

# Cancellation for the simplex Hilbert transform

PAVEL ZORIN-KRANICH

We show that the truncated simplex Hilbert transform enjoys some cancellation in the sense that its norm grows sublinearly in the number of scales retained in the truncation. This extends the recent result by Tao on cancellation for the multilinear Hilbert transform. Our main tool is the Hilbert space regularity lemma due to Gowers, which enables a very short proof.

## 1. Introduction

Consider the  $(n + 1)$ -linear form defined on functions of  $n$  variables

$$(1.1) \quad \Lambda_K(F_0, \dots, F_n) := \int_{\mathbb{R}^{n+1}} \prod_{i=0}^n F_i(x_{(i)}) K \left( \sum_{i=0}^n x_i \right) dx,$$

where  $x_{(i)} = (x_0, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$  denotes the omission of the  $i$ -th coordinate. We are interested in  $K$  being a smooth truncation of a one-dimensional Calderón–Zygmund kernel. Recall that a one-dimensional Calderón–Zygmund kernel is a distribution  $\hat{K}$  that satisfies  $\|\hat{K}\|_\infty \leq 1$  and that coincides, away from 0, with a differentiable function  $K$  such that  $|K(t)| \leq |t|^{-1}$  and  $|K'(t)| \leq |t|^{-2}$ . Our truncations have the form

$$\psi_{\mathbf{S}} = \sum_{k \in \mathbf{S}} \psi_k, \quad \text{where } \psi_k(t) = \phi(2^{-k}t)K(t),$$

$\mathbf{S} \subset \mathbb{Z}$  is an interval, and  $\phi$  is an even, smooth function supported on  $\pm[1, 4]$  such that  $\sum_{k \in \mathbb{Z}} \phi(2^{-k}t) = 1$  for all  $t \neq 0$ . We call the form  $\Lambda_{\psi_{\mathbf{S}}} := \Lambda_{\psi_{\mathbf{S}}}$  the *truncated simplex Hilbert transform*, in analogy to the truncated triangular Hilbert transform, to which this form specializes for  $n = 2$ ,  $K(t) = \frac{1}{t}$ . The eponymous simplex is spanned by the set  $\{0, \dots, n\}$ . Each function  $F_i$  is

---

*2010 Mathematics Subject Classification:* 42B20 (Primary); 47A07 (Secondary).

associated to a side of the simplex and accepts the variables whose indices span that side.

Since  $\|\psi_k\|_1 = O(1)$  and by Hölder’s inequality the estimate

$$(1.2) \quad |\Lambda_{\mathbf{S}}(F_0, \dots, F_n)| \lesssim_n |\mathbf{S}| \prod_{i=0}^n \|F_i\|_{p_i}$$

is immediate for any Hölder tuple of exponents  $1 \leq p_i \leq \infty$ . Our main result, extending a recent result by Tao [4], is the following qualitative improvement over this bound.

**Theorem 1.3.** *Let  $n \geq 1$ . Then for any  $1 < p_i < \infty$  with  $\sum_{i=0}^n p_i^{-1} = 1$  we have*

$$|\Lambda_{\mathbf{S}}(F_0, \dots, F_n)| \leq o_{n,p_0,\dots,p_n}(|\mathbf{S}|) \prod_{i=0}^n \|F_i\|_{p_i}.$$

The restriction to Hölder tuples of exponents is necessary; more in general, if the kernel  $K$  is homogeneous of degree  $d$ , then the form (1.1) can only be bounded on  $L^{p_0} \times \dots \times L^{p_n}$  if  $\sum_{i=0}^n p_i^{-1} = 1 + (1 + d)/n$ . Uniform bounds in  $\mathbf{S}$  remain out of reach of current methods unless  $n = 1$ , in which case  $\Lambda_{\mathbf{S}}$  reduces to (the dual of) a usual truncation of  $K$ .

An appropriate choice of the functions  $F_i$  (see Appendix A) shows that Theorem 1.3 yields some cancellation for the maximally modulated by polynomials of degree  $d$ , linearized,  $(n - d)$ -linear, truncated Calderón–Zygmund operator

$$(1.4) \quad \begin{aligned} & C_{N_1,\dots,N_d}(f_{d+1}, \dots, f_n)(x) \\ &= \text{p.v.} \int_{\mathbb{R}} e^{i \sum_{j=1}^d N_j(x)t^j} \prod_{j=d+1}^n f_j(x - b_j t) \psi_{\mathbf{S}}(t) dt, \end{aligned}$$

where  $b_{d+1}, \dots, b_n$  are distinct non-zero real numbers and  $N_1, \dots, N_d : \mathbb{R} \rightarrow \mathbb{R}$  are measurable linearizing functions, with constants uniform in the choice of  $b_j$ ’s and  $N_j$ ’s. This answers a question from [4], where such cancellation was obtained for  $d = 0$  and commensurable  $b_j$ ’s.

Our main tool is the Hilbert space regularity lemma due to Gowers [1], whose proof is so short that we chose to write it out in full in Section 2. The multidimensional setup allows us to avoid the much harder arithmetic regularity lemma used in [4].

Theorem 1.3 is proved by induction on  $n$ . The case  $n = 1$  follows from the standard theory of truncated Calderón–Zygmund operators, see e.g. [3, §I.7].

In the inductive step we assume that the theorem holds with  $n > 1$  replaced by  $n - 1$ . Multilinear interpolation with the trivial estimate (1.2) shows that it suffices to consider  $p_0, \dots, p_n > n$  and indicator functions  $F_i = 1_{E_i}$ . We make these assumptions throughout Section 3, which contains a single tree estimate, and Section 4, which describes a standard tree selection algorithm.

## 2. The regularity lemma

The material in this section is almost identical to Gowers’s original exposition in [1]. The only difference from the finite-dimensional case is that it turns out convenient to work with *extended seminorms*, that is, functions  $\|\cdot\|$  on a vector space  $H$  taking values in the extended positive reals  $[0, +\infty]$  that are subadditive, homogeneous, and map 0 to 0 (this observation has been previously used to further streamline [6] Walsh’s proof of the multilinear mean ergodic theorem [5]). The reason is that the atomic seminorms  $\|\cdot\|_\Sigma$ , defined below, are typically extended.

**Lemma 2.1.** *Let  $H$  be a Hilbert space and  $\Sigma \subset H$ . Then the formula*

$$\|f\|_\Sigma := \inf \left\{ \sum_t |\lambda_t| : f = \sum_t \lambda_t \sigma_t, \sigma_t \in \Sigma \right\},$$

where sums are finite (possibly empty), and the infimum of an empty set is by convention  $+\infty$ , defines an extended seminorm on  $H$  whose dual extended seminorm is given by

$$\|f\|_\Sigma^* := \sup_{\phi \in H: \|\phi\|_\Sigma \leq 1} |\langle f, \phi \rangle| = \sup_{\sigma \in \Sigma} |\langle f, \sigma \rangle|.$$

Gowers’s Hilbert space regularity lemma reads as follows.

**Theorem 2.2.** *Let  $\delta > 0$  and  $\eta: \mathbb{R}_+ \rightarrow \mathbb{R}_+$  be any function. Let  $H$  be a Hilbert space with norm  $\|\cdot\|_H$  and let  $\|\cdot\|$  be an arbitrary further extended seminorm on  $H$ . Then for every  $f \in H$  with  $\|f\|_H \leq 1$  there exists  $C = O_{\delta, \eta}(1)$  and a decomposition*

$$(2.3) \quad f = \sigma + u + v$$

such that

$$(2.4) \quad \|\sigma\| < C, \quad \|u\|^* < \eta(C), \quad \text{and} \quad \|v\|_H < \delta.$$

The proof uses the following separation lemma.

**Lemma 2.5.** *Let  $V_i, i = 1, \dots, k$ , be convex subsets of a Hilbert space  $H$ , at least one of which is open, and each of which contains 0. Let  $V := c_1V_1 + \dots + c_kV_k$  with  $c_i > 0$  and take  $f \notin V$ . Then there exists a vector  $\phi \in H$  such that  $\langle f, \phi \rangle \geq 1$  and  $\Re \langle v, \phi \rangle < c_i^{-1}$  for every  $v \in V_i$  and every  $i$ .*

*Proof.* By the assumption the set  $V$  is open, convex and does not contain  $f$ . By the Hahn–Banach theorem there exists a  $\phi \in H$  such that  $\langle f, \phi \rangle \geq 1$  and  $\Re \langle v, \phi \rangle < 1$  for every  $v \in V$ . The claim follows.  $\square$

There is also a constructive version of Lemma 2.5 with an  $\epsilon$  loss, in the sense that the conclusion changes to  $\Re \langle v, \phi \rangle < (1 + \epsilon)c_i^{-1}$  (this version still suffices for our purpose). Indeed, since  $V \ni 0$  is open and  $f \notin V$ , we have  $f \notin (1 + \epsilon)^{-1}\bar{V}$ . Let  $g \in (1 + \epsilon)^{-1}\bar{V}$  be the element that minimizes the distance from  $f$  (such  $g$  is unique). One can then take  $\phi = (f - g) / \langle f - g, f \rangle$ .

*Proof of Theorem 2.2.* Let  $r$  be chosen later (depending only on  $\delta$ ) and define

$$(2.6) \quad C_r = 1, \quad C_{i-1} = \max \left\{ C_i, \frac{2}{\eta(C_i)} \right\}.$$

Let  $V_1, V_2, V_3$  be the open unit balls of  $\|\cdot\|, \|\cdot\|^*$ , and  $\|\cdot\|_H$ , respectively. Suppose that the conclusion fails, then for every  $i \in \{1, \dots, r\}$  we have

$$f \notin C_iV_1 + \eta(C_i)V_2 + \delta V_3.$$

Since  $V_3$  is open in  $H$ , Lemma 2.5 applies, and we obtain vectors  $\phi_i \in H$  such that

$$\langle \phi_i, f \rangle \geq 1, \quad \|\phi_i\|^* \leq (C_i)^{-1}, \quad \|\phi_i\|^{**} \leq \eta(C_i)^{-1}, \quad \|\phi_i\| \leq \delta^{-1}.$$

For every pair  $i < j$  by (2.6) we have

$$|\langle \phi_i, \phi_j \rangle| \leq \|\phi_i\|^* \|\phi_j\|^{**} \leq (C_i)^{-1} \eta(C_j)^{-1} \leq (C_{j-1})^{-1} \eta(C_j)^{-1} \leq \frac{1}{2},$$

so that

$$r^2 \leq \langle \phi_1 + \dots + \phi_r, f \rangle^2 \leq \|\phi_1 + \dots + \phi_r\|^2 \leq r\delta^{-2} + \frac{r^2 - r}{2},$$

which is a contradiction if  $r \geq 2\delta^{-2}$ .  $\square$

### 3. The tree estimate

For each  $k \in \mathbb{Z}$  let  $\mathcal{I}_k$  be the collection of the dyadic cubes  $I \subset \mathbb{R}^{n+1}$  of the form

$$I = I_0 \times \cdots \times I_n = 2^k(m_0, \dots, m_n) + [0, 2^k]^n, \quad \sum_i m_i = 0, \quad m_i \in \mathbb{Z}.$$

The *scale* of a dyadic cube  $I \in \mathcal{I}_k$  is defined as  $s(I) := k$ . Let also  $\mathcal{I}_{\mathbf{S}} := \cup_{k \in \mathbf{S}} \mathcal{I}_k$  and  $\mathcal{I} := \mathcal{I}_{\mathbb{Z}}$ . This gives the splitting

$$(3.1) \quad \Lambda_{\mathbf{S}}(F_0, \dots, F_n) = \sum_{I \in \mathcal{I}_{\mathbf{S}}} \Lambda_I(F_0, \dots, F_n),$$

where for each  $I \in \mathcal{I}_k$  we have set

$$\Lambda_I(F_0, \dots, F_n) := \int_{\mathbb{R}^{n+1}} \prod_{i=0}^n F_i(x_{(i)}) \psi_k \left( \sum x \right) \prod_{i=0}^{n-1} 1_{I_i}(x_i) dx.$$

Contrary to what could be expected, our argument would not benefit from using smoother versions of the cutoffs  $1_{I_i}$ . However, this appears to be a limitation rather than a strength of our approach.

We write elements of  $\mathbb{R}^{n+1}$  as  $x = (x', x_n) \in \mathbb{R}^n \times \mathbb{R}$  and dyadic cubes  $I \in \mathcal{I}$  as  $I' \times I_n$ , where  $I'$  is a dyadic cube in  $\mathbb{R}^n$  and  $I_n$  is a dyadic interval in  $\mathbb{R}$ . A *tree* with top  $J \in \mathcal{I}$  is a collection of boxes  $I \in \mathcal{I}$  such that  $I' \subset J'$ . In this section we obtain a gain over the trivial bound (coming from Fubini's theorem) for the restriction of the sum (3.1) to a tree.

**Proposition 3.2.** *For every  $\delta > 0$  there exists  $S_{\delta,n} \in \mathbb{N}$  such that for any functions  $F_0, \dots, F_n : \mathbb{R}^n \rightarrow [0, 1]$  and for every dyadic cube  $J \in \mathcal{I}$  there exists an interval of scales  $\mathbf{S}' \subset \mathbb{Z}$  with  $|\mathbf{S}'| \leq S_{\delta,n}$  and  $\max \mathbf{S}' = s(J)$  such that*

$$\left| \sum_{k \in \mathbf{S}'} \sum_{I \in \mathcal{I}_k: I' \subset J'} \Lambda_I(F_0, \dots, F_n) \right| \lesssim_n \delta 2^{n s(J)} |\mathbf{S}'|.$$

Note that  $\mathbf{S}'$  depends both on the (bounded) functions  $F_i$  and the dyadic square  $J$ , but  $S_{\delta,n}$  does not.

*Proof of Proposition 3.2.* By scaling we may assume  $s(J) = 0$ . Note that

$$(3.3) \quad \sum_{I \in \mathcal{I}_k: I' \subset J'} \Lambda_I(F_0, \dots, F_n) = \int_{\mathbb{R}^{n+1}} \prod_{i=0}^n F_i(x_{(i)}) \psi_k \left( \sum x \right) \prod_{i=0}^{n-1} 1_{J_i}(x_i) dx$$

for every  $k$  and the integrand is supported on  $10J$ , say.

A *dual function* is a function from  $X := \prod_{i=0}^{n-1} 10J_i$  to  $\mathbb{C}$  of the form

$$x \mapsto \prod_{A \subseteq \{0, \dots, n-1\}} f_A(x|_A),$$

where  $f_A : \prod_{i \in A} 10J_i \rightarrow \mathbb{C}$  are functions bounded by 1. Denote the set of dual functions by  $\Sigma$  and apply Theorem 2.2 with  $H = L^2(X)$ ,  $f = F_n|_X$ , the extended seminorm given by Lemma 2.1 and a function  $\eta$  to be chosen later.

To dispose of the  $L^2$  error term note that at each scale  $k \leq 0$  the right-hand side of (3.3) is bounded by

$$\int_{10J} |F_n(x_{(n)})| |\psi_k \left( \sum x \right)| d\vec{x} \lesssim \|F_n\|_{L^1(10J_{(n)})} \|\psi_k\|_1 \lesssim \|F_n\|_{L^2(10J_{(n)})}.$$

The contribution of the uniform term (bounded in  $\|\cdot\|_{\Sigma}^*$ ) is estimated by

$$\sum_{k \in \mathbf{S}'} \|\hat{\psi}_k\|_1 \left| \int_{10J} \prod_i F_i(x_{(i)}) e \left( \xi_k \sum x \right) \prod_{i=0}^{n-1} 1_{J_i}(x_i) dx \right|$$

for some choice of frequencies  $\xi_k$ . Note that  $\|\hat{\psi}_k\|_1 \lesssim 2^{-k}$ . Inside the absolute value, the character splits into a product of functions depending on one variable each. Since  $n > 1$ , each function that depends on only one coordinate  $x_i$  can be absorbed into one of the functions  $F_i$ ,  $i < n$ . Thus for each fixed  $x_n$  the integral above is a pairing of  $F_n$  with a dual function, and we obtain the estimate

$$\sum_{k \in \mathbf{S}'} 2^{-k} \|F_n\|_{\Sigma}^* \lesssim 2^{|\mathbf{S}'|} \|F_n\|_{\Sigma}^*.$$

It remains to treat the structured term. Suppose  $F_n \in \Sigma$ , so that

$$F_n = \prod_{i=0}^{n-1} f_i,$$

where each function  $f_i$  is bounded by 1 and does not depend on the  $i$ -th coordinate. Substituting this into (3.3) we obtain

$$\int_{10J} \prod_{i=0}^{n-1} (F_i f_i)(x_{(i)}) \psi_{\mathbf{S}'} \left( \sum x \right) \prod_{i=0}^{n-1} 1_{J_i}(x_i) dx.$$

This can be written as

$$\int_{10J_n} \int_{\mathbb{R}^n} \prod_{i=0}^{n-1} \left( 1_{10J} F_i 1_{A_i} \prod_{j \neq i, n} 1_{J_j} \right) (x'_{(i)}, x_n) \psi_{\mathbf{S}'} \left( \sum x' + x_n \right) dx' dx_n.$$

Changing variable in the inner integral and applying the inductive hypothesis (Theorem 1.3 with  $n - 1$  in place of  $n$  and  $p_0 = \dots = p_{n-1} = n$ ) we can bound this by

$$c_{n-1}(|\mathbf{S}'|) |\mathbf{S}'|$$

with a monotonically decreasing function  $c_{n-1}$  such that

$$\lim_{|\mathbf{S}'| \rightarrow \infty} c_{n-1}(|\mathbf{S}'|) = 0.$$

Summing the contributions of the three terms given by Theorem 2.2 we obtain

$$\delta |\mathbf{S}'| + 2^{|\mathbf{S}'|} \eta(C) + C c_{n-1}(|\mathbf{S}'|) |\mathbf{S}'|,$$

where  $C = O_{\delta, \eta}(1)$ . Choose a monotonically increasing function  $\tilde{S}_{\delta, n} : \mathbb{R}_+ \rightarrow \mathbb{N}$  such that  $a c_{n-1}(\tilde{S}_{\delta, n}(a)) \leq \delta$  for all  $a$ . Let  $\eta(a) := \delta \tilde{S}_{\delta, n}(a) 2^{-\tilde{S}_{\delta, n}(a)}$ . Then the claim follows with  $|\mathbf{S}'| = \tilde{S}_{\delta, n}(C)$ .  $\square$

**Corollary 3.4.** *Let  $\delta > 0$  and  $S_{\delta, n}$  be the number from Proposition 3.2. Then for every  $J \in \mathcal{I}$  and every interval  $\mathbf{S}' \subset \mathbb{Z}$  with  $\max \mathbf{S}' = s(J)$ , we have*

$$\left| \sum_{k \in \mathbf{S}'} \sum_{I \in \mathcal{I}_k : I' \subset J'} \Lambda_I(F_0, \dots, F_n) \right| \lesssim_n 2^{ns(J)} (\min(|\mathbf{S}'|, S_{\delta, n}) + \delta \max(|\mathbf{S}'| - S_{\delta, n}, 0))$$

for any functions  $F_0, \dots, F_n$  bounded by 1.

*Proof.* By induction on  $|\mathbf{S}'|$ . For  $|\mathbf{S}'| \leq S_{\delta, n}$  the estimate follows from  $|F_i| \leq 1$  and  $\|\psi_k\|_1 \lesssim 1$ .

If  $|\mathbf{S}'| > S_{\delta,n}$ , then by Proposition 3.2 we can find a final interval  $\mathbf{S}'' \subset \mathbf{S}'$  such that the sum over  $\mathbf{S}''$  can be estimated by  $2^{ns(J)}\delta|\mathbf{S}''|$ . The remaining part of the sum splits into sums over subintervals of scale  $s(J) - |\mathbf{S}''|$ , and to those we apply the Corollary with  $\mathbf{S}' \setminus \mathbf{S}''$  in place of  $\mathbf{S}'$ .  $\square$

### 4. Tree selection

For cubes  $I \in \mathcal{I}_k$  write

$$a_I := 2^{-kn} \Lambda_I(F_0, \dots, F_n)$$

The integrand in the definition of  $\Lambda_I$  vanishes outside  $10I$ , say, and by the Loomis–Whitney inequality we can estimate

$$|a_I| \lesssim \prod_{i=0}^n \min_{\pi_{\Delta} I} M_n F_i,$$

where  $\pi_{\Delta} I$  is the subset of the diagonal  $\Delta = \{x \in \mathbb{R}^{n+1} : \sum x = 0\}$  consisting of the points whose first  $n$  coordinates lie in  $I'$  and  $M_n$  is the maximal function  $M_n F(x) = \sup_{Q \ni x} (|Q|^{-1} \int_Q |F|^n)^{1/n}$ . Raising this to a power  $\alpha$  and summing over the squares  $I$  of a given size we obtain

$$\sum_{I \in \mathcal{I}_k} |a_I|^{\alpha} 2^{kn} \lesssim \int_{\Delta} \prod_{i=0}^n M_n F_i(x_{(i)})^{\alpha}.$$

By Hölder’s inequality this is bounded by

$$\prod_{i=0}^n \|(M_n F_i)^{\alpha}\|_{p_i} = \prod_{i=0}^n \|M_n F_i\|_{\alpha p_i}^{\alpha} \lesssim \prod_i |E_i|^{1/p_i}$$

provided  $\alpha p_i > n$  for all  $i$ . It follows from  $p_i > n$  that

$$\begin{aligned} (4.1) \quad \sum_{k \in \mathbf{S}} \sum_{I: |a_I| < \delta, l(I)=k} |a_I| 2^{nk} &\leq \sum_{k \in \mathbf{S}} \sum_{I: l(I)=2^k} |a_I|^{\alpha} \delta^{1-\alpha} 2^{nk} \\ &\lesssim_{\alpha} \delta^{1-\alpha} |\mathbf{S}| \prod_i |E_i|^{1/p_i} \end{aligned}$$

for every  $\alpha$  sufficiently close to 1 and every  $\delta > 0$ .



Let now  $\mathcal{J} \subset \mathcal{I}$  be the collection of maximal cubes  $J$  with  $|a_J| > \delta$ . The union of these cubes cannot be too large. Indeed, we have

$$\pi_\Delta \bigcup \{J : |a_J| > \delta\} \subset \Delta \cap \left\{ \prod_{i=0}^n M_n 1_{E_i} \gtrsim \delta \right\}.$$

The measure of the latter set is bounded by

$$\delta^{-1} \left\| \prod_{i=0}^n M_n 1_{E_i} \right\|_{L^1(\Delta)} \leq \delta^{-1} \prod_{i=0}^n \|M_n 1_{E_i}\|_{L^{p_i}(\Delta)} \lesssim \delta^{-1} \prod_{i=0}^n |E_i|^{1/p_i},$$

where we have again used  $p_i > n$ . Let  $S = S_{\delta^2, n}$  be the number given by Proposition 3.2 with  $\delta^2$  in place of  $\delta$ . For those  $J \in \mathcal{J}$  with  $s(J) > \min \mathbf{S} + \delta^{-2}S$  by Corollary 3.4 we have

$$\left| \sum_{I \in \mathcal{I}_S: I' \subset J'} \Lambda_I(F_0, \dots, F_n) \right| \lesssim_n 2^{ns(J)} \delta^2 |\mathbf{S}|.$$

In particular,

$$\begin{aligned} & \sum_{J \in \mathcal{J}: s(J) > \min \mathbf{S} + \delta^{-2}S} \left| \sum_{I \in \mathcal{I}_S: I' \subset J'} \Lambda_I(F_0, \dots, F_n) \right| \\ & \lesssim_n \delta^2 |\mathbf{S}| \left| \pi_\Delta \bigcup \{J : |a_J| > \delta\} \right| \lesssim \delta |\mathbf{S}| \prod_i |E_i|^{1/p_i}. \end{aligned}$$

On the other hand, by (4.1) with  $\alpha = 0$  we have

$$\sum_{J \in \mathcal{J}: s(J) \leq \min \mathbf{S} + \delta^{-2}S} \sum_{I \in \mathcal{I}_S: I' \subset J'} |\Lambda_I(F_0, \dots, F_n)| \lesssim \delta^{-2}S \prod_i |E_i|^{1/p_i}.$$

Summing the above contributions we obtain the claim of Theorem 1.3 (in the case of characteristic functions).

### Appendix A. Maximally modulated multilinear Hilbert transform

In this appendix we show how to encode some one-dimensional multilinear operators in the simplex Hilbert transform using the ideas from [2, Appendix B]. Consider the family of multilinear forms

$$(A.1) \quad \Lambda_{\vec{\beta}_0, \dots, \vec{\beta}_n}(F_0, \dots, F_n) := \int_{\mathbb{R}^n} \int_{\mathbb{R}} \prod_{i=0}^n F_i(\vec{x} - \vec{\beta}_i t) K(t) dt d\vec{x},$$

where  $\vec{\beta}_i \in \mathbb{R}^n$  are in general position. The main observation is that

$$(A.2) \quad \|\Lambda_K\|_{L^{p_0} \times \dots \times L^{p_n}} = |\det B|^{1/p_0 + \dots + 1/p_n - 1} \|\Lambda_{\vec{\beta}_0, \dots, \vec{\beta}_n}\|_{L^{p_0} \times \dots \times L^{p_n}},$$

where

$$B := \begin{pmatrix} 1 & \dots & 1 \\ \vec{\beta}_0 & \dots & \vec{\beta}_n \end{pmatrix}.$$

In particular, the norm of (A.1) does not depend on the  $\vec{\beta}_i$ 's provided  $\sum_{i=0}^n 1/p_i = 1$ .

*Proof of (A.2).* Consider the change of variables

$$\begin{pmatrix} t \\ \vec{u} \end{pmatrix} = B\vec{x}, \quad \vec{u} \in \mathbb{R}^n, \vec{x} \in \mathbb{R}^{n+1}.$$

If  $\pi_i : \mathbb{R}^{n+1} \rightarrow \mathbb{R}^n$  denotes omission of the  $i$ -th coordinate, then for arbitrary functions  $F_0, \dots, F_n$  we have

$$\begin{aligned} \Lambda(F_0, \dots, F_n) &= \int_{\mathbb{R}^{n+1}} \prod_{i=0}^n F_i(\pi_i(\vec{x})) K(\sum \vec{x}) d\vec{x} \\ &= |\det B|^{-1} \int_{\mathbb{R} \times \mathbb{R}^n} \prod_{i=0}^n F_i(\pi_i B^{-1}(t, \vec{u})) K(t) dt d\vec{u} \\ &= |\det B|^{-1} \int_{\mathbb{R} \times \mathbb{R}^n} \prod_{i=0}^n \tilde{F}_i(\vec{u} - \vec{\beta}_i t) K(t) dt d\vec{u} \\ &= |\det B|^{-1} \Lambda_{\vec{\beta}_0, \dots, \vec{\beta}_n}(\tilde{F}_0, \dots, \tilde{F}_n), \end{aligned}$$

where

$$(A.3) \quad \tilde{F}_i(\vec{u}) := F_i(\pi_i B^{-1}(0, \vec{u})).$$

Here we have used the fact that

$$\pi_i B^{-1} \begin{pmatrix} 1 \\ \vec{\beta}_i \end{pmatrix} = 0.$$

The important observation is now that

$$\|\tilde{F}_i\|_{p_i} = |\det B|^{1/p_i} \|F_i\|_{p_i}.$$

Indeed, the change of variables in the definition of  $\tilde{F}_i$  is given by the  $n \times n$  submatrix of  $B^{-1}$  obtained by crossing out the first column and the  $i$ -th row.

By Cramer’s rule the determinant of that submatrix equals  $\det B^{-1}$  times the  $(1, i)$ -th entry of  $B$ , up to the sign. Since the latter entry of  $B$  is 1, the determinant of the change of variables is  $\pm(\det B)^{-1}$ . The ratio of  $L^{p_i}$  norms equals the absolute value of the determinant to the power  $-1/p_i$ , as required. The claim now follows after taking a supremum over the  $F_i$ ’s.  $\square$

Now we encode the maximally modulated multilinear Hilbert transform (1.4) in (A.1). Let  $\vec{\beta}_0 \in \mathbb{R}^n$  be the origin,  $\vec{\beta}_j = e_j$  for  $j = 1, \dots, d$ ,  $\vec{\beta}_{d+1} = b_{d+1}e_{d+1}$ , and  $\vec{\beta}_j = b_j e_{d+1} + \epsilon e_j$  for  $j = d + 2, \dots, n$ . The identity

$$\sum_{j=0}^k (-1)^{k-j} \binom{k}{j} j^m = \begin{cases} 0 & \text{for } m = 0, 1, \dots, k - 1, \\ k! & \text{for } m = k \end{cases}$$

can be shown by induction on a positive integer  $k$  and its immediate consequence is

$$\sum_{j=0}^d \sum_{k=\max\{j,1\}}^d (-1)^j \frac{1}{k!} \binom{k}{j} N_k(x_{d+1}) \left( \sum_{l=1}^k l x_l - j t \right)^k = \sum_{k=1}^d N_k(x_{d+1}) t^k.$$

It follows that with

$$\begin{aligned} &F_j(x_1, \dots, x_n) \\ &= g_j(x_{d+1}) \prod_{k=\max\{j,1\}}^d \exp \left( i(-1)^j \frac{1}{k!} \binom{k}{j} N_k(x_{d+1}) \left( \sum_{l=1}^k l x_l \right)^k \right) \end{aligned}$$

for  $j = 0, 1, \dots, d$  and

$$F_j(x_1, \dots, x_n) = f_j(x_{d+1})$$

for  $j = d + 1, \dots, n$  the form (A.1) formally becomes

$$\int g_0 g_1 \cdots g_d C_{N_1, \dots, N_d}(f_{d+1}, \dots, f_n).$$

To be precise we should multiply each  $F_j$  by  $\prod_{j \neq d+1} \phi(x_j)$ , where  $\phi$  is a fixed smooth non-negative cutoff function, and then let  $\epsilon \rightarrow 0$ .

## Acknowledgements

I thank Terence Tao and Vjekoslav Kovač for useful discussions about the triangular Hilbert transform and Christoph Thiele for persistent questions that helped to streamline the argument.

## References

- [1] W. T. Gowers, *Decompositions, approximate structure, transference, and the Hahn-Banach theorem*, Bull. Lond. Math. Soc. **42** (2010), no. 4, 573–606.
- [2] V. Kovač, C. Thiele, and P. Zorin-Kranich, *Dyadic triangular Hilbert transform of two general and one not too general function*, Forum Math. Sigma **3** (2015) e25, 27.
- [3] E. M. Stein, *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*, Vol. 43 of Princeton Mathematical Series, Princeton University Press, Princeton, NJ (1993), ISBN 0-691-03216-5. With the assistance of Timothy S. Murphy, Monographs in Harmonic Analysis, III.
- [4] T. Tao, *Cancellation for the multilinear Hilbert transform*, Collect. Math. **67** (2016), no. 2, 191–206.
- [5] M. N. Walsh, *Norm convergence of nilpotent ergodic averages*, Ann. of Math. (2) **175** (2012), no. 3, 1667–1688.
- [6] P. Zorin-Kranich, *Norm convergence of multiple ergodic averages on amenable groups*, J. Analyse Math. **130** (2016), no. 1, 219–241.

UNIVERSITÄT BONN, MATHEMATISCHES INSTITUT  
ENDENICHER ALLEE 60, 53115 BONN, GERMANY  
*E-mail address*: pzorin@uni-bonn.de

RECEIVED AUGUST 4, 2015