

# A semi-Lagrangian time splitting method for the Schrödinger equation with vector potentials\*

Shi Jin<sup>†</sup>, Zhennan Zhou<sup>‡</sup>

September 10, 2013

## Abstract

In this paper, we present a time splitting scheme for the Schrödinger equation in the presence of electromagnetic field in the semi-classical regime, where the wave function propagates  $O(\varepsilon)$  oscillations in space and time. With the operator splitting technique, the time evolution of the Schrödinger equation is divided into three parts: the kinetic step, the convection step and the potential step. The kinetic and the potential steps can be handled by the classical time-splitting spectral method. For the convection step, we propose a semi-Lagrangian method in order to allow large time steps. We prove the unconditional stability conditions with spatially variant external vector potentials, and the error estimate in the  $l^2$  approximation of the wave function. By comparing with the semi-classical limit, the classical Liouville equation in the Wigner framework, we show that this method is able to capture the correct physical observables with time step  $\Delta t \gg \varepsilon$ . We implement this method numerically for both one dimensional and two dimensional cases to verify that  $\varepsilon$ -independent time steps can indeed be taken in computing physical observables.

## 1 Introduction

Many problems in solid state physics and quantum chemistry require the solution to the Schrödinger equation in the presence of electromagnetic field with a small (scaled) Planck constant  $\varepsilon$ ,

$$i\varepsilon\partial_t u^\varepsilon = \frac{1}{2}(-i\varepsilon\nabla_x - \mathbf{A})^2 u^\varepsilon + V u^\varepsilon, \quad t \in \mathbb{R}^+, \quad x \in \mathbb{R}^d; \quad (1.1)$$

$$u^\varepsilon(x, 0) = u_0(x), \quad x \in \mathbb{R}^d \quad (1.2)$$

---

\*This work was partially supported by NSF grants DMS-1114546 and DMS-1107291: NSF Research Network in Mathematical Sciences KI-Net: Kinetic description of emerging challenges in multiscale problems of natural sciences.

<sup>†</sup>Department of Mathematics, Institute of Nature Science, and Ministry of Education Key Laboratory in Scientific and Engineering Computing, Shanghai Jiao Tong University, Shanghai 200240, P. R. China, and Department of Mathematics, University of Wisconsin, Madison, WI 53706, USA(jin@math. wisc.edu)

<sup>‡</sup>Department of Mathematics, University of Wisconsin, Madison, WI 53706, USA (zhou@math. wisc.edu).

where  $u^\varepsilon(x, t)$  is the complex-valued wave function,  $V(x)$  is the scalar potential and  $\mathbf{A}(x)$  is the vector potential.

Mathematically, the electromagnetic field, or respectively, electric field  $\mathbf{E}(x) \in \mathbb{R}^d$  and magnetic field  $\mathbf{B}(x) \in \mathbb{R}^d$  are described by the scalar potential  $V(x) \in \mathbb{R}$  and the vector potential  $\mathbf{A}(x) \in \mathbb{R}^d$  as

$$\mathbf{E} = -\nabla V(x), \quad \mathbf{B} = \nabla \times \mathbf{A}. \quad (1.3)$$

In the dynamic picture, one defines canonical momentum  $\hat{\mathbf{P}} = -i\varepsilon\nabla_x$  and the kinetic momentum is  $\hat{\kappa} = \hat{\mathbf{P}} - \mathbf{A}$  (see [9]). The Schrödinger equation (1.1) can be derived from the one in the absence of the vector potential by local gauge transformation (see [26]).

In fact, one can simplify the potential description by imposing one more condition, namely, specifying the gauge. Due to the fact that the potential fields are not what are observed, while the electric and magnetic fields are, there is freedom to impose conditions on the potentials so long as whatever condition is chosen to impose does not affect the resultant electric and magnetic fields. This freedom is called the gauge freedom. For any choice of a scalar function of position  $\lambda(x) \in \mathbb{R}$ , the potentials can be changed as follows:

$$\mathbf{A}' = \mathbf{A} + \nabla_x \lambda, \quad V' = V. \quad (1.4)$$

One can easily show that electric field  $\mathbf{E}(x) \in \mathbb{R}^d$  and magnetic field  $\mathbf{B}(x) \in \mathbb{R}^d$  do not change at all under this transformation. One natural choice is, choosing  $\lambda$ , so that  $\nabla_x \cdot \mathbf{A}' = 0$ . This is the so-called Coulomb gauge. In this gauge, the vector potential and the canonical momentum operator commute,  $[\mathbf{A}, -i\varepsilon\nabla_x] = 0$ , so that the modified “kinetic” part of the Schrödinger equation (1.1) can be simplified to:

$$\begin{aligned} \frac{1}{2}(-i\varepsilon\nabla_x - \mathbf{A})^2 u^\varepsilon &= -\frac{\varepsilon^2}{2}\Delta_x u^\varepsilon + \frac{i\varepsilon}{2}(\mathbf{A} \cdot \nabla_x + \nabla_x \cdot \mathbf{A})u^\varepsilon + \frac{1}{2}|\mathbf{A}|^2 u^\varepsilon \\ &= -\frac{\varepsilon^2}{2}\Delta_x u^\varepsilon + i\varepsilon\mathbf{A} \cdot \nabla_x u^\varepsilon + \frac{1}{2}|\mathbf{A}|^2 u^\varepsilon. \end{aligned}$$

Previously, many numerical methods have been designed for the semi-classical Schrödinger equation with only scalar potentials. As far as we know, little research has been done for the semi-classical Schrödinger equation with vector potentials in the aspect of numerical simulations. However, dynamics for particles exposed to external electromagnetic field result in many far-reaching consequences in quantum mechanics, such as the Landau level, the Zeeman effect and superconductivity. In the aspect of analysis, the Schrödinger operator with the vector potential has different features in spectral and scattering properties (see [2]). Numerically, it gives new challenges as well, especially in the semi-classical regime.

In the Schrödinger equation, the wave function acts as an auxiliary quantity used to compute

primary physical quantities such as the position density,

$$n(t, x) = |u^\varepsilon(t, x)|^2, \quad (1.5)$$

the current density,

$$\mathbf{I}(t, x) = \varepsilon \operatorname{Im}(\overline{u^\varepsilon}(t, x) \nabla_x u^\varepsilon(t, x)) = \frac{\varepsilon}{2i} (\overline{u^\varepsilon} \nabla_x u^\varepsilon - u^\varepsilon \nabla_x \overline{u^\varepsilon}), \quad (1.6)$$

where  $\bar{f}$  denotes the complex conjugate of  $f$ . As a matter of fact, in the presence of the vector potential, one needs to introduce the modified current density as

$$\mathbf{J}(t, x) = \frac{1}{2} (\overline{u^\varepsilon} (-i\varepsilon \nabla_x - A) u^\varepsilon - u^\varepsilon (-i\varepsilon \nabla_x - A) \overline{u^\varepsilon}), \quad (1.7)$$

so that mass conservation equation is satisfied,

$$\frac{\partial}{\partial t} n + \nabla_x \cdot \mathbf{J} = 0. \quad (1.8)$$

We remark that  $n$  and  $J$  are gauge invariant quantities. Numerically, computing  $\mathbf{I}(t, x)$  and  $\mathbf{J}(t, x)$  face the same challenge.

It is well known that, the semi-classical Schrödinger equation propagates oscillation of wavelength of order  $O(\varepsilon)$  in space and time, so that the wave function  $u^\varepsilon$  does not converge in the strong sense as  $\varepsilon \rightarrow 0$ . In addition, since the macroscopic physical quantities are non-linear transforms of  $u^\varepsilon$ , the classical limit for those physical observables are not guaranteed by the weak convergence. Mathematically, some micro-local analysis methodologies were introduced to explore the so-called semi-classical limit of the Schrödinger equation. The celebrated Wigner transform (see [11, 10, 23]) has been shown to be a very powerful tool to reveal the macroscopic properties of the Schrödinger equation in the semi-classical regime. In this framework, the semi-classical limit of the Schrödinger equation can be derived, which is the classical Liouville equation. This provides an insightful viewpoint to understand the transition from quantum mechanics to classical mechanics.

Numerically, the oscillatory nature of the wave function for the semi-classical Schrödinger equation, in general, gives rise to significant computational burden. As a matter of fact, even for unconditionally stable methods, the numerical results may lead to completely wrong physical observables if the mesh grids fail to completely resolve the  $O(\varepsilon)$  oscillations in space and time. In [20], by utilizing the Wigner transform to study finite difference approximation of the Schrödinger equation, Markowich, Pietra and Pohl have shown for prevailing finite difference method, to obtain correct physical observables, one has to enforce the following meshing strategy

$$\Delta x = o(\varepsilon), \quad \Delta t = o(\varepsilon). \quad (1.9)$$

In the meanwhile, in order to guarantee accurate  $L^2$  approximation of the wave function, even more restrictive conditions have to be satisfied.

In [5], Bao, Jin and Markowich have shown that the time splitting spectral method gives much less restrictive conditions in approximating not only the wave function but also physical observables. By comparing with the semi-classical limit using the Wigner Transform, and presenting extensive numerical experiments, they have shown that the following meshing strategy is sufficient in computing correct physical observables,

$$\Delta x = O(\varepsilon), \quad \Delta t = o(1). \quad (1.10)$$

In other words, one is allowed to take  $\varepsilon$ -independent time steps in computing physical observables. The readers can refer to the recent review [15] on the computation of semi-classical Schrödinger equations by Jin, Markowich and Sparber.

In the presence of the vector potential, in general, the stability constraint in solving the convection part

$$\partial_t u^\varepsilon = A \cdot \nabla_x u^\varepsilon \quad (1.11)$$

requires  $\Delta t = O(\varepsilon)$  by an explicit scheme since one needs to take  $\Delta x = O(\varepsilon)$  in order to resolve spatial oscillations. The primary goal of this paper is to develop a numerical method with meshing strategy (1.10), so that large time steps satisfying  $\Delta t \gg \varepsilon$  are allowed to take in computing the physical observables.

We propose a semi-Lagrangian method to handle the convection step with the goal to allow time steps which are as large as the case without the vector potential term. In this method, one follows the characteristics of the convection equation (1.11) backwards in time from the original grid points to the previous time step, and then use polynomial interpolation to compute the corresponding value of the numerical solution.

The semi-Lagrangian methodology has been extensively studied in atmospheric models. See [28] for a review by Staniforth and Côté. This technique has been extended to general transport equations, for example, by Lin and Rood in [18], and in particular, a lot of research has been done for its applications in Vlasov equation and other kinetic models (see [27, 24]). Previously, the stability study for the semi-Lagrangian method was carried out only for constant coefficient problems (except the case of using the spectral interpolation, see [29]). In this paper, we give rigorous  $l^2$  stability analysis of this method for variable coefficient problems with the polynomial interpolation, and prove the unconditional stability if suitable interpolation points are used.

By analyzing the correspondence between solving the Schrödinger equation and the semi-classical limit, namely the Liouville equation in the Wigner framework, we show that the semi-Lagrangian time splitting method is able to compute correct physical observables even with time step  $\Delta t \gg \varepsilon$ , which is further verified numerically.

The rest of this paper is organized as follows. In Section 2, we present the numerical methods to solve the one dimensional Schrödinger equation with vector potentials based on the time splitting technique and carry out the  $l^2$  stability analysis for arbitrary Courant numbers. In Section 3, we prove the error estimate and corresponding meshing strategy for the semi-Lagrangian time

splitting method in the  $l^2$  approximation of the wave function. In Section 4, by comparing with the semi-classical limit, namely the Liouville equation in the Wigner framework, we prove that the meshing strategy can be much relaxed so that  $\varepsilon$ -independent time step is allowed if one only aims to obtain correct physical observables. We discuss how to extend the method to multidimensional cases in Section 5. Extensive numerical examples are shown in Section 6 to verify the proposed meshing strategy. We give some final remarks and future directions in the last section.

## 2 The description of numerical methods

### 2.1 The time splitting and the spectral approximation.

In this section, we present the numerical method to solve the Schrödinger equation (1.1), (1.2) in one dimension with periodic boundary condition. The extension to multidimensional cases is straightforward, and will be discussed in Section 5.

We assume, on computation domain  $[a, b]$ , a uniform spatial grid  $x_j = a + j\Delta x$ ,  $j = 0, \dots, N-1$ , where  $N = 2^{n_0}$ ,  $n_0$  is an positive integer and  $\Delta x = \frac{b-a}{N}$ . We also assume uniform time steps  $t_k = k\Delta t$ ,  $k = 0, \dots, K$ . The construction of numerical methods is based on the following (first order) operator splitting technique.

With the Coulomb gauge, the Schrödinger equation (1.1) can be formulated as

$$i\varepsilon\partial_t u^\varepsilon = -\frac{\varepsilon^2}{2}\Delta u^\varepsilon + i\varepsilon\mathbf{A} \cdot \nabla u^\varepsilon + \frac{1}{2}|\mathbf{A}|^2 u^\varepsilon + V u^\varepsilon, \quad a < x < b, \quad t \in \mathbb{R}^+; \quad (2.1)$$

$$u^\varepsilon(0, x) = u_0^\varepsilon(x), \quad u^\varepsilon(t, a) = u^\varepsilon(t, b), \quad u_x^\varepsilon(t, a) = u_x^\varepsilon(t, b).$$

By the operator splitting technique, for every time step  $t \in [t_n, t_{n+1}]$ , one solves the kinetic step

$$i\varepsilon\partial_t u^\varepsilon = -\frac{\varepsilon^2}{2}\Delta u^\varepsilon, \quad t \in [t_n, t_{n+1}]; \quad (2.2)$$

followed by the potential step

$$i\varepsilon\partial_t u^\varepsilon = \frac{1}{2}|\mathbf{A}|^2 u^\varepsilon + V u^\varepsilon, \quad t \in [t_n, t_{n+1}], \quad (2.3)$$

and followed by the convection step

$$\partial_t u^\varepsilon = \mathbf{A} \cdot \nabla u^\varepsilon, \quad t \in [t_n, t_{n+1}]. \quad (2.4)$$

For clarity, we rewrite the equation (2.1) as

$$\partial_t u^\varepsilon = (\mathcal{A} + \mathcal{B} + \mathcal{C}) u^\varepsilon \quad (2.5)$$

where

$$\mathcal{A} = \frac{i\varepsilon}{2}\Delta, \quad \mathcal{B} = -\frac{i}{\varepsilon} \left( \frac{1}{2}|\mathbf{A}|^2 + V \right), \quad \mathcal{C} = \mathbf{A} \cdot \nabla.$$

Let  $u^\varepsilon(t_n)$  be the exact solution at  $t = t_n$ , so  $u^\varepsilon(t_{n+1}) = e^{(\mathcal{A}+\mathcal{B}+\mathcal{C})\Delta t}u^\varepsilon(t_n)$ . Let  $U_j^n$  be the numerical approximation of  $u^\varepsilon(x_j, t_n)$  and  $u^{\varepsilon,n}$  be the numerical approximation of  $u^\varepsilon(t_n)$ , which means  $u^{\varepsilon,n}$  has  $U_j^n$  as components. Define the solution obtained by the (first order) operator splitting (without spatial discretization) as

$$w^{n+1} = e^{\mathcal{C}\Delta t}e^{\mathcal{B}\Delta t}e^{\mathcal{A}\Delta t}u^\varepsilon(t_n). \quad (2.6)$$

Note that  $w^{n+1}$  differs from  $u^\varepsilon(t_{n+1})$  due to the operator splitting error.

After operator splitting, the kinetic step can be solved analytically in time in the Fourier space, and the potential step can be solved exactly by direct integration in time:

$$U_j^* = \frac{1}{N} \sum_{l=-N/2}^{N/2-1} e^{-i\varepsilon\Delta t\mu_l^2/2} \hat{U}_l^n e^{i\mu_l(x_j-a)}; \quad (2.7)$$

$$U_j^{**} = e^{-i(\frac{1}{2}|\mathbf{A}|^2(x_j)+V(x_j))\Delta t/\varepsilon} U_j^*; \quad (2.8)$$

where  $\hat{U}_l^n$  are Fourier Coefficients of  $U_j^n$ , defined by

$$\hat{U}_l^n = \sum_{j=0}^{N-1} U_j^n e^{-i\mu_l(x_j-a)}, \quad \mu_l = \frac{2\pi l}{b-a}, \quad l = -\frac{M}{2}, \dots, \frac{M}{2} - 1.$$

But for the convection step, there is no obvious way to solve it analytically based on discrete data for a variable  $\mathbf{A}(x)$ . We propose in the next sub-section a semi-Lagrangian method to solve the convection equation (2.9).

We need to give two remarks here:

*Remark 1.* Even if one doesn't specify the Coulomb gauge, the commutator  $[\nabla_x, \mathbf{A}] = \nabla_x \cdot \mathbf{A}$  appears, and one only needs to add this contribution to operator  $\mathcal{C}$ , namely one modifies  $\mathcal{C} = -i(\frac{1}{2}|\mathbf{A}|^2 + V)/\varepsilon + \frac{1}{2}\nabla_x \cdot \mathbf{A}$ . But, since  $\nabla_x \cdot \mathbf{A}$  is only a slowly varying scalar function, this modification will not introduce any new challenges numerically.

*Remark 2.* The first order operator splitting implies first order convergence in time. One can make use of Strang's splitting to obtain second order time discretization method. If one wants to apply the second order Strang's splitting to three operators, one can firstly group  $\mathcal{A} + \mathcal{B}$  together as a single operator, and apply Strang's splitting to  $\mathcal{A} + \mathcal{B}$  and  $\mathcal{C}$ , while in the steps corresponding to  $\mathcal{A} + \mathcal{B}$ , one also uses Strang's splitting.

## 2.2 A semi-Lagrangian method for the convection step

In this part, we present a semi-Lagrangian method (abbreviated by SL) to solve the following scalar convection equation with periodic boundary conditions

$$\partial_t u^\varepsilon - \mathbf{A} \cdot \nabla_x u^\varepsilon = 0, \quad t \in [t_n, t_{n+1}]. \quad (2.9)$$

Such an approach has been used to solve atmospheric models, Vlasov equation and other transport equations with improved stability condition, see [28, 18, 27, 24]. This method consists of two parts: backward characteristic tracing and interpolation. We compute the data  $U_j^{n+1}$  by firstly tracing backwards along the characteristic line :

$$\frac{dx(t)}{dt} = -\mathbf{A}(x(t)), \quad x(t_{n+1}) = x_j, \quad (2.10)$$

for time interval  $[t_n, t_{n+1}]$ . Denote  $x(t_n) = x_j^0$ , obtained by numerically solving the ODE (2.10) backwards in time as shown in the graph below:

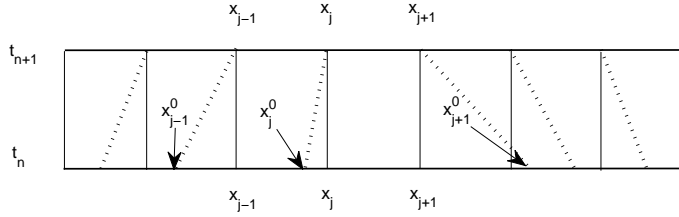


Figure 2.1: Backward Tracing:  $x_j$  are grid points;  $x_j^0$  are the shifted grid points, which are the solutions to problem (2.10) backwards in time at  $t = t^n$ ; dot line "... " indicates characteristics.

We call the point set  $\{x_j^0\}$  the shifted point set. By the method of characteristics,  $U_j^{n+1} = U^n(x_j^0)$ . But,  $U^n(x_j^0)$  in general are not known, since  $x_j^0$  are not necessarily grid points. Therefore, interpolation is needed to approximate  $U_j^{n+1} = U^n(x_j^0)$  based on  $u^{\varepsilon, n}$ . We compare the following two choices: the spectral interpolation and the  $M^{th}$  order polynomial interpolation.

For the spectral approximation, the interpolant  $\Pi_N U^n(x) = \sum_{k=-N/2}^{N/2-1} c_k e^{ikx}$  is a global approximation to  $U^n(x)$  based on  $u^{\varepsilon, n}$ . One needs  $O(N \log N)$  operations to get the Fourier coefficients  $c_k$  via the FFT method. But, one needs  $O(N)$  operations to evaluate the interpolant at each point  $x_j^0$ , since the shifted points  $x_j^0$  are not necessarily the grid points, which means the inverse FFT does not apply. Hence, the total cost is  $O(N^2)$  in each time step. This will make the whole scheme very costly.

But, for the  $M^{th}$  order Lagrange polynomial interpolation, one needs to establish a polynomial interpolant for each shifted point  $x_j^0$  with the discrete data on the closest  $M$  grid points

$x_{j_1}, \dots, x_{j_M}$ . For each shifted point  $x_j^0$ , one uses  $M$  grid points near  $x_j^0$  to form a Lagrange polynomial interpolant to approximate  $U(x_j^0)$  with error of order  $O(\Delta x^M)$ . But, certain stability constraints need to be satisfied for different interpolation methods. We discuss this issue in the next sections. In practice, the fourth order interpolation, namely the cubic polynomial interpolation is widely used (see [28]). The total cost of the semi-Lagrangian method with the polynomial interpolation is  $O(N)$  for each time step. Therefore, we choose to take the polynomial interpolation rather than a spectral interpolation.

We remark that, in [29], a local Chebyshev polynomial approach was proposed to improve the efficiency of the semi-Lagrangian method with the spectral interpolation, which is essentially using a local polynomial interpolation with interpolation points at Chebyshev extrema to approximate the Fourier basis function  $e^{ix\xi}$ . The cost could be reduced to  $O(N(\log N)^2)$ , and the accuracy is restricted by the regularity of the solutions and the order of local Chebyshev polynomials used. Thus, the semi-Lagrangian method proposed in [29] could also apply to this problem.

Recall that, the kinetic step requires  $O(N \log N)$  operations with the spectral approximation. Hence, the overall cost of the semi-Lagrangian time splitting method (abbreviated by SL-TS) for the whole Schrödinger equation is  $O(N \log N)$  for each time step. We prove in the next sections that the SL-TS method is unconditionally stable if suitable interpolation points are used. Therefore, the time step can be taken much larger than  $\Delta x$ , and the constraint  $\Delta t = O(\varepsilon)$  is removed, if only physical observables are needed.

In summary, the semi-Lagrangian method for the convection step is implemented by the following procedures:

1. For each grid point  $x_j$ ,  $j = 1, \dots, N$ , solve equation (2.10) backward for  $\Delta t$  time to obtain the shifted grid points  $x_j^0$ ,  $j = 1, \dots, N$ .
2. For each shifted grid point  $x_j^0$ , find the closest  $M$  grids points  $x_{j_1}, \dots, x_{j_M}$  subject to certain stability requirements and obtain the approximate value  $U_j^{n+1} = U^n(x_j^0)$  by polynomial interpolation.

Here are some remarks for the semi-Lagrangian method:

*Remark 3.* When the vector potential  $A$  is time independent, the backward characteristic tracing step is also independent of time. In other words, one just needs to solve (2.10) for the set of shifted point  $\{x_j^0\}$  once with sufficiently small time step  $\widetilde{\Delta t}$ , and can make use of them for all future time steps. This step can be done in a preprocessed step with great precision. When the vector potential is time dependent, the backward characteristic tracing step needs to be done for every time step with  $O(N)$  operations.

Consider the one dimensional convection equation with periodic boundary conditions:

$$\partial_t u - A(x) \cdot \partial_x u = 0, \quad a < x < b. \quad (2.11)$$

Assume  $A(x) \in C^2([a, b])$ ,  $A(x)$  and its first two derivatives are bounded. With the  $M^{\text{th}}$  order Lagrange polynomial interpolation, the numerical scheme for the convection equation (2.11) can



be written as

$$U_j^{n+1} = \sum_{m=1}^M l_{m+p_j}(x_j^0) U_{m+p_j}^n, \quad (2.12)$$

where  $p_j$  is determined by  $\Delta t$ ,  $\Delta x$  and the velocity  $A(x)$ , and  $l_{m+p_j}$  are the Lagrange basis functions.

If  $A(x)$  is spatially constant, in [6], it has been shown that, when linear interpolation ( $M = 2$ ) is applied, the scheme (2.12) is unconditionally stable when  $x_j^0$  is between  $x_{p_j+1}$  and  $x_{p_j+2}$ , and when quadratic interpolation ( $M = 3$ ) is applied, the scheme (2.12) is unconditionally stable when  $x_j^0$  is between  $x_{p_j+1}$  and  $x_{p_j+3}$ . Similar analysis has been presented in [8]. However, little research has been done in the rigorous stability analysis for the semi-Lagrangian method with spatially variant  $A(x)$ .

We study the stability requirements of the semi-Lagrangian method in the following three subsections. In Section 2.3, we deal with the case when  $\Delta t = O(\Delta x)$ , and the semi-Lagrangian difference operator is treated as a one-parameter family of operators depending on  $\Delta x$ . This was done in [17], where  $\Delta t$  is treated as  $\Delta x$ , but for our purpose, we need to keep track of both  $\Delta x$  and  $\Delta t$  in the analysis. In Section 2.4, we study the case when  $C_c = \|A\|_{L^\infty} \Delta t / \Delta x > 1$ , which together with results in Section 2.3 covers all Courant number cases, in particular when  $\Delta t \gg \Delta x$ , so an unconditional stability is established. In Section 2.5, we derive the specific stability requirement for the semi-Lagrangian method with the fourth order polynomial interpolation.

For numerical comparison, we also introduce a pseudo-spectral method for the convection equation (2.9). We discretize the spatial derivative by the spectral approximation,

$$\partial_t U_j^\varepsilon - \mathbf{A}|_{x=x_j} \cdot D_x U^\varepsilon|_{x=x_j} = 0, \quad (2.13)$$

where

$$D_x U^\varepsilon|_{x=x_j} = \frac{1}{N} \sum_{l=-N/2}^{N/2-1} i\mu_l \hat{U}_l^n e^{i\mu_l(x_j-a)},$$

and use some explicit ODE solver in time discretization. With the spectral approximation, the eigenvalues of the spatial discretization operator are purely imaginary (see [32]). This indicates, the absolute stability region of the time-discretization method used has to cover part of the imaginary axis near origin. Namely, one needs the so-called I-stability studied in [4]. For example, one can use the fourth order Runge-Kutta method or the leap frog method. We name this method time-explicit spectral approximation method (abbreviated by TESP).

Note that the FFT method can be used to compute spatial derivative with the spectral approximation, so the total cost in this step is  $O(N \log(N))$ , which is comparable to the cost of the kinetic step. This method has spectral convergence in space. However, because some explicit time-discretization method is applied, one needs to enforce  $\Delta t / \Delta x = O(1)$  to guarantee stability, which means the overall meshing strategy is  $\Delta x = O(\varepsilon)$ ,  $\Delta t = O(\varepsilon)$ . We remark that, the drawback of the TESP method is, it requires  $\varepsilon$  dependent time step even if only physical observables

are desired.

### 2.3 Stability of the semi-Lagrangian method when $\Delta t = O(\Delta x)$

Define the difference operator  $P_\delta$  depending on a positive parameter  $\delta$  in the following way

$$P_\delta = \sum_{\alpha} p_\alpha(x) T_\alpha, \quad (2.14)$$

where  $\alpha$  is an integer, and  $T_\alpha$  is the shift operator,  $(T_\alpha u)(x) = u(x + \alpha\delta)$ . The natural domain for the difference operator is the space of functions defined on a lattice. It is easy to show that boundedness and positivity of the difference operator over lattice functions of the  $l_2$  space are equivalent to these properties of the difference operator over the  $L^2$  space (see [17]). Next, define the symbol  $p(x, \xi)$  of the one-parameter family  $P_\delta$  as

$$p(x, \xi) = \sum_{\alpha} p_\alpha(x) e^{i\alpha\xi}, \quad (2.15)$$

which is  $2\pi$ -periodic in  $\xi$ . For functions  $f(x)$ , denote by  $|f|_k$  the maximum of the  $L^2$  norm of  $f$  and the  $L^2$  norms of its first  $k$  partial derivatives, namely,

$$|f|_k = \max_{l \in \{0, \dots, k\}} \|\partial_x^l f\|_{L^2}.$$

Then, define the  $(k, l)$  norm of the symbol  $p$  as

$$|p|_{k,l} = \sum_{\alpha} |p_\alpha|_k (1 + |\alpha|)^l. \quad (2.16)$$

Denote by  $C_{k,l}$  the class of symbols with finite  $(k, l)$  norm.

We quote the following lemma from [17], which plays a crucial role in proving stability.

**Lemma 1.** *Let  $P_\delta$  be a one-parameter family of difference operators, whose symbol  $p(x, \xi)$  is in  $C_{2,0} \cap C_{0,2}$ . Suppose  $p(x, \xi)$  is bounded by 1:*

$$|p(x, \xi)| \leq 1, \quad \forall x, \xi \in \mathbb{R}. \quad (2.17)$$

*Then for all  $\delta$ , there exists a constant  $K$ , such that the operator  $P_\delta$  is bounded in the following way:*

$$\|P_\delta\|_{L^2} \leq 1 + K\delta. \quad (2.18)$$

To prove this Lemma, we need to quote two theorems from [17]. The first one is a standard result for pseudo-differential operators:

**Theorem 1.** *Let  $A_\delta, B_\delta$  be one-parameter families of difference operators of the form (2.14) with symbols  $a(x, \xi)$  and  $b(x, \xi)$ , respectively. Denote the product  $ab$  by  $c$ , and the one-parameter family*

of operators with symbol  $c$  by  $C_\delta$ . Then

$$\|A_\delta B_\delta - C_\delta\|_{L^2} \leq \delta |a|_{0,1} |b|_{1,0}.$$

The following theorem is the main result of paper [17]:

**Theorem 2.** *Let  $Q_\delta$  be a one-parameter family of difference operators whose symbol  $q(x, \xi)$  is in  $C_{2,0} \cap C_{0,2}$ . Suppose  $q$  is Hermitian and non-negative definite for every  $x$  and  $\xi$ . Then there exists a constant  $K$  related to  $|q|_{2,0}$  and  $|q|_{0,2}$ , such that the operator  $Q_\delta$  satisfies the inequality*

$$\mathbf{Re}(u, Q_\delta u) \geq -K\delta (u, u),$$

for all  $\delta$ , and arbitrary  $u \in L^2$ .

Now we sketch the proof of Lemma 1 from [17], and later modify the proof to show that the stability constraint in time steps  $\Delta t$  can be relaxed.

*Proof.* For all  $u \in L^2$ ,

$$\|u\|_{L^2}^2 - \|P_\delta u\|_{L^2}^2 = (u, u) - (P_\delta u, P_\delta u) = (u, (I - P_\delta^* P_\delta) u), \quad (2.19)$$

where  $P_\delta^*$  is the adjoint of  $P_\delta$ .

Define the symbol  $q = 1 - p^* p$ , and denote the difference operator with symbol  $q$  by  $Q_\delta$ . By Theorem 1,  $Q_\delta$  differs from  $(I - P_\delta^* P_\delta)$  only by  $O(\delta)$ . In other words, there exists a  $\delta$ -independent constant  $K_1 = |p|_{0,1} |p|_{1,0}$ , such that

$$\|Q_\delta - (I - P_\delta^* P_\delta)\|_{L^2} \leq K_1 \delta. \quad (2.20)$$

From (2.19) and (2.20), one gets

$$-\|u\|_{L^2}^2 + \|P_\delta u\|_{L^2}^2 \leq -(u, Q_\delta u) + K_1 \delta \|u\|_{L^2}^2.$$

Obviously,  $q$  is Hermitian, and according to the assumption (2.17),  $q$  is also non-negative definite. Therefore, according to Theorem 2, there is some constant  $K_2$  such that

$$\mathbf{Re}(u, Q_\delta u) \geq -K_2 \delta (u, u). \quad (2.21)$$

If  $\|u\|_{L^2}^2 - \|P_\delta u\|_{L^2}^2 \leq 0$ , then by (2.20) and (2.21), there is some positive constant  $K$  (for example,  $K = K_1 + K_2$ ), such that

$$-\|u\|_{L^2}^2 + \|P_\delta u\|_{L^2}^2 = \mathbf{Re}\left(-\|u\|_{L^2}^2 + \|P_\delta u\|_{L^2}^2\right) \leq K\delta \|u\|_{L^2}^2,$$

which implies

$$\|P_\delta u\|_{L^2}^2 \leq (1 + K\delta) \|u\|_{L^2}^2.$$

If  $\|u\|_{L^2}^2 - \|P_\delta u\|_{L^2}^2 \geq 0$ , the estimate above is also satisfied. This completes the proof.  $\square$

Therefore, for the semi-Lagrangian method, if one takes  $\delta = \Delta x$ , and  $\Delta t = O(\Delta x)$ , then  $\Delta x$  and  $\Delta t$  can be treated as a single parameter. We take  $M = \Delta t/\Delta x$ , which characterizes the ratio of temporal and spatial mesh. Thus, the numerical scheme can be written as

$$u(x + \Delta t) = P_\delta u(x).$$

As long as one can show  $|p(x, \xi)| \leq 1$  for all  $x$  and  $\xi$ , then by Lemma 1,

$$\begin{aligned} \|u(x, t_0 + n\Delta t)\|_{L^2} &\leq \|P_\delta^n u(x, t_0)\|_{L^2} \leq (1 + K\delta)^n \|u(x, t_0)\|_{L^2} \\ &\leq e^{Kn\Delta x} \|u(x, t_0)\|_{L^2} = e^{\frac{Kn\Delta t}{M}} \|u(x, t_0)\|_{L^2} \leq e^{\frac{KnT}{M}} \|u(x, t_0)\|_{L^2}. \end{aligned} \quad (2.22)$$

This implies, for fixed  $M = O(1)$ , in other words when  $\Delta t = O(\Delta x)$ , the semi-Lagrangian method is stable as long as  $|p(x, \xi)| \leq 1$  for all  $x$  and  $\xi$ . The reason is with fixed  $M = \Delta t/\Delta x$ , the constant  $K$  in (2.18) is  $\delta$ -independent.

Another implication is in the limit  $M = \Delta t/\Delta x \rightarrow 0$ , the above stability proof fails. However, we would like to allow  $\Delta t = o(1)$  despite  $\Delta x = O(\varepsilon)$ , so the Courant number can be as large as  $O(\frac{1}{\varepsilon})$ . We further show in the next subsection the stability conditions for the semi-Lagrangian method with arbitrary Courant number  $C_c = \|A\|_{L^\infty} \Delta t/\Delta x > c_0$ , where  $c_0 > 0$ .

## 2.4 Stability of the semi-Lagrangian method for arbitrary Courant numbers

The proof above for the stability of the semi-Lagrangian method does not apply directly, for arbitrary Courant numbers. Instead, in this subsection we explore the dependence of  $P_\delta$  on both  $\Delta t$  and  $\Delta x$ . Especially, we aim to derive the estimates of  $P_\delta$  which are valid even when  $\Delta t \gg \Delta x$ .

For arbitrary Courant number cases, one expects to derive the requirements such that  $P_\delta$  is bounded in the following way,

$$\|P_\delta\|_{L^2} < 1 + K(\Delta x + \Delta t), \quad (2.23)$$

for some constant  $K$  independent of  $\Delta x$  and  $\Delta t$ . However, when  $\Delta t \gg \Delta x$ , one can no longer treat the constants  $K_1$  in (2.20) and  $K_2$  in (2.21) simply as bounded quantities. Hence, one needs to derive the dependence of those constants on both  $\Delta t$  and  $\Delta x$ .

When  $A(x)$  has zeros, even for bounded  $A(x)$ , there is no obvious way to find desired bounds for certain  $(k, l)$  norms of the symbol of  $P_\delta$  such that (2.23) is satisfied. To simplify the analysis, we firstly introduce a method to decompose the semi-Lagrangian difference operator  $P_\delta$  into a product of a shift operator and another semi-Lagrangian difference operator  $\tilde{P}_\delta$ , which corresponds to a

different characteristic velocity  $-\tilde{A}(x)$ , where  $\tilde{A}$  is positive. We will derive estimates for certain  $(k, l)$  norms of the symbol of  $\tilde{P}_\delta$ , which help to prove the stability of the semi-Lagrangian method with arbitrary Courant numbers.

We now introduce the decomposition of the semi-Lagrangian difference operator. Since  $A(x)$  is bounded, assume  $a_{\min} \leq A(x) \leq a_{\max}$ . We naturally assume  $\|A\|_{L^\infty} > 0$ , because otherwise only the trivial case  $A \equiv 0$  is allowed. Suppose for each point  $x_{\text{in}}$  at time level  $t = t^{n+1}$ , the backward characteristic passing through it hits the point  $x^0$  at time level  $t = t^n$ . Notice,  $\forall m \in \mathbb{Z}$ , one can rewrite the difference operator as

$$P_\delta = \sum_{\alpha} p_{\alpha}(x_{\text{in}}) T_{\alpha} = \left( \sum_{\beta} \tilde{p}_{\beta}(x_{\text{in}}) T_{\beta} \right) T_m = \tilde{P}_{\delta} T_m,$$

where  $\beta = \alpha - m$  and  $\tilde{p}_{\beta}(x_{\text{in}}) = p_{\beta+m}(x_{\text{in}})$ . Actually,  $\forall m \in \mathbb{Z}$ , the shift operator  $T_m$  can be seen as the semi-Lagrangian difference operator corresponding to the exact method of characteristic with the characteristic speed  $-\frac{m\Delta x}{\Delta t}$ . Thus, the operator  $\tilde{P}_{\delta}$  is the semi-Lagrangian difference operator with velocity  $-A(x) + \frac{m\Delta x}{\Delta t}$ . This observation allows one to rewrite the original semi-Lagrangian difference operator  $P_{\delta}$  as a product of a shift operator and another semi-Lagrangian difference operator  $\tilde{P}_{\delta}$  with different velocity field, though the difference in corresponding velocities is only a constant.

Since in Section 2.3, the cases when the Courant number  $C_c = O(1)$  were treated, to complete the analysis for arbitrary Courant number cases, without the loss of generality, we assume  $C_c > 1$ . Next, we choose a special integer  $m$ , so that the characteristic velocity that the semi-Lagrangian difference operator  $\tilde{P}_{\delta}$  corresponds to is negative and bounded away from 0. Since  $C_c = \|A\|_{L^\infty} \Delta t / \Delta x > 1$ , we claim that there exists a constant  $b \in [1, 2]$ , such that  $(a_{\min} - b\|A\|_{L^\infty}) \Delta t / \Delta x = m_0 \in \mathbb{Z}$ . Actually, consider the linear function,

$$y(s) = -(\|A\|_{L^\infty} \Delta t / \Delta x) s + a_{\min} \Delta t / \Delta x$$

with the slope  $-(\|A\|_{L^\infty} \Delta t / \Delta x) < -1$  and the domain  $\mathcal{D} = [1, 2]$ , so the length of the range is larger than 1. Thus, there exists a point  $b \in [1, 2]$ , such that  $y(b)$  equals an integer, and we denote this integer by  $m_0$ .

Therefore, the operator  $T_{m_0}$  can be seen as the semi-Lagrangian difference operator that corresponds to the constant characteristic speed  $-a_{\min} + b\|A\|_{L^\infty}$ , and  $\tilde{P}_{\delta} = P_{\delta} T_{-m_0}$  corresponds to the semi-Lagrangian difference operator with the same  $\Delta x$  and  $\Delta t$ , but the velocity  $-A(x)$  is replaced by  $-\tilde{A}(x) = -(A(x) - a_{\min} + b\|A\|_{L^\infty})$ . We denote the symbol of  $\tilde{P}_{\delta}$  by  $\tilde{p}$ , then obviously,  $\tilde{p} = p e^{-im_0\xi}$ . For different pairs of  $\Delta x$  and  $\Delta t$ , one may have different  $\tilde{P}_{\delta}$ , but the corresponding  $\tilde{A}(x)$  has a uniform lower bound and a uniform upper bound:

$$\tilde{A}(x) = A(x) - a_{\min} + b\|A\|_{L^\infty} \geq b\|A\|_{L^\infty} \geq \|A\|_{L^\infty}; \quad (2.24)$$

$$\tilde{A}(x) = A(x) - a_{\min} + b \|A\|_{L^\infty} \leq 2 \|A\|_{L^\infty} + b \|A\|_{L^\infty} \leq 4 \|A\|_{L^\infty}. \quad (2.25)$$

Also, observe that  $A(x)$  and  $\tilde{A}(x)$  have the same derivatives, so we denote  $\|\partial_x A\|_{L^\infty} = \left\| \partial_x \tilde{A} \right\|_{L^\infty} = L_1$  and  $\|\partial_x^2 A\|_{L^\infty} = \left\| \partial_x^2 \tilde{A} \right\|_{L^\infty} = L_2$ .

Since the polynomial interpolation is applied, the coefficients  $\tilde{p}_\beta(x_{\text{in}})$  of  $\tilde{P}_\delta$  are Lagrange basis functions, which are polynomials of  $r(x_{\text{in}})$ , with

$$r(x_{\text{in}}) = \frac{x^0 - x_l}{\Delta x}, \quad (2.26)$$

where  $x_l$  is the first point on the left hand side of  $x^0$  such that for some integer  $n_0$ ,  $(x_l - x_{\text{in}}) / \Delta x = n_0$ . To give estimates of certain  $(k, l)$  norms of  $\tilde{p}$ , we study the derivative of  $r$  with respect to  $x_{\text{in}}$ .

Consider the characteristic equation

$$\frac{dx(t)}{dt} = -\tilde{A}(x(t)), \quad (2.27)$$

with  $x(\Delta t) = x_{\text{in}}$ . Denote the solution to this initial value problem by  $\mathbf{x}(t; x_{\text{in}})$ . Integrate the equation from  $\Delta t$  to 0, one gets

$$x^0 = x_{\text{in}} + \int_0^{\Delta t} \tilde{A}(\mathbf{x}(s; x_{\text{in}})) ds. \quad (2.28)$$

Since  $\tilde{A}(x) \in C^1$ ,  $\mathbf{x}(t; x_{\text{in}})$  is continuously differentiable with respect to  $x_{\text{in}}$ . Therefore, there exists a constant  $\eta$ , such that  $\left| \frac{d\mathbf{x}(t; x_{\text{in}})}{dx_{\text{in}}} \right| < \eta$ .

Because  $\tilde{A}(x) \geq \|A\|_{L^\infty} > 0$ , there exists  $t' \in [0, \Delta t]$ , such that

$$x_l = x_{\text{in}} + \int_0^{t'} \tilde{A}(\mathbf{x}(s; x_{\text{in}})) ds. \quad (2.29)$$

In other words, one gets

$$n_0 = \frac{x_l - x_{\text{in}}}{\Delta x} = \frac{\int_0^{t'} \tilde{A}(\mathbf{x}(s; x_{\text{in}})) ds}{\Delta x}. \quad (2.30)$$

Rewrite  $r = \int_{t'}^{\Delta t} \tilde{A}(\mathbf{x}(s; x_{\text{in}})) ds / \Delta x$ . Then one gets the following estimate for its derivative with respect to  $x_{\text{in}}$ ,

$$\begin{aligned} \left| \frac{dr}{dx_{\text{in}}} \right| &= \frac{\left| \int_{t'}^{\Delta t} \partial_x \tilde{A}(\mathbf{x}(s; x_{\text{in}})) \frac{d\mathbf{x}}{dx_{\text{in}}} ds \right|}{\Delta x} \\ &\leq \frac{\eta (\Delta t - t')}{\Delta x} \left\| \partial_x \tilde{A} \right\|_{L^\infty} = \frac{\eta (\Delta t - t') L_1}{\Delta x}. \end{aligned}$$

By the definition of  $x_l$ , one has  $0 \leq x^0 - x_l \leq \Delta x$ , so (2.28) and (2.29) imply

$$0 \leq \int_{t'}^{\Delta t} \tilde{A}(\mathbf{x}(s; x_{\text{in}})) ds \leq \Delta x,$$

then the estimate (2.24) implies,

$$(\Delta t - t') \|A\|_{L^\infty} \leq \int_{t'}^{\Delta t} \tilde{A}(\mathbf{x}(s; x_{\text{in}})) ds \leq \Delta x, \quad (2.31)$$

and thus one concludes

$$\left| \frac{dr}{dx_{\text{in}}} \right| \leq \frac{\eta(\Delta t - t') L_1}{\Delta x} \leq \frac{\eta L_1}{\|A\|_{L^\infty}}. \quad (2.32)$$

This means,  $\frac{dr}{dx_{\text{in}}}$  is uniformly bounded with respect to  $\Delta x$  and  $\Delta t$ . Similarly, one can show  $\frac{d^2 r}{dx_{\text{in}}^2}$  is also uniformly bounded. We remark that, without introducing  $\tilde{A}$ , one cannot get the estimate (2.31).

We're ready to prove the following lemma, which gives the estimate of the  $L^2$  norm of the semi-Lagrangian difference operator  $P_\delta$  when the Courant number exceeds 1.

**Lemma 2.** *Let  $P_\delta$  be the semi-Lagrangian difference operator with  $\delta = \Delta x$  and  $C_c = \|A\|_{L^\infty} \Delta t / \Delta x > 1$ , as is defined above. Suppose its symbol  $p(x, \xi)$  is bounded by 1:*

$$|p(x, \xi)| \leq 1, \quad \forall x, \xi \in \mathbb{R}. \quad (2.33)$$

*Then for all  $\delta$ , there exists a constant  $K$ , such that the operator  $P_\delta$  is bounded in the following way:*

$$\|P_\delta\|_{L^2} \leq 1 + K(\Delta x + \Delta t). \quad (2.34)$$

*Proof.* For the semi-Lagrangian difference operators  $P_\delta$  when  $C_c = \|A\|_{L^\infty} \Delta t / \Delta x > 1$ , we introduce the decomposition  $P_\delta = \tilde{P}_\delta T_{m_0}$  as is defined above. Observe, for all  $L^2$  functions  $u$ ,

$$\begin{aligned} \|u\|_{L^2}^2 - \|P_\delta u\|_{L^2}^2 &= (u, (I - P_\delta^* P_\delta) u) \\ &= \left( u, \left( I - \left( \tilde{P}_\delta T_{m_0} \right)^* \tilde{P}_\delta T_{m_0} \right) u \right) \\ &= \left( u, \left( T_{-m_0} T_{m_0} - T_{-m_0} \left( \tilde{P}_\delta \right)^* \tilde{P}_\delta T_{m_0} \right) u \right) \\ &= \left( T_{m_0} u, \left( I - \left( \tilde{P}_\delta \right)^* \tilde{P}_\delta \right) T_{m_0} u \right). \end{aligned}$$

Next, we define the symbol  $\tilde{q}(x, \xi) = 1 - \tilde{p}^* \tilde{p}$ , and denote the difference operator with symbol  $\tilde{q}$  by  $\tilde{Q}_\delta$ .

To make use of Theorem 1 and Theorem 2, we estimate some  $(k, l)$  norms of the symbols  $\tilde{p}$  and  $\tilde{q}$  to be used in the proof. Recall that  $\tilde{p} = \sum_\beta \tilde{p}_\beta(r(x_{\text{in}})) e^{i\beta\xi}$ , since  $\tilde{p}_\beta(r(x_{\text{in}}))$  are polynomials of  $r$ , and the first derivative of  $r$  is bounded uniformly with respect to  $\Delta x$  and  $\Delta t$ , so  $|\tilde{p}_\beta|_0$  and  $|\tilde{p}_\beta|_1$  are bounded on any bounded interval of  $r$ . In practice, the interval of  $r$  is determined so that the condition  $|p(x, \xi)| \leq 1$  is satisfied. In the next subsection, we derive the interval of  $r$  when the fourth order interpolation is applied, so that  $|p(x, \xi)| \leq 1$ .

If the  $M^{\text{th}}$  order polynomial interpolation is applied, one then has  $|\beta| \leq n_0 + M$ , where  $M$  is

fixed and

$$n_0 = \frac{\int_0^{t'} \tilde{A}(\mathbf{x}(s, x_{\text{in}})) ds}{\Delta x} \leq \frac{\int_0^{\Delta t} \tilde{A}(\mathbf{x}(s, x_{\text{in}})) ds}{\Delta x} \leq 4 \|A\|_{L^\infty} \frac{\Delta t}{\Delta x}.$$

Therefore,  $1 + |\beta| \leq K_1 \left(1 + \frac{\Delta t}{\Delta x}\right)$  for some constant  $K_1$ . In conclusion, on some bounded interval of  $r$  and with  $M$  fixed,  $|\tilde{p}|_{1,0} = \sum_\beta |\tilde{p}_\beta|_1$  is bounded and  $|\tilde{p}|_{0,1} = \sum_\beta |\tilde{p}_\beta|_0 (1 + |\beta|) = O\left(1 + \frac{\Delta t}{\Delta x}\right)$ .

Next, with

$$\tilde{q} = 1 - \tilde{p}^* \tilde{p} = \sum_\zeta \tilde{q}_\zeta(r) e^{i\zeta \xi},$$

we now prove that, if  $|\zeta| > 2M$ ,  $\tilde{q}_\zeta = 0$ . Actually

$$\begin{aligned} \tilde{q} &= 1 - \left( \sum_{\gamma_1} \tilde{p}_{\gamma_1}(r) e^{i\gamma_1 \xi} \right) \left( \sum_{\gamma_2} \tilde{p}_{\gamma_2}(r) e^{-i\gamma_2 \xi} \right) \\ &= e^{i0\xi} - \sum_{\gamma_1, \gamma_2} \tilde{p}_{\gamma_1} \tilde{p}_{\gamma_2} e^{i(\gamma_1 - \gamma_2)\xi}. \end{aligned}$$

For any specific  $x$ , if the  $M$ -point interpolation is applied, the set  $\{\tilde{p}_\gamma(x)\}$  has at most  $M$  nonzero elements, we denote the corresponding indices by  $\gamma^{(1)}(x), \dots, \gamma^{(M)}(x)$ , which are  $M$  consecutive integers. Therefore, in the summation  $\sum_{\gamma_1, \gamma_2} \tilde{p}_{\gamma_1} \tilde{p}_{\gamma_2} e^{i(\gamma_1 - \gamma_2)\xi}$ , one has  $\tilde{p}_{\gamma_1}(x) \tilde{p}_{\gamma_2}(x) = 0$  unless  $\gamma_1, \gamma_2 \in \{\gamma^{(1)}(x), \dots, \gamma^{(M)}(x)\}$ . This implies, in this summation, the nonzero contributions occur only when  $|\gamma_1 - \gamma_2| \leq 2M$ . Due to the arbitrariness of  $x$ , one concludes, in the summation,  $\tilde{q}(x, \xi) = \sum_\zeta \tilde{q}_\zeta(x) e^{i\zeta \xi}$ , one has  $|\zeta| \leq 2M$ .

So when  $M$  is fixed,  $1 + |\zeta| \leq 2M + 1$ . Since the coefficients  $\tilde{q}_\zeta(r)$  are also polynomials of  $r$ , whose first and second derivatives with respect to  $x_{\text{in}}$  are bounded on bounded intervals of  $r$ , so  $|\tilde{q}_\beta|_0, |\tilde{q}_\beta|_1$  and  $|\tilde{q}_\beta|_2$  are bounded on bounded intervals of  $r$ . In conclusion, on bounded intervals of  $r$  and with  $M$  fixed,  $|\tilde{q}|_{2,0} = \sum_\beta |\tilde{q}_\beta|_2$  is bounded, and  $|\tilde{q}|_{0,2} = \sum_\zeta |\tilde{q}_\zeta|_0 (1 + |\zeta|)^2$  is also bounded. Therefore, the symbol  $\tilde{q}$  is in  $C_{2,0} \cap C_{0,2}$ .

By Theorem 1, the operator  $\tilde{Q}_\delta$  differs from  $I - (\tilde{P}_\delta)^* \tilde{P}_\delta$  in the  $L^2$  norm at most by  $\delta |\tilde{p}|_{1,0} |\tilde{p}|_{0,1}$ , so there exists a constant  $K_3$ , such that

$$\left\| \tilde{Q}_\delta - \left( I - (\tilde{P}_\delta)^* \tilde{P}_\delta \right) \right\|_{L^2} \leq K_3 \left( 1 + \frac{\Delta t}{\Delta x} \right) \delta = K_3 (\Delta x + \Delta t).$$

Then one gets,

$$\begin{aligned} -\|u\|_{L^2}^2 + \|P_\delta u\|_{L^2}^2 &= -\left( T_{m_0} u, \left( I - (\tilde{P}_\delta)^* \tilde{P}_\delta \right) T_{m_0} u \right) \\ &\leq -\left( T_{m_0} u, \tilde{Q}_\delta T_{m_0} u \right) + \left\| \tilde{Q}_\delta - \left( I - (\tilde{P}_\delta)^* \tilde{P}_\delta \right) \right\|_{L^2} \|T_{m_0} u\|_{L^2}^2 \\ &\leq -\left( T_{m_0} u, \tilde{Q}_\delta T_{m_0} u \right) + K_3 (\Delta x + \Delta t) \|u\|_{L^2}^2. \end{aligned} \tag{2.35}$$

Recall that, we have shown that the symbol  $\tilde{q} \in C_{2,0} \cap C_{0,2}$ . Besides, one can easily verify that



$q$  is Hermitian, non-negative, which implies  $\tilde{q}$  is also Hermitian and non-negative. Actually, by definition,  $\tilde{q} = 1 - \tilde{p}^* \tilde{p}$  is obviously Hermitian. And  $\tilde{q} = 1 - \tilde{p}^* \tilde{p} = 1 - |\tilde{p}|^2 = 1 - |p|^2 \geq 0$ , because  $|p| < 1$  by assumption. Therefore,  $\tilde{q}$  satisfies all the hypotheses in Theorem 2, so there exists a constant  $K_4$  such that

$$\mathbf{Re} \left( T_{m_0} u, \tilde{Q}_\delta T_{m_0} u \right) \geq -K_4 \delta \|T_{m_0} u\|_{L^2}^2 = -K_4 \Delta x \|u\|_{L^2}^2. \quad (2.36)$$

If  $\|u\|_{L^2}^2 - \|P_\delta u\|_{L^2}^2 \leq 0$ , by (2.35) and (2.36), there exists some constant  $K$  (for example, one can take  $K = K_3 + K_4$ ), such that

$$-\|u\|_{L^2}^2 + \|P_\delta u\|_{L^2}^2 = \mathbf{Re} \left( -\|u\|_{L^2}^2 + \|P_\delta u\|_{L^2}^2 \right) \leq K (\Delta x + \Delta t) \|u\|_{L^2}^2,$$

which implies

$$\|P_\delta u\|_{L^2}^2 \leq (1 + K (\Delta x + \Delta t)) \|u\|_{L^2}^2.$$

If  $\|u\|_{L^2}^2 - \|P_\delta u\|_{L^2}^2 \geq 0$ , the estimate above is also satisfied. Hence, we have shown when  $C_c = \|A\|_{L^\infty} \Delta t / \Delta x > 1$ , the condition that  $|p(x, \xi)| \leq 1$  for all  $x$  and  $\xi$  implies that there is some constant  $K$ , such that

$$\|P_\delta\|_{L^2} \leq 1 + K(\Delta x + \Delta t).$$

□

According to Lemma 2, when the Courant number  $C_c = \|A\|_{L^\infty} \Delta t / \Delta x$  exceeds 1, which means  $\Delta x < \|A\|_{L^\infty} \Delta t$ , one gets

$$\begin{aligned} \|u(x, t_0 + n\Delta t)\|_{L^2} &\leq \|P_\delta\|_{L^2}^n \|u(x, t_0)\|_{L^2} \\ &\leq (1 + C(\Delta x + \Delta t))^n \|u(x, t_0)\|_{L^2} \\ &\leq e^{Cn(\Delta x + \Delta t)} \|u(x, t_0)\|_{L^2} \\ &\leq e^{Cn(1 + \|A\|_{L^\infty})\Delta t} \|u(x, t_0)\|_{L^2} \\ &\leq e^{C(1 + \|A\|_{L^\infty})T} \|u(x, t_0)\|_{L^2}, \end{aligned}$$

with  $n\Delta t \leq T$  and  $C$  is independent of  $\Delta x$  and  $\Delta t$ .

Therefore, by Lemma 1 and Lemma 2, we have shown that when the Courant number  $C_c = \|A\|_{L^\infty} \Delta t / \Delta x \geq c_0$  for some  $c_0 > 0$ , the semi-Lagrangian method is stable if  $|p(x, \xi)| \leq 1$  for all  $x$  and  $\xi$ . The stability requirement for the semi-Lagrangian method is basically the norm of the symbol  $p(x, \xi)$  is bounded by 1 with the exception that the proof fails in the limit  $\Delta t / \Delta x \rightarrow 0$  with  $\Delta x$  fixed. In other words, as long as the Courant number  $C_c = \|A\|_{L^\infty} \Delta t / \Delta x \geq c_0$  for some  $c_0 > 0$ , the condition that the norm of the symbol  $p(x, \xi)$  is no greater than 1 is sufficient to prove stability.

We summarize the result in the following theorem.

**Theorem 3.** *Consider the semi-Lagrangian scheme*

$$u(t + \Delta t, x) = P_\delta u(t, x), \quad (2.37)$$

where  $P_\delta$  is the difference operator in the form (2.14), and  $\delta = \Delta x$ . Suppose, there is negligible error in tracing the characteristics, and suppose there exists a positive constant  $c_0$  such that

$$\|A\|_{L^\infty} \Delta t / \Delta x > c_0 > 0. \quad (2.38)$$

Then if for all  $x$  and  $\xi$ , the norm of the symbol  $p(x, \xi)$  is no greater than 1,  $|p(x, \xi)| \leq 1$ , the semi-Lagrangian scheme is stable in the sense that

$$\|u(T)\|_{L^2} \leq C_T \|u(0)\|_{L^2}$$

for all solutions to the scheme (2.37), where  $C_T$  depends on  $T$  and  $c_0$ , but is independent of  $\Delta x$ ,  $\Delta t$ .

## 2.5 Semi-Lagrangian method with the fourth order polynomial interpolation

According to Theorem 3, as long as the Courant number is bounded away from 0, the semi-Lagrangian method is stable when the norm of the symbol  $p(x, \xi)$  is no greater than 1. Therefore, for specific orders of interpolation, one just needs to work out the requirement on the choice of interpolation points such that this condition is satisfied. Since in Section 6 we choose to use cubic polynomial interpolation ( $M = 4$ ) in numerical examples, we carry out the detailed calculation for this choice.

Define  $r = (x_j^0 - x_{p_j+2}) / \Delta x$ , and the scheme can be written as

$$U_j^{n+1} = \sum_{m=1}^4 l_{m+p_j}(r) U_{m+p_j}^n, \quad (2.39)$$

where

$$l_{p_j+1}(r) = -\frac{r(1-r)(2-r)}{6}, \quad l_{p_j+2} = \frac{(1+r)(1-r)(2-r)}{2},$$

$$l_{p_j+3}(r) = \frac{(1+r)r(2-r)}{2}, \quad l_{p_j+4} = -\frac{(1+r)r(1-r)}{6}.$$

To be more specific, one needs to work out the interval of  $r$  for this choice of interpolation, so that the norm of the symbol of this semi-Lagrangian difference operator is no greater than 1.

For simplicity, we drop  $p_j$  in the sub-indices, and instead write  $l_1, l_2, l_3, l_4$ . By direct calculation,

$$l_1(r) + l_4(r) = -\frac{r(1-r)}{2}, \quad l_2(r) + l_3(r) = \frac{(1+r)(2-r)}{2}$$

$$l_1(r) + l_2(r) + l_3(r) + l_4(r) = 1. \quad (2.40)$$

So the symbol of the difference operator is

$$p(x, \xi) = \left( l_1(r)e^{-i3\xi/2} + l_2(r)e^{-i\xi/2} + l_3(r)e^{i\xi/2} + l_4(r)e^{i3\xi/2} \right) e^{i\xi(p_j - j + 2 + \frac{1}{2})}.$$

Multiply by  $\bar{p}$  on both sides, one gets,

$$\begin{aligned} |p(x, \xi)|^2 &= l_1^2 + l_2^2 + l_3^2 + l_4^2 + 2l_1l_4 \cos(3\xi) + 2(l_1l_3 + l_2l_4) \cos(2\xi) \\ &\quad + 2(l_1l_2 + l_2l_3 + l_3l_4) \cos(\xi). \end{aligned}$$

Note, the identity (2.40) implies  $(l_1 + l_2 + l_3 + l_4)^2 = 1$ . So one gets,

$$\begin{aligned} |p(x, \xi)|^2 &= 1 + 2l_1l_4 (\cos(3\xi) - 1) + 2(l_1l_3 + l_2l_4) (\cos(2\xi) - 1) \\ &\quad + 2(l_1l_2 + l_2l_3 + l_3l_4) (\cos(\xi) - 1). \end{aligned}$$

With trigonometric identities  $\cos(2\xi) = 2\cos^2(\xi) - 1$  and  $\cos(3\xi) = 4\cos^3(\xi) - 3\cos(\xi)$ , one gets

$$\begin{aligned} |p(x, \xi)|^2 &= 1 - \frac{1}{3}r(1-r)(1+r)(2-r)(1 - \cos(\xi))^2 \\ &\quad - \frac{2}{9}r^2(1-r)^2(1+r)(2-r)(1 - \cos(\xi))^3 \\ &= 1 + r(1-r)(1+r)(2-r)(1 - \cos(\xi))^2 S(r, \xi), \end{aligned}$$

where,

$$S(r, \xi) = \frac{1}{9} (2(1 - \cos(\xi))r^2 - 2(1 - \cos(\xi))r - 3).$$

$S(r, \xi)$  is a quadratic function in  $r$  with parameter  $\xi$  in the coefficients, we denote its roots by  $r_1$  and  $r_2$ , where  $r_1 \leq r_2$ . When  $\cos(\xi) = 1$ , one gets  $|p(x, \xi)|^2 = 1$ . When  $\frac{1}{4} < \cos(\xi) < 1$ ,  $r_1 < -1$ ,  $r_2 > 2$ , and  $|p(x, \xi)|^2 \leq 1$  means  $r \in [r_1, -1] \cup [0, 1] \cup [2, r_2]$ . When  $\cos(\xi) = \frac{1}{4}$ ,  $r_1 = -1$ ,  $r_2 = 2$ , and  $|p(x, \xi)|^2 \leq 1$  means  $r \in \{-1\} \cup [0, 1] \cup \{2\}$ . When  $-1 \leq \cos(\xi) < \frac{1}{4}$ ,  $-1 < r_1 < 0$ ,  $1 < r_2 < 2$ , and  $|p(x, \xi)|^2 \leq 1$  means  $r \in [-1, r_1] \cup [0, 1] \cup [r_2, 2]$ . Based on the analysis above, one sees the condition  $|p(x, \xi)|^2 \leq 1$  is equivalent to  $r \in \{-1\} \cup [0, 1] \cup \{2\}$ . So in practice, we take the stability interval  $[0, 1]$  for  $r$ .

Therefore, the semi-Lagrangian method with the fourth order polynomial interpolation is stable when the Courant number is bounded away from 0, and  $x_j^0$  is between  $x_{p_j+2}$  and  $x_{p_j+3}$ , namely, when the shifted grid points are between the second points and the third points in interpolation. Similarly, one can derive stability requirements for other polynomial interpolations, for example, see calculations in [6].

We give the following two remarks for stability constraints.

*Remark 4.* For the semi-Lagrangian method with polynomial interpolation, the stability constraint

comes from two parts: solving the characteristic ODE (2.10) and the polynomial interpolation. The latter aspect has been studied in depth so far. On the other hand, solving the characteristic equation (2.10) with time step  $\widetilde{\Delta t}$  requires

$$\|\partial_x A\|_{L^\infty} \widetilde{\Delta t} \leq C, \quad (2.41)$$

for some constant  $C$  to guarantee stability. Actually, one can take sufficiently small  $\widetilde{\Delta t}$  in solving the ODE (2.10) to get a very accurate approximation for the characteristics. This can be done in a preprocessed step, and thus does not affect the computational cost in time evolution.

*Remark 5.* Another widely used necessary stability criterion for the semi-Lagrangian method is the deformation constraint proposed in [18, 27], which means characteristics initiated from adjacent grid points do not intersect in  $\widetilde{\Delta t}$  time. This criterion actually leads to the same constraint (2.41).

So one concludes, for stability,  $\Delta t$  in the semi-Lagrangian method is  $\Delta x$ -independent, and thus  $\varepsilon$ -independent in the sense that it requires that the interpolation points are chosen properly, as is stated in Theorem 3. Especially, one is allowed to take  $\Delta t \gg \Delta x$ . In comparison, the stability constraint for TESP method is  $\|A\|_{L^\infty} \Delta t / \Delta x \leq C$  for some constant  $C$ , which indicates  $\Delta t = O(\Delta x) = O(\varepsilon)$ .

Note that, since the numerical schemes in the kinetic step and the potential step are realized by exact time integration, the numerical methods are unconditionally stable. As is analyzed above, the stability constraint for the semi-Lagrangian method in solving the convection part is  $\Delta x$ -independent, and thus  $\varepsilon$ -independent. Therefore, one concludes the SL-TS method allows  $\Delta t$  to be independent of  $\Delta x$  and  $\varepsilon$ . We will show in later sections, the SL-TS method possesses great advantages in relaxing the time step restriction over TESP time splitting method in computing physical observables.

### 3 Error estimates in the presence of vector potential

In this section, we study the error in approximating the wave function and the meshing strategy of the SL-TS method. We use  $\|\cdot\|_{l^2}$  to denote the discrete  $l^2$  norm

$$\|U\|_{l^2} = \left( \frac{b-a}{N} \sum_{j=0}^{N-1} |U_j|^2 \right)^{\frac{1}{2}}. \quad (3.1)$$

We further assume that, the wave functions are  $\varepsilon$ -oscillatory in space and time but the potentials are not oscillatory. So there are  $t, \varepsilon, x$  independent positive constants  $B_m, C_m, D_m$  so that

$$\left\| \frac{\partial^{m_1+m_2}}{\partial x^{m_1} \partial t^{m_2}} u(t, x) \right\|_{C([0, T]; L^2(a, b))} \leq \frac{1}{\varepsilon^{m_1+m_2}} C_{m_1+m_2},$$

$$\left\| \frac{\partial^m}{\partial x^m} A(x) \right\|_{L^2(a,b)} \leq D_m, \quad \left\| \frac{\partial^m}{\partial x^m} U(x) \right\|_{L^2(a,b)} \leq B_m. \quad (3.2)$$

Note that the differentiation operator is unbounded for general smooth functions, but it is bounded in the subspace of smooth  $L^2$  function which are at most  $\varepsilon$ -oscillatory. We use  $f_I$  to denote the spectral approximation based on the discrete data  $f_j$  or  $f(x_j)$ . Now we are ready to prove the following error estimate for the first order SL-TS method. The proof basically follows Theorem 4.1 in [5], though the situation here is more complicated due to the convection step.

**Theorem 4.** *Let  $u^\varepsilon(t, x)$  be the exact solution of equation (2.1),  $u^{\varepsilon, n}$  be the discrete approximation by the first order SL-TS method. We assume the characteristic equations (2.10) are numerically solved with minimal error in the preprocessed step, the  $M^{\text{th}}$  order polynomial interpolation is taken in the semi-Lagrangian method for the convection step and the corresponding stability condition is satisfied. Under assumption (3.2), we further assume  $\Delta x/\varepsilon = O(1)$  and  $\Delta t/\varepsilon = O(1)$ ; then for all positive integers  $m \geq 1$  and  $t \in [0, T]$ ,*

$$\|u^\varepsilon(t_n) - u_I^{\varepsilon, n}\|_{L^2} \leq G_{m, M} \frac{T}{\Delta t} \left[ \left( \frac{\Delta x}{\varepsilon} \right)^m + \left( \frac{\Delta x}{\varepsilon} \right)^M \right] + \frac{CT\Delta t}{\varepsilon}, \quad (3.3)$$

where  $C$  is a positive constant independent of  $\Delta t$ ,  $\Delta x$ ,  $\varepsilon$ ,  $m$  and  $M$ , and  $G_{m, M}$  are positive constants independent of  $\Delta t$ ,  $\Delta x$  and  $\varepsilon$ .

*Proof.* Recall that, we have defined operator splitting solution (without spatial discretization) (2.6)  $w^{n+1} = e^{\mathcal{A}\Delta t} e^{\mathcal{B}\Delta t} e^{\mathcal{C}\Delta t} u^\varepsilon(t_n)$ . We firstly show, by studying the commutators between three operators in (2.5), when potentials are spatially variant, the local splitting error in equations (2.2), (2.3), (2.4) for equation (2.1) is:

$$\|u^\varepsilon(t_{n+1}) - w^{n+1}\|_{L^2} = O\left(\frac{\Delta t^2}{\varepsilon}\right). \quad (3.4)$$

Clearly, the exact solution to (2.1) at  $t = t_{n+1}$  with initial data  $u^\varepsilon(t_n)$  is given by

$$u^\varepsilon(t_{n+1}) = e^{(\mathcal{A}+\mathcal{B}+\mathcal{C})\Delta t} u^\varepsilon(t_n).$$

The operator splitting error results from the non-commutativity of the operators  $\mathcal{A}$ ,  $\mathcal{B}$  and  $\mathcal{C}$ . In [5], it was shown that

$$[\mathcal{A}\Delta t, \mathcal{B}\Delta t]u^\varepsilon = O\left(\frac{\Delta t^2}{\varepsilon}\right),$$

where  $[\cdot, \cdot]$  denotes the commutator. Similarly, by a direct computation:

$$\begin{aligned} [\mathcal{A}\Delta t, \mathcal{C}\Delta t]u^\varepsilon &= (\Delta t)^2 \frac{i\varepsilon}{2} (\partial_{xx} \mathbf{A} \partial_x u^\varepsilon + \partial_x \mathbf{A} \partial_{xx} u^\varepsilon) = O\left(\frac{\Delta t^2}{\varepsilon}\right); \\ [\mathcal{B}\Delta t, \mathcal{C}\Delta t]u^\varepsilon &= (\Delta t)^2 \left(-\frac{i}{\varepsilon}\right) \left(\mathbf{A} \partial_x \left(\frac{1}{2} |\mathbf{A}|^2 + V\right)\right) u^\varepsilon = O\left(\frac{\Delta t^2}{\varepsilon}\right). \end{aligned}$$

Therefore, we have shown the local operator splitting error is  $O\left(\frac{\Delta t^2}{\varepsilon}\right)$  as in (3.4).

By triangle inequality,

$$\|u^\varepsilon(t_{n+1}) - u_I^{\varepsilon, n+1}\|_{L^2} \leq \|u^\varepsilon(t_{n+1}) - w^{n+1}\|_{L^2} + \|w^{n+1} - w_I^{n+1}\|_{L^2} + \|w_I^{n+1} - u_I^{\varepsilon, n+1}\|_{L^2} \quad (3.5)$$

where  $w_I^{n+1}$  denotes the spectral interpolation approximation of  $w^{n+1}$ . On the right hand side of (3.5), the first term gives the operator splitting error (3.4), and the second term gives the spectral approximation error which is bounded by  $C_m \left(\frac{\Delta x}{\varepsilon}\right)^m$ , where  $m$  can be any positive integer, which depends on the regularity of the solution (see [22]).

Now we focus on the last term on the right hand side of (3.5). With the spectral approximation, we have for any periodic function  $f \in L^2(a, b)$ ,  $\|f_I\|_{L^2} = \|f\|_{l^2}$ . In the SL-TS method, the potential step governed by operator  $\mathcal{B}$  is solved analytically, while the kinetic step and convection step governed by operators  $\mathcal{A}$  and  $\mathcal{C}$  are evolved by numerical approximations, denoted by  $\mathcal{A}_{SP}$  and  $\mathcal{C}_{SL}$  respectively. So once again, by triangle inequality:

$$\begin{aligned} \|w_I^{n+1} - u_I^{\varepsilon, n+1}\|_{L^2} &= \|w^{n+1} - u^{\varepsilon, n+1}\|_{l^2} \\ &= \|e^{\mathcal{C}\Delta t} e^{\mathcal{B}\Delta t} e^{\mathcal{A}\Delta t} u^\varepsilon(t_n) - e^{\mathcal{C}_{SL}\Delta t} e^{\mathcal{B}\Delta t} e^{\mathcal{A}_{SP}\Delta t} u^{\varepsilon, n}\|_{l^2} \\ &\leq \|e^{\mathcal{C}\Delta t} e^{\mathcal{B}\Delta t} e^{\mathcal{A}\Delta t} u^\varepsilon(t_n) - e^{\mathcal{C}\Delta t} e^{\mathcal{B}\Delta t} e^{\mathcal{A}_{SP}\Delta t} u^\varepsilon(t_n)\|_{l^2} \\ &\quad + \|e^{\mathcal{C}\Delta t} e^{\mathcal{B}\Delta t} e^{\mathcal{A}_{SP}\Delta t} u^\varepsilon(t_n) - e^{\mathcal{C}_{SL}\Delta t} e^{\mathcal{B}\Delta t} e^{\mathcal{A}_{SP}\Delta t} u^\varepsilon(t_n)\|_{l^2} \\ &\quad + \|e^{\mathcal{C}_{SL}\Delta t} e^{\mathcal{B}\Delta t} e^{\mathcal{A}_{SP}\Delta t} u^\varepsilon(t_n) - e^{\mathcal{C}_{SL}\Delta t} e^{\mathcal{B}\Delta t} e^{\mathcal{A}_{SP}\Delta t} u^{\varepsilon, n}\|_{l^2}. \end{aligned} \quad (3.6)$$

Note that, the first term on the right hand side of (3.6) measures the spectral approximation of  $u^\varepsilon(t_n)$ , so by Theorem 3 from [22], this term is of order  $O\left(\left(\frac{\Delta x}{\varepsilon}\right)^m\right)$  for any positive integer  $m$ . According to Remark 3, the error in computing the shifted grid points is much smaller than the interpolation error, so the second term on the right hand side is bounded by the  $M^{\text{th}}$  order polynomial interpolation error, which is  $O\left(\left(\frac{\Delta x}{\varepsilon}\right)^M\right)$ . It can easily be shown that the operators  $e^{\mathcal{A}\Delta t}$ ,  $e^{\mathcal{B}\Delta t}$  and  $e^{\mathcal{C}\Delta t}$  (in the Coulomb gauge) are unitary operators with respect to periodic smooth functions in the  $L^2$  norm, which implies  $\|e^{\mathcal{A}\Delta t}\|_{L^2} = \|e^{\mathcal{B}\Delta t}\|_{L^2} = \|e^{\mathcal{C}\Delta t}\|_{L^2} = 1$ . Bao, Jin and Markowich in [5] have proved that  $\mathcal{A}_{SP}$  is also unitary with respect to  $u_I^\varepsilon$ , the spectral approximation of smooth  $L^2$  function. Note,  $e^{\mathcal{C}_{SL}\Delta t}$  is not a unitary operator, but with stability constraint as in Section 2.3, we have  $\|e^{\mathcal{C}_{SL}\Delta t}\|_{L^2}^n \leq C'$  for some constant  $C'$  and  $n\Delta t \leq T$ . So, the last term on the right hand side of (3.6)

$$\begin{aligned} &\|e^{\mathcal{C}_{SL}\Delta t} e^{\mathcal{B}\Delta t} e^{\mathcal{A}_{SP}\Delta t} u^\varepsilon(t_n) - e^{\mathcal{C}_{SL}\Delta t} e^{\mathcal{B}\Delta t} e^{\mathcal{A}_{SP}\Delta t} u^{\varepsilon, n}\|_{l^2} \\ &\leq \|e^{\mathcal{C}_{SL}\Delta t}\|_{L^2} \|e^{\mathcal{B}\Delta t} e^{\mathcal{A}_{SP}\Delta t} u^\varepsilon(t_n) - e^{\mathcal{B}\Delta t} e^{\mathcal{A}_{SP}\Delta t} u^{\varepsilon, n}\|_{l^2} \\ &\leq \|e^{\mathcal{C}_{SL}\Delta t}\|_{L^2} \|u^\varepsilon(t_n) - u_I^{\varepsilon, n}\|_{L^2}. \end{aligned}$$

This leads to

$$\|w_I^{n+1} - u_I^{\varepsilon, n+1}\|_{L^2} \leq \|e^{C_{SL}\Delta t}\|_{L^2} \|u(t_n) - u_I^{\varepsilon, n}\|_{L^2} + C'_m \left(\frac{\Delta x}{\varepsilon}\right)^m + C''_M \left(\frac{\Delta x}{\varepsilon}\right)^M,$$

where  $C'_m$  and  $C''_M$  are some constants, which are  $t, x, \varepsilon$  independent. So now, we have derived a recursive relation

$$\begin{aligned} \|u^\varepsilon(t_{n+1}) - u_I^{\varepsilon, n+1}\|_{L^2} &\leq \|e^{C_{SL}\Delta t}\|_{L^2} \|u^\varepsilon(t_n) - u_I^{\varepsilon, n}\|_{L^2} \\ &\quad + C_1 \left(\frac{\Delta x}{\varepsilon}\right)^m + C_2 \left(\frac{\Delta x}{\varepsilon}\right)^M + C_3 \left(\frac{\Delta t^2}{\varepsilon}\right) \end{aligned}$$

where  $C_1, C_2$  and  $C_3$  are some constants, which are  $t, x, \varepsilon$  independent.

Based on the recursive relation for  $\|u^\varepsilon(t_n) - u_I^{\varepsilon, n}\|_{L^2}$ , by induction, one concludes that,

$$\|u^\varepsilon(t_n) - u_I^{\varepsilon, n}\|_{L^2} \leq G_{m, M} \frac{T}{\Delta t} \left[ \left(\frac{\Delta x}{\varepsilon}\right)^m + \left(\frac{\Delta x}{\varepsilon}\right)^M \right] + \frac{CT\Delta t}{\varepsilon}. \quad (3.7)$$

This completes the proof.  $\square$

We remark that, in practice, if the solution is sufficiently regular, the error introduced by polynomial interpolation is dominant in spatial discretization since  $m$  can be chosen to be fairly large. Then in practice, it suffices to consider the following error bound

$$\|u^\varepsilon(t_n) - u_I^{\varepsilon, n}\|_{L^2} \leq \tilde{G}_M \frac{T}{\Delta t} \left(\frac{\Delta x}{\varepsilon}\right)^M + \frac{CT\Delta t}{\varepsilon}. \quad (3.8)$$

This implies that, if one wants to control the error of the wave function in the  $L^2$  norm so that  $\|u^\varepsilon(t_n) - u_I^{\varepsilon, n}\|_{L^2} < \delta$ , the corresponding meshing strategy is

$$\frac{\Delta t}{\varepsilon} = O(\delta), \quad \frac{\Delta x}{\varepsilon} = O\left(\delta^{1/M} \Delta t^{1/M}\right). \quad (3.9)$$

For higher order operator splitting technique, similar analysis can be done, which is omitted in this paper.

## 4 Computing the physical observables

In general, if one only cares about the physical observables, weaker conditions in the meshing strategy may be sufficient (see [20, 5]). The Wigner transform can be used to illustrate this point. For  $f, g \in L^2(\mathbb{R}^d)$ , the Wigner transform is defined as a phase-space function

$$w^\varepsilon(f, g)(t, x, \xi) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{iy \cdot \xi} \bar{f}\left(x + \frac{\varepsilon}{2}y\right) g\left(x - \frac{\varepsilon}{2}y\right) dy. \quad (4.1)$$

Recall that  $u^\varepsilon(t, x)$  is the exact solution of equation (2.1). Denote  $w^\varepsilon = w^\varepsilon(u^\varepsilon, u^\varepsilon)$ , which satisfies the Wigner equation

$$\partial_t w^\varepsilon + \xi \cdot \nabla_x w^\varepsilon + \Theta[V + |\mathbf{A}|^2/2]w^\varepsilon + \Gamma[\mathbf{A}]w^\varepsilon = 0, \quad (4.2)$$

in which two pseudo-differential operators are defined by

$$\Theta[U]w^\varepsilon := \frac{i}{(2\pi)^d \varepsilon} \int_{R^d} \left( U\left(x + \frac{\varepsilon}{2}\alpha\right) - U\left(x - \frac{\varepsilon}{2}\alpha\right) \right) \hat{w}^\varepsilon(x, \alpha, t) e^{i\alpha \cdot \xi} d\alpha; \quad (4.3)$$

$$\begin{aligned} \Gamma[\mathbf{A}]w^\varepsilon := & -\frac{1}{(2\pi)^d} \int_{R^d} \mathbf{A}\left(x + \frac{\varepsilon}{2}\alpha\right) \cdot \nabla_x \bar{u}\left(x + \frac{\varepsilon}{2}\alpha\right) u\left(x - \frac{\varepsilon}{2}\alpha\right) e^{i\alpha \cdot \xi} d\alpha \\ & -\frac{1}{(2\pi)^d} \int_{R^d} \bar{u}\left(x + \frac{\varepsilon}{2}\alpha\right) \mathbf{A}\left(x - \frac{\varepsilon}{2}\alpha\right) \cdot \nabla_x u\left(x - \frac{\varepsilon}{2}\alpha\right) e^{i\alpha \cdot \xi} d\alpha. \end{aligned} \quad (4.4)$$

By Weyl's Calculus, as  $\varepsilon \rightarrow 0$ , the Wigner measure  $w^0 = \lim_{\varepsilon \rightarrow 0} w^\varepsilon(u^\varepsilon, u^\varepsilon)$  satisfies the classical Liouville equation

$$w_t^0 + (\xi - \mathbf{A}) \cdot \nabla_x w^0 + ((\xi - \mathbf{A}) \cdot \nabla_x \mathbf{A} - \nabla_x U) \cdot \nabla_\xi w^0 = 0, \quad (4.5)$$

with

$$w^0(t=0, x, \xi) = w_I^0(x, \xi) := \lim_{\varepsilon \rightarrow 0} w^\varepsilon(u_0^\varepsilon, u_0^\varepsilon). \quad (4.6)$$

All the limits above are defined in an appropriate weak sense (see [31, 11]).

Now let  $a(x, \xi)$  be a smooth real-valued phase space function with sufficient decay at infinity, called a semi-classical symbol. Then the self-adjoint pseudo-differential operator  $A^\varepsilon := a(x, \varepsilon D)^W$  is called an observable, where  $D = i\nabla_x$ , and  $W$  stands for the Weyl quantization (see [11]). If one specifies a quantum state  $u^\varepsilon(t, x)$ , then the average of this observable in this state is defined as

$$E_a^\varepsilon(t) = \int_{R^d} \bar{u}^\varepsilon(t, x) (a(x, \varepsilon D)^W u^\varepsilon(t, x)) dx. \quad (4.7)$$

One significant property of the Wigner transform is that it establishes the duality identity in the following sense

$$\int_{R^d} \bar{u}^\varepsilon(t, x) (a(x, \varepsilon D)^W u^\varepsilon(t, x)) dx = \int_{R^d \times R^d} w^\varepsilon(t, x, \xi) a(x, \xi) dx d\xi. \quad (4.8)$$

As a consequence,  $E_a^\varepsilon(t)$  can be taken to the semi-classical limit via

$$\lim_{\varepsilon \rightarrow 0} E_a^\varepsilon(t) = \int_{R^d \times R^d} w^0(t, x, \xi) a(x, \xi) dx d\xi. \quad (4.9)$$

These semi-classical limits have been mathematically justified in [30, 23].

Let  $\tilde{w}^\varepsilon$  be the Wigner transform of the numerical approximation solution. One can easily prove



the following inequality (see [5])

$$|E_a^\varepsilon - \tilde{E}_a^\varepsilon| \leq \|a\|_{\mathcal{E}} \cdot \|w^\varepsilon - \tilde{w}^\varepsilon\|_{\mathcal{E}^*} \leq C \|a\|_{\mathcal{E}} \cdot \|u^\varepsilon - \tilde{u}^\varepsilon\|_{L^2(a,b)}, \quad (4.10)$$

where  $\mathcal{E}$  is the Banach space

$$\mathcal{E} = \left\{ \phi \in C_0(\mathbb{R}_x^d \times \mathbb{R}_\xi^d) : (\mathcal{F}_{\xi \rightarrow v} \phi) \in L^1(\mathbb{R}_v^d; C_0(\mathbb{R}_x^d)) \right\}.$$

$\mathcal{F}$  denotes the Fourier transform and  $\mathcal{E}^*$  is the dual space of  $\mathcal{E}$  (see [19]). This inequality is not sharp, but it shows that, the  $L^2$  approximation of the wave function at least implies approximation of mean value of physical observables in the same order.

In each time step  $t \in [t_n, t_{n+1}]$  after operator splitting, the error in the wave function is introduced due to spectral approximation and polynomial approximation. By Theorem 4 and the estimate (4.10), the error in the corresponding Wigner transform can be estimated. Although it might not be optimal, the spatial meshing strategy  $\frac{\Delta x}{\varepsilon} = O(\delta^{1/M} \Delta t^{1/M})$  is sufficient to guarantee an  $O(\delta)$  error in all physical observables caused by spectral and polynomial approximations on the time interval  $[0, T]$ .

The splitting error in computing the physical observables can be understood in the following way. The time splitting in solving the Schrödinger equation corresponds to the time splitting of the Wigner equations: one firstly solves

$$w_t^\varepsilon + \xi \cdot \nabla_x w^\varepsilon = 0, \quad t \in [t_n, t_{n+1}],$$

followed by solving

$$w_t^\varepsilon + \Gamma[\mathbf{A}]w^\varepsilon = 0, \quad t \in [t_n, t_{n+1}],$$

and then followed by solving

$$w_t^\varepsilon + \Theta[V + \frac{1}{2}|\mathbf{A}|^2]w^\varepsilon = 0, \quad t \in [t_n, t_{n+1}].$$

For fixed  $\Delta t$ , one can take the limit  $\varepsilon \rightarrow 0$ , and obtain the time splitting of the classical Liouville equation: one firstly solves

$$w_t^0 + \xi \cdot \nabla_x w^0 = 0, \quad t \in [t_n, t_{n+1}],$$

followed by solving

$$w_t^0 - \mathbf{A} \cdot \nabla_x w^\varepsilon + \xi \cdot \nabla_x \mathbf{A} \cdot \nabla_\xi w^0 = 0, \quad t \in [t_n, t_{n+1}],$$

and then followed by solving

$$w_t^0 - (\mathbf{A} \cdot \nabla_x \mathbf{A} + \nabla_x V) \cdot \nabla_\xi w^0 = 0, \quad t \in [t_n, t_{n+1}].$$

Consider the SL-TS method introduced in Section 2. As is summarized at the end of Section 2, if the interpolation points are chosen properly according to Theorem 3, the whole SL-TS method is stable even when  $\Delta t \gg \varepsilon$ . Next, note that no  $\varepsilon$ -dependent error is introduced by the splitting. In the kinetic step and the potential step, the time integrations are performed exactly. In the convection part, the backward characteristic tracing is done in a preprocessed step with sufficiently fine yet  $\varepsilon$  independent time steps. Therefore, there is no  $\varepsilon$ -dependent error at all in time discretizations.

After all these considerations, we conclude that with the SL-TS method, large time steps satisfying  $\Delta t \gg \varepsilon$  can be taken to capture correct physical observables. In other words, with time step  $\Delta t = O(\delta)$  and spatial meshing strategy (3.9), one gets numerical solutions with  $O(\delta)$  error in the Wigner functions as  $\varepsilon \rightarrow 0$ , and as a result,  $O(\delta)$  error in all the physical observables.

*Remark 6.* When the vector potential  $A$  is time dependent, the above analysis still holds, because  $A$  is  $\varepsilon$ -independent, and thus in the backward characteristic tracing step,  $\widetilde{\Delta t}$  is also  $\varepsilon$ -independent.

## 5 Extension to multidimensional cases

In this section, we discuss how to extend the SL-TS method to the multidimensional cases. The Schrödinger equation (1.1) can be written as

$$i\varepsilon\partial_t u^\varepsilon = -\frac{\varepsilon^2}{2}\Delta u^\varepsilon + i\varepsilon\mathbf{A} \cdot \nabla u^\varepsilon + \frac{1}{2}|\mathbf{A}|^2 u^\varepsilon + V u^\varepsilon, \quad x \in \Pi_{i=1, \dots, d}[a_i, b_i], \quad t \in \mathbb{R}^+; \quad (5.1)$$

$$u^\varepsilon(0, x) = u_0^\varepsilon(x), \quad x \in \Pi_{i=1, \dots, d}[a_i, b_i]$$

with periodic boundary condition. By operator splitting technique, for every time step  $t \in [t_n, t_{n+1}]$ , we solve the kinetic step

$$i\varepsilon\partial_t u^\varepsilon = -\frac{\varepsilon^2}{2}\Delta u^\varepsilon = -\frac{\varepsilon^2}{2} \sum_{l=1}^d \partial_{x_l}^2 u^\varepsilon, \quad t \in [t_n, t_{n+1}]; \quad (5.2)$$

followed by the potential step

$$i\varepsilon\partial_t u^\varepsilon = \frac{1}{2}|\mathbf{A}|^2 u^\varepsilon + V u^\varepsilon, \quad t \in [t_n, t_{n+1}], \quad (5.3)$$

and then followed the convection step

$$\partial_t u^\varepsilon = \mathbf{A} \cdot \nabla u^\varepsilon = \sum_{l=1}^d \mathbf{A}_l(x_1, \dots, x_d) \partial_{x_l} u^\varepsilon, \quad t \in [t_n, t_{n+1}]. \quad (5.4)$$

We remark that the kinetic step and the potential step are exactly the original TSSP method as in [5] without vector potential, while the convection step can be done by dimensional splitting. In fact the dimensional splitting will not introduce  $\varepsilon$  dependent error in time because the classical

Liouville equation is also well separated in the same dimensional splitting for the convection step. Note that, the convection step of the Liouville equation can be reformulated as

$$\partial_t w^0 - \sum_{l=1}^d \mathbf{A}_l \partial_{x_l} w^0 + \sum_{l=1}^d \xi_l \nabla_x \mathbf{A}_l \cdot \nabla_\xi w^0 = 0, \quad t \in [t_n, t_{n+1}]. \quad (5.5)$$

After dimensional splitting one still observes the one-to-one correspondence between the equations (5.4) and (5.5). In addition, the semi-classical limit of the Schrödinger equation in each dimension is exactly the counterpart of the classical Liouville equation. Therefore, all numerical methods can be carried out by the same means and the meshing strategy would stay unchanged.

We remark that, the dimension splitting is only one of possible techniques to use. Multidimensional semi-Lagrangian method is also applicable, see [28, 18, 6].

In Section 6, we implement this method to a particular three dimensional model from physics, which can be reduced to a two dimensional one, to verify the numerical properties of the SL-TS method in higher dimensions.

## 6 Numerical examples

In the series of numerical tests, the reference solution is computed by time splitting method with sufficiently fine mesh grids, but the convection part is numerically evolved by the time-explicit spectral method (TESP), which means we apply the spectral approximation for spatial derivative and an explicit ODE solver (here we use the fourth order Runge-Kutta method) is used in time discretization. The SL-TS method is implemented with the fourth order (four-point) Lagrange polynomial interpolation. In the following sets of examples, we want to test improved stability, convergence in space and time and the ability of capturing correct physical observables with large time step when the vector potential is time dependent, or when caustics are formed. We also test an essentially 2-D problem to verify the SL-TS in higher dimensions.

The error in wave functions is measured in the  $L^2$  norm, but the error in physical observables is measured in the weak sense. We define cumulative distribution function of physical observable  $b(x, t)$  as

$$B(x, t) = \int_{-\infty}^x b(s, t) ds.$$

We compute the  $l^2$  norm of the error in these cumulative distribution functions instead since the physical observables may converge only in the weak sense (see [7, 12]).

**Example 1.** In the construction of the SL-TS scheme, we assume the vector potential is time independent. But, this method can handle time dependent vector potentials with little modification. In this example, we test the following one dimensional problem with a time dependent vector potential. We choose the computation domain  $C = [0, 2\pi]$  and compute from  $t_0 = 0$  till  $T = 0.5$ . The scalar potential is chosen as  $V(x) = (x - \pi)^2$  and the vector potential is  $\mathbf{A}(x, t) = \sin(x - 2t)$ . The initial condition is chosen as  $\varphi_0 = e^{-5(x-\pi)^2} e^{i \cos(x)/\varepsilon}$ .

We firstly test the improved stability condition of the scheme and convergence in time. We choose  $\varepsilon = \frac{1}{128}$ , the reference solution is computed by the TESP method with  $\Delta x = \frac{2\pi}{5120}$  and  $\Delta t = \frac{1}{4096}$ . We test the scheme with the same spatial mesh grids but different time steps:  $\Delta t = \frac{1}{50}, \frac{1}{100}, \frac{1}{200}, \frac{1}{400}, \frac{1}{800}$ . As can be seen from Figure 6.1, even with  $\Delta t \gg \Delta x$ , the scheme is still stable and gives correct first order of convergence in time for wave functions and physical observables.

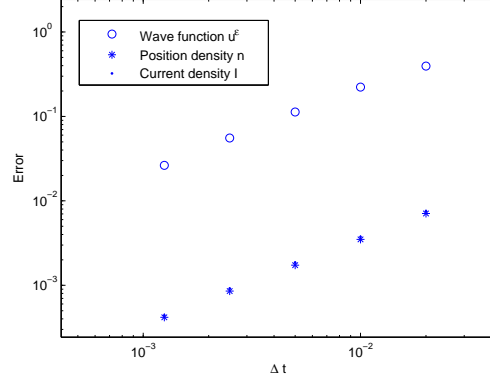


Figure 6.1: Reference solution:  $\Delta x = \frac{2\pi}{5120}$ ,  $\Delta t = \frac{1}{4096}$ . SL-TS method:  $\Delta x = \frac{2\pi}{5120}$ ,  $\Delta t = \frac{1}{50}, \frac{1}{100}, \frac{1}{200}, \frac{1}{400}, \frac{1}{800}$ .  $\circ \circ \circ$ : error in wave function  $u^\varepsilon$ ;  $* * *$ : error in mass density  $n$ ;  $\cdot \cdot \cdot$ : error in flux density  $\mathbf{I}$ .

Next, we test whether one can compute physical observables with the meshing strategy  $\Delta t = O(\sigma)$  and roughly  $\Delta x = O(\varepsilon)$ . We test our scheme for  $\varepsilon = \frac{1}{128}, \frac{1}{256}, \frac{1}{512}, \frac{1}{1024}, \frac{1}{2048}$  with  $\Delta t = 0.01$  and sufficiently fine spatial grids. We plot numerical approximation of mass density  $n$  and current density  $\mathbf{I}$  (in circles) together with the reference solution (in solid lines) for  $\varepsilon = \frac{1}{128}, \frac{1}{2048}$  in Figure 6.2.

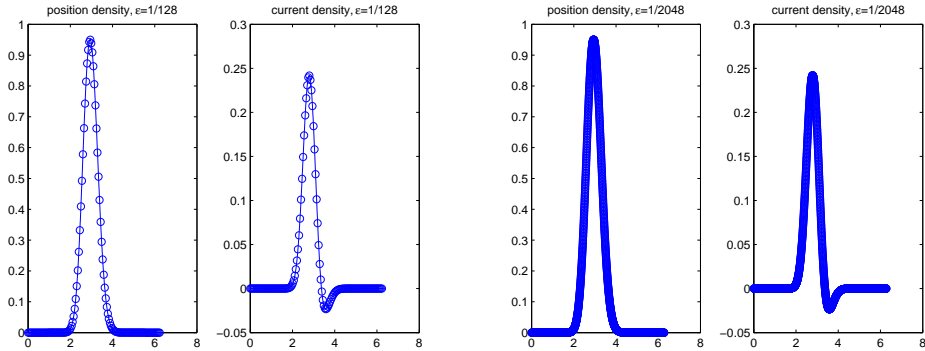


Figure 6.2: Left:  $\varepsilon = \frac{1}{128}$ , reference solution  $\Delta x = \frac{2\pi}{2560}$ ,  $\Delta t = \frac{1}{1280}$ ; SL-TS method  $\Delta x = \frac{2\pi}{2560}$ ,  $\Delta t = 0.01$ . Right:  $\varepsilon = \frac{1}{2048}$ , reference solution  $\Delta x = \frac{2\pi}{40960}$ ,  $\Delta t = \frac{1}{20480}$ ; SL-TS method  $\Delta x = \frac{2\pi}{40960}$ ,  $\Delta t = 0.01$ .

With fixed large time step and  $\Delta x = O(\varepsilon)$ , we expect that error in the wave function increases as  $\varepsilon$  decreases, but the error in physical observables would stay the same order. This set of numerical tests have confirmed the expectation as shown in Figure 6.3.

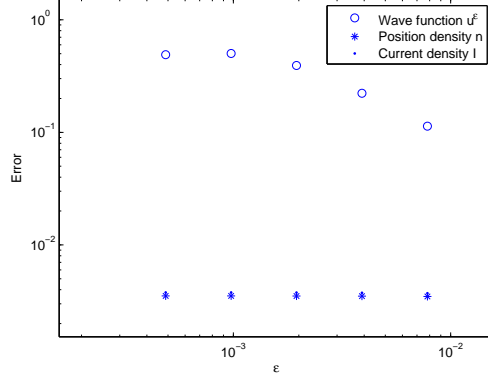


Figure 6.3: SL-TS method: fixed  $\Delta t = 0.01$  for  $\varepsilon = \frac{1}{128}, \frac{1}{256}, \frac{1}{512}, \frac{1}{1024}, \frac{1}{2048}$ , and correspondingly  $\Delta x = \frac{2\pi}{2560}, \frac{2\pi}{5120}, \frac{2\pi}{10240}, \frac{2\pi}{20480}, \frac{2\pi}{40960}$ .

To show that meshing strategy with  $\Delta t = O(\delta)$  gives correct physical observables but not the correct wave function, we can examine the error in wave functions and in physical observables when  $\varepsilon = \frac{1}{2048}$  with the  $\Delta t = 0.01$  and  $\Delta x = \frac{2\pi}{40960}$ . Since  $\Delta t \gg \varepsilon$ , one sees in Figure 6.4  $O(1)$  error in wave function but  $O(\Delta t)$  error in physical observables.

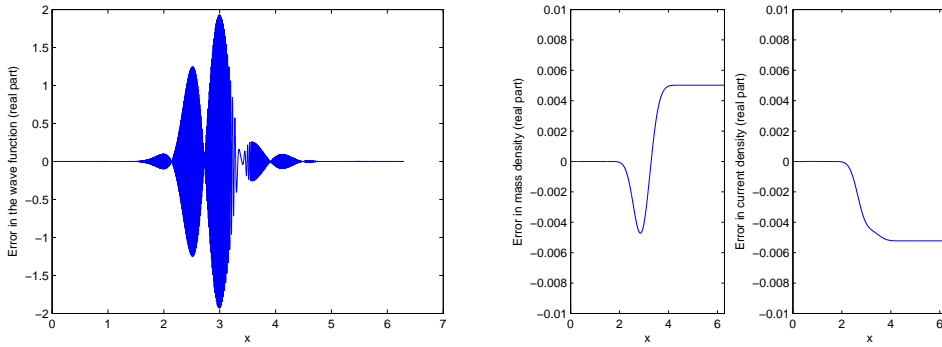


Figure 6.4: Left:  $L^2$  error in wave functions. Right:  $L^2$  error in mass density and current density.

We remark that, one can get similar numerical results for the problem with time-independent vector potentials.

**Example 2.** In this test, we check the solution after caustics formation. As shown in previous research, most prevailing schemes for Schrödinger equation, for example the Crank–Nicolson spectral method (CNSP) and the Crank–Nicolson finite difference method (CNFD), may fail to capture the correct physical observable when wave functions are not resolved either in space or in time (see [5]). In addition, these methods even require finer mesh both in spatial grids and time steps to compute accurate wave functions. The failure in obtaining accurate physical observables with unresolved mesh is most obvious after caustics formation.

In this example, the scalar potential is chosen as  $V(x) = 1$  and the vector potential is  $\mathbf{A}(x) = \sin(x/2\pi)/5 + 1/5$ . Note that, the constant part in vector potential gives purely spatial translation while the varying part affects the profile of the wave functions and therefore modify the profile of physical observables. The initial condition is chosen as in the WKB form, where

$$\varphi_0 = n_0(x)e^{iS_0(x)/\varepsilon}, \quad n_0(x) = e^{-25(x-0.5)^2}, \quad S_0 = -\frac{1}{5} \ln \left( e^{5(x-0.5)} + e^{-5(x-0.5)} \right).$$

Due to the compressive initial velocity  $\frac{d}{dx}S_0(x)$ , caustics will form. We now test whether one obtains accurate physical observables with  $\Delta t = O(\sigma)$  and roughly  $\Delta x = O(\varepsilon)$  meshing strategy. Note that, in order not to affect the caustics too much, we choose scalar and vector potential to be less spatially varying.

We choose the computation domain  $C = [0, 1]$  and compute from  $t_0 = 0$  till  $T = 0.54$ . We test the SL-TS scheme for  $\varepsilon = \frac{1}{128}, \frac{1}{256}, \frac{1}{512}, \frac{1}{1024}, \frac{1}{2048}$  with  $\Delta t = 0.01$  and  $O(\varepsilon)$  spatial grids. We plot numerical approximation of mass density and current density (in dots) together with the reference solution (in solid lines) for  $\varepsilon = \frac{1}{128}, \frac{1}{2048}$  in Figure 6.5, which shows good agreements.

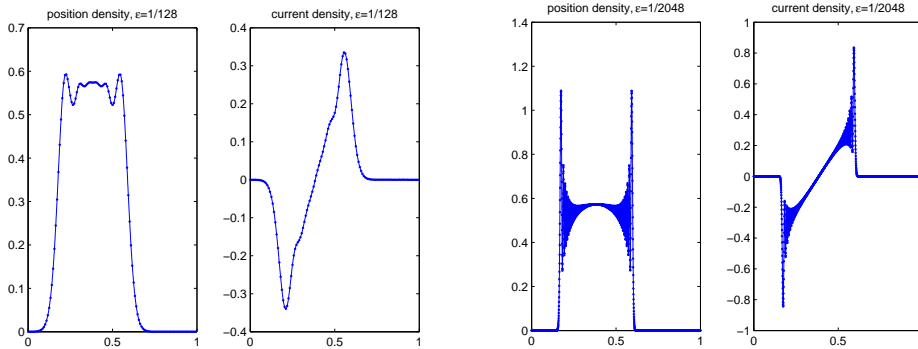


Figure 6.5: Left:  $\varepsilon = \frac{1}{128}$ , reference solution  $\Delta x = \frac{1}{2560}$ ,  $\Delta t = \frac{1}{1280}$ ; SL-TS method  $\Delta x = \frac{1}{2560}$ ,  $\Delta t = 0.01$ . Right:  $\varepsilon = \frac{1}{2048}$ , reference solution  $\Delta x = \frac{1}{40960}$ ,  $\Delta t = \frac{1}{20480}$ ; SL-TS method  $\Delta x = \frac{1}{40960}$ ,  $\Delta t = 0.01$ .

With fixed large time step and  $\Delta x = O(\varepsilon)$ , we expect that error in the wave function increases as  $\varepsilon$  decreases, but the error in physical observables would stay almost the same order. This set of

numerical tests have confirmed the expectation as shown in Figure 6.6.

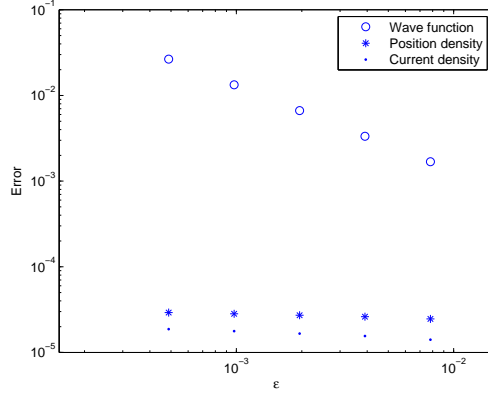


Figure 6.6: SL-TS method: fixed  $\Delta t = 0.01$  for  $\varepsilon = \frac{1}{128}, \frac{1}{256}, \frac{1}{512}, \frac{1}{1024}, \frac{1}{2048}$ , and correspondingly  $\Delta x = \frac{1}{2560}, \frac{1}{5120}, \frac{1}{10240}, \frac{1}{20480}, \frac{1}{40960}$ .

At last, we test spatial convergence. We fix  $\varepsilon = \frac{1}{128}$ , and compute this test problem with fine time steps and  $\Delta x/\varepsilon = \frac{1}{5}, \frac{1}{10}, \frac{1}{20}, \frac{1}{40}$ . The reference solution is computed with sufficiently fine mesh,  $\Delta x = \frac{1}{10240}$  and  $\Delta t = \frac{1}{20480}$ . We plot the reference wave function (both real and imaginary parts) and physical observables in Figure 6.7.

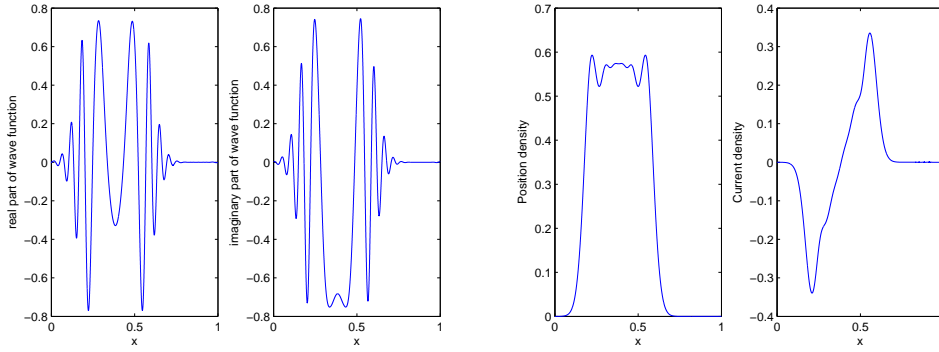


Figure 6.7:  $\varepsilon = \frac{1}{128}$ . Left: the wave function (real part and imaginary part). Right: mass density and current density.

Since we use the four-point Lagrangian interpolation approximation for the convection step, according to meshing strategy (3.9), the convergence order in spatial grids should be slightly worse than  $O(\Delta x^4)$ . Figure 6.13 shows the convergence order in spatial grids is between 3 and 4.

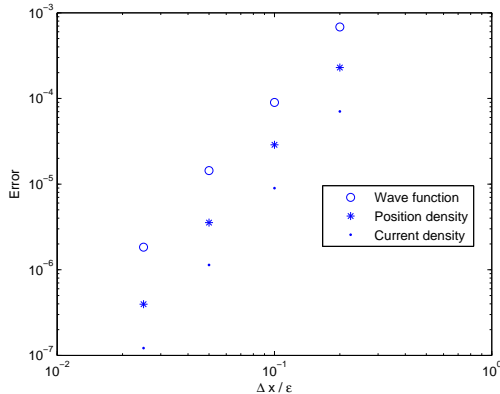


Figure 6.8: Spatial grids convergence for  $\epsilon = \frac{1}{128}$ , with fixed  $\Delta t = \frac{1}{2560}$  and  $\Delta x = \frac{1}{640}, \frac{1}{1280}, \frac{1}{2560}, \frac{1}{5120}$ .

**Example 3.** In this example, we test our method for an essentially 2-D model. This is a commonly-used physics model (see [25]). A charged particle is moving in the three dimensional space when the magnetic field is pointing along  $z$  axis and the scalar potential vanishes, namely  $V = 0$ . We write wave function as  $u^\epsilon(t, \mathbf{X})$ , where  $\mathbf{X} = (x, y, z) \in R^3$ . For simplicity, we assume the components of the vector potential are

$$\mathbf{A}_x = A_1(x, y), \quad \mathbf{A}_y = A_2(x, y), \quad \mathbf{A}_z = 0.$$

Note that, when  $\mathbf{A}_x = -\frac{1}{2}By$ ,  $\mathbf{A}_y = \frac{1}{2}Bx$ ,  $\mathbf{A}_z = 0$ , the vector potential corresponds to uniform magnetic field with magnitude  $B$  along  $z$  direction. Obviously, the Coulomb gauge is satisfied,  $\nabla_X \cdot \mathbf{A} = 0$ . The magnitude of the magnetic field in this simplified model is  $\partial_x A_2 - \partial_y A_1$ . Then the Hamiltonian can be written as

$$H = -\frac{\epsilon^2}{2} (\Delta_x + \Delta_y) + i\epsilon A_1 \partial_x + i\epsilon A_2 \partial_y + \frac{1}{2} (|A_1|^2 + |A_2|^2) - \frac{\epsilon^2}{2} \Delta_z. \quad (6.1)$$

This means that, the particles have only free motion in the  $z$  direction. So it makes perfect sense to consider the motion of the particle on the  $x-y$  plane only by introducing the reduced Hamiltonian,

$$\tilde{H} = -\frac{\epsilon^2}{2} (\Delta_x + \Delta_y) + i\epsilon A_1 \partial_x + i\epsilon A_2 \partial_y + \frac{|A_1|^2 + |A_2|^2}{2}. \quad (6.2)$$

We remark that, this simple model can help to derive the Larmor frequency, which plays an essential role in magnetic or spin resonance.

For numerical simulation, the computation domain is chosen as  $[-\pi, \pi] \times [-\pi, \pi]$ . The vector potential is chosen as  $\mathbf{A} = (A_1, A_2, 0) = (-\frac{1}{2} \cos(y), \frac{1}{2} \cos(x), 0)$  and scalar potential vanishes



$V = 0$ . The initial wave function is well localized at the point  $(0.05, 0.1)$  with  $O(\varepsilon)$  oscillation

$$u_0(x, y) = e^{-20(x-0.05)^2 - 20(y-0.1)^2} e^{i \sin(x) \sin(y) / \varepsilon}.$$

We choose  $\varepsilon = \frac{1}{64}$ ,  $\Delta x = \Delta y = \frac{2\pi}{1280}$ , and compare numerical solutions by fine time step  $\Delta t = \frac{1}{640}$  and by coarse time step  $\Delta t = \frac{1}{20}$  at  $T = 0.2$  and  $T = 0.4$ , respectively. We plot level curves of mass density in each case in Figure 6.9.

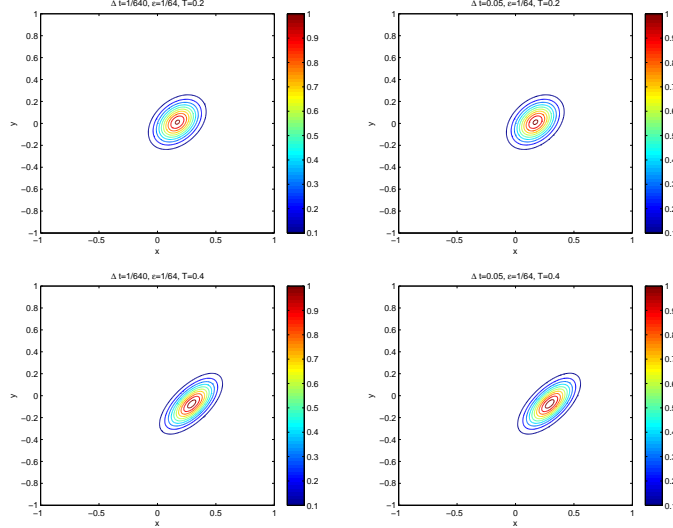


Figure 6.9: Level curves of mass density for  $\varepsilon = \frac{1}{64}$  at  $T = 0.2, 0.4$ . Left:  $\Delta x = \Delta y = \frac{2\pi}{1280}$ ,  $\Delta t = \frac{1}{640}$ . Right:  $\Delta x = \Delta y = \frac{2\pi}{1280}$ ,  $\Delta t = \frac{1}{20}$ .

We choose  $\varepsilon = \frac{1}{128}$  and repeat this test. We choose  $\Delta x = \Delta y = \frac{2\pi}{2560}$ , and compare numerical solutions by fine time step  $\Delta t = \frac{1}{1280}$  and by coarse time step  $\Delta t = \frac{1}{20}$  at  $T = 0.2$  and  $T = 0.4$ , respectively. The level curves of mass density in each case are plotted in Figure 6.10.

In Figures 6.9 and 6.10, the pictures on the left column show particle density when  $\varepsilon$ -dependent, fully resolved time step is taken. The pictures on the right side are plots of particle density when  $\varepsilon$ -independent time step is taken. One can see the good agreements in results from the two sets of tests. This numerical example shows that the SL-TS method can successfully capture correct physical observables when  $\varepsilon$ -independent time step is taken in the multidimensional cases.

## 7 Conclusion

In this paper, we introduced and studied a semi-Lagrangian time splitting scheme for the Schrödinger equation in the presence of electromagnetic field when the semi-classical parameter  $\varepsilon \ll 1$ . Numer-

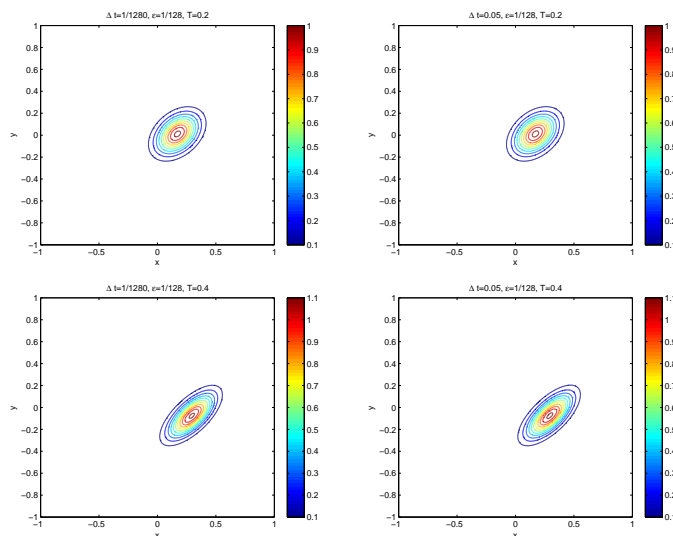


Figure 6.10: Level curves of mass density for  $\varepsilon = \frac{1}{128}$  at  $T = 0.2, 0.4$ . Left:  $\Delta x = \Delta y = \frac{2\pi}{2560}$ ,  $\Delta t = \frac{1}{1280}$ . Right:  $\Delta x = \Delta y = \frac{2\pi}{256}$ ,  $\Delta t = \frac{1}{20}$ .

ically, the burden from  $O(\varepsilon)$  oscillations both in space and time have been reduced by the spectral approximation and the semi-Lagrangian method with polynomial interpolation. This method is easy to implement, can be extended to multidimensional cases and has better stability constraints. We proved the unconditional stability properties when the vector potential is spatially variant, which allows one to take  $\Delta t \gg \varepsilon$ . We established the error estimate for the SL-TS method in the  $L^2$  approximation of the wave function, and derived the corresponding meshing strategy. It was also shown that the meshing strategy can be much relaxed, namely  $\Delta t = o(1)$ ,  $\Delta x = O(\varepsilon)$ , if only physical observables are needed, with the help of the Wigner transform and its classical limit. Many numerical tests have confirmed our numerical analysis results.

In the future, we will explore the extension of this method to the Schrödinger equation with fast varying periodic potentials.

## References

- [1] Y. Aharonov and D. Bohm, Significance of electromagnetic potentials in the quantum theory, Phys. Rev. 115, 485-491 (1959).
- [2] J. Avron, I. Herbst, B. Simon, Schrödinger operators with magnetic fields. I. General interactions, Duke Math. J. Volume 45, Number 4, 847-883 (1978).
- [3] G. Bal, A. Fannjiang, G. Papanicolaou and L. Ryzhik, Radiative transport in a periodic structure, J. Stat. Phys. 95, 479-494 (1999).

- [4] W. Bao and S. Jin, Weakly compressible high-order I-stable central difference schemes for incompressible viscous flows, *Computer methods in applied mechanics and engineering* 190.37, 5009-5026 (2001).
- [5] W. Bao, S. Jin, and P. A. Markowich, On Time-Splitting Spectral Approximations for the Schrödinger Equation in the Semi-classical Regime, *Journal of Computational Physics* 175, 487524 (2002).
- [6] J. R. Bates, A. McDonald, Multiply-Upstream, Semi-Lagrangian Advective Schemes: Analysis and Application to a Multi-Level Primitive Equation Model. *Mon. Wea. Rev.*, 110, 18311842 (1982).
- [7] Y. Brenier, E. Grenier, Sticky particles and scalar conservation laws, *SIAM J. Num. Anal.* 38, 2317-2328 (1998).
- [8] M. Falcone, R. Ferretti, Convergence analysis for a class of high-order semi-Lagrangian advection schemes, *SIAM J. Numer. Anal.*, 35(3), 909940 (1998).
- [9] R. Feynman, R. Leighton, M. Sands, S. Treiman, *The Feynman lectures on physics*, Physics Today (1964).
- [10] I. Gasser and P. A. Markowich, Quantum hydrodynamics, Wigner transforms and the classical limit, *Asymptotic Anal.* 14, 97-116 (1997).
- [11] P. Gérard, P. A. Markowich, N. J. Mauser, and F. Poupaud, Homogenization limits and Wigner transforms, *Comm. Pure Appl. Math.* 50, 323-379 (1997).
- [12] L. Gosse and N. J. Mauser, Multiphase semiclassical approximation of an electron in a one-dimensional crystalline lattice. III. From ab initio models to WKB for Schrödinger- Poisson, *J. Comput. Phys.*, 211, 326-346 (2006).
- [13] D. Gottlieb and S. A. Orszag, *Numerical Analysis of Spectral Methods*, Soc. for Industr. & Appl. Math., Philadelphia (1977).
- [14] S. Gottlieb, C. W. Shu, Total variation diminishing Runge-Kutta schemes, *Math. Comput.*, 67, 73-85 (1998).
- [15] S. Jin, P. Markowich, C. Sparber, Mathematical and computational methods for semiclassical Schrodinger equations, *Acta Numerica*, Volume 20, 121-209 (2011).
- [16] W. Kohn. Theory of Bloch electrons in a magnetic field: the effective Hamiltonian, *Phys. Rev.* 115, 14601478 (1959).
- [17] P. D. Lax, L. Nirenberg, On stability for difference schemes; a sharp form of Gårding's inequality, *Communications on Pure and Applied Mathematics* Volume 19, Issue 4, 473492, (1966).

- [18] S. J. Lin, R. B. Rood, Multidimensional flux-form semi-Lagrangian transport schemes, *Monthly Weather Review*, Vol. 124, No. 9., 2046-2070 (1996).
- [19] P. A. Markowich, N. J. Mauser, and F. Poupaud, A Wigner function approach to semi-classical limits: Electrons in a periodic potential, *J. Math. Phys.* 35, 1066-1094 (1994).
- [20] P. A. Markowich, P. Pietra, and C. Pohl, Numerical approximation of quadratic observables of Schrödinger equations in the semi-classical limit, *Numer. Math.* 81, 595-630 (1999).
- [21] G. Panati, H. Spohn and S. Teufel. Space-adiabatic perturbation theory, *Adv. Theor. Math. Phys.* Volume 7, Number 1, 145-204 (2003).
- [22] J. E. Pasciak, Spectral and pseudo-spectral methods for advection equations, *Math. Comp.* 35, 1081-1092 (1980).
- [23] T. Paul, P. L. Lions, Sur les mesures de Wigner, *Rev. Mat. Iberoamericana* 9, 553618 (1993) .
- [24] J. M. Qiu, A. Christlieb, A conservative high order semi-Lagrangian WENO method for the Vlasov equation, *Journal of Computational Physics*, Volume 229, Issue 4, Pages 11301149 (2010)
- [25] J. Sakurai, S. Tuan, E. Commins, *Modern Quantum Mechanics*, American Journal of Physics (1995).
- [26] M. O. Scully, M. S. Zubairy, *Quantum optics*, Cambridge University Press (1997).
- [27] E. Sonnendrücker, J. Roche, P. Bertrand, The Semi-Lagrangian Method for the Numerical Resolution of the Vlasov Equation, *Journal of Computational Physics*, Volume 149, Issue 2, 201220 (1999).
- [28] A. Staniforth, J. Côté, Semi-Lagrangian integration schemes for atmospheric models-a review, *Monthly Weather Review*, Volume 119, Issue 9, 119, 22062223 (1991).
- [29] E. Süli, A. Ware, A spectral method of characteristics for hyperbolic problems, *SIAM journal on numerical analysis*, 28(2), 423-445 (1991).
- [30] L. Tartar, H-measures: A new approach for studying homogenization, oscillations and concentration effects in partial differential equations, *Proc. Roy. Soc. Edinburgh Sect. A* 115, 193-230 (1990).
- [31] S. Teufel and G. Panati, Propagation of Wigner functions for the Schrödinger equation with a perturbed periodic potential, "Multiscale methods in Quantum Mechanics", Birkhauser, Boston, 207-220 (2004).
- [32] L. N. Trefethen, *Spectral Methods in MATLAB*, SIAM, Philadelphia, 2000.