MAXIMAL FUNCTIONS ASSOCIATED TO FILTRATIONS

MICHAEL CHRIST AND ALEXANDER KISELEV

1. Introduction

Let $(X,\mu),(Y,\nu)$ be arbitrary measure spaces. Denote by χ_E the characteristic function of a set E. To any sequence of measurable subsets $\{\mathcal{Y}_n : n \in \mathbb{Z}\}$ of Y and any operator T defined on $L^p(Y)$ can be associated a maximal operator

$$T^*f(x) = \sup_n |T(f\cdot \chi_{\mathcal{Y}_n})(x)|.$$

The operator norm of $T : L^p(Y) \mapsto L^q(X)$ is denoted by $\|T\|_{p,q}$. By a filtration of Y we will mean any sequence of subsets $\mathcal{Y}_n \subset Y$ which are nested: $\mathcal{Y}_n \subset \mathcal{Y}_{n+1}$ for every n.

The purposes of this paper are firstly, to establish a rather general maximal theorem[1], Theorem 1.1, and secondly, to establish two multilinear variants.

Theorem 1.1. *Let $1 \le p,q \le \infty$, and suppose that $T : L^p(Y) \mapsto L^q(X)$ is a bounded linear operator. Then for any filtration $\{\mathcal{Y}_n : n \in \mathbb{Z}\}$ of Y, the maximal operator T^* is bounded from $L^p(Y)$ to $L^q(X)$, provided that $p < q$. Moreover*

$$\|T^*\|_{p,q} \le (1 - 2^{-(p^{-1}-q^{-1})})^{-1}\|T\|_{p,q}.$$

Because the constant appearing in the conclusion is independent of $\{\mathcal{Y}_n\}$, a corresponding result can be deduced for continuum filtrations, indexed by a real variable, as a direct consequence via a limiting argument. The most natural example is where $Y = \mathbb{R}$ and $\mathcal{Y}_n = (-\infty, n]$ for each $n \in \mathbb{R}$.

Theorem 1.2. *Suppose that $p < q$. Let $T : L^p(\mathbb{R}) \mapsto L^q(\mathbb{R})$ be defined by $Tf(x) = \int_{\mathbb{R}} K(x,y)f(y)\,dy$, where K is locally integrable. Define*

$$T^*f(x) = \sup_{s\in\mathbb{R}} \Big|\int_{y<s} K(x,y)f(y)\,dy\Big|. \tag{1.1}$$

Then $\|T^*f\|_p \le (1-2^{-(p^{-1}-q^{-1})})^{-1}\|f\|_q$ for all continuous functions f having bounded support.

A corollary is that $\tilde{T}f(x) = \int_{y<x} K(x,y)f(y)\,dy$ is bounded from L^p to L^q, with operator norm $\le (1 - 2^{-(p^{-1}-q^{-1})})^{-1}\|T\|_{p,q}$. Except in the trivial cases where $p = 1$ or $q = \infty$, the hypothesis $p < q$ is necessary even for the conclusion of this corollary to hold for general operators. Indeed, applying the corollary to the Hilbert transform

Date: January 21, 2000.
Research supported by NSF grants DMS-9970660 and DMS-9801530.

[1]This theorem was stated and proved in less generality in [1].

would yield the absurd conclusion that convolution with the restriction to the negative half axis of t^{-1} preserves L^p.

A classical example is a variant due to Zygmund [9] (cf. [8], p. 257) of theorems of Menshov [4] and Paley [5]. Let $X = Y = \mathbb{R}^1$ with Lebesgue measure, and let T be the Fourier transform. Then[2] $f \mapsto \sup_{s\in\mathbb{R}} |\int_{-\infty}^{s} e^{-iyx} f(y)\,dy|$ maps L^p to $L^{p'}$ for all $1 \le p < 2$.

We were led to this result in studying the generalized (that is, not L^2) eigenfunctions of Schrödinger operators on the real line; both Theorem 1.2 and the technique underlying its proof were used in [1],[2] to analyze an infinite family of multilinear expressions out of which the generalized eigenfunctions are built. A related result formulated earlier by one of us [3] has elements in common with Theorem 1.2 but is more closely connected with interpolation theory.

Our theorems and their proofs apply without change to functions taking values in Banach spaces. This generalization of the corollary to Theorem 1.2 mentioned above has already proved useful in connection with Strichartz-type estimates in work of Tao [7] on nonlinear evolution equations, and of Smith and Sogge [6] on the obstacle problem.

A multilinear variant of Theorem 1.1 is as follows. Let $T : L^p(\mathbb{R}, dx) \mapsto L^q(\Lambda, d\mu(\lambda))$ be a bounded linear operator with a locally integrable distribution kernel $K(\lambda, x)$. Define

(1.2)
$$\mathcal{M}_n^*(f_1, f_2, \dots, f_n)(\lambda) = \sup_{y\le y'\in\mathbb{R}} \left| \int \cdots \int_{y\le x_1\le x_2\le\cdots\le x_n\le y'} \prod_{i=1}^{n} \big(K(\lambda, x_i) f_i(x_i) dx_i\big) \right| .$$

Theorem 1.3. *Suppose that $p < q$ and that $2 \le q$. Then for every $n \ge 1$, $(f_1, \dots, f_n) \mapsto \mathcal{M}_n^*(f_1, \dots f_n)$ maps $\otimes^n L^p(\mathbb{R})$ boundedly to $L^{q/n}(\Lambda)$, with operator norm $\le B^n \|T\|_{p,q}^n$.*

Here B is a finite universal constant. With B^n replaced by some constant depending on n in an unspecified way, this was proved in our earlier work [1], but the proof there gives an inferior bound for large n.

Our next variant demonstrates a very substantial improvement, in the special case when all the functions f_i are taken to be the same. It has at least one concrete application; it serves in a companion paper [2] as the heart of our analysis of the asymptotic behavior of generalized eigenfunctions of one-dimensional Schrödinger operators.

Theorem 1.4. *Suppose that $p < q$ and that $2 \le q$. Then for every $n \ge 1$ and every $f \in L^p(\mathbb{R})$,*

$$\|\mathcal{M}_n^*(f, f, \dots, f)\|_{L^{q/n}(\Lambda)} \le \frac{B^n \|T\|_{p,q}^n \|f\|_{L^p}^n}{\sqrt{n!}} . \tag{1.3}$$

This bound improves that of the preceding theorem by the factor $1/\sqrt{n!}$, which has just the right dependence on n for our intended application [2]. Even the weaker

[2] The case $p = 2$ is of course also valid by virtue of a theorem of Carleson but lies outside the scope of our theorem or those of Menshov, Paley and Zygmund.

version of the conclusion in which y, y' are fixed and the supremum is removed was not previously known to us. Much of the work in this paper is devoted to this numerical factor. No such factor arises in Theorem 1.3; modulo the factor of B^n, the bound stated cannot be improved.

There are also more general versions of Theorem 1.3 and Theorem 1.4 in the same spirit as Theorem 1.1, but their formulations are more awkward and are therefore omitted.

2. Proof of the Maximal Theorem

It is no loss of generality to assume that Y is divisible in the sense that for any measurable $S \subset Y$ and any $t \in [0, \nu(S)]$ there exists a measurable subset S' of S satisfying $\nu(S') = t$. Indeed, divisibility may be achieved by replacing Y by $Y \times [0,1]$, ν by the product of ν with Lebesgue measure, T by $T \circ \pi$ where $\pi f(y) = \int_0^1 f(y,s)\,ds$, and $\mathcal{Y}_n$ by $\mathcal{Y}_n \times [0,1]$. Boundedness of T^* is then implied by boundedness of $(T \circ \pi)^*$.

Denote by $C_{p,q}$ the constant in the Theorem. It suffices to show that for any measurable function $N : X \mapsto \mathbb{Z}$ with finite range, the linear operator defined by $T^{(N)} f(x) = T(f \cdot \chi_{\mathcal{Y}_{N(x)}})(x)$ satisfies

$$\|T^{(N)} f\|_q \leq C_{p,q} \|T\|_{p,q} \tag{2.1}$$

for all $f \in L^p(Y)$ satisfying $\|f\|_p = 1$. Fix any such f. Define a probability measure λ on Y by $\lambda(S) = \int_S |f|^p \, d\nu$. The following lemma will be proved below.

Lemma 2.1. *There exists a collection $\{B_j^m\}$ of measurable subsets of Y, indexed by $m \in \{0, 1, 2, \dots\}$ and $1 \leq j \leq 2^m$, satisfying*

- *For each m, $\{B_j^m : 1 \leq j \leq 2^m\}$ is a partition of Y into disjoint measurable subsets.*
- *Each B_j^m is a union of precisely two sets $B_{j_1}^{m+1}, B_{j_2}^{m+1}$.*
- *$\lambda(B_j^m) = 2^{-m}$ for all m, j.*
- *Each set $\mathcal{Y}_n$ may be decomposed, modulo λ-null sets, as an empty, finite, or countably infinite union*

$$\mathcal{Y}_n = \bigcup_{i \geq 1} B_{j_i}^{m_i} \quad \textit{with} \quad m_1 < m_2 < \dots. \tag{2.2}$$

The decomposition (2.2) may not be unique, but for each n we choose one such decomposition.

Define $A_n = \{x : N(x) = n\}$; these are pairwise disjoint subsets of Y, all but finitely many of which are empty. Define $\mathcal{R}$ to be the set of all triples (m, j, n) such that B_j^m is one of the sets occurring in the chosen decomposition (2.2) of $\mathcal{Y}_n$. Define

$$D_j^m = \bigcup_{n : (m,j,n) \in \mathcal{R}} A_n. \tag{2.3}$$

For each fixed m, the sets D_j^m are disjoint. For if $D_j^m \cap D_i^m \neq \emptyset$, then because the sets A_n are pairwise disjoint, there must exist n such that A_n intersects, and hence is contained in, both D_j^m and D_i^m. Hence B_j^m, B_i^m both occur in the decomposition (2.2) of $\mathcal{Y}_n$. Since the two superscripts m are equal, this is prohibited by the construction unless $i = j$.

Define $f_j^m = f \cdot \chi_{B_j^m}$. Then $f \cdot \chi_{\mathcal{Y}_n} = \sum_{(m,j):(m,j,n)\in\mathcal{R}} f_j^m$, so

$$T^{(N)} f = \sum_n \chi_{A_n} \cdot T(f \cdot \chi_{\mathcal{Y}_n})$$

$$= \sum_n \sum_{(m,j):(m,j,n)\in\mathcal{R}} \chi_{A_n} T(f_j^m) = \sum_m \sum_j \chi_{D_j^m} \cdot T(f_j^m).$$

Fix m. Because the sets D_j^m are disjoint,

$$\|\sum_j \chi_{D_j^m} \cdot T(f_j^m)\|_q^q \le \sum_j \|T f_j^m\|_q^q \le \|T\|_{p,q}^q \sum_j \|f_j^m\|_p^q$$

$$= \|T\|_{p,q}^q 2^{-m(q-p)/p} \sum_j \|f_j^m\|_p^p = \|T\|_{p,q}^q 2^{-m(q-p)/p} \|f\|_p^p = \|T\|_{p,q}^q 2^{-m(q-p)/p}.$$

Taking the qth root and summing over m concludes the proof of Theorem 1.1.

Proof of Lemma. Without loss of generality we may assume that for all n, $\lambda(\mathcal{Y}_n) > 0$ and $\lambda(\mathcal{Y}_{n+1}\backslash\mathcal{Y}_n) > 0$. We will construct a measurable function $\varphi : Y \mapsto [0,1]$ satisfying $\lambda(\varphi^{-1}([0,t))) = t$ for all $t \in [0,1]$ and $\mathcal{Y}_n = \varphi^{-1}([0,\lambda(\mathcal{Y}_n)))$. In particular, $\lambda(\varphi^{-1}\{t\}) = 0$ for all t. Then define

$$B_j^m = \varphi^{-1}([(j-1)2^{-m}, j2^{-m})).$$

To decompose $\mathcal{Y}_n$ as in (2.2), consider for each n a binary expansion $\lambda(\mathcal{Y}_n) = \sum_m r_m 2^{-m}$, with each r_m equal either to 0 or to 1. If there happens to be more than one such expansion, choose one. There is a corresponding decomposition of the interval $[0, \lambda(\mathcal{Y}_n))$ into a union of disjoint, adjacent dyadic intervals, closed on the left and open on the right, of lengths $2^{-m_1}, 2^{-m_2}, \dots$, where the m_i are those indices for which $r_{m_i} = 1$, listed in increasing order. Applying φ^{-1} to each of these intervals yields one of the sets $B_{j_i}^{m_i}$. This is the required decomposition of $\mathcal{Y}_n$.

Define $\varphi(y) = 1$ for all $y \in Y\backslash \cup_n \mathcal{Y}_n$, and $\varphi(y) = 0$ for all $y \in \cap_n \mathcal{Y}_n$. Each remaining point $y \in Y$ belongs to $\mathcal{Y}_n \backslash \mathcal{Y}_{n-1} = \tilde{\mathcal{Y}}_n$ for a unique index n. By repeated application of the divisibility property of (Y, λ), it is possible to define $\varphi : \tilde{\mathcal{Y}}_n \mapsto [\lambda(\mathcal{Y}_{n-1}), \lambda(\mathcal{Y}_n))$ so that $\lambda(\varphi^{-1}([\lambda(\mathcal{Y}_{n-1}), t))) = t - \lambda(\mathcal{Y}_{n-1})$ for every $t \in [\lambda(\mathcal{Y}_{n-1}), \lambda(\mathcal{Y}_n))$. The resulting function $\varphi : Y \mapsto [0,1]$ has both of the required properties. □

Remark. The proof can be recast in the following manner, which will be used in our analysis of multilinear analogues. Let r be any exponent strictly greater than p. Define

$$Gf(x) = \sum_{m=0}^{\infty} \Big(\sum_{j=1}^{2^m} |T(f_j^m)(x)|^r \Big)^{1/r}.$$

Then $T^*f(x) \le Gf(x)$, and $\|Gf\|_{L^q(X)} \le C_{p,q,r}\|T\|_{p,q}\|f\|_{L^p(Y)}$. Indeed, when $q/r \ge 1$, Minkowski's integral inequality gives

$$\begin{aligned}
[\int \Big(\sum_j |T(f_j^m)(x)|^r\Big)^{q/r}\,dx]^{r/q} &\le \sum_j \|T(f_j^m)\|_{L^q(X)}^r \\
&\le \|T\|_{p,q}^r \sum_j \|f_j^m\|_{L^p(Y)}^r \\
&\le \|T\|_{p,q}^r \max_j \|f_j^m\|_{L^p(Y)}^{r-p} \cdot \sum_j \|f_j^m\|_{L^p(Y)}^p \\
&\le \|T\|_{p,q}^r 2^{-m(r-p)/p}\|f\|_p^r.
\end{aligned}$$

Since G decreases as r increases, we conclude that for any $r > p$ there exists $\delta > 0$ such that

$$\|(\sum_j |T(f_j^m)|^r)^{1/r}\|_q \le 2^{-\delta m}\|T\|_{p,q}\|f\|_p.$$

Summing over $m \ge 0$ completes the proof.

As an application we deduce a theorem of Menshov [4]. Let $\{\phi_n : n \ge 1\}$ be an orthonormal subset of $L^2(X)$ for some measure space X. Then for any $p < 2$ and any sequence of coefficients $c = \{c_n\} \in \ell^p$, the partial sums of the series $\sum_n c_n\phi_n$ converge almost everywhere in X.

Proof. Let $Y = \{1, 2, 3, \dots\}$, equipped with counting measure. Let $\mathcal{Y}_n = \{1, 2, 3, \dots, n\} \subset Y$. Define the operator $T(c) = \sum_n c_n\phi_n$, mapping $L^2(Y)$ boundedly to $L^2(X)$. T maps the smaller space $L^p(Y)$ boundedly to $L^2(X)$, as well. The partial sums are $S_N(c) = T(c \cdot \chi_{\mathcal{Y}_N})$. By our theorem, $\sup_N |S_N(c)|$ defines a bounded operator from ℓ^p to $L^2(X)$. Combined with the obvious fact that almost everywhere convergence holds for the dense class of all finitely supported c, this implies almost everywhere convergence for all $c \in \ell^p$. $\square$

Menshov's theorem fails, in general, for $p = 2$. This is a further indication that our results cannot extend to the endpoint $p = q$, and are far too crude to yield Carleson's theorem.

3. Some Numerical Inequalities

The proofs of our multilinear results, Theorems 1.3 and 1.4, are based on an induction on the degree n of multilinearity. In this section we analyze the asymptotic behavior of the solution of a certain numerical recursion, (3.2), which will be used in the proof of Theorem 1.4.

Lemma 3.1. *There exists $\gamma \in \mathbb{R}^+$ such that the numbers c_k defined by*

$$c_k = k^{-k/2}k^{-\gamma} \qquad \text{for all } k \ge 2 \tag{3.1}$$

satisfy for every $k \geq 6$ *the inequalities*

$$c_k y^k + \sum_{j=2}^{k-2} c_j c_{k-j} x^j y^{k-j} + c_k x^k \leq c_k (x^2 + y^2)^{k/2} \qquad \textit{for all } x, y \geq 0. \tag{3.2}$$

For $k = 2, 3$ the inequality holds, if it is interpreted as the trivial assertion $c_k x^k + c_k y^k \leq c_k(x^2 + y^2)^{k/2}$; this holds for any c_2, c_3. Our argument does not work for $k = 4, 5$, but we will show later that this is irrelevant for the main application. The lack of any terms associated to the indices $j = 1, n-1$ in the inequality is not a typographical error.

Proof. Define

$$\beta_{k,j} = \frac{k^{k/2}}{j^{j/2}(k-j)^{(k-j)/2}} \left(\frac{k}{j(k-j)} \right)^{\gamma} \tag{3.3}$$

for $2 \leq j \leq k-2$, and $\beta_{k,0} = \beta_{k,k} = 1$. Then $c_j c_{k-j}/c_k = \beta_{k,j}$.

Assume that $k \geq 6$. Suppose first that k is even. We aim to majorize the left-hand side of (3.2), divided by c_k, by the binomial series $\sum_{l=0}^{k/2} \binom{k/2}{l} (x^2)^l (y^2)^{\frac{k}{2}-l} = (x^2+y^2)^{k/2}$, via a term-by-term comparison, while bearing in mind that the former series has about twice as many terms as the latter.

Suppose that j is even. We write $a \sim b$ to indicate that the ratios $a/b, b/a$ are bounded above by a positive constant, uniformly over all relevant parameters on which a, b depend. In this notation, Stirling's formula is $n! \sim n^{n+\frac{1}{2}} e^{-n}$. The binomial coefficients are thus

$$\begin{aligned} \binom{k/2}{j/2} &\sim (k/2)^{-k/2} (j/2)^{j/2} ((k-j)/2)^{(k-j)/2} \left(\frac{k}{j(k-j)} \right)^{1/2} \\ &= k^{-k/2} j^{j/2} (k-j)^{(k-j)/2} \left(\frac{k}{j(k-j)} \right)^{1/2}. \end{aligned} \tag{3.4}$$

Thus for any preassigned positive constant ϵ,

$$\beta_{k,j} \binom{k/2}{j/2}^{-1} \sim \left(\frac{k}{j(k-j)} \right)^{\gamma - \frac{1}{2}} \leq \epsilon \tag{3.5}$$

uniformly in j, k satisfying $2 \leq j \leq k-2$, provided that γ is chosen to be sufficiently large and $k \geq 5$, because $k/(j(k-j))$ is then bounded uniformly by a constant strictly less than 1. This breaks down for $k = 4, j = 2$, causing the restriction $k \geq 6$ in the statement of the lemma, since our analysis of $k = 5$ involves a reduction to $k = 4$; see below.

For j odd, the above calculations tell us that

(3.6)
$$\begin{aligned}\binom{k/2}{(j-1)/2} &\sim k^{-k/2}(j-1)^{(j-1)/2}(k-j+1)^{(k-j+1)/2}\left(\frac{k}{(j-1)(k-j+1)}\right)^{1/2}\\ &\sim k^{-k/2}j^{(j-1)/2}(k-j)^{(k-j+1)/2}\left(\frac{k}{j(k-j)}\right)^{1/2}\\ &\sim \beta_{k,j}\left(\frac{k}{j(k-j)}\right)^{\frac{1}{2}-\gamma}\left(\frac{k-j}{j}\right)^{1/2}.\end{aligned}$$

Likewise

(3.7)
$$\begin{aligned}\binom{k/2}{(j+1)/2} &\sim k^{-k/2}(j+1)^{(j+1)/2}(k-j-1)^{(k-j-1)/2}\left(\frac{k}{(j+1)(k-j-1)}\right)^{1/2}\\ &\sim \beta_{k,j}\left(\frac{k}{j(k-j)}\right)^{\frac{1}{2}-\gamma}\left(\frac{k-j}{j}\right)^{-1/2}.\end{aligned}$$

For j even we have a direct majorization

$$\beta_{k,j}x^jy^{k-j} \le \epsilon\binom{k/2}{j/2}x^jy^{k-j} = \epsilon\binom{k/2}{j/2}(x^2)^{j/2}(y^2)^{(k-j)/2}. \tag{3.8}$$

For j odd write

$$\begin{aligned}\beta_{k,j}x^jy^{k-j} &\le \tfrac{1}{2}\left(\frac{j}{k-j}\right)^{-1/2}\beta_{k,j}x^{j-1}y^{k-j+1} + \tfrac{1}{2}\left(\frac{j}{k-j}\right)^{+1/2}\beta_{k,j}x^{j+1}y^{k-j-1}\\ &\le \epsilon\binom{k/2}{(j-1)/2}(x^2)^{(j-1)/2}(y^2)^{(k-j+1)/2} + \epsilon\binom{k/2}{(j+1)/2}(x^2)^{(j+1)/2}(y^2)^{(k+j+1)/2}.\end{aligned} \tag{3.9}$$

Thus for even $k \ge 6$, by choosing $\epsilon = 1/3$ we obtain

$$\sum_{j=2}^{k-2}\frac{c_jc_{k-j}}{c_k}x^jy^{k-j} \le \sum_{l=1}^{(k/2)-1}\binom{k/2}{l}(x^2)^l(y^2)^{(k/2)-l}. \tag{3.10}$$

Adding in the two remaining terms $x^k + y^k$ from the left-hand side of the inequality to be proved, we obtain on the right the full binomial series, whose sum equals $(x^2+y^2)^{k/2}$.

The proof for odd k is similar but involves an extra initial step. Start by majorizing x^jy^{k-j} by the sum of appropriate coefficients times $(x^2+y^2)^{-1/2}x^{j+1}y^{k-j}$ and $(x^2+y^2)^{+1/2}x^{j-1}y^{k-j}$, or of $(x^2+y^2)^{-1/2}x^jy^{k+1-j}$ and $(x^2+y^2)^{+1/2}x^jy^{k-1-j}$. The first alternative can be applied when $j-1\ge 2$ and $k-j \le k-1-2$; thus when $j \ge 3$. The second can be applied when $k-j\ge 3$. Since $k > 6$ and $2 \le j \le k-2$, at least one of these two restrictions is satisfied. This reduces matters to the case where k is even, and the above argument can then be applied. The details, including the choice of the coefficients that replace $[j/(k-j)]^{\pm 1/2}$ in the above reasoning, are left to the reader. □

Observe that for any $B \in \mathbb{R}^+$, if c_k is replaced by $\tilde{c}_k = B^k c_k$ for all k, then the conclusion of the lemma remains valid for $\{\tilde{c}_k\}$. Moreover, the same goes if $B \geq 1$ and $\tilde{c}_k = B^k c_k$ for all $k \geq 4$, while $\tilde{c}_k = c_k$ for all $k < 4$. The significance of this is that for B sufficiently large, $\{\tilde{c}_k\}$ will satisfy the inequality (3.2) for all k. In fact, it is easy to check that $B = 2$ is sufficient. Hence the restriction $k \geq 6$ in the preceding lemma can be eliminated.

Lemma 3.2. *Suppose that the coefficients* $\{c_k : k \geq 2\}$ *satisfy* (3.2) *for every* $k \geq 2$. *Let* $b_1 = 1$, *and let* $b_k = A^k c_k$ *for all* $k \geq 2$. *Then for any nonnegative* x_1, x_2, y_1, y_2, *for any* $k \geq 2$,

(3.11)
$$b_k x_2^k + b_1 b_{k-1} x_2^{k-1} y_1 + \sum_{j=2}^{k-2} b_j b_{k-j} x_1^j x_1^{k-j} + b_1 b_{k-1} x_1^{k-1} y_2 + b_k x_1^k \leq b_k(|x| + |y|)^k.$$

Here $|x| = (x_1^2 + x_2^2)^{1/2}$, $|y| = (y_1^2 + y_2^2)^{1/2}$. Note that x_1, x_2 now play the roles filled by x, y in the preceding lemma.

Proof. Divide both sides of the inequality by b_k. Using the preceding lemma to majorize the sum of all terms not involving y_1, y_2, and noting that $b_1 b_{k-1}/b_k = A^{-1} c_{k-1}/c_k$, we find that the resulting left-hand side is

$$\begin{aligned} (3.12) \quad & \leq |x|^k + A^{-1} \frac{c_{k-1}}{c_k} |y| (x_1^{2k-2} + x_2^{2k-2})^{1/2} \\ & \leq |x|^k + C A^{-1} k^{1/2} |y| (x_1^{2k-2} + x_2^{2k-2})^{1/2} \\ & \leq |x|^k + C A^{-1} k^{1/2} |y| \cdot 2|x|^{k-1}, \end{aligned}$$

using the inequality $\sqrt{a+b} \leq \sqrt{a} + \sqrt{b}$. On the other hand, expanding $(|x| + |y|)^k$ in binomial series and retaining only two terms, we obtain a quantity $\geq |x|^k + k|x|^{k-1}|y|$, which dominates the right-hand side provided that $A \geq 2C$. □

4. Multilinear maximal operators on the diagonal

Let $f \in L^1(\mathbb{R})$, and consider a multilinear expression

$$M_n(f) = \int \cdots \int_{x_1 \leq x_2 \leq \cdots \leq x_n} \prod_{i=1}^n f(x_i) dx_i, \tag{4.1}$$

restricted to the diagonal $f_1 = f_2 = \cdots f_n = f$.

Definition. A martingale structure on a subinterval $I \subset \mathbb{R}$ is a collection of subintervals $\{E_j^m : m \geq 0, 1 \leq j \leq 2^m\}$ of I that satisfy the following conditions, modulo endpoints.

- $I = \cup_j E_j^m$ for every m.
- $E_j^m \cap E_{j'}^m = \emptyset$ for every $j < j'$.
- If $j < j'$, $x \in E_j^m$, and $x' \in E_{j'}^m$, then $x < x'$.
- For every m, j, $E_j^m = E_{(2j-1)}^{m+1} \cup E_{2j}^{m+1}$.

Such a martingale structure is said to be adapted to f in L^p if $\int_{E_j^m} |f|^p = 2^{-m} \int_I |f|^p$ for every m, j.

The second condition is actually a consequence of the third.

To any martingale structure $\{E_j^m\}$ on $\mathbb{R}$, we associate the sublinear operator

$$\mathfrak{g}(f) = \sum_{m=1}^{\infty} (\sum_{j=1}^{2^m} |\int_{E_j^m} f|^2)^{1/2}. \tag{4.2}$$

The absolute value signs are outside the integrals; this is essential in the application [2] for which this machinery is designed.

It is obvious that $|M_n(f)| \le \|f\|_1^n / n!$, with equality for nonnegative f. The following variant is related but seemingly less obvious.

Proposition 4.1. *There exists $B < \infty$ such that for any martingale structure on $\mathbb{R}$, for every $f \in L^1$, and for every $n \ge 1$,*

$$|M_n(f)| \le B^n \mathfrak{g}(f)^n / \sqrt{n!}\ . \tag{4.3}$$

No connection is assumed between f and the martingale structure; both are arbitrary.

Proof. Fix n, f. By Stirling's formula, it suffices to prove that $|M_n(f)| \le b_n \mathfrak{g}(f)^n$, where $b_n \sim n^{-n/2} A^n n^{-\gamma}$ are the constants in Lemma 3.2.

Define $f_{m,j} = |\int_{E_j^m} f|$, and introduce the variants

$$\mathfrak{g}_J^M(f) = \sum_{m>M}^{\infty} (\sum_{j:E_j^m \subset E_J^M} f_{m,j}^2)^{1/2}. \tag{4.4}$$

These will be used in conjunction with the fact that for any element E_j^m of the martingale structure, the collection $\{E_J^M : E_J^M \subset E_j^m\}$ is a martingale structure on E_j^m.

Partition the region of integration $\Omega = \{x : x_1 < x_2 < \cdots < x_n\}$ into regions $S_j = \{x \in \Omega : x_j \in E_1^1 \text{ and } x_{j+1} \in E_2^1\}$, with $0 \le j \le n$, interpreting this to mean that $x \in S_0 \Leftrightarrow x_1 \in E_2^1$ and $x \in S_n \Leftrightarrow x_n \in E_1^1$. For $1 \le j \le n-1$,

$$\int \cdots \int_{S_j} \prod_{i=1}^{n} f(x_i) dx_i = M_j(f \cdot \chi_1^1) \cdot M_{n-j}(f \cdot \chi_2^1), \tag{4.5}$$

while for $j = 0$ it equals $M_n(f \cdot \chi_2^1)$ and for $j = n$ it equals $M_n(f \cdot \chi_1^1)$. Since $M_1(f) = \int f$, for $j = 1$ this simplifies to $\int_{E_1^1} f \cdot M_{n-1}(f \cdot \chi_2^1)$, while for $j = n-1$ it becomes $\int_{E_2^1} f \cdot M_{n-1}(f \cdot \chi_1^1)$. This yields the recursive bound

$$\begin{aligned} |M_n(f)| \le {} & |M_n(f \cdot \chi_1^1)| + |\int_{E_2^1} f| \cdot M_{n-1}(f \cdot \chi_1^1)| \\ & + \sum_{j=2}^{n-2} |M_{n-j}(f \cdot \chi_1^1)| \cdot |M_j(f \cdot \chi_2^1)| + |\int_{E_1^1} f| \cdot M_{n-1}(f \cdot \chi_2^1)| + |M_n(f \cdot \chi_2^1)|\ . \end{aligned} \tag{4.6}$$

We proceed by induction on n. Thus for $2 \le j \le n-2$,

$$|M_j(f \cdot \chi_1^1) \cdot M_{n-j}(f \cdot \chi_2^1)| \le b_j b_{n-j} \mathfrak{g}_1^1(f)^j \mathfrak{g}_2^1(f)^{n-j}. \tag{4.7}$$

For $j=1$ we have the simpler upper bound $f_{1,1} b_{n-1} \mathfrak{g}_2^1(f)^{n-1}$, and for $j=n-1$, the bound $f_{1,2} b_{n-1} \mathfrak{g}_1^1(f)^{n-1}$.

For $j=0,n$ we are in the same situation with which we began, except that f has been replaced by its restriction to either half E_r^1 of the space. To handle these terms we proceed first via circular reasoning, majorizing them by the desired quantities $b_n \mathfrak{g}_1^1(f)^n$ and $b_n \mathfrak{g}_2^1(f)^n$ respectively, and will subsequently explain how this can be justified by restructuring the induction argument to eliminate the circularity.

By summing over all $0 \le j \le n$, we obtain from (4.6) and (4.7)

$$\begin{aligned} |M_n(f)| \le & \sum_{j=2}^{n-2} b_j b_{n-j} \mathfrak{g}_1^1(f)^j \mathfrak{g}_2^1(f)^{n-j} \\ & + b_{n-1} f_{1,1} \mathfrak{g}_2^1(f)^{n-1} + b_{n-1} f_{1,2} \mathfrak{g}_1^1(f)^{n-1} + b_n \mathfrak{g}_1^1(f)^n + b_n \mathfrak{g}_2^1(f)^n \ . \end{aligned} \tag{4.8}$$

Setting $x_t = \mathfrak{g}_t^1(f)$, $y_t = f_{1,t}$ for $t=1,2$, we are in the situation of Lemma 3.2, and hence conclude that

$$|M_n(f)| \le b_n \left([f_{1,1}^2 + f_{1,2}^2]^{1/2} + [\mathfrak{g}_1^1(f)^2 + \mathfrak{g}_2^1(f)^2]^{1/2} \right)^n . \tag{4.9}$$

Now for $t=1,2$ we have $\mathfrak{g}_t^1(f) = \sum_{m \ge 2} B_{m,t}$, where $B_{m,t} = (\sum_{j: E_j^m \subset E_t^1} f_{m,j}^2)^{1/2}$. Thus

$$(\mathfrak{g}_1^1(f)^2 + \mathfrak{g}_2^1(f)^2)^{1/2} = \Big((\sum_{m \ge 2} B_{m,1})^2 + (\sum_{m \ge 2} B_{m,2})^2 \Big)^{1/2} \tag{4.10}$$

is the norm of the vector $(\sum_{m \ge 2} B_{m,1} \, , \ \sum_{m \ge 2} B_{m,2}) \in \mathbb{C}^2$, and hence is majorized by the sum of the norms:

$$\le \sum_{m \ge 2} (B_{m,1}^2 + B_{m,2}^2)^{1/2} = \sum_{m \ge 2} (\sum_{j=1}^{2^m} f_{m,j}^2)^{1/2}. \tag{4.11}$$

Therefore

$$[f_{1,1}^2 + f_{1,2}^2]^{1/2} + [\mathfrak{g}_1^1(f)^2 + \mathfrak{g}_2^1(f)^2]^{1/2} \le \mathfrak{g}(f). \tag{4.12}$$

This completes the proof, except for justifying the treatment of the regions S_0, S_n.

For large integers K define Ω_K to be the set of all $x = (x_1, \dots, x_n) \in \Omega$ such that there exist no $j \in \{1, 2, \dots 2^K\}$ and $1 \le i \le n-1$ such that both x_i, x_{i+1} belong to E_j^K. Since $\Omega_K \uparrow \Omega$ as $K \to \infty$, it suffices to prove that $\int \cdots \int_{\Omega_K} \prod_i f(x_i) dx_i$ satisfies the bound desired for $M_n(f)$, for every K. We generalize the setup, allowing $\mathbb{R}$ to be replaced by an arbitrary subinterval, and $\{E_j^m\}$ to be replaced by an arbitrary martingale structure on that subinterval.

The bound is proved by a double induction on K, n, doing induction on n for each K. When $K=1$, $M_n = 0$ for $n \ge 3$, while $M_2 = \int_{E_1^1} f \cdot \int_{E_2^1} f$; thus the induction is well founded.

The above argument still applies; the factor of the characteristic function of Ω_K causes no disturbance. Moreover, for the contributions of S_0, S_n, K is effectively replaced by $K-1$, if one replaces $\mathbb{R}$ by E_t^1 and considers $\{E_j^m \subset E_t^1\}_{m\ge 2}$ to be a martingale structure on E_t^1. Thus the induction hypothesis applies, and yields the bounds desired for the contributions of these two exceptional regions. □

The same reasoning yields a slightly more general result, in which $M_n(f)$ is modified by replacing some of the factors $f(x_i)$ by their complex conjugates. Nothing is changed in the conclusion, nor in its proof, since $\overline{f}_{m,j} \equiv f_{m,j}$.

Proposition 4.1 has a maximal version. Write $M_n(f)(y,y') = M_n(f\chi_{y,y'})$, and define

$$M_n^*(f) = \sup_{y\le y' \in \mathbb{R}} M_n(f)(y,y'). \tag{4.13}$$

To control it, we employ a variant of $\mathfrak{g}(f)$:

$$\tilde{\mathfrak{g}}(f) = \sum_{m\ge 1} m\Big(\sum_{j=1}^{2^m} |\int_{E_j^m} f|^2\Big)^{1/2}. \tag{4.14}$$

Proposition 4.2. *There exists $B<\infty$ such that for every locally integrable function f and every $n\ge 1$,*

$$M_n^*(f) \le B^n \tilde{\mathfrak{g}}(f)^n/\sqrt{n!}\ . \tag{4.15}$$

Proof. We have shown that

$$|M_n(f)(y,y')| = |M_n(f\chi_{[y,y']})| \le B^n \mathfrak{g}(f\chi_{[y,y']})^n/\sqrt{n!} \tag{4.16}$$

for some constant B, so it suffices to show that $\mathfrak{g}(f\chi_I) \le C\tilde{\mathfrak{g}}(f)$, uniformly for all intervals I. The definition of $\mathfrak{g}$ involves summing over m,j, and the summands are all nonnegative. Any martingale interval E_j^m that is contained in I makes the same contribution to $\mathfrak{g}(f)$ as to $\mathfrak{g}(f\chi_I)$, and any E_j^m contained in the complement of I contributes nothing to $\mathfrak{g}(f\chi_I)$, and something ≥ 0 to $\mathfrak{g}(f)$. Thus it suffices to analyze the contributions of those E_j^m that intersect both I and $\mathbb{R}\backslash I$.

For any particular m, there are at most two such values of j. We discuss only those (m,j) for which E_j^m intersects the left endpoint y of I; the same reasoning applies to the right. For the purposes of this discussion, we may consider the E_j^m to be open intervals, since sets of measure zero have no effect on the quantities to be estimated. We must then majorize $\sum_m |\int_{I\cap E_{i(m)}^m} f|$, where $i(m)$ is the unique index such that $y\in E_i^m$ if such an index exists, and where we understand the m-th term of the sum to be zero if no such index exists. As in the proof of Theorem 1.1, $I\cap E_{i(m)}^m$ can be partitioned, modulo sets of measure zero, in a unique way into a union of sets E_J^M, where M ranges over all integers $> m$ and for each such M, at most two indices J arise for such a partition. Indeed, if $\mathbb{R}$ is identified homeomorphically with $(0,1)$ in such a way that $\{E_j^m\}$ become the dyadic intervals of lengths 2^{-m}, then this partition is simply the usual Whitney decomposition of (the interior of) $I\cap E_{i(m)}^m$. A given E_J^M can arise in the partitions of more than one set $E_{i(m)}^m$, but can so arise only for

$m < M$, Hence for any M, the total number of all E_J^M arising in all such partitions, counted according to multiplicity, is $\le CM$. Majorizing $|\int_{I\cap E_{i(m)}^m} f|$ by $\sum |\int_{E_J^M} f|$, where the sum is over all intervals in the partition we arrive at the desired bound. $\square$

We can now prove our diagonal-multilinear maximal theorem, Theorem 1.4.

Proof of Theorem 1.4. Let $\{E_j^m\}$ be a martingale structure on $\mathbb{R}$, adapted to f in L^p. Define

$$\mathfrak{G}(f)(\lambda) = \sum_{m=1}^{\infty} m \cdot \big(\sum_{j=1}^{2^m} |T(f\chi_j^m)(\lambda)|^2 \big)^{1/2}. \tag{4.17}$$

It suffices now to combine two ingredients. Firstly, $f \mapsto \mathfrak{G}(f)$ maps $L^p(\mathbb{R})$ boundedly to $L^q(\Lambda)$, as in the first remark following the proof of Theorem 1.1; the extra factor of m is harmless because there is a favorable factor of $2^{-\delta m}$ for some $\delta(p,q) > 0$ for the operator norm of the m-th summand in the series defining $\mathfrak{G}(f)$. Secondly, by Proposition 4.2, there is a pointwise bound $\mathcal{M}_n^*(f)(\lambda) \le B^n \mathfrak{G}(f)(\lambda)^n/\sqrt{n!}$. $\square$

5. Proof of Theorem 1.3

The proof of Theorem 1.3 is close to that of Theorem 1.4, but some modifications are needed. With the natural change of notation from $M_n^*(f)$ to $M_n^*(f_1,\dots,f_n)$, we wish to show that for all sufficiently large n,

$$M_n^*(f_1,\dots,f_n) \le A^n n^{-\gamma} \prod_{i=1}^n \tilde{\mathfrak{g}}(f_i). \tag{5.1}$$

We will treat only $M_n(f_1,\dots,f_n) = \left| \int\cdots\int_{y\le x_1\le x_2\le\cdots\le x_n\le y'} \prod_{i=1}^n \big(K(\lambda,x_i)f_i(x_i)dx_i\big)\right|$ where y, y' are fixed; the bound for its maximal relative M_n^* then follows as in the proof of Proposition 4.2. As in the proof of Proposition 4.1, the factor $n^{-\gamma}$ is of no consequence but makes the induction argument work more smoothly.

(5.1) in turn follows from the bound $M_n(f_1,\dots,f_n) \le A^n n^{-\gamma} \prod_{i=1}^n \mathfrak{g}(f_i)$. For small n, that bound is a simple consequence of the same recursive procedure used to derive Proposition 4.1; the details are left to the reader.

Proceeding as in the proof of that proposition, we obtain for all large n

$$\begin{aligned} A^{-n}n^{\gamma}M_n(f_1,\dots,f_n) \le{}& \prod_{j=1}^n \mathfrak{g}_1^1(f_j) + \epsilon \prod_{j=1}^{n-1} \mathfrak{g}_1^1(f_j)\cdot (f_n)_{1,2} \\ &+ \epsilon \sum_{i=2}^{n-2} \Big(\prod_{j=i}^{i} \mathfrak{g}_1^1(f_j) \prod_{j=i+1}^{n} \mathfrak{g}_2^1(f_j)\Big) + \epsilon (f_1)_{1,1} \prod_{j=2}^n \mathfrak{g}_2^1(f_j) + \prod_{j=1}^n \mathfrak{g}_2^1(f_j); \end{aligned} \tag{5.2}$$

the first and last terms are obtained via the same circular reasoning as in the earlier proof. And as in the proof of Proposition 4.1, the constant ϵ may be made as small as desired by choosing γ and the constant A to be sufficiently large, provided $n \ge 6$.

To simplify the notation set $x_t^j = \mathfrak{g}_t^1(f_j)$ for $t = 1, 2$, and set $y_t^j = |\int_{E_t^1} f_j|$. Also set $|x^j| = [(x_1^j)^2 + (x_2^j)^2]^{1/2}$, and likewise define $|y^j|$.

In this notation, $A^{-n}n^{\gamma}M_n(f_1,\ldots,f_n)$ is majorized by

$$\prod_{j=1}^{n} x_1^j + \epsilon\prod_{j=1}^{n-1} x_1^j \cdot y_2^n + \epsilon\sum_{i=2}^{n-2}\prod_{j=1}^{n-i} x_1^j \prod_{j=n-i+1}^{n} x_2^j + \epsilon\prod_{j=2}^{n} x_2^j y_1^1 + \prod_{j=1}^{n} x_2^j. \tag{5.3}$$

We focus attention temporarily on the middle term. Consider first the case where n is even. Then after pulling out a common factor of $\epsilon x_1^1 x_1^2 x_2^{n-1} x_2^n$, we are left with

$$\begin{aligned}\Big(x_2^3x_2^4\cdots x_2^{n-2} + x_1^3x_1^4x_2^5x_2^6\cdots x_2^{n-2} + x_1^3x_1^4x_1^5x_1^6x_2^7x_2^8\cdots x_2^{n-2} + \cdots + x_1^3x_1^4\cdots x_1^{n-2}\Big)\\ + \Big(x_1^3x_2^4\cdots x_2^{n-2} + x_1^3x_1^4x_1^5x_2^6\cdots x_2^{n-2} + \cdots + x_1^3\cdots x_1^{n-3}x_2^{n-2}\Big).\end{aligned} \tag{5.4}$$

We claim that each of the two terms in parentheses is $\le \prod_{j=3}^{n-2}|x^j|$. To see this observe that the sum of the first two terms in the first set of parentheses is $(x_1^3x_1^4 + x_2^3x_2^4)x_2^5\cdots x_2^{n-2} \le |x^3|\cdot|x^4|\cdot x_2^5\cdots x_2^{n-2}$, by Cauchy-Schwarz. In all the other terms in the first set of parentheses, majorize $x_t^3x_t^4$ by $|x^3|\cdot|x^4|$. Now there is one fewer term than when we began; the sum of the contribution of the term just obtained by applying Cauchy-Schwarz and the next term in the sum is $\le |x^3|\cdot|x^4|(x_1^5x_1^6 + x_2^5x_2^6)x_2^7\cdots x_2^n$. Apply Cauchy-Schwarz to bound the factor in parentheses by $|x^5|\cdot|x^6|$, majorize each factor $x_t^5x_t^6$ in the other terms by $|x^5|\cdot|x^6|$, and repeat the process. Eventually the desired bound $\prod_{j=3}^{n-2}|x^j|$ is obtained.

The analysis of the second line of (5.4) proceeds in the same way, except that we begin by majorizing x_t^3 and x_t^{n-2} by $|x^3|, |x^{n-2}|$, respectively, in every term, and then proceed as in the preceding paragraph. The same bound $\prod_{j=3}^{n-2}|x^j|$ is obtained. Finally, the case of odd n is the same except for minor changes in notation. Multiplying by $\epsilon x_1^1x_1^2x_2^{n-2}x_2^n$ and taking $\epsilon \le 1/2$, the sum of the contributions of the two terms in (5.4) is $\le x_1^1x_1^2x_2^{n-2}x_2^n\prod_{j=3}^{n-2}|x^j|$.

The next step is to add in the terms $\prod_{j=1}^{n} x_t^j$ for $t = 1, 2$. Cauchy-Schwarz gives us

$$\Big(\prod_{j=1}^{n} x_1^j + x_1^1x_1^2x_2^{n-2}x_2^n\prod_{j=3}^{n-2}|x^j|\Big) + \prod_{j=1}^{n} x_2^j \le x_1^1x_1^2\prod_{j=3}^{n}|x^j| + x_2^1x_2^2\prod_{j=3}^{n}|x^j|. \tag{5.5}$$

Another application of Cauchy-Schwarz gives the bound $\prod_{j=1}^{n}|x^j|$.

We have thus shown that

$$A^n n^{\gamma} M_n(f_1,\ldots,f_n) \le \prod_{j=1}^{n}|x^j| + y_2^n\prod_{j=1}^{n-1}|x^j| + y_1^1\prod_{j=2}^{n}|x^j|, \tag{5.6}$$

which in turn is plainly

$$\le (y_1^1 + |x^1|)\prod_{j=2}^{n-1}|x^j|(y_2^n + |x^n|) \le \prod_{j=1}^{n}(|y^j| + |x^j|). \tag{5.7}$$

This last product is the desired bound.

References

[1] M. Christ and A. Kiselev, *Absolutely continuous spectrum for one-dimensional Schrödinger operators with slowly decaying potentials: some optimal results*, J. Amer. Math. Soc. 11 (1998), 771-797, MR 99g:34166.
[2] ———, *WKB asymptotic behavior of almost all generalized eigenfunctions for one-dimensional Schrödinger operators with Lebesgue class potentials* , preprint January 2000.
[3] A. Kiselev, *Interpolation theorem related to a.e. convergence of integral operators*, Proc. Amer. Math. Soc. 127 (1999), no. 6, 1781–1788, MR 99h:34124.
[4] D. Menshov, *Sur les series de fonctions orthogonales*, Fund. Math. 10, 375-420 (1927).
[5] R.E.A.C. Paley, *Some theorems on orthonormal functions*, Studia Math. 3 (1931) 226-245.
[6] H. Smith and C. Sogge, *Global Strichartz estimates for nontrapping perturbations of the Laplacian*, Comm. Partial Differential Equations, to appear.
[7] T. Tao, *Spherically averaged endpoint Strichartz estimates for the two-dimensional Schrödinger equation*, Comm. Partial Differential Equations, to appear.
[8] A. Zygmund, *Trigonometric Series*, Vol. 2, Cambridge Univ. Press, Cambridge, 1977.
[9] ———, *A remark on Fourier transforms*, Proc. Camb. Phil. Soc. 32 (1936), 321-327.

Michael Christ, Department of Mathematics, University of California, Berkeley, Berkeley, CA 94720, USA
E-mail address: `mchrist@math.berkeley.edu`

Alexander Kiselev, Department of Mathematics, University of Chicago, Chicago, Ill. 60637, USA
E-mail address: `kiselev@math.uchicago.edu`