



**HAL**  
open science

# A well-balanced and asymptotic-preserving scheme for the one-dimensional linear Dirac equation

Laurent Gosse

► **To cite this version:**

Laurent Gosse. A well-balanced and asymptotic-preserving scheme for the one-dimensional linear Dirac equation. BIT Numerical Mathematics, 2014, pp.??-??. 10.1007/s10543-014-0510-4 . hal-00963984

**HAL Id: hal-00963984**

**<https://hal.science/hal-00963984>**

Submitted on 23 Mar 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# A well-balanced and asymptotic-preserving scheme for the one-dimensional linear Dirac equation

Laurent Gosse

Received: date / Accepted: date

**Abstract** The numerical approximation of one-dimensional relativistic Dirac wave equations is considered within the recent framework consisting in deriving local scattering matrices at each interface of the uniform Cartesian computational grid. For a Courant number equal to unity, it is rigorously shown that such a discretization preserves exactly the  $L^2$  norm despite being explicit in time. This construction is well-suited for particles for which the reference velocity is of the order of  $c$ , the speed of light. Moreover, when  $c$  diverges, that is to say, for slow particles (the characteristic scale of the motion is non-relativistic), Dirac equations are naturally written so as to let a “diffusive limit” emerge numerically, like for discrete 2-velocity kinetic models. It is shown that an asymptotic-preserving scheme can be deduced from the aforementioned well-balanced one, with the following properties: it yields unconditionally a classical Schrödinger equation for free particles, but it handles the more intricate case with an external potential only conditionally (the grid should be such that  $c\Delta x \rightarrow 0$ ). Such a stringent restriction on the computational grid can be circumvented easily in order to derive a seemingly original Schrödinger scheme still containing tiny relativistic features. Numerical tests (on both linear and nonlinear equations) are displayed.

**Keywords** Dirac equation · One-dimensional relativistic quantum mechanics · Asymptotic-preserving and well-balanced numerical methods

**Mathematics Subject Classification (2000)** MSC 35Q41 · 65M06 · 81Q05

## 1 Introduction

The fundamental equation of relativistic quantum mechanics is the Dirac equation which combines the special theory of relativity with quantum effects; see also [36, 42] for an application within the Jackiw-Tetelboim “R=T” model of 1 + 1 general

---

L. Gosse  
Istituto per le Applicazioni del Calcolo, via dei Taurini, 19, 00185 Rome, Italy  
E-mail: l.gosse@ba.iac.cnr.it

relativity. In one space dimension, the dynamics of such spin- $\frac{1}{2}$  particles is described by a 2-component *spinor* wave-function but nevertheless display interesting features.

### 1.1 Scalar and vector potentials in linear 1+1 Dirac systems

In 1D, the Dirac equation for a relativistic particle of mass  $m > 0$  moving in an external potential  $V(x) \in \mathbb{R}^2$  reads, where  $H_0$  stands for the “free Hamiltonian”,

$$i\partial_t\Psi = [H_0 + V(x)]\Psi, \quad H_0 = \left( -ic\sigma_1\partial_x \cdot + \frac{mc^2}{\hbar}\sigma_3 \cdot \right).$$

The vector  $\Psi(t, x) \in \mathbb{C}^2$  is usually called a *Dirac spinor*. By neglecting magnetic fields, the remaining external potential  $V$  is split into:

$$V = V_t + \sigma_1 V_e + \sigma_2 V_p + \sigma_3 V_s,$$

where  $V_e$ , the so-called space-term can always be absorbed by means of a gauge transform. The other components  $V_t$ ,  $V_p$  and  $V_s$  are referred to as the time-potential, scalar and pseudo-scalar terms, respectively. Hereafter, the focus will be drawn onto the case  $V_p \equiv 0$ ,  $V_s, V_t$  are real scalar functions of  $x \in \mathbb{R}$ . To fix ideas, the term  $V_s$  may correspond to a confining potential, like a space-dependent effective mass of an electron [23], whereas  $V_t$ , the so-called Lorentz time-potential cannot have a confining effect because of the Klein paradox [7] (see also [14, 15, 39, 41, 42, 45]). Parameters  $c > 0$  and  $\sigma_i$ ,  $i = 1, 2, 3$ , denote the speed of light and classical Pauli matrices, respectively:

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

### 1.2 Conservation of the $L^2$ -norm and non-relativistic limit

In the context of quantum mechanics applications, the preservation of the  $L^2$  norm is important because it is related to the conservation of the probability of presence for particles. For the linear Dirac equations, the Crank-Nicolson scheme is a popular choice [1, 2] because, by Cayley’s theorem [22, 21], it preserves exactly this  $L^2$  norm: here we show in §2.2 that the well-balanced approach [18] allows to improve on what is explained in [46] by furnishing an original, less costly, *explicit* scheme endowed with the same preservation property. An original MUSCL-type reconstruction [19] is presented, too, in §2.3. This is numerically illustrated on both linear and nonlinear benchmarks: see §4.1 and §4.2, respectively. Usually, such well-balanced discretizations deliver automatically asymptotic-preserving (AP) schemes when the limiting process only retains flux terms, like in [18, 20]: here, by studying the non-relativistic limit of Dirac equations with several potentials [5], we see that this is not true anymore when these zero-order potential terms survive in the limit. Indeed, it is shown in §3.3 and §3.4 that despite the second order finite difference is unconditionally recovered, obtaining the correct contribution of the potential asks for a stringent restriction on the computational grid. The somewhat simpler case of non-relativistic

limit of free particles (including explicit schemes) is studied in §3.3. However, it is possible to circumvent it and the correct asymptotic behavior is achieved. Numerical validations are shown in §4.3. Existing asymptotic-preserving strategies were mostly developed in order to handle the classical limit of both linear and non-linear Schrödinger equations: see [9, 13]. a strong convergence result for free particles is furnished in Appendix A.

## 2 Derivation of $L^2$ -conservative well-balanced approximations

It appears that the presentation advocated in [18], and which relies on the use of local scattering matrices, is closely related to ideas published in other contexts, see [4, 12].

### 2.1 Computation of local scattering matrices

Accordingly, we start with a model slightly more involved than what appears in [46]:

$$\partial_t \Psi + c \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \partial_x \Psi = -\frac{i}{\hbar} \left\{ m(x)c^2 \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + V_t(x) \right\} \Psi. \quad (2.1)$$

Denoting  $\Psi = (\psi_+, \psi_-)$ , one diagonalizes the convective part of (2.1) by means of Riemann invariants, which reduce to diagonal variables by linearity. The resulting problem reads, for  $\hbar = 1$  and  $c^2$  being lumped into the space-dependent mass  $m(x)$ ,

$$\zeta_{\pm} = \frac{1}{2}(\psi_+ \pm \psi_-), \quad \partial_t \zeta_{\pm} \pm c \partial_x \zeta_{\pm} = -i(m(x)\zeta_{\mp} + V_t(x)\zeta_{\pm}), \quad (2.2)$$

In order to derive a set of piecewise-constant numerical approximations  $\zeta_{\pm}^{\Delta x}$ , one deduces from [46] that the right-hand side must be upwinded if the explicit Euler time-integrator is used. Following [18], Chapter 8, the first step for building a well-balanced scheme for (2.2) is to consider the wave equation for which low-order terms are “localized” on a uniform grid with a space-step  $\Delta x$ :

$$\boxed{\partial_t \zeta_{\pm} \pm c \partial_x \zeta_{\pm} = -i\Delta x \sum_{j \in \mathbb{Z}} (m(x)\zeta_{\mp} + V_t(x)\zeta_{\pm}) \delta(x - x_{j-\frac{1}{2}})}, \quad (2.3)$$

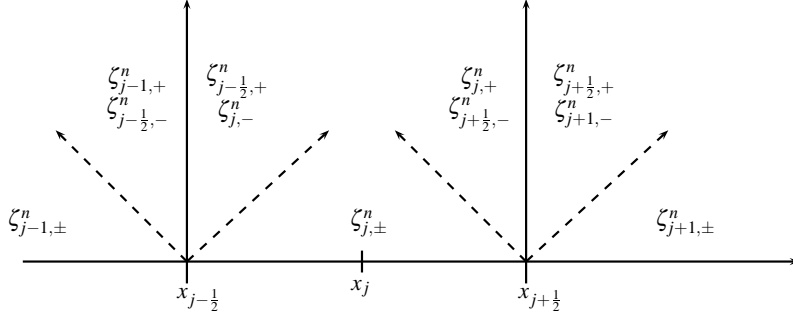
where shorthand notation  $x_j = j\Delta x$  is used for  $j \in \mathbb{Z}$ , together with  $t^n = n\Delta t$ ,  $n \in \mathbb{N}$  and  $\Delta t$  is a time-step constrained by a Courant number  $\nu = \frac{c\Delta t}{\Delta x}$ . Piecewise constant approximations  $\zeta_{\pm}^{\Delta x}$  are defined by values generated by a time-marching process:

$$\zeta_{\pm}^{\Delta x}(t^n, x_j) = \zeta_{j,\pm}^n \in \mathbb{C}.$$

The presence of Dirac masses at each interface  $x_{j-\frac{1}{2}}$  yields the following scheme:

$$\boxed{\zeta_{j,+}^{n+1} = \zeta_{j,+}^n - \frac{c\Delta t}{\Delta x} (\zeta_{j,+}^n - \zeta_{j-\frac{1}{2},+}^n), \quad \zeta_{j,-}^{n+1} = \zeta_{j,-}^n + \frac{c\Delta t}{\Delta x} (\zeta_{j+\frac{1}{2},-}^n - \zeta_{j,-}^n)}. \quad (2.4)$$

Coupling zero-order terms are treated by means of jump relations induced by the



**Fig. 2.1** Schematic view of the scheme (2.4) with Riemann problems for (2.3) and standing waves.

Dirac masses inserted in (2.3). Formal consistency comes from a weak convergence,

$$\Delta x \sum_{j \in \mathbb{Z}} \delta(x - x_{j-\frac{1}{2}}) = \Delta x \sum_{j \in \mathbb{Z}} \delta\left(x - (j - \frac{1}{2})\Delta x\right) \rightarrow 1 \quad \text{as } \Delta x \rightarrow 0.$$

This way, interface values  $\zeta_{j \mp \frac{1}{2}, \pm}^n$  are deduced from  $\zeta_{j \mp 1, \pm}^n$  by means of local scattering matrices (see Fig. 2.1) which expression is given below.

**Lemma 2.1** For any  $j \in \mathbb{Z}$ ,  $\Delta x > 0$ , the smooth solution  $\tilde{\zeta}_{\pm}(x)$  of the boundary-value problem for the stationary equations in  $x \in (0, \Delta x)$ , with  $\tilde{\zeta}_{+}(0) = \tilde{\zeta}_{+}^L$ ,  $\tilde{\zeta}_{-}(\Delta x) = \tilde{\zeta}_{-}^R$ ,

$$\pm c \partial_x \tilde{\zeta}_{\pm} = -i \left( m(x_{j-\frac{1}{2}}) \tilde{\zeta}_{\mp} + V_t(x_{j-\frac{1}{2}}) \tilde{\zeta}_{\pm} \right),$$

satisfies the following relation, with constant values  $V_t = \frac{V_t(x_{j-\frac{1}{2}})}{c}$ ,  $\omega = \frac{m(x_{j-\frac{1}{2}})}{c}$ :

$$\begin{pmatrix} \tilde{\zeta}_{+}(\Delta x) \\ \tilde{\zeta}_{-}(0) \end{pmatrix} = \begin{pmatrix} \frac{\exp(i\Delta x V_t)}{\cosh(\Delta x \omega)} & -i \tanh(\Delta x \omega) \\ -i \tanh(\Delta x \omega) & \frac{\exp(-i\Delta x V_t)}{\cosh(\Delta x \omega)} \end{pmatrix} \begin{pmatrix} \tilde{\zeta}_{+}(0) \\ \tilde{\zeta}_{-}(\Delta x) \end{pmatrix}. \quad (2.5)$$

The  $2 \times 2$  matrix in (2.5) is the *scattering matrix* for (2.4): it relates the *outgoing states*  $\tilde{\zeta}_{\mp}(0/\Delta x)$  to the *incoming ones*  $\tilde{\zeta}_{\pm}(0/\Delta x)$ . These states are said to be incoming because they are prescribed on *incoming characteristics* pointing inside the domain. One may factorize this matrix with  $\text{sech}(\omega \Delta x) = \frac{1}{\cosh(\omega \Delta x)}$ , the hyperbolic secant:

$$S(x) = \text{sech}(\omega \Delta x) \begin{pmatrix} \exp(i\Delta x V_t) & -i \sinh(\Delta x \omega) \\ -i \sinh(\Delta x \omega) & \exp(-i\Delta x V_t) \end{pmatrix}, \quad S_{j \pm \frac{1}{2}} \stackrel{\text{def}}{=} S(x_{j \pm \frac{1}{2}}).$$

*Proof:* The computation is standard because the stationary equations are linear, only the forward/backward boundary data make it differ from a usual ODE system:

$$c \frac{d}{dx} \begin{pmatrix} \tilde{\zeta}_{+} \\ \tilde{\zeta}_{-} \end{pmatrix} = -i \begin{pmatrix} V_t & \omega \\ -\omega & -V_t \end{pmatrix} \begin{pmatrix} \tilde{\zeta}_{+} \\ \tilde{\zeta}_{-} \end{pmatrix}.$$

By means of a diagonalization process and a logarithmic integration, one finds:

$$\begin{aligned}\tilde{\zeta}_+(\Delta x) &= \exp(i\Delta x V_t) \left[ \tilde{\zeta}_+(0) \cosh(\omega \Delta x) - i \tilde{\zeta}_-(0) \sinh(\omega \Delta x) \right], \\ \tilde{\zeta}_-(\Delta x) &= \exp(i\Delta x V_t) \left[ i \tilde{\zeta}_+(0) \sinh(\omega \Delta x) + \tilde{\zeta}_-(0) \cosh(\omega \Delta x) \right].\end{aligned}\quad (2.6)$$

In the second equation, the unknown is  $\tilde{\zeta}_-(0)$ , hence the system becomes:

$$\begin{aligned}\tilde{\zeta}_-(0) &= \frac{\exp(-V_t \Delta x)}{\cosh(\omega \Delta x)} \tilde{\zeta}_-(\Delta x) - i \tanh(\omega \Delta x) \tilde{\zeta}_+(0), \\ \tilde{\zeta}_+(\Delta x) &= \exp(i\Delta x V_t) \tilde{\zeta}_+(0) \left[ \cosh(\omega \Delta x) - \frac{\sinh^2(\omega \Delta x)}{\cosh(\omega \Delta x)} \right] - i \tilde{\zeta}_-(\Delta x) \tanh(\omega \Delta x).\end{aligned}$$

By rearranging terms, one finds easily (2.5).  $\square$

## 2.2 Hyperbolic rotation, $L^2$ -estimate and convergence

In the special case for which  $V_t \equiv 0$ , one observes that the relations (2.6) imply that  $\tilde{\zeta}_\pm(\Delta x)$  (right states) result from the action of a *hyperbolic rotation* onto  $\tilde{\zeta}_\pm(0)$  (left states). Planar hyperbolic rotations are sometimes called ‘‘Lorentz transformations’’.

Going back to the general case  $V_t \neq 0$ , the scheme (2.4) rewrites:

$$\begin{cases} \zeta_{j,+}^{n+1} = (1 - \nu) \zeta_{j,+}^n + \nu \left( \frac{\exp(i\Delta x V_t(x_{j-\frac{1}{2}}))}{\cosh(\Delta x \omega_{j-\frac{1}{2}})} \zeta_{j-1,+}^n - i \tanh(\Delta x \omega_{j-\frac{1}{2}}) \zeta_{j,-}^n \right) \\ \zeta_{j,-}^{n+1} = (1 - \nu) \zeta_{j,-}^n + \nu \left( \frac{\exp(-i\Delta x V_t(x_{j+\frac{1}{2}}))}{\cosh(\Delta x \omega_{j+\frac{1}{2}})} \zeta_{j+1,-}^n - i \tanh(\Delta x \omega_{j+\frac{1}{2}}) \zeta_{j,+}^n \right) \end{cases} \quad (2.7)$$

Let's hereafter denote  $S_{j-\frac{1}{2}}$  the  $2 \times 2$  scattering matrix appearing in (2.4): in the particular case where the Courant number  $\nu = 1$ , which means that the computational grid is chosen so that  $c\Delta t = \Delta x$  (like in [46]),  $\zeta_{j,\pm}^{n+1}$  are the scattering values from the data of  $\zeta_{j\mp 1,\pm}^n$ . Thus a simple criterion for the preservation of the  $L^2$  norm reads:

**Lemma 2.2** *Assume  $V_t(x) \in L^\infty \cap L^2(\mathbb{R})$  is a real function and  $c\Delta t = \Delta x$ , there holds:*

$$\forall n \in \mathbb{N}, \quad \|\zeta^{\Delta x}(t^n, \cdot)\|_{L^2(\mathbb{R})} = \|\zeta^{\Delta x}(0, \cdot)\|_{L^2(\mathbb{R})}.$$

*Remark 2.1* The preservation of the  $L^2$  norm with an explicit scheme like (2.7) is **not** straightforward: indeed, stable upwind discretizations with Courant numbers  $\nu < 1$  always contain a certain amount of *numerical viscosity* which, in the nonlinear case, dissipate enough *entropy* to prevent oscillation onset in the numerical solution. Mathematical entropies being just convex functions, in the context of quantum mechanics applications which ask for conservation of the probability of presence, the scheme must moreover *preserve* (and not dissipate) a particular entropy, the  $L^2$ -norm.

*Proof:* The property is equivalent to stating that for any  $j \in \mathbb{Z}$ , the scattering matrix  $S_{j-\frac{1}{2}}$  is unitary. By multiplying it by the complex conjugate of its transpose, it is straightforward to check that one recovers the identity matrix of  $\mathbb{R}^2$ . Hence:

$$\left( \zeta_{j,+}^{n+1} \zeta_{j-1,-}^{n+1} \right)^* \begin{pmatrix} \zeta_{j,+}^{n+1} \\ \zeta_{j-1,-}^{n+1} \end{pmatrix} = \left( \zeta_{j,+}^n \zeta_{j-1,-}^n \right)^* \left( S_{j-\frac{1}{2}}^* \right)^T S_{j-\frac{1}{2}} \begin{pmatrix} \zeta_{j,+}^n \\ \zeta_{j-1,-}^n \end{pmatrix},$$

which directly yields  $|\zeta_{j,+}^{n+1}|^2 + |\zeta_{j-1,-}^{n+1}|^2 = |\zeta_{j,+}^n|^2 + |\zeta_{j-1,-}^n|^2$ . To derive the uniform estimate, it remains to multiply by  $\Delta x$  and to sum on  $j \in \mathbb{Z}$ .  $\square$

Thanks to the linearity of the system (2.1), the aforementioned conservation of the  $L^2$  norm, which yields weak compactness in  $L^2$ , is enough to ensure the weak convergence of a subsequence of approximate solutions  $\zeta_{\pm}^{\Delta x}$  toward the exact ones [43] for smooth  $V_i$ . Moreover, by uniqueness of the limit, all the sequence converges. More accurate bounds are established in §A where strong  $L^2_{loc}$  convergence is shown.

### 2.3 A MUSCL-type reconstruction compatible with the S-matrix

The  $L^2$ -preservation property rigorously holds only for a Courant number  $\nu = 1$ , thus one may want to make the scheme (2.7) less vulnerable to the damage of numerical dissipation. One standard way to proceed is to set up the so-called MUSCL-reconstruction, relying on a conveniently limited piecewise-linear extrapolation of the cell-average values  $\zeta_{j,\pm}^n$ . One issue is that such piecewise-linear data ask now

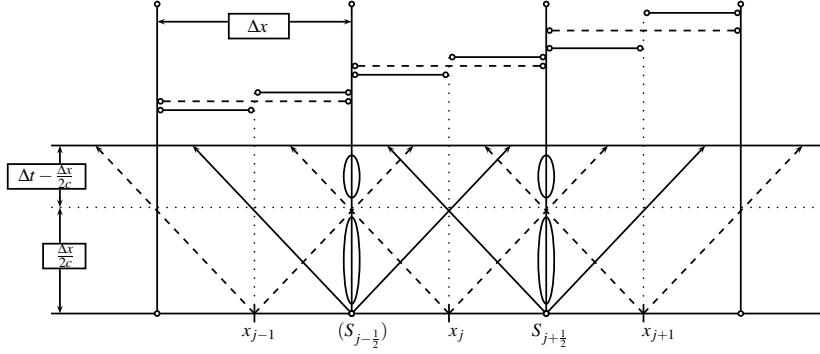


Fig. 2.2 Schematic view of the MUSCL reconstruction and scattering matrices

for the resolution of “Generalized Riemann Problems” (GRP) at each interface  $x_{j\pm\frac{1}{2}}$ , see [6,44]: even in cases where it is doable, it doesn’t fit with treating the right-hand side by means of S-matrices. However, a new type of MUSCL scheme was devised in [19] and, along with heavy notation, it proceeds as follows:

1. **Reconstruction:** given the sequence of spinors  $(\zeta_{j,\pm}^n)_{j \in \mathbb{Z}} \in \mathbb{C}^2$ , one first performs a linear interpolation. At both borders of each cell  $C_j = (x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}})$ , the following values are derived by means of local slopes  $\sigma_j^n$  and a slope-limiter  $\phi : \mathbb{R} \rightarrow [0, 2]$ ,

$$\zeta_{j-\frac{1}{2},\pm}^{n,R} = \zeta_{j,\pm}^n - \frac{\sigma_j^n \Delta x}{2}, \quad \zeta_{j+\frac{1}{2},\pm}^{n,L} = \zeta_{j,\pm}^n + \frac{\sigma_j^n \Delta x}{2}.$$

More precisely, mass conservation is ensured with the choice:

$$\begin{aligned} \zeta_{j-\frac{1}{2},\pm}^{n,R} &= \Re(\zeta_{j,\pm}^n) - \frac{\phi(r_j^n)}{2} \Re(\zeta_{j+1,\pm}^n - \zeta_{j,\pm}^n) \\ &\quad + i\Im(\zeta_{j,\pm}^n) - i\frac{\phi(r_j^n)}{2} \Im(\zeta_{j+1,\pm}^n - \zeta_{j,\pm}^n), \end{aligned} \quad \phi(r) = \frac{r+|r|}{1+|r|},$$

and the ratios to which slope-limiters are applied are,

$$r_j^n = \frac{\Re(\zeta_{j+1,\pm}^n - \zeta_{j,\pm}^n)}{\Re(\zeta_{j,\pm}^n - \zeta_{j-1,\pm}^n)}, \quad \tilde{r}_j^n = \frac{\Im(\zeta_{j+1,\pm}^n - \zeta_{j,\pm}^n)}{\Im(\zeta_{j,\pm}^n - \zeta_{j-1,\pm}^n)}.$$

Finally, the reconstructed interface states read, for each cell  $C_j$ :

$$\begin{cases} \zeta_{j-\frac{1}{2},\pm}^{n,R} = \zeta_{j,\pm}^n - \frac{\phi(r_j^n)}{2} \Re(\zeta_{j+1,\pm}^n - \zeta_{j,\pm}^n) - i\frac{\phi(r_j^n)}{2} \Im(\zeta_{j+1,\pm}^n - \zeta_{j,\pm}^n), \\ \zeta_{j+\frac{1}{2},\pm}^{n,L} = \zeta_{j,\pm}^n + \frac{\phi(r_j^n)}{2} \Re(\zeta_{j+1,\pm}^n - \zeta_{j,\pm}^n) + i\frac{\phi(r_j^n)}{2} \Im(\zeta_{j+1,\pm}^n - \zeta_{j,\pm}^n). \end{cases}$$

2. Fluxes computation: Yet we have discontinuities at both the center  $x_j$  and interfaces  $x_{j\pm\frac{1}{2}}$ : assuming the right-hand side of (2.2) is again localized only at interfaces, the formulation (2.3) still holds. Thus the discontinuities at the center  $x_j$  are simply advected until they reach the interfaces  $x_{j\pm\frac{1}{2}}$ , and there they are scattered by the S-matrix: see Fig. 2.2. Observe that (2.7) can be put in a matrix form as:

$$\begin{pmatrix} \zeta_{j,+}^{n+1} \\ \zeta_{j-1,-}^{n+1} \end{pmatrix} = (1-\nu) \begin{pmatrix} \zeta_{j,+}^n \\ \zeta_{j-1,-}^n \end{pmatrix} + \nu S_{j-\frac{1}{2}} \begin{pmatrix} \zeta_{j-1,+}^n \\ \zeta_{j,-}^n \end{pmatrix}, \quad \nu = \frac{c\Delta t}{\Delta x}. \quad (2.8)$$

So, by taking into account the (new) discontinuity at the cell's center,

$$\begin{pmatrix} \zeta_{j,+}^{n+1} \\ \zeta_{j-1,-}^{n+1} \end{pmatrix} = \begin{pmatrix} \zeta_{j,+}^n \\ \zeta_{j-1,-}^n \end{pmatrix} - \min(\nu, \frac{1}{2}) \left\{ \begin{pmatrix} \zeta_{j+\frac{1}{2},+}^{n,L} \\ \zeta_{j-\frac{1}{2},-}^{n,R} \end{pmatrix} - S_{j-\frac{1}{2}} \begin{pmatrix} \zeta_{j-\frac{1}{2},+}^{n,L} \\ \zeta_{j+\frac{1}{2},-}^{n,R} \end{pmatrix} \right\} \\ + \max(0, \nu - \frac{1}{2}) \left\{ \begin{pmatrix} \zeta_{j+\frac{1}{2},+}^{n,R} \\ \zeta_{j-\frac{1}{2},-}^{n,L} \end{pmatrix} - S_{j-\frac{1}{2}} \begin{pmatrix} \zeta_{j-\frac{1}{2},+}^{n,R} \\ \zeta_{j+\frac{1}{2},-}^{n,L} \end{pmatrix} \right\}.$$

3. Overall algorithm: Hence, by linearity, one can form the following 2 states,

$$\begin{cases} \tilde{\zeta}_{j+\frac{1}{2},+}^n = \min(\nu, \frac{1}{2}) \zeta_{j+\frac{1}{2},+}^{n,L} - \max(0, \nu - \frac{1}{2}) \zeta_{j+\frac{1}{2},+}^{n,R}, \\ \tilde{\zeta}_{j+\frac{1}{2},-}^n = \min(\nu, \frac{1}{2}) \zeta_{j+\frac{1}{2},-}^{n,R} - \max(0, \nu - \frac{1}{2}) \zeta_{j+\frac{1}{2},-}^{n,L}, \end{cases} \quad (2.9)$$

and the MUSCL scheme rewrites in a slightly less involved form:

$$\begin{pmatrix} \zeta_{j,+}^{n+1} \\ \zeta_{j-1,-}^{n+1} \end{pmatrix} = \begin{pmatrix} \zeta_{j,+}^n \\ \zeta_{j-1,-}^n \end{pmatrix} - \left\{ \begin{pmatrix} \tilde{\zeta}_{j+\frac{1}{2},+}^n \\ \tilde{\zeta}_{j-\frac{1}{2},-}^n \end{pmatrix} - S_{j-\frac{1}{2}} \begin{pmatrix} \tilde{\zeta}_{j-\frac{1}{2},+}^n \\ \tilde{\zeta}_{j+\frac{1}{2},-}^n \end{pmatrix} \right\}. \quad (2.10)$$

The scheme (2.9)–(2.10) will be tested in §4.2.2.



### 3 Asymptotic-Preserving property for non-relativistic limit

The numerical time-marching scheme (2.7) is well adapted to the case where  $c$  is of the order of the characteristic scale of the computational grid, that is to say, for relativistic particles. In case the particles (usually represented with wavepackets endowed with a particular *group velocity*) move much slower, the aforementioned computational grids, for which  $c\Delta t \simeq \Delta x$ , don't really fit thus one is interested to explore the behavior of the numerical process (2.7) in the limit  $c \rightarrow +\infty$ .

#### 3.1 Formal non-relativistic limit of 1D Dirac equation

Hereafter we consider the particular case of (2.1) with a mass  $m > 0$  independent of  $x$ . In order to study the non-relativistic limit of a Dirac particle with (positive) energy  $E \simeq mc^2$ , one retains only the slower time-variation like  $\exp(-i(E - mc^2)t/\hbar)$  by considering the rescaled quantities,  $\tilde{\psi}_{\pm}(t, x) = \exp(imc^2t/\hbar)\psi_{\pm}(t, x)$ , which satisfy:

$$\partial_t \tilde{\psi}_+ + c\partial_x \tilde{\psi}_- = \frac{i}{\hbar}V_t(x)\tilde{\psi}_+, \quad \partial_t \tilde{\psi}_- + c\partial_x \tilde{\psi}_+ = 2i\frac{mc^2}{\hbar}\tilde{\psi}_- + \frac{i}{\hbar}V_t(x)\tilde{\psi}_-.$$

When  $c$  diverges, the second equation is dominated by the rule,  $\tilde{\psi}_- = \frac{\hbar}{2imc}\partial_x \tilde{\psi}_+$ , which, once inserted in the first one, yields the classical Schrödinger equation:

$$i\partial_t \tilde{\psi}_+ + \frac{\hbar^2}{2m}\partial_{xx} \tilde{\psi}_+ = V_t(x)\tilde{\psi}_+. \quad (3.1)$$

The rigorous theory of non-relativistic limits of linear multidimensional Dirac equation with external potential is given in [5, 23].

#### 3.2 Rescaled diagonal variables and local scattering matrix

With a slight abuse of notation, we still denote the Riemann invariants associated with the ‘‘rescaled spinor’’  $\zeta_{\pm} = \frac{1}{2}(\tilde{\psi}_+ \pm \tilde{\psi}_-)$ . A new (small) parameter  $\varepsilon = \frac{1}{c}$  is introduced for ease of reading, and a rescaled system is derived:

$$\partial_t \zeta_{\pm} \pm \frac{1}{\varepsilon}\partial_x \zeta_{\pm} = 2i \left( \pm \frac{m}{\varepsilon^2}(\zeta_+ - \zeta_-) + \frac{V_t(x)}{2}\zeta_{\pm} \right). \quad (3.2)$$

By substituting  $\frac{V_t}{2} \rightarrow V_t$  and following the same roadmap as in Lemma 2.1, one gets:

**Lemma 3.1** *Let  $V_t > 0$ ,  $\varepsilon > 0$  and consider the boundary-value problem on  $x \in (0, \Delta x)$ , associated with the stationary equations of (3.2), with incoming data  $\tilde{\zeta}_+(0)$ ,  $\tilde{\zeta}_-(\Delta x)$ : if one defines the following quantity,*

$$\lambda_{\varepsilon} = \sqrt{(\varepsilon V_t + m/\varepsilon)^2 - (m/\varepsilon)^2} \rightarrow \sqrt{2V_t m}, \quad \varepsilon \rightarrow 0,$$

1. the local scattering matrix reads:

$$S = \frac{1}{D_\varepsilon} \begin{pmatrix} -\frac{im}{\varepsilon\lambda_\varepsilon} \sin(2\lambda_\varepsilon\Delta x) & 1 \\ 1 & -\frac{im}{\varepsilon\lambda_\varepsilon} \sin(2\lambda_\varepsilon\Delta x) \end{pmatrix}. \quad (3.3)$$

where  $D_\varepsilon = \cos(2\lambda_\varepsilon\Delta x) - i\frac{\varepsilon V_t + m/\varepsilon}{\lambda_\varepsilon} \sin(2\lambda_\varepsilon\Delta x)$ .

2. Moreover,  $S$  is unitary:  $(S^*)^T S$  is the  $2 \times 2$  identity matrix.

*Proof:* The proof of (i) proceeds by integrating the following ODE system:

$$\frac{1}{\varepsilon} \frac{d}{dx} \begin{pmatrix} \tilde{\zeta}_+ \\ \tilde{\zeta}_- \end{pmatrix} = 2i \begin{pmatrix} \frac{m}{\varepsilon^2} + V_t & -\frac{m}{\varepsilon^2} \\ \frac{m}{\varepsilon^2} & -(\frac{m}{\varepsilon^2} + V_t) \end{pmatrix} \begin{pmatrix} \tilde{\zeta}_+ \\ \tilde{\zeta}_- \end{pmatrix}, \quad x \in (0, \Delta x).$$

By means of a standard diagonalization procedure, one finds that  $\pm\lambda_\varepsilon$  are corresponding eigenvalues and a tedious logarithmic integration delivers:

$$\begin{pmatrix} \tilde{\zeta}_+(\Delta x) \\ \tilde{\zeta}_-(\Delta x) \end{pmatrix} = \begin{pmatrix} \cos(2\lambda_\varepsilon\Delta x) + i\frac{\varepsilon V_t + m/\varepsilon}{\lambda_\varepsilon} \sin(2\lambda_\varepsilon\Delta x) & -\frac{im}{\varepsilon\lambda_\varepsilon} \sin(2\lambda_\varepsilon\Delta x) \\ \frac{im}{\varepsilon\lambda_\varepsilon} \sin(2\lambda_\varepsilon\Delta x) & \cos(2\lambda_\varepsilon\Delta x) - i\frac{\varepsilon V_t + m/\varepsilon}{\lambda_\varepsilon} \sin(2\lambda_\varepsilon\Delta x) \end{pmatrix} \begin{pmatrix} \tilde{\zeta}_+(0) \\ \tilde{\zeta}_-(0) \end{pmatrix}.$$

Since  $\tilde{\zeta}_-(\Delta x)$  belongs to the incoming boundary data, the second line must be inverted in order to derive the expression of the outgoing state  $\tilde{\zeta}_-(0)$ :

$$\tilde{\zeta}_-(0) = \frac{\tilde{\zeta}_-(\Delta x) - \frac{im}{\varepsilon\lambda_\varepsilon} \sin(2\lambda_\varepsilon\Delta x) \tilde{\zeta}_+(0)}{\cos(2\lambda_\varepsilon\Delta x) - i\frac{\varepsilon V_t + m/\varepsilon}{\lambda_\varepsilon} \sin(2\lambda_\varepsilon\Delta x)}.$$

When plugging  $\tilde{\zeta}_-(0)$  into the expression of  $\tilde{\zeta}_+(\Delta x)$ , a simplification occurs:

$$\tilde{\zeta}_+(\Delta x) = \frac{\overbrace{\cos^2(2\lambda_\varepsilon\Delta x) + \left( \frac{(\varepsilon V_t + m/\varepsilon)^2}{\lambda_\varepsilon^2} - \frac{(m/\varepsilon)^2}{\lambda_\varepsilon^2} \right) \sin^2(2\lambda_\varepsilon\Delta x)}^{=1}}{\cos(2\lambda_\varepsilon\Delta x) - i\frac{\varepsilon V_t + m/\varepsilon}{\lambda_\varepsilon} \sin(2\lambda_\varepsilon\Delta x)} \tilde{\zeta}_+(0) - \frac{\frac{im}{\varepsilon\lambda_\varepsilon} \sin(2\lambda_\varepsilon\Delta x)}{\cos(2\lambda_\varepsilon\Delta x) - i\frac{\varepsilon V_t + m/\varepsilon}{\lambda_\varepsilon} \sin(2\lambda_\varepsilon\Delta x)} \tilde{\zeta}_-(\Delta x).$$

It is now straightforward to derive the expression of the scattering matrix. Now, for proving (ii), one just computes  $(S^*)^T S$  and simplifies the expression of  $D_\varepsilon \cdot D_\varepsilon^*$ :

$$(S^*)^T S = \frac{1}{\cos^2(2\lambda_\varepsilon\Delta x) + \underbrace{\frac{(\varepsilon V_t + m/\varepsilon)^2}{\lambda_\varepsilon^2} - \frac{(m/\varepsilon)^2}{\lambda_\varepsilon^2}}_{=1 + \frac{(m/\varepsilon)^2}{\lambda_\varepsilon^2}} \sin^2(2\lambda_\varepsilon\Delta x)} \times \begin{pmatrix} 1 + \frac{(m/\varepsilon)^2}{\lambda_\varepsilon^2} \sin^2(2\lambda_\varepsilon\Delta x) & 0 \\ 0 & 1 + \frac{(m/\varepsilon)^2}{\lambda_\varepsilon^2} \sin^2(2\lambda_\varepsilon\Delta x) \end{pmatrix}.$$

With this expression, one deduces easily that  $(S^*)^T S$  is the  $2 \times 2$  identity matrix.  $\square$

### 3.3 The special case of free particles

In the former calculation, it appears that a source of complication is the interaction of both source terms, namely the particle's mass  $m$  and the time-potential  $V_t$ . However, one may restrict attention to the simpler case of non-relativistic limit of free particles, for which  $V_t \equiv 0$ . Clearly, this assumption implies that  $\lambda_\varepsilon = 0$ , hence  $\cos(2\lambda_\varepsilon \Delta x) = 1$  and  $\frac{\varepsilon V_t + m/\varepsilon}{\lambda_\varepsilon} \sin(2\lambda_\varepsilon \Delta x) = \frac{2\Delta x m}{\varepsilon}$ . The (constant) scattering matrix  $S$  simplifies into:

$$S = \frac{1}{1 - 2im\Delta x/\varepsilon} \begin{pmatrix} 1 & -2im\Delta x/\varepsilon \\ -2im\Delta x/\varepsilon & 1 \end{pmatrix}.$$

#### 3.3.1 Asymptotic-Preserving implicit schemes

One can follow the early computations in [20] dealing with the diffusive regime: it consists in observing that the outgoing states rewrite in the equivalent form,

$$\begin{cases} \tilde{\zeta}_+(\Delta x) = \tilde{\zeta}_-(\Delta x) + \frac{1}{1-2i\Delta x m/\varepsilon} (\tilde{\zeta}_+(0) - \tilde{\zeta}_-(\Delta x)), \\ \tilde{\zeta}_-(0) = \tilde{\zeta}_+(0) + \frac{1}{1-2i\Delta x m/\varepsilon} (\tilde{\zeta}_-(\Delta x) - \tilde{\zeta}_+(0)). \end{cases}$$

Inserting these scattering states into the expression of the scheme (2.4) yields:

$$\boxed{\begin{cases} \zeta_{j,+}^{n+1} = \zeta_{j,+}^n - \frac{\Delta t}{\varepsilon \Delta x} (\zeta_{j,+}^{n+1} - \zeta_{j,-}^{n+1}) + \frac{\Delta t}{\varepsilon \Delta x - 2i\Delta x^2 m} (\zeta_{j-1,+}^{n+1} - \zeta_{j,-}^{n+1}), \\ \zeta_{j,-}^{n+1} = \zeta_{j,-}^n + \frac{\Delta t}{\varepsilon \Delta x} (\zeta_{j,+}^{n+1} - \zeta_{j,-}^{n+1}) + \frac{\Delta t}{\varepsilon \Delta x - 2i\Delta x^2 m} (\zeta_{j+1,-}^{n+1} - \zeta_{j,+}^{n+1}), \end{cases}} \quad (3.4)$$

In order to shed light onto the numerical dynamics of  $\tilde{\psi}_{j,+}^n$ , one sums up both lines of this scheme and unconditionally obtains that, in the limit  $\varepsilon \rightarrow 0$ ,

$$i\tilde{\psi}_{j,+}^{n+1} = i\tilde{\psi}_{j,+}^n - \frac{\Delta t}{2\Delta x^2 m} (\tilde{\psi}_{j+1,+}^{n+1} - 2\tilde{\psi}_{j,+}^{n+1} + \tilde{\psi}_{j-1,+}^{n+1}),$$

a consistent finite-difference approximation of the classical Schrödinger equation, treated implicitly because the explicit Euler scheme is unconditionally unstable [10].

*Remark 3.1* Based on the results of [1, 2, 32, 33], it may be a better choice to use a second-order Crank-Nicolson time-integrator (instead of an implicit Euler method) in (3.4) in order to preserve the  $L^2$ -norm of the (piecewise constant) numerical approximation  $\tilde{\psi}_+^{\Delta x}$ . The computations become more involved, though, as one can see:

$$\begin{cases} \zeta_{j,+}^{n+1} + \frac{\Delta t}{2\varepsilon \Delta x} (\zeta_{j,+}^{n+1} - \zeta_{j,-}^{n+1}) - \frac{\Delta t/2}{\varepsilon \Delta x - 2i\Delta x^2 m} (\zeta_{j-1,+}^{n+1} - \zeta_{j,-}^{n+1}) \\ \quad = \zeta_{j,+}^n - \frac{\Delta t}{\varepsilon 2\Delta x} (\zeta_{j,+}^n - \zeta_{j,-}^n) + \frac{\Delta t/2}{\varepsilon \Delta x - 2i\Delta x^2 m} (\zeta_{j-1,+}^n - \zeta_{j,-}^n), \\ \zeta_{j,-}^{n+1} - \frac{\Delta t}{2\varepsilon \Delta x} (\zeta_{j,+}^{n+1} - \zeta_{j,-}^{n+1}) - \frac{\Delta t/2}{\varepsilon \Delta x - 2i\Delta x^2 m} (\zeta_{j+1,-}^{n+1} - \zeta_{j,+}^{n+1}) \\ \quad = \zeta_{j,-}^n + \frac{\Delta t}{2\varepsilon \Delta x} (\zeta_{j,+}^n - \zeta_{j,-}^n) + \frac{\Delta t/2}{\varepsilon \Delta x - 2i\Delta x^2 m} (\zeta_{j+1,-}^n - \zeta_{j,+}^n). \end{cases} \quad (3.5)$$

### 3.3.2 An explicit and conditionally stable AP scheme ?

The derivation of Asymptotic-Preserving schemes is usually all about the behavior, as a small parameter goes to zero, of discrete derivatives *in the space variable,  $x$* . The discretization in time being kept as the Euler forward method, for instance. In the present context, this raises an issue because applying straightforwardly such an elementary time-integrator leads to unconditional instability: in the former subsection, we thus chose to set up time-implicit schemes for the sake of stability, but a drawback is their CPU cost. Hence one may wonder whether it may be possible to proceed, like in [20], in such a way that a stable time-explicit scheme for the Schrödinger equation can be recovered in the limit  $\varepsilon \rightarrow 0$ . Basically, explicit, or barely implicit (*i.e.* not asking for inversions of linear systems) methods essentially reduce to:

- Leapfrog or Dufort-Frankel methods, which are not self-starting, [3, 33, 47]
- Euler forward method, but perturbed by supplementary dissipation, [10, 34]

Setting up the AP process described in (3.4) within either Leapfrog or Dufort-Frankel time-integration methods appears to produce only numerical instability. However, the second strategy, based on supplementary diffusion terms can lead to stable algorithms, so we present it now, following both [10, 34].

1. the following forward-Euler scheme (compare with (3.4)) is unstable,

$$\begin{cases} \frac{\zeta_{j,+}^{n+1} - \zeta_{j,+}^n}{\Delta t} + \frac{1}{\varepsilon \Delta x} (\zeta_{j,+}^{n+1} - \zeta_{j,-}^{n+1}) = \frac{1}{\varepsilon \Delta x - 2i \Delta x^2 m} (\zeta_{j-1,+}^n - \zeta_{j,-}^n), \\ \frac{\zeta_{j,-}^{n+1} - \zeta_{j,-}^n}{\Delta t} - \frac{1}{\varepsilon \Delta x} (\zeta_{j,+}^{n+1} - \zeta_{j,-}^{n+1}) = \frac{1}{\varepsilon \Delta x - 2i \Delta x^2 m} (\zeta_{j+1,-}^n - \zeta_{j,+}^n), \end{cases} \quad (3.6)$$

at least because, by formally imposing  $\varepsilon = 0$  and summing, an unstable discretization of the free Schrödinger equation arises. However, it has the advantage of being easy to implement since the implicit part can be inverted analytically.

2. one may think that this instability comes exclusively from the change of equation's type in the limit  $\varepsilon \rightarrow 0$ , and so that modifying the time-differentiation like in [10, 34] can fix this issue. Keeping finite-differences in space, one substitutes:

$$\frac{\zeta_{j,\pm}^{n+1} - \zeta_{j,\pm}^n}{\Delta t} \rightsquigarrow \frac{\zeta_{j,\pm}^{n+1} - \zeta_{j,\pm}^n}{\Delta t (1 + \Delta t (\alpha - i\beta))} \quad \text{or} \quad \frac{\zeta_{j,\pm}^{n+1} - \zeta_{j,\pm}^n}{\Delta t + \Delta x^2 (\alpha - i\beta)}.$$

According to [10], the second choice leads to CFL restrictions in  $\Delta t \simeq O(\Delta x^4)$ , which is prohibitively costly, so we restrict ourselves to the first one, less demanding on  $\Delta t$ , even if it can become unstable for  $\Delta t, \Delta x$  becoming very small.

3. within our notation, there is supplementary diffusion if  $\beta > 0$ , but it is needed only for small values of  $\varepsilon$ , so we propose the following modified time-differentiation,

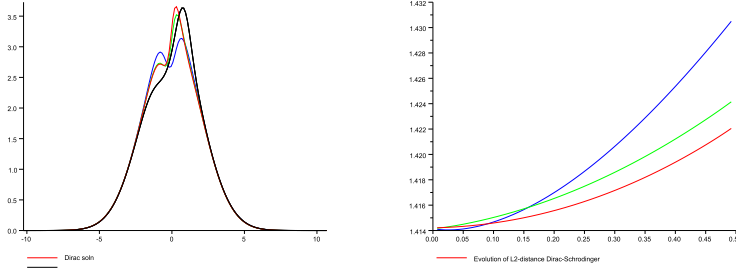
$$\frac{\zeta_{j,\pm}^{n+1} - \zeta_{j,\pm}^n}{\Delta t (1 + \Delta t (\alpha - i\beta) (1 - \varepsilon^\gamma))}, \quad 0 < \gamma \leq 1, \quad (3.7)$$

meaning that diffusion is activated only as  $\varepsilon \ll 1$ . It remains to plug it inside (3.6).

Position densities were computed with the resulting scheme (3.6)–(3.7),  $\beta = 1$  and  $\varepsilon = 1, 0.1, 0.01$ , see Fig. 3.1. Initial data was sampled on a uniform grid of  $2^8$  points,

$$\zeta_{\pm}(t=0, x) = \frac{1}{2} \exp(i \tanh(x) - (x/4)^2), \quad x \in (-16, 16).$$

The data for Schrödinger, treated with the second-order Leapfrog method, is thus  $2\zeta_{\pm}(t=0, \cdot)$ . Numerical position densities become closer as  $\varepsilon$  is decreased, up to  $\varepsilon \simeq 0.001$  where convergence stalls because of diffusion. Accordingly, one cannot



**Fig. 3.1** Position densities generated by (3.6)–(3.7) at time  $t = 0.5$  for  $\varepsilon = 1, 0.1, 0.01$  (left) and time-growth of the  $L^2$  difference between them and Leapfrog approximation for the free Schrödinger equation.

really speak about explicit AP schemes because the scheme obtained in the limit  $\varepsilon \rightarrow 0$  doesn't match exactly the numerical approximation one gets, for instance, out of a Leapfrog discretization: it still contains the supplementary diffusion which is necessary for stability. Hence **we shall not push the study of such conditionally stable explicit Asymptotic-Preserving discretizations any further.**

### 3.4 Difficulties with the complete case

In the general case  $V_t \neq 0$ , the full scattering matrix (3.3) must be considered, and the eigenvalue  $\lambda_\varepsilon$  strongly couples both the source terms  $m/\varepsilon^2$  and  $\varepsilon V_t$ : in order to ease the reading, we shall limit ourselves to the case of a **constant time-potential**  $V_t > 0$  because it already highlights most of the corresponding issues. Thanks to this simplifying assumption, the scattering matrix  $S$  is still independent of the index  $j \in \mathbb{Z}$  (the modifications necessary for handling a space-dependent time-potential are straightforward) and the outgoing states read now:

$$\begin{cases} \tilde{\zeta}_+(\Delta x) = \tilde{\zeta}_-(\Delta x) + \frac{1}{D_\varepsilon} \left( \tilde{\zeta}_+(0) - \left[ \cos(2\lambda_\varepsilon \Delta x) - \frac{i\varepsilon V_t}{\lambda_\varepsilon} \sin(2\lambda_\varepsilon \Delta x) \right] \tilde{\zeta}_-(\Delta x) \right), \\ \tilde{\zeta}_-(0) = \tilde{\zeta}_+(0) + \frac{1}{D_\varepsilon} \left( \tilde{\zeta}_-(\Delta x) - \left[ \cos(2\lambda_\varepsilon \Delta x) - \frac{i\varepsilon V_t}{\lambda_\varepsilon} \sin(2\lambda_\varepsilon \Delta x) \right] \tilde{\zeta}_+(0) \right). \end{cases}$$

Clearly, these expressions split naturally between a first part which generates the second-order divided difference (like in the preceding section), and a second one

which contains the time-potential term. Plugging into the upwind scheme (2.4),

$$\begin{cases} \zeta_{j,+}^{n+1} = \zeta_{j,+}^n - \frac{\Delta t}{\varepsilon \Delta x} (\zeta_{j,+}^{n+1} - \zeta_{j,-}^{n+1}) \\ \quad + \frac{\Delta t}{\varepsilon D_\varepsilon \Delta x} \left( \zeta_{j-1,+}^{n+1} - \cos(2\lambda_\varepsilon \Delta x) \zeta_{j,-}^{n+1} \right) \\ \quad + \frac{i \Delta t}{\varepsilon D_\varepsilon \Delta x} \frac{\varepsilon V_t}{\lambda_\varepsilon} \sin(2\lambda_\varepsilon \Delta x) \zeta_{j,-}^{n+1} \\ \zeta_{j,-}^{n+1} = \zeta_{j,-}^n + \frac{\Delta t}{\varepsilon \Delta x} (\zeta_{j,+}^{n+1} - \zeta_{j,-}^{n+1}) \\ \quad + \frac{\Delta t}{\varepsilon D_\varepsilon \Delta x} \left( \zeta_{j+1,-}^{n+1} - \cos(2\lambda_\varepsilon \Delta x) \zeta_{j,+}^{n+1} \right) \\ \quad + \frac{i \Delta t}{\varepsilon D_\varepsilon \Delta x} \frac{\varepsilon V_t}{\lambda_\varepsilon} \sin(2\lambda_\varepsilon \Delta x) \zeta_{j,+}^{n+1} \end{cases} \quad (3.8)$$

Taking into account that  $\lambda_\varepsilon = O(1)$  uniformly in  $\varepsilon$ , there are 2 singular limits:

1. the term  $\frac{\cos(2\lambda_\varepsilon \Delta x)}{\varepsilon D_\varepsilon \Delta x}$  behaves as follows as  $\varepsilon \rightarrow 0$ ,

$$\frac{\cos(2\lambda_\varepsilon \Delta x)}{\Delta x \left[ \varepsilon \cos(2\lambda_\varepsilon \Delta x) - i \frac{\varepsilon^2 V_t + m}{\lambda_\varepsilon} \sin(2\lambda_\varepsilon \Delta x) \right]} \rightarrow \frac{1}{\Delta x \left[ \varepsilon - 2i \Delta x m \frac{\tan(2\lambda_\varepsilon \Delta x)}{2\lambda_\varepsilon \Delta x} \right]}.$$

Assuming that  $\lambda_\varepsilon \Delta x \rightarrow 0$  as  $\Delta x \rightarrow 0$ , which is immediate thanks to the uniform bound on  $\lambda_\varepsilon$ , one sees that this term is endowed with the correct behavior because

$$\frac{\tan(2\lambda_\varepsilon \Delta x)}{2\lambda_\varepsilon \Delta x} \rightarrow 1, \quad \Delta x \rightarrow 0.$$

2. the term  $\frac{\varepsilon V_t \sin(2\lambda_\varepsilon \Delta x)}{\varepsilon \lambda_\varepsilon D_\varepsilon \Delta x}$  behaves as follows as  $\varepsilon \rightarrow 0$ ,

$$\frac{V_t \sin(2\lambda_\varepsilon \Delta x)}{\left[ \Delta x \lambda_\varepsilon \cos(2\lambda_\varepsilon \Delta x) - i \Delta x \left( \varepsilon V_t + \frac{m}{\varepsilon} \right) \sin(2\lambda_\varepsilon \Delta x) \right]} \rightarrow \frac{2V_t}{\left[ 2\Delta x \lambda_\varepsilon \frac{\cos(2\lambda_\varepsilon \Delta x)}{\sin(2\lambda_\varepsilon \Delta x)} - 2i \frac{\Delta x m}{\varepsilon} \right]}.$$

Now, still assuming that  $\lambda_\varepsilon \Delta x \rightarrow 0$  as  $\Delta x \rightarrow 0$ , one sees that  $\frac{2\Delta x \lambda_\varepsilon}{\tan(2\lambda_\varepsilon \Delta x)} \rightarrow 1$ , as expected. But the imaginary term  $2i \frac{m \Delta x}{\varepsilon}$  doesn't vanish unless  $m \Delta x \leq \varepsilon^{1+\alpha}$ ,  $\alpha > 0$ , which is a stringent restriction on the computational grid. Surprisingly, in order to correctly recover the potential term in the non-relativistic limit, one must impose a restriction on  $\Delta x/\varepsilon$ . This phenomenon seems to be new: it doesn't manifest itself in the diffusive limits previously investigated: see [20, 18].

Fortunately, the issue raised by the imaginary term displaying an incorrect behavior in the limit  $\varepsilon \rightarrow 0$  can be easily circumvented by substituting  $D_\varepsilon$  by  $\Re(D_\varepsilon)$  in the denominator, its real part (which is endowed with the correct behavior as  $\varepsilon \rightarrow 0$  in absence of restrictions on  $\Delta x$ ). Another fix may be the following one:

$$D_\varepsilon \rightsquigarrow \tilde{D}_\varepsilon = \cos(2\lambda_\varepsilon \Delta x) - i \frac{\varepsilon V_t + \max(1, \frac{1}{\varepsilon}) m}{\lambda_\varepsilon} \sin(2\lambda_\varepsilon \Delta x).$$

A simple asymptotic-preserving scheme in the non-relativistic limit of (3.2) is deduced from (3.8) by,

$$\left\{ \begin{array}{l} \zeta_{j,+}^{n+1} = \zeta_{j,+}^n - \frac{\Delta t}{\varepsilon \Delta x} (\zeta_{j,+}^{n+1} - \zeta_{j,-}^{n+1}) \\ \quad + \frac{\Delta t}{\varepsilon D_\varepsilon \Delta x} \left( \zeta_{j-1,+}^{n+1} - \cos(2\lambda_\varepsilon \Delta x) \zeta_{j,-}^{n+1} \right) \\ \quad + \frac{i\Delta t}{\varepsilon \Re(D_\varepsilon) \Delta x} \frac{\varepsilon V_t}{\lambda_\varepsilon} \sin(2\lambda_\varepsilon \Delta x) \zeta_{j,-}^{n+1} \\ \zeta_{j,-}^{n+1} = \zeta_{j,-}^n + \frac{\Delta t}{\varepsilon \Delta x} (\zeta_{j,+}^{n+1} - \zeta_{j,-}^{n+1}) \\ \quad + \frac{\Delta t}{\varepsilon D_\varepsilon \Delta x} \left( \zeta_{j+1,-}^{n+1} - \cos(2\lambda_\varepsilon \Delta x) \zeta_{j,+}^{n+1} \right) \\ \quad + \frac{i\Delta t}{\varepsilon \Re(D_\varepsilon) \Delta x} \frac{\varepsilon V_t}{\lambda_\varepsilon} \sin(2\lambda_\varepsilon \Delta x) \zeta_{j,+}^{n+1} \end{array} \right. \quad (3.9)$$

thanks to the fact that  $\Re(D_\varepsilon)$  retains only the correct (real) part of the denominator.

*Remark 3.2* For the  $L^2$ -norm, a mid-point time-integrator for (3.9) is an appealing choice thanks to the Cayley theorem. If the time-potential is not constant in  $x$ , the scheme (3.9) involves both values  $V_t(x_{j\pm\frac{1}{2}})$  acting on each  $\zeta_{j,\mp}^n$ , respectively. Numerical computations involving a Crank-Nicolson time-integrator like (3.5) are in §4.3.

The formal Schrödinger limit is obtained by adding both equations of (3.9),

$$\begin{aligned} \tilde{\psi}_j^{n+1} &= \tilde{\psi}_j^n - \frac{\Delta t}{\varepsilon D_\varepsilon \Delta x} \left( \tilde{\psi}_{j-1}^{n+1} - 2\cos(2\lambda_\varepsilon \Delta x) \tilde{\psi}_j^{n+1} + \tilde{\psi}_{j-1}^{n+1} \right) \\ &\quad + \frac{i\Delta t \sin(2\lambda_\varepsilon \Delta x)}{\Re(D_\varepsilon) \lambda_\varepsilon \Delta x} \left( \frac{V_t}{2} \tilde{\psi}_j^{n+1} \right), \\ D_\varepsilon &= \cos(2\lambda_\varepsilon \Delta x) - i \left( \frac{\varepsilon V_t + m/\varepsilon}{\lambda_\varepsilon} \right) \sin(2\lambda_\varepsilon \Delta x). \end{aligned}$$

under the innocuous conditions  $\varepsilon \rightarrow 0$ ,  $\lambda_\varepsilon \Delta x \rightarrow 0$  (and after re-substituting  $V_t \rightarrow \frac{V_t}{2}$ ). Such a first-order “implicit Euler” discretization, which appears to be *an original Schrödinger scheme containing tiny relativistic features*, dissipates the  $L^2$  norm.

**Proposition 1** *Assume the time-potential  $V_t \in L^\infty(\mathbb{R})$  is a real smooth function, then the scheme (3.10) is asymptotic-preserving in the non-relativistic limit  $\varepsilon \rightarrow 0$ .*

*Proof:* With the former computations at hand, there are 2 cases to look at:

- Assume first that the time-potential  $V_t(x) \geq 0$ , the main novelty is that both  $\lambda_\varepsilon$  and  $D_\varepsilon$  depend on space and one must consider:

$$\left\{ \begin{array}{l} \lambda_\varepsilon(x_{j-\frac{1}{2}}) = \sqrt{(\varepsilon V_t(x_{j-\frac{1}{2}}) + m/\varepsilon)^2 - (m/\varepsilon)^2} \in \mathbb{R}, \\ D_\varepsilon(x_{j-\frac{1}{2}}) = \cos\left(2\lambda_\varepsilon(x_{j-\frac{1}{2}})\Delta x\right) - i \frac{\varepsilon V_t(x_{j-\frac{1}{2}}) + m/\varepsilon}{\lambda_\varepsilon(x_{j-\frac{1}{2}})} \sin\left(2\lambda_\varepsilon(x_{j-\frac{1}{2}})\Delta x\right). \end{array} \right.$$

The scheme (3.9) is amended as follows:

$$\left\{ \begin{array}{l} \zeta_{j,+}^{n+1} = \zeta_{j,+}^n - \frac{\Delta t}{\varepsilon \Delta x} (\zeta_{j,+}^{n+1} - \zeta_{j,-}^{n+1}) \\ \quad + \frac{\Delta t}{\varepsilon D_\varepsilon(x_{j-\frac{1}{2}}) \Delta x} \left( \zeta_{j-1,+}^{n+1} - \cos(2\lambda_\varepsilon(x_{j-\frac{1}{2}}) \Delta x) \zeta_{j,-}^{n+1} \right) \\ \quad + \frac{i \Delta t}{\varepsilon \Re(D_\varepsilon(x_{j-\frac{1}{2}})) \Delta x} \frac{\varepsilon V_t(x_{j-\frac{1}{2}})}{\lambda_\varepsilon(x_{j-\frac{1}{2}})} \sin(2\lambda_\varepsilon(x_{j-\frac{1}{2}}) \Delta x) \zeta_{j,-}^{n+1} \\ \zeta_{j,-}^{n+1} = \zeta_{j,-}^n + \frac{\Delta t}{\varepsilon \Delta x} (\zeta_{j,+}^{n+1} - \zeta_{j,-}^{n+1}) \\ \quad + \frac{\Delta t}{\varepsilon D_\varepsilon(x_{j+\frac{1}{2}}) \Delta x} \left( \zeta_{j+1,-}^{n+1} - \cos(2\lambda_\varepsilon(x_{j+\frac{1}{2}}) \Delta x) \zeta_{j,+}^{n+1} \right) \\ \quad + \frac{i \Delta t}{\varepsilon \Re(D_\varepsilon(x_{j+\frac{1}{2}})) \Delta x} \frac{\varepsilon V_t(x_{j+\frac{1}{2}})}{\lambda_\varepsilon(x_{j+\frac{1}{2}})} \sin(2\lambda_\varepsilon(x_{j+\frac{1}{2}}) \Delta x) \zeta_{j,+}^{n+1} \end{array} \right. \quad (3.10)$$

Thanks to the smoothness of  $x \mapsto V_t(x)$ , the interface coefficients are still consistent because  $V_t(x_{j \mp \frac{1}{2}}) - V_t(x_j) \rightarrow 0$  strongly as  $\Delta x \rightarrow 0$ .

- In case  $V_t(x)$  has no definite sign, it can happen that  $\lambda_\varepsilon(x_{j-\frac{1}{2}})$  becomes purely imaginary for some  $j \in \mathbb{Z}$ . But this doesn't create any substantial issue because all the (circular) trigonometric relations which were used formerly still hold for (hyperbolic) trigonometric functions: for all  $x \in \mathbb{R}$ ,

$$\begin{aligned} \cos(ix) &= \frac{\exp(i^2x) + \exp(-i^2x)}{2} = \cosh(x), \\ \sin(ix) &= \frac{\exp(i^2x) - \exp(-i^2x)}{2i} = i \sinh(x), \end{aligned}$$

thus one recovers  $\cos^2(ix) + \sin^2(ix) = 1$ , together with all the limits as  $x \rightarrow 0$ .

□

## 4 Numerical results in various settings

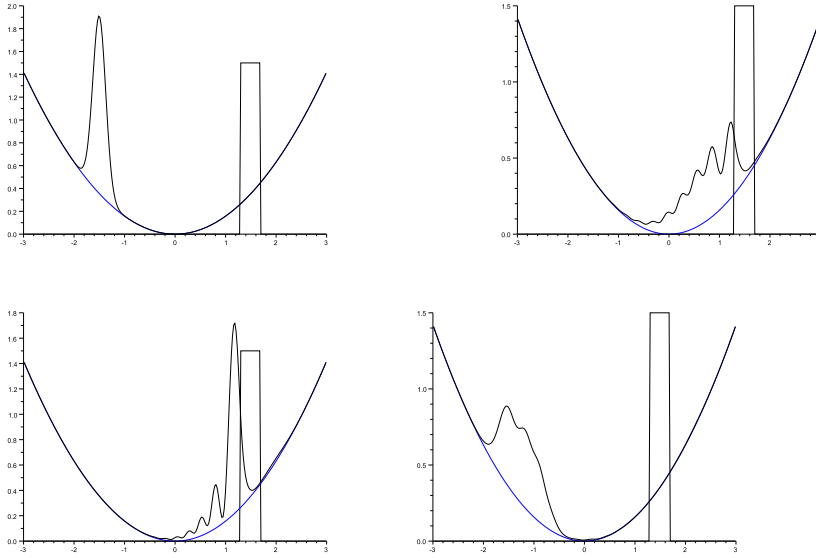
### 4.1 A test-case for linear problems

Here we simply aim at displaying the behavior of the explicit scheme (2.7) on a standard test-case for (2.1) involving both a pseudo-scalar and a time potential. More precisely, the pseudo-scalar term  $V_s(x) = \frac{\alpha x^2}{2}$  has a confinement effect, free from the Klein paradox. The time potential corresponds to a wall modeled by means of an indicator function,  $V_t(x) = \frac{3}{2} \chi_{1.25 < x < 1.75}$ . Initial data for  $\zeta_\pm(t=0, x)$  read

$$\sqrt{\frac{1}{\sigma \sqrt{2\pi}} \exp\left(-\frac{(x+1.5)^2}{2\sigma^2}\right)}, \quad \sigma = 0.125.$$

The computational domain is  $x \in (-3, 3)$  and  $N = 2^8$  points are used to grid it. On Fig. 4.1, results are shown at different times with a Courant number  $\nu = 1$ , allowing for an exact (up to machine accuracy) conservation of the  $L^2$  norm. We show the moduli added to the pseudo-scalar potential,  $V_s(x) + |\zeta_+|^2 + |\zeta_-|^2$  in order to emphasize its effects on the wave packet's dynamics. *Zitterbewegung* is observed especially when the movement is interrupted by the wall, and a small tunneling effect takes place.





**Fig. 4.1** Snapshots of (2.7) at  $t = 0, 1.54, 2.08, 3.54$  of a pulse on a parabola hitting a wall with  $c = 1$ .

## 4.2 Formal extension to nonlinear problems

### 4.2.1 Well-balanced upwind scheme

Following some examples displayed in [8], we intend to test the scheme (2.7) against exact solutions of a nonlinear Dirac equation (compare with (2.1), no time-potential):

$$\partial_t \Psi + \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \partial_x \Psi + i \begin{pmatrix} 1 + |\psi_-|^2 - |\psi_+|^2 \\ -1 + |\psi_+|^2 - |\psi_-|^2 \end{pmatrix} \cdot \Psi = 0. \quad (4.1)$$

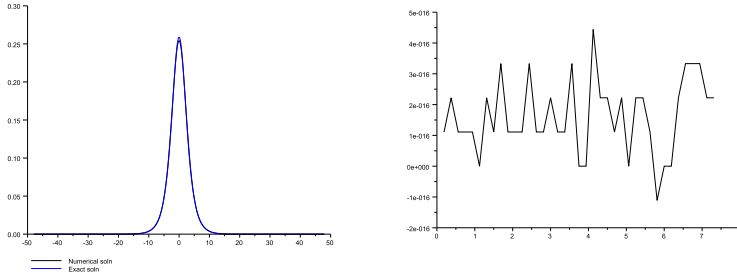
According to [1, 2, 8], this equation admits a family of exact solutions given by:

$$\psi_+(t, x) = \exp(-i\Lambda t)A(x), \quad \psi_-(t, x) = i \exp(-i\Lambda t)B(x),$$

where  $\Lambda \in (-1, 1)$ ,  $\mu = \sqrt{1 - \Lambda^2}$ ,  $\mu_A = \mu \sqrt{2(1 + \Lambda)}$ ,  $\mu_B = \mu \sqrt{2(1 - \Lambda)}$ , and

$$A(x) = \mu_A \frac{\cosh(\mu x)}{1 + \Lambda \cosh(2\mu x)}, \quad B(x) = \mu_B \frac{\sinh(\mu x)}{1 + \Lambda \cosh(2\mu x)}.$$

In order to tackle (4.1) by means of (2.7), it suffices therefore to cancel the time-potential  $V_t$  and to include the nonlinearity  $|\psi_-|^2 - |\psi_+|^2$ , “frozen at each time  $t^n$ ”, in the pseudo-scalar potential  $V_s$ . This yields a “nonlinearly varying mass” instead of the linear space-varying mass  $m(x)$  showing up in (2.1). Numerical results for the modulus of  $\Psi$  with  $N = 2^9$  points gridding the interval  $x \in (-32, 32)$  and a Courant number equal to 1,  $\Delta t = \Delta x$ , are displayed on Fig. 4.2, left. The time  $T = \frac{2\pi}{\Lambda}$  is chosen so that the exact solutions are equal to the initial data. Hence, despite the

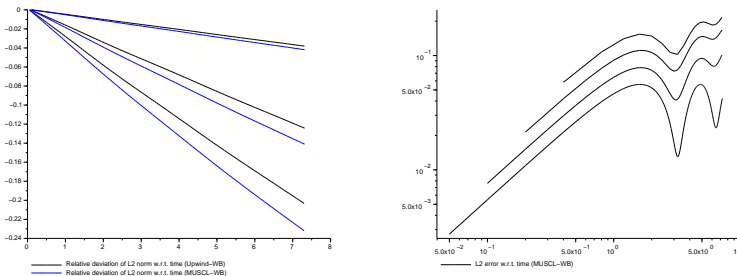


**Fig. 4.2** Solution of (4.1) for  $T = \frac{2\pi}{\Lambda}$ ,  $\Lambda = \sqrt{\frac{3}{4}}$  (left),  $L^2$ -norm change (right).

forementioned derivations don't rigorously extend to nonlinear cases, it is found numerically that the scheme (2.7) behaves rather correctly on these exact solutions, the  $L^2$  norm is conserved too, up to machine accuracy (see Fig. 4.2, right).

#### 4.2.2 Testing the well-balanced (WB) MUSCL extension

As said before, the  $L^2$ -preservation property holds essentially for  $v = 1$ , that is, in the absence of numerical diffusion (Lemma 2.2). One may wonder whether the MUSCL reconstruction (2.9)–(2.10) is able to improve this situation. Of course, as explained in Remark 2.1, the  $L^2$  norm corresponds to an entropy for the semi-linear hyperbolic system (see [48,49] for applications of hyperbolic techniques to Dirac systems) thus it is dissipated by finite-difference methods for the sake of stability in the presence of steep gradients. On the left of Fig. 4.3, we display the sensitivity of the time-evolution



**Fig. 4.3** Relative dissipation of  $L^2$ -norm for  $v = 0.9, 0.65, 0.4$  and  $\Delta x = 2^{-8}$  (left), growth of  $L^2$  errors for 4 decreasing values of  $\Delta x = 2^{-7}, 2^{-8}, 2^{-9}, 2^{-10}$  but  $v = 0.8$  kept constant (right), both functions of time.

of the  $L^2$  norm with respect to the Courant number for both the upwind (2.7) and the MUSCL (2.9)–(2.10) schemes on the semilinear problem (4.1). The improvement in terms of  $L^2$ -norm is of the order of 15% roughly, so this type of numerical scheme is too much dissipative for Courant numbers below 0.8. On the right of Fig. 4.3, we show that, for  $v = 0.8$  the  $L^2$ -error grows linearly in time while displaying some

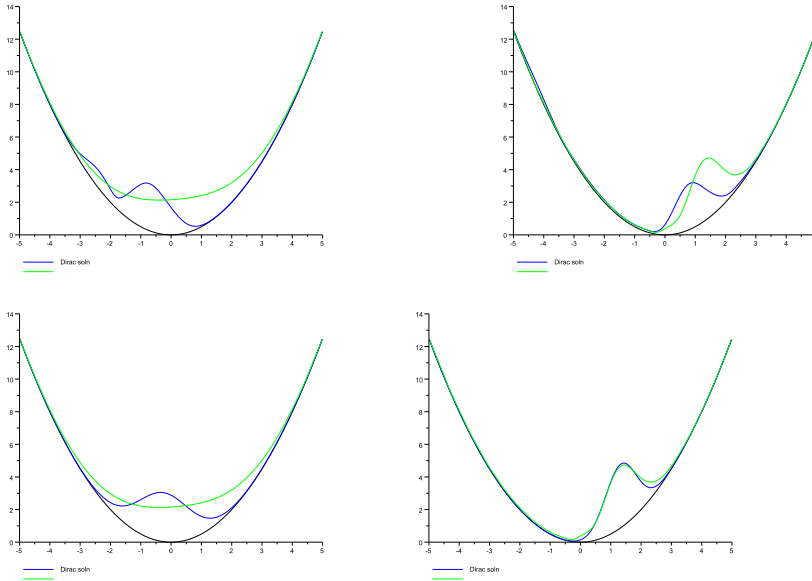
oscillations. Moreover, the decay of these errors when  $\Delta x$  passes from  $2^{-7}$  to  $2^{-10}$  appears to be linear, too. The parameters were kept identical to those of Fig. 4.2.

### 4.3 Non-relativistic limit and Asymptotic-Preserving (AP)

We confront numerically the scheme (3.9) against a direct Schrödinger computation: more precisely, a Crank-Nicolson time-integrator is applied both to (3.2), with the same numerical fluxes leading to (3.9), and to the usual non-relativistic equation,

$$i\partial_t \psi + \frac{1}{2m} \partial_{xx} \psi = V_l(x) \psi, \quad V_l(x) = \frac{x^2}{2}, \quad (4.2)$$

in the domain  $x \in (-5, 5)$  with  $N = 2^9$  grid points. Initial data correspond to a Gaussian pulse centered in  $x = -\frac{5}{4}$  and with wave number  $k = 1$ . Numerical densities of presence are nearly identical as soon as  $\varepsilon < 0.001$  and they roughly match the ones generated by the direct Schrödinger computation (see Fig. 4.4, bottom line). Oppositely, when the Dirac equation is in relativistic regime  $\varepsilon \simeq 1$ , noticeable differences show up as, for small time, the initial pulse splits into 2 components (Fig. 4.4, top, left), and despite the potential's effect, the numerical solutions are quite different at later time (Fig. 4.4, top, right). This was to be expected as the solutions for Dirac and



**Fig. 4.4** Comparison of Crank-Nicolson results on both (3.2) and Schrödinger equation (4.2) at  $T = 1$  (left),  $T = 3$  (right) and for  $\varepsilon = 1$  (top),  $\varepsilon = 0.01$  (bottom).

Schrödinger equations have no reason to look like each other in relativistic regime.

Following Remark 3.2, a Crank-Nicolson approach of the type (3.5) was implemented for this benchmark in order to show its numerical stability even in the presence of a singular limit  $\varepsilon \rightarrow 0$ ; practically, it seems that simulating the “relaxing system” (3.2) is less demanding in terms of numerical boundary conditions. We chose to implement Neumann boundary conditions, even if it is of little importance because the quantum particle is unable to reach the borders in presence of the steep parabolic potential (this situation is similar to the one of Fig. 4.1).

## 5 Conclusion

The classical defect of finite-difference methods in the context of quantum mechanics is their numerical viscosity which generally perturbs the  $L^2$ -norm of the approximate solution (one remedy is the setup of second-order implicit time-integrators [22, 32, 46]). Here we showed that, at least in the relativistic scaling, the well-balanced approach yields a numerical scheme (2.7) on an inhomogeneous linear problem which preserves exactly the  $L^2$ -norm while being completely explicit in time. When it comes to the (singular) non-relativistic limit of this scheme, there are 2 different situations:

1. for free particles (no external time-potential), the well-balanced scheme (2.7) can be reformulated into an asymptotic-preserving one (3.8) by mimicking former computations designed in the context of discrete kinetic models [20],
2. for particles submitted to an external time-potential, the coupling between different source terms leads to a “strange imaginary term” in  $D_\varepsilon$  which produces an incorrect behavior on the potential **only** (the Laplacian term is always correct). However, it is easy to fix this issue by taking the real part of this  $D_\varepsilon$  coefficient.

A very different strategy for quantum mechanics equations, better suited for  $L^2$ -norm preservation, is the split-step methods [17, 26, 33]; these algorithms too can furnish asymptotic-preserving strategies in the non-relativistic limit, see Remark 4.3 in [26]. This scheme has all the strengths and weaknesses which characterize the Fourier-based time-splitting schemes:

- Spectral accuracy and  $L^2$ -norm conservation on periodic smooth solutions;
- Possible issues to reproduce Jost “distorted plane wave” solutions (the ones which constitute the continuous spectrum of the stationary operator) in a scattering environment [16]. Reflection/transmission coefficients may result flawed unless a fine grid is used with potentials  $V(x)$  being either oscillating (see [29], page 183) or sharply-varying (thus inducing high frequencies).

Multi-dimensional problems are usually handled by dimensional splitting, but it is delicate matters in case one wants to reproduce accurately wave interactions in a context of weak solutions. A first step may be to consider the simpler case of so-called “massless Dirac equations” for which the 4-component spinor reduces to a 2-component one, see for instance [37].

**Acknowledgements** The author thanks Dr. S. Malaguti (Ferrara) for having suggested this problem.

## A Discrete $H^1(\mathbb{R})$ -bounds and strong convergence of free particles

**Lemma A.1** Assume  $c\Delta t \leq \Delta x$ ,  $x \mapsto V_t(x), V_s(x)$  being Lipschitz and  $\zeta_{\pm}(t=0, \cdot) \in H^1(\mathbb{R})$ , then numerical approximations generated by (2.7) satisfy:

$$\forall n \in \mathbb{N}, \quad \|\zeta^{\Delta x}(t^n, \cdot)\|_{L^2(\mathbb{R})} \leq \|\zeta^{\Delta x}(0, \cdot)\|_{L^2(\mathbb{R})}.$$

If moreover,  $\partial_x V_t = \partial_x V_s \equiv 0$ , then:

$$\forall n \in \mathbb{N}, \quad \|\partial_x \zeta^{\Delta x}(t^n, \cdot)\|_{L^2(\mathbb{R})} \leq \|\partial_x \zeta^{\Delta x}(0, \cdot)\|_{L^2(\mathbb{R})}.$$

*Proof:* Hereafter we denote  $|\zeta_{\pm}| = \sqrt{|\zeta_+|^2 + |\zeta_-|^2}$  the Euclidian vectorial norm and recall that  $|S_{j+\frac{1}{2}} \zeta_{\pm}|$ .

– As a consequence of the compact writing (2.8), one has

$$\begin{pmatrix} \zeta_{j,+}^{n+1} \\ \zeta_{j-1,-}^{n+1} \end{pmatrix} = (1-\nu) \begin{pmatrix} \zeta_{j,+}^n \\ \zeta_{j-1,-}^n \end{pmatrix} + \nu S_{j-\frac{1}{2}} \begin{pmatrix} \zeta_{j-1,+}^n \\ \zeta_{j,-}^n \end{pmatrix},$$

and deduces easily an  $L^2$  estimate for  $\nu \leq 1$  because each scattering matrix  $S_{j-\frac{1}{2}}$ ,  $j \in \mathbb{Z}$ , is unitary.

$$\begin{aligned} \left( \sum_{j \in \mathbb{Z}} \Delta x \left| \begin{pmatrix} \zeta_{j,+}^{n+1} \\ \zeta_{j-1,-}^{n+1} \end{pmatrix} \right|^2 \right)^{\frac{1}{2}} &\leq (1-\nu) \left\{ \sum_{j \in \mathbb{Z}} \Delta x \left| \begin{pmatrix} \zeta_{j,+}^n \\ \zeta_{j-1,-}^n \end{pmatrix} \right|^2 \right\}^{\frac{1}{2}} + \nu \left\{ \sum_{j \in \mathbb{Z}} \Delta x \left| S_{j-\frac{1}{2}} \begin{pmatrix} \zeta_{j-1,+}^n \\ \zeta_{j,-}^n \end{pmatrix} \right|^2 \right\}^{\frac{1}{2}} \\ &\leq (1-\nu) \left\{ \sum_{j \in \mathbb{Z}} \Delta x \left| \begin{pmatrix} \zeta_{j,+}^n \\ \zeta_{j-1,-}^n \end{pmatrix} \right|^2 \right\}^{\frac{1}{2}} + \nu \left\{ \sum_{j \in \mathbb{Z}} \Delta x \left| \begin{pmatrix} \zeta_{j-1,+}^n \\ \zeta_{j,-}^n \end{pmatrix} \right|^2 \right\}^{\frac{1}{2}} \\ &\leq \left( \sum_{j \in \mathbb{Z}} \Delta x \left| \begin{pmatrix} \zeta_{j,+}^n \\ \zeta_{j-1,-}^n \end{pmatrix} \right|^2 \right)^{\frac{1}{2}} \leq \left( \sum_{j \in \mathbb{Z}} \Delta x \left| \begin{pmatrix} \zeta_{j,+}^{n=0} \\ \zeta_{j,-}^{n=0} \end{pmatrix} \right|^2 \right)^{\frac{1}{2}}. \end{aligned}$$

The first step is just Minkowski's inequality and linearity of the  $L^2$  norm for  $0 \leq \nu \leq 1$ . The second one is the S-matrix being unitary. The third is shifting sum's indexes in order to produce

$$\sum_{j \in \mathbb{Z}} \left| \begin{pmatrix} \zeta_{j-1,+}^n \\ \zeta_{j,-}^n \end{pmatrix} \right|^2 = \sum_{j \in \mathbb{Z}} |\zeta_{j-1,+}^n|^2 + |\zeta_{j,-}^n|^2 = \sum_{j \in \mathbb{Z}} |\zeta_{j,+}^n|^2 + |\zeta_{j-1,-}^n|^2 = \sum_{j \in \mathbb{Z}} \left| \begin{pmatrix} \zeta_{j,+}^n \\ \zeta_{j-1,-}^n \end{pmatrix} \right|^2.$$

This reordering of the sums is specific to the  $L^2$  norm; therefore one needs both the S-matrix being unitary and the ability to rearrange the summation in order to derive the final estimate.

– When the scattering matrix is space-dependent, this doesn't automatically yield an estimate on the divided-differences: indeed, one sees on Fig. 4.1 that the presence of external potentials can increase the Lipschitz constant of the numerical approximations. One begins by "triangulating" as follows:

$$\begin{pmatrix} \zeta_{j+1,+}^{n+1} - \zeta_{j,+}^{n+1} \\ \zeta_{j,-}^{n+1} - \zeta_{j-1,-}^{n+1} \end{pmatrix} = (1-\nu) \begin{pmatrix} \zeta_{j+1,+}^n - \zeta_{j,+}^n \\ \zeta_{j,-}^n - \zeta_{j-1,-}^n \end{pmatrix} + \nu S_{j+\frac{1}{2}} \begin{pmatrix} \zeta_{j,+}^n - \zeta_{j-1,+}^n \\ \zeta_{j+1,-}^n - \zeta_{j,-}^n \end{pmatrix} + \nu [S_{j-\frac{1}{2}} - S_{j+\frac{1}{2}}] \begin{pmatrix} \zeta_{j-1,+}^n \\ \zeta_{j,-}^n \end{pmatrix}.$$

Then, one proceeds as before:

$$\begin{aligned}
\left( \sum_{j \in \mathbb{Z}} \Delta x \left| \begin{pmatrix} \zeta_{j+1,+}^{n+1} - \zeta_{j,+}^{n+1} \\ \zeta_{j,-}^{n+1} - \zeta_{j-1,-}^{n+1} \end{pmatrix} \right|^2 \right)^{\frac{1}{2}} &\leq (1-\nu) \left\{ \sum_{j \in \mathbb{Z}} \Delta x \left| \begin{pmatrix} \zeta_{j+1,+}^n - \zeta_{j,+}^n \\ \zeta_{j,-}^n - \zeta_{j-1,-}^n \end{pmatrix} \right|^2 \right\}^{\frac{1}{2}} \\
&\quad + \nu \left\{ \sum_{j \in \mathbb{Z}} \Delta x \left| S_{j+\frac{1}{2}} \begin{pmatrix} \zeta_{j,+}^n - \zeta_{j-1,+}^n \\ \zeta_{j+1,-}^n - \zeta_{j,-}^n \end{pmatrix} \right|^2 \right\}^{\frac{1}{2}} \\
&\quad + \frac{c\Delta t}{\Delta x} \left\{ \sum_{j \in \mathbb{Z}} \Delta x \left| [S_{j-\frac{1}{2}} - S_{j+\frac{1}{2}}] \begin{pmatrix} \zeta_{j-1,+}^n \\ \zeta_{j,-}^n \end{pmatrix} \right|^2 \right\}^{\frac{1}{2}} \\
&\leq \left\{ \sum_{j \in \mathbb{Z}} \Delta x \left| \begin{pmatrix} \zeta_{j+1,+}^n - \zeta_{j,+}^n \\ \zeta_{j,-}^n - \zeta_{j-1,-}^n \end{pmatrix} \right|^2 \right\}^{\frac{1}{2}} \\
&\quad + c\Delta t \cdot Lip_x(S) \cdot \left( \sum_{j \in \mathbb{Z}} \Delta x \left| \begin{pmatrix} \zeta_{j,+}^{n=0} \\ \zeta_{j,-}^{n=0} \end{pmatrix} \right|^2 \right)^{\frac{1}{2}} \\
&\leq \left\{ \sum_{j \in \mathbb{Z}} \Delta x \left| \begin{pmatrix} \zeta_{j+1,+}^{n=0} - \zeta_{j,+}^{n=0} \\ \zeta_{j,-}^{n=0} - \zeta_{j-1,-}^{n=0} \end{pmatrix} \right|^2 \right\}^{\frac{1}{2}} \\
&\quad + c(n+1)\Delta t \cdot Lip_x(S) \cdot \left( \sum_{j \in \mathbb{Z}} \Delta x \left| \begin{pmatrix} \zeta_{j,+}^{n=0} \\ \zeta_{j,-}^{n=0} \end{pmatrix} \right|^2 \right)^{\frac{1}{2}}.
\end{aligned}$$

We denoted by  $Lip_x(S)$  the Lipschitz constant in  $x$  of the S-matrix, which depends on the Lipschitz constant of both  $V_t(x)$  (the time potential) and  $V_s(x)$  (the pseudo-scalar term). One cannot deduce a discrete  $H^1$  estimate because of that last term. However, if  $V_t$  and  $V_s$  are constant,  $Lip_x(S) \equiv 0$  and:

$$\sum_{j \in \mathbb{Z}} \frac{1}{\Delta x} \left| \begin{pmatrix} \zeta_{j+1,+}^{n+1} - \zeta_{j,+}^{n+1} \\ \zeta_{j,-}^{n+1} - \zeta_{j-1,-}^{n+1} \end{pmatrix} \right|^2 \leq \sum_{j \in \mathbb{Z}} \frac{1}{\Delta x} \left| \begin{pmatrix} \zeta_{j+1,+}^{n=0} - \zeta_{j,+}^{n=0} \\ \zeta_{j,-}^{n=0} - \zeta_{j-1,-}^{n=0} \end{pmatrix} \right|^2. \quad (\text{A.1})$$

□

Lemma A.1 states that the  $L^2$ -norm (a particular entropy) is dissipated in time as soon as the Courant number  $\nu < 1$  by the upwind finite-differencing process. This is generally seen as a drawback for such a method; however, in certain circumstances and for short-time computations, dissipation can be interesting when an initial data oscillating at a frequency close to the grid's cutoff frequency  $\pi/\Delta x$  is prescribed. These data usually induce wave-packets endowed with flawed group velocities (see [35,50] and references therein) of the order of  $\Delta x$  and energy conservation implies that they never disappear. Instead, running a numerical scheme containing a little bit of numerical viscosity allows to progressively smoothen high frequencies and restore wave-packets with group velocities compatible with the computational grid.

**Lemma A.2** *Under the hypotheses of Lemma A.1, assume both  $V_t, V_s$  are constant, then:*

$$\forall n \in \mathbb{N}, \quad \|\zeta^{\Delta x}(t^{n+1}, \cdot) - \zeta^{\Delta x}(t^n, \cdot)\|_{L^2(\mathbb{R})} \leq c\Delta t \left\{ \|\partial_x \zeta^{\Delta x}(0, \cdot)\|_{L^2(\mathbb{R})} + Lip_{\Delta x}(S) \|\zeta^{\Delta x}(0, \cdot)\|_{L^2(\mathbb{R})} \right\}.$$

*Proof:* It consists in evaluating the size of the time-variation and taking advantage of previous  $H^1$  estimates:

$$\begin{aligned} \left( \sum_{j \in \mathbb{Z}} \Delta x \left| \begin{pmatrix} \zeta_{j,+}^{n+1} - \zeta_{j,+}^n \\ \zeta_{j-1,-}^{n+1} - \zeta_{j-1,-}^n \end{pmatrix} \right|^2 \right)^{\frac{1}{2}} &= v \left\{ \sum_{j \in \mathbb{Z}} \Delta x \left| \begin{pmatrix} \zeta_{j,+}^n \\ \zeta_{j-1,-}^n \end{pmatrix} - S_{j-\frac{1}{2}} \begin{pmatrix} \zeta_{j-1,+}^n \\ \zeta_{j,-}^n \end{pmatrix} \right|^2 \right\}^{\frac{1}{2}} \\ &\leq c\Delta t \left\{ \sum_{j \in \mathbb{Z}} \frac{1}{\Delta x} \left| \begin{pmatrix} \zeta_{j,+}^n - \zeta_{j-1,+}^n \\ \zeta_{j-1,-}^n - \zeta_{j,-}^n \end{pmatrix} \right|^2 \right\}^{\frac{1}{2}} \\ &\quad + c\Delta t \left\{ \sum_{j \in \mathbb{Z}} \Delta x \left| \frac{(Id - S_{j-\frac{1}{2}})}{\Delta x} \begin{pmatrix} \zeta_{j-1,+}^n \\ \zeta_{j,-}^n \end{pmatrix} \right|^2 \right\}^{\frac{1}{2}}. \end{aligned}$$

Now, from the expression (2.5), one sees that for any  $j \in \mathbb{Z}$ , if  $\Delta x = 0$ ,  $S_{j-\frac{1}{2}} = Id$ , the identity matrix. Hence  $|Id - S_{j-\frac{1}{2}}| \leq \Delta x \cdot Lip_{\Delta x}(S)$  and this is enough to ensure time-equicontinuity thanks to (A.1).  $\square$

Lemmas A.1 and A.2 allow to conclude that strong convergence holds in  $L_{loc}^2(\mathbb{R}_+ \times \mathbb{R})$  (and almost everywhere) as  $\Delta x \rightarrow 0$ ,  $c\Delta t \leq \Delta x$  for free particles of constant mass  $m$  without external potential.

## References

1. A. Alvarez, *Linearized Crank-Nicolson scheme for nonlinear Dirac equations*, J. Comp. Phys. **99** (1992) 348–350
2. A. Alvarez, Kuo Pen-Yu, L. Vazquez, *The numerical study of a nonlinear one-dimensional Dirac equation*, Applied Math. & Comput. **13** (1983) 1–15.
3. A. Askar A.S. Cakmak, *Explicit integration method for the timedependent Schrödinger equation for collision problems*, J. Chem. Phys. **68** 2794–2798 (1978).
4. Hocine Bahlouli, El Bouâzzaoui Choubabi, Ahmed Jellal, *Solution of one-dimensional Dirac equation via Poincaré map*, EPL **95** (2011) 17009. DOI: 10.1209/0295-5075/95/17009
5. P. Bechouche, N. Mauser, F. Poupaud, *Semi-(non)relativistic limits of the Dirac equation with external time-dependent electromagnetic fields*, Comm. Math Phys. **197** (1998) 405–425.
6. C. Berthou, C. Sarazin, R. Turpault, *Space-time Generalized Riemann Problem Solvers of Order  $k$  for Linear Advection with Unrestricted Time Step*, J. Sci. Comput. **55** 268–308 (2013)
7. S. D. Bosanac, *Solution of Dirac equation for a step potential and the Klein paradox*, J. Phys. A: Math. Gen. **40**(30) (2007) 8991.
8. N. Bournaveas, G.E. Zouraris, *Theory and numerical approximations for a nonlinear Dirac system*, Math. Modell. & Numer. Anal. (M2AN) **46** (2012) 841–874.
9. R. Carles, B. Mohammadi, *Numerical aspects of the nonlinear Schrödinger equation in the semiclassical limit in a supercritical regime*, Math. Modell. & Numer. Anal. (M2AN) **45** (2011) 981–1008
10. T. Chan, D. Lee, L. Shen, *Stable Explicit Schemes for Equations of the Schrödinger Type*, SIAM J. Numer. Anal. **23** (1986) 274.
11. Chen Jing-Bo, Liu Hong, *Two Kinds of Square-Conservative Integrators for Nonlinear Evolution Equations*, Chin. Phys. Lett. **25** (2008) 1168–1171
12. I. Cotaescu, P. Gravila, M. Paulescu, *Applying the Dirac equation to derive the transfer matrix for piecewise constant potentials*, Physics Letters A 366.4 (2007) 363–366
13. P. Degond, S. Gallego, F. Méhats, *An asymptotic preserving scheme for the Schrödinger equation in the semiclassical limit*, C.R. Math. Acad. Sci. Paris **345** (2007) 531–536.
14. F. Dominguez-Adame, M.A. Gonzalez, *Solvable linear potentials in the Dirac equation*, Europhys. Lett. **13**(3) (1990) 193–198.
15. F. Dominguez-Adame, A. Rodriguez, *A one-dimensional relativistic screened Coulomb potential*, Phys. Lett. A **198** (1995) 275–278.
16. V. Duchêne, J.L. Marzuola, M.I. Weinstein, *Wave operator bounds for one-dimensional Schrödinger operators with singular potentials and applications*, J. Math. Phys. **52** (2011) 013505.
17. J. De Frutos, J.M. Sanz-Serna, *Split-Step Spectral Schemes for Nonlinear Dirac Systems*, J. Comp. Phys. **83** (1989) 407–423

18. L. Gosse, **Computing Qualitatively Correct Approximations of Balance Laws**, Springer (2013) ISBN 978-88-470-2891-3
19. L. Gosse, *MUSCL reconstruction and Haar wavelets*, Preprint (2014)
20. L. Gosse, G. Toscani, *An asymptotic-preserving well-balanced scheme for the hyperbolic heat equations*, C.R. Math. Acad. Sci. Paris **334** (2002) 337–342.
21. B.Z. Guo, Hans Zwart, *On the Relation between Stability of Continuous- and Discrete-Time Evolution Equations via the Cayley Transform*, Integr. equ. oper. theory **54** (2006), 349383
22. E. Hairer, Ch. Lubich, G. Wanner, **Geometric numerical integration. Structure-preserving algorithms for ordinary differential equations**, Springer Series in Computational Mathematics. Springer-Verlag, Berlin Heidelberg (2006)
23. J. Hiller, *Solution of the one-dimensional Dirac equation with a linear scalar potential*, Amer. J. Phys. **70**(5) (2002) 522–524
24. Jialin Hong, Chun Li, *Multi-symplectic Runge-Kutta methods for nonlinear Dirac equations*, J. Comp. Phys. **211** (2006) 448–472
25. F. de la Hoz, F. Vadillo, *An integrating factor for nonlinear Dirac equations*, Computer Physics Comm. **181** (2010) 1195–1203
26. Zhongyi Huang, Shi Jin, Peter A. Markowich, Christof Sparber, Chunxiong Zheng, *A time-splitting spectral scheme for the Maxwell–Dirac system*, J. Comp. Phys. **208** (2005) 761–789
27. W. Hunziker, *On the Nonrelativistic Limit of the Dirac Theory*, Comm. Math. Phys. **40** (1975) 215–222
28. S. Jin, *Efficient asymptotic-preserving (AP) schemes for some multiscale kinetic equations*, SIAM J. Sci. Comput. **21** (1999) 441–454.
29. S. Jin, P.A. Markowich, C. Sparber, *Mathematical and computational methods for semiclassical Schrödinger equations*, Acta Numerica **20** (2011) 211–289.
30. Linghua Kong, Ruxun Liu, Xiaohong Zheng, *A survey on symplectic and multi-symplectic algorithms*, Applied Math. & Comput. **186** (2007) 670–684
31. Ph. LeFloch, A.E. Tzavaras, *Representation of weak limits and definition of nonconservative products*, SIAM J. Math. Anal. **30** (1999), 1309 – 1342.
32. Christian Lubich, *Integrators for Quantum Dynamics: A Numerical Analyst’s Brief Review*, in **Quantum Simulations of Complex Many-Body Systems: From Theory to Algorithms**, J. Grotendorst, D. Marx, A. Muramatsu (Eds.), John von Neumann Institute for Computing, Jülich, NIC Series **10** (2002), ISBN 3-00-009057-6, 459-466.
33. P. A. Markowich, P. Pietra, C. Pohl, *Numerical approximation of quadratic observables of Schrödinger-type equations in the semiclassical limit*, Numer. Math. **81** (1999) 595–630.
34. R. Mickens, *Stable explicit schemes for equations of Schrödinger type*, Phys. Rev. A **39** (1989)
35. S. Micu, *Uniform boundary controllability of a semi-discrete 1-D wave equation with vanishing viscosity*, SIAM J. Cont. Optim. **47** (2008), 2857–2885.
36. S M Morsink, R B Mann, *Black hole radiation of Dirac particles in 1+1 dimensions*, Class. Quantum Grav. **8** (1991) 2257
37. S. Noelle, *Hyperbolic systems of conservation laws, the Weyl equation, and multidimensional upwinding*, J. Comput. Phys. **115** (1994) 22–26.
38. A. Sinha, R. Roychoudury, *Dirac equation in (1+1)-dimensional curved space-time*, Intern. J. Theo. Phys. **33**:1511–1522, 1994.
39. D. Solomon, *An exact solution of the Dirac equation for a time-dependent Hamiltonian in 1-1 dimension space-time*, Canad. J. Phys. **88** 137–138 (2010)
40. C. Sparber, P.A. Markowich, *Semiclassical asymptotics for the Maxwell-Dirac system*, J. Math. Phys. **44** (2003) 4555.
41. S. Succi, *Numerical solution of the Schrödinger equation using discrete kinetic theory*, Phys. Rev. E **53** (1996) 1969–1975
42. S. Succi, R. Benzi, *Lattice Boltzmann equation for quantum mechanics*, Physica D **69** (1993) 327–332
43. B. Thaller, **Advanced Visual Quantum Mechanics**, (Chapter 7) New York: Springer (2005) ISBN 0-387-20777-5.
44. E.F. Toro, **Riemann Solvers and Numerical Methods for Fluid Dynamics: A Practical Introduction**, Third Edition, Springer (2009). [pages 427–29]
45. P. Weinberger, *All you need to know about the Dirac equation*, Phil. Magazine **88** 18-20 (2008) 2585-2601
46. P.P.F. Wessels, W.J. Caspers, F.W. Wiegel, *Discretizing the one-dimensional Dirac equation*, Europhys. Lett. **46**(2) (1999) 123–126.



- 
47. L. Wu, *Dufort-Frankel-Type Methods for Linear and Nonlinear Schrödinger Equations*. SIAM J. Numer. Anal. **33** (1996) 1526–1533
  48. Y. Zhang, *Global strong solution to a nonlinear Dirac-type equation in one dimension*. Nonlinear Analysis: Theory, Methods & Applications, **80** (2013) 150-155.
  49. Y. Zhang, *Global solution to a cubic nonlinear Dirac equation in  $1 + 1$  dimensions*, Preprint (2014).
  50. E. Zuazua, *Propagation, observation, and control of waves approximated by finite difference methods*, SIAM Review **47** (2005) 197–243.