

**NUMERICAL METHODS AND
COMPARISONS FOR THE DIRAC EQUATION
IN THE NONRELATIVISTIC LIMIT REGIME**

JIA XIAOWEI

NATIONAL UNIVERSITY OF SINGAPORE

2015

**NUMERICAL METHODS AND
COMPARISONS FOR THE DIRAC EQUATION
IN THE NONRELATIVISTIC LIMIT REGIME**

JIA XIAOWEI

(B.Sc., Tsinghua University)

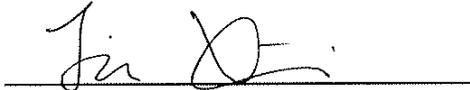
**A THESIS SUBMITTED
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY
DEPARTMENT OF MATHEMATICS
NATIONAL UNIVERSITY OF SINGAPORE
2015**

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.

A handwritten signature in black ink, appearing to read 'Jia Xiaowei', is written over a solid horizontal line.

Jia Xiaowei

11 Aug 2015

Acknowledgements

It is a great pleasure to take this opportunity to thank those who made this thesis possible.

First and foremost, I would like to deliver my deepest gratitude to my supervisor Prof. Bao Weizhu, for his generous support, patient guidance, constructive suggestion, invaluable help and kind encouragement enabled me to conduct such an interesting research project. His advices about life and career will be a great fortune in my future path.

I would like to express my appreciation to my collaborator Dr. Cai Yongyong, for his great contribution to the work. Also, I would like to thank Dr. Tang Qinglin, Dr. Zhao Xiaofei, Dr. Zhang Yong and Ms. Wang Yan for fruitful discussions and suggestions on my research. Specially, sweet memories and thanks goes to Dr. Dong Xuanchun, who was kind, helpful and a great role model during my PhD study. Thank Ms Yin Jia for proofreading the thesis. My sincere thanks also go to all the former colleagues, fellow graduates in our group and all my friends from my office and the Math Department, for the encouragement, emotional support, comradeship and entertainment they offered. I would like to thank NUS for awarding me the Research Scholarship which financially supported me during my Ph.D. candidature.

Last but not least, to my beloved girlfriend and my warm family, thank for their encouragement, steadfast support and unconditional love when it was most needed.

Jia Xiaowei
August 2015

Contents

List of Tables	x
List of Figures	xii
1 Introduction	1
1.1 The Dirac equation	1
1.2 Nondimensionalization and nonrelativistic limit regime	4
1.3 The nonlinear Dirac equation	6
1.4 Purpose and outline of the thesis	8
2 Classical numerical methods	10
2.1 Properties of the Dirac equation	10
2.2 Finite difference methods	14
2.2.1 Finite difference time domain methods	15
2.2.2 Linear stability analysis	18
2.2.3 Mass and energy conservation	20
2.2.4 Main results on error estimates	21
2.2.5 Proof of the error estimates for the CNFD method	23
2.2.6 Proof of the error estimates for the LFFD method	25

2.3	Exponential wave integrator pseudospectral methods	27
2.3.1	The EWI-FP method in 1D	27
2.3.2	Linear stability analysis	30
2.3.3	Convergence analysis	31
2.3.4	Extension to 2D and 3D	35
2.4	Time-splitting Pseudospectral methods	37
2.4.1	The TSFP method in 1D	37
2.4.2	Extension to 2D and 3D	39
2.5	Numerical results	40
2.5.1	Comparison of spatial/temporal resolution	41
2.5.2	Dynamics with the honeycomb potential	46
3	Extension to the nonlinear Dirac equation	52
3.1	Basic properties	52
3.2	Finite difference methods	58
3.2.1	Finite difference time domain methods	58
3.2.2	Linear stability analysis	61
3.2.3	Mass conservation and energy conservation	62
3.2.4	Main results on error estimates	63
3.2.5	Proof of the error estimates for the CNFD method	65
3.2.6	Proof of the error estimates for the LFFD method	68
3.3	Exponential wave integrator pseudospectral methods	71
3.3.1	The EWI-FP method	72
3.3.2	Convergence analysis	75
3.4	Time-splitting Fourier pseudospectral methods	80
3.5	Numerical results	84
4	Fourth order compact time splitting methods	93
4.1	Introduction	93
4.2	Fourth order compact time splitting	95

4.3	Extension to the time-dependent potential	97
4.4	Numerical examples	101
5	A uniformly accurate multiscale method	105
5.1	Introduction	105
5.2	Multiscale decomposition	108
5.3	A multiscale time integrator pseudospectral method	110
5.3.1	The method in 1D	110
5.3.2	A uniform error bound	117
5.3.3	Error analysis	118
5.4	Numerical examples	131
5.4.1	Accuracy test	131
5.4.2	Convergence of Dirac equation to the limiting model	132
6	Concluding remarks and future work	136
	Bibliography	139

Summary

The Dirac equation which was first derived in 1928, is a widely used model in particle physics. It is a relativistic version of the Schrödinger equation that is consistent with both the principles of quantum mechanics and the theory of special relativity. The Dirac equation describes all spin- $1/2$ massive particles, for which parity is a symmetry, such as electrons and quarks. Since 2003, the Dirac equation has regained considerable research interests due to the groundbreaking discovery of graphene, the first two dimensional material. The dynamics of electrons in graphene can be very well described by the Dirac equation. Therefore the computation for the Dirac equation is of significant research value.

The purpose of this thesis is to propose and analyze some efficient numerical methods for solving the Dirac equation in the nonrelativistic limit regime. The numerical implementations here include some classical discretization methods and multiscale decomposition methods. We mainly focus on the comparisons of how the error bounds depend on the nonrelativistic limit parameter in different methods. This would help us to choose proper temporal step size in order to resolve the oscillation of this regime, and finally achieve a uniformly accurate numerical method such that the error bound is independent of the nonrelativistic limit parameter and so is the choice of temporal step size.

This thesis contains three parts. In the first part, several discretization methods are proposed, analyzed and compared for solving the linear and nonlinear Dirac (NLD) equation in the nonrelativistic limit regime, of which the solution is highly oscillatory with a dimensionless parameter $0 < \varepsilon \ll 1$. In fact, solutions in this regime propagate waves with wavelength at $O(\varepsilon^2)$ when $0 < \varepsilon \ll 1$, which would bring a remarkable lot of computational burdens and make the error bounds not uniformly accurate. Frequently used second order finite difference time domain (FDTD) methods are first analyzed and concluded with an optimal error bound with respect to the parameter ε . Exponential wave integrator Fourier pseudospectral (EWI-FP) methods and time splitting Fourier pseudospectral (TSFP) methods are proposed and analyzed afterwards. Rigorous and optimal error estimates with numerical results show that these two kinds of methods overcome the FDTD methods in the nonrelativistic limit regime. However, none of the above second order methods is a uniformly accurate one in solving the linear and nonlinear Dirac equation in the nonrelativistic limit regime.

The second part is devoted to applying fourth order compact splitting operator methods to solve the Dirac equation in the nonrelativistic limit regime. We state the conclusion that with the coefficients of factorization of the evolution operator being purely positive, more accurate numerical solutions could be obtained with larger temporal step size. Several fourth order splitting operator methods with purely positive factorization coefficients are presented and numerical results are shown to support our conclusion.

The last part is to propose and analyze a multiscale time integrator Fourier pseudospectral method (MTI-FP) to solve the Dirac equation in the nonrelativistic limit regime. This method is motivated by the frequency and spectral decomposition. Two rigorous error bounds are established independently via two different mathematical approaches for the MTI-FP as $O\left(h^{m_0} + \frac{\tau^2}{\varepsilon^2}\right)$ and $O(h^{m_0} + \tau^2 + \varepsilon^2)$ with h mesh size, τ time step and $m_0 \geq 2$ depending on the regularity of the solution, which immediately implies that MTI-FP converges uniformly and optimally in space

with exponential rate if the solution is smooth, and uniformly in time with linear convergent rate $O(\tau)$ for all $0 < \varepsilon \leq 1$, and optimal with quadratic convergent rate at $O(\tau^2)$ in the regimes when either $\varepsilon = O(1)$ or $0 < \varepsilon < \tau$. Numerical results are provided at last to confirm the error bounds and the best performance of the MTI-FP method among all the methods analyzed in this thesis for solving the Dirac equation in the nonrelativistic limit regime.

Notations

$i = \sqrt{-1}$	imaginary unit
\hbar	Planck constant
c	speed of light
t	time variable
\mathbb{R}^d	d dimensional Euclidean space
\mathbb{C}^d	d dimensional complex space
$\mathbf{x} = (x_1, \dots, x_d)^T$	spatial variable in \mathbb{R}^d
τ	time step size
h	space mesh size
ε	a dimensionless parameter with its value $0 < \varepsilon \leq 1$
$\Psi := \Psi(\mathbf{x}, t) \in \mathbb{C}^4$	4-component complex wave function
$\Phi := \Phi(\mathbf{x}, t) \in \mathbb{C}^2$	2-component complex wave function
$\sigma_i (i = 1, 2, 3)$	Pauli matrices
∇	gradient operator
$\nabla^2 = \nabla \cdot \nabla, \Delta$	Laplace operator
$A \lesssim B$	$ A \leq C \cdot B$ for some generic constant $C > 0$ independent of τ, h and ε
$\ u\ _p := \ u\ _{L^p(\mathbb{R}^d)}$	L^p ($p \in [1, \infty]$) norm of function $u(\mathbf{x})$,

$\operatorname{Re}(f)$	real part of f
$\operatorname{Im}(f)$	imaginary part of f
\bar{f}	conjugate of function f
Q^T	transpose of matrix Q
Q^*	conjugate transpose of matrix Q
$\hat{f}(\xi) := \int_{\mathbb{R}^d} f(\mathbf{x})e^{-i\mathbf{x}\cdot\xi} d\mathbf{x}$	Fourier transform of $f(\mathbf{x})$
1D	one dimension
2D	two dimension
3D	three dimension
NLDE	nonlinear Dirac equation
CNFD	Crank-Nicolson finite difference
SIFD	semi-implicit finite difference
LFFD	Leap-Frog finite difference
EWI	exponential wave integrator
EWI-FP	exponential wave integrator Fourier pseudospectral
TSFP	time-splitting Fourier pseudospectral
MTI	multiscale time integrator
Fig.	figure
Tab.	table

List of Tables

2.1	Spatial and temporal error analysis of the LFFD method for the Dirac equation in 1D.	42
2.2	Spatial and temporal error analysis of the SIFD1 method for the Dirac equation in 1D.	43
2.3	Spatial and temporal error analysis of the SIFD2 method for the Dirac equation in 1D.	44
2.4	Spatial and temporal error analysis of the CNFD method for the Dirac equation in 1D.	45
2.5	Spatial and temporal error analysis of the EWI-FP method for the Dirac equation in 1D.	46
2.6	Spatial and temporal error analysis of the TSFP method for the Dirac equation in 1D.	49
2.7	Comparison of temporal errors of different numerical methods for the Dirac equation under proper ε -scalability.	50
2.8	Spatial error analysis of the CNFD method for the free Dirac equation with different h	50
2.9	Comparison of properties of different numerical methods for solving the Dirac equation with M being the number of grid points in space.	51

3.1	Spatial and temporal error analysis of the LFFD method for the NLDE.	86
3.2	Spatial and temporal error analysis of the SIFD1 method for the NLDE.	87
3.3	Spatial and temporal error analysis of the SIFD2 method for the NLDE.	88
3.4	Spatial and temporal error analysis of the CNFD method for the NLDE.	89
3.5	Spatial and temporal error analysis of the EWI-FP method for the NLDE.	90
3.6	Spatial and temporal error analysis of the TSFP method for the NLDE.	91
3.7	Comparison of properties of different numerical methods for solving the NLDE with M being the number of grid points in space.	92
5.1	Spatial error analysis of the MD-EWI-FP method for the Dirac equation in 1D.	132
5.2	Temporal error analysis of the MD-EWI-FP method for the Dirac equation.	134

List of Figures

1.1	The solution $\phi_1(t = 1, x)$ and $\phi_1(t, x = 0)$ of the Dirac equation with $d = 1$ for different ε	6
2.1	Dynamics of the densities $\rho_1(t, \mathbf{x}) = \phi_1(t, \mathbf{x}) ^2$ (left) and $\rho_2(t, \mathbf{x}) = \phi_2(t, \mathbf{x}) ^2$ (right) of the Dirac equation in 2D with a honeycomb lattice potential when $\varepsilon = 1$	47
2.2	Dynamics of the densities $\rho_1(t, \mathbf{x}) = \phi_1(t, \mathbf{x}) ^2$ (left) and $\rho_2(t, \mathbf{x}) = \phi_2(t, \mathbf{x}) ^2$ (right) of the Dirac equation in 2D with a honeycomb potential when $\varepsilon = 0.2$	48
3.1	The soliton solution $ \Phi(t = 0, x) $ and $ \Phi(t = 15, x) $ of the Dirac equation (3.1.11) with $\varepsilon = 1$, where $ \Phi(t, x) = \sqrt{ \phi_1(t, x) ^2 + \phi_2(t, x) ^2}$, the speed $v = 0.8$	57
4.1	The computational error of the Dirac equation using different splitting operator methods with $\varepsilon = \frac{1}{4}$	103
4.2	The computational error of the Dirac equation using different splitting operator methods with $\varepsilon = \frac{1}{16}$	104
5.1	Error functions $E_{\text{sch}}(t)$ and $E_{\text{pau}}(t)$ for different ε	133

Introduction

1.1 The Dirac equation

Walking through the history of physics, before the twentieth century, the equations of Newton's law were used to predict what a system would do at any time after the initial conditions. At that time the whole world was surrounded by an atmosphere that the development of physics was meeting its end. When time came to the new century (twentieth century), an evolution happened in physics. Albert Einstein set up the theory of relativity in early twentieth century [35]. Later in 1926, the Austrian physicist Erwin Schrödinger formulated a linear partial differential equation to describe the wave function of a quantum system such as atoms, molecules, and subatomic particles whether free, bound, or localized [85]. In quantum mechanics, Schrödinger equation holds the same important position as Newton's law in classical mechanics. Inspired by these two genius discoveries, one wishes to build relativistic wave equations where quantum mechanics and special relativity simultaneously apply. The Klein-Gorden equation was the first such equation to be obtained, but density of this system may be negative, which seems impossible for a legitimate probability density. To overcome this problem, Dirac thought to try an equation that was first order in both time and space. In 1928, he derived a relativistic wave equation, which in its free form, or including electromagnetic field, could be used

to describe all spin- $\frac{1}{2}$ particles such as electrons and quarks [32–34]. When these particles are moving at an extremely high velocity, or bounded by very strong classical fields, the nonrelativistic modeling based on Schrödinger equation fails and theoretical investigation should be based on the Dirac equation.

Given electromagnetic fields, to deal with the linear one-particle Dirac equation, the most compact form reads

$$(i\hbar\gamma^\eta\partial_\eta - m_0c + e\gamma^\eta A_\eta)\Psi = 0. \quad (1.1.1)$$

Here the unknown Ψ is the 4-component complex wave function of the spinorfield: $\Psi(t, \mathbf{x}) = (\Psi_1, \Psi_2, \Psi_3, \Psi_4)^T \in \mathbb{C}^4$, $x_0 = ct$, $\mathbf{x} = (x_1, x_2, x_3)^T \in \mathbb{R}^3$ with x_0 and \mathbf{x} denoting the time and spatial coordinates in Minkowski space. ∂_η stands for $\frac{\partial}{\partial x_\eta}$, i.e. $\partial_0 = \frac{\partial}{\partial x_0} = \frac{1}{c}\frac{\partial}{\partial t}$, $\partial_k = \frac{\partial}{\partial x_k}$ ($k = 1, 2, 3$), where we consequently adopt notation that Greek letter η denotes 0, 1, 2, 3 and k denotes the 3 spatial dimension indices 1, 2, 3. $\gamma^\eta A_\eta$ stands for the summation $\sum_{\eta=1}^3 \gamma^\eta A_\eta$. The physical constants are: \hbar for the Plank constant, c for the speed of light, m_0 for the electron's mass, and e for the unit charge. $\gamma^\eta \in \mathbb{C}^{4 \times 4}$, $\eta = 0, \dots, 3$ are the 4×4 matrices given by

$$\gamma^0 = \begin{pmatrix} I_2 & \mathbf{0} \\ \mathbf{0} & -I_2 \end{pmatrix}, \quad \gamma^k = \begin{pmatrix} \mathbf{0} & \sigma_k \\ \sigma_k & \mathbf{0} \end{pmatrix}, \quad k = 1, 2, 3, \quad (1.1.2)$$

where I_m (m a positive integer) is the $m \times m$ identity matrix and σ^k ($k = 1, 2, 3$) are the 2×2 Pauli matrices

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

$A_\eta(t, \mathbf{x}) \in \mathbb{R}$, $\eta = 0, \dots, 3$ are the components of the electromagnetic potentials, in particular $V(t, \mathbf{x}) = -A_0(t, \mathbf{x})$ is the electric potential and $\mathbf{A}(t, \mathbf{x}) = (A_1, A_2, A_3)^T$ is the magnetic potential vector. Hence the electric field is given by $\mathbf{E}(t, \mathbf{x}) = \nabla A_0 - \partial_t \mathbf{A}$ and the magnetic field $\mathbf{B}(t, \mathbf{x}) = \text{curl} \mathbf{A} = \nabla \times \mathbf{A}$.

The extreme conditions where relativistic effects are important can be found in

many areas such as relativistic heavy ion collisions, heavy ion spectroscopy, cosmology, astrophysics, and more recently, in laser-matter interactions [82] and condense matter physics [62]. For this reason, the Dirac equation coupled with an electromagnetic field has been studied extensively to evaluate many observables such as electron-positron production, molecule spectra, molecular ionization rates, etc. However, solving this equation remains a very challenging task because of its intricate matrix structure, its unbounded spectrum (the Dirac equation has negative energy states which forbid the use of naive minimization numerical methods) and its multiscale.

Existing approaches to tackle these important problems can usually be classified into three categories. The first one is analytical method, which aims at finding closed-form solutions. Although many important problems were treated in this way [11, 47], it only allows the study of idealized systems. The second approach is the semi-classical approximation which can be used to study more complex configurations than the analytical method [72]. However, it is only valid for a certain range of wave function parameters, which may not be realized in the physical system under study. The last one is based on full numerical approximations, which in principle, can be used to investigate any physical system. But even on the numerical side, finding the solution to the Dirac equation is still a challenging problem: it requires a lot of computer resources and numerical artifacts such as the fermion-doubling problem plague certain numerical schemes. Therefore, special cares have to be taken to resolve these issues in solving the Dirac equation numerically for physical relevant systems.

1.2 Nondimensionalization and nonrelativistic limit regime

In (1.1.1), substitute x_0 by ct , we can obtain the three dimensional Dirac equation describing the time evolution of spin-1/2 massive particles within external time-dependent electromagnetic potentials [32,33]

$$i\hbar\partial_t\Psi(t, \mathbf{x}) = \left[-i\hbar\sum_{j=1}^3\alpha_j\partial_j + mc^2\beta\right]\Psi(t, \mathbf{x}) + e\left[V(t, \mathbf{x})I_4 - \sum_{j=1}^3A_j(t, \mathbf{x})\alpha_j\right]\Psi(t, \mathbf{x}), \quad (1.2.1)$$

where $i = \sqrt{-1}$, $\mathbf{x} = (x_1, x_2, x_3)^T \in \mathbb{R}^3$ (equivalently written as $\mathbf{x} = (x, y, z)^T$) is the spatial coordinate vector. The 4×4 matrices $\alpha_1, \alpha_2, \alpha_3$ and β are defined as

$$\alpha_1 = \begin{pmatrix} \mathbf{0} & \sigma_1 \\ \sigma_1 & \mathbf{0} \end{pmatrix}, \quad \alpha_2 = \begin{pmatrix} \mathbf{0} & \sigma_2 \\ \sigma_2 & \mathbf{0} \end{pmatrix}, \quad \alpha_3 = \begin{pmatrix} \mathbf{0} & \sigma_3 \\ \sigma_3 & \mathbf{0} \end{pmatrix}, \quad \beta = \begin{pmatrix} I_2 & \mathbf{0} \\ \mathbf{0} & -I_2 \end{pmatrix}.$$

In order to scale the Dirac equation (1.2.1), we introduce

$$\begin{aligned} \tilde{t} &= \frac{t}{t_s}, & \tilde{\mathbf{x}} &= \frac{\mathbf{x}}{x_s}, & \tilde{\Psi}(\tilde{t}, \tilde{\mathbf{x}}) &= x_s^{3/2}\Psi(t, \mathbf{x}), \\ \tilde{V}(\tilde{t}, \tilde{\mathbf{x}}) &= \frac{V(t, \mathbf{x})}{A_s}, & \tilde{A}_j(\tilde{t}, \tilde{\mathbf{x}}) &= \frac{A_j(t, \mathbf{x})}{A_s}, & j &= 1, 2, 3, \end{aligned} \quad (1.2.2)$$

where x_s, t_s and A_s are reference length unit, time unit and potential unit, respectively satisfying $t_s = \frac{mx_s^2}{h}$ and $A_s = \frac{mv^2}{e}$ with $v = \frac{x_s}{t_s}$ being the wave speed. Plugging (1.2.2) into (1.2.1), multiplying by $\frac{t_s x_s^{3/2}}{h}$, and then removing all $\tilde{\cdot}$, we obtain the following dimensionless Dirac equation in 3D

$$i\partial_t\Psi(t, \mathbf{x}) = \left[-\frac{i}{\varepsilon}\sum_{j=1}^3\alpha_j\partial_j + \frac{1}{\varepsilon^2}\beta\right]\Psi(t, \mathbf{x}) + \left[V(t, \mathbf{x})I_4 - \sum_{j=1}^3A_j(t, \mathbf{x})\alpha_j\right]\Psi(t, \mathbf{x}), \quad (1.2.3)$$

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

$$0 < \varepsilon := \frac{x_s}{t_s c} = \frac{v}{c} \leq 1. \quad (1.2.4)$$

For the Dirac equation (1.2.3) with $\varepsilon = 1$, i.e. $O(1)$ -speed of light regime, there are extensive analytical and numerical results in the literatures. For the existence

and multiplicity of bound states and/or standing wave solutions, we refer to [36] and references therein. For the analysis of the classical/semiclassical limits via the Wigner transform techniques, we refer to [9, 45] and references therein. For the numerical methods and comparison such as the finite difference time domain (FDTD) methods and the Gaussian beam methods, we refer to [102, 106, 107] and references therein. However, for the Dirac equation (1.2.3) with $0 < \varepsilon \ll 1$, i.e. nonrelativistic limit regime (or the scaled speed of light goes to infinity), analysis and efficient computation of the Dirac equation (1.2.3) are mathematically rather complicated issues. The main difficulty is that the solutions are highly oscillatory in time and the corresponding kinetic energy functionals are indefinite [20, 37] and become unbounded when $\varepsilon \rightarrow 0$. There are extensive mathematical analyses of the (semi)-nonrelativistic limit of the Dirac equation (1.2.3) to the Pauli equation [20, 58] and/or the Schrödinger equation when $\varepsilon \rightarrow 0$ [20]. These rigorous analytical results show that the solutions propagate waves with wavelength $O(\varepsilon^2)$ and $O(1)$ in time and space, respectively, when $0 < \varepsilon \ll 1$. In fact, the oscillatory structure of the solutions to the Dirac equation (1.2.3) when $0 < \varepsilon \ll 1$ can be formally observed from its dispersion relation. To illustrate this further, Fig 1.1 shows the solution of the Dirac equation with $d = 1$, $V(t, x) = \frac{1-x}{1+x^2}$, $A_1(t, x) = \frac{(1+x)^2}{1+x^2}$ and $\Phi_0(x) = (\exp(-x^2/2), \exp(-(x-1)^2/2))^T$ for different ε . This highly oscillatory nature of the solutions to the Dirac equation causes severe numerical burdens in practical computation, making the numerical approximation extremely challenging and costly in the nonrelativistic regime $0 < \varepsilon \ll 1$.

Recently, different numerical methods were proposed and analyzed for the efficient computation of the Klein-Gordon equation in the nonrelativistic limit regime [17] and/or highly oscillatory dispersive partial differential equations (PDEs) [13]. To our knowledge, so far there are few results on the numerics of the Dirac equation in the nonrelativistic limit regime.

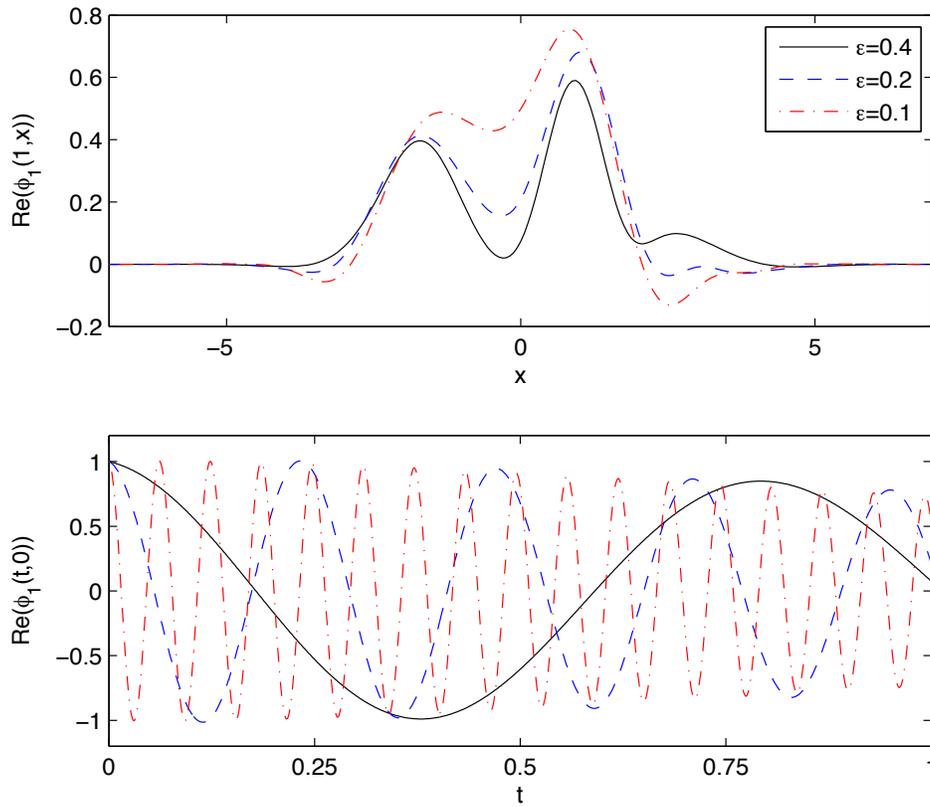


Figure 1.1: The solution $\phi_1(t = 1, x)$ and $\phi_1(t, x = 0)$ of the Dirac equation with $d = 1$ for different ε .

1.3 The nonlinear Dirac equation

Following Dirac's discovery of the linear equation of the electron, there appears the fundamental idea of nonlinear description of an elementary spin- $\frac{1}{2}$ particle which makes it possible basis model for a unified field theory, e.g. the nonlinear Dirac equation (NLDE). A key feature of the NLDE is that it allows solitary wave solutions or particle-like solutions—the stable localized solutions with finite energy and charge. That is, the particles appear as intense localized regions of field which can be recognized as the basic ingredient in the description of extended objects in quantum mechanics. NLDE models attracted wide interest of physicists and mathematicians around the 1970s and 1980s, especially on looking for the solitary wave solutions

and investigation the related physical and mathematical properties.

For the NLDE in (1+1) dimensions (i.e. one time dimension plus one space dimension), [24, 65] derived several analytical solitary wave solutions for the quadric nonlinearity, as well as [69] for fractional nonlinearity and [30, 93] for general nonlinearity by using explicitly the constraint resulting from energy-momentum conservation, and summarized by Mathieu [70]. In contrast, even though mathematicians have claimed existence for various situations, there are few explicit solutions in (1+3) dimensions except for some particular cases shown in [103], and most understandings are based on numerical investigations, eg [3, 79, 97]. Readers are referred to an overview [36] on this topic. Beyond this, the study of the NLDE in (1+1) dimensions could be very helpful for that in (1+3) dimensions since the (1+1) dimensional NLDE corresponds to the asymptotic form of the equation in the physically interesting case of (1+3) dimensions as emphasized by Kaus [63]. That is, some qualitative properties of the NLDE solitary waves could be similar in such two cases.

In the case that theoretical methods are not capable of providing satisfactory results, numerical methods are used to obtain the solitary wave solutions and investigate the stability for the NLDE. Alvarez and Carreras [5] simulated the interaction dynamics between the (1+1)-dimensional NLDE solitary waves of different initial charge for the Soler model by using a second-order accurate Crank-Nicolson (CN) scheme [6]. They first saw there: charge and energy interchange except for some particular initial velocities of the solitary waves; inelastic interaction in binary collisions; and oscillating state production from binary collisions. Inspired by their work, Shao and Tang revisited this interaction dynamics problem in 2005 [86] by employing a fourth-order accurate Runge-Kutta discontinuous Galerkin (RKDG) method [87]. They revealed the collapse in binary and ternary collisions of two-humped NLDE solitary waves [86]; a long-live oscillating state formed with an approximate constant frequency in collisions of two standing waves [87]; full repulsion in binary and ternary collisions of out-of-shape waves [88]. Their numerical results also inferred that two-humped profile could undermine the stability during the scattering

of NLDE solitary waves. Besides the often used CN and RKDG methods, there exist many other numerical schemes for solving the (1+1)-dimensional NLDE: split-step spectral schemes [43], the linearized CN scheme [4], the semi-implicit scheme [22,60], Legendre rational spectral method [101], multi-symplectic Runge-Kutta method [56], adaptive mesh methods [102] etc. The fourth-order accurate RKDG method is very appropriate for investigating the interaction dynamics of the NLDE solitary waves due to their ability to capture the discontinuous or strong gradients without producing spurious oscillations, and thus performs better than the second-order accurate CN scheme [6]. However, the high cost due to the relatively more freedoms used in each cell and the stringent time step constraint reduce its practicality in more realistic simulations where real time and quantitative results are required.

Recently, there has been a magnificent increment of interest in the NLDE models, as they emerge naturally as practical models in physical systems, such as the gap solitons in nonlinear optics, Bose-Einstein condensates in honeycomb optical lattices [48] and matter influencing the evolution of the Universe in cosmology [81]. In view of such new trend, longtime stable, efficient, conservative and high-order accurate numerical methods for solving the NLDE are highly desirable. Finite difference methods, usually as the first try in practice, enable easy coding and debugging and thus are often used by physicists and engineers. However, all of these finite difference methods are often of the second order accuracy and thus sustain fast error growth with respect to time. To achieve relatively slow error growth, sometimes high-order accurate numerical methods are required.

1.4 Purpose and outline of the thesis

This work is devoted to proposing and analyzing efficient and accurate numerical methods for solving the Dirac equation in the nonrelativistic limit regime. Various classes of numerical methods will be proposed and compared, and some of them will be analyzed in the stability and convergence. Rigorous error estimates will be

provided for some of these methods too.

The thesis is organized as follows. In Chapter 2 and 3, we study the numerical methods for the linear and nonlinear Dirac equation with external electromagnetic fields in the nonrelativistic limit regime, respectively. Several second-order finite difference methods are reviewed and their stability and convergence are analyzed in this regime first to illustrate the computational burden brought by the oscillatory solutions in this regime.

In Chapter 4, we propose fourth-order compact splitting operator methods to solve the Dirac equation in the nonrelativistic limit regime. These methods are improvements from the second-order TSFP method, which can reduce the error and obtain a more accurate solution. At the end of this chapter, the numerical results show that the performance of this method is much better than that of TSFP method.

In Chapter 5, we investigate the uniform convergence rate (resp. to ε) for a multiscale time integrator Fourier pseudospectral (MTI-FP) method solving the Dirac equation in the nonrelativistic limit regime. Based on the frequency and spectral decomposition of the Dirac operator, with the help of exponential wave integrator in time and Fourier pseudospectral discretization in space, the MTI-FP method is derived. Via two different mathematical approaches, two different error bounds, $O(h^{m_0} + \frac{\tau^2}{\varepsilon^2})$ and $O(h^{m_0} + \tau^2 + \varepsilon^2)$, are established for this new method. Then a conclusion is drawn that the MTI-FP method is uniformly accurate in the nonrelativistic limit regime. Numerical results are displayed to support this conclusion, and also some numerical results to show the convergence of the Dirac equation to the limit Schrödinger and Pauli type equation are presented.

In Chapter 6, some conclusions are drawn and some possible future works are discussed.

Throughout the paper, we adopt standard notations of Sobolev spaces and their norms, and use the notation $A \lesssim B$ to represent that there exists a generic constant $C > 0$, which is independent of time step τ , mesh size h and ε , such that $|A| \leq C \cdot B$.

Classical numerical methods

In this chapter, the computation for the Dirac equation with external electromagnetic potential in the nonrelativistic limit regime is considered. Several different numerical methods, e.g. finite difference methods, exponential wave integrator methods and time-splitting Fourier pseudospectral methods are applied to this highly oscillatory system and the numerical results are compared at last.

2.1 Properties of the Dirac equation

Similar to the dimension reduction of the nonlinear Schrödinger equation and/or the Schrödinger-Poisson equations with/without anisotropic external potentials [12], when the initial data $\Psi(0, \mathbf{x})$ and the electromagnetic potentials $V(t, \mathbf{x})$ and $\mathbf{A}(t, \mathbf{x})$ are independent of z and thus the wave function Ψ is formally assumed to be independent of z , or when the electromagnetic potentials $V(t, \mathbf{x})$ and $\mathbf{A}(t, \mathbf{x})$ are strongly confined in the z -direction and thus Ψ is formally assumed to be concentrated on the xy -plane, then the 3D Dirac equation (1.2.3) can be reduced to the Dirac equation in 2D with $\mathbf{x} = (x, y)^T \in \mathbb{R}^2$ as

$$i\partial_t\Psi(t, \mathbf{x}) = \left[-\frac{i}{\varepsilon} \sum_{j=1}^2 \alpha_j \partial_j + \frac{1}{\varepsilon^2} \beta\right] \Psi(t, \mathbf{x}) + \left[V(t, \mathbf{x})I_4 - \sum_{j=1}^2 A_j(t, \mathbf{x})\alpha_j\right] \Psi(t, \mathbf{x}). \quad (2.1.1)$$

This 2D Dirac equation has been widely used to model the electron structure and/or dynamical properties of graphene since they share the same dispersive relation on the Dirac points [1,75–77,83]. Similarly, under the proper assumptions on the initial data and the external electromagnetic potentials, the 3D Dirac equation (1.2.3) can be reduced to the Dirac equation in 1D with $\Psi = \Psi(t, x)$ as

$$i\partial_t\Psi(t, x) = \left[-\frac{i}{\varepsilon}\alpha_1\partial_x + \frac{1}{\varepsilon^2}\beta\right]\Psi(t, x) + \left[V(t, x)I_4 - A_1(t, x)\alpha_1\right]\Psi(t, x), \quad x \in \mathbb{R}. \quad (2.1.2)$$

In fact, the Dirac equation in 3D (1.2.3), in 2D (2.1.1) and in 1D (5.1.3) can be written in a unified way in d -dimensions ($d = 1, 2, 3$)

$$i\partial_t\Psi(t, \mathbf{x}) = \left[-\frac{i}{\varepsilon}\sum_{j=1}^d\alpha_j\partial_j + \frac{1}{\varepsilon^2}\beta\right]\Psi(t, \mathbf{x}) + \left[V(t, \mathbf{x})I_4 - \sum_{j=1}^d A_j(t, \mathbf{x})\alpha_j\right]\Psi(t, \mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^d, \quad (2.1.3)$$

and the initial condition for dynamics is given as

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

The Dirac equation (2.1.3) is dispersive and time symmetric. Introducing the position density ρ_j for the j -component ($j = 1, 2, 3, 4$) and the total density ρ

$$\rho(t, \mathbf{x}) = \sum_{j=1}^4 \rho_j(t, \mathbf{x}) = \Psi^*\Psi, \quad \rho_j(t, \mathbf{x}) = |\psi_j(t, \mathbf{x})|^2, \quad 1 \leq j \leq 4 \quad (2.1.4)$$

as well as the current density $\mathbf{J}(t, \mathbf{x}) = (J_1(t, \mathbf{x}), J_2(t, \mathbf{x}), J_3(t, \mathbf{x}))^T$

$$J_l(t, \mathbf{x}) = \frac{1}{\varepsilon}\Psi^*\alpha_l\Psi, \quad l = 1, 2, 3, \quad (2.1.5)$$

where $\Psi^* = \overline{\Psi}^T$ denotes the complex transpose conjugate, then the following conservation law can be obtained from the Dirac equation (2.1.3)

$$\partial_t\rho(t, \mathbf{x}) + \nabla \cdot \mathbf{J}(t, \mathbf{x}) = 0, \quad \mathbf{x} \in \mathbb{R}^d, \quad t \geq 0. \quad (2.1.6)$$

Thus the Dirac equation (2.1.3) conserves the total mass as

$$\|\Psi(t, \cdot)\|^2 := \int_{\mathbb{R}^d} |\Psi(t, \mathbf{x})|^2 d\mathbf{x} = \int_{\mathbb{R}^d} \sum_{j=1}^4 |\psi_j(t, \mathbf{x})|^2 d\mathbf{x} \equiv \|\Psi(0, \cdot)\|^2 = \|\Psi_0\|^2, \quad t \geq 0. \quad (2.1.7)$$

If the electric potential V is perturbed by a constant, e.g. $V(t, \mathbf{x}) \rightarrow V(t, \mathbf{x}) + V^0$ with V^0 being a real constant, then the solution $\Psi(t, \mathbf{x}) \rightarrow e^{-iV^0 t} \Psi(t, \mathbf{x})$ which implies the density of each component ρ_j ($j = 1, 2, 3, 4$) and the total density ρ unchanged. When $d = 1$, if the magnetic potential A_1 is perturbed by a constant, e.g. $A_1(t, \mathbf{x}) \rightarrow A_1(t, \mathbf{x}) + A_1^0$ with A_1^0 being a real constant, then the solution $\Psi(t, \mathbf{x}) \rightarrow e^{iA_1^0 t \alpha_1} \Psi(t, \mathbf{x})$ which implies the total density ρ unchanged; but this property is not valid when $d = 2, 3$. In addition, when the electromagnetic potentials are time-independent, i.e. $V(t, \mathbf{x}) = V(\mathbf{x})$ and $A_j(t, \mathbf{x}) = A_j(\mathbf{x})$ for $j = 1, 2, 3$, the following energy functional is also conserved

$$\begin{aligned} E(t) &:= \frac{1}{2} \int_{\mathbb{R}^d} \left(-\frac{i}{\varepsilon} \sum_{j=1}^d \Psi^* \alpha_j \partial_j \Psi + \frac{1}{\varepsilon^2} \Psi^* \beta \Psi + V(\mathbf{x}) |\Psi|^2 - \sum_{j=1}^d A_j(\mathbf{x}) \Psi^* \alpha_j \Psi \right) d\mathbf{x} \\ &\equiv E(0), \quad t \geq 0. \end{aligned} \quad (2.1.8)$$

Furthermore, if the external electromagnetic potentials are constants, i.e. $V(t, \mathbf{x}) \equiv V^0$ and $A_j(t, \mathbf{x}) \equiv A_j^0$ for $j = 1, 2, 3$, the Dirac equation (2.1.3) 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 $\mathbf{k} = (k_1, \dots, k_d)^T \in \mathbb{R}^d$ satisfy the following

$$\omega \mathbf{B} = \left[\sum_{j=1}^d \left(\frac{k_j}{\varepsilon} - A_j^0 \right) \alpha_j + \frac{1}{\varepsilon^2} \beta + V^0 I_4 \right] \mathbf{B}. \quad (2.1.9)$$

Solving the above eigenvalue problem, we can get the *dispersion relation*

$$\omega = V^0 \pm \sqrt{\frac{1}{\varepsilon^4} + \left(\frac{\mathbf{k}}{\varepsilon} - \mathbf{A}^0 \right)^2} \quad (2.1.10)$$

Plugging (1.1.2) and (1.1.3) into (2.1.1), the 2D Dirac equation (2.1.1) can be decoupled as

$$\begin{aligned} i\partial_t \psi_1 &= -\frac{i}{\varepsilon} (\partial_x - i\partial_y) \psi_4 + \frac{1}{\varepsilon^2} \psi_1 + V(t, \mathbf{x}) \psi_1 - [A_1(t, \mathbf{x}) - iA_2(t, \mathbf{x})] \psi_4, \\ i\partial_t \psi_4 &= -\frac{i}{\varepsilon} (\partial_x + i\partial_y) \psi_1 - \frac{1}{\varepsilon^2} \psi_4 + V(t, \mathbf{x}) \psi_4 - [A_1(t, \mathbf{x}) + iA_2(t, \mathbf{x})] \psi_1, \end{aligned} \quad (2.1.11)$$

$$\begin{aligned} i\partial_t \psi_2 &= -\frac{i}{\varepsilon} (\partial_x + i\partial_y) \psi_3 + \frac{1}{\varepsilon^2} \psi_2 + V(t, \mathbf{x}) \psi_2 - [A_1(t, \mathbf{x}) + iA_2(t, \mathbf{x})] \psi_3, \\ i\partial_t \psi_3 &= -\frac{i}{\varepsilon} (\partial_x - i\partial_y) \psi_2 - \frac{1}{\varepsilon^2} \psi_3 + V(t, \mathbf{x}) \psi_3 - [A_1(t, \mathbf{x}) - iA_2(t, \mathbf{x})] \psi_2. \end{aligned} \quad (2.1.12)$$

Equation (2.1.12) will collapse to (2.1.11) under the transformation $y \rightarrow -y$ and $A_2 \rightarrow -A_2$. Thus, in 2D, the Dirac equation (2.1.1) can be reduced to the following simplified PDEs with $\Phi = \Phi(t, \mathbf{x}) = (\phi_1(t, \mathbf{x}), \phi_2(t, \mathbf{x}))^T \in \mathbb{C}^2$ and $\mathbf{x} \in \mathbb{R}^2$

$$i\partial_t\Phi(t, \mathbf{x}) = \left[-\frac{i}{\varepsilon} (\sigma_1\partial_x + \sigma_2\partial_y) + \frac{1}{\varepsilon^2}\sigma_3 + V(t, \mathbf{x})I_2 - A_1(t, \mathbf{x})\sigma_1 - A_2(t, \mathbf{x})\sigma_2 \right] \Phi(t, \mathbf{x}), \quad (2.1.13)$$

where $\Phi = (\psi_1, \psi_4)^T$ (or $\Phi = (\psi_2, \psi_3)^T$ under the transformation $y \rightarrow -y$ and $A_2 \rightarrow -A_2$). Similarly, in 1D, the Dirac equation (5.1.3) can be reduced to the following simplified PDEs with $\Phi = \Phi(t, x) = (\phi_1(t, x), \phi_2(t, x))^T$

$$i\partial_t\Phi(t, x) = \left[-\frac{i}{\varepsilon}\sigma_1\partial_x + \frac{1}{\varepsilon^2}\sigma_3 \right] \Phi(t, x) + \left[V(t, x)I_2 - A_1(t, x)\sigma_1 \right] \Phi(t, x), \quad x \in \mathbb{R}, \quad (2.1.14)$$

where $\Phi = (\psi_1, \psi_4)^T$ (or $\Phi = (\psi_2, \psi_3)^T$). Again, the Dirac equation in 2D (2.1.13) and in 1D (2.1.14) can be written in a unified way in d -dimensions ($d = 1, 2$)

$$i\partial_t\Phi(t, \mathbf{x}) = \left[-\frac{i}{\varepsilon} \sum_{j=1}^d \sigma_j \partial_j + \frac{1}{\varepsilon^2} \sigma_3 \right] \Phi(t, \mathbf{x}) + \left[V(t, \mathbf{x})I_2 - \sum_{j=1}^d A_j(t, \mathbf{x})\sigma_j \right] \Phi(t, \mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^d, \quad (2.1.15)$$

and the initial condition for dynamics is given as

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

The Dirac equation (2.1.15) is dispersive and time symmetric. By introducing the position density ρ_j for the j -th component ($j = 1, 2$) and the total density ρ

$$\rho(t, \mathbf{x}) = \sum_{j=1}^2 \rho_j(t, \mathbf{x}) = \Phi^* \Phi, \quad \rho_j(t, \mathbf{x}) = |\phi_j(t, \mathbf{x})|^2, \quad (2.1.16)$$

as well as the current density $\mathbf{J}(t, \mathbf{x}) = (J_1(t, \mathbf{x}), J_2(t, \mathbf{x}))^T$

$$J_j(t, \mathbf{x}) = \frac{1}{\varepsilon} \Phi^* \sigma_j \Phi, \quad j = 1, 2, \quad (2.1.17)$$

the conservation law (3.1.5) is also satisfied [23]. In addition, the Dirac equation (2.1.15) conserves the total mass as

$$\|\Phi(t, \cdot)\|^2 := \int_{\mathbb{R}^d} |\Phi(t, \mathbf{x})|^2 d\mathbf{x} = \int_{\mathbb{R}^d} \sum_{j=1}^2 |\phi_j(t, \mathbf{x})|^2 d\mathbf{x} \equiv \|\Phi(0, \cdot)\|^2 = \|\Phi_0\|^2, \quad t \geq 0. \quad (2.1.18)$$

Again, if the electric potential V is perturbed by a constant, e.g. $V(t, \mathbf{x}) \rightarrow V(t, \mathbf{x}) + V^0$ with V^0 being a real constant, the solution $\Phi(t, \mathbf{x}) \rightarrow e^{-iV^0 t} \Phi(t, \mathbf{x})$ which implies the density of each component ρ_j ($j = 1, 2$) and the total density ρ unchanged. When $d = 1$, if the magnetic potential A_1 is perturbed by a constant, e.g. $A_1(t, \mathbf{x}) \rightarrow A_1(t, \mathbf{x}) + A_1^0$ with A_1^0 being a real constant, the solution $\Phi(t, \mathbf{x}) \rightarrow e^{iA_1^0 t \sigma_1} \Phi(t, \mathbf{x})$ implying the total density ρ unchanged; but this property is not valid when $d = 2$. When the electromagnetic potentials are time-independent, i.e. $V(t, \mathbf{x}) = V(\mathbf{x})$ and $A_j(t, \mathbf{x}) = A_j(\mathbf{x})$ for $j = 1, 2$, the following energy functional is also conserved

$$\begin{aligned} E(t) &:= \frac{1}{2} \int_{\mathbb{R}^d} \left(-\frac{i}{\varepsilon} \sum_{j=1}^d \Phi^* \sigma_j \partial_j \Phi + \frac{1}{\varepsilon^2} \Phi^* \sigma_3 \Phi + V(\mathbf{x}) |\Phi|^2 - \sum_{j=1}^d A_j(\mathbf{x}) \Phi^* \sigma_j \Phi \right) d\mathbf{x} \\ &\equiv E(0), \quad t \geq 0. \end{aligned} \quad (2.1.19)$$

Furthermore, if the external electromagnetic potentials are constants, i.e. $V(t, \mathbf{x}) \equiv V^0$ and $A_j(t, \mathbf{x}) \equiv A_j^0$ for $j = 1, 2$, the Dirac equation (2.1.15) admits the plane wave solution as $\Phi(t, \mathbf{x}) = \mathbf{B} e^{i(\mathbf{k} \cdot \mathbf{x} - \omega t)}$, where the time frequency ω , amplitude vector $\mathbf{B} \in \mathbb{R}^2$ and spatial wave number $\mathbf{k} = (k_1, \dots, k_d)^T \in \mathbb{R}^d$ satisfy the following

$$\omega \mathbf{B} = \left[\sum_{j=1}^d \left(\frac{k_j}{\varepsilon} - A_j^0 \right) \sigma_j + \frac{1}{\varepsilon^2} \sigma_3 + V^0 I_2 \right] \mathbf{B}. \quad (2.1.20)$$

Solving the above eigenvalue problem, we can get the *dispersion relation*

$$\omega = V^0 \pm \sqrt{\frac{1}{\varepsilon^4} + \left(\frac{\mathbf{k}}{\varepsilon} - \mathbf{A}^0 \right)^2} \quad (2.1.21)$$

2.2 Finite difference methods

In this section, we apply the commonly used FDTD methods to the Dirac equation (2.1.3) (or (2.1.15)) and analyze their stabilities and convergence in the nonrelativistic limit regime. For simplicity of notations, we shall only present the numerical methods and their analysis for (2.1.15) in 1D. Generalization to (2.1.3) and/or higher dimensions 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. [19, 23, 52, 53, 57, 78, 102, 106] and references therein), in practical computation, we truncate the whole space problem onto an interval $\Omega = (a, b)$ with periodic boundary conditions, which is large enough such that the truncation error is negligible. In 1D, the Dirac equation (2.1.15) with periodic boundary conditions collapses to

$$i\partial_t\Phi(t, x) = \left[-\frac{i}{\varepsilon}\sigma_1\partial_x + \frac{1}{\varepsilon^2}\sigma_3 + V(t, x)I_2 - A_1(t, x)\sigma_1 \right] \Phi(t, x), \quad x \in \Omega, t > 0, \quad (2.2.1)$$

$$\Phi(t, a) = \Phi(t, b), \quad \partial_x\Phi(t, a) = \partial_x\Phi(t, b), \quad t \geq 0, \quad \Phi(0, x) = \Phi_0(x), \quad x \in \overline{\Omega}, \quad (2.2.2)$$

where $\Phi_0(a) = \Phi_0(b)$ and $\Phi'_0(a) = \Phi'_0(b)$.

2.2.1 Finite difference time domain methods

Choose mesh size $h := \Delta x = \frac{b-a}{M}$ with M being an even positive integer, time step $\tau := \Delta t > 0$ and denote the grid points and time steps as:

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

Denote $X_M = \{U = (U_0, U_1, \dots, U_M)^T \mid U_j \in \mathbb{C}^2, j = 0, 1, \dots, M, U_0 = U_M\}$ and we always use $U_{-1} = U_{M-1}$ if it is involved. For any $U \in X_M$, we denote its Fourier representation as

$$U_j = \sum_{l=-M/2}^{M/2-1} \tilde{U}_l e^{i\mu_l(x_j-a)} = \sum_{l=-M/2}^{M/2-1} \tilde{U}_l e^{2ijl\pi/M}, \quad j = 0, 1, \dots, M, \quad (2.2.3)$$

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

$$\mu_l = \frac{2l\pi}{b-a}, \quad \tilde{U}_l = \frac{1}{M} \sum_{j=0}^{M-1} U_j e^{-2ijl\pi/M}, \quad l = -\frac{M}{2}, \dots, \frac{M}{2} - 1. \quad (2.2.4)$$

The standard l^2 -norm in X_M is given as

$$\|U\|_{l^2}^2 = h \sum_{j=0}^{M-1} |U_j|^2, \quad U \in X_M. \quad (2.2.5)$$

Let Φ_j^n be the numerical approximation of $\Phi(t_n, x_j)$ and $V_j^n = V(t_n, x_j)$, $V_j^{n+1/2} = V(t_n + \tau/2, x_j)$, $A_{1,j}^n = A_1(t_n, x_j)$ and $A_{1,j}^{n+1/2} = A_1(t_n + \tau/2, x_j)$ for $0 \leq j \leq M$ and $n \geq 0$. Denote $\Phi^n = (\Phi_0^n, \Phi_1^n, \dots, \Phi_M^n)^T \in X_M$ as the solution vector at $t = t_n$. Introduce the finite difference discretization operators for $j = 0, 1, \dots, M-1$ and $n \geq 0$ as:

$$\begin{aligned} \delta_t^+ \Phi_j^n &= \frac{\Phi_j^{n+1} - \Phi_j^n}{\tau}, & \delta_t \Phi_j^n &= \frac{\Phi_j^{n+1} - \Phi_j^{n-1}}{2\tau}, \\ \delta_x \Phi_j^n &= \frac{\Phi_{j+1}^n - \Phi_{j-1}^n}{2h}, & \Phi_j^{n+\frac{1}{2}} &= \frac{\Phi_j^{n+1} + \Phi_j^n}{2}. \end{aligned}$$

Here we consider several frequently used FDTD methods to discretize the Dirac equation (2.2.1) for $j = 0, 1, \dots, M-1$.

I. Leap-frog finite difference (LFFD) method

$$i\delta_t \Phi_j^n = \left[-\frac{i}{\varepsilon} \sigma_1 \delta_x + \frac{1}{\varepsilon^2} \sigma_3 \right] \Phi_j^n + \left[V_j^n I_2 - A_{1,j}^n \sigma_1 \right] \Phi_j^n, \quad n \geq 1. \quad (2.2.6)$$

II. Semi-implicit finite difference (SIFD1) method

$$i\delta_t \Phi_j^n = -\frac{i}{\varepsilon} \sigma_1 \delta_x \Phi_j^n + \frac{1}{\varepsilon^2} \sigma_3 \frac{\Phi_j^{n+1} + \Phi_j^{n-1}}{2} + \left[V_j^n I_2 - A_{1,j}^n \sigma_1 \right] \frac{\Phi_j^{n+1} + \Phi_j^{n-1}}{2}, \quad n \geq 1. \quad (2.2.7)$$

III. Another semi-implicit finite difference (SIFD2) method

$$i\delta_t \Phi_j^n = \left[-\frac{i}{\varepsilon} \sigma_1 \delta_x + \frac{1}{\varepsilon^2} \sigma_3 \right] \frac{\Phi_j^{n+1} + \Phi_j^{n-1}}{2} + \left[V_j^n I_2 - A_{1,j}^n \sigma_1 \right] \Phi_j^n, \quad n \geq 1. \quad (2.2.8)$$

IV. Crank-Nicolson finite difference (CNFD) method

$$i\delta_t^+ \Phi_j^n = \left[-\frac{i}{\varepsilon} \sigma_1 \delta_x + \frac{1}{\varepsilon^2} \sigma_3 \right] \Phi_j^{n+1/2} + \left[V_j^{n+1/2} I_2 - A_{1,j}^{n+1/2} \sigma_1 \right] \Phi_j^{n+1/2}, \quad n \geq 0. \quad (2.2.9)$$

The initial and boundary conditions in (2.2.2) are discretized as:

$$\Phi_M^{n+1} = \Phi_0^{n+1}, \quad \Phi_{-1}^{n+1} = \Phi_{M-1}^{n+1}, \quad n \geq 0, \quad \Phi_j^0 = \Phi_0(x_j), \quad j = 0, 1, \dots, M. \quad (2.2.10)$$

For the LFFD (2.2.6), SIFD1 (2.2.7) and SIFD2 (2.2.8), the first step can be computed as

$$\Phi_j^1 = \Phi_j^0 + \tau \left[-\frac{1}{\varepsilon} \sigma_1 \Phi_0'(x_j) - i \left(\frac{1}{\varepsilon^2} \sigma_3 + V_j^0 I_2 - A_{1,j}^0 \sigma_1 \right) \Phi_j^0 \right], \quad j = 0, 1, \dots, M-1. \quad (2.2.11)$$

The above four methods are all time symmetric, i.e. they are unchanged under $\tau \leftrightarrow -\tau$ and $n+1 \leftrightarrow n-1$ in the LFFD, SIFD1 and SIFD2 methods or $n+1 \leftrightarrow n$ in the CNFD method, and the memory costs are the same at $O(M)$. The LFFD method (2.2.6) is explicit and its computational cost per step is $O(M)$. In fact, it might be the simplest and most efficient discretization for the Dirac equation when $\varepsilon = 1$ and thus it has been widely used when $\varepsilon = 1$. The SIFD1 method (2.2.7) is implicit, however at each time step for $n \geq 1$, the corresponding linear system is decoupled and can be solved explicitly for $j = 0, 1, \dots, M-1$ as

$$\begin{aligned} \Phi_j^{n+1} &= \left[(i - \tau V_j^n) I_2 - \frac{\tau}{\varepsilon^2} \sigma_3 + \tau A_{1,j}^n \sigma_1 \right]^{-1} \cdot \\ &\quad \left[\left((i + \tau V_j^n) I_2 + \frac{\tau}{\varepsilon^2} \sigma_3 - \tau A_{1,j}^n \sigma_1 \right) \Phi_j^{n-1} - \frac{2i\tau}{\varepsilon} \sigma_1 \delta_x \Phi_j^n \right], \end{aligned} \quad (2.2.12)$$

and thus its computational cost per step is $O(M)$.

The SIFD2 method (2.2.8) is implicit, however at each time step for $n \geq 1$, the corresponding linear system is decoupled in phase (Fourier) space and can be solved explicitly in phase space for $l = -M/2, \dots, M/2 - 1$ as

$$\begin{aligned} (\widetilde{\Phi^{n+1}})_l &= \left(i I_2 - \frac{\tau \sin(\mu_l h)}{\varepsilon h} \sigma_1 - \frac{\tau}{\varepsilon^2} \sigma_3 \right)^{-1} \cdot \\ &\quad \left[\left(i I_2 + \frac{\tau \sin(\mu_l h)}{\varepsilon h} \sigma_1 + \frac{\tau}{\varepsilon^2} \sigma_3 \right) (\widetilde{\Phi^{n-1}})_l + 2\tau (\widetilde{G^n \Phi^n})_l \right], \end{aligned} \quad (2.2.13)$$

where $G^n = (G_0^n, G_1^n, \dots, G_M^n)^T \in X_M$ with $G_j^n = -A_{1,j}^n \sigma_1 + V_j^n I_2$ for $j = 0, 1, \dots, M$, and thus its computational cost per step is $O(M \ln M)$. The CNFD method (2.2.9) is implicit and at each time step for $n \geq 0$, the corresponding linear system is coupled and needs to be solved via either a direct solver or an iterative solver, and thus its computational cost per step depends on the linear system solver, which is usually much larger than $O(M)$, especially in 2D and 3D. Based on the computational cost per time step, the LFFD method is the most efficient one and the CNFD method is the most expensive one.

2.2.2 Linear stability analysis

In order to carry out the linear stability analysis for the FDTD methods via the von Neumann method [90], we assume 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 in the Dirac equation (2.2.1). Then we have the following results for the FDTD methods:

Lemma 2.1 (i) *The LFFD method (2.2.6) is stable under the stability condition*

$$0 < \tau \leq \frac{\varepsilon^2 h}{|V^0| \varepsilon^2 h + \sqrt{h^2 + \varepsilon^2 (1 + \varepsilon h |A_1^0|)^2}}, \quad h > 0, \quad 0 < \varepsilon \leq 1. \quad (2.2.14)$$

(ii) *The SIFD1 method (2.2.7) is stable under the stability condition*

$$0 < \tau \leq \varepsilon h, \quad h > 0, \quad 0 < \varepsilon \leq 1. \quad (2.2.15)$$

(iii) *The SIFD2 method (2.2.8) is stable under the stability condition*

$$0 < \tau \leq \frac{1}{|V^0| + |A_1^0|}, \quad h > 0, \quad 0 < \varepsilon \leq 1. \quad (2.2.16)$$

(iv) *The CNFD method (2.2.9) is unconditionally stable, i.e. it is stable for any $\tau, h > 0$ and $0 < \varepsilon \leq 1$.*

Proof: (i) Plugging

$$\Phi_j^n = \sum_{l=-M/2}^{M/2-1} \xi_l^n (\widetilde{\Phi^0})_l e^{i\mu_l(x_j - a)} = \sum_{l=-M/2}^{M/2-1} \xi_l^n (\widetilde{\Phi^0})_l e^{2ijl\pi/M}, \quad j = 0, 1, \dots, M, \quad n \geq 0, \quad (2.2.17)$$

with $\xi_l \in \mathbb{C}$ and $(\widetilde{\Phi^0})_l$ being the amplification factor and the Fourier coefficient at $n = 0$, respectively, of the l -th mode ($l = -\frac{M}{2}, \dots, \frac{M}{2} - 1$) in the phase space into (2.2.6), using the orthogonality of the Fourier series, we obtain

$$\left| (\xi_l^2 - 1) I_2 - 2i\tau \xi_l \left(A_1^0 \sigma_1 - V^0 I_2 - \frac{1}{\varepsilon^2} \sigma_3 - \frac{\sin(\mu_l h)}{\varepsilon h} \sigma_1 \right) \right| = 0. \quad (2.2.18)$$

Substituting (1.1.3) into (2.2.18), we get that the amplification factor ξ_l satisfies

$$\xi_l^2 - 2i\tau \theta_l \xi_l - 1 = 0, \quad (2.2.19)$$

where

$$\theta_l = -V^0 \pm \frac{1}{\varepsilon^2 h} \sqrt{h^2 + \varepsilon^2 (A_1^0 \varepsilon h - \sin(\mu_l h))^2}.$$

Then the stability condition for the LFFD method (2.2.6) becomes

$$|\xi_l| \leq 1 \iff |\tau \theta_l| \leq 1,$$

which immediately implies the condition (2.2.14).

(ii) Similar to (i), plugging (2.2.17) into the SIFD1 method (2.2.7), we have

$$\left| (\xi_l^2 - 1)I_2 - i\tau(\xi_l^2 + 1) \left(A_1^0 \sigma_1 - V^0 I_2 - \frac{1}{\varepsilon^2} \sigma_3 \right) + \frac{2i\tau \xi_l \sin(\mu_l h)}{\varepsilon h} \sigma_1 \right| = 0. \quad (2.2.20)$$

Noticing (1.1.3), under the condition (2.2.15), we can get $|\xi_l| \leq 1$ for $l = -\frac{M}{2}, \dots, \frac{M}{2} - 1$, and thus it is stable.

(iii) Similar to (i), plugging (2.2.17) into the SIFD2 method (2.2.8), we have

$$\left| (\xi_l^2 - 1)I_2 + i\tau(\xi_l^2 + 1) \left(\frac{1}{\varepsilon^2} \sigma_3 + \frac{\sin(\mu_l h)}{\varepsilon h} \sigma_1 \right) - 2i\tau \xi_l (A_1^0 \sigma_1 - V^0 I_2) \right| = 0. \quad (2.2.21)$$

Noticing (1.1.3), under the condition (2.2.16), we obtain

$$|\xi_l| \leq 1, \quad l = -\frac{M}{2}, \dots, \frac{M}{2} - 1,$$

and thus it is stable.

(iv) Similar to (i), plugging (2.2.17) into the CNFD method (2.2.9), we obtain

$$\left| (\xi_l - 1)I_2 + \frac{i\tau}{2}(\xi_l + 1) \left(\frac{1}{\varepsilon^2} \sigma_3 - A_1^0 \sigma_1 - V^0 I_2 + \frac{\sin(\mu_l h)}{\varepsilon h} \sigma_1 \right) \right| = 0. \quad (2.2.22)$$

Noticing (1.1.3), we have for $l = -\frac{M}{2}, \dots, \frac{M}{2} - 1$,

$$|\xi_l| = \left| \frac{2 + i\tau\theta_l}{2 - i\tau\theta_l} \right| = 1, \quad \theta_l = V^0 \pm \frac{1}{\varepsilon^2 h} \sqrt{h^2 + \varepsilon^2 (A_1^0 \varepsilon h - \sin(\mu_l h))^2}. \quad (2.2.23)$$

Thus it is unconditionally stable.

□

2.2.3 Mass and energy conservation

For the CNFD method (2.2.9), we have the following conservative properties.

Lemma 2.2 *The CNFD (2.2.9) conserves the mass in the discretized level, i.e.*

$$\|\Phi^n\|_{l^2}^2 := h \sum_{j=0}^{M-1} |\Phi_j^n|^2 \equiv h \sum_{j=0}^{M-1} |\Phi_j^0|^2 = \|\Phi^0\|_{l^2}^2 = h \sum_{j=0}^{M-1} |\Phi_0(x_j)|^2, \quad n \geq 0. \quad (2.2.24)$$

Furthermore, if $V(t, x) = V(x)$ and $A_1(t, x) = A_1(x)$ are time independent, the CNFD (2.2.9) conserves the energy as well,

$$\begin{aligned} E_h^n &= -\frac{i}{\varepsilon} h \sum_{j=1}^{M-1} (\Phi_j^n)^* \sigma_1 \delta_x \Phi_j^n + \frac{1}{\varepsilon^2} h \sum_{j=0}^{M-1} (\Phi_j^n)^* \sigma_3 \Phi_j^n + h \sum_{j=0}^{M-1} V_j (\Phi_j^n)^* \sigma_3 \Phi_j^n \\ &\quad - h \sum_{j=0}^{M-1} A_{1,j} (\Phi_j^n)^* \sigma_1 \Phi_j^n \\ &\equiv E_h^0, \quad n \geq 0, \end{aligned} \quad (2.2.25)$$

where $V_j = V(x_j)$ and $A_{1,j} = A_1(x_j)$ for $j = 0, 1, \dots, M$.

Proof: (i) Firstly, we prove the mass conservation (2.2.24). Multiplying both sides of (2.2.9) from left by $h\tau (\Phi_j^{n+1/2})^*$ and taking the imaginary part, we have

$$h|\Phi_j^{n+1}|^2 = h|\Phi_j^n|^2 - \frac{\tau h}{2\varepsilon} \left[(\Phi_j^{n+1/2})^* \sigma_1 \delta_x \Phi_j^{n+1/2} + (\Phi_j^{n+1/2})^T \sigma_1 \delta_x \bar{\Phi}_j^{n+1/2} \right]. \quad (2.2.26)$$

Summing (2.2.26) for $j = 0, 1, \dots, M-1$ and noticing (1.1.3), we get

$$\begin{aligned} \|\Phi^{n+1}\|_{l^2}^2 &= \|\Phi^n\|_{l^2}^2 - \frac{\tau h}{2\varepsilon} \sum_{j=0}^{M-1} \left[(\Phi_j^{n+1/2})^* \sigma_1 \delta_x \Phi_j^{n+1/2} + (\Phi_j^{n+1/2})^T \sigma_1 \delta_x \bar{\Phi}_j^{n+1/2} \right] \\ &= \|\Phi^n\|_{l^2}^2 - \frac{\tau}{2\varepsilon} \sum_{j=0}^{M-1} \left[(\Phi_j^{n+1/2})^* \sigma_1 \Phi_{j+1}^{n+1/2} + (\Phi_j^{n+1/2})^T \sigma_1 \bar{\Phi}_{j+1}^{n+1/2} \right. \\ &\quad \left. - (\Phi_{j+1}^{n+1/2})^* \sigma_1 \Phi_j^{n+1/2} - (\Phi_{j+1}^{n+1/2})^T \sigma_1 \bar{\Phi}_j^{n+1/2} \right] \\ &= \|\Phi^n\|_{l^2}^2, \quad n \geq 0, \end{aligned} \quad (2.2.27)$$

which immediately implies (2.2.24) by induction.

(ii) Secondly, we prove the energy conservation (2.2.25). Multiplying both sides of (2.2.9) from left by $2h(\Phi_j^{n+1} - \Phi_j^n)^*$ and taking the real part, we have

$$\begin{aligned} & -h \operatorname{Re} \left[\frac{i}{\varepsilon} (\Phi_j^{n+1} - \Phi_j^n)^* \sigma_1 \delta_x (\Phi_j^{n+1} + \Phi_j^n) \right] + \frac{h}{\varepsilon^2} [(\Phi_j^{n+1})^* \sigma_3 \Phi_j^{n+1} - (\Phi_j^n)^* \sigma_3 \Phi_j^n] \\ & + h V_j (|\Phi_j^{n+1}|^2 - |\Phi_j^n|^2) - h A_{1,j} [(\Phi_j^{n+1})^* \sigma_1 \Phi_j^{n+1} - (\Phi_j^n)^* \sigma_1 \Phi_j^n] = 0. \end{aligned} \quad (2.2.28)$$

Summing (2.2.28) for $j = 0, 1, \dots, M-1$ and noticing the summation by parts formula, we have

$$\begin{aligned} & h \sum_{j=0}^{M-1} \operatorname{Re} \left(\frac{i}{\varepsilon} (\Phi_j^{n+1} - \Phi_j^n)^* \sigma_1 \delta_x (\Phi_j^{n+1} + \Phi_j^n) \right) \\ & = \frac{ih}{\varepsilon} \sum_{j=0}^{M-1} (\Phi_j^{n+1})^* \sigma_1 \delta_x \Phi_j^{n+1} - \frac{ih}{\varepsilon} \sum_{j=0}^{M-1} (\Phi_j^n)^* \sigma_1 \delta_x \Phi_j^n, \end{aligned}$$

and

$$\begin{aligned} 0 = & -\frac{ih}{\varepsilon} \sum_{j=0}^{M-1} (\Phi_j^{n+1})^* \sigma_1 \delta_x \Phi_j^{n+1} + \frac{ih}{\varepsilon} \sum_{j=0}^{M-1} (\Phi_j^n)^* \sigma_1 \delta_x \Phi_j^n \\ & + \frac{h}{\varepsilon^2} \sum_{j=0}^{M-1} ((\Phi_j^{n+1})^* \sigma_3 \Phi_j^{n+1} - (\Phi_j^n)^* \sigma_3 \Phi_j^n) + h \sum_{j=0}^{M-1} V_j (|\Phi_j^{n+1}|^2 - |\Phi_j^n|^2) \\ & - h \sum_{j=0}^{M-1} A_{1,j} ((\Phi_j^{n+1})^* \sigma_1 \Phi_j^{n+1} - (\Phi_j^n)^* \sigma_1 \Phi_j^n), \end{aligned}$$

which immediately implies (2.2.25). □

2.2.4 Main results on error estimates

Let $0 < T < T^*$ with T^* being the maximal existence time of the solution, and denote $\Omega_T = [0, T] \times \Omega$. Motivated by the nonrelativistic limit of the Dirac equation [20] and the dispersion relation (3.1.18), we assume that the exact solution of (2.2.1) satisfies $\Phi \in C^3([0, T]; (L^\infty(\Omega))^2) \cap C^2([0, T]; (W_p^{1,\infty}(\Omega))^2) \cap C^1([0, T]; (W_p^{2,\infty}(\Omega))^2) \cap C([0, T]; (W_p^{3,\infty}(\Omega))^2)$ and

$$(A) \left\| \frac{\partial^{r+s}}{\partial t^r \partial x^s} \Phi \right\|_{L^\infty([0,T]; (L^\infty(\Omega))^2)} \lesssim \frac{1}{\varepsilon^{2r}}, \quad 0 \leq r \leq 3, \quad 0 \leq r+s \leq 3, \quad 0 < \varepsilon \leq 1, \quad (2.2.29)$$

where $W_p^{m,\infty}(\Omega) = \{u \mid u \in W^{m,\infty}(\Omega), \partial_x^l u(a) = \partial_x^l u(b), l = 0, \dots, m-1\}$ for $m \geq 1$ 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

$$(B) \quad V_{\max} := \max_{(t,x) \in \overline{\Omega}_T} |V(t,x)|, \quad A_{1,\max} := \max_{(t,x) \in \overline{\Omega}_T} |A_1(t,x)|. \quad (2.2.30)$$

Define the grid error function $\mathbf{e}^n = (\mathbf{e}_0^n, \mathbf{e}_1^n, \dots, \mathbf{e}_M^n)^T \in X_M$ as:

$$\mathbf{e}_j^n = \Phi(t_n, x_j) - \Phi_j^n, \quad j = 0, 1, \dots, M, \quad n \geq 0, \quad (2.2.31)$$

with Φ_j^n being the approximations obtained from the FDTD methods.

For the CNFD (2.2.9), we can establish the following error bound.

Theorem 2.1 *Under the assumptions (A) and (B), there exist constants $h_0 > 0$ and $\tau_0 > 0$ sufficiently small and independent of ε , such that for any $0 < \varepsilon \leq 1$, $0 < h \leq h_0$ and $0 < \tau \leq \tau_0$, we have the following error estimate for the CNFD (2.2.9) with (2.2.10)*

$$\|\mathbf{e}^n\|_{l^2} \lesssim \frac{h^2}{\varepsilon} + \frac{\tau^2}{\varepsilon^6}, \quad 0 \leq n \leq \frac{T}{\tau}. \quad (2.2.32)$$

For the LFFD (2.2.6), we assume the stability condition

$$0 < \tau \leq \frac{\varepsilon^2 h}{\varepsilon^2 h V_{\max} + \sqrt{h^2 + \varepsilon^2 (1 + \varepsilon h A_{1,\max})^2}}, \quad h > 0, \quad 0 < \varepsilon \leq 1, \quad (2.2.33)$$

and establish the following error estimate.

Theorem 2.2 *Under the assumptions (A) and (B), there exist constants $h_0 > 0$ and $\tau_0 > 0$ sufficiently small and independent of ε , such that for any $0 < \varepsilon \leq 1$, when $0 < h \leq h_0$ and $0 < \tau \leq \tau_0$ and under the stability condition (2.2.33), we have the following error estimate for the LFFD (2.2.6) with (2.2.10) and (2.2.11)*

$$\|\mathbf{e}^n\|_{l^2} \lesssim \frac{h^2}{\varepsilon} + \frac{\tau^2}{\varepsilon^6}, \quad 0 \leq n \leq \frac{T}{\tau}. \quad (2.2.34)$$

Similar to the proofs of the LFFD and CNFD methods, error estimates for SIFD1 (2.2.7) and SIFD2 (2.2.8) can be derived and the details are omitted here for brevity. For the SIFD2 (2.2.8), we assume the stability condition

$$0 < \tau \leq \frac{1}{V_{\max} + A_{1,\max}}, \quad h > 0, \quad 0 < \varepsilon \leq 1, \quad (2.2.35)$$

and establish the following error estimates.

Theorem 2.3 *Under the assumptions (A) and (B), there exist constants $h_0 > 0$ and $\tau_0 > 0$ sufficiently small and independent of ε , such that for any $0 < \varepsilon \leq 1$, when $0 < h \leq h_0$ and $0 < \tau \leq \tau_0$ and under the stability condition (2.2.15), we have the following error estimate for the SIFD1 (2.2.7) with (2.2.10) and (2.2.11)*

$$\|\mathbf{e}^n\|_{l^2} \lesssim \frac{h^2}{\varepsilon} + \frac{\tau^2}{\varepsilon^6}, \quad 0 \leq n \leq \frac{T}{\tau}.$$

Theorem 2.4 *Under the assumptions (A) and (B), there exist constants $h_0 > 0$ and $\tau_0 > 0$ sufficiently small and independent of ε , such that for any $0 < \varepsilon \leq 1$, when $0 < h \leq h_0$ and $0 < \tau \leq \tau_0$ and under the stability condition (2.2.35), we have the following error estimate for the SIFD2 (2.2.8) with (2.2.10) and (2.2.11)*

$$\|\mathbf{e}^n\|_{l^2} \lesssim \frac{h^2}{\varepsilon} + \frac{\tau^2}{\varepsilon^6}, \quad 0 \leq n \leq \frac{T}{\tau}.$$

Based on Theorems 2.1-2.4, the four FDTD methods studied here share the same temporal/spatial resolution capacity in the nonrelativistic limit regime. In fact, given an accuracy bound $\delta > 0$, the ε -scalability of the four FDTD methods is:

$$\tau = O\left(\varepsilon^3 \sqrt{\delta}\right) = O(\varepsilon^3), \quad h = O\left(\sqrt{\delta \varepsilon}\right) = O(\sqrt{\varepsilon}), \quad 0 < \varepsilon \ll 1.$$

2.2.5 Proof of the error estimates for the CNFD method

Proof of Theorem 2.1. Define the local truncation error $\xi^n = (\xi_0^n, \xi_1^n, \dots, \xi_M^n)^T \in X_M$ of the CNFD (2.2.9) with (2.2.10) as

$$\begin{aligned} \xi_j^n := & i\delta_t^+ \Phi(t_n, x_j) + \frac{i}{\varepsilon} \sigma_1 \frac{\delta_x \Phi(t_{n+1}, x_j) + \delta_x \Phi(t_n, x_j)}{2} \\ & + \left[-\frac{1}{\varepsilon^2} \sigma_3 + A_1(t_{n+1/2}, x_j) \sigma_1 - V(t_{n+1/2}, x_j) I_2 \right] \frac{\Phi(t_{n+1}, x_j) + \Phi(t_n, x_j)}{2}, \end{aligned} \quad (2.2.36)$$

for $0 \leq j \leq M-1$, $n \geq 0$. Applying the Taylor expansion in (2.2.36), noticing (2.2.1) and the assumptions (A) and (B), and using the triangle inequality, for $0 < \varepsilon \leq 1$, we obtain

$$\begin{aligned} |\xi_j^n| &\leq \frac{\tau^2}{24} \|\partial_{ttt}\Phi\|_{L^\infty(\bar{\Omega}_T)} + \frac{h^2}{6\varepsilon} \|\partial_{xxx}\Phi\|_{L^\infty(\bar{\Omega}_T)} + \frac{\tau^2}{8\varepsilon} \|\partial_{xtt}\Phi\|_{L^\infty(\bar{\Omega}_T)} \\ &\quad + \frac{\tau^2}{8} \left(\frac{1}{\varepsilon^2} + V_{\max} + A_{1,\max} \right) \|\partial_{tt}\Phi\|_{L^\infty(\bar{\Omega}_T)} \\ &\lesssim \frac{\tau^2}{\varepsilon^6} + \frac{h^2}{\varepsilon} + \frac{\tau^2}{\varepsilon^5} + \frac{\tau^2}{\varepsilon^6} \lesssim \frac{\tau^2}{\varepsilon^6} + \frac{h^2}{\varepsilon}, \quad j = 0, 1, \dots, M-1, \quad n \geq 0, \end{aligned}$$

which immediately implies

$$\|\xi^n\|_{l^\infty} = \max_{0 \leq j \leq M-1} |\xi_j^n| \lesssim \frac{\tau^2}{\varepsilon^6} + \frac{h^2}{\varepsilon}, \quad \|\xi^n\|_{l^2} \lesssim \|\xi^n\|_{l^\infty} \lesssim \frac{\tau^2}{\varepsilon^6} + \frac{h^2}{\varepsilon}, \quad n \geq 0, \quad 0 < \varepsilon \leq 1. \quad (2.2.37)$$

Subtracting (2.2.9) from (2.2.36), noticing (2.2.31), we get for $n \geq 0$

$$i\delta_t^+ \mathbf{e}_j^n = -\frac{i}{\varepsilon} \sigma_1 \delta_x \mathbf{e}_j^{n+1/2} + \frac{1}{\varepsilon^2} \sigma_3 \mathbf{e}_j^{n+1/2} + \left(V_j^{n+1/2} I_2 - A_{1,j}^{n+1/2} \sigma_1 \right) \mathbf{e}_j^{n+1/2} + \xi_j^n, \quad (2.2.38)$$

with $\mathbf{e}_j^{n+1/2} = \frac{\mathbf{e}_j^{n+1} + \mathbf{e}_j^n}{2}$ for $j = 0, 1, \dots, M$, and the boundary and initial conditions are given as

$$\mathbf{e}_0^n = \mathbf{e}_M^n, \quad \mathbf{e}_{-1}^n = \mathbf{e}_{M-1}^n, \quad n \geq 0, \quad \mathbf{e}_j^0 = \mathbf{0}, \quad j = 0, 1, \dots, M. \quad (2.2.39)$$

Similar to the proof for Lemma 2.2, multiplying (2.2.38) from the left by $h\tau \left(\mathbf{e}_j^{n+1/2} \right)^*$, taking the imaginary part, then summing for $j = 0, 1, \dots, M-1$, using the triangle inequality and Young's inequality, noticing (1.1.3), (2.2.37) and (2.2.39), we get

$$\begin{aligned} \|\mathbf{e}^{n+1}\|_{l^2}^2 - \|\mathbf{e}^n\|_{l^2}^2 &\lesssim \tau h \sum_{j=0}^{M-1} |\xi_j^n| \left(|\mathbf{e}_j^{n+1}| + |\mathbf{e}_j^n| \right) \lesssim \tau \left(\|\xi^n\|_{l^2}^2 + \|\mathbf{e}^{n+1}\|_{l^2}^2 + \|\mathbf{e}^n\|_{l^2}^2 \right) \\ &\lesssim \tau \left(\|\mathbf{e}^{n+1}\|_{l^2}^2 + \|\mathbf{e}^n\|_{l^2}^2 \right) + \tau \left(\frac{h^2}{\varepsilon} + \frac{\tau^2}{\varepsilon^6} \right)^2, \quad n \geq 0. \end{aligned}$$

Summing the above inequality for $n = 0, 1, \dots, m-1$, we get

$$\|\mathbf{e}^m\|_{l^2}^2 - \|\mathbf{e}^0\|_{l^2}^2 \lesssim \tau \sum_{k=0}^m \|\mathbf{e}^k\|_{l^2}^2 + \tau m \left(\frac{h^2}{\varepsilon} + \frac{\tau^2}{\varepsilon^6} \right)^2, \quad 0 \leq m \leq \frac{T}{\tau}.$$

Taking τ_0 sufficiently small, when $0 < \tau \leq \tau_0$, we have

$$\begin{aligned} \|\mathbf{e}^m\|_{l^2}^2 &\lesssim \tau \sum_{k=0}^{m-1} \|\mathbf{e}^k\|_{l^2}^2 + \tau m \left(\frac{h^2}{\varepsilon} + \frac{\tau^2}{\varepsilon^6} \right)^2 \\ &\leq \tau \sum_{k=0}^{m-1} \|\mathbf{e}^k\|_{l^2}^2 + T \left(\frac{h^2}{\varepsilon} + \frac{\tau^2}{\varepsilon^6} \right)^2, \quad 0 \leq m \leq \frac{T}{\tau}. \end{aligned}$$

Using the discrete Gronwall's inequality and noticing $\|\mathbf{e}^0\|_{l^2} = 0$, we obtain

$$\|\mathbf{e}^m\|_{l^2}^2 \lesssim T \left(\frac{h^2}{\varepsilon} + \frac{\tau^2}{\varepsilon^6} \right)^2 \lesssim \left(\frac{h^2}{\varepsilon} + \frac{\tau^2}{\varepsilon^6} \right)^2, \quad 0 \leq m \leq \frac{T}{\tau},$$

which immediately implies the error bound (2.2.32). □

2.2.6 Proof of the error estimates for the LFFD method

Proof of Theorem 2.2. Define the local truncation error $\tilde{\xi}^n = (\tilde{\xi}_0^n, \tilde{\xi}_1^n, \dots, \tilde{\xi}_M^n)^T \in X_M$ of the LFFD (2.2.6) with (2.2.10) and (2.2.11) as follows, for $0 \leq j \leq M-1$,

$$\tilde{\xi}_j^n := i\delta_t \Phi(t_n, x_j) + \frac{i}{\varepsilon} \sigma_1 \delta_x \Phi(t_n, x_j) - \left(\frac{1}{\varepsilon^2} \sigma_3 - V_j^n I_2 + A_{1,j}^n \sigma_1 \right) \Phi(t_n, x_j), \quad n \geq 1, \quad (2.2.40)$$

$$\tilde{\xi}_j^0 := i\delta_t^+ \Phi(0, x_j) + \frac{i}{\varepsilon} \sigma_1 \delta_x \Phi_0(x_j) - \left(\frac{1}{\varepsilon^2} \sigma_3 + V_j^0 I_2 - A_{1,j}^0 \sigma_1 \right) \Phi_0(x_j). \quad (2.2.41)$$

Applying the Taylor expansion in (2.2.40) and (2.2.41), noticing (2.2.1) and the assumptions (A) and (B), Similar to the proof of Theorem 2.1, we obtain

$$|\tilde{\xi}_j^0| \lesssim \frac{\tau}{\varepsilon^4} + \frac{h^2}{\varepsilon}, \quad |\tilde{\xi}_j^n| \lesssim \frac{\tau^2}{\varepsilon^6} + \frac{h^2}{\varepsilon}, \quad j = 0, 1, \dots, M-1, \quad n \geq 1,$$

which immediately implies

$$\|\tilde{\xi}^n\|_{l^\infty} = \max_{0 \leq j \leq M-1} |\tilde{\xi}_j^n| \lesssim \frac{\tau^2}{\varepsilon^6} + \frac{h^2}{\varepsilon}, \quad \|\tilde{\xi}^n\|_{l^2} \lesssim \|\tilde{\xi}^n\|_{l^\infty} \lesssim \frac{\tau^2}{\varepsilon^6} + \frac{h^2}{\varepsilon}, \quad n \geq 1, \quad 0 < \varepsilon \leq 1. \quad (2.2.42)$$

Subtracting (2.2.6) from (2.2.40), noticing (2.2.31), we get

$$i\delta_t \mathbf{e}_j^n = -\frac{i}{\varepsilon} \sigma_1 \delta_x \mathbf{e}_j^n + \frac{1}{\varepsilon^2} \sigma_3 \mathbf{e}_j^n + (V_j^n I_2 - A_{1,j}^n \sigma_1) \mathbf{e}_j^n + \tilde{\xi}_j^n, \quad 0 \leq j \leq M-1, \quad n \geq 1, \quad (2.2.43)$$

where the boundary and initial conditions are given as

$$\mathbf{e}_0^n = \mathbf{e}_M^n, \quad \mathbf{e}_{-1}^n = \mathbf{e}_{M-1}^n, \quad n \geq 0, \quad \mathbf{e}_j^0 = \mathbf{0}, \quad j = 0, 1, \dots, M. \quad (2.2.44)$$

For the first step, we have

$$\|\mathbf{e}^1\|_{l^2} = \tau \|\tilde{\xi}^0\|_{l^2} \lesssim \frac{\tau^2}{\varepsilon^4} + \frac{\tau h^2}{\varepsilon} \lesssim \frac{h^2}{\varepsilon} + \frac{\tau^2}{\varepsilon^6}. \quad (2.2.45)$$

Denote \mathcal{E}^{n+1} for $n = 0, 1, \dots$ as

$$\mathcal{E}^{n+1} = \|\mathbf{e}^{n+1}\|_{l^2}^2 + \|\mathbf{e}^n\|_{l^2}^2 + 2 \operatorname{Re} \left(\tau h \sum_{j=0}^{M-1} (\mathbf{e}_j^{n+1})^* \sigma_1 \delta_x \mathbf{e}_j^n \right) - 2 \operatorname{Im} \left(\frac{\tau h}{\varepsilon^2} \sum_{j=0}^{M-1} (\mathbf{e}_j^{n+1})^* \sigma_3 \mathbf{e}_j^n \right),$$

and under the stability condition (2.2.33), e.g., $\tau \leq \frac{\varepsilon^2 \tau_1 h}{\varepsilon^2 h V_{\max} + \sqrt{h^2 + \varepsilon^2 (1 + \varepsilon h A_{1, \max})^2}}$ with $\tau_1 = \frac{1}{4}$, which implies $\frac{\tau}{h} \leq \frac{1}{4}$ and $\frac{\tau}{\varepsilon^2} \leq \frac{1}{4}$, using Cauchy inequality, we can get that

$$\frac{1}{2} (\|\mathbf{e}^{n+1}\|_{l^2}^2 + \|\mathbf{e}^n\|_{l^2}^2) \leq \mathcal{E}^{n+1} \leq \frac{3}{2} (\|\mathbf{e}^{n+1}\|_{l^2}^2 + \|\mathbf{e}^n\|_{l^2}^2). \quad (2.2.46)$$

It follows from (2.2.45) that

$$\mathcal{E}^1 \lesssim \left(\frac{h^2}{\varepsilon} + \frac{\tau^2}{\varepsilon^6} \right)^2. \quad (2.2.47)$$

Similar to the proof of Theorem 2.1, multiplying both sides of (2.2.43) from the left by $2h\tau (\mathbf{e}_j^{n+1} + \mathbf{e}_j^{n-1})^*$, taking the imaginary part, then summing for $j = 0, 1, \dots, M-1$, using Cauchy inequality, (2.2.42) and (2.2.46), we get for $n \geq 1$,

$$\begin{aligned} \mathcal{E}^{n+1} - \mathcal{E}^n &\lesssim h\tau \sum_{j=0}^{M-1} \left((A_{1, \max} + V_{\max}) |\mathbf{e}_j^n| + |\tilde{\xi}_j^n| \right) (|\mathbf{e}_j^{n+1}| + |\mathbf{e}_j^{n-1}|) \\ &\lesssim \tau (\mathcal{E}^{n+1} + \mathcal{E}^n) + \tau \left(\frac{h^2}{\varepsilon} + \frac{\tau^2}{\varepsilon^6} \right)^2. \end{aligned}$$

Summing the above inequality for $n = 1, 2, \dots, m-1$, we get

$$\mathcal{E}^m - \mathcal{E}^1 \lesssim \tau \sum_{k=1}^m \mathcal{E}^k + m\tau \left(\frac{h^2}{\varepsilon} + \frac{\tau^2}{\varepsilon^6} \right)^2.$$

Taking τ_0 sufficiently small, using the discrete Gronwall's inequality and noticing (3.2.55), we obtain from above equation that

$$\mathcal{E}^m \lesssim \left(\frac{h^2}{\varepsilon} + \frac{\tau^2}{\varepsilon^6} \right)^2, \quad 1 \leq m \leq \frac{T}{\tau},$$

which immediately implies the error bound (2.2.34) in view of (2.2.46).

□

2.3 Exponential wave integrator pseudospectral methods

In this section, we propose an exponential wave integrator Fourier pseudospectral (EWI-FP) method to solve the Dirac equation (2.1.3) (or (2.1.15)) and establish its stability and convergence in the nonrelativistic limit regime. Again, for simplicity of notations, we shall only present the numerical method and its analysis for (2.2.1) in 1D. Generalization to (2.1.3) and/or higher dimensions is straightforward.

2.3.1 The EWI-FP method in 1D

Denote

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

Let $[C_p(\bar{\Omega})]^2$ be the function space consisting of all periodic vector function $U(x) : \bar{\Omega} = [a, b] \rightarrow \mathbb{C}^2$. For any $U(x) \in [C_p(\bar{\Omega})]^2$ and $U \in X_M$, define $P_M : [L^2(\Omega)]^2 \rightarrow Y_M$ as the standard projection operator [89], $I_M : [C_p(\bar{\Omega})]^2 \rightarrow Y_M$ and $I_M : X_M \rightarrow Y_M$ as the standard interpolation operator [89], i.e.

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

with

$$\begin{aligned} \widehat{U}_l &= \frac{1}{b-a} \int_a^b U(x) e^{-i\mu_l(x-a)} dx, \\ \widetilde{U}_l &= \frac{1}{M} \sum_{j=0}^{M-1} U_j e^{-2ij\pi/M}, \quad l = -\frac{M}{2}, -\frac{M}{2} + 1, \dots, \frac{M}{2} - 1, \end{aligned} \quad (2.3.1)$$

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

The Fourier spectral discretization for Dirac equation (2.2.1) is as follows:

Find $\Phi_M(t, x) \in Y_M$, i.e.

$$\Phi_M(t, x) = \sum_{l=-M/2}^{M/2-1} (\widehat{\Phi_M})_l(t) e^{i\mu_l(x-a)}, \quad a \leq x \leq b, \quad t \geq 0, \quad (2.3.2)$$

such that for $a < x < b$ and $t > 0$,

$$i\partial_t \Phi_M(t, x) = \left[-\frac{i}{\varepsilon} \sigma_1 \partial_x + \frac{1}{\varepsilon^2} \sigma_3 \right] \Phi_M(t, x) + P_M(V\Phi_M)(t, x) - \sigma_1 P_M(A_1\Phi_M)(t, x). \quad (2.3.3)$$

Substituting (2.3.2) into (2.3.3), noticing the orthogonality of $\phi_l(x)$, we get for $l = -\frac{M}{2}, -\frac{M}{2} + 1, \dots, \frac{M}{2} - 1$,

$$i \frac{d}{dt} (\widehat{\Phi_M})_l(t) = \left[\frac{\mu_l}{\varepsilon} \sigma_1 + \frac{1}{\varepsilon^2} \sigma_3 \right] (\widehat{\Phi_M})_l(t) + (\widehat{V\Phi_M})_l(t) - \sigma_1 (\widehat{A_1\Phi_M})_l(t), \quad t \geq 0.$$

For each l ($l = -\frac{M}{2}, -\frac{M}{2} + 1, \dots, \frac{M}{2} - 1$), when t is near $t = t_n$ ($n \geq 0$), we rewrite the above ODEs as

$$i \frac{d}{ds} (\widehat{\Phi_M})_l(t_n + s) = \frac{1}{\varepsilon^2} \Gamma_l (\widehat{\Phi_M})_l(t_n + s) + \widehat{F}_l^n(s), \quad s \in \mathbb{R}, \quad (2.3.4)$$

where

$$\widehat{F}_l^n(s) = (\widehat{G\Phi_M})_l(t_n + s), \quad G(t, x) = V(t, x)I_2 - \sigma_1 A_1(t, x), \quad s, t \in \mathbb{R}, \quad (2.3.5)$$

and $\Gamma_l = \mu_l \varepsilon \sigma_1 + \sigma_3 = Q_l D_l (Q_l)^*$ with

$$\Gamma_l = \begin{pmatrix} 1 & \mu_l \varepsilon \\ \mu_l \varepsilon & -1 \end{pmatrix}, \quad Q_l = \begin{pmatrix} \frac{1+\delta_l}{\sqrt{2\delta_l(1+\delta_l)}} & -\frac{\varepsilon\mu_l}{\sqrt{2\delta_l(1+\delta_l)}} \\ \frac{\varepsilon\mu_l}{\sqrt{2\delta_l(1+\delta_l)}} & \frac{1+\delta_l}{\sqrt{2\delta_l(1+\delta_l)}} \end{pmatrix}, \quad D_l = \begin{pmatrix} \delta_l & 0 \\ 0 & -\delta_l \end{pmatrix}. \quad (2.3.6)$$

where $\delta_l = \sqrt{1 + \varepsilon^2 \mu_l^2}$.

Solving the above ODE (2.3.4) via the integrating factor method, we obtain

$$(\widehat{\Phi_M})_l(t_n + s) = e^{-is\Gamma_l/\varepsilon^2} (\widehat{\Phi_M})_l(t_n) - i \int_0^s e^{i(w-s)\Gamma_l/\varepsilon^2} \widehat{F}_l^n(w) dw, \quad s \in \mathbb{R}. \quad (2.3.7)$$

Taking $s = \tau$ in (2.3.7) we have

$$(\widehat{\Phi_M})_l(t_{n+1}) = e^{-i\tau\Gamma_l/\varepsilon^2} (\widehat{\Phi_M})_l(t_n) - i \int_0^\tau e^{\frac{i(w-\tau)\Gamma_l}{\varepsilon^2}} \widehat{F}_l^n(w) dw. \quad (2.3.8)$$

To obtain a numerical method with second order accuracy in time, we approximate the integrals in (2.3.8) via Gautschi-type rules, which have been widely used for integrating highly oscillatory ODEs [17, 44, 51, 54], as

$$\int_0^\tau e^{\frac{i(w-\tau)\Gamma_l}{\varepsilon^2}} \widehat{F}_l^0(w) dw \approx \int_0^\tau e^{\frac{i(w-\tau)\Gamma_l}{\varepsilon^2}} dw \widehat{F}_l^0(0) = -i\varepsilon^2 \Gamma_l^{-1} \left[I_2 - e^{-\frac{i\tau}{\varepsilon^2} \Gamma_l} \right] \widehat{F}_l^0(0), \quad (2.3.9)$$

and

$$\begin{aligned}
& \int_0^\tau e^{\frac{i(w-\tau)}{\varepsilon^2}\Gamma_l} \widehat{F}_l^n(w) dw \\
& \approx \int_0^\tau e^{\frac{i(w-\tau)}{\varepsilon^2}\Gamma_l} \left(\widehat{F}_l^n(0) + w \delta_t^- \widehat{F}_l^n(0) \right) dw \\
& = -i\varepsilon^2 \Gamma_l^{-1} \left[I - e^{-\frac{i\tau}{\varepsilon^2}\Gamma_l} \right] \widehat{F}_l^n(0) + \left[-i\varepsilon^2 \tau \Gamma_l^{-1} + \varepsilon^4 \Gamma_l^{-2} \left(I - e^{-\frac{i\tau}{\varepsilon^2}\Gamma_l} \right) \right] \delta_t^- \widehat{F}_l^n(0), \quad n \geq 1,
\end{aligned} \tag{2.3.10}$$

where we have approximated the time derivative $\partial_t \widehat{F}_l^n(s)$ at $s = 0$ by finite difference as

$$\partial_t \widehat{F}_l^n(0) \approx \delta_t^- \widehat{F}_l^n(0) = \frac{\widehat{F}_l^n(0) - \widehat{F}_l^{n-1}(0)}{\tau}.$$

Now, we are ready to describe our scheme. Let $\Phi_M^n(x)$ be the approximation of $\Phi_M(t_n, x)$ ($n \geq 0$). Choosing $\Phi_M^0(x) = (P_M \Phi_0)(x)$, an *exponential wave integrator Fourier spectral* (EWI-FS) discretization for the Dirac equation (2.2.1) is to update the numerical approximation $\Phi_M^{n+1}(x) \in Y_M$ ($n = 0, 1, \dots$) as

$$\Phi_M^{n+1}(x) = \sum_{l=-M/2}^{M/2-1} \widehat{(\Phi_M^{n+1})}_l e^{i\mu_l(x-a)}, \quad a \leq x \leq b, \quad n \geq 0, \tag{2.3.11}$$

where for $l = -\frac{M}{2}, \dots, \frac{M}{2} - 1$,

$$\widehat{(\Phi_M^{n+1})}_l = \begin{cases} e^{-i\tau\Gamma_l/\varepsilon^2} \widehat{(\Phi_M^0)}_l - i\varepsilon^2 \Gamma_l^{-1} \left[I_2 - e^{-\frac{i\tau}{\varepsilon^2}\Gamma_l} \right] (G(t_0) \widehat{\Phi_M^0})_l, & n = 0, \\ e^{-i\tau\Gamma_l/\varepsilon^2} \widehat{(\Phi_M^n)}_l - iQ_l^{(1)}(\tau) (G(t_n) \widehat{\Phi_M^n})_l - iQ_l^{(2)}(\tau) \delta_t^- (G(t_n) \widehat{\Phi_M^n})_l, & n \geq 1, \end{cases} \tag{2.3.12}$$

with the matrices $Q_l^{(1)}(\tau)$ and $Q_l^{(2)}(\tau)$ given as

$$Q_l^{(1)}(\tau) = -i\varepsilon^2 \Gamma_l^{-1} \left[I - e^{-\frac{i\tau}{\varepsilon^2}\Gamma_l} \right], \quad Q_l^{(2)}(\tau) = -i\varepsilon^2 \tau \Gamma_l^{-1} + \varepsilon^4 \Gamma_l^{-2} \left(I - e^{-\frac{i\tau}{\varepsilon^2}\Gamma_l} \right).$$

The above procedure is not suitable in practice due to the difficulty in computing the Fourier coefficients through integrals in (2.3.1). Here we present an efficient implementation by choosing $\Phi_M^0(x)$ as the interpolant of $\Phi_0(x)$ on the grids $\{x_j, j = 0, 1, \dots, M\}$ and approximate the integrals in (2.3.1) by a quadrature rule.

Let Φ_j^n be the numerical approximation of $\Phi(t_n, x_j)$ for $j = 0, 1, 2, \dots, M$ and $n \geq 0$, and denote $\Phi^n \in X_M$ as the vector with components Φ_j^n . Choosing $\Phi_j^0 =$

$\Phi_0(x_j)$ ($j = 0, 1, \dots, M$), an *EWI Fourier pseudospectral* (EWI-FP) method for computing Φ^{n+1} for $n \geq 0$ reads

$$\Phi_j^{n+1} = \sum_{l=-M/2}^{M/2-1} \widetilde{(\Phi^{n+1})}_l e^{2ijl\pi/M}, \quad j = 0, 1, \dots, M, \quad (2.3.13)$$

where

$$\widetilde{(\Phi^{n+1})}_l = \begin{cases} e^{-i\tau\Gamma_l/\varepsilon^2} \widetilde{(\Phi_0)}_l - i\varepsilon^2 \Gamma_l^{-1} \left[I_2 - e^{-\frac{i\tau}{\varepsilon^2} \Gamma_l} \right] (G(t_0) \widetilde{\Phi_0})_l, & n = 0, \\ e^{-i\tau\Gamma_l/\varepsilon^2} \widetilde{(\Phi^n)}_l - iQ_l^{(1)}(\tau) (G(t_n) \widetilde{\Phi^n})_l - iQ_l^{(2)}(\tau) \delta_t^- (G(t_n) \widetilde{\Phi^n})_l, & n \geq 1. \end{cases} \quad (2.3.14)$$

The EWI-FP (2.3.13)-(2.3.14) is explicit, and can be solved efficiently by the fast Fourier transform (FFT). The memory cost is $O(M)$ and the computational cost per time step is $O(M \log M)$.

2.3.2 Linear stability analysis

To consider the linear stability, we assume that in the Dirac equation (2.2.1), the external potential fields are constants, i.e. $A_1(t, x) \equiv A_1^0$ and $V(t, x) \equiv V^0$ with A_1^0 and V^0 being two real constants. In this case, we adopt the Von Neumann stability requirement that the errors grow exponentially at most. Then we have

Theorem 2.5 *The EWI-FP method (2.3.13)-(2.3.14) and EWI-FS method (2.3.11)-(2.3.12) are stable under the stability condition*

$$0 < \tau \lesssim 1, \quad 0 < \varepsilon \leq 1. \quad (2.3.15)$$

Proof of Theorem 2.5. We shall only prove the EWI-FS case (2.3.11)-(2.3.12), as the EWI-FP method case (2.3.14) is quite the same. Similar to the proof of Theorem 2.1, noticing (2.3.10), we find that,

$$\xi_l^2(\widetilde{\Phi^0})_l = \xi_l e^{-i\tau\Gamma_l/\varepsilon^2} (\widetilde{\Phi^0})_l - i \int_0^\tau e^{i(w-\tau)\Gamma_l/\varepsilon^2} (V^0 I_2 - A_1^0 \sigma_1) \left(\xi_l + \frac{w}{\tau} (\xi_l - 1) \right) (\widetilde{\Phi^0})_l dw. \quad (2.3.16)$$

Denoting $C = |V^0| + |A_1^0|$, taking the l^2 norms of the vectors on both sides of (2.3.16) and then dividing both sides by the l^2 norm of $(\widetilde{\Phi}^0)_l$, in view of the properties of $e^{-is\Gamma_l/\varepsilon^2}$, we get

$$|\xi_l|^2 \leq \left(1 + C\tau + \frac{C}{2}\tau\right) |\xi_l| + \frac{C}{2}\tau, \quad (2.3.17)$$

which implies

$$\left(|\xi_l| - \frac{1 + 3C\tau/2}{2}\right)^2 \leq \frac{1 + 5C\tau + 9C^2\tau^2/4}{4} \leq \frac{(1 + 5C\tau/2)^2}{4}.$$

Thus, we obtain

$$|\xi_l| \leq 1 + 4C\tau, \quad l = -\frac{M}{2}, \dots, \frac{M}{2} - 1,$$

and it follows that EWI-FS (2.3.11)-(2.3.12) is stable. \square

2.3.3 Convergence analysis

In order to obtain an error estimate for the EWI methods (2.3.11)-(2.3.12) and (2.3.13)-(2.3.14), motivated by the results in [21, 26], we assume that there exists an integer $m_0 \geq 2$ such that the exact solution $\Phi(t, x)$ of Dirac equation (2.2.1) satisfies

$$(C) \quad \|\Phi\|_{L^\infty([0, T]; (H_p^{m_0})^2)} \lesssim 1, \quad \|\partial_t \Phi\|_{L^\infty([0, T]; (L^2)^2)} \lesssim \frac{1}{\varepsilon^2}, \quad \|\partial_{tt} \Phi\|_{L^\infty([0, T]; (L^2)^2)} \lesssim \frac{1}{\varepsilon^4},$$

where $H_p^k(\Omega) = \{u \mid u \in H^k(\Omega), \partial_x^l u(a) = \partial_x^l u(b), l = 0, \dots, k-1\}$. In addition, we assume electromagnetic potentials satisfy

$$(D) \quad \|V\|_{W^{2, \infty}([0, T]; L^\infty)} + \|A_1\|_{W^{2, \infty}([0, T]; L^\infty)} \lesssim 1.$$

The following estimate can be established.

Theorem 2.6 *Let $\Phi_M^n(x)$ be the approximation obtained from the EWI-FS (2.3.11)-(2.3.12). Under the assumptions (C) and (D), there exists $h_0 > 0$ and $\tau_0 > 0$ sufficiently small and independent of ε such that, for any $0 < \varepsilon \leq 1$, when $0 < h \leq h_0$ and $0 < \tau \leq \tau_0$, we have the following error estimate*

$$\|\Phi(t_n, x) - \Phi_M^n(x)\|_{L^2} \lesssim \frac{\tau^2}{\varepsilon^4} + h^{m_0}, \quad 0 \leq n \leq \frac{T}{\tau}. \quad (2.3.18)$$

Proof of Theorem 2.6. Define the error function $\mathbf{e}^n(x)$ for $n = 0, 1, \dots$ as

$$\mathbf{e}^n(x) = \begin{pmatrix} e_1^n(x) \\ e_2^n(x) \end{pmatrix} := P_M \Phi(t_n, x) - \Phi_M^n(x) = \sum_{l=-M/2}^{M/2-1} \widehat{\mathbf{e}}_l^n e^{i\mu_l(x-a)}, \quad a \leq x \leq b. \quad (2.3.19)$$

Using the triangular inequality and standard interpolation result, we get

$$\begin{aligned} \|\Phi(t_n, x) - \Phi_M^n(x)\|_{L^2} &\leq \|\Phi(t_n, x) - P_M \Phi(t_n, x)\|_{L^2} + \|\mathbf{e}^n(x)\|_{L^2} \\ &\leq h^{m_0} + \|\mathbf{e}^n(x)\|_{L^2} \quad 0 \leq n \leq \frac{T}{\tau}, \end{aligned} \quad (2.3.20)$$

which means that we only need estimate $\|\mathbf{e}^n(x)\|_{L^2}$.

Define the local truncation error $\xi^n(x) = \sum_{l=-M/2}^{M/2-1} \widehat{\xi}_l^n e^{i\mu_l(x-a)} \in Y_M$ of the EWI-FP (2.3.12) for $n \geq 0$ as

$$\widehat{\xi}_l^n = \begin{cases} (\widehat{\Phi(\tau)})_l - e^{-i\tau\Gamma_l/\varepsilon^2} (\widehat{\Phi(0)})_l + i\varepsilon^2 \Gamma_l^{-1} \left[I_2 - e^{-\frac{i\tau}{\varepsilon^2} \Gamma_l} \right] (G(0)\widehat{\Phi(0)})_l, & n = 0, \\ (\widehat{\Phi(t_{n+1})})_l - e^{-i\tau\Gamma_l/\varepsilon^2} (\widehat{\Phi(t_n)})_l + iQ_l^{(1)}(\tau)(G(t_n)\widehat{\Phi(t_n)})_l \\ + iQ_l^{(2)}(\tau)\delta_t^-(G(t_n)\widehat{\Phi(t_n)})_l, & n \geq 1, \end{cases} \quad (2.3.21)$$

where we write $\Phi(t)$ and $G(t)$ in short for $\Phi(t, x)$ and $G(t, x)$, respectively.

Firstly, we estimate the local truncation error $\xi^n(x)$. Multiplying both sides of the Dirac equation (2.2.1) by $e^{i\mu_l(x-a)}$ and integrating over the interval (a, b) , we easily recover the equations for $(\widehat{\Phi(t)})_l$, which are exactly the same as (2.3.4) with Φ_M being replaced by $\Phi(t, x)$. Replacing Φ_M with $\Phi(t, x)$, we use the same notations $\widehat{F}_l^n(s)$ as in (2.3.5) and the time derivatives of $\widehat{F}_l^n(s)$ enjoy the same properties of time derivatives of $\Phi(t, x)$. Thus, the same representation (2.3.8) holds for $(\widehat{\Phi(t_n)})_l$ for $n \geq 1$. From the derivation of the EWI method, it is clear that the error $\xi^n(x)$ comes from the approximations for the integrals in (2.3.9) and (2.3.10), and we have

$$\widehat{\xi}_l^0 = -i \int_0^\tau e^{\frac{i(s-\tau)}{\varepsilon^2} \Gamma_l} (\widehat{F}_l^0(s) - \widehat{F}_l^0(0)) ds = -i \int_0^\tau \int_0^s e^{\frac{i(s-\tau)}{\varepsilon^2} \Gamma_l} \partial_{s_1} \widehat{F}_l^0(s_1) ds_1 ds, \quad (2.3.22)$$

and for $n \geq 1$

$$\widehat{\xi}_l^n = -i \int_0^\tau e^{\frac{i(s-\tau)}{\varepsilon^2} \Gamma_l} \left(\int_0^s \int_0^{s_1} \partial_{s_2 s_2} \widehat{F}_l^n(s_2) ds_2 ds_1 + s \int_0^1 \int_{\theta_\tau}^\tau \partial_{\theta_1 \theta_1} \widehat{F}_l^{n-1}(\theta_1) d\theta_1 d\theta \right) ds. \quad (2.3.23)$$

For $n = 0$, the above equalities imply $|\widehat{\xi}_l^0| \lesssim \int_0^\tau \int_0^s |\partial_{s_1} \widehat{F}_l^0(s_1)| ds_1 ds$ and by the Bessel inequality and assumptions (C) and (D), we find

$$\begin{aligned} \|\xi^0(x)\|_{L^2}^2 &= (b-a) \sum_{l=-M/2}^{M/2-1} |\widehat{\xi}_l^0|^2 \lesssim (b-a) \tau^2 \int_0^\tau \int_0^s \sum_{l=-M/2}^{M/2-1} |\partial_{s_1} \widehat{F}_l^0(s_1)|^2 ds_1 ds \\ &\lesssim \tau^2 \int_0^\tau \int_0^s \|\partial_{s_1}(G(s_1)\Phi(s_1))\|_{L^2}^2 ds_1 ds \lesssim \frac{\tau^4}{\varepsilon^4}. \end{aligned}$$

Similarly, for $n \geq 1$, we obtain

$$\begin{aligned} \|\xi^n(x)\|_{L^2}^2 &= (b-a) \sum_{l=-M/2}^{M/2-1} |\widehat{\xi}_l^n|^2 \\ &\lesssim \tau^3 \int_0^\tau \int_0^s \int_0^{s_1} \sum_{l=-\frac{M}{2}}^{\frac{M}{2}-1} |\partial_{s_2 s_2} \widehat{F}_l^n(s_2)|^2 ds_2 ds_1 ds \\ &\quad + \tau^3 \int_0^\tau \int_0^1 \int_{\theta_\tau}^\tau s \sum_{l=-\frac{M}{2}}^{\frac{M}{2}-1} |\partial_{\theta_1 \theta_1} \widehat{F}_l^{n-1}(\theta_1)|^2 d\theta_1 d\theta ds \\ &\lesssim \tau^6 \|\partial_{tt}(G(t)\Phi(t))\|_{L^\infty([0,T];(L^2)^2)}^2 \lesssim \frac{\tau^6}{\varepsilon^8}, \end{aligned}$$

where we have used the assumptions (C) and (D). Hence, we derive that

$$\|\xi^0(x)\|_{L^2} \lesssim \frac{\tau^2}{\varepsilon^2}, \quad \|\xi^n(x)\|_{L^2} \lesssim \frac{\tau^3}{\varepsilon^4}, \quad n \geq 1. \quad (2.3.24)$$

Now, we look at the error equations. For each fixed $l = -M/2, \dots, M/2 - 1$, subtracting (2.3.12) from (2.3.21), we obtain the equation for the error vector function as

$$\widehat{\mathbf{e}}_l^{n+1} = e^{-i\tau\Gamma_l/\varepsilon^2} \widehat{\mathbf{e}}_l^n + \widehat{R}_l^n + \widehat{\xi}_l^n, \quad 1 \leq n \leq \frac{T}{\tau} - 1, \quad (2.3.25)$$

where $R^n(x) = \sum_{l=-M/2}^{M/2-1} \widehat{R}_l^n e^{i\mu_l(x-a)} \in Y_M$ for $n \geq 1$ is given by

$$\begin{aligned} \widehat{R}_l^n &= -iQ_l^{(1)}(\tau) \left((G(\widehat{t}_n)\widehat{\Phi}(\widehat{t}_n))_l - (G(\widehat{t}_n)\widehat{\Phi}_M^n)_l \right) \\ &\quad - iQ_l^{(2)}(\tau) \left(\delta_t^-(G(\widehat{t}_n)\widehat{\Phi}(\widehat{t}_n))_l - \delta_t^-(G(\widehat{t}_n)\widehat{\Phi}_M^n)_l \right), \end{aligned} \quad (2.3.26)$$

and $\widehat{\mathbf{e}}_l^0 = \mathbf{0}$, $\widehat{\mathbf{e}}_l^1 = \widehat{\xi}_l^0$.

Using the properties of the matrices $Q_l^{(1)}(\tau)$ and $Q_l^{(2)}(\tau)$, it is easy to verify that

$$\|Q_l^{(1)}(\tau)\|_2 \leq \tau, \quad \|Q_l^{(2)}(\tau)\|_2 \leq \frac{\tau^2}{2}, \quad l = -\frac{M}{2}, \dots, \frac{M}{2} - 1, \quad (2.3.27)$$

where $\|Q\|_2$ denotes the l^2 norm of matrix Q . Combining (2.3.26), (2.3.27) and assumption (D), we get

$$\begin{aligned} \|R^n(x)\|_{L^2}^2 &= (b-a) \sum_{l=-M/2}^{M/2-1} |\widehat{R}_l^n|^2 \\ &\lesssim (b-a)\tau^2 \sum_{l=-M/2}^{M/2-1} \left(\left| (\widehat{\Phi}(t_n))_l - (\widehat{\Phi}_M^n)_l \right|^2 + \left| (\widehat{\Phi}(t_{n-1}))_l - (\widehat{\Phi}_M^{n-1})_l \right|^2 \right) \\ &\lesssim \tau^2 \sum_{k=n-1}^n \|\Phi(t_k, x) - \Phi_M^k(x)\|_{L^2}^2 \\ &\lesssim \tau^2 h^{2m_0} + \tau^2 \|\mathbf{e}^n(x)\|_{L^2}^2 + \tau^2 \|\mathbf{e}^{n-1}(x)\|_{L^2}^2. \end{aligned} \quad (2.3.28)$$

Multiplying both sides of (2.3.25) by $(\widehat{\mathbf{e}}_l^{n+1} + e^{-i\tau\Gamma_l/\varepsilon^2}\widehat{\mathbf{e}}_l^n)^*$ from left, taking the real parts and using Cauchy inequality, we obtain

$$|\widehat{\mathbf{e}}_l^{n+1}|^2 - |\widehat{\mathbf{e}}_l^n|^2 \leq \tau \left(|\widehat{\mathbf{e}}_l^{n+1}|^2 + |\widehat{\mathbf{e}}_l^n|^2 \right) + \frac{|\widehat{R}_l^n|^2}{\tau} + \frac{|\widehat{\xi}_l^n|^2}{\tau}.$$

Multiplying both sides of the above inequality by $(b-a)$ and summing together for $l = -M/2, \dots, M/2 - 1$, in view of the Bessel inequality, we obtain

$$\begin{aligned} \|\mathbf{e}^{n+1}(x)\|_{L^2}^2 - \|\mathbf{e}^n(x)\|_{L^2}^2 &\lesssim \tau (\|\mathbf{e}^{n+1}(x)\|_{L^2}^2 + \|\mathbf{e}^n(x)\|_{L^2}^2) + \frac{1}{\tau} \|R^n(x)\|_{L^2}^2 \\ &\quad + \frac{1}{\tau} \|\xi^n(x)\|_{L^2}^2, \quad n \geq 1. \end{aligned} \quad (2.3.29)$$

Summing (2.3.29) for $n = 1, \dots, m-1$, using (2.3.28) and (2.3.24), we derive

$$\|\mathbf{e}^m(x)\|_{L^2}^2 - \|\mathbf{e}^1(x)\|_{L^2}^2 \lesssim \tau \sum_{k=1}^m \|\mathbf{e}^k(x)\|_{L^2}^2 + \frac{m\tau^5}{\varepsilon^8} + m\tau h^{2m_0}, \quad m \leq \frac{T}{\tau}.$$

Since $\|\mathbf{e}^0(x)\|_{L^2} = 0$ and $\|\mathbf{e}^1(x)\|_{L^2} \lesssim \frac{\tau^2}{\varepsilon^2} \lesssim \frac{\tau^2}{\varepsilon^4}$, the discrete Gronwall's inequality will imply that for sufficiently small τ ,

$$\|\mathbf{e}^m(x)\|_{L^2}^2 \lesssim h^{2m_0} + \frac{\tau^4}{\varepsilon^8}, \quad 1 \leq m \leq \frac{T}{\tau}. \quad (2.3.30)$$

Combining (2.3.20) and (2.3.30), we draw the conclusion (2.3.18). \square

Remark 2.1 *The same error estimate in Theorem 2.6 holds for the EWI-FP (2.3.13)-(2.3.14) and the proof is quite similar to that of Theorem 2.6.*

2.3.4 Extension to 2D and 3D

The EWI-FS (2.3.11)-(2.3.12), EWI-FP (2.3.13)-(2.3.14) can be easily extended to 2D and 3D with tensor grids by modifying the matrices Γ_l in (2.3.6). For the reader's convenience, we present the modifications of Γ_l in (2.3.6) in 2D and 3D as follows.

For the Dirac equation (2.1.15) in 2D, i.e. we take $d = 2$ in (2.1.15). The problem is truncated on $\Omega = (a_1, b_1) \times (a_2, b_2)$ with mesh sizes $h_1 = (b_1 - a_1)/M_1$ and $h_2 = (b_2 - a_2)/M_2$ (M_1, M_2 two even positive integers) in the x - and y -direction, respectively. The wave function Φ is a two-component vector, and the matrix Γ_l in (2.3.6) will be replaced by

$$\Gamma_{jk} = \begin{pmatrix} 1 & \varepsilon\mu_j^{(1)} - i\varepsilon\mu_k^{(2)} \\ \varepsilon\mu_j^{(1)} + i\varepsilon\mu_k^{(2)} & -1 \end{pmatrix}, \quad \mu_j^{(1)} = \frac{2j\pi}{b_1 - a_1}, \quad \mu_k^{(2)} = \frac{2k\pi}{b_2 - a_2}, \quad (2.3.31)$$

where $-\frac{M_1}{2} \leq j \leq \frac{M_1}{2} - 1$, $-\frac{M_2}{2} \leq k \leq \frac{M_2}{2} - 1$, and the Schur decomposition $\Gamma_{jk} = Q_{jk}D_{jk}Q_{jk}^*$ is given as

$$Q_{jk} = \begin{pmatrix} \frac{1+\delta_{jk}}{\sqrt{2\delta_{jk}(1+\delta_{jk})}} & \frac{-\varepsilon\mu_j^{(1)} + i\varepsilon\mu_k^{(2)}}{\sqrt{2\delta_{jk}(1+\delta_{jk})}} \\ \frac{\varepsilon\mu_j^{(1)} + i\varepsilon\mu_k^{(2)}}{\sqrt{2\delta_{jk}(1+\delta_{jk})}} & \frac{1+\delta_{jk}}{\sqrt{2\delta_{jk}(1+\delta_{jk})}} \end{pmatrix}, \quad D_{jk} = \begin{pmatrix} \delta_{jk} & 0 \\ 0 & -\delta_{jk} \end{pmatrix}, \quad (2.3.32)$$

where $\delta_{jk} = \sqrt{1 + \varepsilon^2(\mu_j^{(1)})^2 + \varepsilon^2(\mu_k^{(2)})^2}$.

For the Dirac equation (2.1.3) in 3D, i.e. we take $d = 3$ in (2.1.3). The problem is truncated on $\Omega = (a_1, b_1) \times (a_2, b_2) \times (a_3, b_3)$ with mesh sizes $h_1 = (b_1 - a_1)/M_1$, $h_2 = (b_2 - a_2)/M_2$ and $h_3 = (b_3 - a_3)/M_3$ (M_1, M_2, M_3 three even positive integers) in x -, y - and z -direction, respectively. The wave function Ψ is a four-component

vector, and the matrix Γ_l in (2.3.6) will be replaced by Γ_{jkl} as:

$$\Gamma_{jkl} = \begin{pmatrix} 1 & 0 & \varepsilon\mu_l^{(3)} & \varepsilon\mu_j^{(1)} - i\varepsilon\mu_k^{(2)} \\ 0 & 1 & \varepsilon\mu_j^{(1)} + i\varepsilon\mu_k^{(2)} & -\varepsilon\mu_l^{(3)} \\ \varepsilon\mu_l^{(3)} & \varepsilon\mu_j^{(1)} - i\varepsilon\mu_k^{(2)} & -1 & 0 \\ \varepsilon\mu_j^{(1)} + i\varepsilon\mu_k^{(2)} & -\varepsilon\mu_l^{(3)} & 0 & -1 \end{pmatrix}, \quad (2.3.33)$$

where $-\frac{M_1}{2} \leq j \leq \frac{M_1}{2} - 1$, $-\frac{M_2}{2} \leq k \leq \frac{M_2}{2} - 1$, $-\frac{M_3}{2} \leq l \leq \frac{M_3}{2} - 1$ and

$$\mu_j^{(1)} = \frac{2j\pi}{b_1 - a_1}, \quad \mu_k^{(2)} = \frac{2k\pi}{b_2 - a_2}, \quad \mu_l^{(3)} = \frac{2l\pi}{b_3 - a_3}. \quad (2.3.34)$$

The eigenvalues of Γ_{jkl} are

$$\delta_{jkl}, \delta_{jkl}, -\delta_{jkl}, -\delta_{jkl}, \quad \text{with} \quad \delta_{jkl} = \sqrt{1 + \varepsilon^2 |\mu_j^{(1)}|^2 + \varepsilon^2 |\mu_k^{(2)}|^2 + \varepsilon^2 |\mu_l^{(3)}|^2}.$$

The corresponding eigenvectors are

$$\mathbf{v}_{jkl}^1 = \begin{pmatrix} 1 + \delta_{jkl} \\ 0 \\ \varepsilon\mu_l^{(3)} \\ \varepsilon\mu_j^{(1)} + i\varepsilon\mu_k^{(2)} \end{pmatrix}, \quad \mathbf{v}_{jkl}^2 = \begin{pmatrix} 0 \\ 1 + \delta_{jkl} \\ \varepsilon\mu_j^{(1)} - i\varepsilon\mu_k^{(2)} \\ -\varepsilon\mu_l^{(3)} \end{pmatrix},$$

$$\mathbf{v}_{jkl}^3 = \begin{pmatrix} -\varepsilon\mu_l^{(3)} \\ -\varepsilon\mu_j^{(1)} - i\varepsilon\mu_k^{(2)} \\ 1 + \delta_{jkl} \\ 0 \end{pmatrix}, \quad \mathbf{v}_{jkl}^4 = \begin{pmatrix} -\varepsilon\mu_j^{(1)} + i\varepsilon\mu_k^{(2)} \\ \varepsilon\mu_l^{(3)} \\ 0 \\ 1 + \delta_{jkl} \end{pmatrix}.$$

Then the Schur decomposition $\Gamma_{jkl} = Q_{jkl} D_{jkl} Q_{jkl}^*$ is given as

$$D_{jkl} = \text{diag}(\delta_{jkl}, \delta_{jkl}, -\delta_{jkl}, -\delta_{jkl}), \quad Q_{jkl} = \frac{1}{\sqrt{2\delta_{jkl}(1 + \delta_{jkl})}} (\mathbf{v}_{jkl}^1, \mathbf{v}_{jkl}^2, \mathbf{v}_{jkl}^3, \mathbf{v}_{jkl}^4).$$

where $\lambda(t, \mathbf{x}) = \sqrt{|A_1(t, \mathbf{x})|^2 + |A_2(t, \mathbf{x})|^2 + |A_3(t, \mathbf{x})|^2}$ and

$$\mathbf{u}^1 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{A_3(t, \mathbf{x})}{\sqrt{2}\lambda(t, \mathbf{x})} \\ \frac{A_1(t, \mathbf{x}) + iA_2(t, \mathbf{x})}{\sqrt{2}\lambda(t, \mathbf{x})} \end{pmatrix}, \quad \mathbf{u}^2 = \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ \frac{A_1(t, \mathbf{x}) - iA_2(t, \mathbf{x})}{\sqrt{2}\lambda(t, \mathbf{x})} \\ \frac{-A_3(t, \mathbf{x})}{\sqrt{2}\lambda(t, \mathbf{x})} \end{pmatrix},$$

$$\mathbf{u}^3 = \begin{pmatrix} \frac{-A_3(t, \mathbf{x})}{\sqrt{2}\lambda(t, \mathbf{x})} \\ \frac{-A_1(t, \mathbf{x}) - iA_2(t, \mathbf{x})}{\sqrt{2}\lambda(t, \mathbf{x})} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}, \quad \mathbf{u}^4 = \begin{pmatrix} \frac{-A_1(t, \mathbf{x}) + iA_2(t, \mathbf{x})}{\sqrt{2}\lambda(t, \mathbf{x})} \\ \frac{A_3(t, \mathbf{x})}{\sqrt{2}\lambda(t, \mathbf{x})} \\ 0 \\ \frac{1}{\sqrt{2}} \end{pmatrix}.$$

For the Dirac equation (2.1.3) in 2D, we simply let $\mu_l^{(3)} = 0$, in the above 3D case; and for the Dirac equation (1.2.3) in 1D, we let $\mu_k^{(2)} = \mu_l^{(3)} = 0$ in the above 3D case. Then the EWI-FP (2.3.13)-(2.3.14) can be designed accordingly for the Dirac equation (2.1.3) in 2D and 1D.

2.4 Time-splitting Pseudospectral methods

2.4.1 The TSFP method in 1D

In this section, we present a time-splitting Fourier pseudospectral (TSFP) method to solve the Dirac equation (2.1.3) (or (2.1.15)). Again, for simplicity of notations, we shall only present the numerical method and its analysis for (2.2.1) in 1D. Generalization to (2.1.3) and/or higher dimensions is straightforward and results remain valid without modifications.

From time $t = t_n$ to time $t = t_{n+1}$, the Dirac equation (2.2.1) is splitted into two steps. One solves first

$$i\partial_t \Phi(t, x) = \left[-\frac{i}{\varepsilon} \sigma_1 \partial_x + \frac{1}{\varepsilon^2} \sigma_3 \right] \Phi(t, x), \quad x \in \Omega, \quad (2.4.1)$$

with the periodic boundary condition (2.2.2) for the time step of length τ , followed by solving

$$i\partial_t \Phi(t, x) = [-A_1(t, x) \sigma_1 + V(t, x) I_2] \Phi(t, x), \quad x \in \Omega, \quad (2.4.2)$$

for the same time step. Equation (2.4.1) will be first discretized in space by the Fourier spectral method and then integrated (in phase or Fourier space) in time exactly [19]. For the ODEs (2.4.2), we can integrate *analytically* in time as

$$\Phi(t, x) = e^{i[\int_{t_n}^t A_1(s, x) ds \sigma_1 - \int_{t_n}^t V(s, x) ds I_2]} \Phi(t_n, x), \quad a \leq x \leq b, \quad t_n \leq t \leq t_{n+1}.$$

In practical computation, from time $t = t_n$ to $t = t_{n+1}$, one often combines the splitting steps via the standard Strang splitting [92] – which results in a second order time-splitting Fourier pseudospectral (TSFP) method – as

$$\begin{aligned} \Phi_j^{(1)} &= \sum_{l=-M/2}^{M/2-1} e^{-i\tau\Gamma_l/2\varepsilon^2} (\widetilde{\Phi^n})_l e^{i\mu_l(x_j-a)} = \sum_{l=-M/2}^{M/2-1} Q_l e^{-i\tau D_l/2\varepsilon^2} (Q_l)^* (\widetilde{\Phi^n})_l e^{\frac{2ijl\pi}{M}}, \\ \Phi_j^{(2)} &= e^{-i\int_{t_n}^{t_{n+1}} G(t, x_j) dt} \Phi_j^{(1)} = P e^{-i\int_{t_n}^{t_{n+1}} \Lambda(t, x_j) dt} P^* \Phi_j^{(1)}, \quad j = 0, 1, \dots, M, \quad n \geq 0, \\ \Phi_j^{n+1} &= \sum_{l=-M/2}^{M/2-1} e^{-i\tau\Gamma_l/2\varepsilon^2} (\widetilde{\Phi^{(2)}})_l e^{i\mu_l(x_j-a)} = \sum_{l=-M/2}^{M/2-1} Q_l e^{-i\tau D_l/2\varepsilon^2} (Q_l)^* (\widetilde{\Phi^{(2)}})_l e^{\frac{2ijl\pi}{M}}, \end{aligned} \tag{2.4.3}$$

where $G(t, x) = V(t, x)I_2 - A_1(t, x)\sigma_1 = P \Lambda(t, x) P^*$ and

$$P = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}, \quad \Lambda(t, x) = \begin{pmatrix} V(t, x) - A_1(t, x) & 0 \\ 0 & V(t, x) + A_1(t, x) \end{pmatrix}. \tag{2.4.4}$$

Remark 2.2 *Again, if the definite integrals in $\int_{t_n}^{t_{n+1}} \Lambda(t, x_j) dt$ cannot be evaluated analytically, we can evaluate them numerically via the Simpson's quadrature rule as*

$$\begin{aligned} \int_{t_n}^{t_{n+1}} A_1(t, x_j) dt &\approx \frac{A_1(t_n, x_j) + 4A_1(t_n + \frac{\tau}{2}, x_j) + A_1(t_{n+1}, x_j)}{6}, \\ \int_{t_n}^{t_{n+1}} V(t, x_j) dt &\approx \frac{V(t_n, x_j) + 4V(t_n + \frac{\tau}{2}, x_j) + V(t_{n+1}, x_j)}{6}. \end{aligned}$$

Lemma 2.3 *The TSFP (2.4.3) conserves the mass in the discretized level, i.e.*

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

Proof: The proof is quite standard and similar to that of Lemma 2.2. We omit it here. \square

From Lemma 2.3, we conclude that TSFP (2.4.3) is unconditionally stable. In addition, under proper assumptions of the exact solution $\Phi(t, x)$ and electromagnetic potentials, it is easy to show the following error estimate via the formal Lie calculus introduced in [66],

$$\|I_M(\Phi^n) - \Phi(t_n, x)\|_{L^2} \lesssim h^{m_0} + \frac{\tau^2}{\varepsilon^4}, \quad (2.4.5)$$

where m_0 depends on the regularity of $\Phi(t, x)$. We omit the details here for brevity.

2.4.2 Extension to 2D and 3D

The TSFP (2.4.3) can be easily extended to 2D and 3D with tensor grids by modifying the matrices Γ_l in (2.3.6) and $G(t, x)$ in (3.4.10) in the TSFP case. For the reader's convenience, we only present the modification of $G(t, x)$ in (3.4.10) in 2D and 3D as follows since modification of Γ_l is same as the EWI-FP method.

For the Dirac equation (2.1.15) in 2D, i.e. we take $d = 2$ in (2.1.15). The problem is truncated on $\Omega = (a_1, b_1) \times (a_2, b_2)$ with mesh sizes $h_1 = (b_1 - a_1)/M_1$ and $h_2 = (b_2 - a_2)/M_2$ (M_1, M_2 two even positive integers) in the x - and y -direction, respectively. The wave function Φ is a two-component vector, and the matrix $G(t, x)$ in (3.4.10) becomes $G(t, \mathbf{x})$ and the Schur decomposition $G(t, \mathbf{x}) = P(t, \mathbf{x})\Lambda(t, \mathbf{x})P^*(t, \mathbf{x})$ reads

$$\Lambda(t, \mathbf{x}) = \begin{pmatrix} V(t, \mathbf{x}) - \lambda(t, \mathbf{x}) & 0 \\ 0 & V(t, \mathbf{x}) + \lambda(t, \mathbf{x}) \end{pmatrix},$$

$$P(t, \mathbf{x}) = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{A_1(t, \mathbf{x}) - iA_2(t, \mathbf{x})}{\sqrt{2\lambda(t, \mathbf{x})}} \\ \frac{-A_1(t, \mathbf{x}) + iA_2(t, \mathbf{x})}{\sqrt{2\lambda(t, \mathbf{x})}} & \frac{1}{\sqrt{2}} \end{pmatrix}, \quad (2.4.6)$$

where $\lambda(t, \mathbf{x}) = \sqrt{|A_1(t, \mathbf{x})|^2 + |A_2(t, \mathbf{x})|^2}$.

For the Dirac equation (2.1.3) in 3D, i.e. we take $d = 3$ in (2.1.3). The problem is truncated on $\Omega = (a_1, b_1) \times (a_2, b_2) \times (a_3, b_3)$ with mesh sizes $h_1 = (b_1 - a_1)/M_1$, $h_2 = (b_2 - a_2)/M_2$ and $h_3 = (b_3 - a_3)/M_3$ (M_1, M_2, M_3 three even positive integers) in x -, y - and z -direction, respectively. The wave function Ψ is a four-component vector, and the matrix $G(t, x)$ in (3.4.10) becomes $G(t, \mathbf{x})$ and the Schur decomposition

$G(t, \mathbf{x}) = P(t, \mathbf{x})\Lambda(t, \mathbf{x})P^*(t, \mathbf{x})$ reads

$$\begin{aligned}\Lambda(t, \mathbf{x}) &= V(t, \mathbf{x})I_4 + \lambda(t, \mathbf{x}) \operatorname{diag}(-1, -1, 1, 1), \\ P(t, \mathbf{x}) &= (\mathbf{u}^1(t, \mathbf{x}), \mathbf{u}^2(t, \mathbf{x}), \mathbf{u}^3(t, \mathbf{x}), \mathbf{u}^4(t, \mathbf{x})),\end{aligned}$$

where $\lambda(t, \mathbf{x}) = \sqrt{|A_1(t, \mathbf{x})|^2 + |A_2(t, \mathbf{x})|^2 + |A_3(t, \mathbf{x})|^2}$ and

$$\begin{aligned}\mathbf{u}^1 &= \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{A_3(t, \mathbf{x})}{\sqrt{2}\lambda(t, \mathbf{x})} \\ \frac{A_1(t, \mathbf{x}) + iA_2(t, \mathbf{x})}{\sqrt{2}\lambda(t, \mathbf{x})} \end{pmatrix}, \quad \mathbf{u}^2 = \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ \frac{A_1(t, \mathbf{x}) - iA_2(t, \mathbf{x})}{\sqrt{2}\lambda(t, \mathbf{x})} \\ \frac{-A_3(t, \mathbf{x})}{\sqrt{2}\lambda(t, \mathbf{x})} \end{pmatrix}, \\ \mathbf{u}^3 &= \begin{pmatrix} \frac{-A_3(t, \mathbf{x})}{\sqrt{2}\lambda(t, \mathbf{x})} \\ \frac{-A_1(t, \mathbf{x}) - iA_2(t, \mathbf{x})}{\sqrt{2}\lambda(t, \mathbf{x})} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}, \quad \mathbf{u}^4 = \begin{pmatrix} \frac{-A_1(t, \mathbf{x}) + iA_2(t, \mathbf{x})}{\sqrt{2}\lambda(t, \mathbf{x})} \\ \frac{A_3(t, \mathbf{x})}{\sqrt{2}\lambda(t, \mathbf{x})} \\ 0 \\ \frac{1}{\sqrt{2}} \end{pmatrix}.\end{aligned}$$

For the Dirac equation (2.1.3) in 2D, we simply let $A_3(t, \mathbf{x}) \equiv 0$ in the above 3D case; and for the Dirac equation (1.2.3) in 1D, we let $A_2(t, \mathbf{x}) = A_3(t, \mathbf{x}) \equiv 0$ in the above 3D case. With the modification of Γ_l , then the TSFP (2.4.3) can be designed accordingly for the Dirac equation (2.1.3) in 2D and 1D.

2.5 Numerical results

In this section, we compare the accuracy of different numerical methods including the FDTD, EWI-FP and TSFP methods for the Dirac equation (2.1.15) in 1D in terms of the mesh size h , time step τ and the parameter $0 < \varepsilon \leq 1$. We will pay particular attention to the ε -scalabilities of different methods in the nonrelativistic limit regime, i.e. $0 < \varepsilon \ll 1$. Then we simulate the dynamics of the Dirac equation (2.1.15) in 2D with a honeycomb lattice potential by the TSFP method.

2.5.1 Comparison of spatial/temporal resolution

To test the accuracy, we choose the electromagnetic potentials in the Dirac equation (2.1.15) with $d = 1$ as

$$A_1(t, x) = \frac{(x+1)^2}{1+x^2}, \quad V(t, x) = \frac{1-x}{1+x^2}, \quad x \in \mathbb{R}, \quad t \geq 0,$$

and the initial value as

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

The problem is solved numerically on an interval $\Omega = (-16, 16)$, i.e. $a = -16$ and $b = 16$, with periodic boundary conditions on $\partial\Omega$. The ‘exact’ solution $\Phi(t, x) = (\phi_1(t, x), \phi_2(t, x))^T$ is obtained numerically by using the TSFP method with a very fine mesh size and a small time step, e.g. $h_e = 1/16$ and $\tau_e = 10^{-7}$ to compare with the numerical solutions obtained by EWI-FP and TSFP, and $h_e = 1/4096$ to compare with the numerical solutions obtained by FDTD methods. Denote $\Phi_{h,\tau}^n$ as the numerical solution obtained by a numerical method with mesh size h and time step τ . In order to quantify the convergence, we introduce

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

Tab. 2.1 lists spatial errors $e_{h,\tau_e}(t = 2)$ with different h (upper part) and temporal errors $e_{h_e,\tau}(t = 2)$ with different τ (lower part) for the LFFD method (2.2.6). Tabs. 2.2-2.6 show similar results for the SIFD1 method (2.2.7), SIFD2 method (2.2.8), CNFD method (2.2.9), EWI-FP method (2.3.13)-(2.3.14) and TSFP method (2.4.3), respectively. For the LFFD and SIFD1 methods, due to the stability condition and accuracy requirement, we take

$$\delta_j(\varepsilon) = \begin{cases} \varepsilon^2 & \varepsilon_0/2^j \leq \varepsilon \leq 1, \\ \varepsilon_0^2/4^j & 0 < \varepsilon < \varepsilon_0/2^j, \end{cases} \quad j = 0, 1, \dots$$

in Tables 2.1 and 2.2. For comparison, Tab. 2.7 depicts temporal errors of different numerical methods under different ε -scalability.

Table 2.1: Spatial and temporal error analysis of the LFFD method for the Dirac equation in 1D.

Spatial Errors	$h_0 = 1/8$	$h_0/2$	$h_0/2^2$	$h_0/2^3$	$h_0/2^4$
$\varepsilon_0 = 1$	1.06E-1	2.65E-2	6.58E-3	1.64E-3	4.10E-4
$\varepsilon_0/2$	9.06E-2	2.26E-2	5.64E-3	1.41E-3	3.51E-4
$\varepsilon_0/2^2$	8.03E-2	2.02E-2	5.04E-3	1.25E-3	3.05E-4
$\varepsilon_0/2^3$	9.89E-2	2.47E-2	6.17E-3	1.54E-3	3.85E-4
$\varepsilon_0/2^4$	9.87E-2	2.48E-2	6.18E-3	1.54E-3	3.83E-4
Temporal Errors	$\tau_0 = 0.1$	$\tau_0/8$	$\tau_0/8^2$	$\tau_0/8^3$	$\tau_0/8^4$
$\varepsilon_0 = 1$	<u>1.38E-1</u>	1.99E-3	3.11E-5	4.86E-7	7.59E-9
$\varepsilon_0/2$	unstable	<u>1.14E-2</u>	1.77E-4	2.77E-6	4.32E-8
$\varepsilon_0/2^2$	unstable	4.59E-1	<u>7.01E-3</u>	1.05E-4	1.64E-6
$\varepsilon_0/2^3$	unstable	unstable	4.14E-1	<u>6.42E-3</u>	1.00E-4
$\varepsilon_0/2^4$	unstable	unstable	unstable	4.04E-1	<u>6.00E-3</u>

From Tabs. 2.1-2.7, and additional numerical results not shown here for brevity, we can draw the following conclusions for the Dirac equation by using different numerical methods:

(i). For the discretization error in space, for any fixed $\varepsilon = \varepsilon_0 > 0$, the FDTD methods are second-order accurate, and resp., the EWI-FP and TSFP methods are spectrally accurate (cf. each row in the upper parts of Tabs. 2.1-2.6). For $0 < \varepsilon \leq 1$, the errors are independent of ε for the EWI-FP and TSFP methods (cf. each column in the upper parts of Tabs. 2.5-2.6), and resp., are almost independent of ε for the FDTD methods (cf. each column in the upper parts of Tabs. 2.1-2.4). In general, for any fixed $0 < \varepsilon \leq 1$ and $h > 0$, the EWI-FP and TSFP methods perform much better than the FDTD methods in spatial discretization.

(ii). For the discretization error in time, in the $O(1)$ speed-of-light regime, i.e.

Table 2.2: Spatial and temporal error analysis of the SIFD1 method for the Dirac equation in 1D.

Spatial Errors	$h_0 = 1/8$	$h_0/2$	$h_0/2^2$	$h_0/2^3$	$h_0/2^4$
$\varepsilon_0 = 1$	1.06E-1	2.65E-2	6.58E-3	1.64E-3	4.10E-4
$\varepsilon_0/2$	9.06E-2	2.26E-2	5.64E-3	1.41E-3	3.51E-4
$\varepsilon_0/2^2$	8.03E-2	2.02E-2	5.04E-3	1.25E-3	3.05E-4
$\varepsilon_0/2^3$	9.89E-2	2.47E-2	6.17E-3	1.54E-3	3.85E-4
$\varepsilon_0/2^4$	9.87E-2	2.48E-2	6.18E-3	1.54E-3	3.83E-4
Temporal Errors	$\tau_0 = 0.1$	$\tau_0/8$	$\tau_0/8^2$	$\tau_0/8^3$	$\tau_0/8^4$
$\varepsilon_0 = 1$	<u>1.44E-1</u>	2.09E-3	3.27E-5	5.11E-7	7.98E-9
$\varepsilon_0/2$	unstable	<u>2.99E-2</u>	4.67E-4	7.30E-6	1.14E-7
$\varepsilon_0/2^2$	unstable	8.18E-1	<u>1.54E-2</u>	2.41E-4	3.77E-6
$\varepsilon_0/2^3$	unstable	unstable	7.99E-1	<u>1.31E-2</u>	2.05E-4
$\varepsilon_0/2^4$	unstable	unstable	4.19E-1	7.97E-1	<u>1.26E-2</u>

$\varepsilon = O(1)$, all the numerical methods including FDTD, EWI-FP and TSFP are second-order accurate (cf. the first row in the lower parts of Tabs. 2.1-2.6). In general, the EWI-FP and TSFP methods perform much better than the FDTD methods in temporal discretizations for a fixed time step. In the non-relativistic limit regime, i.e. $0 < \varepsilon \ll 1$, for the FDTD methods, the ‘correct’ ε -scalability is $\tau = O(\varepsilon^3)$ which verifies our theoretical results; for the EWI-FP and TSFP methods, the ‘correct’ ε -scalability is $\tau = O(\varepsilon^2)$ which again confirms our theoretical results. In fact, for $0 < \varepsilon \leq 1$, one can observe clearly second-order convergence in time for the FDTD methods only when $\tau \lesssim \varepsilon^3$ (cf. upper triangles in the lower parts of Tabs. 2.1-2.4), and resp., for the EWI-FP and TSFP methods when $\tau \lesssim \varepsilon^2$ (cf. upper triangles in the lower parts of Tabs. 2.5-2.6). In general, for any fixed $0 < \varepsilon \leq 1$ and $\tau > 0$, the TSFP method performs the best, and the EWI-FP method performs much better than the FDTD methods in temporal discretization (cf. Tab. 2.7).

Table 2.3: Spatial and temporal error analysis of the SIFD2 method for the Dirac equation in 1D.

Spatial Errors	$h_0 = 1/8$	$h_0/2$	$h_0/2^2$	$h_0/2^3$	$h_0/2^4$
$\varepsilon_0 = 1$	1.06E-1	2.65E-2	6.58E-3	1.64E-3	4.10E-4
$\varepsilon_0/2$	9.06E-2	2.26E-2	5.64E-3	1.41E-3	3.51E-4
$\varepsilon_0/2^2$	8.03E-2	2.02E-2	5.04E-3	1.25E-3	3.05E-4
$\varepsilon_0/2^3$	9.89E-2	2.47E-2	6.17E-3	1.54E-3	3.85E-4
$\varepsilon_0/2^4$	9.87E-2	2.48E-2	6.18E-3	1.54E-3	3.83E-4
Temporal Errors	$\tau_0 = 0.1$	$\tau_0/8$	$\tau_0/8^2$	$\tau_0/8^3$	$\tau_0/8^4$
$\varepsilon_0 = 1$	<u>1.72E-1</u>	2.59E-3	4.05E-5	6.33E-7	9.89E-9
$\varepsilon_0/2$	1.69	<u>3.57E-2</u>	5.58E-4	8.72E-6	1.36E-7
$\varepsilon_0/2^2$	2.59	8.66E-1	<u>1.63E-2</u>	2.55E-4	3.98E-6
$\varepsilon_0/2^3$	2.67	2.89	8.43E-1	<u>1.37E-2</u>	2.14E-4
$\varepsilon_0/2^4$	3.07	3.56	5.19E-1	8.37E-1	<u>1.28E-2</u>

(iii). From Tab. 2.6, our numerical results suggest the following error bound for the TSFP method when $\tau \lesssim \varepsilon^2$,

$$\|I_M(\Phi^n) - \Phi(t_n, \cdot)\|_{L^2} \lesssim h^{m_0} + \frac{\tau^2}{\varepsilon^2}, \quad (2.5.1)$$

which is much better than (2.4.5) for the TSFP method in the nonrelativistic limit regime. Rigorous mathematical justification for (2.5.1) is on-going.

From Tabs. 2.1-2.4, in the numerical example, we could not observe numerically the ε -dependence in the spatial discretization error for the FDTD methods, i.e. $\frac{1}{\varepsilon}$ in front of h^2 , which was proven in Theorems 2.1-2.4. In order to investigate the spatial ε -resolution of the FDTD methods, we consider the Dirac equation (2.2.1) on $\Omega = (-1, 1)$ with no electromagnetic potential – the free Dirac equation, i.e.

$$A_1(t, x) \equiv 0, \quad V(t, x) \equiv 0, \quad x \in (-1, 1), \quad t \geq 0.$$

Table 2.4: Spatial and temporal error analysis of the CNFD method for the Dirac equation in 1D.

Spatial Errors	$h_0 = 1/8$	$h_0/2$	$h_0/2^2$	$h_0/2^3$	$h_0/2^4$
$\varepsilon_0 = 1$	1.06E-1	2.65E-2	6.58E-3	1.64E-3	4.10E-4
$\varepsilon_0/2$	9.06E-2	2.26E-2	5.64E-3	1.41E-3	3.51E-4
$\varepsilon_0/2^2$	8.03E-2	2.02E-2	5.04E-3	1.25E-3	3.05E-4
$\varepsilon_0/2^3$	9.89E-2	2.47E-2	6.17E-3	1.54E-3	3.85E-4
$\varepsilon_0/2^4$	9.87E-2	2.48E-2	6.18E-3	1.54E-3	3.83E-4
Temporal Errors	$\tau_0 = 0.1$	$\tau_0/8$	$\tau_0/8^2$	$\tau_0/8^3$	$\tau_0/8^4$
$\varepsilon_0 = 1$	<u>5.48E-2</u>	8.56E-4	1.34E-5	2.09E-7	3.27E-9
$\varepsilon_0/2$	3.90E-1	<u>6.63E-3</u>	1.77E-4	2.77E-6	4.32E-8
$\varepsilon_0/2^2$	1.79	2.27E-1	<u>3.55E-3</u>	1.56E-5	2.44E-7
$\varepsilon_0/2^3$	3.10	4.69E-1	2.06E-1	<u>3.22E-3</u>	5.03E-4
$\varepsilon_0/2^4$	2.34	1.83	8.05 E-1	2.04E-1	<u>3.19E-3</u>

The initial data in (2.2.2) is taken as

$$\phi_1(0, x) = e^{9\pi i(x+1)}, \quad \phi_2(0, x) = e^{9\pi i(x+1)}, \quad -1 \leq x \leq 1.$$

Tab. 2.8 shows the spatial errors $e_{h,\tau_e}(t = 2)$ of the CNFD method with different h . The results for the LFFD, SIFD1 and SIFD2 methods are similar and they are omitted here for brevity. From Tab. 2.8, we can conclude that the error bounds in the Theorems 2.1-2.4 are sharp.

Based on the above comparison, in view of both temporal and spatial accuracies we conclude that the EWI-FP and TSFP methods perform much better than the FDTD methods for the discretization of the Dirac equation, especially in the non-relativistic limit regime. For the reader's convenience, we summarize the properties of different numerical methods in Tab. 2.9.

Table 2.5: Spatial and temporal error analysis of the EWI-FP method for the Dirac equation in 1D.

Spatial Errors	$h_0=2$	$h_0/2$	$h_0/2^2$	$h_0/2^3$	$h_0/2^4$
$\varepsilon_0 = 1$	1.10	2.43E-1	2.99E-3	2.79E-6	1.00E-8
$\varepsilon_0/2$	1.06	1.46E-1	1.34E-3	9.61E-7	5.90E-9
$\varepsilon_0/2^2$	1.11	1.43E-1	9.40E-4	5.10E-7	7.02E-9
$\varepsilon_0/2^3$	1.15	1.44E-1	7.89E-4	3.62E-7	6.86E-9
$\varepsilon_0/2^4$	1.18	1.45E-1	7.63E-4	2.91E-7	8.46E-9
Temporal Errors	$\tau_0=0.1$	$\tau_0/4$	$\tau_0/4^2$	$\tau_0/4^3$	$\tau_0/4^4$
$\varepsilon_0 = 1$	<u>1.40E-1</u>	8.51E-3	5.33E-4	3.34E-5	2.09E-6
$\varepsilon_0/2$	4.11E-1	<u>2.37E-2</u>	1.49E-3	9.29E-5	5.81E-6
$\varepsilon_0/2^2$	6.03	1.88E-1	<u>1.18E-2</u>	7.38E-4	4.62E-5
$\varepsilon_0/2^3$	2.21	3.98	1.60E-1	<u>1.01E-2</u>	6.31E-4
$\varepsilon_0/2^4$	2.16	2.09	3.58	1.53E-1	<u>9.69E-3</u>

2.5.2 Dynamics with the honeycomb potential

Here we study numerically the dynamics of the Dirac equation (2.1.15) in 2D with a honeycomb lattice potential, i.e. we take $d = 2$ and

$$A_1(t, \mathbf{x}) = A_2(t, \mathbf{x}) \equiv 0, \quad V(t, \mathbf{x}) = \cos\left(\frac{4\pi}{\sqrt{3}}\mathbf{e}_1 \cdot \mathbf{x}\right) + \cos\left(\frac{4\pi}{\sqrt{3}}\mathbf{e}_2 \cdot \mathbf{x}\right) + \cos\left(\frac{4\pi}{\sqrt{3}}\mathbf{e}_3 \cdot \mathbf{x}\right),$$

with

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

The initial data in (2.1.16) is taken as

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

The problem is solved numerically on $\Omega = [-10, 10]^2$ by the TSFP method with mesh size $h = 1/16$ and time step $\tau = 0.01$. Figs. 2.1 and 2.2 depict the densities $\rho_j(t, \mathbf{x}) = |\phi_j(t, \mathbf{x})|^2$ ($j = 1, 2$) for $\varepsilon = 1$ and $\varepsilon = 0.2$, respectively.

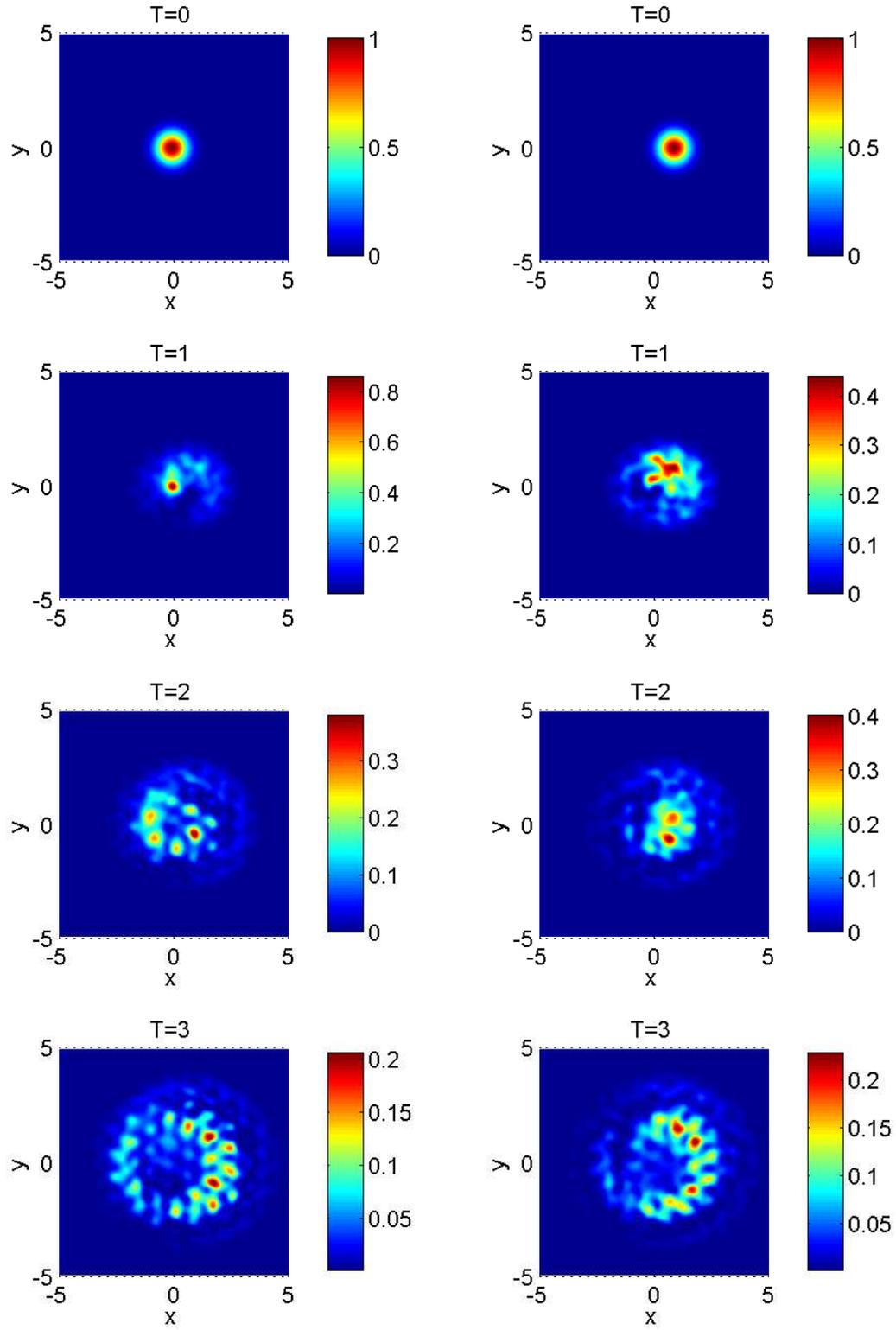


Figure 2.1: Dynamics of the densities $\rho_1(t, \mathbf{x}) = |\phi_1(t, \mathbf{x})|^2$ (left) and $\rho_2(t, \mathbf{x}) = |\phi_2(t, \mathbf{x})|^2$ (right) of the Dirac equation in 2D with a honeycomb lattice potential when $\varepsilon = 1$.

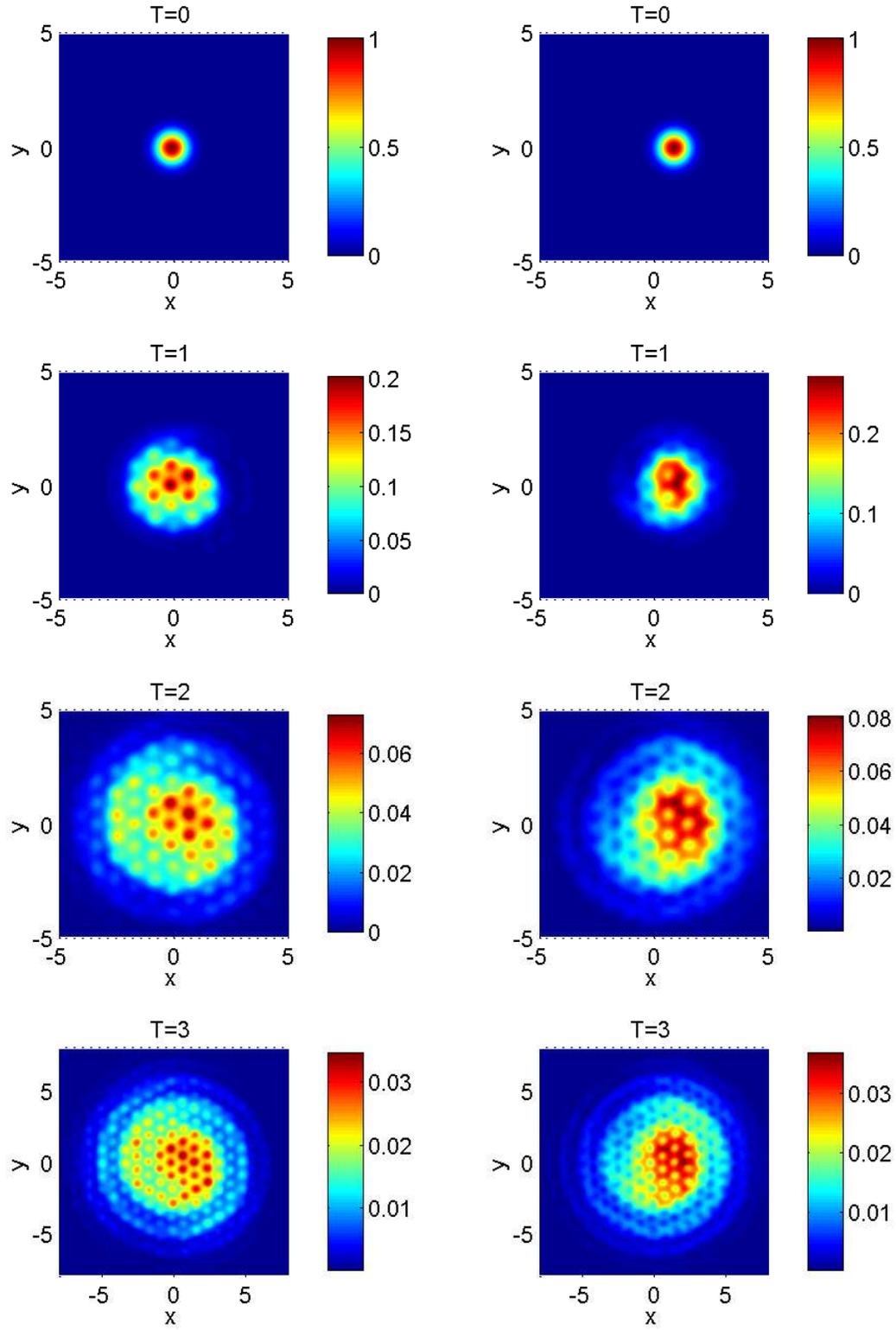


Figure 2.2: Dynamics of the densities $\rho_1(t, \mathbf{x}) = |\phi_1(t, \mathbf{x})|^2$ (left) and $\rho_2(t, \mathbf{x}) = |\phi_2(t, \mathbf{x})|^2$ (right) of the Dirac equation in 2D with a honeycomb potential when $\varepsilon = 0.2$.

Table 2.6: Spatial and temporal error analysis of the TSFP method for the Dirac equation in 1D.

Spatial Errors	$h_0 = 2$	$h_0/2$	$h_0/2^2$	$h_0/2^3$	$h_0/2^4$	
$\varepsilon_0 = 1$	1.10	2.43E-1	2.99E-3	2.79E-6	9.45E-9	
$\varepsilon_0/2$	1.06	1.46E-1	1.34E-3	9.61E-7	5.57E-9	
$\varepsilon_0/2^2$	1.11	1.43E-1	9.40E-4	5.10E-7	6.50E-9	
$\varepsilon_0/2^3$	1.15	1.44E-1	7.89E-4	3.62E-7	6.84E-9	
$\varepsilon_0/2^4$	1.18	1.45E-1	7.62E-4	2.88E-7	7.49E-9	
$\varepsilon_0/2^5$	1.19	1.46E-1	7.53E-4	2.59E-7	7.96E-9	
$\varepsilon_0/2^6$	1.20	1.47E-1	7.49E-4	2.63E-7	6.90E-9	
Temporal Errors	$\tau_0=0.4$	$\tau_0/4$	$\tau_0/4^2$	$\tau_0/4^3$	$\tau_0/4^4$	$\tau_0/4^5$
$\varepsilon_0 = 1$	<u>2.17E-1</u>	1.32E-2	8.22E-4	5.13E-5	3.21E-6	2.01E-7
$\varepsilon_0/2$	1.32	<u>6.60E-2</u>	4.07E-3	2.54E-4	1.59E-5	9.92E-7
$\varepsilon_0/2^2$	2.50	3.33E-1	<u>1.68E-2</u>	1.04E-3	6.49E-5	4.06E-6
$\varepsilon_0/2^3$	1.79	1.97	8.15E-2	<u>4.15E-3</u>	2.57E-4	1.60E-5
$\varepsilon_0/2^4$	1.35	8.27E-1	8.85E-1	2.01E-2	<u>1.03E-3</u>	6.35E-5
$\varepsilon_0/2^5$	8.73E-1	2.25E-1	2.33E-1	2.49E-1	4.98E-3	<u>2.55E-4</u>

From Fig. 2.1 and Fig. 2.2, we find that the dynamics of the Dirac equation depends significantly on ε . In addition, the TSFP method can capture the dynamics very accurately and efficiently.

Table 2.7: Comparison of temporal errors of different numerical methods for the Dirac equation under proper ε -scalability.

$\tau = O(\varepsilon^3)$	$\varepsilon_0 = 1$	$\varepsilon_0/2$	$\varepsilon_0/2^2$	$\varepsilon_0/2^3$	$\varepsilon_0/2^4$
$\tau = O(h)$	$h_0 = 1/8$	$h_0/2$	$h_0/2^2$	$h_0/2^3$	$h_0/2^4$
	$\tau_0 = 0.1$	$\tau_0/8$	$\tau_0/8^2$	$\tau_0/8^3$	$\tau_0/8^4$
LFFD	1.38E-1	1.14E-2	7.01E-3	6.42E-3	6.00E-3
SIFD1	1.44E-1	2.99E-2	1.54E-2	1.31E-2	1.27E-2
$\tau = O(\varepsilon^3)$	$\varepsilon_0 = 1$	$\varepsilon_0/2$	$\varepsilon_0/2^2$	$\varepsilon_0/2^3$	$\varepsilon_0/2^4$
	$\tau_0 = 0.1$	$\tau_0/8$	$\tau_0/8^2$	$\tau_0/8^3$	$\tau_0/8^4$
SIFD2	1.31E-1	2.41E-2	1.45E-2	2.30E-2	1.26E-2
CNFD	5.48E-2	6.63E-3	3.55E-3	3.22E-3	3.19E-3
$\tau = O(\varepsilon^2)$	$\varepsilon_0 = 1$	$\varepsilon_0/2$	$\varepsilon_0/2^2$	$\varepsilon_0/2^3$	$\varepsilon_0/2^4$
	$\tau_0 = 0.1$	$\tau_0/4$	$\tau_0/4^2$	$\tau_0/4^3$	$\tau_0/4^4$
EWI-FP	1.40E-1	2.37E-2	1.18E-2	1.01E-3	9.69E-3
TSFP	1.32E-2	4.07E-3	1.04E-3	2.57E-4	6.35E-5

Table 2.8: Spatial error analysis of the CNFD method for the free Dirac equation with different h .

ε	$\varepsilon_0 = 1$	$\varepsilon_0/2$	$\varepsilon_0/2^2$	$\varepsilon_0/2^3$	$\varepsilon_0/2^4$
$h_0 = 1/256$	1.61E-1	3.21E-1	6.35E-1	1.21	2.07
$h_0/2$	4.03E-2	8.05E-2	1.59E-1	3.07E-1	5.43E-1
$h_0/2^2$	1.01E-2	2.01E-2	3.99E-2	7.69E-2	1.36E-1
$h_0/2^3$	2.52E-3	5.03E-2	9.97E-3	1.92E-2	3.41E-2
$h_0/2^4$	6.30E-4	1.26E-2	2.47E-3	4.95E-3	8.64E-3

Table 2.9: Comparison of properties of different numerical methods for solving the Dirac equation with M being the number of grid points in space.

Method	LFFD	SIFD1	SIFD2
Time symmetric	Yes	Yes	Yes
Mass conservation	No	No	No
Energy conservation	No	No	No
Dispersion Relation	No	No	No
Unconditionally stable	No	No	No
Explicit scheme	Yes	No	No
Temporal accuracy	2nd	2nd	2nd
Spatial accuracy	2nd	2nd	2nd
Memory cost	$O(M)$	$O(M)$	$O(M)$
Computational cost	$O(M)$	$O(M)$	$O(M \ln M)$
Resolution	$h = O(\sqrt{\varepsilon})$	$h = O(\sqrt{\varepsilon})$	$h = O(\sqrt{\varepsilon})$
when $0 < \varepsilon \ll 1$	$\tau = O(\varepsilon^3)$	$\tau = O(\varepsilon^3)$	$\tau = O(\varepsilon^3)$
Method	CNFD	EWI-FP	TSFP
Time symmetric	Yes	No	Yes
Mass conservation	Yes	No	Yes
Energy conservation	Yes	No	No
Dispersion Relation	No	No	Yes
Unconditionally stable	Yes	No	Yes
Explicit scheme	No	Yes	Yes
Temporal accuracy	2nd	2nd	2nd
Spatial accuracy	2nd	Spectral	Spectral
Memory cost	$O(M)$	$O(M)$	$O(M)$
Computational cost	$\gg O(M)$	$O(M \ln M)$	$O(M \ln M)$
Resolution	$h = O(\sqrt{\varepsilon})$	$h = O(1)$	$h = O(1)$
when $0 < \varepsilon \ll 1$	$\tau = O(\varepsilon^3)$	$\tau = O(\varepsilon^2)$	$\tau = O(\varepsilon^2)$

Extension to the nonlinear Dirac equation

This chapter investigates the performance of several numerical methods for solving the nonlinear Dirac equation in the nonrelativistic limit regime, i.e. $0 < \varepsilon \ll 1$. Conventional finite difference time domain (FDTD) discretization, Gautschi type exponential wave integrator (EWI) and time splitting with Fourier pseudospectral discretization in spatial domain are illustrated. Rigorous error bounds are carried out with particular attention on how their optimal error bounds are dependent explicitly on ε . Numerical results are also reported in the last section to support these error bounds.

3.1 Basic properties

In quantum field theory, the nonlinear Dirac equation (NLD) is a self-interacting Dirac fermions model which is widely considered in quantum mechanics as a toy model of self-interaction electrons. Different self-interactions give rise to different NLD models. Based on the scalar self-interaction, we can obtain the widely used Soler model which is a quantum field theory model of Dirac fermion interacting in $(3 + 1)$ dimensions.

$$i\hbar\partial_t\Psi = \left[-i\hbar\sum_{j=1}^3\alpha_j\partial_j+mc^2\beta\right]\Psi+e\left[V(t,\mathbf{x})I_4-\sum_{j=1}^3A_j(t,\mathbf{x})\alpha_j\right]\Psi+\mathbf{F}(\Psi)\Psi, \quad \mathbf{x}\in\mathbb{R}^3, \quad (3.1.1)$$

where $\Psi(t, x) = (\psi_1(t, x), \psi_2(t, x), \psi_3(t, x), \psi_4(t, x))^T \in \mathbb{C}^4$ is a complex four-component spinor, $\mathbf{x} = (x_1, x_2, x_3) \in \mathbb{R}^3$, $\partial_j = \frac{\partial}{\partial x_j}$, c denote the speed of light, $m > 0$ is the mass of the electron, and \hbar denotes the Planck constant. $\alpha_1, \alpha_2, \alpha_3$ and β are 4×4 complex matrices as the same as in Chapter 2. For the simplicity of notations, here we take $\mathbf{F}(\Psi) = g_1(\Psi^*\beta\Psi)\beta + g_2|\Psi|^2I_4$ with $g_1, g_2 \in \mathbb{R}$ two constants and $\Psi^* = \bar{\Psi}^T$, while \bar{f} denotes the complex conjugate of f , which is motivated from the so-called Soler model, e.g. $g_2 = 0$ and $g_1 \neq 0$, in quantum field theory [41, 43, 91, 100] and BECs with a chiral confinement and/or spin-orbit coupling, e.g. $g_1 = 0$ and $g_2 \neq 0$ [24, 48, 49]. We remark here that our numerical methods and their error estimates can be easily extended to the NLDE with other nonlinearities [81, 100]. Following the similar procedure, one can have the dimensionless NLDE

$$i\partial_t\Psi = \left[-\frac{i}{\varepsilon}\sum_{j=1}^d\alpha_j\partial_j + \frac{1}{\varepsilon^2}\beta\right]\Psi + \left[V(t,\mathbf{x})I_4 - \sum_{j=1}^dA_j(t,\mathbf{x})\alpha_j\right]\Psi + \mathbf{F}(\Psi)\Psi, \quad \mathbf{x}\in\mathbb{R}^d, \quad (3.1.2)$$

and

$$\mathbf{F}(\Psi) = \lambda_1(\Psi^*\beta\Psi)\beta + \lambda_2|\Psi|^2I_4, \quad \Psi \in \mathbb{C}^4. \quad (3.1.3)$$

with $\lambda_1 = \frac{g_1}{mv^2x_3^3} \in \mathbb{R}$ and $\lambda_2 = \frac{g_2}{mv^2x_3^3} \in \mathbb{R}$ two dimensionless constants for the interaction strength.

For the dynamics, the initial condition is given as

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

The NLDE (3.1.2) is dispersive and time symmetric. Introducing the position density ρ_j for the j -component ($j = 1, 2, 3, 4$) and the total density ρ as well as the

current density $\mathbf{J}(t, \mathbf{x}) = (J_1(t, \mathbf{x}), J_2(t, \mathbf{x}), J_3(t, \mathbf{x}))^T$

$$\begin{aligned} \rho(t, \mathbf{x}) &= \sum_{j=1}^4 \rho_j(t, \mathbf{x}) = \Psi^* \Psi, \quad \rho_j(t, \mathbf{x}) = |\psi_j(t, \mathbf{x})|^2, \quad 1 \leq j \leq 4; \\ J_l(t, \mathbf{x}) &= \frac{1}{\varepsilon} \Psi^* \alpha_l \Psi, \quad l = 1, 2, 3, \end{aligned} \quad (3.1.4)$$

then the following conservation law can be obtained from the NLDE (3.1.2)

$$\partial_t \rho(t, \mathbf{x}) + \nabla \cdot \mathbf{J}(t, \mathbf{x}) = 0, \quad \mathbf{x} \in \mathbb{R}^d, \quad t \geq 0. \quad (3.1.5)$$

Thus the NLDE (3.1.2) conserves the total mass as

$$\|\Psi(t, \cdot)\|^2 := \int_{\mathbb{R}^d} |\Psi(t, \mathbf{x})|^2 d\mathbf{x} = \int_{\mathbb{R}^d} \sum_{j=1}^4 |\psi_j(t, \mathbf{x})|^2 d\mathbf{x} \equiv \|\Psi(0, \cdot)\|^2 = \|\Psi_0\|^2, \quad t \geq 0. \quad (3.1.6)$$

The NLDE (3.1.2) conserves the total mass as

$$\|\Psi(t, \cdot)\|^2 = \int_{\mathbb{R}^3} |\Psi(t, \mathbf{x})|^2 d\mathbf{x} = \int_{\mathbb{R}^3} \sum_{j=1}^4 |\psi_j(t, \mathbf{x})|^2 d\mathbf{x} \equiv \|\Psi(0, \cdot)\|^2 = \|\Psi(0)\|^2, \quad t \geq 0. \quad (3.1.7)$$

If the electric potential V is perturbed by a constant, e.g. $V(t, \mathbf{x}) \rightarrow V(t, \mathbf{x}) + V^0$ with V^0 being a real constant, then the solution $\Psi(t, \mathbf{x}) \rightarrow e^{-iV^0 t} \Psi(t, \mathbf{x})$ which implies the density of each component $\rho_j(t, \mathbf{x}) = |\psi_j(t, \mathbf{x})|^2$ ($j = 1, 2, 3, 4$) and the total density $\rho(t, \mathbf{x}) = \sum_{j=1}^4 \rho_j(t, \mathbf{x})$ unchanged. Moreover, if the electromagnetic potentials are independent of time, i.e. $V(t, \mathbf{x}) = V(\mathbf{x})$, $A_j(t, \mathbf{x}) = A_j(\mathbf{x})$, then the NLDE (3.1.2) conserves the total energy as

$$\begin{aligned} E(t) &:= \int_{\mathbb{R}^d} \left[-\frac{i}{\varepsilon} \sum_{j=1}^d \Psi^* \alpha_j \partial_j \Psi + \frac{1}{\varepsilon^2} \Psi^* \beta \Psi + V(\mathbf{x}) |\Psi|^2 + G(\Psi) - \sum_{j=1}^d A_j(\mathbf{x}) \Psi^* \alpha_j \Psi \right] d\mathbf{x} \\ &\equiv E(0), \quad t \geq 0, \end{aligned} \quad (3.1.8)$$

where

$$G(\Psi) = \frac{\lambda_1}{2} (\Psi^* \beta \Psi)^2 + \frac{\lambda_2}{2} |\Psi|^4, \quad \Psi \in \mathbb{C}^4. \quad (3.1.9)$$

Furthermore, if the external electromagnetic potentials are constants, i.e. $V(t, \mathbf{x}) \equiv V^0$ and $A_j(t, \mathbf{x}) \equiv A_j^0$ for $j = 1, 2, 3$, the NLDE (3.1.2) 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 $\mathbf{k} = (k_1, \dots, k_d)^T \in \mathbb{R}^d$ satisfy the following *dispersion relation*

$$\omega \mathbf{B} = \left[\sum_{j=1}^d \left(\frac{k_j}{\varepsilon} - A_j^0 \right) \alpha_j + \frac{1}{\varepsilon^2} \beta + V^0 I_4 + \lambda_1 (\mathbf{B}^* \beta \mathbf{B}) \beta + \lambda_2 |\mathbf{B}|^2 I_4 \right] \mathbf{B}. \quad (3.1.10)$$

Again, similarly to the Dirac equation [15], in several applications in one dimension (1D) and two dimensions (2D), the NLDE (3.1.2) can be simplified to the following NLDE in d -dimensions ($d = 1, 2$) with $\Phi := \Phi(t, \mathbf{x}) = (\phi_1(t, \mathbf{x}), \phi_2(t, \mathbf{x}))^T \in \mathbb{C}^2$ [41, 43, 91]

$$i\partial_t \Phi = \left[-\frac{i}{\varepsilon} \sum_{j=1}^d \sigma_j \partial_j + \frac{1}{\varepsilon^2} \sigma_3 \right] \Phi + \left[V(t, \mathbf{x}) I_2 - \sum_{j=1}^d A_j(t, \mathbf{x}) \sigma_j \right] \Phi + \mathbf{F}(\Phi) \Phi, \quad \mathbf{x} \in \mathbb{R}^d, \quad (3.1.11)$$

where

$$\mathbf{F}(\Phi) = \lambda_1 (\Phi^* \sigma_3 \Phi) \sigma_3 + \lambda_2 |\Phi|^2 I_2, \quad \Phi \in \mathbb{C}^2, \quad (3.1.12)$$

with $\lambda_1 \in \mathbb{R}$ and $\lambda_2 \in \mathbb{R}$ two dimensionless constants for the interaction strength.

Again, the initial condition for dynamics is given as

$$\Phi(t = 0, \mathbf{x}) = \Phi_0(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^d. \quad (3.1.13)$$

The NLDE (3.1.11) is dispersive and time symmetric. By introducing the position density ρ_j for the j -th component ($j = 1, 2$) and the total density ρ as well as the current density $\mathbf{J}(t, \mathbf{x}) = (J_1(t, \mathbf{x}), J_2(t, \mathbf{x}))^T$

$$\rho(t, \mathbf{x}) = \sum_{j=1}^2 \rho_j(t, \mathbf{x}) = \Phi^* \Phi, \quad \rho_j(t, \mathbf{x}) = |\phi_j(t, \mathbf{x})|^2, \quad J_j(t, \mathbf{x}) = \frac{1}{\varepsilon} \Phi^* \sigma_j \Phi, \quad j = 1, 2, \quad (3.1.14)$$

the conservation law (3.1.5) is also satisfied [23]. In addition, the Dirac equation (3.1.11) conserves the total mass as

$$\|\Phi(t, \cdot)\|^2 := \int_{\mathbb{R}^d} |\Phi(t, \mathbf{x})|^2 d\mathbf{x} = \int_{\mathbb{R}^d} \sum_{j=1}^2 |\phi_j(t, \mathbf{x})|^2 d\mathbf{x} \equiv \|\Phi(0, \cdot)\|^2 = \|\Phi_0\|^2, \quad t \geq 0. \quad (3.1.15)$$

Again, if the electric potential V is perturbed by a constant, e.g. $V(t, \mathbf{x}) \rightarrow V(t, \mathbf{x}) + V^0$ with V^0 being a real constant, the solution $\Phi(t, \mathbf{x}) \rightarrow e^{-iV^0 t} \Phi(t, \mathbf{x})$ which implies

the density of each component ρ_j ($j = 1, 2$) and the total density ρ unchanged. When $d = 1$, if the magnetic potential A_1 is perturbed by a constant, e.g. $A_1(t, \mathbf{x}) \rightarrow A_1(t, \mathbf{x}) + A_1^0$ with A_1^0 being a real constant, the solution $\Phi(t, \mathbf{x}) \rightarrow e^{iA_1^0 t \sigma_1} \Phi(t, \mathbf{x})$ implying the total density ρ unchanged; but this property is not valid when $d = 2$. When the electromagnetic potentials are time-independent, i.e. $V(t, \mathbf{x}) = V(\mathbf{x})$ and $A_j(t, \mathbf{x}) = A_j(\mathbf{x})$ for $j = 1, 2$, the following energy functional is also conserved

$$\begin{aligned} E(t) &:= \int_{\mathbb{R}^d} \left(-\frac{i}{\varepsilon} \sum_{j=1}^d \Phi^* \sigma_j \partial_j \Phi + \frac{1}{\varepsilon^2} \Phi^* \sigma_3 \Phi + V(\mathbf{x}) |\Phi|^2 - \sum_{j=1}^d A_j(\mathbf{x}) \Phi^* \sigma_j \Phi + G(\Phi) \right) d\mathbf{x} \\ &\equiv E(0), \quad t \geq 0, \end{aligned} \quad (3.1.16)$$

where

$$G(\Phi) = \frac{\lambda_1}{2} (\Phi^* \sigma_3 \Phi)^2 + \frac{\lambda_2}{2} |\Phi|^4, \quad \Phi \in \mathbb{C}^2. \quad (3.1.17)$$

If $V(t, x)$ and $A_1(t, x)$ are independent of time. In addition, if the magnetic potential A_1 is perturbed by a constant, e.g. $A_1(t, \mathbf{x}) \rightarrow A_1(t, \mathbf{x}) + A_1^0$ with A_1^0 being a real constant, then the solution $\Psi(t, \mathbf{x}) \rightarrow e^{iA_1^0 t \alpha_1} \Psi(t, \mathbf{x})$ which implies the total density ρ unchanged; this is not valid in (3+1) dimensions. Furthermore, if the external electromagnetic potentials are constants, i.e. $V(t, \mathbf{x}) \equiv V^0$ and $A_j(t, \mathbf{x}) \equiv A_j^0$ for $j = 1, 2$, the Dirac equation (3.1.11) admits the plane wave solution as $\Phi(t, \mathbf{x}) = \mathbf{B} e^{i(\mathbf{k} \cdot \mathbf{x} - \omega t)}$, where the time frequency ω , amplitude vector $\mathbf{B} \in \mathbb{R}^2$ and spatial wave number $\mathbf{k} = (k_1, \dots, k_d)^T \in \mathbb{R}^d$ satisfy the following *dispersion relation*

$$\omega \mathbf{B} = \left[\sum_{j=1}^d \left(\frac{k_j}{\varepsilon} - A_j^0 \right) \sigma_j + \frac{1}{\varepsilon^2} \sigma_3 + V^0 I_2 + \lambda_1 (\mathbf{B}^* \sigma_3 \mathbf{B}) \beta + \lambda_2 |\mathbf{B}|^2 I_2 \right] \mathbf{B}. \quad (3.1.18)$$

Particularly, when $d = 1$, $\varepsilon = 1$, $V(t, x) \equiv 0$ and $A_1(t, x) \equiv 0$ in (3.1.11) and $\lambda_1 = -1$ and $\lambda_2 = 0$ in (3.1.12), denote

$$A(t, x) = \frac{\sqrt{(1-\sigma^2)(1+\sigma)/\lambda} \cosh(x\sqrt{1-\sigma^2})}{1 + \sigma \cosh(2x\sqrt{1-\sigma^2})} e^{-i\sigma t}, \quad (3.1.19)$$

$$B(t, x) = i \frac{\sqrt{(1-\sigma^2)(1-\sigma)/\lambda} \sinh(x\sqrt{1-\sigma^2})}{1 + \sigma \cosh(2x\sqrt{1-\sigma^2})} e^{-i\sigma t}, \quad (3.1.20)$$

the NLDE (3.1.11) admits the following soliton solution $\Phi(t, x) = (\phi_1(t, x), \phi_2(t, x))^T$ [5]:

$$\begin{aligned} \phi_1(t, x) = & \sqrt{\frac{1+\gamma}{2}} A(\gamma(t - v(x - x_0)), \gamma((x - x_0) - vt)) \\ & + \text{sign}(v) \sqrt{\frac{\gamma-1}{2}} B(\gamma(t - v(x - x_0)), \gamma((x - x_0) - vt)) \end{aligned} \quad (3.1.21)$$

$$\begin{aligned} \phi_2(t, x) = & \sqrt{\frac{1+\gamma}{2}} B(\gamma(t - v(x - x_0)), \gamma((x - x_0) - vt)) \\ & + \text{sign}(v) \sqrt{\frac{\gamma-1}{2}} A(\gamma(t - v(x - x_0)), \gamma((x - x_0) - vt)) \end{aligned} \quad (3.1.22)$$

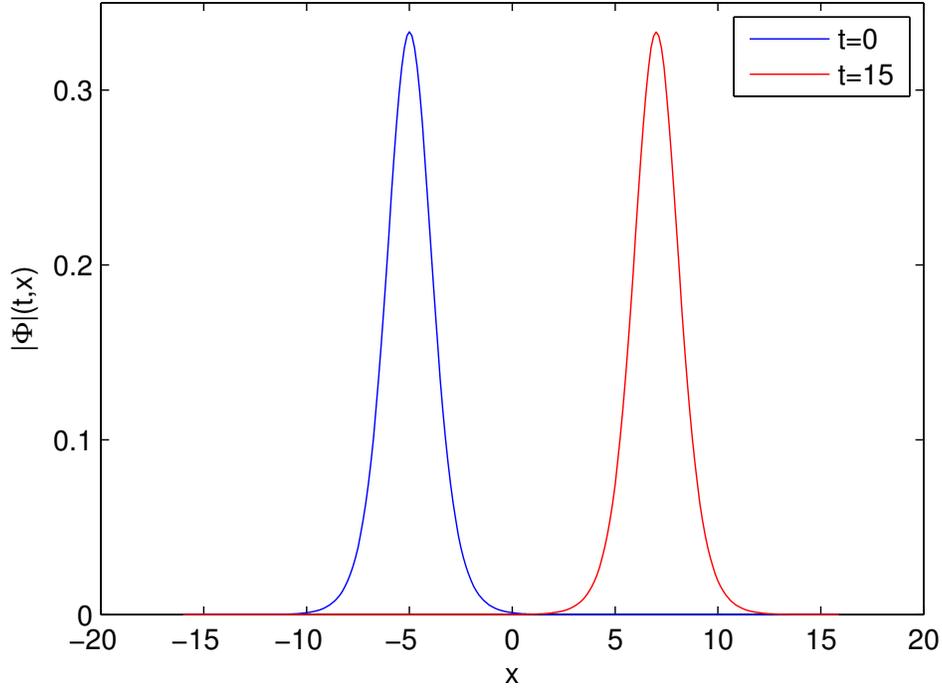


Figure 3.1: The soliton solution $|\Phi(t = 0, x)|$ and $|\Phi(t = 15, x)|$ of the Dirac equation (3.1.11) with $\varepsilon = 1$, where $|\Phi(t, x)| = \sqrt{|\phi_1(t, x)|^2 + |\phi_2(t, x)|^2}$, the speed $v = 0.8$.

3.2 Finite difference methods

In this section, we apply several conventional FDTD methods to the NLDE (3.1.11) (or (3.1.2)) with external electromagnetic field and analyze their stabilities and convergence in the nonrelativistic limit regime. For simplicity of notations, we shall only present the numerical methods and their analysis for (3.1.11) in 1D. Generalization to (3.1.2) and/or higher dimensions is straightforward and results remain valid without modifications. Similar to most works in the literatures for the analysis and computation of the NLDE (cf. [4, 6, 43, 56, 87, 102, 107] and references therein), in practical computation, we truncate the whole space problem onto an interval $\Omega = (a, b)$ with periodic boundary conditions, which is large enough such that the truncation error is negligible. In 1D, the NLDE (3.1.11) with periodic boundary conditions collapses to

$$i\partial_t\Phi(t, x) = \left[-\frac{i}{\varepsilon}\sigma_1\partial_x + \frac{1}{\varepsilon^2}\sigma_3 + V(t, x)I_2 - A_1(t, x) - F(\Phi)\sigma_1 \right] \Phi(t, x), \quad x \in \Omega, \quad t > 0, \quad (3.2.1)$$

$$\Phi(t, a) = \Phi(t, b), \quad \partial_x\Phi(t, a) = \partial_x\Phi(t, b), \quad t \geq 0, \quad \Phi(0, x) = \Phi_0(x), \quad x \in \overline{\Omega}, \quad (3.2.2)$$

where $\Phi_0(a) = \Phi_0(b)$, $\Phi'_0(a) = \Phi'_0(b)$ and $\mathbf{F}(\Phi)$ is given in (3.1.12).

3.2.1 Finite difference time domain methods

Similar as in Chapter 2, choose mesh size $h := \Delta x = \frac{b-a}{M}$ with M being an even positive integer, time step $\tau := \Delta t > 0$ and denote the grid points and time steps as:

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

Denote $X_M = \{U = (U_0, U_1, \dots, U_M)^T \mid U_j \in \mathbb{C}^2, j = 0, 1, \dots, M, U_0 = U_M\}$ and we always use $U_{-1} = U_{M-1}$ if it is involved. For any $U \in X_M$, we denote its Fourier

representation as

$$U_j = \sum_{l=-M/2}^{M/2-1} \tilde{U}_l e^{i\mu_l(x_j-a)} = \sum_{l=-M/2}^{M/2-1} \tilde{U}_l e^{2ijl\pi/M}, \quad j = 0, 1, \dots, M, \quad (3.2.3)$$

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

$$\mu_l = \frac{2l\pi}{b-a}, \quad \tilde{U}_l = \frac{1}{M} \sum_{j=0}^{M-1} U_j e^{-2ijl\pi/M}, \quad l = -\frac{M}{2}, \dots, \frac{M}{2} - 1. \quad (3.2.4)$$

The standard l^2 -norm in X_M is given as

$$\|U\|_2^2 = h \sum_{j=0}^{M-1} |U_j|^2, \quad U \in X_M. \quad (3.2.5)$$

Let Φ_j^n be the numerical approximation of $\Phi(t_n, x_j)$, denote $\Phi^n = (\Phi_0^n, \Phi_1^n, \dots, \Phi_M^n)^T \in X_M$ as the solution vector at $t = t_n$ and $V_j^n = V(t_n, x_j)$, $V_j^{n+1/2} = V(t_n + \tau/2, x_j)$, $A_{1,j}^n = A_1(t_n, x_j)$, $A_{1,j}^{n+1/2} = A_1(t_n + \tau/2, x_j)$, $\mathbf{F}_j^n = \mathbf{F}(\Phi_j^n)$ and $\mathbf{F}_j^{n+1/2} = \frac{1}{2} [\mathbf{F}(\Phi_j^n) + \mathbf{F}(\Phi_j^{n+1})]$ for $0 \leq j \leq M$ and $n \geq 0$. Introduce the finite difference discretization operators for $j = 0, 1, \dots, M-1$ and $n \geq 0$ as:

$$\begin{aligned} \delta_t^+ \Phi_j^n &= \frac{\Phi_j^{n+1} - \Phi_j^n}{\tau}, & \delta_t \Phi_j^n &= \frac{\Phi_j^{n+1} - \Phi_j^{n-1}}{2\tau}, \\ \delta_x \Phi_j^n &= \frac{\Phi_{j+1}^n - \Phi_{j-1}^n}{2h}, & \Phi_j^{n+1/2} &= \frac{\Phi_j^{n+1} + \Phi_j^n}{2}. \end{aligned}$$

Here we consider several frequently used FDTD methods to discretize the NLDE

I. Leap-frog finite difference (LFFD) method

$$i\delta_t \Phi_j^n = \left[-\frac{i}{\varepsilon} \sigma_1 \delta_x + \frac{1}{\varepsilon^2} \sigma_3 \right] \Phi_j^n + [V_j^n I_2 - A_{1,j}^n \sigma_1 + \mathbf{F}_j^n] \Phi_j^n, \quad n \geq 1. \quad (3.2.6)$$

II. Semi-implicit finite difference (SIFD1) method

$$i\delta_t \Phi_j^n = -\frac{i}{\varepsilon} \sigma_1 \delta_x \Phi_j^n + \left[\frac{1}{\varepsilon^2} \sigma_3 + V_j^n I_2 - A_{1,j}^n \sigma_1 + \mathbf{F}_j^n \right] \frac{\Phi_j^{n+1} + \Phi_j^{n-1}}{2}, \quad n \geq 1. \quad (3.2.7)$$

III. Another semi-implicit finite difference (SIFD2) method

$$i\delta_t \Phi_j^n = \left[-\frac{i}{\varepsilon} \sigma_1 \delta_x + \frac{1}{\varepsilon^2} \sigma_3 \right] \frac{\Phi_j^{n+1} + \Phi_j^{n-1}}{2} + [V_j^n I_2 - A_{1,j}^n \sigma_1 + \mathbf{F}_j^n] \Phi_j^n, \quad n \geq 1. \quad (3.2.8)$$

IV. Crank-Nicolson finite difference (CNFD) method

$$i\delta_t^+ \Phi_j^n = \left[-\frac{i}{\varepsilon} \sigma_1 \delta_x + \frac{1}{\varepsilon^2} \sigma_3 + V_j^{n+1/2} I_2 - A_{1,j}^{n+1/2} \sigma_1 + \mathbf{F}_j^{n+1/2} \right] \Phi_j^{n+1/2}, \quad n \geq 0. \quad (3.2.9)$$

The initial and boundary conditions in (3.2.2) are discretized as:

$$\Phi_M^{n+1} = \Phi_0^{n+1}, \quad \Phi_{-1}^{n+1} = \Phi_{M-1}^{n+1}, \quad n \geq 0, \quad \Phi_j^0 = \Phi_0(x_j), \quad j = 0, 1, \dots, M. \quad (3.2.10)$$

For the LFFD (3.2.6), SIFD1 (3.2.7) and SIFD2 (3.2.8), the first step can be computed as

$$\Phi_j^1 = \Phi_j^0 + \tau \left[-\frac{1}{\varepsilon} \sigma_1 \Phi_0'(x_j) - i \left(\frac{1}{\varepsilon^2} \sigma_3 - F^0 \sigma_3 + V_j^0 I_2 - A_{1,j}^0 \sigma_1 \right) \Phi_j^0 \right], \quad j = 0, 1, \dots, M-1. \quad (3.2.11)$$

Among all the three FDTD methods, the LFFD method (3.2.6) is completely explicit and its computational cost per step is $O(M)$. In fact, it might be the simplest and most efficient discretization for the NLDE when $\varepsilon = 1$ and thus it has been widely used when $\varepsilon = 1$. The SIFD1 method (3.2.7) is implicit, however at each time step for $n \geq 1$, the linear system is decoupled and so can be solved explicitly for $j = 0, 1, \dots, M-1$

$$\begin{aligned} \Phi_j^{n+1} = & \left[(i - \tau V_j^n) I_2 + (F(\Phi_j^n) - \frac{1}{\varepsilon^2}) \sigma_3 + \tau A_{1,j}^n \sigma_1 \right]^{-1} \\ & \left[\left((i - \tau V_j^n) I_2 + (F(\Phi_j^n) - \frac{1}{\varepsilon^2}) \sigma_3 + \tau A_{1,j}^n \sigma_1 \right) \Phi_j^{n-1} - \frac{2i\tau}{\varepsilon} \sigma_1 \delta_x \Phi_j^n \right], \end{aligned} \quad (3.2.12)$$

and thus the computational cost per step is also $O(M)$.

The SIFD2 (3.2.8) is implicit, however at each time step for $n \geq 1$, the corresponding linear system is decoupled in phase (Fourier) space and can be solved explicitly in phase space for $l = -M/2, \dots, M/2 - 1$ as

$$\begin{aligned} (\widetilde{\Phi^{n+1}})_l = & \left(i I_2 - \frac{\tau \sin(\mu_l h)}{\varepsilon h} \sigma_1 - \frac{\tau}{\varepsilon^2} \sigma_3 \right)^{-1} \\ & \left[\left(i I_2 + \frac{\tau \sin(\mu_l h)}{\varepsilon h} \sigma_1 + \frac{\tau}{\varepsilon^2} \sigma_3 \right) (\widetilde{\Phi^{n-1}})_l + 2\tau \mathbf{G}(\widetilde{\Phi^n})_l \right], \end{aligned} \quad (3.2.13)$$

where $\mathbf{G}(\Phi^n) = (\mathbf{G}(\Phi^n)_0, \mathbf{G}(\Phi^n)_1, \dots, \mathbf{G}(\Phi^n)_M)^T \in X_M$ with

$$\mathbf{G}(\Phi^n)_j = [V_j^n I_2 - A_{1,j}^n \sigma_1 + \mathbf{F}_j^n] \Phi_j^n \quad j = 0, 1, \dots, M,$$

and thus its computational cost per step is $O(M \ln M)$. The CNFD method (3.2.9) is implicit and at each time step for $n \geq 0$, we need to solve a nonlinear coupled system. It needs to be solved via a solver for nonlinear coupled system, and thus its computational cost per step depends on which nonlinear method to choose, and it is usually much larger than $O(M)$, especially in 2D and 3D. Based on the computational cost per time step, the LFFD method is the most efficient one and the CNFD method is the most expensive one.

3.2.2 Linear stability analysis

Let $0 < T < T^*$ with T^* being the maximal existence time of the solution, and denote $\Omega_T = [0, T] \times \Omega$. In order to carry out the linear stability analysis for the FDTD methods via the von Neumann method [90], we assume the electromagnetic potentials $V \in C(\overline{\Omega_T})$ and $A_1 \in C(\overline{\Omega_T})$ and denote

$$V_{\max} := \max_{(t,x) \in \overline{\Omega_T}} |V(t, x)|, \quad A_{1,\max} := \max_{(t,x) \in \overline{\Omega_T}} |A_1(t, x)|. \quad (3.2.14)$$

Similarly to the linear stability analysis for the FDTD methods to the Dirac equation via the von Neumann method in [15], we can show that the CNFD method (3.2.9) is unconditionally stable, i.e. it is stable for any $\tau > 0$, $h > 0$ and $0 < \varepsilon \leq 1$; the LFFD method (3.2.6), SIFD1 method (3.2.7) and SIFD2 method (3.2.8) are stable under the following stability conditions

$$\text{LFFD : } \quad 0 < \tau \leq \frac{\varepsilon^2 h}{\varepsilon^2 h (V_{\max} + F_{\max}) + \sqrt{h^2 + \varepsilon^2 (1 + \varepsilon h A_{1,\max})^2}}, \quad (3.2.15)$$

$$\text{SIFD1 : } \quad 0 < \tau \leq \varepsilon h, \quad h > 0, \quad 0 < \varepsilon \leq 1, \quad (3.2.16)$$

$$\text{SIFD2 : } \quad 0 < \tau \leq \frac{1}{V_{\max} + A_{1,\max} + F_{\max}}, \quad (3.2.17)$$

where $F_{\max} = (|\lambda_1| + |\lambda_2|) \max_{0 \leq j \leq M, n \geq 0} |\Phi_j^n|^2$.

□

3.2.3 Mass conservation and energy conservation

For the CNFD method (3.2.9), we have the following conservative properties.

Lemma 3.1 *The CNFD (3.2.9) conserves the mass in the discretized level, i.e.*

$$\|\Phi^n\|_{l^2}^2 := h \sum_{j=0}^{M-1} |\Phi_j^n|^2 \equiv h \sum_{j=0}^{M-1} |\Phi_j^0|^2 = \|\Phi^0\|_{l^2}^2 = h \sum_{j=0}^{M-1} |\Phi_0(x_j)|^2, \quad n \geq 0. \quad (3.2.18)$$

Furthermore, if $V(t, x) = V(x)$ and $A_1(t, x) = A_1(x)$ are time independent, the CNFD method (3.2.9) conserves the energy as well,

$$\begin{aligned} E_h^n &= -\frac{i}{\varepsilon} h \sum_{j=1}^{M-1} (\Phi_j^n)^* \sigma_1 \delta_x \Phi_j^n + \frac{h}{\varepsilon^2} \sum_{j=0}^{M-1} (\Phi_j^n)^* \sigma_3 \Phi_j^n + h \sum_{j=0}^{M-1} V_j (\Phi_j^n)^* \sigma_3 \Phi_j^n \\ &\quad - h \sum_{j=0}^{M-1} A_{1,j} (\Phi_j^n)^* \sigma_1 \Phi_j^n + h \sum_{j=0}^{M-1} G(\Phi_j^n) \\ &\equiv E_h^0, \quad n \geq 0, \end{aligned}$$

where $G(\Phi)$ is given in (3.1.17).

Proof: The proof of mass conservation (3.2.18) of the CNFD method is similar to the case of the Dirac equation in [15] and thus it is omitted here for brevity. In order to prove the energy conservation (3.2.19), multiplying both sides of (3.2.9) from left by $2h(\Phi_j^{n+1} - \Phi_j^n)^*$ and taking the real part, noticing (3.1.12) and (3.1.17), we have

$$\begin{aligned} &-h \operatorname{Re} \left[\frac{i}{\varepsilon} (\Phi_j^{n+1} - \Phi_j^n)^* \sigma_1 \delta_x (\Phi_j^{n+1} + \Phi_j^n) \right] + \frac{h}{\varepsilon^2} [(\Phi_j^{n+1})^* \sigma_3 \Phi_j^{n+1} - (\Phi_j^n)^* \sigma_3 \Phi_j^n] \\ &+ G(\Phi_j^{n+1}) - G(\Phi_j^n) + hV(x_j)(|\Phi_j^{n+1}|^2 - |\Phi_j^n|^2) \\ &- hA_1(x_j) [(\Phi_j^{n+1})^* \sigma_1 \Phi_j^{n+1} - (\Phi_j^n)^* \sigma_1 \Phi_j^n] = 0. \end{aligned} \quad (3.2.19)$$

Summing (3.2.19) for $j = 0, 1, \dots, M-1$ and noticing the summation by parts formula

$$h \sum_{j=0}^{M-1} \operatorname{Re} \left(\frac{i}{\varepsilon} (\Phi_j^{n+1} - \Phi_j^n)^* \sigma_1 \delta_x (\Phi_j^{n+1} + \Phi_j^n) \right) = \frac{ih}{\varepsilon} \sum_{j=0}^{M-1} [(\Phi_j^{n+1})^* \sigma_1 \delta_x \Phi_j^{n+1} - (\Phi_j^n)^* \sigma_1 \delta_x \Phi_j^n],$$

we have

$$\begin{aligned}
0 = & -\frac{ih}{\varepsilon} \sum_{j=0}^{M-1} (\Phi_j^{n+1})^* \sigma_1 \delta_x \Phi_j^{n+1} + \frac{h}{\varepsilon^2} \sum_{j=0}^{M-1} (\Phi_j^{n+1})^* \sigma_3 \Phi_j^{n+1} + h \sum_{j=0}^{M-1} G(\Phi_j^{n+1}) \\
& + h \sum_{j=0}^{M-1} [V(x_j) |\Phi_j^{n+1}|^2 - A_1(x_j) (\Phi_j^{n+1})^* \sigma_1 \Phi_j^{n+1}] + \frac{ih}{\varepsilon} \sum_{j=0}^{M-1} (\Phi_j^n)^* \sigma_1 \delta_x \Phi_j^n \\
& - \frac{h}{\varepsilon^2} \sum_{j=0}^{M-1} (\Phi_j^n)^* \sigma_3 \Phi_j^n - h \sum_{j=0}^{M-1} G(\Phi_j^n) - h \sum_{j=0}^{M-1} [V(x_j) |\Phi_j^n|^2 - A_1(x_j) (\Phi_j^n)^* \sigma_1 \Phi_j^n],
\end{aligned}$$

which immediately implies (3.2.19). □

3.2.4 Main results on error estimates

Let $0 < T < T^*$ with T^* being the maximal existence time of the solution, and denote $\Omega_T = [0, T] \times \Omega$. Motivated by the analytical results of the NLDE, we assume that the exact solution of (3.2.1) satisfies $\Phi \in C^3([0, T]; (L^\infty(\Omega))^2) \cap C^2([0, T]; (W_p^{1,\infty}(\Omega))^2) \cap C^1([0, T]; (W_p^{2,\infty}(\Omega))^2) \cap C([0, T]; (W_p^{3,\infty}(\Omega))^2)$ and

$$(A') \quad \left\| \frac{\partial^{r+s}}{\partial t^r \partial x^s} \Phi \right\|_{L^\infty([0, T]; (L^\infty(\Omega))^2)} \lesssim \frac{1}{\varepsilon^{2r}}, \quad 0 \leq r \leq 3, \quad 0 \leq r+s \leq 3, \quad 0 < \varepsilon \leq 1, \quad (3.2.20)$$

where $W_p^{m,\infty}(\Omega) = \{u \mid u \in W^{m,\infty}(\Omega), \partial_x^l u(a) = \partial_x^l u(b), l = 0, \dots, m-1\}$ for $m \geq 1$ 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 $V \in C(\bar{\Omega}_T)$ and $A_1 \in C(\bar{\Omega}_T)$ and denote

$$(B') \quad V_{\max} := \max_{(t,x) \in \bar{\Omega}_T} |V(t, x)|, \quad A_{1,\max} := \max_{(t,x) \in \bar{\Omega}_T} |A_1(t, x)|. \quad (3.2.21)$$

Define the grid error function $\mathbf{e}^n = (\mathbf{e}_0^n, \mathbf{e}_1^n, \dots, \mathbf{e}_M^n)^T \in X_M$ as:

$$\mathbf{e}_j^n = \Phi(t_n, x_j) - \Phi_j^n, \quad j = 0, 1, \dots, M, \quad n \geq 0, \quad (3.2.22)$$

with Φ_j^n being the approximations obtained from the FDTD methods.

For the CNFD (3.2.9), we can establish the following error bound.

Theorem 3.1 *Assume $0 < \tau \lesssim \varepsilon^3 h^{\frac{1}{4}}$, under the assumptions (A') and (B'), there exist constants $h_0 > 0$ and $\tau_0 > 0$ sufficiently small and independent of ε , such that for any $0 < \varepsilon \leq 1$, when $0 < h \leq h_0$ and $0 < \tau \leq \tau_0$ satisfying $0 < h \lesssim \varepsilon^{\frac{2}{3}}$, we have the following error estimate for the CNFD method (3.2.9) with (3.2.10)*

$$\|\mathbf{e}^n\|_{l^2} \lesssim \frac{h^2}{\varepsilon} + \frac{\tau^2}{\varepsilon^6}, \quad \|\Phi^n\|_{l^\infty} \leq 1 + M_0, \quad 0 \leq n \leq \frac{T}{\tau}. \quad (3.2.23)$$

For the LFFD (3.2.6), we assume the stability condition

$$0 < \tau \leq \frac{\varepsilon^2 h}{\varepsilon^2 h (V_{\max} + F_{\max}) + \sqrt{h^2 + \varepsilon^2 (1 + \varepsilon h A_{1,\max})^2}}, \quad h > 0, \quad 0 < \varepsilon \leq 1, \quad (3.2.24)$$

and establish the following error estimate.

Theorem 3.2 *Assume $0 < \tau \lesssim \varepsilon^3 h^{\frac{1}{4}}$, under the assumptions (A') and (B'), there exist constants $h_0 > 0$ and $\tau_0 > 0$ sufficiently small and independent of ε , such that for any $0 < \varepsilon \leq 1$, when $0 < h \leq h_0$ and $0 < \tau \leq \tau_0$ satisfying $0 < \tau \lesssim \min\{h, \varepsilon^2\}$, $0 < h \lesssim \varepsilon^{\frac{2}{3}}$ and the stability condition (3.2.15), we have the following error estimate for the LFFD (3.2.6) with (3.2.10) and (3.2.11)*

$$\|\mathbf{e}^n\|_{l^2} \lesssim \frac{h^2}{\varepsilon} + \frac{\tau^2}{\varepsilon^6}, \quad \|\Phi^n\|_{l^\infty} \leq 1 + M_0, \quad 0 \leq n \leq \frac{T}{\tau}. \quad (3.2.25)$$

Similar to the proofs of the LFFD and CNFD methods, error estimates for SIFD1 (3.2.7) and SIFD2 (3.2.8) can be derived.

Theorem 3.3 *Assume $0 < \tau \lesssim \varepsilon^3 h^{\frac{1}{4}}$, under the assumptions (A') and (B'), there exist constants $h_0 > 0$ and $\tau_0 > 0$ sufficiently small and independent of ε , such that for any $0 < \varepsilon \leq 1$, when $0 < h \leq h_0$ and $0 < \tau \leq \tau_0$ satisfying $0 < \tau \lesssim h$, $0 < h \lesssim \varepsilon^{\frac{2}{3}}$ and the stability condition (3.2.16), we have the following error estimate for the SIFD1 (3.2.7) with (3.2.10) and (3.2.11)*

$$\|\mathbf{e}^n\|_{l^2} \lesssim \frac{h^2}{\varepsilon} + \frac{\tau^2}{\varepsilon^6}, \quad \|\Phi^n\|_{l^\infty} \leq 1 + M_0, \quad 0 \leq n \leq \frac{T}{\tau}.$$

Theorem 3.4 *Assume $0 < \tau \lesssim \varepsilon^3 h^{\frac{1}{4}}$, under the assumptions (A') and (B'), there exist constants $h_0 > 0$ and $\tau_0 > 0$ sufficiently small and independent of ε , such that for any $0 < \varepsilon \leq 1$, when $0 < h \leq h_0$ and $0 < \tau \leq \tau_0$ satisfying $0 < h \lesssim \varepsilon^{\frac{2}{3}}$ and the stability condition (3.2.17), we have the following error estimate for the SIFD2 (3.2.8) with (3.2.10) and (3.2.11)*

$$\|\mathbf{e}^n\|_{l^2} \lesssim \frac{h^2}{\varepsilon} + \frac{\tau^2}{\varepsilon^6}, \quad \|\Phi^n\|_{l^\infty} \leq 1 + M_0, \quad 0 \leq n \leq \frac{T}{\tau}.$$

Based on Theorems 3.1-3.4, the four FDTD methods studied here share the same temporal/spatial resolution capacity in the nonrelativistic limit regime. In fact, given an accuracy bound $\delta > 0$, the ε -scalability of the four FDTD methods is:

$$\tau = O\left(\varepsilon^3 \sqrt{\delta}\right) = O(\varepsilon^3), \quad h = O\left(\sqrt{\delta \varepsilon}\right) = O(\sqrt{\varepsilon}), \quad 0 < \varepsilon \ll 1.$$

Remark 3.1 *The above Theorems are still valid in high dimensions provided that the conditions $0 < \tau \lesssim \varepsilon^3 h^{\frac{1}{4}}$ and $0 < h \lesssim \varepsilon^{\frac{2}{3}}$ are replaced by $0 < \tau \lesssim \varepsilon^3 h^{C_d}$ and $0 < h \lesssim \varepsilon^{\frac{1}{2(1-C_d)}}$, respectively, with $C_d = \frac{d}{4}$ for $d = 1, 2, 3$.*

3.2.5 Proof of the error estimates for the CNFD method

Proof of Theorem 3.1. Comparison to the proof of the CNFD method for the Dirac equation, it is expected that the main difficulty is to show the numerical solution Φ^n is uniformly bounded, i.e. $\|\Phi^n\|_{l^\infty} \lesssim 1$. In order to do so, we adapt the cut-off technique to truncate the nonlinearity $\mathbf{F}(\Phi)$ to a global Lipschitz function with compact support. Here to overcome this problem, we use the cut-off technique to truncate the nonlinearity $\mathbf{F}(\Phi)$ to a global Lipschitz function with compact support [12–14]. Noting the regularity assumption (A'), we have

$$M_0 = \max_{0 \leq t \leq T} \|\Phi(t, x)\|_{(L^\infty(\Omega))^2} \quad (3.2.26)$$

Choose a smooth function $\alpha(\rho) (\rho > 0) \in C^\infty([0, \infty))$ defined as

$$\alpha(\rho) = \begin{cases} 1, & 0 \leq \rho \leq 1 \\ \in [0, 1], & 1 \leq \rho \leq 2 \\ 0, & \rho \geq 2 \end{cases}$$

Denote $M_1 = 2(1 + M_0)^2 > 0$ and define

$$\mathbf{F}_{M_1}(\Phi) = \alpha \left(\frac{|\Phi|^2}{M_1} \right) \mathbf{F}(\Phi), \quad \Phi \in \mathbb{C}^2, \quad (3.2.27)$$

then $\mathbf{F}_{M_1}(\Phi)$ has compact support and is smooth and global Lipschitz, i.e.,

$$\|\mathbf{F}_{M_1}(\Phi_1) - \mathbf{F}_{M_1}(\Phi_2)\| \leq C_{M_1} \left| |\Phi_1| - |\Phi_2| \right| \lesssim \left| |\Phi_1| - |\Phi_2| \right|, \quad \Phi_1, \Phi_2 \in \mathbb{C}^2, \quad (3.2.28)$$

where C_{M_1} is a constant independent of ε , h and τ . Choose $\tilde{\Phi}^n \in X_M$ ($n \geq 0$) such that $\tilde{\Phi}^0 = \Phi^0$ and $\tilde{\Phi}^n$ ($n \geq 1$), with $\tilde{\Phi}^n = (\tilde{\Phi}_0^n, \tilde{\Phi}_1^n, \dots, \tilde{\Phi}_M^n)^T$ and $\tilde{\Phi}_j^n = (\tilde{\phi}_{1,j}^n, \tilde{\phi}_{2,j}^n)^T$ for $j = 0, 1, \dots, M$, be the numerical solution of the following finite difference equation

$$i\delta_t^+ \tilde{\Phi}_j^n = \left[-\frac{i}{\varepsilon} \sigma_1 \delta_x + \frac{1}{\varepsilon^2} \sigma_3 + V_j^{n+1/2} I_2 - A_{1,j}^{n+1/2} \sigma_1 + \mathbf{F}_{M_1,j}^{n+1/2} \right] \tilde{\Phi}_j^{n+1/2}, \quad 0 \leq j \leq M-1, \quad n \geq 0, \quad (3.2.29)$$

where $\tilde{\Phi}_j^{n+1/2} = \frac{1}{2} [\tilde{\Phi}_j^n + \tilde{\Phi}_j^{n+1}]$ and $\mathbf{F}_{M_1,j}^{n+1/2} = \frac{1}{2} [\mathbf{F}_{M_1}(\tilde{\Phi}_j^n) + \mathbf{F}_{M_1}(\tilde{\Phi}_j^{n+1})]$ for $j = 0, 1, \dots, M$. In fact, we can view $\tilde{\Phi}^n$ as another approximation to $\Phi(t_n, x)$. Define the corresponding errors:

$$\tilde{\mathbf{e}}_j^n = \Phi(t_n, x_j) - \tilde{\Phi}_j^n, \quad j = 0, 1, \dots, M, \quad n \geq 0$$

Then the local truncation error $\tilde{\xi}^n \in X_M$ of the scheme (3.2.29) is defined as

$$\begin{aligned} \tilde{\xi}_j^n &:= i\delta_t^+ \Phi(t_n, x_j) \\ &\quad - \left[-\frac{i}{\varepsilon} \sigma_1 \delta_x + \frac{1}{\varepsilon^2} \sigma_3 + V_j^{n+1/2} I_2 - A_{1,j}^{n+1/2} \sigma_1 + \mathbf{W}_j^n(\Phi) \right] \frac{\Phi(t_{n+1}, x_j) + \Phi(t_n, x_j)}{2}, \end{aligned} \quad (3.2.30)$$

where

$$\mathbf{W}_j^n(\Phi) = \frac{1}{2} [\mathbf{F}_{M_1}(\Phi(t_n, x_j)) + \mathbf{F}_{M_1}(\Phi(t_{n+1}, x_j))], \quad j = 0, 1, \dots, M, \quad n \geq 0. \quad (3.2.31)$$

Taking the Taylor expansion in the local truncation error (3.2.30), noticing (3.2.1) and (3.2.27), under the assumptions of (A') and (B'), with the help of triangle

inequality and Cauchy-Schwartz inequality, we have

$$\begin{aligned} |\tilde{\xi}_j^n| &\leq \frac{\tau^2}{24} \|\partial_{ttt}\Phi\|_{L^\infty(\bar{\Omega}_T)} + \frac{h^2}{6\varepsilon} \|\partial_{xxx}\Phi\|_{L^\infty(\bar{\Omega}_T)} + \frac{\tau^2}{8\varepsilon} \|\partial_{xtt}\Phi\|_{L^\infty(\bar{\Omega}_T)} \\ &\quad + \frac{\tau^2}{8} \left(\frac{1}{\varepsilon^2} + 2 + 2(|\lambda_1| + |\lambda_2|)M_0^2 + V_{\max} + A_{1,\max} \right) \|\partial_{tt}\Phi\|_{L^\infty(\bar{\Omega}_T)} \\ &\lesssim \frac{\tau^2}{\varepsilon^6} + \frac{h^2}{\varepsilon} + \frac{\tau^2}{\varepsilon^5} + \frac{\tau^2}{\varepsilon^6} \lesssim \frac{\tau^2}{\varepsilon^6} + \frac{h^2}{\varepsilon}, \quad j = 0, 1, \dots, M-1, \quad n \geq 0. \end{aligned} \quad (3.2.32)$$

Subtracting (3.2.30) from (3.2.29), we can obtain

$$i\delta_t^+ \tilde{\mathbf{e}}_j^n = \left[-\frac{i}{\varepsilon} \sigma_1 \delta_x + \frac{1}{\varepsilon^2} \sigma_3 + V_j^{n+1/2} I_2 - A_{1,j}^{n+1/2} \sigma_1 \right] \tilde{\mathbf{e}}_j^{n+1/2} + \tilde{\xi}_j^n + \tilde{\eta}_j^n, \quad (3.2.33)$$

where $\tilde{\mathbf{e}}_j^{n+1/2} = \frac{1}{2} [\tilde{\mathbf{e}}_j^n + \tilde{\mathbf{e}}_j^{n+1}]$, $0 \leq j \leq M-1$, $n \geq 0$ and

$$\tilde{\eta}_j^n = \frac{1}{2} \mathbf{W}_j^n(\Phi) [\Phi(t_{n+1}, x_j) + \Phi(t_n, x_j)] - \mathbf{F}_{M_1, j}^{n+1/2} \tilde{\Phi}_j^{n+1/2}, \quad 0 \leq j \leq M-1, \quad n \geq 0. \quad (3.2.34)$$

Combining (3.2.34), (3.2.31) and (3.2.28), we get

$$|\tilde{\eta}_j^n| \lesssim |\tilde{\mathbf{e}}_j^{n+1}| + |\tilde{\mathbf{e}}_j^n|, \quad 0 \leq j \leq M-1, \quad n \geq 0. \quad (3.2.35)$$

Multiplying both sides of (3.2.33) by $h(\tilde{\mathbf{e}}_j^{n+1/2})^*$, summing them up for $j = 0, 1, \dots, M-1$, taking imaginary parts and applying the Cauchy inequality, noticing (3.2.32), we can have

$$\begin{aligned} \|\tilde{\mathbf{e}}^{n+1}\|_{l^2}^2 - \|\tilde{\mathbf{e}}^n\|_{l^2}^2 &\lesssim \tau \left(\|\tilde{\xi}^n\|_{l^2}^2 + \|\tilde{\xi}^n\|_{l^2}^2 + \|\tilde{\mathbf{e}}^{n+1}\|_{l^2}^2 + \|\tilde{\mathbf{e}}^n\|_{l^2}^2 \right) \\ &\lesssim \tau \left[\left(\frac{h^2}{\varepsilon} + \frac{\tau^2}{\varepsilon^6} \right)^2 + \|\tilde{\mathbf{e}}^{n+1}\|_{l^2}^2 + \|\tilde{\mathbf{e}}^n\|_{l^2}^2 \right], \quad n \geq 0. \end{aligned} \quad (3.2.36)$$

Summing the above inequality, we obtain

$$\|\tilde{\mathbf{e}}^n\|_{l^2}^2 - \|\tilde{\mathbf{e}}^0\|_{l^2}^2 \lesssim \tau \sum_{l=0}^n \|\tilde{\mathbf{e}}^l\|_{l^2}^2 + \left(\frac{h^2}{\varepsilon} + \frac{\tau^2}{\varepsilon^6} \right)^2, \quad 0 \leq n \leq \frac{T}{\tau}. \quad (3.2.37)$$

Using the discrete Gronwall's inequality and noting $\tilde{\mathbf{e}}^0 = \mathbf{0}$, there exist $0 < \tau_1 \leq \frac{1}{2}$ and $h_1 > 0$ sufficiently small and independent of ε , when $0 < \tau \leq \tau_1$ and $0 < h \leq h_1$, we get

$$\|\tilde{\mathbf{e}}^n\|_{l^2} \lesssim \frac{h^2}{\varepsilon} + \frac{\tau^2}{\varepsilon^6}, \quad 0 \leq n \leq \frac{T}{\tau}. \quad (3.2.38)$$

Applying the inverse inequality in 1D, we have

$$\|\tilde{\mathbf{e}}^n\|_{l^\infty} \lesssim \frac{1}{\sqrt{h}} \|\tilde{\mathbf{e}}^n\|_{l^2} \lesssim \frac{h^{\frac{3}{2}}}{\varepsilon} + \frac{\tau^2}{\varepsilon^6 \sqrt{h}}, \quad 0 \leq n \leq \frac{T}{\tau}. \quad (3.2.39)$$

Under the conditions $0 < \tau \lesssim \varepsilon^3 h^{\frac{1}{4}}$ and $0 < h \lesssim \varepsilon^{2/3}$, there exist $h_2 > 0$ and $\tau_2 > 0$ sufficiently small and independent of ε , when $0 < h \leq h_2$ and $0 < \tau \leq \tau_2$, we get

$$\|\tilde{\Phi}^n\|_{l^\infty} \leq \|\Phi\|_{L^\infty(\Omega_T)} + \|\tilde{\mathbf{e}}^n\|_{l^\infty} \leq 1 + M_0, \quad 0 \leq n \leq \frac{T}{\tau}. \quad (3.2.40)$$

Therefore, under the conditions in Theorem 3.1, the discretization (3.2.29) collapses exactly to the CNFD discretization (3.2.9) for the NLDE if we take $\tau_0 = \min\{1/2, \tau_1, \tau_2\}$ and $h_0 = \min\{h_1, h_2\}$, i.e.

$$\tilde{\Phi}^n = \Phi^n, \quad 0 \leq n \leq \frac{T}{\tau}. \quad (3.2.41)$$

Thus the proof is completed. □

Remark 3.2 *In the proof above we use the inverse inequality in 1D. In fact we can have the inverse inequality in 2D and 3D*

$$\|e^n\|_{l^\infty} \lesssim \frac{1}{C_d(h)} \|e^n\|_{l^2}, \quad C_d(h) = \begin{cases} h^{\frac{1}{2}}, & d = 1, \\ h, & d = 2, \\ h^{\frac{3}{2}}, & d = 3. \end{cases} \quad (3.2.42)$$

and this proof can be easily extended to 2D and 3D cases, with the requirement $\tau \lesssim h$.

3.2.6 Proof of the error estimates for the LFFD method

Proof of Theorem 3.2. Again, comparison to the proof of the LFFD method for the Dirac equation, the main difficulty is to show the numerical solution Φ^n is uniformly bounded, i.e. $\|\Phi^n\|_{l^\infty} \lesssim 1$. In order to do so, we adapt the method of mathematical

induction [12–14]. Define the local truncation error $\hat{\xi}^n = (\xi_0^n, \xi_1^n, \dots, \xi_M^n)^T \in X_M$ of the LFFD (3.2.6) as

$$\begin{aligned} \hat{\xi}_j^0 &:= i\delta_t^+ \Phi(0, x_j) + \frac{1}{\tau} \sin\left(\frac{\tau}{\varepsilon}\right) \sigma_1 \delta_x \Phi_0(x_j) \\ &\quad + i \left(\frac{1}{\tau} \sin\left(\frac{\tau}{\varepsilon^2}\right) \sigma_3 + V_j^0 I_2 - A_{1,j}^0 \sigma_1 + \mathbf{F}(\Phi(0, x_j)) \right) \Phi(0, x_j), \end{aligned} \quad (3.2.43)$$

$$\hat{\xi}_j^n := i\delta_t \Phi(t_n, x_j) - \left[-\frac{i}{\varepsilon} \sigma_1 \delta_x + \frac{1}{\varepsilon^2} \sigma_3 + V_j^n I_2 - A_{1,j}^n \sigma_1 + \mathbf{F}(\Phi(t_n, x_j)) \right] \Phi(t_n, x_j), \quad (3.2.44)$$

for $0 \leq j < M, n \geq 1$. Similarly to the proof of Theorem 3.1, applying the Taylor expansion, we obtain

$$|\hat{\xi}_j^0| \lesssim \frac{\tau}{\varepsilon^4} + \frac{h^2}{\varepsilon}, \quad |\hat{\xi}_j^n| \lesssim \frac{\tau^2}{\varepsilon^6} + \frac{h^2}{\varepsilon}, \quad j = 0, 1, \dots, M-1, \quad n \geq 1. \quad (3.2.45)$$

Subtracting (3.2.11) and (3.2.6) from (3.2.43) and (3.2.44), respectively, we get the error equations

$$i\delta_t \mathbf{e}_j^n = \left[-\frac{i}{\varepsilon} \sigma_1 \delta_x + \frac{1}{\varepsilon^2} \sigma_3 + V_j^n I_2 - A_{1,j}^n \sigma_1 \right] \mathbf{e}_j^n + \eta_j^n + \hat{\xi}_j^n, \quad 0 \leq j \leq M-1, \quad n \geq 1, \quad (3.2.46)$$

$$\mathbf{e}_j^1 = \tau \hat{\xi}_j^0 + \mathbf{e}_j^0, \quad \mathbf{e}_j^0 = \Phi(0, x_j) - \Phi_j^0 = 0, \quad j = 0, 1, \dots, M-1, \quad (3.2.47)$$

where $\eta^n \in X_M$ is given as

$$\eta_j^n = \mathbf{F}(\Phi(t_n, x_j)) \Phi(t_n, x_j) - \mathbf{F}(\Phi_j^n) \Phi_j^n, \quad j = 0, 1, \dots, M, \quad n \geq 1. \quad (3.2.48)$$

From (3.2.6), we know that (3.2.25) is valid for $n = 0$. In addition, noticing (3.2.45) and assume $0 < \tau \leq 1$, we have

$$\|\mathbf{e}^1\|_{l^2} \lesssim \|\mathbf{e}^1\|_{l^\infty} \leq \tau \|\hat{\xi}^0\|_{l^\infty} \lesssim \frac{\tau^2}{\varepsilon^4} + \frac{\tau h^2}{\varepsilon} \lesssim \frac{\tau^2}{\varepsilon^6} + \frac{h^2}{\varepsilon}. \quad (3.2.49)$$

By using the inverse inequality, we get

$$\|\mathbf{e}^1\|_{l^\infty} \lesssim \frac{1}{h^{1/2}} \|\mathbf{e}^1\|_{l^2} \lesssim \frac{\tau^2}{\varepsilon^6 h^{1/2}} + \frac{h^{3/2}}{\varepsilon}. \quad (3.2.50)$$

Thus, under the conditions in Theorem 3.2, there exist $h_1 > 0$ and $\tau_1 > 0$ sufficiently small and independent of ε such that, for $0 < \varepsilon \leq 1$, when $0 < h \leq h_1$ and $0 < \tau \leq \tau_1$, we have

$$\|\Phi^1\|_{l^\infty} \leq \|\Phi(t_1, x)\|_{L^\infty} + \|\mathbf{e}^1\|_{l^\infty} \leq 1 + M_0, \quad (3.2.51)$$

which immediately implies that (3.2.25) is valid for $n = 1$.

Now we assume that (3.2.25) is valid for $0 \leq n \leq m \leq \frac{T}{\tau} - 1$. From (3.2.48), we have

$$\begin{aligned} |\eta_j^l| &= |F(\Phi(t_l, x_j))\Phi(t_l, x_j) - F(\Phi_j^l)\Phi_j^l| \\ &\leq \|F(\Phi(t_l, x_j)) - F(\Phi_j^l)\| |\Phi(t_l, x_j)| + \|F(\Phi_j^l)\| |\Phi(t_l, x_j) - \Phi_j^l| \\ &\lesssim |\mathbf{e}_j^l|, \quad j = 0, 1, \dots, M-1, \quad l = 0, 1, \dots, m. \end{aligned} \quad (3.2.52)$$

Denote \mathcal{E}^l ($l \geq 0$) as

$$\mathcal{E}^l = \|\mathbf{e}^{l+1}\|_{l^2}^2 + \|\mathbf{e}^l\|_{l^2}^2 + 2 \operatorname{Re} \left(\tau h \sum_{j=0}^{M-1} (\mathbf{e}_j^{l+1})^* \sigma_1 \delta_x \mathbf{e}_j^l \right) - 2 \operatorname{Im} \left(\frac{\tau h}{\varepsilon^2} \sum_{j=0}^{M-1} (\mathbf{e}_j^{l+1})^* \sigma_3 \mathbf{e}_j^l \right). \quad (3.2.53)$$

Under the stability condition (3.2.15) and the conditions in Theorem 3.2, for $0 < \varepsilon \leq 1$, when $\tau > 0$ and $h > 0$ satisfying $0 < \frac{\tau}{h} \leq \frac{1}{4}$ and $0 < \frac{\tau}{\varepsilon^2} \leq \frac{1}{4}$, using the Cauchy inequality, we obtain

$$\frac{1}{2} (\|\mathbf{e}^{l+1}\|_{l^2}^2 + \|\mathbf{e}^l\|_{l^2}^2) \leq \mathcal{E}^l \leq \frac{3}{2} (\|\mathbf{e}^{l+1}\|_{l^2}^2 + \|\mathbf{e}^l\|_{l^2}^2). \quad (3.2.54)$$

From (3.2.25) with $n = 0$ and $n = 1$, we have

$$\mathcal{E}^0 \lesssim \left(\frac{h^2}{\varepsilon} + \frac{\tau^2}{\varepsilon^6} \right)^2. \quad (3.2.55)$$

Multiplying both sides of (3.2.46) from the left by $2h\tau (\mathbf{e}_j^{n+1} + \mathbf{e}_j^{n-1})^*$, taking the imaginary part, then summing for $j = 0, 1, \dots, M-1$, using the Cauchy inequality, (3.2.45) and (3.2.54), we get

$$\begin{aligned} \mathcal{E}^l - \mathcal{E}^{l-1} &\lesssim h\tau \sum_{j=0}^{M-1} [(A_{1,\max} + V_{\max})|\mathbf{e}_j^l| + |\eta_j^n| + |\xi_j^l|] (|\mathbf{e}_j^{l+1}| + |\mathbf{e}_j^{l-1}|) \\ &\lesssim \tau (\mathcal{E}^l + \mathcal{E}^{l-1}) + \tau \left(\frac{h^2}{\varepsilon} + \frac{\tau^2}{\varepsilon^6} \right)^2, \quad l \geq 1. \end{aligned}$$

Summing the above inequality for $l = 1, 2, \dots, m$, we get

$$\mathcal{E}^m - \mathcal{E}^0 \lesssim \tau \sum_{l=0}^m \mathcal{E}^l + m\tau \left(\frac{h^2}{\varepsilon} + \frac{\tau^2}{\varepsilon^6} \right)^2.$$

There exist $0 < \tau_2 \leq \frac{1}{2}$ and $h_2 > 0$ sufficiently small and independent of ε , when $0 < \tau \leq \tau_2$ and $0 < h \leq h_2$, using the discrete Gronwall's inequality and noticing (3.2.55), we obtain

$$\|\mathbf{e}^{m+1}\|_{l^2}^2 \leq 2\mathcal{E}^m \lesssim \left(\frac{h^2}{\varepsilon} + \frac{\tau^2}{\varepsilon^6} \right)^2, \quad 1 \leq m \leq \frac{T}{\tau} - 1. \quad (3.2.56)$$

In addition, by using the inverse inequality, we get

$$\|\mathbf{e}^{m+1}\|_{l^\infty} \lesssim \frac{1}{h^{1/2}} \|\mathbf{e}^{m+1}\|_{l^2} \lesssim \frac{\tau^2}{\varepsilon^6 h^{1/2}} + \frac{h^{3/2}}{\varepsilon}. \quad (3.2.57)$$

Thus, under the conditions in Theorem 3.2, there exist $h_3 > 0$ and $\tau_3 > 0$ sufficiently small and independent of ε such that, for $0 < \varepsilon \leq 1$, when $0 < h \leq h_3$ and $0 < \tau \leq \tau_3$, we have

$$\|\Phi^{m+1}\|_{l^\infty} \leq \|\Phi(t_{m+1}, x)\|_{L^\infty} + \|\mathbf{e}^{m+1}\|_{l^\infty} \leq 1 + M_0, \quad (3.2.58)$$

which immediately implies that (3.2.25) is valid for $n = m+1$. Thus we complete the proof of Theorem 3.2 by taking $\tau_0 = \min\{1/4, \tau_1, \tau_2, \tau_3\}$ and $h_0 = \min\{1, h_1, h_2, h_3\}$.

□

Proof of Theorem 3.3 and 3.4. Follow the analogous proofs of Theorem 3.1 and 3.2, we omit the details here for brevity.

3.3 Exponential wave integrator pseudospectral methods

In this section, we propose an exponential wave integrator Fourier pseudospectral (EWI-FP) method to solve the NLDE (3.2.1) and establish its stability and convergence in the nonrelativistic limit regime. Generalization to higher dimensions is similar to Chapter 2 and the results remain valid without modifications.

3.3.1 The EWI-FP method

Denote

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

Let $[C_p(\bar{\Omega})]^2$ be the function space consisting of all periodic vector function $U(x) : \bar{\Omega} = [a, b] \rightarrow \mathbb{C}^2$. For any $U(x) \in [C_p(\bar{\Omega})]^2$ and $U \in X_M$, define $P_M : [L^2(\Omega)]^2 \rightarrow Y_M$ as the standard projection operator [89], $I_M : [C_p(\bar{\Omega})]^2 \rightarrow Y_M$ and $I_M : X_M \rightarrow Y_M$ as the standard interpolation operator, i.e.

$$(P_M U)(x) = \sum_{l=-M/2}^{M/2-1} \hat{U}_l e^{i\mu_l(x-a)}, \quad (I_M U)(x) = \sum_{l=-M/2}^{M/2-1} \tilde{U}_l e^{i\mu_l(x-a)}, \quad a \leq x \leq b,$$

with

$$\begin{aligned} \hat{U}_l &= \frac{1}{b-a} \int_a^b U(x) e^{-i\mu_l(x-a)} dx, \\ \tilde{U}_l &= \frac{1}{M} \sum_{j=0}^{M-1} U_j e^{-2ijl\pi/M}, \quad l = -\frac{M}{2}, -\frac{M}{2} + 1, \dots, \frac{M}{2} - 1, \end{aligned} \quad (3.3.1)$$

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

The Fourier spectral discretization for Dirac equation (3.2.1) is as follows:

Find $\Phi_M(t, x) \in Y_M$, i.e.

$$\Phi_M(t, x) = \sum_{l=-M/2}^{M/2-1} \widehat{(\Phi_M)}_l(t) e^{i\mu_l(x-a)}, \quad a \leq x \leq b, \quad t \geq 0, \quad (3.3.2)$$

such that for $a < x < b$ and $t > 0$,

$$i\partial_t \Phi_M = \left[-\frac{i}{\varepsilon} \sigma_1 \partial_x + \frac{1}{\varepsilon^2} \sigma_3 \right] \Phi_M + P_M [(V(t, x)I_2 - A_1(t, x)\sigma_1 + \mathbf{F}(\Phi_M)) \Phi_M]. \quad (3.3.3)$$

Substituting (3.3.2) into (3.3.3), noticing the orthogonality of $\phi_l(x)$, we get for $l = -\frac{M}{2}, -\frac{M}{2} + 1, \dots, \frac{M}{2} - 1$,

$$i \frac{d}{dt} \widehat{(\Phi_M)}_l(t) = \left[\frac{\mu_l}{\varepsilon} \sigma_1 + \frac{1}{\varepsilon^2} \sigma_3 \right] + \widehat{\mathbf{G}(\Phi_M)}_l(t), \quad t > 0, \quad (3.3.4)$$

where

$$\mathbf{G}(\Phi_M) = (V(t, x)I_2 - A_1(t, x)\sigma_1 + \mathbf{F}(\Phi_M)) \Phi_M, \quad x \in \Omega, \quad t \geq 0. \quad (3.3.5)$$

For $t \geq 0$ and each l ($l = -\frac{M}{2}, -\frac{M}{2} + 1, \dots, \frac{M}{2} - 1$), when t is near $t = t_n$ ($n \geq 0$), we rewrite the above ODEs as

$$i \frac{d}{ds} (\widehat{\Phi_M})_l(t_n + s) = \frac{1}{\varepsilon^2} \Gamma_l (\widehat{\Phi_M})_l(t_n + s) + \mathbf{G}(\widehat{\Phi_M})_l^n(s), \quad s > 0, \quad (3.3.6)$$

where $\Gamma_l = \mu_l \varepsilon \sigma_1 + \sigma_3 = Q_l D_l (Q_l)^*$ with $\delta_l = \sqrt{1 + \varepsilon^2 \mu_l^2}$ and

$$\Gamma_l = \begin{pmatrix} 1 & \mu_l \varepsilon \\ \mu_l \varepsilon & -1 \end{pmatrix}, \quad Q_l = \begin{pmatrix} \frac{1+\delta_l}{\sqrt{2\delta_l(1+\delta_l)}} & -\frac{\varepsilon\mu_l}{\sqrt{2\delta_l(1+\delta_l)}} \\ \frac{\varepsilon\mu_l}{\sqrt{2\delta_l(1+\delta_l)}} & \frac{1+\delta_l}{\sqrt{2\delta_l(1+\delta_l)}} \end{pmatrix}, \quad D_l = \begin{pmatrix} \delta_l & 0 \\ 0 & -\delta_l \end{pmatrix}, \quad (3.3.7)$$

and

$$\mathbf{G}(\widehat{\Phi_M})_l^n(s) = \mathbf{G}(\widehat{\Phi_M})_l(t_n + s), \quad s \geq 0, \quad n \geq 0, \quad (3.3.8)$$

Solving the above ODE (3.3.6) via the integrating factor method, we obtain

$$(\widehat{\Phi_M})_l(t_n + s) = e^{-is\Gamma_l/\varepsilon^2} (\widehat{\Phi_M})_l(t_n) - i \int_0^s e^{i(w-s)\Gamma_l/\varepsilon^2} \mathbf{G}(\widehat{\Phi_M})_l^n(w) dw, \quad s \geq 0. \quad (3.3.9)$$

Taking $s = \tau$ in (3.3.9) we have

$$(\widehat{\Phi_M})_l(t_{n+1}) = e^{-i\tau\Gamma_l/\varepsilon^2} (\widehat{\Phi_M})_l(t_n) - i \int_0^\tau e^{\frac{i(w-\tau)\Gamma_l}{\varepsilon^2}} \mathbf{G}(\widehat{\Phi_M})_l^n(w) dw. \quad (3.3.10)$$

To obtain a numerical method with second order accuracy in time, we approximate the integral in (3.3.10) via the Gautschi-type rule, which has been widely used for integrating highly oscillatory ODEs [17, 44, 51], as

$$\begin{aligned} \int_0^\tau e^{\frac{i(w-\tau)\Gamma_l}{\varepsilon^2}} \mathbf{G}(\widehat{\Phi_M})_l^0(w) dw &\approx \int_0^\tau e^{\frac{i(w-\tau)\Gamma_l}{\varepsilon^2}} dw \mathbf{G}(\widehat{\Phi_M})_l^0(0) \\ &= -i\varepsilon^2 \Gamma_l^{-1} \left[I_2 - e^{-\frac{i\tau}{\varepsilon^2} \Gamma_l} \right] \mathbf{G}(\widehat{\Phi_M})_l^0(0), \end{aligned} \quad (3.3.11)$$

and for $n \geq 1$

$$\begin{aligned} \int_0^\tau e^{\frac{i(w-\tau)\Gamma_l}{\varepsilon^2}} \mathbf{G}(\widehat{\Phi_M})_l^n(w) dw &\approx \int_0^\tau e^{\frac{i(w-\tau)\Gamma_l}{\varepsilon^2}} \left(\mathbf{G}(\widehat{\Phi_M})_l^n(0) + w \delta_t^- \mathbf{G}(\widehat{\Phi_M})_l^n(0) \right) dw \\ &= -i\varepsilon^2 \Gamma_l^{-1} \left[I - e^{-\frac{i\tau}{\varepsilon^2} \Gamma_l} \right] \mathbf{G}(\widehat{\Phi_M})_l^n(0) \\ &\quad + \left[-i\varepsilon^2 \tau \Gamma_l^{-1} + \varepsilon^4 \Gamma_l^{-2} \left(I - e^{-\frac{i\tau}{\varepsilon^2} \Gamma_l} \right) \right] \delta_t^- \mathbf{G}(\widehat{\Phi_M})_l^n(0), \end{aligned} \quad (3.3.12)$$

where we have approximated the time derivative $\partial_t \widehat{\mathbf{G}}(\widehat{\Phi}_M)_l^n(s)$ at $s = 0$ by finite difference as

$$\partial_t \widehat{\mathbf{G}}(\widehat{\Phi}_M)_l^n(0) \approx \delta_t^- \widehat{\mathbf{G}}(\widehat{\Phi}_M)_l^n(0) = \frac{1}{\tau} \left[\widehat{\mathbf{G}}(\widehat{\Phi}_M)_l^n(0) - \widehat{\mathbf{G}}(\widehat{\Phi}_M)_l^{n-1}(0) \right]. \quad (3.3.13)$$

Now, we are ready to describe our scheme. Let $\Phi_M^n(x)$ be the approximation of $\Phi_M(t_n, x)$ ($n \geq 0$). Choosing $\Phi_M^0(x) = (P_M \Phi_0)(x)$, an *exponential wave integrator Fourier spectral* (EWI-FS) discretization for the NLDE (3.2.1) is to update the numerical approximation $\Phi_M^{n+1}(x) \in Y_M$ ($n = 0, 1, \dots$) as

$$\Phi_M^{n+1}(x) = \sum_{l=-M/2}^{M/2-1} (\widehat{\Phi}_M^{n+1})_l e^{i\mu_l(x-a)}, \quad a \leq x \leq b, \quad n \geq 0, \quad (3.3.14)$$

where for $l = -\frac{M}{2}, \dots, \frac{M}{2} - 1$,

$$(\widehat{\Phi}_M^{n+1})_l = \begin{cases} e^{-i\tau\Gamma_l/\varepsilon^2} (\widehat{\Phi}_M^0)_l - i\varepsilon^2 \Gamma_l^{-1} \left[I_2 - e^{-\frac{i\tau}{\varepsilon^2} \Gamma_l} \right] \widehat{\mathbf{G}}(\widehat{\Phi}_M^0)_l, & n = 0, \\ e^{-i\tau\Gamma_l/\varepsilon^2} (\widehat{\Phi}_M^n)_l - iQ_l^{(1)}(\tau) \widehat{\mathbf{G}}(\widehat{\Phi}_M^n)_l - iQ_l^{(2)}(\tau) \delta_t^- \widehat{\mathbf{G}}(\widehat{\Phi}_M^n)_l, & n \geq 1, \end{cases} \quad (3.3.15)$$

with the matrices $Q_l^{(1)}(\tau)$ and $Q_l^{(2)}(\tau)$ given as

$$Q_l^{(1)}(\tau) = -i\varepsilon^2 \Gamma_l^{-1} \left[I - e^{-\frac{i\tau}{\varepsilon^2} \Gamma_l} \right], \quad Q_l^{(2)}(\tau) = -i\varepsilon^2 \tau \Gamma_l^{-1} + \varepsilon^4 \Gamma_l^{-2} \left(I - e^{-\frac{i\tau}{\varepsilon^2} \Gamma_l} \right), \quad (3.3.16)$$

and

$$\mathbf{G}(\Phi_M^n) = (V(t_n, x)I_2 - A_1(t_n, x)\sigma_1 + \mathbf{F}(\Phi_M^n)) \Phi_M^n, \quad n \geq 0. \quad (3.3.17)$$

The above procedure is not suitable in practice due to the difficulty in computing the Fourier coefficients through integrals in (3.3.1). Here we present an efficient implementation by choosing $\Phi_M^0(x)$ as the interpolant of $\Phi_0(x)$ on the grids $\{x_j, j = 0, 1, \dots, M\}$ and approximate the integrals in (3.3.1) by a quadrature rule.

Let Φ_j^n be the numerical approximation of $\Phi(t_n, x_j)$ for $j = 0, 1, 2, \dots, M$ and $n \geq 0$, and denote $\Phi^n \in X_M$ as the vector with components Φ_j^n . Choosing $\Phi_j^0 = \Phi_0(x_j)$ ($j = 0, 1, \dots, M$), an *EWI Fourier pseudospectral* (EWI-FP) method for computing Φ^{n+1} for $n \geq 0$ reads

$$\Phi_j^{n+1} = \sum_{l=-M/2}^{M/2-1} (\widetilde{\Phi}^{n+1})_l e^{2ijl\pi/M}, \quad j = 0, 1, \dots, M, \quad (3.3.18)$$

where

$$\widetilde{(\Phi^{n+1})}_l = \begin{cases} e^{-i\tau\Gamma_l/\varepsilon^2} \widetilde{(\Phi^0)}_l - i\varepsilon^2\Gamma_l^{-1} \left[I_2 - e^{-\frac{i\tau}{\varepsilon^2}\Gamma_l} \right] \widetilde{(W(\Phi^0))}_l, & n = 0, \\ e^{-i\tau\Gamma_l/\varepsilon^2} \widetilde{(\Phi^n)}_l - iQ_l^{(1)}(\tau) \widetilde{(W(\Phi^n))}_l - iQ_l^{(2)}(\tau) \delta_t^- \widetilde{(W(\Phi^n))}_l, & n \geq 1. \end{cases} \quad (3.3.19)$$

The EWI-FP (3.3.18)-(3.3.19) is explicit, and can be solved efficiently by the fast Fourier transform (FFT). The memory cost is $O(M)$ and the computational cost per time step is $O(M \log M)$.

Similarly to the analysis of the EWI-FP method for the Dirac equation in Chapter 2, we can obtain that the EWI-FP for the NLDE is stable under the stability condition (details are omitted here for brevity)

$$0 < \tau \lesssim 1, \quad 0 < \varepsilon \leq 1. \quad (3.3.20)$$

3.3.2 Convergence analysis

In order to obtain an error estimate for the EWI methods (3.3.14)-(3.3.15) and (3.3.18)-(3.3.19), motivated by the results in [21, 26], we assume that there exists an integer $m_0 \geq 2$ such that the exact solution $\Phi(t, x)$ of NLDE (3.2.1) satisfies

$$(C') \quad \|\Phi\|_{L^\infty([0, T]; (H_p^{m_0})^2)} \lesssim 1, \quad \|\partial_t \Phi\|_{L^\infty([0, T]; (L^2)^2)} \lesssim \frac{1}{\varepsilon^2}, \quad \|\partial_{tt} \Phi\|_{L^\infty([0, T]; (L^2)^2)} \lesssim \frac{1}{\varepsilon^4},$$

where $H_p^k(\Omega) = \{u \mid u \in H^k(\Omega), \partial_x^l u(a) = \partial_x^l u(b), l = 0, \dots, k-1\}$. In addition, we assume electromagnetic potentials satisfy

$$(D') \quad \|V\|_{W^{2, \infty}([0, T]; L^\infty)} + \|A_1\|_{W^{2, \infty}([0, T]; L^\infty)} \lesssim 1.$$

Under the above assumptions, the following are well defined,

$$M_0 := \max_{0 \leq t \leq T} \|\Phi(t, x)\|_{L^\infty} \lesssim 1 \quad (3.3.21)$$

The following estimate can be established.

Theorem 3.5 *Let $\Phi_M^n(x)$ be the approximation obtained from the EWI-FS (3.3.14)-(3.3.15). Assume $\tau \lesssim \varepsilon^2 C_d(h)$, under the assumptions (C') and (D'), there exists*

$h_0 > 0$ and $\tau_0 > 0$ sufficiently small and independent of ε such that, for any $0 < \varepsilon \leq 1$, when $0 < h \leq h_0$ and $0 < \tau \leq \tau_0$, we have the following error estimate

$$\|\Phi(t_n, x) - \Phi_M^n(x)\|_{L^2} \lesssim \frac{\tau^2}{\varepsilon^4} + h^{m_0}, \quad \|\Phi_M^n(x)\|_{L^\infty} \leq M_1 + 1, \quad 0 \leq n \leq \frac{T}{\tau}. \quad (3.3.22)$$

Proof of Theorem 3.5. Here the main difficulty is to show that the numerical solution $\Phi_M^n(x)$ is uniformly bounded, i.e. $\|\Phi_M^n(x)\|_{L^\infty} \lesssim 1$, which will be established by the method of mathematical induction [12–14]. Define the error function $\mathbf{e}^n(x) \in Y_M$ for $n \geq 0$ as

$$\mathbf{e}^n(x) = P_M \Phi(t_n, x) - \Phi_M^n(x) = \sum_{l=-M/2}^{M/2-1} \widehat{\mathbf{e}}_l^n e^{i\mu_l(x-a)}, \quad a \leq x \leq b, \quad n \geq 0. \quad (3.3.23)$$

Using the triangular inequality and standard interpolation result, we get

$$\begin{aligned} \|\Phi(t_n, x) - \Phi_M^n(x)\|_{L^2} &\leq \|\Phi(t_n, x) - P_M \Phi(t_n, x)\|_{L^2} + \|\mathbf{e}^n(x)\|_{L^2} \\ &\leq h^{m_0} + \|\mathbf{e}^n(x)\|_{L^2}, \quad 0 \leq n \leq \frac{T}{\tau}. \end{aligned} \quad (3.3.24)$$

Thus we only need estimate $\|\mathbf{e}^n(x)\|_{L^2}$. It is easy to see that (3.3.22) is valid when $n = 0$.

Define the local truncation error $\xi^n(x) = \sum_{l=-M/2}^{M/2-1} \widehat{\xi}_l^n e^{i\mu_l(x-a)} \in Y_M$ of the EWI-FP (3.3.15) for $n \geq 0$ as

$$\widehat{\xi}_l^n = \begin{cases} (\widehat{\Phi(\tau)})_l - e^{-i\tau\Gamma_l/\varepsilon^2} (\widehat{\Phi(0)})_l + i\varepsilon^2 \Gamma_l^{-1} [I_2 - e^{-\frac{i\tau}{\varepsilon^2}\Gamma_l}] \widehat{\mathbf{G}(\Phi)}_l(0), & n = 0, \\ (\widehat{\Phi(t_{n+1})})_l - e^{-i\tau\Gamma_l/\varepsilon^2} (\widehat{\Phi(t_n)})_l + iQ_l^{(1)}(\tau) \widehat{\mathbf{G}(\Phi)}_l(t_n) + iQ_l^{(2)}(\tau) \delta_t^- \widehat{\mathbf{G}(\Phi)}_l(t_n), & n \geq 1, \end{cases} \quad (3.3.25)$$

where we denote $\Phi(t)$ and $\mathbf{G}(\Phi)$ in short for $\Phi(t, x)$ and $\mathbf{G}(\Phi(t, x))$ in (3.3.17), respectively, for the simplicity of notations. In order to estimate the local truncation error $\xi^n(x)$, multiplying both sides of the NLDE (3.2.1) by $e^{i\mu_l(x-a)}$ and integrating over the interval (a, b) , we easily recover the equations for $\widehat{\Phi(t)}_l$, which are exactly the same as (3.3.6) with Φ_M being replaced by $\Phi(t, x)$. Replacing Φ_M with $\Phi(t, x)$, we use the same notations $\widehat{\mathbf{G}(\Phi)}_l^n(s)$ as in (3.3.8) and the time derivatives of $\widehat{\mathbf{G}(\Phi)}_l^n(s)$

enjoy the same properties of time derivatives of $\Phi(t, x)$. Thus, the same representation (3.3.10) holds for $\widehat{\Phi}(t_n)_l$ for $n \geq 1$. From the derivation of the EWI-FS method, it is clear that the error $\xi^n(x)$ comes from the approximations for the integrals in (3.3.11) and (3.3.12). Thus we have

$$\begin{aligned}\widehat{\xi}_l^0 &= -i \int_0^\tau e^{\frac{i(s-\tau)}{\varepsilon^2}\Gamma_l} \left[\widehat{\mathbf{G}(\Phi)_l^0}(s) - \widehat{\mathbf{G}(\Phi)_l^0}(0) \right] ds \\ &= -i \int_0^\tau \int_0^s e^{\frac{i(s-\tau)}{\varepsilon^2}\Gamma_l} \partial_{s_1} \widehat{\mathbf{G}(\Phi)_l^0}(s_1) ds_1 ds,\end{aligned}\quad (3.3.26)$$

and for $n \geq 1$

$$\begin{aligned}\widehat{\xi}_l^n &= -i \int_0^\tau e^{\frac{i(s-\tau)}{\varepsilon^2}\Gamma_l} \int_0^s \int_0^{s_1} \partial_{s_2 s_2} \widehat{\mathbf{G}(\Phi)_l^n}(s_2) ds_2 ds_1 ds \\ &\quad - i \int_0^\tau e^{\frac{i(s-\tau)}{\varepsilon^2}\Gamma_l} s \int_0^1 \int_{\theta\tau}^\tau \partial_{\theta_1 \theta_1} \widehat{\mathbf{G}(\Phi)_l^{n-1}}(\theta_1) d\theta_1 d\theta ds.\end{aligned}\quad (3.3.27)$$

Subtracting (3.3.15) from (3.3.25), we obtain

$$\widehat{\mathbf{e}}_l^{n+1} = e^{-i\tau\Gamma_l/\varepsilon^2} \widehat{\mathbf{e}}_l^n + \widehat{R}_l^n + \widehat{\xi}_l^n, \quad 1 \leq n \leq \frac{T}{\tau} - 1, \quad (3.3.28)$$

$$\widehat{\mathbf{e}}_l^0 = \mathbf{0}, \quad \widehat{\mathbf{e}}_l^1 = \widehat{\xi}_l^0, \quad l = -\frac{M}{2}, \dots, \frac{M}{2} - 1. \quad (3.3.29)$$

where $R^n(x) = \sum_{l=-M/2}^{M/2-1} \widehat{R}_l^n e^{i\mu_l(x-a)} \in Y_M$ for $n \geq 1$ is given by

$$\widehat{R}_l^n = -iQ_l^{(1)}(\tau) \left[\widehat{\mathbf{G}(\Phi(t_n))_l} - \widehat{\mathbf{G}(\Phi_M^n)_l} \right] - iQ_l^{(2)}(\tau) \left[\delta_t^- \widehat{\mathbf{G}(\Phi(t_n))_l} - \delta_t^- \widehat{\mathbf{G}(\Phi_M^n)_l} \right]. \quad (3.3.30)$$

From (3.3.26) and (3.3.29), we have

$$|\widehat{\xi}_l^0| \lesssim \int_0^\tau \int_0^s \left| \partial_{s_1} \widehat{\mathbf{G}(\Phi)_l^0}(s_1) \right| ds_1 ds. \quad (3.3.31)$$

By the Parseval equality and assumptions (C) and (D), we get

$$\begin{aligned}\|\mathbf{e}^1(x)\|_{L^2}^2 &= \|\xi^0(x)\|_{L^2}^2 = (b-a) \sum_{l=-M/2}^{M/2-1} \left| \widehat{\xi}_l^0 \right|^2 \\ &\lesssim (b-a)\tau^2 \int_0^\tau \int_0^s \sum_{l=-M/2}^{M/2-1} \left| \partial_{s_1} \widehat{\mathbf{G}(\Phi)_l^0}(s_1) \right|^2 ds_1 ds \\ &\lesssim \tau^2 \int_0^\tau \int_0^s \|\partial_{s_1}(\mathbf{G}(\Phi(s_1)))\|_{L^2}^2 ds_1 ds \lesssim \frac{\tau^4}{\varepsilon^4} \lesssim \frac{\tau^4}{\varepsilon^8}.\end{aligned}\quad (3.3.32)$$

Thus we have

$$\|\Phi(t_1, x) - \Phi_M^1(x)\|_{L^2} \lesssim h^{m_0} + \|\mathbf{e}^1(x)\|_{L^2} \lesssim h^{m_0} + \frac{\tau^2}{\varepsilon^4}. \quad (3.3.33)$$

By using the inverse inequality, we get

$$\|\mathbf{e}^1(x)\|_{L^\infty} \leq \frac{1}{h^{1/2}} \|\mathbf{e}^1(x)\|_{L^2} \lesssim \frac{\tau^2}{\varepsilon^4 h^{1/2}}, \quad (3.3.34)$$

which immediately implies

$$\begin{aligned} \|\Phi_M^1(x)\|_{L^\infty} &\leq \|\Phi(t_1, x)\|_{L^\infty} + \|\Phi(t_1, x) - P_M \Phi(t_1, x)\|_{L^\infty} + \|\mathbf{e}^1(x)\|_{L^\infty} \\ &\leq M_0 + h^{m_0-1} + \frac{\tau^2}{\varepsilon^4 h^{1/2}}. \end{aligned} \quad (3.3.35)$$

Under the conditions in Theorem 3.5, there exist $h_1 > 0$ and $\tau_1 > 0$ sufficiently small and independent of ε , for $0 < \varepsilon \leq 1$, when $0 < h \leq h_1$ and $0 < \tau \leq \tau_1$, we have

$$\|\Phi_M^1(x)\|_{L^\infty} \leq 1 + M_0, \quad (3.3.36)$$

thus (3.3.22) is valid when $n = 1$.

Now we assume that (3.3.22) is valid for all $0 \leq n \leq m \leq \frac{T}{\tau} - 1$, then we need to show that it is still valid when $n = m + 1$. Similarly to (3.3.31) and (3.3.32), under the assumptions (C) and (D), we obtain

$$|\widehat{\xi}_l^n| \leq \int_0^\tau \left(\int_0^s \int_0^{s_1} \left| \partial_{s_2 s_2} \widehat{\mathbf{G}}(\Phi)_l^n(s_2) \right| ds_2 ds_1 + s \int_0^1 \int_{\theta\tau}^\tau \left| \partial_{\theta_1 \theta_1} \widehat{\mathbf{G}}(\Phi)_l^{n-1}(\theta_1) \right| d\theta_1 d\theta \right) ds, \quad (3.3.37)$$

$$\begin{aligned} \|\xi^n(x)\|_{L^2}^2 &= (b-a) \sum_{l=-M/2}^{M/2-1} |\widehat{\xi}_l^n|^2 \lesssim \tau^3 \int_0^\tau \int_0^s \int_0^{s_1} \sum_{l=-\frac{M}{2}}^{\frac{M}{2}-1} \left| \partial_{s_2 s_2} \widehat{\mathbf{G}}(\Phi)_l^n(s_2) \right|^2 ds_2 ds_1 ds \\ &\quad + \tau^3 \int_0^\tau \int_0^1 \int_{\theta\tau}^\tau s \sum_{l=-\frac{M}{2}}^{\frac{M}{2}-1} \left| \partial_{\theta_1 \theta_1} \widehat{\mathbf{G}}(\Phi)_l^{n-1}(\theta_1) \right|^2 d\theta_1 d\theta ds \\ &\lesssim \tau^6 \|\partial_{tt}(W(\Phi(t)))\|_{L^\infty([0,T];(L^2)^2)}^2 \lesssim \frac{\tau^6}{\varepsilon^8}, \quad n = 0, 1, \dots, m. \end{aligned} \quad (3.3.38)$$

Using the properties of the matrices $Q_l^{(1)}(\tau)$ and $Q_l^{(2)}(\tau)$, it is easy to verify that

$$\|Q_l^{(1)}(\tau)\|_2 \leq \tau, \quad \|Q_l^{(2)}(\tau)\|_2 \leq \frac{\tau^2}{2}, \quad l = -\frac{M}{2}, \dots, \frac{M}{2} - 1. \quad (3.3.39)$$

Combining (3.3.30) and (3.3.39), we get

$$\begin{aligned}
 \frac{1}{b-a} \|R^n(x)\|_{L^2}^2 &= \sum_{l=-M/2}^{M/2-1} |\widehat{R}_l^n|^2 \\
 &\lesssim \tau^2 \sum_{l=-M/2}^{M/2-1} \left[\left| (\widehat{\Phi}(t_n))_l - (\widehat{\Phi}_M^n)_l \right|^2 + \left| (\widehat{\Phi}(t_{n-1}))_l - (\widehat{\Phi}_M^{n-1})_l \right|^2 \right. \\
 &\quad \left. + \left| \widehat{\mathbf{G}}(\widehat{\Phi})_l(t_n) - \widehat{\mathbf{G}}(\widehat{\Phi}_M^n)_l \right|^2 + \left| \widehat{\mathbf{G}}(\widehat{\Phi})_l(t_{n-1}) - \widehat{\mathbf{G}}(\widehat{\Phi}_M^{n-1})_l \right|^2 \right] \\
 &\lesssim \tau^2 \left[\|\Phi(t_n, x) - \Phi_M^n(x)\|_{L^2}^2 + \|\Phi(t_{n-1}, x) - \Phi_M^{n-1}(x)\|_{L^2}^2 \right] \\
 &\lesssim \tau^2 h^{2m_0} + \tau^2 \|\mathbf{e}^n(x)\|_{L^2}^2 + \tau^2 \|\mathbf{e}^{n-1}(x)\|_{L^2}^2 \quad n = 0, 1, \dots, m.
 \end{aligned} \tag{3.3.40}$$

Multiplying both sides of (3.3.28) from left by $(\widehat{\mathbf{e}}_l^{n+1} + e^{-i\tau\Gamma_l/\varepsilon^2}\widehat{\mathbf{e}}_l^n)^*$, taking the real parts and using the Cauchy inequality, we obtain

$$|\widehat{\mathbf{e}}_l^{n+1}|^2 - |\widehat{\mathbf{e}}_l^n|^2 \leq \tau \left(|\widehat{\mathbf{e}}_l^{n+1}|^2 + |\widehat{\mathbf{e}}_l^n|^2 \right) + \frac{|\widehat{R}_l^n|^2}{\tau} + \frac{|\widehat{\xi}_l^n|^2}{\tau}. \tag{3.3.41}$$

Summing the above for $l = -M/2, \dots, M/2 - 1$ and then multiplying it by $(b-a)$, using the Parseval equality, we obtain for $n \geq 1$

$$\begin{aligned}
 \|\mathbf{e}^{n+1}(x)\|_{L^2}^2 - \|\mathbf{e}^n(x)\|_{L^2}^2 &\lesssim \tau \left(\|\mathbf{e}^{n+1}(x)\|_{L^2}^2 + \|\mathbf{e}^n(x)\|_{L^2}^2 \right) \\
 &\quad + \frac{1}{\tau} \left(\|R^n(x)\|_{L^2}^2 + \|\xi^n(x)\|_{L^2}^2 \right).
 \end{aligned} \tag{3.3.42}$$

Summing (3.3.42) for $n = 1, \dots, m$, using (3.3.40), we derive

$$\|\mathbf{e}^{m+1}(x)\|_{L^2}^2 - \|\mathbf{e}^1(x)\|_{L^2}^2 \lesssim \tau \sum_{k=1}^{m+1} \|\mathbf{e}^k(x)\|_{L^2}^2 + \frac{m\tau^5}{\varepsilon^8} + m\tau h^{2m_0}, \quad 1 \leq m \leq \frac{T}{\tau} - 1. \tag{3.3.43}$$

Noticing $\|\mathbf{e}^1(x)\|_{L^2} \lesssim \frac{\tau^2}{\varepsilon^2} \lesssim \frac{\tau^2}{\varepsilon^4}$ and using the discrete Gronwall's inequality, there exist $0 < \tau_2 \leq \frac{1}{2}$ and $h_2 > 0$ sufficiently small and independent of ε such that, for $0 < \varepsilon \leq 1$, when $0 < \tau \leq \tau_2$ and $0 < h \leq h_2$, we get

$$\|\mathbf{e}^{m+1}(x)\|_{L^2}^2 \lesssim h^{2m_0} + \frac{\tau^4}{\varepsilon^8}, \quad 1 \leq m \leq \frac{T}{\tau} - 1. \tag{3.3.44}$$

Thus we have

$$\|\Phi(t_{m+1}, x) - \Phi_M^{m+1}(x)\|_{L^2} \lesssim h^{m_0} + \|\mathbf{e}^{m+1}(x)\|_{L^2} \lesssim h^{m_0} + \frac{\tau^2}{\varepsilon^4}. \quad (3.3.45)$$

By using the inverse inequality, we get

$$\|\mathbf{e}^{m+1}(x)\|_{L^\infty} \leq \frac{1}{h^{1/2}} \|\mathbf{e}^{m+1}(x)\|_{L^2} \lesssim \frac{\tau^2}{\varepsilon^4 h^{1/2}}, \quad (3.3.46)$$

which immediately implies

$$\begin{aligned} \|\Phi_M^{m+1}(x)\|_{L^\infty} &\leq \|\Phi(t_{m+1}, x)\|_{L^\infty} + \|\Phi(t_{m+1}, x) - P_M \Phi(t_{m+1}, x)\|_{L^\infty} + \|\mathbf{e}^{m+1}(x)\|_{L^\infty} \\ &\leq M_0 + h^{m_0-1} + \frac{\tau^2}{\varepsilon^4 h^{1/2}}. \end{aligned} \quad (3.3.47)$$

Under the conditions in Theorem 3.5, there exist $h_3 > 0$ and $\tau_3 > 0$ sufficiently small and independent of ε , for $0 < \varepsilon \leq 1$, when $0 < h \leq h_3$ and $0 < \tau \leq \tau_3$, we have

$$\|\Phi_M^{m+1}(x)\|_{L^\infty} \leq 1 + M_0, \quad (3.3.48)$$

thus (3.3.22) is valid when $n = m + 1$. Then the proof of (3.3.22) is completed by the method of mathematical induction under the choice of $h_0 = \min\{h_1, h_2, h_3\}$ and $\tau_0 = \min\{1/2, \tau_1, \tau_2, \tau_3\}$.

□

Remark 3.3 *The same error estimate in Theorem 3.5 holds for the EWI-FP (3.3.18)-(3.3.19) and the proof is quite similar to that of Theorem 3.5.*

3.4 Time-splitting Fourier pseudospectral methods

In this section, we present a time-splitting Fourier pseudospectral (TSFP) method to solve the NLDE (3.2.1). Again, generalization to higher dimensions is straightforward and results remain valid without modifications.

From time $t = t_n$ to time $t = t_{n+1}$, the Dirac equation (3.2.1) is split into three steps. One solves first

$$i\partial_t\Phi(t, x) = [V(t, x)I_2 - A_1(t, x)\sigma_1 + \mathbf{F}(\Phi(t, x))] \Phi(t, x), \quad x \in \Omega, \quad (3.4.1)$$

with the periodic boundary condition (3.2.2) for the time step of length τ , followed by solving

$$i\partial_t\Phi(t, x) = [V(t, x)I_2 - A_1(t, x)\sigma_1 + \mathbf{F}(\Phi(t, x))] \Phi(t, x), \quad x \in \Omega, \quad (3.4.2)$$

for the same time step. Eq. (3.4.1) will be first discretized in space by the Fourier spectral method and then integrated (in phase or Fourier space) in time *exactly* [15, 19]. For the ODEs (3.4.2), multiplying $\Phi^*(t, x)$ from the left, we get

$$i\Phi^*(t, x)\partial_t\Phi(t, x) = \Phi^*(t, x) [V(t, x)I_2 - A_1(t, x)\sigma_1 + \mathbf{F}(\Phi(t, x))] \Phi(t, x), \quad x \in \Omega. \quad (3.4.3)$$

Taking conjugate to both sides of the above equation, noticing (3.1.12), we obtain

$$-i\partial_t\Phi^*(t, x)\Phi(t, x) = \Phi^*(t, x) [V(t, x)I_2 - A_1(t, x)\sigma_1^* + \mathbf{F}(\Phi(t, x))] \Phi(t, x), \quad x \in \Omega, \quad (3.4.4)$$

where $\sigma_1^* = \overline{\sigma_1}^T$. Summing (3.4.3) and (3.4.4), noticing (3.1.12), $\sigma_1^* = \sigma_1$ and $\sigma_3^* = \sigma_3$, we obtain for $\rho(t, x) = |\Phi(t, x)|^2$

$$\partial_t\rho(t, x) = 0, \quad t_n \leq t \leq t_{n+1}, \quad x \in \Omega, \quad (3.4.5)$$

which immediately implies $\rho(t, x) = \rho(t_n, x)$.

If $A_1(t, x) \equiv 0$, multiplying (3.4.2) from the left by $\Phi^*(t, x)\sigma_3$ and by a similar procedure, we get $\Phi^*(t, x)\sigma_3\Phi(t, x) = \Phi^*(t_n, x)\sigma_3\Phi(t_n, x)$ for $t_n \leq t \leq t_{n+1}$ and $x \in \Omega$. Thus if $\lambda_1 = 0$ or $A_1(t, x) \equiv 0$, we have

$$\mathbf{F}(\Phi(t, x)) = \mathbf{F}(\Phi(t_n, x)), \quad t_n \leq t \leq t_{n+1}, \quad x \in \Omega. \quad (3.4.6)$$

Plugging (3.4.6) into (3.4.2), we obtain

$$i\partial_t\Phi(t, x) = [V(t, x)I_2 - A_1(t, x)\sigma_1 + \mathbf{F}(\Phi(t_n, x))] \Phi(t, x), \quad x \in \Omega, \quad (3.4.7)$$

which can be integrated *analytically* in time as

$$\Phi(t, x) = e^{-i \int_{t_n}^t [V(s, x) I_2 - A_1(s, x) \sigma_1 + \mathbf{F}(\Phi(t_n, x))] ds} \Phi(t_n, x), \quad a \leq x \leq b, \quad t_n \leq t \leq t_{n+1}. \quad (3.4.8)$$

In practical computation, if $\lambda_1 = 0$ or $A_1(t, x) \equiv 0$, from time $t = t_n$ to $t = t_{n+1}$, we often combine the splitting steps via the Strang splitting [92] – which results in a second order TSFP method as

$$\begin{aligned} \Phi_j^{(1)} &= \sum_{l=-M/2}^{M/2-1} e^{-i\tau\Gamma_l/2\varepsilon^2} (\widetilde{\Phi^n})_l e^{i\mu_l(x_j-a)} = \sum_{l=-M/2}^{M/2-1} Q_l e^{-i\tau D_l/2\varepsilon^2} (Q_l)^* (\widetilde{\Phi^n})_l e^{\frac{2ijl\pi}{M}}, \\ \Phi_j^{(2)} &= e^{-i \int_{t_n}^{t_{n+1}} \mathbf{W}(t, x_j) dt} \Phi_j^{(1)} = P_j e^{-i\Lambda_j} P_j^* \Phi_j^{(1)}, \quad j = 0, 1, \dots, M, \quad n \geq 0, \\ \Phi_j^{n+1} &= \sum_{l=-M/2}^{M/2-1} e^{-i\tau\Gamma_l/2\varepsilon^2} (\widetilde{\Phi^{(2)}})_l e^{i\mu_l(x_j-a)} = \sum_{l=-M/2}^{M/2-1} Q_l e^{-i\tau D_l/2\varepsilon^2} (Q_l)^* (\widetilde{\Phi^{(2)}})_l e^{\frac{2ijl\pi}{M}}, \end{aligned} \quad (3.4.9)$$

where

$$\begin{aligned} \int_{t_n}^{t_{n+1}} \mathbf{W}(t, x_j) dt &= V_j^{(1)} I_2 - A_{1,j}^{(1)} \sigma_1 + \tau \mathbf{F}(\Phi_j^{(1)}) \\ &= \left(V_j^{(1)} + \tau \lambda_2 |\Phi_j^{(1)}|^2 \right) I_2 - A_{1,j}^{(1)} \sigma_1 + \tau \lambda_1 (\Phi_j^{(1)})^* \sigma_3 \Phi_j^{(1)} \sigma_3 \\ &= P_j \Lambda_j P_j^* \end{aligned}$$

with $V_j^{(1)} = \int_{t_n}^{t_{n+1}} V(t, x_j) dt$, $A_{1,j}^{(1)} = \int_{t_n}^{t_{n+1}} A_1(t, x_j) dt$, $\Lambda_j = \text{diag}(\Lambda_{j,+}, \Lambda_{j,-})$, and $\Lambda_{j,\pm} = V_j^{(1)} + \tau \lambda_2 |\Phi_j^{(1)}|^2 \pm \tau \lambda_1 (\Phi_j^{(1)})^* \sigma_3 \Phi_j^{(1)}$ and $P_j = I_2$ if $A_{1,j}^{(1)} = 0$, and resp., $\Lambda_{j,\pm} = V_j^{(1)} + \tau \lambda_2 |\Phi_j^{(1)}|^2 \pm A_{1,j}^{(1)}$ and

$$P_j = P^{(0)} := \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}, \quad (3.4.10)$$

if $A_{1,j}^{(1)} \neq 0$ and $\lambda_1 = 0$.

Of course, if $\lambda_1 \neq 0$ and $A_1(t, x) \neq 0$, then $\Phi^*(t, x) \sigma_3 \Phi(t, x)$ is no longer time-independent in the second step (3.4.2) due to that $\sigma_1^* \sigma_3^* = \sigma_1 \sigma_3 \neq \sigma_3 \sigma_1$. In this situation, we will split (3.4.2) into two steps as: one first solves

$$i\partial_t \Phi(t, x) = [V(t, x) I_2 - A_1(t, x) \sigma_1] \Phi(t, x), \quad x \in \Omega, \quad (3.4.11)$$

followed by solving

$$i\partial_t\Phi(t, x) = \mathbf{F}(\Phi(t, x)) \Phi(t, x), \quad x \in \Omega. \quad (3.4.12)$$

Similarly to the Dirac equation [15], Eq. (3.4.11) can be integrated *analytically* in time. For Eq. (3.4.12), both $\rho(t, x)$ and $\Phi^*(t, x)\sigma_3\Phi(t, x)$ are invariant in time, i.e. $\rho(t, x) \equiv \rho(t_n, x)$ and $\Phi^*(t, x)\sigma_3\Phi(t, x) \equiv \Phi^*(t_n, x)\sigma_3\Phi(t_n, x)$ for $t_n \leq t \leq t_{n+1}$ and $x \in \bar{\Omega}$. Thus it collapses to

$$i\partial_t\Phi(t, x) = \mathbf{F}(\Phi(t_n, x)) \Phi(t_n, x), \quad x \in \Omega, \quad (3.4.13)$$

and it can be integrated *analytically* in time too. Similarly, a second-order TSFP method can be designed provided that we replace $\Phi^{(2)}$ in the third step by $\Phi^{(4)}$ and the second step in (3.4.9) by

$$\begin{aligned} \Phi_j^{(2)} &= e^{-\frac{i}{2} \int_{t_n}^{t_{n+1}} \mathbf{F}(\Phi(t_n, x_j)) dt} \Phi_j^{(1)} = e^{-i\Lambda_j^{(1)}} \Phi_j^{(1)}, \\ \Phi_j^{(3)} &= e^{-i \int_{t_n}^{t_{n+1}} [V(t, x_j)I_2 - A_1(t, x_j)\sigma_1] dt} \Phi_j^{(2)} = P_j e^{-i\Lambda_j^{(2)}} P_j^* \Phi_j^{(2)}, \\ \Phi_j^{(4)} &= e^{-\frac{i}{2} \int_{t_n}^{t_{n+1}} \mathbf{F}(\Phi(t_n, x_j)) dt} \Phi_j^{(3)} = e^{-i\Lambda_j^{(1)}} \Phi_j^{(3)}, \quad j = 0, 1, \dots, M, \quad n \geq 0, \end{aligned} \quad (3.4.14)$$

where $\Lambda_j^{(1)} = \text{diag}(\Lambda_{j,+}^{(1)}, \Lambda_{j,-}^{(1)})$ with $\Lambda_{j,\pm}^{(1)} = \frac{\tau}{2} \left[\lambda_2 |\Phi_j^{(1)}|^2 \pm \lambda_1 (\Phi_j^{(1)})^* \sigma_3 \Phi_j^{(1)} \right]$, $\Lambda_j^{(2)} = \text{diag}(\Lambda_{j,+}^{(2)}, \Lambda_{j,-}^{(2)})$ with $\Lambda_{j,\pm}^{(2)} = V_j^{(1)} \pm A_{1,j}^{(1)}$, and $P_j = I_2$ if $A_{1,j}^{(1)} = 0$, and resp., $P_j = P^{(0)}$ if $A_{1,j}^{(1)} \neq 0$ for $j = 0, 1, \dots, M$.

Remark 3.4 *If the above definite integrals cannot be evaluated analytically, we can evaluate them numerically via the Simpson's quadrature rule as*

$$\begin{aligned} V_j^{(1)} &= \int_{t_n}^{t_{n+1}} V(t, x_j) dt \approx \frac{\tau}{6} \left[V(t_n, x_j) + 4V\left(t_n + \frac{\tau}{2}, x_j\right) + V(t_{n+1}, x_j) \right], \\ A_{1,j}^{(1)} &= \int_{t_n}^{t_{n+1}} A_1(t, x_j) dt \approx \frac{\tau}{6} \left[A_1(t_n, x_j) + 4A_1\left(t_n + \frac{\tau}{2}, x_j\right) + A_1(t_{n+1}, x_j) \right]. \end{aligned}$$

Similarly to the TSFP for the Dirac equation in Chapter 2, we can show that the TSFP (3.4.9) for the NLDE conserves the mass in the discretized level with the details omitted here for brevity.

Lemma 3.2 *The TSFP (3.4.9) conserves the mass in the discretized level, i.e.*

$$\|\Phi^n\|_{l^2}^2 := h \sum_{j=0}^{M-1} |\Phi_j^n|^2 \equiv h \sum_{j=0}^{M-1} |\Phi_j^0|^2 = \|\Phi^0\|_{l^2}^2 = h \sum_{j=0}^{M-1} |\Phi_0(x_j)|^2, \quad n \geq 0. \quad (3.4.15)$$

From Lemma 3.2, we conclude that TSFP (3.4.9) is unconditionally stable. In addition, under proper assumptions of the exact solution Φ and electromagnetic potentials as well as conditions on the mesh size h and time step τ , for $0 < \varepsilon \leq 1$, it is easy to show the following error estimate via the formal Lie calculus introduced in [66],

$$\|I_M(\Phi^n) - \Phi(t_n, x)\|_{L^2} \lesssim h^{m_0} + \frac{\tau^2}{\varepsilon^4}, \quad \|\Phi^n\|_{l^\infty} \lesssim 1 + M_0, \quad 0 \leq n \leq \frac{T}{\tau}, \quad (3.4.16)$$

where m_0 depends on the regularity of Φ . Thus the temporal/spatial resolution capacity of the TSFP method for the NLDE in the nonrelativistic limit regime is: $h = O(1)$ and $\tau = O(\varepsilon^2)$. In fact, for a given accuracy bound $\delta_0 > 0$, the ε -scalability of the TSFP is:

$$\tau = O\left(\varepsilon^2 \sqrt{\delta_0}\right) = O(\varepsilon^2), \quad h = O\left(\delta_0^{1/m_0}\right) = O(1), \quad 0 < \varepsilon \ll 1. \quad (3.4.17)$$

It is straightforward to generalize the TSFP to the NLDE (3.1.11) in 2D and (3.1.2) in 1D, 2D and 3D and the details are omitted here for brevity.

3.5 Numerical results

In this section, we compare the accuracy of different numerical methods including the FDTD, EWI-FP and TSFP methods in solving the NLDE (3.1.11) in terms of the mesh size h , time step τ and the parameter $0 < \varepsilon \leq 1$. We will pay particular attention to the ε -scalabilities of different methods in the nonrelativistic limit regime, i.e. $0 < \varepsilon \ll 1$.

To test the accuracy, we take $d = 1$, $\lambda_1 = -1$, $\lambda_2 = 0$ and choose the electromagnetic potentials in the NLDE (3.1.11) as

$$A_1(t, x) = \frac{(x+1)^2}{1+x^2}, \quad V(t, x) = \frac{1-x}{1+x^2}, \quad x \in \mathbb{R}, \quad t \geq 0,$$

and the initial value as

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

The problem is solved numerically on an interval $\Omega = (-16, 16)$, i.e. $a = -16$ and $b = 16$, with periodic boundary conditions on $\partial\Omega$. The ‘exact’ solution $\Phi(t, x) = (\phi_1(t, x), \phi_2(t, x))^T$ is obtained numerically by using the TSFP method with a very fine mesh size and a small time step, e.g. $h_e = 1/16$ and $\tau_e = 10^{-7}$ to compare with the numerical solutions obtained by EWI-FP and TSFP, and $h_e = 1/4096$ to compare with the numerical solutions obtained by FDTD methods. Denote $\Phi_{h,\tau}^n$ as the numerical solution obtained by a numerical method with mesh size h and time step τ . In order to quantify the convergence, we introduce

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

Tab. 3.1 lists spatial errors $e_{h,\tau_e}(t = 2)$ with different h (upper part) and temporal errors $e_{h_e,\tau}(t = 2)$ with different τ (lower part) for the LFFD method (3.2.6). Tabs. 3.2-3.6 show similar results for the SIFD1 method (3.2.7), SIFD2 method (3.2.8), CNFD method (3.2.9), EWI-FP method (3.3.18)-(3.3.19) and TSFP method (3.4.9), respectively. under different ε -scalability.

From Tabs. 3.1-3.6, and additional numerical results not shown here for brevity, we can draw the following conclusions for the Dirac equation by using different numerical methods:

(i). For the discretization error in space, for any fixed $\varepsilon = \varepsilon_0 > 0$, the FDTD methods are second-order accurate, and resp., the EWI-FP and TSFP methods are spectrally accurate (cf. each row in the upper parts of Tabs. 3.1-3.6). For $0 < \varepsilon \leq 1$, the errors are independent of ε for the EWI-FP and TSFP methods (cf. each column in the upper parts of Tabs. 3.5-3.6), and resp., are almost independent of ε for the FDTD methods (cf. each column in the upper parts of Tabs. 3.1-3.4). In general, for any fixed $0 < \varepsilon \leq 1$ and $h > 0$, the EWI-FP and TSFP methods perform much better than the FDTD methods in spatial discretization.

Table 3.1: Spatial and temporal error analysis of the LFFD method for the NLDE.

Spatial Errors	$h_0 = 1/8$	$h_0/2$	$h_0/2^2$	$h_0/2^3$	$h_0/2^4$
$\varepsilon_0 = 1$	8.15E-2	2.02E-2	5.00E-3	1.25E-3	3.12E-4
$\varepsilon_0/2$	9.29E-2	2.30E-2	5.73E-2	1.43E-3	3.58E-4
$\varepsilon_0/2^2$	9.91E-2	2.46E-2	6.12E-3	1.53E-3	3.82E-4
$\varepsilon_0/2^3$	9.89E-2	2.47E-2	6.17E-3	1.54E-3	3.85E-4
$\varepsilon_0/2^3$	9.87E-2	2.48E-2	6.18E-3	1.54E-3	3.83E-4
Temporal Errors	$\tau_0 = 0.1$	$\tau_0/8$	$\tau_0/8^2$	$\tau_0/8^3$	$\tau_0/8^4$
	$h_0 = 1/8$	$h_0/8\delta_1(\varepsilon)$	$h_0/8^2\delta_2(\varepsilon)$	$h_0/8^3\delta_3(\varepsilon)$	$h_0/8^4\delta_4(\varepsilon)$
$\varepsilon_0 = 1$	<u>1.95E-1</u>	2.67E-3	4.16E-5	6.50E-7	1.00E-8
$\varepsilon_0/2$	unstable	<u>2.03E-2</u>	3.14E-4	4.91E-6	7.67E-8
$\varepsilon_0/2^2$	unstable	4.65E-1	<u>7.17E-3</u>	1.11E-4	1.74E-6
$\varepsilon_0/2^3$	unstable	unstable	4.13E-1	<u>6.08E-3</u>	1.01E-4
$\varepsilon_0/2^4$	unstable	unstable	3.48	4.04E-1	<u>6.20E-3</u>

(ii). For the discretization error in time, in the $O(1)$ speed-of-light regime, i.e. $\varepsilon = O(1)$, all the numerical methods including FDTD, EWI-FP and TSFP are second-order accurate (cf. the first row in the lower parts of Tabs. 3.1-3.6). In general, the EWI-FP and TSFP methods perform much better than the FDTD methods in temporal discretization for a fixed time step. In the non-relativistic limit regime, i.e. $0 < \varepsilon \ll 1$, for the FDTD methods, the ‘correct’ ε -scalability is $\tau = O(\varepsilon^3)$ which verifies our theoretical results; for the EWI-FP and TSFP methods, the ‘correct’ ε -scalability is $\tau = O(\varepsilon^2)$ which again confirms our theoretical results. In fact, for $0 < \varepsilon \leq 1$, one can observe clearly second-order convergence in time for the FDTD methods only when $\tau \lesssim \varepsilon^3$ (cf. upper triangles in the lower parts of Tabs. 3.1-3.4), and resp., for the EWI-FP and TSFP methods when $\tau \lesssim \varepsilon^2$ (cf. upper triangles in the lower parts of Tabs. 3.5-3.6). In general, for any fixed $0 < \varepsilon \leq 1$ and $\tau > 0$, the TSFP method performs the best, and the EWI-FP method performs

Table 3.2: Spatial and temporal error analysis of the SIFD1 method for the NLDE.

Spatial Errors	$h_0 = 1/8$	$h_0/2$	$h_0/2^2$	$h_0/2^3$	$h_0/2^4$
$\varepsilon_0 = 1$	8.15E-2	2.02E-2	5.00E-3	1.25E-3	3.12E-4
$\varepsilon_0/2$	9.29E-2	2.30E-2	5.73E-2	1.43E-3	3.58E-4
$\varepsilon_0/2^2$	9.91E-2	2.46E-2	6.12E-3	1.53E-3	3.82E-4
$\varepsilon_0/2^3$	9.89E-2	2.47E-2	6.17E-3	1.54E-3	3.85E-4
$\varepsilon_0/2^3$	9.87E-2	2.48E-2	6.18E-3	1.54E-3	3.83E-4
Temporal Errors	$\tau_0 = 0.1$	$\tau_0/8$	$\tau_0/8^2$	$\tau_0/8^3$	$\tau_0/8^4$
	$h_0 = 1/8$	$h_0/8\delta_1(\varepsilon)$	$h_0/8^2\delta_2(\varepsilon)$	$h_0/8^3\delta_3(\varepsilon)$	$h_0/8^4\delta_4(\varepsilon)$
$\varepsilon_0 = 1$	<u>1.69E-1</u>	2.16E-3	4.08E-5	6.38E-7	9.81E-9
$\varepsilon_0/2$	unstable	<u>3.23E-2</u>	5.04E-4	7.87E-6	1.23E-7
$\varepsilon_0/2^2$	unstable	8.22E-1	<u>1.62E-2</u>	2.05E-4	3.20E-6
$\varepsilon_0/2^3$	unstable	unstable	8.00E-1	<u>1.32E-2</u>	1.97E-4
$\varepsilon_0/2^4$	unstable	ubstable	4.44E-1	7.97E-1	<u>1.27E-2</u>

much better than the FDTD methods in temporal discretization.

(iii). From Tab. 3.6, our numerical results suggest the following error bound for the TSFP method when $\tau \lesssim \varepsilon^2$,

$$\|I_M(\Phi^n) - \Phi(t_n, \cdot)\|_{L^2} \lesssim h^{m_0} + \frac{\tau^2}{\varepsilon^2}, \quad (3.5.1)$$

which is much better than (3.4.16) for the TSFP method in the nonrelativistic limit regime. Rigorous mathematical justification for (3.5.1) is on-going.

For the ε -dependence in the spatial discretization error for the FDTD methods, i.e. $\frac{1}{\varepsilon}$ in front of h^2 , which was proven in Theorems 3.1-3.4, one can refer to Tab.2.8 in Chapter 2.

Based on the above comparison, in view of both temporal and spatial accuracies and resolution capacities, we conclude that the EWI-FP and TSFP methods

Table 3.3: Spatial and temporal error analysis of the SIFD2 method for the NLDE.

Spatial Errors	$h_0 = 1/8$	$h_0/2$	$h_0/2^2$	$h_0/2^3$	$h_0/2^4$
$\varepsilon_0 = 1$	8.15E-2	2.02E-2	5.00E-3	1.25E-3	3.12E-4
$\varepsilon_0/2$	9.29E-2	2.30E-2	5.73E-2	1.43E-3	3.58E-4
$\varepsilon_0/2^2$	9.91E-2	2.46E-2	6.12E-3	1.53E-3	3.82E-4
$\varepsilon_0/2^3$	9.89E-2	2.47E-2	6.17E-3	1.54E-3	3.85E-4
$\varepsilon_0/2^3$	9.87E-2	2.48E-2	6.18E-3	1.54E-3	3.83E-4
Temporal Errors	$\tau_0 = 0.1$	$\tau_0/8$	$\tau_0/8^2$	$\tau_0/8^3$	$\tau_0/8^4$
$\varepsilon_0 = 1$	<u>1.31E-1</u>	2.10E-3	3.27E-5	5.11E-7	7.98E-9
$\varepsilon_0/2$	1.28	<u>2.41E-2</u>	3.78E-4	5.91E-6	9.23E-8
$\varepsilon_0/2^2$	2.34	8.99E-1	<u>1.45E-2</u>	2.30E-4	3.61E-6
$\varepsilon_0/2^3$	2.46	2.94	8.19E-1	<u>1.30E-2</u>	2.04E-4
$\varepsilon_0/2^4$	2.79	3.15	4.28E-1	8.02E-1	<u>1.26E-2</u>

perform much better than the FDTD methods for the discretization of the Dirac equation, especially in the nonrelativistic limit regime. For the reader's convenience, we summarize the properties of different numerical methods in Tab. 3.7.

Table 3.4: Spatial and temporal error analysis of the CNFD method for the NLDE.

Spatial Errors	$h_0 = 1/8$	$h_0/2$	$h_0/2^2$	$h_0/2^3$	$h_0/2^4$
$\varepsilon_0 = 1$	8.15E-2	2.02E-2	5.00E-3	1.25E-3	3.12E-4
$\varepsilon_0/2$	9.29E-2	2.30E-2	5.73E-2	1.43E-3	3.58E-4
$\varepsilon_0/2^2$	9.91E-2	2.46E-2	6.12E-3	1.53E-3	3.82E-4
$\varepsilon_0/2^3$	9.89E-2	2.47E-2	6.17E-3	1.54E-3	3.85E-4
$\varepsilon_0/2^3$	9.87E-2	2.48E-2	6.18E-3	1.54E-3	3.83E-4
Temporal Errors	$\tau_0=0.1$	$\tau_0/8$	$\tau_0/8^2$	$\tau_0/8^3$	$\tau_0/8^4$
$\varepsilon_0 = 1$	<u>7.13E-2</u>	9.76E-4	1.52E-5	2.38E-7	3.65E-9
$\varepsilon_0/2$	4.58E-1	<u>7.75E-3</u>	1.21E-4	1.89E-6	2.95E-8
$\varepsilon_0/2^2$	1.74	2.34E-1	<u>3.86E-3</u>	6.01E-5	9.42E-7
$\varepsilon_0/2^3$	3.13	5.25E-1	2.07E-1	<u>3.49E-3</u>	5.46E-5
$\varepsilon_0/2^4$	2.34	1.84	8.16E-1	2.04E-1	<u>3.42E-3</u>

Table 3.5: Spatial and temporal error analysis of the EWI-FP method for the NLDE.

Spatial Errors	$h_0=2$	$h_0/2$	$h_0/2^2$	$h_0/2^3$	$h_0/2^4$	
$\varepsilon_0 = 1$	1.68	4.92E-1	4.78E-2	1.40E-4	2.15E-9	
$\varepsilon_0/2$	1.48	3.75E-1	1.57E-2	4.24E-5	6.60E-10	
$\varepsilon_0/2^2$	1.21	2.90E-1	4.66E-3	4.91E-6	6.45E-10	
$\varepsilon_0/2^3$	1.37	2.68E-1	2.40E-3	6.00E-7	6.34E-10	
$\varepsilon_0/2^4$	1.41	2.75E-1	1.84E-3	3.06E-7	6.13E-10	
$\varepsilon_0/2^5$	1.45	2.76E-1	1.74E-3	2.37E-7	5.98E-10	
Temporal Errors	$\tau_0=0.1$	$\tau_0/4$	$\tau_0/4^2$	$\tau_0/4^3$	$\tau_0/4^4$	$\tau_0/4^5$
$\varepsilon_0 = 1$	<u>1.62E-1</u>	8.75E-3	5.44E-4	3.40E-5	2.12E-6	1.33E-7
$\varepsilon_0/2$	2.02	<u>2.58E-2</u>	1.59E-3	9.94E-5	6.21E-6	3.88E-7
$\varepsilon_0/2^2$	2.11	2.11	<u>1.12E-2</u>	6.94E-4	4.33E-5	2.71E-6
$\varepsilon_0/2^3$	2.12	2.12	1.52E-1	<u>8.88E-3</u>	5.53E-4	3.45E-5
$\varepsilon_0/2^4$	2.06	2.06	2.06	1.40E-1	<u>8.24E-3</u>	5.13E-4
$\varepsilon_0/2^5$	2.09	2.03	2.03	2.03	1.36E-1	<u>8.01E-3</u>

Table 3.6: Spatial and temporal error analysis of the TSFP method for the NLDE.

Spatial Errors	$h_0=2$	$h_0/2$	$h_0/2^2$	$h_0/2^3$	$h_0/2^4$	
$\varepsilon_0 = 1$	1.68	4.92E-1	4.78E-2	1.40E-4	2.15E-9	
$\varepsilon_0/2$	1.48	3.75E-1	1.57E-2	4.24E-5	6.60E-10	
$\varepsilon_0/2^2$	1.21	2.90E-1	4.66E-3	4.91E-6	6.45E-10	
$\varepsilon_0/2^3$	1.37	2.68E-1	2.40E-3	6.00E-7	6.34E-10	
$\varepsilon_0/2^4$	1.41	2.75E-1	1.84E-3	3.06E-7	6.13E-10	
$\varepsilon_0/2^5$	1.45	2.76E-1	1.74E-3	2.37E-7	5.98E-10	
Temporal Errors	$\tau_0=0.4$	$\tau_0/4$	$\tau_0/4^2$	$\tau_0/4^3$	$\tau_0/4^4$	$\tau_0/4^5$
$\varepsilon_0 = 1$	<u>1.60E-1</u>	9.56E-3	5.95E-4	3.72E-5	2.32E-6	1.54E-7
$\varepsilon_0/2$	8.94E-1	<u>3.91E-2</u>	2.40E-3	1.50E-4	9.35E-6	5.85E-7
$\varepsilon_0/2^2$	2.60	2.18E-1	<u>1.06E-2</u>	6.56E-4	4.09E-5	2.56E-6
$\varepsilon_0/2^3$	2.28	2.33	2.48E-2	<u>2.58E-3</u>	1.60E-4	9.98E-6
$\varepsilon_0/2^4$	1.46	1.28	1.30	1.15E-2	<u>6.19E-4</u>	3.84E-5
$\varepsilon_0/2^5$	1.52	3.27E-1	4.06E-1	4.13E-1	2.83E-3	<u>1.53E-4</u>

Table 3.7: Comparison of properties of different numerical methods for solving the NLDE with M being the number of grid points in space.

Method	LFFD	SIFD1	SIFD2
Mass conservation	No	No	No
Energy conservation	No	No	No
Unconditionally stable	No	No	Yes
Explicit scheme	Yes	No	No
Temporal accuracy	2nd	2nd	2nd
Spatial accuracy	2nd	2nd	2nd
Memory cost	$O(M)$	$O(M)$	$O(M)$
Computational cost	$O(M)$	$O(M)$	$> O(M)$
Resolution	$h = O(\sqrt{\varepsilon})$	$h = O(\sqrt{\varepsilon})$	$h = O(\sqrt{\varepsilon})$
when $0 < \varepsilon \ll 1$	$\tau = O(\varepsilon^3)$	$\tau = O(\varepsilon^3)$	$\tau = O(\varepsilon^3)$
Method	CNFD	EWI-FP	TSFP
Mass conservation	Yes	No	Yes
Energy conservation	Yes	No	No
Unconditionally stable	Yes	No	Yes
Explicit scheme	No	Yes	Yes
Temporal accuracy	2nd	2nd	2nd
Spatial accuracy	2nd	Spectral	Spectral
Memory cost	$O(M)$	$O(M)$	$O(M)$
Computational cost	$\gg O(M)$	$O(M \ln M)$	$O(M \ln M)$
Resolution	$h = O(\sqrt{\varepsilon})$	$h = O(1)$	$h = O(1)$
when $0 < \varepsilon \ll 1$	$\tau = O(\varepsilon^3)$	$\tau = O(\varepsilon^2)$	$\tau = O(\varepsilon^2)$

Fourth order compact time splitting methods

In this chapter, we present a series of fourth order compact splitting methods with purely positive coefficients for solving the time-dependent Dirac equation. For the simplicity of notations, we only consider the 1D Dirac equation

$$i\partial_t\Phi(t, x) = \left[-\frac{i}{\varepsilon}\sigma_1\partial_x + \frac{1}{\varepsilon^2}\sigma_3 + V(t, x)I_2 - A_1(t, x)\sigma_1\right]\Phi(t, x), \quad (4.0.1)$$

while all the results can be easily achieved by some careful calculations.

4.1 Introduction

Our early work [15] shows that even though the second order time splitting Fourier pseudospectral (TSFP) method performs better than time domain finite difference (TDFD) methods, the scalability for TSFP is still dependent on ε , i.e. to obtain an accurate numerical solution, the time step for the Dirac equation in the nonrelativistic limit regime with small ε should also be very small. One possible improvement is to use a higher order variation of time splitting method [46, 98, 99] which has the advantage of being unitary, remain applicable to higher dimensions and easily generalizable to even higher orders. The disadvantage is that the time step size needed for

convergence seemed to be small, and many iterations are required for evolving the system forward in time which will increase the computational burden. In Chin and Chen's work [27] they show that the method of factorizing the evolution operator to fourth order with purely positive coefficients produces algorithms capable of solving the time-dependent Schrödinger equation with time step 5 to 10 times as large as before.

The relativistic quantum state of spin $-\frac{1}{2}$ is evolved forward by the Dirac evolution operator

$$e^{-i\tau H} = e^{-i\tau(T+W)}, \quad (4.1.1)$$

where $T = -\frac{i}{\varepsilon}\sigma_1\partial_x + \frac{1}{\varepsilon^2}\sigma_3$, and $W = V(t, \mathbf{x})I_2 - A_1(t, \mathbf{x})\sigma_1$ are the kinetic and potential energy operators, respectively. In the split operator approach, with the help of Taylor expansion, the short time evolution operator (4.1.1) is factorized to second order in the product form

$$\mathcal{T}^{(2)}(\tau) \equiv e^{-i\frac{\tau}{2}W} e^{-i\tau T} e^{-i\frac{\tau}{2}W} = e^{-i\tau(T+W)+\tau^3 C+\dots}, \quad (4.1.2)$$

where the error term is indicated as $\tau^3 C$. Thus $\mathcal{T}^{(2)}(\tau)$ evolves the system according to the Hamiltonian

$$H^{(2)} = T + W + \tau^2 C + \dots, \quad (4.1.3)$$

which deviates from the original Hamiltonian by an error term of second order in τ . Every occurrence of $e^{-i\tau T}$ requires two fast Fourier transforms (FFTs), one direct and one inverse.

The higher order split operator approach can be easily constructed. For example, we can decompose the evolution operator in the form [31, 94, 108]

$$e^{-i\tau(T+W)} = \Pi_i e^{-ia_i\tau T} e^{-ib_i\tau W} \quad (4.1.4)$$

with coefficients $\{a_i, b_i\}$ determined by the required order of accuracy.

However, any factorization of form (4.1.4) beyond second order must produce some negative coefficients in the set $\{a_i, b_i\}$, corresponding to some steps in which the system is evolved backward in time [95]. [27] shows that for Schrödinger evolution

operator the resulting higher order algorithms converge only for very small values of τ and are far from optimal [29]. In this section, we show that insisting on decomposing the Dirac evolution operator to fourth order with purely positive time steps yield algorithms with good convergent properties at large time steps.

4.2 Fourth order compact time splitting

The Forest-Ruth (FR) scheme [42] is a typical example of fourth order splitting with negative coefficients

$$\mathcal{T}_{FR}^{(4)} = \mathcal{T}^{(2)}(\tilde{\tau})\mathcal{T}^{(2)}(-s\tilde{\tau})\mathcal{T}^{(2)}(\tilde{\tau}) \quad (4.2.1)$$

where $s = 2^{1/3}$ is chosen to cancel the $\tau^3 C$ error term to obtain higher order accuracy and $\tilde{\tau} = \tau/(2 - s)$ rescales the sum of forward-backward-forward time steps back to τ . Identical construction can be applied to generate a $(n + 2)$ th order algorithm $\mathcal{T}^{(n+2)}$ from a triplet product of $\mathcal{T}^{(n)}$,

$$\mathcal{T}^{(n+2)} = \mathcal{T}^{(n)}(\tilde{\tau})\mathcal{T}^{(n)}(-s\tilde{\tau})\mathcal{T}^{(n)}(\tilde{\tau}) \quad (4.2.2)$$

where $s = 2^{1/(n+2)}$. The FR algorithm above requires six FFTs, while the alternative one with interchanged operators T and W would have required eight FFTs per iteration.

Recently, a number of fourth order splitting schemes with only positive coefficients for Schrödinger evolution operator has derived by Chin [25] and Suzuki [94]. Similarly, for Dirac evolution operator (4.1.1), an additional operator is required

$$[W, [T, W]] = -\frac{4}{\varepsilon^2} A_1^2 \sigma_3. \quad (4.2.3)$$

Proof: $T = -\frac{i}{\varepsilon}\sigma_1\partial_x + \frac{1}{\varepsilon^2}\sigma_3$, $W = V(t, x)I_2 - A_1(t, x)\sigma_1$, and it is easy to observe

that

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

Also we have

$$\begin{aligned}
& \left[W, \left[-\frac{i}{\varepsilon} \sigma_1 \partial_x, W \right] \right] \\
& = 2(I_2 - A_1 \sigma_1) \left(-\frac{i}{\varepsilon} \sigma_1 \partial_x \right) (VI_2 - A_1 \sigma_1) \\
& \quad - (VI_2 - A_1 \sigma_1)^2 \left(-\frac{i}{\varepsilon} \sigma_1 \partial_x \right) - \left(-\frac{i}{\varepsilon} \sigma_1 \partial_x \right) (VI_2 - A_1 \sigma_1)^2 \\
& = 2(VI_2 - A_1 \sigma_1) (\partial_x VI_2 - \partial_x A_1 \sigma_1) \left(-\frac{i}{\varepsilon} \sigma_1 \right) \\
& \quad - \left(-\frac{i}{\varepsilon} \sigma_1 \right) 2(VI_2 - A_1 \sigma_1) (\partial_x VI_2 - \partial_x A_1 \sigma_1) \\
& = 0
\end{aligned}$$

and

$$\begin{aligned}
& \left[W, \left[\frac{1}{\varepsilon^2} \sigma_3, W \right] \right] \\
& = 2(VI_2 - A_1 \sigma_1) \frac{1}{\varepsilon^2} \sigma_3 (VI_2 - A_1 \sigma_1) - (VI_2 - A_1 \sigma_1)^2 \frac{1}{\varepsilon^2} \sigma_3 - \frac{1}{\varepsilon^2} \sigma_3 (VI_2 - A_1 \sigma_1)^2 \\
& = 2(VI_2 - A_1 \sigma_1) (VI_2 + A_1 \sigma_1) \frac{1}{\varepsilon^2} \sigma_3 - (VI_2 - A_1 \sigma_1)^2 \frac{1}{\varepsilon^2} \sigma_3 - (VI_2 + A_1 \sigma_1)^2 \frac{1}{\varepsilon^2} \sigma_3 \\
& = -\frac{4}{\varepsilon^2} A_1^2 \sigma_1^2 \sigma_3 \\
& = -\frac{4}{\varepsilon^2} A_1^2 \sigma_3.
\end{aligned}$$

Conclude all above, we can get (4.2.3).

□

With the help of this additional operator, we can obtain several different fourth order splitting schemes with purely positive coefficients.

Scheme A.

$$\mathcal{T}_A^{(4)} \equiv e^{-i\frac{\tau}{6}W} e^{-i\frac{\tau}{2}T} e^{-i\frac{2\tau}{3}\widetilde{W}} e^{-i\frac{\tau}{2}T} e^{-i\frac{\tau}{6}W}, \quad (4.2.4)$$

Scheme B.

$$\mathcal{T}_B^{(4)} \equiv e^{-i\frac{\tau}{2}[1-(1/\sqrt{3})]T} e^{-i\frac{\tau}{2}\overline{W}} e^{-i\frac{\tau}{\sqrt{3}}T} e^{-i\frac{\tau}{2}\overline{W}} e^{-i\frac{\tau}{2}[1-(1/\sqrt{3})]W}, \quad (4.2.5)$$

Scheme C.

$$\mathcal{T}_C^{(4)} \equiv e^{-i\frac{\tau}{6}T} e^{-i\frac{3}{8}\tau W} e^{-i\frac{\tau}{3}T} e^{-i\frac{\tau}{4}\widetilde{W}} e^{-i\frac{\tau}{3}T} e^{-i\frac{3}{8}\tau W} e^{-i\frac{\tau}{6}T}, \quad (4.2.6)$$

Scheme D.

$$\mathcal{T}_D^{(4)} \equiv e^{-i\frac{\tau}{8}\widetilde{W}} e^{-i\frac{\tau}{3}T} e^{-i\frac{3}{8}\tau W} e^{-i\frac{\tau}{3}T} e^{-i\frac{3}{8}\tau W} e^{-i\frac{\tau}{3}T} e^{-i\frac{\tau}{8}\widetilde{W}}, \quad (4.2.7)$$

with \widetilde{W} given by

$$\widetilde{W} = W - \frac{1}{48}\tau^2[W, [T, W]] \quad (4.2.8)$$

and \overline{W} given by

$$\overline{W} = W - \frac{1}{24}(2 - \sqrt{3})\tau^2[W, [T, W]]. \quad (4.2.9)$$

At present, no sixth order factorization with positive coefficients are known. However, one can use the triple construction (4.2.2) to build a sixth order algorithm by iterating on three fourth order algorithms.

4.3 Extension to the time-dependent potential

For $H(t)$ a time-dependent operator, the evolution equation

$$i\partial_t\psi(t) = H(t)\Psi(t), \quad (4.3.1)$$

has an operator solution

$$\Psi(t + \tau) = Q \left(\exp \left(-i \int_t^{t+\tau} H(s) ds \right) \right) \Psi(t). \quad (4.3.2)$$

$Q\left(\exp\left(-i\int_t^{t+\tau} H(s)ds\right)\right)$ is called the time-ordered operator. More generally, the ordered exponential operator is defined as

$$OE[H](t) = Q\left(\exp\left(-i\int_0^t H(s)ds\right)\right) = \sum_{n=0}^{\infty} \frac{i^n}{n!} \int_0^t \dots \int_0^t \mathcal{Q}\{H(s_1)\dots H(s_n)\} ds_1 \dots ds_n, \quad (4.3.3)$$

where \mathcal{Q} is a higher order operation that ensures the exponential is time-ordered, i.e. any product of $h(t)$ that occurs in the expansion of the exponential must be ordered such that the values of t are increasing from right to left.

The ordered exponential is the unique solution of the initial value problem

$$\begin{aligned} i\frac{d}{dt}OE[H](t) &= H(t)OE[H](t) \\ OE[H](0) &= I \end{aligned} \quad (4.3.4)$$

So the ordered exponential is the solution to the integral equation

$$OE[H](t) = I - i\int_0^t H(s)OE[H](s)ds. \quad (4.3.5)$$

Expand this iteratively, we can obtain the conventional expansion for ordered exponential

$$OE[H](t) = 1 - i\left(\int_0^t H(s)ds + \int_0^t \int_0^{s_1} H(s_1)H(s_2)ds_2ds_1 + \dots\right). \quad (4.3.6)$$

Another intuitive interpretation is

$$\begin{aligned} \mathcal{Q}\left(\exp\left(-i\int_t^{t+\tau} H(s)ds\right)\right) &= \lim_{n\rightarrow\infty} \mathcal{Q}\left(e^{-i\tau/n\sum_{i=1}^n H(t+i(\tau/n))}\right) \\ &= \lim_{n\rightarrow\infty} e^{(-i\tau/n)H(t+\tau)} \dots e^{(-i\tau/n)H(t+2\tau/n)} e^{(-i\tau/n)H(t+\tau/n)}. \end{aligned} \quad (4.3.7)$$

There are many ways of solving the time-ordered problem. Here we follow the Suzuki's method [96], which directly implements time ordering without any additional formalism or auxiliary variables.

Let's denote the forward time operator

$$D = i\frac{\overrightarrow{\partial}}{\partial t}, \quad (4.3.8)$$

such that for any two time-dependent functions $F(t)$ and $G(t)$,

$$F(t)E^{-i\tau D}G(t) = F(t + \tau)G(t). \quad (4.3.9)$$

Suzuki's method proved that [96]

$$Q\left(\exp\left(-i\int_t^{t+\tau} H(s)ds\right)\right) = \exp[-i\tau(H(t) + D)]. \quad (4.3.10)$$

Proof: Using Trotter formula

$$e^{x(A+B)} = \lim_{n \rightarrow \infty} \left(e^{\frac{x}{n}A} e^{\frac{x}{n}B}\right)^n \quad (4.3.11)$$

Then we can have

$$\begin{aligned} \exp[-i\tau(H(t) + D)] &= \lim_{n \rightarrow \infty} [e^{-i\frac{\tau}{n}H(t)} e^{-i\frac{\tau}{n}D}]^n \\ &= \lim_{n \rightarrow \infty} e^{-i\frac{\tau}{n}H(t+\tau)} \dots e^{-i\frac{\tau}{n}H(t+2\tau/n)} e^{-i\frac{\tau}{n}H(t+\tau/n)} \\ &= Q\left(\exp\left(-i\int_t^{t+\tau} H(s)ds\right)\right). \end{aligned}$$

□

For the widely used case of $H(t) = T + W(t)$, where only one of the operators is explicitly dependent on time, the short time evolution can be written as

$$\Psi(t + \tau) = e^{-i\tau[\tilde{T} + W(t)]}\Psi(t), \quad (4.3.12)$$

which is just like the time-independent operator case but with an effective $\tilde{T} = T + D$.

This suggests

1. Decompose $e^{-i\tau[\tilde{T} + W(t)]}$ into $e^{-i\tau\tilde{T}}$ and $e^{-i\tau W(t)}$ using any factorization scheme applicable in the time-independent case.

2. Since $[T, D] = 0$ (T is time-independent),

$$e^{-i\tau\tilde{T}} = e^{-i\tau D} e^{-i\tau T} \quad (4.3.13)$$

incorporate all time-dependent requirements by applying (4.3.9). For example, a second order factorization gives

$$\begin{aligned} \mathcal{T}_A^{(2)} &= e^{-i\tau/2\tilde{T}} e^{-i\tau/2W(t)} e^{-i\tau/2\tilde{T}} \\ &= e^{-i\tau/2D} e^{-i\tau/2T} e^{-i\tau/2W(t)} e^{-i\tau/2D} e^{-i\tau/2T} \\ &= e^{-i\tau/2T} e^{-i\tau/2W(t+\tau/2)} e^{-i\tau/2T} \end{aligned} \quad (4.3.14)$$

which is well-known as midpoint algorithm for time-dependent problems. Another one is

$$\begin{aligned}
\mathcal{T}_B^{(2)} &= e^{-i\tau/2W(t)} e^{-i\tau/2\tilde{T}} e^{-i\tau/2W(t)} \\
&= e^{-i\tau/2W(t)} e^{-i\tau D} e^{-i\tau T} e^{-i\tau/2W(t)} \\
&= e^{-i\tau/2W(t+\tau)} e^{-i\tau T} e^{-i\tau/2W(t)}
\end{aligned} \tag{4.3.15}$$

Thus, for $H(t) = T + W(t)$, the effect of time ordering is to increment the time dependence of each potential operator $W(t)$ by the same of the time steps of all the T operators to its right.

For the Dirac equation with a time-dependent potential, the wave function evolves forward in a short time τ by

$$\Psi(\tau) = e^{-i\tau[\tilde{T}+W(t)]}\Psi(0) \tag{4.3.16}$$

where $\tilde{T} = T + D$. Thus the second order algorithms for solving the Dirac equation with time step size τ can be described simply as

$$\mathcal{T}_A^{(2)}(\tau) = e^{-i\tau/2T} e^{-i\tau/2W(\tau/2)} e^{-i\tau/2T}, \tag{4.3.17}$$

$$\mathcal{T}_B^{(2)}(\tau) = e^{-i\tau/2W(\tau)} e^{-i\tau/2T} e^{-i\tau/2W(0)}, \tag{4.3.18}$$

Following two-step approach, we can transcribe any time-dependent factorization algorithm into a time-independent algorithm. For example,

$$\mathcal{T}_A^{(4)} = e^{-i\tau/6W(\tau)} e^{-i\tau/2T} e^{-2i\tau/3\tilde{W}(\tau/2)} e^{-i\tau/2T} e^{-i\tau/6W(0)} \tag{4.3.19}$$

with \tilde{W} defined by

$$\begin{aligned}
\tilde{W}(t) &= W(t) - \frac{1}{48}\tau^2[W(t), [\tilde{T}, W(t)]] \\
&= W(t) - \frac{1}{48}\tau^2[W(t), [T, W(t)]]
\end{aligned} \tag{4.3.20}$$

since $[W(t), [D, W(t)]] = 0$, there is no additional complication to the additional operator term caused by the forward time derivative D.

The other fourth order algorithms are

$$\mathcal{T}_B^{(4)} = e^{-it_2\tau T} e^{-i\tau/2\overline{W}(a_2\tau)} e^{-it_1\tau T} e^{-i\tau/2\overline{W}(a_1\tau)} e^{-it_0\tau T} \quad (4.3.21)$$

where

$$\begin{aligned} t_0 = t_2 &= \frac{1}{2} \left(1 - \frac{1}{\sqrt{3}} \right), & t_1 &= \frac{1}{\sqrt{3}}, \\ a_1 &= \frac{1}{2} \left(1 - \frac{1}{\sqrt{3}} \right), & a_2 &= \frac{1}{2} \left(1 + \frac{1}{\sqrt{3}} \right), \end{aligned}$$

and with \overline{W} given by

$$\overline{W}(t, \mathbf{x}) = W(t, \mathbf{x}) - \frac{1}{24}(2 - \sqrt{3})\tau^2[W(t), [T, W(t)]], \quad (4.3.22)$$

The time-dependent forms of Scheme C and Scheme D are, respectively

$$\mathcal{T}_C^{(4)}(\tau) = e^{-i\tau/6T} e^{-i3\tau/8W(5\tau/6)} e^{-i\tau/3T} e^{-i\tau/4\widetilde{W}(\tau/2)} e^{-i\tau/3T} e^{-i3\tau/8W(\tau/6)} e^{-i\tau/6T}, \quad (4.3.23)$$

$$\mathcal{T}_D^{(4)}(\tau) = e^{-i\tau/8\widetilde{W}(\tau)} e^{-i\tau/3T} e^{-i3\tau/8W(2\tau/3)} e^{-i\tau/3T} e^{-i3\tau/8W(\tau/3)} e^{-i\tau/3T} e^{-i\tau/8\widetilde{W}(0)}, \quad (4.3.24)$$

4.4 Numerical examples

In this section, we compare the accuracy of several different 4th order splitting operator methods including the FR method (4.2.1), Scheme A to Scheme D (4.2.4)-(4.2.7), for the Dirac equation (4.0.1) in terms of the mesh size h , time step τ and the parameter $0 < \varepsilon \leq 1$. We will focus on the convergence rates and comparing largest convergent steps of different methods.

To test the accuracy, we choose the electromagnetic potentials in the Dirac equation (4.0.1) same as in Chapter 2

$$A_1(t, x) = \frac{(x+1)^2}{1+x^2}, \quad V(t, x) = \frac{1-x}{1+x^2}, \quad x \in \mathbb{R}, \quad t \geq 0,$$

and the same initial value

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

The problem is solved numerically on an interval $\Omega = (-16, 16)$, i.e. $a = -16$ and $b = 16$, with periodic boundary conditions on $\partial\Omega$. The ‘exact’ solution $\Phi(t, x) = (\phi_1(t, x), \phi_2(t, x))^T$ is obtained numerically by using the TSFP method with a very fine mesh size and a small time step, e.g. $h_e = 1/16$ and $\tau_e = 10^{-7}$ to compare with the numerical solutions obtained by the fourth order splitting operator methods mentioned above. We use the same discrete l_2 norm as before.

Fig. 4.1 and Fig. 4.2 show the resulting computational error of different splitting operator numerical methods for the Dirac equation (4.0.1) with $\varepsilon = \frac{1}{4}$ and $\varepsilon = \frac{1}{16}$, respectively. We can draw the following conclusions:

(i). For the discretization error in time, for any fixed $\varepsilon = \varepsilon_0 > 0$, the TSSP method is second-order accurate, while all the fourth order splitting operator methods are fourth order accurate (cf. the slope of the lines in Fig. 4.1-Fig. 4.2). Among all the fourth order splitting operator methods, the compact splitting with purely positive coefficient ones, i.e. Scheme A to Scheme D, can produce converged results of conventional fourth order algorithms using time step 4 to 8 times as large, which means that the fourth order error of these new algorithms are roughly 300 times smaller than that of fourth order algorithms with negative coefficients, such as the traditional FR method.

(ii). From Fig. 4.1 to Fig. 4.2 we find out that as ε goes small, the new splitting operator methods with only positive coefficients show fourth order convergence in a much larger time step size than the FR method does.

(iii). Among all the new fourth order splitting methods, it is remarkably notable that Scheme A only requires four FFTs, Scheme B and Scheme D require six FFTs while Scheme C has four T operators corresponding to eight FFTs. Another important observation is that Scheme C and Scheme D give almost identical results which are better than the other two. Obviously then, Scheme D is preferable with two fewer FFTs than Scheme C.

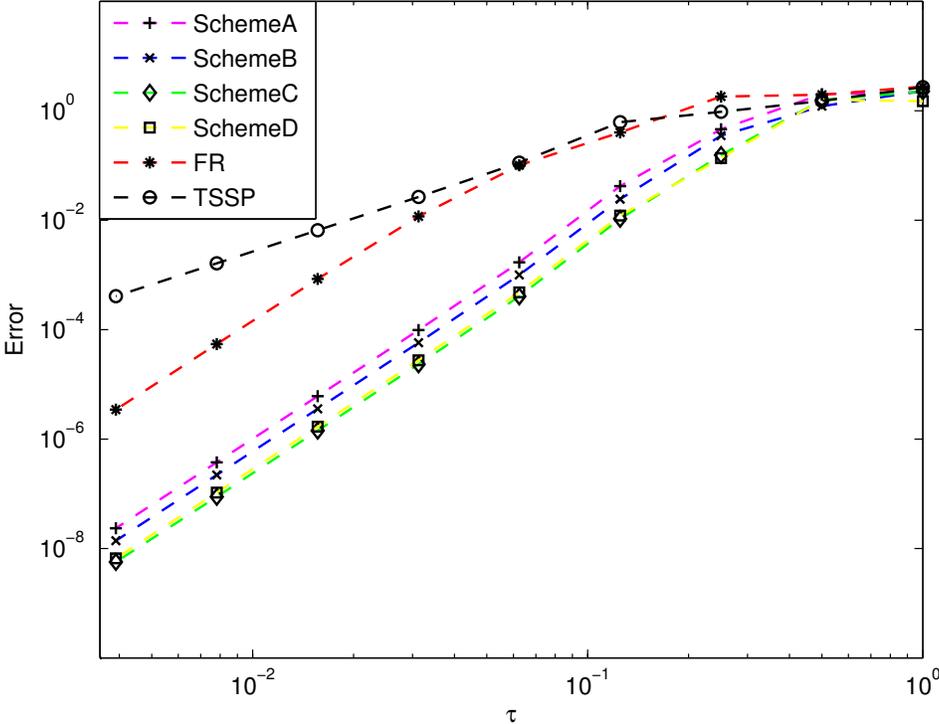


Figure 4.1: The computational error of the Dirac equation using different splitting operator methods with $\varepsilon = \frac{1}{4}$.

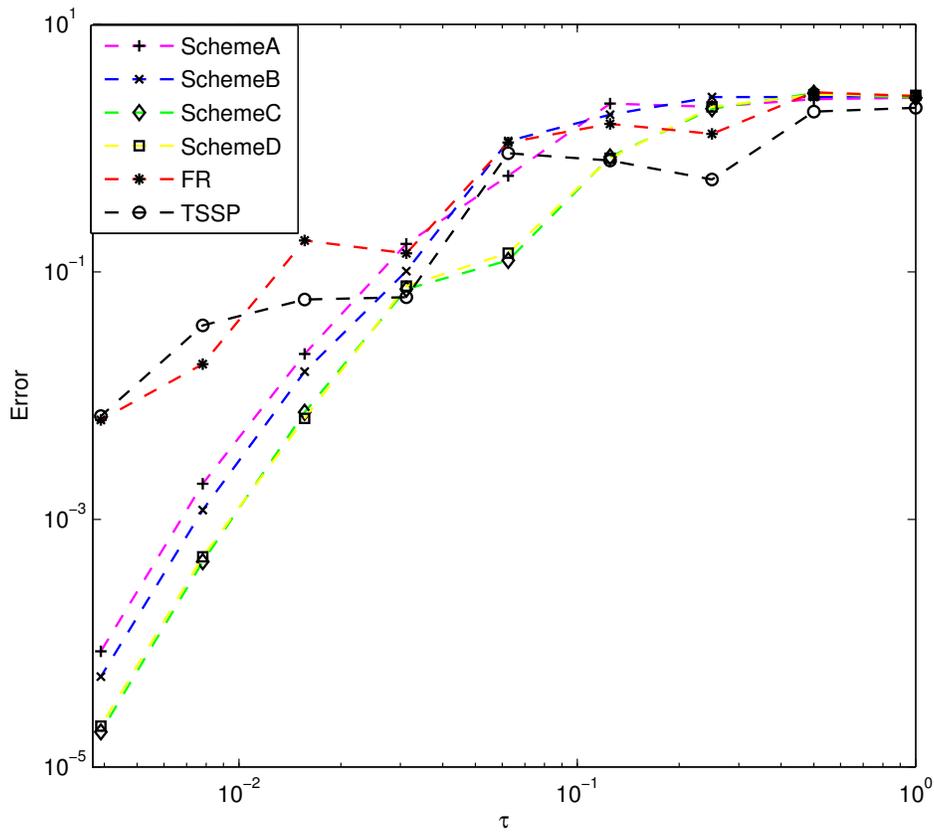


Figure 4.2: The computational error of the Dirac equation using different splitting operator methods with $\varepsilon = \frac{1}{16}$.

A uniformly accurate multiscale method

In this chapter, based on the spectral decomposition, a second order multiscale time integrator Fourier pseudospectral (MTI-FP) method is designed. A rigorous error analysis is proposed to show that this is a uniformly accurate method. Numerical results are also shown at last of this chapter to illustrate the uniform accuracy.

5.1 Introduction

After proper nondimensionlization and dimension reduction, the Dirac equation for the spin-1/2 particles with external electromagnetic potential in d dimensions ($d = 1, 2, 3$) reads [7, 15, 23, 32, 33, 58, 58, 78, 85]:

$$\begin{cases} i\partial_t\Psi(t, \mathbf{x}) = \left[-\frac{i}{\varepsilon} \sum_{j=1}^d \alpha_j \partial_j + \frac{1}{\varepsilon^2} \beta + V(t, \mathbf{x}) I_4 - \sum_{j=1}^d A_j(t, \mathbf{x}) \sigma_j \right] \Psi(t, \mathbf{x}), \\ \Psi(t = 0, \mathbf{x}) = \Psi_0(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^d, \end{cases} \quad (5.1.1)$$

where $\Psi^* = \overline{\Psi^T}$ denotes the conjugate transpose of Ψ . In the nonrelativistic limit, i.e. $\varepsilon \rightarrow 0^+$, as proven in [20, 26, 58, 67, 71, 73, 85, 105], the solution Ψ to the Dirac

equation (5.1.1) will split into the electron part and the positron part, i.e.,

$$\Psi = e^{-it/\varepsilon^2} \begin{pmatrix} \varphi_1 \\ \varphi_2 \\ 0 \\ 0 \end{pmatrix} + e^{it/\varepsilon^2} \begin{pmatrix} 0 \\ 0 \\ \varphi_3 \\ \varphi_4 \end{pmatrix} + O(\varepsilon) = e^{-it/\varepsilon^2} \Phi_e + e^{it\varepsilon^2} \Phi_p + O(\varepsilon), \quad (5.1.2)$$

where both the ‘electron component’ Φ_e and the ‘positron component’ Φ_p satisfy the (different) Schrödinger equation [20, 85]. In addition, a higher order $O(\varepsilon^2)$ approximate model is provided by Pauli’s equation [67, 71] and we refer the readers to the references [20, 67, 71, 85] and references therein for details on the nonrelativistic limit of the Dirac equation (5.1.1). In practice, for lower dimensions $d = 1, 2$, the Dirac equation (5.1.1) consists of two equivalent sets of decoupled equations [15] and thus can be reduced to the following equation for $\Phi := \Phi(t, \mathbf{x}) = (\phi_1(t, \mathbf{x}), \phi_2(t, \mathbf{x}))^T \in \mathbb{C}^2$ as

$$\begin{cases} i\partial_t \Phi(t, \mathbf{x}) = \left[-\frac{i}{\varepsilon} \sum_{j=1}^d \sigma_j \partial_j + \frac{1}{\varepsilon^2} \sigma_3 + V(t, \mathbf{x}) I_2 - \sum_{j=1}^d A_j(t, \mathbf{x}) \sigma_j \right] \Phi(t, \mathbf{x}), \\ \Phi(t=0, \mathbf{x}) = \Phi_0(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^d, \quad d = 1, 2, \end{cases} \quad (5.1.3)$$

where $\Phi = (\psi_1, \psi_4)^T$ (or $\Phi = (\psi_2, \psi_3)^T$ in 1D and under the transformation $x_2 \rightarrow -x_2$ and $A_2 \rightarrow -A_2$ in 2D). As a result of its simplicity compared to (5.1.1), the above Dirac equation (5.1.3) has been widely used when considering the 1D and 2D cases [15, 102, 107]. For the nonrelativistic limit as $\varepsilon \rightarrow 0^+$, the same limit model as (5.1.2) can be obtained for the Dirac equation (5.1.1) and we omit the details here for brevity.

There have been extensive theoretical and numerical results for the Dirac equation (5.1.1) in literatures. Along the analytical front, time independent properties and dynamical properties have been thoroughly investigated, such as the bound states [37], semi-classical limit [45] and nonrelativistic limit [20, 73, 85], etc. Along the numerical front, various finite difference time domain (FDTD) methods [10, 52, 53, 102, 107], time-splitting Fourier spectral (TSFP) methods [19, 57] and Gaussian beam method [106] have been proposed to solve the Dirac equations

(5.1.1) and (5.1.3). However, the most existing numerical methods are designed for the Dirac equations (5.1.1) or (5.1.3) in the parameter regime $\varepsilon = O(1)$. In fact, for the Dirac equation in the nonrelativistic limit, i.e. $0 < \varepsilon \ll 1$, based on the theoretical analysis [20, 26, 58, 67, 71, 73, 85, 105], the solution exhibits propagating waves with wavelength $O(\varepsilon^2)$ and $O(1)$ in time and space, respectively. This rapid oscillation in time brings significant difficulties in designing and analyzing the numerical methods for the Dirac equation (5.1.1) or (5.1.3) when $0 < \varepsilon \ll 1$. In our recent work [15], we have rigorously analyzed and compared the frequently used FDTD methods and TSFP methods in the nonrelativistic limit and shown that the meshing strategy for FDTD methods and TSFP methods should be $h = O(\sqrt{\varepsilon})$, $\tau = O(\varepsilon^3)$ and $h = O(1)$, $\tau = O(\varepsilon^2)$, respectively, where h is the mesh size and τ is the time step. Thus, the existing FDTD and TSFP methods are capable to solve Dirac equation (5.1.1) or (5.1.3) efficiently and accurately in the regime $\varepsilon = O(1)$, and are less efficient and time consuming in the regime $0 < \varepsilon \ll 1$. Our aim is to propose and analyze a uniformly accurate numerical methods for the Dirac equation (5.1.1) w.r.t $\varepsilon \in (0, 1]$.

Recently, a uniformly accurate multiscale time integrator Fourier pseudospectral method was successfully designed and rigorously analyzed for the Klein-Gordon equation in the non-relativistic limit [16]. The key ingredients included a multiscale decomposition of the exact solution [68] and the Gautschi-type exponential wave integrator (EWI), which has been widely explored in highly oscillatory ordinary differential equations (ODEs) [44, 51, 55] and dispersive partial differential equations (PDEs) [13, 17, 18]. For Dirac equation (5.1.1) or (5.1.3) in the nonrelativistic limit, based on a suitable combination of the multiscale decomposition of the exact solution and EWI, we propose a multiscale time integrator Fourier pseudospectral (MTI-FP) method with uniform spectral accuracy in space. The proposed MTI-FP possesses the error bounds $O(\tau^2/\varepsilon^2)$ and $O(\tau^2 + \varepsilon^2)$ in time, which shows that MTI-FP converges in time uniformly at linear rate w.r.t. $\varepsilon \in (0, 1]$ and optimally at quadratic rate when $\varepsilon = O(1)$ or $0 < \varepsilon \leq \tau$. In addition, the MTI-FP is explicit,

efficient and easy to implement.

5.2 Multiscale decomposition

For simplicity of notations, we only consider Dirac equation (5.1.3), while all the notations and results can be easily generalized to (5.1.1). We denote the d -dimensions ($d = 1, 2$) Dirac equation in the form

$$i\partial_t\Phi(t, \mathbf{x}) = \frac{1}{\varepsilon^2}T^\varepsilon\Phi(t, \mathbf{x}) + W(t, \mathbf{x})\Phi(t, \mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^d, \quad (5.2.1)$$

where the wave function vector $\Phi(t, \mathbf{x}) = (\phi_1(t, \mathbf{x}), \phi_2(t, \mathbf{x}))^T \in \mathbb{C}^2$. T^ε is the "free Dirac operator"

$$T^\varepsilon = -\varepsilon i \sum_{j=1}^d \sigma_j \partial_j + \sigma_3, \quad (5.2.2)$$

and $W \equiv W(t, \mathbf{x})$ is the "electromagnetic operator"

$$W(t, \mathbf{x}) = V(t, \mathbf{x})I_2 - \sum_{j=1}^d A_j(t, \mathbf{x})\sigma_j, \quad (5.2.3)$$

and $\sigma_1, \sigma_2, \sigma_3$ are the Pauli matrices.

The initial condition for (5.1.3) is given as

$$\Phi(t = 0, \mathbf{x}) = \Phi_0(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^d. \quad (5.2.4)$$

Considering the spectral problem for the operator T , it is diagonalizable in the phase space and can be decomposed as

$$T^\varepsilon = \sqrt{1 - \varepsilon^2\Delta} \Pi_+ - \sqrt{1 - \varepsilon^2\Delta} \Pi_-, \quad (5.2.5)$$

where Π_+ and Π_- are pseudodifferential projectors defined as

$$\Pi_+ = \frac{1}{2} \left[I_2 + (1 - \varepsilon^2\Delta)^{-1/2} T \right], \quad \Pi_- = \frac{1}{2} \left[I_2 - (1 - \varepsilon^2\Delta)^{-1/2} T \right], \quad (5.2.6)$$

It is easy to verify that $\Pi_+ + \Pi_- = I_2$ and $\Pi_+\Pi_- = \Pi_-\Pi_+ = \mathbf{0}$, $\Pi_\pm^2 = \Pi_\pm$.

In order to design a uniformly accurate numerical method for the Dirac equation (5.1.1) or (5.1.3), from the experience in the uniformly accurate methods for Klein-Gordon equation in the nonrelativistic limit [16], recalling that there exist propagating waves with $O(\varepsilon^2)$ wavelength in time, a multiscale decomposition should possess $O(\varepsilon^2)$ accuracy, so that the first order time derivative of the residue is bounded and a uniformly accurate scheme can be obtained. Thus, the first order Schrödinger decomposition (5.1.2) is inappropriate and the second order model, Pauli's equation (see [20, 67, 73]), might work. However, due to the linearity of the Dirac equation (5.3.1) (or (5.1.1), (5.1.3)), we have a direct and better decomposition by applying the projectors Π_{\pm} to the equation [20].

Choose time step $\tau := \Delta t$ and denote time steps as $t_n := n\tau$ for $n = 0, 1, 2, \dots$. Denote $\Phi^n(\mathbf{x}) = \Phi(t_n, \mathbf{x})$, from t_n to t_{n+1} , the solution $\Phi(t, \mathbf{x}) = \Phi(t_n + s, \mathbf{x})$ to (5.1.3) can be decomposed by different frequencies as

$$\Phi(t_n + s, \mathbf{x}) = e^{-is/\varepsilon^2} \Psi^{1,n}(s, \mathbf{x}) + e^{is/\varepsilon^2} \Psi^{2,n}(s, \mathbf{x}) \quad (5.2.7)$$

and then decompose them by the projectors

$$\Psi_{\pm}^{k,n}(s, \mathbf{x}) = \Pi_{\pm} \Psi^{k,n}(s, \mathbf{x}), \quad k = 1, 2, \quad (5.2.8)$$

we can obtain the entire decomposition

$$\Psi(t_n + s, \mathbf{x}) = e^{-is/\varepsilon^2} (\Psi_+^{1,n}(s, \mathbf{x}) + \Psi_-^{1,n}(s, \mathbf{x})) + e^{is/\varepsilon^2} (\Psi_+^{2,n}(s, \mathbf{x}) + \Psi_-^{2,n}(s, \mathbf{x})). \quad (5.2.9)$$

where $(\Psi_+^{k,n}(s, \mathbf{x}), \Psi_-^{k,n}(s, \mathbf{x}))$ ($k = 1, 2$) solves the coupled system for $\mathbf{x} \in \mathbb{R}^d$ ($d = 1, 2$) as

$$\begin{cases} i\partial_s \Psi_+^{1,n}(s, \mathbf{x}) = \frac{1}{\varepsilon^2} (\sqrt{1 - \varepsilon^2 \Delta} - 1) \Psi_+^{1,n}(s, \mathbf{x}) + \Pi_+ (W^n \Psi_+^{1,n}(s, \mathbf{x}) + W^n \Psi_-^{1,n}(s, \mathbf{x})), \\ i\partial_s \Psi_-^{1,n}(s, \mathbf{x}) = \frac{1}{\varepsilon^2} (-\sqrt{1 - \varepsilon^2 \Delta} - 1) \Psi_-^{1,n}(s, \mathbf{x}) + \Pi_- (W^n \Psi_+^{1,n}(s, \mathbf{x}) + W^n \Psi_-^{1,n}(s, \mathbf{x})), \\ \Psi_+^{1,n}(0, \mathbf{x}) = \Pi_+ \Phi^n(\mathbf{x}), \quad \Psi_-^{1,n}(0, \mathbf{x}) = \mathbf{0}, \end{cases} \quad (5.2.10)$$

and

$$\begin{cases} i\partial_s \Psi_+^{2,n}(s, \mathbf{x}) = \frac{1}{\varepsilon^2} (\sqrt{1 - \varepsilon^2 \Delta} + 1) \Psi_+^{2,n}(s, \mathbf{x}) + \Pi_+ (W^n \Psi_+^{2,n}(s, \mathbf{x}) + W^n \Psi_-^{2,n}(s, \mathbf{x})), \\ i\partial_s \Psi_-^{2,n}(s, \mathbf{x}) = \frac{1}{\varepsilon^2} (-\sqrt{1 - \varepsilon^2 \Delta} + 1) \Psi_-^{2,n}(s, \mathbf{x}) + \Pi_- (W^n \Psi_+^{2,n}(s, \mathbf{x}) + W^n \Psi_-^{2,n}(s, \mathbf{x})), \\ \Psi_+^{2,n}(0, \mathbf{x}) = \mathbf{0}, \quad \Psi_-^{2,n}(0, \mathbf{x}) = \Pi_- \Phi^n(\mathbf{x}), \end{cases} \quad (5.2.11)$$

with $W^n := W(t_n + s, \mathbf{x})$.

Following the analysis in [20], it is easy to verify that $\Psi_+^{1,n}(s, \mathbf{x}) = O(1)$, $\Psi_-^{2,n}(s, \mathbf{x}) = O(1)$, $\Psi_-^{1,n}(s, \mathbf{x}) = O(\varepsilon^2)$, $\Psi_+^{2,n}(s, \mathbf{x}) = O(\varepsilon^2)$, and $\partial_s \Psi_\pm^{k,n} = O(1)$. Thus $\Phi(t_{n+1}, \mathbf{x})$ can be evaluated by numerically solving the coupled systems (5.2.10)-(5.2.11) properly through the decomposition (5.2.9).

5.3 A multiscale time integrator pseudospectral method

5.3.1 The method in 1D

For the simplicity of notations, we shall only present our method and analysis for the Dirac equation (5.1.3) in 1D. Generalizations to high dimensions and/or (5.1.1) are straightforward and our results remain valid without any modifications. As a common practice in the literatures [19, 23, 52, 53, 57, 78, 102, 106] for practical computation, the Dirac equation (5.1.3) is usually truncated on a bounded interval $\Omega = (a, b)$ with periodic boundary conditions for $\Phi := \Phi(t, x) \in \mathbb{C}^2$,

$$\begin{cases} i\partial_t \Phi(t, x) = \frac{1}{\varepsilon^2} T^\varepsilon \Phi(t, x) + W(t, x) \Phi(t, x), & x \in \Omega, \quad t > 0, \\ \Phi(t, a) = \Phi(t, b), \quad \partial_x \Phi(t, a) = \partial_x \Phi(t, b), & t \geq 0, \\ \Phi(0, x) = \Phi_0(x), & x \in \bar{\Omega}, \end{cases} \quad (5.3.1)$$

where \mathcal{T} and $W(t, x)$ collapse to

$$\mathcal{T} = -i\varepsilon \sigma_1 \partial_x + \sigma_3, \quad W(t, \mathbf{x}) = V(t, x) I_2 - A_1(t, x) \sigma_1. \quad (5.3.2)$$

Then the systems (5.2.10)-(5.2.11) for the decomposition (5.2.9) with $x \in \Omega$ becomes

$$\begin{cases} i\partial_s \Psi_+^{1,n}(s, x) = \frac{1}{\varepsilon^2} (\sqrt{1 - \varepsilon^2 \Delta} - 1) \Psi_+^{1,n}(s, x) + \Pi_+ (W \Psi_+^{1,n}(s, x) + W \Psi_-^{1,n}(s, x)), \\ i\partial_s \Psi_-^{1,n}(s, x) = \frac{1}{\varepsilon^2} (-\sqrt{1 - \varepsilon^2 \Delta} - 1) \Psi_-^{1,n}(s, x) + \Pi_- (W \Psi_+^{1,n}(s, x) + W \Psi_-^{1,n}(s, x)), \\ \Psi_{\pm}^{1,n}(s, a) = \Psi_{\pm}^{1,n}(s, b), \quad s \geq 0; \quad \Psi_+^{1,n}(0, x) = \Pi_+ \Phi(t_n, x), \quad \Psi_-^{1,n}(0, x) = \mathbf{0}, \end{cases} \quad (5.3.3)$$

and

$$\begin{cases} i\partial_s \Psi_+^{2,n}(s, x) = \frac{1}{\varepsilon^2} (\sqrt{1 - \varepsilon^2 \Delta} + 1) \Psi_+^{2,n}(s, x) + \Pi_+ (W \Psi_+^{2,n}(s, x) + W \Psi_-^{2,n}(s, x)), \\ i\partial_s \Psi_-^{2,n}(s, x) = \frac{1}{\varepsilon^2} (-\sqrt{1 - \varepsilon^2 \Delta} + 1) \Psi_-^{2,n}(s, x) + \Pi_- (W \Psi_+^{2,n}(s, x) + W \Psi_-^{2,n}(s, x)), \\ \Psi_{\pm}^{2,n}(s, a) = \Psi_{\pm}^{2,n}(s, b), \quad s \geq 0; \quad \Psi_+^{2,n}(0, x) = \mathbf{0}, \quad \Psi_-^{2,n}(0, x) = \Pi_- \Phi(t_n, x). \end{cases} \quad (5.3.4)$$

Based on (5.3.3)-(5.3.4), a uniformly accurate numerical method can be designed as follows. We shall combine the Fourier spectral discretization in space and EWI in time.

Choose mesh size $\Delta x := \frac{b-a}{M}$ with M being a positive even integer and denote $h := \Delta x$ and grid points as $x_j := a + j \Delta x$, $j = 0, 1, \dots, M$. Denote $X_M = \{U = (U_0, U_1, \dots, U_M)^T \mid U_j \in \mathbb{C}^2, j = 0, 1, \dots, M, U_0 = U_M\}$ and we always use $U_{-1} = U_{M-1}$ if it is involved. The l^2 norm in X_M is given by

$$\|U\|_{l^2}^2 = h \sum_{j=0}^{M-1} |U_j|^2, \quad U \in X_M. \quad (5.3.5)$$

Let $[C_p(a, b)]^2$ be the function space consisting of all periodic vector function $U(x) : [a, b] \rightarrow \mathbb{C}^2$. For any $U(x) \in [C_p(a, b)]^2$ and $U \in X_M$, define $P_M : [L^2(a, b)]^2 \rightarrow Y_M$ as the standard projection operator, $I_M : [C_p(a, b)]^2 \rightarrow Y_M$ and $I_M : X_M \rightarrow Y_M$ as the standard interpolation operator, i.e.

$$(P_M U)(x) = \sum_{l=-M/2}^{M/2-1} \widehat{U}_l e^{i\mu_l(x-a)}, \quad (I_M U)(x) = \sum_{l=-M/2}^{M/2-1} \widetilde{U}_l e^{i\mu_l(x-a)}, \quad a \leq x \leq b, \quad (5.3.6)$$

with

$$\widehat{U}_l = \frac{1}{b-a} \int_a^b U(x) e^{-i\mu_l(x-a)} dx, \quad \widetilde{U}_l = \frac{1}{M} \sum_{j=0}^{M-1} U_j e^{-2ijl\pi/M}, \quad \mu_l = 2l\pi/(b-a), \quad (5.3.7)$$

where $l = -\frac{M}{2}, \dots, \frac{M}{2} - 1$ and $U_j = U(x_j)$ when U is a function. The Parseval's identity implies that

$$\|I_M(U)(\cdot)\|_{L^2} = \|U\|_{l^2}, \quad \forall U \in X_M. \quad (5.3.8)$$

The Fourier spectral discretization for (5.3.3-5.3.4) reads:

Find $\Psi_{\pm, M}^{k, n} := \Psi_{\pm, M}^{k, n}(s) = \Psi_{\pm, M}^{k, n}(s, x) \in Y_M$ ($0 \leq s \leq \tau$), i.e.

$$\Psi_{\pm, M}^{k, n}(s, x) = \sum_{l=-M/2}^{M/2-1} \left(\widehat{\Psi_{\pm}^{k, n}} \right)_l(s) e^{i\mu_l(x-a)}, \quad k = 1, 2, \quad a \leq x \leq b, \quad s \geq 0, \quad (5.3.9)$$

such that for $a < x < b$ and $s \geq 0$,

$$\begin{cases} i\partial_s \Psi_{+, M}^{1, n}(s) = \frac{1}{\varepsilon^2} (\sqrt{1 - \varepsilon^2 \Delta} - 1) \Psi_{+, M}^{1, n}(s) + \Pi_+ (W^n \Psi_{+, M}^{1, n} + W^n \Psi_{-, M}^{1, n}), \\ i\partial_s \Psi_{-, M}^{1, n}(s) = \frac{1}{\varepsilon^2} (-\sqrt{1 - \varepsilon^2 \Delta} - 1) \Psi_{-, M}^{1, n}(s) + \Pi_- (W^n \Psi_{+, M}^{1, n}(s) + W^n \Psi_{-, M}^{1, n}(s)), \\ \Psi_{+, M}^{1, n}(0) = P_M (\Pi_+ \Phi(t_n, x)), \quad \Psi_{-, M}^{1, n}(0) = \mathbf{0}, \end{cases} \quad (5.3.10)$$

and

$$\begin{cases} i\partial_s \Psi_{+, M}^{2, n}(s) = \frac{1}{\varepsilon^2} (\sqrt{1 - \varepsilon^2 \Delta} + 1) \Psi_{+, M}^{2, n}(s) + \Pi_+ (W^n \Psi_{+, M}^{2, n}(s) + W^n \Psi_{-, M}^{2, n}(s)), \\ i\partial_s \Psi_{-, M}^{2, n}(s) = \frac{1}{\varepsilon^2} (-\sqrt{1 - \varepsilon^2 \Delta} + 1) \Psi_{-, M}^{2, n}(s) + \Pi_- (W^n \Psi_{+, M}^{2, n}(s) + W^n \Psi_{-, M}^{2, n}(s)), \\ \Psi_{+, M}^{2, n}(0) = \mathbf{0}, \quad \Psi_{-, M}^{2, n}(0) = P_M (\Pi_- \Phi(t_n, x)). \end{cases} \quad (5.3.11)$$

We then obtain the equations for the Fourier coefficients as

$$\begin{cases} i\partial_s \left(\widehat{\Psi_+^{1, n}} \right)_l = \frac{\delta_l^-}{\varepsilon^2} I_2 \left(\widehat{\Psi_+^{1, n}} \right)_l + \Pi_l^+ \left(\widehat{W^n \Psi_{+, M}^{1, n}} \right)_l + \Pi_l^+ \left(\widehat{W^n \Psi_{-, M}^{1, n}} \right)_l, \\ i\partial_s \left(\widehat{\Psi_-^{1, n}} \right)_l = -\frac{\delta_l^+}{\varepsilon^2} I_2 \left(\widehat{\Psi_-^{1, n}} \right)_l + \Pi_l^- \left(\widehat{W^n \Psi_{+, M}^{1, n}} \right)_l + \Pi_l^- \left(\widehat{W^n \Psi_{-, M}^{1, n}} \right)_l, \end{cases} \quad (5.3.12)$$

and

$$\begin{cases} i\partial_s \left(\widehat{\Psi_+^{2, n}} \right)_l = \frac{\delta_l^+}{\varepsilon^2} I_2 \left(\widehat{\Psi_+^{2, n}} \right)_l + \Pi_l^+ \left(\widehat{W^n \Psi_{+, M}^{2, n}} \right)_l + \Pi_l^+ \left(\widehat{W^n \Psi_{-, M}^{2, n}} \right)_l, \\ i\partial_s \left(\widehat{\Psi_-^{2, n}} \right)_l = -\frac{\delta_l^-}{\varepsilon^2} I_2 \left(\widehat{\Psi_-^{2, n}} \right)_l + \Pi_l^- \left(\widehat{W^n \Psi_{+, M}^{2, n}} \right)_l + \Pi_l^- \left(\widehat{W^n \Psi_{-, M}^{2, n}} \right)_l, \end{cases} \quad (5.3.13)$$

where for $l = -\frac{M}{2}, \dots, \frac{M}{2} - 1$,

$$\delta_l = \sqrt{1 + \varepsilon^2 \mu_l^2}, \quad \delta_l^+ = \delta_l + 1, \quad \delta_l^- = \delta_l - 1; \quad (5.3.14)$$

Π_l^+ and Π_l^- are the corresponding Fourier representations of the projectors Π_{\pm} as

$$\Pi_l^+ = \begin{pmatrix} \frac{1+\delta_l}{2\delta_l} & \frac{\varepsilon\mu_l}{2\delta_l} \\ \frac{\varepsilon\mu_l}{2\delta_l} & \frac{\delta_l-1}{2\delta_l} \end{pmatrix}, \quad \Pi_l^- = \begin{pmatrix} \frac{\delta_l-1}{2\delta_l} & -\frac{\varepsilon\mu_l}{2\delta_l} \\ -\frac{\varepsilon\mu_l}{2\delta_l} & \frac{1+\delta_l}{2\delta_l} \end{pmatrix}. \quad (5.3.15)$$

Using variation-of-constant formula, we can write the solution as

$$\begin{aligned} & \widehat{\left(\Psi_+^{1,n}\right)}_l(s) \\ &= e^{-i\delta_l^- s/\varepsilon^2} \widehat{\left(\Psi_+^{1,n}\right)}_l(0) - i \int_0^s e^{-i\delta_l^-(s-w)/\varepsilon^2} \Pi_l^+ \left(\widehat{\left(W^n \Psi_{+,M}^{1,n}\right)}_l(w) + \widehat{\left(W^n \Psi_{-,M}^{1,n}\right)}_l(w) \right) dw, \\ & \widehat{\left(\Psi_-^{1,n}\right)}_l(s) \\ &= e^{i\delta_l^+ s/\varepsilon^2} \widehat{\left(\Psi_-^{1,n}\right)}_l(0) - i \int_0^s e^{i\delta_l^+(s-w)/\varepsilon^2} \Pi_l^- \left(\widehat{\left(W^n \Psi_{+,M}^{1,n}\right)}_l(w) + \widehat{\left(W^n \Psi_{-,M}^{1,n}\right)}_l(w) \right) dw, \\ & \widehat{\left(\Psi_+^{2,n}\right)}_l(s) \\ &= e^{-i\delta_l^+ s/\varepsilon^2} \widehat{\left(\Psi_+^{2,n}\right)}_l(0) - i \int_0^s e^{-i\delta_l^+(s-w)/\varepsilon^2} \Pi_l^+ \left(\widehat{\left(W^n \Psi_{+,M}^{2,n}\right)}_l(w) + \widehat{\left(W^n \Psi_{-,M}^{2,n}\right)}_l(w) \right) dw, \\ & \widehat{\left(\Psi_-^{2,n}\right)}_l(s) \\ &= e^{i\delta_l^- s/\varepsilon^2} \widehat{\left(\Psi_-^{2,n}\right)}_l(0) - i \int_0^s e^{i\delta_l^-(s-w)/\varepsilon^2} \Pi_l^- \left(\widehat{\left(W^n \Psi_{+,M}^{2,n}\right)}_l(w) + \widehat{\left(W^n \Psi_{-,M}^{2,n}\right)}_l(w) \right) dw. \end{aligned}$$

Using the initial conditions, choosing $s = \tau$, we approximate the integrals via Gautschi type quadrature rules [16, 44, 51, 55] or EWI [13, 15–18], using Taylor expansion and the equations (5.3.10)-(5.3.11) to determine the first order derivative, e.g. for the first integral in the above equation, we could derive

$$\begin{aligned}
& -i \int_0^\tau e^{-i\delta_l^-(\tau-w)/\varepsilon^2} \Pi_l^+ \left((\widehat{W^n \Psi_{+,M}^{1,n}})_l(w) + (\widehat{W^n \Psi_{-,M}^{1,n}})_l(w) \right) dw \quad (5.3.16) \\
& \approx -i \int_0^\tau e^{-i\delta_l^-(\tau-w)/\varepsilon^2} \Pi_l^+ \left((\widehat{W^n \Psi_{+,M}^{1,n}})_l(0) + (\widehat{W^n \Psi_{-,M}^{1,n}})_l(0) \right) dw \\
& -i \int_0^\tau e^{-i\delta_l^-(\tau-w)/\varepsilon^2} w \Pi_l^+ \left(\partial_s (\widehat{W^n \Psi_{+,M}^{1,n}})_l(0) + \partial_s (\widehat{W^n \Psi_{-,M}^{1,n}})_l(0) \right) dw \\
& = p_l^-(\tau) \Pi_l^+ (\widehat{W^n \Psi_{+,M}^{1,n}})_l(0) + q_l^-(\tau) \Pi_l^+ \left(\partial_s (\widehat{W^n \Psi_{+,M}^{1,n}})_l(0) + \partial_s (\widehat{W^n \Psi_{-,M}^{1,n}})_l(0) \right), \quad (5.3.17)
\end{aligned}$$

where

$$p_l^-(\tau) = -i\tau e^{\frac{-i\tau\delta_l^-}{2\varepsilon^2}} \operatorname{sinc}\left(\frac{\delta_l^- \tau}{2\varepsilon^2}\right), \quad q_l^-(\tau) = -\frac{\tau\varepsilon^2}{\delta_l^-} \left(1 - e^{\frac{-i\tau\delta_l^-}{2\varepsilon^2}} \operatorname{sinc}\left(\frac{\delta_l^- \tau}{2\varepsilon^2}\right)\right), \quad (5.3.18)$$

and $\operatorname{sinc}(s) = \frac{\sin s}{s}$ with $\operatorname{sinc}(0) = 1$. Note that for the special case $l = 0$, $q_0^-(\tau) = -i\frac{\tau^2}{2}$ and $p_l^-(\tau) = O(\tau)$, $q_l(\tau) = O(\tau^2)$.

The other integrals can be approximated similarly as

$$\begin{aligned}
& -i \int_0^\tau e^{i\delta_l^+(\tau-w)/\varepsilon^2} \Pi_l^- \left((\widehat{W^n \Psi_{+,M}^{1,n}})_l(w) + (\widehat{W^n \Psi_{-,M}^{1,n}})_l(w) \right) dw \quad (5.3.19) \\
& \approx -\overline{p_l^+(\tau)} \Pi_l^- (\widehat{W^n \Psi_{+,M}^{1,n}})_l(0) - \overline{q_l^+(\tau)} \Pi_l^- \left(\partial_s (\widehat{W^n \Psi_{+,M}^{1,n}})_l(0) + \partial_s (\widehat{W^n \Psi_{-,M}^{1,n}})_l(0) \right),
\end{aligned}$$

$$\begin{aligned}
& -i \int_0^\tau e^{-i\delta_l^+(\tau-w)/\varepsilon^2} \Pi_l^+ \left((\widehat{W^n \Psi_{+,M}^{2,n}})_l(w) + (\widehat{W^n \Psi_{-,M}^{2,n}})_l(w) \right) dw \quad (5.3.20)
\end{aligned}$$

$$\approx p_l^+(\tau) \Pi_l^+ (\widehat{W^n \Psi_{-,M}^{2,n}})_l(0) + q_l^+(\tau) \Pi_l^+ \left(\partial_s (\widehat{W^n \Psi_{+,M}^{2,n}})_l(0) + \partial_s (\widehat{W^n \Psi_{-,M}^{2,n}})_l(0) \right),$$

$$\begin{aligned}
& -i \int_0^\tau e^{i\delta_l^-(\tau-w)/\varepsilon^2} \Pi_l^- \left((\widehat{W^n \Psi_{+,M}^{2,n}})_l(w) + (\widehat{W^n \Psi_{-,M}^{2,n}})_l(w) \right) dw \quad (5.3.21)
\end{aligned}$$

$$\approx -\overline{p_l^-(\tau)} \Pi_l^- (\widehat{W^n \Psi_{-,M}^{2,n}})_l(0) - \overline{q_l^-(\tau)} \Pi_l^- \left(\partial_s (\widehat{W^n \Psi_{+,M}^{2,n}})_l(0) + \partial_s (\widehat{W^n \Psi_{-,M}^{2,n}})_l(0) \right),$$

with \bar{c} denoting the complex conjugate of c and

$$p_l^+(\tau) = -i\tau e^{\frac{-i\tau\delta_l^+}{2\varepsilon^2}} \operatorname{sinc}\left(\frac{\delta_l^+ \tau}{2\varepsilon^2}\right), \quad q_l^+(\tau) = -\frac{\tau\varepsilon^2}{\delta_l^+} \left(1 - e^{\frac{-i\tau\delta_l^+}{2\varepsilon^2}} \operatorname{sinc}\left(\frac{\delta_l^+ \tau}{2\varepsilon^2}\right)\right). \quad (5.3.22)$$

Omitting the spatial x variable and writing

$$\begin{aligned} f_{\pm}^{k,n}(s) &= W(t_n + s)\Psi_{\pm,M}^{k,n}(s), \quad \dot{f}_{\pm}^{k,n}(s) = W(t_n)\partial_s\Psi_{\pm,M}^{k,n}(s), \\ g_{\pm}^{k,n}(s) &= \partial_s W(t_n + s)\Psi_{\pm,M}^{k,n}(s), \end{aligned} \quad (5.3.23)$$

we find the solutions should be updated in the order from small component to large component as

$$\begin{cases} \widehat{(\Psi_-^{1,n})}_l(\tau) \approx -\overline{p_l^+(\tau)}\Pi_l^-\widehat{(f_+^{1,n})}_l(0) - \overline{q_l^+(\tau)}\Pi_l^-\widehat{(g_+^{1,n})}_l(0) \\ \quad - \overline{q_l^+(\tau)}\Pi_l^-\left(\widehat{(f_+^{1,n})}_l(0) + \widehat{(f_-^{1,n})}_l(0)\right), \\ \widehat{(\Psi_+^{2,n})}_l(\tau) \approx p_l^+(\tau)\Pi_l^+\widehat{(f_-^{2,n})}_l(0) + q_l^+(\tau)\Pi_l^+\widehat{(g_-^{2,n})}_l(0) \\ \quad + q_l^+(\tau)\Pi_l^+\left(\widehat{(f_+^{2,n})}_l(0) + \widehat{(f_-^{2,n})}_l(0)\right), \end{cases}$$

with initial values and derivatives determined from (5.3.10)-(5.3.11) as

$$\begin{aligned} \widehat{(\Psi_+^{1,n})}_l(0) &= \Pi_l^+(\widehat{(\Phi(t_n))}_l), \quad \widehat{(\Psi_-^{1,n})}_l(0) = \widehat{(\Psi_+^{2,n})}_l(0) = 0, \quad \widehat{(\Psi_-^{2,n})}_l(0) = \Pi_l^-(\widehat{(\Phi(t_n))}_l), \\ \widehat{(\partial_s\Psi_{+,M}^{1,n})}_l(0) &\approx -i\frac{2\sin(\mu_l^2\tau/2)}{\delta_l^+\tau}\widehat{(\Psi_+^{1,n})}_l(0) - i\Pi_l^+(W(t_n)\widehat{\Psi_{+,M}^{1,n}}(0))_l, \\ \widehat{(\partial_s\Psi_{-,M}^{1,n})}_l(0) &= -i\Pi_l^-(W(t_n)\widehat{\Psi_{+,M}^{1,n}}(0))_l, \quad \widehat{(\partial_s\Psi_{+,M}^{2,n})}_l(0) = -i\Pi_l^+(W(t_n)\widehat{\Psi_{-,M}^{2,n}}(0))_l, \\ \widehat{(\partial_s\Psi_{-,M}^{2,n})}_l(0) &\approx i\frac{2\sin(\mu_l^2\tau/2)}{\delta_l^+\tau}\widehat{(\Psi_-^{2,n})}_l(0) - i\Pi_l^-(W(t_n)\widehat{\Psi_{-,M}^{2,n}}(0))_l, \end{aligned}$$

To avoid loss of accuracy, the derivatives $\partial_s\Psi_{+,M}^{1,n}(0)$ and $\partial_s\Psi_{-,M}^{2,n}(0)$ are approximated using filters $2\sin(\mu_l^2\tau/2)/\tau$ instead of μ_l^2 , followed by

$$\begin{cases} \widehat{(\Psi_+^{1,n})}_l(\tau) \approx e^{-i\delta_l^-\tau/\varepsilon^2}\widehat{(\Psi_+^{1,n})}_l(0) + p_l^-(\tau)\Pi_l^+\widehat{(f_+^{1,n})}_l(0) + q_l^-(\tau)\Pi_l^+\widehat{(g_+^{1,n})}_l(0) \\ \quad + q_l^-(\tau)\Pi_l^+\left(\widehat{(f_+^{1,n})}_l(0) + \widehat{(f_-^{1,n})}_l(0)\right) \\ \widehat{(\Psi_-^{2,n})}_l(\tau) \approx e^{i\delta_l^-\tau/\varepsilon^2}\widehat{(\Psi_-^{2,n})}_l(0) - \overline{p_l^-(\tau)}\Pi_l^-\widehat{(f_-^{2,n})}_l(0) - \overline{q_l^-(\tau)}\Pi_l^-\widehat{(g_-^{2,n})}_l(0) \\ \quad - \overline{q_l^-(\tau)}\Pi_l^-\left(\widehat{(f_+^{2,n})}_l(0) + \widehat{(f_-^{2,n})}_l(0)\right), \end{cases}$$

with $\widehat{(\partial_s\Psi_-^{1,n})}_l(0)$ and $\widehat{(\partial_s\Psi_+^{2,n})}_l(0)$ approximated in another way as

$$\widehat{(\partial_s\Psi_-^{1,n})}_l(0) \approx \widehat{(\Psi_-^{1,n})}_l(\tau)/\tau, \quad \widehat{(\partial_s\Psi_+^{2,n})}_l(0) \approx \widehat{(\Psi_+^{2,n})}_l(\tau)/\tau,$$

where the derivatives are approximated using filters to avoid loss of accuracy.

Based on the above discussions, a multiscale time integrator Fourier pseudospectral (MTI-FP) method for solving 1D Dirac equation (5.3.1) is designed as follows. Let $W_j^n = W(t_n, x_j)$, $\Phi_j^n \in X_M$ be the numerical approximation of exact solution $\Phi(t_n, x_j)$ to the Dirac equation (5.3.1); $\Psi_{\pm,j}^{k,n+1}$ be the numerical approximation of exact solution $\Psi_{\pm}^{k,n}(\tau, x_j)$ for $k = 1, 2$, $j = 0, 1, \dots, M$ and $n \geq 0$. Choosing $\Phi_j^0 = \Phi_0(x_j)$, then the scheme reads for $n \geq 0$ and $j = 0, 1, \dots, M$ as:

$$\Phi_j^{n+1} = e^{-i\tau/\varepsilon^2} (\Psi_{+,j}^{1,n+1} + \Psi_{-,j}^{1,n+1}) + e^{i\tau/\varepsilon^2} (\Psi_{+,j}^{2,n+1} + \Psi_{-,j}^{2,n+1}) = \sum_{l=-M/2}^{M/2-1} (\widetilde{\Phi^{n+1}})_l e^{i\mu_l(x_j-a)}, \quad (5.3.24)$$

where

$$\Psi_{\pm,j}^{k,n+1} = \sum_{l=-M/2}^{M/2-1} (\widetilde{\Psi_{\pm}^{k,n+1}})_l e^{i\mu_l(x_j-a)}, \quad k = 1, 2, \quad (5.3.25)$$

with

$$\left\{ \begin{array}{l} (\widetilde{\Psi_-^{1,n+1}})_l = -\overline{p_l^+(\tau)} \Pi_l^-(\widetilde{f_+^1})_l - \overline{q_l^+(\tau)} \Pi_l^-(\widetilde{g_+^1})_l - \overline{q_l^+(\tau)} \Pi_l^- \left(\left(\widetilde{f_+^1} \right)_l + \left(\widetilde{f_-^1} \right)_l \right), \\ (\widetilde{\Psi_+^{2,n+1}})_l = p_l^+(\tau) \Pi_l^+(\widetilde{f_-^2})_l + q_l^+(\tau) \Pi_l^+(\widetilde{g_-^2})_l + q_l^+(\tau) \Pi_l^+ \left(\left(\widetilde{f_+^2} \right)_l + \left(\widetilde{f_-^2} \right)_l \right), \\ (\widetilde{\Psi_+^{1,n+1}})_l = e^{-i\frac{\delta_l^- \tau}{\varepsilon^2}} (\widetilde{\Psi_+^1})_l + p_l^-(\tau) \Pi_l^+(\widetilde{f_+^1})_l \\ \quad + q_l^-(\tau) \Pi_l^+(\widetilde{g_+^1})_l + q_l^-(\tau) \Pi_l^+ \left(\left(\widetilde{f_+^1} \right)_l + \left(\widetilde{f_-^{1,*}} \right)_l \right), \\ (\widetilde{\Psi_-^{2,n+1}})_l = e^{i\frac{\delta_l^- \tau}{\varepsilon^2}} (\widetilde{\Psi_-^2})_l - \overline{p_l^-(\tau)} \Pi_l^-(\widetilde{f_-^2})_l - \overline{q_l^-(\tau)} \Pi_l^-(\widetilde{g_-^2})_l \\ \quad - \overline{q_l^-(\tau)} \Pi_l^- \left(\left(\widetilde{f_+^{2,*}} \right)_l + \left(\widetilde{f_-^2} \right)_l \right), \end{array} \right. \quad (5.3.26)$$

and

$$\left\{ \begin{array}{l} f_{\pm,j}^k = \sum_{l=-M/2}^{M/2-1} (\widetilde{f_{\pm}^k})_l e^{i\mu_l(x_j-a)}, \quad \dot{f}_{\pm,j}^k = \sum_{l=-M/2}^{M/2-1} (\widetilde{\dot{f}_{\pm}^k})_l e^{i\mu_l(x_j-a)}, \\ g_{\pm,j}^k = \sum_{l=-M/2}^{M/2-1} (\widetilde{g_{\pm}^k})_l e^{i\mu_l(x_j-a)}, \\ \dot{f}_{-,j}^{1,*} = \sum_{l=-M/2}^{M/2-1} (\widetilde{\dot{f}_-^{1,*}})_l e^{i\mu_l(x_j-a)} \quad \dot{f}_{+,j}^{2,*} = \sum_{l=-M/2}^{M/2-1} (\widetilde{\dot{f}_+^{2,*}})_l e^{i\mu_l(x_j-a)}, \end{array} \right. \quad (5.3.27)$$

with

$$\left\{ \begin{array}{l} \widetilde{(\Psi_+^1)}_l = \Pi_l^+(\widetilde{\Phi^n})_l, \quad \widetilde{(\Psi_-^1)}_l = \mathbf{0}, \quad \widetilde{(\Psi_+^2)}_l = \mathbf{0}, \quad \widetilde{(\Psi_-^2)}_l = \Pi_l^-(\widetilde{\Phi^n})_l, \\ f_{\pm,j}^k = W_j^n \Psi_{\pm,j}^k, \quad \dot{f}_{\pm,j}^k = W_j^n \dot{\Psi}_{\pm,j}^k, \quad g_{\pm,j}^k = \partial_t W(t_n, x_j) \Psi_{\pm,j}^k, \\ \widetilde{(\dot{\Psi}_+^1)}_l = -i \frac{2 \sin(\mu_l^2 \tau / 2)}{\delta_l^+ \tau} \widetilde{(\Psi_+^1)}_l - i \Pi_l^+(\widetilde{f_+^1})_l, \\ \widetilde{(\dot{\Psi}_-^1)}_l = -i \Pi_l^-(\widetilde{f_+^1})_l, \quad \widetilde{(\dot{\Psi}_+^2)}_l = -i \Pi_l^+(\widetilde{f_-^2})_l, \\ \widetilde{(\dot{\Psi}_-^2)}_l = i \frac{2 \sin(\mu_l^2 \tau / 2)}{\delta_l^+ \tau} \widetilde{(\Psi_-^2)}_l - i \Pi_l^-(\widetilde{f_-^2})_l, \\ \dot{f}_{-,j}^{1,*} = W_j^n (\Psi_{-,j}^{1,n+1}) / \tau, \quad \dot{f}_{+,j}^{2,*} = W_j^n (\Psi_{+,j}^{2,n+1}) / \tau. \end{array} \right. \quad (5.3.28)$$

Note that the small component $\Psi_{-,j}^{1,n+1}$ and $\Psi_{+,j}^{2,n+1}$ are evaluated at the first step, then finite difference approximations replacing the time derivatives for the small component should be used in the evaluations of the large component $\Psi_{+,j}^{1,n+1}$ and $\Psi_{-,j}^{2,n+1}$.

5.3.2 A uniform error bound

In order to obtain an error estimate for the MTI-FP (5.3.24) in the time interval $0 < t < T < \infty$, motivated by the results in [20, 26], we make the following assumptions on the electromagnetic potentials

$$(A'') \quad \|V\|_{W^{2,\infty}([0,T];(W_p^{m_0,\infty})^2)} + \|A_1\|_{W^{2,\infty}([0,T];(W_p^{m_0,\infty})^2)} \lesssim 1, \quad m_0 \geq 4,$$

and the exact solution $\Phi := \Phi(t, x)$ of Dirac equation (5.3.1)

$$(B'') \quad \|\Phi\|_{L^\infty([0,T];(H_p^{m_0})^2)} \lesssim 1, \quad \|\partial_t \Phi\|_{L^\infty([0,T];(H_p^{m_0-2})^2)} \lesssim \frac{1}{\varepsilon^2}, \quad \|\partial_{tt} \Phi(t, x)\|_{L^\infty([0,T];(L^2)^2)} \lesssim \frac{1}{\varepsilon^4},$$

where $H_p^k(\Omega) = \{u \mid u \in H^k(\Omega), \partial_x^l u(a) = \partial_x^l u(b), l = 0, \dots, m-1\}$ and $W_p^{k,\infty}(\Omega) = \{u \mid u \in W^{k,\infty}(\Omega), \partial_x^l u(a) = \partial_x^l u(b), l = 0, \dots, m-1\}$. We remark here that assumption (B'') is equivalent to the initial value $\Phi_0(x) \in (H_p^{m_0})^2$ [20, 73] under the assumption (A'').

Theorem 5.1 *Let $\Phi^n \in X_M$ and $\Phi_I^n(x) = I_M(\Phi^n)(x) \in Y_M$ be the numerical approximation obtained from MTI-FP (5.3.24). Under assumptions (A'') and (B''),*

there exists constants $0 < \tau_0, h_0 \leq 1$ independent of ε , such that if $0 < \tau \leq \tau_0$ and $0 < h \leq h_0$, we have

$$\|\Phi(t_n, \cdot) - \Phi_I^n(\cdot)\|_{L^2} \lesssim h^{m_0} + \frac{\tau^2}{\varepsilon^2}, \quad \|\Phi(t_n, \cdot) - \Phi_I^n(\cdot)\|_{L^2} \lesssim h^{m_0} + \tau^2 + \varepsilon^2, \quad (5.3.29)$$

which yields the uniform error bound by taking the minimum $\min\{\varepsilon^2, \tau^2/\varepsilon^2\}$,

$$\|\Phi(t_n, \cdot) - \Phi_I^n(\cdot)\|_{L^2} \lesssim h^{m_0} + \tau. \quad (5.3.30)$$

Remark 5.1 From the analysis point of view, we remark that $W_p^{m_0, \infty}$ assumption in (A'') is necessary for the exact solution $\Phi(t, x)$ remaining in $(H_p^{m_0})^2$, which would give the spectral accuracy. In practice, as long as the solution is well localized such that the error from the periodic truncation of potential term $W(t, x)\Phi(t, x)$ is negligible, the error estimates in the above theorem hold.

5.3.3 Error analysis

From now on, we will write the exact solution $\Phi(t, x)$ as $\Phi(t)$ for short. Define the error function $\mathbf{e}^n(x) = \sum_{l=-M/2}^{M/2-1} (\widetilde{\mathbf{e}^n})_l e^{i\mu_l(x-a)} \in Y_M$ for $n \geq 0$ as

$$\mathbf{e}^n(x) = P_M(\Phi(t_n))(x) - \Phi_I^n(x) = P_M(\Phi(t_n))(x) - I_M(\Phi^n)(x), \quad x \in \Omega. \quad (5.3.31)$$

Using assumption (B'') , triangle inequality and standard Fourier projection properties, we find

$$\|\Phi(t_n, \cdot) - \Phi_I^n(\cdot)\|_{L^2} \leq \|\Phi(t_n, \cdot) - P_M(\Phi(t_n))(\cdot)\|_{L^2} + \|P_M(\Phi(t_n))(\cdot) - \Phi_I^n(\cdot)\|_{L^2} \quad (5.3.32)$$

$$\lesssim h^{m_0} + \|\mathbf{e}^n(\cdot)\|_{L^2}, \quad 0 \leq n \leq \frac{T}{\tau}.$$

Hence, we only need estimate $\|\mathbf{e}^n(\cdot)\|_{L^2}$. To this purpose, local truncation error will be studied as the first step. Since MTI-FP (5.3.24) is designed by the multiscale decomposition, the following properties of the decomposition (5.3.3)-(5.3.4) are essential for the error analysis.

From t_n to t_{n+1} , let $\Psi_{\pm}^{k,n}(s, x)$ ($s \geq 0$, $k = 1, 2$) be the solution to the system (5.3.3)-(5.3.4), and the decomposition (5.2.9) holds as

$$\Phi(t_n + s, x) = e^{-is/\varepsilon^2} (\Psi_+^{1,n}(s, x) + \Psi_-^{1,n}(s, x)) + e^{is/\varepsilon^2} (\Psi_+^{2,n}(s, x) + \Psi_-^{2,n}(s, x)). \quad (5.3.33)$$

Then the error $\mathbf{e}^{n+1}(x)$ ($n \geq 0$) (5.3.31) can be decomposed as

$$\mathbf{e}^{n+1}(x) = e^{-i\tau/\varepsilon^2} (\mathbf{z}_+^{1,n+1}(x) + \mathbf{z}_-^{1,n+1}(x)) + e^{i\tau/\varepsilon^2} (\mathbf{z}_+^{2,n+1}(x) + \mathbf{z}_-^{2,n+1}(x)), \quad x \in \Omega \quad (5.3.34)$$

with

$$\mathbf{z}_{\pm}^{k,n+1}(x) = \sum_{l=-M/2}^{M/2-1} \widetilde{(\mathbf{z}_{\pm}^{k,n+1})}_l e^{i\mu_l(x-a)} = P_M(\Psi_{\pm}^{k,n}(\tau))(x) - I_M(\Psi_{\pm}^{k,n+1})(x), \quad k = 1, 2. \quad (5.3.35)$$

By the same arguments in [20], we can establish the regularity results.

Lemma 5.1 *Under the assumptions (A'') and (B''), the exact solutions $\Psi_{\pm}^{k,n}(s, x)$ ($s \geq 0$, $k = 1, 2$, $0 \leq n \leq \frac{T}{\tau} - 1$) to the system (5.3.3)-(5.3.4) satisfy*

$$\|\Psi_{\pm}^{k,n}\|_{L^\infty([0,\tau];(H_p^{m_0})^2)} \lesssim 1, \quad \|\partial_{ss}\Psi_+^{1,n}\|_{L^\infty([0,\tau];(H_p^{m_4})^2)} + \|\partial_{ss}\Psi_-^{2,n}\|_{L^\infty([0,\tau];(H_p^{m_4})^2)} \lesssim 1, \quad (5.3.36)$$

$$\|\partial_s\Psi_{\pm}^{k,n}\|_{L^\infty([0,\tau];(H_p^{m_2})^2)} \lesssim 1, \quad \|\partial_{ss}\Psi_+^{2,n}\|_{L^\infty([0,\tau];(H_p^{m_4})^2)} + \|\partial_{ss}\Psi_-^{1,n}\|_{L^\infty([0,\tau];(H_p^{m_4})^2)} \lesssim \frac{1}{\varepsilon^2}, \quad (5.3.37)$$

$$\|\Psi_+^{2,n}\|_{L^\infty([0,\tau];(H_p^{m_1})^2)} + \|\Psi_-^{1,n}\|_{L^\infty([0,\tau];(H_p^{m_1})^2)} \lesssim \varepsilon^2, \quad m_k = m_0 - k, \quad k = 1, 2, 4. \quad (5.3.38)$$

Proof: Noticing the properties of projectors Π_{\pm} and assumption (B''), the initial data $\Psi_{\pm}^n(0, x) \in (H_p^{m_0})^2$, $\partial_s\Psi_{\pm}^n(0, x) \in H_p^{m_0-2}$ with uniform bounds. The estimates for $\Psi_{\pm}^{k,n}$ and $\partial_s\Psi_{\pm}^{k,n}$ have been derived in [20], where one only need to replace the whole space Fourier transform to the Fourier series on torus. Thus, the proof is omitted here for brevity. It remains to estimate $\partial_{ss}\Psi_{\pm}^{k,n}$. Here, we show the case $k = 1$, while $k = 2$ case is quite the same. Differentiating (5.3.3) once with respect

to time s , we obtain for $\partial_{ss}\Psi_{\pm}^{1,n}(s)$,

$$\begin{aligned} i\partial_{ss}\Psi_{+}^{1,n}(s) &= - (1 - \varepsilon^2\Delta)^{-1/2} \Delta\partial_s\Psi_{+}^{1,n}(s) + \Pi_{+}\partial_s (W\Psi_{+}^{1,n}(s, x) + W\Psi_{-}^{1,n}(s, x)), \\ i\partial_{ss}\Psi_{-}^{1,n}(s) &= - \frac{\sqrt{1 - \varepsilon^2\Delta} + 1}{\varepsilon^2} \partial_s\Psi_{-}^{1,n}(s) + \Pi_{+}\partial_s (W\Psi_{+}^{1,n}(s, x) + W\Psi_{-}^{1,n}(s, x)). \end{aligned}$$

Since for any $\Phi \in (H_p^m)^2$, we have

$$\left\| (1 - \varepsilon^2\Delta)^{-1/2} \Delta\Phi \right\|_{H_p^{m-2}} \leq \|\Phi\|_{H_p^m}, \quad \left\| \frac{\sqrt{1 - \varepsilon^2\Delta} + 1}{\varepsilon^2} \Phi \right\|_{H_p^{m-1}} \lesssim \frac{1}{\varepsilon^2} \|\Phi\|_{H_p^m}, \quad (5.3.39)$$

which implies the bounds (5.3.37) for $\partial_{ss}\Psi_{\pm}^{1,n}(s)$, in view of the estimates for $\Psi_{\pm}^{1,n}$ and assumption (A'') . □

Having Lemma 5.1 and decomposition (5.3.34), we are able to define the local truncation error $\xi_{\pm}^{k,n}(x) = \sum_{l=-M/2}^{M/2-1} (\widehat{\xi_{\pm}^{k,n}})_l e^{i\mu_l(x-a)}$ ($x \in \Omega$, $k = 1, 2$, $n \geq 0$) for MTI-FP scheme (5.3.24)-(5.3.28) as

$$\left\{ \begin{aligned} (\widehat{\xi_{-}^{1,n}})_l &= (\widehat{\Psi_{-}^{1,n}}(\tau))_l + \overline{p_l^+(\tau)} \Pi_l^- (\widehat{f_+^1(0)})_l + \overline{q_l^+(\tau)} \Pi_l^- (\widehat{g_+^1(0)})_l \\ &\quad + \overline{q_l^+(\tau)} \Pi_l^- \left((\widehat{f_+^1(0)})_l + (\widehat{f_-^1(0)})_l \right), \\ (\widehat{\xi_{+}^{2,n}})_l &= (\widehat{\Psi_{+}^{2,n}}(\tau))_l - p_l^+(\tau) \Pi_l^+ (\widehat{f_-^2(0)})_l - q_l^+(\tau) \Pi_l^+ (\widehat{g_-^2(0)})_l \\ &\quad - q_l^+(\tau) \Pi_l^+ \left((\widehat{f_+^2(0)})_l + (\widehat{f_-^2(0)})_l \right), \\ (\widehat{\xi_{+}^{1,n}})_l &= (\widehat{\Psi_{+}^{1,n}}(\tau))_l - e^{-i\frac{\delta_l^- \tau}{\varepsilon^2}} (\widehat{\Psi_{+}^{1,n}}(0))_l - p_l^-(\tau) \Pi_l^+ (\widehat{f_+^1(0)})_l - q_l^-(\tau) \Pi_l^+ (\widehat{g_+^1(0)})_l \\ &\quad - q_l^-(\tau) \Pi_l^+ \left((\widehat{f_+^1(0)})_l + (\widehat{f_-^{1,*}(\tau)})_l \right), \\ (\widehat{\xi_{-}^{2,n}})_l &= (\widehat{\Psi_{-}^{2,n}}(\tau))_l - e^{i\frac{\delta_l^- \tau}{\varepsilon^2}} (\widehat{\Psi_{-}^{2,n}}(0))_l + \overline{p_l^-(\tau)} \Pi_l^- (\widehat{f_-^2(0)})_l + \overline{q_l^-(\tau)} \Pi_l^- (\widehat{g_-^2(0)})_l \\ &\quad + \overline{q_l^-(\tau)} \Pi_l^- \left((\widehat{f_+^{2,*}(\tau)})_l + (\widehat{f_-^2(0)})_l \right), \end{aligned} \right. \quad (5.3.40)$$

with

$$\begin{cases} \widehat{(\Psi_+^{1,n})}_l(0) = \Pi_l^+(\widehat{\Phi}(t_n))_l, \quad \widehat{(\Psi_-^{1,n})}_l(0) = \mathbf{0}, \quad \widehat{(\Psi_+^{2,n})}_l(0) = \mathbf{0}, \quad \widehat{(\Psi_-^{2,n})}_l(0) = \Pi_l^-(\widehat{\Phi}(t_n))_l, \\ f_\pm^k(s) = W(t_n + s)\Psi_\pm^{k,n}(s), \quad \dot{f}_\pm^k(s) = W(t_n)\dot{\Psi}_\pm^{k,n}(s), \quad g_\pm^k(s) = \partial_t W(t_n)\Psi_\pm^{k,n}(s), \\ \dot{f}_-^{1,*}(s) = W(t_n)(\Psi_-^{1,n+1}(s))/s \quad \dot{f}_{+,j}^{2,*} = W(t_n)(\Psi_{+,j}^{2,n+1}(s))/s. \end{cases} \quad (5.3.41)$$

and

$$\dot{\Psi}_\pm^{k,n}(s, x) = \sum_{l=-M/2}^{M/2} \widetilde{(\dot{\Psi}_+^{k,n}(s))}_l e^{i\mu_l(x-a)}, \quad k = 1, 2, \quad s \geq 0, \quad (5.3.42)$$

given by

$$\begin{cases} \widetilde{(\dot{\Psi}_+^{1,n}(s))}_l = -i \frac{2 \sin(\mu_l^2 \tau / 2)}{\delta_l^+ \tau} \widetilde{(\Psi_+^{1,n}(s))}_l - i \Pi_l^+(\widetilde{f_+^1(s)})_l, \\ \widetilde{(\dot{\Psi}_-^{1,n}(s))}_l = -i \Pi_l^-(\widetilde{f_+^1(s)})_l, \quad \widetilde{(\dot{\Psi}_+^{2,n}(s))}_l = -i \Pi_l^+(\widetilde{f_-^2(s)})_l, \\ \widetilde{(\dot{\Psi}_-^{2,n}(s))}_l = i \frac{2 \sin(\mu_l^2 \tau / 2)}{\delta_l^+ \tau} \widetilde{(\Psi_-^{2,n}(s))}_l - i \Pi_l^-(\widetilde{f_-^2(s)})_l. \end{cases} \quad (5.3.43)$$

We have the following estimates for the above local truncation error.

Lemma 5.2 *Under assumption (A'') and (B''), the local truncation error $\xi_\pm^{k,n} \in Y_M$ (5.3.40) for $n = 0, 1, \dots, \frac{T}{\tau} - 1$ satisfies*

$$\|\xi_\pm^{k,n}\|_{L^2} \lesssim h^{m_0} + \frac{\tau^2}{\varepsilon^2}, \quad \|\xi_\pm^{k,n}\|_{L^2} \lesssim h^{m_0} + \tau^2 + \varepsilon^2, \quad k = 1, 2. \quad (5.3.44)$$

Proof: We will only prove the estimates (5.3.44) for $k = 1$, as the proof for $k = 2$ is the same. Using the fact $\delta_l^+ \geq 1$ and the definitions of $p_l^\pm(\tau)$ and $q_l^\pm(\tau)$ ($l = -M/2, \dots, M/2 - 1$), we notice that

$$|p_l^\pm(\tau)| \lesssim \tau, \quad |q_l^\pm(\tau)| \lesssim \tau^2, \quad |p_l^+(\tau)| \lesssim \varepsilon^2, \quad |q_l^+(\tau)| \lesssim \tau \varepsilon^2. \quad (5.3.45)$$

Multiplying both sides of the equations in the system (5.3.3) by $e^{-i\mu_l(x-a)}$ and integrating over Ω , we easily recover the equations for $\widetilde{(\Psi_\pm^{k,n})}_l(s)$, which are exactly the same as (5.3.12)-(5.3.13) with $\Psi_{\pm,M}^{k,n}$ being replaced by $\Psi_\pm^{k,n}$.

Following the derivation of the MTI-FP scheme, it is easy to find that the local truncation error comes from the approximations in the integrals (5.3.16), (5.3.19),

(5.3.20) and (5.3.21). In particular, for $l = -M/2, \dots, M/2 - 1$, we have

$$\begin{aligned} (\widehat{\xi_-^{1,n}})_l &= -i \int_0^\tau e^{i\delta_l^+(\tau-s)/\varepsilon^2} \Pi_l^- \left((\widehat{f_+^1(s)})_l + (\widehat{f_-^1(s)})_l \right) ds + \overline{p_l^+(\tau)} \Pi_l^- (\widehat{f_+^1(0)})_l \\ &\quad + \overline{q_l^+(\tau)} \Pi_l^- (\widehat{g_+^1(0)})_l + \overline{q_l^+(\tau)} \Pi_l^- \left((\widehat{f_+^1(0)})_l + (\widehat{f_-^1(0)})_l \right), \end{aligned} \quad (5.3.46)$$

$$\begin{aligned} (\widehat{\xi_+^{1,n}})_l &= -i \int_0^\tau e^{-i\delta_l^-(\tau-s)} \Pi_l^+ \left((\widehat{f_+^1(s)})_l + (\widehat{f_-^1(s)})_l \right) ds - p_l^-(\tau) \Pi_l^+ (\widehat{f_+^1(0)})_l \\ &\quad - q_l^-(\tau) \Pi_l^+ (\widehat{g_+^1(0)})_l - q_l^-(\tau) \Pi_l^+ \left((\widehat{f_+^1(0)})_l + (\widehat{f_-^{1,*}(\tau)})_l \right). \end{aligned} \quad (5.3.47)$$

Type I estimates. Firstly, we prove the first kind estimates in (5.3.44). Using Taylor expansion, we have

$$\begin{aligned} (\widehat{\xi_-^{1,n}})_l &= -i \int_0^\tau \int_0^s \int_0^{s_1} e^{i\delta_l^+(\tau-s)/\varepsilon^2} \Pi_l^- \left((\partial_{s_2 s_2} \widehat{f_+^1(s_2)})_l + (\partial_{s_2 s_2} \widehat{f_-^1(s_2)})_l \right) ds_2 ds_1 ds \\ &\quad + (\widehat{\eta_-^1})_l, \end{aligned} \quad (5.3.48)$$

$$\begin{aligned} (\widehat{\xi_+^{1,n}})_l &= -i \int_0^\tau \int_0^s \int_0^{s_1} e^{-i\delta_l^-(\tau-s)/\varepsilon^2} \Pi_l^+ \left((\partial_{s_2 s_2} \widehat{f_+^1(s_2)})_l + (\partial_{s_2 s_2} \widehat{f_-^1(s_2)})_l \right) ds_2 ds_1 ds \\ &\quad + (\widehat{\eta_+^1})_l, \end{aligned} \quad (5.3.49)$$

where $\eta_\pm^1(x) = \sum_{l=-M/2}^{M/2-1} (\widehat{\eta_\pm^1})_l e^{i\mu_l(x-a)}$ with

$$\begin{aligned} (\widehat{\eta_-^1})_l &= \overline{p_l^+(\tau)} \Pi_l^- \left(-(\widehat{f_+^1(0)})_l + (\widehat{f_+^1(0)})_l \right) + \overline{q_l^+(\tau)} \Pi_l^- \left(-(\widehat{g_+^1(0)})_l + (\widehat{g_+^1(0)})_l \right) \\ &\quad + \overline{q_l^+(\tau)} \Pi_l^- \left(-(\widehat{f_+^{1,n}(0)})_l + (\widehat{f_+^1(0)})_l - (\widehat{f_-^{1,n}(0)})_l + (\widehat{f_-^1(0)})_l \right), \\ (\widehat{\eta_+^1})_l &= p_l^-(\tau) \Pi_l^+ \left((\widehat{f_+^1(0)})_l - (\widehat{f_+^1(0)})_l \right) + q_l^-(\tau) \Pi_l^+ \left((\widehat{g_+^1(0)})_l - (\widehat{g_+^1(0)})_l \right) \\ &\quad + q_l^-(\tau) \Pi_l^+ \left((\widehat{f_+^{1,n}(0)})_l - (\widehat{f_+^1(0)})_l + (\widehat{f_-^{1,n}(0)})_l - (\widehat{f_-^{1,*}(\tau)})_l \right), \end{aligned}$$

and $f_\pm^{1,n}(s)$ is given in (5.3.23) with $\Psi_{\pm,M}^{k,n}$ being replaced by $\Psi_\pm^{k,n}$. Since $\|\Pi_l^\pm\|_{l^2} \leq 1$ ($l = -\frac{M}{2}, \dots, \frac{M}{2} - 1$) with $\|Q\|_{l^2}$ being the standard l^2 norm of the matrix Q , using

(5.3.45) and triangle inequality, we obtain

$$\begin{aligned} |(\widehat{\eta_-^1})_l| &\lesssim \tau \left| (\widehat{f_+^1(0)})_l - (\widetilde{f_+^1(0)})_l \right| + \tau^2 \left| (\widehat{g_+^1(0)})_l - (\widetilde{g_+^1(0)})_l \right| + \tau^2 \left| (\widehat{\dot{f}_+^{1,n}(0)})_l - (\widetilde{\dot{f}_+^{1,n}(0)})_l \right| \\ &\quad + \tau^2 \left| (\widehat{\dot{f}_+^{1,n}(0)})_l - (\widetilde{\dot{f}_+^{1,n}(0)})_l \right| + \tau^2 \left| (\widehat{\dot{f}_-^{1,n}(0)})_l - (\widetilde{\dot{f}_-^{1,n}(0)})_l \right| \\ &\quad + \tau^2 \left| (\widehat{\dot{f}_-^{1,n}(0)})_l - (\widetilde{\dot{f}_-^{1,n}(0)})_l \right|, \end{aligned}$$

and Parseval's theorem then implies

$$\begin{aligned} \|\eta_-^1(\cdot)\|_{L^2}^2 &\lesssim \tau^2 \|P_M(f_+^1(0)) - I_M(f_+^1(0))\|_{L^2}^2 + \tau^4 \|P_M(g_+^1(0)) - I_M(g_+^1(0))\|_{L^2}^2 \\ &\quad + \tau^4 \|P_M(\dot{f}_+^{1,n}(0)) - I_M(\dot{f}_+^{1,n}(0))\|_{L^2}^2 + \tau^4 \|P_M(\dot{f}_-^{1,n}(0)) - I_M(\dot{f}_-^{1,n}(0))\|_{L^2}^2 \\ &\quad + \tau^4 \|I_M(\dot{f}_-^{1,n}(0)) - I_M(\dot{f}_-^1(0))\|_{L^2}^2 + \tau^4 \|I_M(\dot{f}_-^{1,n}(0)) - I_M(\dot{f}_-^1(0))\|_{L^2}^2. \end{aligned} \tag{5.3.50}$$

Recalling assumption (A'') , (B'') and Lemma 5.1, we have

$$\begin{aligned} f_\pm^1(0) &= W(t_n) \Psi_\pm^{1,n}(0) \in H_p^{m_0}, \quad g_\pm^1(0) = \partial_t W(t_n) \Psi_\pm^{1,n}(0) \in H_p^{m_0}, \\ \dot{f}_\pm^{1,n}(0) &= W(t_n) \partial_s \Psi_\pm^{1,n}(0) \in H_p^{m_0-2}. \end{aligned}$$

Employing (5.3.8) and Cauchy inequality further, for $m_0 \geq 4$, we can bound $\|\eta_-^1(\cdot)\|_{L^2}$ from (5.3.50) as

$$\begin{aligned} \|\eta_-^1(\cdot)\|_{L^2} &\lesssim \tau h^{m_0} + \tau^2 (h^{m_0} + h^{m_0-2}) \\ &\quad + \tau^2 \sqrt{h \sum_{j=0}^{M-1} \left| W(t_n, x_j) (\partial_s \Psi_+^{1,n}(0, x_j) - \dot{\Psi}_+^{1,n}(0, x_j)) \right|^2} \\ &\quad + \tau^2 \sqrt{h \sum_{j=0}^{M-1} \left| W(t_n, x_j) (\partial_s \Psi_-^{1,n}(0, x_j) - \dot{\Psi}_-^{1,n}(0, x_j)) \right|^2} \\ &\lesssim \tau (\tau^2 + h^{m_0}) \\ &\quad + \tau^2 \left(\|I_M(\partial_s \Psi_+^{1,n}(0)) - \dot{\Psi}_+^{1,n}(0)\|_{L^2} + \|I_M(\partial_s \Psi_-^{1,n}(0)) - \dot{\Psi}_-^{1,n}(0)\|_{L^2} \right) \\ &\lesssim \tau (\tau^2 + h^{m_0}) \\ &\quad + \tau^2 \left(\|P_M(\partial_s \Psi_+^{1,n}(0)) - \dot{\Psi}_+^{1,n}(0)\|_{L^2} + \|P_M(\partial_s \Psi_-^{1,n}(0)) - \dot{\Psi}_-^{1,n}(0)\|_{L^2} \right). \end{aligned} \tag{5.3.51}$$

Using equation (5.3.10), we get

$$\begin{aligned} (\widehat{\partial_s \Psi_+^{1,n}})_l(0) - (\widetilde{\Psi_+^{1,n}})_l(0) &= -i \frac{2 \sin(\mu_l^2 \tau / 2)}{\delta_l^+ \tau} \left((\widehat{\Psi_+^{1,n}(0)})_l - (\widetilde{\Psi_+^{1,n}(0)})_l \right) \\ &\quad - i \Pi_l^+ \left((\widehat{f_+^1(0)})_l - (\widetilde{f_+^1(0)})_l \right) \\ &\quad - i \left(\delta_l^- - \frac{2 \sin(\mu_l^2 \tau / 2)}{\delta_l^+ \tau} \right) (\widehat{\Psi_+^{1,n}(0)})_l, \\ (\widehat{\partial_s \Psi_-^{1,n}})_l(0) - (\widetilde{\Psi_-^{1,n}})_l(0) &= -i \Pi_l^- \left((\widehat{f_+^1(0)})_l - (\widetilde{f_+^1(0)})_l \right), \end{aligned}$$

and

$$\|P_M(\partial_s \Psi_-^{1,n}(0)) - I_M(\dot{\Psi}_-^{1,n}(0))\|_{L^2} \leq \|P_M(f_+^1(0)) - I_M(f_+^1(0))\|_{L^2} \lesssim h^{m_0}.$$

Noticing that $|\sin(s) - s| \leq \frac{s^2}{2}$ ($s \in \mathbb{R}$), we have

$$\left| \delta_l^- - \frac{2 \sin(\mu_l^2 \tau / 2)}{\delta_l^+ \tau} \right| = \frac{2}{\delta_l^+} \left| \frac{1}{2} \mu_l^2 - \frac{\sin(\mu_l^2 \tau / 2)}{\tau} \right| \lesssim \mu_l^4 \tau, \quad l = -M/2, \dots, M/2 - 1,$$

which leads to

$$\begin{aligned} \left| (\widehat{\partial_s \Psi_+^{1,n}})_l(0) - (\widetilde{\Psi_+^{1,n}})_l(0) \right| &\lesssim \frac{1}{\tau} \left| (\widehat{\Psi_+^{1,n}(0)})_l - (\widetilde{\Psi_+^{1,n}(0)})_l \right| + \left| (\widehat{f_+^1(0)})_l - (\widetilde{f_+^1(0)})_l \right| \\ &\quad + \tau \mu_l^4 \left| (\widehat{\Psi_+^{1,n}(0)})_l \right|, \end{aligned}$$

and for $m_0 \geq 4$,

$$\begin{aligned} \|P_M(\partial_s \Psi_+^{1,n}(0)) - I_M(\dot{\Psi}_+^{1,n}(0))\|_{L^2} &\lesssim \frac{1}{\tau} \|P_M(\Psi_+^{1,n}(0)) - I_M(\Psi_+^{1,n}(0))\|_{L^2} \\ &\quad + \tau \|P_M(\Psi_+^{1,n}(0))\|_{H^4} \\ &\quad + \|P_M(f_+^1(0)) - I_M(f_+^1(0))\|_{L^2} \\ &\lesssim h^{m_0} + \tau + h^{m_0} / \tau. \end{aligned}$$

Combing the above estimates with (5.3.51), we derive

$$\|\eta_-^1(\cdot)\|_{L^2} \lesssim \tau(h^{m_0} + \tau^2) + \tau^2(h^{m_0} + h^{m_0}/\tau + \tau) \lesssim \tau(h^{m_0} + \tau^2). \quad (5.3.52)$$

By the same procedure, $\|\eta_+^1(\cdot)\|_{L^2}$ can be bounded as

$$\begin{aligned} \|\eta_+^1(\cdot)\|_{L^2} &\lesssim \tau(\tau^2 + h^{m_0}) + \tau^2 \|P_M(\partial_s \Psi_-^{1,n}(0)) - P_M(\Psi_-^{1,n}(\tau)) / \tau\|_{L^2} \\ &\quad + \tau \|P_M(\Psi_-^{1,n}(\tau)) - I_M(\Psi_-^{1,n}(\tau))\|_{L^2} \end{aligned}$$

where Taylor expansion gives

$$\partial_s \Psi_-^{1,n}(0) - \Psi_-^{1,n}(\tau)/\tau = -\tau \int_0^1 \int_0^s \partial_{s_1 s_1} \Psi_-^{1,n}(s_1 \tau) ds_1 ds.$$

Thus, recalling Lemma 5.1, we estimate

$$\|\eta_+^1(\cdot)\|_{L^2} \lesssim \tau(\tau^2 + h^{m_0}) + \tau^3 \|\partial_{ss} \Psi_-^{1,n}\|_{L^\infty([0,\tau];(L^2)^2)} \lesssim \tau(\tau^2 + h^{m_0} + \tau^2/\varepsilon^2). \quad (5.3.53)$$

Now, Lemma 5.1 together with (5.3.48), (5.3.49), (5.3.50) and (5.3.51) implies

$$\begin{aligned} \|\xi_\pm^{1,n}(\cdot)\|_{L^2} &\lesssim \tau^3 \|\partial_{ss}(W(t_n + s)\Psi_+^{1,n}(s))\|_{L^\infty([0,\tau];(L^2)^2)} \\ &\quad + \tau^3 \|\partial_{ss}(W(t_n + s)\Psi_-^{1,n}(s))\|_{L^\infty([0,\tau];(L^2)^2)} + \|\eta_\pm^1(\cdot)\|_{L^2} \\ &\lesssim \tau \left(\frac{\tau^2}{\varepsilon^2} + h^{m_0} \right). \end{aligned} \quad (5.3.54)$$

Type II estimates. Next, we prove the second estimates for $\xi_\pm^{1,n}(x)$ in (5.3.44). Starting from (5.3.46) and (5.3.47), we treat the terms involving $f_+^1(s)$, $\dot{f}_+^1(s)$ and $g_+^1(s)$ in the same way as in proving (5.3.54), and leave the rest terms as

$$\begin{aligned} \widehat{(\xi_-^{1,n})}_l &= \widehat{(\zeta_-^1)}_l - i \int_0^\tau e^{i\delta_l^+(\tau-s)/\varepsilon^2} \Pi_l^- \widehat{(f_-^1(s))}_l ds + \overline{q_l^+(\tau)} \Pi_l^- \widehat{(f_-^1(0))}_l, \\ \widehat{(\xi_+^{1,n})}_l &= \widehat{(\zeta_+^1)}_l - i \int_0^\tau e^{-i\delta_l^-(\tau-s)} \Pi_l^+ \widehat{(f_-^1(s))}_l ds - q_l^-(\tau) \Pi_l^+ \widehat{(f_-^{1,*}(\tau))}_l, \end{aligned}$$

with $\zeta_\pm^1(x) = \sum_{l=-M/2}^{M/2-1} \widehat{(\zeta_\pm^1)}_l e^{i\mu_l(x-a)}$ satisfying

$$\|\zeta_\pm^1(\cdot)\|_{L^2} \lesssim \tau(h^{m_0} + \tau^2).$$

The proof of the above decomposition and the corresponding error bounds for $\zeta_\pm^1(x)$ is identical to the proof of (5.3.54), and we omit it here for brevity. Applying triangle inequality and (5.3.45), we have

$$\begin{aligned} \left| \widehat{(\xi_-^{1,n})}_l \right| &\leq \left| \widehat{(\zeta_-^1)}_l \right| + \int_0^\tau \left| \widehat{(f_-^1(s))}_l \right| ds + \tau \varepsilon^2 \left| \widehat{(f_-^1(0))}_l \right|, \\ \left| \widehat{(\xi_+^{1,n})}_l \right| &= \left| \widehat{(\zeta_+^1)}_l \right| + \int_0^\tau \left| \widehat{(f_-^1(s))}_l \right| ds + \tau^2 \left| \widehat{(f_-^{1,*}(\tau))}_l \right|, \end{aligned}$$

Recalling Lemma 5.1 which implies $\|\Psi_-^{1,n}(s)\|_{H_p^{m_0-1}} \lesssim \varepsilon^2$, we know $\|f_-^1(s)\|_{L^2} \lesssim \varepsilon^2$, $\|\dot{f}_-^{1,*}(\tau)\|_{H_p^{m_0-1}} \lesssim \varepsilon^2/\tau$ and

$$\|f_-^1(0)\|_{H^{m_0-1}} \lesssim \|\dot{\Psi}_-^{1,n}(0)\|_{H^{m_0-1}} \lesssim \|I_M(f_+^1(0))\|_{H^{m_0-1}} \lesssim \|f_+^1(0)\|_{H^{m_0}} \lesssim 1.$$

Hence, using Parseval's theorem, we find

$$\begin{aligned} \|\xi_-^{1,n}(\cdot)\|_{L^2} &\lesssim \|\zeta_-^1(\cdot)\|_{L^2} + \tau \|f_-^1\|_{L^\infty([0,\tau];(L^2)^2)} + \tau \varepsilon^2 \|I_M(\dot{f}_-^1(0))\|_{L^2} \lesssim \tau(\tau^2 + h^{m_0} + \varepsilon^2), \\ \|\xi_+^{1,n}(\cdot)\|_{L^2} &\lesssim \|\zeta_+^1(\cdot)\|_{L^2} + \tau \|f_-^1\|_{L^\infty([0,\tau];(L^2)^2)} + \tau^2 \|I_M(\dot{f}_-^{1,*}(\tau))\|_{L^2} \lesssim \tau(\tau^2 + h^{m_0} + \varepsilon^2), \end{aligned}$$

which completes the proof for (5.3.45).

Thus, we have established error bounds (5.3.44) for $\xi_\pm^{1,n}$. $\xi_\pm^{2,n}$ can be controlled in the same way and the proof is omitted. \square

Subtracting (5.3.26) from (5.3.40), noticing (5.3.27) and (5.3.41), we get the error equations for $\mathbf{z}_\pm^{k,n+1}(x)$ (5.3.35) for $k = 1, 2$ as

$$\begin{cases} (\widetilde{\mathbf{z}}_-^{1,n+1})_l = (\widetilde{\mathcal{F}}_-^{1,n})_l + (\widetilde{\xi}_-^{1,n})_l, & (\widetilde{\mathbf{z}}_+^{2,n+1})_l = (\widetilde{\mathcal{F}}_+^{2,n})_l + (\widetilde{\xi}_+^{2,n})_l, \\ (\widetilde{\mathbf{z}}_+^{1,n+1})_l = e^{-i\frac{\delta_l^- \tau}{\varepsilon^2}} \Pi_l^+ (\widetilde{\mathbf{e}}^n)_l + (\widetilde{\mathcal{F}}_+^{1,n})_l + (\widetilde{\xi}_+^{1,n})_l, \\ (\widetilde{\mathbf{z}}_-^{2,n+1})_l = e^{i\frac{\delta_l^- \tau}{\varepsilon^2}} \Pi_l^- (\widetilde{\mathbf{e}}^n)_l + (\widetilde{\mathcal{F}}_-^{2,n})_l + (\widetilde{\xi}_-^{2,n})_l, \end{cases} \quad (5.3.55)$$

where $\mathcal{F}_\pm^{k,n}(x) = \sum_{l=-M/2}^{M/2-1} (\mathcal{F}_\pm^{k,n})_l e^{i\mu_l(x-a)}$ ($k = 1, 2$) is given by

$$\begin{cases} (\widetilde{\mathcal{F}}_-^{1,n})_l &= -\overline{p_l^+}(\tau) \Pi_l^- (\widetilde{F}_+^{1,n})_l - \overline{q_l^+}(\tau) \Pi_l^- (\widetilde{G}_+^{1,n})_l - \overline{q_l^+}(\tau) \Pi_l^- \left((\widetilde{\dot{F}}_+^{1,n})_l + (\widetilde{\dot{F}}_-^{1,n})_l \right), \\ (\widetilde{\mathcal{F}}_+^{2,n})_l &= p_l^+(\tau) \Pi_l^+ (\widetilde{F}_-^{2,n})_l + q_l^+(\tau) \Pi_l^+ (\widetilde{G}_-^{2,n})_l + q_l^+(\tau) \Pi_l^+ \left((\widetilde{\dot{F}}_+^{2,n})_l + (\widetilde{\dot{F}}_-^{2,n})_l \right), \\ (\widetilde{\mathcal{F}}_+^{1,n})_l &= p_l^-(\tau) \Pi_l^+ (\widetilde{F}_+^{1,n})_l + q_l^-(\tau) \Pi_l^+ (\widetilde{G}_+^{1,n})_l + q_l^-(\tau) \Pi_l^+ \left((\widetilde{\dot{F}}_+^{1,n})_l + (\widetilde{\dot{F}}_-^{1,*})_l \right), \\ (\widetilde{\mathcal{F}}_-^{2,n})_l &= -\overline{p_l^-}(\tau) \Pi_l^- (\widetilde{F}_-^{2,n})_l - \overline{q_l^-}(\tau) \Pi_l^- (\widetilde{G}_-^{2,n})_l - \overline{q_l^-}(\tau) \Pi_l^- \left((\widetilde{\dot{F}}_+^{2,*})_l + (\widetilde{\dot{F}}_-^{2,n})_l \right), \end{cases} \quad (5.3.56)$$

with

$$\begin{aligned}\dot{F}_{\pm}^{k,n}(x) &= \sum_{l=-M/2}^{M/2-1} \widetilde{(\dot{F}_{\pm}^{k,n})}_l e^{i\mu_l(x-a)} \in Y_M \quad k = 1, 2, \\ G_{\pm}^{k_{\pm},n}(x) &= \sum_{l=-M/2}^{M/2-1} \widetilde{(G_{\pm}^{k_{\pm},n})}_l e^{i\mu_l(x-a)} \in Y_M, \\ F_{\pm}^{k_{\pm},n}(x) &= \sum_{l=-M/2}^{M/2-1} \widetilde{(F_{\pm}^{k_{\pm},n})}_l e^{i\mu_l(x-a)} \in Y_M \\ \dot{F}_{\pm}^{k_{\mp},*}(x) &= \sum_{l=-M/2}^{M/2-1} \widetilde{(\dot{F}_{\pm}^{k_{\mp},*})}_l e^{i\mu_l(x-a)} \in Y_M\end{aligned}$$

($k_+ = 1, k_- = 2$) defined as

$$\begin{aligned}\widetilde{(F_{\pm}^{k_{\pm},n})}_l &= \widetilde{(f_{\pm}^{k_{\pm}}(0))}_l - \widetilde{(f_{\pm}^{k_{\pm}})}_l, \quad \widetilde{(\dot{F}_{\pm}^{k,n})}_l = \widetilde{(\dot{f}_{\pm}^k(0))}_l - \widetilde{(\dot{f}_{\pm}^k)}_l, \quad \widetilde{(G_{\pm}^{k_{\pm},n})}_l = \widetilde{(g_{\pm}^{k_{\pm}}(0))}_l - \widetilde{(g_{\pm}^{k_{\pm}})}_l, \\ \widetilde{(\dot{F}_{-}^{1,*})}_l &= \widetilde{(f_{-}^{1,*}(0))}_l - \widetilde{(f_{-}^{1,*})}_l, \quad \widetilde{(\dot{F}_{+}^{2,*})}_l = \widetilde{(f_{+}^{2,*}(0))}_l - \widetilde{(f_{+}^{2,*})}_l.\end{aligned}\tag{5.3.57}$$

For the electromagnetic error part $\mathcal{F}_{\pm}^{k,n}(x)$ ($k = 1, 2, 0 \leq n \leq \frac{T}{\tau} - 1$), we have the lemma below.

Lemma 5.3 *Under assumption (A'') and (B''), the electromagnetic error part $\mathcal{F}_{\pm}^{k,n}(x) \in Y_M$ ($k = 1, 2, 0 \leq n \leq \frac{T}{\tau} - 1$) defined in (5.3.56) with (5.3.57) satisfies the bounds as*

$$\begin{aligned}\|F_{\pm}^{k_{\pm},n}(\cdot)\|_{L^2} + \|G_{\pm}^{k_{\pm},n}(\cdot)\|_{L^2} + \|\dot{F}_{\pm}^{3-k_{\pm},n}(\cdot)\|_{L^2} &\lesssim h^{m_0} + \|\mathbf{e}^n(\cdot)\|_{L^2}, \quad k_+ = 1, \quad k_- = 2, \\ \|\dot{F}_{+}^{1,n}(\cdot)\|_{L^2} + \|\dot{F}_{-}^{2,n}(\cdot)\|_{L^2} &\lesssim \frac{1}{\tau}(h^{m_0} + \|\mathbf{e}^n(\cdot)\|_{L^2}), \\ \|\dot{F}_{\pm}^{k_{\mp},*}(\cdot)\|_{L^2} &\lesssim \frac{1}{\tau}(h^{m_0} + \|\mathbf{z}_{\pm}^{k_{\mp},n+1}(\cdot)\|_{L^2}),\end{aligned}$$

which implies that

$$\begin{aligned}\|\mathcal{F}_{\pm}^{k_{\pm},n}(\cdot)\|_{L^2} &\lesssim \tau(h^{m_0} + \|\mathbf{z}_{\mp}^{k_{\pm},n+1}(\cdot)\|_{L^2} + \|\mathbf{e}^n(\cdot)\|_{L^2}), \\ \|\mathcal{F}_{\pm}^{k_{\mp},n}(\cdot)\|_{L^2} &\lesssim \tau(h^{m_0} + \|\mathbf{e}^n(\cdot)\|_{L^2}).\end{aligned}\tag{5.3.58}$$

Proof: Recalling assumptions (A'') and (B'') , Lemma 5.1, (5.3.57), (5.3.41), (5.3.43), (5.3.27) and (5.3.28), applying Parseval's theorem, we can derive that

$$\begin{aligned} \|F_+^{1,n}(\cdot)\|_{L^2}^2 &\leq \|I_M(f_+^1(0)) - I_M(f_+^1)\|_{L^2}^2 = h \sum_{j=0}^{M-1} |W(t_n, x_j)(\Psi_+^{1,n}(0, x_j) - \Psi_{+,j}^k)|^2 \\ &\lesssim h \sum_{j=0}^{M-1} |\Psi_+^{1,n}(0, x_j) - \Psi_{+,j}^1|^2 \lesssim h^{2m_0} + \|P_M(\Phi(t_n)) - I_M(\Phi^n)\|_{L^2}^2 \\ &\lesssim h^{2m_0} + \|\mathbf{e}^n(\cdot)\|_{L^2}^2, \end{aligned}$$

and similarly we have $\|F_-^{2,n}(\cdot)\|_{L^2} \lesssim h^{m_0} + \|\mathbf{e}^n(\cdot)\|_{L^2}$. Using the same idea, we can obtain

$$\|G_+^{1,n}(\cdot)\|_{L^2} + \|G_-^{2,n}(\cdot)\|_{L^2} \lesssim h^{m_0} + \|\mathbf{e}^n(\cdot)\|_{L^2},$$

and

$$\begin{aligned} \|F_-^{1,*}(\cdot)\|_{L^2} &\lesssim \|I_M(f_-^{1,*}(\tau)) - I_M(f_-^{1,*})\|_{L^2} \lesssim \frac{1}{\tau} \|I_M(\Psi_-^{1,n}(\tau)) - I_M(\Psi_-^{1,n+1})\|_{L^2} \\ &\lesssim \frac{1}{\tau} (h^{m_0} + \|\mathbf{z}_-^{1,n+1}\|_{L^2}), \\ \|F_+^{2,*}(\cdot)\|_{L^2} &\lesssim \|I_M(f_+^{2,*}(\tau)) - I_M(f_+^{2,*})\|_{L^2} \lesssim \frac{1}{\tau} (h^{m_0} + \|\mathbf{z}_+^{2,n+1}\|_{L^2}). \end{aligned}$$

It remains to estimate $\|\dot{F}_\pm^{k,n}(\cdot)\|_{L^2}$. Again, in the same spirit of the above arguments, we arrive at

$$\|\dot{F}_\pm^{k,n}(\cdot)\|_{L^2} \lesssim \|I_M(\dot{\Psi}_\pm^{k,n}(0)) - I_M(\dot{\Psi}_\pm^k)\|_{L^2}. \quad (5.3.59)$$

Comparing (5.3.43) with (5.3.28), noticing the properties of δ_t^\pm and the arguments in the above proof, we find

$$\begin{aligned} \|I_M(\dot{\Psi}_-^{1,n}(0)) - I_M(\dot{\Psi}_-^1)\|_{L^2} &\lesssim \|I_M(f_+^1(0)) - I_M(f_+^1)\|_{L^2} \lesssim h^{m_0} + \|\mathbf{e}^n(\cdot)\|_{L^2}, \\ \|I_M(\dot{\Psi}_+^{2,n}(0)) - I_M(\dot{\Psi}_+^2)\|_{L^2} &\lesssim \|I_M(f_-^2(0)) - I_M(f_-^2)\|_{L^2} \lesssim h^{m_0} + \|\mathbf{e}^n(\cdot)\|_{L^2}, \\ \|I_M(\dot{\Psi}_+^{1,n}(0)) - I_M(\dot{\Psi}_+^1)\|_{L^2} &\lesssim \frac{1}{\tau} \|I_M(\Psi_+^{1,n}(0)) - I_M(\Psi_+^1)\|_{L^2} + \|I_M(f_+^1(0)) - I_M(f_+^1)\|_{L^2} \\ &\lesssim \frac{1}{\tau} (h^{m_0} + \|\mathbf{e}^n(\cdot)\|_{L^2}) + (h^{m_0} + \|\mathbf{e}^n(\cdot)\|_{L^2}), \\ \|I_M(\dot{\Psi}_-^{2,n}(0)) - I_M(\dot{\Psi}_-^2)\|_{L^2} &\lesssim \frac{1}{\tau} \|I_M(\Psi_-^{2,n}(0)) - I_M(\Psi_-^2)\|_{L^2} + \|I_M(f_-^2(0)) - I_M(f_-^2)\|_{L^2} \\ &\lesssim \frac{1}{\tau} (h^{m_0} + \|\mathbf{e}^n(\cdot)\|_{L^2}) + (h^{m_0} + \|\mathbf{e}^n(\cdot)\|_{L^2}), \end{aligned}$$

which implies the bounds for $\|\dot{F}_\pm^{k,n}(\cdot)\|_{L^2}$ in view of (5.3.59). Combine all the results above, in view of (5.3.57) and properties of the coefficients $p_l^\pm(\tau), q_l^\pm(\tau)$ (5.3.45), we conclude that (5.3.58) holds. \square

Now, we are ready to prove the main theorem.

Proof of Theorem 5.1. Recalling the decomposition (5.3.34) and error equation (5.3.55), we immediately get

$$\begin{aligned} (\widetilde{\mathbf{e}^{n+1}})_l &= e^{-i\tau/\varepsilon^2} \left(\widetilde{\mathbf{z}_+^{1,n+1}} + \widetilde{\mathbf{z}_-^{1,n+1}} \right) + e^{i\tau/\varepsilon^2} \left(\widetilde{\mathbf{z}_+^{2,n+1}} + \widetilde{\mathbf{z}_-^{2,n+1}} \right) \\ &= \left(e^{-i\delta_l/\varepsilon^2} \Pi_l^+ (\widetilde{\mathbf{e}^n})_l + e^{i\delta_l/\varepsilon^2} \Pi_l^- (\widetilde{\mathbf{e}^n})_l \right) + (\widetilde{\chi^n})_l, \end{aligned} \quad (5.3.60)$$

with $\chi^n(x) = \sum_{l=-M/2}^{M/2-1} (\widetilde{\chi^n})_l e^{i\mu_l(x-a)} \in Y_M$ given as

$$(\widetilde{\chi^n})_l = \sum_{k=1,2} e^{i\tau(2k-3)/\varepsilon^2} \left((\widetilde{\mathcal{F}_+^{k,n}})_l + (\widetilde{\mathcal{F}_-^{k,n}})_l + (\widetilde{\xi_+^{k,n}})_l + (\widetilde{\xi_-^{k,n}})_l \right), \quad l = -\frac{M}{2}, \dots, \frac{M}{2} - 1. \quad (5.3.61)$$

In particular $\|\mathbf{e}^0\|_{L^2} = \|P_M(\Phi_0) - I_M(\Phi_0)\|_{L^2} \lesssim h^{m_0}$.

Taking the l^2 norm of the vectors in the error equation (5.3.55), then summing together for $l = -M/2, \dots, M/2 - 1$, utilizing Lemma 5.3 and Parserval's theorem, there holds for $\tau \leq 1$,

$$\begin{aligned} \|\mathbf{z}_-^{1,n+1}(\cdot)\|_{L^2} &\lesssim \|\mathcal{F}_-^{1,n}(\cdot)\|_{L^2} + \|\xi_-^{1,n}(\cdot)\|_{L^2} \lesssim \tau(h^{m_0} + \|\mathbf{e}^n(\cdot)\|_{L^2}) + \|\xi_-^{1,n}(\cdot)\|_{L^2}, \\ \|\mathbf{z}_+^{2,n+1}(\cdot)\|_{L^2} &\lesssim \|\mathcal{F}_+^{2,n}(\cdot)\|_{L^2} + \|\xi_+^{2,n}(\cdot)\|_{L^2} \lesssim \tau(h^{m_0} + \|\mathbf{e}^n(\cdot)\|_{L^2}) + \|\xi_+^{2,n}(\cdot)\|_{L^2}, \end{aligned}$$

and so

$$\begin{aligned} \|\chi^n(\cdot)\| &\lesssim \tau(h^{m_0} + \|\mathbf{e}^n(\cdot)\|_{L^2}) + \|\mathbf{z}_-^{1,n+1}(\cdot)\|_{L^2} + \|\mathbf{z}_+^{2,n+1}(\cdot)\|_{L^2} \\ &\quad + \sum_{k=1,2} \left(\|\xi_+^{k,n}(\cdot)\|_{L^2} + \|\xi_-^{k,n}(\cdot)\|_{L^2} \right) \\ &\lesssim \tau(h^{m_0} + \|\mathbf{e}^n(\cdot)\|_{L^2}) + \sum_{k=1,2} \left(\|\xi_+^{k,n}(\cdot)\|_{L^2} + \|\xi_-^{k,n}(\cdot)\|_{L^2} \right). \end{aligned}$$

From Lemma 5.2 on local truncation error $\xi_{\pm}^{k,n}(x)$, we get

$$\|\chi^n(\cdot)\|_{L^2} \lesssim \tau \|\mathbf{e}^n(\cdot)\|_{L^2} + \tau(h^{m_0} + \tau^2/\varepsilon^2), \quad 0 \leq n \leq \frac{T}{\tau} - 1, \quad (5.3.62)$$

$$\|\chi^n(\cdot)\|_{L^2} \lesssim \tau \|\mathbf{e}^n(\cdot)\|_{L^2} + \tau(h^{m_0} + \tau^2 + \varepsilon^2), \quad 0 \leq n \leq \frac{T}{\tau} - 1. \quad (5.3.63)$$

Now, taking the l^2 norm of the vectors on both sides of (5.3.60), making using of the orthogonal properties of Π_l^{\pm} where $|e^{i\theta_1}\Pi_l^+\mathbf{v} + e^{i\theta_2}\Pi_l^-\mathbf{v}|^2 = |\Pi_l^+\mathbf{v}|^2 + |\Pi_l^-\mathbf{v}|^2 = |\mathbf{v}|^2$ for all $\mathbf{v} \in \mathbb{C}^2$, $\theta_1, \theta_2 \in \mathbb{R}$, we can get

$$\begin{aligned} \left| \widetilde{(\mathbf{e}^{n+1})}_l \right|^2 &= \left| e^{-i\delta_l/\varepsilon^2} \Pi_l^+ \widetilde{(\mathbf{e}^n)}_l + e^{i\delta_l/\varepsilon^2} \Pi_l^- \widetilde{(\mathbf{e}^n)}_l \right|^2 + |\widetilde{(\chi^n)}_l|^2 \\ &\quad + 2\operatorname{Re} \left(e^{-i\delta_l/\varepsilon^2} \Pi_l^+ \widetilde{(\mathbf{e}^n)}_l + e^{i\delta_l/\varepsilon^2} \Pi_l^- \widetilde{(\mathbf{e}^n)}_l \right)^* \widetilde{(\chi^n)}_l \\ &= |\widetilde{(\mathbf{e}^n)}_l|^2 + |\widetilde{(\chi^n)}_l|^2 + 2\operatorname{Re} \left(e^{-i\delta_l/\varepsilon^2} \Pi_l^+ \widetilde{(\mathbf{e}^n)}_l + e^{i\delta_l/\varepsilon^2} \Pi_l^- \widetilde{(\mathbf{e}^n)}_l \right)^* \widetilde{(\chi^n)}_l \end{aligned}$$

where $v^* = \overline{v^T}$ is the complex conjugate of vector v and $\operatorname{Re}(c)$ denotes the real part of complex number c . Applying Cauchy inequality, we find

$$\left| \widetilde{(\mathbf{e}^{n+1})}_l \right|^2 - \left| \widetilde{(\mathbf{e}^n)}_l \right|^2 \lesssim \tau |\widetilde{(\mathbf{e}^n)}_l|^2 + \frac{1}{\tau} |\widetilde{(\chi^n)}_l|^2, \quad l = -\frac{M}{2}, \dots, \frac{M}{2} - 1. \quad (5.3.64)$$

Summing (5.3.64) together for $l = -\frac{M}{2}, \dots, \frac{M}{2} - 1$ and using Parseval's theorem, we obtain

$$\|\mathbf{e}^{n+1}(\cdot)\|_{L^2}^2 - \|\mathbf{e}^n(\cdot)\|_{L^2}^2 \lesssim \tau \|\mathbf{e}^n(\cdot)\|_{L^2}^2 + \frac{1}{\tau} \|\chi^n(\cdot)\|_{L^2}^2, \quad 0 \leq n \leq \frac{T}{\tau} - 1. \quad (5.3.65)$$

Summing (5.3.65) for indices $1, 2, \dots, n$ and using (5.3.62)-(5.3.63), we conclude that for $0 \leq n \leq \frac{T}{\tau} - 1$,

$$\|\mathbf{e}^{n+1}(\cdot)\|_{L^2}^2 - \|\mathbf{e}^0(\cdot)\|_{L^2}^2 \lesssim \tau \sum_{m=1}^n \|\mathbf{e}^m(\cdot)\|_{L^2}^2 + n\tau(h^{m_0} + \tau^2/\varepsilon^2)^2, \quad (5.3.66)$$

$$\|\mathbf{e}^{n+1}(\cdot)\|_{L^2}^2 - \|\mathbf{e}^0(\cdot)\|_{L^2}^2 \lesssim \tau \sum_{m=1}^n \|\mathbf{e}^m(\cdot)\|_{L^2}^2 + n\tau(h^{m_0} + \tau^2 + \varepsilon^2)^2. \quad (5.3.67)$$

Since $\|\mathbf{e}^0(\cdot)\|_{L^2} \lesssim h^{m_0}$, Gronwall's inequality will lead to the conclusion that for sufficiently small τ ,

$$\|\mathbf{e}^{n+1}(\cdot)\|_{L^2}^2 \lesssim (h^{m_0} + \tau^2/\varepsilon^2)^2, \quad \|\mathbf{e}^{n+1}(\cdot)\|_{L^2}^2 \lesssim (h^{m_0} + \tau^2 + \varepsilon^2)^2, \quad 0 \leq n \leq \frac{T}{\tau} - 1. \quad (5.3.68)$$

In view of (5.3.32), we conclude that (5.3.29) holds. \square

5.4 Numerical examples

In this section, we present numerical tests on our MTI-FP scheme (5.3.24) and use MTI-FP to study the convergence of the Dirac equation (5.1.3) to the limiting Schrödinger model (5.1.2) and second order model of the Pauli's equation kind. To this purpose, we choose the electromagnetic potentials in the Dirac equation (5.1.3) with $d = 1$ as

$$A_1(t, x) = \frac{(x+1)^2}{1+x^2}, \quad V(t, x) = \frac{1-x}{1+x^2}, \quad x \in \mathbb{R}, \quad t \geq 0,$$

and the initial value as

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

5.4.1 Accuracy test

The problem is solved numerically on an interval $\Omega = (-16, 16)$, i.e. $a = -16$ and $b = 16$, with periodic boundary conditions on $\partial\Omega$. The 'exact' solution $\Phi(t, x) = (\phi_1(t, x), \phi_2(t, x))^T$ is obtained numerically by using the TSFP method with a very fine mesh size and a small time step, e.g. $h_e = 1/32$ and $\tau_e = 10^{-7}$. Denote $\Phi_{h,\tau}^n$ as the numerical solution obtained by a numerical method with mesh size h and time step τ . In order to quantify the convergence, we introduce

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

The spatial errors and temporal errors are displayed in Tabs. 5.1 and 5.2, respectively. From Tabs. 5.1-5.2 and additional numerical results not shown here for brevity, we can draw the following conclusions for the MTI-FP method:

(i) For spatial discretization error, the MTI-FP is uniformly spectral accurate for all $\varepsilon \in (0, 1]$ (cf. Tab. 5.1).

(ii) For temporal discretization error, the MTI-FP is uniformly convergent with linear rate. When time step τ is small (upper triangle part of Tab. 5.2), second order

Table 5.1: Spatial error analysis of the MD-EWI-FP method for the Dirac equation in 1D.

Spatial Errors	$h_0 = 2$	$h_0/2$	$h_0/2^2$	$h_0/2^3$	$h_0/2^4$
$\varepsilon_0 = 1$	1.65	5.74E-1	7.08E-2	7.00E-5	8.53E-10
$\varepsilon_0/2$	1.39	3.45E-1	7.06E-3	6.67E-6	9.71E-10
$\varepsilon_0/2^2$	1.18	1.67E-1	1.71E-3	1.43E-6	1.10E-9
$\varepsilon_0/2^3$	1.13	1.46E-1	1.03E-3	6.77E-7	9.16E-10
$\varepsilon_0/2^4$	1.15	1.45E-1	8.52E-4	4.86E-7	1.33E-9

convergence is clear; when ε is small (lower triangle part of Tab. 5.2), second order convergence is also clear; near the diagonal part where $\tau \sim \varepsilon^2$ (cf. the underlined diagonal part of 5.2), degeneracy of the convergence rate and the uniform linear convergence rate for the temporal error are observed. In particular, the underlined errors in Tab. 5.2 degenerate in the parameter regime $\tau \sim \varepsilon^2$ predicted by our error estimates (5.3.29).

5.4.2 Convergence of Dirac equation to the limiting model

Similar to (5.1.2), the Schrödinger first order model for the Dirac equation (5.1.3) as $\varepsilon \rightarrow 0^+$ reads

$$\Phi(t, x) = e^{-it/\varepsilon^2} \phi_e \mathbf{e}_1 + e^{it/\varepsilon^2} \phi_p \mathbf{e}_2 + O(\varepsilon), \quad \mathbf{e}_1 = (1, 0)^T, \quad \mathbf{e}_2 = (0, 1)^T, \quad (5.4.1)$$

where $\phi_e := \phi_e(t, x) \in \mathbb{C}$ and $\phi_p := \phi_p(t, x) \in \mathbb{C}$ satisfy the Schrödinger equations [20, 73, 85, 105],

$$i\partial_t \phi_e = \left[-\frac{1}{2} \Delta + V(t, x) \right] \phi_e, \quad i\partial_t \phi_p = \left[\frac{1}{2} \Delta + V(t, x) \right] \phi_p, \quad x \in \mathbb{R}, \quad (5.4.2)$$

and the initial data is determined through (5.4.1).

To obtain a second order model of Pauli's equation type, we formally drop the

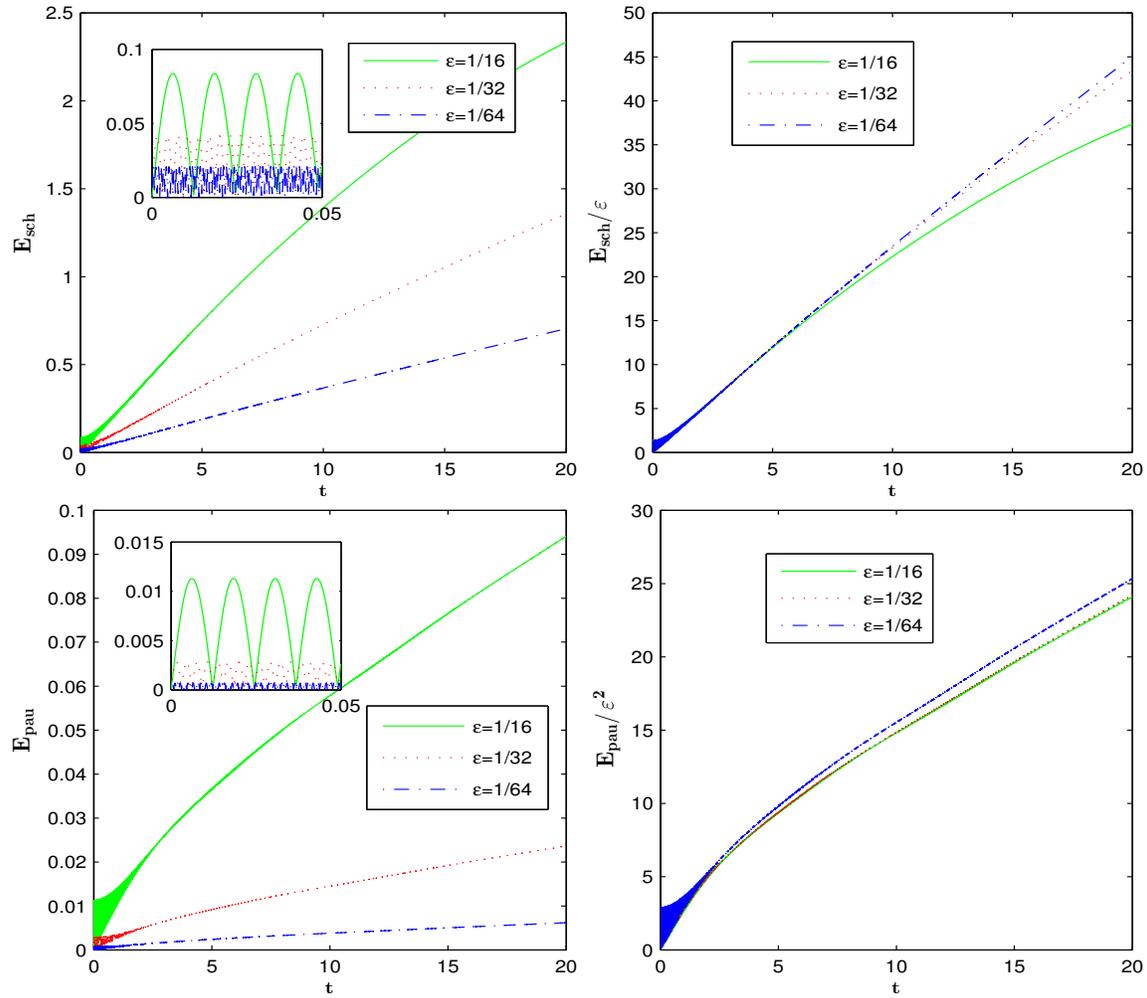
Figure 5.1: Error functions $E_{\text{sch}}(t)$ and $E_{\text{pau}}(t)$ for different ε .

Table 5.2: Temporal error analysis of the MD-EWI-FP method for the Dirac equation.

Table 1 Temporal error at $T=2$ with $\tau_0 = 0.1/16$

Mesh	$\tau_0 = 0.1$	$\tau_0/2$	$\tau_0/2^2$	$\tau_0/2^3$	$\tau_0/2^4$	$\tau_0/2^5$	$\tau_0/2^6$
$\varepsilon_0 = 1$	3.69E-2	9.18E-3	2.29E-3	5.73E-4	1.43E-4	3.58E-5	8.94E-6
$\varepsilon_0/2$	5.98E-2	1.51E-2	3.77E-3	9.45E-4	2.36E-4	5.90E-5	1.48E-5
$\varepsilon_0/2^2$	<u>1.91E-1</u>	5.67E-2	1.47E-2	3.74E-3	9.39E-4	2.35E-4	5.87E-5
$\varepsilon_0/2^3$	7.12E-2	<u>7.17E-2</u>	<u>4.90E-2</u>	1.48E-2	3.89E-3	9.84E-4	2.47E-4
$\varepsilon_0/2^4$	1.78E-2	1.76E-2	1.80E-2	<u>1.82E-2</u>	<u>1.22E-2</u>	3.73E-3	9.79E-4
$\varepsilon_0/2^5$	7.11E-3	3.30E-3	4.07E-3	4.43E-3	4.53E-3	<u>4.56E-3</u>	<u>3.05E-3</u>
$\varepsilon_0/2^6$	7.19E-3	1.99E-3	5.10E-4	6.84E-4	1.02E-3	1.10E-3	1.13E-3
$\varepsilon_0/2^7$	7.07E-3	1.70E-3	4.49E-4	2.61E-4	8.81E-5	1.68E-4	2.54E-4
$\varepsilon_0/2^8$	7.05E-3	1.71E-3	4.23E-4	1.09E-4	3.91E-5	6.01E-5	2.18E-5
$\varepsilon_0/2^9$	7.05E-3	1.71E-3	4.22E-4	1.05E-4	2.61E-5	1.37E-5	6.98E-6

small components in (5.3.3)-(5.3.4) to get

$$\Phi(t, x) = e^{-it/\varepsilon^2} \Psi_e(t, x) + e^{it/\varepsilon^2} \Psi_p(t, x) + O(\varepsilon^2), \quad (5.4.3)$$

where $\Psi_e := \Psi_e(t, x) \in \mathbb{C}^2$ and $\Psi_p := \Psi_p(t, x) \in \mathbb{C}^2$ satisfy the equations

$$i\partial_t \Psi_e = \frac{1}{\varepsilon^2} \mathcal{D} \Psi_e + \Pi_+ (W \Psi_e), \quad i\partial_t \Psi_p = -\frac{1}{\varepsilon^2} \mathcal{D} \Psi_p + \Pi_- (W \Psi_p), \quad (5.4.4)$$

with $\mathcal{D} = \sqrt{1 - \varepsilon^2 \Delta} - 1$ and initial value as

$$\Psi_e(0, x) = \Pi_+ \Phi(0, x), \quad \Psi_p(0, x) = \Pi_- \Phi(0, x). \quad (5.4.5)$$

To investigate the convergence order of the above limiting models (5.4.1) and (5.4.3) numerically, we solve the Schrödinger equation (5.4.2) to obtain (ϕ_e, ϕ_p) and Pauli type equation (5.4.4) to get (Ψ_e, Ψ_p) , by TSFP method [12] and EWI-FP method [15], respectively. Then, the solution Φ to the Dirac equation (5.1.3) is computed by the MTI-FP method and we can study the convergence rate of Dirac equation (5.1.3) to (5.4.1) or (5.4.3). All the computations are done on the bounded interval

$\Omega = [-512, 512]$ with fine mesh $h = 1/16$ and time step $\tau = 10^{-4}$. In order to quantify the convergence, we introduce the error functions

$$E_{\text{sch}}(t) = \left\| \Phi(t, \cdot) - e^{-it/\varepsilon^2} \phi_e(t, \cdot) \mathbf{e}_1 - e^{it/\varepsilon^2} \phi_p(t, \cdot) \mathbf{e}_2 \right\|_{L^2},$$

$$E_{\text{pau}}(t) = \left\| \Phi(t, \cdot) - e^{-it/\varepsilon^2} \Psi_e(t, \cdot) - e^{it/\varepsilon^2} \Psi_p(t, \cdot) \right\|_{L^2}.$$

Fig. 5.1 depicts the evolution of the errors $E_{\text{sch}}(t)$ and $E_{\text{pau}}(t)$, and we can conclude that Schrödinger model (5.4.1) is linearly $O(\varepsilon)$ accurate, while the (5.4.3) model is quadratically $O(\varepsilon^2)$ accurate. In particular, both the errors $E_{\text{sch}}(t)$ and $E_{\text{pau}}(t)$ are observed to grow linearly in time, i.e.

$$E_{\text{sch}}(t) \leq C_1(1+t)\varepsilon, \quad E_{\text{pau}}(t) \leq C_2(1+t)\varepsilon^2.$$

We find that (5.4.4) is the same second order approximate model as Pauli equation [67, 73] for the Dirac equation (5.1.3) in the nonrelativistic limit.

Chapter 6

Concluding remarks and future work

This thesis is devoted to study efficient and accurate numerical methods for solving the Dirac equation in the nonrelativistic limit regime with focus on proposing and analyzing the multiscale time integrator methods. The study focus on dynamics, the rigorous error bounds and how they related to the nonrelativistic limit parameter $0 < \varepsilon \leq 1$. In this regime, the solution propagates waves with wavelength $O(\varepsilon^2)$ and amplitude $O(1)$ in time, which will cost greatly in computations.

In Chapter 2, three types of numerical methods based on different time integrations were analyzed rigorously and compared numerically for simulating the Dirac equation in the nonrelativistic limit regime, i.e. $0 < \varepsilon \ll 1$ or the speed of light goes to infinity. The first class consists of the second order standard FDTD methods, including energy conservative/ nonconservative and implicit/semi-implicit/explicit ones. In the nonrelativistic limit regime, the error estimates of the FDTD methods were rigorously analyzed, which suggest that the ε -scalability of the FDTD methods is $\tau = O(\varepsilon^3)$ and $h = O(\sqrt{\varepsilon})$. The second class applies the Fourier spectral discretization in space and Gautschi-type integration in time, resulting in an EWI-FP method. Rigorous error bounds for the EWI-FP method were derived, which show that the ε -scalability of the EWI-FP method is $\tau = O(\varepsilon^2)$ and $h = O(1)$. The last class combines the Fourier spectral discretization in space and splitting technique in time, which leads to a TSFP method. Based on the rigorous error analysis,

the ε -scalability of the TSFP method is $\tau = O(\varepsilon^2)$ and $h = O(1)$, which is similar to the EWI-FP method. From the error analysis and numerical results, the EWI-FP and TSFP methods perform much better than the FDTD methods, especially in the nonrelativistic limit regime. Extensive numerical results indicate that the TSFP method is superior than the EWI-FP in terms of accuracy and efficiency, and thus the TSFP method is favorable for solving the Dirac equation directly, especially in the nonrelativistic limit regime. Finally, we studied the dynamics of the Dirac equation in 2D with a honeycomb lattice potential and observed some interesting dynamics for different ε .

In Chapter 3, the three types of numerical methods mentioned in Chapter 2 were extended to solving the NLD equation in the nonrelativistic limit regime. Linear stability and convergence analyses were carried out as well as the conservative properties among different methods. Rigorous error bounds showed the ε -resolutions for FDTD, EWI-FP and TSFP are similar as that for the linear case in Chapter 2. Extensive numerical results confirmed this conclusion for the NLD in the nonrelativistic limit regime.

Chapter 4 studied a new class of 4th order compact time splitting methods for solving the Dirac equation. This class of methods is characterized by factorized with purely positive coefficients for the Dirac evolution operator. Numerical comparisons among second order Strang splitting, fourth order Forest-Ruth methods and these new ones were presented, showed that the fourth order compact operator splitting methods enjoys a larger convergence step and smaller error bounds than the general fourth order operator splitting methods, such as FR method.

In the remaining Chapter, a multiscale time integrator Fourier pseudospectral method (MTI-FP) was proposed and rigorously analyzed for the Dirac equation involving a dimensionless parameter $\varepsilon \in (0, 1]$. It aimed to overcome the difficulty of highly oscillatory waves with $O(\varepsilon^2)$ wavelength in time in the nonrelativistic limit regime, i.e. $\varepsilon \rightarrow 0^+$. Rigorous error analysis showed that the MTI-FP is uniformly convergent in spatial discretization with spectral accuracy, and uniformly

convergent in temporal discretization with linear order for $\varepsilon \in (0, 1]$, while the temporal accuracy is optimal with quadratic convergence rate for $\varepsilon = O(1)$ or $\varepsilon \leq \tau$. This result significantly improves the error bounds of the existing numerical methods for the Dirac equation in the nonrelativistic limit regime. The key ideas included a proper multiscale decomposition of the Dirac equation and the use of Gautschi type EWI in time. Numerical results confirmed the error estimates and suggested our error bounds are sharp. Convergence rates of the Dirac equation to its limiting first order Schrödinger equation model and second order Pauli type equation model were also verified at last.

The present work was focusing on reviewing existing numerical methods and designing a uniformly accurate one for solving the Dirac equation in the nonrelativistic limit regime. However, a uniformly accurate method for NLD equation is still lacking. In future, we will study the limit equation of the NLDE in the nonrelativistic limit regime and investigate suitable uniformly convergent numerical methods. Another possible future work is to apply these efficient numerical methods in simulation of dynamics of electrons in graphene, which are conducted by the Dirac equation.

Bibliography

- [1] D. A. Abanin, S. V. Morozov, L. A. Ponomarenko, R. V. Gorbachev, A. S. Mayorov, M. I. Katsnelson, K. Watanabe, T. Taniguchi, K. S. Novoselov, L. S. Levito and A. K. Geim, *Giant nonlocality near the Dirac point in graphene*, Science, 332 (2011) 328–330.
- [2] M. J. Ablowitz and Y. Zhu, *Nonlinear waves in shallow honeycomb lattices*, SIAM J. Appl. Math., 72 (2012) 240–260.
- [3] A. Alvarez, *Spinorial solitary wave dynamics of a (1+3)-dimensional model*, Phys. Rev. D, 31 (1985) 2701-2703.
- [4] A. Alvarez, *Linearized Crank-Nicolson scheme for nonlinear Dirac equations*, J. Comput. Phys., 99 (1992) 348-350.
- [5] A. Alvarez, B. Carreras, *Interaction dynamics for the solitary waves of a nonlinear Dirac model*, Phys. Lett. A, 86 (1981) 327-332.
- [6] A. Alvarez, P.Y. Kuo, L. Vázquez, *The numerical study of a nonlinear one-dimensional Dirac equation*, Appl. Math. Comput., 13 (1983) 1-15.
- [7] C. D. Anderson, *The positive electron*, Phys. Rev., 43 (1933) 491–498.

-
- [8] X. Antonia, W. Bao and C. Besse, *Computational methods for the dynamics of the nonlinear Schrödinger/Gross-Pitaevskii equations*, Comput. Phys. Commun., 184 (2013) 2621-2633.
- [9] A. Arnold and H. Steinrück, *The ‘electromagnetic’ Wigner equation for an electron with spin*, ZAMP, 40 (1989) 793–815.
- [10] X. Antoine, E. Lorin, J. Sater, F. Fillion-Gourdeau and A.D. Bandrauk, *Absorbing boundary conditions for relativistic quantum mechanics equations*, J. Comput. Phys., 277 (2014) 268–304.
- [11] V. Bagrov, D. Gitman, *Exact Solutions of Relativistic Wave Equations, Mathematics and Its Applications*, Soviet series, Kluwer Academic Publishers, 1990.
- [12] W. Bao and Y. Cai, *Mathematical theory and numerical methods for Bose-Einstein condensation*, Kinet. Relat. Mod., 6 (2013) 1–135.
- [13] W. Bao, Y. Cai, *Optimal error estimates of finite difference methods for the Gross-Pitaevskii equation with angular momentum rotation*, Math. Comp. 82 (2013) 99–128.
- [14] 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., 52 (2014), 1103–1127.
- [15] W. Bao, Y. Cai and X. Jia, *Numerical methods and comparison for the Dirac equation in the nonrelativistic limit regime*, preprint.
- [16] W. Bao, Y. Cai and Y. Zhao, *A uniformly accurate multiscale time integrator pseudospectral method for the Klein-Gordon equation in the nonrelativistic limit regime*, SIAM J. Numer. Anal., 52 (2014) 2488–2511.
- [17] W. Bao and X. Dong, *Analysis and comparison of numerical methods for the Klein-Gordon equation in the nonrelativistic limit regime*, Numer. Math., 120 (2012) 189–229.

-
- [18] W. Bao, X. Dong and X. Zhao, *Uniformly correct multiscale time integrators for highly oscillatory second order differentiation equations*, J. Math. Study, 47 (2014) 111–150.
- [19] W. Bao and X. Li, *An efficient and stable numerical method for the Maxwell-Dirac system*, J. Comput. Phys., 199 (2004) 663–687.
- [20] P. Bechouche, N. Mauser and F. Poupaud, *(Semi)-nonrelativistic limits of the Dirac equation with external time-dependent electromagnetic field*, Commun. Math. Phys., 197 (1998) 405–425.
- [21] P. Bechouche, N. Mauser and S. Selberg, *On the asymptotic analysis of the Dirac-Maxwell system in the nonrelativistic limit*, J. Hyper. Differential Equations, 2 (2005) 129–182.
- [22] N. Bournaveas and G.E. Zouraris, *Split-step spectral scheme for nonlinear Dirac systems*, ESAIM: M2AN, 46 (2012) 841–874.
- [23] 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., 257 (2014) 318–332.
- [24] S.J. Chang, S.D. Ellis and B.W. Lee, *Chiral confinement: an exact solution of the massive Thirring model*, Phys. Rev. D, 11 (1975) 3572–2582.
- [25] S.A. Chin, *Symplectic integrators from composite operator factorizations*, Phys. Lett. A, 226 (1997) 344–348.
- [26] R. J. Cirincione and P. R. Chernoff, *Dirac and Klein Gordon equations: convergence of solutions in the nonrelativistic limit*, Commun. Math. Phys., 79 (1981) 33–46.
- [27] S.A. Chin and C.R. Chen, *Fourth order gradient symplectic integrator methods for solving the time-dependent Schrödinger equation*, J. Chem. Phys., 114 (2001) 7338–7341.

-
- [28] S.A. Chin and C.R. Chen, *Gradient symplectic algorithms for solving the Schrödinger equation with time-dependent potential*, J. Chem. Phys., 117 (2002) 1409-1415.
- [29] S.A. Chin and D.W. Kidwell, *Higher-order force gradient symplectic algorithms*, Phys. Rev. E, 62 (2000) 8746-8752.
- [30] F. Cooper, A. Khare, B. Mihaila and A. Saxena, *Solitary waves in the nonlinear Dirac equation with arbitrary nonlinearity*, Phys. Rev. E, 82 (2010) 036604.
- [31] M. Creutz and A. Gocksch, Phys. Rev. Lett., *Higher-order hybrid Monte Carlo algorithms*, 63 (1993) 9-12.
- [32] P. A. M. Dirac, *The quantum theory of the electron*, Proc. R. Soc. Lond. A, 117 (1928) 610–624.
- [33] P. A. M. Dirac, *A theory of electrons and protons*, Proc. R. Soc. Lond. A, 126 (1930) 360–365.
- [34] P. A. M. Dirac, *Principles of Quantum Mechanics*, Oxford University Press, London, 1958.
- [35] A. Einstein, *Relativity: the special and general theory*, 1916 (Translation 1920), New York: H. Holt and Company.
- [36] M. Esteban and E. Séré, *Existence and multiplicity of solutions for linear and nonlinear Dirac problems*, Partial Differential Equations and Their Applications, 1997.
- [37] M. Esteban and E. Séré, *An overview on linear and nonlinear Dirac equations*, Discrete Contin. Dyn. Syst., 8 (2002) 381–397.
- [38] C. L. Fefferman and M. I. Weinstein, *Honeycomb lattice potentials and Dirac points*, J. Amer. Math. Soc., 25 (2012) 1169–1220.

-
- [39] C. L. Fefferman and M. I. Weinstein, *Wave packets in honeycomb structures and two-dimensional Dirac equations*, Commun. Math. Phys., 326 (2014) 251–286.
- [40] R. Finkelstein, C. Fronsdal and P. Kaus, *Nonlinear spinor field*, Phys. Rev., 103 (1956) 1571-1579.
- [41] R. Finkelstein, R. Lelevier and M. Ruderman, *Nonlinear spinor fields*, Phys. Rev., 83 (1951) 326-332.
- [42] E. Forest and R.D. Ruth, *Fourth order symplectic integration*, Physica D, 43 (1990) 105-117.
- [43] J. De Frutos and J.M. Sanz-Serna, *Split-step spectral scheme for nonlinear Dirac systems*, J. Comput. Phys., 83 (1989) 407-423.
- [44] W. Gautschi, *Numerical integration of ordinary differential equations based on trigonometric polynomials*, Numer. Math., 3 (1961) 381–397.
- [45] P. Gérard, P. A. Markowich, N. J. Mauser and F. Poupaud, *Homogenization limits and Wigner transforms*, Comm. Pure Appl. Math., 50 (1997) 321–377.
- [46] S.K. Gray and D.E. Manolopoulos, *Symplectic integrators tailored to the time-dependent Schrödinger equation*, J. Chem. Phys., 104 (1996) 7099-7113
- [47] W. Greiner, *Relativistic quantum mechanics: wave equations*, Springer, 1987.
- [48] L.H. Haddad and L.D. Carr, *The nonlinear Dirac equation in Bose-Einstein condensates: foundation and symmetries*, Physica D, 238 (2009), 1413-1421.
- [49] L. H. Haddad, C. M. Weaver, L. D. Carr. *The nonlinear Dirac equation in Bose-Einstein condensates: I. Relativistic solitons in armchair nanoribbon optical lattices*. arXiv: 1305.6532
- [50] C.R. Hagen, *New solutions of the Thirring model*, II Nuovo Cimento, 51 (1967) 169-186.

-
- [51] E. Hairer, C. Lubich and G. Wanner, *Geometric Numerical Integration*, Springer-Verlag, 2002.
- [52] R. Hammer, W. Pötz and A. Arnold, *Single-cone real-space finite difference scheme for the time-dependent Dirac equation*, J. Comput. Phys., 265 (2014) 50–70.
- [53] 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., 256 (2014) 728–747.
- [54] M. Hochbruck and Ch. Lubich, *A Gautschi-type method for oscillatory second-order differential equations*, Numer. Math., 83 (1999) 402–426.
- [55] M. Hochbruck and A. Ostermann, *Exponential integrators*, Acta Numer., 19 (2000) 209–286.
- [56] J.L. Hong and C. Li, *Multi-symplectic Runge-Kutta methods for nonlinear Dirac equations*, J. Comput. Phys., 211 (2006) 448–472.
- [57] 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., 208 (2005) 761–789.
- [58] W. Hunziker, *On the nonrelativistic limit of the Dirac theory*, Commun. Math. Phys., 40 (1975) 215–222.
- [59] D.D. Ivanenko, *Notes to the theory of interaction particles*, Zhurn. Exp. Teoret. Fiz., 8 (1938) 260–266.
- [60] S. Jiménez, *Derivation of the discrete conservation laws for a family of finite difference schemes*, Appl. Math. Comput., 64 (1994) 13–45.
- [61] K. Johnson, *Solution of the equations for the Green's functions of a two dimensional relativistic field theory*, II Nuovo Cimento, 20 (1961) 773–790.

-
- [62] M.I. Katsnelson, K.S. Novoselov and A.K. Geim, *Chiral tunnelling and the Klein paradox in graphene*, Nat. Phys, 2 (2006) 620-625.
- [63] P. Kaus, *Comment on relation of two-dimensional to asymptotic four-dimensional solitons*, Phys. Rev. D, 14 (1976) 1722-1724.
- [64] V.E. Korepin, *Dirac calculation of the S matrix in the massive Thirring model*, Theoretical and Mathematical Physics, 41 (1979) 953-967.
- [65] S.Y. Lee, T.K. Kuo and A. Gavrielides, *Exact localized solutions of two-dimensional field theories of massive fermions with Fermi interactions*, Phys. Rev. D, 12 (1975) 2249-2253.
- [66] C. Lubich, *On splitting methods for Schrödinger-Poisson and cubic nonlinear Schrödinger equations*, Math. Comp., 77 (2008) 2141–2153.
- [67] N. Masmoudi and N. J. Mauser, *The selfconsistent Pauli equation*, Monatsh. Math., 132 (2001) 19–24.
- [68] N. Masmoudi and K. Nakanishi, *From nonlinear Klein-Gordon equation to a system of coupled nonlinear Schrödinger equations*, Math. Ann., 324 (2002), pp. 359-389.
- [69] P. Mathieu, *Compact solitons, bags, and radial excitations*, Phys. Rev. D, 32 (1985) 3288-3293.
- [70] P. Mathieu, *Soliton solutions for Dirac equations with homogeneous non-linearity in (1+1) dimensions*, J. Phys. A: Math. Gen., 18 (1985) L1061-L1066.
- [71] N. J. Mauser, *Rigorous derivation of the Pauli equation with time-dependent electromagnetic field*, VLSI Design, 9 (1999) 415–426.
- [72] N. Milosevic, V.P. Krainov and T. Brabec, *Semiclassical Dirac theory of tunnel ionization*, Phys. Rev. Lett., 89 (2002) 193001.

- [73] B. Najman, *The nonrelativistic limit of the nonlinear Dirac equation*, Ann. Inst. Henri Poincaré, 9 (1992) 3–12.
- [74] A. H. C. Neto, F. Guinea, N. M. R. Peres, K. S. Novoselov and A. K. Geim, *The electronic properties of the graphene*, Rev. Mod. Phys., 81 (2009) 109–162.
- [75] K. Novoselov, A. Geim, S. Morozov, D. Jiang, M. I. Katsnelson, I. V. Grigorieva, S. V. Dubonos and A. A. Firsov, *Two-dimensional gas of massless Dirac fermions in graphene*, Nature, 438 (2005) 197–200.
- [76] K. Novoselov, A. Geim, S. Morozov, D. Jiang, Y. Zhang, S. Dubonos, I. Grigoriev and A. Firsov, *Electric field effect in atomically thin carbon films*, Science, 306 (2004) 666–669.
- [77] K. Novoselov, Z. Jiang, Y. Zhang, S. Morozov, H. Stormer, U. Zeitler, J. Maan, G. Boebinger, P. Kim and A. Geim, *Room-temperature quantum Hall effect in graphene*, Science, 315 (2007) 1379.
- [78] J. W. Nraun, Q. Su and R. Grobe, *Numerical approach to solve the time-dependent Dirac equation*, Phys. Rev. A, 59 (1999) 604–612.
- [79] J. Rafelski, *Soliton solutions of a selfinteracting Dirac field in three space dimensions*, Phys. Lett. B, 66 (1977) 262-266.
- [80] A.F. Rañada, *Classical nonlinear Dirac field models of extended particles*, in: A. O. Barut (Ed.), Quantum Theory, Groups, Fields and Particles, Springer, New York, 1983, 271-291.
- [81] B. Saha, *Nonlinear spinor fields and its role in cosmology*, Int. J. Theor. Phys., 51 (2012) 1812-1837.
- [82] Y.I. Salamin, S. Hu, K.Z. Hatsagortsyan and C.H. Keitel, *Relativistic high-power laser-matter interactions*, Phys. Rep., 427 (2-3) (2006) 41-155.

-
- [83] F. Schedin, A. Geim, S. Morozov, E. Hill, P. Blake, M. Katsnelson and K. Novoselov, *Detection of individual gas molecules absorbed on graphene*, *Nature Materials*, 6 (2007) 652–655.
- [84] A. Y. Schoene, *On the nonrelativistic limits of the Klein-Gordon and Dirac equations*, *J. Math. Anal. Appl.*, 71 (1979), 36–74.
- [85] E. Schrödinger, *An undulatory theory of the mechanics of atoms and molecules*, *Phys. Rev.*, 28 (1926), pp 1049-1070.
- [86] S.H. Shao and H.Z. Tang, *Interaction for the solitary waves of a nonlinear Dirac model*, *Phys. Lett. A*, 345 (2005) 119-128.
- [87] 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*, 6 (2006) 623-640.
- [88] S.H. Shao and H.Z. Tang, *Interaction of solitary waves with a phase shift in a nonlinear Dirac model*, *Commun. Comput. Phys.*, 3 (2008) 950-967.
- [89] J. Shen and T. Tang, *Spectral and High-Order Methods with Applications*, Science, Beijing, 2006.
- [90] G. D. Smith, *Numerical Solution of Partial Differential Equations: Finite Difference Methods*, Clarendon Press, 1985.
- [91] M. Soler, *Classical, stable, nonlinear spinor field with positive rest energy*, *Phys. Rev. D*, 1 (1970) 2766-2769.
- [92] G. Strang, *On the construction and comparison of difference schemes*, *SIAM J. Numer. Anal.*, 5 (1968) 505–517.
- [93] J. Stubbe, *Exact localized solutions of a family of two-dimensional nonlinear spinor fields*, *J. Math. Phys.*, 27 (1986) 2561-2567.

-
- [94] M. Suzuki, *Fractal decomposition of exponential operators with applications to many-body theories and Monte Carlo simulations*, Phys. Lett. A, 146 (1990) 319-323.
- [95] M. Suzuki, *General theory of fractal path integrals with applications to manybody theories and statistical physics*, J. Math. Phys., 32 (1991) 400-407.
- [96] M. Suzuki, *General decomposition theory of ordered exponentials*, Proc. Jpn. Acad., Ser. B: Phys. Biol. Sci., 69 (1993) 161-166.
- [97] K. Takahashi, *Soliton solutions of nonlinear Dirac equations*, J. Math. Phys., 20 (1979) 1232-1238.
- [98] K. Takahashi and K. Ikeda, *Applicability of symplectic integrator to classically unstable quantum dynamics*, J. Chem. Phys., 96 (1993) 8680-8694.
- [99] K. Takahashi and K. Ikeda, *Application of symplectic integrator to stationary reactive-scattering problems: Inhomogeneous Schrödinger equation approach*, J. Chem. Phys., 106 (1997) 4463-4480.
- [100] W.E. Thirring, *A soluble relativistic field theory*, Ann. Phys., 3 (1958) 91-112.
- [101] Z.Q. Wang and B.Y. Guo, *Modified Legendre rational spectral method for the whole line*, J. Comput. Math., 22 (2004) 457-472.
- [102] H. Wang and H.Z. Tang, *An efficient adaptive mesh redistribution method for a nonlinear Dirac equation*, J. Comput. Phys., 222 (2007) 176-193.
- [103] J. Werle, *Non-linear spinor equations with localized solutions*, Lett. Math. Phys., 2 (1977) 109-114.
- [104] H. Weyl, *A remark on the coupling of gravitation and electron*, Phys. Rev., 77 (1950) 699-701.
- [105] G. B. White, *Splitting of the Dirac operator in the nonrelativistic limit*, Ann. Inst. Henri Poincaré, 53 (1990) 109-121.

-
- [106] H. Wu, Z. Huang, S. Jin and D. Yin, *Gaussian beam methods for the Dirac equation in the semi-classical regime*, Commun. Math. Sci., 10 (2012) 1301–1315.
- [107] J. Xu, S. H. Shao and H. Z. Tang, *Numerical methods for nonlinear Dirac equation*, J. Comput. Phys., 245 (2013) 131–149.
- [108] H. Yoshida, *Construction of higher order symplectic integrators*, Phys. Lett. A, 150 (1990) 262-268.

