L^p-IMPROVING PROPERTIES OF AVERAGES ON POLYNOMIAL CURVES AND RELATED INTEGRAL ESTIMATES

Philip T. Gressman

ABSTRACT. In the combinatorial method proving of L^p-improving estimates for averages along curves pioneered by Christ [8], it is desirable to estimate the average modulus (with respect to some uniform measure on a set) of a polynomial-like function from below using only the value of the function or its derivatives at some prescribed point. In this paper, it is shown that there is always a relatively large set of points (independent of the particular function to be integrated) for which such estimates are possible. Inequalities of this type are then applied to extend the results of Tao and Wright [29] to obtain endpoint restricted weak-type estimates for averages over curves given by polynomials.

The purpose of this paper is twofold. First, a somewhat surprising inequality for $L^1(\mu)$-norms of polynomial-type functions on the real line will be established. The most important special case of this inequality is as follows:

Theorem 1. *Suppose that $K \subset \mathbb{R}$ is measurable. For any positive integer n and any $0 < \epsilon < 1$, there is an interval I with $|K \cap I| \geq \frac{1-\epsilon}{n}|K|$ (here $|\cdot|$ acting on sets denotes Lebesgue measure) and a constant $c_{n,\epsilon}$ such that*

$$\text{(1)} \qquad \int_K |p(t)|dt \geq c_{n,\epsilon}|K|^{j+1} \sup_{t \in I} |p^{(j)}(t)|$$

for any polynomial p of degree n or less and any $j = 0, \ldots, n$.

It is fairly trivial to show that such an interval I can always be found when p is given, but somewhat unexpected that I exists independently of p. It is also fairly straightforward to construct a family of counterexamples to this inequality if one takes $\epsilon = 0$, for example (which suggests that there must be a certain amount of subtlety involved in proving the positive result). This theorem and its generalization to probability measures μ will be taken up in the first section.

The second major purpose of this paper is to apply the inequality (1) to answer a question of Tao and Wright [29] concerning L^p-improving bounds for averages along curves. In that work, inequalities of the form (1) appear naturally while carrying out the combinatorial methods of Christ [8]. Tao and Wright were able to establish (1) in the case when K has some additional structure (K was taken to be a central set of width w, see their lemma 7.3); the price of requiring such structure was certain losses in exponents which made it impossible to obtain L^p-improving inequalities on the boundary of the type set. Interpreted in the framework of that paper, theorem 1 indicates that the set K does not need any significant structure (namely, K does not need to be a central set of width w) for the desirable inequality (1) to hold. As a consequence, theorem 1 alone makes it possible to prove the full range of restricted

---

weak-type estimates for one-dimensional averaging operators given by polynomial curves (that is, when both the averaging operator and the dual operator can be described by polynomial functions in appropriate coordinate systems).

This result is formulated in the standard bilinear way as follows: suppose U is an open ball in $\mathbb{R}^{d+1}$ and one is given projections $\pi_1 : U \to \mathbb{R}^d$ and $\pi_2 : U \to \mathbb{R}^d$ such that the differentials $d\pi_1$ and $d\pi_2$ are surjective at every point. The Radon-like operator R associated to these projections is defined by duality as

$$(2) \qquad \int_{\mathbb{R}^d} Rf(y)g(y)dy := \int_U f(\pi_1(x))g(\pi_2(x))\psi(x)dx$$

where ψ is some bounded cutoff function (not necessarily smooth) supported in U. Next, let X_1 and X_2 be vector fields on U which are nonvanishing and satisfy $d\pi_1(X_1) = 0 = d\pi_2(X_2)$, and suppose that for all words $w = (w_1, w_2, \ldots, w_k)$ (each w_j equals 1 or 2) of sufficient length, the commutator

$$X_w := [X_{w_1}, [X_{w_2}, [\cdots [X_{w_{k-1}}, X_{w_k}] \cdots]]]$$

vanishes identically on U. This geometric condition guarantees that both R and R^* are given by averages over polynomial curves. Under these conditions the following theorem holds:

Theorem 2. *Suppose that the vector fields X_1, X_2 are as assumed above. Let $x_0 \in U$, and consider the mapping*

$$\Phi_{x_0}(t_1, \ldots, t_{d+1}) := \exp(t_1 X_1) \circ \cdots \circ \exp(t_{d+1}X_{d+1})(x_0)$$

(where the periodicity convention $X_{j+2} = X_j$ is used). Let $J_{x_0}(t)$ be the Jacobian determinant of this mapping (as a function of the parameters t). If $\partial_t^\alpha J_{x_0}(t) \neq 0$ at $t = 0$ for some multiindex $\alpha = (\alpha_1, \ldots, \alpha_{d+1})$, then the averaging operator R given by (2) satisfies a restricted weak-type estimate

$$\left| \int_U \chi_F(\pi_1(x))\chi_G(\pi_2(x))\psi(x)dx \right| \leq C|F|^{\frac{1}{p_1}}|G|^{\frac{1}{p_2}}$$

when the support of ψ is a sufficiently small neighborhood of x_0 and the exponents p_1 and p_2 are given by $\frac{1}{p_1} := \frac{A_1}{A_1+A_2-1}, \frac{1}{p_2} := \frac{A_2}{A_1+A_2-1}$, where

$$A_1 := \left\lceil \frac{d+1}{2} \right\rceil + \alpha_1 + \cdots + \alpha_{2\lceil \frac{d+1}{2} \rceil - 1}, \quad A_2 := \left\lfloor \frac{d+1}{2} \right\rfloor + \alpha_2 + \cdots + \alpha_{2\lfloor \frac{d+1}{2} \rfloor}.$$

It should be noted that the hypotheses of theorem 2 (regarding the derivatives of the Jacobian determinant) are, in fact, equivalent to the Hörmander-type hypotheses used by Tao and Wright (see lemma 9.1 in [29] or sections 9 and 10 of the paper of Christ, Nagel, Stein, and Wainger [11] in which the double fibration curvature formulation $(\mathcal{C}_\Lambda)$ is shown to be equivalent to the Jacobian determinant formulation $(\mathcal{C}_J)$). In light of this equivalence, theorem 2 successfully establishes restricted weak-type estimates on the boundary of the type set of the operator (2) which were just missed in [29].

Regarding integral estimates and the related issue of sublevel sets, earlier results of particular interest to the problem at hand include the work of Carbery and Wright [6], Carbery, Christ, and Wright [5], Phong, Stein, and Sturm [24], and Phong and Sturm [25], as well as many others. In this paper, the attention will be exclusively focused

on one-dimensional estimates, using methods similar to those employed by Carbery, Christ and Wright [5] who built upon ideas of Arhipov, Karacuba and Čubarikov [1]. Similar one-dimensional methods have also been employed by Rogers [26] to obtain sharp constants for sublevel set estimates and van der Corput's lemma.

In the case of Radon-like transforms and averaging operators, the reader is referred to the papers of Tao and Wright [29] and Christ, Nagel, Stein, and Wainger [11] for more complete lists of references. In this paper, the argument to be followed was originally devised by Christ [8]. Tao and Wright [29] made important additions to the Christ argument which will, of course, be necessary to use here as well. More recently, these ideas have been adapted to a variety of other contexts by Christ and Erdoğan [9], [10], Bennett, Carbery, Christ, and Tao [3], Bennett, Carbery, and Wright [4], Erdoğan and R. Oberlin [12], and many others. Earlier approaches to L^p-improving estimates for averaging operators based on oscillatory integral estimates, beginning with Littman [16], Phong and Stein [20], [21], [22], [23], including Greenleaf and Seeger [14], [15], Seeger [27], [28], and D. Oberlin [17], [18], [19], will not be discussed here.

1. Estimation of integrals by pointwise values

To begin this section, a number of definitions are in order. First, suppose K is a closed set contained in an open interval I. For each nonnegative integer n, a function $f \in C^n(I)$ is said to be of polynomial type n on (K, I) when $f^{(n)}$ does not change sign (i.e., is nonnegative or nonpositive) and there exists a finite constant C for which $\sup_{t \in I} |f^{(n)}(t)| \leq C \inf_{t \in K} |f^{(n)}(t)|$. Any polynomial of degree n on I is, of course, polynomial type n on (K, I) for any closed set K contained in I.

In general, if a probability measure μ is supported on the closed set K, it will necessary to consider functions which are of polynomial type on (K_ϵ, I) for some set K_ϵ slightly larger than K (since, if K has Lebesgue measure zero, the values of f on K are largely independent of the values of the derivatives of f on the same set K). To that end, given any closed set K and any ϵ, let K_ϵ be the union of K and the sets K_L and K_R given by

$$K_R := \{t \in I \mid \inf \{d \geq 0 \mid t + d \in K\} \leq \epsilon \inf \{d \geq 0 \mid t - d \in K\}\},$$
$$K_L := \{t \in I \mid \inf \{d \geq 0 \mid t - d \in K\} \leq \epsilon \inf \{d \geq 0 \mid t + d \in K\}\}$$

(K_R and K_L are the points which are bounded on both sides by the set K but are proportionately much closer to K on one side than the other; note that the Lebesgue measures of K_L and K_R are bounded by $\epsilon |I \setminus K|$).

The final definition needed to begin this section is a notion of the length of the set on which μ is supported: given a probability measure μ supported on an open interval $I \subset \mathbb{R}$ (it will be assumed throughout that all Borel sets are measurable), a positive integer n, and an $\epsilon \in (0, 1)$, let $|\mu|_{n,\epsilon}$ be the infimum of $\sum_{j=1}^{n} |I_j|$ over all collections of closed intervals $\{I_1, \ldots, I_n\}$ which satisfy $\mu(\bigcup_{j=1}^{n} I_j) \geq 1 - \epsilon$. Note that $|\mu|_{n,\epsilon}$ is decreasing in n, increasing in ϵ, and $|\mu_K|_{n,\epsilon} \geq (1 - \epsilon)|K|$ when, for example, μ_K is normalized Lebesgue measure on the set K.

A remark about notation: the two parameters having already appeared, namely n and ϵ, appear in essentially every inequality to come; in particular, most proportionality constants will vary as these parameters vary. When the nature of these constants

is uninteresting or unimportant, the notation $A \lesssim B$ will be used to indicate that there is a constant $C_{n,\epsilon}$ completely determined by ϵ and n such that $A \leq C_{n,\epsilon}B$.

The main theorems of this section can now be stated. The first is as follows:

Theorem 3. *Let K be a closed set contained in an open interval $I \subset \mathbb{R}$ (possibly infinite). For any $\epsilon \in (0,1)$, let K_ϵ be the union of K with K_R and K_L where For any positive integer n, any Borel probability measure μ supported on K, and any $f \in C^n(I)$ for which $f^{(n)}$ does not change sign, it must be true that*

$$\int |f(t)|d\mu(t) \gtrsim |\mu|_{n,\epsilon}^n \inf_{t \in K_\epsilon} |f^{(n)}(t)|.$$

Theorems of this type are not new; see, for example Carbery, Christ, and Wright [5], Arhipov, Karacuba and Čubarikov [1], or Rogers [26]. The main new feature is the presence of $|\mu|_{n,\epsilon}$ on the right-hand side; in most previous cases μ is assumed to be the uniform measure on some set K and $|\mu|_{n,\epsilon}$ is replaced by $|K|$. To prove the full uniform estimate (theorem 4 and its corollaries), it is necessary to distinguish the length $|\mu|_{n,\epsilon}$ from the measure of the support of μ.

The second theorem of this section establishes uniform integral estimates from below by a supremum of the function on a set E which depends only on the class of functions to which f belongs. This is, of course, the most difficult task necessary to establish any result along the lines of theorem 1:

Theorem 4. *Suppose μ is a Borel probability measure supported on some closed $K \subset I$; fix some positive integer n and $\epsilon \in (0,1)$. There exists a closed set $E \subset I$ with at most n connected components for which $\mu(E) \geq 1 - \epsilon$ and*

$$\int |f(t)|d\mu(t) \gtrsim c_{n,\epsilon}C^{-1} \sup_{t \in E} |f(t)|$$

for any function f which is of polynomial type n on (K_ϵ, I) with constant C.

In particular, an immediate corollary of theorems 3 and 4 is the following:

Corollary 1. *Given a Borel probability measure μ supported on $K \subset I$ and ϵ and n as above, there exists a closed interval I' (possibly a single point) with $\mu(I') \geq \frac{1-\epsilon}{n}$ such that, for any function f which is polynomial type n on (K_ϵ, I) with constant C,*

$$\int |f(t)|d\mu(t) \gtrsim C^{-1} \min\{|I'|^j, |\mu|_{n,\epsilon}^j\} \sup_{t \in I'} |f^{(j)}(t)|$$

for any $j = 0, \ldots, n$.

Theorem 1 from the introduction follows from this corollary when f is polynomial of degree n and μ is the uniform measure on K.

1.1. Combinatorial considerations. In what follows, let V_n be the Vandermonde polynomial in n variables, i.e., $V_n(t_1, \ldots, t_n) := \prod_{j>i}(t_j - t_i)$. The proof begins with a more detailed look at a standard idea: the estimation of higher derivatives of a function f via sampling at a finite number of points. This is typically carried out using Lagrange interpolating polynomials; in the process, the structure of the "remainder term" can explored as well. The result of this work is contained in the following proposition; it should be noted that essentially the same result appears in

recent work of Bak, Oberlin, and Seeger (equation (2.10)) in work on the Fourier restriction problem [2].

Proposition 1. *Let $t_1, \ldots, t_{n+1}$ be distinct points in some interval I. Given these points, there exists a nonnegative function $\psi_t(s)$, supported on $[t_1, t_{n+1}]$, with total integral 1 such that*

$$(3) \qquad \sum_{i=1}^{n+1} (-1)^{n+1-i} f(t_i) \frac{V_n(t_1, \ldots, \hat{t}_i, \ldots, t_{n+1})}{V_{n+1}(t_1, \ldots, t_{n+1})} = \frac{1}{n!} \int f^{(n)}(s) \psi_t(s) ds$$

for any $f \in C^{(n)}(I)$ (here $\hat{t}_i$ indicates that t_i is omitted). This function ψ_t has the following properties: (1) ψ_t is supported on the convex hull of the t_j's and has integral 1, (2) ψ_t is a polynomial of degree at most $n-1$ on each interval containing none of the t_j's, and (3) $\psi_t \in C^{(n-2)}(I)$ when $n \geq 2$.

Proof. Let $a \in I$ be such that $a \leq \min_j t_j$. If f is any n-times continuously differentiable function on I, let

$$g(t) := \int_a^t \frac{(t-s)^{n-1}}{(n-1)!} f^{(n)}(s) ds.$$

It is straightforward to check that g is also n-times differentiable when $t > a$ and $g^{(n)}(t) = f^{(n)}(t)$ there (g is nothing more than the remainder term for the degree $n-1$ Taylor polynomial at a). This implies that $f - g$ is a polynomial of degree $n-1$ (the Taylor polynomial at a); therefore it must be the case that the determinant

$$(4) \qquad \begin{vmatrix} f(t_1) - g(t_1) & 1 & t_1 & \cdots & t_1^{n-1} \\ f(t_2) - g(t_2) & 1 & t_2 & \cdots & t_2^{n-1} \\ \vdots & & \vdots & \vdots & \ddots & \vdots \\ f(t_{n+1}) - g(t_{n+1}) & 1 & t_{n+1} & \cdots & t_{n+1}^{n-1} \end{vmatrix}$$

vanishes. Using Cramer's rule and the fact that the minors corresponding to entries in the first column are Vandermonde matrices, it follows that

$$\sum_{i=1}^{n+1} (-1)^i (f(t_i) - g(t_i)) V_n(t_1, \ldots, \hat{t}_i, \ldots, t_{n+1}) = 0.$$

since the n-th derivative of the difference vanishes at every point of $[t_1, t_{n+1}]$. But this implies that

$$\sum_{i=1}^{n+1} (-1)^i f(t_i) V_n(t_1, \ldots, \hat{t}_i, \ldots, t_{n+1}) = \sum_{i=1}^{n+1} (-1)^i g(t_i) V_n(t_1, \ldots, \hat{t}_i, \ldots, t_{n+1})$$

$$= \int_a^\infty f^{(n)}(s) \frac{1}{(n-1)!} \sum_{t_i > s} (-1)^i (t_i - s)^{n-1} V_n(t_1, \ldots, \hat{t}_i, \ldots, t_{n+1}) ds.$$

When $s \geq \max_j t_j$, the sum inside the integral vanishes because no terms are included. When $s \leq \min_j t_j$, on the other hand, the sum again vanishes since it represents the determinant of a Vandermonde-type matrix whose first column is given by $(t_i - s)^{n-1}$

(which vanishes just like (4)). To compute the integral of this function, it suffices to plug in $f(s) := \frac{s^n}{n!}$:

$$\sum_{i=1}^{n+1} (-1)^i \frac{t_i^n}{n!} V_n(t_1, \ldots, \hat{t}_i, \ldots, t_{n+1}) = \frac{(-1)^{n+1}}{n!} V_{n+1}(t_1, \ldots, t_{n+1}).$$

Thus, one is led to define

$$\psi_t(s) := \frac{n(-1)^{n+1}}{V_{n+1}(t_1, \ldots, t_{n+1})} \sum_{t_i > s} (-1)^i (t_i - s)^{n-1} V_n(t_1, \ldots, \hat{t}_i, \ldots, t_{n+1})$$

$$= \sum_{t_i > s} \frac{n(t_i - s)^{n-1}}{\prod_{j \neq i}(t_i - t_j)}.$$

This ψ_t has integral one and satisfies the correct integral identity. Furthermore, it follows directly from this definition that ψ_t is piecewise a polynomial of degree at most $n-1$ on all intervals not containing any t_j's. It is also immediate that $\psi_t \in C^{(n-2)}(I)$ when $n \geq 2$ because it is a finite linear combination of such functions (i.e., functions equal to $(t_i - s)^{n-1}$ when $s \leq t_i$ and equal to 0 otherwise). It remains to show that ψ_t is nonnegative. If this were not the case, it would be possible to find a function $f \in C^{(n)}(I)$ such that $f^{(n)}$ is strictly positive on I but the right-hand side of (3) is zero. Examining (4), this is possible only when there is a polynomial of degree $n-1$ which agrees with this function f at $t_1, \ldots, t_{n+1}$. Repeated applications of Rolle's theorem shows that this cannot be the case (i.e., the n-th derivative of f must vanish at some point if f agrees with a polynomial of degree $n-1$ at $n+1$ distinct points). $\square$

It is perhaps worth noting that the three properties of ψ_t (support and normalization, piecewise polynomial of degree $n-1$, and global $C^{(n-2)}$ regularity) uniquely determine ψ_t. Even when the normalization condition is dropped, there is still only a one-dimensional family of such functions, namely, multiples of ψ_t; this is to say that there are no nontrivial piecewise functions satisfying these conditions and having integral 0. The proof of this fact proceeds inductively on n in a fairly standard way.

The advantage gained in finding ψ_t more-or-less explicitly is that it allows one to improve upon the trivial estimate from below on $\int f^{(n)}(s)\psi_t(s)ds$ to exploit the fact that, while ψ_t is not supported at the points t_j, it must always, in fact, have some positive proportion of its mass which lies near the t_j's. In other words, if the t_j's happen to be separated by some large distance, it never occurs that an overwhelming fraction of the mass of ψ_t is concentrated inside that gap. This fact will be made precise during the proof of theorem 3.

Proposition 2. *For any Borel probability measure μ on an interval $I \subset \mathbb{R}$ and any positive integer n, let*

$$(\ell_n(\mu))^n := \frac{\int |V_{n+1}(t_1, \ldots, t_{n+1})| d\mu(t_1) \cdots d\mu(t_{n+1})}{\int |V_n(t_1, \ldots, t_n)| d\mu(t_1) \cdots d\mu(t_n)}.$$

The quantity $\ell_n(\mu)$ is zero if and only if μ is supported on a set of n or fewer points. Furthermore, given any $\epsilon > 0$, there exists a finite collection of at most n closed, disjoint intervals I_j (possibly length zero) such that $\mu(\bigcup_j I_j) \geq 1 - \epsilon$, $|\bigcup_j I_j| \lesssim \ell_n(\mu)$

(that is, the Lebesgue measure of the union) and $\mu(I_j) \gtrsim 1$. These intervals will be called the (n, ϵ)-children of μ.

Proof. First of all, it is necessarily true that $\ell_n(\mu) = 0$ if the mass of μ is supported on a finite set of n or fewer points, since in this case the Vandermonde polynomial V_{n+1} vanishes almost everywhere on the $(n+1)$−fold product of μ. In all other cases, the distribution function $\mu((-\infty, t] \cap I)$ must take at least $n + 1$ distinct, nonzero values, meaning that the interval I may be partitioned into at least $n + 1$ disjoint pieces, each of which has nonzero μ-measure. Since V_{n+1} does not vanish when each t_i belongs to a distinct element of the partition, the integral cannot be zero.

Assuming now that μ is not supported on a set of n points, consider the ratio

$$\frac{|V_{n+1}(t_1, \ldots, t_{n+1})|}{|V_n(t_1, \ldots, t_n)|} = \prod_{i=1}^{n} |t_n - t_i|.$$

If $E_{t_1, \ldots, t_n}$ is the set $\{ t_{n+1} \mid \prod_{i=1}^{n} |t_n - t_i| \leq 2\epsilon^{-1}(\ell_n(\mu))^n \}$, it follows that

$$\int \frac{|V_{n+1}(t_1, \ldots, t_{n+1})|}{|V_n(t_1, \ldots, t_n)|} d\mu(t_{n+1}) \geq (1 - \mu(E_{t_1, \ldots, t_n})) 2\epsilon^{-1}(\ell_n(\mu))^n$$

for each possible ensemble $t_1, \ldots, t_n$ of distinct points. Multiplying both sides by $|V_n(t_1, \ldots, t_n)|$ and integrating $d\mu(t_1) \cdots d\mu(t_n)$, it follows that

$$(\ell_n(\mu))^n \geq \inf_{t_1, \ldots, t_n \in I} \epsilon^{-1}(1 - \mu(E_{t_1, \ldots, t_n}))(\ell_n(\mu))^n$$

so there must be a choice of $t_1, \ldots, t_n$ for which $\mu(E_{t_1, \ldots, t_n}) \geq 1 - \frac{\epsilon}{2}$. This sublevel set consists of at most n closed connected components. The Lebesgue measure of the sublevel set is at most $2n\ell_n(\mu)$ (since the sublevel condition requires, in particular, that t_{n+1} must be within distance $\ell_n(\mu)$ of at least one of the t_j's for $j = 1, \ldots, n$). Let the (n, ϵ)-children of μ be the connected components of $E_{t_1, \ldots, t_n}$ whose μ measure is at least $\frac{\epsilon}{2n}$. Clearly they are disjoint, have bounded lengths, and the μ measure of the union is at least $1 - \epsilon$. $\qquad\square$

Proposition 3. *Let $t_1, \ldots, t_N$ be points in some interval I. For each positive integer n, there is a closed set $E_n \subset I$ which consists of no more than n connected components, contains t_j for $j = 1, \ldots, N$ and satisfies*

$$\sup_{t \in E_n} |f(t)| \leq (n+1)2^n \max_{j=1,\ldots,N} |f(t_j)|$$

for any $f \in C^n(I)$ whose n-th derivative does not change sign.

Proof. Without loss of generality, it suffices to assume that $t_1 < t_2 < \cdots < t_N$ and that $N \geq n + 1$. It is also permissible to assume that $f^{(n)}$ is nonnegative.

Consider first the case $N = n + 1$. Fix $t_1 < t_2 < \cdots < t_{n+1}$ and fix k to be the index which maximizes $V_n(t_1, \ldots, \hat{t_k}, \ldots, t_{n+1})$; notice that the index can never equal 1 or $n + 1$ (since omitting t_2 or t_n, respectively, will always increase the product). Let I' be the shorter interval of $[t_{k-1}, t_k]$ or $[t_k, t_{k+1}]$ (if they have the same length, either choice is acceptable). Given $f \in C^{(n)}(I)$, let

$$\tilde{f}(t) := f(t) - \sum_{j \neq k} f(t_j) \prod_{i \neq j, k} \frac{t - t_i}{t_j - t_i}.$$

Clearly $\tilde{f}(t_j) = 0$ for $j \neq k$. It suffices to work with $\tilde{f}$ rather than f since the values of f and $\tilde{f}$ do not differ appreciably on I' or at t_k. More precisely, at $t = t_k$ one has

$$|\tilde{f}(t_k)| \leq |f(t_k)| + \sum_{j \neq k} |f(t_k)| \frac{|V_n(t_1, \ldots, \hat{t}_j, \ldots, t_{n+1})|}{|V_n(t_1, \ldots, \hat{t}_k, \ldots, t_{n+1})|} \leq (n+1) \max_j |f(t_j)|.$$

Next, if $t \in I'$, then $\prod_{i \neq j,k} |t - t_i| \leq 2^{n-1} \prod_{i \neq j,k} |t_k - t_i|$ since $|t - t_i| \leq |t - t_k| + |t_k - t_i| \leq \min_j\{|t_k - t_j|\} + |t_k - t_i|$. It must therefore be the case that

$$|f(t)| \leq |\tilde{f}(t)| + n 2^{n-1} \max_j |f(t_j)|;$$

hence if $|\tilde{f}(t)| \leq C \max_j |\tilde{f}(t_j)|$, then $|f(t)| \leq (C(n+1) + n 2^{n-1}) \max_j |f(t_j)|$.

Now (3) implies that, for $t \in I'$, $(-1)^{n+1-k} f(t) \geq 0$ when $f^{(n)}$ is nonnegative on I by virtue of the fact that the Vandermonde polynomials are positive and $\tilde{f}(t_i)$ vanishes for $i \neq k$. Fix some $t \in I'$; let $t_1' < t_2' < \cdots < t_{n+1}'$ be the sequence of numbers obtained by replacing t_{k-1} with t when $I' = [t_{k-1}, t_k]$ or replacing t_{k+1} with t when $I' = [t_k, t_{k+1}]$ (so that $t_j' = t_j$ for all but one value of j). For this collection of points, (3) implies that

$$\sum_{j=1}^{n+1} (-1)^{n+1-j} \tilde{f}(t_j') \frac{V_n(t_1', \ldots, \hat{t}_j', \ldots, t_{n+1}')}{V_{n+1}(t_1', \ldots, t_{n+1}')} \geq 0$$

as well. Here all but two terms must vanish (by virtue of the vanishing of $\tilde{f}$). Thus

$$(-1)^{n+1-k} \tilde{f}(t_k) V_n(t_1', \ldots, \hat{t}_k', \ldots, t_{n+1}') \geq (-1)^{n+1-k} \tilde{f}(t) V_n(t_1', \ldots, \hat{t}, \ldots, t_{n+1}')$$

(where $\hat{t}$ is properly interpreted as $\widehat{t_{k-1}'}$ or $\widehat{t_{k+1}'}$ depending on which of those indices had its corresponding value replaced by t). Both sides are nonnegative, and just as before, the Vandermonde polynomial on the left-hand side is at most a factor of 2^{n-1} larger than the corresponding polynomial on the right-hand side. It must therefore be the case that $|\tilde{f}(t)| \leq 2^{n-1} |\tilde{f}(t_k)|$ and hence $|f(t)| \leq (n+1) 2^n \max_j |f(t_j)|$.

For the case of general N, suppose that there were more than n intervals of the form $[t_k, t_{k+1}]$ for which

$$\sup_{t \in [t_j, t_{j+1}]} |f(t)| \geq (n+1) 2^n \max_{j=1,\ldots,N} |f(t_j)|.$$

If $s_1, \ldots, s_{n+1}$ are the leftmost endpoints of such intervals, then there must exist an interval I' on which

$$\sup_{t \in I'} |f(t)| \leq (n+1) 2^n \max_{j=1,\ldots,n+1} |f(s_j)|$$

for any function f whose n-th derivative does not change sign. The leftmost endpoint of this interval coincides with one of the s_j's, hence I' must contain $[t_j, t_{j+1}]$, giving a contradiction. $\qquad \square$

1.2. The proofs of theorems 3 and 4.

Proof of theorem 3. The proof proceeds in two parts; first for any $f \in C^n(I)$,

$$(5) \qquad \int |f(t)| d\mu(t) \geq \frac{(\ell_n(\mu))^n}{(n+1)!} \inf_{t \in I} |f^{(n)}(t)|.$$

Following this the general situation is considered: if the support of μ is contained in some closed set K and $K_\epsilon \supset K$ is as defined in theorem 3 then

$$(6) \qquad \int |f(t)|d\mu(t) \gtrsim (\ell_n(\mu))^n \inf_{t \in K_\epsilon} f^{(n)}(t)$$

provided that $f^{(n)}$ does not change sign on I.

Regarding (5), for any function f one has the trivial inequality

$$(n+1) \int |f(t)|d\mu(t) \int |V_n(t_1, \ldots, t_n)|d\mu(t_1) \cdots d\mu(t_n)$$
$$\geq \int \left| \sum_{i=1}^{n+1} (-1)^{n+1-i} f(t_i) V_n(t_1, \ldots, \hat{t}_i, \ldots, t_{n+1}) \right| d\mu(t_1) \cdots d\mu(t_{n+1}).$$

If one supposes further that $f \in C^{(n)}(I)$, equation (3) from proposition 1 gives that

$$(n+1) \int |f(t)|d\mu(t) \int |V_n(t_1, \ldots, t_n)|d\mu(t_1) \cdots d\mu(t_n)$$
$$\geq \frac{1}{n!} \int \left| \int f^{(n)}(s)\psi_t(s)ds \right| |V_{n+1}(t_1, \ldots, t_{n+1})|d\mu(t_1) \cdots d\mu(t_{n+1}).$$

Since $|\int f^{(n)}(s)\psi_t(s)ds| \geq \inf_{t \in S} |f^{(n)}(t)|$, the inequality (5) must be true.

The proof of (6) and, hence, theorem 3 follows from a closer estimation of the integral $\int f^{(n)}(s)\psi_t(s)ds$. Let $t_1, \ldots, t_{n+1}$ be taken from some closed set $K \subset I$. Fix any $\epsilon > 0$ and let K_ϵ (as in theorem 3) be the closed set of points t for which $t_+ := \inf \{s \in K \mid s \geq t\}$ and $t_- := \sup \{s \in K \mid s \leq t\}$ both exist and satisfy either $|t_+ - t| \leq \epsilon|t_+ - t_-|$ or $|t_- - t| \leq \epsilon|t_+ - t_-|$. It suffices to show that, for any continuous function g on I which does not change sign,

$$\int g(s)\psi_t(s)ds \gtrsim \inf_{s \in K_\epsilon} g(s).$$

To that end, let $p_c(t) := \sum_{i=0}^{n-1} c_i t^i$ where the c_i are real coefficients whose squares sum to 1. The ratio

$$\frac{\int_0^\epsilon |p_c(t)|dt + \int_{1-\epsilon}^\epsilon |p_c(t)|dt}{\int_0^1 |p_c(t)|dt}$$

is never equal to zero for any polynomial p of degree at most $n-1$; therefore compactness of the unit sphere and homogeneity imply that there exists a constant $C_{n,\epsilon}$ such that $\int_0^\epsilon |p(t)|dt + \int_{1-\epsilon}^\epsilon |p(t)|dt \geq C_{n,\epsilon} \int_0^1 |p(t)|dt$ for any polynomial p of degree $n-1$. By a suitable change of variables, one has the integral of $|p|$ over the ends (each of length ϵ times the length of the whole interval) of any interval is bounded below by a constant times the integral over the whole interval.

Consider now the integral of g against ψ_t. Clearly there exist a countable number of open, disjoint intervals I_j in the convex hull of K such that

$$\int g(s)\psi_t(s)ds = \int_K g(s)\psi_t(s)ds + \sum_j \int_{I_j} g(s)\psi_t(s)ds.$$

Since $K \subset K_\epsilon$, $\int_K g(s)\psi_t(s)ds \geq (\inf_{s \in K_\epsilon} g(s)) \int_K \psi_t(s)ds$. As for each I_j, the ends of these intervals are in K_ϵ as well (and g is nonnegative on the interior region of

I_j). Furthermore, ψ_t is a polynomial of degree at most $n-1$ on I_j since this interval contains no points of K (hence none of the t_j's). Thus

$$\int g(s)\psi_t(s)ds \geq \inf_{s \in K_\epsilon} g(s) \left(\int_K \psi_t(s)ds + C_{n,\epsilon} \sum_j \int_{I_j} \psi_t(s)ds \right).$$

Summing these finishes the proof.   □

Proof of theorem 4. Fix some ϵ' and positive integer n. Let $\mathcal{C}^{(1)}$ be the collection of (n, ϵ')-children of μ given by proposition 2. Next let $\mathcal{C}^{(2)}$ be the collection of all $(n-1, \epsilon')$-children of intervals $I \in \mathcal{C}^{(1)}$, where the children of an interval I are understood as the children of the measure $\mu_I := \mu(I)^{-1} \mu|_I$ (note that this will always be well-defined since the μ-measures of children always have a minimal amount of mass as controlled by ϵ'). Continue in this manner until the collection $\mathcal{C}^{(n)}$ (the 1-children of the collection $\mathcal{C}^{(n-1)}$) is obtained.

Suppose that f is of polynomial type n on (K_ϵ, I) with constant C. This implies by (the proof of) theorem 3, that

$$(7) \qquad \int |f(t)|d\mu(t) \gtrsim C^{-1}(\ell_n(\mu))^n \sup_{t \in I} |f^{(n)}(t)|.$$

Now, given some interval $I_0 \in \mathcal{C}^{(n)}$, let I_j be the unique element of $\mathcal{C}^{(n-j)}$ containing I_1. For convenience, let $I_{-1} := I_0$ and $I_n := I$ (the interval on which μ is supported). Fix $c^{-1} = \log \frac{3}{2}$, and choose the j in $0, \ldots, n$ which maximizes

$$(8) \qquad c^j |I_{j-1}|^j \sup_{t \in I_j} |f^{(j)}(t)|$$

(if j is not unique, choose the largest such j). If $j = n$, then inequality (7) has as an immediate consequence that

$$\int |f(t)|d\mu(t) \gtrsim C^{-1}|I_{j-1}|^j \sup_{t \in I_j} |f^{(j)}(t)|$$

for $j = 0, \ldots, n$. Suppose instead that the maximizing index j is not equal to n. In this case, let $s_0 \in I_j$ be the point where the supremum is obtained. It follows that, for any $s \in I_j$,

$$c^j |I_{j-1}|^j |f^{(j)}(s) - f^{(j)}(s_0)| \leq \sum_{k=j+1}^{n} \frac{c^j |I_{j-1}|^j |s - s_0|^{k-j}}{(k-j)!} \sup_{t \in I_k} |f^{(k)}(t)|$$

$$\leq \sum_{k=j+1}^{n} \frac{c^{j-k}}{(k-j)!} c^k |I_{k-1}|^k \sup_{t \in I_k} |f^{(k)}(t)|$$

$$\leq (e^{\frac{1}{c}} - 1)c^j |I_{j-1}|^j \sup_{t \in I_j} |f^{(j)}(t)|.$$

It must therefore be the case that $\sup_{t \in I_j} |f^{(j)}(t)| \leq 2 \inf_{t \in I_j} |f^{(j)}(t)|$, meaning that f is of polynomial type j on (I_j, I_j) with constant 2. Therefore theorem 3 and

proposition 2 guarantee that

$$\int_{I_j} |f(t)|d\mu(t) \gtrsim \mu(I_j)(\ell_j(\mu_{I_j}))^j \sup_{t\in I_j} |f^{(j)}(t)|$$

$$\gtrsim |I_{j-1}|^j \sup_{t\in I_j} |f^{(j)}(t)|$$

$$\gtrsim |I_{k-1}|^k \sup_{t\in I_k} |f^{(k)}(t)| \;\; \forall k = 0, \ldots, n.$$

In particular, there is now a collection of intervals I', namely $\mathcal{C}^{(n)}$ such that the integral $\int |f|d\mu(t) \gtrsim C^{-1} \sup_{t\in I'} |f(t)|$ for any function f which is polynomial type n on (K_ϵ, I) with constant C. By proposition 3, these intervals can be joined together independently of f so that they cover the same set as before but consist of no more than n connected components. In particular, the μ-measure is at least $(1-\epsilon')^n$, which can be made greater than $1 - \epsilon$ for suitably-chosen ϵ'. $\qquad\square$

As for the remaining corollary, choose the interval I' to be a connected component of E as given by theorem 4 which has μ-measure at least $\frac{1-\epsilon}{n}$. In this case, the conclusion of the corollary follows immediately from the combined conclusions of theorems 3 and 4:

Proposition 4. *Suppose that $f \in C^n(I')$. For any $j = 0, \ldots, n$,*

$$\min\{|I'|^j, \ell^j\} \sup_{t\in I'} |f^{(j)}(t)| \lesssim \sup_{t\in I'} |f(t)| + \ell^j \sup_{t\in I'} |f^{(n)}(t)|.$$

Proof. As in the proof of theorem 4, let j be the index out of $0, \ldots, n$ which maximizes $2^j \min\{|I|^j, \ell^j\} \sup_{t\in I} |f^{(j)}(t)|$. If $j = n$, then there is nothing else to prove. Otherwise, for any $t, s \in I$, the mean-value theorem assures that

$$|f^{(j)}(t) - f^{(j)}(s)| \leq |t - s| \sup_{u\in I'} |f^{(j+1)}(u)|.$$

Provided that the inequality $|t-s| \leq \min\{|I|, \ell\}$, the right-hand side will be no greater than $\frac{1}{2} \sup_{u\in I'} |f^{(j)}(u)|$. In particular, if I'' is any interval of length $\min\{|I'|, \ell\}$ containing the point where $f^{(j)}$ achieves its maximum, then the supremum of $f^{(j)}$ on that interval is bounded by twice the infimum. But in this case equation (3) guarantees that $\min\{|I'|^j, \ell^j\} \sup_{t\in I'} |f^{(j)}(t)|$ is bounded below by a constant (depending only on n) times the supremum of f (to apply equation (3), simply choose evenly-spaced points of I''). $\qquad\square$

2. L^p-improving estimates for polynomial curves

This section is devoted to the proof of theorem 2. The proof is itself divided into two parts. The first is the main argument, relying on integral estimates and refinements (and, in particular, relying on theorem 1). With the one-dimensional integral estimates already established, the main portion of the proof of theorem 2 is remarkably short. The second part of the proof deals with counting solutions of the iterated flows of X_1 and X_2. As is customary, this boils down to an application of Bézout's theorem; the difference here is that the vector fields X_1 and X_2 must first be lifted to a nilpotent Lie group (as was done by Christ, Nagel, Stein, and Wainger [11]) to produce a setting in which the flows correspond to polynomial mappings.

2.1. Refinements, re-centering, and integral estimates. The main innovation of the work of Tao and Wright over the original paper of Christ was the observation that, under certain circumstances, the integral of a function over the flow $\exp(tX)(x_0)$ of a vector field X can be estimated from below by the value of that function or its derivatives evaluated at $t = 0$ (the "central" part of central sets of a fixed width). Of course, it is not always possible to make an estimate of this sort (that is, it is easy to construct examples of functions which happen for particular x_0 to be much larger at $t = 0$ than at the other values of t which form the support of the integral). Tao and Wright circumvent this problem by introducing the notion of a set with width w; more recently, Christ [7] avoids this problem by introducing (ϵ, δ)-generic sets. The problem with the construction of Tao and Wright is that, in the process, unavoidable small losses are encountered in various exponents which lead to less-than-sharp restricted weak-type results. One way to avoid this problem, at least in the case of polynomial curves, is to use theorem 1 instead of introducing central sets of a fixed width. The application of theorem 1 comes in the following lemma which describes the set of x_0's for which this re-centering can be accomplished. This new set is called a refinement of the original:

Proposition 5. *Let $U' \subset \mathbb{R}^{d+1}$ be open and $\pi : U \to \mathbb{R}^d$ have surjective differential at every point; let X be a nonvanishing vector field on U' for which $d\pi(X) = 0$. Let $U \subset U'$ be open and bounded and fix a positive integer n. There exists a nonzero constant c depending on n and the bounded subset U such that, for any measurable $\Omega \subset U$, there is a refinement $\Omega' \subset \Omega$ with $|\Omega'| \geq c|\Omega|$ such that, for any $x_0 \in \Omega'$, the integral estimate*

$$\int |f(t, x_0)| \chi_\Omega(\exp(tX)(x_0)) dt \geq$$

$$c \max_{j=0,\ldots,n} \left\{ \left(\frac{|\Omega|}{|\pi(\Omega)|} \right)^{j+1} \left| \frac{\partial^j f}{\partial t^j}(0, x_0) \right| \right\}$$

holds for any function $f(t, x_0)$ which is a polynomial of degree at most n for each fixed x_0.

Proof. It suffices to restrict attention to the portion of Ω which lies on a particular integral curve of X and show that a positive proportion of such points can be taken to lie in Ω'. In this case, one can change variables so that X simply coincides with a coordinate direction and $\exp(tX)$ is simply translation by t in that particular direction. For a first approximation, Ω' is taken to be the set of all points x_0 such that

$$\int \chi_\Omega(\exp(tX)(x_0)) dt \geq c \frac{|\Omega|}{|\pi(\Omega)|}$$

for some small c; Fubini's theorem guarantees that the set $\Omega \setminus \Omega'$ is necessarily only a small fraction of the set Ω. Now the set Ω' as defined is still slightly too big. However, restricting attention to the intersection of a fiber of π with the set Ω', it suffices to prove that, for any set $K \subset \mathbb{R}$, there is a subset $K' \subset K$ with $|K'| \geq c|K|$ for which

$$(9) \qquad \int |f(t, s)| \chi_K(t + s) dt \geq c \max_{j=0,\ldots,n} \left\{ |K|^{j+1} \left| \frac{\partial^j f}{\partial t^j}(0, s) \right| \right\}$$

whenever $s \in K'$. But this inequality follows directly from theorem 1 when, for example, s lies inside the interval I given by that theorem. Thus, if Ω' is further reduced to contain only those points in each fiber which lie in the corresponding interval I given by theorem 1, the proposition follows. $\qquad\qquad\square$

This refinement Ω' is contained in Ω and has $|\Omega'| \geq c|\Omega|$; it follows that

$$\frac{|\Omega'|}{|\tilde{\pi}(\Omega')|} \geq c\frac{|\Omega|}{|\tilde{\pi}(\Omega)|}$$

for any projection $\tilde{\pi}$ (which may or may not be the same as the projection used for refining). Thus when proposition 5 is applied iteratively, it is always possible for the *original* Ω to appear on the right-hand side of (9) at the price of a slightly worse constant (which is not a problem as long as the iterations terminate after a uniformly bounded number of steps).

The proof of theorem 2 now proceeds exactly as in the work of Christ [8] or Tao and Wright [29]. For each x_0, consider the mapping

$$\Phi_{x_0}(t_1, \ldots, t_{d+1}) := \exp(t_1 X_1) \circ \cdots \circ \exp(t_{d+1} X_{d+1})(x_0).$$

In the next section, it will be established that, for fixed x_0, this mapping has finite multiplicity everywhere except for some set of times $(t_1, \ldots, t_{d+1})$ which has $(d+1)$-dimensional Lebesgue measure zero. Thus it follows that, for any (measurable) set $\Omega \subset U$,

$$|\Omega| \geq c \int \chi_\Omega(\Phi_{x_0}(t_1, \ldots, t_{d+1}))|J_{x_0}(t_1, \ldots, t_{d+1})|dt_1 \cdots dt_d$$

where c is the reciprocal of the maximum multiplicity and $J_{x_0}(t)$ is the Jacobian determinant of the mapping $\Phi_{x_0}(t)$. If Ω' is the refinement via the previous proposition with respect to the mapping π_1 and vector field X_1, it follows that

$$\int \chi_\Omega(\Phi_{x_0}(t_1, \ldots, t_{d+1}))|J_{x_0}(t_1, \ldots, t_{d+1})|dt_1 \cdots dt_{d+1}$$
$$\geq c'\left(\frac{|\Omega|}{|\pi_1(\Omega)|}\right)^{j+1} \int \chi_{\Omega'}(\Phi_{x_0}(0, t_2, \ldots, t_d))\left|\frac{\partial^j J_{x_0}}{\partial t_1^j}(0, t_2, \ldots, t_d)\right|dt_2 \cdots dt_{d+1}$$

for any $j = 0, \ldots, n$. But this new integral can, in turn, be estimated in exactly the same way by refining Ω' with respect to the mapping π_2 and the vector field X_2 and so on. The end result is that, for any multiindex α there is a constant c_α such that

$$(10) \qquad\qquad |\Omega| \geq c\prod_{i=1}^{d+1}\left(\frac{|\Omega|}{|\pi_i(\Omega)|}\right)^{\alpha_i+1}\left|\frac{\partial^\alpha J_{x_0}}{\partial t^\alpha}(0)\right|$$

for all x_0 in some iterated refinement of Ω (which, in particular, has nonzero measure). Notice, however, that when Ω is defined by taking $\chi_\Omega(x) := \chi_F(\pi_1(x))\chi_G(\pi_2(x))$, this inequality may be manipulated to give theorem 2.

It is also worth noting that when equation (10) is summed over all multiindices α, one obtains the rather interesting geometric inequality that

$$|\Omega| \geq c\left|B_0\left(x_0, \frac{|\Omega|}{|\pi_1(\Omega)|}, \frac{|\Omega|}{|\pi_2(\Omega)|}\right)\right|$$

where $B_0(x_0, \delta_1, \delta_2)$ is the image of the set $[-\delta_1, \delta_1] \times [-\delta_2, \delta_2] \times \cdots \times [-\delta_{d+1}, \delta_{d+1}]$ (with the usual periodicity convention) under the mapping Φ_{x_0}. The measure of this set is, in turn, comparable to the measure of the two-parameter Carnot-Carathéodory ball $B(x_0; \delta_1, \delta_2)$ of Tao and Wright. Thus equation (10) gives a rather direct proof of the improved version of Tao and Wright's equation (66) mentioned in the second remark at the end of the paper.

2.2. Lifting as related to polynomial curves. In this section, it remains to show that Φ_{x_0} has bounded multiplicity outside some exceptional set and that the Jacobian determinant $J_{x_0}(t)$ is (up to a factor bounded away from 0) a polynomial function of the t parameters. The main idea of the proof of these facts is a lifting argument involving the Baker-Campbell-Hausdorff formula. The reader is referred to the paper of Christ, Nagel, Stein, and Wainger [11] for a thorough treatment of this topic. In the proof at hand, this previous must be improved slightly (to obtain exact formulas rather than asymptotic ones), but there is not any added difficulty; in fact, it will suffice to only reproduce a few very small pieces of this much larger work.

To that end, let $\mathcal{N}$ be the collection of all words w for which X_w (the commutator as defined at the beginning of the paper) does not vanish identically. For any $s \in \mathbb{R}^{\mathcal{N}}$, let

$$s \cdot X := \sum_{w \in \mathcal{N}} s_w X_w.$$

Fix a bounded open set $U \subset \mathbb{R}^{d+1}$ on which X_1 and X_2 are defined and fix $x_0 \in U$. Let $\tilde{U} \subset \mathbb{R}^{\mathcal{N}}$ be the collection of all s for which $\exp(\theta s \cdot X)(x_0) \in U$ for all $\theta \in [0, 1]$ (the inclusion of $\theta < 1$ guarantees that for any s, the associated integral curves used to define $\exp(s \cdot X)(x_0)$ remain in U). Suppose for the moment that it is possible to establish the following two facts:

(1) For any $x \in U$, the fiber $\exp(s \cdot X)(x_0) = x$ of the mapping $\exp(s \cdot X)(x_0)$ (as a function from $\tilde{U}$ to U) can be parametrized by a polynomial function, i.e., there exists a mapping $N_x(u)$ which parametrizes the fiber, has coordinate functions which are polynomials in the u variables, and has surjective differential.

(2) There exists a lifting $\tilde{\Phi}_{x_0}(t)$ of $\Phi_{x_0}(t)$ which is also polynomial, that is, $\exp(\tilde{\Phi}_{x_0}(t) \cdot X)(x_0) = \Phi_{x_0}(t)$ and the coordinate functions of $\tilde{\Phi}_{x_0}$ are polynomial functions of t.

These facts combined allow one to use Bézout's theorem, just as was employed in the original paper of Christ [8]. Specifically, it follows that the equation $\Phi_{x_0}(t) = x$ has a solution for some t only when the equations $\tilde{\Phi}_{x_0}(t) = N_x(u)$ have a solution for the same t and some value of u. In the usual manner, an additional parameter v can be added to these equations in such a way that the resulting system of equations is homogeneous in (t, u, v) and reduces to the original system when $v = 1$. Now Bézout's theorem guarantees that the number of irreducible components (in complex projective space) of the variety determined by these equations is at most the product of the degrees (and, in particular, does not depend on the particular choice of x). See Fulton [13], chapter 8, section 4 (and, in particular, example 8.4.6) for this version of Bézout's theorem.

Now The Jacobian determinant $J_{x_0}(t)$ is nonzero at t_0 only when the graph of $\tilde{\Phi}_{x_0}(t)$ is transverse to the fibers of $\exp(s \cdot X)(x_0)$ at the point $\tilde{\Phi}_{x_0}(t_0)$. Thus solutions to $\Phi_{x_0}(t) = x$ (for real t) at which the Jacobian determinant $J_{x_0}(t)$ is nonvanishing arise only when there is an isolated solution of the system $\tilde{\Phi}_{x_0}(t) = N_x(u)$ in (t, u) space; that is, when the Jacobian determinant with respect to (t, u) of the mapping $\tilde{\Phi}_{x_0}(t) - N_x(u)$ is also nonzero. This, in turn, guarantees that the solution (t, u) remains isolated amongst complex solutions as well. But any isolated solution corresponds to an irreducible component of the zero set of the homogeneous equations in complex projective space; it therefore follows that there is a uniform bound on the number of solutions to $\Phi_{x_0}(t) = x$ which occur where the Jacobian determinant is nonvanishing.

As for any solutions at which the Jacobian determinant may vanish, Sard's lemma guarantees that the set of times t at which the Jacobian determinant does vanish has $(d+1$-dimensional) measure zero. In particular, this also means that the set of points $x \in U$ for which there can exist a solution to $\Phi_{x_0}(t) = x$ with vanishing Jacobian determinant is also a set of measure zero in U. Thus, except for an exceptional set of x's of measure zero, the system of equations $\Phi_{x_0}(t) = x$ has a uniformly bounded number of solutions, and the Jacobian determinant of Φ at each such solution is nonzero. Thus the usual change-of-variables formula gives at once the desired inequality

$$|\Omega| \geq c \int \chi_\Omega(\Phi_{x_0}(t)) |J_{x_0}(t)| dt$$

for any set $\Omega \subset U$.

A bit of notation is in order; given vector fields A and B, the vector denoted by $d \exp(A)(B)|_y$ is meant to be the vector at the point y obtained by transporting B via the exponential mapping $\exp(A)$ (so in particular, it is the vector B at $\exp(-A)(y)$ transported to y).

To establish the necessary properties of the lifting of $\Phi_{x_0}(t)$, two facts from geometry are required. The first is the following: suppose that A is a vector field on U which depends smoothly on some parameter h. For any $x_0 \in U$ and any s sufficiently small, the tangent vector of the curve $\gamma(h) := \exp(sA(h))(x_0)$ (as a function of h for s and x_0 fixed) is given by

$$(11) \qquad \frac{\partial}{\partial h}\gamma(h) = \left(\int_0^1 d\exp((1-\theta)sA(h)) \left(\frac{\partial A}{\partial h} \right) d\theta \right) \Bigg|_{\gamma(h)}.$$

In other words, the tangent vector of γ at $\gamma(h)$ is given by transporting

$$(12) \qquad \left(\int_0^1 d\exp(-\theta s A(h)) \left(\frac{\partial A}{\partial h} \right) d\theta \right) \Bigg|_{x_0}$$

to the point $\gamma(h)$ via the exponential map $\exp(sA(h))$. The second fact needed is that, for any vector fields A and B on U; if s is sufficiently small then

$$(13) \qquad \frac{\partial}{\partial s} \left(d\exp(sA)(B)|_{x_0} \right) = d\exp(sA)([B, A])|_{x_0}$$

Note that equation (13) is nothing more than a computation of the Lie derivative of B with respect to A and can be found in Warner [30], for example. It is only (11) which

requires a bit more explanation. To that end, let $\Gamma(s) := \exp(-sA(h)) \circ \exp(sA(h + \Delta h))(x_0)$.

$$\frac{\partial}{\partial s}g(\Gamma(s)) = -\left.(A(h)g)\right|_{\Gamma(s)} + \left.(d\exp(-sA(h))(A(h+\Delta h))g)\right|_{\Gamma(s)}$$
$$= \left.(d\exp(-sA(h))(A(h+\Delta h) - A(h))g)\right|_{\Gamma(s)}.$$

It therefore follows that

$$\frac{g(\Gamma(s)) - g(\Gamma(0))}{\Delta h} = \int_0^1 \left.\left(d\exp(-s\theta A(h))\left(\frac{A(h+\Delta h) - A(h)}{\Delta h}\right)g\right)\right|_{\Gamma(s\theta)} d\theta.$$

As $\Delta h \to 0$, note that $\Gamma(s) \to x_0$. For fixed s, h, let $g(x) = f(\exp(sA(h))(x))$ and let $\Delta h \to 0$. The result is that

$$\frac{\partial}{\partial h}f(\exp(sA(h))(x_0)) = \left.\left(\int_0^1 d\exp(-\theta sA(h))\left(\frac{\partial A}{\partial h}\right)d\theta\right)\right|_{x_0} f(\exp(sA(h))(x_0))$$

which is precisely what it means for the curve $\gamma(h) = \exp(sA(h))(x_0)$ to have a tangent vector obtained by transporting the vector (12) via the map $\exp(sA(h))$.

Consider the mapping $\varphi : \tilde{U} \to U$ given by $\varphi(s) := \exp(s \cdot X)(x_0)$. Taking a Taylor expansion of the integrand (11) with respect to θ (computing these derivatives via (13)) allows one to easily compute derivatives of φ with respect to the parameters s_w:

$$\frac{\partial}{\partial s_w}\varphi(s) = \left.\left(\int_0^1 d\exp((1-\theta)s \cdot X)(X_w)\,d\theta\right)\right|_{\varphi(s)}$$
$$= \left.\left(\sum_{j=0}^{\infty} \frac{(-1)^j[(s \cdot X)^j X_w]}{(j+1)!}\right)\right|_{\varphi(s)}$$

where $[(s \cdot X)^j X_w]$ is the repeated commutator given by $[(s \cdot X)^0 X_w] := X_w$ and $[(s \cdot X)^{j+1} X_w] = [s \cdot X, [(s \cdot X)^j X_w]]$ for each $j \geq 0$. Note that this sum is, in fact, a finite sum by virtue of the vanishing commutator condition. Also note that the coefficients of the sum are precisely the Taylor coefficients of $\frac{e^{-x}-1}{-x}$.

Now let $\tilde{c}_w(s)$ be the coefficient of X_w when the sum

$$\sum_{j=0}^{\infty} \frac{(-1)^j B_j[(s \cdot X)^j X_1]}{j!}$$

is expanded into a linear combination of words (where the B_j's are the Bernoulli numbers; note that these are chosen so that the coefficients in j are equal to the Taylor coefficients of $\frac{-x}{e^{-x}-1}$). It follows that

$$\left(\sum_{w \in \mathcal{N}} \tilde{c}_w(s)\frac{\partial}{\partial s_w}\right)\varphi(s) = \left.X_1\right|_{\varphi(s)}.$$

The vector field $\sum_{w \in \mathcal{N}} \tilde{c}_w(s)\frac{\partial}{\partial s_w}$ is thus a lifting of X_1; moreover, since the coefficient $\tilde{c}_w(s)$ involves only those parameters $s_{w'}$ for which the length of w' is less than the length of w, it follows that the integral curves of this lifted vector field in $\mathbb{R}^{\mathcal{N}}$ will be given by a polynomial function of the time parameter. Thus composing these flows

(X_2 may be lifted in precisely the same way) establishes the fact that $\Phi_{x_0}(t)$ has a lifting $\tilde{\Phi}_{x_0}(t)$ which is given coordinate-wise by polynomial functions.

To parametrize the fibers of $\varphi(s)$, the alternative formulation

$$\frac{\partial}{\partial s_w} f(\varphi(s)) = \left(\sum_{j=0}^{\infty} \frac{[(s \cdot X)^j X_w]}{(j+1)!} \right)\Bigg|_{x_0} f(\exp(s \cdot X)(x_0))$$

is used (here the Taylor series expansion of the integrand of (12) is taken instead of (11)). Suppose that constants u_w are chosen so that $\sum_{w \in \mathcal{N}} u_w X_w$ equals the zero vector *at the particular point* x_0 (the sum may or may not equal zero elsewhere). If one defines coefficients $\tilde{d}_w(s, u)$ as before by expanding

$$\sum_{w \in \mathcal{N}} \tilde{d}_w(s, u) X_w := \sum_{j=0}^{\infty} \frac{B_j [(s \cdot X)^j (u \cdot X)]}{j!}$$

formally, it follows that the vector field $\sum_{w \in \mathcal{N}} \tilde{d}_w(s, u) \frac{\partial}{\partial s_w}$ will satisfy the property that

$$\left(\sum_{w \in \mathcal{N}} \tilde{d}_w(s, u) \frac{\partial}{\partial s_w} \right) \varphi(s) = 0.$$

In other words, for any appropriately chosen values of u, the corresponding vector field will be tangent to the fibers of $\varphi(s)$. The integral curves of these fibers will be given by polynomial functions of u for the same reason that the coefficients $\tilde{d}_w$ only depend on $s_{w'}$ for w' of shorter length. Now a simple dimension-counting guarantees that the fibers can, in fact, be smoothly parametrized by the flows of these vector fields, that is, the exponential flow of these vector fields based at x_0 has surjective differential with respect to the u variables (one needs only apply the implicit function theorem; note that the curvature condition on the vector fields X_1 and X_2 guarantees that the differential $d\varphi$ is surjective). Thus the earlier counting arguments hold, and in particular, $\Phi_{x_0}(t) = x$ has boundedly many solutions for all x outside a set of measure zero.

To complete the proof of theorem 2, one fact remains to be established: namely, that the Jacobian determinant $J_{x_0}(t)$ is, up to a nonvanishing factor, a polynomial function of the parameters t. Notice that the vector $\frac{\partial}{\partial t_i} \Phi_{x_0}(t)$ is equal to $d \exp(t_1 X_1) \circ \cdots \circ d \exp(t_{i-1} X_{i-1})(X_i)$ evaluated at the point $\Phi_{x_0}(t)$. Up to a bounded, nonvanishing factor, the Jacobian determinant can be evaluated by transporting these vectors to the point x_0 and then computing a determinant. In that case, it follows that $J_{x_0}(t)$ is proportional to

$$\det(d \exp(-t_{d+1} X_{d+1}) \circ \cdots \circ d \exp(-t_2 X_2)(X_1),$$
$$d \exp(-t_{d+1} X_{d+1}) \circ \cdots \circ d \exp(-t_3 X_3)(X_2),$$
$$\ldots, X_{d+1})|_{x_0}.$$

But now (13) guarantees that each vector in this expression is a polynomial function of t, and multilinearity of the determinant establishes the desired property of $J_{x_0}(t)$.

References

[1] G. I. Arhipov, A. A. Karacuba, and V. N. Čubarikov, *Trigonometric integrals*, Izv. Akad. Nauk SSSR Ser. Mat. **43** (1979), no. 5, 971–1003, 1197.

[2] J.-G. Bak, D. M. Oberlin, and A. Seeger, *Restriction of Fourier transforms to curves, II: Some classes with vanishing torsion*, J. Aust. Math. Soc. **85** (2008)

[3] J. Bennett, A. Carbery, M. Christ, and T. Tao, *The Brascamp-Lieb inequalities: finiteness, structure and extremals*, Geom. Funct. Anal. **17** (2008), no. 5, 1343–1415.

[4] J. Bennett, A. Carbery, and J. Wright, *A non-linear generalization of the Loomis-Whitney inequality and applications*, Math. Rest. Lett. **12** (2005) 443–457.

[5] A. Carbery, M. Christ, and J. Wright, *Multidimensional van der Corput and sublevel set estimates*, J. Amer. Math. Soc. **12** (1999), no. 4, 981–1015.

[6] A. Carbery and J. Wright, *What is van der Corput's lemma in higher dimensions?*, in Proceedings of the 6th International Conference on Harmonic Analysis and Partial Differential Equations (El Escorial, 2000), Vol. Extra, 13–26 (2002).

[7] M. Christ, *Lebesgue space bounds for one-dimensional generalized Radon transforms.* (preprint).

[8] ———, *Convolution, curvature, and combinatorics: a case study*, Internat. Math. Res. Notices (1998), no. 19, 1033–1048.

[9] M. Christ and M. B. Erdoğan, *Mixed norm estimates for a restricted X-ray transform*, J. Anal. Math. **87** (2002) 187–198.

[10] M. Christ and M. B. Erdoğan, *Mixed norm estimates for certain generalized Radon transforms*, Trans. Amer. Math. Soc. **360** (2008), no. 10, 5477–5488.

[11] M. Christ, A. Nagel, E. M. Stein, and S. Wainger, *Singular and maximal Radon transforms: analysis and geometry*, Ann. of Math. (2) **150** (1999), no. 2, 489–577.

[12] M. B. Erdoğan and R. Oberlin, *Estimates for the X-ray transform restricted to 2-manifolds.* (preprint).

[13] W. Fulton, Intersection theory, Vol. 2 of *Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]*, Springer-Verlag, Berlin (1984).

[14] A. Greenleaf and A. Seeger, *Fourier integral operators with fold singularities*, J. Reine Angew. Math. **455** (1994) 35–56.

[15] ———, *Fourier integral operators with cusp singularities*, Amer. J. Math. **120** (1998), no. 5, 1077–1119.

[16] W. Littman, $L^p - L^q$-*estimates for singular integral operators arising from hyperbolic equations*, in Partial differential equations (Proc. Sympos. Pure Math., Vol. XXIII, Univ. California, Berkeley, Calif., 1971), 479–481, Amer. Math. Soc., Providence, R.I. (1973).

[17] D. M. Oberlin, *Convolution estimates for some measures on curves*, Proc. Amer. Math. Soc. **99** (1987), no. 1, 56–60.

[18] ———, *A convolution estimate for a measure on a curve in* $\mathbf{R}^4$, Proc. Amer. Math. Soc. **125** (1997), no. 5, 1355–1361.

[19] ———, *A convolution estimate for a measure on a curve in* $\mathbf{R}^4$. *II*, Proc. Amer. Math. Soc. **127** (1999), no. 1, 217–221.

[20] D. H. Phong and E. M. Stein, *Hilbert integrals, singular integrals, and Radon transforms. II*, Invent. Math. **86** (1986), no. 1, 75–113.

[21] ———, *Radon transforms and torsion*, Internat. Math. Res. Notices (1991), no. 4, 49–60.

[22] ———, *Models of degenerate Fourier integral operators and Radon transforms*, Ann. of Math. (2) **140** (1994), no. 3, 703–722.

[23] ———, *The Newton polyhedron and oscillatory integral operators*, Acta Math. **179** (1997), no. 1, 105–152.

[24] D. H. Phong, E. M. Stein, and J. Sturm, *Multilinear level set operators, oscillatory integral operators, and Newton polyhedra*, Math. Ann. **319** (2001), no. 3, 573–596.

[25] D. H. Phong and J. Sturm, *Algebraic estimates, stability of local zeta functions, and uniform estimates for distribution functions*, Ann. of Math. (2) **152** (2000), no. 1, 277–329.

[26] K. M. Rogers, *Sharp van der Corput estimates and minimal divided differences*, Proc. Amer. Math. Soc. **133** (2005), no. 12, 3543–3550 (electronic).

[27] A. Seeger, *Degenerate Fourier integral operators in the plane*, Duke Math. J. **71** (1993), no. 3, 685–745.

[28] ———, *Radon transforms and finite type conditions*, J. Amer. Math. Soc. **11** (1998), no. 4, 869–897.

[29] T. Tao and J. Wright, L^p *improving bounds for averages along curves*, J. Amer. Math. Soc. **16** (2003), no. 3, 605–638.

[30] F. W. Warner, Foundations of differentiable manifolds and Lie groups, Scott, Foresman and Co., Glenview, Ill.-London (1971).

Department of Mathematics, University of Pennsylvania, 209 S 33rd Street, Philadelphia, PA 19104

E-mail address: `gressman@math.upenn.edu`