

## THE QUARTILE OPERATOR AND POINTWISE CONVERGENCE OF WALSH SERIES

CHRISTOPH THIELE

ABSTRACT. The bilinear Hilbert transform is given by

$$H(f, g)(x) := p.v. \int f(x-t)g(x+t) \frac{dt}{t}.$$

It satisfies estimates of the type

$$\|H(f, g)\|_r \leq C(s, t) \|f\|_s \|g\|_t.$$

In this paper we prove such estimates for a discrete model of the bilinear Hilbert transform involving the Walsh Fourier transform. We also reprove the well-known fact that the Walsh Fourier series of a function in  $L^p[0, 1]$ , with  $1 < p$  converges pointwise almost everywhere. The purpose of this exposition is to clarify the connection between these two results and to present an easy approach to recent methods of time-frequency analysis.

### 1. INTRODUCTION

Time-frequency analysis deals with decompositions of functions into pieces which are well localized in time and frequency. The most restrictive definition for a function in one real parameter, which we call time, to be localized in a (bounded) time interval  $I$  is that it is supported in this interval. Being localized in a frequency interval  $\omega$  then means that the Fourier transform is supported in  $\omega$ . However, with this definition there is no nonzero function which is localized both in time and frequency.

We describe a weaker notion of localization. Fix a function  $f$  in the Schwartz class. We define it to be localized in the time interval  $[0, 1)$  and the frequency interval  $[0, 1)$ . Then, for any real numbers  $k, n, l$ , we define the function

$$f_{k,n,l}(x) := f(2^k x - x_0) e^{2\pi i \xi_0 (2^k x - x_0)}$$

to be localized in the time interval

$$[2^{-k} \xi_0, 2^k (\xi_0 + 1)).$$

Thus we have for each pair  $(I, \omega)$  of intervals with  $|I| |\omega| = 1$  a function that is localized in the time interval  $I$  and the frequency interval  $\omega$ . Observe that this notion of localization comes with uniform estimates of the localized functions and their Fourier transforms away from the intervals  $I$  and  $\omega$ .

---

Received by the editors September 25, 1997.

2000 *Mathematics Subject Classification*. Primary 42A50, 42A20, 42C10.

*Key words and phrases*. Bilinear Hilbert transform, convergence of Fourier series, Walsh functions.

Sometimes it is convenient to view the pair  $(I, \omega)$  as a rectangle  $I \times \omega$  in an abstract plane, the phase plane. We take this point of view.

In the applications we have in mind, i.e., proofs of boundedness of the bilinear Hilbert transform [8], [9], and almost everywhere convergence of Fourier series ([4], [5]), time-frequency localization is used in the following way. We write all functions that are involved in the desired estimates in various ways as sums of time-frequency localized functions. Then we use the fact that two time-frequency localized functions that are localized in disjoint rectangles in the phase plane are almost orthogonal.

A simplifying approach to these applications is to pretend that functions, which are localized in disjoint rectangles, are strictly orthogonal, and thus provide heuristic proofs, which of course fail to be strict proofs. In this paper we choose a slightly different approach and define a discrete model of Fourier analysis, in which we do have functions which are localized both in time and frequency in the strict sense, but nevertheless all other aspects of Fourier analysis are essentially preserved. In addition we hope that maybe some reader will be interested in the resulting model theorems by themselves.

Define the Walsh functions recursively on the unit interval  $[0, 1)$  by  $W_0 := 1$  and, for every nonnegative integer  $n$ ,

$$W_{2n}(x) := \begin{cases} W_n(x) & \text{if } x < 0.5, \\ W_n(2x - 1) & \text{if } x \geq 0.5, \end{cases}$$

$$W_{2n+1}(x) := \begin{cases} W_n(2x) & \text{if } x < 0.5, \\ -W_n(2x - 1) & \text{if } x \geq 0.5. \end{cases}$$

These functions form an orthonormal bases of  $L^2([0, 1))$ , much alike the Fourier basis. They originate in [14], where the enumeration of the Walsh functions is different, a modification that is inessential for our discussion. We extend these functions to the real line by having them vanish outside the unit interval. Then we define for integers  $j, l, n$  and  $n \geq 0$ :

$$w_{n,l,j}(x) := 2^{-j/2} W_n(2^{-j}x - l).$$

This function is localized in the rectangle

$$[2^j l, 2^j(l+1)) \times [2^{-j}n, 2^{-j}(n+1))$$

in the strict sense that it is supported in the interval  $[2^j l, 2^j(l+1))$  and its Walsh Fourier transform is supported in the interval  $[2^{-j}n, 2^{-j}(n+1))$ . In particular two such functions are orthogonal, if they are localized in disjoint rectangles. We will prove this orthogonality later without reference to the Walsh Fourier transform. We would like to emphasize that the collection  $w_{n,j,l}$  of functions is obviously not a basis of  $L^2(\mathbb{R})$ , but rather an infinite union of bases.

We will reprove the following theorem, which is the analogue of the famous Carleson Hunt theorem. Its proof goes back to [2] and [11], a nice exposition is given in [6].

**Theorem 1.1.** *For every function  $f \in L^p[0, 1)$  with  $1 < p \leq \infty$ , the partial sums*

$$\sum_{n=0}^N (f, W_n) W_n(x)$$

*converge to  $f(x)$  as  $N \rightarrow \infty$  for almost every  $x$  in  $[0, 1)$ .*

A *finite dyadic step function* is a compactly supported function on the real line which is constant on dyadic intervals of length  $2^{-N}$  for some  $N$  depending on  $f$ . The *quartile operator*  $H_W$  is a bilinear operator defined originally for finite dyadic step functions by

$$H_W(f_1, f_2) := \sum_{j,l=-\infty}^{\infty} \sum_{n=0}^{\infty} 2^{-j/2} (f_1, w_{j,l,4n}) (f_2, w_{j,l,4n+1}) w_{j,l,4n+2}.$$

We will prove the following theorem:

**Theorem 1.2.** *Let  $1 < r_1, r_2 \leq \infty$  and  $\frac{2}{3} < r < \infty$ , such that  $1/r = 1/r_1 + 1/r_2$ . Then we have for all finite dyadic step functions  $f_1, f_2$ :*

$$\|H_W(f_1, f_2)\|_r \leq C \|f_1\|_{r_1} \|f_2\|_{r_2}$$

*with a constant  $C$  depending only on  $r_1$  and  $r_2$ . Hence the operator  $H_W$  extends to a bounded operator  $L^{r_1} \times L^{r_2} \rightarrow L^r$ .*

The results of this paper are essentially taken from the author's Ph.D. thesis at Yale University, 1995. The exposition has been revised, since joint work with M. Lacey has improved the understanding of the subject. The author is grateful for having learned analysis in the Walsh phase plane through collaboration with L. Villemoes, which has resulted in the paper [13]. Moreover the author is particularly grateful to his thesis advisor R. Coifman for suggesting this subject and for generous and continued support.

## 2. THE WALSH PHASE PLANE

**2.1. Geometric properties of the Walsh phase plane.** A *dyadic interval* is an interval of the form  $[2^j n, 2^j(n+1))$  with integers  $j, n$ . A main property of dyadic intervals is that given any two of them, they are either disjoint or one is contained in the other.

The *Walsh phase plane* is the half plane  $\Omega := \mathbb{R} \times [0, \infty)$ . We refer to the first coordinate of this set as time and the second coordinate as frequency. A rectangle  $I \times \omega \subset \Omega$  such that  $|I| |\omega| = 1$  and both  $I$  and  $\omega$  are dyadic intervals is called a *tile*. We will denote tiles by lower case letters  $p, q, u, l$  etc., and write  $I_p$  for the time interval and  $\omega_p$  for the frequency interval of a tile  $p$ . Similarly we define *bitiles* and *quartiles* to be dyadic rectangles in  $\Omega$  of area 2 and 4 resp. Bitiles and quartiles will generally be denoted by capital letters  $P, Q$  etc.

If  $P = [x_0, x_1) \times [\xi_0, \xi_1)$  is a bitile, we call the tile

$$u_P := [x_0, x_1) \times [\xi_0, (\xi_0 + \xi_1)/2)$$

the *upper son* of  $P$ . Similarly we define the *lower*, *left*, and *right sons* to be the tiles

$$\begin{aligned} l_P &:= [x_0, x_1) \times [(\xi_0 + \xi_1)/2, \xi_1), \\ l'_P &:= [x_0, (x_0 + x_1)/2) \times [\xi_0, \xi_1), \\ r_P &:= [(x_0 + x_1)/2, x_1) \times [\xi_0, \xi_1). \end{aligned}$$

**Lemma 2.1.** *Let  $P$  and  $Q$  be different bitiles. Then at least one of the intersections  $u_P \cap u_Q$  and  $l_P \cap l_Q$  is empty.*

To prove the lemma, assume that both intersections are nonempty. Pick two points  $(x, \xi) \in u_P \cap u_Q$  and  $(x', \xi') \in l_P \cap l_Q$ . It is easy to see that the frequency interval of  $P$  is the smallest dyadic interval that contains  $\xi$  and  $\xi'$ . Since the same holds for the frequency interval of  $Q$ , the frequency intervals of  $P$  and  $Q$  are equal. This implies that the time intervals of both bitiles have the same length, and as the bitiles have nonempty intersection, the time intervals must be equal. Hence  $P = Q$ , which proves the lemma.

We define a tile  $p$  to be *less than* a tile  $p'$ , and we write  $p \leq p'$ , if  $I_p \subset I_{p'}$  and  $\omega_{p'} \subset \omega_p$ . This defines a partial order on the set of all tiles. Observe that two tiles are *comparable*, i.e.,  $p \leq p'$  or  $p' \leq p$ , if and only if they have nonempty intersection. If  $\mathbf{p}$  is a set of tiles, we denote by  $\mathbf{p}^{\min}$  the set of minimal tiles in  $\mathbf{p}$  and by  $\mathbf{p}^{\max}$  the set of maximal tiles in  $\mathbf{p}$ .

**Lemma 2.2.** *Let  $\mathbf{p}$  be a finite set of pairwise disjoint tiles. Let  $\tilde{\mathbf{p}}$  be the set of all tiles  $\tilde{p}$  with the property*

$$\tilde{p} \subset \bigcup_{p \in \mathbf{p}} p.$$

*Then either  $\mathbf{p} = \mathbf{p}^{\min}$  or there is a bitile  $P$  such that  $u_P \in \mathbf{p}$  and  $l_P \in \mathbf{p}$ .*

*Proof.* The tiles in  $\mathbf{p}^{\min}$  are pairwise incomparable and therefore pairwise disjoint. Area considerations give that the cardinality of  $\mathbf{p}$  is greater than or equal to the cardinality of  $\mathbf{p}^{\min}$ . If  $\mathbf{p} \neq \mathbf{p}^{\min}$ , we can therefore assume that there is a  $p \in \mathbf{p}$  which is not minimal in  $\tilde{\mathbf{p}}$ . We pick a tile in  $\mathbf{p} \setminus \mathbf{p}^{\min}$  with maximal time interval  $I_p$ . By symmetry we can assume that this tile is the upper son  $u_P$  of a bitile  $P$ . Pick a tile  $\tilde{p} \in \tilde{\mathbf{p}}$  which is strictly less than  $u_p$ . It follows that  $E := \tilde{p} \cap l_P$  is not empty.

Since  $E$  is covered by  $\mathbf{p}$ , there is a tile  $q \in \mathbf{p}$  with  $q \cap E \neq \emptyset$ . It is easy to see that  $|\omega_q| \leq \omega_{u_p}$ , since  $q$  and  $p$  are disjoint. Moreover  $q \in \mathbf{p} \setminus \mathbf{p}^{\min}$  since  $\tilde{p} \leq q$ , and by choice of  $p$  we have  $|I_q| \leq |I_{u_p}|$ . Together with the estimate of the frequency intervals this implies that  $u_p$  and  $q$  are congruent. Since  $q \cap l_P \neq \emptyset$  we have  $q = l_P$ . This finishes the proof.  $\square$

**Corollary 2.3.** *Let  $\mathbf{p}$  and  $\tilde{\mathbf{p}}$  be as in Lemma 2.2. Then*

$$\bigcup_{p \in \mathbf{p}} p = \bigcup_{\tilde{p} \in \tilde{\mathbf{p}}^{\min}} \tilde{p}.$$

If  $\mathbf{p} \neq \mathbf{p}^{\min}$ , we find a bitile  $P$  such that  $u_P, l_P \in \mathbf{p}$ . Replacing these two tiles by the left and right son of  $P$  gives a new collection  $\tilde{\mathbf{p}}$  which covers the same set of the Walsh phase plane. Since  $\tilde{\mathbf{p}}$  is obviously finite, iteration of this procedure must stop after finitely many steps. This implies the corollary.

**Corollary 2.4.** *Let  $\mathbf{p}$  and  $\tilde{\mathbf{p}}$  be as in Lemma 2.2. Then for each  $\tilde{p}$  in  $\tilde{\mathbf{p}}$  there is a set  $\mathbf{p}'$  of pairwise disjoint tiles with  $\tilde{p} \in \mathbf{p}'$  and*

$$\bigcup_{p \in \mathbf{p}'} p = \bigcup_{p \in \mathbf{p}} p.$$

By the previous corollary we can assume that  $\mathbf{p} = \mathbf{p}^{\min}$ . We can also assume that each tile in  $\mathbf{p}^{\min}$  intersects  $\tilde{p}$  nontrivially, otherwise we remove all those, which do not. Then it is easy to see that  $\tilde{p}$  is maximal in  $\tilde{\mathbf{p}}$ . Pick  $\mathbf{p}'$  to be the set of maximal tiles in  $\tilde{\mathbf{p}}$ . By the dual version to the previous corollary this proves the corollary.

The following lemma gives a criterion for a subset of the phase plane to be decomposable into a disjoint union of tiles. We define for bitiles the same partial order as for tiles. Then a set  $\mathbf{P}$  of bitiles is called *convex*, if for all ordered triples  $P \leq P' \leq P''$  of bitiles we have the property that if  $P$  and  $P''$  are in  $\mathbf{P}$ , then  $P'$  is also in  $\mathbf{P}$ .

**Lemma 2.5.** *Let  $\mathbf{P}$  be a convex set of bitiles. Then there is a set  $\mathbf{p}$  of pairwise disjoint tiles such that*

$$\bigcup_{p \in \mathbf{p}} p = \bigcup_{P \in \mathbf{P}} P.$$

We use induction on the number of bitiles. Clearly the lemma is true for the empty set of bitiles. Given a nonempty convex set of bitiles we pick a minimal bitile  $P$  in the set. Let  $P'$  be the unique bitile such that  $u_P = l'_{P'}$  or  $u_P = r_{P'}$ . If  $u_P$  is not disjoint from all bitiles in  $\mathbf{P} \setminus \{P\}$ , pick a  $P'' \in \mathbf{P} \setminus \{P\}$  which intersects  $u_P$  nontrivially. Clearly we have  $P \leq P' \leq P''$ , and by convexity we conclude  $P' \in \mathbf{P}$ . This means that  $u_P$  is contained in the union of bitiles in  $\mathbf{P} \setminus \{P\}$ . By symmetry the same argument is true for  $l_P$ . Hence the induction hypothesis implies the statement of the lemma.

**2.2. Basic algebraic properties of the Walsh phase plane.** As in the introduction, we associate to each tile in the Walsh phase plane a *Walsh wave packet*: A tile  $p$  is determined by three integer parameters  $j, l, n$  with  $n \geq 0$ , such that

$$p = [2^j l, 2^j(l+1)) \times [2^{-j} n, 2^{-j}(n+1)).$$

We associate to this tile the wave packet

$$w_p(x) = w_{j,l,n}(x) = 2^{-j/2} W_n(2^{-j}x - l).$$

Observe that the wave packet  $w_p$  is supported on the time interval  $I_p$  of the tile  $p$ , and that its absolute value is constant on this interval.

**Lemma 2.6.** *Let  $P$  be a bitile and let  $w_{u_P}, w_{l_P}, w_{l'_P}$ , and  $w_{r_P}$  be the wave packets corresponding to upper, lower, left and right son of  $P$ . Then we have*

$$\begin{aligned} w_{l_P} &= \frac{1}{\sqrt{2}}(w_{l'_P} + w_{r_P}), \\ w_{u_P} &= \frac{1}{\sqrt{2}}(w_{l'_P} + w_{r_P}). \end{aligned}$$

After unwinding the definitions this is nothing but a reformulation of the recursion relations for the Walsh functions.

**Corollary 2.7.** *Let  $\mathbf{p}$  and  $\mathbf{p}'$  be two finite sets of pairwise disjoint tiles such that*

$$\bigcup_{p \in \mathbf{p}} p = \bigcup_{p' \in \mathbf{p}'} p'.$$

*Then the corresponding sets of wave packets span the same vector subspace of  $L^2(\mathbf{R})$ .*

By Corollary 2.3 we can assume that  $\mathbf{p}' = \tilde{\mathbf{p}}^{\min}$ , where  $\tilde{\mathbf{p}}$  is the set of tiles  $\tilde{p}$  which satisfy

$$\tilde{p} \subset \bigcup_{p \in \mathbf{p}} p.$$

If  $\mathbf{p}$  is not equal  $\tilde{\mathbf{p}}^{\min}$ , we find a bitile  $P$  as in Lemma 2.2. Replacing  $l_P$  and  $u_P$  by  $l'_P$  and  $r_P$  we observe from the previous lemma that the new collection of wavepackets spans the same subspace of  $L^2(\mathbf{R})$ . Proceeding successively, this proves the corollary.

Call a subset of the Walsh phase plane *good*, if it is the disjoint union of a set of tiles. The previous corollary assigns to each good set  $S$ , a unique subspace of  $L^2(\mathbf{R})$ . We denote by  $\Pi_S$  the orthogonal projection in  $L^2(\mathbf{R})$  onto this subspace.

**Corollary 2.8.** *Let  $S$  be a good set and let  $p$  be a tile with  $p \subset S$ . Then the wave packet  $w_p$  is contained in the image of  $\Pi_S$ .*

This is an immediate consequence of Corollaries 2.4 and 2.7.

**Lemma 2.9.** *If two tiles  $p$  and  $p'$  are disjoint, then the corresponding wave packets are orthogonal in  $L^2(\mathbf{R})$ .*

If the time intervals of the two tiles are disjoint, the two wave packets have disjoint support, which implies orthogonality. If the time intervals are not disjoint, then the frequency intervals must be disjoint. Without loss of generality assume  $I_p \subset I_{p'}$ . Let  $\mathbf{p}$  be the set of tiles of the form  $I_{p'} \times \omega$  which intersect  $p$ . Clearly we have  $p' \notin \mathbf{p}$ . The wave packet corresponding to a tile  $I_{p'} \times \omega$  is a Walsh function scaled to the interval  $I_{p'}$ , and two different Walsh functions are known to be orthogonal. Hence by the previous corollary the two wave packets in question are contained in orthogonal subspaces of  $L^2(\mathbf{R})$ , which proves the lemma.

**2.3. Carleson measure estimates.** Let  $f$  be a finite dyadic step function and  $\mathbf{p}$  a set of pairwise disjoint tiles in the phase plane. Lemma 2.9 says that the corresponding wavepackets are pairwise orthogonal, so we obtain the Bessel inequality

$$\sum_{p \in \mathbf{p}} |(f, w_p)|^2 \leq \|f\|_2^2.$$

We will be concerned with sets of pairwise orthogonal tiles such that

$$\|\Pi_p f\|_\infty = \frac{|(f, w_p)|}{\sqrt{|I_p|}} \geq 2^k$$

for some fixed  $k$ . Then the previous inequality becomes

$$\sum_{p \in \mathbf{p}} |I_p| \leq C \frac{\|f\|_2^2}{2^{2k}}.$$

We can write the left-hand side of this inequality as

$$\sum_{p \in \mathbf{p}} |I_p| = \int_{\mathbf{R}} N_{\mathbf{p}}(x) dx,$$

where  $N_{\mathbf{p}}$  is the function which assigns to each time  $x$  the number of tiles in  $\mathbf{p}$  whose time intervals contain  $x$ .

We call a bound as above on the sum of time intervals or the function  $N_{\mathbf{p}}$  a *Carleson measure estimate*. The following lemma describes Carleson measure estimates involving  $L^p$  norms of  $f$ .

**Lemma 2.10.** *Let  $f$  be a finite dyadic step function. Denote the conjugate exponent of  $r$  by  $r'$ , i.e.,  $1/r' + 1/r = 1$ . For each integer  $k$  let  $\mathbf{p}_k$  be a set of pairwise disjoint tiles  $p$  such that  $\|\Pi_p f\|_\infty \geq 2^k$ . Then the following are true:*

1. If  $1 < r < 2$ , then

$$\|N_{\mathbf{p}_k}^{1/(r'+\varepsilon)}\|_r \leq C(\varepsilon, r) \frac{\|f\|_r}{2^k}$$

for all  $\varepsilon > 0$ .

2. If  $r = 2$ , then

$$\sum_{p \in \mathbf{p}_k} |I_p| \leq C \frac{\|f\|_2^2}{2^{2k}}.$$

3. If  $2 < r < \infty$ ,

$$\sum_{k=-\infty}^{\infty} 2^{rk} \sum_{p \in \mathbf{p}_k} |I_p| \leq C(r) \|f\|_r^r.$$

We have already seen the lemma for  $r = 2$ . For  $r > 2$  we use that the function  $f$  is locally square integrable. Denote by  $Mf$  the maximal function

$$Mf(x) = \sup_{I \text{ dyadic}; x \in I} \frac{1}{|I|} \int_I |f(y)| dy.$$

Clearly the time interval of a tile in  $\mathbf{p}_k$  must be contained in the set

$$\{x : Mf(x) \geq 2^k\}.$$

Hence we can apply our  $L^2$  estimate to the restriction of  $f$  to this set. We claim that

$$\int_{\{x: Mf(x) \geq 2^k\}} |f(y)|^2 dy \leq C 2^{2k} |\{x : Mf(x) \geq 2^k\}|.$$

Multiplying this claimed inequality by  $2^{rk-2k}$  and summing over  $k$  gives the lemma in the case  $r > 2$ , because the right hand side becomes  $\|Mf\|_r^r$ , which by the maximal theorem is bounded by  $\|f\|_r^r$ .

To see the claim, observe that on each maximal dyadic subinterval  $I$  of the set  $\{x : |Mf(x)| > 2^{k-1}\}$  we have

$$\frac{1}{2|I|} \int_I |f(y)|^2 dy \leq C 2^{2k}$$

because the next larger dyadic interval contains a point  $x$  where the maximal function  $Mf(x)$  is less than  $2^k$ . Multiplying by  $|I|$  and summing over all maximal dyadic subintervals shows the claim.

To show the lemma for  $r < 2$  we need to invoke interpolation of vector valued functions. Let  $p$  be any collection of pairwise disjoint tiles. For each function  $g$  on  $\mathbb{R}$  we can form the vector valued function  $G$  on  $\mathbb{R}$  with values in  $\mathbf{R}^p$  as follows:

$$G(x) = (\Pi_p g(x))_{p \in \mathbf{p}}.$$

The  $L^s(l^t)$ -norm of the function  $G$  is defined as

$$\|(G(x))_{p \in \mathbf{p}}\|_{L^s(l^t)} := \left( \int \left( \sum_{p \in \mathbf{p}'} |G(x)|^t \right)^{s/t} dx \right)^{1/s}.$$

We intend to use interpolation for the linear operator which maps  $g$  to  $G$ . The endpoints of the interpolation are the obvious bounds

$$\|(\Pi_p g(x))_{p \in \mathbf{P}}\|_{L^2(l^2)}^2 = \sum_{p \in \mathbf{P}} \|\Pi_p g(x)\|_2^2 \leq \|g\|_2^2$$

and, for some small  $\varepsilon > 0$ ,

$$\begin{aligned} \|(\Pi_p g(x))_{p \in \mathbf{P}}\|_{L^{1+\varepsilon}(l^\infty)}^{1+\varepsilon} &= \int \left( \sup_{p \in \mathbf{P}} |\Pi_p g(x)| \right)^{1+\varepsilon} dx \\ &\leq C \int M g(x)^{1+\varepsilon} dx \leq C \|g\|_{1+\varepsilon}^{1+\varepsilon}. \end{aligned}$$

Complex interpolation [1] gives for  $1 < r < 2$ :

$$\|(\Pi_p g(x))_{p \in \mathbf{P}}\|_{L^r(l^{r'+\varepsilon})} \leq C \|g\|_r$$

with a new  $\varepsilon > 0$ , which can be made arbitrarily small by appropriate choice of the previous  $\varepsilon$ .

Now we apply this estimate to the function  $f$  and the set of tiles  $\mathbf{p}_k$ . Together with the observation

$$\begin{aligned} \|N_{\mathbf{p}_k}^{1/(r'+\varepsilon)}\|_r^r &= \int N_{\mathbf{p}_k}(x)^{r/(r'+\varepsilon)} dx \\ &\leq 2^{-rk} \int \left( \sum_{p \in \mathbf{p}_k} |\Pi_p f(x)|^{r'+\varepsilon} \right)^{r/(r'+\varepsilon)} dx \end{aligned}$$

this finishes the proof of the lemma.

**Corollary 2.11.** *The Carleson measure estimates of Lemma 2.10 are also true, if the sets  $\mathbf{p}_k$  are replaced by sets  $\mathbf{P}_k$  of pairwise disjoint bitiles (quartiles).*

We prove the corollary for bitiles, the case of quartiles being similar. If  $P$  is a bitile such that  $\|\Pi_P f\|_\infty \geq 2^k$ , then we can pick  $p \in \{l_P, u_P\}$  such that  $\|\Pi_p(2f)\|_\infty \geq 2^k$ . We pick for each bitile in the set  $\mathbf{P}_k$  such a tile and call this set of tiles  $\mathbf{p}_k$ . Applying the lemma to the sets  $\mathbf{p}_k$  gives the corollary.

### 3. THE QUARTILE OPERATOR

**3.1. Introduction.** In this section  $\mathbf{P}$  denotes the set of all quartiles, i.e., all dyadic rectangles of area four, in the half plane  $\mathbb{R} \times [0, \infty)$ . Each quartile  $P$  is the union of four tiles  $a_P$ ,  $b_P$ ,  $c_P$ , and  $d_P$ , which have the same time interval as the quartile.

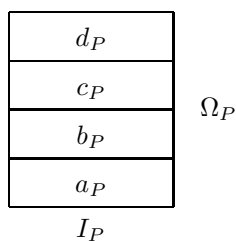


FIGURE 1. Subdivision of quartiles



The quartile operator can be written as

$$(3.1) \quad H_W(f, g) = \sum_{P \in \mathbf{P}} \frac{1}{\sqrt{|I_P|}} (w_{a_P}, f)(w_{b_P}, g) w_{c_P}.$$

The particular choice of three out of four tiles in this definition is inessential for our purpose. For example the operator

$$H'_W(f, g) = \sum_{P \in \mathbf{P}} \frac{1}{\sqrt{|I_P|}} (w_{d_P}, f)(w_{c_P}, g) w_{a_P}$$

can be estimated in the same way. It is important though that the three quartiles in this definition are different, since otherwise the operator was unbounded. For example one easily shows that for  $f$  and  $g$  equal to the characteristic function of  $[0, 1)$  and each  $x \in [0, 1)$  the formal sum

$$H''_W(f, g)(x) = \sum_{P \in \mathbf{P}} \frac{1}{\sqrt{|I_P|}} (w_{a_P}, f)(w_{a_P}, g) w_{b_P}(x)$$

has infinitely many summands of modulus 1.

Since we assume that  $f$  and  $g$  are constant on small dyadic intervals, only finitely many summands in the definition (3.1) are nonzero. Indeed, the Walsh transform of  $g$  is compactly supported and there are only finitely many quartiles  $P$  for which the frequency support of  $b_P$  intersects the frequency support of  $g$ .

We will prove Theorem 1.2 in two steps. The second step in fact suffices to prove the whole theorem, but the first one shows more clearly the symmetry of the bilinear operator and we therefore present it here.

The symmetry is expressed by pairing the quartile operator with a third function and thereby passing to the trilinear form

$$(3.2) \quad T(f_1, f_2, f_3) = \sum_{P \in \mathbf{P}} \frac{1}{\sqrt{|I_P|}} (w_{a_P}, f_1)(w_{b_P}, f_2)(w_{c_P}, f_3).$$

We will show the bound

$$|T(f_1, f_2, f_3)| \leq C_{r_1, r_2, r_3} \|f_1\|_{r_1} \|f_2\|_{r_2} \|f_3\|_{r_3}$$

for  $2 < r_1, r_2, r_3 < \infty$  and  $1/r_1 + 1/r_2 + 1/r_3 = 1$ . By duality this proves the Theorem 1.2 in the case  $1 < r < 2$  and  $2 < s, t < \infty$ .

**3.2. The bound on the trilinear form.** We will actually show

$$(3.3) \quad \sum_{P \in \mathbf{P}} \frac{1}{\sqrt{|I_P|}} |(w_{a_P}, f_1)(w_{b_P}, f_2)(w_{c_P}, f_3)| \leq C_{r_1, r_2, r_3} \|f_1\|_{r_1} \|f_2\|_{r_2} \|f_3\|_{r_3}.$$

Fix exponents  $r_1, r_2, r_3$  as above and dyadic step functions  $f_1, f_2, f_3$ . Without loss of generality we can assume that

$$\|f_1\|_{r_1} = \|f_2\|_{r_2} = \|f_3\|_{r_3} = 1.$$

We organize the set of quartiles into a hierarchy of subsets. For this we use the previously defined partial order of the set of quartiles:

$$P \leq P' \quad \text{if } P \cap P' \neq \emptyset \quad \text{and} \quad I_P \subset I_{P'}.$$

Define the *density* of the quartile  $P$  with respect to a function  $f$  as

$$(3.4) \quad \delta(P, f) := \sup_{P' \geq P} \|\Pi_{P'} f\|_\infty,$$

where  $\Pi_P$  denotes as before the orthogonal projection onto the space spanned by  $w_{a_P}$ ,  $w_{b_P}$ ,  $w_{c_P}$ , and  $w_{d_P}$ . Define for all integers  $k$  the sets

$$(3.5) \quad \mathbf{P}_k(f) := \{P \in \mathbf{P} : 2^k \leq \delta(P, f) < 2^{k+1}\}$$

and the sets

$$\mathbf{P}_{k_1, k_2, k_3} = \bigcap_{i=1,2,3} \mathbf{P}_{k_i}(f_i).$$

The sets  $\mathbf{P}_{k_1, k_2, k_3}$  form a partition of the set of all quartiles for which none of the three densities vanish, i.e., all quartiles which contribute to the value of the sum in (3.3). Recall that there are only finitely many such quartiles.

Let  $\mathbf{P}_{k_1, k_2, k_3}^{\max}$  denote the set of maximal quartiles in  $\mathbf{P}_{k_1, k_2, k_3}$ . We enumerate the quartiles in  $\mathbf{P}_{k_1, k_2, k_3}^{\max}$  as  $P_{k_1, k_2, k_3}^1$ ,  $P_{k_1, k_2, k_3}^2$ , and so on, and we define  $\mathbf{P}_{k_1, k_2, k_3}^i$  to be the set of quartiles in  $\mathbf{P}_{k_1, k_2, k_3}$  which are less than  $P_{k_1, k_2, k_3}^i$  but not less than any of the quartiles  $P_{k_1, k_2, k_3}^{i'}$  with  $i' < i$ . The sets  $\mathbf{P}_{k_1, k_2, k_3}^i$  are called *trees*, and the maximal quartiles  $P_{k_1, k_2, k_3}^i$  *tree tops*. Observe that the trees form a partition of  $\mathbf{P}_{k_1, k_2, k_3}$ .

We prove the following estimate for a single tree  $\mathbf{P}_{k_1, k_2, k_3}^i$ :

$$(3.6) \quad \left| \sum_{P \in \mathbf{P}_{k_1, k_2, k_3}^i} \frac{1}{\sqrt{|I_P|}} (w_{a_P}, f_1)(w_{b_P}, f_2)(w_{c_P}, f_3) \right| \leq 2^{k_1+k_2+k_3} |I_{P_{k_1, k_2, k_3}^i}|.$$

To simplify notation we write  $\mathbf{Q} := \mathbf{P}_{k_1, k_2, k_3}^i$  for the tree and  $Q := P_{k_1, k_2, k_3}^i$  for the tree top. Fix a frequency  $\xi$  in the frequency interval of the tree top and split the tree into four subsets,  $\mathbf{Q}_a$ ,  $\mathbf{Q}_b$ ,  $\mathbf{Q}_c$ , and  $\mathbf{Q}_d$ , which we will refer to as subtrees, where  $\mathbf{Q}_a$  contains all quartiles  $P$  in  $\mathbf{Q}$  such that  $\xi$  lies in the frequency interval of  $a_P$ , etc. We prove the estimate only for  $\mathbf{Q}_a$ , the other cases are done similarly.

The first observation is that the tiles  $b_P$  and  $b_{P'}$  are disjoint for any two different quartiles  $P$  and  $P'$  in  $\mathbf{Q}_a$ . This follows from Lemma 2.1 and the fact that if the time intervals of  $P$  and  $P'$  have nonempty intersection, then the tiles  $a_P$  and  $a_{P'}$  have nonempty intersection.

We conclude that as  $P$  runs through the quartiles of  $\mathbf{Q}_a$ , the wave packets  $w_{b_P}$  form an orthonormal set. The same is true for the tiles  $c_P$ . Therefore we can apply the Cauchy-Schwarz inequality to obtain

$$(3.7) \quad \left| \sum_{P \in \mathbf{Q}} \frac{(w_{a_P}, f_1)}{\sqrt{|I_P|}} (w_{b_P}, f_2)(w_{c_P}, f_3) \right| \leq 2^{k_1} \sum_{P \in \mathbf{Q}} |(w_{b_P}, f_2)| |(w_{c_P}, f_3)| \leq 2^{k_1} \|f_2\|_2 \|f_3\|_2.$$

The second observation is that we can replace in this estimate the functions  $f_2$  and  $f_3$  by their projections onto the subspace corresponding to  $A_{\mathbf{Q}} = \bigcup_{P \in \mathbf{Q}} P$ . This projection  $\Pi_{A_{\mathbf{Q}}}$  is well defined by Lemma 2.5, because by construction the tree  $\mathbf{Q}$  is a convex set of quartiles. The observation is obvious since by Corollary 2.8 the left-hand side of (3.7) does not change under this replacement.

We estimate the norms of the functions  $\Pi_{A_{\mathbf{Q}}} f_i$  for  $i = 2, 3$ . Clearly these functions are supported on  $I_Q$ . Furthermore, if  $x \in I_Q$ , and if  $P$  is the minimal quartile of the tree such that  $x \in I_P$ , then  $\Pi_{A_{\mathbf{Q}}} f_i(x) = \Pi_P f_i(x)$ , because the difference

of the two projections is easily seen to be supported outside  $x$ . But  $\Pi_P f_i(x)$  is bounded by  $2^{k_i}$  by construction of the tree, hence  $\Pi_{A_Q} f_i$  is uniformly bounded by  $2^{k_i}$ . Now we have the estimate

$$2^{k_1} \|\Pi_{A_Q} f_2\|_2 \|\Pi_{A_Q} f_3\|_2 \leq 2^{k_1+k_2+k_3} |I_Q|,$$

which finishes the proof of the tree estimate (3.6).

Next we prove for  $j = 1, 2, 3$ :

$$(3.8) \quad \sum_{P' \in \mathbf{P}_{k_1, k_2, k_3}^{\max}} |I_{P'}| \leq \sum_{P \in \mathbf{P}_{k_j}^{\max}(f_j)} |I_P|.$$

Here  $\mathbf{P}_{k_j}^{\max}(f_j)$  is the set of maximal quartiles in the set  $\mathbf{P}_{k_j}(f_j)$  defined above. Fix a quartile  $P$  in  $\mathbf{P}_{k_j}^{\max}(f_j)$  and consider the collection  $\mathbf{P}_P$  of quartiles in  $\mathbf{P}_{k_1, k_2, k_3}^{\max}$  which are less than  $P$ . The quartiles in  $\mathbf{P}_P$  are pairwise disjoint since they are all maximal in  $\mathbf{P}_{k_1, k_2, k_3}$ . But the frequency intervals of all these quartiles contain the frequency interval of  $P$ , hence the time intervals of the quartiles in  $\mathbf{P}_P$  are pairwise disjoint. This implies the estimate.

$$\sum_{P' \in \mathbf{P}_P} |I_{P'}| \leq |I_P|.$$

If we sum these inequalities for all  $P \in \mathbf{P}_{k_j}^{\max}(f_j)$ , each quartile in  $\mathbf{P}_{k_1, k_2, k_3}^{\max}$  occurs at least once on the left hand side, because it is majorized by at least one quartile in  $\mathbf{P}_{k_j}^{\max}(f_j)$ . This finishes the proof of (3.8).

Now it is easy to finish the proof of (3.3). We have

$$\begin{aligned} & \sum_{P \in \mathbf{P}} \frac{1}{\sqrt{|I_P|}} |(w_{a_P}, f_1)(w_{b_P}, f_2)(w_{c_P}, f_3)| \\ & \leq \sum_{k_1, k_2, k_3} 2^{k_1+k_2+k_3} \sum_i |I_{P_{k_1, k_2, k_3}^i}|. \end{aligned}$$

To proceed with the calculation we pick for each triple  $k_1, k_2, k_3$  the index  $j \in \{1, 2, 3\}$  for which  $r_j k_j$  is maximal and use estimate (3.8) for the choice of  $j$ . Thus

$$\dots \leq \sum_{j=1}^3 \left( \sum_{k_1, k_2, k_3: r_j k_j \geq r_l k_l, l \neq j} 2^{k_1+k_2+k_3} \sum_{P \in \mathbf{P}_{k_j}^{\max}(f_j)} |I_P| \right).$$

For fixed  $j$  we sum first over  $k_l$  with  $l \neq j$ . This gives a geometric series, the dominant term being  $\frac{r_l}{r_j} k_j$ . Using that the sum of reciprocals of the  $r_j$  is equal to 1, we can estimate the previously displayed expression by

$$\dots \leq C \sum_{j=1}^3 \left( \sum_{k_j} 2^{r_j k_j} \sum_{P \in \mathbf{P}_{k_j}^{\max}(f_j)} |I_P| \right).$$

Now the latter sum is bounded by Lemma 2.10, since the set  $\mathbf{P}_{k_j}^{\max}(f_j)$  is a set of pairwise disjoint quartiles. This finishes the proof of (3.3).

**3.3. The decomposition of the bilinear operator.** To prove the remaining estimates of the theorem we will argue directly for the quartile operator and prove the weak type estimates:

$$(3.9) \quad |\{x : |H_W(f_1, f_2)(x)| > \lambda\}| \leq C \frac{(\|f_1\|_{r_1} \|f_2\|_{r_2})^2}{\lambda^r}$$

with  $1/r = 1/r_1 + 1/r_2$ ,  $1 < r_1 < 2$ ,  $2 < r_2 < \infty$ . By real interpolation we obtain the corresponding strong type inequalities [7]. By symmetry we also have the estimates with the conditions on  $r_1$  and  $r_2$  interchanged, and complex interpolation with these estimates [3] gives the strong type estimates with  $1/r = 1/r_1 + 1/r_2$ ,  $1 < r_1$ ,  $r_2 < \infty$ ,  $2/3 < r < 2$ . To obtain the remaining exponents we observe that fixing one argument of the bilinear operator and considering the adjoint operator for the remaining linear operator gives operators which are of the same form as  $H_W$ . Thus by duality we obtain more estimates, in fact we obtain all remaining estimates of the theorem in this way.

We proceed with the proof of (3.9). Fix exponents  $r, r_1, r_2$  and functions  $f_1, f_2$ . By linearity and scaling we can assume that

$$\|f_1\|_{r_1} = \|f_2\|_{r_2} = \lambda = 1.$$

We consider  $H_W(f_1, g)$  as a linear operator in the second argument  $g$ . For simplicity we write  $f := f_1$ .

We consider density of quartiles as before (3.4), but only with respect to the function  $f$ . To simplify notation we write

$$\mathbf{P}_k := \mathbf{P}_k(f).$$

The maximal quartiles in  $\mathbf{P}_k$  are denoted by  $\mathbf{P}_k^{\max}$ . For each quartile in  $\mathbf{P}_k$  we count how many quartiles of  $\mathbf{P}_k^{\max}$  are greater than the quartile. This leads to the following definition: for each nonnegative integer  $i$ ,

$$\mathbf{P}_{k,i} = \{P \in \mathbf{P}_k : 2^i \leq |\{P' \in \mathbf{P}_k^{\max} : P' \geq P\}| < 2^{i+1}\}.$$

The sets  $\mathbf{P}_{k,i}$  are called forests.

Define the *forest operator*  $T_{\mathbf{P}_{k,i}}$  as

$$T_{\mathbf{P}_{k,i}}g := \sum_{P \in \mathbf{P}_{k,i}} \frac{1}{|I_P|} (w_{a_P}, f)(w_{b_P}, g) w_{c_P}.$$

Then the basic decomposition of the quartile operator is

$$H(f, g) = \sum_{k=-\infty}^0 \sum_{i=0}^{\infty} T_{\mathbf{P}_{k,i}}(g).$$

In the next section we will prove the following bound of the forest operator:

$$(3.10) \quad \|T_{\mathbf{P}_{k,i}}g\|_2 \leq C2^k \|g\|_2.$$

For now we take this estimate for granted and proceed to bound the quartile operator. Fix a number  $m > r'_2$ . Let  $A$  be the set of all forests  $\mathbf{P}_{k,i}$  such that  $k < 0$  and  $i < m|k|$ , and let  $B$  be the set of all other forests.

For a time  $x$  to be contained in the time interval of a quartile of the forest  $\mathbf{P}_{k,i}$ , it must be contained in the time interval of at least  $2^i$  quartiles of  $\mathbf{P}_k^{\max}$ . Hence, if

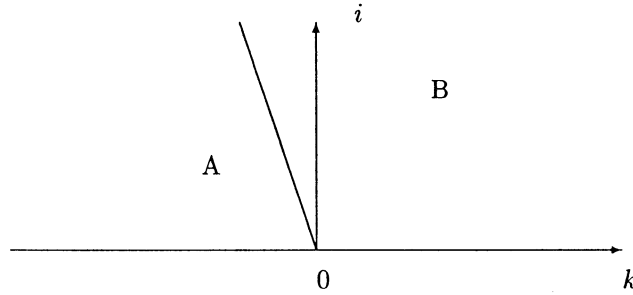


FIGURE 2. Decomposition of the set of forests

$\mathbf{P}_{k,i}$  is a forest in  $B$ , then the support of  $T_{\mathbf{P}_{k,i}}g$  is a subset of

$$E_k = \begin{cases} \{x : N_k(x) \geq 2^{m|k|}\} & \text{if } k \leq 0, \\ \{x : N_k(x) \geq 1\} & \text{if } k > 0. \end{cases}$$

Here  $N_k$  is the counting function of  $\mathbf{P}_k^{\max}$ :

$$N_k := \sum_{P \in \mathbf{P}_k^{\max}} 1_{I_P}.$$

The size of the set  $E_k$  is controlled by the Carleson measure estimate of Lemma 2.10:

$$|E_k| \leq \begin{cases} C2^{-r_1 k} 2^{r_1 m|k|(1/r_1' - \varepsilon)} & \text{if } k \geq 0, \\ C2^{-r_1 k} & \text{if } k < 0. \end{cases}$$

The number  $m$  is chosen such that these estimates are summable over  $k$ . We obtain for  $E = \bigcup_{k=-\infty}^{\infty} E_k$  the bound

$$|E| \leq C.$$

Outside the set  $E$  we only have to consider the forests in  $A$ . Let  $\mathbf{P}_{k,i}$  be such a forest, then the time intervals of all quartiles in the forest are contained in the set  $E_k^0 := \{x : N_k(x) \geq 1\}$ . Applying the forest operator to  $g$  gives the same as applying it to the restriction of  $g$  to the set  $E_k^0$ . For this restriction we have, using the Carleson measure estimate again:

$$\|f_2 1_{E_k^0}\|_2 \leq |E_k^0|^{1/2-1/r_2} \leq 2^{-r_1 k(1/2-1/r_2)}.$$

Now we use (3.10) to obtain for the quartile operator outside the set  $E$ :

$$\begin{aligned} \|H_W(f_1, f_2) 1_{E^c}\|_2 &\leq \left\| \sum_{\mathbf{P}_{k,i} \in A} T_{\mathbf{P}_{k,i}} f_2 \right\|_2 \\ &\leq \sum_{k \leq 0} C \|f_2 1_{E_k^0}\|_2 2^k m|k| \\ &\leq \sum_{k \leq 0} C 2^{-r_1 k(1/2-1/r_2)} 2^k m|k|. \end{aligned}$$

Now since  $1/r_1 + 1/r_2 > 1/2$ , this is a convergent series.

This gives the weak type estimate

$$|\{H_W(f_1, f_2)(x) > 1\}| \leq \|H_W(f_1, f_2)1_{E^c}\|_2^2 + |E| \leq C.$$

This finishes the proof of Theorem 1.2 up to the forest estimate.

**3.4. The forest estimate.** We decompose the forests further. For each quartile  $Q$  in  $\mathbf{P}_{k,i}^{\max}$  we consider the set of all quartiles in  $\mathbf{P}_{k,i}$  which are less than  $Q$  and call this set the *tree* with *tree top*  $Q$ .

Clearly the union of the trees in the forest gives back the forest. It is a beautiful observation of Fefferman in [5] that two such trees are disjoint and, as a consequence of that, cover disjoint subsets of the phase plane. For a proof of disjointness of the trees assume to the contrary that a quartile  $P \in \mathbf{P}_{k,i}$  lies below two maximal quartiles  $Q$  and  $Q'$ . This implies that  $Q$  and  $Q'$  have disjoint frequency intervals. By definition of  $\mathbf{P}_{k,i}^{\max}$  the quartiles  $Q$  and  $Q'$  are less than at least  $2i$  maximal quartiles in  $\mathbf{P}_k$ , and by the disjointness of frequency intervals these two collections of maximal quartiles are disjoint. Hence  $P$  is less than at least  $2^{k+1}$  maximal quartiles in  $\mathbf{Q}_k$ , a contradiction to  $P \in \mathbf{P}_{k,i}$ .

We conclude that whenever two quartiles in a forest are comparable, they are both comparable to a common maximal quartile, and therefore they are in the same tree. This means that quartiles from different trees are incomparable. This implies that the *tree operators*

$$\sum_{P \in \mathbf{P}_{k,i}, P \leq Q} \frac{1}{|I_P|} (w_{a_P}, f)(w_{b_P}, g) w_{c_P}$$

are mutually orthogonal as  $Q$  runs through  $\mathbf{P}_{k,i}^{\max}$ . Here we have used that the trees are convex and therefore cover good subsets of the phase plane. Therefore it suffices to prove the desired  $L^2$  estimate for a single tree operator.

We decompose as before the tree  $\mathbf{Q}$  into four subtrees  $\mathbf{Q}_a$ ,  $\mathbf{Q}_b$ ,  $\mathbf{Q}_c$ , and  $\mathbf{Q}_d$  by fixing a frequency in the frequency interval of each tree top. The bound for the trees  $\mathbf{Q}_a$  and  $\mathbf{Q}_d$  can be done in the same way as we did the tree estimate for the trilinear form. The adjoint of the tree operator of  $\mathbf{Q}_c$  is essentially of the same form as the tree operator of  $\mathbf{Q}_b$ , hence by analogy we are done once we have shown the bound for the tree operator of  $\mathbf{Q}_b$ .

As in the proof of the tree estimate for the trilinear form we can replace  $f$  by its projection  $\Pi_{A_Q} f$  onto the space defined by the area  $A_Q$  of the tree, and again  $\Pi_{A_Q} f$  is pointwise bounded by  $C2^k$ .

We calculate, using that the tiles  $c_P$  run over a disjoint collection of tiles

$$\begin{aligned} & \left\| \sum_{P \in \mathbf{Q}_b} \frac{1}{\sqrt{|I_P|}} (w_{a_P}, f)(w_{b_P}, g) w_{c_P} \right\|_2 \\ &= \sqrt{\sum_{P \in \mathbf{Q}_b} \frac{1}{|I_P|} |(w_{a_P}, f)|^2 |(w_{b_P}, g)|^2}. \end{aligned}$$

Now we view  $\mathbf{Q}_b$  as a measure space, where each quartile has the measure

$$\mu(\{P\}) = |(w_{a_P}, f)|^2.$$

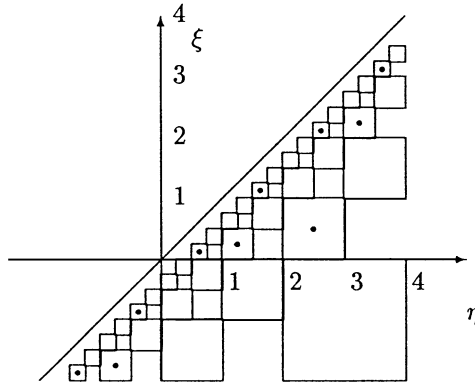


FIGURE 3. Whitney decomposition of the lower triangle

Then the previous expression can be written as

$$\sqrt{\int_0^\infty \mu \left( \left\{ P : \frac{1}{|I_P|} |(w_{b_P}, g)|^2 \geq \lambda \right\} \right) d\lambda}.$$

We estimate the measure of the set inside the integral. The time intervals of the quartiles occurring in the set are obviously contained in the set

$$E := \{x : Mg(x)^2 \geq \lambda\},$$

where  $M$  is the Hardy Littlewood maximal function. The tiles  $a_P$  are pairwise disjoint, hence, by Bessel's inequality, the measure of the above set is bounded by the square of the  $L^2$ -norm of the function  $\Pi_{A_Q} f$ , restricted to the set  $E$ . This norm is bounded by  $C2^k \sqrt{|E|}$ . We continue the above sequence of inequalities, using the maximal theorem, to get

$$\begin{aligned} \sqrt{\int_0^\infty C2^{2k} |\{x : Mg(x)^2 \geq \lambda\}| d\lambda} \\ = C2^k \|Mg\|_2 \leq C2^k \|g\|_2. \end{aligned}$$

This proves the bound for the forest operator.

### 3.5. The quartile operator as a model for the bilinear Hilbert transform.

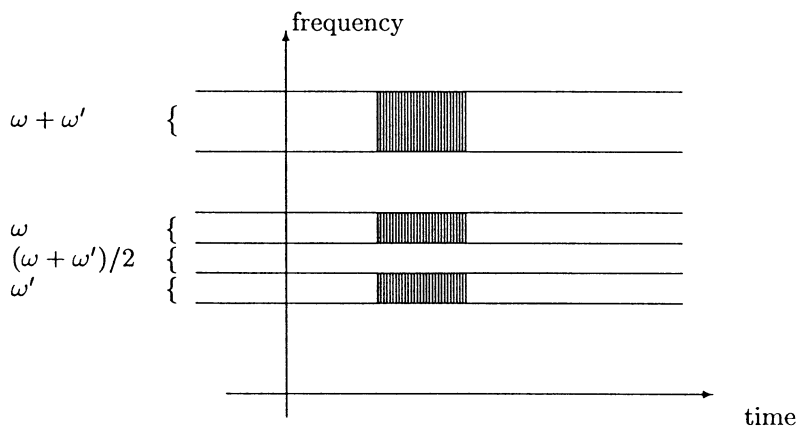
This section gives some heuristic arguments and therefore is not meant to be rigorous. A formal calculation using the Fourier transform gives for the bilinear Hilbert transform

$$\begin{aligned} H(f, g)(x) &:= p.v. \int_{\mathbb{R}} f(x-t)g(x+t) \frac{dt}{t} \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}} \hat{f}(\eta) \hat{g}(\xi) e^{i(\eta+\xi)x} (-i) \operatorname{sgn}(\eta - \xi) d\eta d\xi. \end{aligned}$$

Consider the  $\eta, \xi$  plane and a Whitney decomposition into squares of the area below the diagonal  $\eta = \xi$ , as in Figure 3.

This Whitney decomposition is chosen such that for each dyadic interval

$$\omega = [2^j n, 2^j(n+1))$$

FIGURE 4. Time-frequency diagram for  $(f * 1_\omega^\vee)(x)(g * 1_{\omega'}^\vee)$ 

there is a square  $\omega \times \omega'$  in the decomposition with

$$\omega' = [2^j(n-2), 2^j(n-1)).$$

We fix the notation that  $\omega'$  denotes the dyadic interval which is of the same size as  $\omega$  and two steps below  $\omega$ . Let  $\Omega$  be the set of dyadic intervals of the form

$$\left[2^j \left(n + \frac{2}{4}\right), 2^j \left(n + \frac{3}{4}\right)\right)$$

with integers  $j$  and  $n$ , i.e. those intervals which occur as frequency intervals of third tiles  $c_P$  of quartiles  $P$ . The squares  $\omega \times \omega'$  with  $\omega \in \Omega$  are marked with a dot in Figure 3.

We consider the following modified bilinear operator, where  $1_\omega$  denotes the characteristic function of the interval  $\omega$  and  $1_\omega^\vee$  its inverse Fourier transform.

$$\sum_{\omega \in \Omega} \int_R \int_R \hat{f}(\eta) \hat{g}(\xi) e^{i(\eta+\xi)x} 1_\omega(\eta) 1_{\omega'}(\xi) d\eta d\xi = \sum_{\omega \in \Omega} (f * 1_\omega^\vee)(x)(g * 1_{\omega'}^\vee)(x)$$

To recover the bilinear Hilbert transform from this expression, we have to take a linear combination with the corresponding sum for the area above the diagonal and to average over translated and dilated dyadic grids in order to get rid of the gaps we have left over in the Whitney decomposition. Therefore boundedness properties of the modified operator will translate into boundedness properties of the bilinear Hilbert transform.

The expression  $f * f_\omega^\vee$  can be understood as taking the frequency band  $\omega$  of the function  $f$ . In the Walsh model this corresponds to

$$\sum_{I: \text{dyadic}, |I|=1} (f, w_{I \times \omega}) w_{I \times \omega}.$$

In the continuous case, the frequency band  $\omega$  of  $f$  is multiplied to the frequency band  $\omega'$  of  $g$  to get a function in the frequency band  $\omega + \omega'$ . This situation is shown in Figure 4. In contrast to that, the product



$$\begin{aligned}
 (3.11) \quad & \sum_{I:\text{dyadic}, |I||\omega|=1} (f, w_{I \times \omega}) w_{I \times \omega} \sum_{I:\text{dyadic}, |I||\omega|=1} (g, w_{I \times \omega'}) w_{I \times \omega'} \\
 &= \sum_{I:\text{dyadic}, |I||\omega|=1} \frac{1}{\sqrt{|I|}} (f, w_{I \times \omega}) (g, w_{I \times \omega'}) w_{I \times \{\lfloor \omega \rfloor, 2\lfloor \omega \rfloor\}}
 \end{aligned}$$

gives a frequency band which depends only on  $I$ . This does not reflect the orthogonality properties of the different frequency bands arising in the continuous case. Thus we abandon the pointwise product in the Walsh case and replace (3.11) by

$$(3.12) \quad \sum_{I:\text{dyadic}, |I||\omega|=1} \frac{1}{\sqrt{|I|}} (f, w_{I \times \omega}) (g, w_{I \times \omega'}) w_{I \times ((\omega + \omega')/2)}.$$

Here the factor  $\frac{1}{2}$  has been introduced to simplify the geometric picture. It does not effect orthogonality properties, since the intervals  $(\omega + \omega')/2$  and  $(\tilde{\omega} + \tilde{\omega}')/2$  are disjoint if and only if the intervals  $\omega + \omega'$  and  $\tilde{\omega} + \tilde{\omega}'$  are disjoint.

The summands in (3.12) are the same as those of the quartile operator except for a permutation of the three frequency bands, which in view of the symmetry under taking adjoints is inessential to the theory. Hence summing (3.12) over all dyadic  $\omega$  gives the quartile operator.

#### 4. A.E. CONVERGENCE OF WALSH FOURIER SERIES

**4.1. The partial sum operator in the phase plane.** For a dyadic step function  $f$  on  $[0, 1)$  consider the maximal partial sum operator

$$S^{\max} f(x) = \sup_n \left| \sum_{k=0}^n (f, W_k) W_k(x) \right|,$$

where  $W_k$  are the Walsh functions defined in the introduction. We will prove the weak type estimate

$$|\{x : S^{\max} f(x) > \lambda\}| \leq C \frac{\|f\|_r^r}{\lambda^r}$$

for  $1 < r \leq 2$  and a constant  $C$  depending only on  $r$ . It is well known that this estimate implies Theorem 1.1, [12].

Fix a dyadic step function  $n : [0, 1) \mapsto \mathbb{N}_0$ . Define the linearized maximal partial sum operator as

$$S^n(x) := \sum_{k=0}^{n(x)} (f, W_k) W_k(x).$$

To prove the above weak type estimate it suffices to show

$$|\{x : S^n f(x) > \lambda\}| \leq C \frac{\|f\|_r^r}{\lambda^r}$$

with a constant  $C$  independent of the function  $n$ .

We decompose the linearized maximal partial sum operator into a sum of operators associated to *bitiles*. Throughout this section  $\mathbf{P}$  will denote the set of bitiles. Recall that each bitile  $P$  is the union of a lower son  $l_P$  and an upper son  $u_P$ . Define for each bitile  $P$  the operator  $T_P$  by

$$T_P f(x) = \begin{cases} (f, w_{l_P}) w_{l_P}(x) & \text{if } (x, n(x)) \in u_P, \\ 0 & \text{if } (x, n(x)) \notin u_P. \end{cases}$$

Since  $n$  is defined only on  $[0, 1)$ , this definition so far makes sense only for  $I_P \subset [0, 1)$ . However, later we will extend  $n$  to all of  $\mathbb{R}$ . We claim that

$$(4.1) \quad S^n f(x) = \sum_{P \in \mathbf{P}: I_P \subset [0, 1)} T_P f(x).$$

Observe that since  $n$  is bounded there are only finitely many summands on the right hand side of the equation which are nonzero.

Fix an  $x_0 \in [0, 1)$ . Then the rectangle  $R = [0, 1) \times [0, n(x_0))$  defines a projection  $\Pi_R$ . Since this rectangle is the disjoint union of tiles of the form  $[0, 1) \times [k, k+1)$ , and the wavepackets corresponding to these tiles are the Walsh functions  $W_k$ , we have

$$S^n f(x_0) = \Pi_R f(x_0).$$

Now we partition the rectangle  $R$  into a different collection of tiles. For each point  $(x, \xi)$  in the rectangle there is a unique bitile whose lower son contains  $(x, \xi)$  and whose upper son contains  $(x, n(x_0))$ . Indeed, the frequency interval  $\omega$  of this bitile is the minimal dyadic interval which contains  $\xi$  and  $n(x_0)$ , and the time interval of this bitile is the dyadic interval  $I$  which contains  $x_0$  and satisfies  $|I| |\omega| = 2$ . The rectangle  $R$  is therefore the disjoint union of the lower sons of all bitiles whose upper son intersects the set  $[0, 1) \times \{n(x_0)\}$ . Hence

$$\Pi_R f = \sum_{P: I_P \subset [0, 1), u_P \cap ([0, 1) \times \{n(x_0)\}) \neq \emptyset} (f, w_{l_p}) w_{l_p}.$$

Evaluating this expression at  $x_0$  we can restrict the sum to those bitiles, whose time interval contains  $x_0$ . Hence

$$\Pi_R f(x_0) = \sum_{P: I_P \subset [0, 1), (x_0, n(x_0)) \in u_P} (f, w_{l_p}) w_{l_p}(x_0).$$

This proves the claim.

**4.2. The bound of the partial sum operator.** Instead of considering (4.1), we let  $n$  be a dyadic step function from  $\mathbb{R}$  to  $\mathbb{N}_0$  and prove a bound on

$$\sum_{P \in \mathbf{P}} T_P f(x).$$

Here the sum is over all bitiles, which allows for scaling. Clearly this will imply the bound on (4.1).

Hence we fix  $n, r$ , and  $f$  with  $\|f\|_r = 1$ , and we have to show

$$\left| \left\{ x : \sum_{P \in \mathbf{P}} T_P f(x) \geq 1 \right\} \right| \leq C.$$

We organize the set of bitiles into a hierarchy of subsets, as we have done before for the quartile operator. For any subset  $\mathbf{Q} \subset \mathbf{P}$  we write  $T_{\mathbf{Q}}$  for the operator  $\sum_{P \in \mathbf{Q}} T_P$ .

Fix a finite dyadic step function. Recall the partial order of the set of bitiles:

$$P \leq P' \quad \text{if } P \cap P \cap P' \neq \emptyset \quad \text{and} \quad I_P \subset I_{P'}.$$

Define the density of a bitile  $P$  by

$$\delta(P) := \sup_{P' \geq P} \|\Pi_{P'} f\|_{\infty}.$$

Define

$$\mathbf{P}_k := \mathbf{P}_k(f) = \{P \in \mathbf{P} : 2^k \leq \delta(P) < 2^{k+1}\},$$

and

$$\mathbf{P}_{k,i} := \{P \in \mathbf{P}_k : 2^i \leq |\{P' \in \mathbf{P}_k^{\max} : P' \geq P\}| < 2^{i+1}\}.$$

As before we call the sets  $\mathbf{P}_{k,i}$  forests.

With that we simply proceed as before and split the set of forests into the set  $A$  of all forests  $\mathbf{P}_{k,i}$  such that  $k < 0$  and  $i < m|k|$  and the set  $B$  of all other forests, with  $m$  large enough. The forests in  $B$  are supported in the set  $E = \bigcup_K E_k$ , where

$$E_k := \begin{cases} \{x : N_k(x) \geq 2^{m|k|}\} & \text{if } k \leq 0, \\ \{x : N_k(x) \geq 1\} & \text{if } k > 0, \end{cases}$$

and we have the bound

$$|E| \leq C.$$

Let  $\mathbf{P}_{k,i}$  be a forest in the set  $A$ . We need an estimate of  $T_{\mathbf{P}_{k,i}}f$  on the complement of  $E$ , where as before we have defined

$$T_{\mathbf{Q}} := \sum_{P \in \mathbf{Q}} T_P$$

for any set  $\mathbf{Q}$  of bitiles. Define

$$\mathbf{Q}_{k,i} := \{P \in \mathbf{P}_{k,i} : I_P \notin E\}.$$

In the next section we will prove for  $1 < s < \infty$  the following estimate:

$$(4.2) \quad \|T_{\mathbf{Q}_{k,i}}f\|_s \leq C2^k \left( \sum_{P \in \mathbf{Q}_{k,i}^{\max}} |I_P| \right)^{1/s}.$$

Using the Carleson measure estimates of Lemma 2.10 we obtain

$$\begin{aligned} \sum_{P \in \mathbf{Q}_{k,i}^{\max}} |I_P| &\leq \int_R \min(N_k(x), 2^i) dx \\ &\leq \int_{\mathbf{P}} (N_k(x))^{\frac{r}{r'+\varepsilon}} (2^{m|k|})^{1-\frac{r}{r'+\varepsilon}} dx \\ &\leq C2^{-rk}2^{m|k|(1-\frac{r}{r'+\varepsilon})}. \end{aligned}$$

We substitute this into the forest estimate and sum over all forests in the set  $A$  to get

$$\|T_{\mathbf{P}}f1_{E^c}\|_s \leq C \sum_{k=-\infty}^0 m|k|2^k \left( 2^{-rk}2^{m|k|(1-\frac{r}{r'+\varepsilon})} \right)^{1/s}.$$

For large  $s$  this is a summable series and gives

$$\|T_{\mathbf{P}}f1_{E^c}\|_s \leq C.$$

This finishes the proof except for estimate (4.2).

**4.3. The forest estimate.** We have seen before that each forest  $\mathbf{P}_{k,i}$  is the union of convex trees such that two different trees cover subsets of the Walsh phase plane which are disjoint. The same is true for the sets  $\mathbf{Q}_{k,i}$ , and it is easy to see that we have

$$\sum_{P \in \mathbf{Q}_{k,i}^{\max}} 1_{I_P} \leq \sum_{P \in \mathbf{P}_{k,i}^{\max}} 1_{I_P}.$$

If  $\mathbf{Q}$  is a tree, then the support of the function  $T_{\mathbf{Q}}f$  is contained in the set of all  $x$  such that  $(x, n(x))$  is contained in  $\bigcup_{Q \in \mathbf{Q}} Q$ . Hence for two different trees  $\mathbf{Q}$  and  $\mathbf{Q}'$  in the forest  $\mathbf{Q}_{k,i}$  we conclude that  $T_{\mathbf{Q}}f$  and  $T_{\mathbf{Q}'}f$  are supported on disjoint sets. It therefore suffices to show for  $1 < s < \infty$ :

$$\|T_{\mathbf{Q}_{k,i}}f\|_s \leq C2^k |I_Q|$$

for any convex tree  $\mathbf{Q} \subset \mathbf{P}_{k,i}$  with tree top  $Q$ .

Let the frequency interval of the tree top be  $[\xi, \xi']$ . We split the tree  $\mathbf{Q}$  into two subtrees  $\mathbf{Q}_l$  and  $\mathbf{Q}_u$ , the first one containing all bitiles of the tree such that the frequency interval of the lower son contains  $\xi$ , and the second one containing all bitiles of the tree such that the frequency interval of the upper son contains  $\xi$ . It suffices to show the estimate for both subtrees separately.

Consider first  $\mathbf{Q}_l$ . Then the upper sons of the bitiles in  $\mathbf{Q}_l$  are pairwise disjoint by Lemma 2.1. That means for a given  $x$  there is at most one bitile in  $\mathbf{Q}_l$  whose upper son contains  $(x, n(x))$ . Hence  $T_{\mathbf{Q}_l}f(x)$  is given by  $\Pi_{I_P}f(x)$ , which is bounded by  $2^k$  because the projection  $\Pi_P f(x)$  is bounded by  $2^k$ , by definition of  $\mathbf{P}_{k,i}$ .

Hence  $T_{\mathbf{Q}_l}f(x)$  is pointwise bounded by  $2^k$ . Since it is supported on the time interval of the tree top, this proves the desired estimate for  $\mathbf{Q}_l$ .

Now consider the subtree  $\mathbf{Q}_u$ . Let  $I = I_Q$  be the time interval of the tree top  $Q$  of  $\mathbf{Q}$ . All frequency intervals we will mention in this paragraph have the property that their Cartesian product with  $I$  is a rectangle which is a disjoint union of tiles and therefore defines a projection. We will show below that for each  $x$  there is an interval  $\omega(x)$ , which is not necessarily dyadic, such that

$$T_{\mathbf{Q}_u}f(x) = \Pi_{I \times \omega(x)}f(x).$$

Moreover we will show that this interval can be written as

$$(4.3) \quad \omega(x) = (\omega_1(x) \setminus \omega_2(x)) \cap \omega,$$

where  $\omega_1(x)$  and  $\omega_2(x)$  are dyadic intervals and  $\omega = [0, \xi)$  with  $\xi$  as above. By partitioning the whole strip  $I \times [0, \infty)$  into tiles with time interval  $I$ , we see that this gives the operator identity

$$\Pi_{I \times \omega(x)} = (\Pi_{I \times \omega_1(x)} - \Pi_{I \times \omega_2(x)})\Pi_{I \times \omega}.$$

The operator  $\Pi_{I \times \omega_1(x)}$  is pointwise bounded by the maximal operator. This is seen from partitioning the rectangle  $I \times \omega_1(x)$  into tiles, which have frequency interval  $\omega_1(x)$ . Then at each point the projection operator coincides with the projection onto a single tile, which clearly is bounded by the maximal operator. The same is true for  $\Pi_{I \times \omega_2(x)}$ . This gives

$$|T_{\mathbf{Q}_u}f(x)| \leq |\Pi_{I \times \omega(x)}f(x)| \leq 2M\Pi_{I \times \omega}f(x).$$

The operator  $\Pi_{I \times \omega}$  is nothing but the ordinary partial sum operator

$$\sum_{k=0}^n (f, W_k) W_k$$

with a certain  $n$ , rescaled to the interval  $I$ . This one can see from partitioning the rectangle into tiles with time interval  $I$ . By Paley's theorem [10] this operator is bounded in  $L^r$  for  $1 < r < \infty$  with a bound not depending on  $n$ . This and the maximal theorem applied to the previous inequality gives the desired estimate for the tree operator.

It remains to show the existence of the frequency interval  $\omega(x)$  with the stated properties. If no bitile in  $\mathbf{Q}_u$  contains the point  $(x, n(x))$ , then there is nothing to prove because then the operator vanishes at  $x$ . Otherwise pick  $\omega_1(x)$  to be the frequency interval of the minimal bitile in  $\mathbf{Q}_u$  which contains the point  $(x, n(x))$ , and let  $\omega_2(x)$  be the frequency interval of the upper son of the maximal bitile in  $\mathbf{Q}_u$  which contains the point  $(x, n(x))$ . By equation (4.3) this determines the interval  $\omega(x)$ .

Let  $P$  be any bitile in  $\mathbf{Q}_u$  which contributes to  $T_{Q_u} f(x)$ , i.e., whose upper son contains  $(x, n(x))$ . We show that the lower son  $l_P$  is contained in the rectangle  $I \times \omega$ . The time interval of  $l_P$  is certainly contained in the time interval  $I$  of the tree top. As to the frequency interval of  $l_P$ , we observe that it lies below the frequency  $\xi$  by definition of  $\mathbf{Q}_u$ , and hence is contained in  $\omega$ . Furthermore, it is contained in  $\omega_1(x)$ , because of the extremality of the latter. But it is disjoint from  $\omega_2(x)$ , because by extremality the interval  $\omega_2(x)$  must be contained in the frequency interval of  $u_P$ , which is certainly disjoint from the frequency interval of  $l_P$ .

By Lemma 2.1 the lower sons of the bitiles in  $\mathbf{Q}_u$  whose upper sons contain  $(x, n(x))$  are disjoint. To show that the projection onto the union of these tiles coincides at the point  $x$  with the projection onto the rectangle  $I \times \omega(x)$ , it remains to show that those tiles cover each point of the line  $\{x\} \times \omega(x)$ , because then the difference between the two projections is supported outside  $x$ .

So let  $(x, \xi'')$  be a point on this line. Then there is a unique bitile whose upper son contains  $(x, n(x))$  and whose lower son contains  $(x, \xi'')$ . By reversing the arguments above we see that it is squeezed between two bitiles of  $\mathbf{Q}_u$  w.r.t. the partial order of bitiles. By convexity of  $\mathbf{Q}$  the bitile is in  $\mathbf{Q}$ . Moreover the frequency interval of its upper son contains  $\omega_2(x)$  and hence contains the point  $(x, \xi)$ . Hence the bitile is in  $\mathbf{Q}_u$ . This finishes the proof of the forest estimate.

## REFERENCES

1. J. Bergh and J. Löfström, *Interpolation spaces, an introduction*, Springer-Verlag, Berlin, Heidelberg, New York, 1976. MR **58**:2349
2. P. Billard, *Sur la convergence presque partout des series de Fourier Walsh des fonctions de l'espace  $L^2(0, 1)$* , Studia Math. **28** (1967), 363–388. MR **36**:599
3. A. P. Calderon and A. Zygmund, *A note on the interpolation of linear operators*, Studia Math. **12** (1951), 194–204. MR **13**:754e
4. L. Carleson, *On convergence and growth of partial sums of Fourier series*, Acta Math. **116** (1966), 135–157. MR **33**:7774
5. C. Fefferman, *Pointwise convergence of Fourier series*, Ann. of Math. **98** (1973), 551–571. MR **49**:5676
6. R. Hunt, *Almost everywhere convergence of Walsh-Fourier series of  $L^2$ -functions*, Actes, Congrès Internat. Math., vol. 2, 1970, pp. 655–661. MR **58**:23330
7. S. Janson, *On interpolation of multi-linear operators*, (Proceedings Lund, 1986), Springer Lecture Notes in Math., vol. 1302, 1988. MR **89g**:46114

8. M. Lacey and C. Thiele,  $L^p$  estimates on the bilinear Hilbert transform for  $2 < p < \infty$ , Ann. of Math. (2) **146** (1997), 693–724. MR **99b**:42014
9. M. Lacey and C. Thiele, *Bounds for the bilinear Hilbert transform on  $L^p$* , Proc. Nat. Acad. Sci. **94** (1997), 33–35.
10. R.E.A.C. Paley, *A remarkable series of orthogonal functions* (I), Proc. London Math. Soc. **34** (1932), 241–264.
11. P. Sjölin, *An inequality of Paley and convergence a.e. of Walsh-Fourier series*, Ark. Mat. **7** (1968), 551–570. MR **39**:3222
12. E. Stein, *Singular Integrals and Differentiability Properties of Functions*, Princeton University Press, Princeton, N.J., 1970. MR **44**:7280
13. C. M. Thiele and L. F. Villemoes, *A fast algorithm for adapted time frequency tilings*, Appl. Comput. Harmon. Anal. **3** (1996), 91–99. MR **98g**:65134
14. J. L. Walsh, *A closed set of normal orthogonal functions*, Amer. J. Math. **45** (1923), 5–24.

DEPARTMENT OF MATHEMATICS, YALE UNIVERSITY, NEW HAVEN, CONNECTICUT 06511

*Current address:* Department of Mathematics, University of California, Los Angeles, California 90095-1555

*E-mail address:* thiele@math.ucla.edu