UNIFORM $L^p$-Smothing for Weighted Averages on Curves

Betsy Stovall

Abstract. 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

In this article we consider weighted versions of multilinear generalized Radon transforms of the form

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

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

In [20, 18], near endpoint estimates of the form

$$|M_0(f_1, \ldots, f_k)| \lesssim \prod_{i=1}^k \|f_i\|_{L^{p_i}}$$

were established for $M_0$ under the assumption that the $\pi_i$ satisfy a certain finite type condition on the support of $a$, and the bounds established therein depend on this ‘type.’ These results are nearly sharp in the sense that if the type of the $\pi_i$ 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 on the boundary of the support of $a$ 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 $\pi_1, \ldots, \pi_k$ and the cutoff function $a$. Further, the weights we employ transform naturally under various transformations of the $\pi_i$, and may thus reasonably be viewed as generalizations of the affine arclength measure on curves $\mathbb{R}^d$. A number of recent articles have been devoted to establishing uniform estimates for operators weighted by affine arclength measure, and these results provide much of the motivation for this article.

1.1. A motivating example. Stating the main results of this article, or even the results of [20, 18] 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. Consider the operator
\[ T_0 f(x) = \int_{\mathbb{R}} f(x - \gamma(t)) a(t) \, dt, \]
where \( \gamma : \mathbb{R} \to \mathbb{R}^d \) is a smooth curve and \( a \) is a continuous cutoff function. By duality, \( T_0 : L^p \to L^q \) if and only if
\[ \left| \int_{\mathbb{R}^d} \int_{\mathbb{R}} f(x - \gamma(t)) g(x) a(t) \, dt \right| \lesssim \|f\|_{L^p} \|g\|_{L^q'}, \]
which is an estimate of the form (1.2).

The curve \( \gamma \) is said to be of type (at most) \( N \) when the torsion, \( \det(\gamma'(t), \ldots, \gamma^{(d)}(t)) \), vanishes to order at most \( N \) at any point. The results of [20] imply that if \( \gamma \) 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 interior of the trapezoid with vertices
\[ (0, 0), \quad (1, 1), \quad (p_N^{-1}, q_N^{-1}) := \left( \frac{1 + N + \frac{d(d-1)}{2}}{N + \frac{d(d+1)}{2}}, \frac{1}{N + d + 1} \right), \quad (1 - q_N^{-1}, 1 - p_N^{-1}). \] (1.3)

Further, if \( N \) is the minimal type of \( T_0 \) on \( \{ t : a(t) \neq 0 \} \), \( T_0 \) does not map \( L^p \) boundedly into \( L^q \) if \( (p^{-1}, q^{-1}) \) does not lie in the (closure of) this trapezoid. If \( \gamma \) is not of finite type, \( T_0 \) satisfies no \( L^p \to L^q \) estimates off the line \( \{ p = q \} \).

It is now known that if \( \gamma \) is a polynomial curve, convolution with affine arclength measure on \( \gamma \), which is the operator
\[ Tf(x) = \int_{\mathbb{R}} f(x - \gamma(t)) |\det(\gamma'(t), \ldots, \gamma^{(d)}(t))|^{\frac{2}{p+d+1}} \, dt, \]
maps \( L^p \) boundedly into \( L^q \) if and only if \( (p^{-1}, q^{-1}) \) lies on the line segment joining \( (p_0^{-1}, q_0^{-1}), (1 - q_0^{-1}, 1 - p_0^{-1}) \) above, \([12, 6, 17]\). Further, the operator norms established in \([12, 6, 17]\) 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 by Oberlin in the non-polynomial case in \([13]\). The above mentioned results are essentially optimal, both in terms of the exponents involved and in terms of pointwise estimates on the weight, \([14]\) (cf. Proposition 2.2). Analogous results are also known for the restricted X-ray transform, \([8]\). 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 \([1, 7, 9]\).

Our goal in this article is to address the gap between the general results of \([20, 18]\) and the type-independent results of \([12, 6, 17, 8]\) 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 \([20, 18]\) to the case when the \( \pi_i \) 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. To state the main results will require some notation. We denote the nonnegative integers by \( \mathbb{Z}_0 \). We consider the partial order \( \preceq \) on \( \mathbb{Z}_0^k \) defined by
$b_1 \leq b_2$ if $b_1^i \leq b_2^i$, $1 \leq i \leq k$. We say $b_1 < b_2$ if at least one of these inequalities is strict. If $B \subset \mathbb{Z}_0^k$, is any set, we define a polytope

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

where ‘ch’ denotes the convex hull.

We will consider vector fields $X_1, \ldots, X_k$, defined and smooth on the closure of an open set $U$. A word $w$ is an element of $W = \bigcup_{n=1}^{\infty} \{1, \ldots, k\}^n$. To each word is associated a vector field $X_w$, defined recursively by $X_{(i)} = X_i$, $1 \leq i \leq k$ and $X_{(w,i)} = [X_w, X_i]$, for $w \in W$ and $1 \leq i \leq k$. The degree of $w \in W$ 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 $w, w' \in W$,

$$[X_w, X_{w'}] = \sum_{\deg \hat{w} = \deg w + \deg w'} C_{w, w', \hat{w}} X_{\hat{w}}, \quad (2.1)$$

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

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

and so (2.1) is easily obtained by inducting on $\|\deg w'\|_i$ (this observation was previously made in e.g. [11]). 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 $I = (w_1, \ldots, w_d)$ is a $d$-tuple of words, we define $\deg(I) = \sum_{i=1}^d \deg w_i$ and

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

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

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

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

$$P_A = \text{ch}(\bigcup_{x \in A} P_x).$$

**Results.** Let $U \subset \mathbb{R}^d$ be an open set and let $\pi_1, \ldots, \pi_k : \overline{U} \to \mathbb{R}^{d-1}$ be smooth submersions (i.e. having surjective differentials). Define vector fields

$$X_j = *(d\pi_j^1 \wedge \cdots \wedge d\pi_j^{d-1}), \quad 1 \leq j \leq k, \quad (2.2)$$

where $*$ denotes the Hodge star operator, which maps $(d-1)$-forms to vector fields. Let $a$ be a continuous function with compact support contained in $U$.

Let $b_0$ be an extreme point of $P_{\text{supp } a}$ and choose a $d$-tuple of words $I_0 = (w_1, \ldots, w_d)$ with $\deg I_0 = b_0$. Define the generalized affine arclength

$$\rho = \rho_{I_0} = |\det(X_{w_1}, \ldots, X_{w_d})|^{\frac{1}{|\pi_j|}}, \quad (2.3)$$

where $|b|_p$ denotes the $\ell_p$ norm, $1 \leq p \leq \infty$, and define a $k$-linear form $M : [C^0(\mathbb{R}^d)]^k \to \mathbb{C}$ by

$$M(f_1, \ldots, f_k) = \int_{\mathbb{R}^d} \prod_{j=1}^k f_j \circ \pi_j(x) \rho(x) a(x) \,dx. \quad (2.4)$$

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

$$q(b) = \left(\frac{b^1}{|b|_1 - 1}, \ldots, \frac{b^k}{|b|_1 - 1}\right). \quad (2.5)$$
It is easy to check that $q$ equals its own inverse.

The following is our main theorem.

**Theorem 2.1.** Let $(p_1, \ldots, p_k) \in [1, \infty]^k$ and assume that $p_j^{-1} \leq q_j(b)$, $1 \leq j \leq k$, with strict inequality when $b_0 \neq 0$. Then

$$|M(f_1, \ldots, f_k)| \lesssim \prod_{j=1}^k \|f_j\|_{L^{p_j}(\mathbb{R}^{d-1})}, \quad (2.6)$$

for all continuous $f_1, \ldots, f_k$, where the implicit constant depends on the $\pi_j$, $\alpha$, $(p_1^{-1}, \ldots, p_k^{-1})$ and $b_0$, but not on the $f_j$. Thus $M$ extends to a bounded $k$-linear form on $\prod_{j=1}^k L^{p_j}(\mathbb{R}^{d-1})$.

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

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

$$M_\mu(\chi_{E_1}, \ldots, \chi_{E_k}) := \int_{\mathbb{R}^d} \prod_{j=1}^k \chi_{E_j} \circ \pi_j \, d\mu \leq C_\mu \prod_{j=1}^k \|E_j\|_{\mathcal{P}^{\frac{1}{p_j}}} \quad (2.7)$$

holds for all Borel sets $E_1, \ldots, E_k \subset \mathbb{R}^{d-1}$. If $\mu \neq 0$, $(p_1, \ldots, p_k) \in [1, \infty]$. If $\sum_j p_j^{-1} > 1$, let $b_p = q(p_1^{-1}, \ldots, p_k^{-1})$. Then $\mu(\{x : b_p \notin \mathcal{P}_x\}) = 0$. If in addition, $b_p$ is an extreme point of $\mathcal{P}_{\text{supp } \mu}$, $\mu$ is absolutely continuous with respect to Lebesgue measure, and its Radon–Nikodym derivative satisfies

$$\frac{du}{dx} \leq C_{d,p} C_\mu \sum_{\deg J = b_p} |\lambda_J|^{\frac{1}{\sum_j p_j - 1}} \quad (2.8)$$

In the translation invariant case, a similar result is due to D. Oberlin in [14] (cf. [8] 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 made here, as can be seen by considering the example of convolution with affine arclength on $\gamma(t) = (t, e^{-1/t} \sin(\frac{\pi}{t}))$, $t > 0$, for $k$ sufficiently large. (This example is due to Sjölin in [15].)

The proofs of Theorem 2.1 and Proposition 2.2 will rely on a more general result about smooth vector fields $X_1, \ldots, X_k$ on $\mathbb{R}^d$ (with no further assumptions). To state this result, we need some additional terminology.

Let $J \in \{1, \ldots, k\}^d$. 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 multi-index, we define $\deg J(\alpha)$ to be the $k$-tuple whose $i$-th entry is $\sum_{J_i = i} \alpha_i$. We define

$$\Psi^J_{x_0}(t_1, \ldots, t_d) = \exp(t_1 X_{J_1}) \circ \cdots \circ \exp(t_1 X_{J_1})(x_0). \quad (2.9)$$

We define another polytope,

$$\tilde{\mathcal{P}}_{x_0} = \mathcal{P}(\{\deg J + \deg J(\alpha) : J \in \{1, \ldots, k\}^d \text{ and } \alpha \in (\mathbb{Z}_0)^d \text{ satisfy } \partial \det D\Psi^J_{x_0}(0) \neq 0\}).$$
Proposition 2.3. For each \( x_0 \in U \), \( \overline{P}_{x_0} = P_{x_0} \). Furthermore, for each extreme point \( b \) of \( P_{x_0} \),
\[
\sum_{\deg t = b} |\lambda_I(x_0)| \sim \sum_{J \in \{1, \ldots, k\}^d} \sum_{\alpha \in (\mathbb{Z}_0)^d} |\partial^\alpha_I \det D\Psi^J_{x_0}(0)|. \quad (2.10)
\]
The implicit constants may be taken to depend only on \( d \) and \( b \), 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 \( \gamma \), 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 \( \gamma \) does not vanish to infinite order at any point. We note in particular that
\[
|\det(X_1, X_2, X_{(1,2)}, \ldots, X_{(1,\ldots, 2)})| = |\det(X_1, X_2, X_{(2,1)}, \ldots, X_{(2,\ldots, 2)})| = |\det(\gamma', \ldots, \gamma(d))|,
\]
and if \( U \) is any open set, the only extreme points of \( \mathcal{P}_U \) (unless \( \mathcal{P}_U \) is empty) are
\[
(\frac{d(d-1)}{2} + 1, d), \quad (d, \frac{d(d-1)}{2} + 1).
\]
Thus the affine arclength in this case is defined in the usual way:
\[
\rho(t, x) = |\det(\gamma'(t), \ldots, \gamma(d)(t))|^{\frac{2}{d+1}}.
\]
By Theorem 2.1, for any smooth \( \gamma : \mathbb{R} \to \mathbb{R}^d \), and any continuous compactly supported \( a \), the convolution operator
\[
Tf(x) = \int f(x - \gamma(t)) |\det(\gamma'(t), \ldots, \gamma(d))(t)|^{\frac{2}{d+1}} dt
\]
maps \( L^p \) into \( L^q \) whenever \((p^{-1}, q^{-1})\) lies in the interior of the trapezoid with vertices as in (1.3) in the case \( N = 0 \). Without further assumptions on \( \gamma \), this result seems to be new, but as mentioned in the introduction, is already known in some special cases.

Restricted X-ray transforms. Let \( \gamma : \mathbb{R} \to \mathbb{R}^{d-1} \) be a smooth map and for \((s, t, x) \in \mathbb{R}^{1+d-1} \), 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 \((w_1, \ldots, w_{d+1})\) with \( \det(X_{w_1}, \ldots, X_{w_{d+1}}) \neq 0 \) are those satisfying
\[
w_1 = 1, \quad w_2 = 2, \quad w_i = (1, 2, \cdots, 2), \quad 3 \leq i \leq d + 1
\]
after reordering. Thus the only extreme point of the Newton polytope is \((d, 1 + \frac{d(d-1)}{2})\), and

\[
\rho(s, t, x) = |\det(\gamma'(t), \ldots, \gamma^{(d-1)}(t))|^\frac{1}{d+1},
\]

which is a power of the usual affine arclength. Theorem 2.1 thus gives a partial generalization of the results of [8], wherein a sharp strong type bound for the X-ray transform restricted to polynomial curves with affine arclength was established.

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

\[
\left| \int_{\mathbb{R}^d} \prod_{i=1}^d f_i \circ \pi_i(x) \, \det(X_1, \ldots, X_d(x))^{\frac{1}{d(2d-1)}} \, a(x) \, dx \right| \lesssim \prod_{i=1}^d \|f_i\|_{L^{d-i+\varepsilon}(\mathbb{R}^{d-1})},
\]

for all \(\varepsilon > 0\). In the case when the \(X_i\) do span at every point of the support of \(a\), the endpoint estimate was proved in [2]. (The classical Loomis–Whitney inequality is the endpoint estimate when the \(\pi_i\) 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 [19] and using 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, Proposition 6.1, which implies Theorem 2.1. Our techniques for the proof of the main theorem are essentially those of [20, 18, 3], with some modifications to handle the potential failure of the Hörmander condition.

**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. This project was supported in part by NSF DMS-0902667.

### 3. Invariants of the affine arclengths

Let \(U, \pi_1, \ldots, \pi_k,\) and \(X_1, \ldots, X_k\) be as defined above. For \(1 \leq j \leq k\), 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 \(\deg I' \prec \deg I_0, \lambda_I \equiv 0\). (This minimality is essential.) Define \(\rho\) as in (2.3).

**Proposition 3.1.** Let \(F : U \to \mathbb{R}^d\) and \(G_j : V_j \to \mathbb{R}^{d-1}, 1 \leq j \leq k,\) be smooth maps. Define \(\tilde{\pi}_j = G_j \circ \pi_j \circ F, 1 \leq j \leq k,\) and let \(\tilde{X}_j, \tilde{\rho}\) 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 \(q\) is defined as in (2.5).
In the notation above, let $a$ be a continuous, compactly supported function with $\text{supp } a \subset U$, and define
\[
\tilde{M}(f_1, \ldots, f_k) = \int_U \prod_{j=1}^k f_j \circ \tilde{\pi}_j(x) \tilde{\rho}(x) a \circ F(x) \, dx.
\]
Proposition 3.1 implies that if each $G_j$ is equal to the identity and $F$ is one-to-one, then
\[
\tilde{M}(f_1, \ldots, f_k) = M(f_1, \ldots, f_k).
\]
If we simply assume that $F$ and all of the $G_j$’s are one-to-one, the proposition implies that for $(p_1^{-1}, \ldots, p_k^{-1}) = (b_0)$,
\[
\sup_{f_1, \ldots, f_k \neq 0} \frac{\tilde{M}(f_1, \ldots, f_k)}{\prod_{j=1}^k \|f_j\|_{L^p_j}} = \sup_{f_1, \ldots, f_k \neq 0} \frac{M(f_1, \ldots, f_k)}{\prod_{j=1}^k \|f_j\|_{L^p_j}}.
\]
We stress, however, that our theorem covers only the non-endpoint cases satisfying $(p_1^{-1}, \ldots, p_k^{-1}) \neq (b_0)$ and $b_0$ extreme, so it is not known that either side is finite except in certain cases (cf. [2, 6, 8, 12, 17]).

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

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

**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
\[
\tilde{X}_j = (\det DF)F^*X_j,
\]
where $F^*$ is the pullback by $F$, given by
\[
F^*X = (DF)^{-1}X \circ F. \tag{3.2}
\]
For $1 \leq i \leq k$, let $Y_i = F^*X_i$. Then by naturality of the Lie bracket, $Y_w = F^*X_w$, $w \in \mathcal{W}$. Thus for any $1 \leq i, j \leq k$,
\[
\tilde{X}_{(i,j)} = (\det DF)^2Y_{(i,j)} + (\det DF)\{[(DF)^{-1}X_i \circ F](\det DF)\}Y_j
\]
\[
- (\det DF)\{[(DF)^{-1}X_j \circ F](\det DF)\}Y_i.
\]
By induction, for each $w \in \mathcal{W}$,
\[
\tilde{X}_w = (\det DF)^{|\deg w|1}Y_w + \sum_{\deg w' < \deg w} f_{w,w'}Y_{w'}, \tag{3.3}
\]
where the $f_{w,w'}$ are smooth functions.

By (3.3), (3.2), and our minimality assumption,
\[
\det(\tilde{X}_{w_1}, \ldots, \tilde{X}_{w_d})
\]
\[
= (\det DF)^{|b_0|1} \det(Y_{w_1}, \ldots, Y_{w_d}) + \sum_{b' < b_0} \sum_{\deg F' = \deg b'} f_{1,F'} \det(Y_{w_1}', \ldots, Y_{w_d}')
\]
\[
= (\det DF)^{|b_0|1} \det(X_{w_1}, \ldots, X_{w_d}) \circ F + 0.
\]
This completes the proof in the first case.
In the second case, when $F$ is the identity, it is easy to compute $\tilde{X}_j = [(\det DG_j) \circ \pi_j] X_j$, and it can be shown using the product rule and minimality of $b_0$ (as above) that

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

which implies (3.1).

\[\square\]

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

This section will largely be devoted to the proof of Proposition 2.3. Throughout this section, we fix a point $b_0 \in [0, \infty)^k$, which will correspond to either an extreme point or an element in the exterior of a polytope. We will say that an object (such as $\varepsilon > 0$, $v_0 \in (0, 1]^k$, a set $A \subset \mathbb{Z}_0^k$, etc.) is admissible (or may be chosen admissibly) if it may be chosen from a finite collection $C_{b_0,d}$ of such objects.

The proofs of Propositions 2.3 and 2.2 will rely on the following compactness result about polytopes with vertices in $\mathbb{Z}_0^k$.

**Proposition 4.1.** Let $B \subset \mathbb{Z}_0^k$ and assume that $b_0 \notin P(B)$. 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 $p \in P(B)$

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

Moreover, $\varepsilon, v_0, A$ may be chosen admissibly.

We will prove this proposition in the Appendix. Assuming that it holds for now, we turn to the proof of Proposition 2.3. For ease of notation, we assume that $x_0 = 0$ and fix a neighborhood $V$ of $0, V \subset U$. We note that it suffices to prove the theorem in the case when $X_i = \partial_i$, $1 \leq i \leq d$; in particular, this implies that $P_0 \neq \emptyset$. Indeed, if the theorem does hold under this assumption, then for any smooth $X_1, \ldots, X_k$, the theorem holds for $\partial_1, \ldots, \partial_d, X_1, \ldots, X_k$, with $k + d$ replacing $k$. We may then transfer the result back to $X_1, \ldots, X_k$ by restricting to those $b \in [0, \infty)^{k+d}$ with $b_1 = \ldots = b_d = 0$.

Assume that $b_0$ is an extreme point for $P_0$. Recalling that $W = \bigcup_{n=1}^\infty \{1, \ldots, k\}^n$ is the set of all words, fix a $d$-tuple

$$I_0 = (w_1, \ldots, w_d) \in W^d$$

with $\deg I_0 = b_0$ such that

$$|\lambda_{I_0}(0)| = \max_{\deg I = b_0} |\lambda_I(0)|.$$  \hfill (4.2)

This $d$-tuple $I_0$ is admissible.

By the definition of $P_0$, $\lambda_{I_0}(0) \neq 0$. Since the $X_i$ are smooth, the $X_w$ are continuous, so we may assume that $V$ is so small that

$$\frac{1}{2}|\lambda_{I_0}(0)| \leq \frac{1}{2}\deg I = b_0 |\lambda_I(x)| \leq |\lambda_I(x)| \leq 2|\lambda_{I_0}(0)|, \quad x \in V. \hfill (4.3)$$

By Proposition 4.1, there exist admissible parameters $v_0 \in (0, 1]^k$ and $\varepsilon > 0$ such that

$$v_0 \cdot b_0 + \varepsilon < v_0 \cdot p, \quad p \in P_0 \cap \mathbb{Z}_0^k \setminus \{b_0\}. \hfill (4.4)$$

Let

$$W_0 = \{w \in W : v_0 \cdot \deg w \leq d\}. \hfill (4.5)$$

Since $v_0 \in (0, 1]^k$ is admissible, $W_0$ is finite, admissible, and contains all of the single-letter words, $(1), \ldots, (k)$. Furthermore, since $(1, \ldots, 1, 0, \ldots, 0) \in P_0$ by
construction, \( v_0 \cdot b_0 + \varepsilon < d \), so \( \deg w \leq b_0 \) implies that \( w \in \mathcal{W}_0 \). We will prove the following.

**Lemma 4.2.** For all \( m \geq 0 \), there exists \( \delta_m > 0 \), depending on \( X_1, \ldots, X_k, m, b_0 \), such that for all \( 0 < \delta \leq \delta_m \), \( I \in \mathcal{W}_0^d \), and \( w, w' \in \mathcal{W}_0 \),

\[
|\delta_{v_0 \cdot \deg w} I \lambda_I(0)| < \delta^{v_0 \cdot \deg w} |\lambda_I(0)|, \quad I \in \mathcal{W}_1, \quad \deg I \neq b_0
\]

(4.6)

\[
\|\delta_{v_0 \cdot \deg w} X_w\|_{C^m(V)} \leq \frac{1}{d} \text{dist}(0, \partial V), \quad \|\delta_{v_0 \cdot \deg w} X_w\|_{C^m(V)} \leq 1,
\]

(4.7)

\[
\left[ \delta_{v_0 \cdot \deg w} X_w, \delta_{v_0 \cdot \deg w'} X_{w'} \right] = \sum_{\tilde{w} \in \mathcal{W}_0} c_{w, w', \tilde{w}} \delta_{v_0 \cdot \deg w} X_{\tilde{w}},
\]

(4.8)

with

\[
\|c_{w, w', \tilde{w}}\|_{C^m(V)} \lesssim 1,
\]

(4.9)

for an implicit constant depending only on \( b_0, d \).

**Proof.** The \( X_w \) are all smooth, \( \mathcal{W}_0 \) is a finite set, and \( v_0 \cdot \deg w > 0 \) for all \( w \in \mathcal{W} \), so the inequalities in (4.7) are all trivial for sufficiently small \( \delta \). Similarly, by (4.4), inequality (4.6) also holds for sufficiently small \( \delta \).

Let \( w, w' \in \mathcal{W}_0 \) and \( 0 < \delta \leq 1 \). By the Jacobi identity (cf. (2.1)),

\[
\left[ \delta_{v_0 \cdot \deg w} X_w, \delta_{v_0 \cdot \deg w'} X_{w'} \right] = \sum_{\deg \tilde{w} = \deg w + \deg w'} C_{w, w', \tilde{w}} \delta_{v_0 \cdot \deg \tilde{w}} X_{\tilde{w}},
\]

for constants \( C_{w, w', \tilde{w}} \), which are admissible since \( \mathcal{W}_0 \) is an admissible set. If \( v_0 \cdot (\deg w + \deg w') \leq d \), each \( \tilde{w} \) in the above sum lies in \( \mathcal{W}_0 \), and (4.8-4.9) are automatic. Otherwise, each \( \tilde{w} \) in the sum is an element of

\[
\mathcal{W}_1 = \{ w \in \mathcal{W} : d < v_0 \cdot \deg w \leq 2d \},
\]

which is an admissible finite subset of \( \mathcal{W} \). Recalling that \( X_i = \partial_i, 1 \leq i \leq d \),

\[
\delta_{v_0 \cdot \deg \tilde{w}} = \sum_{i=1}^{d} \delta_{v_0 \cdot \deg \tilde{w} - v_0^i} X_0^i (\delta_{v_0} X_i).
\]

For \( \tilde{w} \in \mathcal{W}_1, v_0 \cdot \deg \tilde{w} - v_0^i > 0 \), and (4.8-4.9) follow for sufficiently small \( \delta_m \) by smoothness of the \( X_i \) and finiteness of \( \mathcal{W}_1 \). This completes the proof of the lemma. \( \square \)

It is now just a matter of converting notation to apply Theorem 5.1 of [19]. Rather than restate that theorem in its full generality, we will summarize precisely the result we need.

**Lemma 4.3.** For each \( m \geq 0 \), there exists \( \delta_m > 0 \), satisfying the conclusions of Lemma 4.2, and depending on \( m, b_0, X_1, \ldots, X_k \), such that for all \( 0 < \delta \leq \delta_m \), the map

\[
\Phi^\delta(y_1, \ldots, y_d) = \exp(y_1 \delta_{v_0 \cdot w_1} X_{w_1} + \cdots + y_d \delta_{v_0 \cdot w_d} X_{w_d})(0)
\]

is defined and one-to-one on the unit ball \( B(1) \) and satisfies

\[
|\det D\Phi^\delta(y)| \sim \delta^{v_0 \cdot b_0} |\lambda_I(0)|, \quad y \in B(1).
\]

(4.10)

Furthermore, if we define \( Y_w^\delta \) to be the pullback of \( \delta_{v_0 \cdot \deg w} X_w \) by \( \Phi^\delta \):

\[
Y_w^\delta = (D\Phi^\delta)^{-1} \delta_{v_0 \cdot \deg w} X_w \circ \Phi^\delta,
\]

the \( Y_w^\delta \) are smooth on \( B(1) \) and satisfy

\[
\|Y_w^\delta\|_{C^m(B(1))} \lesssim_m 1, \quad w \in \mathcal{W}_0
\]

(4.11)
The terms \( \text{rem } 5.1 \) from \cite{19} are met, we provide a short dictionary to translate the notation. Then for \( \delta \) follows from (4.11).

Thus the lemma holds with (4.12). The lower bound takes a little more work. Our argument is inspired by \cite{20, 3, 5}, though care must be taken to obtain uniform results. The \( A \) priori, the results of \cite{19} only guarantee that for each \( m \geq 0 \), there exists an admissible constant \( \eta > 0 \) such that the conclusions hold on \( B(\eta) \), whenever \( 0 < \delta \leq \delta_M \) (for a slightly larger \( M = M(m, d, k) \)). To see that this implies the lemma, define

\[
D^\eta_{v_0, t_0}(t_1, \ldots, t_d) = (\eta^{v_0\deg w_1} t_1, \ldots, \eta^{v_0\deg w_d} t_d),
\]

and observe that

\[
\tilde{\Phi}^{\eta \delta} = \Phi^{\delta} \circ D^\eta_{v_0, t_0}, \quad Y^\delta_w = (D^\eta_{v_0, t_0})^{-1} \eta^{v_0\deg w} Y_w \circ D^\eta_{v_0, t_0}.
\]

Thus the lemma holds with \( \delta_m = \eta \delta_M \).

**Lemma 4.4.** For \( \delta > 0 \), define

\[
\tilde{\Psi}^{J, \delta}(t_1, \ldots, t_d) = e^{t_1 Y_{J_1}^\delta} \circ \cdots \circ e^{t_d Y_{J_d}^\delta}(0).
\]

Then for \( 0 < \delta \leq \delta_0 := \delta |b_0| + 2 \),

\[
\max_{J \in \{1, \ldots, k\}^d} \| \det D\tilde{\Psi}^{J, \delta} \|_{C^0(\overline{B(1)})} \sim 1,
\]

(4.13)

with implicit constants depending only on \( b_0, d \). Here \( Y^\delta_w \) and \( \delta_m, m \geq 1 \), are as in Lemma 4.3.

**Proof.** Let \( m = |b_0| + 2 \). The upper bound,

\[
\| \det D\tilde{\Psi}^{J, \delta} \|_{C^0(\overline{B(1)})} \lesssim 1
\]

follows from (4.11).

The lower bound takes a little more work. Our argument is inspired by \cite{20, 3, 5}, though care must be taken to obtain uniform results. The \( d \)-tuple \( I_0 \) in (4.1), (4.2) and the set \( \mathcal{W}_0 \) defined by (4.5) are admissible, so it suffices to prove the lemma under the assumption that these are fixed parameters.

Consider for a moment an arbitrary collection \( Y_1, \ldots, Y_k \) satisfying

\[
\| Y_w \|_{C^m(\overline{B(1)})} \lesssim 1, \quad w \in \mathcal{W}_0, \quad |\det(Y_{w_1}, \ldots, Y_{w_d})| \sim 1, \text{ on } B(1).
\]

(4.14)

Our first step is to show that the left side of (4.13) (omitting the \( \delta \)'s) is nonzero. For \( 1 \leq i \leq d \) and \( J \in \{1, \ldots, k\}^d \), define

\[
\tilde{\Psi}_i^J(t_1, \ldots, t_i) = e^{t_1 Y_{J_1}} \circ \cdots \circ e^{t_i Y_{J_i}}(0).
\]

If

\[
\max_{J \in \{1, \ldots, k\}^d} \| \det D\tilde{\Psi}_i^J \|_{C^0(\overline{B(1)})} \neq 0
\]

(4.15)

then we’re done. Otherwise, we can choose \( i \) so that \( i + 1 \) is the minimal index with

\[
\max_{J \in \{1, \ldots, k\}^{i+1}} \| \partial_{t_1} \tilde{\Psi}_{i+1}^J \wedge \cdots \wedge \partial_{t_{i+1}} \tilde{\Psi}_{i+1}^J \|_{C^0(\overline{B(1)})} = 0.
\]
By (4.14), the $Y_i$ can’t all vanish at 0, so
\[
\max_{J \in \{1, \ldots, k\}^d} \| \frac{\partial}{\partial \xi} \tilde{\Psi}^J \|_{C^0(B(1))} \neq 0,
\]
which implies that this minimal $i$ is at least 1 (but less than $d$, by assumption).

By minimality of $i$, there exist $J \in \{1, \ldots, k\}^d$, $t_0 \in B(1) \subset \mathbb{R}^d$, and $\varepsilon > 0$ such that $\tilde{\Psi}_t^J$ is an injective immersion on $B_{t_0}(\varepsilon) = \{ t \in \mathbb{R}^d : |t - t_0| < \varepsilon \}$. Let $\Sigma = \tilde{\Psi}_t^J(B_{t_0}(\varepsilon))$. By our assumption and the definition of exponentiation, for $1 \leq j \leq k$, and $(t_1, \ldots, t_i) \in B(1)$,
\[
0 = (\partial_{t_1} \tilde{\Psi}_t^{J(j,1)} \wedge \cdots \wedge \partial_{t_{i+1}} \tilde{\Psi}_t^{J(j,1)})(t_1, \ldots, t_i, 0) = (\partial_{t_1} \tilde{\Psi}_t^J \wedge \cdots \wedge \partial_{t_1} \tilde{\Psi}_t^J)(t_1, \ldots, t_i) + Y_j(\tilde{\Psi}_t^J(t_1, \ldots, t_i)).
\]
Therefore $Y_1, \ldots, Y_k$ are tangent to $\Sigma$, as must be any Lie brackets that are defined (so all of those up to order $m$). Since $m \geq |b_0|$, this contradicts (4.14), so (4.15) must hold.

To prove the existence of a uniform lower bound, we again argue by contradiction. Assume that there is no uniform lower bound
\[
\max_{J \in \{1, \ldots, k\}^d} \| \det D\Psi^J \|_{C^0(B(1))} \gtrsim 1,
\]
holding for $0 < \delta \leq \delta_0$. Then there exist sequences $Y_1^{(n)}(1), \ldots, Y_k^{(n)}$, coming from vector fields $X_1^{(n)}, \ldots, X_k^{(n)}$ and parameters $0 < \delta(n) \leq \delta_0(n)$, such that
\[
\|Y_w^{(n)}\|_{C^0(B(1))} \lesssim m, \quad w \in \mathcal{W}_0, \quad |\det(Y_{w_1}^{(n)}(1), \ldots, Y_{w_d}^{(n)}(1))| \sim 1 \text{ on } B(1),
\]
but such that
\[
\max_{J \in \{1, \ldots, k\}^d} \| \det D\Psi^J \|_{C^0(B(1))} \to 0, \quad (4.16)
\]
where $\Psi^{J,(n)}(t_1, \ldots, t_d) = e^{t_1 Y_1^{(n)}} \circ \cdots \circ e^{t_d Y_d^{(n)}}(0)$. By Arzela–Ascoli, after passing to a subsequence,
\[
Y_w^{(n)} \to Y_w \text{ in } C^m(B(1)), \quad w \in \mathcal{W}_0, \quad (4.17)
\]
where for $w \notin \{1, \ldots, k\}$, $w$ is interpreted as a formal index. Since $m \geq 1$, $Y_w^{(n)} = [Y_w^{(n)} Y_i^{(n)}] \to [Y_w Y_i]$, so $Y_{w_1}^{(n)} = Y_{w_1}^{(n)}$ for all $w \in \mathcal{W}_0$, which implies $Y_w^{(n)} = Y_w$ for all $w \in \mathcal{W}_0$. Furthermore, the $Y_i$ satisfy (4.14). In this case, (4.16) and (4.17) contradict (4.15). Therefore (4.13) must hold with a uniform lower bound. This completes the proof. \hfill $\square$

For $J \in \{1, \ldots, k\}^d$, we recall that
\[
\Psi^J(t_1, \ldots, t_d) = e^{t_1 X_1 \circ \cdots \circ e^{t_d X_d}}(0).
\]
In the next lemma, we transfer the inequality in Lemma 4.4 over to $\Psi^J$.

**Lemma 4.5.** If $J \in \{1, \ldots, k\}^d$ and $\alpha$ is a multiindex with
\[
\nu_0 \cdot (\deg J + \deg J \alpha) < \nu_0 \cdot b_0,
\]
we have $\partial^\alpha \det D\Psi^J(0) = 0$. Additionally,
\[
\sum_{J \in \{1, \ldots, k\}^d} \sum_{\alpha \in \mathbb{Z}_0^d} |\partial^\alpha \det D\Psi^J(0)| \sim |\lambda_{b_0}(0)|. \quad (4.18)
\]
Proof. For \( \delta > 0 \), define

\[ \Psi^{J,\delta} = \Psi^J \circ D^{\delta}, \]

where \( D^{\delta}(t_1, \ldots, t_d) = (\delta^{v_0}t_1, \ldots, \delta^{v_d}t_d) \).

By naturality of the Lie bracket,

\[ \Psi^{J,\delta} = \Phi^\delta \circ \Psi^{J,\delta}. \]

Let \( M = M(b_0, d) \) be an integer, sufficiently large to satisfy the requirements below, and let \( \delta_0 = \delta M \). We may assume that \( M \geq |b_0| + 2 \), so by Lemma 4.4 and inequality (4.8), for \( 0 < \delta \leq \delta_0 \),

\[ \max_{J \in \{1, \ldots, k\}^d} \| \det D\Psi^{J,\delta} \|_{C^0(B(1))} \sim \delta^{v_0 - b_0} |\lambda_{I_0}(0)|. \] (4.19)

Let \( P^{J,\delta} \) be the degree \( M - 2 \) Taylor polynomial of \( \det D\Psi^{J,\delta} \), centered at 0. Then

\[ \| P^{J,\delta} - \det D\Psi^{J,\delta} \|_{C^0(B(1))} = (\frac{\delta}{\delta_0})^{v_0 - \deg J} \| P^{J,\delta_0} - \det D\Psi^{J,\delta_0} \|_{C^0(D^{J/\delta_0}(B(1)))} \]

\[ \lesssim M \left( \frac{\delta}{\delta_0} \right)^{v_0 - \deg J} \min v_0(M-1) \| D^{M-1} \det D\Psi^{J,\delta_0} \|_{C^0(D^{J/\delta_0}(B(1)))} \]

\[ \lesssim M \left( \frac{\delta}{\delta_0} \right)^{v_0 - \deg J + \min v_0(M-1)}, \]

where the last inequality follows from (4.7).

We may further assume that \( M \) is so large that \( v_0 \cdot b_0 < (M-1) \min v_0 \), and such that for every multiindex \( \alpha \) with \( \deg J + \deg J \alpha = b_0 \) for some \( J \in \{1, \ldots, k\}^d \), \( |\alpha| \leq M - 2 \). By the equivalence of all norms on the finite dimensional vector space of degree \( M - 2 \) polynomials in \( d \) variables,

\[ \| P^{J,\delta} \|_{C^0(B(1))} \sim_{M,d} \sum_{|\alpha| \leq M-2} |\partial^\alpha P^{J,\delta}(0)| = \delta^{v_0 - (\deg J + \deg J \alpha)} \sum_{|\alpha| \leq M-2} |\partial^\alpha \det D\Psi^J(0)|. \]

If \( J \in \{1, \ldots, k\}^d \) and \( \alpha \) is a multiindex with \( v_0 \cdot (\deg J + \deg J \alpha) \leq v_0 \cdot b_0 \), then

\[ \delta^{v_0 - b_0} |\partial^\alpha \det D\Psi^J(0)| \lesssim \| P^{J,\delta} \|_{C^0(B(1))} \]

\[ \lesssim \| \det D\Psi^{J,\delta} \|_{C^0(B(1))} + \left( \frac{\delta}{\delta_0} \right)^{v_0 - \deg J + \min v_0(M-1)} \]

\[ \lesssim \delta^{v_0 - b_0} |\lambda_{I_0}(0)| + \left( \frac{\delta}{\delta_0} \right)^{v_0 - \deg J + \min v_0(M-1)}. \]

Letting \( \delta \to 0 \), we learn two things:

\[ |\partial^\alpha \det D\Psi^J(0)| = 0, \text{ if } v_0 \cdot (\deg J + \deg J \alpha) < v_0 \cdot b_0, \] (4.20)

\[ |\partial^\alpha \det D\Psi^J(0)| \lesssim |\lambda_{I_0}(0)|, \text{ if } v_0 \cdot (\deg J + \deg J \alpha) = v_0 \cdot b_0, \] (4.21)

where the implicit constant in (4.21) is admissible.

By (4.19) and the fact that there are only finitely many \( J \in \{1, \ldots, k\}^d \), there exists \( J \in \{1, \ldots, k\}^d \) and a sequence \( \delta_n \to 0 \) such that

\[ \| \det D\Psi^{J,\delta_n} \|_{C^0(B(1))} \gtrsim \delta_n^{v_0 - b_0} |\lambda_{I_0}(0)|. \]

From this and (4.20)

\[ \delta_n^{v_0 - b_0} |\lambda_{I_0}(0)| - \left( \frac{\delta}{\delta_0} \right)^{v_0 - \deg J + \min v_0(M-1)} \lesssim \| P^{J,\delta_n} \|_{C^0(B(1))} \]

\[ \lesssim \sum_{|\alpha| \leq M-2} \delta_n^{v_0 - (\deg J + \deg J \alpha) |\partial^\alpha \det D\Psi^J(0)|} \]
Letting \( n \to \infty \), and applying (4.20),
\[
|\lambda_{b_0}(0)| \lesssim \sum_{v_0 \cdot (\deg J + \deg_j \alpha) = v_0 \cdot b_0} |\partial^\alpha \det D\Psi^J(0)|.
\]
This completes the proof of (4.18).

To complete the proof of Proposition 2.3, it suffices prove the following.

**Lemma 4.6.** \( \mathcal{P}_0 = \overline{\mathcal{P}_0} \).

Indeed, assuming the lemma, by our choice of \( v_0 \) (so that \( b \cdot v_0 \) is minimized at \( b_0 \)), the only multiindices appearing in the sum on the right of (4.18) are those with \( \deg J + \deg_j \alpha \), so Proposition 2.3 is proved.

**Proof of Lemma 4.6.** Suppose that \( b_0 \notin \overline{\mathcal{P}_0} \) is an extreme point of the convex hull \( \text{ch}(\mathcal{P}_0 \cup \overline{\mathcal{P}_0}) \). By Proposition 4.1, there exist \( \varepsilon > 0 \), \( v_0 \in (\varepsilon, 1]^k \) such that \( b_0 \cdot v_0 + \varepsilon < b \cdot v_0 \) for every \( b \in ((\mathcal{P}_0 \cup \overline{\mathcal{P}_0}) \cap \mathbb{Z}_0^k) \setminus \{b_0\} \). By Lemmas 4.2-4.5, there exist \( J \in \{1, \ldots, k\}^d \) and a multiindex \( \alpha \) with \( \deg J + \deg_j \alpha \in \overline{\mathcal{P}_0} \) and \( v_0 \cdot (\deg J + \deg_j \alpha) = v_0 \cdot b_0 \), a contradiction. Therefore every extreme point of \( \text{ch}(\mathcal{P}_0 \cup \overline{\mathcal{P}_0}) \) is contained in \( \overline{\mathcal{P}_0} \), so \( \mathcal{P}_0 \subset \overline{\mathcal{P}_0} \).

Now suppose that \( b_0 \notin \mathcal{P}_0 \) is an extreme point of \( \text{ch}(\mathcal{P}_0 \cup \overline{\mathcal{P}_0}) \). By Proposition 4.1, there exist \( \varepsilon_0 > 0 \) and \( v_0 \in (\varepsilon_0, 1]^k \) such that \( v_0 \cdot b_0 + \varepsilon_0 < v_0 \cdot b \) for every \( b \in (\mathcal{P}_0 \cup \overline{\mathcal{P}_0}) \cap \mathbb{Z}_0^k \setminus \{b_0\} \). Let
\[
\mathcal{B} = \{b \in \mathcal{P}_0 \cap \mathbb{Z}_0^k : v_0 \cdot b = \min_{b' \in \mathcal{P}_0} v_0 \cdot b'\}.
\]
Since \( \mathcal{P}_0 \neq \emptyset \) by assumption, \( \mathcal{B} \neq \emptyset \). Let \( b_1 \) be an extreme point of \( \mathcal{P}(\mathcal{B}) \) be extreme; then \( b_1 \) is also an extreme point of \( \mathcal{P}_0 \), so by Proposition 4.1, there exist \( \varepsilon_1 > 0 \) and \( v_1 \in (\varepsilon_1, 1]^k \) such that \( v_1 \cdot b_1 + \varepsilon_1 < v_1 \cdot b \) for every \( b \in \mathcal{P}_0 \cap \mathbb{Z}_0^k \setminus \{b_1\} \).

For \( \delta > 0 \), let \( v_\delta = v_0 + \delta v_1 \). By construction, \( v_\delta \cdot b_1 + \delta \varepsilon < v_\delta \cdot b \), for every \( b \in \mathcal{P}_0 \cap \mathbb{Z}_0^k \setminus \{b_1\} \). Thus by Lemmas 4.2–4.5, if \( J \in \{1, \ldots, k\}^d \) and \( \alpha \) is a multiindex with \( v_\delta \cdot (\deg J + \deg_j \alpha) < v_\delta b_1 \), \( \partial^\alpha \det D\Psi^J(0) \neq 0 \). On the other hand, if \( \delta < \frac{\varepsilon}{v_\delta \cdot b_0} \),
\[
v_\delta \cdot b_0 < v_\delta \cdot b_0 + \frac{\varepsilon}{2} < v_0 \cdot b_0 < v_\delta \cdot b_1,
\]
and by definition of \( \overline{\mathcal{P}_0} \) and our assumption that \( b_0 \) is an extreme point, there must exist \( J \in \{1, \ldots, k\}^d \) and a multiindex \( \alpha \) such that \( \deg J + \deg_j \alpha = b_0 \). This gives a contradiction, and so we must have \( \overline{\mathcal{P}_0} \subset \mathcal{P}_0 \). This completes the proof of Lemma 4.6 and thus of Proposition 2.3.

**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 \( X_1, \ldots, X_d \). Using the approximation \( \exp(tX) = \sum_{n=0}^{\infty} \frac{t^n}{n!} X^{n-1} + O(|t|^N) \), which may be found in [5], and somewhat tedious computations, one can show that
\[
\partial^\alpha_{t_i} \big|_{t=0} \det D_t(e^{t_1 X_1} \circ \cdots \circ e^{t_d X_d})(x_0) = \pm \sum_{w_1, \ldots, w_d} \prod_{i=1}^d \left( \frac{\alpha_i}{\deg_i w_{i+1}, \ldots, \deg_i w_d} \right) \deg(X_{w_1}, X_{w_2}, \ldots, X_{w_d}),
\]
where the sum is taken over those words \( w_i = (w_i^1, \ldots, w_i^n) \) satisfying \( \sum \deg w_i = \alpha + (1, \ldots, 1) \) and \( w_i^1 = i > w_i^2 \geq \cdots \geq w_i^n \) (in particular, \( w_1^1 = 1 \)). Replacing \( X_i \) above with \( X_I \), gives an alternative proof that the right side of (2.10) is bounded by the left, but using this formula to bound the left of (2.10) by the right seems nontrivial.

It is not true that the left of (2.10) is bounded by the right if no assumptions are made on the vector fields or the point \( b_0 \). To see this, let \( \gamma(t) = (t, \ldots, t^d) \) and define vector fields on \( \mathbb{R}^{d+1} \) by \( X_0 = \partial_t, \quad X_i = \partial_t - \gamma'(t) \cdot \nabla x, \quad 1 \leq i \leq d \). Then

\[
I = (1, 2, (1, 3), \ldots, (1, \ldots, 1, d))
\]

is a minimal spanning \( d \)-tuple of words with degree \( b = (\frac{d(d-1)}{2} + 1, 1, \ldots, 1) \), but the quantity on the right of (2.10) is zero.

Less uniform versions of (2.10) may be found in [20, 18, 5]. Let \( X_1, \ldots, X_k \) be vector fields which satisfy the Hörmander condition uniformly on \( U \), let \( \varepsilon > 0 \), and let \( \delta_1, \ldots, \delta_k \) be vector fields satisfying the smallness and weak comparability conditions

\[
\delta_i \leq K, \quad \delta_i \leq K \delta_j, \quad 1 \leq i, j \leq k.
\]

Then [20, 18] establish

\[
\sum_{|\deg I| \leq N} (\prod_{i=1}^k \delta_i^{\deg I}) |\lambda_I(x_0)| \sim \sum_{J \in \{1, \ldots, k\}^d \alpha : \deg J - \deg \alpha \leq N'} \sum_{i=1}^k (\prod_{i=1}^k \delta_i^{\deg J + \deg J, i}) |\partial_i^\alpha \deg D_t \Psi_{x_0}(0)|,
\]

for \( x_0 \in U \). It is not shown, however, how to remove the dependence of the implicit constant on \( \varepsilon, K \), or the \( X_i \), or, in particular, how to remove the assumption that the Hörmander condition holds uniformly.

5. Proof 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 \( \mu \) has compact support. Fix an open set \( V \Subset U \) with \( \text{supp} \mu \Subset V \). Suppose that \( p_{j_0} < 1 \). For each \( \varepsilon > 0 \), we may cover \( \pi_j(V) \) by \( O(\varepsilon^{-(d-1)}) \) balls \( B_i \) of radius \( \varepsilon \) with finite overlap. Then by our assumption,

\[
\mu(U) \leq \sum_j \mu(\pi_j^{-1}(B_j)) = \sum_j \int_{\mathbb{R}^d} \chi_{B_{j_0}} \circ \pi_{j_0}(x) \, d\mu \lesssim \varepsilon^{(\frac{1}{p_{j_0}} - 1)(d-1)},
\]

and letting \( \varepsilon \to 0 \), we see that \( \mu \equiv 0 \).

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

\[
Z = \{x \in V : b_p \notin P_x\}
\]
\[
\Omega = \{x \in V : b_p \text{ is an extreme point of } P_x\},
\]

or the quantity on the right of (2.8).

The proposition will follow from the following two lemmas.

**Lemma 5.1.** \( \mu(Z) = 0 \).
Lemma 5.2. Let \( \rho = \sum_{t-b_p} \lambda_1 \) and let
\[
\Omega_n = \{ x \in \Omega : 2^n \leq \rho(x) \leq 2^{n+1} \}, \quad n \in \mathbb{Z}.
\]
If \( \Omega' \subset \Omega_n \) is compact, \( \mu(\Omega') \leq C_{d,p} 2^n |\Omega'| \).

Proof of Lemma 5.1. By Proposition 4.1,
\[
Z = \bigcup_{i=1}^{c_{p,d}} Z_{A_i},
\]
where each \( A_i \subset \mathbb{Z}_{k}^n \) is a finite set satisfying
\[
b_p \notin \mathcal{P}(A_i), \quad \mathcal{P}_x \subset \mathcal{P}(A_i), \quad x \in Z_{A_i}, \quad 1 \leq i \leq c_{p,d}.
\]
Let \( A = A_i \) be fixed for the remainder of the argument and let \( \Omega' = Z_{A_i} \); we will show that \( \mu(\Omega') = 0 \).

Choose \( \varepsilon > 0 \) and \( v \in (\varepsilon, 1]^k \) such that
\[
v \cdot b_p + \varepsilon < v \cdot b, \quad \text{for } b \in \mathcal{P}(A),
\]
and let
\[
W_0 = \{ w \in W : v \cdot \text{deg } w \leq d \}.
\]
Then \( \{ X_w : w \in W_0 \} \) contains the coordinate vector fields. Let \( N = N_{d,k,p} \) be an integer whose size will be determined in a moment and which is, in particular, larger than \( \frac{d}{\varepsilon} \). Since \( V \Subset U \) and the \( X_i \) are smooth on \( U \), there exists \( \delta_0 > 0 \), depending on the \( X_i \) and on \( p \), such that for all \( 0 < \delta \leq \delta_0 \) and \( I \in W_0 \) satisfying \( \text{deg } I \in \mathcal{P}(A) \), and all \( w, w' \in W_0 \),
\[
|\delta^{v\text{-deg }I} \lambda_I(x)| < \delta^v \delta^{v\cdot b_p}, \quad (5.1)
\]
\[
|\delta^{v\text{-deg }w} X_w|_{C^0(V)} \leq \frac{1}{\delta} \text{dist}(V, \partial U), \quad |\delta^{v\text{-deg }w} X_w|_{C^N(V)} \leq 1
\]
\[
[\delta^{v\text{-deg }w} X_w, \delta^{v\text{-deg }w'} X_{w'}] = \sum_{\tilde{w}, \tilde{w}' \in W_0} c_{\tilde{w}, \tilde{w}', \partial} \delta^{v\text{-deg }\tilde{w}} X_{\tilde{w}},
\]
with
\[
||c_{\tilde{w}, \tilde{w}', \partial}||_{C^N(V)} \lesssim 1,
\]
where the implicit constant above depends on admissible quantities, and hence ultimately on \( d, p \). (Indeed, this follows in much the same way as Lemma 4.2.)

For \( x \in \Omega' \) and \( 0 < \delta \leq \delta_0 \), choose \( I^\delta_x \in W_0 \) such that
\[
\delta^{v\text{-deg }I^\delta_x} |\lambda_{I^\delta_x}(x)| = \max_{I \in W_0} \delta^{v\text{-deg }I} |\lambda_I(x)|.
\]
Let
\[
\Phi^\delta_x(t_1, \ldots, t_d) = \exp(t_1 \delta^{v\text{-deg }w_1} X_{w_1} + \cdots + t_d \delta^{v\text{-deg }w_d} X_{w_d})(x)
\]
\[
B(x, \delta) = \{ \Phi^\delta_x(t) : |t| < 1 \},
\]
where \( I^\delta_x = (w_1, \ldots, w_d) \).

By the results of [19], provided \( N = N_{d,p} \) is sufficiently large, these balls are doubling in the sense of [16], with constants depending only on \( d, p, N \). Furthermore,
\[
|B(x, \delta)| \sim \delta^{v\text{-deg }I^\delta_x} |\lambda_{I^\delta_x}(x)| \quad (5.3)
\]
\[
\exp(tX_i)(y) \in B(x, C\delta) \quad \text{whenever } y \in B(x, \delta), \quad |t| < \delta^{v'}.
\]

(5.4)
where $C = C_{d,p}$. By (5.3), the doubling property, and the fact that the $X_i$ are nonvanishing,

$$
|B(x, \frac{\delta}{2})| \sim |B(x, C\delta)| \sim \int_{\pi_i(B(x, C\delta))} \mathcal{H}^1(\pi_i^{-1}\{y\} \cap B(x, C\delta))
\geq \int_{\pi_i(B(x, \delta))} \mathcal{H}^1(\pi_i^{-1}\{y\} \cap B(x, \delta))
\geq \delta^v|\pi_i(B(x, \delta))|.
$$

(5.5)

By standard covering arguments and the doubling property, for each $0 < \delta \leq \delta_0$, there exists \(\{x_j\}_{j=1}^{M_s} \subset \mathbb{Z}'\) such that $Z' \subset \bigcup_{j=1}^{M_s} B(x_j, \delta)$ and such that the balls $B(x_j, C^{-1}\delta)$ are pairwise disjoint.

Putting this all together, we conclude that

$$
\mu(V) \leq \sum_{j=1}^{M_s} \mu(B(x_j, \delta)) \lesssim \sum_{j=1}^{k} \prod_{i=1}^{k} |\pi_i(B(x_j, \delta))|^{\frac{1}{\pi_i}}
\lesssim \sum_{j=1}^{M_s} |B(x_j, C\delta)|^{\sum_{i=1}^{k} \frac{1}{\pi_i}} \prod_{i} \delta^{-\frac{1}{\pi_i}}
\lesssim \sum_{j=1}^{M_s} |B(x_j, C\delta)|(\delta^{v-deg I_j} |\lambda_{I_j}(x_j)|)^{\sum_{i=1}^{k} \frac{1}{\pi_i}} \delta^{-v-b_p(\sum_{i=1}^{k} \frac{1}{\pi_i})}
\lesssim \sum_{j=1}^{M_s} |B(x_j, C^{-1}\delta)|(\delta^{v}(\sum_{i=1}^{k} \frac{1}{\pi_i})^{-1}).
$$

Since the $B(x_j, C^{-1}\delta)$ are disjoint, $\sum_j |B(x_j, C^{-1}\delta)| \leq |\bigcup_j B(x_j, C^{-1}\delta)| \lesssim 1$, and the lemma follows by sending $\delta$ to 0.

**Proof of Lemma 5.2.** The proof is similar to that of Lemma 5.1.

Fix $n$, and observe that on $\Omega_n$, $\max_{deg I=b_p} |\lambda_I(x)| \sim 2^{n(l(b_p)+1)-1}$. By Proposition 4.1, it suffices to consider $\Omega' \subset \Omega_n$ for which there exists a finite set $A \subset \mathbb{Z}_0^k$ such that

$$
b_p \notin \mathcal{P}(A), \quad \mathcal{P}_x \subset \mathcal{P}(A \cup \{b_p\}), \quad x \in \Omega'.
$$

(5.6)

Choose $\varepsilon > 0$, $v \in (\varepsilon, 1]^k$ to satisfy

$$
v \cdot b_p + \varepsilon < v \cdot b, \quad b \in \mathcal{P}(A \cup \{b_p\}) \cap \mathbb{Z}_0^k \setminus \{b_p\},
$$

and let

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

Since $(1, \ldots, 1, 0, \ldots, 0) \in \mathcal{P}_x$, $x \in U$, $(1, \ldots, 1, 0, \ldots, 0) \in \mathcal{P}(A \cup \{b_p\})$. Therefore $v \cdot b_p < \sum_{i=1}^{d} v^i \leq d$, so deg $I = b_p$ implies that $I \in \mathcal{W}_0$. Again let $N = N_{d,p}$ be a large integer. As before, there exists $\delta_n > 0$ such that $0 < \delta \leq \delta_n$ implies the following:

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

for all $I \in \mathcal{W}_0$ satisfying $\deg I \neq b_p$,

$$
\|\delta^{v-deg w} X_w\|_{C^0(V)} \leq \frac{1}{2} \text{dist}(V, \partial U), \quad \|\delta^{v-deg w} X_w\|_{C^N(V)} \leq 1,
$$

for all $w \in \mathcal{W}_0$, and

$$
[\delta^{v-deg w} X_w, \delta^{v-deg w'} X_w] = \sum_{\bar{w}\in \mathcal{W}_0} c_{w,w',\delta} \delta^{v-deg \bar{w}} X_{\bar{w}},
$$
with
\[ \|c_{w,w'}^{w,w'}\|_{C^N(V)} \leq C_{d,p}, \]
for all \( w, w' \in W_0 \). In particular, if \( \delta_n \) is sufficiently small,
\[ \delta^{v-\text{deg} l} |\alpha_l(x)| = \max_{I \in W_0^d} \delta^{v-\text{deg} I} |\alpha_I(x)| \sim \delta^{v-b_p} 2^n (|b_p| - 1) \]
for some \( d \)-tuple \( I^d \in W_0^d \) satisfying \( \text{deg} I^d = b_p \). Thus, considering the balls \( B(x, \delta) \) defined in (5.2), for \( x \in \Omega' \), \( 0 < \delta \leq \delta_n \),
\[ |B(x, \delta)| \sim 2^{n(|b_p| - 1)} \delta^{v-b_p} = 2^{\frac{n}{p^* - 1}} \delta^{v-b_p}. \]

Since the balls \( B(x, \delta) \) are doubling, for each \( \eta > 0 \), there exists \( \delta > 0 \), \( 0 < \delta \leq \delta_n \), and \( \{x_j\}_{j=1}^{M_k} \subset \Omega' \) such that
\[ \Omega' \subset \bigcup_{j=1}^{M_k} B(x_j, \delta), \quad |\bigcup_{j=1}^{M_k} B(x_j, \delta)| \leq |\Omega'| + \eta, \]
and such that the \( B(x_j, C^{-1}\delta) \) are pairwise disjoint.

Arguing as in the proof of Lemma 5.1,
\[ \mu(\Omega') \leq \sum_{j=1}^{M_k} \mu(B(x_j, \delta)) \lesssim \sum_j |B(x_j, C\delta)||B(x_j, C\delta)|^{\frac{1}{p^* - 1}} \delta^{-v-b_p} (\sum_i \frac{1}{p_i} - 1) \]
\[ \sim \sum_j |B(x_j, C\delta)|^{2^\alpha} \lesssim 2^n (|\Omega'| + \eta), \]
and the implicit constants depend only on \( d, p \). Letting \( \eta \to 0 \) completes the proof. \( \square \)

**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 \( \mathcal{P}_x \), where \( \theta p = (\theta p_1, \ldots, \theta p_k) \). 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 [18], if \( a \) is continuous with compact support in \( U \),
\[ |\int \prod_{j=1}^{k} f_j \circ \pi_j(x) a(x) \, dx| \lesssim \prod_{j=1}^{k} \|f_j\|_{L^{p_j}}. \]
Additionally,
\[ |\int \prod_{j=1}^{k} f_j \circ \pi_j(x) \log |x - x_0| a(x) \, dx| \lesssim \prod_{j=1}^{k} \|f_j\|_{L^{\infty}}. \]
Thus by interpolation,
\[ |\int \prod_{j=1}^{k} f_j \circ \pi_j(x) \log |x - x_0|^{1-\theta} a(x) \, dx| \lesssim \prod_{j=1}^{k} \|f_j\|_{L^{p_j}}. \]

For the unweighted bilinear operator in the ‘polynomial-like’ case, the endpoint restricted weak type bounds are known and due to Gressman in [10]; in the multilinear case, the corresponding estimates follow by combining his techniques with arguments in [18]. The deduction of endpoint bounds from the arguments in [10] does not seem to be immediate in the weighted case, and so these questions remain
open except in certain special configurations (such as convolution or restricted X-ray transform along polynomial curves).

6. Proof of the main theorem: Theorem 2.1

We will prove a somewhat more general result. Fix \( J_0 \in \{1, \ldots, k\}^d \) and for \( x \in U \), define \( \Psi_x^J(t) \) as in (2.9). Let \( \beta \) be some fixed multiindex, let \( b_0 = \deg J_0 + \deg J_0 \beta_0 \), and consider

\[
\tilde{\rho}(x) = |\partial^{\beta_0}_{x_0}|_{\beta_0 = 0} \det D_t \Psi_x^J(t)|^{\frac{1}{|\beta_0|}}.
\]  

(6.1)

Let \( a \) be a continuous cutoff function with compact support contained in \( U \) and define the multilinear form

\[
\tilde{\mathcal{M}}(f_1, \ldots, f_k) = \int_{\mathbb{R}^d} \prod_{j=1}^k f_j \circ \pi_j(x) \tilde{\rho}(x) a(x) \, dx.
\]

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

**Proposition 6.1.** Let \( (p_1, \ldots, p_k) \in [1, \infty)^k \) satisfy \( (p_1^{-1}, \ldots, p_k^{-1}) \prec q(b_0) \), with \( p_i^{-1} < q_i(b_0) \) when \( b_i \neq 0 \). Then

\[
|\tilde{\mathcal{M}}(f_1, \ldots, f_k)| \lesssim \prod_{j=1}^k \|f_j\|_{L^{p_j}},
\]

for all continuous \( f_1, \ldots, f_k \).

Since \( J_0 \) and \( \beta_0 \) are fixed, we will henceforth drop the tildes from our notation, with the understanding that we are using (6.1) instead of (2.3). Our first step is to reduce the proof of Proposition 6.1 to a simpler restricted weak type estimate.

If \( \Omega \subset \text{supp } a \) is measurable, let \( \rho(\Omega) = \int_{\Omega} \rho(x) \, dx \).

**Lemma 6.2.** Let \( \delta, \varepsilon > 0 \). Let \( \Omega \subset \text{supp } a \) be a Borel set, and define

\[
\alpha_j = |\pi_j(\Omega)|^{-1} \rho(\Omega), \quad 1 \leq j \leq k.
\]

Assume that \( \frac{1}{2} \rho_0 \leq \rho(x) \leq 2 \rho_0 \) for \( x \in \Omega \), that \( \alpha_j < \frac{1}{2} \), \( 1 \leq j \leq k \), and that

\[
\rho_0 \geq c_{\delta, \varepsilon}^{-1} \max_j (\alpha_j)^{1-\delta} \quad \text{and} \quad \text{diam}(\Omega) \leq c_{\delta, \varepsilon} (\min_j \alpha_j)^{\delta}.
\]  

(6.2)

Then if \( \delta, \varepsilon \) are sufficiently small and \( c_{\delta, \varepsilon} \) is sufficiently small, depending on \( \delta, \varepsilon, d, b_0 \), and the \( \pi_j \), we have

\[
\prod_{j=1}^k \alpha_j^{b_j + \varepsilon} \lesssim \rho(\Omega).
\]  

(6.3)

The implicit constant depends on \( \varepsilon, \delta, b_0 \), and the the \( \pi_j \), but is independent of \( \rho_0, \Omega \).

**Proof of Proposition 6.1 from Lemma 6.2.** It obviously suffices to prove (2.6) for nonnegative \( f_1, \ldots, f_k \). In this case, by Hölder’s inequality and our assumptions on the \( p_j \),

\[
\mathcal{M}(f_1, \ldots, f_k) \leq \prod_{\{j: b_j = 0\}} \|f_j\|_{L^\infty} \int_{\mathbb{R}^d} \prod_{\{j: b_j \neq 0\}} f_j \circ \pi_j(x) \rho(x) a(x) \, dx
\]
\begin{align*}
\leq \prod_{\{j : p_j = \infty\}} \|f_j\|_{L^{p_j}} \int_{\mathbb{R}^d} \prod_{\{j : p_j \neq \infty\}} f_j \circ \pi_j (x) \rho(x) a(x) \, dx,
\end{align*}
so we may further assume that \( b_0^j \neq 0, p_j \neq \infty, j = 1, \ldots, k \).

We are only considering the non-endpoint case, so by real interpolation with the trivial (by Hölder) inequality

\begin{align*}
M(f_1, \ldots, f_k) \lesssim \prod_{j=1}^{k} \|f_j\|_{L^{p_j}}, \quad \sum_{j=1}^{k} p_j^{-1} \leq 1,
\end{align*}

it suffices (recalling the definition (2.5) of \( q \) and the assumptions on the \( p_j \)) to prove that for all Borel sets \( E_1, \ldots, E_k \) and sufficiently small \( \varepsilon > 0 \),

\begin{align*}
\int_{\mathbb{R}^d} \prod_{j=1}^{k} \chi_{E_j} \circ \pi_j (x) \rho(x) a(x) \, dx \lesssim \prod_{j=1}^{d} |E_j|^{q_j(b_0)^{-\varepsilon}}. \tag{6.4}
\end{align*}

Letting

\begin{align*}
\Omega = \left( \bigcup_{j=1}^{k} \pi_j^{-1}(E_j) \right) \cap \text{supp} \, a,
\end{align*}

(6.4) follows from

\begin{align*}
\rho(\Omega) \lesssim \prod_{j=1}^{d} |\pi_j(\Omega)|^{q_j(b_0)^{-\varepsilon}}, \tag{6.5}
\end{align*}

By a bit of arithmetic, (6.5) follows from

\begin{align*}
\prod_{j=1}^{d} \alpha_j^{q_j(q(b_0)^{-\varepsilon}), \ldots, \varepsilon)} \lesssim \rho(\Omega),
\end{align*}

which in turn is implied by

\begin{align*}
\prod_{j=1}^{d} \alpha_j^{b_j^j + \varepsilon} \lesssim \rho(\Omega), \tag{6.6}
\end{align*}

with a slightly smaller \( \varepsilon \).

By the coarea formula,

\begin{align*}
\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}
\end{align*}

We recall that since the \( \pi_j \) are submersions, the \( X_j \) never vanish, and since the support of \( a \) is compact, \( \frac{\rho}{|X_j|} \) is bounded. Thus (6.7) implies that

\begin{align*}
\alpha_j \lesssim \text{diam}(\Omega) \leq \text{diam}(\text{supp} \, a).
\end{align*}

By taking a partition of unity of the support of \( a \), we may assume that \( \text{diam}(\Omega) \) is so small that \( \alpha_j < \frac{1}{2}, j = 1, \ldots, k \). Reordering if necessary, \( \alpha_1 \leq \cdots \leq \alpha_k \).

Let \( \delta > 0 \) be a small constant which will be determined (depending on \( \varepsilon, b_0, d \)) later on. Let \( c_{\delta, \varepsilon} \) be a small constant which will also be chosen later. Cover \( \Omega \) by \( O(c_{\delta, \varepsilon}^{-d} \alpha_1^{-b}) \) balls of radius \( c_{\delta, \varepsilon} \alpha_1^{\delta} \). Then the intersection \( \Omega' \) of \( \Omega \) with one of these balls must satisfy

\begin{align*}
\rho(\Omega) \geq c_{\delta, \varepsilon}^{d} \alpha_1^{d} \rho(\Omega). \tag{6.8}
\end{align*}

For \( n \in \mathbb{Z} \), let

\begin{align*}
\Omega'_n = \{ x \in \Omega' : 2^n \leq \rho(x) < 2^{n+1} \}.
\end{align*}
Therefore
\[ \Omega' = \bigcup_{n=-\infty}^{C} \Omega'_n, \]
where \( C < \infty \) depends on the \( \pi_j \).

By the coarea formula,
\[
\rho(\Omega'_n) = \int_{\pi_1(\Omega'_n)} \int_{\pi_j^{-1}(y)} \chi_{\Omega'_n}(x) \rho(x) \frac{1}{|X_j(x)|} d\mathcal{H}^1(x) dy \\
\sim \int_{\pi_1(\Omega'_n)} \int_{\pi_j^{-1}(y)} \chi_{\Omega'_n}(x) 2^n d\mathcal{H}^1(x) dy \\
\lesssim 2^n |\pi_1(\Omega'_n)| \leq 2^n |\pi_1(\Omega)|.
\]

Therefore
\[
\log \alpha_1 + \log(c_{\delta,\varepsilon}\alpha_1^{1+\delta d}) - C \\
\sum_{n=-\infty}^{\log(c_{\delta,\varepsilon}\alpha_1^{1+\delta d})} \rho(\Omega'_n) \lesssim 2^{-C} c_{\delta,\varepsilon}^d \alpha_1^{1+\delta d} \rho(\Omega) = 2^{-C} c_{\delta,\varepsilon}^d \alpha_1^{1+\delta d} \rho(\Omega),
\]
so if \( C \) is sufficiently large, depending on \( d \) and the \( \pi_j \), we have by (6.8) that
\[
\rho \left( \bigcup_{n=-\infty}^{\log(c_{\delta,\varepsilon}\alpha_1^{1+\delta d})} \Omega'_n \right) < \frac{1}{2} \rho(\Omega').
\]

Thus we may fix a value of \( n, \log(c_{\delta,\varepsilon}\alpha_1^{1+\delta d}) - C \leq n \leq C \), so that
\[
\rho(\Omega'_n) \gtrsim \frac{1}{\log(c_{\delta,\varepsilon}\alpha_1^{1+\delta d})} \rho(\Omega') \gtrsim c_{\delta,\varepsilon}^d \alpha_1^{1+\delta d} \rho(\Omega).
\]

Let \( \rho_0 \) denote this value of \( 2^n \).

Let \( \alpha_j' = |\pi_j(\Omega'_n)|^{-1} \rho(\Omega'_n) \), \( 1 \leq j \leq k \). Then
\[
\alpha_j' \gtrsim c_{\delta,\varepsilon}^d \alpha_1 \alpha_j, \quad 1 \leq j \leq k,
\]
which implies that
\[
\alpha_1 = \min_j \alpha_j^{1+2\delta d} \lesssim c_{\delta,\varepsilon}^{-1} \min_j (\alpha_j')^{\frac{1}{1+2\delta d}}.
\]

Therefore
\[
\text{diam}(\Omega'_n) \leq \alpha_j^{1+2\delta d} \lesssim c_{\delta,\varepsilon}^{-1} \min_j (\alpha_j')^{\frac{1}{1+2\delta d}}
\]
\[
\alpha_j' \leq \max_{y \in \pi_j(\Omega'_n)} \int_{\pi_j^{-1}(y)} \chi_{\Omega'_n}(x) \rho(x) d\mathcal{H}^1(x) \lesssim \text{diam}(\Omega'_n) \rho_0 \lesssim c_{\delta,\varepsilon}^{1-2\delta d} (\alpha_j')^{\frac{1-2\delta d}{1+2\delta d}} \rho_0.
\]

Thus provided \( c_{\delta,\varepsilon} \) is chosen sufficiently small, \( \Omega'_n \) satisfies the hypotheses of the lemma, with a slightly larger power of \( \delta \) (which nonetheless can be made as small as required). It remains to check that (6.6) follows from the conclusions of the lemma.

Since \( \rho(\Omega'_{n}) \leq \rho(\Omega) \), (6.6) follows from
\[
\rho(\Omega'_n) \gtrsim \prod_{j=1}^{k} \alpha_j^{b_j + \varepsilon},
\]
Lemma 6.3. If \( \alpha \geq \alpha_1^{2d} \alpha_j \), this follows from

\[
\rho(\Omega'_n) \gtrsim \alpha_1^{b_0^j + 2d + \varepsilon} \prod_{j=2}^{k} (\alpha_j)^{b_0^j + \varepsilon}.
\]

Finally, since \( (\alpha'_1)^{\frac{1}{1+2d}} \gtrsim \alpha_1 \), provided \( \delta = \delta_{\varepsilon,b_0,d,k} \) is sufficiently small, (6.6) ultimately follows from

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

where \( \varepsilon' \) is a positive number, smaller than \( \varepsilon \).

This completes the proof of Theorem 2.1 from Lemma 6.2. \( \square \)

We devote the remainder of this section to the proof of Lemma 6.2. We use the method of refinements, which originated in [4]; the specifics here largely follow from arguments that have appeared in [20, 3, 18].

Recalling (6.1),

\[
|g^{\lambda_0} \det D\Psi^b_{x_0}(0)| \sim \rho_0^{\lambda_1} \rho(\Omega) =: \lambda_0, \quad x_0 \in \Omega. \quad (6.9)
\]

Reordering if necessary, we may assume that \( \alpha_1 \leq \ldots \leq \alpha_k \).

Assume that we are given a fixed constant \( c_x > 0 \). Then, as in [20], for \( w > 0 \), we say that a set \( S \subseteq [-w, w] \) is a central set of width \( w \) if for any interval \( I \subseteq [-w, w] \),

\[
|I \cap S| \leq c_x (\frac{|I|}{w})^\varepsilon |S|.
\]

**Lemma 6.3.** If \( c_x \) is sufficiently small, for each subset \( \Omega' \subseteq \Omega \) with \( \rho(\Omega') \gtrsim \alpha_1^{C_x} \rho(\Omega) \) and each \( 1 \leq j \leq k \), there exists a refinement \( \langle \Omega'_j \rangle \subseteq \Omega' \) with \( \rho(\Omega'_j) \gtrsim \alpha_1^{C_x} \rho(\Omega') \), such that for each \( x \in \Omega'_j \),

\[
\mathcal{F}_j(x, \Omega'_j) := \{ t : |t| \lesssim \alpha_1^\delta \text{ and } e^{tX_j}(x) \in \Omega'_j \}
\]

is a central set whose width \( w_j \) and measure satisfy

\[
\rho_0^{-1} \alpha_1^{C_x} \alpha_j \lesssim w_j \lesssim c_x \alpha_1^\delta \quad \text{and} \quad |\mathcal{F}_j(x, \Omega'_j)| \gtrsim \rho_0^{-1} \alpha_1^{C_x} \alpha_j. \quad (6.10)
\]

The constant \( C_x \) depends on \( C \) but is independent of \( \varepsilon \).

For the proof of this lemma, we refer the reader to [20, Lemma 8.2]. The only change needed is notational; because \( \rho(x) \sim \rho_0 \) on \( \Omega \), measures taken with respect to \( \rho \) are, after dividing by \( \rho_0 \), comparable to Lebesgue measure.

Write \( J_0 = (j_1, \ldots, j_d) \). With \( \Omega_0 = \Omega \), for \( 1 \leq i \leq d \), we define

\[
\Omega_i = \langle \Omega_{i-1} \rangle_{j_{d-i+1}}.
\]

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

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 \( 2 \leq i \leq d \), let

\[
F_i = \{ (t_1, \ldots, t_i) : (t_1, \ldots, t_{i-1}) \in F_{i-1}, \ t_i \in \mathcal{F}_{j_i}(x_{i-1}(t_1, \ldots, t_{i-1}), \Omega_{d-i+1}) \}
\]

\[
x_i(t_1, \ldots, t_i) = e^{tX_{j_i}}(x_{i-1}(t_1, \ldots, t_{i-1})).
\]

For each \( i \) and each \( (t_1, \ldots, t_i) \in F_i \),

\[
x_i(t_1, \ldots, t_i) \in \Omega_{d-i+1} \subseteq \Omega_d.
\]
so $\mathcal{F}_{j+1}(x_i(t_1, \ldots, t_i), \Omega_{d-i})$ is a central set whose width and measure satisfy (6.10) (with $j_{i+1}$ in place of $j$). Therefore

$$\Psi_{x_i}^j(F_d) \subset \Omega \quad \text{and} \quad |F_d| \gtrsim \rho_0^{-d} \alpha_1^{C_\varepsilon} \alpha_1^{\deg J}.$$ \hfill (6.11)

Let $\Psi^N_{x_0}$ be the degree $N$ Taylor polynomial of $\Psi^N_{x_0}$, where $N$ is a large integer, to be determined in a moment. Let $Q_w = \prod_{i=1}^d [-w_i, w_i]$. Assuming $N \geq |b_0|_1$, by the equivalence of the $C^0(Q_{(1, \ldots, 1)})$ and $C^N(Q_{(1, \ldots, 1)})$ norms on the degree $N$ polynomials in $d$ variables, scaling, and (6.9),

$$\| \det \Psi^N_{x_0}(w) \|_{C^0(Q_w)} = \|(\det \Psi^N_{x_0})(w_1, \ldots, w_d)\|_{C^0(Q_{(1, \ldots, 1)})} \sim_{N,d} w_1^{\beta_0} \det \Psi^N_{x_0}(0) \sim w_1^{\beta_0} \lambda_0.$$

Thus by (6.10) and the definitions of $\beta_0, \lambda_0$,

$$\| \det \Psi^N_{x_0}(w) \|_{C^0(Q_w)} \gtrsim C_{N,d,b_0} \rho_0^{-d} \alpha_1^{C_\varepsilon} \alpha_1^{\deg J} \beta_0.$$ \hfill (6.12)

The next lemma follows from Lemma 6.2 of [3]; or alternatively by applying Lemma 7.3 of [20] iteratively.

**Lemma 6.4.** Let $T_i \subset \mathbb{R}^i$, $1 \leq i \leq d$ be sets satisfying $T_{i+1} \subset T_i \times [-1,1]$, $1 \leq i \leq d-1$. Assume that the sets

$$\tau_1 := T_1, \quad \tau_i(t) := \{s \in [-1,1] : (t, x) \in T_i\}, 2 \leq i \leq d, t \in T_{i-1}$$

are central sets of width $w_i$. Then if $P$ is a degree $N$ polynomial on $\mathbb{R}^d$, there exists a subset $T'_d \subset T_d$ of measure $|T'_d| \gtrsim_{N,\varepsilon} |T_d|$ such that

$$|P(x)| \gtrsim_{d,\varepsilon,\deg P} |P|_{C^0(Q_w)}, \quad x \in T'_d.$$

In particular,

$$| \det \Psi^N_{x_0}(t) | \gtrsim C_{d,N,\varepsilon,b_0} \rho_0^{-d} \alpha_1^{C_\varepsilon} \alpha_1^{\deg J} \gamma_0$$

on some refinement $F'_d$ of $F_d$. We observe in addition that

$$\| \det \Phi^N_{x_0} - \det \Phi^N_{x_0} \|_{C^0(Q_w)}$$

\begin{equation}
\lesssim \| \Phi^N_{x_0} - \Phi^N_{x_0} \|_{C^0(Q_w)} \left( \| \Phi^N_{x_0} \|_{C^0(Q_w)}^{-1} + \| \Phi^N_{x_0} \|_{C^0(Q_w)} \right) \lesssim \max_i w_i^{N-1} \| \Phi^N_{x_0} \|_{C^0(Q_w)} \lesssim C_{N,d}(\varepsilon_\delta, \alpha_1)^{N-1} \cdot N. \tag{6.13}
\end{equation}

By (6.2) and the fact that $\alpha_j \leq 1$, $1 \leq j \leq k$, we may choose $N$ sufficiently large depending on $\delta, \varepsilon, b_0$, and then $c_{\delta, \varepsilon}$ sufficiently small depending on $N, \delta, \varepsilon$, to obtain

$$| \det \Psi^N_{x_0}(t) | \gtrsim C_{d,N,\varepsilon,b_0} \rho_0^{-d} \alpha_1^{C_\varepsilon} \alpha_1^{\deg J} \beta_0$$

on a refinement $F'_d$ of $F_d$, which satisfies $|F'_d| \gtrsim |F_d|$. Thus by (6.11),

$$\rho_0 \int_{F_d} | \det \Psi^N_{x_0}(t) | \, dt \gtrsim C_{d,N,\varepsilon,b_0} \rho_0^{d} \alpha_1^{C_\varepsilon} \alpha_1^{\deg J} \gamma_0 \gtrsim \alpha_1^{C_\varepsilon} \alpha_1^{b_0}. \tag{6.14}
$$

The proof will be complete once we establish a lower bound for the volume of $\Psi^N_{x_0}(F_d)$ in terms of the left hand side of (6.14). Such a lower bound would be immediate if $\Psi^N_{x_0}$ was $N$-to-one off a set of measure zero, but this condition does not follow from our hypotheses.
It follows from (6.13) and (6.10) that by taking $N$ larger and $c_{\delta, \varepsilon}$ smaller if necessary, we may assume that

$$\| \det D\Psi_{x_0}^b - \det D\Psi_{x_0}^N \|_{C^0(Q_w)} < c_{d,N}(\prod_{i=1}^d w_i) \| \det D\Psi_{x_0}^b \|_{C^0(Q_w)}.$$  

By Lemma 7.1 of [3] (the proof is also summarized in [18]) and scaling, we thus have

$$\rho(\Psi_{x_0}^b(F_t)) \sim \rho_0|\Psi_{x_0}^b(F_t)| \geq N_{d,\varepsilon} \rho_0 \int_{F_t} |\det D\Psi_{x_0}^b(t)| \, dt.$$  

This completes the proof of Proposition 6.1, and hence of Theorem 2.1.

7. Appendix: The proof of Proposition 4.1

In this section we prove Proposition 4.1, which was used in proving Propositions 2.2 and 2.3. Conclusions (i) and (ii) are actually equivalent, as shown by the following lemmas.

Lemma 7.1. If $A \subset \mathbb{Z}_0^k$ is a finite set and $b_0 \notin \mathcal{P}(A)$, 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 $p \in \mathcal{P}(A)$.

Lemma 7.2. If $v_0 \in (0, 1]^k$, there exists a finite set $A \subset \mathbb{Z}_0^k$ such that $b_0 \notin \mathcal{P}(A)$ and

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

Proof of Lemma 7.1. If $b_0 = (0, \ldots, 0)$, the result is trivial (take $v_0 = (1, \ldots, 1)$ and $\varepsilon = \frac{1}{2}$), so we may assume that $b_0 \neq (0, \ldots, 0)$. Since $b_0 \notin \mathcal{P}(A)$, $b_0$ is an extreme point of $\mathcal{P}(A \cup \{b_0\})$, so there exists $v_1 \in \mathbb{R}^k$ such that $v_1 \cdot b_0 < v_1 \cdot p$ for every $p \in \mathcal{P}(A \cup \{b_0\}) \setminus \{b_0\}$. Since $b_0 + [0, \infty)^k \subset \mathcal{P}(A \cup \{b_0\})$, $v_1 \in [0, \infty)^k$, and so, multiplying $v_1$ by a positive constant if necessary, we may assume that $v_1 \in [0, 1]^k$. Let

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

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

$$b \cdot v_2 = v_1 \cdot b_0 + v_1 \cdot (b - b_0) + \delta |b_1| \geq v_2 \cdot b_0 + \delta |b_0| \geq v_2 \cdot b_0 + \delta.$$  

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

Proof of Lemma 7.2. Let $\varepsilon = \min_i v_0^i$ and let $N = [k \varepsilon^{-1}(b_0 \cdot v_0 + 1)]$. If $p \in \mathbb{Z}_0^k$ and $|p| \geq N$,

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

so the conclusion holds with

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

The following lemma implies that the conclusions of Proposition 4.1 would hold if $\mathcal{B}$ had cardinality $\# \mathcal{B} \leq k$.

Lemma 7.3. Let $\mathcal{B} = \{b_1, \ldots, b_m\} \subset \mathbb{Z}_0^k$, where $1 \leq m \leq k$, and assume that $b_0 \notin \mathcal{P}(\mathcal{B})$. Then there exist admissible $\varepsilon > 0$, $v_0 \in (\varepsilon, 1]^k$ such that $b \cdot v_0 > b_0 \cdot v_0 + \varepsilon$ for every $p \in \mathcal{P}(\mathcal{B})$. 


Proof. We note that by Lemma 7.1 the lemma would hold if \( \{b_1, \ldots, b_m\} \) was an admissible set. We will reduce to this case.

If \(|b_i| > |b_0|_1, 1 \leq i \leq m\), the conclusion holds with \( v_0 = (1, \ldots, 1), \varepsilon = \frac{1}{2}(|b_0|_1 + 1) - 1\). Thus, reindexing if necessary, we may assume that \(|b_1|_1 \leq |b_0|_1\). Therefore \( \{b_1\} \) is admissible.

Assume that \( \{b_1, \ldots, b_j\} \) is admissible for some \( j < m \). By assumption, \( b_0 \notin \mathcal{P}(\{b_1, \ldots, b_j\}) \), so by Lemma 7.1, 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 \( 1 \leq i \leq j \). 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 \( v_j \cdot b_{j+1} \leq v_j \cdot b_0 \). Therefore \( b_{j+1} \) is admissible, and hence \( \{b_1, \ldots, b_{j+1}\} \) is admissible as well. The procedure must terminate after at most \( m \) \( (\leq k) \) steps, and so the lemma is proved.

Lemma 7.3 has the following corollary.

**Lemma 7.4.** Under the hypotheses of Lemma 7.3, there exists an admissible \( \varepsilon > 0 \) such that if

\[
\frac{b}{\theta} = \sum_{i=1}^{m} \theta_i b_i
\]

is any convex combination of \( b_1, \ldots, b_m \), there exists an \( i, 1 \leq i \leq k \) such that \( \frac{b}{\theta}^{i} (\theta) \geq \frac{b}{\theta}_0^{i} + \varepsilon \).

Proof. By Lemma 7.3, there exist admissible \( \varepsilon > 0, v_0 \in ([\varepsilon, 1]^k \) such that

\[
\varepsilon < \left( \frac{b}{\theta} - b_0 \right) \cdot v_0 \leq \left( \sum_{i=1}^{k} v_{\theta}^i \right) \max_{1 \leq i \leq k} \left( \frac{b}{\theta}^i (\theta) - \frac{b}{\theta}_0^i \right) \leq \max_{1 \leq i \leq k} \left( \frac{b}{\theta}^i (\theta) - \frac{b}{\theta}_0^i \right).
\]

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 in a moment; \( C \) will be admissible. Define \( \mathcal{A} = \mathcal{B}' \cup \mathcal{B}'' \), where

\[
\mathcal{B}' = \{ b \in \mathcal{B} : |b|_1 \leq C \}
\]

\[
\mathcal{B}'' = \{ C e_i : 1 \leq i \leq k \}
\]

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

Assume that \( b_0 \in \mathcal{P}(\mathcal{A}) \). Since \( b_0 \notin \mathcal{A} \), by standard arguments (see, for instance, \[21\]), \( b_0 \geq \sum_{i=1}^{k} \theta_i a_i \), for some \( a_1, \ldots, a_k \in \mathcal{A} \) and \( 0 \leq \theta_i \leq 1 \) satisfying \( \sum \theta_i = 1 \). Reindexing if necessary,

\[
b_0 \geq \sum_{l=1}^{j} \theta_l C e_i + \sum_{l=j+1}^{k} \theta_l b_l, \quad (7.1)
\]

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

Let \( b(\theta) = \left( \sum_{l=j+1}^{k} \theta_l \right)^{-1} \sum_{l=j+1}^{k} \theta_l b_l \).
By Lemma 7.4, there exists an $i$, $1 \leq i \leq k$ such that $b_i(\theta) \geq b_0 + \varepsilon$, where $\varepsilon > 0$ is admissible. By (7.1),

$$b_0 \geq \left( \sum_{l=j+1}^{k} \theta_j \right) b(\theta),$$

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

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

so

$$\sum_{l=1}^{j} \theta_j \geq 1 - \frac{|b_0|_{\infty}}{|b_0|_{\infty} + \varepsilon} = \frac{\varepsilon}{|b_0|_{\infty} + \varepsilon}. \quad (7.2)$$

On the other hand, comparing the $i_l$-th coordinates in (7.1), $\theta_l \leq \frac{|b_0|_{\infty}}{C}$, $1 \leq l \leq j$, so $\sum_{l=1}^{j} \theta_j \leq \frac{k |b_0|_{\infty}}{C}$. For $C = C(\varepsilon, b_0)$ sufficiently large (admissible since $\varepsilon$ is), this contradicts (7.2), and the proof of Proposition 4.1 is complete. \qed

REFERENCES


Department of Mathematics, 480 Lincoln Dr., Madison, WI 53706–1325

E-mail address: stovall@math.wisc.edu