PARAMETRICES AND DISPERSIVE ESTIMATES FOR SCHROEDINGER OPERATORS WITH VARIABLE COEFFICIENTS

DANIEL TATARU

ABSTRACT. In this article we consider variable coefficient time dependent Schrödinger evolutions in $\mathbb{R}^n$. For this we use phase space methods to construct outgoing parametrices and to prove global in time Strichartz type estimates. This is done in the context of C^2 metrics which satisfy a weak asymptotic flatness condition at infinity.

1. INTRODUCTION

Consider first solutions to the homogeneous Schroedinger equation in $\mathbb{R} \times \mathbb{R}^n$

$$(i\partial_t - \Delta)u = 0 \qquad u(0) = u_0$$

Their energy is preserved,

$$\|u(t)\|_{L^2} = \|u(0)\|_{L^2}.$$

At the same time due to the dispersion of the waves there is uniform decay for spatially localized initial data,

ccdisperse
$$\|u(t)\|_{L^\infty} \lesssim t^{-\frac{n}{2}} \|u(0)\|_{L^1} \tag{1}$$

One way of thinking of this is as a straightforward consequence of uniform bounds for the fundamental solution,

$$K(t,x) = c_n t^{-\frac{n}{2}} e^{i\frac{x^2}{4}t}$$

As a consequence of (1) one can also obtain time averaged decay estimates for merely L^2 initial data. These are called Strichartz estimates, and have the form

jg
$$\|u\|_{L^p(L^q)} \lesssim \|\nabla u_0\|_{L^2} \tag{2}$$

This holds for all pairs (p,q) satisfying the relations $2 \leq p \leq \infty$, $2 \leq q \leq \infty$ and

pq
$$\frac{2}{p} + \frac{n}{q} \leq \frac{n}{2} \tag{3}$$

with the exception of the forbidden endpoint $(2,\infty)$ in dimension $n = 2$. In the sequel such pairs are called Strichartz pairs.

Date: February 28, 2006 .

The author was partially supported by NSF grants DMS0354539 and DMS 0301122 and also by MSRI for Fall 2005.

A straightforward consequence of (2) is an estimate for solutions to the inhomogeneous problem

$$(i\partial_t - \Delta)u = f \qquad u(0) = 0 \quad u_t(0) = 0$$

namely

jga (4)
$$\|u\|_{L^p(L^q)} \lesssim \|f\|_{L^1(L^2)}$$

The simplest case of (4) is the well-known energy estimate

jgb (5)
$$\|\nabla u\|_{L^\infty(L^2)} \leq \|f\|_{L^1(L^2)}$$

However, there is a larger family of estimates for solutions to the inhomogeneous wave equation where we also vary the norms in the right hand side,

jgc (6)
$$\|u\|_{L^p(0,T;L^q)} \lesssim \|f\|_{L^{p_1'}(L^{q_1'})}$$

This holds for all Strichartz pairs (p, q), (p_1, q_1). For more information we refer the reader to the expository article [9]. The endpoint estimate $(p, q) = (2, \frac{2n}{n-2})$ was obtained only later in [11] $(n \geq 3)$.

In this article we are interested in the variable coefficient case of these estimates, where we replace $-\Delta$ by a second order elliptic operator of the form

$$A(t, x, D) = D_i a^{ij}(t, x) D_j$$

Thus we consider evolutions of the form

maineq (7)
$$Pu = f, \qquad u(0) = u_0$$

where

$$P = D_t + A(t, x, D)$$

This is a considerably more delicate problem, which has several new features tied to the nontrivial behaviour of its Hamilton flow.

The first of these is that dispersive estimates such as (1) do not hold in general, even if we restrict ourselves to coefficients a^{ij} which are sufficiently small smooth compactly supported perturbations of the identity. This is because even a small perturbation of the identity suffices in order to refocus a group of Hamilton flow rays originating at the same point. This produces some caustics-like concentration for the fundamental solution.

A second feature is related to the long time behavior of the bicharacteristics. In the flat case all bicharacteristics are straight so they escape to infinity both forward and backward in time. However, in the variable coefficient case it is possible to have trapped rays, which are confined to a bounded spatial region. These correspond to singularities which are largely confined to a bounded region, and destroy not only the dispersive estimates (1) but also the Strichartz estimates in (2). On the positive side, the existence of trapped rays is a more stable phenomena; in particular, it cannot happen for small perturbations of the identity.

The first work in this direction [17] considers the case of a C^2 compactly supported perturbation of the identity, subject to a nontrapping condition. Then Strichartz estimates are proved locally in time. An essential part of the argument is to take advantage of the local smoothing estimates for variable coefficient Schrödinger equations. These allow one to stably split the estimates in two, one part which is localized to a compact set and another which lives on a flat background. In the simplest form (see [6]) they are stated as

$$\|\langle x\rangle^{-\frac{1}{2}+\epsilon}\langle D\rangle^{\frac{1}{2}}u\|_{L^2([0,1]\times\mathbb{R}^n)} \lesssim \|u(0)\|_{L^2}$$

Hence they give a gain of 1/2 derivative within a compact spatial region. Heuristically this is a reflection of the fact that waves with high frequency λ move at high speed $O(\lambda)$ and thus spend a short time $O(\lambda^{-1})$ within a bounded spatial region. Square averaging in time one then obtains the half derivative gain $\lambda^{-\frac{1}{2}}$.

The results in [17] are based on a phase space analysis of the spatially localized part of the Schrödinger waves, following earlier work of Smith [16] and the author [18], [19] on the similar problem for the wave equation.

In the meantime this type local analysis has been recast in a semiclassical language in [3], who further considered various properties of Schrödinger evolutions on compact manifolds. Another partial result was obtained in [2].

Simplified presentations of localized wave packet type parametrix constructions are now available in [12], [21]. These apply to evolutions of the form

$$(D_t + a^w(t,x,D))u = 0, \qquad u(0) = u_0$$

on the unit time scale, for symbols a which satisfy a partial S^0_{00} type condition

$$|\partial_x^\alpha \partial_x^\beta a(t,x,\xi)| \leq c_{\alpha\beta}, \qquad |\alpha| + |\beta| \geq 2$$

These parametrices are often useful in rescaled forms. However due to their finite time horizon they cannot be directly applied to obtain optimal results for metrics which are not compactly supported perturbations of the identity.

More recently, two versions of parametrix constructions have been obtained for metrics which are asymptotically flat; both imply local in time Strichartz estimates.

Robbiano and Zuily [13] consider smooth asymptotically flat metrics in $\mathbb{R}^n$ of the short range type and which satisfy a nontrapping assumption. Their approach uses a parametrix which is a Fourier integral operator with complex phase and relies considerably on Sjöstrand's theory of the FBI transform.

Hassell-Tao-Wunsch MR2131050 [10] instead have a more direct parametrix construction emulating the model of the constant coefficient fundamental solution. A sharper version of the localized energy estimates is then used to control the errors. Their setup is of smooth asymptotically conic manifolds with short range scattering metrics, extended shortly afterward to long range scattering metrics.

In different directions, time independent compactly supported perturbations of the Laplacian are considered globally in time in RT [14]. Also short time bounds for time independent long range perturbations of the Laplacian outside compact sets are obtained in BT [1].

In the present article we consider global in time parametrices and Strichartz estimates for metrics in $\mathbb{R}^n$ which are merely of class C^2 and which are asymptotically flat only in a very weak sense. Due to the global nature of the result it is convenient to consider scale invariant assumptions on the coefficients. Such a scale invariant assumption is

$$|a - I_n| + |a^{-1} - I_n| + |x||\partial_x a(x,t)| + |x|^2(|\partial_x^2 a(x,t)| + |\partial_t a(x,t)|) \leq C$$

If C is small this prevents trapping, but some heuristic computations seem to indicate that the sharp pointwise decay of outgoing waves may fail because of repeated caustics formation along geodesics. Thus it is conceivable that one might be able to construct solutions which are localized along certain geodesics for a long time.

Thus we are led to introduce a slightly stronger assumption, namely

coeff
$$\sum_{j\in\mathbb{Z}} \sup_{A_j} |x|^2(|\partial_x^2 a(t,x)| + |\partial_t a(x,t)|) + |x||\partial_x a(t,x)| + |a(t,x) - I_n| \leq \epsilon \tag{8}$$

where A_j is the dyadic region

$$A_j = \mathbb{R} \times \{2^j \leq |x| \leq 2^{j+1}\}$$

If ϵ is small enough then this precludes the existence of trapped rays, while for arbitrary ϵ it restricts the trapped rays to finitely many dyadic regions.

Because of the reduced coefficient regularity for small x, it seems virtually impossible to control the Hamilton flow and to construct parametrices along incoming rays i.e. which approach the origin. However, the situation improves considerably in the case of outgoing rays. Thus the main part of the article is devoted to an outgoing parametrix construction. This suffices in order to capture the full behaviour of the Schrödinger equation due to the nontrapping assumption, which guarantees that each ray can be split into two parts, one of which is outgoing forward in time while the other is outgoing backward in time.

Our parametrix construction is based on the use of a time dependent FBI transform. However we do not use Sjöstrand's theory [15]. Instead, we take advantage of the simpler approach introduced by the author in [18], [19], [20], [21]; the latter is recommended to the reader as a good starting point. In this analysis the FBI transform is used to turn the equation into a degenerate parabolic evolution in the phase space. Bounds for this evolution are then obtained using the maximum principle.

For more information about phase space transforms we refer to [8] and [5]. One of the main starting points in the phase space analysis of pde's is Fefferman's article [7].

Even though our parametrix is very precise, there are still errors which need to be controlled and this is done using localized energy estimates, otherwise known as local smoothing estimates. We prove such estimates in the case when the parameter ϵ in (8) is sufficiently small. If ϵ is large then nontrapping may fail, and thus the localized energy estimates may fail. With a nontrapping assumption it is likely that the localized energy estimates hold locally in time, but it is not clear what happens globally in time. To avoid being distracted from the main purpose of this paper we have decided to brush aside this problem and simply use the localized energy estimates as an assumption for large ϵ.

Scaling plays an essential role in our analysis. Modulo rescaling and Littlewood-Paley theory all our analysis is reduced to waves which have fixed frequency of size $O(1)$. Such waves have a propagation speed of size $O(1)$, therefore our study of outgiong waves can be largely localized to cones of the form $\{|x| \approx |t|\}$. Certainly the exact flow cannot have a precise localization of this type due to the uncertainty principle. To compensate for this we introduce an artificial damping term which produces rapid decay of waves which do not have the above localization. This allows us to restrict our attention to the above cone modulo rapidly decreasing errors.

Before we state our main results we need to introduce the function spaces for the localized energy estimates. We consider a dyadic partition of unity in frequency,

$$1 = \sum_{k=-\infty}^{\infty} S_k(D)$$

and for each $k \in \mathbb{Z}$ we measure functions of frequency 2^k using the norm

$$\|u\|_{X_k} = 2^k \|u\|_{L^2(A_{<-k})} + 2^{\frac{k}{2}} \sup_{j \geq -k} \|(|x| + 2^{-k})^{-\frac{1}{2}} u\|_{L^2(A_j)}$$

where

$$A_{<j} = \mathbb{R} \times \{|x| \leq 2^j\}$$

To measure the regularity of solutions to the Schrodinger equation we use the global localized energy space X defined by the norm

$$\|u\|_X^2 = \sum_{k=-\infty}^{\infty} \|S_k u\|_{X_k}^2$$

If $n \geq 3$ then one can think of this as a space of distributions, and the following Hardy type inequality holds:

$$\||x|^{-1} u\|_{L^2} \lesssim \|u\|_X \tag{9}$$

This can be used to connect the present work with estimates proved for the flat Laplacian with inverse square potentials, see MR2106340 [4] and references therein.

On the other hand if $n = 1, 2$ one has a BMO type structure, i.e. X is a space of distributions modulo time dependent constants.

For the inhomogeneous term in the equation, on the other hand, we use the dual space X' with norm

$$\|f\|_{X'}^2 = \sum_{k=-\infty}^{\infty} \|S_k f\|_{X_k'}^2$$

Following the discussion above, if $n \geq 3$ then X' is dense in $S'(\mathbb{R}^n)$ and

$$\|u\|_{X'} \lesssim \|xu\|_{L^2}$$

If $n = 1, 2$ then functions in X' must satisfy the cancellation condition

$$\int_{\mathbb{R}}^{n} f(x)dx = 0$$

Definition 1. *We say that the operator P satisfies the localized energy estimates if for each initial data $u_0 \in L^2$ and each inhomogeneous term $f \in L^1L^2 \cap X'$ there exists an unique solution u to (maineq 7) which satisfies the bound*

$$\|u\|_{L^\infty L^2 \cap X} \lesssim \|u_0\|_{L^2} + \|f\|_{L^1L^2 + X'} \tag{10}$$

The localized energy estimates hold under the assumption that the coefficients a^{ij} are a small perturbation of the identity:

Theorem 2. *Assume that the coefficients a^{ij} satisfy (coeff 8) with an ϵ which is sufficiently small.*
12 *Then the operator P satisfies the localized energy estimates globally in time.*

This leads to our main scale invariant Strichartz estimate:

Theorem 3. *Assume that the coefficients a^{ij} satisfy (8) with an ϵ which is sufficiently small. Let (p_1, q_1) and (p_2, q_2) be two Strichartz pairs. Then the solution u to (7) satisfies*

$$\|u\|_{L^{p_1}(L^{q_1})\cap X} \lesssim \|u_0\|_{L^2} + \|f\|_{L^{p_2'}(L^{q_2'})+X'} \tag{11}$$

If ϵ is large then any localized energy estimates require an additional nontrapping condition. Even then the nontrapping can at most guarantee local in time bounds. However, we can still prove a conditional result:

Theorem 4. *a) Let $-\infty \leq T^- < T^+ \leq \infty$. Assume that the coefficients a^{ij} satisfy (8) in (T^-, T^+). Then for every Strichartz pair (p, q) we have*

$$\|u\|_{L^pL^q} \lesssim \|u\|_X + \|Pu\|_{X'} \tag{12}$$

In addition there is a parametrix K for P which satisfies

$$\|Kf\|_{L^{p_1}L^{q_1}\cap X} + \|(PK-I)f\|_{X'} \lesssim \|f\|_{L^{p_2'}(L^{q_2'})} \tag{13}$$

for any two Strichartz pairs (p_1, q_1) and (p_2, q_2).

b) Assume that in addition the operator P satisfies the localized energy estimates in $[T^-, T^+]$. Then the solution u to (7) satisfies the full Strichartz estimates in (11).

Finally, we comment on possible lower order terms in the equation. Suppose we add first and zero order terms to P,

$$P = D_t + D_i a^{ij} D_j + b^i D_i + c$$

Consistent with (8) we introduce the following condition on the coefficients b and c:

$$\sum_{j\in\mathbb{Z}} \sup_{A_j} |x|^2 |\partial_x b(t,x)| + |x||b(t,x)| \leq \epsilon \tag{14}$$

$$\sup_{\mathbb{R}\times\mathbb{R}^n} |x|^2 |c(t,x)| \leq \epsilon \tag{15}$$

Then we have

Remark 5. *If $n \geq 3$ and b, c satisfy (14), (15) then the following estimate holds:*

$$\|(b^i D_i + c)u\|_{X'} \lesssim \epsilon \|u\|_X \tag{16}$$

Consequently the results in Theorems 2, 3, 4 remain valid when such lower order terms are added to P.

If $n = 1, 2$ then (16) cannot hold for any nonzero b, c due to the BMO type structure of X. One can still add lower order terms to the equation but these must have some decay in time if one is to be able to take advantage of the dispersive estimates.

In applications one might be concerned that the condition (8) imposes the nontrivial restriction $a(t, 0) = I_n$. This is true, but it is needed only because we are allowing the derivatives of the coefficients to be singular at 0. Otherwise, such a restriction is unnecessary:

Remark 6. *Assume that the the condition (8) on the coefficients a^{ij} is modified for $|x| < 1$ to*

$$\sup_{|x|<1} (|\partial_x^2 a(t,x)| + |\partial_t a(x,t)|) + |\partial_x a(t,x)| + |a(t,x) - I_n| \leq \epsilon$$

Assume also that for $k > 0$ the definition of the space X_k is changed to

$$\|u\|_{X_k} = 2^{\frac{k}{2}} \|u\|_{L^2(A_{<-0})} + 2^{\frac{k}{2}} \sup_{j\geq 0} \|(|x| + 2^{-k})^{-\frac{1}{2}} u\|_{L^2(A_j)}$$

Then the results in Theorems 2,3,4 remain valid. Their proofs are essentially identical with only a few obvious changes.

The paper is structured as follows. After introducing some notations in the next section, in Section 3 we consider the paradifferential calculus associated to our problem. More precisely, we show that without any loss we are allowed to mollify the coefficients a^{ij} on a suitable x dependent scale. We also prove the bound (16) for the lower order terms. This allows us to reduce our analysis to problems which are frequency localized in dyadic regions.

Section 4 contains the proof of the localized energy estimates in Theorem 2. The main step of the proof is carried out in a frequency localized context and involves a Morawetz type mutiplier technique.

In Section 5 we state our main result on the existence of frequency localized outgoing parametrices, namely Proposition 10. Using this result we conclude the proof of Theorems 3,4.

The rest of the paper is devoted to the parametrix construction. In Section 6 we introduce the pseudodifferential operators and the phase space transforms. An important role is played by the results concerning the conjugation of pdo's with respect to the phase space transform, for which we use some results from [20], [21]. The parametrix is first constructed in Section 7 in the case of evolutions governed by a pseudodifferential operator a^w whose symbol satisfies a suitable smallness condition uniformly in x. This construction is then transferred in Section 8 to small perturbations of Δ via conjugation with respect to the flat Schroedinger flow. Finally to arrive at the desired setup we need to insure that the parametrix is localized in outgoing propagation cones. This is done in the last section by means of choosing a suitable damping term in the equation.

2. Notations

notations

We consider a smooth spatial Littlewood-Paley decomposition

$$1 = \sum_{j=-\infty}^{\infty} \chi_j(x) \qquad \text{supp } \chi_j \subset \{2^{j-1} < |x| < 2^{j+1}\}$$

We also set

$$\chi_{<j} = \sum_{k<j} \chi_k$$

Given ϵ as in (coeff 8) we can find a sequence $\epsilon_j \in l^1$ so that

ej
$$\sup_{A_j} |x|^2 |\partial_x^2 a(t,x)| + |x||\partial_x a(t,x)| + |a(t,x) - I_n| \leq \epsilon_j \tag{17}$$

and

$$\sum \epsilon_j \lesssim \epsilon$$

Without any restriction in generality we can assume that ϵ_j is slowly varying, say

$$|\ln \epsilon_j - \ln \epsilon_{j-1}| \leq 2^{-10}$$

We also choose a function ϵ with the property that

$$\epsilon_j < \epsilon(s) < 2\epsilon_j \qquad 2^j < s < 2^{j+1}$$

and so that

$$\epsilon_j'(s) \leq 2^{-5} s^{-1} \epsilon_j(s)$$

This implies that

$$\int_{\mathbb{R}} \frac{\epsilon(s)}{s} \approx \epsilon$$

We define the bounded function $e(s)$ by

$$e(s) = \epsilon^{-1} \int_{-\infty}^{s} \frac{\epsilon(\sigma)}{\sigma} d\sigma$$

We consider a frequency Littlewood-Paley decomposition

$$1 = \sum_{j=-\infty}^{\infty} S_j(D)$$

where

$$\text{supp } s_j \subset \{2^{j-1} < |\xi| < 2^{j+1}\}$$

We also use the related notations $S_{<k}$, $S_{>k}$, etc.

We say that a function f is localized at frequency 2^k if $\hat{f}$ is supported in $\{2^{j-1} < |\xi| < 2^{j+1}\}$. An operator K is localized at frequency 2^k if for any f localized at frequency 2^k its image Kf is frequency localized in $\{2^{j-10} < |\xi| < 2^{j+10}\}$.

para

3. The paradifferential calculus

In order to reduce the problem to a frequency localized context and to simplify the parametrix construction it is convenient to localize the coefficients in frequency. This is somewhat more complicated than usual because the frequency localization scale needs to depend on the spatial scale.

Given a frequency scale k we define the regularized coefficients

$$a^{ij}_{(k)} = + \sum_{l<k-4} S_{<l}\chi_{<k-2l} S_l a^{ij}$$

Correspondingly we define the mollified operators

$$A_{(k)} = D_i a^{ij}_{(k)} D_j$$

which are used on functions of frequency 2^k. Roughly speaking, their coefficients are frequency localized in the region

$$|\xi| \ll 2^k (1 + 2^k|x|)^{-\frac{1}{2}}$$

We also introduce a global mollified operator

$$\tilde{A} = \sum_{k=-\infty}^{\infty} A_{(k)} S_k$$

Due to (17) and to the fact that the e_j's are slowly varying it follows that the dyadic parts of the coefficients will satisfy the bounds

aijl
$$|S_l a^{ij}(t,x)| \lesssim \left\{ \begin{array}{ll} 2^{-2m-2l}\epsilon_m & 2^m < |x| < 2^{m+1},\ m+l \geq 0 \\ \epsilon_{-l} & |x| < 2^{-l} \end{array} \right. \tag{18}$$

This also allows us to obtain bounds on the coefficients of $A_{(k)}$,

coeffak
$$\begin{aligned} |\partial^\alpha a^{ij}_{(k)}(x)| &\leq c_\alpha \epsilon 2^{|\alpha| k}(1 + 2^k|x|)^{-\alpha} \qquad |\alpha| \leq 2 \\ |\partial^\alpha a^{ij}_{(k)}(x)| &\leq c_\alpha \epsilon 2^{|\alpha| k}(1 + 2^k|x|)^{-1-\frac{|\alpha|}{2}} \qquad |\alpha| \geq 2 \end{aligned} \tag{19}$$

The main result of this section shows that we can freely replace A by $\tilde{A}$ in Theorems 2, 4(a). It also shows that at frequency 2^k the operators $\tilde{A}$ and $A_{(k)}$ are interchangeable.

Proposition 7. *Assume that the coefficients a^{ij} satisfy (8). Then the following estimates hold:*

amta
$$\|(A - \tilde{A})u\|_{X'} \lesssim \epsilon \|u\|_X \tag{20}$$

tamak (21) $$\|(\tilde{A}-A_{(k)})S_l u\|_{X_k'} \lesssim \epsilon \|S_l u\|_{X_k}, \qquad |l-k|\leq 2$$

comak (22) $$\|[A_{(k)}, S_k]u\|_{X_k'} \lesssim \epsilon \|u\|_{X_k}$$

Proof. To prove (amta 20) we write the difference $A-\tilde{A}$ in the form

$$A-\tilde{A} = A_{low} + A_{mid} + A_{high}$$

where

$$A_{low} = \sum_{k=-\infty}^{\infty} D_i \left(\sum_{l<k-4} S_{<l}\chi_{>k-2l} S_l a^{ij} \right) D_j S_k$$

$$A_{mid} = \sum_{k=-\infty}^{\infty} \sum_{l=k-4}^{k+4} D_i (S_l a^{ij}) D_j S_k$$

$$A_{high} = \sum_{k=-\infty}^{\infty} \sum_{l>k+4} D_i (S_l a^{ij}) D_j S_k$$

The operator A_{low} keeps the frequency localization, i.e. takes frequencies of size 2^k into frequencies of size 2^k. The operators D_i and D_j yield 2^k factors, so we only need to show that

dplow (23) $$\left\| \sum_{l<k-4} S_{<l}\chi_{>k-2l} S_l a^{ij} v \right\|_{X_k'} \lesssim \epsilon 2^{-2k} \|v\|_{X_k}$$

For the mollified cutoff we use the trivial bound

$$|S_{<l}\chi_{>k-2l}(x)| \leq \begin{cases} 2^{4l-4k} & |x| < 2^{k-2l-2} \\ 1 & |x| \geq 2^{k-2l-2} \end{cases}$$

Suppose that $2^m < |x| < 2^{m+1}$. For $m \geq -k$ we combine this with (aijl 18) to obtain

$$\left| \sum_{l<k-4} S_{<l}\chi_{>k-2l} S_l a^{ij} \right| \lesssim \sum_{l=-\infty}^{-m-1} 2^{4l-4k}\epsilon_{-l} + \sum_{l=-m}^{\frac{k-m}{2}-1} \epsilon_m 2^{-2m-2l} 2^{4l-4k} + \sum_{l=\frac{k-m}{2}}^{m-4} \epsilon_m 2^{-2m-2l}$$
$$\lesssim 2^{-m-k}\epsilon_m$$

A similar argument (or the uncertainty principle) shows also that

$$\left| \sum_{l<k-4} S_{<l}\chi_{>k-2l} S_l a^{ij} \right| \lesssim \epsilon_{-k} \qquad |x| < 2^{-k}$$

Together the last two bounds imply (dplow 23).

In the case of A_{mid} we can have output at all frequencies less than or equal to 2^k. To simplify the notations we take $l=k$. Then it suffices to obtain the off-diagonal decay

$$\|S_l D_i (S_k a^{ij} D_j S_k v)\|_{X_l'} \lesssim \epsilon 2^{\frac{l-k}{2}} \|v\|_{X_k}, \qquad l \leq k+2$$

which reduces to

$$\|S_l(S_k a^{ij} v)\|_{X_l'} \lesssim 2^{-\frac{l}{2}-\frac{3k}{2}} \epsilon \|v\|_{X_k}$$

We make a dyadic decomposition with respect to the spatial variable,

$$S_l(S_k a^{ij} v) = S_l(\chi_{<-k} S_k a^{ij} v) + \sum_{m=-k}^{\infty} S_l(\chi_m S_k a^{ij} v)$$

Then it suffices to prove that for $m > -k$ we have

$$\|S_l(\chi_m S_k a^{ij}) v)\|_{X_l'} \lesssim 2^{-\frac{l}{2}-\frac{3k}{2}} \epsilon_m \|v\|_{X_k}$$

with the obvious modification when $m = k$. We freeze the weights using (aijl 18) and the definition of X_k. Then the above bound reduces to

lmk
$$\|S_l(\chi_m v)\|_{X_l'} \lesssim 2^{-\frac{l}{2}+k+\frac{3m}{2}} \|v\|_{L^2} \tag{24}$$

For this we consider two cases. If $m + l \geq 0$ then

$$\|S_l(\chi_m v)\|_{L^2} \lesssim \|v\|_{L^2}, \qquad \|x S_l(\chi_m v)\|_{L^2} \lesssim 2^m \|v\|_{L^2}$$

Hence interpolating

fs
$$\|S_l(\chi_m v)\|_{X_l'} \lesssim 2^{\frac{m-l}{2}} \tag{25}$$

which suffices for (lmk 24) since

$$\frac{m-l}{2} \leq -\frac{l}{2} + k + \frac{3m}{2}$$

In the second case $m + l < 0$ we have a better L^2 bound

$$\|S_l(\chi_m v)\|_{L^2} \lesssim 2^{\frac{n(m+l)}{2}} \|v\|_{L^2}, \qquad \|x S_l(\chi_m v)\|_{L^2} \lesssim 2^{-l} 2^{\frac{n(m+l)}{2}} \|v\|_{L^2}$$

Interpolating and taking the worst dimension $n = 1$ we obtain again (fs 25) and conclude as before.

Finally we consider the contribution of A_{high}, for which we obtain again an off-diagonal decay

$$\|S_l D_i (S_l a^{ij} D_j S_k v)\|_{X_l'} \lesssim \epsilon 2^{\frac{k-l}{2}} \|v\|_{X_k}$$

We can replace D_i and D_j with factors of 2^l, respectively 2^k and then drop the multipliers to reduce this to

$$\|S_l a^{ij} v\|_{X_l'} \lesssim \epsilon 2^{-\frac{k}{2}-\frac{3l}{2}} \|v\|_{X_k}$$

Localizing to dyadic spatial regions this follows from

$$\|\chi_m S_l a^{ij}\|_{L^\infty} \lesssim \epsilon_m 2^{-m} 2^{-l} \qquad m \geq -l$$

and

$$\|\chi_{<-l} S_l a^{ij}\|_{L^\infty} \lesssim \epsilon_{-l}$$

which are in turn consequences of (18). This concludes the proof of (20).

The proof of (21) is similar but simpler, as only terms as in A_{low} can occur. Finally, for (22) we can factor out the derivatives and reduce it to

$$\|[S_k, a^{ij}_{(k)}]v\|_{X_k'} \lesssim \epsilon 2^{-2k}\|v\|_{X_k}$$

Then this follows directly from (19) with $\alpha = 1$.

$\square$

In a similar manner we prove the bound (16) which shows that in high dimension we can completely dispense with lower order terms.

Proof of (16). As a consequence of (14), (15) we note the following bounds on dyadic pieces of the coefficients:

$$\begin{aligned} |S_k b(x)| &\lesssim 2^{-k}\epsilon(2^{-k}+|x|)(2^{-k}+|x|)^{-2}, \quad |S_{<k} b(x)| \lesssim 2^{-k}\epsilon(2^{-k}+|x|)(2^{-k}+|x|)^{-2} \\ |S_k c(x)| &\lesssim (2^{-k}+|x|)^{-2}, \qquad |S_{<k} c(x)| \lesssim (2^{-k}+|x|)^{-2} \end{aligned} \tag{26}$$

To prove (16) we first do a dyadic decomposition of the two factors,

$$(b^i\partial_i + c)u = \sum_{k>j+4}(S_k b^i\partial_i + S_k c)S_j u + \sum_{|k-j|\le 4}(S_k b^i\partial_i + S_k c)S_j u + \sum_j (S_{<j-4} b^i\partial_i + S_{<j-4} c)S_j u$$

We consider the three cases. Begin with the high-low interactions in the first sum. The output is at frequency 2^k so we measure it in X_k'. The derivative yelds at most a 2^k factor. Using also (26) we compute

$$\begin{aligned} \|(S_k b^i\partial_i + S_k c)S_j u\|_{X_k'} &\lesssim \|(2^{-k}+|x|)^{-2} S_j u\|_{X_k'} \\ &\lesssim \|\chi_{>-j}(2^{-j}+|x|)^{-2} S_j u\|_{X_k'} + \|\chi_{\le j}(2^{-k}+|x|)^{-2} S_j u\|_{X_k'} \\ &\lesssim 2^{\frac{j-k}{2}}\|S_j u\|_{X_j} + \|\chi_{\le j}(2^{-k}+|x|)^{-2}\|_{X_k'}\|S_j u\|_{L^\infty(A_{<j})} \\ &\lesssim \|S_j u\|_{X_j}(2^{\frac{j-k}{2}} + 2^{\frac{(n-2)j}{2}}\|\chi_{\le j}(2^{-k}+|x|)^{-2}\|_{X_k'}) \end{aligned}$$

The outcome of the last computation depends on the dimension. If $n > 3$ then we get the constant $2^{\frac{j-k}{2}}$ which decays away from the diagonal so it suffices for the summation with respect to j,k. If $n = 3$ we have an additional logarithmic loss so we obtain the slightly weaker constant $|k-j|2^{\frac{j-k}{2}}$.

The second case corresponds to high-high frequency interactions, and the output must be at the same frequency or lower. This is dual to the first case.

Finally we consider the last case, of low-high interactions. Here the output is at frequency 2^j, therefore the square summability of the sum with respect to j is inherited from $S_j u$.

Hence it suffices to estimate

$$\|(S_{<j-4}b^i\partial_i + S_{<j-4}c)S_j u\|_{X_j'} \lesssim \|(2^{-j}+|x|)^{-2}S_j u\|_{X_j'} \lesssim \|S_j u\|_{X_j}$$

□

mora

4. Localized energy estimates

Here we prove Theorem l2 2. This is done via a positive commutator method. Let $(\alpha_m)_{m\in\mathbb{Z}}$ be a positive slowly varying sequence with $\sum \alpha_k = 1$. Correspondingly we define the space $X_{k,\alpha}$ with norm

$$\|u\|^2_{X_{k,\alpha}} = \sum_{j>-k} \alpha_j \|(|x|+2^{-k})^{-\frac{1}{2}}u\|_{L^2(A_j)}$$

and the dual space

$$\|u\|^2_{X_{k,\alpha}} = \sum_{j>-k} \alpha_j^{-1} \|(|x|+2^{-k})^{\frac{1}{2}}u\|_{L^2(A_j)}$$

The key step in the proof of Theorem l2 2 is the following frequency localized estimate:

Proposition 8. *Assume that ϵ is sufficiently small. Then the bound*

locl2 (27)
$$\|u\|_{L^\infty L^2\cap X_{k,\alpha}} \lesssim \|u(0)\|_{L^2} + \|(D_t + A_{(k)})u\|_{L^1L^2 + X'_{k,\alpha}}$$

holds for all functions $u \in L^\infty L^2 \cap X_{k,\alpha}$ localized at frequency 2^k uniformly with respect to all slowly varying sequences (α_m) with

suma (28)
$$\sum_{k=-\infty}^{\infty} \alpha_m = 1$$

l2loc

Proof. Let Q be an L^2 bounded selfadjoint operator in $\mathbb{R}^n$. Then the operator

$$C = i[A_{(k)}, Q]$$

is also selfadjoint and we have

$$2Im\langle A_{(k)}u, Qu\rangle = \langle Cu, u\rangle.$$

For the Schrodinger equation we obtain

$$\frac{d}{dt}\langle u, Qu\rangle = -2\Im\langle (D_t + A_{(k)})u, Qu\rangle + \langle Cu, u\rangle$$

When $Q = I$ this gives the energy estimate

$$\frac{d}{dt}\|u\|^2_{L^2} = -2\Im\langle (D_t + A_{(k)})u, u\rangle$$

If δ is a small parameter then the modified energy

$$E(u) = \|u\|^2_{L^2} - \delta\langle u, Qu\rangle$$

is positive definite and satisfies

$$\frac{d}{dt}E(u) = 2\Im\langle (D_t + A_{(k)})u, (1-\delta Q)u\rangle - \delta\langle Cu, u\rangle$$

Integrating in time we obtain

$$\|u\|^2_{L^\infty L^2} + \delta\langle Cu, u\rangle \lesssim \|u(0)\|^2_{L^2} + \|(D_t + A_{(k)})u\|_{L^1L^2 + X'_{k,\alpha}} \|(1-\delta Q)u\|_{L^\infty L^2 \cap X_{k,\alpha}}$$

Then the conclusion of the proposition follows from the Cauchy-Schwartz inequality and the next lemma:

Lemma 9. *For each $k \in \mathbb{Z}$ and each slowly varying sequence (α_m) satisfying (suma 28) there exist an L^2 bounded selfadjoint operator Q so that*

$$\langle Cu, u\rangle \gtrsim \|u\|^2_{X_{k,\alpha}} \tag{29}$$

$$\|Qu\|_{X_{k,\alpha}} \lesssim \|u\|_{X_{k,\alpha}} \tag{30}$$

for all functions u localized at frequency 2^k.

We now prove the lemma. By rescaling we can assume that $k = 0$. Let δ be a small parameter which will be chosen later. We increase the sequence (α_m) so that it remains slowly varying and the following properties hold:

aam
$$\begin{cases} \alpha_m = 1 \qquad \text{for } \ m \leq 0 \\ \displaystyle\sum_{m>0} \alpha_m \approx 1 \\ \epsilon_m \leq \delta\alpha_{m+\log_2\delta} \end{cases} \tag{31}$$

The last property requires $\epsilon < \delta$, and (after δ is fixed) determines the allowable range for the parameter ϵ in Theorem 12 2. To the sequence α_m we associate a slowly varying function α with the property that

$$\alpha(s) \approx \alpha_m, \qquad s \approx 2^m$$

With this notation the last part of (aam 31) can be rewritten in the form

ed
$$\epsilon(s) \lesssim \delta\alpha(\delta s) \tag{32}$$

We consider an even function ϕ with the following properties:

(i) $\phi(s) \approx (1+s)^{-1}$ for $s > 0$.

(ii) $\phi(s) + s\phi'(s) \approx \alpha(s)$ for $s > 0$.

(iii) $\phi(|x|)$ is localized at frequency at most $O(1)$.

The construction of r satisfying (i) and (ii) is easy and is left to the reader. The condition (iii) is obtained simply by truncating the function $\phi(|x|)$ in frequency on a sufficiently large scale so that the first two properties are not affected.

The operator Q is the differential operator

$$Q(x, D) = \delta(Dx\phi(\delta|x|) + \phi(\delta|x|)xD)$$

The choice of the function ϕ insures that Q is L^2 bounded at frequency 1,

$$\|QS_0\|_{L^2 \to L^2} \lesssim 1, \qquad |k| \leq 4$$

Since the weight describing the $X_{k,\alpha}$ norm is slowly varying on the unit spatial scale, we also obtain the boundedness

$$\|QS_0\|_{X_{0,\alpha} \to X_{0,\alpha}} \lesssim 1, \qquad |k| \leq 4$$

It remains to compute the commutator

$$C = i[A_{(1)}, Q]$$

We have

$$\begin{aligned} i[A_{(0)}, Q] = {} & 4\delta D_i \phi(\delta|x|) a^{ij}_{(0)} D_j + 2\delta(Dx\delta x^i |x|^{-1} \phi'(\delta|x|) a^{ij}_{(0)} D_j + D_i \delta x^j |x|^{-1} \phi'(\delta|x|) a^{ij}_{(0)} xD) \\ & + \delta D_i \phi(\delta|x|)(x_k \partial_k a^{ij}_{(0)}) D_j + \partial_i (a^{ij}_{(0)} (\partial_j \partial(\delta x \phi(\delta|x|)))) \end{aligned}$$

Given the regularity of the coefficients a^{ij}, the last two terms are small on frequency 1 functions. Precisely,

$$|\delta\phi(\delta|x|)(x_k \partial_k a^{ij}_{(0)})| \lesssim \frac{\delta^2 \alpha(\delta|x|)}{1 + \delta|x|}$$

and

$$|\partial_i (a^{ij}_{(0)} (\partial_j \partial(\delta x \phi(\delta|x|)))| \lesssim \frac{\delta^3}{(1 + \delta|x|)^3}$$

therefore for functions u localized at frequency 1 we have the bounds

$$\langle \delta D_i \phi(\delta|x|)(x_k \partial_k a^{ij}_{(0)}) D_j u, u \rangle \lesssim \delta \langle \frac{\delta\alpha(\delta|x|)}{1 + \delta|x|} u, u \rangle$$

respectively

$$\langle \partial_i (a^{ij}_{(0)} (\partial_j \partial \delta x \phi(\delta|x|)) u, u \rangle \lesssim \delta \langle \frac{\delta\alpha(\delta|x|)}{1 + \delta|x|} u, u \rangle$$

It remains to show that the contribution of the first two terms is large. By (32) we can replace the matrix a^{ij} by I_n at the expense of an error similar to the last right hand side above. Then we are left with the expression

$$C_0 = 4\delta D_i \phi(\delta|x|) D_i + 4\delta D \frac{x}{|x|} \delta|x| \phi'(\delta|x|) \frac{x}{|x|} D$$

which satisfies

$$\langle C_0 u, u\rangle \geq 4\delta \langle (\phi(\delta|x|) + \delta|x|\phi'(\delta|x|))\nabla u, \nabla u\rangle \approx \langle \frac{\delta\alpha(\delta|x|)}{1+\delta|x|} u, u\rangle$$

for all u localized at frequency 1. If δ is small enough this dominates all the error terms, therefore we obtain

$$\langle Cu, u\rangle \gtrsim \langle \frac{\delta\alpha(\delta|x|)}{1+\delta|x|} u, u\rangle$$

This concludes the proof of the lemma, and also the proof of Proposition 8. $\square$

We conclude now the proof of Theorem 2. Let (β_m) be another slowly varying sequence with

$$\sum_m \beta_m = 1$$

Applying Proposition 8 with α_m replaced by $\alpha_m + \beta_m$ we obtain the bound

$$\|u\|_{L^\infty L^2 \cap X_{k,\alpha+\beta}} \lesssim \|u(0)\|_{L^2} + \|(D_t + A_{(k)})u\|_{L^1 L^2 + X'_{k,\alpha+\beta}}$$

for all u localized at frequency 2^k. This implies the weaker estimate

$$\|u\|_{L^\infty L^2 \cap X_{k,\alpha}} \lesssim \|u(0)\|_{L^2} + \|(D_t + A_{(k)})u\|_{L^1 L^2 + X'_{k,\beta}}$$

Since any l^1 sequence is dominated by a slowly varying l^1 sequence, we can drop the assumption that α and β are slowly varying. Then we maximize the left hand side with respect to $\alpha \in l^1$ and minimize the right hand side with respect to $\beta \in l^1$. This yields

$$\|u\|_{L^\infty L^2 \cap X_k} \lesssim \|u(0)\|_{L^2} + \|(D_t + A_{(k)})u\|_{X'_k + L^1 L^2} \tag{33}$$

For an arbitrary function $u \in X'$ we apply this bound to $S_k u$. We have

$$(D_t + A_{(k)})S_k u = S_k(D_t + \tilde{A})u + [A_{(k)}, S_k]u + S_k(A_{(k)} - \tilde{A})$$

The last two terms are frequency localized and can be estimated by (21) and (22),

$$\|[A_{(k)}, S_k]u + S_k(A_{(k)} - \tilde{A})u\|_{X'_k} \lesssim \epsilon \sum_{|k-l|\leq 2} \|S_l u\|_{X_l}$$

Then after summation we obtain

$$\begin{aligned}
\|u\|^2_{L^\infty L^2 \cap X} &\lesssim \sum_k \|S_k u\|^2_{L^\infty L^2 \cap X_k} \\
&\lesssim \sum_k \|D_t + A_{(k)}) S_k u\|^2_{L^1 L^2 + X_k'} + \|S_k u(0)\|^2_{L^2} \\
&\lesssim \|u(0)\|^2_{L^2} + \sum_k \|S_k (D_t + \tilde{A}) u\|^2_{L^1 L^2 + X_k'} + \|[A_{(k)}, S_k] u + S_k (A_{(k)} - \tilde{A}) u\|^2_{X_k'} \\
&\lesssim \|u(0)\|^2_{L^2} + \|(D_t + \tilde{A}) u\|^2_{L^1 L^2 + X'} + \epsilon \|u\|^2_X \\
&\lesssim \|u(0)\|^2_{L^2} + \|(D_t + A) u\|^2_{L^1 L^2 + X'} + \epsilon \|u\|^2_X
\end{aligned}$$

For small ϵ we can neglect the last right hand side term to obtain

dualbd
$$\|u\|^2_{L^\infty L^2 \cap X} \lesssim \|u(0)\|_{L^2} + \|(D_t + A) u\|^2_{L^1 L^2 + X'} \tag{34}$$

which holds in any time interval containing 0.

Since the operator $D_t + A$ is selfadjoint, by duality this shows that for any $f \in L^1 L^2 + X'$ there is a solution v to

$$(D_t + A) v = f, \qquad v(0) = v_0$$

with

$$\|v\|_{X \cap L^\infty L^2} \lesssim \|f\|_{L^1 L^2 + X'} + \|v(0)\|_{L^2}$$

By (34) this solution is unique. The proof of Theorem 2 is concluded.

tts
5. Parametrices and Strichartz estimates

Here we reduce the proof of Theorem 3 to the construction of a suitable parametrix for $D_t + A_{(0)}$. Our main result concerning parametrices is

Proposition 10. *Assume that ϵ is sufficiently small. Then there is a parametrix K_0 for $D_t + A_{(0)}$ which is localized at frequency 1 and has the following properties:*

(i) L^2 bound:

ees
$$\|K_0(t,s)\|_{L^2 \to L^2} \lesssim 1 \tag{35}$$

(ii) Error estimate:

l2error
$$\|(1+|x|)^N (D_t + A_{(0)}) K_0(t,s)\|_{L^2 \to L^2} \lesssim (1 + |t-s|)^{-N} \tag{36}$$

(iii) Jump condition: $K_0(s+0,s)$ and $K_0(s-0,s)$ are $S^0_{1,0}$ type pseudodifferential operators satisfying

$$(K_0(s+0,s) - K_0(s-0,s)) S_0 = S_0$$

(iv) Outgoing parametrix:

outl2
$$\|1_{\{|x|<2^{-10}|t-s|\}} K_0(t,s)\|_{L^2\to L^2} \lesssim (1+|t-s|)^{-N} \tag{37}$$

(v) Pointwise decay:

$$\|K_0(t,s)\|_{L^1\to L^\infty} \lesssim (1+|t-s|)^{-\frac{n}{2}} \tag{38}$$

K0

We leave the proof of this result for later sections, and we show that it implies Theorems 3,4. As an intermediate step we prove the following localized Strichartz estimates for the parametrix:

Proposition 11. *The parametrix K_0 given by Proposition 10 has the following properties:*

(i) (regularity) For any Strichartz pairs (p_1,q_1) respectively (p_2,q_2) with $q_1 \le q_2$ we have

lpbd
$$\|K_0 f\|_{L^{p_1}L^{q_1}\cap X_0} \lesssim \|f\|_{L^{p_2'}L^{q_2'}} \tag{39}$$

(ii) (error estimate) For any Strichartz pair (p,q) we have

lperror
$$\|[(D_t + A_{(0)})K_0 - 1]f\|_{X_0'} \lesssim \|f\|_{L^{p'}L^{q'}} \tag{40}$$

k0lp

Proof. By interpolation we obtain the bound

$$\|K_0(t,s)\|_{L^{q'}\to L^q} \lesssim (1+|t-s|)^{-\frac{n}{q}}, \qquad 2 \le q \le \infty$$

By the Hardy-Littlewood-Sobolev inequality this implies that

ppp
$$\|K_0\|_{L^{p'}L^{q'}\to L^pL^q} \lesssim 1 \tag{41}$$

for all Strichartz pairs (p,q) with $q>2$. The case $q=2$ is handled as in [11].

We obtain the similar estimate with different Strichartz pairs (p_1,q_1) respectively (p_2,q_2) by interpolating with $L^{p'}L^{q'} \to L^2$ bounds, namely

lp2
$$\|K_0(t,\cdot)\|_{L^{p'}L^{q'}\to L^2} \lesssim 1 \tag{42}$$

By the TT^* argument this is equivalent to

$$\|K_0^*(\cdot,t)K_0(t,\cdot)\|_{L^{p'}L^{q'}\to L^pL^q} \lesssim 1$$

This follows from (41), Sobolev embeddings and the following Lemma:

Lemma 12. *The parametrix K_0 given by Proposition 10 satisfies*

$$\|K_0^*(s_1,t)K_0(t,s_2) - K_0^*(s_1+0,s_1)K_0(s_1,s_2)\|_{L^2\to L^2} \lesssim (1+|s_1-s_2|)^{-N} \qquad t > s_1 > s_2$$

Proof. Compute

$$D_t(K_0^*(s_1,t)K_0(t,s_2)) = -[(D_t + A_{(0)})K_0(t,s_1)]^*K_0(t,s_2) + K_0^*(s_1,t)[(D_t + A_{(0)})K_0(t,s_2)]$$

In the first term we use (36) for the first factor and the energy estimate (35) combined with (37) for the second to obtain

$$\begin{aligned}\|[(D_t + A_{(0)})K_0(t,s_1)]^*K_0(t,s_2)\|_{L^2\to L^2} \lesssim & \|(1+|x|)^N(D_t + A_{(0)})K_0(t,s_1)\|_{L^2\to L^2} \\ & \|(1+|x|)^{-N}K_0(t,s_2)\|_{L^2\to L^2} \\ \lesssim & (1+|t-s_1|)^{-N}(1+|t-s_2|)^{-N}\end{aligned}$$

The second term is estimated in a similar manner. Hence we obtain

$$\|D_t(K_0(s_1,t)^*K_0(t,s_2))\|_{L^2\to L^2} \lesssim (1+|t-s_1|)^{-N}(1+|t-s_2|)^{-N}$$

We integrate this relation between s_1 and t to obtain the conclusion of the Lemma. □

The $L^{p'}L^{q'} \to X_0$ bound can be rewritten in the form

$$\|1_{\{|x|<R\}}K_0\|_{L^{p'}L^{q'}\to L^2} \lesssim |R|^{\frac{1}{2}}$$

To prove this we split

$$K_0(t,s) = 1_{\{|t-s|<2^{10}R\}}K_0(t,s) + 1_{\{|t-s|>2^{10}R\}}K_0(t,s)$$

For the first part we use (42) and Holder's inequality, while for the second part we use the rapid decay in (37) combined with Sobolev embeddings.

Finally, the error estimates in (40) follow easily from the ones in (36).

□

Proposition 11 is useful only if ϵ is small. However, we can prove a similar result even if ϵ is not small:

Proposition 13. *Assume that the coefficients a^{ij} satisfy (8) in a time interval $[T^-,T^+]$. Then there is a parametrix K_0 for $A_{(0)}$ localized at frequency 1 and which satisfies*

(i) (regularity) For any Strichartz pairs (p_1,q_1) respectively (p_2,q_2) with $q_1 \le q_2$ we have

lpbdl (43)
$$\|K_0 f\|_{L^{p_1}L^{q_1}\cap X_0} \lesssim \|f\|_{L^{p_2'}L^{q_2'}}$$

(ii) (error estimate) For any Strichartz pair (p,q) we have

lperrorl (44)
$$\|[(D_t + A_{(0)})K_0 - 1]f\|_{X_0'} \lesssim \|f\|_{L^{p'}L^{q'}}$$

k0lplarge

Proof. If either T^- or T^+ is finite then we extend the coefficients to all of $\mathbb{R} \times \mathbb{R}^n$. This can be done for instance using reflections in time. Hence in what follows we assume that $T^- = -\infty$, $T^+ = \infty$.

Let δ be a small constant chosen so that Proposition 10 applies for $\epsilon \lesssim \delta$. We consider the slowly varying sequence ϵ_j attached to (8) as in (17). Since in $A_{(0)}$ the coefficients are truncated at frequency less than 1, we only need the ϵ_j's for $j \geq 0$. We partition the set $\mathbb{N}$ of indices into intervals

$$N = \bigcup_{j \in J} I_j$$

so that for each k either of the following two properties holds:

(a) (intervals of first kind)

$$\sum_{k \in I_j} \epsilon_k \lesssim \delta \tag{45}$$

(b) (intervals of the second kind)

$$|I_j| = 1$$

We only need finitely many such intervals,

$$|J| \lesssim \frac{\epsilon}{\delta}$$

Given an interval $I = [j,k] \subset \mathbb{N}$ we denote $\tilde{I} = [j-2, k+2]$. We also define an associated dyadic cutoff function

$$\chi_I(x) = \sum_{i \in I} \chi_i(x)$$

In order to keep the frequency localization we assume that these cutoff functions are frequency localized at frequency $\ll 1$. Hence they will have some tails in other dyadic regions. However, these tails are rapidly decreasing away from the initial support. Thus they are harmless, and will be neglected in the sequel.

We seek the parametrix K_0 of the form

$$K_0 = \sum_{j \in J} \chi_{\tilde{I}_j} K^{I_j} \chi_{I_j}$$

where each of the terms is essentially localized to $\tilde{I}_j$. Since the above sum has finitely many terms, it suffices to construct the K^I's so that each of these terms satisfies the estimates

lpii
$$\| \chi_{\tilde{I}_j} K^{I_j} \chi_{I_j} \|_{L^{p_1} L^{q_1} \cap X_0} \lesssim \|f\|_{L^{p_2'} L^{q_2'}} \tag{46}$$

respectively

eele
$$\| ((D_t + A_{(0)}) \chi_{\tilde{I}_j} K^{I_j} \chi_{I_j} - \chi_{I_j}) f \|_{X_0'} \lesssim \|f\|_{L^{p'} L^{q'}} \tag{47}$$

We consider the two cases as above. The index j for the interval is ommited in the sequel.

a) Suppose I is an interval of the first kind. We restrict the coefficients a of A to the annulus $\{|x| \in 2^{\tilde I}\}$. Then we consider an extension a^I to $\mathbb{R} \times \mathbb{R}^n$ which satisfies (8) with $\epsilon \lesssim \delta$.

For this extension we let K^I be the parametrix given by Proposition 10. By Proposition 11 we can estimate

$$\|\chi_{\tilde I_j} K^{I_j} \chi_{I_j} f\|_{L^{p_1}L^{q_1}\cap X_0} \lesssim \|K^{I_j}\chi_{I_j} f\|_{L^{p_1}L^{q_1}\cap X_0} \lesssim \|\chi_{I_j} f\|_{L^{p_2'}L^{q_2'}} \lesssim \|f\|_{L^{p_2'}L^{q_2'}}$$

which gives the corresponding part of (43).

For the error estimate (47) we write

$$(D_t + A_{(0)})\chi_{\tilde I} K^I \chi_I - \chi_I = \chi_{\tilde I}((D_t + A^I_{(0)})K^I - 1)\chi_I + (A_{(0)} - A^I_{(0)})\chi_{\tilde I} K^I \chi_I + [A^I_{(0)}, \chi_{\tilde I}] K^I \chi_I$$

The first term is bounded directly using (40) for K^I. For the second we can neglect the derivatives due to the frequency 1 localization. If it were not for the frequency localization, the functions $a - a^I$ and $\chi_{\tilde I}$ would have supports with dyadic separation. Given the frequency localization, there is minimal overlapping,

$$|\partial^\alpha (a_{(0)} - a^I_{(0)})\partial^\beta \chi_{\tilde I}| \lesssim (1+|x|)^{-N}$$

This combined with the X_0 bound in (39) gives the estimate for the second term. Finally, for the third term we use again (39); then it remains to prove the commutator estimate

$$\|[A^I_{(0)}, \chi_{\tilde I}] v\|_{X_0'} \lesssim \|v\|_{X_0}$$

where v is localized at frequency 1. But this is simply a consequence of the fact that the derivatives of $\chi_{\tilde I}$ are essentially concentrated in two dyadic spatial regions and satisfy the bound

$$|\nabla \chi_{\tilde I}(x)| \lesssim (1+|x|)^{-1}$$

b) Suppose that I is an interval of the second kind. To fix the notations assume that in $\{|x| \in 2^{\tilde I}\}$ we have $|x| \approx R$ and

$$|\partial^\alpha a| \lesssim M R^{-\alpha}, \qquad |\alpha| = 1, 2$$

Denote

$$r = \delta M^{-1} R$$

We partition the spacetime set $\mathbb{R} \times \{|x| \in 2^{\tilde I}\}$ into cubes of size r,

$$\{|x| \in 2^{\tilde I}\} \subset \bigcup_{Q \in \mathcal{Q}} Q$$

To this we associate a local partition of unit

$$\chi_{\tilde{I}} = \sum_{Q \in \mathcal{Q}} \chi_Q \chi_{\tilde{I}}, \qquad \text{supp } \chi_Q \subset 2Q$$

and also cutoff functions $\tilde{\chi}_Q$ with slightly larger support in $4Q$. The center of a cube Q is denoted by (t_Q, x_Q). We restrict the coefficients a to $5Q$ and then reextend them to a^Q so that they are constant and equal to (t_Q, x_Q) outside $6Q$. This can be done in such a way that the following bounds hold:

$$|a^Q - a^Q(t_Q, x_Q)| + r|\partial_x a^Q| + r^2(|\partial_x^2 a^Q| + |\partial_t a^Q|) \lesssim \delta \tag{48}$$

For each Q we choose new affine coordinates in $\mathbb{R}^n$ in which

$$6Q \subset \{|x| \approx r\}, \qquad a^Q(t_Q, x_Q) = I_n$$

Then the coefficients a^Q satisfy (8) with $\epsilon = \delta$ therefore we can find a parametrix K^Q as in Proposition 11. Consequently we define the parametrix K_I as

$$K_I = \sum_{Q \in \mathcal{Q}} \tilde{\chi}_Q K^Q \chi_Q$$

Now the bounds (46) and (47) follow as in case (a) from (39) and (40) for K^Q.

We note that this second part of the parametrix construction is not new. If we rescale Q to a cube of size $r^{-1} \times 1^n$ then we match the setup considered in [17], [3]. Rescaling Q to a cube of size $1 \times (\sqrt{r})^n$ we match the parametrix constructions in [21], [12]. All of these results employ a time independent scale for the parametrix. The only reason we prefer here to use the stronger result in Proposition 11 is to make this article self contained.

□

Proof of Theorems 3, 4. In what follows we work in a time interval $[T^-, T^+]$, possibly infinite. By (21) we can replace the operator $\tilde{A}_{(0)}$ by $\tilde{A}$ in Propositions 11, 13. Rescaling this result we obtain similar parametrices K_j at any dyadic frequency 2^j. We first assemble these dyadic parametrices and set

$$K = \sum_{K=-\infty}^{\infty} K_j S_j$$

The properties of K are summarized in the next lemma.

Lemma 14. *The parametrix K for $D_t + A$ has the following properties:*

(i) (regularity) For any Strichartz pairs (p_1, q_1) respectively (p_2, q_2) with $q_1 \leq q_2$ we have

$$\|Kf\|_{L^{p_1} L^{q_1} \cap X} \lesssim \|f\|_{L^{p_2'} L^{q_2'}} \tag{49}$$

(ii) (error estimate) For any Strichartz pair (p,q) we have

lperror2b (50)
$$\|((D_t + A)K - I)f\|_{X'} \lesssim \|f\|_{L^{p'}L^{q'}}$$

k

Part (i) follows directly from the Littlewood-Paley theory. Similarly we get part (ii) but with $\tilde{A}$ instead of A. However, by (20) we can freely interchange A and $\tilde{A}$.

A second step is to use duality to establish an $L^2 \to L^pL^q$ bound.

Lemma 15. *If there is a parametrix K for $D_t + A$ as in Lemma 14 and (p,q) is a Strichartz pair then*

xxp (51)
$$\|u\|_{L^pL^q} \lesssim \|u\|_{L^\infty L^2 \cap X} + \|(D_t + A)u\|_{X'}$$

Proof. Without any restriction in generality we assume that T^- and T^+ are finite but prove the bound with constants which are independent of T^+ and T^-. For $f \in L^{p'}L^{q'}$ we use integration by parts and the fact that A is selfadjoint to write

$$\langle u, f\rangle = \langle (D_t + A)u, Kf\rangle - \langle u, [(D_t + A)K - 1]f\rangle + i\langle u(t), Kf(t)\rangle|_{T^-}^{T^+}$$

Then by (49) and (40) we obtain

$$|\langle u, f\rangle| \lesssim \|f\|_{L^{p'}L^{q'}}(\|u\|_{L^\infty L^2 \cap X} + \|(D_t + A)u\|_{X'})$$

The conclusion follows. □

Now we prove (11). Without any restriction in generality we assume that $q_1 \leq q_2$; the opposite case follows by duality. If

$$(D_t + A)u = f + g, \qquad f \in L^{p_2'}L^{q_2'}, \ g \in X'$$

then we write

$$u = Kf + v,$$

We use (49) to bound Kf in $L^{p_1}L^{q_1} \cap X$. It remains to bound v, which solves

$$(D_t + A)v = (1 - (D_t + A)K)f + g$$

In the case of Theorem 3 we use succesively (51) and Theorem 2 we obtain

$$\begin{aligned}
\|v\|_{L^{p_1}L^{q_1}} &\lesssim \|v\|_{L^\infty L^2 \cap X} + \|(D_t + A)v\|_{X'} \\
&\lesssim \|v(0)\|_{L^2} + \|(D_t + A)v\|_{X'} \\
&\lesssim \|u(0)\|_{L^2} + \|Kf\|_{L^\infty L^2} + \|(1 - (D_t + A)K)f\|_{X'} + \|g\|_{X'} \\
&\lesssim \|u(0)\|_{L^2} + \|f\|_{L^{p_2'}L^{q_2'}} + \|g\|_{X'}
\end{aligned}$$

Then (11) follows.

In the case of Theorem tlargee 4 the argment is similar, but instead of using Theorem 12 2 we assume that the localized energy estimates hold.

□

pdo

6. Pseudodifferential operators and phase space transforms

In preparation for the outgoing parametrix construction in the following sections we introduce here the required microlocal analysis setup.

The simplest class of pseudodifferential operators which we use is S^0_{00} and some variations of it $S^{0(k)}_{00}$ defined by

$$a \in S^{0,(k)}_{00} \equiv |\partial_x^\alpha \partial_\xi^\beta a(x,\xi)| \leq c_{\alpha\beta} \qquad |\alpha| + |\beta| \geq k$$

These correspond to the euclidean metric in $\mathbb{R}^{2n}$,

$$g = dx^2 + d\xi^2$$

The indices 0 and 00 are not useful here so we simply drop them, and use instead the shorter notation $S^{(k)}$. In the sequel k will take only the values 0, 1 and 2.

In our analysis we have to work on a varying time dependent scale. Thus for $t > 0$ we introduce the classes $S^{(k)}_t$ defined by

$$a \in S^{(k)}_t \equiv |\partial_x^\alpha \partial_\xi^\beta a(x,\xi)| \leq c_{\alpha\beta} t^{\frac{|\beta|-|\alpha|}{2}} \qquad |\alpha| + |\beta| \geq k$$

These are obtained from $S^{(k)}$ by rescaling, so all the results we need are quickly transferred from $S^{(k)}$ to $S^{(k)}_t$. The correspond to the metric

$$g_t = t^{-1} dx^2 + t d\xi^2$$

The distance with respect to this metric is denoted by d_t.

Finally we also use time dependent pseudodifferential operators in $\mathbb{R} \times T^*\mathbb{R}^n$, with the corresponding symbol classes

$$a \in l^1 S^{(k)} \iff \sum_j 2^{j(1+\frac{|\alpha|-|\beta|}{2})} \|\partial_x^\alpha \partial_\xi^\beta a(t,x,\xi)\|_{L^\infty(\{t \approx 2^j\})} \leq c_{\alpha\beta}, \qquad |\alpha| + |\beta| \geq k \tag{52}$$

We note that these symbol classes are invariant with respect to the parabolic scaling

$$a(t,x,\xi) \to \lambda^2 a(\lambda^2 t, \lambda x, \lambda^{-1} t)$$

A special role in our analysis is played by the class $l^1 S^{(2)}$. In this context we introduce a notion of smallness. Precisely, for a small parameter $\epsilon > 0$ we define $l^1 S^{(2)}_\epsilon$ to consist of those $l^1S^{(2)}$ symbols for which

se2
$$\sum_j 2^{j(1+\frac{|\alpha|-|\beta|}{2})}\|\partial_x^\alpha\partial_\xi^\beta a(t,x,\xi)\|_{L^\infty(\{t\approx 2^j\})} \leq \epsilon, \qquad |\alpha|+|\beta|=2 \tag{53}$$

We can bound the terms in the above sum by a slowly varying sequence ϵ_j as in Section notations 2 and we construct the corresponding functions $\epsilon(t)$ and $e(t)$ for $t>0$. Then (se2 53) can be rewritten as

abbd
$$|\partial_x^\alpha\partial_\xi^\beta a(t,x,\xi)| \lesssim \frac{\epsilon(t)}{t} t^{\frac{|\beta|-|\alpha|}{2}} \qquad |\alpha|+|\beta|=2 \tag{54}$$

Given a large index N we can assume that there is a similar bound for the higher derivatives

abbd1
$$|\partial_x^\alpha\partial_\xi^\beta a(t,x,\xi)| \lesssim \epsilon^{-1}\frac{\epsilon(t)}{t} t^{\frac{|\beta|-|\alpha|}{2}} \qquad 3\leq |\alpha|+|\beta| \leq N \tag{55}$$

perhaps for a modified choice of the sequence ϵ_j. The large index N will be neglected in what follows, and we will argue as if $N=\infty$, keeping in mind that each of the results we prove requires only finitely many seminorms.

To study the microlocal regularity of solutions to Schrödinger type equations we use phase space transforms. Corresponding to the unit scale we have the Bargmann transform

$$Tu(x,\xi) = c_n \int e^{-\frac{(x-y)^2}{2}} e^{i\xi(x-y)} u(y)dy$$

The value $Tu(x,\xi)$ roughly measures how much of the function u is concentrated near position x and frequency ξ on the unit scale. This is an isometry from $L^2(\mathbb{R}^n)$ into $L^2(\mathbb{R}^{2n})$, which implies the inversion formula

$$T^*T = I$$

However T is not onto; its range consists of those L^2 functions which satisfy a Cauchy-Riemann type equation,

$$i\partial_\xi T = (\partial_x - i\xi)T$$

The corresponding transform on the g_t scale is obtained by rescaling, and is sometimes called the FBI transform:

$$T_{\frac{1}{t}}u(x,\xi) = c_n t^{-\frac{n}{4}} \int e^{-\frac{(x-y)^2}{2t}} e^{i\xi(x-y)} u(u)dy$$

The Cauchy-Riemann type equation has now the form

cr
$$\frac{i}{t}\partial_\xi T_{\frac{1}{t}} = (\partial_x - i\xi)T_{\frac{1}{t}} \tag{56}$$

The main idea in our approach to long time dynamics for Schrödinger type evolutions is to use a time dependent phase space transform to turn the equation into an evolution equation in the phase space. This requires results on conjugating pseudodifferential operators with respect to phase space transforms. Such results were first proved in nlw [18], cs [19], MR1944027 [20]. However, for what is needed here we refer the reader to the expository paper [21]. For convenience the results below are stated including the parameter t. However, by rescaling they all reduce to the case when $t = 1$.

Given a pseudodifferential operator in the Weyl calculus $a^w \in OPS_t^{(k)}$ we define its phase space image

$$\tilde{A} = T_{\frac{1}{t}} a^w T_{\frac{1}{t}}^*$$

The kernel of $\tilde{A}$ is called the phase space kernel of a^w. We begin our discussion with the case $k = 0$.

Proposition 16. *a) Let $A : S(\mathbb{R}^n) \to S^*(\mathbb{R}^n)$. Then $A \in OPS_t^{(0)}$ iff its phase space kernel K is rapidly decreasing away from the diagonal,*

$$|K(x_1, \xi_1; x_2, \xi_2)| \leq c_N (1 + d_t((x_1, \xi_1), (x_2, \xi_2)))^{-N} \tag{57}$$

b) Let $a \in S_t^{(0)}$ be a symbol supported in a set D. Then its phase space kernel K satisfies the stronger bound

$$|K(x_1, \xi_1; x_2, \xi_2)| \leq c_N (1 + d_t((x_1, \xi_1), (x_2, \xi_2)) + d_t((x_1, \xi_1), D)^{-N} \tag{58}$$

Part (a) is proved in [21]. Part (b) is an easy variation on the same theme which is left for the reader. As a consequence of part (a) one obtains that $OPS_t^{(0)}$ operators are L^2 bounded, which is the Calderon-Vaillancourt theorem.

For $OPS_t^{(1)}$ the L^2 boundedness is lost. However the next result asserts that modulo $OPS_t^{(0)}$ such operators can be replaced with the multiplication by their symbol in the phase space.

Proposition 17. *a) Let $a \in S_t^{(1)}$. Then we have the conjugation result*

$$T_{\frac{1}{t}} a^w = (a + E) T_{\frac{1}{t}} \tag{59}$$

where the kernel K_e of E satisfies (57).

b) Assume in addition that a is supported in a set D. Then its phase space kernel K satisfies

$$|K(x_1, \xi_1; x_2, \xi_2)| \leq c_N (|a(x_1, \xi_1)| + (1 + d_t((x_1, \xi_1), D)^{-N})(1 + d_t((x_1, \xi_1), (x_2, \xi_2)))^{-N} \tag{60}$$

If $D = \mathbb{R}^{2n}$ then part (b) follows from part (a) which is proved in [21]. Otherwise it is again a fairly straightforward variation on the same theme. A direct consequence of part (a) is the sharp Garding inequality,

Corollary 18. *Let $a \in S_t^{(1)}$ be a real nonnegative symbol. Then*

$$\langle a^w u, u\rangle \geq -C\|u\|_{L^2}^2 \tag{61}$$

garding

Finally in the case $k = 2$ we have (see [21]):

Proposition 19. *a) Let $a \in S_t^{(2)}$. Then we have the conjugation result*

$$T_{\frac{1}{t}} a^w = (a + i(a_\xi(\partial_x - i\xi) - a_x\partial_\xi) + E)T_{\frac{1}{t}} \tag{62}$$

k=2 *where the kernel K_e of E satisfies (57).*

Last but not least we consider an evolution equation which is a good model for short time Schrödinger dynamics,

pdoev
$$(D_t + a^w(t,x,D))u = 0, \qquad u(0) = u_0 \tag{63}$$

where a is a real symbol in $S^{(2)}$, uniformly in $t \in [0,1]$. For the next results we refer the reader to [21] and also [12] to some extent. We begin with the corresponding Hamilton flow,

$$\left\{ \begin{array}{l} \dot{x} = a_\xi(t,x,\xi) \\ \dot{\xi} = -a_x(t,x,\xi) \end{array} \right.$$

We denote the time evolution maps by $\chi(t,s)$. These are characterized by

Proposition 20. *Assume that a is a real symbol in $S^{(2)}$, uniformly in $t \in [0,1]$. Then $\chi(t,s)$ are bilipschitz symplectic maps.*

Now we turn our attention to the evolution (63).

Proposition 21. *Assume that a is a real symbol in $S^{(2)}$, uniformly in $t \in [0,1]$. Then (63) is L^2 well-posed forward and backward in time.*

We denote by $S(t,s)$ the corresponding evolution operators. These are characterized using the Bargmann transform as follows:

Proposition 22. *Assume that a is a real symbol in $S^{(2)}$, uniformly in $t \in [0,1]$. Then the phase space kernels $K(t,s)$ of $S(t,s)$ satisfy*

$$|K(t,x,\xi,s,y,\eta)| \leq c_N(1 + |(x,\xi) - \chi(t,s)(y,\eta)|)^{-N}$$

flatphase

In the terminology of phasespace [21] we say that $S(t,s)$ is an $S^{(0)}$ type FIO associated to the canonical transformation $\chi(t,s)$. We also have a corresponding Egorov theorem. Given a pdo $q^w(0)$ at the initial time we define its conjugates along the flow by

$$q^w(t) = S(t,0)q^w(0)S(0,t)$$

Then

Proposition 23. *Assume that a is a real symbol in $S^{(2)}$, uniformly in $t \in [0,1]$.*
a) Let $q(0) \in S^{(0)}$. Then $q(t) \in S^{(0)}$ uniformly in t.
b) Let $q(0) \in S^{(1)}$. Then $q(t) \in S^{(1)}$ uniformly in t, and

$$q(t,x,\xi) - q(0) \circ \chi(0,t) \in S^{(0)}$$

conj01

We need to improve this result in a special case. This is not considered in phasespace [21], so we will prove it here.

Proposition 24. *Assume that $a(t,x,\xi) = \xi^2$. Let $q(0) \in S^{(2)}$. Then $q(t) \in S^{(0)}$ uniformly in t and*

$$q(t,x,\xi) - q(0) \circ \chi(0,t) \in S^{(0)}$$

conj02

Proof. We compute directly

$$\chi(t,s)(x,\xi) = (x + 2(t-s)\xi, \xi)$$

Then we want to show that

$$r^w(t,x,D) = e^{-itD^2} q^w(x,D) e^{itD^2} - q^w(x+2tD, D) \in OPS^{(0)}$$

uniformly in t. Compute

$$\frac{d}{dt} e^{-itD^2} r^w(t,x,D) e^{itD^2} = e^{-itD^2} r_1^w(t,x,D) e^{itD^2}$$

where

$$r_1^w(s,x,D) = [iD^2, q^w(x+2sD,D)] - \frac{d}{ds} q^w(x+2sD,D)$$

Then using the Weyl calculus we get

$$r_1(s,x,\xi) = i\Delta_x q(x+2s\xi,\xi) \in S^{(0)}$$

By Proposition conj01 23 conjugation by $e^{\pm itD^2}$ leaves the $S^{(0)}$ class unchanged so the conclusion follows. $\square$

modeleq

7. A long time phase space parametrix

In this section we construct global in time parametrices for a class of equations governed by pseudodifferential operators $a^w(t,x,D)$ satisfying a smallness condition,

$$a \in l^1 S^{(2)}_\epsilon, \qquad \epsilon \ll 1$$

This class does not include the operator $A_{(0)}$ which we are interested in. However, it does include the operator $-\Delta - A_{(0)}$ in the phase space region

$$\{|\xi| \approx 1,\ |x| \approx |t|,\ t \geq 1\}$$

This will allow us in the next section to make the transition to $A_{(0)}$ via a conjugation with respect to the flat Schrödinger flow.

In our analysis we add a damping term to the L^2 conservative equation. Its role will ultimately be to kill all the waves which stray away from the above phase space region.

Thus we consider the forward evolution equation

abevl (64)
$$(D_t + a^w(t,x,D) - ib^w(t,x,D) + c^w(t,x,D))u = 0, \qquad t > 0$$

where $a \in l^1 S^{(2)}_\epsilon$ respectively $b \in l^1 S^{(1)}$ are real symbols with $b \geq 0$ while $c \in l^1 S^{(0)}$ is a complex symbol. We think of a^w as the operator driving the evolution while b^w is a damping term and c^w is a negligible error. As discussed in the previous section, we can assume that a satisfies (54), (55) with $\epsilon(t)$ as in Section 2. Without any restriction on generality we can also assume that b and c satisfy (54) for $|\alpha| + |\beta| \geq 1$, respectively $|\alpha| + |\beta| \geq 0$.

It is fairly easy to study this evolution in L^2:

Proposition 25. *Assume that $a \in \epsilon l^1 S^{(2)}$ and $b \in \epsilon l^1 S^{(1)}$ are real symbols with $b \geq 0$, while $c \in \epsilon l^1 S$. Then the equation (64) is forward well-posed in $L^2(\mathbb{R}^n)$, and the corresponding evolution operators satisfy*

$$\|S(t,s)\|_{L^2 \to L^2} \lesssim 1, \qquad 0 < s < t$$

Proof. To establish energy estimates for solutions u to (64) it suffices to compute

$$\frac{d}{dt}\|u(t)\|^2_{L^2} = -\langle b^w(t,x,D)u(t), u(t)\rangle - \Im\langle c^w(t,x,D)u(t), u(t)\rangle$$

Garding's inequality in Corollary 18 applied to b yields

$$\langle b^w(t,x,D)u(t), u(t)\rangle \geq -C\epsilon^{-1}\frac{\epsilon(t)}{t}\|u\|^2_{L^2}$$

while for c^w we have L^2 bounds

$$\|c^w(t,x,D)u\|_{L^2} \leq C\epsilon^{-1}\frac{\epsilon(t)}{t}\|u\|_{L^2}$$

Hence by Gronwall's inequality we get

$$\|u(t)\|_{L^2}^2 \leq e^{2C(e(t)-e(s))}\|u(s)\|_{L^2}^2, \qquad t > s$$

$\square$

We are interested in obtaining much more precise bounds on the phase space localization of the solutions. The phase space image of the evolution $S(t,s)$ is the family of evolution operators

$$\tilde{S}(t,s) = T_{\frac{1}{t}} S(t,s) T_{\frac{1}{s}}^*$$

Our goal is to obtain precise bounds on the phase space kernels of $\tilde{S}(t,s)$. These are described in terms of two geometric quantities:

(i) The Hamilton flow of $D_t + a^w$. This is described by the ode's

$$\begin{cases} \dot{x} = a_\xi(t,x,\xi) \\ \dot{\xi} = -a_x(t,x,\xi) \end{cases}$$

We denote the trajectories of the Hamilton flow by

$$t \to (x_t, \xi_t)$$

and the flow map by $\chi(t,s)$. The regularity of the flow is computed using the linearized equations:

Proposition 26. *If $a \in l^1 S^{(2)}_\epsilon$ and $t > s$ then the Hamilton flow has the Lipschitz regularity*

$$\frac{\partial(x_t,\xi_t)}{\partial(x_s,\xi_s)} = \begin{pmatrix} I_n + \epsilon O(\frac{t}{s}) & \epsilon O(\frac{1}{s}) \\ \epsilon O(t) & I_n + \epsilon O(1) \end{pmatrix} \tag{65}$$

respectively

$$\frac{\partial(x_s,\xi_s)}{\partial(x_t,\xi_t)} = \begin{pmatrix} 1 + \epsilon O(1) & \epsilon O(\frac{1}{s}) \\ \epsilon O(t) & I_n + \epsilon O(\frac{t}{s}) \end{pmatrix} \tag{66}$$

flowreg

We note that if ϵ is small and $s < t$ then for fixed x_t the map $\xi_s \to \xi_t$ is a diffeomorphism. Then it is more convenient to parametrize the graph of $\chi(t,s)$ using the variables (x_t,ξ_s). This choice of independent variables yields the better relation

xsxt

$$\frac{\partial(x_t,\xi_s)}{\partial(x_s,\xi_t)} = \begin{pmatrix} I_n + \epsilon O(1) & \epsilon O(\frac{1}{s}) \\ \epsilon O(t) & I_n + \epsilon O(1) \end{pmatrix} \tag{67}$$

(ii) The exponential decay along the flow determined by the damping. Along each trajectory (x_t,ξ_t) we define the weight function

$$\psi(t,x_t,\xi_t) = \int_1^t b(s,x_s,\xi_s)ds$$

Heuristically we expect $e^{-\psi(t,x_t,\xi_t)}$ to describe the behavior of the energy along the flow. The lower limit of integration is set arbitrarily to 1. In our analysis we only care about the differences

$$\psi(x_t,\xi_t) - \psi(x_s,\xi_s)$$

Their Lipschitz dependence on the (x_s,ξ_t) variables is described in the following

Proposition 27. *If $a \in l^1 S^{(2)}_\epsilon$ with ϵ small, $b \in l^1 S^{(1)}$ and $t > s$ then*

psir

$$\frac{\partial(\psi(x_t,\xi_t) - \psi(x_s,\xi_s))}{\partial(x_s,\xi_t)} = (O(s^{-\frac{1}{2}}), O(t^{\frac{1}{2}})) \tag{68}$$

psireg

The proof uses again the linearization of the Hamilton flow. The argument is routine and is left for the reader.

The main result of this section is a sharp pointwise bound on the kernel of the phase space operator $\tilde{S}(t,s)$.

Theorem 28. *Let $a \in l^1 S^{(2)}_\epsilon$, $b \in l^1 S^{(1)}$ be real symbols with $b \geq 0$ and $c \in l^1 S^{(0)}$. Then for $s < t$ the kernel K of the operator $\tilde{S}(t,s)$ satisfies the bound*

kbd

$$|K(t,x,\xi_t,s,x_s,\xi)| \lesssim t^{-\frac{n}{4}} s^{\frac{n}{4}} \left(1 + (\psi(x_s,\xi_s) - \psi(x_t,\xi_t))^2 + \frac{(x-x_t)^2}{t} + s(\xi-\xi_s)^2\right)^{-N} \tag{69}$$

lt1

Proof. If u is the forward solution to (abevl 64) with initial data

$$u(s,y) = c_n s^{-\frac{1}{4}} e^{-\frac{(y-x_s)^2}{2s}} e^{i\xi(y-x_s)}$$

then the kernel K is given by

$$K(t,x,\xi_t,s,x_s,\xi) = (T_{\frac{1}{t}} u(t))(x,\xi_t)$$

At time $t = s$ a direct computation gives an initial data for K,

$$K(s,x,\xi_s,s,x_s,\xi) = c_n e^{-\frac{(x-x_s)^2}{4s}} e^{-\frac{(\xi-\xi_s)^2}{4s}} e^{i\frac{1}{2}(x-x_s)(\xi+\xi_s)}$$

From (abevl 64) we have

$$0 = T_{\frac{1}{t}}(\partial_t + i a^w(t,x,D) + b^w(t,x,\xi) + i c^w(t,x,D))u$$

To obtain an equation for K we need to conjugate the above pseudodifferential operators with respect to the phase space transform $T_{\frac{1}{t}}$.

For the time derivative a direct computation yields

$$\partial_t T_{\frac{1}{t}} = (-\frac{n}{4t} - \frac{1}{2t^2}\partial_\xi^2) T_{\frac{1}{t}}$$

Using the Cauchy-Riemann type equation (56) this can be rewritten in the form

$$\partial_t T_{\frac{1}{t}} = \left(\frac{n}{4t} + \frac{1}{2}(\partial_x - i\xi)^2 \right) T_{\frac{1}{t}}$$

For the pseudodifferential operators a^w,b^w and c^w we use the conjugation results in Propositions 16,17,19. Adding the pieces together we can write an equation for the phase space function $K(t) = T_{\frac{1}{t}} u(t)$:

$$\left(\partial_t + ia + b(t,x,\xi) - a_x \partial_\xi + a_\xi(\partial_x - i\xi) - \frac{n}{4t} - \frac{1}{2}(\partial_x - i\xi)^2 + E \right) K(t,x,\xi) = 0$$

where E stands for an error term with good kernel bounds,

$$|E(t,x_1,\xi_1,x_2,\xi_2)| \lesssim \frac{\epsilon(t)}{t\epsilon}(1 + t^{-1}(x-x_1)^2)^{-N}(1+t(\xi-\xi_1)^2)^{-N}$$

We prove the bound for K using the maximum principle for the function $|K|$. Compute

$$\begin{aligned} \partial_t |K| &= |K|^{-1} \Re(K_t \bar{K}) \\ &\leq (a_x\partial_\xi - a_\xi \partial_x - b + \frac{n}{4t})|K| + |K|^{-1}\frac{1}{2}\Re(\partial_x - i\xi)^2 K\bar{K} + E|K| \end{aligned}$$

With $Z = Ke^{-ix\xi}$ we write the second order term as

$$\begin{aligned} \Re(\partial_x - i\xi)^2 K\bar{K} &= \Re \partial_x^2 Z \bar{Z} \\ &= \frac{1}{2}\partial_x^2 |Z|^2 - |\nabla Z|^2 \\ &= |Z|\partial_x^2|Z| + |\nabla|Z||^2 - |\nabla Z|^2 \\ &\lesssim |K|\partial_x^2 |K| \end{aligned}$$

This leads to

$$L|K| \leq E|K|, \qquad L = \partial_t + b - \frac{n}{4t} + a_\xi \partial_x - a_x \partial_\xi - \frac{1}{2}\partial_x^2 \tag{70}$$

Since L is a degenerate parabolic operator, it satisfies the maximum principle. Hence in order to obtain bounds for $|K|$ it suffices to construct an appropriate positive supersolution. To better motivate the construction we do this in several steps.

We begin with the case when $a = 0,\ b = 0$ when we can use the fundamental solution to the heat equation,

$$W_0(t,x) = t^{-\frac{n}{2}} \phi_0\left(\frac{x^2}{2t}\right), \qquad \phi_0(s) = e^{-s}$$

In order to allow for a and b we need to replace this exact solution with a more robust supersolution. Precisely, for a large constant C we define

$$W_1(t,x) = e^{C\epsilon(t)}(t + \epsilon^{-1}t\epsilon(t))^{-\frac{n}{2}} \phi_0\left(\frac{x^2}{2(t+\epsilon^{-1}t\epsilon(t))}\right)$$

For this we compute

gauss
$$
\begin{aligned}
(\partial_t - \frac{1}{2}\Delta_x)W_1(t,x) &= (Ce'(t) + \epsilon^{-1}\frac{\epsilon(t)+t\epsilon'(t)}{t+\epsilon^{-1}t\epsilon(t)}\left(-\frac{n}{2}+\frac{x^2}{2(t+t\epsilon(t))}\right)W_1(t,x) \\
&\approx \epsilon^{-1}\frac{\epsilon(t)}{t}\left(C+\frac{x^2}{2t}\right)W_1(x,t)
\end{aligned}
\tag{71}
$$

This allows us to construct suitable supersolutions for the operator L.

Lemma 29. *Given* (s, x_s, ξ) *we define the function*

$$
W_{s,x_s,\xi}(t,x,\xi_t) = s^{\frac{1}{4}}t^{\frac{1}{4}}(1+s(\xi-\xi_s)^2)^{-N}e^{-(\psi(t,x_t,\xi_t)-\psi(s,x_s,\xi_s))}W_1(t,x-x_t), \qquad t \geq s
$$

Then

$$
LW_{s,x_s,\xi} > 0
$$

noe

Proof. Since x_s is fixed, it follows that x_t and ξ_s are functions of ξ_t and t. The transport equation for ψ shows that

$$
(\partial_t - a_x(t,x_t,\xi_t)\partial_{\xi_t})\psi(t,x_t,\xi_t) = b(t,x_t,\xi_t)
$$

Hence for the exponential we obtain

$$
\begin{aligned}
&(\partial_t + b(t,x,\xi_t) + a_\xi(t,x,\xi_t)\partial_x - a_x(t,x,\xi_t)\partial_{\xi_t})e^{-\psi(t,x_t,\xi_t)} = \\
&[b(t,x,\xi_t) - b(t,x_t,\xi_t) + (a_x(t,x,\xi_t) - a_x(t,x_t,\xi_t))\partial_{\xi_t}\psi(t,x_t,\xi_t)]e^{-\psi(t,x_t,\xi_t)}
\end{aligned}
$$

which by (psir 68) for ψ, and (abbd 54) for both a and b yields the bound

w1
$$
(\partial_t + b(t,x,\xi_t) + a_\xi(t,x,\xi_t)\partial_x - a_x(t,x,\xi_t)\partial_{\xi_t})e^{-\psi(t,x_t,\xi_t)} = O\left(\frac{\epsilon(t)}{t\epsilon}\right)\frac{|x-x_t|}{t^{\frac{1}{2}}}e^{-\psi(t,x_t,\xi_t)} \tag{72}
$$

A similar computation applies for the first factor, namely

$$
\begin{aligned}
&(\partial_t + a_\xi(t,x,\xi_t)\partial_x - a_x(t,x,\xi_t)\partial_{\xi_t})(1+s(\xi-\xi_s)^2)^{-N} = \\
&(a_x(x,\xi_t) - a_x(x_t,\xi_t))\partial_{\xi_t}(1+s(\xi-\xi_s)^2)^{-N}
\end{aligned}
$$

By (xsxt 67), ξ_s is a Lipschitz function of ξ_t. Hence for $t > s$ we obtain

w2
$$
\begin{aligned}
&(\partial_t + a_\xi(t,x,\xi_t)\partial_x - a_x(t,x,\xi_t)\partial_{\xi_t})(1+s(\xi-\xi_s)^2)^{-N} \\
&= O\left(\frac{\epsilon(t)}{t}\right)\frac{|x-x_t|}{t^{\frac{1}{2}}}(1+s(\xi-\xi_s)^2)^{-N}
\end{aligned}
\tag{73}
$$

Finally for the W_1 component we need to use the Hamilton flow equations to obtain

$$
\partial_t x_t = a_\xi(t,x_t,\xi_t) - a_x(t,x_t,\xi_t)\partial_{\xi_t}x_t
$$

Then we can compute

$$
\begin{aligned}
&(\partial_t + a_\xi(t,x,\xi_t)\partial_x - a_x(t,x,\xi_t)\partial_{\xi_t} - \frac{1}{2}\Delta_x)W_1(t,x-x_t)\\
=&[(\partial_t - \frac{1}{2}\Delta_x)W_1](t,x-x_t) - (\partial_t x_t - a_\xi(t,x,\xi_t) + a_x(t,x,\xi_t)\partial_{\xi_t}x_t)[\partial_x W_1](t,x-x_t)\\
=&[(\partial_t - \frac{1}{2}\Delta_x)W_1](t,x-x_t)\\
+&(a_\xi(t,x,\xi_t) - a_\xi(t,x_t,\xi_t) - (a_x(t,x,\xi_t) - a_x(t,x_t,\xi_t))\partial_{\xi_t}x_t)[\partial_x W_1](t,x-x_t)
\end{aligned}
$$

By (67) we have $|\partial_{\xi_t}x_t| \lesssim t$. Using also (54) for a we obtain

$$
\begin{aligned}
&(\partial_t + a_\xi(t,x,\xi_t)\partial_x - a_x(t,x,\xi_t)\partial_{\xi_t} - \frac{1}{2}\Delta_x)W_1(t,x-x_t)\\
&= [(\partial_t - \frac{1}{2}\Delta_x)W_1](t,x-x_t) + O\left(\frac{\epsilon(t)}{t}\right)\frac{(x-x_t)^2}{t}W_1(t,x-x_t)
\end{aligned}
\tag{74}
$$

Putting together the results in (72), (73) and (74) we obtain

$$
\begin{aligned}
LW_{s,x_s,\xi}(t,x,\xi_t) =& s^{\frac{1}{4}}t^{\frac{1}{4}}(1+s(\xi-\xi_s)^2)^{-N}e^{-(\psi(t,x_t,\xi_t)-\psi(s,x_s,\xi_s))}[(\partial_t - \frac{1}{2}\Delta_x)W_1](t,x-x_t)\\
&+ O\left(\frac{\epsilon(t)}{t\epsilon}\right)\left(\frac{|x-x_t|}{t^{\frac{1}{2}}} + \epsilon\frac{(x-x_t)^2}{t}\right)W_{s,x_s,\xi}(t,x,\xi_t)
\end{aligned}
$$

Then the conclusion of the lemma follows from (71). $\square$

To conclude the proof of the Theorem we need to also allow for the error term E. Since the kernel of E is merely of rapid decrease at infinity, we expect E to replace the Gaussian with a rapidly decreasing factor on the same scale. To achieve this we use the functions $W_{s,x_s,\xi}(t,x,\xi_t)$ in a manner similar to the variation of the parameters formula.

This argument involves multiple bicharacteristic rays. Hence in order to avoid confusion we introduce some new notations. Given a position x at time s and a frequency ξ at time $t > s$ we denote by

$$\sigma \to (x_\sigma(x,\xi), \xi_\sigma(x,\xi))$$

the bicharacteristic satisfying

$$x_s(x,\xi) = x, \qquad \xi_t(x,\xi) = \xi$$

The variables s,t are ommitted from this notation, but they will always be clear from the context.

For $t \geq s$ we define the modified function

$$\tilde{W}_{s,x_s,\xi_s}(t,x,\xi) = t^{-\frac{n}{2}}F_{s,t}(x - x_t(x_s,\xi), \xi_s(x_s,\xi) - \xi_s, \psi(x_t(x_s,\xi),\xi) - \psi(x_s,\xi_s(x_s,\xi))) \tag{75}$$

where

$$F_{s,t}(y,\eta,\psi) = \left(1+\frac{y^2}{t}\right)^{-N_0}\left(1+\psi^2+\frac{y^2}{t}\right)^{-N_1}\left(1+s\left(\eta^2+\frac{y^2}{t^2}\right)\right)^{-N_2}$$

It is easy to see that

$$W_{s,x_s,\xi_s} \lesssim \tilde{W}_{s,x_s,\xi_s}$$

Then the counterpart of Lemma 29 is

Lemma 30. *Assume that $N_0 > \frac{n}{2}$, $N_1, N_2 \geq 0$. Then there is a function Z_{s,x_s,ξ_s} satisfying the bounds*

zw0

$$Z_{s,x_s,\xi_s}(s) \approx \tilde{W}_{s,x_s,\xi_s}(s) \tag{76}$$

and

zw

$$W_{s,x_s,\xi_s}(t) \lesssim Z_{s,x_s,\xi_s}(t) \lesssim \tilde{W}_{s,x_s,\xi_s}(t), \qquad t > s \tag{77}$$

and which is a strong supersolution for L, namely

supere

$$LZ_{s,x_s,\xi_s} \gtrsim E\tilde{W}_{s,x_s,\xi_s} \tag{78}$$

zl

Proof. We define the function Z in a manner inspired by the Duhamel formula,

$$\begin{aligned} Z_{s,x_s,\xi_s}(t,x,\xi) = &\int_{\mathbb{R}^{2n}} W_{s,y,\eta}(t,x,\xi)\tilde{W}_{s,x_s,\xi_s}(s,y,\eta)dyd\eta d\sigma \\ &+ \epsilon^{-1}\int_s^t \frac{\epsilon(\sigma)}{\sigma}\int_{\mathbb{R}^{2n}} W_{\sigma,y,\eta}(t,x,\xi)\tilde{W}_{s,x_s,\xi_s}(\sigma,y,\eta)dyd\eta d\sigma \end{aligned}$$

The initial data for Z is

$$Z_{s,x_s,\xi_s}(s,x,\xi) = \int_{\mathbb{R}^{2n}} W_{s,y,\eta}(s,x,\xi)\tilde{W}_{s,x_s,\xi_s}(s,y,\eta)dyd\eta d\sigma$$

The kernel $W_{s,y,\eta}(s,x,\xi)$ is rapidly decreasing on the $s^{\frac{1}{2}} \times s^{-\frac{1}{2}}$ scale away from the diagonal, (i.e. it satisfies (57) with $t = s$). Since W is temperate on the g_t scale, the bound (76) follows.

To verify (78) we use Lemma 29 to compute

$$LZ_{s,x_s,\xi_s}(t,x,\xi) \gtrsim \epsilon^{-1}\frac{\epsilon(t)}{t}\int_{\mathbb{R}^{2n}} W_{t,y,\eta}(t,x,\xi)\tilde{W}_{s,x_s,\xi_s}(t,y,\eta)dyd\eta$$

But $W_{t,y,\eta}(t,x,\xi)$ is a rapidly decaying kernel on the $t^{\frac{1}{2}} \times t^{-\frac{1}{2}}$ scale,

$$|W_{t,y,\eta}(t,x,\xi)| \lesssim (1+t^{-1}(x-y)^2+t(\xi-\eta)^2)^{-N}$$

while, by (67) and (68), $\tilde{W}_{s,x_s,\xi_s}(t,y,\eta)$ is temperate on the same scale. Hence we obtain

$$LZ_{s,x_s,\xi_s}(t,x,\xi) \gtrsim \epsilon^{-1}\frac{\epsilon(t)}{t}\tilde{W}_{s,x_s,\xi_s}(t,x,\xi) \gtrsim E\tilde{W}_{s,x_s,\xi_s}(t,x,\xi)$$

Therefore to conclude the proof of the lemma it remains to prove the second part of (77), i.e. that $Z \lesssim \tilde{W}$. Since

$$\int \frac{\epsilon(t)}{t} \approx \epsilon$$

this reduces to the fixed time bound

$$\int_{\mathbb{R}^{2n}} W_{\sigma,y,\eta}(t,x,\xi)\tilde{W}_{s,x_s,\xi_s}(\sigma,y,\eta)dyd\eta \lesssim \tilde{W}_{s,x_s,\xi_s}(t,x,\xi) \qquad s<\sigma<t$$

By rescaling we can take $\sigma = 1$. The powers of s and t cancel, so it remain to prove the integral bound

$$I \lesssim F_{s,t}(x - x_t(x_s,\xi), \xi_s(x_s,\xi) - \xi_s, \psi(t,x_t(x_s,\xi),\xi) - \psi(s,x_s,\xi_s(x_s,\xi))) \tag{79}$$

where

$$I = \int_{\mathbb{R}^{2n}} e^{-\frac{(x-x_t(y,\xi))^2}{t}}(1+(\eta-\xi_1(y,\xi))^2)^{-N}e^{-\psi(t,x_t(y,\xi),\xi)+\psi(1,y,\xi_1(y,\xi))}$$
$$F_{s,1}(y - x_1(x_s,\eta), \xi_s(x_s,\eta)-\xi_s, \psi(1,x_1(x_s,\eta),\eta) - \psi(s,x_s,\xi_s(x_s,\eta)))dyd\eta$$

Using the trivial bound

$$(1+p^2)^{-N_1}e^{-q} \lesssim (1+(p+q)^2)^{-N_1}$$

we include the second exponential in F,

$$I \lesssim \int_{\mathbb{R}^{2n}} e^{-\frac{(x-x_t(y,\xi))^2}{t}}(1+(\eta-\xi_1(y,\xi))^2)^{-N}F_{s,1}(y-x_1(x_s,\eta),\xi_s(x_s,\eta)-\xi_s,\delta\psi(y,\eta))dyd\eta$$

where

$$\delta\psi(y,\eta) = -\psi(t,x_t(y,\xi),\xi) + \psi(1,y,\xi_1(y,\xi) - \psi(1,x_1(x_s,\eta),\eta) + \psi(s,x_s,\xi_s(x_s,\eta))$$

By Proposition 26 the expressions $x_1(x_s,\eta)$ and $\xi_s(x_s,\eta)$ are Lipschitz in η. By Proposition 27 the expression $\delta\psi(y,\eta)$ is also Lipschitz with respect to η. Hence for large enough N the integration with respect to η is trivial, and we obtain

$$I \approx \int_{\mathbb{R}^n} e^{-\frac{(x-x_t(y,\xi))^2}{t}}F_{s,1}(y-x_1(x_s,\xi_1(y,\xi)),\xi_s(x_s,\xi_1(y,\xi))-\xi_s,\delta\psi(y,\xi_1(y,\xi)))dy$$

Denote

$$\alpha = x_t(x_s,\xi) - x_t(y,\xi)$$

Then repeatedly using Proposition 26 we obtain

$$y - x_1(x_s,\xi_1(y,\xi)) = (1+O(\epsilon))(x_s(y,\xi)-x_s) = (1+O(\epsilon))(x_t(x_s,\xi)-x_t(y,\xi))$$

namely

alpha1
$$y - x_1(x_s, \xi_1(y,\xi)) = (1 + O(\epsilon))\alpha \tag{80}$$

Similarly we obtain

$$\xi_s(x_s, \xi_1(y,\xi)) - \xi_s(x_s,\xi) = (1+O(\epsilon))(\xi_1(y,\xi) - \xi_1(x_s,\xi)) = O(\epsilon)(x_t(y,\xi) - x_t(x_s,\xi))$$

i.e.

alpha2
$$\xi_s(x_s, \xi_1(y,\xi)) - \xi_s(x_s,\xi) = O(\epsilon)\alpha \tag{81}$$

On the other hand by Proposition psireg 27 the expression $\delta\psi(y,\eta)$ is Lipschitz in both arguments, therefore

$$\begin{aligned}
\delta\psi(y, \xi_1(y,\xi)) &= \delta\psi(x_1(x_s,\xi), \xi_1(x_s,\xi)) + O(|y - x_1(x_s,\xi)| + |\xi_1(y,\xi) - \xi_1(x_s,\xi)|) \\
&= \delta\psi(x_1(x_s,\xi), \xi_1(x_s,\xi)) + O(|\alpha|)
\end{aligned}$$

The middle terms cancel in the expression for $\delta\psi(x_1(x_s,\xi),\xi_1(x_s,\xi))$ and we obtain

alpha3
$$\delta\psi(y,\xi_1(y,\xi)) = -\psi(x(x_s,\xi),\xi) + \psi(x_s, \xi_s(x_s,\xi)) + O(|\alpha|) \tag{82}$$

Using (alpha1 80),(alpha2 81) and (alpha3 82) in the expression for $F_{s,1}$ yields

$$\begin{aligned}
F_{s,1}(y - x_1(x_s,\xi_1(y,\xi)), \xi_s(x_s,\xi_1(y,\xi)) - \xi_s, \delta\psi(y,\xi_1(y,\xi))) \approx \\
F_{s,1}(\alpha, \xi_s - \xi_s(x_s,\xi), \psi(x(x_s,\xi),\xi) - \psi(x_s,\xi_s(x_s,\xi)))
\end{aligned}$$

By Proposition flowreg 26 there is a bilipschitz correspondence between y and α. Hence we can change the variable of integration to α to obtain

$$I \lesssim \int_{\mathbb{R}^n} e^{-\frac{(x - x_t(x_s,\xi) - \alpha)^2}{t}} F_{s,1}(\alpha, \xi_s - \xi_s(x_s,\xi), \psi(x(x_s,\xi),\xi) - \psi(x_s,\xi_s(x_s,\xi)))d\alpha$$

Then (iii 79) reduces to

fstbd
$$\int_{\mathbb{R}^n} e^{-\frac{(x-\alpha)^2}{t}} F_{s,1}(\alpha,\xi,\psi)d\alpha \lesssim \Psi_{s,t}(x,\xi,\psi), \qquad s \leq 1 \leq t. \tag{83}$$

Expanding this is written as

$$\begin{aligned}
\int_{\mathbb{R}^n} e^{-\frac{(x-\alpha)^2}{t}} (1+\alpha^2)^{-N_0}(1+\alpha^2+\psi^2)^{-N_1}(1+s(\xi^2+\alpha^2))^{-N_2} d\alpha \lesssim \\
(1+t^{-1}x^2)^{-N_0}(1+x^2+\psi^2)^{-N_1}(1+s(\xi^2+t^{-2}x^2))^{-N_2}
\end{aligned}$$

After a dyadic decomposition with respect to the size of $|x-\alpha|$ this is equivalent to

$$\begin{aligned}
\int_{|x-\alpha|<\sqrt{t}} (1+\alpha^2)^{-N_0}(1+\alpha^2+\psi^2)^{-N_1}(1+s(\xi^2+\alpha^2))^{-N_2} d\alpha \lesssim \\
(1+t^{-1}x^2)^{-N_0}(1+x^2+\psi^2)^{-N_1}(1+s(\xi^2+t^{-2}x^2))^{-N_2}
\end{aligned}$$

We consider two cases. If $|x| > 2\sqrt{t}$ then the integrand has constant size on the domain of integration, so the integral has size

$$t^{\frac{n}{2}}(1+x^2)^{-N_0}(1+x^2+\psi^2)^{-N_1}(1+s(\xi^2+x^2))^{-N_2} \leq$$
$$t^{\frac{n}{2}-N_1}(1+t^{-1}x^2)^{-N_0}(1+t^{-1}x^2+\psi^2)^{-N_1}(1+s(\xi^2+t^{-2}x^2))^{-N_2}$$

If on the other hand $|x| \leq 2\sqrt{t}$ then it suffices to bound the integral by

$$\int_{|\alpha|<3\sqrt{t}} (1+\alpha^2)^{-N_0}(1+\alpha^2+\psi^2)^{-N_1}(1+s(\xi^2+\alpha^2))^{-N_2} d\alpha \lesssim (1+\psi^2)^{-N_1}(1+s\xi^2)^{-N_2}$$

This concludes the proof of (83) and therefore the proof of the lemma. $\square$

For the proof of Theorem 28 we need a slightly better control of the constants than in the previous lemma:

Lemma 31. *Let Z_{s,x_s,ξ_s} be as in Lemma 30. Then for any $M \geq 2$ we have*

$$L(e^{Me(t)}Z_{s,x_s,\xi_s}) \gtrsim c_N (\ln M)^N E(e^{Me(t)}Z_{s,x_s,\xi_s})$$

Proof. We have

$$L(e^{Me(t)}Z_{s,x_s,\xi_s}) = e^{Me(t)}(Me'(t)Z_{s,x_s,\xi_s} + LZ_{s,x_s,\xi_s})$$

Using Lemma 30 for the second term, it remains to prove that

$$Me'(t)Z_{s,x_s,\xi_s} + E\tilde{W}_{s,x_s,\xi_s} \gtrsim (\ln M)^N EZ_{s,x_s,\xi_s}$$

But $\tilde{W}_{s,x_s,\xi_s}$ is g_t temperate, therefore

$$E\tilde{W}_{s,x_s,\xi_s} \approx \frac{\epsilon(t)}{t\epsilon}\tilde{W}_{s,x_s,\xi_s}$$

Elliminating the common factor $\epsilon(t)/t\epsilon$ the bound above reduces to

$$E_t Z_{s,x_s,\xi_s} \lesssim MZ_{s,x_s,\xi_s} + (\ln M)^{-N}\tilde{W}_{s,x_s,\xi_s} \tag{84}$$

where E_t is the convolution with the rapidly decreasing kernel $(1+t^{-1}x^2+t\xi^2)^{-4N}$, with N sufficiently large.

By rescaling we can take $t=1$. To compare the averages of Z with its pointwise values we need some control on its derivatives. We claim that for all R sufficiently large we have

$$|\nabla Z_{s,x_s,\xi_s}(1,x,\xi)| \leq RZ_{s,x_s,\xi_s}(1,x,\xi) + Ce^{-\frac{R^2}{3}}\tilde{W}_{s,x_s,\xi_s}(1,x,\xi) \tag{85}$$

We first use this to prove (84). Gronwall's inequality together with the fact that $\tilde{W}_{s,x_s,\xi_s}$ is temperate yield

$$Z_{s,x_s,\xi_s}(1,x_1,\xi_1) \lesssim e^{Rd}Z_{s,x_s,\xi_s}(1,x,\xi) + e^{Rd-\frac{R^2}{4}}\tilde{W}_{s,x_s,\xi_s}(1,x,\xi) \qquad d = d(x,\xi,x_1,\xi_1)$$

This is useful in a ball of radius $\frac{R}{8}$. Outside we have the trivial bound $Z_{s,x_s,\xi_s} \lesssim \tilde{W}_{s,x_s,\xi_s}$; since $\tilde{W}$ is temperate, for large enough N we gain powers of R. Hence we obtain

$$E_1 Z_{s,x_s,\xi_s}(1,x,\xi) \lesssim e^{\frac{R^2}{8}} Z_{s,x_s,\xi_s}(1,x,\xi) + R^{-2N} \tilde{W}_{s,x_s,\xi_s}(1,x,\xi)$$

which gives (84) for $M = e^{\frac{R^2}{8}}$.

It remains to prove (85). For this we differentiate $W_{s,x_s,\xi}$,

$$|\nabla_{x,\xi_t} W_{s,x_s,\xi}(1,x,\xi_t)| \leq (C + |x - x_t|) W_{s,x_s,\xi}(1,x,\xi_t)$$

For $|x - x_t| < R$ we bound this in terms of $W_{s,x_s,\xi}(1,x,\xi_t)$, otherwise we contend ourselves with some rapid decay:

$$|\nabla_{x,\xi_t} W_{s,x_s,\xi}(1,x,\xi_t)| \leq (R+C) W_{s,x_s,\xi}(1,x,\xi_t) + C e^{-\frac{R^2}{2}} R^{2N} (1+(x-x_t)^2)^{-N} (1+s(\xi-\xi_s)^2)^{-N}$$

We use this in the expression for ∇Z and integrate. Observe that in (83) the exponential can be harmlessly replaced by a sufficiently large polynomial decay, therefore the contribution of the second term is still bounded by $\tilde{W}$. This yields

$$|\nabla_{x,\xi_t} W_{s,x_s,\xi}(1,x,\xi_t)| \lesssim (R+C) W_{s,x_s,\xi}(1,x,\xi_t) + C e^{-\frac{R^2}{2}} R^{2N} \tilde{W}_{s,x_s,\xi}(1,x,\xi_t)$$

Redenoting $R := R + C$ the bound (85) follows.

$\square$

Now we can conclude the proof of Theorem 28 using the maximum principle for $L-E$. By (70) the function $|K|$ is a subsolution for $L-E$, while by Lemma 31 the function $e^{Me(t)} Z_{s,x_s,\xi_s}$ is a supersolution for large enough M. Hence we obtain

$$|K(t,x,\xi_t)| \lesssim e^{Me(t)} Z_{s,x_s,\xi_s} \lesssim \tilde{W}_{s,x_s,\xi}(t,x,\xi_t)$$

$\square$

8. A perturbation of the Schrodinger equation

Here we consider the evolution equation

$$(D_t - \Delta + a_0^w(t,x,D) - i b_0^w(t,x,D))u = 0$$

where $a_0 \in l^1 S^{(2)}_\epsilon$, $b_0 \in l^1 S^{(1)}$ are real symbols with $b \geq 0$. This will serve as the model for our outgoing parametrix. We denote by $S_0(t,s)$ the L^2 evolution generated by the above equation, and by $\tilde{S}_0(t,s)$ its phase space image

$$\tilde{S}_0(t,s) = T_{\frac{1}{t}} S_0(t,s) T^*_{\frac{1}{s}}$$

We want to obtain bounds on the kernel of $\tilde{S}_0(t,S)$ which are similar to the ones in Theorem 28. As a preliminary step we need to study the regularity of the associated Hamilton flow which we denote by $\chi_0(t,s)$. This can be done directly, but for our purposes it is more convenient to reduce it to the case considered in the previous section.

At each time t we consider the simplectic map μ defined by

$$\mu_t(x,\xi) = (x + 2t\xi, \xi)$$

This extends to a space-time simplectic map

$$\mu(t,\tau,x,\xi) = (t, \tau - \xi^2, x + 2t\xi, \xi)$$

If p_0 is the symbol

$$p_0(t,\tau,x,\xi) = \tau + \xi^2 + a_0(t,x,\xi)$$

then its image through μ is

$$p(\mu(t,\tau,x,\xi)) = \tau + a(\mu(t,\tau,x,\xi)), \qquad a(t,x,\xi) = a_0(t, x + 2t\xi, \xi)$$

Hence the conjugate of the Hamilton flow $\chi_0(t,s)$ for $\tau + \xi^2 + a_0$ with respect to μ_t is the Hamilton flow $\chi(t,s)$ for $\tau + a(t,x,\xi)$,

$$\chi_0(t,s) = \mu_t \circ \chi(t,s) \circ \mu_s^{-1}$$

We note that $a \in \epsilon l^1 S^{(2)}$ iff $a_0 \in \epsilon l^1 S^{(2)}$. Hence from (xsxt 67) we obtain its counterpart for the χ_0 flow,

Proposition 32. *If $a_0 \in l^1 S^{(2)}_\epsilon$ with ϵ sufficiently small and $t > s$ then the Hamilton flow $\chi_0(t,s)$ has the Lipschitz regularity*

xsxt1
$$\frac{\partial(x_t,\xi_s)}{\partial(x_s,\xi_t)} = \begin{pmatrix} I_n + \epsilon O(1) & \epsilon O(\frac{1}{s}) \\ 2tI_n + \epsilon O(t) & I_n + \epsilon O(1) \end{pmatrix} \tag{86}$$

flowreg1

We proceed in a similar manner with b_0 and set

$$b(t,x,\xi) = b_0(t, x + 2t\xi, \xi)$$

Then the integral ψ_0 of b_0 along the χ_0 flow is the μ conjugate of the integral ψ of b along the χ flow. Hence we also trivially obtain the analog of Proposition psireg 27, namely

Proposition 33. *If $a_0 \in l^1 S^{(2)}_\epsilon$ with ϵ sufficiently small and $b_0 \in l^1 S^{(1)}$ then for $t > s$ we have*

psir1
$$\frac{\partial(\psi_0(x_t,\xi_t) - \psi_0(x_s,\xi_s))}{\partial(x_s,\xi_t)} = (O(s^{-\frac{1}{2}}), O(t^{\frac{1}{2}})) \tag{87}$$

psireg1

Now we can state our main result:

Theorem 34. *Let $a_0 \in l^1 S^{(2)}_\epsilon$, $b_0 \in l^1 S^{(2)}$ be real symbols with $b_0 \geq 0$ with ϵ sufficiently small. Then for $s < t$ the kernel K_0 of the operator $\tilde{S}_0(t,s)$ satisfies the bound*

kbd1 (88)

$$|K_0(t,x,\xi_t,s,x_s,\xi)| \lesssim t^{-\frac{n}{4}} s^{\frac{n}{4}} \left(1 + (\psi_0(x_s,\xi_s) - \psi_0(x_t,\xi_t))^2 + \frac{(x-x_t)^2}{t} + s(\xi-\xi_s)^2\right)^{-N}$$

tphase1

Proof. We use Theorem 28 via a conjugation with respect to the flat Schrödinger flow, whih corresponds to the cannonical transformations μ_t. Denote

$$S(t,s) = e^{-itD^2} S_0(t,s) e^{isD^2}$$

Then we compute

$$\frac{d}{dt} S(t,s) = e^{-itD^2}(a_0^w(t,x,D) - ib_0^w(t,x,D)) e^{itD^2} S(t,s)$$

Hence $S(t,s)$ is the evolution associated to the pseudodifferential operator

$$e^{-itD^2}(a_0^w(t,x,D) - ib_0^w(t,x,D)) e^{itD^2}$$

Using rescaled versions of Proposition (23),(24) his operator can be expressed in the form

$$a^w(t,x,D) - ib^w(t,x,D) + c^w(t,x,D)$$

where the remainder term satisfies $c \in l^1 S^{(0)}$. Hence the phace space kernel of $S(t,s)$ satisfies the bounds given by Theorem 28.

Returning to the original equation, for the phase space evolution $\tilde{S}_0(t,s)$ we can write

$$\begin{aligned}
\tilde{S}_0(t,s) &= T_{\frac{1}{t}} e^{itD^2} S(t,s) e^{-isD^2} T^*_{\frac{1}{s}} \\
&= T_{\frac{1}{t}} e^{itD^2} T^*_{\frac{1}{t}} T_{\frac{1}{t}} S(t,s) T^*_{\frac{1}{s}} T_{\frac{1}{s}} e^{-isD^2} T^*_{\frac{1}{s}} \\
&= (T_{\frac{1}{t}} e^{itD^2} T^*_{\frac{1}{t}}) \tilde{S}(t,s) (T_{\frac{1}{s}} e^{-isD^2} T^*_{\frac{1}{s}})
\end{aligned}$$

By a rescaled version of Proposition 22 the kernel of the first factor $T_{\frac{1}{t}} e^{itD^2} T^*_{\frac{1}{t}}$ is rapidly decreasing on the $t^{\frac{1}{2}} \times t^{-\frac{1}{2}}$ scale away from the graph of μ_t, while the kernel of the last factor $T_{\frac{1}{s}} e^{-isD^2} T^*_{\frac{1}{s}}$ is rapidly decreasing on the $s^{\frac{1}{2}} \times s^{-\frac{1}{2}}$ scale away from the graph of μ_s^{-1}. Hence the composition simply replaces the Hamilton flow associated to a by the Hamilton flow associated to a_0 and the function ψ with ψ_0 in the kernel bounds. Thus (69) implies (88), and the proof is concluded. $\square$

9. The parametrix construction

In this section we prove Proposition KO 10. We begin with a dyadic partition of the initial data with respect to the distance from the origin. At frequency 1 we consider a smooth partition of unit in the phase space

$$s_{-1}(\xi) + s_0(\xi) + s_1(\xi) = \sum_{\pm} \sum_{j \geq 0} p_j^{\pm}(x, \xi)$$

where $p_j^{\pm}$ have the support properties:

$$\operatorname{supp} p_j^{\pm} \subset \{2^{j-1} < |x| < 2^{j+1}, \quad \pm x\xi \geq -2^{-5}|x|\}$$

The signs $\pm$ correspond to waves which are outgoing forward, respectively backward in time.

In order for the corresponding pseudodifferential operators to preserve the frequency 1 localization we mollify these symbols in x, replacing

$$p_j^{\pm}(x, \xi) \to S_{<-10}(D_x) p_j^{\pm}(x, \xi)$$

This adds rapidly decreasing tails on the unit scale away from the initial localization region; however these tails play no role in the sequel so we simply disregard and keep the same notation for the symbols.

We construct the parametrix K_0 as a sum of the form

$$K_0(t, s) = \begin{cases} \displaystyle\sum_{j=1}^{\infty} S_j^-(t, s)(p_j^-)^w(x, D) & t < s \\ \displaystyle\sum_{j=1}^{\infty} S_j^+(t, s)(p_j^+)^w(x, D) & t > s \end{cases}$$

The following Proposition summarizes the properties of the K_j's:

Proposition 35. *Assume that ϵ is sufficiently small. Then for each $s \in \mathbb{R}$ there is an outgoing parametrix S_j^+ for $D_t + A_{(0)}$ in $\{t > s\}$ which is localized at frequency 1 and has the following properties:*

(i) L^2 bound:

$$\|S_j^+(t, s)\|_{L^2 \to L^2} \lesssim 1 \tag{89}$$

(ii) Error estimate:

l2errorj
$$\|x^{\alpha}(D_t + \tilde{A}_{(0)}) S_j^+(t, s) P_j^+\|_{L^2 \to L^2} \lesssim (2^j + |t - s|)^{-N} \tag{90}$$

(iii) Initial data:

$$S_j^+(s + 0, s) = I$$

(iv) Outgoing parametrix:

outl2j (91)
$$\|1_{\{|x|<2^{-10}(|t-s|+2^j)\}}S_j^+(t,s)P_j^+\|_{L^2\to L^2} \lesssim (|t-s|+2^j)^{-N}$$

(v) Finite speed:

finitel2j (92)
$$\|x^\alpha 1_{\{|x|>2^{10}(|t-s|+2^j)\}}S_j^+(t,s)P_j^+\|_{L^2\to L^2} \lesssim (|t-s|+2^j)^{-N}$$

(vi) Pointwise decay:

(93)
$$\|S_j^+(t,s)P_j^+\|_{L^1\to L^\infty} \lesssim (1+|t-s|)^{-\frac{n}{2}}$$

Kj

With obvious modifications the same holds for P_j^-. It is easy to verify that by summation this implies Proposition 10. We proceed to prove the above proposition.

We note that the hypothesis is translation invariant in time. In order to place ourselves in the context considered in Sections 7,8 we assume without any restriction in generality that $s=2^j$. By slightly increasing the ϵ_k's we can also assume that

eej (94)
$$\epsilon_j \approx \epsilon$$

The condition (8) insures that

$$a_{(0)}(x,\xi)-\xi^2 \in l^1 S_\epsilon^{(2)}, \qquad |x|\approx t,\ |\xi|\approx 1$$

We would like to have this satisfied for all x,ξ. This can be easily achieved due to properties (iv),(v) above. These allow us to truncate the coefficients $a_{(0)}^{ij}-\delta^{ij}$ at the expense of a negligible error in (90). Since we also seek a frequency localized parametrix, we can also freely modify the symbol $a_{(0)}(x,\xi)$ at higher frequencies. Thus we have reduced Proposition 35 to the case of an evolution governed by a symbol of the form

$$\xi^2 + a_0(t,x,\xi), \qquad a_0 \in l^1 S_\epsilon^{(2)}$$

In effect it is easy to see that the symbol a_0 can be chosen to have even better regularity,

extraa (95)
$$|\partial_x^\alpha \partial_\xi^\beta a_0(t,x,\xi)| \lesssim \epsilon(t)t^{-|\alpha|} \qquad |\alpha|\leq 2$$
$$|\partial_x^\alpha \partial_\xi^\beta a_0(t,x,\xi)| \lesssim \epsilon(t)t^{-1-\frac{|\alpha|}{2}} \qquad |\alpha|\geq 2$$

We note that we can weaken the requirement that the parametrix P_j^+ is frequency localized to bounds of the form

lowfreq (96)
$$\|x^\alpha P_{<-5}(D)S_j^+(t,s)P_j^+\|_{L^2\to L^2} \lesssim (|t-s|+2^j)^{-N}$$

highfreq (97)
$$\|x^\alpha \partial^\beta P_{>5}(D)S_j^+(t,s)P_j^+\|_{L^2\to L^2} \lesssim (|t-s|+2^j)^{-N}$$

Indeed, if P_j^+ satisfies these estimates then we can replace it by its frequency truncation

$$P_{-7<\cdot<7} P_j^+$$

at the expense of negligible errors in (90).

At this point we are already allowed to apply the result in Theorem 34. However, such a direct application would fail to give the rapid decay in time in (91), (92), (96), (97). To gain this decay we introduce an artificial damping term into the equation. Precisely, we define $S_j^+(t,s)$ to be the forward evolution operator associated to the equation

$$(i\partial_t + \Delta - a_0^w(t,x,D))u = -ib_0^w(t,x,D)u, \qquad u(2^j) = u_0 \tag{98}$$

where b_0 is a nonnegative symbol in $l^1 S^{(1)}$. The effect of the damping is to give the decay we want, but there is a price to pay, namely that we have to be able to estimate $b_0^w(t,x,D)u$ as an error term,

$$\|x^\alpha b_0^w(t,x,D) S_j^+(t,s) P_j^+\|_{L^2 \to L^2} \lesssim (|t-s| + 2^j)^{-N} \tag{99}$$

With this choice of $S_j^+(t,s)$ the properties (i), (iii) are trivial. Next we show how to prove the pointwise decay (vi). We consider three cases:

Case 1: $|t-s| \geq s$**.** For this we use the bounds in Theorem 34 and neglect the damping. Take $u(s) = \delta_0$. Then

$$Tu(s, x_s, \xi_s) = s^{-\frac{n}{4}} e^{-\frac{x_s^2}{2s}} e^{ix_s\xi_s}$$

By Theorem 34 we obtain

$$|Tu(t,x_t,\xi_t)| \lesssim t^{-\frac{n}{4}} \int (1 + t^{-1}(x_t - x_t(\xi_t, x_s))^2)^{-N} (1 + s(\xi_s - \xi_s(x_s,\xi_t))^2)^{-N} e^{-\frac{x_s^2}{2s}} dx_s d\xi_s$$

Integrating with respect to ξ_s we obtain

$$|Tu(t,x_t,\xi_t)| \lesssim t^{-\frac{n}{4}} s^{-\frac{n}{2}} \int (1 + t^{-1}(x_t - x_t(\xi_t, x_s))^2)^{-N} e^{-\frac{x_s^2}{2s}} dx_s d\xi_s$$

Since $x_t(\xi_t, x_s)$ is a Lipschitz function of x_s the integration with respect to x_s is also straightforward, and we obtain

$$|Tu(t,x_t,\xi_t)| \lesssim t^{-\frac{n}{4}} (1 + t^{-1}(x_t - x_t(\xi_t, 0))^2)^{-N}$$

Inverting the phase space transform we have

$$|u(t,y)| \lesssim t^{-\frac{n}{2}} \int (1 + t^{-1}(x_t - x_t(\xi_t,0))^2)^{-N} e^{-\frac{(x-x_t)^2}{2t}} dx_t d\xi_t$$

We integrate with respect to x_t,

$$|u(t,y)| \lesssim \int (1 + t^{-1}(y - x_t(\xi_t,0))^2)^{-N} d\xi_t$$

If $|t-s| \gtrsim s$ then the map $\xi_t \to x_t(\xi_t, 0)$ has inverse Lipschitz constant t^{-1}. Hence integration with respect to ξ_t yields

$$|u(t,y)| \lesssim t^{-\frac{n}{2}}$$

Case 2: $1 \leq |t-s| \leq s$**.** The difficulty we encounter here is that the $s^{\frac{1}{2}} \times s^{-\frac{1}{2}}$ phase space scale in our parametrix representation is too rough to allow $s^{\frac{1}{2}} \times s^{-\frac{1}{2}}$ packets starting together at time s to separate before time t. Hence instead of pointwise phase space bounds for the solution we would also have to exploit cancellations coming from stationary phase. To avoid this difficulty we reinitialize the time scale. For this we impose an additional condition on the regularity of b_0 for $|t-s| < 2^j$, namely

$$|\partial_x^\alpha \partial_\xi^\beta b_0(t,x,\xi)| \lesssim t^{-\frac{1}{2}-|\alpha|} \quad |\alpha| \leq 1$$

(100)

$$|\partial_x^\alpha \partial_\xi^\beta b_0(t,x,\xi)| \lesssim t^{-1-\frac{|\alpha|}{2}} \quad |\alpha| \geq 1$$

Together with (95) this implies that a_0 and b_0 have enough additional regularity so that the hypothesis of Theorem 34 remains valid after the time translation which resets the value of the initial time s to $t-s$. Then the above computation still applies in the new coordinates.

Case 3: $0 \leq |t-s| \leq 1$**.** This is the easiest, because our initial data is localized at frequency 1. Then we can use Sobolev embeddings combined with L^2 bounds,

$$\|u(t)\|_{L^\infty} \lesssim \sum_{|\beta|<N} \|\partial^\beta u(t)\|_{L^2} \lesssim \|P_j^+ u(0)\|_{L^2} \lesssim \|u(0)\|_{L^1}$$

For the rest of the proof we need to know more about the properties of b. We allow b_0 to vary between 0 and $t^{-\frac{3}{4}}$, and require that it satisfy the following properties:

(b1) $t^{\frac{3}{4}} b_0$ is nonincreasing along the Hamilton flow for $D_t + D_x^2 + a^w$, and

$$0 < t^{\frac{3}{4}} b_0(x_t, \xi_t) < 1 \implies b_0(x_{2t}, \xi_{2t}) = 0.$$

(b2) At the initial time we have

$$b(2^j, x, \xi) = 0 \qquad in \qquad \{2^{-3} < |\xi| < 2^3,\ 2^{j-2} < |x| < 2^{j+2},\ x\xi > -2^{-4}|x|\}$$

(b3) At any time $t \geq 2^j$ we have

$$b(t,x,\xi) = t^{-\frac{3}{4}} \qquad outside \qquad \{2^{-4} < |\xi| < 2^4\} \cap \{2^{-6}t < |x| < 2^6 t\}$$

The power $\frac{3}{4}$ is somewhat arbitrary, anything between $\frac{1}{2}$ and 1 works.

We verify that a symbol $b \in l^1 S^{(1)}$ which has the above properties leads to the correct conclusion.

Lemma 36. *Assume that the symbol $b \in l^1 S^{(1)}$ satisfies the properties (b1),(b2),(b3) above. Then the bounds (91), (92), (96), (97) and (99) hold.*

Proof. This proof is based on pointwise bounds for the phase space transform of the solution u to the equation

$$(i\partial_t + \Delta - a^w(t,x,D))u = -ib^w(t,x,D)u, \qquad u(2^j) = P_j^+ u_0$$

At the initial time $s = 2^j$ we know that the symbol of P_j^+ is supported in the region

$$D_s = \{2^{-2} < |\xi| < 2^2,\ 2^{j-1} < |x| < 2^{j+1},\ x\xi > -2^{-5}|x|\}$$

Then the phase space transform of $P_j^+ u_0$ decays rapidly outside this region,

$$|(T_{\frac{1}{s}} P_j^+ u_0)(x,\xi)| \lesssim (1 + d_s((x,\xi), D_s)^2)^{-N}$$

The phase space transform of $u(t)$ is expressed in terms of the phase space kernel $K_j^+(t,s)$ of $S_j^+(t,s)$, namely

$$(T_{\frac{1}{t}} u)(t,x,\xi) = \int K_j^+(t,x,\xi_t,s,x_s,\xi) T_{\frac{1}{s}} u(s,x_s,\xi) dx_s d\xi$$

For the phase space kernel we use the bounds in Theorem tphase1 34. Hence we obtain

$$|(T_{\frac{1}{t}} u)(t,x,\xi_t)| \lesssim s^{\frac{n}{4}} t^{-\frac{n}{4}} \int \left(1 + \frac{(x-x_t)^2}{t}\right)^{-N} \left(1 + s(\xi-\xi_s)^2 + (\psi(x_t,\xi_t) - \psi(x_s,\xi_s)^2\right)^{-N}$$
$$(1 + d_s((x_s,\xi), D_s)^2)^{-N} dx_s d\xi$$

The integration with respect to ξ is trivial. Also, by (xsxt 67), the variables x_s and x_t are in a bilipschitz correspondence. Hence we obtain

$$|(T_{\frac{1}{t}} u)(t,x,\xi_t)| \lesssim s^{-\frac{n}{4}} t^{+\frac{n}{4}} \int t^{-\frac{n}{2}} \left(1 + \frac{(x-x_t)^2}{t}\right)^{-N} W(x_t,\xi_t) dx_t$$

where

$$W(x_t,\xi_t) = \left(1 + (\psi(x_t,\xi_t) - \psi(x_s,\xi_s))^2\right)^{-N} \left(1 + d_s((x_s,\xi_s), D_s)^2\right)^{-N}$$

If we simply bound W by 1 then this gives the global estimate

decay1 (101)
$$|(T_{\frac{1}{t}} u)(t,x_t,\xi_t)| \lesssim \left(\frac{t}{s}\right)^{\frac{n}{4}}$$

On the other hand, we claim it also implies a rapid decay in the support of b, more precisely

decay2 (102)
$$|(T_{\frac{1}{t}} u)(t,x_t,\xi_t)| \lesssim \left(\frac{t}{s}\right)^{\frac{n}{4}} (t^{\frac{1}{2}} + t^{-2}|x_t|^2 + |\xi_t|^2)^{-N} (1 + d_t((x_t,\xi_t), \text{supp } b(t))^2)^N$$

It suffices to prove that W satisfies a similar bound,

decay3 (103)
$$|W(t,x_t,\xi_t)| \lesssim \left(\frac{t}{s}\right)^{\frac{n}{4}} (t^{\frac{1}{2}} + t^{-2}|x_t|^2 + |\xi_t|^2)^{-N} (1 + d_t((x_t,\xi_t), \text{supp } b(t))^2)^N$$

We begin with the simpler case when (x_t, ξ_t) lies inside the support of b, i.e. $b(t, x_t, \xi_t) > 0$. Then $b(s, x_s, \xi_s) > 0$, which by (b2) implies that

$$d_s((x_s, \xi_s), D_s) \gtrsim s^{\frac{1}{2}}$$

On the other hand we have

$$|\xi_s| \approx |\xi_t|, \qquad |x_s - x_t| \lesssim t|\xi_t|$$

which give

$$1 + d_s((x_s, \xi_s), D_s) \gtrsim |\xi_t| + t^{-1}|x_t|$$

Summing up the two bounds we get

$$d_s((x_s, \xi_s), D_s) \gtrsim s^{\frac{1}{2}} + |\xi_t| + t^{-1}|x_t| \tag{104}$$

This suffices for (103) provided that $t < 4s$. For larger t we use also (b1) to conclude that

$$b(\sigma, x_\sigma, \xi_\sigma) = \sigma^{-\frac{3}{4}} \qquad s \leq \sigma \leq t/2.$$

Then

$$\psi(x_t, \xi_t) - \psi(x_s, \xi_s) \gtrsim t^{\frac{1}{4}}$$

which together with (104) leads again to (103).

Next we consider the case when $(x_t, \xi_t) \notin \text{supp } b(t)$. Then $|x_t| \approx t$ and $|\xi_t| \approx 1$. Due to (101) we can assume without any restriction in generality that

$$d_t((x_t, \xi_t), \text{supp } b(t)) \ll t^{-\frac{1}{4}} \tag{105}$$

We consider two cases as before. If $t < 4s$ this implies that

$$d_s((x_s, \xi_s), \text{supp } b(s)) \ll s^{-\frac{1}{4}}$$

therefore

$$d_s((x_s, \xi_s), D_s) \gtrsim s^{\frac{1}{2}}$$

and (103) follows.

If $t > 4s$ then we need a more detailed analysis of the Lipschitz regularity of the decay factor $\psi(x_t, \xi_t) - \psi(x_s, \xi_s)$. By (b1), the symbol $t^{\frac{1}{2}} b_1$ is either constant along a bicharacteristic, or it changes from 1 to 0 with the transition occuring within a single dyadic time interval. Suppose this interval is around time $\sigma > 2s$. Then on one hand we have

$$\psi(x_t, \xi_t) - \psi(x_s, \xi_s) \approx \sigma^{\frac{1}{4}}$$

while, on the other hand, (87) leads to the better bounds

$$\frac{\partial(\psi(x_t, \xi_t) - \psi(x_s, \xi_s))}{\partial(x_t, \xi_t)} = (O(\sigma^{-\frac{1}{2}}, O(t^{\frac{1}{2}}))$$

which we rewrite in the form

$$|\nabla(\psi(x_t,\xi_t)-\psi(x_s,\xi_s))|_{g_t} \lesssim t^{\frac{1}{2}}\sigma^{-\frac{1}{2}}$$

We summarize this result as

$$\boxed{\texttt{ode}}\ (106)\quad \psi(x_t,\xi_t)-\psi(x_s,\xi_s) \gtrsim s^{\frac{1}{4}} \implies |\nabla(\psi(x_t,\xi_t)-\psi(x_s,\xi_s))|_{g_t} \lesssim t^{\frac{1}{2}}(\psi(x_t,\xi_t)-\psi(x_s,\xi_s))^{-2}$$

Suppose (105) holds. Then we can find a second bicharacteristic $\sigma \to (\tilde{x}_\sigma, \tilde{\xi}_\sigma)$ with

$$b(t,\tilde{x}_t,\tilde{\xi}_t) > 0, \qquad d_t((x_t,\xi_t),(\tilde{x}_t,\tilde{\xi}_t)) \ll t^{-\frac{1}{4}}$$

The first relation implies that

$$\psi(\tilde{x}_t,\tilde{\xi}_t)-\psi(\tilde{x}_s,\tilde{\xi}_s) \approx t^{\frac{1}{4}}$$

The the second relation combined with (106) gives

$$\psi(x_t,\xi_t)-\psi(x_s,\xi_s) \approx t^{\frac{1}{4}}$$

and concludes the proof of (103).

Together, the bounds (101) and (102) imply (91), (92), (96) and (97) since the support of b contains the localization regions in each of these bounds.

To prove (99) we use the properties of the phase space kernel of b^w, given by Proposition 17. This gives

$$\begin{aligned}|T_{\frac{1}{t}}b^w(t,x,D)u(t,x,\xi)| \lesssim& (b(t,x,\xi)+(1+d_t((x,\xi),\text{supp }b(t))^2)^{-2N}\int(1+d_t((x,\xi),(y,\eta))^2)^{-2N}\\ &(t+t^{-2}|y|^2+|\eta|^2)^{-N}(1+d_t((y,\eta),\text{supp }b(t))^2)^N dyd\eta\\ \lesssim&(t+t^{-2}|x|^2+|\xi|^2)^{-N}(1+d_t((x,\xi),\text{supp }b(t))^2)^{-N}\end{aligned}$$

Then (99) follows.

$\square$

It remains to construct a symbol b with the desired properties.

Lemma 37. *There exists a symbol $b \in l^1S^{(1)}$ which satisfies the properties (b1), (b2) and (b3) and has the additional regularity (100).*

Proof. Let ϕ be a smooth cutoff function which equals 0 in $(-\infty,0)$ and 1 in $(1,\infty)$. We define the symbol b as

$$b(t,x,\xi) = t^{-\frac{3}{4}}(1-\phi(b_1)\phi(b_2)\phi(b_3)\phi(b_4)\phi(b_5))$$

Here $\phi(b_1)$ selects the frequencies which are not too large,

$$b_1(t,\xi) = \frac{2^4+Ce(t)-|\xi|}{\epsilon(t)}$$

where C is a fixed large constant. The symbol $\phi(b_2)$ selects the frequencies which are not too small,

$$b_2(t,\xi) = \frac{|\xi| - 2^{-4} + Ce(t)}{\epsilon(t)}$$

$\phi(b_3)$ selects the outgoing waves,

$$b_3(t,x,\xi) = \frac{2^{-3}|x||\xi| + x\xi}{2^{-12}|x|}.$$

Finally $\phi(b_4)$ selects a spatial region not too far from the origin

$$b_4(t,x) = \frac{2^6 t - |x|}{t}$$

while $\phi(b_5)$ selects a spatial region not too close too the origin

$$b_5(t,x,\xi) = \frac{|x||\xi| - 2^{-5}t|\xi| + x\xi}{2^{-10}t}.$$

We note that

$$\{2^{-3} < |\xi| < 2^3\} \cap \{2^{-5}t < |x| < 2^5 t\} \cap \{x\xi > -2^{-4}|x|\} = D_t \subset \{b = 0\}$$

while

$$\{b < 1\} \subset E_t = \{2^{-4} < |\xi| < 2^4\} \cap \{2^{-6}t < |x| < 2^6 t\} \cap \{x\xi > -2^{-3}|x|\}$$

so the conditions (b2) and (b3) are easily satisfied.

To prove (b1) it suffices to study the behavior of b along the Hamilton flow within E_t and show that for each b_j we have

bj (107)
$$\frac{d}{dt} b_j(t,\xi_t) \approx \frac{2}{t} \qquad in\ E_t \cap \{0 \leq b_j \leq 1\}$$

For $(x_t,\xi_t) \in E_t$ we have

$$\frac{d}{dt}\xi_t = O\left(\frac{\epsilon(t)}{t}\right), \qquad \frac{d}{dt}x_t = 2\xi_t + O(\epsilon(t))$$

Then an easy computation shows that

$$\frac{d}{dt} b_1(t,\xi_t) \approx \frac{1}{\epsilon t} \qquad in\ \{0 \leq b_1 \leq 1\}$$

with the main contribution coming from the derivative of $e(t)$. The computation for b_2 is identical. For b_3 we have

$$\frac{d}{dt} b_3(t,x_t,\xi_t) = \frac{2^{-2}(\xi_t^2 x_t^2 - (x_t\xi_t)^2)}{2^{-12}|x|^3} + \frac{O(\epsilon(t))}{t} \geq \frac{2}{t} \qquad in\ E_t \cap \{0 \leq b_3 \leq 1\}$$

For b_4 we compute

$$\frac{d}{dt} b_4(t,x_t) = \frac{|x_t|^2 + 2t x_t \xi_t}{t^2|x_t|} + \frac{O(\epsilon(t))}{t} > \frac{2^5}{t} \qquad in\ E_t \cap \{0 \leq b_4 \leq 1\}$$

Finally for b_5 we also estimate

$$\begin{aligned}\frac{d}{dt}b_5(t,x_t,\xi_t) =&\frac{2|x_t|^{-1}|\xi_t|x_t\xi_t+2\xi_t^2}{2^{-10}t}-\frac{|x_t||\xi_t|+x_t\xi_t}{2^{-10}t^2}+\frac{O(\epsilon(t))}{t}\\ \geq&\frac{\xi_t^2}{2^{-10}t}-\frac{2^{-5}|\xi_t|}{2^{-10}t}\\ \geq&\ \frac{2}{t}\end{aligned}$$

within the set $E_t \cap \{0 \leq b_5 \leq 1\}$. □

References

BT [1] Jean-Marc Bouclet and Nikolay Tzvetkov. Strichartz estimates for long range perturbations. preprint.

MR2066943 [2] N. Burq. Smoothing effect for Schrödinger boundary value problems. *Duke Math. J.*, 123(2):403–427, 2004.

MR2058384 [3] N. Burq, P. Gérard, and N. Tzvetkov. Strichartz inequalities and the nonlinear Schrödinger equation on compact manifolds. *Amer. J. Math.*, 126(3):569–605, 2004.

MR2106340 [4] Nicolas Burq, Fabrice Planchon, John G. Stalker, and A. Shadi Tahvildar-Zadeh. Strichartz estimates for the wave and Schrödinger equations with potentials of critical decay. *Indiana Univ. Math. J.*, 53(6):1665–1680, 2004.

MR93i:35010 [5] Jean-Marc Delort. *F.B.I. transformation.* Springer-Verlag, Berlin, 1992. Second microlocalization and semilinear caustics.

MR1795567 [6] Shin-ichi Doi. Smoothing effects for Schrödinger evolution equation and global behavior of geodesic flow. *Math. Ann.*, 318(2):355–389, 2000.

MR85f:35001 [7] Charles L. Fefferman. The uncertainty principle. *Bull. Amer. Math. Soc. (N.S.)*, 9(2):129–206, 1983.

MR92k:22017 [8] Gerald B. Folland. *Harmonic analysis in phase space.* Princeton University Press, Princeton, NJ, 1989.

MR1151250 [9] J. Ginibre and G. Velo. Smoothing properties and retarded estimates for some dispersive evolution equations. *Comm. Math. Phys.*, 144(1):163–188, 1992.

MR2131050 [10] Andrew Hassell, Terence Tao, and Jared Wunsch. A Strichartz inequality for the Schrödinger equation on nontrapping asymptotically conic manifolds. *Comm. Partial Differential Equations*, 30(1-3):157–205, 2005.

MR1646048 [11] Markus Keel and Terence Tao. Endpoint Strichartz estimates. *Amer. J. Math.*, 120(5):955–980, 1998.

MR2094851 [12] Herbert Koch and Daniel Tataru. Dispersive estimates for principally normal pseudodifferential operators. *Comm. Pure Appl. Math.*, 58(2):217–284, 2005.

RZ [13] L. Robbiano and C. Zuily. Strichartz estimates for the schrodinger equation with variable coefficients. Preprint.

RT [14] Igor Rodnianski and Terence Tao. Long time decay estimates for the schrödinger equation on manifolds. preprint.

Sj1 [15] Johannes Sjöstrand. Singularités analytiques microlocales. In *Astérisque, 95*, pages 1–166. Soc. Math. France, Paris, 1982.

MR1644105 [16] Hart F. Smith. A parametrix construction for wave equations with $C^{1,1}$ coefficients. *Ann. Inst. Fourier (Grenoble)*, 48(3):797–835, 1998.

MR1924470 [17] Gigliola Staffilani and Daniel Tataru. Strichartz estimates for a Schrödinger operator with nonsmooth coefficients. *Comm. Partial Differential Equations*, 27(7-8):1337–1372, 2002.

nlw [18] Daniel Tataru. Strichartz estimates for operators with nonsmooth coefficients and the nonlinear wave equation. *Amer. J. Math.*, 122(2):349–376, 2000.

cs [19] Daniel Tataru. Strichartz estimates for second order hyperbolic operators with nonsmooth coefficients. II. *Amer. J. Math.*, 123(3):385–423, 2001.

MR1944027 [20] Daniel Tataru. On the Fefferman-Phong inequality and related problems. *Comm. Partial Differential Equations*, 27(11-12):2101–2138, 2002.

phasespace [21] Daniel Tataru. Phase space transforms and microlocal analysis. In *Phase space analysis of partial differential equations. Vol. II*, Pubbl. Cent. Ric. Mat. Ennio Giorgi, pages 505–524. Scuola Norm. Sup., Pisa, 2004.

Department of Mathematics, University of California at Berkeley