ANALYSIS & PDE

Volume 7 No. 5 2014

BETSY STOVALL

UNIFORM L^p-IMPROVING FOR WEIGHTED AVERAGES ON CURVES

msp

ANALYSIS AND PDE
Vol. 7, No. 5, 2014
dx.doi.org/10.2140/apde.2014.7.1109

UNIFORM L^p-IMPROVING FOR WEIGHTED AVERAGES ON CURVES

BETSY STOVALL

We define variable parameter analogues of the affine arclength measure on curves and prove near-optimal L^p-improving estimates for associated multilinear generalized Radon transforms. Some of our results are new even in the convolution case.

1. Introduction

We consider weighted versions of multilinear generalized Radon transforms of the form

$$M_0(f_1, \dots, f_k) := \int_{\mathbb{R}^d} \prod_{i=1}^{k} f_i \circ \pi_i(x) a(x)\, dx, \tag{1-1}$$

where a is a continuous cutoff function and the $\pi_i : \mathbb{R}^d \to \mathbb{R}^{d-1}$ are smooth submersions.

In [Tao and Wright 2003; Stovall 2011], near endpoint estimates of the form

$$|M_0(f_1, \dots, f_k)| \le C \prod_{i=1}^{k} \|f_i\|_{L^{p_i}(\mathbb{R}^{d-1})}, \tag{1-2}$$

with $C = C(\pi_1, \dots, \pi_k, p_1, \dots, p_k)$, were established for M_0 under the assumption that the π_i satisfy a certain finite-type condition on the support of a. In particular, it was found that the exponents on the right in (1-2) depend on this type. These results are nearly sharp in the sense that if the type of the π_i degenerates anywhere on the set where $a \neq 0$, then the corresponding near endpoint estimates also fail. It is not, however, known in general what happens when the type degenerates at some point where $a \neq 0$ (for instance, on the boundary of the support) or the rate at which the constants in (1-2) blow up as the type degenerates.

Our goal is to quantify and counteract the failure of (1-2) in such situations by replacing M_0 by an appropriately weighted operator, for which we will establish near-optimal Lebesgue space bounds. The exponents (though not the implicit constants) in these bounds will be independent of the choice of $\pi_1, \dots, \pi_k$ and the cutoff function a. Further, the weights we employ transform naturally under changes of coordinates, so they may reasonably be viewed as generalizations of the affine arclength measure on curves in $\mathbb{R}^d$. A number of recent articles (such as [Bak et al. 2009; Dendrinos et al. 2009; Dendrinos and Müller 2013; Dendrinos and Stovall 2012; Dendrinos and Wright 2010; Drury and Marshall 1987; Oberlin 2002; 2003; 2010; Sjölin 1974; Stovall 2010]) have been devoted to establishing uniform estimates for

MSC2010: 42B99.
Keywords: generalized Radon transform, affine arclength.

operators weighted by affine arclength measure, and these results provide much of the motivation for this article.

A motivating example. Stating the main results of this article, or even the results of [Tao and Wright 2003; Stovall 2011] requires some notation, so we postpone this until the next section. By way of background and motivation, we will spend the remainder of the introduction describing a concrete case about which much is known, and which provides the inspiration for the more general operators considered in this article. Let $\gamma : \mathbb{R} \to \mathbb{R}^d$ be a smooth curve and a a continuous cutoff function. Consider the operator

$$T_0 f(x) := \int_{\mathbb{R}} f(x-\gamma(t))a(t)\,dt, \qquad f \in C_0^0(\mathbb{R}^d).$$

By duality, $T_0 : L^p(\mathbb{R}^d) \to L^q(\mathbb{R}^d)$ if and only if, for all $f \in L^p(\mathbb{R}^d)$ and $g \in L^q(\mathbb{R}^d)$,

$$\left| \int_{\mathbb{R}^d} \int_{\mathbb{R}} f(x-\gamma(t))g(x)a(t)\,dt \right| \le C(\gamma, p, q)\|f\|_{L^p(\mathbb{R}^d)}\|g\|_{L^{q'}(\mathbb{R}^d)};$$

this may be compared with (1-2).

The curve γ is said to be of type (at most) N when $\det(\gamma'(t), \dots, \gamma^{(d)}(t))$ vanishes to order at most N at any point. The results of [Dendrinos and Stovall 2014] imply that if γ is of type N on the support of a, $\|T_0\|_{L^p \to L^q} < \infty$ if (p^{-1}, q^{-1}) lies in the trapezoid with vertices

$$(0,0), \quad (1,1), \quad (p_N^{-1}, q_N^{-1}) := \left(\frac{d}{N+d(d+1)/2}, \frac{d-1}{N+d(d+1)/2} \right), \quad (1-q_N^{-1}, 1-p_N^{-1}). \tag{1-3}$$

(The nonendpoint result was due to Tao and Wright [2003].) Further, if N is the maximal type of T_0 on $\{t : a(t) \neq 0\}$, this is sharp. If γ is not of finite type, T_0 satisfies no $L^p(\mathbb{R}^d) \to L^q(\mathbb{R}^d)$ estimates off the line $\{p = q\}$.

It was first noticed in [Sjölin 1974] and [Drury and Marshall 1985] that affine, as opposed to Euclidean, arclength has a uniformizing effect on the bounds for convolution and Fourier restriction operators associated to possibly degenerate curves. It is now known that, for a polynomial curve γ, the convolution operator with affine arclength measure on γ,

$$Tf(x) := \int_{\mathbb{R}} f(x-\gamma(t))|\det(\gamma'(t), \dots, \gamma^{(d)}(t))|^{2/(d(d+1))}\,dt,$$

maps $L^p(\mathbb{R}^d)$ boundedly into $L^q(\mathbb{R}^d)$ if and only if (p^{-1}, q^{-1}) lies on the line segment joining (p_0^{-1}, q_0^{-1}), $(1-q_0^{-1}, 1-p_0^{-1})$, with p_0, q_0 defined as above (provided $T \not\equiv 0$) [Oberlin 2002; Dendrinos et al. 2009; Stovall 2010]. Further, the operator norms these papers established depend only on the degree of the polynomial; for this, it is crucial that the affine arclength transforms nicely under reparametrizations and affine transformations. Further investigations have been carried out in [Oberlin 2010; Dendrinos and Stovall 2014] in the nonpolynomial case. The above mentioned results are essentially optimal, both in terms of the exponents involved and in terms of pointwise estimates on the weight [Oberlin 2003] (see Proposition 2.2). Analogous results are also known for the restricted X-ray transform [Dendrinos and Stovall 2012; 2014]. There have also been a number of recent articles aimed at establishing uniform estimates for Fourier restriction to curves with affine arclength measure, for instance [Bak et al. 2009; Dendrinos and Müller 2013; Dendrinos and Wright 2010; Stovall 2014].

Our goal in this article is to address the gap between the general results of [Tao and Wright 2003; Stovall 2011] and the type-independent results of [Dendrinos et al. 2009; Dendrinos and Stovall 2012; Oberlin 2002; Stovall 2010] by introducing a generalization of the affine arclength measure, well-suited to (1-1). We will also prove near endpoint bounds for the weighted operator and, in particular, will generalize the results of [Tao and Wright 2003; Stovall 2011] to the case when the π_i completely fail to be of finite type on the support of a. Some of our results are new even in the translation-invariant case.

2. Basic notions and statements of the main results

Notation. Throughout the article, we will use the now-standard notation $A \lesssim B$ to mean that $A \leq CB$ for some innocuous implicit constant C. The value of this constant will be allowed to change from line to line. The meaning of "innocuous" will be specified at the beginning of most sections, though in this section it will be specified *in situ*, and in the next it does not arise. Additionally, $A \gtrsim B$ if $B \lesssim A$, and $A \sim B$ if $A \lesssim B$ and $B \lesssim A$. We denote the nonnegative integers by $\mathbb{Z}_0$. If ℓ is any integer, δ is an ℓ-tuple of real numbers and $\beta \in \mathbb{Z}_0^\ell$ is a multiindex, we denote by δ^β the quantity $\delta_1^{\beta_1} \cdots \delta_\ell^{\beta_\ell}$.

We will also use some less standard notation. We consider the partial order $\preceq$ on $\mathbb{Z}_0^k$ defined by $b_1 \preceq b_2$ if $b_1^i \leq b_2^i$ for $1 \leq i \leq k$. We say $b_1 \prec b_2$ if at least one of these inequalities is strict. If $\mathcal{B} \subseteq \mathbb{Z}_0^k$ is any set, we define a polytope

$$\mathcal{P}(\mathcal{B}) := \operatorname{ch} \bigcup_{b \in \mathcal{B}} ([0,\infty)^k + \{b\}),$$

where "ch" denotes the convex hull.

Fix a dimension d and an integer $k \geq 2$; k may exceed d. We will consider vector fields $X_1, \dots, X_k$, defined and smooth on the closure of an open set U. A word w is an element of $\mathcal{W} := \bigcup_{n=1}^\infty \{1, \dots, k\}^n$. To each word is associated a vector field X_w, defined recursively by $X_{(i)} := X_i$ for $1 \leq i \leq k$ and $X_{(w,i)} := [X_w, X_i]$ for $w \in \mathcal{W}$ and $1 \leq i \leq k$. The degree of $w \in \mathcal{W}$ is the k-tuple, $\deg w$, whose i-th entry is the number of occurrences of i in w.

All brackets of such vector fields lie in the span of the X_w: if $w, w' \in \mathcal{W}$,

$$[X_w, X_{w'}] = \sum_{\deg \widetilde{w} = \deg w + \deg w'} C_{w,w'}^{\widetilde{w}} X_{\widetilde{w}}, \tag{2-1}$$

where $C_{w,w'}^{\widetilde{w}}$ is an integer. Indeed, by the Jacobi identity,

$$[X_w, [X_{w'}, X_i]] = [[X_w, X_{w'}], X_i] - [X_{(w,i)}, X_{w'}],$$

so (2-1) is easily obtained by inducting on $\|\deg w'\|_{\ell^1}$ [Hörmander 1967]. We note that for each $b \in \mathbb{N}^k$ there are only finitely many words w with $\deg w = b$, so the sum in (2-1) is finite.

If $I = (w_1, \dots, w_d)$ is a d-tuple of words, we define $\deg I := \sum_{i=1}^d \deg w_i$ and

$$\lambda_I := \det(X_{w_1}, \dots, X_{w_d}).$$

The Newton polytope of the vector fields $X_1, \dots, X_k$ at the point $x_0 \in U$ is defined to be

$$\mathcal{P}_{x_0} := \mathcal{P}(\{\deg I : I \text{ is a } d\text{-tuple of words satisfying } \lambda_I(x_0) \neq 0\}),$$

and we define the Newton polytope of a set $A \subseteq U$ to be

$$\mathcal{P}_A := \operatorname{ch}\Bigl(\bigcup_{x \in A} \mathcal{P}_x\Bigr).$$

The Hörmander condition is the statement that $\mathcal{P}_{x_0} \neq \varnothing$ for each $x_0 \in U$. When the X_i are nonvanishing vector fields tangent to the fibers of the π_i, this is the finite-type hypothesis in [Tao and Wright 2003; Stovall 2011].

Results. Let $U \subseteq \mathbb{R}^d$ be an open set and let $\pi_1, \dots, \pi_k : \overline{U} \to \mathbb{R}^{d-1}$ be smooth submersions (i.e., they have surjective differentials). Letting $\star$ denote the composition of the Hodge-star operator, which maps $(d-1)$-forms to 1-forms, with the natural identification of 1-forms with vectors via the Euclidean metric, we define vector fields

$$X_j := \star(d\pi_j^1 \wedge \cdots \wedge d\pi_j^{d-1}), \quad 1 \leq j \leq k. \tag{2-2}$$

Let a be a continuous function with compact support contained in U.

Fix a d-tuple of words $I_0 = (w_1, \dots, w_d)$ and define the generalized affine arclength

$$\rho = \rho_{I_0} := |\det(X_{w_1}, \dots, X_{w_d})|^{1/(|\deg I_0|_1 - 1)}, \tag{2-3}$$

where $|b|_1$ denotes the ℓ_1-norm. Define a k-linear form $M : [C^0(\mathbb{R}^d)]^k \to \mathbb{C}$ by

$$M(f_1, \dots, f_k) := \int_{\mathbb{R}^d} \prod_{j=1}^k f_j \circ \pi_j(x) \rho(x) a(x)\, dx. \tag{2-4}$$

For $b \in \mathbb{R}^k$ with $|b|_1 > 1$, define

$$\boldsymbol{q}(b) := \frac{b}{|b|_1 - 1}. \tag{2-5}$$

It is easy to check that $\boldsymbol{q}$ equals its own inverse. The following is our main theorem.

Theorem 2.1. *Assume that* $\deg I_0$ *is an extreme point of* $\mathcal{P}_{\operatorname{supp} a}$. *Then, for all* $\boldsymbol{p} \in [1, \infty]^k$ *satisfying* $(p_1^{-1}, \dots, p_k^{-1}) \preceq \boldsymbol{q}(b)$ *and* $p_j^{-1} < q_j(b)$ *when* $(\deg I_0)_j \neq 0$, *we have the estimate*

$$|M(f_1, \dots, f_k)| \lesssim \prod_{j=1}^k \|f_j\|_{L^{p_j}(\mathbb{R}^{d-1})}, \tag{2-6}$$

for all continuous $f_1, \dots, f_k$. *The implicit constant depends on the* π_j, a, $\boldsymbol{p}$ *and* b_0, *but not on the* f_j. *Thus* M *extends to a bounded* k*-linear form on* $\prod_{j=1}^k L^{p_j}(\mathbb{R}^{d-1})$.

The extremality hypothesis seems natural by analogy with the translation-invariant case; it also leads to certain invariants of the weight, as we will discuss below. However, we ultimately prove a more general result, Theorem 6.1, which does not require extremality. (We postpone stating the latter because it requires more notation.)

With the given weight, the above theorem is nearly sharp. Indeed, under the hypotheses and notation above, we have the following.

Proposition 2.2. *Let μ be a nonnegative Borel measure whose support is contained in U, and assume that the bound*

$$M_\mu(\chi_{E_1}, \dots, \chi_{E_k}) := \int_{\mathbb{R}^d} \prod_{j=1}^{k} \chi_{E_j} \circ \pi_j \, d\mu \le A(\mu) \prod_{j=1}^{k} |E_j|^{1/p_j} \tag{2-7}$$

holds for all Borel sets $E_1, \dots, E_k \subseteq \mathbb{R}^{d-1}$ and some constant $A(\mu) < \infty$. If $\mu \not\equiv 0$, $(p_1, \dots, p_k) \in [1, \infty]^k$. If $\sum_j p_j^{-1} > 1$, let $b_p := \boldsymbol{q}(p_1^{-1}, \dots, p_k^{-1})$. Then $\mu(\{x : b_p \notin \mathcal{P}_x\}) = 0$. If in addition b_p is an extreme point of $\mathcal{P}_{\operatorname{supp}\mu}$, then μ is absolutely continuous with respect to Lebesgue measure and its Radon–Nikodym derivative satisfies

$$\frac{d\mu}{dx} \lesssim A(\mu) \sum_{\deg I = b_p} |\lambda_I|^{1/(|b_p|_1 - 1)}. \tag{2-8}$$

The implicit constant in (2-8) may be chosen to depend only on d and p; $A(\mu)$ has the same value in (2-7) and (2-8).

In the translation-invariant case, a similar result is due to Oberlin [2003] (see [Dendrinos and Stovall 2012] for the restricted X-ray transform). The final statement in the proposition only applies in the endpoint case, which is not otherwise addressed in this article. The endpoint version of Theorem 2.1 is known to fail without further assumptions on the X_i than those made here, as can be seen by considering the example of convolution with affine arclength on $\gamma(t) = (t, e^{-1/t} \sin(1/t^k))$, $t > 0$, for k sufficiently large [Sjölin 1974].

The proofs of Theorem 2.1 and Proposition 2.2 will rely on a more general result about smooth vector fields $X_1, \dots, X_k$ on $\mathbb{R}^d$. To state this result, we need some additional terminology.

Let $J \in \{1, \dots, k\}^d$. We define $\deg J$ to be the k-tuple whose i-th entry is the number of occurrences of i in J. If $\alpha \in \mathbb{Z}_0^d$ is a multiindex, we define $\deg_J \alpha$ to be the k-tuple whose i-th entry is $\sum_{\ell : J_\ell = i} \alpha_\ell$. We define

$$\Psi_{x_0}^J(t_1, \dots, t_d) := \exp(t_d X_{J_d}) \circ \cdots \circ \exp(t_1 X_{J_1})(x_0). \tag{2-9}$$

We define another polytope,

$$\widetilde{\mathcal{P}}_{x_0} := \mathcal{P}\big(\{\deg J + \deg_J \alpha : J \in \{1, \dots, k\}^d \text{ and } \alpha \in (\mathbb{Z}_0)^d \text{ satisfy } \partial_t^\alpha \det D\Psi_{x_0}^J(0) \ne 0\}\big).$$

Proposition 2.3. *For each $x_0 \in U$, $\widetilde{\mathcal{P}}_{x_0} = \mathcal{P}_{x_0}$. Furthermore, for each extreme point b_0 of $\mathcal{P}_{x_0}$,*

$$\sum_{\deg I = b_0} |\lambda_I(x_0)| \sim \sum_{J \in \{1,\dots,k\}^d} \sum_{\substack{\alpha \in (\mathbb{Z}_0)^d : \\ \deg J + \deg_J \alpha = b_0}} |\partial_t^\alpha \det D\Psi_{x_0}^J(0)|. \tag{2-10}$$

The implicit constants may be taken to depend only on d and b_0, and in particular may be chosen to be independent of the X_i.

Examples. We take a moment to discuss a few concrete cases where these results apply.

The translation-invariant case. Let $\gamma : \mathbb{R} \to \mathbb{R}^d$ be a smooth map and for $(t, x) \in \mathbb{R}^{1+d}$ define $\pi_1(t, x) = x$, $\pi_2(t, x) = x - \gamma(t)$. Thus the unweighted operator M_0 in (1-1) is essentially convolution with Euclidean arclength measure on γ, paired with a test function.

Using the definition above, $X_1 = \partial_t$, $X_2 = \partial_t + \gamma' \cdot \nabla_x$. If w is any word of length $n \geq 2$ and if the first two letters of w are 1 and 2, $X_w(t, x) = \gamma^{(n)}(t)$. If $d \geq 2$, the Hörmander condition is equivalent to the statement that the torsion of γ does not vanish to infinite order at any point. We note in particular that

$$|\det(X_1, X_2, X_{(1,2)}, \dots, X_{(1,\dots,1,2)})| = |\det(X_1, X_2, X_{(2,1)}, \dots, X_{(2,\dots,2,1)})| = |\det(\gamma', \dots, \gamma^{(d)})|$$

and, if U is any open set, the only extreme points of $\mathcal{P}_U$ (unless $\mathcal{P}_U$ is empty) are

$$\big(\tfrac{1}{2}d(d-1)+1, d\big), \quad \big(d, \tfrac{1}{2}d(d-1)+1\big).$$

Thus the affine arclength in this case is defined in the usual way:

$$\rho(t, x) = |\det(\gamma'(t), \dots, \gamma^{(d)}(t))|^{2/(d(d+1))}.$$

By Theorem 2.1, for any smooth $\gamma : \mathbb{R} \to \mathbb{R}^d$ and any continuous cutoff function a, the convolution operator

$$Tf(x) = \int f(x - \gamma(t)) |\det(\gamma'(t), \dots, \gamma^{(d)}(t))|^{2/(d(d+1))} a(t)\, dt$$

maps $L^p(\mathbb{R}^d)$ into $L^q(\mathbb{R}^d)$ whenever (p^{-1}, q^{-1}) lies in the interior of the trapezoid with vertices as in (1-3) in the case $N = 0$. For general smooth curves this result is new but, as mentioned in the introduction, even stronger results are known in some special cases.

Restricted X-ray transforms. Let $\gamma : \mathbb{R} \to \mathbb{R}^{d-1}$ be a smooth map and, for $(s, t, x) \in \mathbb{R}^{1+1+d-1}$, define $\pi_1(s, t, x) := (t, x)$, $\pi_2(s, t, x) := (s, x - s\gamma(t))$. Then the operator M_0 in (1-1) is the restricted X-ray transform

$$Xf(t, x) = \int_{\mathbb{R}} f(s, x - s\gamma(t)) a(s, t)\, ds,$$

paired with a test function. Using the above definition,

$$X_1 = \partial_s, \quad X_2 = \partial_t + s\gamma'(t) \cdot \nabla_x.$$

If $d \geq 3$, the only $(d{+}1)$-tuples of words $(w_1, \dots, w_{d+1})$ with $\det(X_{w_1}, \dots, X_{w_{d+1}}) \not\equiv 0$ are, after reordering, those satisfying

$$w_1 = 1, \qquad w_2 = 2, \qquad w_i = (1, 2, \dots, 2), \quad 3 \leq i \leq d+1.$$

Thus the only extreme point of the Newton polytope is $(d, 1 + \frac{1}{2}d(d-1))$, and

$$\rho(s, t, x) = |\det(\gamma'(t), \dots, \gamma^{(d-1)}(t))|^{2/(d(d+1))},$$

which is a power of the usual affine arclength. Theorem 2.1 thus gives a partial generalization of the results of [Dendrinos and Stovall 2012], wherein a sharp strong-type bound for the X-ray transform restricted to polynomial curves with affine arclength was established.

Generalized Loomis–Whitney. Let $\pi_1, \dots, \pi_d : \mathbb{R}^d \to \mathbb{R}^{d-1}$ be smooth submersions. The point $(1, \dots, 1)$ is always extreme or in the exterior of the Newton polytope, so for $\varepsilon > 0$

$$\left| \int_{\mathbb{R}^d} \prod_{i=1}^{d} f_i \circ \pi_i(x) |\det(X_1, \dots, X_d)(x)|^{1/(d-1)} a(x)\, dx \right| \lesssim \prod_{i=1}^{d} \| f_i \|_{L^{d-1+\varepsilon}(\mathbb{R}^{d-1})},$$

with the implicit constant depending on the π_i and ε. In the case when the X_i do span at every point of the support of a, the endpoint estimate was proved in [Bennett et al. 2005]. (The classical Loomis–Whitney inequality is the endpoint estimate when the π_i are linear and $a \equiv 1$.)

Outline. In Section 3, we show that the weights we employ satisfy certain natural invariants; this makes them reasonable generalizations of the usual affine arclength measure. In Section 4, we prove Proposition 2.3 by employing the results of [Street 2011] and a compactness argument; we also use a combinatorial lemma, whose proof is postponed to the Appendix. In Section 5, we prove the optimality result, Proposition 2.2. Finally, in Section 6, we prove a more general result, Theorem 6.1, which implies Theorem 2.1. Our techniques for the proof of the main theorem are essentially those of [Christ 2008; Tao and Wright 2003; Stovall 2011], with some modifications to handle the potential failure of the Hörmander condition.

3. Invariants of the affine arclengths

Let $U, \pi_1, \dots, \pi_k$, and $X_1, \dots, X_k$ be as defined above. For $1 \le j \le k$, let $V_j := \pi_j(U)$. Fix a d-tuple of words I_0, and assume that $b_0 := \deg I_0$ is minimal in the sense that if $\deg I' \prec \deg I_0$, then $\lambda_I \equiv 0$. (This minimality is essential.) Define ρ as in (2-3).

Proposition 3.1. *Let $F : U \to \mathbb{R}^d$ and $G_j : V_j \to \mathbb{R}^{d-1}$, $1 \le j \le k$, be smooth maps. Define $\widetilde{\pi}_j := G_j \circ \pi_j \circ F$ for $1 \le j \le k$, and let $\widetilde{X}_j$, $\tilde{\rho}$ be defined as in (2-2), (2-3), with tildes inserted. Then*

$$\tilde{\rho} = \left(\prod_{j=1}^{k} |(\det DG_j) \circ \pi_j|^{q_j(b_0)} \right) |\det DF| \rho \circ F, \tag{3-1}$$

where $\boldsymbol{q}$ is defined as in (2-5).

In the notation above, let a be a continuous, compactly supported function with $\operatorname{supp} a \subseteq U$, and define

$$\widetilde{M}(f_1, \dots, f_k) := \int_U \prod_{j=1}^{k} f_j \circ \widetilde{\pi}_j(x) \tilde{\rho}(x) a \circ F(x)\, dx.$$

Proposition 3.1 implies that if each G_j is equal to the identity and F is one-to-one, then

$$\widetilde{M}(f_1, \dots, f_k) = M(f_1, \dots, f_k).$$

If we simply assume that F and all of the G_j are one-to-one, the proposition implies that

$$\sup_{f_1,\dots,f_k\neq 0}\frac{\widetilde{M}(f_1,\dots,f_k)}{\prod_{j=1}^k \|f_j\|_{L^{p_j}(\mathbb{R}^{d-1})}} = \sup_{f_1,\dots,f_k\neq 0}\frac{M(f_1,\dots,f_k)}{\prod_{j=1}^k \|f_j\|_{L^{p_j}(\mathbb{R}^{d-1})}} \quad \text{for } (p_1^{-1},\dots,p_k^{-1}) := \boldsymbol{q}(b_0).$$

We stress, however, that our theorem covers only the nonendpoint cases satisfying $(p_1^{-1},\dots,p_k^{-1}) \neq \boldsymbol{q}(b_0)$ and b_0 extreme, so it is not known that either side is finite except in certain cases; see [Bennett et al. 2005; Dendrinos et al. 2009; Dendrinos and Stovall 2012; Oberlin 2002; Stovall 2010].

If we fix j, we may consider the family of curves $\gamma_j^{\underline{x}}(t) := \pi_j(\underline{x}, t)$. For any smooth one-to-one function $\phi : \mathbb{R} \to \mathbb{R}$, $(\underline{x}, t) \mapsto (\underline{x}, \phi(t))$ is also smooth and one-to-one and has Jacobian determinant $\phi'(t)$. Thus we obtain:

Corollary 3.2. *The generalized affine arclength defines a parametrization-invariant measure on each of the curves* $\gamma_j^{\underline{x}} = \pi_j(\underline{x}, t)$.

Proof of Proposition 3.1. We will prove the proposition first when the G_j are equal to the identity and then when F is. The general case follows by taking compositions.

In the first case, it suffices by simple approximation arguments to prove the identity when $\det DF \neq 0$. In this case, careful computations reveal that

$$\widetilde{X}_j = (\det DF) F^* X_j,$$

where F^* is the pullback by F, given by

$$F^* X := (DF)^{-1} X \circ F. \tag{3-2}$$

For $1 \le i \le k$, let $Y_i = F^* X_i$. Then, by naturality of the Lie bracket, $Y_w = F^* X_w$ for $w \in \mathcal{W}$. By induction (with base case $w = (j)$), the coordinate expression for the Lie bracket $[X, X'] = X(X') - X'(X)$, and the product rule, for each $w \in \mathcal{W}$,

$$\widetilde{X}_w = (\det DF)^{|\deg w|_1} Y_w + \sum_{\deg w' \prec \deg w} f_{w,w'} Y_{w'}, \tag{3-3}$$

where the $f_{w,w'}$ are smooth functions.

By (3-3), (3-2) and our minimality assumption,

$$\begin{aligned}\det(\widetilde{X}_{w_1},\dots,\widetilde{X}_{w_d}) &= (\det DF)^{|b_0|_1}\det(Y_{w_1},\dots,Y_{w_d}) + \sum_{b' \prec b_0}\sum_{\deg I' = b'} f_{I,I'} \det(Y_{w_1'},\dots,Y_{w_d'}) \\ &= (\det DF)^{|b_0|_1 - 1}\det(X_{w_1},\dots,X_{w_d}) \circ F + 0.\end{aligned}$$

This completes the proof in the first case.

In the second case, when F is the identity, it is easy to compute $\widetilde{X}_j = [(\det DG_j) \circ \pi_j] X_j$, and it can be shown using the product rule and minimality of b_0 (as above) that

$$\det(\widetilde{X}_{w_1},\dots,\widetilde{X}_{w_d}) = \prod_{j=1}^k [(\det DG_j) \circ \pi_j]^{b_0^j} \det(X_{w_1},\dots,X_{w_d}),$$

which implies (3-1). □

4. Equivalence of the two polytopes: the proof of Proposition 2.3

Fix a point $b_0 \in [0, \infty)^k$. We say that an object (such as a constant, vector, or set) is admissible if it may be chosen from a finite collection, depending only on b_0 and d, of such objects. In particular, all implicit constants in this section will be admissible.

The proof of Proposition 2.3 will rely on a compactness result about polytopes with vertices in $\mathbb{Z}_0^k$:

Proposition 4.1. *Let $\mathcal{B} \subseteq \mathbb{Z}_0^k$ and assume that $b_0 \notin \mathcal{P}(\mathcal{B})$. There exist*

(i) *$\varepsilon > 0$ and $v_0 \in (\varepsilon, 1]^k$ such that $v_0 \cdot b_0 + \varepsilon < v_0 \cdot p$ for every $p \in \mathcal{P}(\mathcal{B})$, and*

(ii) *a finite set $\mathcal{A} \subseteq \mathbb{Z}_0^k$ such that $b_0 \notin \mathcal{P}(\mathcal{A})$ and $\mathcal{P}(\mathcal{B}) \subseteq \mathcal{P}(\mathcal{A})$.*

Moreover, ε, v_0, $\mathcal{A}$ are admissible.

Note that this also applies when b_0 is an extreme point of $\mathcal{P}(\mathcal{B})$, since in this case $b_0 \notin \mathcal{P}(\mathcal{B} \setminus \{b_0\})$.

Assuming the validity of Proposition 4.1 for now (it will be proved in the Appendix), we devote the remainder of the section to the proof of Proposition 2.3.

We may of course assume that $x_0 = 0$ and that U is a bounded neighborhood of 0. Furthermore, we may assume that $k > d$ and $X_i = \partial_i$, $1 \le i \le d$. Indeed, if the proposition holds under this assumption, it holds for $\partial_1, \dots, \partial_d, X_1, \dots, X_k$, with $k + d$ replacing k. We may then transfer the result back to $X_1, \dots, X_k$ by restricting to those $b \in [0, \infty)^{k+d}$ with $b^1 = \cdots = b^d = 0$. By this assumption, $\mathcal{P}_0 \neq \varnothing$, and it suffices to prove that if b_0 is an extreme point of $\mathcal{P}_{x_0}$ then (2-10) holds, and if $b_0 \notin \mathcal{P}_{x_0}$ then $b_0 \notin \widetilde{\mathcal{P}}_{x_0}$.

We begin with the case when b_0 is an extreme point of $\mathcal{P}_0$. Fix a neighborhood V of 0, sufficiently small for later purposes, with $\overline{V} \subseteq U$. Choose a d-tuple $I_0 = (w_1, \dots, w_d) \in \mathcal{W}^d$ with $\deg I_0 = b_0$ and

$$|\lambda_{I_0}(0)| = \max_{\deg I = b_0} |\lambda_I(0)|. \tag{4-1}$$

(Note that I_0 is admissible, since only finitely many d-tuples of words give rise to this degree.) By smoothness of the X_j, we may assume that V is so small that

$$\tfrac{1}{4}|\lambda_{I_0}(0)| \le \tfrac{1}{2} \max_{\deg I = b_0} |\lambda_I(x)| \le |\lambda_{I_0}(x)| \le 2|\lambda_{I_0}(0)|, \quad \text{for all } x \in V.$$

By Proposition 4.1, we may choose admissible $v_0 = (v_0^1, \dots, v_0^k) \in (0, 1]^k$ and $\varepsilon > 0$ such that $v_0 \cdot b_0 + \varepsilon < v_0 \cdot p$ for every $p \in \mathcal{P}_0 \cap \mathbb{Z}_0^k \setminus \{b_0\}$.

Lemma 4.2. *For each $m \ge 1$, there exists $\delta(m) > 0$, depending on m, b_0, $X_1, \dots, X_k$ such that, for all $0 < \delta < \delta(m)$, the map*

$$\Phi^\delta(y_1, \dots, y_d) := \exp\bigl(y_1 \delta^{v_0 \cdot \deg w_1} X_{w_1} + \cdots + y_d \delta^{v_0 \cdot w_d} X_{w_d}\bigr)(0) \tag{4-2}$$

and the pullbacks

$$Y_j^\delta := (\Phi^\delta)^* \delta^{v_0^j} X_j = (D\Phi^\delta)^{-1} \delta^{v_0^j} X_j \circ \Phi^\delta \tag{4-3}$$

satisfy these properties: Φ^δ is a diffeomorphism of the unit ball $B(1)$ onto a neighborhood of 0 in V,

$$|\det D\Phi^\delta(y)| \sim \delta^{v_0 \cdot b_0} |\lambda_{I_0}(0)|, \quad y \in B(1), \tag{4-4}$$

$$\|Y_j^\delta\|_{C^m(B(1))} \lesssim 1, \quad 1 \le j \le k, \tag{4-5}$$

$$|\det(Y_{w_1}^\delta(y), \dots, Y_{w_d}^\delta(y))| \sim 1, \quad y \in B(1). \tag{4-6}$$

Proof. Recall that $\mathscr{W}$ is the set of all words. Let

$$\mathscr{W}_0 := \{w \in \mathscr{W} : \deg w \cdot v_0 \le d\} \quad \text{and} \quad \mathscr{W}_1 := \{w \in \mathscr{W} : d < \deg w \cdot v_0 \le 2d\}. \tag{4-7}$$

Since v_0 is an admissible element of $(0, 1]^k$, these are admissible, finite sets, and $\mathscr{W}_0$ contains the one-letter words $(1), (2), \dots, (k)$. Furthermore, $\mathscr{W}_0$ contains b_0 since our choice of v_0 and assumption that $X_j = \partial_j$ for $1 \le j \le d$ imply that

$$v_0 \cdot b_0 \le v_0 \cdot (1, \dots, 1, 0, \dots, 0) = (v_0)_1 + \dots + (v_0)_d \le d.$$

The vector fields X_w are all smooth, $\mathscr{W}_0 \cup \mathscr{W}_1$ is a finite set, and each coefficient of v_0 is positive. Thus for each $M \ge 0$, for all sufficiently small $\delta > 0$ and all $w \in \mathscr{W}_0 \cup \mathscr{W}_1$,

$$\|\delta^{v_0 \cdot \deg w} X_w\|_{C^0(V)} \le \frac{1}{d} \operatorname{dist}(0, \partial V), \quad \|\delta^{v_0 \cdot \deg w} X_w\|_{C^M(V)} \le 1. \tag{4-8}$$

Additionally, by our choice of v_0 and ε,

$$|\delta^{v_0 \cdot \deg I} \lambda_I(0)| < \delta^\varepsilon |\delta^{v_0 \cdot b_0} \lambda_{I_0}(0)|, \quad I \in (\mathscr{W}_0 \cup \mathscr{W}_1)^d, \quad \deg I \ne b_0. \tag{4-9}$$

By the Jacobi identity, if $w, w' \in \mathscr{W}_0$,

$$[\delta^{v_0 \cdot \deg w} X_w, \delta^{v_0 \cdot \deg w'} X_{w'}] = \sum_{\deg \widetilde{w} = \deg w + \deg w'} C_{w,w'}^{\widetilde{w}} (\delta^{v_0 \cdot \deg \widetilde{w}} X_{\widetilde{w}}), \tag{4-10}$$

for constants $C_{w,w'}^{\widetilde{w}}$ that are admissible because $\mathscr{W}_0$ is. If $v_0 \cdot (\deg w + \deg w') \le d$, each $\widetilde{w}$ in the sum is an element of $\mathscr{W}_0$. If not, each $\widetilde{w}$ is in $\mathscr{W}_1$, and we can expand

$$\delta^{v_0 \cdot \deg \widetilde{w}} X_{\widetilde{w}} = \sum_{j=1}^d \delta^{v_0 \cdot \deg \widetilde{w}} X_{\widetilde{w}}^j \partial_j = \sum_{j=1}^d (\delta^{v_0 \cdot \deg \widetilde{w} - v_0^j} X_{\widetilde{w}}^j)(\delta^{v_0^j} X_j).$$

Note that $v_0 \cdot \deg \widetilde{w} - v_0^j > 0$ for $\widetilde{w} \in \mathscr{W}_1$. Using (4-10) to put the pieces back together, for sufficiently small $\delta > 0$ and any $w, w' \in \mathscr{W}_0$,

$$[\delta^{v_0 \cdot \deg w} X_w, \delta^{v_0 \cdot \deg w'} X_{w'}] = \sum_{\widetilde{w} \in \mathscr{W}_0} c_{w,w'}^{\widetilde{w},\delta} \delta^{v_0 \cdot \deg \widetilde{w}} X_{\widetilde{w}},$$

with

$$\left\|c_{w,w'}^{\widetilde{w},\delta}\right\|_{C^M(V)} \lesssim 1. \tag{4-11}$$

The conclusion of the lemma is now a direct application of [Street 2011, Theorem 5.3], whose (lengthy) proof uses compactness arguments and Gronwall's inequality, among other tools. For the convenience of the reader wishing to verify this, we provide a short dictionary to translate the notation. Let M be sufficiently large (depending on m, d, I_0) and choose $\delta(m) > 0$ sufficiently small that (4-8), (4-9) and (4-11) all hold. Then the terms

$$\{X_1, \dots, X_q\},\ \{d_1, \dots, d_q\},\ \mathscr{A},\ (\delta^d X),\ n_0(x, \delta)$$

from [Street 2011] are, in our notation,

$$\{X_w\}_{w \in \mathscr{W}_0},\ \{\deg w\}_{w \in \mathscr{W}_0},\ \{(\delta^{v_0^1}, \dots, \delta^{v_0^k}) : 0 < \delta \le \delta(m)\},\ (\delta^{v_0 \cdot \deg w} X_w)_{w \in \mathscr{W}_0},\ d.$$

A priori, the results of [Street 2011] only guarantee that for each $m \ge 0$ there exists an admissible constant $\eta > 0$ such that the conclusions hold on $B(\eta)$. We want $\eta = 1$, but this is just a matter of rescaling. Define

$$D^\eta_{v_0, I_0}(t_1, \dots, t_d) := (\eta^{v_0 \cdot \deg w_1} t_1, \dots, \eta^{v_0 \cdot \deg w_d} t_d);$$

then

$$\Phi^{\eta\delta} = \Phi^\delta \circ D^\eta_{v_0, I_0}, \qquad Y^{\eta\delta}_w = (D^\eta_{v_0, I_0})^{-1} \eta^{v_0 \cdot \deg w} Y_w \circ D^\eta_{v_0, I_0}.$$

Thus the lemma holds with a slightly smaller (η times the original) value of $\delta(M)$. □

Lemma 4.3. *Let m be a sufficiently large admissible integer, and let $Y_1, \dots, Y_k$ be vector fields with the properties that*

$$\|Y_j\|_{C^m(B(1))} \lesssim 1, \tag{4-12}$$

$$|\det(Y_{w_1}, \dots, Y_{w_d})| \sim 1 \quad \textit{on } B(1); \tag{4-13}$$

here we recall that $(w_1, \dots, w_d) = I_0$. For $J \in \{1, \dots, k\}^d$, define

$$\widetilde{\Psi}^J(t_1, \dots, t_d) := e^{t_d Y_{J_d}} \circ \cdots \circ e^{t_1 Y_{J_1}}(0).$$

Then

$$\max_{J \in \{1, \dots, k\}^d} \|\det D\widetilde{\Psi}^J\|_{C^0(B(c_0))} \sim 1 \tag{4-14}$$

for some admissible constant $c_0 > 0$; in particular, $\widetilde{\Psi}^J$ is defined on the ball $B(c_0)$.

Proof. There are similar results in [Christ 2008; Christ et al. 1999; Stovall 2011; Tao and Wright 2003], but without the uniformity, so we give a complete proof.

The upper bound $\|\det D\widetilde{\Psi}^J\|_{C^0(B(c_0))} \sim 1$ is an immediate consequence of (4-12) for $m \ge 2$, by Picard's existence theorem.

For the lower bound, we first show that if $m \ge |b_0|_1 + 2$, the left side of (4-14) is nonzero. For $1 \le i \le d$ and $J \in \{1, \dots, k\}^i$, define

$$\widetilde{\Psi}^J_i(t_1, \dots, t_i) := e^{t_i Y_{J_i}} \circ \cdots \circ e^{t_1 Y_{J_1}}(0);$$

$\widetilde{\Psi}^J_i \in C^{m+1}(B(c_0))$ for admissible $c_0 > 0$ by standard ODE existence results. Supposing that the left side of (4-14) is zero, there exists some minimal $i \in \{0, \dots, d-1\}$ such that

$$\max_{J \in \{1, \dots, k\}^{i+1}} \|\partial_{t_1} \widetilde{\Psi}^J_{i+1} \wedge \cdots \wedge \partial_{t_{i+1}} \widetilde{\Psi}^J_{i+1}\|_{C^0(B(c_0))} = 0.$$

By (4-13), the Y_j cannot all vanish at zero, so this i is at least 1.

By minimality of i, there exist $J \in \{1, \dots, k\}^i$, $t_0 \in \mathbb{R}^i$ with $|t_0| < c_0$, and $\varepsilon > 0$ such that $\widetilde{\Psi}_i^J$ is an injective immersion on $\{t \in \mathbb{R}^i : |t - t_0| < \varepsilon\} =: B_{t_0}(\varepsilon)$. Our assumption and the definition of exponentiation imply that, for all $1 \le j \le k$ and $(t_1, \dots, t_i) \in B(c_0)$,

$$0 = \big(\partial_{t_1} \widetilde{\Psi}_{i+1}^{(J,j)} \wedge \cdots \wedge \partial_{t_{i+1}} \widetilde{\Psi}_{i+1}^{(J,j)}\big)(t_1, \dots, t_i, 0) = \big(\partial_{t_1} \widetilde{\Psi}_i^J \wedge \cdots \wedge \partial_{t_i} \widetilde{\Psi}_i^J\big)(t_1, \dots, t_i) \wedge Y_j\big(\widetilde{\Psi}_i^J(t_1, \dots, t_i)\big).$$

Therefore $Y_1, \dots, Y_k$ are tangent to $\widetilde{\Psi}_i^J(B_{c_0}(\varepsilon))$, as must be any Lie brackets that are defined, in particular all of those up to order m. Since $m \ge |b_0|_1$, this contradicts (4-13). Tracing back, we see that we must have $\det \widetilde{\Psi}^J \not\equiv 0$ on $B(c_0)$ for some $J \in \{1, \dots, k\}^d$.

Now we prove that there is a uniform lower bound for $m := |b_0|_1 + 3$. If not, there exists a sequence $\big(Y_1^{(n)}, \dots, Y_k^{(n)}\big)$ satisfying hypotheses (4-12) and (4-13), but with

$$\max_{J \in \{1,\dots,k\}^d} \|\det D\widetilde{\Psi}^{(n),J}\|_{C^0(B(c_0))} \to 0,$$

where $\widetilde{\Psi}^{(n),J}(t_1, \dots, t_d) := \exp\big(t_d Y_{J_d}^{(n)}\big) \circ \cdots \circ \exp\big(t_1 Y_{J_1}^{(n)}\big)(0)$. By Arzelà–Ascoli, after passing to a subsequence, each $\big(Y_j^{(n)}\big)$ converges in $C^{m-1}(B(1))$ to some vector field Y_j. Thus for $|\deg w|_1 \le m-1$, $Y_w^{(n)} \to Y_w$, and by standard ODE results, for each J, the sequence $(\widetilde{\Psi}^{(n),J})$ converges to $\widetilde{\Psi}^J$ in $C^m(B(c_0))$. So $Y_1, \dots, Y_k$ satisfy hypotheses (4-12) and (4-13) (the former with $m = |b_0|_1 + 2$), but $\det D\widetilde{\Psi}^J \equiv 0$ on $B(c_0)$ for all $J \in \{1, \dots, k\}^d$. This is impossible, so the lower bound in (4-14) must hold. □

We return to a consideration of the vector fields $X_1, \dots, X_k$ in the next lemma, where we transfer the inequality in Lemma 4.3 from $\widetilde{\Psi}^J$ to Ψ^J.

Lemma 4.4. *For $J \in \{1, \dots, k\}^d$ and $\alpha \in \mathbb{Z}_0^d$, if $v_0 \cdot (\deg J + \deg_J \alpha) < v_0 \cdot b_0$, then $\partial^\alpha \det D\Psi^J(0) = 0$. Furthermore,*

$$\sum_{J \in \{1,\dots,k\}^d} \; \sum_{\substack{\alpha \in (\mathbb{Z}_0)^d \\ v_0 \cdot (\deg J + \deg_J \alpha) = v_0 \cdot b_0}} |\partial^\alpha \det D\Psi^J(0)| \sim |\lambda_{I_0}(0)|. \tag{4-15}$$

Proof. For $J \in \{1, \dots, k\}^d$, let

$$\Psi^{J,\delta} := \Psi^J \circ D_J^\delta, \quad \text{where } D_J^\delta(t_1, \dots, t_d) := (\delta^{v_0^{J_1}} t_1, \dots, \delta^{v_0^{J_d}} t_d),$$
$$\widetilde{\Psi}^{J,\delta} := e^{t_d Y_{J_d}^\delta} \circ \cdots \circ e^{t_1 Y_{J_1}^\delta}(0),$$

with $Y_1^\delta, \dots, Y_k^\delta$ as in (4-3). By naturality of exponentiation, $\Psi^{J,\delta} = \Phi^\delta \circ \widetilde{\Psi}^{J,\delta}$, where Φ^δ is defined in (4-2). Hence by Lemmas 4.2 and 4.3,

$$\max_{J \in \{1,\dots,k\}^d} \|\det D\Psi^{J,\delta}\|_{C^0(B(c_0))} \sim \delta^{v_0 \cdot b_0} |\lambda_{I_0}(0)|, \quad 0 < \delta < \delta(m), \tag{4-16}$$

where $m = m(b_0, d)$ is sufficiently large and $\delta(m)$ is the (inadmissible) constant from Lemma 4.2. As we will see, the lemma follows by sending $\delta \searrow 0$.

Let $M = M(b_0, d)$ be a sufficiently large integer, let $J \in \{1, \dots, k\}^d$, and let $P^{J,\delta}$ be the degree M Taylor polynomial of $\det D\Psi^{J,\delta}$, centered at 0. Then

$$\begin{aligned}\|P^{J,\delta} - \det D\Psi^{J,\delta}\|_{C^0(B(c_0))} &= \Big(\frac{\delta}{\delta(m)}\Big)^{v_0\cdot \deg J} \|P^{J,\delta(m)} - \det D\Psi^{J,\delta(m)}\|_{C^0(D^{\delta/\delta(m)}B(c_0))} \\ &\lesssim i\Big(\frac{\delta}{\delta(m)}\Big)^{v_0\cdot \deg J+(M+1)\min_i v_0^i} \|\det D\Psi^{J,\delta(m)}\|_{C^0(D^{\delta/\delta(m)}B(c_0))} \\ &\lesssim \Big(\frac{\delta}{\delta(m)}\Big)^{v_0\cdot \deg J+(M+1)\min_i v_0^i}, \end{aligned} \tag{4-17}$$

where the first inequality is by Taylor's theorem and admissibility of M, and the second is from (4-8), if m is sufficiently large, depending on M. Motivated by this inequality, we assume that $v_0 \cdot b_0 < M \min_i v_0^i$.

By the equivalence of all norms on the space of polynomials of d variables of degree at most M,

$$\|P^{J,\delta}\|_{C^0(B(c_0))} \sim \sum_{|\alpha|_1 \le M} |\partial^\alpha P^{J,\delta}(0)| = \sum_{|\alpha|_1\le M} \delta^{v_0\cdot(\deg J + \deg_J \alpha)} |\partial^\alpha \det D\Psi^J(0)|. \tag{4-18}$$

If $\alpha \in \mathbb{Z}_0^d$ and $v_0 \cdot (\deg J + \deg_J \alpha) \le v_0 \cdot b_0$, then $|\alpha|_1 \le (v_0 \cdot \deg_J \alpha)/\min_i v_0^i \le M$, and

$$\begin{aligned}\delta^{v_0\cdot(\deg J+\deg_J\alpha)}|\partial^\alpha \det D\Psi^J(0)| = |\partial^\alpha P^{J,\delta}(0)| &\lesssim \|P^{J,\delta}\|_{C^0(B(c_0))} \\ &\lesssim \|\det D\Psi^{J,\delta}\|_{C^0(B(c_0))} + \Big(\frac{\delta}{\delta(m)}\Big)^{v_0\cdot \deg J+(M+1)\min_i v_0^i} \\ &\lesssim \delta^{v_0\cdot b_0}|\lambda_{I_0}(0)| + \Big(\frac{\delta}{\delta(m)}\Big)^{v_0\cdot \deg J+(M+1)\min_i v_0^i}.\end{aligned}$$

Sending $\delta \searrow 0$, we see that

$$\partial^\alpha \det D\Psi^J(0) = 0 \quad \text{whenever } v_0\cdot(\deg J + \deg_J\alpha) < v_0 \cdot b_0, \tag{4-19}$$

$$|\partial^\alpha \det D\Psi^J(0)| \lesssim |\lambda_{I_0}(0)| \quad \text{if } v_0\cdot(\deg J+\deg_J \alpha) = v_0\cdot b_0. \tag{4-20}$$

Now for the lower bound. By (4-16) and the fact that there are only finitely many choices for J, there exist $J \in \{1, \dots, k\}^d$ and a sequence $\delta_n \searrow 0$ such that

$$\|\det D\Psi^{J,\delta_n}\|_{C^0(B(c_0))} \gtrsim \delta_n^{v_0\cdot b_0}|\lambda_{I_0}(0)|. \tag{4-21}$$

Since $M \min_i v_0^i > v_0 \cdot b_0$ and $\lambda_{I_0}(0) \neq 0$, (4-21), (4-17) and (4-18) imply that for δ_n sufficiently (inadmissibly) small,

$$\delta_n^{v_0\cdot b_0}|\lambda_{I_0}(0)| \lesssim \|P^{J,\delta_n}\|_{C^0(B(c_0))} \lesssim \sum_{|\alpha|_1\le M} \delta_n^{v_0\cdot(\deg J+\deg_J\alpha)}|\partial^\alpha \det D\Psi^J(0)|.$$

Applying (4-19) and letting $n \to \infty$,

$$|\lambda_{I_0}(0)| \lesssim \sum_{v_0\cdot(\deg J+\deg_J\alpha)=v_0\cdot b_0} |\partial^\alpha \det D\Psi^J(0)|.$$

This completes the proof of (4-15), and thus of Lemma 4.4. □

By our choice of v_0, (4-15) is just (2-10), so to complete the proof of Proposition 2.3, it suffices to prove the following.

Lemma 4.5. $\mathcal{P}_0 = \widetilde{\mathcal{P}}_0$.

Proof. By (2-10), $\widetilde{\mathcal{P}}_0$ contains the extreme points of $\mathcal{P}_0$, so $\mathcal{P}_0 \subseteq \widetilde{\mathcal{P}}_0$. Now suppose that $b_0 \notin \mathcal{P}_0$. Then there exist $v_0 \in (0,1]^k$ and $\varepsilon > 0$ such that $v_0 \cdot b_0 + \varepsilon < v_0 \cdot p$, for all $p \in \mathcal{P}_0$. At least one extreme point b of $\mathcal{P}_0$ satisfies $v_0 \cdot b = \max_{p \in \mathcal{P}_0} v_0 \cdot p$; perturbing v_0 slightly, we may assume that there exists $b_1 \in \mathcal{P}_0$ such that

$$v_0 \cdot b_0 < v_0 \cdot b_1 < v_0 \cdot p, \quad \text{for all } p \in \mathcal{P}_0 \text{ with } p \neq b_1.$$

By Lemma 4.4, $\partial^\alpha \det D\Psi^J(0) = 0$ whenever $(\deg J + \deg_J \alpha) \cdot v_0 < v_0 \cdot b_1$, so $b_0 \notin \widetilde{\mathcal{P}}_0$. Thus $\mathcal{P}_0 \subseteq \widetilde{\mathcal{P}}_0$, and we are done. □

Remarks. A more direct argument, using the Baker–Campbell–Hausdorff formula, should be possible, but the author has not been able to carry this out. Let $k = d$ and consider vector fields $X_1, \dots, X_d$. Using the approximation $\exp(tX) = \sum_{n=0}^{N} (t^n/n!) X^{n-1}(X) + O(|t|^N)$ [Christ et al. 1999], the formula for the Lie derivative of a determinant of d vector fields, and somewhat tedious computations, one can show that

$$\partial_t^\alpha|_{t=0} \det D_t\big(e^{t_d X_d} \circ \cdots \circ e^{t_1 X_1}\big)(x_0) = \pm \sum_{w_1,\dots,w_d}^{*} \prod_{i=1}^{d} \binom{\alpha_i}{\deg_i w_{i+1}, \dots, \deg_i w_d} \det(X_{w_1}, X_{w_2}, \dots, X_{w_d}),$$

where $*$ indicates the sum is over those words $w_i = (w_i^1, \dots, w_i^{n_i})$ that satisfy $\sum_i \deg w_i = \alpha + (1, \dots, 1)$ and $w_i^1 = i > w_i^2 \geq \cdots \geq w_i^{n_i}$ (in particular, $w_1 = (1)$). Replacing X_i above with X_{J_i} gives an alternative proof that the right (Jacobian) side of (2-10) is bounded by the left (determinant) side, but using this formula to bound the left of (2-10) by the right seems nontrivial.

The estimate (2-10) may fail if b is not extreme (even if it is minimal). To see this, let $\gamma(t) := (t, \dots, t^d)$ and define $X_0 := \partial_t$, $X_i := \partial_t - \gamma'(t) \cdot \nabla_x$, $1 \leq i \leq d$, and take $b := (1 + \frac{1}{2}d(d-1), 1, \dots, 1)$. In this case, the only I with $\deg I = b$ and $\lambda_I \not\equiv 0$ are those of the form

$$I = ((1), (j_1), (1, j_2), \dots, (1, \dots, 1, j_d)),$$

with the j_i distinct. Thus the left side of (2-10) is a nonzero dimensional constant. On the other hand, simple combinatorial considerations show that the right side of (2-10) must be identically zero.

Less uniform versions of (2-10) may be found in [Christ et al. 1999; Stovall 2011; Tao and Wright 2003]. Let $X_1, \dots, X_k$ be smooth vector fields and assume that there exists a d-tuple $I = (w_1, \dots, w_d)$ such that $|\lambda_I| \geq 1$ on U. Let $\delta_1, \dots, \delta_k$ be scalars satisfying the smallness and weak comparability conditions

$$\delta_i \leq K, \quad \delta_i \leq K \delta_j^\varepsilon, \quad 1 \leq i, j \leq k.$$

Then [Tao and Wright 2003; Stovall 2011] prove that there exist $N \geq |\deg I|_1$ and N' (depending on I) such that

$$\sum_{|\deg I|_1 \leq N} \Big(\prod_{i=1}^{k} \delta_i^{(\deg I)_i}\Big) |\lambda_I(x_0)| \sim \sum_{J \in \{1,\dots,k\}^d} \sum_{\substack{\alpha \in (\mathbb{Z}_0)^d \\ \deg J + \deg_J \alpha \leq N'}} \Big(\prod_{i=1}^{k} \delta_i^{\deg J + \deg_J \alpha}\Big) |\partial_t^\alpha \det D_t \Psi_{x_0}^J(0)|, \quad x_0 \in U,$$

with inadmissible implicit constants. It is not shown, however, how to remove the dependence of the implicit constant on ε, K, or the X_i, or, in particular, how to remove the assumption that the Hörmander condition holds uniformly.

5. Proof of the optimality result: Proposition 2.2

The entirety of this section will be devoted to the proof of Proposition 2.2. It suffices to prove the proposition when $\operatorname{supp}\mu \subseteq V$, and V and W are bounded open subsets of U with $\overline{V} \subseteq W$, $\overline{W} \subseteq U$. (Recall that U is the set on which the π_i, and hence the X_i, are defined.) By (2-7) with $E_i = \pi_i(V)$ for $1 \le i \le k$, $\mu(V) < \infty$.

Throughout this section, an object will be said to be admissible if it depends (or it is taken from a finite set depending) only on d and $p = (p_1, \dots, p_k)$. All implicit constants will be admissible. The constant $A(\mu)$ will always represent precisely the quantity in (2-7), and in particular will not be allowed to change from line to line.

First suppose that $p_{j_0} < 1$. Without loss of generality, $j_0 = 1$. We may cover $\pi_1(V)$ by $C_{V,\pi_1}\varepsilon^{-(d-1)}$ balls B_i of radius ε, so

$$\mu(V) \le \sum_i \int \chi_{B_1} \circ \pi_1 \prod_{j=2}^{k} \chi_{\pi_j(V)} \circ \pi_j \, d\mu \le A(\mu) \sum_i |B_1|^{1/p_1} \prod_{j=2}^{k} |\pi_j(V)|^{1/p_j}$$
$$\le C(\mu, d, p, V, \pi_2, \dots, \pi_k)\varepsilon^{(d-1)(1/p_1-1)}.$$

Letting $\varepsilon \to 0$, we see that $\mu \equiv 0$.

We now turn to the case when $\sum_j p_j^{-1} > 1$. Replacing $\{X_1, \dots, X_k\}$ with $\{\partial_1, \dots, \partial_d, X_1, \dots, X_k\}$, $(p_1, \dots, p_k)$ with $(\infty, \dots, \infty, p_1, \dots, p_k)$, and k with $d + k$ if necessary, we may assume that $X_i = \partial_i$, $1 \le i \le d$, without affecting either of the sets

$$Z := \{x \in V : b_p \notin \mathcal{P}_x\}, \quad \Omega := \{x \in V : b_p \text{ is an extreme point of } \mathcal{P}_x\},$$

or the quantity on the right of (2-8).

The proposition will follow from the next two lemmas.

Lemma 5.1. $\mu(Z) = 0$.

Lemma 5.2. *If* $\rho := \sum_{\deg I = b_p} |\lambda_I|^{1/(|b_p|_1 - 1)}$ *and*

$$\Omega_n := \{x \in \Omega : 2^n \le \rho(x) \le 2^{n+1}\}, \quad n \in \mathbb{Z},$$

then $\mu(\Omega') \lesssim A(\mu) 2^n |\Omega'|$ *for any Borel set* $\Omega' \subseteq \Omega_n$.

Proof of Lemma 5.1. By Proposition 4.1, there exist admissible, finite sets $\mathcal{A}_i$, $i = 1, \dots, C_{p,d}$ such that $b_p \notin \mathcal{P}(\mathcal{A}_i)$ for any i and, for each $x \in Z$, there exists an i such that $\mathcal{P}_x \subseteq \mathcal{P}(\mathcal{A}_i)$. For the remainder of the proof of the lemma, we let $\mathcal{A} = \mathcal{A}_i$ be fixed and define

$$Z' := \{x \in Z : \mathcal{P}_x \subseteq \mathcal{P}(\mathcal{A})\}.$$

It suffices to show that $\mu(Z') = 0$.

Choose admissible $\varepsilon > 0$ and $v \in (\varepsilon, 1]^k$ such that

$$v \cdot b_p + \varepsilon < v \cdot b, \quad \text{for } b \in \mathcal{P}(\mathcal{A}).$$

Define

$$\mathcal{W}_0 := \{w \in \mathcal{W} : v \cdot \deg w \leq d\}.$$

Let $N = N_{d,p}$ be an integer whose size will be determined in a moment and which is, in particular, larger than d/ε. Since $\overline{W}$ is compact and contained in U, the X_i are smooth on U and $\{X_w : w \in \mathcal{W}_0\}$ contains the coordinate vector fields, it follows that there exists $\delta_0 > 0$, depending on the π_i, p and W, such that for all $0 < \delta \leq \delta_0$, $I \in \mathcal{W}_0^d$ satisfying $\deg I \in \mathcal{P}(\mathcal{A})$, $x \in W$, and $w, w' \in \mathcal{W}_0$,

$$|\delta^{v \cdot \deg I} \lambda_I(x)| < \delta^{\varepsilon} \delta^{v \cdot b_p}, \tag{5-1}$$

$$\|\delta^{v \cdot \deg w} X_w\|_{C^0(W)} \leq \frac{1}{d} \operatorname{dist}(V, \partial W), \quad \|\delta^{v \cdot \deg w} X_w\|_{C^N(W)} \leq 1, \tag{5-2}$$

$$[\delta^{v \cdot \deg w} X_w, \delta^{v \cdot \deg w'} X_{w'}] = \sum_{\widetilde{w} \in \mathcal{W}_0} c_{w,w'}^{\widetilde{w},\delta} \delta^{v \cdot \deg \widetilde{w}} X_{\widetilde{w}},$$

with

$$\|c_{w,w'}^{\widetilde{w},\delta}\|_{C^N(W)} \lesssim 1.$$

We omit the details since they are essentially the same as arguments found in the proof of Lemma 4.2.

For $x \in Z'$ and $0 < \delta \leq \delta_0$, choose $I_x^\delta \in \mathcal{W}_0^d$ such that

$$\delta^{v \cdot \deg I_x^\delta} |\lambda_{I_x^\delta}(x)| = \max_{I \in \mathcal{W}_0^d} \delta^{v \cdot \deg I} |\lambda_I(x)|.$$

Let

$$\begin{aligned} \Phi_x^\delta(t_1, \dots, t_d) &:= \exp\bigl(t_1 \delta^{v \cdot \deg w_1} X_{w_1} + \cdots + t_d \delta^{v \cdot \deg w_d} X_{w_d}\bigr)(x), \\ B(x, \delta) &:= \{\Phi_x^\delta(t) : |t| < 1\}, \end{aligned} \tag{5-3}$$

where $I_x^\delta = (w_1, \dots, w_d)$. Then $B(x, \delta) \subseteq W$ by (5-2) and the fact that $x \in Z' \subseteq V$.

By the results of [Street 2011], provided $N = N_{d,p}$ is sufficiently large, these balls are doubling in the sense that $|B(x, \delta)| \sim |B(x, 2\delta)|$, for all $x \in Z'$ and $0 < \delta \leq \delta_0$. (Here we are using the fact that ε and v are admissible.) Furthermore, for $x \in V$,

$$|B(x, \delta)| \sim \delta^{v \cdot \deg I_x^\delta} |\lambda_{I_x^\delta}(x)|, \tag{5-4}$$

$$\exp(tX_i)(y) \in B(x, C\delta) \quad \text{whenever } y \in B(x, \delta),\ |t| < \delta^{v^i}, \tag{5-5}$$

where $C = C_{d,p}$. By the doubling property, the change of variables formula and (5-5), if $\sigma_i : \pi_i(W) \to \mathbb{R}^d$ is any smooth section of π_i (i.e., $\sigma_i \circ \pi_i$ is the identity) with $\sigma_i(\pi_i(V)) \subseteq W$, then

$$\begin{aligned} |B(x, \delta)| \sim |B(x, C\delta)| &= \int_{\pi_i(B(x, C\delta))} \int_{\mathbb{R}} \chi_{B(x, C\delta)}(e^{tX_i}(\sigma_i(y)))\, dt\, dy \\ &\geq \int_{\pi_i(B(x, \delta/2))} \int_{\mathbb{R}} \chi_{B(x, C\delta)}(e^{tX_i}(\sigma_i(y)))\, dt\, dy \gtrsim \delta^{v^i} |\pi_i(B(x, \delta))|. \end{aligned} \tag{5-6}$$

By the Vitali covering lemma (as stated in [Stein 1993], for instance), for each $0 < \delta \le \delta_0$ there exists a collection of points $\{x_j\}_{j=1}^{M_\delta} \subseteq Z'$ such that $Z' \subseteq \bigcup_{j=1}^{M_\delta} B(x_j, \delta)$ and such that the balls $B(x_j, C^{-1}\delta)$ are pairwise disjoint. By this, (2-7) and the fact that $\chi_{B(x_j,\delta)} \le \prod_{i=1}^{k} \chi_{\pi_i(B(x_j,\delta))} \circ \pi_i$, (5-6), (5-4) and the definition of b_p, the doubling property and (5-1), and, finally, disjointness of the $B(x_j, \delta)$,

$$\begin{aligned}\mu(Z') &\le \sum_{j=1}^{M_\delta} \mu(B(x_j,\delta)) \le A(\mu) \sum_j \prod_{i=1}^{k} |\pi_i(B(x_j,\delta))|^{1/p_i} \\ &\lesssim A(\mu) \sum_j |B(x_j, C\delta)|^{\sum_i 1/p_i} \prod_i \delta^{-v^i/p_i} \\ &\sim A(\mu) \sum_j |B(x_j, C\delta)| (\delta^{v\cdot \deg I_{x_j}^\delta - v\cdot b_p} |\lambda_{I_{x_j}^\delta}(x_j)|)^{\sum_i 1/p_i - 1} \\ &\lesssim A(\mu) \sum_j |B(x_j, C^{-1}\delta)| \delta^{\varepsilon \sum_i 1/p_i - 1} \le A(\mu) |W| \delta^{\varepsilon \sum_i 1/p_i - 1}.\end{aligned}$$

The lemma follows by sending δ to 0. □

Proof of Lemma 5.2. The proof is similar to that of Lemma 5.1. Fix n and $\Omega' \subseteq \Omega_n$. Let $x \in \Omega'$. Since $\Omega' \subseteq \Omega$, b_p is an extreme point of $\mathcal{P}_x$. By the definition of ρ, $\max_{\deg I = b_p} |\lambda_I(x)| \sim 2^{n(|b_p|_1 - 1)}$.

By Proposition 4.1 and a covering argument, we may assume that there exists a finite set $\mathcal{A} \subseteq \mathbb{Z}_0^k$ such that $b_p \notin \mathcal{P}(\mathcal{A})$ and for each $x \in \Omega'$, $\mathcal{P}_x \subseteq \mathcal{P}(\mathcal{A} \cup \{b_p\})$. Choose $\varepsilon > 0$, $v \in (\varepsilon, 1]^k$ such that $v \cdot b_p + \varepsilon < v \cdot b$ for each $b \in \mathcal{P}(\mathcal{A} \cup \{b_p\}) \cap \mathbb{Z}_0^k \setminus \{b_p\}$, and let

$$\mathcal{W}_0 := \{w \in \mathcal{W} : v \cdot \deg w \le d\}.$$

Since $(1, \dots, 1, 0, \dots, 0) \in \mathcal{P}_x$ for each $x \in U$, $(1, \dots, 1, 0, \dots, 0) \in \mathcal{P}(\mathcal{A} \cup \{b_p\})$. Therefore we have $v \cdot b_p \le \sum_{i=1}^{d} v^i \le d$, so $\deg I = b_p$ implies that $I \in \mathcal{W}_0^d$.

Let $N = N_{d,p}$ be a large integer. As before, there exists $\delta_n > 0$, which depends on n, the π_i and p, such that for all $0 < \delta \le \delta_n$, $x \in \Omega'$, $I \in \mathcal{W}_0^d$ with $\deg I \ne b_p$, and $w, w' \in \mathcal{W}_0$,

$$|\delta^{v\cdot \deg I} \lambda_I(x)| < \delta^\varepsilon \max_{\deg I' = b_p} \delta^{v \cdot \deg I'} |\lambda_{I'}(x)|,$$

$$\|\delta^{v\cdot \deg w} X_w\|_{C^0(W)} \le \frac{1}{d} \operatorname{dist}(V, \partial W), \qquad \|\delta^{v\cdot \deg w} X_w\|_{C^N(W)} \le 1,$$

$$[\delta^{v\cdot \deg w} X_w, \delta^{v\cdot \deg w'} X_{w'}] = \sum_{\widetilde{w} \in \mathcal{W}_0} c_{w,w'}^{\widetilde{w},\delta} \delta^{v\cdot \deg \widetilde{w}} X_{\widetilde{w}},$$

with

$$\|c_{w,w'}^{\widetilde{w},\delta}\|_{C^N(W)} \le C_{d,p},$$

for all $w, w' \in \mathcal{W}_0$. In particular, we may choose δ_n sufficiently small that for each $x \in \Omega'$ and $0 < \delta \le \delta_n$, there exists a d-tuple $I_x^\delta \in \mathcal{W}_0^d$ such that $\deg I_x^\delta = b_p$ and

$$\delta^{v\cdot \deg I_x^\delta} |\lambda_{I_x^\delta}(x)| = \max_{I \in \mathcal{W}_0^d} \delta^{v\cdot \deg I} |\lambda_I(x)| \sim \delta^{v\cdot b_p} 2^{n(|b_p|_1 - 1)}.$$

Thus, considering the balls $B(x,\delta)$ (defined in (5-3)) for $x \in \Omega'$ and $0 < \delta \le \delta_n$,

$$|B(x,\delta)| \sim 2^{n(|b_p|_1-1)}\delta^{v\cdot b_p} = 2^{n/(\sum_i 1/p_i-1)}\delta^{v\cdot b_p}.$$

Since the balls $B(x,\delta)$ are doubling, for each $\eta > 0$ there exist a collection $\{x_j\}_{j=1}^{M_\delta} \subseteq \Omega'$ and a parameter $0 < \delta \le \delta_n$ such that

$$\Omega' \subseteq \bigcup_{j=1}^{M_\delta} B(x_j,\delta), \qquad \Bigl|\bigcup_{j=1}^{M_\delta} B(x_j,\delta)\Bigr| \le |\Omega'| + \eta,$$

and such that the $B(x_j, C^{-1}\delta)$ are pairwise disjoint.

Arguing as in the proof of Lemma 5.1,

$$\begin{aligned}\mu(\Omega') &\le \sum_{j=1}^{M_\delta} \mu(B(x_j,\delta)) \lesssim A(\mu)\sum_j |B(x_j,\delta)||B(x_j,\delta)|^{\sum_i 1/p_i-1}\delta^{-v\cdot b_p(\sum_i 1/p_i-1)}\\ &\sim A(\mu)\sum_j |B(x_j,\delta)|2^n \lesssim A(\mu)2^n(|\Omega'|+\eta).\end{aligned}$$

Letting $\eta \to 0$ completes the proof. □

Remarks. The pointwise upper bound (2-8) is false if no assumptions are made on b_p. Indeed, if b_p lies in the interior of $\mathcal{P}_{x_0}$, then for some $\theta < 1$, $b_{\theta p}$ lies in the interior of $\mathcal{P}_{x_0}$, where $\theta p = (\theta p_1, \dots, \theta p_k)$. Thus for some neighborhood U of x_0, $b_{\theta p}$ lies in the interior of $\mathcal{P}_x$ for every $x \in U$. Hence by the main result in [Stovall 2011], if a is continuous with compact support in U,

$$\Bigl|\int \prod_{j=1}^k f_j \circ \pi_j(x) a(x)\,dx\Bigr| \lesssim \prod_{j=1}^k \|f_j\|_{L^{\theta p_j}}.$$

Additionally,

$$\Bigl|\int \prod_{j=1}^k f_j \circ \pi_j(x)\bigl|\log|x-x_0|\bigr| a(x)\,dx\Bigr| \lesssim \prod_{j=1}^k \|f_j\|_{L^\infty}.$$

Thus by interpolation,

$$\Bigl|\int \prod_{j=1}^k f_j \circ \pi_j(x)\bigl|\log|x-x_0|\bigr|^{1-\theta} a(x)\,dx\Bigr| \lesssim \prod_{j=1}^k \|f_j\|_{L^{p_j}}.$$

For the unweighted bilinear operator in the "polynomial-like" case, the endpoint-restricted weak-type bounds are known and are due to Gressman [2009]; in the multilinear case, the corresponding estimates follow by combining his techniques with arguments in [Stovall 2011]. The deduction of endpoint bounds from the arguments in [Gressman 2009] does not seem to be immediate in the weighted case, and so these questions remain open except for certain special configurations (such as convolution or restricted X-ray transform along polynomial curves).

6. Proof of the main theorem: Theorem 2.1

In this section, undecorated constants and implicit constants (C, c, $\lesssim$, $\gtrsim$, $\sim$) will be allowed to depend on a cutoff function a (specifically, on upper bounds for $\operatorname{diam}(\operatorname{supp} a)$ and $\|a\|_{L^\infty}$), a point $b_0 \in \mathbb{Z}_0^k$, and exponents $p_1, \dots, p_k$ (all of which will be given in a moment), as well as the π_j. Other parameters (namely ε, δ, N) that depend on b_0, $p_1, \dots, p_k$ will arise later on, so implicit constants may depend on these quantities as well. Unless otherwise stated, decorated constants and implicit constants (c_d, $\lesssim_{N,d}$, etc.) will only be allowed to depend on the objects in their subscripts.

Let $J_0 \in \{1, \dots, k\}^d$ and for $x \in U$ define $\Psi_x^{J_0}(t)$ as in (2-9). Let β_0 be a multiindex and define $b_0 := \deg J_0 + \deg_{J_0} \beta_0$. Let

$$\tilde{\rho}(x) := \left|\partial_t^{\beta_0}|_{t=0} \det D_t \Psi_x^{J_0}(t)\right|^{1/(|b_0|_1-1)}. \tag{6-1}$$

Let a be continuous and compactly supported in U, and define the multilinear form

$$\widetilde{M}(f_1, \dots, f_k) := \int_{\mathbb{R}^d} \prod_{j=1}^k f_j \circ \pi_j(x) \tilde{\rho}(x) a(x)\, dx.$$

In light of Proposition 2.3, the following more general result (we need not assume that b_0 is extreme) implies Theorem 2.1.

Theorem 6.1. *Let $(p_1, \dots, p_k) \in [1, \infty)^k$ satisfy $(p_1^{-1}, \dots, p_k^{-1}) \prec \boldsymbol{q}(b_0)$, with $p_i^{-1} < \boldsymbol{q}_i(b_0)$ when $b_0^i \neq 0$. Then*

$$|\widetilde{M}(f_1, \dots, f_k)| \lesssim \prod_{j=1}^k \|f_j\|_{L^{p_j}}, \tag{6-2}$$

for all continuous $f_1, \dots, f_k$.

Since J_0 and β_0 are fixed, we will henceforth drop the tildes from our notation, with the understanding that we are using (6-1) instead of (2-3) to define ρ.

It suffices to prove (6-2) when the f_j are nonnegative. Suppose that $b_j = 0$ for some j. Then π_j plays no role in the definition of ρ, and $p_j = \infty$ so, by Hölder's inequality, we may ignore f_j entirely. Thus we may assume that $b_j \neq 0$ for each j. In fact, we may assume that, for each j, $p_j < \infty$, since $\|f_j\|_{L^{p_j}(\pi_j(\operatorname{supp} a))} \lesssim \|f_j\|_{L^\infty}$, by the compact support of a.

We only claim a nonendpoint result, so by real interpolation with the trivial (by Hölder) inequalities of the form

$$M(f_1, \dots, f_k) \lesssim \prod_{j=1}^k \|f_j\|_{L^{\tilde{p}_j}}, \quad \sum_{j=1}^k p_j^{-1} \leq 1,$$

it suffices to prove that, for all Borel sets $E_1, \dots, E_k$ and some sufficiently small $\varepsilon > 0$,

$$\int_{\mathbb{R}^d} \prod_{j=1}^k \chi_{E_j} \circ \pi_j(x) \rho(x) a(x)\, dx \lesssim \prod_{j=1}^k |E_j|^{\boldsymbol{q}_j(b_0)-\varepsilon}. \tag{6-3}$$

Letting $\Omega := \operatorname{supp} a \cap \bigcap_{j=1}^{k} \pi_j^{-1}(E_j)$, (6-3) will follow from

$$\rho(\Omega) \lesssim \prod_{j=1}^{k} |\pi_j(\Omega)|^{\boldsymbol{q}_j(b_0)-\varepsilon}. \tag{6-4}$$

If we define

$$\alpha_j := \frac{\rho(\Omega)}{|\pi_j(\Omega)|}, \tag{6-5}$$

a bit of arithmetic shows that (6-4) is equivalent to

$$\prod_{j=1}^{k} \alpha_j^{\boldsymbol{q}_j(\boldsymbol{q}(b_0)-(\varepsilon,\dots,\varepsilon))} \lesssim \rho(\Omega),$$

which in turn would be implied by

$$\prod_{j=1}^{k} \alpha_j^{b_0^j+\varepsilon} \lesssim \rho(\Omega), \tag{6-6}$$

with a slightly smaller ε. (We recall that $\boldsymbol{q}$ equals its own inverse.)

By the coarea formula,

$$\alpha_j = |\pi_j(\Omega)|^{-1} \int_{\pi_j(\Omega)} \int_{\pi_j^{-1}\{y\}} \chi_\Omega(x)\rho(x) \frac{1}{|X_j(x)|} \, d\mathcal{H}^1(x)\, dy. \tag{6-7}$$

Since π_j is a submersion, $|X_j| \gtrsim 1$ and $\mathcal{H}^1(\pi_j^{-1}\{y\}) \lesssim 1$ for all $y \in \pi_j(\Omega)$. Since $\rho \lesssim 1$ by smoothness of the π_j, (6-7) implies that

$$\alpha_j \lesssim \operatorname{diam}(\Omega) \le \operatorname{diam}(\operatorname{supp} a). \tag{6-8}$$

By taking a partition of unity, we may assume that the α_j are as small as we like, in particular, that they are smaller than $\frac{1}{2}$. Reordering if necessary, $\alpha_1 \le \cdots \le \alpha_k$.

For $n \in \mathbb{Z}$, let $\Omega_n = \{x \in \Omega : 2^n \le \rho(x) < 2^{n+1}\}$. Then for C sufficiently large, $\Omega_n = \varnothing$ for all $n > C$. On the other hand, since π_1 is a submersion and $\operatorname{supp} a$ is compact,

$$\sum_{n \le \log \alpha_1 - C} \rho(\Omega_n) \lesssim \sum_{n \le \log \alpha_1 - C} 2^n |\pi_1(\Omega)| \lesssim 2^{-C} \alpha_1 |\pi_1(\Omega)| = 2^{-C} \rho(\Omega).$$

Thus, for C sufficiently large,

$$\rho\Bigl(\bigcup_{n \le \log \alpha_1 - C} \Omega_n\Bigr) < \tfrac{1}{2}\alpha_1 |\pi_1(\Omega)| = \tfrac{1}{2}\rho(\Omega).$$

By pigeonholing, there exists n with $\log \alpha_1 - C \le n \le C$ such that

$$\rho(\Omega_n) \ge \bigl(2(|\log \alpha_1| + 2C)\bigr)^{-1} \rho(\Omega) \gtrsim \alpha_1^{\varepsilon} \rho(\Omega). \tag{6-9}$$

Define

$$\alpha_{n,j} := \frac{\rho(\Omega_n)}{|\pi_j(\Omega_n)|}, \qquad j = 1, \dots, k.$$

By (6-9) and the triviality $\rho(\Omega_n) \le \rho(\Omega)$, together with the proof of (6-8) and the small diameter of $\operatorname{supp} a$,

$$\alpha_1^{\varepsilon} \alpha_j \lesssim \alpha_{n,j} \le \tfrac{1}{2}.$$

Therefore (6-6) follows from

$$\rho(\Omega_n) \gtrsim \prod_{j=1}^{k} (\alpha_{n,j})^{b_0^j+\varepsilon}, \tag{6-10}$$

with a slightly smaller value of ε. Henceforth, we let $\rho_0 := 2^n$ (for this value of n) and drop the n from the notation in (6-10). We note that $\rho(\Omega) \sim \rho_0|\Omega|$. Reordering again, we may continue to assume that $\alpha_1 \le \cdots \le \alpha_k$.

Let $\delta > 0$ be a small constant (depending on ε, b_0, d), which will be determined later on. Cover Ω by $c_d\alpha_1^{-\delta d}$ balls of radius α_1^{δ}. By pigeonholing, there exists $\Omega' \subseteq \Omega$ with

$$\rho(\Omega') \gtrsim \alpha_1^{\delta d}\rho(\Omega).$$

Arguing as above, the parameters $\alpha_j' := |\pi_j(\Omega')|^{-1}\rho(\Omega')$ satisfy

$$\alpha_1^{1+\delta d} \le \alpha_1^{\delta d}\alpha_j \lesssim \alpha_j' \lesssim \operatorname{diam}(\Omega') \le \alpha_1^{\delta}. \tag{6-11}$$

Thus, for δ sufficiently small, (6-10) would follow from

$$\rho(\Omega') \gtrsim \prod_{j=1}^{k} (\alpha_j')^{b_0^j+\varepsilon},$$

with a slightly smaller value of ε.

Since $\alpha_j' \lesssim \operatorname{diam}(\operatorname{supp} a)$, we may assume that the α_j' are as small as we like (depending on the π_j, ε and δ). Thus (6-11) implies that, for each $1 \le j \le k$,

$$\operatorname{diam}(\Omega') \le c(\alpha_j')^{\delta},$$

for some slightly smaller value of δ and with c as small as we like. By the same argument as for (6-8),

$$\alpha_j' \lesssim \rho_0 \operatorname{diam}(\Omega') \lesssim \rho_0(\alpha_j')^{\delta},$$

whence $\rho_0 \ge c^{-1}(\alpha_j')^{1-\delta}$, again with a slightly smaller value of δ.

In summary, to complete the proof of Theorem 6.1 (and thereby that of Theorem 2.1) it suffices to prove the following.

Lemma 6.2. *Let $\varepsilon > 0$ be sufficiently small depending on b_0 and $\delta > 0$ be sufficiently small depending on ε, b_0. Let $\Omega \subseteq \operatorname{supp} a$ be a Borel set, and define $\alpha_1, \dots, \alpha_k$ as in (6-5). Assume that $\alpha_1 \le \dots \le \alpha_k$, that*

$$\rho_0 \le \rho(x) \le 2\rho_0 \quad \textit{for all } x \in \Omega,$$

and that

$$\alpha_k < c, \quad \rho_0 \ge c^{-1}\alpha_k^{1-\delta}, \quad \operatorname{diam}(\Omega) \le c\alpha_1^{\delta}. \tag{6-12}$$

Then for c sufficiently small, depending on the π_j, b_0, ε, δ, we have

$$\prod_{j=1}^{k} \alpha_j^{b_0^j+\varepsilon} \lesssim \rho(\Omega). \tag{6-13}$$

We note in particular that all constants and implicit constants are independent of ρ_0, Ω, and the α_j.

We devote the remainder of this section to the proof of Lemma 6.2. We use the method of refinements, which originated in [Christ 1998] and was further developed in similar contexts in [Christ 2008; Tao and Wright 2003].

Recalling (6-1),

$$|\partial^{\beta_0} \det D\Psi_{x_0}^{J_0}(0)| \sim \rho_0^{|b_0|_1 - 1} =: \lambda_0, \quad \text{for } x_0 \in \Omega. \tag{6-14}$$

As in [Tao and Wright 2003], for $w > 0$, we say that a set $S \subseteq [-w, w]$ is a central set of width w if, for any interval $I \subseteq [-w, w]$,

$$|I \cap S| \lesssim \left(\frac{|I|}{w}\right)^{\varepsilon} |S|.$$

Lemma 6.3. *For each subset $\Omega' \subseteq \Omega$ with $\rho(\Omega') \gtrsim \alpha_1^{C\varepsilon}\rho(\Omega)$ and each $1 \le j \le k$, there exists a refinement $\langle \Omega' \rangle_j \subseteq \Omega'$ with $\rho(\langle \Omega' \rangle_j) \gtrsim \alpha_1^{2C\varepsilon}\rho(\Omega')$ such that, for each $x \in \langle \Omega' \rangle_j$, there is a central set*

$$\mathcal{F}_j(x, \langle \Omega' \rangle_j) \subseteq \{t : |t| \lesssim \alpha_1^{\delta} \text{ and } e^{tX_j}(x) \in \langle \Omega' \rangle_j\} \tag{6-15}$$

whose width w_j and measure satisfy

$$\rho_0^{-1}\alpha_1^{2C\varepsilon}\alpha_j \lesssim w_j \le c\alpha_1^{\delta} \quad \text{and} \quad |\mathcal{F}_j(x, \langle \Omega' \rangle_j)| \gtrsim \rho_0^{-1}\alpha_1^{2C\varepsilon}\alpha_j. \tag{6-16}$$

This lemma has essentially the same proof as [Tao and Wright 2003, Lemma 8.2], but we sketch the argument for the convenience of the reader.

Sketch proof of Lemma 6.3. First we discard shorter-than-average π_j fibers in Ω', leaving a subset $\Omega'' \subseteq \Omega'$ with $\rho(\Omega'') \gtrsim \rho(\Omega')$ such that, for each $x \in \Omega''$,

$$\left|\{t : |t| \lesssim \alpha_1^{\delta} \text{ and } e^{tX_j}(x) \in \Omega''\}\right| \gtrsim \frac{|\Omega'|}{|\pi_j(\Omega')|} \gtrsim \alpha_1^{C\varepsilon}\rho_0^{-1}\alpha_j.$$

Next, if $S \subseteq [-c\alpha_1^{\delta}, c\alpha_1^{\delta}]$ is a measurable set, it contains a translate S' of a central set of measure at least $|S|^{1+2\varepsilon}$ and width at most $c\alpha_1^{\delta}$. Indeed, take $S' = S \cap I'$, where I' is a minimal length dyadic interval with $|S \cap I'| \ge (|I'|/\alpha_1^{\delta})^{\varepsilon}|S|$.

Using the exponential map, each π_j fiber in Ω'' is naturally associated to a set $S \subseteq [-c\alpha_1^{\delta}, c\alpha_1^{\delta}]$; S can be refined to a translate S' of a central set, and S' is then a fiber of the set $\langle \Omega' \rangle_j$. By the definition of exponentiation, for $x \in \langle \Omega' \rangle_j$ the set $\mathcal{F}_j(x, \langle \Omega' \rangle_j)$ in (6-15) contains 0, and it is easy to see that a 0-containing translate of a central set of width w is a central set of width $2w$. Finally, by pigeonholing, we can select only those fibers having the most popular dyadic width (there are at most $\log \alpha_1$ options). □

Write $J_0 = (j_1, \dots, j_d)$. With $\Omega_0 := \Omega$, for $1 \le i \le d$ we define

$$\Omega_i := \langle \Omega_{i-1} \rangle_{j_{d-i+1}}.$$

By Lemma 6.3, for each i, $\rho(\Omega_i) \gtrsim \alpha_1^{C\varepsilon}\rho(\Omega)$.

Fix $x_0 \in \Omega_d$. Let

$$F_1 := \mathcal{F}_{j_1}(x_0, \Omega_d), \quad x_1(t) := e^{tX_{j_1}}(x_0),$$

and for $2 \le i \le d$, let

$$F_i := \{(t_1, \dots, t_i) : (t_1, \dots, t_{i-1}) \in F_{i-1},\ t_i \in \mathscr{F}_{j_i}(x_{i-1}(t_1, \dots, t_{i-1}),\ \Omega_{d-i+1})\}$$
$$x_i(t_1, \dots, t_i) := e^{t_i X_{j_i}} x_{i-1}(t_1, \dots, t_{i-1}).$$

By construction, for each i and each $(t_1, \dots, t_i) \in F_i$,

$$x_i(t_1, \dots, t_i) \in \Omega_{d-i+1} \subseteq \Omega_{d-i},$$

so $\mathscr{F}_{j_{i+1}}(x_i(t_1, \dots, t_i), \Omega_{d-i})$ is a central set whose width and measure satisfy (6-16) (with j_{i+1} in place of j). Furthermore,

$$\Psi_{x_0}^{J_0}(F_d) \subseteq \Omega \quad \text{and} \quad |F_d| \gtrsim \rho_0^{-d} \alpha_1^{C\varepsilon} \alpha^{\deg J_0}; \tag{6-17}$$

here we recall that $\deg J$ is the k-tuple whose i-th entry is the number of appearances of i in the d-tuple J.

Let $\Psi_{x_0}^N$ be the degree N Taylor polynomial of $\Psi_{x_0}^{J_0}$, where $N \ge |b_0|_1 + 1$ is a large integer to be chosen later. Let $Q_w = \prod_{i=1}^d [-w_i, w_i]$ and let $Q_1 = Q_{(1,\dots,1)}$. By scaling, the equivalence of all norms on the degree N polynomials in d variables, and (6-14),

$$\begin{aligned}\|\det D\Psi_{x_0}^N\|_{C^0(Q_w)} = \sup_{t \in Q_1} |\det D\Psi_{x_0}^N(w_1 t_1, \dots, w_d t_d)| &\sim_{N,d} \sum_\beta w^\beta |\partial^\beta \det D\Psi_{x_0}^N(0)| \\ &\ge w^{\beta_0} |\partial^{\beta_0} \det D\Psi_{x_0}^N(0)| \sim w^{\beta_0} \lambda_0.\end{aligned}$$

Thus, by (6-16), the definition of λ_0, and some arithmetic,

$$\|\det D\Psi_{x_0}^N\|_{C^0(Q_w)} \gtrsim \rho_0^{d-1} \alpha_1^{C\varepsilon} \alpha^{\deg_{J_0} \beta_0}. \tag{6-18}$$

(We recall that $\deg_J \beta$ is the k-tuple whose i-th entry equals $\sum_{\ell : J_\ell = i} \beta_\ell$.)

Lemma 6.4. *If P is any degree N polynomial on $\mathbb{R}^d$, there is a subset $F_d' \subseteq F_d$ such that $|F_d'| \gtrsim_{N,\varepsilon,d} |F_d|$ and*

$$|P(t)| \gtrsim_{N,\varepsilon,d} \|P\|_{C^0(Q_w)} \quad \textit{for } t \in F_d'.$$

The lemma follows from [Christ 2008, Lemma 6.2] or [Tao and Wright 2003, Lemma 7.3]. Roughly, if S is a central set of width w_0 and p is a degree N polynomial, p is close to $\|p\|_{C^0([-w_0,w_0])}$ on most of S. This is because the set where p is small is the union of at most N small intervals. Recalling how our set F_d was constructed (from a "tower" of central sets), it is possible to iterate d times to obtain the lemma.

Now we use $\Psi_{x_0}^N$ to control $\Psi_{x_0}^{J_0}$ via the following lemma, which just paraphrases [Christ 2008, Lemma 7.1]. We recall that Q_1 is the unit cube.

Lemma 6.5. *Let $N, C_1, c_2, c_3 > 0$. There exists a constant $c_0 > 0$, depending on C_1, c_2, c_3, N and d, such that the following holds. Let $\Psi : Q_1 \to \mathbb{R}^d$ be twice continuously differentiable and let $\Psi^N : \mathbb{R}^d \to \mathbb{R}^d$ be a degree N polynomial. Set $J_\Psi := \|\det D\Psi\|_{C^0(Q_1)}$ and assume that*

$$\|\Psi\|_{C^0(Q_1)} \le C_1, \quad \|\Psi - \Psi^N\|_{C^2(Q_1)} \le c_0 \mathscr{J}_\Psi^2. \tag{6-19}$$

Let $G \subseteq Q_1$ be a Borel set with the property that, for any degree N^d polynomial $P : \mathbb{R}^d \to \mathbb{R}$,

$$\bigl|\{t \in G : |P(t)| \geq c_2 \|P\|_{C^0(Q_1)}\}\bigr| \geq c_3 |G|. \tag{6-20}$$

Then

$$|\Psi(G)| \geq c_0 |G| \|\det D\Psi^N\|_{C^0(Q_1)}.$$

For the complete details, see [Christ 2008]. We give a quick sketch of that argument here.

Sketch proof of Lemma 6.5. Let $P = \det D\Psi^N$ and let G' denote the set on the left of (6-20). By (6-19),

$$|\det D\Psi(t)| \sim |P(t)| \sim \|P\|_{C^0(Q_1)} \sim \mathcal{J}_\Psi, \text{ for } t \in G', \quad \text{and} \quad \|\Psi^N\|_{C^2(Q_1)} \leq 2C_1. \tag{6-21}$$

This first series of inequalities above imply that

$$\int_{G'} |\det D\Psi| \geq c_0^{1/2} |G| \|\det D\Psi^N\|_{C^0(Q_1)}.$$

It remains to show that Ψ is finite-to-one on G', so that $|\Psi(G')| \gtrsim \int_{G'} |\det D\Psi|$.

First the local case. For c_0 sufficiently small and B any ball with radius $c_0^{1/2} \mathcal{J}_\Psi$ and center in G', Ψ, Ψ^N may be shown to be one-to-one on $10B$ and to satisfy

$$|\det D\Psi(t)| \sim |P(t)| \sim \mathcal{J}_\Psi, \quad t \in 10B. \tag{6-22}$$

We cover G' by a finitely overlapping collection of such balls B.

Globally, we know (it is an application of Bezout's theorem) that Ψ^N is at most $C_{N,d}$-to-one on G'. Thus a point $x \in \mathbb{R}^d$ lies in $\Psi^N(10B)$ for at most $C_{N,d}$ balls $B \in \mathcal{B}$. We are done if we can show that $\Psi(B) \subseteq \Psi^N(10B)$. By the mean value theorem (applied to $(\Psi^N)^{-1}$), then Cramer's rule, (6-21) and (6-22),

$$\operatorname{dist}\bigl(\Psi^N(B), (\Psi^N(10B))^c\bigr) \geq \operatorname{dist}(B, (10B)^c) \|(D\Psi^N)^{-1}\|_{C^0(10B)}^{-1} > c_0^{1/2} \mathcal{J}_\Psi \operatorname{diam}(B).$$

The right side is just $c_0 \mathcal{J}_\Psi^2 \geq \operatorname{dist}(\Psi(B), \Psi^N(B))$, so we are done. □

Let D_w denote the dilation $D_w(t_1, \dots, t_d) = (w_1 t_1, \dots, w_d t_d)$. We will apply Lemma 6.5 with $\Psi = \Psi_{x_0}^{J_0} \circ D_w$, $\Psi^N = \Psi_{x_0}^N \circ D_w$ and $G = D_w F_d$. By Lemma 6.4, we just need to verify (6-19).

Since $w_j \leq 1$ for each j, $\|\Psi\|_{C^2(Q_1)} \leq \|\Psi_{x_0}^{J_0}\|_{C^2(Q_w)} \lesssim 1$. For the error bound,

$$\|\Psi_{x_0}^{J_0} - \Psi_{x_0}^N\|_{C^2(Q_w)} \lesssim \max_i w_i^{N-1} \|\Psi_{x_0}^{J_0}\|_{C^{N+1}(Q_w)} \lesssim (c\alpha_1^\delta)^N, \tag{6-23}$$

where c is as in (6-12). (Recall that implicit constants do not depend on c.) We choose N larger than $\delta^{-1}(10 \deg_{J_0} \beta_0 + 10d)$ and then choose c sufficiently small. Combining (6-23), (6-12) and (6-18),

$$\|\Psi_{x_0}^{J_0} - \Psi_{x_0}^N\|_{C^2(Q_w)} \leq c_0 \Bigl(\prod_j w_j\Bigr)^2 \|\det D\Psi_{x_0}^N\|_{C^0(Q_w)}^2.$$

For c_0 sufficiently small, this implies that

$$\|\det D\Psi_{x_0}^{J_0} - \det D\Psi_{x_0}^N\|_{C^0(Q_w)} < \tfrac{1}{2} \|\det D\Psi_{x_0}^N\|_{C^0(Q_w)},$$

so $\|\det D\Psi_{x_0}^{J_0}\|_{C^0(Q_w)} \geq \frac{1}{2}\|\det D\Psi_{x_0}^{N}\|_{C^0(Q_w)}$. Rescaling gives us (6-19).

Applying Lemma 6.5, inequality (6-18), and $b_0 = \deg J_0 + \deg_{J_0} \beta_0$,

$$|\Omega| \geq |\Psi_{x_0}^{J_0}(F_d)| \gtrsim |F_d|\rho_0^{d-1}\alpha_1^{C\varepsilon}\alpha^{\deg_{J_0}\beta_0} \gtrsim \rho_0^{-1}\alpha_1^{2C\varepsilon}\alpha^{b_0}.$$

The proof of Theorem 2.1 is finally complete.

Appendix: proof of Proposition 4.1

In this section we prove Proposition 4.1, which was used in proving Propositions 2.2 and 2.3. We fix, for the remainder of this section, a point $b_0 \in [0,\infty)^k$. An object is admissible if it may be chosen from a finite collection, depending only on b_0, of such objects, and all implicit constants will be admissible (i.e., depending only on b_0).

The following two lemmas show that conclusions (i) and (ii) of Proposition 4.1 are equivalent.

Lemma A.1. *If $\mathscr{A} \subseteq \mathbb{Z}_0^k$ is a finite set and $b_0 \notin \mathscr{P}(\mathscr{A})$, there exist $\varepsilon > 0$ and $v_0 \in (\varepsilon, 1]^k$ such that $v_0 \cdot b_0 + \varepsilon < v_0 \cdot p$ for every $p \in \mathscr{P}(\mathscr{A})$.*

Lemma A.2. *If $v_0 \in (0,1]^k$, there exists a finite set $\mathscr{A} \subseteq \mathbb{Z}_0^k$ such that $b_0 \notin \mathscr{P}(\mathscr{A})$ and*

$$\{b \in \mathbb{Z}_0^k : v_0 \cdot b_0 < v_0 \cdot b\} \subseteq \mathscr{P}(\mathscr{A}).$$

Proof of Lemma A.1. We may assume that $b_0 \neq (0,\dots,0)$ and $\mathscr{A} \neq \varnothing$; otherwise, the result is trivial. Since $b_0 \notin \mathscr{P}(\mathscr{A})$, there exists $v_1 \in \mathbb{R}^k$ such that $v_1 \cdot b_0 < v_1 \cdot p$ for every $p \in \mathscr{P}(\mathscr{A})$. Since $\mathscr{P}(\mathscr{A})$ contains a translate of $[0,\infty)^k$, $v_1 \in [0,\infty)^k$. We may assume that $v_1 \in [0,1]^k$. Let

$$\delta := \tfrac{1}{2}|b_0|_1^{-1} \min_{b \in \mathscr{A}} v_1 \cdot (b - b_0).$$

Since $\mathscr{A}$ is finite, $\delta > 0$. Let $v_2 := v_1 + (\delta,\dots,\delta)$. Then $v_2 \in [\delta, 1+\delta]^k$. If $b \in \mathscr{A}$,

$$b \cdot v_2 = v_1 \cdot b_0 + v_1 \cdot (b - b_0) + \delta|b|_1 \geq v_2 \cdot b_0 + \delta|b_0|_1 \geq v_2 \cdot b_0 + \delta.$$

The conclusion thus holds with $\varepsilon := \frac{1}{2}\delta/(1+\delta)$, $v_0 := v_2/(1+\delta)$. □

Proof of Lemma A.2. Let $\varepsilon := \min_i v_0^i$ and let $N := \lceil k\varepsilon^{-1}(b_0 \cdot v_0 + 1)\rceil$. If $p \in \mathbb{Z}_0^k$ and $|p|_1 \geq N$,

$$v_0 \cdot p \geq \min_j v_0^j \max_i p^i \geq \varepsilon\left(\frac{N}{k}\right) \geq b_0 \cdot v_0 + 1,$$

so the conclusion holds with

$$\mathscr{A} := \{b \in \mathbb{Z}_0^k : |b|_1 \leq N \text{ and } v_0 \cdot b > v_0 \cdot b_0\}. \qquad \square$$

The following lemma implies that the conclusions of Proposition 4.1 hold whenever $\mathscr{B}$ is a finite set with $\#\mathscr{B} \leq k+1$.

Lemma A.3. *Let $\mathscr{B} \subseteq \mathbb{Z}_0^k$ be a finite set. Assume that $\#\mathscr{B} \leq k+1$ and that $b_0 \notin \mathscr{P}(\mathscr{B})$. Then there exist admissible $\varepsilon > 0$ and $v_0 \in (\varepsilon, 1]^k$ such that $b \cdot v_0 > b_0 \cdot v_0 + \varepsilon$ for every $p \in \mathscr{P}(\mathscr{B})$.*

The same proof shows that, for any finite $\mathcal{B}$ with $b_0 \notin \mathcal{P}(\mathcal{B})$, there exist $\varepsilon > 0$ and $v_0 \in (\varepsilon, 1]^k$, taken from a finite list that depends only on b_0 and m, such that $b \cdot v_0 > b_0 \cdot v_0 + \varepsilon$ for every $p \in \mathcal{P}(\mathcal{B})$, but for simplicity we only prove the version that we use.

Proof. The conclusion is trivial if $\mathcal{B} = \varnothing$, so we write $\mathcal{B} = \{b_1, \dots, b_m\}$ with $m \le k+1$. By Lemma A.1, the conclusion is trivial if $\{b_1, \dots, b_m\}$ is admissible; we will reduce to this case.

If $|b_i|_1 > |b_0|_1$, $1 \le i \le m$, the conclusion holds with $v_0 = (1, \dots, 1)$, $\varepsilon = \frac{1}{2}(\lceil |b_0|_1 + 1 \rceil - 1)$. Reindexing if necessary, we may assume that $|b_1|_1 \le |b_0|_1$, in which case $\{b_1\}$ is admissible.

Assume that for some $j < m$, $\{b_1, \dots, b_j\}$ is admissible. By assumption, $b_0 \notin \mathcal{P}(\{b_1, \dots, b_j\})$, so by Lemma A.1 there exist admissible $\varepsilon_j > 0$, $v_j \in (\varepsilon_j, 1]^k$ such that $v_j \cdot b_0 + \varepsilon_j < v_j \cdot b_i$ for $1 \le i \le j$. If $v_j \cdot b_0 + \varepsilon_j < v_j \cdot b_i$ for every i, the conclusion of the lemma holds with $\varepsilon = \varepsilon_j$, $v_0 = v_j$. Otherwise, after reindexing, we may assume that $v_j \cdot b_{j+1} \le v_j \cdot b_0$. Therefore b_{j+1} is admissible, and hence $\{b_1, \dots, b_{j+1}\}$ is admissible as well. The procedure must terminate after at most $m \le k+1$ steps, and so the lemma is proved. □

Lemma A.3 has the following corollary.

Lemma A.4. *Under the hypotheses of Lemma A.3, there exists an admissible $\varepsilon > 0$ such that if*

$$b(\theta) := \sum_{i=1}^{m} \theta_i b_i$$

is any convex combination of $b_1, \dots, b_m$, there exists an i, $1 \le i \le k$ such that $b^i(\theta) \ge b_0^i + \varepsilon$.

Proof. By Lemma A.3, there exist admissible $\varepsilon > 0$, $v_0 \in (\varepsilon, 1]^k$ such that

$$\varepsilon < (b(\theta) - b_0) \cdot v_0 \le \left(\sum_{i=1}^{k} v_0^i \right) \max_{1 \le i \le k} (b^i(\theta) - b_0^i) \le \max_{1 \le i \le k} (b^i(\theta) - b_0^i). \qquad \square$$

Finally, we are ready to complete the proof of Proposition 4.1.

Proof of Proposition 4.1. Let $C > |b_0|_1$ be a large constant, to be determined (admissibly) in a moment. Define $\mathcal{A} := \mathcal{B}' \cup \mathcal{B}''$, where

$$\mathcal{B}' := \{b \in \mathcal{B} : |b|_1 \le C\},$$
$$\mathcal{B}'' := \{Ce_i : 1 \le i \le k\}.$$

Here e_i denotes the i-th standard basis vector. Then, since $\mathcal{P}(\mathcal{B}'') = \mathcal{P}(\{b \in \mathbb{Z}_0^k : |b|_1 \ge C\})$, $\mathcal{P}(\mathcal{B}) \subseteq \mathcal{P}(\mathcal{A})$. It remains to show that, for C sufficiently large, $b_0 \notin \mathcal{P}(\mathcal{A})$.

Assume that $b_0 \in \mathcal{P}(\mathcal{A})$. By Carathéodory's theorem from combinatorics (see, for instance, [Ziegler 1995, p. 46]), $b_0 \succeq \sum_{l=1}^{k+1} \theta_l a_l$, for some $a_1, \dots, a_{k+1} \in \mathcal{A}$ and $0 \le \theta_l \le 1$ satisfying $\sum_l \theta_l = 1$. Reindexing if necessary,

$$b_0 \succeq \sum_{l=1}^{j} \theta_l C e_{i_l} + \sum_{l=j+1}^{k+1} \theta_l b_l, \tag{A-1}$$

where $b_{j+1}, \dots, b_{k+1} \in \mathcal{B}'$. Since $C > |b_0|_1$, $\sum_{l=j+1}^{k+1} \theta_l > 0$ and, since $b_0 \notin \mathcal{P}(\mathcal{B}') \subseteq \mathcal{P}(\mathcal{B})$, $\sum_{l=1}^{j} \theta_l > 0$.

Let

$$b(\theta) := \Big(\sum_{l=j+1}^{k+1} \theta_l\Big)^{-1} \sum_{l=j+1}^{k+1} \theta_l b_l.$$

By Lemma A.4, there exists an i, $1 \le i \le k+1$ such that $b^i(\theta) \ge b_0^i + \varepsilon$, where $\varepsilon > 0$ depends only on b_0 (crucially, not on C). By (A-1),

$$b_0 \succeq \Big(\sum_{l=j+1}^{k+1} \theta_j\Big) b(\theta),$$

so, comparing the i-th coordinates, we see that

$$\sum_{l=j+1}^{k+1} \theta_j \le \frac{b_0^i}{b_0^i + \varepsilon} \le \frac{|b_0|_\infty}{|b_0|_\infty + \varepsilon},$$

so

$$\sum_{l=1}^{j} \theta_j \ge 1 - \frac{|b_0|_\infty}{|b_0|_\infty + \varepsilon} = \frac{\varepsilon}{|b_0|_\infty + \varepsilon}. \tag{A-2}$$

On the other hand, by (A-1) and the fact that all coordinates of the b_i are nonnegative, $\sum_{l=1}^{j} \theta_j \le |b_0|_1/C$. For $C = C(\varepsilon, b_0)$ sufficiently large (admissible since ε is), this contradicts (A-2), and the proof of Proposition 4.1 is complete. □

Acknowledgements

The idea for this project came from a conversation with Michael Christ, and Terence Tao provided many valuable suggestions during the early stages of this work. The author is very grateful for these discussions. She would also like to thank the anonymous referee for a detailed and enormously helpful report. This project was supported in part by NSF DMS-0902667 and 1266336.

References

[Bak et al. 2009] J.-G. Bak, D. M. Oberlin, and A. Seeger, "Restriction of Fourier transforms to curves and related oscillatory integrals", *Amer. J. Math.* **131**:2 (2009), 277–311. MR 2010b:42012 Zbl 1166.42006

[Bennett et al. 2005] J. Bennett, A. Carbery, and J. Wright, "A non-linear generalisation of the Loomis–Whitney inequality and applications", *Math. Res. Lett.* **12**:4 (2005), 443–457. MR 2006m:26039 Zbl 1106.26020

[Christ 1998] M. Christ, "Convolution, curvature, and combinatorics: A case study", *Int. Math. Res. Not.* **1998**:19 (1998), 1033–1048. MR 2000a:42026 Zbl 0927.42008

[Christ 2008] M. Christ, "Lebesgue space bounds for one-dimensional generalized Radon transforms", preprint, 2008, Available at http://math.berkeley.edu/~mchrist/Papers/radon.pdf.

[Christ et al. 1999] M. Christ, A. Nagel, E. M. Stein, and S. Wainger, "Singular and maximal Radon transforms: Analysis and geometry", *Ann. of Math.* (2) **150**:2 (1999), 489–577. MR 2000j:42023 Zbl 0960.44001

[Dendrinos and Müller 2013] S. Dendrinos and D. Müller, "Uniform estimates for the local restriction of the Fourier transform to curves", *Trans. Amer. Math. Soc.* **365**:7 (2013), 3477–3492. MR 3042592 Zbl 1277.42009

[Dendrinos and Stovall 2012] S. Dendrinos and B. Stovall, "Uniform estimates for the X-ray transform restricted to polynomial curves", *J. Funct. Anal.* **262**:12 (2012), 4986–5020. MR 2916059 Zbl 1253.44002

[Dendrinos and Stovall 2014] S. Dendrinos and B. Stovall, "Uniform bounds for convolution and restricted X-ray transforms along degenerate curves", preprint, 2014, Available at http://www.math.wisc.edu/∼stovall/research/DSdegenerate.pdf.

[Dendrinos and Wright 2010] S. Dendrinos and J. Wright, "Fourier restriction to polynomial curves I: A geometric inequality", *Amer. J. Math.* **132**:4 (2010), 1031–1076. MR 2011e:42035 Zbl 1203.42015

[Dendrinos et al. 2009] S. Dendrinos, N. Laghi, and J. Wright, "Universal L^p improving for averages along polynomial curves in low dimensions", *J. Funct. Anal.* **257**:5 (2009), 1355–1378. MR 2010f:42040 Zbl 1177.42009

[Drury and Marshall 1985] S. W. Drury and B. P. Marshall, "Fourier restriction theorems for curves with affine and Euclidean arclengths", *Math. Proc. Cambridge Philos. Soc.* **97**:1 (1985), 111–125. MR 87b:42019 Zbl 0567.42009

[Drury and Marshall 1987] S. W. Drury and B. P. Marshall, "Fourier restriction theorems for degenerate curves", *Math. Proc. Cambridge Philos. Soc.* **101**:3 (1987), 541–553. MR 88e:42026 Zbl 0645.42015

[Gressman 2009] P. T. Gressman, "L^p-improving properties of averages on polynomial curves and related integral estimates", *Math. Res. Lett.* **16**:6 (2009), 971–989. MR 2011d:42042 Zbl 1204.42022

[Hörmander 1967] L. Hörmander, "Hypoelliptic second order differential equations", *Acta Math.* **119** (1967), 147–171. MR 36 #5526 Zbl 0156.10701

[Oberlin 2002] D. M. Oberlin, "Convolution with measures on polynomial curves", *Math. Scand.* **90**:1 (2002), 126–138. MR 2002m:28004 Zbl 1034.42012

[Oberlin 2003] D. M. Oberlin, "Affine dimension: Measuring the vestiges of curvature", *Michigan Math. J.* **51**:1 (2003), 13–26. MR 2004a:28013 Zbl 1035.53023

[Oberlin 2010] D. M. Oberlin, "Convolution with measures on flat curves in low dimensions", *J. Funct. Anal.* **259**:7 (2010), 1799–1815. MR 2011i:44007 Zbl 1203.42024

[Sjölin 1974] P. Sjölin, "Fourier multipliers and estimates of the Fourier transform of measures carried by smooth curves in R^2", *Studia Math.* **51** (1974), 169–182. MR 52 #6299 Zbl 0291.42003

[Stein 1993] E. M. Stein, *Harmonic analysis: Real-variable methods, orthogonality, and oscillatory integrals*, Princeton Mathematical Series **43**, Princeton University Press, 1993. MR 95c:42002 Zbl 0821.42001

[Stovall 2010] B. Stovall, "Endpoint $L^p \to L^q$ bounds for integration along certain polynomial curves", *J. Funct. Anal.* **259**:12 (2010), 3205–3229. MR 2011g:42049 Zbl 1209.44003

[Stovall 2011] B. Stovall, "L^p improving multilinear Radon-like transforms", *Rev. Mat. Iberoam.* **27**:3 (2011), 1059–1085. MR 2895344 Zbl 1230.42016

[Stovall 2014] B. Stovall, "Uniform estimates for Fourier restriction to polynomial curves in $\mathbb{R}^d$", preprint, 2014, Available at http://www.math.wisc.edu/∼stovall/research/curve_rest_apr_29.pdf. To appear in *Amer. J. Math.*

[Street 2011] B. Street, "Multi-parameter Carnot–Carathéodory balls and the theorem of Frobenius", *Rev. Mat. Iberoam.* **27**:2 (2011), 645–732. MR 2012m:53064 Zbl 1222.53036

[Tao and Wright 2003] T. Tao and J. Wright, "L^p improving bounds for averages along curves", *J. Amer. Math. Soc.* **16**:3 (2003), 605–638. MR 2004j:42005 Zbl 1080.42007

[Ziegler 1995] G. M. Ziegler, *Lectures on polytopes*, Grad. Texts in Math. **152**, Springer, New York, 1995. MR 96a:52011 Zbl 0823.52002

Received 27 Nov 2013. Accepted 9 Jun 2014.

BETSY STOVALL: stovall@math.wisc.edu
Department of Mathematics, University of Wisconsin, 480 Lincoln Dr., Madison, WI 53706-1325, United States

Analysis & PDE

msp.org/apde

See inside back cover or msp.org/apde for submission instructions.

The subscription price for 2014 is US $180/year for the electronic version, and $355/year (+$50, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues from the last three years and changes of subscribers address should be sent to MSP.

Analysis & PDE (ISSN 1948-206X electronic, 2157-5045 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840, is published continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

APDE peer review and production are managed by EditFLOW® from Mathematical Sciences Publishers.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

ANALYSIS & PDE

Volume 7 No. 5 2014