ON THE CONVERGENCE OF FROZEN GAUSSIAN APPROXIMATION FOR LINEAR NON-STRICLY HYPERBOLIC SYSTEMS

LIHUI CHAI†, JAMES C. HATELEY‡, EMMANUEL LORIN§, AND XU YANG¶

Abstract. Frozen Gaussian approximation (FGA) has been applied and numerically verified as an efficient tool to compute high-frequency wave propagation modeled by non-strictly hyperbolic systems, such as the elastic wave equations [J.C. Hateley, L. Chai, P. Tong and X. Yang, Geophys. J. Int., 216, 1394–1412, 2019] and the Dirac system [L. Chai, E. Lorin and X. Yang, SIAM Numer. Anal., 57, 2383–2412, 2019]. However, the theory of convergence is still incomplete for non-strictly hyperbolic systems, where the latter can be interpreted as a diabatic (or more) coupling. In this paper, we establish the convergence theory for FGA for linear non-strictly hyperbolic systems, with an emphasis on the elastic wave equations and the Dirac system. Unlike the convergence theory of FGA for strictly linear hyperbolic systems, the key estimate lies in the boundness of intraband transitions in diabatic coupling.

Keywords. Frozen Gaussian approximation, Convergence, Non-strictly hyperbolic, Elastic wave equations, Dirac equation

AMS subject classifications. 65M12,81Q05

1. Introduction

The goal of this paper is to provide the convergence theory of frozen Gaussian approximation (FGA) for linear non-strictly hyperbolic systems. In [13], the authors study in detail the convergence and the accuracy of FGA’s applied to linear strictly hyperbolic systems in high frequency regime. However, several fundamental hyperbolic systems are not strictly hyperbolic, such as the elastic wave equations or the Dirac system modeling in particular quantum relativistic particles, see [18]. The paper aims to precisely study the boundness of intraband transitions in the diabatic coupling, which is specific to non-strictly hyperbolic systems.

For the sake of clarity, we then shall first consider the two fundamental examples mentioned above: elastic wave equations and the Dirac system; then we will extend the arguments for general non-strictly hyperbolic systems by mainly focusing on the technical consequences due to the multiplicity of some eigenvalues of Jacobian matrices, and will refer to the appropriate references in the strictly hyperbolic case.

We start by introducing the elastic wave equations in three dimensions which models elastic wave propagation, as in [6]. We define the elastic wave system:

\[
\begin{align*}
\text{Elastic wave system (EWS):} & \quad \left\{ \begin{array}{l}
(\rho(x)\partial_t^2 - \mathcal{L})u(t,x) = 0, \\
u(0,x) = u_0(x), \\
\partial_t u(0,x) = u_1(x),
\end{array} \right.
\end{align*}
\] (1.1)
where the operator $\mathcal{L}$ is given by

$$
\mathcal{L}u(t, x) = (\lambda(x) + 2\mu(x))\nabla(\nabla \cdot u(t, x)) - \mu(x) \nabla \times (\nabla \times u(t, x)),
$$

with the differential operators taken in the spatial variables, $\lambda, \mu : \mathbb{R}^3 \to \mathbb{R}$ being the first and second Lamé parameters and $\rho : \mathbb{R}^3 \to \mathbb{R}$ is a material density. We remark that the P-, S- wave speeds (e.g., $\omega$) respectively. In addition, if one considers the EWS and defines the following quantities

$$
\Theta(t, x) = \nabla \cdot u(t, x), \quad \psi(t, x) = \nabla \times u(t, x), \quad \nu(t, x) = \partial_t u(t, x),
$$

with $\nu = (v_1, v_2, v_3)^T$ and $\Psi = (\Theta_1, \Psi_2, \Psi_3)^T$ with $X = (v_1, v_2, v_3, \Theta, \Psi_1, \Psi_2, \Psi_3)^T$, then the elastic wave equations can be written as a matrix system,

$$
\partial_t X = M_x \partial_x X + M_y \partial_y X + M_z \partial_z X,
$$

where, using sparse notation; e.g., $M_{ij} = v$ is denoted $(i, j, v)$, $M_x, M_y, M_z$ are as follows:

$$
M_x : (1,4,c_2^p),(2,7,c_2^p),(3,6,-c_2^s),(4,1,1),(7,2,1),(6,3,-1),
$$

$$
M_y : (1,7,-c_2^s),(2,4,c_2^p),(3,5,c_2^s),(4,2,1),(5,3,1),(7,1,-1),
$$

$$
M_z : (1,6,c_2^p),(2,5,-c_2^s),(3,4,c_2^p),(4,3,1),(6,1,1),(5,2,-1).
$$

It can be seen that eq. (1.5) is a non-strictly hyperbolic system; the eigenvalues of $M_x + M_y + M_z$ are $\pm c_p, 0, \pm c_s$, where $\pm c_s$ have a multiplicity of 2.

Another fundamental non-strictly hyperbolic system that we shall study is the Dirac system, usually referred in the Physics literature as the **Dirac equation**:

**Dirac System (DS):**

$$
\begin{align*}
\varepsilon \partial_t \psi^\varepsilon(t, x, \sigma) &= (-i\varepsilon \sigma \cdot \nabla - \hat{\sigma} \cdot A(x) + m\beta c^2 + V(x)) \psi^\varepsilon(t, x, \sigma), \\
\psi^\varepsilon(x, 0, \sigma) &= \varphi^\varepsilon(x) = \omega_I(x) \exp \left( \frac{i}{\varepsilon} S_I(x) \right),
\end{align*}
$$

where $\psi^\varepsilon = (\psi_1^\varepsilon, \psi_2^\varepsilon, \psi_3^\varepsilon, \psi_4^\varepsilon)^T$ which takes its values in $\mathbb{C}^4$ is a 4-spinor, $S_I$ is the initial phase, $\omega_I$ the initial amplitude with $\varphi_I \in L^2(\mathbb{R}^d; \mathbb{C}^4)$. The Dirac matrices $\hat{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$, $\beta$ are defined as follows. For $\gamma = x, y, z$,

$$
\alpha_\gamma = \begin{bmatrix} 0 & \sigma_\gamma \\ \sigma_\gamma & 0 \end{bmatrix}, \quad \beta = \begin{bmatrix} \mathbb{I}_2 & 0 \\ 0 & -\mathbb{I}_2 \end{bmatrix}.
$$

The $\sigma_\gamma$’s are the $2 \times 2$ Pauli matrices defined as

$$
\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \quad \text{and} \quad \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},
$$

and $\mathbb{I}_2$ is the $2 \times 2$ unit matrix. The momentum operator is denoted $p = -i\nabla$. The speed of light $c$ and fermion mass $m$ are kept explicit. This equation models a relativistic electron of mass $m$ subject to an interaction potential $V$ and an electromagnetic field $A$, and where for fixed time $t$, $\psi^\varepsilon(t, x) \in L^2(\mathbb{R}^d) \otimes \mathbb{C}^4$ is the coordinate $(x = (x, y, z))$ dependent 4-spinor. We set

$$
B = -\hat{\sigma} \cdot A + m\beta c^2 + V,
$$

(1.10)
and then the Dirac operator can be written as
\[ D = -i\varepsilon \hat{\sigma} \cdot \nabla + B, \]  
\[ (1.11) \]
and its corresponding semi-classical symbol which reads
\[ D(q,p) = \hat{\sigma} \cdot (pc - A(q)) + m\hat{\sigma}_0 c^2 + V(q) = \hat{\sigma} \cdot pc + B(q), \]  
\[ (1.12) \]
is a Hermitian matrix which has two double eigenvalues
\[ h_{\pm}(q,p) = \sqrt{|pc - A(q)|^2 + c^4 + V(q)}, \]  
\[ (1.13) \]
with the corresponding normalized eigenvectors denoted as \( \Upsilon_{\pm 1} \) and \( \Upsilon_{\pm 2} \). For more details of the computation for the eigenvalues please refer to [2].

The main result of this paper is the proof that FGA for both EWS [6] and DS [2] are first order convergent, as for strictly hyperbolic systems [14]. The proof will follow the same machinery from [14], with however more careful estimates for the boundness of intraband transitions in diabatic coupling. The main theorem for which will provide a detailed proof, reads as follows,

**Main Theorem 1.** Let \( \{u_0^\varepsilon\} \) be a family of asymptotically high frequency initial conditions. Let \( u : [0,T] \times \mathbb{R}^d \) satisfies respective hyperbolic system (EWS) (1.1) or (DS) (1.7).

- *(For elastic wave system).* Let \( u_0^\varepsilon \in H^1_0(\mathbb{R}^d) \) uniformly bounded, i.e. \( \|u_0^\varepsilon\|_{H^1_0} < M \), and let \( u_F \) be the FGA to (1.1), then
  \[ \sup_{t \in [0,T]} \|u(t,\cdot) - u_F(t,\cdot)\|_E \leq \varepsilon C_T, \]
  where the norm \( \|\cdot\|_E \) is a scaled semi-norm defined in (3.53).

- *(For Dirac system).* Let \( u_0^\varepsilon \in L^2(\mathbb{R}^d) \) uniformly bounded, and \( u_F \) is the FGA to (1.7), then
  \[ \sup_{t \in [0,T]} \|u(t,\cdot) - u_F(t,\cdot)\|_{L^2} \leq \varepsilon C_T. \]

In each case, \( C_T \) is a constant depending on the final time \( T \).

**Related works.** The FGA, introduced originally in quantum mechanics as the Herman-Kluk (HK) propagator [7–9], was used to approximate the solution of the Schrödinger equation in the semi-classical regime. The mathematical analysis was then proposed in [16,17] to show the accuracy and efficiency of the HK ansatz, in particular, when the initial data are localized in phase space. HK formalism was later developed for several types of partial differential equations, such as the wave equations [12], linear hyperbolic systems of conservation laws [13], elastic wave equations, and seismic tomography [3, 4, 6]. The FGA for the elastic wave equations has been used to train neural networks for seismic interface and pocket detection [5]. Some applications and analysis on the Schrödinger equations were also proposed in [10,11,19].

**Organization of the paper.** In Section 2, we introduce the necessary notations and preliminaries needed for phase plane analysis. We present the full convergence analysis for the elastic wave equations in Section 3, and for the Dirac system in Section 4. In Section 5, we provide the key arguments which allow for generalizing the convergence statements to any linear non-strictly hyperbolic systems. In Section 6, we propose some concluding remarks.
2. Notations and preliminaries

In this section, we introduce some notations and preliminary results that are needed for the convergence analysis of both the elastic wave equations and the Dirac system.

2.1. Notation We denote $x,y \in \mathbb{R}^d$ as spatial variables, $(q,p) \in \mathbb{R}^{2d}$ for the position and momentum variables, respectively, in the phase space.

We use hereafter the notation $O(\varepsilon^\infty) \colon A^\varepsilon = O(\varepsilon^\infty)$ meaning for any $k \in \mathbb{N}$

$$\lim_{\varepsilon \to 0} \varepsilon^{-k}|A^\varepsilon| = 0. \tag{2.1}$$

Notation $C$ will be used as a general positive constant, that can vary from line to line. The explicit value of this finite constant is however irrelevant in the analysis. We will use subscripts to denote constant dependence, e.g. $C_T$, is a constant that depends on the parameter $T$. We will respectively denote by $S$, $C^\infty$ and $C^\infty_c$, the Schwartz class, smooth and compactly supported smooth function spaces. For generality, we will often use $\mathbb{R}^d$ for a $d$-dimensional Euclidean space; however, for the actual equations and computations we set $d = 3$, as we deal with these differential operators on $\mathbb{R}^3$.

2.2. Wave packet decomposition For any $(q,p) \in \mathbb{R}^{2d}$, we define $\phi^\varepsilon_{q,p}$ as

$$\phi^\varepsilon_{q,p}(x) = (-2\pi \varepsilon)^{-d/2} \exp\left(i p \cdot (x - q)/\varepsilon - |x - q|^2/(2\varepsilon)\right). \tag{2.2}$$

We recall that the Fourier-Bros-Iagolnitzer (FBI) transform on $\mathcal{S}(\mathbb{R}^d)$ [15], is defined as

$$(\mathcal{F}^\varepsilon f)(q,p) = (\pi \varepsilon)^{-d/4} \langle \psi^\varepsilon_{q,p}, f \rangle$$

$$= 2^{-d/2}(\pi \varepsilon)^{-3d/4}\int_{\mathbb{R}^d} \exp\left(i p \cdot (x - q)/\varepsilon - |x - q|^2/(2\varepsilon)\right)f(x) \, dx.$$  

The inverse transform $(\mathcal{F}^\varepsilon)^*$ defined on $\mathcal{S}(\mathbb{R}^{2d})$ is given by

$$\left((\mathcal{F}^\varepsilon)^* f \right)(x) = 2^{-d/2}(\pi \varepsilon)^{-3d/4}\int_{\mathbb{R}^{2d}} \exp\left(i p \cdot (x - q)/\varepsilon - |x - q|^2/(2\varepsilon)\right)g(q,p) \, dq \, dp. \tag{2.3}$$

The following is a standard result from microlocal analysis, e.g., see [1].

PROPOSITION 2.1. For Schwartz class functions, the FBI transform is an isometry on $\mathbb{R}^d$, i.e., for any $f \in \mathcal{S}(\mathbb{R}^d)$,

$$\|\mathcal{F}^\varepsilon f\|_{L^2} = \|f\|_{L^2}. \tag{2.4}$$

Furthermore; $(\mathcal{F}^\varepsilon)^* \mathcal{F}^\varepsilon = \text{Id}_{L^2(\mathbb{R}^d)}$. By standard density arguments, this implies that the domain of $\mathcal{F}^\varepsilon$ and $(\mathcal{F}^\varepsilon)^*$ can be extended to $L^2(\mathbb{R}^d)$ and $L^2(\mathbb{R}^{2d})$ respectively.

DEFINITION 2.1. Let $\{u^\varepsilon\} \subset L^1(\mathbb{R}^d)$ be a family of functions which is uniformly bounded. Given $\delta > 0$, $\{u^\varepsilon\}$ is asymptotically high frequency with cut off $\delta$, if

$$\int_{\mathbb{R}^{2d} \setminus K_\delta} |(\mathcal{F}^\varepsilon u^\varepsilon)(q,p)|^2 \, dq \, dp = O(\varepsilon^\infty), \tag{2.5}$$

as $\varepsilon \to 0$.

DEFINITION 2.2. For $M_n \in L^\infty(\mathbb{R}^{2d};\mathbb{C}^{N \times N})$ and a Schwartz function $u \in \mathcal{S}(\mathbb{R}^d;\mathbb{C}^n)$, for each $n = 1,\ldots,N$ we define the Fourier integral operator (FIO) $T^\varepsilon_n(t,M)u$ as

$$(T^\varepsilon_n(t,M)u)(x) = (2\pi \varepsilon)^{-3d/2}\int_{\mathbb{R}^{3d}} G^\varepsilon_n(t,x,y,p,q)M(q,p)u(y) \, dq \, dp \, dy, \tag{2.6}$$
with
\[ G_n^{\varepsilon}(t, x, y, p, q) = e^{i\phi_n(t, x, y, p, q)/\varepsilon}, \]
where the phase function is defined as
\[ \phi_n(t, x, y, p, q) = \frac{i}{2} |y - q|^2 - p \cdot (y - q) + \frac{i}{2} |x - Q_n(t, q, p)|^2 + P_n(t, q, p) \cdot (x - Q_n(t, q, p)). \]

Proposition 2.2. If \( M \in L^\infty(\mathbb{R}^{2d}; \mathbb{C}^{N \times N}) \), for any \( t \) and each \( n = 1, \ldots, N \), \( I_n^\varepsilon(t, M) \) can be extended to a bounded linear operator on \( L^2(\mathbb{R}^{2d}; \mathbb{C}^N) \) with the bound
\[ \|I_n^\varepsilon(t, M)\|_{L^2(\mathbb{R}^{2d}; \mathbb{C}^N)} \leq 2^{-d/2} \|M\|_{L^\infty(\mathbb{R}^{2d}; \mathbb{C}^{N \times N})}. \]
This is Proposition 3.7 in [13], with a more general version proved in [17, Theorem 2].

3. Convergence analysis for the elastic wave equations
In this section, we first introduce the Hamiltonian flow associated to the FGA formulation of the elastic wave equations (EWS) described in (1.1), and related boundedness estimates on the quantities used in the FGA formulation. Then we compute the asymptotic corrections, and prove that these correction terms are bounded in proper norms, which eventually implies the convergence results.

3.1. Hamiltonian flow for EWS and some estimates
According to the results in [6], the Hamiltonian associated with \( \Theta, \Psi \) in the FGA formulation for both P- and S-waves are
\[ H_{p\pm,s\pm}(t, Q, P) = \pm c_{p,s}(Q_{p\pm,s\pm}(t, q, p)) P_{p\pm,s\pm}(t, q, p), \]
where the wave speeds \( c_{p,s} \) are given in (1.3). The corresponding flows are given by
\[ \begin{cases} 
\frac{dQ_{p\pm,s\pm}}{dt}(0, q, p) = \pm c_{p,s}(Q_{p\pm,s\pm}(t, q, p)) P_{p\pm,s\pm}(t, q, p), \\
\frac{dP_{p\pm,s\pm}}{dt}(0, q, p) = \mp \partial_q c_{p,s}(Q_{p\pm,s\pm}(t, q, p)) |P_{p\pm,s\pm}(t, q, p)|.
\end{cases} \]
with initial conditions
\[ Q_{p\pm,s\pm}(0, q, p) = q \text{ and } P_{p\pm,s\pm}(0, q, p) = p. \]

We remark that
\[ |p \cdot \partial_q H(t, q, p)| \lesssim |p|^2, \quad \text{and} \quad |q \cdot \partial_p H(t, q, p)| \lesssim |q|^2, \]
so that the global Lipschitz assumption A in [14] is satisfied.

For \( \delta > 0 \), define the closed set \( K_\delta \subset \mathbb{R}^{2d} \)
\[ K_\delta = \left\{(q, p) \in \mathbb{R}^{2d}: |q| \leq 1/\delta, \quad \delta \leq |p| \leq 1/\delta \right\}. \]
For all practical purposes, the set \( K_\delta \) is bounding the position and the magnitude for the direction of propagation of the wave packets. The upper bound conditions on \( |q| \) and \( |p| \) are reasonable as any computational domain will be a finite domain. Also, \( p \)
bounded away from zero is reasonable as if \( \mathbf{p} = 0 \) the wave packet does not propagate and the Hamiltonian system is degenerate i.e., \( H = 0 \).

**Proposition 3.1.** For \( T > 0 \) and \( \delta > 0 \), there is a constant \( \delta_T > 0 \), such that

\[
(J_{p,s}(q,p) = \begin{pmatrix}
(\partial_q Q_{p,s})^T(q,p) & (\partial_p Q_{p,s})^T(q,p) \\
(\partial_q P_{p,s})^T(q,p) & (\partial_p P_{p,s})^T(q,p)
\end{pmatrix},
\]

for any \((q,p) \in K_\delta\) and \( t \in [0,T] \).

**Proof.** This is Proposition 3.1 in [14], and the proof requires the bound in (3.4) and Gronwall’s inequality. For the following definition we shall omit the branch subscript.

**Definition 3.1.** A map \( \kappa_{p,s} : (q,p) \to (Q_{p,s}(q,p), P_{p,s}(q,p)) \) is called a canonical transformation if the associated Jacobian matrix is symplectic, i.e., for any \((q,p)\)

\[
J_{p,s}(q,p) = \begin{pmatrix}
(\partial_q Q_{p,s})^T(q,p) & (\partial_p Q_{p,s})^T(q,p) \\
(\partial_q P_{p,s})^T(q,p) & (\partial_p P_{p,s})^T(q,p)
\end{pmatrix},
\]

is such that

\[
J_{p,s}^{-1} = \begin{pmatrix}
0 & \text{Id}_3 \\
\text{Id}_3 & 0
\end{pmatrix},
\]

where \( \text{Id}_3 \) is a 3 \times 3 identity matrix.

**Proposition 3.2.** The map \( \kappa_{p,s} \) is a canonical transform for any \( T, \delta > 0 \); furthermore it has a bounded sup-norm.

**Proof.** This is Proposition 3.4 in [14]. For a canonical transform \( \kappa_{p,s} \) define the quantity \( Z^{\kappa_{p,s} \cdot t}(q,p) \) for \(|p| > 0\) as

\[
Z^{\kappa_{p,s} \cdot t} = \partial_z (Q(q,p) + iP(q,p)),
\]

with \( \partial_z = (\partial_q - i\partial_p) \). Dropping the superscript \( \kappa_{p,s} \),

\[
Z = (i\text{Id}_3 \text{Id}_3) \begin{pmatrix}
\partial_q Q & \partial_q P \\
\partial_p Q & \partial_p P
\end{pmatrix} \begin{pmatrix}
-i\text{Id}_3 \\
\text{Id}_3
\end{pmatrix}.
\]

The following compact notation will be useful hereafter.

**Definition 3.2.** For \( \mathbf{a} \in C^\infty(\Omega, \mathbb{C}) \), define for \( k \in \mathbb{N} \)

\[
\Lambda_k, \Omega(\mathbf{a}) := \max_{|\alpha_q| + |\alpha_p| = k} \sup_{(q,p) \in \Omega} |\partial_q^{\alpha_q} \partial_p^{\alpha_p} \mathbf{a}(q,p)|,
\]

with \( \alpha_q \), and \( \alpha_p \) being multi-indices corresponding to \( q \) and \( p \) respectively. By convention, we denote \( \Lambda_k = \Lambda_k, \mathbb{R}^{2d} \). We will also need the following technical lemma.

**Lemma 3.1.** \( Z^{\kappa_{p,s} \cdot t} \) is invertible for \((q,p) \in \mathbb{R}^{2d} \) with \(|p| > 0 \). Furthermore, for any \( k \geq 0 \) and \( \delta > 0 \), there exist constants \( C_{k, \delta} > 0 \) such that

\[
\Lambda_k (\kappa_{p,s} \cdot t) ((Z^{\kappa_{p,s} \cdot t}(q,p))^{-1}) \leq C_{k, \delta}.
\]

**Proof.** The proof is based on the property of symplectic transforms to bound the eigenvalues of \( Z^{\kappa_{p,s} \cdot t}(Z^{\kappa_{p,s} \cdot t})^\ast \); see Lemma 5.1 in [13] for details. In the following, we introduce the notation, \( f \sim g \)

\[
\int_{\mathbb{R}^{2d}} f(y)G^e(t,x,y,q,p)dydqdy = \int_{\mathbb{R}^{2d}} g(y)G^e(t,x,y,q,p)dydqdy.
\]

(3.12)
We then have

**Lemma 3.2.** For any vector $a(y,q,p) = (a_j)$, matrix $M(y,q,p) = (M_{jk})$, and tensor $T(y,q,p) = (T_{ijk})$ in Schwartz class, one has the following integration by parts formula in the componentwise form, with $\partial_z = (\partial_{z_1}, \partial_{z_2}, \partial_{z_3})$,

\[
a_j(x - Q)_j \sim - \varepsilon \partial_{z_m} (a_j Z_{jm}^{-1}),
\]

\[
(x - Q)_j(x - Q)_k M_{jk} \sim \varepsilon \partial_{z_n} Q_j M_{jk} Z_{kn}^{-1} + \varepsilon^2 \partial_{z_m} (\partial_{z_n} (M_{jk} Z_{kn}^{-1}) Z_{jm}^{-1}),
\]

\[
(x - Q)_i(x - Q)_j(x - Q)_k T_{ijk} \sim - \varepsilon^2 \partial_{z_n} (\partial_{z_l} Q_j T_{ijk} Z_{il}^{-1} Z_{kn}^{-1})
\]

\[
- \varepsilon^2 \partial_{z_m} (\partial_{z_n} Q_j \partial_{z_l} (T_{ijk} Z_{il}^{-1}) Z_{kn}^{-1})
\]

\[
- \varepsilon^3 \partial_{z_m} (\partial_{z_n} (\partial_{z_l} (T_{ijk} Z_{il}^{-1}) Z_{kn}^{-1}) Z_{jm}^{-1})
\].

**Proof.** The proof requires integration by parts and invertibility of $Z$ from Lemma (3.1). This is a special case of Lemma 5.2 in [13], and it is also Lemma 3.2 in [12], where the reader could find detailed proofs. So far, we have introduced some important notations and relatively standard estimates for the analysis of FGA. In the following, we present new estimates valid for the non-strictly hyperbolic elastic wave equations.

**3.2. Next order corrections of FGA for EWS**

According to [6], the first order FGA is

\[
u_{F,0}(t,x) = (2\pi \varepsilon)^{-3d/2} \int_{\mathbb{R}^{3d}} \sum_{b=\pm} a_{p,b,0}(t,y,q,p) G^\varepsilon_{p,b}(t,x,y,q,p) dy dp dq.
\]

(3.13)

As the computations for the branches and P-, S- wavefields are similar, in the following we will either omit the subscript, or will simply subscript using the index $n = \{1, \ldots, 6\}$ instead of $(p, \pm, sh, sh)$ or $(p, \pm, s, s)$. With this notation, we can define eq. (3.13) more compactly, as

\[
u_{F,0}(t,x) = (2\pi \varepsilon)^{-3d/2} \int_{\mathbb{R}^{3d}} \sum_{n=1}^6 a_{n,0}(t,y,q,p) G^\varepsilon_n(t,x,y,q,p) dy dp dq.
\]

(3.14)

For $k > 1$, define the $k$-th ordered FGA with a correction term as

\[
u_{F,k}(t,x) = \nu_{F,0}(t,x) + (2\pi \varepsilon)^{-3d/2} \int_{\mathbb{R}^{3d}} \sum_{j=1}^{k,6} \varepsilon^j (a_{n,j}(t,y,q,p) + a_{n,j}^+(t,y,q,p))
\]

\[\times G^\varepsilon_n(t,x,y,q,p) dy dp dq,
\]

(3.15)

where the terms $a_{n,1}$ will be defined later. We next define a standard smooth cutoff function $\chi_\delta : \mathbb{R}^{2d} \to [0,1]$ for the set $K_\delta$ as

\[
\chi_\delta(q,p) = \begin{cases} 
0, & (q,p) \in \mathbb{R} \setminus K_\delta/2, \\
1, & (q,p) \in K_\delta . 
\end{cases}
\]

(3.16)

and such that for any $k \in \mathbb{N}$, there exists a constant $C_{K_\delta}$ such that

\[
\Lambda_k \left( \chi_\delta(q,p) \right) < C_{K_\delta}.
\]

(3.17)
We define the filtered version of the FGA, as follows for $k \in \mathbb{N}$,

$$
\bar{u}_{\mathcal{F},k}(t,x) = (2\pi \varepsilon)^{-3d/2} \int_{\mathbb{R}^{3d}} \chi_{\delta}(q,p) \sum_{j=0, n=1}^{k,6} \varepsilon^j (a_{n,j}(t,y,q,p) + a_{n,j}^+(t,y,q,p)) \times G_n^\varepsilon(t,x,y,q,p) dy dp dq,
$$

(3.18)

with $a_{n,0}^+ = 0$. In the following we detail the construction of $a_{n,1}^+$. Define the unit vectors $\hat{N}_{p\pm}, \hat{N}_{sv\pm}, \hat{N}_{sh\pm}$ that points in the direction $P$, SV, or SH waves respectively. Then $a_{n,0}(t,y,q,p)$ is defined as follows,

$$
a_{n,0}(t,y,q,p) = a_{n,0}(t,q,p) \alpha_n^\varepsilon(y,q,p) \hat{N}_n(t,q,p),
$$

(3.19)

where $\hat{N}_n(0,q,p) = \hat{n}_n$ and $\alpha^\varepsilon$ incorporates the initial conditions,

$$
\alpha_n^\varepsilon(y,q,p) = \frac{1}{2c_n|P|^2} (u_0^\varepsilon(y)|P| \pm i \varepsilon u_1^\varepsilon(y)) \cdot \hat{n}_n.
$$

(3.20)

The scalar functions $a_{n,0}$, with $n$ representing for $(p\pm, sv\pm, sh\pm)$ satisfy the following evolution equations [6],

$$
\frac{d a_p}{dt} = a_p \left( \pm \frac{\partial Q_p c_p \cdot P_p}{|P_p|^2} + \frac{1}{2} \text{Tr} \left( Z_{p}^{-1} \frac{d Z_{p}}{dt} \right) \right),
$$

(3.21)

$$
\frac{d a_{sv}}{dt} = a_{sv} \left( \pm \frac{\partial Q_{sv} c_s \cdot P_s}{|P_s|^2} + \frac{1}{2} \text{Tr} \left( Z_{sv}^{-1} \frac{d Z_{sv}}{dt} \right) \right) - a_{sh} \frac{d \hat{N}_{sh}}{dt} \cdot \hat{N}_{sv},
$$

(3.22)

$$
\frac{d a_{sh}}{dt} = a_{sh} \left( \pm \frac{\partial Q_{sh} c_s \cdot P_s}{|P_s|^2} + \frac{1}{2} \text{Tr} \left( Z_{sh}^{-1} \frac{d Z_{sh}}{dt} \right) \right) + a_{sv} \frac{d \hat{N}_{sh}}{dt} \cdot \hat{N}_{sv}.
$$

(3.23)

Eq. (3.20) is derived from $u_0(x), u_1(x)$ written in terms of FBI an inverse FBI transforms, i.e.,

$$
u_0(x) = (2\pi \varepsilon)^{-3d/2} \int_{\mathbb{R}^{3d}} u_0(y) G^\varepsilon(0,x,y,q,p) dy dp dq,
$$

(3.24)

and decomposing the integrand in terms of the basis $\{ \hat{n}_p, \hat{n}_{sv}, \hat{n}_{sh} \}$.

**Remark 3.1.** It is easy to check that $u_{\mathcal{F}}(0,x) = u_0(x)$ as $\mathcal{F}^*(\mathcal{F}(u_0)) = u_0$.

Following the same derivation strategy as in [6], and after lengthy calculations, one arrives at, in component form and omitting the $n$ index,

$$
2c|P| \frac{da_1}{dt} = 2a_1 c \partial c_j P_j |P| + a_1 c |P| \left( Z_{jk}^{-1} \frac{d Z_{jk}}{dt} \right)
$$

$$
- i(a_0)_{lt} - ia_0 - \partial_k \left( 4 \left( a_0 \frac{P_l}{|P|} \right)_t P_l \partial c_j Z_{jk}^{-1} \right) \frac{P_t}{|P|}
$$

$$
- 2(a_0)_{lt} |P| \left( 3 \frac{\partial c_j \partial c_k}{c} - \partial^2 c_{jk} \right) Z_{jk}^{-1} \partial_l(Q_k)
$$

$$
+ \partial_k \left( a_0 \frac{P_t}{|P|} M_{jk} Z_{jk}^{-1} \right) \frac{P_l}{|P|} + a_0 N_{jk} Z_{jk}^{-1} \partial_l(Q_k)
$$

$$
+ 2i \partial_n \left( a_0 \frac{P_l}{|P|} \left( \partial c_l \partial c_j \partial c_k |P|^2 \right) + 2i \partial c_l \partial c_k P_j - c \partial c_l P_j \frac{P_k}{|P|^2} \right) Z_{jm}^{-1} \partial_m(Q_k) Z_{ln}^{-1} \frac{P_t}{|P|},
$$

(3.25)
where

\[ M_{jk} = 2\partial_{c_j} |P| + 2i\frac{P_j}{|P|} + \partial_{c_j} \partial_{c_k} P_k + c \partial^2_{c_k} P_k + i\frac{2c \partial_{c_k} P_k P_j}{|P|^2} + ic \partial_{c_j}, \] (3.25)

\[ N_{jk} = 4 \frac{\partial_{c_j} \partial_{c_k}}{c} |P| + 4i \frac{\partial_{c_j} P_j}{|P|} - \partial^2_{c_j k} \partial_c n P_n - ic \partial^2_{c_j k} \]

\[ + \frac{\partial_{c_j} \partial_{c_k} \partial_c n P_n}{c} + 2\partial_{c_j} \partial^2_{c_n k} P_n + 4i \frac{\partial_{c_j} \partial_c n P_j P_n}{|P|^2} + i\partial_{c_j} \partial_{c_k}. \] (3.26)

For the convenience of estimating the intraband transitions, we define the operators \( \mathcal{L}_0, \mathcal{L}_1, \mathcal{L}_2 \) acting on \( A \) as

\[ \mathcal{L}_0(A) := -\rho A (P \cdot Q_t)^2 + (\lambda + \mu)(A \cdot P) P + \mu (P \cdot P) A, \] (3.27)

\[ \mathcal{L}_1(A) := \lambda A + 2i A_t P \cdot Q_t - i\rho A (P_t - iQ_t) \cdot Q_t, \]

\[ -i(\lambda + \mu) \partial_z : (Z^{-T} A P) - i(\lambda + \mu) \partial_z : (A \cdot P Z^{-T}) \]

\[ + \partial_z : (Z^{-T} \partial_Q \rho A (P \cdot Q_t)^2) - 2\mu i \partial_z : (Z^{-T} P A) \]

\[ - \rho A \text{Tr} ((P_t - iQ_t) \otimes (P_t - iQ_t) Z^{-1} \partial_z Q) \]

\[ - (\lambda + \mu) Z^{-1} : (\partial_z Q A) - \mu A \text{Tr} (Z^{-1} \partial_z) \]

\[- \frac{1}{2} \partial^2_{QQ} \rho A (P \cdot Q_t)^2 \text{Tr}(Z^{-1} \partial_z), \] (3.28)

\[ \mathcal{L}_2(A) := \rho A_{tt} - \rho A_t P \cdot Q_t - \varepsilon \rho \partial_z : (\rho Z^{-T} M_1) + \partial_z : (Z^{-T} \partial_Q \rho (P_t - iQ_t) \cdot Q_t) A \]

\[ + \partial_z : (\rho A (P_t - iQ_t) \otimes (P_t - iQ_t) Z^{-1}) \]

\[ - (\lambda + \mu) \partial_z : (\partial_z (Q A) Z^{-1}) \]

\[- \mu \partial_z : (\partial_z (Z^{-1}) \partial_Q \rho M_t \text{Tr} (Z^{-1} \partial_z) \]

\[- \frac{1}{2} \partial^2_{QQ} \rho A (P_t - iQ_t) \cdot Q_t \text{Tr} (Z^{-1} \partial_z) \]

\[ - \partial_z : (Z^{-1} \partial^2_{QQ} \rho A) (P \cdot Q_t)^2 Z^{-1} \]

\[ + \partial_z : \left( \text{Tr} (\partial^2_{QQ} \rho Z^{-1} \partial_z Q) (P \cdot Q_t) \right) \left( \left( P_t - iQ_t \right) A \right). \] (3.29)

Now \((\partial^2_t - \mathcal{L}) u_{F,1}\) can be written as, with \( \mathcal{L} \) given by (1.2),

\[(\partial^2_t - \mathcal{L}) u_{F,1} = (2\pi \varepsilon)^{-3/2} \sum_n \int_{\mathbb{R}^{3d}} (\varepsilon^{-2} L_{n,0}(a_{n,0} + \varepsilon a_{n,1}) \]

\[\varepsilon^{-1} L_{n,1}(a_{n,0} + \varepsilon a_{n,1}) + L_{n,2}(a_{n,0} + \varepsilon a_{n,1}) G^{\varepsilon}_{n} d\mathbf{y} dq dp.\] (3.30)

Substituting the dynamics for \( L_{n,0} \) reveals that \( L_{n,0}(a_{n,0}) = 0. \) Looking at the \( O(1/\varepsilon) \) term and equating to zero gives

\[ L_{n,1}(a_{n,0}) = -L_{n,0}(a_{n,1}). \] (3.31)

Now \( L_{n,0} \) is defined as

\[ L_{n,0} = (\mu |P_n|^2 - \rho (P_n \cdot Q_n)^2) \text{Id}_3 + (\lambda + \mu) P_n \otimes P_n, \] (3.32)
which is a symmetric matrix with eigenvalues
\[ \beta_{n,1} = (\lambda + 2\mu)|P_n| - \rho|P \cdot \partial_t Q_n|^2, \]
\[ \beta_{n,2} = \mu|P|^2 - \rho|P_n \cdot \partial_t Q_n|^2, \]
\[ \beta_{n,3} = \mu|P|^2 - \rho|P_n \cdot \partial_t Q_n|^2, \]
and the corresponding eigenvectors
\[ P_n = (p_{n,1}, p_{n,1}, p_{n,1}), \]
\[ d_{n,1} = (-p_{n,2}, p_{n,1}, 0), \]
\[ d_{n,2} = (-p_{n,3}, 0, p_{n,1}). \]
For the P-wave, \( n = p \), taking inner product of with the eigenvectors brings
\[ \langle P_p, L_{p,0}(a_{n,1}) \rangle = -\langle P_p, L_{p,1}(a_{n,0}) \rangle, \]
which yields
\[ \langle L^*_{p,0}(P_p), a_{n,1} \rangle = \langle L_{p,0}(P), a_{p,1} \rangle = ((\lambda + 2\mu)|P|^2 - \rho|P_p \cdot \partial_t Q_p|^2)\langle P_p, a_{p,1} \rangle = 0. \]
After plugging in (3.2) one can recover the equation (3.21) by
\[ \langle P_p, L_{p,1}(a_{p,0}) \rangle = 0. \]
Considering \( d_{1,2} \),
\[ \langle d_{1,2}, L_{p,0}(a_{n,1}) \rangle = -\langle d_{1,2}, L_{p,1}(a_{p,0}) \rangle. \]
Then
\[ \langle L^*_{p,0}(d_{1,2}), a_{p,1} \rangle = \langle L_{p,0}(d_{1,2}), a_{p,1} \rangle = (\mu|P|^2 - \rho|P \cdot \partial_t Q|^2)\langle d_{1,2}, a_{p,1} \rangle. \]
Plugging in the Hamiltonian flow (3.2) gives
\[ \langle d_{1,2}, a_{p,1} \rangle = \frac{1}{\rho(c_s^2 - c_p^2)}\langle P \rangle \langle d_{1,2}, L_{p,1}(a_{p,0}) \rangle. \]
Define the pseudo-inverse, for \( v \in \mathcal{S}(\mathbb{R}^2) \),
\[ L_{p,0}^{-1}(v) = \frac{1}{\rho(c_s^2 - c_p^2)}\langle P \rangle \left( \langle d_1, v \rangle \hat{d}_1 + \langle d_2, v \rangle \hat{d}_2 \right), \]
and define
\[ a_{p,1}^+ v = L_{p,0}^{-1}((\text{Id} - \Pi_p)L_{p,1}(v)), \]
where \( \Pi_p \) is projection onto \( P_p \).
For the S-wave, with \( n = sv, sh \), from (3.31) one has
\[ L_{s,1}(a_{s,1}) = -L_{s,0}(a_{sh,0} + a_{sv,0}). \]
Let \( d_{s,1} = \tilde{N}_{sh} \), taking inner product with (3.47) gives
\[ \langle L_{s,0}(\tilde{N}_{sv}), a_{s,1} \rangle = (\mu|P_s|^2 - \rho|P_s \cdot \partial_t Q_s|^2)\langle \tilde{N}_{sv}, a_{s,1} \rangle = 0. \]
which is zero when the dynamics are substituted. From this one can get
\[ \langle \hat{N}_{sv}, L_{s,1}a_{sv,0} \rangle = -\langle \hat{N}_{sv}, L_{s,1}a_{sh,0} \rangle, \] (3.49)
which gives us eq. (3.22). Note that eq. (3.23) can be recovered in a similar manor. Taking inner product with \( P_s \) of (3.47) leads to
\[ \langle P, a_{s,1} \rangle = -\frac{1}{(\lambda + \mu)|P|^2} \langle P, L_{s,1}(a_{sv,0} + a_{sh,0}) \rangle. \] (3.50)
Define the pseudo-inverse, for \( v \in S(\mathbb{R}^3) \),
\[ L_{s,0}^{-1}(v) = -\frac{1}{(\lambda + \mu)|P|^2} \langle \hat{P}_s, v \rangle \hat{P}_s, \] (3.51)
and then define
\[ a_{s,1}^\perp v = L_{s,0}^{-1}((\text{Id} - \Pi_s)L_{s,1}(v)), \] (3.52)
with \( \Pi_s \) a projection onto the spanned space by \( d_{s,1} \) and \( d_{s,2} \).

In the next section, we will bound these pseudo-inverses and then further reach the convergence results.

3.3. Error Estimates and Main Result for EWS

Thanks to the above computation in particular the explicit expression of \( a_p^\perp, a_s^\perp, L_{p,0}^{-1}, \) and \( L_{s,0}^{-1} \), we derive some error estimates and eventually conclude on the convergence of FGA for the EWS.

DEFINITION 3.3. Define the scaled semi-norm
\[ \| u(t, \cdot) \|_E = \varepsilon (\| \partial_t u(t, \cdot) \|_{L^2} + \| \nabla \cdot u(t, \cdot) \|_{L^2} + \| \nabla \times u(t, \cdot) \|_{L^2}). \] (3.53)

PROPOSITION 3.3. Let \( a_n = a_{sv}a_{sv}\hat{N}_{sv} + a_{sh}a_{sh}\hat{N}_{sh} \) and \( a_p = a_p\alpha_p\hat{N}_p \). The terms \( a_p, a_s \) are bounded in the \( L^2 \) sense; furthermore,
\[ \| u_{F,1} - u_{F,0} \|_E \leq \varepsilon C_{T,\delta}. \] (3.54)

Proof. First, we remark that
\[ \| a_n(t, \cdot) \|_{L^2} \leq \| a_n \|_{L^2} \| a_n(t, \cdot) \|_{L^\infty} \text{ and } \| a_n(t, \cdot) \|_{L^\infty} \leq \| a_n(t, \cdot) \|_{L^\infty}. \] (3.55)
From the definitions we have an immediate bound
\[ \| u_{F,1}(t, \cdot) - u_{F,0}(t, \cdot) \|_E \leq (2\pi \varepsilon)^{-3d/2} \sum_n \varepsilon \| \int_{\mathbb{R}^3} a_{n,1}^\perp + a_{n,1}G_n dy dq dp \|_E. \] (3.56)
Applying the derivatives with Proposition 2.2, we have the estimate
\[ \| u_{F,1}(t, \cdot) - u_{F,0}(t, \cdot) \|_E \leq \varepsilon C \sum_n \| a_{n,1}^\perp(t, \cdot) + a_{n,1}(t, \cdot) \|_{L^\infty}. \] (3.57)
The estimate of (3.56) then follows directly from Proposition 2.2. We need to bound the prefactor terms, we note that on the compact set \( K_\delta \) the bound for the prefactor terms fall from Lemma 5.4 in [13]. We go through several of the bounds here, starting with
the P-wave and dropping the subscripts as the calculations are the same and setting $P = P_p, Q = Q_p$,

$$
\partial_t a_{p,0} = a_{p,0} \left( \frac{\partial Q c_p \cdot P}{|P|} + \frac{1}{2} \text{Tr} \left( Z^{-1} \partial_t Z \right) \right),
$$

(3.58)

$$
\partial_t a_{p,1} = a_{p,1} \left( \frac{\partial Q c_p \cdot P}{|P|} + \frac{1}{2} \text{Tr} \left( Z^{-1} \partial_t Z \right) \right) + F_p(a_{p,0}, \partial_z a_{p,0}, Q, P, c_p).
$$

(3.59)

With $F_p$ being a continuously differentiable function in its arguments for $P, Q \in K_{\delta T}$. Equation (3.58) immediately implies:

$$
\partial_t |a_{p,0}| \leq |a_{p,0}| \left| \frac{\partial Q c_p \cdot P}{|P|} + \frac{1}{2} \text{Tr} \left( Z^{-1} \partial_t Z \right) \right|. \tag{3.60}
$$

An application of Gronwall’s inequality gives

$$
\sup_{t \in [0, T]} \Lambda_0, K_{\delta/2} (a_{p,0}(t, q, p)) \leq C_{\delta, T}. \tag{3.61}
$$

To bound eq. (3.59), $\partial_z a_{p,0}$ needs to be bounded, but with partial $z$ of (3.58) using a similar inequality as (3.60) and taking Gronwall’s inequality we have,

$$
\sup_{t \in [0, T]} \Lambda_1, K_{\delta/2} (a_{p,0}(t, q, p)) \leq C_{\delta, T}. \tag{3.62}
$$

The function $F_p(a_{p,0}, \partial_z a_{p,0}, Q, P, c_p)$ is differentiable with differentiable arguments on the compact set $K_{\delta/2}$. Combining (3.61), (3.62) and Gronwall’s inequality to (3.59) we see that $\|a_p P\|_{L^2}$ is bounded on $[0, T] \times K_{\delta/2}$.

For the S-wave terms, again using the short notation $P = P_s, Q = Q_s$ and dropping the brach subscript, we can write the system ((3.22), (3.23)) as

$$
\frac{da}{dt} = M(t) a, \tag{3.65}
$$

where $m_\pm = \partial_t \hat{N}_{sh} \cdot \hat{N}_{sv}$ and

$$
h_\pm = 2 \partial Q s_{s, \pm} c_s \cdot \hat{N}_{sh} + a^s \text{Tr} \left( Z^{-1} \partial_t Z^s \right). \tag{3.64}
$$

Denote $M$ as the matrix in eq. (3.63), and $a = (a^{sv}, a^{sh})^T$. Then the system can be recast as

$$
\frac{da}{dt} = M(t) a. \tag{3.65}
$$

Solving for the eigenvalues:

$$
\lambda_{sh, sv}(t) = -\partial Q c_s \cdot \hat{N}_{sh, sv} - \frac{1}{2} \text{Tr} \left( Z_s^{-1} \frac{dZ_s}{dt} \right) \pm i \frac{d\hat{N}_{sv, sv}}{dt} \cdot \hat{N}_{sh, sv}. \tag{3.66}
$$

To see that the latter are bounded, simply note that a smooth $\{ \hat{N}_p, \hat{N}_sh, \hat{N}_{sh} \}$ form an orthonormal frame, and hence the last term in (3.66) is bounded for all $t \geq 0$. 

Note that
\[
\text{Tr}(Z_s^{-1} \frac{dZ_s}{dt}) = \frac{1}{\text{det}(Z_s)} \frac{d\text{det}(Z_s)}{dt},
\]
then by (3.1) we have a bound for \(\text{det}(Z_s)\) so eq. (3.67) is bounded for all \(t \geq 0\). Notice that
\[
\frac{\partial_q H_s \cdot \partial_p H_s}{H_s} = -\frac{\partial Q_s c_s \cdot P_s}{|P_s|},
\]
then with (3.4), eq. (3.68) is bounded for all \(t \geq 0\). Now the eigenvalues in eq. (3.66) are bounded for all \(t \geq 0\). So we have
\[
\sup_{t \in [0,T]} \Lambda_{0, K_s/2}(a_{s,0}(t, q, p)) \leq C_{\delta, T}.
\]
For \(\partial_t a_{s,1}\), we can write the system
\[
\frac{d}{dt} \begin{pmatrix} a_{sv,1} \\ a_{sh,1} \end{pmatrix} = \frac{1}{2} \begin{pmatrix} h & m \\ -m & h \end{pmatrix} \begin{pmatrix} a_{sv}^+ \\ a_{sh}^+ \end{pmatrix} + F_s(a_{s,0}, \partial_s a_{s,0}, Q, P, c_s),
\]
with \(F_s\) being a continuously differentiable function in its arguments for \(P, Q \in K_{\delta, T}\). The bounds follow in a similar fashion from previous work, we arrive at
\[
\sup_{t \in [0,T]} \Lambda_{1, K_s/2}(a_{s,1}(t, q, p)) \leq C_{\delta, T},
\]
which gives the needed result. \(\square\)

**Proposition 3.4.** For any \(T > 0\) and \(t \in [0,T]\)
\[
\|u_{F,1}(t, \cdot) - \tilde{u}_{F,1}(t, \cdot)\|_E = O(\varepsilon^\infty).
\]

**Proof.** Starting from the definition of \(u_{F,1}\) and \(\tilde{u}_{F,1}\), we get
\[
\|u_{F,1} - \tilde{u}_{F,1}\|_E 
\leq (2\pi\varepsilon)^{-3d/2} \int_{\mathbb{R}^{3d}} \left\| \sum_n (1 - \chi_{\delta})(a_{n,0} + \varepsilon a_{n,1}) e^{\pm p_n} \right\|_E \, dp \, dq 
\leq 2^{-d/2} \sum_n \left\| (1 - \chi_{\delta})(a_{n,0} + \varepsilon a_{n,1}) \tilde{N}_n F a_n \right\|_E
\leq 2^{-d/2} \left( \|F^e u_0^\varepsilon\|_{L^2(\mathbb{R}^n \setminus K_{\delta})} + \varepsilon \|F^e u_1^\varepsilon\|_{L^2(\mathbb{R}^n \setminus K_{\delta})} \right) \sum_n \left\| (1 - \chi_{\delta})(a_{n,0} + \varepsilon a_{n,1}) \tilde{N}_n \right\|_L \leq C_{\delta, T} \left( \|F^e u_0^\varepsilon\|_{E(\mathbb{R}^n \setminus K_{\delta})} + \varepsilon \|F^e u_1^\varepsilon\|_{E(\mathbb{R}^n \setminus K_{\delta})} \right) 
\leq C_{\delta, T} \left( \|F^e u_0^\varepsilon\|_{L^2(\mathbb{R}^n \setminus K_{\delta})} + \varepsilon \|F^e u_1^\varepsilon\|_{L^2(\mathbb{R}^n \setminus K_{\delta})} \right) = O(\varepsilon^\infty).
\]

Where the second equality is from Proposition 2.2, the third inequality is by similar arguments found in Proposition 3.3. Also from eq. (3.20), with the direct bound
\[
\|a_n\|_E \leq C(\|u_0^\varepsilon\|_E + \varepsilon \|u_1^\varepsilon\|_E).
\]
The last inequality is justified noticing the derivatives do not affect the initial conditions only \(G_n^\varepsilon\) in \(F^e u_0^\varepsilon\) and \(F^e u_1^\varepsilon\). \(\square\)
Proposition 3.5. The operators $\mathcal{L}_0, \mathcal{L}_1, \mathcal{L}_2$ are bounded. That is, for a given $T$ and any $t \in [0, T]$, $a \in C^\infty([0, T]) \times S(\mathbb{R}^{2d})$ and for $k = 0, 1, j = 0, 1, 2$, 
\[
\sup_{t \in [0, T]} \Lambda_{k, K_s}(\mathcal{L}_j(a)) < C_{T, K_s} \text{ and } \|\mathcal{L}_j(a(t, \cdot))\|_{L^\infty} < C_{T, \delta}.
\] (3.77)

Proof. Notice $\mathcal{L}_{n,j}$ depend on $P_n, Q_n, Z_n^{-1}$ and its derivatives, which are all bounded on $[0, T] \times K_\delta$, and this gives the result. □

Proposition 3.6. For a given $T$ and any $t \in [0, T]$, for $k = 0, 1$ and $a \in C^\infty([0, T]) \times S(\mathbb{R}^{2d})$, we have 
\[
\Lambda_{k, K_s}(a_{n,1}^+(a(t, \cdot))) < C_{T, K_s} \text{ and } \|a_{n,1}^+(a(t, \cdot))\|_{L^\infty} < C_{T, \delta}.
\] (3.78)

Proof. For both $a_p^+$ and $a_s^+$, the pseudo-operators $\mathcal{L}_{p,0}^{-1}$ and $\mathcal{L}_{s,0}^{-1}$ are bounded on $K_\delta$ as $|P| > 0$ and $|\mathcal{L}_{n,0}^{-1}|_{L^\infty} \leq C\delta^{-1}$. Then by Proposition 3.5 we have the following result. □

Proposition 3.7. Consider the Elastic wave equations with a forcing term, $x \in \mathbb{R}^d$, 
\[
\begin{aligned}
\rho(x)\partial_t^2 u - (\lambda + 2\mu)\nabla(\nabla \cdot u) + \mu \nabla \times \nabla \times u &= F(t, x), \\
u(0, x) &= u_0^0, \\
u(t, x) &= u_1^t.
\end{aligned}
\] (3.79)

Let $T > 0$, and let $u_0(t, x) \in C^\infty([0, T]) \times H_0^1(\mathbb{R}^d)$. For each $t \in [0, T]$, we have the following estimate: 
\[
\|\rho \partial_t u\|_{L^2} + \|\lambda + 2\mu\|_{L^2} + \|\mu \nabla \times u\|_{L^2} \leq C_T \left(\frac{1}{\varepsilon}\|u(0, \cdot)\|_E + \int_0^t \|F(s, \cdot)\|_{L^2} \, ds\right).
\] (3.80)

In particular, 
\[
\sup_{t \in [0, T]} \|u(t, \cdot)\|_E \leq C_T \left(\|u_0^0\|_E + \varepsilon \int_0^T \|F(s, \cdot)\|_{L^2} \, ds\right).
\] (3.81)

Proof. This is a standard estimate. Dotting eq. (3.80) with $\partial_t u$ and integrating over space we have 
\[
\frac{1}{2} \int_{\mathbb{R}^d} \rho(x)|\partial_t u|^2 + (\lambda + 2\mu)|\nabla \cdot u|^2 + \mu|\nabla \times u|^2 \, dx \leq \int_{\mathbb{R}^d} |\partial_t u \cdot F| \, dx.
\] (3.82)

The right-hand side can then be estimated by 
\[
\int_{\mathbb{R}^d} |\partial_t u \cdot F| \, dx \leq \frac{1}{2} \int_{\mathbb{R}^d} |\partial_t u|^2 + |F|^2 \, dx.
\] (3.83)

Adding the missing terms to apply Gronwall’s inequality gives the bound 
\[
e^t\left(\|\rho(x)u_1\|^2_{L^2} + \|\lambda + \mu\|_{L^2} \cdot \|u_0\|_{L^2} + \|\mu \nabla \times u\|_{L^2}^2\right) + \int_0^t e^{t-s} \int_{\mathbb{R}^d} |F(s, x)|^2 \, dx \, ds.
\] (3.84)

Taking the maximum over $\rho, \lambda, \mu$ and over $T$, we arrive at the estimate. □

Proposition 3.8. We have 
\[
\|(\partial_t^2 - \mathcal{L})\bar{u}_F, 1\|_E \leq C_{T, \delta}.
\] (3.85)
Proof. Plugging $\tilde{u}_{F,1}$ into (1.1) gives,

$$(\partial_t^2 - \mathcal{L})\tilde{u}_{F,1}(t, x)$$

$$=(2\pi \varepsilon)^{-3d/2} \int_{\mathbb{R}^{3d}} \chi_\delta(q, p) \sum_{n, m} \varepsilon^{m-2} \mathcal{L}_{n, m} \left( a_{n, 0} + \varepsilon(a_{n, 1} + \varepsilon a_{n, 1}^\perp) \right) G_n^\varepsilon dy dp dq. \quad (3.86)$$

Expanding and simplifying the above equation yield

$$(\partial_t^2 - \mathcal{L})\tilde{u}_{F,1}(t, x) = (2\pi \varepsilon)^{-3/2} \int_{\mathbb{R}^{3d}} \chi_\delta(q, p) G_n^\varepsilon dy dq dp \varepsilon^{-2} \mathcal{L}_{n, 0}(a_{n, 0})$$

$$+ \varepsilon^{-1} \mathcal{L}_{n, 0}(a_{n, 1}) + \varepsilon^{-1} \mathcal{L}_{n, 1}(a_{n, 0}) + \mathcal{L}_{n, 2}(a_{n, 0} + \varepsilon a_{n, 1})$$

$$- (\varepsilon^{-1} \mathcal{L}_{n, 0} + \mathcal{L}_{n, 1} + \varepsilon \mathcal{L}_{n, 2}) \mathcal{L}_{n, 0}^{-1}((\text{Id} - \Pi_n)(\mathcal{L}_{n, 1}(a_{n, 0} \tilde{N}_n))). \quad (3.87)$$

Direct cancellation gives

$$(\partial_t^2 - \mathcal{L})\tilde{u}_{F,1}(t, x) = (2\pi \varepsilon)^{-3/2} \int_{\mathbb{R}^{3d}} G_n^\varepsilon dy dq dp R_0(t, p, q) + \varepsilon R_1(t, p, q). \quad (3.88)$$

where

$$R_0(t, p, q) = \mathcal{L}_{n, 1}(a_{n, 1}) + \mathcal{L}_{n, 2}(a_{n, 0}) - \mathcal{L}_{n, 1} a_{n, 1}^\perp, \quad (3.89)$$

$$R_1(t, p, q) = \mathcal{L}_{n, 2}(a_{n, 1}) - \mathcal{L}_{n, 2} a_{n, 1}^\perp. \quad (3.90)$$

Then by Propositions 2.2 and 3.6,

$$\|(\partial_t^2 - \mathcal{L})\tilde{u}_{F,1}(t, \cdot)\|_{L^2} \leq C_{T, \delta}(\|R_0(t, \cdot)\|_{L^\infty} + \varepsilon \|R_1(t, \cdot)\|_{L^\infty}). \quad (3.91)$$

By Propositions 3.5 and 3.6, $R_0$ and $R_1$ are hence bounded. □

PROPOSITION 3.9. Let $u$ solve the Cauchy problem (1.1). If $u_{F,0}$ is the first order FGA (3.13), then we have the following estimate on the initial conditions:

$$\|u(0, x) - u_{F}(0, x)\|_E \leq \varepsilon C_T. \quad (3.92)$$

Proof. First computing the following

$$\partial_t a_n(0, y, q, p) = \alpha(y, q, p) \left( \frac{\partial c(q) \cdot p}{|p|} - d \right) \tilde{n}_n. \quad (3.93)$$

For estimating $u(0, x) - u_{F,0}(0, x)$ in the energy norm, we can write $\partial_t u(0, x) = u_1^\varepsilon(x)$. This gives

$$\int_{\mathbb{R}^{3d}} u_0^\varepsilon(y) \frac{1}{\varepsilon} \Phi_n(0, x, y, q, p) G_n^\varepsilon(0, x, y, q, p) dy dq dp$$

$$= \int_{\mathbb{R}^{3d}} u_1^\varepsilon(y) G_n^\varepsilon(0, x, y, q, p) dy dq dp. \quad (3.94)$$

Then
\[ |u_1^\varepsilon(x) - \partial_t u_{F,0}(0,x)| = (2\pi \varepsilon)^{-d/2} \left| \sum_n \int_{\mathbb{R}^d} \left( \frac{1}{2} (u_1^\varepsilon(y) \cdot \hat{n}) \hat{n} - \partial_t a_n(0, y, q, p) - a_n \frac{i}{\varepsilon} \Phi_n(0, x, y, q, p) \right) G_n^\varepsilon(0, x, y, q, p) \text{ dy dq dp} \right|. \] (3.95)

For one of the right terms, it becomes, after plugging in (3.94),

\[ \int_{\mathbb{R}^d} \left( u_0^\varepsilon(y) \frac{i}{2\varepsilon} \Phi_n(0, x, y, q, p) - \alpha(y, q, p) \left( \frac{\partial c(q) \cdot p}{|p|} - d \right) \right) \hat{n} \]
\[ - 2^{d/2} \alpha(y, q, p) \hat{n} \frac{i}{\varepsilon} \Phi_n(0, x, y, q, p) \times G_n^\varepsilon(0, x, y, q, p) \text{ dy dq dp}. \] (3.96)

Using (3.20) and summing over the wavefields and branches gives

\[ u_1^\varepsilon(x) - \partial_t u_{F,0}(0,x) = - \sum_n \int_{\mathbb{R}^d} \frac{1}{2 c_n(p)^2} \left( u_0^\varepsilon(y) c_n(p) \pm i \varepsilon u_1^\varepsilon(y) \right) \cdot \hat{n} \]
\[ \times \left( \frac{\partial c(q) \cdot p}{|p|} - d \right) \hat{n} G_n^\varepsilon(0, x, y, q, p) \text{ dy dq dp}. \] (3.97)

By Proposition (2.2), we then have the estimate

\[ \| u_1^\varepsilon - \partial_t u_{F,0}(0, \cdot) \|_{L^2} \leq C_T. \] (3.98)

For the Div term,

\[ \nabla \cdot u_F(0, x) = (2\pi \varepsilon)^{-d/2} \sum_n \int_{\mathbb{R}^d} \nabla \cdot \left( a_n(0, y, q, p) G_n^\varepsilon(0, x, y, q, p) \right) \text{ dy dq dp} \]
\[ = (2\pi \varepsilon)^{-d/2} \sum_n \int_{\mathbb{R}^d} \left( \frac{i}{\varepsilon} a_n \cdot p_n + (x - Q_n) \right) G_n^\varepsilon(0, x, y, q, p) \text{ dy dq dp}. \] (3.99)

Applying the operators and integration by parts gives, for one term,

\[ \int_{\mathbb{R}^d} \left( \frac{i}{\varepsilon} a_n \cdot p_n - \partial_z (Z^{-1} a) \right) G_n^\varepsilon(0, x, y, q, p) \text{ dy dq dp}. \] (3.100)

Writing the difference \( \nabla \cdot u_0^\varepsilon - \nabla \cdot u_{F,0}(t, x) \) in terms of the FIO (eq. (2.2)), leads to

\[ (2\pi \varepsilon)^{-d/2} \int_{\mathbb{R}^d} \left( \frac{i}{\varepsilon} (u_0^\varepsilon(y) - a(0, y, q, p)) \cdot p_n + \partial_z (Z^{-1} (u_0^\varepsilon(y) - a(0, y, q, p))) \right) \]
\[ \times G_n^\varepsilon(0, x, y, q, p) \text{ dy dq dp}. \] (3.101)

With \( a(0, y, q, p) = \alpha_n(y, q, p) \hat{n} \) and summing over \( n \) we have

\[ \nabla \cdot u_0^\varepsilon(x) - \nabla \cdot u_{F,0}(0, x) \]
\[ = (2\pi \varepsilon)^{-d/2} \sum_n \int_{\mathbb{R}^d} \partial_z (Z^{-1} (u_0^\varepsilon(y) - \alpha_n(y, q, p) \hat{n})) G_n^\varepsilon(0, x, y, q, p) \text{ dy dq dp}. \] (3.102)

Again, by Proposition (2.2) we arrive at the estimate

\[ \| \nabla \cdot u_0^\varepsilon - \nabla \cdot u_{F,0}(t, \cdot) \|_{L^2} \leq C_T. \] (3.103)
The Curl term has a similar estimate as the Div term. These three estimates show the result. □

THEOREM 3.1. Let \( \{u_0^\pm\} \) be a family of asymptotically high frequency initial conditions, and let \( u \) be solution to the Cauchy problem (1.1). If \( u_{F,0} \) is the first order FGA (3.13), then for a given \( T \) and any \( t \in [0,T] \), \( \delta > 0 \) and sufficiently small \( \varepsilon \), we have

\[
\sup_{t \in [0,T]} \|u(t,\cdot) - u_{F,0}(t,\cdot)\|_E \leq \varepsilon C_{T,\delta}.
\] (3.104)

Proof. By the triangle inequality,

\[
\|u - u_{F,0}\|_E \leq \|u - \tilde{u}_{F,1}\|_E + \|\tilde{u}_{F,1} - u_{F,1}\|_E + \|u_{F,1} - u_{F,0}\|_E.
\] (3.105)

For the first term, we define the quantity \( e = u - \tilde{u}_{F,1} \) and by Propositions 3.7 and 3.8

\[
\|e\|_E \leq C_{T,\delta}(\|e(0,\cdot)\|_E + \varepsilon \int_0^T \|R_0\|_{L^2} ds) + O(\varepsilon^2).
\] (3.106)

Proposition 3.9 then shows that \( \|e(0,\cdot)\|_E \leq \varepsilon C_{T,\delta} \), and Propositions 3.5 and 3.6 show that \( \|R_0\|_{L^2} \leq C_{T,\delta} \). Thus, in eq. (3.105), the first term is estimated at the correct order, the second term is a \( O(\varepsilon^\infty) \) thanks to Proposition 3.4, and the last term is estimated to the desired order by Proposition 3.3. □

4. Convergence analysis for the Dirac system

Since the derivation of the boundedness estimates related to the Hamiltonian flow associated to the FGA formulation of EWS is essentially the same as the FGA formulation of the Dirac system (DS), we shall omit this derivation in this section, and only present the asymptotic corrections and the convergence results for the Dirac system.

4.1. Next order corrections of FGA for the Dirac system

According to [2], the FGA is of the form

\[
u_{F,0}(t,x) = (2\pi\varepsilon)^{-3d/2} \int_{\mathbb{R}^d} \sum_{\pm} a_{\pm,0}(t,y,q,p) G_{\pm}(t,x,y,q,p) dy dp dq,
\] (4.1)

where \( a_{\pm,0} = a_{\pm,1} + a_{\pm,2} \), \( \pm \) indicates the positive/negative eigenvalue. For \( k > 1 \), define the \( k \)-th order FGA with a correction term as

\[
u_{F,k}(t,x) = (2\pi\varepsilon)^{-3d/2} \int_{\mathbb{R}^d} \sum_{j=0}^{k} \sum_{\pm} \varepsilon^j \tilde{a}_{\pm,j}(t,y,q,p) G_{\pm}(t,x,y,q,p) dy dp dq,
\] (4.2)

where \( \tilde{a}_{\pm,j} = \tilde{a}_{\pm,1} + \tilde{a}_{\pm,2} \), \( \tilde{a}_{\pm,m,0} = a_{\pm,m,0} \), and the terms \( \tilde{a}_{\pm,m,j} = a_{\pm,m,j} + a_{\pm,m,j}^\pm \) will be defined later for \( j \geq 1 \). Let \( \Upsilon_{\pm 1} \) and \( \Upsilon_{\pm 2} \) be the normalized eigenvectors corresponding to the eigenvalue \( h_{\pm} \) of the Dirac symbol defined in (1.12). Then \( a_{\pm,0}(t,y,q,p) \) is defined as follows,

\[
a_{\pm,0}(t,y,q,p) = a_{\pm,m,j}(t,q,p) \alpha_{\pm,m}^\varepsilon(y,q,p) \Upsilon_{\pm,m}(t,q,p),
\] (4.3)

where \( \alpha^\varepsilon \) incorporates the initial conditions,

\[
\alpha_{\pm,m}^\varepsilon(y,q,p) = \varphi^{\dagger}_\varepsilon(y) \cdot \Upsilon_{\pm,m}(0,q,p).
\] (4.4)
The computations for the ± branches will be similar, so that we either omit the ± subscript or subscript by \( m \) instead of ±\( m \). The scalar functions \( a_{m,0} \), with \( m=1,2 \), satisfy the following evolution equations [2]

\[
\frac{d}{dt} \begin{pmatrix} a_{1,0} \\ a_{2,0} \end{pmatrix} + \Xi \begin{pmatrix} a_{1,0} \\ a_{2,0} \end{pmatrix} = 0,
\]

where \( \Xi \) is 2 by 2 matrix with elements

\[
\Xi_{mn} = \delta_{mn} \frac{dY_n}{dt} - \delta_{mn} \frac{i}{2} \partial_{z_k} Q_l \partial_{Q_l} \partial_{Q_j} V Z^{-1}_{jk} + \partial_{z_k} Y^\dagger_n F^n_j Z^{-1}_{jk},
\]

and

\[
F^n_j = (\partial_{P_l} h(Q, P) - c\hat{\sigma}_j - i\partial_{Q_j} h(Q, P) + i\partial_{Q_j} B) Y_n.
\]

We define the operators \( \mathcal{L}_{\pm,0}, \mathcal{L}_{\pm,1}, \mathcal{L}_{\pm,2} \) acting on \( a \) as

\[
\mathcal{L}_{\pm,0} a = i(c\hat{\sigma} \cdot P \pm B(Q \pm) + \partial_t S \pm - P \pm \cdot \partial_t Q \pm) a,
\]

\[
\mathcal{L}_{\pm,1} a = \partial_t a - \partial_{z_k} \left[ (\partial_t (Q_{\pm,j} + i P_{\pm,j}) - c\hat{\sigma}_j + i\partial_{Q_j} B) Z^{-1}_{jk} a \right] + \frac{i}{2} \partial_{z_k} Q_{\pm,l} \partial_{Q_l} \partial_{Q_j} B Z^{-1}_{jk} a,
\]

\[
\mathcal{L}_{\pm,2} a = \frac{1}{3i} \partial_{z_m} (\partial_{z_l} Q_j \partial_{Q_j} \partial_{Q_k} B Z^{-1}_{il} Z^{-1}_{kn} a) + \frac{1}{6i} \partial_{z_m} Q_j \partial_{z_l} (\partial_{Q_j} \partial_{Q_k} B Z^{-1}_{il} a) Z^{-1}_{kn} + \frac{\varepsilon}{6i} \partial_{z_m} (\partial_{z_l} (\partial_{Q_j} \partial_{Q_k} B Z^{-1}_{il} a) Z^{-1}_{kn}) Z^{-1}_{jm}.
\]

Looking at the \( O(\varepsilon^0) \) terms, and equating to zero gives

\[
\mathcal{L}_{\pm,0} a_{\pm,0} = 0,
\]

thus

\[
\partial_t S \pm - P \pm \cdot \partial_t Q \pm = -h \pm,
\]

then, together with the Hamiltonian flow, one recovers the evolution equation for the action:

\[
\frac{d}{dt} S = P \cdot \partial_P h(Q, P) - h(Q, P).
\]

Looking at the \( O(\varepsilon^1) \) terms and equating to zero gives

\[
\mathcal{L}_{\pm,1} a_{\pm,0} + \mathcal{L}_{\pm,0} \tilde{a}_{\pm,1} = 0.
\]

Let us take the + branch for example and inner product the above equation with \( Y_{\pm m} \),

\[
Y^\dagger_{\pm m} \mathcal{L}_{\pm,1} a_{\pm,0} = -Y^\dagger_{\pm m} \mathcal{L}_{\pm,0} \tilde{a}_{\pm,1} = (\mathcal{L}_{\pm,0} Y_{\pm m})^\dagger \tilde{a}_{\pm,1} = 0,
\]

from which one recovers eq. (4.5).

Define

\[
\mathcal{L}_{\pm,0}^{-1} = \frac{i}{h_\pm - h_\mp} \left( Y_{\pm 1} Y^\dagger_{\pm 1} + Y_{\pm 2} Y^\dagger_{\pm 2} \right),
\]
which is a pseudo-inverse of $L_{\pm,0}$. We then define
\begin{equation}
\mathbf{a}_{\pm,1}^\dagger = -L_{\pm,0}^{-1}L_{\pm,1}\mathbf{a}_{\pm,0}.
\end{equation}
Looking at the $O(\varepsilon^2)$ terms and equating to zero give
\begin{equation}
L_{\pm,1}(\mathbf{a}_{\pm,1} + \mathbf{a}_{\pm,1}^\dagger) + L_{\pm,2}\mathbf{a}_{\pm,0} + L_{\pm,0}\tilde{\mathbf{a}}_{\pm,2} = 0,
\end{equation}
which implies
\begin{equation}
\frac{d}{dt} \left( \begin{array}{c} a_{1,1} \\ a_{2,1} \end{array} \right) + \Xi \left( \begin{array}{c} a_{1,0} \\ a_{2,0} \end{array} \right) + \left( \begin{array}{c} \Upsilon_{\mp 1} \\ \Upsilon_{\mp 2} \end{array} \right) \left( L_{\pm,1}\mathbf{a}_{\pm,1}^\dagger + L_{\pm,2}\mathbf{a}_{\pm,0} \right) = 0.
\end{equation}

4.2. Error estimate and main result for the Dirac system

**Lemma 4.1.** The pseudo-inverse operator defined in (4.13) is bounded, that is, for each $k \in \mathbb{N}$, there exists a constant $C_k$
\begin{equation}
\Lambda_k \left( L_{\pm,0}^{-1}(q,p) \right) \leq C_k.
\end{equation}

*Proof.* Recall
\begin{equation}
L_{\pm,0}^{-1}(q,p) = \frac{i}{h_\pm(q,p) - h_\mp(q,p)} \left( \Upsilon_{\mp 1}(q,p)\Upsilon_{\mp 1}^\dagger(q,p) + \Upsilon_{\mp 2}(q,p)\Upsilon_{\mp 2}^\dagger(q,p) \right).
\end{equation}
Since $h_+ - h_- \geq 2c^2$, and $h_\mp$ and $\Upsilon_{\mp 1,\pm 2}$ are smooth functions of $(q,p)$, the estimate follows easily. \(\Box\)

**Lemma 4.2.** For any $T > 0$, $k \in \mathbb{N}$, $j = 0, 1$, there exists a constant $C_{k,j,T}$
\begin{equation}
\sup_{0 \leq t \leq T} \Lambda_k(\tilde{\mathbf{a}}_{\pm,j}) \leq C_{k,j,T},
\end{equation}
\begin{equation}
\sup_{0 \leq t \leq T} \Lambda_k(\partial_t \tilde{\mathbf{a}}_{\pm,j}) \leq C_{k,j,T}.
\end{equation}

*Proof.* Noticing that $\Xi$, $\Upsilon$'s, and $L$'s are smooth and bounded, the estimates follow from (4.5), (4.16), and Grownwall’s lemma. \(\Box\)

**Proposition 4.1.** Let $u_{F,0}$ and $u_{F,1}$ be the zeroth and first order FGA solution in (4.1) and (4.2), then for any $T > 0$, there exists a constant $C_T$, such that
\begin{equation}
\sup_{0 \leq t \leq T} \| u_{F,1} - u_{F,0} \|_{L^2} \leq \varepsilon C_T.
\end{equation}

*Proof.* From the definitions we have
\begin{equation}
\| u_{F,1} - u_{F,0} \|_{L^2} \leq (2\pi \varepsilon)^{3d/2} \sum_\pm \varepsilon \left\| \int_{\mathbb{R}^d} \tilde{\mathbf{a}}_{\pm,1}(t,y,q,p)G_\pm^\varepsilon(t,x,y,q,p)dydpdq \right\|_{L^2},
\end{equation}
then by Proposition 2.2,
\begin{equation}
\| u_{F,1} - u_{F,0} \|_{L^2} \leq \varepsilon C \sum_\pm \| \tilde{\mathbf{a}}_{\pm,1}(t,y,q,p) \|_{L^\infty}.
\end{equation}
Therefore by Lemma 4.2, we arrive at the estimate (4.20). □

**Lemma 4.3.** For any $T > 0$, there exists a constant $C_T$, such that for any $\varepsilon > 0$,
\[
\|(\varepsilon \partial_t - D) u_{F,1}\|_{L^2} \leq \varepsilon^2 C_T.
\]

**Proof.** Plugging $u_{F,1}$ into DS gives
\[
(i\varepsilon \partial_t - D) u_{F,1} = (2\pi\varepsilon)^{-3d/2} \int_{\mathbb{R}^d} \sum_{\pm} (\mathcal{L}_{\pm,0} a_{\pm,0} + \varepsilon \mathcal{L}_{\pm,0} \tilde{a}_{\pm,1} + \varepsilon \mathcal{L}_{\pm,1} a_{\pm,0} + \varepsilon^2 \mathcal{L}_{\pm,1} \tilde{a}_{\pm,1}) G_\varepsilon^d dy dq dp.
\]

(4.24)

Note that $\mathcal{L}_{\pm,0} a_{\pm,0} = 0$ since the action equation (4.10) and $a_{\pm,0}$ is in the eigenspace of the Dirac symbol.

To show $\mathcal{L}_{\pm,0} \tilde{a}_{\pm,1} + \mathcal{L}_{\pm,1} a_{\pm,0} = 0$, taking the + for example, it is suffices to show that, for each $m = \pm 1, \pm 2$, we have
\[
\mathcal{Y}^\dagger_m (\mathcal{L}_{\pm,0} \tilde{a}_{\pm,1} + \mathcal{L}_{\pm,1} a_{\pm,0}) = 0.
\]

(4.25)

When $m$ is negative, since $\tilde{a}_{\pm,1} = a_{\pm,1} + a_{\pm,1}^\perp$, and $a_{\pm,1} \in \ker \mathcal{L}_{\pm,0}$, (4.25) is equivalent to
\[
\mathcal{Y}^\dagger_m (\mathcal{L}_{\pm,0} a_{\pm,1}^\perp + \mathcal{L}_{\pm,1} a_{\pm,0}) = 0,
\]

(4.26)

which is valid by the construction of $a_{\pm,1}^\perp$. When $m$ is positive, (4.25) is valid by the evolutionary equation of $a_{m,0}$.

Therefore,
\[
\|(i\varepsilon \partial_t - D) u_{F,1}\|_{L^2} \leq \varepsilon^2 C_T (\|\mathcal{L}_{\pm,1} \tilde{a}_{\pm,1} + \mathcal{L}_{\pm,2} a_{\pm,0}\|_{L^\infty} + \varepsilon \|\mathcal{L}_{\pm,2} \tilde{a}_{\pm,1}\|_{L^\infty}).
\]

(4.27)

By Lemma 4.2, $\|\mathcal{L}_{\pm,1} \tilde{a}_{\pm,1} + \mathcal{L}_{\pm,2} a_{\pm,0}\|_{L^\infty}$ and $\|\mathcal{L}_{\pm,2} \tilde{a}_{\pm,1}\|_{L^\infty}$ are bounded. □

**Proposition 4.2.** Let $u$ be the solution of the DS (1.7) and $u_{F,1}$ be the corresponding FGA solution, then for any $T > 0$, there exists a constant $C_T$, so that for any $\varepsilon > 0$
\[
\sup_{0 \leq t \leq T} \|u - u_{F,1}\|_{L^2} \leq \varepsilon C_T.
\]

(4.28)

**Proof.** Let $e = u - u_{F,1}$, and then by Lemma 4.3,
\[
\|e(t,\cdot)\|_{L^2} \leq \|e(0,\cdot)\|_{L^2} + \varepsilon^{-1} \int_0^t \|(i\varepsilon \partial_t - D) u_{F,1}(s,\cdot)\|_{L^2} ds \leq \varepsilon C_T.
\]

(4.29)

\[\square\]

Noticing that $\|u - u_{F,0}\|_{L^2} \leq \|u - u_{F,1}\|_{L^2} + \|u_{F,1} - u_{F,0}\|_{L^2}$, from Propositions 4.1 and 4.2, we can then state our main theorem on the accuracy of the FGA for the Dirac system.

**Theorem 4.1.** Let $u$ be the solution of the DS (1.7) and $u_{F,0}$ be the corresponding FGA solution, then for any $T > 0$, there exists a constant $C_T$, so that for any $\varepsilon > 0$
\[
\sup_{0 \leq t \leq T} \|u - u_{F,0}\|_{L^2} \leq \varepsilon C_T.
\]

(4.30)

**Proof.** The result follows from Propositions 4.1 and 4.2, and
\[
\|u - u_{F,0}\|_{L^2} \leq \|u - u_{F,1}\|_{L^2} + \|u_{F,1} - u_{F,0}\|_{L^2}.
\]

(4.31)

\[\square\]
5. Generalization to linear non-strictly hyperbolic systems In the previous sections, we have precisely analyzed the order of convergence of the FGA for two fundamental examples of linear non-strictly hyperbolic systems. The analysis for strictly hyperbolic systems was proposed in [13]; rather than a complete analysis of convergence of the FGA for linear non-strictly hyperbolic systems which would require reiteration of results from [13], we discuss hereafter the extension to non-strictly hyperbolic systems using the same arguments as the ones used in Sections 3 and 4. Consider a linear hyperbolic system

\[ \partial_t u + \sum_{i=1}^d A_i(x) \partial_{x_i} u = 0, \]

for \( u : \mathbb{R}^d \to \mathbb{R}^N \), with \( A_i \) smooth, and \( \sum_{i=1}^d p_i A_i(q) \) having eigenvalues not all distinct \( \{H_n\}_{n=1}^N \). The general strategy for proving the convergence consists in estimating

\[ \sup_{t \in [0,T]} \| u(t, \cdot) - u_{F,0}(t, \cdot) \|_E \leq \varepsilon C_{T,\delta}, \tag{5.1} \]

where \( u \) is the exact solution and \( u_{F,0} \) the FGA at order 0. In order to estimate (5.1), one must estimate the following 3 terms:

\[ \| u - \tilde{u}_{F,1} \|_{L^2}, \| \tilde{u}_{F,1} - u_{F,1} \|_{L^2}, \| u_{F,1} - u_{F,0} \|_{L^2}, \tag{5.2} \]

where \( \tilde{u}_{F,1} \) is a filtered version of FGA, and \( u_{F,1} \) is the first order FGA. This is in particular the strategy which is used in [13, Theorem 4.1] (the corresponding notation in the latter reference, are \( \mathcal{P}_t u_0^\varepsilon \), \( \mathcal{P}_{t,K,\delta}^\varepsilon u \), \( \mathcal{P}_{t,K,\delta}^\varepsilon u_0^\varepsilon \)). However, in the non-strictly hyperbolic case, the FGA and filtered FGA possess asymptotic correction terms \( a_n^\perp \) which allow for dealing with the multiplicity of eigenvalues. This correction term, \( a_n^\perp \), can be explicitly evaluated thanks to bounded pseudo-inverse operators \( L_n^{-1} \), using similar compactness arguments on the eigenvalues and eigenvectors as in Proposition 3.6 for elastic wave system and Lemma 4.1 for the Dirac system. Then proving \( \| u - u_{F,0} \|_{L^2} \leq \varepsilon C_{T,\delta} \) is identical in the strict and non-strict hyperbolic cases as in [13, Theorem 4.1]. The estimates on the other terms in (5.2), is a consequence on the boundness of the correction terms and arguments from [13], which were reiterated in Sections 3 and 4 for the elastic wave equations and the Dirac system respectively.

6. Conclusion In this paper, we established the convergence theory of FGA for elastic wave system (EWS) and Dirac system (DS), which has been numerically verified as an efficient tool to compute high-frequency wave propagation. Unlike the convergence theory of FGA for strictly linear hyperbolic systems [13], we needed to analyze the boundness of intraband transitions in diabatic coupling, which only appears when the system is non-strictly hyperbolic. The techniques we have developed for proving the convergence of FGA for both EWS and DS can be straightforwardly used to prove the convergence of FGA for any non-strictly hyperbolic systems.

Acknowledgments. L.C. was partially supported by the NSFC grant 11901601. J.C.H. and X.Y. were partially supported by the NSF grant DMS-1818592. E.L. was partially supported through the NSERC Discovery Grant program.


