ON MAXIMAL FUNCTIONS WITH CURVATURE

BEN KRAUSE

ABSTRACT. We exhibit a class of “relatively curved” \( \gamma(t) := (\gamma_1(t), \ldots, \gamma_n(t)) \), so that the pertaining multi-linear maximal function satisfies the sharp range of Hölder exponents,

\[
\left\| \frac{1}{r} \int_0^r \prod_{i=1}^n |f_i(x - \gamma_i(t))| \, dt \right\|_{L^p(\mathbb{R})} \leq C \cdot \prod_{i=1}^n \| f_i \|_{L^{p_i}(\mathbb{R})}
\]

whenever \( \frac{1}{p} = \sum_{j=1}^n \frac{1}{p_j} \), where \( p_j > 1 \) and \( p \geq p_\gamma \), where \( 1 \geq p_\gamma > 1/n \) for certain curves.

For instance, \( p_\gamma = 1/n^+ \) for the case of fractional monomials, \( \gamma(t) = (t^{\alpha_1}, \ldots, t^{\alpha_n}) \), \( \alpha_1 < \cdots < \alpha_n \).

Two sample applications of our method are as follows:

For any measurable \( u_1, \ldots, u_n : \mathbb{R}^n \to \mathbb{R} \), with \( u_i \) independent of the \( i \)th coordinate vector, and any relatively curved \( \gamma \),

\[
\lim_{r \to 0} \frac{1}{r} \int_0^r F(x_1 - u_1(x) \cdot \gamma_1(t), \ldots, x_n - u_n(x) \cdot \gamma_n(t)) \, dt = F(x_1, \ldots, x_n), \text{ a.e.}
\]

for every \( F \in L^p(\mathbb{R}^n) \), \( p > 1 \).

Every appropriately normalized set \( A \subset [0, 1] \) of sufficiently large Hausdorff dimension contains the progression,

\( \{x, x - \gamma_1(t), \ldots, x - \gamma_n(t)\} \subset A \),

for some \( t \geq c_\gamma > 0 \) strictly bounded away from zero, depending on \( \gamma \).

1. Introduction

The study of bilinear maximal functions along curves,

\[
B_\gamma(f, g)(x) := \sup_{r > 0} \frac{1}{r} \int_0^r |f(x - t)| \cdot |g(x - \gamma(t))| \, dt,
\]

where \( \gamma \) is thought of as having some “curvature,” grew out of X. Li’s work on the singular integral formulation [14],

\[
H_\gamma(f, g)(x) := \int f(x - t) \cdot g(x - \gamma(t)) \, \frac{dt}{t}.
\]

Date: April 13, 2020.
Aside from its intrinsic interest, understanding the oscillation of the bilinear Euclidean averages (say),

\[ \left\{ \frac{1}{r} \int_0^r f(x - t) \cdot g(x - t^2) \, dt : r \right\}, \quad f, g \in L^2(\mathbb{R}) \]

is an important first step in understanding the analogous dynamical systems problem

\[ \left\{ \frac{1}{r} \sum_{n=1}^r T^n f(x) \cdot T^n g(x) : r \right\}, \quad f, g \in L^2(X) \]

where \( T : X \to X \) is a measure-preserving transformation on a probability space, \( (X, \mu) \). The author will address the dynamical situation in forthcoming work, joint with M. Mirek and T. Tao [11].

The first non-trivial estimates for (1) were established by X. Li [15], in the case where \( \gamma \) is a polynomial, and was continued with the work of A. Gaitan and V. Lie, [6]. In particular, the following norm estimates were obtained.

**Theorem 1.1** ([15]). Suppose that \( \gamma \) is a polynomial of degree \( d \) which vanishes to at least degree 2 at the origin, and that

\[ \frac{1}{p} + \frac{1}{q} = \frac{1}{r} \]

for some \( r > \frac{d-1}{d} \). Then

\[ \| B_{\gamma}(f, g) \|_{L^r(\mathbb{R})} \lesssim_{d, r, p, q} \| f \|_{L^p(\mathbb{R})} \cdot \| g \|_{L^q(\mathbb{R})}. \]

This result is sharp up to the endpoint.

We refer the reader to the subsection on notation below for the definition of \( \lesssim_{d, r, p, q} \).

Using different methods, these results were extended in [6] to handle the more general case where \( \gamma \) is “non-flat,” see [16, Section 2] for a precise definition, and the treatment of (2) under this non-flat assumption. Good representative examples of non-flat curves are monomials

\[ \gamma(t) = |t|^\alpha, \quad \alpha \neq 0, 1, \]

and accordingly the restriction that the bilinear maximal function maps into \( L^1 \) is in fact sharp for the full class of non-flat curves, which is quite robust.

In both approaches, the role of curvature was fundamental, as subtle oscillatory integral techniques interposed crucially in the study of the relevant bilinear multipliers.

The purpose of this note is to exploit this curvature phenomenon in a more efficient manner: we consider the multi-linear maximal functions,

\[ B_{\vec{\gamma}}(f_1, \ldots, f_n)(x) := \sup_{r > 0} \frac{1}{r} \int_0^r \prod_{j=1}^n |f_j(x - \gamma_j(t))| \, dt, \]

under an appropriate relative curvature condition.
Definition 1.2. A smooth curve $\gamma : \mathbb{R}_+ \to \mathbb{R}$ is $\alpha$-regular at 0 if there exists some $c \neq 0$, so that
$$\gamma(t) = t^\alpha \cdot (c + \varphi(t)), \quad t \to 0^+$$
where
$$|\partial^j \varphi(t)| = o_{t \to 0^+}(t^{-j}),$$
for each $j \geq 0$, and $A$-regular at $\infty$ if there exists some $C \neq 0$ so that
$$\gamma(t) = t^A \cdot (C + \Phi(t)), \quad t \to \infty$$
where
$$|\partial^j \Phi(t)| = o_{t \to \infty}(t^{-j}).$$
$\gamma$ is $(\alpha, A)$-regular if $\gamma$ is both $\alpha$-regular at 0, and $A$-regular at $\infty$.

The class of curves we consider will be regular component wise, but will have pairwise distinct orders of vanishing/growth near the origin and infinity.

Definition 1.3. Given a curve $\vec{\gamma}(t) := (\gamma_1(t), \ldots, \gamma_n(t))$, we say that $\vec{\gamma}$ is relatively curved if the $\gamma_i$ are $(\alpha_i, A_i)$ regular, and that $\alpha_i \neq \alpha_j$ for any $i \neq j$, and similarly $A_i \neq A_j$ for $i \neq j$.

Good examples of relatively curved $\vec{\gamma}$ are (fractional) polynomial curves are of the following form:
$$\gamma_i(t) = c_i \cdot t^{\alpha_i} + P_i(t) + C_i t^{A_i}, \quad c_i, C_i \neq 0$$
where $P_i$ are (not necessarily distinct) fractional polynomials, all of whose monomials have degrees $\alpha_i < \beta < A_i$.

More exotic examples can be constructed, for instance
$$\gamma_i(t) = \log(1 + t^{\alpha_i}) \cdot e^{\sqrt{\log(10 + t^{\rho_i})}} + t^{A_i},$$
where $\alpha_i < A_i$, $\rho_i \geq 0$, provided that the $\{\alpha_i\}$ are all distinct, as are the $\{A_i\}$.

Another perspective on the requirement of regular curvature is the following: Define
$$\text{Leading Order}(\vec{\gamma}(t)) := (t^{\beta_1}, \ldots, t^{\beta_n}), \quad \beta_i := \begin{cases} \alpha_i & \text{if } t \to 0^+ \\ A_i & \text{if } t \to \infty \end{cases}.$$ 

The class of relatively curved $\vec{\gamma}$ are regular component-wise, so that
$$\{\xi \in \mathbb{R}^n : \text{Leading Order}(\vec{\gamma}(t)) \cdot \xi = 0\} = \{0\}$$
gives the zero frequency a distinguished role, which allows us to bring classical real variable techniques to bear. Note that in the bilinear setting, relative curvature forces $\gamma$ to vanish to degree 2 at the origin.

The point of this restriction is that – in the language of time frequency analysis – we preclude any modulation invariance from our analysis. Indeed, the purely modulation invariant case, when each $\gamma_i = c_i \gamma$ are scalar multiples, is covered by time-frequency methods in [13] in the bilinear setting, and by [3] in greater generality.
In the special case where $\vec{\gamma} = \text{Leading Order}(\vec{\gamma})$ are fractional monomials, we have the following theorem.

**Theorem 1.4.** Suppose that $\vec{\gamma} = (t^{\alpha_1}, \ldots, t^{\alpha_n})$ is a fractional monomial curve, where $0 < \alpha_1 < \cdots < \alpha_n$ are distinct. Then for each $q > 1/n$ so that

$$\frac{1}{q} = \frac{1}{p_1} + \cdots + \frac{1}{p_n}, \quad p_i > 1$$

the following estimate holds:

$$\|B_{\vec{\gamma}}(f_1, \ldots, f_n)\|_q \lesssim_{\vec{\gamma}, q, p_i} \prod_{i=1}^n \|f_i\|_{p_i}.$$ 

For more general curves, we are unable to push exponents down to $q > 1/n$, but we have the following substitute, in the particular case where the our curves $\gamma_i$ are all polynomials whose monomials have pairwise different degrees:

$$\gamma_i(t) = \sum_{l \in V_i} a_l t^l$$

where $V_i \cap V_j = \emptyset$ for $i \neq j$. Call such $\vec{\gamma}$ polynomially curved.

**Theorem 1.5.** Suppose that $\vec{\gamma}$ is polynomially curved of degree $d$. Then for each $q > \frac{d-1}{d}$ so that

$$\frac{1}{q} = \frac{1}{p_1} + \cdots + \frac{1}{p_n}, \quad p_i > 1$$

the following estimate holds:

$$\|B_{\vec{\gamma}}(f_1, \ldots, f_n)\|_q \lesssim_{\vec{\gamma}, q, p_i} \prod_{i=1}^n \|f_i\|_{p_i}.$$ 

Moreover, for each $d \geq 2$ there exist polynomially curved $\vec{\gamma}$ of degree $d$ that do not map below $L^{\frac{d}{d-1}}$.

In greatest generality, we have the following result.

**Theorem 1.6.** Suppose that $\vec{\gamma}$ is relatively curved. Then for each $q \geq 1$ so that

$$\frac{1}{q} = \frac{1}{p_1} + \cdots + \frac{1}{p_n}, \quad p_i > 1$$

the following estimate holds:

$$\|B_{\vec{\gamma}}(f_1, \ldots, f_n)\|_q \lesssim_{\vec{\gamma}, q, p_i} \prod_{i=1}^n \|f_i\|_{p_i}.$$ 

1.1. **Applications.** There are two applications of our method: first, to the theory of variable coefficient maximal functions, and second, to Euclidean Ramsey theory.
1.1.1. Variable Coefficient Maximal Functions. One important question in the theory of variable coefficient maximal functions concerns the weakest conditions on curves

\[ \Gamma(x, t) : \mathbb{R}^n_x \times \mathbb{R}_t \to \mathbb{R}, \]

where \( t \mapsto \Gamma(x, t) \) has some “curvature,” so that the maximal function

\[ M_{\Gamma} F(x) := \sup_{r > 0} \frac{1}{r} \int_0^r |F(x - \Gamma(x, t))| \, dt \]

is bounded on \( L^p(\mathbb{R}^n) \). For instance, the work of [7] handled the case where

\[ \Gamma(x, t) := \Gamma(x_1, x_2, t) := (t, u(x_1) \cdot t^\alpha), \quad 0 < \alpha \neq 1 \]

for measurable \( u \), in which case full \( L^p(\mathbb{R}^2) \), \( p > 1 \) estimates for \( M_{\Gamma} \) were established independent of \( u \), [7, Corollary 1.3].

This result was extended in [17], for instance to the case where

\[ \Gamma(x_1, x_2, t) := (t, \sum_{i=1}^n u_i(x_1) \cdot t^{\alpha_i}), \quad 0 < \alpha_1 < \cdots < \alpha_n, \alpha_i \neq 1, \]

\( u_i \) measurable, see [17, Theorem 1]. For a more thorough perspective on these results, we refer the reader to [7] and [17].

Using the methods developed in this paper, we are able to extend these results to the higher dimensional setting under the assumption of relative curvature.

For \( \vec{u} := (u_1, \ldots, u_n) \) so that \( u_i \) is independent of the \( i \)th coordinate variable, define the variable coefficient maximal function along \( \vec{\gamma} \),

\[ M_{\vec{\gamma}, \vec{u}} F(x) := \sup_{r > 0} \frac{1}{r} \int_0^r |F(x_1 - u_1(x) \cdot \gamma_1(t), \ldots, x_n - u_n(x) \cdot \gamma_n(t))| \, dt, \]

as well as its local counterpart,

\[ M_{\vec{\gamma}, \vec{u}, r_{\vec{\gamma}}} F(x) := \sup_{r_0 > r > 0} \frac{1}{r} \int_0^r |F(x_1 - u_1(x) \cdot \gamma_1(t), \ldots, x_n - u_n(x) \cdot \gamma_n(t))| \, dt. \]

**Theorem 1.7** (Variable Coefficient Maximal Functions). Suppose that \( F \in L^p(\mathbb{R}^n) \), and that \( \vec{\gamma} \) is component-wise homogeneous and relatively curved, or alternatively \( \vec{u} \) depends only on a single \( x_i \).

Then the following estimates hold:

\[ \| M_{\vec{\gamma}, \vec{u}} F \|_{L^p(\mathbb{R}^n)} \lesssim_{\vec{\gamma}, p, n} \| F \|_{L^p(\mathbb{R}^n)}, \]

for any \( p > 1 \).

In the case where \( \vec{\gamma} \) is only relatively curved, there exists \( r_{\vec{\gamma}} > 0 \) so that

\[ \| M_{\vec{\gamma}, \vec{u}, r_{\vec{\gamma}}} F \|_{L^p(\mathbb{R}^n)} \lesssim_{\vec{\gamma}, p, n} \| F \|_{L^p(\mathbb{R}^n)}. \]

By a standard density argument, the following corollary then presents.
Corollary 1.8. For any relatively curved $\vec{\gamma}$, and any measurable $u_1, \ldots, u_n$ so that $u_i$ is independent of the $i$th coordinate vector,

$$\lim_{r \to 0} \frac{1}{r} \int_0^r F(x_1 - u_1(x) \cdot \gamma_1(t), \ldots, x_n - u_n(x) \cdot \gamma_n(t)) \, dt = F(x_1, \ldots, x_n), \text{ a.e.}$$

for every $F \in L^p(\mathbb{R}^n)$, $p > 1$.

1.1.2. Euclidean Ramsey Theory. A second application of our method is to problems in Euclidean Ramsey Theory, in particular to the detection of polynomial (or relatively curved) progressions inside “sparse” sets. The detection of three term arithmetic progressions inside of suitable fractals was first conducted in [12], and the case of longer linear progressions was addressed in [2]. In both cases, the fractals considered were by necessity “nice,” as there exist full-dimensional subsets $A \subset [0,1]$, $\dim_H(A) = 1$, that do not contain any three-term arithmetic progressions, see [8, 9]. More precisely, the work of [12, 2] addressed the situation where $A$ had sufficiently large Fourier dimension, $\dim_F(A) > 1 - \epsilon$. These two notions of dimension are linked through the Fourier transform, and one always has $\dim_F(A) \leq \dim_H(A)$. On the other hand, there are sets of large Hausdorff dimension which have zero Fourier dimension, [5].

In the setting of curved (polynomial) progressions, preliminary results on three-term progressions, see [10], suggest that no restriction on Fourier dimension is needed. In particular, the focus of our work here will be the deduction of a non-linear Szemeredi-type theorem for sets $A \subset [0,1]$ of sufficiently large Hausdorff dimension,

$$1 - \epsilon < \dim_H(A) \leq 1.$$

Before stating our result in this direction, we need to normalize our sets in question:

For fractals $A$ of $\dim_H A \geq \beta$, we will say that $A$ is $(\Lambda, \beta)$ normalized if there exists some Frostman measure supported on $A$, $\mu$, with

$$\mu(I) \leq \Lambda \cdot |I|^\beta,$$

for each interval $I \cap A \neq \emptyset$.

\begin{equation}
\mu(I) \leq \Lambda \cdot |I|^\beta,
\end{equation}

for each interval $I \cap A \neq \emptyset$.

\textbf{Theorem 1.9} (Euclidean Ramsey Theory). Suppose that $\vec{\gamma}$ is relatively curved, and that $A \subset [0,1]$ is $(\Lambda, \beta)$ normalized, where $\beta$ is sufficiently large,

$$1 - c_{\vec{\gamma},\Lambda} < \beta \leq \dim_H A \leq 1$$

for some absolute constant $c_{\vec{\gamma},\Lambda} > 0$.

Then there exists an absolute constant, $c_{\vec{\gamma},\beta} > 0$, so that one may find the non-linear progression

$$\{x, x - \gamma_1(t), \ldots, x - \gamma_n(t)\} \in A$$

\footnote{There exist $\beta$-dimensional fractals which cannot be $(\Lambda, \beta)$ normalized for any $\Lambda < \infty$; on the other hand, for any $\beta$-dimensional $A$, for every $\kappa < \beta$, there exist $\Lambda_\kappa < \infty$ so that $A$ may be $(\Lambda_\kappa, \kappa)$ normalized.}
for some $t \geq c_{\gamma, \beta} > 0$ bounded away from 0.

The structure of the paper is as follows:

- In §2 we prove non-trivial $L^1$ estimates for certain “oscillatory” kernels; these estimates form the key quantitative input for the paper;
- In §3 we prove single-scale estimates, which will be used in proving Theorems 1.5 and 1.6;
- In §4 we combine the foregoing discussion to complete the proofs of Theorems 1.4, 1.5, and 1.6;
- §5 contains our application to variable coefficient maximal functions, Theorem 1.7;
- §6 contains our application to Euclidean Ramsey theory, Theorem 1.9.

1.2. Notation. Here and throughout, $e(t) := e^{2\pi it}$. Throughout, $C$ will be a large number which may change from line to line. $M_{HL}$ will denote the Hardy-Littlewood maximal function, and for functions of multiple variables, we let

\[(7) \quad M''_{HL} F(x) := \sup_{r>0} \frac{1}{2r} \int_{-r}^{r} |F(x_1, \ldots, x_i - t, \ldots, x_n)| \, dt\]

denote the Hardy-Littlewood maximal function in the $i$th variable. We will also often use the so-called shifted maximal functions,

\[(8) \quad M_s f(x) := \sup_{r>0} \sup_{|u| \leq r} \frac{1}{2r} \int_{-r}^{r} |f(x - u - t)| \, dt,\]

and define $M''_s$ analogously. For $\vec{s} = (s_1, \ldots, s_n)$, we will reserve

\[(9) \quad s_0 := \max_i s_i.\]

Throughout, we let $\varphi$ denote various mean-one Schwartz functions, normalized in some sufficiently large semi-norm. The precise choice of $\varphi$ might differ from line to line. Similarly, we use $\psi$ to denote a similar function, but with mean zero. Often, we will assume that such functions may be chosen to have Fourier transform supported in an annulus away from the origin. We let

$\phi_t(x) := 2^{-t} \phi(2^{-t} x)$.

We will make use of the modified Vinogradov notation. We use $X \lesssim Y$, or $Y \gtrsim X$, to denote the estimate $X \leq CY$ for an absolute constant $C$. We use $X \approx Y$ as shorthand for $Y \lesssim X \lesssim Y$. We also make use of big-O notation: we let $O(Y)$ denote a quantity that is $\lesssim Y$. We let $f(t) := o_t \to a(X(t))$ denote a quantity so that $\frac{|f(t)|}{X(t)} \to 0$ as $t \to a$.

If we need $C$ to depend on a parameter, we shall indicate this by subscripts, thus for instance $X \lesssim_p Y$ denotes the estimate $X \leq C_p Y$ for some $C_p$ depending on $p$. We analogously define $O_p(Y)$. 
2. Kernel Computations

By non-negativity, there is no loss in re-defining the maximal function

\[ B(f_1, \ldots, f_n)(x) := B_\gamma(f_1, \ldots, f_n)(x) := \sup_k |B_k(f_1, \ldots, f_n)(x)|, \]

where

\[ B_k(f_1, \ldots, f_n)(x) := \int \prod_{i=1}^n f_i(x - \gamma_i(t)) \cdot \rho_k(t) \, dt, \]

where \( \rho_k(t) := 2^{-k} \rho(2^{-k}t) \) for some \( L^1 \) normalized \( \rho \) supported in \( \{ t \approx 1 \} \).

The key to our argument is an appropriate analysis of the following kernel. In what follows, we will assume that \( |k| \gg 1 \) is so large that \( \vec{\gamma} \) is well approximated by its Taylor expansion near 0 and \( \infty \). For concreteness, we will assume that \( \gamma_i \) is \( (\alpha_i, A_i) \)-regular.

We will only consider the local case, when \( t \to 0^+ \), so \( k \to -\infty \), as the opposite case, when \( t, k \to \infty \), is similar.

So, with \( k \gg 1 \), consider the function of \( n \) variables,

\[ K_{\vec{s},k}(v_1, \ldots, v_n) := \int \prod_{i=1}^n \psi_{-s_i - k\alpha_i}(v_i - \gamma_i(t)) \cdot \rho_{-k}(t) \, dt, \quad \vec{s} = (s_1, \ldots, s_n). \]  

With

\[ D_k G(v_1, \ldots, v_n) := 2^k \sum \alpha_i \cdot G(v_1 \cdot 2^{k\alpha_1}, \ldots, v_n \cdot 2^{k\alpha_n}) \]

the \( L^1 \)-normalized dilation, observe that

\[ K_{\vec{s},k}(v_1, \ldots, v_n) =: D_k T_{\vec{s},k}(v_1, \ldots, v_n), \]

where

\[ T_{\vec{s},k}(v_1, \ldots, v_n) = \int \prod_{i=1}^n \psi_{-s_i}(v_i - 2^{k\alpha_i} \gamma_i(2^{-k}t)) \cdot \rho(t) \, dt \]

is increasingly “independent of \( k \)” for \( |k| \to \infty \), by the definition of relative curvature.

Now, let \( c_0 \) be a small parameter, and set \( V_i \) to be the collection of points \( v_i \) so that for some \( t \approx 1 \),

\[ (1 + 2^{s_i + k\alpha_i} |v_i - \gamma_i(2^{-k}t)|) \lesssim 2^{c_0 s_0} \]

in \( V_i \). Notice that if we set \( w_i := 2^{k\alpha_i} v_i \), then we see that

\[ (1 + 2^{s_i} |w_i - 2^{k\alpha_i} \gamma_i(2^{-k}t)|) \lesssim 2^{c_0 s_0}, \]

which says that \( w_i \) lives in a \( 2^{c_0 s_0 - s_i} \) neighborhood of (say) a \( C^1 \) curve that lives at unit scales,

\[ t \mapsto 2^{k\alpha_i} \cdot \gamma_i(2^{-k}t), \quad t \approx 1. \]

In particular,

\[ |\{ w_i : (13) \}| \lesssim 2^{c_0 s_0 - s_i}, \]
and thus
\begin{equation}
|V_i| := |\{v_i : (12)\}| \lesssim 2^{c_0 s_0 - s_i - k\alpha_i}.
\end{equation}

Collect
\begin{equation}
X_{\vec{s},k} := \{v = (v_1, \ldots, v_n) : v_i \in V_i\},
\end{equation}
and note that
\begin{equation}
|X_{\vec{s},k}| \lesssim 2^{nc_0 s_0 - S - kA},
\end{equation}
by (14). Here and throughout
\begin{equation}
S := \sum_i s_i, \quad A := \sum_i \alpha_i.
\end{equation}

On the other hand, if $v \notin X_{\vec{s},k}$, for each $t \approx 1$, and any $N \gg 1$ sufficiently large,
\begin{equation}
\prod_{i=1}^n \psi_{-s_i - k\alpha_i}(v_i - \gamma_i(2^{-k}t))
\end{equation}
\begin{equation}
\lesssim_N 2^{S+kA} \cdot \min \left\{ 2^{-N^2c_0} \prod_{i=1}^n (1 + 2^{s_i + k\alpha_i} |v_i - \gamma_i(2^{-k}t)|)^{-N^2} \right\}
\end{equation}
\begin{equation}
\lesssim 2^{-Ns_0} \cdot 2^{S+kA} \cdot \prod_{i=1}^n (1 + 2^{s_i + k\alpha_i} |v_i - \gamma_i(2^{-k}t)|)^{-N}
\end{equation}
for $N$ sufficiently large.

**Lemma 2.1.** There exists an absolute constant $1 > c \gg c_0 > 0$ so that the following pointwise estimate holds:
\begin{equation}
|K_{\vec{s},k}(v_1, \ldots, v_n)|
\end{equation}
\begin{equation}
\lesssim 2^{-Ns_0} \cdot 2^{S+kA} \cdot \int \prod_{i=1}^n (1 + 2^{s_i + k\alpha_i} |v_i - \gamma_i(2^{-k}t)|)^{-N} \rho(t) \, dt
\end{equation}
\begin{equation}
+ 2^{S+kA} \cdot 2^{-cs_0} \cdot 1_{X_{\vec{s},k}}(v_1, \ldots, v_n).
\end{equation}

**Proof.** By (20), we may assume that $v \in X_{\vec{s},k}$. We use Fourier inversion to express
\begin{equation}
K_{\vec{s},k}(v_1, \ldots, v_n) = \int e \left( \sum_i \xi_i v_i \right) \cdot e \left( -\sum_i \xi_i \gamma_i(t) \right) \cdot \prod_i \hat{\psi}(2^{-s_i - k\alpha_i} \xi_i) \cdot \rho_k(t) \, dt d\xi
\end{equation}
\begin{equation}
= 2^{S+kA} \cdot \int I_{\vec{s},k}(\xi_1, \ldots, \xi_n) \cdot \prod_{i=1}^n e(2^{s_i + k\alpha_i} \xi_i v_i) \hat{\psi}(\xi_i) \, d\xi,
\end{equation}
where
\[
I_{\vec{s},k}(\xi_1, \ldots, \xi_n) := \int e \left( - \sum_i 2^{s_i} \cdot \xi_i \cdot 2^{k\alpha_i} \gamma_i(2^{-k}t) \right) \cdot \rho(t) \, dt.
\]

Thus, it suffices to exhibit a gain,
\[
|K_{\vec{s},k}(v_1, \ldots, v_n)| \lesssim 2^{S+k \cdot A} \int |I_{\vec{s},k}(\xi_1, \ldots, \xi_n)| \cdot \prod_{i=1}^n |\hat{\psi}(\xi_i)| \, d\xi
\]
\[
\lesssim 2^{S+k \cdot A} \cdot 2^{-cs_0}
\]
for some \( c > 0 \).

The key observation is that by \((\alpha_i, A_i)\)-regularity implies that one may decompose
\[
\Xi_{\vec{s},k}(t) := \sum_i 2^{s_i} \cdot \xi_i \cdot 2^{k\alpha_i} \cdot \gamma_i(2^{-k}t) = \sum_i 2^{s_i} \cdot \xi_i \cdot t^{\alpha_i} + \Phi(t),
\]
where
\[
\|\Phi(t)\|_{C^N} = o_{|k| \to \infty} \left( \sum_i 2^{s_i} \cdot |\xi_i| \right)
\]
for each \( N \) sufficiently large. In particular, there exists some absolute \( N = N(\vec{\gamma}) \) so that
\[
\min_{t \approx 1} |\partial_t^M \Xi_{\vec{s},k}(t)| \gtrsim \sum_i 2^{s_i} \cdot |\xi_i|
\]
for some \( M \leq N \).

By standard stationary phase estimates, see for instance [19], there exists some \( c_2 > 0 \) so that we may dominate
\[
|I_{\vec{s},k}(\xi_1, \ldots, \xi_n)| \lesssim (1 + \sum_i 2^{s_i} \cdot |\xi_i|)^{-c_2},
\]
and substituting this appropriately yields the estimate (27). \( \square \)

**Proposition 2.2.** Suppose that \( \frac{1}{p_1} + \cdots + \frac{1}{p_n} = 1 \), \( p_j \geq 1 \) are Hölder conjugate. There exists a constant \( c > 0 \) so that for any \( k \geq 0 \)
\[
\left\| \int K_{\vec{s},k}(x-v_1, \ldots, x-v_n) \prod_{i=1}^n f_i(v_i) \, dv \right\|_1 \lesssim 2^{-cs_0} \prod_{j=1}^n \|f_j\|_{p_j}.
\]

**Proof.** By interpolation, it suffices to prove estimates with all \( p_i = 1 \) or \( \infty \). But this just follows from Fubini’s theorem, since for each \( k \gg 1 \)
\[
\|K_{\vec{s},k}\|_1 \lesssim 2^{-Ns_0} + 2^{S+k \cdot A} \cdot 2^{-cs_0} \cdot |X_{\vec{s},k}| \lesssim 2^{-(c-n c_0)s_0} \lesssim 2^{-c/2 \cdot s_0}
\]
for some absolute \( c > 0 \), provided we have chosen \( c_0 > 0 \) sufficiently small. See (16) and Lemma 2.1. \( \square \)
We will upgrade this result to $L^q$ spaces for $q < 1$; the key to this step is the pointwise majorization,

\[
| \int K_{\gamma,k}(x-v_1, \ldots, x-v_n) \prod_{i=1}^n f_i(v_i) \, dv | \lesssim \prod_{i=1}^n M_{s_0} f_i(x),
\]

see (8). Since the shifted maximal functions have $L^p$ norms that grow at most linearly in $s$ by a Calderón-Zygmund endpoint argument,

\[
\| M_s f \|_{L^{1,\infty}(\mathbb{R})} \lesssim s \cdot \| f \|_{L^1(\mathbb{R})},
\]

see [18, Theorem 4.5], or [7, Lemma 3.2] for a stronger statement, we will be able to interpolate between our two estimates, in the full $L^r$, $r > 1/n$ range.

The arguments in this section are essentially sufficient to prove Theorem 1.4, in the case where $\tilde{\gamma}$ is component-wise homogeneous, i.e.

Leading Order($\tilde{\gamma}$) = $\tilde{\gamma}$.

To prove Theorems 1.5 and 1.6 we will require a further ingredient, which will be addressed in the following section.

3. Single Scale Estimates

The estimates that we proved in the previous section relied on the validity of certain Taylor expansions holding near the origin or near infinity, see (29) and (30). For the case of pure monomials, these expansions are exact; in general, we have the following single-scale substitute, which we will be prepared to use \( \lesssim \) many times. The argument in this section is essentially lifted from [15, §3].

Proposition 3.1. Suppose that $\tilde{\gamma}$ is a polynomially curved of degree $d$. Then the following estimate holds,

\[
\| B_k(f_1, \ldots, f_n) \|_q \lesssim \prod_{i=1}^n \| f_i \|_{p_i},
\]

whenever $q > \frac{d-1}{d}$ and $\frac{1}{q} = \frac{1}{p_1} + \cdots + \frac{1}{p_n}$, and this is sharp up to the end-point.

In the more general relatively curved setting, the same estimate holds for $q \geq 1$.

Proof. The second statement follows by convexity, so we turn to the first. We remark that for $r < 1$ we have the inequality,

\[
\| \sum_n a_n \|_r^r \leq \sum_n \| a_n \|_r^r,
\]

which follows from the sub-additivity of $t \mapsto t^r$, $0 \leq r < 1$. 

Turning to the proof, by interpolation we may reduce to the case where $p_i = \infty$ for all $i \neq j_1, j_2$; without loss of generality these will be $j_1 = 1, j_2 = 2$. We will let
\[
\|P\| := \left\| \sum_j a_j t^j \right\| := \sum_{j=1}^n |a_j|
\]
denote the coefficient norm of $P$, and will use the well-known level set estimate, see for instance [21, Proposition 2.2],
\[
|\{ |t| \lesssim 1 : |P(t)| \lesssim \lambda \}| \lesssim (\lambda/\|P\|)^{-1/d}
\]
for $P$ a polynomial of degree $d$.

In what follows, set
\[
B_{k;X}(f_1, f_2)(x) := \int_X f_1(x - \gamma_1(2^k t)) \cdot f_2(x - \gamma_2(2^k t)) \rho(t) \, dt.
\]
If we consider the change of variables
\[
(37) \quad u_i = x - \gamma_i(2^k t),
\]
then our Jacobian is exactly
\[
(38) \quad \left| \frac{\partial(x, t)}{\partial(u_1, u_2)} \right| = \frac{1}{|\partial_t(\gamma_2(2^k t) - \gamma_1(2^k t))|}.
\]
In particular, if we let
\[
\Gamma := \sup_{i \in X} \left| \frac{\partial(x, t)}{\partial(u_1, u_2)} \right|
\]
then we may always estimate
\[
(39) \quad \|B_{k;X}(f_1, f_2)\|_1 \lesssim J_X \cdot \|f_1\|_1 \cdot \|f_2\|_1.
\]
Accordingly, we will pigeon-hole $\text{supp} \ \rho \subset \{ |t| \approx 1 \}$ so that
\[
\left| \frac{\partial(x, t)}{\partial(u_1, u_2)} \right|
\]
is approximately constant.

In particular, for each $m_i$, set
\[
E_{m_i} := \{ |t| \approx 1 : |\partial_t(\gamma_i(2^k t))| \approx 2^{m_i} \},
\]
and set $X_{m_1, m_2} := E_{m_1} \cap E_{m_2}$, and let
\[
\Gamma_i := \{ \gamma_i(s) : s \in X_{m_1, m_2} \}.
\]
We have the following restriction:
\[
(40) \quad m_i \leq \begin{cases} \frac{dk - C}{2} & \text{if } k \leq 0 \\ \frac{dk + C}{2} & \text{if } k \geq 0. \end{cases}
\]
Note that $\Gamma_i = \bigcup_{J_i} I$ are unions of $\lesssim d$ intervals, each of length
\[
\max_{I \in J_i} |I| \lesssim 2^{m_i} \cdot |X_{m_1, m_2}|. 
\]
In what follows, we will assume that $m_1 \geq m_2$, so that we may assume that our spatial variable, $x$, lives inside $P$, an interval of length $|P| \lesssim 2^{m_1} \cdot |E_{m_1}|$, and that this holds uniformly for all $m_2 \leq m_1$.

Moving forward, there are two cases we consider, according whether $m_1 \geq m_2 + C_\gamma$, in which (38) has size approximately $2^{-m_1}$, or $m_2 \leq m_1 \leq m_2 + C_\gamma$, in which slightly more detailed analysis is needed.

Beginning with the first case, $m_1 \geq m_2 + C_\gamma$, we set
\[
X = \bigcup_{m_2 \leq m_1 - C_\gamma} X_{m_1, m_2}
\]
and use Cauchy-Schwartz to dominate
\[
\|B_{k; X}(f_1, f_2)\|_{1/2} = |P| \cdot \|B_{k; X}(f_1, f_2)\|_1 \lesssim 2^{m_1} \cdot |E_{m_1}| \cdot \|B_{k; X}(f_1, f_2)\|_1 \lesssim |E_{m_1}| \cdot \|f_1\|_1 \cdot \|f_2\|_1,
\]
by (39).

It suffices now to show that $|E_{m_1}|$ is geometrically decaying in $m_1$ restricted by (40).

Using (36), we may estimate
\[
|E_{m_1}| \lesssim |\{ |t| \leq 1 : |2^k (\gamma_1)'(2^k t)| \lesssim 2^{m_1} \}| \lesssim \left( \frac{2^{m_1}}{\|2^k (\gamma_1)'(2^k t)\|} \right)^{-\frac{1}{r-1}}
\]
which decays appropriately, by splitting according to the sign of $k$, see (40) above. By interpolating the trivial estimates at $L^\infty$, we deduce that for each $p_i$ so that
\[
\frac{1}{r} = \frac{1}{p_1} + \frac{1}{p_2}
\]
with $r > 1/2$, $p_1, p_2 > 1$, we have the estimate
\[
\| \sum_{m_1 \geq m_2 + C_\gamma} B_{k; X_{m_1, m_2}}(f_1, f_2) \|_r \lesssim \|f_1\|_{p_1} \cdot \|f_2\|_{p_2},
\]
since we can use the geometric decay in $m$ to sum, see (35).

It remains to handle the remaining case; by paying a constant factor, it suffices to assume that $m_1 = m_2$, and we will re-label $X_m := X_{m, m}$. 


We will show that for each \( m \), and each \( r > \frac{d-1}{d} \), there exists some \( \theta_r > 0 \) so that
\[
\| B_{k;X_m,l}(f_1, f_2) \|_r \lesssim 2^{-\theta_r \cdot \frac{r}{d-1}} \cdot (2^m/\|2^k \gamma_1(2^k t)\|)^{-\theta_r \cdot \frac{r}{d-1}} \cdot \| f_1 \|_{p_1} \cdot \| f_2 \|_{p_2}
\]
whenever \( \frac{1}{r} = \frac{1}{p_1} + \frac{1}{p_2}, \ p_i > 1 \). Indeed, given (41), we will be able to sum the appropriate geometric series, and estimate
\[
\| \sum_{m,l} B_{k;X_m,l}(f_1, f_2) \|_r \lesssim \| f_1 \|_{p_1} \cdot \| f_2 \|_{p_2}.
\]
Guided by (39), we further decompose the level sets of the Jacobian
\[
\left| \frac{\partial(x, t)}{\partial(u_1, u_2)} \right|^{-1} = |\partial_t (\gamma_2(2^k t) - \gamma_1(2^k t))| \approx 2^{m-l}, \ l \geq 0
\]
i.e. we let
\[
X_{m;l} := \left\{ t \in X_m : \left| \frac{\partial(x, t)}{\partial(u_1, u_2)} \right| \approx 2^{l-m} \right\},
\]
see (37). Note that we may assume that \( B_{k;X_m,l} \) is supported on \( O_d(1) \) many intervals of length \( \lesssim 2^m \cdot |X_{m;l}| \). This yields an \( L^1 \) estimate
\[
\| B_{k;X_m,l}(f_1, f_2) \|_1 \lesssim 2^{l-m} \cdot \| f_1 \|_1 \cdot \| f_2 \|_1,
\]
so that by Cauchy-Schwartz
\[
\| B_{k;X_m,l}(f_1, f_2) \|_{1/2} \lesssim 2^l \cdot |X_{m;l}| \cdot \| f_1 \| \cdot \| f_2 \|.
\]
On the other hand, for any \( \frac{1}{p_1} + \frac{1}{p_2} = 1 \), we may estimate
\[
\| B_{k;X_m,l}(f_1, f_2) \|_{1} \lesssim \prod_{i=1}^{2} \left( \int \int |f_i(x - \gamma_i(t))|^{p_i} 1_{X_{m;l}}(t) \ dx \ dx \right)^{1/p_i}
\]
\[
\lesssim |X_{m;l}| \cdot \prod_i \| f_i \|_{p_i}.
\]
To estimate \( |X_{m;l}| \), we first observe that on \( X_{m;l} \), we have the lower bound
\[
\| \partial_t (\gamma_2(2^k t) - \gamma_1(2^k t)) \| \gtrsim 2^m,
\]
since the degrees of the monomials appearing \( \gamma_1 \) and \( \gamma_2 \) are all distinct, and we have the lower bound
\[
2^m \gtrsim |\partial_t (\gamma_i(2^k t))| \lesssim \| \partial_t (\gamma_i(2^k t)) \|.
\]
Consequently, we obtain the desired estimate,
\[
|X_{m;l}| \lesssim \min \{ 2^{-\frac{m}{d-1}}, |X_m| \} \lesssim 2^{-\frac{m}{d-1}} \cdot (2^m/\|2^k \gamma_1(2^k t)\|)^{-\frac{m}{d-1}},
\]
where \( 0 < \kappa = \kappa(r) \ll 1 \) is sufficiently small. Interpolating yields (41).

As far as sharpness is concerned, one simply uses [15, §3.2] at the single (unit) scale level, with all but two functions equal to (say) \( 1_{|x| \leq 10} \). \( \square \)
4. The Argument

In light of the single scale estimates developed in the previous section, throughout the following discussion we will assume that $|k| \gg \gamma$ is sufficiently large, see (29) and (30), and we will show that the maximal function restricted to these scales maps into $L^r$ for $r > 1/n$. For notational ease, we will re-label

$$B_{\gamma}(f_1, \ldots, f_n) := \sup_{|k| \geq C_{\gamma}} |B_k(f_1, \ldots, f_n)|.$$ 

In what follows, we will let

$$\beta_i := \left\{ \begin{array}{ll}
\alpha_i & \text{if } k \leq -C_{\gamma} \\
A_i & \text{if } k \geq C_{\gamma} \end{array} \right..$$

We begin by introducing a “high-frequency” decomposition of our maximal operator.

To this end, set

$$g_i = f_i - \varphi_k \beta_i * f_i$$

and decompose

$$B_k(g_1, \ldots, g_k) := \sum_{s \geq 1} \left( \sum_{s, s_0 = s} B_k(\psi_{k\beta_1-s_1} * g_1, \ldots, \psi_{k\beta_n-s_n} * g_n) \right)$$

$$=: \sum_{s \geq 1} B_{k,s}(g_1, \ldots, g_n),$$

and set

$$B_{\gamma,s}(g_1, \ldots, g_n) := \sup_k |B_{k,s}(g_1, \ldots, g_n)|, \ s \geq 0.$$ 

Note that we can assume that the sum over $s_i \leq s_0$ is truncated below at $s_i \geq 0$.

**Proposition 4.1.** For any $q > \frac{1}{n}$, and any $p_1, \ldots, p_n > 1$ so that

$$\frac{1}{q} = \frac{1}{p_1} + \cdots + \frac{1}{p_n}$$

there exists some absolute $c > 0$ so that

$$\|B_{\gamma,s}(g_1, \ldots, g_n)\|_q \lesssim 2^{-cs} \cdot \prod_{i=1}^n \|f_i\|_{p_i}.$$ 

The $g_i$ are defined in (43).

**Proof.** By (34), we have an upper bound of $s^n$ at $1/n < q_0 < q$, so our task is to interpolate this with something better.
At \( q = 1 \), when each \( p_j \geq 2 \) we may use Hölder’s inequality and the boundedness of the Littlewood-Paley square function to estimate

\[
\sum_{1 \leq s_i \leq s} \sum_k \|B_k(\psi_{k\beta_1-s_1} * f_1, \ldots, \psi_{k\beta_n-s_n} * f_n)\|_1
\]

\[
\lesssim 2^{-cs} \sum_{1 \leq s_i \leq s} \sum_k \prod_{i=1}^n \|\psi_{k\beta_i-s_i} * f_i\|_{p_i}
\]

\[
\lesssim 2^{-cs} \sum_{1 \leq s_i \leq s} \prod_{i=1}^n \left( \sum_k \|\psi_{k\beta_i-s_i} * f_i\|_{p_i}^{p_i} \right)^{1/p_i}
\]

\[
= 2^{-cs} \sum_{1 \leq s_i \leq s} \prod_{i=1}^n \left( \sum_i \|\psi_{k\beta_i-s_i} * f_i\|_{p_i}^{p_i} \right)^{1/p_i}
\]

\[
\leq s^n \cdot 2^{-cs} \cdot \prod_{i=1}^n \|Sf_i\|_{p_i},
\]

where

\[
Sf^2 := \sum_n |\psi_n * f|^2,
\]

is the Littlewood-Paley square function, and we have used that each \( p_i \geq 2 \) to pass to the final inequality. Since \( S \) is bounded on each \( L^p, 1 < p < \infty \), we have arrived at the estimate

\[
\|B_{\gamma,s}(g_1, \ldots, g_n)\|_1 \lesssim s^n \cdot 2^{-cs} \cdot \prod_{i=1}^n \|f_i\|_{p_i},
\]

which allows us to interpolate. \( \square \)

We now combine the uncertainty principle with Proposition 4.1 to prove our Theorems 1.4, 1.5, and 1.6; note that as far as Theorem 1.6, by interpolation it suffices to prove that \( B_{\gamma} \) maps into \( L^1 \).

**Lemma 4.2.** Suppose \( l \geq 0 \). One may express

\[
B_k(f_1, \ldots, \varphi_{k\beta_1+l} * f_i, \ldots, f_n) = \varphi_{k\beta_1+l} * f_i \cdot B_{k\neq i}(f_1, \ldots, f_n) + E_l,
\]

where

\[
|E_l| \lesssim 2^{-l} \cdot M_{HL} f_i \cdot \sup_k B_{k\neq i}(|f_1|, \ldots, |f_n|).
\]

Here,

\[
B_{k, \neq i}(f_1, \ldots, f_n)(x) := \int \prod_{j \neq i} f_j(x - \gamma_j(t)) \rho_k(t) \, dt.
\]
Proof. We set
\[ E_i := B_k(f_1, \ldots, \varphi_{k\beta_i+t} \ast f_i, \ldots, f_n) - \varphi_{k\beta_i+t} \ast f_i \cdot B_{k, \neq i}(f_1, \ldots, f_n), \]
and use the mean-value theorem. More precisely,
\[ B_k(f_1, \ldots, \varphi_{k\beta_i+t} \ast f_i, \ldots, f_n)(x) \]
\[ = \int \left( \int \prod_{j \neq i} f_j(x - \gamma_j(t)) \cdot \varphi_{k\beta_i+t}(x - \gamma_i(t) - y) \cdot \rho_k(t) \, dt \right) f(y) \, dy \]
\[ = \left( \int \prod_{j \neq i} f_j(x - \gamma_j(t)) \rho_k(t) \, dt \right) \cdot \int \varphi_{k\beta_i+t}(x - y) f(y) \, dy \]
\[ + O \left( 2^{-l} \cdot M_{\mathcal{H}L} f_i \cdot \sup_k B_{k, \neq i}(|f_1|, \ldots, |f_n|) \right), \]
since
\[ \sup_{|u| \leq 2^{k\beta_i}} |\varphi_{k\beta_i+t}(x - u) - \varphi_{k\beta_i+t}(x)| \lesssim 2^{-l} \cdot 2^{-k\beta_i - l} \cdot (1 + 2^{-k\beta_i - l} \cdot |x|)^{-100} \]
by the mean-value theorem. \( \square \)

By recursion and induction on \( n \geq 1 \), each of Theorem 1.4, 1.5, 1.6 follows from Proposition 4.1.

The argument is complete.

5. Application One: Variable Coefficient Maximal Functions

In this section, we assume that \( \gamma(t) \) is relatively curved, and that we have re-defined our maximal operator
\[ M_{\vec{\gamma}, \vec{u}} F(x) := \sup_k |M_k F(x)|, \]
where
\[ M_k F(x) := \int F(x_1 - u_1(x) \cdot \gamma_1(t), \ldots, x_n - u_n(x) \cdot \gamma_n(t)) \cdot \rho_k(t) \, dt. \]

By the non-negativity of our maximal operators, there is no harm in making this replacement. One restriction that we will impose is that \( |k| \gg \vec{\gamma} \) 1 sufficiently large that (29) and (30) hold; in the fractional monomial setting, no such excision is necessary.

We begin with some notation: for a function of one variable, \( g \), we use
\[ g \ast_i F(x) := \int F(x_1, \ldots, x_i - t, \ldots, x_n) \cdot g(t) \, dt \]
to denote convolution in the \( i \)th variable. We let
\[ \otimes g_i \ast F(x) := \int F(x - t_1, \ldots, x - t_n) \cdot \prod_i g_i(t_i) \, dt \]
for \( g_1, \ldots, g_n \) one-variable functions.

Define 

\[ 2^{\omega_i(x)} = |u_i(x)|, \quad u_i(x) \neq 0 \]

so that \( \omega_i(x) \) is independent of the \( i \)th coordinate vector.

We will let 

\[ M_{k, \neq i}F(x) := \int F(x_1 - u_1(x) \cdot \gamma_1(t), \ldots, x_i, \ldots, x_n - u_n(x) \cdot \gamma_n(t)) \cdot \rho_k(t) \, dt, \]

denote the lower dimensional averaging operator. Note that whenever \( u_i(x) = 0 \), 

\[ M_k F(x) = M_{k, \neq i} F(x). \]

By induction on the dimension, it therefore suffices to prove Theorem 1.7 under the assumption that \( u_i(x) \neq 0 \) never vanishes, i.e. \( \omega_i(x) \) is (almost) everywhere defined.

Then we have the following lemma, in analogy with Lemma 4.2; its proof is the same.

**Lemma 5.1.** The following identity holds:

\begin{align}
(56) \quad & M_k(\varphi_{k\beta_i + \omega_i(x)} + t \ast_i F)(x) \\
(57) \quad & = \varphi_{k\beta_i + \omega_i(x)} + t \ast_i M_{k, \neq i} F(x) + O \left( 2^{-l} \cdot \sup_k M_{k, \neq i} M_{H_L}^i |F|(x) \right).
\end{align}

\( \beta_i \) are defined in (42).

In particular, by recursion and induction, \( L^p \) estimates on (54) will follow from summing the following estimate over \( s_0 \geq 1 \).

**Proposition 5.2.** For each \( 1 < p \leq \infty \), there exists an absolute \( c_p > 0 \) so that for each \( s_i \geq 0 \)

\[ \| \sup_k |M_k \left( \bigotimes \varphi_{k\beta_i + \omega_i(x) - s_i} \ast_i F \right)(x)\|_{L^p} \lesssim 2^{-c_p s_0} \cdot \|F\|_p. \]

**Proof.** For each \( k \), we have

\begin{align}
(58) \quad & M_k \left( \bigotimes \varphi_{k\beta_i + \omega_i(x) - s_i} \ast_i F \right) \\
(59) \quad & = \int F(x_1 - v_1, \ldots, x_n - v_n) K_{s,k}(x; v) \, dv,
\end{align}

where 

\[ K_{s,k}(x; v) := \int \prod_{i=1}^n \varphi_{k\beta_i + \omega_i(x) - s_i}(v_i - u_i(x) \cdot \gamma_i(t)) \rho_k(t) \, dt, \]

so that we have the pointwise estimate

\[ |M_k \left( \bigotimes \varphi_{k\beta_i + \omega_i(x) - s_i} \ast_i F \right)| \lesssim M_{s_0}^1 \circ \cdots \circ M_{s_0}^n F, \]

where \( M_{s_0}^i \) denotes the shifted maximal \( M_{s_0} \) in the \( i \)th coordinate vector; notice how crucially we used that \( u_i \) was independent of \( x_i \).
We will use this estimate at $L^p_0$, where $p > p_0 > 1$ is small:
$$
\| \sup_k |M_k (\otimes \psi_k \beta_i + \omega_i(x) - s_i \ast F) \|_p \lesssim s^n \cdot \| F \|_{p_0}.
$$

On the other hand, at the $L^\infty$ endpoint, we deduce a power savings,
$$
\| \sup_k |M_k (\otimes \psi_k \beta_i + \omega_i(x) - s_i \ast F) \|_\infty \lesssim 2^{-cs} \cdot \| F \|_\infty.
$$

Indeed, for any $k, x$, we have the uniform estimate
$$
|M_k (\otimes \psi_k \beta_i + \omega_i(x) - s_i \ast F)(x)| \leq \| F \|_\infty \cdot \int |K_{s,k}(x; v)| \, dv \lesssim 2^{-cs_0} \cdot \| F \|_\infty,
$$
by Proposition 2.2.

Interpolation completes the proof.

In particular, this concludes Theorem 1.7 in the case where $\vec{\gamma}$ is component-wise homogeneous, or when one restricts attention to the local maximal function, (5), or alternatively when one only considers sufficiently large scales.

To deal with finitely many exceptional scales, we need to restrict $\vec{u}$ to a single-variable dependence:

**Lemma 5.3.** Suppose that $\vec{u}$ depends only on $x_i$. Then for any $k$, and any $1 \leq p \leq \infty$,
$$
\| \int F(x_1 - u_1(x) \cdot \gamma_1(t), \ldots, x_n - u_n(x) \cdot \gamma_n(t)) \cdot \rho_k(t) \, dt \|_p \lesssim \| F \|_p.
$$

**Proof.** One uses Fubini to bound

(60) $\| F(x_1 - u_1(x) \cdot \gamma_1(t), \ldots, x_n - u_n(x) \cdot \gamma_n(t)) \|_{L^p_x}$

(61) $= \| \| F(x_1 - u_1(x) \cdot \gamma_1(t), \ldots, x_n - u_n(x) \cdot \gamma_n(t)) \|_{L^p_x \times \mathbb{R}^n} = \| F \|_p,$

independent of $t$. Here
$$
\| G \|_{L^p_x \times \mathbb{R}^n} := \int \| F(x_1, \ldots, x_i, \ldots, x_n) \|_p \, dx_1 \ldots dx_n
$$
where the integration is not taken over the $i$th coordinate vector.

The proof of Theorem 1.7 is complete.

6. **Application Two: Euclidean Ramsey Theory**

We turn to the proof of Theorem 1.9 which we accomplish in steps. For ease of presentation, we will excise $C \lesssim 1$ many scales; in particular, in what follows, we shall assume that (29) and (29) hold for all $k \geq 1$.

Our first order of business is to establish Theorem 1.9 in the case when our set $A$ has positive measure; this follows by specializing $f = 1_A$ in the Proposition 6.1 below.
This proposition is a multi-linear analogue of Bourgain’s work \[1\] on non-linear Roth theorems and its subsequent extension \[4\].

**Proposition 6.1.** Suppose that \( \gamma \) is relatively curved and that \( 0 \leq f \leq 1_{[0,1]} \) satisfies

\[
\int f \geq \epsilon > 0.
\]

Then there exists some \( k \leq \epsilon^{-10n^2} \) so that

\[
\int f_0 \cdot B_{-k}(f_1, \ldots, f_n) \gtrsim \epsilon^{n+1}
\]

for any \( f_i = \varphi_{-t_i} * f, \infty \geq t_i > 0 \).

Let

\[
(62) \quad B_0(f_1, \ldots, f_n)(x) := \int_0^1 \prod_{i=1}^n f_i(x - t^i) \, dt
\]

denote the multi-linear averaging operator with a rough cut-off, \( 1_{[0,1]} \), instead of the bump function, \( \rho \). By re-scaling the previous proposition, we arrive at the following corollary, which we will use below.

**Corollary 6.2.** Suppose that \( 0 \leq f \leq M \), and that \( \int f = 1 \). Then

\[
\int f_0 \cdot B_0(f_1, \ldots, f_n) \gtrsim 2^{-M10n^2}
\]

whenever \( f_i = f * \varphi_{-t_i} \) as above.

The key ingredient in our proof of Proposition 6.1 and thus Corollary 6.2 is the following lemma, which we will iterate at many scales.

**Lemma 6.3.** Let \( k \geq 0 \), and suppose that \( 0 \leq f_0, f_1, \ldots, f_n \leq 1_{[0,1]} \) have \( \int f_i \geq \epsilon \), but that

\[
\int f_0 \cdot B_{-k}(f_1, \ldots, f_n) \ll \epsilon^{n+1}.
\]

Further, assume that

\[
\int f_0 \cdot \prod_{i=1}^n \varphi_{s-ka_i} * f_i \gtrsim \epsilon^{n+1}
\]

for \( s \lesssim \log(1/\epsilon) \).

Then there exists some index \( 1 \leq i \leq n \) and some \( \|l\| \lesssim \log(1/\epsilon) \) so that

\[
\|\psi_{l-ka_i} * f_i\|_n \gtrsim \log^{-n}(1/\epsilon) \cdot \epsilon^{n+1}.
\]

\(^2\)We interpret \( \varphi_{-\infty} = \delta \), the point-mass at the origin.
Proof. Suppose that

$$
\int f_0 \cdot B_{-k}(f_1, \ldots, f_n) \ll \epsilon^{n+1}.
$$

With $s = C \cdot \log(1/\epsilon)$ for some sufficiently large $C$, for any $0 \leq x \leq 1$, we may expand

$$
B_{-k}(f_1, \ldots, f_n)(x) = \prod_{i=1}^{n} \varphi_{s - k \alpha_i} * f_i(x)
+ \sum_{|s_i| \leq s} B_{-k}(\psi_{s_1 - k \alpha_1} * f_1, \ldots, \psi_{s_n - k \alpha_n} * f_n)(x)
+ O(\epsilon^C),
$$

where we have used Lemma 4.2 to bound the pointwise error by $\lesssim \epsilon^C$.

Using our assumption and subtracting appropriately yields the lower bound on

$$
\epsilon^{n+1} \lesssim \sum_{|s_i| \leq s} \int |f_0| \cdot |B_{-k}(\psi_{s_1 - k \alpha_1} * f_1, \ldots, \psi_{s_n - k \alpha_n} * f_n)|
 \lesssim \sum_{|s_i| \leq s} \|B_{-k}(\psi_{s_1 - k \alpha_1} * f_1, \ldots, \psi_{s_n - k \alpha_n} * f_n)\|_1
$$

from which the result follows from pigeon-holing. \hfill \square

We also need the following technical lemma.

Lemma 6.4. For any $0 \leq f \leq 1_{[0,1]}$, any $k_i \geq 0$ and any $n$

$$
\int \prod_{i=0}^{n} \varphi_{-k_i} * f \gtrsim \left( \int f \right)^{n+1}.
$$

Proof. Following the approach of [4], we minorize the convolution operators $\varphi_k * f$ with the conditional expectation operators

$$
\mathbb{E}_k f := \sum_{|I| = 2^k \text{ dyadic}} \left( \frac{1}{I} \int_I f \right) 1_I,
$$

at which point the result follows by induction. \hfill \square

Proof of Proposition 6.1. We proceed by contradiction, and assume that for each $k \leq K$ no lower bound held

$$
\int f_0 \cdot B_{-k}(f_1, \ldots, f_n) \ll \epsilon^{n+1};
$$

we will show that $K \lesssim \epsilon^{-10n^2}$. 

ON MAXIMAL FUNCTIONS WITH CURVATURE

21
To do so, we use Lemma 6.4 and Lemma 6.3 to extract an index $1 \leq i \leq n$, and a sparse subset of scales $X \subset \{k \leq K\}$ separated by $\gg \log(1/\epsilon)$, of size

$$|X| \gtrsim \frac{K}{n \cdot \log(1/\epsilon)},$$

with the following property:

For each $k \in X$ there exists some perturbation $|s_k| \lesssim \log(1/\epsilon)$ so that

$$\|\psi_{s_k - k\alpha_i} \ast f_i\|_n \gtrsim \epsilon^{n+1}/\log^n(1/\epsilon).$$

In particular, taking a $\ell^n$ sum of (64) yields the upper bound

$$\epsilon^{-n^2} \cdot |X| \lesssim \sum_{k \in X} \|\psi_{s_k - k\alpha_i} \ast f_i\|_n \lesssim \|Sf_i\|_n \lesssim \|f_i\|_n \lesssim 1,$$

by the boundedness of the Littlewood-Paley square function, (49). In particular, we have exhibited the desired upper bound, $K \lesssim \epsilon^{-10n^2}$. $\square$

We now turn to the proof of Theorem 1.9. Before doing so, we recall that whenever a set $A$ has Hausdorff dimension $> \beta$,

$$\dim_H(A) > \beta,$$

one may find a $\beta$-dimensional Frostman measure, $\mu$, supported on $A$: a probability measure supported on $A$, so that $\mu$ satisfies the ball growth condition,

$$\mu(I) \lesssim |I|^\beta$$

for each interval $I \subset A \cap [0, 1]$. By our normalizing assumption (6), we may select one such $\mu$ so that for each $k$,

$$\|\varphi_{-k} \ast \mu\|_{\infty} \lesssim \Lambda \cdot 2^k(1-\beta)$$

for each $k \geq 0$; the implicit constant is determined only by the Schwartz normalization that we have imposed on $\{\varphi\}$, see the subsection on notation, §1.2.

The Proof of Theorem 1.9. With $\mu$ an appropriate Frostman measure, set

$$f := \varphi_{-J} \ast \mu$$

for some sufficiently large $J$. It suffices to exhibit upper and lower bounds – independent of $J$ – for

$$\int f : \overline{B}_0(f, \ldots, f),$$

where $\overline{B}_0$ is the multi-linear averaging operator with the rough cut-off, (62).
To do so, with \( l \lesssim 1 \) a sufficiently large integer to be determined, decompose

\[
B_0(f, \ldots, f) = \sum_{s > l} \left( \sum_{s_0 = s} B_0(\psi_{-s_1} * f, \ldots, \psi_{-s_n} * f) \right) + B_0(\varphi_{-l} * f, \ldots, \varphi_{-l} * f)
\]

\[
= \sum_{s > l} B_s + B_l,
\]

so that upon taking inner products, we may express (65) as

\[
(65) = \langle \varphi_{-2n \cdot l} * f, B_l \rangle + \sum_{s > l} \langle \varphi_{-2n \cdot s} * f, B_s \rangle + \sum_{s \geq l} \langle f - \varphi_{-2n \cdot s} * f, B_s \rangle.
\]

By Proposition 2.2 we may sum

\[
\sum_{s > l} |\langle \varphi_{-2n \cdot s} * f, B_s \rangle| \lesssim \sum_{s > l} \|\varphi_{-2n \cdot s} * f\|_\infty \cdot \|B_s\|_1
\]

\[
\lesssim \sum_{s > l} 2^{2n \cdot s(1-\beta)} \cdot 2^{-cs} \cdot \|\varphi_{-s} * f\|_n^n
\]

\[
\lesssim \sum_{s > l} 2^{2n \cdot s(1-\beta)} \cdot 2^{-cs} \cdot \|\varphi_{-s} * f\|_n^{n-1} \cdot \|\varphi_{-s} * f\|_1
\]

\[
\lesssim \sum_{s > l} 2^s \left( n \cdot 2n(1-\beta) - c \right)
\]

\[
\lesssim 2^{-c/2l},
\]

where \( c \) is the constant appearing in the diagonal case of the Proposition, and we assume that \( \beta \) is sufficiently close to 1 that

\[
n^2 \cdot 2 \cdot (1 - \beta) < c/2.
\]

In particular, there exists \( c \) so that for all \( \beta \) sufficiently close to 1 we have exhibited

\[
(65) = \langle \varphi_{-2n \cdot l} * f, B_l \rangle + O \left( 2^{-c/2l} \right) + \sum_{s \geq l} \langle f - \varphi_{-2n \cdot s} * f, B_s \rangle.
\]

We claim that

\[
\langle f - \varphi_{-2n \cdot s} * f, B_s \rangle = O(2^{-cs}),
\]

so that the final sum in (66) is \( O(2^{-cl}) \). We just estimate

\[
|\langle f - \varphi_{-2n \cdot s} * f, B_s \rangle| \leq \sum_{r \geq 2n \cdot s} \left| \prod_{i=0}^n |f(v_i)| \int \psi_{-r}(x - v_0)K_{\vec{s}}(x - v_1, \ldots, x - v_n) \, dx \right| \, dv
\]

\[
\leq \sum_{r \geq 2n \cdot s} \sup_{|v_1|, \ldots, |v_n| \leq 1} \left| \int \psi_{-r}(x)K_{\vec{s}}(x - v_1, \ldots, x - v_n) \, dx \right|.
\]
By Fourier inversion, we dominate
\[
\left| \int \psi_r(x) \cdot K_s(x - v_1, \ldots, x - v_n) \, dx \right|
\leq \int \left| \hat{\psi} \left( -2^{-r} \cdot \sum_i \xi_i \right) \right| \cdot \prod_{i=1}^n |\psi(2^{-s_i} \xi_i)| \, d\xi
\leq 2((n+1)s - r) \leq 2^{-r/10},
\]
since \((n + 1) \cdot s \leq 9r/10\). The mean-zero nature of \(\psi\) presents in the low-frequency estimate,
\[
\hat{\psi}(\zeta) = \int e(-\zeta x) \cdot \psi(x) \, dx = \int (1 - e(-\zeta x)) \cdot \psi(x) \, dx = O(|\zeta|)
\]
by the mean-value theorem; the implicit constant is uniform over appropriately normalized \(\psi\).

In particular, we may refine (66) to
\[
\langle \varphi_{-2n-l} \ast f, B_l \rangle = O \left( 2^{-c/2l} \right).
\]

As far as convergence is concerned, an upper bound for the first term is given by
\[
\| \varphi_{-2n-l} \ast f \|_\infty \lesssim \Lambda \cdot 2^{2n-l(1-\beta)},
\]
which is in particular bounded independent of \(J\). As for our lower bound, we have
\[
\langle \varphi_{-2n-l} \ast f, B_l \rangle \gtrsim 2^{-\Lambda^{10n^2} \cdot 2^{20n^3} \cdot (1-\beta) \cdot l}
\]
by applying Corollary 6.2.

In particular, we have shown that
\[
\langle \varphi_{-2n-l} \ast f, B_l \rangle \gtrsim 2^{-\Lambda^{10n^2} \cdot 2^{20n^3} \cdot (1-\beta) \cdot l} - O(2^{-c/2l});
\]
since
\[
2^{-\Lambda^{10n^2}} = \lim_{\beta \to 1^-} 2^{-\Lambda^{10n^2} \cdot 2^{20n^3} \cdot (1-\beta) \cdot l} \gg 2^{-c/2l},
\]
we conclude a lower bound on (65) for \(\beta < \dim_H A\) sufficiently close to 1, provided we choose \(l \gg \Lambda 1\) sufficiently large.

The proof is complete. \(\square\)

6.1. Final Remarks. The methods described allow one to treat (maximal truncations of) the singular integral formulation of Theorems 1.4, 1.5, 1.6, 1.7, as the same scale type decay persists in the high-frequency regime, see Lemma 2.4 and Proposition 2.2. In the instance when at least one function is smoother than the appropriate spatial scale, a downwards induction is available, as per the uncertainty principle Lemma 4.2.
ON MAXIMAL FUNCTIONS WITH CURVATURE

References

[1] J. Bourgain. A nonlinear version of Roth’s theorem for sets of positive density in the real line. J. Analyse Math. 50 (1988), 169181.

[2] V. Chang, I. Laba, M. Pramanik. Finite configurations in sparse sets. J. Anal. Math. 128 (2016), 289335.

[3] C. Demeter, T. Tao, C. Thiele. Maximal multilinear operators. Trans. Amer. Math. Soc. 360 (2008), no. 9, 49895042.

[4] P. Durcik, S. Guo, J. Roos. A polynomial Roth theorem on the real line. Trans. Amer. Math. Soc. 371 (2019), no. 10, 69736993.

[5] P. Erdos. On a family of symmetric Bernoulli convolutions, Amer. J. Math. 61 (1939), 974-975.

[6] A. Gaitan, V. Lie. The Boundedness of the (Sub-)Bilinear Maximal Function Along “Non-Flat” Smooth Curves. Preprint, https://arxiv.org/pdf/1902.03807.pdf.

[7] S. Guo, J. Hickman, V. Lie, J. Roos. Maximal operators and Hilbert transforms along variable non-flat homogeneous curves. Proc. Lond. Math. Soc. (3) 115 (2017), no. 1, 177219.

[8] T. Keleti A 1-dimensional subset of the reals that intersects each of its translates in at most a single point. Real Anal. Exchange, 24(2):843844, 1998/99.

[9] T. Keleti. Construction of one-dimensional subsets of the reals not containing similar copies of given patterns. Anal. PDE, 1(1):2933, 2008.

[10] B. Krause. A Non-Linear Roth Theorem for Fractals of Sufficiently Large Dimension. Preprint, https://arxiv.org/abs/1904.10562.

[11] B. Krause, M. Mirek, T. Tao. On pointwise convergence of bilinear ergodic averages along polynomial orbits. In preparation.

[12] I. Laba, M. Pramanik. Arithmetic progressions in sets of fractional dimension. Geom. Funct. Anal. 19 (2009), no. 2, 429456.

[13] M. Lacey. The bilinear maximal functions map into $L^p$ for $2/3 < p \leq 1$. Ann. of Math. (2) 151 (2000), no. 1, 3557.

[14] X. Li. Bilinear Hilbert transforms along curves, I, Analysis & PDE Vol. 6 (2013), No. 1, 197-220.

[15] X. Li. L. Xiao. Uniform estimates for bilinear Hilbert transforms and bilinear maximal functions associated to polynomials. Amer. J. Math. 138 (2016), no. 4, 907962.

[16] V. Lie. On the boundedness of the Bilinear Hilbert transform along non-flat smooth curves. American Journal of Mathematics, 137(2):313363, 2015.

[17] V. Lie. A Unified Approach to Three Themes in Harmonic Analysis. Preprint, https://arxiv.org/pdf/1902.03807.pdf

[18] C. Muscalu, W. Schlag. Classical and multilinear harmonic analysis. Vol. II. Cambridge Studies in Advanced Mathematics, 138. Cambridge University Press, Cambridge, 2013. xvi+324 pp.

[19] E. Stein, E Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals. Princeton Mathematical Series, 43. Monographs in Harmonic Analysis, III. Princeton University Press, Princeton, NJ, 1993.

[20] E. Stein. Oscillatory integrals related to Radon-like transforms. Proceedings of the Conference in Honor of Jean-Pierre Kahane (Orsay, 1993). J. Fourier Anal. Appl. 1995, Special Issue, 535551.

[21] E. Stein and S. Wainger. Oscillatory integrals related to Carleson’s theorem. Math. Res. Lett., 8(5-6):789–800, 2001.

Department of Mathematics, Princeton, Princeton, NJ, 08540

E-mail address: benkrause2323@gmail.com