

SMOOTHING PROPERTIES OF EVOLUTION EQUATIONS VIA CANONICAL TRANSFORMS AND COMPARISON

MICHAEL RUZHANSKY AND MITSURU SUGIMOTO

ABSTRACT. The paper describes a new approach to global smoothing problems for dispersive and non-dispersive evolution equations based on the global canonical transforms and the underlying global microlocal analysis. For this purpose, the Egorov-type theorem is established with canonical transformations in the form of a class of Fourier integral operators, and their weighted L^2 -boundedness properties are derived. This allows us to globally reduce general dispersive equations to normal forms in one or two dimensions. Then, several comparison principles are introduced to relate different smoothing estimates by comparing certain expressions involving their symbols. As a result, it is shown that the majority of smoothing estimates for different equations are equivalent to each other. Moreover, new estimates as well as several refinements of known results are obtained. The proofs are considerably simplified. A comprehensive analysis is presented of smoothing estimates for homogeneous and inhomogeneous, dispersive and also non-dispersive equations with constant coefficients. Applications are given to the detailed description of smoothing properties of the Schrödinger, relativistic Schrödinger, wave, Klein-Gordon, and other equations. Critical cases of some estimates and their relation to the trace estimates are discussed.

CONTENTS

1. Introduction	2
2. Comparison principle	12
3. Equivalent model estimates	16
4. Canonical transforms	21
5. Smoothing estimates for dispersive equations	25
6. Invariant estimates and sharpness	31
7. Smoothing estimates for non-dispersive equations	35
8. Relativistic Schrödinger, wave, and Klein-Gordon equations	41
9. Model estimates for inhomogeneous equations	45
10. Smoothing estimates for dispersive inhomogeneous equations	48
11. Trace theorems	52
References	56

Date: January 14, 2007.

This work was completed with the aid of “UK-Japan Joint Project Grant” by “The Royal Society (UK)” and “Japan Society for the Promotion of Science”. The first author was also supported by the Leverhulme Research Fellowship.

1. INTRODUCTION

In the last two decades, since the independent pioneering works by Ben-Artzi and Devinatz [BD2], Constantin and Saut [CS], Sjölin [Sj] and Vega [V], the local, and then global smoothing effects of Schorödinger equations, or more generally, those of dispersive equations has been intensively investigated. Similar smoothing effects have been observed for different equations of great importance in mathematical physics (for example, smoothing for generalized Korteweg-de Vries equations was already studied by Kato [Ka2], several other equations were studied in a series of papers by Kenig, Ponce and Vega [KPV1]–[KPV5]), etc). Over the years, several techniques to understand these smoothing properties through the Fourier analysis, functional analysis, spectral theory and harmonic analysis have been developed. The analysis of such smoothing estimates is particularly important in applications to nonlinear evolution equations, especially to those with derivatives in the potential or in the nonlinearity. Over the last three decades two major approaches, Strichartz and smoothing estimates, proved to be two extremely efficient tools for dealing with nonlinear equations. The smoothing effect is crucial in allowing to recover the loss of derivatives in the equation making these estimates a very good substitute for the Strichartz estimates that are normally used for semilinear equations.

The objective of this paper is to provide a new approach leading to a comprehensive understanding of the effect of global smoothing, together with new results, through two novel ideas. It will allow us not only to recover existing and to prove new estimates, but to effectively show that the smoothing phenomenon for equations describing often completely different physical processes (like wave, Klein-Gordon, Schrödinger, relativistic Schrödinger, KdV, Benjamin-Ono, Davey-Stewartson, Shrira, Zakharov-Schulman, and many other equations) is of essentially the same nature. For this, we will provide a way to show the equivalence of smoothing properties for very different equations by introducing two new ideas for the subject.

The first is the idea of *canonical transformations*. Let us mention immediately that will not concentrate specifically on the local smoothing since it is contained in its global version. Although canonical transforms are well known in microlocal analysis of local problems, we will apply them in a global setting here to globally reduce problems to normal forms in lower dimensions. As it turns out, it is then possible to carry out the pointwise analysis in these model problems. However, there are some essential differences with the microlocal case. On one hand, we will still be able to reduce elliptic operators to one dimensional models. On the other hand, in the case of dispersive operators (or operators of real principal type) the global reduction will be made to models in two dimensions, in difference with the well-known microlocal constructions of Duistermaat and Hörmander [DH].

Another important idea presented and explored in this paper is an introduction of a certain *comparison principle* for evolution equations which will allow us to derive new estimates for dispersive (and non-dispersive) equations from known ones, as well as compare estimates for different equations. The idea here is that we can compare certain expressions involving symbols and weights for different estimates and conclude that one estimate implies the other if an inequality between these expressions holds. In particular, it will imply that smoothing estimates are equivalent if certain

expression involving symbols are equivalent. This will, for example, show that a variety of global smoothing estimates for Schrödinger equations are simply equivalent to the corresponding estimates for the relativistic Schrödinger, wave, Klein–Gordon, linearised KdV, Benjamin–Ono, and other equations. In addition, it will show that the local smoothing effect for Schrödinger equations that was established by Sjölin [Sj] and Vega [V] is equivalent to the energy conservation of a travelling wave in one dimension. The gain of $1/2$ -derivative corresponds to the Jacobian of the frequency transformation between Schrödinger and a one-dimensional wave in the radial direction, and the $1/2$ -derivative smoothing for Schrödinger is the energy estimate for this wave, which in turn is just the translation invariance property of the Lebesgue measure on the real line (see (1.17) and the discussion around it). The local gain of one derivative for Korteweg-de Vries equation was also observed by Kato [Ka2], whose proof used the algebraic properties of the symbol and the fact that the situation is one-dimensional. Again, by the comparison principle we will immediately recover this result (and its global version) from the $1/2$ -smoothing for Schrödinger, or from the energy conservation for the wave equation.

The methods and ideas described in the paper seem to apply very efficiently to the area of smoothing estimates. These estimates are usually immediately applied to the well-posedness of nonlinear equations and many methods are developed for this purpose. From this point of view, results presented here have immediate consequences for the well-posedness problems for rather general classes of nonlinear evolution equations. Moreover, it can be expected that they can be also applied to a variety of other problems where weighted estimates and phase space analysis are of importance.

Let us mention that there has already been a lot of literature on the subject of global smoothing estimates from different points of view. See, Ben-Artzi and Devinatz [BD1, BD2], Ben-Artzi and Klainerman [BK], Chihara [Ch], Hoshiro [Ho2], Kato and Yajima [KY], Kenig, Ponce and Vega [KPV1, KPV2, KPV3, KPV4], Linares and Ponce [LP], Simon [Si], Sugimoto [Su1, Su2], Walther [Wa1, Wa2], and many others. The two most commonly used techniques are based on the limiting absorption principle from the spectral theory or on the restriction theorems in harmonic analysis. They are both better adapted to study elliptic equations and often do not work in the dispersive non-elliptic cases (let alone non-dispersive equations). Indeed, there may be lack of information about the spectrum, or level surfaces of a non-elliptic symbol may be non-compact, may have singularities or vanishing Gaussian curvature. The described method of canonical transforms works equally well in all dispersive situations (elliptic and non-elliptic). As such, methods developed here may be used to producing their counterpart restriction theorems for corresponding level surfaces. Moreover, since one can make full use of the microlocal analysis, most results can be extended to different types of symbols, e.g. to quasi-homogeneous symbols, just by microlocalising in appropriate directions and applying results of this paper. Because of this we will not give a full treatment of quasi-homogeneous symbols here, mostly restricting ourselves to homogeneous symbols with lower order terms for the clarity of the exposition.

Moreover, we will obtain the corresponding results on the global smoothing for solutions to inhomogeneous problems. There are considerably less results on this topic available in the literature. Mostly the Schrödinger equation was treated (e.g.

Linares and Ponce [LP], Kenig, Ponce and Vega [KPV5]), or the one dimensional case (Kenig, Ponce and Vega [KPV3, KPV4] or Laurey [La]). Some more general results on the local smoothing for dispersive operators were obtained by Chichara [Ch] and Hoshiro [Ho2], and for dispersive differential operators by Koch and Saut [KoSa]. In this paper we will extend these results in two directions: we will establish the global smoothing for rather general dispersive equations of different orders in all dimensions. Moreover, using the presented methods further results can be obtained for some classes of non-dispersive equations as well. In turn, many of the established results can be also further extended to include small perturbations by terms of the order up to the established smoothing (e.g. for small magnetic potentials in the case of the Schrödinger equation as in Georgiev and Tarulli [GT], or for the wave or Dirac equations as in e.g. D’Ancona and Fanelli [DF], etc.) In non-dispersive cases they may involve certain structural conditions on perturbations corresponding to the invariant form of smoothing estimates established in this paper. Moreover, global versions of the analytic microlocal smoothing (as considered locally by using FBI transforms by e.g. Robbiano and Zuily [RZ] for Schrödinger, or Takuwa [Ta] for dispersive equations, with a preceding work in the smooth setting by Craig, Kappeler and Strauss [CKS]) can be also expected to be obtained by the presented methods. Smoothing of solutions to Schrödinger equations has been also analysed by Kapitanski and Safarov [KaSa], where a relation between propagators of Schrödinger and wave equations was established. In fact, in our setting that relation can be also viewed as a composition of a canonical transform and a comparison argument of this paper (in particular using relation (1.16) below). In this paper we will concentrate on the case of the case of equations with “constant coefficients”, for the clarity and comprehensiveness of the exposition. Further applications to evolution equations for time and space dependent operators will appear elsewhere.

Now we will explain the essence of the approach and relate it to the known techniques. The main idea is that instead of establishing smoothing estimates for different classes of equations, we will instead relate such estimates to each other. In particular, we will be able to relate the majority of such estimates to a simple pointwise estimate for a travelling wave in lower dimensions, which will follow by a simple Fourier analysis argument. This will shed some light on the nature of such estimates in a wider context, exhibiting a quantitative smoothing phenomenon of the same type for equations describing very different physical processes. A comprehensive understanding of smoothing for equations with constant coefficients should also make an impact on problems with potentials and problems of perturbations of such equations, an area which saw an amazingly rapid development over the recent years. In particular, having an explicit relation between smoothing estimates for constant coefficient equations can suggest the corresponding admissible classes of potentials, etc. Moreover, it should also influence the understanding of smoothing–Strichartz estimates, a combined technique which also proved to be very effective in nonlinear problems.

We will also suggest some invariant forms of the smoothing estimates which we expect to continue to hold in non-dispersive cases as well, and we will give several results to justify this expectation (these are estimates (1.19)–(1.21)). We note that despite their natural appearance in many problems, quite limited results are available for non-dispersive equations while it is known that some usual estimates fail in those

case. In Section 6 we will discuss what smoothing estimates are natural for such equations and Section 7 will be devoted to the analysis of the non-dispersive cases.

As one of the simplest cases, let us first consider the following Schrödinger equation:

$$(1.1) \quad \begin{cases} (i\partial_t + \Delta_x) u(t, x) = 0 & \text{in } \mathbb{R}_t \times \mathbb{R}_x^n, \\ u(0, x) = \varphi(x) & \text{in } \mathbb{R}_x^n. \end{cases}$$

We know that the solution operator $e^{it\Delta_x}$ preserves the L^2 -norm for each fixed $t \in \mathbb{R}$. On the other hand, the extra gain of regularity of order $1/2$ in x can be observed if we integrate the solution in t . For example, in the case $n = 1$, we have

$$(1.2) \quad \left\| |D_x|^{1/2} u(\cdot, x) \right\|_{L^2(\mathbb{R}_t)} \leq C \|\varphi\|_{L^2(\mathbb{R})},$$

for all $x \in \mathbb{R}$. This result was given by e.g. Kenig, Ponce and Vega [KPV1]. Again, an application of the comparison principle in Section 2 will allow us to compare this estimate to the one-dimensional wave equation and will show that (1.2) is nothing else but the energy conservation for the wave. In the one dimensional case this is again just the translation invariance of the Lebesgue measure.

In the higher dimensional case $n \geq 2$, similar global smoothing properties are of importance:

$$(1.3) \quad \|Au\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)},$$

where A is one of the following:

- (1) $A = \langle x \rangle^{-s} |D_x|^{1/2}$; $s > 1/2$,
- (2) $A = |x|^{\alpha-1} |D_x|^\alpha$; $1 - n/2 < \alpha < 1/2$,
- (3) $A = \langle x \rangle^{-s} \langle D_x \rangle^{1/2}$; $s \geq 1$ ($s > 1$ if $n = 2$).

We use the standard notation

$$\langle x \rangle = (1 + |x|^2)^{1/2} \quad \text{and} \quad \langle D_x \rangle = (1 - \Delta_x)^{1/2}.$$

The type (1) was given by Ben-Artzi and Klainerman [BK] ($n \geq 3$), and Chihara [Ch] ($n \geq 2$). The type (2) was given by Kato and Yajima [KY] ($n \geq 3$, $0 \leq \alpha < 1/2$ or $n = 2$, $0 < \alpha < 1/2$), and Sugimoto [Su1] ($n \geq 2$). Watanabe [W] showed that it is not true for $\alpha = 1/2$. The type (3) was given by Kato and Yajima [KY] ($n \geq 3$), and Walther [Wa1] ($n \geq 2$) who also showed that it is not true for $s < 1$ ($s \leq 1$ if $n = 2$). The type (4) estimate with homogeneous weight and $\langle D_x \rangle^{1/2}$ was not considered much but it will be discussed in Sections 5 and 6.

Each proof was carried out by proving one of the following estimates (or their variants):

$$(1.4) \quad \left\| \widehat{A^* f} \Big|_{\rho\mathbb{S}^{n-1}} \right\|_{L^2(\rho\mathbb{S}^{n-1})} \leq C \sqrt{\rho} \|f\|_{L^2(\mathbb{R}^n)} \quad (\text{Restriction theorem}),$$

where, $\rho\mathbb{S}^{n-1} = \{\xi \in \mathbb{R}^n : |\xi| = \rho\}$, ($\rho > 0$), or

$$(1.5) \quad \sup_{\text{Im } \zeta > 0} |(R(\zeta)A^* f, A^* f)| \leq C \|f\|_{L^2(\mathbb{R}^n)}^2 \quad (\text{Resolvent estimate}),$$

where $R(\zeta) = (-\Delta_x - \zeta)^{-1}$. Estimate (1.4) implies the dual one of estimate (1.3). Estimate (1.5) implies (1.3) since the resolvent $R(\zeta)$ is the Laplace transform of the

solution operator $e^{it\Delta_x}$ of equation (1.1):

$$R(\zeta) = \frac{1}{i} \int_0^\infty e^{it\Delta_x} e^{i\zeta t} dt \quad (\text{Im } \zeta > 0).$$

The fact that (1.5) implies (1.4) is due to the formula

$$\text{Im} (R(\rho^2 + i0)f, f) = \frac{1}{4(2\pi)^{n-1}\rho} \left\| \widehat{f}|_{\rho\mathbb{S}^{n-1}} \right\|_{L^2(\rho\mathbb{S}^{n-1})}^2,$$

see e.g. Hörmander [H, Corollary 14.3.10].

In this paper we introduce several new ideas to prove estimate (1.3). The main two proposed methods (canonical transforms and comparison principles) are centred at comparing different estimates rather than looking at them individually. This approach will allow us to actually relate most of estimates to each other as well as to their normal forms. For example, we will show that estimates (1.3) with A as in (1), (2), or (3), are equivalent to some simple one dimensional estimates. To explain this idea, let us first recall that operators other than the Schrödinger operator have also attracted much attention for their smoothing properties. For example, relativistic Schrödinger equations have been investigated in [BN] and [Wa2], wave and Klein–Gordon equations in [Be], Kortevég–de Vries equations in [KPV2], Benjamin–Ono equations in [KPV4], Davey–Stewartson systems in [LP], certain dispersive polynomial equations in [BD2], third order differential equations in [KoSa], to mention a few, and they can be expressed in the general form

$$(1.6) \quad \begin{cases} (i\partial_t + a(D_x)) u(t, x) = 0 & \text{in } \mathbb{R}_t \times \mathbb{R}_x^n, \\ u(0, x) = \varphi(x) & \text{in } \mathbb{R}_x^n, \end{cases}$$

where $a(\xi)$ is a real-valued function of $\xi = (\xi_1, \dots, \xi_n)$ with the growth of order m , and $a(D_x)$ is the corresponding operator. Equations of this type have been extensively studied under the ellipticity ($a(\xi) \neq 0$ for $\xi \neq 0$) or the dispersiveness ($\nabla a(\xi) \neq 0$ for $\xi \neq 0$) conditions. Under such conditions, various global smoothing estimates have been established for solutions $u(t, x) = e^{ita(D_x)}\varphi(x)$ in many papers, in both differential and pseudo-differential cases ([BN], [BD2], [Ch], [CS], [Ho1], [Ho2], [KY], [KPV1], [RS1], [Wa2], etc.). The dispersiveness condition was shown to be necessary for certain types of estimates (see Hoshiro [Ho2]), but we will show how to get around that. Now, suppose that we want to establish a weighted smoothing estimate of the form

$$(1.7) \quad \left\| w(x)\rho(D_x)e^{ita(D_x)}\varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)},$$

giving a smoothing of type $\rho(D_x)$ with some weight $w(x)$. The rough idea of the canonical transform method is to use certain operators T for which we have the relations

$$a(D_x) \circ T = T \circ \tilde{a}(D_x) \quad \text{and} \quad \rho(D_x) \circ T = T \circ \tilde{\rho}(D_x),$$

for some other operators $\tilde{a}(D_x)$ and $\tilde{\rho}(D_x)$. Then we also have $e^{ita(D_x)} \circ T = T \circ e^{it\tilde{a}(D_x)}$. We now substitute $T\varphi$ for φ in estimate (1.7), and have

$$\left\| w(x)\rho(D_x)e^{ita(D_x)}T\varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|T\varphi\|_{L^2(\mathbb{R}_x^n)}.$$

Using the above identities we can conclude that estimate (1.7) is equivalent to the estimate

$$(1.8) \quad \left\| w(x) T \tilde{\rho}(D_x) e^{it\tilde{a}(D_x)} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|T\varphi\|_{L^2(\mathbb{R}_x^n)}.$$

If now operators T and T^{-1} are bounded in $L^2(\mathbb{R}_x^n)$ and in weighted $L^2(\mathbb{R}_x^n)$ with weight $w(x)$ respectively, we can remove them from (1.8) to finally conclude that weighted smoothing estimate (1.7) is equivalent to

$$(1.9) \quad \left\| w(x) \tilde{\rho}(D_x) e^{it\tilde{a}(D_x)} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)}.$$

This idea can be used in a variety of ways. Not only can we derive one smoothing estimate from another, but we can also consider equivalent classes of smoothing estimates and find their normal forms where the analysis would follow from some straightforward argument. Thus, Section 3 will be devoted to estimates in such model cases, while Section 4 will be devoted to weighted estimates for necessary operators T . Moreover, in Sections 2 we will develop comparison principles which will also allow us to relate the model estimates from different classes among each other. All of the arguments will be invertible, thus establishing a more or less complete set of relations among different types of smoothing estimates (dispersive in Section 5 and non-dispersive in Section 7). In addition, in Section 3 we will also relate estimates with different weights.

As for transformation operators T and T^{-1} , we will consider Fourier integral operators, or rather operators which can be globally written in the form

$$(1.10) \quad Tu(x) = (2\pi)^{-n} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} e^{i\Phi(x,y,\xi)} p(x,y,\xi) u(y) dy d\xi \quad (x \in \mathbb{R}^n),$$

where $p(x,y,\xi)$ is an amplitude function and $\Phi(x,y,\xi)$ is a real phase function (not always positively homogeneous in ξ in our applications). Especially, if $p(x,y,\xi) = 1$ and $\Phi(x,y,\xi)$ satisfies the graph condition

$$\begin{aligned} \Lambda &= \{(x, \Phi_x, y, -\Phi_y); \Phi_\xi = 0\} \\ &= \{(x, \xi), \chi(x, \xi)\} \subset T^*\mathbb{R}^n \times T^*\mathbb{R}^n, \end{aligned}$$

we have the relation

$$\begin{aligned} T \circ A(X, D_x) \circ T^* &= B(X, D_x) + (\text{lower order terms}), \\ B(x, \xi) &= (A \circ \chi)(x, \xi), \end{aligned}$$

for pseudo-differentiable operators $A(X, D_x)$ and $B(X, D_x)$. In this way, Fourier integral operators are recognised as a tool of the realisation of the canonical transformation. This fact is well-known microlocally as Egorov's theorem, and by taking phase function appropriately, properties of the operator $B(X, D_x)$ can be extracted from those of the operator $A(X, D_x)$. By using Egorov's theorem, many *qualitative* properties of solutions of partial differential equations (propagation of singularities, construction of parametrises, etc.) have been investigated. Our main interest is to establish *quantitative* properties as well (global L^2 -property for example) by the same idea. In this paper, we take

$$(1.11) \quad \Phi(x, y, \xi) = x \cdot \xi - y \cdot \psi(\xi)$$

and use the exact relation

$$(1.12) \quad T \circ \sigma(D_x) = a(D_x) \circ T, \quad a(D_x) = (\sigma \circ \psi)(D_x),$$

for translation invariant pseudo-differential operators $\sigma(D_x)$ and $a(D_x)$. For example, the Laplacian $\Delta_x = \partial_{x_1}^2 + \cdots + \partial_{x_n}^2$ can be transformed to $\partial_{x_n}^2$ by choosing an appropriate $\psi(\xi)$, and hence we will be able to reduce the smoothing estimate for Schrödinger equation (1.1) to the one dimensional estimate (1.2). We note that since we will be working with operators with constant coefficients we are able to perform the exact global calculus, in comparison to the calculus modulo lower order or smoothing terms provided by the Egorov's theorem. Moreover, we will be using the exact inverse T^{-1} rather than the adjoint T^* . The global L^2 -boundedness of operators (1.10) has been investigated before, for example by Asada and Fujiwara [AF], Kumano-go [Ku] and Boukhemair [Bo1, Bo2]. Unfortunately, in all these papers an assumption was made for the second order derivatives matrix $\nabla_\xi^2 \Phi(x, y, \xi)$ to be globally bounded in all variables, which clearly fails for the phase (1.11). However, the global L^2 and also weighted L^2 boundedness theorems for Fourier integral operators without such assumption are required for our analysis. Some of these results have been established by the authors in [RS2] and some will be proved in Section 4.

It is remarkable that the method of canonical transformations described above allows us to carry out a global microlocal reduction of equation (1.6) to the model cases $|\xi_n|^m$ (elliptic case) or $\xi_1 |\xi_n|^{m-1}$ (non-elliptic case) under the dispersiveness condition. For example, for equation (1.6) Chihara [Ch] used involved spectral and harmonic analysis and established the estimate

$$(1.13) \quad \left\| \langle x \rangle^{-s} |D_x|^{(m-1)/2} e^{ita(D_x)} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)} \quad (s > 1/2)$$

in the case when $a(\xi)$ is positively homogeneous of order $m > 1$. With canonical transforms, this estimate is easily reduced to low dimensional pointwise estimates

$$(1.14) \quad \left\| |D_x|^{(m-1)/2} e^{it|D_x|^m} \varphi(x) \right\|_{L^2(\mathbb{R}_t)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x)},$$

$$(1.15) \quad \left\| |D_y|^{(m-1)/2} e^{itD_x |D_y|^{m-1}} \varphi(x, y) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_y)} \leq C \|\varphi\|_{L^2(\mathbb{R}_{x,y}^2)},$$

for all $x \in \mathbb{R}$, respectively. Note that estimate (1.14) with $m = 2$ is estimate (1.2) for the Schrödinger equation in one dimension. By establishing (1.14) and (1.15) directly, we will be able to immediately obtain (1.13) for $m > 0$, thus also including the hyperbolic case $m = 1$, which will be important for further analysis, in particular for the understanding of the meaning of various estimates in terms of the finite speed of propagation of singularities, etc. The results which will be thus obtained on this path generalise and extend many known results in the literature mentioned above. Moreover, this new idea gives us a clear comprehensive understanding of the smoothing effects of dispersive equations.

In addition, we will introduce another technique with which we can show that the comparison of the symbols implies the same comparison of corresponding operators. For example, in the one dimensional case, if we have

$$\frac{|\sigma(\xi)|}{|f'(\xi)|^{1/2}} \leq A \frac{|\tau(\xi)|}{|g'(\xi)|^{1/2}}$$

then we have automatically estimate

$$\|\sigma(D_x)e^{itf(D_x)}\varphi(x)\|_{L^2(\mathbb{R}_t)} \leq A\|\tau(D_x)e^{itg(D_x)}\varphi(x)\|_{L^2(\mathbb{R}_t)},$$

for all $x \in \mathbb{R}$. This will, in turn, imply a variety of weighted estimates. It will also allow us to relate normal forms of estimates for operators of different orders. As an example, let us mention the following consequence for $n = 1$ and $l, m > 0$:

$$(1.16) \quad \left\| |D_x|^{(m-1)/2} e^{it|D_x|^m} \varphi(x) \right\|_{L^2(\mathbb{R}_t)} = \sqrt{\frac{l}{m}} \left\| |D_x|^{(l-1)/2} e^{it|D_x|^l} \varphi(x) \right\|_{L^2(\mathbb{R}_t)}$$

for every $x \in \mathbb{R}$, assuming that $\text{supp } \widehat{\varphi} \subset [0, +\infty)$ or $(-\infty, 0]$. We will introduce this kind of comparison principles in more general settings, which will prove to be another strong tool to induce general estimates from simple ones. Particularly, if we use the comparison principle in both directions, we can show the equivalence of many different smoothing estimates. For example, using (1.16) with $l = 1$, we can show that estimate (1.14) or (1.15) is equivalent to the same estimate but just in the special case $m = 1$. This fact means that these two standard estimates can in turn be derived from the equality

$$(1.17) \quad \left\| e^{itD_x} \varphi(x) \right\|_{L^2(\mathbb{R}_t)} = \|\varphi\|_{L^2(\mathbb{R}_x)}$$

in the case $n = 1$, which is just the conservation of energy for the travelling wave in one dimension. In this way, smoothing estimates for dispersive equation (1.6) can be surprisingly reduced to just a simple equality (1.17), which is a straightforward consequence of the trivial fact $e^{itD_x}\varphi(x) = \varphi(x + t)$. Thus, we can immediately recover the gain of 1/2-derivatives for the Schrödinger and of one derivative for the Korteweg- de Vries equations (as in e.g. Kato [Ka2]). In this way we can actually reduce all dispersive smoothing estimates to those for model hyperbolic, Schrödinger, relativistic, KdV, or other equations (whichever we prefer), or we can show that they are all equivalent to each other. In addition, we will find some explicit best constants based on a constant found by Simon [Si] using Kato's theory [Ka1].

On the other hand, coupled dispersive equations are of immense importance in applications while with only limited analysis available. To give an example, let $v(t, x)$ and $w(t, x)$ solve the following coupled system of Schrödinger equations:

$$(1.18) \quad \begin{cases} i\partial_t v = \Delta_x v + b(D_x)w, \\ i\partial_t w = \Delta_x w + c(D_x)v, \\ v(0, x) = v_0(x), w(0, x) = w_0(x). \end{cases}$$

This is the simplest example of Schrödinger equations coupled through linearised operators $b(D_x), c(D_x)$. Such equations appear in many areas in physics. For example, this is a model of wave packets with two modes (in the presence of resonances), see Tan and Boyd [TB]. In fiber optics they appear to describe certain types of a pair of coupled modulated wave-trains (see e.g. Manganaro and Parker [MP]). They also describe the field of optical solitons in fibres (see Zen and Elim [ZE]) as well as Kerr dispersion and stimulated Raman scattering for ultrashort pulses transmitted through fibres. In these cases the linearised operators b and c would be of zero order. In models of optical pulse propagation of birefringent fibres and in wavelength-division-multiplexed systems they are of the first order (see Pelinovsky

and Yang [PY]). They may be of higher orders as well, for example in models of optical solitons with higher order effects (see Nakkeeran [Na]). A by now standard way to tackle nonlinear versions of (1.18) are Strichartz and smoothing estimates. We will give some examples of such approach based on the critical case of one of the smoothing estimates established in this paper. For example, we will apply it to the global in time well-posedness of derivative nonlinear Schrödinger equations with some structural conditions. The details fall outside the scope of this paper and will appear elsewhere.

Suppose now that we are in the simplest situation when system (1.18) can be diagonalised. Its eigenvalues are $a_{\pm}(\xi) = -|\xi|^2 \pm \sqrt{b(\xi)c(\xi)}$ and the system uncouples into scalar equations of type (1.6) with operators $a(D_x) = a_{\pm}(D_x)$. Since the structure of operators $b(D_x), c(D_x)$ may be quite involved, this motivates the study of scalar equations (1.6) with operators $a(D_x)$ of rather general form. Not only the presence of lower order terms is important in time global problems, the principal part may be rather general since we may have $\nabla a_{\pm} = 0$ at some points. In such situation we microlocalise around such points and lose the structure (but not the properties) of the symbol completely.

The combination of the proposed two new methods (canonical transformations and the comparison principles) however has a good power on the occasion of this analysis. Besides the simplification of the proofs of smoothing estimates for standard dispersive equations, we have an advantage in treating rather general dispersive equations where $a(\xi)$ admits lower order terms, and also non-dispersive equations where the dispersiveness condition $\nabla a(\xi) \neq 0$ breaks (Section 7). The inclusion of lower order terms in the analysis is essential here since time global properties are dominated by the low frequency part $|\xi| \leq R$ ($R > 0$) of $a(\xi)$. This fact is true even for the ordinary Schrödinger equation, with the homogeneous Laplacian. In such low frequency case we can not talk about the principal part of an operator, so operators with lower order terms appear naturally, and will be considered as condition (L) in Section 5.

We also suggest an invariant form of smoothing estimates which remain valid also in some areas without dispersion, where standard smoothing estimates are known to fail. Let us observe the following form of estimates of types (1)–(3) for (1.3). The first estimate may be rewritten in the form

$$(1.19) \quad \left\| \langle x \rangle^{-s} |\nabla a(D_x)|^{1/2} e^{ita(D_x)} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)} \quad (s > 1/2).$$

An analogous invariant forms for the other smoothing estimates are estimate

$$(1.20) \quad \left\| |x|^{\alpha-m/2} |\nabla a(D_x)|^{\alpha/(m-1)} e^{ita(D_x)} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)} \quad (m \neq 1),$$

which we expect to hold for $(m-n)/2 < \alpha < (m-1)/2$, with $m > 0$ being the order of $a(D_x)$, and estimate

$$(1.21) \quad \left\| \langle x \rangle^{-m/2} \langle \nabla a(D_x) \rangle^{1/2} e^{ita(D_x)} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)},$$

for $n > m > 1$ because of the low frequency contribution. Such estimates have a number of advantages which we would like to point out taking (1.19) as an example:

- in the dispersive case it is equivalent to the usual estimate (1.22) below;

- it does continue to hold for a variety of non-dispersive equations, where $\nabla a(\xi)$ may become zero on some set and when (1.22) fails;
- it does take into account zeros of the gradient $\nabla a(\xi)$, which is also responsible for the interface between dispersive and non-dispersive zone (e.g. how quickly the gradient vanishes);
- it is invariant under canonical transforms of the equation;
- the proposed estimates are scaling invariant or “almost” invariant;
- the estimates are also “sharp” (see Section 6).

We observe the estimate

$$(1.22) \quad \|\langle x \rangle^{-s} |D_x|^{(m-1)/2} e^{ita(D_x)} \varphi(x)\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)}$$

which is known from e.g. Theorems 5.1 and Corollary 5.5. An additional advantage of using estimate (1.19) rather than (1.22) is that (1.19) takes into account possible zeros of the gradient $\nabla a(\xi)$ in the non-dispersive case. Thus, in the one dimensional or in the radially symmetric cases of Section 2 we will see that estimate (1.19) is still valid in non-dispersive cases. In Section 7 we will show examples of this estimate in other non-dispersive cases. In particular, we will justify the invariant estimates above in several situations using the following ideas:

- we can microlocalise around non-dispersive points and apply the canonical transforms there (Theorem 7.3);
- in radially symmetric cases we can use the comparison principle (Theorems 2.5, 7.1, and Corollary 5.7) ;
- in some quasi-homogeneous cases or when the symbol can be represented as a sum of one dimensional monomials, we can use the comparison principle as well (Examples 7.2, 7.3 and 7.4);
- in the homogeneous case with some information on the Hessian we can use canonical transforms to reduce the general case to the previous situation (Theorem 7.2).

In several situations estimate (1.21) will be weaker than (1.19) and (1.20), so we may mostly concentrate on these two. Moreover, in Sections 4, 5, and 6, we will argue that in various complicated situations (like in general non-dispersive cases) estimates (1.19)–(1.21) are the (sharp) smoothing estimates that one can hope to obtain.

We will explain the organisation of this paper. In Section 2, we give the precise statements of the comparison principle. In Section 3, we prove important model estimates and also the equivalence of them by using the comparison principle. We will also apply the comparison principle again to compare many estimates with the estimates given here, and get secondary comparison results. In Section 4, we introduce and show the fundamental properties of our main tools which originate in the idea of canonical transformation. In Section 5, we list results which extend and explain estimate (1.3) with types (1)–(3), which were partially announced by the authors in [RS1] and [RS4]. Especially, these kinds of time-global estimate for the operator $a(D_x)$ with lower order terms are new results provided by the new method. We also explain how general cases can be reduced to the model estimates given in Section 3. Additional arguments with the idea of canonical transformation are also presented there. A second comparison result for radially symmetric case is also given there. In Section 6 we will propose and discuss an invariant form of smoothing estimates which remains to

hold in non-dispersive situations as well, and we also discuss the sharpness of all the estimates. In Section 7, we will establish invariant estimates for several case of non-dispersive equations by using the second comparison results. In Section 8 we apply the second comparison result further to the relativistic Schrödinger, Klein–Gordon, and wave equations. Sections 9 and 10 are devoted to non-homogeneous problems as a counterpart of Sections 3 and 5, respectively. Section 11 is devoted to related problems, including critical cases of some of the estimates, and the corresponding trace theorems.

Finally we comment on the notation used in this paper. As usual, we will denote $D_{x_j} = -i\partial_{x_j}$ and view operators $a(D_x)$ as Fourier multipliers. Constants denoted by letter C in estimates are always positive and may differ on different occasions, but will still be denoted by the same letter.

2. COMPARISON PRINCIPLE

In this section we will introduce a useful tool to derive new smoothing estimates from known ones and to relate different estimates for solutions to different equations with each other. That is a comparison principle for solutions $u(t, x) = e^{itf(D_x)}\varphi(x)$ and $v(t, x) = e^{itg(D_x)}\varphi(x)$ to evolution equations with operators $f(D_x)$ and $g(D_x)$, where $t \in \mathbb{R}$ and $x \in \mathbb{R}^n$:

$$\left\{ \begin{array}{l} (i\partial_t + f(D_x))u(t, x) = 0, \\ u(0, x) = \varphi(x), \end{array} \right. \quad \text{and} \quad \left\{ \begin{array}{l} (i\partial_t + g(D_x))v(t, x) = 0, \\ v(0, x) = \varphi(x). \end{array} \right.$$

In the sequel, we write $x = (x_1, \dots, x_n)$, $\xi = (\xi_1, \dots, \xi_n)$, and $D_x = (D_1, D_2, \dots, D_n)$ where D_j denotes $D_{x_j} = \frac{1}{i}\frac{\partial}{\partial x_j}$, ($j = 1, 2, \dots, n$).

First we note the following fundamental result:

Theorem 2.1. *Let $f \in C^1(\mathbb{R}^n)$ be a real-valued function such that, for almost all $\xi' = (\xi_2, \dots, \xi_n) \in \mathbb{R}^{n-1}$, $f(\xi_1, \xi')$ is strictly monotone in ξ_1 on the support of a measurable function σ on \mathbb{R}^n . Then we have*

$$(2.1) \quad \left\| \sigma(D_x) e^{itf(D_x)} \varphi(x_1, x') \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_{x'}^{n-1})}^2 = (2\pi)^{-n} \int_{\mathbb{R}^n} |\widehat{\varphi}(\xi)|^2 \frac{|\sigma(\xi)|^2}{|\partial f / \partial \xi_1(\xi)|} d\xi$$

for all $x_1 \in \mathbb{R}$, where $x' = (x_2, \dots, x_n) \in \mathbb{R}^{n-1}$.

Proof. Let $\eta = \Phi(\xi)$ and $\xi = \Phi^{-1}(\eta)$ be changes of variables defined by

$$\Phi(\xi) = (f(\xi), \xi'); \quad \Phi^{-1}(\eta) = (s(\eta), \eta'),$$

where we write $\eta = (\eta_1, \eta')$, $\eta' = (\eta_2, \dots, \eta_n)$. We assume that all the integrals below make sense which can be justified in an usual manner using the assumption and Sard's theorem. In view of this we perform calculations on the set $|\partial\Phi(\xi)| = |\partial f / \partial \xi_1(\xi)| \neq 0$.

We have

$$\begin{aligned}
& \sigma(D_x)e^{itf(D_x)}\varphi(x) \\
&= (2\pi)^{-n} \int_{\mathbb{R}^n} e^{itf(\xi)} e^{ix \cdot \xi} \sigma(\xi) \widehat{\varphi}(\xi) d\xi \\
&= (2\pi)^{-n} \int_{\Phi(\mathbb{R}^n)} e^{i(t\eta_1 + x' \cdot \eta')} e^{ix_1 s(\eta)} \sigma(\Phi^{-1}(\eta)) \widehat{\varphi}(\Phi^{-1}(\eta)) |\partial\Phi^{-1}(\eta)| d\eta,
\end{aligned}$$

where we used the substitution $\xi = \Phi^{-1}(\eta)$ on the support of χ . Using Plancherel's identity, we get

$$\begin{aligned}
& \|\sigma(D_x)e^{itf(D_x)}\varphi(x)\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^{n-1})}^2 \\
&= (2\pi)^{-n} \int_{\Phi(\mathbb{R}^n)} |\sigma(\Phi^{-1}(\eta)) \widehat{\varphi}(\Phi^{-1}(\eta))|^2 |\partial\Phi^{-1}(\eta)|^2 d\eta \\
&= (2\pi)^{-n} \int_{\mathbb{R}^n} |\sigma(\xi) \widehat{\varphi}(\xi)|^2 |\partial\Phi^{-1}(\Phi(\xi))|^2 |\partial\Phi(\xi)| d\xi \\
&= (2\pi)^{-n} \int_{\mathbb{R}^n} |\widehat{\varphi}(\xi)|^2 \frac{|\sigma(\xi)|^2}{|\partial f / \partial \xi_1(\xi)|} d\xi,
\end{aligned}$$

where we have used the substitution $\eta = \Phi(\xi)$ and the identity $|\partial\Phi^{-1}(\Phi(\xi))| = |\partial\Phi(\xi)|^{-1} = |\partial f / \partial \xi_1(\xi)|^{-1}$. Note that this quantity is independent of x_1 , finishing the proof of (2.1). \square

The following comparison principle is a straightforward consequence of Theorem 2.1:

Corollary 2.2. *Let $f, g \in C^1(\mathbb{R}^n)$ be real-valued functions such that, for almost all $\xi' = (\xi_2, \dots, \xi_n) \in \mathbb{R}^{n-1}$, $f(\xi_1, \xi')$ and $g(\xi_1, \xi')$ are strictly monotone in ξ_1 on the support of a measurable function χ on \mathbb{R}^n . Let $\sigma, \tau \in C^0(\mathbb{R}^n)$ be such that, for some $A > 0$, we have*

$$(2.2) \quad \frac{|\sigma(\xi)|}{|D_1 f(\xi)|^{1/2}} \leq A \frac{|\tau(\xi)|}{|D_1 g(\xi)|^{1/2}}$$

for all $\xi \in \text{supp } \chi$ satisfying $D_1 f(\xi) \neq 0$ and $D_1 g(\xi) \neq 0$. Then we have

$$(2.3) \quad \begin{aligned} & \|\chi(D_x)\sigma(D_x)e^{itf(D_x)}\varphi(x_1, x')\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^{n-1})} \\ & \leq A \|\chi(D_x)\tau(D_x)e^{itg(D_x)}\varphi(\tilde{x}_1, x')\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^{n-1})} \end{aligned}$$

for all $x_1, \tilde{x}_1 \in \mathbb{R}$, where $x' = (x_2, \dots, x_n) \in \mathbb{R}^{n-1}$. Consequently, for any measurable function w on \mathbb{R} we have

$$(2.4) \quad \begin{aligned} & \|w(x_1)\chi(D_x)\sigma(D_x)e^{itf(D_x)}\varphi(x)\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \\ & \leq A \|w(x_1)\chi(D_x)\tau(D_x)e^{itg(D_x)}\varphi(x)\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \end{aligned}$$

Moreover, if $\chi \in C^0(\mathbb{R}^n)$ and $w \neq 0$ on a set of \mathbb{R} with positive measure, the converse is true, namely, if we have estimate (2.3) for all φ , for some $x_1, \tilde{x}_1 \in \mathbb{R}$, or if we

have estimate (2.4) for all φ , and the norms are finite, then we also have inequality (2.2).

We remark that the last inequality in Corollary 2.2 gives the comparison between different weighted estimates. The reason to introduce function χ into the estimates is that the relation between symbols may be different for different regions of the frequencies ξ , (for example this is the case for the relativistic Schrödinger and for the Klein-Gordon equations which will be discussed in Section 8), so we have freedom to choose different σ for different types of behaviour of f' . The assumption $\sigma, \tau \in C^0(\mathbb{R}^n)$ made there is for the clarity of the exposition and can clearly be relaxed. We will not need it in this paper, but if σ and τ are simply measurable, satisfy (2.2) almost everywhere, and if all the integrals make sense, the conclusion of Corollary 2.2 and subsequent results continue to hold.

In the case $n = 1$, we neglect $x' = (x_2, \dots, x_n)$ in a natural way and just write $x = x_1$, $\xi = \xi_1$, and $D_x = D_1$. Similarly in the case $n = 2$, we use the notation $(x, y) = (x_1, x_2)$, $(\xi, \eta) = (\xi_1, \xi_2)$, and $(D_x, D_y) = (D_1, D_2)$. In both cases, we write $\tilde{x} = \tilde{x}_1$ in notation of Corollary 2.2. Then we have the following corollaries:

Corollary 2.3. *Suppose $n = 1$. Let $f, g \in C^1(\mathbb{R})$ be real-valued and strictly monotone on the support of a measurable function χ on \mathbb{R} . Let $\sigma, \tau \in C^0(\mathbb{R})$ be such that, for some $A > 0$, we have*

$$(2.5) \quad \frac{|\sigma(\xi)|}{|f'(\xi)|^{1/2}} \leq A \frac{|\tau(\xi)|}{|g'(\xi)|^{1/2}}$$

for all $\xi \in \text{supp } \chi$ satisfying $f'(\xi) \neq 0$ and $g'(\xi) \neq 0$. Then we have

$$(2.6) \quad \|\chi(D_x)\sigma(D_x)e^{itf(D_x)}\varphi(x)\|_{L^2(\mathbb{R}_t)} \leq A\|\chi(D_x)\tau(D_x)e^{itg(D_x)}\varphi(\tilde{x})\|_{L^2(\mathbb{R}_t)}$$

for all $x, \tilde{x} \in \mathbb{R}$. Consequently, for general $n \geq 1$ and for any measurable function w on \mathbb{R}^n , we have

$$(2.7) \quad \|w(x)\chi(D_j)\sigma(D_j)e^{itf(D_j)}\varphi(x)\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq A\|w(x)\chi(D_j)\tau(D_j)e^{itg(D_j)}\varphi(x)\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)},$$

where $j = 1, 2, \dots, n$. Moreover, if $\chi \in C^0(\mathbb{R})$ and $w \neq 0$ on a set of \mathbb{R}^n with positive measure, the converse is true, namely, if we have estimate (2.6) for all φ , for some $x, \tilde{x} \in \mathbb{R}$, or if we have estimate (2.6) for all φ , and the norms are finite, then we also have inequality (2.5).

Corollary 2.4. *Suppose $n = 2$. Let $f, g \in C^1(\mathbb{R}^2)$ be real-valued functions such that, for almost all $\eta \in \mathbb{R}$, $f(\xi, \eta)$ and $g(\xi, \eta)$ are strictly monotone in ξ on the support of a measurable function χ on \mathbb{R}^2 . Let $\sigma, \tau \in C^0(\mathbb{R}^2)$ be such that, for some $A > 0$, we have*

$$(2.8) \quad \frac{|\sigma(\xi, \eta)|}{|\partial f / \partial \xi(\xi, \eta)|^{1/2}} \leq A \frac{|\tau(\xi, \eta)|}{|\partial g / \partial \xi(\xi, \eta)|^{1/2}}$$

for all $(\xi, \eta) \in \text{supp } \chi$ satisfying $\partial f / \partial \xi(\xi, \eta) \neq 0$ and $\partial g / \partial \xi(\xi, \eta) \neq 0$. Then we have

$$(2.9) \quad \begin{aligned} & \|\chi(D_x, D_y)\sigma(D_x, D_y)e^{itf(D_x, D_y)}\varphi(x, y)\|_{L^2(\mathbb{R}_t \times \mathbb{R}_y)} \\ & \leq A\|\chi(D_x, D_y)\tau(D_x, D_y)e^{itg(D_x, D_y)}\varphi(\tilde{x}, y)\|_{L^2(\mathbb{R}_t \times \mathbb{R}_y)} \end{aligned}$$

for all $x, \tilde{x} \in \mathbb{R}$. Consequently, for general $n \geq 2$ and for any measurable function w on \mathbb{R}^{n-1} we have

$$(2.10) \quad \begin{aligned} & \|w(\tilde{x}_k)\chi(D_j, D_k)\sigma(D_j, D_k)e^{itf(D_j, D_k)}\varphi(x)\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \\ & \leq A\|w(\tilde{x}_k)\chi(D_j, D_k)\tau(D_j, D_k)e^{itg(D_j, D_k)}\varphi(x)\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)}, \end{aligned}$$

where $j \neq k$ and $\tilde{x}_k = (x_1, \dots, x_{k-1}, x_{k+1}, \dots, x_n)$. Moreover, if $\chi \in C^0(\mathbb{R}^2)$ and $w \neq 0$ on a set of \mathbb{R}^{n-1} with positive measure, the converse is true, namely, if we have estimate (2.9) for all φ , for some $x, \tilde{x} \in \mathbb{R}$, or if we have estimate (2.9) for all φ , and the norms are finite, then we also have inequality (2.8).

By the same argument as used in the proof of Theorem 2.1 and Corollary 2.2, we have a comparison result for radially symmetric case. Below, we denote the set of the positive real numbers $(0, \infty)$ by \mathbb{R}_+ .

Theorem 2.5. *Let $f, g \in C^1(\mathbb{R}_+)$ be real-valued and strictly monotone on the support of a measurable function χ on \mathbb{R}_+ . Let $\sigma, \tau \in C^0(\mathbb{R}_+)$ be such that, for some $A > 0$, we have*

$$(2.11) \quad \frac{|\sigma(\rho)|}{|f'(\rho)|^{1/2}} \leq A \frac{|\tau(\rho)|}{|g'(\rho)|^{1/2}}$$

for all $\rho \in \text{supp } \chi$ satisfying $f'(\rho) \neq 0$ and $g'(\rho) \neq 0$. Then we have

$$(2.12) \quad \|\chi(|D_x|)\sigma(|D_x|)e^{itf(|D_x|)}\varphi(x)\|_{L^2(\mathbb{R}_t)} \leq A\|\chi(|D_x|)\tau(|D_x|)e^{itg(|D_x|)}\varphi(x)\|_{L^2(\mathbb{R}_t)}$$

for all $x \in \mathbb{R}^n$. Consequently, for any measurable function w on \mathbb{R}^n , we have

$$(2.13) \quad \begin{aligned} & \|w(x)\chi(|D_x|)\sigma(|D_x|)e^{itf(|D_x|)}\varphi(x)\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \\ & \leq A\|w(x)\chi(|D_x|)\tau(|D_x|)e^{itg(|D_x|)}\varphi(x)\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)}. \end{aligned}$$

Moreover, if $\chi \in C^0(\mathbb{R}_+)$ and $w \neq 0$ on a set of \mathbb{R}^n with positive measure, the converse is true, namely, if we have estimate (2.12) for all φ , for some $x \in \mathbb{R}^n$, or if we have estimate (2.13) for all φ , and the norms are finite, then we also have inequality (2.11).

Proof. Below, we will write $\xi = \rho\omega$, where $\rho > 0$ and $\omega \in \mathbb{S}^{n-1}$. As usual we perform calculations on the set $f'(\rho) \neq 0$, where the inverse of f is differentiable. We have

$$\begin{aligned} & \chi(|D_x|)\sigma(|D_x|)e^{itf(|D_x|)}\varphi(x) \\ & = (2\pi)^{-n} \int_{\mathbb{R}^n} e^{itf(|\xi|)} e^{ix \cdot \xi} (\chi\sigma)(|\xi|) \widehat{\varphi}(\xi) d\xi \\ & = (2\pi)^{-n} \int_{\mathbb{R}_+} \int_{\mathbb{S}^{n-1}} e^{itf(\rho)} e^{i\rho x \cdot \omega} (\chi\sigma)(\rho) \widehat{\varphi}(\rho\omega) \rho^{n-1} d\rho d\omega \\ & = (2\pi)^{-n} \int_{f(\mathbb{R}_+)} \int_{\mathbb{S}^{n-1}} e^{it\eta} e^{if^{-1}(\eta)x \cdot \omega} (\chi\sigma)(f^{-1}(\eta)) \widehat{\varphi}(f^{-1}(\eta)\omega) f^{-1}(\eta)^{n-1} |(f^{-1})'(\eta)| d\omega d\eta, \end{aligned}$$

where we used a substitution $\rho = f^{-1}(\eta)$ on the support of χ . Using Plancherel's identity, we get

$$\begin{aligned}
& \|\chi(|D_x|)\sigma(|D_x|)e^{itf(|D_x|)}\varphi(x)\|_{L^2(\mathbb{R}_t)}^2 \\
&= (2\pi)^{-2n+1} \int_{f(\mathbb{R}_+)} d\eta \times \\
(2.14) \quad & \times \left| \int_{\mathbb{S}^{n-1}} e^{if^{-1}(\eta)x \cdot \omega} (\chi\sigma)(f^{-1}(\eta)) \widehat{\varphi}(f^{-1}(\eta)\omega) f^{-1}(\eta)^{n-1} |(f^{-1})'(\eta)| d\omega \right|^2 \\
&= (2\pi)^{-2n+1} \int_{\mathbb{R}_+} \left| \int_{\mathbb{S}^{n-1}} e^{i\rho x \cdot \omega} (\chi\sigma)(\rho) \widehat{\varphi}(\rho\omega) \rho^{n-1} |(f^{-1})'(f(\rho))| d\omega \right|^2 |f'(\rho)| d\rho \\
&= (2\pi)^{-2n+1} \int_{\mathbb{R}_+} \left| \int_{\mathbb{S}^{n-1}} e^{i\rho x \cdot \omega} \widehat{\varphi}(\rho\omega) d\omega \right|^2 \rho^{2(n-1)} |\chi(\rho)|^2 \frac{|\sigma(\rho)|^2}{|f'(\rho)|} d\rho,
\end{aligned}$$

where we have used the substitution $\eta = f(\rho)$ again and the identity $(f^{-1})'(f(\rho)) = f'(\rho)^{-1}$. From assumption (2.11) it follows that

$$\begin{aligned}
& \|\chi(|D_x|)\sigma(|D_x|)e^{itf(|D_x|)}\varphi(x)\|_{L^2(\mathbb{R}_t)}^2 \\
& \leq (2\pi)^{-2n+1} A^2 \int_{\mathbb{R}_+} \left| \int_{\mathbb{S}^{n-1}} e^{i\rho x \cdot \omega} \widehat{\varphi}(\rho\omega) d\omega \right|^2 \rho^{2(n-1)} |\chi(\rho)|^2 \frac{|\tau(\rho)|^2}{|g'(\rho)|} d\rho \\
& = A^2 \|\chi(|D_x|)\tau(|D_x|)e^{itg(|D_x|)}\varphi(x)\|_{L^2(\mathbb{R}_t)}^2,
\end{aligned}$$

finishing the proof of (2.12). Estimate (2.13) follows from it immediately. The converse is also obtained from equality (2.14) which holds for any (radially symmetric) function φ . \square

3. EQUIVALENT MODEL ESTIMATES

Let us now give important examples of the use of the comparison principle described in Section 2. We still use the same notation as in Section 2. That is, denoting the dimension of the variable x by n , we write $x = (x_1, \dots, x_n)$ and $D_x = (D_1, D_2, \dots, D_n)$. We just write $x = x_1$, $D_x = D_1$ in the case $n = 1$, and $(x, y) = (x_1, x_2)$, $(D_x, D_y) = (D_1, D_2)$ in the case $n = 2$.

If both sides in expression (2.2) in Corollary 2.2 are equivalent, we can use the comparison in two directions, from which it follows that norms on both sides in (2.3) are equivalent. The same is true for Corollaries 2.3, 2.4 and Theorem 2.5. In particular, we can conclude that many smoothing estimates for the Schrödinger type equations of different orders are equivalent to each other. Indeed, applying Corollary 2.3 in two directions, we immediately obtain that for $n = 1$ and $l, m > 0$, we have

$$(3.1) \quad \left\| |D_x|^{(m-1)/2} e^{it|D_x|^m} \varphi(x) \right\|_{L^2(\mathbb{R}_t)} = \sqrt{\frac{l}{m}} \left\| |D_x|^{(l-1)/2} e^{it|D_x|^l} \varphi(x) \right\|_{L^2(\mathbb{R}_t)}$$

for every $x \in \mathbb{R}$, assuming that $\text{supp } \widehat{\varphi} \subset [0, +\infty)$ or $(-\infty, 0]$. Applying Corollary 2.4, we similarly obtain that for $n = 2$ and $l, m > 0$, we have

$$(3.2) \quad \left\| |D_y|^{(m-1)/2} e^{itD_x|D_y|^{m-1}} \varphi(x, y) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_y)} \\ = \left\| |D_y|^{(l-1)/2} e^{itD_x|D_y|^{l-1}} \varphi(x, y) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_y)}$$

for every $x \in \mathbb{R}$. On the other hand, in the case $n = 1$, we have easily

$$(3.3) \quad \left\| e^{itD_x} \varphi(x) \right\|_{L^2(\mathbb{R}_t)} = \|\varphi\|_{L^2(\mathbb{R}_x)} \quad \text{for all } x \in \mathbb{R},$$

which is a straightforward consequence of the fact $e^{itD_x} \varphi(x) = \varphi(x + t)$. By using equality (3.3), we can estimate the right hand sides of equalities (3.1) and (3.2) with $l = 1$, and as a result, we have easily the following variety of pointwise estimates in low dimensions:

Theorem 3.1. *Suppose $n = 1$ and $m > 0$. Then we have*

$$(3.4) \quad \left\| |D_x|^{(m-1)/2} e^{it|D_x|^m} \varphi(x) \right\|_{L^2(\mathbb{R}_t)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x)}$$

for all $x \in \mathbb{R}$. Suppose $n = 2$ and $m > 0$. Then we have

$$(3.5) \quad \left\| |D_y|^{(m-1)/2} e^{itD_x|D_y|^{m-1}} \varphi(x, y) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_y)} \leq C \|\varphi\|_{L^2(\mathbb{R}_{x,y}^2)}$$

for all $x \in \mathbb{R}$. Each estimate above is equivalent to itself with $m = 1$ which is a direct consequence of equality (3.3). In particular, we have equalities (3.1) and (3.2).

Estimates (3.4) and (3.5) in Theorem 3.1 in the special case $m = 2$ were shown by Kenig, Ponce and Vega [KPV1, p.56] and by Linares and Ponce [LP, p.528], respectively. Theorem 3.1 shows that these results, together with their generalisation to other orders m , are in fact just corollaries of the elementary one dimensional fact $e^{itD_x} \varphi(x) = \varphi(x + t)$ once we apply the comparison principle.

By using the comparison principle in the radially symmetric case, we have also another type of equivalence of smoothing estimates. In fact, by Theorem 2.5, we immediately obtain

$$\left\| |x|^{\beta-1} |D_x|^\beta e^{it|D_x|^2} \varphi \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} = \sqrt{\frac{m}{2}} \left\| |x|^{\beta-1} |D_x|^{m/2+\beta-1} e^{it|D_x|^m} \varphi \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \\ = \sqrt{\frac{m}{2}} \left\| |x|^{\alpha-m/2} |D_x|^\alpha e^{it|D_x|^m} \varphi \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)},$$

where $m > 0$ and $\alpha = m/2 + \beta - 1$. On the other hand, we know the estimate

$$(3.6) \quad \left\| |x|^{\beta-1} |D_x|^\beta e^{it|D_x|^2} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)} \quad (1 - n/2 < \beta < 1/2),$$

which was given by Sugimoto [Su1, Theorem 1.1]. Noticing that $1 - n/2 < \beta < 1/2$ is equivalent to $(m - n)/2 < \alpha < (m - 1)/2$, we have the estimate

$$(3.7) \quad \left\| |x|^{\alpha-m/2} |D_x|^\alpha e^{it|D_x|^m} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)} \\ (m > 0, \quad (m - n)/2 < \alpha < (m - 1)/2).$$

We note that estimate (3.6) is a special case ($m = 2$) of estimate (3.7), but the comparison principle of Section 2 shows that they are equivalent to each other.

We remark that estimate (3.6) is implied from its restricted version

$$(3.8) \quad \left\| |x|^{\beta-1} |D_x|^\beta e^{it|D_x|^2} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)} \quad (1/2 - \varepsilon \leq \beta < 1/2),$$

where $\varepsilon > 0$ is sufficiently small. (The case $0 < \varepsilon < 1/2$ is the result of Kato and Yajima [KY], and the critical case of this estimate with $\varepsilon = 0$ was given in [Su2] and explained geometrically in [RS3]). In fact, estimate (3.6) with $1 - n/2 < \beta < 1/2 - \varepsilon$ can be reduced to the one with $\beta = 1/2 - \varepsilon$ if we use the estimate

$$\left\| |x|^{\beta-1} |D_x|^\beta v \right\|_{L^2(\mathbb{R}^n)} \leq C \left\| |x|^{(1/2-\varepsilon)-1} |D_x|^{1/2-\varepsilon} v \right\|_{L^2(\mathbb{R}^n)}$$

which is a consequence of the following lemma.

Lemma 3.1 ([SW], Theorem B*). *Suppose $k < n/2$, $l < n/2$, $0 < m < n$, and $k + l + m = n$. Then the operator $|x|^{-l} |D_x|^{m-n} |x|^{-k}$ is $L^2(\mathbb{R}^n)$ -bounded.*

Furthermore, we can show that in fact estimate (3.7) is also equivalent to estimate

$$(3.9) \quad \left\| \langle x \rangle^{-m/2} e^{it|D_x|^m} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)} \quad (n > m > 1)$$

for all φ such that $\text{supp } \widehat{\varphi} \subset \{\xi \in \mathbb{R}^n : |\xi| \leq 1\}$

(given by Walther in [Wa2, Theorem 4.1]). In fact, estimate (3.9) is a direct consequence of estimate (3.7) with $\alpha = 0$ if we notice a trivial inequality $\langle x \rangle^{-m/2} \leq |x|^{-m/2}$. Note also that the assumption $n > m > 1$ assures $(m - n)/2 < \alpha = 0 < (m - 1)/2$. On the other hand, by Theorem 2.5 we have

$$\begin{aligned} & \left\| \langle x \rangle^{\alpha-m/2} \chi(|D_x|) |D_x|^\alpha e^{it|D_x|^m} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \\ &= \sqrt{\frac{\mu}{m}} \left\| \langle x \rangle^{\alpha-m/2} \chi(|D_x|) |D_x|^{\alpha+(\mu-m)/2} e^{it|D_x|^\mu} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \\ &= \sqrt{\frac{m-2\alpha}{m}} \left\| \langle x \rangle^{-\mu/2} \chi(|D_x|) e^{it|D_x|^\mu} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)}, \end{aligned}$$

where $m > 0$ and $\mu = m - 2\alpha > 0$. Hence, from estimate (3.9) with $m = \mu$, we obtain

$$\left\| \langle x \rangle^{\alpha-m/2} \chi(|D_x|) |D_x|^\alpha e^{it|D_x|^m} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)},$$

where $n > m - 2\alpha > 1$, or equivalently $(m - n)/2 < \alpha < (m - 1)/2$. Here we take a cut-off function $\chi(\rho) \in C_0^\infty([0, 1])$ such that $\chi(\rho) \equiv 1$ for $\rho \leq 1/2$. From this estimate, we obtain estimate (3.7). In fact, we have the equality

$$\begin{aligned} & \left\| |x|^{\alpha-m/2} |D_x|^\alpha e^{it|D_x|^m} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \\ &= \lim_{\lambda \searrow 0} \left\| \lambda^{\alpha-m/2} \langle x/\lambda \rangle^{\alpha-m/2} \chi(\lambda|D_x|) |D_x|^\alpha e^{it|D_x|^m} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)}, \end{aligned}$$

and noticing the identities $\|g(t, x)\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} = \lambda^{m/2+n/2} \|g(\lambda^m t, \lambda x)\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)}$ and $(m(\lambda D_x)\varphi)(\lambda x) = m(D_x)(\varphi(\lambda \cdot))(x)$, we have

$$\begin{aligned} & \left\| |x|^{\alpha-m/2} |D_x|^\alpha e^{it|D_x|^m} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \\ & \leq \sup_{\lambda>0} \left\| \langle x \rangle^{\alpha-m/2} \chi(|D_x|) |D_x|^\alpha e^{it|D_x|^m} \varphi_\lambda(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)}, \end{aligned}$$

where $\varphi_\lambda(x) = \lambda^{n/2} \varphi(\lambda x)$. Note also that $\|\varphi_\lambda\|_{L^2(\mathbb{R}_x^n)} = \|\varphi\|_{L^2(\mathbb{R}_x^n)}$.

Finally we remark that the last inequality implies

$$\left\| |x|^{\alpha-m/2} |D_x|^\alpha e^{it|D_x|^m} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq \sup_{\lambda>0} \left\| \langle x \rangle^{\alpha-m/2} |D_x|^\alpha e^{it|D_x|^m} \varphi_\lambda(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)}$$

by the comparison principle Theorem 2.5. Thus we can conclude the following:

Theorem 3.2. *We have equivalent estimates (3.6), (3.7), and (3.9). Furthermore, they are equivalent to estimate (3.8) with sufficiently small $\varepsilon > 0$. In particular, for $m > 0$ (and any α, β) we have the following relations (which are finite for α, β as in the above estimates)*

$$\begin{aligned} & \left\| |x|^{\beta-1} |D_x|^\beta e^{it|D_x|^2} \varphi \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} = \sqrt{\frac{m}{2}} \left\| |x|^{\beta-1} |D_x|^{m/2+\beta-1} e^{it|D_x|^m} \varphi \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)}, \\ & \left\| \langle x \rangle^{\alpha-m/2} |D_x|^\alpha e^{it|D_x|^m} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq \left\| |x|^{\alpha-m/2} |D_x|^\alpha e^{it|D_x|^m} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \\ & \leq \sup_{\lambda>0} \left\| \langle x \rangle^{\alpha-m/2} |D_x|^\alpha e^{it|D_x|^m} \varphi_\lambda(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)}, \end{aligned}$$

where $\varphi_\lambda(x) = \lambda^{n/2} \varphi(\lambda x)$, and we take $\alpha \leq m/2$ in the last estimate. The operator norms of operators $\langle x \rangle^{\alpha-m/2} |D_x|^\alpha e^{it|D_x|^m}$ and $|x|^{\alpha-m/2} |D_x|^\alpha e^{it|D_x|^m}$ as mappings from $L^2(\mathbb{R}^n)$ to $L^2(\mathbb{R}_t \times \mathbb{R}_x^n)$ are equal.

As a nice consequence, for $n \geq 3$ and $m > 0$ we can conclude also the estimate

$$(3.10) \quad \left\| |x|^{-1} |D_x|^{m/2-1} e^{it|D_x|^m} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq \sqrt{\frac{2\pi}{m(n-2)}} \|\varphi\|_{L^2(\mathbb{R}_x^n)},$$

where the constant $\sqrt{\frac{2\pi}{m(n-2)}}$ is sharp. This follows from the first equality in Theorem 3.2 with $\beta = 0$ and the fact that the constant $C = \sqrt{\frac{\pi}{n-2}}$ is sharp in (3.6) with $\beta = 0$, as shown by Simon [Si] as a consequence of constants in Kato's theory [Ka1].

In general, best constants in the radially symmetric case can be obtained by changing to spherical harmonics and looking at the appearing one dimensional integral. Thus, if $n \geq 2$ and f is injective and differentiable on $(0, \infty)$, the best constant in the inequality

$$\left\| w(|x|) \sigma(|D_x|) e^{itf(|D_x|)} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)}$$

is given by

$$C = (2\pi)^{(n+1)/2} \left(\sup_{\substack{\rho>0 \\ k \in \mathbb{N}}} \left\{ \rho \sigma(\rho)^2 f'(\rho)^{-1} \int_0^\infty J_{\nu(k)}(r\rho)^2 w(r)^2 r dr \right\} \right)^{1/2},$$

where for $\lambda > -1/2$ the Bessel function J_λ of order λ is given by

$$J_\lambda(\rho) = \frac{\rho^\lambda}{2^\lambda \Gamma(\lambda + 1/2) \Gamma(1/2)} \int_{-1}^1 e^{i\rho r} (1 - r^2)^{\lambda-1/2} dr,$$

and $\nu(k) = n/2 + k - 1$. This expression was obtained by Walther [Wa2], and it can be used to analyse estimates for radially symmetric equations by carefully looking at the asymptotic behaviour of Bessel functions and subsequent integrals.

The estimates listed in Theorems 3.1 and 3.2 will act as model ones later. In the subsequent sections, further smoothing results will be derived from them, hence from simple estimates (3.3) and (3.8), by the (introduced further) method of canonical transformations or some combination use of it and the comparison principle. The following are straightforward results of Theorems 3.1 and 3.2:

Corollary 3.3. *Suppose $n \geq 1$, $m > 0$, and $s > 1/2$. Then we have*

$$(3.11) \quad \left\| \langle x_n \rangle^{-s} |D_n|^{(m-1)/2} e^{it|D_n|^m} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)}.$$

Suppose $n \geq 2$, $m > 0$, and $s > 1/2$. Then we have

$$(3.12) \quad \left\| \langle x_1 \rangle^{-s} |D_n|^{(m-1)/2} e^{itD_1|D_n|^{m-1}} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)}.$$

Proof. Use first the square integrability of $\langle x_n \rangle^{-s}$ in one dimension, then apply estimate (3.4) in x_n to obtain estimate (3.11). Similarly estimate (3.12) is obtained from estimate (3.5). \square

Corollary 3.4. *Suppose $m > 0$ and $(m - n)/2 < \alpha < (m - 1)/2$. Then we have*

$$(3.13) \quad \left\| |x|^{\alpha-m/2} |D_x|^\alpha e^{it|D_x|^m} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)}.$$

Suppose $m > 0$ and $(m - n + 1)/2 < \alpha < (m - 1)/2$. Then we have

$$(3.14) \quad \left\| |x|^{\alpha-m/2} |D'|^\alpha e^{it(|D_1|^m - |D'|^m)} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)},$$

where $D' = (D_2, \dots, D_n)$.

Proof. Estimate (3.13) is the same one as estimate (3.7). From estimate (3.13) in $x' \in \mathbb{R}^{n-1}$, where $x' = (x_2, \dots, x_n)$, and Plancherel's theorem in x_1 , we obtain estimate (3.14) if we notice the trivial inequality $|x|^{\alpha-m/2} \leq |x'|^{\alpha-m/2}$. \square

By using the comparison principle again, we can compare many estimates with the model estimates stated above, which have been also induced by the comparison principle from the trivial estimate (3.3) and so on. For example, in notation of Corollary 2.3, setting $\tau(\xi) = |\xi|^{(m-1)/2}$ and $g(\xi) = |\xi|^m$, we have $|\tau(\xi)|/|g'(\xi)|^{1/2} = m^{-1/2}$. Similarly in notation of Corollary 2.4, setting $\tau(\xi, \eta) = |\eta|^{(m-1)/2}$ and $g(\xi, \eta) = \xi|\eta|^{m-1}$, we have $|\tau(\xi, \eta)|/|\partial g/\partial \xi(\xi, \eta)|^{1/2} = 1$. Hence, noticing that $\chi(D_x)$ is L^2 -bounded for $\chi \in L^\infty$, we obtain the following *secondary comparison* results from Corollary 3.3.

Corollary 3.5. *Suppose $n \geq 1$ and $s > 1/2$. Let $\chi \in L^\infty(\mathbb{R})$. Let $f \in C^1(\mathbb{R})$ be real-valued and strictly monotone on $\text{supp } \chi$. Let $\sigma \in C^0(\mathbb{R})$ be such that, for some $A > 0$, we have*

$$|\sigma(\xi)| \leq A |f'(\xi)|^{1/2}$$

for all $\xi \in \text{supp } \chi$. Then we have

$$\left\| \langle x_j \rangle^{-s} \chi(D_j) \sigma(D_j) e^{itf(D_j)} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)},$$

where $j = 1, 2, \dots, n$.

Corollary 3.6. *Suppose $n \geq 2$ and $s > 1/2$. Let $\chi \in L^\infty(\mathbb{R}^2)$. Let $f \in C^1(\mathbb{R}^2)$ be a real-valued function such that, for almost all $\eta \in \mathbb{R}$, $f(\xi, \eta)$ is strictly monotone in ξ on $\text{supp } \chi$. Let $\sigma \in C^0(\mathbb{R}^2)$ be such that for some $A > 0$ we have*

$$|\sigma(\xi, \eta)| \leq A \left| \frac{\partial f}{\partial \xi}(\xi, \eta) \right|^{1/2}$$

for all $(\xi, \eta) \in \text{supp } \chi$. Then we have

$$\left\| \langle x_j \rangle^{-s} \chi(D_j, D_k) \sigma(D_j, D_k) e^{itf(D_j, D_k)} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)},$$

where $j \neq k$.

Applications of these secondary comparison results will be given in Section 7. We can also have results for radially symmetric operators if we compare them with estimate (3.13) of Corollary 3.4 by using Theorem 2.5. We will discuss it later in the last part of Section 5 together with more different type of results.

4. CANONICAL TRANSFORMS

Based on the argument in the introduction, we will now introduce the main tool to reduce general operators to normal forms. That is the canonical transformation which changes the equation

$$\begin{cases} (i\partial_t + a(D_x)) u(t, x) = 0, \\ u(0, x) = \varphi(x), \end{cases} \quad \text{to} \quad \begin{cases} (i\partial_t + \sigma(D_x)) v(t, x) = 0, \\ v(0, x) = g(x), \end{cases}$$

where $a(D_x)$ and $\sigma(D_x)$ are related with each other as in the relation (1.12) in the introduction, i.e. we have $a(\xi) = (\sigma \circ \psi)(\xi)$. If the initial data $\varphi(x)$ is the corresponding transform of $g(x)$, then the solution $u(t, x) = e^{ita(D_x)} \varphi(x)$ is the corresponding transform of $v(t, x) = e^{it\sigma(D_x)} g(x)$. In this way, we will reduce general smoothing estimates to model ones listed in Section 3.

Now we will describe this more precisely. Let $\Gamma, \tilde{\Gamma} \subset \mathbb{R}^n$ be open sets and $\psi : \Gamma \rightarrow \tilde{\Gamma}$ be a C^∞ -diffeomorphism (we do not assume them to be cones since we do not require homogeneity of phases). We always assume that

$$(4.1) \quad C^{-1} \leq |\det \partial\psi(\xi)| \leq C \quad (\xi \in \Gamma),$$

for some $C > 0$. We set formally

$$(4.2) \quad \begin{aligned} I_\psi u(x) &= \mathcal{F}^{-1} [\mathcal{F}u(\psi(\xi))] (x) = (2\pi)^{-n} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} e^{i(x \cdot \xi - y \cdot \psi(\xi))} u(y) dy d\xi, \\ I_\psi^{-1} u(x) &= \mathcal{F}^{-1} [\mathcal{F}u(\psi^{-1}(\xi))] (x) = (2\pi)^{-n} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} e^{i(x \cdot \xi - y \cdot \psi^{-1}(\xi))} u(y) dy d\xi. \end{aligned}$$

The operators I_ψ and I_ψ^{-1} can be justified by using cut-off functions $\gamma \in C^\infty(\Gamma)$ and $\tilde{\gamma} = \gamma \circ \psi^{-1} \in C^\infty(\tilde{\Gamma})$ which satisfy $\text{supp } \gamma \subset \Gamma$, $\text{supp } \tilde{\gamma} \subset \tilde{\Gamma}$. We set

$$(4.3) \quad \begin{aligned} I_{\psi, \gamma} u(x) &= \mathcal{F}^{-1} [\gamma(\xi) \mathcal{F}u(\psi(\xi))] (x) \\ &= (2\pi)^{-n} \int_{\mathbb{R}^n} \int_{\Gamma} e^{i(x \cdot \xi - y \cdot \psi(\xi))} \gamma(\xi) u(y) dy d\xi, \\ I_{\psi, \gamma}^{-1} u(x) &= \mathcal{F}^{-1} [\tilde{\gamma}(\xi) \mathcal{F}u(\psi^{-1}(\xi))] (x) \\ &= (2\pi)^{-n} \int_{\mathbb{R}^n} \int_{\tilde{\Gamma}} e^{i(x \cdot \xi - y \cdot \psi^{-1}(\xi))} \tilde{\gamma}(\xi) u(y) dy d\xi. \end{aligned}$$

In the case that $\Gamma, \tilde{\Gamma} \subset \mathbb{R}^n \setminus 0$ are open cones, we may consider the homogeneous ψ and γ which satisfy $\text{supp } \gamma \cap \mathbb{S}^{n-1} \subset \Gamma \cap \mathbb{S}^{n-1}$ and $\text{supp } \tilde{\gamma} \cap \mathbb{S}^{n-1} \subset \tilde{\Gamma} \cap \mathbb{S}^{n-1}$, where $\mathbb{S}^{n-1} = \{\xi \in \mathbb{R}^n : |\xi| = 1\}$. Then we have the expressions for compositions

$$(4.4) \quad I_{\psi, \gamma} = \gamma(D_x) \cdot I_\psi = I_\psi \cdot \tilde{\gamma}(D_x), \quad I_{\psi, \gamma}^{-1} = \tilde{\gamma}(D_x) \cdot I_\psi^{-1} = I_\psi^{-1} \cdot \gamma(D_x),$$

and the identities

$$(4.5) \quad I_{\psi, \gamma} \cdot I_{\psi, \gamma}^{-1} = \gamma(D_x)^2, \quad I_{\psi, \gamma}^{-1} \cdot I_{\psi, \gamma} = \tilde{\gamma}(D_x)^2.$$

We have also the formula

$$(4.6) \quad I_{\psi, \gamma} \cdot \sigma(D_x) = (\sigma \circ \psi)(D_x) \cdot I_{\psi, \gamma}, \quad I_{\psi, \gamma}^{-1} \cdot (\sigma \circ \psi)(D_x) = \sigma(D_x) \cdot I_{\psi, \gamma}^{-1}.$$

We also introduce the weighted L^2 -spaces. For the weight function $w(x)$, let $L_w^2(\mathbb{R}^n; w)$ be the set of measurable functions $f : \mathbb{R}^n \rightarrow \mathbb{C}$ such that the norm

$$\|f\|_{L^2(\mathbb{R}^n; w)} = \left(\int_{\mathbb{R}^n} |w(x) f(x)|^2 dx \right)^{1/2}$$

is finite. Then we have the following fundamental theorem:

Theorem 4.1. *Assume that the operator $I_{\psi, \gamma}$ defined by (4.3) is $L^2(\mathbb{R}^n; w)$ -bounded. Suppose that we have the estimate*

$$(4.7) \quad \|w(x) \rho(D_x) e^{it\sigma(D_x)} \varphi(x)\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)}$$

for all φ such that $\text{supp } \hat{\varphi} \subset \text{supp } \tilde{\gamma}$, where $\tilde{\gamma} = \gamma \circ \psi^{-1}$. Assume also that the function

$$(4.8) \quad q(\xi) = \frac{\gamma \cdot \zeta}{\rho \circ \psi}(\xi)$$

is bounded. Then we have

$$(4.9) \quad \|w(x) \zeta(D_x) e^{ita(D_x)} \varphi(x)\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)}$$

for all φ such that $\text{supp } \hat{\varphi} \subset \text{supp } \gamma$, where $a(\xi) = (\sigma \circ \psi)(\xi)$.

Proof. Substituting $I_{\psi,q}^{-1}\varphi$ for φ in (4.7), where $I_{\psi,q}^{-1} = I_{\psi}^{-1} \cdot q(D_x)$, we have

$$\|w(x)I_{\psi,q}^{-1}(\rho \circ \psi)(D_x)e^{ita(D_x)}\varphi(x)\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C\|I_{\psi,q}^{-1}\varphi\|_{L^2(\mathbb{R}_x^n)}$$

for φ such that $\text{supp } \widehat{\varphi} \subset \text{supp } \gamma$. Here we have noticed (4.6). Then we have

$$\|w(x)I_{\psi,\gamma}^{-1}\zeta(D_x)e^{ita(D_x)}\varphi(x)\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C\|I_{\psi,q}^{-1}\varphi\|_{L^2(\mathbb{R}_x^n)}.$$

By Plancherel's theorem, we have the L^2 -boundedness of $I_{\psi,q}^{-1}$ if we notice the assumption (4.1) and the boundedness of $q(\xi)$ given by (4.8). On the other hand, $I_{\psi,\gamma}$ is $L^2(\mathbb{R}^n; w)$ -bounded by the assumption, and we obtain (4.9) if we notice (4.5). \square

As for the $L^2(\mathbb{R}^n; w)$ -boundedness of the operator $I_{\psi,\gamma}$, we have criteria for some special weight functions. For $\kappa \in \mathbb{R}$, let $L_\kappa^2(\mathbb{R}^n)$, $\dot{L}_\kappa^2(\mathbb{R}^n)$ be the set of measurable functions f such that the norm

$$\|f\|_{L_\kappa^2(\mathbb{R}^n)} = \left(\int_{\mathbb{R}^n} |\langle x \rangle^\kappa f(x)|^2 dx \right)^{1/2}, \quad \|f\|_{\dot{L}_\kappa^2(\mathbb{R}^n)} = \left(\int_{\mathbb{R}^n} \|x\|^\kappa |f(x)|^2 dx \right)^{1/2}$$

is finite, respectively.

The following theorem is a simplified version of [RS2, Theorem 1.1] given by the authors, where the L_κ^2 -boundedness for more general x -dependent Fourier integral operators was treated under less restrictive conditions, with exact expressions for the numbers of derivatives, etc. These weighted boundedness results played an important role in the critical case of some of the smoothing estimates in [RS3]. They will be of crucial importance here as well.

Theorem 4.2. *Suppose $\kappa \in \mathbb{R}$. Assume that all the derivatives of entries of the $n \times n$ matrix $\partial\psi$ and those of γ are bounded. Then the operators $I_{\psi,\gamma}$ and $I_{\psi,\gamma}^{-1}$ defined by (4.3) are $L_\kappa^2(\mathbb{R}^n)$ -bounded.*

For homogeneous ψ and γ , we have another type of weighted boundedness result:

Theorem 4.3. *Let $\Gamma, \tilde{\Gamma} \subset \mathbb{R}^n \setminus 0$ be open cones. Suppose $|\kappa| < n/2$. Assume $\psi(\lambda\xi) = \lambda\psi(\xi)$, $\gamma(\lambda\xi) = \gamma(\xi)$ for all $\lambda > 0$ and $\xi \in \Gamma$. Then the operators $I_{\psi,\gamma}$ and $I_{\psi,\gamma}^{-1}$ defined by (4.3) are $L_\kappa^2(\mathbb{R}^n)$ -bounded and $\dot{L}_\kappa^2(\mathbb{R}^n)$ -bounded.*

We remark that the boundedness in Theorem 4.3 with the case $\kappa \leq 0$ is equivalent to the one with $\kappa \geq 0$ by the duality argument. In fact, the formal adjoint of I_ψ can be given by

$$\begin{aligned} I_\psi^* u(x) &= (2\pi)^{-n} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} e^{-i(y \cdot \xi - x \cdot \psi(\xi))} u(y) dy d\xi, \\ &= (2\pi)^{-n} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} e^{i(x \cdot \xi - y \cdot \psi^{-1}(\xi))} |\det \partial\psi^{-1}(\xi)| u(y) dy d\xi, \\ &= (2\pi)^{-n} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} e^{i(x \cdot \xi - y \cdot \psi^{-1}(\xi))} |\det \partial\psi(\psi^{-1}(\xi))|^{-1} u(y) dy d\xi, \\ &= I_\psi^{-1} \cdot |\det \partial\psi(D_x)|^{-1} u(x), \end{aligned}$$

from which we obtain the formula

$$I_{\psi,\gamma}^* = I_{\psi,d}^{-1}; \quad d(\xi) = |\det \partial\psi(\xi)|^{-1} \gamma(\xi).$$

Note that $d(\xi)$ satisfies the same property as that of $\gamma(\xi)$ in virtue of (4.1).

We also remark that the $L_\kappa^2(\mathbb{R}^n)$ -boundedness in Theorem 4.3 is equivalent to the $\dot{L}_\kappa^2(\mathbb{R}^n)$ -boundedness. In fact, the $L_\kappa^2(\mathbb{R}^n)$ -boundedness is a straightforward consequence of the $\dot{L}_\kappa^2(\mathbb{R}^n)$ -boundedness in the case $\kappa \geq 0$. On the other hand, the $L_\kappa^2(\mathbb{R}^n)$ -boundedness induces the $\dot{L}_\kappa^2(\mathbb{R}^n)$ -boundedness by the scaling argument because we have $I_{\psi,\gamma} D_\lambda = D_\lambda I_{\psi,\gamma}$, and also have

$$\lambda^{n/2+k} \|D_\lambda u\|_{L_\kappa^2(\mathbb{R}^n)} = \|(\lambda^2 + |x|^2)^{k/2} u(x)\|_{L^2(\mathbb{R}^n)} \rightarrow \|u\|_{\dot{L}_\kappa^2(\mathbb{R}^n)} \quad (\lambda \searrow 0),$$

where D_λ denotes the dilation operator $D_\lambda : u(x) \mapsto u(\lambda x)$.

We prepare a few lemmas which will be used to prove Theorem 4.3. The following two results are due to Kurtz and Wheeden [KW, Theorem 3], and Stein and Weiss [SW, Theorem B*] (see also Lemma 3.1), respectively.

Lemma 4.1. *Suppose $|\kappa| < n/2$. Assume that $m(\xi) \in C^n(\mathbb{R}^n \setminus 0)$ and all the derivative of $m(\xi)$ satisfies $|\partial^\gamma m(\xi)| \leq C_\gamma |\xi|^{-|\gamma|}$ for all $\xi \neq 0$ and $|\gamma| \leq n$. Then $m(D_x)$ is $L_\kappa^2(\mathbb{R}^n)$ and $\dot{L}_\kappa^2(\mathbb{R}^n)$ -bounded.*

Lemma 4.2. *Suppose $1 - n/2 < \kappa < n/2$. Then the operator $|D_x|^{-1}$ is $L_\kappa^2(\mathbb{R}^n)$ - $L_{\kappa-1}^2(\mathbb{R}^n)$ -bounded and $\dot{L}_\kappa^2(\mathbb{R}^n)$ - $\dot{L}_{\kappa-1}^2(\mathbb{R}^n)$ -bounded.*

We remark that, in Lemma 4.1, the $L_\kappa^2(\mathbb{R}^n)$ -boundedness is equivalent to the $\dot{L}_\kappa^2(\mathbb{R}^n)$ -boundedness, and the $L_\kappa^2(\mathbb{R}^n)$ - $L_{\kappa-1}^2(\mathbb{R}^n)$ -boundedness in Lemma 4.2 is also equivalent to the $\dot{L}_\kappa^2(\mathbb{R}^n)$ - $\dot{L}_{\kappa-1}^2(\mathbb{R}^n)$ -boundedness, by essentially the same argument as in the above remark.

Proof of Theorem 4.3. In view of the remarks below Theorem 4.3, it suffices to show the L_κ^2 -boundedness of $I_{\psi,\gamma}$ in the case $0 \leq \kappa < n/2$.

First we assume $n \geq 3$. If we note

$$e^{ix \cdot \xi} = \frac{1 - ix \cdot \partial_\xi}{\langle x \rangle^2} e^{ix \cdot \xi},$$

we can justify, by integration by parts,

$$\begin{aligned} I_{\psi,\gamma} u(x) &= (2\pi)^{-n} \int \int e^{i(x \cdot \xi - y \cdot \psi(\xi))} \gamma(\xi) u(y) dy d\xi \\ &= (2\pi)^{-n} \int \int e^{i(x \cdot \xi - y \cdot \psi(\xi))} \left(\frac{\gamma(\xi) + x \gamma(\xi)^t \partial \psi(\xi)^t y + ix \cdot \partial \gamma(\xi)}{\langle x \rangle^2} \right) u(y) dy d\xi, \end{aligned}$$

and have the formula

$$(4.10) \quad I_{\psi,\gamma} = \frac{1}{\langle x \rangle^2} I_{\psi,\gamma} + \frac{x}{\langle x \rangle^2} {}^t \partial \psi(D_x) I_{\psi,\gamma} {}^t x + i \frac{x}{\langle x \rangle^2} \cdot I_{\psi,\eta} |D_x|^{-1},$$

where $\eta(\xi) = |\psi(\xi)| \partial \gamma(\xi)$, and it satisfies the same assumption of the theorem as that of $\gamma(\xi)$. Assume that $I_{\psi,\gamma}$ is $L_{\kappa-1}^2$ -bounded under the assumption of the theorem.

Then, by the formula (4.10) and Lemmas 4.1 and Lemma 4.2, $I_{\psi,\gamma}$ is also L_κ^2 -bounded if $1 - n/2 < \kappa < n/2$. On the other hand, by Plancherel's theorem and assumption (4.1), we have the L^2 -boundedness of $I_{\psi,\gamma}$ under the assumption of the theorem. Then, by induction and the interpolation, we have the L_κ^2 -boundedness of $I_{\psi,\gamma}$ with $0 \leq \kappa \leq k_0$, where k_0 is the largest integer less than $n/2$. As for $k_0 < \kappa < n/2$, we have $0 < \kappa - 1 < k_0$ in the case $n \geq 3$. Hence, from the $L_{\kappa-1}^2$ -boundedness of $I_{\psi,\gamma}$, we obtain the L_κ^2 -boundedness.

In the cases $n = 1, 2$, we can construct a (C^1 -)diffeomorphism $\psi_e : \mathbb{R}^n \setminus 0 \rightarrow \mathbb{R}^n \setminus 0$ which is an extension of $\psi : \Gamma \rightarrow \tilde{\Gamma}$ satisfying $C^{-1} \leq |\det \partial \psi_e(\xi)| \leq C$ ($\xi \in \mathbb{R}^n \setminus 0$) for some $C > 0$. (In fact, it is trivial in the case $n = 1$. In the case $n = 2$, because of the homogeneity of $\psi(\xi)$, we have only to extend the function on the arc $\Gamma \cap \mathbb{S}^1$ to \mathbb{S}^1 keeping the diffeomorphism. It can be carried out by an elementary argument and we will omit the details.) Then, instead of (4.10), we have

$$I_{\psi,\gamma} = \gamma(D_x)I_{\psi_e}, \quad I_{\psi_e} = \frac{1}{\langle x \rangle^2} I_{\psi_e} + \frac{x}{\langle x \rangle^2} {}^t \partial \psi_e(D_x) I_{\psi_e} {}^t x.$$

From this formula, together with the L^2 -boundedness of I_{ψ_e} and that of all the entries of $\partial \psi_e(D_x)$, we obtain similarly the L_κ^2 -boundedness of I_{ψ_e} with $0 \leq \kappa \leq 1$. Since we have the L_κ^2 -boundedness of $\gamma(D_x)$ for $|\kappa| < n/2$ by Lemma 4.1, we can conclude that $I_{\psi,\gamma}$ is L_κ^2 -bounded with $0 \leq \kappa < n/2$. \square

5. SMOOTHING ESTIMATES FOR DISPERSIVE EQUATIONS

As an application of the canonical transformations described in Section 4, we can derive smoothing estimates for general dispersive equations from model estimates listed in Section 3. Note that the estimates that we will present are derived from just two simple estimates (3.3) and (3.8) in virtue of the comparison principle. The results which will be thus obtained in this section generalise many known results of the form (1.3) in the introduction. For the optimality of orders, see Section 6.

Let us consider the solution

$$u(t, x) = e^{ita(D_x)} \varphi(x)$$

to the equation

$$\begin{cases} (i\partial_t + a(D_x)) u(t, x) = 0 & \text{in } \mathbb{R}_t \times \mathbb{R}_x^n, \\ u(0, x) = \varphi(x) & \text{in } \mathbb{R}_x^n, \end{cases}$$

where we always assume that function $a(\xi)$ is real-valued. Let $a_m(\xi) \in C^\infty(\mathbb{R}^n \setminus 0)$, the *principal* part of $a(\xi)$, be a positively homogeneous function of order m , that is, satisfy $a_m(\lambda\xi) = \lambda^m a_m(\xi)$ for all $\lambda > 0$ and $\xi \neq 0$.

We sometimes decompose the initial data φ into the sum of the *low frequency* part φ_l and the *high frequency* part φ_h , where $\text{supp } \widehat{\varphi}_l \subset \{\xi : |\xi| < 2R\}$ and $\text{supp } \widehat{\varphi}_h \subset \{\xi : |\xi| > R\}$ with sufficiently large $R > 0$. Each part can be realised by multiplying $\chi(D_x)$ or $(1 - \chi)(D_x)$ to $\varphi(x)$, hence to $u(t, x)$, where $\chi \in C_0^\infty(\mathbb{R}^n)$ is an appropriate cut-off function.

First we consider the case that $a(\xi)$ has no lower order terms, and assume that $a(\xi)$ is *dispersive*:

$$(H) \quad a(\xi) = a_m(\xi), \quad \nabla a_m(\xi) \neq 0 \quad (\xi \in \mathbb{R}^n \setminus 0),$$

where $\nabla = (\partial_1, \dots, \partial_n)$ and $\partial_j = \partial_{\xi_j}$. A typical example is $a(\xi) = a_m(\xi) = |\xi|^m$. Especially, $a(\xi) = a_2(\xi) = |\xi|^2$ is the case of the Schrödinger equation.

The following result is derived from Corollary 3.3 and it is a generalisation of the result by Ben-Artzi and Klainerman [BK] which treated the case $a(\xi) = |\xi|^2$ and $n \geq 3$ (using spectral methods):

Theorem 5.1. *Assume (H). Suppose $n \geq 1$, $m > 0$, and $s > 1/2$. Then we have*

$$(5.1) \quad \left\| \langle x \rangle^{-s} |D_x|^{(m-1)/2} e^{ita(D_x)} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)}.$$

Chihara [Ch] proved Theorem 5.1 in the case $m > 1$, by proving the restriction theorem (1.4) or the resolvent estimates (1.5). We will, however, give a simpler proof by reducing estimate (5.1) for elliptic $a(\xi)$ to one dimensional model estimate (3.11) and non-elliptic $a(\xi)$ to two dimensional (3.12) in Corollary 3.3. Recall that these model estimates are a corollary of estimates (3.4) and (3.5) in Theorem 3.1, which is a direct consequence of just a trivial estimate (3.3). We also note that $m = 1$ is the case of the wave equation and is important for reducing the estimates to the model energy conservation case (3.3).

We also get a scaling invariant estimate for homogeneous weights $|x|^{-s}$ instead of non-homogenous ones $\langle x \rangle^{-s}$. The following result is derived from Corollary 3.4 and it is a generalisation of the result by Kato and Yajima [KY] which treated the case $a(\xi) = |\xi|^2$ with $n \geq 3$ and $0 \leq \alpha < 1/2$, or with $n = 2$ and $0 < \alpha < 1/2$. Ben-Artzi and Klainerman [BK] gave an alternative proof of the case $a(\xi) = |\xi|^2$ with $n \geq 3$ and $0 \leq \alpha < 1/2$, based on the estimate with a non-homogeneous weight and spectral decompositions. Our extension of these results is as follows:

Theorem 5.2. *Assume (H). Suppose $m > 0$ and $(m - n + 1)/2 < \alpha < (m - 1)/2$, or $m > 0$ and $(m - n)/2 < \alpha < (m - 1)/2$ in the elliptic case $a(\xi) \neq 0$ ($\xi \neq 0$). Then we have*

$$(5.2) \quad \left\| |x|^{\alpha-m/2} |D_x|^\alpha e^{ita(D_x)} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)}.$$

Sugimoto [Su1] proved Theorem 5.2 for elliptic $a(\xi)$ of order $m = 2$ and $1 - n/2 < \alpha < 1/2$, $n \geq 2$. We note that in general we can not allow $\alpha = (m - 1)/2$ in estimate (5.2), see Section 6. However, a sharp version of this estimate is still possible if one cut-off the main global singularity of the solution $u(t, x) = e^{ita(D_x)} \varphi(x)$. The location of this singularity is at the set of all classical trajectories corresponding to the operators $a(D_x)$. Such results and their sharpness have been discussed in authors' paper [RS3]. We note that this case has deep implications clarifying the null-form structure for derivative nonlinear Schrödinger equations and equations of similar type.

We have another type of smoothing estimate replacing $|D_x|^{(m-1)/2}$ by $\langle D_x \rangle^{(m-1)/2}$. The following result is a direct consequence of Theorems 5.1 and 5.2, and it also extends the result by Kato and Yajima [KY] which treated the case $a(\xi) = |\xi|^2$ and $n \geq 3$:

Corollary 5.3. *Assume (H). Suppose $n - 1 > m > 1$, or $n > m > 1$ in the elliptic case $a(\xi) \neq 0$ ($\xi \neq 0$). Then we have*

$$(5.3) \quad \left\| \langle x \rangle^{-m/2} \langle D_x \rangle^{(m-1)/2} e^{ita(D_x)} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)}.$$

Proof of Corollary 5.3. Theorem 5.1 implies the stronger estimate for the high frequency part of estimate (5.3) replacing the weight $\langle x \rangle^{-m/2}$ by $\langle x \rangle^{-s}$ with $s > 1/2$. Theorem 5.2 with $\alpha = 0$ also implies the stronger estimate for the low frequency part replacing the weight $\langle x \rangle^{-m/2}$ by $|x|^{-m/2}$. \square

We remark that Walther [Wa2] used spherical harmonics and asymptotics of Bessel functions to prove the result of Corollary 5.3 directly in the radially symmetric case of $a(\xi) = |\xi|^m$ (this satisfies assumption (H) and the ellipticity). In the elliptic case with $m = 2$, Walther's result was extended to the non-radially symmetric case by the authors [RS2]. Corollary 5.3 is the development of that analysis allowing non-elliptic operators as well. We may also look at the other type of global smoothing of the form (5.3), but with the weight $\langle x \rangle^{-m/2}$ replaced by homogeneous ones. However, this follows from the previous types. For example, we can observe that estimate (5.2) trivially implies

$$\left\| |x|^{\alpha-m/2} \langle D_x \rangle^\alpha e^{ita(D_x)} \varphi_h(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi_h\|_{L^2(\mathbb{R}_x^n)},$$

for high frequency parts, while for low frequency part we get

$$\left\| |x|^{-m/2} e^{ita(D_x)} \varphi_l(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi_l\|_{L^2(\mathbb{R}_x^n)}$$

as a special case of (5.2) with $\alpha = 0$.

The main idea to prove Theorems 5.1 and 5.2 is to reduce them to Corollaries 3.3 and 3.4 by using Theorem 4.1. If some estimate for $e^{it\sigma(D_x)}$ is listed there, then all our task is to find $\psi(\xi)$ such that $a(\xi) = (\sigma \circ \psi)(\xi)$ and verify all the boundedness assumptions we need. We will use the notation $\xi = (\xi_1, \dots, \xi_n)$, $\eta = (\eta_1, \dots, \eta_n)$, and $D_x = (D_1, \dots, D_n)$ as used there.

We assume (H). Let $\Gamma \subset \mathbb{R}^n \setminus 0$ be a sufficiently small conic neighbourhood of $e_n = (0, \dots, 0, 1)$, and take a cut-off function $\gamma(\xi) \in C^\infty(\Gamma)$ which is positively homogeneous of order 0 and satisfies $\text{supp } \gamma \cap \mathbb{S}^{n-1} \subset \Gamma \cap \mathbb{S}^{n-1}$. By the microlocalisation and the rotation of the initial data φ , we may assume $\text{supp } \widehat{\varphi} \subset \text{supp } \gamma$. The dispersive assumption $\nabla a_m(e_n) \neq 0$ in this direction implies the following two possibilities:

- (i): $\partial_n a_m(e_n) \neq 0$. Then, by Euler's identity $a_m(\xi) = (1/m) \nabla a_m(\xi) \cdot \xi$, we have $a_m(e_n) \neq 0$. Hence, in this case, we may assume that $a(\xi) (> 0)$ and $\partial_n a(\xi)$ are bounded away from 0 for $\xi \in \Gamma$.
- (ii): $\partial_n a_m(e_n) = 0$. Then there exists $j \neq n$ such that $\partial_j a_m(e_n) \neq 0$, say $\partial_1 a_m(e_n) \neq 0$. Hence, in this case, we may assume $\partial_1 a(\xi)$ is bounded away from 0 for $\xi \in \Gamma$. We remark $a(e_n) = 0$ by Euler's identity.

Proof of Theorem 5.1. The estimate with the case $n = 1$ is given by estimate (3.11) in Corollary 3.3. In fact, we have $a(\xi) = a(1)|\xi|^m$ for $\xi > 0$ in this case. Hence we may assume $n \geq 2$. We remark that it is sufficient to show theorem with $1/2 < s < n/2$ because the case $s \geq n/2$ is easily reduced to this case.

In the case (i), we take

$$(5.4) \quad \sigma(\eta) = |\eta_n|^m, \quad \psi(\xi) = (\xi_1, \dots, \xi_{n-1}, a(\xi)^{1/m}).$$

Then we have $a(\xi) = (\sigma \circ \psi)(\xi)$ and

$$(5.5) \quad \det \partial\psi(\xi) = \begin{vmatrix} E_{n-1} & 0 \\ * & (1/m)a(\xi)^{1/m-1}\partial_n a(\xi) \end{vmatrix},$$

where E_{n-1} is the identity matrix of order $n - 1$. We remark that (4.1) is satisfied since $\det \partial\psi(e_n) = (1/m)a(e_n)^{1/m-1}\partial_n a(e_n) \neq 0$. By estimate (3.11) in Corollary 3.3, we have estimate (4.7) in Theorem 4.1 with $\sigma(D_x) = |D_n|^m$, $w(x) = \langle x \rangle^{-s}$, and $\rho(\xi) = |\xi_n|^{(m-1)/2}$. Note here the trivial inequality $\langle x \rangle^{-s} \leq \langle x_n \rangle^{-s}$. If we take $\zeta(\xi) = |\xi|^{(m-1)/2}$, then $q(\xi) = \gamma(\xi)(|\xi|/a(\xi)^{1/m})^{(m-1)/2}$ defined by (4.8) is a bounded function. On the other hand, $I_{\psi, \gamma}$ is L^2_{-s} -bounded for $1/2 < s < n/2$ by Theorem 4.3. Hence, by Theorem 4.1, we have estimate (4.9), that is, estimate (5.1).

In the case (ii), we take

$$\sigma(\eta) = \eta_1 |\eta_n|^{m-1}, \quad \psi(\xi) = (a(\xi)|\xi_n|^{1-m}, \xi_2, \dots, \xi_n)$$

Then we have $a(\xi) = (\sigma \circ \psi)(\xi)$ and

$$\det \partial\psi(\xi) = \begin{vmatrix} \partial_1 a(\xi)|\xi_n|^{1-m} & * \\ 0 & E_{n-1} \end{vmatrix}.$$

Since $\det \partial\psi(e_n) = \partial_1 a(e_n) \neq 0$, (4.1) is satisfied. Similarly to the case (i), the estimate for $\sigma(D_x) = D_1 |D_n|^{m-1}$ is given by estimate (3.12) in Corollary 3.3, which implies estimate (5.1) again by Theorem 4.1. \square

Proof of Theorem 5.2. In the case (i), which is the only possibility for the elliptic $a(\xi) \neq 0$ ($\xi \neq 0$), we take

$$\sigma(\eta) = |\eta|^m, \quad \psi(\xi) = \left(\xi_1, \dots, \xi_{n-1}, \sqrt{a(\xi)^{2/m} - (\xi_1^2 + \dots + \xi_{n-1}^2)} \right).$$

Then we have $a(\xi) = (\sigma \circ \psi)(\xi)$ and

$$\det \partial\psi(\xi) = \begin{vmatrix} E_{n-1} & 0 \\ * & (1/m)a(\xi)^{2/m-1}\partial_n a(\xi) / \sqrt{a(\xi)^{2/m} - (\xi_1^2 + \dots + \xi_{n-1}^2)} \end{vmatrix}.$$

Since $\det \partial\psi(e_n) = (1/m)a(e_n)^{1/m-1}\partial_n a(e_n) \neq 0$, (4.1) is satisfied. The estimate for $\sigma(D_x) = |D|^m$ is given by estimate (3.13) in Corollary 3.4. In the case (ii), we take

$$\sigma(\eta) = |\eta_1|^m - (\eta_2^2 + \dots + \eta_n^2)^{m/2}, \quad \psi(\xi) = \left((a(\xi) + (\xi_2^2 + \dots + \xi_n^2)^{m/2})^{1/m}, \xi_2, \dots, \xi_n \right)$$

Then we have $a(\xi) = (\sigma \circ \psi)(\xi)$ and

$$\det \partial\psi(\xi) = \begin{vmatrix} (1/m)(a(\xi) + (\xi_2^2 + \dots + \xi_n^2)^{m/2})^{1/m-1}\partial_1 a(\xi) & * \\ 0 & E_{n-1} \end{vmatrix}.$$

Since $\det \partial\psi(e_n) = (1/m)\partial_1 a(e_n) \neq 0$, (4.1) is satisfied. The estimate for $\sigma(D_x) = |D_1|^m - (D_2^2 + \dots + D_n^2)^{m/2}$ is given by estimate (3.14) in Corollary 3.4. By the same argument as used in the proof of Theorem 5.1, we have Theorems 5.2. \square

As another advantage of the new method, we can also consider the case that $a(\xi)$ has lower order terms, and assume that $a(\xi)$ is dispersive in the following sense:

$$(L) \quad \begin{aligned} & a(\xi) \in C^\infty(\mathbb{R}^n), \quad \nabla a(\xi) \neq 0 \quad (\xi \in \mathbb{R}^n), \quad \nabla a_m(\xi) \neq 0 \quad (\xi \in \mathbb{R}^n \setminus 0), \\ & |\partial^\alpha(a(\xi) - a_m(\xi))| \leq C_\alpha |\xi|^{m-1-|\alpha|} \quad \text{for all multi-indices } \alpha \text{ and all } |\xi| \geq 1. \end{aligned}$$

We note that $a(\xi) = |\xi|^m$ does not satisfy (L) because $\nabla a(\xi)$ vanishes at the origin $\xi = 0$, while it satisfies (H). On the other hand, $a(\xi) = a_3(\xi) + \xi_1$ satisfies (L) with $m = 3$, where $a_3(\xi) = \xi_1^3 + \xi_2^3 + \dots + \xi_n^3$ and $\xi = (\xi_1, \xi_2, \dots, \xi_n)$. As will be discussed soon, the ability to include the lower order terms and conditions on them is very important in global problems. In fact, it is known that low frequencies are often responsible for the orders of decay of the solutions and their smoothing property for large times. However, the difference between the principal part and the lower order terms becomes extinct in the low frequency part, and one has to look at the properties of the full symbol. Thus, if we want to have the dispersive behaviour of the problem we need to look at the dispersiveness of the full symbol in assumption (L). For large ξ conditions $\nabla a(\xi) \neq 0$ and $\nabla a_m(\xi) \neq 0$ are clearly equivalent, while for small ξ condition $\nabla a_m(\xi) \neq 0$ is not necessary (but it is satisfied anyway due to the homogeneity of a_m). Thus, condition (L) may be formulated also in the following way

$$(L) \quad \begin{aligned} & a(\xi) \in C^\infty(\mathbb{R}^n), \quad |\nabla a(\xi)| \geq C \langle \xi \rangle^{m-1} \quad (\xi \in \mathbb{R}^n) \quad \text{for some } C > 0, \\ & |\partial^\alpha(a(\xi) - a_m(\xi))| \leq C_\alpha |\xi|^{m-1-|\alpha|} \quad \text{for all multi-indices } \alpha \text{ and all } |\xi| \geq 1. \end{aligned}$$

The last line of this assumption simply amounts to saying that the principal part a_m of a is positively homogeneous of order m for $|\xi| \geq 1$.

The following result is also derived from Corollary 3.3:

Theorem 5.4. *Assume (L). Suppose $n \geq 1$, $m > 0$, and $s > 1/2$. Then we have*

$$(5.6) \quad \left\| \langle x \rangle^{-s} \langle D_x \rangle^{(m-1)/2} e^{ita(D_x)} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)}.$$

Thus, by Theorem 5.4, we can have better estimate than that in Corollary 5.3 even under weaker conditions on m and n if we assume (L) instead of (H). This fact does not contradict to the optimality of Corollary 5.3 with the case $a(\xi) = |\xi|^m$ (see the remark below Corollary 5.3) because it does not satisfy assumption (L). This does emphasise once again the importance of the dispersiveness assumption $\nabla a \neq 0$.

Note that the following result is a straightforward consequence of Theorem 5.4 and the L^2 -boundedness of $|D_x|^{(m-1)/2} \langle D_x \rangle^{-(m-1)/2}$ with $m \geq 1$, which is an analog of Theorem 5.1 for $a(D_x)$ with lower order terms (assumption $m \geq 1$ is natural to be able to talk about lower order terms):

Corollary 5.5. *Assume (L). Suppose $n \geq 1$, $m \geq 1$ and $s > 1/2$. Then we have*

$$(5.7) \quad \left\| \langle x \rangle^{-s} |D_x|^{(m-1)/2} e^{ita(D_x)} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)}.$$

Proof of Theorem 5.4. We decompose the initial data φ into the sum of the high frequency part and the low frequency part. For high frequency part, the same argument as in the proof of Theorem 5.1 is valid. (Furthermore, we can use Theorem 4.2 instead of Theorem 4.3 to assure the boundedness of $I_{\psi, \gamma}$, hence we need not assume $n \geq 2$.) We show how to get the estimates for low frequency part. Because of the compactness of it, we may assume $\partial_j a(\xi) \neq 0$ with some j , say $j = n$, on a bounded set $\Gamma \subset \mathbb{R}^n$ and $\text{supp } \widehat{\varphi} \subset \Gamma$. Since we have $a(\xi) + c > 0$ on Γ with some constant $c > 0$ and

$$\left\| \langle x \rangle^{-s} \langle D_x \rangle^{(m-1)/2} e^{ita(D_x)} \varphi \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} = \left\| \langle x \rangle^{-s} \langle D_x \rangle^{(m-1)/2} e^{it(a(D_x)+2c)} \varphi \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)},$$

we may assume $a(\xi) \geq c > 0$ on Γ without loss of generality. We take a cut-off function $\gamma(\xi) \in C_0^\infty(\Gamma)$, and choose $\psi(\xi)$ and $\sigma(\eta)$ in the same way as (5.4). Assumption (4.1) is also verified if we notice (5.5). By estimate (3.11) in Corollary 3.3, we have estimate (4.7) in Theorem 4.1 with $\sigma(D_x) = |D_n|^m$, $w(x) = \langle x \rangle^{-s}$ ($s > 1/2$), and $\rho(\xi) = |\xi_n|^{(m-1)/2}$ as in the proof of Theorem 5.1. If we take $\zeta(\xi) = \langle \xi \rangle^{(m-1)/2}$, then $q(\xi) = \gamma(\xi) (\langle \xi \rangle / a(\xi)^{1/m})^{(m-1)/2}$ defined by (4.8) is a bounded function. On the other hand, $I_{\psi, \gamma}$ is L^2_{-s} -bounded for all $s > 1/2$ by Theorem 4.2. Hence, by Theorem 4.1, we have estimate (4.9), that is, estimate (5.6). \square

Recall that assumption (L) in Theorem 5.4 requires the condition $\nabla a(\xi) \neq 0$ ($\xi \in \mathbb{R}^n$) for the full symbol, besides the same one $\nabla a_m(\xi) \neq 0$ ($\xi \neq 0$) for the principal term. We will now introduce an intermediate assumption between (H) and (L), and discuss what happens if we do not have the condition $\nabla a(\xi) \neq 0$:

$$(HL) \quad \begin{aligned} a(\xi) &= a_m(\xi) + r(\xi), \quad \nabla a_m(\xi) \neq 0 \quad (\xi \in \mathbb{R}^n \setminus 0), \quad r(\xi) \in C^\infty(\mathbb{R}^n) \\ |\partial^\alpha r(\xi)| &\leq C \langle \xi \rangle^{m-1-|\alpha|} \quad \text{for all multi-indices } \alpha. \end{aligned}$$

In view of the proof of Theorem 5.4, we see that Theorems 5.1, 5.2, and Corollary 5.3 remain valid if we replace assumption (H) by (HL) and functions $\varphi(x)$ in the estimates by its (sufficiently large) high frequency part $\varphi_h(x)$. However we cannot control the low frequency part $\varphi_l(x)$, and so have only the time local estimates on the whole:

Theorem 5.6. *Assume (HL). Suppose $n \geq 1$, $m > 0$, $s > 1/2$, and $T > 0$. Then we have*

$$\int_0^T \left\| \langle x \rangle^{-s} \langle D_x \rangle^{(m-1)/2} e^{ia(D_x)} \right\|_{L^2(\mathbb{R}_x^n)}^2 dt \leq C \|\varphi\|_{L^2(\mathbb{R}^n)}^2,$$

where $C > 0$ is a constant depending on $T > 0$.

Proof of Theorem 5.6. We decompose φ into the sum of low and high frequency parts. For the high frequency part, the same arguments as in the proof of Theorems 5.1 and 5.4 are valid (and furthermore we can have the estimate with $T = \infty$). The estimate for the low frequency part is trivial. In fact, if $\text{supp } \mathcal{F}\varphi \subset \{\xi : |\xi| \leq R\}$, we have

$$\begin{aligned} \int_0^T \left\| \langle x \rangle^{-s} \langle D_x \rangle^{(m-1)/2} e^{ita(D_x)} \varphi(x) \right\|_{L^2(\mathbb{R}_x^n)}^2 dt &\leq \int_0^T \left\| \langle D_x \rangle^{(m-1)/2} e^{ita(D_x)} \varphi(x) \right\|_{L^2(\mathbb{R}_x^n)}^2 dt \\ &\leq CT \left\| \langle \xi \rangle^{(m-1)/2} \widehat{\varphi}(\xi) \right\|_{L^2(\mathbb{R}^n)}^2 \\ &\leq CT \langle R \rangle^{m-1} \|\varphi\|_{L^2(\mathbb{R}^n)}^2 \end{aligned}$$

by Plancherel's theorem. \square

We remark that Theorem 5.4 is the time global version (that is, the estimate with $T = \infty$) of Theorem 5.6, and the extra assumption $\nabla a(\xi) \neq 0$ is needed for that. Since the assumption $\nabla a(\xi) \neq 0$ for large ξ is automatically satisfied by assumption (HL), Theorem 5.4 means that the condition $\nabla a(\xi) \neq 0$ for small ξ assures the time global estimate. In this sense, the low frequency part have a responsibility for the time global smoothing.

Finally, we will state a secondary comparison result, similarly to Corollaries 3.5 and 3.6 in Section 3, especially for radially symmetric operators, which will play various important roles in later sections (Sections 7 and 8). In notation of Theorem 2.5, setting $\tau(\rho) = \rho^{(m-1)/2}$ and $g(\rho) = \rho^m$, we have $|\tau(\rho)|/|g'(\rho)|^{1/2} = m^{-1/2}$. If we take $\tau(\rho) = \rho^\alpha$ and $g(\rho) = \rho^2$ instead, we have $|\tau(\rho)|/|g'(\rho)|^{1/2} = 2^{-1/2} \rho^{\alpha-1/2}$. Then we obtain the following results from Theorem 5.1 with $a(\xi) = |\xi|^m$ and Theorem 5.2 with $a(\xi) = |\xi|^2$, that is, estimate (3.6) in Section 3:

Corollary 5.7. *Suppose $n \geq 1$, $s > 1/2$, and $1 - n/2 < \alpha < 1/2$. Let $\chi \in L^\infty(\mathbb{R}_+)$. Let $f \in C^1(\mathbb{R}_+)$ be real-valued and strictly monotone on $\text{supp } \chi$. Let $\sigma \in C^0(\mathbb{R}_+)$ be such that for some $A > 0$ we have*

$$(5.8) \quad |\sigma(\rho)| \leq A|f'(\rho)|^{1/2}$$

for all $\rho \in \text{supp } \chi$. Then we have

$$(5.9) \quad \left\| \langle x \rangle^{-s} \chi(|D_x|) \sigma(|D_x|) e^{itf(|D_x|)} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)},$$

$$(5.10) \quad \left\| |x|^{\alpha-1} \chi(|D_x|) |D_x|^{\alpha-1/2} \sigma(|D_x|) e^{itf(|D_x|)} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)}.$$

6. INVARIANT ESTIMATES AND SHARPNESS

Let us now suggest an invariant form of smoothing estimates for the solution $u(t, x) = e^{ita(D_x)} \varphi(x)$ to the equation

$$(6.1) \quad \begin{cases} (i\partial_t + a(D_x)) u(t, x) = 0 & \text{in } \mathbb{R}_t \times \mathbb{R}_x^n, \\ u(0, x) = \varphi(x) & \text{in } \mathbb{R}_x^n, \end{cases}$$

which remain valid also in some areas without dispersion $\nabla a(\xi) \neq 0$, where standard smoothing estimates are known to fail. We will discuss these estimates in this section and then will establish them in a variety of situations in Section 7.

We can equivalently rewrite estimates in Section 5 in the form

$$(6.2) \quad \left\| w(x) \zeta(|\nabla a(D_x)|) e^{ita(D_x)} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)},$$

where w is a weight function of the form $w(x) = |x|^\delta, \langle x \rangle^\delta$ and ζ is a function on \mathbb{R}_+ of the form $\zeta(\rho) = \rho^\eta, (1 + \rho^2)^{\eta/2}$ with some $\delta, \eta \in \mathbb{R}$. For example, we can rewrite estimate (5.1) of Theorem 5.1 as well as estimate (5.7) of Corollary 5.5 for the dispersive equations in the form

$$(6.3) \quad \left\| \langle x \rangle^{-s} |\nabla a(D_x)|^{1/2} e^{ita(D_x)} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)}.$$

Similarly we can rewrite estimate (5.2) of Theorem 5.2 in the form

$$(6.4) \quad \left\| |x|^{\alpha-m/2} |\nabla a(D_x)|^{\alpha/(m-1)} e^{ita(D_x)} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)} \quad (m \neq 1),$$

and estimate (5.3) of Corollary 5.3 ($s = -m/2$) as well as estimate (5.6) of Theorem 5.4 in the form

$$(6.5) \quad \left\| \langle x \rangle^{-s} \langle \nabla a(D_x) \rangle^{1/2} e^{ita(D_x)} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)}.$$

Indeed, under assumption (H) we clearly have $|\nabla a(\xi)| \geq c|\xi|^{m-1}$, so the equivalence between estimate (6.3) and estimate (5.1) in Theorem 5.1 follows from the fact that $|\nabla a(D_x)|^{1/2} |D_x|^{-(m-1)/2}$ and $|\nabla a(D_x)|^{-1/2} |D_x|^{(m-1)/2}$ are bounded in $L^2(\mathbb{R}^n)$. Under assumption (L) the same argument works for large frequencies, while for small frequencies both $\langle \xi \rangle^{(m-1)/2}$ and $|\nabla a(\xi)|^{1/2}$ are bounded away from zero. Thus we have the equivalence between estimate (6.3) and estimate (5.7) in Corollary 5.5. The same is true for the other equivalences. We may also look at the other type of global smoothing of the form (6.5), but with homogeneous weight functions. However, this follows from the type (6.4) as was also explained in the remark below Corollary 5.3.

Estimate (6.2), hence estimates (6.3) – (6.5) are invariant under canonical transformations by Theorem 4.1. More precisely, we have the following theorem:

Theorem 6.1. *Let ζ be a function on \mathbb{R}_+ of the form $\zeta(\rho) = \rho^\eta$ or $(1 + \rho^2)^{\eta/2}$ with some $\eta \in \mathbb{R}$. Assume that the operators $I_{\psi, \gamma}$ and $I_{\psi, \gamma}^{-1}$ defined by (4.3) are $L^2(\mathbb{R}^n; w)$ -bounded. Then the following two estimates*

$$\begin{aligned} \left\| w(x) \zeta(|\nabla a(D_x)|) e^{ita(D_x)} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} &\leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)} \quad (\text{supp } \widehat{\varphi} \subset \text{supp } \gamma), \\ \left\| w(x) \zeta(|\nabla \sigma(D_x)|) e^{it\sigma(D_x)} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} &\leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)} \quad (\text{supp } \widehat{\varphi} \subset \text{supp } \widetilde{\gamma}) \end{aligned}$$

are equivalent to each other, where $a = \sigma \circ \psi \in C^1$ on $\text{supp } \gamma$ and $\widetilde{\gamma} = \gamma \circ \psi^{-1}$.

Proof. Note that $\nabla a(\xi) = \nabla \sigma(\psi(\xi)) D\psi(\xi)$ and $C|\nabla a(\xi)| \leq |\nabla \sigma(\psi(\xi))| \leq C'|\nabla a(\xi)|$ on $\text{supp } \gamma$ with some $C, C' > 0$, which is assured by the assumption (4.1). Then the result is obtained from Theorem 4.1. \square

On account of these argument, we will call estimate (6.2) an *invariant estimate*, and indeed we expect invariant estimates (6.3), (6.4), and (6.5) to hold for $s > 1/2$, $(m - n)/2 < \alpha < (m - 1)/2$, and $s = -m/2$ ($n > m > 1$), respectively in ordinally settings (elliptic case for example), with $m > 0$ being the order of $a(D_x)$.

Let us briefly indicate that invariant estimate (6.3) with $s > 1/2$ is also a refinement of another known estimate for non-dispersive equations. If operator $a(D_x)$ has real-valued symbol $a(\xi) \in C^1(\mathbb{R}^n)$ which is positively homogeneous of order $m \geq 1$ and no dispersiveness assumption is made, Hoshiro [Ho1] showed the estimate

$$(6.6) \quad \left\| \langle x \rangle^{-s} \langle D_x \rangle^{-s} |a(D)|^{1/2} e^{ita(D_x)} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)} \quad (s > 1/2).$$

But once we prove (6.3) with $s > 1/2$, we can have better estimate

$$\left\| \langle x \rangle^{-s} \langle D_x \rangle^{-1/2} |a(D)|^{1/2} e^{ita(D_x)} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)} \quad (s > 1/2)$$

with respect to the number of derivatives. In fact, using the Euler's identity $ma(\xi) = \xi \cdot \nabla a(\xi)$, we see that this estimate trivially follows from

$$\left\| \langle x \rangle^{-s} \langle D_x \rangle^{-1/2} |D_x|^{1/2} |\nabla a(D)|^{1/2} e^{ita(D_x)} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)} \quad (s > 1/2),$$

which in turn follows from (6.3) with $s > 1/2$ because $|D_x|^{1/2} \langle D_x \rangle^{-1/2}$ is $L^2(\mathbb{R}^n)$ -bounded. In fact, estimate (6.6) holds only because of the homogeneity of a , since in this case by Euler's identity zeros of a contain zeros of ∇a . In general, estimate (6.6) cuts off too much, and therefore does not reflect the nature of the problem for non-homogeneous symbols, as (6.3) still does.

In terms of invariant estimates, we can also give another explanation to the reason why we do not have time global estimate in Theorem 5.6. The problem is that the symbol of the smoothing operator $\langle D_x \rangle^{(m-1)/2}$ does not vanish where the symbol of $\nabla a(D_x)$ vanishes, as should be anticipated by the invariant estimate (6.3). If zeros of $\nabla a(D_x)$ are not taken into account, the weight should change to the one as in estimate (5.3).

Sharpness of smoothing estimates. Let us now discuss the scaling invariance and sharpness properties of estimates (6.3)–(6.5) taking liberty of also referring to results that will be established in the sequel. Let us restrict to the case when $a(\xi) \in C^\infty(\mathbb{R}^n \setminus 0)$ is elliptic and positively homogeneous of order $m > 0$. Then it is easy to see that estimate (6.4) is scaling invariant with respect to the natural scaling $u_\lambda(t, x) = u(\lambda^m t, \lambda x)$ to the solution of equation (6.1). If $a(\xi)$ is dispersive, that is if $\nabla a(\xi) \neq 0$ ($\xi \neq 0$), estimate (6.4) holds for $(m - n)/2 < \alpha < (m - 1)/2$ by Theorem 5.2. Also, the validity of this estimate for some value of α implies the validity of the estimate for smaller α 's (see the proof of this given just before Lemma 3.1). Thus, the critical case of this estimate is for the largest value $\alpha = (m - 1)/2$. In the case of the Schrödinger equation ($m = 2$) this is the critical case of Kato–Yajima's estimate and it was shown to fail in the critical case $\alpha = 1/2$ by Watanabe [W] (although quite implicitly).

We will now give a more direct explicit argument for the failure of this and other critical estimates. We note that in the critical case $\alpha = (m - 1)/2$ estimate (6.4)

(which we will show to fail) becomes

$$(6.7) \quad \left\| |x|^{-1/2} |\nabla a(D_x)|^{1/2} e^{ita(D_x)} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)}.$$

Such an estimate would be very useful for the well-posedness analysis of derivative nonlinear equations or equations with magnetic potentials since the recovery of the loss of regularity would be sharp, so one wants to repair it. One way is to locate and then cut-off the main singularity. This was done by the authors in [RS3] and is briefly discussed in Section 11. The other way is to first observe that this estimate is equivalent to a weaker estimate

$$(6.8) \quad \left\| \langle x \rangle^{-1/2} |\nabla a(D_x)|^{1/2} e^{ita(D_x)} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)}.$$

Indeed, (6.8) follows from (6.7) by the trivial inequality $\langle x \rangle^{-1/2} \leq |x|^{-1/2}$, while (6.7) follows from (6.8) by the scaling argument (similar to the one just before Theorem 3.2). Now, for dispersive $a(\xi)$ by using the canonical transform method of Section 5, estimate (6.8) is equivalent to its normal form. For example, in the case of elliptic $a(D_x)$, it is equivalent to the one dimensional estimate

$$(6.9) \quad \left\| \langle x \rangle^{-1/2} |D_x|^{(m-1)/2} e^{it|D_x|^m} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x)}.$$

Now, by the comparison principle of Section 2, it is equivalent to its special case with $m = 1$, which is estimate

$$(6.10) \quad \left\| \langle x \rangle^{-1/2} e^{it|D_x|} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x)}.$$

If $\text{supp } \widehat{\varphi} \subset [0, \infty)$, we have $e^{it|D_x|} \varphi(x) = \varphi(x + t)$, and so, finally, (6.10) is equivalent to

$$(6.11) \quad \left\| \langle x \rangle^{-1/2} \varphi(x + t) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x)}.$$

The last estimate clearly fails since $\langle x \rangle^{-1/2}$ is not in $L^2(\mathbb{R}_x^1)$, thus implying that all the estimates (6.7)–(6.11) fail. Note that we may talk about equivalence of (false) estimates here since both the canonical transform method and the comparison principle apply to expressions on the left hand side of these estimates and these arguments are of equivalence, showing that estimates hold or fail simultaneously.

Now, we can try to repair (6.7), or rather (6.8), by taking a stronger weight $\langle x \rangle^{-s}$ for $s > 1/2$. In this way we arrive at the “almost” scaling invariant estimate

$$(6.12) \quad \left\| \langle x \rangle^{-s} |\nabla a(D_x)|^{1/2} e^{ita(D_x)} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)} \quad (s > 1/2),$$

which is invariant estimate (6.3). We remark that, for high frequencies, this estimate implies another type of invariant estimate (6.5) in some cases, for example under assumption (H) of $a(\xi)$ being positively homogeneous of order $m \geq 1$.

Let us discuss the third invariant estimate (6.5) with $s = -m/2$ ($n > m > 1$). For large frequencies it is weaker than (6.12), so we may restrict ourselves to bounded frequencies, in which case (6.5) is equivalent to the estimate

$$\left\| \langle x \rangle^{-m/2} e^{ita(D_x)} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)}.$$

By Theorem 3.2 and especially the scaling argument preceding it, we can conclude that this is in turn equivalent to the estimate

$$\| |x|^{-m/2} e^{ita(D_x)} \varphi(x) \|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)}.$$

But this estimate is scaling invariant (it is a special case of (6.4) with $\alpha = 0$), which justifies the sharpness of the order $-m/2$ of the weight. Thus, the expected orders of the weights in invariant estimates (6.3)–(6.5) are sharp.

A similar argument can be used to justify the optimality of the smoothing operator $|\nabla a(D_x)|^{1/2}$ in estimate (6.3). For example, in the case of elliptic $a(D_x)$, the weighted estimate (6.3) for $|\nabla a(D_x)|^{1/2} e^{ita(D_x)} \varphi(x)$ will be reduced (by the canonical transform method) to the weighted estimate for the model case $|D_x|^{(m-1)/2} e^{it|D_x|^m} \varphi(x)$. This, in turn, by the comparison principle, can be reduced to the pointwise estimate for its special case $m = 1$, that is, to the L^2 -estimate for $e^{it|D_x|} \varphi(x) = \varphi(x + t)$, with $\text{supp } \widehat{\varphi} \subset [0, \infty)$. Since there is no smoothing of a travelling wave, operator $|\nabla a(D_x)|^{1/2}$ in (6.3) is sharp. Similar arguments apply to non-elliptic dispersive $a(\xi)$ by reducing to models in two dimension, and to non-homogeneous symbols $a(\xi)$ by using assumption (L) in Section 5.

7. SMOOTHING ESTIMATES FOR NON-DISPERSIVE EQUATIONS

The most important application of the secondary comparison results Corollaries 3.5, 3.6, and 5.7 which were stated in Sections 3 and 5 is to the smoothing estimates for the equations

$$\begin{cases} (i\partial_t + a(D_x)) u(t, x) = 0 & \text{in } \mathbb{R}_t \times \mathbb{R}_x^n, \\ u(0, x) = \varphi(x) & \text{in } \mathbb{R}_x^n, \end{cases}$$

where real-valued function $a(\xi)$ fails to satisfy dispersive assumption (H) or (L) in Section 5. In Corollary 5.7 for example, even if we lose the dispersiveness assumption at zeros of f' , the estimate is still valid because σ must vanish at the same points with the order determined by condition (5.8). The same is true in other comparison results Corollaries 3.5 and 3.6. In this section, we will treat the smoothing estimates of non-dispersive equations based on this observation.

The following result states that we still have estimate (6.3) of invariant form suggested in Section 6 for *non-dispersive equations* in a general setting of the radially symmetric case:

Theorem 7.1. *Suppose $n \geq 1$ and $s > 1/2$. Let $a(\xi) = f(|\xi|)$, where $f \in C^1(\mathbb{R}_+)$ is real-valued. Assume that f' has only finitely many zeros. Then we have*

$$\| \langle x \rangle^{-s} |\nabla a(D_x)|^{1/2} e^{ita(D_x)} \varphi(x) \|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)}.$$

Proof. Noticing $|\nabla a(\xi)| = |f'(|\xi|)|$, use Corollary 5.7 for $\sigma(\rho) = |f'(\rho)|^{1/2}$ in each interval where f is strictly monotone. \square

Example 7.1. As a consequence of Theorem 7.1, we have the estimate of invariant form (6.3) if $a(\xi)$ is a real polynomial of $|\xi|$. For example, let $a(\xi) = f(|\xi|^2)^2$, with $f(\rho)$ being a non-constant polynomial on \mathbb{R} , The principal part $a_m(\xi)$ of $a(\xi)$ is a power of $|\xi|^2$ multiplied by a constant, hence it satisfies $\nabla a_m(\xi) \neq 0$ ($\xi \neq 0$). If $f(\rho)$

is a homogeneous polynomial, then $a(\xi)$ satisfies assumption (H) in Section 5 and we have estimate (6.3) by Theorem 5.1. In the case when $f(\rho)$ is not homogeneous, trivially $a(\xi)$ does not satisfy (H). Furthermore $a(\xi)$ does not satisfy assumption (L) in Section 5 either since $\nabla a(\xi) = 4f(|\xi|^2)f'(|\xi|^2)\xi$ vanishes on the set $|\xi|^2 = c$ such that $f(c) = 0$ or $f'(c) = 0$ as well as at the origin $\xi = 0$. Hence Corollary 5.5 does not assure the estimate (6.3), but even in this case, we have it by Theorem 7.1.

If we use Corollaries 3.5 and 3.6, we can obtain estimate (6.3) for non-dispersive equations in the non-radially symmetric case, as well. We will not try to exhaust the general case, but give some typical examples of the case when assumption (L) in Section 5 breaks. Below, we use the notation $x = (x_1, \dots, x_n)$, $\xi = (\xi_1, \dots, \xi_n)$, and $D_x = (D_1, \dots, D_n)$ as in Section 2:

Example 7.2. Let $a(\xi) = \xi_1^2 + \xi_2^2 + \xi_1$ in \mathbb{R}^2 , so that $a(\xi)$ fails to satisfy $\nabla a(\xi) \neq 0$ at $(\xi_1, \xi_2) = (-1/2, 0)$. By taking χ in Corollaries 3.5 and 3.6 to be characteristic functions of appropriate sets and $s > 1/2$, we obtain estimates

$$\begin{aligned} \left\| \langle x_2 \rangle^{-s} |D_2|^{1/2} e^{itD_2^2} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^2)} &\leq C \|\varphi\|_{L^2(\mathbb{R}_x^2)}, \\ \left\| \langle x_1 \rangle^{-s} |2D_1 + 1|^{1/2} e^{it(D_1^2 + D_1)} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^2)} &\leq C \|\varphi\|_{L^2(\mathbb{R}_x^2)}, \end{aligned}$$

which imply

$$\left\| \langle x \rangle^{-s} (|2D_1 + 1|^{1/2} + |D_2|^{1/2}) e^{ita(D_x)} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^2)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x^2)}$$

by trivial inequalities $\langle x \rangle^{-s} \leq \langle x_k \rangle^{-s}$ ($k = 1, 2$) and Plancherel's theorem for partial Fourier transforms. Substituting $\eta(D_x)\varphi$ for φ in the estimate, where

$$\eta(\xi) = |\nabla a(\xi)|^{1/2} (|2\xi_1 + 1|^{1/2} + |\xi_2|^{1/2})^{-1},$$

we have estimate (6.3) if we note the boundedness of $\eta(\xi)$ and use Plancherel's theorem again.

Example 7.3. Let $a(\xi) = \xi_1^3 + \xi_2^3 + \xi_3^2$ in \mathbb{R}^3 , so that $a(\xi)$ fails to satisfy $\nabla a(\xi) \neq 0$ at $(\xi_1, \xi_2, \xi_3) = (0, 0, 0)$. Choosing appropriate cut-off functions χ in Corollary 3.5, we get the estimates

$$\begin{aligned} \left\| \langle x_k \rangle^{-s} |D_k| e^{itD_k^3} \varphi(x) \right\|_{L_{t,x}^2} &\leq C \|\varphi\|_{L^2(\mathbb{R}_x^3)} \quad (k = 1, 2), \\ \left\| \langle x_3 \rangle^{-s} |D_3|^{1/2} e^{itD_3^2} \varphi(x) \right\|_{L_{t,x}^2} &\leq C \|\varphi\|_{L^2(\mathbb{R}_x^3)} \end{aligned}$$

for $s > 1/2$, which implies

$$\left\| \langle x \rangle^{-s} (|D_1| + |D_2| + |D_3|^{1/2}) e^{ita(D_x)} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^3)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x^3)},$$

hence we get estimate (6.3) in the same way as in Example 7.2.

Example 7.4. Equations of the third order often appear in applications to KdV and nonlinear Schrödinger equation. For example, the Shrira equation [Sh] describing the propagation of a three-dimensional packet of weakly nonlinear internal gravity waves leads to third order polynomials in two dimensions. Strichartz estimates for the corresponding solutions have been analysed by e.g. Ghidaglia and Saut [GS] and

by Ben-Artzi, Koch and Saut [BKS] by reducing the equations to pointwise estimates for operators in normal forms given by

$$a(\xi) = \xi_1^3 + \xi_2^3, \quad a(\xi) = \frac{1}{6}\xi_1^3 + \frac{1}{2}\xi_2^2 \quad \text{and} \quad a(\xi) = \frac{1}{2}(\xi_1^2 + \xi_1\xi_2^2).$$

By the same argument of Example 7.3, we obtain estimate (6.3) for the first two polynomials ($s > 1/2$). For the third polynomial, we use Corollary 3.6 to obtain the estimates

$$\begin{aligned} \left\| \langle x_1 \rangle^{-s} |D_1 + D_2^2/2|^{1/2} e^{ita(D_1, D_2)} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^2)} &\leq C \|\varphi\|_{L^2(\mathbb{R}_x^2)}, \\ \left\| \langle x_2 \rangle^{-s} |D_1 D_2|^{1/2} e^{ita(D_1, D_2)} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^2)} &\leq C \|\varphi\|_{L^2(\mathbb{R}_x^2)} \end{aligned}$$

for $s > 1/2$, which imply estimate (6.3) by the same argument as in Example 7.2.

Now we will present two more approaches to treat non-dispersive equations. Recall that, in Section 5, the method of canonical transformation effectively works to reduce smoothing estimates for dispersive equations to standard estimates. We explain here that this strategy works for non-dispersive cases as well.

We will however look at the rank of the Hessian $\nabla^2 a(\xi)$, instead of the principal type assumption $\nabla a(\xi) \neq 0$. Assume now that $a(\xi) \in C^\infty(\mathbb{R}^n \setminus 0)$ is real-valued and positively homogeneous of order two. It can be noted that from Euler's identity we obtain

$$(7.1) \quad \nabla a(\xi) = \xi \nabla^2 a(\xi)$$

since $\nabla a(\xi)$ is homogeneous of order one (here ξ is viewed as a row). Then the condition $\text{rank } \nabla^2 a(\xi) = n$ implies $\nabla a(\xi) \neq 0$ ($\xi \neq 0$), and as we have already explained, we have estimates (6.3) and (6.4) by Theorems 5.1 and 5.2 in this favourable case. We will show that in the non-dispersive situation the rank of $\nabla^2 a(\xi)$ still has a responsibility for smoothing properties. We assume

$$(7.2) \quad \text{rank } \nabla^2 a(\xi) \geq k \quad \text{whenever} \quad \nabla a(\xi) = 0 \quad (\xi \neq 0)$$

with some $1 \leq k \leq n - 1$. We note that condition (7.2) is invariant under the canonical transformation in the following sense:

Lemma 7.1. *Let $a = \sigma \circ \psi$, with $\psi : U \rightarrow \mathbb{R}^n$ satisfying $\det D\psi(\xi) \neq 0$ on an open set $U \subset \mathbb{R}^n$. Then, for each $\xi \in U$, $\nabla a(\xi) = 0$ if and only if $\nabla \sigma(\psi(\xi)) = 0$. Furthermore the ranks of $\nabla^2 a(\xi)$ and $\nabla^2 \sigma(\psi(\xi))$ are equal on Γ whenever $\xi \in U$ and $\nabla a(\xi) = 0$.*

Proof. Differentiation gives $\nabla a(\xi) = \nabla \sigma(\psi(\xi)) D\psi(\xi)$ and we have the first assertion. Another differentiation gives $\nabla^2 a(\xi) = \nabla^2 \sigma(\psi(\xi)) D\psi D\psi$ when $\nabla a(\xi) = 0$. This implies the second assertion. \square

To fix the notation, we assume

$$(7.3) \quad \nabla a(e_n) = 0 \quad \text{and} \quad \text{rank } \nabla^2 a(e_n) = k \quad (1 \leq k \leq n - 1),$$

where $e_n = (0, \dots, 0, 1)$. Then we have $a(e_n) = 0$ by Euler's identity $2a(\xi) = \xi \cdot \nabla a(\xi)$. We claim that there exists a conic neighbourhood $\Gamma \subset \mathbb{R}^n \setminus 0$ of e_n and a homogeneous

C^∞ -diffeomorphism $\psi : \Gamma \rightarrow \tilde{\Gamma}$ (satisfying $\psi(\lambda\xi) = \lambda\psi(\xi)$ for all $\lambda > 0$ and $\xi \in \Gamma$) as appeared in Section 4 such that we have the form

$$(7.4) \quad a(\xi) = (\sigma \circ \psi)(\xi), \quad \sigma(\eta) = c_1\eta_1^2 + \cdots + c_k\eta_k^2 + r(\eta_{k+1}, \dots, \eta_n),$$

where $\eta = (\eta_1, \dots, \eta_n)$ and $c_j = \pm 1$ ($j = 1, 2, \dots, k$). We remark that r must be real-valued and positively homogeneous of order two.

We will prove the existence of such ψ that will satisfy (7.4). By (7.1), (7.3), and the symmetricity, all the entries of the matrix $\nabla^2 a(e_n)$ are zero except for the (perhaps) non-zero upper left $(n-1) \times (n-1)$ corner matrix. Moreover, by a linear transformation involving only the first $(n-1)$ variables of $\xi = (\xi_1, \dots, \xi_{n-1}, \xi_n)$, we may assume $\partial^2 a / \partial \xi_1^2(e_n) \neq 0$. We remark that (7.3) still holds under this transformation. Then, by the Malgrange preparation theorem, we can write

$$(7.5) \quad a(\xi) = \pm c(\xi)^2(\xi_1^2 + a_1(\xi')\xi_1 + a_2(\xi')), \quad \xi' = (\xi_2, \dots, \xi_n).$$

locally in a neighbourhood of e_n , where $c(\xi) > 0$ is some strictly positive function, while function a_1 and a_2 are smooth and real valued. Restricting this expression to the hyperplane $\xi_n = 1$, and using the homogeneity

$$a(\xi) = \pm \xi_n^2 a(\xi_1/\xi_n, \dots, \xi_{n-1}/\xi_n, 1),$$

we can extend the expression (7.5) to a conic neighbourhood Γ of e_n , so that functions $c(\xi)$, $a_1(\xi')$ and $a_2(\xi')$ are positively homogeneous of orders zero, one, and two, respectively. Let us define $\psi_0(\xi) = c(\xi)\xi$ and $\tau(\eta) = \pm(\eta_1^2 + a_1(\eta')\eta_1 + a_2(\eta'))$, so that $a(\xi) = (\tau \circ \psi_0)(\xi)$, where we write $\eta = (\eta_1, \eta')$, $\eta' = (\eta_2, \dots, \eta_n)$. Furthermore, let us define $\psi_1(\xi) = (\xi_1 + \frac{1}{2}a_1(\xi'), \xi')$, so that $\tau(\xi) = (\sigma \circ \psi_1)(\xi)$ with $\sigma(\eta) = \eta_1^2 + r(\eta')$, where $r(\eta') = a_2(\eta') - \frac{1}{4}a_1(\eta')^2$ is positively homogeneous of degree two. Then we have $a = \sigma \circ \psi$, where $\psi = \psi_1 \circ \psi_0$, and thus we have the expression (7.4) with $k = 1$.

We note that, by the construction, we have $\psi(e_n) = c(e_n)(\frac{1}{2}a_1(e'_n), e'_n)$, where $c(e_n) > 0$ and $e'_n = (0, \dots, 0, 1) \in \mathbb{R}^{n-1}$. Then we can see that the function $r(\eta')$ of $(n-1)$ -variables is defined on a conic neighbourhood of e'_n in \mathbb{R}^{n-1} . On account of this fact and Lemma 7.1, we can apply the same argument above to $r(\eta')$, and repeating the process k -times, we have the expression (7.4).

To complete the proof, we check that $\det D\psi_0(\xi) = c(\xi)^n$, which clearly implies $\det D\psi(\xi) = c(\xi)^n$, and assures that it does not vanish on a sufficiently narrow Γ . We observe first that $D\psi_0(\xi) = c(\xi)I_n + {}^t\xi\nabla c(\xi)$, where I_n is the identity n by n matrix. We note that if we consider the matrix $A = (\alpha_i\beta_j)_{i,j=1} = {}^t\alpha\beta$, where $\alpha = (\alpha_1, \dots, \alpha_n)$, $\beta = (\beta_1, \dots, \beta_n)$, then A has rank one, so its eigenvalues are $n-1$ zeros and some λ . But $\text{Tr } A$ is also the sum of the eigenvalues, hence $\lambda = \text{Tr } A$. Now, let $\alpha = \xi$, $\beta = \nabla c(\xi)$, and $A = {}^t\alpha\beta$. Since $c(\xi)$ is homogeneous of order zero, by Euler's identity we have $\text{Tr } A = \xi \cdot \nabla c(\xi) = 0$, hence all eigenvalues of A are zero. It follows now that there is a non-degenerate matrix S such that $S^{-1}AS$ is strictly upper triangular. But then $\det D\psi_0(\xi) = \det(c(\xi)I_n + S^{-1}AS)$, where matrix $c(\xi)I_n + S^{-1}AS$ is upper triangular with n copies of $c(\xi)$ at the diagonal. Hence $\det D\psi_0(\xi) = c(\xi)^n$.

On account of the above observation, we have the following result which states that estimates (6.3) and (6.4) with $m = 2$ still holds for a class of non-dispersive equations. It is an illustrations of invariant estimates (1.19) and (1.20) with $m = 2$.

Theorem 7.2. *Let $a \in C^\infty(\mathbb{R}^n \setminus 0)$ be real-valued and satisfy $a(\lambda\xi) = \lambda^2 a(\xi)$ for all $\lambda > 0$ and $\xi \neq 0$. Assume that $\text{rank } \nabla^2 a(\xi) \geq n - 1$ whenever $\nabla a(\xi) = 0$ and $\xi \neq 0$. Suppose $n \geq 2$ and $s > 1/2$. Then we have*

$$\left\| \langle x \rangle^{-s} |\nabla a(D_x)|^{1/2} e^{ita(D_x)} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)}.$$

Suppose $(4 - n)/2 < \alpha < 1/2$, or $(3 - n)/2 < \alpha < 1/2$ in the elliptic case $a(\xi) \neq 0$ ($\xi \neq 0$). Then we have

$$\left\| |x|^{\alpha-1} |\nabla a(D_x)|^\alpha e^{ita(D_x)} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)}.$$

Proof. By microlocalisation and an appropriate rotation, we may assume $\text{supp } \widehat{\varphi} \subset \Gamma$, where $\Gamma \subset \mathbb{R}^n \setminus 0$ is a sufficiently narrow conic neighbourhood of the direction $e_n = (0, \dots, 0, 1)$. Since everything is all right in the dispersive case $\nabla a(e_n) \neq 0$ by Theorems 5.1 and 5.2, we assume $\nabla a(e_n) = 0$. We may also assume $n \geq 2$ since $\nabla a(e_n) = 0$ implies $\nabla a(\xi) = 0$ for all $\xi \neq 0$ in the case $n = 1$. Then we have $\text{rank } \nabla^2 a(e_n) \neq n$ by the relation (7.1), hence $\text{rank } \nabla^2 a(e_n) = n - 1$ by the assumption $\text{rank } \nabla^2 a(\xi) \geq n - 1$. In the setting (7.3) and (7.4) above, we have

$$(7.6) \quad \text{rank } \nabla^2 \widetilde{r}(\psi(e_n)) = 0$$

by Lemma 7.1, where $\widetilde{r}(\eta) = r(\eta_{k+1}, \dots, \eta_n)$. Since $k = n - 1$ in our case, we can see that r is a function of one variable and r'' vanishes identically by (7.6) and the homogeneity of r . Then r is a polynomial of order one, but is also positively homogeneous of order two. Hence we can conclude that $r = 0$ identically, and have the relation

$$a(\xi) = (\sigma \circ \psi)(\xi), \quad \sigma(\eta) = c_1 \eta_1^2 + \dots + c_{n-1} \eta_{n-1}^2.$$

Now, we have the estimates

$$\begin{aligned} \left\| \langle x \rangle^{-s} |\nabla \sigma(D_x)|^{1/2} e^{it\sigma(D_x)} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} &\leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)}, \\ \left\| |x|^{\alpha-1} |\nabla \sigma(D_x)|^\alpha e^{it\sigma(D_x)} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} &\leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)}, \end{aligned}$$

if we use the trivial inequalities $\langle x \rangle^{-s} \leq \langle x' \rangle^{-s}$ and $|x|^{\alpha-1} \leq |x'|^{\alpha-1}$, Theorems 5.1 and 5.2 with respect to x' , and the Plancherel's theorem in x_n , where $x = (x', x_n)$ and $x' = (x_1, \dots, x_{n-1})$. On account of Theorem 6.1 and the L^2_{-s} , $\dot{L}^2_{\alpha-1}$ -boundedness of the operators $I_{\psi, \gamma}$ and $I_{\psi, \gamma}^{-1}$ for $(1/2 <) s < n/2$, $-n/2 < \alpha - 1 (< -1/2)$ (see Theorem 4.3), we have the conclusion. \square

Example 7.5. The function $a(\xi) = b(\xi)^2$ satisfies condition (7.2) with $k = 1$, where $b(\xi)$ is a positively homogeneous function of order one such that $\nabla b(\xi) \neq 0$ ($\xi \neq 0$). Indeed, if $b(\xi)$ is elliptic, then $\nabla a(\xi) = 2b(\xi)\nabla b(\xi) \neq 0$ ($\xi \neq 0$). If $b(\xi_0) = 0$ at a point $\xi_0 \neq 0$, then $\nabla a(\xi_0) = 0$ and further differentiation immediately yields $\nabla^2 a(\xi_0) = 2^t \nabla b(\xi_0) \nabla b(\xi_0)$, and clearly we have $\text{rank } \nabla^2 a(\xi_0) \geq 1$. Especially in the case $n = 2$, $a(\xi)$ meets the condition in Theorem 7.2. As an example, we consider

$$a(\xi) = \frac{\xi_1^2 \xi_2^2}{\xi_1^2 + \xi_2^2},$$

where we write $\xi = (\xi_1, \xi_2)$. Setting $b(\xi) = \xi_1 \xi_2 / |\xi|$, we clearly have $a(\xi) = b(\xi)^2$ and

$$\nabla b(\xi) = \left(\frac{\xi_2^3}{|\xi|^3}, \frac{\xi_1^3}{|\xi|^3} \right),$$

hence $\nabla b(\xi) \neq 0$ ($\xi \neq 0$). Although $\nabla a(\xi) = 0$ on the lines $\xi_1 = 0$ and $\xi_2 = 0$, we have estimate (6.3) in virtue of Theorem 7.2. This is an illustration of a smoothing estimate for the Cauchy problem for an equation like

$$i\partial_t \Delta u + D_1^2 D_2^2 u = 0,$$

which can be reduced to the second order non-dispersive pseudo-differential equation with symbol $a(\xi)$ above. Similarly, we have estimates (6.3) and (6.4) for more general case

$$a(\xi) = \frac{\xi_1^2 \xi_2^2}{\xi_1^2 + \xi_2^2} + \xi_3^2 + \cdots + \xi_n^2,$$

where we write $\xi = (\xi_1, \dots, \xi_n)$ since we obtain $\text{rank } \nabla^2 a(\xi) \geq n - 1$ from the observation above.

Next we consider more general operators $a(\xi)$ of order m which may have some lower order terms. Then even the most favourable case $\det \nabla^2 a(\xi) \neq 0$ does not imply the dispersive assumption $\nabla a(\xi) \neq 0$. The method of canonical transforms, however, can also allow us to treat this problem by obtaining localised estimates near points ξ where $\nabla a(\xi) = 0$.

Assume that ξ_0 is a non-degenerate critical point of $a(\xi)$, that is, that we have $\nabla a(\xi_0) = 0$ and $\det \nabla^2 a(\xi_0) \neq 0$. Let us microlocalise around ξ_0 , so that we only look at what happens around ξ_0 . In this case, the order of the symbol $a(\xi)$ does not play any role and we do not distinguish between the main part and lower order terms. Let Γ denote a sufficiently small open bounded neighbourhood of ξ_0 so that ξ_0 is the only critical point of $a(\xi)$ in Γ . Since $\nabla^2 a(\xi_0)$ is symmetric and non-degenerate, we may assume $\nabla^2 a(\xi_0) = \text{diag}\{\pm 1, \dots, \pm 1\}$ by a linear transformation. By Morse lemma for $a(\xi)$, there exists a diffeomorphism $\psi : \Gamma \rightarrow \tilde{\Gamma} \subset \mathbb{R}^n$ with an open bounded neighbourhood of the origin such that

$$a(\xi) = (\sigma \circ \psi)(\xi), \quad \sigma(\eta) = c_1 \eta_1^2 + \cdots + c_n \eta_n^2,$$

where $\eta = (\eta_1, \dots, \eta_n)$ and $c_j = \pm 1$ ($j = 1, 2, \dots, n$). From Theorems 5.1 and 5.2 applied to operator $\sigma(D_x)$, we obtain the estimates

$$(7.7) \quad \left\| \langle x \rangle^{-s} |\nabla \sigma(D_x)|^{1/2} e^{it\sigma(D_x)} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)} \quad (s > 1/2).$$

Hence by Theorem 6.1, together with the L^2_{-s} -boundedness of the operators $I_{\psi, \gamma}$ and $I_{\psi, \gamma}^{-1}$ (which is assured by Theorem 4.2), we have estimate (7.7) with $\sigma(D_x)$ replaced by $a(D_x)$ assuming $\text{supp } \hat{\varphi} \subset \Gamma$.

On the other hand, we have the same estimate for general φ by Corollary 5.5 if we assume (L) (see the argument after estimate (6.3)). The above argument, however, assures that the following weak assumption is also sufficient if $a(\xi)$ has finitely many critical points and they are non-degenerate:

$$(L') \quad \begin{aligned} a(\xi) &\in C^\infty(\mathbb{R}^n), & |\nabla a(\xi)| &\geq C\langle \xi \rangle^{m-1} \quad (\text{for large } \xi \in \mathbb{R}^n) \quad \text{for some } C > 0, \\ |\partial^\alpha(a(\xi) - a_m(\xi))| &\leq C_\alpha |\xi|^{m-1-|\alpha|} \quad \text{for all multi-indices } \alpha \text{ and all } |\xi| \gg 1. \end{aligned}$$

Thus, we have established the following result:

Theorem 7.3. *Let $a \in C^\infty(\mathbb{R}^n)$ be real-valued and assume that it has finitely many critical points, all of which are non-degenerate. Assume also (L'). Suppose $n \geq 1$, $m \geq 1$, and $s > 1/2$. Then we have*

$$\|\langle x \rangle^{-s} |\nabla a(D_x)|^{1/2} e^{ita(D_x)} \varphi(x)\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)}.$$

8. RELATIVISTIC SCHRÖDINGER, WAVE, AND KLEIN–GORDON EQUATIONS

In Section 5, we gave a criteria Corollary 5.7 for smoothing estimates to hold in the radially symmetric case. Such subject has been also investigated by Walther [Wa2], and he derived another type of criteria based on certain integrals involving Bessel functions and their asymptotics. However, the approach presented in this paper applies to such estimates in an essentially different way in the sense that instead of verifying convergence of infinitely many integrals involving expressions based on special functions we simply compare the estimate we want to have to one that we already know to hold (in a model case or otherwise).

A typical direct application of Corollary 5.7 is to the relativistic Schrödinger type equations

$$(Relativistic Schrödinger) \quad \begin{cases} (i\partial_t - \sqrt{1 - \Delta_x}) u(t, x) = 0, \\ u(0, x) = \varphi(x). \end{cases}$$

In [BN], Ben-Artzi and Nemirovsky proved the following results. Suppose first that $h \in C^1(\mathbb{R}_+)$ is real valued, $h' > 0$, and h' is locally Hölder continuous. Then, it follows that $h(-\Delta_x)$ is self-adjoint in $L^2(\mathbb{R}^n)$ and its spectrum is absolutely continuous and satisfies $\sigma(h(-\Delta_x)) = [\overline{h(0)}, h(\infty)]$, where $h(\infty) = \lim_{\theta \rightarrow \infty} h(\theta)$. Suppose further that $h'(\theta)$ satisfies a uniform Hölder condition near $\theta = 0$ and that $h'(0) > 0$. We remark that then we have

$$(8.1) \quad h'(\theta) \geq C \quad \text{as } \theta \searrow 0$$

for some $C > 0$. Assuming also $n \geq 3$ and

$$(8.2) \quad h'(\theta) \geq \frac{C}{\sqrt{\theta}} \quad \text{as } \theta \rightarrow +\infty$$

for some $C > 0$, Ben-Artzi and Nemirovsky proved the estimate

$$(8.3) \quad \|\langle x \rangle^{-1} e^{-ith(-\Delta_x)} \varphi\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)}.$$

for the solution $u(t, x) = e^{-ith(-\Delta_x)} \varphi$ to the equation

$$(8.4) \quad \begin{cases} (i\partial_t - h(-\Delta_x)) u(t, x) = 0, \\ u(0, x) = \varphi(x). \end{cases}$$

In particular, for $h(\theta) = \sqrt{1 + \theta}$, this leads to the time global estimate for the relativistic Schrödinger equation:

$$(8.5) \quad \|\langle x \rangle^{-1} e^{-it\sqrt{1-\Delta_x}} \varphi\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)}.$$

We remark that the order of the weight $\langle x \rangle^{-1}$ in estimate (8.5) is sharp (see Walther [Wa2], for example, or Section 6). However, it can still be refined, and this will be done in Theorem 8.1.

The proof of [BN] is based on the limiting absorption principle for the resolvent of the operator $h(-\Delta_x)$. But the comparison principle also allows us to get a simple proof of several refinements of estimate (8.3). Now we remark that, by looking at invariant estimates (6.3) and (6.4) in Section 6 for equation (8.4), we should expect the estimates of the form

$$\begin{aligned} & \|\langle x \rangle^{-s} |D_x h'(-\Delta_x)|^{1/2} e^{-ith(-\Delta_x)} \varphi(x)\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)}, \\ & \||x|^{\alpha-m/2} |D_x h'(-\Delta_x)|^{\alpha/(m-1)} e^{-ith(-\Delta_x)} \varphi(x)\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)} \quad (m \neq 1), \end{aligned}$$

where the order m of the operator $h(-\Delta_x)$ has a different meaning for low frequency ($m = 2$) and high frequency ($m = 1$). In fact, these estimate can be shown using the comparison principle in Theorem 2.5 for radially symmetric operators, also without assumptions (8.1) and (8.2). In a special case with conditions (8.1) and (8.2), we get the following realisation of these estimates:

Theorem 8.1. *Suppose $n \geq 1$, $s > 1/2$, and $1 - n/2 < \alpha < 1/2$. Let $h \in C^1(\mathbb{R}_+)$ be a real-valued and strictly increasing function which satisfies (8.1) and (8.2). Let $\chi \in C_0^\infty(\mathbb{R}^n)$ be equal to one in a neighbourhood of the origin. Then we have*

$$(8.6) \quad \|\langle x \rangle^{-s} |D_x|^{1/2} e^{-ith(-\Delta_x)} \varphi_l(x)\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi_l\|_{L^2(\mathbb{R}_x^n)},$$

$$(8.7) \quad \|\langle x \rangle^{-s} e^{-ith(-\Delta_x)} \varphi_h(x)\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi_h\|_{L^2(\mathbb{R}_x^n)},$$

$$(8.8) \quad \||x|^{\alpha-1} |D_x|^\alpha e^{-ith(-\Delta_x)} \varphi_l(x)\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi_l\|_{L^2(\mathbb{R}_x^n)},$$

$$(8.9) \quad \||x|^{\alpha-1} |D_x|^{\alpha-1/2} e^{-ith(-\Delta_x)} \varphi_h(x)\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi_h\|_{L^2(\mathbb{R}_x^n)},$$

where $\varphi_l = \chi(D_x)\varphi$ and $\varphi_h = (1 - \chi(D_x))\varphi$. Consequently, if $n \geq 3$, we have

$$(8.10) \quad \|\langle x \rangle^{-1} e^{-ith(-\Delta_x)} \varphi(x)\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)}.$$

If $n = 2$ and $r > 1$, we have

$$(8.11) \quad \|\langle x \rangle^{-r} e^{-ith(-\Delta_x)} \varphi(x)\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)}.$$

We note straight away that estimates (8.6) and (8.8) improve Ben-Artzi and Nirovsky's estimate (8.3) for the low frequency part, while (8.7) also improves the weight given in (8.3) for the high frequency part. From this point of view, we can see that estimate (8.3) does only capture estimate (8.8) with $\alpha = 0$ for the low frequency part of the smoothing. In fact, (8.8) with $\alpha = 0$ improves the low frequency part of (8.3) to the better weight $|x|^{-1}$ in (8.8), compared to $\langle x \rangle^{-1}$ in (8.3).

Proof. Taking $f(\rho) = -h(\rho^2)$, condition (8.1) implies that $|f'(\rho)| \geq C\rho$ as $\rho \searrow 0$. At the same time, condition (8.2) implies that $|f'(\rho)| = 2\rho h'(\rho^2) \geq C$ as $\rho \rightarrow +\infty$. It follows that we can take $\sigma(\rho)$ to be $\sigma(\rho) = \rho^{1/2}$ for small ρ and $\sigma(\rho) = 1$ for large ρ to meet condition (5.8) in Corollary 5.7. Then estimates (5.9) and (5.10) imply estimates (8.6)–(8.9). Estimate (8.10) is just a consequence estimate (8.7) and (8.8) with $\alpha = 0$, which we can take to meet $1 - n/2 < \alpha < 1/2$ if $n \geq 3$. In the case $n = 2$, instead of (8.9) with $\alpha = 0$, we alternatively use the estimate

$$\left\| \langle x \rangle^{-r} e^{-ith(-\Delta_x)} \varphi_l(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi_l\|_{L^2(\mathbb{R}_x^n)} \quad (r > 1),$$

which can be easily given by the comparison (use Theorem 2.5) with the estimate

$$\left\| \langle x \rangle^{-r} e^{it\Delta_x} \varphi_l(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi_l\|_{L^2(\mathbb{R}_x^n)} \quad (r > 1)$$

for Schrödinger equations in the case $n = 2$. This type of estimate can be found in Ben-Artzi and Klainerman [BK] or Walther [Wa1]. \square

Taking $h(\theta) = \sqrt{1 + \theta}$ in Theorem 8.1 as a special case, we obtain estimates (8.6)–(8.11) for solutions to the relativistic Schrödinger equation. For example, estimate (8.5) is a special case of estimate (8.10) (that is, estimate (8.3)). We can also observe the refinement of the weight in (8.5) for both high and low frequencies, given by (8.7) and (8.8) with $\alpha = 0$, to $\langle x \rangle^{-s}$ and $|x|^{-1}$, respectively. We also remark that by the comparison principle for radially symmetric operators, all of these estimates are equivalent to corresponding estimates for Schrödinger or wave equation, which can be also derived from pointwise estimates in one dimension as was explained in Section 5. More precisely, by Theorem 2.5, we have the equalities

$$(8.12) \quad \begin{aligned} \left\| e^{-it\sqrt{1-\Delta_x}} \varphi(x) \right\|_{L^2(\mathbb{R}_t)} &= \sqrt{2} \left\| \langle D_x \rangle^{1/2} e^{it\Delta_x} \varphi(x) \right\|_{L^2(\mathbb{R}_t)} \\ &= \left\| |D_x|^{-1/2} \langle D_x \rangle^{1/2} e^{\pm it\sqrt{-\Delta_x}} \varphi(x) \right\|_{L^2(\mathbb{R}_t)} \end{aligned}$$

for almost all $x \in \mathbb{R}^n$. In fact, since $f(\rho) = -\sqrt{1 + \rho^2}$ and $g(\rho) = -\rho^2$ satisfy $1/|f'(\rho)|^{1/2} = |2f(\rho)|^{1/2}/|g'(\rho)|^{1/2}$, we have the first equality. The proof of the second one is similar. Then multiplying appropriate weight functions to the both sides of equalities (8.12) and integrating them in x imply the equivalence of the estimates.

For example, by (8.12), we have the equivalence of the estimate

$$(8.13) \quad \left\| \langle x \rangle^{-1} \langle D_x \rangle^{1/2} e^{it\Delta_x} \varphi \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)}$$

for the standard Schrödinger equation and estimate (8.5) for the relativistic Schrödinger equation. We remark that Corollary 5.3 also assures estimate (8.13) in the case $n \geq 3$, so we have

Theorem 8.2. *Let $n \geq 3$. Then we have equivalent estimates (8.5) and (8.13). We also have the equality*

$$(8.14) \quad \left\| \langle x \rangle^{-1} e^{-it\sqrt{1-\Delta_x}} \varphi \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} = \sqrt{2} \left\| \langle x \rangle^{-1} \langle D_x \rangle^{1/2} e^{it\Delta_x} \varphi \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)}.$$

Such equivalence as in Theorem 8.2 was shown by Walther [Wa2] (but without equivalence nor without $\sqrt{2}$), who used an explicit calculation using spherical harmonics and Bessel functions, specific for the radially symmetric case, but it is easy to

see it if we use the comparison method. Similar equivalence between the relativistic Schrödinger equation and the wave equation can be also given by equality (8.12):

$$\begin{aligned}
& \left\| \langle x \rangle^{-s} e^{\pm it\sqrt{-\Delta_x}} \varphi_l(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \sim \left\| \langle x \rangle^{-s} |D_x|^{1/2} e^{-it\sqrt{1-\Delta_x}} \varphi_l(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)}, \\
& \left\| \langle x \rangle^{-s} e^{\pm it\sqrt{-\Delta_x}} \varphi_h(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \sim \left\| \langle x \rangle^{-s} e^{-it\sqrt{1-\Delta_x}} \varphi_h(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)}, \\
& \left\| |x|^{\beta-1/2} |D_x|^\beta e^{\pm it\sqrt{-\Delta_x}} \varphi_l(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \sim \left\| |x|^{\beta-1/2} |D_x|^{\beta+1/2} e^{-it\sqrt{1-\Delta_x}} \varphi_l(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)}, \\
& \left\| |x|^{\beta-1/2} |D_x|^\beta e^{\pm it\sqrt{-\Delta_x}} \varphi_h(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \sim \left\| |x|^{\beta-1/2} |D_x|^\beta e^{-it\sqrt{1-\Delta_x}} \varphi_h(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)}.
\end{aligned}$$

As another consequence of Theorem 8.1, we have the estimates

$$\begin{aligned}
(8.15) \quad & \left\| \langle x \rangle^{-s} e^{\pm it\sqrt{-\Delta_x}} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)} \quad (s > 1/2), \\
& \left\| |x|^{\beta-1/2} |D_x|^\beta e^{\pm it\sqrt{-\Delta_x}} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)} \quad ((1-n)/2 < \beta < 0)
\end{aligned}$$

for $n \geq 1$. Indeed, we also obtain the first estimate from Theorem 5.1 and the second estimate from Theorem 5.2, or from (3.7) with $m = 1$. We note that contrary to the relativistic Schrödinger equation, here we get the same estimates for low and high frequencies. The critical case of the second estimate with $\beta = 0$ was analysed by the authors in [RS3] and it was shown that its modification still holds by introducing the Laplace-Beltrami operator on the sphere into the estimate. In fact, that analysis was done for general second order strictly hyperbolic equations with homogeneous symbols with critical sets associated to some sets related to the classical orbits. See Section 11 for further information.

Now we apply estimate (8.15) to the wave equation

$$(\text{Wave Equation}) \quad \begin{cases} \partial_t^2 u - \Delta u = 0, \\ u(0, x) = u_0(x), \\ \partial_t u(0, x) = v_0(x). \end{cases}$$

Then we have estimates

$$\begin{aligned}
(8.16) \quad & \left\| \langle x \rangle^{-s} u \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C (\|u_0\|_{L^2(\mathbb{R}_x^n)} + \||D_x|^{-1} v_0\|_{L^2(\mathbb{R}_x^n)}), \\
& \left\| |x|^{\beta-1/2} |D_x|^\beta u \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C (\|u_0\|_{L^2(\mathbb{R}_x^n)} + \||D_x|^{-1} v_0\|_{L^2(\mathbb{R}_x^n)}),
\end{aligned}$$

where we can take any $n \geq 1$, $s > 1/2$, and $(1-n)/2 < \beta < 0$. These estimates have been previously established for $n \geq 3$ and $-1 < \beta < 0$ (see Ben-Artzi [Be], where spectral methods were used). These estimates follow now from the smoothing estimates for propagators $e^{\pm it\sqrt{-\Delta_x}}$, which can be obtained by the comparison principle. We note that the usual way of relating smoothing estimates of wave and Schrödinger equation goes via a change of variables in the corresponding restriction theorems (see, for example, [RS3]). Now we can relate them directly by the comparison principle in Theorem 2.5. We also note that in the case of $n \geq 3$ and $\beta = -1/2$ the best constant $\sqrt{\frac{2\pi}{n-2}}$ in the second inequality is given by (3.10) with $m = 1$.

Let us finally state smoothing estimates for the Klein–Gordon equation

$$(8.17) \quad \begin{cases} \partial_t^2 u - \Delta u + \mu^2 u = 0, \\ u(0, x) = u_0(x), \\ \partial_t u(0, x) = v_0(x), \end{cases}$$

for $\mu > 0$. In the case $n \geq 3$ the estimate

$$(8.17) \quad \|\langle x \rangle^{-1} u\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C(\|u_0\|_{L^2(\mathbb{R}_x^n)} + \|(\mu^2 - \Delta)^{-1/2} v_0\|_{L^2(\mathbb{R}_x^n)})$$

was given in [Be]. Since propagators here are of the form $e^{\pm it\sqrt{\mu^2 - \Delta_x}}$, we can apply Theorem 8.1 with $h(\theta) = \sqrt{\mu^2 + \theta}$. In particular, this implies estimate (8.17), as well as all of its refinements given by Theorem 8.1. In particular, we get the weight $\langle x \rangle^{-s}$ with $s > 1$ in the case of $n = 2$, and better weights for high frequencies in all dimensions $n \geq 1$.

9. MODEL ESTIMATES FOR INHOMOGENEOUS EQUATIONS

We now turn to deal with inhomogeneous equations, for which we also have similar smoothing estimates. Such estimates are necessary for nonlinear applications, and they can be obtained by further developments of the presented methods. Note that

$$u(t, x) = -i \int_0^t e^{i(t-\tau)a(D_x)} f(\tau, x) d\tau$$

solves the equation

$$\begin{cases} (i\partial_t + a(D_x)) u(t, x) = f(t, x) & \text{in } \mathbb{R}_t \times \mathbb{R}_x^n, \\ u(0, x) = 0 & \text{in } \mathbb{R}_x^n. \end{cases}$$

We will give model estimates for it below, where we write $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ and $D_x = (D_1, D_2, \dots, D_n)$. We also write $x = x_1$, $D_x = D_1$ in the case $n = 1$, and $(x, y) = (x_1, x_2)$, $(D_x, D_y) = (D_1, D_2)$ in the case $n = 2$.

Theorem 9.1. *Suppose $n = 1$ and $m > 0$. Let $a(\xi) \in C^\infty(\mathbb{R} \setminus 0)$ be a real-valued function which satisfies $a(\lambda\xi) = \lambda^m a(\xi)$ for all $\lambda > 0$ and $\xi \neq 0$. Then we have*

$$(9.1) \quad \left\| a'(D_x) \int_0^t e^{i(t-\tau)a(D_x)} f(\tau, x) d\tau \right\|_{L^2(\mathbb{R}_t)} \leq C \int_{\mathbb{R}} \|f(t, x)\|_{L^2(\mathbb{R}_t)} dx$$

for all $x \in \mathbb{R}$. Suppose $n = 2$ and $m > 0$. Then we have

$$(9.2) \quad \left\| |D_x|^{m-1} \int_0^t e^{i(t-\tau)|D_x|^{m-1} D_y} f(\tau, x, y) d\tau \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x)} \leq C \int_{\mathbb{R}} \|f(t, x, y)\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x)} dy$$

for all $y \in \mathbb{R}$.

Corollary 9.2. *Suppose $n \geq 1$, $m > 0$, and $s > 1/2$. Let $a(\xi) \in C^\infty(\mathbb{R} \setminus 0)$ be a real-valued function which satisfies $a(\lambda\xi) = \lambda^m a(\xi)$ for all $\lambda > 0$ and $\xi \neq 0$. Then we have*

$$\left\| \langle x_n \rangle^{-s} a'(D_n) \int_0^t e^{i(t-\tau)a(D_n)} f(\tau, x) d\tau \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\langle x_n \rangle^s f(t, x)\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)}.$$

Suppose $n \geq 2$, $m > 0$, and $s > 1/2$. Then we have

$$\left\| \langle x_1 \rangle^{-s} |D_n|^{m-1} \int_0^t e^{i(t-\tau)D_1|D_n|^{m-1}} f(\tau, x) d\tau \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\langle x_1 \rangle^s f(t, x)\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)}.$$

Theorem 9.1 with the case $n = 1$ is a unification of the results by Kenig, Ponce and Vega who treated the cases $a(\xi) = \xi^2$ ([KPV3, p.258]), $a(\xi) = |\xi|\xi$ ([KPV4, p.160]), and $a(\xi) = \xi^3$ ([KPV2, p.533]). Corollary 9.2 is a straightforward result of Theorem 9.1 and Cauchy–Schwarz’s inequality.

Since we unfortunately do not know the comparison principle for inhomogeneous equations, we cannot reduce Theorem 9.1 to more elementary estimates as we can successfully do that for homogeneous equations in Section 3. Hence we will give a direct proof to Theorem 9.1. Note that we have another expression of the solution to inhomogeneous equation

$$\begin{cases} (i\partial_t + a(D_x)) u(t, x) = f(t, x) & \text{in } \mathbb{R}_t \times \mathbb{R}_x^n, \\ u(0, x) = 0 & \text{in } \mathbb{R}_x^n, \end{cases}$$

using the weak limit $R(\tau \pm i0)$ of the resolvent $R(\tau \pm i\varepsilon)$ as $\varepsilon \searrow 0$, where $R(\lambda) = (a(D_x) - \lambda)^{-1}$:

$$(9.3) \quad \begin{aligned} u(t, x) &= -i \int_0^t e^{i(t-\tau)a(D_x)} f(\tau, x) d\tau \\ &= \mathcal{F}_\tau^{-1} R(\tau - i0) \mathcal{F}_t f^+ + \mathcal{F}_\tau^{-1} R(\tau + i0) \mathcal{F}_t f^- \end{aligned}$$

(see Sugimoto [Su1] and Chihara [Ch]). Here \mathcal{F}_t denotes the Fourier Transformation in t and \mathcal{F}_τ^{-1} its inverse, and $f^\pm(t, x) = f(t, x)Y(\pm t)$ is the characteristic function $Y(t)$ of the set $\{t \in \mathbb{R} : t > 0\}$.

Proof of Estimate (9.1). Let us use a variant of the argument of Chihara [Ch, Section 4]. We set $R(\lambda) = (a(D_x) - \lambda)^{-1}$ and show the estimate

$$|a'(D_x)R(s \pm i0)g(x)| \leq C \int_{\mathbb{R}} |g(x)| dx,$$

where $C > 0$ is a constant independent of $s \in \mathbb{R}$, $x \in \mathbb{R}$ and $g \in L^1(\mathbb{R})$. Then, on account of the expression (9.3), Plancherel’s theorem, and Minkowski’s inequality, we have the desired result. For this purpose, we consider the kernel

$$k(s, x) = \mathcal{F}^{-1} [a'(\xi)(a(\xi) - (s \pm i0))^{-1}] (x)$$

and show its uniform boundedness. By the scaling argument, everything is reduced to show the estimates

$$\sup_{x \in \mathbb{R}} |k(\pm 1, x)| \leq C \quad \text{and} \quad \sup_{x \in \mathbb{R}} |k(0, x)| \leq C.$$

By using an appropriate partition of unity $\widehat{\phi}_1(\xi) + \widehat{\phi}_2(\xi) + \widehat{\phi}_3(\xi) = 1$, we split $k(\pm 1, x)$ into the corresponding three parts $k = k_1 + k_2 + k_3$, where $\widehat{\phi}_1$ has its support near the origin, $\widehat{\phi}_2$ near the point $\xi^m = \pm 1$, and $\widehat{\phi}_3$ away from these points. The estimate for k_1 is trivial. The other estimates are reduced to the boundedness of

$$(9.4) \quad k_0^\pm(x) = \mathcal{F}^{-1}[(\xi \pm i0)^{-1}](x) = \mp i\sqrt{2\pi}Y(\pm x).$$

In fact,

$$k_2(\pm 1, x) = \mathcal{F}^{-1}\left[(\xi - (\alpha \pm i0))^{-1}\widehat{\psi}(\xi)\right](x) = (e^{i\alpha x}k_0^\mp) * \psi(x)$$

where $\alpha \in \mathbb{R}$ is a point which solves $a(\alpha) = \pm 1$, and

$$\widehat{\psi}(\xi) = a'(\xi)\frac{\xi - \alpha}{a(\xi) - (\pm 1)}\widehat{\phi}_2(\xi) \in C_0^\infty(\mathbb{R}).$$

Furthermore, if we notice

$$\frac{a'(\xi)}{a(\xi) - s} = m\left(\frac{s}{(a(\xi) - s)\xi} + \frac{1}{\xi}\right),$$

we have

$$\frac{1}{m}k_3(\pm 1, x) = \pm \mathcal{F}^{-1}\left[\frac{\widehat{\phi}_3(\xi)}{(a(\xi) \mp 1)\xi}\right](x) + k_0^\pm(x) - k_0^\pm * (\phi_1(x) + \phi_2(x)).$$

It is easy to deduce the estimates for k_2 and k_3 . It is also easy to verify

$$\frac{a'(\xi)}{a(\xi) \pm i0} = \frac{m}{\xi \pm i0} + c\delta$$

with a constant c and Dirac's delta function δ , and have the estimate for $k(0, x)$. \square

Proof of Estimate (9.2). We set $R(\lambda) = (|D_x|^{m-1}D_y - \lambda)^{-1}$ and show the estimate

$$\| |D_x|^{m-1}R(s \pm i0)g(x, y) \|_{L^2(\mathbb{R}_x)} \leq C \int \|g(x, y)\|_{L^2(\mathbb{R}_x)} dy,$$

where $C > 0$ is a constant independent of $s \in \mathbb{R}$, $y \in \mathbb{R}$ and $g \in L^1(\mathbb{R}^2)$. Then, by the expression (9.3), Plancherel's theorem, and Minkowski's inequality again, we have the desired result.

First we note, we may assume $\widehat{g}(\xi, \eta) = 0$ for $\xi < 0$. Then we have

$$\begin{aligned} & |D_x|^{m-1}R(s \pm i0)g(x, y) \\ &= (2\pi)^{-2} \int_0^\infty \int_{-\infty}^\infty e^{i(x\xi+y\eta)} |\xi|^{m-1} (|\xi|^{m-1}\eta - (s \pm i0))^{-1} \widehat{g}(\xi, \eta) d\xi d\eta \\ &= (2\pi)^{-2} \int_0^\infty \int_{-\infty}^\infty e^{ix\xi} |\xi|^{m-1} (|\xi|^{m-1}\eta - (s \pm i0))^{-1} \widehat{g}_y(\xi, \eta) d\xi d\eta \\ &= (2\pi)^{-2} \int_{-\infty}^\infty \int_0^\infty e^{ixb} (a - (s \pm i0))^{-1} \widehat{g}_y(b, ab^{-(m-1)}) dadb \\ &= (2\pi)^{-2} \int_{-\infty}^\infty \int_0^\infty e^{ixb} \mathcal{F}_a[(a - (s \pm i0))^{-1}] \mathcal{F}_a^{-1}[\widehat{g}_y(b, ab^{-(m-1)})] dadb \\ &= (2\pi)^{-1} \int_{-\infty}^\infty \int_0^\infty e^{ixb} e^{-isa} k_0^\mp(-a) b^{m-1} \widetilde{g}_y(b, ab^{m-1}) dadb, \end{aligned}$$

hence we have

$$\mathcal{F}_x[|D_x|^{m-1}R(s \pm i0)g(x, y)](b) = \int_{-\infty}^{\infty} e^{-isa} k_0^\mp(-a) b^{m-1} \tilde{g}_y(b, ab^{m-1}) da$$

for $b \geq 0$, and it vanishes for $b < 0$. Here $g_y(x, \cdot) = g(x, \cdot + y)$, and \tilde{g}_y denotes its partial Fourier transform with respect to the first variable. We have also used here the change of variables $a = \xi^{m-1}\eta$, $b = \xi$ and Parseval's formula. Note that $\partial(a, b)/\partial(\xi, \eta) = b^{m-1}$ and k_0^\mp is a bounded function defined by (9.4). Then we have the estimate

$$\begin{aligned} |\mathcal{F}_x[|D_x|^{m-1}R(s \pm i0)g(x, y)](b)| &\leq \sqrt{2\pi} \int_{-\infty}^{\infty} |b^{m-1} \tilde{g}_y(b, ab^{m-1})| da \\ &= \sqrt{2\pi} \int_{-\infty}^{\infty} |\tilde{g}_y(b, a)| da, \end{aligned}$$

and, by Plancherel's theorem and Minkowski's inequality, we have

$$\begin{aligned} \||D_x|^{m-1}R(s \pm i0)g(x, y)\|_{L^2(\mathbb{R}_x)} &\leq \sqrt{2\pi} \int_{-\infty}^{\infty} \|g_y(x, a)\|_{L^2(\mathbb{R}_x)} da \\ &= \sqrt{2\pi} \int_{-\infty}^{\infty} \|g(x, y)\|_{L^2(\mathbb{R}_x)} dy, \end{aligned}$$

which is the desired estimate. \square

10. SMOOTHING ESTIMATES FOR DISPERSIVE INHOMOGENEOUS EQUATIONS

Let us consider the inhomogeneous equation

$$\begin{cases} (i\partial_t + a(D_x)) u(t, x) = f(t, x) & \text{in } \mathbb{R}_t \times \mathbb{R}_x^n, \\ u(0, x) = 0 & \text{in } \mathbb{R}_x^n, \end{cases}$$

where we always assume that function $a(\xi)$ is real-valued. Let the principal part $a_m(\xi) \in C^\infty(\mathbb{R}^n \setminus 0)$, be a positively homogeneous function of order m . Recall the dispersive conditions we used in Section 5:

$$\text{(H)} \quad a(\xi) = a_m(\xi), \quad \nabla a_m(\xi) \neq 0 \quad (\xi \in \mathbb{R}^n \setminus 0),$$

$$\text{(L)} \quad \begin{aligned} &a(\xi) \in C^\infty(\mathbb{R}^n), \quad \nabla a(\xi) \neq 0 \quad (\xi \in \mathbb{R}^n), \quad \nabla a_m(\xi) \neq 0 \quad (\xi \in \mathbb{R}^n \setminus 0), \\ &|\partial^\alpha(a(\xi) - a_m(\xi))| \leq C_\alpha |\xi|^{m-1-|\alpha|} \quad \text{for all multi-indices } \alpha \text{ and all } |\xi| \geq 1. \end{aligned}$$

The following is a counterpart of Theorem 5.1 which treated homogeneous equations:

Theorem 10.1. *Assume (H). Suppose $m > 0$ and $s > 1/2$. Then we have*

$$(10.1) \quad \left\| \langle x \rangle^{-s} |D_x|^{m-1} \int_0^t e^{i(t-\tau)a(D_x)} f(\tau, x) d\tau \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\langle x \rangle^s f(t, x)\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)}$$

in the case $n \geq 2$, and

$$(10.2) \quad \left\| \langle x \rangle^{-s} a'(D_x) \int_0^t e^{i(t-\tau)a(D_x)} f(\tau, x) d\tau \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x)} \leq C \|\langle x \rangle^s f(t, x)\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x)}$$

in the case $n = 1$.

Chihara [Ch] proved Theorem 10.1 with $m > 1$ under the assumption (H). As was pointed out in [Ch, p.1958], we cannot replace $a'(D_x)$ by $|D_x|^{m-1}$ in estimate (10.2) for the case $n = 1$, but there is another explanation for this obstacle. If we decompose $f(t, x) = \chi_+(D_x)f(t, x) + \chi_-(D_x)f(t, x)$, where $\chi_\pm(\xi)$ is a characteristic function of the set $\{\xi \in \mathbb{R} : \pm\xi \geq 0\}$, then we easily obtain

$$\begin{aligned} & \left\| \langle x \rangle^{-s} |D_x|^{(m-1)/2} \int_0^t e^{i(t-\tau)a(D_x)} f(\tau, x) d\tau \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x)} \\ & \leq C \left(\left\| \langle x \rangle^s |D_x|^{-(m-1)/2} f_+(t, x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x)} + \left\| \langle x \rangle^s |D_x|^{-(m-1)/2} f_-(t, x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x)} \right) \end{aligned}$$

from Theorem 10.1. But we cannot justify the estimate

$$(10.3) \quad \left\| \langle x \rangle^s |D_x|^{-(m-1)/2} f_\pm(t, x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x)} \leq \left\| \langle x \rangle^s |D_x|^{-(m-1)/2} f(t, x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x)}$$

for $s > 1/2$ by Lemma 4.1 because it requires $s < n/2$ and it is impossible for $n = 1$.

Similarly, as a counterpart of Theorem 5.4, we have

Theorem 10.2. *Assume (L). Suppose $n \geq 1$, $m > 0$, and $s > 1/2$. Then we have*

$$(10.4) \quad \left\| \langle x \rangle^{-s} \langle D_x \rangle^{m-1} \int_0^t e^{i(t-\tau)a(D_x)} f(\tau, x) d\tau \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \left\| \langle x \rangle^s f(t, x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)}.$$

The following result is a straightforward consequence of Theorem 10.2 and the L_s^2 -boundedness of $|D_x|^{(m-1)/2} \langle D_x \rangle^{-(m-1)/2}$ with $(1/2 <) s < n/2$ and $m \geq 1$ (which is assured by Lemma 4.1):

Corollary 10.3. *Assume (L). Suppose $n \geq 2$, $m \geq 1$, and $s > 1/2$. Then we have*

$$\left\| \langle x \rangle^{-s} |D_x|^{m-1} \int_0^t e^{i(t-\tau)a(D_x)} f(\tau, x) d\tau \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \left\| \langle x \rangle^s f(t, x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)}.$$

We remark that the same argument of canonical transformations as used for homogeneous equations in Section 5 works for inhomogeneous ones, as well. That is, the proofs of Theorems 10.1 and 10.2 are carried out by reducing them to model estimates in Corollary 9.2. We omit the details because the argument is essentially the same, but we just remark that we use the following slight modification of Theorem 4.1. The only difference is that we need the weighted L^2 -boundedness of the operator $I_{\psi, q}^{-1}$ instead of just the L^2 -boundedness of it induced by the boundedness of $q(\xi)$:

Theorem 10.4. *Assume that the operator $I_{\psi, \gamma}$ defined by (4.3) is $L^2(\mathbb{R}^n; w)$ -bounded. Suppose that we have the estimate*

$$\left\| w(x) \rho(D_x) \int_0^t e^{i(t-\tau)\sigma(D_x)} f(\tau, x) d\tau \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \left\| v(x) f(t, x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)}$$

for all f such that $\text{supp } \mathcal{F}_x f(t, \cdot) \subset \text{supp } \tilde{\gamma}$, where $\tilde{\gamma} = \gamma \circ \psi^{-1}$. Also assume that the operator $I_{\psi, q}^{-1}$ defined by (4.3) with $q(\xi) = (\gamma \cdot \zeta)/(\rho \circ \psi)(\xi)$ is $L^2(\mathbb{R}^n; \nu)$ -bounded. Then we have

$$\left\| w(x) \zeta(D_x) \int_0^t e^{i(t-\tau)a(D_x)} f(\tau, x) d\tau \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|v(x) f(t, x)\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)}$$

for all f such that $\text{supp } \mathcal{F}_x f(t, \cdot) \subset \text{supp } \gamma$, where $a(\xi) = (\sigma \circ \psi)(\xi)$.

The following is a counterpart of Theorem 5.6:

Theorem 10.5. *Assume (HL). Suppose $n \geq 1$, $m > 0$, $s > 1/2$, and $T > 0$. Then we have*

$$\begin{aligned} \int_0^T \left\| \langle x \rangle^{-s} \langle D_x \rangle^{m-1} \int_0^t e^{i(t-\tau)a(D_x)} f(\tau, x) d\tau \right\|_{L^2(\mathbb{R}_x^n)}^2 dt \\ \leq C \int_0^T \|\langle x \rangle^s f(t, x)\|_{L^2(\mathbb{R}_x^n)}^2 dt, \end{aligned}$$

where $C > 0$ is a constant depending on $T > 0$.

Proof. By multiplying $\chi(D_x)$ and $(1-\chi)(D_x)$ to $f(t, x)$, we decompose it into the sum of low frequency part and high frequency part, where $\chi(\xi)$ is an appropriate cut-off function. As in the proof of Theorem 5.4, the estimate for the high frequency part can be reduced to Corollaries 3.3 and 9.2 by using Theorems 4.1 and 10.4, together with the boundedness result Theorem 4.2. Here we note that, for $t \in [0, T]$,

$$\int_0^t e^{i(t-\tau)a(D_x)} f(\tau, x) d\tau = \int_0^t e^{i(t-\tau)a(D_x)} \chi_{[0, T]}(\tau) f(\tau, x) d\tau,$$

where $\chi_{[0, T]}$ denotes the characteristic function of the interval $[0, T]$. The estimate for the low frequency part is trivial. In fact, if $\text{supp}_\xi \mathcal{F}_x f(t, \xi) \subset \{\xi; |\xi| \leq R\}$, we have

$$\begin{aligned} \int_0^T \left\| \langle x \rangle^{-s} \langle D_x \rangle^{m-1} \int_0^t e^{i(t-\tau)a(D_x)} f(\tau, x) d\tau \right\|_{L^2(\mathbb{R}_x^n)}^2 dt \\ \leq \int_0^T \left(\int_0^T \|\langle D_x \rangle^{m-1} e^{i(t-\tau)a(D_x)} f(\tau, x)\|_{L^2(\mathbb{R}_x^n)} d\tau \right)^2 dt \\ \leq CT^2 \langle R \rangle^{2(m-1)} \int_0^T \|\langle x \rangle^s f(t, x)\|_{L^2(\mathbb{R}_x^n)}^2 dt. \end{aligned}$$

by Plancherel's theorem. □

If we combine Theorem 10.1 with Theorem 5.1, we have a result for the equation

$$(10.5) \quad \begin{cases} (i\partial_t + a(D_x)) u(t, x) = f(t, x) & \text{in } \mathbb{R}_t \times \mathbb{R}_x^n, \\ u(0, x) = \varphi(x) & \text{in } \mathbb{R}_x^n. \end{cases}$$

Corollary 10.6. *Assume (H). Suppose $m > 0$ and $s > 1/2$. Then the solution u to equation (10.5) satisfies*

$$\begin{aligned} & \left\| \langle x \rangle^{-s} |D_x|^{-(m-1)/2} a'(D_x) u(t, x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x)} \\ & \leq C \left(\|\varphi\|_{L^2(\mathbb{R})} + \left\| \langle x \rangle^s |D_x|^{-(m-1)/2} f(t, x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x)} \right) \end{aligned}$$

in the case $n = 1$, and

$$\begin{aligned} & \left\| \langle x \rangle^{-s} |D_x|^{(m-1)/2} u(t, x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \\ & \leq C \left(\|\varphi\|_{L^2(\mathbb{R}^n)} + \left\| \langle x \rangle^s |D_x|^{-(m-1)/2} f(t, x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \right) \end{aligned}$$

in the case $n \geq 2$.

If we combine Theorem 10.2 with Theorem 5.4, we have the following:

Corollary 10.7. *Assume (L). Suppose $n \geq 1$, $m > 0$, and $s > 1/2$. Then the solution u to equation (10.5) satisfies*

$$\begin{aligned} & \left\| \langle x \rangle^{-s} \langle D_x \rangle^{(m-1)/2} u(t, x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x)} \\ & \leq C \left(\|\varphi\|_{L^2(\mathbb{R})} + \left\| \langle x \rangle^s \langle D_x \rangle^{-(m-1)/2} f(t, x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x)} \right). \end{aligned}$$

If we combine Theorem 10.5 with Theorem 5.6, we have the following:

Corollary 10.8. *Assume (HL). Suppose $n \geq 1$, $m > 0$, $s > 1/2$, and $T > 0$. Then the solution u to equation (10.5) satisfies*

$$\begin{aligned} & \int_0^T \left\| \langle x \rangle^{-s} \langle D_x \rangle^{(m-1)/2} u(t, x) \right\|_{L^2(\mathbb{R}_x^n)}^2 dt \\ & \leq C \left(\|\varphi\|_{L^2(\mathbb{R}^n)}^2 + \int_0^T \left\| \langle x \rangle^s \langle D_x \rangle^{-(m-1)/2} f(t, x) \right\|_{L^2(\mathbb{R}_x^n)}^2 dt \right), \end{aligned}$$

where $C > 0$ is a constant depending on $T > 0$.

Corollary 10.8 is an extension of the result by Hoshiro [Ho2], which treated the case that $a(\xi)$ is a polynomial. The proof relied on Mourre's method, which is known in spectral and scattering theories. Here we use the argument of canonical transformations, extending the result and simplifying the proof.

11. TRACE THEOREMS

Another consequence of the proposed method of canonical transforms is that we can carry out the geometric analysis of the smoothing estimates leading to results relating the symbols with the location of the main singularities for solutions. To exemplify this and to give an idea of how to use it in problems at hand let us briefly mention the result that was established by the authors in [RS3] partly relying on a variant of the method developed here.

This concerns the critical case ($\alpha = 1/2$) of the Kato–Yajima’s estimate

$$(11.1) \quad \left\| |x|^{\alpha-1} |D_x|^\alpha e^{it\Delta_x} \varphi(x) \right\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)},$$

which holds for $1 - n/2 < \alpha < 1/2$. In Section 3 we argued that this estimate for values of α close to $1/2$ implies the same estimate for smaller α (see (3.8)). Thus, the critical case of this estimate with $\alpha = 1/2$ is important, especially since it can be applied to the well-posedness problems of the derivative nonlinear Schrödinger equations (see [RS5]). However, the estimate fails in the critical case (see Watanabe [W], or Chapter 6 for more general results) and it is known that it is necessary to cut-off the radial derivatives for the estimate to hold in the critical case as well (see [Su2]). This can be done by replacing operator $|D_x|^\alpha$ by the Laplace-Beltrami operator on the sphere at the level $\alpha = 1/2$. The method of canonical transforms allows one to use any operator there as long as its symbol vanishes on a certain set related to the symbol of the Laplace operator (the sphere is this case).

To explain this precisely, let us formulate it for the equation

$$(11.2) \quad \begin{cases} (i\partial_t + a(D_x)) u(t, x) = 0, \\ u(0, x) = \varphi(x) \in L^2(\mathbb{R}_x^n), \end{cases}$$

where real-valued function $a(\xi) \in C^\infty(\mathbb{R}^n \setminus 0)$ is elliptic and positively homogeneous of order two, that is, it satisfies $a(\xi) > 0$ and $a(\lambda\xi) = \lambda^2 a(\xi)$ for $\lambda > 0$ and $\xi \neq 0$. We remark that these condition assure assumption (H) with $m = 2$ in Section 5 since we have $\nabla a(\xi) \neq 0$ by the Euler’s identity $a(\xi) = 1/2 \nabla a(\xi) \cdot \xi$ and the ellipticity of $a(\xi)$. The case $a(\xi) = |\xi|^2$ corresponds to the usual Laplacian $a(D_x) = -\Delta_x$.

Let us define $\{(x(t), y(t)) : t \in \mathbb{R}\}$ to be the classical orbit, that is, the solution of the ordinary differential equation

$$\begin{cases} \dot{x}(t) = (\nabla_\xi a)(\xi(t)), & \dot{\xi}(t) = 0, \\ x(0) = 0, & \xi(0) = \xi_0, \end{cases}$$

and consider the set of the paths of all classical orbits

$$\begin{aligned} \Gamma_a &= \{(x(t), \xi(t)) : t \in \mathbb{R}, \xi_0 \in \mathbb{R}^n \setminus 0\} \\ &= \{(\lambda \nabla a(\xi), \xi) : \xi \in \mathbb{R}^n \setminus 0, \lambda \in \mathbb{R}\}. \end{aligned}$$

Let pseudo-differential operator $\sigma(X, D)$ have symbol $\sigma(x, \xi)$ which is smooth in $x \neq 0$, $\xi \neq 0$, and which is positively homogeneous of order $-1/2$ with respect to x , and of order $1/2$ with respect to ξ . Suppose also the structure condition

$$(11.3) \quad \sigma(x, \xi) = 0 \quad \text{if} \quad (x, \xi) \in \Gamma_a \quad \text{and} \quad x \neq 0.$$

Then it was shown in [RS3] that the solution $u = e^{ita(D_x)}\varphi$ to (11.2) satisfies

$$(11.4) \quad \|\sigma(X, D_x)e^{ita(D_x)}\varphi(x)\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C\|\varphi\|_{L^2(\mathbb{R}^n)}$$

if $n \geq 2$ and the Gaussian curvature of the hypersurface

$$(11.5) \quad \Sigma_a = \{\xi \in \mathbb{R}^n : a(\xi) = 1\}$$

never vanishes. The typical example for such critical operator $\sigma(X, D_x)$ is given by the elements of

$$(11.6) \quad \Omega_1 = |x|^{-1/2} \left(\frac{x}{|x|} \wedge \frac{\nabla a(D_x)}{|\nabla a(D_x)|} \right) |D_x|^{1/2},$$

where the outer product $p \wedge q$ of vectors $p = (p_1, p_2, \dots, p_n)$ and $q = (q_1, q_2, \dots, q_n)$ is defined by $p \wedge q = (p_i q_j - p_j q_i)_{i < j}$. Another interesting example is the element of

$$(11.7) \quad \Omega_2 = |x|^{-1/2} \left(\frac{\nabla a^*(x)}{|\nabla a^*(x)|} \wedge \frac{D_x}{|D_x|} \right) |D_x|^{1/2},$$

where $a^*(x)$ is the *dual function* of $a(\xi)$ which is positively homogeneous of order two and is characterised by the relation $a^*(\nabla a(\xi)) = 1$. We remark that the sum of the squares of all elements of Ω_2 forms the main factor of the homogeneous extension of the Laplace-Beltrami operator on the *dual hypersurface* $\Sigma_a^* = \{\nabla a(\xi) : \xi \in \Sigma_a\}$. The dual function $a^*(x)$ can be also determined by the relation $\Sigma_{a^*} = \Sigma_a^*$.

The proof of this result relies on the critical case of the limiting absorption principle which can be proved by reducing its statement to a model situation by the canonical transform method combined with weighted estimates for the transform operators. On the other hand, it can be reduced to a corresponding smoothing estimate for the Laplace operator with *any* critical operator, for example to the homogeneous extension of the Laplace-Beltrami operator on the sphere, recovering, in particular, the result of [Su2]. This result has been extended to include small perturbations by Barceló, Bennett and Ruiz [BBR]. For further details on these arguments we refer to authors' paper [RS3]. On the other hand, the set Γ_a corresponds to the Hamiltonian flow of $a(D_x)$, which is known to play a role in such problems also in a more general setting of manifolds. There, non-trapping conditions also enter (e.g. Doi [Do1, Do2] in the case of Schrödinger operators on manifolds, using Egorov theorem, or Burq [Bu] and Burq, Gérard and Tzvetkov [BGT] in the case of Schrödinger boundary value problems, using propagation properties of Wigner measures), and such conditions can be also expressed in terms of properties of the set Γ_a . In our case this simply corresponds to the dispersiveness of $a(D_x)$.

One important topic related to this argument are the Fourier restriction and trace theorems. Below, we frequently quote the description in [RS3, Section 5] which the reader may consult for the details. First note that the formal adjoint $T^* : \mathcal{S}(\mathbb{R}_t \times \mathbb{R}_x^n) \rightarrow \mathcal{S}'(\mathbb{R}_x^n)$ of the solution operator

$$T = e^{ita(D_x)} : \mathcal{S}(\mathbb{R}_x^n) \rightarrow \mathcal{S}'(\mathbb{R}_t \times \mathbb{R}_x^n)$$

to equation (11.2) is expressed as

$$(11.8) \quad T^*[v(t, x)] = \mathcal{F}_\xi^{-1} [(\mathcal{F}_{t,x} v)(a(\xi), \xi)].$$

Then, for any operator $A = A(X, D_x)$ acting on the variable x , the estimate

$$(11.9) \quad \|Ae^{ita(D_x)}\varphi\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)} \leq C \|\varphi\|_{L^2(\mathbb{R}_x^n)}$$

implies the estimate

$$(11.10) \quad \left\| \widehat{A^*f}|_{\Sigma_a} \right\|_{L^2(\Sigma_a; d\omega/|\nabla a|)} \leq C \|f\|_{L^2(\mathbb{R}_x^n)},$$

where $d\omega$ is the standard surface element of the hypersurface Σ_a defined by (11.5). Indeed, by (11.8) and Plancherel's theorem, we have for $v(t, x) = g(t)f(x)$

$$\begin{aligned} \|T^*A^*v\|_{L^2(\mathbb{R}^n)}^2 &= (2\pi)^{-n} \|(\mathcal{F}_{t,x}A^*v)(a(\xi), \xi)\|_{L^2(\mathbb{R}_\xi^n)}^2 \\ &= (2\pi)^{-n} \int_0^\infty \left(\int_{\Sigma_a} |(\mathcal{F}_{t,x}A^*v)(\rho^2, \rho\omega)|^2 \frac{2\rho^{n-1}d\omega}{|\nabla a(\omega)|} \right) d\rho \\ &= (2\pi)^{-n} \int_0^\infty |\widehat{g}(\rho^2)\sqrt{\rho}|^2 \left(\int_{\Sigma_a} \left| \frac{1}{\sqrt{\rho}} (\widehat{A^*f})(\rho\omega) \right|^2 \frac{2\rho^{n-1}d\omega}{|\nabla a(\omega)|} \right) d\rho. \end{aligned}$$

Here we have used the change of variables $\xi \mapsto \rho\omega$ ($\rho > 0, \omega \in \Sigma_a$). At the same time, by (11.9), we have

$$\|T^*A^*v\|_{L^2(\mathbb{R}^n)}^2 \leq C \|v\|_{L^2(\mathbb{R}_t \times \mathbb{R}_x^n)}^2 = C \|g\|_{L^2(\mathbb{R})}^2 \|f\|_{L^2(\mathbb{R}^n)}^2.$$

Note that we have by Plancherel's theorem

$$\|g\|_{L^2(\mathbb{R})}^2 = \frac{1}{2\pi} \|\widehat{g}\|_{L^2(\mathbb{R})}^2 = \frac{1}{4\pi} \int_0^\infty |\widehat{g}(\rho^2)\sqrt{\rho}|^2 d\rho,$$

if $\text{supp } \widehat{g} \subset [0, \infty)$. Combining all these relations and taking arbitrary g , we have estimate

$$(11.11) \quad \left\| \widehat{A^*f}|_{\rho\Sigma_a} \right\|_{L^2(\rho\Sigma_a; \rho^{n-1}d\omega/|\nabla a|)} \leq C \sqrt{\rho} \|f\|_{L^2(\mathbb{R}_x^n)},$$

where $\rho > 0$, $\rho\Sigma_a = \{\rho\omega : \omega \in \Sigma_a\}$. Taking $\rho = 1$, we have estimate (11.10). We remark that, conversely, estimate (11.11) implies estimate (11.9).

In this paper we have already obtained the examples of operators A which satisfy smoothing estimate (11.9), hence the Fourier restriction estimate (11.10). For example, by Theorems 5.1 and 5.2, we can take

$$(11.12) \quad \begin{aligned} A_1 &= \langle x \rangle^{-s} |D_x|^{1/2} \quad (s > 1/2), \\ A_2 &= |x|^{\alpha-1} |D_x|^\alpha \quad (1 - n/2 < \alpha < 1/2). \end{aligned}$$

We can also take $A = \sigma(X, D_x)$ which appeared in estimate (11.4), especially the elements of the operators Ω_1 or Ω_2 defined by (11.6) or (11.7), but in this case we also need the non-degenerate Gaussian curvature condition on the hypersurface Σ_a defined by (11.5), which is equivalent to $\det \nabla^2 a(\xi) \neq 0$ ($\xi \neq 0$) (see Miyachi [Mi], for example). Their formal adjoints are given by

$$\begin{aligned} A_1^* &= |D_x|^{1/2} \langle x \rangle^{-s} \quad (s > 1/2), \quad A_2^* = |D_x|^{1-s} |x|^{-s} \quad (1/2 < s < n/2), \\ \Omega_1^* &= |D_x|^{1/2} \left(\frac{\nabla a(D_x)}{|\nabla a(D_x)|} \wedge \frac{x}{|x|} \right) |x|^{-1/2}, \quad \Omega_2^* = |D_x|^{1/2} \left(\frac{D_x}{|D_x|} \wedge \frac{\nabla a^*(x)}{|\nabla a^*(x)|} \right) |x|^{-1/2}. \end{aligned}$$

Note that we have $|\nabla a(\xi)| \geq C > 0$ on Σ_a since $\nabla a(\xi) \neq 0$ ($\xi \neq 0$) in our case. From the construction, we have the same property for a^* , as well. We also note that $\|f\|_{L^2_s(\mathbb{R}^n)} = \|\widehat{f}\|_{H^s(\mathbb{R}^n)}$ and $\|f\|_{\dot{L}^2_s(\mathbb{R}^n)} = \|\widehat{f}\|_{\dot{H}^s(\mathbb{R}^n)}$, where $H^s(\mathbb{R}^n)$ and $\dot{H}^s(\mathbb{R}^n)$ are (homogeneous) Sobolev spaces with the norms $\|g\|_{H^s(\mathbb{R}^n)} = \|\langle D_x \rangle^s g\|_{L^2(\mathbb{R}^n)}$ and $\|g\|_{\dot{H}^s(\mathbb{R}^n)} = \||D_x|^s g\|_{L^2(\mathbb{R}^n)}$ respectively. Then we can conclude the following trace results:

Theorem 11.1. *Let $a(\xi) \in C^\infty(\mathbb{R}^n \setminus 0)$ be real-valued and satisfy $a(\xi) > 0$ and $a(\lambda\xi) = \lambda^2 a(\xi)$ for $\lambda > 0$ and $\xi \neq 0$. Let $\Sigma_a = \{\xi \in \mathbb{R}^n : a(\xi) = 1\}$. Suppose $s > 1/2$. Then we have*

$$(11.13) \quad \|f|_{\Sigma_a}\|_{L^2(\Sigma_a; d\omega)} \leq C \|f\|_{H^s(\mathbb{R}^n)}.$$

Moreover, for $1/2 < s < n/2$, we have

$$(11.14) \quad \|f|_{\Sigma_a}\|_{L^2(\Sigma_a; d\omega)} \leq C \|f\|_{\dot{H}^s(\mathbb{R}^n)}.$$

If we in addition assume that the Gaussian curvature of Σ_a is non-vanishing, then we have also

$$(11.15) \quad \left\| \left(\frac{\nabla a(x)}{|\nabla a(x)|} \wedge \frac{D_x}{|D_x|} \right) f|_{\Sigma_a} \right\|_{L^2(\Sigma_a; d\omega)} \leq C \|f\|_{\dot{H}^{1/2}(\mathbb{R}^n)}$$

and

$$(11.16) \quad \left\| \left(\frac{x}{|x|} \wedge \frac{\nabla a^*(D_x)}{|\nabla a^*(D_x)|} \right) f|_{\Sigma_a} \right\|_{L^2(\Sigma_a; d\omega)} \leq C \|f\|_{\dot{H}^{1/2}(\mathbb{R}^n)},$$

where $a^*(x)$ is the dual function of $a(\xi)$.

The third and fourth estimates (11.15) and (11.16) in Theorem 11.1 say that we can attain the critical order $s = 1/2$ in the first and second estimates (11.13) and (11.14) under a structure condition. In fact, we get a complete range of operators for the critical smoothing if we use the restriction estimate (11.10) with $A = \sigma(X, D_x)$ satisfying (11.3). Estimates (11.15) and (11.16) are the interesting special cases of them.

We note finally, that the results on the global smoothing imply the growth rates of the restriction norms. For example, smoothing with operator A_1 in (11.12) implies the uniform trace estimate

$$\|f|_{\rho\Sigma_a}\|_{L^2(\rho\Sigma_a, \rho^{n-1}d\omega)} \leq C \|f\|_{H^s(\mathbb{R}^n)} \quad (s > 1/2)$$

on account of (11.11). If we use A_2 in (11.12) instead, we get

$$\|f|_{\rho\Sigma_a}\|_{L^2(\rho\Sigma_a, \rho^{n-1}d\omega)} \leq C \rho^{s-1/2} \|f\|_{\dot{H}^s(\mathbb{R}^n)} \quad (n/2 > s > 1/2).$$

However, in the critical cases, in addition to (11.15) and (11.16), we obtain

$$\begin{aligned} & \left\| \left(\frac{\nabla a(x)}{|\nabla a(x)|} \wedge \frac{D_x}{|D_x|} \right) f|_{\rho\Sigma_a} \right\|_{L^2(\rho\Sigma_a, \rho^{n-1}d\omega)} \leq C \|f\|_{\dot{H}^{1/2}(\mathbb{R}^n)}, \\ & \left\| \left(\frac{x}{|x|} \wedge \frac{\nabla a^*(D_x)}{|\nabla a^*(D_x)|} \right) f|_{\rho\Sigma_a} \right\|_{L^2(\rho\Sigma_a, \rho^{n-1}d\omega)} \leq C \|f\|_{\dot{H}^{1/2}(\mathbb{R}^n)}. \end{aligned}$$

REFERENCES

- [AF] K. Asada and D. Fujiwara, *On some oscillatory integral transformations in $L^2(\mathbb{R}^n)$* , Japan. J. Math. (N.S.) **4** (1978), 299–361.
- [BBR] J.A. Barceló, J.M. Bennett and A. Ruiz, *Spherical perturbations of Schrödinger equations*, J. Fourier Anal. Appl. **12** (2006), 269–290.
- [Be] M. Ben-Artzi, *Regularity and smoothing for some equations of evolution*, Nonlinear partial differential equations and their applications. College de France Seminar, Vol. XI (Paris, 1989–1991), 1–12, Pitman Res. Notes Math. Ser., 299, Longman Sci. Tech., Harlow, 1994.
- [BD1] M. Ben-Artzi and A. Devinatz, *The limiting absorption principle for partial differential operators*, Mem. Amer. Math. Soc. **66** (1987).
- [BD2] M. Ben-Artzi and A. Devinatz, *Local smoothing and convergence properties of Schrödinger type equations*, J. Funct. Anal. **101** (1991), 231–254.
- [BK] M. Ben-Artzi and S. Klainerman, *Decay and regularity for the Schrödinger equation*, J. Analyse Math. **58** (1992), 25–37.
- [BKS] M. Ben-Artzi, H. Koch and J. C. Saut, *Dispersion estimates for third order equations in two dimensions*, Comm. Partial Differential Equations **28** (2003), 1943–1974.
- [BN] M. Ben-Artzi and J. Nirenberg, *Remarks on relativistic Schrödinger operators and their extensions*, Ann. Inst. H. Poincaré Phys. Theor. **67** (1997), 29–39.
- [Bo1] A. Boulkhemair, *Estimations L^2 précises pour des intégrales oscillantes*, Comm. Partial Differential Equations **22** (1997), 165–184.
- [Bo2] A. Boulkhemair, *L^2 continuity for Fourier integral operators with nonregular phases*, Asymptot. Anal. **40** (2004), 147–161.
- [Bu] N. Burq, *Smoothing effect for Schrödinger boundary value problems*, Duke Math. J. **123** (2004), 403–427.
- [BGT] N. Burq, P. Gérard and N. Tzvetkov, *On nonlinear Schrödinger equations in exterior domains*, Ann. Inst. H. Poincaré Anal. Non Linéaire **21** (2004), 295–318.
- [Ch] H. Chihara, *Smoothing effects of dispersive pseudodifferential equations*, Comm. Partial Differential Equations **27** (2002), 1953–2005.
- [CS] P. Constantin and J. C. Saut, *Local smoothing properties of dispersive equations*, J. Amer. Math. Soc. **1** (1988), 413–439.
- [CKS] W. Craig, T. Kappeler and W. Strauss, *Microlocal dispersive smoothing for the Schrödinger equation*, Comm. Pure Appl. Math. **48** (1995), 769–860.
- [DF] P. D’Ancona and L. Fanelli, *Decay estimates for the wave and Dirac equations with a magnetic potential*, preprint, to appear in Comm. Pure Appl. Math.
- [Do1] S. Doi, *Smoothing effects of Schrödinger evolution groups on Riemannian manifolds*, Duke Math. J. **82** (1996), 679–706.
- [Do2] S. Doi, *Smoothing effects for Schrödinger evolution equation and global behavior of geodesic flow*, Math. Ann. **318** (2000), 355–389.
- [DH] J. J. Duistermaat and L. Hörmander, *Fourier integral operators. II*, Acta Math. **128** (1972), 183–269.
- [GT] V. Georgiev and M. Tarulli, *Scale invariant energy smoothing estimates for the Schrödinger equation with small magnetic potential*, Asymptot. Anal. **47** (2006), 107–138.
- [GS] J. M. Ghidaglia and J. C. Saut, *Nonelliptic Schrödinger equations*, J. Nonlinear Sci. **3** (1993), 169–195.
- [H] L. Hörmander, *The Analysis of Linear Partial Differential Operators II*, Springer-Verlag, Berlin-New York, 1983.
- [Ho1] T. Hoshiro, *Mourre’s method and smoothing properties of dispersive equations*, Comm. Math. Phys. **202** (1999), 255–265.
- [Ho2] T. Hoshiro, *Decay and regularity for dispersive equations with constant coefficients*, J. Anal. Math. **91** (2003), 211–230.
- [KaSa] L. Kapitanski and Y. Safarov, *Dispersive smoothing for Schrödinger equations*, Math. Res. Lett. **3** (1996), 77–91.

- [Ka1] T. Kato, *Wave operators and similarity for some non-selfadjoint operators*, Math. Ann. **162** (1965/1966), 258–279.
- [Ka2] T. Kato, *On the Cauchy problem for the (generalized) Korteweg-de Vries equation*, Studies in applied mathematics, 93–128, Adv. Math. Suppl. Stud., 8, Academic Press, New York, 1983.
- [KY] T. Kato and K. Yajima, *Some examples of smooth operators and the associated smoothing effect*, Rev. Math. Phys. **1** (1989), 481–496.
- [KPV1] C. E. Kenig, G. Ponce and L. Vega, *Oscillatory integrals and regularity of dispersive equations*, Indiana Univ. Math. J. **40** (1991), 33–69.
- [KPV2] C. E. Kenig, G. Ponce and L. Vega, *Well-posedness and scattering results for the generalized Korteweg-de Vries equation via the contraction principle*, Comm. Pure Appl. Math. **46** (1993), 527–620.
- [KPV3] C. E. Kenig, G. Ponce and L. Vega, *Small solutions to nonlinear Schrödinger equations*, Ann. Inst. H. Poincaré Anal. Non Linéaire **10** (1993), 255–288.
- [KPV4] C. E. Kenig, G. Ponce and L. Vega, *On the generalized Benjamin-Ono equation*, Trans. Amer. Math. Soc. **342** (1994), 155–172.
- [KPV5] C. E. Kenig, G. Ponce and L. Vega, *On the Zakharov and Zakharov-Schulman systems*, J. Funct. Anal. **127** (1995), 204–234.
- [KoSa] H. Koch and J. C. Saut, *Local smoothing and local solvability for third order dispersive equations*, preprint.
- [Ku] H. Kumano-go, *A calculus of Fourier integral operators on \mathbb{R}^n and the fundamental solution for an operator of hyperbolic type*, Comm. Partial Differential Equations **1** (1976), 1–44.
- [KW] D. S. Kurtz and R. L. Wheeden, *Results on weighted norm inequalities for multipliers*, Trans. Amer. Math. Soc. **255** (1979), 343–362.
- [La] C. Laurey, *The Cauchy problem for a third order nonlinear Schrödinger equation*, Nonlinear Anal. **29** (1997), 121–158.
- [LP] F. Linares and G. Ponce, *On the Davey-Stewartson systems*, Ann. Inst. H. Poincaré Anal. Non Linéaire **10** (1993), 523–548.
- [MP] N. Manganaro and D.F. Parker, *Similarity reductions for variable-coefficient coupled nonlinear Schrödinger equations*, J. Phys. A: Math. Gen. **26** 4093–4106.
- [Mi] A. Miyachi, *On some estimates for the wave equation in L^p and H^p* , J. Fac. Sci. Univ. Tokyo Sect. IA Math. **27** (1980), 331–354.
- [Na] K. Nakkeeran, *Optical solitons in erbium doped fibers with higher order effects*, Physics Letters A **275** (2000), 415–418.
- [PY] D. Pelinovsky and J. Yang, *Instabilities of multi-hump vector solitons in coupled nonlinear Schrödinger equations*, Stud. Appl. Math. **115** (2005), 109–137.
- [RZ] L. Robbiano and C. Zuily, *Microlocal analytic smoothing effect for the Schrödinger equation*, Duke Math. J. **100** (1999), 93–129.
- [RS1] M. Ruzhansky and M. Sugimoto, *A new proof of global smoothing estimates for dispersive equations*, Advances in pseudo-differential operators, 65–75, Oper. Theory Adv. Appl., **155**, Birkhäuser, Basel, 2004.
- [RS2] M. Ruzhansky and M. Sugimoto, *Global L^2 -boundedness theorems for a class of Fourier integral operators*, Comm. Partial Differential Equations **31** (2006), 547–569.
- [RS3] M. Ruzhansky and M. Sugimoto, *A smoothing property of Schrödinger equations in the critical case*, Math. Ann. **335** (2006), 645–673.
- [RS4] M. Ruzhansky and M. Sugimoto, *Global smoothing estimates for dispersive equations with non-polynomial symbols*, Proceedings of The 12th International Conference on Finite or Infinite Dimensional Complex Analysis and Applications, Kyushu University Press, Fukuoka.
- [RS5] M. Ruzhansky and M. Sugimoto, *A smoothing property of Schrödinger equations and a global existence result for derivative nonlinear equations*, in Advances in Analysis 315–320, World Sci. Publ., Hackensack, NJ, 2005.
- [Sh] V. I. Shrira, *On the propagation of a three-dimensional packet of weakly nonlinear internal gravity waves*, Internat. J. Non-Linear Mech. **16** (1981), 129–138.

- [Si] B. Simon, *Best constants in some operator smoothness estimates*, J. Funct. Anal. **107** (1992), 66–71.
- [Sj] P. Sjölin, *Regularity of solutions to the Schrödinger equation*, Duke Math. J. **55** (1987), 699–715.
- [SW] E. M. Stein and G. Weiss, *Fractional integrals on n -dimensional Euclidean space*, J. Math. Mech. **7** (1958), 503–514.
- [Su1] M. Sugimoto, *Global smoothing properties of generalized Schrödinger equations*, J. Anal. Math. **76** (1998), 191–204.
- [Su2] M. Sugimoto, *A Smoothing property of Schrödinger equations along the sphere*, J. Anal. Math. **89** (2003), 15–30.
- [Ta] H. Takuwa, *Microlocal analytic smoothing effects for operators of real principal type*, Osaka J. Math. **43** (2006), 13–62.
- [TB] B. Tan and J. Boyd, *Coupled-Mode Envelope Solitary Waves in a Pair of Cubic Schrödinger Equations with Cross Modulation: Analytical Solution and Collisions*, Chaos, Solitons and Fractals (1999).
- [V] L. Vega, *Schrödinger equations: Pointwise convergence to the initial data*, Proc. Amer. Math. Soc. **102** (1988), 874–878.
- [Wa1] B. G. Walther, *A sharp weighted L^2 -estimate for the solution to the time-dependent Schrödinger equation*, Ark. Mat. **37** (1999), 381–393.
- [Wa2] B. G. Walther, *Regularity, decay, and best constants for dispersive equations*, J. Funct. Anal. **189** (2002), 325–335.
- [W] K. Watanabe, *Smooth perturbations of the selfadjoint operator $|\Delta|^{\alpha/2}$* , Tokyo J. Math. **14** (1991), 239–250.
- [ZE] F. P. Zen and H. I. Elim, *Multi-soliton solution of the integrable coupled nonlinear Schrödinger equation of Manakov type*, arXiv:solv-int/9901010.

MICHAEL RUZHANSKY:
 DEPARTMENT OF MATHEMATICS
 IMPERIAL COLLEGE LONDON
 180 QUEEN'S GATE, LONDON SW7 2AZ, UK
E-mail address m.ruzhansky@imperial.ac.uk

MITSURU SUGIMOTO:
 DEPARTMENT OF MATHEMATICS, GRADUATE SCHOOL OF SCIENCE
 OSAKA UNIVERSITY
 MACHIKANAYAMA-CHO 1-16, TOYONAKA, OSAKA 560-0043, JAPAN
E-mail address sugimoto@math.wani.osaka-u.ac.jp