

# DISPERSIVE ESTIMATES FOR PRINCIPALLY NORMAL PSEUDODIFFERENTIAL OPERATORS

HERBERT KOCH AND DANIEL TATARU

## 1. INTRODUCTION

Dispersive estimates are  $L^q$  estimates for nonelliptic partial differential operators which are a consequence of the decay properties of their fundamental solutions. These decay properties follow from spatial spreading of the singularities of the solutions. Since solutions propagate in directions conormal to the characteristic set of the operator, this spreading can be related to nonzero curvatures of the characteristic set. Dispersive estimates for constant coefficient operators are closely related to the restriction theorem in harmonic analysis.

Various types of dispersive estimates are known to be true for operators such as the wave operator, the Schrödinger operator and the linear KdV, see Ginibre-Velo [4], Keel-Tao [11]. They have proved to be useful in the study of nonlinear problems, as well as of problems with unbounded potentials.

More recently, similar estimates have been obtained for wave operators with variable coefficients, beginning with the smooth case in Kapitanskii [10], Mockenhaupt, Seeger and Sogge [14], up to operators with  $C^2$  coefficients in Smith [15] and Tataru [21], [23]. Similar results were obtained for the Schrödinger equation in Staffilani-Tataru [19] ( $C^2$  coefficients) and in Burq-Gerard-Tzvetkov [1] (smooth coefficients). In the variable coefficient elliptic case one should also mention Sogge's  $L^q$  eigenfunction bounds on compact manifolds, see [18].

All the above examples are operators with real symbols. On the other hand, in unique continuation problems one is interested in Carleman estimates. These are uniform weighted estimates with respect to a family of exponential weights,

$$\|e^{\tau\phi}u\|_{L^q} \leq \|e^{\tau\phi}P(x, D)u\|_{L^r}, \quad \tau > \tau_0$$

With the substitution  $v = e^{\tau\phi}u$  these can be rewritten as

$$\|v\|_{L^q} \leq \|P(x, D + i\tau\nabla\phi)v\|_{L^r}, \quad \tau > \tau_0$$

Even if  $P$  has constant coefficients and real symbol, the conjugated operator

$$P_\phi = P(x, D + i\tau\nabla\phi)$$

will have complex symbol and variable coefficients. In order for any such estimates to hold  $P_\phi$  must satisfy a so-called pseudoconvexity condition. It is known that such a condition implies  $L^2$  estimates. However,  $L^r \rightarrow L^q$  estimates are considerably more difficult to obtain. The first results in this direction were obtained in Jerison-Kenig [9] for the Laplacian with a polynomial weight. Later these were extended by Sogge [17] to second order elliptic operators with smooth coefficients. Further work of Wolff [25] and of the authors [13] addresses also the case of Lipschitz coefficients, and  $L^p$  gradient potentials.

To this one should add the work of Kenig-Ruiz-Sogge [12] for second order constant coefficient operators, and the work of Sogge [16] for parabolic operators with smooth coefficients. The counterpart of the Jerison-Kenig estimates for second order constant coefficient parabolic operators is proved in Escauriaza [3].

All of the above mentioned results take advantage of the special form of the operator in one way or another. On the other hand, it is clear that only the geometry of the characteristic set should matter.

Motivated by problems in unique continuation and in local solvability, in the present article we consider the problem of obtaining dispersive estimates for operators which are principally normal. However, of independent interest is our parametrix construction for principally normal operators, as well as the corresponding pointwise estimates for the kernel of the parametrix. We only make assumptions on the geometry of the characteristic set, and we also seek to use minimal regularity for the symbols/coefficients.

An obstacle in applying our results to obtain Carleman estimates for unique continuation problems is that the conjugated operator  $P_\phi$  introduced above does not satisfy the principal normality condition. Fortunately the  $L^2$  estimates which follow from the pseudoconvexity condition are strong enough so that they allow spatial localization on a much smaller scale. This scale turns out to be precisely the largest scale on which the principal normality survives.

To give the reader some idea of the results we obtain in this article, we present some very simple examples. All these examples have constant coefficients, however our results apply as well to operators with variable coefficients. In what follows  $u$  is supported in the unit ball in  $\mathbb{R}^n$  and  $\lambda > 1$ . In many cases the support restriction is easily removed by scaling.

- As a consequence of Theorem 2 we have

$$\|u\|_{L^{\frac{2(n+1)}{n-1}}} \lesssim \lambda^{-\frac{2}{n+1}} \|(\Delta + \lambda^2)u\|_{L^{\frac{2(n+1)}{n+3}}}.$$

- By Theorem 3 for any differential operator  $Q(D)$  with constant coefficients and real symbol we have

$$\|u\|_{L^{\frac{2(n+1)}{n-1}}} \lesssim \lambda^{-\frac{2}{n+1}} \|(\Delta + \lambda^2 + i\lambda^2 Q(D/\lambda))u\|_{L^{\frac{2(n+1)}{n+3}}}.$$

- Theorem 4 implies

$$\|u\|_{L^{\frac{2(n+2)}{n}}} \lesssim \lambda^{-\frac{4}{n+2}} \|(\Delta + \lambda^2 + \lambda\partial_1)u\|_{L^{\frac{2(n+2)}{n+4}}}.$$

- By Theorem 5, for  $|\delta| \leq 1$

$$\|u\|_{L^{\frac{2(n+2)}{n}}} \lesssim \delta^{-\frac{2}{n+2}} \lambda^{-\frac{4}{n+2}} \|(\Delta + \lambda^2 + \delta\lambda\partial_1)u\|_{L^{\frac{2(n+2)}{n+4}}}.$$

These results are slightly stronger than stated in our theorems. However, there are obvious  $L^2$  estimates for  $\Delta + \lambda^2$  which allow to improve the estimates to the form stated above.

A new obstacle which we face in this analysis is that for principally normal operators the characteristic set is a codimension 2 manifold. Its curvature properties are not as easy to describe as in the codimension 1 case. The propagation no longer occurs along rays, but instead along two dimensional surfaces in the phase space. Even in the case when the curvature of the codimension 2 characteristic set is “non-degenerate”, the spatial projections of these two dimensional surfaces through a point must overlap substantially in some directions. Thus, unlike in the case of operators with real symbols, there is no hope to obtain uniformly strong kernel decay estimates in all directions for the parametrix. Instead, there will always be a lower dimensional set of directions with a weaker kernel decay, and the geometry of this set can be quite intricate. Two extreme examples are  $-\Delta_{\mathbb{R}^n} - 1 + iD_1$ , respectively  $D_2 - D_1^2 + i(D_3 - D_1^2)$ . In the first case the characteristic set is the intersection of the sphere  $|\xi| = 1$  with the plane  $\xi_1 = 0$ ; all two dimensional planes which are normal to it intersect on a line, where the kernel decays like  $|x|^{-1}$  compared to  $|x|^{-n/2}$  in the other directions. In the second case, the characteristic set is the curve  $\xi_1 \rightarrow \gamma(\xi_1) = (\xi_1, \xi_1^2, \xi_1^3)$ , which has nonzero curvature and torsion. This time the directions of bad decay of are those perpendicular to both  $\dot{\gamma}$  and  $\ddot{\gamma}$ ; they are spread on a cone, but in exchange the kernel is not as large in those directions as in the first case. The generic decay of the low frequency part of the parametrix is like  $|x|^{-3/2}$  compared to  $|x|^{-4/3}$  in the bad directions.

Fortunately we are able to produce a factorization of the parametrix which allows us to establish the dispersive estimates without having to

study the kernel decay in the bad region. Instead, all we need is to prove this decay in the good region, where focusing does not occur. Some related results were independently proved by Dos Santos Ferreira [2]. However, his estimates are somewhat weaker as they are based on the worst decay rate instead of the generic one.

Another feature of this work is that we seek to obtain our estimates with minimal regularity assumptions on the coefficients. Precisely, in the case of the principally normal operators it turns out that we need to have  $C^2$  coefficients. This gets even better for some of the applications to unique continuation, where we have better spatial localization.

The structure of the article is as follows. In the next section we state our assumptions and results in a dyadic setting. The advantage of doing it this way is that there is more than one interesting case in which the dyadic results can be applied. In Section 3 we use elliptic arguments to reduce the problems to some canonical formulation.

Next we consider a set of increasingly complex problems. We begin in Section 4 with the case of operators with real symbols. First we construct a wave packet type representation of the fundamental solution, then we use curvature assumptions to prove pointwise bounds for it. This in turn yields the dispersive estimates.

In Section 5 we turn to the construction of parametrices in the symplectic case. This has an operator theory flavor but is also borrows some ideas from the Littlewood-Paley theory. Our construction and the  $L^2$  type estimates for the parametrix apply in an abstract setup for operators

$$D_t - A(t) + iB(t)$$

where  $A(t)$ ,  $B(t)$  are selfadjoint operators in a Hilbert space, satisfying the fixed time commutator estimate

$$\|[D_t - A, B]u\| \lesssim \|Bu\| + \|u\|$$

In Section 6 we combine these  $L^2$  bounds with the dispersive estimates for operators with real symbols. As a first consequence we obtain the same results for principally normal operators as for the selfadjoint part. The second case we consider is the involutive case, when both the real and the imaginary part of the symbols are of principal type, with transversal characteristic sets. The dispersive estimates follow from bounds for operators with real symbols and the  $L^2$  bounds for the parametrix. Then we study the degenerate involutive case, where the imaginary part is still of principal type but small, say of size  $\delta$ . In this case we use the wave packet representation of the fundamental solution in the real case to derive a similar representation for the parametrix. This gives pointwise kernel decay in the good directions,

while in the bad decay directions we fall back on the approach for non-degenerate operators. We obtain estimates with a sharp dependence of the constants on  $\delta$ .

The first application we consider is to local solvability problems. Principally normal operators are known to be locally solvable with loss of one derivative. Here we consider instead principally normal operators with unbounded potentials, and use dispersive estimates in order to prove similar results.

The last part of the article is devoted to applications to unique continuation problems; more precisely, we obtain  $L^q$  Carleman estimates for elliptic and parabolic operators with  $C^1$  coefficients, and for (non-elliptic) principally normal operators with  $C^2$  coefficients. Our strategy is as follows. On the unit spatial scale we use only on the  $L^2$  Carleman estimates; these allow us to localize the  $L^q$  estimates on a much smaller spatial scale, see Lemmas 8.4, 8.5. On the smaller spatial scale we are able to use the dispersive estimates for principally normal operators.

## 2. THE FIXED FREQUENCY RESULTS

For  $\lambda > 1$  and  $j = 0, 1, \dots$  we consider a class of symbols, denoted by  $S_\lambda^j$ , which satisfy the conditions

$$(1) \quad \begin{aligned} |\partial_x^\alpha \partial_\xi^\beta a(x, \xi)| &\leq c_{\alpha, \beta} \lambda^{-|\beta|} & |\alpha| \leq j \\ |\partial_x^\alpha \partial_\xi^\beta a(x, \xi)| &\leq c_{\alpha, \beta} \lambda^{\frac{|\alpha| - j}{2} - |\beta|} & |\alpha| \geq j \end{aligned}$$

The parameter  $\lambda$  plays the role of the frequency. The symbols in  $S_\lambda^j$  are bounded, and the corresponding operators are  $L^2$  bounded. Both the space of symbols  $S_\lambda^j$  and the corresponding class of pseudodifferential operators are algebras. One can see that the first  $j$  derivatives of the symbols behave as for  $S_{1,0}$  symbols, while the rest are as for  $S_{\frac{1}{2},1}$  symbols. This suffices in order for the usual calculus of pseudodifferential operators to apply. In this article we use the Weyl calculus, but this is not essential.

Later on we use symbols in  $S_\lambda^j$  to describe the frequency  $\lambda$  part of order 0 pseudodifferential operators. For operators of order  $k \neq 0$  we use the notation

$$\lambda^k S_\lambda^j = \{\lambda^k a; a \in S_\lambda^j\}$$

and we denote the corresponding class of operators by  $\lambda^k OPS_\lambda^j$ . All estimates in this paper require only control of a finite number of derivatives of the symbols. We do not keep track of the number of derivatives needed. Nevertheless we can and do consider the symbol spaces and

the spaces of operators as Banach spaces. The statement that a certain operator in  $\lambda^k OPS_\lambda^j$  is bounded as a linear map from  $L^q$  to  $L^r$  means that its operator norm is controlled by the norm of its symbol and possibly by the constants defined in the assumptions below. It is however crucial to obtain the correct dependence of the constants on  $\lambda$  and possibly other parameters.

Given real symbols  $p_{re}, p_{im} \in \lambda S_\lambda^2$  and  $0 \leq \delta \leq 1$  we seek estimates from below for the operator

$$p^w(x, D) = p_{re}^w(x, D) + i\delta p_{im}^w$$

of the form

$$\|\chi^w u\|_{L^r} \leq c_1 \|p^w u\|_{L^q} + c_2 \|u\|_{L^2}$$

combining various pairs of exponents  $q$  and  $r$ . The constants  $c_1, c_2$  possibly depend on  $\lambda, \delta$ , finitely many of the constants  $c_{\alpha\beta}$  in (1), and on the geometry of the real symbols  $p_{re}$  and  $p_{im}$  of the symbol in a quantitative way which will be described below. To keep the notation concise we introduce the following

**Definition 2.1.** *Given a Banach space  $X$  and  $\rho > 0$  we denote by  $\rho X$  the same vector space equipped with the norm*

$$\|u\|_{\rho X} = \rho^{-1} \|u\|_X$$

*Thus the unit ball in  $\rho X$  is  $\rho$  times the unit ball in  $X$ .*

The main condition which connects  $p_{re}$  and  $p_{im}$  is a principal normality condition:

(A1) The operator  $p^w$  is principally normal, i.e.

$$(2) \quad |\{p_{re}, p_{im}\}| \lesssim |p_{re}| + |p_{im}| + 1$$

As proved in [22], this condition guarantees that one has good  $L^2$  estimates from below for  $p^w$  even for  $p$  in a larger class than  $\lambda S_\lambda^2$ :

**Theorem 1.** *Let  $p = p_{re} + ip_{im}$  be a symbol which satisfies*

$$|\partial_x^\alpha \partial_\xi^\beta p(x, \xi)| \leq c_{\alpha, \beta} \lambda^{\frac{|\alpha| - |\beta|}{2}} \quad |\alpha| \geq 2$$

*If (A1) holds then*

$$(3) \quad \|p_{re}^w u\|_{L^2}^2 + \|p_{im}^w u\|_{L^2}^2 \lesssim \|p^w u\|_{L^2}^2 + \|u\|_{L^2}^2$$

Here we are interested in obtaining  $L^r$  estimates, and in order to do this we need to restrict ourselves to a region of spatial size 1 and frequency size  $\lambda$ . After a translation we can assume this is centered at the origin, therefore we set

$$B_\lambda = \{(x, \xi); |x| < 1, |\xi| < \lambda\}$$

We let  $\chi \in S_\lambda^0$  be a symbol which is compactly supported in  $B_\lambda$ . For comparison we consider first the case when  $p$  is elliptic, i.e.  $p \in \lambda S_\lambda^0$ ,  $p^{-1} \in \lambda^{-1} S_\lambda^0$ .

**Proposition 2.2.** *a) If  $a \in S_\lambda^0$  is supported in  $B_\lambda$  and  $1 \leq r \leq q \leq \infty$  then*

$$\|a^w u\|_{L^q} \lesssim \lambda^{\frac{n}{r} - \frac{n}{q}} \|u\|_{L^r}.$$

*b) If  $p \in S_\lambda^0$  is elliptic then*

$$\|\chi^w u\|_{L^q} \lesssim \lambda^{\frac{n}{r} - \frac{n}{q} - 1} \|p^w u\|_{L^r} + \lambda^{-N} \|u\|_{L^2}, \quad N > 0$$

For principally normal operators  $p \in \lambda S_\lambda^2$  we seek to prove estimates of the form

$$(4) \quad \|\chi^w u\|_{\lambda^{\rho(q)} L^q} \lesssim \|p^w u\|_{L^2 + \lambda^{-\rho(q)} L^{q'}} + \|u\|_{L^2}$$

The exponent  $\rho$  is chosen as in the above elliptic estimates,

$$\rho(q) = \frac{n-1}{2} - \frac{n}{q}$$

In many examples this relation can also be derived from scaling considerations. If (4) holds for some  $q$  then it is easy to see that it must hold for all larger  $q$ . However, if  $p$  is not elliptic then (4) cannot hold for  $q = 2$ . Our goal is to find the lowest value of  $q$  for which (4) holds.

A dual form of (4) applies to the adjoint operator  $(p^w)^* = \bar{p}^w$ . However, since the class of operators we work with is invariant with respect to taking adjoints, it suffices to state it for  $p^w$ . We seek a parametrix  $K$  for the operator  $P$  with the following properties:

$$(5) \quad \|Kf\|_{L^2} + \|K\chi^w f\|_{\lambda^{\rho(q)} L^q} \lesssim \|f\|_{L^2 + \lambda^{-\rho(q)} L^{q'}}$$

$$(6) \quad \|(p^w K - I)\chi^w f\|_{L^2} \lesssim \|f\|_{L^2 + \lambda^{-\rho(q)} L^{q'}}$$

We also want to obtain mixed norm estimates. For this we split the coordinates  $x = (x_1, x')$  and we use the notation  $L^q L^r = L_{x_1}^q L_{x'}^r$ . Then we seek estimates of the form

$$(7) \quad \|\chi^w u\|_{\lambda^{\rho(q_1, r_1)} L^{q_1} L^{r_1}} \lesssim \|p^w u\|_{L^2 + \lambda^{-\rho(q_2, r_2)} L^{q'_2} L^{r'_2}} + \|u\|_{L^2}$$

Again the exponent  $\rho$  is determined by comparison with elliptic operators,

$$\rho(q, r) = \frac{n-1}{2} - \frac{1}{q} - \frac{n-1}{r}$$

Correspondingly, we want a parametrix  $K$  with the mapping properties

$$(8) \quad \|K\chi^w f\|_{L^2 \cap \lambda^{\rho(q_1, r_1)} L^{q_1} L^{r_1}} \lesssim \|f\|_{L^2 + \lambda^{-\rho(q_2, r_2)} L^{q'_2} L^{r'_2}}$$

$$(9) \quad \|(p^w K - I)\chi^w f\|_{L^2} \lesssim \|f\|_{L^2 + \lambda^{-\rho(q_2, r_2)} L^{q'_2} L^{r'_2}}$$

Most of the work in this article is devoted to the construction of parametrices for principally normal operators. The parametrix bounds can be easily connected to the direct estimates using the following result.

**Proposition 2.3.** *Assume that there is a parametrix  $K$  for  $\bar{p}^w$  which satisfies (8) and (9). Then (7) holds with  $(q_1, r_1)$  and  $(q_2, r_2)$  interchanged. In particular, if there is a parametrix for  $\bar{p}^w$  satisfying (5) and (6) then (4) holds for  $p^w$ .*

*Proof.* Let  $K$  be the parametrix for  $\bar{p}^w$ . Given  $g \in \lambda^{-\rho(q_1, r_1)} L^{q'_1} L^{r'_1}$  we consider the decomposition

$$\chi^w g = g_1 + \bar{p}^w v, \quad v = K \chi^w g$$

The estimates (8) and (9) for  $\bar{p}^w$  show that

$$\|g_1\|_{L^2} + \|v\|_{L^2 \cap \lambda^{\rho(q_2, q_2)} L^{q_2} L^{r_2}} \lesssim \|g\|_{\lambda^{-\rho(q_1, r_1)} L^{q'_1} L^{r'_1}}.$$

Now we compute

$$\langle \chi^w u, g \rangle = \langle u, \chi^w g \rangle = \langle u, g_1 \rangle + \langle u, \bar{p}^w v \rangle = \langle u, g_1 \rangle + \langle p^w u, v \rangle$$

Using the previous inequality we estimate

$$|\langle \chi^w u, g \rangle| \lesssim (\|u\|_{L^2} + \|p^w u\|_{L^2 + \lambda^{-\rho(q_2, q_2)} L^{q'_2} L^{r'_2}}) \|g\|_{\lambda^{-\rho(q_1, r_1)} L^{q'_1} L^{r'_1}}$$

This implies (7) for  $p^w$ .  $\square$

To obtain such estimates we need some geometric information on the characteristic sets of  $p_{re}$  and  $p_{im}$ . We denote

$$\begin{aligned} \Sigma^{re} &= \text{char } p_{re} \cap B_\lambda, & \Sigma^{im} &= \text{char } p_{im} \cap B_\lambda \\ \Sigma &= \text{char } p \cap B_\lambda. \end{aligned}$$

Our first results use only the geometric information for  $p_{re}$ , which is introduced next.

**(A2)**  $p_{re}$  is of real principal type, i.e.

$$(10) \quad |\nabla_\xi p_{re}(x, \xi)| \gtrsim 1 \quad \text{in } \Sigma^{re}$$

This implies that for each  $x$  the fiber  $\Sigma_x^{re}$  of the characteristic set  $\Sigma^{re}$  is a smooth codimension one hypersurface. Since  $p \in S_\lambda^2$ , the second fundamental form of  $\Sigma^{re}$  has size  $O(\lambda^{-1})$ . Then we impose the following curvature condition on  $\Sigma^{re}$ .

**(A3)** The characteristic set  $\Sigma^{re}$  has  $n-1-k$  nonvanishing curvatures, i.e. for each  $x$  the second fundamental form of  $\Sigma_x^{re}$  has rank at least  $n-1-k$ . More precisely there exists a  $n-1-k$  minor  $M$  of the second fundamental form of  $\Sigma_x^{re}$  with

$$(11) \quad |\det M| \gtrsim \lambda^{k-n+1}.$$



Here  $0 \leq k \leq n - 2$ , but the most interesting cases for applications are  $k = 0, 1$ . The first corresponds to the Schrödinger equation, while the second arises in the study of the wave equation.

Finally, we prove mixed norm estimates, and this requires a choice of coordinates.

**(A4)** The level sets of  $x_1$  are noncharacteristic for  $p_{re}$ , i.e.

$$(12) \quad |\partial_{\xi_1} p_{re}| \gtrsim 1 \quad \text{in } \Sigma^{re}$$

We begin with a result for operators with real symbols:

**Theorem 2.** *a) Let  $p \in \lambda S_\lambda^2$  be a real symbol satisfying (A2-3). Let*

$$(13) \quad q = \frac{2(n+1-k)}{n-1-k}, \quad \rho(q) = \frac{n-1+k}{2(n+1-k)}$$

*Then (4) holds for  $p^w$ , and there is a parametrix  $K$  for  $p^w$  which satisfies (5) and (6).*

*b) If in addition the coordinates are chosen so that (A4) holds and  $(q_1, r_1), (q_2, r_2)$  satisfy*

$$(14) \quad \frac{2}{q} + \frac{n-1-k}{r} = \frac{n-1-k}{2}, \quad 2 \leq q, r \leq \infty, \quad (q, r) \neq (2, \infty)$$

*then (7) holds for  $p^w$ , and there is a parametrix  $K$  for  $p^w$  which satisfies (8) and (9).*

As it turns out, the conclusion of the previous theorem often remains valid if we add a small imaginary component to the above real symbol. We subject such perturbations to a slightly strengthened form of the principal normality condition **(A1)**, namely

**(A1)'** There exist symbols  $r_1 \in \lambda S_\lambda^2$ ,  $r_2, r_3 \in S_\lambda^1$  and  $r_4 \in S_\lambda^0$  such that

$$(15) \quad \{p_{re}, p_{im}\} = r_1 + r_2 p_{re} + r_3 p_{im} + r_4, \quad |r_1| \lesssim |p|$$

**Theorem 3.** *a) Let  $p \in \lambda S_\lambda^2$  be a symbol satisfying (A1)', (A2) and (A3). Let  $q$  be as in (13). Then (4) holds for  $p^w$ , and there is a parametrix  $K$  for  $p^w$  which satisfies (5) and (6).*

*b) Assume in addition that the coordinates are chosen so that (A4) also holds and at least one of the following two conditions applies:*

$$(B1) \quad |d_\xi p_{re} \wedge d_\xi p_{im}| \ll 1 \quad \text{in } \Sigma$$

$$(B2) \quad |\partial_{\xi_1} p_{im}| \ll 1, \quad |\partial_\xi \partial_{\xi_1} p_{im}| \ll \lambda^{-1} \quad \text{in } \Sigma$$

*If  $(q_1, r_1), (q_2, r_2)$  satisfy (14) and  $q_1, q_2 > 2$  then (7) holds for  $p^w$ , and there is a parametrix  $K$  for  $p^w$  which satisfies (8) and (9).*

We had hoped to prove the result with **(A1)'** replaced by **(A1)** but this seems to cause certain difficulties in the proof. It would be interesting to know if this can be done.

We believe that part (b) should hold in general, without any of the assumptions **(B1)** and **(B2)**. Note that **(B1)** is not so harmful, because if it does not hold then **(A2)'** below must hold so we can place ourselves in the setup of the stronger Theorem 4 below. Unfortunately the case (b) of Theorem 4 contains the assumption<sup>1</sup> **(A5)'**, which rules out some of the examples we wish to consider later on. To at least partially compensate for that we have also added the case **(B2)** to the theorem.

Another possible improvement to Theorem 3 would be to also prove it when  $q_1 = q_2 = 2$ . It may be possible to modify our argument to allow one of  $q_1, q_2$  to equal 2, but we do not know how to allow both of them to be 2.

Our next result uses geometric information for both  $p_{re}$  and  $p_{im}$ ; this allows for an improved range of indices in the estimates. Begin with

**(A2)'**  $p_{re}$  and  $p_{im}$  are of real principal type and their characteristic sets are transversal, i.e.

$$(16) \quad |d_\xi p_{re} \wedge d_\xi p_{im}| \gtrsim 1 \quad \text{in } \Sigma$$

This implies that for each  $x \in \mathbb{R}^n$  the  $x$  section  $\Sigma_x$  of the characteristic set of  $p$  is a smooth codimension two submanifold of  $\mathbb{R}^n$ . At each  $\xi \in \Sigma_x$  the two dimensional normal space  $N\Sigma_x$  is generated by  $\partial_\xi p_{re}(x, \xi)$  and  $\partial_\xi p_{im}(x, \xi)$ . Its second fundamental form  $S_{x, \xi}$  maps  $N\Sigma_x \times T\Sigma_x$  into  $T\Sigma_x$ . If we consider it as a quadratic form in  $T\Sigma_x$  depending on the parameter  $\nu \in N\Sigma_x$ , its rank may well depend on  $\nu$ . In particular is not possible to have  $S_{x, \xi}(\nu)$  nondegenerate for all  $\nu \in N\Sigma_x$ . However if it is nondegenerate for some  $\nu$  then it must be nondegenerate for all  $\nu$  except for at most  $n - 2$  values. In this case the directions  $\nu$  for which  $S_{x, \xi}(\nu)$  is degenerate are precisely the directions in which the kernel of the parametrix for  $p^w$  has less decay. Our contention is that for the purpose of proving dispersive estimates we can neglect the bad directions and instead use only an assumption on the generic behavior. Consequently, the curvature condition we impose on  $\Sigma$  is as follows:

**(A3)'** The characteristic set  $\Sigma$  has  $n - 2 - k$  nonvanishing curvatures, i.e. for each  $(x, \xi) \in \Sigma$  there is  $\nu \in N_\xi \Sigma_x$  so that the second fundamental form  $S_{x, \xi}(\nu)$  has rank (at least)  $n - 2 - k$ . More precisely

---

<sup>1</sup>This is likely to be unnecessary, see the discussion after Theorem 4.

we assume that there is a  $n - 2 - k$  minor  $M$  of  $S_{x,\xi}(\nu)$  with

$$(17) \quad |\det M| \gtrsim \lambda^{k-n+2}.$$

In order to obtain mixed norm estimates we need to impose some restriction on how coordinates are chosen. Thus we add the replacement of **(A4)**, namely

**(A4)'** The level sets of  $x_1$  are noncharacteristic for  $p^w$ , i.e.

$$(18) \quad |p_{\xi_1}| \gtrsim 1 \quad \text{in } \Sigma$$

For technical reasons we also invoke a last condition which essentially says that there are no bad directions for the second fundamental form which are tangent to the level sets of  $x_1$ . More precisely, the condition **(A4)'** guarantees that  $N_\xi \Sigma_x$  is transversal to the planes  $x_1 = \text{const}$ . Hence for each  $(x, \xi) \in \Sigma$  there is a unique direction  $\nu_0 \in N_\xi \Sigma_x$  which is tangent to  $x_1 = \text{const}$ .

**(A5)'** There exists a  $n - 2 - k$  minor  $M$  of  $S_{x,\xi}(\nu_0)$  which satisfies (17).

**Theorem 4.** *a) Let  $p \in \lambda S_\lambda^2$  be a symbol satisfying **(A1)**, **(A2)'** and **(A3)'**. Let  $q$  satisfy*

$$(19) \quad q = \frac{2(n+2-k)}{n-k}, \quad \rho(q) = \frac{n-2+k}{2(n+2-k)}$$

*Then (4) holds for  $p^w$ , and there is a parametrix  $K$  for  $p^w$  which satisfies (5) and (6).*

*b) Assume that in addition the coordinates are chosen so that **(A4)'**, **(A5)'** hold and  $(q_1, r_1), (q_2, r_2)$  satisfy*

$$(20) \quad \frac{2}{q} + \frac{n-k}{r} = \frac{n-k}{2} \quad 2 \leq q, r \leq \infty, \quad (q, r) \neq (2, \infty)$$

*with  $q_1$  and  $q_2$  not both equal to 2. Then (7) holds for  $p^w$ , and there is a parametrix  $K$  for  $p^w$  which satisfies (8) and (9).*

We expect Theorem 4 to be valid without the condition **(A5)'**, which insures good kernel decay for the parametrix near  $x_1$  slices. The primary obstruction in this result should come from low kernel decay in the  $x_1$  direction, not along  $x_1$  slices. We were able to prove partial results with logarithmic losses, but not fully remove **(A5)'**.

Finally, we also consider a degenerate case. For  $0 < \delta < 1$  we consider the operator  $p = p_{re} + i\delta p_{im}$  and seek an estimate similar to (4), but with the correct control of the constants as a function of  $\delta$ .

As  $\delta$  approaches 0 one expects the parametrix for  $p^w$  to concentrate closer to the Hamilton flow of  $p_{re}$ . Then it is natural to assume that the direction of the Hamilton flow of  $p_{re}$  is not one of the bad directions:

(A6)' At each point in  $\Sigma$  the second fundamental form of  $\Sigma^{re}$  restricted to  $T\Sigma$  has rank at least  $n - 2 - k$ , i.e. it has a  $n - 2 - k$  minor  $M$  which satisfies (17).

**Theorem 5.** a) Let  $p_{re}^w, p_{im}^w \in S_\lambda^2$  be real symbols satisfying (A1), (A2)', (A3) with  $k$  replaced by  $k + 1$ , (A3)' and (A6)'. Let  $q$  satisfy (19). Then (4) holds for  $p^w$  with  $\lambda^{\rho(q)}$  replaced by  $\delta^{-\frac{1}{n+2-k}} \lambda^{\rho(q)}$ ; also there is a parametrix  $K$  for  $p^w$  which satisfies (5) and (6) with a similar substitution.

b) If in addition the coordinates are chosen so that (A4), (A5)' are satisfied and  $(q_1, r_1), (q_2, r_2)$  are as in Theorem 4 then (7) holds for  $p^w$  with  $\lambda^{\rho(q,s)}$  replaced by  $\delta^{-\frac{2}{q(n-k)}} \lambda^{\rho(q,s)}$ ; also there is a parametrix  $K$  for  $p^w$  which satisfies (8) and (9) with a similar substitution.

### 3. CANONICAL MODELS

Our aim is to reduce the operator  $p^w$  to a canonical form. We decompose the coordinates in the form  $x = (x_1, x')$  and  $\xi = (\xi_1, \xi')$  with  $x_1, \xi_1 \in \mathbb{R}$  and  $x', \xi' \in \mathbb{R}^{n-1}$ .

**Definition 3.1.** We say that the pair of symbols  $p_{re}, p_{im} \in \lambda S_\lambda$  are in canonical form if there are real symbols  $a, b \in \lambda S_\lambda$  so that

$$p_{re}(x, \xi) = \xi_1 + a(x, \xi'), \quad p_{im}(x, \xi) = b(x, \xi')$$

The main result of this section asserts that it suffices to prove the results in the paper when  $p, q$  are in the canonical form. If  $p, q$  are in canonical form then we will prove the estimate for a cutoff symbol  $\chi$  of the form

$$\chi = \chi(x, \xi')$$

supported in

$$B'_\lambda = \{(x, \xi') : |x| \leq 1, |\xi'| \leq \lambda\}$$

This strengthens the estimates.

**Proposition 3.2.** a) Assume that the results in Theorems 2, 4, 5 hold for  $p_{re}, p_{im}$  in canonical form. Then they hold in general.

b) Assume that the result in Theorem 3 holds for  $p_{re}, p_{im}$  in canonical form with the weaker hypothesis that the curvature assumption (A3) applies to  $\xi_1 + a(x, \xi') + \alpha b(x, \xi)$  for some real  $\alpha$  (instead of  $\xi_1 + a(x, \xi')$ ). Then it holds in general.

*Proof.* We prove a series of lemmas, which imply the assertion. First we note that operators localized in  $B_\lambda$  are bounded in all  $L^q L^r$  spaces.

**Lemma 3.3.** Let  $\eta \in S_\lambda^0$  be supported in  $B_\lambda$ . Then  $\eta^w$  is bounded in  $L^q L^r$  for all  $1 \leq q, r \leq \infty$ .

This follows from the translation invariant kernel bound

$$|k(x, y)| \leq c_N \lambda^n (1 + \lambda|x - y|)^{-N}$$

where  $k(x, y)$  is the kernel of  $\eta^w$ . As an immediate application we show that one can localize better the output of the parametrix  $K$ .

**Lemma 3.4.** *Suppose that  $K$  is a parametrix satisfying (8) and (9). Let  $\tilde{\chi}$  be supported in  $B_\lambda$  and identically 1 in the support of  $\chi$ . Then  $\tilde{K} = \tilde{\chi}^w K$  satisfies (8) and (9).*

*Proof.* Lemma 3.3 implies the estimate (8) for  $\tilde{K}$ . For the error estimate we write

$$(p^w \tilde{\chi}^w K - I)\chi^w = \tilde{\chi}^w (p^w K - I)\chi^w + (I - \tilde{\chi}^w)\chi^w + [p^w, \tilde{\chi}^w]K\chi^w$$

In the first term we use the  $L^2$  boundedness of  $\tilde{\chi}^w$ . The second is negligible because  $\chi$  and  $I - \tilde{\chi}$  have disjoint supports. The commutator in the third term is in  $OPS_\lambda^0$  and therefore  $L^2$  bounded.  $\square$

Next we prove that the estimates (8) and (9) do not change if we multiply  $p$  by a zero order elliptic symbol.

**Lemma 3.5.** *Let  $e \in S_\lambda^2$  be an elliptic symbol. Then the conclusion of either of the Theorems 2,3,4,5 holds for  $p$  if and only if it holds for  $\tilde{p} = ep$ .*

*Proof.* Let  $K$  be a parametrix of  $p^w$  which satisfies (8) and (9). By Lemma 3.4 we may replace  $K$  by  $\tilde{\chi}^w K$ . By Lemma 2.3 it suffices to show that  $(e^{-1})^w \tilde{\chi}^w K$  is a parametrix for  $\tilde{p}^w$  which again satisfies (8) and (9). The inequality (8) holds for  $(e^{-1})^w \tilde{\chi}^w K$  because, by Lemma 3.3, the operator  $(e^{-1})^w \tilde{\chi}$  is bounded on  $L^q(L^r)$ . The inequality (9) follows from

$$e^w p^w (e^{-1})^w \tilde{\chi}^w K - I = p^w \tilde{\chi}^w K - I + (e^w p^w (e^{-1})^w - p^w) \tilde{\chi}^w K$$

where  $(e^w p^w (e^{-1})^w - p^w) \in OPS_\lambda^0$  is  $L^2$  bounded.  $\square$

We now consider the issue of localization. The next lemma asserts that the estimates in Theorems 2,3,4,5 are microlocalizable. This allows us to carry out the reduction to the canonical form locally in the phase space.

**Lemma 3.6.** *Let  $0 < \varepsilon < 1$ . Assume that the conclusion of either of Theorems 2,3,4,5 holds for  $\chi$  compactly supported in any ball  $\varepsilon B_\lambda$  contained in  $B_\lambda$ . Then the same holds for  $\chi$  supported in  $B_\lambda$ .*

*Proof.* We consider a finite covering

$$\text{supp } \chi \subset \bigcup_j \varepsilon B_\lambda^j$$

and a corresponding partition of unity

$$1 = \sum \chi_j \text{ in supp } \chi, \quad \text{supp } \chi_j \subset \varepsilon B_\lambda^j.$$

We denote by  $K_j$  the parametrix for  $p^w$  associated to  $\varepsilon B_\lambda^j$ . We define a parametrix  $K$  for  $p^w$  in  $B_\lambda$  by

$$K = \sum K_j \chi_j^w.$$

Then (8) is verified directly, while for (9) we compute

$$(p^w K - id)\chi^w = \sum (p^w K_j - I)\chi_j^w \chi^w + (I - \sum \chi_j^w)\chi^w.$$

The first term is estimated using the hypothesis for  $\varepsilon B_\lambda^j$ , and the second is a smoothing operator since the supports of  $\chi$  and  $(1 - \sum \chi_j)$  are separated.  $\square$

Our next concern is the choice of coordinates. These are uniquely determined in part (b) of the theorems, but we have a choice to make in part (a). The next lemma asserts that we can always choose coordinates so that part (a) is a special case of part (b).

**Lemma 3.7.** *Assume that Theorems 2,3,4,5 are true under the additional hypothesis (A4), and provided that all the assumptions in part (b) also hold for part (a). Then Theorems 2,3,4,5 are true as stated.*

*Proof.* There is nothing to do for part (b) of Theorems 2,4,5. For Theorem 4 (b) we locally get (A4) from (A4)' if we multiply  $p$  by a suitably chosen complex number.

Consider now part (a) of Theorems 2,3,4,5. In Theorem 2(a) we take some  $(x_0, \xi_0) \in \Sigma^{re}$ . By (A1)  $\partial_\xi p_{re}(x_0, \xi_0) \neq 0$ , therefore we can choose coordinates so that  $\partial_{\xi_1} p_{re}(x_0, \xi_0) \neq 0$ . Then the same must be true in an  $\varepsilon B_\lambda$  neighborhood of  $(x_0, \xi_0)$ , i.e. (A4) holds there.

In Theorem 3(a) we take  $(x_0, \xi_0) \in \Sigma^{re}$  and consider three cases.

- If  $|p_{im}(x_0, \xi_0)| \gtrsim \lambda$  then we are in the elliptic region and all our estimates are straightforward.
- If  $|p_{im}(x_0, \xi_0)| \ll \lambda$  and  $|\partial_\xi p_{im}(x_0, \xi_0)| \ll 1$  then (B1) holds in a neighborhood, therefore we can proceed as in the case of Theorem 2.
- If  $|p_{im}(x_0, \xi_0)| \ll 1$  and  $|\partial_\xi p_{im}(x_0, \xi_0)| \gtrsim 1$  then  $\xi_0$  is close to  $\Sigma_{x_0}$ , therefore we can move it slightly and assume that  $(x_0, \xi_0) \in \Sigma$ .
  - If (B1) holds there then we proceed as in the case of Theorem 2.
  - If (B1) does not hold at  $(x_0, \xi_0)$  then (A2)' holds there. If  $\Sigma_{x_0}^{re}$  has  $n - k - 1$  nonvanishing curvatures at  $\xi_0$  then  $\Sigma_{x_0}$  has at least  $n - k - 3$  nonvanishing curvatures at  $\xi_0$ . Hence

we can use Theorem 4(a) to get a result which is at least as good as Theorem 3(a).

In the case of Theorem 4(a) we note that **(A5)'** includes **(A4)'**. Given  $(x_0, \xi_0) \in \Sigma$  we need to choose coordinates near it so that **(A5)'** holds. This only requires that  $dx^1$  is not orthogonal to the set of finitely many directions of bad decay associated to  $(x_0, \xi_0)$ , so it can always be achieved.

For Theorem 5(a) we argue as above, with the only difference that now **(A6)'** requires **(A4)**, namely that  $dx^1$  not be orthogonal to  $\partial_\xi p_{re}(x_0, \xi_0)$ .  $\square$

We now proceed to the main result of the section, which asserts that we can always choose an elliptic symbol  $e$  so that  $ep$  is in the canonical form.

**Lemma 3.8.** (a) Let  $p \in \lambda S_\lambda^2$  be a real symbol which satisfies **(A4)**. Then near each  $(x, \xi) \in \Sigma^{re}$  there is a real elliptic symbol  $e \in S_\lambda^2$  and a real symbol  $a \in \lambda S_\lambda^2$  so that

$$e(x, \xi)p(x, \xi) = \xi_1 + a(x, \xi')$$

(bc) Let  $p = p_{re} + ip_{im} \in \lambda S_\lambda^2$  be a symbol which satisfies **(A4)**. Then near each  $(x, \xi) \in \Sigma^{re}$  there is an elliptic symbol  $e \in S_\lambda^2$  and real symbols  $a, b \in \lambda S_\lambda^2$  so that

$$e(x, \xi)p(x, \xi) = \xi_1 + a(x, \xi') + ib(x, \xi').$$

(d) Let  $p = p_{re} + i\delta p_{im}$ , with  $p_{re}, p_{im} \in \lambda S_\lambda^2$  symbols which satisfy **(A4)** and  $\delta \in [0, 1]$ . Then near each  $(x, \xi) \in \Sigma^{re}$  there is an elliptic symbol  $e \in S_\lambda^2$  with  $\Im e \in \delta S_\lambda^2$  and real symbols  $a, b \in \lambda S_\lambda^2$  so that

$$e(x, \xi)p(x, \xi) = \xi_1 + a(x, \xi') + i\delta b(x, \xi')$$

with the bounds for  $e, a, b$  independent of  $\delta$ .

*Proof.* (a) Since  $p$  is real and  $\partial_{\xi_1} p \approx 1$  it follows that the zero set of  $p$  can be expressed as

$$\{p(x, \xi) = 0\} = \{\xi_1 + a_0(x, \xi') = 0\} \quad a \in S_\lambda^2$$

We make a change of variable  $\xi_1 \rightarrow \xi_1 + a_0(x, \xi')$  which leaves unchanged the classes of symbols we work with. This reduces the problem to the case when

$$\{p(x, \xi) = 0\} = \{\xi_1 = 0\}$$

and we can take

$$e(x, \xi) = \frac{\xi_1}{p(x, \xi)}.$$

It is easy to verify that  $e$  has the desired regularity.

(bc) If  $p_{im} \neq 0$  then we begin by reducing the problem as before to the case when  $p_{re} = \xi_1$ . We want to find an elliptic symbol  $e \in S_\lambda^2$  and real symbols  $a, b \in \lambda S_\lambda^2$  such that

$$e(\xi_1 + ip_{im}) = \xi_1 + a(x, \xi') + ib(x, \xi')$$

First we produce a formal series with this property:

**Lemma 3.9.** *Let*

$$p_{im}(x, \xi_1, \xi') \approx \sum_{k \geq 0} q_k(x, \xi') \xi_1^k, \quad q_k \in \lambda^{1-k} S_\lambda^2$$

*be the formal Taylor series for  $p_{im}$  at  $\xi_1 = 0$ . Then there are formal series*

$$\sum_{k, l \geq 0} e_{k, l}(x, \xi') \xi_1^k q_0^l, \quad \sum_{l \geq 1} a_l(x, \xi') q_0^l, \quad \sum_{l \geq 1} b_l(x, \xi') q_0^l$$

*with coefficients  $e_{k, l} \in \lambda^{-k-l} S_\lambda^2$ ,  $a_l, b_l \in \lambda^{1-l} S_\lambda^2$  whose partial sums*

$$e^N = \sum_{k+l \leq N} e_{k, l}(x, \xi') \xi_1^k q_0^l, \quad a^N = \sum_{l \leq N} a_l(x, \xi') q_0^l, \quad b^N = \sum_{l \leq N} b_l(x, \xi') q_0^l$$

*satisfy*

$$e^{N-1}(\xi_1 + ip_{im}(x, \xi)) = \xi_1 + a^N(x, \xi) + ib^N(x, \xi) + O(\lambda^{-N}(|\xi_1| + |q_0|)^{N+1})$$

*Proof.* The coefficients  $e_{k, l}$ ,  $a_l$ ,  $b_l$  are uniquely determined inductively. We begin with  $N = 1$ , where we must have

$$e_{00}(\xi_1 + iq_0 + iq_1 \xi_1) = \xi_1 + a_1 q_0 + ib_1 q_0.$$

We check this first at points where  $q_0 = 0$  and second where  $\xi_1 = 0$ . This yields

$$e_{00} = (1 + iq_1)^{-1}, \quad a_1 = -\Im(1 + iq_1)^{-1}, \quad b_1 = \Re(1 + iq_1)^{-1}$$

For the induction step we must have

$$\sum_{k+l=N-1} e_{k, l} \xi_1^k q_0^l (\xi_1(1+iq_1) + iq_0) = (a_N + ib_N) q_0^N - \sum_{k+l < N-1} e_{k, l} q_{N-k-l} \xi_1^{N-l} q_0^l$$

Then the coefficients  $e_{k, l}$  are obtained by polynomially dividing the sum on the right by  $\xi_1(1+iq_1) + iq_0$  as polynomials in  $\xi_1$ ; finally, and  $(a_N + ib_N) q_0^N$  is the remainder.  $\square$

The second step is to find smooth functions which match the formal series up to any order. This is a classical argument, see e.g. Hörmander [6], Proposition 18.1.3.



**Lemma 3.10.** *There are symbols  $e \in S_\lambda^2$ ,  $a, b \in \lambda S_\lambda^2$  so that*

$$e - e^N = O(\lambda^{-N-1}(|\xi_1| + |q_0|)^{N+1}), \quad a - a^N, b - b^N = O(\lambda^{-N}|q_0|^{N+1})$$

Now we continue the proof of Lemma 3.8. Combining Lemma 3.9 and Lemma 3.10, we obtain

$$e(\xi_1 + iq(x, \xi)) = \xi_1 + a(x, \xi') + ib(x, \xi') + r(x, \xi', \xi_1, q_0)$$

where the remainder term  $r$  vanishes of infinite order at  $\xi_1 = 0, q_0 = 0$ . Finally we eliminate  $r$  with the substitution

$$e := e + r(x, \xi', \xi_1, q_0)(\xi_1 + iq(x, \xi))^{-1}$$

It is clear that the second right hand side term is a smooth symbol.

(d) This follows simply by replacing  $q_0$  by  $\delta q_0$  in the argument above.  $\square$

To conclude the proof of Proposition 3.2 it remain to study how the hypothesis of our theorems is modified by the multiplication with the elliptic symbol constructed in Lemma 3.8. We begin with

**Lemma 3.11.** *a) The hypotheses of Theorem 2 remain unchanged if we multiply  $p$  by any real elliptic symbol  $e \in S_\lambda^2$ .*

*(b) The hypotheses of Theorem 3 remain unchanged if we multiply  $p$  by any real elliptic symbol  $e \in S_\lambda^2$ .*

*(c) The hypotheses of Theorem 4 remain unchanged if we multiply  $p$  by any elliptic symbol  $e \in S_\lambda^2$ .*

*(d) The hypotheses of Theorem 5 remain unchanged if we multiply  $p$  by any symbol  $e \in S_\lambda^2$  with  $\Re e$  elliptic and  $\Im e \in \delta S_\lambda^2$ .*

The proof of the lemma is straightforward. It completes the proof of Proposition 3.2 (a). However, Proposition 3.2 (b), which refers to Theorem 3, requires a more detailed discussion because the symbol  $e$  does not necessarily satisfy the condition in part (b) of Lemma 3.11.

To understand what happens we consider each of the steps in the proof of Lemma 3.8(bc). By Lemma 3.11(b) the initial multiplication by an elliptic symbol changes nothing.

We fix some  $(x_0, \xi_0)$  on  $\Sigma^{re}$ . The change of variable  $\xi_1 + a_0(x, \xi') \rightarrow \xi_1$  does not affect either of **(B1)** or **(B2)**. If we have  $|q_0(x_0, \xi'_0)| \gtrsim \lambda$  then  $p$  is elliptic at  $x_0, \xi_0$  and all our estimates are straightforward. Hence in what follows we assume that  $|q_0(x_0, \xi'_0)| \ll \lambda$ . We consider two cases:

*Case 1.* Suppose **(B1)** holds near  $(x_0, \xi_0)$ . Then we must have

$$(21) \quad |q_0(x_0, \xi'_0)| \ll \lambda, \quad |\partial_\xi q_0(x_0, \xi'_0)| \ll 1.$$

The symbols  $a(x, \xi')$ ,  $b(x, \xi')$  obtained in Lemma 3.8(bc) can be written in the form

$$(22) \quad a(x, \xi') = a_0(x, \xi') + a_1(x, \xi')q_0(x, \xi') + f(x, \xi')q_0^2(x, \xi'), \quad f \in \lambda^{-1}S_\lambda^2$$

$$(23) \quad b(x, \xi) = b_1(x, \xi')q_0(x, \xi') + g(x, \xi')q_0^2(x, \xi'), \quad g \in \lambda^{-1}S_\lambda^2$$

Here  $b_1 \in S_\lambda^2$  is elliptic. Since the symbol  $a_1(x, \xi') \in S_\lambda^2$  is not necessarily small, near  $(x_0, \xi_0)$  we write

$$a(x, \xi') = a_0(x, \xi') + \frac{a_1(x_0, \xi'_0)}{b_1(x_0, \xi'_0)}b(x, \xi') + \tilde{a}(x, \xi')$$

where, due to (21), the remainder  $\tilde{a}(x, \xi')$  satisfies

$$(24) \quad |\partial_\xi \tilde{a}(x_0, \xi'_0)| \ll 1, \quad |\partial_\xi^2 \tilde{a}(x_0, \xi'_0)| \ll \lambda^{-1}$$

Since  $\xi_1 + a_0(x, \xi')$  satisfies the curvature condition **(A2)** this implies that so must  $\xi_1 + a(x, \xi') - \alpha b(x, \xi')$  near  $(x_0, \xi_0)$ , where  $\alpha = \frac{a_1(x_0, \xi'_0)}{b_1(x_0, \xi'_0)}$ .

*Case 2.* Suppose **(B1)** does not hold near  $(x_0, \xi_0)$ . Then we must have

$$|q_0(x_0, \xi_0)| \ll \lambda \quad |\partial_\xi q_0(x_0, \xi_0)| \gtrsim 1.$$

Hence we can shift  $\xi_0$  slightly to arrive at the case when

$$|q_0(x_0, \xi_0)| = 0 \quad |\partial_\xi q_0(x_0, \xi_0)| \gtrsim 1.$$

Since **(B1)** does not hold, **(B2)** must hold at  $(x_0, \xi_0)$ . This implies that

$$|q_1(x_0, \xi_0)| \ll 1, \quad |\partial_\xi q_1(x_0, \xi_0)| \ll \lambda^{-1}$$

Hence the same must hold for the symbol  $a_1 = \Im(1 + iq_1)^{-1}$ . Going back to (22), we obtain a relation of the form

$$a(x, \xi') = a_0(x, \xi') + \tilde{a}(x, \xi')$$

where  $\tilde{a}$  satisfies (24). Then  $\xi_1 + a(x, \xi')$  satisfies the curvature condition **(A2)** near  $(x_0, \xi_0)$ . This concludes the proof of Proposition 3.2.  $\square$

#### 4. THE REAL CASE

**4.1. The parametrix construction.** In this section we denote  $x_1$  by  $t$ , set  $d = n - 1$  and redenote  $x'$  by  $x \in \mathbb{R}^d$ . We shall construct a phase space representation of the fundamental solution for the initial value problem

$$(25) \quad (D_t + a^w(t, x, D))u = f \quad u(0) = u_0$$

where  $t \in [0, 1]$  and  $x \in \mathbb{R}^d$ . If the symbol  $a$  is real then  $a^w$  is selfadjoint, therefore it generates an isometric evolution operator  $S(t, s)_{t, s \in [0, 1]}$  in  $L^2(\mathbb{R}^d)$ .

To keep the argument simple we work in a normalized setup where all scales are of order 1. Thus we assume that the symbol  $a$  is measurable in  $t$  and that it satisfies the bounds

$$(26) \quad |\partial_x^\alpha \partial_\xi^\beta a(t, x, \xi)| \leq c_{\alpha, \beta} \quad |\alpha| + |\beta| \geq 2$$

We first outline a simple parametrix construction based on the FBI transform, following ideas in [21]. There the equation is conjugated with respect to the FBI transform and replaced by a simpler transport equation in the phase space.

The FBI transform<sup>2</sup> is an isometry from  $L^2(\mathbb{R}^d)$  to  $L^2(\mathbb{R}^{2d})$  which is defined by

$$Tu(x, \xi) = 2^{-\frac{d}{2}} \pi^{-\frac{3d}{4}} \int e^{-i\xi(x-y) - \frac{1}{2}(x-y)^2} f(y) dy$$

We approximate the conjugated operator by

$$\tilde{A} = a(x, \xi) + a_\xi \left( \frac{1}{i} \partial_x - \xi \right) - \frac{1}{i} a_x \partial_\xi$$

and we use the error estimate in [22], Theorem 6:

**Lemma 4.1.** *Assume that the symbol  $a$  satisfies (26). Then we have*

$$\|Ta^w - \tilde{A}T\|_{L^2 \rightarrow L^2} \lesssim 1$$

and the dual estimate

$$\|T^* \tilde{A} - a^w T^*\|_{L^2 \rightarrow L^2} \lesssim 1$$

The operator  $\tilde{A}$  is selfadjoint, therefore it generates an isometric evolution operator  $\tilde{S}(t, s)$  in  $L^2(\mathbb{R}^{2n})$ . Then a natural choice for a forward parametrix is the operator

$$K(t, s) = 1_{t \geq s} T^* \tilde{S}(t, s) T$$

Given the above error estimates, it is straightforward to prove that this provides a good approximate solution in the  $L^2$  sense:

**Proposition 4.2.** *Assume that the symbol  $a$  satisfies (26). Then the operator  $K(t, s)$  satisfies*

$$\begin{aligned} \|K(t, s)\|_{L^2 \rightarrow L^2} &\leq 1 \\ \|(D_t + a^w)K(t, s)\|_{L^2 \rightarrow L^2} &\lesssim 1 \\ \lim_{t \rightarrow s+} K(t, s) &= 1_{L^2} \end{aligned}$$

---

<sup>2</sup>Often one adds the factor  $e^{\frac{1}{2}\xi^2}$  in the formula; this would generate some obvious changes in what follows.

This result is strong enough in order to solve the original equation (25) iteratively for  $f \in L^1 L^2$  and  $u_0 \in L^2$ . The kernel of the parametrix  $K$  is easy to describe explicitly. To do this we begin with the evolution operator  $\tilde{S}(t, s)$  in the phase space. It corresponds to the transport type operator

$$D_t + \tilde{A} = -i(\partial_t + a_\xi \partial_x - a_x \partial_\xi) + a(x, \xi) - \xi a_\xi$$

Solutions are transported along the Hamilton flow for  $D_t + a^w$ , namely

$$\dot{x} = a_\xi \quad \dot{\xi} = -a_x$$

We denote its solution by  $x^t(x, \xi)$  and  $\xi^t(x, \xi)$  where  $x$  and  $\xi$  are the initial data at time  $t = 0$ . There is also a phase shift. We define the real phase function  $\psi$  by

$$(27) \quad \dot{\psi} = -a + \xi a_\xi, \quad \psi(0, \bar{x}, \bar{\xi}) = 0$$

where  $\dot{\psi}$  denotes the differentiation along the flow. Then  $\tilde{S}(t, s)$  is given by

$$(\tilde{S}(t, s)u)(x^t, \xi^t) = u(x^s, \xi^s) e^{i(\psi(t, x, \xi) - \psi(s, x, \xi))}$$

and the parametrix  $K$  has the kernel

$$K(t, y, s, \tilde{y}) = 2^{-d} \pi^{-\frac{3d}{2}} \int_{\mathbb{R}^{2n}} e^{-\frac{1}{2}(y-x^t)^2 - \frac{1}{2}(\tilde{y}-x^s)^2 + i\xi^t(y-x^t) - i\xi^s(\tilde{y}-x^s)} e^{i(\psi(t, x, \xi) - \psi(s, x, \xi))} dx d\xi.$$

One can use this directly to prove the dispersive estimates, as in [21].

Here we prefer to use a different approach and work with exact instead of approximate solutions. Inspired by the above parametrix, we seek to obtain a similar representation for the solution.

**Proposition 4.3.** *The kernel  $K$  of the fundamental solution operator  $D_t + a^w$  can be represented in the form*

$$(28) \quad K(t, y, s, \tilde{y}) = \int_{\mathbb{R}^{2n}} e^{-\frac{1}{2}(\tilde{y}-x^s)^2} e^{-i\xi^s(\tilde{y}-x^s)} e^{i(\psi(t, x, \xi) - \psi(s, x, \xi))} e^{i\xi^t(y-x^t)} G(t, s, x, \xi, y) dx d\xi$$

where the function  $G$  satisfies

$$(29) \quad |(x^t - y)^\gamma \partial_x^\alpha \partial_\xi^\beta \partial_y^\nu G(t, s, x, \xi, y)| \lesssim c_{\gamma, \alpha, \beta, \nu}$$

*Proof.* Without any restriction in generality take  $s = 0$  and drop the argument  $s$  in the notations. Given  $u_0 \in L^2$  we need to find the solution  $u = S(t, 0)u_0$  for the equation

$$(D_t + a^w)u = 0, \quad u(0) = u_0$$

We use the FBI transform to decompose  $u_0$  into coherent states, and write

$$u = S(t, 0)T^*Tu_0 = \int S(t, s)\phi_{x,\xi}Tu(x, \xi)dx d\xi$$

where the coherent states  $\phi_{x,\xi}$  are given by

$$\phi_{x,\xi}(y) = 2^{-\frac{d}{2}}\pi^{-\frac{3d}{4}}e^{i\xi(x-y)-\frac{1}{2}(x-y)^2}.$$

Then we can define the function  $G$  by

$$G(t, x, \xi, y) = 2^{-\frac{d}{2}}\pi^{-\frac{3d}{4}}e^{-i\xi^t(y-x^t)}e^{-i\psi(t,x,\xi)}(S(t, 0)\phi_{x,\xi})(y)$$

so that (28) holds. It remains to prove that  $G$  satisfies the bounds (29). For this we need to study the regularity of the Hamilton flow and of the phase function.

**Lemma 4.4.** *Assume that the symbol  $A$  satisfies (26). Then*

$$(30) \quad |\partial_x^\alpha \partial_\xi^\beta x^t| + |\partial_x^\alpha \partial_\xi^\beta \xi^t| \leq c_{\alpha,\beta} \quad |\alpha| + |\beta| \geq 1$$

and, for  $|\alpha| + |\beta| \geq 1$  and  $x_0, \xi_0 \in \mathbb{R}^d$ ,

$$(31) \quad \left| \partial_x^\alpha \partial_\xi^\beta [\xi^t(y - x^t) + \psi(t, x, \xi) + \xi_0(x - x_0)] \right|_{x=x_0, \xi=\xi_0} \leq c_{\alpha,\beta}(1 + |y - x_0^t|)$$

*Proof.* The bound (30) is a consequence of the structure of the Hamiltonian equations.

For (31) we note that if any derivatives fall on  $\xi^t$  then the corresponding term can be easily estimated using (30). It remains to consider the quantity

$$e = -\xi^t \partial_x^\alpha \partial_\xi^\beta x^t + \partial_x^\alpha \partial_\xi^\beta \psi(t, x, \xi) + \xi_0 \partial_x^\alpha \partial_\xi^\beta (x - x_0)$$

At  $t = 0$  and  $(x, \xi) = (x_0, \xi_0)$  we trivially have  $e = 0$  therefore it suffices to bound its derivative along the flow. But

$$\dot{e} = a_x(x_t, \xi_t) \partial_x^\alpha \partial_\xi^\beta x^t - \xi^t \partial_x^\alpha \partial_\xi^\beta a_\xi(x_t, \xi_t) - \partial_x^\alpha \partial_\xi^\beta a(x_t, \xi_t) + \partial_x^\alpha \partial_\xi^\beta (\xi^t a_\xi(x_t, \xi_t))$$

We use Leibniz's rule for the last expression. All terms are bounded except the ones where all derivatives fall on  $\xi^t$ , respectively  $a_\xi(x_t, \xi_t)$ . The latter cancels the second expression in the above formula, and we are left with

$$\dot{e} = a_x(x_t, \xi_t) \partial_x^\alpha \partial_\xi^\beta x^t - \partial_x^\alpha \partial_\xi^\beta a(x_t, \xi_t) + a_\xi(x_t, \xi_t) \partial_x^\alpha \partial_\xi^\beta \xi^t + O(1)$$

For the middle expression we use the chain rule. The terms where at least two derivatives fall on the symbol  $a$  are bounded. But this leaves us with only two terms which cancel the first and the last in the above formula.  $\square$

We now continue the proof of Proposition 4.3. Using (31) one easily obtains at  $x = x_0, \xi = \xi_0$

$$(32) \quad \left| \partial_y^\nu \partial_x^\alpha \partial_\xi^\beta [\xi^t(y - x^t) + \psi(t, x, \xi) + \xi_0(x - x_0) - \xi_0^t(y - x_0^t) - \psi(t, x_0, \xi_0)] \right| \leq c_{\alpha, \beta, \nu}(1 + |y - x_0^t|)$$

Then it suffices to prove the estimates (29) at  $(x_0, \xi_0)$  for the modified function

$$G_1(t, x, \xi, y) = e^{-i\xi_0^t(y - x_0^t)} e^{-i\psi(t, x_0, \xi_0)} (S(t, 0)(e^{i\xi_0(x - x_0)} \phi_{x, \xi})) (y)$$

We translate  $G_1$  to the origin by setting

$$G_2(t, x, \xi, y) = G_1(t, x_0 + x, \xi_0 + \xi, x_0^t + y)$$

The  $x$  and  $\xi$  variables are translated so that they are now centered at the origin. We do not change the notations since at this point  $x$  and  $\xi$  appear only in the initial data for  $G_1$  and  $G_2$ .

The relation (29) at  $(x_0, \xi_0)$  is replaced by a similar relation at  $(0, 0)$ ,

$$(33) \quad \left| y^\gamma \partial_x^\alpha \partial_\xi^\beta \partial_y^\nu G_2(t, x, \xi, y) \right|_{x=0, \xi=0} \lesssim c_{\gamma, \alpha, \beta, \nu}$$

A routine computation shows that the function  $G_2$  solves the modified equation

$$(D_t + a_2^w(t, y, D_y))G_2 = 0, \quad G_2(0) = \phi_{x, \xi}$$

where

$$a_2(t, y, \eta) = a(t, x_0^t + y, \xi_0^t + \eta) - a(t, x_0^t, \xi_0^t) - y a_x(t, x_0^t, \xi_0^t) - \eta a_\xi(t, x_0^t, \xi_0^t)$$

still satisfies (26) but in addition vanishes of second order at  $0 \in \mathbb{R}^{2d}$ . To differentiate it with respect to  $x, \xi$  it suffices to differentiate the initial data. But the functions

$$\partial_x^\alpha \partial_\xi^\beta \phi_{x, \xi}(y)|_{x=0, \xi=0}$$

are Schwartz functions in  $y$ . Hence it suffices to consider the problem

$$(D_t + a_2^w(t, y, D))v = 0, \quad v(0) = v_0$$

where the initial data  $v_0$  is a Schwartz function, and prove that the solution  $v(t)$  is also a Schwartz function. This follows if we can prove energy estimates for the functions  $y^\alpha \partial^\beta v$ , which we do by induction over  $k = |\alpha| + |\beta|$ . If  $k = 0$  then we trivially have

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

For  $k = 1$  we compute the equations for  $yv$  and  $\partial_y v$ :

$$(D_t + a_2^w(t, y, D))(yv) = -i(\partial_\eta a)^w(t, y, D)v$$

$$(D_t + a_2^w(t, y, D))(\partial_y v) = i(\partial_y a)^w(t, y, D)v$$

To bound the right hand side we need the next lemma for the symbol  $b = \partial_y a_2$  and  $b = \partial_\xi a_2$ . This is a special case of Theorem 3 in [22].

**Lemma 4.5.** *Suppose that the symbol  $b(x, \xi)$  satisfies*

$$\|\partial_y^\alpha \partial_\eta^\beta b(y, \eta)\| \leq c_{\alpha, \beta} \quad |\alpha| + |\beta| \geq 1$$

*and also  $b(0, 0) = 0$ . Then*

$$\|b^w(y, D)u\|_{L^2} \lesssim \|yu\|_{L^2} + \|\partial u\|_{L^2} + \|u\|_{L^2}$$

Using Lemma 4.5 and the Gronwall inequality we conclude that

$$\|yv(t)\|_{L^2} + \|\partial v(t)\|_{L^2} \lesssim \|yv_0\|_{L^2} + \|\partial v_0\|_{L^2} + \|v_0\|_{L^2}$$

It remains to do the induction step. We denote by  $L_k$  all operators of the form  $x^\alpha \partial^\beta$  with  $|\alpha| + |\beta| = k$ . Suppose that

$$\sum_{j \leq k} \|L_j v(t)\|_{L^2} \leq c_k \sum_{j \leq k} \|L_j v_0\|_{L^2}$$

The functions  $L_{k+1}v$  solve a weakly coupled system of the form

$$(D_t - a_2^w)L_{k+1}v = (\partial_{y, \eta} a_2)^w L_k v + \sum_{i \geq 2}^{i+j \leq k+1} (\partial_{y, \eta}^i a_2)^w L_j v$$

For this we use energy estimates and Gronwall's inequality. The first right hand side term is estimated using Lemma 4.5 and the second using the induction hypothesis.  $\square$

We conclude the section with a supplementary result where we establish some time regularity of the function  $G$  in Proposition 4.3. This will not be needed until Section 6.4.

**Proposition 4.6.** *Let  $a$  be a symbol satisfying (26), and  $G$  be as in Proposition 4.3. Suppose that  $a$  is smooth in  $t$  near some  $t_0 \in [0, 1]$  and satisfies the additional relations*

$$(34) \quad |a_x(t_0, x, \xi)| + |a_\xi(t_0, x, \xi)| \leq \mu$$

$$(35) \quad |\partial_t^\sigma \partial_x^\alpha \partial_\xi^\beta a(t_0, x, \xi)| \leq c_{\alpha, \beta, \sigma} \mu^{\sigma+1}, \quad \sigma \geq 1, \quad |\alpha| + |\beta| \geq 1$$

*where  $\mu > 1$  is a large parameter. Then  $G$  satisfies the additional bound*

$$(36) \quad |(x^t - y)^\gamma \partial_t^\sigma \partial_x^\alpha \partial_\xi^\beta \partial_y^\nu G(t_0, s, x, \xi, y)| \lesssim c_{\gamma, \alpha, \beta, \nu, \sigma} \mu^\sigma$$

*Proof.* We begin with an analysis of the time derivatives of the Hamilton flow. The counterpart of (30) is

$$(37) \quad |\partial_t^\sigma \partial_x^\alpha \partial_\xi^\beta x^t| + |\partial_t^\sigma \partial_x^\alpha \partial_\xi^\beta \xi^t| \leq c_{\alpha,\beta,\sigma} \mu^\sigma, \quad \sigma + |\alpha| + |\beta| \geq 1, \quad t = t_0$$

This can be proved by induction with respect to  $\sigma$ . The details are left for the reader. An immediate consequence of it is

$$(38) \quad |\partial_t^\sigma \partial_x^\alpha \partial_\xi^\beta a_x(t, x^t, \xi^t)| + |\partial_t^\sigma \partial_x^\alpha \partial_\xi^\beta a_\xi(t, x^t, \xi^t)| \leq c_{\alpha,\beta,\sigma} \mu^{\sigma+1}, \quad t = t_0$$

Next we examine the steps in the proof of Proposition 4.3 and track the time derivatives of  $G$ .

(i) *The phase correction.* Here we need to strengthen (32) to a form which also includes time derivatives,

$$(39) \quad \left| \partial_t^\sigma \partial_y^\nu \partial_x^\alpha \partial_\xi^\beta [\xi^t(y - x^t) + \psi(t, x, \xi) + \xi_0(x - x_0) - \xi_0^t(y - x_0^t) - \psi(t, x_0, \xi_0)] \right| \leq c_{\alpha,\beta,\nu,\sigma} (1 + |y - x_0^t|) \mu^\sigma, \quad t = t_0.$$

We know this for  $\sigma = 0$ , and it is trivial if  $|\alpha| + |\beta| = 0$ . Then we denote by  $e$  the expression which is differentiated in the above formula, and we first compute

$$\begin{aligned} \partial_{x,\xi} \partial_t e &= \partial_{x,\xi} \partial_t (\xi^t(y - x^t) + \psi(t, x, \xi)) \\ &= \partial_{x,\xi} (-a_x(t, x^t, \xi^t)(y - x^t) - a(t, x^t, \xi^t)) \\ &= -(y - x^t) \partial_{x,\xi} a_x(t, x^t, \xi^t) - a_\xi(t, x^t, \xi^t) \partial_{x,\xi} \xi^t \end{aligned}$$

The proof of (39) is completed using Leibniz's rule and (37), (38).

(ii) *The coordinate change.* The function  $G_2$  is obtained from  $G_1$  after a time dependent translation of  $x_0^t$ . This translation preserves the bounds (36) on  $G$  since by (37) we know that

$$|\partial_t^\sigma x_0^t| \leq c_\sigma \mu^\sigma \quad t = t_0$$

(iii) *The localized evolution.* The bounds for  $\partial_t v$  are obtained directly from the equation. To obtain bounds for higher order time derivatives of  $v$  we repeatedly differentiate the equation. For this we need to control the derivatives of the symbol  $a_2$ . We know that

$$a_2(t, 0, 0) = 0, \quad \partial_{x,\xi} a_2(t, 0, 0) = 0$$

Then the same will apply to the time derivatives of  $a_2$ . It remains to show that

$$|\partial_t^\sigma \partial_x^\alpha \partial_\xi^\beta a(t, x + x_0^t, \xi + \xi_0^t)| \leq c_{\alpha,\beta,\sigma} \mu^{1+\sigma}, \quad |\alpha| + |\beta| \geq 2, \quad \sigma \geq 1, \quad t = t_0$$

which is easily done using the chain rule and (37), (38).  $\square$



**4.2. Fixed time estimates.** Here we combine the above representation of the fundamental solution for  $D_t + a^w$  with the curvature condition in order to obtain pointwise bounds for its kernel.

**Proposition 4.7.** *Let  $a \in S_\lambda^2$  and  $0 \leq k \leq d$  so that for each  $(t, x, \xi) \in B_\lambda$  there exists an  $d-k$  nondegenerate minor  $M$  of  $\partial_\xi^2 a(t, x, \xi)$  satisfying*

$$|\det M| \gtrsim \lambda^{-(d-k)}.$$

*Then there exists  $T > 0$  so that for all  $|t - s| < T$  we have*

$$(40) \quad \|S(t, s)\chi^w u_0\|_{L^\infty} \lesssim \lambda^{\frac{d+k}{2}} |t - s|^{-\frac{-d+k}{2}} \lesssim \|u_0\|_{L^1}$$

*Proof.* Without any restriction in generality we take  $s = 0$ . We fix  $t_0$  in  $[0, 1]$  and seek to prove the above estimate when  $t = t_0$ . The result is trivial if  $t_0 < \lambda^{-1}$ . Hence in the sequel we assume that  $t_0 \geq \lambda^{-1}$ .

We rescale the problem to reduce it to an estimate for  $t = 1$ . If  $u = S(t, 0)\chi^w u_0$  then we set

$$v(t, x) = u\left(\frac{t}{t_0}, \frac{x\sqrt{t_0}}{\sqrt{\lambda}}\right)$$

The function  $v$  solves the equation

$$(D_t + \tilde{a}^w(t, x, D))v = 0, \quad v(0) = \tilde{\chi}(t, x, D)v_0$$

where

$$\tilde{a}(t, x, \xi) = t_0 a\left(\frac{t}{t_0}, \frac{x\sqrt{t_0}}{\sqrt{\lambda}}, \frac{\xi\sqrt{\lambda}}{\sqrt{t_0}}\right), \quad \tilde{\chi}(t, x, \xi) = \chi^w\left(\frac{x\sqrt{t_0}}{\sqrt{\lambda}}, \frac{\xi\sqrt{\lambda}}{\sqrt{t_0}}\right)$$

The new frequency scale is  $\mu = \sqrt{t_0\lambda}$ . The rescaled version of (40) has the form

$$(41) \quad \|v(1)\|_{L^\infty} \lesssim \mu^k \|v_0\|_{L^1}$$

It is easy to verify that  $\tilde{a}$  satisfies (26), therefore we can use the parametrix in Proposition 4.3,

$$v(t, y) = \int_{\mathbb{R}^{3d}} G(t, x, \xi, y) e^{-\frac{1}{2}(\tilde{y}-x)^2 + i\xi^t(y-x^t) - i\xi(y-x)} e^{i\psi(t, x, \xi)} (\tilde{\chi}v_0)(\tilde{y}) dx d\xi d\tilde{y}$$

where the symbol  $\tilde{\chi}$  is compactly supported in

$$\tilde{B} = \{|x| \leq \mu t_0^{-1}, |\xi| \leq \mu\}$$

and it is smooth on the scale of  $\tilde{B}$ . The contribution of the complement of  $\tilde{B}$  to the above integral is negligible. More precisely we can write

$$v(t, y) = \int_{\tilde{B}} G(t, x, \xi, y) e^{-\frac{1}{2}(\tilde{y}-x)^2 + i\xi^t(y-x^t) - i\xi(y-x)} e^{i\psi(t, x, \xi)} (\tilde{\chi}^w v_0)(\tilde{y}) dx d\xi d\tilde{y} \\ + O(\mu^{-\infty})$$

Using an  $L^1$  bound for  $\tilde{\chi} v_0$  and a trivial estimate for the kernel, the inequality (41) would follow from

$$\int_{\tilde{B}} |G(1, x, \xi, y)| d\xi \lesssim \mu^k$$

Given the bounds (29) for  $G$ , this reduces to

$$(42) \quad \int_{\tilde{B}} (1 + |x^1 - y|)^{-N} d\xi \lesssim \mu^k, \quad N \text{ large}$$

The key factor here is the dependence of  $x^1$  on  $\xi$ . We study this using the linearization of the Hamilton flow. The functions

$$X = \frac{\partial x^t}{\partial \xi}, \quad \Xi = \frac{\partial \xi^t}{\partial \xi}$$

solve the ordinary differential equation along the Hamilton flow

$$\begin{cases} \dot{X} = \tilde{a}_{\xi x} X + \tilde{a}_{\xi \xi} \Xi \\ \dot{\Xi} = -\tilde{a}_{xx} X - \tilde{a}_{x\xi} \Xi \end{cases} \quad \begin{cases} X(0) = 0 \\ \Xi(0) = I \end{cases}$$

Since  $\tilde{a}_{\xi x}, \tilde{a}_{xx}, \tilde{a}_{x\xi} = O(\sqrt{t_0})$  we obtain

$$\dot{X} = \tilde{a}_{\xi \xi} + O(\sqrt{t_0})$$

We can also compute

$$\dot{\tilde{a}}_{\xi \xi} = \tilde{a}_{t\xi\xi} + \tilde{a}_{\xi\xi x} \tilde{a}_\xi - \tilde{a}_{\xi\xi\xi} \tilde{a}_x = O(\sqrt{t_0})$$

Hence we obtain

$$X(t) = t(\tilde{a}_{\xi \xi}(0, x, \xi) + O(\sqrt{t_0}))$$

which at time 1 gives

$$(43) \quad \frac{\partial x^1}{\partial \xi} = \tilde{a}_{\xi \xi}(0, x, \xi) + O(\sqrt{t_0})$$

Given  $\xi_0 \in \tilde{B}$  we choose coordinates

$$\xi = (\xi', \xi''), \quad \xi' = (\xi_1, \dots, \xi_{d-k})$$

so that the matrix  $(\partial_{\xi}^2 \tilde{a}(0, x, \xi_0))$  is nondegenerate. Since  $|\partial_{\xi}^3 \tilde{a}(x, \xi)| \lesssim \mu^{-1}$ , it follows that the same must hold for  $\xi \in B(\xi_0, \delta\mu)$  for small fixed  $\delta$ . To prove (42) we split  $\tilde{B}$  into balls of radius  $\delta\mu$ . Since

$$\int_{B(\xi_0, \delta\mu)} (1 + |x^1 - y|)^{-N} d\xi \lesssim \mu^k \sup_{\xi''} \int_{B'(x_0, \delta\mu)} (1 + |x^1 - y|)^{-N} d\xi'$$

it suffices to show that

$$\int_{B(\xi_0, \delta\mu)} (1 + |x^1 - y|)^{-N} d\xi' \lesssim 1$$

But in this region (43) shows that  $\partial_{\xi'} x^1$  is a small perturbation of the nondegenerate matrix  $\partial_{\xi}^2 \tilde{a}(x, \xi_0)$ . Hence the above estimate follows.  $\square$

**4.3. Mixed norm estimates.** Here we use the fixed time bounds obtained before in order to derive space-time estimates in mixed norm spaces.

**Proposition 4.8.** *Assume that  $D_t + a^w$  satisfies **(A2)**, **(A3)** in  $B_\lambda$ . Let  $\chi, \tilde{\chi} \in S_\lambda^2$  be symbols which are supported in  $B_\lambda$ . Let  $u$  solve*

$$(D_t + a^w)u = \tilde{\chi}^w(x, D)f_1 + f_2 \quad u(0) = u_0$$

in  $[0, 1]$ . Then for  $(r, s)$  as in (14) we have

$$(44) \quad \|u\|_{L^\infty L^2} + \|\chi^w(x, D)u\|_{\lambda^{\rho(r,s)} L^r L^s} \lesssim \|f_1\|_{\lambda^{-\rho(r,s)} L^{r'} L^{s'}} + \|f_2\|_{L^1 L^2} + \|u_0\|_{L^2}$$

*Proof.* It suffices to prove this in a sufficiently small time interval, as we can iterate it and obtain it in the full interval  $[0, 1]$ .

Besides the trivial energy estimates need to show that

$$(45) \quad \begin{aligned} \chi^w S(t, s) &: L^2 \rightarrow \lambda^{\rho(r,s)} L^r L^s \\ S(t, s) \chi^w &: \lambda^{-\rho(r,s)} L^{r'} L^{s'} \rightarrow L^2 \\ 1_{t>s} \chi^w S(t, s) \tilde{\chi}^w &: \lambda^{-\rho(r,s)} L^{r'} L^{s'} \rightarrow \lambda^{\rho(r,s)} L^r L^s \end{aligned}$$

The first two statements are dual. Using a  $TT^*$  argument they reduce to the third without  $1_{t>s}$ . But the third with  $1_{t>s}$  is dual to the third with  $1_{t<s}$  instead, therefore it implies the first two. It remains to prove the third.

The energy estimates yield the trivial bound

$$\|S(t, s)\|_{L^2 \rightarrow L^2} \leq 1$$

On the other hand the decay estimates in Proposition 4.7 show that

$$\|\chi^w S(t, s) \tilde{\chi}^w\|_{L^1 \rightarrow L^\infty} \lesssim |t - s|^{-\frac{d-k}{2}} \lambda^{\frac{d+k}{2}}$$

If  $r > 2$  then (45) follows from the two estimates above by interpolation and the Hardy-Littlewood-Sobolev inequality. The case  $r = 2$  needs some extra work, and it can be obtained as in Keel-Tao [11].  $\square$

**Corollary 4.9.** *Assume that  $D_t + \tilde{a}$  satisfies (A2), (A3) in  $B_\lambda$ . Let  $\chi \in S_\lambda^2$  be a symbol which is supported in  $B_\lambda$ . Then<sup>3</sup>*

$$(46) \quad \|\chi^w(x, D)u\|_{L^\infty(L^2) \cap \lambda^{\frac{n+k-2}{2(n-k)}} L^2(L^{\frac{2(n-k)}{n-k-2}})} \lesssim \|u\|_{L^2} + \|(D_t + A)u\|_{L^2}.$$

This follows easily from the previous proposition applied to  $\tilde{\chi}^w u$ , where the symbol  $\tilde{\chi}$  is chosen to equal 1 in a neighborhood of the support of  $\chi$ .

## 5. THE PARAMETRIX IN THE GENERAL CASE

In this section we construct two parametrices for operators in canonical form. These constructions do not use much of the structure of our operators, so we prefer to write it in a more abstract setup.

Thus, given selfadjoint operators  $A(t)$ ,  $B(t)$  in a Hilbert space  $X$  for  $t \in [0, 1]$ , we seek to construct a parametrix for the operator

$$D_t + A + iB$$

For simplicity we assume that both  $A(t)$  and  $B(t)$  are bounded and smooth as functions of  $t$ , but the constants in our estimates are independent of any such bounds.

Making a slight abuse of notation we use simply  $\|\cdot\|$  both for the norm of  $X$  and for the operator norm in  $L(X)$  through this section. All other norms will be indicated with a subscript. We also abbreviate the notation  $\|\cdot\|_{L^q} := \|\cdot\|_{L^q(0,1;X)}$ .

The main relation connecting  $A$  and  $B$  is the fixed time commutator estimate

$$(47) \quad \|[D_t + A, B]u\| \lesssim \|Bu\| + \|u\|$$

We do not use this directly, instead we first obtain a simple consequence of it. Denote by  $S(t, s)$  the unitary evolution generated by  $D_t + A$  in  $X$ .

**Lemma 5.1.** *Assume that (47) holds in  $[0, 1]$ . Then*

$$(48) \quad \|(B(t)S(t, s) - S(t, s)B(s))u\| \lesssim |t - s|(\|B(s)u\| + \|u\|)$$

---

<sup>3</sup>If  $k = n - 2$  then the  $L^2 L^\infty$  bound is not quite true, but all the intermediate bounds are still valid.

*Proof.* First we compute the equation for  $B(t)S(t, s)u$ ,

$$(D_t + A)B(t)S(t, s)u = [D_t + A, B]S(t, s)u$$

Using energy estimates and (47) we obtain

$$\|B(t)S(t, s)u\| \lesssim \|B(s)u\| + \int_s^t \|B(r)S(r, s)u\| + \|S(r, s)u\| dr$$

Applying Gronwall's lemma this yields

$$(49) \quad \|B(t)S(t, s)u\| \lesssim \|B(s)u\| + \|u\|$$

Then we write

$$(D_t + A)(B(t)S(t, s) - S(t, s)B(s)) = [D_t + A, B]S(t, s)$$

which implies that

$$(50) \quad \|(B(t)S(t, s) - S(t, s)B(s))u\| \lesssim \int_s^t \|B(r)S(r, s)u\| + \|S(r, s)u\| dr$$

To obtain (48) it suffices to estimate the right hand side using (49).  $\square$

**5.1. A simple parametrix.** Assume first that  $D_t + A$  commutes with  $B$ . Then we can produce an exact parametrix for  $D_t + A + iB$ , namely

$$H(t, s) = 1_{(t-s)B(t) < 0} e^{(t-s)B(t)} S(t, s) = S(t, s) 1_{(t-s)B(s) < 0} e^{(t-s)B(s)}$$

where the  $B$  dependent part is interpreted in the sense of operator calculus for selfadjoint operators.

In the noncommuting case we can use either of these two expressions as a parametrix, but they are no longer equal. Set

$$(51) \quad H(t, s) = S(t, s) 1_{(t-s)B(s) < 0} e^{(t-s)B(s)}$$

**Proposition 5.2.** *Suppose that (48) holds for  $t, s \in [0, 1]$ . Then for  $t, s$  in a bounded interval we have the fixed time estimates*

$$(52) \quad \|H(t, s)\|_{X \rightarrow X} \leq 1, \quad \|H(t, s)B(s)\|_{X \rightarrow X} \leq |t - s|^{-1},$$

$$(53) \quad \|B(t)H(t, s)\|_{X \rightarrow X} \lesssim |t - s|^{-1}, \quad \|B(t)H(t, s)B(s)\|_{X \rightarrow X} \lesssim |t - s|^{-2}$$

*In addition, the following space time error estimate holds:*

$$(54) \quad \|(D_t + A + iB)H - I\|_{L^1 \rightarrow L^\infty} \lesssim 1$$

*Proof.* The bounds in (52) are trivial. So are the ones in (53) provided that  $D_t + A$  and  $B$  commute. Otherwise, they follow from (48).

It remains to prove (54). A simple computation shows that

$$[(D_t + A + iB)H - I](t, s) = i(B(t)S(t, s) - S(t, s)B(s))1_{(t-s)B(s) < 0}e^{(t-s)B(s)}$$

so (54) follows also from (48).  $\square$

**5.2. A more robust parametrix.** While the above parametrix is quite simple, it is not clear whether one can use it to show that the operator  $D_t + A + iB$  inherits most of the dispersive estimates from  $D_t + A$ . To do this we use a modified version of the above parametrix, which is somewhat reminiscent of the Littlewood-Paley theory. We consider a dyadic partition of the unity

$$1 = \sum_{j=0}^{\infty} \kappa_j^2$$

where the functions  $\kappa_j$  are supported in  $\{2^j \leq \max\{|\xi|, 1\} \leq 2^{j+2}\}$  and are smooth on the scale of their support. For  $j > 0$  we denote by  $\kappa_j^+$  respectively  $\kappa_j^-$  the parts of  $\kappa_j$  supported in  $[0, \infty)$ , respectively  $(-\infty, 0]$ . For  $j = 0$  we set  $\kappa_0^+ = \kappa_0$  and  $\kappa_0^- = 0$ . Using the functional calculus for selfadjoint operators we define the dyadic operators  $\kappa_j(B(t))$ . Then the modified parametrix  $H$  has the form

$$(55) \quad H(t, s) = 1_{t>s} \sum_j \kappa_j^-(B(t))S(t, s)\kappa_j^-(B(s))e^{(t-s)B(t)} \\ - 1_{t<s} \sum_j \kappa_j^+(B(t))S(t, s)\kappa_j^+(B(s))e^{(t-s)B(t)}$$

To measure the regularity of this parametrix we introduce some function spaces which depend only on  $A$  and not on  $B$ . This will allow us later to transfer the dispersive estimates from  $D_t - A$  to  $D_t - A + iB$ . We begin with the energy space  $L^\infty$ , and the “classical” solutions for  $D_t + A$  which are in the space  $W_A^{1,1}$  of  $X$  valued functions with norm

$$\|u\|_{W_A^{1,1}} = \|u\|_{L^\infty} + \|(D_t + A)u\|_{L^1}.$$

In between these spaces we define the space  $V_A^2$  of functions with bounded 2-variation along the  $D_t + A$  flow, with norm

$$\|u\|_{V_A^2}^2 = \|u(0)\|^2 + \sup_{(t_j) \in \mathcal{T}} \sum_j \|u(t_{j+1}) - S(t_{j+1}, t_j)u(t_j)\|^2$$

where  $\mathcal{T}$  is the set of finite increasing sequences in  $[0, 1]$ . Functions in  $V_A^2$  have at most countably many discontinuities. To eliminate functions in  $V_A^2$  which are zero a.e. we assume that all functions in  $V_A^2$  are

right continuous. The  $V_A^2$  space satisfies

$$(56) \quad W_A^{1,1} \subset V_A^2 \subset L^\infty$$

The closure of the space of smooth  $X$  valued functions in  $V_A^2$  is  $V_A^2 \cap C$ , i.e. the subspace of continuous functions in  $V_A^2$ .

Making a slight abuse of notation we denote the dual space of  $V_A^2 \cap C$  by  $(V_A^2)^*$ . This is a space of distributions which has an atomic structure. There are two kinds of atoms in  $(V_A^2)^*$ :

$$\begin{aligned} (L^1 \text{ type}) \quad & f = f_0(x)\delta_{t_0}, \quad \|f_0\| = 1 \\ (2\text{-variation type}) \quad & f = \sum_j S(t_{j+1}, t_j) f_j \delta_{t_{j+1}} - f_j \delta_{t_j}, \quad \sum_j \|f_j\|^2 = 1 \end{aligned}$$

where  $(t_j) \in \mathcal{T}$ . This has to be understood in the sense that the atomic space generated by these atoms is a weakly\* dense subspace of  $(V_A^2)^*$  with an equivalent norm. It follows from (56) that

$$L^1 \subset (V_A^2)^* \subset (D_t + A)L^\infty$$

Following is the main result of this section, which describes the mapping properties of the parametrix  $H$  in (55) in terms of the  $V_A^2$  and the  $(V_A^2)^*$  spaces.

**Proposition 5.3.** *Assume that  $A, B$  are selfadjoint operators which satisfy (48). Then the parametrix  $H$  for  $D_t + A + iB$  in (55) satisfies the estimates*

$$(57) \quad H : (V_A^2)^* \rightarrow V_A^2$$

$$(58) \quad (D_t + A + iB)H - I : (V_A^2)^* \rightarrow L^\infty$$

**Remark 5.4.** *The same result and proof apply if we construct a parametrix for the operator  $D_t + A + \alpha B + iB$ ,  $\alpha \in \mathbb{R}$  using the same formula but with  $e^{(t-s)B(s)}$  replaced with  $e^{(t-s)(1-i\alpha)B(s)}$ .*

*Proof.* We first observe that we can conjugate the result with respect to the group of isometries generated by  $A$  and reduce the problem to the case when  $A = 0$ . We omit  $A$  in the notation  $V^2 := V_0^2$  and  $(V^2)^* := (V_0^2)^*$ . The condition (48) implies that

$$(59) \quad \|(B(t) - B(s))u\| \lesssim |t - s|(\|B(s)u\| + \|u\|)$$

Next we consider the operators  $\kappa_j(B(t))$ . They depend on  $t$ , but the next lemma shows that this dependence is mild.

**Lemma 5.5.** *Suppose that  $B$  satisfies (59). Then for the operators  $\kappa_i(B(t))$  defined above we have*

*a) (bound for low modes of  $B$ )*

$$(60) \quad \|[\kappa_j(B(t)) - \kappa_j(B(s))]f\| \lesssim |t - s|2^{-j}\|B(s)f\|$$

*b) (almost orthogonality)*

$$(61) \quad \|\kappa_i(B(t))\kappa_j(B(s))\| \lesssim |t - s|2^{-|i-j|} \quad |i - j| \geq 3$$

*c) (Lipschitz bound)*

$$(62) \quad \|[\kappa_j(B(t)) - \kappa_j(B(s))]f\| \lesssim |t - s|\|f\|$$

*d) (bound for high modes of  $B$ )*

$$(63) \quad \|B(t)[\kappa_j(B(t)) - \kappa_j(B(s))]f\| \lesssim |t - s|2^j\|f\|$$

*The same estimates hold if we replace  $\kappa_j$  by  $\kappa_j^\pm$  or any other bump functions on the same scale and with similar supports.*

*Proof.* a) We use the representation

$$\kappa_j(B(t)) = \int e^{i\tau B(t)} \widehat{\kappa}_j(\tau) d\tau$$

Since  $\kappa_j$  is an integrable bump function on the  $2^{-j}$  scale, the estimate (60) follows if we prove that

$$\|(e^{i\tau B(t)} - e^{i\tau B(s)})f\| \lesssim |\tau||t - s|\|B(s)f\|$$

For this we compute

$$\frac{d}{d\tau} (e^{i\tau B(t)} e^{-i\tau B(s)}) = e^{i\tau B(t)} (B(t) - B(s)) e^{-i\tau B(s)}$$

which by (59) gives

$$\|(e^{i\tau B(t)} - e^{i\tau B(s)})f\| \lesssim \int_0^\tau \|(B(t) - B(s))e^{i\theta B(s)}f\| d\theta \lesssim |\tau||t - s|\|B(s)f\|$$

b) By duality we can assume without any restriction in generality that  $i - j \geq 3$ . Then by (60) we get

$$\begin{aligned} \|\kappa_i(B(t))\kappa_j(B(s))f\| &= \|(\kappa_i(B(t)) - \kappa_i(B(s)))\kappa_j(B(s))f\| \\ &\leq 2^{-i}|t - s|\|B(s)\kappa_j(B(s))f\| \\ &\lesssim 2^{j-i}|t - s|\|f\| \end{aligned}$$

c) We write

$$\begin{aligned} \kappa_j(B(t)) - \kappa_j(B(s)) &= (\kappa_j(B(t)) - \kappa_j(B(s))) \sum_{i < j+3} \kappa_i(B(s)) \\ &\quad + \sum_{i \geq j+3} \kappa_j(B(t))\kappa_i(B(s)) \end{aligned}$$



For the first term we use (60), while for the second we use (61).

d) Compute

$$\begin{aligned} B(t)(\kappa_j(B(t)) - \kappa_j(B(s))) &= [B(t)\kappa_j(B(t)) - B(s)\kappa_j(B(s))] \\ &\quad + (B(t) - B(s))\kappa_j(B(s)) \end{aligned}$$

For the first term we use (62) while for the second we use (59).  $\square$

We want to replace the estimates (57) and (58) by their dyadic counterparts. For this we need the following

**Lemma 5.6.** *Suppose  $B$  satisfies (59). Then we have the estimate*

$$\left\| \sum_{j \in \mathbb{N}} \kappa_j(B) u_j \right\|_{V^2}^2 \lesssim \sum_{j \in \mathbb{N}} \|u_j\|_{V^2}^2$$

and its dual

$$\sum_{j \in \mathbb{N}} \|\kappa_j(B) f\|_{(V^2)^*}^2 \lesssim \|f\|_{(V^2)^*}^2.$$

The same holds if we replace  $\kappa_j$  by any other symbols with similar support, size and regularity.

*Proof.* Denote

$$u = \sum_{j \in \mathbb{N}} \kappa_j(B) u_j.$$

Then the first estimate follows from the definition of the  $V^2$  norm and the inequality

$$(64) \quad \|u(t) - u(s)\|^2 \lesssim \sum_{j \in \mathbb{N}} \|u_j(t) - u_j(s)\|^2 + |t - s|(\|u_j\|_{L^\infty}^2 + \|v_j\|_{L^\infty}^2)$$

To prove this we write

$$\begin{aligned} u(t) - u(s) &= \sum_{j \in \mathbb{N}} \kappa_j(B(t)) u_j(t) - \kappa_j(B(s)) u_j(s) \\ &= \sum_{j \in \mathbb{N}} \kappa_j(B(t)) (u_j(t) - u_j(s)) \\ &\quad + \sum_{j \in \mathbb{N}} (\kappa_j(B(t)) - \kappa_j(B(s))) u_j(s) \end{aligned}$$

The terms in the first sum are clearly almost orthogonal, and are estimated by the first right hand side term in (64). To estimate the second sum by the second right hand side term in (64) we need to prove that its terms are almost orthogonal as well. Precisely, it would suffice to show that we have the off-diagonal decay

$$\|(\kappa_j(B(t)) - \kappa_j(B(s)))(\kappa_k(B(t)) - \kappa_k(B(s)))\| \lesssim 2^{-|k-j|} |t - s|$$

For  $|i - j| \leq 3$  this follows from (62). For  $|i - j| > 3$  we have

$$\begin{aligned} & (\kappa_j(B(t)) - \kappa_j(B(s)))(\kappa_k(B(t)) - \kappa_k(B(s))) = \\ & -\kappa_j(B(t))\kappa_k(B(s)) - \kappa_j(B(s))\kappa_k(B(t)) \end{aligned}$$

and for each of the two terms we use (61). □

We now return to the proof of the theorem. Without any restriction in generality we consider only the forward part of the parametrix  $H$ . Due to Lemma 5.6, the bound (57) for  $H$  will follow from the dyadic estimates

$$(65) \quad \|H_j f\|_{V^2} \lesssim \|f\|_{(V^2)^*}$$

where

$$H_j(t, s) = 1_{t > s} \kappa_j^-(B(t)) \kappa_j^-(B(s)) e^{(t-s)B(t)}$$

Since  $(V^2)^*$  is an atomic space it suffices to prove (65) when  $f$  is a  $(V^2)^*$  atom. We begin with an  $L^1$  type atom  $f = f_0 \delta_{t_0}$  for which (65) is a consequence of the next simple lemma:

**Lemma 5.7.** *For a fixed  $t_0 \in [0, 1]$  set*

$$v(t) = H_j(t, t_0) f_0$$

*Then*

$$\|e^{2^j(t-t_0)} v\|_{V^2} \lesssim \|f_0\|$$

The additional gain provided by the exponential factor is not needed for  $L^1$  type atoms, but we will need it later for the 2-variation type atoms.

It remains to prove (65) for an atom  $f \in (V^2)^*$  of the form

$$(66) \quad f = \sum_k (\delta_{t_{k+1}} - \delta_{t_k}) f_k$$

where  $(t_k) \in \mathcal{T}$  is an increasing finite sequence. Denote

$$u = H_j f, \quad u_k = H_j(f_k \delta_{t_{k+1}} - f_k \delta_{t_k})$$

or, more explicitly,

$$(67) \quad u_k(t) = H_j(t, t_{k+1}) f_k - H_j(t, t_k) f_k$$

For each  $k$  the function  $u_k$  is supported in  $[t_k, 1]$  and decays exponentially in time on the  $2^{-j}$  time scale. We decompose  $u$  into three parts,

$$u = v_1 + v_2 + v_3$$

where

$$\begin{aligned} v_1 &= \sum_k 1_{[t_k, t_{k+1}]} u_k, & v_2 &= \sum_{\{k: t_{k+1} - t_k > 2^{-j}\}} 1_{t > t_{k+1}} u_k, \\ v_3 &= \sum_{\{k: t_{k+1} - t_k \leq 2^{-j}\}} 1_{t > t_{k+1}} u_k. \end{aligned}$$

The terms in  $v_1$  have disjoint supports, and the square summability with respect to  $k$  is inherited from  $f_k$ . Hence it suffices to consider a single  $f_k$ , for which the bound follows from Lemma 5.7.

The terms in  $v_2$  do not have disjoint supports. However, they decay in time on the  $2^{-j}$  scale while their starting points  $t_k$  are at least  $2^{-j}$  separated because the intervals  $[s_k, t_k]$  are disjoint. Hence they are almost orthogonal, and again it suffices to consider a single  $f_k$ . But then we can use again Lemma 5.7. Note that in this case there is no significant cancellation between the inputs at times  $s_k$  and  $t_k$ .

The terms in  $v_3$  also decay exponentially on the  $2^{-j}$  scale. However, they correspond to intervals  $[t_k, t_{k+1}]$  of size less than  $2^{-j}$  which can be closer than  $2^{-j}$ , so we lose the orthogonality with respect to  $k$ . We partition the unit interval in subintervals of length  $2^{-j}$ , and group the intervals  $[t_k, t_{k+1}]$  together in bunches contained in single  $2^{-j}$  subintervals. The outputs of different bunches are almost orthogonal, so we only need to worry about a single bunch.

Within a single bunch the orthogonality is lost. However, the intervals are disjoint so within each  $2^{-j}$  subinterval we retain control of the sum of the lengths of  $[t_k, t_{k+1}]$ . Another redeeming feature is that now there is some cancellation between the input at times  $t_k$  and  $t_{k+1}$ .

Then it suffices to show that

**Lemma 5.8.** *Let  $u_k$  be as in (67). Then*

$$\|1_{t > t_{k+1}} u_k\|_{V^2} \lesssim |t_{k+1} - t_k| 2^j \|f_k\|.$$

This lemma is only interesting if  $|t_{k+1} - t_k| \leq 2^{-j}$ , otherwise it follows from Lemma 5.7. To obtain (57) it remains to prove Lemma 5.7 and Lemma 5.8.

*Proof of Lemma 5.7:* Recall that

$$v(t) = 1_{t > t_0} \kappa_j^-(B(t)) \kappa_j^-(B(t_0)) e^{(t-t_0)B(s)} f_0$$

and note the trivial bound,

$$(68) \quad \|e^{2^j(t-t_0)} v(t)\| \lesssim \|f_0\|$$

To obtain the conclusion of the lemma we prove a stronger result, which asserts that  $v$  is Lipschitz on the  $2^{-j}$  scale and decays exponentially on the same scale. More precisely, we claim that

$$\|v(\tau_1) - v(\tau_2)\| \lesssim |\tau_1 - \tau_2| 2^j e^{2^j(t_0 - \tau_1)} \|f_0\| \quad t_0 \leq \tau_1 \leq \tau_2$$

Indeed,

$$\begin{aligned} v(\tau_1) - v(\tau_2) &= [\kappa_j^-(B(\tau_1)) - \kappa_j^-(B(\tau_2))] \kappa_j^-(B(t_0)) e^{(\tau_1 - t_0)B(t_0)} f_0 \\ &\quad + \kappa_j^-(B(\tau_2)) \kappa_j^-(B(t_0)) [e^{(\tau_1 - t_0)B(s)} - e^{(\tau_2 - t_0)B(s)}] f_0 \end{aligned}$$

For the first term we use (62), while the bound for the second term is trivial since for  $\xi \in [-2^{j+2}, -2^j]$  we have

$$|\kappa_j^-(\xi) [e^{(\tau_1 - t_0)\xi} - e^{(\tau_2 - t_0)\xi}]| \lesssim 2^j |\tau_1 - \tau_2| e^{2^j(t_0 - \tau_1)}$$

□

*Proof of Lemma 5.8:* As before, begin with a pointwise estimate for  $t \geq t_{k+1}$ ,

$$(69) \quad \|u_k(t)\| \lesssim |t_{k+1} - t_k| 2^j e^{2^j(t_{k+1} - t)} \|f_k\|$$

To prove it we write

$$u_k = w_1 + w_2$$

where

$$\begin{aligned} w_1(t) &= \kappa_j^-(B(t)) \kappa_j^-(B(t_{k+1})) (e^{(t - t_{k+1})B(t_{k+1})} - e^{(t - t_k)B(t_{k+1})}) f_k \\ w_2(t) &= \kappa_j^-(B(t)) [\phi_j(B(t_{k+1})) - \phi_j(B(t_k))] f_k \end{aligned}$$

where

$$\phi_j(b) = \kappa_j^-(b) e^{(t - t_k)b}.$$

The bound for  $w_1$  follows from the inequality

$$(70) \quad |e^{(t - t_{k+1})\xi} - e^{(t - t_k)\xi}| \lesssim |t_{k+1} - t_k| 2^j e^{2^j(t_{k+1} - t)}, \quad \xi \in [-2^{j+2}, -2^j]$$

The estimate for  $w_2(t)$  follows from (62) applied to  $\phi_j$ .

Next we seek a similar Lipschitz bound for  $t_{k+1} < \tau_1 < \tau_2$ , namely

$$(71) \quad \|u_k(\tau_1) - u_k(\tau_2)\| \lesssim |\tau_1 - \tau_2| |t_{k+1} - t_k| 2^{2j} e^{2^j(t_k - \tau_1)} \|f_k\|$$

We split  $u_k$  as above,  $u_k = w_1 + w_2$ . For  $w_1$  this bound is obtained as in Lemma 5.7, using (62) and symbol bounds. It remains to prove it for  $w_2$ . We denote

$$\phi_j(\xi) = \kappa_j^-(\xi) e^{(\tau_1 - t_k)\xi}, \quad \psi_j(\xi) = \kappa_j^-(\xi) e^{(\tau_1 - t_k)\xi} (1 - e^{(\tau_1 - \tau_2)\xi})$$

and represent

$$w_2(\tau_1) - w_2(\tau_2) = w_3 + w_4$$

where

$$w_3 = (\kappa_j^-(B(\tau_1)) - \kappa_j^-(B(\tau_2)))[\phi_j(B(t_{k+1})) - \phi_j(B(t_k))]f_k$$

$$w_4 = \kappa_j^-(B(\tau_2))(\psi_j(B(t_{k+1})) - \psi_j(B(t_k)))f_k$$

For  $w_3$  we use (62) twice together with the fact that  $\phi_j$  is a bump function on the  $2^j$  scale, of size  $e^{2^j(t_{k+1}-\tau_1)}$ . Finally, the bound for  $w_4$  needs (62) for  $\psi_j$ , which is a bump function on the  $2^j$  scale and of size  $|\tau_1 - \tau_2|2^j e^{2^j(t_{k+1}-\tau_1)}$ .  $\square$

We now continue with the proof of Proposition 5.3, (58). Recall that we have reduced the problem to the case when  $A = 0$ , and denote

$$L = -i[I - (D_t + iB)H]$$

It suffices to look at the forward part of  $L$ ,

$$\begin{aligned} L(t, s) &= 1_{t>s} \sum_j \kappa_j^-(B(t))(B(t) - B(s))\kappa_j^-(B(s))e^{(t-s)B(s)} \\ &\quad + \{\partial_t \kappa_j^-(B(t))\}\kappa_j^-(B(s))e^{(t-s)B(s)}. \end{aligned}$$

We need to prove that

$$(72) \quad \|Lf\|_{L^\infty} \lesssim \|f\|_{(V^2)^*}.$$

It suffices to do this in the special case when  $f$  is an atom. We denote

$$\begin{aligned} L_j^1(t, s) &= 1_{\{t>s\}} \kappa_j^-(B(t))(B(t) - B(s))\kappa_j^-(B(s))e^{(t-s)B(s)} \\ L_j^2(t, s) &= 1_{\{t>s\}} \{\partial_t \kappa_j^-(B(t))\}\kappa_j^-(B(s))e^{(t-s)B(s)} \end{aligned}$$

The difference between these two components is that  $L_j^1$  keeps the size of the frequency, but  $L_j^2$  does not, so we need to gain some decay off the diagonal. Arguing exactly as in the case of (57), the problem reduces to the two counterparts of the estimates (68) and (69) in Lemma 5.7, respectively Lemma 5.8. These are stated in the next two lemmas. The first lemma implies (72) for  $L^1$  atoms, and also for 2-variation type atoms with  $t_{k+1} - t_k \geq 2^{-j}$ .

**Lemma 5.9.** *For  $t > s$  we have*

$$(73) \quad \|L_j^1(t, s)g\| \lesssim e^{2^j(s-t)}\|g\|$$

$$(74) \quad \|\kappa_i(B(t))L_j^2(t, s)g\| \lesssim 2^{-|i-j|}e^{2^j(s-t)}\|g\|.$$

The second lemma allows us to prove (72) for 2-variation type atoms with  $t_{k+1} - t_k \leq 2^{-j}$ .

**Lemma 5.10.** Suppose  $|t_{k+1} - t_k| \leq 2^{-j}$  and  $t > t_{k+1}$ . Then

$$(75) \quad \|(L_j^1(t, t_{k+1}) - L_j^1(t, t_k))g\| \lesssim |t_{k+1} - t_k| 2^j e^{2^j(t_{k+1}-t)} \|g\|$$

and

$$(76) \quad \|\kappa_i(B(t))(L_j^2(t, t_{k+1}) - L_j^2(t, t_k))g\| \lesssim |t_{k+1} - t_k| 2^j 2^{-|i-j|} e^{2^j(t_{k+1}-t)} \|g\|.$$

*Proof of Lemma 5.9:* The bound (73) follows from (59),

$$\begin{aligned} \|L_j^1(t, s)g\| &\lesssim |t - s| (\|B(s)\kappa_j^-(B(s))e^{(t-s)B(s)}g\| \\ &\quad + \|\kappa_j^-(B(s))e^{(t-s)B(s)}g\|) \\ &\lesssim 2^j |t - s| e^{2^j(s-t)} \|g\|. \end{aligned}$$

For (74) we note that

$$\|\kappa_j^-(B(s))e^{(t-s)B(s)}\| \lesssim e^{2^j(s-t)}.$$

Then it remains to show that

$$(77) \quad \|\kappa_i(B(t))\partial_t \kappa_j^-(B(t))\| \lesssim 2^{-|i-j|}.$$

We consider two cases. If  $i > j$  then we use (63):

$$\|\kappa_i(B(t))\partial_t \kappa_j^-(B(t))g\| \lesssim 2^{-i} \|B(t)(\partial_t \kappa_j^-(B(t)))g\| \lesssim 2^{j-i} \|g\|.$$

If  $i \leq j$  then we use duality and (60):

$$\|\partial_t \kappa_j^-(B(t))\kappa_i(B(t))g\| \lesssim 2^{-j} \|B(t)\kappa_i(B(t))g\| \lesssim 2^{i-j} \|g\|.$$

□

*Proof of Lemma 5.10:* For (75) we write

$$(L_j^1(t, t_{k+1}) - L_j^1(t, t_k))g = w_1 + w_2 + w_3$$

where

$$\begin{aligned} w_1 &= \kappa_j^-(B(t))(B(t) - B(t_{k+1}))(\phi_j(B(t_{k+1})) - \phi_j(B(t_k)))g \\ w_2 &= \kappa_j^-(B(t))(B(t) - B(t_{k+1}))\phi_j(B(t_k))(1 - e^{(t_{k+1}-t_k)B(t_k)})g \\ w_3 &= \kappa_j^-(B(t))(B(t_{k+1}) - B(t_k))\kappa_j^-(B(t_k))e^{(t-t_k)B(t_k)}g \end{aligned}$$

with

$$\phi_j(\xi) = \kappa_j^-(\xi)e^{(t-t_{k+1})\xi}.$$

For  $w_1$  we use (59) to get

$$(78) \quad \|\kappa_j^-(B(t))(B(t) - B(t_{k+1}))\| \lesssim 2^j |t - t_{k+1}|$$

and (62) to obtain

$$\|\phi_j(B(t_{k+1})) - \phi_j(B(t_k))\| \lesssim 2^j |t_{k+1} - t_k| e^{2^j(t_{k+1}-t)}.$$

For  $w_2$  and  $w_3$  we combine (78) with straightforward symbol bounds.

It remains to prove (76). Set

$$\kappa_i(B(t))(L_j^2(t, t_{k+1}) - L_j^2(t, t_k))g = w_4 + w_5$$

where

$$w_4 = \kappa_i(B(t))(\partial_t \kappa_j^-(B(t)))(\phi_j(B(t_{k+1})) - \phi_j(B(t_k)))g$$

$$w_5 = \kappa_i(B(t))(\partial_t \kappa_j^-(B(t)))\phi_j(B(t_k))(1 - e^{(t_{k+1}-t_k)B(t_k)})g$$

To estimate  $w_4$  we use (77) for the first two factors combined with (62) for the last. For  $w_5$  we combine (77) with a symbol bound for the rest.  $\square$

Now the proof of estimate (58) and hence the proof of Proposition 5.3 are complete.  $\square$

## 6. THE DISPERSIVE ESTIMATES

In this section we prove Theorems 3,4,5 using the parametrices constructed in the previous section. By Lemma 3.4 it suffices to prove the estimates (8) and (9) for the range of exponents given in the Theorems.

**6.1. The parametrices.** In order to prove Theorems 3,4,5 we need to use the parametrix in Proposition 5.2 or the one in Proposition 5.3. In either case we have to verify that the estimate (47) holds. In what follows we assume that the operator  $P$  is in canonical form,

$$P = D_t + a^w(t, x, D_x) + ib^w(t, x, D_x)$$

**Lemma 6.1.** *Suppose that  $p$  is in canonical form and that either **(A1)** and **(A2)'** hold or **(A1)'** holds. Then the following fixed time estimate is valid:*

$$\|[D_t + a^w, b^w]u\|_{L^2} \lesssim \|b^w u\|_{L^2} + \|u\|_{L^2}$$

*Proof.* We first show that **(A1)** and **(A2)'** imply **(A1)'**. Suppose that  $p = \tau + a(t, x, \xi)$  and  $q = b(t, x, \xi)$ . The principal normality condition **(A1)** takes the form

$$|\{\tau + a, b\}| = |\{a, b\} + b_t| \lesssim |\tau + a| + |b| + 1$$

Setting  $\tau = -a$  this reduces to

$$(79) \quad |\{\tau + a, b\}| = |\{a, b\} + b_t| \lesssim |b| + 1$$

If  $\tau + a, b \in \lambda S_\lambda^2$  then  $\{\tau + a, b\} \in \lambda S_\lambda^1$ . Since  $b$  is of principal type the stronger condition **(A1)'** holds by a simple division argument.

Suppose now that **(A1)'** holds. Then

$$\begin{aligned} \{a, b\} + b_t &= r_1(t, x, \tau, \xi) + r_2(t, x, \tau, \xi)(\tau + a(t, x, \xi)) \\ &+ r_3(t, x, \tau, \xi)b(t, x, \xi) + r_4(t, x, \tau, \xi) \end{aligned}$$

where

$$|r_1(t, x, \xi, \tau)| \lesssim |\tau + a(t, x, \xi)| + |b(t, x, \xi)| + 1$$

The left hand side is independent of  $\tau$ . Setting  $\tau = -a(t, x, \xi)$  on the right we obtain a similar relation of the form

$$\{a, b\} + b_t = r_1(t, x, \xi) + r_3(t, x, \xi)b(t, x, \xi) + r_4(t, x, \xi)$$

where

$$|r_1(t, x, \xi)| \lesssim |b(t, x, \xi)| + 1, \quad r_1 \in \lambda S_\lambda^2$$

Hence, after rescaling, by Theorem 3 of [22], we obtain

$$\|r_1^w(t, x, D)\|_{L^2} \lesssim \|b^w(t, x, D)u\|_{L^2} + \|u\|_{L^2}.$$

Moreover the operator

$$(r_3b)^w(t, x, D) - r_3^w(t, x, D)b^w(t, x, D)$$

is bounded in  $L^2$ . This implies the conclusion of Lemma 6.1.  $\square$

**6.2. The general case: Proof of Theorem 3.** Recall that in this case the canonical form of  $p$  is

$$p = \tau + a(t, x, \xi) + \alpha b(t, x, \xi) + ib(t, x, \xi) \quad \alpha \in \mathbb{R}$$

where the symbol  $\tau + a(t, x, \xi)$  satisfies the curvature condition **(A2)**.

We prove that we can find a parametrix  $K$  which satisfies (8) and (9). We choose a second cutoff multiplier  $\tilde{\chi}$ , identically 1 in the support of  $\chi$ . Then we define

$$K(t, s) = \tilde{\chi}^w(x, D)H(t, s).$$

where  $H$  is the parametrix in Proposition 5.3, modified as described in Remark 5.4 if  $\alpha \neq 0$ .

Given the estimates (57) and (58), in order to prove (8) and (9) it suffices to show that for  $(r, s)$  satisfying (14) we have

$$(80) \quad \chi^w : \lambda^{-\rho} L^{q'} L^{r'} \rightarrow (V_A^2)^*, \quad \tilde{\chi}^w : V_A^2 \rightarrow \lambda^\rho L^q L^r,$$

where  $A = a^w$ . These are dual estimates, and Lemma 6.2 below asserts that they are a consequence of the dispersive estimates for  $D_t + a^w$ .



**Proposition 6.2.** *Let  $q > 2$ . Suppose that the estimate (46) in Corollary 4.9 holds for the operator  $D_t + a^w$ . Then for any symbol  $\chi \in S_\lambda^2$  with support in  $B_\lambda$  we have the following microlocal embeddings:*

$$(81) \quad \chi : V_A^2 \rightarrow \lambda^{\rho(r,s)} L^q L^r, \quad \chi : \lambda^{-\rho(r,s)} L^{q'} L^{r'} \rightarrow (V_A^2)^*$$

*Proof.* It suffices to prove the first embedding, the second follows by duality. We do not use the full strength of (7), instead we only use  $L^2$  norms on the right hand side. Using the canonical form of  $p$  we write it as

$$\|\chi^w(t, x, D_x)u\|_{\lambda^{\rho(q,r)} L^q L^r} \lesssim \|(D_t + a^w(t, x, D_x))u\|_{L^2} + \|u\|_{L^2}$$

We further specialize this to solutions to the homogeneous equation,

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

for which the  $L^2$  norm of the solutions is preserved in time. Then

$$(82) \quad \|\chi^w(t, x, D_x)u\|_{\lambda^{\rho(q,r)} L^q L^r} \lesssim \|u_0\|_{L^2}$$

We define the atomic space  $U_A^q \subset L^\infty$  whose atoms have the form

$$u = \sum_{(t_k) \in \mathcal{T}} 1_{[t_k, t_{k+1})} S(t, 0) u_k, \quad \sum_k \|u_k\|_{L^2}^q = 1$$

Thus the atoms are step functions where each step is a solution to the homogeneous equation.

**Lemma 6.3.** *Suppose that (82) holds. Then*

$$\chi^w : U_A^q \rightarrow \lambda^{\rho(q,r)} L^q L^r$$

The proof of the Lemma is straightforward. It suffices to prove it for each  $U_A^q$  atom. But then we apply (82) to each step of the atom and then sum up the  $q$ 'th power of the results.

To conclude the proof of Proposition 6.2 we still need a second result, namely

**Lemma 6.4.** *Suppose that  $q > 2$ . Then  $V_A^2 \subset U_A^q$ .*

*Proof.* We can conjugate by the evolution operator  $S(t, s)$  associated to  $A$  and reduce the problem to the case when  $A = 0$ . Hence we replace  $V_A^2$  by  $V^2$  and  $U_A^q$  by  $U^q$ . Recall also that according to our convention, all  $V^2$  functions are right continuous. The same holds for all  $U^q$  functions because each atom is right continuous.

Let  $u \in V^2$  with norm 1. For each nonnegative integer  $j$  we inductively construct functions  $u_j$  and  $v_j$  and a finite disjoint partition  $\mathcal{I}_j$  of the time interval  $[0, 1]$  with the following properties:

- (a) The functions  $u_j$  are right continuous,  $\|u_j\|_{L^\infty} \leq 2^{-j}$  and  $u - u_j$  is constant on any interval  $I \in \mathcal{I}_j$ .
- (b) The functions  $v_j$  are right continuous step functions associated to the partition  $\mathcal{I}_j$ .
- (c) We have  $u_{j+1} = u_j - v_{j+1}$ .
- (d) For each  $j$ ,  $\mathcal{I}_{j+1}$  is a subpartition of  $\mathcal{I}_j$ .

This partition is constructed as follows. We initialize  $u_0 = u$ ,  $v_0 = 0$ ,  $\mathcal{I}_0 = \{[0, 1]\}$ . It remains to do the inductive step. Suppose we have  $u_j$  and  $\mathcal{I}_j$ . We partition each interval  $I \in \mathcal{I}_j$  according to the following criteria. Begin with the left endpoint  $t_I^0$ . Then choose the next point  $t_I^1$  minimal with the property that  $\|u_j(t_I^0) - u_j(t_I^1)\| \geq 2^{-j-1}$ , and continue until no such point can be found (i.e. we have reached the right end of  $I$ ). This process ends after finitely many steps, as (i) shows that  $u_j \in V^2(I)$ .

The finer partition of  $[0, 1]$  obtained in this way is denoted by  $\mathcal{I}_{j+1}$ . The function  $v_{j+1}$  is defined by

$$v_{j+1}(t) = u_j(t_I^k), \quad t \in [t_I^k, t_I^{k+1}]$$

Then we set

$$u_{j+1} = u_j - v_{j+1}$$

It is clear that the properties (a)-(d) are satisfied by construction. By (a) and (c) we obtain the representation

$$u = \sum_{j=1}^{\infty} v_j$$

which converges in  $L^\infty$  since

$$\|v_j\|_{L^\infty} \lesssim \|u_{j-1}\|_{L^\infty} + \|u_j\|_{L^\infty} \lesssim 2^{1-j} + 2^{-j}$$

Next we measure  $v_j$  in  $U^q$  as multiples of atoms. We obtain

$$\|v_j\|_{U^q} \lesssim 2^{-j} n_j^{\frac{1}{q}}$$

where  $n_j$  is the number of intervals in  $\mathcal{I}_j$ . To estimate  $n_j$  we compute the 2-variation of  $u$  with respect to the  $\mathcal{I}_j$  partition. This is where we use the inductive choice of the partition. Precisely, take  $I \in \mathcal{I}_{j-1}$ . By (a) we know that for all intervals  $J \in \mathcal{I}_j$  with  $J \subset I$ , except possibly for the last one, the variation of  $u$  between its endpoints equals the variation of  $u_j$  between its endpoints, which is at least  $2^{-j}$ . Hence we obtain

$$1 = \|u\|_{V^2}^2 \geq (n_j - n_{j-1})2^{-2j}$$

Therefore  $n_j - n_{j-1} \leq 2^{2j}$ , which after summation leads to  $n_j \lesssim 2^{2j}$ . Going back to  $v_j$  this yields

$$\|v_j\|_{U^q} \lesssim 2^{(\frac{2}{q}-1)j},$$

which in turn implies that

$$\|u\|_{U^q} \lesssim 1$$

□

This concludes the proof of Proposition 6.2. □

**6.3. The involutive case: Proof of Theorem 4.** We begin with a discussion of the geometric conditions. The condition **(A3)'** guarantees that for each  $(x, \xi) \in \Sigma$  there exist real  $\alpha, \beta$  such that the Hessian  $\partial_\xi^2(\alpha p_{re}(x, \xi) + \beta p_{im}(x, \xi))$  restricted to  $T\Sigma_x$  has rank at least  $n - 2 - k$ . For operators in canonical form this says that if  $b(t, x, \xi) = 0$  then the Hessian  $\partial_\xi^2(\alpha a(t, x, \xi) + \beta b(t, x, \xi))$  restricted to the orthogonal complement of  $b_\xi(t, x, \xi)$  has rank at least  $n - 2 - k$ . The stronger condition **(A5)'** says that for operators in canonical form the same holds with  $\alpha = 0$  and  $\beta = 1$ . This is the same as saying that the characteristic set of  $b$  has at least  $n - 2 - k$  nonvanishing curvatures. This will allow us to use Theorem 2 for  $b^w$ .

Here the dispersive estimates will follow using the simpler parametrix defined in (51), the  $L^2$  estimates of Proposition 5.2 and the dispersive estimates for operators with real symbols in Theorem 2.

We consider a multiplier  $\tilde{\chi}$  supported in  $B_\lambda$  and whose symbol equals 1 near the support of  $\chi$ . Then we define the localized parametrix  $K$  by

$$K = \tilde{\chi}^w H.$$

where  $H$  is defined in (51) with  $A = a^w$  and  $B = b^w$ . We begin with fixed time estimates.

**Proposition 6.5.** *Assume that **(A3-5)'** hold. Then the parametrix  $K$  defined above satisfies the bounds*

$$\begin{aligned} \|K(t, s)\|_{L^2 \rightarrow L^2} &\lesssim 1 \\ \|K(t, s)\chi^w\|_{L^{\frac{2(n-k)}{n-k+2}} \rightarrow L^2} &\lesssim |t-s|^{-\frac{1}{2}} \lambda^{\frac{n+k-2}{2(n-k)}}, \\ (83) \quad \|K(t, s)\|_{L^2 \rightarrow L^{\frac{2(n-k)}{n-k-2}}} &\lesssim |t-s|^{-\frac{1}{2}} \lambda^{\frac{n+k-2}{2(n-k)}} \\ \|K(t, s)\chi^w\|_{L^{\frac{2(n-k)}{n-k+2}} \rightarrow L^{\frac{2(n-k)}{n-k-2}}} &\lesssim |t-s|^{-1} \lambda^{\frac{n+k-2}{n-k}} \end{aligned}$$

$$(84) \quad \|[(D_t + A + iB)K](t, s)\chi^w\|_{L^{\frac{2(n-k)}{n-k+2}} \rightarrow L^2} \lesssim |t-s|^{-1/2} \lambda^{\frac{n+k-2}{2(n-k)}}$$

*Proof.* For each  $t$  and each  $1 \leq \mu \leq \lambda^{\frac{1}{2}}$  we define the Hilbert space

$$X_\mu(t) = \{u \in L^2; Bu \in L^2\}, \quad \|u\|_{X_\mu(t)}^2 = \mu \|u\|_{L^2}^2 + \mu^{-1} \|Bu\|_{L^2}^2$$

Its dual is given by

$$\begin{aligned} X_\mu(t)^* &= \{u = f_1 + B(t)u_2; f_1, f_2 \in L^2\}, \\ \|f\|_{X_\mu(t)^*}^2 &= \inf_{f=f_1+B(t)f_2} \mu^{-1} \|f_1\|_{L^2}^2 + \mu \|f_2\|_{L^2}^2 \end{aligned}$$

Set  $\mu = |t - s|^{-\frac{1}{2}}$ . The  $L^2 \rightarrow L^2$  estimates for  $H$  in Proposition 5.2 lead to

$$\begin{aligned} \|H(t, s)\|_{L^2 \rightarrow X_\mu(t)} &\lesssim |t - s|^{-\frac{1}{2}}, \quad \|H(t, s)\|_{X_\mu(t)^* \rightarrow L^2} \lesssim |t - s|^{-\frac{1}{2}}, \\ \|H(t, s)\|_{X_\mu(s)^* \rightarrow X_\mu(t)} &\lesssim |t - s|^{-1} \end{aligned}$$

A short computation using Lemma 5.1 (see also Lemma 3.4) also shows that

$$\|[(D_t + A + iB)\tilde{\chi}^w H](t, s)\|_{X_\mu(t)^* \rightarrow L^2} \lesssim |t - s|^{-\frac{1}{2}},$$

Then the conclusion of the proposition follows from the next lemma:

**Lemma 6.6.** *Assume that (A3-5)' hold. Then*

$$(85) \quad \chi^w : X_\mu \rightarrow \lambda^{\frac{n+k-2}{2(n-k)}} L^{\frac{2(n-k)}{n-k-2}}$$

*Proof.* We need to prove the estimate

$$\lambda^{-\frac{n-1}{n-k}} \|\chi^w u\|_{L^{\frac{2(n-k)}{n-k-2}}} \lesssim \mu \|u\|_{L^2} + \mu^{-1} \|B(t)u\|_{L^2} \quad 1 \leq \mu \leq \lambda^{\frac{1}{2}}$$

We can localize spatially on the  $\mu^{-2}$  scale. Then we rescale back to scale 1. After doing this we have reduced the problem to a similar problem but with  $\mu = 1$  and  $B(x, \xi) := \mu^{-2} B(\mu^{-2}x, \mu^2\xi)$ . Then the conclusion follows from Theorem 2 with  $\lambda := \mu^{-2}\lambda$ .  $\square$

This completes the proof of Proposition 6.5 and hence the proof of Theorem 4.  $\square$

**6.4. The degenerate involutive case: Proof of Theorem 5.** As before, we consider a multiplier  $\tilde{\chi}$  supported in  $B_\lambda$  and whose symbol equals 1 near the support of  $\chi$ . Then we define the localized parametrix  $K$  by

$$K = \tilde{\chi}^w H$$

where  $H$  is the better parametrix introduced in (55) with  $A = a^w$  and  $B = b^w$ . Because of Lemma 6.1 and Theorem 4.3, this is a good  $L^2$  parametrix. We need to show that it satisfies (8) and (9).

Our strategy is as follows. Due to the presence of the small parameter  $\delta$ , one expects that most of the kernel of the parametrix  $K$  is concentrated in phase space in a small angular neighborhood of the

Hamilton flow for  $D_t + a^w$ . Our assumption **(A6)'** shows that this is not a degenerate decay direction, therefore this part of the parametrix should satisfy good pointwise estimates. The rest of the parametrix, on the other hand, may contain bad decay directions. However, it has the redeeming quality that it is small, i.e. satisfies the same bounds as the parametrix we have constructed in the nondegenerate case.

As in the nondegenerate case, we begin with a discussion of the geometric assumptions. From **(A3)'** we know that, given  $(t, x, \xi)$  with  $b(t, x, \xi) = 0$ , there are real  $\alpha, \beta$  so that the Hessian  $\partial_\xi^2(\alpha a(t, x, \xi) + \beta b(t, x, \xi))$  restricted to the orthogonal complement of  $b_\xi(t, x, \xi)$  has rank at least  $n - 2 - k$ . The condition **(A5)'** says that the same holds with  $\alpha = 0$  and  $\beta = 1$ . **(A6)'**, on the other hand, says that we can also choose  $\alpha = 1$  and  $\beta = 0$ . In other words,  $\partial_\xi^2 a(t, x, \xi)$  must have rank  $n - 2 - k$  on the orthogonal complement of  $b_\xi$ .

**6.4.1. A pointwise fixed time bound.** Here we set up the first building block of our estimates for  $K$ , namely the pointwise estimates in directions which are close to the Hamilton flow of  $D_t + a^w$ . In the following proposition the notation  $B_1^{1,1}$  stands for a Besov space.

**Proposition 6.7.** *Let  $\varepsilon > 0$ , small. Let  $\kappa^l, \kappa^r$  be symbols whose Fourier transforms satisfy*

$$\|\hat{\kappa}^l\|_{B_1^{1,1}} \leq 1, \quad \|\hat{\kappa}^r\|_{L^1} \leq 1, \quad \text{supp } \hat{\kappa}^l, \hat{\kappa}^r \subset [-\varepsilon, \varepsilon]$$

*Then*

$$(86) \quad \|\tilde{\chi}^w \kappa^l((t-s)b^w(t))S(t,s)\kappa^r((t-s)b^w(s))\chi^w\|_{L^1 \rightarrow L^\infty} \lesssim \lambda^{\frac{n+k-2}{2}}|t-s|^{-\frac{n-k}{2}}$$

*Proof.* We change both the temporal and spatial scale, exactly as in the proof of Proposition 4.7. In the new rescaled setting the spatial and frequency scales are both equal to

$$\mu = \sqrt{|t-s|\lambda},$$

$s$  becomes 0 and  $t$  becomes 1. The estimate (86) changes to

$$(87) \quad \|\tilde{\chi}^w \kappa^l(b^w(1))S(1,0)\kappa^r(b^w(0))\chi^w\|_{L^1 \rightarrow L^\infty} \lesssim \mu^{k-1}$$

The rescaled symbols  $a, b$  satisfy (26), and in addition

$$(88) \quad |b_x| + |b_\xi| \lesssim \mu$$

The symbols  $\tilde{\chi}$  and  $\chi$  are compactly supported inside a ball

$$B_\mu = \{|x| \leq \mu, |\xi| \leq \mu\}$$

and are smooth on the scale of  $B_\mu$ .

We rewrite the above operator in the form

$$\int \hat{\kappa}^l(\theta) \tilde{\chi}^w e^{i\theta b^w(1)} S(1, 0) e^{i h b^w(0)} \chi^w \hat{\kappa}^r(h) d\theta dh$$

Since  $\hat{\kappa}^r$  is integrable we can neglect the  $h$  integration and seek a bound for

$$H_h(t, s) = \int \hat{\kappa}^l(\theta) \tilde{\chi}^w e^{i\theta b^w(1)} S(1, 0) e^{i h b^w(0)} \chi^w d\theta$$

To obtain pointwise bounds we need some representation for the kernel  $W_\theta(y, \tilde{y})$  of  $\tilde{\chi}^w e^{i\theta b^w(1)} S(t, s) e^{i h b^w(0)} \chi^w$ . In the phase space this means we move from  $(x, \xi)$  to  $(x_h, \xi_h)$  along the  $D_h - b^w(0)$  flow, to  $(x_{h,1}, \xi_{h,1})$  along the  $D_t + a^w$  flow and then further to  $(x_{h,1,\theta}, \xi_{h,1,\theta})$  the Hamilton flow for  $D_\theta - b^w(1)$ . This motivates the following

**Lemma 6.8.** *Let  $a, b$  be real symbols satisfying (26) and (88). The kernel of  $\tilde{\chi}^w e^{i\theta b^w(1)} S(1, 0) e^{i h b^w(0)} \chi^w$  has the form*

$$W_\theta(y, \tilde{y}) = \int_{B_\mu} G(\theta, x, \xi, y) e^{i\Psi} e^{-\frac{1}{2}(\tilde{y}-x)^2} e^{i\xi(x-\tilde{y})} dx d\xi + O(\mu^{-\infty})$$

with

$$\begin{aligned} \Psi &= -\xi_{h,1,\theta}(x_{h,1,\theta} - y) + \xi(x - \tilde{y}) + \psi_{-B(0)}(h, x, \xi) \\ &\quad + \psi_A(1, x_h, \xi_h) + \psi_{-B(1)}(\theta, x_{h,1}, \xi_{h,1}) \end{aligned}$$

where  $G$  is smooth, bounded, compactly supported in  $(x, \xi) \in B_\mu$ , rapidly decreasing away from  $y = x_{h,1,\theta}$  and

$$(89) \quad \left| (y - x_{h,1,\theta})^\beta \frac{\partial^\alpha}{\partial \theta^\alpha} G(\theta, x, \xi, y) \right| \lesssim c_{\alpha,\beta} \mu^{|\alpha|}$$

*Proof.* We interpret  $e^{i\theta b^w(1)} S(1, 0) e^{i h b^w(0)}$  as a single evolution operator where the generator is  $\pm b^w(0)$  up to time 0,  $a^w(t)$  for  $t \in [0, 1]$  respectively  $\pm b^w(1)$  beyond time 1. Then the representation is the one given by Proposition 4.3. In addition, in order to obtain (89) we also need the supplementary result in Proposition 4.6, which we can apply due to (88). The role of the operators  $\tilde{\chi}$  and  $\chi$  is simply to restrict the nontrivial part of  $G$  to a compact subset of  $B_\mu$ .  $\square$

We return to the proof of Proposition 6.7. We have

$$H_h(y, \tilde{y}) = \int \hat{\kappa}^l(\theta) W_\theta(y, \tilde{y}) d\theta$$

and we want to prove that

$$(90) \quad |H_h(y, \tilde{y})| \lesssim \mu^{k-1}$$

Here  $\hat{\kappa}^l(\theta) \in B_1^{1,1}$ , with support in  $[-\varepsilon, \varepsilon]$ . But  $B_1^{1,1}$  can be thought of as an atomic space where an atom  $\omega_j$  is a bounded function supported in an interval  $I$  of size  $2^{-j}$  and which is smooth on the same scale. The size of  $j$  is limited by the fact that  $I$  must be contained in  $[-\varepsilon, \varepsilon]$ . Without any restriction in generality we can assume that  $\hat{\kappa}^l(\theta)$  is such an atom,

$$\hat{\kappa}^l(\theta) = \omega_j(\theta), \quad \text{supp } \omega_j \subset [\theta_0 - 2^{-j}, \theta_0 + 2^{-j}] \subset [-\varepsilon, \varepsilon]$$

We need to consider three cases depending on the size of  $j$ .

**I. The case  $2^j \leq \mu$ .** Then we first need to integrate by parts with respect to  $\theta$ . The derivative of the phase with respect to  $\theta$  is

$$\begin{aligned} \frac{d}{d\theta} [\xi_{h,1,\theta}(x_{h,1,\theta} - y) + \psi_{-B(1)}(\theta, x_{h,1}, \xi_{h,1})] = \\ b(x_{h,1}, \xi_{h,1}) + O(\mu)(1 + |(x_{h,1,\theta} - y)|) \end{aligned}$$

while for higher order derivatives we get

$$\left| \frac{d^\alpha}{d\theta^\alpha} [\xi_{h,1,\theta}(x_{h,1,\theta} - y) + \psi_{-B(1)}(\theta, x_{h,1}, \xi_{h,1})] \right| \leq c_\alpha \mu^{|\alpha|} (1 + |(x_{h,1,\theta} - y)|)$$

Also we can use (89) for  $G$ . Hence in the region where  $|b| \gg \mu$  we can integrate by parts and get a rapidly decaying contribution. At this point we simply take absolute values and write

$$\begin{aligned} |H_h(y, \tilde{y})| \lesssim \int_{B_\mu} (1 + \mu^{-1}|b(x_{h,1}, \xi_{h,1})|)^{-N} (1 + |y - x_{h,1,\theta}|)^{-N} \\ (1 + |\tilde{y} - x|)^{-N} dx d\xi d\theta \end{aligned}$$

Since  $x_{h,1,\theta}$  is a Lipschitz function of  $x$ , it follows that the integration with respect to  $x$  is trivial, so we get

$$|H_h(y, \tilde{y})| \lesssim \int_{B_\mu} (1 + \mu^{-1}|b(\tilde{y}_{h,1}, \xi_{h,1})|)^{-N} (1 + |y - \tilde{y}_{h,1,\theta}|)^{-N} d\xi d\theta$$

Since  $|\partial_\xi b| \approx \mu$  in  $\{b = 0\}$ , the first factor in the integrand essentially restricts  $\xi$  to a neighborhood of size 1 of  $\{b = 0\}$ . Then we can evaluate the above integral by a similar integral on  $\{b = 0\}$ ,

$$|H_h(y, \tilde{y})| \lesssim \int_{\{b(\tilde{y}, \xi)=0\} \cap B_\mu} (1 + |y - \tilde{y}_{h,1,\theta}|)^{-N} d\mathcal{H}^{n-2}(\xi) d\theta$$

Also on the set  $b = 0$  we can use the principal normality condition to absorb  $h$  into  $\theta$ . More precisely, we can write

$$\tilde{y}_{h,1,\theta} = \tilde{y}_{0,1,\theta+\tilde{\theta}(h,y,\xi)} \quad \tilde{\theta} \approx h$$

Thus without any restriction in generality we assume that  $h = 0$ . Dropping the subscript  $h$ , it remains to prove that

$$\int_{\{b(\tilde{y}, \xi)=0\} \cap B_\mu} (1 + |y - \tilde{y}_{1,\theta}|)^{-N} d\mathcal{H}^{n-2}(\xi) d\theta \lesssim \mu^{k-1}$$

We also rescale  $\theta$  by a  $\mu$  factor, so that it varies on the  $\mu$  scale (same as for  $\xi$ ). Then the desired estimate becomes

$$\int_{\{b(\tilde{y}, \xi)=0\} \cap B_\mu} (1 + |y - \tilde{y}_{1,\mu^{-1}\theta}|)^{-N} d\mathcal{H}^{n-2}(\xi) d\theta \lesssim \mu^k$$

Hence we need to investigate the map

$$\xi, \theta \rightarrow \tilde{y}_{1,\mu^{-1}\theta}\xi \quad \text{in } \text{char}_{\tilde{y}} B$$

As in the proof of Proposition 4.7 one shows that this is a small Lipschitz perturbation of

$$\xi, \theta \rightarrow a_\xi + \mu^{-1}\theta b_\xi$$

The differential of this map is given by the matrix

$$\begin{pmatrix} a_{\xi\xi} + \mu^{-1}\theta b_{\xi\xi}, & \mu^{-1}b_\xi \end{pmatrix}$$

acting from  $b_\xi^\perp \times \mathbb{R}$  to  $\mathbb{R}^n$ . Since  $\theta$  is small it suffices to study this at  $\theta = 0$ . Then we need this matrix to have rank  $n - k - 1$ . Equivalently,  $a_{\xi\xi}$  must have rank  $n - k - 2$  as a quadratic form acting on the orthogonal complement of  $b_\xi$ . But this follows from our assumption **(A6)**'.

**II. The case  $\mu \leq 2^j \leq \mu^2$ .** For this range of  $j$  we can still integrate by parts with respect to  $\theta$  but only in the region where  $|b| \gg 2^j$ . Neglecting all other oscillations as well as the  $x$  integration this leads to

$$|H_h(y, \tilde{y})| \lesssim \int_{I_j} \int_{B_\mu} (1 + 2^{-j}|b(\tilde{y}_{h,1}, \xi_{h,1})|)^{-N} (1 + |y - \tilde{y}_{h,1,\theta}|)^{-N} d\xi d\theta$$

Since  $|\theta - \theta_0| \leq \mu^{-1}$  in  $I_j$  it follows that  $|y_{h,1,\theta} - y_{h,1,\theta_0}| \lesssim 1$ . Then the  $\theta$  integration is also trivial and we obtain

$$(91) \quad |H_h(y, \tilde{y})| \lesssim 2^{-j} \int_{B_\mu} (1 + 2^{-j}|b(\tilde{y}, \xi)|)^{-N} (1 + |y - \tilde{y}_{h,1,\theta_0}|)^{-N} d\xi$$

We split this integral into two regions, one where  $b$  is small,  $|b| \ll \mu^2$ , and one where  $b$  is large,  $|b| \gtrsim \mu^2$ .

In the first region the level sets of  $b$  are nondegenerate and close to zero level sets. Then we can use the coarea formula to reduce (90) in this region to bounds for integrals over level sets of  $b$ :

$$(92) \quad \int_{\{b(\tilde{y}, \xi)=b_0\} \cap B_\mu} (1 + |y - \tilde{y}_{h,1,\theta_0}|)^{-N} d\mathcal{H}^{n-2}(\xi) \lesssim \mu^k, \quad |b_0| \ll \mu^2$$



Hence we need to consider the map

$$\{b(\tilde{y}, \xi) = b_0\} \ni \xi \rightarrow \tilde{y}_{h,1,\theta_0}$$

Since  $h, \theta_0 \ll 1$ , this is a small  $\dot{C}^1$  perturbation of the map  $\xi \rightarrow a_\xi$  on  $\{b(\tilde{y}, \xi) = b_0\}$ . The differential of this map is

$$\eta \rightarrow a_{\xi\xi}\eta, \quad \eta \perp b_\xi$$

Since  $b_0 \ll \mu^2$ , this is a small  $\dot{C}^1$  perturbation of a similar map for  $b_0 = 0$ , and by **(A6)**, it has rank at least  $n - k - 2$ . The domain of integration is an  $n - 2$  dimensional cube of size  $\mu$ , so (92) follows.

It remains to consider the large values of  $|b|$ , i.e. where  $b \gtrsim \mu^2$ . For this the right hand side in (91) is largest when  $2^j = \mu^2$ , in which case we need to prove that

$$(93) \quad \int_{B_\mu} (1 + |y - \tilde{y}_{h,1,\theta_0}|)^{-N} d\xi \lesssim \mu^{k+1}$$

Now we need to consider the map

$$B_\mu \ni \xi \rightarrow \tilde{y}_{h,1,\theta_0}.$$

Since  $h, \theta_0 \ll 1$ , this is a small small  $\dot{C}^1$  perturbation of the map  $\xi \rightarrow a_\xi$ . The domain of integration is  $n - 1$  dimensional, therefore we need  $a_{\xi\xi}$  to have rank at least  $n - k - 2$ .

**III. The case  $\mu^2 \leq 2^j$ .** This is the easiest case, as there is no integration by parts. We freeze  $\theta$  as in the previous case and the estimate quickly reduces to (93).  $\square$

6.4.2. *A dyadic fixed time  $L^{r'} \rightarrow L^r$  bound.* Here we show how to combine the above pointwise bounds with the  $L^{r'} \rightarrow L^r$  previously obtained in the nondegenerate case.

**Proposition 6.9.** *Suppose that  $\phi_j, \psi_j$  are smooth bump functions on the  $2^j$  scale. Let  $2 \leq r \leq \frac{2(n-k)}{n-k-2}$  and  $q$  subject to (20). Then we have*

$$\|\tilde{\chi}^w \phi_j(b^w(t)) S(t, s) \psi_j(b^w(s)) \chi^w\|_{L^{r'} \rightarrow L^r} \lesssim \lambda^{2\rho(q,r)} 2^{\frac{2j}{q}} (1 + 2^j |t - s|)^{-\frac{2(n-k-2)}{q(n-k)}}$$

*Proof.* Let  $\varepsilon > 0$ , small. Let  $\kappa$  be a smooth function supported in  $[-\varepsilon, \varepsilon]$  which equals 1 near the origin. We define modified functions  $\tilde{\phi}_j, \tilde{\psi}_j$  by

$$\hat{\phi}_j(\theta) = \begin{cases} 0 & |t - s| < 2^{-j} \\ \kappa(|t - s|^{-1}\theta) \hat{\phi}_j(\theta) & \text{otherwise} \end{cases}$$

and similarly for  $\tilde{\psi}_j$ . The difference satisfies

$$|\phi_j(b) - \tilde{\phi}_j(b)| \lesssim (1 + 2^{-j}|b|)^{-N} (1 + 2^j|t - s|)^{-N}$$

We want to substitute  $\phi_j, \psi_j$  by  $\tilde{\phi}_j, \tilde{\psi}_j$ . To do this we consider the error term

$$E(t, s) = \phi_j(b^w(t))S(t, s)\psi_j(b^w(s)) - \tilde{\phi}_j(b^w(t))S(t, s)\tilde{\psi}_j(b^w(s))$$

The above difference bound easily leads the estimates stronger than (52), (53), namely

$$\begin{aligned} \|E(t, s)\|_{L^2 \rightarrow L^2} &\lesssim (1 + 2^j|t - s|)^{-N}, \\ \|b^w(t)E(t, s)\|_{L^2 \rightarrow L^2} &\lesssim 2^j(1 + 2^j|t - s|)^{-N}, \\ \|E(t, s)b^w(s)\|_{L^2 \rightarrow L^2} &\lesssim 2^j(1 + 2^j|t - s|)^{-N} \\ \|b^w(t)E(t, s)b^w(s)\|_{L^2 \rightarrow L^2} &\lesssim 2^{2j}(1 + 2^j|t - s|)^{-N}, \end{aligned}$$

Using these relations the  $L^{r'} \rightarrow L^r$  bound for  $\tilde{\chi}^w E(t, s)\chi^w$  follows from the curvature condition on the characteristic set of the symbol  $b$  simply by repeating the arguments in the nondegenerate case in the proof of Proposition 6.5.

It remains to prove the estimate for the operator

$$\tilde{\chi}^w \tilde{\phi}_j(b^w(t))S(t, s)\tilde{\psi}_j(b^w(s))\chi^w$$

which is zero if  $|t - s| < 2^{-j}$ . For this we interpolate the trivial  $L^2 \rightarrow L^2$  bound with an  $L^1 \rightarrow L^\infty$  bound derived from Proposition 6.7. If we set

$$\tilde{\phi}_j(b) = \kappa_j^l((t - s)b), \quad \tilde{\psi}_j(b) = \kappa_j^r((t - s)b),$$

then by definition both  $\kappa_j^l$  and  $\kappa_j^r$  are bump functions on the  $2^j|t - s|$  scale and are supported in  $[-\varepsilon, \varepsilon]$ . Then

$$\|\hat{\kappa}_j^l\|_{L^1} \lesssim 1, \quad \|\hat{\kappa}_j^r\|_{B_1^{1,1}} \lesssim 2^j|t - s|,$$

therefore Proposition 6.7 yields

$$\|\tilde{\chi}^w \tilde{\phi}_j(b^w(t))S(t, s)\tilde{\psi}_j(b^w(s))\chi^w\|_{L^1 \rightarrow L^\infty} \lesssim 2^j|t - s|\lambda^{\frac{n+k-2}{2}}|t - s|^{-\frac{n-k}{2}}$$

which is exactly what we need since  $|t - s| > 2^{-j}$  and the following relations hold:

$$2\rho(q, r) = \frac{n + k - 2}{2} \left( \frac{1}{r'} - \frac{1}{r} \right), \quad (n - k) \left( \frac{1}{r'} - \frac{1}{r} \right) = \frac{4}{q}.$$

□

6.4.3. *A dyadic  $L^2 \rightarrow L^q L^r$  bound.* Here we use a  $TT^*$  argument to derive an  $L^2 \rightarrow L^q L^r$  bound from the above fixed time bound.

**Proposition 6.10.** *Let  $\phi_j(t, b)$  be smooth bump functions on the  $2^j$  scale. Let  $l \leq j$ . Then*

$$(94) \quad \|1_{|t-s| < 2^{l-j}} \chi^w \phi_j(t, b^w(t)) S(t, s)\|_{L^2 \rightarrow L^q L^r} \lesssim 2^{\frac{2l}{q(n-k)}} \lambda^{\rho(q, r)}$$

*Proof.* Using a  $TT^*$  argument this reduces to a bound

$$\|\chi^w \phi_j(t, b^w(t)) S(t, s) \phi_j(s, b^w(s)) \chi^w\|_{L^{q'} L^{r'} \rightarrow L^q L^r} \lesssim 2^{\frac{4l}{r(n-k)}} \lambda^{2\rho(r, s)}$$

where  $t, s$  are restricted to a  $2^{l-j}$  interval. For this we use the fixed time bound in Proposition 6.9, which yield

$$\begin{aligned} & \|\chi^w \phi_j(t, b^w(t)) S(t, s) \phi_j(s, b^w(s)) \chi^w\|_{L^{q'} L^{r'} \rightarrow L^q L^r} \\ & \lesssim \lambda^{2\rho(q, r)} 2^{\frac{2j}{q}} \left( \int_0^{2^{l-j}} (1 + 2^j |t|)^{-\frac{n-k-2}{n-k}} dt \right)^{\frac{2}{q}} \\ & \approx \lambda^{2\rho(q, r)} 2^{\frac{2j}{q}} 2^{\frac{2(l-j)}{q}} 2^{-l \frac{2(n-k-2)}{q(n-k)}} = 2^{\frac{4l}{q(n-k)}} \lambda^{2\rho(q, r)}. \end{aligned}$$

We note that in effect one could obtain better bounds for  $l < 0$ , but they are not needed here.  $\square$

6.4.4. *The parametrix bound (8).* We recall that the global parametrix  $H$  is given by

$$\begin{aligned} H(t, s) &= \sum_j 1_{t < s} \kappa_j^-(\delta b^w(t)) S(t, s) \kappa_j^-(\delta b^w(s)) e^{\delta(t-s)b^w(s)} \\ &\quad - 1_{t > s} \kappa_j^+(\delta b^w(t)) S(t, s) \kappa_j^+(\delta b^w(s)) e^{\delta(t-s)b^w(s)} \end{aligned}$$

It suffices to consider the first term. We set

$$H_j(t, s) = \kappa_j^-(\delta b^w(t)) S(t, s) \kappa_j^-(\delta b^w(s)) e^{\delta(t-s)b^w(s)}$$

and apply Proposition 6.9 for each  $j$ . The symbols  $\kappa_j^-(\delta \cdot)$  are bump functions on the  $\delta^{-1} 2^j$  scale, while the exponential factor provides exponential decay on the  $2^{-j}$  time scale. Hence we obtain

$$\begin{aligned} & \|\tilde{\chi}^w H_j(t, s) \chi^w\|_{L^{r'} \rightarrow L^r} \\ & \lesssim \lambda^{2\rho(q, r)} (\delta^{-1} 2^j)^{\frac{2}{q}} (1 + \delta^{-1} 2^j |t-s|)^{-\frac{2(n-k-2)}{q(n-k)}} (1 + 2^j |t-s|)^{-N} \\ & \lesssim \lambda^{2\rho(q, r)} \delta^{-\frac{4}{q(n-k)}} 2^{\frac{2j}{q}} (2^j |t-s|)^{-\frac{2(n-k-2)}{q(n-k)}} (1 + 2^j |t-s|)^{-N} \end{aligned}$$

Next we sum this up with respect to  $j$ . For fixed  $t-s$  the largest contribution comes from the indices  $j$  which satisfies  $2^j |t-s| \approx 1$ . This gives

**Proposition 6.11.** *For  $2 \leq r \leq \frac{2(n-k)}{n-k-2}$  the operator  $H$  satisfies the estimate*

$$(95) \quad \|\tilde{\chi}^w H(t, s) \chi^w\|_{L^{r'} \rightarrow L^r} \lesssim \delta^{-\frac{4}{q(n-k)}} \lambda^{2\rho(q,r)} |t-s|^{-\frac{2}{q}}$$

If we have this then we use the Hardy-Littlewood-Sobolev inequality to obtain the  $L^{q'} L^{r'} \rightarrow L^q L^q$  bounds for  $H$ .

Next we prove the  $L^2 \rightarrow L^r L^s$  bounds for  $H$ . The  $L^{r'} L^{s'} \rightarrow L^2$  bounds are essentially dual, and their proof is similar. We decompose the operator  $H$  as

$$H = \sum_l H^l$$

where

$$H^l(t, s) = \sum_j 1_{2^{l-1-j} \leq |t-s| \leq 2^{l-j}} H_j(t, s)$$

We apply Proposition 6.10 to the terms on the right. The symbols  $\kappa_j^-(\delta \cdot)$  are smooth bumps on the  $\delta^{-1} 2^j$  scale, and the exponential provides rapid decay on the  $2^{-j}$  time scale. Then we obtain

$$\|1_{2^{l-1-j} \leq |t-s| \leq 2^{l-j}} \chi^w H_j(t, s)\|_{L^2 \rightarrow L^q L^r} \lesssim \delta^{-\frac{2}{q(n-k)}} 2^{\frac{2l}{q(n-k)}} \lambda^{\rho(q,r)} (1+2^l)^{-N}$$

We sum this with respect to  $j$  using the fact that we have orthogonality which gives square summability in  $j$  on the  $L^2$  side, while on the  $L^q L^r$  side we only need  $l^q$  summability in  $j$  because the functions we add live in disjoint time intervals. Thus we obtain

$$\|\chi^w H^l\|_{L^2 \rightarrow L^q L^r} \lesssim \delta^{-\frac{2}{q(n-k)}} 2^{\frac{2l}{q(n-k)}} \lambda^{\rho(q,r)} (1+2^l)^{-N}$$

Adding these estimates with respect to  $l$  we finally obtain

$$\|\chi^w H\|_{L^2 \rightarrow L^q L^r} \lesssim \delta^{-\frac{2}{q(n-k)}} \lambda^{\rho(q,r)}.$$

6.4.5. *The error estimates* (9). Here we prove that for each  $t \in [0, 1]$  the parametrix  $H$  satisfies the error estimates

$$(96) \quad \|(D_t + a^w + i\delta b^w)H(t, s)\tilde{\chi}\|_{L^{q'} L^{r'} \rightarrow L^2} \lesssim \delta^{-\frac{2}{q(n-k)}} \lambda^{\rho(q,r)}$$

*Proof.* It suffices to look at the forward part of the error. Then for  $t > s$  we denote

$$E(t, s) = -i(D_t - a^w + i\delta b^w)H(t, s) = \sum_j E_j(t, s)$$

where

$$E_j(t, s) = \delta \kappa_j^-(b^w(t))(b^w(t)S(t, s) - S(t, s)b^w(s))\kappa_j^-(b^w(s))e^{\delta(t-s)b^w(s)}$$

Let  $\tilde{\kappa}_j(\eta)$  be a symbol which equals  $2^{-j}\eta$  in the support of  $\kappa_j^-$  and has slightly larger support. Then

$$E_j(t, s) = 2^j \kappa_j^-(\delta b^w(t)) (\tilde{\kappa}_j(\delta b^w(t)) S(t, s) - S(t, s) \tilde{\kappa}_j(\delta b^w(s))) \kappa_j^-(\delta b^w(s)) e^{\delta(t-s)b^w(s)}$$

which we rewrite as

$$E_j(t, s) = \int_s^t 2^j \kappa_j^-(\delta b^w(t)) S(t, h) [D_t + a^w, \tilde{\kappa}_j(\delta b^w)](h) S(h, s) \kappa_j^-(\delta b^w(s)) e^{(t-s)\delta b^w(s)} dh$$

We further split the  $E_j$ 's into dyadic pieces based on the distance  $t - s$ ,

$$E_j^l(t, s) = 1_{\{2^{l-j-1} \leq t-s \leq 2^{l-j}\}} E_j(t, s)$$

Then we add them back interchanging the order of summation,

$$E^l = \sum_{j>l} E_j^l, \quad E = \sum_l E^l$$

To obtain an  $L^{q'} L^{r'} \rightarrow L^2$  bound for  $1_{t>s} E_j^l(t, s) \chi^w$  we first use the fact that the operators  $\kappa_j^-(\delta b^w(t))$ ,  $S(t, h)$  are  $L^2$  bounded. By (62) the operator  $[D_t + a^w, \tilde{\kappa}_j(\delta b^w)](h)$  is also bounded. In addition, the  $h$  integration occurs on an interval of size  $2^{l-j}$ . Then we have

$$\begin{aligned} & \|E_j^l(t, \cdot) \chi^w\|_{L^{q'} L^{r'} \rightarrow L^2} \\ & \lesssim 2^l \sup_{h<t} \|1_{s \in I^j(t, h)} S(h, s) \kappa_j^-(\delta b^w(s)) e^{(t-s)\delta b^w(s)} \chi^w\|_{L^{q'} L^{r'} \rightarrow L^2} \end{aligned}$$

where

$$I^j(t, h) = [t - 2^{l-j-1}, \min\{t - 2^{l-j}, h\}]$$

For the term on the right we use the dual of the estimate (6.10). The symbol  $\kappa_j^-(\delta \eta) e^{(t-s)\delta \eta}$  is a smooth bump function on the  $\delta^{-1} 2^j$  scale, and whose size can be bounded by  $(1 + 2^j |t - s|)^{-N}$ . Hence we obtain

$$\begin{aligned} & \|1_{s \in I^j(t, h)} S(h, s) \kappa_j^-(\delta b^w(s)) e^{(t-s)\delta b^w(s)} \chi^w\|_{L^{q'} L^{r'} \rightarrow L^2} \\ & \lesssim \delta^{-\frac{2}{q(n-k)}} \lambda^{\rho(q, r)} (1 + 2^l)^{-N} \end{aligned}$$

which implies that

$$\|E_j^l \chi^w\|_{L^{q'} L^{r'} \rightarrow L^2} \lesssim \delta^{-\frac{2}{q(n-k)}} \lambda^{\rho(q, r)} 2^l (1 + 2^l)^{-N}$$

Summing up with respect to  $j$  we obtain

$$\|E^l \chi^w\|_{L^{q'} L^{r'} \rightarrow L^2} \lesssim \delta^{-\frac{2}{q(n-k)}} \lambda^{\rho(q, r)} 2^l (1 + 2^l)^{-N}$$

There is no loss in the summation. On one hand the inputs come from disjoint time intervals, so we gain an  $\ell^{q'}$  summation with respect to  $j$ .

On the other hand the outputs are almost  $L^2$  orthogonal, so we only need an  $\ell^2$  summation in  $j$ .

The last step is to perform the summation with respect to  $l$ , which is trivial.  $\square$

## 7. APPLICATIONS TO LOCAL SOLVABILITY

In local solvability problems one considers a partial differential operator or a pseudodifferential operator  $P(x, D)$  and seeks to find a local solution  $u$  to the equation

$$P(x, D)u = f$$

for  $f$  with sufficiently small support. Local solvability is known to hold for principally normal pseudodifferential operators with  $S_{1,0}$  type symbols, see Hörmander [5] and also for low regularity symbols of type  $C^2S_{1,0}$ , see Tataru [22]. We first state the  $L^2$  solvability result proved in [22].

**Theorem 6.** *Let  $p \in C^2S_{1,0}^1$  be a principally normal pseudodifferential operator. Then  $P(x, D)$  is locally solvable with loss of one derivative, in the sense that for sufficiently small  $\varepsilon > 0$ , any ball  $B_\varepsilon$  of radius  $\varepsilon$  and any  $f \in L^2$  supported in  $B_\varepsilon$  there is  $u \in L^2(\mathbb{R}^n)$  so that  $P(x, D)u = f$  in  $B_\varepsilon$ .*

By duality this theorem reduces to proving an  $L^2$  bound from below for the adjoint operator, namely

$$(97) \quad \|v\|_{L^2} \lesssim \varepsilon \|P(x, D)^*v\|_{L^2}, \quad \text{supp } v \subset B_\varepsilon$$

This estimate is stable with respect to  $L^2$  bounded perturbations of  $P(x, D)$ . An immediate consequence of it is the following

**Corollary 7.1.** *Under the same assumptions as in Theorem 6, local solvability holds for  $P(x, D) + V$  for any potential  $V \in L^\infty$ .*

In this section we consider principally normal pseudodifferential operators  $P$  and use geometric information about their characteristic sets in order to derive local solvability for operators of the form  $P(x, D) + V$  where  $V$  is an unbounded potential. Our main result is

**Theorem 7.** *a) Let  $q, \rho(q)$  be as in (13). Let  $k \geq 0$  and  $p \in C^2S_{1,0}^{1+2\rho(q)}$  be a symbol whose restriction to frequency  $\lambda$  satisfies  $(A1)', (A2), (A3)$ . Let  $V \in L^s$  where*

$$\frac{1}{s} = \frac{1}{q'} - \frac{1}{q}$$

*Then  $P(x, D) + V$  is locally solvable with loss of one derivative in the sense that for sufficiently small  $\varepsilon > 0$ , any ball  $B_\varepsilon$  of radius  $\varepsilon$  and any*

$f \in H^{-\rho(q)} + L^{q'}$  supported in  $B_\varepsilon$  there is  $u \in H^{\rho(q)} \cap L^q$  so that  $p^w u = f$  in  $B_\varepsilon$ .

b) The same result holds if  $q, \rho(q)$  are as in (19) and the restriction of  $p$  to frequency  $\lambda$  satisfies (A1), (A2)', (A3)'.

Results of the same type but with different assumptions have been obtained also by Dos Santos [2].

**Remark 7.2.** *Implicit in this theorem is the assumption that  $P$  has order  $1 + 2\rho(q)$ . However, this is the most interesting case. If the order of  $P$  is different then one should consider the two possibilities:*

a) *If the order is larger then  $p \in C^2 S_{1,0}^{1+2\rho(\tilde{q})}$  for some  $\tilde{q} > q$ . In this situation the above result holds with  $q$  replaced by  $\tilde{q}$ . To prove it one simply needs to relax the dispersive estimates using Sobolev embeddings.*

b) *If the order of  $P$  is smaller then, under the same assumptions as in Theorem 7, one can prove that the result obtained by formally interpolating the conclusions of Corollary 7.1 and Theorem 7 is true.*

*Proof of Theorem 7.* We prove part (a), which uses Theorem 3. The proof of part (b) is similar, with the only difference that it uses Theorem 4 instead.

By duality the theorem reduces to proving a bound from below for the adjoint operator. Our main estimate is

$$(98) \quad \|v\|_{\varepsilon^{\frac{1}{8}} H^{\rho(q)} \cap L^q} \lesssim \|P(x, D)^* v\|_{\varepsilon^{-\frac{1}{8}} H^{-\rho(q)} + L^{q'}}, \quad \text{supp } v \subset B_\varepsilon$$

Multiplication by  $V$  maps  $L^q$  into  $L^{q'}$ . If  $\varepsilon$  is sufficiently small then  $V$  must be small in  $B_\varepsilon$ . Hence we can freely replace  $P(x, D)^*$  by  $P(x, D)^* + V$  and conclude the proof of the theorem.

Another useful observation is that we can modify  $P(x, D)^*$  by any perturbation which is bounded from  $H^{\rho(q)}$  into  $H^{-\rho(q)}$ . We make use of this in order to truncate the symbol  $p(x, \xi)$  at frequency  $\leq \sqrt{\lambda}$  with respect to the  $x$  variable. After this reduction, the frequency  $\lambda$  part of  $p$  belongs to  $\lambda^{1+2\rho(q)} S_\lambda^2$ .

We fix a ball  $B_r$  of fixed sufficiently small radius  $r$ , and we assume that  $B_\varepsilon \subset \frac{1}{2}B_r$ . We consider a locally finite covering of the frequency space with balls

$$\mathbb{R}^n = \bigcup_{j=0}^{\infty} B_j$$

where  $B_0 = B(0, 1)$  while for each  $j > 0$  there exists some  $\lambda > 1$  so that  $B_j \subset \{|\xi| \approx \lambda\}$ , the radius of  $B_j$  is comparable with  $\lambda$  and Theorem 3(b) can be applied in  $B_r \times B_j$  with respect to a suitable coordinate system (which may depend on  $j$ ). This is always possible

if  $r$  is sufficiently small, as discussed in Section 3. Correspondingly we choose a smooth partition of unity in the frequency space

$$1 = \chi_0(\xi) + \sum_j \chi_j(\xi)$$

Let  $\chi(x, \xi)$  be a smooth symbol supported in  $B_r \times B_j$  and which equals 1 in  $\frac{1}{2}B_r \times \text{supp } \chi_j$ . We also choose symbols  $p_j \in \lambda^{1+2\rho(q)} S_\lambda^2$  which agree with  $p$  in  $B_r \times B_j$ . From Theorem 3(b) we obtain

$$\|\chi^w w\|_{\lambda^{-\rho(q)} L^\infty L^2 \cap L^q} \lesssim \|\bar{p}_j^w(x, D)w\|_{\lambda^{\rho(q)} L^2 + \chi^w(\lambda^{\rho(q)} L^1 L^2 + L^{q'})} + \|w\|_{\lambda^{-\rho(q)} L^2}$$

We apply the above inequality to  $w = \chi_j(D)v$  with  $v$  supported in  $\frac{1}{2}B_r$ . After including some rapidly decreasing tails in the last right hand side term we obtain

$$\begin{aligned} \|\chi_j(D)v\|_{\lambda^{-\rho(q)} L^\infty L^2 \cap L^q} &\lesssim \|\bar{p}_j^w(x, D)\chi_j(D)v\|_{\lambda^{\rho(q)} L^2 + \lambda^{\rho(q)} L^1 L^2 + L^{q'}} \\ &\quad + \|\chi_j(D)v\|_{\lambda^{-\rho(q)} L^2} + \lambda^{-N} \|v\|_{L^2} \end{aligned}$$

We take a new multiplier  $\tilde{\chi}_j$  with slightly larger support, and which equals 1 in the support of  $\chi_j$ . Replacing  $\chi_j(D)$  by  $\chi_j(D)\tilde{\chi}_j(D)$  in the right hand side, after some commutations we get

$$\begin{aligned} \|\chi_j(D)v\|_{\lambda^{-\rho(q)} L^\infty L^2 \cap L^q} &\lesssim \|\chi_j(D)\bar{p}_j^w(x, D)\tilde{\chi}_j(D)v\|_{\lambda^{\rho(q)} L^2 + \lambda^{\rho(q)} L^1 L^2 + L^{q'}} \\ &\quad + \|\tilde{\chi}_j(D)v\|_{\lambda^{-\rho(q)} L^2} + \lambda^{-N} \|v\|_{L^2} \end{aligned}$$

We can also replace  $\bar{p}_j^w(x, D)$  first by  $P_j(x, D)^*$  and then by  $P(x, D)^*$  to obtain

$$\begin{aligned} \|\chi_j(D)v\|_{\lambda^{-\rho(q)} L^\infty L^2 \cap L^q} &\lesssim \|\chi_j(D)P(x, D)^*\tilde{\chi}_j(D)v\|_{\lambda^{\rho(q)} L^2 + \lambda^{\rho(q)} L^1 L^2 + L^{q'}} \\ &\quad + \|\tilde{\chi}_j(D)v\|_{\lambda^{-\rho(q)} L^2} + \lambda^{-N} \|v\|_{L^2} \end{aligned}$$

Finally we drop  $\tilde{\chi}_j$  in the last term at the expense of a rapidly decreasing contribution,

$$\begin{aligned} \|\chi_j(D)v\|_{\lambda^{-\rho(q)} L^\infty L^2 \cap L^q} &\lesssim \|\chi_j(D)P(x, D)^*v\|_{\lambda^{\rho(q)} L^2 + \lambda^{\rho(q)} L^1 L^2 + L^{q'}} \\ &\quad + \|\tilde{\chi}_j(D)v\|_{\lambda^{-\rho(q)} L^2} + \lambda^{-N} \|v\|_{L^2} \end{aligned}$$

At this point we use the assumption on the support of  $v$ . The kernels of  $\chi(D)v$  and of  $\chi(D)P(D, x)$  decay rapidly on the  $\lambda$  scale. Hence if  $\lambda^{-1} < \varepsilon^{\frac{1}{2}}$  then all the tails beyond the  $\varepsilon^{-\frac{1}{4}}$  scale are negligible. Then we use Holder's inequality to turn the  $L^\infty L^2$  and the  $L^1 L^2$  norms into  $L^2$  norms and obtain

$$\begin{aligned} \|\chi_j(D)v\|_{\varepsilon^{\frac{1}{8}} \lambda^{-\rho(q)} L^2 \cap L^q} &\lesssim \|\chi_j(D)P(x, D)^*v\|_{\varepsilon^{-\frac{1}{8}} \lambda^{\rho(q)} L^2 + L^{q'}} \\ &\quad + \|\tilde{\chi}_j(D)v\|_{\lambda^{-\rho(q)} L^2} + \lambda^{-N} \|v\|_{L^2} \end{aligned}$$



Summing up using Littlewood-Paley theory yields

$$\|\chi_{>\varepsilon^{-\frac{1}{2}}}(D)v\|_{\varepsilon^{\frac{1}{8}}H^{\rho(q)}\cap L^q} \lesssim \|P(x, D)^*v\|_{\varepsilon^{-\frac{1}{8}}H^{-\rho(q)}+L^{q'}} + \|v\|_{H^{\rho(q)}}$$

If the multiplier  $\chi_{>\varepsilon^{-\frac{1}{2}}}(D)$  were not there then for sufficiently small  $\varepsilon$  we could absorb the second right hand side term into the left hand side and conclude the proof. As it is, we also need a bound for the low frequencies. This we get since  $v$  has very small support, therefore most of its energy has to be concentrated at high frequencies:

$$\|\chi_{<\varepsilon^{-\frac{1}{2}}}v\|_{H^\rho} \ll \|v\|_{H^\rho}, \quad 0 \leq \rho < \frac{n}{2}$$

Indeed, if  $H^\rho \subset L^r$  is a sharp Sobolev embedding then

$$\|\chi_{<\varepsilon^{-\frac{1}{2}}}v\|_{H^\rho} \lesssim \varepsilon^{-\frac{\rho}{2}}\|v\|_{L^2} \lesssim \varepsilon^{\frac{\rho}{2}}\|v\|_{L^r} \lesssim \varepsilon^{\frac{\rho}{2}}\|v\|_{H^\rho}$$

The proof is even easier if  $\rho = 0$ . □

## 8. APPLICATIONS TO UNIQUE CONTINUATION

Consider a partial differential operator  $P(x, D)$  of order  $m$  in  $\mathbb{R}^n$ . Let  $\Gamma$  be an oriented hypersurface in  $\mathbb{R}^n$ , which can be represented as a nondegenerate level set of a smooth function,  $\Gamma = \{\phi = 0\}$ . The sign of  $\phi$  away from  $\Gamma$  determines the orientation of  $\Gamma$ . Denote the two sides of  $\Gamma$  by  $\Gamma^+ = \{\phi > 0\}$  and  $\Gamma^- = \{\phi < 0\}$ . Then we define the unique continuation property across  $\Gamma$  for solutions to  $P(x, D)u = 0$  as follows:

**Definition 8.1.** *We say that unique continuation property across  $\Gamma$  holds for the operator  $P(x, D)$  if for each  $x_0 \in \Gamma$  there exists a neighborhood  $V$  of  $x_0$  such that the following holds: Let  $u$  be a solution for  $P(x, D)u = 0$  in  $V$  so that  $u = 0$  in  $\Gamma^+ \cap V$ . Then  $u = 0$  near  $x_0$ .*

In other words, the values of a solution  $u$  to  $Pu = 0$  on one side of  $\Gamma$  (i.e. in  $\Gamma^+$ ) uniquely determine its values on the other side (i.e. in  $\Gamma^-$ ) near  $\Gamma$ . One can also reinterpret this as an uniqueness result for the Cauchy problem for  $P(x, D)$  in  $\Gamma^+$  with initial data on  $\Gamma$ .

Whether the unique continuation property holds depends on the geometry of the surface  $\Gamma$  relative to the operator  $P$ . One naturally introduces the pseudoconvexity condition to describe this. We let  $p(x, \xi)$  be the principal symbol of  $P$  and introduce the notation

$$p_\phi(x, \xi, \tau) = p(x, \xi + i\tau\nabla\phi)$$

**Definition 8.2.** *We say that the surface  $\Gamma$  is strongly pseudoconvex with respect to  $P$  if either*

a)  $P$  is elliptic and

$$(99) \quad \{\Re p_\phi, \Im p_\phi\} > 0 \quad \text{on } T_\Gamma^* \mathbb{R}^n \cap \{p_\phi = \{p_\phi, \phi\} = 0\}, \quad \tau > 0$$

b)  $P$  has real principal symbol and both

$$(100) \quad \{p, \{p, \phi\}\} > 0 \quad \text{on } T_\Gamma^* \mathbb{R}^n \cap \{\{p, \phi\} = p = 0\}$$

and (99) hold.

Note that the property of pseudo-convexity only depends on  $\Gamma$  and its orientation and not on  $\phi$ . There is also a version of this which applies to principally normal operators, but here we choose to keep things simple. Note also that for anisotropic operators such as the heat or the Schrödinger operator one has to make some obvious adjustments in the definition of the principal symbol and of the Poisson bracket.

The pseudoconvexity condition does not preclude surfaces from being characteristic at least at some points. However, here we assume for simplicity that this not is the case, namely

$$p(x, \nabla \phi) \neq 0$$

Now we can state the main result (see Hörmander [7] and references therein, and also Isakov [8] for the anisotropic case):

**Theorem 8.** *Let  $P$  be an operator with  $C^1$  coefficients which is either elliptic or has real principal symbol. Suppose that the oriented surface  $\Gamma$  is strongly pseudoconvex with respect to  $P$ . Then unique continuation across  $\Gamma$  holds for  $P$ .*

A main tool in proving unique continuation results is provided by the Carleman estimates. To describe them we need to introduce the notion of pseudoconvex functions.

**Definition 8.3.** *We say that the function  $\phi$  is strongly pseudoconvex with respect to  $P$  if either*

a)  $P$  is elliptic and

$$(101) \quad \{\Re p_\phi, \Im p_\phi\} > 0 \quad \text{on } \{p_\phi = 0\}, \quad \tau > 0$$

b)  $P$  has real principal symbol and both

$$(102) \quad \{p, \{p, \phi\}\} > 0 \quad \text{on } \{p = 0\}$$

and (101) hold.

The nondegenerate level sets of pseudoconvex functions are pseudoconvex surfaces. Conversely, any pseudoconvex surface is a nondegenerate level set of some pseudoconvex function.

The  $L^2$  Carleman estimates below imply the above unique continuation result via a standard argument.

**Theorem 9** ( $L^2$  Carleman estimates). *Let  $P$  be an operator with  $C^1$  coefficients which is either elliptic or has real principal symbol. Suppose that  $\phi$  is strongly pseudoconvex with respect to  $P$  in some bounded open  $\Omega \subset \mathbb{R}^n$ . Given any compact subset  $K$  of  $\Omega$  there are  $c, \tau_0 > 0$  so that for all functions  $u$  supported in  $K$  we have:*

(a) *If  $P$  is elliptic:*

$$(103) \quad \tau^{-1} \|e^{\tau\phi} u\|_{H_\tau^m}^2 \leq c \|e^{\tau\phi} P(x, D)u\|_{L^2}^2, \quad \tau \geq \tau_0$$

(b) *If  $P$  has real principal symbol:*

$$(104) \quad \tau \|e^{\tau\phi} u\|_{H_\tau^{m-1}}^2 \leq c \|e^{\tau\phi} P(x, D)u\|_{L^2}^2, \quad \tau \geq \tau_0$$

Here and below, the spaces  $H_\tau^k$  are defined like the usual Sobolev but giving to  $\tau$  the same weight as a derivative. Precisely,

$$\|u\|_{H_\tau^k} \approx \|(|D| + \tau)^k u\|_{L^2}$$

A similar meaning is associated to the notation  $H_\tau^{k,p}$ .

Our interest lies in replacing the  $L^2$  estimates with  $L^p$  estimates. This is useful in problems with unbounded potentials. These can also arise as linearizations of nonlinear problems.

**Theorem 10** ( $L^p$  Carleman estimates, elliptic case). *Let  $P$  be an elliptic operator of order  $m$  with  $C^1$  coefficients. Let  $\phi$  be a strongly pseudoconvex with respect to  $P$  in some compact set  $\Omega \subset \mathbb{R}^n$ . Assume that the characteristic set of  $p_\phi$  has  $n - 2 - k$  nonvanishing curvatures, and let  $r, \rho(r)$  be as in (19). Then there are  $c, \tau_0 > 0$  so that for all functions  $u$  supported in  $\Omega$  we have:*

$$(105) \quad \|e^{\tau\phi} u\|_{\tau^{\frac{1}{4}} H_\tau^m \cap H_\tau^{m-\frac{1}{2}-\rho(r),r}}^2 \leq c \|e^{\tau\phi} P(x, D)u\|_{\tau^{-\frac{1}{4}} L^2 + H_\tau^{\frac{1}{2}+\rho(r),r'}}^2, \quad \tau \geq \tau_0$$

**Theorem 11** ( $L^p$  Carleman estimates, real case). *Let  $P$  be an operator of order  $m$  with real principal symbol and  $C^2$  coefficients. Let  $\phi$  be a strongly pseudoconvex with respect to  $P$  in some compact set  $\Omega \subset \mathbb{R}^n$ . Assume that the characteristic set of  $p$  has  $n - 1 - k$  nonvanishing curvatures and the characteristic set of  $p_\phi$  has  $n - 2 - k$  nonvanishing curvatures, and let  $r, \rho(r)$  be as in (13). Then there are  $c, \tau_0 > 0$  so that for all (smooth) functions  $u$  supported in  $\Omega$  we have:*

$$(106) \quad \|e^{\tau\phi} u\|_{\tau^{-\frac{1}{4}} H_\tau^{m-1} \cap H_\tau^{m-1-\rho(r),r}}^2 \leq c \|e^{\tau\phi} P(x, D)u\|_{\tau^{\frac{1}{4}} L^2 + H_\tau^{\rho(r),r'}}^2, \quad \tau \geq \tau_0$$

In most cases one can also obtain mixed norm estimates. However, we preferred to have simpler statements for general operators. Later on when we consider examples we also state mixed norm estimates for equations where this is relevant such as the heat equation and the Schrödinger equation.

Also at this point we contend ourselves with the Carleman estimates. We state the corresponding unique continuation statements only for the examples we consider below, and we leave to the reader the task of deriving the corresponding unique continuation results for other problems of interest.

Observe that, given any estimate of the form

$$\|e^{\tau\phi}u\|_X^2 \leq c\|e^{\tau\phi}P(x, D)u\|_Y$$

the substitution  $v = e^{\tau\phi}u$  transforms it into

$$\|v\|_X^2 \leq c\|P_\phi(x, D, \tau)v\|_Y$$

Hence we need to understand the geometry of the operators  $P_\phi$ . Note that only the principal part of  $P_\phi$  is important, all the lower order terms are negligible due to the  $L^2$  part of the Carleman estimates.

The operator  $P_\phi$  is an operator with complex symbol, therefore we would like to apply our results for principally normal operators. Hence we want  $p_\phi$  to be principally normal. However, the pseudoconvexity condition shows that this is not the case, more precisely

$$\{\Re p_\phi, \Im p_\phi\} > 0 \quad \text{in} \quad p_\phi = 0$$

To overcome this difficulty we use a two scale approach which begins with the observation that the  $L^2$  Carleman estimates allow a localization to the  $\tau^{-\frac{1}{2}}$  scale. Hence, on one hand we apply our dispersive estimates on the  $\tau^{-\frac{1}{2}}$  spatial scale. This scale turns out to be sufficiently small so that the commutator between  $\Re p_\phi$  and  $\Im p_\phi$  becomes negligible, i.e. we gain the principal normality. On the other hand, in order to combine these localized results we use the global  $L^2$  estimate.

The part of the proof of Theorem 10 which is obtained by assembling together spatially localized estimates on the  $\tau^{-\frac{1}{2}}$  is contained in the following two lemmas.

**Lemma 8.4.** *Under the assumptions in Theorem 10 there is a parametrix  $K$  for  $P_\phi$  which satisfies the bounds*

$$(107) \quad \|Kf\|_{\tau^{\frac{1}{4}}H_\tau^m \cap H_\tau^{m-\frac{1}{2}-\rho(r),r}} \lesssim \|f\|_{\tau^{-\frac{1}{4}}L^2 + H_\tau^{\frac{1}{2}+\rho(r),r'}}$$

$$(108) \quad \|(I - P_\phi K)f\|_{\tau^{-\frac{1}{4}}L^2} \lesssim \|f\|_{\tau^{-\frac{1}{4}}L^2 + H_\tau^{\frac{1}{2}+\rho(r),r'}}$$

**Lemma 8.5.** *Under the assumptions in Theorem 10 the operator  $P_\phi$  satisfies*

$$(109) \quad \|w\|_{H_\tau^{m-\frac{1}{2}-\rho(r),r}} \lesssim \|w\|_{\tau^{\frac{1}{4}}H_\tau^m} + \|P_\phi w\|_{\tau^{-\frac{1}{4}}L^2}$$

Before proving the lemmas we show how they can be combined with the  $L^2$  Carleman estimates to prove Theorem 10.

*Proof of Theorem 10.* We need to prove that

$$\|v\|_{\tau^{\frac{1}{4}}H_\tau^m \cap H_\tau^{m-\frac{1}{2}-\rho(r),r}}^2 \leq c \|P_\phi(x, D, \tau)v\|_{\tau^{-\frac{1}{4}}L^2 + H_\tau^{\frac{1}{2}+\rho(r),r'}}^2,$$

Let  $K$  be as in Lemma 8.4. We decompose  $v$  into

$$v = w + KP_\phi v.$$

The function  $KP_\phi v$  satisfies the correct bounds by (107) while

$$P_\phi w = (I - P_\phi K)P_\phi v.$$

Using (108) we bound the right hand side in  $L^2$ :

$$\|P_\phi w\|_{\tau^{-1/4}L^2} \leq \|P_\phi v\|_{\tau^{-\frac{1}{4}}L^2 + H_\tau^{\frac{1}{2}+\rho(r),r'}}$$

But the  $L^2$  Carleman estimate (105) allows us to also obtain an  $L^2$  estimate for  $w$ ,

$$\|w\|_{\tau^{\frac{1}{4}}H_\tau^m} \lesssim \|P_\phi w\|_{\tau^{-\frac{1}{4}}L^2}.$$

Hence we can use Lemma 8.5 to obtain the correct estimate for  $w$  and conclude the proof of Theorem 10.  $\square$

*Proof of Lemmas 8.4, 8.5.* Without any restriction in generality we assume that  $\tau^{-\frac{1}{2}} \ll d(K, \partial\Omega)$ . Then we claim that it suffices to prove both lemmas in a ball of radius  $\tau^{-\frac{1}{2}}$ . For this we consider a locally finite covering of  $\Omega$  with balls of radius  $\tau^{-\frac{1}{2}}$ ,

$$\Omega \subset \bigcup_j B_j$$

Correspondingly we consider a smooth partition of unity

$$1 = \sum_j \chi_j, \quad \text{supp } \chi_j \subset B_j$$

Suppose we know that for each  $j$  there is a parametrix  $K_j$  so that for  $f$  supported in  $B_j$  the function  $K_j f$  is supported in  $2B_j$  and the estimates (107), (108) hold. Then we can construct a parametrix  $K$  for  $P_\phi$  by

$$K = \sum_j K_j \chi_j$$

To obtain the estimates (107), (108) for  $K$  we need to verify that we can sum up the bounds for  $K_j$  in  $l^2$ ,

$$\sum \|\chi_j f\|_{\tau^{-\frac{1}{4}}L^2 + H_\tau^{\frac{1}{2} + \rho(r), r'}}^2 \lesssim \|f\|_{\tau^{-\frac{1}{4}}L^2 + H_\tau^{\frac{1}{2} + \rho(r), r'}}^2$$

respectively

$$\left\| \sum K_j \chi_j f \right\|_{\tau^{\frac{1}{4}}H_\tau^m \cap H_\tau^{m - \frac{1}{2} - \rho(r), r}}^2 \lesssim \sum \|K_j \chi_j f\|_{\tau^{\frac{1}{4}}H_\tau^m \cap H_\tau^{m - \frac{1}{2} - \rho(r), r}}^2$$

Similarly, if (109) holds for  $w$  supported in  $B_j$  then we can conclude it holds in general using the previous inequality and the additional estimate

$$\sum_j \|\chi_j w\|_{\tau^{\frac{1}{4}}H_\tau^m}^2 + \|P_\phi \chi_j w\|_{\tau^{-\frac{1}{4}}L^2}^2 \lesssim \|w\|_{\tau^{\frac{1}{4}}H_\tau^m}^2 + \|P_\phi w\|_{\tau^{-\frac{1}{4}}L^2}^2$$

These three inequalities are easy exercises which are left for the reader. We only observe that while the first two can be localized further down to the  $\tau^{-1}$  spatial scale, this would be useless as a parametrix satisfying the right bounds cannot be constructed on a smaller scale. On the other hand for the third bound the  $\tau^{-\frac{1}{2}}$  scale is optimal because the commutators  $[P_\phi, \chi_j]$  have to be controlled.

To prove the lemmas in a ball of radius  $\tau^{-\frac{1}{2}}$  we rescale it to the unit ball. The rescaled operator is

$$\tilde{P}_\phi(x, D, \mu) = P_\phi\left(\frac{x}{\mu}, D, \mu\right), \quad \mu = \tau^{\frac{1}{2}}$$

Then for Lemma 8.4 we need a parametrix  $\tilde{K}$  for  $\tilde{P}_\phi$  which satisfies

$$(110) \quad \|\tilde{K}f\|_{\mu^{\frac{1}{2}}H_\mu^m \cap H_\mu^{m - \frac{1}{2} - \rho(r), r}} \lesssim \|f\|_{\mu^{-\frac{1}{2}}L^2 + H_\mu^{\frac{1}{2} + \rho(r), r}}$$

$$(111) \quad \|(I - \tilde{P}_\phi K)f\|_{\mu^{-\frac{1}{2}}L^2} \lesssim \|f\|_{\mu^{-\frac{1}{2}}L^2 + H_\mu^{\frac{1}{2} + \rho(r), r'}}$$

while for Lemma 8.5 we need the estimate

$$(112) \quad \|w\|_{H_\tau^{m - \frac{1}{2} - \rho(r), r}} \lesssim \|w\|_{\mu^{\frac{1}{2}}H_\mu^m} + \|P_\phi w\|_{\mu^{-\frac{1}{2}}L^2}$$

Within the unit ball the coefficients of  $\tilde{P}_\phi$  vary by  $O(\mu^{-1})$ . Then

$$\|(\tilde{P}_\phi(x, D, \mu) - \tilde{P}_\phi(0, D, \mu))w\|_{\mu^{-\frac{1}{2}}L^2} \lesssim \|w\|_{\mu^{\frac{1}{2}}H_\mu^m}$$

therefore without any loss we can freeze the coefficients of  $P_\phi$  and replace  $\tilde{P}_\phi(x, D, \mu)$  by  $\tilde{P}_\phi(0, D, \mu)$ . This is principally normal by default.

The symbol

$$\tilde{p}_\phi(x, \xi, \mu) = p(0, \xi + i\mu\nabla\phi(0))$$

is elliptic in the region  $\tau \ll |\xi|$ . Then the only region in frequency where the problem is nontrivial is  $\{|\xi| \lesssim \tau\}$ . This is where we use

the curvature condition. For low frequencies  $|\xi| \lesssim \mu$  we can use the parametrix  $K$  given by Theorem 4, while at higher frequencies  $\tilde{P}_\phi(0, D, \mu)$  is elliptic. More precisely, we consider a large enough constant  $C$  so that  $\tilde{P}_\phi(0, D, \mu)$  is elliptic in  $\{|\xi| > C\mu\}$ . Then we denote by  $\chi$  a symbol supported in  $\{|\xi| > 2C\mu\}$  and which equals 1 in the region  $\{|\xi| < C\mu\}$ . The curvature condition in the hypothesis of Theorem 10 implies that we can use Theorem 4 to produce a parametrix  $K_\mu$  in the region  $\{|\xi| < 2C\mu\}$ .

Then we define the parametrix  $K$  for  $P_\phi(0, D, \mu)$  by

$$K = K_\mu \chi(D) + P_\phi(0, D, \mu)^{-1} (1 - \chi(D))$$

The bounds for  $K$  follow easily from the similar bounds in Theorem 4 for  $K_\mu$ . This concludes the proof of Lemma 8.4.

For Lemma 8.5 we use the same setup. The bound for  $(1 - \chi(D))w$  follows from ellipticity, while the bound for  $\chi(D)w$  is nothing but (4) in the context of Theorem 4.  $\square$

An argument as the one in the proof of Theorem 10 shows that Theorem 11 is a consequence of the following counterparts of Lemmas 8.4, 8.5:

**Lemma 8.6.** *Under the assumptions in Theorem 11 there is a parametrix  $K$  for  $P_\phi$  which satisfies the bounds*

$$(113) \quad \|Kf\|_{\tau^{-\frac{1}{4}} H_\tau^{m-1} \cap H_\tau^{m-1-\rho(r),r}} \lesssim \|f\|_{\tau^{\frac{1}{4}} L^2 + H_\tau^{\rho(r),r}}.$$

$$(114) \quad \|(I - P_\phi K)f\|_{\tau^{\frac{1}{4}} L^2} \lesssim \|f\|_{\tau^{\frac{1}{4}} L^2 + H_\tau^{\rho(r),r'}}.$$

**Lemma 8.7.** *Under the assumptions in Theorem 11 the operator  $P_\phi$  satisfies*

$$(115) \quad \|w\|_{H_\tau^{m-1-\rho(r),r}} \lesssim \|w\|_{\tau^{-\frac{1}{4}} H_\tau^{m-1}} + \|P_\phi w\|_{\tau^{\frac{1}{4}} L^2}$$

*Proof of Lemmas 8.6, 8.7.* As in the proof of Lemmas 8.4, 8.5 we can still localize on the  $\tau^{-\frac{1}{2}}$  scale, and then rescale it back to the unit ball. However, now we can no longer freeze the coefficients of the rescaled operator

$$\tilde{P}_\phi(x, D, \mu) = P_\phi\left(\frac{x}{\mu}, D, \mu\right),$$

because its characteristic set is nontrivial at high frequencies. We divide the Fourier space in dyadic regions

$$D_\lambda = \{|\xi| \leq 2\mu\}, \quad \lambda = \mu$$

$$D_\lambda = \left\{ \frac{\lambda}{2} \leq |\xi| \leq 2\lambda \right\}, \quad \lambda = 2^j \mu, \quad j > 0$$

Correspondingly we consider a partition of unity in frequency

$$1 = \sum_{\lambda=2^j \mu}^{j \geq 0} \chi_\lambda(\xi), \quad \text{supp } \chi_\lambda \subset D_\lambda$$

We also consider symbols  $\tilde{\chi}_\lambda$  with slightly larger support which equal 1 near the support of  $\chi_\lambda$ . Write the principal part of  $P_\phi$  in the form

$$P_\phi = \sum_{|\alpha|=m} c_\alpha(x) (D, \tau)^\alpha$$

where  $\tau$  carries the same weight as a derivative. Then

$$\tilde{P}_\phi = \sum_{|\alpha|=m} c_\alpha(x/\mu) (D, \mu)^\alpha$$

For  $\lambda = 2^j \mu$  we define the regularized coefficients

$$c_{\alpha, \lambda} = S_{< \lambda^{\frac{1}{2}}} c_\alpha$$

where  $S_{< \lambda^{\frac{1}{2}}}$  is a multiplier with smooth symbol which selects the frequencies  $\lambda^{\frac{1}{2}}$  and smaller. These do not differ much from the original coefficients,

$$(116) \quad \|c_{\alpha, \lambda} - c_\alpha\|_{L^\infty} \lesssim \lambda^{-1} \|c_\alpha\|_{C^2}$$

We also introduce the modified operators,

$$\tilde{P}_{\phi, \lambda} = \sum_{|\alpha|=m} c_{\alpha, \lambda}(x/\mu) (D, \mu)^\alpha$$

somewhat in the spirit of the paradifferential calculus.

In the region  $D_\lambda$  these symbols are in  $\lambda^m S_\lambda^2$ . They also satisfy the principal normality condition. Indeed, since  $P$  has real coefficients, it follows that all terms in  $\Im \tilde{P}_{\phi, \lambda}$  contain at least one power of  $\mu$ . Then

$$|\{\Re \tilde{P}_{\phi, \lambda}, \Im P_{\phi, \lambda}\}| \lesssim \lambda^{2(m-1)} \quad \text{in } D_\lambda$$

By (116) the curvature condition in  $D_\lambda$  is easily transferred from  $\tilde{P}_\phi$  to  $P_{\phi, \lambda}$ . If  $\lambda \approx \mu$  then this is given in the hypothesis of the Theorem 11. If  $\lambda \gg \mu$  then  $\Re \tilde{P}_\phi$  is a small perturbation of  $\tilde{P} = P(x/\mu, D)$ , therefore we can use the curvature condition for  $P$ .

Now we are in a position to use Theorem 4 for  $\lambda \approx \mu$  and Theorem 3 for  $\lambda \gg \mu$ . We obtain parametrices  $K_\lambda$  which satisfy

$$\begin{aligned} \lambda^{m-1} \|K_\lambda \chi_\lambda f\|_{L^2 \cap \lambda^{\rho(r)} L^r} &\lesssim \|f\|_{L^2 + \lambda^{-\rho(r)} L^{r'}} \\ \|(I - \tilde{P}_{\phi, \lambda} K_\lambda) \chi_\lambda f\|_{L^2} &\lesssim \|f\|_{L^2 + \lambda^{-\rho(r)} L^{r'}} \end{aligned}$$



Now we can define the parametrix  $K$  for  $\tilde{P}_\phi$  as

$$Kf = \sum_{\lambda} \tilde{\chi}_{\lambda} K_{\lambda} \chi_{\lambda} f$$

It remains to prove that  $K$  satisfies the desired bounds. The bound for  $K$  follows from the bound for  $K_j$  and Littlewood-Paley theory. Consider now the error estimates. We have

$$\begin{aligned} I - \tilde{P}_{\phi} K &= \sum_{\lambda} I - \tilde{P}_{\phi, \lambda} \tilde{\chi}_{\lambda} K_{\lambda} \chi_{\lambda} \\ &= \sum_{\lambda} (\tilde{P}_{\phi} - \tilde{P}_{\phi, \lambda}) \tilde{\chi}_{\lambda} K_{\lambda} \chi_{\lambda} + [\tilde{P}_{\phi, \lambda}, \tilde{\chi}_{\lambda}] K_{\lambda} \chi_{\lambda} \\ &\quad + \sum_{\lambda} \tilde{\chi}_{\lambda} (I - \tilde{P}_{\phi, \lambda} K_{\lambda}) \chi_{\lambda} \end{aligned}$$

For the first term we use (116), for the second an  $L^2$  commutator estimate. Finally the bound for the third is given by the similar estimates for  $K_{\lambda}$ . For each term we also need to use the Littlewood-Paley theory. This concludes the proof of Lemma 8.6.

For Lemma 8.7 we use Theorem 4 for  $\lambda \approx \mu$  respectively Theorem 3 for  $\lambda \gg \mu$  (In both case we use the theorems in the form of estimate (4)). These imply that

$$\lambda^{m-1} \|\chi_{\lambda} w\|_{\lambda^{\rho(r)} L^r} \lesssim \lambda^{m-1} \|\tilde{\chi} w\|_{L^2} + \|\tilde{P}_{\phi, \lambda} \tilde{\chi} w\|_{L^2}$$

which after some commuting we can square and sum up using the Littlewood-Paley theory.  $\square$

**8.1. The Laplace equation.** Consider a second order elliptic operator in  $\mathbb{R}^n$ ,

$$P = \partial_j g^{jk}(x) \partial_k$$

Any surface  $\Gamma$  is strongly pseudoconvex with respect to  $P$ . If  $\phi$  is a strongly pseudoconvex function with respect to  $P$  with  $\nabla \phi \neq 0$  then

$$p_{\phi}(x, \xi, \tau) = g^{jk} \xi_j \xi_k - \tau^2 g^{jk} \partial_j \phi \partial_k \phi + 2i\tau g^{jk} \xi_j \partial_k \phi$$

The characteristic set of the real part is an ellipsoid centered at the origin, while the characteristic set of the imaginary part is a plane through the origin. The characteristic set of the full operator is the intersection of the ellipsoid with the plane, and has  $n - 2$  nonvanishing curvatures. Therefore, we can use Theorem 10 with  $k = 0$ :

**Theorem 12.** *Let  $P$  be a second order elliptic operator with  $C^1$  coefficients. Let  $\phi$  be a strongly pseudoconvex function with respect to  $P$ .*

Then for compactly supported  $u$  we have

$$\|e^{\tau\phi}u\|_{H_{\tau}^{\frac{2}{n+2}, \frac{2(n+2)}{n}} \cap \tau^{\frac{1}{4}}H_{\tau}^1} \lesssim \|e^{\tau\phi}Pu\|_{H_{\tau}^{-\frac{2}{n+2}, \frac{2(n+2)}{n+4}} + \tau^{-\frac{1}{4}}H_{\tau}^{-1}}, \quad \tau \geq \tau_0$$

Note that this is not precisely in the form stated in (105) but it can be easily obtained from it by conjugating  $P_{\phi}$  with a first order elliptic multiplier because  $P$  is in divergence form. We prefer the above formulation because of its symmetry. Similar adjustments are made in all the other examples we consider.

Applied to unique continuation problems, this yields

**Theorem 13.** *Let  $P$  be a second order elliptic operator with  $C^1$  coefficients. Let  $\Gamma$  be a smooth surface. Then unique continuation for  $P + V$  across  $\Gamma$  holds for all potentials  $V$  which have the multiplicative mapping property*

$$V : H^{\frac{2}{n+2}, \frac{2(n+2)}{n}} \rightarrow H^{-\frac{2}{n+2}, \frac{2(n+2)}{n+4}}$$

This includes the case  $V \in L^{\frac{n}{2}}$ , for which the result was proved by Wolff [24]. This can be relaxed to a slightly larger Morrey space.

For strong unique continuation problems and problems involving gradient potentials we refer the reader to Wolff [25], the authors paper [13], and the references therein.

**8.2. The wave equation.** Consider a second order hyperbolic operator in  $\mathbb{R}^{d+1}$ ,

$$P = \partial_j g^{jk}(x) \partial_k$$

where the matrix  $g^{ij}$  has signature  $(d, 1)$ . Which noncharacteristic surfaces  $\Gamma$  are strongly pseudoconvex with respect to the  $P$ ? One needs to distinguish between space like surfaces ( $g^{jk} \partial_j \phi \partial_k \phi < 0$ ) and time-like surfaces ( $g^{jk} \partial_j \phi \partial_k \phi > 0$ ). All space-like surfaces are pseudoconvex, since the Cauchy problem with initial data on a space-like surface is well-posed. For time-like surfaces, on the other hand, the condition (99) is trivially fulfilled but (100) may or may not hold.

If  $\phi$  is a strongly pseudoconvex function with respect to  $P$  with  $\nabla \phi \neq 0$  then

$$p_{\phi}(x, \xi, \tau) = g^{jk} \xi_j \xi_k - \tau^2 g^{jk} \partial_j \phi \partial_k \phi + 2i\tau g^{jk} \xi_j \partial_k \phi$$

The characteristic set of  $p$  is a cone which has  $d - 1$  nonvanishing curvatures. The characteristic set of the real part of  $p_{\phi}$  is a hyperboloid, which has  $d$  nonvanishing curvatures. The characteristic set of the full operator  $p_{\phi}$  is the intersection of the hyperboloid with a plane, and has  $d - 1$  nonvanishing curvatures. Hence we can use Theorem 11 with  $k = 1$ ,  $n = d + 1$ :

**Theorem 14.** *Let  $P$  be a second order hyperbolic operator with  $C^2$  coefficients. Let  $\phi$  be a strongly pseudoconvex function with respect to  $P$ . Then for compactly supported  $u$  we have*

$$\|e^{\tau\phi}u\|_{L^{\frac{2(d+1)}{d-1}} \cap \tau^{-\frac{1}{4}}H_{\tau}^{\frac{1}{2}}} \lesssim \|e^{\tau\phi}Pu\|_{L^{\frac{2(d+1)}{d+3}} + \tau^{\frac{1}{4}}H_{\tau}^{-\frac{1}{2}}}, \quad \tau > \tau_0$$

Applied to unique continuation problems, this gives

**Theorem 15.** *Let  $P$  be a second order hyperbolic operator with  $C^2$  coefficients. Let  $\Gamma$  be a smooth surface which is strongly pseudoconvex with respect to  $P$ . Then unique continuation for  $P + V$  across  $\Gamma$  holds for all potentials  $V \in L^{\frac{d+1}{2}}$ .*

This improves an earlier result in [20]. One can also produce versions of this with potentials in mixed norm spaces.

**8.3. The heat equation.** Consider a second order parabolic operator in  $\mathbb{R} \times \mathbb{R}^d$ ,

$$P = \partial_t - \partial_j g^{jk}(t, x) \partial_k$$

We denote by  $\sigma$  and  $\xi$  the time, respectively the space Fourier variable. Then the symbol of  $P$  is

$$p(t, x, \sigma, \xi) = -i\sigma + g^{jk}\xi_j\xi_k$$

This vanishes only at  $\sigma = 0$ ,  $\xi = 0$  so we should treat  $P$  as an elliptic operator. However, the results in Theorem 10 cannot be applied directly due to the different scaling associated to the heat operator. Instead one needs to adapt that setup to the current problem. This is discussed in what follows.

First we note that a time derivative is roughly equivalent with two time derivatives. Hence the size of the frequency is now  $(|\sigma|^2 + |\xi|^4)^{\frac{1}{4}}$  and the dyadic regions in frequency correspond to  $(|\sigma|^2 + |\xi|^4)^{\frac{1}{4}} \approx 2^j$ . The weighted Sobolev spaces are redefined accordingly,

$$\|u\|_{H_{\tau}^k} = \|(|D_t|^2 + |D_x|^4 + \tau^4)^{\frac{k}{4}} \hat{u}\|_{L^2}$$

and similarly for  $H_{\tau}^{k,p}$ .

The principal symbol of the conjugated operator  $P_{\phi}$  now has the form

$$\begin{aligned} p_{\phi}(t, x, \sigma, \xi, \tau) &= p(t, x, \sigma, \xi + i\tau\nabla\phi) \\ &= g^{jk}\xi_j\xi_k - \tau^2 g^{jk} \partial_j \phi \partial_k \phi + i(\sigma + 2\tau g^{jk} \xi_j \partial_k \phi) \end{aligned}$$

Observe that the time derivatives of  $\phi$  do not appear in this formula, as the terms containing them are lower order terms. By the same token,

time derivatives are also excluded from the definition of the Poisson bracket, namely

$$\{\Re p_\phi, \Im p_\phi\} = (\Re p_\phi)_x (\Im p_\phi)_\xi - (\Re p_\phi)_\xi (\Im p_\phi)_x$$

Another adjustment one needs to make concerns the scale of the localization in the  $L^2$  Carleman estimates, which now can be done on parabolic balls of size  $\tau^{-1} \times (\tau^{-\frac{1}{2}})^d$ . Because of this less time regularity for the coefficients is needed.

Taking all these considerations into account the analysis proceeds very much like in the case of Theorem 10. The operator  $P_\phi$  is elliptic at all frequencies larger than  $\tau$ , so the analysis must concentrate on the frequency region

$$|\sigma| + |\xi|^2 \lesssim \tau^2$$

A short computation shows that any surface  $\Gamma$  which is not tangent to the time slices is strongly pseudoconvex with respect to  $P$ . If  $\phi$  is a strongly pseudoconvex function with respect to  $P$  with  $\nabla_x \phi \neq 0$  then for fixed  $t$  and  $x$  the characteristic set of  $\Re p_\phi$  is a cylinder on top of an ellipsoid, while the characteristic set of  $\Im p_\phi$  is an oblique plane. Then the characteristic set of the full operator is an  $d-1$  dimensional ellipsoid which is the intersection of the cylinder with the plane, and has  $d-1$  nonvanishing curvatures. Therefore, we use (an adapted version of) Theorem 10 in  $d+1$  dimensions with  $k=0$  and  $n=d+1$ .

**Theorem 16.** *Let  $P$  be a second order parabolic operator whose coefficients are  $C^1$  in  $x$  and  $C^{\frac{1}{2}}$  in time. Let  $\phi$  be a strongly pseudoconvex function with respect to  $P$  with  $\nabla_x \phi \neq 0$ . Then for compactly supported  $u$  we have*

$$\|e^{\tau\phi} u\|_{H_\tau^{\frac{1}{d+3}, \frac{2(d+3)}{d+1}} \cap \tau^{\frac{1}{4}} H_\tau^1} \lesssim \|e^{\tau\phi} P u\|_{H_\tau^{-\frac{1}{d+3}, \frac{2(d+3)}{d+5}} + \tau^{-\frac{1}{4}} H_\tau^{-1}}, \quad \tau > \tau_0$$

Applied to unique continuation problems, this yields

**Theorem 17.** *Let  $P$  be a second order parabolic operator whose coefficients are  $C^1$  in  $x$  and  $C^{\frac{1}{2}}$  in time. Let  $\Gamma$  be a smooth surface. Then unique continuation for  $P+V$  across  $\Gamma$  holds for all potentials  $V$  which have the multiplicative mapping property*

$$V : H^{\frac{1}{d+3}, \frac{2(d+3)}{d+1}} \rightarrow H^{-\frac{1}{d+3}, \frac{2(d+3)}{d+5}}$$

This includes the case  $V \in L^{\frac{d+2}{2}}$ , for which the result is new. This can be relaxed to a slightly larger Morrey space. Applications of these ideas to strong unique continuation problems are contained in a forthcoming paper of the authors.

**Remark 8.8.** By using the mixed norm estimates in part (b) of Theorem 3 we can also obtain versions of these results with potentials in mixed norm spaces. For simplicity we state a weaker form of the Carleman estimates,

$$\|e^{\tau\phi}u\|_{L^q L^r \cap \tau^{\frac{1}{4}} H_\tau^1} \lesssim \|e^{\tau\phi}Pu\|_{L^{q'} L^{r'} + \tau^{-\frac{1}{4}} H_\tau^{-1}}, \quad \tau > \tau_0$$

which holds whenever<sup>4</sup>

$$\frac{2}{q} + \frac{d}{r} = \frac{d}{2}, \quad 2 \leq q \leq \infty$$

This yields the unique continuation result for  $P+V$  provided that  $V \in L_t^{\tilde{q}} L_x^{\tilde{r}}$  where the exponents  $\tilde{q}$  and  $\tilde{r}$  satisfy the scaling relation<sup>5</sup>

$$\frac{2}{\tilde{q}} + \frac{d}{\tilde{r}} = 2, \quad 1 \leq \tilde{q}, \tilde{r} \leq \infty$$

**8.4. The Schrödinger equation.** Here we consider the second order Schrödinger operator in  $\mathbb{R} \times \mathbb{R}^d$ ,

$$P = i\partial_t - \partial_j g^{jk}(x) \partial_k$$

For this one needs to use the same setup as in the case of parabolic equations.

We say that a surface  $\Gamma$  is noncharacteristic if it is not tangent to the time slices. As in the case of the wave equation, the condition (99) is always satisfied, but the condition (100) may or may not hold.

If  $\phi$  is a strongly pseudoconvex function with respect to  $P$  with  $\nabla_x \phi \neq 0$  then

$$p_\phi(t, x, \sigma, \xi, \tau) = \sigma - g^{jk} \xi_j \xi_k - \tau^2 g^{jk} \partial_j \phi \partial_k \phi + 2i\tau g^{jk} \xi_j \partial_k \phi$$

The characteristic set of  $p$  is a paraboloid which has  $d$  nonvanishing curvatures. The characteristic set of the real part of  $p_\phi$  is a paraboloid, which has  $d$  nonvanishing curvatures. The characteristic set of the full operator  $p_\phi$  is an  $d-1$  dimensional ellipsoid which is the intersection of the paraboloid with a vertical plane, and has  $d-1$  nonvanishing curvatures. Then (a variant of) Theorem 11 gives

**Theorem 18.** Let  $P$  be a second order Schrödinger operator whose coefficients are  $C^2$  in  $x$  and  $C^1$  in time. Let  $\phi$  be a strongly pseudoconvex function with respect to  $P$  with  $\nabla_x \phi \neq 0$ . Then for compactly supported  $u$  we have

$$\|e^{\tau\phi}u\|_{L^{\frac{2(d+2)}{d}} \cap \tau^{-\frac{1}{4}} H_\tau^{\frac{1}{2}}} \lesssim \|e^{\tau\phi}Pu\|_{L^{\frac{2(d+2)}{d+4}} + \tau^{\frac{1}{4}} H_\tau^{-\frac{1}{2}}}, \quad \tau > \tau_0$$

<sup>4</sup>The exponent  $q = 2$  is now allowed because the pair  $(q, r)$  is no longer the endpoint.

<sup>5</sup>For  $d = 1$  one needs the additional restriction  $\tilde{q} \geq 2$

Applied to unique continuation problems, this yields

**Theorem 19.** *Let  $P$  be a second order Schrödinger operator whose coefficients are  $C^2$  in  $x$  and  $C^1$  in time. Let  $\Gamma$  be a smooth surface. Then unique continuation for  $P + V$  across  $\Gamma$  holds for all potentials  $V \in L^{\frac{d+2}{2}}$ .*

One can also produce versions of this result involving mixed norm spaces.

**Remark 8.9.** *Using the mixed norm version of Theorems 3, 4 one obtains the Carleman estimates*

$$\|e^{\tau\phi}u\|_{L^q L^r \cap \tau^{\frac{1}{4}} H_\tau^1} \lesssim \|e^{\tau\phi}Pu\|_{L^{q'} L^{r'} + \tau^{-\frac{1}{4}} H_\tau^{-1}}, \quad \tau > \tau_0$$

which holds whenever<sup>6</sup>

$$\frac{2}{q} + \frac{d}{r} = \frac{d}{2}, \quad 2 < q \leq \infty$$

This yields the unique continuation result for  $P + V$  provided that  $V \in L_t^{\tilde{q}} L_x^{\tilde{r}}$  provided the exponents  $\tilde{q}$  and  $\tilde{r}$  satisfy the scaling relation<sup>7</sup>

$$\frac{2}{\tilde{q}} + \frac{d}{\tilde{r}} = 2, \quad 1 < \tilde{q}, \tilde{r} \leq \infty$$

## REFERENCES

- [1] Nicolas Burq, Patrick Gérard, and Nikolay Tzvetkov. The Schrödinger equation on a compact manifold: Strichartz estimates and applications. In *Journées “Équations aux Dérivées Partielles” (Plestin-les-Grèves, 2001)*, pages Exp. No. V, 18. Univ. Nantes, Nantes, 2001.
- [2] David Dos Santos Ferreira. Inegalites de carleman  $L^p$  pour des indices critiques. Thesis, Universite de Rennes I.
- [3] Luis Escauriaza. Carleman inequalities and the heat operator. *Duke Math. J.*, 104(1):113–127, 2000.
- [4] J. Ginibre and G. Velo. Generalized Strichartz inequalities for the wave equation. *J. Funct. Anal.*, 133(1):50–68, 1995.
- [5] Lars Hörmander. Differential operators of principal type. *Math. Ann.*, 140:124–146, 1960.
- [6] Lars Hörmander. *The analysis of linear partial differential operators. III*, volume 274 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1985. Pseudo-differential operators.
- [7] Lars Hörmander. *The analysis of linear partial differential operators. IV*. Springer-Verlag, Berlin, 1985. Fourier integral operators.
- [8] Victor Isakov. Carleman type estimates in an anisotropic case and applications. *J. Differential Equations*, 105(2):217–238, 1993.

<sup>6</sup>It would be interesting to see if the exponent  $q = 2$  is admissible.

<sup>7</sup>For  $d = 1$  one needs the additional restriction  $\tilde{q} \geq 2$

- [9] David Jerison and Carlos E. Kenig. Unique continuation and absence of positive eigenvalues for Schrödinger operators. *Ann. of Math. (2)*, 121(3):463–494, 1985. With an appendix by E. M. Stein.
- [10] L.V. Kapitanskij. Norm estimates in Besov and Lizorkin-Triebel spaces for the solutions of second-order linear hyperbolic equations. *J. Sov. Math.*, 56(2):2348–2389, 1991.
- [11] Markus Keel and Terence Tao. Endpoint Strichartz estimates. *Amer. J. Math.*, 120(5):955–980, 1998.
- [12] C. E. Kenig, A. Ruiz, and C. D. Sogge. Uniform Sobolev inequalities and unique continuation for second order constant coefficient differential operators. *Duke Math. J.*, 55(2):329–347, 1987.
- [13] Herbert Koch and Daniel Tataru. Carleman estimates and unique continuation for second-order elliptic equations with nonsmooth coefficients. *Comm. Pure Appl. Math.*, 54(3):339–360, 2001.
- [14] Gerd Mockenhaupt, Andreas Seeger, and Christopher D. Sogge. Local smoothing of Fourier integral operators and Carleson-Sjölin estimates. *J. Amer. Math. Soc.*, 6(1):65–130, 1993.
- [15] Hart F. Smith. A parametrix construction for wave equations with  $C^{1,1}$  coefficients. *Ann. Inst. Fourier (Grenoble)*, 48(3):797–835, 1998.
- [16] C. D. Sogge. A unique continuation theorem for second order parabolic differential operators. *Ark. Mat.*, 28(1):159–182, 1990.
- [17] Christopher D. Sogge. Strong uniqueness theorems for second order elliptic differential equations. *Amer. J. Math.*, 112(6):943–984, 1990.
- [18] Christopher D. Sogge. *Fourier integrals in classical analysis*, volume 105 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1993.
- [19] Gigliola Staffilani and Daniel Tataru. Strichartz estimates for a Schrödinger operator with nonsmooth coefficients. *Comm. Partial Differential Equations*, 27(7-8):1337–1372, 2002.
- [20] Daniel Tataru. The  $X_\theta^s$  spaces and unique continuation for solutions to the semilinear wave equation. *Comm. Partial Differential Equations*, 21(5-6):841–887, 1996.
- [21] Daniel Tataru. Strichartz estimates for second order hyperbolic operators with nonsmooth coefficients. II. *Amer. J. Math.*, 123(3):385–423, 2001.
- [22] Daniel Tataru. On the Fefferman-Phong inequality and related problems. *Comm. Partial Differential Equations*, 27(11-12):2101–2138, 2002.
- [23] Daniel Tataru. Strichartz estimates for second order hyperbolic operators with nonsmooth coefficients. III. *J. Amer. Math. Soc.*, 15(2):419–442 (electronic), 2002.
- [24] T.H. Wolff. A property of measures in  $\mathbf{R}^N$  and an application to unique continuation. *Geom. Funct. Anal.*, 2(2):225–284, 1992.
- [25] Thomas H. Wolff. Recent work on sharp estimates in second order elliptic unique continuation problems. In *Fourier analysis and partial differential equations (Miraflores de la Sierra, 1992)*, pages 99–128. CRC, Boca Raton, FL, 1995.