
$\varepsilon_0/2^2$	3.92E-2	3.13E-2	1.63E-2	6.50E-4	4.66E-4	**1.78E-5**
order	–	0.32	0.94	4.65	0.48	**4.71**
$\varepsilon_0/2^5$	3.87E-2	8.31E-2	3.50E-3	7.28E-4	4.95E-4	2.23E-4
order	–	-1.10	4.57	2.26	0.56	1.15
$\varepsilon_0/2^8$	4.10E-3	1.06E-3	5.13E-4	1.64E-4	5.96E-5	3.20E-5
order	–	1.95	1.05	1.64	1.46	0.90
$\varepsilon_0/2^{11}$	3.42E-3	5.01E-4	6.78E-6	2.76E-5	2.36E-5	4.50E-6
order	–	2.77	6.21	-2.03	0.23	2.39
$\varepsilon_0/2^{14}$	3.40E-3	1.49E-4	1.34E-5	1.22E-5	1.29E-5	1.25E-5
order	–	4.51	3.47	0.13	-0.08	0.05
$\max\limits_{0<\varepsilon\leq 1} e^{\varepsilon,\tau}(t=4)$	3.65E-1	1.07E-1	1.95E-2	5.82E-3	2.28E-3	8.66E-4
order	–	1.77	2.45	1.74	1.35	1.40

bold lines), there is always fourth order convergence. The lower bold lines as in the case of S_1 and S_2 are not observable here, but we believe when $\tau \gtrsim \sqrt[3]{\varepsilon}$, there will also be fourth order convergence. These two diagonal lines bring about the two error bounds $\tau^4+\tau\varepsilon$, and $\tau^4+\tau^4/\varepsilon^5$, as Theorem 4.8 indicates.

Through these results, we find out that for the fourth-order splitting methods $S_{4\mathrm{c}}$ and $S_{4\mathrm{RK}}$ to Dirac and nonlinear Dirac equations respectively, we still have uniform convergence in time, and the error bounds can be improved under non-resonant time steps, which is a similar property to S_1 and S_2.

Chapter 5

Finite Difference Time Domain (FDTD) Methods for the Dirac Equation in the Semiclassical Regime

In this chapter, we study the dynamics of the Dirac equation in the semiclassical regime, i.e., we take $\varepsilon = \nu = 1$ in (1.1.17) for $d = 1,2$, or in (1.1.7) for $d = 1,2,3$.

Here we focus on the study of finite difference time domain (FDTD) methods, which have been extensively studied for linear and nonlinear Schrödinger equation [10, 15], Klein-Gordon equation [20, 36], and Gross-Pitaevskii equation [12] previously. We implement different FDTD methods, and find out relation of the error bounds to the mesh size h, time step size τ, as well as the small parameter δ. The performance of different methods is compared through numerical examples.

5.1 The FDTD methods

For simplicity, here we consider the two-component form (1.1.17), which could be expressed as ($d = 1,2$)

$$i\delta\partial_t\Phi = \left(-i\delta\sum_{j=1}^{d}\sigma_j\partial_j + \sigma_3\right)\Phi + \left(V(t,\mathbf{x})I_2 - \sum_{j=1}^{d}A_j(t,\mathbf{x})\sigma_j\right)\Phi, \quad \mathbf{x}\in\mathbb{R}^d, \qquad (5.1.1)$$

with proper initial condition $\Phi(0,\mathbf{x}) = \Phi_0(\mathbf{x})$ for $\mathbf{x}\in\mathbb{R}^d$.

In this section, we apply four commonly used FDTD methods to the Dirac equation in the semiclassical regime (5.1.1) and analyze their stabilitiy conditions. To make the notations

simple, we only show the numerical methods and related analysis in 1D. Generalization to 2D and the four-component expression (1.1.7) is straightforward and results remain valid without modifications. Similar to most works in the literatures for the analysis and computation of the Dirac equation (cf. [14, 15, 16, 17, 27] and references therein), in practical computation, we truncate the whole space problem onto a large enough interval $\Omega = (a,b)$ such that the truncation error is negligible, and assert periodic boundary conditions. In 1D, the Dirac equation (5.1.1) with periodic boundary conditions collapses to

$$i\delta\partial_t\Phi = (-i\delta\sigma_1\partial_x + \sigma_3)\Phi + (V(t,x)I_2 - A_1(t,x)\sigma_1)\Phi, \quad x \in \Omega, \quad t > 0, \tag{5.1.2}$$

$$\Phi(t,a) = \Phi(t,b),\ \partial_x\Phi(t,a) = \partial_x\Phi(t,b),\ t \geq 0; \quad \Phi(0,x) = \Phi_0(x),\ x \in \overline{\Omega}, \tag{5.1.3}$$

where $\Phi := \Phi(t,x)$, $\Phi_0(a) = \Phi_0(b)$ and $\Phi_0'(a) = \Phi_0'(b)$.

5.1.1 The methods

Choose mesh size $h := \triangle x = \frac{b-a}{M}$ with M being an even positive integer, time step size $\tau := \triangle t > 0$ and represent the grid points and time steps as:

$$x_j := a + jh, \quad j = 0,1,...,M; \quad t_n := n\tau, \quad n = 0,1,2,... \tag{5.1.4}$$

Denote $X_M = \{U = (U_0,U_1,...,U_M)^T | U_j \in \mathbb{C}^2, j = 0,1,...,M, U_0 = U_M\}$, and take $U_{-1} = U_{M-1}$, $U_{M+1} = U_1$ if they are involved. For any $U \in X_M$, its Fourier representation can be expressed as

$$U_j = \sum_{l=-M/2}^{M/2-1} \widetilde{U}_l e^{i\mu_l(x_j-a)} = \sum_{l=-M/2}^{M/2-1} \widetilde{U}_l e^{2ijl\pi/M}, \quad j = 0,1,...,M, \tag{5.1.5}$$

where μ_l and $\widetilde{U}_l \in \mathbb{C}^2$ are defined as

$$\mu_l = \frac{2l\pi}{b-a}, \quad \widetilde{U}_l = \frac{1}{M}\sum_{j=0}^{M-1} U_j e^{-2ijl\pi/M}, \quad l = -\frac{M}{2},...,\frac{M}{2}-1. \tag{5.1.6}$$

In X_M, the standard l^2-norm is given as

$$\|U\|_{l^2}^2 = h\sum_{j=0}^{M-1} |U_j|^2, \quad U \in X_M. \tag{5.1.7}$$

Let Φ_j^n be the numerical approximation of $\Phi(t_n,x_j)$, $V_j^n = V(t_n,x_j)$, $V_j^{n+1/2} = V(t_n+\tau/2,x_j)$, $A_{1,j}^n = A_1(t_n,x_j)$ and $A_{1,j}^{n+1/2} = A_1(t_n+\tau/2,x_j)$ for $0 \le j \le M$, and $n \ge 0$. Denote $\Phi^n = (\Phi_0^n,\Phi_1^n,...,\Phi_M^n)^T \in X_M$ as the solution vector at $t = t_n$. Introduce the finite difference discretization operators for $j = 0,1,...,M$ and $n \ge 0$ as:

$$\delta_t^+\Phi_j^n = \frac{\Phi_j^{n+1}-\Phi_j^n}{\tau}, \quad \delta_t\Phi_j^n = \frac{\Phi_j^{n+1}-\Phi_j^{n-1}}{2\tau}, \quad \delta_x\Phi_j^n = \frac{\Phi_{j+1}^n-\Phi_{j-1}^n}{2h}, \tag{5.1.8}$$

and also

$$\Phi_j^{n+\frac{1}{2}} = \frac{\Phi_j^{n+1}+\Phi_j^n}{2}, \tag{5.1.9}$$

we could establish the FDTD methods.

Here we consider four frequently used FDTD methods to discretize the Dirac equation (5.1.2) for $j = 0,1,...,M-1$.

I. Leap-frog finite difference (LFFD) method, for $n \ge 1$,

$$i\delta_t\Phi_j^n = -i\sigma_1\delta_x\Phi_j^n + \frac{1}{\delta}\left(\sigma_3 + V_j^n I_2 - A_{1,j}^n\sigma_1\right)\Phi_j^n. \tag{5.1.10}$$

II. Semi-implicit finite difference (SIFD1) method, for $n \ge 1$,

$$i\delta_t\Phi_j^n = -i\sigma_1\delta_x\Phi_j^n + \frac{1}{\delta}\left(\sigma_3 + V_j^n I_2 - A_{1,j}^n\sigma_1\right)\frac{\Phi_j^{n+1}+\Phi_j^{n-1}}{2}. \tag{5.1.11}$$

III. Another semi-implicit finite difference (SIFD2) method, for $n \ge 1$,

$$i\delta_t\Phi_j^n = \left(-i\sigma_1\delta_x + \frac{1}{\delta}\sigma_3\right)\frac{\Phi_j^{n+1}+\Phi_j^{n-1}}{2} + \frac{1}{\delta}\left(V_j^n I_2 - A_{1,j}^n\sigma_1\right)\Phi_j^n. \tag{5.1.12}$$

IV. Crank-Nicolson finite difference (CNFD) method, for $n \ge 0$,

$$i\delta_t^+\Phi_j^n = -i\sigma_1\delta_x\Phi_j^{n+\frac{1}{2}} + \frac{1}{\delta}\left(\sigma_3 + V_j^{n+\frac{1}{2}} I_2 - A_{1,j}^{n+\frac{1}{2}}\sigma_1\right)\Phi_j^{n+\frac{1}{2}}. \tag{5.1.13}$$

The initial and boundary conditions (5.1.3) for these FDTD methods are discretized as:

$$\Phi_M^{n+1} = \Phi_0^{n+1}, \quad \Phi_{-1}^{n+1} = \Phi_{M-1}^{n+1}, \quad n \ge 0; \quad \Phi_j^0 = \Phi_0(x_j), \quad j = 0,1,...,M. \tag{5.1.14}$$

Besides, by applying Taylor expansion and noticing the equation (5.1.2), we can design the first step for the LFFD (5.1.10), SIFD1 (5.1.11) and SIFD2 (5.1.12) methods as ($j = 0,1,...,M$)

$$\Phi_j^1 = \Phi_0(x_j) - \tau\sigma_1\Phi_0'(x_j) - i\sin\left(\frac{\tau}{\delta}\right)\left(\sigma_3 + V_j^0 I_2 - A_{1,j}^0\sigma_1\right)\Phi_0(x_j). \tag{5.1.15}$$

In the above, we adopt $\frac{1}{\tau}\sin\left(\frac{\tau}{\delta}\right)$ instead of $\frac{1}{\delta}$ such that (5.1.15) is second order in terms of τ for any fixed $0<\delta\leq 1$ and $\|\Phi^1\|_\infty := \max_{0\leq j\leq M}|\Phi_j^1| \lesssim 1$ for $0<\delta\leq 1$. We remark here that it can be simply replaced by 1 when $\delta = 1$.

The above four methods are all time-symmetric, i.e. they are unchanged under $\tau \leftrightarrow -\tau$ and $n+1 \leftrightarrow n-1$ for the LFFD, SIFD1 and SIFD2 methods or $n+1 \leftrightarrow n$ for the CNFD method, and their memory costs are all $O(M)$. The LFFD method (5.1.10) is explict and its computational cost per step is $O(M)$. Actually, it might be the simplest and most efficient method for the Dirac equation when $\delta = 1$. The SIFD1 method (5.1.11) is implicit, however, at each time step for $n \geq 1$, the corresponding linear system is decoupled and can be solved explicitly as

$$\Phi_j^{n+1} = \left[\left(i-\frac{\tau}{\delta}V_j^n\right)I_2 - \frac{\tau}{\delta}\left(\sigma_3 - A_{1,j}^n\sigma_1\right)\right]^{-1} H_j^n, \quad j = 0,1,...,M-1, \tag{5.1.16}$$

with $H_j^n = -2i\tau\sigma_1\delta_x\Phi_j^n + \left[\left(i+\frac{\tau}{\delta}V_j^n\right)I_2 + \frac{\tau}{\delta}\left(\sigma_3 - A_{1,j}^n\sigma_1\right)\right]\Phi_j^{n-1}$, and thus its computational cost per step is also $O(M)$.

The SIFD2 method (5.1.12) is implicit, but at each time step for $n \geq 1$, the corresponding linear system can be decoupled in the phase (Fourier) space and thus it can be solved explicitly in phase space as

$$(\widetilde{\Phi^{n+1}})_l = \left(iI_2 - \frac{\tau\sin(\mu_l h)}{h}\sigma_1 - \frac{\tau}{\delta}\sigma_3\right)^{-1} L_l^n, \quad l = -\frac{M}{2},...,\frac{M}{2}-1, \tag{5.1.17}$$

where

$$L_l^n = \left(iI_2 + \frac{\tau\sin(\mu_l h)}{h}\sigma_1 + \frac{\tau}{\delta}\sigma_3\right)(\widetilde{\Phi^{n-1}})_l + \frac{2\tau}{\delta}(\widetilde{G^n\Phi^n})_l, \tag{5.1.18}$$

and $G^n = (G_0^n, G_1^n, ...G_M^n)^T \in X_M$ with $G_j^n = V_j^n I_2 - A_{1,j}^n\sigma_1$ for $j = 0,1,...,M$, and thus its computational cost per step is $O(M\ln M)$. The CNFD method (5.1.13) is implicit and at each time step for $n \geq 0$, the corresponding linear system is coupled so that it needs to be solved through either a direct solver or an iterative solver. As a result, its computational cost per step depends heavily on the solver, and it is usually much larger than $O(M)$, especially in 2D and 3D. From the analysis on the computational cost per time step here, the LFFD method is the most efficient among the four methods and the CNFD is the most expensive one.

5.1.2 Mass and energy conservation

For the CNFD method (5.1.13), we have the following conservative properties.

Lemma 5.1. *The CNFD method* (5.1.13) *conserves the mass in the discretized level, i.e., for* $n \geq 0$*,*

$$\|\Phi^n\|_{l^2}^2 := h\sum_{j=0}^{M-1}|\Phi_j^n|^2 \equiv h\sum_{j=0}^{M-1}|\Phi_j^0|^2 = \|\Phi^0\|_{l^2}^2 = h\sum_{j=0}^{M-1}|\Phi_0(x_j)|^2. \tag{5.1.19}$$

Furthermore, if $V(t,x) = V(x)$ *and* $A_1(t,x) = A_1(x)$ *are time-independent, the CNFD method* (5.1.13) *conserves the energy as well, that is*

$$\begin{aligned} E_h^n :=& h\sum_{j=0}^{M-1}\left[-i\delta(\Phi_j^n)^*\sigma_1\delta_x\Phi_j^n + (\Phi_j^n)^*\sigma_3\Phi_j^n + V_j|\Phi_j^n|^2 - A_{1,j}(\Phi_j^n)^*\sigma_1\Phi_j^n\right] \\ \equiv& E_h^0, \quad n \geq 0, \end{aligned} \tag{5.1.20}$$

where $V_j = V(x_j)$ *and* $A_{1,j} = A_1(x_j)$ *for* $j = 0,1,...,M$*.*

Proof. (i) We first prove the mass conservation (5.1.19). Multiply both sides of (5.1.13) from left by $h\tau(\Phi_j^{n+\frac{1}{2}})^*$ and take the imaginary part, we have for $j = 0,1,...,M-1$

$$h|\Phi_j^{n+1}|^2 = h|\Phi_j^n|^2 - \frac{\tau h}{2}\left[(\Phi_j^{n+\frac{1}{2}})^*\sigma_1\delta_x\Phi_j^{n+\frac{1}{2}} + (\Phi_j^{n+\frac{1}{2}})^T\sigma_1\delta_x\overline{\Phi}_j^{n+\frac{1}{2}}\right]. \tag{5.1.21}$$

Summing up (5.1.21) for $j = 0,1,...,M-1$ and noticing the expression of Pauli matrices, we get

$$\begin{aligned} \|\Phi^{n+1}\|_{l^2}^2 =& \|\Phi^n\|_{l^2}^2 - \frac{\tau h}{2}\sum_{j=0}^{M-1}\left[(\Phi_j^{n+\frac{1}{2}})^*\sigma_1\delta_x\Phi_j^{n+\frac{1}{2}} + (\Phi_j^{n+\frac{1}{2}})^T\sigma_1\delta_x\overline{\Phi}_j^{n+\frac{1}{2}}\right] \\ =& \|\Phi^n\|_{l^2}^2 - \frac{\tau}{4}\sum_{j=0}^{M-1}\Big[(\Phi_j^{n+\frac{1}{2}})^*\sigma_1\Phi_{j+1}^{n+\frac{1}{2}} + (\Phi_j^{n+\frac{1}{2}})^T\sigma_1\overline{\Phi}_{j+1}^{n+\frac{1}{2}} \\ & - (\Phi_{j+1}^{n+\frac{1}{2}})^*\sigma_1\Phi_j^{n+\frac{1}{2}} - (\Phi_{j+1}^{n+\frac{1}{2}})^T\sigma_1\overline{\Phi}_j^{n+\frac{1}{2}}\Big] \\ =& \|\Phi^n\|_{l^2}^2, \quad n \geq 0, \end{aligned} \tag{5.1.22}$$

which immediately implies (5.1.19) by induction.

(ii) We further move on to prove the energy conservation (5.1.20). Multiply both sides of (5.1.13) from left by $2h(\Phi_j^{n+1} - \Phi_j^n)^*$ and take the real part, we have

$$\begin{aligned} 0 =& -h\mathrm{Re}\left[i(\Phi_j^{n+1} - \Phi_j^n)^*\delta_x(\Phi_j^n + \Phi_j^{n+1})\right] + \frac{h}{\delta}\left[(\Phi_j^{n+1})^*\sigma_3\Phi_j^{n+1} - (\Phi_j^n)^*\sigma_3\Phi_j^n\right] \\ & + \frac{hV_j}{\delta}(|\Phi_j^{n+1}|^2 - |\Phi_j^n|^2) - \frac{hA_{1,j}}{\delta}\left[(\Phi_j^{n+1})^*\sigma_1\Phi_j^{n+1} - (\Phi_j^n)^*\sigma_1\Phi_j^n\right]. \end{aligned} \tag{5.1.23}$$

Summing up (5.1.23) for $j = 0, 1, ..., M-1$, and noticing the summation by parts formula, we have

$$\begin{aligned} & h\sum_{j=0}^{M-1} \mathrm{Re}\left[i(\Phi_j^{n+1}-\Phi_j^n)^*\delta_x(\Phi_j^n+\Phi_j^{n+1})\right] \\ = \quad & ih\sum_{j=0}^{M-1}(\Phi_j^{n+1})^*\sigma_1\delta_x\Phi_j^{n+1} - ih\sum_{j=0}^{M-1}(\Phi_j^n)^*\sigma_1\delta_x\Phi_j^n, \end{aligned} \tag{5.1.24}$$

and

$$\begin{aligned} & -ih\sum_{j=0}^{M-1}(\Phi_j^{n+1})^*\sigma_1\delta_x\Phi_j^{n+1} + ih\sum_{j=0}^{M-1}(\Phi_j^n)^*\sigma_1\delta_x\Phi_j^n + \frac{hV_j}{\delta}\sum_{j=0}^{M-1}(|\Phi_j^{n+1}|^2-|\Phi_j^n|^2) \\ & -\frac{hA_{1,j}}{\delta}\sum_{j=0}^{M-1}\left[(\Phi_j^{n+1})^*\sigma_1\Phi_j^{n+1}-(\Phi_j^n)^*\sigma_1\Phi_j^n\right] \\ & +\frac{h}{\delta}\sum_{j=0}^{M-1}\left[(\Phi_j^{n+1})^*\sigma_3\Phi_j^{n+1}-(\Phi_j^n)^*\sigma_3\Phi_j^n\right] = 0, \end{aligned} \tag{5.1.25}$$

which directly implies (5.1.20). □

5.1.3 Linear stability conditions

In order to carry out the linear stability analysis for the FDTD methods via the von Neumann method [110], we assume in the Dirac equation (5.1.2) that $A_1(t,x) \equiv A_1^0$ and $V(t,x) \equiv V^0$ with A_1^0 and V^0 being two real constants. Then we have the following results for the FDTD methods:

Lemma 5.2. *(i) The LFFD method* (5.1.10) *is stable under the stability condition*

$$0 < \tau \le \frac{\delta h}{|V^0|h + \sqrt{h^2+(\delta+h|A_1^0|)^2}}, \quad h > 0, \quad 0 < \delta \le 1. \tag{5.1.26}$$

(ii) The SIFD1 method (5.1.11) *is stable under the stability condition*

$$0 < \tau \le h, \quad h > 0, \quad 0 < \delta \le 1. \tag{5.1.27}$$

(iii) The SIFD2 method (5.1.12) *is stable under the stability condition*

$$\tau \le \frac{\delta}{|V^0|+|A_1^0|}, \quad h > 0, \quad 0 < \delta \le 1. \tag{5.1.28}$$

(iv) The CNFD method (5.1.13) *is unconditionally stable, i.e. it is stable for any* τ, $h>0$ *and* $0<\delta\leq 1$.

Proof. (i)Plugging

$$\Phi_j^n=\sum_{l=-M/2}^{M/2-1}\xi_l^n(\widetilde{\Phi^0})_l e^{i\mu_l(x_j-a)}=\sum_{l=-M/2}^{M/2-1}\xi_l^n(\widetilde{\Phi^0})_l e^{2ijl\pi/M},\quad 0\leq j\leq M, \tag{5.1.29}$$

with $\xi_l\in\mathbb{C}$ and $(\widetilde{\Phi^0})_l$ being the amplification factor and the Fourier coefficient at $n=0$ of the l-th mode in the phase space, respectively, into (5.1.10), using the orthogonality of the Fourier series, we obtain for $l=-\frac{M}{2},...,\frac{M}{2}-1$,

$$\left|(\xi_l^2-1)I_2+2i\tau\xi_l\left[\frac{1}{\delta}(\sigma_3+V^0I_2-A_1^0\sigma_1)+\frac{\sin(\mu_l h)}{h}\sigma_1\right]\right|=0. \tag{5.1.30}$$

Substituting the Pauli matrices (1.1.3) into (5.1.30), we get that the amplification factor ξ_l satisfies

$$\xi_l^2+2i\tau\theta_l\xi_l-1=0,\quad l=-\frac{M}{2},...,\frac{M}{2}-1, \tag{5.1.31}$$

where

$$\theta_l=-\frac{V^0}{\delta}\pm\frac{1}{\delta h}\sqrt{h^2+(\delta\sin(\mu_l h)-hA_1^0)^2},\quad l=-\frac{M}{2},...,\frac{M}{2}-1. \tag{5.1.32}$$

Then the stability condition for the LFFD method (5.1.10) becomes

$$|\xi_l|\leq 1\Leftrightarrow|\tau\theta_l|\leq 1,\quad l=-\frac{M}{2},...,\frac{M}{2}-1, \tag{5.1.33}$$

which gives the condition (5.1.41).

(ii) As the implicit part is automatically stable, we only need to focus on the explicit part $i\delta_t\Phi_j^n=-i\sigma_1\delta_x\Phi_j^n$. Similar to (i), plugging (5.1.29) into this part, we have for $l=-\frac{M}{2},...,\frac{M}{2}-1$,

$$\xi_l^2\pm\frac{2i\tau\sin(\mu_l h)}{h}\xi_l-1=0. \tag{5.1.34}$$

So the stability requires $\left|\frac{\tau}{h}\sin(\mu_l h)\right|\leq 1$, which suggests $0<\tau\leq h$.

(iii) Similar to (ii), we just need to concentrate on $i\delta_t\Phi_j^n=\frac{1}{\delta}(V_j^0I_2-A_{1,j}^0\sigma_1)\Phi_j^n$. Plugging (5.1.29) into it, we have

$$\xi_l^2+\frac{2i\tau}{\delta}(V^0\pm A_1^0)-1=0,\quad l=-\frac{M}{2},...,\frac{M}{2}-1, \tag{5.1.35}$$

which gives

$$\left|\frac{\tau}{\delta}(V^0 \pm A_1^0)\right| \leq 1. \tag{5.1.36}$$

As a result, we could get the corresponding stability condition.

(iv) Similar to (i), plugging (5.1.29) into (5.1.13), we obtain for $l = -\frac{M}{2}, ..., \frac{M}{2} - 1$,

$$\left|(\xi_l - 1)I_2 + \frac{i\tau}{2}(\xi_l + 1)\left(\frac{\sin(\mu_l h)}{h}\sigma_1 + \frac{1}{\delta}(\sigma_3 + V^0 I_2 - A_1^0\sigma_1)\right)\right| = 0. \tag{5.1.37}$$

Take

$$\theta_l = -V^0 \pm \frac{1}{h}\sqrt{h^2 + (\delta\sin(\mu_l h) - hA_1^0)^2}, \quad l = -\frac{M}{2}, ..., \frac{M}{2} - 1, \tag{5.1.38}$$

then we could solve out

$$\xi_l = \frac{2\delta + i\tau\theta_l}{2\delta - i\tau\theta_l}, \quad l = -\frac{M}{2}, ..., \frac{M}{2} - 1, \tag{5.1.39}$$

which indicates $|\xi_l| = 1$ for $l = -\frac{M}{2}, ..., \frac{M}{2} - 1$, so the method is unconditionally stable. □

Remark 5.1. *For the case where the electromagnetic potentials are not constants, take*

$$V_{max} := \max_{(t,x)\in\overline{\Omega}_T} |V(t,x)|, \quad A_{1,\max} := \max_{(t,x)\in\overline{\Omega}_T} |A_1(t,x)|, \tag{5.1.40}$$

then the stability condition for LFFD becomes

$$0 < \tau \leq \frac{\delta h}{V_{max}h + \sqrt{h^2 + (\delta + hA_{1,\max})^2}}, \quad h > 0, \quad 0 < \delta \leq 1, \tag{5.1.41}$$

and the stability condition for SIFD2 becomes

$$\tau \leq \frac{\delta}{V_{max} + A_{1,\max}}, \quad h > 0, \quad 0 < \delta \leq 1, \tag{5.1.42}$$

while the stability condition for SIFD1 and CNFD remain unchanged.

5.2 Error estimates

Let $0 < T < T^*$ with T^* being the maximal existence time of the solution, and denote $\Omega_T = [0,T] \times \Omega$. To get proper error estimates, we need to assume that the exact solution of (5.1.2) satisfies

$$\begin{aligned}\Phi \in & C^3([0,T];(L^\infty(\Omega))^2) \cap C^2([0,T];(W_p^{1,\infty}(\Omega))^2) \\ & \cap C^1([0,T];(W_p^{2,\infty}(\Omega))^2) \cap C([0,T];(W_p^{3,\infty}(\Omega))^2),\end{aligned} \tag{5.2.1}$$

and

$$(E) \qquad \left\| \frac{\partial^{r+s}}{\partial t^r \partial x^s} \Phi \right\|_{L^\infty([0,T];(L^\infty(\Omega))^2)} \lesssim \frac{1}{\delta^{r+s}},\ 0 \le r \le 3,\ 0 \le r+s \le 3,\ 0 < \delta \le 1,$$

where $W_p^{m,\infty}(\Omega) = \{u | u \in W^{m,\infty}(\Omega),\ \partial_x^l u(a) = \partial_x^l u(b),\ l = 0, ..., m-1\}$ for $m \ge 1$ and here the boundary values are understood in the trace sense. In the subsequent discussion, we will omit Ω when referring to the space norm taken on Ω. In addition, we assume the electromagnetic potentials $V \in C(\overline{\Omega}_T)$ and $A_1 \in C(\overline{\Omega}_T)$ and denote

$$(F) \qquad V_{\max} := \max_{(t,x)\in\overline{\Omega}_T} |V(t,x)|, \quad A_{1,\max} := \max_{(t,x)\in\overline{\Omega}_T} |A_1(t,x)|,$$

then we could come up with the following error estimates.

5.2.1 The main results

Define the grid error function $\mathbf{e}^n = (\mathbf{e}_0^n, \mathbf{e}_1^n, ..., \mathbf{e}_M^n)^T \in X_M$ as:

$$\mathbf{e}_j^n := \Phi(t_n, x_j) - \Phi_j^n, \quad j = 0, 1, ..., M, \quad n \ge 0, \tag{5.2.2}$$

with Φ_j^n being the numerical approximations obtained from the FDTD methods, then we could prove the following error estimates the under respective stability conditions for each method.

Theorem 5.1. *Under the assumptions (E) and (F), there exist constants $h_0 > 0$ and $\tau_0 > 0$ sufficiently small and independent of δ, such that for any $0 < \delta \le 1$, when $0 < h \le h_0$, $0 < \tau \le \tau_0$ and under the stability condition* (5.1.41)*, we have the following error estimate for the LFFD method* (5.1.10) *with* (5.1.14) *and* (5.1.15)

$$\|\mathbf{e}^n\|_{l^2} \lesssim \frac{\tau^2}{\delta^3} + \frac{h^2}{\delta^3}, \quad 0 \le n \le \frac{T}{\tau}. \tag{5.2.3}$$

Theorem 5.2. *Under the assumptions (E) and (F), there exist constants $h_0 > 0$ and $\tau_0 > 0$ sufficiently small and independent of δ, such that for any $0 < \delta \le 1$, when $0 < h \le h_0$, $0 < \tau \le \tau_0$ and under the stability condition*(5.1.27)*, we have the following error estimate for the SIFD1 method* (5.1.11) *with* (5.1.14) *and* (5.1.15)

$$\|\mathbf{e}^n\|_{l^2} \lesssim \frac{\tau^2}{\delta^3} + \frac{h^2}{\delta^3}, \quad 0 \le n \le \frac{T}{\tau}. \tag{5.2.4}$$

Theorem 5.3. *Under the assumptions* (E) *and* (F)*, there exist constants* $h_0 > 0$ *and* $\tau_0 > 0$ *sufficiently small and independent of* δ*, such that for any* $0 < \delta \leq 1$*, when* $0 < h \leq h_0$*,* $0 < \tau \leq \tau_0$ *and under the stability condition* (5.1.42)*, we have the following error estimate for the SIFD2 method* (5.1.12) *with* (5.1.14) *and* (5.1.15)

$$\|\mathbf{e}^n\|_{l^2} \lesssim \frac{\tau^2}{\delta^3} + \frac{h^2}{\delta^3}, \quad 0 \leq n \leq \frac{T}{\tau}. \tag{5.2.5}$$

Theorem 5.4. *Under the assumptions* (E) *and* (F)*, there exist constants* $h_0 > 0$ *and* $\tau_0 > 0$ *sufficiently small and independent of* δ*, such that for any* $0 < \delta \leq 1$*, when* $0 < h \leq h_0$ *and* $0 < \tau \leq \tau_0$*, we have the following error estimate for the CNFD method* (5.1.13) *with* (5.1.14) *and* (5.1.15)

$$\|\mathbf{e}^n\|_{l^2} \lesssim \frac{\tau^2}{\delta^3} + \frac{h^2}{\delta^3}, \quad 0 \leq n \leq \frac{T}{\tau}. \tag{5.2.6}$$

Based on Theorem 5.1 to Theorem 5.4, the four FDTD methods studied here share the same temporal/spatial resolution capacity in the semiclassical regime. In fact, given an accuracy bound $\kappa > 0$, the δ-scalability of the four FDTD methods is:

$$\tau = O(\sqrt{\delta^3 \kappa}) = O(\delta^{3/2}), \quad h = O(\sqrt{\delta^3 \kappa}) = O(\delta^{3/2}), \quad 0 < \delta \ll 1. \tag{5.2.7}$$

Moreover, for observables like the total probability density and the current density, we can derive error estimates as follows.

Corollary 5.1. *Under the assumptions* (E) *and* (F)*, with the initial and boundary conditions* (5.1.14)*,* (5.1.15) *and respective stability conditions for LFFD, SIFD1, SIFD2, and CNFD, there exist constants* $h_0 > 0$ *and* $\tau_0 > 0$ *sufficiently small and independent of* δ*, such that for any* $0 < \delta \leq 1$*, when* $0 < h \leq h_0$ *and* $0 < \tau \leq \tau_0$*, the following error estimate on the total probability density holds for the FDTD methods* (5.1.10)-(5.1.13)

$$\|\rho^n - \rho(t_n, \cdot)\|_{l^2} \lesssim \frac{\tau^2}{\delta^3} + \frac{h^2}{\delta^3}, \quad 0 \leq n \leq \frac{T}{\tau}, \tag{5.2.8}$$

where ρ^n *is obtained from the wave function* Φ^n *through* (2.3.9) *with* $d = 1$*.*

Corollary 5.2. *Under the assumptions* (E) *and* (F)*, with the initial and boundary conditions* (5.1.14)*,* (5.1.15) *and respective stability conditions for LFFD, SIFD1, SIFD2, and CNFD, there exist constants* $h_0 > 0$ *and* $\tau_0 > 0$ *sufficiently small and independent of* δ*, such that for*

any $0<\delta\leq 1$*, when* $0<h\leq h_0$ *and* $0<\tau\leq\tau_0$*, the following error estimate on the current density holds for the FDTD methods* (5.1.10)-(5.1.13)

$$\|\mathbf{J}^n-\mathbf{J}(t_n,\cdot)\|_{l^2}\lesssim\frac{\tau^2}{\delta^3}+\frac{h^2}{\delta^3},\quad 0\leq n\leq\frac{T}{\tau},\tag{5.2.9}$$

where $\mathbf{J}^n$ *is obtained from the wave function* Φ^n *through* (2.3.10) *with* $d=1$.

5.2.2 Proof for Theorem 5.1 to Theorem 5.4

In this section, we will prove Theorem 5.1 to Theorem 5.4.

Proof for Theorem 5.1 for the LFFD method

Define the local truncation error $\widetilde{\xi}^n=(\widetilde{\xi}_0^n,\widetilde{\xi}_1^n,...,\widetilde{\xi}_M^n)^T\in X_M$ of the LFFD method (5.1.10) with (5.1.14) and (5.1.15) as follows, for $0\leq j\leq M-1$ and $n\geq 1$,

$$\widetilde{\xi}_j^n:=i\delta_t\Phi(t_n,x_j)+i\sigma_1\delta_x\Phi(t_n,x_j)-\frac{1}{\delta}\left(\sigma_3+V_j^nI_2-A_{1,j}^n\sigma_1\right)\Phi(t_n,x_j),\tag{5.2.10}$$

$$\widetilde{\xi}_j^0:=i\delta_t^+\Phi_0(x_j)+i\sigma_1\delta_x\Phi_0(x_j)-\frac{1}{\delta}\left(\sigma_3+V_j^0I_2-A_{1,j}^0\sigma_1\right)\Phi_0(x_j).\tag{5.2.11}$$

Applying the Taylor expansion in (5.2.10) and (5.2.11), we obtain for $j=0,1,...,M-1$ and $n\geq 1$,

$$\widetilde{\xi}_j^0=\frac{i}{2}\tau\partial_{tt}\Phi(\tau',x_j)+\frac{i}{6}h^2\sigma_1\partial_{xxx}\Phi_0'(x_j),\tag{5.2.12}$$

$$\widetilde{\xi}_j^n=\frac{i}{6}\tau^2\partial_{ttt}\Phi(t_n',x_j)+\frac{i}{6}h^2\sigma_1\partial_{xxx}\Phi(t_n,x_j'),\tag{5.2.13}$$

where $0<\tau'<\tau$, $t_{n-1}<t_n'<t_{n+1}$ and $x_{j-1}<x_j'<x_{j+1}$. Noticing (5.1.2) and the assumptions (E) and (F), we have

$$|\widetilde{\xi}_j^0|\lesssim\frac{\tau}{\delta^2}+\frac{h^2}{\delta^3},\quad |\widetilde{\xi}_j^n|\lesssim\frac{\tau^2}{\delta^3}+\frac{h^2}{\delta^3},\quad j=0,1,...,M-1,\quad n\geq 1,\tag{5.2.14}$$

which immediately implies for $n\geq 1$

$$\|\widetilde{\xi}^n\|_{l^\infty}=\max_{0\leq j\leq M-1}|\widetilde{\xi}_j^n|\lesssim\frac{\tau^2}{\delta^3}+\frac{h^2}{\delta^3},\quad \|\widetilde{\xi}^n\|_{l^2}\lesssim\|\widetilde{\xi}^n\|_{l^\infty}\lesssim\frac{\tau^2}{\delta^3}+\frac{h^2}{\delta^3}.\tag{5.2.15}$$

Subtracting (5.1.10) from (5.2.10), noticing (5.2.2), we get for $0\leq j\leq M-1$ and $n\geq 1$,

$$i\delta_t\mathbf{e}_j^n=-i\sigma_1\delta_x\mathbf{e}_j^n+\frac{1}{\delta}\left(\sigma_3+V_j^nI_2-A_{1,j}^n\sigma_1\right)\mathbf{e}_j^n+\widetilde{\xi}_j^n,\tag{5.2.16}$$

where the boundary and initial conditions are given as

$$\mathbf{e}_0^n = \mathbf{e}_M^n, \quad \mathbf{e}_{-1}^n = \mathbf{e}_{M-1}^n, \quad n \geq 0, \quad \mathbf{e}_j^0 = \mathbf{0}, \quad j = 0, 1, ..., M. \tag{5.2.17}$$

For the first step, we have $\frac{i}{\tau}\mathbf{e}_j^1 = \widetilde{\xi}_j^0$ for $j = 0, 1, ..., M$, so

$$\|\mathbf{e}^1\|_{l^2} = \tau \|\widetilde{\xi}_j^0\|_{l^2} \lesssim \frac{\tau^2}{\delta^2} + \frac{\tau h^2}{\delta^3} \lesssim \frac{\tau^2}{\delta^3} + \frac{h^2}{\delta^3}. \tag{5.2.18}$$

Furthermore, multiply $2\tau h \left(\mathbf{e}_j^{n+1} + \mathbf{e}_j^{n-1}\right)^*$ from left on both sides to (5.2.16), then sum up from $j = 0$ to $j = M-1$, and take the imaginary part, we have

$$\mathscr{E}^{n+1} - \mathscr{E}^n = 2\tau h \mathrm{Im}\left[\sum_{j=0}^{M-1} \left(\mathbf{e}_j^{n+1} + \mathbf{e}_j^{n-1}\right)^* \widetilde{\xi}_j^n\right], \tag{5.2.19}$$

where $\mathscr{E}^n$ for $n = 0, 1, ...$ is denoted as

$$\begin{aligned}\mathscr{E}^{n+1} =& \|\mathbf{e}^{n+1}\|_{l^2}^2 + \|\mathbf{e}^n\|_{l^2}^2 + \frac{\tau}{h}\mathrm{Re}\left(h\sum_{j=0}^{M-1}(\mathbf{e}_j^{n+1})^* \sigma_1 (\mathbf{e}_{j+1}^n - \mathbf{e}_{j-1}^n)\right) \\ & - \frac{2\tau}{\delta}\mathrm{Im}\left(h\sum_{j=0}^{M-1}(\mathbf{e}_j^{n+1})^* \left(\sigma_3 + V_j^n I_2 - A_{1,j}^n \sigma_1\right)\mathbf{e}_j^n\right).\end{aligned} \tag{5.2.20}$$

Consequently, we have

$$\begin{aligned}\mathscr{E}^{n+1} - \mathscr{E}^n \lesssim & \tau h \sum_{j=0}^{M-1}\left(|\mathbf{e}_j^{n+1}| + |\mathbf{e}_j^{n-1}|\right)|\widetilde{\xi}_j^n| \lesssim \tau\left(\|\mathbf{e}^n\|_{l^2}^2 + \|\mathbf{e}^{n-1}\|_{l^2}^2\right) + \tau\|\widetilde{\xi}^n\|_{l^2}^2 \\ \lesssim & \tau(\mathscr{E}^{n+1} + \mathscr{E}^n) + \tau\left(\frac{\tau^2}{\delta^3} + \frac{h^2}{\delta^3}\right)^2,\end{aligned} \tag{5.2.21}$$

by noticing (5.2.15). Summing the above inequality for $n = 1, 2, ..., m-1$, we get

$$\mathscr{E}^m - \mathscr{E}^1 \lesssim \tau \sum_{k=1}^{m} \mathscr{E}^k + m\tau\left(\frac{h^2}{\delta^3} + \frac{\tau^2}{\delta^3}\right)^2, \quad 1 \leq m \leq \frac{T}{\tau}. \tag{5.2.22}$$

Under the stability condition (5.1.41) $\tau \leq \frac{\delta\tau_1 h}{|V^0|h + \sqrt{h^2 + (\delta + h|A_1^0|)^2}}$, if we take $\tau_1 = \frac{1}{4}$, we could derive $\frac{\tau}{h} \leq \frac{1}{4}$, and $\frac{\tau}{\delta}(1 + |V^0| + |A_1^0|) \leq \frac{1}{4}$, which gives

$$\frac{1}{2}\left(\|\mathbf{e}^{n+1}\|_{l^2}^2 + \|\mathbf{e}^n\|_{l^2}^2\right) \leq \mathscr{E}^{n+1} \leq \frac{3}{2}\left(\|\mathbf{e}^{n+1}\|_{l^2}^2 + \|\mathbf{e}^n\|_{l^2}^2\right), \quad n \geq 0, \tag{5.2.23}$$

by using Cauchy inequality. Then from (5.2.18), we have

$$\mathscr{E}^1 \lesssim \left(\frac{h^2}{\delta^3}+\frac{\tau^2}{\delta^3}\right)^2. \tag{5.2.24}$$

So if we take τ_0 sufficiently small, the under the discrete Gronwall's inequality for (5.2.22), it can be obtained that

$$\mathscr{E}^m \lesssim \left(\frac{h^2}{\delta^3}+\frac{\tau^2}{\delta^3}\right)^2, \quad 1 \le m \le \frac{T}{\tau}, \tag{5.2.25}$$

which immediately implies the error bound (5.2.3) in view of (5.2.23). □

The ideas of proof for Theorem 5.2 to Theorem 5.4 are similar to the proof for Theorem 5.1, so for brevity, here we only show an outline of the proof.

The outline of proof for Theorem 5.2 for the SIFD1 method

The local truncation error for the first step is the same as (5.2.12), and by using Taylor expansion, we also have for $n \ge 1$

$$\widetilde{\xi}_j^n = \frac{i\tau^2}{6}\partial_{ttt}\Phi(t_n',x_j)+\frac{ih^2}{6}\sigma_1\partial_{xxx}\Phi(t_n,x_j')-\frac{\tau^2}{2\delta}\left(\sigma_3+V_j^nI_2-A_{1,j}^n\sigma_1\right)\partial_{tt}\Phi(t_n'',x_j),$$

where $t_{n-1}<t_n',t_n''<t_{n+1}$, $x_{j-1}<x_j'<x_{j+1}$. Noticing (5.1.2) and the assumptions (E) and (F), we can have for $n \ge 1$

$$\|\widetilde{\xi}^n\|_{l^\infty} = \max_{0\le j\le M-1}|\widetilde{\xi}_j^n| \lesssim \frac{\tau^2}{\delta^3}+\frac{h^2}{\delta^3}, \quad \|\widetilde{\xi}^n\|_{l^2} \lesssim \|\widetilde{\xi}^n\|_{l^\infty} \lesssim \frac{\tau^2}{\delta^3}+\frac{h^2}{\delta^3}. \tag{5.2.26}$$

Noticing (5.2.2), we get for the error function with $0 \le j \le M-1$ and $n \ge 1$,

$$i\delta_t\mathbf{e}_j^n = -i\sigma_1\delta_x\mathbf{e}_j^n+\frac{1}{2\delta}\left(\sigma_3+V_j^nI_2-A_{1,j}^n\sigma_1\right)(\mathbf{e}_j^{n+1}+\mathbf{e}_j^{n-1})+\widetilde{\xi}_j^n, \tag{5.2.27}$$

where the boundary and initial conditions are taken as before, and for the first step, we still have

$$\|\mathbf{e}^1\|_{l^2} = \tau\|\widetilde{\xi}_j^0\|_{l^2} \lesssim \frac{\tau^2}{\delta^2}+\frac{\tau h^2}{\delta^3} \lesssim \frac{\tau^2}{\delta^3}+\frac{h^2}{\delta^3}. \tag{5.2.28}$$

Denote

$$\mathscr{E}^{n+1} = \|\mathbf{e}^{n+1}\|_{l^2}^2+\|\mathbf{e}^n\|_{l^2}^2+\frac{\tau}{h}\mathrm{Re}\left(h\sum_{j=0}^{M-1}(\mathbf{e}_j^{n+1})^*\sigma_1(\mathbf{e}_{j+1}^n-\mathbf{e}_{j-1}^n)\right), \tag{5.2.29}$$

for $n \geq 0$, multiply $2\tau h\left(\mathbf{e}_j^{n+1}+\mathbf{e}_j^{n-1}\right)^*$ from left on both sides to (5.2.27), then sum up from $j=0$ to $j=M-1$, and take the imaginary part, we have

$$\begin{aligned}\mathscr{E}^{n+1}-\mathscr{E}^n &= 2\tau h \mathrm{Im}\left[\sum_{j=0}^{M-1}\left(\mathbf{e}_j^{n+1}+\mathbf{e}_j^{n-1}\right)^*\widetilde{\xi}_j^n\right]\\ &\lesssim \tau(\mathscr{E}^{n+1}+\mathscr{E}^n)+\tau\left(\frac{\tau^2}{\delta^3}+\frac{h^2}{\delta^3}\right)^2, \end{aligned} \tag{5.2.30}$$

by noticing (5.2.26). Under the stability condition (5.1.27), if we take $\frac{\tau}{h}\leq\frac{1}{2}$, we could derive

$$\frac{1}{2}\left(\|\mathbf{e}^{n+1}\|_{l^2}^2+\|\mathbf{e}^n\|_{l^2}^2\right)\leq\mathscr{E}^{n+1}\leq\frac{3}{2}\left(\|\mathbf{e}^{n+1}\|_{l^2}^2+\|\mathbf{e}^n\|_{l^2}^2\right),\quad n\geq 0, \tag{5.2.31}$$

by using Cauchy inequality as before. Then following the same process as in the proof for Theorem 5.1, it can be obtained that for sufficiently small τ_0

$$\mathscr{E}^m\lesssim\left(\frac{h^2}{\delta^3}+\frac{\tau^2}{\delta^3}\right)^2,\quad 1\leq m\leq\frac{T}{\tau}, \tag{5.2.32}$$

which immediately implies the error bound (5.2.4) in view of (5.2.23). □

The outline of proof for Theorem 5.3 for the SIFD2 method

The local truncation error for the first step is the same as (5.2.12), and by using Taylor expansion, we also have for $n\geq 1$

$$\begin{aligned}\widetilde{\xi}_j^n =& \frac{i\tau^2}{6}\partial_{ttt}\Phi(t_n',x_j)+\frac{i\tau^2}{2}\sigma_1\partial_{xtt}\Phi(t_n'',x_j)+\frac{ih^2}{12}\sigma_1\partial_{xxx}\Phi(t_{n+1},x_j')\\ &+\frac{ih^2}{12}\sigma_1\partial_{xxx}\Phi(t_{n-1},x_j'')+\frac{\tau^2}{2\delta}\sigma_3\partial_{tt}\Phi(t_n''',x_j),\end{aligned} \tag{5.2.33}$$

where $t_{n-1}<t_n',t_n'',t_n'''<t_{n+1}$, $x_{j-1}<x_j',x_j''<x_{j+1}$. Noticing (5.1.2) and the assumptions (E) and (F), we can have for $n\geq 1$

$$\|\widetilde{\xi}^n\|_{l^\infty}=\max_{0\leq j\leq M-1}|\widetilde{\xi}_j^n|\lesssim\frac{\tau^2}{\delta^3}+\frac{h^2}{\delta^3},\quad \|\widetilde{\xi}^n\|_{l^2}\lesssim\|\widetilde{\xi}^n\|_{l^\infty}\lesssim\frac{\tau^2}{\delta^3}+\frac{h^2}{\delta^3}. \tag{5.2.34}$$

Noticing (5.2.2), we get for the error function with $0\leq j\leq M-1$ and $n\geq 1$,

$$\begin{aligned}i\delta_t\mathbf{e}_j^n = & -\frac{i}{2}\sigma_1\delta_x\left(\mathbf{e}_j^{n+1}+\mathbf{e}_j^{n-1}\right)+\frac{1}{2\delta}\sigma_3\left(\mathbf{e}_j^{n+1}+\mathbf{e}_j^{n-1}\right)\\ &+\frac{1}{\delta}\left(V_j^nI_2-A_{1,j}^n\sigma_1\right)\mathbf{e}_j^n+\widetilde{\xi}_j^n,\end{aligned} \tag{5.2.35}$$

where the boundary and initial conditions are taken as before, and for the first step, we still have

$$\|\mathbf{e}^1\|_{l^2} = \tau\|\widetilde{\xi}_j^0\|_{l^2} \lesssim \frac{\tau^2}{\delta^2} + \frac{\tau h^2}{\delta^3} \lesssim \frac{\tau^2}{\delta^3} + \frac{h^2}{\delta^3}. \tag{5.2.36}$$

Denote

$$\begin{aligned}\mathscr{E}^{n+1} =& \|\mathbf{e}^{n+1}\|_{l^2}^2 + \|\mathbf{e}^n\|_{l^2}^2 + \frac{\tau h}{2}\mathrm{Re}\left((\mathbf{e}_j^{n+1})^*\sigma_1\delta_x\mathbf{e}_j^{n+1} + (\mathbf{e}_j^n)^*\sigma_1\delta_x\mathbf{e}_j^n\right) \\ & - \frac{2\tau}{\delta}\mathrm{Im}\left(h\sum_{j=0}^{M-1}(\mathbf{e}_j^{n+1})^*\left(V_j^n I_2 - A_{1,j}^n\sigma_1\right)\mathbf{e}_j^n\right), \quad n \geq 0,\end{aligned} \tag{5.2.37}$$

multiply $2\tau h\left(\mathbf{e}_j^{n+1} + \mathbf{e}_j^{n-1}\right)^*$ from left on both sides to (5.2.35), then sum up from $j = 0$ to $j = M-1$, and take the imaginary part, we have

$$\begin{aligned}\mathscr{E}^{n+1} - \mathscr{E}^n &= 2\tau h\mathrm{Im}\left[\sum_{j=0}^{M-1}\left(\mathbf{e}_j^{n+1} + \mathbf{e}_j^{n-1}\right)^*\widetilde{\xi}_j^n\right] \\ &\lesssim \tau(\mathscr{E}^{n+1} + \mathscr{E}^n) + \tau\left(\frac{\tau^2}{\delta^3} + \frac{h^2}{\delta^3}\right)^2,\end{aligned} \tag{5.2.38}$$

by noticing (5.2.34). Under the stability condition (5.1.42), if we take $\frac{\tau}{h} \leq \frac{1}{2}$, and $\frac{\tau}{\delta}(|V^0| + |A_1^0|) \leq \frac{1}{4}$, we could derive

$$\frac{1}{2}\left(\|\mathbf{e}^{n+1}\|_{l^2}^2 + \|\mathbf{e}^n\|_{l^2}^2\right) \leq \mathscr{E}^{n+1} \leq \frac{3}{2}\left(\|\mathbf{e}^{n+1}\|_{l^2}^2 + \|\mathbf{e}^n\|_{l^2}^2\right), \quad n \geq 0, \tag{5.2.39}$$

by using Cauchy inequality as before. Then following the same process as in the proof for Theorem 5.1, it can be obtained that for sufficiently small τ_0

$$\mathscr{E}^m \lesssim \left(\frac{h^2}{\delta^3} + \frac{\tau^2}{\delta^3}\right)^2, \quad 1 \leq m \leq \frac{T}{\tau}, \tag{5.2.40}$$

which immediately implies the error bound (5.2.5) in view of (5.2.23). □

The outline of proof for Theorem 5.4 for the CNFD method

By using Taylor expansion, we can get the local truncation error for $n \geq 0$

$$\begin{aligned}\widetilde{\xi}_j^n =& \frac{i\tau^2}{6}\partial_{ttt}\Phi(t_n', x_j) + \frac{ih^2}{12}\sigma_1\partial_{xxx}\Phi(t_n, x_j') + \frac{ih^2}{12}\sigma_1\partial_{xxx}\Phi(t_{n+1}, x_j'') \\ & + \frac{i\tau^2}{4}\partial_{xtt}\Phi(t_n'', x_j) - \frac{\tau^2}{4\delta}\left(\sigma_3 + V_j^{n+\frac{1}{2}}I_2 - A_{1,j}^{n+\frac{1}{2}}\sigma_1\right)\partial_{tt}\Phi(t_n''', x_j),\end{aligned} \tag{5.2.41}$$

where $t_{n-1} < t_n', t_n'', t_n''' < t_{n+1}$, $x_{j-1} < x_j', x_j'' < x_{j+1}$. Noticing (5.1.2) and the assumptions (E) and (F), we can have for $n \geq 1$

$$\|\widetilde{\xi}^n\|_{l^\infty} = \max_{0\leq j\leq M-1} |\widetilde{\xi}_j^n| \lesssim \frac{\tau^2}{\delta^3} + \frac{h^2}{\delta^3}, \quad \|\widetilde{\xi}^n\|_{l^2} \lesssim \|\widetilde{\xi}^n\|_{l^\infty} \lesssim \frac{\tau^2}{\delta^3} + \frac{h^2}{\delta^3}. \tag{5.2.42}$$

Noticing (5.2.2), we get for the error function with $0 \leq j \leq M-1$ and $n \geq 0$,

$$i\delta_t^+ \mathbf{e}_j^n = -i\sigma_1 \delta_x \mathbf{e}_j^{n+1/2} + \frac{1}{\delta}\left(\sigma_3 + V_j^{n+1/2} I_2 - A_{1,j}^{n+1/2}\sigma_1\right)\mathbf{e}_j^{n+1/2} + \widetilde{\xi}_j^n, \tag{5.2.43}$$

where the boundary and initial conditions are taken as before.
Multiply $\tau h\left(\mathbf{e}_j^{n+1} + \mathbf{e}_j^n\right)^*$ from left on both sides to (5.2.43), then sum up from $j = 0$ to $j = M-1$, and take the imaginary part, we have

$$\begin{aligned} \|\mathbf{e}^{n+1}\|_{l^2}^2 - \|\mathbf{e}^n\|_{l^2}^2 &= \tau h \mathrm{Im}\left[\sum_{j=0}^{M-1}\left(\mathbf{e}_j^{n+1} + \mathbf{e}_j^n\right)^* \widetilde{\xi}_j^n\right] \\ &\lesssim \tau(\|\mathbf{e}^{n+1}\|_{l^2}^2 + \|\mathbf{e}^n\|_{l^2}^2) + \tau\left(\frac{\tau^2}{\delta^3} + \frac{h^2}{\delta^3}\right)^2, \end{aligned} \tag{5.2.44}$$

by noticing (5.2.42). Then summing up from $n = 0$ to $n = m-1$, by applying the discrete Gronwall's inequality, for sufficiently small τ_0, we could obtain

$$\|\mathbf{e}^m\|_{l^2}^2 \lesssim \left(\frac{h^2}{\delta^3} + \frac{\tau^2}{\delta^3}\right)^2, \quad 1 \leq m \leq \frac{T}{\tau}, \tag{5.2.45}$$

which is the error bound (5.2.6) in view of (5.2.23). □

5.3 Numerical results

In this section, we study numerically the spatial and temporal resolution of the FDTD methods for the Dirac equation in the semiclassical regime, where the solution propagates waves with wavelength at $O(\delta)$ in both space and time. In the example, we take $d = 1$, and the electromagnetic potentials to be

$$V(t,x) = \frac{1-x}{1+x^2}, \quad A_1(t,x) = \frac{(x+1)^2}{1+x^2}, \quad x \in \mathbb{R}. \tag{5.3.1}$$

To quantify the numerical errors, we use the following representations of relative errors of the wave function Φ, the total probability density ρ and the current density $\mathbf{J}$

$$\begin{aligned} e^r_{\Phi}(t_n) =& \frac{\|\Phi^n - \Phi(t_n,\cdot)\|_{l^2}}{\|\Phi(t_n,\cdot)\|_{l^2}}, \quad e^r_{\rho}(t_n) = \frac{\|\rho^n - \rho(t_n,\cdot)\|_{l^2}}{\|\rho(t_n,\cdot)\|_{l^2}}, \\ & e^r_{\mathbf{J}}(t_n) = \frac{\|\mathbf{J}^n - \mathbf{J}(t_n,\cdot)\|_{l^2}}{\|\mathbf{J}(t_n,\cdot)\|_{l^2}}, \end{aligned} \tag{5.3.2}$$

where ρ^n and $\mathbf{J}^n$ can be computed from the numerical solution of the wave function at the nth time step Φ^n via (2.3.9) and (2.3.10) with $d = 1$, respectively.

For the initial condition, here we take

$$\begin{aligned} \phi_1(0,x) &= \frac{1}{2} e^{-4x^2} e^{iS_0(x)/\delta} \left(1 + \sqrt{1 + S_0'(x)^2}\right), \\ \phi_2(0,x) &= \frac{1}{2} e^{-4x^2} e^{iS_0(x)/\delta} S_0'(x), \quad x \in \mathbb{R}, \end{aligned} \tag{5.3.3}$$

for $\delta \in (0,1]$, with

$$S_0(x) = \frac{1}{40}\big(1 + \cos(2\pi x)\big), \qquad x \in \mathbb{R}. \tag{5.3.4}$$

As previously stated, the problem is solved numerically on a bounded domain $\Omega = (-16,16)$, i.e. with $a = -16$ and $b = 16$. Moreover, because the exact solution is not known, here we use the fourth-order compact time-splitting (S_{4c}) Fourier pseudospectral method put forward in Chapter 2 with a very fine mesh size $h = h_e = 1/4096$ and a very small time step size $\tau = \tau_e = 10^{-4}$ to get the numerical 'exact' solution for comparison.

In Table 5.3.1 to Table 5.3.4, relative spatial and temporal errors of the wave function $e^r_{\Phi}(t = 2)$ using the four finite difference methods LFFD (5.1.10), SIFD1 (5.1.11), SIFD2 (5.1.12), and CNFD (5.1.13) are presented respectively. Here for simplicity and considering the stability conditions, we let the mesh size h and time step size τ decrease simultaneously.

From these tables, we can observe second order convergence in space and time for all the four methods LFFD, SIFD1, SIFD2, and CNFD with any $\delta \in (0,1]$ (cf. each row in Table 5.3.1 to Table 5.3.4). The δ-resolution of these methods are all $h = O(\delta^{3/2})$ and $\tau = O(\delta^{3/2})$, which is verified through the upper triangles of each table above the bold diagonal line. This corresponds well with our error estimates in Theorem 5.1 to Theorem 5.4. Moreover, the numerical solutions from LFFD and SIFD2 are unstable with small δ and relative large τ, because in stability conditions (5.1.41) and (5.1.42), the restrictions on τ become more strict

Table 5.3.1: Discrete l^2 relative spatial and temporal errors for the wave function $e^r_\Phi(t=2)$ using the LFFD method.

$e^r_\Phi(t=2)$	$\tau_0 = 0.1$ $h_0 = 1/8$	$\tau_0/4$ $h_0/4$	$\tau_0/4^2$ $h_0/4^2$	$\tau_0/4^3$ $h_0/4^3$	$\tau_0/4^4$ $h_0/4^4$
$\delta_0 = 1$	**2.83E-1**	1.13E-2	7.17E-4	4.49E-5	2.81E-6
order	**–**	2.32	1.99	2.00	2.00
$\delta_0/4^{2/3}$	Unstable	**5.43E-2**	3.28E-3	2.05E-4	1.28E-5
order	–	**–**	2.02	2.00	2.00
$\delta_0/4^{4/3}$	Unstable	Unstable	**1.79E-2**	1.11E-3	6.92E-5
order	–	–	**–**	2.01	2.00
$\delta_0/4^2$	Unstable	Unstable	Unstable	**1.05E-2**	6.57E-4
order	–	–	–	**–**	2.00
$\delta_0/4^{8/3}$	Unstable	Unstable	Unstable	1.38E-1	**8.48E-3**
order	–	–	–	–	**2.01**

Table 5.3.2: Discrete l^2 relative spatial and temporal errors for the wave function $e^r_\Phi(t=2)$ using the SIFD1 method.

$e^r_\Phi(t=2)$	$\tau_0 = 0.1$ $h_0 = 1/8$	$\tau_0/4$ $h_0/4$	$\tau_0/4^2$ $h_0/4^2$	$\tau_0/4^3$ $h_0/4^3$	$\tau_0/4^4$ $h_0/4^4$
$\delta_0 = 1$	**1.85E-1**	1.04E-2	6.41E-4	4.01E-5	2.50E-6
order	**–**	2.08	2.01	2.00	2.00
$\delta_0/4^{2/3}$	9.16E-1	**6.66E-2**	4.12E-3	2.57E-4	1.61E-5
order	–	**1.89**	2.01	2.00	2.00
$\delta_0/4^{4/3}$	1.70	8.17E-1	**5.54E-2**	3.47E-3	2.17E-4
order	–	0.53	**1.94**	2.00	2.00
$\delta_0/4^2$	1.69	1.11	8.19E-1	**5.49E-2**	3.43E-3
order	–	0.30	0.22	**1.95**	2.00
$\delta_0/4^{8/3}$	1.44	1.58	1.40	8.26E-1	**5.51E-2**
order	–	-0.07	0.09	0.38	**1.95**

Table 5.3.3: Discrete l^2 relative spatial and temporal errors for the wave function $e^r_\Phi(t=2)$ using the SIFD2 method.

$e^r_\Phi(t=2)$	$\tau_0 = 0.1$ $h_0 = 1/8$	$\tau_0/4$ $h_0/4$	$\tau_0/4^2$ $h_0/4^2$	$\tau_0/4^3$ $h_0/4^3$	$\tau_0/4^4$ $h_0/4^4$
$\delta_0 = 1$	3.82E-1	**3.95E-2**	2.55E-3	1.60E-4	9.98E-6
order	–	**1.64**	1.98	2.00	2.00
$\delta_0/4^{2/3}$	7.72E-1	1.21E-1	**8.01E-3**	5.01E-4	3.13E-5
order	–	1.33	**1.96**	2.00	2.00
$\delta_0/4^{4/3}$	Unstable	4.72E-1	4.21E-2	**2.66E-3**	1.66E-4
order	–	–	1.74	**1.99**	2.00
$\delta_0/4^2$	Unstable	1.24	3.14E-1	2.09E-2	**1.31E-3**
order	–	–	0.99	1.95	**2.00**
$\delta_0/4^{8/3}$	Unstable	Unstable	1.11	2.68E-1	1.69E-2
order	–	–	–	1.02	1.99

Table 5.3.4: Discrete l^2 relative spatial and temporal errors for the wave function $e^r_\Phi(t=2)$ using the CNFD method.

$e^r_\Phi(t=2)$	$\tau_0 = 0.1$ $h_0 = 1/8$	$\tau_0/4$ $h_0/4$	$\tau_0/4^2$ $h_0/4^2$	$\tau_0/4^3$ $h_0/4^3$	$\tau_0/4^4$ $h_0/4^4$
$\delta_0 = 1$	**2.82E-1**	2.41E-2	1.52E-3	9.53E-5	5.96E-6
order	**–**	1.77	1.99	2.00	2.00
$\delta_0/4^{2/3}$	6.77E-1	**8.08E-2**	5.12E-3	3.20E-4	2.00E-5
order	–	**1.53**	1.99	2.00	2.00
$\delta_0/4^{4/3}$	1.18	4.20E-1	**3.09E-2**	1.93E-3	1.21E-4
order	–	0.74	**1.88**	2.00	2.00
$\delta_0/4^2$	1.14	9.62E-1	3.20E-1	**2.06E-2**	1.29E-3
order	–	0.12	0.79	**1.98**	2.00
$\delta_0/4^{8/3}$	1.08	1.14	9.10E-1	3.03E-1	**1.92E-2**
order	–	-0.03	0.16	0.79	**1.99**

with smaller δ. Comparatively, SIFD1 and CNFD do not suffer from the stability problem, in that CNFD is unconditionally stable, and the stability condition for SIFD1 only requires that $0 < \tau \leq h$ (5.1.27), which is always satisfied in our computation (cf. each column in Table 5.3.1 to Table 5.3.4).

We also test the relative spatial and temporal errors of the total probability $e_\rho^r(t=2)$ and of the current density $e_{\mathbf{J}}^r(t=2)$ using the four methods. As the results for these methods are similar, we only show the errors obtained by using the CNFD method as follows.

Table 5.3.5: Discrete l^2 relative spatial and temporal errors for the total probability $e_\rho^r(t=2)$ using the CNFD method.

$e_\rho^r(t=2)$	$\tau_0=0.1$ $h_0=1/8$	$\tau_0/4$ $h_0/4$	$\tau_0/4^2$ $h_0/4^2$	$\tau_0/4^3$ $h_0/4^3$	$\tau_0/4^4$ $h_0/4^4$
$\delta_0=1$	**3.87E-1**	3.07E-2	2.04E-3	1.28E-4	8.02E-6
order	**–**	1.83	1.96	2.00	2.00
$\delta_0/4^{2/3}$	8.18E-1	**8.54E-2**	5.54E-3	3.46E-4	2.16E-5
order	–	**1.63**	1.97	2.00	2.00
$\delta_0/4^{4/3}$	1.34	2.80E-1	**1.81E-2**	1.13E-3	7.06E-5
order	–	1.13	**1.98**	2.00	2.00
$\delta_0/4^2$	1.42	1.24	1.33E-1	**8.27E-3**	5.17E-4
order	–	0.10	1.61	**2.00**	2.00
$\delta_0/4^{8/3}$	1.23	1.38	1.21	9.91E-2	**6.18E-3**
order	–	-0.08	0.09	1.81	**2.00**

Table 5.3.6: Discrete l^2 relative spatial and temporal errors for the current density $e_{\mathbf{J}}^r(t=2)$ using the CNFD method.

$e_{\mathbf{J}}^r(t=2)$	$\tau_0=0.1$ $h_0=1/8$	$\tau_0/4$ $h_0/4$	$\tau_0/4^2$ $h_0/4^2$	$\tau_0/4^3$ $h_0/4^3$	$\tau_0/4^4$ $h_0/4^4$
$\delta_0=1$	**4.48E-1**	3.67E-2	2.42E-3	1.52E-4	9.48E-6
order	**–**	1.80	1.96	2.00	2.00
$\delta_0/4^{2/3}$	9.90E-1	**1.07E-1**	6.93E-3	4.33E-4	2.71E-5
order	–	**1.61**	1.97	2.00	2.00
$\delta_0/4^{4/3}$	1.29	3.60E-1	**2.47E-2**	1.54E-3	9.64E-5
order	–	0.92	**1.93**	2.00	2.00
$\delta_0/4^2$	1.22	1.22	1.52E-1	**9.56E-3**	5.98E-4
order	–	0.00	1.50	**1.99**	2.00
$\delta_0/4^{8/3}$	1.16	1.25	1.21	1.05E-1	**6.54E-3**
order	–	-0.06	0.02	1.77	**2.00**

Table 5.3.5 and Table 5.3.6 respectively display the relative errors for the total probability and the current density. We can observe that the results in both cases have similar patterns with the relative errors for the wave function using the CNFD method (cf. Table 5.3.4). More specifically, there is always second order convergence in space and time for $\delta \in (0,1]$ (cf. each row in Table 5.3.5 and Table 5.3.6); and the δ-scalability for both total probability and current density is $h = O(\delta^{3/2})$ and $\tau = O(\delta^{3/2})$ (cf. the upper triangles above the bold diagonal lines), which coincides with the Corollaries 5.1 and 5.2. As mentioned before, the other three finite difference methods LFFD, SIFD1 and SIFD2 will generate similar results.

From the numerical results presented in this section, we successfully justify our error estimates for wave function using the finite difference methods in Theorem 5.1 to Theorem 5.4, as well as the error estimates for total probability and current density in Corollaries 5.1 and 5.2 for Dirac equation in the semiclassical regime.

Chapter 6

Conclusion and future work

This thesis focuses on multiscale methods and corresponding analysis for the Dirac and nonlinear Dirac equation. Different regimes of the equations are taken into consideration, and we study time-splitting as well as finite difference methods in solving the dynamics. The main work in the thesis is summarized as follows.

1. Propose a new fourth-order compact time-splitting method for the Dirac equation.

To improve the performance of fourth-order splitting methods, S_{4c} is designed and applied to the Dirac equation. It reduces the computational cost by introducing a double commutator between two operators, and because there is no backward sub-step, the accuracy of S_{4c} is also better than other fourth-order methods. The method still performs much better in higher dimensions if there is no external magnetic potential. The spatial and temporal resolution of S_{4c} for different regimes are studied as well.

2. Study super-resolution of the time-splitting methods.

In the absence of magnetic potential, there is super-resolution for time-splitting methods in solving the Dirac and nonlinear Dirac equation. S_1 and S_2 are examined thoroughly in the thesis. The uniform error bounds could be improved when the time steps are taken to be non-resonant. For each case, rigorous proof is carried out, and numerical results are presented to validate the error estimates.

3. Examine the finite difference methods for the Dirac equation in the semiclassical

regime.

Four frequently used finite difference finite domain (FDTD) methods are applied to the Dirac equation in the semiclassical regime, and their stability conditions, as well as error estimates are examined in detail. It is found out that all the FDTD methods share the same spatial and temporal resolution. The comparison among them shows that LFFD is most efficient with the most strict stability condition, while CNFD is unconditionally stable, but could be very time-consuming.

Some future work is listed below:

- Apply the exponential wave integrator Fourier pseudospectral (EWI-FP) method, and the time-splitting Fourier pseudospectral method (TSFP) to the Dirac/nonlinear Dirac equation in the semiclassical regime, and study their error bounds. Moreover, it is challenging to put forward a uniform accurate method in this regime.

- Propose suitable numerical methods, including the finite difference, EWI and time-splitting methods to solve equations related to the Dirac equation, such as the Weyl and the Majorana equation. Different regimes of these equations could also be considered.

- Study the Dirac/nonlinear Dirac equation for many-body systems, which may bring about more insight into physical systems.

- Find out the possible application of our research in physics, such as in graphene and other 2D materials. The link could likely be found from the relation of the Dirac/nonlinear Dirac equation to the lattice/nonlinear lattice Schrödinger equation [59, 61].

Bibliography

[1] D. A. Abanin, S. V. Morozov, L. A. Ponomarenko, R. V. Gorbachev, M. I. K. A. S. Mayorov, K. Watanabe, T. Taniguchi, K. S. Novoselov, L. S. Levitov, and A. K. Geim, "Giant nonlocality near the Dirac point in graphene", *Science*, vol. 332, pp. 328–330, 2011.

[2] M. J. Ablowitz and Y. Zhu, "Nonlinear waves in shallow honeycomb lattices", *SIAM J. Appl. Math.*, vol. 72, pp. 240–260, 2012.

[3] X. Antoine, W. Bao, and C. Besse, "Computational methods for the dynamics of the nonlinear Schrödinger/Gross-Pitaevskii equations", *Comput. Phys. Commun.*, vol. 184, pp. 2621–2633, 2013.

[4] X. Antoine and E. Lorin, "Computational performance of simple and efficient sequential and parallel Dirac equation solvers", Hal-01496817, 2017.

[5] X. Antoine, E. Lorin, J. Sater, F. Fillion-Gourdeau, and A. D. Bandrauk, "Absorbing boundary conditions for relativistic quantum mechanics equations", *J. Comput. Phys.*, vol. 277, pp. 268–304, 2014.

[6] A. Arnold and H. Steinrück, "The 'electromagnetic' Wigner equation for an electron with spin", *ZAMP*, vol. 40, pp. 793–815, 1989.

[7] P. Bader, A. Iserles, K. Kropielnicka, and P. Singh, "Effective approximation for the linear time-dependent Schrödinger equation", *Found. Comp. Math.*, vol. 14, pp. 689–720, 2014.

[8] M. Balabane, T. Cazenave, A. Douady, and M. Frank, "Existence of excited states for a nonlinear Dirac field", *Commun. Math. Phys.*, vol. 119, pp. 153–176, 1988.

[9] M. Balabane, T. Cazenave, and L. Vazquez, "Existence of standing waves for Dirac fields with singular nonlinearities", *Commun. Math. Phys.*, vol. 133, pp. 53–74, 1990.

[10] W. Bao and Y. Cai, "Uniform error estimates of finite difference methods for the nonlinear Schrödinger equation with wave operator", *SIAM J. Numer. Anal.*, vol. 50, pp. 492–521, 2012.

[11] W. Bao and Y. Cai, "Mathematical theory and numerical methods for Bose-Einstein condensation", *Kinet. Relat. Mod.*, vol. 6, pp. 1–135, 2013.

[12] W. Bao and Y. Cai, "Optimal error estimates of finite difference methods for the Gross-Pitaevskii equation with angular momentum rotation", *Math. Comp.*, vol. 82, pp. 99–128, 2013.

[13] W. Bao and Y. Cai, "Uniform and optimal error estimates of an exponential wave integrator sine pseudospectral method for the nonlinear Schrödinger equation with wave operator", *SIAM J. Numer. Anal.*, vol. 52, pp. 1103–1127, 2014.

[14] W. Bao, Y. Cai, X. Jia, and Q. Tang, "A uniformly accurate multiscale time integrator pseudospectral method for the Dirac equation in the nonrelativistic limit regime", *SIAM J. Numer. Anal.*, vol. 54, pp. 1785–1812, 2016.

[15] W. Bao, Y. Cai, X. Jia, and Q. Tang, "Numerical methods and comparison for the Dirac equation in the nonrelativistic limit regime", *J. Sci. Comput.*, vol. 71, pp. 1094–1134, 2017.

[16] W. Bao, Y. Cai, X. Jia, and J. Yin, "Error estimates of numerical methods for the nonlinear Dirac equation in the nonrelativistic limit regime", *Sci. China Math.*, vol. 59, pp. 1461–1494, 2016.

[17] W. Bao, Y. Cai, and J. Yin, "Super-resolution of the time-splitting methods for the Dirac equation in the nonrelativisitic limit regime", *arXiv e-prints*, arXiv: 1811.02174, 2018.

[18] W. Bao, Y. Cai, and J. Yin, "Uniform error bounds of time-splitting methods for the nonlinear Dirac equation in the nonrelativistic limit regime", *arXiv e-prints*, arXiv: 1906.11101, 2019.

[19] W. Bao, Y. Cai, and X. Zhao, "A uniformly accurate multiscale time integrator pseudospectral method for the Klein-Gordon equation in the nonrelativistic limit regime", *SIAM J. Numer. Anal.*, vol. 52, pp. 2488–2511, 2014.

[20] W. Bao and X. Dong, "Analysis and comparison of numerical methods for Klein-Gordon equation in nonrelativistic limit regime", *Numer. Math.*, vol. 120, pp. 189–229, 2012.

[21] W. Bao, S. Jin, and P. A. Markowich, "On time-splitting spectral approximation for the Schrödinger equation in the semiclassical regime", *J. Comput. Phys.*, vol. 175, pp. 487–524, 2002.

[22] W. Bao, S. Jin, and P. A. Markowich, "Numerical study of time-splitting spectral discretizations of nonlinear Schrödinger equations in the semi-classical regimes", *SIAM J. Sci. Comput.*, vol. 25, pp. 27–64, 2003.

[23] W. Bao and X. Li, "An efficient and stable numerical methods for the Maxwell-Dirac system", *J. Comput. Phys.*, vol. 199, pp. 663–687, 2004.

[24] W. Bao and J. Shen, "A fourth-order time-splitting Laguerre-Hermite pseudo-spectral method for Bose-Einstein condensates", *SIAM J. Sci. Comput.*, vol. 26, pp. 2010–2028, 2005.

[25] W. Bao and F. Sun, "Efficient and stable numerical methods for the generalized and vector Zakharov system", *SIAM J. Sci. Comput.*, vol. 26, pp. 1057–1088, 2005.

[26] W. Bao, F. Sun, and G. W. Wei, "Numerical methods for the generalized Zakharov system", *J. Comput. Phys.*, vol. 190, pp. 201–228, 2003.

[27] W. Bao and J. Yin, "A fourth-order compact time-splitting Fourier pseudospectral method for the Dirac equation", *Res. Math. Sci.*, vol. 6, article 11, 2019.

[28] T. Bartsch and Y. Ding, "Solutions of nonlinear Dirac equations", *J. Diff. Eq.*, vol. 226, pp. 210–249, 2006.

[29] P. Bechouche, N. Mauser, and F. Poupaud, "(semi)-nonrelativistic limits of the Dirac equaiton with external time-dependent electromagnetic field", *Commun. Math. Phys.*, vol. 197, pp. 405–425, 1998.

[30] S. Blanesa and P. C. Moan, "Practical symplectic partitioned Runge-Kutta and Runge-Kutta-Nyström methods", *J. Comput. Appl. Math.*, vol. 142, pp. 313–330, 2002.

[31] O. Boada, A. Celi, J. I. Latorre, and M. Lewenstein, "Dirac equation for cold atoms in artificial curved spacetimes", *New J. Phys.*, vol. 13, article 035002, 2011.

[32] J. Bolte and S. Keppeler, "A semiclassical approach to the Dirac equation", *Ann. Phys.*, vol. 274, pp. 125–162, 1999.

[33] N. Bournaveas and G. E. Zouraris, "Split-step spectral scheme for nonlinear Dirac systems", *ESAIM: M2AN*, vol. 46, pp. 841–874, 2012.

[34] D. Brinkman, C. Heitzinger, and P. A. Markowich, "A convergent 2D finite-difference scheme for the Dirac-Poisson system and the simulation of graphene", *J. Comput. Phys.*, vol. 257, pp. 318–332, 2014.

[35] Y. Cai and Y. Wang, "Uniformly accurate nested Picard iterative integrators for the Dirac equation in the nonrelativistic limit regime", *SIAM J. Numer. Anal.*, vol. 57, no. 4, pp. 1602–1624, 2019.

[36] Y. Cai and W. Yi, "Error estimates of finite difference time domain methods for the Klein–Gordon–Dirac system in the nonrelativistic limit regime", *Commun. Math. Sci.*, vol. 16, no. 5, pp. 1325–1346, 2018.

[37] M. Caliari, A. Ostermann, and C. Piazzola, "A splitting approach for the magnetic Schrödinger equation", *J. Comput. Appl. Math.*, vol. 316, pp. 74–85, 2017.

[38] E. Carelli, E. Hausenblas, and A. Prohl, "Time-splitting methods to solve the stochastic incompressible Stokes equation", *SIAM J. Numer. Anal.*, vol. 50, pp. 2917–2939, 2012.

[39] R. Carles, "On Fourier time-splitting methods for nonlinear Schrödinger equations in the semi-classical limit", *SIAM J. Numer. Anal.*, vol. 51, pp. 3232–3258, 2013.

[40] R. Carles and C. Gallo, "On Fourier time-splitting methods for nonlinear Schrödinger equations in the semi-classical limit II. Analytic regularity", *Numer. Math.*, vol. 136, pp. 315–342, 2017.

[41] T. Cazenave and L. Vazquez, "Existence of localized solutions for a classical nonlinear Dirac field", *Commun. Math. Phys.*, vol. 105, pp. 34–47, 1986.

[42] S. J. Chang, S. D. Ellis, and B. W. Lee, "Chiral confinement: An exact solution of the massive Thirring model", *Phys. Rev. D*, vol. 11, pp. 3572–3582, 1975.

[43] S. A. Chin, "Symplectic integrators from composite operator factorizations", *Phys. Lett. A*, vol. 226, pp. 344–348, 1997.

[44] S. A. Chin and C. R. Chen, "Fourth order gradient symplectic integrator methods for solving the time-dependent Schrödinger equation", *J. Chem. Phys*, vol. 114, pp. 7338–7341, 2001.

[45] S. A. Chin and C. R. Chen, "Gradient symplectic algorithms for solving the Schrödinger equation with time-dependent potentials", *J. Chem. Phys*, vol. 117, pp. 1409–1415, 2002.

[46] F. Cooper, A. Khare, B. Mihaila, and A. Saxena, "Solitary waves in the nonlinear Dirac equation with arbitrary nonlinearity", *Phys. Rev. E*, vol. 82, article 036604, 2010.

[47] A. Das, "General solutions of Maxwell–Dirac equations in $1+1$ dimensional space-time and spatial confined solution", *J. Math. Phys.*, vol. 34, pp. 3986–3999, 1993.

[48] A. Das and D. Kay, "A class of exact plane wave solutions of the Maxwell–Dirac equations", *J. Math. Phys.*, vol. 30, pp. 2280–2284, 1989.

[49] A. S. Davydov, "Quantum Mechanics", Pergamon Press, 1976.

[50] P. A. M. Dirac, "The quantum theory of the electron", *Proc. R. Soc. Lond. A*, vol. 117, pp. 610–624, 1928.

[51] P. A. M. Dirac, "A theory of electrons and protons", *Proc. R. Soc. Lond. A*, vol. 126, pp. 360–365, 1930.

[52] P. A. M. Dirac, "The Principles of Quantum Mechanics", Oxford University Press, 1947.

[53] J. Dolbeault, M. J. Esteban, and E. Séré, "On the eigenvalues of operators with gaps: Applications to Dirac operator", *J. Funct. Anal.*, vol. 174, pp. 208–226, 2000.

[54] M. J. Esteban and Séré, "Stationary states of the nonlinear Dirac equation: a variational approach", *Commun. Math. Phys.*, vol. 171, pp. 323–350, 1995.

[55] M. J. Esteban and Séré, "An overview on linear and nonlinear Dirac equations", *Discrete Contin. Dyn. Syst.*, vol. 8, pp. 381–397, 2002.

[56] M. J. Esteban and E. Séré, "Partial Differential Equations and Their Applications", 1997.

[57] D. Fang, S. Jin, and C. Sparber, "An efficient time-splitting method for the Ehrenfest dynamics", *Multiscale Model. Simul.*, vol. 16, pp. 900–921, 2018.

[58] C. L. Fefferman, J. P. Lee-Thorp, and M. I. Weinstein, "Honeycomb Schrödinger operators in the strong binding regime", *Commun. Pur. Appl. Math.*, vol. 71, pp. 1178–1270, 2018.

[59] C. L. Fefferman and M. I. Weistein, "Honeycomb lattice potentials and Dirac points", *J. Am. Math. Soc.*, vol. 25, pp. 1169–1220, 2012.

[60] C. L. Fefferman and M. I. Weistein, "Waves in honeycomb structures", *Journées équations aux dérivées partielles*, pp. 1–12, 2012.

[61] C. L. Fefferman and M. I. Weistein, "Wave packets in honeycomb structures and two-dimensional Dirac equations", *Commun. Math. Phys.*, vol. 326, pp. 251–286, 2014.

[62] A. Ferreira, J. V. Gomes, J. Nilsson, E. R. Mucciolo, N. M. R. Peres, and A. H. C. Neto, "Unified description of the dc-conductivity of monolayer and bilayer graphene at finite densities based on resonant scatterers", *Phys. Rev. B*, vol. 83, article 165402, 2011.

[63] F. Fillion-Gourdeau, E. Lorin, and A. D. Bandrauk, "Resonantly enhanced pair production in a simple diatomic models", *Phys. Rev. Lett.*, vol. 110, article 013002, 2013.

[64] R. Finkelstein, R. Lelevier, and M. Ruderman, "Nonlinear spinor fields", *Phys. Rev.*, vol. 83, pp. 326–332, 1951.

[65] E. Forest and R. D. Ruth, "Fourth-order symplectic integration", *Physica D: Nonlinear Phenomena*, vol. 43, pp. 105–117, 1990.

[66] J. D. Frutos and J. M. Sanz-Serna, "Split-step spectral scheme for nonlinear Dirac systems", *J. Comput. Phys.*, vol. 83, pp. 407–423, 1989.

[67] W. I. Fushchich and W. M. Shtelen, "On some exact solutions of the nonlinear Dirac equation", *J. Phys. A: Math. Gen.*, vol. 16, pp. 271–277, 1983.

[68] L. Gauckler, "On a splitting method for the Zakharov system", *Numer. Math.*, vol. 139, pp. 349–379, 2018.

[69] S. Geng, "Syplectic partitioned Runge-Kutta methods", *J. Comput. Math.*, vol. 11, pp. 365–372, 1993.

[70] P. Gérad, P. A. Markowich, N. J. Mauser, and F. Poupaud, "Homogenization limits and Wigner transforms", *Comm. Pure Appl. Math.*, vol. 50, pp. 321–377, 1997.

[71] F. Gesztesy, H. Grosse, and B. Thaller, "A rigorious approach to relativistic corrections of bound state energies for spin-1/2 particles", *Ann. Inst. Henri Poincaré Phys. Theor.*, vol. 40, pp. 159–174, 1984.

[72] L. Gross, "The Cauchy problem for the coupled Maxwell and Dirac equations", *Commun. Pure Appl. Math.*, vol. 19, pp. 1–15, 1966.

[73] L. H. Haddad and L. D. Carr, "The nonlinear Dirac equation in Bose-Einstein condensates: Foundation and symmetries", *Physica D*, vol. 238, pp. 1413–1421, 2009.

[74] L. H. Haddad, C. M. Weaver, and L. D. Carr, "The nonlinear Dirac equation in Bose-Einstein condensates: I. Relativistic solitons in armchair nanoribbon optical lattices", *New J. Phys.*, vol. 17, article 063033, 2015.

[75] C. R. Hagen, "New solutions of the Thirring model", *Nuovo Cimento*, vol. 51, pp. 169–186, 1967.

[76] E. Hairer, G. Wanner, and C. Lubich, "Geometric Numerical Integration", Springer-Verlag, 2002.

[77] R. Hammer, W. Pötz, and A. Arnold, "A dispersion and norm preserving finite difference scheme with transparent boundary conditions for the Dirac equation in (1+1)D", *J. Comput. Phys.*, vol. 256, pp. 728–747, 2014.

[78] W. Heisenberg, "Quantum theory of fields and elementary particles", *Rev. Mod. Phys.*, vol. 29, pp. 269–278, 1957.

[79] J. L. Hong and C. Li, "Multi-symplectic Runge-Kutta methods for nonlinear Dirac equations", *J. Comput. Phys.*, vol. 211, pp. 448–472, 2006.

[80] Z. Huang, S. Jin, P. A. Markowich, C. Sparber, and C. Zheng, "A time-splitting spectral scheme for the Maxwell-Dirac system", *J. Comput. Phys.*, vol. 208, pp. 761–789, 2005.

[81] W. Hunziker, "On the nonrelativistic limit of the Dirac theory", *Commun. Math. Phys.*, vol. 40, pp. 215–222, 1975.

[82] D. D. Ivanenko, "Notes to the theory of interaction via particles", *Zh. Eksp. Teor. Fiz.*, vol. 8, pp. 260–266, 1938.

[83] J. M. Jauch and F. Rohrlich, "The Theory of Photons and Electrons", Springer-Verlag, 1976.

[84] S. Jiang, L. Greengard, and W. Bao, "Fast and accurate evaluation of nonlocal Coulomb and dipole-dipole interactions via the nonuniform FFT", *SIAM J. Sci. Comput.*, vol. 36, pp. B777–B794, 2014.

[85] S. Jin, P. A. Markowich, and C. Zheng, "Numerical simulation of a generalized Zakharov system", *J. Comput. Phys.*, vol. 201, pp. 376–395, 2004.

[86] S. Jin, P. Markowich, and C. Sparber, "Mathematical and numerical methods for semiclassical Schrödinger equations", *Acta Numer.*, vol. 20, pp. 121–209, 2011.

[87] S. Jin and C. Zheng, "A time-splitting spectral method for the generalized Zakharov system in multi-dimensions", *J. Sci. Comput.*, vol. 26, pp. 127–149, 2006.

[88] A. Komech, "Golbal attraction to solitary waves for a nonlinear Dirac equation with mean field interaction", *SIAM J. Math. Anal.*, vol. 42, pp. 2944–2964, 2010.

[89] V. E. Korepin, "Dirac calculation of the S matrix in the massive Thirring model", *Theor. Math. Phys.*, vol. 41, pp. 953–967, 1979.

[90] M. Lemou, F. Méhats, and X. Zhao, "Uniformly accurate numerical schemes for the nonlinear Dirac equation in the nonrelativistic limit regime", *Commun. Math. Sci.*, vol. 15, no. 4, pp. 1107–1128, 2017.

[91] S. Li, X. Li, and F. Shi, "Time-splitting methods with charge conservation for the nonlinear Dirac equation", *Numer. Meth. Part. D. E.*, vol. 33, pp. 1582–1602, 2017.

[92] C. Lubich, "On splitting methods for Schrödinger-Poisson and cubic nonlinear Schrödinger equations", *Math. Comp.*, vol. 77, pp. 2141–2153, 2008.

[93] P. Mathieu, "Soliton solutions for Dirac equations with homogeneous non-linearity in (1+1) dimensions", *J. Phys. A: Math. Gen.*, vol. 18, pp. L1061–L1066, 1985.

[94] R. I. McLachlan and G. R. W. Quispel, "Splitting methods", *Acta Numer.*, pp. 341–434, 2002.

[95] A. H. C. Neto, F. Guinea, N. M. R. Peres, K. S. Novoselov, and A. K. Geim, "The electronic properties of graphene", *Rev. Mod. Phys.*, vol. 81, pp. 109–162, 2009.

[96] K. S. Novoselov, A. K. Geim, S. Morozov, D. Jiang, Y. Zhang, S. V. Dubonos, I. V. Grigorieva, and A. A. Firsov, "Electric field effect in atomically thin carbon films", *Science*, vol. 306, pp. 666–669, 2004.

[97] J. W. Nraun, Q. Su, and R. Grobe, "Numerical approach to solve the time-dependent Dirac equation", *Phys. Rev. A*, vol. 59, pp. 604–612, 1999.

[98] T. Ohlsson, "Relativistic Quantum Physics: From Advanced Quantum Mechanics to Introductory Quantum Field Theory", Cambridge University Press, 2011.

[99] P. B. Pal, "Dirac, Majorana, and Weyl fermions", *Am. J. Phys.*, vol. 79, pp. 485–498, 2011.

[100] J. Rafelski, "Soliton solutions of a selfinteracting Dirac field in three space dimensions", *Phys. Lett. B*, vol. 66, pp. 262–266, 1977.

[101] P. Ring, "Relativistic mean field theory in finite nuclei", *Prog. Part. Nucl. Phys.*, vol. 37, pp. 193–263, 1996.

[102] B. Saha, "Nonlinear spinor fields and its role in cosmology", *Int. J. Theor. Phys.*, vol. 51, pp. 1812–1837, 2012.

[103] E. Schrödinger, "An undulatory theory of the mechanics of atoms and molecules", *Phys. Rev.*, vol. 28, no. 6, pp. 1049–1070, 1926.

[104] C. E. Shannon, "Communication in the presence of noise", *Proceedings of the Institute of Radio Engineers*, vol. 37, pp. 10–21, 1949.

[105] S. H. Shao, N. R. Quintero, F. G. Mertens, F. Cooper, A. Khare, and A. Saxena, "Stability of solitary waves in the nonlinear Dirac equation with arbitrary nonlinearity", *Phys. Rev. E*, vol. 90, article 032915, 2014.

[106] S. H. Shao and H. Z. Tang, "Interaction for the solitary waves of a nonlinear Dirac model", *Phys. Lett. A*, vol. 345, pp. 119–128, 2005.

[107] S. H. Shao and H. Z. Tang, "Higher-order accurate Runge-Kutta discontinuous Galerkin methods for a nonlinear Dirac model", *Discrete Cont. Dyn. Syst. B*, vol. 6, pp. 623–640, 2006.

[108] S. H. Shao and H. Z. Tang, "Interaction of solitary waves with a phase shift in a nonlinear Dirac model", *Commun. Comput. Phys.*, vol. 3, pp. 950–967, 2008.

[109] J. Shen and T. Tang, "Spectral and High-Order Methods with Applications", Science Press, 2006.

[110] G. D. Smith, "Numerical Solution of Partial Differential Equations: Finite Difference Methods", Clarendon Press, 1985.

[111] M. Soler, "Classical, stable, nonlinear spinor field with positive rest energy", *Phys. Rev. D*, vol. 1, pp. 2766–2769, 1970.

[112] H. Spohn, "Semiclassical limit of the Dirac equation and spin precession", *Annal. Phys.*, vol. 282, pp. 420–431, 2000.

[113] G. Strang, "On the construction and comparison of difference schemes", *SIAM J. Numer. Anal.*, vol. 5, pp. 507–517, 1968.

[114] J. Stubbe, "Exact localized solutions of a family of two-dimensional nonliear spinor fields", *J. Math. Phys.*, vol. 27, pp. 2561–2567, 1986.

[115] M. Suzuki, "General theory of fractal path integrals with applications to many-body theories and statistical physics", *J. Math. Phys.*, vol. 32, pp. 400–407, 1991.

[116] M. Suzuki, "General decompositon theory of ordered exponentials", *Proc. Japan Acad.*, vol. 69, pp. 161–166, 1993.

[117] M. Suzuki, "New scheme of hybrid exponential product formulas with applications to quantum Monte-Carlo simulations", *Springer Proc. Phys.*, vol. 80, pp. 169–174, 1995.

[118] M. Suzukiv, "Fractal decomposition of exponential operators with applications to many-body theories and Monte Carlo simulations", *Phys. Lett. A*, vol. 146, pp. 319–323, 1990.

[119] K. Takahashi, "Soliton solutions of nonlinear Dirac equations", *J. Math. Phys.*, vol. 20, pp. 1232–1238, 1979.

[120] M. Thalhammer, "Convergence analysis of high-order time-splitting pseudo-spectral methods for nonlinear Schrödinger equations", *SIAM J. Numer. Anal.*, vol. 50, pp. 3231–3258, 2012.

[121] B. Thaller, "The Dirac Equation", Springer, 1992.

[122] W. E. Thirring, "A soluble relativistic field theory", *Ann. Phys.*, vol. 3, pp. 91–112, 1958.

[123] H. F. Trotter, "On the product of semi-groups of operators", *Proc. Amer. Math. Soc.*, vol. 10, pp. 545–551, 1959.

[124] L. Verlet, "Computer 'experiments' on classical fluids, I: Thermodynamical properties of Lennard-Jones molecules", *Phys. Rev.*, vol. 159, pp. 98–103, 1967.

[125] H. Wang and H. Z. Tang, "An efficient adaptive mesh redistribution method for a nonlinear Dirac equation", *J. Comput. Phys.*, vol. 222, pp. 176–193, 2007.

[126] H. Wu, Z. Huang, S. Jin, and D. Yin, "Gaussian beam methods for the Dirac equation in the semi-classical regime", *Commun. Math. Sci.*, vol. 10, pp. 1301–1305, 2012.

[127] J. Xu, S. H. Shao, and H. Z. Tang, "Numerical methods for nonlinear Dirac equationn", *J. Comput. Phys.*, vol. 245, pp. 131–149, 2013.

[128] J. Xu, S. H. Shao, H. Z. Tang, and D. Y. Wei, "Multi-hump solitary waves of a nonlinear Dirac equation", *Commun. Math. Sci.*, vol. 13, no. 5, pp. 1219–1242, 2015.

[129] S. Xu, I. Belopolski, N. Alidoust, M. Neupane, G. Bian, C. Zhang, R. Sankar, G. Chang, Z. Yuan, C. Lee, S. Huang, H. Zheng, J. Ma, D. S. Sanchez, B. Wang, A. Bansil, F. Chou, P. P. Shibayev, H. Lin, S. Jia, and M. Z. Hasan, "Discovery of a Weyl fermion semimetal and topological Fermi arcs", *Science*, vol. 349, pp. 613–617, 2015.

[130] H. Yoshida, "Construction of higher order symplectic integrators", *Phys. Lett. A*, vol. 150, pp. 262–268, 1990.

List of Publications

1. Error estimates of numerical methods for the nonlinear Dirac equation in the nonrelativistic limit regime (with W. Bao, Y. Cai and X. Jia), *Sci. China Math.*, Vol. 59 (2016), pp. 1461-1494.

2. A fourth-order compact time-splitting Fourier pseudospectral method for the Dirac equation (with W. Bao), *Res. Math. Sci.*, Vol. 6 (2019), article 11.

3. Super-resolution of time-splitting methods for the Dirac equation in the nonrelativistic limit regime (with W. Bao and Y. Cai), arXiv: 1811.02174, submitted to *Math. Comput.*, under revision.

4. Simple high-order boundary conditions for computing rogue waves in the nonlinear Schrödinger equation (with T. Wang and Z. Xu), submitted to *Comput. Phys. Commun.*, under revision.

5. Error bounds of the finite difference time domain methods for the Dirac equation in the semiclassical regime (with Y. Ma), *J. Sci. Comput.*, to appear.

6. Uniform error bounds of time-splitting methods for the nonlinear Dirac equation in the nonrelativistic limit regime (with W. Bao, and Y. Cai), arXiv: 1906.11101v1.

7. The fourth-order compact time-splitting Fourier pseudospectral method for the Dirac equation with time-dependent potentials, in preparation.

8. Full-discretized uniform error bounds of time-splitting methods for the Dirac/nonlinear Dirac equation in the nonrelativistic regime, in preparation.