ON THE ASYMPTOTIC ANALYSIS OF THE DIRAC-MAXWELL SYSTEM IN THE NONRELATIVISTIC LIMIT

PHILIPPE BECHOUCHE, NOBERT J. MAUSER, AND SIGMUND SELBERG

ABSTRACT. We study the behavior of solutions of the Dirac-Maxwell system (DM) in the nonrelativistic limit $c \to \infty$, where c is the speed of light. DM is a nonlinear system of PDEs obtained by coupling the Dirac equation for a 4-spinor to the Maxwell equations for the self-consistent field created by the moving charge of the spinor. The limit $c \to \infty$, sometimes also called post-Newtonian, yields a Schrödinger-Poisson system, where the spin and magnetic field no longer appear.

We prove that DM is locally well-posed for H^1 data (for fixed c), and that as $c \to \infty$ the existence time grows at least as fast as $log(c)$, provided the data are uniformly bounded in H^1. Moreover, if the datum for the Dirac spinor converges in H^1, then the solution of DM converges, modulo a phase correction, in $C([0,T]; H^1)$ to a solution of a Schrödinger-Poisson system. Our results also apply to a mixed state formulation of DM.

The proof relies on modifications of the bilinear null form estimates of Klainerman and Machedon, and extends our previous work on the nonrelativistic limit of the Klein-Gordon-Maxwell system.

1. INTRODUCTION

In this paper we study the behavior of solutions to the Dirac-Maxwell (DM) system in the limit $c \to \infty$, where c is the speed of light. Coupled to the Coulomb gauge condition, this system has the form

$$\left(i\gamma^\mu \partial_\mu - M + g\gamma^\mu A_\mu\right)\psi = 0, \qquad \partial^\nu F_{\mu\nu} = J_\mu / c, \qquad \partial^j A_j = 0. \tag{1.1}$$

Here the unknowns are the spinor field $\psi(t,x) \in \mathbb{C}^4$, regarded as a column vector, and the electromagnetic potential $A_\mu(t,x) \in \mathbb{R}$, $\mu = 0,1,2,3$. Further, $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is the electromagnetic field tensor, and

$$J^\mu = c\left\langle \gamma^0 \gamma^\mu \psi, \psi \right\rangle_{\mathbb{C}^4}$$

is the 4-current density. On the Minkowski spacetime $\mathbb{R}^{1+3}$ we use relativistic coordinates $x^0 = ct \in \mathbb{R}$, $x = (x^1, x^2, x^3) \in \mathbb{R}^3$. ∂_μ stands for $\frac{\partial}{\partial x^\mu}$. Thus, $\partial_0 = \frac{1}{c}\partial_t$, where $\partial_t = \frac{\partial}{\partial t}$. We also write $\nabla = (\partial_1, \partial_2, \partial_3)$, $\Delta = \partial_1^2 + \partial_2^2 + \partial_3^2$ and $D^s = (-\Delta)^{s/2}$ for $s \in \mathbb{R}$. Indices are raised and lowered using the metric $(\eta_{\mu\nu}) = \operatorname{diag}(-1,1,1,1)$. The Einstein summation convention is in effect. Thus, repeated greek indices $\mu, \nu, \dots$ are summed over $0,1,2,3$, and repeated roman indices $j, k, \dots$ over $1,2,3$. For example, $\Delta = \partial_j \partial^j$. We denote by $\langle \cdot, \cdot \rangle_{\mathbb{C}^n}$ the standard inner product on $\mathbb{C}^n$.

The physical constants are $M = m_0 c/\hbar$, $g = e/\hbar c$, where m_0 is the spinor's rest mass, $\hbar$ is the Planck constant and e is the unit charge. By γ^μ, $\mu = 0,1,2,3$, we denote the

---

2000 *Mathematics Subject Classification.* 35Q40, 35L70.

Key words and phrases. Dirac equation; Maxwell equations; well-posedness; nonrelativistic limit.

4×4 Dirac matrices, given in 2×2 block form by

$$\gamma^0 = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}, \qquad \gamma^j = \begin{pmatrix} 0 & \sigma^j \\ -\sigma^j & 0 \end{pmatrix},$$

where $\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $\sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$ and $\sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ are the Pauli matrices. The following related matrices occur frequently:

$$\alpha^j := \gamma^0\gamma^j = \begin{pmatrix} 0 & \sigma^j \\ \sigma^j & 0 \end{pmatrix}, \qquad S^m := i\gamma^k\gamma^l = \begin{pmatrix} \sigma^m & 0 \\ 0 & \sigma^m \end{pmatrix},$$

where (k, l, m) is any cyclic permutation of $(1, 2, 3)$. Note the identities

$$\alpha^j\alpha^k = -\alpha^k\alpha^j + 2\delta^{jk} I = \delta^{jk} I + i\epsilon^{jkl} S_l. \tag{1.2}$$

The first equation in (1.1) is the Dirac equation. Multiplying it on the left by γ^0 and taking the imaginary part of its $\mathbb{C}^4$ inner product with ψ yields the conservation law $\partial_\mu J^\mu = 0$. Thus, the "charge" is conserved:

$$\int \langle \psi(t), \psi(t) \rangle_{\mathbb{C}^4} \, dx = \|\psi(t)\|_{L^2}^2 = const. \tag{1.3}$$

The second equation in (1.1) is the Maxwell equation. We split A_μ into its temporal part A_0, the electric potential, and its spatial part $\mathbf{A} = (A^1, A^2, A^3)$, the magnetic potential. Hence the electric field is given by $\mathbf{E} = \nabla A_0 - \partial_0 \mathbf{A}$ and the magnetic field by $\mathbf{B} = \nabla \times \mathbf{A}$, and the second equation in (1.1) is seen to be equivalent to the Maxwell system in classical form, with charge density $\rho = J^0/c$ and current density $\mathbf{J} = \{J^k\}_{k=1,2,3}$.

The third equation in (1.1) is the Coulomb gauge condition $\operatorname{div} \mathbf{A} = 0$. The reason for this choice of gauge is explained below. It is equivalent to $\mathbb{P}\mathbf{A} = \mathbf{A}$, where $\mathbb{P}$ is the projection onto divergence free vector fields in $\mathbb{R}^3_x$. The second and third equations in (1.1) are then equivalent to $\Delta A_0 = c^{-1} J^0$ and $\left(c^{-2}\partial_t^2 - \Delta\right)\mathbf{A} = c^{-1}\mathbb{P}\mathbf{J}$ provided the initial data of $\mathbf{A}$ are divergence free. Thus, when properly rescaled (see [1], [22]), the system (1.1) is conveniently expressed in terms of a small dimensionless parameter

$$\varepsilon \simeq \frac{1}{c}$$

as follows:

$$i\partial_t\psi^\varepsilon = -i\varepsilon^{-1}\alpha^k\partial_k\psi^\varepsilon + \varepsilon^{-2}\gamma^0\psi^\varepsilon - A_k^\varepsilon\alpha^k\psi^\varepsilon - A_0^\varepsilon\psi^\varepsilon, \tag{1.4a}$$

$$\Delta A_0^\varepsilon = \rho^\varepsilon, \tag{1.4b}$$

$$\Box_\varepsilon \mathbf{A}^\varepsilon = \varepsilon \mathbb{P}\mathbf{J}^\varepsilon, \tag{1.4c}$$

where we have put in superscripts to emphasize the dependence on ε. Here

$$\Box_\varepsilon = \varepsilon^2\partial_t^2 - \Delta$$

and

$$\rho^\varepsilon = \left\langle \psi^\varepsilon, \psi^\varepsilon \right\rangle_{\mathbb{C}^4}, \qquad \mathbf{J}^\varepsilon = \varepsilon^{-1}\left\langle \vec{\alpha}\psi^\varepsilon, \psi^\varepsilon \right\rangle_{\mathbb{C}^4}. \tag{1.4d}$$

Here $\langle \vec{\alpha}\psi, \psi \rangle$ denotes the 3-vector with components $\left\langle \alpha^k\psi, \psi \right\rangle$. We consider the Cauchy problem for (1.4) with "finite energy" initial data

$$\psi^\varepsilon|_{t=0} = \psi_0^\varepsilon \in H^1, \qquad (\mathbf{A}^\varepsilon, \partial_t\mathbf{A}^\varepsilon)|_{t=0} = (\mathbf{a}_0^\varepsilon, \mathbf{a}_1^\varepsilon) \in \mathbb{P}\dot{H}^1 \times \mathbb{P}L^2. \tag{1.5}$$

We prove three types of results for this system as $\varepsilon \to 0$. First, local well-posedness (abbreviated l.w.p.) with a logarithmic lower bound on the existence time. Second, convergence in the nonrelativistic limit if the initial datum of ψ converges. Third, we prove some more precise results on the asymptotic behavior of the Dirac spinor under various smallness assumptions on its "positron part". These results are described in detail in the next three subsections.

Our results generalize to so-called "mixed states" of quantum statistical mechanics; see Sect. 14. This formulation is indispensable when dealing with the semiclassical limit $\hbar \to 0$, since it allows for specific assumptions on the occupation probabilities as an additional initial data. The combined nonrelativistic/semiclassical limit from Dirac-Maxwell to Vlasov-Poisson is dealt with in a forthcoming paper [21].

1.1. **Local existence.**

There are two issues here: (i) l.w.p. for ε fixed, and (ii) the nature of the ε-dependence of the local existence time as $\varepsilon \to 0$.

Concerning (i), the main difficulty is that one cannot directly estimate the bilinear term $A_j \alpha^j \psi$ in the Dirac equation when ψ is at much higher frequency than $\mathbf{A}$, due to the failure of the Strichartz estimate for $\|\mathbf{A}\|_{L_t^2 L_x^\infty}$ in $1+3$ dimensions. The crucial fact proved here is that when the Dirac equation is squared, the bilinear terms resulting from this dangerous term can all be expressed in terms of null bilinear forms, provided the Coulomb gauge condition is used, and this enables us to prove l.w.p. of DM in the data space $(\psi(0), \mathbf{A}(0)) \in H^1 \times H^1$, a result entirely analogous to that of Klainerman and Machedon [13] for the Klein-Gordon-Maxwell (KGM) system. (The square of the Dirac eq. is similar to the Klein-Gordon eq., but contains some additional bilinear terms due to the presence of spin.)

L.w.p. of DM was first proved by Gross [10] for sufficiently regular data. Bournaveas [5] studied the problem at low regularity, and proved l.w.p. in $H^{1/2+\delta} \times H^{1+\delta}$ for $\delta > 0$, using linear Strichartz estimates. This was improved to $H^{1/2} \times H^1$ in [23], by partially employing the null structure of the squared Dirac equation (in fact, using the null structure of the Klein-Gordon terms, not of the additional bilinear terms due to the spin; the null structure of the latter terms are however crucial for the $H^1 \times H^1$ result obtained in the present work).

Of course, the $H^{1/2} \times H^1$ result proved in [23] implies l.w.p. in $H^1 \times H^{3/2}$, but it does not imply our $H^1 \times H^1$ result. The point is that $\mathbf{A}$ is kept at the same regularity in both cases. In fact, DM is l.w.p. in $H^s \times H^1$ for $1/4 < s \leq 1$; this is proved in a forthcoming paper by the third author. (Scaling suggests that $L^2 \times H^{1/2}$ should be the optimal result.)

The question of global existence and uniqueness for DM remains largely open[1] (but see Georgiev [9] for a small data result), however, we prove—and this brings us to the second issue mentioned above—that as $\varepsilon \to 0$ the local existence time goes to infinity at a certain rate, subject to the initial assumptions

$$\left\|\psi_0^\varepsilon\right\|_{H^1} = O(1), \qquad \left\|\mathbf{a}_0^\varepsilon\right\|_{\dot H^1} + \varepsilon \left\|\mathbf{a}_1^\varepsilon\right\|_{L^2} = O\left(\frac{1}{\varepsilon^\Lambda}\right) \quad \text{as} \quad \varepsilon \longrightarrow 0, \tag{1.6}$$

where

$$0 < \Lambda < \frac{1}{2} \tag{1.7}$$

will be kept fixed throughout the paper. (The upper bound $1/2$ is explained by the factor $\varepsilon^{1/2}$ appearing in the L^2 bilinear estimates discussed in Sect. 3; this quantifies the fact that the magnetic field is a relativistic effect.)

---

[1] This in contrast to the situation for KGM; see [13]. The crucial point is that KGM has a definite energy density, unlike DM.

Theorem 1.1. *(H^1 **l.w.p. of DM.**) The initial value problem (1.4), (1.5) is locally well posed for fixed ε, with an existence time $T_\varepsilon > 0$ depending only on ε and the size of the norms of the data. Moreover, if (1.6) holds, then*

$$T_\varepsilon \geq r \log \frac{1}{\varepsilon} \quad \text{as} \quad \varepsilon \longrightarrow 0, \tag{1.8}$$

where r is a constant depending on $\sup_{\varepsilon>0} \left\| \psi_0^\varepsilon \right\|_{L^2}$ but independent of ε. Moreover,

$$\left\| \psi^\varepsilon(t) \right\|_{H^1} = O(1), \qquad \left\| \mathbf{A}^\varepsilon(t) \right\|_{\dot{H}^1} + \varepsilon \left\| \partial_t \mathbf{A}^\varepsilon(t) \right\|_{L^2} = O\left(\frac{1}{\varepsilon^\Lambda} \right) \tag{1.9}$$

uniformly in every finite time interval as $\varepsilon \to 0$.

We remark that the $H^{1/2} \times H^1$ l.w.p. result in [23] does not give any rate of growth of the existence time as $c \to \infty$ (i.e., as $\varepsilon \to 0$), but only existence in a uniform time interval whose size depends on the size of the data norm. To be precise, the data space used for ψ in [23] is $c^{-1/2} H^{1/2} + H^1$ (which at fixed c is the same as $H^{1/2}$) and they prove existence in a uniform time interval provided that

$$\| \psi_{\text{low}}(0) \|_{H^1} + c^{1/2} \left\| \psi_{\text{high}}(0) \right\|_{H^{1/2}} = O(1)$$

and that the H^1 norm of the data for the magnetic potential is $O(1)$. Here ψ_{low} and ψ_{high} denote the low ($\leq c \simeq 1/\varepsilon$) and high ($\geq c \simeq 1/\varepsilon$) frequency parts of $\psi(t)$.

We remark that for the purpose of the combined nonrelativistic/semiclassical limit from DM to Vlasov-Poisson [21], the l.w.p. result in [23] is not strong enough, as it only gives existence on a uniform time interval but not any growth rate in c.

The key idea which allows us to get a growth rate for T_ε is already present in our earlier paper [2] on KGM, but is not used to full effect there, since for KGM local well-posedness in H^1 in any case becomes global by conservation of energy.

Our idea can be summed up in the following two steps.

(i) One expects that DM approaches a Schrödinger-Poisson system (SP) as $\varepsilon \to 0$. As is well-known, SP is l.w.p. in L^2, and this extends to a global result due to conservation of the L^2 norm. For example, this can be proved (see [2]) using the Strichartz norm

$$Z(\psi) = \| \psi \|_{L_t^\infty L_x^2} + \| \psi \|_{L_t^2 L_x^6}$$

for the Schrödinger equation in $1+3$ dimensions. Heuristically, one may expect that the same argument should apply to KGM/DM as $\varepsilon \to 0$, modulo some error converging to zero, and this is exactly what we prove; we obtain closed estimates for ψ and $\mathbf{A}$ in certain spacetime norms, and in particular the Z norm of the low frequency part ψ_{low}. We are able to fine tune the estimates so that terms not containing the Z-norm have some positive power of ε in front of them, and thus lose their influence as $\varepsilon \to 0$.

We then use a bootstrap argument to prove existence up to a time T^* *depending only on the L^2 norm of ψ_0^ε*, provided ε is sufficiently small, depending on the size of (1.6).

To see this argument in its clearest form we refer the reader to Sects. 2 and 5 in [2], where it is done for KGM. The estimates for DM, which we deal with here, are considerably more difficult, due to the new bilinear terms which appear. We prove that the extra terms have a null structure, and we prove some bilinear spacetime estimates which are needed to control these terms.

(ii) On account of the conservation of charge (1.3) for the Dirac equation, the time T^* is a constant, and we can iterate our existence argument a number $N \sim \log\left(\frac{1}{\varepsilon}\right)$ times to obtain the long time result. See Sect. 7. The point here is that for each iteration, i.e., each time we advance the time by T^*, the bound (1.6), which dictates how small ε must be taken, may grow by some fixed factor. Thus, to make N iterations, i.e., to reach a time $T \sim N$, we need $\varepsilon \le \delta C^{-N}$, where δ depends on the initial data norm, and $C > 1$ is an absolute constant. Taking the logarithm gives the relationship $T \sim N \sim \log\left(\frac{1}{\varepsilon}\right)$ as $\varepsilon \to 0$.

A key tool used in this paper is L^2 spacetime bilinear estimates, related to those first proved by Klainerman and Machedon. The estimates we use are discussed in Sect. 3, but it seems worthwhile to include a brief heuristic discussion here. For the moment, take $\varepsilon = 1$ to simplify.

A crucial problem is to control bilinear interactions between magnetic potential and spinor. Since we work with "$X^{s,b}$ spaces", it suffices (cf. Theorem 4.1 in Sect. 4) to consider interactions between a free wave $\square u = 0$ and a solution v of the free Dirac equation, which is then also a solution of the Klein-Gordon equation $\square v + v = 0$. The spacetime Fourier transforms $\widehat{u}(\tau, \xi)$ and $\widehat{v}(\tau, \xi)$ are measures supported, respectively, on the light cone $\tau = \pm|\xi|$ and on the hyperboloid (of two sheets) $\tau = \pm\left(1 + |\xi|^2\right)^{1/2}$. The key observation is that for large frequencies, the hyperboloid looks like a cone, and so the interaction between u and v is essentially like that between two free waves, which was the case considered by Klainerman and Machdeon. On the other hand, for ξ in a neighborhood of the origin, the hyperboloid it has more curvature than a cone, which provides additional smoothing.

Let us be a bit more precise. In fact, for the purpose of the nonrelativistic limit $\varepsilon \to 0$, it is crucial to (i) split $v = v_+ + v_-$ into positive and negative energy parts (in Fourier space this corresponds to restricting to the upper and lower sheets of the hyperboloid) and (ii) to subtract the rest energy, which corresponds to a translation in Fourier space, to the surfaces $\tau = \pm\left[\left(1 + |\xi|^2\right)^{1/2} - 1\right]$. The expression inside the square brackets is $\sim |\xi|^2$ for $|\xi| \lesssim 1$ (giving a paraboloid-like surface) and $\sim |\xi|$ for $|\xi| \gtrsim 1$ (a cone-like surface), suggesting that $v_\pm$ should behave like solutions of the free Schrödinger equation at low frequency, and like solutions of the free wave equation at high frequency. These points are discussed in more detail in later sections.

We stress the fact that no convergence assumption was made on the data in Theorem 1.1—all we need is the uniform bound (1.6). However, if we *do* assume that ψ_0^ε converges in H^1, then we can pass to the nonrelativistic limit, which we discuss next.

1.2. **Nonrelativistic limit.**

The nonrelativistic limit of the linear Dirac equation with a given time-dependent electromagnetic potential was treated in [1] (earlier papers, see e.g. [6], dealt only with the static case, i.e. time-independent potential). There are also some results on the nonlinear Dirac and Klein-Gordon equations in the literature, see e.g. [24] and recently [19], but for the coupled nonlinear KGM and DM systems there are no results previous to our work (i.e. the present paper as well as [2, 3]) and the independent and parallel work of Masmoudi and Nakanishi [23].

Let us now state our result. We split the Dirac spinor into its upper and lower components:

$$\psi^\varepsilon = \begin{pmatrix} \tilde{\chi}^\varepsilon \\ \tilde{\eta}^\varepsilon \end{pmatrix} \tag{1.10}$$

where $\tilde{\chi}$ and $\tilde{\eta}$ are 2-spinors, i.e. column vectors in $\mathbb{C}^2$. Before one can pass to the limit $\varepsilon \to 0$, the rest energy must be subtracted, which for the upper "positive energy" component means multiplication by e^{it/ε^2} and for the lower "negative energy" component multiplication by e^{-it/ε^2}.

Theorem 1.2. ***(Nonrelativistic limit of DM.)*** *Consider the solution of (1.4), (1.5) obtained in Theorem 1.1, with data satisfying:*

(i) $v_0 := \lim_{\varepsilon\to 0} \psi_0^\varepsilon$ *exists in* H^1,

(ii) $\left\|\mathbf{a}_0^\varepsilon\right\|_{\dot{H}^1} + \varepsilon \left\|\mathbf{a}_1^\varepsilon\right\|_{L^2} = O\left(\frac{1}{\varepsilon^\Lambda}\right)$ *as* $\varepsilon \to 0$.

Denote the upper and lower 2-spinors of v_0 *by* v_0^+ *and* v_0^- *respectively, and let* (u, v_+, v_-) *be the solution of the Schrödinger-Poisson system*[2]

$$\Delta u = n, \qquad n = |v_+|^2 + |v_-|^2, \qquad \left(i\partial_t \pm \frac{\Delta}{2}\right) v_\pm + u v_\pm = 0, \tag{1.11}$$

with initial data $v_\pm|_{t=0} = v_0^\pm$. *Then as* $\varepsilon \to 0$,

$$\psi^\varepsilon = e^{-it/\varepsilon^2}\begin{pmatrix} v_+ \\ 0 \end{pmatrix} + e^{+it/\varepsilon^2}\begin{pmatrix} 0 \\ v_- \end{pmatrix} + o(1) \qquad \text{in} \quad H^1, \tag{1.12a}$$

$$A_0^\varepsilon = u + o(1) \qquad \text{in} \quad \dot{H}^1, \tag{1.12b}$$

$$\rho^\varepsilon = n + o(1) \qquad \text{in} \quad L^p, \ 1 \le p \le 3, \tag{1.12c}$$

uniformly in every finite time interval. Moreover, the relativistic current density converges as follows: Let

$$\begin{aligned} \mathbf{J}^0 &= \operatorname{Im}\langle \nabla v_+, v_+ \rangle_{\mathbb{C}^2} - \operatorname{Im}\langle \nabla v_-, v_- \rangle_{\mathbb{C}^2} \\ &\quad + \frac{1}{2}\nabla \times \langle \vec{\sigma} v_+, v_+ \rangle_{\mathbb{C}^2} - \frac{1}{2}\nabla \times \langle \vec{\sigma} v_-, v_- \rangle_{\mathbb{C}^2} \end{aligned} \tag{1.13}$$

where $\langle \nabla v_\pm, v_\pm \rangle$ *and* $\langle \vec{\sigma} v_\pm, v_\pm \rangle$ *are the vectors with components* $\left\langle \partial^j v_\pm, v_\pm \right\rangle$ *and* $\left\langle \sigma^j v_\pm, v_\pm \right\rangle$, *respectively, for* $j = 1, 2, 3$. *Then*

$$\mathbf{J}^\varepsilon \longrightarrow \mathbf{J}^0 \quad \text{in} \quad \left[C_c^1(\mathbb{R}_t \times \mathbb{R}_x^3)\right]' \quad \text{weak} \ast \tag{1.14}$$

as $\varepsilon \to 0$. *(The first line in the right hand side of* (1.13) *is the conserved current associated to the limiting system* (1.11), *whereas the second line consists of the well-known divergence-free additional terms due to the interaction spin-magnetic field; see* [18].*)*

We remark that the corresponding result where in the assumption (i) and the conclusion (1.12a) one replaces the space H^1 by $c^{-1/2}H^{1/2} + H^1$, was proved in [23]. While the space $c^{-1/2}H^{1/2} + H^1$ is larger than H^1, it also has a smaller norm, so although in [23] the initial convergence assumption is less strong, then so is the conclusion.

We can improve the convergence rate to $O(\varepsilon)$ by strengthening the initial assumptions. Thus, we shall prove:

Theorem 1.3. *Strengthen the hypotheses of Theorem 1.2 by assuming*

$$\left\|\psi_0^\varepsilon\right\|_{H^2} = O(1), \quad \left\|\nabla \mathbf{a}_0^\varepsilon\right\|_{H^1} + \varepsilon \left\|\mathbf{a}_1^\varepsilon\right\|_{H^1} = O\left(\frac{1}{\varepsilon^\Lambda}\right),$$

$$\psi_0^\varepsilon = \begin{pmatrix} v_0^+ \\ 0 \end{pmatrix} + O(\varepsilon) \qquad \text{in} \quad H^1$$

---

[2] This system is globally well posed for L^2 data.

as $\varepsilon \longrightarrow 0$*. Moreover, assume* $v_0^+ \in H^5$*. Then*

$$\psi^\varepsilon = e^{-it/\varepsilon^2}\begin{pmatrix} v_+ \\ 0 \end{pmatrix} + O(\varepsilon) \qquad \text{in} \quad H^1 \quad \text{as} \quad \varepsilon \longrightarrow 0 \tag{1.15}$$

uniformly in every finite time interval. Furthermore, the convergence in (1.12b) and (1.12c) is also $O(\varepsilon)$*.*

Remark 1.4. The hypotheses are not strong enough to guarantee strong convergence of the current density $\mathbf{J}^\varepsilon$ locally uniformly in time. In fact, a simple counterexample is given by the initial datum $\psi_0^\varepsilon = [v_0^+, \varepsilon v_0^+]^T$. Then $\mathbf{J}^\varepsilon$ initially has vector components $2\,\mathrm{Re}\left\langle \sigma^j v_0^+, v_0^+ \right\rangle$, which does not agree with the weak limit $\mathbf{J}^0$ given by (1.13).

It is instructive to compare the last theorem to the formal derivation of the nonrelativistic limit usually reproduced in physics textbooks, the basic premise of which is a smallness assumption on the lower component $\tilde{\eta}^\varepsilon$ of the spinor. The idea is to define[3]

$$\phi^\varepsilon = \begin{pmatrix} \chi^\varepsilon \\ \eta^\varepsilon \end{pmatrix} := e^{it/\varepsilon^2}\psi^\varepsilon. \tag{1.16}$$

Then (1.15) can be restated

$$\chi^\varepsilon = v_+ + O(\varepsilon), \qquad \eta^\varepsilon = O(\varepsilon) \qquad \text{in} \quad H^1 \quad \text{as} \quad \varepsilon \longrightarrow 0. \tag{1.17}$$

The Dirac equation (1.4a) gives

$$iD_0\chi^\varepsilon = -i\sigma^j D_j \eta^\varepsilon, \qquad iD_0\eta^\varepsilon = -i\sigma^j D_j \chi^\varepsilon - \frac{2}{\varepsilon}\eta^\varepsilon, \tag{1.18}$$

where we write $D_0 = \varepsilon\partial_t - i\varepsilon A_0^\varepsilon$ and $D_j = \partial_j - i\varepsilon A_j^\varepsilon$. Thus,

$$\eta^\varepsilon = -\varepsilon\frac{1}{2}i\sigma^j\partial_j\chi^\varepsilon - \varepsilon^2\frac{1}{2}\left\{i\partial_t\eta^\varepsilon + A_0^\varepsilon\eta^\varepsilon + A_j^\varepsilon\sigma^j\chi^\varepsilon\right\}, \tag{1.19}$$

and substituting this in the first equation in (1.18) gives, after some algebra,

$$i\partial_t\chi^\varepsilon = \frac{1}{2}\left(i\nabla + \varepsilon\mathbf{A}^\varepsilon\right)^2\chi^\varepsilon - A_0^\varepsilon\chi^\varepsilon - \frac{1}{2}\varepsilon B_j^\varepsilon\sigma^j\chi^\varepsilon - \varepsilon r^\varepsilon$$

where $r^\varepsilon = \frac{1}{2}\sigma^j D_j\left(\partial_t\eta^\varepsilon - iA_0^\varepsilon\eta^\varepsilon\right)$ and $\mathbf{B}^\varepsilon = \nabla \times \mathbf{A}^\varepsilon$. Then by formal considerations of magnitude, in particular assuming $\partial_t\eta^\varepsilon = O(1)$, one obtains a Schrödinger equation in the limit $\varepsilon \to 0$. It is possible to make this argument rigorous, but it has a fundamental weakness which limits its usefulness, namely that $\partial_t\eta^\varepsilon$ can be no better than $O(1/\varepsilon)$ unless one adds a further constraint on the initial data. In fact, it is clear from (1.19) that $\partial_t\eta^\varepsilon = O(1)$ in L^2 initially if and only if the constraint

$$\eta^\varepsilon = -\varepsilon\frac{1}{2}i\sigma^j\partial_j\chi^\varepsilon + O(\varepsilon^2) \tag{1.20}$$

holds in L^2 at time $t = 0$, assuming the data (1.5) are $O(1)$.

However, the constraint (1.20) is not needed in Theorem 1.3, the reason being that instead of the simple splitting into upper and lower components as in (1.10), we apply the eigenspace projections of the "Dirac operator" $\mathcal{Q}(D) = -i\alpha^k\partial_k + \gamma^0$ whose Fourier

---

[3] Here we break the symmetry of the signs in (1.12a), i.e. between "electrons" and "positrons", but this is not important since the lower component is in any case expected to vanish.

symbol is a 4×4 matrix with eigenvalues $\pm\lambda(\xi)$, where $\lambda(\xi) = \sqrt{1 + |\xi|^2}$. As in [1] we use the spectral decomposition $\mathcal{Q}(\varepsilon D) = \lambda(\varepsilon D)\Pi_+(\varepsilon D) - \lambda(\varepsilon D)\Pi_-(\varepsilon D)$, where

$$\lambda(\varepsilon D) = \sqrt{1 - \varepsilon^2 \Delta}, \qquad \Pi_\pm(\varepsilon D) = \frac{1}{2}\left(I \pm [\lambda(\varepsilon D)]^{-1}\mathcal{Q}(\varepsilon D)\right). \tag{1.21}$$

Since the positive and negative eigenvalues $\pm\lambda(\varepsilon D)$ correspond to positive and negative energies of a free Dirac particle, the spectral decomposition is related to electrons and positrons ([7]). The formal limit $\varepsilon \to 0$ of $\Pi_\pm(\varepsilon D)$ yields the operators

$$\Pi^0_\pm = \frac{1}{2}(I \pm \gamma^0), \qquad \Pi^0_+ = \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}, \qquad \Pi^0_- = \begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix}.$$

The following basic lemma shows that $\Pi^0_\pm$ is the leading order term in a series expansion of $\Pi_\pm(\varepsilon D)$ in powers of ε, and moreover that (1.20) is basically equivalent to $\Pi_-(\varepsilon D)\psi^\varepsilon = O(\varepsilon^2)$, a condition which resurfaces in the next subsection.

Lemma 1.5. *For all $s \in \mathbb{R}$, $\Pi_\pm(\varepsilon D)$ is bounded from $H^s \to H^s$ uniformly in ε. Moreover,*

$$\Pi_\pm(\varepsilon D) = \Pi^0_\pm \mp \varepsilon \mathcal{R}^\varepsilon_1 \tag{1.22}$$

$$= \Pi^0_\pm \mp i\varepsilon \frac{1}{2}\alpha^k \partial_k \mp \varepsilon^2 \mathcal{R}^\varepsilon_2 \tag{1.23}$$

where $\mathcal{R}^\varepsilon_j$ denotes an operator bounded from $H^s \to H^{s-j}$ uniformly in ε.

Proof. This follows immediately from

$$\Pi_\pm(\varepsilon D) - \Pi^0_\pm = \mp\frac{1}{2}[\lambda(\varepsilon D)]^{-1} i\varepsilon\alpha^k\partial_k \mp \frac{1}{2}\left(1 - [\lambda(\varepsilon D)]^{-1}\right)\gamma^0 \tag{1.24}$$

and the fact that the Fourier symbol of $1 - [\lambda(\varepsilon D)]^{-1}$ satisfies the inequalities

$$0 \le 1 - \frac{1}{\sqrt{1 + \varepsilon^2 |\xi|^2}} \le \min\left\{1, \varepsilon|\xi|, \varepsilon^2 |\xi|^2\right\} \tag{1.25}$$

where ξ is the Fourier variable corresponding to x. □

Before moving on, we prove that the initial data assumption (i) in Theorem 1.2 implies something stronger, namely the convergence of $\Pi_\pm(\varepsilon D)\psi^\varepsilon_0$.

Lemma 1.6. *If $\lim_{\varepsilon\to 0} \psi^\varepsilon_0 = v_0 = \begin{pmatrix} v^+_0 \\ v^-_0 \end{pmatrix}$ exists in H^1, then*

$$\lim_{\varepsilon\to 0} \Pi_+(\varepsilon D)\psi^\varepsilon_0 = \begin{pmatrix} v^+_0 \\ 0 \end{pmatrix} \quad \textit{and} \quad \lim_{\varepsilon\to 0} \Pi_-(\varepsilon D)\psi^\varepsilon_0 = \begin{pmatrix} 0 \\ v^-_0 \end{pmatrix}$$

in H^1.

Proof. It suffices to prove $\left(\Pi_\pm(\varepsilon D) - \Pi^0_\pm\right)\psi^\varepsilon_0 \to 0$ in H^1. But the proof of Lemma 1.5 shows that the Fourier symbol of $\Pi_\pm(\varepsilon D) - \Pi^0_\pm$ is bounded in absolute value by $\min\{1, \varepsilon|\xi|\}$. Thus $\left(\Pi_\pm(\varepsilon D) - \Pi^0_\pm\right)(\psi^\varepsilon_0 - v_0) \to 0$ in H^1, and by Plancherel's theorem and dominated convergence, $\left(\Pi_\pm(\varepsilon D) - \Pi^0_\pm\right)v_0 \to 0$ in H^1. □

1.3. **Semi-nonrelativistic limit.** As in [1], by the "semi-nonrelativistic limit" we understand the approximation of the upper component of the Dirac equation by the Pauli equation for a 2-spinor, which reads

$$i\partial_t \chi_P^\varepsilon = \frac{1}{2}\left(i\nabla + \varepsilon \mathbf{A}^\varepsilon\right)^2 \chi_P^\varepsilon - A_0^\varepsilon \chi_P^\varepsilon - \frac{1}{2}\varepsilon B_j^\varepsilon \sigma^j \chi_P^\varepsilon \tag{1.26a}$$

with initial condition

$$\chi_P^\varepsilon|_{t=0} = \chi_{P0}^\varepsilon \in H^1. \tag{1.26b}$$

Note that the naive "upper and lower components" approach in (1.18)–(1.19) can give at best an $O(\varepsilon)$ approximation to the Pauli equation, assuming the initial constraint (1.20), which as remarked is essentially equivalent to $\Pi_-(\varepsilon D)\psi^\varepsilon = O(\varepsilon^2)$.

In contrast, by using the Dirac projections $\Pi_\pm(\varepsilon D)$ instead of just $\Pi_\pm^0$, we can prove an $O(\varepsilon^2)$ approximation, with the same initial constraint. In fact, we have the following result:

Theorem 1.7. *Consider the solution of (1.4), (1.5) obtained in Theorem 1.1. Define* χ^ε *as in (1.16) and let* χ_P^ε *be the solution of the Pauli equation (1.26). Assume the initial conditions*

(i) $\left\|\psi_0^\varepsilon\right\|_{H^5} = O(1), \quad \left\|\nabla \mathbf{a}_0^\varepsilon\right\|_{H^4} + \varepsilon \left\|\mathbf{a}_1^\varepsilon\right\|_{H^4} = O(1),$

(ii) $\left\|\Pi_-(\varepsilon D)\psi_0^\varepsilon\right\|_{H^2} = O(\varepsilon),$

(iii) $\left\|\Pi_-(\varepsilon D)\psi_0^\varepsilon\right\|_{H^1} = O(\varepsilon^2),$

as $\varepsilon \to 0$. *Then if*

$$\left\|\chi^\varepsilon - \chi_P^\varepsilon\right\|_{H^1} = O(\varepsilon^2) \tag{1.27}$$

holds at time $t = 0$, *it also holds uniformly in every finite time interval. For the current density we then have*

$$\mathbf{J}^\varepsilon = \mathbf{J}_P^\varepsilon + \frac{1}{2}\nabla \times \left\langle \vec{\sigma}\chi_P^\varepsilon, \chi_P^\varepsilon \right\rangle_{\mathbb{C}^2} + O(\varepsilon) \quad \text{in} \quad L_x^1, \tag{1.28}$$

uniformly in every finite time interval, where $\mathbf{J}_P^\varepsilon = \operatorname{Im}\left\langle (\nabla - i\varepsilon\mathbf{A}^\varepsilon)\chi_P^\varepsilon, \chi_P^\varepsilon \right\rangle_{\mathbb{C}^2}$ *is the current density associated to the Pauli equation.*

The remainder of this paper is organized as follows: In the following section we square the Dirac equation and reinterpret it in terms of the projections $\Pi_\pm(\varepsilon D)\psi^\varepsilon$ of the spinor, and we prove that the main bilinear terms can be expressed in terms of null forms. Then in Sect. 3 we discuss the linear and bilinear spacetime estimates of Strichartz type that are used in this paper. The proofs of those estimates that are not already in the literature can be found in Sect. 12. In Sect. 4 we define the function spaces that we use, and recall their main properties. The main estimates for the nonlinear terms are proved in Sect. 5, which is the heart of the paper. Then in Sects. 7–11 these estimates are applied to prove the main theorems.

To close this section we introduce some notational conventions which will be in effect throughout:

- For function spaces we use the following notation. If X is a Banach space of functions on $\mathbb{R}_x^3$, we denote by $L_t^p X$ the space with norm

$$\|u\|_{L_t^p X} = \left(\int_{-\infty}^{\infty} \|u(t,\cdot)\|_X^p \, dt\right)^{1/p},$$

  with the usual modification if $p = \infty$. The localization of this norm to a time slab $S_T = (0,T) \times \mathbb{R}^3$ is denoted $\|u\|_{L_t^p X(S_T)}$.

- In estimates, we use the notation $\lesssim$ to mean $\le$ up to multiplication by a positive constant C independent of ε. Moreover, in estimates over a time slab S_T, C is also understood to be independent of T.
- For exponents, we use the convenient shorthand p^+ (resp. p^-) for $p+\zeta$ (resp. $p-\zeta$) with $\zeta > 0$ sufficiently small, independently of ε. The notation ∞^- stands for a sufficiently large, positive exponent.
- We denote by $f(x) \mapsto \widehat{f}(\xi)$ and $u(t,x) \mapsto \widehat{u}(\tau,\xi)$ the Fourier transforms on $\mathbb{R}^3$ and $\mathbb{R}^{1+3}$, respectively. As in [2] we split functions $f(x)$ into their low ($|\xi| \lesssim 1/\varepsilon$) and high ($|\xi| \gtrsim 1/\varepsilon$) frequency parts,
$$f = f_{\text{low}} + f_{\text{high}}, \tag{1.29}$$
corresponding to a smooth partition of unity in Fourier space.

2. PRELIMINARIES

As already mentioned, our approach to the Dirac equation is to square it and apply techniques similar to those used for KGM in [2]. It is therefore convenient to work with the "KG splitting"

$$\psi_\pm^\varepsilon = \begin{pmatrix} \tilde{\chi}_\pm^\varepsilon \\ \tilde{\eta}_\pm^\varepsilon \end{pmatrix} := \frac{1}{2}\left\{ \psi^\varepsilon \pm \varepsilon^2 [\lambda(\varepsilon D)]^{-1} \bigl(i\partial_t \psi^\varepsilon + A_0^\varepsilon \psi^\varepsilon \bigr) \right\} \tag{2.1}$$

as used in [2]. In order to compare this to the Dirac projections (1.21), observe that if ψ^ε solves the Dirac equation (1.4a), then

$$\psi_\pm^\varepsilon = \Pi_\pm(\varepsilon D)\psi^\varepsilon \mp \frac{1}{2}\varepsilon^2 [\lambda(\varepsilon D)]^{-1} (A_k^\varepsilon \alpha^k \psi^\varepsilon). \tag{2.2}$$

But using the estimate

$$\left\| [\lambda(\varepsilon D)]^{-1} f \right\|_{H^\sigma} \le \varepsilon^{-r} \left\| f \right\|_{H^{\sigma-r}} \quad \text{for} \quad 0 \le r \le 1, \tag{2.3}$$

followed by Hölder's inequality and Sobolev embedding, we see that

$$\varepsilon^2 \left\| [\lambda(\varepsilon D)]^{-1} (A_k^\varepsilon \alpha^k \psi^\varepsilon) \right\|_{H^\sigma} \lesssim \varepsilon^{2-\sigma} \left\| \mathbf{A}^\varepsilon \right\|_{\dot{H}^1} \left\| \psi^\varepsilon \right\|_{H^1} \quad \text{for} \quad 0 \le \sigma \le 1, \tag{2.4}$$

so the right hand side of (2.2) is $O(\varepsilon^{1-\Lambda})$ in H^1 at time t if the bound (1.9) in Theorem 1.1 holds. As far as proving Theorem 1.2 is concerned, it is therefore immaterial whether we use $\psi_\pm^\varepsilon$ or $\Pi_\pm(\varepsilon D)\psi^\varepsilon$.

For later use we note the following consequences of (2.2) and (2.4). First,

$$\left\| \psi_\pm^\varepsilon \right\|_{H^\sigma} \lesssim \left\| \psi^\varepsilon \right\|_{H^\sigma} + \varepsilon^{2-\sigma} \left\| \mathbf{A}^\varepsilon \right\|_{\dot{H}^1} \left\| \psi^\varepsilon \right\|_{H^1} \quad \text{for} \quad 0 \le \sigma \le 1, \tag{2.5}$$

using the uniform boundedness of $\Pi_\pm(\varepsilon D)$. Second,

$$\left\| \tilde{\chi}_-^\varepsilon \right\|_{L_x^2} + \left\| \tilde{\eta}_+^\varepsilon \right\|_{L_x^2} \lesssim \varepsilon \left\| \psi^\varepsilon \right\|_{H^1} + \varepsilon^2 \left\| \mathbf{A}^\varepsilon \right\|_{\dot{H}^1} \left\| \psi^\varepsilon \right\|_{H^1}, \tag{2.6}$$

where we used (1.22) and the orthogonality of Π_+^0 and Π_-^0.

Let us now restate the system (1.4) in terms of the splitting (2.1) of the spinor. First we subtract the rest energy, defining

$$\phi_\pm^\varepsilon = \begin{pmatrix} \chi_\pm^\varepsilon \\ \eta_\pm^\varepsilon \end{pmatrix} := e^{\pm it/\varepsilon^2} \psi_\pm^\varepsilon. \tag{2.7}$$

Thus

$$\psi^\varepsilon = \psi_+^\varepsilon + \psi_-^\varepsilon = e^{-it/\varepsilon^2} \phi_+^\varepsilon + e^{+it/\varepsilon^2} \phi_-^\varepsilon. \tag{2.8}$$

Lemma 2.1. *In terms of the splitting* (2.8)*, defined via* (2.1) *and* (2.7)*, the Dirac equation (1.4a) is equivalent to a system of two equations*

$$L_+^\varepsilon \phi_+^\varepsilon = -A_0^\varepsilon \phi_+^\varepsilon + \frac{1}{2} e^{it/\varepsilon^2} R^\varepsilon, \quad L_-^\varepsilon \phi_-^\varepsilon = -A_0^\varepsilon \phi_-^\varepsilon - \frac{1}{2} e^{-it/\varepsilon^2} R^\varepsilon, \tag{2.9}$$

provided the constraint (2.2) *is satisfied at time* $t = 0$*, or equivalently that the Dirac equation is satisfied at* $t = 0$*. Here*

$$L_\pm^\varepsilon = i\partial_t \mp \frac{\lambda(\varepsilon D) - 1}{\varepsilon^2} \tag{2.10}$$

and R^ε *is given by*[4]

$$\begin{aligned}\lambda(\varepsilon D) R^\varepsilon = \varepsilon \left\{ 2i\mathbf{A}^\varepsilon \cdot \nabla + i \operatorname{div} \mathbf{A}^\varepsilon + i E_k^\varepsilon \alpha^k - B_k^\varepsilon S^k \right\} \psi^\varepsilon \\ + \varepsilon^2 (\mathbf{A}^\varepsilon)^2 \psi^\varepsilon - [A_0^\varepsilon, \lambda(\varepsilon D)](\psi_+^\varepsilon - \psi_-^\varepsilon).\end{aligned} \tag{2.11}$$

Further, $[\cdot,\cdot]$ *denotes the commutator and*

$$\mathbf{E}^\varepsilon = (E_1^\varepsilon, E_2^\varepsilon, E_3^\varepsilon) := \nabla A_0^\varepsilon - \varepsilon \partial_t \mathbf{A}^\varepsilon, \qquad \mathbf{B}^\varepsilon = (B_1^\varepsilon, B_2^\varepsilon, B_3^\varepsilon) := \nabla \times \mathbf{A}^\varepsilon. \tag{2.12}$$

Remark 2.2. An obvious, but rather important fact is that in the commutator we may replace $\lambda(\varepsilon D)$ by the better behaved $\lambda(\varepsilon D) - 1$.

Proof. Squaring the Dirac equation (1.4a) yields (cf. [7, Sect. 70])

$$\left\{ \varepsilon^2 (i\partial_t + A_0^\varepsilon)^2 + (\nabla - i\varepsilon \mathbf{A}^\varepsilon)^2 - \varepsilon^{-2} - i\varepsilon E_j^\varepsilon \alpha^j + \varepsilon B_j^\varepsilon S^j \right\} \psi^\varepsilon = 0. \tag{2.13}$$

Applying $i\partial_t + A_0^\varepsilon$ to both sides of (2.1) and using (2.13) and $\varepsilon^2 (i\partial_t + A_0^\varepsilon)\psi^\varepsilon = \lambda(\varepsilon D)(\psi_+^\varepsilon - \psi_-^\varepsilon)$ (this follows from (2.1)), one easily obtains (2.9). Reversing these steps, one finds that (2.9) implies the squared Dirac equation (2.13). But the latter implies the Dirac equation, since we assume that (2.2) holds initially, which amounts to saying that the Dirac equation is satisfied initially. ☐

Let us make a brief, heuristic comparison of (2.9) with the expected limit (1.11). As it turns out, R^ε vanishes in the limit, so (2.9) tends to the Schrödinger equation in (1.11). In fact, the Fourier symbol of $(\lambda(\varepsilon D) - 1)/\varepsilon^2$ is

$$h_\varepsilon(\xi) := \frac{|\xi|^2}{1 + \sqrt{1 + \varepsilon^2 |\xi|^2}} \sim \begin{cases} |\xi|^2/2 & \text{for } |\xi| \lesssim 1/\varepsilon, \\ |\xi|/\varepsilon & \text{for } |\xi| \gtrsim 1/\varepsilon, \end{cases} \tag{2.14}$$

so $L_\pm$ tends to the Schrödinger operator $i\partial_t \pm \Delta/2$ as $\varepsilon \to 0$. Moreover, the charge and current densities (1.4d) are given in terms of the fields (2.7) by

$$\begin{aligned}\rho^\varepsilon = \langle \chi_+^\varepsilon, \chi_+^\varepsilon \rangle + \langle \chi_-^\varepsilon, \chi_-^\varepsilon \rangle + \langle \eta_+^\varepsilon, \eta_+^\varepsilon \rangle + \langle \eta_-^\varepsilon, \eta_-^\varepsilon \rangle \\ + 2 \operatorname{Re} \left\{ e^{-2it/\varepsilon^2} \langle \chi_+^\varepsilon, \chi_-^\varepsilon \rangle + e^{-2it/\varepsilon^2} \langle \eta_+^\varepsilon, \eta_-^\varepsilon \rangle \right\},\end{aligned} \tag{2.15}$$

$$\begin{aligned}\mathbf{J}^\varepsilon = \frac{2}{\varepsilon} \operatorname{Re} \Big\{ \left\langle \sigma^j \chi_+^\varepsilon, \eta_+^\varepsilon \right\rangle + \left\langle \sigma^j \chi_-^\varepsilon, \eta_-^\varepsilon \right\rangle \\ + e^{2it/\varepsilon^2} \left\langle \sigma^j \chi_-^\varepsilon, \eta_+^\varepsilon \right\rangle + e^{-2it/\varepsilon^2} \left\langle \sigma^j \chi_+^\varepsilon, \eta_-^\varepsilon \right\rangle \Big\}_{j=1,2,3}.\end{aligned} \tag{2.16}$$

We expect [cf. (2.6)] that $\chi_-^\varepsilon, \eta_+^\varepsilon \to 0$. Thus, in the right hand side of (2.15) only the first and fourth terms are of importance, and $\Delta A_0^\varepsilon = \rho^\varepsilon$ tends to the Poisson equation in (1.11).

---

[4] Note that the term $\operatorname{div} \mathbf{A}^\varepsilon$ vanishes if $(\psi^\varepsilon, A_\mu^\varepsilon)$ is a solution of the full system (1.4), (1.5).

For later use we note the estimate

$$0 \le |\xi|/\varepsilon - h_\varepsilon(\xi) \lesssim \varepsilon^{-2}. \tag{2.17}$$

This reduces to $r - \alpha(r) \lesssim 1$, where

$$\alpha(r) := \frac{r^2}{1+\sqrt{1+r^2}}. \tag{2.18}$$

But $r - \alpha(r) = r + 1 - s = 1 - \frac{1}{r+s}$, where $s = \sqrt{1+r^2}$.

We now turn to the problem of obtaining closed estimates for the modified DM system (2.9), (1.4b), (1.4c). A serious obstacle to estimating the bilinear terms in (2.11) is the failure of the endpoint Strichartz estimate for the wave equation in dimension $1+3$. The salient feature of the Coulomb gauge, however, is that these problematic terms can be expressed in terms of the null bilinear forms

$$Q_0(u,v) = \partial_0 u \partial_0 v - \nabla u \cdot \nabla v, \qquad Q_{\alpha\beta}(u,v) = \partial_\alpha u \partial_\beta v - \partial_\beta u \partial_\alpha v, \tag{2.19}$$

where ∂_0 denotes $\varepsilon\partial_t$ and $0 \le \alpha < \beta \le 3$. These bilinear forms enjoy better regularity properties than generic products of derivatives.

We emphasize that in the following result ψ does not have to solve the Dirac equation.

Lemma 2.3. ***(Null structure.)*** *Given a potential $\{A_\mu(t,x)\}$ satisfying the Coulomb condition* $\operatorname{div}\mathbf{A} = 0$, *define $\mathbf{E} = \nabla A_0 - \varepsilon\partial_t \mathbf{A}$ and $\mathbf{B} = \nabla \times \mathbf{A}$, and consider the linear operator $\psi \to \left\{2i\mathbf{A}\cdot\nabla + iE_j\alpha^j - B_jS^j\right\}\psi$ appearing in (2.11). We have the following identities:*

$$2\mathbf{A}\cdot\nabla\psi = -Q_{jk}(D^{-1}a^{jk},\psi) \tag{2.20}$$

and

$$\begin{aligned}&\left\{i(E_j - \partial_j A_0)\alpha^j - B_jS^j\right\}\psi \\ &\quad = Q_{jk}(D^{-1}\varepsilon\partial_t a^{jk}, U) - Q_{jk}(D^{-1}\partial_l a^{jk}, \alpha^l U) \\ &\quad\quad + Q_0(A_j, \alpha^j U) + Q_{0j}(A_k, \alpha^j\alpha^k U) - \frac{i}{2}Q_{jk}(A_m, \epsilon^{jkl}S_l\alpha^m U)\end{aligned} \tag{2.21}$$

where $a_{jk} = R_jA_k - R_kA_j$, $R_j = D^{-1}\partial_j$ and $U = U(\psi)$ is given by

$$\Box_\varepsilon U = -i\left(\varepsilon\partial_t + \alpha^j\partial_j\right)\psi, \qquad U|_{t=0} = 0, \qquad i\varepsilon\partial_t U|_{t=0} = \psi_0. \tag{2.22}$$

Here ψ_0 denotes $\psi|_{t=0}$.

Proof. (2.20) goes back to the work of Klainerman and Machedon [13] on KGM, so we concentrate on the new identity (2.21). Define $\partial_\pm = \varepsilon\partial_t \pm \alpha^j\partial_j$ and observe that

$$(\partial_- A_j)\alpha^j = -(E_j - \partial_j A_0)\alpha^j - iB_jS^j, \tag{2.23}$$

where we used the second identity in (1.2) and the assumption $\operatorname{div}\mathbf{A} = 0$. By the first identity in (1.2), $\partial_+\partial_- = \Box_\varepsilon$. Thus (2.22) implies that $w = \psi - i\partial_- U$ satisfies $\partial_+ w = 0$ with $w|_{t=0} = 0$, whence $\psi = i\partial_- U$. Apply (2.23) to this and use $\partial_+(\alpha^j U) = \alpha^j\partial_- U + 2\partial^j U$ [by (1.2)] to rewrite the left hand side of (2.21) as

$$(\partial_- A_j)\partial_+(\alpha^j U) - 2(\partial_- A_j)\partial^j U.$$

To the last term we apply the identity (2.20); this we can do since $\operatorname{div} \partial_\mu \mathbf{A} = 0$ for $\mu = 0, 1, 2, 3$ by assumption. To the first term we apply the following general formula, obtained using the second identity in (1.2),

$$(\partial_-\phi)(\partial_+U) = Q_0(\phi, U) + Q_{0j}(\phi, \alpha^j U) - \frac{i}{2} Q_{jk}(\phi, \epsilon^{jkl} S_l U),$$

where ϕ is a function and U a 4-spinor. This last formula is due to Klainerman and Machedon [15]; we remark that it was used in [5] to reveal a null structure in the equation for the magnetic potential. □

3. BILINEAR SPACETIME ESTIMATES

The main technical tools used in this paper are spacetime estimates of Strichartz type for solutions of the free initial value problems

$$\begin{aligned} &\square_\varepsilon u = 0, && u|_{t=0} = f, && \partial_t u|_{t=0} = 0, \\ &L^\varepsilon_\pm v = 0, && v|_{t=0} = g, && \end{aligned} \tag{3.1}$$

on $\mathbb{R}^{1+3}$. Let us first describe the new L^2 product estimates that are proved in this paper, and then we recall the estimates proved in [2]. All these estimates are analogues of bilinear spacetime estimates for two solutions of the free wave equation, which were first investigated by Klainerman and Machedon. A systematic discussion of such estimates can be found in the article of Foschi and Klainerman [8]. In our case, one of the free waves is replaced by a solution of $L^\varepsilon_\pm v = 0$. As discussed in the Introduction, v behaves like a free wave in the high frequency range, and we are then able to modify the usual proofs of the bilinear estimates, as given in [8]. In the low frequency range, on the other hand, there is additional smoothing, and the proof is less delicate; in particular, the null condition plays no role at low frequency.

Let μ and λ be dyadic numbers of the form 2^j, $j \in \mathbb{Z}$. Denote by Δ_μ the Littlewood-Paley operator given by $\big(\Delta_\mu f\big)\widehat{\ }(\xi) = \beta(\xi/\mu)\widehat{f}(\xi)$, where β is a bump function supported in $|\xi| \sim 1$ such that $\sum_{j\in\mathbb{Z}} \beta(\xi/2^j) = 1$ for $\xi \neq 0$. We write $f_\mu = \Delta_\mu f$ and similarly for g, u, v. Thus $f = \sum_\mu f_\mu$ etc.

We shall prove the following:

Theorem 3.1. *The solutions u, v of (3.1) satisfy the following dyadic spacetime estimates:*

(i) $\left\| \Delta_\mu(u_\lambda v_\lambda)\right\|_{L^2_{t,x}} \lesssim \varepsilon^{1/2}\mu \left\| f_\lambda\right\|_{L^2} \left\|g_\lambda\right\|_{L^2}$ *if* $\mu \lesssim \lambda \lesssim 1/\varepsilon$.

(ii) $\left\| \Delta_\mu(u_\lambda v_\lambda)\right\|_{L^2_{t,x}} \lesssim \varepsilon^{1/2}\mu^{1/2}\lambda^{1/2} \left\| f_\lambda\right\|_{L^2} \left\|g_\lambda\right\|_{L^2}$ *if* $\mu \lesssim \lambda$, $\lambda \gg 1/\varepsilon$.

(iii) $\left\|u_\mu v_\lambda\right\|_{L^2_{t,x}} \lesssim \varepsilon^{1/2}\min(\mu, \lambda) \left\| f_\mu\right\|_{L^2} \left\|g_\lambda\right\|_{L^2}$ *for all* μ, λ.

See [8, Thm. 12.1] for the analogous estimates when u and v both solve the free wave equation. Cf. the remarks above.

By decomposing the product uv into dyadic pieces, then applying Theorem 3.1 and finally exploiting the orthogonality properties in Fourier space to sum up, one obtains the following corollary. (The complete argument can be found in [8, Sect. 12].)

Corollary 3.2. $\left\|D^{-\sigma}(uv)\right\|_{L^2_{t,x}} \le C_{s_1,s_2}\,\varepsilon^{1/2} \left\| f\right\|_{\dot H^{s_1}} \left\|g\right\|_{\dot H^{s_2}}$ *provided that* $s_1, s_2 < 1$, $\sigma < \frac{1}{2}$ *and* $s_1 + s_2 + \sigma = 1$.

Estimates of this type for the case where u and v both solve the free wave equation were first investigated by Klainerman and Machedon. The case $(s_1, s_2, \sigma) = (0, 1, 0)$ is excluded, a fact related to the false endpoint case of the Strichartz estimates for the wave equation in $1+3$ dimensions. However, by assuming a little extra regularity one can easily sum the dyadic pieces and one obtains the following nonsharp bilinear estimate.

Corollary 3.3. *$\|uv\|_{L^2_{t,x}} \le C_\delta \varepsilon^{1/2} \|f\|_{L^2} \|g\|_{H^{1+\delta}}$ for all $\delta > 0$.*

Proof. It suffices to prove the sharp estimate

$$\|uv_\lambda\|_{L^2_{t,x}} \lesssim \varepsilon^{1/2} \|f\|_{L^2} \, \lambda \, \|g_\lambda\|_{L^2} . \tag{3.2}$$

Write $u = \sum_\mu u_\mu$ and consider the cases $\mu \lesssim \lambda$ and $\mu \gg \lambda$. In the first case,

$$\left\| \left(\sum\nolimits_{\mu \lesssim \lambda} u_\mu \right) v_\lambda \right\|_{L^2} \lesssim \sum\nolimits_{\mu \lesssim \lambda} \|u_\mu v_\lambda\|_{L^2} \lesssim \varepsilon^{1/2} \left(\sum\nolimits_{\mu \lesssim \lambda} \frac{\mu}{\lambda} \|f_\mu\|_{L^2} \right) \lambda \, \|g_\lambda\|_{L^2} ,$$

where we used Theorem 3.1(iii) to get the last inequality. In the second case we have, by orthogonality in Fourier space,

$$\left\| \left(\sum\nolimits_{\mu \gg \lambda} u_\mu \right) v_\lambda \right\|^2_{L^2} \lesssim \sum\nolimits_{\mu \gg \lambda} \|u_\mu v_\lambda\|^2_{L^2} ,$$

and by Theorem 3.1(iii) we dominate this by $\varepsilon \|f\|^2_{L^2} \lambda^2 \|g_\lambda\|^2_{L^2}$. □

Here we could also take f in $H^{1+\delta}$ and $g \in L^2$, but we shall not need this. However, for null bilinear forms one can get the sharp result (i.e. $\delta = 0$). Thus, we recall the following, proved in [2, Proposition 4]:

$$\left\| Q_{ij}(u, v) \right\|_{L^2_{t,x}} \lesssim \varepsilon^{1/2} \|f\|_{\dot H^2} \|g\|_{\dot H^1} \tag{3.3}$$

where Q_{ij} is given by (2.19). We remark that this is the analogoue of an estimate for two solutions of the free wave equation proved by Klainerman and Machedon.

Since we will prove part (ii) of Theorem 3.1 by a reduction to linear Strichartz estimates, let us recall these (for $1+3$ dimensions). We say that a pair (q, r) of Lebesgue exponents is *wave admissible* if $(q, r) \neq (2, \infty)$ and $1/q + 1/r \le 1/2$, and *sharp wave admissible* if the last inequality is an equality.

For the free wave u in (3.1) one has the well-known estimate

$$\|u\|_{L^q_t L^r_x} \le C_{q,r} \, \varepsilon^{1/q} \|f\|_{\dot H^s} , \tag{3.4}$$

for wave admissible (q, r) and $s = 3/2 - 3/r - 1/q$. As proved in [17], this can be improved if the Fourier support of f is small. Thus, if $\widehat{f}$ is supported in a cube with side length $\sim \mu$ and at distance $\sim \lambda$ from the origin, where $\mu \ll \lambda$, then

$$\|u\|_{L^q_t L^r_x} \le C_{q,r} \, \varepsilon^{1/q} \left(\frac{\mu}{\lambda} \right)^{1/2 - 1/r} \|f\|_{\dot H^s} , \tag{3.5}$$

for q, r, s as above.

For v satisfying (3.1) we have, as proved in [2, Proposition 1],

$$\|v\|_{L^q_t L^r_x} \le C_{q,r} \left(\|g_{\text{low}}\|_{\dot H^{1/q}} + \varepsilon^{1/q} \left\| g_{\text{high}} \right\|_{\dot H^{2/q}} \right) \tag{3.6}$$

for sharp wave admissible (q, r). (Then one can use Sobolev embedding to obtain estimates for all wave admissible pairs.) In order to prove Theorem 3.1(ii) we need the analogue of (3.5) in this context. Thus, we shall prove:

Proposition 3.4. *Let v be as in (3.1), and suppose $\widehat{g}$ is supported in a cube with side length $\sim \mu$ and at distance $\sim \lambda$ from the origin, where $\mu \ll \lambda$. Then*

$$\|v\|_{L_t^q L_x^r} \le C_{q,r} \left(\frac{\mu}{\lambda}\right)^{1/2-1/r} \left(\|g_{\text{low}}\|_{\dot{H}^{1/q}} + \varepsilon^{1/q} \left\|g_{\text{high}}\right\|_{\dot{H}^{2/q}}\right) \tag{3.7}$$

for sharp wave admissible (q, r).

Finally, recalling the basic heuristic that $L_\pm^\varepsilon$ behaves like a Schrödinger operator at low frequencies, it is not surprising that we have the following Schrödinger type estimates, proved in [2]. We say that a pair (q, r) is *Schrödinger admissible* if $q, r \ge 2$ and $2/q + 3/r = 3/2$.

Proposition 3.5. *Let (q, r) and $(\widetilde{q}, \widetilde{r})$ be any two Schrödinger admissible pairs. Then for the solution of $L_\pm^\varepsilon v = F$ with data $v|_{t=0} = f$ we have*

$$\|v_{\text{low}}\|_{L_t^q L_x^r(S_T)} + \|v_{\text{low}}\|_{L_t^\infty L_x^2(S_T)} \lesssim \|f_{\text{low}}\|_{L^2} + \|F_{\text{low}}\|_{L_t^{\widetilde{q}'} L_x^{\widetilde{r}'}(S_T)},$$

where $\frac{1}{\widetilde{q}} + \frac{1}{\widetilde{q}'} = 1$ and $\frac{1}{\widetilde{r}} + \frac{1}{\widetilde{r}'} = 1$.

4. FUNCTION SPACES

We shall use the following spaces of functions on $\mathbb{R}^{1+3}$ with weighted norms defined in Fourier space:

- $H_\varepsilon^{s,\theta}$ with norm $\left\|\langle\xi\rangle^s \langle|\tau| - \varepsilon^{-1}|\xi|\rangle^\theta \widehat{u}(\tau,\xi)\right\|_{L^2_{\tau,\xi}}$.
- $\dot{H}_\varepsilon^{s,\theta}$ with norm $\left\||\xi|^s \langle|\tau| - \varepsilon^{-1}|\xi|\rangle^\theta \widehat{u}(\tau,\xi)\right\|_{L^2_{\tau,\xi}}$.
- $\mathcal{H}_\varepsilon^{s,\theta}$ with norm $\|u\|_{H_\varepsilon^{s,\theta}} + \varepsilon\|\partial_t u\|_{H_\varepsilon^{s-1,\theta}}$.
- $\dot{\mathcal{H}}_\varepsilon^{1,\theta}$ with norm $\|u\|_{\dot{H}_\varepsilon^{1,\theta}} + \varepsilon\|\partial_t u\|_{H_\varepsilon^{0,\theta}}$.
- $X^{s,\theta}_{\tau=\pm h_\varepsilon(\xi)}$ with norm $\left\|\langle\xi\rangle^s \langle\tau \mp h_\varepsilon(\xi)\rangle^\theta \widehat{u}(\tau,\xi)\right\|_{L^2_{\tau,\xi}}$ and h_ε as in (2.14).

Here $\langle\cdot\rangle$ stands for $1 + |\cdot|$. These spaces are by now standard, and we will recall their main properties without proofs. For more details and further references to the literature, the reader may consult e.g. [26], [16].

It will be convenient to introduce the notation

$$U^\varepsilon(t) = e^{it(\lambda(\varepsilon D)-1)/\varepsilon^2} = e^{ith_\varepsilon(D)}, \quad S(t) = e^{it\Delta/2}, \quad W^\varepsilon(t) = e^{itD/\varepsilon} \tag{4.1}$$

for the propagators associated to, respectively, the operators $L_\pm^\varepsilon$ defined in Lemma 2.1, the Schrödinger operator and the wave operator.

(i) *Superposition principle.* A fundamental property of the so-called "Wave Sobolev space" $H_\varepsilon^{s,\theta}$ is that any function in this space can be written as a superposition (H^s-valued integral over the real line) of solutions of the free wave equation with initial data in H^s. (See [16, Proposition 3.4] for the precise statement.) This, in effect, replaces Duhamel's principle in the framework of the Wave Sobolev spaces, and it has the following simple but extremely useful consequence (see [16]):

Theorem 4.1. ***(Transfer Principle.)*** *Suppose T is a multilinear operator $(f_1(x), \dots, f_k(x)) \mapsto T(f_1, \dots, f_k)(x)$ acting in x-space. If T satisfies an estimate*

$$\left\|T\left(W^\varepsilon(\pm t) f_1, \dots, W^\varepsilon(\pm t) f_k\right)\right\|_{L_t^q L_x^r} \le C\varepsilon^{1/q}\|f_1\|_{H^{s_1}} \cdots \|f_k\|_{H^{s_k}},$$

for all combinations of signs, then

$$\|T(u_1, \dots, u_k)\|_{L_t^q L_x^r} \le C_\theta \varepsilon^{1/q}\|u_1\|_{H_\varepsilon^{s_1,\theta}} \cdots \|u_k\|_{H_\varepsilon^{s_k,\theta}}$$

holds for all $u_j \in H_\varepsilon^{s_j,\theta}$*, provided* $\theta > 1/2$*. Moreover, the same statement holds with* H^s *and* $H_\varepsilon^{s_j,\theta}$ *replaced by their homogeneous counterparts.*

Remark 4.2. The spaces $X^{s,\theta}_{\tau=\pm h_\varepsilon(\xi)}$ are related to the equation $L^\varepsilon_\pm v = 0$ in the same way that the Wave Sobolev spaces are related to the free wave equation. Thus, we have a superposition principle and hence a transfer principle for these spaces as well. To be precise, in the above theorem, one can replace any one of the $W^\varepsilon(\pm t)$ by $U^\varepsilon(\pm t)$ and correspondingly $H_\varepsilon^{s_j,\theta}$ by $X^{s_j,\theta}_{\tau=\pm h_\varepsilon(\xi)}$.

Applying the transfer principle to estimates from the previous section, we have, for $\theta > 1/2$,

$$\left\| u_{\text{high}} \right\|_{L_t^q L_x^r} \lesssim \varepsilon^{1/2+1/r} \left\| u_{\text{high}} \right\|_{X^{1,\theta}_{\tau=\pm h_\varepsilon(\xi)}} \quad \text{for sharp wave adm. } (q,r), \tag{4.2}$$

$$\left\| u_{\text{low}} \right\|_{L_t^q L_x^r} \lesssim \left\| u_{\text{low}} \right\|_{X^{0,\theta}_{\tau=\pm h_\varepsilon(\xi)}} \quad \text{for Schrödinger adm. } (q,r), \tag{4.3}$$

$$\left\| u_{\text{low}} \right\|_{L_t^2 L_x^\infty} \lesssim \left\| u_{\text{low}} \right\|_{X^{1,\theta}_{\tau=\pm h_\varepsilon(\xi)}} . \tag{4.4}$$

Here (4.3) follows from Proposition 3.5 with $F = 0$. By Sobolev embedding we reduce (4.4) to the case $(q,r) = (2,6)$ of (4.3). Finally, (4.2) holds by virtue of (3.6) and the trivial estimate

$$\left\| f_{\text{high}} \right\|_{H^s} \lesssim \varepsilon^\sigma \left\| f_{\text{high}} \right\|_{H^{s+\sigma}} \quad \text{for} \quad \sigma > 0. \tag{4.5}$$

(ii) *Embeddings.* The most basic embeddings are

$$H_\varepsilon^{s,\theta}, \ X^{s,\theta}_{\tau=\pm h_\varepsilon(\xi)} \hookrightarrow C_b(\mathbb{R}; H^s), \qquad \dot{H}_\varepsilon^{s,\theta} \hookrightarrow C_b(\mathbb{R}; \dot{H}^s), \tag{4.6}$$

which hold *uniformly in* ε for any $\theta > 1/2$. Also uniform in ε are

$$L_t^p L_x^2 \hookrightarrow H_\varepsilon^{0,\theta-1}, \ X^{0,\theta-1}_{\tau=\pm h_\varepsilon(\xi)} \quad \text{for} \quad \frac{1}{\frac{3}{2}-\theta} < p \le 2, \quad \frac{1}{2} < \theta < 1. \tag{4.7}$$

In fact, the dual statement $H_\varepsilon^{0,1-\theta}$, $X^{0,1-\theta}_{\tau=\pm h_\varepsilon(\xi)} \hookrightarrow L_t^{p'} L_x^2$ follows by interpolation between the trivial case $p = 2$, $\theta = 1$ and (4.6). We shall also need

$$\left\| e^{\pm it/\varepsilon^2} u \right\|_{X^{0,\theta-1}_{\tau=\pm h_\varepsilon(\xi)}} \lesssim \varepsilon^{-2(1-\theta)} \left\| u \right\|_{H_\varepsilon^{0,\theta-1}} . \tag{4.8}$$

This is obvious if $\widehat{u}(\tau,\xi)$ is supported in $\left| |\tau| - |\xi|/\varepsilon \right| \lesssim \varepsilon^{-2}$; then we can in fact replace the left hand side by $\|u\|_{L^2}$. On the other hand, if $\widehat{u}(\tau,\xi)$ is supported in $\left| |\tau| - |\xi|/\varepsilon \right| \gg \varepsilon^{-2}$, then (4.8) follows from (2.17).

(iii) *Time cut-off.* We denote by $H_\varepsilon^{s,\theta}(S_T)$ the restriction to

$$S_T = (0,T) \times \mathbb{R}^3.$$

The norm $\|u\|_{H_\varepsilon^{s,\theta}(S_T)} = \inf\left\{ \|v\|_{H_\varepsilon^{s,\theta}} : v = u \text{ on } S_T \right\}$ makes this a complete space. Norms on the other restriction spaces are similarly defined.

The idea behind the following “cut-off lemmas” originates in the work of Bourgain [4] on the Schrödinger and KdV equations, and was developed further by Kenig-Ponce-Vega [11] in their work on KdV and by Klainerman-Machedon [14] and the last author [25] for the wave equation. In fact, the argument given in [11] applies to $X^{s,\theta}$ spaces in general, and in particular proves the following.

Lemma 4.3. *Suppose* $L^{\varepsilon}_{\pm} v = F$ *on* S_T *with* $v|_{t=0} = f$. *Let* $\theta > 1/2$. *Then for* $0 \le T \le 1$,

$$\|v\|_{X^{s,\theta}_{\tau=\pm h_\varepsilon(\xi)}(S_T)} \le C_\theta \left(\|f\|_{H^s} + \|F\|_{X^{s,\theta-1}_{\tau=\pm h_\varepsilon(\xi)}(S_T)} \right)$$

where C_θ *is independent of* T *and* ε.

By rescaling $x \to \varepsilon x$ we reduce the next result to the case $\varepsilon = 1$, which in turn follows from estimates proved in [14].

Lemma 4.4. *Suppose* $\Box_\varepsilon u = F$ *on* S_T *with* $(u, \partial_t u)|_{t=0} = (f, g)$. *Let* $\theta > 1/2$. *Then for* $0 \le T \le 1$,

$$\|u\|_{\dot{\mathcal{H}}^{1,\theta}_\varepsilon(S_T)} \le C_\theta \left(\|f\|_{\dot H^1} + \varepsilon \|g\|_{L^2} + \frac{1}{\varepsilon} \|F\|_{H^{0,\theta-1}_\varepsilon(S_T)} \right) \tag{4.9}$$

where C_θ *is independent of* T *and* ε. *Also, for* $M \in \mathbb{N}$ *large enough,*

$$\|u\|_{\mathcal{H}^{s,\theta}_\varepsilon(S_T)} \le C_\theta \varepsilon^{-M} \left(\|f\|_{H^s} + \|g\|_{H^{s-1}} + \|F\|_{\mathcal{H}^{s-1,\theta-1}_\varepsilon(S_T)} \right). \tag{4.10}$$

The last inequality is not sharp w.r.t. ε, but it will only be used in a situation where powers of ε are not important. In order to estimate the Dirac current density we shall need the following “integration by parts”-version of Lemma 4.4.

Lemma 4.5. *If* $\Box_\varepsilon u = e^{it/\varepsilon^2} F$ *on* S_T *with vanishing data, then*

$$\|u\|_{\dot{\mathcal{H}}^{1,\theta}_\varepsilon(S_T)} \lesssim \varepsilon \|\partial_t F\|_{L^2(S_T)} + \|F\|_{L^2_t H^1(S_T)} + \varepsilon^{1-2\theta} \|F\|_{L^2(S_T)}$$

for $1/2 < \theta \le 1$ *and all* $0 \le T \le 1$.

Proof. Write $e^{it/\varepsilon^2} F = (\varepsilon^2/i)\, \partial_t \big[e^{it/\varepsilon^2} F \big] - (\varepsilon^2/i) e^{it/\varepsilon^2} \partial_t F$ and $u = u_1 + u_2$ accordingly. By (4.9), $\|u_2\|_{\dot{\mathcal{H}}^{1,\theta}_\varepsilon(S_T)} \lesssim \varepsilon \big\| \partial_t F \big\|_{L^2(S_T)}$.

Now let F_{ext} be any extension of F to all of $\mathbb{R}^{1+3}$, and set $G = e^{it/\varepsilon^2} F_{\mathrm{ext}}$. Split $G = G_1 + G_2$ such that $\widehat{G}_1(\tau, \xi)$ is supported in $|\tau| \lesssim |\xi|/\varepsilon$ while $\widehat{G}_2(\tau, \xi)$ is supported in $|\tau| \gg |\xi|/\varepsilon$, and write $u_1 = u_{1,1} + u_{1,2}$ accordingly; i.e., $\Box_\varepsilon u_{1,j} = (\varepsilon^2/i)\, \partial_t G_j$ on S_T, with vanishing data.

By (4.9), $\big\| u_{1,1} \big\|_{\dot{\mathcal{H}}^{1,\theta}_\varepsilon(S_T)} \lesssim \varepsilon \|\partial_t G_1\|_{L^2}$, and by Plancherel’s theorem and the assumptions on the Fourier support,

$$\|\partial_t G_1\|_{L^2} \lesssim \frac{1}{\varepsilon} \big\| D G_1 \big\|_{L^2} \le \frac{1}{\varepsilon} \big\| D G \big\|_{L^2} = \frac{1}{\varepsilon} \big\| D F_{\mathrm{ext}} \big\|_{L^2}.$$

Hence $\big\| u_{1,1} \big\|_{\dot{\mathcal{H}}^{1,\theta}_\varepsilon(S_T)} \lesssim \|F\|_{L^2_t H^1(S_T)}$.

Finally, to estimate $u_{1,2}$ we first observe that it has an extension to all of $\mathbb{R}^{1+3}$, defined in Fourier space by

$$\widehat{u}_{1,2}(\tau, \xi) = \frac{1}{-\varepsilon^2 \tau^2 + \left| \xi^2 \right|} \Big[(\varepsilon^2/i)\, \partial_t G_2 \Big]^{\widehat{}}(\tau, \xi).$$

Thus $\big| \widehat{u}_{1,2}(\tau, \xi) \big| \sim \frac{1}{|\tau|} \big| \widehat{G}_2(\tau, \xi) \big|$, and since

$$\big\| u_{1,2} \big\|_{\dot{\mathcal{H}}^{1,\theta}_\varepsilon(S_T)} \le \big\| u_{1,2} \big\|_{\dot{\mathcal{H}}^{1,\theta}_\varepsilon(\mathbb{R}^{1+3})} \lesssim \big\| (\varepsilon |\tau| + |\xi|) \langle |\tau| - |\xi|/\varepsilon \rangle^\theta \, \widehat{u}_{1,2}(\tau, \xi) \big\|_{L^2_{\tau,\xi}}$$

we conclude that $\big\| u_{1,2} \big\|_{\dot{\mathcal{H}}^{1,\theta}_\varepsilon(S_T)}$ is dominated by

$$\varepsilon \big\| \langle \tau \rangle^\theta \, \widehat{G}_2(\tau, \xi) \big\|_{L^2_{\tau,\xi}} \lesssim \varepsilon \big\| \langle \tau + 1/\varepsilon^2 \rangle^\theta \, \widehat{F}_{\mathrm{ext}}(\tau, \xi) \big\|_{L^2_{\tau,\xi}} \lesssim \varepsilon \|\partial_t F_{\mathrm{ext}}\|_{L^2} + \frac{\|F_{\mathrm{ext}}\|_{L^2}}{\varepsilon^{2\theta-1}},$$

where we used the assumption $\theta \le 1$ to get the first term on the right hand side. Since ∂_t is a local operator, taking the infimum over all extensions F_{ext} gives the desired estimate. $\square$

5. MAIN ESTIMATES

Here we prove the main *a priori* estimates for the nonlinear terms in DM. Throughout this section we shall assume $0 < T \le 1$. We define norms

$$\begin{gathered} X_T^\varepsilon(\cdot) = \varepsilon^\Lambda \left\| \cdot \right\|_{\dot{\mathcal{H}}_\varepsilon^{1,\theta}(S_T)}, \quad Y_{T,\pm}^\varepsilon(\cdot) = \left\| \cdot \right\|_{X^{1,\theta}_{\tau=\pm h_\varepsilon(\xi)}(S_T)}, \\ Z_T^\varepsilon(\cdot) = \left\| (\cdot)_{\text{low}} \right\|_{L_t^2 L_x^6 \cap L_t^\infty L_x^2(S_T)}, \end{gathered} \tag{5.1}$$

for a fixed $\theta > 1/2$ satisfying [cf. (1.7)]

$$\Lambda + 1 - 2\theta > 0. \tag{5.2}$$

In estimates where the choice of sign in $Y_{T,\pm}^\varepsilon$ is open, we will simply write Y_T^ε.

5.1. **Estimates for A_0.**

Lemma 5.1. *Consider the operator* $A_0[\psi] = \Delta^{-1} \langle \psi, \psi \rangle$. *Split* $\psi = \psi_{\text{low}} + \psi_{\text{high}}$ *and write*

$$A_0[\psi] = A_0^{\varepsilon\prime}[\psi] + A_0^{\varepsilon\prime\prime}[\psi] \tag{5.3}$$

where $A_0^{\varepsilon\prime}$ *corresponds to “low-low” interactions:* $A_0^{\varepsilon\prime}[\psi] = \Delta^{-1} \langle \psi_{\text{low}}, \psi_{\text{low}} \rangle$. *For these operators, the following estimates hold:*

$$\left\| \Delta A_0[\psi] \right\|_{L_t^\infty L_x^r(S_T)} \lesssim Y_T^\varepsilon(\psi)^2 \qquad \text{for} \quad 1 \le r \le 3, \tag{5.4}$$

$$\left\| \Delta A_0[\psi] \right\|_{L_t^2 L_x^r(S_T)} \lesssim Y_T^\varepsilon(\psi)^2 \qquad \text{for} \quad 1 \le r \le 6. \tag{5.5}$$

$$\left\| \Delta A_0^{\varepsilon\prime}[\psi] \right\|_{L_t^2 L_x^r(S_T)} \lesssim Z_T^\varepsilon(\psi)^2 \qquad \text{for} \quad 1 \le r \le 3/2, \tag{5.6}$$

$$\left\| \Delta A_0^{\varepsilon\prime\prime}[\psi] \right\|_{L_t^2 L_x^r(S_T)} \lesssim \varepsilon Y_T^\varepsilon(\psi)^2 \qquad \text{for} \quad 1 \le r \le 3/2, \tag{5.7}$$

$$\left\| \Delta A_0^{\varepsilon\prime}[\psi] \right\|_{L_t^p L_x^{(3/2)+}(S_T)} \lesssim Z_T^\varepsilon(\psi)^2 \qquad \text{for} \quad 1 \le p < 2, \tag{5.8}$$

$$\left\| \Delta A_0^{\varepsilon\prime\prime}[\psi] \right\|_{L_t^p L_x^{(3/2)+}(S_T)} \lesssim \varepsilon^{1^-} Y_T^\varepsilon(\psi)^2 \qquad \text{for} \quad 1 \le p < 2. \tag{5.9}$$

Proof. (5.4), hence (5.5) for $r \le 3$, follow from Hölder’s inequality and Sobolev embedding, while (5.5) for larger r follows from the estimate $\|\psi\|_{L_t^4 L_x^{2r}(S_T)} \lesssim Y_T^\varepsilon(\psi)$ for $2 \le r \le 6$; by Sobolev embedding and the transfer principle (see Theorem 4.1 and the remark following it), the latter reduces to the $L_t^4 L_x^4$ Strichartz estimate (3.6). Let us now prove (5.8) and (5.9); the proofs of (5.6) and (5.7) are similar. Write

$$\left\| \langle \psi, \psi \rangle \right\|_{L_x^{(3/2)+}} \le \|\psi\|_{L_x^6} \|\psi\|_{L_x^{2+}} \lesssim \|\psi\|_{L_x^6}^{1^+} \|\psi\|_{L_x^2}^{1^-}. \tag{5.10}$$

For $\psi = \psi_{\text{low}}$ the $L_t^{2^-}$ norm of this is clearly dominated by the right hand side of (5.8). On the other hand, if at least one ψ_{high} is present, then we dominate by the right hand side of (5.9) using the $\dot{H}^1 \hookrightarrow L_x^6$ Sobolev embedding and the estimate (4.5). $\square$

In connection with the above estimates we shall need the embeddings

$$\|f\|_{L_x^\infty} \lesssim \|\Delta f\|_{L_x^{(3/2)-}} + \|\Delta f\|_{L_x^{(3/2)+}}, \tag{5.11}$$

$$\|\nabla f\|_{L_x^\infty} \lesssim \|\Delta f\|_{L_x^{3-}} + \|\Delta f\|_{L_x^{3+}}. \tag{5.12}$$

In particular, these imply the estimates (not involving time)

$$\|A_0 \cdot \psi\|_{H^1} \lesssim \left(\|\Delta A_0\|_{L_x^{(3/2)-}} + \|\Delta A_0\|_{L_x^{(3/2)+}}\right) \|\psi\|_{H^1}, \tag{5.13}$$

$$\|\nabla A_0 \cdot \psi\|_{H^1} \lesssim \left(\|\Delta A_0\|_{L_x^{3-}} + \|\Delta A_0\|_{L_x^{3+}}\right) \|\psi\|_{H^1}. \tag{5.14}$$

Next, we consider estimates for the commutator term in (2.11).

Lemma 5.2. *With notation as in Lemma 5.1, we have the estimates*

$$\left\|\left[A_0'[\psi], \lambda(\varepsilon D) - 1\right] \psi'\right\|_{L^2(S_T)} \lesssim \varepsilon^{2^-} Y_T^\varepsilon(\psi)^2 Y_T^\varepsilon(\psi'), \tag{5.15}$$

$$\begin{aligned} &\left\|\left[A_0''[\psi], \lambda(\varepsilon D) - 1\right] \psi'\right\|_{L^2(S_T)} \\ &\qquad \lesssim \varepsilon \left(\left\|A_0^{\varepsilon\prime\prime}[\psi]\right\|_{L_t^2 L_x^\infty(S_T)} + \left\|\nabla A_0^{\varepsilon\prime\prime}[\psi]\right\|_{L_t^2 L_x^3(S_T)}\right) Y_T^\varepsilon(\psi'), \end{aligned} \tag{5.16}$$

Proof. As proved in [2, Lemma 9],

$$\left\|[\Delta^{-1}(fg), \lambda(\varepsilon D) - 1]h\right\|_{L_x^2} \le \varepsilon^2 C_\rho \|f\|_{H^{1+\rho}} \|g\|_{H^{1+\rho}} \|h\|_{H^1} \quad \text{(for all } \rho > 0\text{)},$$

yielding (5.15). For (5.16) we simply expand the commutator, and use the fact that $\lambda(\varepsilon D) - 1$ can in effect be replaced by $\varepsilon\nabla$. To be precise,

$$\|(\lambda(\varepsilon D) - 1) f\|_{L^2} \lesssim \varepsilon \|f\|_{\dot{H}^1}, \tag{5.17}$$

in view of (2.14). □

5.2. **Estimates for the main bilinear terms.** Here we prove estimates for the key bilinear terms appearing in the "remainder" $e^{\pm it/\varepsilon^2} R^\varepsilon$ appearing in Lemma 2.1. We make use of the null structure inherent in these terms (cf. Lemma 2.3) and the following null form estimate.

Theorem 5.3. *Let* $1/2 < \theta < 1$. *Then*

$$\|Q(u,v)\|_{H_\varepsilon^{0,\theta-1}} \le C_\theta \|u\|_{\dot{\mathcal{H}}_\varepsilon^{1,\theta}} \|v\|_{\mathcal{H}_\varepsilon^{(1+\theta)+,1}}$$

holds on $\mathbb{R}^{1+3}$ *for all null forms* Q *in (2.19). Moreover, if* $Q = Q_{ij}$, *then the norm* $\|u\|_{\dot{\mathcal{H}}_\varepsilon^{1,\theta}}$ *in the the right hand side can be replaced by* $\|u\|_{\dot{H}_\varepsilon^{1,\theta}}$.

By a standard procedure we reduce this to well-known bilinear estimates for the homogeneous wave equation; the proof is in Sect. 13.

Theorem 5.4. *Fix* $\varepsilon > 0$. *Given a spinor* $\phi(t,x)$ *and a potential*[5] $\{A_\mu(t,x)\}$ *satisfying the Coulomb condition* $\operatorname{div} \mathbf{A} = 0$, *we define* $\mathbf{E} = \nabla A_0 - \varepsilon \partial_t \mathbf{A}$ *and* $\mathbf{B} = \nabla \times \mathbf{A}$. *Let* ψ *be the solution of the IVP*

$$\left\{i\varepsilon\partial_t + i\alpha^k\partial_k - (1/\varepsilon)\gamma^0\right\}\psi = -\varepsilon A_k \alpha^k \phi - \varepsilon A_0 \phi, \qquad \psi|_{t=0} = \psi_0, \tag{5.18}$$

in the fixed time slab S_T. *Then for the bilinear expression [cf. (2.11)]*

$$\mathcal{B} := [\lambda(\varepsilon D)]^{-1} \varepsilon \left\{2i\mathbf{A} \cdot \nabla + i E_k \alpha^k - B_k S^k\right\} \psi$$

---

[5] The fields are assumed to be sufficiently smooth. The precise regularity is not important since we shall derive a priori estimates.

we have the estimates [recall (5.1)]

$$\left\| e^{\pm it/\varepsilon^2} \mathcal{B} \right\|_{X^{1,\theta-1}_{\tau=\pm h_\varepsilon(\xi)}(S_T)} \lesssim \varepsilon^{(1/2-\Lambda)^-} C \left(Y^\varepsilon_T(\psi) + Y^\varepsilon_T(\phi) \right), \tag{5.19}$$

$$\|\mathcal{B}\|_{L^2(S_T)} \lesssim \varepsilon C \left(Y^\varepsilon_T(\psi) + Y^\varepsilon_T(\phi) \right), \tag{5.20}$$

where C depends polynomially (with coefficients independent of ε) on the norms

$$X^\varepsilon_T(\mathbf{A}), \quad \|\Delta A_0\|_{L^\infty_t L^{(3/2)\pm}_x(S_T)} \quad \text{and} \quad \|\Delta A_0\|_{L^2_t L^{3\pm}_x(S_T)}. \tag{5.21}$$

Proof. We first prove (5.19). Using (2.3), (5.3) and (4.7) we reduce to estimating

$$N_1 = \left\| e^{\pm it/\varepsilon^2} \left\{ i(E_k - \partial_k A_0)\alpha^k - B_k S^k \right\} \psi \right\|_{X^{0,\theta-1}_{\tau=\pm h_\varepsilon(\xi)}(S_T)},$$

$$N_2 = \|\mathbf{A} \cdot \nabla \psi\|_{L^2(S_T)}, \quad N_3 = \left\| \varepsilon(\partial_k A_0)\alpha^k \psi \right\|_{L^2_t H^1(S_T)}.$$

The estimate for N_3 is obtained directly from (5.14). For N_2 we apply Lemma 2.3 and (3.3) via the transfer principle (Theorem 4.1 and the subsequent remark); this gives $N_2 \lesssim \varepsilon^{1/2-\Lambda} X^\varepsilon_T(\mathbf{A}) Y^\varepsilon_T(\psi)$.

For N_1 we write $\left\{ i(E_k - \partial_k A_0)\alpha^k - B_k S^k \right\} \psi = \sum I_\mu$ where

$$I_\mu = \left\{ i(E_k - \partial_k A_0)\alpha^k - B_k S^k \right\} \Delta_\mu \psi$$

and the sum is over all dyadic numbers μ of the form 2^j, $j \in \mathbb{Z}$. Here Δ_μ is the Littlewood-Paley operator defined in Sect. 3. We split into the cases (i) $\mu \le 1/\varepsilon$ and (ii) $\mu > 1/\varepsilon$.

Case (i). By (4.7), we can reduce to proving

$$\left\| \sum\nolimits_{\mu \le 1/\varepsilon} I_\mu \right\|_{L^2(S_T)} \lesssim \varepsilon^{(1/2-\Lambda)^-} X^\varepsilon_T(\mathbf{A}) Y^\varepsilon_T(\psi), \tag{5.22}$$

but this follows from Corollary 3.3 via the transfer principle (see Theorem 4.1 and the remark following it). Note that we used the assumption $\mu \le 1/\varepsilon$ to recover the small loss of regularity (the δ) in Corollary 3.3.

Case (ii). Using (4.7) we write

$$\left\| e^{\pm it/\varepsilon^2} I_\mu \right\|_{X^{0,\theta-1}_{\tau=\pm h_\varepsilon(\xi)}(S_T)} \lesssim \left\| I_\mu \right\|^{1-\sigma}_{L^2(S_T)} \left\| e^{\pm it/\varepsilon^2} I_\mu \right\|^{\sigma}_{X^{0,\theta-1}_{\tau=\pm h_\varepsilon(\xi)}(S_T)} \tag{5.23}$$

where $0 < \sigma \ll 1$ will be chosen later. Proceeding as in case (i), but using the sharp estimate (3.2), we obtain $\left\| I_\mu \right\|_{L^2(S_T)} \lesssim \varepsilon^{1/2-\Lambda} X^\varepsilon_T(\mathbf{A}) Y^\varepsilon_T(\psi)$. Further, we claim there exist $\zeta > 0$ and $M \in \mathbb{N}$, both independent of ε and μ, such that

$$\left\| e^{\pm it/\varepsilon^2} I_\mu \right\|_{X^{0,\theta-1}_{\tau=\pm h_\varepsilon(\xi)}(S_T)} \lesssim \varepsilon^{-M} \mu^{-\zeta} C' \left[Y^\varepsilon_T(\psi) + Y^\varepsilon_T(\phi) \right], \tag{5.24}$$

where C' depends polynomially on (the first two) norms in (5.21). Granting this claim for the moment, we see that by choosing σ sufficiently small in (5.23), depending on M, and summing the inequalities over $\mu > 1/\varepsilon$, we get the desired bound (5.19).

Let us prove the claim. In the notation of Lemma 2.3,

$$\begin{aligned} I_\mu = {} & Q_{jk}(D^{-1}\varepsilon\partial_t a^{jk}, \Delta_\mu U) - Q_{jk}(D^{-1}\partial_l a^{jk}, \alpha^l \Delta_\mu U) \\ & + Q_0(A_j, \alpha^j \Delta_\mu U) + Q_{0j}(A_k, \alpha^j \alpha^k \Delta_\mu U) - \frac{i}{2} Q_{jk}(A_m, \epsilon^{jkl} S_l \alpha^m \Delta_\mu U) \end{aligned}$$

where

$$\square_\varepsilon U = -\varepsilon^{-1}\gamma^0 \psi + \varepsilon A_k \alpha^k \phi + \varepsilon A_0 \phi, \quad U|_{t=0} = 0, \quad i\varepsilon\partial_t U|_{t=0} = \psi_0. \tag{5.25}$$

Thus, using (4.8) and Theorem 5.3, we reduce (5.24) to

$$\left\| \Delta_\mu U \right\|_{\mathcal{H}_\varepsilon^{2-2\zeta,1}(S_T)} \le C_\zeta \varepsilon^{-M} \mu^{-\zeta} C' \left[Y_T^\varepsilon(\psi) + Y_T^\varepsilon(\phi) \right] \tag{5.26}$$

for $\zeta > 0$ such that $1 + \theta < 2 - 2\zeta$. Clearly, it suffices to show

$$\| U \|_{\mathcal{H}_\varepsilon^{2-\zeta,1}(S_T)} \le C_\zeta \varepsilon^{-M} C' \left[Y_T^\varepsilon(\psi) + Y_T^\varepsilon(\phi) \right],$$

but using (4.10) and (5.25) we reduce this to (5.13) and

$$\left\| A_k \alpha^k \phi \right\|_{L_t^2 H^{1-\zeta}(S_T)} \le C_\zeta \varepsilon^{1/2-\Lambda} X_T^\varepsilon(\mathbf{A}) Y_T^\varepsilon(\phi),$$

which follows from Corollary 3.2 and the transfer principle. This concludes the proof of (5.19).

The left hand side of (5.20) is $\lesssim \varepsilon^{1^-} \widetilde{N}_1 + \varepsilon N_2 + N_3$, where N_2 and N_3 are as above, while $\widetilde{N}_1 = \left\| i(E_k - \partial_k A_0)\alpha^k \psi - B_k S^k \psi \right\|_{L_t^2 H^{0^-}(S_T)}$; here we used (2.3) with $r = 0^+$. Write $\widetilde{N}_1 \le \widetilde{N}_{1,1} + \widetilde{N}_{1,2}$ corresponding to $\psi = \psi_{\text{low}} + \psi_{\text{high}}$. For the low frequency case we can use (5.22). By Sobolev embedding and Hölder's inequality,

$$\widetilde{N}_{1,2} \lesssim \left(\| \nabla \mathbf{A} \|_{L_t^\infty L_x^2(S_T)} + \varepsilon \, \| \partial_t \mathbf{A} \|_{L_t^\infty L_x^2(S_T)} \right) \left\| \psi_{\text{high}} \right\|_{L_t^{2+} L_x^{\infty-}(S_T)},$$

where $(2^+, \infty^-)$ is chosen to be sharp wave admissible. Now apply (4.2). □

5.3. **Estimates for the current density.** Let $F \mapsto \square_\varepsilon^{-1} F$ be the solution operator of $\square_\varepsilon u = F$ with vanishing initial data.

Lemma 5.5. *Consider the bilinear operator (cf. (1.4c))*

$$\mathbf{A}^\varepsilon[\phi, \psi] = \square_\varepsilon^{-1} \mathbb{P} \langle \vec{\alpha}\phi, \psi \rangle .$$

Recalling the splitting (1.29) into low and high spatial frequencies, as well as the definition (5.1), we have the estimates

$$X_T^\varepsilon \left(\mathbf{A}^\varepsilon[\phi_{\text{low}}, \psi] \right) \lesssim \varepsilon^{\Lambda-1} Y_T^\varepsilon(\phi) \, \| \psi \|_{L_t^\infty L_x^2(S_T)}, \tag{5.27}$$

$$X_T^\varepsilon \left(\mathbf{A}^\varepsilon[\phi_{\text{high}}, \psi] \right) \lesssim \varepsilon^{\Lambda} Y_T^\varepsilon(\phi) Y_T^\varepsilon(\psi). \tag{5.28}$$

Moreover, recalling the definition 2.10,

$$\begin{aligned} X_T^\varepsilon \left(\mathbf{A}^\varepsilon[e^{it/\varepsilon^2}\phi_{\text{low}}, e^{-it/\varepsilon^2}\psi_{\text{low}}] \right) &\lesssim \varepsilon^{\Lambda+1-2\theta} Y_T^\varepsilon(\phi) Y_T^\varepsilon(\psi) \\ &\quad + \varepsilon^{\Lambda+(1/2)^-} \left(\left\| L_\pm^\varepsilon \phi \right\|_{L^2(S_T)} Y_T^\varepsilon(\psi) + Y_T^\varepsilon(\phi) \left\| L_\pm^\varepsilon \psi \right\|_{L^2(S_T)} \right), \end{aligned} \tag{5.29}$$

where the signs in $L_\pm^\varepsilon \phi$ and $L_\pm^\varepsilon \psi$ can be chosen completely arbitrarily.

Proof. We remark that $\mathbb{P}$ is simply "thrown away" here; it is uniformly bounded on all the spaces we work with.

For (5.27), we apply (4.9), followed by (4.4). For (5.28), apply (4.9) again, followed by the general estimates

$$\left\| u_{\text{high}} v_{\text{low}} \right\|_{L^2} \lesssim \left\| u_{\text{high}} \right\|_{L_t^\infty L_x^2} \| v_{\text{low}} \|_{L_t^2 L_x^\infty} \lesssim \varepsilon \left\| u_{\text{high}} \right\|_{X^{1,\theta}_{\tau=\pm h_\varepsilon(\xi)}} \| v_{\text{low}} \|_{X^{1,\theta}_{\tau=\pm h_\varepsilon(\xi)}},$$

$$\left\| u_{\text{high}} v_{\text{high}} \right\|_{L^2} \lesssim \left\| u_{\text{high}} \right\|_{L^4} \left\| v_{\text{high}} \right\|_{L^4} \lesssim \varepsilon^{3/2} \left\| u_{\text{high}} \right\|_{X^{1,\theta}_{\tau=\pm h_\varepsilon(\xi)}} \left\| v_{\text{high}} \right\|_{X^{1,\theta}_{\tau=\pm h_\varepsilon(\xi)}};$$

these follow from (4.2) and (4.4).

By Lemma 4.5, (5.29) reduces to proving the same estimate for

$$\varepsilon^{\Lambda+1} \left\| \partial_t F^\varepsilon \right\|_{L^2(S_T)} + \varepsilon^{\Lambda} \left\| F^\varepsilon \right\|_{L_t^2 H^1(S_T)} + \varepsilon^{\Lambda+1-2\theta} \left\| F^\varepsilon \right\|_{L^2(S_T)}, \tag{5.30}$$

where $F^\varepsilon = \langle \vec{\alpha}\phi_{\text{low}}, \psi_{\text{low}} \rangle$. For the first term, write $i\varepsilon\partial_t = \varepsilon L^\varepsilon_\pm \pm (\lambda(\varepsilon D) - 1)/\varepsilon$ and make use of (5.17) and the estimate $\|u_{\text{low}}\|_{L^\infty(S_T)} \lesssim \varepsilon^{-(1/2)^+} Y^\varepsilon_T(u)$, which holds by Sobolev embedding and (4.6). For the second term in (5.30), we write

$$\left\|F^\varepsilon\right\|_{L_t^2 H^1(S_T)} \lesssim \|\phi\|_{L_t^\infty H^1(S_T)} \|\psi_{\text{low}}\|_{L_t^2 L_x^\infty(S_T)} + \|\phi_{\text{low}}\|_{L_t^2 L_x^\infty(S_T)} \|\psi\|_{L_t^\infty H^1(S_T)},$$

using (4.4) and (4.6). The same estimate is used for the last term in (5.30). □

6. ITERATION SCHEME AND LOCAL EXISTENCE

For fixed ε we have the following local existence theorem :

Theorem 6.1. *For fixed ε, the Dirac-Maxwell-Coulomb system* (1.4) *is locally well posed for initial data in the space* (1.5)*. The existence time $T > 0$ only depends on ε and the size of the norms of the data, and the solution is in the space*

$$\psi^\varepsilon \in H^{1,\theta}_\varepsilon(S_T), \qquad \mathbf{A}^\varepsilon \in \dot{\mathcal{H}}^{1,\theta}_\varepsilon(S_T), \qquad A^\varepsilon_0 \in C([0,T];\dot{H}^1), \tag{6.1}$$

for all $1/2 < \theta < 1$. Moreover, the solution is unique in this class, and we have $\phi^\varepsilon_\pm \in X^{1,\theta}_{\tau=\pm h_\varepsilon(\xi)}(S_T)$, where $\phi^\varepsilon_\pm$ is defined by (2.1) *and* (2.7).

We shall prove this by Picard iteration, using the scheme

$$\left\{i\varepsilon\partial_t + i\alpha^k\partial_k - (1/\varepsilon)\gamma^0\right\}\psi^{\varepsilon,m+1} = -\varepsilon A^{\varepsilon,m}_k \alpha^k \psi^{\varepsilon,m} - \varepsilon A^{\varepsilon,m}_0 \psi^{\varepsilon,m}, \tag{6.2a}$$

$$\Delta A^{\varepsilon,m}_0 = \rho^{\varepsilon,m}, \tag{6.2b}$$

$$\Box_\varepsilon \mathbf{A}^{\varepsilon,m+1} = \varepsilon\mathbb{P}\mathbf{J}^{\varepsilon,m}, \tag{6.2c}$$

with initial data as in (1.5). Here $\rho^{\varepsilon,m}$ and $\mathbf{J}^{\varepsilon,m}$ are given by (1.4d) with ψ^ε replaced by its m-th iterate $\psi^{\varepsilon,m}$. Note that A_0 is not really iterated; (6.2b) simply defines $A^{\varepsilon,m}_0$ in terms of $\psi^{\varepsilon,m}$. Observe also that $\operatorname{div}\mathbf{A}^{\varepsilon,m} = 0$ for all m, since $w = \mathbf{A}^{\varepsilon,m} - \mathbb{P}\mathbf{A}^{\varepsilon,m}$ satisfies $\Box w = 0$ with vanishing initial data.

By convention we start the iteration at $m = -1$ and set all iterates identically equal to zero there. Thus, the iterates $\psi^{\varepsilon,0}, \mathbf{A}^{\varepsilon,0}$ are just solutions of the free Dirac and wave equations with data (1.5). Define [cf. (2.1) and (2.7)]

$$\psi^{\varepsilon,m+1}_\pm = \frac{1}{2}\left\{\psi^{\varepsilon,m+1} \pm \varepsilon^2[\lambda(\varepsilon D)]^{-1}\left(i\partial_t\psi^{\varepsilon,m+1} + A^{\varepsilon,m}_0\psi^{\varepsilon,m}\right)\right\}, \tag{6.3a}$$

$$\phi^{\varepsilon,m}_\pm = \begin{pmatrix} \chi^{\varepsilon,m}_\pm \\ \eta^{\varepsilon,m}_\pm \end{pmatrix} := e^{\pm it/\varepsilon^2}\psi^{\varepsilon,m}_\pm. \tag{6.3b}$$

Proceeding as in the proof of Lemma 2.1 one finds

$$L^\varepsilon_\pm\phi^{\varepsilon,m+1}_\pm = -A^{\varepsilon,m}_0\phi^{\varepsilon,m}_\pm \pm \frac{1}{2}e^{\pm it/\varepsilon^2}R^{\varepsilon,m}, \tag{6.4}$$

where

$$R^{\varepsilon,m} = [\lambda(\varepsilon D)]^{-1}\left\{\varepsilon\mathcal{B}^{\varepsilon,m} + \varepsilon^2\mathcal{C}^{\varepsilon,m} - \left[A^{\varepsilon,m}_0, \lambda(\varepsilon D)\right]\left(\psi^{\varepsilon,m}_+ - \psi^{\varepsilon,m}_-\right)\right\}, \tag{6.5}$$

$$\mathcal{B}^{\varepsilon,m} = \left\{2i\mathbf{A}^{\varepsilon,m}\cdot\nabla + iE^{\varepsilon,m}_k\alpha^k - B^{\varepsilon,m}_k S^k\right\}\psi^{\varepsilon,m}, \tag{6.6}$$

$$\mathcal{C}^{\varepsilon,m} = \left\{A^{\varepsilon,m}_k A^{\varepsilon,m-1}_l \alpha^k\alpha^l + A^{\varepsilon,m}_k A^{\varepsilon,m-1}_0 \alpha^k - A^{\varepsilon,m}_0 A^{\varepsilon,m-1}_k \alpha^k\right\}\psi^{\varepsilon,m-1}, \tag{6.7}$$

and $E^{\varepsilon,m}_k$, $B^{\varepsilon,m}_k$ are given by (2.12) with A^ε_μ replaced by $A^{\varepsilon,m}_\mu$. Note also that

$$\left\|\psi^{\varepsilon,m}_\pm(t)\right\|_{H^1} \lesssim \left\|\psi^{\varepsilon,m}(t)\right\|_{H^1} + \varepsilon\left\|\mathbf{A}^{\varepsilon,m-1}(t)\right\|_{\dot{H}^1}\left\|\psi^{\varepsilon,m-1}(t)\right\|_{H^1}; \tag{6.8}$$

this is just the analogue of (2.5) for the iterates.

We now turn to the proof of Theorem 6.1. By standard arguments, this reduces to proving closed estimates for the iterates in the space (6.1). Set

$$B_T^m = \left\| \psi^{\varepsilon,m} \right\|_{H_\varepsilon^{1,\theta}(S_T)} + \left\| \mathbf{A}^{\varepsilon,m} \right\|_{\dot{\mathcal{H}}_\varepsilon^{1,\theta}(S_T)}, \tag{6.9}$$

and denote by B_0 the norm of the data (1.5). Then it suffices to prove

$$B_T^{m+1} \le CP(B_0) + CT^\delta P\left(B_T^m + B_T^{m-1} \right), \tag{6.10}$$

for some constants $C, \delta > 0$ and a polynomial P with $P(0) = 0$. Here C and P may depend on ε, but since the latter is fixed here, we do not indicate this explicitly. In what follows, C, δ and P may change from line to line. (Observe also that since all the nonlinear terms in DM are in fact multilinear, the same arguments then give estimates for a difference of two iterates.)

By Lemma 4.4,

$$\left\| \mathbf{A}^{\varepsilon,m+1} \right\|_{\dot{\mathcal{H}}_\varepsilon^{1,\theta}(S_T)} \le C\left(\left\| \mathbf{a}_0^\varepsilon \right\|_{\dot{H}^1} + \varepsilon \left\| \mathbf{a}_1^\varepsilon \right\|_{L^2} \right) + CT^\delta \left\| \psi^{\varepsilon,m} \right\|_{L^4(S_T)}^2, \tag{6.11}$$

where the T^δ comes from applying (4.7) and Hölder's inequality in time. Apply (3.6) via the transfer principle to see that $\|u\|_{L^4(S_T)} \lesssim \|u\|_{H_\varepsilon^{1,\theta}(S_T)}$ in general. In order to estimate $\psi^{\varepsilon,m+1}$ we use the splitting (2.8) and the embedding $\|u\|_{H_\varepsilon^{s,\theta}} \lesssim \varepsilon^{-2\theta} \|u\|_{X^{s,\theta}_{\tau=\pm h_\varepsilon(\xi)}}$, which follows from (2.17). Thus, using also Lemma 4.3 and (4.7), we see that

$$\begin{aligned} \left\| \psi^{\varepsilon,m+1} \right\|_{H_\varepsilon^{1,\theta}(S_T)} \lesssim \left\| \phi_\pm^{\varepsilon,m}(0) \right\|_{H^1} &+ T^\delta \left\| A_0^{\varepsilon,m} \phi_\pm^{\varepsilon,m} \right\|_{L_t^\infty H^1(S_T)} \\ &+ \left\| e^{\pm it/\varepsilon^2} R^{\varepsilon,m} \right\|_{X^{1,\theta-1}_{\tau=\pm h_\varepsilon(\xi)}(S_T)}. \end{aligned} \tag{6.12}$$

The first term on the right hand side is covered by (6.8); for the second term we use (5.13) and

$$\left\| \Delta A_0^{\varepsilon,m} \right\|_{L_t^\infty L_x^r(S_T)} \lesssim \left\| \psi^{\varepsilon,m} \right\|_{H_\varepsilon^{1,\theta}(S_T)}^2 \quad \text{for} \quad 1 \le r \le 3, \tag{6.13}$$

where we used Sobolev embedding followed by (4.6). For the third term we claim that

$$\begin{aligned} &\left\| e^{\pm it/\varepsilon^2} R^{\varepsilon,m} \right\|_{X^{1,\theta-1}_{\tau=\pm h_\varepsilon(\xi)}(S_T)} \\ &\qquad \le CT^\delta P\left(B_T^m + B_T^{m-1} \right) \left(\left\| \psi^{\varepsilon,m} \right\|_{H_\varepsilon^{1,\theta}(S_T)} + \left\| \psi^{\varepsilon,m-1} \right\|_{H_\varepsilon^{1,\theta}(S_T)} \right). \end{aligned} \tag{6.14}$$

For the bilinear terms (6.6), this follows from Theorem 5.4 and (5.5); the key observation is that the Y-norms appearing in those estimates can be replaced by $H_\varepsilon^{1,\theta}$-norms. Indeed, the bilinear estimates in Corollaries 3.2 and 3.3, as well as the null form estimate (3.3), are valid also in the case where both u and v solve the homogeneous wave equation, so we can apply the transfer principle for the $H_\varepsilon^{s,\theta}$ spaces instead of $X^{s,\theta}_{\tau=\pm h_\varepsilon(\xi)}$.

It should also be remarked that the proof of Theorem 5.4 shows that one can insert factors T^δ in the right hand sides of (5.19) and (5.20); the observation here is that if (4.7) is used with $p < 2$, and we use Hölder's inequality in time to pass to $p = 2$, then we gain a factor T^δ.

For the cubic terms (6.7), apply [this holds by (4.7) and (2.3)]

$$\left\| e^{\pm it/\varepsilon^2} [\lambda(\varepsilon D)]^{-1} u \right\|_{X^{1,\theta-1}_{\tau=\pm h_\varepsilon(\xi)}(S_T)} \lesssim \frac{1}{\varepsilon} \|u\|_{L_t^{2-} L_x^2(S_T)}, \tag{6.15}$$

followed by $\|fgh\|_{L_x^2} \le \|f\|_{L_x^6} \|g\|_{L_x^6} \|h\|_{L_x^6} \lesssim \|f\|_{H^1} \|g\|_{H^1} \|h\|_{H^1}$.

For the commutator term in (6.5), we use Lemma 5.2 (here too, the Y-norms can be replaced by $H_\varepsilon^{1,\theta}$-norms), followed by (5.11) and Sobolev embedding.

7. UNIFORM H^1 BOUNDS AND LONG TIME EXISTENCE

Setting

$$\mathcal{I}^\varepsilon(t) = \|\psi^\varepsilon(t)\|_{H^1} + \varepsilon^\Lambda\big(\|\mathbf{A}^\varepsilon(t)\|_{\dot H^1} + \varepsilon\,\|\partial_t\mathbf{A}^\varepsilon(t)\|_{L^2}\big),$$

we have the following result.

Theorem 7.1. *Consider the solution $(\psi^\varepsilon, A_\mu^\varepsilon)$ of (1.4), (1.5) from Theorem 6.1, existing up to a time $T_\varepsilon > 0$ and belonging to the space (6.1) over this time interval. There exist*

- (i) *a time $T^* > 0$ depending only on $\sup_{\varepsilon>0}\|\psi_0^\varepsilon\|_{L^2}$,*
- (ii) *constants $C, M, \varepsilon_0 > 0$ independent of ε,*

such that if

$$\mathcal{I}^\varepsilon(0) \le B \quad \textit{for all} \quad \varepsilon, \tag{7.1}$$

where we assume $B \ge 1$, then

$$\mathcal{I}^\varepsilon(T) \le CB \quad \textit{for} \quad 0 < \varepsilon \le \frac{\varepsilon_0}{B^M} \quad \textit{and} \quad 0 \le T \le \min(T^*, T_\varepsilon). \tag{7.2}$$

Moreover, there is a polynomial P with $P(0) = 0$, independent of ε, such that

$$Z_T^\varepsilon(\phi_\pm^\varepsilon) \le C\,\|\psi_0^\varepsilon\|_{L^2} + \varepsilon P\big(\mathcal{I}^\varepsilon(0)\big), \tag{7.3}$$

for T, ε as in (7.2)*. [See (5.1) for the definition of Z_T^ε.]*

Remark 7.2. It follows from the proof (see (7.9) in Corollary 7.4), that the precise dependence of T^* on $R = \sup_{\varepsilon>0}\|\psi_0^\varepsilon\|_{L^2}$ can be taken to be of the form $T^*(R) = \frac{\eta}{1+R^K}$, where $\eta > 0$ and $K \in \mathbb{N}$ are constants independent of ε.

We claim that this result, together with Theorem 6.1, implies Theorem 1.1. To see this, first observe that Theorem 6.1 implies $T_\varepsilon \ge T^*$ for ε as in (7.2). In view of the conservation of charge (1.3) for the Dirac equation, we can now iterate the argument any number of times (T^* will not change), obtaining

$$T_\varepsilon \ge NT^* \quad \text{and} \quad \sup_{0\le T\le NT^*} \mathcal{I}^\varepsilon(T) \le C^N B \quad \text{for} \quad 0 < \varepsilon \le \frac{\varepsilon_0}{(C^{N-1}B)^M} \tag{7.4}$$

for all $N \in \mathbb{N}$. But from the last inequality we have

$$\log\frac{1}{\varepsilon} \ge \log\left(\frac{B^M}{\varepsilon_0}\right) + N\log\left(C^M\right).$$

Thus, as $\varepsilon \to 0$ we may choose N proportional to $\log(1/\varepsilon)$, and since $T_\varepsilon \ge NT^*$ we get the key estimate (1.8) in Theorem 1.1.

In the proof of Theorem 7.1 we use the iterates from Sect. 6. Set [recall (5.1)]

$$X_T^{\varepsilon,m} = X_T^\varepsilon\left(\mathbf{A}^{\varepsilon,m}\right), \quad Y_T^{\varepsilon,m} = Y_{T,+}^\varepsilon\big(\phi_+^{\varepsilon,m}\big) + Y_{T,-}^\varepsilon\big(\phi_-^{\varepsilon,m}\big),$$
$$Z_T^{\varepsilon,m} = Z_T^\varepsilon\big(\phi_+^{\varepsilon,m}\big) + Z_T^\varepsilon\big(\phi_-^{\varepsilon,m}\big),$$

and also

$$X_0^\varepsilon = \varepsilon^\Lambda\left(\|\mathbf{a}_0^\varepsilon\|_{\dot H^1} + \varepsilon\,\|\mathbf{a}_1^\varepsilon\|_{L^2}\right), \quad Y_0^\varepsilon = \|\psi_0^\varepsilon\|_{H^1} \quad \text{and} \quad Z_0^\varepsilon = \|\psi_0^\varepsilon\|_{L^2} \tag{7.5}$$

for initial data (1.5). Then we have:

Proposition 7.3. *There exist $C, \gamma, \delta > 0$ and a polynomial P with $P(0) = 0$, all independent of ε, such that the estimates*

$$X_T^{\varepsilon,m+1} \le C X_0^\varepsilon + \varepsilon^\gamma P_T^{\varepsilon,m}, \tag{7.6a}$$

$$Y_T^{\varepsilon,m+1} \le C Y_0^\varepsilon + C T^\delta \left[Z_T^{\varepsilon,m} \right]^2 Y_T^{\varepsilon,m} + \varepsilon^\gamma P_T^{\varepsilon,m}, \tag{7.6b}$$

$$Z_T^{\varepsilon,m+1} \le C Z_0^\varepsilon + C T^\delta \left[Z_T^{\varepsilon,m} \right]^2 Z_T^{\varepsilon,m} + \varepsilon P_T^{\varepsilon,m}, \tag{7.6c}$$

hold for $0 \le T \le 1$ and $m \ge -1$, where

$$P_T^{\varepsilon,m} := \begin{cases} P(X_0^\varepsilon + Y_0^\varepsilon) & \text{for} \quad m = -1, \\ P\left(X_T^{\varepsilon,m} + X_T^{\varepsilon,m-1} + Y_T^{\varepsilon,m} + Y_T^{\varepsilon,m-1} \right) & \text{for} \quad m \ge 0. \end{cases} \tag{7.7}$$

In fact, these estimates hold for [recall (5.2)]

$$\gamma \le \Lambda + 1 - 2\theta. \tag{7.8}$$

The proof is deferred to the end of this section.

Corollary 7.4. *There exist $C, \delta > 0$ and a polynomial Q, all independent of ε, such that if $\gamma > 0$ is sufficiently small depending on Λ, and if $T, \varepsilon > 0$ are taken so small that*

$$2CT^\delta \left[2C \left\| \psi_0^\varepsilon \right\|_{L^2} + 1 \right]^2 \le 1, \qquad 2\varepsilon^{\gamma/2} Q(X_0^\varepsilon + Y_0^\varepsilon) \le 1, \tag{7.9}$$

then

$$X_T^{\varepsilon,m} \le 2C X_0^\varepsilon + \varepsilon^{\gamma/2}(X_0^\varepsilon + Y_0^\varepsilon), \tag{7.10a}$$

$$Y_T^{\varepsilon,m} \le 2C Y_0^\varepsilon + \varepsilon^{\gamma/2}(X_0^\varepsilon + Y_0^\varepsilon), \tag{7.10b}$$

$$Z_T^{\varepsilon,m} \le 2C Z_0^\varepsilon + \varepsilon^{1-\gamma/2}(X_0^\varepsilon + Y_0^\varepsilon), \tag{7.10c}$$

for $m \ge 0$.

Proof. This is a simple induction. Since $P(0) = 0$ in Proposition 7.3, there is a polynomial $Q(r)$ such that $P(4[C + 1]r) \le r Q(r)$ for $r \ge 0$. Then

$$P_T^{\varepsilon,m} \le Q(X_0^\varepsilon + Y_0^\varepsilon) \cdot (X_0^\varepsilon + Y_0^\varepsilon) \tag{7.11}$$

holds for $m = -1$, in view of the definition (7.7). Hence (7.10) for $m = 0$ follows from (7.6a)–(7.6c) and the fact that the iterates at $m = -1$ all vanish. Now assume (7.10) holds for $0 \le m \le m_0$. Then (7.11) holds for such m, and using (7.9) and (7.6) we obtain (7.10) for $m = m_0 + 1$. □

We are now in a position to prove Theorem 7.1. Indeed, from the proof of Theorem 6.1 we know that the iterates $\phi_\pm^{\varepsilon,m}$ converge in the Y-norms to $\phi_\pm^\varepsilon$. We can therefore pass to the limit $m \to \infty$ in Corollary 7.4. Thus, from (7.10a) and (7.10b) we get (7.2), since $\mathcal{I}^\varepsilon(0) = X_0^\varepsilon + Y_0^\varepsilon$ and $\mathcal{I}^\varepsilon(t) \lesssim X_T^\varepsilon + Y_T^\varepsilon$ for $0 \le t \le T$, by (4.6). Further, from (7.10c) we get

$$Z_T^\varepsilon(\phi_\pm^\varepsilon) \le 2C \left\| \psi_0^\varepsilon \right\|_{L^2} + 1. \tag{7.12}$$

Substituting this in the second term in the right hand side of (7.6c) in the limit $m \to \infty$, we then obtain (7.3). This proves Theorem 7.1.

Proof of Proposition 7.3. We claim that [cf. (6.5)]

$$\left\| e^{\pm it/\varepsilon^2} R^{\varepsilon,m} \right\|_{X^{1,\theta-1}_{\tau=\pm h_\varepsilon(\xi)}(S_T)} \lesssim \varepsilon^{(1/2-\Lambda)^-} P_T^{\varepsilon,m} \tag{7.13}$$

$$\left\| R^{\varepsilon,m} \right\|_{L^2(S_T)} \lesssim \varepsilon P_T^{\varepsilon,m} \tag{7.14}$$

For the bilinear terms (6.6), this follows from Theorem 5.4 and estimate (5.5) in Lemma 5.1; for the cubic terms in (6.7), we proceed exactly as in the proof of Theorem 6.1. Finally, consider the commutator term in (6.5); for (7.14) we apply Lemma 5.2 followed by (5.11) and (5.4); for (7.13), we apply (6.15) followed by Lemma 5.2 (but with L_t^2 in (5.16) replaced by $L_t^{2^-}$), then (5.11) and finally (5.9). This proves (7.13) and (7.14).

To prove (7.6a), we write $\mathbf{A}^{\varepsilon,m+1} = \mathbf{A}^{\varepsilon,0} + \square_\varepsilon^{-1} \varepsilon \mathbb{P} \mathbf{J}^{\varepsilon,m}$; the first term in the right hand side is covered by Lemma 4.4; to handle the other term, we shall decompose $\mathbf{J}^{\varepsilon,m}$ suitably and apply the estimates in Lemma 5.5. In fact, we decompose it as in (2.16), using (6.3); then we further decompose each of the χ's and η's into their low and high frequency parts, as in (1.29). For any term in this expansion containing at least one high frequency factor, we use (5.28) in Lemma 5.5. The remaining terms are all low-low interactions; those among them that contain χ_- or η_+ can be estimated using (5.27) and

$$\left\| \chi_-^{\varepsilon,m} \right\|_{L_x^2} + \left\| \eta_+^{\varepsilon,m} \right\|_{L_x^2} \lesssim \varepsilon \left\| \psi^{\varepsilon,m} \right\|_{H^1} + \varepsilon^2 \left\| \mathbf{A}^{\varepsilon,m-1} \right\|_{\dot{H}^1} \left\| \psi^{\varepsilon,m-1} \right\|_{H^1},$$

which is simply the analogue of (2.6) for iterates. This leaves us with a single low-low term, corresponding to the very last term in (2.16), and this can be estimated using (5.29) in Lemma 5.5, provided we can estimate the terms in the right hand side of (6.4) in $L^2(S_T)$; but this is easy, using Lemma (5.1) and (7.14). This proves (7.6a).

Next, applying Lemma 4.3 to the equation (6.4) and using the embedding (4.7) and the estimate (7.13), as well as (2.5) at $t = 0$, we reduce (7.6b) to

$$\left\| A_0^{\varepsilon,m} \phi_\pm^{\varepsilon,m} \right\|_{L_t^{2^-} H^1(S_T)} \lesssim T^\delta \left[Z_T^{\varepsilon,m} \right]^2 Y_T^{\varepsilon,m} + \varepsilon^{1^-} \left[Y_T^{\varepsilon,m} \right]^3. \tag{7.15}$$

But this follows from (5.13), (5.8) and (5.9). The factor T^δ comes from applying Hölder's inequality in time.

Now consider (7.6c). By Proposition 3.5 applied to (6.4), and (2.5) at $t = 0$,

$$Z_T^{\varepsilon,m+1} \lesssim Z_0^\varepsilon + \varepsilon^{2-\Lambda} X_T^{\varepsilon,m} Y_T^{\varepsilon,m} + \sum_\pm \left\| A_0^{\varepsilon,m} \phi_\pm^{\varepsilon,m} \right\|_{L_t^{1^+} L_x^{2^-}(S_T)} + \left\| R^{\varepsilon,m} \right\|_{L_t^1 L_x^2(S_T)}.$$

The last term is covered by (7.14). By (5.6) and (5.7),

$$\left\| A_0^{\varepsilon,m} \phi_\pm^{\varepsilon,m} \right\|_{L_t^{1^+} L_x^{2^-}(S_T)} \lesssim T^\delta \left[Z_T^{\varepsilon,m} \right]^2 \left\| \phi_\pm^{\varepsilon,m} \right\|_{L_t^\infty L_x^2(S_T)} + \varepsilon \left[Y_T^{\varepsilon,m} \right]^3.$$

Then (7.6c) follows, since $\left\| \phi_\pm^{\varepsilon,m} \right\|_{L_t^\infty L_x^2(S_T)} \lesssim Z_T^{\varepsilon,m} + \varepsilon Y_T^{\varepsilon,m}$, by (4.5). ☐

8. HIGHER ORDER BOUNDS

The local well-posedness of DM was established in Sect. 6, and a standard argument shows that higher regularity persists, i.e., if the norm

$$\mathcal{I}_m^\varepsilon(t) = \left\| \psi^\varepsilon(t) \right\|_{H^{m+1}} + \varepsilon^\Lambda \left(\left\| \nabla \mathbf{A}^\varepsilon(t) \right\|_{H^m} + \varepsilon \left\| \partial_t \mathbf{A}^\varepsilon(t) \right\|_{H^m} \right)$$

is initially finite, then it stays finite in the interval of existence $0 \le T \le T_\varepsilon$, for fixed ε. Here we concentrate on proving bounds which are uniform in ε.

Proposition 8.1. *If $\mathcal{I}_m^\varepsilon(0) = O(1)$, then $\mathcal{I}_m^\varepsilon(t) = O(1)$ in every finite time interval.*

More precisely, we prove the following. Given $T < \infty$, let N be the smallest natural number such that $NT^* \geq T$, with T^* as in Theorem 7.1. Assume that

$$\mathcal{I}_m^\varepsilon(0) \leq B \quad \text{for all} \quad \varepsilon. \tag{8.1}$$

Then $\mathcal{I}_m^\varepsilon(t) \leq P_m\big(C^N B\big)$ for all $0 \leq t \leq T$, provided ε is sufficiently small [to be precise, provided it satisfies the last inequality in (7.4)]. Here C is a universal constant and P_m is a polynomial depending only on m.

Proof. Step 1. We introduce norms

$$\begin{aligned} X_{T,k}^\varepsilon &= \varepsilon^\Lambda \sum_{|\alpha|\leq k} X_T^\varepsilon\big(\partial_x^\alpha \mathbf{A}^\varepsilon\big), \qquad X_{0,k}^\varepsilon = \varepsilon^\Lambda\left(\left\|\nabla \mathbf{a}_0^\varepsilon\right\|_{H^k} + \varepsilon\left\|\mathbf{a}_1^\varepsilon\right\|_{H^k}\right), \\ Y_{T,k}^\varepsilon &= \sum_{|\alpha|\leq k} \sum_{\pm} Y_{T,\pm}^\varepsilon\big(\partial_x^\alpha \phi_\pm^\varepsilon\big), \quad Y_{0,k}^\varepsilon = \left\|\psi_0^\varepsilon\right\|_{H^{k+1}}, \end{aligned}$$

and claim there exist $C, \delta, \gamma > 0$ and polynomials Q_k—all independent of ε—such that for $0 \leq T \leq 1$ and $k \geq 1$

$$\begin{aligned} X_{T,k}^\varepsilon &\leq C X_{0,k}^\varepsilon + \varepsilon^\gamma Q_0\left(X_0^\varepsilon + Y_0^\varepsilon\right) \cdot \left\{X_{T,k}^\varepsilon + Y_{T,m}^\varepsilon\right\} \\ &\quad + Q_k\big(X_{T,k-1}^\varepsilon + Y_{T,k-1}^\varepsilon\big), \\ Y_{T,k}^\varepsilon &\leq C Y_{0,k}^\varepsilon + C T^\delta \left\{1 + 2C\left\|\psi_0^\varepsilon\right\|_{L^2}\right\}^2 Y_{T,k}^\varepsilon \\ &\quad + \varepsilon^\gamma Q_0\left(X_0^\varepsilon + Y_0^\varepsilon\right) \cdot \left\{X_{T,k}^\varepsilon + Y_{T,k}^\varepsilon\right\} + Q_k\big(X_{T,k-1}^\varepsilon + Y_{T,k-1}^\varepsilon\big). \end{aligned}$$

To prove this, we apply $\sum_{|\alpha|\leq k} \partial_x^\alpha$ to the system, and imitate the proof of the estimates in Proposition 7.3 (which correspond to $k = 0$). We single out the top order terms where k derivatives fall on one of the fields $\mathbf{A}$, ψ or $\phi_\pm$; these are estimated exactly like in the case $k = 0$. All other terms are lumped together and yield the term $Q_k\big(X_{T,k-1}^\varepsilon + Y_{T,k-1}^\varepsilon\big)$. We skip the straightforward but tedious details of this argument.

Step 2. Let N be the smallest natural number such that $NT^* \geq T$, where T^* is as in Theorem 7.1. For $n = 1, \ldots, N$ let $X_{n,k}^\varepsilon$ and $Y_{n,k}^\varepsilon$ be the norms defined as in Step 1, but taken over the time interval $[(n-1)T^*, nT^*]$. Now set

$$\omega_{n,k} = \begin{cases} 0 & \text{if } k = -1, \\ X_{n,k}^\varepsilon + Y_{n,k}^\varepsilon & \text{if } 0 \leq n \leq N \text{ and } 0 \leq k \leq m. \end{cases}$$

Then by the estimates in Step 1, and since T^* was chosen so that (7.9) holds,

$$\omega_{n,k} \leq C\omega_{n-1,k} + Q_k\big(\omega_{n,k-1}\big) \quad \text{for } 1 \leq n \leq N \text{ and } 0 \leq k \leq m.$$

Induction on n gives, using (8.1), $\gamma_k \leq C^N B + C^N Q_k(\gamma_{k-1})$ for $0 \leq k \leq m$, where $\gamma_k = \max_{0\leq n\leq N} \omega_{n,k}$. Induction on k, using the fact that $\gamma_{-1} = 0$, then yields $\gamma_k \leq P_k(C^N B)$ for $0 \leq k \leq m$. □

9. ESTIMATES FOR THE SMALL COMPONENT

In this section we prove that if the "positron part" $\Pi_-(\varepsilon D)\psi^\varepsilon$ is small initially, then it stays small uniformly in every finite time interval, where "small" means either $O(\varepsilon)$ or $O(\varepsilon^2)$.

Theorem 9.1. (i) *Assume (8.1) holds for some $m \geq 0$, and that at $t = 0$,*

$$\left\|\Pi_-(\varepsilon D)\psi^\varepsilon\right\|_{H^m} = O(\varepsilon). \tag{9.1}$$

Then the same estimate holds uniformly in every finite time interval.

Moreover, if we strengthen (8.1) by assuming that at $t = 0$*,*

$$\left\| \nabla \mathbf{A}^\varepsilon \right\|_{H^m} + \varepsilon \left\| \partial_t \mathbf{A}^\varepsilon \right\|_{H^m} = O(1), \tag{9.2}$$

then the same estimate holds uniformly in every finite time interval.

(ii) *Assume that (8.1), (9.1) and (9.2) hold for some* $m \geq 1$ *(by part (i) it suffices to assume they hold initially). Then if, at* $t = 0$*,*

$$\left\| \Pi_-(\varepsilon D)\psi^\varepsilon \right\|_{H^{m-1}} = O(\varepsilon^2), \tag{9.3}$$

the same estimate holds uniformly in every finite time interval.

Let us interpret this result in terms of η^ε, the lower component of $e^{it/\varepsilon^2}\psi^\varepsilon$, as in (1.16). Firstly, (9.1) is equivalent to

$$\left\| \eta^\varepsilon \right\|_{H^m} = O(\varepsilon). \tag{9.4}$$

This follows from (1.22), since $\|\psi^\varepsilon\|_{H^{m+1}} = O(1)$ on account of Proposition 8.1. Secondly, (9.3) implies

$$\left\| \partial_t \eta^\varepsilon \right\|_{H^{m-1}} = O(1), \tag{9.5}$$

while the converse implication holds if we also assume (9.1) [i.e., (9.4)]. To see this, note that by (1.23), (9.3) is equivalent to

$$\sigma^k \partial_k \eta^\varepsilon = O(\varepsilon), \qquad \eta^\varepsilon + i\varepsilon\frac{1}{2}\sigma^k\partial_k\chi^\varepsilon = O(\varepsilon^2) \qquad \text{in} \quad H^{m-1} \tag{9.6}$$

where χ^ε is the upper component of $e^{it/\varepsilon^2}\psi^\varepsilon$, as in (1.16). But by the second equation in (1.18),

$$i\varepsilon^2\partial_t\eta^\varepsilon = \eta^\varepsilon + i\varepsilon\frac{1}{2}\sigma^k\partial_k\chi^\varepsilon + O(\varepsilon^2) \qquad \text{in} \quad H^{m-1} \tag{9.7}$$

where we used (9.2).

Proof of Theorem 9.1(i). Step 1. Set

$$\tilde{Z}^\varepsilon_{T,k} = \sum_{|\alpha|\leq k} \left\| \partial_x^\alpha (\phi^\varepsilon_-)_{\text{low}} \right\|_{L_t^2 L_x^6 \cap L_t^\infty L_x^2(S_T)}.$$

There exist $C, \delta > 0$ and a polynomial P, independent of ε, such that for $0 \leq T \leq 1$ and $0 \leq k \leq m$,

$$\tilde{Z}^\varepsilon_{T,k} \leq C\tilde{Z}^\varepsilon_{0,k} + CT^\delta \left\{ C \left\| \psi^\varepsilon_0 \right\|_{L^2} + 1 \right\}^2 \tilde{Z}^\varepsilon_{T,k} + \left\{ \varepsilon + \tilde{Z}^\varepsilon_{T,k-1} \right\} P\left(X^\varepsilon_{T,k} + Y^\varepsilon_{T,k} \right),$$

where by convention $\tilde{Z}^\varepsilon_{T,k} = 0$ for $k = -1$, and where $X^\varepsilon_{T,k}, Y^\varepsilon_{T,k}$ are as in Step 1 of the proof of Proposition 8.1. This estimate follows by a straightforward modification of the proof of the estimate for the Z-norm in Proposition 7.3, taking into account the bound (7.12).

Step 2. Fix a time $T < \infty$, and let N be the smallest natural number such that $NT^* \geq T$, where T^* is as in Theorem 7.1. For $n = 1, \dots, N$ let $\tilde{Z}^\varepsilon_{n,k}$ be the norm defined as in Step 1, but taken over the time interval $[(n-1)T^*, nT^*]$. Now set

$$\theta_{n,k} = \begin{cases} 0 & \text{if } k = -1, \\ \tilde{Z}^\varepsilon_{n,k} & \text{if } 0 \leq n \leq N \text{ and } 0 \leq k \leq m. \end{cases}$$

Then by the estimate in Step 1, and (the proof of) Proposition 8.1, we have

$$\theta_{n,k} \leq C\theta_{n-1,k} + \left(\varepsilon + \theta_{n,k-1}\right) P(C^N B) \quad \text{for } 1 \leq n \leq N \text{ and } 0 \leq k \leq m,$$

provided ε is sufficiently small [to be precise, provided it satisfies the last inequality in (7.4)]. In what follows, C and P may change from line to line. Induction on n gives

$$\eta_k \leq C^N \left\| (\phi_-^\varepsilon)_{\mathrm{low}}(t=0) \right\|_{H^k} + \left(\varepsilon + \eta_{k-1}\right) P(C^N B) \quad \text{for } 0 \leq k \leq m,$$

where $\eta_k = \max_{0 \leq n \leq N} \theta_{n,k}$. Induction on k then yields

$$\left\| (\phi_-^\varepsilon)_{\mathrm{low}}(t) \right\|_{H^m} \leq \left(\varepsilon + \left\| (\phi_-^\varepsilon)_{\mathrm{low}}(t=0) \right\|_{H^m} \right) P(C^N B) \tag{9.8}$$

for $0 \leq t \leq T$, provided ε satisfies the last inequality in (7.4). In view of (2.2)–(2.4) and (4.5), we can replace $(\phi_-^\varepsilon)_{\mathrm{low}}$ by $\Pi_-(\varepsilon D)\psi^\varepsilon$. This proves (9.1) in the interval $[0, T]$.

Step 3. We prove (9.2), assuming it holds initially. In view of Lemma 4.4, this reduces to proving $\|\mathbf{J}^\varepsilon\|_{L_t^2 H^m(S_T)} = O(1)$, where $\varepsilon \mathbf{J}^\varepsilon = 2 \operatorname{Re} \langle \vec{\sigma} \chi^\varepsilon, \eta^\varepsilon \rangle$. If χ^ε is replaced by $\chi_{\mathrm{high}}^\varepsilon$, we can proceed as in the proof of (5.28) in Lemma 5.5; if χ^ε is replaced by $\chi_{\mathrm{low}}^\varepsilon$, we write

$$\left\| \partial_x^\alpha \mathbf{J}^\varepsilon \right\|_{L^2} \leq \frac{1}{\varepsilon} \sum_{\beta+\gamma=\alpha} c_{\alpha\beta} \left\| \partial_x^\beta \chi_{\mathrm{low}}^\varepsilon \right\|_{L_t^2 L_x^\infty} \left\| \partial_x^\gamma \eta^\varepsilon \right\|_{L_t^\infty L_x^2}, \tag{9.9}$$

for $|\alpha| \leq m$, and use (4.4), (the proof of) Proposition 8.1 and (9.4). This gives, with notation as in Step 2,

$$\left\| \nabla \mathbf{A}^\varepsilon(t) \right\|_{H^m} + \varepsilon \left\| \partial_t \mathbf{A}^\varepsilon(t) \right\|_{H^m} \leq P(C^N B) \tag{9.10}$$

for $0 \leq t \leq T$, provided ε satisfies the last inequality in (7.4). □

Proof of Theorem 9.1(ii). In this proof, all estimates are understood to hold uniformly in $0 \leq t \leq T$, where $T = NT^*$ and T^* is as in Theorem 7.1. Choose B so large that, in addition to (8.1), we have initial bounds εB, B and $\varepsilon^2 B$ in (9.1), (9.2) and (9.3), respectively. In what follows, C (a universal constant) and P (a polynomial) may change from line to line.

In view of the hypotheses, we have by Proposition 8.1 and Theorem 9.1(i),

$$\left\| \psi^\varepsilon(t) \right\|_{H^{m+1}} + \left\| \nabla \mathbf{A}^\varepsilon(t) \right\|_{H^m} + \varepsilon \left\| \partial_t \mathbf{A}^\varepsilon(t) \right\|_{H^m} \leq P(C^N B), \tag{9.11}$$

as well as (9.8), provided ε satisfies the last inequality in (7.4), which we assume henceforth. Then by (2.2),

$$\left\| \psi_\pm^\varepsilon - \Pi_\pm(\varepsilon D)\psi^\varepsilon \right\|_{H^{m-1}} \leq \varepsilon^2 P(C^N B), \tag{9.12}$$

where $\psi_\pm^\varepsilon$ are given by (2.1); thus, it suffices to prove $\psi_-^\varepsilon = O(\varepsilon^2)$ in H^{m-1}. Setting[6] $\phi_\pm^\varepsilon = e^{it/\varepsilon^2} \psi_\pm^\varepsilon$, we have by an obvious modification of Lemma 2.1,

$$i\partial_t \phi_+^\varepsilon - \frac{\lambda(\varepsilon D) - 1}{\varepsilon^2} \phi_+^\varepsilon = -A_0^\varepsilon \phi_+^\varepsilon + \frac{1}{2} e^{it/\varepsilon^2} R^\varepsilon, \tag{9.13}$$

$$i\partial_t \phi_-^\varepsilon + \frac{\lambda(\varepsilon D) + 1}{\varepsilon^2} \phi_-^\varepsilon = -A_0^\varepsilon \phi_-^\varepsilon + \frac{1}{2} e^{it/\varepsilon^2} R^\varepsilon, \tag{9.14}$$

with R^ε given by

$$\begin{aligned} \lambda(\varepsilon D) R^\varepsilon = \varepsilon &\left\{ 2i\mathbf{A}^\varepsilon \cdot \nabla + iE_k^\varepsilon \alpha^k - B_k^\varepsilon S^k \right\} \left(\psi_+^\varepsilon + \psi_+^\varepsilon \right) \\ &+ \varepsilon^2 (\mathbf{A}^\varepsilon)^2 \left(\psi_+^\varepsilon + \psi_+^\varepsilon \right) - \varepsilon^2 \left[A_0^\varepsilon, \frac{\lambda(\varepsilon D) - 1}{\varepsilon^2} \right] (\psi_+^\varepsilon - \psi_-^\varepsilon). \end{aligned} \tag{9.15}$$

---

[6] This definition breaks with (2.7), and must be considered local to this proof.

Note also that (9.12) and (9.11) imply

$$\left\|\phi^\varepsilon_\pm\right\|_{H^{m+1}} = \left\|\psi^\varepsilon_\pm\right\|_{H^{m+1}} \le P(C^N B), \tag{9.16}$$

since the $\Pi^\varepsilon_\pm$ are uniformly bounded on H^s. From this and (9.11) it is easy to show that the terms in the right hand sides of (9.13) and (9.14) are $O(1)$ [bounded by $P(C^N B)$, in fact] in H^m; for the terms $\left\|A^\varepsilon_0\phi^\varepsilon_\pm\right\|_{H^m}$, this only requires Leibniz' rule and Hölder's inequality in combination with various Sobolev embeddings; ditto for $\|R^\varepsilon\|_{H^m}$, but then we first use (2.3) with $r = 1$. We omit the details.

To estimate the second term in the left hand side of (9.13), we use (5.17) and (9.16); similarly for the corresponding term in (9.14), but there we write $\frac{\lambda(\varepsilon D)+1}{\varepsilon^2} = \frac{\lambda(\varepsilon D)-1}{\varepsilon^2} + \frac{1}{\varepsilon^2}$ and use also (9.8). Finally, then, we conclude that

$$\left\|\partial_t\phi^\varepsilon_\pm\right\|_{H^m} \le \frac{1}{\varepsilon}P(C^N B). \tag{9.17}$$

From (9.14), $\phi^\varepsilon_- = \varepsilon^2[\lambda(\varepsilon D) + 1]^{-1}\left(-i\partial_t\phi^\varepsilon_- - A^\varepsilon_0\phi^\varepsilon_- + \frac{1}{2}e^{it/\varepsilon^2}R^\varepsilon\right)$, so to prove $\left\|\phi^\varepsilon_-\right\|_{H^{m-1}} \le \varepsilon^2 P(C^N B)$ it suffices to show $\left\|\partial_t\phi^\varepsilon_-\right\|_{H^{m-1}} \le P(C^N B)$. We first show this for $m = 1$. The higher order cases then follow by induction.

Apply a time derivative to (9.14), take the $\mathbb{C}^4$ inner product with $\partial_t\phi^\varepsilon_-$, then take imaginary parts, and finally integrate in x, making use of the self-adjointness of $\frac{\lambda(\varepsilon D)+1}{\varepsilon^2}$ and A^ε_0 (as a multiplication operator); the result is that

$$\left\|\partial_t\phi^\varepsilon_-(t)\right\|_{L^2} \le \left\|\partial_t\phi^\varepsilon_-(t=0)\right\|_{L^2} + \int_0^t \left(\left\|[\partial_t A^\varepsilon_0]\phi^\varepsilon_-\right\|_{L^2} + \left\|\partial_t\widetilde{R}^\varepsilon\right\|_{L^2}\right)\,dt',$$

where $\widetilde{R}^\varepsilon = \frac{1}{2}e^{it/\varepsilon^2}R^\varepsilon$. The first term on the right hand side is bounded by $P(C^N B)$; this follows from (9.14) [and (9.12)], since (9.3) is assumed to hold initially. Thus, it suffices to prove that the two terms inside the integral are bounded by $P(C^N B)$.

For the first term inside the integral, we use the fact that $\partial_t A^\varepsilon_0$ enjoys the same bounds as $c\nabla A^\varepsilon_0$, since $\Delta\partial_t A^\varepsilon_0 = -\operatorname{div}\mathbf{J}^\varepsilon$ by (1.4b) and the conservation law $\partial_t\rho^\varepsilon + \operatorname{div}\mathbf{J}^\varepsilon = 0$. Using (9.8) we then get $\left\|[\partial_t A^\varepsilon_0]\phi^\varepsilon_-\right\|_{L^2} \le P(C^N B)$.

It remains to prove that $\left\|\partial_t\widetilde{R}^\varepsilon\right\|_{L^2} \le P(C^N B)$. The key observation is that

$$\begin{aligned}\lambda(\varepsilon D)\widetilde{R}^\varepsilon = \varepsilon\left\{2i\mathbf{A}^\varepsilon\cdot\nabla + iE^\varepsilon_j\alpha^j - B^\varepsilon_j S^j\right\}\left(\phi^\varepsilon_+ + \phi^\varepsilon_+\right)& \\ + \varepsilon^2(\mathbf{A}^\varepsilon)^2\left(\phi^\varepsilon_+ + \phi^\varepsilon_+\right) - \varepsilon^2\left[A^\varepsilon_0, \frac{\lambda(\varepsilon D)-1}{\varepsilon^2}\right](\phi^\varepsilon_+ - \phi^\varepsilon_-)&.\end{aligned}$$

(Recall that $\phi^\varepsilon_\pm = e^{it/\varepsilon^2}\psi^\varepsilon_\pm$ in this proof; if we had used the symmetric definition (2.7), we would not have got rid of the phase factor e^{it/ε^2}.) Thus, $\left\|\partial_t\widetilde{R}^\varepsilon\right\|_{L^2}$ is dominated by a sum of terms; using the estimates which are now at our disposal (for the moment we assume $m = 1$), namely, (9.2), (9.16) and (9.17), it is an easy matter to show that all these terms are bounded by $P(C^N B)$. The details are left to the interested reader, but we point out that when the time derivative falls on the electric potential E^ε_j, we get terms involving $\partial_t^2\mathbf{A}^\varepsilon$; then one must use the wave equation satisfied by $\mathbf{A}^\varepsilon$. Recall also that $\partial_t A^\varepsilon_0$ enjoys the same bounds as $c\nabla A^\varepsilon_0$, as remarked above. This concludes the case $m = 1$. To handle $m \ge 2$, one proceeds by induction. Thus, we apply $\partial_t\partial_x^\alpha$, where $|\alpha| \le m-1$, to the equation (9.14), and we take the imaginary part of its inner product with $\partial_t\partial_x^\alpha\phi^\varepsilon_-$ and integrate in x. Then, if one takes into account the estimates proved in the previous induction steps, it is easy to modify the argument we gave for $m = 1$. We omit the easy details of this argument. $\square$

10. NONRELATIVISTIC LIMIT

We first prove Theorem 1.2, then we discuss the modifications needed to prove Theorem 1.3.

Proof of (1.12a). This can be restated:

$$e^{it/\varepsilon^2}\Pi_+^0\psi^\varepsilon \to \begin{pmatrix} v_+ \\ 0 \end{pmatrix}, \quad e^{-it/\varepsilon^2}\Pi_-^0\psi^\varepsilon \to v_- = \begin{pmatrix} 0 \\ v_- \end{pmatrix} \quad \text{in } H^1 \text{ as } \varepsilon \to 0 \tag{10.1}$$

locally uniformly in time. We claim it suffices to prove

$$e^{it/\varepsilon^2}\Pi_+(\varepsilon D)\psi^\varepsilon \to \begin{pmatrix} v_+ \\ 0 \end{pmatrix}, \quad e^{-it/\varepsilon^2}\Pi_-(\varepsilon D)\psi^\varepsilon \to v_- = \begin{pmatrix} 0 \\ v_- \end{pmatrix} \quad \text{in } H^1 \text{ as } \varepsilon \to 0. \tag{10.2}$$

To prove the claim, write $\Pi_\pm(\varepsilon D)\psi^\varepsilon = \Pi_\pm^0\psi^\varepsilon \pm r^\varepsilon$. By the orthogonality between Π_+^0 and Π_-^0, we get $\Pi_+^0 r^\varepsilon = -\Pi_+^0\Pi_-(\varepsilon D)\psi^\varepsilon$ and $\Pi_-^0 r^\varepsilon = \Pi_-^0\Pi_+(\varepsilon D)\psi^\varepsilon$; but if (10.2) holds, then the right hand sides converge to zero in H^1. Thus $r^\varepsilon = o(1)$ in H^1 and we have proved that (10.2) implies (10.1). In the remainder of the proof we drop the superscript ε on the fields, to simplify the notation.

Using (2.2) and (2.4) we reduce (10.2) to proving

$$\phi_+ \longrightarrow \begin{pmatrix} v_+ \\ 0 \end{pmatrix}, \quad \phi_- \longrightarrow \begin{pmatrix} 0 \\ v_- \end{pmatrix} \quad \text{in} \quad H^1 \quad \text{as} \quad \varepsilon \longrightarrow 0, \tag{10.3}$$

uniformly in any given time interval $[0, T]$. By (the proof of) Theorem 1.1, the solution exists in this time interval for all sufficiently small $\varepsilon > 0$, and

$$X_T^\varepsilon(\mathbf{A}^\varepsilon) + \sum_\pm Y_{T,\pm}^\varepsilon(\phi_\pm^\varepsilon) = O(1), \qquad \left\| e^{\pm it/\varepsilon^2} R^\varepsilon \right\|_{X^{1,\theta-1}_{\tau=\pm h_\varepsilon(\xi)}(S_T)} = o(1) \tag{10.4}$$

as $\varepsilon \to 0$. Note that (10.3) holds at time $t = 0$, by Lemma 1.6. Thus, it suffices to prove that there exist $K, \delta > 0$, depending on T and $X_T^\varepsilon(\mathbf{A}^\varepsilon) + \sum_\pm Y_{T,\pm}^\varepsilon(\phi_\pm^\varepsilon)$, but independent of ε, such that for every time interval $I = [t_0, t_1] \subset [0, T]$,

$$f(I) \le K f(\{t_0\}) + K\,|I|^\delta\, f(I) + o(1) \tag{10.5}$$

as $\varepsilon \to 0$, where

$$f(I) = \left\| \phi_+ - \begin{pmatrix} v_+ \\ 0 \end{pmatrix} \right\|_{L_t^\infty H^1(I\times\mathbb{R}^3)} + \left\| \phi_- - \begin{pmatrix} 0 \\ v_- \end{pmatrix} \right\|_{L_t^\infty H^1(I\times\mathbb{R}^3)}. \tag{10.6}$$

W.l.o.g. we take $I = [0, T]$ and only estimate the first term in (10.6). Write

$$\begin{aligned} \phi_+(t) &= U^\varepsilon(t)\phi_0^+ + \int_0^t U^\varepsilon(t-s)\left[L_+^\varepsilon\phi_+(s)\right] ds, \\ v_+(t) &= S(t)v_0^+ - \int_0^t S(t-s)\left[(uv_+)(s)\right] ds, \end{aligned}$$

where ϕ_0^+, v_0^+ are the data of ϕ_+, v_+ and $U^\varepsilon(t)$, $S(t)$ are given by (4.1). Thus

$$\begin{aligned} \phi_+(t) - \begin{pmatrix} v_+ \\ 0 \end{pmatrix}(t) &= U^\varepsilon(t)\left[\phi_0^+ - \begin{pmatrix} v_0^+ \\ 0 \end{pmatrix}\right] + \left[U^\varepsilon(t) - S(t)\right]\begin{pmatrix} v_0^+ \\ 0 \end{pmatrix} \\ &\quad + \int_0^t U^\varepsilon(t-s)\left[\begin{pmatrix} uv_+ \\ 0 \end{pmatrix}(s) + L_+^\varepsilon \phi_+(s)\right] ds \\ &\quad + \int_0^t \left[S(t-s) - U^\varepsilon(t-s)\right]\begin{pmatrix} uv_+ \\ 0 \end{pmatrix}(s)\, ds \\ &= I_1 + I_2 + I_3 + I_4. \end{aligned} \tag{10.7}$$

Clearly, $\|I_1\|_{L_t^\infty H^1(S_T)} \lesssim \|\phi_0^+ - [v_0^+, 0]^T\|_{H^1}$, and as in [2, Sect. 5], $\|I_j\|_{L_t^\infty H^1(S_T)} = o(1)$ for $j = 2, 4$, by the dominated convergence theorem and the fact that

$$\|\nabla u\|_{L_x^3} + \|u\|_{L_x^\infty} \lesssim \|v_+\|_{H^1}^2 + \|v_-\|_{H^1}^2 < \infty \tag{10.8}$$

uniformly in every finite time interval. It remains to consider I_3. By Lemma 4.3 and the embeddings (4.6) and (4.7),

$$\|I_3\|_{L_t^\infty H^1(S_T)} \lesssim \left\| \begin{pmatrix} uv_+ \\ 0 \end{pmatrix} - A_0\phi_+ \right\|_{L_t^2 H^1(S_T)} + \left\| e^{\pm it/\varepsilon^2} R^\varepsilon \right\|_{X^{1,\theta-1}_{\tau=\pm h_\varepsilon(\xi)}(S_T)}. \tag{10.9}$$

The second term on the right hand side is $o(1)$ by (10.4), and the first term is bounded by

$$T^{1/2}\left(\left\| u\left\{ \begin{pmatrix} v_+ \\ 0 \end{pmatrix} - \phi_+ \right\} \right\|_{L_t^\infty H^1(S_T)} + \|(u - A_0)\phi_+\|_{L_t^\infty H^1(S_T)} \right).$$

But using Leibniz' rule, Hölder's inequality and Sobolev embedding, it is easy to see that the terms inside the parentheses are dominated by $Kf(I)$, where K depends on the size of $X_T^\varepsilon(\mathbf{A}^\varepsilon) + \sum_\pm Y_{T,\pm}^\varepsilon(\phi_\pm^\varepsilon)$ and (10.8). □

Proof of (1.12b) and (1.12c). Using Sobolev embedding we reduce (1.12b) to (1.12c). To prove the latter, observe that (10.3) implies

$$\chi_+ \longrightarrow v_+, \quad \chi_- \longrightarrow 0, \quad \eta_+ \longrightarrow 0, \quad \eta_- \longrightarrow v_- \quad \text{in } H^1 \text{ as } \varepsilon \longrightarrow 0, \tag{10.10}$$

locally uniformly in time. Thus (1.12c) follows immediately from (2.15) using Hölder's inequality and Sobolev embedding. □

Proof of (1.14). Multiply (2.16) by a C^1 compactly supported test function $G(t,x)$ and integrate in t, x. W.l.o.g. assume G is real-valued. The integrals corresponding to the last two terms in the right hand side of (2.16) are $O(\varepsilon)$ in absolute value. To see this, integrate by parts in time and use

$$\left| \int fg\, dx \right| \le \|f\|_{H^{-1}} \|g\|_{H^1} \tag{10.11}$$

and the bound, locally uniform in time,

$$\|\partial_t \phi_\pm\|_{H^{-1}} = O(1). \tag{10.12}$$

The latter is easily reduced to the uniform bounds from Theorem 7.1, using Lemma 2.1, Sobolev embedding and Hölder's inequality.

Next, fix $1 \le j \le 3$ and consider $I_\pm := \frac{2}{\varepsilon}\operatorname{Re}\int \left\langle \sigma^j \chi_\pm, \eta_\pm \right\rangle G\,dt\,dx$. In view of (10.10) and (2.6),

$$I_- = \frac{2}{\varepsilon}\operatorname{Re}\int \left\langle \sigma^j \chi_-, \upsilon_- \right\rangle G\,dt\,dx + o(1). \tag{10.13}$$

By (2.2),(2.4) and (10.4),

$$\frac{1}{\varepsilon}\begin{pmatrix}\chi_-\\ 0\end{pmatrix} = e^{-it/\varepsilon^2}\frac{1}{\varepsilon}\Pi_+^0\Pi_-(\varepsilon D)\psi + O(\varepsilon^{2-\Lambda}) \quad \text{in} \quad L_x^2 \tag{10.14}$$

locally uniformly in time. But by (1.24),

$$\frac{1}{\varepsilon}\Pi_+^0\Pi_-(\varepsilon D)\psi = \frac{i}{2}[\lambda(\varepsilon D)]^{-1}\begin{pmatrix}\sigma^k\partial_k\tilde{\eta}\\ 0\end{pmatrix} + \frac{1}{2\varepsilon}\left(1-[\lambda(\varepsilon D)]^{-1}\right)\begin{pmatrix}\tilde{\chi}\\ 0\end{pmatrix}.$$

In view of (1.12a) and the bound (1.25), it follows that

$$\begin{aligned} &e^{-it/\varepsilon^2}\frac{1}{\varepsilon}\Pi_+^0\Pi_-(\varepsilon D)\psi \\ &\quad = \frac{i}{2}[\lambda(\varepsilon D)]^{-1}\begin{pmatrix}\sigma^k\partial_k\upsilon_-\\ 0\end{pmatrix} + \frac{1}{2\varepsilon}\left(1-[\lambda(\varepsilon D)]^{-1}\right)\begin{pmatrix}e^{-2it/\varepsilon^2}\upsilon_+\\ 0\end{pmatrix} + o(1)\end{aligned} \tag{10.15}$$

in L_x^2. Moreover, by dominated convergence,

$$[\lambda(\varepsilon D)]^{-1}\sigma^k\partial_k\upsilon_- = \sigma^k\partial_k\upsilon_- + o(1) \quad \text{in} \quad H^{-1}. \tag{10.16}$$

Using (10.13)–(10.16) and either Hölder's inequality or (10.11), we conclude that

$$I_- = \operatorname{Re}\int i\left\langle \sigma^j\sigma^k\partial_k\upsilon_-, \upsilon_-\right\rangle G\,dt\,dx + I'_- + o(1)$$

where $I'_- = \frac{1}{\varepsilon}\operatorname{Re}\int e^{-2it/\varepsilon^2}\left\langle \sigma^j\left(1-[\lambda(\varepsilon D)]^{-1}\right)\upsilon_+, \upsilon_-\right\rangle G\,dt\,dx$. But the latter is $O(\varepsilon)$ in absolute value (integrate by parts in time and use the analogue of (10.12) for $\upsilon_\pm$). Using (1.2) we finally conclude that

$$I_- = \int \left\{ -\operatorname{Im}\left\langle \partial_j\upsilon_-, \upsilon_-\right\rangle - \frac{1}{2}\epsilon^{jkl}\partial_k\left\langle \sigma_l\upsilon_-, \upsilon_-\right\rangle \right\} G\,dt\,dx + o(1).$$

A similar calculation can be done for I_+, and this proves (1.14). □

Next, we prove Theorem 1.3. By hypothesis, (9.4), or equivalently (9.1), holds initially and therefore also uniformly in every finite time interval, by Theorem 9.1. Next observe that since (1.15) holds initially, we have $\Pi_\pm(\varepsilon D)\psi^\varepsilon = \psi_\pm^\varepsilon + O(\varepsilon)$ in H^1 locally uniformly in time. In fact, this follows from (2.2)–(2.4), using Proposition 8.1 with $m = 1$. We conclude that it suffices to prove (1.15) with ψ^ε replaced by ψ_+^ε. We proceed as in the proof of Theorem 1.2, but now the remainder term in (10.5) must be improved from $o(1)$ to $O(\varepsilon)$, and $f(I)$ is given by the first term in the right hand side of (10.6). Again we reduce to estimating the terms $I_1, \dots, I_4$ as given by (10.7).

The term I_1 is estimated exactly as before, but is now $O(\varepsilon)$ since (1.15) is assumed to hold initially. Using the fact that $U^\varepsilon(t) - S(t) = \varepsilon^2\mathcal{R}_4^\varepsilon(t)$, where $\mathcal{R}_4^\varepsilon(t)$ is bounded from $H^{s+4} \to H^s$ uniformly in ε and $0 \le t \le T$, and the assumption that the initial datum of υ_+ is in H^5, we get $\left\|I_j\right\|_{L_t^\infty H^1(S_T)} = O(\varepsilon^2)$ for $j = 2, 4$. For I_3 we use again (10.9), but now the last term is $O(\varepsilon)$; this follows from (7.14), using Proposition 8.1 with $m = 1$. The first term in the right hand side of (10.9) is estimated exactly as before. This proves (1.15), and then it follows immediately that (1.12b) and (1.12c) are improved to $O(\varepsilon)$.

11. SEMI-NONRELATIVISTIC LIMIT

Here we prove Theorem 1.7. In view of the initial assumptions (i)–(iii), Proposition 8.1 applies with $m = 4$, while (9.1)–(9.5) hold with $m = 2$. We write

$$\phi^\varepsilon = \begin{pmatrix} \chi^\varepsilon \\ \eta^\varepsilon \end{pmatrix} := e^{it/\varepsilon^2} \psi^\varepsilon, \qquad \phi_\pm^\varepsilon := e^{it/\varepsilon^2} \psi_\pm^\varepsilon,$$

with $\psi_\pm^\varepsilon$ defined as in (2.1). Also, we denote by $\chi_\pm^\varepsilon$ the upper component of $\phi_\pm^\varepsilon$. Observe that $\Pi_\pm(\varepsilon D)\psi^\varepsilon = \psi_\pm^\varepsilon + O(\varepsilon^2)$ in H^1, in view of (2.2)–(2.4) and (9.11) for $m = 2$. On account of (9.3), we may therefore replace χ^ε in (1.27) by χ_+^ε. By Lemma 2.1,

$$\left(i\partial_t - \frac{\lambda(\varepsilon D) - 1}{\varepsilon^2} \right) \chi_+^\varepsilon + A_0^\varepsilon \chi_+^\varepsilon = \tilde{R}^\varepsilon,$$

where

$$\begin{aligned} \tilde{R}^\varepsilon &= \varepsilon i \mathbf{A}^\varepsilon \cdot \nabla \chi^\varepsilon - \frac{1}{2}\varepsilon B_j^\varepsilon \sigma^j \chi^\varepsilon + \frac{1}{2}\varepsilon^2 \left(\mathbf{A}^\varepsilon\right)^2 \chi^\varepsilon \\ &\quad - \frac{1}{2}\left(1 - [\lambda(\varepsilon D)]^{-1}\right) \varepsilon \left\{ 2i\mathbf{A}^\varepsilon \cdot \nabla \chi^\varepsilon - B_j^\varepsilon \sigma^j \chi^\varepsilon \right\} + \frac{1}{2}[\lambda(\varepsilon D)]^{-1} \varepsilon \left\{ iE_j^\varepsilon \sigma^j \eta^\varepsilon \right\} \\ &\quad - \frac{1}{2}\left(1 - [\lambda(\varepsilon D)]^{-1}\right) \varepsilon^2 \left\{ \left(\mathbf{A}^\varepsilon\right)^2 \chi^\varepsilon \right\} - \frac{1}{2}\varepsilon^2 [\lambda(\varepsilon D)]^{-1} \left[A_0^\varepsilon, \frac{\lambda(\varepsilon D) - 1}{\varepsilon^2} \right] \left(\chi_+^\varepsilon - \chi_-^\varepsilon \right). \end{aligned}$$

Recalling the bound (1.25) on the symbol of $1 - [\lambda(\varepsilon D)]^{-1}$, and using Proposition 8.1, (9.4) and (9.2) with $m = 2$, we conclude that

$$\tilde{R}^\varepsilon = \varepsilon i \mathbf{A}^\varepsilon \cdot \nabla \chi^\varepsilon - \frac{1}{2}\varepsilon B_j^\varepsilon \sigma^j \chi^\varepsilon + \frac{1}{2}\varepsilon^2 \left(\mathbf{A}^\varepsilon\right)^2 \chi^\varepsilon + O(\varepsilon^2) \quad \text{in} \quad H^1,$$

locally unformly in time.

Then, using Proposition 8.1 with $m = 4$ and noting that $\frac{\lambda(\varepsilon D) - 1}{\varepsilon^2} = \frac{\Delta}{2} + \varepsilon^2 \mathcal{R}_4^\varepsilon$, where $\mathcal{R}_4^\varepsilon$ is bounded from $H^{s+4} \to H^s$ uniformly in ε, we further conclude that

$$i\partial_t \chi_+^\varepsilon = \frac{1}{2}\left(i\nabla + \varepsilon \mathbf{A}^\varepsilon\right)^2 \chi_+^\varepsilon - A_0^\varepsilon \chi_+^\varepsilon - \frac{1}{2}\varepsilon B_j^\varepsilon \sigma^j \chi_+^\varepsilon + \varepsilon^2 r^\varepsilon \tag{11.1}$$

where $r^\varepsilon = O(1)$ in H^1 locally unformly in time. Comparing (11.1) to the Pauli equation (1.26) via the energy inequality for the self-adjoint "Pauli operator", $P^\varepsilon = \frac{1}{2}\left(i\nabla + \varepsilon \mathbf{A}^\varepsilon\right)^2 - \frac{1}{2}\varepsilon B_j^\varepsilon \sigma^j$, one finds that

$$f(I) \leq f(\{t_0\}) + K\,|I|\, f(I) + O(\varepsilon^2)$$

as $\varepsilon \to 0$, where $f(I) = \left\| \chi_+^\varepsilon - \chi_P^\varepsilon \right\|_{L_t^\infty H^1(I \times \mathbb{R}^3)}$ for time intervals $I = [t_0, t_1] \subset [0, T]$, and where K depends on T but not on ε. In fact, K depends on the $O(1)$ bound (9.11), which holds for $m = 2$, as we recall. We conclude that $f([0, T]) = O(\varepsilon^2)$, and this proves (1.27).

Observe that (1.20) holds in H^1 locally uniformly in time, in view of (1.19) and the fact that (9.4), (9.5) and (9.2) hold for $m = 2$. Substituting (1.20) into $\mathbf{J}^\varepsilon = \varepsilon^{-1} 2\, \mathrm{Re}\, \langle \vec{\sigma} \chi^\varepsilon, \eta^\varepsilon \rangle_{\mathbb{C}^2}$ and using (1.27) yields (1.28).

12. PROOFS OF THE SPACETIME ESTIMATES

Here we prove Theorem 3.1 and Proposition 3.4.

Proof of Proposition 3.4. Let Q be a cube with side length $\sim \mu$ centered at ξ_0, where $|\xi_0| \sim \lambda$, and let $\chi_Q(\xi)$ be a smooth cut-off function equal to 1 on Q. For example, we can take

$$\chi_Q(\xi) := \eta\left(\frac{\xi - \xi_0}{\mu}\right), \tag{12.1}$$

where η is a smooth bump function equal to 1 on a neighborhood of the origin. Then by the TT^* method, we reduce (3.7) to the decay estimate

$$\left|K_{\varepsilon,Q}(t,x)\right| \lesssim \begin{cases} \mu\,|t|^{-1} & \text{for } \lambda \lesssim 1/\varepsilon, \\ \varepsilon\mu\lambda\,|t|^{-1} & \text{for } \lambda \gg 1/\varepsilon, \end{cases} \tag{12.2}$$

for the convolution kernel $K_{\varepsilon,Q}(t,x) := \int_{\mathbb{R}^3} e^{ix\cdot\xi} e^{ith_\varepsilon(\xi)} \chi_Q(\xi)\,d\xi$, with h_ε given by (2.14). In view of the scaling identity $K_{\varepsilon,Q}(t,x) = \varepsilon^{-3} K_{1,\varepsilon Q}(\varepsilon^{-2}t, \varepsilon^{-1}x)$, it suffices to prove (12.2) for $\varepsilon = 1$. To simplify the notation we write K_Q instead of $K_{1,Q}$. Thus,

$$K_Q(t,x) = \int_0^\infty \int_{S^2} e^{irx\cdot\omega} e^{it\alpha(r)} \chi_Q(r\omega) r^2\,d\sigma(\omega)\,dr \tag{12.3}$$

where σ is surface measure on S^2 and α is given by (2.18). Note that

$$\alpha'(r) = \frac{r}{\sqrt{1+r^2}} \quad \text{and} \quad \alpha''(r) = \frac{1}{(1+r^2)^{3/2}}. \tag{12.4}$$

We split the problem into the following cases:

(i) $\lambda \lesssim 1$ and $|x| \gtrsim \lambda\,|t|$,
(ii) $\lambda \lesssim 1$ and $|x| \ll \lambda\,|t|$,
(iii) $\lambda \gg 1$ and $|x| \gtrsim |t|$,
(iv) $\lambda \gg 1$ and $|x| \ll |t|$.

Rewrite (12.3) as $K_Q(t,x) = \int_0^\infty e^{it\alpha(r)} a(r,x) r^2\,dr$ where

$$a(r,x) := \int_{S^2} e^{irx\cdot\omega} \chi_Q(r\omega)\,d\sigma(\omega).$$

We shall need the following:

Lemma 12.1. $|a(r,x)| \lesssim (r\,|x|)^{-1}\chi_I(r)$, *where* χ_I *is the characteristic function of an interval* I *of length* $\sim \mu$ *and centered at a distance* $\sim \lambda$ *from the origin.*

Proof. The statement about the r-support of $a(r,x)$ is obvious, and the decay statement follows from the fact that $\left|\int_{S^2} e^{ix\cdot\omega}\gamma(\omega)\,d\sigma(\omega)\right| \le 4\pi/|x|$ for any function such that $|\gamma(\omega)| \le 1$. To prove this fact, first note that by rotational invariance of $d\sigma$ we may assume $x = (0,0,|x|)$. Then passing to spherical coordinates $\omega(\theta,\phi) = (\cos\theta\sin\phi, \sin\theta\sin\phi, \cos\phi)$ we have

$$\int_{S^2} e^{ix\cdot\omega}\gamma(\omega)\,d\sigma(\omega) = \int_0^{2\pi}\int_0^\pi e^{i|x|\cos\phi}\gamma\left(\omega(\theta,\phi)\right)\sin\phi\,d\phi\,d\theta.$$

Changing variables $\phi \to u = |x|\cos\phi$ and taking absolute values then gives the desired estimate. □

Thus, $\left|K_Q(t,x)\right| \lesssim \int_I \left(r/|x|\right) dr \sim \mu\lambda/|x|$, and this covers cases (i) and (iii).

To handle the remaining cases we write (12.3) as $K_Q(t,x) = \int_{S^2} b(\omega)\,d\sigma(\omega)$, where

$$b(\omega) = \int_0^\infty \frac{d}{dr}\left[e^{i(t\alpha(r)+rx\cdot\omega)}\right] \frac{\chi_Q(r\omega)r^2}{i\left(t\alpha'(r) + x\cdot\omega\right)}\,dr.$$

Integrate by parts and write

$$-\frac{d}{dr}\left[\frac{\chi_Q(r\omega)r^2}{i\big(t\alpha'(r)+x\cdot\omega\big)}\right]=\frac{\chi_Q(r\omega)r^2t\alpha''(r)}{i\big(t\alpha'(r)+x\cdot\omega\big)^2}-\frac{\frac{d}{dr}\big[\chi_Q(r\omega)r^2\big]}{i\big(t\alpha'(r)+x\cdot\omega\big)}.$$

Correspondingly we split $b=b_1+b_2$. Observe that the r-support of $\chi_Q(r\omega)$ is contained in an interval I of length $\sim\mu$ and centered at a distance λ from the origin, while the ω-support is contained in a set given by

$$\angle(\omega,\omega_0)\lesssim\mu/\lambda \tag{12.5}$$

for some $\omega_0\in S^2$. Moreover, in view of (12.1) we have

$$\left|\tfrac{d}{dr}\chi_Q(r\omega)\right|\lesssim 1/\mu. \tag{12.6}$$

Now consider case (iv). Then on account of (12.4) we have $\alpha'(r)\sim 1$ and $\alpha''(r)\sim\lambda^{-3}$ for $r\in I$, so $\left|t\alpha'(r)+x\cdot\omega\right|\gtrsim|t|$. Thus

$$|b_1(\omega)|\lesssim\big(1/\lambda\,|t|\big)\int_I dr\lesssim\mu/\lambda\,|t|\,,$$

which is more than good enough. Next, using (12.6) we have

$$|b_2(\omega)|\lesssim\big(1/\,|t|\big)\int_I\Big(r+r^2/\mu\Big)\,dr\lesssim(\lambda\mu+\lambda^2)/\,|t|\lesssim\lambda^2/\,|t|\,.$$

But integrating this over the region (12.5) on S^2 gives us a bound $\mu^2/\,|t|$, which again is more than good enough.

Finally, consider case (ii). Then $\alpha'(r)\sim\lambda$ and $\alpha''(r)\sim 1$ for $r\in I$, so $\left|t\alpha'(r)+x\cdot\omega\right|\gtrsim\lambda\,|t|$. Thus

$$|b_1(\omega)|\lesssim\big(1/\,|t|\big)\int_I dr\lesssim\mu/\,|t|\,,$$

$$|b_2(\omega)|\lesssim\big(1/\lambda\,|t|\big)\int_I\Big(r+r^2/\mu\Big)\,dr\lesssim(\mu+\lambda)/\,|t|\lesssim\lambda/\,|t|\,.$$

Taking into account (12.5) we thus get the desired bound, and this concludes the proof of Proposition 3.4. □

Proof of Theorem 3.1(ii). If $\mu\sim\lambda$, this reduces to part (iii) of the theorem, so we may assume $\mu\ll\lambda$ (and $\lambda\gg 1/\varepsilon$). But then by an orthogonality argument (see, e.g., the proof of the analogous estimate in Theorem 12.1 of [8]) we reduce to proving $\|uv\|_{L^2_{t,x}}\lesssim\varepsilon^{1/2}\mu^{1/2}\lambda^{1/2}\,\|f\|_{L^2}\,\|g\|_{L^2}$ in the case where the Fourier transforms of f,g are supported in (diametrically opposite) cubes with side length $\sim\mu$ and at distance $\sim\lambda$ from the origin. But this follows from Hölder's inequality and the estimates (3.5) and (3.7) with $(q,r)=(4,4)$. □

Proof of Theorem 3.1(i). If $\mu\sim\lambda$, this reduces to part (iii) of the theorem, so we may assume $\mu\ll\lambda\lesssim 1/\varepsilon$. By orthogonality, we reduce to proving

$$\|uv\|_{L^2_{t,x}}\lesssim\varepsilon^{1/2}\mu\,\|f\|_{L^2}\,\|g\|_{L^2} \tag{12.7}$$

in the case where $\widehat{f},\widehat{g}$ are supported in opposite cubes $Q,-Q$ with side length $\sim\mu$ and at distance $\sim\lambda$ from the origin. By rescaling $t\to t/\varepsilon$ we further reduce to proving (12.7) without the $\varepsilon^{1/2}$ in the right hand side, and with u,v given by

$$[u(t)]\widehat{\ }(\xi)=e^{it|\xi|}\widehat{f}(\xi),\qquad [v(t)]\widehat{\ }(\xi)=e^{\pm it\varepsilon^{-1}\alpha(\varepsilon|\xi|)}\widehat{g}(\xi). \tag{12.8}$$

Here α is given by (2.18). Then by a standard Cauchy-Schwarz argument, see e.g. [2, Sect. 3.4], we finally reduce to proving that

$$\int \chi_{\{\eta:\eta\in Q\}\cap\{\eta:\eta-\xi\in Q\}}(\eta)\delta\big(\tau-|\eta|\pm k(|\xi-\eta|)\big)\,d\eta \lesssim \mu^2 \tag{12.9}$$

where

$$k(\rho) := \varepsilon^{-1}\alpha(\varepsilon\rho). \tag{12.10}$$

(Here and in what follows we use the notation χ_A for the characteristic function of a set A.) Then in view of (12.4) there is an absolute constant c_0 such that

$$\big|k'(\rho)\big| \le c_0 < 1 \quad \text{for all} \quad \rho \lesssim 1/\varepsilon,\ 0<\varepsilon<1. \tag{12.11}$$

Denote by $I_\pm(\tau,\xi)$ the integral in (12.9). In polar coordinates $\eta = r\omega$, $r>0$, $\omega\in S^2$, we have $I_\pm(\tau,\xi) = \int_{S^2} a_\pm(\tau,\xi;\omega)\,d\sigma(\omega)$, where

$$a_\pm(\tau,\xi;\omega) := \int_0^\infty \chi_{\{r:r\omega\in Q\}\cap\{r:r\omega-\xi\in Q\}}(r)\delta\big(\tau-r\pm k(|\xi-r\omega|)\big)r^2\,dr.$$

Observe that the ω-support of a is contained in a set given by (12.5), so it suffices to prove that $a_\pm \lesssim \lambda^2$. Observe also that in the integral defining $a_\pm$, the variable r is restricted to an interval I of length $\sim\mu$ and centered at a distance λ from the origin.

We shall use the following fact: If $f:\mathbb{R}\to\mathbb{R}$ is differentiable with $\big|f'(r)\big|>0$, and f has a zero at r_0, then

$$\delta\big(f(r)\big)\,dr = \frac{\delta(r-r_0)\,dr}{|f'(r_0)|}. \tag{12.12}$$

Take

$$f(r) := \tau - r \pm k(|\xi-r\omega|), \tag{12.13}$$

for fixed τ,ξ,ω. Then for r such that $r\omega-\xi\in Q$,

$$\big|f'(r)\big| = 1 \mp k'(|\xi-r\omega|)\frac{(\xi-r\omega)\cdot\omega}{|\xi-r\omega|} \ge 1-c_0 \gtrsim 1, \tag{12.14}$$

where we used (12.11) and the assumption $\lambda\lesssim 1/\varepsilon$. On account of (12.12) and (12.14), we then get $a_\pm\lesssim\lambda^2$ as desired. This concludes the proof of part (i) of Theorem 3.1. □

Proof of Theorem 3.1(iii). This reduces to proving

$$\int \chi_{\{\eta:|\eta|\sim\mu\}\cap\{\eta:|\xi-\eta|\sim\lambda\}}(\eta)\delta\big(\tau-|\eta|\pm k(|\xi-\eta|)\big)\,d\eta \lesssim \big[\min(\mu,\lambda)\big]^2, \tag{12.15}$$

for k defined by (12.10). Let us denote the above integral by $I_\pm(\tau,\xi)$. Passing to polar coordinates we have $I_\pm(\tau,\xi)=\int_{S^2} a_\pm(\tau,\xi;\omega)\,d\sigma(\omega)$, where now

$$a_\pm(\tau,\xi;\omega) := \int_0^\infty \chi_{\{r:r\sim\mu\}\cap\{r:|r\omega-\xi|\sim\lambda\}}(r)\delta\big(\tau-r\pm k(|\xi-r\omega|)\big)r^2\,dr.$$

We split into the cases (a) $\lambda\lesssim 1/\varepsilon$ and (b) $\lambda\gg 1/\varepsilon$.

Case (a). Then in view of (12.12) and (12.14) with $f(r)$ given by (12.13), we have $a_\pm\lesssim\mu^2$. Now integrate over S^2, taking into account the fact that on the support of $a_\pm$,

$$\angle(\omega,\xi) = \angle(\eta,\xi) \lesssim \lambda/\mu \quad \text{if} \quad \mu\gg\lambda. \tag{12.16}$$

Case (b). By rotational symmetry we may assume $\xi = (|\xi|,0,0)$. Now parametrize the sphere S^2 by $(y,\theta)\mapsto\omega=\Big(y,\sqrt{1-y^2}\,\nu(\theta)\Big)$, $\nu(\theta)=(\cos\theta,\sin\theta)$. Then surface measure $d\sigma(\omega)$ on S^2 becomes $dy\,d\theta$. Again we use (12.12) with $f(r)$ given by (12.13). Observe that f depends implicitly on y but not on θ. Denote by $A = A(\tau,\xi)$ the set of $y \in (-1,1)$ such that $f(r)$ given by (12.13) has a zero $r_0 = r_0(y) > 0$. Since $\left|f'(r)\right| > 0$, the implicit function theorem guarantees that A is open and $r_0 : A \to (0,\infty)$ is a smooth function. Differentiating $f\big(r_0(y)\big) = 0$ gives

$$0 = f'(r_0)r_0'(y) \mp k'(|\xi - r_0\omega|)\frac{r_0\,|\xi|}{|\xi - r_0\omega|}, \tag{12.17}$$

where we used $\xi \cdot \partial_y\omega = \xi_1 = |\xi|$ and $\omega \cdot \partial_y\omega = 0$.

Let us suppress the subscript and write $r(y)$ instead of $r_0(y)$ from now on. Solving (12.17) for $r'(y)$ and using the fact that $f'(r) < 0$, we see that $\partial r/\partial y$ is either strictly negative or strictly positive, depending on whether we have the $+$ sign or the $-$ sign in (12.15). The function $r(y)$ is therefore a change of variables.

With this information in hand, we solve (12.17) for $f'(r)$ and substitute into (12.12), thus arriving at the identity

$$\int F(\eta)\delta\big(\tau - |\eta| \pm k(|\xi - \eta|)\big)\,d\eta = \iint F(r\omega)\frac{r\,|\xi - r\omega|}{|\xi|\,k'(|\xi - r\omega|)}\left|\frac{\partial r}{\partial y}\right|\,dy\,d\theta.$$

Changing variables $y \to r$ finally gives

$$\int F(\eta)\delta\big(\tau - |\eta| \pm k(|\xi - \eta|)\big)\,d\eta = \iint F(r\omega)\frac{r\,|\xi - r\omega|}{|\xi|\,k'(|\xi - r\omega|)}\,dr\,d\theta, \tag{12.18}$$

where ω is now a function of r and θ. We apply this with

$$F(\eta) := \chi_{\{\eta:|\eta|\sim\mu\}\cap\{\eta:|\xi-\eta|\sim\lambda\}}(\eta).$$

Since $\lambda \gg 1/\varepsilon$, we see from (12.4) that $k'(|\xi - r\omega|) \sim 1$, whence

$$F(r\omega)\frac{r\,|\xi - r\omega|}{|\xi|\,k'(|\xi - r\omega|)} \sim \frac{\mu\lambda}{|\xi|}F(r\omega). \tag{12.19}$$

We now split into the subcases (b1) $\mu \ll \lambda$, (b2) $\mu \sim \lambda$ and (b3) $\mu \gg \lambda$.

Case (b2). In this case we can prove the estimate in Theorem 3.1(iii) directly, by applying Hölder's inequality followed by the linear Strichartz estimate (3.6) with $(q,r) = (4,4)$. (This works because we are at high frequency, i.e. $\gg 1/\varepsilon$.)

Case (b1). Then $|\xi| \sim \lambda$, so the desired estimate (12.15) follows readily from (12.19) and (12.18).

Case (b3). Then $|\xi| \sim \mu$, so (12.19) and (12.18) imply

$$I_\pm(\tau,\xi) \lesssim \lambda \iint \chi_{\{r:r\sim\mu\}\cap\{r:|\xi-r\omega|\sim\lambda\}}(r)\,dr\,d\theta.$$

Recall that ω is now a function of (r,θ). However, $|\xi - r\omega|$ is independent of θ, so by a slight abuse of notation we will simply write $\omega = \omega(r)$ and integrate out θ, leaving us with

$$\lambda \int \chi_{\{r:r\sim\mu\}\cap\{r:|\xi-r\omega(r)|\sim\lambda\}}(r)\,dr.$$

Clearly it suffices to prove that the support of the integrand is contained in an interval of length $\sim \lambda$. Let us assume there is no such interval, and obtain a contradiction. Fix a point r_0 in the support, and write $r = r_0 + \kappa$ for a general point r in the support. In view of our assumption, κ varies on a scale $\gg \lambda$. Thus, if we can show that

$$|\xi - r\omega(r)|^2 = a + \kappa^2 + O(\lambda\kappa + \lambda^2), \tag{12.20}$$

for some constant a, it follows that $|\xi - r\omega(r)|$ also varies on a scale $\gg \lambda$, and we have the contradiction we seek, since $|\xi - r\omega(r)| \sim \lambda$ on the support.

To prove (12.20), write $|\xi - r\omega|^2 = |\xi|^2 + (r^2 - 2r\,|\xi|) + 2r(1-\omega_1)\,|\xi|$. On account of (12.16) we have $1 - \omega_1 \lesssim (\lambda/\mu)^2$, so the last term on the right hand side is $O(\lambda^2)$. For the second term we calculate

$$r^2 - 2r\,|\xi| = (r_0^2 - 2r_0\,|\xi|) + 2(r_0 - |\xi|)\kappa + \kappa^2.$$

But $\big|r_0 - |\xi|\big| \le |r_0\omega(r_0) - \xi| \sim \lambda$, so we conclude that (12.20) holds. This ends the proof of Theorem 3.1. $\square$

13. PROOF OF THEOREM 5.3

As remarked, by a standard procedure this reduces to some well-known bilinear estimates for the homogeneous wave equation. The first observation is that by rescaling $x \to \varepsilon x$ we can reduce to the case $\varepsilon = 1$. Thus we suppress the subscript on $H^{s,\theta}$ etc. from now on.

Some notation: For $s \in \mathbb{R}$, let D^s, D_+^s and D_-^s be the Fourier multipliers $\big(D^s u\big)\widehat{\ } = |\xi|^s\,\widehat{u}$, $\big(D_+^s u\big)\widehat{\ } = \big(|\tau| + |\xi|\big)^s\widehat{u}$ and $\big(D_-^s u\big)\widehat{\ } = \big||\tau| - |\xi|\big|^s\widehat{u}$. The notation $u \precsim v$ means $\big|\widehat{u}\big| \lesssim \widehat{v}$. We are concerned with bilinear operators $B(u,v)$ of the form

$$[B(u,v)]\widehat{\ }(\tau,\xi) = \int b(\tau-\lambda,\xi-\eta;\lambda,\eta)\widehat{u}(\tau-\lambda,\xi-\eta)\widehat{v}(\lambda,\eta)\,d\lambda\,d\eta,$$

where $b(\tau,\xi;\lambda,\eta)$ is the *symbol* of B. The symbols of the null forms Q_0, Q_{ij} and Q_{0j} are, respectively,

$$q_0(\tau,\xi;\lambda,\eta) = \tau\lambda - \xi\cdot\eta, \tag{13.1a}$$

$$q_{ij}(\tau,\xi;\lambda,\eta) = -\xi_i\eta_j + \xi_j\eta_i, \tag{13.1b}$$

$$q_{0j}(\tau,\xi;\lambda,\eta) = -\tau\eta_j + \lambda\xi_j. \tag{13.1c}$$

Since we rely on estimates for the *absolute values* of these symbols, and since all norms involved only depend on the absolute value of the Fourier transform, we may assume $\widehat{u},\widehat{v} \ge 0$ henceforth.

For $s \in \mathbb{R}$, let R^s be the bilinear operator with symbol r^s, where

$$r(\tau,\xi;\lambda,\eta) = \begin{cases} |\xi| + |\eta| - |\xi+\eta| & \text{if} \quad \tau\lambda \ge 0, \\ |\xi+\eta| - \big||\xi| - |\eta|\big| & \text{if} \quad \tau\lambda < 0. \end{cases}$$

We shall need the estimate, for $\theta > 1/2$,

$$\big\| R^{1/2}(u,v) \big\|_{L^2} \lesssim \|u\|_{H^{0,\theta}}\,\|v\|_{H^{3/2,\theta}}\,, \tag{13.2}$$

which derives from an estimate for the homogeneous wave equation via the transfer principle; see [16] for the details. We also need

$$R^s(u,v) \precsim D_-^s(uv) + (D_-^s u)v + uD_-^s v \quad \text{for} \quad s \ge 0. \tag{13.3}$$

This follows easily from the triangle inequality, if one keeps track of the signs of τ and λ as in the proof of the following lemma, which is more or less standard.

Lemma 13.1. *The following estimates hold:*

$$Q_{ij}(u,v) \precsim R^{1/2}(Du, D^{1/2}v) + R^{1/2}(D^{1/2}u, Dv) \tag{13.4a}$$

$$\precsim R^{(1/2)^-}(Du, D^{(1/2)^+}v) + R^{(1/2)^-}(D^{(1/2)^+}u, Dv), \tag{13.4b}$$

$$Q_{0j}(u,v) \precsim [\textit{r.h.s.}(13.4\text{a})] + Du \cdot D_-v + D_-u \cdot Dv, \tag{13.4c}$$

$$Q_0(u,v) \precsim [\textit{r.h.s.}(13.4\text{a})] + D_+u \cdot D_-v + D_-u \cdot D_+v. \tag{13.4d}$$

Proof. All these statements reduce to estimates on the absolute values of the symbols (13.1). First, by [8, Lemma 13.2] we have

$$\left| q_{ij}(\tau,\xi;\lambda,\eta) \right| \le |\xi \times \eta| \le |\xi|^{1/2} |\eta|^{1/2} |\xi+\eta|^{1/2} [r(\tau,\xi;\lambda,\eta)]^{1/2},$$

where r is the symbol of R as defined above. Then (13.4a) and (13.4b) follow, in view of the fact that

$$r(\tau,\xi;\lambda,\eta) \le 2\min(|\xi|,|\eta|). \tag{13.5}$$

To prove (13.4c), write

$$q_{0j}(\tau,\xi;\lambda,\eta) = (\epsilon_1|\xi| - \tau)\eta_j + (\lambda - \epsilon_2|\eta|)\xi_j - \epsilon_1(|\xi|\,\eta_j - \epsilon_1\epsilon_2|\eta|\,\xi_j),$$

where ϵ_1 and ϵ_2 are the signs of τ and λ, respectively. That is, $\epsilon_1\tau = |\tau|$ and $\epsilon_2\lambda = |\lambda|$. Now take absolute values and use the fact (see [8, Lemma 13.2]) that

$$\left| |\xi|\,\eta_j \pm |\eta|\,\xi_j \right| \le |\xi|^{1/2} |\eta|^{1/2} (|\xi|+|\eta|)^{1/2} [r(\tau,\xi;\lambda,\eta)]^{1/2}$$

holds for all τ,ξ,λ,η. (The sign in the left hand side is independent of the signs of τ,λ.) This proves (13.4c). The proof of (13.4d) is similar. Write

$$q_0(\tau,\xi;\lambda,\eta) = (\tau - \epsilon_1|\xi|)\lambda + (\lambda - \epsilon_2|\eta|)\epsilon_1|\xi| + \epsilon_1\epsilon_2|\eta|\,|\xi| - \xi\cdot\eta$$

and use (see [8, Lemma 13.2]) $\left| |\eta|\,|\xi| - \xi\cdot\eta \right| \le (|\xi|+|\eta|) r(\tau,\xi;\lambda,\eta)$ and (13.5). ☐

Finally, we need the estimate (here $s_1, s_2, \theta_1, \theta_2 \ge 0$)

$$\|uv\|_{L^2} \lesssim \|u\|_{H^{s_1,\theta_1}} \|u\|_{H^{s_2,\theta_2}} \quad \text{for} \quad s_1+s_2 > \tfrac{3}{2}, \quad \theta_1+\theta_2 > \tfrac{1}{2}. \tag{13.6}$$

See [16, Proposition A.1] for the simple proof of this fact.

We are now ready to prove Theorem 5.3. By interpolation, we reduce to

$$\|Q(u,v)\|_{L^2} \lesssim \|u\|_{\dot{\mathcal{H}}^{1,\theta}} \|v\|_{\mathcal{H}^{2,1}} \tag{13.7}$$

$$\|Q(u,v)\|_{H^{0,(-1/2)^-}} \lesssim \|u\|_{\dot{\mathcal{H}}^{1,\theta}} \|v\|_{\mathcal{H}^{(3/2)^+,1}} \tag{13.8}$$

where $\|u\|_{\dot{\mathcal{H}}^{1,\theta}}$ in the right hand side can be replaced by $\|u\|_{\dot{H}^{1,\theta}}$ if $Q = Q_{ij}$.

Proof of (13.7). First observe that for the last two terms in the right hand sides of (13.4c) and (13.4d), the estimate reduces to special cases of (13.6), since we can always replace D_- by $D_-^{\theta} D_+^{1-\theta}$. Thus, it only remains to prove the estimate for the right hand side of (13.4a), but this reduces to (13.2). ☐

Proof of (13.8). First consider Q_{ij}. Applying (13.3) to (13.4b), we reduce to

$$\|uv\|_{L^2} \lesssim \|u\|_{H^{0,(1/2)+}} \|v\|_{H^{1+,1}} , \tag{13.9a}$$

$$\|uv\|_{L^2} \lesssim \|u\|_{\dot{H}^{(1/2)-,(1/2)+}} \|v\|_{H^{(1/2)+,1}} , \tag{13.9b}$$

$$\|uv\|_{H^{0,(-1/2)-}} \lesssim \|u\|_{L^2} \|v\|_{H^{1+,1}} , \tag{13.9c}$$

$$\|uv\|_{H^{0,(-1/2)-}} \lesssim \|u\|_{\dot{H}^{(1/2)-,0}} \|v\|_{H^{(1/2)+,1}} , \tag{13.9d}$$

$$\|uv\|_{H^{0,(-1/2)-}} \lesssim \|u\|_{H^{0,(1/2)+}} \|v\|_{H^{1+,(1/2)+}} , \tag{13.9e}$$

$$\|uv\|_{H^{0,(-1/2)-}} \lesssim \|u\|_{\dot{H}^{(1/2)-,(1/2)+}} \|v\|_{H^{(1/2)+,(1/2)+}} . \tag{13.9f}$$

Via duality and the transfer principle, these reduce to the estimates in Corollaries 3.2 and 3.3, which are valid in the case where u, v are both solutions of the homogeneous wave equation, as remarked in Sect. 3.

It remains to consider the second and third terms in the right hand sides of (13.4c) and (13.4d). For the second term we can apply (13.6) directly, while for the third term we replace D_- by $D_-^{(1/2)^-} D_+^{(1/2)^+}$, thus reducing to (13.9d). $\square$

14. GENERALIZATION TO MIXED STATES

Our aim here is to generalize the main results to a mixed quantum state, where we have countably many spinor fields $\{\psi_j^\varepsilon\}_{j\in\mathbb{N}}$, each satisfying the Dirac equation. This generalization is straightforward, due to the structure of the system (the key facts being that the Dirac equation is linear in the spinor, whereas the densities are quadratic in the spinor). We limit our attention to the existence results from Sects. 6–9; however, the results on the nonrelativistic limit also generalize without difficulty.

The Dirac-Maxwell-Coulomb system for a mixed state reads:

$$\left.\begin{aligned} &i\partial_t \psi_j^\varepsilon = -i\varepsilon^{-1}\alpha^k \partial_k \psi_j^\varepsilon + \varepsilon^{-2}\gamma^0 \psi_j^\varepsilon - A_k^\varepsilon \alpha^k \psi_j^\varepsilon - A_0^\varepsilon \psi_j^\varepsilon, \quad j \in \mathbb{N},\\ &\Delta A_0^\varepsilon = \rho^\varepsilon,\\ &\square_\varepsilon \mathbf{A}^\varepsilon = \varepsilon \mathbb{P}\mathbf{J}^\varepsilon, \end{aligned}\right\} \tag{14.1}$$

where

$$\rho^\varepsilon = \sum_j \mu_j^\varepsilon \left\langle \psi_j^\varepsilon, \psi_j^\varepsilon \right\rangle_{\mathbb{C}^4}, \qquad \mathbf{J}^\varepsilon = \varepsilon^{-1} \sum_j \mu_j^\varepsilon \left\langle \vec{\alpha}\psi_j^\varepsilon, \psi_j^\varepsilon \right\rangle_{\mathbb{C}^4} . \tag{14.2}$$

Here $\left\{\mu_j^\varepsilon\right\}_{j\in\mathbb{N}}$ is a given nonnegative sequence in $l^1(\mathbb{N})$, depending on the parameter ε, and we have initial data

$$\psi_j^\varepsilon|_{t=0} = \psi_{j,0}^\varepsilon \in H^1, \quad j \in \mathbb{N}; \qquad \sum_j \mu_j^\varepsilon \left\| \psi_{j,0}^\varepsilon \right\|_{H^1}^2 < \infty; \tag{14.3}$$

$$(\mathbf{A}^\varepsilon, \partial_t \mathbf{A}^\varepsilon)|_{t=0} = (\mathbf{a}_0^\varepsilon, \mathbf{a}_1^\varepsilon) \in \mathbb{P}\dot{H}^1 \times \mathbb{P}L^2. \tag{14.4}$$

The main asymptotic assumption (1.6) is now replaced by

$$\sum_j \mu_j^\varepsilon \left\| \psi_{j,0}^\varepsilon \right\|_{H^1}^2 = O(1), \quad \left\| \mathbf{a}_0^\varepsilon \right\|_{\dot{H}^1} + \varepsilon \left\| \mathbf{a}_1^\varepsilon \right\|_{L^2} = O\left(\frac{1}{\varepsilon^\Lambda} \right) \quad \text{as} \quad \varepsilon \longrightarrow 0. \tag{14.5}$$

Then we have the following generalization of Theorem 1.1.

Theorem 14.1. *The initial value problem (14.1), (14.2), (14.3), (14.4) is locally well posed for fixed* ε*, with an existence time* $T_\varepsilon > 0$ *depending only on* ε *and the size of the norms of the data. Moreover, if (14.5) holds, then*

$$T_\varepsilon \geq r \log \frac{1}{\varepsilon} \quad \textit{as} \quad \varepsilon \longrightarrow 0,$$

where r *is a constant depending on* $\sup_{\varepsilon>0} \sum_j \mu_j^\varepsilon \left\| \psi_{j,0}^\varepsilon \right\|_{L^2}^2$*, but not on* ε*. Further, (14.5) continues to hold uniformly in every finite time interval as* $\varepsilon \to 0$.

Since $\sum_j \mu_j^\varepsilon \left\| \psi_j^\varepsilon(t) \right\|_{L^2}^2$ is conserved in time [(1.3) holds for each ψ_j^ε], this theorem reduces (by the argument given after Remark 7.2), to the following analogues of Theorems 6.1 and 7.1.

Let $d\mu^\varepsilon$ denote counting measure on $\mathbb{N}$ multiplied by the weights μ_j^ε. Thus, if X is a Banach space of functions, the notation $\{\phi_j\} \in l^2(d\mu^\varepsilon; X)$ means that $\sum \mu_j^\varepsilon \left\| \phi_j \right\|_X^2 < \infty$.

Theorem 14.2. *(Cf. Theorem 6.1.) For fixed* ε*, the Dirac-Maxwell-Coulomb system (14.1), (14.2) is locally well posed for initial data in the space (14.3), (14.4). The existence time* $T > 0$ *only depends on* ε *and the size of the norms of the data, and the solution belongs to and is unique in the class*

$$\{\psi_j^\varepsilon\}_{j\in\mathbb{N}} \in l^2\left(d\mu^\varepsilon; H_\varepsilon^{1,\theta}(S_T)\right), \quad \mathbf{A}^\varepsilon \in \dot{\mathcal{H}}_\varepsilon^{1,\theta}(S_T), \quad A_0^\varepsilon \in C([0,T]; \dot{H}^1), \tag{14.6}$$

for all $1/2 < \theta < 1$*. Moreover,* $\{\phi_{\pm,j}^\varepsilon\}_{j\in\mathbb{N}} \in l^2\left(d\mu^\varepsilon; X_{\tau=\pm h_\varepsilon(\xi)}^{1,\theta}(S_T)\right)$.

The last statement requires some explanation: We consider $\psi^\varepsilon = \{\psi_j^\varepsilon\}_{j\in\mathbb{N}}$ to be a vector-valued solution of the Dirac equation, and the considerations in Sect. 2, in particular the definitions of $\psi_\pm^\varepsilon$ and $\phi_\pm^\varepsilon$, then apply *componentwise*.

The same remark applies to the iteration scheme in Sect. 6. Thus, we have a vector-valued sequence of iterates $\psi^{\varepsilon,m} = \{\psi_j^{\varepsilon,m}\}_{j\in\mathbb{N}}$ satisfying (6.2a), and (6.3)–(6.7) are understood to be vector equations, while (6.8) holds componentwise, i.e., with index j on the $\psi_\pm$ and ψ fields. The iterates of the potentials may conveniently be written

$$A_0^{\varepsilon,m} = \sum \mu_j^\varepsilon A_{0,j}^{\varepsilon,m}, \quad \mathbf{A}^{\varepsilon,m} = \mathbf{A}^{\varepsilon,0} + \sum \mu_j^\varepsilon \mathbf{A}_j^{\varepsilon,m}, \tag{14.7}$$

where $A_{0,j}^{\varepsilon,m} = \Delta^{-1}\langle \psi_j^{\varepsilon,m}, \psi_j^{\varepsilon,m} \rangle$, $\mathbf{A}^{\varepsilon,0}$ is the solution of the homogeneous wave equation with initial data (14.4) and $\mathbf{A}_j^{\varepsilon,m} = \mathbf{A}[\psi_j^{\varepsilon,m}, \psi_j^{\varepsilon,m}]$, with notation as in Sect. 5.3.

The proof of Theorem 14.2 then follows exactly that of Theorem 6.1, with a few obvious modifications which we now list. In (6.9) we replace the $H_\varepsilon^{1,\theta}$ norm of $\psi^{\varepsilon,m}$ by the $l^2\big(d\mu^\varepsilon; H_\varepsilon^{1,\theta}(S_T)\big)$ norm. (6.11) holds with index j on the $\mathbf{A}$ and ψ fields, and without the data norm in the right hand side; multiplying both sides by μ_j^ε and summing gives the synthetic estimate

$$\left\| \mathbf{A}^{\varepsilon,m+1} \right\|_{\dot{\mathcal{H}}_\varepsilon^{1,\theta}(S_T)} \leq C \left(\left\| \mathbf{a}_0^\varepsilon \right\|_{\dot{H}^1} + \varepsilon \left\| \mathbf{a}_1^\varepsilon \right\|_{L^2} \right) + T^\delta \sum \mu_j^\varepsilon \left\| \psi_j^{\varepsilon,m} \right\|_{L^4(S_T)}^2,$$

and (6.13) is similarly interpreted. Further, (6.12) and (6.14) hold with index j on the ψ, $\phi_\pm$ and R fields (but not on A_0); squaring both sides, multiplying by μ_j^ε and summing gives the corresponding synthetic estimates. Taken all together, this proves (6.10), hence Theorem 14.2.

Similar considerations show that the results in Sects. 7–9 generalize to mixed states; moreover, the proofs remain valid, if one observes the following simple rules: (i) For any quantity F estimated in some norm $\|F\|$, and which is now *vector-valued*, i.e., $F = \{F_j\}_{j\in\mathbb{N}}$, we substitute $\|F\|$ by $\left(\sum \mu_j^\varepsilon \|F_j\|^2\right)^{1/2}$; the pure state estimates can be applied componentwise, then squared and summed with weights μ_j^ε—this works because the Dirac equation is linear in the spinor field. (ii) For the potentials, the pure state estimates can be used via the decompositions in (14.7), then summed with weights μ_j^ε; here the bilinearity of the potentials with respect to the spinor field comes into play. Thus, we content ourselves with stating the key results in generalized form, leaving the details of the proofs to the interested reader.

Defining

$$\mathcal{I}^\varepsilon(t) = \left(\sum \mu_j^\varepsilon \left\|\psi_j^\varepsilon(t)\right\|_{H^1}^2\right)^{\frac{1}{2}} + \varepsilon^\Lambda\left(\left\|\mathbf{A}^\varepsilon(t)\right\|_{\dot{H}^1} + \varepsilon\left\|\partial_t\mathbf{A}^\varepsilon(t)\right\|_{L^2}\right),$$

we have:

Theorem 14.3. *(Cf. Theorem 7.1.) Consider the solution* $(\psi^\varepsilon = \{\psi_j^\varepsilon\}_{j\in\mathbb{N}}, A_\mu^\varepsilon)$ *of (14.1), (14.2), (14.3), (14.4) from Theorem 6.1, existing up to a time* $T_\varepsilon > 0$ *and belonging to (14.6) over this time interval. There exist* $T^* > 0$*, depending only on* $\sup_{\varepsilon>0}\sum_j \mu_j^\varepsilon \left\|\psi_{j,0}^\varepsilon\right\|_{L^2}^2$*, and constants* $C, M, \varepsilon_0 > 0$*, independent of* ε*, such that if*

$$\mathcal{I}^\varepsilon(0) \le B \quad \text{for all} \quad \varepsilon,$$

then $\mathcal{I}^\varepsilon(t) \le CB$ *for* $0 < \varepsilon \le \varepsilon_0 B^{-M}$ *and* $0 \le t \le \min(T^*, T_\varepsilon)$.

Next, we consider bounds on higher derivatives. Set

$$\mathcal{I}_m^\varepsilon(t) = \left(\sum \mu_j^\varepsilon \left\|\psi_j^\varepsilon(t)\right\|_{H^{m+1}}^2\right)^{\frac{1}{2}} + \varepsilon^\Lambda\left(\left\|\nabla\mathbf{A}^\varepsilon(t)\right\|_{H^m} + \varepsilon\left\|\partial_t\mathbf{A}^\varepsilon(t)\right\|_{H^m}\right)$$

for $m \ge 0$. Then we have:

Proposition 14.4. *(Cf. Proposition 8.1.) Assume* $\mathcal{I}_m^\varepsilon(0) \le B$ *uniformly in* ε*. Given* $T < \infty$*, let* N *be the smallest natural number such that* $NT^* \ge T$*, with* T^* *as in Theorem 14.3. Then*

$$\mathcal{I}_m^\varepsilon(t) \le P_m\left(C^N B\right) \quad \text{for} \quad 0 \le t \le T, \quad \text{provided} \quad 0 < \varepsilon \le \frac{\varepsilon_0}{(C^{N-1}B)^M}.$$

Here C, M, ε_0 *are universal constants and* P_m *is a polynomial.*

Finally, we have the "small positron" result.

Theorem 14.5. *(Cf. Theorem 9.1.) Let* $T < \infty$ *be given, and let* N *be the smallest natural number such that* $NT^* \ge T$*, with* T^* *as in Theorem 14.3.*

(i) *Suppose* $\mathcal{I}_m^\varepsilon(0) \le B$ *and* $\left(\sum \mu_j^\varepsilon \left\|\Pi_-(\varepsilon D)\psi_j^\varepsilon(t=0)\right\|_{H^m}^2\right)^{\frac{1}{2}} \le \varepsilon B$ *for all* ε*. Then*

$$\left(\sum \mu_j^\varepsilon \left\|\Pi_-(\varepsilon D)\psi_j^\varepsilon(t)\right\|_{H^m}^2\right)^{\frac{1}{2}} \le \varepsilon P_m(C^N B)$$

for all $0 \le t \le T$*, provided* $0 < \varepsilon \le \varepsilon_0(C^{N-1}B)^{-M}$*. Here (and below)* C, M, ε_0 *are universal constants and* P_m *is a polynomial.*

Moreover, if we strengthen the assumption $\mathcal{I}_m^\varepsilon(0) \le B$ *by requiring that* $\left\|\nabla\mathbf{a}_0^\varepsilon\right\|_{H^m} + \varepsilon\left\|\mathbf{a}_1^\varepsilon\right\|_{H^m} \le B$ *for all* ε*, then*

$$\left\|\nabla\mathbf{A}^\varepsilon(t)\right\|_{H^m} + \varepsilon\left\|\partial_t\mathbf{A}^\varepsilon(t)\right\|_{H^m} \le P_m(C^N B)$$

for all $0 \le t \le T$*, provided* ε *satisfies the same condition as above.*

(ii) *Assume that the hypotheses of part (i) are satisfied, for some* $m \ge 1$*. If, in addition,*

$$\left(\sum \mu_j^\varepsilon \left\| \Pi_-(\varepsilon D)\psi_j^\varepsilon(t=0) \right\|_{H^{m-1}}^2 \right)^{\frac{1}{2}} \le \varepsilon^2 B.$$

Then

$$\left(\sum \mu_j^\varepsilon \left\| \Pi_-(\varepsilon D)\psi_j^\varepsilon(t) \right\|_{H^{m-1}}^2 \right)^{\frac{1}{2}} \le \varepsilon^2 P_m(C^N B)$$

for all $0 \le t \le T$*, provided* ε *satisfies the same condition as in part (i).*

ACKNOWLEDGMENTS

We are indebted to the referee for a number of useful suggestions which improved our presentation. Financial support by the Austrian START project “Nonlinear Schrödinger and quantum Boltzmann equations” (FWF Y137-TEC) of N.J.M. and by the European network HYKE (HPRN-CT-2002-00282) as well as by the OeAD (“acciones integradas”) is acknowledged. The last author was supported in part by the project “PDE and Harmonic Analysis”(160192/V30) funded by the Norwegian Research Council.

REFERENCES

[1] P. Bechouche, N.J. Mauser and F. Poupaud, (Semi)-nonrelativistic limits of the Dirac equation with external time-dependent electromagnetic field, *Comm. Math. Phys.* **197** (1998), no. 2, 405–425

[2] P. Bechouche, N.J. Mauser and S. Selberg, Nonrelativistic limit of Klein-Gordon-Maxwell to Schrödinger-Poisson, *Amer. J. Math.* **126** (2004), 31–64

[3] P. Bechouche, N.J. Mauser and S. Selberg, Derivation of Schrödinger-Poisson as the nonrelativistic limit of Klein-Gordon-Maxwell, *Hyperbolic problems: theory, numerics, applications,* 357–367, Springer, Berlin, 2003

[4] J. Bourgain, Fourier transform restriction phenomena for certain lattice subsets and applications to nonlinear evolution equations I: Schrödinger equations, and II: The KdV equation, *Geom. Funct. Anal.* **3** (1993), 107–156 and 209–262

[5] N. Bournaveas, Local existence for the Maxwell-Dirac equations in three space dimensions, *Comm. Partial Differential Equations* **21** (1996), no. 5-6, 693–720

[6] R. Cirincione and P.R. Chernoff, Dirac and Klein-Gordon equations: convergence of solutions in the non-relativistic limit, *Comm. Math. Phys.* **79** (1981), no. 1, 33–46

[7] P.A.M. Dirac, *Principles of Quantum Mechanics, 4th ed.,* Oxford University Press, London, 1958

[8] D. Foschi and S. Klainerman, Homogeneous L^2 bilinear estimates for wave equations. *Ann. Scient. ENS* 4^e *serie* **23** (2000), 211–274

[9] V. Georgiev, Small amplitude solutions of the Maxwell-Dirac equations. *Indiana Univ. Math. J.* **40** (1991), no. 3, 845–883

[10] L. Gross, The Cauchy problem for the coupled Maxwell and Dirac equations, *Comm. Pure Appl. Math.* **19** (1966), 1–15

[11] C. Kenig, G. Ponce and L. Vega, The Cauchy problem for the KdV equation in Sobolev spaces of negative indices, *Duke Math. J.* **71** (1994), 1–21

[12] S. Klainerman and M. Machedon, Space-time estimates for null forms and the local existence theorem, *Comm. Pure Appl. Math.* **46** (1993), 1221–1268

[13] S. Klainerman and M. Machedon, On the Maxwell-Klein-Gordon equation with finite energy, *Duke Math. J.* **74** (1994), 19–44

[14] S. Klainerman and M. Machedon, Smoothing estimates for null forms and applications, *Duke Math. J.* **81** (1995), 99–133

[15] S. Klainerman and M. Machedon, personal communication

[16] S. Klainerman and S. Selberg, Bilinear estimates and applications to nonlinear wave equations, *Commun. Contemp. Math.* **4** (2002), no. 2, 223–295

[17] S. Klainerman and D. Tataru, On the optimal local regularity for Yang-Mills equations in $\mathbb{R}^{4+1}$, *J. Amer. Math. Soc.* **12** (1999), 93–116

[18] L.D. Landau and E.M. Lifschitz, *Quantenmechanik, Vol. III, 2nd ed.,* Akademie-Verlag, Berlin, 1971

[19] S. Machihara, K. Nakanishi and T. Ozawa , Nonrelativistic limit in the energy space for nonlinear Klein-Gordon equations, *Math. Ann.* **322** (2002), 603–621

[20] N.J. Mauser, Semi-relativistic approximations of the Dirac equation: first and second order corrections, *Trans. Theor. Stat. Phys.* **29** (2000), 122-137

[21] N.J. Mauser and S. Selberg, Convergence of the Dirac-Maxwell system to the Vlasov-Poisson system, submitted (2004)

[22] N. Masmoudi and N.J. Mauser, The selfconsistent Pauli equation, *Mathematische Monatshefte* **132** (2001), 19-24

[23] N. Masmoudi and K. Nakanishi, From Maxwell-Klein-Gordon and Maxwell-Dirac to Poisson-Schrödinger, *Internat. Math. Res. Notices* (2003), no. 13, 697–734

[24] B. Najman, The nonrelativistic limit of the nonlinear Dirac equation, *Ann. Inst. Henri Poincaré Anal. Non Lineaire*, **9** (1992), 3–12

[25] S. Selberg, On an estimate for the wave equation and applications to nonlinear problems, *Differential and Integral Equations* **2** (2002), 213–236

[26] T. Tao, Multilinear weighted convolution of L^2 functions, and applications to nonlinear dispersive equations, *Amer. J. Math.* **123** (2001) , no. 5, 839–908

DPT. MATEMÁTICA APLICADA, UNIVERSIDAD DE GRANADA, C/ SEVERO OCHOA S/N, 18071 GRANADA, SPAIN
E-mail address: `phbe@ugr.es`

WOLFGANG PAULI INSTITUTE C/O INST. F. MATH. UNIV. WIEN, NORDBERGSTR. 15, 1090 WIEN, AUSTRIA
E-mail address: `norbert.mauser@univie.ac.at`

DEPT. OF MATH. SCIENCES, NORWEGIAN UNIV. OF SCIENCE AND TECHNOLOGY, ALFRED GETZ' V. 1, 7491 TRONDHEIM, NORWAY
E-mail address: `sigmund.selberg@math.ntnu.no`