

Endpoint bilinear restriction theorems for the cone, and some sharp null form estimates

Terence Tao

Department of Mathematics, UCLA, Los Angeles, CA 90024, USA
(e-mail: tao@math.ucla.edu)

Received: 24 September 1999 / in final form: 25 April 2000 /
Published online: 4 May 2001 – © Springer-Verlag 2001

Abstract. Recently Wolff [25] obtained a nearly sharp L^2 bilinear restriction theorem for bounded subsets of the cone in general dimension. We obtain the endpoint of Wolff's estimate and generalize to the case when one of the subsets is large. As a consequence, we are able to deduce some nearly-sharp L^p null form estimates.

Mathematics Subject Classification (1991): 42B15, 35L05

1. Introduction

Let $n \geq 2$ be a fixed integer. We say that a function $\phi : \mathbf{R}^{n+1} \rightarrow H$ is a *red wave* if it takes values in a finite dimensional complex Hilbert space H , and its space-time Fourier transform $\hat{\phi}$ is an L^2 measure on the set

$$2^k \Sigma^{red} := \{(\xi, |\xi|) : \angle(\xi, e_1) \leq \frac{\pi}{8}, 2^k \leq |\xi| \leq 2^{k+1}\}$$

for some integer k , where e_1 is a fixed basis vector. Similarly, we say that $\psi : \mathbf{R}^{n+1} \rightarrow H'$ is a *blue wave* if it takes values in a finite dimensional complex Hilbert space H' and $\hat{\psi}$ is an L^2 measure on

$$2^k \Sigma^{blue} := \{(\xi, -|\xi|) : \angle(\xi, e_1) \leq \frac{\pi}{8}, 2^k \leq |\xi| \leq 2^{k+1}\}$$

for some integer k . In both cases we call 2^k the *frequency* of the waves ϕ , ψ .

Red and blue waves both solve the free wave equation, but propagate along different sets of characteristics. Note that blue waves are the time reversal of red waves. Also, these waves are automatically smooth and bounded

thanks to the frequency localization. The vector valued formulation will be convenient for inductive reasons, but our implicit constants shall always be independent of H and H' .

We define the *energy* of ϕ, ψ by

$$E(\phi) := \|\phi(t)\|_2^2, \quad E(\psi) := \|\psi(t)\|_2^2 \tag{1}$$

where $t \in \mathbf{R}$ is arbitrary. This definition is independent of the choice of t , and is related to the standard notion of energy by the formula

$$2^{2k} E(\phi) \sim \frac{1}{2} \int_{\mathbf{R}^n} |\phi_t(0)|^2 + |\nabla \phi(0)|^2 dx.$$

Throughout the paper, $p_0 = p_0(n)$ will denote the exponent

$$p_0 := \frac{n + 3}{n + 1}. \tag{2}$$

The main result of this paper is the following bilinear estimate.

Theorem 1.1. *Let ϕ be a red wave of frequency 1, and ψ be a blue wave of frequency 2^k for some $k \geq 0$. Then we have*

$$\|\phi\psi\|_p \lesssim 2^{k(\frac{1}{p} - \frac{1}{2} + \varepsilon)} E(\phi)^{1/2} E(\psi)^{1/2} \tag{3}$$

for all $2 \geq p \geq p_0, \varepsilon > 0$. In particular, ϕ, ψ have frequency 1 then

$$\|\phi\psi\|_p \lesssim E(\phi)^{1/2} E(\psi)^{1/2}. \tag{4}$$

In the above theorem and in the sequel, the implicit constants may depend on ε but are independent of H, H' , and $\phi\psi : \mathbf{R}^{n+1} \rightarrow H \otimes H'$ denotes the tensor product $\phi \otimes \psi$ of ϕ and ψ .

The estimate (4) solves a conjecture of Machedon and Klainerman. The restriction $p \geq p_0$ is sharp; see e.g. [7], [22], [21]. For $2 \leq p \leq \infty$ the theory is much simpler, and the best possible estimate is

$$\|\phi\psi\|_p \lesssim 2^{kn(\frac{1}{2} - \frac{1}{p})} E(\phi)^{1/2} E(\psi)^{1/2}.$$

This is easily proved from the $p = 2$ case and Sobolev embedding.

The estimate (4) is a genuinely bilinear estimate and cannot be proven directly from linear estimates. Indeed, the Strichartz estimate [19] combined with the Hölder inequality only yields the range $p \geq (n + 1)/(n - 1)$, while Plancherel’s theorem and Cauchy-Schwarz only gives the range $p \geq 2$ (see e.g. [11]). In the $n = 2$ case, the fact that one could go below $p = 2$ was first shown by Bourgain [2], and in [22] a concrete range of p was given, namely $p > 2 - \frac{8}{121}$. More recently, Wolff [25] obtained the range $p > p_0$ for all dimensions n . Thus (4) is the endpoint of that in [25].

The generalization (3) of (4) is necessary in order to develop sharp null form estimates, as we shall see in Sect. 17. The estimate (3) is sharp except for the ε . To see this, let ψ be a blue wave whose Fourier transform is supported in a unit “square” in $2^k \Sigma_{blue}$, and which is comparable to 1 on a $2^k \times 1$ tube oriented in a blue null direction, and let $\phi = (\phi_i)_{i=1}^{2^k}$ be a vector-valued red wave of frequency 1 such that ϕ_i is comparable to 1 on B_i , where the B_i are a family of 2^k unit balls that cover the above tube. The ε term can probably be removed when $p > p_0$. When $p = p_0$ the author conjectures that ε can still be removed, but this seems to require an extremely delicate analysis and some new Keakeya estimates for null rays. However, it should be fairly straightforward to replace the $2^{\varepsilon k}$ factor by a polynomial k^C in this case.

Broadly, our strategy to prove Theorem 1.1 is as follows. We shall localize the estimate (3) to a cube Q of side-length $R \gg 2^k$, and obtain a bound independent of R . This will be obtained by induction on R , as follows.

If ϕ and ψ are dispersed fairly evenly throughout Q , we shall decompose Q into sub-cubes of side-length $2^{-C_0} R$ and decompose ϕ and ψ into smaller waves, each of which is concentrated on one of these sub-cubes. By an interpolation between bilinear L^1 and L^2 estimates as in Wolff [25], the cross-terms are well controlled, and one can replace ϕ and ψ by a “quilt” of waves on the $2^{-C_0} R$ -cubes. One then applies the induction hypothesis to each sub-cube and sums up.

This tactic works well when ϕ and ψ are dispersed, but there is a problem when almost all the energy of ϕ and ψ simultaneously concentrate in a disk D of radius $r \ll R$. By Huygens’ principle the wave $\phi\psi$ is then concentrated in the double light cone generated by D . Restricting to this smaller set, we can exploit a more favourable bilinear L^1 estimate (Corollary 13.2) than the more trivial bilinear L^1 estimate (27) used in the non-concentrated case. One can then repeat the non-concentrated argument, and localize ϕ and ψ all the way down to the scale of r , at which point the waves become non-concentrated and one obtains enough of a gain to close the induction.

The proof is unfortunately rather complex. In an attempt to give the reader a sense of the full argument without drowning in the details, the author has abstracted the argument into several sections. We first give the top-level argument, in which Theorem 1.1 is deduced from several major propositions. Then, we give the medium-level argument, in which these major propositions are deduced from some elementary estimates and a key estimate, Proposition 4.1. Finally, we devote several sections to the proof of Proposition 4.1. We have tried to make each section as self-contained as possible, so that the arguments in each section rely only on the Propositions and Lemmata of the previous sections, and not on the method of arguments or on notation specific to a single section.

The author wishes to thank Damiano Foschi, Sergiu Klainerman, Christoph Thiele, and Tom Wolff for many helpful discussions, and for making their preprints readily available. This work was conducted at UCLA and at UNSW, with additional support from NSF grant DMS-9706764 and grants from the Sloan and Packard foundations. The author is also indebted to the referee for finding some errors in the first version of this paper (which claimed (3) without the $2^{\varepsilon k}$ factor).

2. Some notation

In this section we list some notation which will be used throughout the argument.

We fix $N \gg 1$ to be a large integer depending only on n ($N = 2^{n^{10}}$ will suffice); the disclaimer “assuming N is sufficiently large” will be implicit throughout our arguments. We also let $0 < \varepsilon \ll 1$ be an arbitrary small number. We also let C_0 denote an integer much larger than N , $1/\varepsilon$; for instance,

$$C_0 := 2^{\lfloor N/\varepsilon \rfloor^{10}} \quad (5)$$

will suffice. Generally speaking, we use N as a large exponent, and C_0 as a very large constant (large enough to dominate any reasonable quantity arising from N or ε). As a rule of thumb, any term containing an r^{-N} or R^{-N} factor may be ignored for all practical purposes; these error terms only arise because one cannot quite simultaneously localize in both space and frequency.

We use the notation $A := B$ to indicate that A is being defined to equal B . We shall extend this notation in several ways, for instance $\hat{f} := g$ indicates that f is being defined via the Fourier transform.

We let C denote various large numbers that vary from line to line (possibly depending on N , ε , but will not depend explicitly on C_0), and let $A \lesssim B$ or $A = O(B)$ denote the estimate $A \leq CB$ where C depends only on n and ε . Similarly we use $A \ll B$ to denote $A \leq C^{-1}B$. Generally speaking, we shall use the \lesssim notation to control error terms, but will need the more precise \leq notation for the main terms due to the inductive nature of the argument. In particular, we will be dealing with many estimates of the form $A \leq B + CE$, $A \leq (1 + Cc)B + c^{-C}E$, or $A \leq (1 + CNc)B + c^{-C}E$, where B is the main term, E is the error term, and $0 < c \ll 1$ is some small parameter which we may optimize in later. For such estimates it will be very important that the factor in front of B is either 1 or very close to 1, as we will be unable to close the induction otherwise.

Note that our constants C are independent of the dimension of the spaces H , H' . This will be important for the induction argument.

If $\phi(x, t)$ is a function of both space and time, we use $\phi(t)$ to denote the spatial function $\phi(t)(x) := \phi(x, t)$.

We use the hat notation for the spatial Fourier transform

$$\hat{f}(\xi) := \int_{\mathbf{R}^n} e^{-2\pi i x \cdot \xi} f(x) dx$$

as well as the spacetime Fourier transform

$$\hat{\phi}(\xi, \tau) := \int_{\mathbf{R}^{n+1}} e^{-2\pi i(x \cdot \xi + t\tau)} \phi(x, t) dx dt,$$

with the meaning being clear from context. We define the *frequency support* $\text{supp}(\hat{f})$ of a function f to be the support of the Fourier transform \hat{f} .

We shall always treat \mathbf{R}^{n+1} as endowed with the Euclidean metric and never with the Minkowski metric, so that terms such as length $|x|$, angle $\angle(x, y)$, etc. retain their usual meaning in \mathbf{R}^{n+1} . On the other hand, we will employ the Lorentz transforms on occasion, especially when we derive null form estimates in Sect. 17.

We now define some familiar geometric objects, namely disks, cubes, cones, and conic neighbourhoods.

A *disk* will be any subset D of \mathbf{R}^{n+1} of the form

$$D = D(x_D, t_D; r_D) := \{(x, t_D) : |x - x_D| \leq r_D\}$$

for some $(x_D, t_D) \in \mathbf{R}^{n+1}$ and $r_D > 0$. The reader should note that disks are n -dimensional objects even though they reside in \mathbf{R}^{n+1} . We call t_D the *time co-ordinate* of the disk D . If D is a disk, we define the cutoff function $\tilde{\chi}_D$ on \mathbf{R}^{n+1} by

$$\tilde{\chi}_D(x, t) := \left(1 + \frac{|x - x_D|}{r_D}\right)^{-N^{10}}. \tag{6}$$

If $D = D(x_D, t_D; r_D)$ is a disk and $c > 0$, we define cD to be the disk $cD = D(x_D, t_D; cr_D)$, and the *disk exterior* $D^{ext} = D^{ext}(x_D, t_D; r_D)$ to be the region

$$D^{ext}(x_D, t_D; r_D) := \{(x, t_D) : |x - x_D| > r_D\}.$$

We endow disks and disk exteriors with spatial Lebesgue measure dx .

We define $Q(x_Q, t_Q; r_Q)$ to be the $n + 1$ -dimensional cube in \mathbf{R}^{n+1} centered at (x_Q, t_Q) with side-length r_Q and with sides parallel to the axes. We call the interval $[t_Q - r_Q/2, t_Q + r_Q/2]$ the *lifespan* of Q . If $Q = Q(x_Q, t_Q; r_Q)$ is a cube and $c > 0$, we use cQ to denote the cube $cQ := Q(x_Q, t_Q; cr_Q)$. Finally, we define the cubical annuli $Q^{ann}(x_Q, t_Q; r_1, r_2)$ by

$$Q^{ann}(x_Q, t_Q; r_1, r_2) := Q(x_Q, t_Q; r_2) \setminus Q(x_Q, t_Q; r_1).$$

Let $\underline{\Sigma} = \underline{\Sigma}_n$ denote the spatial region

$$\underline{\Sigma} := \left\{ \xi : \frac{1}{2} \leq |\xi| \leq 4, \angle(\xi, e_1) \leq \frac{\pi}{4} \right\}.$$

If $(x_0, t_0) \in \mathbf{R}^{n+1}$, define the *red cone* with vertex (x_0, t_0) to be the set

$$C^{red}(x_0, t_0) := \{(x_0 + r\omega, t_0 - r) : r \in \mathbf{R}, \omega \in S^{n-1} \cap \underline{\Sigma}\},$$

and the *blue cone* at this vertex to be

$$C^{blue}(x_0, t_0) := \{(x_0 + r\omega, t_0 + r) : r \in \mathbf{R}, \omega \in S^{n-1} \cap \underline{\Sigma}\}.$$

For any $r \gg 1$, we define $C^{red}(x_0, t_0; r)$, $C^{blue}(x_0, t_0; r)$ to be the r -neighbourhoods of $C^{red}(x_0, t_0)$, $C^{blue}(x_0, t_0)$ respectively. Finally, we define the combined conic neighbourhood $C^{purple}(x_0, t_0; r)$ to be

$$C^{purple}(x_0, t_0; r) := C^{red}(x_0, t_0; r) \cup C^{blue}(x_0, t_0; r).$$

For any integer j , let \mathbf{D}_j denote the dilation operator

$$\mathbf{D}_j \phi(x, t) := \phi(2^j x, 2^j t),$$

and \mathbf{T} to be the time reversal operator

$$\mathbf{T} \phi(x, t) := \phi(x, -t).$$

The operator \mathbf{T} maps red waves onto blue waves and vice versa, while the operator \mathbf{D}_j maps waves of frequency 1 onto waves of frequency 2^j .

In our induction argument we shall frequently be decomposing red and blue waves into smaller waves. Unfortunately, these decompositions often enlarge the Fourier support of the waves slightly, which is a potential obstruction to closing the induction. To get around this we introduce the notion of the *margin* $\text{margin}(\phi)$ of a wave ϕ . More precisely, if ϕ is a red wave of frequency 1, we define $\text{margin}(\phi)$ to be the non-negative real number

$$\text{margin}(\phi) := \text{dist}(\text{supp}(\hat{\phi}), \Sigma^+ \setminus \Sigma^{red})$$

where Σ^\pm denotes the light cones

$$\Sigma^\pm := \{(\xi, \pm|\xi|) : \xi \in \mathbf{R}^n\}.$$

We extend this definition to general red or blue waves by the formulae

$$\text{margin}(\mathbf{T}\mathbf{D}_k \phi) := \text{margin}(\mathbf{D}_k \phi) := \text{margin}(\phi).$$

Thus, for instance, if ψ is a blue wave of frequency 2^k , then

$$\text{margin}(\psi) := 2^{-k} \text{dist}(\text{supp}(\hat{\psi}), 2^k(\Sigma^- \setminus \Sigma^{blue})).$$

The concept of margin is only needed to overcome the technical obstruction mentioned earlier, and otherwise plays no role of importance.

3. The top-level proof of Theorem 1.1

In this section we give the top-level proof of Theorem 1.1, in which we state the key propositions and give the inductive argument which yields the Theorem from these propositions. We shall prove these propositions in later sections.

Throughout the proof $n \geq 2$, $k \geq 0$, $\varepsilon > 0$ will be fixed. It suffices to verify the case $k \geq 2^{C_0^{10}}$, since the $k < 2^{C_0^{10}}$ case follows by applying a Lorentz transform. We shall thus assume $k \geq 2^{C_0^{10}}$ throughout the argument. In particular, any factor depending on k will dwarf any factor depending on C_0 , which in turn dwarfs any factor depending on N , which in turn dwarfs any absolute constants.

We can also assume that k is a multiple of C_0 . This is a purely notational convenience that we shall only use in Sect. 16.

It is known (see e.g. [2], [11], [22], [16], [7], or Lemma 14.2 below) that

$$\|\phi\psi\|_2 \lesssim E(\phi)^{1/2}E(\psi)^{1/2}. \tag{7}$$

Thus Theorem 1.1 holds for $p = 2$. By interpolation it thus suffices to prove the theorem when

$$p = p_0 = \frac{n + 3}{n + 1}, \tag{8}$$

and we shall implicitly assume this throughout the proof.

Definition 3.1. For any $R \geq C_0 2^k$, we define $A(R)$ to be the best constant for which the inequality

$$\|\phi\psi\|_{L^p(Q_R)} \leq A(R)E(\phi)^{1/2}E(\psi)^{1/2} \tag{9}$$

holds for all spacetime cubes Q_R of side-length R , red waves ϕ of frequency 1, and blue waves ψ of frequency 2^k such that one has the margin requirement

$$\text{margin}(\phi), \text{margin}(\psi) \geq 1/100 - (2^k/R)^{1/N}. \tag{10}$$

The waves ϕ, ψ may take values in arbitrary finite-dimensional Hilbert spaces.

Note that it suffices to verify (9) for those ϕ, ψ which obey the normalization

$$E(\phi) = E(\psi) = 1. \tag{11}$$

From (7) it is clear that $A(R)$ is finite for each R . The margin requirements on $A(R)$ are technicalities which are needed for the induction on R to work properly, as many of our decompositions will decrease the margin of ϕ and ψ slightly. However, we may remove the margin requirements by

a finite partition of space and frequency, and some mild Lorentz transforms to obtain the crude estimate

$$\|\phi\psi\|_{Q_R} \lesssim A(R')E(\phi)^{1/2}E(\psi)^{1/2} \tag{12}$$

for any cube Q_R of side-length R , any $R' \sim R$, and any red and blue waves ϕ, ψ of frequency $1, 2^k$ respectively. In particular to prove (3) it suffices to show that

$$A(R) \lesssim 2^{CC_0}2^{\varepsilon k}2^{k(\frac{1}{p}-\frac{1}{2})} \tag{13}$$

uniformly for all $R \geq C_02^k$.

Because of the increasingly strict margin requirements as $R \rightarrow \infty$ we see that $A(R)$ is not necessarily increasing in R . We therefore define the auxiliary quantity $\bar{A}(R)$ for all $R \geq C_02^k$ by

$$\bar{A}(R) := \sup_{C_02^k \leq r \leq R} A(r).$$

We now begin the proof of (13). In Sect. 7 we shall prove (13) when R is close to 2^k :

Proposition 3.2. *For any $R \geq C_02^k$, we have the bound*

$$A(R) \lesssim 2^{CC_0}(R/2^k)^C2^{k(\frac{1}{p}-\frac{1}{2})}.$$

This proposition is needed to begin the inductive argument. The exact power of $(R/2^k)$ is unimportant as we shall soon improve this bound substantially.

In Sect. 8, we adapt the localization ideas from [25] to prove the following recursive inequality on $A(R)$.

Proposition 3.3. *Let $R \geq 2C_02^k$ and $0 < c \leq 2^{-C_0}$, and let ϕ, ψ be red and blue waves of frequency 1 and 2^k respectively which obey the relaxed margin requirement*

$$\text{margin}(\phi), \text{margin}(\psi) \geq 1/100 - 2(2^k/R)^{1/N}. \tag{14}$$

Then for any cube Q_R of side-length R we have

$$\|\phi\psi\|_{L^p(Q_R)} \leq [(1 + Cc)\bar{A}(R/2) + c^{-C}2^{k(\frac{1}{p}-\frac{1}{2})}]E(\phi)^{1/2}E(\psi)^{1/2}. \tag{15}$$

In particular, we have

$$A(R) \leq (1 + CNc)\bar{A}(R/2) + c^{-C}2^{k(\frac{1}{p}-\frac{1}{2})}.$$

The slight relaxation of margin requirements in Proposition 3.3 (as compared to (10)) is important for technical reasons, but should be ignored for a first reading.

The loss of CNc in the main term comes about because we need to divide an R -cube into sub-cubes at various scales, and then shrink those cubes further by about c in order to create some separation between those sub-cubes (this is necessary, otherwise the error terms can blow up). In practice we can always optimize c so that its impact on the argument is negligible.

By setting $c = (2^k/R)^{\frac{\varepsilon}{cN}}$ in Proposition 3.3, we obtain

$$A(R) \leq (1 + CN(2^k/R)^{\frac{\varepsilon}{cN}})\bar{A}(R/2) + C(2^k/R)^{\varepsilon/N}2^{k(\frac{1}{p}-\frac{1}{2})}$$

for all $R \geq 2C_02^k$. Iterating this and using Proposition 3.2 when $R \sim 2^k$, we obtain

$$A(R) \lesssim 2^{CC_0}(R/2^k)^{\varepsilon/N}2^{k(\frac{1}{p}-\frac{1}{2})} \tag{16}$$

for any $R \geq C_02^k$. This can then be used to obtain Theorem 1.1 for $p > p_0$ (cf. [25], [23], [2]), but we shall not do so here.

From the preceding discussion we observe that (13) is already proven for $R \leq 2^{C_0k}$. Thus we may assume that $R > 2^{C_0k}$. R is now very large, dominating most quantities which depend only on k, C_0, N , and ε .

In the large R case we need to introduce the notion of energy concentration.

Definition 3.4. For any $r > 0$, spacetime cube Q , red wave ϕ , and blue wave ψ , we define the *energy concentration* $E_{r,Q}(\phi, \psi)$ to be the quantity

$$E_{r,Q}(\phi, \psi) := \max \left(\frac{1}{2}E(\phi)^{1/2}E(\psi)^{1/2}, \sup_D \|\phi\|_{L^2(D)}\|\psi\|_{L^2(D)} \right)$$

where D ranges over all disks of radius r whose time co-ordinate t_D is contained in the lifespan of Q .

Clearly we have $E_{r,Q}(\phi, \psi) \sim E(\phi)^{1/2}E(\psi)^{1/2}$ for any r , with equality or near-equality only occurring when ϕ, ψ simultaneously concentrate in a disk of radius r . However, the choice of whether to use $E_{r,Q}(\phi, \psi)$ instead of $E(\phi)^{1/2}E(\psi)^{1/2}$ will be crucial to make a certain induction work.

Roughly speaking, the strategy for the $R \geq 2^{C_0k}$ case is to show that Proposition 3.3 can be improved slightly unless there is substantial energy concentration. This slight improvement will be enough to close the induction, however, one must still deal with the concentrated case. In this case we use Huygens principle to restrict ourselves to a small region of a double cone C^{purple} , in which case one can obtain an improvement of Proposition 3.3 by other means.

To make this strategy more precise we shall need the following technical variant of $A(R)$ which is weighted slightly to exploit the gain in the non-concentrated case.

Definition 3.5. For any $R \geq 2^{C_0k}$ and any $r', r > 0$, we define $A(R, r, r')$ to be the best constant for which the inequality

$$\begin{aligned} & \|\phi\psi\|_{L^p(Q_R \cap C^{purple}(x_0, t_0; r'))} \\ & \leq A(R, r, r')(E(\phi)^{1/2}E(\psi)^{1/2})^{1/p} E_{r, C_0Q_R}(\phi, \psi)^{1/p'} \end{aligned} \quad (17)$$

holds for all spacetime cubes Q_R of side-length R , all $(x_0, t_0) \in \mathbf{R}^{n+1}$, red waves ϕ of frequency 1, and blue waves ψ of frequency 2^k such that one has the margin requirement (10) and the energy bound (11).

As before, it suffices to verify (17) assuming the energy normalization (11). It is important that the right-hand side if (17) is exactly as stated, and not (for instance) the comparable quantity $A(R, r, r')E(\phi)^{1/2}E(\psi)^{1/2}$. In practice r and r' shall usually be comparable in size.

The preceding heuristics regarding concentration can be formalized in the following Proposition, which we prove in Sect. 11. This Proposition is basically an application of Huygens’ principle, combined with some more sophisticated arguments to deal with the highly concentrated case $r \lesssim R^{1/2+4/N}$.

Proposition 3.6. For any $R \geq 2^{C_0k}$, we have

$$A(R) \leq (1 - C_0^{-C}) \sup_{\substack{2^{C_0k} \leq \tilde{R} \leq R \\ \tilde{R}^{1/2+4/N} \leq r}} A(\tilde{R}, r, C_0(1+r)) + 2^{CC_0} 2^{\varepsilon k} 2^{k(\frac{1}{p} - \frac{1}{2})}.$$

The requirement $\tilde{R}^{1/2+4/N} \leq r$ is somewhat difficult to obtain, but it is necessary to do so because the tools we shall develop to control $A(R, r, r')$ have a spatial uncertainty of about \sqrt{R} and so one cannot effectively exploit any concentration effects near or below this scale. This uncertainty of \sqrt{R} is responsible for all the powers of $r^{1/N}$ and $R^{1/N}$ which appear in the arguments; these powers should be ignored for a first reading. The key point to observe in Proposition 3.6 is that we have somehow wrested a small gain $(1 - C_0^{-C})$ from the main term on the right-hand side, thanks to the beneficial effects of non-concentration in (17). This gain will allow us to absorb all error terms and $(1 + Cc)$ factors in the other Propositions in this section, thus closing the induction.

To use this inequality inductively we need to bound $A(R, r, r')$ in terms of $A(R)$. From (12) it is easy to show that $A(R, r, r') \leq CA(R)$, but this is too crude to close the induction, and one must take some more care with the constant in the leading term.

The bounding of $A(R, r, r')$ by $A(R)$ can be split into two stages. First, we shall use finite speed of propagation in Sect. 12 to observe that one can obtain the desired bound in the non-concentrated case $r \geq C_0^C R$:

Proposition 3.7. *For any $R \geq 2^{C_0 k}$, $r \geq C_0^C R$, $r' > 0$, and $0 < c \leq 2^{-C_0}$, we have*

$$A(R, r, r') \leq (1 + Cc)\overline{A(R)} + c^{-C}2^{k(\frac{1}{p}-\frac{1}{2})}.$$

For the concentrated case we iterate the following Proposition, which is proven in 9.

Proposition 3.8. *For any $R \geq 2^{C_0 k}$ and $C_0^C R \geq r > R^{1/2+3/N}$, we have*

$$A(R, r, r') \leq (1 + CNc)A(R/C_0, r(1 - Cr^{-1/3N}), r') + c^{-C} \left(1 + \frac{R}{2^k r}\right)^{-1/N} 2^{k(\frac{1}{p}-\frac{1}{2})}$$

for any $0 < c \leq 2^{-C_0}$.

The decay of $(1 + \frac{R}{2^k r})^{-1/N}$ in the error term is crucial for this endpoint result as it allows us to avoid losing a logarithmic factor $\log(R/r)$ in the iteration, which would otherwise be fatal to the proof of the endpoint. This decay ultimately arises from the improved energy estimates on cones as encoded in Lemma 13.1. However, the presence of the 2^k means that we still lose a factor of k when summing over scales R , and this is the main source of the $2^{\varepsilon k}$ loss in (3).

These two Propositions combine to give

Corollary 3.9. *For any $R \geq 2^{C_0 k}$ and $r \geq R^{1/2+4/N}$, we have*

$$A(R, r, C_0(1 + r)) \leq (1 + CNc)\overline{A(R)} + c^{-C}2^{\varepsilon k}2^{k(\frac{1}{p}-\frac{1}{2})}$$

for any $0 < c \leq 2^{-C_0}$.

Proof. We may assume that $r < C_0^C R$ since the claim follows from Proposition 3.7 otherwise. Let J be the first integer such that $r \geq 2^{-J}C_0^C R$; from the hypotheses we have $J \lesssim \log(r)$. Define $r =: r_0 > r_1 > \dots > r_J$ inductively by $r_{j+1} := r_j(1 - Cr_j^{1-1/3N})$. One can verify inductively that $r_j = r + O(jr^{-1/4N})$ for all j , and in particular that $r_j \sim r$.

From Proposition 3.8 we have

$$A(R/2^j, r_j, C_0(1 + r)) \leq (1 + CNc_j)A(R/2^{j+1}, r_{j+1}, C_0(1 + r)) + c_j^{-C} \left(1 + \frac{R}{2^j 2^k r}\right)^{-1/N} 2^{k(\frac{1}{p}-\frac{1}{2})}$$

for any $0 < c_j \leq 2^{-C_0}$.

Observe that

$$1 + \frac{R}{2^j 2^k r} \gtrsim C_0^{-C} 2^{(J-j-k)_+}.$$

If we thus set $c_j := k^{-1} c 2^{-(J-j-k)_+/(2CN)}$ for a suitable constant C , we thus obtain

$$\begin{aligned} &A(R/2^j, r_j, C_0(1+r)) \\ &\leq (1 + CNk^{-1} c 2^{-(J-j-k)_+/(2CN)}) A(R/2^{j+1}, r_{j+1}, C_0(1+r)) \\ &\quad + k^C c^{-C} 2^{-(J-j-k)_+/2N} 2^{k(\frac{1}{p}-\frac{1}{2})}. \end{aligned}$$

Iterating this, one obtains

$$A(R, r, C_0(1+r)) \leq (1 + CNc) A(R/2^J, r_J, C_0(1+r)) + k^C c^{-C} 2^{k(\frac{1}{p}-\frac{1}{2})}.$$

The claim then follows from Proposition 3.7. □

Combining this with (3.6) and setting $c := 2^{-C_0}$, we thus obtain

$$A(R) \leq (1 - C_0^{-C}) \overline{A(R)} + 2^{CC_0} 2^{\varepsilon k} 2^{k(\frac{1}{p}-\frac{1}{2})}$$

for all $R \geq 2^{C_0 k}$. Combining this with (16) we thus see that this inequality thus holds for all $R \geq C_0 2^k$. Taking suprema and using the monotonicity of $\overline{A(R)}$ we thus obtain

$$\overline{A(R)} \leq (1 - C_0^{-C}) \overline{A(R)} + 2^{CC_0} 2^{\varepsilon k} 2^{k(\frac{1}{p}-\frac{1}{2})}$$

for all $R \geq C_0 2^k$, which implies

$$\overline{A(R)} \leq 2^{CC_0} 2^{\varepsilon k} 2^{k(\frac{1}{p}-\frac{1}{2})},$$

and (13) follows for all $R \geq C_0 2^k$ as desired.

It remains to prove Propositions 3.2-3.8. It turns out that most of these propositions follow as consequences of a localization property for red and blue waves, Proposition 4.1, which we shall state in the next section, after some notation. In Sections 7-12 we show how this Proposition implies Propositions 3.2-3.8. Finally, in Sections 13-16 we give a proof of Proposition 4.1.

4. The main proposition

In this section we state the main proposition of the argument, Proposition 4.1. As we shall see in the next few sections, this Proposition will be the main tool used to prove Propositions 3.2, 3.3, and 3.8, and also plays a minor role in the proof of Proposition 3.7.

Proposition 4.1 involves the localization of waves ϕ, ψ on a large cube Q into smaller waves localized on sub-cubes of Q . To make this precise we must introduce some notation.

If Q is a cube of side-length R , and $j \geq 0$ is an integer, we may partition Q into $2^{(n+1)j}$ cubes of side-length $2^{-j}R$; we use $\mathcal{Q}_j(Q)$ to denote the collection of these cubes.

If Q is a cube and $j \geq 0$ is an integer, we define a *red wave table* ϕ on Q with depth j to be any red wave with the vector form

$$\phi =: (\phi^{(q)})_{q \in \mathcal{Q}_j(Q)},$$

where the components $\phi^{(q)}$ may themselves be vector-valued. If $0 \leq j' < j$, and $q' \in \mathcal{Q}_{j'}(Q)$, we define $\phi^{(q')}$ to be the red wave table on q' with depth $j - j'$ given by

$$(\phi^{(q')})^{(a)} := \phi^{(a)} \text{ for all } q \in \mathcal{Q}_{j-j'}(q').$$

Note that

$$E(\phi) = \sum_{q' \in \mathcal{Q}_{j'}(Q)} E(\phi^{(q')})$$

for all $0 \leq j' \leq j$.

If ϕ is a red wave table on Q of depth j and $0 \leq j' \leq j$, we define the j' -quilt $[\phi]_{j'}$ of ϕ to be the non-negative function

$$[\phi]_{j'} := \sum_{q \in \mathcal{Q}_{j'}(Q)} |\phi^{(q)}| \chi_q.$$

Note we have the pointwise estimates

$$|\phi^{(q)}| \chi_q \leq [\phi]_j \leq [\phi]_{j-1} \dots \leq [\phi]_0 = |\phi| \chi_Q \tag{18}$$

for all $q \in \mathcal{Q}_j(Q)$. We define blue wave tables and their quilts analogously.

The estimates (18) are of course very crude, and we shall frequently be exploiting various improvements to this estimate in the sequel.

If Q is a cube, $k \geq 0$ is an integer, and $0 \leq c \ll 1$, we define the (c, k) -interior $I^{c,k}(Q)$ of Q by

$$I^{c,k}(Q) := \bigcup_{q \in \mathcal{Q}_k(Q)} (1 - c)q. \tag{19}$$

The advantage of working with $I^{c,k}(Q)$ instead of Q is that the sub-cubes of $I^{c,k}(Q)$ have some significant separation properties.

With these notational preliminaries, we are now ready to state Proposition 4.1.

Proposition 4.1. *Let $R \geq C_0 2^k$, $0 < c \leq 2^{-C_0}$, and let ϕ, ψ be red and blue waves with frequency 1 and 2^k respectively, which obey the energy normalization (11) and the relaxed margin requirement (14). For any cube Q , define the set $X(Q) \subset Q$ by*

$$X(Q) := \bigcap_{j=C_0}^k I^{c 2^{-(k-j)/N}, j}(Q). \tag{20}$$

Then for any cube Q of side-length CR , we can find a red wave table Φ on Q with depth k and frequency 1, and a blue wave table Ψ on Q with depth C_0 and frequency 2^k , such that we have the margin estimates

$$\text{margin}(\Phi), \text{margin}(\Psi) \geq 1/100 - (2^{k+C_0}/R)^{1/N}, \tag{21}$$

we have the energy estimates

$$E(\Phi), E(\Psi) \leq 1 + Cc, \tag{22}$$

and we have the inequality

$$\|\phi\psi\|_{L^p(X(Q))} \leq \|[\Phi]_k[\Psi]_{C_0}\|_{L^p(X(Q))} + c^{-C} 2^{k(\frac{1}{p}-\frac{1}{2})}. \tag{23}$$

For any cone $C^{\text{purple}}(x_0, t_0; r)$ with $r > 1$, we may improve (23) to

$$\begin{aligned} \|\phi\psi\|_{L^p(X(Q) \cap C^{\text{purple}}(x_0, t_0; r))} &\leq \|[\Phi]_k[\Psi]_{C_0}\|_{L^p(X(Q) \cap C^{\text{purple}}(x_0, t_0; r))} \\ &\quad + c^{-C} 2^{k(\frac{1}{p}-\frac{1}{2})} \left(1 + \frac{R}{2^k r}\right)^{-1/N}. \end{aligned} \tag{24}$$

Furthermore, we have the persistence of non-concentration

$$E_{r(1-C_0 r^{-1/4N}), 2Q}(\Phi, \Psi) \leq E_{r, 2Q}(\phi, \psi) + Cc + c^{-C} R^{C-N/2} \tag{25}$$

for all $r \gtrsim R^{1/2+3/N}$.

The reason why we use $X(Q)$ instead of Q is that the sub-cubes of $X(Q)$ have some non-zero separation between them, which will let us avoid some unpleasantness in Sect. 15. To return from $X(Q)$ to Q we shall use an averaging lemma, Lemma 6.1; this causes the loss of Cc in many of the Propositions stated previously. The estimates (22), (25) are stating that Φ, Ψ are “smaller than or equal to” ϕ, ψ in energy norm, while (23), (24) state that the quilts $[\Phi]_k, [\Psi]_{C_0}$ are good approximations for ϕ, ψ respectively.

This replacement of waves by quilts of essentially equal or lesser energy allows one to induct efficiently (providing that one has constants close to 1 in the main terms).

Note that the margin requirements on ϕ, ψ are slightly weaker than those in Definition 3.5. The gain of $(1 + \frac{R}{2^k r})^{-1/N}$ in (24) over (23) is responsible for the corresponding gain in Proposition 3.8 as compared against similar estimates such as Proposition 3.3. This gain is essential as it needs to compensate both for the Cc loss in (22) and for the $\log(R)$ loss that would otherwise arise from an induction on scale.

Interestingly, the low-frequency wave ϕ can be localized much further than the high-frequency wave ψ ; the former can be localized to cubes of 2^{-k} the side-length, whereas the high-frequency waves can only be localized by about 2^{-C_0} before the error estimates begin to deteriorate logarithmically. This behaviour is also seen in the example given in the introduction demonstrating that the power of k in (3) is essentially sharp. In applications we shall only need to localize to 2^{-C_0} , however in the proof of Proposition 4.1 we shall need to localize ϕ all the way down to 2^{-k} before one can begin to localize ψ .

The proof of Proposition 4.1 is the longest part of the argument, and is basically a “pigeonhole-free” version of the arguments in Wolff [25]. Since it is the statement of this Proposition, rather than the method of proof, which are required to prove the Propositions of the previous section, we shall defer the proof of Proposition 4.1 to sections 14-16. For now, we devote ourselves to the question of how this Proposition, combined with some other more elementary tools, can be used to prove Propositions 3.2–3.8.

5. Energy estimates

In this section we record some fairly easy energy estimates which will be used throughout the paper.

Let ϕ, ψ be red and blue waves. By integrating (1) along the life-span of a cube we see that

$$\|\phi\|_{L^2(Q_R)} \lesssim R^{1/2} E(\phi)^{1/2}, \quad \|\psi\|_{L^2(Q_R)} \lesssim R^{1/2} E(\psi)^{1/2} \tag{26}$$

for any cube Q_R of side-length R . By Hölder’s inequality we thus have

$$\|\phi\psi\|_{L^1(Q_R)} \lesssim R E(\phi)^{1/2} E(\psi)^{1/2}. \tag{27}$$

To compensate for the loss of R in (27) we shall often seek variants of (7) in which one has a gain of $R^{-(n-1)/4}$. This gain exactly balances the loss of R when interpolated at the exponent $p = p_0$. One such example of this gain is Lemma 14.2; see also (54), (108).

To tackle the case $k \gg 1$ of widely differing frequencies we shall need the following improvement of (27) when ϕ is replaced by a quilt.

Lemma 5.1. *Let Q_R be a cube of side-length $R > 0$, $j \geq 0$ is an integer, ϕ be a red wave table on Q_R with depth at least j , and ψ be a blue wave. Then*

$$\|[\phi]_j\|_2 \lesssim 2^{-j/2} R^{1/2} E(\phi)^{1/2},$$

and

$$\|[\phi]_j \psi\|_1 \lesssim 2^{-j/2} R E(\phi)^{1/2} E(\psi)^{1/2}.$$

Proof. Let $q \in \mathcal{Q}_j(Q_R)$. From (26) for $\phi^{(q)}$ we have

$$\|\phi^{(q)}\|_{L^2(q)} \lesssim 2^{-j/2} R^{1/2} E(\phi^{(q)})^{1/2}.$$

Square summing this in q we obtain the first estimate, and the second estimate follows from (26) for ψ and Hölder’s inequality. \square

This improvement over (27) is one demonstration of the gain involved when passing from a wave to a quilt. The other major advantage of quilts is that they allow one to localize estimates on large cubes to estimates on small cubes so efficiently that one can make induction on scale arguments work.

6. An averaging lemma

The following averaging lemma will be needed in the proofs of Proposition 3.2, 3.3, 3.8. It allows one to average out a bounds on $X(Q)$ to obtain a bound on Q .

Lemma 6.1. *Let $R > 0$, $0 < c \leq 2^{-C_0}$, and F be an arbitrary smooth function. Let Q_R be a cube of side-length R . Then there exists a cube Q of side-length CR contained in C^2Q_R such that*

$$\|F\|_{L^p(Q_R)} \leq (1 + CNc)\|F\|_{L^p(X(Q))},$$

where $X(Q)$ is the set in (20).

Proof. By the pigeonhole principle it suffices to show that

$$\|F\|_{L^p(Q_R)}^p \leq \frac{1}{|Q_R|} \int_{Q_R} (1 + Cc)^p \|F\|_{L^p(Q_R \cap (X(Q(x_0, t_0; CR))))}^p dx_0 dt_0.$$

From the symmetry and translation covariance of $X(Q)$ we have the identity

$$\begin{aligned} & \int_{Q_R} \|F\|_{L^p(Q_R \cap (X(Q(x_0, t_0; CR))))}^p dx_0 dt_0 \\ &= \int_{Q_R} |F(x, t)|^p |X(Q(x, t; CR)) \cap Q_R| dx dt. \end{aligned}$$

On the other hand, from (19) we have

$$|Q(x_0, t_0; CR) \setminus I^{c,k}(Q(x_0, t_0; CR))| \lesssim c|Q(x_0, t_0; CR)|$$

so from (20) we have

$$|Q(x, t; CR) \setminus X(Q(x, t; CR))| \lesssim Nc|Q(x, t; CR)|$$

and thus that

$$|Q_R| \leq (1 + CNc)|X(Q(x, t; CR)) \cap Q_R|.$$

The claim then follows. \square

7. Proof of Proposition 3.2

We now have enough machinery to prove Propositions 3.2, 3.3, and 3.8. We begin with Proposition 3.2.

Fix $R \geq C_0 2^k$, and let ϕ, ψ be red and blue waves of frequency 1 and 2^k respectively, obeying the margin requirement (10) and the energy normalization (11). Let Q_R be a cube of side-length R . To prove Proposition 3.2 it suffices to show that

$$\|\phi\psi\|_{L^p(Q_R)} \lesssim 2^{CC_0} (R/2^k)^C 2^{k(\frac{1}{p}-\frac{1}{2})}.$$

Set $c := 2^{-C_0}$, so that $1 + CNc \sim 1$. From Lemma 6.1 with $F := \phi\psi$, we may find a cube Q of side-length CR such that

$$\|\phi\psi\|_{L^p(Q_R)} \lesssim \|\phi\psi\|_{L^p(X(Q))}.$$

Let Φ, Ψ be as in Proposition 4.1. By (23) and (18), we thus have

$$\|\phi\psi\|_{L^p(Q_R)} \lesssim \|[\Phi]_k \Psi\|_p + 2^{CC_0} 2^{k(\frac{1}{p}-\frac{1}{2})}.$$

By (18), (7) and (22) we have

$$\|[\Phi]_k \Psi\|_2 \leq \|\Phi\Psi\|_2 \lesssim E(\Phi)^{1/2} E(\Psi)^{1/2} \lesssim 1,$$

while from Lemma 5.1 and (22) we have

$$\|[\Phi]_k \Psi\|_1 \leq 2^{-k/2} R E(\Phi)^{1/2} E(\Psi)^{1/2} \lesssim 2^{-k/2} R = (R/2^k) 2^{k/2}.$$

Interpolating between the two, one obtains

$$\|[\Phi]_k \Psi\|_p \lesssim (R/2^k)^C 2^{k(\frac{1}{p}-\frac{1}{2})},$$

and the claim follows. \square

8. Proof of Proposition 3.3

Let ϕ, ψ, Q_R, c be as in Proposition 3.3. We may assume the normalization (10). We may assume that $R \geq 2^{CC_0}2^k$ for some large C , since the claim (15) follows from Proposition 3.2 otherwise.

From Lemma 6.1 with $F := \phi\psi$, we may find a cube Q of side-length CR such that

$$\|\phi\psi\|_{L^p(Q_R)} \leq (1 + CNc)\|\phi\psi\|_{L^p(X(Q))}.$$

Let Φ, Ψ be as in Proposition 4.1. By (23), we thus have

$$\|\phi\psi\|_{L^p(Q_R)} \leq (1 + CNc)\|[\Phi]_k[\Psi]_{C_0}\|_{L^p(X(Q))} + c^{-C}2^{k(\frac{1}{p}-\frac{1}{2})}.$$

By (18) and the hypothesis $k \gg C_0$, we thus have

$$\|\phi\psi\|_{L^p(Q_R)} \leq (1 + CNc)\|[\Phi]_{C_0}[\Psi]_{C_0}\|_{L^p(Q)} + c^{-C}2^{k(\frac{1}{p}-\frac{1}{2})}. \tag{28}$$

We expand

$$\|[\Phi]_{C_0}[\Psi]_{C_0}\|_{L^p(Q)} = \left(\sum_{q \in \mathcal{Q}_{C_0}(Q)} \|\Phi^{(q)}\Psi^{(q)}\|_{L^p(q)}^p \right)^{1/p}.$$

By (9) and (21), then the inclusion $l^p \subset l^1$, followed by Cauchy-Schwarz and then (22) we thus have

$$\begin{aligned} \|[\Phi]_{C_0}[\Psi]_{C_0}\|_{L^p(Q)} &\leq A(2^{-C_0}R) \left(\sum_{q \in \mathcal{Q}_{C_0}(Q)} E(\Phi^{(q)})^{p/2} E(\Psi^{(q)})^{p/2} \right)^{1/p} \\ &\leq \bar{A}(R/2) \sum_{q \in \mathcal{Q}_{C_0}(Q)} E(\Phi^{(q)})^{1/2} E(\Psi^{(q)})^{1/2} \\ &\leq \bar{A}(R/2) \left(\sum_{q \in \mathcal{Q}_{C_0}(Q)} E(\Phi^{(q)}) \right)^{1/2} \left(\sum_{q \in \mathcal{Q}_{C_0}(Q)} E(\Psi^{(q)}) \right)^{1/2} \\ &= \bar{A}(R/2) E(\Phi)^{1/2} E(\Psi)^{1/2} \\ &\leq (1 + Cc)\bar{A}(R/2). \end{aligned}$$

Inserting this back into (28) we obtain the desired estimate (15). □

9. Proof of Proposition 3.8

We now prove Proposition 3.8. This will be a reprise of the argument used to prove Proposition 3.3, but will use non-concentration to improve upon the inclusion $l^p \subset l^1$ used previously.

Fix $R \geq 2^{C_0 k}$, $0 < c \leq 2^{-C_0}$, $r' > 0$, and $C_0^C R \geq r > R^{1/2+3/N}$. Let ϕ, ψ be red and blue waves of frequency 1, 2^k respectively obeying (14). We may assume the normalization (11). Let Q_R be a spacetime cube of side-length R , and let (x_0, t_0) be a point in spacetime. To prove Proposition 3.8 it suffices to show that

$$\begin{aligned} \|\phi\psi\|_{L^p(Q_R \cap C^{purple}(x_0, t_0; r'))} &\leq (1 + CNc)A(R/C_0, \tilde{r}, r')E_{r, C_0 Q_R}(\phi, \psi)^{1/p} \\ &\quad + c^{-C} \left(1 + \frac{R}{2^{kr'}}\right)^{-1/N} 2^{k(\frac{1}{p} - \frac{1}{2})} \end{aligned}$$

where $\tilde{r} := r(1 - C_0 r^{-1/3N})$. Here we have used the fact that $E_{r, C_0 Q_R}(\phi, \psi)^{1/p} \sim 1$.

Applying Lemma 6.1 with

$$F := \phi\psi\chi_{C^{purple}(x_0, t_0; r')}$$

we see that there exists a cube Q of side-length CR contained in $C^2 Q_R$ such that

$$\|\phi\psi\|_{L^p(Q_R \cap C^{purple}(x_0, t_0; r'))} \leq (1 + CNc)\|\phi\psi\|_{L^p(X(Q) \cap C^{purple}(x_0, t_0; r'))}.$$

Let Φ, Ψ be as in Proposition 4.1. By (24) and the inclusion $2Q \subset C_0 Q_R$, it thus suffices to show that

$$\begin{aligned} \|[\Phi]_k[\Psi]_{C_0}\|_{L^p(X(Q) \cap C^{purple}(x_0, t_0; r'))} \\ \leq (1 + Cc)A(R/C_0, \tilde{r}, r')E_{r, 2Q}(\phi, \psi)^{1/p'} + c^{-C} R^{C-N/2}, \end{aligned}$$

since the latter factor can easily be absorbed into $c^{-C}(1 + \frac{R}{2^{kr'}})^{-1/N} 2^{k(\frac{1}{p} - \frac{1}{2})}$ since $R > 2^{C_0 k}$.

By (18) we can estimate $[\Phi]_k$ by $[\Phi]_{C_0}$. By (25) it thus suffices to show that

$$\begin{aligned} \|[\Phi]_{C_0}[\Psi]_{C_0}\|_{L^p(Q \cap C^{purple}(x_0, t_0; r'))} \\ \leq (1 + Cc)A(R/C_0, \tilde{r}, r')E_{\tilde{r}, 2Q}(\Phi, \Psi)^{1/p'}. \end{aligned}$$

Here we have used some trivial bound in $A(R/C_0, \tilde{r}, r')$ which is polynomial in R , e.g. by (7) and Hölder.

Raising both sides to the p^{th} power and expanding, we reduce to showing that

$$\sum_{q \in \mathcal{Q}_{C_0}(Q)} \|\Phi^{(q)}\Psi^{(q)}\|_{L^p(q)}^p \leq (1 + Cc)A(R/C_0, \tilde{r}, r')^p E_{\tilde{r}, 2Q}(\Phi, \Psi)^{p/p'}. \tag{29}$$

The cube q has side-length $2^{-C_0}CR \ll R/C_0$, so in particular $C_0q \subset 2Q$. By (17) and (21) we thus have

$$\|\Phi^{(q)}\Psi^{(q)}\|_{L^p(q)} A(R/C_0, \tilde{r}, r')(E(\Phi^{(q)})^{1/2}E(\Psi^{(q)})^{1/2})^{1/p} E_{\tilde{r}, 2Q}(\Phi^{(q)}, \Psi^{(q)})^{1/p'}.$$

Also, by (22) we have

$$E_{\tilde{r}, 2Q}(\Phi^{(q)}, \Psi^{(q)}) \leq (1 + Cc)E_{\tilde{r}, 2Q}(\Phi, \Psi).$$

Comparing these estimates to (29), it suffices to show that

$$\sum_{q \in \mathcal{Q}_{C_0}(Q)} E(\Phi^{(q)})^{1/2}E(\Psi^{(q)})^{1/2} \leq 1 + Cc.$$

But this follows from Cauchy-Schwarz and (22) (cf. the proof of Proposition 3.3). □

10. Spatial localization, and Huygens’ principle

We shall shortly begin the proof of Propositions 3.6 and 3.7, but we must first develop some machinery concerning the localization of a wave to a disk, and the resulting estimates arising from Huygens’ principle. This machinery will also be useful in proving the persistence of non-concentration estimate (56) in Sect. 15.

To localize a wave to a disk we shall need a bump function η_0 and an evolution operator $U(t)$, which we now construct.

Let η_0 to denote a fixed non-negative Schwarz function on \mathbf{R}^n with total mass 1 and whose Fourier transform is supported on the unit disk; such a function can be constructed for instance by $\eta_0 = \hat{\varphi} * \hat{\varphi}$, where φ is a real even bump function supported near the origin such that $\varphi(0) = 1$. For any $r > 0$ we define η_r by

$$\eta_r(x) := r^{-n}\eta_0\left(\frac{x}{r}\right). \tag{30}$$

The function η_r is thus an L^1 -normalized Schwarz function concentrating on an r -disk with good frequency localization properties. We shall use this

cutoff function both in this section and in the construction of the wave packet decomposition in Appendix I.

Fix $a(\xi)$ to be a bump function supported on $\underline{\Sigma}$ which equals 1 on the spatial projection of Σ^{red} and Σ^{blue} . We define the evolution operator $U(t) = U^{red}(t)$ by

$$\widehat{U(t)f}(\xi) := a(\xi)e^{2\pi it|\xi|} \hat{f}(\xi) \tag{31}$$

and the kernel K_t by

$$K_t(x) := \int a(\xi)e^{2\pi i(x \cdot \xi + t|\xi|)} d\xi.$$

We have the propagation law

$$\phi(t) = U(t)\phi(0) = \phi(0) * K_t \tag{32}$$

for all red waves ϕ with frequency 1, and all times t_1, t_2 . A stationary phase computation gives the pointwise estimate

$$|K_t(x)| \lesssim (1 + \text{dist}((x, t), C^{red}(0, 0)))^{-N^{10}}. \tag{33}$$

We shall also use the operator $U(t)$ in the proof of an energy estimate on cones (Lemma 13.1), and in the proof of wave packet decomposition (Lemma 15.2) in Appendix I.

Definition 10.1. If $D = D(x_D, t_D; r)$ is a disk, and ϕ is a red wave of frequency 1, we define $P_D\phi$ at time t_D by

$$P_D\phi(t_D) := (\chi_D * \eta_{r^{1-1/N}})\phi(t_D)$$

and at other times t by

$$P_D\phi(t) := U(t - t_D)P_D\phi(t_D).$$

Here η_r . For red waves of frequency 2^j we define P_D by the formula

$$P_D\mathbf{D}_j\phi := \mathbf{D}_jP_D\phi,$$

and for blue waves we define P_D by the formula

$$P_D\mathbf{T}\phi := \mathbf{T}P_D\phi.$$

The operator P_D localizes a wave to the disk D at time t_D , while $1 - P_D$ similarly localizes to D^{ext} . More precisely:

Lemma 10.2. *Let j be an integer, $r \geq C_0 2^{-j}$. Let $D = D(x_D, t_D; r)$ be a disk with radius r , and let ϕ be a red wave with frequency 2^j and margin $\text{margin}(\phi) \geq C_0(2^j r)^{-1+1/N}$. Then $P_D \phi$ is a red wave of frequency 2^j which satisfies margin estimate*

$$\text{margin}(P_D \phi) \geq \text{margin}(\phi) - C_0(2^j r)^{-1+1/N}$$

and the energy estimates

$$\|\tilde{\chi}_D^{-N} P_D \phi\|_{L^2(D_+^{ext})} \lesssim (2^j r)^{-N} E(\phi)^{1/2} \tag{34}$$

$$\|(1 - P_D)\phi\|_{L^2(D_-)} \lesssim (2^j r)^{-N} E(\phi)^{1/2} \tag{35}$$

$$E(P_D \phi) \leq \|\phi\|_{L^2(D_+)}^2 + C(2^j r)^{-N} E(\phi) \tag{36}$$

$$E((1 - P_D)\phi) \leq \|\phi\|_{L^2(D_-^{ext})}^2 + C(2^j r)^{-N} E(\phi) \tag{37}$$

$$E(P_D \phi), E((1 - P_D)\phi) \leq E(\phi) \tag{38}$$

where D_-, D_+ are the disks

$$D_{\pm} := D(x_D, t_D; r(1 \pm (2^j r)^{-1/2N})).$$

Note in particular that

$$\frac{1}{2}D \subset D_- \subset D \subset D_+ \subset 2D.$$

Proof. By scaling it suffices to verify this when $j = 0$. The claims follow directly from the easily verified estimates

$$0 \leq \chi_D * \eta_{r^{1-1/N}}(x) \leq 1,$$

$$\chi_D * \eta_{r^{1-1/N}}(x) \lesssim r^{-N} \text{ for } x \in D_-,$$

$$1 - Cr^{-N} \leq \chi_D * \eta_{r^{1-1/N}}(x) \text{ for } x \in D_+^{ext}.$$

□

Applying **T** we see that similar statements hold for blue waves.

We now investigate the localization properties of P_D for times other than t_D . Heuristically speaking, $P_D \phi$ is supported on the cone $C^{red}(x_D, t_D; Cr)$, while $(1 - P_D)\phi$ vanishes on the cube $Q(x_D, t_D; C^{-1}r)$. More precisely, we have

Lemma 10.3. *Let D be a disk of radius $r \geq C_0$, and let ϕ be a red wave of frequency 1 and margin $\text{margin}(\phi) \geq C_0 r^{-1+1/N}$. Let $1 \leq q \leq 2$, $R \gtrsim r$.*

– If ψ is a blue wave of arbitrary frequency, we have the finite speed of propagation law

$$\|((1 - P_D)\phi)\psi\|_{L^q(Q(x_D, t_D, C^{-1}r))} \lesssim r^{C-N} E(\phi)^{1/2} E(\psi)^{1/2} \quad (39)$$

and the Huygens' principle

$$\begin{aligned} \|(P_D\phi)\psi\|_{L^q(Q(x_D, t_D; R) \setminus C^{red}(x_D, t_D; Cr+R^{1/N}))} \\ \lesssim R^{C-N} E(\phi)^{1/2} E(\psi)^{1/2}. \end{aligned} \quad (40)$$

– If ψ is a blue wave of frequency 2^k for some $k \geq 0$ such that $\text{margin}(\psi) \gg (2^k r)^{-1+1/N}$, then we have

$$\begin{aligned} \|(P_D\phi)(P_D\psi)\|_{L^q(Q^{ann}(x_D, t_D; Cr+CR^{1/N}, R))} \\ \lesssim R^{C-N} E(\phi)^{1/2} E(\psi)^{1/2}. \end{aligned} \quad (41)$$

We will also need to use (39), (40), (41) with the roles of ϕ , ψ and red and blue reversed. If one repeats the arguments below, one finds that one loses a factor of 2^{Ck} on the right-hand sides; however, if $r \geq 2^k$ one can absorb these losses into the factors of r^{C-N} or R^{C-N} . Thus we may freely interchange ϕ and ψ and red and blue in the $r \geq 2^k$ case.

Proof. By Hölder it suffices to verify the claims when $q = 2$.

To prove (39) we first use (32), (33), and (35) to obtain

$$\|(1 - P_D)\phi\|_{L^\infty(Q(x_D, t_D; C^{-1}r))} \lesssim r^{C-N} E(\phi)^{1/2},$$

and (39) follows from (26). A similar argument gives

$$\|P_D\phi\|_{L^\infty(Q(x_D, t_D; R) \setminus C^{red}(x_D, t_D; Cr+R^{1/N}))} \lesssim R^{C-N} E(\phi)^{1/2}, \quad (42)$$

and (40) follows from (26).

We now prove (41). From (40) with ψ replaced by $P_D\psi$, we have

$$\begin{aligned} \|(P_D\phi)(P_D\psi)\|_{L^2(Q(x_D, t_D; R) \setminus C^{red}(x_D, t_D; Cr+R^{1/N}))} \\ \lesssim R^{C-N} E(\phi)^{1/2} E(\psi)^{1/2}. \end{aligned}$$

By applying (42) with R , r replaced by $2^k R$, $2^k r$, and then applying \mathbf{TD}_k , we obtain

$$\|P_D\psi\|_{L^\infty(Q(x_D, t_D; R) \setminus C^{blue}(x_D, t_D; Cr+R^{1/N}))} \lesssim R^{C-N} E(\psi)^{1/2},$$

so by (26) we have

$$\begin{aligned} \|(P_D\phi)(P_D\psi)\|_{L^2(Q(x_D, t_D; R) \setminus C^{blue}(x_D, t_D; Cr+R^{1/N}))} \\ \lesssim R^{C-N} E(\phi)^{1/2} E(\psi)^{1/2}. \end{aligned}$$

Combining this with the previous estimate we obtain (41). \square

The estimate (41) is one way of exploiting the transversality of red and blue null directions. We shall also rely on transversality in other ways, notably in Lemma 13.1 and in Lemma 14.3.

11. Proof of Proposition 3.6

We now have enough machinery to prove Propositions 3.6 and 3.7, except for a somewhat difficult estimate, Lemma 11.1, which we defer to Appendix II.

Fix $R \geq 2^{C_0 k}$, and let ϕ, ψ be red and blue waves of frequency 1 and 2^k respectively, obeying the margin requirement (10) and the energy normalization (11). Let Q_R be a cube of side-length R . To prove Proposition 3.6 we have to show that

$$\|\phi\psi\|_{L^p(Q_R)} \leq (1 - C_0^{-C}) \sup_{\substack{2^{C_0 k} \leq \tilde{R} \leq R \\ \tilde{R}^{1/2+4/N} \leq r}} A(\tilde{R}, r, C_0(1+r)) + 2^{CC_0} 2^{\varepsilon k} 2^{k(\frac{1}{p} - \frac{1}{2})}. \tag{43}$$

We may of course assume that ϕ, ψ are nearly extremal in the sense that

$$\|\phi\psi\|_{L^p(Q_R)} \sim A(R). \tag{44}$$

We may also assume that

$$A(R) \geq 2^{CC_0} 2^{k(\frac{1}{p} - \frac{1}{2})} \tag{45}$$

since (43) is trivial otherwise.

Let $0 < \delta < 1/2$ be a small number to be chosen later. Let r be the supremum of all radii $r \geq C_0 2^k$ such that

$$E_{r, C_0 Q_R}(\phi, \psi) \leq 1 - \delta, \tag{46}$$

or $r = C_0 2^k$ if no such radius exists.

Since ϕ, ψ are smooth and have finite energy, it is easy to see that r is well defined, and that one can find a disk $D = D(x_0, t_0; r)$ of radius r with t_0 in the lifespan of $C_0 Q_R$ for which

$$\min(\|\phi\|_{L^2(D)}^2, \|\psi\|_{L^2(D)}^2) \geq 1 - 2\delta. \tag{47}$$

Fix the disk $D(x_0, t_0; r)$, and define the disk D' by

$$D' := C_0^{1/2} D = D(x_0, t_0; C_0^{1/2} r)$$

and define the double conic region Ω by

$$\Omega := Q_R \cap C^{purple}(x_0, t_0; C_0(1+r)).$$

We now show that in Q_R , the expression $\phi\psi$ is mostly concentrated in Ω .

We split ϕ and ψ into red and blue waves respectively as

$$\phi = (1 - P_{D'})\phi + P_{D'}\phi, \quad \psi = (1 - P_{D'})\psi + P_{D'}\psi$$

where the projection operator $P_{D'}$ is as in the previous section.

From (47), (11), and (37) we have

$$E((1 - P_{D'})\phi), E((1 - P_{D'})\psi) \lesssim \delta + C_0^{-C}.$$

From (12) we thus see that

$$\|((1 - P_{D'})\phi)(1 - P_{D'})\psi\|_{L^p(Q_R)} \lesssim (\delta + C_0^{-C})A(R). \tag{48}$$

On the other hand, from (40) and its analogue with the roles of red and blue reversed (using the assumption $r \geq 2^k$; see the remarks following Lemma 10.3) we have

$$\|(P_{D'}\phi)\psi\|_{L^p(Q_R \setminus \Omega)}, \|((1 - P_{D'})\phi)P_{D'}\psi\|_{L^p(Q_R \setminus \Omega)}, \lesssim C_0^{-C}.$$

Combining these estimates using the triangle inequality we obtain

$$\|\phi\psi\|_{L^p(Q_R \setminus \Omega)} \lesssim (\delta + C_0^{-C})A(R).$$

We thus see from (44) that

$$\|\phi\psi\|_{L^p(Q_R \setminus \Omega)} \lesssim (\delta + C_0^{-C})\|\phi\psi\|_{L^p(Q_R)}.$$

Raising this to the p^{th} power and re-arranging, we see that

$$\|\phi\psi\|_{L^p(\Omega)} \geq (1 - C(\delta + C_0^{-C})^p)^{1/p} \|\phi\psi\|_{L^p(Q_R)}. \tag{49}$$

We now divide into two cases. First suppose that we are in the medium or low concentration case $r \geq R^{1/2+4/N}$. In particular we have $r > C_0 2^k$, so from the definition of r we see that (46) holds. From the definition of $A(R, r)$ (and the assumption $\delta \ll 1/2$) we thus have

$$\|\phi\psi\|_{L^p(\Omega)} \leq (1 - \delta)^{1/p'} A(R, r).$$

Combining these estimates and setting δ to some small negative power of C_0 we obtain the desired estimate (43) as desired (provided that C_0 is large enough; (5) suffices).

Now suppose instead that we are in the highly concentrated case $r \leq C_0 R^{1/2+4/N}$. Define \tilde{R} by the formula

$$\tilde{R} := \max(2^{C_0 k}, (r)^{\frac{1}{1/2+4/N}}).$$

thus $2^{C_0 k} \leq \tilde{R} \leq R$. In analogy to the previous case, we observe that

$$\|\phi\psi\|_{L^p(Q(x_0, t_0; \tilde{R}) \cap \Omega)} \leq (1 - \delta)^{1/p'} A(\tilde{R}, r)$$

when $\tilde{R} > 2^{C_0 k}$. When $\tilde{R} = 2^{C_0 k}$ one can use (13) (which has already been proven for this value of \tilde{R}) and (12) to obtain

$$\|\phi\psi\|_{L^p(Q(x_0, t_0; \tilde{R}) \cap \Omega)} \leq 2^{CC_0} 2^{\epsilon k} 2^{k(\frac{1}{p} - \frac{1}{2})}.$$

In either case, we see that if we could show that

$$\|\phi\psi\|_{L^p(Q_R \setminus Q(x_0, t_0; \tilde{R}))} \lesssim (\delta + C_0^{-C})A(R) \tag{50}$$

then we could repeat the argument in the first case and obtain (43) as desired.

It remains to prove (50). By (48) and the triangle inequality it suffices to show that

$$\|(P_{D'}\phi)\psi\|_{L^p(Q_R \setminus Q(x_0, t_0; \tilde{R}))} \lesssim (\delta + C_0^{-C})A(R)$$

and

$$\|((1 - P_{D'})\phi)P_{D'}\psi\|_{L^p(Q_R \setminus Q(x_0, t_0; \tilde{R}))} \lesssim (\delta + C_0^{-C})A(R).$$

By a dyadic decomposition and (45) this will follow from

Lemma 11.1. *Let $R \geq 2^{C_0k}$, and let $C_02^k < r \leq R^{1/2+4/N}$. Let $D = D(x_0, t_0, C_0^{1/2}r)$ be a disk. Then for any red wave ϕ and blue wave ψ of frequency 1, 2^k respectively with*

$$\text{margin}(\phi), \text{margin}(\psi) \geq 1/200$$

we have

$$\begin{aligned} &\|(P_D\phi)\psi\|_{L^p(Q^{ann}(x_0, t_0; R, 2R))}, \|\phi(P_D\psi)\|_{L^p(Q^{ann}(x_0, t_0; R, 2R))} \\ &\lesssim E(\phi)^{1/2}E(\psi)^{1/2}R^{-1/C}. \end{aligned}$$

The proof of this lemma shall be postponed to Appendix II, as it requires some machinery which is developed later in this paper, and the arguments used to prove this lemma are not used elsewhere. The estimate in Lemma 11.1 is not of an endpoint type, and could probably be proven by more elementary means than the methods used in the Appendix.

Assuming Lemma 11.1, the proof of Proposition 3.6 is thus complete. \square

12. Proof of Proposition 3.7

Let $R \geq 2^{C_0k}$, $0 < c \leq 2^{-C_0}$, and $r > C_0R$. Let ϕ, ψ be red and blue waves of frequency 1 and 2^k respectively obeying (11) and (10). Let Q_R be a cube of side-length R . To prove Proposition 3.7, it suffices to show that

$$\|\phi\psi\|_{L^p(Q_R)} \leq E_{r, C_0Q_R}(\phi, \psi)^{1/p}(1 + Cc)\overline{A(R)} + c^{-C}2^{k(\frac{1}{p} - \frac{1}{2})}.$$

Let (x_{Q_R}, t_{Q_R}) be the center of Q_R , and let D denote the disk

$$D := D(x_{Q_R}, t_{Q_R}; r/2).$$

We then divide

$$\phi = (1 - P_D)\phi + P_D\phi, \quad \psi = (1 - P_D)\psi + P_D\psi.$$

From Lemma 10.2 (especially (36)), Definition 3.4 and the hypothesis $R \geq 2^{C_0k}$ we see that $P_D\phi, P_D\psi$ obey the relaxed margin requirements (14) and the energy estimates

$$E(P_D\phi), E(P_D\psi) \leq E_{r, C_0Q_R}(\phi, \psi) + CR^{C-N}.$$

From Proposition 3.3 we thus have

$$\begin{aligned} & \| (P_D\phi)(P_D\psi) \|_{L^p(Q_R)} \\ & \leq (1 + Cc)(E_{r, C_0Q_R}(\phi, \psi) + CR^{C-N})\overline{A(R)} + c^{-C}C_0^C 2^{k(\frac{1}{p}-\frac{1}{2})}. \end{aligned}$$

The R^{C-N} term can easily be absorbed into the $c^{-C}C_0^C 2^{k(\frac{1}{p}-\frac{1}{2})}$ term by some trivial bound on $\overline{A(R)}$ (Proposition 3.2 will do). By the triangle inequality, we will thus be done if we can show

$$\| ((1 - P_D)\phi)\psi \|_{L^p(Q_R)}, \| (P_D\phi)(1 - P_D)\psi \|_{L^p(Q_R)} \leq c^{-C}C_0^C 2^{k(\frac{1}{p}-\frac{1}{2})}.$$

But these estimates follow easily from (39) and the counterpart with red and blue reversed (this is legitimate by the hypothesis $R \geq 2^{C_0k}$ and the remarks following Lemma 10.3). \square

13. Energy estimates on light cones of opposite color

To prove Theorem 1.1 it remains only to prove Proposition 4.1. Of course, we shall not use Propositions 3.2-3.8 in the proof of this Proposition.

In this section and the next, we derive two basic estimates needed to prove Proposition 4.1; these are ‘‘pigeonhole-free’’ versions of similar estimates in [25]. The first estimate is an extension of (26) from a cube to a conic neighbourhood of opposite color. The second is a strengthening of (7) when at least one of the waves ϕ, ψ has low frequency dispersion. Then, in Sect. 15, we combine these two estimates with a wave packet decomposition to derive a key ingredient (Proposition 15.1) in the proof of Proposition 4.1.

We begin with the energy estimate on cones:

Lemma 13.1. *Let ϕ be a red wave of frequency 2^j . Then for any $(x_0, t_0) \in \mathbf{R}^{n+1}$ and $R \gtrsim 2^{-j}$ we have*

$$\|\phi\|_{L^2(C^{blue}(x_0, t_0; R))} \lesssim R^{1/2} E(\phi)^{1/2}.$$

Proof. By translation invariance we may take $(x_0, t_0) = 0$, and by scaling we may take $j = 0$.

The blue cone is a null surface, and so standard energy estimates will not prove this estimate. However this can be salvaged since ϕ is red, so that the characteristics of ϕ will be transverse to the blue cone.

We turn to the details. Let $U(t)$ be the wave evolution operator defined in (31). By (32), it suffices to show that

$$\left(\int \|\chi_{S_t} U(t) f\|_{L^2(\mathbf{R}^n)}^2 dt\right)^{1/2} \lesssim R^{1/2} \|f\|_{L^2(\mathbf{R}^n)}$$

for all $f \in L^2(\mathbf{R}^n)$, where S_t is the slice $S_t := \{x: (x, t) \in C^{blue}(0, 0; R)\}$.

We shall apply the TT^* method. By duality the above estimate is equivalent to

$$\| \int U(t)^* (\chi_{S_t} F(t)) dt \|_{L^2(\mathbf{R}^n)} \lesssim R^{1/2} \|F\|_{L^2(\mathbf{R}^{n+1})}.$$

We square this as

$$\int \int \langle \chi_{S_s} U(s) U(t)^* (\chi_{S_t} F(t)), F(s) \rangle dt ds \lesssim R \|F\|_2^2$$

where $\langle f, g \rangle = \int_{\mathbf{R}^n} f(x) \overline{g(x)} dx$ is the usual complex inner product. By Cauchy-Schwarz and Young's inequality it suffices to show that

$$| \langle \chi_{S_s} U(s) U(t)^* (\chi_{S_t} F(t)), F(s) \rangle | \lesssim (1 + |s - t|/R)^{-N} \|F(t)\|_2 \|F(s)\|_2$$

for all s, t .

When $|s - t| \lesssim R$ this follows from Cauchy-Schwarz and the L^2 boundedness of the operator $\chi_{S_s} U(s) U(t)^* \chi_{S_t}$, so we may assume that $|s - t| \gg R$. The convolution operator $U(s) U(t)^*$ behaves essentially like $U(s - t)$, and its kernel $K_{s,t}(x)$ satisfies a similar kernel estimate to (33), namely

$$|K_{s,t}(x)| \lesssim (1 + \text{dist}((x, s - t), C^{red}(0, 0)))^{-N^{10}}.$$

The claim then follows from the transversality of $C^{red}(0, 0)$ and $C^{blue}(x_0, t_0)$ and crude estimates. □

Applying **T**, we see that a similar estimate holds with the roles of red and blue reversed. As a corollary of both Lemma 13.1 and its time-reversed counterpart, we obtain the following improvement to (27) when one restricts to a double cone.

Corollary 13.2. *Let ϕ, ψ be red and blue waves of frequency 1, 2^k respectively. Let $R > r \gg 1$, $(x_0, t_0) \in \mathbf{R}^{n+1}$, and Q_R be a cube of side-length R . Then*

$$\|\phi\psi\|_{L^2(C^{purple}(x_0, t_0; r) \cap Q_R)} \lesssim r^{1/2} R^{1/2} E(\phi)^{1/2} E(\psi)^{1/2}.$$

Proof. The contribution of $C^{blue}(x_0, t_0; r) \cap Q_R$ is acceptable by using Lemma 13.1 to control ϕ and (26) to control ψ . Similarly for the contribution of $C^{red}(x_0, t_0; r) \cap Q_R$. □

14. Bilinear L^2 estimates in the low dispersion case.

To prove Proposition 4.1 we shall need to study waves with rather low frequency dispersion. To make this concept precise, we introduce the following notation, which we shall re-use in later sections.

Definition 14.1. If ϕ is a red or blue wave (of any frequency), we define the *angular dispersion* of ϕ to be the quantity

$$\text{diam} \left\{ \frac{\xi}{|\xi|} : (\xi, \tau) \in \text{supp}(\hat{\phi}) \right\}.$$

A key observation (implicit in [16], and exploited explicitly in [25]) is that there is an improvement to (7) when either ϕ or ψ has low dispersion.

Lemma 14.2. [16, 25] *Let ϕ be a red wave of frequency 1, and let ψ be a blue wave. If ϕ has angular dispersion $O(1/r)$ for some $r \gg 1$, then*

$$\|\phi\psi\|_{L^2(\mathbf{R}^{n+1})} \lesssim r^{-(n-1)/2} E(\phi)^{1/2} E(\psi)^{1/2}. \tag{51}$$

Proof. Let ψ have frequency 2^k . From hypothesis, the frequency support of ϕ is contained in a sector Γ of width $O(1/r)$. By Plancherel’s theorem it thus suffices to show that

$$\|fd\sigma_1 * gd\sigma_2\|_2 \lesssim r^{-(n-1)/2} \|f\|_{L^2(\Gamma)} \|g\|_{L^2(2^k \Sigma^{blue})}$$

for all f and g , where $d\sigma_1$ and $d\sigma_2$ denote surface measure on Γ and $2^k \Sigma^{blue}$ respectively. By Young’s inequality we have

$$\|fd\sigma_1 * gd\sigma_2\|_1 \lesssim \|f\|_1 \|g\|_1$$

so it suffices by interpolation to show that

$$\|fd\sigma_1 * gd\sigma_2\|_\infty \lesssim r^{-(n-1)} \|f\|_\infty \|g\|_\infty,$$

which by positivity reduces to showing that

$$\|d\sigma_1 * d\sigma_2\|_\infty \lesssim r^{-(n-1)}. \tag{52}$$

But this follows from the observation that the intersection of Γ and $x - 2^k \Sigma^{blue}$ is always transverse and has $n - 1$ -dimensional measure at most $O(r^{-(n-1)})$ for any $x \in \mathbf{R}^{n+1}$. \square

This estimate is global in spacetime, however for our purposes it will be convenient to work with a localized form of this estimate.

Lemma 14.3. *Let j be an integer and $r \gg 1, 2^{-j}$. Let ϕ be a red wave with frequency 1 and angular dispersion $O(1/r)$, and let ψ be a blue wave with frequency 2^j such that*

$$\text{margin}(\phi), \text{margin}(\psi) \geq 1/200.$$

Then we have

$$\|\phi\psi\|_{L^2(Q)} \lesssim r^{-(n+1)/2} \|\phi\|_{L^2(CQ)} \|\psi\|_{L^2(CQ)} + r^{C-N} E(\phi)^{1/2} E(\psi)^{1/2} \tag{53}$$

for all cubes Q of side-length r .

Proof. The idea is to combine Lemma 13.1 with the localization machinery developed in Sect. 10.

Let D be the disk $D := D(x_Q, r_Q; C^{1/2}r)$. If C is sufficiently large, then from (39) and (38) we have

$$\|((1 - P_D)\phi)\psi\|_{L^2(Q)} \lesssim r^{C-N} E(\phi)^{1/2} E(\psi)^{1/2}$$

and

$$\|P_D\phi(1 - P_D)\psi\|_{L^2(Q)} \lesssim r^{C-N} E(\phi)^{1/2} E(\psi)^{1/2}.$$

From the triangle inequality we thus have

$$\|\phi\psi\|_{L^2(Q)} \leq \|(P_D\phi)(P_D\psi)\|_{L^2(\mathbf{R}^{n+1})} + Cr^{C-N} E(\phi)^{1/2} E(\psi)^{1/2}.$$

From Lemma 10.2 we see that $P_D\phi$ has dispersion $O(1/r)$. By Lemma 14.2 and (38) we thus have

$$\|\phi\psi\|_{L^2(Q)} \lesssim r^{-(n+1)/2} \|\phi\|_{L^2(CD)} \|\psi\|_{L^2(CD)} + r^{C-N} E(\phi)^{1/2} E(\psi)^{1/2}.$$

The claim then follows by letting t_D range over the lifespan of Q , averaging, and applying Cauchy-Schwarz. \square

An analogue of this lemma exists with the roles of red and blue reversed, but we shall not invoke this analogue explicitly.

15. Wave packet decomposition

Using the results of the previous section, we can now prove the main tool used to derive Proposition 4.1.

Proposition 15.1. *Let j, j' be integers, and let $R \gg 2^{-j}, 2^{j-2j'}, 0 < c \leq 2^{-C_0}$. Let Q be a spacetime cube of side-length R , ϕ be a red wave of frequency 2^j with $\text{margin}(\phi) \geq (2^j R)^{-1/2}$, and ψ be a blue wave of frequency $2^{j'}$.*

Then there exists a red wave table $\Phi = \Phi_c(\phi, \psi; Q)$ on Q with depth C_0 , frequency 2^j with margin

$$\text{margin}(\Phi) \geq \text{margin}(\phi) - C(2^j R)^{-1/2}$$

such that the following properties hold.

– ($[\Phi]_{C_0}$ approximates ϕ) *We have*

$$\|(|\phi| - [\Phi]_{C_0})\psi\|_{L^2(I^c, C_0(Q))} \lesssim c^{-C} \left(\frac{2^j}{R}\right)^{(n-1)/4} E(\phi)^{1/2} E(\psi)^{1/2}. \tag{54}$$

– (Bessel inequality) *We have*

$$E(\Phi) \leq (1 + Cc)E(\phi) \tag{55}$$

– (Persistence of non-concentration) *For any $r \gtrsim R^{1/2+1/N} 2^{-j(1/2-1/N)}$, we have*

$$E_{r(1-C_0(2^j r)^{-1/2N}), C_0 Q}(\Phi, \psi) \leq (1 + Cc)E_{r, C_0 Q}(\phi, \psi) \tag{56}$$

Roughly speaking, $\Phi^{(q)}$ is the portion of ϕ which concentrates in q . The Proposition is phrased using frequencies 2^j and $2^{j'}$ rather than 1 and 2^k in order to make it symmetric in ϕ and ψ , as we shall also need the time reversal of Proposition 15.1.

Proof. The quantities r, R have the units of length, while the frequencies $2^j, 2^{j'}$ have the units of inverse length. Finally, c and the margins are dimensionless. One can then verify that the entire Proposition is dimensionally correct and thus scale invariant. By scaling we may thus set $j = 0$. By translation invariance we may assume Q is centered at the origin. Set $r := 2^{-J} R$, where J is chosen so that $r \sim \sqrt{R}$.

We may assume that $c > R^{-1/10n}$ (for instance), since otherwise the claim is trivial by setting $\Phi = 0$ (and using some trivial bound such as (7) to show (54).

Let \mathcal{E} be a maximal $1/r$ -separated subset of $S^{n-1} \cap \underline{\Sigma}$, and let L denote the lattice $L := c^{-2}r\mathbf{Z}^n$. We define a *red tube* to be any set $T = T(\omega_T, x_T)$ of the form

$$T = \{(x, t) \in \mathbf{R}^{n+1} : |x - (x_T + \omega_T t)| \leq r\}$$

where $x_T \in L$ and $\omega_T \in \mathcal{E}$. We let $\mathbf{T} = \mathbf{T}^{red}$ denote the set of all red tubes. If T is a red tube, we define the cutoff function $\tilde{\chi}_T$ on \mathbf{R}^{n+1} by

$$\tilde{\chi}_T(x, t) := \tilde{\chi}_{D(x_T + \omega_T t, t; r)}(x). \tag{57}$$

We shall need the following careful decomposition of ϕ into wave packets which are concentrated on tubes in \mathbf{T} .

Lemma 15.2. *With the above notation, one can find for each $T \in \mathbf{T}$ a red wave ϕ_T with frequency 1 and angular dispersion at most $CR^{-1/2}$, such that*

- We have the margin estimate

$$\text{margin}(\phi_T) \geq \text{margin}(\phi) - CR^{-1/2}. \tag{58}$$

- The map $\phi \mapsto \phi_T$ is linear for each T , and

$$\phi = \sum_{T \in \mathbf{T}} \phi_T. \tag{59}$$

- We have

$$E(\phi_T) \lesssim c^{-C} \|\tilde{\chi}_T(t)\phi(t)\|_2 \tag{60}$$

for all $T \in \mathbf{T}$ and t in the lifespan of C_0Q .

- If $\text{dist}(T, Q) \geq C_0R$ then

$$\|\phi_T\|_{L^\infty(Q)} \lesssim \text{dist}(T, Q)^{C-N} E(\phi)^{1/2}. \tag{61}$$

- We have

$$\sum_{T \in \mathbf{T}} \sup_{q \in \mathcal{Q}_J(Q)} \tilde{\chi}_T(x_q, t_q)^{-3} \|\phi_T\|_{L^2(Cq)}^2 \lesssim c^{-C} r E(\phi). \tag{62}$$

- We have the Bessel inequality

$$\left(\sum_{q_0} E\left(\sum_{T \in \mathbf{T}} m_{q_0, T} \phi_T \right) \right)^{1/2} \leq (1 + Cc) E(\phi)^{1/2} \tag{63}$$

whenever q_0 runs over a finite index set and the $m_{q_0, T}$ are non-negative numbers such that

$$\sum_{q_0} m_{q_0, T} = 1 \tag{64}$$

for all $T \in \mathbf{T}$.

Roughly speaking, ϕ_T is the portion of ϕ which has frequency support in the sector of width $1/r$ and direction $(\omega_T, 1)$, and is spatially concentrated in T . A naive microlocalization to this region of space and frequency, taking some care to ensure that the ϕ_T are still waves, would obtain most of the above properties, but would probably need to replace the $(1 + Cc)$ factor in (63) by a larger constant, which would then cause a similar unacceptable loss in (22) and then destroy the induction. This necessitates a delicate construction of the ϕ_T based on averaging. Because the details of the proof are technical and not particularly relevant to the rest of the argument, we defer the proof of this Lemma to Appendix I, and continue with the proof of Proposition 15.1.

Using Lemma 15.2, we can now define Φ by

$$\Phi^{(q_0)} := \sum_{T \in \mathbf{T}} \frac{m_{q_0, T}}{m_T} \phi_T \tag{65}$$

for all $q_0 \in \mathcal{Q}$, where

$$m_{q_0, T} := \|\psi \tilde{\chi}_T\|_{L^2(q_0)}^2 + R^{-10n} E(\psi), \tag{66}$$

and

$$m_T := \sum_{q_0 \in \mathcal{Q}_{C_0}(Q)} m_{q_0, T} = \|\psi \tilde{\chi}_T\|_{L^2(Q)}^2 + R^{-10n} 2^{(n+1)C_0} E(\psi). \tag{67}$$

The $R^{-10n} E(\psi)$ factor is only present in (66) to ensure that m_T does not completely degenerate to zero. One can think of $\Phi^{(q_0)}$ as consisting of those wave packets ϕ_T such that $\psi \tilde{\chi}_T$ concentrates in q_0 .

It is clear that Φ is a red wave with

$$\text{margin}(\Phi) \geq \text{margin}(\phi) - CR^{-1/2}$$

and that

$$\phi = \sum_{q_0 \in \mathcal{Q}_{C_0}(Q)} \Phi^{(q_0)}. \tag{68}$$

The estimate (55) follows from (63).

We now show (56). Let $D = D(x_0, t_0; r(1 - C_0 r^{-1/2N}))$ be any disk of radius $r(1 - C_0 r^{-1/2N})$ with t_0 in the lifespan of $C_0 Q$. By (55) it then suffices to show that

$$\|\Phi\|_{L^2(D)} \|\psi\|_{L^2(D)} \leq (1 + Cc) E_{r', C_0 Q}(\phi, \psi). \tag{69}$$

Let D' denote the slightly larger disk $D' := D(x_0, t_0, r - \frac{C_0}{2} r^{-1/2N})$, and let D'' denote the even larger disk $D'' := D(x_0, t_0, r)$. We may divide

$$\phi = P_{D'} \phi + (1 - P_{D'}) \phi =: \phi_1 + \phi_2.$$

The map $\phi \mapsto \Phi$ is linear, so we may write $\Phi =: \Phi_1 + \Phi_2$ accordingly. From (55) and (36) we have

$$\begin{aligned} \|\Phi_1\|_{L^2(D)} &\leq E(\Phi_1)^{1/2} \leq (1 + Cc)E(\phi_1)^{1/2} \\ &\leq (1 + Cc)\|\phi\|_{L^2(D'')} + r^{C-N}E(\phi). \end{aligned} \tag{70}$$

Next, we claim that

$$\|\Phi_2\|_{L^2(D)} \lesssim r^{C-N}E(\phi). \tag{71}$$

To see this, we first consider the tubes T which do not intersect CQ . By (65), (61), and Hölder, their contribution is acceptable. Thus we need only consider those tubes which intersect CQ . By the triangle inequality it suffices to show

$$\|(\phi_2)_T\|_{L^2(D)} \lesssim r^{C-N}E(\phi)$$

for each tube T .

First suppose that $\text{dist}(T, D) \leq R^{1/2+1/100N}$. Then by (35), (38), and the assumption $r \gtrsim R^{1/2+1/N}$ we see that

$$\|\tilde{\chi}_T(t_0)\phi_2(t_0)\|_2 \lesssim r^{C-N}E(\phi).$$

The claim then follows from (60).

Now suppose that $\text{dist}(T, D) \geq R^{1/2+1/100N}$. By (62), the hypothesis $c > R^{-1/10}$ and Bernstein’s inequality (or Sobolev embedding) we see that

$$\|(\phi_2)_T\|_{L^\infty(q)} \lesssim r^{C-N}E(\phi_2) \lesssim r^{C-N}E(\phi)$$

whenever $\text{dist}(q, T) \geq R^{1/2+1/200N}$. The claim then follows. This completes the proof of (71).

Combining (70) and (71), we see that

$$\|\Phi\|_{L^2(D)} \leq (1 + Cc)\|\phi\|_{L^2(D'')} + r^{C-N}E(\phi),$$

and (69) easily follows (using the trivial inequality $\|\psi\|_{L^2(D)} \leq \|\psi\|_{L^2(D'')}$). This proves (56).

We now turn to the proof of (54). By (68) we have

$$(|\phi| - [\Phi]_{C_0})|\psi| \leq \sum_{q_0 \in \mathcal{Q}_{C_0}(Q)} |\Phi^{(q_0)}\psi|(1 - \chi_{q_0}),$$

so by the triangle inequality it suffices to show that

$$\|\Phi^{(q_0)}\psi\|_{L^2(I^{c,C_0}(Q) \setminus q_0)} \lesssim c^{-C}r^{-(n-1)/2}E(\phi)^{1/2}E(\psi)^{1/2}. \tag{72}$$

for each q in $\mathcal{Q}_{C_0}(Q)$.

Fix q_0 . We shall use a (heavily disguised) version of the arguments in Wolff [25]. If $q \in \mathcal{Q}_J(Q)$ intersects $I^{c, C_0}(Q) \setminus q_0$, then $\text{dist}(q, q_0) \gtrsim cR$. By squaring (72), we thus reduce to showing that

$$\sum_{q \in \mathcal{Q}_J(Q): \text{dist}(q, q_0) \gtrsim cR} \|\Phi^{(q_0)}\psi\|_{L^2(q)}^2 \lesssim c^{-C} r^{-(n-1)} E(\phi)E(\psi). \tag{73}$$

Consider a single summand from (73). From (65) and the triangle inequality we have

$$\|\Phi^{(q_0)}\psi\|_{L^2(q)} \leq \sum_{T \in \mathbf{T}} \frac{m_{q_0, T}}{m_T} \|\phi_T \psi\|_{L^2(q)}. \tag{74}$$

Consider the tubes T which do not intersect CQ . By (61) and (26), their total contribution to (74) is $O(R^{C-N} E(\phi)^{1/2} E(\psi)^{1/2})$. As for the tubes T which do intersect CQ , we may apply Lemma 14.3 (since $r \gg 1, 2^{-j'}$) to obtain

$$\begin{aligned} \|\phi_T \psi\|_{L^2(q)} &\lesssim r^{-(n+1)/2} \|\psi\|_{L^2(Cq)} \|\phi_T\|_{L^2(Cq)} \\ &\quad + R^{C-N} E(\phi_T)^{1/2} E(\psi)^{1/2}. \end{aligned}$$

The total contribution of the error term to (74) is $O(R^{C-N} E(\phi)^{1/2} E(\psi)^{1/2})$ by (63) and the fact that there are only $O(R^C)$ tubes being summed here. Combining all these estimates we obtain

$$\begin{aligned} &\|\Phi^{(q_0)}\psi\|_{L^2(q)} \\ &\lesssim r^{-(n+1)/2} \|\psi\|_{L^2(Cq)} \sum_{T \in \mathbf{T}} \frac{m_{q_0, T}}{m_T} \|\phi_T\|_{L^2(Cq)} + R^{C-N} E(\phi)^{1/2} E(\psi)^{1/2}. \end{aligned}$$

Inserting this back into (73), we see that it suffices to show that

$$\begin{aligned} \sum_{q \in \mathcal{Q}_J(Q): \text{dist}(q, q_0) \gtrsim cR} \|\psi\|_{L^2(Cq)}^2 \left(\sum_{T \in \mathbf{T}} \frac{m_{q_0, T}}{m_T} \|\phi_T\|_{L^2(Cq)} \right)^2 \\ \lesssim c^{-C} r^2 E(\phi)E(\psi). \tag{75} \end{aligned}$$

Using the trivial estimate

$$\frac{m_{q_0, T}}{m_T} \leq \frac{m_{q_0, T}^{1/2}}{m_T^{1/2}}$$

followed by Cauchy-Schwarz, we have

$$\left(\sum_{T \in \mathbf{T}} \frac{m_{q_0, T}}{m_T} \|\phi_T\|_{L^2(Cq)} \right)^2$$

$$\leq \left(\sum_{T \in \mathbf{T}} \frac{\|\phi_T\|_{L^2(Cq)}^2}{m_T \tilde{\chi}_T(x_q, t_q)} \right) \left(\sum_{T \in \mathbf{T}} m_{q_0, T} \tilde{\chi}_T(x_q, t_q) \right).$$

By (66), we have

$$\sum_{T \in \mathbf{T}} m_{q_0, T} \tilde{\chi}_T(x_q, t_q) \lesssim \|\psi\chi\|_2^2 + R^{-5n} E(\psi)$$

where

$$\chi := \left(\sum_{T \in \mathbf{T}} \tilde{\chi}_T(x_q, t_q) \tilde{\chi}_T^2 \right)^{1/2} \chi_{q_0}.$$

Since $\text{dist}(q, q_0) \gtrsim cR$, we see from elementary geometry that

$$\chi(x, t) \lesssim c^{-C} \left(1 + \frac{\text{dist}((x, t), C^{\text{red}}(x_q, t_q))}{r} \right)^{-10n}.$$

From Lemma 13.1 and the triangle inequality we thus have

$$\|\psi\chi\|_2^2 \lesssim c^{-C} r E(\psi).$$

Inserting all these estimates back into (75), we see that it will suffice to show that

$$\sum_{q \in \mathcal{Q}_J(Q)} \|\psi\|_{L^2(Cq)}^2 \sum_{T \in \mathbf{T}} \frac{\|\phi_T\|_{L^2(Cq)}^2}{m_T \tilde{\chi}_T(x_q, t_q)} \lesssim c^{-C} r E(\phi). \tag{76}$$

We re-arrange the left-hand side as

$$\sum_{T \in \mathbf{T}} \sum_{q \in \mathcal{Q}_J(Q)} \tilde{\chi}_T(x_q, t_q)^{-3} \|\phi_T\|_{L^2(Cq)}^2 \frac{\|\psi\|_{L^2(Cq)}^2 \tilde{\chi}_T(x_q, t_q)^2}{m_T}$$

and estimate this by

$$\sum_{T \in \mathbf{T}} \left(\sup_{q \in \mathcal{Q}_J(Q)} \tilde{\chi}_T(x_q, t_q)^{-3} \|\phi_T\|_{L^2(Cq)}^2 \right) \sum_{q \in \mathcal{Q}_J(Q)} \frac{\|\psi\|_{L^2(Cq)}^2 \tilde{\chi}_T(x_q, t_q)^2}{m_T}$$

Since the inner sum is $O(1)$ by (67), the desired estimate (76) then follows from (62). This proves (54). \square

We define the wave table $\Psi_c(\phi, \psi; Q)$ in analogy to $\Phi_c(\phi, \psi; Q)$ by time reversal:

$$\Psi_c(\phi, \psi; Q) := \mathbf{T}\Phi_c(\mathbf{T}\psi, \mathbf{T}\phi; \mathbf{T}Q).$$

Of course, one has the analogue of Proposition 15.1 for Ψ but with the roles of red and blue reversed.

16. Proof of Proposition 4.1

We are now ready to prove Proposition 4.1, in which $\phi\psi$ is replaced by a quilted analogue $[\Phi]_k[\Psi]_{C_0}$.

We begin with the construction of Φ ; this will be achieved by iterating Proposition 15.1.

We define recursively the wave tables ϕ_j on Q of depth j for all $0 \leq j \leq k$ which are multiples of C_0 (recall that we assumed k to be a multiple of C_0 in Sect. 3). Set ϕ_0 to be the wave table on Q of depth 0

$$(\phi_0)^{(Q)} := \phi$$

and then define inductively

$$\phi_{j+C_0}^{(q)} := \Phi_{c2^{-(k-j)/N}}(\phi_j^{(q)}, \psi; q)$$

for all $q \in \mathcal{Q}_j(Q)$. We then choose Φ to be $\Phi := \phi_k$.

By induction we see that

$$\text{margin}(\phi_j) \geq 1/100 - (2^k/R)^{1/N} - C(2^j/R)^{1/2}$$

for each j . This gives the desired margin requirements on Φ since $R \gg 2^k$. From (55) we have

$$E(\phi_j) \leq (1 + Cc2^{-(k-j)/N})E(\phi_{j-1})$$

for all $C_0 < j \leq k$. Telescoping this, we obtain

$$E(\phi_j) \leq (1 + Cc)E(\phi) = 1 + Cc \tag{77}$$

for all j . Thus (22) holds for Φ . A similar iteration of (56) (absorbing any factor of k which appears into a small power of $r^{1/N}$) yields

$$E_{r(1-r^{-1/3N}), C_0Q}(\Phi, \psi) \leq (1 + Cc)E_{r, C_0Q}(\phi, \psi) \tag{78}$$

for all $r \gtrsim R^{1/2+2/N}$.

Let $C_0 \leq j < k$ be a multiple of C_0 , and let q^* be a cube in $\mathcal{Q}_j(Q)$. From (54) with R replaced by $2^{1-j}R$, Q replaced by q^* , and ϕ replaced by $\phi_j^{(q^*)}$, we see that

$$\begin{aligned} & \|([\phi_j]_j - [\phi_{j+C_0}]_{j+C_0})\psi\|_{L^2(X(Q) \cap q^*)} \\ & \lesssim e^{-C}2^{C(k-j)/N}(2^{1-j}R)^{-\frac{n-1}{4}}E(\phi_j^{(q^*)})^{1/2}. \end{aligned}$$

Square-summing this in q^* and using (77) we obtain

$$\|([\phi_j]_j - [\phi_{j+C_0}]_{j+C_0})\psi\|_{L^2(X(Q))} \lesssim e^{-C}2^{C(k-j)/N}(2^{-j}R)^{-\frac{n-1}{4}}. \tag{79}$$

On the other hand, from Lemma 5.1 with R replaced by $2^k R$ and ϕ replaced by ϕ_j , we have

$$\|[\phi_j]_j \psi\|_{L^1(X(Q))} \lesssim 2^{-j/2} R.$$

Replacing j by $j + C_0$ and then subtracting, we obtain

$$\|([\phi_j]_j - [\phi_{j+C_0}]_{j+C_0}) \psi\|_{L^1(X(Q))} \lesssim 2^{CC_0} 2^{-j/2} R.$$

Interpolating this with (79) using (8), we obtain

$$\begin{aligned} \|([\phi_j]_j - [\phi_{j+C_0}]_{j+C_0}) \psi\|_{L^p(X(Q))} \\ \lesssim c^{-C} 2^{C(k-j)/N} 2^{-j(\frac{1}{p}-\frac{1}{2})} = c^{-C} 2^{-\alpha(k-j)} 2^{k(\frac{1}{p}-\frac{1}{2})} \end{aligned}$$

for some constant $\alpha > 0$. Telescoping this with the triangle inequality, we obtain

$$\|\phi \psi\|_{L^p(X(Q))} \leq \|[\Phi]_k \psi\|_{L^p(X(Q))} + c^{-C} 2^{k(\frac{1}{p}-\frac{1}{2})}. \tag{80}$$

Having constructed Φ , we now define Ψ as

$$\Psi := \Psi_c(\Phi, \psi; Q).$$

The estimate (22) for Ψ follows from the time reversal of (55). It remains to prove (23), (24), and (25).

To prove (23), we observe from the analogue of (54) that

$$\|\Phi(|\psi| - [\Psi]_{C_0})\|_{L^2(X(Q))} \lesssim c^{-C} 2^{CC_0} (R/2^k)^{-(n-1)/4} E(\Phi)^{1/2} E(\psi)^{1/2}.$$

From (22) and (18) we thus have

$$\|[\Phi]_k(|\psi| - [\Psi]_{C_0})\|_{L^2(X(Q))} \lesssim c^{-C} 2^{CC_0} (R/2^k)^{-(n-1)/4}. \tag{81}$$

From Lemma 5.1 we have

$$\|[\Phi]_k\|_{L^2(X(Q))} \lesssim 2^{-k} R^{1/2}. \tag{82}$$

From (22) we thus we have

$$\|[\Phi]_k \psi\|_{L^1(X(Q))}, \|[\Phi]_k \Psi\|_{L^1(X(Q))} \lesssim 2^{-k/2} R$$

so from (18) and the triangle inequality we have

$$\|[\Phi]_k(|\psi| - [\Psi]_{C_0})\|_{L^1(X(Q))} \lesssim 2^{-k/2} R.$$

Interpolating this with (81), we obtain

$$\|[\Phi]_k(|\psi| - [\Psi]_{C_0})\|_{L^p(X(Q))} \lesssim c^{-C} 2^{CC_0} 2^{k(\frac{1}{p}-\frac{1}{2})}, \tag{83}$$

and (23) follows from this, (80), and the triangle inequality.

Now we prove (24). We shall assume that $2^k r < R$ since (24) follows from (23) otherwise. For brevity let Ω denote the region

$$\Omega := X(Q) \cap C^{purple}(x_0, t_0; r).$$

We will need to localize (80) and (83) to Ω . We begin with the localization of (80). By (18) and Corollary 13.2 we have

$$\|[\phi_j]_j \psi\|_{L^1(\Omega)} \leq \|\phi_j \psi\|_{L^1(\Omega)} \lesssim r^{1/2} R^{1/2} \leq (2^k r/R)^{1/2} 2^{-j} R.$$

Replacing j by $j + C_0$ and then subtracting, we obtain

$$\|([\phi_j]_j - [\phi_{j+C_0}]_{j+C_0}) \psi\|_{L^1(\Omega)} \lesssim (2^k r/R)^{1/2} 2^{-j} R.$$

Interpolating this with (79) using (8), we obtain

$$\|([\phi_j]_j - [\phi_{j+C_0}]_{j+C_0}) \psi\|_{L^p(\Omega)} \lesssim c^{-C} 2^{C(k-j)/N} 2^{-j(\frac{1}{p}-\frac{1}{2})} (2^k r/R)^{1/N}.$$

Telescoping this with the triangle inequality, we obtain

$$\|\phi \psi\|_{L^p(\Omega)} \leq \|[\Phi]_k \psi\|_{L^p(\Omega)} + C c^{-C} 2^{CC_0} 2^{k(\frac{1}{p}-\frac{1}{2})} (2^k r/R)^{1/N}. \quad (84)$$

Now we localize (83). From (18) and Corollary 13.2 we have

$$\|[\Phi]_k \psi\|_{L^1(\Omega)}, \|[\Phi]_k \Psi\|_{L^1(\Omega)} \lesssim r^{1/2} R^{1/2} = (2^k r/R)^{1/2} 2^{-k/2} R.$$

so from (18) and the triangle inequality we have

$$\|[\Phi]_k(|\psi| - [\Psi]_{C_0})\|_{L^1(\Omega)} \lesssim (2^k r/R)^{1/2} 2^{-k/2} R.$$

Interpolating this with (81), we obtain

$$\|[\Phi]_k(|\psi| - [\Psi]_{C_0})\|_{L^p(\Omega)} \lesssim (2^k r/R)^{1/N} c^{-C} 2^{CC_0} 2^{k(\frac{1}{p}-\frac{1}{2})},$$

and (23) follows from this, (80), and the triangle inequality.

Finally, we show (25). From the time-reversed version of (56), and conceding various powers of 2^k and $r^{1/N}$, we see that

$$E_{r(1-r^{-1/3N}), C_0 Q}(\Phi, \Psi) \leq (1 + Cc) E_{r(1-r^{1/2N}), C_0 Q}(\Phi, \psi)$$

for all $r \gtrsim R^{1/2+3/N}$. Combining this with (78) we obtain (25) as desired. This completes the proof of Proposition 4.1, and hence of Theorem 1.1. \square

17. Null form estimates

In this section we now apply Theorem 1.1 to obtain some nearly sharp null form estimates. The arguments in this section do not require the methods used above to prove Theorem 1.1.

Let D_0, D_+, D_- denote the Fourier multipliers

$$\begin{aligned} \widehat{D_0\phi}(\xi, \tau) &:= |\xi|\widehat{\phi}(\xi, \tau) \\ \widehat{D_+\phi}(\xi, \tau) &:= (|\xi| + |\tau|)\widehat{\phi}(\xi, \tau) \\ \widehat{D_-\phi}(\xi, \tau) &:= \||\xi| - |\tau|\|\widehat{\phi}(\xi, \tau). \end{aligned}$$

If ϕ is a solution to the free wave equation and $s \in \mathbf{R}$, we define $\phi[0]$ to be the vector $(\phi(0), D_0^{-1}\phi_t(0))$. In particular, we have

$$\|\phi[0]\|_{\dot{H}^s} = (\|\phi(0)\|_{\dot{H}^s}^2 + \|\phi_t(0)\|_{\dot{H}^{s-1}}^2)^{1/2}$$

where $\dot{H}^s := D_0^{-s}L^2$ is the homogeneous Sobolev space of order s .

In [7] the following problem was considered:

Problem 17.1. Determine the set of all exponents $(p, \beta_0, \beta_+, \beta_-, \alpha_1, \alpha_2)$ such that one has the estimates

$$\|D_0^{\beta_0} D_+^{\beta_+} D_-^{\beta_-}(\phi\psi)\|_{L^p(\mathbf{R}^{n+1})} \lesssim \|\phi[0]\|_{\dot{H}^{\alpha_1}} \|\psi[0]\|_{\dot{H}^{\alpha_2}} \tag{85}$$

for all vector-valued solutions ϕ, ψ to the free wave equation (not necessarily red or blue).

This problem was resolved in [7] for $p = 2$, but is largely open otherwise (with some partial results in [23], [13]). A successful resolution to this problem (and of its generalization to mixed Lebesgue norms $L_t^q L_x^r$ and to when ϕ, ψ solve the inhomogeneous wave equation) is likely to have application to the low-regularity behaviour of non-linear wave equations.

To see the connection between Problem 17.1 and Theorem 1.1, we first observe the consequence of Theorem 1.1, which addresses Problem 17.1 in the frequency-localized case.

Proposition 17.2. Let β be a real number, and $k, l \geq 0$ be integers. Let ϕ and ψ be waves which have frequency supports in the sectors

$$\{(\xi_1, \xi_2, \xi'', \tau) : \tau = |\xi|, \xi_1 \sim 1, \xi_2 \sim 2^{-l}, |\xi''| \lesssim 2^{-l}\} \tag{86}$$

and

$$2^k\{(\xi_1, \xi_2, \xi'', \tau) : \tau = -|\xi|, -\xi_1 \sim 1, \xi_2 \sim 2^{-l}, |\xi''| \lesssim 2^{-l}\} \tag{87}$$

respectively. Then we have

$$\begin{aligned} & \| |\square|^\beta(\phi\psi) \|_p, \| |\square|^\beta(\phi\bar{\psi}) \|_p \\ & \lesssim 2^{\beta(k-2l)} 2^l \left(\frac{n+1}{p} - (n-1) \right) 2^{k \left(\frac{1}{p} - \frac{1}{2} + \varepsilon \right)} E(\phi)^{1/2} E(\psi)^{1/2} \end{aligned} \tag{88}$$

for all $2 \geq p \geq p_0$, $\varepsilon > 0$, where $|\square|$ is the multiplier $|\square| := D_+ D_-$, the energy $E(\cdot)$ is defined as in (1), and the implicit constants may depend on β, ε .

Proof. Consider the Minkowski-conformal linear transformation $L : \mathbf{R}^{n+1} \rightarrow \mathbf{R}^{n+1}$ given by

$$L(\xi_1, \xi', \tau) := \left(\frac{\xi_1 + \tau}{2} + 2^{-2l} \frac{\xi_1 - \tau}{2}, 2^{-l} \xi', \frac{\xi_1 + \tau}{2} - 2^{-2l} \frac{\xi_1 - \tau}{2} \right),$$

and define the associated operator T_L by

$$T_L \phi := \phi \circ L^*,$$

where L^* is the adjoint of L . A routine computation shows that $T_L |\square| = 2^{-2j} |\square| T_L$, so that

$$T_L (|\square|^\beta \phi \psi) = 2^{-2j\beta} |\square|^\beta (T_L \phi) (T_L \psi).$$

Also, we observe from Plancherel that $E(T_L \phi) \sim 2^{(n-1)j} E(\phi)$ and $E(T_L \psi) \sim 2^{(n-1)j} E(\psi)$. Finally, we have $\|T_L F\|_p = 2^{-(n+1)l/p} \|F\|_p$. Combining all these facts we see that to prove (88) it suffices to do so when $l = 0$.

Set $l = 0$. We can write

$$\begin{aligned} & |\square|^\beta(\phi\psi)(x, t) \\ & = 2^{nk} 2^{\beta k} \int \int e^{2\pi i x \cdot (2^k \eta + \xi)} e^{2\pi i t (|\xi| - 2^k |\eta|)} m(\xi, \eta) \hat{f}(\xi) \hat{g}(2^k \eta) \, d\xi d\eta \end{aligned}$$

where $f(x) := \phi(x, 0)$ and $g(x) := \psi(x, 0)$, and m is the symbol

$$m(\xi, \eta) := C a(\xi) a(\eta) (|\eta| |\xi| + \langle \eta, \xi \rangle)^\beta.$$

Since m is smooth and compactly supported, we may decompose m as a Fourier series

$$m(\xi, \eta) = \sum_{j, j' \in L} c_{j, j'} e^{-2\pi i j \cdot \xi} e^{-2\pi i j' \cdot \eta}$$

for $|\xi|, |\eta| \lesssim 1$, where L is some discrete lattice and $c_{j, j'}$ are rapidly decreasing co-efficients (uniformly in k). This implies that

$$D_-^{\beta-}(\phi\psi) = \sum_{j, j' \in L} c_{j, j'} \phi_j \psi_{j'}$$

where

$$\phi_j(x, t) := \phi(x - j, t), \quad \psi_{j'}(x, t) := \psi(x - 2^{-k}j', t).$$

The first claim of (88) then follows from this decomposition, the triangle inequality, (3), and the observation that $\phi_j, \psi_{j'}$ have the same energy as ϕ, ψ respectively. The second claim is proven similarly but also uses the observation that $\phi_j\psi_{j'}$ has the same magnitude as $\phi_j\psi_{j'}$. \square

We make the technical remark that the above estimates continue to hold if one replaces β by the complex number $\beta + it$, with constants that grow at most exponentially in t .

We now consider the general setting of Problem 17.1, where no frequency restrictions are assumed on ϕ or ψ .

By considering a large number of key examples, it was shown in [7] that the conditions

$$\beta_0 + \beta_+ + \beta_- = \alpha_1 + \alpha_2 + \frac{n + 1}{p} - n \tag{89}$$

$$p \geq p_0 \tag{90}$$

$$\beta_- \geq \frac{n + 1}{2p} - \frac{n - 1}{2} \tag{91}$$

$$\beta_0 \geq \frac{n + 1}{p} - n$$

$$\beta_0 \geq \frac{n + 3}{p} - (n + 1) \tag{92}$$

$$\alpha_1 + \alpha_2 \geq \frac{1}{p} \tag{93}$$

$$\alpha_1 + \alpha_2 \geq \frac{n + 3}{p} - n \tag{94}$$

$$\alpha_i \leq \beta_- + \frac{n}{2}$$

$$\alpha_i \leq \beta_- + \frac{n - 1}{2} + \frac{n + 1}{2} \left(\frac{1}{2} - \frac{1}{p} \right)$$

$$\alpha_i \leq \beta_- + \frac{n - 1}{2} + (n + 2) \left(\frac{1}{2} - \frac{1}{p} \right) \tag{95}$$

were necessary for (85), where $i = 1, 2$. In that paper it was also shown that the above conditions were also necessary when $p = 2$, except for the endpoint cases when one has equality in (92), [(91) and (93)], or [(91) and (95) for some $i = 1, 2$], with (85) being false in these cases. It was then conjectured that for the above conditions are similarly necessary when $p \neq 2$, except perhaps for some endpoints.

By combining Proposition 17.2 with dyadic decomposition arguments as in [23], we can obtain the following progress on this conjecture in the $p \leq 2$ case.

Theorem 17.3. *Let $p_0 \leq p \leq 2$. Then (85) holds whenever (89) holds, (91), (92), (95) hold with strict inequality for $i = 1, 2$, and the conditions (93), (94) are replaced by the more strict*

$$\alpha_1 + \alpha_2 > \frac{1}{2} + \frac{n + 3}{n - 1} \left(\frac{1}{p} - \frac{1}{2} \right). \tag{96}$$

If $p = p_0$, then one can let (92) be obeyed with equality.

This is sharp except for endpoints when $n = 2$ (since (96) nearly matches (94) in this case), and is somewhat sharp for $n > 2$. One can replace the multipliers D_0, D_+, D_- by other symbols which satisfy the same types of regularity and decay estimates, but we shall not pursue this matter here.

Proof. In order to apply the complex interpolation method we shall allow the indices $\beta_0, \beta_+, \beta_-$ to acquire an imaginary part *it*. For notational simplicity we keep ourselves to the case $t = 0$, but the reader may easily verify that the following argument also works for arbitrary t with constants which are at most exponential in t .

When $p = 2$ these claims were proven in [7] (see also [23]). Since we have replaced (93), (94) with the linear condition (96), we see by complex interpolation that it suffices to verify the claims when $p = p_0$. In this case (92) becomes $\beta_0 \geq 0$; since the operator $D_0^{\beta_0} D_+^{-\beta_0}$ is bounded in L^p by standard multiplier theory we may assume that $\beta_0 = 0$.

Let ϕ, ψ be solutions to the free wave equation. By a finite decomposition and time reversal symmetry we may assume that $\hat{\phi}$ is supported in the upper light cone Σ^+ and that $\hat{\psi}$ is supported on either the upper or the lower light cone Σ^\pm .

Write $\phi =: \sum_j \phi_j, \psi =: \sum_k \psi_k$, where ϕ_j, ψ_k have frequency supports on the region $D_+ \sim 2^j, D_+ \sim 2^k$ respectively. We now rewrite (85) as

$$\begin{aligned} & \left\| \sum_{j,k} D_+^{\beta_+ - \beta_-} |\square|^{\beta_-} (\phi_j \psi_k) \right\|_p \\ & \lesssim \left(\sum_j 2^{2\alpha_1 j} E(\phi_j) \right)^{1/2} \left(\sum_k 2^{2\alpha_2 k} E(\psi_k) \right)^{1/2}. \end{aligned}$$

It suffices to show that

$$\|D_+^{\beta_+ - \beta_-} |\square|^{\beta_-} (\phi_j \psi_k)\|_p \lesssim 2^{-\varepsilon|j-k|} (2^{2\alpha_1 j} E(\phi_j))^{1/2} (2^{2\alpha_2 k} E(\psi_k))^{1/2}$$

for all j, k and some $\varepsilon > 0$, since the claim then follows from the triangle inequality, Cauchy-Schwarz, and Young’s inequality for sums.

By symmetry we may take $j \leq k$; by scaling and (89) we may take $j = 0$. Having used (89), our arguments will no longer require this condition and we shall discard it. Since we are assuming (95) to hold with strict inequality we may absorb the ε into the α_i factors, and reduce ourselves to showing that

$$\|D_+^{\beta_+ - \beta_-} |\square|^{\beta_-} (\phi_0 \psi_k)\|_p \lesssim 2^{\varepsilon k} 2^{\alpha_2 k} E(\phi_0)^{1/2} E(\psi_k)^{1/2}.$$

Fix k . For any $l \geq 0$, decompose the double light cone into finitely overlapping projective sectors Γ of angular width 2^{-l} , and let $\phi_0 = \sum_{\Gamma} \phi_{0,\Gamma}$, $\psi_k = \sum_{\Gamma} \psi_{k,\Gamma}$ be a Fourier decomposition of ϕ_0, ψ_k subordinate to these sectors. It suffices to show that

$$\begin{aligned} \left\| \sum_{\Gamma, \Gamma': \angle(\Gamma, \Gamma') \sim 2^{-l}} D_+^{\beta_+ - \beta_-} |\square|^{\beta_-} (\phi_{0,\Gamma} \psi_{k,\Gamma'}) \right\|_p \\ \lesssim 2^{-\varepsilon l} 2^{\varepsilon k} 2^{\alpha_2 k} E(\phi_0)^{1/2} E(\psi_k)^{1/2} \end{aligned}$$

for all $l \geq 0$, since the claim follows by summing in l using the bilinear partition of unity based on angular separation (see e.g. [21], [23]). By the triangle inequality and Cauchy-Schwarz as before, it suffices to show that

$$\|D_+^{\beta_+ - \beta_-} |\square|^{\beta_-} (\phi_{0,\Gamma} \psi_{k,\Gamma'})\|_p \lesssim 2^{-\varepsilon l} 2^{\varepsilon k} 2^{\alpha_2 k} E(\phi_{0,\Gamma})^{1/2} E(\psi_{k,\Gamma'})^{1/2} \tag{97}$$

for each Γ, Γ' with angular separation 2^{-l} . By a spatial rotation we may assume that ϕ has frequency support in (86), and that either ψ or $\bar{\psi}$ has frequency support in (87).

Now suppose $k \gg 1$. In this case $D_+^{\beta_+ - \beta_-}$ is equal to 2^k times a harmless multiplier on the frequency support of $\phi_{0,\Gamma} \psi_{k,\Gamma'}$, so by Proposition 17.2 the left-hand side of (97) is majorized by

$$2^{k(\beta_+ - \beta_-)} 2^{\beta_- (k - 2l)} 2^{l\left(\frac{n+1}{p} - (n-1)\right)} 2^{k\left(\frac{1}{p} - \frac{1}{2} + \varepsilon\right)} E(\phi_{0,\Gamma})^{1/2} E(\psi_{k,\Gamma'})^{1/2}.$$

The claim (97) then follows after some algebra from (95), (89), and the assumption that (91) holds with strict inequality.

It remains to consider the case $k = O(1)$; by a mild Lorentz transformation we may make $k = 0$. (This affects D_+ slightly, but this change is irrelevant). Since Γ and Γ' differ in angle by 2^{-l} , some geometry shows that D_+ is at least 2^{-l} on the frequency support of $\phi\psi$. We may thus majorize $D_+^{\beta_+ - \beta_-}$ in L^p by $1 + 2^{l(\beta_+ - \beta_-)}$, and by Proposition 17.2 again the left-hand side of (97) is bounded by

$$(1 + 2^{l(\beta_+ - \beta_-)}) 2^{\beta_- (-2l)} 2^{l\left(\frac{n+1}{p} - (n-1)\right)} E(\phi_{0,\Gamma})^{1/2} E(\psi_{0,\Gamma'})^{1/2}.$$

The claim (97) then follows after some algebra from (96), (2), (89), and the assumption that (91) holds with strict inequality. \square

It seems plausible that many of the missing endpoints in the above result could be obtained if one was willing to prove a generalization of Proposition 17.2 which considered the interaction of multiple scales and multiple angular separations using more sophisticated tools than the triangle inequality, and in which the $2^{\varepsilon k}$ loss was eliminated.

Apart from the issue of endpoints, there is still the unsatisfactory gap between the condition (96) in Theorem 17.3, and the necessary conditions (93), (94) conjectured in [7], when $n > 2$. In particular, to resolve the conjecture we would need to consider exponents near the case

$$p = \frac{n + 2}{n} = p_0(n - 1), \quad \alpha_1 + \alpha_2 = 1/p, \tag{98}$$

which is the intersection of (93) and (94). This case turns out to be related to the open (and quite difficult)

Conjecture 17.4 (Machedon-Klainerman for the Schrödinger equation). Let ϕ, ψ be functions on \mathbf{R}^{n+1} which have frequency supports in

$$\left\{ \left(\xi, \frac{1}{2}|\xi|^2 \right) : \xi \in \underline{\Sigma} \right\}$$

and

$$\left\{ \left(\xi, \frac{1}{2}|\xi|^2 \right) : \xi \in -\underline{\Sigma} \right\}$$

respectively. Then (4) holds for all $p \geq p_0(n)$.

Proposition 17.5. *Suppose that (85) holds for some set of exponents satisfying (98). Then Conjecture 17.4 holds with n replaced by $n - 1$.*

Proof. We use the method of descent, exploiting the fact that the paraboloid in \mathbf{R}^n is a conic section of the light cone in \mathbf{R}^{n+1} .

To prove Conjecture 17.4 for $n - 1$, it suffices to verify it for $p = p_0(n - 1) = \frac{n+2}{n}$, since the other endpoint $p = \infty$ is trivial. From Plancherel’s theorem it suffices to show that

$$\begin{aligned} \left\| \int_{\underline{\Sigma}_{n-1}} \int_{\underline{\Sigma}_{n-1}} e^{2\pi i x \cdot (\xi - \eta)} e^{2\pi i t \cdot \frac{1}{2}(|\xi|^2 + |\eta|^2)} f(\xi)g(\eta) \, d\xi d\eta \right\|_{L^p(\mathbf{R}^{n-1+1})} \\ \lesssim \|f\|_2 \|g\|_2 \end{aligned} \tag{99}$$

for all C^∞ functions f, g on $\underline{\Sigma}_{n-1}$.

Fix f, g . Let $R \gg 1$ be a large number, and consider the functions ϕ_R, ψ_R defined on \mathbf{R}^{n+1} by

$$\phi_R(x_1, x', t) := \int_{-1}^1 \int e^{2\pi i(t+x_1)(R+s)} e^{2\pi i(t-x_1)\frac{|\xi|^2}{4(R+s)}} e^{2\pi i x' \cdot \xi} f(\xi) \, d\xi ds$$

and

$$\psi_R(x_1, x', t) := \int_{-1}^1 \int e^{-2\pi i(t+x_1)(R+s)} e^{2\pi i(t-x_1)\frac{|\xi|^2}{4(R+s)}} e^{2\pi i x' \cdot \xi} g(\xi) \, d\xi ds.$$

One can easily verify that ϕ_R and ψ_R solve the wave equation, and that

$$\|\phi_R[0]\|_{\dot{H}^{\alpha_1}} \sim R^{\alpha_1} \|f\|_2, \quad \|\psi_R[0]\|_{\dot{H}^{\alpha_2}} \sim R^{\alpha_2} \|g\|_2.$$

By the assumption that (85) holds at (98), we have

$$\|D_0^{\beta_0} D_-^{\beta_-} D_+^{\beta_+}(\phi\psi)\|_p \lesssim R^{1/p} \|f\|_2 \|g\|_2$$

for some $\beta_0, \beta_-, \beta_+$. However, a computation of the frequency support of $\phi\psi$ shows that D_0, D_-, D_+ are all given by smooth functions comparable to 1 on this support, and so we have

$$\|\phi\psi\|_p \lesssim R^{1/p} \|f\|_2 \|g\|_2$$

Making the change of variables $X = t + x_1, T = (t - x_1)/2R$, we have

$$\begin{aligned} \phi\psi(X, x', T) &= e^{4\pi i X R} \int_{-1}^1 \int_{-1}^1 \iint e^{2\pi i X(s+s')} e^{2\pi i T \left(\frac{|\xi|^2}{2+s/2R} + \frac{|\eta|^2}{2+s'/2R} \right)} \\ &\quad e^{2\pi i x' \cdot (\xi+\eta)} f(\xi) g(\eta) \, d\xi d\eta ds ds' \end{aligned}$$

and so we have

$$\|\phi\psi\|_{L_X^p L_{x'}^p L_T^p} \lesssim \|f\|_2 \|g\|_2,$$

the $R^{1/p}$ factor having cancelled against the Jacobian term.

The phase factor $2^{4\pi i X R}$ can be discarded. Letting $R \rightarrow \infty$ and taking limits, and then evaluating the s, s' integrations, we thus see that

$$\begin{aligned} \left\| \left(\frac{\sin 2\pi X}{X} \right)^2 \iint e^{2\pi i T \left(\frac{|\xi|^2}{2} + \frac{|\eta|^2}{2} \right)} e^{2\pi i x' \cdot (\xi+\eta)} f(\xi) \, d\xi d\eta \right\|_{L_X^p L_{x',T}^p} \\ \lesssim \|f\|_2 \|g\|_2, \end{aligned}$$

and (99) follows since the X behaviour is trivial. □

The Machedon-Klainerman conjecture for the Schrödinger equation looks very similar to the results proved in Theorem 1.1 with $k = 0$, however the method of proof breaks down because the strong Huygens’ principle ((40) and (41)) totally fails for this equation. The arguments in [25] fail for similar reasons (basically, the tubes T are no longer constrained to point in null directions). Only partial progress is known for this problem; for instance, when $n = 2$ this conjecture should hold for all $p \geq 2 - \frac{1}{3}$, but the best result achieved to date is $p > 2 - \frac{2}{17}$, see [22]. A resolution of this conjecture would also yield new results for the very difficult restriction and Bochner-Riesz problems for the paraboloid [21], [6] as well as the problem of pointwise convergence of the Schrödinger equation to the initial data [23].

The estimate (85) at the endpoint (98) shares some features in common with the conjectures (29), (30) in [25], and the resolutions to these problems may well be related.

18. Appendix I: Proof of Lemma 15.2

We now give the proof of Lemma 15.2. The main technical difficulty is to obtain a good constant in the main term of (63); this will be obtained by using characteristic functions to decompose the Fourier domain, followed by an averaging over rotations to smooth things out spatially.

Partition

$$S^{n-1} \cap \underline{\Sigma} = \bigcup_{\omega \in \mathcal{E}} A_\omega$$

where A_ω consists of those points in $S^{n-1} \cap \underline{\Sigma}$ which are closer to ω than any other element of \mathcal{E} . Thus A_ω is in the $O(1/r)$ -neighbourhood of ω .

Let $G \subset SO(n)$ denote the set of all rotations in \mathbf{R}^n which differ from the identity by $O(1/r)$. Let $d\Omega$ be a smooth compactly supported probability measure on the interior of G .

For each $\Omega \in G$ and $\omega \in \mathcal{E}$, we define the Fourier projection operators $P_{\Omega,\omega}$ for test functions f on \mathbf{R}^n by

$$\widehat{P_{\Omega,\omega} f}(\xi) := \chi_{\Omega(A_\omega)}(\xi/|\xi|) \hat{f}(\xi).$$

Note that

$$\phi(0) = \sum_{\omega \in \mathcal{E}} P_{\Omega,\omega} \phi(0) \tag{100}$$

for all $\Omega \in G$ and red waves ϕ of frequency 1.

The decomposition (100) is well-behaved in frequency, but has terrible spatial localization properties. To get around this we will now average in $d\Omega$ and then apply a spatial cutoff.

Recall the function η_0 constructed in Sect. 10. Write

$$\eta^{x_0}(x) := \eta_0 \left(\frac{c^2}{r}(x - x_0) \right)$$

for all $x_0 \in L$. From the Poisson summation formula we have

$$1 = \sum_{x_0 \in L} \eta^{x_0}.$$

We define the functions ϕ_T at time 0 by

$$\phi_T(x, 0) := \eta^{x_T}(x) \int (P_{\Omega, \omega_T} \phi(0))(x) d\Omega \tag{101}$$

and at other times by

$$\phi_T(t) := U(t)\phi_T(0),$$

where $U(t)$ was the evolution operator defined in (31). One may verify that ϕ_T is a red wave with angular dispersion $O(1/r)$, that the map $\phi \mapsto \phi_T$ is linear, and that (58), (59) hold.

We now show (60). From (101) and the rapid decay of η^{x_T} we have the pointwise estimate

$$\phi_T(0) \lesssim c^{-C} \tilde{\chi}_T(0)^4 \int |P_{\Omega, \omega_T} \phi(0)| d\Omega, \tag{102}$$

so it suffices by Minkowski and the L^2 boundedness of P_{Ω, ω_T} to show that

$$\|\tilde{\chi}_T(0)P_{\Omega, \omega_T} \phi(0)\|_2 \lesssim C_0^C \|\tilde{\chi}_T(t)P_{\Omega, \omega_T} \phi(t)\|_2.$$

for all $\Omega \in G$, since the C_0^C factor can be absorbed into the c^{-C} factor by the hypothesis $c \leq 2^{-C_0}$. In fact, we will show the stronger

$$\|\phi\|_{L^2(D(x_D, t_D; r))} \lesssim C_0 \|\tilde{\chi}_{D(x_D - (t - t_D)\omega_T, t; r)}^3 \phi(t)\|_2 \tag{103}$$

for all disks D of radius r and time co-ordinate $t_D = O(R)$, and any red wave ϕ with frequency support in the sector $\{(\xi, |\xi|) : |\xi| \sim 1, \angle(\xi, \omega_T) \lesssim C/r\}$. The previous claim follows by breaking up $\tilde{\chi}_T(0)$ into various disks D .

We now show (103). We have the identity

$$\phi(x, t_D) = \int e^{2\pi i(x \cdot \xi + (t - t_D)|\xi|)} \varphi(\xi) \widehat{\phi(t)}(\xi) d\xi$$

where φ is an arbitrary bump function which equals one on the sector $\{\xi \in \underline{\Sigma} : \angle(\xi, \omega_T) \lesssim 1/r\}$ and is adapted to a slight enlargement of this sector. By standard stationary phase estimates and the observation that $t_D - t = O(C_0 r^2)$, we thus have

$$\phi(t_D) = \phi(t) * K$$

where the kernel K satisfies the estimates

$$|K(x)| \lesssim C_0^C r^{1-n} (1+|x-(t-t_D)\omega_T|/r)^{-N^{10}} (1+|(x-(t-t_D)\omega_T)\cdot\omega_T|)^{-N^{10}}.$$

The claim (103) then follows from a dyadic decomposition of K around the point $(t-t_D)\omega_T$, followed by Young’s inequality. This proves (60).

We now show (61). Write $R' := \text{dist}(T, Q)$. Let D denote the disk $D = D(0, 0; C^{-1}R')$. Since $R' \gg R$, it suffices from (32), (33), and crude estimates to show that

$$\|\phi_T\|_{L^2(D)} \lesssim R'^{-N} E(\phi)^{1/2}.$$

But this is clear from (101) and the estimates

$$\|\eta^{x_T}\|_{L^\infty(D)} \lesssim R'^{-N}, \quad \|P_{\Omega, \omega_T} \phi(0)\|_2 \leq E(\phi)^{1/2}.$$

This proves (61).

We now show (62). From (103) the left-hand side of (62) is majorized by

$$r \sum_T \sup_q \tilde{\chi}_T(x_q, t_q)^{-3} \|\tilde{\chi}_{D(x_q+t_q\omega_T, 0; r)}^3 \phi_T\|_2^2.$$

From (102) this is majorized by

$$c^{-C} r \sum_T \sup_q \tilde{\chi}_T(x_q, t_q)^{-3} \|\tilde{\chi}_{D(x_q+t_q\omega_T, 0; r)}^3 \tilde{\chi}_T(0)\|^4 \int |P_{\Omega, \omega_T} \phi(0)| \, d\Omega\|_2^2.$$

From (57), (6) we have the elementary inequality

$$\tilde{\chi}_{D(x_q+t_q\omega_T, 0; r)}(x) \tilde{\chi}_T(x, 0) \lesssim \tilde{\chi}_T(x_q, t_q)$$

so we may bound the previous by

$$c^{-C} r \sum_T \|\tilde{\chi}_T(0)\| \int |P_{\Omega, \omega_T} \phi(0)| \, d\Omega\|_2^2.$$

Summing in x_T , we reduce to showing that

$$\sum_{\omega \in \mathcal{E}} \left\| \int |P_{\Omega, \omega} \phi(0)| \, d\Omega \right\|_2^2 \lesssim E(\phi).$$

But this follows from Minkowski’s inequality followed by Plancherel’s theorem.

We now show (63). We expand the left-hand side as

$$\left(\sum_{q_0} \left\| \int \sum_{\omega \in \mathcal{E}} \sum_{x_0 \in L} m_{q_0, T(\omega, x_0)} \eta^{x_0} P_{\Omega, \omega}(\phi(0)) \, d\Omega \right\|_2^2 \right)^{1/2}.$$

By Minkowski’s inequality this is less than or equal to

$$\int \left(\sum_{q_0} \left\| \sum_{\omega} \sum_{x_0} m_{q_0, T(\omega, x_0)} \eta^{x_0} P_{\Omega, \omega}(\phi(0)) \right\|_2^2 \right)^{1/2} d\Omega. \tag{104}$$

Define the set $Y \subset S^{n-1}$ by

$$Y := \bigcup_{\omega \in \mathcal{E}} \{ \alpha \in A_{\omega} : \text{dist}(\alpha, S^{n-1} \setminus A_{\omega}) > Cc^2/r \};$$

this set Y plays a similar role here to the sets $I^{c,k}(Q)$ used in Proposition 4.1. and define $P_{\Omega(Y)}$ to be the multiplier

$$\widehat{P_{\Omega(Y)} f}(\xi) := \chi_{\Omega(Y)} \left(\frac{\xi}{|\xi|} \right) \hat{f}(\xi).$$

By the triangle inequality, (104) is less than the sum of

$$\int \left(\sum_{q_0} \left\| \sum_{\omega} \sum_{x_0} m_{q_0, T(\omega, x_0)} \eta^{x_0} P_{\Omega, \omega} P_{\Omega(Y)} \phi(0) \right\|_2^2 \right)^{1/2} d\Omega. \tag{105}$$

and

$$\int \left(\sum_{q_0} \left\| \sum_{\omega} \sum_{x_0} m_{q_0, T(\omega, x_0)} \eta^{x_0} P_{\Omega, \omega} (1 - P_{\Omega(Y)}) \phi(0) \right\|_2^2 \right)^{1/2} d\Omega. \tag{106}$$

Consider the quantity (105). The contributions of each ω are orthogonal as ω varies, so we can rewrite (105) as

$$\int \left(\sum_{q_0} \sum_{\omega} \left\| \sum_{x_0} m_{q_0, T(\omega, x_0)} \eta^{x_0} P_{\Omega, \omega} P_{\Omega(Y)} \phi(0) \right\|_2^2 \right)^{1/2} d\Omega,$$

which can be re-arranged as

$$\int \left(\sum_{\omega} \int |(P_{\Omega, \omega} P_{\Omega(Y)} \phi(0))(x)|^2 \sum_{q_0} \left(\sum_{x_0} m_{q_0, T(\omega, x_0)} \eta^{x_0}(x) \right)^2 dx \right)^{1/2} d\Omega.$$

This is clearly less than or equal to

$$\int \left(\sum_{\omega} \int |(P_{\Omega, \omega} P_{\Omega(Y)} \phi(0))(x)|^2 \left(\sum_{q_0} \sum_{x_0} m_{q_0, T(\omega, x_0)} \eta^{x_0}(x) \right)^2 dx \right)^{1/2} d\Omega.$$

Summing in q_0 and then in x_0 , this simplifies to

$$\int \left(\sum_{\omega} \int |(P_{\Omega,\omega} P_{\Omega(Y)} \phi(0))(x)|^2 dx \right)^{1/2} d\Omega.$$

By the orthogonality of the $P_{\Omega,\omega}$, we thus have

$$(105) \leq \int \|P_{\Omega(Y)} \phi(0)\|_2 d\Omega \leq E(\phi)^{1/2}.$$

Now consider (106). Repeating the above argument, but noting that we must use almost orthogonality instead of orthogonality, we obtain

$$(106) \lesssim \int \|(1 - P_{\Omega(Y)})\phi(0)\|_2 d\Omega.$$

By Cauchy-Schwarz we thus have

$$(106) \lesssim \left(\int \|(1 - P_{\Omega(Y)})\phi(0)\|_2^2 d\Omega \right)^{1/2};$$

by Plancherel we thus have

$$(106) \lesssim \left(\int |\hat{\phi}(\xi, 0)|^2 \left(\int \chi_{\Omega(Y)^c} \left(\frac{\xi}{|\xi|} \right) d\Omega \right) d\xi \right)^{1/2}.$$

The inner integral can be seen to be $O(c^2)$ for all ξ , so by Plancherel again we have

$$(106) \lesssim cE(\phi)^{1/2}.$$

Combining our estimates for (105) and (106) we obtain (63). □

19. Appendix II: Proof of Lemma 11.1

To complete the proof of Theorem 1.1 we have to prove Lemma 11.1. This shall be a straightforward application of the wave packet decomposition Lemma 15.2 and the energy estimates on cones developed in Sect. 13. However, our arguments are not as delicate as those in the rest of the proof of Theorem 1.1, as Lemma 11.1 is not an endpoint estimate. For instance, we will be able to lose powers of $R^{1/N}$ in our estimates.

We turn to the details. Let $R \geq 2^{C_0 k}$, and let ϕ, ψ be red and blue waves of frequency 1, 2^k respectively and margins at least $1/200$. We may assume the energy normalization (11). By translation invariance it suffices to show that

$$\|(P_D \phi)\psi\|_{L^p(Q^{ann}(0,0;R,2R))}, \|\phi(P_D \psi)\|_{L^p(Q^{ann}(0,0;R,2R))} \lesssim R^{-1/C}$$

for any disk D centered at the origin with radius $2^k \leq r \leq C_0 R^{1/2+4/N}$. We shall only prove the estimate for $(P_D \phi)\psi$; the estimate for $\phi(P_D \psi)$ is proven similarly, observing that any additional powers of 2^k which appear can be absorbed in to the $R^{-1/C}$ factor.

As in previous arguments, we obtain the L^p estimate by interpolation between an L^1 estimate and an L^2 estimate, namely

$$\|(P_D \phi)\psi\|_{L^1(Q^{ann}(0,0;R,2R))} \lesssim R^{C/N} R^{3/4} \tag{107}$$

and

$$\|(P_D \phi)\psi\|_{L^2(Q^{ann}(0,0;R,2R))} \lesssim R^{C/N} R^{-(n-1)/4}. \tag{108}$$

By interpolation and some algebra involving (8) one obtains the desired result if N is sufficiently small. The point is that (107) improves over (27) by a substantial power of R .

We first prove (107). The contribution outside of the set $C^{red}(0, 0; R^{1/2})$ is acceptable by (40), and the contribution inside this set is acceptable by Corollary 13.2. This proves (107).

Now we prove (108). Set $Q := Q(0, 0; 2R)$. We use Lemma 15.2 with this cube Q and some arbitrary value of c (say $c = 2^{-C_0}$) to decompose $(P_D \phi)$ as

$$(P_D \phi) = \sum_{T \in \mathbf{T}} (P_D \phi)_T.$$

By (61) and crude estimates (e.g. (7) or Lemma 14.2) the contribution of those tubes T for which $\text{dist}(T, 0) \gg R$ is acceptable, so we can restrict ourselves to those tubes T for which $\text{dist}(T, 0) \lesssim R$. There are only $O(R^C)$ of these tubes.

We now dispose of those remaining tubes T for which $R^{1/2+1/N} \lesssim \text{dist}(T, 0) \lesssim R$. For such tubes T we see from (60) with $t = 0$ and (34) that

$$E((P_D \phi)_T) \lesssim \|\tilde{\chi}_T(0)P_D \phi(0)\|_2^2 \lesssim R^{-N} E(\phi) = R^{-N},$$

and one can dispose of the contribution of these tubes by crude estimates.

It remains to deal with those tubes for which $\text{dist}(T, 0) \lesssim R^{1/2+1/N}$. In this case we want to exploit the fact that $(P_D \phi)_T$ is concentrated on $R^{1/N}T$, the dilate of T around its center by $R^{1/N}$. Indeed, from (62) we see that

$$\|(P_D \phi)_T\|_{L^2(Q \setminus R^{1/N}T)} \lesssim R^{-N},$$

and so the contribution outside $R^{1/N}T$ is acceptable by crude estimates. It thus remains to control the expression

$$\| \sum_T (P_D \phi)_T \psi \chi_{R^{1/N}T} \|_{L^2(Q^{ann}(0,0;R,2R))}.$$

We now observe the geometric fact that as T ranges over all tubes in \mathbf{T} with $\text{dist}(T, 0) \lesssim R^{1/2+1/N}$, the tubes $\chi_{R^{1/N}T}$ have an overlap of at most $O(R^{C/N})$ in $Q^{\text{ann}}(0, 0; R, 2R)$. By almost orthogonality, it thus suffices to show that

$$\left(\sum_T \|(P_D \phi)_T \psi\|_2^2 \right)^{1/2} \lesssim R^{-(n-1)/4}.$$

Applying Lemma 14.2, we can bound the left-hand side by

$$R^{-(n-1)/4} \left(\sum_T E((P_D \phi)_T) E(\psi) \right)^{1/2}.$$

But this is acceptable by (55) and (38). \square

References

1. M. Beals, Self-Spreading and strength of Singularities for solutions to semilinear wave equations, *Annals of Math* **118** (1983), 187–214
2. J. Bourgain, Estimates for cone multipliers, *Operator Theory: Advances and Applications*, **77** (1995), 41–60
3. J. Bourgain, Some new estimates on oscillatory integrals, *Essays in Fourier Analysis in honor of E. M. Stein*, Princeton University Press (1995), 83–112
4. J. Bourgain, Global well-posedness of defocusing critical non-linear Schrödinger equation in the radial case, to appear, *J. Amer. Math. Soc.*
5. J. Bourgain, On the dimension of Kakeya sets and related maximal inequalities, *Geom. Funct. Anal.* **9** (1999), 256–282
6. A. Carbery, Restriction implies Bochner-Riesz for paraboloids., *Math. Proc. Cambridge Philos. Soc.* **111** (1992), no. 3, 525–529
7. D. Foschi, S. Klainerman, Homogeneous L^2 bilinear estimates for wave equations, to appear, *Les Annales Scientifiques et L'Ecole Normale Supérieure*
8. N. Katz, I. Łaba, T. Tao, An improved bound on the Minkowski dimension of Besicovitch sets in \mathbf{R}^3 , to appear, *Annals of Math.* math.CA/9903166
9. N. Katz, T. Tao, A new bound on partial sum-sets and difference sets, and applications to the Kakeya conjecture, *Math Res. Letters* **6** (1999), 625–630. math.CO/9906097
10. S. Klainerman, M. Machedon, Space-time estimates for null forms and the local existence theorem, *Comm. Pure Appl. Math.* **46** (1993), no. 9, 1221–1268
11. S. Klainerman, M. Machedon, Remark on Strichartz-type inequalities. With appendices by Jean Bourgain and Daniel Tataru. *Internat. Math. Res. Notices* **5** (1996), 201–220
12. S. Klainerman, M. Machedon, On the regularity properties of a model problem related to wave maps, *Duke Math. J.* **87** (1997), 553–589
13. S. Klainerman, D. Tataru, On the optimal regularity for Yang-Mills equations in \mathbf{R}^{4+1} , *J. Amer. Math. Soc.* **12** (1999), 93–116
14. A. Moyua, A. Vargas, L. Vega, Schrödinger Maximal Function and Restriction Properties of the Fourier transform, *International Math. Research Notices* **16** (1996)
15. A. Moyua, A. Vargas, L. Vega, Restriction theorems and Maximal operators related to oscillatory integrals in \mathbf{R}^3 , to appear, *Duke Math. J.*
16. G. Mockenhaupt, A note on the cone multiplier, *Proc. AMS* **117** (1993), 145–152

17. E. M. Stein, *Singular Integrals and Differentiability Properties of Functions*, Princeton University Press, 1970
18. E. M. Stein, *Harmonic Analysis*, Princeton University Press, 1993
19. R. S. Strichartz, Restriction of Fourier Transform to Quadratic Surfaces and Decay of Solutions of Wave Equations, *Duke Math. J.*, **44** (1977), 705–774
20. C. D. Sogge, *Lectures on Nonlinear Wave Equations*, Monographs in Analysis II, International Press, 1995
21. T. Tao, A. Vargas, L. Vega, A bilinear approach to the restriction and Keakeya conjectures, *J. Amer. Math. Soc.* **11** (1998), pp. 967–1000
22. T. Tao, A. Vargas, A bilinear approach to cone multipliers I. Restriction Estimates, to appear, *GAFA*
23. T. Tao, A. Vargas, A bilinear approach to cone multipliers II. Applications, to appear, *GAFA*
24. T. H. Wolff, An improved bound for Keakeya type maximal functions, *Revista Mat. Iberoamericana.* **11** (1995). 651–674
25. T. Wolff, A sharp bilinear restriction estimate, to appear, *IMRN*