# A TWO-DIMENSIONAL DIRAC SYSTEM AND ENVELOPE SOLUTIONS TO THE NONLINEAR SCHRÖDINGER EQUATION

David C. Webb

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Mathematics.

Chapel Hill
2017

Approved by:

Jeremy Marzuola

Hans Christianson

Jason Metcalfe

Michael Taylor

Mark Williams

ABSTRACT

DAVID C. WEBB: A Two-Dimensional Dirac System and Envelope Solutions to the Nonlinear Schrödinger Equation
(Under the direction of Jeremy Marzuola)

In studying the cubic nonlinear Schrödinger (NLS) equation with hexagonal lattice potential, Ablowitz, Nixon, and Zhu [1] and Fefferman and Weinstein [31] have used ansatz solutions of a periodic, cubic NLS equation to derive two similar two-dimensional Dirac equations with cubic nonlinearity.

Chapters 1 and 2 of this thesis deal with solutions and lifespans of solutions for the linear and nonlinear Dirac equations. We establish local and almost global existence results as well as ill-posedness below the critical regularity, $\dot{H}^{1/2}$. We leave as an open question whether solutions blow up in finite time or if a global existence result can be found.

The third chapter modifies the machinery of [32] and [31] to explore an open question posed by Fefferman and Weinstein, [31]. We prove that an envelope of solutions to the slowly modulated Dirac equation provides a good approximation for solutions to the nonlinear Schrödinger equation with hexagonal lattice. The NLS solution is shown to exist for long times with the nonlinear Dirac dynamics affecting the solution on any constant timescale. The same timescale is also proven for an ansatz envelope proposed by Ablowitz, Nixon, and Zhu. The timescale is also extended in an intermediate regime by slightly weakening the nonlinearity of the governing Dirac equation.

The final chapter discusses future work focusing on some numerical simulations for the Dirac and Schrödinger equations.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

# LIST OF FIGURES

## Introduction

The majority of this dissertation will be given over to studying the behavior of solutions to a particular two-dimensional system of Dirac equations with a cubic nonlinearity,

$$i\partial_t u - (\partial_x - i\partial_y)v = |u|^2 u$$

(0.1)

$$i\partial_t v + (\partial_x + i\partial_y)u = |v|^2 v$$

with initial data $u(0, \mathbf{x}) = f(\mathbf{x}), v(0, \mathbf{x}) = g(\mathbf{x})$. We will denote $U = \begin{pmatrix} u \\ v \end{pmatrix}$ a solution to the equation. This equation will be written in several alternate forms throughout this dissertation. The most common will be by defining the differential operator $P$ as follows:

(0.2) $$P \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} i\partial_t & -\partial_x + i\partial_y \\ \partial_x + i\partial_y & i\partial_t \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} |u|^2 u \\ |v|^2 v \end{pmatrix}.$$

We also will define the function $F$. When studying the linear equation, $F$ will merely be the inhomogeneity

$$F(t, x, y) = \begin{pmatrix} F_1(t, x, y) \\ F_2(t, x, y) \end{pmatrix}.$$

When studying the nonlinear equation $F$ will be the nonlinearity

$$F(u, v) = \begin{pmatrix} |u|^2 u \\ |v|^2 v \end{pmatrix}.$$

Thus the Dirac system can be written in the condensed form $PU = F$.

To get another way we will represent the equation we define the operator

$$\tilde{P} = \begin{pmatrix} i\partial_t & \partial_x - i\partial_y \\ -\partial_x - i\partial_y & i\partial_t \end{pmatrix}.$$

This operator is important because $\tilde{P}P = -\Box\text{Id}$ where $\Box$ is the wave equation operator. Thus we can act on the Dirac equation by $\tilde{P}$ to get

$$
\begin{pmatrix} -\Box u \\ -\Box v \end{pmatrix} = \begin{pmatrix} i\partial_t & \partial_x - i\partial_y \\ -\partial_x - i\partial_y & i\partial_t \end{pmatrix} \begin{pmatrix} F_1 \\ F_2 \end{pmatrix}
$$

(0.3)

$$
= \begin{pmatrix} i\partial_t F_1 + \partial_x F_2 - i\partial_y F_2 \\ i\partial_t F_2 - \partial_x F_1 - i\partial_y F_1 \end{pmatrix}
$$

with initial data $U(0, \mathbf{x}) = \begin{pmatrix} f(\mathbf{x}) \\ g(\mathbf{x}) \end{pmatrix}$ and $U_t(0, \mathbf{x}) = \begin{pmatrix} -i|f|^2 f - ig_x - g_y \\ -i|g|^2 g + if_x - f_y \end{pmatrix}$.

Thus solutions to the Dirac equation will also solve the above wave equation. We will see later that the fundamental solution operator to the Dirac equation has many similarities to the wave equation solution operator. Considering both of these facts, the large body of research in wave equations becomes immediately useful.

We will also work with the similar Dirac equation derived in [**31**] by Fefferman and Weinstein:

(0.4)

$$
\partial_t u + \overline{\lambda_\#}\big(\partial_x + i\partial_y\big)v = -ig\big(\beta_1|u|^2 + 2\beta_2|v|^2\big)u
$$

$$
\partial_t v + \lambda_\#\big(\partial_x - i\partial_y\big)u = -ig\big(2\beta_2|u|^2 + \beta_1|v|^2\big)v
$$

where $\lambda_\#$, $\beta_j$, and $g$ are constants defined in [**31**]. Throughout Chapters 1 and 2 we work primarily with (0.1) as it will be easier to deal with notationally. However, the arguments will apply identically to (0.4). In Chapter 3 we will work more closely with Equation 0.4 as it works more readily with the machinery used in that chapter.

**0.1. Motivation for the Equation.** The above Dirac equation is derived from optical physics in Ablowitz, Nixon, Zhu [**1**] as well as in Ablowitz, Zhu [**2**]. They are studying the behavior of light passing through the material optical graphene, which is an optical material with a hexagonal lattice. In particular they are studying what happens when the wavelength of light used corresponds to certain coordinates (called diabolical points) of the lattice.

They begin by using the nonlinear Schrödinger (NLS) equation,

(0.5)
$$
i\Psi_t + \Delta\Psi - V(\mathbf{x})\Psi + \sigma|\Psi|^2\Psi = 0,
$$

to model the behavior of light passing through an optical medium. As optical graphene has a hexagonal lattice they use the honeycomb lattice potential $V(\mathbf{x}) = V_0|e^{ik_0\mathbf{b}_1\cdot\mathbf{x}} + e^{ik_0\mathbf{b}_2\cdot\mathbf{x}} + e^{ik_0\mathbf{b}_3\cdot\mathbf{x}}|$

where $\mathbf{b}_1 = (0,1), \mathbf{b}_2 = (-\frac{\sqrt{3}}{2}, -\frac{1}{2}), \mathbf{b}_3 = (\frac{\sqrt{3}}{2}, -\frac{1}{2})$ represent the nearest neighbors to any particular node on the lattice.

Ablowitz, Nixon, and Zhu [1] create an ansatz for the NLS for which solutions to the Dirac system serve as an appropriate envelope within the ansatz, (0.1). However, it is important to note that their derivation of the Dirac system is not mathematically rigorous. It is a reasonable guess for a model to describe a behavior of optical graphene. Part of our exploration of this Dirac system will be to determine how well this model matches the experimental and numerical data as well as investigating on what time scales the model may be a good fit.

In [1] the authors ran numerical simulations inputting Gaussian initial data into the Schrödinger equation. In their simulations there appear to be 3 major qualities we see in the solution. While the solutions do not appear to be perfectly rotationally invariant, it does seem that $|\Psi(r, \theta_1)| \sim |\Psi(r, \theta_2)|$ for any choice of $\theta_i$. In other words, $r$ appears to have a much greater impact on the size of the solution than does $\theta$. Also, the solution appears to decay as you move away from the light cone. Most interesting is the fact that for their choice of initial data it appears that the solution has finite speed of propagation. This last result is quite notable as the Schrödinger equation is known to have infinite propagation speed.

The authors also ran numerical simulations for the derived linear Dirac system. In one of these simulations they input initial data that is Gaussian in one coordinate and zero in the other. This result also shows decay away from the light cone and finite speed of propagation. However, it also appears that $|U|$ is perfectly rotationally invariant showing some variance from the Schrödinger model. In another simulation, both components of the initial data were Gaussian. In this case, the decay off the light cone and finite propagation speed still appear, but the rotational invariance has disappeared completely. Not all of these numerical results are particularly notable. However, we include them here as we will analyticaly confirm these observations in Chapter 1.

The numerical simulations seem to indicate that the Dirac model is a reasonable yet imperfect model for the system considering the radial symmetry exhibited by the Dirac solutions which is not fully present in solutions to the Schrödinger equation. However, we will include a small error term in the ansatz solutions to account for the small discrepancy. In the Linear Results section of this dissertation we will analytically approach rotational invariance, decay away from the light cone, and finite speed of propagation to confirm the numerical results seen for the Dirac equation.

In the nonlinear section we will consider several function spaces and determine on what time scale solutions to the Dirac equation will exist in those spaces.

The Dirac equation derived by Fefferman and Weinstein, (0.4), is found by working with an ansatz solution to the periodic potential Schrödinger equation (0.5). They scale unknown functions $u$ and $v$ by a small constant, $\delta$, and show that the error term in the ansatz can be held small on a large timescale if $u$ and $v$ are components of the solution to the Dirac equation (0.4).

Their choice of ansatz is motivated by their work in [32] in which they prove a similar result for the corresponding linear Schrödinger and Dirac equations. Their work in [31] poses an open question using the solutions to (0.4) as ansatz envelopes for solutions to the NLS equation. They discuss some of the key changes and steps of the proof, but they do not actually provide a full proof for the ansatz solution nor do they give a timescale for the ansatz. In Chapter 3, we will modify the ansatz proposed by Fefferman and Weinstein and prove a timescale for which the ansatz solutions hold. We will also adapt the argument to an ansatz proposed by Ablowitz and Zhu, [2], which is more complicated because of different time and space scalings in their ansatz and certain assumptions they make, in particular that the potential is tight binding.

**0.2. Background.** As we have already shown that results pertaining to the wave equation are relevant to the Dirac equation in question we will begin by looking at established results for the wave equation particularly with cubic nonlinearity in two dimensions.

The case of the cubic wave equation in two dimensions is quite interesting. Li, Zhou showed in [78] that small data global solutions exist for all nonlinearities of power greater than three. Also, there are several examples, [3], [57] , [84] of cubic nonlinearities that result in nonexistence of global solutions even in small data. For a general cubic nonlinearity the existence time has only been shown to have a lower bound $T_* \geq c\epsilon^{-6}$ with $\epsilon$ the size of the initial data, [77].

Almost global results have been obtained where $\epsilon$ represents the size of the initial data. Kovalyov [54] proved a timescale of $T_* \geq \exp\left(c\epsilon^{-2}\right)$ for a class of nonlinearities that only depend on first and second derivatives of the solution. Li, Zhou [77] proved the same timescale by assuming $\partial_u^3 F(0,0,0) = \partial_u^4 F(0,0,0) = 0$ where $F(u, Du, D_x Du)$ is the nonlinearity. Katayama [47] dropped the assumption $\partial_u^4 F(0,0,0) = 0$ and proved the timescale $T_* \geq \epsilon^{-18}$. However, he also proved another almost global result for when the nonlinearity satisfies an almost-null condition.

There are also global existence results for the cubic nonlinearity. Godin [37] proved small data global existence for when the nonlinearity has quadratic and cubic parts which both satisfy

4

a null condition, and Katayama, [**46**] proved the same when only assuming the cubic part of the nonlinearity satisfies a null condition. In [**48**] Katayama proved a small data global result for a weaker null condition analagous to that of Alinhac [**6**] for three dimensions. Hoshiga [**40**] was able to prove a small data global existence result for a class of nonlinearity that does not satisfy a null condition. Also important to mention is the work of Alinhac [**5**] where he proves small data global existence for a large class of cubic derivative nonlinearities which satisfy two null conditions.

In general there is not as much literature studying the Dirac equation (massive or massless), and only a small amount of that literature deals with the two-dimensional case. In the one-dimensional case Pelinovsky [**67**] was able to prove small data global existence in $H^1$ for the massive Dirac equation with cubic nonlinearity. His proof relies on conservation laws for the equation and shows that a time dependent upper bound for the equation may continually grow but always remains finite.

In the three-dimensional case Escobedo and Vega [**27**] prove local well-posedness for the massive cubic Dirac equation in $H^{1+}$. This result was extended to small data global well-posedness by Machihara, Nakanishi, and Ozawa, [**62**], in the same space. Machihara, Nakamura, Nakanishi, and Ozawa, [**61**], were able to improve the regularity of the global result to $H^1$. Also in three dimensions Bejenaru and Herr [**16**] proved a small data global result for the massive equation with a different cubic nonlinearity in the space $H^1$ relying on the endpoint Strichartz estimates for the Klein-Gordon equation.

In the two-dimensional case Pecher [**66**] was able to show local well-posedness in $H^{\frac{1}{2}+}$ for a massive Dirac equation with a nonlinearity satisfying a null form. In a recent article Bejenaru and Herr [**17**] have improved this to small data global existence for the critical space $H^{\frac{1}{2}}$ by once again using the endpoint Strichartz estimates for the Klein-Gordon equation. Bournaveas and Candy [**20**] have been able to prove a similar small data global result in $\dot{H}^{\frac{1}{2}}$ for a cubic, massless equation with a null form.

CHAPTER 1

# Linear Results

## 1.1. Basic Results.

**1.1.1. Solution Operator.** Before we can move on to more interesting results, we need to first establish the solution operator of the Dirac equation. This will be found by applying the Fourier Transform to the equation, diagonalizing the resulting operator, exponentiating the diagonalized operator, reverting to the undiagonalized form, and applying the Inverse Fourier Transform. As the title of this sections implies, the results of this section are very basic. Because everything in this paper relies on these basic results, we cover this initial section in great detail in the interest of thoroughness.

By applying the Fourier Transform to the homogeneous equation, we get

$$i\partial_t \hat{u} - (i\xi_1 + \xi_2)\hat{v} = 0$$

(1.1)

$$i\partial_t \hat{v} + (i\xi_1 - \xi_2)\hat{u} = 0.$$

Defining $z = \xi_2 + i\xi_1$ We can rewrite this as

(1.2)
$$i\partial_t \begin{pmatrix} \hat{u} \\ \hat{v} \end{pmatrix} = \begin{pmatrix} 0 & z \\ \bar{z} & 0 \end{pmatrix} \begin{pmatrix} \hat{u} \\ \hat{v} \end{pmatrix} := \hat{\mathcal{H}} \begin{pmatrix} \hat{u} \\ \hat{v} \end{pmatrix}.$$

Notice that to do this we have assumed $z \neq 0$. In the case $z = 0$ we also have $\xi = 0$ and the equation reduces to an ordinary differential equation with solution $U = (f, g)$ which is what we get if we replace $\xi = 0$ in the solution operator we find below.

Now we can find the solution operator in Fourier Space, $e^{it\hat{\mathcal{H}}}$, by diagonalizing $\hat{\mathcal{H}}$. We find eigenvectors of $\hat{\mathcal{H}}$ to be $p_1 = \begin{pmatrix} z \\ |z| \end{pmatrix}$ and $p_2 = \begin{pmatrix} z \\ -|z| \end{pmatrix}$. Thus to diagonalize $\hat{\mathcal{H}}$ we conjugate it by

(1.3)
$$A = \begin{pmatrix} z & z \\ |z| & -|z| \end{pmatrix}.$$

Thus

$$D_{\hat{\mathcal{H}}} = A^{-1}\hat{\mathcal{H}}A$$

(1.4)
$$= \begin{pmatrix} |z| & 0 \\ 0 & -|z| \end{pmatrix}.$$

Now, since $D_{\hat{\mathcal{H}}}$ is a diagonal matrix, it is easy to compute

(1.5)
$$e^{-itD_{\hat{\mathcal{H}}}} = \begin{pmatrix} e^{-it|z|} & 0 \\ 0 & e^{it|z|} \end{pmatrix}.$$

Then conjugating backwards results in

$$e^{-it\hat{\mathcal{H}}} = Ae^{-itD_{\hat{\mathcal{H}}}}A^{-1}$$

(1.6)
$$= \begin{pmatrix} \cos(t|z|) & \frac{iz}{|z|}\sin(t|z|) \\ \frac{i\bar{z}}{|z|}\sin(t|z|) & \cos(t|z|) \end{pmatrix}.$$

Thus we have the solution operator $\begin{pmatrix} \hat{u} \\ \hat{v} \end{pmatrix} = e^{-it\hat{\mathcal{H}}}\begin{pmatrix} \hat{f} \\ \hat{g} \end{pmatrix}$. By applying the inverse Fourier Transform we find

(1.7)
$$\begin{pmatrix} u \\ v \end{pmatrix} = \mathcal{F}^{-1}\begin{pmatrix} \cos(t|z|)\hat{f} + \frac{iz}{|z|}\sin(t|z|)\hat{g} \\ \frac{i\bar{z}}{|z|}\sin(t|z|)\hat{f} + \cos(t|z|)\hat{g} \end{pmatrix}.$$

Then noting that $|z| = |\xi|$ we substitute into the above equation giving the solution operator

(1.8)
$$\begin{pmatrix} u \\ v \end{pmatrix} = \mathcal{F}^{-1}\begin{pmatrix} \left(\cos(t|\xi|)\hat{f}\right) + \left(\frac{-\xi_1+i\xi_2}{|\xi|}\sin(t|\xi|)\hat{g}\right) \\ \left(\frac{\xi_1+i\xi_2}{|\xi|}\sin(t|\xi|)\hat{f}\right) + \left(\cos(t|\xi|)\hat{g}\right) \end{pmatrix}.$$

**1.1.2. Global Well-Posedness.** In this section we will prove global well-posedness for the inhomogeneous linear equation. Having already found the solution operator for the homogeneous equation we will use Duhamel's principle to find the solution to the inhomogeneous equation. Then we will express an upper bound of the Sobolev norm of this solution in terms of the Sobolev norms of the initial conditions $f, g$ and the inhomogeneities $F_1, F_2$. Global well-posedness will easily follow from this inequality.

1.1.2.1. *Solving the inhomogeneous equation.* The solution of the homogeneous equation has been found to be

(1.9)
$$U = \begin{pmatrix} u \\ v \end{pmatrix} = \mathcal{F}^{-1} \begin{pmatrix} \cos(t|\xi|)\hat{f} + \frac{-\xi_1 + i\xi_2}{|\xi|}\sin(t|\xi|)\hat{g} \\ \frac{\xi_1 + i\xi_2}{|\xi|}\sin(t|\xi|)\hat{f} + \cos(t|\xi|)\hat{g} \end{pmatrix}.$$

Duhamel's principle tells us that the solution to the inhomogeneous case with zero initial conditions is in fact an integral of the solution to the homogeneous case with initial conditions equal to the original inhomogeneity. In other words, the solution to

(1.10)
$$i\partial_t u - (\partial_x - i\partial_y)v = F_1(t, x, y)$$

$$i\partial_t v + (\partial_x + i\partial_y)u = F_2(t, x, y)$$

with $u(0, x, y) = v(0, x, y) = 0$ is given by

(1.11)
$$\begin{pmatrix} u \\ v \end{pmatrix} = -i \int_0^t \mathcal{F}^{-1} \begin{pmatrix} \cos[(t-\tau)|\xi|]\hat{F}_1(\tau, \xi) + \frac{-\xi_1 + i\xi_2}{|\xi|}\sin[(t-\tau)|\xi|]\hat{F}_2(\tau, \xi) \\ \frac{\xi_1 + i\xi_2}{|\xi|}\sin[(t-\tau)|\xi|]\hat{F}_1(\tau, \xi) + \cos[(t-\tau)|\xi|]\hat{F}_2(\tau, \xi) \end{pmatrix} d\tau.$$

The solution for the inhomogeneous equation with nonzero initial data is merely the sum of the solutions to the two above cases:

(1.12)
$$U = \begin{pmatrix} u \\ v \end{pmatrix} = \mathcal{F}^{-1} \begin{pmatrix} \cos(t|\xi|)\hat{f} + \frac{-\xi_1 + i\xi_2}{|\xi|}\sin(t|\xi|)\hat{g} \\ \frac{\xi_1 + i\xi_2}{|\xi|}\sin(t|\xi|)\hat{f} + \cos(t|\xi|)\hat{g} \end{pmatrix}$$
$$+ \mathcal{F}^{-1} \begin{pmatrix} -i \int_0^t \left( \cos[(t-\tau)|\xi|]\hat{F}_1(\tau) + \frac{-\xi_1 + i\xi_2}{|\xi|}\sin[(t-\tau)|\xi|]\hat{F}_2(\tau) \right) d\tau \\ -i \int_0^t \left( \frac{\xi_1 + i\xi_2}{|\xi|}\sin[(t-\tau)|\xi|]\hat{F}_1(\tau) + \cos[(t-\tau)|\xi|]\hat{F}_2(\tau) \right) d\tau \end{pmatrix}.$$

1.1.2.2. *Bounding the Solution.* Assume that $f, g \in \mathcal{S}$ and $F_1, F_2 \in C^\infty([0, \infty], \mathcal{S})$. Then we can drop the inverse Fourier transform inside the $L^2$ norm without changing the value. Then, by applying the triangle inequality we can see that

(1.13)
$$\|u\|_{L^2} \le \left\| \cos(t|\xi|)\hat{f} \right\|_{L^2} + \left\| \frac{-\xi_1 + i\xi_2}{|\xi|}\sin(t|\xi|)\hat{g} \right\|_{L^2}$$
$$+ \int_0^t \left\| \cos[(t-\tau)|\xi|]\hat{F}_1(\tau, \xi) \right\|_{L^2} + \left\| \frac{-\xi_1 + i\xi_2}{|\xi|}\sin[(t-\tau)|\xi|]\hat{F}_2(\tau, \xi) \right\|_{L^2} d\tau.$$

Then we note that $|\cos(t|\xi|)|, |\sin(t|\xi|)|, |\frac{-\xi_1 + i\xi_2}{|\xi|}| \le 1$, and the above immediately becomes

(1.14)
$$\|u\|_{L^2} \le \|\hat{f}\|_{L^2} + \|\hat{g}\|_{L^2} + \int_0^t \left( \left\| \hat{F}_1(\tau, \cdot) \right\|_{L^2} + \left\| \hat{F}_2(\tau, \cdot) \right\|_{L^2} \right) d\tau.$$

8

Once again since in Schwartz space the $L^2$ norm of a function is the same as the $L^2$ norm of its Fourier transform, this becomes

(1.15)
$$\|u\|_{L^2} \leq \|f\|_{L^2} + \|g\|_{L^2} + \int_0^t \left( \left\| F_1(\tau, \cdot) \right\|_{L^2} + \left\| F_2(\tau, \cdot) \right\|_{L^2} \right) d\tau.$$

The exact same bound is found for $\|v\|_{L^2}$ with the same procedure leading to the bounds

(1.16)
$$\|U\|_{L^2} \leq C \left( \|f\|_{L^2} + \|g\|_{L^2} + \int_0^t (\|F_1(\tau)\|_{L^2} + \|F_2(\tau)\|_{L^2}) d\tau \right)$$
$$\|U\|_{H^s} \leq C \left( \|f\|_{H^s} + \|g\|_{H^s} + \int_0^t (\|F_1(\tau)\|_{H^s} + \|F_2(\tau)\|_{H^s}) d\tau \right)$$

where the $H^s$ norms are possibly infinite because we have not assumed initial data and inhomogeneities in the appropriate spaces.

1.1.2.3. *Global Well-Posedness in $H^s$.* Now assume that $f, g \in H^s$ and $F_1, F_2 \in L^1([0, \infty], H^s)$. Because of the density of $\mathcal{S}$ in $H^s$ we can choose sequences $f_j, g_j \in \mathcal{S}$ and $F_{1,j}, F_{2,j} \in L^1([0, \infty], \mathcal{S})$ such that as $j \to \infty$ each converges to its obvious limit. In other words,

$$\|f_j - f\|_{H^s} \to 0, \quad \|g_j - g\|_{H^s} \to 0, \quad \|F_{1,j} - F_1\|_{H^s} \to 0, \quad \|F_{2,j} - F_2\|_{H^s} \to 0.$$

We will denote $U_j$ to be the corresponding solutions to the Dirac equation with initial data $f_j, g_j$ and inhomogeneity $F_{1,j}, F_{2,j}$. Then the work we have done for Schwarz initial data gives us

(1.17) $\quad \|U_j - U_k\|_{H^s} \lesssim \|f_j - f_k\|_{H^s} + \|g_j - g_k\|_{H^s} + \int_0^t \|F_{1,j} - F_{1,k}\|_{H^s} + \|F_{2,j} - F_{2,k}\|_{H^s} d\tau.$

Thus the sequence $U_j$ is Cauchy in the Banach Space $H^s$, and hence converges to some $U$. But $U_j \to U$ in the distributional sense as well. Thus $PU_j \to PU$ where $P$ is our differential operator. But

(1.18)
$$PU_j = \begin{pmatrix} F_{1,j} \\ F_{2,j} \end{pmatrix} \to \begin{pmatrix} F_1 \\ F_2 \end{pmatrix}.$$

Thus $PU = \begin{pmatrix} F_1 \\ F_2 \end{pmatrix}$ meaning $U$ is solves the Dirac equation.

Since the bound for $\|U_j\|_{H^s}$ holds for every $j$, the same bound holds for the limit, $\|U\|_{H^s}$. Thus we have

$$
\begin{aligned}
(1.19) \qquad \|U\|_{H^s} &\leq C\Big(\|f\|_{H^s} + \|g\|_{H^s} + \int_0^t (\|F_1(\tau)\|_{H^s} + \|F_2(\tau)\|_{H^s}) d\tau\Big) \\
&\leq C\Big(\|f\|_{H^s} + \|g\|_{H^s} + \|F_1(\tau)\|_{L_t^1 H_{\mathbf{x}}^s} + \|F_2(\tau)\|_{L_t^1 H_{\mathbf{x}}^s}\Big)
\end{aligned}
$$

for $f, g \in H^s$ and $F_1, F_2 \in L^1([0, \infty], H^s)$.

Since $f, g \in H^s$ and $F_1, F_2 \in L^1([0, \infty], H^s)$, Inequality (1.19) immediately shows that $U \in H^s$ since $\|U\|_{H^s}$ has a uniform bound for all time.

For uniqueness since we are in the linear case it is sufficient to show that when $f = g = F_1 = F_2 = 0$ then $U$ must also be 0. This is trivial from (1.19).

To show that the solution depends continuously on initial data let us assume $U^*$ is the solution for initial data $f^*$ and $g^*$ with the same inhomogeneity as the equation for $U$. Then

$$
(1.20) \qquad \|U - U^*\|_{H^s} \leq C\Big(\|f - f^*\|_{H^s} + \|g - g^*\|_{H^s}\Big)
$$

which is the desired statement.

## 1.2. Analytic Support of Numerical and Experimental Observations.

The purpose of this section is to support some of the qualitative observations made by the numerical work in [1]. The subsection on Gaussian initial data supports the radial symmetry seen in [1], but is not strictly applicable to the later work in this dissertation. The subsection establishing decay away from the light cone is interesting as it quantifies the Huygens-like behavior seen in [1], but, once again, we do not make further use of it. The finite speed of propagation is vital for several of the nonlinear results in Chapters 2 and 3.

### 1.2.1. Gaussian Initial Data.

The experimental and numerical work done in [1] features radially symmetric initial data. Therefore we will analyze what happens when a Gaussian initial condition is input into the homogeneous linear Dirac equation. When discussing rotational invariance we will be working with the homogeneous Dirac equation

$$
i\partial_t u - (\partial_x - i\partial_y)v = 0
$$

$$
i\partial_t v + (\partial_x + i\partial_y)u = 0.
$$

One important property of the Gaussian function that we will use regularly is that the Fourier transform of a Gaussian, $e^{-k_1|\mathbf{x}|^2}$, is a different Gaussian, $e^{-k_2|\xi|^2}$. For the sake of simplicity we will just normalize the coefficient $k_1 = 1, k_2 = k$.

1.2.1.1. *The Magnitude of the Solution.* We will now assume that $f = e^{-|\mathbf{x}|^2}$ and $g = 0$. Then substituting in the initial conditions we find that

$$
\begin{aligned}
u(\mathbf{x},t) &= \mathcal{F}^{-1}\left(\cos(t|\xi|)e^{-k|\xi|^2}\right) \\
&:= \psi_1(\mathbf{x},t) \\
v(\mathbf{x},t) &= (-i\partial_x + \partial_y)\mathcal{F}^{-1}\left(\frac{1}{|\xi|}\sin(t|\xi|)e^{-k|\xi|^2}\right) \\
&:= (-i\partial_x + \partial_y)\psi_2(\mathbf{x},t).
\end{aligned}
$$

(1.21)

Now we consider the absolute value of our solution, $|U| = \sqrt{|u|^2 + |v|^2}$. Defining $\psi_1, \psi_2$ as above we notice that both are inverse Fourier transforms of even, real-valued functions and are thus real valued. Thus the real and imaginary parts of $u$ and $v$ are easy to calculate, and we find $|u|^2 = \psi_1^2$ and $|v|^2 = \psi_{2,x}^2 + \psi_{2,y}^2$. Thus

(1.22)
$$
|U| = (\psi_1^2 + \psi_{2,x}^2 + \psi_{2,y}^2)^{1/2}.
$$

It is important to note that $\psi_2$ was defined as the sine term of the solution *after* the $\xi_1, \xi_2$ have been pulled outside of the Inverse Fourier Transform as derivatives.

1.2.1.2. *Rotational Invariance.* Now we must stop and show that $\psi_1(\mathbf{x},t)$ and $\psi_2(\mathbf{x},t)$ are both rotationally invariant.

In order to do this we will first convert both into polar coordinates. Then, by doing a carefully chosen change of variables we can eliminate the dependence on $\theta$ in the equations. Once there is no dependence on $\theta$, it is trivial to say the function is rotationally invariant.

We first show this for $\psi_1$,

(1.23)
$$
\begin{aligned}
\psi_1(\mathbf{x},t) &= C\int e^{i\mathbf{x}\cdot\xi}\cos(t|\xi|)e^{-k|\xi|^2}d\xi \\
\psi_1(r,\theta,t) &= C\int e^{i(r\cos(\theta)\xi_1 + r\sin(\theta)\xi_2)}\cos(t|\xi|)e^{-k|\xi|^2}d\xi \\
&= C\int e^{i[r(\xi_1\cos\theta + \xi_2\sin\theta)]}\cos(t|\xi|)e^{-k|\xi|^2}d\xi.
\end{aligned}
$$

Then we do a simple change of variables rotating $\xi$ by $-\theta$, $\eta_1 = \xi_1 \cos\theta + \xi_2 \sin\theta$ and $\eta_2 = -\xi_1 \sin\theta + \xi_2 \cos\theta$. Thus $|\eta| = |\xi|$ and $d\eta = d\xi$.

$$(1.24) \qquad \psi_1(r,\theta,t) = C \int e^{ir\eta_1} \cos(t|\eta|)e^{-k|\eta|^2} d\eta.$$

We can now see that $\psi_1$ has no dependence on $\theta$ and is therefore rotationally invariant. Also we can proceed similarly doing the same change of variables to show that $\psi_2$ is rotationally invariant.

$$\psi_2(\mathbf{x},t) = C \int e^{i\mathbf{x}\cdot\xi} \sin(t|\xi|)|\xi|^{-1} e^{-k|\xi|^2} d\xi$$

$$\psi_2(r,\theta,t) = C \int e^{i(r\cos(\theta)\xi_1 + r\sin(\theta)\xi_2)} \sin(t|\xi|)|\xi|^{-1} e^{-k|\xi|^2} d\xi$$

$$(1.25)$$

$$= C \int e^{i[r(\xi_1\cos\theta + \xi_2\sin\theta)]} \sin(t|\xi|)|\xi|^{-1} e^{-k|\xi|^2} d\xi$$

$$= C \int e^{ir\eta_1} \sin(t|\eta|)|\eta|^{-1} e^{-k|\eta|^2} d\eta.$$

Here we now see that $\psi_2$ is also rotationally invariant since it has no dependence on $\theta$.

Now, we will rewrite (1.22) using the identity $\psi_{2,x}^2 + \psi_{2,y}^2 = \psi_{2,r}^2 + \frac{1}{r^2}\psi_{2,\theta}^2$. However, in our case $\psi_2$ is rotationally invariant, so $\psi_{2,\theta} = 0$. Thus, $|U| = \sqrt{\psi_1^2 + \psi_{2,r}^2}$, and there is no dependence on $\theta$. Thus we have confirmed that with the above Gaussian and zero initial conditions the solution to the linear Dirac problem maintains rotational invariance.

It is important to note that the above argument does not hold if the initial data $f, g$ are both Gaussian. In fact, if we assume that $f$ and $g$ are both Gaussian, then the above will verify that the solutions do not have radial symmetry, as indicated in the numerical work of [1].

**1.2.2. Decay Away from the Light Cone.** For the wave equation in odd spatial dimensions Huygens' Principle holds stating that the value of the solution at a point $(\mathbf{x}, t)$ depends only on the values of the initial condition on $\partial B(\mathbf{x}, t)$. In other words, solutions at a point are determined by the values of the initial data on a light cone centered at that point.

As we are working with a Dirac equation in 2 dimensions we do not expect this to hold. However, in the experimental and numerical work done in [1], the solutions appear to be concentrated near the light cone. In this section we analytically support that observation.

We will appeal to the Klainerman-Sobolev inequality originally developed by Klainerman in [52]. We will use the inequality as stated in [47]:

LEMMA 1.2.1 (Klainerman-Sobolev Inequality).

(1.26)
$$(1 + t + |\mathbf{x}|)^{\frac{1}{2}}(1 + |t - |\mathbf{x}||)^{\frac{1}{2}}|v(t, \mathbf{x})| \leq C\|v(t, \cdot)\|_{2,2}$$

*for any function $v$ with finite right-hand side.*

However, we need to define the norm on the right-hand side. To do this we introduce the following vector fields: $\Gamma_0 = t\partial_t + x\partial_x + y\partial_y, \Omega_{01} = t\partial_x + x\partial_t, \Omega_{02} = t\partial_y + y\partial_t, \Omega_{12} = x\partial_y - y\partial_x, \partial_t, \partial_x, \partial_y$. We index these vector fields by $\Gamma = \{\Gamma_i\}_{i=0}^{6}$. We discuss these more thoroughly in the Almost Global section in Chapter 2. For now we just need to define

(1.27)
$$|v(t, \mathbf{x})|_s = \sum_{|\alpha| \leq s} |\Gamma^\alpha v(t, \mathbf{x})|$$

$$\|v(t, \cdot)\|_{s,p} = \left\| |v(t, \cdot)|_s \right\|_{L^p}.$$

Using the Klainerman-Sobolev inequality, we can immediately state that for $U$, a solution to (0.1), we have

(1.28)
$$|U(t, \mathbf{x})| \leq (1 + t + |\mathbf{x}|)^{-\frac{1}{2}}(1 + |t - |\mathbf{x}||)^{-\frac{1}{2}}\|U(t, \cdot)\|_{2,2}$$

$$\leq (1 + |t - |\mathbf{x}||)^{-\frac{1}{2}}\|U(t, \cdot)\|_{2,2}.$$

If the norm on the right-hand side is finite, then we see that there is decay away from the light cone. However, this statement is vacuous if the norm on the right-hand side is not finite.

For the linear, inhomogeneous Dirac equation it is a quick result that the norm on the right-hand side is finite. If $U$ solves the linear, inhomogeneous Dirac equation, then $U$ also solves the linear, inhomogeneous wave equation (0.3). As we mention more carefully in Section 2.3, each of these vector fields commutes well with the wave equation operator, $\Box$, (except for $\Gamma_0$ which gives an extra copy of $\Box$ when commuting with $\Box$). Thus, $\|U(t, \cdot)\|_{2,2}$ can be bounded:

(1.29)
$$\|U(t, \cdot)\|_{2,2} \leq C\left(\|f\|_{2,2} + \|g\|_{2,2} + \int_0^t (\|F_1(\tau)\|_{2,2} + \|F_2(\tau)\|_{2,2})d\tau\right).$$

As long as we restrict the initial data to be in the 2,2-norm space and the inhomogeneities to be $L^1$ in time and be in the 2,2-norm space in space, then the right-hand side will stay finite. This analytically confirms the decay off the light cone seen in [1].

Even for the nonlinear Dirac equation, one consequence of our result from Section 2.3 will be that $\|U(t, \cdot)\|_{2,2}$ remains finite on an exponential time scale if the initial data is small enough and

13

compactly supported. Thus for the nonlinear Dirac equation, we technically only prove decay off the light cone for small enough initial data with compact support (since that is when we are able to prove that the norm above is finite). However, Ablowitz, Nixon, and Zhu run their simulations for a short enough time scale ($t = 10.8$) that this gives us reasonable analytical support for the qualitative observations made in [**1**] even in the nonlinear case (in addition to confirming the observation for the linear case).

**1.2.3. Finite Propagation Speed.** In this section we will show that the Dirac equation has finite speed of propagation (FPS).

As we have already shown in the introduction, solutions to the Dirac equation also solve a particular wave equation. As it is well known that the wave equation has finite speed of propagation, this property easily transmits to the Dirac equation.

1.2.3.1. *The Homogeneous Case.* If we assume $U$ is a solution to the homogeneous Dirac equation, then we can simply apply the operator $\tilde{P}$ to the equation to conclude

$$(1.30) \qquad \begin{pmatrix} -\Box u \\ -\Box v \end{pmatrix} = \begin{pmatrix} -\Box & 0 \\ 0 & -\Box \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \tilde{P} P \begin{pmatrix} u \\ v \end{pmatrix} = \tilde{P} \cdot 0 = 0.$$

Thus $U$ also solves the homogeneous wave equation.

Among many places, a proof for finite propagation speed of the wave equation is given in [**28**]. It states that if $u$ solves the wave equation and $u = u_t = 0$ on $B(x_0, t_0) \times \{t = 0\}$, then $u = 0$ within the backward light cone $C = \{(x, t) \mid 0 \le t \le t_0, |x - x_0| \le t_0 - t\}$. This is equivalent to saying that if $u_0$ is compactly supported on $K$, then for any time $t > 0$ we have that $u(\cdot, t)$ is supported on $B(K, t)$, all the points within a distance, $t$, from the support, $K$.

Thus $u, v$ both have finite propagation speed when they are solutions to the homogeneous Dirac problem.

1.2.3.2. *The Inhomogeneous Case.* To solve the inhomogeneous case we must look back at how we found the inhomogeneous solution. Namely, we must use Duhamel's Formula again. We already have FPS for solutions to the homogeneous case with nonzero initial conditions. Therefore if we can show FPS for solutions to the inhomogeneous case with zero initial conditions we will be done since the general solution is merely the sum of the two.

Thus we are looking for a solution to

$$(1.31) \qquad P\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} F_1 \\ F_2 \end{pmatrix}, \qquad u_0 = v_0 = 0.$$

Using Duhamel we first let $u(x, t; \tau)$ be a solution to

$$(1.32) \qquad P\begin{pmatrix} u(\mathbf{x}, t; \tau) \\ v(\mathbf{x}, t; \tau) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{on} \quad \mathbb{R}^2 \times (\tau, \infty)$$

with initial conditions

$$(1.33) \qquad \begin{pmatrix} u(\mathbf{x}, t; \tau) \\ v(\mathbf{x}, t; \tau) \end{pmatrix} = \begin{pmatrix} F_1(\mathbf{x}, t; \tau) \\ F_2(\mathbf{x}, t; \tau) \end{pmatrix} \quad \text{on} \quad \mathbb{R}^2 \times \{t = \tau\}.$$

Then by Duhamel's Formula the solution to (1.31) is merely $u(\mathbf{x}, t) = i \int_0^t u(\mathbf{x}, t; \tau) d\tau$ and similarly for $v$. This can be solved out further (which is done in the above section for Global Well-Posedness), but for the purposes of showing FPS this is all we need.

Notice that (1.32) is a homogeneous equation where the starting point is now $t = \tau$ instead of $t = 0$. Thus $F_1(\tau)$ having compact support on $K_\tau$ implies that $u(x, t; \tau)$ has compact support on $B(K_\tau, t - \tau)$.

Now we will consider the case when $F_1(\tau)$ has compact support, $K_\tau$, for all $\tau$. Then $u(x, t) = i \int_0^t u(x, t; \tau) d\tau$ is supported on the union of the balls, $B(K_\tau, t - \tau)$. Notice that

$$\cup_{\tau=0}^t B(K_\tau, t - \tau) \subset B\left(\cup_{\tau=0}^t K_\tau, t\right).$$

This yields FPS as long as $\cup K_\tau$ is bounded. However, this is trivial as it is just a union over a compact interval of compact sets. Notice also that $v$ and $F_2(\tau)$ are treated the same way.

In particular, for $F_1(\tau)$ compactly supported on $K_\tau$ for each $\tau$, We have that $u(x, t)$ is supported on $B(\cup_{\tau=0}^t K_\tau, t)$, which is a bounded set.

**1.3. Strichartz Estimates.** Strichartz estimates are used to bound the mixed space-time (Strichartz) norm (and potentially the Sobolev norm) of the solution by Sobolev norms of the initial data and a Strichartz norm of the inhomogeneities. They can give a broad range of Strichartz spaces on the right-hand side of the inequality giving the potential to bootstrap more optimal estimates than might otherwise be achieved.

15

We prove the following Strichartz estimates to allow us to get an improved local existence result for solutions to the nonlinear equation when working with Strichartz spaces. However, neither the Strichartz estimates nor the Strichartz space local existence result are needed in proving the main result in Chapter 3. Thus, these Strichartz estimates are mainly included for the sake of completeness and are found by making modifications to the wave equation Strichartz estimates.

It is also important to note that we will be reducing the solution operator to the half-wave operator. Thus, the Strichartz estimates are essentially the same as the well known Strichartz estimates for the wave equation with some modifications of the gap condition. Although originaly developed by Strichartz in [75], we will primarily make use of the proof in Lawrie, [56], to influence the overall strategy of the proof. In particular, his proof will work for the desired estimates with the addition of a Lemma from Stein, [73], and carefully tracking how the difference in the inhomogeneous terms of the fundamental solution changes the gap condition. Also of use in motivating this section were Sogge, [72], and Selberg, [70]. Strichartz estimates for the wave equation were also proved in [35, 50, 51, 59].

We will work with the inhomogeneous Dirac equation

(1.34)
$$i\partial_t u - (\partial_x - i\partial_y)v = n(t,x,y)$$
$$i\partial_t v + (\partial_x + i\partial_y)u = m(t,x,y)$$

with initial data $u(0,\cdot) = f(\cdot)$ and $v(0,\cdot) = g(\cdot)$ and inhomogeneities $n$ and $m$.

**1.3.1. Statement of Desired Strichartz Estimate.** Before getting started we must discuss what particular mixed space-time norms will work for the Strichartz estimates. To do so we define the following:

DEFINITION 1. We say that the pair $(q,r)$ is wave admissible if

(1.35)
$$2 \leq q \leq \infty, \quad 2 \leq r < \infty \quad \text{and} \quad \frac{2}{q} \leq \frac{n-1}{2}\left(1 - \frac{2}{r}\right).$$

In particular, since our Dirac equation is in 2 dimensions, wave admissible means

(1.36)
$$\frac{2}{q} \leq \frac{1}{2} - \frac{1}{r}.$$

In particular, we notice that this further restricts $q > 4$.

16

We will assume that $(q, r)$ and $(\tilde{q}, \tilde{r})$ are both wave admissible. We define $S_T = [0, T] \times \mathbb{R}^2$. Then we claim the following Strichartz Estimates for the above Dirac equation:

THEOREM 1.3.1. *Let $(q, r)$ be wave admissible. Assuming that $U$ solves Equation 1.34 with $f, g \in \dot{H}^s(\mathbb{R}^2)$ and $n = m = 0$ (i.e. the homogeneous equation), then*

$$(1.37) \qquad \|U\|_{L_t^q L_x^r (S_T)} \lesssim \|f\|_{\dot{H}^s} + \|g\|_{\dot{H}^s}$$

*where $s, q, r$ satisfy $\frac{2}{r} + \frac{1}{q} = 1 - s$.*

Notice that it suffices to prove this estimate for $u$ (since the proof for $v$ will be identical), and it is trivial to distribute using the triangle inequality and reduce to the half-wave operator

$$(1.38) \qquad \|u\|_{L_t^q L_x^r} \leq \left\| \mathcal{F}^{-1} \left( e^{it|\xi|} \hat{f} \right) \right\|_{L_t^q L_x^r} + \left\| \mathcal{F}^{-1} \left( \frac{-\xi_1 + i\xi_2}{|\xi|} e^{it|\xi|} \hat{g} \right) \right\|_{L_t^q L_x^r}.$$

We also state the desired Strichartz estimates for the inhomogeneous equation.

THEOREM 1.3.2. *Let $(q, r)$ and $(\tilde{q}, \tilde{r})$ both be wave admissible. Assuming that $U$ solves Equation 1.34 with $f, g \in \dot{H}^s(\mathbb{R}^2)$ and $n, m \in L_t^{\tilde{q}'} L_x^{\tilde{r}'}(S_T)$, then*

$$(1.39) \qquad \|U\|_{L_t^q L_x^r (S_T)} \lesssim \|f\|_{\dot{H}^s} + \|g\|_{\dot{H}^s} + \|n\|_{L_t^{\tilde{q}'} L_x^{\tilde{r}'}(S_T)} + \|m\|_{L_t^{\tilde{q}'} L_x^{\tilde{r}'}(S_T)}$$

*where $s, q, r, \tilde{q}, \tilde{r}$ satisfy the gap condition $\frac{2}{r} + \frac{1}{q} = 1 - s = \frac{2}{\tilde{r}'} + \frac{1}{\tilde{q}'} - 1$.*

Again, it suffices to prove this estimate for $u$ (since the proof for $v$ will be identical), and it is trivial to distribute using the triangle inequality and reduce to the half-wave operator

$$(1.40) \qquad \begin{aligned} \|u\|_{L_t^q L_x^r} &\leq \left\| \mathcal{F}^{-1} \left( e^{it|\xi|} \hat{f} \right) \right\|_{L_t^q L_x^r} + \left\| \mathcal{F}^{-1} \left( \frac{-\xi_1 + i\xi_2}{|\xi|} e^{it|\xi|} \hat{g} \right) \right\|_{L_t^q L_x^r} \\ &+ \left\| \int_0^t \mathcal{F}^{-1} \left( e^{i(t-\tau)|\xi|} \hat{n}(\tau) \right) d\tau \right\|_{L_t^q L_x^r} \\ &+ \left\| \int_0^t \mathcal{F}^{-1} \left( \frac{-\xi_1 + i\xi_2}{|\xi|} e^{i(t-\tau)|\xi|} \hat{m}(\tau) \right) d\tau \right\|_{L_t^q L_x^r}. \end{aligned}$$

However, the proof for Theorem 1.3.1 will already establish the needed bounds for both homogeneous terms. Thus, we only need to find bounds for the last two terms on the right-hand side of (1.40).

In the proof we will show how the gap condition is the proper scaling for the estimate on the regularity term to work out. This will be the primary (if small) difference between these Strichartz

17

estimates and the Strichartz estimates for the wave equation. The other being using a lemma from [73] to deal with the $\frac{-\xi_1+i\xi_2}{|\xi|}$ multipliers in the second and fourth terms.

However, the gap condition in Theorem 1.3.2 along with wave admissability reduces to $q = \tilde{q} = \infty$ and $r = \tilde{r} = 2$ which is just the $L^2$ energy inequality for the equation. Thus, on its own, Theorem 1.3.2 is essentially trivial. We include it to show how the standard wave equation Strichartz arguments apply for the Dirac equation. We also include it so that we may introduce regularity on the right-hand side of the inequality (while using only minor modifications to the proof) to widen the condition governing the allowable pairs in the estimate. In order to do so we introduce the following theorem

THEOREM 1.3.3. *Let $(q,r)$ and $(\tilde{q},\tilde{r})$ both be wave admissible. Assuming that $U$ solves Equation 1.34 with $f,g \in \dot{H}^s(\mathbb{R}^2)$ and $n,m \in L_t^{\tilde{q}'} L_x^{\tilde{r}'}(S_T)$, then*

$$(1.41) \qquad \|U\|_{L_t^q L_x^r(S_T)} \lesssim \|f\|_{\dot{H}^s} + \|g\|_{\dot{H}^s} + \|n\|_{L_t^{\tilde{q}'} \dot{W}_x^{s,\tilde{r}'}(S_T)} + \|m\|_{L_t^{\tilde{q}'} \dot{W}_x^{s,\tilde{r}'}(S_T)}$$

*and $s$ satisfies the condition $\frac{2}{r} + \frac{1}{q} = 1 - s = \frac{2}{\tilde{r}'} + \frac{1}{\tilde{q}'} - 1 - s$.*

After the proofs, we will show that Theorem 1.3.3 can be extended to the near endpoint case $q > 4, r = \infty$ by appealing to a result of Fang and Wang, [29].

**1.3.2. Proving the Strichartz Estimates.**

1.3.2.1. *Proof of Theorem 1.3.1.* For the homogeneous estimates we rely entirely on the proof provided in Lawrie, [56], with the addition of a minor modification to deal with the Fourier multiplier in the second term.

Lawrie uses the following pointwise estimate lemma

LEMMA 1.3.4. *Let $f \in \mathcal{S}$ be such that $\operatorname{supp}(\hat{f}) \subset \{\frac{1}{2} \leq |\xi| \leq 2\}$. Then,*

$$(1.42) \qquad \begin{aligned} \left\| e^{\pm it\sqrt{-\Delta}} f \right\|_{L_x^\infty} &\leq C\langle t \rangle^{-\frac{n-1}{2}} \|f\|_{L^1} \\ \left\| e^{\pm it\sqrt{-\Delta}} f \right\|_{L^2} &\leq \|f\|_{L^2}. \end{aligned}$$

Lawrie uses these estimates and the standard $TT^*$ argument to prove that

$$(1.43) \qquad \left\| e^{\pm it\sqrt{-\Delta}} f \right\|_{L_t^q L_x^r} \leq C\|f\|_{\dot{H}^s}$$

for any $f$ such that $\text{supp}(\hat{f}) \subset \{\frac{1}{2} \le |\xi| \le 2\}$. Using the restriction, $\frac{2}{r} + \frac{1}{q} = 1 - s$, he then extends the inequality to the case where $\text{supp}(\hat{f}) \subset \{2^{j-1} \le |\xi| \le 2^{j+1}\}$.

He then extends the result to $f$ without any assumption on the support of $\hat{f}$ by using a Littlewood-Paley expansion

$$(1.44) \qquad \left\| e^{\pm it\sqrt{-\Delta}} f \right\|_{L_t^q L_x^r}^2 \le \sum_j \| F_j \|_{L_t^q L_x^r}^2$$

where $F_j$ is $e^{\pm it\sqrt{-\Delta}} f$ frequency localized on the diadic intervals $\{2^{j-1} \le |\xi| \le 2^{j+1}\}$. In other words, $F_j = \hat{\psi}_j * e^{\pm it\sqrt{-\Delta}} f$ where $\{\psi_j\}_{j=1}^\infty$ is a diadic partition of unity. Since each $F_j$ is frequency localized on $\{2^{j-1} \le |\xi| \le 2^{j+1}\}$ he can apply (1.43) to get

$$(1.45) \qquad \left\| e^{\pm it\sqrt{-\Delta}} f \right\|_{L_t^q L_x^r}^2 \le C \sum_j \| \hat{\psi}_j * f \|_{\dot{H}^s}^2 .$$

Using Littlewood-Paley theory again, he shows

$$(1.46) \qquad \sum_j \| \hat{\psi}_j * f \|_{\dot{H}^s}^2 \le \| f \|_{\dot{H}^s}^2$$

and thus

$$(1.47) \qquad \left\| e^{\pm it\sqrt{-\Delta}} f \right\|_{L_t^q L_x^r} \le C \| f \|_{\dot{H}^s}$$

giving the desired bound for the first term of (1.38).

To prove the desired bound for the second term of (1.38) we recognize that we can rewrite it as

$$(1.48) \qquad \left\| e^{\pm it\sqrt{-\Delta}} \left( \mathcal{F}^{-1} \frac{-\xi_1 + i\xi_2}{|\xi|} \hat{g} \right) \right\|_{L_t^q L_x^r} .$$

Replacing $f$ in (1.47) by $\mathcal{F}^{-1} \frac{-\xi_1 + i\xi_2}{|\xi|} \hat{g}$ immediately gives us

$$(1.49) \qquad \left\| e^{\pm it\sqrt{-\Delta}} \left( \mathcal{F}^{-1} \frac{-\xi_1 + i\xi_2}{|\xi|} \hat{g} \right) \right\|_{L_t^q L_x^r} \le C \left\| \mathcal{F}^{-1} \frac{-\xi_1 + i\xi_2}{|\xi|} \hat{g} \right\|_{\dot{H}^s} .$$

Using the fact that $\left| \frac{-\xi_1 + i\xi_2}{|\xi|} \right| = 1$ and that the norm on the right-hand side is an $L^2$-based norm we see that the norm on the right-hand side is equal to $\| g \|_{\dot{H}^s}$. Thus we have the desired estimate for the second term of (1.38)

$$(1.50) \qquad \left\| e^{\pm it\sqrt{-\Delta}} \left( \mathcal{F}^{-1} \frac{-\xi_1 + i\xi_2}{|\xi|} \hat{g} \right) \right\|_{L_t^q L_x^r} \le C \| g \|_{\dot{H}^s} .$$

19

This concludes the proof for Theorem 1.3.1.

1.3.2.2. *Proof of Theorem 1.3.2.* As mentioned after the statement of the theorem, the desired bounds for the first two terms of (1.40) are established by the proof for Theorem 1.3.1. Thus we only need to prove the desired bounds for the third and fourth terms of (1.40). We can rewrite the third term as

$$(1.51) \qquad \left\| \int_0^t e^{\pm i(t-\tau)\sqrt{-\Delta}} n(\cdot,\tau)d\tau \right\|_{L_t^q L_x^r}.$$

Looking back at the proof in [**56**] we find a result stating that for any $n$ such that $\text{supp}(\hat{n}) \subset \{\frac{1}{2} \le |\xi| \le 2\}$ and wave admissable pairs $(q,r)$ and $(\tilde{q},\tilde{r})$ we have the bound

$$(1.52) \qquad \left\| \int_0^t e^{\pm i(t-\tau)\sqrt{-\Delta}} n(\cdot,\tau)d\tau \right\|_{L_t^q L_x^r} \le C\|n\|_{L_t^{\tilde{q}'} L_x^{\tilde{r}'}}.$$

Lawrie proceeds to extend this to prove the bound for the inhomogeneous term of the wave equation in the case where $\text{supp}(\hat{n}) \subset \{2^{j-1} \le |\xi| \le 2^{j+1}\}$ by using the gap condition for the wave equation. However, since the inhomogeneous term of the fundamental solution is different than that of the wave equation, his result does not immediately hold in this case. Instead, we mimic his proof on page 23 of [**56**] to get the desired result when $\text{supp}(\hat{n}) \subset \{2^{j-1} \le |\xi| \le 2^{j+1}\}$.

Remembering that $\{\psi_j\}$ is a diadic partition of unity we expand

$$(1.53) \qquad \begin{aligned} \int_0^t e^{\pm i(t-\tau)\sqrt{-\Delta}} n(\cdot,\tau)d\tau &= \sum_{j\in\mathbb{Z}} \int\int e^{i(\pm(t-\tau)|\xi|+\mathbf{x}\cdot\xi)}\psi_j(\xi)\hat{n}(\xi,\tau)d\xi d\tau \\ &:= \sum_{j\in\mathbb{Z}} K_j(n). \end{aligned}$$

Rewriting $\psi_j(\xi) = \psi_0(2^{-j}\xi)$ and performing some changes of variables we get

$$(1.54) \qquad \begin{aligned} K_j(n)(\mathbf{x},t) &= \int\int e^{i(\pm(t-\tau)|\xi|+\mathbf{x}\cdot\xi)}\psi_0(2^{-j}\xi)\hat{n}(\xi,\tau)d\xi d\tau \\ &= \int\int e^{i(\pm(2^j t-2^j\tau)|\xi|+2^j\mathbf{x}\cdot\xi)}\psi_0(\xi)2^{2j}\hat{n}(2^j\xi,\tau)d\xi d\tau \\ &= 2^{-j}\int\int e^{i(\pm(2^j t-\tau)|\xi|+2^j\mathbf{x}\cdot\xi)}\psi_0(\xi)2^{2j}\hat{n}(2^j\xi,2^{-j}\tau)d\xi d\tau \\ &= 2^{-j}\int\int e^{i(\pm(2^j t-\tau)|\xi|+2^j\mathbf{x}\cdot\xi)}\psi_0(\xi)\hat{n}_{2^{-j}}(\xi,\tau)d\xi d\tau \\ &= 2^{-j}K_0(n_{2^{-j}})(2^j\mathbf{x},2^j t) \end{aligned}$$

where $n_{2^{-j}}(\mathbf{x}, t) = n(2^{-j}\mathbf{x}, 2^{-j}t)$. In the equivalent steps for the inhomogeneous term of the wave equation there would be an additional factor, $2^{-j}$, resulting from the change of variables and the $|\xi|^{-1}$ present in the inhomogeneous term of the fundamental solution of the wave equation.

Taking the norm of $K_j(n)$, scaling the variables, and applying (1.52) we get

$$
\begin{aligned}
\|K_j(n)\|_{L_t^q L_x^r} &= 2^{j(-1-\frac{1}{q}-\frac{2}{r})} \|K_0(n_{2^{-j}})\|_{L_t^q L_x^r} \\
&\leq C 2^{j(-1-\frac{1}{q}-\frac{2}{r})} \|n_{2^{-j}}\|_{L_t^{\tilde{q}'} L_x^{\tilde{r}'}} \\
&= C 2^{j(-1-\frac{1}{q}-\frac{2}{r}+\frac{1}{\tilde{q}'}+\frac{2}{\tilde{r}'})} \|n\|_{L_t^{\tilde{q}'} L_x^{\tilde{r}'}} \\
&= C \|n\|_{L_t^{\tilde{q}'} L_x^{\tilde{r}'}} \qquad \text{(by the gap condition)}.
\end{aligned}
$$
(1.55)

This means that if we take a more general $n$ without assuption on the supp($\hat{n}$), then

$$
(1.56) \qquad \left\| \int_0^t e^{\pm i(t-\tau)\sqrt{-\Delta}} \left( \hat{\psi}_j(\cdot) * n(\cdot, \tau) \right) d\tau \right\|_{L_t^q L_x^r} \leq C \left\| \hat{\psi}_j * n \right\|_{L_t^{\tilde{q}'} L_x^{\tilde{r}'}}
$$

where have to use the gap condition.

Then we define $N_j := \int_0^t e^{\pm i(t-\tau)\sqrt{-\Delta}} \left( \hat{\psi}_j(\cdot) * n(\cdot, \tau) \right) d\tau$ to allow us to write our Littlewood-Paley expansion:

$$
(1.57) \qquad \int_0^t e^{\pm i(t-\tau)\sqrt{-\Delta}} \hat{n}(\cdot, \tau) d\tau = \sum_{j \in \mathbb{N}} N_j(\cdot, t).
$$

Taking the square of the norm of both sides and applying the Minkowski triangle inequality we get

$$
(1.58) \qquad \left\| \int_0^t e^{\pm i(t-\tau)\sqrt{-\Delta}} \hat{n}(\cdot, \tau) d\tau \right\|_{L_t^q L_x^r}^2 \leq \sum_j \|N_j(\cdot, t)\|_{L_t^q L_x^r}^2
$$

where we note that we do not require any restrictions on supp($\hat{n}$).

Then we can apply (1.56) and Littlewood-Paley again to get

$$
\begin{aligned}
\sum_j \|N_j(\cdot, t)\|_{L_t^q L_x^r}^2 &\leq C \sum_j \left\| \hat{\psi}_j * n \right\|_{L_t^{\tilde{q}'} L_x^{\tilde{r}'}}^2 \\
&\leq C \|n\|_{L_t^{\tilde{q}'} L_x^{\tilde{r}'}}^2.
\end{aligned}
$$
(1.59)

Combining (1.58) and (1.59) we get the desired result for the third term of (1.40),

$$
(1.60) \qquad \left\| \int_0^t e^{\pm i(t-\tau)\sqrt{-\Delta}} \hat{n}(\cdot, \tau) d\tau \right\|_{L_t^q L_x^r} \leq C \|n\|_{L_t^{\tilde{q}'} L_x^{\tilde{r}'}}.
$$

The fourth term of (1.40) is simple to deal with using (1.60) and a lemma from Stein, [**73**]. We notice that the fourth term of (1.40) can be written as

$$(1.61) \qquad \left\| \int_0^t e^{\pm i(t-\tau)\sqrt{-\Delta}} \mathcal{F}^{-1} \left( \frac{-\xi_1 + i\xi_2}{|\xi|} \hat{m}(\cdot, \tau) \right) d\tau \right\|_{L_t^q L_x^r}^2.$$

Applying (1.60) to the above immediately gives us

$$(1.62) \qquad \begin{aligned} &\left\| \int_0^t e^{\pm i(t-\tau)\sqrt{-\Delta}} \mathcal{F}^{-1} \left( \frac{-\xi_1 + i\xi_2}{|\xi|} \hat{m}(\cdot, \tau) \right) d\tau \right\|_{L_t^q L_x^r} \\ &\qquad \leq C \left\| \mathcal{F}^{-1} \left( \frac{-\xi_1 + i\xi_2}{|\xi|} \hat{m}(\cdot, \tau) \right) \right\|_{L_t^{\tilde{q}'} L_x^{\tilde{r}'}}. \end{aligned}$$

However, this is easily dealt with. According to Stein [**73**], if $\Lambda f := \mathcal{F}^{-1}\big(m(\xi)\hat{f}(\xi)\big)$ such that $m$ is a multiplier satisfying

$$\left| \left( \frac{\partial}{\partial \xi} \right)^\alpha m(\xi) \right| \leq C_\alpha |\xi|^{-\alpha}$$

for all multiindices $\alpha$, then $\Lambda$ mapping $L^p$ to $L^p$ for some $p$ implies $\Lambda$ also maps $L^{p^*}$ to $L^{p^*}$ for all $1 < p^* \leq p$.

A quick calculation verifies that $m(\xi) := (-\xi_1 + i\xi_2)/|\xi|$ satisfies this condition. We note that $(-\xi_1 + i\xi_2)/|\xi|$ will be invariant in $L_t^{q'}$ since it is independent of $t$. Also we see that this $\Lambda$ maps $L_x^2 \to L_x^2$ since $L^2$ is invariant under $\mathcal{F}, \mathcal{F}^{-1}$ and $|(-\xi_1 + i\xi_2)/|\xi|| = 1$. Thus we have that

$$(1.63) \qquad \left\| \mathcal{F}^{-1} \left( \frac{-\xi_1 + i\xi_2}{|\xi|} m \right) \right\|_{L_t^{\tilde{q}'} L_x^{\tilde{r}'}} \lesssim \|m\|_{L_t^{\tilde{q}'} L_x^{\tilde{r}'}}$$

since $1 < r' \leq 2$. Then combining (1.62) and (1.63) we get the desired estimate for the final term of (1.40).

Therefore we have the desired result for Theorem 1.3.2:

$$(1.64) \qquad \|U\|_{L_t^q L_x^r(S_T)} \lesssim \|f\|_{\dot{H}^s} + \|g\|_{\dot{H}^s} + \|n\|_{L_t^{\tilde{q}'} L_x^{\tilde{r}'}(S_T)} + \|m\|_{L_t^{\tilde{q}'} L_x^{\tilde{r}'}(S_T)}.$$

1.3.2.3. *Proof of Theorem 1.3.3.* The proof for this theorem is identical to the proof of Theorem 1.3.2 except for the scaling considerations necessitating the gap condition. Clearly the homogeneous terms behave identically because they are unchanged. For the inhomogeneous terms the main change is that the norm $L_x^{\tilde{r}'}$ is changed to $\dot{W}_x^{s,\tilde{r}'}$ wherever it appears. This results in an extra factor of $|\xi|^s$ on the Fourier side of each of the associated integrals.

In the case of (1.52) when $\text{supp}(\hat{n}) \subset \{\frac{1}{2} \leq |\xi| \leq 2\}$ this doesn't change anything because $|\xi| \approx 1$.

In (1.55), the second line would have an extra factor, $|\xi|^s$, on the Fourier side, and in (1.55), $|\xi| \approx 2^j$. Thus, the factor in the third line, $2^{j(-1-\frac{1}{q}-\frac{2}{r}+\frac{1}{\tilde{q}'}+\frac{2}{\tilde{r}'})}$, would become $2^{j(-1-\frac{1}{q}-\frac{2}{r}+\frac{1}{\tilde{q}'}+\frac{2}{\tilde{r}'}+s)}$. Then, the condition on $s$ in Theorem 1.3.3 would give the same result as (1.55) (except for using the $\dot{W}_x^{s,\tilde{r}'}$ norm).

The rest of the proof proceeds identically giving us the desired result for Theorem 1.3.3.

**1.3.3. Extending to $r = \infty$.** Motivated by the work of Fang and Wang [29], we will show that Theorem 1.3.3 can be extended to the Strichartz space $L_t^q L_x^\infty$ when $q > 4$. This is not trivial because the proof of the Strichartz estimates required a Littlewood-Paley decomposition, and the inequalities used in the decomposition do not hold in $L_x^\infty$. We will have to use some other technique to find a bound for the $L_x^\infty$ norm.

The main step of extending to $L_t^q L_x^\infty$ is done by the following generalized Gagliardo-Nirenberg inequality

LEMMA 1.3.5. *Let $a, c \in (1, \infty), \kappa, \mu \in (0, n)$ and $\kappa a < n < \mu c$, then*

$$\|f\|_{L^\infty} \lesssim \|D^\kappa f\|_{L^a}^\theta \|D^\mu f\|_{L^c}^{1-\theta}$$

*where $\theta = \left(\frac{\mu}{n} - \frac{1}{c}\right) / \left(\frac{\mu}{n} - \frac{1}{c} + \frac{1}{a} - \frac{\kappa}{n}\right)$.*

The lemma is proven by Escobedo and Vega in [27].

To apply this to our desired Strichartz estimates $n = 2$ and we will let $a = 2, \kappa = s = 1 - \frac{1}{q}$, and $\mu = \frac{3}{2c} + \frac{1}{4} - \frac{1}{q}$. Notice that as long as $c$ is large then $n < \mu c$ (this would not be possible if $q = 4$). Then we choose $c$ large enough so that the pair $((1-\theta)q, c)$ is admissable.

We will define a new operator, $\mathcal{T}$, by

(1.65) $$\mathcal{T}f(t,x) = \int e^{ix\cdot\xi + it|\xi|} \hat{f}(\xi) d\xi = \mathcal{F}^{-1}\left(e^{it|\xi|}\hat{f}(\xi)\right).$$

For ease of notation we will represent $\mathcal{T}f = e^{itD}f$. When combined with the Strichartz estimates already established this will allow us to find a useful bound.

We recognize that the first term in (1.40) is $\mathcal{T}f$, and

$$\|\mathcal{T}f\|_{L_t^q L_x^\infty} \lesssim \left( \|D^s e^{itD} f\|_{L_x^2}^\theta \|D^\mu e^{itD} f\|_{L_x^c}^{1-\theta} \right)_{L_t^q}$$

(1.66)

$$\lesssim \|e^{itD} D^s f\|_{L_t^\infty L_x^2}^\theta \|e^{itD} D^\mu f\|_{L_t^{q(1-\theta)} L_x^c}^{1-\theta}$$

$$\lesssim \|D^s f\|_{L^2}^\theta \|D^\mu f\|_{\dot{H}^{s-\mu}}^{1-\theta}$$

$$\lesssim \|f\|_{\dot{H}^s}$$

where the third line comes from the already established Strichartz estimates for wave admissable pairs.

By an easy calculation we can determine that $\mathcal{T}^* F = \int \mathcal{F}^{-1} \left( e^{-i\tau|\xi|} \hat{F}(t,\xi) \right) d\tau$ and $\mathcal{T}\mathcal{T}^* F = \int \mathcal{F}^{-1} \left( e^{i(t-\tau)|\xi|} \hat{F}(t,\xi) \right) d\tau$. In other words, the third term from (1.40) is $\mathcal{T}\mathcal{T}^* n$ with the time integration restricted to $[0,T]$.

A similar calculation to (1.66) shows that

(1.67)
$$\|\mathcal{T}\mathcal{T}^* n\|_{L_t^q L_x^\infty} \lesssim \|n\|_{L_t^{\tilde{q}'} \dot{W}_x^{s,\tilde{r}'}}$$

where the gap condition and the previously established inhomogeneous Strichartz estimates are used to get the third line just as was done in (1.66).

Then applying the Christ-Kiselev lemma, [**24**], gives the desired bound for the third term in (1.40):

(1.68)
$$\left\| \int_0^t e^{\pm i(t-\tau)\sqrt{-\Delta}} \hat{n}(\cdot,\tau) d\tau \right\|_{L_t^q L_x^\infty} \leq C \|n\|_{L_t^{\tilde{q}'} \dot{W}_x^{s,\tilde{r}'}} .$$

This just leaves the second and fourth term in (1.40) which have the fraction multiplier in them. However, the multiplier can be handled exactly as it was in the proof of the above Strichartz estimates. For the second term we eliminate the $\frac{-\xi_1 + i\xi_2}{|\xi|}$ factor using the $L^2$ invariance of $\mathcal{F}$ and $\mathcal{F}^{-1}$. In the fourth term we eliminate the $\frac{-\xi_1 + i\xi_2}{|\xi|}$ by using the lemma from [**73**].

Thus we are able to extend the Strichartz estimates in Theorem 1.3.3 from only wave admissable pairs $(q,r)$ with $r \neq \infty$ to allowing $r = \infty$ as long as $q > 4$.

**1.3.4. Conclusion of Strichartz Estimates.** Thus by extending the allowable range of $q, r$ in Theorem 1.3.1 we have established the Strichartz estimates

(1.69)
$$\|U\|_{L_t^q L_x^r(S_T)} \lesssim \|f\|_{\dot{H}^s} + \|g\|_{\dot{H}^s} + \|n\|_{L_t^{\tilde{q}'} \dot{W}_x^{s,\tilde{r}'}(S_T)} + \|m\|_{L_t^{\tilde{q}'} \dot{W}_x^{s,\tilde{r}'}(S_T)}$$

24

for $4 < q, \tilde{q} \leq \infty$, $2 \leq r \leq \infty$, $2 \leq \tilde{r} < \infty$ with $2/q \leq 1/2 - 1/r$, $2/\tilde{q} \leq 1/2 - 1/\tilde{r}$, and the condition,
$\frac{2}{r} + \frac{1}{q} = 1 - s = \frac{2}{\tilde{r}'} + \frac{1}{\tilde{q}'} - 1 - s$.

In our application of these Strichartz estimates to local existence in Chapter 2 we will make the choices $r = \infty$, $\tilde{q} = \infty$, and $\tilde{r} = 2$. Thus we can use

(1.70) $$\|U\|_{L_t^q L_x^\infty(S_T)} \lesssim \|f\|_{\dot{H}^s} + \|g\|_{\dot{H}^s} + \|n\|_{L_t^1 \dot{H}_x^s(S_T)} + \|m\|_{L_t^1 \dot{H}_x^s(S_T)}$$

where $q > 4$ can be as close to 4 as desired and $s = 1 - \frac{1}{q}$.

# Nonlinear Results

## 2.1. Remarks on Conservation Laws, Scaling, and Criticality.

Before exploring more advanced results for the nonlinear Dirac equation we first look at a few quick results for the nonlinear equation.

### 2.1.1. Nonlinear Conservation Laws.

In this section we will use multiplier methods to establish some basic conservation laws for the system $PU = CF(U)$ where we let $C = 0$ denote the linear homogeneous case and $C = 1$ denote the nonlinear case. Note, we only show this for Dirac equation derived by Ablowitz, Nixon, and Zhu, [**1**], as Fefferman and Weinstein establish conserved quantities for the equation they derived in [**31**]. We also assume $U$ is a solution in a functions space with good decay at infinity such as $\mathcal{S}(\mathbb{R}^2), L^2(\mathbb{R}^2)$, or $H^s(\mathbb{R}^2)$.

These laws will be Conservation of Mass/Charge

$$(2.1) \qquad \int |u|^2 + |v|^2 dxdy,$$

Conservation of Hamiltonian

$$(2.2) \qquad \int 2\Re(\bar{v}u_x) - 2\Im(\bar{v}u_y) - \frac{C}{2}\left(|u|^4 + |v|^4\right) dxdy,$$

and Conservation of Momentum,

$$(2.3) \qquad \int \Im(\bar{u}u_x + \bar{v}v_x)dxdy.$$

2.1.1.1. *Finding Conservation of Mass.* To find the Law for Conservation of Mass (or Charge) consider the product

$$\begin{pmatrix} \bar{u} \\ \bar{v} \end{pmatrix} \cdot \begin{pmatrix} iu_t - (v_x - iv_y) \\ iv_t + (u_x + iu_y) \end{pmatrix} = C \begin{pmatrix} \bar{u} \\ \bar{v} \end{pmatrix} \cdot \begin{pmatrix} |u|^2 u \\ |v|^2 v \end{pmatrix}.$$

Expanding the dot product we get:

$$C(|u|^4 + |v|^4) = i\bar{u}u_t - \bar{u}v_x + i\bar{u}v_y + i\bar{v}v_t + \bar{v}u_x + i\bar{v}u_y$$

(2.4)

$$iC(|u|^4 + |v|^4) + \bar{u}u_t + \bar{v}v_t = i(\bar{v}u_x - \bar{u}v_x) - (\bar{u}v_y + \bar{v}u_y).$$

Taking the Real Component of both sides and applying the product rule we get

(2.5)
$$\Re(\bar{u}u_t + \bar{v}v_t) = -\Im(\partial_x(\bar{v}u)) - \Re(\partial_y(\bar{u}v)).$$

Notice the contribution of the nonlinearity vanishes in this step since $|u|^4 + |v|^4$ is purely real.

We also note the trivial identity $\partial_t(|u|^2 + |v|^2) = 2\Re(\bar{u}u_t + \bar{v}v_t)$.

Thus we get,

$$\partial_t \frac{1}{2} \int |u|^2 + |v|^2 dxdy = \int -\Im(\partial_x(\bar{v}u)) - \Re(\partial_y(\bar{u}v)) dxdy$$

(2.6)
$$= -\int \partial_x(\Im(\bar{v}u)) dxdy - \int \partial_y(\Re(\bar{u}v)) dydx$$

$$= 0$$

where the final equality is achieved by the Fundamental Theorem of Calculus and decay assumptions of solutions $u, v$ at infinity. This proves the Conservation of Mass stated above.

It is interesting to note that since the contribution from the nonlinearity vanishes during the proof this is the same conservation law for the purely linear problem.

2.1.1.2. *Finding Conservation of Hamiltonian.* To find the Law for Conservation of Hamiltonian consider the following

$$\int \begin{pmatrix} \bar{u}_t \\ \bar{v}_t \end{pmatrix} \cdot \begin{pmatrix} iu_t - (v_x - iv_y) \\ iv_t + (u_x + iu_y) \end{pmatrix} dxdy = C \int \begin{pmatrix} \bar{u}_t \\ \bar{v}_t \end{pmatrix} \cdot \begin{pmatrix} |u|^2 u \\ |v|^2 v \end{pmatrix}.$$

Expanding the dot product of the left side we get

(2.7)
$$\int i|u_t|^2 - \bar{u}_t v_x + i\bar{u}_t v_y + i|v_t|^2 + \bar{v}_t u_x + i\bar{v}_t u_y dxdy.$$

Adding the above equation to its complex conjugate gives

$$\int -(\bar{u}_t v_x + u_t \bar{v}_x) + i(\bar{u}_t v_y - u_t \bar{v}_y) + (\bar{v}_t u_x + v_t \bar{u}_x) + i(\bar{v}_t u_y - v_t \bar{u}_y) dxdy$$

$$= \int -2\Re(\bar{u}_t v_x) - 2\Im(\bar{u}_t v_y) + 2\Re(\bar{v}_t u_x) - 2\Im(\bar{v}_t u_y) dxdy$$

(2.8)

$$= \int 2\Re(\bar{v}_t u_x + u_{xt}\bar{v}) - 2\Im(\bar{v}_t u_y + u_{yt}\bar{v}) dxdy$$

$$= \partial_t \int 2\Re(\bar{v}u_x) - 2\Im(\bar{v}u_y) dxdy.$$

Doing the same with the right side of the equation gives us

$$C \int |u|^2 u\bar{u}_t + |u|^2 \bar{u}u_t + |v|^2 v\bar{v}_t + |v|^2 \bar{v}v_t dxdy$$

(2.9)

$$= \frac{C}{2}\partial_t \int |u|^4 + |v|^4 dxdy.$$

Putting the two sides back together we get

$$\partial_t \int 2\Re(\bar{v}u_x) - 2\Im(\bar{v}u_y) dxdy = \frac{C}{2}\partial_t \int |u|^4 + |v|^4 dxdy$$

(2.10)

$$0 = \partial_t \left( \int 2\Re(\bar{v}u_x) - 2\Im(\bar{v}u_y) - \frac{C}{2}\left(|u|^4 + |v|^4\right) dxdy \right)$$

yielding the desired Conservation of Hamiltonian.

2.1.1.3. *Finding Conservation of Momentum.* To find the Law for Conservation of Momentum we define the operators $\partial_- = \partial_x - i\partial_y$ and $\partial_+ = \partial_x + i\partial_y$. Then we consider the following

$$\int \begin{pmatrix} \partial_-^* \bar{u} \\ \partial_+^* \bar{v} \end{pmatrix} \cdot \begin{pmatrix} iu_t - (v_x - iv_y) \\ iv_t + (u_x + iu_y) \end{pmatrix} dxdy = C \int \begin{pmatrix} \partial_-^* \bar{u} \\ \partial_+^* \bar{v} \end{pmatrix} \cdot \begin{pmatrix} |u|^2 u \\ |v|^2 v \end{pmatrix} dxdy.$$

A quick calculation verifies that $\partial_-^* = -\partial_+$ and $\partial_+^* = -\partial_-$. By expanding the dot product we get:

$$C \int -u^2\bar{u}\bar{u}_x - iu^2\bar{u}\bar{u}_y - v^2\bar{v}\bar{v}_x + iv^2\bar{v}\bar{v}_y dxdy$$

(2.11)

$$= \int (-i\bar{u}_x u_t + \bar{u}_x v_x - i\bar{u}_x v_y + \bar{u}_y v_t + i\bar{u}_y v_x + \bar{u}_y v_y$$

$$- i\bar{v}_x v_t - \bar{v}_x u_x - i\bar{v}_x u_y - \bar{v}_y v_t + i\bar{v}_y u_x - \bar{v}_y u_y) dxdy.$$

Doing the same with

$$\int \begin{pmatrix} \partial_-^* u \\ \partial_+^* v \end{pmatrix} \cdot \begin{pmatrix} -i\bar{u}_t - (\bar{v}_x + i\bar{v}_y) \\ -i\bar{v}_t + (\bar{u}_x - i\bar{u}_y) \end{pmatrix} dxdy = C \int \begin{pmatrix} \partial_-^* u \\ \partial_+^* v \end{pmatrix} \cdot \begin{pmatrix} |u|^2 \bar{u} \\ |v|^2 \bar{v} \end{pmatrix} dxdy$$

28

gives us

$$C \int -\bar{u}^2 u u_x - i\bar{u}^2 u u_y - \bar{v}^2 v v_x + i\bar{v}^2 v v_y dx dy$$

(2.12)
$$= \int (iu_x \bar{u}_t + u_x \bar{v}_x + iu_x \bar{v}_y - u_y \bar{v}_t + iu_y \bar{v}_x - u_y \bar{v}_y$$

$$+ iv_x \bar{v}_t - v_x \bar{u}_x + iv_x \bar{u}_y + v_y \bar{v}_t + iv_y \bar{u}_x + v_y \bar{u}_y) dx dy.$$

Adding (2.11) and (2.12) as well as the conjugate of both equations we get

$$-C \int 2(u^2 \bar{u} \bar{u}_x + \bar{u}^2 u u_x) + 2(v^2 \bar{v} \bar{v}_x + \bar{v}^2 v v_x) dx dy$$

(2.13)
$$= 2i \int (\bar{u}_t u_x - u_t \bar{u}_x) + (\bar{v}_t v_x - v_t \bar{v}_x) + (\bar{u}_y v_x - v_y \bar{u}_x) + (\bar{v}_y u_x - u_y \bar{v}_x)$$

$$= 2i \int i\Im(\bar{u}_t u_x + \bar{u}_t u_x) + i\Im(\bar{v}_t v_x + \bar{v}_t v_x)$$

$$+ (\bar{u}_y v_x - v_y \bar{u}_x) + (\bar{v}_y u_x - u_y \bar{v}_x) dx dy.$$

We note here that $\partial_x(u^2 \bar{u}^2) = 2(u^2 \bar{u} \bar{u}_x + \bar{u}^2 u u_x)$, and likewise for $v$. Thus the integral on the left-hand side of the equation vanishes by the Fundamental Theorem of Calculus. On the left-hand side, integrating by parts selectively and taking a complex conjugate inside each $\Im$ then yields

$$0 = 2i \int i\Im(\bar{u}_t u_x + u_{xt} \bar{u}) + i\Im(\bar{v}_t v_x + v_{xt} \bar{v})$$

$$+ (\bar{u}_y v_x + v_{xy} \bar{u}) + (\bar{v}_y u_x + u_{xy} \bar{v}) dx dy$$

(2.14)
$$= 2i \int i\partial_t \Im(\bar{u} u_x) + i\partial_t \Im(\bar{v} v_x) + \partial_y(\bar{u} v_x) + \partial_y(\bar{v} u_x) dx dy$$

$$= -2\partial_t \int \Im(\bar{u} u_x + \bar{v} v_x) dx dy$$

which implies the desired Conservation of Momentum.

**2.1.2. Scaling of Solutions to Nonlinear Equation.** We will also take a brief look at scaling solutions to the equation to determine an allowable scaling and investigate the effect that scaling has on the $L^2$ norm.

2.1.2.1. *Finding an allowable scaling.* Assuming $U = (u, v)$ is a solution we define $\tilde{U} = \lambda^\alpha U(\lambda^{\beta_1} t, \lambda^{\beta_2} x, \lambda^{\beta_3} y)$. Thus, if $\tilde{U}$ is also a solution we have

(2.15)
$$i\lambda^{\alpha+\beta_1} u_t(\#) - \lambda^{\alpha+\beta_2} v_x(\#) + i\lambda^{\alpha+\beta_3} v_y(\#) = \lambda^{3\alpha} |u(\#)|^2 u(\#)$$

where $\#$ denotes $(\lambda^{\beta_1} t, \lambda^{\beta_2} x, \lambda^{\beta_3} y)$.

However, since $u$ is a solution this becomes

$$\begin{aligned}(2.16) \quad i\lambda^{\alpha+\beta_1}u_t(\#) - \lambda^{\alpha+\beta_2}v_x(\#) + i\lambda^{\alpha+\beta_3}v_y(\#) &= \lambda^{3\alpha}\big(iu_t(\#) - v_x(\#) + iv_y(\#)\big) \\ &= \lambda^{3\alpha}iu_t(\#) - \lambda^{3\alpha}v_x(\#) + i\lambda^{3\alpha}v_y(\#).\end{aligned}$$

Thus we determine that $\beta_1 = \beta_2 = \beta_3 = 2\alpha$. Therefore, we will set $\alpha = \frac{1}{2}$ and work with the allowable scaling

$$\tilde{U} = \lambda^{\frac{1}{2}}U(\lambda t, \lambda x, \lambda y).$$

2.1.2.2. *How Scaling Affects the $L_t^\infty L_\mathbf{x}^2$ Norm of Solutions.* We start with $\tilde{U}$ as determined in the above section, and we consider its $L^2$ norm. However for the sake of simplicity we will actually just consider the $L^2$ norm of $\tilde{u}$.

$$\begin{aligned}(2.17) \quad \|\tilde{u}\|_{L_t^\infty L_\mathbf{x}^2} &= \sup_t \left(\int |\tilde{u}|^2\right)^{\frac{1}{2}} \\ &= \lambda^{\frac{1}{2}} \sup_t \left(\int\int |u(\lambda t, \lambda x, \lambda y)|^2 dx dy\right)^{\frac{1}{2}} \\ &= \lambda^{\frac{1}{2}} \sup_t \left(\int\int \lambda^{-2}|u(\lambda t, q, r)|^2 dq dr\right)^{\frac{1}{2}} \\ &= \lambda^{-\frac{1}{2}} \|u\|_{L_t^\infty L_\mathbf{x}^2}.\end{aligned}$$

An identical argument gives the same result for $v$. Thus it must also be true that $\|\tilde{U}\|_{L^2} = \lambda^{-\frac{1}{2}}\|U\|_{L^2}$.

Thus, since scaling a solution $U$ with parameter $\lambda$ results in the $L^2$ norm being multiplied by a negative power of $\lambda$ we can conclude that the cubic nonlinearity is mass supercritical.

Note that the same scaling works for the Dirac nonlinearity found by Fefferman and Weinstein, [**31**]. The only difference in the proof is there are two nonlinear terms instead of one.

**2.2. Local Existence.** In this section we will establish some local existence results for the cubic Dirac equation. The first being in a simple energy space using an energy estimate, and the second using a Strichartz space to allow us to use the Strichartz estimates to bootstrap the result into a longer existence time. For both of these results we write out the proof for the Ablowitz, Nixon, and Zhu, [**1**], Dirac equation as the nonlinearity is notationally easier to work with. However, the proof for the Fefferman and Weinstein, [**31**], Dirac equation differs only in the fact that the nonlinearity has two cubic terms which will have factors of both $u$ and $v$, not necessarily only one of them.

30

As the fundamental solution of the Dirac equation, (0.1), is very similar to that for the wave equation (and in fact reduces to linear combinations of the half-wave operator), we use standard wave equation techniques to prove local existence for the Dirac equation. In particular, the methods we use for these proofs are motivated by Lindblad and Sogge, [**59**], as well as Sogge, [**72**]. One can find local existence for even quasilinear wave equations dating back to Hughes, Kato, and Marsden, [**41**], and Kato, [**49**].

**2.2.1. An Energy Existence Result.** In this section we will establish some energy estimates and an existence result in the space $X_E = C([0,T], H^s(\mathbb{R}^2))$ where $s > 1$. In other words the case when $s > n/2$, and therefore $H^s$ is an algebra.

We claim the following energy estimate

THEOREM 2.2.1. *Given $U$ a solution to the linear Dirac equation with initial conditions $f = u(0,\cdot)$, $g = v(0,\cdot)$ and inhomogeneities $F_1, F_2$ we have the energy estimates*

$$(2.18) \qquad \|U(t,\cdot)\|_{H^s} \le E_s + \int_0^t \|F_1(\tau)\|_{H^s} + \|F_2(\tau)\|_{H^s} d\tau$$

*and also*

$$(2.19) \qquad \|U\|_{X_E} \le E_s + T(\|F_1\|_{X_E} + \|F_2\|_{X_E})$$

*where $s > 1$ and $E_s = \|f\|_{H^s} + \|g\|_{H^s}$*

PROOF. It suffices to prove the estimate for $u$ instead of $U$. Consider $\|\mathcal{F}^{-1}\cos(t|\xi|)\hat{f}\|_{H^s}$.

We start by rewriting as an $L^2$ norm

$$\left\|\mathcal{F}^{-1}\cos(t|\xi|)\hat{f}\right\|_{H^s} = \left\|\mathcal{F}^{-1}(1+|\xi|^2)^{s/2}\cos(t|\xi|)\hat{f}\right\|_{L^2}$$

$$= \left\|(1+|\xi|^2)^{s/2}\cos(t|\xi|)\hat{f}\right\|_{L^2}$$

$$(2.20) \qquad \le \left\|(1+|\xi|^2)^{s/2}\hat{f}\right\|_{L^2}$$

$$= \left\|\mathcal{F}^{-1}(1+|\xi|^2)^{s/2}\hat{f}\right\|_{L^2}$$

$$= \|f\|_{H^s}.$$

Notice that the sine term works the exact same way since $\left|\frac{-\xi_1+i\xi_2}{|\xi|}\right| = 1$ and the inhomogeneous terms work out the same way because $\|\int(\cdot)d\tau\|_{H^s} \le \int\|\cdot\|_{H^s}d\tau$. This proves (2.18).

To prove (2.19) we notice that the norm on $C[0,T]$ is the sup norm. Thus taking the sup of $t$ over $[0,T]$ in (2.18) yields (2.19). □

Now that we have an energy estimate we will define our Picard Iteration. We start by defining $U_{-1} = 0$. Then iteratively we will define

$$PU_{j+1} = \begin{pmatrix} |u_j|^2 u_j \\ |v_j|^2 v_j \end{pmatrix}$$

where $P$ is the operator in our Dirac Equation (0.2).

First we need to show that this sequence is well defined in $X_E$. Trivially $U_{-1} \in X_E$, and we will assume that $U_j \in X_E$ by induction. Then applying the energy estimate to $U_{j+1}$ we get

$$\|U_{j+1}(t,\cdot)\|_{H^s} \le E_s + \int_0^t \left\||u_j(\tau,\cdot)|^2 u_j(\tau,\cdot)\right\|_{H^s} + \left\||v_j(\tau,\cdot)|^2 v_j(\tau,\cdot)\right\|_{H^s} d\tau.$$

However, for $s > \frac{n}{2}$ we have $H^s \subset L^\infty$ and therefore $H^s$ is an algebra. Thus it is closed under multiplication and we get

$$\|U_{j+1}(t,\cdot)\|_{H^s} \le E_s + \int_0^t C\|u_j(\tau,\cdot)\|_{H^s}^3 + \|v_j(\tau,\cdot)\|_{H^s}^3 d\tau$$

$$\le E_s + \int_0^t C\|U_j(\tau,\cdot)\|_{H^s}^3$$

where the constant, $C$, comes from the Sobolev embedding $H^s(\mathbb{R}^2) \subset L^\infty(\mathbb{R}^2)$ for $s > 1$.

Taking the supremum in time results in

(2.21) $$\|U_{j+1}\|_{X_E} \le C\left(E_s + T\|U_j\|_{X_E}^3\right).$$

Thus the sequence is well defined in $X_E$.

To show that the sequence converges we will prove that it is uniformly bounded and Cauchy.

We define $R = 2CE_s$ and assume inductively that $\|U_j\|_{X_E} \le R$. Now we want to show $\|U_{j+1}\|_{X_E} \le R$. Plugging in the bound for $U_j$ into (2.21) we get

(2.22)
$$\|U_{j+1}\|_{X_E} \le C\left(E_s + TR^3\right)$$
$$= C\left(E_s + T(8C^3 E_s^3)\right).$$

Now we make our local in time assumption and restrict $T < (16C^3 E_s^2)^{-1}$. This gives us

$$\|U_{j+1}\|_{X_E} \leq C \left( E_s + \frac{1}{2} E_s \right)$$

(2.23)

$$\leq R.$$

Thus the sequence is uniformly bounded by $R$.

Now, to show the sequence is Cauchy we will notice that $U_{j+1} - U_j$ solves the equation with zero initial conditions and inhomogeneities $|u_j|^2 u_j - |u_{j-1}|^2 u_{j-1}$ and $|v_j|^2 v_j - |v_{j-1}|^2 v_{j-1}$. Thus we can apply the energy inequality

$$\|U_{j+1} - U_j\|_{H^s} \leq \int_0^t \left\| |u_j|^2 u_j - |u_{j-1}|^2 u_{j-1} \right\|_{H^s} + \left\| |v_j|^2 v_j - |v_{j-1}|^2 v_{j-1} \right\|_{H^s} d\tau$$

$$\leq C \int_0^t \|u_j - u_{j-1}\|_{H^s} \left( \|u_j\|_{H^s}^2 + \|u_{j-1}\|_{H^s}^2 \right)$$

$$+ \|v_j - v_{j-1}\|_{H^s} \left( \|v_j\|_{H^s}^2 + \|v_{j-1}\|_{H^s}^2 \right) d\tau$$

(2.24)

$$\leq C \int_0^t \left( \|u_j - u_{j-1}\|_{H^s} + \|v_j - v_{j-1}\|_{H^s} \right) \left( \|U_j\|_{H^s}^2 + \|U_{j-1}\|_{H^s}^2 \right) d\tau$$

$$\leq 2CR^2 \int_0^t \|u_j - u_{j-1}\|_{H^s} + \|v_j - v_{j-1}\|_{H^s} d\tau$$

$$= 8C^3 E_s^2 \int_0^t \|u_j - u_{j-1}\|_{H^s} + \|v_j - v_{j-1}\|_{H^s} d\tau.$$

Taking the supremum in $t$ gives

$$\|U_{j+1} - U_j\|_{X_E} \leq 8C^3 E_s^2 T \left( \|u_j - u_{j-1}\|_{X_E} + \|v_j - v_{j-1}\|_{X_E} \right)$$

(2.25)

$$\leq \frac{1}{2} \left( \|u_j - u_{j-1}\|_{X_E} + \|v_j - v_{j-1}\|_{X_E} \right)$$

$$= \frac{1}{2} \|U_j - U_{j-1}\|_{X_E}.$$

Thus the sequence $(u_j)_{j=-1}^{\infty}$ is Cauchy and uniformly bounded. Therefore the sequence converges in $X_E$.

Now we claim that $|u_j|^2 u_j \to |u|^2 u$. To see this observe

$$\left\| |u_j|^2 u_j - |u|^2 u \right\|_{H^s} \leq C \|u_j - u\|_{H^s} \left( \|u_j\|_{H^s}^2 + \|u\|_{H^s}^2 \right)$$

(2.26)

$$\leq 2CR^2 \|u_j - u\|_{H^s}$$

33

and thus

$$(2.27) \qquad \left\| |u_j|^2 u_j - |u|^2 u \right\|_{X_E} \le 2CR^2 \|u_j - u\|_{X_E}.$$

Since $R$ is a uniform constant the right side must converge to zero. Thus $|u_j|^2 u_j \to |u|^2 u$. Likewise we see that $|v_j|^2 v_j \to |v|^2 v$.

Thus we conclude that $U_j \to U$ and $F(U_j) \to F(U)$. However, since $U_j \to U$ in $X_E$ it also converges distributionally and we conclude that $PU_j \to PU$. But $PU_j = F(U_{j-1}) \to F(U)$. By uniqueness of the limit we conclude that $PU = F(U)$. Thus $U$ is a local solution in $X_E$, and the time scale has order of $\epsilon^{-2}$ where $\epsilon$ is the size of the initial conditions since our only assumption on time was $T < (16C^3 E_s^2)^{-1}$.

We also want to show that the solution, $U$, depends continuously on the initial data. To do that, consider $U$ a solution to (0.1) with initial data $u(0) = f$ and $v(0) = g$. Likewise, consider $U^*$ a solution to (0.1) with initial data $u_0^* = f^*$ and $v_0^* = g^*$. Then $U - U^*$ solves the same Dirac equation except with nonlinearities $|u|^2 u - |u^*|^2 u^*$ and $|v|^2 v - |v^*|^2 v^*$ and initial data $f - f^*$ and $g - g^*$.

Applying the energy estimate we get

$$(2.28) \qquad \begin{aligned} \|U - U^*\|_{X_E} \le \|f - f^*\|_{H^s} &+ \|g - g^*\|_{H^s} + \int_0^t \big( \left\| |u|^2 u - |u^*|^2 u^* \right\|_{H^s} \\ &+ \left\| |v|^2 v - |v^*|^2 v^* \right\|_{H^s} \big) d\tau. \end{aligned}$$

Proceeding similarly to (2.24) and (2.25) we get that

$$(2.29) \qquad \int_0^t \big( \left\| |u|^2 u - |u^*|^2 u^* \right\|_{H^s} + \left\| |v|^2 v - |v^*|^2 v^* \right\|_{H^s} \big) d\tau \le \frac{1}{2} \|U - U^*\|_{X_E}$$

by expanding the difference of cubes, applying the uniform bound, and applying the time bound.

Thus we have

$$(2.30) \qquad \begin{aligned} \|U - U^*\|_{X_E} &\le \|f - f^*\|_{H^s} + \|g - g^*\|_{H^s} + \frac{1}{2} \|U - U^*\|_{X_E} \\ \frac{1}{2} \|U - U^*\|_{X_E} &\le \|f - f^*\|_{H^s} + \|g - g^*\|_{H^s}. \end{aligned}$$

This proves that the solutions depend continuously on the initial data.

It is also easy to show uniqueness of solutions using the argument above as we did for global existence of the linear problem in Chapter 1.

We include this elementary result even though the next subsection proves a stronger local existence result because in Chapter 3 we only need to use existence of the Dirac equation in $H^s$ spaces with $s > 1$.

**2.2.2. Improvement Using Bootstrap Space.** Now we will prove a local existence result for a lower regularity space ($H^1$) with an extended time scale (approaching $\epsilon^{-4}$) by using a bootstrap space instead of just energy. We will be using the space $X = C_t H_x^1 \cap L_t^q L_x^\infty$ where $q > 4$. Restricting $q > 4$ is necessary for our method as it will allow us to use the Strichartz estimates, (1.70), with the $L^\infty$ norm in space.

We use the same Picard Iteration as before with $f$ and $g$ the initial conditions.

First we need to show that this sequence is well defined in $X$. Trivially $U_{-1} \in X$, and we will assume that $U_j \in X$ by induction. We start by expanding out

(2.31)
$$\|U_{j+1}\|_X \leq \|U_{j+1}\|_{L^\infty H^1} + \|U_{j+1}\|_{L^q L^\infty}$$
$$= \|U_{j+1}\|_{L^\infty L^2} + \|DU_{j+1}\|_{L^\infty L^2} + \|U_{j+1}\|_{L^q L^\infty}.$$

Then we notice that $L^\infty L^2$ and $L^q L^\infty$ are allowable norms for the Strichartz estimates. For the pair $(\infty, 2)$ we can just use the Strichartz estimate from Theorem 1.3.2 with $q = \tilde{q} = \infty$, $r = \tilde{r} = 2$, and $s = 0$. For the pair $(q, \infty)$ we use Theorem 1.3.3 with $r = \infty, \tilde{q} = \infty, \tilde{r} = 2$, and $s = 1 - \frac{1}{q}$ where $q > 4$ is close to 4. Notice we will freely use the fact that $H^s \subset \dot{H}^s$ and $H^s \subset H^t$ for $s > t$ to move from homogeneous Sobolev spaces to inhomogeneous Sobolev spaces and to larger inhomogeneous Sobolev spaces.

35

Since $s$, the regularity required by the Strichartz pairs, is less than 1 for each of these pairs, letting $E_1$ be the $H^1$ norm of the initial conditions, we get

$$\|U_{j+1}\|_X \lesssim E_1 + \big\||u_j|^2 u_j\big\|_{L^1 L^2} + \big\|D\left(|u_j|^2 u_j\right)\big\|_{L^1 L^2} + \big\||u_j|^2 u_j\big\|_{L^1 \dot{H}^{1-\frac{1}{q}}}$$

$$+ \big\||v_j|^2 v_j\big\|_{L^1 L^2} + \big\|D\left(|v_j|^2 v_j\right)\big\|_{L^1 L^2} + \big\||v_j|^2 v_j\big\|_{L^1 \dot{H}^{1-\frac{1}{q}}}$$

$$\lesssim E_1 + \|u_j\|_{L^4 L^\infty}^2 \|u_j\|_{L^2 L^2} + \|u_j\|_{L^4 L^\infty}^2 \|Du_j\|_{L^2 L^2}$$

$$+ \|u_j\|_{L^4 L^\infty}^2 \|D^{1-\frac{1}{q}} u_j\|_{L^2 L^2}$$

$$+ \|v_j\|_{L^4 L^\infty}^2 \|v_j\|_{L^2 L^2} + \|v_j\|_{L^4 L^\infty}^2 \|Dv_j\|_{L^2 L^2}$$

(2.32)
$$+ \|v_j\|_{L^4 L^\infty}^2 \|D^{1-\frac{1}{q}} v_j\|_{L^2 L^2}$$

$$\lesssim E_1 + \|u_j\|_{L^4 L^\infty}^2 \|u_j\|_{L^2 H^1} + \|v_j\|_{L^4 L^\infty}^2 \|v_j\|_{L^2 H^1}$$

$$\lesssim E_1 + \left(T^{\frac{q-4}{4q}} \|u_j\|_{L^q L^\infty}\right)^2 T^{\frac{1}{2}} \|u_j\|_{L^\infty H^1}$$

$$+ \left(T^{\frac{q-4}{4q}} \|v_j\|_{L^q L^\infty}\right)^2 T^{\frac{1}{2}} \|v_j\|_{L^\infty H^1}$$

$$= E_1 + T^{\frac{q-2}{q}} \left(\|u_j\|_{L^q L^\infty}^2 \|u_j\|_{L^\infty H^1} + \|v_j\|_{L^q L^\infty}^2 \|v_j\|_{L^\infty H^1}\right)$$

$$\lesssim E_1 + T^{\frac{q-2}{q}} \|U_j\|_X^3$$

where the second inequality requires Proposition 3.1 in Christ and Weinstein, [**23**], to deal with the fractional derivatives. The proposition states for $\alpha \in (0,1)$

(2.33)
$$\|D^\alpha F(u)\|_{L^r} \leq C\|F'(u)\|_{L^p}\|D^\alpha u\|_{L^q} \quad \text{for } r^{-1} = p^{-1} + q^{-1}.$$

Thus we conclude

(2.34)
$$\|U_{j+1}\|_X \leq C\left(E_1 + T^{\frac{q-2}{q}} \|U_j\|_X^3\right),$$

and the sequence is well defined. To show that the sequence converges we will prove that it is uniformly bounded and Cauchy.

We define $R = 2CE_1$ and assume inductively that $\|U_j\|_X \leq R$. Once again $u_{-1} = 0$ and trivially has this bound. Now we want to show $\|U_{j+1}\|_X \leq R$. Making our local in time assumption,

$T^{\frac{q-2}{q}} < (16C^3 E_1^2)^{-1}$, and plugging in the bound for $U_j$ into (2.34) we get

$$\|U_{j+1}\|_X \leq C\left(E_1 + T^{\frac{q-2}{q}} R^3\right)$$

$$= C\left(E_1 + T^{\frac{q-2}{q}}(8C^3 E_1^3)\right)$$

(2.35)

$$\leq C\left(E_1 + \frac{1}{2}E_1\right)$$

$$\leq R.$$

Thus the sequence is uniformly bounded by $R$.

Now, to show the sequence is Cauchy we will notice that $U_{j+1} - U_j$ solves the equation with zero initial conditions and inhomogeneities $|u_j|^2 u_j - |u_{j-1}|^2 u_{j-1}$ and $|v_j|^2 v_j - |v_{j-1}|^2 v_{j-1}$. Factoring out the cubic terms, distributing the derivative, and applying Hölder's Inequality we get

$$\|U_{j+1} - U_j\|_X \lesssim \left\||u_j|^2 u_j - |u_{j-1}|^2 u_{j-1}\right\|_{L^1 L^2} + \left\|D_1(|u_j|^2 u_j - |u_{j-1}|^2 u_{j-1})\right\|_{L^1 L^2}$$

$$+ \left\||u_j|^2 u_j - |u_{j-1}|^2 u_{j-1}\right\|_{L^1 \dot{H}^{1-\frac{1}{q}}} + \left\||v_j|^2 v_j - |v_{j-1}|^2 v_{j-1}\right\|_{L^1 L^2}$$

$$+ \left\|D_1(|v_j|^2 v_j - |v_{j-1}|^2 v_{j-1})\right\|_{L^1 L^2} + \left\||v_j|^2 v_j - |v_{j-1}|^2 v_{j-1}\right\|_{L^1 \dot{H}^{1-\frac{1}{q}}}$$

$$\lesssim \|u_j - u_{j-1}\|_{L^2 L^2}\left(\|u_j\|_{L^4 L^\infty}^2 + \|u_{j-1}\|_{L^4 L^\infty}^2\right)$$

(2.36)

$$+ \|u_j - u_{j-1}\|_{L^2 \dot{H}^1}\left(\|u_j\|_{L^4 L^\infty}^2 + \|u_{j-1}\|_{L^4 L^\infty}^2\right)$$

$$+ \|u_j - u_{j-1}\|_{L^4 L^\infty}\left(\|u_j\|_{L^4 L^\infty}\|u_j\|_{L^2 \dot{H}^1} + \|u_{j-1}\|_{L^4 L^\infty}\|u_{j-1}\|_{L^2 \dot{H}^1}\right)$$

$$+ \|u_j - u_{j-1}\|_{L^2 \dot{H}^{1-\frac{1}{q}}}\left(\|u_j\|_{L^4 L^\infty}^2 + \|u_{j-1}\|_{L^4 L^\infty}^2\right)$$

$$+ \|u_j - u_{j-1}\|_{L^4 L^\infty}\left(\|u_j\|_{L^4 L^\infty}\|u_j\|_{L^2 \dot{H}^{1-\frac{1}{q}}} + \|u_{j-1}\|_{L^4 L^\infty}\|u_{j-1}\|_{L^2 \dot{H}^{1-\frac{1}{q}}}\right)$$

$$+ \text{ the equivalent terms for } v$$

$$\lesssim T^{\frac{q-2}{q}}\|U_j - U_{j-1}\|_X(\|U_j\|_X^2 + \|U_{j-1}\|_X^2).$$

Applying the uniform bound we get

$$\|U_{j+1} - U_j\|_X \leq C 2R^2 T^{\frac{q-2}{q}}\|U_j - U_{j-1}\|_X$$

(2.37)

$$= 8C^3 E_1^2 T^{\frac{q-2}{q}}\|U_j - U_{j-1}\|_X$$

$$\leq \frac{1}{2}\|U_j - U_{j-1}\|_X.$$

Thus the sequence $(U_j)_{j=-1}^\infty$ is Cauchy and uniformly bounded. Therefore the sequence converges to some $U \in X$. Note that the same uniform bound will hold for $U$.

Now we claim that $|U_j|^2 U_j \to |U|^2 U$. Proceeding similarly to (2.36), but starting on the right-hand side of the first line we get

(2.38)
$$\left\||U_j|^2 U_j - |U|^2 U\right\|_{L^1 H^1} \le C \|U_j - U\|_X \left(\|U_j\|_X^2 + \|U\|_X^2\right)$$
$$\le 2CR^2 T^{\frac{1}{2}} \|U_j - U\|_X.$$

Since $R$ is a uniform constant the right side must converge to zero. Thus $|U_j|^2 U_j \to |U|^2 U$.

Thus we conclude that $U_j \to U$ and $F(U_j) \to F(U)$. However, since $U_j \to U$ in $X$ it also converges distributionally and we conclude that $PU_j \to PU$. But $PU_j = F(U_{j-1}) \to F(U)$. By uniqueness of the limit we conclude that $PU = F(U)$ and thus $U$ is a solution in $X$.

The improvement in proven lifespan compared to the purely energy case comes from the fact that we bound $T^{\frac{q-2}{q}} < (16C^3 E_1^2)^{-1}$ instead of $T < (16C^3 E_1^2)^{-1}$. Thus we can extend our existence time close to $\epsilon^{-4}$ by choosing $q$ very close to 4.

While the result using Strichartz estimates does provide a better timescale for solutions to the Dirac equation and in a lower regularity space, in the proof of the main result in Chapter 3 we will just make use of the results for the purely energy-based local existence. In fact, we will only make use of the fact that the $C([0,T], H^s(\mathbb{R}^2))$ norm is bounded for any finite time, $T$, as long as the initial data is chosen small enough.

**2.3. Almost Global Existence.** In this section we will prove an almost global existence result for initial data $u(0, \mathbf{x}) = \epsilon f(\mathbf{x})$ and $v(0, \mathbf{x}) = \epsilon g(\mathbf{x})$. We write the equation in the form $PU = \begin{pmatrix} |u|^2 u \\ |v|^2 v \end{pmatrix}$ where P is the appropriate differential operator. Then by acting on the equation with the operator $\tilde{P}$ we find that $\tilde{P}P = -\Box \text{Id}$ and we get

(2.39)
$$\begin{pmatrix} -\Box u \\ -\Box v \end{pmatrix} = \begin{pmatrix} i\partial_t & \partial_x - i\partial_y \\ -\partial_x - i\partial_y & i\partial_t \end{pmatrix} \begin{pmatrix} |u|^2 u \\ |v|^2 v \end{pmatrix}$$
$$= \begin{pmatrix} i\partial_t(|u|^2 u) + \partial_x(|v|^2 v) - i\partial_y(|v|^2 v) \\ i\partial_t(|v|^2 v) - \partial_x(|u|^2 u) - i\partial_y(|u|^2 u) \end{pmatrix}$$

with initial data $U(0, \mathbf{x}) = \epsilon \begin{pmatrix} f(\mathbf{x}) \\ g(\mathbf{x}) \end{pmatrix}$ and $U_t(0, \mathbf{x}) = \epsilon \begin{pmatrix} -i\epsilon^2 |f|^2 f - ig_x - g_y \\ -i\epsilon^2 |g|^2 g + if_x - f_y \end{pmatrix}$.

38

For ease of notation we will define $m(\mathbf{x}), n(\mathbf{x})$ such that $U_t(0, \mathbf{x}) := \epsilon \begin{pmatrix} m(\mathbf{x}) \\ n(\mathbf{x}) \end{pmatrix}$.

We assume $f, g \in C_0^\infty(\mathbb{R}^2)$ and $\epsilon$ is small, and thus $m, n$ will also be in $C_0^\infty(\mathbb{R}^2)$. We will also define $R$ such the compact support of the initial data is within a ball around the origin of radius, $R$.

We will be looking for solutions to the indicated 2-D wave equation which we will see is a system of equations analogous to one of the equations studied by Katayama [47]. Thus we will approach the problem by using vector fields to get some useful estimates.

As in the last section, we give the detailed proof for the Dirac equation with the notationally simpler nonlinearity given in [1]. If we instead used the equation derived by Fefferman and Weinstein, $PU = \begin{pmatrix} \beta_1|u|u^2 + 2\beta_2|v|^2 u \\ 2\beta_2|u|^2 v + \beta_1|v|^2 v \end{pmatrix}$, the right-hand side of (2.39) would become

$$(2.40) \quad \begin{pmatrix} i\partial_t(\beta_1|u|^2 u + 2\beta_2|v|^2 u) + \partial_x(2\beta_2|u|^2 v + \beta_1|v|^2 v) - i\partial_y(2\beta_2|u|^2 v + \beta_1|v|^2 v) \\ i\partial_t(2\beta_2|u|^2 v + \beta_1|v|^2 v) - \partial_x(\beta_1|u|u^2 + 2\beta_2|v|^2 u) - i\partial_y(\beta_1|u|u^2 + 2\beta_2|v|^2 u) \end{pmatrix}.$$

We will make a remark at the key point in the proof where this nonlinearity would cause a difference.

THEOREM 2.3.1. *Given $U$ a solution to* (2.39) *with the above assumptions and maximal lifespan,* $T^*$, *there exists $\epsilon_0$ and $c$ such that $T^* \geq \exp(c\epsilon^{-2})$ for all $\epsilon \leq \epsilon_0$.*

Our argument for the proof stems from Theorem 1.2 in [47] where the same lifespan is proven for a class of wave equations in divergence form (plus a nonlinear part satisfying a null form). Katayama mentions that his lifespan is sharp as certain examples are known which contradict having a longer lifespan. However, at this point we have not concluded that the lifespan in Theorem 2.39 is sharp as we have not yet been able to prove that this equation is one such example.

**2.3.1. The Vector Fields.** We introduce the following vector fields: $\Gamma_0 = t\partial_t + x\partial_x + y\partial_y, \Omega_{01} = t\partial_x + x\partial_t, \Omega_{02} = t\partial_y + y\partial_t, \Omega_{12} = x\partial_y - y\partial_x, \partial_t, \partial_x, \partial_y$. We index these vector fields by $\Gamma = \{\Gamma_i\}_{i=0}^6$. In actual practice we will frequently be using $\Gamma_i\text{Id}$, but as this behaves identically to $\Gamma_i$ on its own this does not give any problems. It is well known that $\Box$ commutes with all of these vector fields except $[\Box, \Gamma_0] = 2\Box$. Thus $\Box\Gamma^\alpha u = \tilde{\Gamma}^\alpha\Box u$ where $\tilde{\Gamma}$ is just $\Gamma$ with $\Gamma_0 + 2$ instead of $\Gamma_0$ and $\alpha$ is just multi-index notation. Also the commutator of any two vector fields is just a linear combination of the vector fields.

We also use the notation

$$|v(t, \mathbf{x})|_s = \sum_{|\alpha| \leq s} |\Gamma^\alpha v(t, \mathbf{x})|$$

(2.41)

$$\|v(t, \cdot)\|_{s,p} = \left\| |v(t, \cdot)|_s \right\|_{L^p}.$$

Now we list some Lemmas from Katayama that we will need. Note that in all these cases we are working with $n = 2$

We start with the Klainerman-Sobolev inequality originally from [52]. However, we use the form of it written in [47]:

LEMMA 2.3.2 (Klainerman-Sobolev Inequality).

(2.42)
$$(1 + t + |\mathbf{x}|)^{\frac{1}{2}} (1 + |t - |\mathbf{x}||)^{\frac{1}{2}} |v(t, \mathbf{x})| \leq C \|v(t, \cdot)\|_{2,2}$$

for any function v with finite right-hand side.

The following decay estimate is proven in Glassey [36] (Lemma 1) among other places.

LEMMA 2.3.3. Let $f, g \in C_0^\infty(\mathbb{R}^2)$ both supported within the ball of radius $R$. If $v$ is a smooth solution to $\Box v = 0$ with $v(0, \mathbf{x}) = f(\mathbf{x})$ and $v_t(0, \mathbf{x}) = g(\mathbf{x})$ then

(2.43)
$$|v(t, \mathbf{x})| \leq C_R (1 + t + |\mathbf{x}|)^{-\frac{1}{2}} (1 + |t - |\mathbf{x}||)^{-\frac{1}{2}} (\|f\|_{W^{1,1}} + \|g\|_{L^1}).$$

We also have an estimate for an equation in divergence form proved by Katayama in [45] for the two dimensional case. His proof is based on Lindblad, [58], where it is proven in three dimensions.

LEMMA 2.3.4. Let v be a smooth solution to

(2.44)
$$\Box v(t, \mathbf{x}) = \sum_{a=0}^{2} \partial_a \Phi_a(t, \mathbf{x})$$

with zero initial data and the support of $\Phi_a(t, \cdot)$ contained in the ball of radius $t + R$ for any a or t. Then for any integer $s \geq 0$ we have

(2.45)
$$\|v(t, \cdot)\|_{s,2} \leq C_s \int_0^t \sum_{a=0}^{2} \|\Phi_a(\tau, \cdot)\|_{s,2} d\tau + C_{s,R} \sqrt{\log(2 + t)} \sum_{a=0}^{2} \|\Phi_a(0, \cdot)\|_{W^{s,1}}.$$

Also, for ease of notation we define

(2.46)
$$W_0(t, \mathbf{x}) = (1 + t + |\mathbf{x}|)^{-\frac{1}{2}} (1 + |t - |\mathbf{x}||)^{-\frac{1}{2}}.$$

40

It will also be useful to define

$$(2.47) \qquad\qquad W_-(t, \mathbf{x}) = (1 + |t - |\mathbf{x}||)^{-\frac{1}{2}}.$$

### 2.3.2. Almost Global Existence. We define an energy

$$(2.48) \qquad\qquad E(T) = \sup_{0 \leq t < T} \left( \sup_{\mathbf{x}} \{W_0(t, \mathbf{x})^{-1} |U(t, \mathbf{x})|_2\} + \|U(t, \cdot)\|_{4,2} \right).$$

Since $\epsilon$ is small we can assume $E(0) \ll 1$.

It is worth mentioning here that our definition of the energy, (2.48), is the primary departure from [**47**]. Compared to the energy he defines, our energy has a different number of vector fields on the two terms and does not have an additional derivative inside the $4, 2$-norm. However, having defined a similar energy, the fact that we can write our equation as a wave equation system in divergence form allows us to use the arguments in [**47**]. We are also able to use a simpler version of some of his arguments because we have a semilinear equation and because we do not need to do the more complicated arguements showing that $\|DU\|_{4,2}$ is bounded nicely. After having figured out the proper energy to use, (2.48), the main difficulty compared to the work in [**47**] is the fact that we will have more components to deal with when writing out our bounds.

Rewriting the equation we get

$$(2.49) \qquad\qquad \Box U = \sum_{a=0}^{2} \partial_a F^{(a)}(U)$$

with $F^{(0)}(U) = \begin{pmatrix} -i|u|^2 u \\ -i|v|^2 v \end{pmatrix}$, $F^{(1)}(U) = \begin{pmatrix} -|v|^2 v \\ |u|^2 u \end{pmatrix}$, and $F^{(2)}(U) = \begin{pmatrix} i|v|^2 v \\ i|u|^2 u \end{pmatrix}$.

REMARK 1. This is where we see the difference in how we would handle the Fefferman and Weinstein nonlinearity. Each $F^{(a)}$ would simply be made of two terms instead of one, and the upcoming expansions of the cubic terms would be notationally more complicated (as there would be additional terms with mixed $u$ and $v$ factors) but strategically the same.

We also note that within this section we are working with complex-valued functions $U$ and $F^{(a)}(U)$. However, throughout the proof, it will be trivial to replace each with its complex conjugate because we proceed by finding upper bounds for Lebesgue norms and vector field norms, (2.41), of these functions. Since each of these norms involves taking an absolute value, the conjugate will trivially have the same bound.

LEMMA 2.3.5. *For $U$ a solution to (2.39), the energy, $E(T)$ satisfies*

(2.50) $$E(T) \leq C \left( \epsilon + \sqrt{\log(2+T)}\epsilon^3 + \log(2+T)E(T)^3 \right).$$

PROOF. We split up $U$ into two parts $U = V + W$ with

(2.51) $$\Box V = \sum_{a=0}^{2} \partial_a F^{(a)}(U)$$

with zero initial data and

$$\Box W = 0$$

(2.52) $$W(0, \cdot) = \begin{pmatrix} f \\ g \end{pmatrix} \quad W_t(0, \cdot) = \begin{pmatrix} m \\ n \end{pmatrix}$$

Finding the appropriate bound for $W$ is very simple since we can just apply Lemma 2.3.3 to get

(2.53) $$W_0(t, \mathbf{x})^{-1} \|W(t, \cdot)\|_{2,\infty} \leq C_R \epsilon$$

where $\epsilon$ represents the size of the initial data.

The bound for $V$ will take some more work. We apply Lemma 2.3.4 to Equation (2.51) to get

(2.54) $$\|V\|_{4,2} \leq C \int_0^t \sum_{a=0}^{2} \|F^{(a)}(\tau, \cdot)\|_{4,2} d\tau + C_R \sqrt{\log(2+t)} \sum_{a=0}^{2} \|F^{(a)}(0, \cdot)\|_{W^{4,1}}.$$

Now we state an inequality which is easily verified by just tracking how the vector fields can be distributed among the factors

$$|F^{(a)}|_4 \leq C|u|_2^2|u|_4 + C|v|_2^2|v|_4.$$

Applying this estimate we get

(2.55) $$\begin{aligned}
\|F^{(a)}(\tau, \cdot)\|_{4,2} &\lesssim \left\||u|_2^2|u|_4\right\|_{L^2} + \left\||v|_2^2|v|_4\right\|_{L^2} \\
&= \left\|W_-^{-2}|u|_2^2 W_-^2|u|_4\right\|_{L^2} + \left\|W_-^{-2}|v|_2^2 W_-^2|v|_4\right\|_{L^2} \\
&\leq \left\|W_-^{-1}|u|_2\right\|_{L^\infty}^2 \left\|W_-^2|u|_4\right\|_{L^2} \\
&\quad + \left\|W_-^{-1}|v|_2\right\|_{L^\infty}^2 \left\|W_-^2|v|_4\right\|_{L^2}.
\end{aligned}$$

We can immediately get

$$\|W_-^{-1}|u|_2\|_{L^\infty} \leq (1+\tau)^{-\frac{1}{2}}E(T)$$

and likewise for $v$.

For the other factor notice that $W_-^2 \leq 1$ which means

$$\|W_-^2|u|_4\|_{L^2} \leq \|u\|_{4,2} \leq E(T)$$

and likewise for $v$.

Combining both of these we get

(2.56)
$$\|F^{(a)}(\tau,\cdot)\|_{4,2} \leq C(1+\tau)^{-1}E(T)^3.$$

Also, the $\|F^{(a)}(0,\cdot)\|_{W^{4,1}}$ in (2.54) is bounded by $\epsilon^3$ since we are just evaluating the norms of the initial conditions.

Thus we have

(2.57)
$$\|V(t,\cdot)\|_{4,2} \leq C_R\left(\sqrt{\log(2+t)}\epsilon^3 + \log(2+t)E(T)^3\right).$$

When we also apply Klainerman's Inequality we get

(2.58)
$$W_0(t,\mathbf{x})^{-1}\|V(t,\cdot)\|_{2,\infty} \leq C_R\left(\sqrt{\log(2+t)}\epsilon^3 + \log(2+t)E(T)^3\right).$$

Combining our estimates for $V$ and $W$ we get

(2.59)
$$W_0(t,\mathbf{x})^{-1}|U(t,\mathbf{x})|_2 \leq C_R\left(\epsilon + \sqrt{\log(2+t)}\epsilon^3 + \log(2+t)E(T)^3\right).$$

Now we establish a bound for $\|U(t,\cdot)\|_{4,2}$ by using the $L^2$ energy inequality for the wave equation and the fact that the vector fields commute with $\Box$.

(2.60)
$$\begin{aligned}
\|U(t,\cdot)\|_{4,2} &\leq C\|U(0,\cdot)\|_{4,2} + C\int_0^t \left\|\sum_{a=0}^2 F^{(a)}(U(\tau,\cdot))\right\|_{4,2} d\tau \\
&\leq C\|U(0,\cdot)\|_{4,2} + C\int_0^t \sum_{a=0}^2 \left\|F^{(a)}(U(\tau,\cdot))\right\|_{4,2} d\tau \\
&\leq C\|U(0,\cdot)\|_{4,2} + C\int_0^t (1+\tau)^{-1}E(T)^3 d\tau
\end{aligned}$$

where the last line follows from (2.56).

By evaluating the integral we immediately get

(2.61) $$\|U(t,\cdot)\|_{4,2} \leq C\epsilon + C\log(2+t)E(T)^3.$$

Then, combining our bounds for both terms of the energy we get the desired

(2.62) $$E(T) \leq C_R \left( \epsilon + \sqrt{\log(2+T)}\epsilon^3 + \log(2+T)E(T)^3 \right).$$

$\square$

We define a Picard iteration by $U_{-1} = 0$ and

(2.63) $$\Box U_{j+1} = \sum_{a=0}^{2} \partial_a F^{(a)}(U_j).$$

We also denote $E_j(T)$ to be the energy $E(T)$ of the $j$th Picard iterate. The same proof for Lemma 2.3.5 also immediately proves the inequality

(2.64) $$E_{j+1}(T) \leq C_R \left( \epsilon + \sqrt{\log(2+T)}\epsilon^3 + \log(2+T)E_j(T)^3 \right).$$

We note that the constant, $C_R$, depends only on the size of the support of the initial data and does not change for different values of $j$.

As $E_{-1}(T) = 0$, we fix $C_R = C$ and notice that $E_0 \leq 2C\epsilon$ for small enough initial data, and we assume inductively that $E_j(T) \leq 2C\epsilon$. Then we get

(2.65) $$E_{j+1}(T) \leq C \left( \epsilon + \log(2+T)\epsilon^3 + \log(2+T)8C^3\epsilon^3 \right).$$

If $\log(2+T) \leq \frac{1}{1+8C^3}\epsilon^{-2}$ then $E_{j+1}(T) \leq 2C\epsilon$. Thus the sequence $\{E_j(T)\}$ is uniformly bounded. In other words $\|U_j\|_{E(T)}$ is uniformly bounded where $\|\cdot\|_{E(T)}$ indicates the norm defined by the energy, $E(T)$.

To show that the sequence $U_j$ is Cauchy we consider $\|U_{j+1} - U_j\|_{E(T)}$. We notice that $U_{j+1} - U_j$ solves (2.63) with zero initial data and $F^{(a)}(U_j - U_{j-1}) := F^{(a)}(U_j) - F^{(a)}(U_{j-1})$. Following the

same proof as Lemma 2.3.5 and factoring the difference of cubes we arrive at

$$\|U_{j+1} - U_j\|_{E(T)} \leq C \log(2+T)\|U_j - U_{j-1}\|_{E(T)} \left( \|U_j\|^2_{E(T)} + \|U_{j-1}\|^2_{E(T)} \right)$$

$$\leq C \log(2+T)\|U_j - U_{j-1}\|_{E(T)} \left( E_j(T)^2 + E_{j-1}(T)^2 \right)$$

(2.66)

$$\leq 8C^3\epsilon^2 \log(2+T)\|U_j - U_{j-1}\|_{E(T)}$$

$$\leq \frac{8C^3\epsilon^2}{1 + 8C^3\epsilon^2}\|U_j - U_{j-1}\|_{E(T)}.$$

Since $\frac{8C^3\epsilon^2}{1+8C^3\epsilon^2} < 1$, this implies that the sequence $U_j$ is Cauchy.

Since $U_j$ is Cauchy and uniformly bounded, it must converge to some $U$. It is also easy to conclude that the limit, $U$, must be a solution to (2.39).

To summarize, for small enough initial data of size $\epsilon$, we have shown that there exists $k$ such that for any $T + 2 \leq e^{k\epsilon^{-2}}$ there exists a solution to (2.39) at least up to time, $T$. Thus, if $T^*$ is the maximal lifespan for such solutions, there also exists a $c$ such that $T^* \geq e^{c\epsilon^{-2}}$. This proves Theorem 2.3.1.

## 2.4. Ill-Posedness Below $\dot{H}^{\frac{1}{2}}$.

We have already established local well-posedness for the equation in an $H^1$ space. In this section we demonstrate ill-posedness in $\dot{H}^s$ for $s < \frac{1}{2}$ by construction of initial data that will result in a contradiction to local well-posedness as we iterate through successive approximations.

We have the fundamental solution to the corresponding inhomogeneous linear problem

(2.67)
$$\hat{U} = \mathcal{T}_t \begin{pmatrix} \hat{f} \\ \hat{g} \end{pmatrix} + i \int_0^t \mathcal{T}_{t-\tau} \begin{pmatrix} \hat{F}_1(\tau) \\ \hat{F}_2(\tau) \end{pmatrix} d\tau$$

where

$$\mathcal{T}_t = \begin{pmatrix} \cos(t|\xi|) & \frac{-\xi_1+i\xi_2}{|\xi|}\sin(t|\xi|) \\ \frac{\xi_1+i\xi_2}{|\xi|}\sin(t|\xi|) & \cos(t|\xi|) \end{pmatrix} := \begin{pmatrix} L_1(t,\xi) & L_2(t,\xi) \\ L_3(t,\xi) & L_1(t,\xi) \end{pmatrix},$$

the initial data is $U(0,\cdot) = \begin{pmatrix} f \\ g \end{pmatrix}$, and $F_1, F_2$ are the components of the inhomogeneity. Also note for future simplification that $L_3 = -\bar{L}_2$.

### 2.4.1. Choice of Initial Data.

Motivated by [21] we choose $\hat{g} = 0$, and define $f$ by

(2.68)
$$\hat{f} = \frac{1}{N^{s+1}}\chi_{2N \leq |\xi| \leq 3N}\chi_{|\theta| \leq \delta}$$

45

where $N$ is a large parameter and $\delta > 0$ is small. Thus our initial data is already frequency localized and, in particular, exists only in a narrow, high frequency sector. We will frequently abuse notation by letting $f$ also refer to the vector $\begin{pmatrix} f \\ 0 \end{pmatrix}$.

Proceeding along the usual iteration with this initial data doing the first iteration we find

$$(2.69) \qquad \hat{U}_0(t,\xi) = \mathcal{T}_t \begin{pmatrix} \hat{f} \\ 0 \end{pmatrix} = \begin{pmatrix} \cos(t|\xi|)\hat{f} \\ \frac{\xi_1 + i\xi_2}{|\xi|} \sin(t|\xi|)\hat{f} \end{pmatrix} := \begin{pmatrix} \hat{u}_1 \\ \hat{v}_1 \end{pmatrix}.$$

Thus the second iteration gives

$$(2.70) \qquad \hat{U}_1(t,\xi) = \begin{pmatrix} L_1(t,\xi)\hat{f} \\ L_3(t,\xi)\hat{f} \end{pmatrix} + i \int_0^t \mathcal{T}_{t-\tau} \begin{pmatrix} \widehat{|u_1|^2 u_1}(\tau,\xi) \\ \widehat{|v_1|^2 v_1}(\tau,\xi) \end{pmatrix} d\tau.$$

We notice that the cubic Duhamel term is in fact $A_3(f)$ defined in Section 3 of Bejenaru and Tao [18] (since the iterate, $u_1$, is defined in terms of $f$). Using the notation of [18] we define the spaces $D = \dot{H}^s$ and $S = C_t \dot{H}^s$. Theorem 3 of Bejenaru and Tao states that if the equation is well-posed and $f \in B_D(0,\epsilon) := \{f \in \dot{H}^s : \|f\|_{\dot{H}^s} < \epsilon\}$ for $\epsilon$ small with $A_3 : D \to S$, then

$$(2.71) \qquad \|A_3(f)\|_{C_t \dot{H}^s} \leq C\|f\|_{\dot{H}^s}^3.$$

We will prove a contradictory lower bound on $\|A_3(f)\|_{C_t \dot{H}^s}$.

### 2.4.2. Establishing a Lower Bound.
We now work to establish a lower bound to

$$(2.72) \qquad A_3 = i \int_0^t \mathcal{T}_{t-\tau} \begin{pmatrix} \widehat{|u_1|^2 u_1}(\tau,\xi) \\ \widehat{|v_1|^2 v_1}(\tau,\xi) \end{pmatrix} d\tau.$$

We can use the convolution theorem to expand $\widehat{|u_1|^2 u_1} = \hat{u}_1 * \hat{\bar{u}}_1 * \hat{u}_1$ and likewise for the second component. Doing this inside of $A_3$ we get

$$A_3 = i \int_0^t \int \mathcal{T}_{t-\tau} \begin{pmatrix} \hat{u}_1(\tau,\sigma)\hat{\bar{u}}_1(\tau,\xi-\eta)\hat{u}_1(\tau,\eta-\sigma) \\ \hat{v}_1(\tau,\sigma)\hat{\bar{v}}_1(\tau,\xi-\eta)\hat{v}_1(\tau,\eta-\sigma) \end{pmatrix} d\sigma d\eta d\tau$$

$$(2.73)$$

$$:= \begin{pmatrix} B_3 \\ C_3 \end{pmatrix}.$$

46

Trivially we know $\|A_3\|_{C_t \dot{H}^s} \geq \|B_3\|_{C_t \dot{H}^s}$. Thus we will only consider its first component

$$B_3 = i \int_0^t \int L_1(t - \tau, \xi)\hat{u}_1(\tau, \sigma)\hat{\bar{u}}_1(\tau, \xi - \eta)\hat{u}_1(\tau, \eta - \sigma)$$

$$+ L_2(t - \tau, \xi)\hat{v}_1(\tau, \sigma)\hat{\bar{v}}_1(\tau, \xi - \eta)\hat{v}_1(\tau, \eta - \sigma)d\sigma d\eta d\tau.$$

(2.74)

Plugging in the definition of $u_1$ and $v_1$ (and changing the order of complex conjugation and the Fourier transform in the middle factor) we get

$$B_3 = i \int_0^t \int \Big( L_1(t - \tau, \xi)L_1(\tau, \sigma)L_1(\tau, \xi - \eta)L_1(\tau, \eta - \sigma)\hat{f}(\sigma)\hat{f}(\eta - \xi)\hat{f}(\eta - \sigma)$$

$$- L_2(t - \tau, \xi)L_3(\tau, \sigma)L_2(\tau, \eta - \xi)L_3(\tau, \eta - \sigma)\hat{f}(\sigma)\hat{f}(\eta - \xi)\hat{f}(\eta - \sigma) \Big) d\sigma d\eta d\tau.$$

(2.75)

Now we want to determine what $\xi$ looks like in the support of $\hat{f}(\sigma)\hat{f}(\eta - \xi)\hat{f}(\eta - \sigma)$. Considering the fact that $\xi = \sigma + (\eta - \sigma) - (\eta - \xi)$ the triangle inequality gives us that $|\xi| \leq 9N$. We also want to establish a lower bound for $|\xi|$. To do this let us assume that $\delta$ is small enough such that $\cos(\theta) > \frac{7}{8}$ for all $|\theta| \leq \delta$. Thus the x-component of $\sigma$ and $\eta - \sigma$, are at least $2N(\frac{7}{8})$ while the x-component of $\eta - \xi$ is at most $3N$. We determine the x-component of $\xi$ is at least $2N(\frac{7}{8}) + 2N(\frac{7}{8}) - 3N = \frac{N}{2}$. Thus $\frac{1}{2}N \leq |\xi| \leq 9N$. Thus in the support for the integrand we have that $|\xi|, |\sigma|, |\xi - \eta|, |\eta - \sigma| \sim N$.

Now we will also assume that we are only considering a very small timescale, $t < \frac{1}{100N}$, and establish some useful bounds for each $L_i$.

For the following inequalities we assume $\omega$ is one of our frequency variables. Then our restriction on time gives us the following inequalities

$$L_1(t, \omega) = \cos(t|\omega|) \geq \cos\left(\frac{1}{100N}|\omega|\right) > \frac{3}{4}$$

(2.76)

$$|L_2(t, \omega)| = \left|\frac{-\omega_1 + i\omega_2}{|\omega|}\sin(t|\omega|)\right| \leq \left|\sin\left(\frac{1}{100N}|\omega|\right)\right| < \frac{1}{4}$$

$$|L_3(t, \omega)| = \left|\frac{\omega_1 + i\omega_2}{|\omega|}\sin(t|\omega|)\right| \leq \left|\sin\left(\frac{1}{100N}|\omega|\right)\right| < \frac{1}{4}.$$

The same bounds hold if we replace $t$ by $t - \tau$ since it will still be less than $\frac{1}{100N}$.

Now we rewrite the integrand of $B_3$ as $(W_1 - W_2)\hat{f}(\sigma)\hat{f}(\eta - \xi)\hat{f}(\eta - \sigma)$ where

$$W_1 = L_1(t - \tau, \xi)L_1(\tau, \sigma)L_1(\tau, \xi - \eta)L_1(\tau, \eta - \sigma)$$

and

$$W_2 = L_2(t - \tau, \xi)L_3(\tau, \sigma)L_2(\tau, \eta - \xi)L_3(\tau, \eta - \sigma).$$

Then $|W_1| \geq (\frac{3}{4})^4$ and $|W_2| \leq (\frac{1}{4})^4$. This implies that

$$(2.77) \qquad \left|(W_1 - W_2)\hat{f}(\sigma)\hat{f}(\eta - \xi)\hat{f}(\eta - \sigma)\right| > \frac{5}{16}|\hat{f}(\sigma)||\hat{f}(\eta - \xi)||\hat{f}(\eta - \sigma)|.$$

Now let us work out a series of inequalities

$$
\begin{aligned}
\|A_3\|_{C_t \dot{H}^s} &\geq \|B_3\|_{C_t \dot{H}^s} \\
&\gtrsim \sup_{[0,T]} \left\| |\xi|^s \int_0^t \int |\hat{f}(\tau,\sigma)||\hat{f}(\tau,\eta - \xi)||\hat{f}(\tau,\eta - \sigma)| d\sigma d\eta d\tau \right\|_{L^2} \\
(2.78) \qquad &\gtrsim \left\| N^s \int_0^{\frac{1}{100N}} N^{-3s-3}(N^2\delta)^2 \chi_{2N \leq |\xi| \leq 3N} d\tau \right\|_{L^2} \\
&\gtrsim N^s N^{-1} N^{-3s-3}(N^2\delta)^2(N\delta^{\frac{1}{2}}) \\
&= N^{-2s+1}\delta^{\frac{5}{2}}.
\end{aligned}
$$

Notice that we used the fact that $f$ is supported on a sector of a circle with radius, $N$, and angle, $\delta$, thus the size of its support is effectively $N^2\delta$.

Thus we finish the section with the conclusion that

$$(2.79) \qquad \|A_3(f)\|_{C_t \dot{H}^s} \gtrsim N^{-2s+1}\delta^{\frac{5}{2}}.$$

**2.4.3. The Contradiction.** From [**18**] we know that if the problem is well-posed then $\|A_3(g)\|_{C_t \dot{H}^s} \leq C\|g\|_{\dot{H}^s}^3$ for arbitrary initial data, $g$. However for our initial data $f$ it is simple to calculate $\|f\|_{\dot{H}^s}^3 \sim \delta^{\frac{3}{2}}$.

Thus if we assume our problem is well-posed we have

$$(2.80) \qquad \|A_3(g)\|_{C_t \dot{H}^s} \lesssim \delta^{\frac{3}{2}}.$$

However, if we choose our small parameter $\delta = N^{-\epsilon}$ equation (2.79) gives us

$$(2.81) \qquad \|A_3\|_{C_t \dot{H}^s} \gtrsim N^{-2s+1-\epsilon}\delta^{\frac{3}{2}}.$$

Considering $N$ can be as large as we want this is a contradiction to (2.80) if $-2s + 1 - \epsilon > 0$. Thus we obtain a contradiction to well-posedness whenever $s < \frac{1}{2}$.

**2.4.4. Extending to Another Nonlinearity.** It is also worthwhile to see if this argument can extend to other cubic nonlinearities. Here we look at a particular nonlinearity which is known to

satisfy a null condition, $-\langle \gamma U, U \rangle \gamma U$ where $\gamma = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Thus the nonlinearity can be written

$$\begin{pmatrix} -|u|^2 u + |v|^2 u \\ |u|^2 v - |v|^2 v \end{pmatrix}.$$

This is an example of the class of nonlinearities given in Pecher [**66**], among other places, of the particle-model Dirac equation. It is proved in [**66**] that the massive cubic Dirac equation with this nonlinearity is well-posed in $H^s$ for $s > 1/2$ which is the optimal regularity up to the endpoint. Thus this section verifies the ill-posedness below the regularity, $\frac{1}{2}$, as discussed in [**66**]. We include this to demonstrate how initial data of this type could prove ill-posedness for more cubic nonlinearities.

The argument is almost identical to our previous nonlinearity. The only difference is that there will be four terms to deal with instead of merely two terms.

When we expand out the Duhamel term with this nonlinearity we end up defining

$$(2.82) \qquad A_3 = i \int_0^t \mathcal{T}_{t-\tau} \begin{pmatrix} -\widehat{|u_1|^2 u_1}(\tau, \xi) + \widehat{|v_1|^2 u_1}(\tau, \xi) \\ \widehat{|u_1|^2 v_1}(\tau, \xi) - \widehat{|v_1|^2 v_1}(\tau, \xi) \end{pmatrix} d\tau.$$

When we pick out only the first component as before we get

$$
\begin{aligned}
B_3 = i \int_0^t \int &- L_1(t - \tau, \xi)\hat{u}_1(\tau, \sigma)\hat{\bar{u}}_1(\tau, \xi - \eta)\hat{u}_1(\tau, \eta - \sigma) \\
&+ L_1(t - \tau, \xi)\hat{v}_1(\tau, \sigma)\hat{\bar{v}}_1(\tau, \xi - \eta)\hat{u}_1(\tau, \eta - \sigma) \\
&+ L_2(t - \tau, \xi)\hat{u}_1(\tau, \sigma)\hat{\bar{u}}_1(\tau, \xi - \eta)\hat{v}_1(\tau, \eta - \sigma) \\
&- L_2(t - \tau, \xi)\hat{v}_1(\tau, \sigma)\hat{\bar{v}}_1(\tau, \xi - \eta)\hat{v}_1(\tau, \eta - \sigma)\, d\sigma d\eta d\tau.
\end{aligned}
$$

(2.83)

Substituting in for $u_1, v_1$ as before we get

$$
\begin{aligned}
B_3 = i \int_0^t \int \Big( &- L_1(t - \tau, \xi)L_1(\tau, \sigma)L_1(\tau, \xi - \eta)L_1(\tau, \eta - \sigma) \\
&- L_1(t - \tau, \xi)L_3(\tau, \sigma)L_2(\tau, \xi - \eta)L_1(\tau, \eta - \sigma) \\
&+ L_2(t - \tau, \xi)L_1(\tau, \sigma)L_1(\tau, \eta - \xi)L_3(\tau, \eta - \sigma) \\
&+ L_2(t - \tau, \xi)L_3(\tau, \sigma)L_2(\tau, \eta - \xi)L_3(\tau, \eta - \sigma) \Big) \\
&\hat{f}(\sigma)\hat{f}(\eta - \xi)\hat{f}(\eta - \sigma)\, d\sigma d\eta d\tau.
\end{aligned}
$$

(2.84)

Keeping the same restriction as before on $t$ we also have the same bounds for $L_i$. Also using the trivial bound $L_1 \leq 1$ we get that the integrand of $B_3$ is greater than

$$\left((3/4)^4 - (1/4)^2 - (1/4)^2 - (1/4)^4\right) \hat{f}(\sigma)\hat{f}(\eta - \xi)\hat{f}(\eta - \sigma) = \frac{3}{16}\hat{f}(\sigma)\hat{f}(\eta - \xi)\hat{f}(\eta - \sigma).$$

The rest of the proof works out identically because we have a lower bound for $B_3$ which differs only by a constant from the case of the previous nonlinearity.

**2.5. Discussion of Global Existence Versus Blow-Up.** In this section we do not prove any results but discuss several techniques and results dealing with global existence or finite time blow-up for Dirac and wave equations and how they do (or do not) apply for the Dirac equations, (0.1) and (0.4). We first look at two particular null conditions and discuss why they does not apply to either of our nonlinearities.

Next we investigate the spacetime resonance sets of (0.1) to see if there is compelling evidence for the existence of some other null structure which could lead to global existence [**34**]. We see that there is not the necessary cancellation within each of the spacetime resonance sets to indicate such an undiscovered null structure.

We conclude this section by investigating whether we can use any ordinary differential equation (ODE) blow-up techniques to prove that the solution does blow-up in finite time. However, we are unable to reformulate (0.1) in terms of an ODE, and thus the techniques are not helpful.

**2.5.1. Lack of the Null Form.** In studying the Dirac and wave literature we find many examples of nonlinearities which fail to have a null condition and blowup in finite time [**37, 47, 54, 77, 57, 58, 67, 81, 84**] as well as nonlinearities which satisfy a null condition and have global existence [**16, 17, 20, 40, 48, 45, 46, 53, 62, 61, 66, 67**]. We start by trying to determine whether (0.1) satisies any of these known null conditions.

We first consider the existence (or nonexistence) of the null form when written as the wave equation. When we consider the null condition of Klainerman, [**53**], in two dimensions we get the following definition.

DEFINITION 2. The nonlinearity, $N(u, Du, D_x Du)$, satisfies the null condition if $N(a, b(X_i)_{i=0,1,2}, c(X_j X_i)_{j=1,2,i}$
0 for all $a, b, c \in \mathbb{R}^2$ and null vectors $X = (X_0, X_1, X_2)$ (i.e. $X_0^2 - X_1^2 - X_2^2 = 0$ or $X$ on the light cone).

When plugging in a null vector $X$ into the nonlinearity of the wave equation formulation we get $-3ia_1^2 b_1 X_0 - 3a_2^2 b_2 X_1 + 3ia_2^2 b_2 X_2$. Clearly plugging in $X_0^2 = X_1^2 + X_2^2$ does not result in this quantity vanishing. Thus, the equation does not satisfy the Klainerman null condition. It works out similarly if we use the nonlinearity in (0.4)

However, the purpose of the null form is to establish a decay estimate for the bilinear or trilinear forms in the nonlinearity that is superior to what can be obtained from a generic bilinear or trilinear form. Thus it is possible to have null form behavior even without satisfying the particular Klainerman null condition. To see an example of this we look at the Dirac equations studied in [**16, 17, 20, 62, 61, 66, 67**] all of which have a null form which is used to achieve their global results. The nonlinearities are similar to $\langle \beta U, U \rangle \beta U$ where $\beta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ where the change in sign from $\beta$ results in some cancellation preventing the solutions from blowing up. In our nonlinearity each component is a monomial so there is no obvious cancellation to keep the solutions finite as time goes to infinity.

To look at this in more detail we look at Bejenaru and Herr [**16, 17**]. They work with the massive Dirac equation with the nonlinearity $\langle \psi, \beta \psi \rangle \beta \psi$. They prove global existence for their Dirac equation by projecting to the eigenspaces of the operator with projection $\Pi_\pm$. They define $\Pi_\pm \psi = \psi_\pm$ and use the identity

$$(2.85) \qquad\qquad \Pi_\pm(D)\beta = \beta \Pi_\mp(D) \mp \langle D \rangle^{-1}$$

where $\langle D \rangle$ has the symbol $\sqrt{|\xi|^2 + 1}$.

They expand out $\langle \psi, \beta \psi \rangle$ in the nonlinearity by using $\psi = \psi_+ + \psi_-$ and $\Pi_\pm \psi_\pm = \psi_\pm$ to get

$$
\begin{aligned}
(2.86) \qquad \langle \psi, \beta \psi \rangle &= \langle \Pi_+(D)\psi_+, \beta \Pi_+(D)\psi_+ \rangle + \langle \Pi_-(D)\psi_-, \beta \Pi_-(D)\psi_- \rangle \\
&\quad + \langle \Pi_+(D)\psi_+, \beta \Pi_-(D)\psi_- \rangle + \langle \Pi_-(D)\psi_-, \beta \Pi_+(D)\psi_+ \rangle.
\end{aligned}
$$

The nonlinearity being written in vector form with an inner product is essential in their proof of the null form as it will allow the projection operators to move between sides of the inner product. Without loss of generality we investigate the first term above and look at it on the Fourier side $\langle \Pi_+(\xi)\hat{\psi}_+(\xi), \beta \Pi_+(\eta)\hat{\psi}_+(\eta) \rangle$. Using the identity mentioned above and writing as a bilinear form

Bejenaru and Herr show

$$\langle \Pi_+(\xi)\hat{\psi}_1(\xi), \beta\Pi_+(\eta)\hat{\psi}_2(\eta)\rangle$$

(2.87)
$$= \langle \Pi_+(\xi)\hat{\psi}_1(\xi), \Pi_-(\eta)\beta\hat{\psi}_2(\eta)\rangle - \langle \Pi_+(\xi)\hat{\psi}_2(\xi), \frac{1}{\langle\eta\rangle}\hat{\psi}_2(\eta)\rangle$$

$$= \langle \Pi_-(\eta)\Pi_+(\xi)\hat{\psi}_1(\xi), \beta\hat{\psi}_2(\eta)\rangle - \langle \Pi_+(\xi)\hat{\psi}_1(\xi), \frac{1}{\langle\eta\rangle}\hat{\psi}_2(\eta)\rangle.$$

They further use a lemma stating that $\Pi_+(\xi)\Pi_-(\eta) = \mathcal{O}(\angle(\xi,\eta)) + \mathcal{O}(\langle\xi\rangle^{-1} + \langle\eta\rangle^{-1})$. Thus except for the presence of $\mathcal{O}(\angle(\xi,\eta))$ the right-hand side of the bilinear equality can have a $\mathcal{O}(\langle\xi\rangle^{-1} + \langle\eta\rangle^{-1})$ factored out of both terms. However, it is known that the strongest potential contribution to blowup of a solution is when $\angle(\xi,\eta) \approx 0$. To make use of this we can separate into the regions where $|\angle(\xi,\eta)| > \epsilon$ and $|\angle(\xi,\eta)| \leq \epsilon$ for $\epsilon$ small. When $|\angle(\xi,\eta)| > \epsilon$ we do not need extra decay, and when $|\angle(\xi,\eta)| \leq \epsilon$ the immediately above equality results in an extra $\langle\xi\rangle^{-1} + \langle\eta\rangle^{-1}$ decay.

This null structure cannot be found in the Dirac equation we study. The primary issue is the fact that the nonlinearity cannot be written in vector form and must be written in component form. If we expand out the $|u|^2$ in the first component as Bejenaru and Herr expanded out $\langle\psi, \beta\psi\rangle$ we only get

(2.88)
$$|u|^2 = |u_+|^2 + |u_-|^2 + u_+\bar{u}_- + \bar{u}_+u_-.$$

We cannot even write $u_+ = \Pi_+ u_+$ since $\Pi_+$ must act on a vector. Even more so, we cannot even define $u_\pm$ directly with the projection operator. We have to define $u_\pm$ as the first component of $\Pi_\pm U$. So, there is no way to find a null form by manipulating the projection operators since the projection operators cannot even be directly introduced to the equations. Even more telling is that the null form in other articles essentially reduces to an estimate bounding the product of two projection operators by the angle between Fourier variables [20].

Also of interest is the work of Fefferman and Weinstein [31]. They start with the same nonlinear Schrödinger equation as Ablowitz, Nixon, and Zhu [1] but use a different derivation to arrive at a similar Dirac equation ((0.4) instead of (0.1)). The primary difference in the Fefferman and Weinstein Dirac equation is that the nonlinearity, $F$, is replaced by $F(U) = \begin{pmatrix} (\beta_1|u|^2 + 2\beta_2|v|^2)u \\ (2\beta_2|u|^2 + \beta_1|v|^2)v \end{pmatrix}$ where $\beta_1, \beta_2$ are nonnegative constants. Following the same steps as above for the Ablowitz, Nixon, and Zhu nonlinearity we are able to conclude that the Fefferman-Weinstein nonlinearity also does

not meet the Klainerman null condition or the null condition discussed by Bejenaru and Herr, [**16, 17**].

Next we rewrite the nonlinearity by factoring out $\beta_1$ resulting in $F(U) = \beta_1 \begin{pmatrix} (|u|^2 + k|v|^2)u \\ (k|u|^2 + |v|^2)v \end{pmatrix}$ where $k = 2\beta_2/\beta_1$. By expanding $k|u|^2 = -|u|^2 + (k+1)|u|^2$ (and similarly for $v$) the nonlinearity can be written as

$$\frac{1}{\beta_1}F(U) = \langle \beta U, U \rangle \beta U + (k+1)\begin{pmatrix} |v|^2 u \\ |u|^2 v \end{pmatrix}.$$

However, by substituting $|v|^2 u = |u|^2 u - \langle \beta U, U \rangle u$ (and similarly for $|u|^2 v$) we see the Fefferman-Weinstein nonlinearity can be written

$$(2.89) \qquad \frac{1}{\beta_1}F(U) = -k\langle \beta U, U \rangle \beta U + (k+1)\begin{pmatrix} |u|^2 u \\ |v|^2 v \end{pmatrix}.$$

This is significant as Bournaveas and Candy [**20**] show that the nonlinearity $F(U) = \langle \beta U, U \rangle \beta U$ has null structure and that the two-dimensional massless Dirac equation with this nonlinearity has small data global existence. Thus the Fefferman-Weinstein nonlinearity can be broken down into a nonlinearity which has null structure for which small data global existence results have been shown and another component which has the same structure as our nonlinearity coming from Ablowitz, Nixon, and Zhu.

**2.5.2. Resonances of the Equation.** There is also a link between the existence of a null form and the behavior of the resonant sets of an equation. In general, there are ways to find good decay estimates as long as you are away from the intersection of space resonances and time resonances. One technique explained by Germain [**34**] tells you to integrate by parts in the time or space variables, as appropriate, similarly to the stationary phase method. The contribution to blowup will normally occur on the set of spacetime resonances (the intersection). Essentially, cancellation within each of the spacetime resonance sets can indicate the presence of some null form which could be used to prove global existence.

In order to determine the resonance sets we will start by looking at the profile of $U$ instead, which we can define by

$$(2.90) \qquad \hat{G}(t,\xi) = \begin{pmatrix} G_1 \\ G_2 \end{pmatrix} = \begin{pmatrix} \cos(t|\xi|) & m_1(\xi)\sin(t|\xi|) \\ m_2(\xi)\sin(t|\xi|) & \cos(t|\xi|) \end{pmatrix}^{-1} \hat{U}$$

where $m_1 = \frac{-\xi_1 + i\xi_2}{|\xi|}$ and $m_2 = \frac{\xi_1 + i\xi_2}{|\xi|}$. Thus we find that

(2.91) $\qquad \hat{G}(t,\xi) = \hat{U}(0,\xi) + i \int_0^t \begin{pmatrix} \cos(s|\xi|) & -m_1(\xi)\sin(s|\xi|) \\ -m_2(\xi)\sin(s|\xi|) & \cos(s|\xi|) \end{pmatrix} \begin{pmatrix} \widehat{|u|^2 u(s)} \\ \widehat{|v|^2 v(s)} \end{pmatrix} ds.$

Next we take advantage of the fact that $\widehat{|u|^2 u(s)} = \widehat{u\bar{u}u(s)} = \hat{u}(s) \star \hat{\bar{u}}(s) \star \hat{u}(s)$. Expanding out the convolution we get

$$\hat{G}(t,\xi) = \hat{U}(0,\xi) + i \int_0^t \int \begin{pmatrix} \cos(s|\xi|) & -m_1(\xi)\sin(s|\xi|) \\ -m_2(\xi)\sin(s|\xi|) & \cos(s|\xi|) \end{pmatrix}$$

(2.92)

$$\times \begin{pmatrix} \hat{u}(s,\sigma)\hat{\bar{u}}(s,\xi-\eta)\hat{u}(s,\eta-\sigma) \\ \hat{v}(s,\sigma)\hat{\bar{v}}(s,\xi-\eta)\hat{v}(s,\eta-\sigma) \end{pmatrix} d\sigma d\eta ds.$$

Although to write the equation fully as a profile we want to replace $u,v$ by the components $G_1, G_2$. To do this we note

$$\hat{U}(t,\xi) = \begin{pmatrix} \cos(t|\xi|) & m_1(\xi)\sin(t|\xi|) \\ m_2(\xi)\sin(t|\xi|) & \cos(t|\xi|) \end{pmatrix} \hat{G}(t,\xi)$$

(2.93)

$$\begin{pmatrix} \hat{u} \\ \hat{v} \end{pmatrix}(t,\xi) = \begin{pmatrix} \cos(t|\xi|)G_1(\xi) + m_1(\xi)\sin(t|\xi|)G_2(\xi) \\ m_2(\xi)\sin(t|\xi|)G_1(\xi) + \cos(t|\xi|)G_2(\xi) \end{pmatrix}.$$

By expanding the $\sin(t|\xi|)$ and $\cos(t|\xi|)$ in terms of the half-wave operator using and fully expanding out the matrix multiplication the integrand becomes a linear combination of terms of the form

(2.94) $\qquad e^{is(\pm|\xi|\pm|\xi-\eta|\pm|\eta-\sigma|\pm|\sigma|)}[m_i(\xi)m_i(\xi-\eta)m_i(\eta-\sigma)m_i(\sigma)]\hat{G}_j(\sigma)\hat{G}_j(\eta-\sigma)\hat{\bar{G}}_j(\xi-\eta)$

where $j = 1, 2$, $i = 0, 1, 2$, $m_0$ is uniformly 1, and all the subscripts are independent of each other. In fact, every combination of $\pm$ is attained in the half-wave term. However it is sufficient to consider only the cases with $-|\xi|$.

Thus we have the phase function $\varphi = \varphi_{\pm,\pm,\pm} = -|\xi| \pm |\xi-\eta| \pm |\eta-\sigma| \pm |\sigma|$

First we will consider the set of space resonances of the phase, $\mathcal{S}$. To do that we need to take the $\sigma$ and $\eta$ derivatives of $\varphi$. Thus we get

$$\partial_\sigma \varphi_{\cdot,\pm,\pm} = \mp \frac{\eta - \sigma}{|\eta - \sigma|} \pm \frac{\sigma}{|\sigma|}$$

(2.95)

$$\partial_\eta \varphi_{\pm,\pm,\cdot} = \mp \frac{\xi - \eta}{|\xi - \eta|} \pm \frac{\eta - \sigma}{|\eta - \sigma|}.$$

The set of space resonances is where $\varphi_\sigma = \varphi_\eta = 0$. Setting the above equations equal to 0 we conclude for each phase function $\varphi_{\pm,\pm,\pm}$ the set of space resonances will require that $\sigma, \eta, \xi$ are all collinear. In other words $\mathcal{S} = \{\eta = a\sigma, \xi = b\sigma\}$, and each particular phase function has its own restrictions on $a, b$ based on whether $\sigma, \eta - \sigma$, and $\xi - \eta$ have the same or opposite directions.

For example, let us consider $\varphi_{+,+,+}$. On its set of space resonances $\frac{\eta - \sigma}{|\eta - \sigma|} = \frac{\sigma}{|\sigma|}$ and $\frac{\xi - \eta}{|\xi - \eta|} = \frac{\eta - \sigma}{|\eta - \sigma|}$. In other words, $\sigma, \eta - \sigma, \xi - \eta$ are all collinear AND have the same direction. Thus $|\eta| > |\sigma|$ and $|\xi| > |\eta|$ meaning $b > a > 1$.

The restrictions on $a, b$ for the other phase functions are included in the table below.

The set of time resonances, $\mathcal{T}$, is very easy to find since it is merely when $\varphi = 0$. We also include those in the table.

To find the set of spacetime resonances, $\mathcal{R}$, we merely plug in $\eta = a\sigma$ and $\xi = b\sigma$ into the restriction for the time resonances.

| Phase | $\mathcal{S} = \{\eta = a\sigma, \xi = b\sigma\}$ | $\mathcal{T}$ | $\mathcal{R}$ |
|---|---|---|---|
| $\varphi_{+,+,+}$ | $b > a > 1$ | $|\xi| = |\xi - \eta| + |\eta - \sigma| + |\sigma|$ | $\mathcal{S}$ |
| $\varphi_{+,+,-}$ | $b < a < 1$ | $|\xi| + |\sigma| = |\xi - \eta| + |\eta - \sigma|$ | $b \leq 0, a < 1$ |
| $\varphi_{-,+,+}$ | $a > 1, b$ | $|\xi| + |\xi - \eta| = |\eta - \sigma| + |\sigma|$ | $a > 1, b > 0$ |
| $\varphi_{+,-,+}$ | $a < 1, b$ | $|\xi| + |\eta - \sigma| = |\xi - \eta| + |\sigma|$ | $\mathcal{S}$ |
| $\varphi_{-,-,-}$ | $b > a > 1$ | $\xi = \eta = \sigma = 0$ | $\xi = \eta = \sigma = 0$ |
| $\varphi_{-,-,+}$ | $b < a < 1$ | $|\xi| + |\xi - \eta| + |\eta - \sigma| = |\sigma|$ | $0 \leq b < a < 1$ |
| $\varphi_{+,-,-}$ | $a > 1, b$ | $|\xi| + |\eta - \sigma| + |\sigma| = |\xi - \eta|$ | $b \leq 0, a > 1$ |
| $\varphi_{-,+,-}$ | $a < 1, b$ | $|\xi| + |\xi - \eta| + |\sigma| = |\eta - \sigma|$ | $a < b \leq 0$ |

When we look at these resonance sets we see that for $\varphi_{+,+,-}, \varphi_{+,-,-}$ and $\varphi_{-,+,-}$ one of the vectors $\xi, \sigma, \eta$ points in the opposite direction of the others This could allow for cancellation and the existence of a null form. However, for $\varphi_{+,+,+}, \varphi_{-,+,+}$ and $\varphi_{-,-,+}$ the vectors $\xi, \sigma, \eta$ all point in

the same direction. We also note that in $\varphi_{+,+,+}$ and $\varphi_{-,+,+}$ at least one of $a, b$ is bounded greater than 1. Thus, on those resonance sets, no null structure is particularly evident. Also, for $\varphi_{+,-,+}$ the vectors will sometimes point in the same direction and sometimes have one in the opposite direction. Since the resonant sets for the different phases sometimes lie in the same direction and sometimes in the opposite direction it is impossible for cancellation to occur in each of the sets.

This lack of cancellation within each of the spacetime resonance sets fails to indicate the presence of a null form. It is important to note that this does not mean the absence of a null form. This merely means that we lack additional evidence for the existence of an, as of yet, undiscovered null form.

### 2.5.3. Blowup in the Quadratic Three-Dimensional Dirac Equation.

Just as the cubic nonlinearity is critical in two space dimensions for the Dirac equation, so is the quadratic nonlinearity critical in three space dimensions. Tzvetkov in [**81**] provides a blowup solution for just such a Dirac equation. He is working with the Dirac equation $\mathcal{D}\psi = F(\psi)$ where $\mathcal{D} = i \sum \gamma^\mu \partial_\mu$ and $\gamma^\mu$ are the Dirac matrices. In his blowup result, Tzvetkov essentially assumes $F(\psi) = |\psi|\gamma^0\psi$.

The presence of the Dirac matrix in the nonlinearity is essential to his proof as he begins the proof of his blowup result by multiplying each side of the Dirac equation by $-\gamma^0\psi$. Along with assuming compact support of the initial data this allowed him to determine

$$(2.96) \qquad \partial_t \int_{\mathbf{x} \leq t+R} |\psi(t,\mathbf{x})|^2 dx = \int_{\mathbf{x} \leq t+R} |\psi(t,\mathbf{x})|^3 dx$$

where $R$ is the bound on the support of the initial data. We note that this is possible because Tzvetkov's equation does not conserve mass. The essential element (which is lacking in our two-dimensional Dirac equation) is the positivity of the integrand on the right-hand side. By applying Hölder to the above inequality Tzvetkov gets

$$(2.97) \qquad \partial_t \int_{\mathbb{R}^3} |\psi(t,\mathbf{x})|^2 dx \geq c(t+R)^{-3/2} \left( \int_{\mathbb{R}^3} |\psi(t,\mathbf{x})|^2 dx \right)^{\frac{3}{2}}.$$

When defining $y(t) = \|\psi\|_{L^2}$ this becomes the ODE

$$(2.98) \qquad y' \geq \frac{cy^{3/2}}{(t+R)^{3/2}}.$$

The solution to this ODE blows up in finite time assuming $R < c^2 y_0$. Thus by picking the support of the initial data to be small enough, Tzvetkov showed that the solution of the Dirac equation blows up in finite time.

The proof used by Tzvetkov fails to work for our Dirac equation as we cannot get copies of an appropriate $y$ on both sides of the inequality. The positivity Tzvetkov got in the right-hand side of (2.96) allowed him to use Hölder's inequality to get a copy of $y$ on the right-hand side. However, with the particular structure of our nonlinearity that positivity does not exist on the right-hand side while simultaneously having the left-hand side result in a pure time derivative. Another issue with using this approach in our equation is the fact that for our equation the $L^2$ norm is a conserved quantity. Thus $\partial_t \int |U|^2 d\mathbf{x} = 0$ immediately. I have tried working with introducing multipliers, $\partial_t \int m(\mathbf{x}) |U|^2 d\mathbf{x}$, which would avoid the problem of the conservation law. However, none of these have resulted in a right-hand side that can be tied back into an ODE.

Another reason Tzvetkov's work fails to apply in our case is that his Dirac operator is not the same as ours. Tzvetkov's equation does not have $L^2$ conservation as can be seen in (2.96). The lack of conservation dramatically changes the dynamics of the system. However, it is still useful to consider how ODE blow-up techniques might be applied.

Concluding this section, to this point, we lack sufficient evidence to indicate that solutions to the Dirac equation, (0.1), either blow up in finite time or have some kind of global existence result. This is left as an open question.

# CHAPTER 3

# Application to the Schrödinger Ansatz

Now we will use the results discovered for solutions to the nonlinear Dirac system to examine how these yield solutions to the nonlinear Schrödinger equation from which it is was derived. Primarily in this thesis we have concerned ourselves with the particular form of the Dirac equation derived by Ablowitz, Nixon, and Zhu [1]. However, for this section we will mainly be working with the formulation derived by Fefferman and Weinstein [31]. Fefferman and Weinstein derive their formulation for the nonlinear Dirac equation by assuming an ansatz solution for the cubic nonlinear Schrödinger equation with periodic lattice potential. The ansatz they use takes a linear combination of components of solutions to the Dirac system multiplied by eigenfunctions of the periodic Schrödinger operator along with an error term. In [32], they do the same for the linear Schrödinger equation with periodic lattice and rigorously prove that the error term, $\eta$, is small on a large timescale. In [31], Fefferman and Weinstein plug in the ansatz to derive a differential equation describing the behavior of $\eta$. Using the linear behavior of $\eta$ (and setting aside the nonlinearity in $\eta$) they use a similar process as they used in [32] to determine the nonlinear Dirac equation that allows them to cancel out resonant interactions allowing $\eta$ to remain small. However, they do not include the rigorous detail as they did in [32], and their argument does not yet account for the nonlinear behavior of $\eta$.

In this section, we will provide the rigorous details similar to those used in [32] with an emphasis on where the proofs differ as well as provide a bootstrapping argument to account for the nonlinear behavior of $\eta$. This will result in a smaller timescale than Fefferman and Weinstein had in [32] for the linear Schrödinger equation but more restriction on the actual size of $\eta$.

For ease of reference, during the course of this section we will switch to the notation used in [31, 32].

In Fefferman Weinstein [31] they seek solutions to the nonlinear Schrödinger/Groff-Pitaevskii equation, NLS-GP, (or just NLS for simplicity)

$$(3.1) \qquad i\partial_t \psi = (-\Delta + V(x))\psi + g|\psi|^2\psi$$

of the form

$$(3.2) \qquad \psi^\delta(\mathbf{x}, t) = \left( \delta^{1/2} \sum_{j=1}^2 \alpha_j(\mathbf{X}, T)\Phi_j(\mathbf{x}) + \eta^\delta(\mathbf{x}, t) \right) e^{-i\mu_* t}$$

with initial data

$$(3.3) \qquad \psi^\delta(\mathbf{x}, 0) = \delta^{1/2} \sum_{j=1}^2 \alpha_j(\mathbf{X}, 0)\Phi_j(\mathbf{x})$$

where $\alpha = (\alpha_j)_{j=1}^2$ is an unknown slowly modulating function, $\Phi_1, \Phi_2, V(\mathbf{x})$, and $\mu_*$ to be defined in the upcoming section, $\eta^\delta$ is the error term (normally written simply as $\eta$), and $\mathbf{X} = \delta\mathbf{x}$, $T = \delta t$.

### 3.1. Survey of the Relevant Floquet-Bloch Theory.
Before proceeding any further, as in [32] and [9], we list some of the needed results from Floquet-Bloch theory.

We consider the lattice, $\Lambda$, created by the span of the vectors

$$(3.4) \qquad \mathbf{v}_1 = a \begin{pmatrix} \frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{pmatrix}, \quad \mathbf{v}_2 = a \begin{pmatrix} \frac{\sqrt{3}}{2} \\ -\frac{1}{2} \end{pmatrix}, \quad a > 0.$$

The honeycomb structure we consider is formed by the union of two sub-lattices created by applying integer multiples of $\mathbf{v}_i$ to the starting points $A, B$ chosen such that the result is invariant under rotations by $2\pi/3$. We also consider the dual lattice, $\Lambda^*$, formed by the vectors

$$(3.5) \qquad \mathbf{k}_1 = b \begin{pmatrix} \frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{pmatrix}, \quad \mathbf{k}_2 = a \begin{pmatrix} \frac{1}{2} \\ -\frac{\sqrt{3}}{2} \end{pmatrix}, \quad q = \frac{4\pi}{a\sqrt{3}}.$$

The fundamental period cell of the lattice is given by

$$(3.6) \qquad \Omega = \{\theta_1\mathbf{v}_1 + \theta_2\mathbf{v}_2 : 0 \le \theta_j \le 1\}.$$

We also make a choice of the fundamental unit cell of the dual lattice centered at $\mathbf{K} := \frac{1}{3}(\mathbf{k}_1 + \mathbf{k}_2)$ which we call the Brillouin zone, $\mathcal{B}_h$. It is shown in [32] that $\mathbf{K}$ is actually a Dirac point (which we will define shortly).

We call a function, $f$, $\Lambda$-periodic if $f(\mathbf{x} + \mathbf{v}) = f(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^2$ and $\mathbf{v} \in \Lambda$ and denote $f \in L^2_{per,\Lambda}$. We also call a function $\mathbf{k}$-pseudo periodic if $f(\mathbf{x} + \mathbf{v}) = e^{i\mathbf{k}\cdot\mathbf{v}} f(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^2$ and $\mathbf{v} \in \Lambda$ and denote $f \in L^2_{k,\Lambda}$. If the choice of lattice is clear, then the $\Lambda$ is dropped from the notation.

We also define precisely our potential function, $V(\mathbf{x})$.

DEFINITION 3. Let $V \in C^\infty(\mathbb{R}^2 \to \mathbb{R})$. We call $V$ a *honeycomb lattice potential* if there exists $\mathbf{x}_0 \in \mathbb{R}^2$ such that $V_{\mathbf{x}_0}(\mathbf{x}) = V(\mathbf{x} - \mathbf{x}_0)$ has the following properties:

(1) $V_{\mathbf{x}_0}$ is $\Lambda$-periodic.

(2) $V_{\mathbf{x}_0}(-\mathbf{x}) = V_{\mathbf{x}_0}(\mathbf{x})$

(3) $V_{\mathbf{x}_0}$ is invariant under the $2\pi/3$ rotation matrix, $\mathcal{R}$.

From now on we will assume that $V$ is such a potential.

We define the operator $H = -\Delta + V(\mathbf{x})$ and consider the $\mathbf{k}$-pseudo periodic Floquet-Bloch eigenvalue problem,

(3.7)
$$H\Phi(\mathbf{y}; \mathbf{k}) = \mu(\mathbf{k})\Phi(\mathbf{y}; \mathbf{k}) \quad \mathbf{x} \in \mathbb{R}^2$$
$$\Phi(\mathbf{y} + \mathbf{v}; \mathbf{k}) = e^{i\mathbf{k}\cdot\mathbf{v}}\Phi(\mathbf{y}; \mathbf{k}) \quad \mathbf{v} \in \Lambda.$$

The above problem has a discrete, real-valued spectrum given by:

(3.8)
$$\mu_1(\mathbf{k}) \le \mu_2(\mathbf{k}) \le \mu_3(\mathbf{k}) \le \dots$$

along with their associated solutions, $\Phi_j(\mathbf{x}, \mathbf{k})$ for $j \in \mathbb{N}$. It is also known that $\mu_b(\mathbf{k}) \approx b$ for large $b$ uniformly for $\mathbf{k} \in \mathcal{B}_h$. The solutions, $\Phi_j$, are referred to as Bloch modes, and the graphs of $\mu_j(\mathbf{k})$ within the dual unit cell are call dispersion surfaces. Furthermore, the Bloch modes, $\Phi_j$, are orthonormal to each other and are bounded.

Fefferman and Weinstein [32] also gives us a useful way to write the $H^s$ norm in terms of Floquet-Bloch theory:

(3.9)
$$\|f\|^2_{H^2(\mathbb{R}^2)} \approx \sum_{b \ge 1}(1 + b^2)^{\frac{s}{2}} \int_{\mathcal{B}_h} |\tilde{f}_b|^2 \, d\mathbf{k}.$$

We now define what we mean by a Dirac point

DEFINITION 4. Let $V$ be a smooth honeycomb lattice potential. We call $\mathbf{K} \in \mathcal{B}_h$ a Dirac point if there exists an integer $m_1 \ge 1$, a real number $\mu_*$, and $\lambda, \delta > 0$ such that:

(1) $\mu_*$ is a degenerate eigenvalue of $H$ with $\mathbf{K}$-pseudo periodic boundary conditions.

(2) $\dim \text{Nullspace}\,(H - \mu_*) = 2$.

(3) $\text{Nullspace}\,(H - \mu_*) = \text{span}\{\Phi_1(\mathbf{x}), \Phi_2(\mathbf{x})\}$ where $\Phi_1 \in L^2_{\mathbf{K},\tau}$ and $\Phi_2(\mathbf{x}) = \overline{\Phi_1(-\mathbf{x})}$ (where the $\tau$ indicates that $\mathcal{R}\Phi_1 = \tau\Phi_1$ where $\tau = e^{i2\pi/3}$).

(4) There exists Lipschitz functions $\mu_\pm$,

$$\mu_{m_1}(\mathbf{k}) = \mu_-(\mathbf{k}), \quad \mu_{m_1+1}(\mathbf{k}) = \mu_+(\mathbf{k}), \quad \mu_\pm(\mathbf{K}) = \mu_*,$$

and $E_\pm(\mathbf{k})$, defined for $|\mathbf{k} - \mathbf{K}| < \delta$, and $(\mathbf{k})$-pseudo periodic eigenfunctions of $H$, $\Phi_\pm(\mathbf{x}, \mathbf{k})$, with corresponding eigenvalues $\mu_\pm(\mathbf{k})$ such that

(3.10)
$$\mu_+(\mathbf{k}) - \mu_* = +\lambda|\mathbf{k} - \mathbf{K}|(1 + E_+(\mathbf{k}))$$

$$\mu_-(\mathbf{k}) - \mu_* = -\lambda|\mathbf{k} - \mathbf{K}|(1 + E_-(\mathbf{k}))$$

where $|E_\pm| \leq C|\mathbf{k} - \mathbf{K}|$ for some $C > 0$.

Fefferman and Weinstein, [**32**], define the constant

(3.11)
$$\lambda_\# := \sum_{\mathbf{m} \in \mathcal{S}} c(\mathbf{m})^2 \begin{pmatrix} 1 \\ i \end{pmatrix} \cdot (\mathbf{K}_* + m_1\mathbf{k}_1 + m_2\mathbf{k}_2)$$

where $\mathcal{S} \subset \mathbb{Z}^2$, $c(\mathbf{m})$ are the Fourier coefficients of $\Phi_1(\mathbf{x})$, and $\mathbf{K}_*$ is a vertex of $\mathcal{B}_h$. They prove that if Conditions 1, 2, and 3 above are met and $\lambda = |\lambda_\#| \neq 0$, then Condition 4 is also met.

Fefferman and Weinstein [**32**] also prove the existence of Dirac points for the operator, $H$, and that the $\mu_\pm$ actually correspond to the two lowest lying dispersion surface, $\mu_1(\mathbf{k})$ and $\mu_2(\mathbf{k})$ along with the corresponding Bloch modes $\Phi_1, \Phi_2$ equaling $\Phi_\pm$ respectively. It is also shown by Fefferman and Weinstein that the intersection of $\mu_\pm(\mathbf{k})$ at the Dirac point $\mathbf{K}$ is the only intersection between these two dispersion surfaces and there are no intersections between these two and any higher lying dispersion surfaces.

Another fact about the dispersion surfaces that we need from Fefferman and Weinstein is that there exists $\kappa_1, C_1$ such that for all $\mathbf{k} \in \mathcal{B}_h$ satisfying $|\mathbf{k} - \mathbf{K}| < \kappa_1$ and all $b \notin \{+, -\}$ we have $|\mu_b(\mathbf{k}) - \mu_*| \geq C_1$. In this chapter, $\kappa_1$ always refers to this choice of $\kappa_1$.

We also define the functions $p_\pm(\mathbf{x}; \mathbf{k})$ as in [**32**]

(3.12)
$$\Phi_\pm(\mathbf{x}; \mathbf{k}) = e^{i\mathbf{k} \cdot \mathbf{x}} p_\pm(\mathbf{x}; \mathbf{k}), \quad \langle p_a(\cdot; \mathbf{k}), p_b(\cdot; \mathbf{k}) \rangle = \delta_{ab}, \quad a, b \in \{+, -\}.$$

Letting $\mathbf{K}$ be a Dirac point, defining $\kappa = \mathbf{k} - \mathbf{K}$, and assuming that $|\kappa|$ is small, we can describe the behavior of $\Phi_\pm$ near a Dirac point by

$$(3.13) \qquad p_\pm(\mathbf{x}; \mathbf{k}) = \frac{1}{\sqrt{2}} \frac{\kappa_1 + i\kappa_2}{|\kappa|} p_1(\mathbf{x}) \pm \frac{1}{\sqrt{2}} p_2(\mathbf{x}) + \mathcal{O}_{H^2(\mathbb{R}^2/\Omega)}(|\kappa|)$$

where

$$(3.14) \qquad p_1(\mathbf{x}) = e^{-i\mathbf{K}_* \cdot \mathbf{x}} \Phi_1(\mathbf{x}), \quad p_2(\mathbf{x}) = e^{-i\mathbf{K}_* \cdot \mathbf{x}} \Phi_2(\mathbf{x}).$$

## 3.2. Further Preliminaries.

In [31], Fefferman and Weinstein consider the nonlinear Schrödinger equation

$$(3.15) \qquad i\partial_t \psi = (-\Delta + V(x))\psi + |\psi|^2 \psi$$

with ansatz

$$(3.16) \qquad \psi(\mathbf{x}, t) = \left( \delta^{\frac{1}{2}} \sum_{j=1}^{2} \alpha_j(\delta\mathbf{x}, \delta t) \Phi_j(\mathbf{x}) + \eta^\delta(\mathbf{x}, t) \right) e^{-i\mu_* t}$$

By following a similar procedure to [32] (but without the rigorous justification), Fefferman and Weinstein plug in the ansatz (3.16) into (3.15) and use the linear behavior of the resulting equation in $\eta^\delta$ to conclude that in order for the ansatz to solve (3.15) the functions, $\alpha$, must solve

$$(3.17) \qquad \begin{aligned} \partial_t \alpha_1 &= -\overline{\lambda_\#}(\partial_{x_1} + i\partial_{x_2})\alpha_2 - ig\left(\beta_1|\alpha_1|^2 + 2\beta_2|\alpha_2|^2\right)\alpha_1 \\ \partial_t \alpha_2 &= -\lambda_\#(\partial_{x_1} - i\partial_{x_2})\alpha_1 - ig\left(2\beta_2|\alpha_1|^2 + \beta_1|\alpha_2|^2\right)\alpha_2 \end{aligned}$$

where $\beta_1$, and $\beta_2$ are finite constants defined as integrals of the Bloch modes $\Phi_1, \Phi_2$. We discovered that some of the Floquet-Bloch machinery given in [32] fails to give the needed bounds with this ansatz (although the governing Dirac equation is correct). We will have to make use of a multiscale ansatz to adjust for this.

In [2], Ablowitz and Zhu derive a similar Dirac equation in the tight-binding setting. Here, we introduce the nonlinear Schrödinger equation and ansatz of Ablowitz and Zhu to the analytic framework of Fefferman and Weinstein. The Schrödinger equation can be written as

$$(3.18) \qquad i\partial_t \psi = (-\Delta + V(x))\psi + \delta|\psi|^2 \psi$$

62

with an ansatz of the form

$$(3.19) \qquad \psi^\delta(\mathbf{x}, t) = \left( \sum_{j=1}^{2} \alpha_j(\delta^{1-a}\mathbf{x}, \delta t)\Phi_j(\mathbf{x}) + \eta^\delta(\mathbf{x}, t) \right) e^{-i\mu_* t}$$

with initial data

$$(3.20) \qquad \psi^\delta(\mathbf{x}, 0) = \sum_{j=1}^{2} \alpha_j(\delta^{1-a}\mathbf{x}, 0)\Phi_j(\mathbf{x})$$

where $0 < a < 1$. Also, they assume $\alpha$ is a solution to a slightly different Dirac equation than the one that we have more directly dealt with in the rest of this section.

In other words, Ablowitz and Zhu use a $\delta$ factor on the Schrödinger nonlinearity to create decay instead of on the amplitude of the ansatz, and they also use a different scaling between the space variable and the time variable within the ansatz. This is done because they work in the tight binding setting when the $\lambda_\#$ from Fefferman and Weinstein gets close to 0.

Another relevant piece of work to discuss is by Arbunich and Sparber, [9]. In that work, they use a different framework and approach the NLS from a macroscopic reference trying to find an ansatz solution to the equation,

$$(3.21) \qquad i\delta\partial_t\Psi^\delta = -\delta^2\Delta\Psi^\delta + V\left(\frac{\mathbf{x}}{\delta}\right)\Psi^\delta \pm \delta|\Psi^\delta|^2\Psi^\delta,$$

which is equivalent to (3.1) with $g = 1$ once adjusted for the different reference scaling. Arbunich and Sparber perform a multiscale expansion for their ansatz. The first order expansion they use is the same as the ansatz proposed by Fefferman and Weinstein in [31] where $\alpha$ is a solution of (3.17). They also use a second order expansion giving another amplitude, $\beta$, which is defined by a formally similar Dirac equation which is both massive and linear in $\beta$. They use some of the results from [32] to arrive at their main theorem which states their multiscale ansatz solution is accurate to an order of $\delta^2$ on any constant timescale for which the nonlinear Dirac solutions hold.

We will use the refined Fourier analysis from [32] along with the multiscale expansion strategy used in [9] to show that we only need to project off of a smaller component within the multiscale expansion than is done in [9].

We consider solutions of the nonlinear Schrödinger equation

$$(3.22) \qquad i\partial_t\psi = (-\Delta + V(x))\psi + |\psi|^2\psi$$

with an ansatz of the form

$$\psi(\mathbf{x}, t) = \left( \delta^{\frac{1}{2}} \sum_{j=1}^{2} \alpha_j(\delta\mathbf{x}, \delta t)\Phi_j(\mathbf{x}) + \delta^{\frac{3}{2}} \sum_{j=1}^{2} \beta_j(\delta\mathbf{x}, \delta t)\Phi_j(\mathbf{x}) \right.$$

(3.23)

$$\left. + \delta^{\frac{3}{2}}\alpha^{\perp}(\mathbf{x}, t) + \eta^{\delta}(\mathbf{x}, t) \right) e^{-i\mu_* t}$$

with initial data

(3.24)
$$\psi(\mathbf{x}, 0) = \delta^{\frac{1}{2}} \sum_{j=1}^{2} \alpha_j(\delta\mathbf{x}, 0)\Phi_j(\mathbf{x}) + \delta^{\frac{3}{2}} \sum_{j=1}^{2} \beta_j(\delta\mathbf{x}, 0)\Phi_j(\mathbf{x}) + \delta^{\frac{3}{2}}\alpha^{\perp}(\mathbf{x}, 0)$$

where, similarly to in [**31**], $\alpha$ will be assumed to be a solution to the Dirac equation

(3.25)
$$\partial_t \alpha_1 = -\overline{\lambda_\#}(\partial_{x_1} + i\partial_{x_2})\alpha_2 - i\left(\beta_1|\alpha_1|^2 + 2\beta_2|\alpha_2|^2\right)\alpha_1$$

$$\partial_t \alpha_2 = -\lambda_\#(\partial_{x_1} - i\partial_{x_2})\alpha_1 - i\left(2\beta_2|\alpha_1|^2 + \beta_1|\alpha_2|^2\right)\alpha_2$$

where we assume $|\lambda_\#| \gg 0$ (which is true when we assume $V$ is not a tight binding potential).

Before we can define exactly what $\alpha^{\perp}$ is we need to define

(3.26)
$$\gamma_1(\mathbf{x}, t) = i\partial_t\alpha_1(\mathbf{x}, t), \quad \Psi_1(\mathbf{x}) = \Phi_1(\mathbf{x})$$

$$\gamma_2(\mathbf{x}, t) = i\partial_t\alpha_2(\mathbf{x}, t), \quad \Psi_2(\mathbf{x}) = \Phi_2(\mathbf{x})$$

$$\gamma_3(\mathbf{x}, t) = 2\nabla\alpha_1(\mathbf{x}, t), \quad \Psi_3(\mathbf{x}) = \nabla\Phi_1(\mathbf{x})$$

$$\gamma_4(\mathbf{x}, t) = 2\nabla\alpha_2(\mathbf{x}, t), \quad \Psi_4(\mathbf{x}) = \nabla\Phi_2(\mathbf{x})$$

and $\gamma_5, \Psi_5 \ldots \gamma_{13}, \Psi_{13}$ are a reindexing of $\sum_{j,k,l} \alpha_j\alpha_k\overline{\alpha_l}\Phi_j\Phi_k\overline{\Phi_l}$. For example, $\gamma_5 = \alpha_1\alpha_1\overline{\alpha_1}$ and $\Psi_5 = \Phi_1\Phi_1\overline{\Phi_1}$. Although for ease of use we will ignore the summation and work with the general

(3.27)
$$\gamma_5(\mathbf{x}, t) = \alpha_j(\mathbf{x}, t)\alpha_k(\mathbf{x}, t)\overline{\alpha_l}(\mathbf{x}, t), \quad \Psi_5(\mathbf{x}) = \Phi_j(\mathbf{x})\Phi_k(\mathbf{x})\overline{\Phi_l}(\mathbf{x})$$

similar to Einstein notation. We also define $\mathcal{P}_r(\mathbf{x}) = e^{-i\mathbf{K}_* \cdot \mathbf{x}}\Psi_r(\mathbf{x})$.

Now we can define

$$\alpha^{\perp}(\mathbf{x}, t) := (P_* - 1)(H - \mu_*)^{-1}(P_* - 1)(\alpha^{\perp, D} + \alpha^{\perp, D^c}),$$

(3.28)
$$\alpha^{\perp, D} = \sum_{\{+, -\}} \int_{\mathcal{B}_h} \left( \Phi_\pm(\mathbf{x}; \mathbf{k}) \sum_r \delta^{-2}\hat{\gamma}_r\left(\frac{\mathbf{k} - \mathbf{K}}{\delta}, \delta t\right) \int_\Omega \overline{p_\pm(\mathbf{y}, \mathbf{k})}\mathcal{P}_r(\mathbf{k})d\mathbf{y} \right) d\mathbf{k}$$

$$\alpha^{\perp, D^c} = \sum_{b \notin \{+, -\}} \int_{\mathcal{B}_h} \chi(|\mathbf{k} - \mathbf{K}| < \kappa_1)\left\langle \Phi_b(\cdot; \mathbf{k}), \sum_r \gamma_r(\delta\cdot, \delta t)\Psi_r(\cdot) \right\rangle_{L^2(\mathbb{R}^2)} \Phi_b(\mathbf{x}; \mathbf{k})d\mathbf{k}$$

where $P_*$ is the projection onto the eigenspace of $\mu_*$. We know that $\alpha^\perp$ is well defined because we have projected away from the degenerate eigenspace of $\mu_*$ implying the solvability of $(H - \mu_*)f = (P_* - 1)(\alpha^{\perp,D} + \alpha^{\perp,D^c})$ (for proof see Theorem 4.2.1 in Kuchment, [55]). We note that this choice of $\alpha^\perp$ is equivalent to the perpendicular component of the multiscale expansion in [9] when it is carefully localized to certain regions (to be shown in the proof). In other words, when compared to [9], we will use $\alpha^\perp$ to cancel out fewer terms of our expansion and instead rely on the Floquet-Bloch theory in [32] to show that the remaining terms are well behaved.

We will also assume that $\beta$ is a solution to the forced, inhomogeneous Dirac equation

(3.29)
$$\partial_t \beta_1 = -\overline{\lambda_\#}(\partial_{x_1} + i\partial_{x_2})\beta_2 + i\Delta\alpha_1 + i\langle \Phi_1, \partial_T \alpha^\perp \rangle_{L^2(\Omega)}$$
$$\partial_t \beta_2 = -\lambda_\#(\partial_{x_1} - i\partial_{x_2})\beta_1 + i\Delta\alpha_2 + i\langle \Phi_2, \partial_T \alpha^\perp \rangle_{L^2(\Omega)}.$$

Before stating the main theorem of this section we fully expand the ansatz (3.23) into the NLS equation (3.22) to get

$$i\partial_t \eta - (H - \mu_*)\eta =$$

$$- \delta^{\frac{3}{2}}\Big( \sum_{j=1}^{2} i\partial_T \alpha_j \Phi_j + 2\sum_{j=1}^{2} \nabla_{\mathbf{X}}\alpha_j \cdot \nabla_{\mathbf{x}}\Phi_j - \sum_{j,k,l=1}^{2} \alpha_j \alpha_k \overline{\alpha_l} \Phi_j \Phi_k \overline{\Phi_l} + (H-\mu_*)\alpha^{\perp} \Big)$$

$$- \delta^{\frac{5}{2}}\Big( \sum_{j=1}^{2} i\partial_T \beta_j \Phi_j + 2\sum_{j=1}^{2} \nabla_{\mathbf{X}}\beta_j \cdot \nabla_{\mathbf{x}}\Phi_j + \sum_{j=1}^{2} \Delta\alpha_j \Phi_j + \sum_{j=1}^{2} \partial_T \alpha^{\perp} \Big)$$

$$+ \delta\Big( 2\eta \sum_{j,k=1}^{2} \alpha_j \overline{\alpha_k} \Phi_j \overline{\Phi_k} + \overline{\eta} \sum_{j,k=1}^{2} \alpha_j \alpha_k \Phi_j \Phi_k \Big)$$

$$+ \delta^{\frac{1}{2}}\Big( 2|\eta|^2 \sum_{j=1}^{2} \alpha_j \Phi_j + (\eta)^2 \sum_{j=1}^{2} \overline{\alpha_j}\overline{\Phi_j} \Big) \quad + \quad |\eta|^2\,\eta$$

(3.30)
$$+ \delta^{\frac{5}{2}}\Big( 2\sum_{j,k,l=1}^{2} \alpha_j \overline{\alpha_k}\beta_l \Phi_j \overline{\Phi_k}\Phi_l + \sum_{j,k,l=1}^{2} \alpha_j \alpha_k \overline{\beta_l}\Phi_j \Phi_k \overline{\Phi_l} \Big)$$

$$+ \delta^{2}\Big( \sum_{j,k,l=1}^{2} \alpha_j \overline{\beta_k}\zeta_l \Phi_j \overline{\Phi_k} + \sum_{j,k,l=1}^{2} \overline{\alpha_j}\beta_k \zeta_l \overline{\Phi_j}\Phi_k + \sum_{j,k,l=1}^{2} \alpha_j \beta_k \overline{\zeta_l}\Phi_j \Phi_k \Big)$$

$$+ \delta^{\frac{3}{2}}\Big( 2\sum_{j,k,l=1}^{2} \beta_j \zeta_k \overline{\zeta_l}\Phi_j + \sum_{j,k,l=1}^{2} \overline{\beta_k}\zeta_k \zeta_l \overline{\Phi_k} \Big)$$

$$+ \delta^{\frac{7}{2}}\Big( 2\sum_{j,k,l=1}^{2} \beta_j \overline{\beta_k}\alpha_l \Phi_j \overline{\Phi_k}\Phi_l + \sum_{j,k,l=1}^{2} \beta_j \beta_k \overline{\alpha_l}\Phi_j \Phi_k \overline{\Phi_l} \Big)$$

$$+ \delta^{3}\Big( 2\sum_{j,k,l=1}^{2} \beta_j \overline{\beta_k}\zeta_l \Phi_j \overline{\Phi_k} + \sum_{j,k,l=1}^{2} \beta_j \beta_k \overline{\zeta_l}\Phi_j \Phi_k \Big)$$

$$+ \delta^{\frac{9}{2}}\Big( 3\sum_{j,k,l=1}^{2} \beta_j \overline{\beta_k}\beta_l \Phi_j \overline{\Phi_k}\Phi_l \Big).$$

where $\zeta_1 = \eta, \zeta_2 = \delta^{\frac{3}{2}}\alpha^{\perp}$, $\eta(\mathbf{x},0) = 0$, each $\alpha_j$ and $\beta_j$ are evaluated at $(\delta\mathbf{x}, \delta t)$, and each $\Phi_j$ is evaluated at $(\mathbf{x})$.

Now we state the main theorem of this section:

THEOREM 3.2.1. *Assume that $\alpha$ is a solution to (3.25) and $\beta$ is a solution to (3.29) with zero initial data. Consider the NLS equation (3.22) where $V$ is a honeycomb lattice potential. Fix $A_0 > 12$. Assume initial conditions, $\psi_0$, of the form (3.24) where $\alpha(\cdot, 0) \in H^s$ for large $s$ and has compact support. Then for any $\rho > 0$ there exists $\delta_0$ such that there is a unique solution to (3.22)*

*of the form* (3.23) *with*

$$(3.31) \qquad \sup_{0 \le t \le \rho\delta^{-1}} \|\eta^\delta(\mathbf{x}, t)\|_{H^2(\mathbb{R}^2)} = \mathcal{O}(\delta), \ \textit{for } \delta \to 0$$

*for all* $\delta \le \delta_0$.

We note that the constant implied by $\mathcal{O}$ only depends on $\rho$, $A_0$, the size of the initial data, and the size of the support of the initial data.

REMARK 2. As we discuss in the first paragraph of the proof, the envelopes, $\alpha$ are only evaluated in time up to the constant time, $\rho$. This means we are implicitly assuming that $\rho$ is within the lifespan of the Dirac solutions, $\alpha$. If we assume the size of the initial data is small enough that the almost global results hold, then that allows $\rho$ to be exponentially large. Otherwise we are limited by the local existence lifespans discussed in Chapter 2.

REMARK 3. Using the same argument as we use in the upcoming proof, we were unable to prove the equivalent theorem without the multiscale expansion. In particular, we could not prove the theorem without an $\alpha^\perp$ which cancels out potentially harmful interactions near the Dirac point on the $\mu_*$-eigenspace. In other words we were unable to show that nonlinearity *cannot* pump energy from the first two Bloch modes into other neighboring Bloch modes. By using $\alpha^\perp$ we are able to prove that this potentially harmful pumping of energy to other Blochs would have to come from the components eliminated by $\alpha^\perp$.

**3.3. Beginning the Proof.** We start by considering the lifespan of solutions to the Dirac equation. Within the ansatz, $\alpha$ is evaluated at $\delta t$, and by assumption in Theorem 3.2.1 we have $t \le \rho\delta^{-1}$. Thus the time values for which $\alpha$ is evaluated stay less than $\delta\rho\delta^{-1} = \rho$. We will choose the value $T^* = \rho$ to serve as an upper bound on the Dirac lifespan. Thus, solutions to the Dirac equation, $\alpha$, only have finite time, $\rho$, to grow, and $\|\alpha\|_{H^s}$ is bounded by some constant, $C$, depending only on the size of the initial data, the uniform time bound, $\rho$, and the regularity, $s$ (within the proof we will be using $A_0$ to define the regularity, so the $s$ dependence will be replaced by $A_0$).

Next, we clearly define a Picard iteration for $\eta$. In doing so we will also use the notation $\mathbf{X} = \delta\mathbf{x}$ and $T = \delta t$. We define $\eta_0 = 0$, and instead of working with the Picard iteration of $\eta$ in the form a

differential equation we use the integral equation corresponding to (3.30) and define

$$
i\eta_{n+1}(\cdot, t) =
$$

$$
\begin{aligned}
&- \delta^{\frac{3}{2}} \sum_{j=1}^{2} \int_0^t e^{-i(H-\mu_*)(t-s)} \bigg( i\partial_T \alpha_j(\delta\cdot, \delta s)\Phi_j + 2\nabla_{\mathbf{X}}\alpha_j(\delta\cdot, \delta s) \cdot \nabla_{\mathbf{x}}\Phi_j \\
&\qquad\qquad - \sum_{k,l=1}^{2} \alpha_j(\delta\cdot, \delta s)\alpha_k(\delta\cdot, \delta s)\overline{\alpha_l}(\delta\cdot, \delta s)\Phi_j\Phi_k\overline{\Phi_l} + (H-\mu_*)\alpha^\perp(\mathbf{x}, t) \bigg) ds \\
&- \delta^{\frac{5}{2}} \sum_{j=1}^{2} \int_0^t e^{-i(H-\mu_*)(t-s)} \bigg( i\partial_T \beta_j(\delta\cdot, \delta s)\Phi_j + 2\nabla_{\mathbf{X}}\beta_j(\delta\cdot, \delta s) \cdot \nabla_{\mathbf{x}}\Phi_j \\
&\qquad\qquad - \Delta_{\mathbf{X}}\alpha_j(\delta\cdot, \delta s)\Phi_j - \partial_T \alpha^\perp \bigg) ds \\
&+ \delta \sum_{j,k=1}^{2} \int_0^t e^{-i(H-\mu_*)(t-s)} \bigg( 2\eta_n(\cdot, s)\alpha_j(\delta\cdot, \delta s)\overline{\alpha_k}(\delta\cdot, \delta s)\Phi_j\overline{\Phi_k} \\
&\qquad\qquad + \overline{\eta_n}(\cdot, s)\alpha_j(\delta\cdot, \delta s)\alpha_k(\delta\cdot, \delta s)\Phi_j\Phi_k \bigg) ds \\
&+ \delta^{\frac{1}{2}} \sum_{j=1}^{2} \int_0^t e^{-i(H-\mu_*)(t-s)} \bigg( 2|\eta_n(\cdot, s)|^2 \alpha_j(\delta\cdot, \delta s)\Phi_j \\
&\qquad\qquad + \eta_n(\cdot, s)^2\overline{\alpha_j}(\delta\cdot, \delta s)\overline{\Phi_j} \bigg) ds \\
&+ \int_0^t e^{-i(H-\mu_*)(t-s)} |\eta_n(\cdot, s)|^2 \eta_n(\cdot, s) ds + \text{ Lower order terms} \\
&:= -A_1 - A_2 + A_{3,n+1} + A_{4,n+1} + A_{5,n+1} + \text{ Lower order terms.}
\end{aligned}
$$

(3.32)

The first two terms of (3.32) are independent of $\eta$ and will be handled very similarly to Fefferman and Weinstein's handling of the linear Schrödinger equation in [**32**]. The more significant changes from their method will be more thoroughly explained and typically arise from the fact that for solutions to the nonlinear Dirac equation $|\alpha(\hat{\xi}, t)| \neq |\alpha(\hat{\xi}, 0)|$.

The final three terms listed in (3.32) are dependent on $\eta_n$ and are actually quite easy to deal with. We will show the $H^2$ norms of these terms can be conveniently bounded.

The lower order terms which are not listed have sufficient decay that simply using the closure of $H^2(\mathbb{R}^2)$ under multiplication and scaling the $H^2$ norm gives the needed bounds. These bounds are left out in the proof, but follow as simple corollaries to how we deal with the $\eta$ dependent terms listed.

The terms, $A_1$ and $A_2$ appear in each $\eta_n$ but since there is no $\eta$ dependence in those terms they are handled identically each time. However, the terms $A_{3,n+1}, A_{4,n+1}$, and $A_{5,n+1}$ depend on the previous iterations of $\eta_n$ and have to be handled differently. We will bound $\|\eta_n\|_{H^2(\mathbb{R}^2)}$ by induction.

## 3.4. Bounding the $\eta$ Independent Terms.

An important departure from the proof in [32] is due to the fact that $\|\alpha(x,t)\|_{H^s} \neq \|\alpha(x,0)\|_{H^s}$ when $\alpha$ is a solution for the nonlinear Dirac equation (equality of the two quantities is needed in the proof for the linear problem in [32]). However, the digression is very brief as $\|\alpha(x,0)\|_{H^s}$ merely serves the role of a finite bound. To find another bound to take its place we notice that $\|\alpha_j\|_{C([0,t],H^s(\mathbb{R}^2))}$ is bounded by $\|\alpha_j\|_{C([0,T^*],H^s(\mathbb{R}^2))}$. Within the proof we will find that the $A_0$ fixed in the statement of the theorem gives us sufficient regularity. Thus $\|\alpha_j\|_{C([0,T^*],H^{A_0}(\mathbb{R}^2))}$ will replace the role of $\|\alpha(x,0)\|_{H^s}$ as a uniform bound in our proof and is itself bounded by a constant dependent on $\rho$ and $A_0$.

For the term $A_1$ we seek to show that

$$\|A_1\|_{H^2} \leq C_1(\delta^{\frac{1}{2}} + \delta^{\frac{3}{2}}t).$$

We can consider the term $A_1$ as the solution to the initial value problem

$$i\partial_t f(\mathbf{x},t) - (H - \mu_*)f(\mathbf{x},t) = \delta^{\frac{3}{2}}\left(\sum_r \gamma_r(\delta\mathbf{x},\delta t)\Psi_r(\mathbf{x}) + (H - \mu_0)\alpha^\perp(\mathbf{x},t)\right)$$

(3.33)

$$f(\mathbf{x},0) = 0$$

where we remember the definition

$$
\begin{aligned}
\gamma_1(\mathbf{x},t) &= i\partial_t\alpha_1(\mathbf{x},t), & \Psi_1(\mathbf{x}) &= \Phi_1(\mathbf{x}) \\
\gamma_2(\mathbf{x},t) &= i\partial_t\alpha_2(\mathbf{x},t), & \Psi_2(\mathbf{x}) &= \Phi_2(\mathbf{x}) \\
\gamma_3(\mathbf{x},t) &= 2\nabla\alpha_1(\mathbf{x},t), & \Psi_3(\mathbf{x}) &= \nabla\Phi_1(\mathbf{x}) \\
\gamma_4(\mathbf{x},t) &= 2\nabla\alpha_2(\mathbf{x},t), & \Psi_4(\mathbf{x}) &= \nabla\Phi_2(\mathbf{x}) \\
\gamma_5(\mathbf{x},t) &= \alpha_j(\mathbf{x},t)\alpha_k(\mathbf{x},t)\overline{\alpha_l}(\mathbf{x},t), & \Psi_5(\mathbf{x}) &= \Phi_j(\mathbf{x})\Phi_k(\mathbf{x})\overline{\Phi_l}(\mathbf{x}).
\end{aligned}
$$

(3.34)

LEMMA 3.4.1. *For solutions, $f$, to (3.33) we can write*

$$f(\mathbf{x},t) = \sum_b \int_{\mathcal{B}_h} \tilde{f}_b(\mathbf{k},t)\Phi_b(\mathbf{x};\mathbf{k})d\mathbf{k} + (P_* - 1)\sum_{\{+,-\}}\int_{\mathcal{B}_h}\tilde{\alpha}^{\perp,D}(\mathbf{k},t)\Phi_\pm(\mathbf{x};\mathbf{k})d\mathbf{k}$$

(3.35)

$$+ (P_* - 1)\sum_{b\notin\{+,-\}}\int_{\mathcal{B}_h}\tilde{\alpha}^{\perp,D^c}(\mathbf{k},t)\Phi_b(\mathbf{x};\mathbf{k})d\mathbf{k}$$

69

*where the summation is finite because we have projected onto the $\mu_*$-eigenspace and*

$$\tilde{f}_b = -\delta^{\frac{3}{2}} \int_0^t e^{-i(\mu_b(\mathbf{k})-\mu_*)(t-s)} \langle \Phi_b(\cdot;\mathbf{k}), \sum_r \gamma_r(\delta\cdot,\delta s)\Psi_r(\cdot)\rangle_{L^2(\mathbb{R}^2)}ds.$$

$$\tilde{\alpha}^{\perp,D} = -\delta^{\frac{3}{2}} \int_0^t e^{-i(\mu_b(\mathbf{k})-\mu_*)(t-s)} \sum_r \delta^{-2}\hat{\gamma}_r\left(\frac{\mathbf{k}-\mathbf{K}}{\delta},\delta s\right)\int_\Omega \overline{p_\pm(\mathbf{y},\mathbf{k})}\mathcal{P}_r(\mathbf{k})d\mathbf{y}ds$$

(3.36)

$$\tilde{\alpha}^{\perp,D^c} = -\delta^{\frac{3}{2}} \int_0^t e^{-i(\mu_b(\mathbf{k})-\mu_*)(t-s)}\chi(|\mathbf{k}-\mathbf{K}|<\kappa_1)$$
$$\times \left\langle \Phi_b(\cdot;\mathbf{k}), \sum_r \gamma_r(\delta\cdot,\delta t)\Psi_r(\cdot)\right\rangle_{L^2(\mathbb{R}^2)}ds.$$

Note that typically, solutions, $f$, can be broken up along the summation into frequency components $\mu_+(\mathcal{B}_h), \mu_-(\mathcal{B}_h)$ which intersect at the degenerate eigenvalue $\mu_*$, and all other frequency components:

$$f_D(\mathbf{x},t) = \sum_{b\in\{+,-\}} \int_{\mathcal{B}_h} \tilde{f}_b(\mathbf{k},t)\Phi_b(\mathbf{x};\mathbf{k})d\mathbf{k}$$

$$f_{D^c}(\mathbf{x},t) = \sum_{b\notin\{+,-\}} \int_{\mathcal{B}_h} \tilde{f}_b(\mathbf{k},t)\Phi_b(\mathbf{x};\mathbf{k})d\mathbf{k}$$

(3.37)

$$\alpha_D^\perp(\mathbf{x},t) = (P_*-1) \sum_{\{+,-\}} \int_{\mathcal{B}_h} \tilde{\alpha}^{\perp,D}(\mathbf{k},t)\Phi_\pm(\mathbf{x};\mathbf{k})d\mathbf{k}$$

$$\alpha_{D^c}^\perp(\mathbf{x},t) = (P_*-1) \sum_{b\notin\{+,-\}} \int_{\mathcal{B}_h} \tilde{\alpha}^{\perp,D^c}(\mathbf{k},t)\Phi_b(\mathbf{x};\mathbf{k})d\mathbf{k}$$

where $f_D(\mathbf{x},t)$ and $\alpha_D^\perp(\mathbf{x},t)$ are summed over the same components but kept seperate in the notation, and the same is true for $f_{D^c}$ and $\alpha_{D^c}^\perp$. We also note that $\alpha_{D^c}^\perp$ is already projected off of the $\mu_*$-eigenspace. Thus the $(P_*-1)$ within the definition of $\alpha_{D^c}^\perp$ can just be replaced by $(-1)$.

Note that $\mu_b(\mathbf{k})$ and $\Phi(\mathbf{x},\mathbf{k})$ are functions of $\mathbf{k}\in\mathcal{B}_h$, the Brillouin zone. When we just write $\mu_b$ and $\Phi(\mathbf{x})$ we are assuming that $\mathbf{k}=\mathbf{K}$ where $\mathbf{K}$ is a Dirac point.

Before we begin estimating $f_D$ we further decompose into frequencies pieces which are close to and far from the Dirac point, $K$:

$$f_D(\mathbf{x},t) = \sum_{b\in\{+,-\}} \int_{\mathcal{B}_h} \chi\left(|\mathbf{k}-\mathbf{K}|<\delta^\tau\right)\tilde{f}_b(\mathbf{k},t)\Phi_b(\mathbf{x};\mathbf{k})d\mathbf{k}$$

(3.38)
$$+ \sum_{b\in\{+,-\}} \int_{\mathcal{B}_h} \chi\left(\delta^\tau\le|\mathbf{k}-\mathbf{K}|\right)\tilde{f}_b(\mathbf{k},t)\Phi_b(\mathbf{x};\mathbf{k})d\mathbf{k}$$

$$:= f_{I,D}(\mathbf{x},t) + f_{II,D}(\mathbf{x},t)$$

70

where we choose $\frac{3}{4} < \tau < \frac{A_0-3}{A_0}$. We also decompose

$$\alpha_D^\perp(\mathbf{x},t) = (P_* - 1) \sum_{b \in \{+,-\}} \int_{\mathcal{B}_h} \chi\Big(|\mathbf{k} - \mathbf{K}| < \delta^\tau\Big) \tilde{\alpha}^{\perp,D}(\mathbf{k},t) \Phi_b(\mathbf{x};\mathbf{k}) d\mathbf{k}$$

(3.39)
$$+ (P_* - 1) \sum_{b \in \{+,-\}} \int_{\mathcal{B}_h} \chi\Big(\delta^\tau \le |\mathbf{k} - \mathbf{K}|\Big) \tilde{\alpha}^{\perp,D}(\mathbf{k},t) \Phi_b(\mathbf{x};\mathbf{k}) d\mathbf{k}$$

$$:= \alpha_{I,D}^\perp(\mathbf{x},t) + \alpha_{II,D}^\perp(\mathbf{x},t).$$

Next we state a lemma which is proven without the $\alpha^\perp$ terms in Fefferman and Weinstein (Equation 7.7 in [**32**]) relying on (3.9) and the fact that $f_D$ is only a summation over finitely many $b$, in particular $b = \pm$.

LEMMA 3.4.2. *For $f$ defined and decomposed as above we have that*

(3.40)
$$\|f(\cdot,t)\|_{H^2(\mathbb{R}^2)}^2 \approx \|f_{I,D}(\cdot,t) + \alpha_{I,D}^\perp(\cdot,t)\|_{L^2(\mathbb{R}^2)}^2 + \|f_{II,D}(\cdot,t) + \alpha_{II,D}^\perp(\cdot,t)\|_{L^2(\mathbb{R}^2)}^2$$

$$+ \|f_{D^c}(\cdot,t) + \alpha_{D^c}^\perp(\mathbf{x},t)\|_{H^2(\mathbb{R}^2)}^2.$$

The proof relies on (3.9) and the fact that $f_D$ has only two spectral bands to sum over in order to now have $L^2$ norms on the $F_D$ terms on the right-hand side. Since $\alpha_D^\perp$ also only has two spectral bands to sum over, the proof is no different here.

The last thing we need before handling the two terms separately is Proposition 7.1 from [**32**]:

PROPOSITION 3.4.3. *Let $\Gamma(\mathbf{x},t)$ be a Schwartz class function of $\mathbf{x}$, varying smoothly in $t$. Denote, $\widehat{\Gamma}(\xi,t)$ its Fourier transform with respect to $\mathbf{x}$. Let $\Psi(\mathbf{x}) = e^{i\mathbf{K}\cdot\mathbf{x}}\mathcal{P}(\mathbf{x})$ where $\mathcal{P}(\mathbf{x}+\mathbf{v}) = \mathcal{P}(\mathbf{x})$ when $\mathbf{v}$ is a lattice point, and $\Phi_b(\mathbf{x};\mathbf{k}) = e^{i\mathbf{k}\cdot\mathbf{x}}p(\mathbf{x};\mathbf{k})$ where $p(\mathbf{x}+\mathbf{v};\mathbf{k}) = p(\mathbf{x};\mathbf{k})$ when $\mathbf{v}$ is a lattice point. Then,*

$$\langle \Phi_b(\cdot;\mathbf{k}), \Gamma(\delta\cdot,\delta s)\Psi(\cdot)\rangle_{L^2(\mathbb{R}^2)} =$$

(3.41)
$$\int_\Omega \overline{p_b(\mathbf{y};\mathbf{k})}\Big[\delta^{-2}\sum_{\mathbf{m}\in\mathbb{Z}^2} e^{i\mathbf{m}\cdot\mathbf{y}}\widehat{\Gamma}\Big(\frac{m_1\mathbf{k}_1 + m_2\mathbf{k}_2 + (\mathbf{k}-\mathbf{K})}{\delta}, \delta s\Big)\Big]\mathcal{P}(\mathbf{y})d\mathbf{y}.$$

The assumption of $\Gamma$ being Schwartz class can be relaxed (as done in [**32**]) as long as the Fourier transform and inverse Fourier transform of $\Gamma$ are defined.

**3.4.1. Estimation of** $\|f_{I,D}(\cdot,t) + \alpha_{I,D}^\perp(\cdot,t)\|_{L^2(\mathbb{R}^2)}$. In this part, we will prove that

$$\|f_{I,D}(\cdot,t) + \alpha_{I,D}^\perp(\cdot,t)\|_{L^2(\mathbb{R}^2)} \lesssim \delta^{\frac{5}{2}}t.$$

As in [**32**] we get

$$\|f_{I,D}(\cdot,t) + \alpha^{\perp}_{I,D}(\cdot,t)\|^2_{L^2(\mathbb{R}^2)} \leq \sum_{b\in\{+,-\}} \int_{\mathcal{B}_h} \chi\Big(|\mathbf{k}-\mathbf{K}|<\delta^{\tau}\Big)|\tilde{f}_b(\mathbf{k},t)|^2 d\mathbf{k}$$

(3.42)

$$+ \sum_{b\in\{+,-\}} \int_{\mathcal{B}_h} \chi\Big(|\mathbf{k}-\mathbf{K}|<\delta^{\tau}\Big)|(P_* - 1)\tilde{\alpha}^{\perp,D}(\mathbf{k},t)| d\mathbf{k}$$

where, as before,

$$\tilde{f}_{\pm} = -\delta^{\frac{3}{2}} \int_0^t e^{-i(\mu_{\pm}(\mathbf{k})-\mu_*)(t-s)} \langle \Phi_{\pm}(\cdot;\mathbf{k}), \sum_r \gamma_r(\delta\cdot,\delta s)\Psi_r(\cdot)\rangle_{L^2(\mathbb{R}^2)} ds$$

(3.43)

$$\tilde{\alpha}^{\perp,D} = -\delta^{\frac{3}{2}} \int_0^t e^{-i(\mu_b(\mathbf{k})-\mu_*)(t-s)} \sum_r \delta^{-2}\hat{\gamma}_r\Big(\frac{\mathbf{k}-\mathbf{K}}{\delta},\delta s\Big) \int_{\Omega} \overline{p_{\pm}(\mathbf{y},\mathbf{k})}\mathcal{P}_r(\mathbf{k}) d\mathbf{y} ds.$$

Now we can use Proposition 3.4.3 to expand the inner product, and further separate the summation by pulling out the $\mathbf{m} = 0$ term:

$$\langle \Phi_{\pm}(\cdot;\mathbf{k}), \sum_r \gamma_r(\delta\cdot,\delta s)\Psi_r(\cdot)\rangle_{L^2(\mathbb{R}^2)}$$

$$= \delta^{-2} \sum_r \hat{\gamma}_r\Big(\frac{(\mathbf{k}-\mathbf{K})}{\delta},\delta s\Big) \int_{\Omega} \overline{p_{\pm}(\mathbf{y},\mathbf{k})}\mathcal{P}_r(\mathbf{y}) d\mathbf{y}$$

(3.44)

$$+ \sum_r \int_{\Omega} \delta^{-2} E_{r,\delta^c}\Big(\frac{\mathbf{y}}{\delta},\delta s;\mathbf{k}\Big)\overline{p_{\pm}(\mathbf{y},\mathbf{k})}\mathcal{P}_r(\mathbf{y}) d\mathbf{y}$$

$$:= \sum_r f_{1,r} + \sum_r f_{2,r}$$

where

(3.45) $$\quad E_{r,\delta}\Big(\frac{\mathbf{y}}{\delta},\delta s;\mathbf{k}\Big) = \sum_{m\in\mathbb{Z}^2\setminus\{(0,0)\}} e^{i\mathbf{m}\cdot\mathbf{y}}\hat{\gamma}_r\Big(\frac{m_1\mathbf{k}_1 + m_2\mathbf{k}_2 + (\mathbf{k}-\mathbf{K})}{\delta},\delta s\Big).$$

Let us stop for a moment and denote $f^{1,r}_{I,D}$ to be the $(\sum_r f_{1,r})$-component of $f_{I,D}$. We can notice that $(P_* - 1)f^{1,r}_{I,D} = \alpha^{\perp}_{I,D}$. Most importantly, this implies that $f^{1,r}_{I,D} + \alpha^{\perp}_{I,D} = P_*(f^{1,r}_{I,D})$. Thus, moving forward we will absorb the $\alpha^{\perp}_{I,D}$ term into the $f^{1,r}_{I,D}$ term and just work with $P_*(\sum_r f_{1,r})$ instead of $\sum_r f_{1,r}$. In other words, the addition of the $\alpha^{\perp}_{I,D}$ term allows us to work with only the projection of $\sum_r f_{1,r}$ onto the $\mu_*$-eigenspace.

Now we state and prove a proposition replacing Proposition 7.2 in [**32**] which does not hold directly when $\alpha$ a solution to the nonlinear Dirac equation (instead of the linear Dirac equation).

PROPOSITION 3.4.4. *Let $\gamma_r$ be as previously defined in (3.26). Then:*

(1) *For all $\kappa \in \mathbb{R}^2$ and $0 < A \leq A_0$ we have*

(3.46)
$$|\hat{\gamma}_r(\kappa, t)| \lesssim |\kappa|^{-A}$$

(2) *For all $\mathbf{k} \in \mathcal{B}_h$ such that $|\mathbf{k} - \mathbf{K}| < \delta^\tau$ and $2 < A \leq A_0$ we have*

(3.47)
$$|E_{r,\delta}\left(\delta^{-1}\mathbf{y}, \delta t; \mathbf{k}\right)| \lesssim \delta^A$$

*where $t \leq \rho\delta^{-1}$.*

PROOF. Notice first that the $i\partial_t \alpha_j$ terms can be written as a linear combination of the $2\partial_{x_i} \alpha_j$ and $\alpha_j \alpha_k \overline{\alpha_l}$ terms. Thus we need only prove the proposition for $\gamma_r = \partial_x \alpha_j$ and $\gamma_r = \alpha_j \alpha_k \overline{\alpha_l}$.

Case 1: $\left(\gamma_r = \partial_x \alpha_j\right)$. We take the absolute value of $\hat{\gamma}_r$ and integrate by parts on the Fourier side, using $H^2 \subset L^\infty$, to get

$$
\begin{aligned}
|\hat{\gamma}_r(\kappa, t)| &= |\widehat{\partial_x \alpha_j}(\kappa, t)| \\
&\leq \frac{1}{|\kappa|^A} \|\widehat{\partial_x^{A+1} \alpha_j}\|_{C([0,T^*], L^\infty(\mathbb{R}^2))} \\
&= \frac{1}{|\kappa|^A} \sup_{0 \leq t \leq T^*} \|\widehat{\partial_x^{A+1} \alpha_j}(\cdot, t)\|_{L^\infty(\mathbb{R}^2)} \\
&\lesssim \frac{1}{|\kappa|^A} \sup_{0 \leq t \leq T^*} \|\widehat{\partial_x^{A+1} \alpha_j}(\cdot, t)\|_{H^2(\mathbb{R}^2)}.
\end{aligned}
$$

(3.48)

Now we need to state a brief lemma taking advantage of the finite speed of propagation for the Dirac solution and the compact support of the initial data.

LEMMA 3.4.5. *Let $f$ be a function such that $f(\mathbf{x}) = 0$ for all $|\mathbf{x}| \geq M$. Then,*

$$\|\hat{f}\|_{H^2(\mathbb{R}^2)} \lesssim M^2 \|f\|_{L^2(\mathbb{R}^2)}.$$

PROOF. Without loss of generality we can let $M > 1$. We expand

$$
\begin{aligned}
\|\hat{f}\|_{H^2(\mathbb{R}^2)} = & \|\hat{f}\|_{L^2(\mathbb{R}^2)} + \sum_{i=1}^{2} \|\partial_{\xi_i}\hat{f}\|_{L^2(\mathbb{R}^2)} + \sum_{i,j}^{2} \|\partial_{\xi_i}\partial_{\xi_j}\hat{f}\|_{L^2(\mathbb{R}^2)} \\
= & \|f\|_{L^2(\mathbb{R}^2)} + \sum_{i=1}^{2} \|x_i f\|_{L^2(\mathbb{R}^2)} + \sum_{i,j}^{2} \|x_i x_j f\|_{L^2(\mathbb{R}^2)} \\
\leq & \|f\|_{L^2(\mathbb{R}^2)} + 2M\|f\|_{L^2(\mathbb{R}^2)} + 4M^2\|f\|_{L^2(\mathbb{R}^2)} \\
\lesssim & M^2\|f\|_{L^2(\mathbb{R}^2)}.
\end{aligned}
$$
(3.49)

This concludes the lemma. $\qquad\square$

From the section in Chapter 1 on Finite Propagation Speed we know that the support of $\alpha(\mathbf{x}, t)$ is determined by the support of the initial conditions and the time, $t$, that the support has had to grow. Thus for each time $t$ we get

$$
(3.50) \qquad |\hat{\gamma}_r(\kappa, t)| \lesssim \frac{1}{|\kappa|^A} \sup_{0 \leq t \leq T^*} M_t^2 \|\partial_x^{A+1}\alpha_j(\cdot, t)\|_{L^2(\mathbb{R}^2)}
$$

where $M_t$ can grow as $t$ grows.

However, for all $t \leq T^*$, we know that $M_t \leq M_{T^*}$. Thus, instead of using the variable-time bound, $M_t^2$, we can use the uniform bound, $M_{T^*}^2$, which can be absorbed by the $\lesssim$ as follows,

$$
\begin{aligned}
|\hat{\gamma}_r(\kappa, t)| \lesssim & \frac{1}{|\kappa|^A} \sup_{0 \leq t \leq T^*} \|\partial_x^{A+1}\alpha_j(\cdot, t)\|_{L^2(\mathbb{R}^2)} \\
\leq & \frac{1}{|\kappa|^A} \sup_{0 \leq t \leq T^*} \|\alpha(\cdot, t)\|_{H^{A+1}(\mathbb{R}^2)},
\end{aligned}
$$
(3.51)

concluding Case 1 for the proof.

Case 2: $\left(\gamma_r = \alpha_j\alpha_k\overline{\alpha_l}\right)$. Following the same steps as Case 1 we get

$$
\begin{aligned}
|\hat{\gamma}_r(\kappa, t)| = & |\widehat{\alpha_j\alpha_k\overline{\alpha_l}}(\kappa, t)| \\
\lesssim & \frac{1}{|\kappa|^A} \sup_{0 \leq t \leq T} \|\left(\alpha_j\alpha_k\overline{\alpha_l}\right)(\cdot, t)\|_{H^A(\mathbb{R}^2)}.
\end{aligned}
$$
(3.52)

Since $H^m(\mathbb{R}^2)$ is an algebra for $m > 1$ we get that $\|\alpha_j\alpha_k\overline{\alpha_l}\|_{H^A(\mathbb{R}^2)} \lesssim \|\alpha\|_{H^A(\mathbb{R}^2)}^3$.

Thus we have completed the proof for Part 1 of the theorem.

The proof for Part 2 is identical to that in the proof of Proposition 7.2 in [32] except for substituting our decay estimate from Part 1 instead of their decay estimate. $\qquad\square$

We notice that the proof does require use of the size of the support of $\alpha$. However, since $T^*$ is an upper bound on the existence time of $\alpha(\cdot, t)$ and solutions, $\alpha$, have finite speed of propagation, we really only need the size of the support of $\alpha(\cdot, 0)$, which is assumed to be compact in the statement of the theorem. Thus the constant implied by $\lesssim$ in the lemma truly is a constant.

Applying Proposition 3.4.4 (with $A = A_0$) to the formula for $f_{2,r}$ immediately gives us

$$(3.53) \qquad |f_{2,r}| \lesssim \delta^{A_0 - 2}.$$

Since $A_0 \geq 12$ we get that $f_{2,r}$ has a contribution to $\|f_{I,D}\|_{L^2}$ of size $\delta^{\frac{5}{2}}$

The term $f_{1,r}$ requires a fair amount of extra work. In fact, we actually bound the sum, $\sum f_{1,r}$, because individually they do not give particularly useful bounds. In fact, as we discussed earlier, we are working with $P_*(\sum f_{1,r})$, the projection onto the $\mu_*$-eigenspace. We will see that $P_*(\sum f_{1,r})$ gives zero contribution to $\|f_{I,D}\|_{L^2}$. However, in order for the argument to more readily adapt to our term, $A_2$, we will not yet simplify the projection.

In Fefferman and Weinstein, [32], the summation is only 4 terms. Conveniently those 4 terms behave very similarly. When working in the region $|\mathbf{k} - \mathbf{K}| < \delta^\tau$, Equation 7.25 in [32] states

$$
(3.54) \quad
\begin{aligned}
\hat{\gamma}_r & \left( \frac{\mathbf{k} - \mathbf{K}}{\delta}, \delta s \right) \cdot \langle p_\pm(\cdot; \mathbf{k}), \mathcal{P}_r(\cdot) \rangle_{L^2(\Omega)} \\
&= \frac{1}{\sqrt{2}} \frac{\kappa_1 + i\kappa_2}{|\kappa|} \langle \Phi_1, \hat{\gamma}_r \left( \frac{\kappa}{\delta}, \delta s \right) \cdot \Psi_r \rangle_{L^2(\Omega)} \pm \frac{1}{\sqrt{2}} \langle \Phi_2, \hat{\gamma}_r \left( \frac{\kappa}{\delta}, \delta s \right) \cdot \Psi_r \rangle_{L^2(\Omega)} \\
&\quad + \mathcal{O}\left( |\kappa| \cdot \|\alpha_0\|_{W^{2,1}(\mathbb{R}^2)} \right).
\end{aligned}
$$

However, this used the $L^\infty$ preservation of $\hat{\alpha}$ which we do not have in the nonlinear case. To get the $W^{2,1}$ norm in the final term they used

$$
\left| \hat{\gamma}_r \left( \frac{\mathbf{k} - \mathbf{K}}{\delta}, \delta s \right) \right| = \left| \hat{\gamma}_r \left( \frac{\mathbf{k} - \mathbf{K}}{\delta}, 0 \right) \right| \lesssim \|\alpha_0\|_{W^{2,1}}.
$$

We can adjust for this problem by instead using $H^2 \subset L^\infty$ just as we did in Proposition 3.4.4 only without using integration by parts to get $\kappa$ decay. This results in:

$$(3.55) \qquad \left| \hat{\gamma}_r \left( \frac{\mathbf{k} - \mathbf{K}}{\delta}, \delta s \right) \right| \lesssim 1$$

for $\gamma_r = 2\partial_{x_i} \alpha_j$ and for $\gamma_r = \alpha_j \alpha_k \overline{\alpha_l}$. So, $|\hat{\gamma}_r|$ is shown to uniformly bounded because of the time constraint in Theorem 3.2.1.

Thus, following [**31**], for the nonlinear equation we get

$$\hat{\gamma}_r\left(\frac{\mathbf{k}-\mathbf{K}}{\delta}, \delta s\right) \cdot \langle p_\pm(\cdot; \mathbf{k}), \mathcal{P}_r(\cdot)\rangle_{L^2(\Omega)} =$$

(3.56)
$$\frac{1}{\sqrt{2}}\frac{\kappa_1 + i\kappa_2}{|\kappa|}\langle \Phi_1, \hat{\gamma}_r\left(\frac{\kappa}{\delta}, \delta s\right) \cdot \Psi_r\rangle_{L^2(\Omega)} \pm \frac{1}{\sqrt{2}}\langle \Phi_2, \hat{\gamma}_r\left(\frac{\kappa}{\delta}, \delta s\right) \cdot \Psi_r\rangle_{L^2(\Omega)}$$

$$+ \mathcal{O}\big(|\kappa|\big)$$

where we use $\kappa = \mathbf{k} - \mathbf{K}$. At this point we can note that the term $\mathcal{O}\big(|\kappa|\big)$ is actually $0$ because we are projecting onto the $\mu_*$-eigenspace and the contributions to that term come from the other eigenspaces.

As these inner products are fully expanded in Proposition 7.3 of [**32**] for $1 \leq r \leq 4$, we explore the case when $r = 5$. For most choices of $j, k,$ and $l$ the inner products $\langle \Phi_1, \hat{\gamma}_5\left(\frac{\kappa}{\delta^c}, \delta^c s\right) \cdot \Psi_5\rangle_{L^2(\Omega)}$ and $\langle \Phi_2, \hat{\gamma}_5\left(\frac{\kappa}{\delta^c}, \delta^c s\right) \cdot \Psi_5\rangle_{L^2(\Omega)}$ will disappear. We first note that in this inner product the inputs of $\hat{\gamma}_r$ are not in $\Omega$. Thus we can pull out $\hat{\gamma}_5$ to get $\langle \Phi_1, \hat{\gamma}_5\left(\frac{\kappa}{\delta^c}, \delta^c s\right) \cdot \Psi_5\rangle_{L^2(\Omega)} = \hat{\gamma}_5\left(\frac{\kappa}{\delta^c}, \delta^c s\right)\langle \Phi_1, \Psi_5\rangle_{L^2(\Omega)}$. Several inner products $\langle \Phi_1, \Phi_j \Phi_k \overline{\Phi_l}\rangle_{L^2(\Omega)}$ disappear as follows:

(3.57)
$$\langle \Phi_1, \Phi_1\Phi_1\overline{\Phi_2}\rangle_{L^2(\Omega)} = \int_\Omega \overline{\Phi_1(\mathbf{x})} \cdot \Phi_1(\mathbf{x})\Phi_1(\mathbf{x})\overline{\Phi_2(\mathbf{x})}d\mathbf{x}$$

$$= \int_{R\Omega} \overline{\omega\Phi_1(\mathbf{y})} \cdot \omega\Phi_1(\mathbf{y}) \cdot \omega\Phi_1(\mathbf{y}) \cdot \overline{\bar{\omega}\Phi_2(\mathbf{y})}d\mathbf{y}$$

$$= \omega^2 \int_\Omega \overline{\Phi_1(\mathbf{x})} \cdot \Phi_1(\mathbf{x})\Phi_1(\mathbf{x})\overline{\Phi_2(\mathbf{x})}d\mathbf{x}$$

$$= \omega^2\langle \Phi_1, \Phi_1\Phi_1\overline{\Phi_2}\rangle_{L^2(\Omega)}$$

where $\omega$ in the second line is from a rotation a third of the way around the fundamental period cell of the lattice. Thus the inner product $\langle \Phi_1, \Phi_1\Phi_1\overline{\Phi_2}\rangle_{L^2(\Omega)}$ must be $0$. It works out similarly for several other choices of $j, k, l$. These disappearing inner products are discussed at the end of page 9 in [**31**]. The only inner products which don't disappear in this manner are as follows,

(3.58)
$$\langle \Phi_1, \widehat{\alpha_1\alpha_1\overline{\alpha_1}} \cdot \Phi_1\Phi_1\overline{\Phi_1}\rangle = \beta_1|\alpha_1|^2\alpha_1 \qquad \langle \Phi_2, \widehat{\alpha_2\alpha_2\overline{\alpha_2}} \cdot \Phi_2\Phi_2\overline{\Phi_2}\rangle = \beta_1|\alpha_2|^2\alpha_2$$

$$\langle \Phi_1, \widehat{\alpha_2\alpha_1\overline{\alpha_2}} \cdot \Phi_2\Phi_1\overline{\Phi_2}\rangle = \beta_2|\alpha_2|^2\alpha_1 \qquad \langle \Phi_2, \widehat{\alpha_2\alpha_1\overline{\alpha_1}} \cdot \Phi_2\Phi_1\overline{\Phi_1}\rangle = \beta_2|\alpha_1|^2\alpha_2$$

$$\langle \Phi_1, \widehat{\alpha_1\alpha_2\overline{\alpha_2}} \cdot \Phi_1\Phi_2\overline{\Phi_2}\rangle = \beta_2|\alpha_2|^2\alpha_1 \qquad \langle \Phi_2, \widehat{\alpha_1\alpha_2\overline{\alpha_1}} \cdot \Phi_1\Phi_2\overline{\Phi_1}\rangle = \beta_2|\alpha_1|^2\alpha_2$$

where $\beta_1 = \int_\Omega |\Phi_1(\mathbf{x})|^4 d\mathbf{x} = \int_\Omega |\Phi_2(\mathbf{x})|^4 d\mathbf{x}$ and $\beta_2 = \int_\Omega |\Phi_1(\mathbf{x})|^2|\Phi_2(\mathbf{x})|^2 d\mathbf{x}$.

Now we can actually perform the summation over $r$. Summing the first 4 terms as done in [**32**] and adding in the 6 above inner products we get:

$$(3.59)\quad \begin{aligned} \sum_r \hat{\gamma}_r\left(\frac{\kappa}{\delta}, \delta s\right) \cdot \langle p_{\pm}(\cdot; \mathbf{k}), \mathcal{P}_r(\cdot) \rangle \\ = \frac{i}{\sqrt{2}} \frac{\kappa_1 + i\kappa_2}{|\kappa|} \Big[ \widehat{\partial_t \alpha_1} + \overline{\lambda_\#}\big(\widehat{\partial_{x_1}\alpha_2} + i\widehat{\partial_{x_2}\alpha_2}\big) \\ + i\big(\beta_1 \widehat{|\alpha_1|^2\alpha_1} + 2\beta_2 \widehat{|\alpha_2|^2\alpha_1}\big)\Big]\left(\frac{\kappa}{\delta}, \delta s\right) \\ \pm \frac{i}{\sqrt{2}} \Big[ \widehat{\partial_t \alpha_2} + \lambda_\#\big(\widehat{\partial_{x_1}\alpha_1} - i\widehat{\partial_{x_2}\alpha_1}\big) \\ + i\big(2\beta_2 \widehat{|\alpha_1|^2\alpha_2} + \beta_1 \widehat{|\alpha_2|^2\alpha_2}\big)\Big]\left(\frac{\kappa}{\delta}, \delta s\right). \end{aligned}$$

This is where the specific formulation of the Dirac equation, (3.25), most obviously comes into play. Note that since $\alpha$ is a solution to the nonlinear Dirac equation we have $\widehat{\partial_t\alpha_1} + \overline{\lambda_\#}\big(\widehat{\partial_{x_1}\alpha_2} + i\widehat{\partial_{x_2}\alpha_2}\big) + i\big(\beta_1\widehat{|\alpha_1|^2\alpha_1} + 2\beta_2\widehat{|\alpha_2|^2\alpha_1}\big) = 0$ and $\widehat{\partial_t\alpha_2} + \lambda_\#\big(\widehat{\partial_{x_1}\alpha_1} - i\widehat{\partial_{x_2}\alpha_1}\big) + i\big(2\beta_2\widehat{|\alpha_1|^2\alpha_2} + \beta_1\widehat{|\alpha_2|^2\alpha_2}\big) = 0$. Thus

$$(3.60)\quad \sum_r \hat{\gamma}_r\left(\frac{\kappa}{\delta}, \delta s\right) \cdot \langle p_{\pm}(\cdot; \mathbf{k}), \mathcal{P}_r(\cdot) \rangle = 0,$$

and by substituting the above equality into the definition of $f_{1,r}$, (3.44), we get

$$(3.61)\quad \sum_r f_{1,r} = 0.$$

Substituting (3.61) and (3.53) into the definition for $\tilde{f}_{I,D}$, (3.38), we get

$$(3.62)\quad \tilde{f}_{I,D}(\mathbf{k}, t) = 0 + \mathcal{O}\left(\delta^{\frac{5}{2}}|t| \cdot \chi(|\kappa| < \delta^\tau)\right),$$

and we have the desired bound:

$$(3.63)\quad \|f_{I,D} + \alpha_{I,D}^{\perp}(\cdot, t)\|_{L^2}^2 = \mathcal{O}(\delta^{\frac{5}{2}} t).$$

**3.4.2. Estimation of $\|f_{II,D}(\cdot, t) + \alpha_{II,D}^{\perp}(\cdot, t)\|_{L^2(\mathbb{R}^2)}$.** Recalling the definition of $f_{II,D}$ we get

$$(3.64)\quad \begin{aligned} \|f_{II,D}(\cdot, t) + \alpha_{II,D}^{\perp}(\cdot, t)\|_{L^2(\mathbb{R}^2)}^2 = \sum_{b \in \{+,-\}} \int_{\mathcal{B}_h} \chi\Big(|\mathbf{k} - \mathbf{K}| \geq \delta^\tau\Big) |\tilde{f}_b(\mathbf{k}, t)|^2 d\mathbf{k} \\ + \sum_{b \in \{+,-\}} \int_{\mathcal{B}_h} \chi\Big(|\mathbf{k} - \mathbf{K}| \geq \delta^\tau\Big) |(P_* - 1)\tilde{\alpha}^{\perp, D}(\mathbf{k}, t)| d\mathbf{k} \end{aligned}$$

where

$$\tilde{f}_{\pm} = -\delta^{\frac{3}{2}} \int_0^t e^{-i(\mu_{\pm}(\mathbf{k}) - \mu_*)(t-s)} \langle \Phi_{\pm}(\cdot; \mathbf{k}), \sum_r \gamma_r(\delta \cdot, \delta s) \Psi_r(\cdot) \rangle_{L^2(\mathbb{R}^2)} ds$$

(3.65)

$$\tilde{\alpha}^{\perp, D} = -\delta^{\frac{3}{2}} \int_0^t e^{-i(\mu_b(\mathbf{k}) - \mu_*)(t-s)} \sum_r \delta^{-2} \hat{\gamma}_r \left( \frac{\mathbf{k} - \mathbf{K}}{\delta}, \delta s \right) \int_{\Omega} \overline{p_{\pm}(\mathbf{y}, \mathbf{k})} \mathcal{P}_r(\mathbf{k}) d\mathbf{y} ds.$$

However, as we saw in (3.44) and the paragraph immediately following, $(P_* - 1)\tilde{\alpha}^{\perp, D}$ is a projection of a component of $\tilde{f}_{\pm}$. Thus the second term of the right-hand side of (3.64) is bounded by the first term, and we only need to find a bound for the first term.

Applying Proposition 3.4.3 to the inner product we get

$$\langle \Phi_{\pm}(\cdot; \mathbf{k}), \gamma_r(\delta \cdot, \delta s) \Psi_r(\cdot) \rangle_{L^2(\mathbb{R}^2)} =$$

(3.66)

$$\int_{\Omega} \overline{p_{\pm}(\mathbf{y}; \mathbf{k})} \left[ \delta^{-2} \sum_{\mathbf{m} \in \mathbb{Z}^2} e^{i\mathbf{m} \cdot \mathbf{y}} \hat{\gamma}_r \left( \frac{m_1 \mathbf{k}_1 + m_2 \mathbf{k}_2 + (\mathbf{k} - \mathbf{K})}{\delta}, \delta s \right) \right] \mathcal{P}_r(\mathbf{y}) d\mathbf{y}.$$

We will make use of Lemma 7.4 in [**32**]:

LEMMA 3.4.6. *Let* $\mathbf{k} \in \mathcal{B}_h$ *and assume* $|\mathbf{k} - \mathbf{K}| \geq \delta^{\tau}$. *Then, there exists a constant,* $c_1$, *such that for all* $\mathbf{m} = (m_1, m_2) \in \mathbb{Z}^2$

(3.67)
$$|m_1 \mathbf{k}_1 + m_2 \mathbf{k}_2 + (\mathbf{k} - \mathbf{K})| \geq c_1 \delta^{\tau} (1 + |\mathbf{m}|).$$

Applying Proposition 3.4.4 (with $A = A_0$) along with the above Lemma gives us

(3.68)

$$\left| \delta^{-2} \sum_{\mathbf{m} \in \mathbb{Z}^2} e^{i\mathbf{m} \cdot \mathbf{y}} \hat{\gamma}_r \left( \frac{m_1 \mathbf{k}_1 + m_2 \mathbf{k}_2 + (\mathbf{k} - \mathbf{K})}{\delta}, \delta s \right) \right|$$

$$\leq C \sum_{\mathbf{m} \in \mathbb{Z}^2} \frac{1}{(1 + |\mathbf{m}|)^{A_0}} \delta^{-2} (\delta^{1-\tau})^{A_0}.$$

Since $A_0 > 12 > 2$ the summation converges, and we get

(3.69)
$$|\tilde{f}_{\pm}(\mathbf{k}, t)| \leq C|t|\delta^{-\frac{1}{2}} (\delta^{1-\tau})^{A_0}.$$

Since $\tau < \frac{A_0 - 3}{A_0}$ we get

(3.70)
$$\|f_{II,D}(\cdot, t) + \alpha_{II,D}^{\perp}(\cdot, t)\|_{L^2(\mathbb{R}^2)} = \mathcal{O}(\delta^{\frac{5}{2}} t).$$

**3.4.3. Estimation of** $\|f_{D^c}(\cdot,t) + \alpha_{D^c}^{\perp}(\cdot,t)\|_{H^2(\mathbb{R}^2)}$**.** First, we recall the definitions:

$$f_{D^c}(\mathbf{x},t) = \sum_{b\notin\{+,-\}} \int_{\mathcal{B}_h} \tilde{f}_b(\mathbf{k},t)\Phi_b(\mathbf{x};\mathbf{k})d\mathbf{k}$$

(3.71)

$$\alpha_{D^c}^{\perp}(\mathbf{x},t) = - \sum_{b\notin\{+,-\}} \int_{\mathcal{B}_h} \tilde{\alpha}^{\perp,D^c}(\mathbf{k},t)\Phi_b(\mathbf{x};\mathbf{k})d\mathbf{k}$$

where

$$\tilde{f}_b = -\delta^{\frac{3}{2}} \int_0^t e^{-i(\mu_b(\mathbf{k})-\mu_*)(t-s)}\langle\Phi_b(\cdot;\mathbf{k}), \sum_r \gamma_r(\delta\cdot,\delta s)\Psi_r(\cdot)\rangle_{L^2(\mathbb{R}^2)}ds$$

(3.72)
$$\tilde{\alpha}^{\perp,D^c} = -\delta^{\frac{3}{2}} \int_0^t e^{-i(\mu_b(\mathbf{k})-\mu_*)(t-s)}\chi(|\mathbf{k}-\mathbf{K}| < \kappa_1)$$

$$\times \left\langle\Phi_b(\cdot;\mathbf{k}), \sum_r \gamma_r(\delta\cdot,\delta t)\Psi_r(\cdot)\right\rangle_{L^2(\mathbb{R}^2)} ds.$$

We decompose $f_{D^c}$ as we did $f_D$ to get

$$f_{D^c}(\mathbf{x},t) = \sum_{b\notin\{+,-\}} \int_{\mathcal{B}_h} \chi\Big(|\mathbf{k}-\mathbf{K}| < \kappa_1\Big)\tilde{f}_b(\mathbf{k},t)\Phi_b(\mathbf{x};\mathbf{k})d\mathbf{k}$$

(3.73)

$$+ \sum_{b\notin\{+,-\}} \int_{\mathcal{B}_h} \chi\Big(|\mathbf{k}-\mathbf{K}| \geq \kappa_1\Big)\tilde{f}_b(\mathbf{k},t)\Phi_b(\mathbf{x};\mathbf{k})d\mathbf{k}$$

$$:= f_{I,D^c}(\mathbf{x},t) + f_{II,D^c}(\mathbf{x},t).$$

At this point we notice that $f_{I,D^c} = -\alpha_{D^c}^{\perp}$. Thus, those two terms cancel out and we only need to bound the term $f_{II,D^c}$. However, in order to show why we needed this component of $\alpha^{\perp}$ to cancel out $f_{I,D^c}$, we will continue for now as if there was not an $\alpha_{D^c}^{\perp}$ term to cancel it out. On its own, the term $f_{I,D^c}$ would give the most restrictive bounds of our expansion and thus limit the lifespan. Also, we include this part as the argument is needed when bounding the $A_2$ terms since there is not a term present to cancel out the corresponding $g_{I,D^c}$ component.

3.4.3.1. *Estimation of* $\|f_{I,D^c}(\cdot,t)\|_{H^2(\mathbb{R}^2)}$. Once again, following the proof in [**32**], we use (3.9) to get

(3.74)
$$\|f_{I,D^c}(\cdot,t)\|_{H^2(\mathbb{R}^2)}^2 \approx \sum_{b\notin\{+,-\}} (1+|b|)^2 \int_{\mathbf{k}\in\mathcal{B}_h:|\mathbf{k}-\mathbf{K}|<\kappa_1} |\tilde{f}_b(\mathbf{k},t)|^2 d\mathbf{k}$$

along with the equilavent of Equation 7.50 in [**32**],

$$(3.75) \qquad \begin{aligned} |\tilde{f}_b(\mathbf{k}, t)| &\lesssim \delta^{\frac{3}{2}} \left| \left\langle \Phi_b(\cdot; \mathbf{k}), \sum_r \gamma_r(\delta \cdot, 0) \Psi_r(\cdot) \right\rangle_{L^2(\mathbb{R}^2)} \right| \\ &\quad + \delta^{\frac{5}{2}} \cdot t \max_{0 \leq s \leq T} \left| \left\langle \Phi_b(\cdot; \mathbf{k}), \sum_r \partial_T \gamma_r(\delta \cdot, \delta s) \Psi_r(\cdot) \right\rangle_{L^2(\mathbb{R}^2)} \right|, \end{aligned}$$

and the equivalent of Equation 7.53 in [**32**],

$$(3.76) \qquad \begin{aligned} &\left| \left\langle \Phi_b(\cdot; \mathbf{k}), \sum_r \partial_T \gamma_r(\delta \cdot, \delta s) \Psi_r(\cdot) \right\rangle_{L^2(\mathbb{R}^2)} \right| \\ &\qquad \lesssim (1 + |b|)^{-M} \sum_m \| \gamma_m^*(\delta \cdot, \delta s) \|_{H^{2M+1}(\mathbb{R}^2)} \end{aligned}$$

where we can let $M = 2$ and where $\gamma_r^*$ is the same as $\gamma_r$ except for one additional term to account for the fifth order $\alpha$ term now present because of the additional time derivative term being expanded as spatial derivative terms and cubic terms. We also note that the integration by parts in (3.75) did require us to use the fact that $|\mu_b(\mathbf{k}) - \mu_*|$ is uniformly bounded below for $b \neq \pm$ when $\mathbf{k}$ is close enough to the Dirac point, $\mathbf{K}$. (This is true even in the upcoming tight binding setting in Section 3.8)

Scaling by $\delta$, substituting $M = 2$, and using $H^5$ is an algebra we get

$$(3.77) \qquad \left| \left\langle \Phi_b(\cdot; \mathbf{k}), \sum_r \partial_T \gamma_r(\delta \cdot, \delta s) \Psi_r(\cdot) \right\rangle_{L^2(\mathbb{R}^2)} \right| \lesssim (1 + |b|)^{-2} \delta^{-1}.$$

Plugging this bound into (3.75) we get

$$(3.78) \qquad |\tilde{f}_b(\mathbf{k}, t)|^2 \lesssim (1 + |b|)^{-4} (\delta^1 + \delta^3 \cdot t^2).$$

Now, substituting into (3.74) we have

$$(3.79) \qquad \begin{aligned} \| f_{I,D^c}(\cdot, t) \|_{H^2(\mathbb{R}^2)} &\approx \left( \sum_{b \notin \{+,-\}} (1 + |b|)^{-2} \right)^{\frac{1}{2}} (\delta^{\frac{1}{2}} + \delta^{\frac{3}{2}} t) \\ &\lesssim (\delta^{\frac{1}{2}} + \delta^{\frac{3}{2}} t) \end{aligned}$$

where the summation converges because of our choice, $M = 2$.

Thus we would have the desired bound

$$(3.80) \qquad \| f_{I,D^c}(\cdot, t) \|_{H^2(\mathbb{R}^2)} = \mathcal{O}(\delta^{\frac{1}{2}} + \delta^{\frac{3}{2}} t).$$

However as we discussed right before this subsection, $f_{I,D^c}$ is entirely cancelled out by $\alpha_{D^c}^\perp$. Thus we actually get the contribution

$$(3.81) \qquad \|f_{I,D^c}(\cdot,t) + \alpha_{D^c}^\perp(\cdot,t)\|_{H^2(\mathbb{R}^2)} = 0.$$

3.4.3.2. *Estimation of* $\|f_{II,D^c}(\cdot,t)\|_{H^2(\mathbb{R}^2)}$. As above, we get

$$(3.82) \qquad \|f_{II,D^c}(\cdot,t)\|_{H^2(\mathbb{R}^2)}^2 \approx \sum_{b \notin \{+,-\}} (1+|b|)^2 \int_{\mathbf{k} \in \mathcal{B}_h : |\mathbf{k}-\mathbf{K}| > \kappa_1} |\tilde{f}_b(\mathbf{k},t)|^2 d\mathbf{k}.$$

Substituting in (3.72) we find

$$(3.83) \qquad \begin{aligned} & \int_{\mathcal{B}_h} \chi\left(|\mathbf{k}-\mathbf{K}| \geq \kappa_1\right) |\tilde{f}_b(\mathbf{k},t)|^2 d\mathbf{k} \\ & \leq \delta^3 t^2 \max_{0 \leq s \leq t} \int_{\mathcal{B}_h} \chi\left(|\mathbf{k}-\mathbf{K}| \geq \kappa_1\right) \left| \sum_r \langle \Phi_b(\cdot;\mathbf{k}), \gamma_r(\delta\cdot,\delta s)\Psi_r(\cdot)\rangle_{L^2(\mathbb{R}^2)} \right|^2 d\mathbf{k}. \end{aligned}$$

As is done on page 281 of [**32**], we choose $\tilde{C} > 0$ such that $\tilde{C} + H$ is strictly positive. We also let $a_1, \ldots, a_{j_{max}}, b_1, b_2 \in \mathbb{Z}_+^2$ such that $2j_{max} + \sum_{j=1}^{j_{max}} |a_j| + |b_1| + |b_2| \leq 2M$ where $M$ is a nonnegative integer, and $\widetilde{\sum}$ indicates the sum over all such $a_1, \ldots, a_{j_{max}}, b_1, b_2$. Within the inner product we integrate by parts and proceed as in [**32**] to get

$$(3.84) \qquad \begin{aligned} & \langle \Phi_b(\cdot;\mathbf{k}), \gamma_r(\delta\cdot,\delta s)\Psi_r(\cdot)\rangle_{L^2(\mathbb{R}^2)} \\ & = \frac{1}{\left(\tilde{C} + \mu_b(\mathbf{k})\right)^M} \widetilde{\sum_{a,b_1,b_2}} \delta^{|b_1|} \langle \Phi_b(\cdot;\mathbf{k}), \Pi_{j=1}^{j_{max}} \partial_{\mathbf{x}}^{a_j} V(\cdot)\partial_{\mathbf{X}}^{b_1}\gamma_r(\delta\cdot,\delta s)\partial_{\mathbf{x}}^{b_2}\Psi_r(\cdot)\rangle_{L^2(\mathbb{R}^2)}. \end{aligned}$$

From here we proceed similarly to how we bounded $f_{II,D}$. By choosing $\Gamma(\mathbf{X},T) = \partial_{\mathbf{X}}^{b_1}\nu_r(\delta\mathbf{x},\delta s)$ and $\Psi(\mathbf{x}) = \Pi_{j=1}^{j_{max}} \partial_{\mathbf{x}}^{a_j} V(\mathbf{x})\partial_{\mathbf{x}}^{b_2}\Psi_r(\mathbf{x})$ we are able to apply Proposition 3.4.3 as done by Fefferman and Weinstein. We know that there is a constant $c > 0$ such that $|m_1\mathbf{k}_1 + m_2\mathbf{k}_2 + (\mathbf{k}-\mathbf{K})| \geq c(1+|\mathbf{m}|)$ since we are in the region $|\mathbf{k}-\mathbf{K}| \geq \kappa_1$. Now we can apply Proposition 3.4.7 and bound the inner product

$$(3.85) \qquad \left| \langle \Phi_b(\cdot;\mathbf{k}), \Pi_{j=1}^{j_{max}} \partial_{\mathbf{x}}^{a_j} V(\cdot)\partial_{\mathbf{X}}^{b_1}\gamma_r(\delta\cdot,\delta s)\partial_{\mathbf{x}}^{b_2}\Psi_r(\cdot)\rangle_{L^2(\mathbb{R}^2)} \right| \lesssim \delta^{A_0-2} \sum_{\mathbf{m} \in \mathbb{Z}^2} \frac{1}{(1+|\mathbf{m}|)^{A_0}}$$

where the summation converges because $A_0 > 12$.

Using the fact that $\mu_b(\mathbf{k}) \approx b$ for large $b$ uniformly in $\mathbf{k} \in \mathcal{B}_h$ we substitute (3.83), (3.84), and (3.85) into (3.82) to get

(3.86)
$$\|f_{II,D^c}(\cdot,t)\|^2_{H^2(\mathbb{R}^2)} \lesssim \delta^3 t^2 \delta^{2A_0-4} \sum_b (1+b)^{2-2M}.$$

Choosing $M = 2$ (and thus we were integrating by parts twice in (3.84)) allows the summation to converge. Remembering that $A_0 > 12$, we see that we have at least a factor, $\delta^{23}$, on the right-hand side. Thus we have

(3.87)
$$\|f_{II,D^c}(\cdot,t)\|_{H^2(\mathbb{R}^2)} = \mathcal{O}(\delta^{\frac{23}{2}}t).$$

Substituting (3.63), (3.70), (3.73), (3.81), and (3.87) into (3.40) we get the desired result for $A_1$:

(3.88)
$$\|A_1\|_{H^2} \leq C_1 \delta^{\frac{5}{2}} t.$$

**3.4.4. Estimating the $\beta$ term, $A_2$.** Similar to before, we can consider the term $A_2$ as the solution to the initial value problem

(3.89)
$$i\partial_t g(\mathbf{x},t) - (H - \mu_*)g(\mathbf{x},t) = \delta^{\frac{5}{2}} \sum_r \mathcal{W}_r \nu_r(\delta \mathbf{x}, \delta t) \Psi_r(\mathbf{x})$$
$$g(\mathbf{x},0) = 0$$

where we define

(3.90)
$$
\begin{aligned}
\nu_1(\mathbf{x},t) &= i\partial_t \beta_1(\mathbf{x},t), & \Psi_1(\mathbf{x}) &= \Phi_1(\mathbf{x}) \\
\nu_2(\mathbf{x},t) &= i\partial_t \beta_2(\mathbf{x},t), & \Psi_2(\mathbf{x}) &= \Phi_2(\mathbf{x}) \\
\nu_3(\mathbf{x},t) &= 2\nabla \beta_1(\mathbf{x},t), & \Psi_3(\mathbf{x}) &= \nabla \Phi_1(\mathbf{x}) \\
\nu_4(\mathbf{x},t) &= 2\nabla \beta_2(\mathbf{x},t), & \Psi_4(\mathbf{x}) &= \nabla \Phi_2(\mathbf{x}) \\
\nu_5(\mathbf{x},t) &= \Delta \alpha_j(\mathbf{x},t), & \Psi_5(\mathbf{x}) &= \Phi_j(\mathbf{x}) \\
\nu_6(\mathbf{x},t) &= i\partial_t^2 \alpha_j^!(\mathbf{x},t), & \Psi_6(\mathbf{x}) &= \Phi_j(\mathbf{x}) \\
\nu_7(\mathbf{x},t) &= 2\partial_t \nabla \alpha_j^!(\mathbf{x},t), & \Psi_7(\mathbf{x}) &= \nabla \Phi_j(\mathbf{x}) \\
\nu_8(\mathbf{x},t) &= \partial_t \left( \alpha_j^!(\mathbf{x},t)\alpha_k^!(\mathbf{x},t)\overline{\alpha}_l^!(\mathbf{x},t) \right), & \Psi_8(\mathbf{x}) &= \Phi_j(\mathbf{x})\Phi_k(\mathbf{x})\overline{\Phi}_l(\mathbf{x}),
\end{aligned}
$$

and $\mathcal{W}_r = 1$ for $1 \leq r \leq 5$ and $\mathcal{W}_r = (1 - P_*)(H - \mu_*)^{-1}(1 - P_*)$ for $6 \leq r \leq 8$. We note that $(1 - P_*)(H - \mu_*)^{-1}(1 - P_*)$ is a bounded operator since it is projected away from the $\mu_*$-eigenspace

82

(Kuchment [**55**] Section 4.2). Thus any uniform bound we find for $\sum_r \nu_r(\delta\mathbf{x}, \delta t)\Psi_r(\mathbf{x})$ will be equivalent to a bound for $\sum_r \mathcal{W}_r\nu_r(\delta\mathbf{x}, \delta t)\Psi_r(\mathbf{x})$ modulo a bounding constant depending on the bound of the operator $\mathcal{W}_r$. For that reason we suppress the $\mathcal{W}_r$ moving forward. Notice, that for $\nu_6, \nu_7$ and $\nu_8$ we introduce an exponent, "!", to indicate that those three terms are the time derivative applied to the appropriately localized components of $\partial_t\alpha_j, \nabla\alpha_j$, and $\alpha_j\alpha_k\overline{\alpha}_l$ as indicated in the definition of $\alpha^\perp$, (3.28), and the expansions seen throughout the proof of the bound for $A_1$. This does not affect finding upper bounds for $\nu_6, \nu_7$, and $\nu_8$ because the size of a component is trivially bounded by the size of the whole; it only ensures that the appropriate terms vanish in the full expansion of the term, $g_{1,r}$.

This expansion is more intimidating than the case was for $A_1$. However, it will be handled in a similar fashion. Also note, that $\|\beta\|_{H^{A_0}}$ is uniformly bounded just as $\|\alpha\|_{H^{A_0}}$ is because of the uniform time bound $T^* = \rho$.

We remind ourselves the statement of (3.40) rewritten in terms of $g$ and without the terms of $\alpha^\perp$:

(3.91) $$\|g(\cdot, t)\|^2_{H^2(\mathbb{R}^2)} \approx \|g_{I,D}(\cdot, t)\|^2_{L^2(\mathbb{R}^2)} + \|g_{II,D}(\cdot, t)\|^2_{L^2(\mathbb{R}^2)} + \|g_{D^c}(\cdot, t)\|^2_{H^2(\mathbb{R}^2)}.$$

Note that in this section we are defining $g_{I,D}, g_{II,d}, g_{1,r}, g_{2,r}$ and $g_{D^c}$ equivalently to their $f$ counterparts in proving the bound for $A_1$ (but with a factor $\delta^{\frac{5}{2}}$ instead of $\delta^{\frac{3}{2}}$).

Proposition 3.4.3 still holds as nothing in it was unique to the case of $A_1$, and Proposition 3.4.4 can be easily modified as follows:

PROPOSITION 3.4.7. *Let $\nu_r$ be as previously defined in (3.90). Then:*

*(1) For all $\kappa \in \mathbb{R}^2$ and $0 < A \leq A_0$ we have*

(3.92) $$|\hat{\nu}_r(\kappa, t)| \lesssim |\kappa|^{-A}.$$

*(2) For all $\mathbf{k} \in \mathcal{B}_h$ such that $|\mathbf{k} - \mathbf{K}| < \delta^\tau$ and $2 < A \leq A_0$ we have*

(3.93) $$|E_{r,\delta}\left(\delta^{-1}\mathbf{y}, \delta t; \mathbf{k}\right)| \lesssim \delta^A$$

*where $t \leq \rho\delta^{-1}$ and the $\hat{\gamma}$ within the definition of $E_{r,\delta}$ has been replaced by $\hat{\nu}$.*

The proof of this proposition is a simple corollary of the proof of Proposition 3.4.4. In that proof we wrote the $\partial_t \alpha$ terms as a linear combination of the spatial derivative terms and the cubic terms. This time we have two time derivatives to expand, but we still get a uniform bound.

Just as before Proposition 3.4.7 applied to the formula for $g_{2,r}$ immediately gives

$$(3.94) \qquad |g_{2,r}| \lesssim \delta^{A_0 - 2},$$

and $g_{2,r}$ has a contribution to $\|g_{I,D}\|_{L^2}$ of size bounded by $\delta^{\frac{5}{2}}$.

Also as before, we expand the inner product within the definition of $\sum_r g_{1,r}$ to get

$$\hat{\nu}_r \left( \frac{\mathbf{k} - \mathbf{K}}{\delta}, \delta s \right) \cdot \langle p_\pm(\cdot; \mathbf{k}), \mathcal{P}_r(\cdot) \rangle_{L^2(\Omega)} =$$

$$(3.95) \qquad \frac{1}{\sqrt{2}} \frac{\kappa_1 + i\kappa_2}{|\kappa|} \langle \Phi_1, \hat{\nu}_r \left( \frac{\kappa}{\delta}, \delta s \right) \cdot \Psi_r \rangle_{L^2(\Omega)} \pm \frac{1}{\sqrt{2}} \langle \Phi_2, \hat{\nu}_r \left( \frac{\kappa}{\delta}, \delta s \right) \cdot \Psi_r \rangle_{L^2(\Omega)}$$

$$+ \mathcal{O}(|\kappa|).$$

The difference here is that the term $\mathcal{O}(|\kappa|)$ does not disappear because we are not projected onto the $\mu_*$-eigenspace.

Expanding the resultant inner product as before we get

$$\sum_r \hat{\nu}_r \left( \frac{\kappa}{\delta}, \delta s \right) \cdot \langle p_\pm(\cdot; \mathbf{k}), \mathcal{P}_r(\cdot) \rangle$$

$$= \frac{i}{\sqrt{2}} \frac{\kappa_1 + i\kappa_2}{|\kappa|} \left[ \widehat{\partial_t \beta_1} + \overline{\lambda_\#} \left( \widehat{\partial_{x_1} \beta_2} + i \widehat{\partial_{x_2} \beta_2} \right) \right.$$

$$(3.96) \qquad \left. - \widehat{\Delta \alpha_1} - \langle \Phi_1, \widehat{\partial_T \alpha^\perp} \rangle_{L^2(\Omega)} \right] \left( \frac{\kappa}{\delta}, \delta s \right)$$

$$\pm \frac{i}{\sqrt{2}} \left[ \widehat{\partial_t \beta_2} + \lambda_\# \left( \widehat{\partial_{x_1} \beta_1} - i \widehat{\partial_{x_2} \beta_1} \right) \right.$$

$$\left. - \widehat{\Delta \alpha_2} - \langle \Phi_2, \widehat{\partial_T \alpha^\perp} \rangle_{L^2(\Omega)} \right] \left( \frac{\kappa}{\delta}, \delta s \right) + \mathcal{O}(|\kappa|).$$

Since $\beta$ solves (3.29) the first two terms disappear and we are left with

$$(3.97) \qquad \sum_r \hat{\nu}_r \left( \frac{\kappa}{\delta}, \delta s \right) \cdot \langle p_\pm(\cdot; \mathbf{k}), \mathcal{P}_r(\cdot) \rangle = \mathcal{O}(|\kappa|),$$

and by substituting the above equality into the definition of $g_{1,r}$ we get

$$(3.98) \qquad \sum_r g_{1,r} = \mathcal{O}\left( \delta^{-2} |\kappa| \chi(|\kappa| < \delta^\tau) \right).$$

84

Substituting (3.98) and (3.94) into the definition for $\tilde{g}_{I,D}$ we get

$$(3.99) \qquad \tilde{g}_{I,D}(\mathbf{k},t) = \mathcal{O}\big(\delta^{\frac{1}{2}}|t| \cdot |\kappa| \cdot \chi(|\kappa| < \delta^\tau)\big) + \mathcal{O}\big(\delta^{\frac{5}{2}}|t| \cdot \chi(|\kappa| < \delta^\tau)\big)$$

Similarly to [**32**], the first term in (3.62) gives a contribution to $\|f_{I,D}\|_{L^2}^2$ bounded by

$$(3.100) \qquad \delta t^2 \int_{\mathcal{B}_h} |\mathbf{k} - \mathbf{K}|^2 \, \chi\big(|\mathbf{k} - \mathbf{K}| < \delta^\tau\big) d\mathbf{k} \lesssim \delta^1 t^2 \delta^{4\tau} = t^2 \delta^{4\tau+1}.$$

Since $\tau > \frac{3}{4}$ we get $\delta^{4\tau+1} < \delta^4$. Thus we have the desired bound

$$(3.101) \qquad \|g_{I,D}\|_{L^2} = \mathcal{O}(\delta^2 t).$$

The bound for $g_{II,D}$ proceeds identically to the bound for $f_{II,D}$ other than using Proposition 3.4.7 instead of Proposition 3.4.4. Thus we have the bound

$$(3.102) \qquad \|g_{II,D}\|_{L^2} = \mathcal{O}(\delta^{\frac{5}{2}} t)$$

**3.4.5. Estimation of $\|g_{D^c}(\cdot,t)\|_{H^2(\mathbb{R}^2)}$.** The proof for finding the bounds for $\|g_{D^c}(\cdot,t)\|_{H^2(\mathbb{R}^2)}$ is identical to the one for $\|f_{D^c}(\cdot,t)\|_{H^2(\mathbb{R}^2)}$ other than using Proposition 3.4.7 instead of Proposition 3.4.4, starting with different powers of $\delta$, and the fact that there is not an $\alpha_{D^c}^{\perp}$ to cancel out the $g_{I,D^c}$ term. Following the same proof gives us the desired bounds

$$(3.103) \qquad \begin{aligned} \|g_{I,D^c}(\cdot,t)\|_{H^2(\mathbb{R}^2)} &= \mathcal{O}(\delta^{\frac{3}{2}} + \delta^{\frac{5}{2}} t) \\ \|g_{II,D^c}(\cdot,t)\|_{H^2(\mathbb{R}^2)} &= \mathcal{O}(\delta^{\frac{25}{2}} t). \end{aligned}$$

and

Using Equations (3.91), (3.101), (3.102), and (3.103) we get the desired result for term $A_2$,

$$(3.104) \qquad \|A_2\|_{H^2} = \mathcal{O}(\delta^{\frac{3}{2}} + \delta^2 t).$$

Looking back at our bounds for $\|A_1\|_{H^2}$ and $\|A_2\|_{H^2}$ we can rewrite them as

$$(3.105) \qquad \begin{aligned} \|A_1\|_{H^2} &\le C_1 \rho \delta^{\frac{3}{2}} \\ \|A_2\|_{H^2} &\le C_2 \rho \delta \end{aligned}$$

where have substituted our time upper bound, $t \le \rho \delta^{-1}$.

### 3.5. Bounding the $\eta$ Dependent Terms.

We remind ourselves of the definition of the terms dependent on $\eta$:

$$
\text{(3.106)} \qquad A_{3,n+1} = \delta \sum_{j,k=1}^{2} \int_0^t e^{-i(H-\mu_*)(t-s)} \bigg( 2\eta_n(\cdot,s)\alpha_j(\delta\cdot,\delta s)\overline{\alpha_k}(\delta\cdot,\delta s)\Phi_j(\cdot)\overline{\Phi_k}(\cdot)
$$

$$
+ \overline{\eta_n}(\cdot,s)\alpha_j(\delta\cdot,\delta s)\alpha_k(\delta\cdot,\delta s)\Phi_j(\cdot)\Phi_k(\cdot) \bigg) ds
$$

$$
\text{(3.107)} \qquad A_{4,n+1} = \delta^{\frac{1}{2}} \sum_{j=1}^{2} \int_0^t e^{-i(H-\mu_*)(t-s)} \bigg( 2|\eta_n(\cdot,s)|^2 \alpha_j(\delta\cdot,\delta s)\Phi_j(\cdot)
$$

$$
+ \eta_n(\cdot,s)^2 \overline{\alpha_j}(\delta\cdot,\delta s)\overline{\Phi_j}(\cdot) \bigg) ds
$$

$$
\text{(3.108)} \qquad A_{5,n+1} = \int_0^t e^{-i(H-\mu_*)(t-s)} |\eta_n(\cdot,s)|^2 \, \eta_n(\cdot,s) ds.
$$

### 3.5.1. Estimating $\|A_{3,n+1}(\cdot,t)\|_{H^2(\mathbb{R}^2)}$.

In finding our estimate for $\|A_{3,n+1}(\cdot,t)\|_{H^2(\mathbb{R}^2)}$ we start by taking the $H^2$ norm of both sides and applying the Minkowski triangle inequality to get

$$
\text{(3.109)} \qquad \|A_{3,n+1}(\cdot,t)\|_{H^2(\mathbb{R}^2)} \leq \delta \int_0^t \sum_{j,k=1}^{2} 2\|\eta_n(\cdot,s)\alpha_j(\delta\cdot,\delta s)\overline{\alpha_k}(\delta\cdot,\delta s)\|_{H^2(\mathbb{R}^2)}
$$

$$
+ \|\overline{\eta_n}(\cdot,s)\alpha_j(\delta\cdot,\delta s)\alpha_k(\delta\cdot,\delta s)\|_{H^2(\mathbb{R}^2)} ds
$$

where we got rid of the $\Phi_j$ terms via Hölder's inequality and the boundedness of $\Phi_j$.

In order to better estimate the $H^2$ norms on the right-hand side of the equation we need a simple lemma.

LEMMA 3.5.1. *Given two-dimensional functions $f, g$, and $h$ we can bound*

$$
\text{(3.110)} \qquad \|f \cdot g \cdot h\|_{H^2} \lesssim \|f\|_{H^2} \|g\|_{W^{2,\infty}} \|h\|_{W^{2,\infty}}
$$

The proof is simple. The derivatives on the product, $f \cdot g \cdot h$, are distributed via the product rule, and Hölder's inequality is applied.

We apply Lemma 3.5.1 to (3.109) to get

$$
\text{(3.111)} \qquad \|A_{3,n+1}(\cdot,t)\|_{H^2(\mathbb{R}^2)} \lesssim \delta t \sup_{0 \leq s \leq t} \|\eta_n(\cdot,s)\|_{H^2(\mathbb{R}^2)} \|\alpha(\delta\cdot,\delta s)\|_{W^{2,\infty}}^2.
$$

We move forward by noticing that $\|\partial_x^m \alpha(\delta\cdot, \delta s)\|_{L^\infty} = \delta^{|m|}\|\partial_x^m \alpha(\cdot, \delta s)\|_{L^\infty}$ where $m$ is a multiindex with $0 \leq |m| \leq 2$. Since $\delta^{|m|} \leq 1$ we get $\|\alpha(\delta\cdot, \delta s)\|_{W^{2,\infty}} \lesssim \|\alpha(\cdot, \delta s)\|_{W^{2,\infty}}$. Next we can use the Sobolev embedding $H^2(\mathbb{R}^2) \subset L^\infty(\mathbb{R}^2)$ to get $\|\alpha(\cdot, \delta s)\|_{W^{2,\infty}} \lesssim \|\alpha(\cdot, \delta s)\|_{H^4}$.

Noticing that $\|\alpha(\cdot, \delta s)\|_{H^4} \leq \|\alpha(\cdot, \delta s)\|_{H^{A_0}} \leq C$ we can estimate

$$(3.112) \qquad \|A_{3,n+1}(\cdot, t)\|_{H^2(\mathbb{R}^2)} \leq C_3 \delta \int_0^t \|\eta_n(\cdot, s)\|_{H^2(\mathbb{R}^2)} ds.$$

**3.5.2. Estimating $\|A_{4,n+1}(\cdot, t)\|_{H^2(\mathbb{R}^2)}$.** In finding our estimate for $\|A_{4,n+1}(\cdot, t)\|_{H^2(\mathbb{R}^2)}$ we start the same way as the previous subsection, by taking the $H^2$ norm of both sides and applying the Minkowski triangle inequality to get

$$(3.113) \qquad \|A_{4,n+1}(\cdot, t)\|_{H^2(\mathbb{R}^2)} = \delta^{\frac{1}{2}} \int_0^t \sum_{j=1}^2 2\||\eta_n(\cdot, s)|^2 \alpha_j(\delta\cdot, \delta s)\|_{H^2(\mathbb{R}^2)}$$

$$+ \|\eta_n(\cdot, s)^2 \overline{\alpha_j}(\delta\cdot, \delta s)\|_{H^2(\mathbb{R}^2)} ds.$$

First we use the fact that $H^2(\mathbb{R}^2)$ is an algebra to estimate the first term inside the summation

$$(3.114) \qquad \||\eta_n(\cdot, s)|^2 \alpha_j(\delta\cdot, \delta s)\|_{H^2(\mathbb{R}^2)} \lesssim \|\eta_n(\cdot, s)\|_{H^2(\mathbb{R}^2)} \|\eta_n(\cdot, s)\alpha_j(\delta\cdot, \delta s)\|_{H^2(\mathbb{R}^2)}.$$

Notice that the second term inside the summation behaves similarly. We then apply Lemma 3.5.1 with $f(\cdot) = \eta_n(\cdot, s), g(\cdot) = \alpha_j(\delta\cdot, \delta s)$, and $h(\cdot) = 1$ to the second factor to get

$$(3.115) \qquad \|\eta_n(\cdot, s)^2 \alpha_j(\delta\cdot, \delta s)\|_{H^2(\mathbb{R}^2)} \lesssim \|\eta_n(\cdot, s)\|_{H^2(\mathbb{R}^2)}^2 \|\alpha_j(\delta\cdot, \delta s)\|_{W^{2,\infty}}.$$

As before $\|\alpha(\delta\cdot, \delta s)\|_{W^{2,\infty}} \leq \|\alpha(\cdot, \delta s)\|_{H^4} \leq \|\alpha(\cdot, \delta s)\|_{H^{A_0}} \leq C$. Substituting in to (3.113) we get

$$(3.116) \qquad \|A_{4,n+1}(\cdot, t)\|_{H^2(\mathbb{R}^2)} \leq C_4 \delta^{\frac{1}{2}} \int_0^t \|\eta_n(\cdot, s)\|_{H^2(\mathbb{R}^2)}^2 ds.$$

**3.5.3. Estimating $\|A_{5,n+1}(\cdot, t)\|_{H^2(\mathbb{R}^2)}$.** We take the $H^2$ norm of both sides, apply the triangle inequality, and use the fact that $H^2(\mathbb{R}^2)$ is an algebra to get

$$(3.117) \qquad \|A_{5,n+1}(\cdot, t)\|_{H^2(\mathbb{R}^2)} \leq C_5 \int_0^t \|\eta_n(\cdot, s)\|_{H^2(\mathbb{R}^2)}^3 ds.$$

**3.6. Bootstrapping $\eta$ to Existence.** Using the bounds we have now proven, we substitute (3.105), (3.112), (3.116), and (3.117) into the $H^2$ norm of the Picard iteration (3.32) to get

$$
\|\eta_{n+1}(\cdot,t)\|_{H^2} \leq C\rho\delta + C\int_0^t \delta\|\eta_n(\cdot,s)\|_{H^2(\mathbb{R}^2)} + \delta^{\frac{1}{2}}\|\eta_n(\cdot,s)\|_{H^2(\mathbb{R}^2)}^2
$$

(3.118)

$$
+ \|\eta_n(\cdot,s)\|_{H^2(\mathbb{R}^2)}^3 ds
$$

where we have defined $C = 2\max\limits_{m=1}^{5}\{C_m\}$.

Now we will prove Theorem 3.2.1 using the bootstrapping principle. We define the function $\phi(T) = \sup\limits_{0\leq t\leq T}\|\eta(\cdot,t)\|_{H^2(\mathbb{R}^2)}$. Now we define the hypothesis statement, $H(T)$, and the conclusion statement, $C(T)$.

- The statement, $H(T)$, is "there exists $\delta_0$ such that $\phi(T) < 2Ce^{2C\rho}\rho\delta$ for all $\delta < \delta_0$."

- The statement, $C(T)$, is "there exists $\delta_0$ such that $\phi(T) < Ce^{2C\rho}\rho\delta$ for all $\delta < \delta_0$."

We begin by showing there is a time $T \in [0, \rho\delta^{-1}]$ such that $H(T)$ is true. To do that we first need the following proposition.

PROPOSITION 3.6.1. *There exists $\delta_0$ such that for each step of the Picard iteration, $\eta_n(\mathbf{x},t)$, we have*

(3.119)
$$
\sup_{0\leq t\leq 1}\|\eta_n(\cdot,t)\|_{H^2(\mathbb{R}^2)} < 2C\rho\delta
$$

*for all $\delta \leq \delta_0$.*

PROOF. We prove the proposition by induction. To prove the base case observe that

(3.120)
$$
\sup_{0\leq t\leq 1}\|\eta_1(\cdot,t)\|_{H^2(\mathbb{R}^2)} \leq C\rho\delta.
$$

This proves the base case.

Now we assume that the proposition holds for $\eta_n$ and will show that it must also hold for $\eta_{n+1}$. We pick $\delta_0 < (2C + 4C^2\rho + 8C^3\rho^2)^{-1}$ and restrict $\delta \leq \delta_0$. Substituting in the inductive hypothesis into (3.118) we get

$$
\sup_{0\leq t\leq 1}\|\eta_{n+1}(\cdot,t)\|_{H^2(\mathbb{R}^2)} \leq C\rho\delta + C\int_0^1 \delta(2C\rho\delta) + \delta^{\frac{1}{2}}(2C\rho\delta)^2 + (2C\rho\delta)^3 ds
$$

(3.121)

$$
\leq C\rho\delta + C\rho\delta\left(2C + 4C^2\rho + 8C^3\rho^2\right)\delta
$$

$$
< 2C\rho\delta \qquad \text{since } \delta \leq \delta_0.
$$

This proves the proposition. □

Notice this shows that the sequence $(\eta_n)_{n=1}^\infty$ is uniformly bounded when $t \leq 1$ and $\delta \leq \delta_0$. If the sequence is also Cauchy, then it must converge.

To show the sequence is Cauchy we observe

$$(3.122) \qquad \eta_{n+1}(\cdot,t) - \eta_n(\cdot,t) = (A_{3,n+1} - A_{3,n}) + (A_{4,n+1} - A_{4,n}) + (A_{5,n+1} - A_{5,n})$$

since $A_1 - A_1 = A_2 - A_2 = 0$.

Substituting in the definitions (3.106), (3.107), and (3.108) we get

$$|A_{3,n+1}(\cdot,t) - A_{3,n}(\cdot,t)| \leq 3\delta \sum_{j,k=1}^2 \int_0^t \Big| (\eta_n(\cdot,s) - \eta_{n-1}(\cdot,s))$$
$$\times\, \alpha_j(\delta\cdot,\delta s)\alpha_k(\delta\cdot,\delta s)\Phi_j(\cdot)\Phi_k(\cdot)\Big| ds$$

$$|A_{4,n+1}(\cdot,t) - A_{4,n}(\cdot,t)| \leq 3\delta^{\frac12} \sum_{j=1}^2 \int_0^t \Big| (\eta_n(\cdot,s)^2 - \eta_{n-1}(\cdot,s)^2)$$
$$\times\, \alpha_j(\delta\cdot,\delta s)\Phi_j(\cdot)\Big| ds$$

$$|A_{5,n+1}(\cdot,t - A_{5,n}(\cdot,t)| \leq \int_0^t \Big| (\eta_n(\cdot,s)^3 - \eta_{n-1}(\cdot,s)^3)\Big| ds.$$

Taking the $H^2(\mathbb{R}^2)$ norm and the supremum over our timescale of both sides of (3.122) gives

$$(3.123) \qquad \begin{aligned} \sup_{t\leq 1}\|\eta_{n+1}(\cdot,t) - \eta_n(\cdot,t)\|_{H^2(\mathbb{R}^2)} &\leq \sup_{t\leq 1}\|A_{3,n+1} - A_{3,n}\|_{H^2(\mathbb{R}^2)} \\ &+ \sup_{t\leq 1}\|A_{4,n+1} - A_{4,n}\|_{H^2(\mathbb{R}^2)} \\ &+ \sup_{t\leq 1}\|A_{5,n+1} - A_{5,n}\|_{H^2(\mathbb{R}^2)}. \end{aligned}$$

We proceed as we did when bounding the nonlinear terms starting with a similar result to (3.112)

$$(3.124) \qquad \|A_{3,n+1}(\cdot,t) - A_{3,n}(\cdot,t)\|_{H^2(\mathbb{R}^2)} \leq C_3\delta t \sup_{0\leq s\leq t}\|\eta_n(\cdot,s) - \eta_{n-1}(\cdot,s)\|_{H^2(\mathbb{R}^2)}.$$

89

Taking the supremum over time and using our upper bound for $\delta \leq \delta_0$ gives

$$(3.125) \qquad \sup_{t \leq 1} \|A_{3,n+1}(\cdot, t) - A_{3,n}(\cdot, t)\|_{H^2(\mathbb{R}^2)} \leq \frac{C}{C^*} \sup_{t \leq 1} \|\eta_n(\cdot, s) - \eta_{n-1}(\cdot, s)\|_{H^2(\mathbb{R}^2)}$$

where $C^* = 2C + 4C^2\rho + 8C^3\rho^2$.

We also get a similar result to (3.116)

$$\|A_{4,n+1}(\cdot, t) - A_{4,n}(\cdot, t)\|_{H^2(\mathbb{R}^2)} \leq C_4 \delta^{\frac{1}{2}} t \sup_{0 \leq s \leq t} \|\eta_n(\cdot, s)^2 - \eta_{n-1}(\cdot, s)^2\|_{H^2(\mathbb{R}^2)}$$

$$(3.126) \qquad \leq C_4 \delta^{\frac{1}{2}} t \sup_{0 \leq s \leq t} \|\eta_n(\cdot, s) - \eta_{n-1}(\cdot, s)\|_{H^2(\mathbb{R}^2)}$$

$$\times \left( \|\eta_n(\cdot, s)\|_{H^2(\mathbb{R}^2)} + \|\eta_{n-1}(\cdot, s)\|_{H^2(\mathbb{R}^2)} \right),$$

and taking the supremum over time along with the upper bound for $\delta \leq \delta_0$ gives

$$\sup_{t \leq 1} \|A_{4,n+1}(\cdot, t) - A_{4,n}(\cdot, t)\|_{H^2(\mathbb{R}^2)}$$

$$(3.127) \qquad \leq C \delta^{\frac{1}{2}} \sup_{t \leq \rho \delta^{-1}} \|\eta_n(\cdot, s) - \eta_{n-1}(\cdot, s)\|_{H^2(\mathbb{R}^2)} \left( 4C\rho\delta \right)$$

$$\leq 4C^2 \rho \delta^{\frac{3}{2}} \sup_{t \leq \rho \delta^{-1}} \|\eta_n(\cdot, s) - \eta_{n-1}(\cdot, s)\|_{H^2(\mathbb{R}^2)}$$

$$\leq \frac{4C^2\rho}{C^*} \sup_{t \leq \rho \delta^{-1}} \|\eta_n(\cdot, s) - \eta_{n-1}(\cdot, s)\|_{H^2(\mathbb{R}^2)}.$$

Moving on to the next term we get a similar result to (3.117)

$$\|A_{5,n+1}(\cdot, t) - A_{5,n}(\cdot, t)\|_{H^2(\mathbb{R}^2)} \leq C_5 t \sup_{0 \leq s \leq t} \|\eta_n(\cdot, s)^3 - \eta_{n-1}(\cdot, s)^3\|_{H^2(\mathbb{R}^2)}$$

$$(3.128) \qquad \leq C_5 t \sup_{0 \leq s \leq t} \|\eta_n(\cdot, s) - \eta_{n-1}(\cdot, s)\|_{H^2(\mathbb{R}^2)}$$

$$\times \left( \|\eta_n(\cdot, s)\|_{H^2(\mathbb{R}^2)}^2 + \|\eta_{n-1}(\cdot, s)\|_{H^2(\mathbb{R}^2)}^2 \right),$$

and taking the supremum over time along with our upper bound for $\delta \leq \delta_0$ gives

$$\sup_{t \leq 1} \|A_{5,n+1}(\cdot, t) - A_{5,n}(\cdot, t)\|_{H^2(\mathbb{R}^2)}$$

$$(3.129) \qquad \leq C \sup_{t \leq \rho \delta^{-1}} \|\eta_n(\cdot, s) - \eta_{n-1}(\cdot, s)\|_{H^2(\mathbb{R}^2)} \left( 8C^2\rho^2\delta^2 \right)$$

$$\leq \frac{8C^3\rho^2}{C^*} \sup_{t \leq \rho \delta^{-1}} \|\eta_n(\cdot, s) - \eta_{n-1}(\cdot, s)\|_{H^2(\mathbb{R}^2)}.$$

Finally, substituting (3.125), (3.127), and (3.129) into (3.123) we have

$$(3.130) \qquad \sup_{t \leq \rho \delta^{-1}} \|\eta_{n+1}(\cdot, t) - \eta_n(\cdot, t)\|_{H^2(\mathbb{R}^2)} \leq \frac{C^* - C}{C^*} \sup_{t \leq \rho \delta^{-1}} \|\eta_n(\cdot, s) - \eta_{n-1}(\cdot, s)\|_{H^2(\mathbb{R}^2)}$$

showing that the sequence is Cauchy.

Since the sequence $(\eta_n)$ is Cauchy and uniformly bounded, it must converge to a solution, $\eta$. We can also use the fact that

$$\sup_{0 \leq t \leq 1} \|\eta_n(\cdot, t)\|_{H^2(\mathbb{R}^2)} \leq 2C\delta$$

for each $n$ to conclude that

$$(3.131) \qquad \sup_{0 \leq t \leq 1} \|\eta(\cdot, t)\|_{H^2(\mathbb{R}^2)} \leq 2C\delta \leq 2Ce^{2C\rho}\delta.$$

Thus there exists a time, $T$, in $[0, \rho\delta^{-1}]$ such that $H(T)$ is true.

Now we want to show that $H(T)$ implies $C(T)$ for any $T \in [0, \rho\delta^{-1}]$. In other words, we assume there exists a $\delta_0^*$ such that $\phi(T) < 2Ce^{2C\rho}\rho\delta$ and show that there exists $\delta_0$ such that $\phi(T) < Ce^{2C\rho}\rho\delta$.

Now we know we can take the limit as $n \to \infty$ of (3.118) to get

$$(3.132) \qquad \|\eta(\cdot, t)\|_{H^2} \leq C\rho\delta + C \int_0^t \delta \|\eta(\cdot, s)\|_{H^2(\mathbb{R}^2)} + \delta^{\frac{1}{2}} \|\eta(\cdot, s)\|_{H^2(\mathbb{R}^2)}^2 + \|\eta(\cdot, s)\|_{H^2(\mathbb{R}^2)}^3 ds$$

Remembering our definition $\phi(T) = \sup_{0 \leq t \leq T} \|\eta(\cdot, t)\|_{H^2}$, we get

$$(3.133) \qquad \phi(T) \leq C\rho\delta + C \int_0^t \delta\phi(T) + \delta^{\frac{1}{2}}\phi(T)^2 + \phi(T)^3 ds.$$

Assume there exists a $\delta_0^*$ such that $\phi(T) < 2Ce^{2C\rho}\rho\delta$. Restrict $\delta \leq \delta_0^*$. Thus we have

$$
\begin{aligned}
(3.134) \qquad \phi(T) &\leq C\rho\delta + C \int_0^t \delta\phi(T) + \delta^{\frac{1}{2}}\phi(T)(2Ce^{2C\rho}\rho\delta) + \phi(T)(2Ce^{2C\rho}\rho\delta)^2 ds \\
&= C\rho\delta + C\delta \int_0^t \left( 1 + 2Ce^{2C\rho}\rho\delta^{\frac{1}{2}} + 4C^2 e^{4C\rho}\rho^2\delta \right) \phi(T) ds.
\end{aligned}
$$

Now we choose $\delta_0$ such that $2Ce^{2C\rho}\rho\delta^{\frac{1}{2}} + 4C^2 e^{4C\rho}\rho^2\delta < 1$ for all $\delta \leq \delta_0$. Now we restrict $\delta \leq \delta_0$ and get

$$(3.135) \qquad \phi(T) < C\rho\delta + \int_0^t 2C\delta\phi(T) ds$$

for all $\delta \leq \delta_0$. Applying Gronwall's inequality gives us

$$\phi(T) \leq C\rho\delta e^{2C\delta t}$$

(3.136)

$$\leq Ce^{2C\rho}\rho\delta.$$

Thus, we have shown that $H(T)$ implies $C(T)$.

Lastly we want to show that $C(T)$ implies $H(T')$ for all $T'$ on an open neighborhood of $T$. By assumption, there exists $\delta_0$ such that $\phi(T) < Ce^{2C\rho}\rho\delta$ for all $\delta \leq \delta_0$. Thus, by continuity, we can extend $T$ to some $T' > T$ such that $\phi(T') < 2Ce^{2C\rho}\rho\delta$ with the same choice of $\delta_0$. This $T'$ defines such an open neighborhood.

Thus, by the bootstrapping principle, $C(T)$ holds for all $T \in [0, \rho\delta^{-1}]$. This concludes the proof of Theorem 3.2.1.

This shows that we have existence times $t \leq \rho\delta^{-1}$ for solutions to the nonlinear Schrödinger equation of the form given by the ansatz. In particular, the timescale of the solutions grows as we take smaller values of $\delta$. However, when we look at the dynamics of the Dirac solution within the ansatz, we see that it only has timescale $\delta(\rho\delta^{-1}) = \rho$. Thus, while the timescale of the Schrödinger solutions grows as $\delta \to 0$, the timescale of the Dirac dynamics stays at the constant, $\rho$. However, as stated in Theorem 3.2.1, $\rho$ can be any positive number. Thus this timescale for the nonlinear Dirac dynamics in the ansatz is the same as was found by Arbunich and Sparber in [**9**].

**3.7. Weakening the Nonlinearity.** In Theorem 3.2.1 we proved an ansatz solution for a growing timescale of the Schrödinger equation (3.22) as similar to one posed in an open question from [**32**]. We are able to prove the nonlinear Dirac dynamics exist for up to any constant time. However, we would like to be able to show the timescale of the nonlinear Dirac dynamics can come into play for a growing timescale as $\delta$ goes to zero. In this section, we will modify the Schrödinger equation by unbalancing the nonlinear part of the Schrödinger equation. In doing so, we will also need to unbalance the nonlinearity in the associated Fefferman and Weinstein-type Dirac equation in the same manner. We will conclude that weakening the nonlinearity slightly will allow us to lengthen the timescale established previously in this chapter sufficiently that the associated Dirac dynamics are effective for a growing timescale as well.

To unbalance the nonlinearity we will introduce a factor, $\delta^k$, onto the nonlinearity of the Schrödinger equation. This will allow us to allow us to use $k$ as a free variable giving additional decay on the $\eta$-dependent terms of the Picard iteration. For reasons to be shown later, we will

restrict $0 < k \leq 1$. We seek solutions of

$$(3.137) \qquad i\partial_t \psi = (-\Delta + V(x))\psi + \delta^k |\psi|^2 \psi$$

with a similar ansatz as for Theorem 3.2.1

$$
\begin{aligned}
\psi(\mathbf{x}, t) = \Bigg( & \delta^{\frac{1}{2}} \sum_{j=1}^{2} \alpha_j(\delta\mathbf{x}, \delta t)\Phi_j(\mathbf{x}) + \delta^{\frac{3}{2}} \sum_{j=1}^{2} \beta_j(\delta\mathbf{x}, \delta t)\Phi_j(\mathbf{x}) \\
& + \delta^{\frac{3}{2}}\alpha^{\perp}(\mathbf{x}, t) + \eta^{\delta}(\mathbf{x}, t) \Bigg) e^{-i\mu_* t}
\end{aligned}
$$

(3.23)

with initial data

$$(3.24) \qquad \psi(\mathbf{x}, 0) = \delta^{\frac{1}{2}} \sum_{j=1}^{2} \alpha_j(\delta\mathbf{x}, 0)\Phi_j(\mathbf{x}) + \delta^{\frac{3}{2}} \sum_{j=1}^{2} \beta_j(\delta\mathbf{x}, 0)\Phi_j(\mathbf{x}) + \delta^{\frac{3}{2}}\alpha^{\perp}(\mathbf{x}, 0).$$

The only differences ate that we now require the $\alpha$ to be a solutions of the unbalanced, $\delta$-dependent Dirac equation,

$$
\begin{aligned}
\partial_t \alpha_1 &= -\overline{\lambda_{\#}}(\partial_{x_1} + i\partial_{x_2})\alpha_2 - i\delta^k \left(\beta_1|\alpha_1|^2 + 2\beta_2|\alpha_2|^2\right)\alpha_1 \\
\partial_t \alpha_2 &= -\lambda_{\#}(\partial_{x_1} - i\partial_{x_2})\alpha_1 - i\delta^k \left(2\beta_2|\alpha_1|^2 + \beta_1|\alpha_2|^2\right)\alpha_2
\end{aligned}
$$

(3.138)

and before we redefine $\alpha^{\perp}$ we need to redefine the $\gamma_r$ with $\gamma_r^*$ for $1 \leq 5$ as follows.

$$
\begin{aligned}
\gamma_1^*(\mathbf{x}, t) &= \gamma_1(\mathbf{x}, t) \\
\gamma_2^*(\mathbf{x}, t) &= \gamma_2(\mathbf{x}, t) \\
\gamma_3^*(\mathbf{x}, t) &= \gamma_3(\mathbf{x}, t) \\
\gamma_4^*(\mathbf{x}, t) &= \gamma_4(\mathbf{x}, t) \\
\gamma_5^*(\mathbf{x}, t) &= \delta^k \cdot \gamma_5(\mathbf{x}, t).
\end{aligned}
$$

(3.139)

Then we redefine $\alpha^{\perp}$, (3.28), by simply replacing $\gamma_r$ with $\gamma_r^*$ within the definition.

When we expand the ansatz into the modified Schrödinger equation (3.137) we get a very similar equation for $\eta$:

$$i\partial_t\eta - (H - \mu_*)\eta = -\delta^{\frac{3}{2}}\Big(\sum_{j=1}^{2} i\partial_T\alpha_j\Phi_j + 2\sum_{j=1}^{2}\nabla_{\mathbf{X}}\alpha_j\cdot\nabla_{\mathbf{x}}\Phi_j$$

$$-\delta^k\sum_{j,k,l=1}^{2}\alpha_j\alpha_k\overline{\alpha_l}\Phi_j\Phi_k\overline{\Phi_l} + (H-\mu_*)\alpha^{\perp}\Big)$$

$$-\delta^{\frac{5}{2}}\Big(\sum_{j=1}^{2} i\partial_T\beta_j\Phi_j + 2\sum_{j=1}^{2}\nabla_{\mathbf{X}}\beta_j\cdot\nabla_{\mathbf{x}}\Phi_j + \sum_{j=1}^{2}\Delta\alpha_j\Phi_j + \sum_{j=1}^{2}\partial_T\alpha^{\perp}\Big)$$

$$+\delta^{1+k}\Big(2\eta\sum_{j,k=1}^{2}\alpha_j\overline{\alpha_k}\Phi_j\overline{\Phi_k} + \overline{\eta}\sum_{j,k=1}^{2}\alpha_j\alpha_k\Phi_j\Phi_k\Big)$$

$$+\delta^{\frac{1}{2}+k}\Big(2|\eta|^2\sum_{j=1}^{2}\alpha_j\Phi_j + (\eta)^2\sum_{j=1}^{2}\overline{\alpha_j}\overline{\Phi_j}\Big) \quad + \delta^k\quad |\eta|^2\eta$$

(3.140)
$$+\delta^{\frac{5}{2}+k}\Big(2\sum_{j,k,l=1}^{2}\alpha_j\overline{\alpha_k}\beta_l\Phi_j\overline{\Phi_k}\Phi_l + \sum_{j,k,l=1}^{2}\alpha_j\alpha_k\overline{\beta_l}\Phi_j\Phi_k\overline{\Phi_l}\Big)$$

$$+\delta^{2+k}\Big(\sum_{j,k,l=1}^{2}\alpha_j\overline{\beta_k}\zeta_l\Phi_j\overline{\Phi_k} + \sum_{j,k,l=1}^{2}\overline{\alpha_j}\beta_k\zeta_l\overline{\Phi_j}\Phi_k + \sum_{j,k,l=1}^{2}\alpha_j\beta_k\overline{\zeta_l}\Phi_j\Phi_k\Big)$$

$$+\delta^{\frac{3}{2}+k}\Big(2\sum_{j,k,l=1}^{2}\beta_j\zeta_k\overline{\zeta_l}\Phi_j + \sum_{j,k,l=1}^{2}\overline{\beta_k}\zeta_k\zeta_l\overline{\Phi_k}\Big)$$

$$+\delta^{\frac{7}{2}+k}\Big(2\sum_{j,k,l=1}^{2}\beta_j\overline{\beta_k}\alpha_l\Phi_j\overline{\Phi_k}\Phi_l + \sum_{j,k,l=1}^{2}\beta_j\beta_k\overline{\alpha_l}\Phi_j\Phi_k\overline{\Phi_l}\Big)$$

$$+\delta^{3+k}\Big(2\sum_{j,k,l=1}^{2}\beta_j\overline{\beta_k}\zeta_l\Phi_j\overline{\Phi_k} + \sum_{j,k,l=1}^{2}\beta_j\beta_k\overline{\zeta_l}\Phi_j\Phi_k\Big)$$

$$+\delta^{\frac{9}{2}+k}\Big(3\sum_{j,k,l=1}^{2}\beta_j\overline{\beta_k}\beta_l\Phi_j\overline{\Phi_k}\Phi_l\Big).$$

where $\zeta_1 = \eta, \zeta_2 = \delta^{\frac{3}{2}}\alpha^{\perp}$, $\eta(\mathbf{x},0) = 0$, each $\alpha_j$ and $\beta_j$ are evaluated at $(\delta\mathbf{x},\delta t)$, and each $\Phi_j$ is evaluated at $(\mathbf{x})$ as before. The difference is that each piece originating from the Schrödinger nonlinearity now has an additional $\delta^k$ factor.

This leads us to our modification of Theorem 3.2.1 with longer existence times.

THEOREM 3.7.1. *Assume that $\alpha$ is a solution to (3.138) and $\beta$ is a solution to (3.29) with zero initial data and $\alpha^{\perp}$ defined by as above. Consider the NLS equation (3.137) where $V$ is a*

*honeycomb lattice potential. Fix $A_0 > 12$. Assume initial conditions, $\psi_0$, of the form (3.24) where $\alpha(\cdot, 0) \in H^s$ for $s$ large and has compact support. Restrict $0 < k \leq 1$ and $\epsilon > 0$ small enough such that $-\frac{k}{3} + \epsilon < 0$. Then there exists $\delta_0$ such that*

$$
(3.141) \qquad \sup_{0 \leq t \leq \delta^{-1-\frac{k}{3}+\epsilon}} \|\eta^\delta(\mathbf{x}, t)\|_{H^2(\mathbb{R}^2)} \leq \delta^{\frac{1}{2}-\frac{k}{3}}
$$

*for all $\delta \leq \delta_0$.*

Similarly to Theorem 3.2.1, we start by defining a Picard iteration for the integral equation for $\eta$:

$$
i\eta_{n+1}(\cdot, t) =
$$

$$
-\delta^{\frac{3}{2}} \sum_{j=1}^{2} \int_0^t e^{-i(H-\mu_*)(t-s)} \bigg( i\partial_T \alpha_j(\delta\cdot, \delta s)\Phi_j + 2\nabla_{\mathbf{X}}\alpha_j(\delta\cdot, \delta s) \cdot \nabla_{\mathbf{x}}\Phi_j
$$

$$
- \delta^k \sum_{k,l=1}^{2} \alpha_j(\delta\cdot, \delta s)\alpha_k(\delta\cdot, \delta s)\overline{\alpha_l}(\delta\cdot, \delta s)\Phi_j\Phi_k\overline{\Phi_l} + (H - \mu_*)\alpha^\perp(\mathbf{x}, t) \bigg) ds
$$

$$
- \delta^{\frac{5}{2}} \sum_{j=1}^{2} \int_0^t e^{-i(H-\mu_*)(t-s)} \bigg( i\partial_T \beta_j(\delta\cdot, \delta s)\Phi_j + 2\nabla_{\mathbf{X}}\beta_j(\delta\cdot, \delta s) \cdot \nabla_{\mathbf{x}}\Phi_j
$$

$$
- \Delta_{\mathbf{X}}\alpha_j(\delta\cdot, \delta s)\Phi_j - \partial_T \alpha^\perp \bigg) ds
$$

$$
(3.142) \qquad + \delta^{1+k} \sum_{j,k=1}^{2} \int_0^t e^{-i(H-\mu_*)(t-s)} \bigg( 2\eta_n(\cdot, s)\alpha_j(\delta\cdot, \delta s)\overline{\alpha_k}(\delta\cdot, \delta s)\Phi_j\overline{\Phi_k}
$$

$$
+ \overline{\eta_n}(\cdot, s)\alpha_j(\delta\cdot, \delta s)\alpha_k(\delta\cdot, \delta s)\Phi_j\Phi_k \bigg) ds
$$

$$
+ \delta^{\frac{1}{2}+k} \sum_{j=1}^{2} \int_0^t e^{-i(H-\mu_*)(t-s)} \bigg( 2|\eta_n(\cdot, s)|^2\alpha_j(\delta\cdot, \delta s)\Phi_j
$$

$$
+ \eta_n(\cdot, s)^2\overline{\alpha_j}(\delta\cdot, \delta s)\overline{\Phi_j} \bigg) ds
$$

$$
+ \delta^k \int_0^t e^{-i(H-\mu_*)(t-s)}|\eta_n(\cdot, s)|^2\, \eta_n(\cdot, s)ds + \text{ Lower order terms}
$$

$$
:= -\mathcal{A}_1 - \mathcal{A}_2 + \mathcal{A}_{3,n+1} + \mathcal{A}_{4,n+1} + \mathcal{A}_{5,n+1} + \text{ Lower order terms.}
$$

**3.7.1. Finding the Corresponding Bounds.** If we define $\tilde{\alpha}(\mathbf{x}, t) = \delta^{\frac{k}{2}} \cdot \alpha(\mathbf{x}, t)$ we can see that $\alpha$ being a solution to (3.138) is equivalent to $\tilde{\alpha}$ being a solution to (3.25). If we assume that $\alpha(\mathbf{x}, 0)$ is of size 1, then $\tilde{\alpha}(\mathbf{x}, 0)$ is of size $\delta^{\frac{k}{2}}$. Thus $\tilde{\alpha}$ (and $\alpha$ by extension) has existence time of order $\delta^{-k}$

by our local existence results. This will give us enough existence time within the ansatz envelope to choose a $T^*$ to serve as a maximal existence time as in the proof of Theorem 3.2.1. Observing that, within the ansatz, $\alpha$ is evaluated at $\delta t$, we can use the timescale assumption in the statement of Theorem 3.7.1 to find $\delta t \leq \delta \delta^{-1-k/3+\epsilon} = \delta^{-k/3+\epsilon}$. Since $\frac{k}{3} - \epsilon \leq k$, for small $\delta$ we get $\delta^{-k/3+\epsilon} \leq \delta^{-k}$. As discussed above, $\alpha$ has a lifespan of order $\delta^{-k}$. This allows us to stay within the lifespan of $\alpha$, and we can thus choose $T^* = \delta^{-k/3+\epsilon}$ to be our uniform time bound.

We want to prove Theorem 3.7.1 by closely mirroring the proof of Theorem 3.2.1. Notice that $\mathcal{A}_2$ is formally identical to $A_2$ other than the small change to the definition of $\alpha^\perp$ and the different power of $\delta$. Following the same proof as for the bound of $\|A_2\|_{H^2(\mathbb{R}^2)}$ results in the bound

$$\|\mathcal{A}_2\|_{H^2(\mathbb{R}^2)} \leq C_2\left(\delta^{\frac{3}{2}} + \delta^2 t\right). \tag{3.143}$$

Also, notice that formally, $\mathcal{A}_{j,n+1} = \delta^k A_{j,n+1}$ for $j = 3, 4, 5$. Thus we follow the same proof as for when bounding $\|A_{3,n+1}\|_{H^2(\mathbb{R}^2)}, \|A_{4,n+1}\|_{H^2(\mathbb{R}^2)}$, and $\|A_{5,n+1}\|_{H^2(\mathbb{R}^2)}$ to get

$$\|\mathcal{A}_{3,n+1}(\cdot,t)\|_{H^2(\mathbb{R}^2)} \leq C_3 \delta^{1+k} t \sup_{0 \leq s \leq t} \|\eta_n(\cdot,s)\|_{H^2(\mathbb{R}^2)} \tag{3.144}$$

$$\|\mathcal{A}_{4,n+1}(\cdot,t)\|_{H^2(\mathbb{R}^2)} \leq C_4 \delta^{\frac{1}{2}+k} t \sup_{0 \leq s \leq t} \|\eta_n(\cdot,s)\|_{H^2(\mathbb{R}^2)}^2 \tag{3.145}$$

$$\|\mathcal{A}_{5,n+1}(\cdot,t)\|_{H^2(\mathbb{R}^2)} \leq C_5 \delta^k t \sup_{0 \leq s \leq t} \|\eta_n(\cdot,s)\|_{H^2(\mathbb{R}^2)}^3. \tag{3.146}$$

This leaves the bound for $\|\mathcal{A}_1\|_{H^2(\mathbb{R}^2)}$ where we need to take more care to verify that the proof can mirror that of Theorem 3.2.1. Similarly to our handling of the term $A_1$, starting with Equation 3.33, we can consider $\infty_{1,n+1}$ as the solution to the initial value problem

$$i\partial_t f(\mathbf{x},t) - (H - \mu_*) f(\mathbf{x},t) = \delta^{\frac{3}{2}}\left(\sum_r \gamma_r^*(\delta\mathbf{x}, \delta t)\Psi_r(\mathbf{x}) + (H - \mu_0)\alpha^\perp(\mathbf{x},t)\right)$$
$$f(\mathbf{x},0) = 0 \tag{3.147}$$

where we have defined

$$\gamma_1^*(\mathbf{x}, t) = \gamma_1(\mathbf{x}, t)$$

$$\gamma_2^*(\mathbf{x}, t) = \gamma_2(\mathbf{x}, t)$$

(3.148)
$$\gamma_3^*(\mathbf{x}, t) = \gamma_3(\mathbf{x}, t)$$

$$\gamma_4^*(\mathbf{x}, t) = \gamma_4(\mathbf{x}, t)$$

$$\gamma_5^*(\mathbf{x}, t) = \delta^k \cdot \gamma_5(\mathbf{x}, t)$$

referencing Equation 3.26.

Clearly, the $\gamma_j^*$ for $1 \leq r \leq 4$ can be handled just as the corresponding $\gamma_j$'s are handled in the proof of Theorem 3.2.1. Since the first four terms are identical to the previous proof, if we can show that $\gamma_5^*$ meets all the corresponding bounds of $\gamma_5$ then the same proof will hold for our $\delta$-modified case.

The first bound applied to $\gamma_5$ was Proposition 3.4.4. However, it is very simple to show that Proposition 3.4.4 still holds in the case of Theorem 3.7.1 (after making the obvious modifications such as changing the upper bound on $t$). The bound for the proposition is proven with Equation 3.52. However, simply noticing that

(3.149)
$$|\gamma_5^*| = \delta^k |\gamma_5| \leq |\gamma_5|$$

combined with Equation 3.52 gives us the same bound. Thus Proposition 3.4.4 still holds for this section. This also tell us the bounds for terms $f_{II,D}$ holds since its merely uses Proposition 3.4.4 to bound $\gamma_5$.

Other than while bounding $P_*(\sum_r f_{1,r})$, every other time the term $\gamma_5^*$ appears can be handled by simply using the inequality $\delta^k \leq 1$.

This just leaves us to deal with $P_*(\sum_r f_{1,r})$. Proceeding along the same line as in the proof of Theorem 3.2.1 we encounter the corresponding equation to Equation 3.56,

(3.150)
$$\widehat{\gamma_r^*}\left(\frac{\mathbf{k} - \mathbf{K}}{\delta}, \delta s\right) \cdot \langle p_\pm(\cdot; \mathbf{k}), \mathcal{P}_r(\cdot) \rangle_{L^2(\Omega)} =$$
$$\frac{1}{\sqrt{2}} \frac{\kappa_1 + i\kappa_2}{|\kappa|} \langle \Phi_1, \widehat{\gamma_r^*}\left(\frac{\kappa}{\delta}, \delta s\right) \cdot \Psi_r \rangle_{L^2(\Omega)} \pm \frac{1}{\sqrt{2}} \langle \Phi_2, \widehat{\gamma_r^*}\left(\frac{\kappa}{\delta}, \delta s\right) \cdot \Psi_r \rangle_{L^2(\Omega)}$$
$$+ \mathcal{O}\left(C_{0,\alpha}|\kappa|\right),$$

97

giving us several inner products of the form, $\langle \Phi, \widehat{\gamma_r^*}\left(\frac{\kappa}{\delta}, \delta s\right) \cdot \Psi_r \rangle_{L^2(\Omega)}$, with which we have to be careful. As $\gamma_r^*$ is a scalar we can pull it outside of the inner product. Then just as before, the terms $1 \leq r \leq 4$ are expanded in [**32**] and many of other $r = 5$ terms vanish due to symmetry considerations. The remaining terms, analogous to Equation 3.58, are

(3.151)
$$\begin{aligned}
\langle \Phi_1, \delta^k \widehat{\alpha_1 \alpha_1 \overline{\alpha_1}} \Phi_1 \Phi_1 \overline{\Phi_1} \rangle = \delta^k \beta_1 |\alpha_1|^2 \alpha_1 \qquad & \langle \Phi_2, \delta^k \widehat{\alpha_2 \alpha_2 \overline{\alpha_2}} \Phi_2 \Phi_2 \overline{\Phi_2} \rangle = \delta^k \beta_1 |\alpha_2|^2 \alpha_2 \\
\langle \Phi_1, \delta^k \widehat{\alpha_2 \alpha_1 \overline{\alpha_2}} \Phi_2 \Phi_1 \overline{\Phi_2} \rangle = \delta^k \beta_2 |\alpha_2|^2 \alpha_1 \qquad & \langle \Phi_2, \delta^k \widehat{\alpha_2 \alpha_1 \overline{\alpha_1}} \Phi_2 \Phi_1 \overline{\Phi_1} \rangle = \delta^k \beta_2 |\alpha_1|^2 \alpha_2 \\
\langle \Phi_1, \delta^k \widehat{\alpha_1 \alpha_2 \overline{\alpha_2}} \Phi_1 \Phi_2 \overline{\Phi_2} \rangle = \delta^k \beta_2 |\alpha_2|^2 \alpha_1 \qquad & \langle \Phi_2, \delta^k \widehat{\alpha_1 \alpha_2 \overline{\alpha_1}} \Phi_1 \Phi_2 \overline{\Phi_1} \rangle = \delta^k \beta_2 |\alpha_1|^2 \alpha_2.
\end{aligned}$$

Plugging these into the equation corresponding to (3.56) we get the corresponding equation to Equation 3.59

(3.152)
$$\begin{aligned}
\sum_r \hat{\Gamma}_r &\left(\frac{\kappa}{\delta}, \delta s\right) \cdot \langle p_\pm(\cdot; \mathbf{k}), \mathcal{P}_r(\cdot) \rangle \\
&= \frac{i}{\sqrt{2}} \frac{\kappa_1 + i\kappa_2}{|\kappa|} \Big[ \widehat{\partial_t \alpha_1} + \overline{\lambda_\#} \left( \widehat{\partial_{x_1} \alpha_2} + i \widehat{\partial_{x_2} \alpha_2} \right) \\
&\qquad + i\delta^k \left( \beta_1 \widehat{|\alpha_1|^2 \alpha_1} + 2\beta_2 \widehat{|\alpha_2|^2 \alpha_1} \right) \Big] \left(\frac{\kappa}{\delta}, \delta s\right) \\
&\quad \pm \frac{i}{\sqrt{2}} \Big[ \widehat{\partial_t \alpha_2} + \lambda_\# \left( \widehat{\partial_{x_1} \alpha_1} - i \widehat{\partial_{x_2} \alpha_1} \right) \\
&\qquad + i\delta^k \left( 2\beta_2 \widehat{|\alpha_1|^2 \alpha_2} + \beta_1 \widehat{|\alpha_2|^2 \alpha_2} \right) \Big] \left(\frac{\kappa}{\delta}, \delta s\right).
\end{aligned}$$

Using the fact that $\alpha$ is now a solution to (3.138), both terms vanish leaving us with the result

(3.153)
$$\sum_r \hat{\Gamma}_r \left(\frac{\kappa}{\delta}, \delta s\right) \cdot \langle p_\pm(\cdot; \mathbf{k}), \mathcal{P}_r(\cdot) \rangle = 0$$

and

(3.154)
$$\sum_r f_{1,r} = 0$$

which is the same as we had in the proof for Theorem 3.2.1. After the proof is completed we will revisit Equation 3.152 to show that the above two bounds must be replaced by weaker bounds when $\alpha$ is a solution to the linear Dirac equation, (3.25).

This establishes that we can mirror the proof finding the bounds for $A_1$ giving us

(3.155)
$$\|\mathcal{A}_1\|_{H^2} \leq C_1 \delta^{\frac{5}{2}} t.$$

**3.7.2. Bootstrapping $\eta$ to a Longer Timescale.** We can now substitute (3.155), (3.143), (3.144), (3.145), and (3.146) into the norm of the Picard iteration (3.142). This gives us the bound

$$
\begin{aligned}
\|\eta_{n+1}(\cdot,t)\|_{H^2(\mathbb{R}^2)} \leq & \, C\delta^{\frac{5}{2}}t + C(\delta^{\frac{3}{2}} + \delta^2 t) + C\delta^{1+k}t \sup_{0\leq s\leq t} \|\eta_n(\cdot,s)\|_{H^2(\mathbb{R}^2)} \\
& + C\delta^{\frac{1}{2}+k}t \sup_{0\leq s\leq t} \|\eta_n(\cdot,s)\|^2_{H^2(\mathbb{R}^2)} + C\delta^k t \sup_{0\leq s\leq t} \|\eta_n(\cdot,s)\|^3_{H^2(\mathbb{R}^2)}
\end{aligned}
$$
(3.156)

where we have defined $C = \max\limits_{m=1}^{5}\{C_m\}$.

We will show that this the sequence $(\eta_n)_{j=1}^{\infty}$ is uniformly bounded and Cauchy. Then we will know it must converge.

Now we choose $\delta_0$ small enough such that $\delta_0^{\epsilon} \leq \frac{1}{6C}$.

PROPOSITION 3.7.2. *For each step of the Picard iteration, $\eta_n(\mathbf{x},t)$, and any $\delta \leq \delta_0$ we have*

(3.157)
$$
\sup_{0\leq t\leq \delta^{-1-\frac{k}{3}+\epsilon}} \|\eta_n(\cdot,t)\|_{H^2(\mathbb{R}^2)} \leq \delta^{1-\frac{k}{3}}.
$$

PROOF. We prove the proposition by induction. To prove the base case observe that

$$
\begin{aligned}
\|\eta_1(\cdot,t)\|_{H^2(\mathbb{R}^2)} \leq & \, C\delta^{\frac{5}{2}}t + C(\delta^{\frac{3}{2}} + \delta^2 t) \\
\leq & \, C\delta^{\frac{3}{2}-\frac{k}{3}}\delta^{\epsilon} + C\delta^{\frac{3}{2}-\epsilon}\delta^{\epsilon} + C\delta^{1-\frac{k}{3}}\delta^{\epsilon} \\
\leq & \, 3C\delta^{\frac{1}{2}-\frac{k}{3}}\delta^{\epsilon} \\
\leq & \, 3C\delta^{\frac{1}{2}-\frac{k}{3}}\left(\frac{1}{6C}\right) \\
< & \, \delta^{\frac{1}{2}-\frac{k}{3}}.
\end{aligned}
$$
(3.158)

Thus the base case for Proposition 3.7.2 holds.

Now we assume that the proposition holds for $\eta_j$ and will show that it must also hold for $\eta_{n+1}$. Substituting the inductive hypothesis into (3.156) we get

$$
\begin{aligned}
\|\eta_{n+1}(\cdot,t)\|_{H^2(\mathbb{R}^2)} \leq\, & C\delta^{\frac{5}{2}}t + C(\delta^{\frac{3}{2}} + \delta^2 t) + C\delta^{1+k}t(\delta^{1-\frac{k}{3}}) \\
& + C\delta^{\frac{1}{2}+k}t(\delta^{1-\frac{k}{3}})^2 + C\delta^k t(\delta^{1-\frac{k}{3}})^3 \\
\leq\, & C\delta^{\frac{5}{2}}t + C\delta^{\frac{3}{2}} + C\delta^2 t + C\delta^{2+\frac{2k}{3}}t \\
& + C\delta^{\frac{5}{2}+\frac{k}{3}}t + C\delta^3 t \\
\leq\, & C\delta^{\frac{3}{2}-\frac{k}{3}}\delta^\epsilon + C\delta^{\frac{3}{2}-\epsilon}\delta^\epsilon + C\delta^{1-\frac{k}{3}}\delta^\epsilon + C\delta^{1+\frac{k}{3}}\delta^\epsilon \\
& + C\delta^{\frac{3}{2}}\delta^\epsilon + C\delta^{2-\frac{k}{3}}\delta^\epsilon \\
\leq\, & 6C\delta^{1-\frac{k}{3}}\delta^\epsilon \\
\leq\, & 6C\delta^{1-\frac{k}{3}}\left(\frac{1}{6C}\right) \\
=\, & \delta^{1-\frac{k}{3}}.
\end{aligned}
$$

(3.159)

This proves the proposition. □

Thus the sequence $(\eta_n)_{j=1}^\infty$ is uniformly bounded.

To show the sequence is Cauchy we first remember that we have

(3.160)
$$
|\mathcal{A}_{j,n+1}| = \delta^k |A_{j,n+1}| \leq |A_{j,n+1}|
$$

for $j = 3, 4, 5$. Using this, we can adjust the corresponding section of Theorem 3.2.1.

As before, since the sequence $(\eta_n)_{j=1}^\infty$ is Cauchy and uniformly bounded, it must converge to a solution, $\eta$. We can also use the fact that

$$
\sup_{0 \leq t \leq \delta^{-1-\frac{k}{3}+\epsilon}} \|\eta_n(\cdot,t)\|_{H^2(\mathbb{R}^2)} \leq \delta^{1-\frac{k}{3}}
$$

for each $n$ to conclude that the same bound holds for $\eta$:

(3.161)
$$
\sup_{0 \leq t \leq \delta^{-1-\frac{k}{3}+\epsilon}} \|\eta(\cdot,t)\|_{H^2(\mathbb{R}^2)} \leq \delta^{1-\frac{k}{3}}.
$$

This concludes the proof of Theorem 3.7.1

This result means that we are able to obtain a slight increase in the proven lifespan of the ansatz solutions and the Dirac dynamics when we are willing to unbalance the Dirac equation by making

100

the Dirac dynamics mainly linear and using a small nonlinear correction. As we discussed, the unbalanced Dirac equation will have solutions on a timescale of $\delta^{-k}$ whereas we are only showing that the Dirac dynamics only affect the NLS ansatz on the timescale, $\delta^{-\frac{k}{3}+\epsilon}$. We also note that while this unbalanced formulation does grant an extended timescale, it does so at the cost of having less control over the size of the error term when compared to Theorem 3.2.1. This intermediate regime lives somewhere between using the linear Dirac and the balanced nonlinear Dirac equations and in doing so gives small, measurable changes to both the solution and the timescale for which the error term is provably small.

**3.7.3. Why Does the Linear Dirac Not Suffice?** A logical question to ask would be: Why can we not simply let $\alpha$ be a solution to the linear Dirac equation? In fact, at only one step in the above section did we use the fact that $\alpha$ is a solution to the nonlinear Dirac instead of the linear Dirac. That is just after Equation 3.152 (which we duplicate below).

$$\sum_r \hat{\Gamma}_r\left(\frac{\kappa}{\delta}, \delta s\right) \cdot \langle p_\pm(\cdot; \mathbf{k}), \mathcal{P}_r(\cdot)\rangle$$

(3.152)
$$= \frac{i}{\sqrt{2}} \frac{\kappa_1 + i\kappa_2}{|\kappa|} \left[ \widehat{\partial_t \alpha_1} + \overline{\lambda_\#}\left(\widehat{\partial_{x_1}\alpha_2} + i\widehat{\partial_{x_2}\alpha_2}\right) \right.$$
$$\left. + ig\delta^k \left(\beta_1 \widehat{|\alpha_1|^2\alpha_1} + 2\beta_2 \widehat{|\alpha_2|^2\alpha_1}\right)\right]\left(\frac{\kappa}{\delta}, \delta s\right)$$
$$\pm \frac{i}{\sqrt{2}}\left[\widehat{\partial_t\alpha_2} + \lambda_\#\left(\widehat{\partial_{x_1}\alpha_1} - i\widehat{\partial_{x_2}\alpha_1}\right)\right.$$
$$\left. + ig\delta^k\left(2\beta_2\widehat{|\alpha_1|^2\alpha_2} + \beta_1\widehat{|\alpha_2|^2\alpha_2}\right)\right]\left(\frac{\kappa}{\delta}, \delta s\right).$$

The fact that $\alpha$ was a solution to (3.138) allowed us to completely eliminate both terms of the right-hand side. If we instead assumed that $\alpha$ was a solution to the linear, homogeneous Dirac equation we could still eliminate part of those terms, but we would be left with the cubic subterms within both terms.

However, in general that is not an insurmountable problem. The cubic terms are easily bounded (as we have bounded many similar terms in the proof above). In fact the (now re-existent) cubic terms would give us

(3.162)
$$\sum_r \hat{\Gamma}_r\left(\frac{\kappa}{\delta}, \delta s\right) \cdot \langle p_\pm(\cdot; \mathbf{k}), \mathcal{P}_r(\cdot)\rangle = \mathcal{O}(\delta^k)$$

and

(3.163)
$$\sum_r f_{1,r} = \mathcal{O}\Big(\delta^{-2+k}\chi(|\kappa| < \delta^\tau)\Big).$$

Thus the term $\sum_r f_{1,r}$ does not have zero contribution to the norm of $f_{I,D}$. This would replace Equation 3.62 with

(3.164)
$$\tilde{f}_{I,D}(\mathbf{k}, t) = \mathcal{O}\big(\delta^{-\frac{1}{2}+k} \cdot |t| \cdot \chi(|\kappa| < \delta^\tau)\big) + \mathcal{O}\big(\delta^{\frac{5}{2}}|t| \cdot \chi(|\kappa| < \delta^\tau)\big).$$

The first term of this equation would give a contribution to $\|f_{I,D}\|_{L^2}^2$ bounded by

(3.165)
$$\delta^{-1+2k}t^2 \int_{\mathcal{B}_h} \chi\big(|\mathbf{k} - \mathbf{K}| < \delta^\tau\big)d\mathbf{k} \lesssim \delta^{-1+2k}t^2\delta^{2\tau} = t^2\delta^{-1+2k+2\tau}.$$

Now we use the restrictions $k \leq 1$ and $\tau < 1$ to realize that even if we took $k = 1$ this would give a contribution to $\|f_{I,D}\|_{L^2}$ of size $\delta^{\frac{1}{2}+\tau}t$ which is itself greater than $\delta^{\frac{1}{2}+\frac{A_0-3}{A_0}}t$. Since we fixed $A_0$ in the statement of the theorem, this would not even allow the lifespan $t \leq \delta^{-1}$ so we could not even achieve constant time Dirac dynamics for the solutions.

REMARK 4. By keeping $k \leq 1$, (3.165) would result in a contribution to $\|f_{I,D}\|_{L^2}$ of size strictly worse than $\delta^{\frac{3}{2}}t$. Since this was one of the bounds restricting the bootstrapping, a worse bound would result in a smaller lifespan for which the bootstrapping would work. Thus the nonlinear Dirac equation provides a better provable lifespan than does the linear Dirac equation for the ansatz.

3.7.3.1. *Why does the linear Dirac equation not suffice in Theorem 3.2.1.* This could also lead us to ask: Why does it not suffice to have $\alpha$ be a solution to the linear Dirac for Theorem 3.2.1?

Let us consider Equation 3.59 from Theorem 3.2.1.

(3.59)
$$\sum_r \hat{\gamma}_r\Big(\frac{\kappa}{\delta}, \delta s\Big) \cdot \langle p_\pm(\cdot; \mathbf{k}), \mathcal{P}_r(\cdot)\rangle$$
$$= \frac{i}{\sqrt{2}}\frac{\kappa_1 + i\kappa_2}{|\kappa|}\Big[\widehat{\partial_t\alpha_1} + \overline{\lambda_\#}\big(\widehat{\partial_{x_1}\alpha_2} + i\widehat{\partial_{x_2}\alpha_2}\big)$$
$$+ i\big(\beta_1\widehat{|\alpha_1|^2\alpha_1} + 2\beta_2\widehat{|\alpha_2|^2\alpha_1}\big)\Big]\Big(\frac{\kappa}{\delta}, \delta s\Big)$$
$$\pm \frac{i}{\sqrt{2}}\Big[\widehat{\partial_t\alpha_2} + \lambda_\#\big(\widehat{\partial_{x_1}\alpha_1} - i\widehat{\partial_{x_2}\alpha_1}\big)$$
$$+ i\big(2\beta_2\widehat{|\alpha_1|^2\alpha_2} + \beta_1\widehat{|\alpha_2|^2\alpha_2}\big)\Big]\Big(\frac{\kappa}{\delta}, \delta s\Big).$$

As above, if we take $\alpha$ to be a solution to the linear Dirac we would get

(3.166) $$\tilde{f}_{I,D}(\mathbf{k},t) = \mathcal{O}\big(\delta^{-\frac{1}{2}} \cdot |t| \cdot \chi(|\kappa| < \delta^\tau)\big) + \mathcal{O}\big(\delta^{\frac{5}{2}}|t| \cdot \chi(|\kappa| < \delta^\tau)\big).$$

The first term of this equation would give a contribution to $\|f_{I,D}\|_{L^2}^2$ (and thus $\|A_1\|_{L^2}^2$) bounded by

(3.167) $$\delta^{-1}t^2 \int_{\mathcal{B}_h} \chi\big(|\mathbf{k} - \mathbf{K}| < \delta^\tau\big)d\mathbf{k} \lesssim \delta^{-1}t^2\delta^{2\tau} = t^2\delta^{2\tau-1}.$$

Thus, since $\tau < \frac{A_0-3}{A_0} < 1$, the resulting would be strictly worse than order $\delta^{\frac{1}{2}}t$ which is significantly more restrictive than the current bound of order $\delta^{\frac{5}{2}}t$. This would result in a dramatically shorter lifespan for which the bootstrapping could work.

## 3.8. Tight Binding Regime.

In this section we discuss how well the above machinery applies to the tight binding situation discussed in [2]. As a reminder, in that paper Ablowitz and Zhu discuss the NLS equation.

(3.168) $$i\partial_t\psi = (-\Delta + V(x))\psi + \delta|\psi|^2\psi$$

with ansatz solutions of the form

(3.169) $$\psi^\delta(\mathbf{x},t) = \left(\sum_{j=1}^2 \alpha_j\left(\frac{\delta}{C}\mathbf{x}, \delta t\right)\Phi_j(\mathbf{x}) + \eta^\delta(\mathbf{x},t)\right)e^{-i\mu_* t}$$

with initial data

(3.170) $$\psi^\delta(\mathbf{x},0) = \sum_{j=1}^2 \alpha_j\left(\frac{\delta}{C}\mathbf{x},0\right)\Phi_j(\mathbf{x})$$

where $V_0$, the strength of the potential $V(\mathbf{x})$, is very large. This is what we mean by the "tight binding" setting. If $V_0 \to \infty$ that is what is called the tight binding limit (which is not discussed in this paper). This section deals with the case where $V_0$ is large but finite.

In the above ansatz, the constant $C$ is the same as that in Equation 15 of Ablowitz and Zhu [2]. This is particularly important to note because when we compare Equation 15 in [2] to Equation 3.1 in [32] we see that $C$ in Ablowitz and Zhu is the same as $|\lambda_\#|$ in Fefferman and Weinstein. Most important about this is that in the appendix of [2] there is a result stating that as $V_0$ gets large, $C$ is close to $V_0 e^{-\sqrt{V_0}}$. In other words, for very strong potentials, $V(\mathbf{x})$, we conclude that

$C = |\lambda_\#|$ gets very small. In particular, for a $\delta$ sufficiently small we can pick an $a$ very small such that $\delta^a = C = |\lambda_\#|$.

We also need a slightly different Dirac equation:

$$\partial_t \alpha_1 = -\frac{\overline{\lambda_\#}}{|\lambda_\#|}(\partial_{x_1} + i\partial_{x_2})\alpha_2 - i\left(b_1|\alpha_1|^2 + 2b_2|\alpha_2|^2\right)\alpha_1$$

(3.171)

$$\partial_t \alpha_2 = -\frac{\lambda_\#}{|\lambda_\#|}(\partial_{x_1} - i\partial_{x_2})\alpha_1 - i\left(2b_2|\alpha_1|^2 + b_1|\alpha_2|^2\right)\alpha_2$$

where the $\lambda_\#$ terms have been normalized. We use this normalization of $\lambda_\#$ because otherwise a strong potential, $V(\mathbf{x})$, could unbalance the linear and nonlinear dynamics of the equation. Normalizing takes care of that potential problem.

However, if we were proceed with the particular ansatz proposed by Ablowitz and Zhu, we would come across the same problem as we would have done trying to use the exact ansatz proposed by Fefferman and Weinstein in the non tight binding setting. Thus we will compensate by using a multiscale expansion similar to what we did in the previous sections.

We create the multiscale ansatz:

$$\psi(\mathbf{x},t) = \left( \sum_{j=1}^{2} \alpha_j(\delta^{1-a}\mathbf{x}, \delta t)\Phi_j(\mathbf{x}) + \delta^{1-2a}\sum_{j=1}^{2} \beta_j(\delta^{1-a}\mathbf{x}, \delta t)\Phi_j(\mathbf{x}) \right.$$

(3.172)

$$\left. + \delta^{1-2a}\alpha^\perp(\mathbf{x},t) + \eta^\delta(\mathbf{x},t) \right)e^{-i\mu_* t}$$

with initial data

(3.173) $$\psi(\mathbf{x},0) = \sum_{j=1}^{2}\alpha_j(\delta^{1-a}\mathbf{x},0)\Phi_j(\mathbf{x}) + \delta^{1-a}\sum_{j=1}^{2}\beta_j(\delta^{1-a}\mathbf{x},0)\Phi_j(\mathbf{x}) + \delta^{1-a}\alpha^\perp(\mathbf{x},0)$$

where $\alpha$ will be assumed to be a solution to (3.171).

Again, in order to redefine $\alpha^\perp$ we first need to redefine $\gamma_r$ for $1 \leq r \leq 5$.

$$\gamma_1^\#(\mathbf{x},t) = \delta^a \cdot \gamma_1(\mathbf{x},t)$$

$$\gamma_2^\#(\mathbf{x},t) = \delta^a \cdot \gamma_2(\mathbf{x},t)$$

(3.174) $$\gamma_3^\#(\mathbf{x},t) = \gamma_3(\mathbf{x},t)$$

$$\gamma_4^\#(\mathbf{x},t) = \gamma_4(\mathbf{x},t)$$

$$\gamma_5^\#(\mathbf{x},t) = \delta^a \cdot \gamma_5(\mathbf{x},t)$$

referencing (3.26). Then we redefine $\alpha^\perp$, (3.28), by simply replacing $\gamma_r$ inside its definition by $\gamma_r^\#$.

We also assume that $\beta$ solves the forced, linear Dirac equation

(3.175)
$$\partial_t \beta_1 = -\frac{\overline{\lambda_\#}}{|\lambda_\#|}(\partial_{x_1} + i\partial_{x_2})\beta_2 + i\Delta\alpha_1 + i\langle\Phi_1, \partial_T\alpha^\perp\rangle_{L^2(\Omega)}$$

$$\partial_t \beta_2 = -\frac{\lambda_\#}{|\lambda_\#|}(\partial_{x_1} - i\partial_{x_2})\beta_1 + i\Delta\alpha_2 + i\langle\Phi_2, \partial_T\alpha^\perp\rangle_{L^2(\Omega)}.$$

In all of these equations by letting $\delta$ be small enough and carefully choosing small $a$ we have replaced $C$ within the ansatz proposed in [**2**] by $\delta^a$.

When we expand the ansatz (3.172) into the NLS equation (3.168) we get

$$i\partial_t\eta - (H - \mu_*)\eta = -\delta^{1-a}\bigg(\sum_{j=1}^2 i\delta^a\partial_T\alpha_j\Phi_j + 2\sum_{j=1}^2 \nabla_{\mathbf{X}}\alpha_j \cdot \nabla_{\mathbf{x}}\Phi_j$$

$$-\sum_{j,k,l=1}^2 \delta^a\alpha_j\alpha_k\overline{\alpha_l}\Phi_j\Phi_k\overline{\Phi_l} + (H - \mu_*)\alpha^\perp\bigg)$$

$$-\delta^{2-3a}\bigg(\sum_{j=1}^2 i\delta^a\partial_T\beta_j\Phi_j + 2\sum_{j=1}^2 \nabla_{\mathbf{X}}\beta_j \cdot \nabla_{\mathbf{x}}\Phi_j + \sum_{j=1}^2 \delta^a\Delta\alpha_j\Phi_j + \sum_{j=1}^2 \delta^a\partial_T\alpha^\perp\bigg)$$

$$+\delta\bigg(2\eta\sum_{j,k=1}^2 \alpha_j\overline{\alpha_k}\Phi_j\overline{\Phi_k} + \overline{\eta}\sum_{j,k=1}^2 \alpha_j\alpha_k\Phi_j\Phi_k\bigg)$$

$$+\delta\bigg(2|\eta|^2\sum_{j=1}^2 \alpha_j\Phi_j + (\eta)^2\sum_{j=1}^2 \overline{\alpha_j\Phi_j}\bigg) \quad + \quad \delta|\eta|^2\eta$$

(3.176)
$$+\delta^{2-2a}\bigg(2\sum_{j,k,l=1}^2 \alpha_j\overline{\alpha_k}\beta_l\Phi_j\overline{\Phi_k}\Phi_l + \sum_{j,k,l=1}^2 \alpha_j\alpha_k\overline{\beta_l}\Phi_j\Phi_k\overline{\Phi_l}\bigg)$$

$$+\delta^{2-2a}\bigg(\sum_{j,k,l=1}^2 \alpha_j\overline{\beta_k}\zeta_l\Phi_j\overline{\Phi_k} + \sum_{j,k,l=1}^2 \overline{\alpha_j}\beta_k\zeta_l\overline{\Phi_j}\Phi_k + \sum_{j,k,l=1}^2 \alpha_j\beta_k\overline{\zeta_l}\Phi_j\Phi_k\bigg)$$

$$+\delta^{2-2a}\bigg(2\sum_{j,k,l=1}^2 \beta_j\zeta_k\overline{\zeta_l}\Phi_j + \sum_{j,k,l=1}^2 \overline{\beta_k}\zeta_k\zeta_l\overline{\Phi_k}\bigg)$$

$$+\delta^{3-4a}\bigg(2\sum_{j,k,l=1}^2 \beta_j\overline{\beta_k}\alpha_l\Phi_j\overline{\Phi_k}\Phi_l + \sum_{j,k,l=1}^2 \beta_j\beta_k\overline{\alpha_l}\Phi_j\Phi_k\overline{\Phi_l}\bigg)$$

$$+\delta^{3-4a}\bigg(2\sum_{j,k,l=1}^2 \beta_j\overline{\beta_k}\zeta_l\Phi_j\overline{\Phi_k} + \sum_{j,k,l=1}^2 \beta_j\beta_k\overline{\zeta_l}\Phi_j\Phi_k\bigg)$$

$$+\delta^{4-6a}\bigg(3\sum_{j,k,l=1}^2 \beta_j\overline{\beta_k}\beta_l\Phi_j\overline{\Phi_k}\Phi_l\bigg).$$

105

where $\zeta_1 = \eta, \zeta_2 = \delta^{1-2a}\alpha^{\perp}$, $\eta(\mathbf{x}, 0) = 0$, each $\alpha_j$ and $\beta_j$ are evaluated at $(\delta^{1-a}\mathbf{x}, \delta t)$, and each $\Phi_j$ is evaluated at $(\mathbf{x})$.

Once working with the corresponding integral equation of the upcoming Picard iteration we will be working with the time evolution operator, $e^{-it(H-\mu_*)}$, where $H = -\Delta + V(\mathbf{x})$. In the previous sections we did not need to worry about this because it is already established in [**32**] that the operator maps $H^2 \to H^2$. In other words $\|e^{-it(H-\mu_*)}f\|_{H^2} \lesssim \|f\|_{H^2}$ and the operator easily goes away in our estimates. We need to be a little careful since we are now assuming that $V(\mathbf{x})$ is very large. However, by working in Floquet-Bloch terms it is simple to show that $e^{-it(H-\mu_*)}$ behaves just like a phase (and thus effectively of size 1). As discussed in [**32**] we know from completeness of the Bloch modes that

$$(3.177) \qquad e^{-i(H-\mu_*)t}f = \sum_b \int_{\mathcal{B}_h} e^{-i(\mu_b(\mathbf{k})-\mu_*)t}\langle \Phi_b(\cdot; \mathbf{k}), f(\cdot)\rangle_{L^2(\mathbb{R}^2)}\Phi_b(\mathbf{x}, \mathbf{k})d\mathbf{k}.$$

Where the fact that $\Phi_b$ are eigenfunctions of $H$ allows us to turn $e^{-i(H-\mu_*)t}$ into $e^{-i(\mu_b(\mathbf{k})-\mu_*)t}$ on the Floquet-Bloch side of the equation. Most importantly, $|e^{-i(\mu_b(\mathbf{k})-\mu_*)t}| = 1$ and we can conclude that $|e^{-i(H-\mu_*)t}f| \approx |f|$. Thus we don't need to worry about the larger size of $V$ within $e^{-it(H-\mu_*)}$ in the tight binding section because it gets automatically dealt with by the Floquet-Bloch theory.

Now we state our modification of Theorem 3.2.1 for the tight binding regime

THEOREM 3.8.1. *Assume that $\alpha$ is a solution to* (3.171) *and $\beta$ is a solution to* (3.175) *with zero initial conditions. Fix $A_0 > 12$. Consider the NLS equation* (3.168) *where $V$ is a honeycomb lattice potential with $V_0$ very large as described above. Assume initial conditions, $\psi_0$, of the form* (3.173) *where $\alpha(\cdot, 0) \in H^s$ for large $s$ and has compact support. For all $\delta$ define $a = a_\delta = \frac{\ln(|\lambda_\#|)}{\ln(\delta)}$. Then for any $\rho > 0$ there exists $\delta_0$ such that there is a unique solution to* (3.168) *of the form* (3.172) *with*

$$(3.178) \qquad \sup_{0 \le t \le \rho\delta^{-1}} \|\eta^\delta(\mathbf{x}, t)\|_{H^2(\mathbb{R}^2)} = \mathcal{O}(\delta^{\frac{1}{2}})$$

*for all $\delta \le \delta_0$.*

Note that, as in Theorem 3.2.1, the constant implied by $\mathcal{O}$ depends only on $\rho$, $A_0$, the size of the initial data, and the size of the support of the initial data.

Of particular importance in this theorem (compared to Theorems 3.2.1 and 3.7.1) is that the scaling of the spatial and time variables differ. We also need to make the a priori assumption that $\delta$ is small enough relative to $|\lambda_\#|$ such that our choice $a = \frac{\ln(|\lambda_\#|)}{\ln(\delta)}$ is also very small.

106

Defining the Picard iteration and writing as an integral equation we get

$$
i\eta_{n+1}(\cdot, t) =
$$

$$
-\delta^{1-a} \sum_{j=1}^{2} \int_0^t e^{-i(H-\mu_*)(t-s)} \bigg( i\delta^a \partial_T \alpha_j(\delta^{1-a}\cdot, \delta s)\Phi_j + 2\nabla_{\mathbf{X}}\alpha_j(\delta^{1-a}\cdot, \delta s) \cdot \nabla_{\mathbf{x}}\Phi_j
$$

$$
- \sum_{k,l=1}^{2} \delta^a \alpha_j(\delta^{1-a}\cdot, \delta s)\alpha_k(\delta^{1-a}\cdot, \delta s)\overline{\alpha_l}(\delta^{1-a}\cdot, \delta s)\Phi_j\Phi_k\overline{\Phi_l}
$$

$$
+ (H-\mu_*)\alpha^\perp(\mathbf{x}, t) \bigg) ds
$$

$$
- \delta^{2-3a} \sum_{j=1}^{2} \int_0^t e^{-i(H-\mu_*)(t-s)} \bigg( i\delta^a \partial_T \beta_j(\delta^{1-a}\cdot, \delta s)\Phi_j + 2\nabla_{\mathbf{X}}\beta_j(\delta^{1-a}\cdot, \delta s) \cdot \nabla_{\mathbf{x}}\Phi_j
$$

(3.179)
$$
- \delta^a \Delta_{\mathbf{X}}\alpha_j(\delta^{1-a}\cdot, \delta s)\Phi_j - \delta^a \partial_T \alpha^\perp \bigg) ds
$$

$$
+ \delta \sum_{j,k=1}^{2} \int_0^t e^{-i(H-\mu_*)(t-s)} \bigg( 2\eta_n(\cdot, s)\alpha_j(\delta^{1-a}\cdot, \delta s)\overline{\alpha_k}(\delta^{1-a}\cdot, \delta s)\Phi_j\overline{\Phi_k}
$$

$$
+ \overline{\eta_n}(\cdot, s)\alpha_j(\delta^{1-a}\cdot, \delta s)\alpha_k(\delta^{1-a}\cdot, \delta s)\Phi_j\Phi_k \bigg) ds
$$

$$
+ \delta \sum_{j=1}^{2} \int_0^t e^{-i(H-\mu_*)(t-s)} \bigg( 2|\eta_n(\cdot, s)|^2 \alpha_j(\delta^{1-a}\cdot, \delta s)\Phi_j
$$

$$
+ \eta_n(\cdot, s)^2\overline{\alpha_j}(\delta^{1-a}\cdot, \delta s)\overline{\Phi_j} \bigg) ds
$$

$$
+ \delta \int_0^t e^{-i(H-\mu_*)(t-s)} |\eta_n(\cdot, s)|^2 \eta_n(\cdot, s) ds + \text{ Lower order terms}
$$

$$
:= -B_1 - B_2 + B_{3,n+1} + B_{4,n+1} + B_{5,n+1} + \text{ Lower order terms.}
$$

**3.8.1. Finding the Corresponding Bounds.** Observing that $\alpha$ is evaluated within the ansatz at $\delta t$ we can plug in our timescale bound from the statement of the theorem to see that $\alpha$ only gets evaluated on the timescale $\delta t \leq \delta\rho\delta^{-1} = \rho$. Thus our $\alpha$ timescale is actually bounded by $\rho$ just as in the proof for Theorem 3.2.1.

Following through the proof of Theorem 3.2.1 every step follows exactly the same other than a few modifications. The most obvious is that we have to be careful about having different powers of $\delta$ present, but that is a trivial (if tedious) modification. The second is that, like we did in the proof of Theorem 3.7.1, we make use of the fact that $\delta^a \leq 1$ to allow the proof to transfer over. The most technical modification is for the equations corresponding to (3.59) and (3.96).

The bounds for the $\eta$ dependent terms follow the exact same argument as in Theorem 3.2.1:

$$(3.180) \qquad \|B_{3,n+1}(\cdot,t)\|_{H^2(\mathbb{R}^2)} \leq C_3\delta \int_0^t \|\eta_n(\cdot,s)\|_{H^2(\mathbb{R}^2)} ds$$

$$(3.181) \qquad \|B_{4,n+1}(\cdot,t)\|_{H^2(\mathbb{R}^2)} \leq C_4\delta \int_0^t \|\eta_n(\cdot,s)\|^2_{H^2(\mathbb{R}^2)} ds$$

$$(3.182) \qquad \|B_{5,n+1}(\cdot,t)\|_{H^2(\mathbb{R}^2)} \leq C_5\delta \int_0^t \|\eta_n(\cdot,s)\|^3_{H^2(\mathbb{R}^2)} ds.$$

As mentioned above, using the fact that $\delta^a \leq 1$, the bounds for $\|B_1\|_{H^2}$ and $\|B_2\|_{H^2}$ follow the exact same argument is did $\|A_1\|_{H^2}$ and $\|A_2\|_{H^2}$ other than the terms $f_{1,r}$ and $g_{1,r}$. Paralleling Equation 3.33, we can consider $B_1$ as the solution to the initial value problem

$$
i\partial_t f(\mathbf{x},t) - (H - \mu_*)f(\mathbf{x},t) = \delta^{1-a}\left(\sum_r \gamma_r^{\#}(\delta\mathbf{x},\delta t)\Psi_r(\mathbf{x}) + (H - \mu_0)\alpha^{\perp}(\mathbf{x},t)\right)
$$
$$(3.183)$$
$$
f(\mathbf{x},0) = 0
$$

where we have defined

$$
\gamma_1^{\#}(\mathbf{x},t) = \delta^a \cdot \gamma_1(\mathbf{x},t)
$$
$$
\gamma_2^{\#}(\mathbf{x},t) = \delta^a \cdot \gamma_2(\mathbf{x},t)
$$
$$(3.184) \qquad \gamma_3^{\#}(\mathbf{x},t) = \gamma_3(\mathbf{x},t)$$
$$
\gamma_4^{\#}(\mathbf{x},t) = \gamma_4(\mathbf{x},t)
$$
$$
\gamma_5^{\#}(\mathbf{x},t) = \delta^a \cdot \gamma_5(\mathbf{x},t)
$$

referencing (3.26). Also, we can consider $B_2$ as the solution to the initial value problem

$$
i\partial_t g(\mathbf{x},t) - (H - \mu_*)g(\mathbf{x},t) = \delta^{2k-3a}\sum_r \nu_r^{\#}(\delta\mathbf{x},\delta t)\Psi_r(\mathbf{x})
$$
$$(3.185)$$
$$
g(\mathbf{x},0) = 0
$$

where we define

$$\nu_1^{\#}(\mathbf{x}, t) = \delta^a \cdot \nu_1(\mathbf{x}, t)$$

$$\nu_2^{\#}(\mathbf{x}, t) = \delta^a \cdot \nu_2(\mathbf{x}, t)$$

$$\nu_3^{\#}(\mathbf{x}, t) = \nu_3(\mathbf{x}, t)$$

$$\nu_4^{\#}(\mathbf{x}, t) = \nu_4(\mathbf{x}, t)$$

(3.186)

$$\nu_5^{\#}(\mathbf{x}, t) = \delta^a \cdot \nu_5(\mathbf{x}, t)$$

$$\nu_6^{\#}(\mathbf{x}, t) = \delta^a \cdot \nu_6(\mathbf{x}, t)$$

$$\nu_7^{\#}(\mathbf{x}, t) = \nu_7(\mathbf{x}, t)$$

$$\nu_8^{\#}(\mathbf{x}, t) = \delta^a \cdot \nu_8(\mathbf{x}, t)$$

referencing (3.90).

When we proceed just as in the proof of Theorem 3.2.1 we arrive at

$$\sum_r \widehat{\gamma_r^{\#}}\left(\frac{\kappa}{\delta^{1-a}}, \delta s\right) \cdot \langle p_{\pm}(\cdot; \mathbf{k}), \mathcal{P}_r(\cdot)\rangle$$

$$= \frac{i}{\sqrt{2}} \frac{\kappa_1 + i\kappa_2}{|\kappa|} \Big[ \delta^a \widehat{\partial_t \alpha_1} + \overline{\lambda_{\#}}\big(\widehat{\partial_{x_1} \alpha_2} + i\widehat{\partial_{x_2} \alpha_2}\big)$$

(3.187)

$$+ i\delta^a \big(b_1 \widehat{|\alpha_1|^2 \alpha_1} + 2b_2 \widehat{|\alpha_2|^2 \alpha_1}\big)\Big]\left(\frac{\kappa}{\delta}, \delta s\right)$$

$$\pm \frac{i}{\sqrt{2}} \Big[ \delta^a \widehat{\partial_t \alpha_2} + \lambda_{\#}\big(\widehat{\partial_{x_1} \alpha_1} - i\widehat{\partial_{x_2} \alpha_1}\big)$$

$$+ i\delta^a \big(2b_2 \widehat{|\alpha_1|^2 \alpha_2} + b_1 \widehat{|\alpha_2|^2 \alpha_2}\big)\Big]\left(\frac{\kappa}{\delta}, \delta s\right)$$

corresponding to (3.59). We also arrive at

$$\sum_r \widehat{\nu_r^{\#}}\left(\frac{\kappa}{\delta^{1-a}}, \delta s\right) \cdot \langle p_{\pm}(\cdot; \mathbf{k}), \mathcal{P}_r(\cdot)\rangle$$

$$= \frac{i}{\sqrt{2}} \frac{\kappa_1 + i\kappa_2}{|\kappa|} \Big[ \delta^a \widehat{\partial_t \alpha_1} + \overline{\lambda_{\#}}\big(\widehat{\partial_{x_1} \alpha_2} + i\widehat{\partial_{x_2} \alpha_2}\big)$$

(3.188)

$$- \delta^a \widehat{\Delta \alpha_1} - \delta^a \langle \Phi_1, \widehat{\partial_T \alpha^{\perp}}\rangle_{L^2(\Omega)}\Big]\left(\frac{\kappa}{\delta}, \delta s\right)$$

$$\pm \frac{i}{\sqrt{2}} \Big[ \delta^a \widehat{\partial_t \alpha_2} + \lambda_{\#}\big(\widehat{\partial_{x_1} \alpha_1} - i\widehat{\partial_{x_2} \alpha_1}\big)$$

$$- \delta^a \widehat{\Delta \alpha_2} - \delta^a \langle \Phi_2, \widehat{\partial_T \alpha^{\perp}}\rangle_{L^2(\Omega)}\Big]\left(\frac{\kappa}{\delta}, \delta s\right) + \mathcal{O}\left(|\kappa|\right)$$

corresponding to (3.96).

However, recalling that $\delta^a = |\lambda_\#|$ we are able to factor out a $\delta^a$ from within each set of brackets leaving $\frac{\lambda_\#}{|\lambda_\#|}$ in place of $\lambda_\#$ (and likewise for $\overline{\lambda_\#}$). Thus, considering $\alpha$ solves (3.171) and $\beta$ solves (3.175), both terms in (3.187) and the first two terms in (3.188) vanish as happened in (3.59) and (3.96). From that point the proof continues exactly as in the proof of Theorem 3.2.1 and we get the bounds

$$(3.189) \qquad\qquad \|B_1(\cdot,t)\|_{H^2(\mathbb{R}^2)} \leq C_1\rho\delta$$

$$(3.190) \qquad\qquad \|B_2(\cdot,t)\|_{H^2(\mathbb{R}^2)} \leq C_2\rho\delta^{\frac{1}{2}}.$$

Note that to do this we have assumed that $a$ is very small. This assumption is reasonable because of our a priori assumption that $\delta$ is assumed to be much smaller than $|\lambda_\#|$.

**3.8.2. Bootstrapping.** Substituting in the bounds we just found, (3.189), (3.190), (3.180), (3.181), and (3.182) to the Picard iteration we get

$$(3.191) \qquad \begin{aligned} \|\eta_{n+1}(\cdot,t)\|_{H^2} \leq{} & C\rho\delta^{\frac{1}{2}} + C\int_0^t \delta\|\eta_n(\cdot,s)\|_{H^2(\mathbb{R}^2)} + \delta\|\eta_n(\cdot,s)\|_{H^2(\mathbb{R}^2)}^2 \\ & + \delta\|\eta_n(\cdot,s)\|_{H^2(\mathbb{R}^2)}^3 \, ds \end{aligned}$$

where $C = 2\max\limits_{m=1}^{5}\{C_m\}$.

Just as we did for Theorem 3.2.1, we will prove Theorem 3.8.1 using the bootstrapping principle. We define the function $\phi(T) = \sup\limits_{0 \leq t \leq T} \|\eta(\cdot,t)\|_{H^2(\mathbb{R}^2)}$. Now we define the hypothesis statement, $H(T)$, and the conclusion statement, $C(T)$.

- The statement, $H(T)$, is "there exists $\delta_0$ such that $\psi(T) < 2Ce^{2C\rho}\rho\delta^{\frac{1}{2}}$ for all $\delta < \delta_0$."
- The statement, $C(T)$, is "there exists $\delta_0$ such that $\psi(T) < Ce^{2C\rho}\rho\delta^{\frac{1}{2}}$ for all $\delta < \delta_0$."

With the above hypothesis and conclusion statements the proof is identical to the bootstrapping proof for Theorem 3.2.1.

Thus we have now proved equivalent results in the tight binding setting using a multiscale expansion with the scalings suggested in [**2**] and in the non tight binding setting using a multiscale expansion with the scalings suggested in [**32**]. In both of these multiscale expansions we used a projectino onto a smaller perpendicular component than was done in the multiscale expanions of [**9**]. Both of these have the same existence time for the ansatz (growing as $\delta \to 0$) and for the nonlinear Dirac dynamics inside the envelope (up to any constant time, $\rho$).

# Numerical Simulations and Future Work

This final chapter will be broken into two main parts. The first part will demonstrate a numerical scheme for running simulations on the Schrödinger equation and the Dirac equation. Several aspects of these numerical simulations will be explored, but none of them will be particularly in depth. This is mainly meant to establish that we have a numerical scheme with which we can continue further investigations into the NLS and Dirac equations. The second part will be a brief discussion of future work that can be done.

**4.1. Strang Splitting Methods.** We will discuss the theory involved with setting up these simulations in the first subsection before presenting some of the basic results we have obtained in the following subsection.

**4.1.1. Introduction and Theory.** We start by establishing this method for a Schrödinger equation and Dirac equation which are more general than the specific ones we have used thus far. We consider the NLS equation

$$iu_t = (-\Delta + V)u + uf(|u|^2) \tag{4.1}$$

where $f : \mathbb{R} \to \mathbb{R}$ and $V \in L^2_{per}(\mathbb{R}^2)$ is a honeycomb lattice potential as in the previous chapter. If we take $f(|u|^2) = |u|^2$ this is precisely the NLS equation discussed in previous chapters.

We also consider the Dirac equation

$$iu_t - (\partial_x - i\partial_y)v = f_1(|u|^2, |v|^2)u$$
$$iv_t + (\partial_x + i\partial_y)u = f_2(|u|^2, |v|^2)v \tag{4.2}$$

where $f_1, f_2 : \mathbb{R}^2 \to \mathbb{R}$. As above for the NLS equation, a careful choice of $f_1$ and $f_2$ yields both formulations of the Dirac equation (Ablowitz-Nixon-Zhu and Fefferman-Weinstein) discussed in previous chapters (if we let $\lambda_\# = 1$).

Consider $i\partial_t |v|^2 = i\bar{v}\partial_t v + iv\partial_t \bar{v}$. Consider the nonlinear part of (4.1),

$$(4.3) \qquad\qquad iv_t = vf(|v|^2).$$

Taking the conjugate of (4.3),

$$(4.4) \qquad\qquad -i\bar{v}_t = \bar{v}f(|v|^2),$$

and substituting for $\partial_t v$ and $\partial_t \bar{v}$ everything on the right-hand side cancels, and we are left with

$$(4.5) \qquad\qquad i\partial_t |v|^2 = 0.$$

Thus we see that the amplitude is conserved which will be important for our numerical scheme.

We consider a Strang splitting method for the NLS equation. In other words, we will simulate (4.1) with a composition of the flows of the differential equations

$$(4.6) \qquad\qquad i\partial_t u = -\Delta u$$

and

$$(4.7) \qquad\qquad i\partial_t u = \left(V + f(|u|^2)\right) u.$$

More specifically we will numerically approximate $u(t_n)$ where $t_n = n\tau$ where $\tau$ is a step size greater than 0 and $n$ is the number of steps using the scheme

$$(4.8) \qquad
\begin{aligned}
u_{n+1/2}^{-} &= e^{\frac{i}{2}\tau\Delta} u_n \\
u_{n+1/2}^{+} &= e^{-i\tau\left(V + f(|u_{n+1/2}^{-}|^2)\right)} u_{n+1/2}^{-} \\
u_{n+1} &= e^{\frac{i}{2}\tau\Delta} u_{n+1/2}^{+}.
\end{aligned}$$

Because of the amplitude preserving property, we can conclude that this scheme is explicit and symmetric. The spatial discretization can be handled using a Fourier pseudo-spectral method. Thus the time flow $e^{\frac{i}{2}\tau\Delta}$ can be calculated well using the fast Fourier transform (FFT). Calculating the flow for the nonlinear part amounts to changing the phase of the solution at each point of the mesh.

The Strang splitting scheme [**74**] and higher order splitting schemes (e.g. [**76, 83**]) have been widely applied to a variety of nonlinear Schrödinger equations, mainly semilinear Schrödinger equations, modeling monochromatic light in nonlinear optics, Bose-Einstein condensates, as well as envelope solutions for surface wave trains in fluids. See for example [**7, 8, 11, 12, 13, 14, 15, 22, 30, 38, 43, 64, 65, 68, 71, 82**]. In this section we mainly focus on applying the Strang splitting method to a particular semilinear Schrödinger equation. However, there are many other time discretization approaches to solve non-linear evolution equations: Crank-Nicholson type schemes [**69**] (also [**4**], [**39**] for applications), Magnus expansion approaches [**63**], [**19**], exponential time-differencing schemes [**25, 44**], implicit-explicit methods [**10**], the comparison study in [**79**], and many others. The question of the convergence of the Strang splitting method has been thoroughly analyzed [**26, 33, 42, 60, 71, 80**].

To set up a Strang splitting scheme for the Dirac equation we consider flows of the differential equations:

(4.9)
$$iu_t = (\partial_x - i\partial_y)v$$
$$iv_t = -(\partial_x + i\partial_y)u$$

and

(4.10)
$$iu_t = f_1(|u|^2, |v|^2)u$$
$$iv_t = f_2(|u|^2, |v|^2)v.$$

Looking back at our fundamental solution operator of the linear Dirac equation in Chapter 1 we can see how the time flow acts for (4.9). Composing these flows as done for the Schrödinger equation above gives us

$$u^-_{n+1/2} = \cos\left(\frac{\tau}{2}\sqrt{-\Delta}\right) u_n + i\frac{\partial_x + i\partial_y}{|\partial_x + i\partial_y|} \sin\left(\frac{\tau}{2}\sqrt{-\Delta}\right) v_n$$

$$v^-_{n+1/2} = \cos\left(\frac{\tau}{2}\sqrt{-\Delta}\right) v_n + i\frac{\partial_x - i\partial_y}{|\partial_x - i\partial_y|} \sin\left(\frac{\tau}{2}\sqrt{-\Delta}\right) u_n$$

(4.11)
$$u^+_{n+1/2} = e^{-i\tau f_1\left(|u^-_{n+1/2}|^2,|v^-_{n+1/2}|^2\right)} u^-_{n+1/2}$$

$$v^+_{n+1/2} = e^{-i\tau f_2\left(|u^-_{n+1/2}|^2,|v^-_{n+1/2}|^2\right)} v^-_{n+1/2}$$

$$u_{n+1} = \cos\left(\frac{\tau}{2}\sqrt{-\Delta}\right) u^+_{n+1/2} + i\frac{\partial_x + i\partial_y}{|\partial_x + i\partial_y|} \sin\left(\frac{\tau}{2}\sqrt{-\Delta}\right) v^+_{n+1/2}$$

$$v_{n+1} = \cos\left(\frac{\tau}{2}\sqrt{-\Delta}\right) v^+_{n+1/2} + i\frac{\partial_x - i\partial_y}{|\partial_x - i\partial_y|} \sin\left(\frac{\tau}{2}\sqrt{-\Delta}\right) u^+_{n+1/2}.$$

Having set up the Strang splitting for the Dirac equation, it can be numerically calculated in the same manner as the Schrödinger equation.

**4.2. Numerical Results.** Now that we have discussed how the method is set up, we will include some preliminary numerical demonstrations. These will serve to support the analytical results from previous chapters as well as demonstrate that the method works and can be used for future work.

Throughout this section we will make the choices of $f, f_1, f_2$ to match the nonlinearities discussed in the previous chapters. In particular, we will work with $f(x) = x$ and choose $f_1, f_2$ to match the Ablowitz-Nixon-Zhu type nonlinearity.

**4.2.1. Dispersion Surfaces.** In order to compute the dispersion surfaces of the problem, we must first choose a specific potential, $V(\mathbf{x})$. We make the same choice as [1],

(4.12)
$$V(\mathbf{x}) = V_0|e^{ik_0\mathbf{b}_1\cdot\mathbf{x}} + e^{ik_0\mathbf{b}_2\cdot\mathbf{x}} + e^{ik_0\mathbf{b}_3\cdot\mathbf{x}}|$$

where $\mathbf{b}_1 = (0,1), \mathbf{b}_2 = (-\frac{\sqrt{3}}{2}, \frac{-1}{2}), \mathbf{b}_3 = (\frac{\sqrt{3}}{2}, \frac{-1}{2}), k_0 = \frac{32}{\sqrt{3}}$, and $V_0$ is the lattice intensity. We will make different choices for $V_0$ to demonstrate the dispersion surfaces in tight binding and non-tight binding regimes.

Using Floquet-Bloch theory as we did in Chapter 3, we are able to work in a unit cell in physical space and Fourier space instead of having to use the entire plane. Actually calculating the dispersion surfaces is done by using the *eigs* function in *MATLAB*.

Plots are given of the dispersion surfaces for $V_0 = 10, 100$, and $500$ in Figure 4.1. As expected, in the plots of the first two dispersion surfaces we can see that they touch at exactly two places for

114

(A) $V_0 = 10$        (B) $V_0 = 100$        (C) $V_0 = 500$

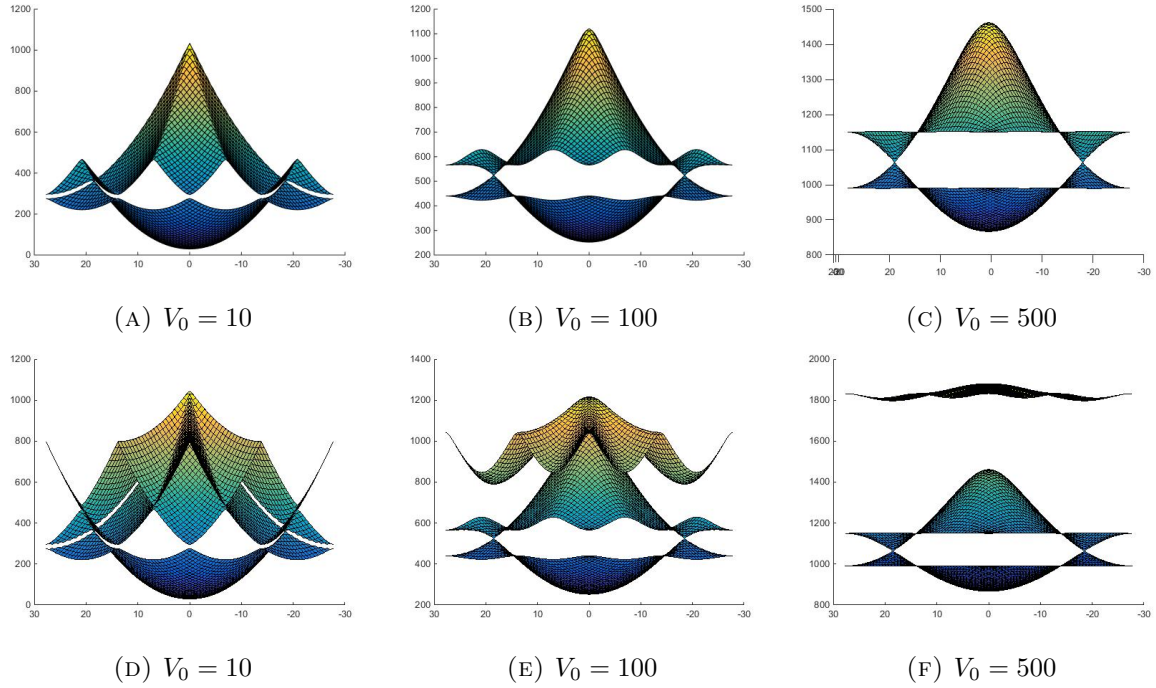(D) $V_0 = 10$        (E) $V_0 = 100$        (F) $V_0 = 500$

FIGURE 4.1. Plots of dispersion surfaces

(A) Plots of the dispersion surfaces for the NLS equation with $V_0 = 10, 100, 500$. The first row plots the first two dispersion surfaces (corresponding to $\mu_\pm(\mathbf{k})$ from Chapter 3). The second row additionally plots the third dispersion surface.

each choice of $V_0$. This point is the Dirac point, $\mathbf{K}$, discussed in Chapter 3 and [**32**]. With $V_0 = 10$ (non-tight binding) we see that the first two dispersion surfaces are very close together at many points in addition to the two points at which they agree. This can lead to much more "leakage" between Bloch modes in the solution and much more potential for poorly behaved solutions (as is to be expected in a non-tight binding setting). For $V_0 = 100, 500$ we see that the first two dispersion surfaces are farther apart (which is to expected as larger values of $V_0$ are more tight binding). We also see from the plots in the second row that as $V_0$ gets larger, the third dispersion surface gets farther away from the first two. This further limits the potential for "leakage" between Bloch modes.

**4.2.2. Dirac Experiments.** For our preliminary runs performing the Dirac simulations we will do two experiments. For both of these experiments we will use Gaussian initial data. The first is to demonstrate that we have radial symmetry when one input is Gaussian and the other zero, but that we do not have radial symmetry when both are Gaussian (just as it was for the linear equation in Chapter 1). The second is to explore how the size of the initial data affects the solutions.

(A) single-Gaussian, $t = 0$    (B) single-Gaussian, $t = 5$    (C) single-Gaussian, $t = 10$

(D) double-Gaussian, $t = 0$    (E) double-Gaussian, $t = 5$    (F) double-Gaussian, $t = 10$
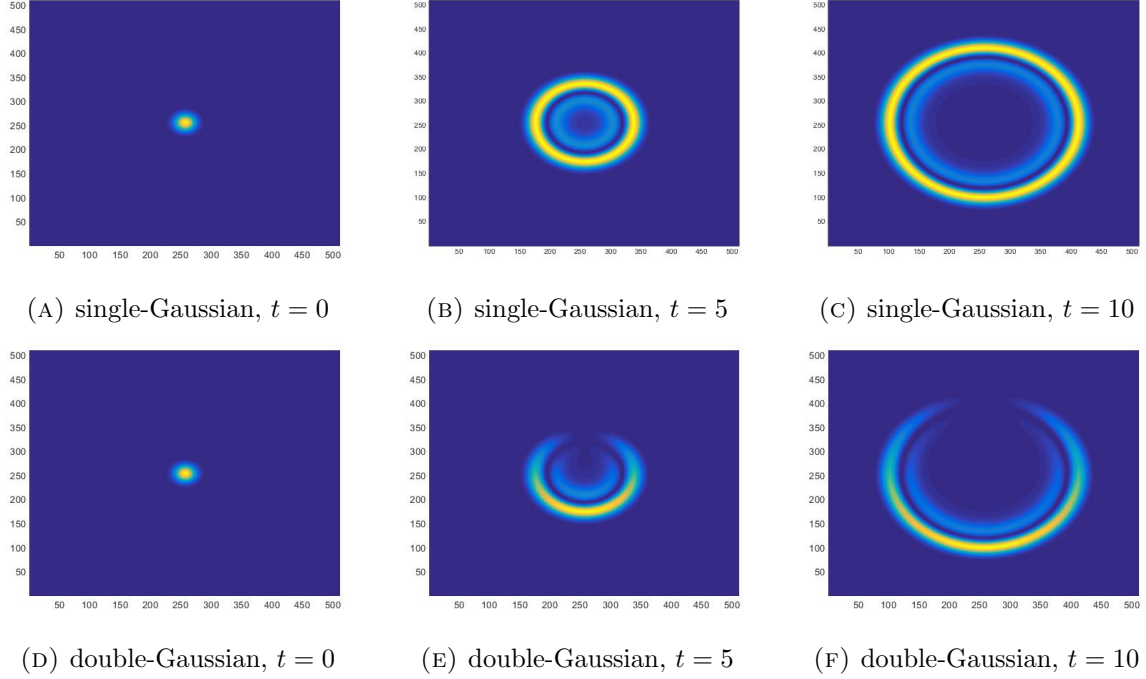
FIGURE 4.3. Evolution of the Dirac equation

(A) The first row demonstrates the evolution of single-Gaussian initial data. The second row demonstrates the evolution of double-Gaussian initial data.

4.2.2.1. *Radial Symmetry.* In [1], Ablowitz, Nixon, and Zhu demonstrate the radial symmetry properties of the linear Dirac equation. Here we will demonstrate the same radial symmetry results for the nonlinear Dirac equation.

We consider the case of the single-Gaussian initial data:

$$(4.13) \qquad u_0(\mathbf{x}) = A(2\pi)^{-1} e^{-\frac{1}{2}(x^2 + y^2)}, \quad v_0(\mathbf{x}) = 0$$

and the case of the double-Gaussian initial data

$$(4.14) \qquad u_0(\mathbf{x}) = v_0(\mathbf{x}) = A(2\pi)^{-1} e^{-\frac{1}{2}(x^2 + y^2)}$$

where $A$ is the initial data parameter. In this section $A$ is taken small enough that solutions behave nicely.

We see the results in Figure 4.3. We can clearly see the radial symmetry in the case of the single-Gaussian initial data. The dynamics appear identical to the results found in [1] for the linear Dirac. In addition to the radial symmetry, we have conical diffraction and two bright rings separated by Poggendorff's dark ring.

116

(A) $A = 5$



(B) $A = 25$


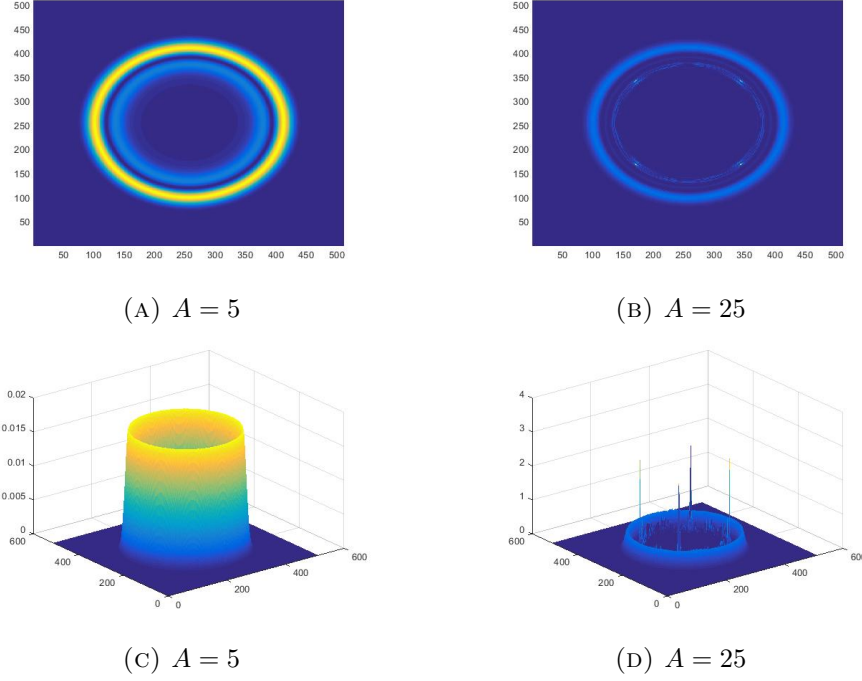
(C) $A = 5$



(D) $A = 25$

FIGURE 4.5. The effect of larger initial data on the simulation

(A) Results of running the Dirac simulation with single-Gaussian initial data of size $A = 5$ and $A = 25$. Each graph is at $t = 10$.

The case of double-Gaussian initial data is also very similar to the linear Dirac in [1]. We have the same behavior as the single-Gaussian case except for the radial symmetry. The introduction of a second, equal Gaussian to the initial data results in decay on one side of the ring to the extent that there is a visible "break" in the ring. Further simulations (not shown here) demonstrate that having two unequal Gaussians for initial data will still create decay on one side of the ring, but not to the extent of having a visible "break" as happens when the initial data are equal.

4.2.2.2. *Size of the Initial Data.* To explore the effects of different sizes of initial data on the simulation we choose Gaussian initial data:

$$(4.15) \qquad\qquad u_0(\mathbf{x}) = A(2\pi)^{-1}e^{-\frac{1}{2}(x^2+y^2)}$$

where $A$ is the initial data parameter, and $v_0(\mathbf{x}) = 0$.

We see the results of these simulations in Figure 4.5. With $A = 5$ we see that the solution behaves well. The plots agree with those found in [1] and [2] for the nonlinear Dirac equation. However, by increasing the size of the initial data to $A = 25$ we see that much of the dynamics breaks down. In the two-dimensional plot we see that overall conical diffraction still holds. However,

117
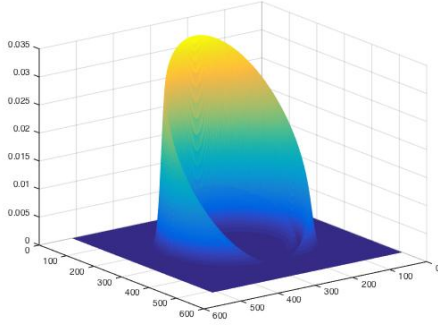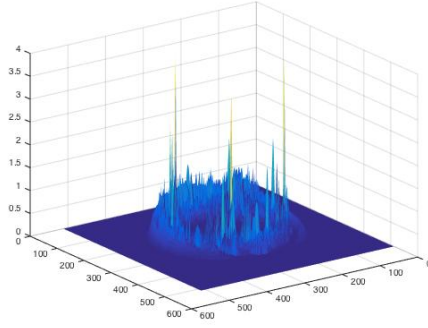
(A) Small data                                     (B) Large data

FIGURE 4.7. Effects of larger, double-Gaussian initial data

the three-dimensional plot shows that the relative growth of the inner and outer ring changes dramatically, and the radial symmetry observed for $A = 5$ does not hold for $A = 25$. By looking at the $z$-axis scaling of the plots we notice that the peaks in the three-dimensional plot for the larger initial data are two orders of magnitude higher than the maximum for the smaller initial data. That is much more than the difference in the size of the initial data. At this point it is hard to determine if this is a numerical error or something more to do with solutions of the Dirac problem.

We also run a simulation with equal, double-Gaussian initial data ($u_0 = v_0$) and $A$ scaled to maintain the same overall size.

These results are shown in Figure 4.7. We see that the small data yields a very nice conical diffraction weighted towards one side of the ring and decaying on the other side of the ring. However, the large data yields a result with all the problems visible in the single-Gaussian case as well as not being weighted towards one side of the ring as should be expected from the double-Gaussian case.

**4.2.3. Schrödinger Experiments.** In simulating solutions for the Schrödinger equation we will work with initial data like

$$(4.16) \qquad u_0(\mathbf{x}) = A\delta^{1/2} \sum_{j=1}^{2} e^{-\left((\delta x)^2 + (\delta y)^2\right)} \Phi_j(\mathbf{x})$$

where $A$ is again a size parameter, $\Phi_j$ are the Bloch modes discussed in Chapter 3, and the $\delta$ is included to match the setup from Chapter 3.

In actuality, the Gaussian is normalized differently to make it behave more nicely in the simulations. Also, it is difficult to compute the exact Bloch modes $\Phi_j$. It is much easier to find two
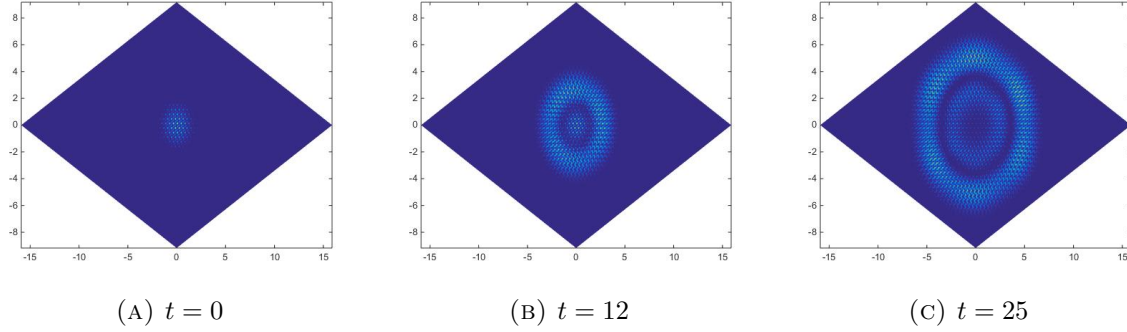
118

(A) $t = 0$      (B) $t = 12$      (C) $t = 25$

FIGURE 4.8. Small data, tight binding NLS evolution

(A) The evolution of the NLS equation, (4.16), displays clear Dirac dynamics with in the tight binding setting with small initial data ($A = 0.007$).

linearly independent eigenvectors of the eigenspace and use those. Making different choices of those eigenvectors is one way to vary the simulations and potentially get better results in future work.

In Chapter 3 we made the assumption that the initial data had compact support. Thus, the choice of Gaussian is not ideal for matching the simulations to the analytic framework of Chapter 3. However, we chose Gaussian initial data for these initial simulations for sake of simplicity. For some future work we would like to run simulations involving initial data with compact support as in Chapter 3.

We do three different simulations in this section. The first one is a small data simulation in the tight binding setting ($V_0 = 100$) where we observe behavior similar to the Dirac evolution numerically supporting the results seen in Chapter 3. In the second simulation we increase the size of the initial data and see the Dirac dynamics break down in the solution. For the final simulation we work with a very non-tight binding potential ($V_0 = 10$) and see that the Dirac dynamics are again absent.

**4.2.4. Tight Binding Small Data.** For this section we choose $V_0 = 100$ and $A = 0.007$. We run these simulations up to $t = 0.25$ with 500 timesteps to ensure high quality images.

The results are seen in Figure 4.8. These results display much of the properties seen in the Dirac simulations. The Gaussian initial data grows outward forming two concentric rings with a dark region (Poggendorff's Dark Ring) between them. Radial symmetry does not quite exist as we can see that the rings are slightly brighter at the vertices of a hexagon (corresponding to the locations of the vertices of the lattice). Also, the ring is slightly deformed. Although, based on the work in [1] we expect that we can get more circular results by changing the initial data or the

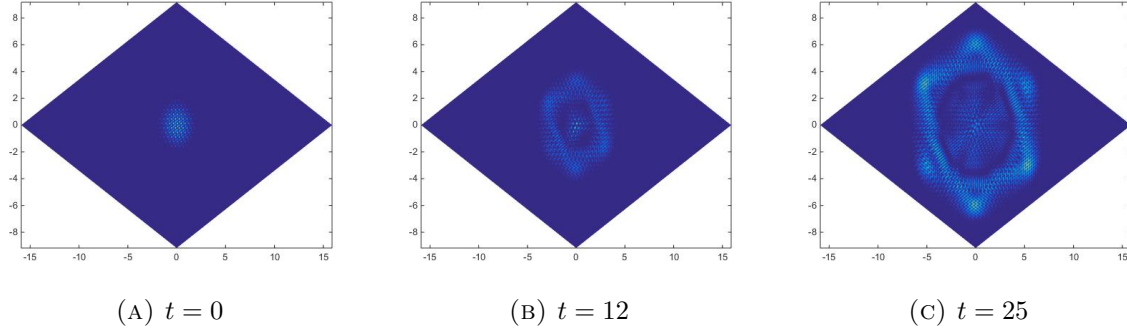(A) $t = 0$             (B) $t = 12$             (C) $t = 25$

FIGURE 4.10. Large data, tight binding NLS evolution

(A) The evolution of the NLS equation, (4.16), with tight binding potential ($V_0 = 100$) and large initial data ($A = 1$). The Dirac dynamics are still in evidence, but the Schrödinger dynamics have more control over the evolution.

potential function, $V$. Overall, the Dirac dynamics are very evident in the results of this run, and we think that in future work we can obtain results where the Dirac behavior is even more precisely displayed.

**4.2.5. Tight Binding Large Data.** For tight binding and large initial data we keep everything the same as the previous subsection except take $A = 1$.

The results of this run are displayed in Figure 4.10. We can still see evidence of an inner and an outer "ring." However the rings are now hexagonal as would be expected from Schrödinger dynamics (not Dirac dynamics). Clearly the lattice potential in the Schrödinger equation has taken over the dynamics of the evolution and the Dirac dynamics play a much smaller role. As we can see, a solution seems to exist on this timescale, but it is not governed by the Dirac dynamics as in the small data case.

Another large data simulation was run with larger data ($A = 10$). However, that evolution blew up on an extremely fast time scale indicating either very rapid blow up of solutions or even non-existence of solutions.

**4.2.6. Non-Tight Binding Small Data.** To simulate a small initial data run in the non-tight binding setting we kept $A = 0.007$, and we ran these simulations up to $t = 0.25$ with 500 timesteps to ensure high quality images. The only change was choosing $V_0 = 10$.

The results of the non-tight binding simulation are displayed in Figure 4.12. In this case, we see that the evolution seems to be entirely controlled by the Schrödinger dynamics. Based on Theorem 3.2.1 this is surprising because we know that the Dirac envelope provides a good ansatz solution

120

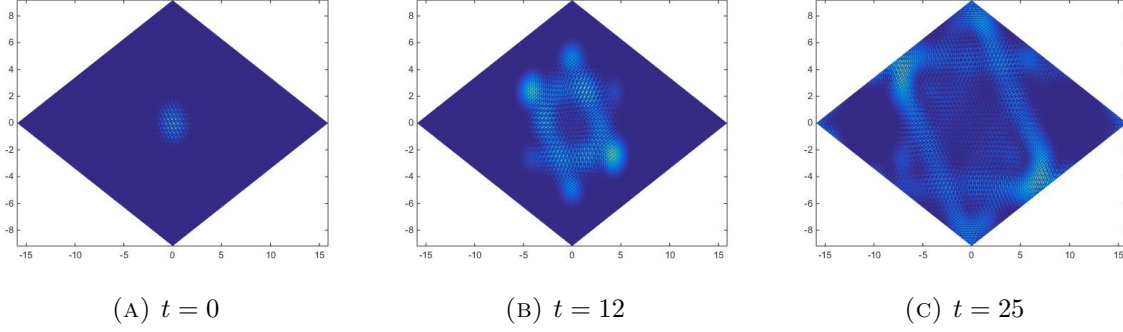(A) $t = 0$        (B) $t = 12$        (C) $t = 25$

FIGURE 4.12. Non-tight binding NLS evolution

(A) The evolution of the NLS equation, (4.16), with non-tight binding potential ($V_0 = 10$) and small initial data ($A = 0.007$).

to the Schrödinger equation. We expect that it is possible to obtain the Dirac dynamics in the non-tight binding setting by changing the initial data and/or the scaling of the equation. Another possibility is that letting $\lambda_\# = 1$ is a bad assumption numerically, and that constant needs to be calculated and included in the simulation. However, due to the closeness of the dispersion surfaces seen in Figure 4.2a, it is possible that these results are consistent with Theorem 3.2.1 based on the differences in the Bloch modes between the tight binding and non-tight binding settings. This will be left to future work.

**4.3. Conclusions and Future Work.** Chapter 1 of this dissertation focused primarily on establishing some of the basic properties of the linear Dirac system concluding with establishing Strichartz-type estimates using wave-theory.

Chapter 2 looked at the nonlinear Dirac problem and established several results. We proved a elementary local existence results in energy space as well as an extended-time, lower regularity local existence result making use of the Strichartz estimates proved in Chapter 1. We also proved an almost-global existence by making use of the wave equation form of the Dirac problem. We also established ill-posedness below the critical regularity space, $\dot{H}^{\frac{1}{2}}$, by showing that the solution is not continuously dependent on the initial data via contradiction. We concluded the chapter by discussing evidence for blow up of solutions, but leave the proof of a blow up result as future work.

Chapter 3 contains the main theorems of this dissertation. We continue the work of Fefferman and Weinstein [**32**] to show that the non linear Schrödinger equation can be well approximated to any constant time by an ansatz containing an envelope of solutions to the nonlinear Dirac equation. We extend the machinery modified from Fefferman and Weinstein to the ansatz proposed

by Ablowitz and Zhu in [**2**] to also prove the result specifically in the tight binding setting. We also apply this machinery to the situation where the Dirac equation has a weakened nonlinearity to extend the time beyond constant time to a timescale growing in $\delta$.

In Chapter 4 we do a few basic numerical simulations to support the results from previous chapters as well as hint at future work that can be done.

**4.3.1. Future Work.** Some potential future work has already been mentioned in previous chapters and sections including a finite time blow up result and improved numerical simulations as discussed previously in this section. Other potential topics of future work could include:

- Studying the lifespan of solutions to the Dirac equation with random initial data.

- Studying nonlinearities of non-local Hartree type as proposed by Arbunich and Sparber, [**9**]. In particular trying to apply the machinery from Chapter 3 to those nonlinearities.

- Utilize improved Schrödinger dispersive estimates to obtain longer existence and/or better error estimates for the Chapter 3 results (either tight binding or non-tight binding)

- More detailed study on the convergence of the numerical schemes used in this chapter. In particular the scheme used for the Dirac equation.

- Further numerical study of Dirac equations with large initial data.

- Further numerical study of longer existence times for both the Schrödinger and Dirac schemes with small initial data.

- More quantitative study analyzing the Dirac dynamics in the tight binding, small data NLS simulation. Directly comparing the NLS simulation with the Dirac simulation results applied as the envelope discussed in Chapter 3.

# REFERENCES

[1] M. Ablowitz, S. Nixon, and Y. Zhu. Conical diffraction in honeycomb lattices. *Physical Review A*, 79(5):053830, 2009.

[2] M. Ablowitz and Y. Zhu. Evolution of Bloch-mode envelopes in two-dimensional generalized honeycomb lattices. *Physical Review A*, 82(1):013840, 2010.

[3] R. Agemi. Blow-up of solutions to nonlinear wave equations in two space dimensions. *manuscripta mathematica*, 73(1):153–162, 1991.

[4] G. Akrivis, V. Dougalis, O. Karakashian, and W. McKinney. Numerical approximation of blow-up of radially symmetric solutions of the nonlinear Schrödinger equation. *SIAM Journal on Scientific Computing*, 25(1):186–212, 2003.

[5] S. Alinhac. The null condition for quasilinear wave equations in two space dimensions i. *Inventiones mathematicae*, 145(3):597–618, 2001.

[6] S. Alinhac. Semilinear hyperbolic systems with blowup at infinity. *Indiana University mathematics journal*, pages 1209–1232, 2006.

[7] X. Antoine, W. Bao, and C. Besse. Computational methods for the dynamics of the nonlinear Schrödinger/Gross–Pitaevskii equations. *Computer Physics Communications*, 184(12):2621–2633, 2013.

[8] H. Appel and E. Gross. Static and time-dependent many-body effects via density-functional theory. *NIC Series*, 10, 2002.

[9] J. Arbunich and C. Sparber. Rigorous derivation of nonlinear dirac equations for wave propagation in honeycomb structures. *arXiv preprint arXiv:1610.08542*, 2016.

[10] U. Ascher, S. Ruuth, and B. Wetton. Implicit-explicit methods for time-dependent partial differential equations. *SIAM Journal on Numerical Analysis*, 32(3):797–823, 1995.

[11] W. Bao and Y. Cai. Mathematical theory and numerical methods for Bose-Einstein condensation. *Kinetic and Related Methods*, 6(1):1–135, 2013.

[12] W. Bao, D. Jaksch, and P. Markowich. Numerical solution of the Gross–Pitaevskii equation for Bose–Einstein condensation. *Journal of Computational Physics*, 187(1):318–342, 2003.

[13] W. Bao, S. Jin, and P. Markowich. Numerical study of time-splitting spectral discretizations of nonlinear Schrödinger equations in the semiclassical regimes. *SIAM Journal on Scientific Computing*, 25(1):27–64, 2003.

[14] W. Bao, N. Mauser, H. Stimming, et al. Effective one particle quantum dynamics of electrons: A numerical study of the Schrodinger-Poisson-xalpha model. *Communications in Mathematical Sciences*, 1(4):809–828, 2003.

[15] W. Bao and J. Shen. A fourth-order time-splitting Laguerre–Hermite pseudospectral method for Bose–Einstein condensates. *SIAM Journal on Scientific Computing*, 26(6):2010–2028, 2005.

[16] I. Bejenaru and S. Herr. The cubic Dirac equation: Small initial data in $H^1(\mathbb{R}^3)$. *Communications in Mathematical Physics*, 335(1):43–82, 2015.

[17] I. Bejenaru and S. Herr. The cubic Dirac equation: Small initial data in $H^{\frac{1}{2}}(\mathbb{R}^2)$. *Communications in Mathematical Physics*, 343(2):515–562, 2016.

[18] I. Bejenaru and T. Tao. Sharp well-posedness and ill-posedness results for a quadratic nonlinear Schrödinger equation. *Journal of functional analysis*, 233(1):228–259, 2006.

[19] S. Blanes, F. Casas, J. Oteo, and J. Ros. The Magnus expansion and some of its applications. *Physics Reports*, 470(5):151–238, 2009.

[20] N. Bournaveas and T. Candy. Global well-posedness for the massless cubic Dirac equation. *International Mathematics Research Notices*, 2016(22):6735–6828, 2016.

[21] R. M. Chen, J. L. Marzuola, D. Spirn, and J. D. Wright. On the regularity of the flow map for the gravity-capillary equations. *Journal of Functional Analysis*, 264(3):752–782, 2013.

[22] S. A. Chin. Higher-order splitting algorithms for solving the nonlinear Schrödinger equation and their instabilities. *Physical Review E*, 76(5):056708, 2007.

[23] F. Christ and M. I. Weinstein. Dispersion of small amplitude solutions of the generalized Korteweg-de Vries equation. *Journal of functional analysis*, 100(1):87–109, 1991.

[24] M. Christ and A. Kiselev. Maximal functions associated to filtrations. *Journal of Functional Analysis*, 179(2):409–425, 2001.

[25] S. M. Cox and P. C. Matthews. Exponential time differencing for stiff systems. *Journal of Computational Physics*, 176(2):430–455, 2002.

[26] S. Descombes and M. Thalhammer. The Lie–Trotter splitting for nonlinear evolutionary problems with critical parameters: a compact local error representation and application to nonlinear Schrödinger equations in the semiclassical regime. *IMA Journal of Numerical Analysis*, 33(2):722–745, 2013.

[27] M. Escobedo and L. Vega. A semilinear Dirac equation in $H^s(\mathbb{R}^3)$ for $s > 1$. *SIAM Journal on Mathematical Analysis*, 28(2):338–362, 1997.

[28] L. C. Evans. *Partial differential equations*. American Mathematical Society, 2010.

[29] D. Fang and C. Wang. Some remarks on Strichartz estimates for homogeneous wave equation. *Nonlinear Analysis: Theory, Methods & Applications*, 65(3):697–706, 2006.

[30] E. Faou, V. Gradinaru, and C. Lubich. Computing semiclassical quantum dynamics with Hagedorn wavepackets. *SIAM Journal on Scientific Computing*, 31(4):3027–3041, 2009.

[31] C. L. Fefferman and M. I. Weinstein. Waves in honeycomb structures. *Journes quations aux drives partielles*, pages 1–12, 2012.

[32] C. L. Fefferman and M. I. Weinstein. Wave packets in honeycomb structures and two-dimensional Dirac equations. *Communications in Mathematical Physics*, 326(1):251–286, 2014.

[33] L. Gauckler. Convergence of a split-step Hermite method for the Gross–Pitaevskii equation. *IMA Journal of Numerical Analysis*, 31(2):396–415, 2011.

[34] P. Germain. Space-time resonances. *Journées Équations aux dérivées partielles*, pages 1–10, 2010.

[35] J. Ginibre and G. Velo. Generalized Strichartz inequalities for the wave equation. *Journal of Functional Analysis*, 133(1):50–68, 1995.

[36] R. T. Glassey. Existence in the large for $\Box u = f(u)$ in two space dimensions. *Mathematische Zeitschrift*, 178(2):233–261, 1981.

[37] P. Godin. Lifespan of solutions of semilinear wave equations in two space dimensions. *Communications in partial differential equations*, 18(5-6):895–916, 1993.

[38] R. Hardin and F. Tappert. Applications of the split-step Fourier method to the numerical solution of nonlinear and variable coefficient wave equations. *Siam Rev*, 15(2):423, 1973.

[39] J. Holmer, J. Marzuola, and M. Zworski. Soliton splitting by external delta potentials. *Journal of Nonlinear Science*, 17(4):349–367, 2007.

[40] A. Hoshiga et al. The existence of the global solutions to semilinear wave equations with a class of cubic nonlinearities in 2-dimensional space. *Hokkaido Mathematical Journal*, 37(4):669–688, 2008.

[41] T. J. Hughes, T. Kato, and J. E. Marsden. Well-posed quasi-linear second-order hyperbolic systems with applications to nonlinear elastodynamics and general relativity. *Archive for Rational Mechanics and Analysis*, 63(3):273–294, 1977.

[42] T. Jahnke and C. Lubich. Error bounds for exponential operator splittings. *BIT Numerical Mathematics*, 40(4):735–744, 2000.

[43] S. Jin, P. Markowich, and C. Sparber. Mathematical and computational methods for semiclassical Schrödinger equations. *Acta Numerica*, 20:121–209, 2011.

[44] A.-K. Kassam and L. N. Trefethen. Fourth-order time-stepping for stiff PDEs. *SIAM Journal on Scientific Computing*, 26(4):1214–1233, 2005.

[45] S. Katayama. Global existence for systems of nonlinear wave equations in two space dimensions. *Publ. RIMS, Kyoto Univ.*, 29:1021–1041, 1993.

[46] S. Katayama. Global existence for systems of nonlinear wave equations in two space dimensions, ii. *Publ. RIMS, Kyoto Univ.*, 31:645–665, 1995.

[47] S. Katayama. Lifespan of solutions for two space dimensional wave equations with cubic non-linearity. *Commun. in Part. Dirr. Eqs.*, 2001.

[48] S. Katayama. Global existence and asymptotic behavior of solutions to systems of semilinear wave equations in two space dimensions. *Hokkaido Math. J.*, 37(4):689–714, 2008.

[49] T. Kato. Quasi-linear equations of evolution, with applications to partial differential equations. *Spectral theory and differential equations*, pages 25–70, 1975.

[50] T. Kato. An $L^{q,r}$-theory for nonlinear Schrödinger equations. *Spectral and scattering theory and applications*, 23:223–238, 1994.

[51] M. Keel and T. Tao. Endpoint Strichartz estimates. *American Journal of Mathematics*, 120(5):955–980, 1998.

[52] S. Klainerman. Uniform decay estimates and the Lorentz invariance of the classical wave equation. *Communications on pure and applied mathematics*, 38(3):321–332, 1985.

[53] S. Klainerman. The null condition and global existence to nonlinear wave equations. *Lect. Appl. Math*, 23:293–326, 1986.

[54] M. Kovalyov. Long-time behaviour of solutions of a system of nonlinear wave equations. *Communications in partial differential equations*, 12(5):471–501, 1987.

[55] P. A. Kuchment. *Floquet theory for partial differential equations*, volume 60. Birkhäuser, 2012.

[56] A. Lawrie. Nonlinear wave equations, 2009.

[57] H. Lindblad. Blow-up for solutions of $\Box u = |u|^p$ with small initial data. *Communications in Partial Differential Equations*, 15(6):757–821, 1990.

[58] H. Lindblad. On the lifespan of solutions of nonlinear wave equations with small initial data. *Communications on Pure and Applied Mathematics*, 43(4):445–472, 1990.

[59] H. Lindblad and C. D. Sogge. On existence and scattering with minimal regularity for semi-linear wave equations. *Journal of Functional Analysis*, 130(2):357–426, 1995.

[60] C. Lubich. On splitting methods for Schrödinger-Poisson and cubic nonlinear Schrödinger equations. *Mathematics of computation*, 77(264):2141–2153, 2008.

[61] S. Machihara, M. Nakamura, K. Nakanishi, and T. Ozawa. Endpoint Strichartz estimates and global solutions for the nonlinear Dirac equation. *Journal of Functional Analysis*, 219(1):1–20, 2005.

[62] S. Machihara, K. Nakanishi, T. Ozawa, et al. Small global solutions and the nonrelativistic limit for the nonlinear Dirac equation. *Revista Matemática Iberoamericana*, 19(1):179–194, 2003.

[63] W. Magnus. On the exponential solution of differential equations for a linear operator. *Communications on pure and applied mathematics*, 7(4):649–673, 1954.

[64] R. I. McLachlan and G. R. W. Quispel. Splitting methods. *Acta Numerica*, 11:341–434, 2002.

[65] D. Pathria and J. L. Morris. Pseudo-spectral solution of nonlinear Schrödinger equations. *Journal of Computational Physics*, 87(1):108–125, 1990.

[66] H. Pecher. Local well-posedness for the nonlinear dirac equation in two space dimensions. *arXiv preprint arXiv:1303.1699*, 2013.

[67] D. Pelinovsky. Survey on global existence in the nonlinear Dirac equations in one dimension. *Harm. Anal. and Nonlinear Part. Dirr. Eqs., RIMS Kokyuroku Bessatsu*, B26:37–50, 2011.

[68] V. M. Pérez-Garcıa and X.-y. Liu. Numerical methods for the simulation of trapped nonlinear Schrödinger systems. *Applied Mathematics and Computation*, 144(2):215–235, 2003.

[69] J. Sanz-Serna. Methods for the numerical solution of the nonlinear Schrödinger equation. *mathematics of computation*, 43(167):21–27, 1984.

[70] S. Selberg. Lecture notes math 632 pde, 2001.

[71] J. Shen and Z.-Q. Wang. Error analysis of the Strang time-splitting Laguerre–Hermite/Hermite collocation methods for the Gross–Pitaevskii equation. *Foundations of Computational Mathematics*, 13(1):99–137, 2013.

[72] C. D. Sogge. *Lectures on non-linear wave equations*, volume 2. International Press Cambridge, MA, 1995.

[73] E. Stein. *Harmonic Analysis: Real-variable Methods, Orthogonality, and Oscillatory Integrals*. Princeton University Press, 1993.

[74] G. Strang. On the construction and comparison of difference schemes. *SIAM Journal on Numerical Analysis*, 5(3):506–517, 1968.

[75] R. S. Strichartz et al. Restrictions of Fourier transforms to quadratic surfaces and decay of solutions of wave equations. *Duke Mathematical Journal*, 44(3):705–714, 1977.

[76] M. Suzuki. Improved Trotter-like formula. *Physics Letters A*, 180(3):232–234, 1993.

[77] L. Ta-Tsien and Z. Yi. Life-span of classical solutions to nonlinear wave equations in two space dimensions, ii. *J. Part. Diff. Equations*, 6:17–38, 1993.

[78] L. Ta-Tsien and Z. Yi. Life-span of classical solutions to nonlinear wave equations in two space dimensions. *J. Math. Pures. Appl.*, pages 223–249, 1994.

[79] T. R. Taha and M. I. Ablowitz. Analytical and numerical aspects of certain nonlinear evolution equations. ii. numerical, nonlinear Schrödinger equation. *Journal of Computational Physics*, 55(2):203–230, 1984.

[80] M. Thalhammer. High-order exponential operator splitting methods for time-dependent Schrödinger equations. *SIAM Journal on Numerical Analysis*, 46(4):2022–2038, 2008.

[81] N. Tzvetkov. Existence of global solutions to nonlinear massless Dirac system and wave equation with small data. *Tsukuba journal of mathematics*, pages 193–211, 1998.

[82] J. Weideman and B. Herbst. Split-step methods for the solution of the nonlinear Schrödinger equation. *SIAM Journal on Numerical Analysis*, 23(3):485–507, 1986.

[83] H. Yoshida. Construction of higher order symplectic integrators. *Physics Letters A*, 150(5-7):262–268, 1990.

[84] Y. Zhou and W. Han. Sharpness on the lower bound of the lifespan of solutions to nonlinear wave equations. *Chinese Annals of Mathematics-Series B*, 32(4):521–526, 2011.