Title

Part I: Uniform estimates for operators involving polynomial curves. Part II: Decoupling estimates for fractal and product sets.

Author

de Dios Pont, Jaume

Publication Date

2023

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Part I: Uniform estimates for operators involving polynomial curves.
Part II: Decoupling estimates for fractal and product sets.

A dissertation submitted in partial satisfaction

of the requirements for the degree

Doctor of Philosophy in Mathematics

by

Jaume de Dios Pont

2023

ABSTRACT OF THE DISSERTATION

Part I: Uniform estimates for operators involving polynomial curves.

Part II: Decoupling estimates for fractal and product sets.

by

Jaume de Dios Pont

Doctor of Philosophy in Mathematics

University of California, Los Angeles, 2023

Professor Terence Chi-Shen Tao, Chair

The first part of the thesis focuses on the uniformity of harmonic analysis estimates on curves. We first show a decomposition theorem for polynomial curves on local fields as a bounded number of perturbations of monomial curves. Using this theorem, we extend uniform restriction estimates for real curves to the endpoint case, show uniform decoupling for those curves, and show novel uniform restriction estimates for curves over $\mathbb{C}$, and $\mathbb{Q}_p$. We then show uniform estimates for the discrete analog to this problem in a restricted range of exponent.

The second part focuses on decoupling estimates for sets with a product or self-similar structure. A recurring phenomenon for those sets is that functions with constant Fourier transform on their support are far from extremizers. As applications we will show a decoupling estimate for fractal subsets of the parabola, and study subsets of cubes with high additive energy compared to their cardinality.

The dissertation of Jaume de Dios Pont is approved.

John B. Garnett

Rowan Killip

Monica Visan

Terence Chi-Shen Tao, Committee Chair

University of California, Los Angeles

2023

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

First and foremost, I am extremely grateful to Professor Tao for his invaluable guidance during the Ph.D. program, and his constant encouragement and support.

I am deeply grateful to John Garnett, whose mentorship during my bachelor's thesis was instrumental in my decision to pursue this Ph.D. in mathematics.

I spent almost a year of the Ph.D. program back at home in Terrassa, during the Covid-19 pandemic. I am extremely grateful to my my parents and my brother for making those very uncertain days much better, for making home feel like home again - and, of course, for the amazing food. I'm also thankful to Alan, Rachel, Asgar, Zane and Jose for many, many laughs over Zoom during that confusing time

I am really grateful to all the people who made my time in the department more enjoyable – and made me look forward to coming back every time I left. I am really grateful for many insightful conversations with José and Rachel (allegedly my Ph.D. advisor), and Ben and James - amongst many others - in the lounge, and, more importantly, later in the evenings at Barney's.

And, most of all, I'm extremely grateful to Irina, for being there all the times it mattered the most and managing to be there every day even from almost 2500 miles away.

Part of the work leading to this dissertation was done during the 2021 Harmonic Analysis and Analytic Number Theory trimester program and the 2022 Interactions between Geometric measure theory, Singular integrals, and PDE trimester program, both at the Hausdorff Institute of Mathematics. Part of the work was done while visiting New York University, in Summer 2019 and Spring 2023, while visiting Joan Bruna. I am grateful for the hospitality and support from these institutions.

I am grateful to Spyros Dendrinos, Zane Li, Jose Madrid, Conor Meade, Helge J. Sammuelsen, Andreas Seeger, and Sang Truong for helpful comments leading to the work that has been included in the first part of the paper. The second part of this thesis contains work that was joint with Alan Chang, Rachel Greenfeld, Paata Ivanisvili, Asgar Jamneshan and Jose Madrid. I am grateful to them, and to Tim Gowers and the anonymous referees for helpful improvements on the content of those papers.

I am grateful for the financial support through "La Caixa" Fellowship LCF/BQ/AA17/11610013, and a UCLA Dissertation year Fellowsihp during the years of the PhD.

VITA

2012–2017	B.S. (Mathematics), Universitat Autònoma de Barcelona
2012–2017	B.S. (Physics), Universitat Autònoma de Barcelona
2017–2018	M.S. (Mathematics) Eidgenössische Technische Hochschule Zürich
2018-2020	"La Caixa" fellow, University of California, Los Angeles
2020-21	Dissertation Year Fellowship, University of California, Los Angeles

PUBLICATIONS

Sinho Chewi, Jaume de Dios Pont, Jerry Li, Chen Lu, Shyam Narayanan. *Query lower bounds for log-concave sampling.* Preprint: arXiv:2304.02599.

Jaume de Dios Pont, Helge Jørgen Samuelsen. *Uniform Fourier Restriction Estimate for Simple Curves of Bounded Frequency.* Preprint: arXiv:2303.11693.

Jaume de Dios Pont, Paata Ivanisvili, José Madrid. *A new proof of the description of the convex hull of space curves with totally positive torsion.* Preprint: arXiv:2201.12932.

Jaume de Dios Pont, Rachel Greenfeld, Paata Ivanisvili, José Madrid. *Additive energies on discrete cubes To appear in* Discrete Analysis. (arXiv:2112.09352)

Alan Chang, Jaume de Dios Pont, Rachel Greenfeld, Asgar Jamneshan, Zane Kun Li, José Madrid. *Decoupling for fractal subsets of the parabola* Math. Z. 301, 1851–1879 (2022). (arXiv:2012.11458)

Jaume de Dios Pont. *A geometric lemma for complex polynomial curves with applications in Fourier restriction theory.* Preprint: arXiv:2003.14140.

Jane Carlen, Jaume de Dios Pont, Cassidy Mentus, Shyr-Shea Chang, Stephanie Wang, Mason A. Porter. *Role Detection in Bicycle-Sharing Networks Using Multilayer Stochastic Block Models.* Network Science 10.1 (2022): 46-81. (arXiv:1908.09440)

Part I

Uniformity for operators defined by polynomial curves

CHAPTER 1

Introduction

This thesis explores questions in the harmonic analysis of curves, focusing on the derivation of novel estimates for polynomial curves. The thesis extends and builds upon existing results in two different directions. The first research direction aims to understand the dependency of estimates on the specific geometry of a curve, with the primary objective being to identify large natural classes of curves where uniform estimates can be obtained. The second direction seeks to extend these findings from real curves, defined as the images of smooth maps from $\mathbb{R} \to \mathbb{R}^d$, to curves in $\mathbb{C}$ and the $p-$adic numbers $\mathbb{Q}_p$. This extension is motivated by the desire to understand estimates for two-dimensional objects, particularly over $\mathbb{C}$, or in the case of $\mathbb{Q}_p$, due to its applications to number theory and discrete harmonic analysis[42], or the simplification of specific proof techniques [34].

A crucial geometric input for obtaining uniform estimates in current results on polynomial curves is the Dendrinos-Wright decomposition [24], which provides a lower bound to certain differential forms. A substantial portion of this work is dedicated to discovering improved versions of this decomposition and generalizing it to local fields. The original proof heavily relies on the ordering of the reals, rendering it inapplicable to higher dimensions or other fields. Transitioning to the local field setup forces the usage of proofs based on compactness, which offer a more refined geometric description. This refined geometric description has applications in the real case, both by simplifying existing proofs and by establishing new uniform estimates without necessitating additional analytic insight.

This part of the thesis is structured as follows:

- In Chapter 2, we establish a decomposition theorem for polynomial curves into a controlled number of pieces. Intuitively, this allows for the decomposition ofDy known only for $\mathbb{R}^d$ and $\mathbb{C}^3$ – as well as uniform decoupling and endpoint restricted uniform restriction for curves in $\mathbb{R}^d$.

- In Chapter 4, we shift our focus to the discrete analogues of the aforementioned problems, wherein $\mathbb{K}^d$ is replaced with $\mathbb{Z}^d$, integrals with sums, and Fourier transforms with Fourier series. We establish uniform estimates for the same operators, albeit within a considerably more constrained (and non-optimal) range.

1.0.1 The operators

The operators that will concern us in this work are three: the restriction of a function's Fourier transform to a curve, the averaging of a function along a curve, and the decoupling of a function with Fourier support in the neighborhood of a curve. Each of these problems can be understood by determining the bounds of an associated operator, with the ultimate objective being to identify classes of curves for which these operators are bounded from L^p to L^q. The model example is the case of the moment curve, denoted as $\mu(t) = (t, t^2, \dots, t^d)$.

We will consider not only curves as maps $\mathbb{R} \to \mathbb{R}^d$, but polynomial curves mapping a local field $\mathbb{K}$ of characteristic zero (that is, $\mathbb{R}, \mathbb{C}$ or a finite extension of $\mathbb{Q}_p$) to $\mathbb{K}^d$. In the case where $\mathbb{K} = \mathbb{R}$, these definitions revert to the standard, well-known definitions.

The restriction operator

Definition 1.0.1 (Restriction operator)**.** *Given a curve $\gamma : \mathbb{K} \to \mathbb{K}^d$, we define the restriction operator $R_{\gamma,dt} : C_c^0(\mathbb{K}^d) \to C^0(\mathbb{K})$ as*

$$(R_{\gamma,dt} f)(t) := \hat{f}(\gamma(t)). \tag{1.1}$$

Here the sub-index (γ, dt) refers to the fact that we consider $\mathbb{K}$, the domain of γ, as a measure space with the Haar (Lebesgue) measure, and will in general refer to a measure in $\mathbb{K}$.

The model result for the restriction operator (when γ is the moment curve and $\mathbb{K} = \mathbb{R}^d$) was originally established by Drury [26] in the full range using an iterative method of *offpsring curves* (see Section 3.1), and states

Theorem 1.0.2 (Drury [26], $L^p \to L^q$ bounds, Bak-Oberlin-Seeger[3], endpoint.)**.** *Let $\mu : \mathbb{R} \to \mathbb{R}^d$, $d \geq 2$, be the moment curve $\mu(t) := (t, t^2, \ldots, t^d)$. Let $R_{\gamma,dx}$ be the Fourier restriction operator defined in 1.0.1. Then, for any pair (p, q) satisfying*

$$p' = \frac{d(d+1)}{2} q, \quad q > \frac{d^2 + d + 2}{d^2 + d}. \tag{1.2}$$

it holds that $\|R_{\gamma,dx}\|_{\mathrm{Op}\left(L^p(\mathbb{K}^d) \to L^q(\gamma; d\lambda_\gamma)\right)} < C_{p,q,d,\mathbb{K}}$. Moreover [3] , if $d \geq 3$, and $p = \frac{d^2+d+2}{d^2+d}$, the restricted endpoint estimate $\|R_{\gamma,dx}\|_{\mathrm{Op}\left(L^{p,1}(\mathbb{R}^d) \to L^p(\gamma; d\lambda_\gamma)\right)} < C_d$ holds.

This result, moreover, is sharp, in the Lorentz range as shown in [1] by studying the decay of the Fourier transform of the measure $\gamma_*(dx)$. In $d = 2$, there is no endpoint restricted-weak-type estimate due to a Kakeya type construction [6].

When establishing estimates about the restriction operator it is useful to consider its formal adjoint, the extension operator.

Definition 1.0.3 (Extension operator)**.** *Given a curve $\gamma : \mathbb{K} \to \mathbb{K}^d$, we define the restriction operator $E_{\gamma,dt} : C_c^0(\mathbb{K}) \to C^0(\mathbb{K}^d)$ as*

$$(E_{\gamma,dt} f)(x) := \int_{\mathbb{K}} e^{2\pi i x \gamma(t)} f(t) dt = \mathcal{F}^{-1}\{\gamma_*(f dt)\} \tag{1.3}$$

The sub-index (γ, dt) refers to the fact that we integrate with respect to the Haar (Lebesgue) dt measure, and $\gamma_*(f dt)$ refers to the pushforward of the $f dt$ measure by γ.

By duality, finding $L^p \to L^q$ estimates for the restriction operator is equivalent to finding $L^{q'} \to L^{p'}$ estimates for the extension operator.

The averaging operator

Definition 1.0.4 (Averaging operator)**.** *Given a curve $\gamma : \mathbb{K} \to \mathbb{K}^d$, we define the averaging operator $T_{\gamma,dt} : C_c^0(\mathbb{K}) \to C^0(\mathbb{K}^d)$ as*

$$(T_{\gamma,dt}f)(x) := \int_{\mathbb{K}} f(x - \gamma(t))dt = \gamma_*(dt) * f \tag{1.4}$$

The sub-index (γ, dt) refers to the fact that we integrate with respect to the Haar (Lebesgue) dt measure, and $\gamma_*(fdt)$ refers to the pushforward of the fdt measure by γ.

In this case, the sharp bounds were proven in [67], and state:

Theorem 1.0.5 (Stovall [67] ($d \geq 4$, endpoints), Christ [16] ($d \geq 4$, non-endpoints), Oberlin [57] ($d = 3$), Littman [53] ($d = 2$))**.** *Let $\mu : \mathbb{R} \to \mathbb{R}^d$ be the moment curve $\mu(t) := (t, t^2, \ldots, t^d)$. Let $T_{\gamma,dx}$ be the averaging operator defined in 1.0.4. Let $p_d := \frac{d+1}{2}$ and $q_d := \frac{d+1}{2}\frac{d}{d-1}$. Then, for any pair (p, q) satisfying*

$$(p, q) = \lambda(p_d, q_d) + (1 - \lambda)(q_d', p_d') \quad \lambda \in [0, 1] \tag{1.5}$$

it holds that $\|T_{\gamma,dx}\|_{\mathrm{Op}\left(L^p(\mathbb{K}^d)\to L^q(\mathbb{K}^d)\right)} < C_{p,q,d,\mathbb{K}}$.

Moreover, T_γ maps the Lorentz space $L^{p_d,u}(\mathbb{K}^d)$ boundedly into $L^{q_d,v}(\mathbb{K}^d)$ and $L^{q_d',v'}(\mathbb{K}^d)$ into $L^{p_d',u'}(\mathbb{K}^d)$ whenever $u < q_d$, $v > p_d$, and $u < v$.

This result is sharp up to Lorentz-space endpints [67]. Since the adjoint of $T_{\gamma,dt}$ is $T_{-\gamma,dt}$ and L^p bounds for both $T_{\gamma,dt}$ and $T_{-\gamma,dt}$ should be the same, $T_{\gamma,dt}$ maps L^p to L^q if and only if $T_{\gamma,dt}$ maps $L^{q'}$ to $L^{p'}$. This is reflected in (1.5), which is invariant under exchanging (p, q) with (q', p').

In this work we will not focus on extending results for this operator, but rather a certain *discrete analogue* of it. One would expect, however, that the results and proof strategies in this section would allow one to construct uniform analogues to this operator as well. For particular forms of curves over $\mathbb{C}$, using a weaker version of the main decomposition theorem of this thesis, this has been achieved by Meade [54].

Decoupling estimates

Definition 1.0.6. *Let $\mathcal{U} = \{U_1, \dots U_m\}$ be a family of open sets $\subseteq \mathbb{K}^d$. We define the decoupling constant $\mathrm{Dec}_{l^p L^q}(\mathcal{U})$ as the best constant so that the inequality*

$$\left\| \sum_{j=1}^m f_j \right\|_{L^q(\mathbb{R}^d)} \leq \mathrm{Dec}_{l^p L^q}(\mathcal{U}) \left(\sum_{j=1}^m \|f_j\|_q^p \right)^{1/p} \tag{1.6}$$

holds for all functions f_i so that $\operatorname{supp} \hat{f}_i \subseteq U_i$.

The $\mathcal{U}$ are, in practice, chosen to be neighbourhoods of geometric objects of interest. In our case we will consider

Definition 1.0.7. *Given a smooth curve $\gamma : \mathbb{R} \to \mathbb{R}^d$ we define $\mathcal{N}_\delta(\gamma) ;= \{U_\gamma(x, \delta)\}_{x \in \delta\mathbb{Z} \cap [0,1]}$ where each U_i is the parallelepiped*

$$U_\gamma(x, \delta) := \left\{ \gamma(x) + \sum_{j=1,\dots d} \delta^j \gamma^{(j)}(\alpha_j), \text{ for all } |\alpha_k| \leq 1, k = 1 \dots d \right\}$$

These parallelepipeds are *adapted neighborhoods* to the moment curve (cf [35, Section 1]). Up to a scaling factor (depending on the curve), they are equivalent to the convex hull of $\gamma([x - \delta, x + \delta])$. The celebrated theorem of Bourgain-Demeter-Guth is an almost-sharp (up to the $\epsilon-$loss) bound for the decoupling constant for a moment curve.

Theorem 1.0.8 ([12], see [35] for this formulation)**.** *Let $\mu : \mathbb{R} \to \mathbb{R}^d$ be the moment curve $\mu(t) := (t, t^2, \dots, t^d)$. Let $\mathrm{Dec}_{l^2 L^p}(N_\delta(\mu))$ be the decoupling constant defined in 1.0.6 for $N_\delta(\mu)$ defined in 1.0.7. Then, for any $p \in [2, \infty]$ and any $\epsilon > 0$ it holds that*

$$\mathrm{Dec}_{l^2 L^p}(N_\delta(\mu)) \leq C_{\epsilon,d} \delta^{-\epsilon} \left(1 + \delta^{-\frac{1}{2}\left(1 - \frac{d^2+d}{p}\right)} \right) \tag{1.7}$$

While we are describing the decoupling estimates as *estimates*, they have an associated operator, which can be used to interpolate between them:

Definition 1.0.9. *Let $\mathcal{U} := \{U_1, \dots U_m\}$ be a family of open sets $\subseteq \mathbb{K}^d$, and $\mathcal{M} := \{M_1, \dots M_m\}$ be a family of functions with $M_j|_{U_j} \equiv 1$. We define the decoupling operator $D_{\mathcal{M}} : l^p([m])L^q(\mathbb{R}^d) \to L^q(\mathbb{R}^d)$ as*

$$D_{\mathcal{M}}((f_1, \dots f_m)) := \sum_{k=1}^{m} \mathcal{F}^{-1}(M_j \hat{f}_j) \tag{1.8}$$

If each f_j has support on U_j, then $\mathcal{F}^{-1}(M_j \hat{f}_j) = f_j$. In particular one has the bound $\mathrm{Dec}_{l^p L^q}(\mathcal{U}) \leq \|D_{\mathcal{M}}\|_{\mathrm{Op}(l^p L^q \to L^q)}$. If the multiplier operators have norm $L^p \to L^p$ bounded by $\lesssim 1$, one can use interpolation theory for the decoupling operator to interpolate decoupling estimates. This cannot be done in general. If, for example, the projection operators are unbounded, or have very large norms, the decoupling operator will give no useful information about the decoupling estimate. Indeed, decoupling estimates for general open sets may not satisfy interpolation inequalities at all.

Multipliers for parallelepipeds are uniformly bounded (with bounds depending only on the dimension). In particular, one can deduce Theorem 1.0.8 from the $p = d(d+1)$ and the trivial cases $p = 2$ (Plancherel) and $p = \infty$ (Holder).

The operators described in this section can be extended, and have been extensively studied in the case when one substitutes the curve γ for a general manifold (see e.g. [20]), or other geometric objects, such as fractals. The *uniformity* question has been studied for much more restricted classes in the higher-dimensional cases as well, where the geometry becomes much more subtle [56, 50, 49]

1.0.2 Discrete analogues to the operators

The Restriction and Averaging operators have natural analogues for curves (or general subsets) of $\mathbb{Z}^d$, where integrals are exchanged for sums, and the $\mathbb{R} \to \mathbb{R}$ Fourier transform is exchanged for the $\mathbb{R} \to \mathbb{Z}$ Fourier series. In the Fourier setting, for example, one can define

Definition 1.0.10 (Discrete analogues)**.** *For $A \subseteq \mathbb{Z}^d$ define the discrete extension operator*

$E_A : \ell^1(A) \to C^0(\mathbb{T}^d)$ *by*

$$\mathrm{DE}_A(f)(x) := \sum_{n \in A} \exp(2\pi i x \cdot n) f(n) \tag{1.9}$$

which has as an adjoint the discrete restriction operator

$$\mathrm{DE}_A(f)(n) := \mathbb{1}_A(n) \int_{\mathbb{T}^d} \exp(-2\pi i x \cdot n) f(x) dx. \tag{1.10}$$

Similarly, one defines the averaging operator as

$$\mathrm{DT}_A(f)(n) := \frac{1}{|A|} \sum_{k \in A} f(n-k) \tag{1.11}$$

As a consequence of Theorem 1.0.8, by letting each $\hat{f}_i$ converge to a Dirac delta at a point and performing a suitable rescaling, one obtains a discrete restriction estimate.

Theorem 1.0.11 ([12], discrete restriction version)**.** *Let* $\mu : \mathbb{R} \to \mathbb{R}^d$ *be the moment curve* $\mu(t) := (t, t^2, \ldots, t^d)$*, and let* $\mu([N])$ *be the image of* $\{1, \ldots N\}$ *by* μ*. Then, for any* $p \in [2, \infty]$ *and any* $\epsilon > 0$ *it holds that*

$$\|E_{\mu([N])}\|_{\mathrm{Op}(\ell^2(\mu([N])) \to L^p(\mathbb{T}^d))} \leq C_{\epsilon,d} N^{\epsilon} \left(N + N^{\frac{1}{2}\left(1 - \frac{d^2+d}{p}\right)} \right) \tag{1.12}$$

This estimate was later used in [38] (together with an ϵ-removal argument for the case when $f \equiv 1$, shown in [12, Section 5]) to provide an averaging estimate for the moment curve

Corollary 1.0.12 ([38, Theorem 1.14])**.** *Let* $\mu(t) = (t, t^2, \ldots t^d)$*,* $d \geq 2$*. Let* $p^* := 2 - \frac{2}{d^2+d+1}$ *Then, for any* $p^* < p < 2$ *it holds that*

$$\| \mathrm{DT}_{\mu([N])} \|_{\mathrm{Op}(\ell^p \to \ell^{p'})} \lesssim_d N^{-\frac{d^2+d}{2}\left(\frac{1}{p} - \frac{1}{p'}\right)}. \tag{1.13}$$

The proof is a direct application Theorem 1.0.11 using the product-convolution law for the Fourier transform by bounding

$$\|f * \mathbb{1}_{\mu([N])}\|_{\ell^{p'}} = \|\hat{f}\widehat{\mathbb{1}_{\mu([N])}}\|_{L^p(\mathbb{T}^d)} \leq \|\hat{f}\|_{L^{p'}(\mathbb{T}^d)} \|\widehat{\mathbb{1}_{\mu([N])}}\|_{L^{\frac{p}{2-p}}(\mathbb{T}^d)} \leq \|f\|_{\ell^p(\mathbb{Z}^d)} \|\widehat{\mathbb{1}_{\mu([N])}}\|_{L^{\frac{p}{2-p}}(\mathbb{T}^d)}$$

and then applying Theorem 1.0.11 (with an $\epsilon-$removal) to estimate the $\widehat{\mathbb{1}_{\mu([N])}} = E_{\mu([N])}(\mathbb{1}_{\mu([N])})$ term. For $p < p^* < 2$ this estimate is sharp, as seen by taking $f = \mathbb{1}_{[-N,N]\times[-N^2,N^2]\times\cdots\times[-N^d,N^d]}$. Outside of this range, the results obtained by interpolating with elementary (Holder-type) estimates do not coincide with the lower bounds arising from $f = \mathbb{1}_{[-N,N]\times[-N^2,N^2]\times\cdots\times[-N^d,N^d]}$.

A significant difference between the continuous statements and their discrete counterparts arises when one considers other polynomial curves different than the moment curve. In the continuous case, near every non-degenerate point, every curve behaves like a moment curve after an affine transformation. This, plus a dilation-invariance argument (Section 1.1.1), shows that one cannot expect better estimates for a generic curve than for the moment curve. This does not hold in the discrete case, where one can obtain better estimates for curves with higher degrees [21, 44].

1.1 Uniform estimates

The theorems shown in the previous section considered multiple operators that could be studied for general curves, in the particular case of the moment curve. Considerable work since then has been devoted to generalizing these results for larger families of curves. A first, natural class is that of curves defined on an interval $[a, b]$ that are quantitatively close (in, say, the C^∞ topology). Theorems 1.0.2, 1.0.5, 1.2.5 extend to these classes with a slight modification of the original proofs. In fact, considering such a larger family of curves can be a necessary part of the proof, such as in proofs of the decoupling theorem [12, 35].

The situation is markedly different for curves with vanishing torsion (such as the curve (t, t^3) near $t = 0$). The torsion for the moment curve is constant, and these curves cannot be studied as perturbations of the moment curve. Not only that, but the naive generalization of the theorems for the moment curve does not hold. There is a family of counterexamples, known as Knapp-type examples (studied in its greater generality in [58]), that impose lower bounds to the curvature of the curve.

1.1.1 Curvature and Knapp examples

Let $\gamma : \mathbb{R} \to \mathbb{R}^d$ be a C^d monomial curve of the form $\gamma(t) := (t^{a_1}, \dots, t^{a_d})$, with $a_{i+1} > a_i$ integers. By a Taylor expansion, $\gamma(t)$ is, after an affine transformation, the leading order approximation to any curve analytic at any point is of this form. The torsion of this curve is

$$\tau_\gamma(t) := \det(\gamma'(t), \dots \gamma^{(d)}(t)) = C_{a_1,\dots,a_d} t^{\sum_{i=1}^d a_i - \frac{d^2+d}{2}}$$

which vanishes at the origin. Consider, for $\delta > 0$, the set

$$K_\delta := [-\delta, \delta]^{a_1} \times \cdots \times [-\delta, \delta]^{a_d}.$$

This set is the parallelepiped *most adapted* to $\gamma([-\delta, \delta])$, in the sense that every parallelepiped, or convex set, containing $\gamma([-\delta, \delta])$ will have volume $\gtrsim$ that of K_δ.

If $f(x) = \mathbb{1}_{K_\delta^*}$, the characteristic function of the dual (polar) set of $K_{\delta/10}$, then, for any $w \in K_\delta$, the integrand of $\int_{\mathbb{R}^d} \mathbb{1}_{K_\delta^*} e^{-iwx} dx$ will have no cancellation, and in particular, if $|t| < \delta$, $\gamma(t) \in K_\delta$ and,

$$|\widehat{\mathbb{1}_{K_\delta^*}}(\gamma(t))| \approx \int_{\mathbb{R}^d} \mathbb{1}_{K_\delta^*} dx \approx \delta^{-\sum \alpha_i}.$$

Let μ_γ be a measure supported on γ. One gets two inequalities:

$$\|\widehat{\mathbb{1}_{K_\delta^*}}(w)\|_q \geq \left(\int_{-\delta}^{\delta} |\widehat{\mathbb{1}_{K_\delta^*}}(w)|^q d\mu_\gamma(t) \right)^{1/q} \gtrsim_d \mu_\gamma([-\delta, \delta])^{1/q} \delta^{-\sum_{i=1}^d \alpha_i}$$

and

$$\|\mathbb{1}_{K_\delta^*}\|_p \approx_d \delta^{-\frac{1}{p} \sum_{i=1}^d \alpha_i}$$

If one wants to find a measure μ_γ so that the restriction estimate $\|R_{\gamma, d\mu_\gamma}\|_{\mathrm{Op}\left(L^p(\mathbb{K}^d) \to L^q(\gamma; d\mu_\gamma)\right)}$ holds, this imposes the constraint

$$\mu([-\delta, \delta]) \lesssim \delta^{\frac{q}{p'} \sum_{i=1}^d \alpha_i}. \tag{1.14}$$

The restriction theorem of Drury (Theorem 1.0.2) holds in the regime $\frac{q}{p'} = \frac{2}{d(d+1)}$, and is sharp, essentially by the example just shown, for a nondegenerate curve. If one wants such a theorem to hold in a degenerate case, when $\sum_{i=1}^{d} \alpha_i > \frac{d^2+d}{2}$, that forces $\mu([-\delta,\delta])$ go to zero faster than $|\delta|$. This motivates the definition of weights $w_\gamma(t)$ so that the measure $w_\gamma(t)dt$ satisfies (1.14). A particular choice of weights that makes (1.14) hold, and in fact with an approximate equality, is that of the affine arclength measure.

1.1.2 The affine arclength measure

In order to prove uniform boundedness for operators and overcome the Knapp condition, one must endow the curves with a measure that is related to the curvature (or torsion in higher dimensions), with a degree of vanishing at degenerate points preventing the Knapp family of counter-examples. A natural choice is the *affine arclength measure*. This work will rely on the natural generalization of the concept to characteristic zero local fields, inspired by the definition for the real numbers.

1.1.2.1 Real affine arclength measure

For a real C^d curve $\gamma : [0,1] \mapsto \mathbb{R}^d$ we define the real affine measure of $\gamma(t)$ as a weighted pushworward of the dt measure:

$$d\lambda_\gamma := \frac{1}{d!}\gamma_* \left(|\det[\gamma'(t), \gamma''(t), \ldots, \gamma^{(d)}(t)]|^{\frac{2}{d^2+d}} dt \right). \tag{1.15}$$

Equivalently by duality, for a function g in $C_0(\mathbb{R}^d;\mathbb{R})$,

$$\int_{\mathbb{R}^d} g(x) d\lambda_\gamma(x) := \int_0^1 g(\gamma(t)) \left|\det[\gamma'(t), \gamma''(t), \ldots, \gamma^{(d)}(t)]\right|^{\frac{2}{d^2+d}} dt \tag{1.16}$$

The potentially suitability of this measure (which vanishes at all the points where the torsion $\det[\gamma'(z), \gamma''(z), \ldots, \gamma^{(d)}(z)]$ vanishes) to control the potential singularities of a restriction estimate was considered as early as in [27], and this is the measure used in, e.g. the main theorem of [68]. There are several properties of this measure that make it particularly suitable:

1. It is parametrization invariant: if $\phi : [0,1] \to [0,1]$ is a diffeomorphism, then $\lambda_\gamma = \lambda_{\gamma\circ\phi}$. This follows from an application of the chain rule, and the fact that for a reparametrization $\phi : \mathbb{R} \to R$,

$$|\det[(\gamma\circ\phi)'(t), \ldots, (\gamma\circ\phi)^{(d)}(t)]| = |\phi'(t)|^{\frac{d^2+d}{2}} |\det[\gamma'(\phi(t)), \ldots, \gamma^{(d)}(\phi(t))]|.$$

   This is the motivation for the choice of the $\frac{2}{d^2+d}$ exponent

2. It is translation covariant: Let $\tau_z : \mathbb{R}^d \to \mathbb{R}^d$ be the translation $\tau_z(x) = x + z$, then $\lambda_{\tau_z\circ\gamma} = (\tau_z)_*\lambda_\gamma$.

3. It is $SU(n)$-covariant: If A is an element of $SU(N)$, then $\lambda_{A\circ\gamma} = A_*\lambda_\gamma$. If A is not an element of $SU(N)$, a power of $|\det A|^{2/(d^2+d)}$is gained.

4. The measure λ_γ (or more precisely its Radon derivative $\frac{d\lambda_\gamma}{d\gamma}$ with respect to the arclength $d\gamma$) vanishes at all the points where the torsion vanishes. As we have seen, this is particularly relevant because restriction and convolution theorems in the full range fail if one uses the Hausdorff measure at the neighborhood of a point where the torsion of γ vanishes. The degree of vanishing, moreover, makes (1.14) an approximate equality.

Weighting with the affine arclength measure, however, is not a sufficient condition, as shown by Sjolin's example, $(t, \sin(t^{-1})\exp(-t^{-1}))$, for $t \in (0,1)$ [64]. This rules out, for example, potential candidate families for *uniform estimates*, such as the class of all polynomial curves: If one had a degree-independent restriction estimate for polynomial curves, by an approximation argument, one would be able to show a restriction estimate for Drury's counterexample

1.1.2.2 Complex affine arclength measure

The affine measure has a natural complex analogue, which has previously been defined in the literature, and used e.g. by Ham-Bak in [2] to prove non-uniform (local) restriction estimates for certain complex polynomial curves, and later by Ham and Chung [18] to prove uniform estimates for the same class of curves.

Inspired by the real affine arclength measure, the affine arclength measure associated with a $d-$dimensional complex analytic curve γ defined on an open set $D \subseteq \mathbb{C}$ (i.e $\gamma(z) : D \to \mathbb{C}^d$) is defined as the push-forward of the Lebesgue measure weighed by a power ($\frac{4}{d^2+d}$) of the torsion of the curve:

$$d\lambda_\gamma = \frac{1}{d!}\gamma_* \left(\det[\gamma'(z), \gamma''(z), \ldots, \gamma^{(d)}(z)]^{\frac{4}{d^2+d}} |dz| \right) \tag{1.17}$$

The properties that were outlined in the real set-up extend to the complex case with the following minor modifications:

1. The measure λ_γ is covariant both under local re-parametrization of z (i.e., if $\phi : D' \to D$ is a conformal map, $\lambda_\gamma = \lambda_{\gamma\circ\phi}$).The factor 4 in the exponent $\frac{4}{d^2+d}$ in the definition (1.17) (in comparison with the factor of 2 in (1.15)) must be introduced to ensure reparametrization still holds.

2. The measure λ_γ is covariant under unitary maps applied on $\mathbb{C}^d$ (i.e. if $L \in SU(\mathbb{C}; d)$, then $d\lambda_{L\circ\gamma} = L_* d\lambda_\gamma$).

Bak and Ham [2] show that this measure is optimal for the Fourier restriction problem, in the sense that any measure supported on the image of γ for which Theorem 1.2.4 below holds in its full range of exponents must be absolutely continuous with respect to $d\lambda_\gamma$.

1.1.2.3 Affine arclength measure associated to a local field

The definition for a general local field is analogous to that of the real and complex case, generalizing the factor of 2 (when $\mathbb{K} = \mathbb{R}$) or 4 (when $\mathbb{K} = \mathbb{C}$) to twice the doubling exponent of the field. Let $\mathbb{K}$ be a local field, and let $m_{\mathbb{K}}$ be the Haar measure associated with $(\mathbb{K}, +)$. Define the doubling exponent of a measure on a metric space as

$$d_{\mathbb{K}} := \lim_{r \to 0} \frac{\log(m_{\mathbb{K}}(B_r))}{\log(r)},$$

where $B_r(0)$ denotes a ball of radius r (by translation invariance of the Haar measure, all balls are equivalent), whenever the limit exists. The doubling exponent of the real numbers is equal to one, the one of the complex to two, the one of any $p-$adic field is equal to 1 as well, and the one of any $p-$adic finite degree extension is equal to the degree of the extension. With this definition in hand, our definition of the affine-invariant measure for a $\mathcal{C}^d(\mathbb{K})$ curve is:

$$d\lambda_\gamma = \frac{1}{d!}\gamma_* \left(\det[\gamma'(z), \gamma''(z), \ldots, \gamma^{(d)}(z)]^{\frac{2d_{\mathbb{K}}}{d^2+d}} |dm_{\mathbb{K}}| \right) \tag{1.18}$$

This definition is compatible with the definitions over $\mathbb{R}, \mathbb{K}$ that we have already given, and, again, the $d_{\mathbb{K}}$ factor ensures the (local) reparametrization invariance. The other properties described above ($SU(n)$ and translation covariance, and vanishing when there is no curvature) transfer as well to the general field case.

1.2 Contributions and comparison to previous work

The key contribution of this part is a geometric decomposition theorem for polynomials, which will be shown in Section 2.1. Informally, when reduced to the case $\mathbb{K} = \mathbb{R}$ it states that

Theorem 1.2.1 (Simplified version of Theorem 2.1.13 for $\mathbb{K} = \mathbb{R}$). *Let $J \subseteq \mathbb{R}^d$ be a finite interval, and $\epsilon > 0$, let γ be a polynomial curve of degree $\leq N$. Then there exists*

1. *A partition* $J = I_1 \cup \cdots \cup I_m$ *(with* $m \leq M(N, \epsilon)$*) of* J *into nonoverlapping intervals.*

2. *A family of nondegenerate monomial curves* $\mu_1, \ldots, \mu_m$ *of degree* $\leq N$*.*

3. *A family of affine maps* $A_1, \ldots, A_m$*,* $A_j : \mathbb{R}^d \to \mathbb{R}^d$*, and affine bijections* $b_j : [\delta_j, 1] \to I_i$*, with* $0 \leq \delta_j < \frac{1}{2}$

so that for point in the interval interval $[\delta_j/2, 2]$ *it holds that*

$$|(A_j \circ \gamma \circ b_j)_k - (\mu_j)_k| \leq \epsilon |(\mu_j)_k|. \tag{1.19}$$

where for a curve γ*,* $(\cdot)_k$ *represents the* $k-th$ *coordinate.*

In the statement of the theorem, J is only finite for simplicity of the proof, as it avoids dealing with unbounded intervals (see Theorem 2.1.13). The implied bounds do not depend on J. This is enough for most applications after a limiting argument. The estimate (1.19) implies estimates that one could expect, such as $|\gamma^{(d)}|$ being comparable to $|\mu_i^{(d)}|$, but also gives bounds in the spirit of (1.19) for multilinear forms with a significant amount of cancellation, such as $\det(\gamma'(x_1), \ldots \gamma'(x_n))$. This will be the content of Theorem 2.4.5.

Theorem 1.2.2 (Simplified version of Theorem 2.4.5 for $\mathbb{K} = \mathbb{R}$)**.** *Let* $\epsilon > 0$*, let* $N, d > 1$*. There is* $\epsilon_D = \epsilon_D(N, d) > 0$ *so that the following holds. Let* $\gamma : \mathbb{R} \to \mathbb{R}^n$ *be a polynomial curve of degree* N*, and* μ *a monomial curve so that in the interval* $[\delta_j, 2]$*, the inequality*

$$|(\gamma_k) - (\mu)_k| \leq \epsilon_D |(\mu)_k| \tag{1.20}$$

holds for all $k = 1, \ldots, d$*. Then, for* $1 \leq k \leq d$*, and* $t_1, \ldots t_k \in [\delta_j, 2]$*, it holds that*

$$\|\gamma'(t_1) \wedge \cdots \wedge \gamma'(t_k) - \mu'(t_1) \wedge \cdots \wedge \mu'(t_k)\| \leq \epsilon \|\mu'(t_1) \wedge \cdots \wedge \mu'(t_k)\|. \tag{1.21}$$

As a particular application of Theorem 2.4.5, when $\mathbb{K} = \mathbb{R}$ one can extend an inequality and decomposition of Dendrinos-Wright [24] to local fields, by proving it explicitly for monomial curves, and using Eq. (1.21).

Lemma 1.2.3 (Dendrinos-Wright, proven in [24], see Proposition 2.4.7)**.** *Let γ be a polynomial curve of degree N, then, and assume that $\Lambda^{(d)}[\gamma](x) := \det(\gamma'(t), \gamma''(t), \ldots, \gamma^{(d)}(t)) \not\equiv 0$. Then one can decompose $\mathbb{R}$ as a disjoint union of intervals*

$$\mathbb{R} = \bigcup_{j=1}^{M_{N,d}} I_j$$

so that for $t \in I_j$

$$|\gamma'(t)| \sim_{N,d} A_j |t - b_j|^{\alpha_j}, \qquad |L_\gamma(t)| \sim_{N,d} A_j |t - b_j|^{\alpha_j}$$

and for $t_1, \ldots, t_d \in I_j$,

$$|\det(\gamma'(t_1), \ldots, \gamma'(t_d))| \gtrsim_{N,d} \prod_{0<i<d} |L_\gamma(t_i)|^{1/d} \prod_{0<j<i} |t_i - t_j|$$

The Dendrinos-Wright decomposition has been a key geometric input to most of the theory for uniform estimates for general polynomial curves (see e.g [23, 68, 54]). Theorem 2.1.13. Theorems 2.1.13 and 2.4.5 not only generalize it to complex and $p-$adic numbers, but give much finer control in the decomposition. In particular, it has consequences even for the $\mathbb{K} = \mathbb{R}$ case. In Section 3.2 we will use it to prove a uniform restricted endpoint estimate for polynomial curves. The proof is essentially a consequence of the geometric decomposition and follow from estimates shown for proof for monomial curves in [5].

The original proof of the Dendrions-Wright inequality relies strongly on the fact that $\mathbb{R}$ is an ordered field, and is proven by induction using a series of iterated integrals that allow one to compute $\det(\gamma'(t_1), \ldots, \gamma'(t_d))$ as a series of integrals of functions of the torsion. The decomposition is then chosen so that all those integrals have constant signs. This makes it particularly hard to extend that approach to a higher dimensional case, or a case with other fields, where one does not have access to a sign.

1.2.1 Proof strategy of the geometric decomposition theorem

The geometric decomposition theorem will be proven in three steps:

Reduction to a one dimensional problem:

Let $\tilde{\tau}$ be the torsion map, mapping polynomial curves of degree up to N, up to affine transformations, to polynomials of degree up to $(N-d)^d$, up to a multiplicative constant:

$$\tilde{\tau} : \mathrm{Aff}(\mathbb{K}^d) \Big\backslash (\mathbb{K}[x]_{\leq N})^d \to \mathbb{K} \Big\backslash \mathbb{K}[x]_{\leq (N-d)^d} \tag{1.22}$$

Where both quotients act on the left by multiplication. A careful analysis shows that the domain and the target of $\tilde{\tau}$ have the same dimension. The key insight of the dimensional reduction is that for polynomial curves for which $\tilde{\tau}\gamma \neq 0$ (i.e. nondegenerate curves), the map $\tilde{\tau}$ is almost injective: The for a fixed polynomial $p(x)$, the number of polynomial curves γ that satisfy to $\tilde{\tau}(\gamma) = p$, is finite. In Section 2.2 we will show a stable version of this result: If $\tilde{\tau}(\gamma)$ is very close to a monomial on a large set, γ must be, after an affine transformation, very close to a monomial curve on a slightly smaller set. This reduces the problem of finding decomposition for γ to that of finding a decomposition for $\tau(\gamma)$, which is a one dimensional curve.

This phenomeon is unique to polynomial curves: for smooth curves, one can generate infinitely many curves with the same torsion, which behave quite differently.

Factorizing the torsion to find the one-dimensional decomposition

The next step is to find a partition for one dimensional polynomials into subsets where, after a possible translation and rescaling, they behave as monomials (in the sense that there is a power n so that $|p(x) - x^n| < \epsilon x^n$). We do that by lifting to a splitting field of $p(x)$, and using the factorization of $p(x)$. We do that by showing there is constant $C = C(\epsilon)$ such that, if all the zeros of $p(x)$ have norm less than $C^{-1}r$ or more than CR, then $p(x)$ behaves like a monomial in the set of points $r \leq |x| \leq R$. Applying this argument around *clusters* of zeros of $p(x)$ shows the result.

Transferring Jacobian inequalities from monomial curves to their perturbations

The motivation of the geometric decomposition is to translate the original proofs of boundedness from monomials to general polynomials uniformly, by reducing to the case of perturbations of monomial proofs. In those proofs, one makes extensive use of quantities with a significant amount of cancellation, such as $\det(\gamma'(t_1), \ldots, \gamma'(t_d))$. Sections 2.4 to 2.6 are devoted to showing the transferring theorem. The theorem will be proven again by compactness: The set of curves that are similar to a given curve on an interval are a compact set. The main new ingredient will be a series of *no-cancellation* properties arising from the explicit form of wedge products for monomial curves. For monomial curves, the polynomials

$$\frac{1}{v(t_1, \ldots t_d)} \det(\mu'(t_1), \ldots, \mu'(t_k)) \tag{1.23}$$

are Schur polynomials, which are sums of monomials with nonnegative integer coefficients. This will prevent quotients of the form "0/0" from arising when studying terms of the form

$$\frac{\det(\gamma'(t_1), \ldots, \gamma'(t_k))}{\det(\mu'(t_1), \ldots, \mu'(t_k))} \tag{1.24}$$

unless some of the t_k go to zero. The case when some of the t_k go to zero can be treated by induction using a transversality condition that will show that if $t_1, \ldots, t_j$ are much closer to zero than $t_{j+1}, \ldots, t_k$ then

$$\frac{|\det(\mu'(t_1), \ldots, \mu'(t_k))|}{|v(t_1, \ldots t_d)|} \approx \frac{\|\mu'(t_1) \wedge \cdots \wedge \mu'(t_j))\|}{|v(t_1, \ldots t_j)|} \cdot \frac{\|\mu'(t_{j+1}) \wedge \cdots \wedge \mu'(t_k))\|}{|v(t_{j+1}, \ldots t_k)|} \tag{1.25}$$

1.2.2 Applications of the geometric decomposition theorem

Once the decomposition result in Theorem 2.1.13 is proven, the general proof strategy of uniform boundedness for an operator T_γ (restriction, convolution, decoupling...) is as follows:

1. Using Theorem 2.1.13, decompose $\gamma = \bigsqcup_{j \in J} \gamma_j$, where γ_j is the restriction of γ to a subset A_j, where it is similar to an affine translation of monomial curve. Since, by Theorem 2.1.13 $|J| \leq M_{d,n,\mathbb{K}}$. We can decompose the operator T_γ as

$$\|T_\gamma\|_{L^p\to L^q} = \left\|\sum_{j\in J} T_{\gamma_j}\right\|_{L^p\to L^q} \leq M_{d,n,\mathbb{K}} \max_{j\in J} \|T_{\gamma_j}\|_{L^p\to L^q}.$$

We will refer to the uniform bounds to $\|T_{\gamma_j}\|_{L^p\to L^q}$, where γ_j is similar to a monomial curve as *perturbative estimates near a monomial curve.* In some situations such as in averaging [66] or endpoint Fourier restriction [5], one is able to show these estimates directly. If that is not the case, one can perform further reductions:

2. After an affine transformation, assume that γ_j is $\epsilon-$similar to a monomial curve. One can decompose $\gamma_j = \bigcup_{k\in K} \gamma_{j,k}$ by restricting into dyadic scales. Using an (operator-specific) Littlewood–Paley type estimate, one shows that

$$\|T_{\gamma_j}\|_{L^p\to L^q} \leq \tilde{C}_{d,n,\mathbb{K}} \max_{k\in K} \|T_{\gamma_{jk}}\|_{L^p\to L^q}.$$

Such an estimate requires some sort of *global transversality conditions* to hold for γ_j, which are deduced from the decomposition theorem.

3. A monomial curve on a dyadic interval can be split into finitely many pieces that are $\epsilon-$similar to a monomial curve (up to an affine transformation). This fact can be uniformly transferred (in Lemma 2.7.2) to curves that are $\epsilon-$similar to a moment curve. That means, one can split γ_{jk} into a union of γ_{jkl} on smaller intervals, each of which is $\epsilon-$similar to a standard moment curve. After affine rescaling, this reduces the bounds of $\|T_\gamma\|_{L^p\to L^q}$ to uniform bounds to

$$\|T_{\tilde{\mu}}\|_{L^p\to L^q}$$

which are uniform over all perturbations ($\tilde{\mu}$) of the standard moment curve the unit ball.

The decomposition result of Theorem 2.1.13, combined with the strategy above will allow us to deduce the following global uniform estimates from their perturbative analogues for monomial curves.

Restriction estimates

Theorem 1.2.4. *Let $\gamma : \mathbb{K} \to \mathbb{K}^d$ be a polynomial of degree at most N in $\mathbb{K}^d$. Let R_γ be the Fourier restriction operator defined in 1.0.1. Then, for any pair (p, q) satisfying*

$$p' = \frac{d(d+1)}{2}q, \quad q > \frac{d^2+d+2}{d^2+d} =: p_d. \tag{1.26}$$

it holds that $\|R_\gamma\|_{Op\left(L^p(\mathbb{K}^d) \to L^q(\gamma; d\lambda_\gamma)\right)} < C_{p,q,N,\mathbb{K}}$. Moreover if $\mathbb{K} = \mathbb{R}$ and $d \geq 3$, the estimate $\|R_{\gamma,dx}\|_{\mathrm{Op}\left(L^{p_d,1}(\mathbb{K}^d) \to L^{p_d}(\gamma; d\lambda_\gamma)\right)} < C_{d,N}$ holds.

The range of (p, q) is known to be the best attainable, at least when $\mathbb{K} = \mathbb{R}$, and when $\mathbb{K} = \mathbb{C}$ [74].

The first uniform $L^p \to L^q$ restriction estimate for curves can be traced back to Sjölin, who showed an estimate of the form of Theorem 1.2.4 for C^2 convex curves in $\mathbb{R}^2$. The first results in higher dimensions were due to Prestini [60], and Christ [17], who showed the first results in the case of degenerate curves. The full range of exponents was shown by Drury [26] with a proof that extends to curves that are sufficiently close to a moment curve in the C^{d+1} topology. Sjölin's result implies Definition 1.0.1 in the case $p = 2$. In higher dimensions, partial progress was made in [4, 3, 5, 24] for restricted classes of functions. Using the geometric lemma proven by Dendrinos-Wright, Stovall extended this result in the case $\mathbb{K} = \mathbb{R}$ to the whole non-endpoint range.

In the case for local fields, for $\mathbb{K} = \mathbb{C}$, Bak and Ham [2] considered the moment curve, as well as curves of the form $(z, z^2, \phi(z))$. The general (non-endpoint) three-dimensional case was settled by Meade [54]. The situation for general local fields, in the case of the moment curve, was studied by Hickman [42] (see also [41]). The proof of Theorem 1.2.4 follows the strategy of the original proof by Stoall, which only needs a *local* version of the theorem as an input, essentially proven in [42], once one has the right polynomial decomposition.

The endpoint estimate was shown by Bak-Oberlin-Seeger [3], who then generalized it [5] to certain *simple* cases of monomial curves, such as curves of the form $(t, t^2, \ldots, t^{d-1}, p(t))$, or general monomial curves. The result in [5], in fact, together with the geometric decomposition theorem, essentially implies the endpoint part of Theorem 1.2.4, using ideas from [27].

Decoupling estimates

Theorem 1.2.5. *Let $\gamma : \mathbb{R} \to \mathbb{R}^d$ be a polynomial of degree at most N in $\mathbb{R}^d$. Let $\mathcal{N}_\delta(\gamma)$ be a union of δ-parallelepipeds adapted to $\gamma([0,1])$ (in the sense of Definition 1.0.7). Let $\mathrm{Dec}_{l^2L^p}(U_\gamma)$ be the decoupling constant associated with $U_\gamma(\delta)$.*

Then, for any $\epsilon > 0, p \in [2,\infty]$ it holds that

$$\mathrm{Dec}_{l^2L^p}(U_\gamma) \leq C_{\epsilon,p,d,N}|U_\gamma|$$

where $|U_\gamma|$ denotes the number of elements (parallelepipeds) in U_γ.

The uniformity theory for decoupling estimates is significantly less developed. A result to highlight is that of Yang [75], which establishes a variation of Theorem 1.2.5 under a slightly different partition. The proof of Theorem 1.2.5 is very robust, as characteristic of Decoupling estimates, and in particular, recovers the main result of [75] (see Section 3.3).

1.2.3 Discrete analogues

We will show, for a restricted range, uniform estimates for discrete analogues to two of the problems considered in the previous section, namely

Theorem 1.2.6. *Let $\mathrm{DE}_{\gamma([N])}$ be the discrete extension operator associated to a curve $\gamma(t) \in \mathbb{Z}[x]^n$ of degrees $d = (d_1, \dots d_n)$ (with $d_{i+1} > d_i$, and $d_n \geq 2$) (cf Definition 1.0.10). Let $p_0 := d_n^2 + d_n$. Then, for any $p_0 \leq p < \infty$ it holds that:*

$$\| \mathrm{DE}_{\gamma([N])} \|_{\mathrm{Op}(l^2(\mathbb{Z}^d)\to L^p)} \leq C_{d,p,\epsilon} N^{-\left(\frac{1}{2}-\frac{|d|}{p}\right))+\epsilon}$$

where $|d| = \sum_{i=1}^n d_i$. Moreover, this result is sharp for $p_0 < p < \infty$.

Theorem 1.2.7. *If* $\mathrm{DT}_{\gamma[N]}$ *is the discrete averaging operator associated to a curve* $\gamma \in \mathbb{Z}[x]^d$ *of degrees* $d = (d_1, \dots d_n)$ *(with* $d_i < d_{i+1}$*, and* $d_n \geq 2$*), let* $p_0 := 2 - \frac{2}{d_n^2 + d_n + 1}$*. Then, for any* $p_0 < p < 2$ *it holds that:*

$$\| \mathrm{DT}_{\gamma([])} \|_{\mathrm{Op}(l^p \to l^{p'})} \leq C_{d,p} N^{-|d|(\frac{1}{p} - \frac{1}{p'})}$$

where $|d| = \sum d_i$*.*

The main ingredient for both results will be the discrete restriction for the moment curve arising from the Decoupling theorem of Bourgain, Demeter and Guth (Theorem 1.0.11). We will *project down* this theorem to lower dimensions to prove discrete analogues to both the averaging and Fourier restriction questions.

This is not a new technique, even for the theorems at hand. Theorem 1.0.11 was used in [48] to prove a discrete restriction estimate for monomial curves (essentially Theorem 1.2.6 restricted to monomial curves), and in [38] to prove certain l^p-improving averages (essentially Theorem 1.2.7 restricted to 1−dimensional polynomials, with a constant depending on the polynomial). The use of (variations of) this technique in number theory is well known, and dates as back as [43]. The interest of the results above lies in the fact that the results are uniform over the polynomials of a given degree. In the averaging case, whether the results held uniformly was posed in [38].

The main drawback of this technique of *projection to lower dimensions* is that one does not expect to obtain sharp ranges for p from it, even after interpolation. In other words, while the power loss in N is sharp for the p in the given range, the range of p in the Theorem 1.2.6 not sharp but arises from the limitation of the proof. In the averaging case, for example, estimates that cannot be deduced from Theorem 1.2.7 can be found in [22], and in [44] for a particular case of the restriction statement, the case (t, t^3).

CHAPTER 2

A geometric decomposition for polynomial curves

2.1 Preliminary definitions and the decomposition theorem

2.1.1 Geometry of the decomposition

In the applications it is necessary to have some control over the shape of the sets forming the partition of the polynomial curve. This does not pose a significant challenge in the case $\mathbb{K} = \mathbb{R}$, in which all of the sets are intervals. In the general field case, the general sets will be constructed as the intersection of two base sets: annuli and sectors.

We start with the definition of sectors. Let K_+ be $\mathbb{R}_{>0}$ if $\mathbb{K} = \mathbb{R}, \mathbb{C}$, and $K_+ := p^{\mathbb{Z}}$ if $\mathbb{K} \geq \mathbb{Q}_p$.

Definition 2.1.1. *Given a locally compact field $\mathbb{K}$, and $K \subseteq \mathbb{K}$ the closure of $\mathbb{Q}$ in $\mathbb{K}$, we define the sector $\Sigma_\epsilon^{\mathbb{K}}$ of amplitude $\epsilon \in \mathbb{R}$ centered around 1 as the set*

$$\Sigma_\epsilon^{\mathbb{K}} := \{x \in \mathbb{K} : d(x, K_+) < \epsilon |x|\}.$$

Whenever $\mathbb{K}$ is clear by the context we will write Σ_ϵ. For an element $t \in \mathbb{K} \setminus \{0\}$ we will denote by $t\Sigma_\epsilon$ the set $\{tx, x \in \Sigma_\epsilon\} = \{x : d(t^{-1}x, K_+) < |t|^{-1}|x|\}$. We will call $t\Sigma_\epsilon$ a sector of amplitude ϵ in the direction of t.

In the case $\mathbb{K} = \mathbb{C}$ sectors correspond to the common angular sectors (see Figure 2.1). This definition is motivated by the fact that within a sector there is an approximate reverse triangle inequality: The norm of the sum of n elements belonging to the same sector is comparable to the sum of the norms (see Appendix A.1 for a more precise statement).

Another quantitiy to be controlled is the norm (or distance to a point), both from above and from below. We will do so by restricting to annuli, defined on a field as follows:

Definition 2.1.2. *An annulus with center* $z_0 \in \mathbb{K}$ *and outer and inner radius* $0 \leq r < R \leq \infty$ *will be denoted by* $A^{\mathbb{K}}_{r,R}(z_0) := \{z \in \mathbb{K}, |z - z_0| \in (r, R)\}$. *When* $\mathbb{K}$ *is clear from the context,the field* $\mathbb{K}$ *will be dropped and the annulus will be denoted by* $A_{r,R}(z_0)$.

We will use the term *truncated sector* to refer to sets A which are the intersection of a sector and an annuli, as well as translations of those sets. The amplitude of a truncated sector will be the amplitude of the associated sector.

2.1.2 Zoom-ins, Canonical Forms, ϵ-similarity

A type of curve playing a key role in the decomposition theorem is that of a monomial curve.

Definition 2.1.3. *For* $\mathbf{n} = (n_1, \ldots, n_d)$, *with* $1 \geq n_i < n_{i-1}$, *we define the monomial curve* $\mu_{\mathbf{n}}$ *to be the curve of the form* $\mu_{\mathbf{n}} : (t) = (t^{n_1}, \ldots t^{n_d})$.

The relevance of these curves is twofold: On one hand, they are the easiest model case to study. Most importantly, we will show that any polynomial curve can be uniformly approximated piecewise by affine transformations of generalized moment curves.

The proof of the main geometric result (Theorem 2.1.13) is based in understanding different regions of the curve γ by studying the different affine transformations of γ. In order to do so we will associate to each polynomial curve $\gamma : \mathbb{K} \to \mathbb{K}^d$ of degree at most N a $d \times N$ matrix $M[\gamma] \in \mathcal{M}_{d\times N}(\mathbb{K})$ defined as:

$$\gamma'(z)_i = \sum_{j=0}^{N} M[\gamma]_{i,j} z^j. \tag{2.1}$$

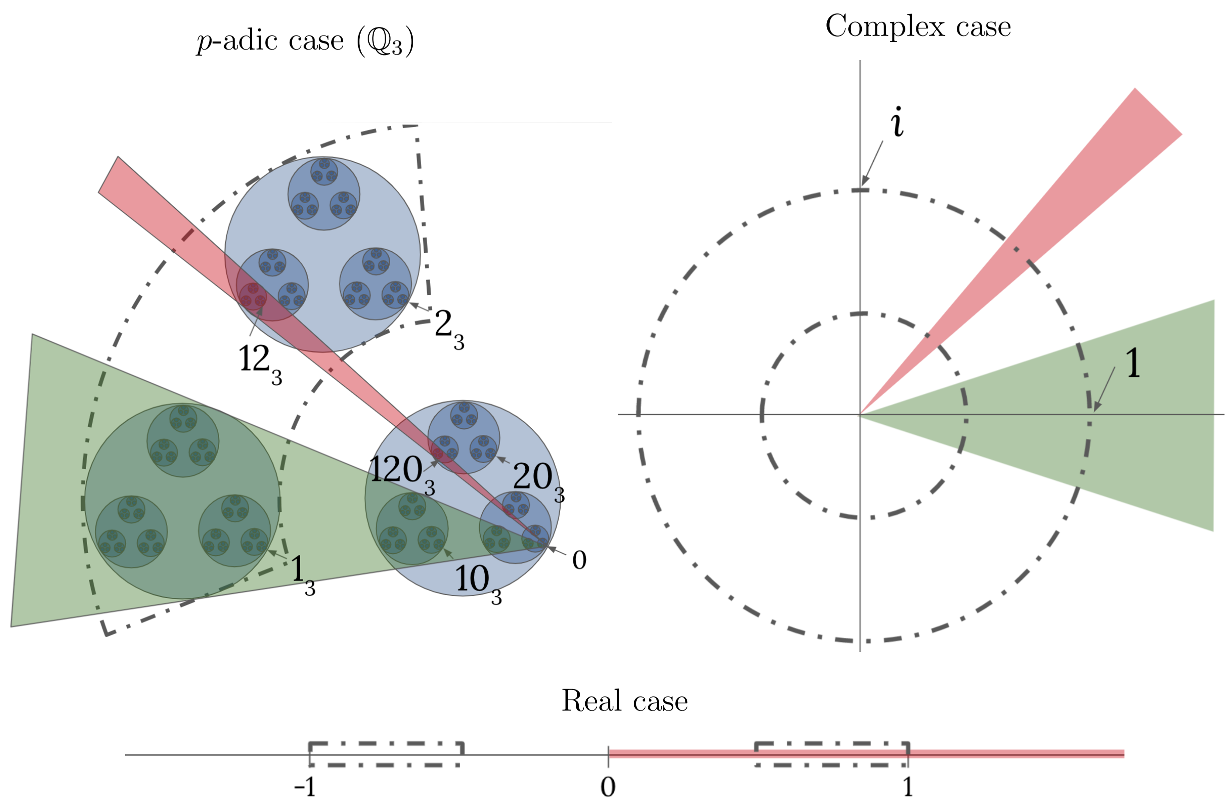


Figure 2.1: Representation of annuli and sectors over complex, real and 3-adic fields.

3-adic case ($\mathbb{Q}_3$): The triangle corresponds to $\Sigma_3^{\mathbb{Q}_3}$. The triangle corresponds to $5\Sigma_{1/27}^{\mathbb{Q}_3}$ (note that 5 in base 10 is 12_3 in base 3. All the numbers in the 3−adic picture are in base 3). The set enclosed by corresponds to $A_{[\frac{1}{2},1]}^{\mathbb{Q}_3}$.

Complex case ($\mathbb{C}$): The triangle corresponds to $\Sigma_{1/4}^{\mathbb{C}}$. The triangle corresponds to $(1+i)\Sigma_{1/10}^{\mathbb{C}}$. The annulus bounded by corresponds to $A_{[\frac{1}{2},1]}^{\mathbb{C}}$.

Real case ($\mathbb{R}$): The set represents the sectors $\Sigma_\alpha^{\mathbb{R}}$ for any $\alpha < 2$ (in the real case all these sectors degenerate to half-lines pointing either to the left or to the right). The set enclosed by corresponds to $A_{[\frac{1}{2},1]}^{\mathbb{R}}$.

In other words, $M[\gamma]_{i,j}$ is the coefficient of degree j of the i-th component of $\gamma'(z)$. Note that γ is not degenerate (in the sense that its not contained in a hyperplane, and its torsion is nonzero at at least one point) if and only if $M[\gamma]$ has rank d. If A is an affine map, then $M[A \circ \gamma] = D(A) \cdot M[\gamma]$, where we are identifying the differential $D(A)$ of the map A with the associated matrix in the canonical basis.

Example 2.1.4. *We will keep the curve $\gamma(t) = (t^4 - t^3 + t + 10, t^3 - t + 5)$ (or, later on, affine transformations of it) as a running example. For this curve we have:*

$$M[\gamma] = \begin{pmatrix} 1 & 0 & -1 & 1 \\ -1 & 0 & 1 & 0 \end{pmatrix}$$

note that $M[\gamma]$ does not keep track of the degree zero (translation) terms of γ.

Definition 2.1.5. *We will say that γ is in a canonical form of degrees $\mathbf{n} = (n_1, \ldots n_d)$ with $0 < n_1 < \cdots < n_d < N$ if $M[\gamma]_{i,n_j} = \delta_{i,j}$. We will say that $\tilde{\gamma}$ is a canonical form of γ of degrees $\mathbf{n}$ if $\tilde{\gamma}$ is in canonical form of degrees $\mathbf{n}$ and $\tilde{\gamma} = A \circ \gamma$ for some invertible affine transformation A.*

Note that if there is a canonical form of degrees $\mathbf{n}$ of γ then it is unique, that for a given multi-index $\mathbf{n}$ there may not be a canonical form at all, and that if γ is non-degenerate, there is at least one multi-indiex $\mathbf{n}$ for which there is a canonical form.

There are two canonical forms which are particularly relevant, the canonical form *at zero*, and the canonical form *at infinity*.

Definition 2.1.6. *The canonical form at zero has degree $\mathbf{n}^{(0)}$, identified by the fact that if γ has a canonical form of degree $\mathbf{n}'$ then $n_i^{(0)} \leq n_i'$. Similarly, the canonical form at infinity has degree $\mathbf{n}^{(\infty)}$, and for any other canonical form of γ of degree n' it holds that $n_i^{(\infty)} \geq n_i'$. The existence of $\mathbf{n}^{(0)}$ follows by row-reducing $M[\gamma]$ into reduced row echelon form, and the existence of $\mathbf{n}^{\infty}$ follows analogously.*

Example 2.1.7. *Keeping γ as in Example 2.1.4, we have that $\tilde{\gamma}_{1,4} := (t-t^3, t^4)$ is a canonical form of degrees $(1,4)$, and is the canonical form at zero. We have that $\tilde{\gamma}_{3,4} = (-t+t^3, t^4)$ is a canonical form of γ of degres $(3,4)$ and is the canonical form at infinity. In this example there are no other canonical forms.*

The degrees of the canonical forms (other than the one at infinity*) are not invariant (or even covariant) by reparametrization. As an example that will be useful in the following sections we consider $\gamma_1(t) = \gamma(t-1) := (-t^3+3t^2-3t, t^4-6t^3+4t^2-6t+1)$. This curve is affine-equivalent to $\tilde{\gamma}_1(t) := (t-t^2+\frac{1}{3}t, t^3-\frac{1}{8}t^4)$, which is in canonical form at zero with exponents $(1,3)$.*

As we will see in the following sections, for any point s not equal to $0, 1, -1$, the canonical form at zero of $\gamma_s := \gamma(t-s)$ has exponents $(1,2)$.

The relevance of the canonical form at zero and at infinity comes from the following fact:

Lemma 2.1.8. *Let $\gamma : \mathbb{K} \to \mathbb{K}^d$ be a polynomial curve with $\gamma(0) = 0$. Let $(\lambda_i)_{i=1}^{\infty} \in \mathbb{K}^{\times}$, be a sequence, with $\lambda_i \to 0$. Let A_i be a sequence of invertible linear maps in $\mathbb{K}^d$.*

1. *If $\gamma_i^0 := A_i \circ \gamma(\lambda_i z)$ converges pointwise to a non-degenerate curve $\tilde{\gamma}$, then $\tilde{\gamma}$ is affine equivalent to a generalized moment curve with exponents n^0.*

2. *If $\gamma_i^{\infty} := A_i \circ \gamma(\lambda_i^{-1} z)$ converges pointwise to a non-degenerate curve $\tilde{\gamma}$, then $\tilde{\gamma}$ is affine equivalent to a generalized moment curve with exponents n^{∞}.*

3. *The torsion $\Lambda^{(d)}[\gamma](0)$ does not vanish if and only if $n^{(0)} = (1, 2, \ldots, d)$*

In other words, the canonical form at zero and at infinity (and their associated exponents) describe the behavior of γ near zero and near infinity respectively.

Proof. For polynomials in local fields of characteristic zero, pointwise convergence of a sequence of polynomials of bounded degree implies convergence of the coefficients and therefore locally uniform convergence. We will show *(1)* only, as *(2)* follows exactly from the same arguments.

Without loss of generality one may assume that γ is already in its canonical form at zero. Let $L_\lambda := \mathrm{diag}(\lambda^{n_1}, \ldots, \lambda^{n_d})$. Let $\gamma_i := L_\lambda \circ \gamma(\lambda_i^{-1} x)$, and $\tilde{A}_i := A_i \cdot L_i^{-1}$. A computation in the matrix representation shows that

$$M[\gamma_i] = L_\lambda \cdot M[\gamma] \cdot \mathrm{diag}(\lambda^{-1}, \ldots \lambda^{-n}). \tag{2.2}$$

Since $M[\gamma]$ is in reduced row echelon form, $M[\gamma_i]$ is as well, and converges to the generalized moment curve $\mu_{\mathbf{n}^0}$. The matrix $M[\gamma_i]$ has full rank, and, for i large enough, so does $M[\gamma_i^0]$ (because of the hypothesis that γ_i^0 converges to a non-degenerate curve). In particular, one may recover (in a continuous manner) the linear maps $\tilde{A}_i$ from $M[\gamma_i]$ and $M[\gamma_i^0]$ once i large enough. Taking a limit of the $\tilde{A}_i$ shows the affine equivalence.

If $\Lambda^{(k)}[\gamma](0)$, the first d derivatives of γ at zero must be linearly independent, and therefore the first d columns of $M[\gamma]$ must form a rank d matrix. This shows *(3)*. □

This motivates the following definition:

Definition 2.1.9 (Zoom-in)**.** *Let γ be a polynomial curve in canonical form with exponents $\mathbf{n}$. Then we define the zoom in of γ at scale $\lambda \in \mathbb{K}$ as the curve $Z_\lambda[\gamma](z) = \mathrm{diag}(\lambda^{n_1}, \ldots \lambda^{n_d}) \circ \gamma(\lambda^{-1} z)$*

The last step will be to quantify the similarity between two polynomials in a certain region, and between a curve and a generalized moment curve.

Definition 2.1.10. *We will say that a polynomial $p : \mathbb{K} \to \mathbb{K}$, $p(z) = \sum_{i \leq \deg p} p_i z^i$, is $\epsilon-$similar to a homogeneous polynomial $q(z) = z^k$ in the annulus $A_{r,R} := \{x \in \mathbb{K}, |x| \in (r, R)\}$ if:*

- $p_k = 1$
- $|p_i| \leq \epsilon R^{d-|I|}$ *whenever* $|I| \geq d$.
- $|p_i| \leq \epsilon r^{d-|I|}$ *whenever* $|I| \leq d$.

We define p *being* $\epsilon-$*similar to* $q(z) = (z-a)^d$ *in the annulus* $A_{r,R}+a := \{x+a, x \in \mathbb{K}, |x| \in (r,R)\}$ *by translation. If* $r > R$ *(and in particular* $A_{[r,R)}$ *is empty) we vacuously say that any polynomial is* $\epsilon-$*similar to any other polynomial in* $A_{r,R}$.

The definition for polynomial curves will be a component-wise generalization of the defintion above:

Definition 2.1.11. *We will say that a curve* $\gamma : \mathbb{K} \to \mathbb{K}^d$ *is affine* $\epsilon-$*similar (or just close enough, when* ϵ *is implicit) to a monomial curve* $\mu_{\mathbf{n}}$ *in the annulus* $A_{r,R} := \{x \in \mathbb{K}, |x| \in (r,R)\}$ *if*

- *The curve* γ *is in canonical form of degrees* $\mathbf{n}$.
- $|M[\tilde{\gamma}]_{i,j}| \leq \epsilon R^{j-n_i}$ *whenever* $j > n_i$.
- $|M[\tilde{\gamma}]_{i,j}| \leq \epsilon r^{n_i-j}$ *whenever* $j < n_i$.

As in the single polynomial case, if $r > R$ *we vacuously say that any polynomial curve is* $\epsilon-$*similar to any other polynomial curve in* $A_{r,R}$.

Example 2.1.12. *Using the same curves as in our previous examples (see Example 2.1.7), the curve* $\tilde{\gamma}_{1,4} = (t - t^3, t^4)$ *is* $\epsilon-$*similar to* (t, t^4) *in* $A_{[0,\sqrt{\epsilon}]}$, *and* $\tilde{\gamma}_{3,4} = (t^3 - t, t^4)$ *is* ϵ*-similar to* (t^3, t^4) *in* $A_{[\epsilon^{-1/2},\infty)}$.

If a curve γ is ϵ-similar to $\mu_{\mathbf{n}}$ then each polynomial $\tilde{\gamma}_i$ is dominated by the monomial x^{n_i}, and, for our purposes (see Theorem 2.4.5) γ and $\mu_{\mathbf{n}}$ will become essentially equivalent for ϵ small enough. As we shall see in the sequel, if γ is ϵ-similar to μ, τ_γ (the torsion) is comparable to τ_μ.

Perhaps more surprisingly, the converse is true as well. Section 2.2 shows that if τ_γ is $\epsilon-$close to z^d in an annulus $A_{C^{-1}r,CR}$ then, after an affine transformation A, the curve $A \circ \gamma$ is $\epsilon-$close to a moment curve in the annulus $A_{r,R}$. This fact is the main insight in the proof of Theorem 2.1.13.

The definition of concepts given in this section is naturally extended around a point $z \neq 0$ by applying the concepts to $\gamma(\cdot - z)$.

2.1.3 The decomposition theorem

We have now all the tools needed to state the decomposition theorem:

Theorem 2.1.13 (Polynomial decomposition). *Let $\epsilon_D > 0$. Let $\mathbb{K}$, be a local field of characteristic zero. Let $\gamma : \mathbb{K} \to \mathbb{K}^d$ be a nondegenerate polynomial curve of degree N with coefficients in $\mathbb{K}$. Then there exists a decomposition of $\mathbb{K}$ as the union of the the closure of annuli*

$$\mathbb{K} = \bigcup_{j=1}^{O(1)} \overline{I}_j = \bigcup_{j=1}^{O(1)} \overline{[A_{r_j,R_j}(c_j)]}$$

with $R_j/r_j \geq 2$ such that the following holds:

To each I_j there is an associated generalized moment curve $\mu_j = (t^{n_{j,1}}, \ldots, t^{n_{j,d}})$ and an affine transformation $L_j : \mathbb{K}^d \to \mathbb{K}^d$ such that the curve $\gamma_j(z) := L_j \circ \gamma(z - c_j)$ is ϵ_D-similar (in the sense of Definition 2.1.11) to $\mu_j(z)$ in the annulus $A_{r_j R_j}$.

All the implicit constants depend only on $\mathbb{K}$, ϵ_D, N, and d.

Outline of the proof: Section 2.2 reduces the problem of finding the annuli that make part of Theorem 2.1.13 hold for a general to a decomposition statement for the torsion. This is the content of Lemma 2.2.2. This one-dimensional statement is proven in Section 2.3.

2.2 Reducing to a one-dimensional problem

The aim of this and the following section is to show show that the approximation property described in Theorem 2.1.13 holds for certain annuli that are *far away* from the zeros of the torsion of γ. We will call these annuli *admissible sets*. Our definition of admissible sets will only depend on the torsion, and this will reduce the problem to a one-dimensional question that will be adressed in Section 2.3.

To keep the notation consistent with a general class of differential forms defined in the sequel in Definition 2.4.1, we will denote the torsion of a curve γ by $\Lambda^{(d)}[\gamma]$.

Definition 2.2.1 (Torsion, see Definition 2.4.1)**.** *We will denote the torsion of* γ*, the determinant* $\det(\gamma'(z), \ldots, \gamma^{(d)}(z))$ *as* $\Lambda^{(d)}[\gamma]$.

Lemma 2.2.2 (The torsion $\Lambda^{(d)}[\gamma]$ encodes γ)**.** *Let* $\epsilon > 0$*. Then there exists* δ, C *depending only on* $d, N, \epsilon, \mathbb{K}$) *with the following property:*

Let γ *be a non-degenerate polynomial curve in* $\mathbb{K}^d$ *of degree at most* N*. If* $\Lambda^{(d)}[\gamma]$ *is* $\delta-$*close to* $p(z) = z^k$ *in an annuli* $A_{(r,R)}$ $(0 \leq r \leq R \leq \infty)$*. Then, for some multi-index* $\mathbf{n} = (n_1, \ldots n_d)$ *with* $\sum_{i=1}^{d} n_i - i = k$*, the curve* γ *has a canonical form* $\tilde{\gamma}$ *of degrees* $\mathbf{n}$ *and* $\tilde{\gamma}$ *is* $\epsilon-$*close to* $\mu_{\mathbf{n}}$ *in* $A_{Cr, C^{-1}R}$.

Note that the constants C, δ do not depend on r, R, in particular, result holds as well when $r = 0$ or $R = \infty$.

The importance of Lemma 2.2.2 lies in the fact that the polynomial $\Lambda^{(d)}[\gamma]$ is, up to a multiplicative constant, an affine invariant, but a priori contains much less information than the notion of ϵ-similarity (which is not affine-invariant). Lemma 2.2.2 provides an affine co-ordinate system addapted to each sector that will allow for computations in the following parts of Theorem 2.1.13.

The idea of controlling γ from $\Lambda^{(d)}[\gamma]$ is not completely new: Lemma 2.2 in [68] shares similarities, both in the statement and in the style of the proof, with the result in Lemma 2.2.2. It roughly states (in the $\mathbb{K} = \mathbb{R}$ setting, and translated to the notation of this work) that if $\Lambda^{(d)}[\gamma] \approx 1$ is approximately constant on a small interval I, then there is a transformation $A \in SU(\mathbb{R}; d)$ such that $\|A\gamma\|_{C^N(I)} \lesssim 1$ (where N is the degree of γ).

Remark 2.2.3. *The statement of Lemma 2.2.2 might seem counter-intiutive at first, especially because it has no counterpart in the class of smooth functions (this in it self should not be a surprise, a decomposition theorem for such curves cannot hold). Lemma 2.2.2* compresses *the information of a $d-$dimensional polynomial curve into a one dimensional polynomial. The following back-of-the-envelope computation sheds light on why the situation is different in the polynomial case.*

The space of polynomial curves of degrees $\mathbf{n} = (n_1, \ldots, n_d)$, with $n_1 < n_2 < \cdots < n_d$, considered up to translation, has dimension $\mathbf{n} = \sum_i n_i$. The set of linear maps that acting preserve $n_1 < n_2 < \cdots < n_d$ is the set of lower triangular maps, which has dimension $\frac{d^2+d}{2}$. The torsion of a polynomial of degree $\mathbf{n}$ can have degree up to $|\mathbf{n}| - \frac{d^2+d}{2}$, and therefore, the space of possible torsions up to scaling has dimension $|\mathbf{n}| - \frac{d^2+d}{2}$ as well. Therefore τ actually is a map between two spaces of the same dimension, polnomial curves up to affine transformation and polynomials of the right degree up to scaling.

Lemma 2.2.2 will be proven as a compactness-type result (and will, consequently, be ineffective in the implicit constants). The goal is to extract the knowledge of a pointwise result and make it uniform using compactnes. The pointwise fact is as follows:

Lemma 2.2.4. *If a polynomial curve γ is such that $\Lambda^{(k)}[\gamma](z) = \kappa z^n$ for some degree n and some $\kappa \neq 0$, then γ is an affine transformation of a moment curve.*

Proof of Lemma 2.2.4. First note that by composing with a linear deformation by the matrix $\mathrm{diag}(\kappa^{-1}, 1, \ldots, 1)$, one may assume that $\kappa = 1$. A curve γ is an affine transformation of a monomial curve if and only if the following two properties hold:

1. There exist d natural numbers $n_1 < n_2 < \cdots < n_d$ such that each component γ_i of γ is a linear combination of $z^{n_1}, \ldots, z^{n_d}$.

2. $M[\gamma]$ has rank d.

This is equivalent to stating that γ is non-degenerate and the canonical forms at zero (with degrees $\mathbf{n}^0$) and at infinity (with degrees $\mathbf{n}^\infty$) have the same degrees, and are thus the same.

Without loss of generality one may assume γ is in canonical form at zero and $\gamma(0) = 0$. Let $\mathbf{n}^0$ be the exponents of the canonical form at zero. Then the degree of vanishing of $\Lambda^{(d)}[\gamma]$ at zero is $\sum(n^0)_i - i$. In order to see this[1], consider the curves $\gamma^0_\lambda = \mathrm{diag}(\lambda^{-n^0_1}, \ldots \lambda^{-n^0_d})\gamma(\lambda^{-1}z)$ parametrized by γ. These curves converge uniformly to $\mu_{\mathbf{n}^0}$, and they all dilations of γ. In particular $\Lambda^{(d)}[\gamma]$ must have the same order of vanishing at zero as $\mu_{\mathbf{n}^0}$.

A direct computation shows that the degree of $\Lambda^{(d)}[\gamma]$ is $\sum(\mathbf{n}^\infty)_i - i$. The only way for both of them to be equal is that $\mathbf{n}_0 = \mathbf{n}_\infty$. $\square$

With this preliminary pointwise result, we can turn into the proof of Lemma 2.2.2. The proof is (essentially) a more quantitative version of the proof of Lemma 2.2.4, and uses Lemma 2.2.4 as a stepping stone.

Proof of Lemma 2.2.2. Without loss of generality the reader may want to assume that $\delta = C^{-1}$*, and* $\delta_n = C_n^{-1}$ *through the proof.*

The proof will go by contradiction. Fix ϵ. A negation of 2.2.2 is that there is a sequence of $(\gamma_n, r_n, R_n$ such that:

- The determinants $\Lambda^d[\gamma_n]$ are δ_n-similar to x^n in A_{r_n,R_n}, for $\delta_n \to 0$.

[1] One can proceed by direct observation of the coefficients in the determinant definition of $\Lambda^{(d)}[\gamma]$ as well.

- There exists constants $C_n \to \infty$ such that γ_n does not have a canonical form $\tilde{\gamma}_n$ that is $\epsilon-$similar to a generalized moment curve in $A_{C_n r_n, C_n^{-1} R_n}$.

Note that that already implies that $\frac{R_n}{r_n} \to \infty$ (for otherwise as soon as $C_n^2 \geq R_n/R_n$, ϵ-similarity would be vacuously true). By reparametrization invariance we will assume that $r_n = R_n^{-1}$. Our goal is then to find a subsequence such that $\tilde{\gamma}_{k_n}$ has a canonical form that $\epsilon-$similar to a generaized moment curve in $A_{C_n r_n, C_n^{-1} R_n}$.

Step 1: Finding a model curve:

Our first step will be to find a subsequence of the γ_n and a sequence of affine maps A_n such that $A_n \circ \gamma_n$ converges in the smooth topology in compact sets to a non-degenerate curve γ. In order to prove this, we will use the following version of the Gramm-Schmidt orthogonalization theorem:

Given a $d \times N-1$ $\mathbb{K}$-valued nonsingular matrix M there exists a $d \times d$ invertible matrix A such that the rows of $A \cdot M$ are orthogonal to each other, and the maximum of each row (in absolute value) is ~ 1.

Without loss of generality, by affine invariance of the statement, we can assume that the matrices $M[\gamma_n]$ have this property. By compactness we can pass to a convergent subsequence of the $M[\gamma_n]$ (resp. γ_n, because one can go back and forth from the matrix to the curve representation up to a translation). Let, in an abuse of notation, γ_n be this convergent subsequence. Now the all coefficients of each polynomial $(\gamma_n)_i$ converge, and thus the γ_n converge in the locally smooth topology to a polynomial curve γ. The Gramm-Schmidth procedure ensures γ is non-degenerate, and by construction $\Lambda^{(d)}[\gamma]$ is a monomial. Therefore, by Lemma 2.2.4, γ must be affine-equivalent to a generalized moment curve.

Step 2: Showing the model curve is $\epsilon-$close to γ

Again, by affine invariance, assume without loss of generality that $\gamma = \mu_{\mathbf{n}}$ is actually a moment curve of degrees $\mathbf{n}$ and that the γ_n that converge to γ are in canonical form converging to this γ. In particular, there are $q_n \to 0, Q_n \to \infty$ such that γ_n is ϵ-close to γ in

A_{q_n,Q_n} (but not for smaller values of q_n or larger values of Q_n).

We will show that $q_n \approx r_n$, and identical proof shows that $Q_n \approx R_n$. By a rescaling argument, define $\hat{\gamma}_n(z) = \text{diag}(q_1^{n_1}, \dots q_d^{n_d})\gamma_n(q_n^{-1}z)$. Now the $\hat{\gamma}_n$ are ϵ-similar to $\mu_{\mathbf{n}}$ in the annulus $A_{(1,Q_n q_n^{-1})}$ and no further. Let $\hat{r}_n = r_n/q_n$, $\hat{R}_n = R_n/q_n$. Assume by contradiction that $\hat{r}_n \to 0$. Since $\hat{R}_n \to \infty$ we have that $\Lambda^{(d)}[\gamma_n] \to z^k$ locally uniformly, and thus $\gamma_n \to \mu_{\tilde{\mathbf{n}}}$, which is affine-equivalent to a moment curve.

Since the $\hat{\gamma}_n$ are ϵ-close to μ_n in the annulus $A_{(1,2)}$, if $\epsilon < \frac{1}{10}$ they must have a subsequence that converges to a nondegenerate curve $\tilde{\gamma}$. On the other hand the $\Lambda^{(d)}[\gamma_n]$ converge locally uniformly to z^k, and thus the $\tilde{\gamma}$ must be a moment curve. On the other hand, the $\tilde{\gamma}$ is ϵ-close to μ_n in the annulus $A_{(1,2)}$, if $\epsilon < \frac{1}{10}$.

By definition, if $\epsilon < \frac{1}{10}$ no two different moment curves are affine ϵ-close to each other in $A_{(1,2)}$. Thus $\tilde{\gamma}$ is affine equivalent to $\mu_{\mathbf{n}}$. But all of the γ_n were in canonical position at $\mathbf{n}$, and thus $\tilde{\gamma} = \mu_{\mathbf{n}}$. This contradicts the fact that γ_n were not $\epsilon-$close to $\mu_{\mathbf{n}}$ in the annulus $A_{(1/2,1)}$.

An exactly analogous argument works with Q_n, R_n, finishing the proof. □

This has reduced the original problem to a one-dimensional problem. If we can cover $\mathbb{K}$ by a finite set of annuli A_i of center a_i such that $\Lambda^{(d)}[\gamma]$ is $\epsilon-$close to $(z-a_i)^{k_i}$ on each annulus, and can do that such that the number of covers only depends on the degree of $\Lambda^{(d)}[\gamma]$ we will have proven part Theorem 2.1.13. This is done in the next section.

Example 2.2.5. *We can study the statement above for $\mathbb{K} = \mathbb{R}$ and $\gamma(t) := (t^4 - t^3 + t, t^3 - t)$ (cf. Examples 2.1.4, 2.1.7 and 2.1.12 up to translations). The proof in the previous section is ineffective in the constants, and therefore specific computations cannot mirror the structure of the proof. Figure 2.2 contains pictures of the decomposition for this curve. It also contains a cover of $\mathbb{R}$ with intervals where*

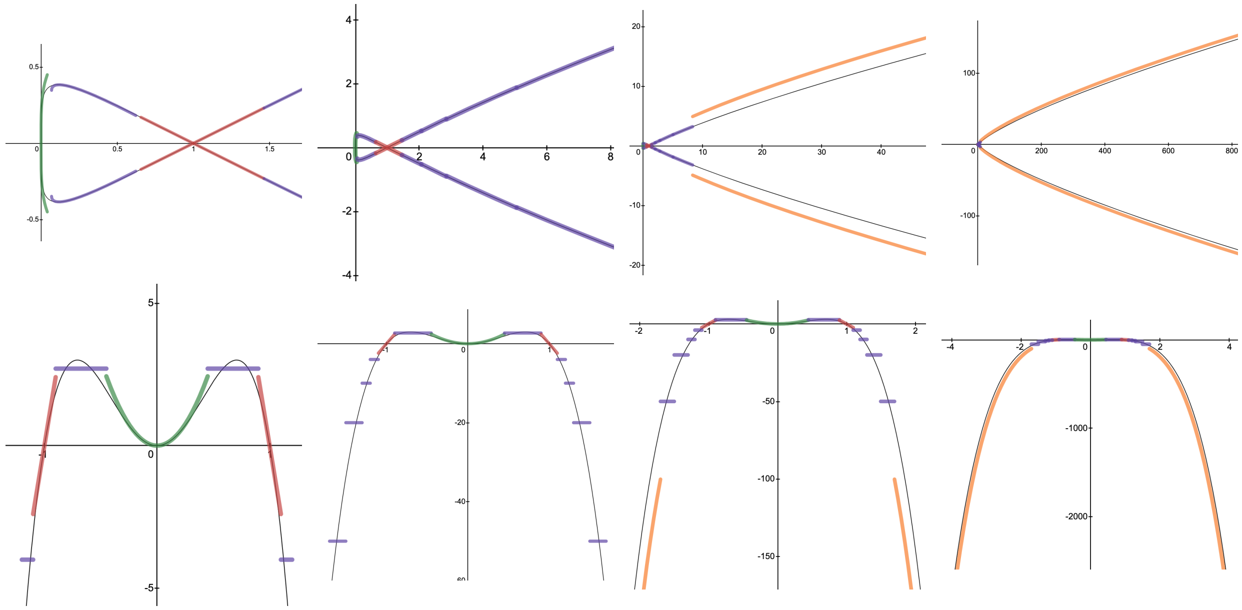


Figure 2.2: **(Top)** The curve $\gamma(t) := (t^4 - t^3 + t, t^3 - t)$ at different zoom levels (black). Affine transformations of generalized moment curves that approximate (are ϵ-similar) to $\gamma(t)$ in color. Up to affine transformations: (t, t^4) in green, (t, t^2) in purple, (t, t^3) in red, (t^3, t^4) in orange. **(Bottom)** The polynomial $\Lambda^{(2)}[\gamma]$, the torsion of γ. Translated monomials that approximate $\Lambda^{(2)}[\gamma]$ in color. They correspond to the same intervals depicted on the top. Up to horizontal translation and scaling: t^2 in green, $t^0 = 1$ in purple, t in red, t^4 in orange

2.3 Solving the one-dimensional case

The goal of this section is to prove the following one dimensional theorem:

Lemma 2.3.1. *Let $p(x)$ be a polynomial of degree N in $\mathbb{K}$, let $\delta > 0$, $C > 0$. Then one may cover $\mathbb{K}$ as a union of*

1. *A set of translated open annuli $A_1, \ldots A_G$, with $A_i = A_{r_i,R_i} + a_i$, and such that $\alpha_i^{-1} p(x)$ is $\delta-$similar to $(x - a_i)^{n_i}$ on the extended annulus $A_i = A_{C^{-1}r_i,CR_i} + a_i$*
2. *A set of singletons $\{b_1\}, \ldots \{b_B\}$ which are centers of of the annuli A_i that have inner radius r_i equal to zero.*

Moreover, G, B are bounded from above by a constant depending only on $N, \mathbb{K}, \delta, C$

Before proving Lemma 2.3.1, we will show that Theorem 2.1.13 follows from Lemma 2.3.1:

Proof of Theorem 2.1.13. Applying Lemma 2.3.1 to $p(x) = \Lambda^{(d)}[\gamma]$ provides the open annuli where the hypotheses of Lemma 2.2.2 hold, and therefore where Theorem 2.1.13 holds. By enlarging the value of C in Lemma 2.3.1 by a factor of 2 if necessary, one can assume that the condition $R_i/r_i > 2$ of Theorem 2.1.13 holds as well. $\square$

An common characteristic of all the fields we consider $\mathbb{K}$ is that for any number $N \geq 0$ there is a field extension $\mathbb{K}' \geq K$ where all polynomials of degree $\leq N$ with coefficiont in $\mathbb{K}$ split. If $\mathbb{K} \geq \mathbb{R}$ this follows from the fact that $\mathbb{C}$ is algebraically closed, If $\mathbb{K} \geq \mathbb{Q}_p$ it follows from the fact that $\mathbb{Q}_p$ only has finitely many extensions of a given degree. Fix N and let $\mathbb{K}'$ be such field. We will show Lemma 2.3.6 substituting $\mathbb{K}$ with this field $\mathbb{K}'$. Lemma 2.3.2 shows that a decomposition in $\mathbb{K}'$ can be used to construct a decomposition in $\mathbb{K}$ with a comparable number of sets. On a first read one may want to consider the case $\mathbb{K} = \mathbb{R} \leq \mathbb{C} = \mathbb{K}'$.

Lemma 2.3.2. *Let $L(p; \mathbb{K}, C, \epsilon)$ be the sum of $G + B$ as defined in Lemma 2.3.1 for a field $\mathbb{K}$, a polynomial $p \in K[x]$and a given C, ϵ. Fix a field extension $\mathbb{K} \leq \mathbb{K}'$. Then $L(p, \mathbb{K}, \epsilon, 10C) \lesssim_{\mathbb{K}', \mathbb{K}, \epsilon, C, \deg(p)} L(p, \mathbb{K}', 2\epsilon, C)$ as long as ϵ is small enough.*

Proof. The strategy of the proof is to decompose each annulus A_j of the decomposition induced by $L(\mathbb{K}', \epsilon, C)$ into further annuli $A_{j,k}$ (and keep all the bad points $\{b_i\}$ in the covering of $\mathbb{K}'$ that belong to $\mathbb{K}$).

Let A_j be one of the annulus in the decomposition of $\mathbb{K}'$, with center a_j, and without loss of generality assume $\alpha_j = 1$. Assume A_j intersects $\mathbb{K}$ (otherwise we may discard A_j) and $a_j \notin \mathbb{K}$ (otherwise we may keep $A_j \cap \mathbb{K}$). Let d_j be the distance between a_j and $\mathbb{K}$, and let $a'_j \in \mathbb{K}$ such that $|a_j - a'_j| = d_j$ (which must exist by local compactness). Without loss of generality $a'_j = 0$, so $d = |a_j|$.

For $S > \epsilon^{-1}N$ the polynomial $(z-a_j)^{n_j}$ is $\epsilon-$close to the polynomial z^j in the annuli $A_{S|a_j|,\infty}$. By adding the annuli $A_{j,0} = A_j \cap A_{S|a_j|,\infty}$ to the list, we can reduce to the case in which $R_j \leq 2\epsilon^{-1}|a_j|$. Without loss of generality, by rescaling, assume $|a_j| \approx 1$.

If δ is small enough (depending on $\deg p, \mathbb{K}', \mathbb{K}$) then $(x-a_j)^{n_j} \cdot (w-a_j)^{-n_n}$ is $\epsilon-$close to the constant polynomial $q(z) = 1$ in an annulus $A_{0,\delta} + w$ for any $w \in A_j \cap \mathbb{K}$. One can cover $w \in A_j \cap \mathbb{K}$ with $O(1)$ annuli of the form $A_{0,C^{-1}\delta_i} + w_i$ because we have explicit bounds of the form $r_j \sim R_j \sim 1$.

The result follows by the triangle inequality for $\epsilon-$closedness. □

From now on, therefore, we will assume that the polynomial p splits into linear factors in $\mathbb{K}$. We will use the zeros of p to build the annuli required of Lemma 2.3.1

Definition 2.3.3. *Let $w_1, \ldots w_k \in \mathbb{K}$ be a list of points in the plane. Let K be a large positive fixed constant. We will say that a set is K-*admissible* for $w_1, \ldots w_k$ if it is either:*

1. *A circle with center z and radius r such that $KR < |w_i - z|$*
2. *A circle with center w_j and radius R such that $KR < |w_i - z|$ for all for $i \neq j$. The radius R may be infinity if $k = 1$.*
3. *An annulus centered at a point z with inner and outer radii r, R such that for all $i = 1 \ldots k$, either w_i belongs to $\mathbb{K} \setminus B_{K^{-1} \cdot r}(z)$ or it belongs to to $B_{K \cdot R}(z)$. The outer radius R is allowed to be infinity.*

The motivation for this definition is that if $w_1, \ldots w_k$ are the zeros of $\Lambda^{(d)}[\gamma]$, then there is a constant $K(d, N, \epsilon)$ such that if a set S is $K-$admissible then $\Lambda^{(d)}[\gamma]$ is ϵ-close to a moment curve in the associated annulus. We can make this precise as the following fact:

Proposition 2.3.4. *Let K large enough (depending on $\mathbb{K}, d, N, \epsilon$), and let γ be a degree n nondegenerate polynomial curve such that $\Lambda^{(d)}[\gamma]$ splits into linear factors over $\mathbb{K}$. Let*

$w_1, \dots w_m$ be the zeros of $\Lambda^{(d)}[\gamma]$. If A is K-admissible for $w_1, \dots w_m$ then conclusion (1) of Lemma 2.3.1 holds in A.

Proof. Without loss of generality, assume $A = A_{r,R}$ is centered at zero. We will relabel the zeros as v_j if they are inside the annulus, and w_j if they are outside. In other words, $\Lambda^{(d)}[\gamma]$ is of the form:

$$\Lambda^{(d)}[\gamma] = \alpha \prod_{j=1}^{m}(x - v_j) \prod_{j=1}^{n} \left(\frac{x}{w_j} - 1 \right) \tag{2.3}$$

with $v_j < K^{-1}r$, and $w_j > KR$. Again, by scale invariance (using the α_i in conclusion *(1)*) one may assume $\alpha = 1$. We will show that this polynomial is $\epsilon-$similar to the polynomial x^m in A. If we expand the product in equation (2.3) , there is a term of the form x^m, so we will only have to show that all the other terms are negligible.

Any other term contributing to the monomial x^{m+k} (for $k > 1$) will be a product of s terms of the form w_j^{-1} and t terms of the form v_j, with $t - s = k$. Applying the bounds to the v_j, w_j, this leads to a bound in this term of the form $(KR)^{-k} \left(\frac{r}{RK^2}\right)^s$, which is bounded by ϵR^{-k} if K is large enough. A parallel approach gives the same result when $k < 1$.

When $k = 0$ (the x^m monomial) we will have one term when $t = s = 0$ (the one giving the main contribution) and multiple contributions where $t = s > 0$. Each of them will be bounded by $\left(\frac{r}{K^2R}\right)^t$, which is a negligible contribution as well. □

Thus, what remains to be done is to show that $\mathbb{K}$ can be covered by $O_{K,k}(1)$ $K-$admissible subsets. That is:

Lemma 2.3.5. *Let $w_1, \dots w_k$, K as in 2.3.3. There is a family $\{A_i\}_{i=1}^{M}$ of admissible subsets of $\mathbb{K}$ such that $\mathbb{K} = \bigcup_{i=1}^{M} A_i$. The value of M depends only on k and K, but not on the specific value of the w_i.*

Before proving this, we will need a uniform covering lemma for compact subsets of a compact metric space:

Lemma 2.3.6 (Uniform Covering Lemma)**.** *Let (X, d) be a compact metric space and $\epsilon > 0$. Then there exists $N := N(\epsilon, X, d)$ such that for any closed subset $K \subset X$ for any open cover $K \subseteq \bigcup_{x \in I} B_\epsilon(x)$ made of balls $B_\epsilon(x_i)$ there is a finite subfamily indexed by I' with $|I'| \leq N$ such that $K \subseteq \bigcup_{x \in I'} B_{2\epsilon}(x)$. In particular, there is a "covering number" $N_c := N_c(\epsilon, X, d) \leq N(\epsilon, X, d)$ such that any compact subset of K can be covered with at most N_c balls or radius ϵ with center in K .*

Proof (of Lemma 2.3.6). Assume there exists a sequence (K_n, I_n) that witnesses a contradiction to the Lemma above. Let $I_n = (x_{n;j})_{j=1}^{m(n)}$. Without loss of generality (by passing to a subsequence), one may assume that $m(n)$ is strictly increasing, K_n converges to a compact subset K in the Hausdorff metric for compact sets, and $(x_{n;j})_{n \to \infty}$ converges as well (for each n) to an element x_j of X.

Let $I := (x_j)_{j=1}^{\infty}$, then (by compactness) there exists a finite subfamily $I' \subseteq I$ such that $K \subseteq \bigcup_{x \in I} B_\epsilon(x)$. For all $n > n_0$ large enough, since $|I'| < \infty$, it must hold that $d(x_{n:j}, x_j) < \epsilon/4$, and $d_H(K_n, K) < \epsilon/4$. Now (still, for $n > n_0$) define $I'_n := \{x_{n;j} : x_j \in I'\}$. By the triangle inequality,$K_n \subseteq \bigcup_{x \in I'_n} B_{2\epsilon}(x_n)$. □

With this in hand, we can proceed to the proof of Lemma 2.3.5

Proof (of Lemma 2.3.5). One of the key facts is the affine invariance of the covering problem. If we rescale the w_i, then the set of K-admissible sets is equally rescaled.

Note that the case $k = 1$ is trivial by the second type of admissible set. Before starting the induction, we will remove a neighborhood of infinity:

By rescaling and translating, we may assume that $w_1 = 0$, all the w_k are contained in the ball of radius $\frac{1}{2}K^{-1}$ and that at least one of the w_k has absolute value $|w_k| \approx_{\mathbb{K}} \frac{1}{2}K^{-1}$. The set $\mathbb{K} \setminus \mathbb{B}_1(0)$ can be covered with an admissible annulus of inner radius 1 and outer radius ∞.

Now, by Lemma 2.3.6 the annulus $B_1(0) \setminus B_{K^{-1}}(0)$ can be covered with $O_K(1)$ admissible balls of radius $\frac{1}{2}K^{-1}$. By rescaling again, it suffices to solve the case when that all the w_i are in the unit ball, one of them has absolute value $\approx_{\mathbb{K}} \frac{1}{2}$ we want to cover the unit ball, and can only use balls or annuli of radius ≤ 1 with center inside the unit ball. We will solve this by induction on k.

In this case one can find a finite amount m (at most k) of radii and centers $\{z_j\}_{j=1}^m$, $\{R_j\}_{j=1}^m$, such that

$$\frac{1}{10} \max |w_i - w_j| > R_j \geq \frac{1}{10} \max |w_i - w_j| (20K^2 k)^{-k},$$

with the following property:

Each w_i belongs exactly to one of the balls $B_{R_j}(z_j)$, and if w_i belongs to $B_{10K^2 R_j}(z_j)$ then it belongs to $B_{R_j}(z_j)$ as well. This is a pigeonholing argument proven with detail as Lemma 2.3.7 below.

One may assume as well without loss of generality that the centers z_j are one of the w_i. Imposing $R_j < \frac{1}{C_0}$ for C_0 large enough, together with the fact that $w_1 = 0$ and $|w_i| \approx_{\mathbb{K}} \frac{1}{2}$ for some i ensures that all the $B_{R_j}(z_j)$ contain strictly less than k elements. This will allow for the induction on k.

We can then cover $B_1(0) \setminus \bigcup_{j=1}^m B_{3KR_j}(z_j)$ with $O_{K,k}(1)$ admissible balls of radius $\sim (2k)^{-K}$. Now it suffices to show that we can cover each of the $B_{3R_j}(z_j)$ with $O_{k,K}(1)$ admissible sets. Fix a z_j, R_j. Let $s_j = \max_{w_i \in B_{R_j}(w_j)} |w_i - z_j| < R_j$.

By construction, $Ks_j < KR_j$, and the annulus $A_{[Ks_j, 3KR_j]} + z_j$ is admissible. By an affine transformation and rescaling, we can reduce ourselves to the situation when $z_j = 0$ and $s_j \approx_{\mathbb{K}} \frac{1}{2}$. By hypothesis z_j was equal to one of the w_i in $B_{R_j}(z_j)$, and the number of poitns w_j inside $B_{R_j}(z_j)$ is strictly less than k. Therefore can cover $B_{Kbs_j}(z_j)$ with $O_{k,K}(1)$ admissible sets by the induction hypothesis. $\square$

Lemma 2.3.7 (Logarithmic clustering). *Let $(X, d(\cdot,\cdot))$ be a metric space (in our application X is $B_1(0) \subseteq \mathbb{C}$), let $w_1, \dots w_k \in X$, and let $T \in \mathbb{R}^+$ (in our application $T = 10K^2$). Let $M := \max_{i,j} d(w_i, w_j)$. Then there exist $s \leq k$ balls $B_{x_i}(r_i)$ of X (where each x_i can be chosen to be one of the w_j), with $r_i \leq M/10$, such that each w_j belongs to exactly one of the balls $B_{x_i}(r_i)$, and such that if $w_j \in B_{x_i}(Tr_i)$ then $w_j \in B_{x_i}(r_i)$. Moreover, $\frac{M}{10} \geq r_i \geq \frac{M}{10}(2Tk)^{-k}$ for all $i = 1 \dots s$.*

Proof. Without loss of generality assume $M = 1$. Define

$$\delta_i := \max_{x \leq 1/10} \{x : \forall j = 1 \dots k, d(w_i, w_j) \notin [x, 2kTx]\}.$$

It must be that $\delta_i > \frac{1}{10}(2Tk)^{-k}$ because there are at most k unique distances. Let $\sim$ be the (transitive) equivalence relation generated by $d(w_i, w_j) < \delta_i \implies w_i \sim w_j$. Now we construct the balls explicitly. From each equivalence class $[w_i]$, choose a representiative w_i such that the associated δ_i is maximized amongst the equivalence class. Without loss of generality assume those repersentatives are $w_1, \dots w_s$, and let $r_1, \dots r_s$ be defined by $r_i = k\delta_i$. A routine computation now shows that the balls $B_{w_s}(r_s)$ have the desired properties. □

This concludes the proof of Theorem 2.1.13.

2.4 The transfer theorem

In applications of Theorem 2.1.13, one wants to *pretend* that the outcome of the decompositon is a monomial curve. In particular, that a set of particularly useful inequalities for monomial curves hold for these approximate monomial curves. Theorem 2.4.5 allows us to transfer inequalities from the monomial curves to their approximations. In order to state it and prove it, we must define a family of differential forms associated to a curve.

2.4.1 Affine-covariant differential forms

In the following sections we will not only consider the affine measure, but a set of related differential forms on the domain of γ:

Definition 2.4.1. *For $0 < k \leq d$ define:*

$$\Lambda^{(k)}[\gamma](z) := \gamma'(z) \wedge \cdots \wedge \gamma^{(k+1)}(z) \tag{2.4}$$

$$\Lambda[\gamma](z_1, \ldots, z_k) := \gamma'(z_1) \wedge \cdots \wedge \gamma'(z_k) \tag{2.5}$$

Note that $\Lambda[\gamma]$ is a function with variable arity (which will be clear by the context) that has an element of $\mathbb{K}^k$ as an input and returns a $k-$form as an output. We will denote the Vandermonde determinant by

$$v(z_1, \ldots z_k) := \prod_{i<j}(z_i - z_j) \tag{2.6}$$

Comparing with the notation $\tau_\gamma = L_\gamma,\ J_\gamma$ used in [24, 68, 2] for some special cases of $\Lambda[\gamma]$, we obtain the examples

$$L_\gamma(z) := \frac{1}{d!}\Lambda^{(d)}[\gamma](z) = \frac{1}{d!}\det[\gamma'(z), \gamma''(z), \ldots, \gamma^{(d)}(z)], \tag{2.7}$$

which leads to $\lambda[\gamma] = \tau_\gamma^{2d_{\mathbb{K}}/d^2+d} = L_\gamma^{2d_{\mathbb{K}}/d^2+d}$. We also have $J_\gamma(x_1, \ldots x_d) = \Lambda[\gamma](x_1, \ldots x_d)$.

Example 2.4.2 (Differential forms associated to curves)**.** *Let $\mu(z) := (z, z^2, z^3)$ be the standard moment curve. Let e_1, e_2, e_3 be the canonical co-ordinate basis on the 1-forms $\Lambda^1(\mathbb{C}^3)$. Then*

$$\begin{aligned} \Lambda^{(1)}[\mu](z) =& e_1 + 2ze_2 + 3z^2e_3 \\ \Lambda^{(2)}[\mu](z) =& \Lambda^{(1)}[\mu](z) \wedge (2e_2 + 6ze_3) \\ =& 2e_1 \wedge e_2 + 6z^2 e_2 \wedge e_3 - 6ze_3 \wedge e_1 \\ \Lambda^{(3)}[\mu](z) =& \Lambda^{(2)}[\mu](z) \wedge (6dw_3) = 12e_1 \wedge e_2 \wedge e_3, \end{aligned} \tag{2.8}$$

and

$$\begin{aligned}
\Lambda[\mu](z) =& e_1 + 2ze_2 + 3z^2e_3 \\
\Lambda[\mu](z_1, z_2) =& \mu'(z_1) \wedge \mu'(z_2) \\
=& (z_2 - z_1)(2e_1 \wedge e_2 + 6z_1z_2e_2 \wedge e_3 \\
& - 3(z_1 + z_2)e_3 \wedge e_1) \\
\Lambda[\mu](z_1, z_2, z_3) =& \det(\mu'(z_1), \mu'(z_2), \mu'(z_3))e_1 \wedge e_2 \wedge e_3 \\
=& 6v(z_1, z_2, z_3)e_1 \wedge e_2 \wedge e_3.
\end{aligned} \tag{2.9}$$

Note that in the preceding example, the form $\Lambda[\gamma](z_1, z_2, z_3)$ *is divisible by the Vandermonde determinant* $v(z_1, z_2, z_3) := (z_1 - z_2)(z_1 - z_3)(z_2 - z_3)$ *(analogously,* $\Lambda[\gamma](z_1, z_2)$ *is divisible by* $v(z_1, z_2)$*). This is not by chance, if two points* z_i, z_j *are the same then the vectors associated to them are aligned, and the associated form must vanish.*

This example can be generalized to moment curves, and to generalized moment curves, as shown in the sequel. Let $\mathbf{n}$ *be an increasing multi-index of length* d*. Let* $\gamma_{\mathbf{n}}$ *be the* generalized moment curve of degree $\mathbf{n}$, $\gamma_{\mathbf{n}} := (z^{\mathbf{n}_1}, \ldots, z^{\mathbf{n}_d})$*. Define* $\delta(\mathbf{n}) := (\mathbf{n}_1 - 1, \mathbf{n}_2 - \mathbf{n}_1 - 1, \mathbf{n}_3 - \mathbf{n}_2 - 1, \ldots, \mathbf{n}_k - \mathbf{n}_{k-1} - 1)$.

Define S_δ*, the Schur polynomial of index* δ*, to be the polynomial obeying the identity* $S_\delta(z_1, \ldots z_k)v(z_1, \ldots z_k) = \det[(z_i^{n_j})_{i,j=1,\ldots n}]$*. For a given element* λ *of the exterior algebra of order* k*, and a canonical basis element* $e_{j_1} \wedge \cdots \wedge e_{j_k}$*, let denote* $\lambda|_e$ *the coefficient of* w *in the canonical basis. Carefully unpacking the definitions shows that*

$$\Lambda[\mu_{\mathbf{n}}](z_1, \ldots, z_k)|_{e_{j_1} \wedge \cdots \wedge e_{j_k}} = \det[(z_i^{n_{j_s}})_{i,s=1,\ldots k}] = S_{\delta(n_{j_1}, \ldots n_{j_k})}(z_1, \ldots z_k)v(z_1, \ldots z_k). \tag{2.10}$$

As we shall see in section 2.6 for a general curve γ *we can recover estimates for* $\Lambda^{(k)}(z)$ *from estimates for* $\Lambda(z_1, \ldots z_k)$*, using the equality*

$$\Lambda^{(k)}(z) = \lim_{(z_1,\ldots,z_k)\to(z,\ldots z)} c_k \left[\frac{\Lambda(z_1, \ldots z_k)}{v(z_1, \ldots z_k)}\right]$$

with $c_k = \prod_{i=1}^{k-1} k!$*. Showing that the form* $\left[\frac{\Lambda(z_1,\dots z_k)}{v(z_1,\dots z_k)}\right]$ *extends continuously to the zero set of* $v(z_1,\dots,z_k)$ *for general curves other than the generalized moment curve, and understanding this limit is a key step of the proofs in section 2.6.*

The fact that, as in the example from the moment curve, the quotient of a differential form by the Candermonde determinant is well-defined, motivates the following definition.

Definition 2.4.3 (Corrected multilinear form)**.** *For* $\gamma : \mathbb{K} \to \mathbb{K}^n$ *and* $\mathbf{z} \in \mathbb{K}^n$ *we define:*

$$\tilde{\Lambda}[\gamma](\mathbf{z}) = \frac{\Lambda[\gamma](\mathbf{z})}{v(\mathbf{z})} \tag{2.11}$$

We will prove that map $\tilde{\Lambda}[\cdot](\cdot)$ is continuous in both variables of domain $\mathbb{K}^d \times P_N(\mathbb{K})^d$.

Example 2.4.4. *Again for moment curve in dimension 3 (as in Example 2.4.2), we see that*

$$\tilde{\Lambda}[\mu](z_1,z_2) = 2de_1 \wedge de_2 + 6z_3z_2de_2 \wedge de_3 - 3(z_1+z_2)de_3 \wedge de_1$$

(which, for $z_1 = z_2$ *happens to be equal to* $\Lambda^{(2)}[\mu](z)$*, as expected from Lemma 2.5.2 with the multiplicative* $C_2 = \frac{1}{0!1!} = 1$*).*

The next relevant example (as in Example 2.4.2 again) is the case of generalized moment curves $\mu_{\mathbf{n}}$*. In that case the coefficients of* $\tilde{\Lambda}[\mu_{\mathbf{n}}](z_1,\dots,z_k)$ *are Schur polynomials. In particular, the coefficient associated to the basis element* $e_{j_1} \wedge \dots \wedge e_{j_k}$ *is (up to a constant* $\prod_{i=1}^{k} \mathbf{n}_{j_i}$ *arising from the differentiating the monomials* $z_i^{\mathbf{n}_i}$*) the Schur polynomial* $S_{\delta(\mathbf{n}_{j_1},\dots,\mathbf{n}_{j_k})}$*.*

2.4.2 The transfer theorem

Theorem 2.4.5 (Transfer theorem)**.** *Let* $n, D > 0$*. Let* $\mathbb{K}$ *be a local field of characteristic zero or sufficiently high characteristic. Let* $\epsilon > 0$*. Then there exists* $\epsilon_D = \epsilon_D(n,d,\mathbb{K},\epsilon)$ *so that the following holds:*

If $\gamma : \mathbb{K} \to \mathbb{K}^d$ *is a nondegenerate curve polynomial curve of degree* $\leq N$ *that is* $\epsilon-$*similar to a monomial curve* μ_0 *on an annulus* $A_{r,R}(0)$*, with* $R \geq 2r$*, then, for any angular sector*

$t\Sigma_\epsilon^{\mathbb{K}}$*, and any points* $z_1, \dots z_k \in (\Sigma_\epsilon^{\mathbb{K}} \cap A_{r,R})$*, it holds that*

$$|\tilde{\Lambda}[\gamma](z_1, \dots, z_k) - \tilde{\Lambda}[\mu_0](z_1, \dots, z_k)| < \epsilon_D |\tilde{\Lambda}[\mu_0](z_1, \dots, z_k)|. \tag{2.12}$$

Remark 2.4.6. *Note that by multiplying both sides by* $v(z_1, \dots, z_k)$ *one gets the same inequality for* $\Lambda[\gamma]$.

Note also that the outer radius R *can be exactly infinity*[2]*, and the inner radius* r *can be exactly zero.*

The *transfer theorem* is named as such because it allows the transfer inequalities from monomial curves to the curves arising from the partition of Theorem 2.1.13. As a first application before [showing this result, we will prove the decomposition result of Dendrinos-Wright [24] for local fields of sufficiently high characteristic:

Proposition 2.4.7 (Dendrinos-Wright for local fields)**.** *Let* $N, d > 0$*. Let* $\mathbb{K}$*, be a local field of characteristic zero or sufficiently high characteristic. Let* $\gamma : \mathbb{K} \to \mathbb{K}^d$ *a polynomial curve of degree* N *such that* $\Lambda^{(d)}[\gamma](z) \not\equiv 0$*. Then there exists a decomposition of* $\mathbb{K}$ *as the union of the the closure of truncated sectors*

$$\mathbb{K} = \bigcup_{j=1}^{O(1)} \overline{I}_j = \bigcup_{j=1}^{O(1)} \overline{\left[(t_j \Sigma_\epsilon \cap A_{r_j R_j}) + c_j\right]}$$

of amplitude $\delta = \delta(N, d, \mathbb{K}))$ *and with* $R_j/r_j \geq 2$ *such that for all* $(z_1, \dots, z_d) \in I_j^d$*,*

$$|\Lambda[\gamma](z_1, \dots, z_d)| \geq C_{d,N,\mathbb{K}} |v(t_1, \dots t_d)| \prod_{k=1}^{d} |\Lambda^{(d)}[\gamma](z_k)|^{1/d}. \tag{DW}$$

Proof. The inequality (DW) holds for a curve $\gamma(x)$ if and only if it holds for an invertible affine transformation $A\gamma(z-t)$ of the curve. In particular, by the decomposition theorem (Theorem 2.1.13), it suffices to show (DW) holds on sectors of the form $t\Sigma_\delta \cap A_{(r,R)}$ for polynomial curves γ which are $\epsilon-$similar to a moment curve μ in $A_{(r,R)}$, for ϵ small enough.

[2] *Or arbitrarily large, for what the matters of the proof matters there's no difference: if one can find a* δ *that makes the lemma true for any* $R > R_0$ *then that same* δ *also works setting* $R = \infty$*. The same argument shows the analogous result for* $r = 0$.

Section 2.5, as a consequence of Lemma 2.5.5 shows that (DW) holds on $t\Sigma_\delta$ when γ is exactly a moment curve and δ is small enough (depending on the degree, the dimension, and the field only). Now the result follows from Theorem 2.4.5, with ϵ_D small enough so that $\epsilon \leq \frac{1}{2}$.

By Lemma 2.5.2, $\Lambda^{(d)}[\gamma](z) = C_d\tilde{\Lambda}[\gamma](z,\dots,z)$, and in particular, (2.12) holds for $\Lambda^{(d)}[\gamma]$ as well. From here,

$$|\Lambda[\gamma](z_1,\dots,z_d)| \approx |\Lambda[\mu](z_1,\dots,z_d)| \tag{2.13}$$

$$\geq C_{d,N,\mathbb{K}}|v(z_1,\dots,z_d)|\prod_{k=1}^{d}|\Lambda^{(d)}[\mu](z_k)|^{1/d} \tag{2.14}$$

$$\approx C_{d,N,\mathbb{K}}|v(z_1,\dots,z_d)|\prod_{k=1}^{d}|\Lambda^{(d)}[\gamma](z_k)|^{1/d}. \tag{2.15}$$

The middle step was the fact that (DW) holds for monomial curves, which we will see in the following section by an explicit computation. □

Strategy of the proof of Theorem 2.4.5 Section 2.5 shows that inequality (DW) of the geometric theorem holds when γ is exactly a moment curve, as well as certain *no cancellation* properties for $\tilde{\Lambda}[\mu]$, for monomial curves μ (essentially, that $\tilde{\Lambda}[\mu]$ is a sum of monomials with coefficients in the nnatural numbers).

This fact is then used in section Section 2.6 to show Theorem 2.4.5. The proof is compactness-based.

2.5 Differential forms associated to a moment curve

This section deals both with showing that Proposition 2.4.7 holds for monomial curves. To prove Proposition 2.4.7 for monomial curves we will use that, for a monomial curve $\mu_{\mathbf{n}}$ with exponents $\mathbf{n} := (n_1 < \cdots < n_d)$ (that is $\mu_{\mathbf{n}}(z) := (z^{n_1},\dots,z^{n_d})$), and for $\mathbf{z} := (z_1,\dots,z_d)$ it holds that:

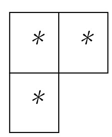
$$\frac{\Lambda[\mu_{\mathbf{n}}](\mathbf{z})}{v(\mathbf{z})} = \pi_{\mathbf{n}} S_{\delta(\mathbf{n})}(\mathbf{z}) \tag{2.16}$$

where $\pi_{\mathbf{n}}$ is the product of all the n_i, $v(z_1, \dots z_d) := \prod_{i<j}(z_i - z_j)$ is the Vandermonde determinant, $\delta(\mathbf{n}) = (n_1 - 1, n_2 - 2, \dots, n_d - d)$ is the excess degree of the generalized moment curve, and $S_{\mathbf{k}}$, for a general non-decreasing multi-index k is the Schur polynomial associated to $\mathbf{k}$. This equality can be taken as a definition of the Schur polynomials (if we compare it with the usual definition of $S_{\delta(\mathbf{n})}$, the factor $\pi_{\mathbf{n}}$ appears when taking the derivatives). A classical result in algebraic combinatorics (see, for example, [31]) states that one may express the Schur polynomials as the sum

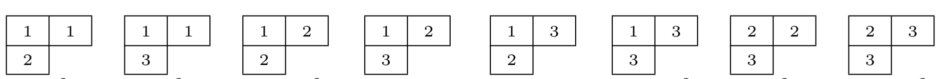
$$S_{\mathbf{k}}(z_1, \dots, z_d) = \sum_{(t_i) \in T_{\mathbf{k}}} z_1^{t_1}, \dots, z_d^{t_d} \tag{2.17}$$

where $T_{\mathbf{k}}$ is the set of semistandard Young tableaux of shape $\mathbf{k} = (k_1, k_2, \dots k_d)$, and (t_i) are the weights associated to each tableaux t.

Example 2.5.1. *Let* $\mathbf{n} = (1, 3, 5)$. *Then the excess degree is* $(0, 1, 2)$. *The Young diagram associated to this partition is of the form*



where the row of length 0 is not drawn. Thus, counting over all possible semi-standard Young Tableaux (ways of filling the Young diagram with the indices $1, 2, 3$ $(1, \dots, d$ *in general) that are strictly increasing vertically and weakly increasing horizontally) gives the polynomial:*



$$S_{(0,1,2)} = z_1^2 z_2 + z_1^2 z_3 + z_1 z_2^2 + z_1 z_2 z_3 + z_1 z_2 z_3 + z_1 z_3^2 + z_2^2 z_3 + z_2 z_3^2$$

This relationship will give us all the control needed for the term $\Lambda(\mu_{\mathbf{n}})(z_1, \dots z_d)$. The other term involved in the Dendrinos-Wright inequality (DW) we want to show is a term of the form $\Lambda^{(d)}[\gamma](z) = \gamma'(z) \wedge \cdots \wedge \gamma^{(d)}(z) = \det(\gamma'(z), \ldots, \gamma^{(d)}(z))$. The following lemma relates $\Lambda^{(d)}$ to Λ, allowing to express $\Lambda^{(d)}$ in terms of the Schur polynomial:

Lemma 2.5.2. *Let $\mathbf{z} \in \mathbb{K}^d$, with $z_i \neq z_j$ for $i \neq z$, let $s \in \mathbb{K}$, and $\gamma : \mathbb{K} \to \mathbb{K}^d$ a polynomial curve, then:*

$$\Lambda^{(d)}[\gamma[(s) = C_d \lim_{\substack{|\lambda| \to 0 \\ \lambda \in \mathbb{K}}} \frac{\Lambda[\gamma](\lambda \mathbf{z} + s)}{v(\lambda \mathbf{z})} \tag{2.18}$$

and, in particular, in the case when γ is a moment curve of exponent $\mathbf{n}$,

$$\Lambda^{(d)}[\gamma](s) = C_d S_{\delta(\mathbf{n})}(s, \ldots, s) \tag{2.19}$$

Proof. By Taylor expansion we have:

$$\gamma_i'(s + \lambda \mathbf{z}_j) = \sum_{k=1}^{d} \frac{1}{(k-1)!} \gamma_i^{(k)}(s) \lambda^{k-1} \mathbf{z}_j^{k-1} + O(\lambda^d). \tag{2.20}$$

Now, defining the matrices $\Gamma'_{ij} = \gamma_i(s + \lambda \mathbf{z}_j)$, $Z_{kj} = (\lambda \mathbf{z}_j)^{k-1}$, and $(T_\gamma)_{ik} = \frac{1}{(k-1)!} \gamma_i^{(k)}(s)$ the equation above can be rewritten as:

$$\Gamma = T_\gamma Z + O(\lambda^d). \tag{2.21}$$

Since the determinant of Z is $v(\lambda \mathbf{z})$, the lemma follows from the multiplicative property of the determinant:

$$\frac{\Lambda[\gamma](\lambda \mathbf{z} + s)}{v(\lambda \mathbf{z})} = \frac{\det \Gamma}{\det Z} = \det[T_\gamma + Z^{-1} O(\lambda^d)] \overset{\lambda \to 0}{\to} \det T_\gamma = C_d \Lambda^{(d)}[\gamma](s). \tag{2.22}$$

The fact that $Z^{-1} = o(\lambda^{-d})$ (and thus we can eliminate the term as $\lambda \to 0$) follows from the adjoint formula for the inverse. Note that the value of $C_d = \prod_{i=1}^{d} \frac{1}{(k-1)!}$ is in fact explicit. □

Remark 2.5.3. *The computations above use the fact that the definitions of derivative, determinant and Taylor expansion are the same over all the fields considered in the paper, and that, in particular, the derivatives of polynomials are the same over all the fields.*

Remark 2.5.4. *The same argument works as well for* $\Lambda^{(k)}[\gamma]$, $1 \leq k < d$, *since each component of* $\Lambda^{(k)}[\gamma]$ *is a determinant of some components of the polynomial, giving rise to the more general equality:*

$$\Lambda^{(k)}[\gamma](s) = C_k \lim_{\lambda \to 0} \frac{\Lambda[\gamma](\lambda \mathbf{z} + s)}{v(\lambda \mathbf{z})} \tag{2.23}$$

for $\mathbf{z}$ *in* $\mathbb{K}^k$ *without repeated components.*

The first relevant example of Lemma 2.5.2 above is the case where $\mu_{\mathbf{n}}$ is a moment curve. In this case (see Example 2.4.2) Lemma 2.5.2 implies that the torsion $\Lambda^{(d)}[\mu_{\delta(\mathbf{n})}(z)]$ is (up to a constant depending on $\mathbf{n}$ and d) equal to the Schur polynomial $S_{\delta(\mathbf{n})}(z, \ldots, z)$ (where, again $\delta(\mathbf{n}) = (\mathbf{n}_1 - 1, \mathbf{n}_2 - 2, \ldots, \mathbf{n}_d - d)$.

Since the Schur polynomials are a sum of monomials with coefficient 1, we can apply a reverse triangle inequality (Lemma A.1.6) to them. This, together with the compactness result in Lemma A.1.4, shows the following result holds:

Lemma 2.5.5. *Let* $\mathbb{K}$ *an admissible field. Let* $S_{\mathbf{n}}$ *be a Schur polynomial in* d *variables. Then one may cover* $\mathbb{K}^{\times}$ *with a finite number of sectors* $\mathbb{K} = \bigcup_{i=1}^{m} \Sigma_i$ *of amplitude* $\epsilon = \epsilon(\mathbb{K}, S_{\mathbf{n}})$ *such that if* $z \in \Sigma_i^d$

$$|S_{\mathbf{n}}(z)| \approx_{n,d} S_{\mathbf{n}}(|z|) \tag{2.24}$$

We now have all the tools we need to prove Theorem 2.1.13 part 3 when γ is a generalized moment curve:

Proof (of (DW)*, generalized moment curve case).* Let μ be a generalized moment curve of exponents $\mathbf{n}$. Decompose $\mathbb{K} = \bigcup W_i$ into finitely many sectors $\Sigma_i = \{z : |\arg z - \theta_i| < \epsilon\}$ of aperture ϵ small enough (depending on the exponents) centered at 0. Local compactness ensures that only finitely many sectors are needed. Now, for $\mathbf{z} = (\mathbf{z}_1, \ldots, \mathbf{z}_d) \in W_i^d$

$$|S_{\mathbf{n}}(\mathbf{z})| \gtrsim_{\mathbf{n}} S_{\mathbf{n}}(|\mathbf{z}|) \gtrsim_{\mathbf{n}} |\mathbf{z}_1 \cdot \mathbf{z}_2 \cdot \dots \cdot \mathbf{z}_d|^{\frac{\deg S_{\mathbf{n}}}{d}} = K_{\mathbf{n}} \left| \prod_{i=1}^{d} S_{\mathbf{n}}(\mathbf{z}_i, \dots, \mathbf{z}_i) \right|^{1/d} \tag{2.25}$$

for some $K_{\mathbf{n}} > 0$. The first inequality is Lemma 2.5.5, the second one AM-GM inequality for all the monomials of $S_{\mathbf{n}}(|\mathbf{z}|)$. The last equality follows from the fact that $S_{\mathbf{n}}(\mathbf{z}_i, \dots, \mathbf{z}_i) = C_{\mathbf{n}} \mathbf{z}_i^{\deg S_{\mathbf{n}}}$ for some positive integer $C_{\mathbf{n}}$. Now the result follows from equation (2.19) from Lemma 2.5.2 (which states that $\Lambda_\gamma^{(d)}(s) = C_d S_{\mathbf{n}}(s, \dots, s)$). □

Lemma 2.5.2 motivates the definition of corrected multilinear forms Definition 2.4.3. The first application of Definition 2.4.3 is a property of the generalized moment curve that will be used in the sequel:

Lemma 2.5.6 (Transversality of the corrected multilinear form for moment curves)**.** *Let μ be a moment curve, and let W be a sector of an admissible field $\mathbb{K}$ of aperture $\epsilon > 0$ (depending on μ, $\mathbb{K}$) small enough. Let $\{\mathbf{w}^{(j)}\}_{j=1}^{\infty}$ be a sequence of elements $\mathbf{w}^{(k)} = (\mathbf{w}_1^{(j)}, \dots, \mathbf{w}_s^{(j)})$ in W^s, let $\{\mathbf{z}^{(j)}\}_{j=1}^{\infty}$ a sequence in W^t, with $k := s+t \leq d$, assume $|\mathbf{z}_i^{(j)}| = O(1)$, and $\mathbf{w}^{(j)} \to 0$. Then*

$$\|\tilde{\Lambda}[\mu](\mathbf{z}_1^{(j)}, \dots, \mathbf{z}_t^{(j)}, \mathbf{w}_1^{(j)}, \dots, \mathbf{w}_s^{(j)})\| \approx_\mu \|\tilde{\Lambda}[\mu](\mathbf{z}_1^{(j)}, \dots, \mathbf{z}_t^{(j)})\| \|\tilde{\Lambda}[\mu](\mathbf{w}_1^{(j)}, \dots, \mathbf{w}_s^{(j)})\| \tag{2.26}$$

as $j \to \infty$.

From a geometric point of view, Lemma 2.5.6 states that if $z_1, \dots z_k$ and $w_1, \dots w_s$ (with $k + s \leq d$) belong to very different scales, then the linear spaces $\langle \mu'(z_1), \dots, \mu'(z_k) \rangle$ and $\langle \mu'(w_1), \dots, \mu'(w_k) \rangle$ must be transverse to each other.

Proof. The $\lesssim_\mu$ direction in (2.26) follows from the fact that, for forms $\|a \wedge b\| \leq \|a\| \|b\|$.

For the converse, let $e_{n_1} \wedge e_{n_2} \wedge \dots \wedge e_{n_k}$ be one of the co-ordinates on the LHS of (2.26). By restricting the curve to those co-ordinates, it can be assumed that $k = d$ (and the result will follow by summing for each co-ordinate by the triangle inequality). The term

$\|\tilde{\Lambda}[\mu](z_1^{(k)}, \ldots, z_t^{(k)})\|$ is a the absolute value of a Schur polynomial with all coefficents on a sector of small aperture (and therefore is $O(1)$ by 2.5.5), so we can omit it in the estimates.

By using the Young tableau decomposition of the Schur polynomials again, it suffices to show that each monomial in any of the coordinates $\tilde{\Lambda}[\mu](w_1^{(k)}, \ldots, w_s^{(k)})$ is dominated by a monomial of $\tilde{\Lambda}[\mu](z_1^{(k)}, \ldots, z_t^{(k)}, w_1^{(k)}, \ldots, w_s^{(k)})$. By positivity (using the same arguments as in Lemma 2.5.5) we can assume without loss of generality that all z_i, w_i are positive rational numbers using the *reverse triangle inequality* of Lemma 2.5.5.

Before showing the domination at a monomial leve, we motivate the proof by an example.

Example 2.5.7. *Assume* $t = 2, s = 3$, *and* $\gamma = (t, t^2, t^4, t^6, t^7)$. *That means the associated tableaux will have shape* $(0, 0, 1, 2, 2)$.

Let's look at the element $e_2 \wedge e_3 \wedge e_4$. *The curve* γ *restricted to the co-ordinates* $(2, 3, 4)$ *is* (t^2, t^4, t^6) *that has associated tableaux* $(1, 2, 3)$. *The important fact here is that elementwise* $(1, 2, 3)$ *is bigger than the smallest three rows of the diagram for the full dimension* $(0, 0, 1)$. *Therefore, given a tableau (for example):*

w_3	w_2	w_1
w_2	w_1	
w_1		

we can now remove the shaded squares to turn it into a $(0, 0, 1)$ *tableau. We can extend this tableau to a* $(0, 0, 1, 2, 2)$ *tableau with an associated monomial that dominates the monomial* $w_1^3 w_2^2 w_3$*:*

z_2	z_2
z_1	z_1
w_3	

The general proof is analogous to the example: given two Young diagrams T, T' of the same height (counting the rows of length zero) we say $T \leq T'$ if each row of T has at most as many elements as the row in T'. Let T be the diagram coming from a component $e_{i_1} \wedge \cdots \wedge e_{i_t}$ (with row lengths $\mathbf{n}_{i_1} - 1, \dots \mathbf{n}_{i_t} - t$) and let T' be the bottom t rows of the diagram associated to the full determinant $e_1 \wedge \cdots \wedge e_k$ (which has lengths $\mathbf{n}_1 - 1, \dots \mathbf{n}_t - t$). The strict monotonicity of n_i implies that $T \leq T'$. Therefore there is a a restriction of tableaux in T' to tableaux in T by removing the extra elements. Now we can turn any tableau in T into a tableau in the full diagram with lengths $\mathbf{n}_1 - 1, \dots \mathbf{n}_k - k$ by filling each new row $i + t$ with w_i.

□

2.6 Relationship between the decomposition and the torsion

The aim of this section is to show that Theorem 2.4.5 holds, using a compactness argument.

Lemma 2.6.1 (Convergence to the model in the non-degenerate set-up)**.** *For any* $0 < k \leq d$ *the function* $\tilde{\Lambda}[\mu](\mathbf{z})$ *is continuous in* $(\mathbf{z}, \mu) \in \mathbb{K}^k \times P_n(\mathbb{K})^d$*, where* $P_n(\mathbb{K})$ *denotes the set of polynomials of degree at most* n*.*

Proof. Consider both the numerator and denominator of $\tilde{\Lambda}[\mu](\mathbf{z})$ as a polynomial in the components of μ and z. The polynomial $\Lambda[\mu](\mathbf{z})$ on the numerator vanishes on the zero set $\mathcal{Z}(v(z_1, \dots, z_k))$, and since $v(z_1, \dots, z_k)$ splits into linear factors which are not repeated, $v(z_1, \dots, z_k)$ must divide the numerator by the Nullstellensatz. □

From this lemma one could directly deduce a local version of Proposition 2.4.7 around points where the Jacobian does not degenerate:

Proposition 2.6.2. *Let* γ *be a polynomial curve in* $\mathbb{K}^d$ *in canonical form at* 0*, such that* $\Lambda^{(d)}[\gamma](0) \neq 0$*. Then there is a neighborhood* $B_\epsilon(0)$*, with* $\epsilon = \epsilon(\gamma) > 0$ *where* (DW) *holds with constant depending only on the dimension and* $\mathbb{K}$*.*

Proof. By the affine invariance of (DW), assume without loss of generality that γ is at canonical form at zero. Consider a sequence $\gamma_i(z) := \text{diag}(\lambda_i^{-1}, \dots \lambda_i^{-d}) \circ \gamma(\lambda_i z)$ of zoom-ins (see section 2.1.2 for further details on zoom-ins) of γ parametrized by λ_i that converge to the moment curve. Therefore, it will suffice to show that, for some λ_i small enough, the Proposition 2.6.2 is true for γ_i in $B_1(0)$ that is:

$$|\tilde{\Lambda}[\gamma_i](z_1, \dots, z_d)| = \left| \frac{\Lambda[\gamma_i](z_1, \dots, z_d)}{v(z_1, \dots, z_d)} \right| \gtrsim_N \prod_{i=1}^{d} \Lambda^{(d)}[\gamma](z_i)^{1/d} \tag{2.27}$$

for $\lambda \leq \epsilon$ small enough and $z_i \in B_1(0)$. By re-scaling back (undoing the zoom-in) equation (2.27) will imply the inequality (DW) for γ and $B_\epsilon(0)$.

For the moment curve (the case $\lambda \to 0$) inequality 2.27 is true, and reads:

$$\tilde{\Lambda}[\mu](z_1, \dots, z_d) \gtrsim 1. \tag{2.28}$$

Since both sides of the inequality converge locally uniformly as $\lambda \to 0$ (the LHS by Lemma 2.6.1 and the RHS because it is the d^{th} root of a sequence of converging polynomials), the inequality is true for λ small enough in the zoom-in. □

A similar (albeit more careful) argument can be used to show that the inequality holds near a point where the torsion vanishes. That, however, would not be enough to show uniform estimates, and instead one needs Theorem 2.4.5.

The proof of Theorem 2.4.5 uses sequential compactness, and therefore is easier to state the contrapositive of Lemma Theorem 2.4.5 instead. We may also assume, by scale invariance of the statements, that $R = 1$. That is, it remains to prove the following statement:

Lemma 2.6.3 (Theorem 2.4.5, convergence verison)**.** *Let $\gamma_m \to \mu$ be a sequence of polynomial curves in canonical form of degree $\mathbf{n}$, with $\mu = \mu_{\mathbf{n}}$ a nondegenerate generalized moment curve. Let r_m define a sequence of annuli $A_{r_m,1}$, so that γ_m is ϵ-close to μ in A_m. Let $\mathbf{z}_m \in (A_m \cap \Sigma)^k$, where Σ is a sector of small enough aperture depending of $m, d, \mathbb{K}$ only. Then:*

$$\lim_{m\to\infty}\left(\frac{\Lambda[\gamma_m](\mathbf{z}_m)-\Lambda[\mu](\mathbf{z}_m)}{|\Lambda[\mu](\mathbf{z}_m)|}\right)=0 \tag{2.29}$$

Proof. First note that it suffices to prove that Lemma 2.6.3 is true for a subsequence of the $(\gamma_m, r_m, \mathbf{z}_m)$. The claim will follow if we can prove that, for any fixed coordinate $e = e_{l_1}\wedge\cdots\wedge e_{l_k}$ we have:

$$\lim_{j\to\infty}\frac{\Lambda_{\gamma_m}(\mathbf{z}_m)|_e}{\Lambda_{\mu}(\mathbf{z}_m)|_e}=1 \tag{2.30}$$

using the notation $w|_e$ to denote the e^{th} co-ordinate of the form w. By restricting the problem to the co-ordinates $(e_{l_1},\dots,e_{l_k})$ we may assume $k=d$, and then it suffices to show, in the same set-up of the lemma, that:

$$\lim_{j\to\infty}\frac{\tilde{\Lambda}_{\gamma_m}(\mathbf{z}_m)}{\tilde{\Lambda}_{\mu}(\mathbf{z}_m)}=1 \tag{2.31}$$

We will prove this by induction. By passing to a subsequence if necessary, assume without loss of generality that the vector $\mathbf{z}_m$ has a limit. In the base case none of the components of $\mathbf{z}_m$ has limit zero. In that case, the denominator converges to a non-zero number (since the denominator is a Schur polynomial in the components of $\mathbf{z}_m$, which does not vanish on a small enough sector by Lemma 2.5.5) and the result follows.

First nontrivial case is when all the components go to zero. In this case, by doing a further zoom-in and passing to a further subsequence if necessary, one can reduce to the case where not all the components of $\mathbf{z}_m$ go to zero. Thus assume that some (but not all) of the components of $\mathbf{z}_m$ go to zero. The fact that the new, zoomed in r_n will still go to zero slower than the new $\max_{j<n_i}(|M[\gamma_m]_{ij}|)^{1/(n_i-j)}$ is preserved by the zoom in procedure, and the fact that the $\max_{j>n_i}(|M[\gamma_m]_{ij}|)^{1/(n_i-j)}\to 0$ will still be true (at an even faster rate in fact).

Without loss of generality assume it is the first $0<k'<k$ components that go to zero. Let $\mathbf{z'}_m := ((\mathbf{z}_m)_1,\dots,(\mathbf{z}_m)_{k'})$ be the sequence made by the first k' components of each $\mathbf{z}_m$, and $\mathbf{z}''_m$ the sequence made by the remaining components. Then,

$$\frac{\tilde{\Lambda}[\gamma_m](\mathbf{z}_m)}{\tilde{\Lambda}[\mu](\mathbf{z}_m)} = \frac{\sum_{e' \wedge e'' = e} \tilde{\Lambda}[\gamma_m](\mathbf{z}'_m)|_{e'} \cdot \tilde{\Lambda}[\gamma_m](\mathbf{z}''_m)|_{e''}}{\sum_{e' \wedge e'' = e} \tilde{\Lambda}[\mu](\mathbf{z}'_m)|_{e'} \cdot \tilde{\Lambda}[\mu](\mathbf{z}''_m)|_{e''}} \tag{2.32}$$

we know by the induction hypothesis that each of the terms in the sum in the numerator converges to the corresponding term in the denominator (in the sense that their quotient goes to 1). Therefore, the result will follow if we can prove there is not much cancellation on the denominator, namely that:

$$\limsup_{j \to \infty} \frac{\sum_{e' \wedge e'' = e} |\tilde{\Lambda}[\mu](\mathbf{z}'_m)|_{e'} \cdot \tilde{\Lambda}[\mu](\mathbf{z}''_m)|_{e''}|}{|\tilde{\Lambda}[\mu](\mathbf{z}_m)|_e|} < \infty \tag{2.33}$$

but this is a consequence of Lemma 2.5.6, because we can bound each of the elements in the sum by $|\tilde{\Lambda}[\mu](\mathbf{z'}_m)| \cdot |\tilde{\Lambda}[\mu](\mathbf{z''}_m)| \lesssim |\tilde{\Lambda}[\mu](\mathbf{z}_m)|$, by equation (2.26).

□

2.7 Further properties arising from the partition

In this section we outline certain properties of the restriction of γ to its partition. These properties will be used in different parts the applications, but are written in further generality accounting for use cases in the litertaure of the original theorem of Dendrinos-Wright not showcase in the applications. In all the theorems below, I_j, ϵ_D are given in the context of Theorems 2.1.13, 2.4.5, and Proposition Proposition 2.4.7.

2.7.1 Sum-maps:

An recurrent object in the literature is the sum-map

$$\Phi[\gamma](t_1, \ldots, t_d) := \sum_{j=1}^{k} \epsilon_j \gamma(t_j), \qquad \epsilon_i = \pm 1$$

or variations of it. This map usually appears in the context of change of variables for integrals, and therefore one is interested in bounding the number of pre-images of $\Phi[\gamma]$. The following proposition guarantees that, outside of a certain lower-dimensional variety that can be neglected in applications, the number of solutions is uniformly bounded:

Proposition 2.7.1. *Assume $\epsilon(\mathbb{K}, d, N)$ in the statement of Theorem 2.4.5 is chosen small enough. Then for each sector Σ_j in Theorem 2.1.13, any $1 \leq k \leq d$ there exists a lower-dimensional variety $B_{j,k} \subset (\mathbb{K})^k$ with the following property: For any $\epsilon \in \{-1, 1\}^k$, define $\Phi^\epsilon[\gamma](t_1, \ldots, t_d) := \sum_{j=1}^k \epsilon_j \gamma(t_j)$. Then for each $x \in \mathbb{K}^d$, the cardinality of*

$$\big((\Phi^\epsilon[\gamma])^{-1}(\{t\}) \cap \Sigma_j^k\big) \setminus B_{j,k}$$

has cardinality $O_{d,N}(1)$.

Proof. The set $B_{j,k}$ will be the zero locus of the Vandermonde determinant. The key realization is that Φ^ϵ is a map from $\mathbb{K}^n$ to itself, and that the Jacobian determinant $|J(\Phi^\epsilon)(z_1, \ldots z_d)| = \Lambda[\gamma](z_1, \ldots z_d)$. By the inverse function theorem this implies that any nontivial variety of $|[\Phi^\epsilon]^{-1}(s)|$ must be contained in the zero locus of the Vandermonde determinant, which has codimension one.

In particular, by Bézout's theorem, there's at most finitely many solutions (depending on the dimensions and degrees) to $J(\Phi^\epsilon)(z_1, \ldots z_d) = S$ outside of $B_{j,k}$. ☐

2.7.2 Local features of γ in the decomposition sets

A monomial curve $\mu_{\mathbf{n}}$ behaves locally like a moment curve far from the origin ($t = 0$). This fact can, under certain circumstances (such as in [68, 75]), be used to extend a result known only for curves which are $\epsilon-$close to a moment cuurve to a general senario, using orthogonality and Littlewood–Paley to deal with the situation near the degeneracy point. The following proposition states this fact in sufficient generality for our purposes

Lemma 2.7.2. *Let $\tilde{\mu}$ be a monomial curve of maximum degree $\leq N$, let γ be ϵ_0-close to $\tilde{\mu}$ (for ϵ_0 small enough depending on N) in $A_{[r,R]}$ with $R = 2r \sim 1$. Let $\epsilon' > 0$. Then there* *exsists $\rho \in \mathbb{K}$ (depending on all the variables but γ) such that for each point $s \in A_{[r,R]}$, there is an affine transformation A_s such that $\gamma_s = A_s \circ \gamma(\rho x + s)$ is ϵ'-close to the standard moment curve μ in $A_{[0,1]}$.*

Proof. Define $\tilde{M}[\gamma]$ by $M[\gamma] = M[\mu] + \epsilon_0 \tilde{M}[\gamma]$. By construction (by the definition of ϵ_0-closeness, $\|\tilde{M}[\gamma]\| \lesssim 1$). Define the $N \times N$ matrix:

$$(L_s)_{ij} = \binom{j}{i} s^i \tag{2.34}$$

so that $M[\gamma(\cdot + s)] = M[\gamma] \cdot L_s$. Define the matrix

$$(V_s) = \begin{pmatrix} \binom{n_1}{0} s^d & \dots & \binom{n_1}{d} s^d \\ \vdots & \ddots & \vdots \\ \binom{n_d}{0} s^d & \dots & \binom{n_d}{d} s^d \end{pmatrix} \tag{2.35}$$

as the first d rows of $M[\mu] \cdot L_s$. This matrix is invertible if and only if $\mu(\cdot + s)$ has a canonical form at zero of degrees $(1, 2, \dots, d)$, which, by part 3 of Lemma 2.1.8 happens only if $\Lambda^{(d)}[\mu](s) \neq 0$, which holds as long as $s \neq 0$ because $\Lambda^{(d)}[\mu](s)$ is a monomial.

In particular, the map $s \mapsto (V_s)^{-1}$ is continuous on $\mathbb{K} \backslash 0$, and by compactness bounded in $A_{[r,R]}$. Moreover, by construction, the map $(V_S)^{-1}$ turns $\mu(\cdot + s)$ into a curve with canonical form at zero, and we can write $(V_s)^{-1} M[\mu] L_S = [\mathrm{Id}_{d\times d} | \tilde{M}_s]$ (with $\|\tilde{M}_s\| \lesssim 1$). This lets us write:

$$(V_s)^{-1} \cdot M[\gamma] \cdot L_s = [\mathrm{Id}_{d\times d}, 0] + \epsilon_0 (V_s)^{-1} \cdot \tilde{M}[\gamma] \cdot L_s. \tag{2.36}$$

Split $(V_s)^{-1} \cdot \tilde{M}[\gamma] \cdot L_s = [M_1 | M_2]$, with M_1 square, and $\|M_1\|, \|M_2\| \lesssim 1$. Choose ϵ_0 small enough to ensure that $(1 + M_1)$ is invertible and $\|(1 + M_1)^{-1}\| \lesssim 1$. Then

$$(1 - M_1)^{-1} (V_s)^{-1} \cdot M[\gamma] \cdot L_s = [Id | \hat{M}] \tag{2.37}$$

for some $\hat{M}$ with $\hat{M} \lesssim 1$. By the definition of L_s, this means that:

$$(1 - M_1)^{-1}(V_s)^{-1} \cdot M[\gamma(\cdot + s)] = [Id|\hat{M}] \tag{2.38}$$

Now choosing ρ with $|\rho|$ small enough shows the result. □

The nondegenerate of the lemma above gives a self-improving result for $\epsilon-$similarity:

Lemma 2.7.3. *Let γ be ϵ_0-close to the moment curve μ (for ϵ_0 small enough depending on N) in $A_{[0,R]}$ with $R =\sim 1$. Let $\epsilon' > 0$. Then there exsists $\rho \in \mathbb{K}$ (depending on all the variables but γ) such that for each point $s \in B_R$, there is an affine transformation A_s such that $\gamma_s = A_s \circ \gamma(\rho x + s)$ is ϵ'-close to the standard moment curve μ in $A_{[0,1]}$.*

CHAPTER 3

Uniform estimates for perturbations of monomial curves

3.1 Uniform restriction near monomial curves

Fix some N, d for the section. Let ϵ, ϵ_0 be small constants depending on N, d. We will denote by $\mathcal{R}(p \to q, \mu)$ the norm of restriction operator from $L^p(\mathbb{K}^d, dx)$ to $L^q(\mathbb{K}^d, \mu)$. A uniform restriction theorem then corresponds to uniformly bounding $\mathcal{R}(p \to q, \lambda_\gamma)$. The triangle inequality ensures that

$$\mathcal{R}(p \to q, \mu_1 + \mu_2) \leq \mathcal{R}(p \to q, \mu_1) + \mathcal{R}(p \to q, \mu_2).$$

This fact, together with the decomposition theorem (Theorem 2.1.13), shows that uniform restriction (Theorem 1.2.4) will follow from a perturbative estimate of the form:

Proposition 3.1.1. *There exists $\epsilon_0(N, d, \mathbb{K})$ so that the following holds. Let γ is a polynomial curve of degree at most N in $\mathbb{K}^d$ that is ϵ_0-similar to a monomial curve μ on $A_{[r,R]}$. Let $\theta \in \mathbb{K}$. Then, for (p, q) satisfying*

$$p' = \frac{d(d+1)}{2} q, q > \frac{d^2 + d + 2}{d^2 + d}. \tag{3.1}$$

it holds that $\mathcal{R}\left(p \to q; \lambda_\gamma \mathbf{1}_{\theta\Sigma_\epsilon \cap A_{[r,R]}})\right) < C_{p,q,\mathbb{K}}$

Note that by *rotation* invariance[1] we can assume $\theta = 1$. The key of the proof is to split $A_{[r,R]}$ into dyadic scales, which are then joined back using Littlewood–Paley, and a

[1] In the case $\mathbb{K} = \mathbb{C}$ one can assume $\theta \in S^1$, and this statement is indeed rotation invariance. For other fields, it can be understood as a particular case of $\mathbb{K}$-*dilation* invariance.

transversality argumetn of Stovall. By scale invariance of the problem, after dyadic splitting one may assume that $R = 2r \sim 1$. In order to further reduce the problem, we will further partition $A_{[r,R]} \cap \Sigma_\epsilon$ into small balls where γ is very close to the regular moment curve, following Lemma 2.7.2. This process will reduce Theorem 3.1.1 to a uniform local restriction estimate for parturbations to the standard moment curve.

Following this idea, the first step will be to show that Theorem 3.1.1 holds on each of those balls, or, equivalently, that it holds when the monomial curve μ is exactly the moment curve. We will do so following the *method of offspring curves*.

For a vector $h \in \mathbb{K}^K$, we define an associated offspring curve as $\gamma_h(x) = \frac{1}{K}\sum_{i=1}^{K} \gamma(x+h_i)$. The proof of Theorem 3.1.1 is inductive in nature, and uses certain restriction estimates for the curves γ_h in order to prove estimates for γ. In order to do so uniformly, we msut see that the hypotheses of Theorem 3.1.1 are preserving when changing γ to γ_h.

Lemma 3.1.2. *Let ϵ_0 small enough depending on N. Let $\mu(x) = (x, x^2, \dots x^d)$ be the moment curve, and let γ be ϵ_0-close to in $B_1(0) = A_{[0,1]}$. Then there exsists $\delta \in \mathbb{R}_{>0}$ with the following property:*

For any $x \in B_1(0)$, any $h_1, \dots h_K \in B_\delta(x)$, the curve $\gamma_h(z) = \frac{1}{K}\sum \gamma(z + h_i)$ is (after an affine transformation) ϵ_0-close to μ in $A_{[0,\delta]} = B_\delta(x)$. In particular, if ϵ_0 is small enough, by Proposition 2.4.7 the estimate (DW) *holds in $B_\delta(x)$ for γ_h.*

Proof. By Lemma 2.7.3, γ is ϵ-similar to μ in a δ_0 neighborhood of x. After translating and rescaling γ, we may therefore assume $x = 0$.

Let $\gamma_{h_i} = \gamma(\cdot + h_i)$. Let $M_0[\gamma] := M[\gamma_{h_i}][1 : d, 1 : d]$ the $d \times d$ leftmost matrix of $M[\gamma]$. If $|h_i| < \delta$ is small enough, $\|M_0[\gamma_{h_i} - Id]\| < O(\delta)$, and by the triangle inequality, $\|M_0[\gamma_h - Id]\| < O(\delta)$, and thus $\|M_0[\gamma_h]^{-1}\| = 1 + O(\delta)$. The matrix $M_0[\gamma_h]^{-1}M[\gamma_h]$ is equal identity on its fist d columns, and therefore $M_0[\gamma_h]^{-1}\gamma_h$ is ϵ_0-close to the identity. □

We have all the tools to prove 3.1.1 in the case when $R = 2r \sim 1$. From here on the proof proceeds as in [68], up to minor necessary technical modifications. We provide an outline for completeness.

3.1.1 Proof of Theorem 3.1.1 when $R \sim 2r$ (Local estimate)

This proof follows closely the proof in [68]. The proof is a combination of a local uniform proof for perturbations of the moment curve, which follows the original proof of Drury. Drury's proof was generalized in the work of Hickman [42] for the exact case of the moment curve, but essentially works for perturbations of the moment curve as well, as presented here.

In order to stay closer to the notation in [68], which we follow closely, we let $L_\gamma := \Lambda^{(d)}[\gamma](z)$, $J_\gamma(z_1, \dots z_d) = \Lambda[\gamma](z_1, \dots z_d)$, and $\lambda_\gamma = L_\gamma^{\frac{2d_{\mathbb{K}}}{d^2+d}}$. As a reminder, $d_{\mathbb{K}}$ (c.f. section 1.1.2.3) is the doubling exponent associated to the Haar measure on $\mathbb{K}$, so that $m_{\mathbb{R}} = 1$, $d_{\mathbb{C}} = 2$.

Lemma 3.1.3 (Bootstrapping for the extension operator)**.** *Let* $1 \leq p_0 < \frac{d^2+d+2}{2}$*,* ϵ_0 *small enough, and assume that there is a constant* $C_{d,N,p}$ *such that:*

$$\mathcal{R}^* \left(p_0 \to \frac{d(d+1)}{2} p_0'; (\lambda_\gamma)_*(B_r \right) < C_{d,N,p} \tag{3.2}$$

for all curves γ ϵ_0*-close to the moment curve* μ *on* $B_r(0)$ *(with* $r \leq 1$*) then, for all* p *satisfying* $p^{-1} > \frac{2}{d(d+2)} + \frac{d-2}{d(d+2)} p_0^{-1}$ *there exists a* $C'_{N.d,p}$ *so that*

$$\mathcal{R}^* \left(p \to \frac{d(d+1)}{2} p'; (\lambda_\gamma)_*(B_r) \right) \leq C'_{d,N,p} \tag{3.3}$$

for all curves γ ϵ_0*-close to the moment curve* μ *on* $B_r(0)$ *(with* $r \leq 1$*).*

Sketch of the proof. The proof follows the argument originally in [26], with a notation much closer to a sketch of the same proof provided in [68]. For the exact case of the moment curve, this proof can be found in [2], and for $\mathbb{Q}_p$ in [42]. This is a minor variation of these proofs, where we consider perturbations of the moment curves as well.

The strategy of the proof is to use the convolution-product relationship of the Fourier transform. Proving the theorem for all offspring curves simultaneously will allow us to split the convolution as an integral over offspring curves, where we will use the hypotheses.

By scale invariance of the problem, we will assume $r \sim 1$. We will restrict ourselves to prove a weaker estimate, namely

$$\mathcal{R}^*\left(p \to \frac{d(d+1)}{2}p'; (\lambda_\gamma)_*(B_\delta(x))\right) \leq C'_{d,N,p} \tag{3.4}$$

where $B_\delta(x)$ is any of the balls arising from 3.1.2. By covering $B_1 = \bigcup_{i=1}^{N_\delta} B_\delta(x_i)$, the result will follow.

Let $\tilde{\gamma}$ be ϵ_0-close to a moment curve, and $f \in L^p(\lambda_\gamma)$. Since by hypothesis $L_\gamma \approx 1$, we can neglect the factor of $|L_\gamma|^{\frac{2d_{\mathbb{K}}}{d^2+d}}$ in λ_γ, and replace norms of the form $\|f\|_{L^s(\lambda_\gamma)}$ with the unweighted $\|f\|_{L^s}$. Moreover since $J_\gamma(z_1, \dots z_d) \approx v(z_1, \dots z_d)$, we will exchange them freely.

If we define $g(\xi) := (\mu * \dots_{d \text{ times}} \dots * \mu)(d \cdot \xi)$, a change of variables computation shows:

$$g\left(\frac{1}{d}\sum_{i=1}^{d} \tilde{\gamma}(t_d)\right) = \frac{c_d}{|J_{\tilde{\gamma}}(t_1, \dots t_d)|^{d_{\mathbb{K}}}} f(t_1) \dots f(t_d). \tag{3.5}$$

This motivates the definition (in order to use the offspring curve hypothesis) for $h = (0, h') \in \{0\} \times B_\delta(x)^{d-1}$:

$$G(t; h) := g\left(\frac{1}{d}\sum_{i=1}^{d} \gamma(t + h_i)\right). \tag{3.6}$$

We can write now

$$\hat{g}(x) = C_d \int_{\mathbb{C}^{d-1}} \int_{\tilde{B}_{(0,h')}} e^{ix\tilde{\gamma}_h(t)} g(\tilde{\gamma}_h(t)) |J_{\tilde{\gamma}}(t + h_1, \dots t + h_d)|^{d_{\mathbb{K}}} dt dh' \tag{3.7}$$

From here, by Plancherel we obtain:

$$\|\hat{g}\|_q \lesssim \|G\|_{L^2_{h'}}(L^q_t; |v(0,h')|^{d_{\mathbb{K}}}) \tag{3.8}$$

Now we can use the induction hypothesis because $\tilde{\gamma}$ is ϵ_0-close to the identity on $B_\delta(x)$ (and Hölder's inequality - using that $|\tilde{B}_1| \lesssim 1$) we get:

$$\|\hat{g}\|_q \lesssim \|G\|_{L^1_{h'}(L^p_t; |v(0,h')|^2)}, \ \ 1 \le p \le p_0, q = \frac{d(d+1)}{2} p. \tag{3.9}$$

Interpolating between these results, we obtain:

$$\|\hat{g}\|_q \lesssim \|G\|_{L^a_{h'}(L^b_t; |v(0,h')|^{d_{\mathbb{K}}})} \tag{3.10}$$

for (a^{-1}, b^{-1}) in the triangle with vertices $(1/2, 1/2), (1, 1, (1, p_0^{-1}))$. Therefore it suffices to bound:

$$\|G\|_{L^a_{h'}(L^b_t; |v(0,h')|^2)} \sim \left[\int dh' |v(h)|^{d_{\mathbb{K}}(a-1)} \left(f(t+h_1)\right)^{a/b} \right]^{1/a} \tag{3.11}$$

where we used that $v(t+h_1, \dots t+h_d) = v(h_1, \dots h_d)$ to take the $|v(h)|^{-m_{\mathbb{K}}}$ term from the inner to the outer integral. Now, an explicit computation shows that $|v(h)|^{d_{\mathbb{K}}(a-1)} \in L^q_{h'}$ for $q < \frac{2}{d(a-1)}$ whenever $0 < a-1 < \frac{2}{d}$ (An application of Lemma 3.1.7 with $p = \infty$ and $f_i = 1$ can be used as well). Interpolating and using Hölder, one can see that:

$$\|G\|_{L^a_{h'}(L^b_t; |v(0,h')|^2)} \lesssim \|f\|^d_{L^p_t} \tag{3.12}$$

whenever $a < b < \frac{2a}{d+2-da}$, and $\frac{d}{p} < \frac{(d+2)(d-1)}{2} \frac{1}{a} + \frac{1}{b} - \frac{d(d-1)}{2}$.

Choosing (a^{-1}, b^{-1}) arbitrarly close to $\left(\frac{d}{d+2}, \frac{d}{d+2} + \frac{1}{p_0} \frac{d+2}{d-2}\right)$ and computing the resulting exponents finishes the proof.

□

We can now conclude the proof of the local restriction estimate 3.1.1 when $R \sim 2r$.

Proof of Theorem 3.1.1 when $R \sim 2r$. Without loss of generality, by scale invariance, $R \sim 2r \sim 1$. Iterating Lemma 3.1.3 (cf. [68]) shows that the conclusion of Lemma 3.1.1 holds on balls where γ is close to a moment curve. Lemma 2.7.2 lets us cover $A_{r,R} \cap \Sigma_\epsilon$ by $O_{\mathbb{K},N,d}(1)$ balls where γ is close to a moment curve, so 3.1.1 holds in each of these balls. By triangle inequality it holds in the whole of $A_{r,R} \cap \Sigma_\epsilon$. □

3.1.2 Almost orthogonality

Our goal is now to extend the result from annuli sectors of the form $A_{[r,\sim 2r]} \cap \Sigma_\epsilon$ to general annuli $A_{[r,R]} \cap \Sigma_\epsilon$ obtained from the decomposition in Theorem 2.1.13. A routine Littlewood–Paley application (which holds in local fields, see e.g [69, Theorem 1.6 on Chapter 6]) gives us the following:

Lemma 3.1.4. *Let $\gamma : \mathbb{K} \to \mathbb{K}^d$ be a complex polynomial curve, and let $T_n = A_{[r,R]} \cap A_{[2^n,2^{n+1}]} \cap \Sigma_\epsilon$ Then for each (p,q) satisfying $q = \frac{d(d+1)}{2}p'$ and $\infty > q > \frac{d^2+d+2}{2}$ and $f \in L^p(d\lambda_\gamma)$ we have:*

$$\|\mathcal{E}_\gamma(\chi T_j f)\|_{L^q(\mathbb{R}^{2d})} \lesssim \left\| \left(\sum_n |\mathcal{E}_\gamma(\chi T_n f)|^2 \right)^{1/2} \right\|_{L^q(\mathbb{R}^{2d})} \tag{3.13}$$

In order to be able to sum the different pieces $\mathcal{E}_\gamma(\chi_{T_{j,n}} f)$ one must bound the interactions arising from the different terms. This idea encapsulated in the following multilinear restriction Lemma, analogous to Lemma 4.1 in [68]. Let, to ease the notation, $D = \frac{d^2+d}{2}$. Assume as well that $q \leq 2D$ (we can then extend to the remaining cases by interpolation with the $L^1 \to L^\infty$ trivial global result).

Lemma 3.1.5. *There exists $\epsilon(N,d,p) > 0$ such that, if $n_1 \leq \cdots \leq n_d$, with $n_d - n_1 > 2d$ and f_i is Schwartz and supported in T_{n_i} (as defined in Lemma 3.1.4) we have:*

$$\left\| \prod_{i=1}^d \mathcal{E}_\gamma[f_i] \right\|_{L^{q/D}} \lesssim 2^{\epsilon(n_D - n_1)} \prod_{i=1}^d \|f_i\|_{L^p(d\lambda)}. \tag{3.14}$$

for $2D \geq q = dp'$.

Before proving Lemma 3.1.5 we will finish the proof of Theorem 3.1.1.

$$\|\mathcal{E}_\gamma\|_{L^q}^q \lesssim \int \prod_{j=1}^{D} \left| \sum_{n_j} |\mathcal{E}_\gamma(\chi T_{n_j} f)|^2 \right|^{q/2D} dx \sim \sum_{n_1 \leq n_2 \leq n_D} \int \prod_{j=1}^{D} |\mathcal{E}_\gamma(\chi T_{n_j} f)|^{q/D} dx \tag{3.15}$$

This sets us precisely in the multilinear situation of Lemma 3.1.5, allowing us to see that:

$$\|\mathcal{E}_\gamma\|_{L^q}^q \lesssim \sum_{n_1 \leq \cdots \leq n_d} 2^{-\epsilon|n_d - n_1|} \prod_{j=1}^{D} \|\chi_{T_{n_j}} f\|_p^{q/D} \lesssim \sum_{m \geq 0} \sum_{\substack{n_1 \leq \cdots \leq n_d \\ n_d = n_1 + m}} 2^{-\epsilon m} \prod_{j=1}^{D} \|\chi_{T_{n_j}} f\|_p^{q/D} \tag{3.16}$$

Let us define $T_{[i,j]} = A_{[2^i, 2^j]} \cap A_{[r,R]} \cap \Sigma_\epsilon$.

$$\|\mathcal{E}_\gamma\|_{L^q}^q \lesssim \sum_{m \geq 0} 2^{-\epsilon m} \sum_{n} \|\chi_{T_{[n:n+m]}} f\|_p^q \lesssim \sum_{m \geq 0} 2^{-\epsilon m} m^{O_q(1)} \sum_{n} \|\chi_{T_{[nm:nm+1]}} f\|_p^q \tag{3.17}$$

since $p \leq q$, we can use that $\sum_n \|\chi_{T_{[nm:nm+1]}} f\|_p^q \leq \left(\sum_n \|\chi_{T_{[nm:nm+1]}} f\|_p^p \right)^{q/p}$, and by the disjointness of supports

$$\|\mathcal{E}_\gamma\|_{L^q}^q \lesssim \|f\|_p (\sum_m 2^{-m} m^{O(1)}) \lesssim \|f\|_p \tag{3.18}$$

Proof of Lemma 3.1.5. By Hausdorff-Young,

$$\left\| \prod_{i=1}^{d} \mathcal{E}_\gamma[f_i] \right\|_{L^{d+1}} \leq \|d\mu_1 * \ldots d\mu_d\|_{L^{\frac{d+1}{d}}} \tag{3.19}$$

where $d\mu_i := \gamma_*(f(t)\lambda_\gamma(t)dt)$, and, if we define $\Phi(\mathbf{t}) := \sum_{i=1}^d \gamma(t_i)$,

$$\begin{aligned} [d\mu_1 * \cdots * d\mu_d](\phi) = & \int_{\mathbb{C}^d} \phi\left(\sum_{i=1}^{d} \gamma(t_i) \right) \prod_{i=1}^{d} f_i(t_i)\lambda_\gamma(t_i) d\mathbf{t} \\ = & \Phi_* \left[\prod_{i=1}^{d} f_i(t_i)\lambda_\gamma(t_i) d\mathbf{t} \right] (\phi) \end{aligned}$$

If Φ was a one-to-one map (as we know it is in the real case) the change of variables rule would give a direct relationship between the L^p norm of the density associated to $d\mu_1 * \cdots * d\mu_d$ and weighted L^p norm of $\prod_{i=1}^d f_i(t_i)\lambda_\gamma(t_i)$. Whenever the map is $O(1)$-to-1, the following lemma serves as an alternative (see the appendix for a proof):

Lemma 3.1.6. *Let $\Omega \subset \mathbb{K}^k$ be an open set, and $\Phi : \Omega \to \mathbb{K}^k$ a smooth map, with non-vanishing Jacobian J_Φ. Assume that Φ is at most n-to-1 (i.e., each point has at most n pre-images). Let $f : \Omega \to \mathbb{K}^d$ be a test function, and let $\Phi_* f := \frac{d(\Phi_* f \cdot \mathrm{Haar}_{(\mathbb{K}^k)})}{d\,\mathrm{Haar}_{(\mathbb{K}^k)}}$ be the measure pushforward of f. Then,*

$$[\Phi_* f](x) = \sum_{\xi : \Phi(\xi) = x} f(\xi) |J_\Phi(\xi)|^{-1}, \ \text{and} \tag{3.20}$$

$$\|[\Phi_* f](x)\|_{L^p(dx)} \lesssim_n \|f |J_\Phi(\xi)|^{1/p-1}\|_{L^p(dy)} \tag{3.21}$$

.

Lemma 3.1.6 (with $\Omega = \prod T_{n_i} \setminus B$, B as in Proposition 2.7.1), together with (3.19) imply the estimate

$$\left\| \prod_{i=1}^d \mathcal{E}_\gamma[f_i] \right\|_{L^{d+1}} \lesssim \left\| \prod_{i=1}^d f_i(z_i)\lambda_\gamma(t_i) J_\gamma^{-\frac{d_{\mathbb{K}}}{d+1}}(\mathbf{z}) \right\|_{L^{\frac{d+1}{d}}(\mathbf{z})} \tag{3.22}$$

$$\lesssim \left\| \prod_{i=1}^d f_i(z_i)\lambda_\gamma(t_i)^{1/2} v(\mathbf{z})^{-\frac{d_{\mathbb{K}}}{d+1}} \right\|_{L^{\frac{d+1}{d}}(\mathbf{z})} \tag{3.23}$$

For l smooth functions $g_1, \dots g_l$ define[2]

$$T(g_1, \dots g_l) := \| \prod_{i=1}^l f_i(z_i)\lambda_\gamma(t_i)^{1/2} v(z_1, \dots z_l)^{-\frac{2}{d+1}} \|_{L^{\frac{d+1}{d}}(\mathbb{C}^l)}.$$

[2]Note that this definition differs from the definition given in [68] by a $\frac{d+1}{d}$ exponent. The definition given in [68] is $T_{\mathrm{Stovall}}(g_1, \dots g_l) := \| \prod_{i=1}^l f_i(z_i)\lambda_\gamma(t_i)^{1/2} v(z_1, \dots z_l)^{-\frac{2}{d+1}} \|_{L^{\frac{d+1}{d}}(\mathbb{K}^l)}^{\frac{d+1}{d}}$.

By the pigeonhole principle there is an index k such that $n_{k+1}-n_i \geq \frac{n_d-n_1}{d}$, and in particular, $n_{k+1}-n_i \geq 2$. In that case:

$$|v(\mathbf{z})| \sim \prod_{1\leq i<j\leq k} |t_i-t_j| \prod_{1\leq i\leq k<j} 2^{n_j} \prod_{1\leq i\leq k<j} |t_i-t_j| \tag{3.24}$$

since all the coupling between variables in (3.23) comes from the Vandermonde determinant, this allows us to split the norm in (3.23) as a product of two norms

$$T(f_1,\ldots,f_d) \lesssim 2^{-\frac{2k}{d+1}\sum_{i>k} n_i} T(f_1,\ldots f_k) T(f_{k+1},\ldots f_d). \tag{3.25}$$

In order to control this terms, following Stovall, we use a lemma by Christ. The original formulation of lemma considers functions only in the local domain, but the argument transfers without change to the domains considered in this work. We provide a sketch of the proof in the appendix for completeness:

Lemma 3.1.7. *Let $f_i, g_{i,j}$ be test functions, then:*

$$\int_{\mathbb{K}^l} \prod_{1\leq i\leq l} f_i(z_i) \prod_{1\leq i<j\leq l} g_{i,j}(z_i-z_j)dz \lesssim \prod_{i=1}^{l} \|f_i\|_p \prod_{1\leq i<j\leq l} \|g_{i,j}\|_{q,\infty} \tag{3.26}$$

whenever $2l = 2lp^{-1} + \frac{2l(l-1)}{2}q^{-1}$ and $p > l > 1$.

Now by Lemma 3.1.7 (see the appendix) and Hölder's inequality (using our control on the size of the supports of f_i) we can bound

$$T(f_1,\ldots f_k) \lesssim \prod_{i=1}^{k} \|f_i\|_{L^{\frac{2d+2}{2d-k+1}}} \lesssim \prod_{i=1}^{k} 2^{n_i\frac{d-k}{d+1}} \|f_i\|_2 \tag{3.27}$$

and

$$T(f_{k+1},\ldots f_d) \lesssim \prod_{i=k+1}^{d} \|f_i\|_{L^{\frac{2d+2}{d+k+1}}} \lesssim \prod_{i=k+1}^{d} 2^{n_i\frac{k}{d+1}} \|f_i\|_2. \tag{3.28}$$

Joining the estimates above, and using the hypothesis that $n_{k+1}-n_k \geq \frac{1}{d}(n_d-n_1)$, we get:

$$\left\|\prod_{i=1}^{d}\mathcal{E}_\gamma[f_i]\right\|_{L^{d+1}} \lesssim 2^{\frac{d-k}{d+1}(n_1+\dots n_k-n_{k+1}-\dots n_d)}\prod_{i=1}^{d}\|f\|_2 \tag{3.29}$$

$$\lesssim 2^{\epsilon_d(n_d-n_1)}\prod_{i=1}^{d}\|f\|_2 \tag{3.30}$$

which finishes the proof. □

3.2 Uniform endpoint restriction near real monomial curves

The goal of this section is to show the following result:

Proposition 3.2.1. *Let $\epsilon(d,N)>0$ be small enough. Let $d\geq 3$ and $0<r<R<\infty$. Let $\mathfrak{C}_{r,R,N,d}$ be the set of polynomial curves $\gamma:\mathbb{R}\to\mathbb{R}^d$ which are ϵ-similar to a monomial curve in (r,R). Then, for all $\gamma\in\mathfrak{C}_{r,R,N,d}$, the estimate*

$$\|R_{\gamma,dx}\|_{\mathrm{Op}\left(L^{p,1}(\mathbb{R}^d)\to L^{p,1}(\gamma;d\lambda_\gamma)\right)} < C_{N,d}. \tag{3.31}$$

Note that the bounding constant does not depend on (r,R), and in particular they be taken to zero or infinity by a limiting argument. In particular, the endpoint part of Theorem 1.2.4 follows from Proposition 3.2.1 and the decomposition theorem (Theorem 2.1.13). The proof of Proposition 3.2.1 follows from the one of the main results in [5]. In order to state that result, we need the following list of hypotheses

Hypotheses 3.2.2 (Hypotheses 4.1 in [5]). *Let $\mathfrak{C}$ be a class of curves with base interval $I_\circ$. For $\gamma\in\mathfrak{C}$ defined on $I\subset I_\circ$ let*

$$E=\{(t_1,\dots,t_d): t_1\in I,\ t_d\in I,\ t_1<t_2<\dots<t_d\}. \tag{3.32}$$

(1) There is $N_1\geq 1$ so that for every $\gamma\in\mathfrak{C}$ the map $\Phi_\gamma:E\to\mathbb{R}^d$ with

$$\Phi_\gamma(t_1,\dots,t_d)=\sum_{j=1}^{d}\gamma(t_j) \tag{3.33}$$

is of multiplicity at most N_1.

(2) Let J_{Φ_γ} denote the Jacobian of Φ_γ,

$$J_{\Phi_\gamma}(t_1, \dots, t_d) = \det\left(\gamma'(t_1), \dots, \gamma'(t_d)\right).$$

Then there is $c_1 > 0$ such that for every $(t_1, \dots, t_d) \in I^d$ with $t_1 < \dots < t_d$ we have the inequality

$$|J_{\Phi_\gamma}(t_1, \dots, t_d)| \geq c_1 \left(\prod_{i=1}^{d} \tau_\gamma(t_i)\right)^{1/d} \prod_{1 \leq j < k \leq d} (t_k - t_j). \tag{3.34}$$

(3) Every offspring curve of a curve in $\mathfrak{C}$ is (after possible reparametrization) the affine image of a curve in $\mathfrak{C}$.

(4x) There is $c_2 > 0$ so that for every $\gamma \in \mathfrak{C}$ and every offspring curve γ_κ of γ we have the inequality

$$|\tau_{\gamma_\kappa}(t)| \geq c_2 \max_{j=1,\dots,d} |\tau_\gamma(t + \kappa_j)| \tag{3.35}$$

where $\tau_\gamma := \det(\gamma'(x), \dots, \gamma^{(d)}(x))$

These are essentially the hypotheses a class of curves has to satisfy in order to satisfy uniform restriction estimates at the endpoint:

Theorem 3.2.3 ([5, Theorem 4.2])**.** *Let $\mathfrak{C}$ be a clas of curves satisfying hypotheses 3.2.2, and such that*

$$\sup_{\gamma \in C} \|R_{\gamma,dx}\|_{\mathrm{Op}\left(L^{p,1}(\mathbb{R}^d) \to L^{p,1}(\gamma; d\lambda_\gamma)\right)} < C_{\mathfrak{C}}. \tag{3.36}$$

Then, one has the bound

$$\sup_{\gamma \in C} \|R_{\gamma,dx}\|_{\mathrm{Op}\left(L^{p,1}(\mathbb{R}^d) \to L^{p,1}(\gamma; d\lambda_\gamma)\right)} < C_{d,N,1,C_1,C_2}. \tag{3.37}$$

In other words, if we show that the hypotheses are satisfied for $\mathbb{C}_{r,R,N,d}$ with constants independent of r, R, and can find bounds for the restriction operator on $\mathcal{C}_{r,R,N,d}$ depending on r, R, we can remove that dependency.

By Lemma 2.7.2, any polynomial curve of degree $\leq N$ that is $\epsilon-$similar to a monomial curve on (r, R) can be decomposed into $O_{r,R,N,d}(1)$ intervals where it is ϵ-similar to the standard moment curve. In particular, by the triangle inequality and the local estimates of [3] for the nondegenerate case, estimate (3.36) holds (with a constant depending on r, R, N, d). Therefore, the proof of Proposition 3.2.1 reduces to showing that Hypotheses 3.2.2 hold. We will in fact show that they hold for a slightly related class, the class $\tilde{\mathfrak{C}}_{N,d,r,R}$ of functions of the form $\gamma(\exp(t))$, for $t \in J \subset [\log r, \log R]$, where γ is a polynomial curve of degree at most N that is $\epsilon-$similar to a monomial curve, in J. This *exponential reparametrization trick* originates in the work of Drury and Marshall [28], and was used subsequently in Bak-Oberlin-Seeger to show Proposition 3.2.1 when γ is a monomial curve, or a *simple* curve of the form $(t, t^2, \ldots, t^{d-1}, p(t))$. Reparametrizing a curve does not change its affine restriction estimates, as the affine arclength is reparametrization invariant. It does, however, change its *offspring curves*, which become significantly more tractable.

3.2.1 Proof of Proposition 3.2.1

In this section we will show that Hypotheses 1-4 hold for the class $\tilde{\mathfrak{C}}_{N,d,r,R}$ if $\epsilon = \epsilon(N, d)$ (implicit in the definition) is small enough.

Fix a nondegenerate monomial curve $\mu(x) = (x^{n_1}, \ldots, x^{n_d})$, and let $\delta = \epsilon(\mu, N) > 0$, to be chosen later. Let γ be a polynomial curve. Let $M[\gamma]_{k,m}$ be the degree m coefficient of the $k-$th component of γ, γ_k (as i Section 2.1.2). For $a > b$, then, for a polynomial curve γ of degree at most N, $\gamma(e^t) : [\log a, \log b] \to \mathbb{R}^d$ belongs to $\tilde{\mathfrak{C}}_{N,d,r,R}$ if (See Definition 2.1.11):

1. If $m = n_k$, $M[\gamma]_{k,m} = 1$.

2. If $m < n_k$, $M[\gamma]_{k,m} \cdot a^{m-n_k} \leq \delta$

3. If $m > n_k$, $M[\gamma]_{k,m} \cdot a^{m-n_k} \leq \delta$

Hypothesis 1: Injectivity of the sum-map

Note that injectivity of the sum-map $\Phi_{\gamma(e^t)}$ is equivalent to injectivity of the sum-map Φ_γ, so we will show the later.

By Theorem 2.4.5, using that $\tilde{\Lambda}[(\gamma)_{\leq j}](\underbrace{x, x, \ldots, x}_{j \text{ times}}) = n!\tau_{(\gamma)_{\leq j}}(x)$ (see Lemma 2.5.2), we see that if ϵ is small enough, $|\tau_{(\gamma)_{\leq j}}(x) - \tau_{(\mu)_{\leq j}}(x)| \leq \frac{1}{10}|\tau_{(\mu)_{\leq j}}(x)|$. Since $\tau_{(\mu)_{\leq j}}(x)$ is a positive multiple of a monomial in $(0, \infty)$, $\tau_{(\gamma)_{\leq j}}(x)$ must be strictly positive for $x > 0$. This shows that all the top-left principal minors of the matrix $(\gamma', \gamma'', \ldots . \gamma^{(d)})$ are positive on $(0, \infty)$. This implies, by an result dating back to Steinig [65] (see [24, Proposition 6.1]) that Φ_γ is in fact injective.

Alternatively, one could use Bezout's theorem to bound the number of solutions.

Hypothesis 2: Jacobian inequality

Hypothesis 2 is essentially the content of the proof of Proposition 2.4.7. Using the transfer theorem (Theorem 2.4.5) one can immediately transfer the Jacobian-Torsion lower bounds for each interval from the analogous inequality for monomial curves:

$$
\begin{aligned}
\frac{J_\Phi(x_1, \ldots, x_d)}{v(x_1, \ldots x_d)} &= \tilde{\Lambda}[\gamma](x_1, \ldots x_d) && (3.38)\\
&\approx \tilde{\Lambda}[\mu](x_1, \ldots x_d) && (3.39)\\
&\gtrsim \prod_{k=1}^{d} \tilde{\Lambda}[\mu](x_k, \ldots x_k)^{1/d} && (3.40)\\
&\approx \prod_{k=1}^{d} \tilde{\Lambda}[\gamma](x_k, \ldots x_k)^{1/d} && (3.41)\\
&= d! \prod_{k=1}^{d} \tau_\gamma(x_k)^{1/d} && (3.42)
\end{aligned}
$$

The inequality between (3.39) and (3.40) follows from the explicit computation of $\tilde{\Lambda}[\gamma](x_1, \ldots x_d)$ as a Schur polynomial, as in (2.25).

Hypothesis 3. Offspring-stability.

Assume that the model monomial curve has degrees $(0 < n_1 < \cdots < n_d \leq N)$. Consider the polynomial curve

$$\gamma_{*h} := \mathrm{diag}_k \left(\sum_{i=1}^{s} \exp(n_k h_i) \right)^{-1} ((\gamma \circ \exp(t))_h) \circ (\log x). \tag{3.43}$$

The curve γ_{*h} is the result of, switching to exponential parametrization performing the offspring operation, returning to the regular parametrization, and normalizing so that the operation stays the identity for the moment curve. In other words, if A is the linear transform given by $A = \mathrm{diag}_k \left(\sum_{i=1}^{s} \exp(n_k h_i)\right)^{-1}$, and $(*)_\kappa$ denotes the offspring operation,

$$\gamma_{*h}(e^t) = A^{-1}(\gamma \circ e)_\kappa.$$

In particular, Hypothesis 3 is equivalent to showing that γ_{*h} is ϵ-similar to μ on its domain.

One has the explicit computation:

$$M[\gamma_{*h}]_{k,n} = M[\gamma_{*h}]_{k,n} \cdot \frac{\sum_{i=1}^{d} \exp(n h_i)}{\sum_{i=1}^{d} \exp(n_k h_i)}(t) \tag{3.44}$$

If the domain of γ was $[a, b]$, the domain of reparametrized gamma will be $[\log a, \log b]$ and the domain of the offspring curve will be $[\log a - \min h_i, \log b - \max h_i]$. The domain of γ_{*h}, back in regular reparametrization, will be the interval $[ae^{-\min h_i}, be^{-\max h_i}]$. In particular, $\epsilon-$similarity will follow the following two estimates for all k, n:

$$M[\gamma_{*h}]_{k,n}(be^{-\max h_i})^{n-n_k} \leq M[\gamma]_{k,n}(b)^{n-n_k} \qquad \text{if } n > n_k \tag{3.45}$$

and

$$M[\gamma_{*h}]_{k,n} \left(ae^{-\min h_i}\right)^{n-n_k} \leq M[\gamma]_{k,n} \left(a\right)^{n-n_k} \qquad \text{if } n < n_k. \tag{3.46}$$

By using (3.44), these are equivalent to

$$\frac{\sum_{i=1}^{d} \exp(n h_i)}{\sum_{i=1}^{d} \exp(n_k h_i)} e^{-(n-n_k) \max_{i=1,\dots d} h_i} \leq 1 \qquad \text{if } n > n_k \tag{3.47}$$

and

$$\frac{\sum_{i=1}^{d} \exp(n h_i)}{\sum_{i=1}^{d} \exp(n_k h_i)} (e^{-(n-n_k) \min_{i=1,\dots d} h_i}) \leq 1 \qquad \text{if } n < n_k. \tag{3.48}$$

respectively. Both inequalities are equivalent to

$$\frac{\sum_{i=1}^d \exp(nh_i)}{\sum_{i=1}^d \exp(n_k h_i)} \leq \max_{i=1\dots d} \frac{\exp(nh_i)}{\exp(n_k h_i)}. \tag{3.49}$$

To show this inequality, note that

$$\frac{\sum_{i=1}^d \exp(nh_i)}{\sum_{i=1}^d \exp(n_k h_i)} = \sum_{i=1}^n \alpha_i \frac{\exp(nh_i)}{\exp(n_k h_i)} \tag{3.50}$$

where

$$\alpha_j := \frac{\exp(n_k h_j)}{\sum_{i=1}^d \exp(n_k h_i)}, \qquad \sum_{j=1\dots d} \alpha_j = 1. \tag{3.51}$$

in particular, the left hand side of (3.49) is a convex combination of the terms $\frac{\exp(nh_i)}{\exp(n_k h_i)}$ appearing in the maximum of the right hand side of (3.49).

Hypothesis 4. Torsion lower bounds.

Keeping the definition $A := \mathrm{diag}_k\left(\sum_{i=1}^s \exp(n_k h_i)\right)$, one has, by an explicit computation for the moment curve $\mu(t) = (t^{n_1}, \dots, t^{n+d})$, that $A^{-1}\mu_{*h}(e^t) = \mu(e^t)$. In Hypothesis 3 we showed that $A^{-1}\gamma_{*h}$ was ϵ-similar to μ. Since reparametrizing $(t \mapsto e^t)$ and multiplying by A change the torsion by a constant, we get that, if ϵ is small enough, (again by Theorem 2.4.5)

$$\frac{1}{2} \leq \frac{\tau_{(\gamma\circ e^t)_\kappa}}{\tau_{(\mu\circ e^t)_\kappa}} \leq 2 \qquad \text{and} \qquad \frac{1}{2} \leq \frac{\tau_{\gamma\circ e^t}}{\tau_{\mu\circ e^t}} \leq 2.$$

In particular,

$$\tau_{(\gamma\circ e^t)_\kappa}(t) \gtrsim \tau_{(\mu\circ e^t)_\kappa}(t) \gtrsim \max_j \tau_{(\mu\circ e^t)}(t+\kappa_j) \gtrsim \max_j \tau_{(\gamma\circ e^t)}(t+\kappa_j) \tag{3.52}$$

where the middle inequality is an explicit inequality for monomial curves, see e.g. [5, Section 6].

This shows that endpoint weak restriction estimates hold independently of (r, R) for curves of $\tilde{\mathfrak{C}}_{N,d,r,R}$, and therefore in the union of all of them (including, by taking limits, the cases $r = 0$ or $R = \infty$). By Theorem 2.1.13, every polynomial curve can be decomposed into $C_{N,d}$ curves of this class, finishing the proof of the endpoint part of Theorem 1.2.4.

3.3 Uniform decoupling near real monomial curves

In [75], Yang showed that the constant C_ϵ in Theorem 1.2.5 can be taken uniform when decoupling over any polynomial curve γ of the form $(x, p(x))$ of bounded degree, and the center points (which in the moment curve are of the form $(\delta \cdot j)$) are adapted to the curve in a certain sense. We will generalize this to arbitrary dimension, and (in $d = 2$) to general polynomial curves.

As a first ingredient to the uniformity-type estimates we will show a series of upgrading steps to the decoupling theorem (Theorem 1.0.8). The first step is seeing that one may dilate the paralellepipeds $\{U_{x,\delta}(\gamma)\}_{x\in\delta\mathbb{Z}\cap[0,1]}$ to larger versions $\{U_{x,M\delta}(\gamma)\}_{x\in\delta\mathbb{Z}\cap[0,1]}$ by losing a factor of $M^{O(1)}$ in the estimate. This is because, if M is an integer

$$\{U_{x,M\delta}(\gamma)\}_{x\in\delta\mathbb{Z}\cap[0,1]} = \bigcup_{k=1,\dots M} \{U_{x,M\delta}(\gamma)\}_{x\in(M\delta+k\delta)\mathbb{Z}\cap[0,1]}. \tag{3.53}$$

and one can apply the decoupling estimate to each of those. This allows one to further upgrade to $\{U_{x,M\delta}(\gamma)\}_{x\in W}$, where W is a $\delta-$separated subset of $[0,1]$.

This is an ingredient to upgrade Theorem 1.0.8 to a local result for curves $\gamma \sim_\epsilon \mu$. Unlike in restriction problems, results for μ can be upgraded to results for $\gamma \sim_\epsilon \mu$. Via the Pramanik–Seeger type iteration, one can upgrade the decoupling theorem Theorem 1.0.8 to a local result:

Corollary 3.3.1 (Pramanik–Seeger improvement)**.** *Let $d > 1$, and $\epsilon > 0$. Let $p := d(d+1)$. There exists $\epsilon_D = \epsilon_D(d) > 0$ so that the following holds for all $\gamma(t)$ such that $\|\gamma(t) - \mu(t)\|_{C^{d+1}([0,1])} \le \epsilon_D$. Let $\delta > 0$, $M > 1$. Let W be a $\delta-$separated subset of $[0,1]$. Let $f_w : \mathbb{R}^d \to \mathbb{R}^d$, be functions such that $\hat{f}_j$ is supported on $Q_{M\delta,\mu}(w)$. Then*

$$\left\|\sum_{w\in W} f_w\right\|_{L^p} \le C_{\epsilon,d}\delta^{-\epsilon} M^{O_d(1)} \left|\sum_{w\in W} \|f_j\|_{L^p}^2\right|^{1/2} \tag{3.54}$$

Proof. Without loss of generality (by enlarging M if necessary) one assume $w = \delta\mathbb{Z} \cap [0,1]$. Let $\epsilon > 0$. Applying Theorem 1.2.5 with $\epsilon/2$ instead of ϵ, there is δ_0 (solving $\delta_0^{\epsilon/2} C_{\epsilon/2,d} = 1$)

such that for all $0 < \delta < \delta_0$

$$\left\|\sum f_j\right\|_{L^p} \le \delta_0^{-\epsilon}\left|\sum \|f_j\|_{L^p}^2\right|^{1/2} \tag{3.55}$$

for all functions f_j supported on $Q_{10M\delta,\mu}(\delta j)$. If $\epsilon_D < \delta_0^d$, we know that $Q_{10M\delta_0,\mu} \supseteq Q_{M\delta_0,\gamma}$, and therefore, (3.55) also holds for all functions f_i supported on $Q_{M\delta_0,\gamma}$.

This shows the result for some $\delta = \delta_0$ small enough. We will now iterate the statement to show it for $\delta = \delta_0^n$, $n \ge 0$. The result will then follow for other intermediate values of δ by the triangle inequality. We will use the following fact:

If $\|\mu - \gamma\|_{C^\infty([0,1])} < \epsilon_d$, and $\delta > 0$, for any $x \in [0, 1-\delta]$ there exists an affine transformation A such that, for $\gamma_x(z) := A\gamma(\delta z - x)$, $\|\mu - \gamma_x\|_{C^\infty([0,1])} < \epsilon_d$ as well.

By induction, we assume (3.55) holds for functions f_j with $\hat{f}_j$ supported $Q_{M\delta_0^{-(n-1)},\gamma_x}(\delta_0^{-(n-1)}j)$ for $j = 1, \dots \delta_0^{-(n-1)}$. By affine invariance of decoupling constants, the same holds if $\hat{f}_j$ is supported on $Q_{M\delta_0^{-n},\gamma}(\delta_0^{-n}j + \delta_0 k)$ for $j = 1, \dots \delta^{-(n-1)}$ and fixed $k = 0 \dots \delta^{-1}$. On the other hand, if $\epsilon_d < \frac{1}{10}$ and $M > 10$ (which we may assume),

$$\bigcup_{j=0}^{\delta^{-(n-1)}} Q_{M\delta_0^{-n},\gamma}(\delta_0^{-n}j + \delta_0 k) \subseteq Q_{M\delta_0^{-1},\gamma}(\delta_0 k)$$

and therefore the result follows from the monotonicity of decoupling. □

This implies a local version of Theorem 1.2.5:

Theorem 3.3.2. *Let $d > 1$, and $\gamma(t) : \mathbb{R} \to \mathbb{R}^d$ a nondegenerate polynomial curve of degree at most N. Let ϵ_0 (as in the hypothesis of Theorem 2.1.13) be small enough as required in Corollary 3.3.1. Let $A_{[r,R]}$ be one of the (real) annuli obtained in Theorem 2.1.13 (which, over the real numbers, reduce to intervals). Assume $[1,2] \subseteq [r,R]$.Let $\delta > 0$, $M > 1$. Let W be a $\delta-$separated subset of $[0,1]$. Let $f_w : \mathbb{R}^d \to \mathbb{R}^d$, be functions such that $\hat{f}_j$ is supported on $Q_{M\delta,\mu}(w)$. Then*

$$\left\| \sum f_w \right\|_{L^p} \leq C_{\epsilon,d,N} \delta^{-\epsilon} M^{O_d(1)} \left| \sum \|f_w\|_{L^p}^2 \right|^{1/2} \tag{3.56}$$

Proof. By Lemma 2.7.2, the interval $[1, 2]$ can be decomposed into $O_{N,d}(1)$ intervals where the hypotheses of the previous theorem 3.3.1 hold after an affine transformation. By restricting ourselves to each of these intervals, applying theorem 3.3.1, and applying the triangle inequality, the result follows. □

We are now ready to state the main result, which implies Theorem 1.2.5 by the geometric decomposition of Theorem 2.1.13. The idea is to split the curve γ into intervals such that each interval contains the same amount of *curvature*. The statement is slightly more technical, as it contains an extra free parameter s.

Theorem 3.3.3. *Let $d > 1$, and $\gamma(t)$ a polynomial curve. Assume that γ is ϵ_0-similar to a moment curve $\mu_{\mathbf{n}}$ on $A_{r,R}$ for ϵ_0 small enough and $R \geq 2r$. Let $s \geq 0$.*
Let $\delta > 0$. Let $\Delta_s(w) := \delta |\Lambda^{(d)}[\gamma](w)|^{-s}$. Let W be a subset of the interval $[r, R]$ such that if w, w' both belong to W, then $|w - w'| > \Delta_s(w) + \Delta_s(w')$. For each $w \in E$, let f_w be such that $\hat{f}_w$ is supported on is supported on $Q_{\Delta_s(w),\gamma}(w)$. Then

$$\left\| \sum_{w\in W} f_w \right\|_{L^p} \leq C_{\epsilon,d,s} \left|1 + |W|\right|^{\epsilon} \left| \sum_{w\in W} \|f_w\|_{L^p}^2 \right|^{1/2} \tag{3.57}$$

Setting $s = 0$ recovers Theorem 1.2.5. Theorem 3.3.3, however, shows that one does not need a particularly adapted family of cube centers (in terms of scaling with respect to the torsion) to obtain a uniform decoupling estimate.

Remark 3.3.4. *The motivation of $\Delta_s(w)$ is to bound (from below) the amount of torsion of γ inside each adapted parallellepiped. It is reasnoable in applications (such as in [75]) to consider rectangles containing equal amounts of curvature, which can be done by settin s to be the right positive exponent.*

In potential applications, the number of points in W in an integral can be bounded by an integral of the torsion, namely

Proposition 3.3.5. *In the context of 3.3.3, if $R \geq M \geq 2m \geq 2r \geq 0$ then*

$$|W \cap [m, M]| \lesssim 1 + \delta^{-1} \int_m^M |\Lambda^{(d)}[\gamma](w)|^s dw,$$

Proof. Without loss of generality assume $|W| \geq 2$. By the separation condition it suffices to show that

$$\int_{[w-\Delta_s(w), w+\Delta_s(w)]\cap[m,M]} |\Lambda^{(d)}[\gamma](x)|^s dx \gtrsim \delta \tag{3.58}$$

since $\gamma \sim \mu_{\mathbf{m}}$ it suffices to check that

$$\int_{[w-\Delta_s(w), w+\Delta_s(w)]\cap[m,M]} |\Lambda^{(d)}[\mu_{\mathbf{n}}](x)|^s dx \gtrsim \delta \tag{3.59}$$

For every $w \in W$ either $[w - \Delta_s(w), w] \subseteq [m, M]$ or $[w, w - \Delta_s(w)] \subseteq [m, M]$, and therefore the result follows from the fact that $\int_a^b |x|^s \gtrsim |a - b|$ □

Proof of Theorem 3.3.3. We will separate the proof into two cases. In the first case, the moment model moment curve on $\mu_{\mathbf{m}}$ is the canonical moment curve, and therefore $|\Lambda^{(d)}[\gamma](w)|^s \sim 1$. The result follows by Corollary 3.3.1.

In the second case, it is a moment curve of degree $N > d$. Let ϵ as in Theorem 3.3.3. Assume without loss of generality by reparametrization that $\delta \leq 1, r \leq 1 \leq 2 \leq R$. By removing one interval, which will not alter the decoupling constant more than a fixed constant, we can assume that there is no interval $[w - \Delta_s(w), w + \Delta_s(w)]$ intersecting 0. The first step is to remove a small amount of intervals that may still be very close to zero:

Removing bad certain bad points:

By hypothesis $|\Lambda^{(d)}[\gamma](w)|^s \approx |\Lambda^{(d)}[\mu_{\mathbf{n}}](w)|^s \approx w^\beta$ for some $\beta \geq 0$ depending only on $\mathbf{n}, s$. Therefore, for all the remaining points, since $w - \Delta_s(w) > 0$ we see that $w \gtrsim \delta^{1/(1+\beta)}$ (where all the implicit constants depend on N and s only). Let $\delta^{1/(1+\beta)} \lesssim w_1 \leq w_2 \leq \cdots \leq w_N$ be the points in w. By the separation hypothesis $w_{n+1} \geq w_n + c\delta w_n^{-\beta}$, for some c depending on

$\mathbf{n}, s$. By chaining these inequalities, one obtains that $w_k \geq kc\delta w_n^{-\beta} \approx k\Delta_s(w)$. This shows that the set $B = \{w \in W, \Delta_s(w) \geq \frac{w}{10}\}$ has finitely many points. By losing an $O(1)$ factor, one can remove those points.

Reducing back to the moment curve

Under this reduction, each interval $[w - \Delta_s(w), w + \Delta_s(w)]$ intersects at most two dyadic intervals $[2^{n-1}, 2^n]$. Since there are $O(\log \delta^{-1} R)$ such partitions containing points in W (all points in W have the property $\delta^{1/(1+\beta)} \lesssim w \leq R$), by a further application of triangle inequality one may assume that all the points w in W are contained in one of such intervals (one could use a Littlewood–Paley type estimate as well, but this log-term is not the dominant term). Again, without loss of generality, by rescaling invariance of the estimates, assume that the interval is the interval $[1, 2]$. In particular now $|\Lambda^{(d)}[\gamma](w)| \sim 1$, and therefore all the points in W are $O(\delta)$-separated.

By Proposition 2.7.2, after splitting the interval $[1, 2]$ into finitely many pieces, we can reduce ourselves to the hypotheses of Corollary 3.3.1. □

Remark: Taking $s = \frac{1}{2}$ and restricting to the class of polynomials therein, one recovers the main result in [75].

Remark: A suitable variation of this proof proof would imply uniform decoupling estimates if decoupling for the moment curve in such fields was known. To the author knowledge such results have not yet been explicitly established but in principle should follow from an argument analogous to that in [35] in, which avoids the use of field-specific techniques like Kakeya, or modular arithmetic.

CHAPTER 4

Uniformity for discrete analogues

4.1 Uniformity for discrete restriction on polynomial curves

The goal of this section is to show Theorem 1.2.6. The proof is similar in spirit to the proof of the main result in [48]. In fact, the proof of Theorem 1.2.6 reduces to that in [48] when γ is a monomial curve, and Lemma 4.1.1 becomes superfluous.

The first ingredient involves a change of variables identity for the discrete extension operator.

Lemma 4.1.1. *Let $T : \mathbb{Z}^d \to \mathbb{Z}^d$ be an injective linear map. For $f : \mathbb{Z}^d \to \mathbb{C}$, define $T_*f := \sum_{y\in\mathbb{Z}^d} f(y)\mathbb{1}_{Ty}$, the push-forward of f. Then the following hold:*

1. $\widehat{T_*f}(x) = \hat{f}(T^t x)$.
2. $\|\widehat{T_*f}\|_p = \|\hat{f}\|_p$

Remark 4.1.2. *If one substitutes $\mathbb{Z}$ for $\mathbb{R}$, then $T_*f(x) = |\det T|^{-1} f(T^{-1}f)$, recovering the classical change-of-variables formula for the Fourier transform.*

Proof of 4.1.1. The first equality is an algebraic identity:

$$\widehat{T_*f}(x) = \sum_{n\in\mathbb{Z}^d}\left(\sum_{y\in\mathbb{Z}^d} f(y)\mathbb{1}_{Ty}(n)\right)\exp(2\pi i n\cdot x) = \sum_{y\in\mathbb{Z}^d} f(y)\exp(2\pi i(Ty)\cdot x) = \hat{f}(T^t x).$$

The second equality follows from the first one since for any function ϕ on $\mathbb{T}$ and any linear map T mapping $\mathbb{Z}^d$ to itself, $\int_{\mathbb{T}^d}\phi(x)dx = \int_{\mathbb{T}^d}\phi(Tx)dx$. $\square$

Lemma 4.1.3. *Let $\gamma = (\gamma_1, \dots \gamma_n)$ be a polynomial curve of degrees $d_1 < d_2 < \dots d_n$. Let μ be the d_n-dimensional moment curve, and T an injective transformation such that $T \circ \mu$ is of the form $(\gamma_1(t), \dots \gamma_n(t), t^{s_1}, \dots t^{s_{d-n}})$, where $(s_1, \dots, s_{d_n - n})$ are the numbers in $1, 2, \dots, d_n$ not appearing in $(d_1, \dots d_n)$. Then:*

$$\|E_{\gamma([N])}\|_{\mathrm{Op}(l^p \to L^q)} \leq C_n N^{\frac{1}{q}(s_1 + \dots + s_{d_n - n})} \|E_{T\circ\mu([N])}\|_{\mathrm{Op}(l^p \to l^q)}$$

Proof. Let $f : \gamma([N]) \to \mathbb{C}$. Split $\mathbb{T}^{d_n} = \mathbb{T}^n \times \mathbb{T}^{n-d_n}$, and write, in a slight abuse of ntation, $E_{T\circ\mu([N])}(f)$ as a two-variable function, $E_N^{T\mu}(f)(x, y)$, with $(x, y) \in \mathbb{T}^n \times \mathbb{T}^{n-d_n}$. Then one has the equality

$$E_{\gamma([N])}(f)(x) = E_{T\circ\mu([N])}(f)(x, 0).$$

and therefore

$$\|E_{\gamma([N])}(f)(x)\|_{L^q(\mathbb{T}^n)} \leq \|E_{T\circ\mu([N])}(f)(x, y)\|_{L^q(\mathbb{T}^n)L^\infty(\mathbb{T}^{n-d_n})}.$$

The L^∞ term on the right hand side can be bounded by an L^q norm using reverse Hölder for each fixed x, as

$$\|E_{T\circ\mu([N])}(f)(x, y)\|_{L^\infty(y)} \lesssim_n N^{1/q\cdot(s_1 + \dots + s_{d_n - n})} \|E_{T\circ\mu([N])}(f)(x, y)\|_{L^q(y)}$$

which holds for $1 < p < \infty$ for functions with Fourier support on rectangles of shape $N^{s_1} \times N^{s_2} \times \dots \times N^{s_{n-d}}$ (applied to $y \mapsto E_{T\circ\mu([N])}(f)(x, y)$). Taking L^q norms in x of this last inequality finishes the proof. □

These are all the essential ingredients to prove Theorem 1.2.7.

Proof of Theorem 1.2.7. Given a polynomial curve γ, one can always find an injective linear map T as in Lemma 4.1.3, to show that

$$\|E_{\gamma([N])}\|_{\mathrm{Op}(l^2 \to L^p)} \leq C_n N^{\frac{1}{p}(s_1 + \dots + s_{d_n - n})} \|E_{T\circ\mu([N])}\|_{\mathrm{Op}(l^2 \to l^p)}$$

by Lemma 4.1.1, $\|E_{T\circ\mu([N])}\|_{\mathrm{Op}(l^2\to l^p)} = \|E_{\mu([N])}\|_{\mathrm{Op}(l^2\to l^p)}$. But Theorem 1.0.11 for the moment curve of degree d_n shows that, for $p > p_0$

$$\|E_{\mu([N])}\|_{\mathrm{Op}(l^2\to l^p)} \leq C_{\epsilon,d_n} N^{\epsilon+\frac{1}{2}-\frac{1+2+\cdots+d_n}{p}}$$

the result now follows from the fact that $d_1 + \ldots d_n + s_1 + \cdots + s_{d_n-n} = 1 + \cdots + d_n$. □

4.2 Uniformity for discrete averages on polynomial curves

The result of this section arose after conversations with José Madrid, the author is grateful for these conversations.

The goal of this section is to show Theorem 1.2.7 on uniform discrete averages. The proof follows a slightly more careful analysis of the work of [38] therein, using ideas from [21], which shows a result in the spirit of Theorem 1.2.7 in the sub-critical case. The main theorem in [38] considers polynomials $\mathbb{Z} \to \mathbb{Z}^d$ we consider only polynomials in $\mathbb{Z}[x]$ instead of polynomials mapping $\mathbb{Z}$ to $\mathbb{Z}^d$. The results are extendable without much effort to the general case[1].

Much like Theorem 1.2.6 followed from Theorem 1.0.11, the proof of Theorem 1.2.7 will follow from Corollary 1.0.12. The proof follows strategy is essentially the same, and one could think of Lemmas 4.2.1 and Lemma 4.2.2 as dual counterparts to Lemmas 4.1.1 and Lemma 4.1.3.

Lemma 4.2.1. *Let $T : \mathbb{Z}^d \to \mathbb{Z}^d$ be an injective linear map. Let γ be a polynomial curve in $\mathbb{Z}[x]^d$. Let $q > p$. Then*

$$\|A_N^{T\circ\gamma}\|_{\mathrm{Op}(l^p\to l^q)} = \|A_N^{\gamma}\|_{\mathrm{Op}(l^p\to l^q)}$$

[1]For each fixed degree, polynomials in $P_{\mathbb{Z}}$ of that degree belong to $(M^{-1}\mathbb{Z})(x)$ for M a large integer, and one may then use the main result in that lattice.

Proof. There is a canonical bijection $\eta : \mathbb{Z}^d \to \mathbb{Z}^d \times \operatorname{coker} T$ (in **Set**) induced by T. Let $\tilde{A}_{\gamma[N]} : l^p(\mathbb{Z}^d \times \operatorname{coker} T) \to l^q(\mathbb{Z}^d \times \operatorname{coker} T)$ be the averaging operator acting independently on each of the copies of $\mathbb{Z}^d$. Then we have:

$$\begin{aligned}
\|A_N^{T\circ\gamma} f\|_{l^q(\mathbb{Z}^d)} =& \|\tilde{A}_N^{\gamma}(f \circ \eta^{-1})\|_{l^q(\operatorname{coker} T) l^q(\mathbb{Z}^d)} && (4.1)\\
\leq& \|\mathcal{A}_N^{\gamma}\|_{\operatorname{Op}(l^p \to l^q)} \|(f \circ \eta^{-1})\|_{l^q(\operatorname{coker} T) l^p(\mathbb{Z}^d)} && (4.2)\\
\leq& \|\mathcal{A}_N^{\gamma}\|_{\operatorname{Op}(l^p \to l^q)} \|(f \circ \eta^{-1})\|_{l^p(\operatorname{coker} T \times \mathbb{Z}^d)} && (4.3)\\
=& \|\mathcal{A}_N^{\gamma}\|_{\operatorname{Op}(l^p \to l^q)} \|f\|_{l^p(\mathbb{Z}^d)} && (4.4)
\end{aligned}$$

This shows the $\leq$ direction. The $\geq$ direction (which is the one we don't need, but it's also true so we may as well show it) is in fact the easy one: given a f that almost-extremizes the RHS, we can obtain $T_* f$ wich gives the same lower bound to the operater in the LHS. □

The last remaining piece to show the main theorem is the following:

Lemma 4.2.2. *Let $\gamma = (\gamma_1, \ldots \gamma_n)$ be a polynomial curve of degrees $d_1 < d_2 < \ldots d_n$. Let μ be the d_n-dimensional moment curve, and T an injective transformation such that $T\mu$ is of the form $(\gamma_1(t), \ldots \gamma_n(t), t^{s_1}, \ldots t^{s_{d-n}})$ (where $(s_1, \ldots, s_{d_n - n})$ are the numbers on $1, \ldots d_n$ not appearing in $(d_1, \ldots d_n)$). Then:*

$$\|A_N^{\gamma}\|_{\operatorname{Op}(l^p \to l^q)} \leq C_d N^{\left(\frac{1}{p} - \frac{1}{q}\right)(s_1 + \ldots s_{d_n - n})} \|A_N^{T\circ\mu}\|_{\operatorname{Op}(l^p \to l^q)}$$

Proof. Given $g \in l^p(\mathbb{Z}^n)$, let $\tilde{g} \in l^p(\mathbb{Z}^d)$ defined as

$$\tilde{g} := g \otimes \chi_{[0, 2N^{s_1}]} \otimes \cdots \otimes \chi_{[0, 2N^{s_{d-n}}]}.$$

Then it holds that:

$$\begin{aligned}\|A_N^\gamma g\|_{l^q} \lesssim_d & N^{-(s_1+\cdots+s_{d_n-n})\frac{1}{q}}\|A_N^{T\mu}\tilde{g}\|_{l^q}\\ \leq & N^{-(s_1+\cdots+s_{d_n-n})\frac{1}{q}}\|A_N^{T\circ\mu}\|_{\mathrm{Op}(l^p\to l^q)}\|\tilde{g}\|_{l^p}\\ \leq & N^{\left(\frac{1}{p}-\frac{1}{q}\right)(s_1+\ldots s_{d_n-n})}\|A_N^{T\circ\mu}\|_{\mathrm{Op}(l^p\to l^q)}\|\tilde{g}\|_{l^p}\end{aligned}$$

□

Now we have all the ingredients to prove Theorem 1.2.7

Proof of Theorem 1.2.7: Note that the transformation T in Lemma 4.2.2 always exists. Combining Lemmas 4.2.1 and 4.2.2, if $p < p'$ (i.e. $p < 2$, as in the hypothesis),

$$\|A_N^\gamma\|_{\mathrm{Op}(l^p\to l^{p'})} \leq C_d N^{\left(\frac{1}{p}-\frac{1}{p'}\right)(s_1+\ldots s_{d-n})}\|A_N^\mu\|_{\mathrm{Op}(l^p\to l^{p'})}$$

Let $|s| = s_1 + \ldots s_{d_n-n}$. Now, Theorem 1.0.12 for the moment curve of degree n, plus the fact that $|d| + |s| = 1 + 2 + \cdots + n$ lets us estimate $\|A_N^\mu\|_{\mathrm{Op}(l^p\to l^{p'})}$, from which we obtain

$$\|A_N^\gamma\|_{\mathrm{Op}(l^p\to l^{p'})} \leq C_d N^{\left(\frac{1}{p}-\frac{1}{p'}\right)|s|} N^{-\left(\frac{1}{p}-\frac{1}{p'}\right)(|s|+|d|)} \leq N^{-\left(\frac{1}{p}-\frac{1}{p'}\right)(|d|)}$$

□

Remark 4.2.3. *If one didn't mind acquiring an ϵ-loss in Theorem 1.2.7, one could use Theorem 1.2.6 to show Theorem 1.2.7, as in the proof of Corollary 1.0.12. Alternatively, one could show as well that in the case of $f = \mathbb{1}_\gamma$, the ϵ-loss can be removed from the exponent in Theorem 1.2.6, following the same proof.*

Part II

Decoupling, Cantor sets and additive combinatorics

CHAPTER 5

Introduction

The common goal of this part is to further explore the relationship between decoupling estimates, discrete restriction estimates, and additive combinatorics. Decoupling is intimately related to estimates in additive number theory: the motivation of Burgain-Demeter-Guth to prove their decoupling estimate for the moment curve was number-theoretic in nature. Their goal was in fact to count the number of solutions to the system

$$\begin{cases} x_1 + \cdots + x_n & = y_1 + \cdots + y_n \\ x_1^2 + \cdots + x_n^2 & = y_1^2 + \cdots + y_n^2 \\ \quad \cdots & = \quad \cdots \\ x_1^n + \cdots + x_n^d & = y_1^n + \cdots + y_n^d \\ |x_i|, |y_i| & < N \end{cases} \tag{5.1}$$

Bounding the number of solutions by $\max\left(N^n, N^{2n-\frac{d^2+d}{2}}\right)$ was known as the *main conjecture* of Vinogradov's Mean Value Theorem (MVMT). In this part we will use decoupling estimates to study two related of problems:

1. In Chapter 6 we study a variation of (5.1) when $k = 1, 2$, and but the digits of each x_i in base p are restricted to a strict subset of $\{0, \dots, p-1\}$. This problem, originally studied by Biggs [8, 7] using number-theoretic tools developed by Wooley [73], is closely related to a problem of decoupling for fractal subsets in a parabola/moment curve. The case $k = 2$ will follow from a careful modification of known proofs for Theorem 1.0.8

([71, 51]), inducting from $k = 1$. The case $k = 1$ will requires different ideas. In Section 2.3 we will build quantitative approximations between those cantor sets and some associated product sets, and compute the decoupling constants for those product sets using the *tensor trick*.

- In Chapter 7 we will study subsets of the hypercube, or a general product set, which have high additive energy, or suitable generalizations of additive energy. The additive energy $E_2(X)$ of a set X is the number of solutions to the equation $a + b = c + d$, for $a, b, c, d \in X$. We will be concerned with finding the best power c so that $E_2(X) \leq |X|^c$ for all subsets X of a product set A^n, for all n. In order to solve this question, we will first *relax* it to an equivalent, continuous version, to which we will be able to apply the exact same *tensor trick* from the Cantor set decoupling problem. In fact, the product sets that we will recover are exactly those from the Cantor set problem.

5.1 Decoupling and solution counting

In order to study and connect these questions, we will consider a hierarchy of four different problems: Decoupling problems, discrete restriction problems, subset solution-counting problems and solution-counting problems. These estimates are ordered by strength: Estimates for the first problem can be transferred to the second, and so forth. In the case of the Vinogradov mean value theorem, the solution-counting lower bounds to problem 4. correspond, up to an ϵ-loss to the upper bounds, obtained to problem 1. by decoupling.

Proposition 5.1.1. *Let $S \subseteq Z^d$, and $\mathcal{U}_S := \{U_s\}_{s \in S}$ a family of open sets so that for all $s \in S$, $s \in U_s$. Let n be a natural number. Then, each of these estimates implies the following (in the sense that* 1. $\implies$ 2. $\implies$ 3. $\implies$ 4.*):*

1. $\mathrm{Dec}_{\ell^p L^{2n}}(\mathcal{U}) \leq C$

2. $\|DE(S)\|_{\ell^p \to L^{2n}} \leq C$

3. For all subsets $A \subseteq S$,

$$\#\left\{(x_1,\dots x_n, y_1, \dots y_n) \in A^n \ s.t \sum_{i=1}^n x_i = \sum_{i=1}^n y_i\right\} \leq C^{2n}|A|^{2n/p} \tag{5.2}$$

4. Estimate (5.2) *holds when $A = S$.*

Proof.

(2. $\Longrightarrow$ 3.) One has the equality

$$\begin{aligned}\int_{\mathbb{T}^d}\left|\sum_{s\in S} a_s e^{2\pi i s w}\right|^{2n} dw &= \sum_{\substack{s_1,\dots s_n\in S\\ t_1,\dots,t_n\in S}} \int_{\mathbb{T}^d} e^{2\pi i w\cdot\left(\sum_{i=1}^n x_i - \sum_{i=1}^n y_i\right)} \prod_{i=1}^n a_{s_i}\bar{a}_{t_i} dw \\ &= \sum_{\substack{s_1,\dots s_n\in S\\ t_1,\dots,t_n\in S}} \mathbb{1}_{\sum_{i=1}^n x_i = \sum_{i=1}^n y_i} \prod_{i=1}^n a_{s_i}\bar{a}_{t_i} dx\end{aligned} \tag{5.3}$$

The left hand side of (5.3) is a discrete extension estimate, and the right hand side is a (weighted) solution counting problem. Now the result follows by taking $a_s = \mathbb{1}_A(s)$, and the fact that $\|\mathbb{1}_A(s)\|_{\ell^p}^{2n} = |A|^{2n/p}$ to estimate the discrete extension estimate.

(1. $\Longrightarrow$ 2.) Let ϕ be a smooth, rapidly decaying function with $\hat{\phi}$ supported on B_δ, with $\delta < \min\left(\frac{1}{2n}, \min_{s\in S}(d(s, U_s^c)\right)$. For $s \in S$, let $\hat{f}_s(w) := \hat{\phi}(w - s)$, or, equivalently, $f_s(w) = a_s e^{2\pi i s w}\phi(w)$. Then, one has the equality:

$$\int_{\mathbb{T}^d}\left|\sum_{s\in S} a_s e^{2\pi i s w}\phi(w)\right|^{2n} dw = \|\phi\|_{L^{2n}}^{2n} \sum_{\substack{s_1,\dots s_n\in S\\ t_1,\dots,t_n\in S}} \mathbb{1}_{\sum_{i=1}^n x_i = \sum_{i=1}^n y_i} \prod_{i=1}^n a_{s_i}\bar{a}_{t_i} dx \tag{5.4}$$

the Fourier support of $|\phi(x)|^s e^{2\pi i x\cdot\left(\sum_{i=1}^n x_i - \sum_{i=1}^n y_i\right)}$ is contained in $\sum_{i=1}^n x_i - \sum_{i=1}^n y_i + B_{2n\delta}$, and in particular contains the origin if and only if $\sum_{i=1}^n x_i - \sum_{i=1}^n y_i = 0$. This shows that

$$\int_{\mathbb{T}^d}\left|\sum_{s\in S} a_s e^{2\pi i s w}\phi(w)\right|^{2n} dw = \sum_{\substack{s_1,\dots s_n\in S\\ t_1,\dots,t_n\in S}} \int_{\mathbb{T}^d} |\phi(x)|^s e^{2\pi i x\cdot\left(\sum_{i=1}^n x_i - \sum_{i=1}^n y_i\right)} \prod_{i=1}^n a_{s_i}\bar{a}_{t_i} dx \tag{5.5}$$

and comparing with (5.3), one gets

$$\int_{\mathbb{T}^d}\left|\sum_{s\in S} a_s e^{2\pi i s w}\phi(w)\right|^{2n} dw = \|\phi\|_{L^{2n}}^{2n} \int_{\mathbb{T}^d}\left|\sum_{s\in S} a_s e^{2\pi i s w}\right|^{2n} dw \|\phi\|^{2n} \tag{5.6}$$

The decoupling estimate estimates the left-hand side of (5.6) by

$$C^{2n}\left(\sum_{s\in S} a_s^p \|\phi\|_{L^{2n}}\right)^{2n/p}$$

so one recovers 2. from 1.

The remaining (3. $\implies$ 4.) is a tautology.

□

Using the result above one may, for example, show that decoupling for a neighborhood the moment curve gives almost-sharp bounds to (5.1). None of these steps are generally sharp, however.

Decoupling for a neighborhood of the (t, t^3) curve (using the Pramanik–Seeger iteration, as in Corollary 3.3.1) cannot be any better than that of (t, t^2) curve, for which the best estimates are, up to an ϵ loss, sharp. Discrete restriction estimates (for example) shown in Theorem 1.2.6, however, are strictly better for $p \geq 12$.

The difference between 2. and 3. is in fact, minimal. By a pigeonholing argument, the best constants differ by at most a factor of $\log |A|$ (or even a small power of $\log A$, see [55]).

Sections 2.3 and 7.4 are essentially concerned about the difference between 4 and the cases 2,3. Problems 3. and 4. turn out to be equivalent only in situations of extreme symmetry, such as the $\{0, 1\}^n$ hypercube, a single arithmetic progression, or the middle-thirds cantor set, but have a gap in cases like $\{0, 1, 2\}^n$, and one has to study the discrete restriction problem to understand the solution-counting problem.

5.2 Decoupling type estimates tensorize

A remarkably useful property of Decoupling estimates is that they tensorize. Namely, if $\mathcal{U}$ is a family of subsets of $\mathbb{R}^a$ and $\mathcal{V}$ is a family of subsets in $\mathbb{R}^b$, and we define $\mathcal{V}\times\mathcal{U} := \{U\times V\}_{\substack{U\in\mathcal{U},\\ V\in\mathcal{V}}}$, then, if $p \leq q$

$$\mathrm{Dec}_{\ell^p \to L^q}(\mathcal{U} \times \mathcal{V}) = \mathrm{Dec}_{\ell^p \to L^q}(\mathcal{U})\, \mathrm{Dec}_{\ell^p \to L^q}(\mathcal{V}) \,. \tag{5.7}$$

The $\geq$ step follows by checking against tensor products, and the $\leq$ step is a standard application of Holder's inequality (see Propositions 6.3.3 and 7.4.5): If f_{UV} are functions with Fourier support on $U \times V$, then

$$\begin{aligned}
\| \sum_{\substack{U \in \mathcal{U} \\ V \in \mathcal{V}}} f_{UV}(x,y) \|_{L^q(y) L^q(x)} &\leq \mathrm{Dec}_{\ell^p \to L^q}(\mathcal{V}) \| \sum_{V \in \mathcal{V}} f_{UV}(x,y) \|_{L^q(y) \ell^p(\mathcal{U}) L^q(x)} && (5.8)\\
&\leq \mathrm{Dec}_{\ell^p \to L^q}(\mathcal{V}) \| \sum_{V \in \mathcal{V}} f_{UV}(x,y) \|_{\ell^p(\mathcal{U}) L^q(x) L^q(y)} && (5.9)\\
&\leq \mathrm{Dec}_{\ell^p \to L^q}(\mathcal{V})\, \mathrm{Dec}_{\ell^p \to L^q}(\mathcal{U}) \, \| f_{UV}(x,y) \|_{\ell^p(\mathcal{U}) L^q(x) \ell^p(\mathcal{U}) L^q(y)} && (5.10)\\
&\leq \mathrm{Dec}_{\ell^p \to L^q}(\mathcal{V})\, \mathrm{Dec}_{\ell^p \to L^q}(\mathcal{U}) \, \| f_{UV}(x,y) \|_{\ell^p(\mathcal{U} \times \mathcal{V}) L^q(x,y)} && (5.11)
\end{aligned}$$

The $p \leq q$ hypothesis is neede to exchange orders of the norms. A similar estimate holds for discrete extension, where, for $p < q$, one has

$$\| DE(A \times B) \|_{\ell^p \to L^q} = \| DE(A) \|_{\ell^p \to L^q} \cdot \| DE(B) \|_{\ell^p \to L^q}$$

CHAPTER 6

Decoupling for Cantor Subsets of the parabola

This section reproduces the paper Decoupling for fractal subsets of the parabola [15], *by Alan Chang, Jaume de Dios Pont, Rachel Greenfeld, Asgar Jamneshan, Zane Kun Li and José Madrid with minor changes.*

Abstract: *(From the published version)* We consider decoupling for a fractal subset of the parabola. We reduce studying l^2L^p decoupling for a fractal subset on the parabola $\{(t,t^2) : 0 \leq t \leq 1\}$ to studying $l^2L^{p/3}$ decoupling for the projection of this subset to the interval $[0,1]$. This generalizes the decoupling theorem of Bourgain-Demeter in the case of the parabola. Due to the sparsity and fractal like structure, this allows us to improve upon Bourgain-Demeter's decoupling theorem for the parabola. In the case when $p/3$ is an even integer we derive theoretical and computational tools to explicitly compute the associated decoupling constant for this projection to $[0,1]$. Our ideas are inspired by the recent work on ellipsephic sets by Biggs [7, 8] using nested efficient congruencing.

6.1 Introduction

Fix an integer $q \geq 3$, not necessarily a prime, and let $\delta(i) := 1/q^i$, $i \geq 0$. Let $C_0 := [0,1]$. To construct level i, we partition C_{i-1} into intervals of length $\delta(i)$, remove some of them, and denote by $N(i)$ the number of unremoved intervals. We associate $C = \bigcap_{i\geq 0} C_i$ with its levels C_i. For an interval I with $|I| = \delta(i)$, $\delta(i) > \delta(j)$, $P_{\delta(j)}(I \cap C_j)$ will denote the collection of intervals that make up C_j which are contained in I. We also let $P_{\delta(i)}(C_i) = P_{\delta(i)}([0,1] \cap C_i)$

be the collection of intervals of length $\delta(i)$ that make up C_i and so $N(i) = \#P_{\delta(i)}(C_i)$.

We call $C = \bigcap_{i\geq 0} C_i$ a *generalized Cantor set* and C_i a *generalized Cantor set of level* i, when the following three conditions are satisfied:

- $N(i+j) = N(i)N(j)$.
- $C_i \subset C_{i-1}$.
- The level C_i is similar to level C_{i-1}. More precisely, for every interval $I \in P_{\delta(i-1)}(C_{i-1})$, the set $I \cap C_i$ is a translate of $q^{-1}C_{i-1}$.

By multiplicativity of $N(\cdot)$, given an $I \in P_{\delta(i)}(C_i)$ and $i < j$, the number of intervals in $P_{\delta(j)}(C_j)$ that are contained in I is $N(j-i)$. Additionally,

$$\delta(i)^{-\dim(C)} = N(i) \tag{6.1}$$

where $\dim(C)$ is the Hausdorff dimension of C. Note that in our definition, it is possible to let $N(i) = q^i$ and so C_i is the partition of $[0,1]$ into intervals of length $1/q^i$.

The traditional middle-thirds Cantor set has $q = 3$ and $N(i) = 2^i$. To avoid writing *generalized Cantor set* repeatedly, we will just call the above constructed set C, a Cantor set and C_i, a level of Cantor set. A simple modification of our argument also allows it to work with asymmetric Cantor sets, however in order to simplify the arguments notation-wise, we do not pursue such a goal here.

Given a level of a Cantor set C_i, for each interval $I \in P_{\delta(i)}(C_i)$, let ℓ_I denote the left endpoint of I and

$$\Omega_I := \{\xi \in \mathbb{R}^2 : \ell_I \leq \xi_1 \leq \ell_I + \delta(i), |\xi_2 - (2\ell_I + \delta(i))(\xi_1 - \ell_I) - \ell_I^2| \leq \delta(i)^2\}.$$

Note that Ω_I is a $O(\delta(i)) \times O(\delta(i)^2)$ parallelogram that covers and is covered by a $O(\delta(i)^2)$ neighborhood of the piece of parabola above I.

For an interval I and $f : \mathbb{R} \to \mathbb{R}$, let f_I be defined such that $\widehat{f_I} = \widehat{f} 1_I$. Next for a region θ and $f : \mathbb{R}^2 \to \mathbb{R}$, let f_θ be defined such that $\widehat{f_\theta} = \widehat{f} 1_\theta$.

6.1.1 Decoupling for C_i on the parabola

Fix a Cantor set C and its levels C_i. For $p \geq 2$, let $D_p(\delta(i))$ be the best constant such that

$$\|\sum_{J\in P_{\delta(i)}(C_i)} f_{\Omega_J}\|_{L^p(\mathbb{R}^2)} \leq D_p(\delta(i))(\sum_{J\in P_{\delta(i)}(C_i)} \|f_{\Omega_J}\|^2_{L^p(\mathbb{R}^2)})^{1/2}$$

for all Schwartz functions f which are Fourier supported in $\bigcup_{J\in P_{\delta(i)}(C_i)} \Omega_J$.

In the case when the Cantor set C is the whole interval $[0,1]$ and C_i is the partition of $[0,1]$ into intervals of length $\delta(i)$, we see that $D_p(\delta(i))$ is just the regular l^2L^p decoupling constant for the parabola considered by Bourgain-Demeter in [11, 10] and so we immediately have $D_p(\delta(i)) \lesssim_\varepsilon \delta(i)^{-\varepsilon}(1+\delta(i)^{-(\frac{1}{2}-\frac{3}{p})})$. Our main result is the following generalization of Bourgain-Demeter's parabola decoupling theorem.

Theorem 6.1.1. *Fix $p \geq 2$ and a Cantor set C and its levels. Let $\kappa_p(C)$ be the smallest number such that*

$$\|\sum_{J\in P_{\delta(i)}(C_i)} f_J\|_{L^p(\mathbb{R})} \lesssim_{p,\varepsilon,\dim(C),N(1)} N(i)^{\kappa_p(C)+\varepsilon}(\sum_{J\in P_{\delta(i)}(C_i)} \|f_J\|^2_{L^p(\mathbb{R})})^{1/2} \tag{6.2}$$

for all Schwartz functions $f : \mathbb{R} \to \mathbb{R}$ and all i. Then the l^2L^{3p} decoupling constant for C is such that for every $\varepsilon > 0$,

$$D_{3p}(\delta(i)) \lesssim_{p,\varepsilon,\dim(C),N(1)} N(i)^{\kappa_p(C)+\varepsilon}.$$

This theorem is proven in Section 6.2. The case of $p = 2$ is just an immediate application Bourgain-Demeter's result on the parabola and (6.1). For $p > 2$, due to the sparsity and fractal structure of C, we can do better than directly applying Bourgain-Demeter (see the examples summarized later or alternatively written in more detail in Section 6.3.3).

In the case when C is the whole interval, Theorem 6.1.1 gives a sharp theorem for decoupling for the parabola. However, whether Theorem 6.1.1 is sharp for arbitrary Cantor sets C is an area to be explored. Note that even if the $\lesssim_{p,\varepsilon,\dim(C),N(1)}$ can be replaced with

$\lesssim_{p,\varepsilon}$ (as is the case with our examples in Section 6.3.3), the proof of Theorem 6.1.1 adds in implicit constants that depends on $\dim(C)$ and $N(1)$.

The proof of Theorem 6.1.1 is inspired from [8], in particular one can think of [8, (1.2)] as an l^2L^{2t} decoupling theorem on the line for which we then upgrade to an l^2L^{6t} decoupling theorem on the parabola. However, Theorem 6.1.1 is more general than [8] since it is valid for arbitrary Cantor sets as defined on the first page rather than ellipsephic sets. Additionally, similar to the relation between [7] and [8], given a Cantor set C and its levels, one can use ideas from [36] to write a version of Theorem 6.1.1 which upgrades l^2L^p decoupling on the line to $l^2L^{k(k+1)p/2}$ decoupling on the moment curve $\xi \mapsto (\xi, \xi^2, \ldots, \xi^k)$. However in this paper we only consider the case of the parabola.

Analogous to how [8] is related to Wooley's nested efficient congruencing [73], the proof of Theorem 6.1.1 is similar in style to the proof of decoupling for the parabola found in [36, 51] though here we more closely follow Tao's exposition [71] based off these two papers. For more discussion on decoupling interpretations of efficient congruencing, see [33, 36, 51] which are decoupling interpretations of the efficient congruencing papers [40], [73], and [59, Section 4.3], respectively.

Demeter in [19] generalized decoupling for the parabola in a different way. He considered the partition that arises from the set $\mathcal{C}_{\alpha,n} = \{0, \alpha\} + \{0, \alpha^2\} + \cdots + \{0, \alpha^n\}$ for $0 \leq \alpha \leq 1/2$ and proved l^2L^p, $2 < p < 6$ decoupling estimates for the parabola decoupling question associated to this partition. The case $\alpha = 1/2$ corresponds to the uniform partition of $[0, 1]$ into intervals of length 2^{-n}. More precisely, he showed that the decoupling constant is $O_\varepsilon(2^{n\varepsilon})$ uniform in α. The difference between Demeter's result and our work here is that he starts with the whole interval $[0, 1]$ and decouples into a self similar partition of $[0, 1]$ built from $\mathcal{C}_{\alpha,n}$ while in our work we start with a sparse subset of $[0, 1]$ and decouple into its individual pieces. Additionally, the intervals in his partition have varying lengths while here our intervals all have the same length. See also [39] for a much stronger square function estimate for a lacunary partition of $[0, 1]$, the same comments on [19] also apply here.

6.1.2 Decoupling for C_i on $[0,1]$

Theorem 6.1.1 reduces studying $D_{3p}(\delta(i))$ to studying (6.2). We accomplish this in Section 6.3 for even integer p and specific Cantor sets C related to ellipsephic sets.

6.1.2.1 Discrete restriction and decoupling

First we define a discrete restriction for subsets $S \subset \mathbb{Z}^m$ and decoupling constants for $\Omega \subset [0,1]$. For $S \subset \mathbb{Z}^m$, let $\mathrm{DE}_{\ell^2\to L^p}(S)$ be the best constant such that

$$\|\sum_{\ell\in S} a(\ell)e(\ell\cdot x)\|_{L^p([0,1]^m)} \leq \mathrm{DE}_{\ell^2\to L^p}(S)(\sum_{\ell\in S}|a(\ell)|^2)^{1/2}$$

for all $a : S \to \mathbb{R}_{\geq 0}$. Next for a subset $\Omega \subset [0,1]$ partitioned into intervals I of equal length, let $K_p(\Omega)$ be the best constant such that

$$\|\sum_I f_I\|_{L^p(\mathbb{R})} \leq K_p(\Omega)(\sum_I \|f_I\|^2_{L^p(\mathbb{R})})^{1/2}$$

for all Schwartz functions $f : \mathbb{R} \to \mathbb{R}$.

Since we plan to discuss multiple different S and S will be related to Ω, we have chosen to emphasize the dependence of $\mathrm{DE}_{\ell^2\to L^p}(S)$ and $K_p(\Omega)$ on S and Ω rather than just the scale that comes naturally with Ω. This is different from what we did in the definition of $D_p(\delta(i))$ above with C_i being associated naturally with the scale $\delta(i)$.

6.1.2.2 Arithmetic Cantor sets and ellipsephic sets

We define an *arithmetic Cantor set* of base q with digits $0 \leq d_1 < \ldots < d_k < q \in \mathbb{N}$ to be the set of fixed points of the iterated function system generated by the functions $\{f_{d_j} = (x \mapsto q^{-1}(x + d_j))\}_{j=1,\ldots,k}$. This is a self-similar compact subset of $[0,1]$ with Hausdorff dimension $\frac{\log k}{\log q}$. We will denote it by $C_q^{\{d_1,\ldots,d_k\}}$.

Denote by $[C_q^{\{d_1,\ldots,d_k\}}]_j$ the $j-$th level of $C_q^{\{d_1,\ldots,d_k\}}$, that is

$$[C_q^{\{d_1,\ldots,d_k\}}]_j := \bigcup_{(s_1,\ldots,s_j)\in\{d_1,\ldots,d_k\}^j} (f_{s_1}\circ\cdots\circ f_{s_j})([0,1]).$$

For brevity of notation, the intervals of length q^{-j} in $P_{q^{-j}}([C_q^{\{d_1,\ldots,d_k\}}]_j)$ will be denoted by $[\mathscr{C}_q^{\{d_1,\ldots,d_k\}}]_j$. In particular, observe that

$$[\mathscr{C}_q^{\{d_1,\ldots,d_k\}}]_j = \{(f_{s_1}\circ\cdots\circ f_{s_j})([0,1]) : (s_1,\ldots,s_j)\in\{d_1,\ldots,d_k\}^j\}.$$

The standard middle thirds Cantor set is the arithmetic Cantor set $C_3^{\{0,2\}}$. Note also that $C_3^{\{0,1\}}$ and $C_3^{\{0,2\}}$ are dilated copies of each other.

There is also a close connection between arithmetic Cantor sets and ellipsephic sets defined in [8]. An *ellipsephic set* of base q with digits $0\le d_1<\cdots<d_k<q\in\mathbb{N}$ is the set of integers of the form $\sum_{s=0}^{j-1} a_s q^s$ (with $a_s\in\{d_1,\ldots,d_k\}$) for some $j\ge 1$. We will denote it by $\mathcal{E}_q^{\{d_1,\ldots,d_k\}}$. We will use $[\mathcal{E}_q^{\{d_1,\ldots,d_k\}}]_j$ to mean the set $\mathcal{E}_q^{\{d_1,\ldots,d_k\}}\cap[0,q^j)$. Comparing the definitions of an arithmetic Cantor set and an ellipsephic set, we easily observe that

$$[C_q^{\{d_1,\ldots,d_k\}}]_j = q^{-j}\left([\mathcal{E}_q^{\{d_1,\ldots,d_k\}}]_j + [0,1]\right).$$

Using the convenience that $2n$ is even and expanding the L^{2n} norm (Proposition 6.3.1), allows use to show Proposition 6.3.4

$$K_{2n}([\mathscr{C}_q^{\{d_1,\ldots,d_k\}}]_j) \sim \mathrm{DE}_{\ell^2\to L^{2n}}([\mathcal{E}_q^{\{d_1,\ldots,d_k\}}]_j) \tag{6.3}$$

(where the implied constant is absolute) which connects decoupling and discrete restriction constants.

When we study $\mathrm{DE}_{\ell^2\to L^{2n}}([\mathcal{E}_q^{\{d_1,\ldots,d_k\}}]_j)$, we will say $\mathcal{E}_q^{\{d_1,\ldots,d_k\}}$ has no carryover if $nd_k<q$. In particular, this definition depends on the n in question. Additionally note that we will say that $\mathcal{E}_q^{\{d_1,\ldots,d_k\}}$ has carryover if $nd_k\ge q$. This terminology was inspired from the proof of [8, Lemma 2.2]. Using Freiman isomorphisms, we have the following nice proposition which simplifies greatly discrete restriction for ellipsephic sets when we have no carryover (see Proposition 6.3.5 for a more precise statement).

Proposition 6.1.2. *If $\mathcal{E}_q^{\{d_1,\ldots,d_k\}}$ is an ellipsephic set without carryover, then*

$$\mathrm{DE}_{\ell^2\to L^{2n}}([\mathcal{E}_q^{\{d_1,\ldots,d_k\}}]_j) = \mathrm{DE}_{\ell^2\to L^{2n}}([\mathcal{E}_q^{\{d_1,\ldots,d_k\}}]_1)^j.$$

Remark 6.1.3. *Laba and Wang in [47] consider a restriction estimate for a certain kind of fractal measure in $\mathbb{R}^d$. The main ingredient in the proof of their main theorem is a decoupling estimate for a particular type of Cantor set on the line built out of a $\Lambda(p)$-set (see Lemma 5, Section 4, and Proposition 1 of [47] for more details, see also [9] for the existence of $\Lambda(p)$ sets). The techniques by which they upgrade a $\Lambda(p)$ set to a Cantor set multiscale decoupling theorem on the line can probably also be applied in our case, though here the point of view we take is more algebraic and is closer in spirit to the number theoretic side of things.*

6.1.3 Examples

As an illustration of the the tools developed above we can consider the case when $n = 2$ and then very explicitly study $\mathrm{DE}_{\ell^2\to L^4}([\mathcal{E}_q^{\{d_1,\ldots,d_k\}}]_j)$ as Proposition 6.3.1 turns such study into an optimization problem subject to a quadratic constraint which we can very explicitly compute. This combined with (6.3) allows us to upgrade l^2L^4 discrete restriction for an ellipsephic set to l^2L^4 decoupling for an arithmetic Cantor set. In particular, below is a summary of Examples 6.3.9-6.3.13 we derived in Section 6.3.3.

C_i	$\delta(i)$	$N(i)$	$K_4(C_i)$
$[C_q^{\{0,1\}}]_i, q > 2$	q^{-i}	2^i	$\sim (2^i)^{\frac{1}{4}\log_2(3/2)}$
$[C_3^{\{0,2\}}]_i$	3^{-i}	2^i	$\sim (2^i)^{\frac{1}{4}\log_2(3/2)}$
$[C_q^{\{0,1,2\}}]_i, q > 4$	q^{-i}	3^i	$\sim (3^i)^{\frac{1}{4}\log_3(15/7)}$
$[C_q^{\{0,1,3\}}]_i, q > 6$	q^{-i}	3^i	$\sim (3^i)^{\frac{1}{4}\log_3(5/3)}$
$[C_q^{\{0^2,1^2,\ldots,\lfloor\sqrt{q}\rfloor^2\}}]_i, q \geq \exp(\exp(O(\frac{1}{\varepsilon})))$	q^{-i}	$(\lfloor\sqrt{q}\rfloor + 1)^i$	$\lesssim_\varepsilon N(i)^\varepsilon$

Note that from the proof of these examples in Section 6.3.3, the implied constants do not depend on $\dim(C)$ or $N(1)$. We only studied the $n = 2$ case out for convenience to demonstrate our methods but it is not a serious constraint.

Remark 6.1.4. *The ellipsephic set associated to the Cantor set in the last row of the table above was considered by Biggs in [8, Corollary 1.4]. The result in that row should be read as follows: Fix an arbitrary $\varepsilon > 0$. Choose an integer $q \geq \exp(\exp(O(1/\varepsilon)))$ and consider $[C_q^{\{0^2,1^2,\ldots,\lfloor\sqrt{q}\rfloor^2\}}]_i$. Note that here the Cantor set depends on q and so also ε. Then we showed that the l^2L^4 decoupling constant for level i of this Cantor set is $\lesssim_\varepsilon N(i)^\varepsilon$ where $N(i) = (\lfloor q\rfloor + 1)^i$.*

Remark 6.1.5. *The example in the second row of the table above is associated to the ellipsephic set $[\mathcal{E}_3^{\{0,2\}}]_j$ which does have carryover. However, the map $x \mapsto x/2$ is a Freiman isomorphism between $[\mathcal{E}_3^{\{0,2\}}]_j$ and $[\mathcal{E}_3^{\{0,1\}}]_j$ and the latter ellipsephic set does not have carryover. Since Freiman isomorphisms do not change numerology (see the equality case of* (6.25)*), the numerology of the second row is the same as that of the first row.*

Remark 6.1.6. *Note that $C_q^{\{0,1,2\}}$ and $C_q^{\{0,1,3\}}$ for $q > 6$ have the same Hausdorff dimension but their associated l^2L^4 decoupling constants are different. In Proposition 6.3.6 we show* *that given a Hausdorff dimension $d = \log_s r$ with $0 < d < 1$ and $r, s \in \mathbb{N}$, there exists an arithmetic Cantor set C such that the associated decoupling exponent $\kappa_{2n}(C)$ as defined in (6.2) is as large as possible. This means that for arbitrary arithmetic Cantor sets $K_{2n}(C)$ does not just depend on the Cantor set, but rather also on arithmetic properties of the set.*

Remark 6.1.7. *A careful look at the proof of Example 6.3.11 (the third row in the table above) shows curiously that the optimizer of discrete restriction for $[\mathcal{E}_q^{\{0,1,2\}}]_1$, $q > 4$ (and hence also $[\mathcal{E}_q^{\{0,1,2\}}]_j$ by Proposition 6.3.5 because of lack of carryover). This is different from the other examples in Section 6.3.3 and the observation that the choice of $a : \{1, \ldots, \mathbb{N}\} \to \mathbb{R}_{\geq 0}$ being the constant function below witnesses the case of equality of the estimates*

$$\| \sum_{1 \leq \ell \leq N} a(\ell) e(\ell x) \|_{L^{2n}([0,1])} \leq N^{\frac{1}{2} - \frac{1}{n}} (\sum_{1 \leq \ell \leq N} |a(\ell)|^2)^{1/2}$$

and

$$\| \sum_{1 \leq n \leq N} a(n) e(nx + n^2 t) \|_{L^6([0,1]^2)} \lesssim_\varepsilon N^\varepsilon (\sum_{1 \leq n \leq N} |a(n)|^2)^{1/2}$$

for all $\{a(n)\} \in \ell^2(\mathbb{N})$. This example suggests potential differences between discrete restriction and solution counting problems in certain cases.

In Table 6.1 below we feed our results into Theorem 6.1.1. Each row should be compared to the estimate that $D_{12}(\delta(i)) \lesssim_\varepsilon \delta(i)^{-1/4-\varepsilon}$ obtained from a direct application of Bourgain-Demeter's decoupling theorem for the parabola.

Note that in the first four rows we have $N(1) \sim 1$ while in the second and last row we have $\dim(C) \sim 1$. Whether our estimates for $D_{12}(\delta(i))$ above are sharp remain an area to be explored (in other words, for example, is there an f Fourier supported in $\bigcup_{J \in [\mathscr{C}_3^{\{0,2\}}]_i} \Omega_J$ such that $D_{12}(\delta(i)) \gtrsim (2^i)^{\frac{1}{4} \log_2(3/2)}$). Continuing the discussion in Remark 6.1.4, the last row in the table above should be compared to [8, Corollary 1.4].

Finally the above methods are very efficient in studying the case when the ellipsephic set does not have carryover and some cases with carryover but which are Freiman isomorphic

Table 6.1: Decoupling estimates for certain fractal sets

C_i	$\delta(i)$	$N(i)$	Applying Theorem 6.1.1
$[C_q^{\{0,1\}}]_i, q > 2$	q^{-i}	2^i	$D_{12}(\delta(i)) \lesssim_{\varepsilon,\dim(C)} (2^i)^{\frac{1}{4}\log_2(3/2)+\varepsilon}$
$[C_3^{\{0,2\}}]_i$	3^{-i}	2^i	$D_{12}(\delta(i)) \lesssim_{\varepsilon} (2^i)^{\frac{1}{4}\log_2(3/2)+\varepsilon}$
$[C_q^{\{0,1,2\}}]_i, q > 4$	q^{-i}	3^i	$D_{12}(\delta(i)) \lesssim_{\varepsilon,\dim(C)} (3^i)^{\frac{1}{4}\log_3(15/7)+\varepsilon}$
$[C_q^{\{0,1,3\}}]_i, q > 6$	q^{-i}	3^i	$D_{12}(\delta(i)) \lesssim_{\varepsilon,\dim(C)} (3^i)^{\frac{1}{4}\log_3(5/3)+\varepsilon}$
$[C_q^{\{0^2,1^2,\ldots,\lfloor\sqrt{q}\rfloor^2\}}]_i, q \geq \exp(\exp(O(\frac{1}{\varepsilon})))$	q^{-i}	$(\lfloor\sqrt{q}\rfloor+1)^i$	$D_{12}(\delta(i)) \lesssim_{\varepsilon,N(1)} N(i)^{\varepsilon}$

to a case which has no carryover. To study the case when the ellipsehic set has carryover we develop an approximation (Proposition 6.3.7) which allows us to numerically approximate the l^2L^{2n} decoupling constant on $[0,1]$ for a given arithmetic Cantor set (see Section 6.3.4 for more details).

6.1.4 Application to solution counting

We end with some applications of our estimates to number theory, in particular to solution counting in Vinogradov systems.

6.1.4.1 The Cantor set $C_3^{\{0,1\}}$

Consider $[C_3^{\{0,1\}}]_j$ and the associated ellipsephic set $[\mathcal{E}_3^{\{0,1\}}]_j$. Note $\#[\mathcal{E}_3^{\{0,1\}}]_j \sim 2^i$. We first obtained that $K_4([C_3^{\{0,1\}}]_j) \sim A_4([\mathcal{E}_3^{\{0,1\}}]_j) \sim (3/2)^{j/4}$. This immediately implies that the number of 4-tuples to

$$x_1 + x_2 = x_3 + x_4$$

with $1 \le x_i \le 3^j$ and $x_i \in [\mathcal{E}_3^{\{0,1\}}]_j$ is $(3/2)^j 2^{2j} = 6^j$. This should be compared to solving $x_1 + x_2 = x_3 + x_4$ where $1 \le x_i \le 2^j$ which would give 8^j such 4-tuples. The 6 in 6^j can be explained by the fact that since $\mathcal{E}_3^{\{0,1\}}$ in this case has no carryover ($2 \cdot 1 < 3$), we can look one digit at a time and there are 6 solutions to $a + b = c + d$ where $a, b, c, d \in \{0, 1\}$.

Next we obtained that $D_{12}(\delta(j)) \lesssim_\varepsilon (3/2)^{j/4+\varepsilon}$ where $\delta(j) = 3^{-j}$. Using the standard reduction from decoupling estimates to solving Vinogradov [12] we see that the number of solutions to the system

$$\begin{aligned} x_1 + x_2 + \cdots + x_6 &= y_1 + y_2 + \cdots + y_6 \\ x_1^2 + x_2^2 + \cdots + x_6^2 &= y_1^2 + y_2^2 + \cdots + y_6^2 \end{aligned} \tag{6.4}$$

where $1 \le x_i, y_i \le 3^j$ and $x_i, y_i \in [\mathcal{E}_3^{\{0,1\}}]_j$ is $\lesssim_\varepsilon (\frac{3}{2})^{3j+\varepsilon} 2^{6j} = 6^{3j+O(\varepsilon)}$. This should be compared to the lower bound of $O(2^{6j})$ coming from the diagonal solutions.

6.1.4.2 The Cantor set $C_q^{\{0^2,1^2,\ldots,\lfloor\sqrt{q}\rfloor^2\}}$

Fix arbitrary $\varepsilon > 0$. Choose q an integer (not necessarily prime) such that $q \ge \exp(\exp(O(1/\varepsilon)))$ and consider the ellipsephic set $[\mathcal{E}_q^{\{0^2,1^2,\ldots,\lfloor\sqrt{q}\rfloor^2\}}]_j$ associated to the Cantor set $[C_q^{\{0^2,1^2,\ldots,\lfloor\sqrt{q}\rfloor^2\}}]_j$. Then the estimate that $D_{12}(\delta(j)) \lesssim_{\varepsilon,N(1)} N(j)^\varepsilon$ implies that the number of solutions to the system (6.4) where $1 \le x_i, y_i \le q^j$ and $x_i, y_i \in [\mathcal{E}_q^{\{0^2,1^2,\ldots,\lfloor\sqrt{q}\rfloor^2\}}]_j$ is $\lesssim_{\varepsilon,N(1)} N(j)^{6+\varepsilon}$. This rederives the implication obtained in [8, Corollary 1.4] (where our $N(j)$ is her Y).

Remark 6.1.8. *In the system considered in Section 6.1.4.1, our upper bound is quite large compared to the lower bound of 2^{6N} which come from the diagonal contribution. In the following, we argue that given an ellipsephic set (whose associated Cantor set has dimension d), then when the number of variables is sufficiently large depending on d, then the contribution of the non-diagonal solutions will be greater than that of the diagonal solutions.*

More precisely, fix an arbitrary arithmetic Cantor set $C_q^{\{d_1,\ldots,d_k\}}$ with Hausdorff dimension $d \in (0,1)$ and consider the associated ellipsephic set $\mathcal{E}_X := [\mathcal{E}_q^{\{d_1,\ldots,d_k\}}]_j$ where we have written $X = q^j$. Then $\#\mathcal{E}_X \sim X^d$. We consider the question of how many solutions are there to the *system*

$$\begin{aligned} x_1 + x_2 + \cdots + x_s &= x_{s+1} + x_{s+2} + \cdots + x_{2s} \\ x_1^2 + x_2^2 + \cdots + x_s^2 &= x_{s+1}^2 + x_{s+2}^2 + \cdots + x_{2s}^2 \end{aligned} \tag{6.5}$$

where $x_i \in \mathcal{E}_X$. *The contribution from the diagonal solutions is* $O(X^{sd})$. *We claim that for sufficiently large* s *there will always be more than* $O(X^{sd})$ *many solutions.*

Consider the map

$$\begin{aligned} \Sigma : (\mathcal{E}_X)^s &\longrightarrow [-sX, sX] \times [-sX^2, sX^2] \\ (a_1, a_2, \ldots, a_s) &\longmapsto (a_1 + \cdots + a_s, a_1^2 + \cdots + a_s^2) \end{aligned}$$

The map Σ *goes from a set of cardinality* $O(X^{sd})$ *to a set of cardinality* $O(s^2X^3)$. *For notational convenience let* $A_X = [-sX, sX] \times [-sX^2, sX^2]$. *The number of solutions* $J_s(X)$ *to* (6.5) *is bounded below by:*

$$\begin{aligned} J_s(X) = &\sum_{(n_1,n_2)\in A_X} \Big(\sum_{\substack{a_1^j+\cdots+a_s^j=n_j \\ a_i\in(\mathcal{E}_X)^s, j=1,2}} 1\Big)^2 \\ \geq &|A_X|^{-1}\Big(\sum_{(n_1,n_2)\in A_X} \sum_{\substack{a_1^j+\cdots+a_s^j=n_j \\ a_i\in(\mathcal{E}_X)^s, j=1,2}} 1\Big)^2 \\ = &(O(s^2X^3))^{-1} \cdot (O(X^{sd}))^2 = O(X^{2sd-3}/s^2) \end{aligned}$$

Therefore the number of solutions to (6.5) *is at least* $O(X^{2sd-3}/s^2)$. *Comparing this to the number of diagonal solutions* $O(X^{sd})$ *shows that for* s *sufficiently large (depending on Hausdorff dimension), the contribution of the off-diagonal solutions are more than the diagonal solutions.*

6.2 Proof of Theorem Theorem 6.1.1

Fix a Cantor set C (and its levels). Much like the proof of decoupling for the parabola in [51], the proof of Theorem 6.1.1 reduces to four lemmas: parabolic rescaling, bilinear reduction, the key estimate, and Hölder's inequality.

6.2.1 Parabolic rescaling and bilinear reduction

We first start with the parabolic rescaling lemma. The proof is fairly standard, but we include it here for convenience.

Lemma 6.2.1 (Parabolic rescaling). *Suppose* $0 \leq \delta(j) \leq \delta(i) \leq 1$ *and* $I \in P_{\delta(i)}(C_i)$*. Then*

$$\|\sum_{J \in P_{\delta(j)}(I \cap C_j)} f_{\Omega_J}\|_{L^p(\mathbb{R}^2)} \leq D_p(\delta(j-i)) \left(\sum_{J \in P_{\delta(j)}(I \cap C_j)} \|f_{\Omega_J}\|^2_{L^p(\mathbb{R}^2)} \right)^{1/2}. \tag{6.6}$$

Proof. Write $I = [a, a + \delta(i)]$. Consider the "Galilean transform" $S_I : \mathbb{R}^2 \to \mathbb{R}^2$ represented by the matrix

$$\begin{pmatrix} \delta(i)^{-1} & 0 \\ 0 & \delta(i)^{-2} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -2a & 1 \end{pmatrix}.$$

The key geometric observation is that since C_i is a level of a Cantor set (and Cantor set levels are similar), we have a bijection $P_{\delta(j)}(I \cap C_j) \to P_{\delta(j-i)}(C_{j-i})$ given by $J \mapsto J' = \delta(i)^{-1}(J - a)$, and furthermore,

$$S_I(\Omega_J - (a, a^2)) = \Omega_{J'}. \tag{6.7}$$

Define $g_I(y) := f(S_I^\top y) e(-S_I(a, a^2) \cdot y)$, so that $\widehat{g_I}(\eta) = \delta(i)^3 \widehat{f}(S_I^{-1}\eta + (a, a^2))$. With J, J' as above, we have

$$\begin{aligned} f_{\Omega_J}(x) &= \int_{\Omega_J} \widehat{f}(\xi) e(\xi \cdot x) \, d\xi \\ &= e(x \cdot (a, a^2)) \int_{\Omega_{J'}} \widehat{g_I}(\eta) e(\eta \cdot (S_I^{-1})^\top x) \, d\eta = e(x \cdot (a, a^2)) (g_I)_{\Omega_{J'}}((S_I^{-1})^\top x) \end{aligned}$$

where in the second equality we made the change of variables $\eta = S_I(\xi - (a, a^2))$ and used (6.7). Therefore,

$$|\sum_{J \in P_{\delta(j)}(I \cap C_j)} f_{\Omega_J}(x)| = |\sum_{J' \in P_{\delta(j-i)}(C_{j-i})} (g_I)_{\Omega_{J'}}((S_I^{-1})^\top x)|$$

and hence

$$\begin{aligned}\|\sum_{J \in P_{\delta(j)}(I \cap C_j)} f_{\Omega_J}\|_{L^p(\mathbb{R}^2)} &= \delta(i)^{-3/p} \|\sum_{J' \in P_{\delta(j-i)}(C_{j-i})} (g_I)_{\Omega_{J'}}\|_{L^p(\mathbb{R}^2)} \\ &\leq \delta(i)^{-3/p} D_p(\delta(j-i)) (\sum_{J' \in P_{\delta(j-i)}(C_{j-i})} \|(g_I)_{\Omega_{J'}}\|_{L^p(\mathbb{R}^2)}^2)^{1/2}.\end{aligned}$$

Reversing all the change of variables then obtains the right hand side of (6.6). □

Parabolic rescaling implies the following immediate corollary.

Corollary 6.2.2 (Almost multiplicativity)**.** *We have*

$$D_p(\delta(i+j)) \leq D_p(\delta(i)) D_p(\delta(j)).$$

Next we define the following bilinear constant. Let $0 \leq \delta(j) \leq \delta(i_1), \delta(i_2) \leq \delta(k) \leq 1$. Let $M_p(j, k, i_1, i_2)$ to be the best constant such that one has the estimate

$$\begin{aligned}&\int_{\mathbb{R}^2} |\sum_{J_1 \in P_{\delta(j)}(I_1 \cap C_j)} f_{\Omega_{J_1}}|^p |\sum_{J_2 \in P_{\delta(j)}(I_2 \cap C_j)} g_{\Omega_{J_2}}|^{2p} \\ &\quad \leq M_p(j, k, i_1, i_2)^{3p} (\sum_{J_1 \in P_{\delta(j)}(I_1 \cap C_j)} \|f_{\Omega_{J_1}}\|_{L^{3p}(\mathbb{R}^2)}^2)^{p/2} (\sum_{J_2 \in P_{\delta(j)}(I_2 \cap C_j)} \|g_{\Omega_{J_2}}\|_{L^{3p}(\mathbb{R}^2)}^2)^p\end{aligned}$$

for all $I_1 \in P_{\delta(i_1)}(C_{i_1})$ and $I_2 \in P_{\delta(i_2)}(C_{i_2})$ such that $d(I_1, I_2) \geq \delta(k)$ and all Schwartz functions f with Fourier support on $\bigcup_{J_1 \in P_{\delta(j)}(I_1 \cap C_j)} \Omega_{J_1}$ and Schwartz functions g with Fourier support on $\bigcup_{J_2 \in P_{\delta(j)}(I_2 \cap C_j)} \Omega_{J_2}$. Note that from Hölder,

$$M_p(j, k, i_1, i_2) \leq D_{3p}(\delta(j - i_1))^{1/3} D_{3p}(\delta(j - i_2))^{2/3}. \tag{6.8}$$

Lemma 6.2.3 (Bilinear reduction)**.** *If* $0 \leq \delta(j) \leq \delta(i) \leq 1$*, then*

$$D_{3p}(\delta(j)) \lesssim D_{3p}(\delta(j-i)) + N(i)^{O(1)} M_p(j, i, i, i). \tag{6.9}$$

Proof. Fix a Schwartz function f with Fourier support in $\bigcup_{J\in P_{\delta(j)}(C_j)}\Omega_J$. We have

$$\|\sum_{J\in P_{\delta(j)}(C_j)} f_{\Omega_J}\|^2_{L^{3p}(\mathbb{R}^2)} = \|\sum_{I_1,I_2\in P_{\delta(i)}(C_i)}\Big(\sum_{J_1\in P_{\delta(j)}(I_1\cap C_j)} f_{\Omega_{J_1}}\sum_{J_2\in P_{\delta(j)}(I_2\cap C_j)} f_{\Omega_{J_2}}\Big)\|_{L^{3p/2}(\mathbb{R}^2)}$$
$$\leq \|\sum_{\substack{I_1,I_2\in P_{\delta(i)}(C_i)\\ d(I_1,I_2)\leq\delta(i)}}(\cdots)\|_{L^{3p/2}(\mathbb{R}^2)} + \|\sum_{\substack{I_1,I_2\in P_{\delta(i)}(C_i)\\ d(I_1,I_2)\geq\delta(i)}}(\cdots)\|_{L^{3p/2}(\mathbb{R}^2)} \qquad (6.10)$$

By multiple applications of the Cauchy-Schwarz inequality, the first term of (6.10) is

$$\leq \sum_{\substack{I_1,I_2\in P_{\delta(i)}(C_i)\\ d(I_1,I_2)\leq\delta(i)}} \|\sum_{J_1\in P_{\delta(j)}(I_1\cap C_j)} f_{\Omega_{J_1}}\|_{L^{3p}(\mathbb{R}^2)}\|\sum_{J_2\in P_{\delta(j)}(I_2\cap C_j)} f_{\Omega_{J_2}}\|_{L^{3p}(\mathbb{R}^2)}$$
$$\leq \Big(\sum_{I_1\in P_{\delta(i)}(C_i)}\|\sum_{J_1\in P_{\delta(j)}(I_1\cap C_j)} f_{\Omega_{J_1}}\|^2_{L^{3p}(\mathbb{R}^2)}\Big)^{1/2}\times$$
$$\Big(\sum_{I_1\in P_{\delta(i)}(C_i)}\Big(\sum_{\substack{I_2\in P_{\delta(i)}(C_i)\\ d(I_1,I_2)\leq\delta(i)}}\|\sum_{J_2\in P_{\delta(j)}(I_2\cap C_j)} f_{\Omega_{J_2}}\|_{L^{3p}(\mathbb{R}^2)}\Big)^2\Big)^{1/2}$$
$$\lesssim \sum_{I\in P_{\delta(i)}(C_i)}\|\sum_{J\in P_{\delta(j)}(I\cap C_j)} f_{\Omega_J}\|^2_{L^{3p}(\mathbb{R}^2)}$$
$$\leq D_{3p}(\delta(j-i))^2\sum_{J\in P_{\delta(j)}(C_j)}\|f_{\Omega_J}\|^2_{L^{3p}(\mathbb{R}^2)}.$$

In the third inequality above, we used the fact that for a fixed I_1, the number of I_2 satisfying $d(I_1,I_2)\leq\delta(i)$ is $\lesssim 1$. In the last inequality above, we applied the definition of $D_{3p}(\delta(j-i))$. This gives the first term on the right hand side of (6.9). The second term of (6.10) is

$$\lesssim N(i)^{O(1)}\max_{\substack{I_1,I_2\in P_{\delta(i)}(C_i)\\ d(I_1,I_2)\geq\delta(i)}}\|\Big(\sum_{J_1\in P_{\delta(j)}(I_1\cap C_j)} f_{\Omega_{J_1}}\Big)\Big(\sum_{J_2\in P_{\delta(j)}(I_2\cap C_j)} f_{\Omega_{J_2}}\Big)\|_{L^{3p/2}(\mathbb{R}^2)}. \qquad (6.11)$$

For any two nonnegative functions F, G, we have $\int F^{3p/2}G^{3p/2}\leq(\int F^pG^{2p})^{1/2}(\int F^{2p}G^p)^{1/2}$ by Cauchy-Schwarz. Using this observation and applying the definition of $M_p(j,i,i,i)^{3p}$ gives that (6.11) is

$$\leq N(i)^{O(1)} M_p(j,i,i,i)^2 \times$$
$$\max_{\substack{I_1, I_2 \in P_{\delta(i)}(C_i) \\ d(I_1,I_2)\geq \delta(i)}} \Big(\sum_{J_1 \in P_{\delta(i)}(I_1\cap C_j)} \|f_{\Omega_{J_1}}\|_{L^{3p}(\mathbb{R}^2)}^2\Big)^{1/2} \Big(\sum_{J_2 \in P_{\delta(i)}(I_2\cap C_j)} \|f_{\Omega_{J_2}}\|_{L^{3p}(\mathbb{R}^2)}^2\Big)^{1/2}$$
$$\leq N(i)^{O(1)} M_p(j,i,i,i)^2 \Big(\sum_{J \in P_{\delta(j)}(C_j)} \|f_{\Omega_J}\|_{L^{3p}(\mathbb{R}^2)}^2\Big).$$

This gives the second term of the right hand side of (6.9) and thus completes the proof of the lemma. □

6.2.2 Key estimate

The main idea of this section is that while the key estimate for the proof of decoupling for the parabola in [51] follows from Plancherel (see [36, Lemma 3.8] with $k = 2$, [51, Remark 4], or [71, Proposition 19]), the key estimate here will follow from (6.2).

Lemma 6.2.4 (Key estimate)**.** *If* $0 \leq \delta(j) \leq \delta(i_1), \delta(i_1'), \delta(i_2) \leq \delta(k) \leq 1$ *with* $\delta(i_2)^2 \leq \delta(i_1') \leq \delta(i_1)$*, then for any* $\varepsilon > 0$*,*

$$M_p(j,k,i_1,i_2) \lesssim_{p,\varepsilon,\dim(C),N(1)} \delta(k)^{-O(1)} M_p(j,k,i_1',i_2) N(i_1'-i_1)^{\kappa_p(C)/3+\varepsilon/3}$$

where $\kappa_p(C)$ *is defined in* (6.2).

Proof. Fix arbitrary $\varepsilon > 0$ and arbitrary $I_1 \in P_{\delta(i_1)}(C_{i_1})$ and $I_2 \in P_{\delta(i_2)}(C_{i_2})$ such that $d(I_1, I_2) \geq \delta(k)$. Next fix arbitrary Schwartz functions f and g with Fourier support in $\bigcup_{J_1 \in P_{\delta(j)}(I_1\cap C_j)} \Omega_{J_1}$ and $\bigcup_{J_2 \in P_{\delta(j)}(I_2\cap C_j)} \Omega_{J_2}$, respectively. We may normalize f and g so that

$$\sum_{J_1 \in P_{\delta(j)}(I_1\cap C_j)} \|f_{\Omega_{J_1}}\|_{L^{3p}(\mathbb{R}^2)}^2 = \sum_{J_2 \in P_{\delta(j)}(I_2\cap C_j)} \|g_{\Omega_{J_2}}\|_{L^{3p}(\mathbb{R}^2)}^2 = 1. \tag{6.12}$$

Thus we need to show that

$$\int_{\mathbb{R}^2} \Big|\sum_{J_1 \in P_{\delta(j)}(I_1\cap C_j)} f_{\Omega_{J_1}}\Big|^p \Big|\sum_{J_2 \in P_{\delta(j)}(I_2\cap C_j)} g_{\Omega_{J_2}}\Big|^{2p}$$
$$\lesssim_{p,\varepsilon,\dim(C),N(1)} \delta(k)^{-O(p)} N(i_1'-i_1)^{p\kappa_p(C)+p\varepsilon} M_p(j,k,i_1',i_2)^{3p}.$$

Write $I_1 := [a, a+\delta(i_1)]$ and $I_2 := [b, b+\delta(i_2)]$. Assume that I_2 is to the left of I_1 and so $a - b > \delta(k)$; the case when I_2 is to the right of I_1 is similar.

We now essentially reduce to the case when $b = 0$. To see this, let $T_{I_2} = \left(\begin{smallmatrix} 1 & 0 \\ -2b & 1 \end{smallmatrix}\right)$, $\widetilde{f}_{I_2}(y) := f(T_{I_2}^\top y)e(-y \cdot T_{I_2}(b, b^2))$, and $\widetilde{g}_{I_2}(y) := g(T_{I_2}^\top y)e(-y \cdot T_{I_2}(b, b^2))$. By a similar argument as in the proof of Lemma 6.2.1, it suffices to show that

$$\int_{\mathbb{R}^2} |\sum_{J_1 \in P_{\delta(j)}((I_1 - b)\cap(C_j - b))} (\widetilde{f}_{I_2})_{\Omega_{J_1}}|^p |\sum_{J_2 \in P_{\delta(j)}([0,\delta(i_2)]\cap(C_j - b))} (\widetilde{g}_{I_2})_{\Omega_{J_2}}|^{2p}$$
$$\lesssim_{p,\varepsilon,\dim(C),N(1)} \delta(k)^{-O(p)} N(i_1' - i_1)^{p\kappa_p(C)+p\varepsilon} M_p(j,k,i_1',i_2)^{3p} \tag{6.13}$$

where

$$\sum_{J_1 \in P_{\delta(j)}((I_1 - b)\cap(C_j - b))} \|(\widetilde{f}_{I_2})_{\Omega_{J_1}}\|_{L^{3p}(\mathbb{R}^2)}^2 = \sum_{J_2 \in P_{\delta(j)}([0,\delta(i_2)]\cap(C_j - b))} \|(\widetilde{g}_{I_2})_{\Omega_{J_2}}\|_{L^{3p}(\mathbb{R}^2)}^2 = 1$$

since $\det T_{I_2} = 1$.

Let

$$G := \sum_{J_2 \in P_{\delta(j)}([0,\delta(i_2)]\cap(C_j - b))} (\widetilde{g}_{I_2})_{\Omega_{J_2}}.$$

Then G (and hence G^2) is Fourier supported in an $O(\delta(i_2)) \times O(\delta(i_2)^2 + \delta(j))$ rectangle centered at the origin. For each $J \in P_{\delta(i_1')}((I_1 - b)\cap(C_{i_1'} - b))$, let

$$F_J := \sum_{J_1 \in P_{\delta(j)}(J\cap(C_j - b))} (\widetilde{f}_{I_2})_{\Omega_{J_1}}.$$

The Fourier transform of F_J is supported in the horizontal strip $\{(\xi_1, \xi_2) : \xi_2 = \gamma_J^2 + O(\delta(i_1'))\}$ where γ_J is the center of J and γ_J is a distance $\gtrsim \delta(k)$ away from the origin. Since $\delta(j), \delta(i_2)^2 \leq \delta(i_1')$, $F_J G^2$ has Fourier transform supported in the horizontal strip $\{(\xi_1, \xi_2) : \xi_2 = \gamma_J^2 + O(\delta(i_1'))\}$ as well.

Using this notation, showing (6.13) is equivalent to showing that

$$\int_{\mathbb{R}^2} |\sum_{J \in P_{\delta(i_1')}((I_1 - b)\cap(C_{i_1'} - b))} F_J G^2|^p \lesssim_{p,\varepsilon,\dim(C),N(1)} \delta(k)^{-O(p)} N(i_1' - i_1)^{p\kappa_p(C)+p\varepsilon} M_p(j,k,i_1',i_2)^{3p}. \tag{6.14}$$

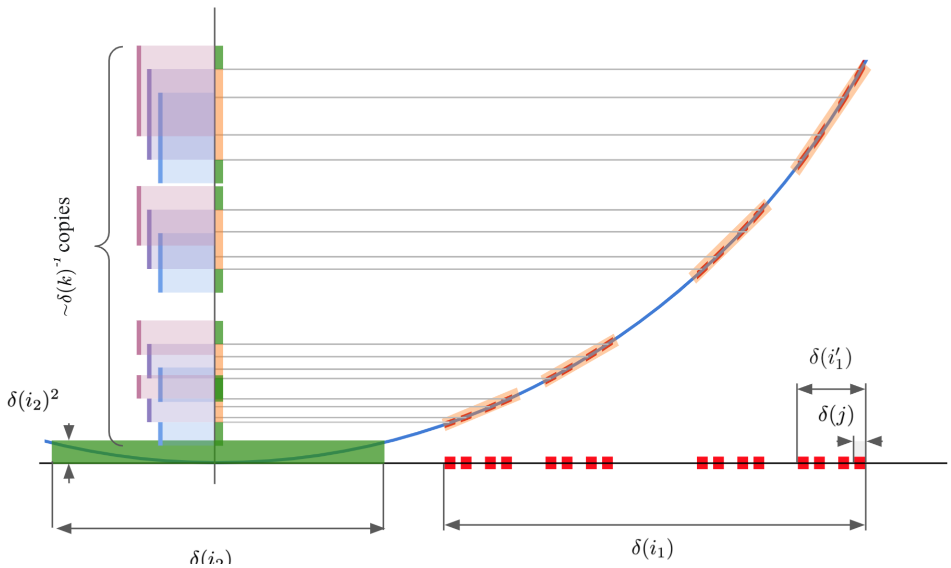


Figure 6.1: Scheme of the key estimate. Since I_1 is away from the origin and the parabola is Lipschitz on I_1 with Lipschitz constant $\gtrsim \delta(k)^{-O(1)}$, we know we can decouple vertically. The fact that we are multiplying by G^2, on the Fourier side amounts to convolving against $\widehat{G} * \widehat{G}$, which adds an uncertainty of size $O(\delta(i_2)^2)$ on each vertical level. This is acceptable because, we can cover the overlap by $\delta(k)^{-1}$ many copies of the orange sets (these copies are in shades of blue, purple and maroon in the picture).

We now claim that

$$\begin{aligned}&\|\sum_{J\in P_{\delta(i_1')}((I_1-b)\cap(C_{i_1'}-b))}F_JG^2\|_{L^p(\mathbb{R}^2)}\\&\qquad\lesssim_{p,\varepsilon,\dim(C),N(1)}\delta(k)^{-O(1)}N(i_1'-i_1)^{\kappa_p(C)+\varepsilon}(\sum_{J\in P_{\delta(i_1')}((I_1-b)\cap(C_{i_1'}-b))}\|F_JG^2\|^2_{L^p(\mathbb{R}^2)})^{1/2}\end{aligned}\tag{6.15}$$

which, as we will show, follows from an application of Cantor set decoupling for the line given by (6.2). Let us see how to use (6.15) to prove (6.14). Reversing the change of variables used to obtain (6.13) and applying the definition of $M_p(j,k,i_1',i_2)$ along with the normalization of g in (6.12) gives

$$\|F_JG^2\|_{L^p(\mathbb{R}^2)}\leq M_p(j,k,i_1',i_2)^3(\sum_{J_1\in P_{\delta(j)}((J+b)\cap C_j)}\|f_{\Omega_{J_1}}\|^2_{L^{3p}(\mathbb{R}^2)})^{1/2}\tag{6.16}$$

for each $J\in P_{\delta(i_1')}((I_1-b)\cap(C_{i_1'}-b))$. Combining (6.15) with (6.16) and using our normalization of f in (6.12) then proves (6.14). Thus it remains to prove (6.15).

First since $p\geq 2$, by Minkowski's inequality, it suffices to prove that for fixed $x\in\mathbb{R}^2$,

$$\begin{aligned}&\int_{\mathbb{R}}|\sum_{J\in P_{\delta(i_1')}((I_1-b)\cap(C_{i_1'}-b))}F_J(x,y)G(x,y)^2|^p\,dy\\&\lesssim_{p,\varepsilon,\dim(C),N(1)}\delta(k)^{-O(p)}N(i_1'-i_1)^{p\kappa_p(C)+p\varepsilon}(\sum_{J\in P_{\delta(i_1')}((I_1-b)\cap(C_{i_1'}-b))}(\int_{\mathbb{R}}|F_J(x,y)G(x,y)^2|^p\,dy)^{2/p})^{p/2}.\end{aligned}\tag{6.17}$$

Indeed, once we obtain the above inequality, we can prove (6.15) by just integrating in x. For fixed x, the Fourier transform in y of $F_J(x,y)G(x,y)^2$ is supported on an interval of length $O(\delta(i_1'))$ centered at γ_J^2 where $\gamma_J\gtrsim\delta(k)$ is the center of the interval $J\in P_{\delta(i_1')}((I_1-b)\cap(C_{i_1'}-b))$. Note that the implied constant in $O(\delta(i_1'))$ is independent of J.

Now suppose $F_{J_1}G^2$ and $F_{J_2}G^2$ had overlapping Fourier supports. Then $\gamma_{J_1}^2=\gamma_{J_2}^2+O(\delta(i_1'))$ and hence $\gamma_{J_1}=\gamma_{J_2}+O(\delta(i_1')\delta(k)^{-O(1)})$ since $\gamma_{J_1},\gamma_{J_2}\gtrsim\delta(k)$. Thus (6.17) now follows if we can show that

$$\int_{\mathbb{R}} |\sum_{J \in P_{\delta(i_1')}((I_1-b)\cap(C_{i_1'}-b))} f_{cJ}(y)|^p\, dy$$
$$\lesssim_{p,\varepsilon,\dim(C),N(1)} \delta(k)^{-O(p)} N(i_1'-i_1)^{p\kappa_p(C)+p\varepsilon}(\sum_{J \in P_{\delta(i_1')}((I_1-b)\cap(C_{i_1'}-b))} (\int_{\mathbb{R}} |f_{cJ}(y)|^p\, dy)^{2/p})^{p/2}$$

for $1 \leq c \lesssim \delta(k)^{-O(1)}$ and for arbitrary Schwartz functions f. Here, cJ denotes the interval having the same center as J but of length $c|J|$. By rescaling I_1 and using the fact that decoupling constants are translation invariant, this then reduces to showing that

$$\|\sum_{J \in P_{\delta(i)}(C_i)} f_{cJ}\|_{L^p(\mathbb{R})} \lesssim_{p,\varepsilon,\dim(C),N(1)} cN(i)^{\kappa_p(C)+\varepsilon}(\sum_{J \in P_{\delta(i)}(C_i)} \|f_{cJ}\|^2_{L^p(\mathbb{R})})^{1/2} \tag{6.18}$$

for $c \geq 1$ and for arbitrary Schwartz functions f. (Here $i = i_1' - i_1$.)

To show (6.18), we can assume that $c \geq 1$ is an integer. We can find translations $\{\tau_k : 1 \leq k \leq c\}$ such that for any $J \in P_{\delta(i)}(C_i)$, the interval cJ is covered by the union of $\{\tau_k(J) : 1 \leq k \leq c\}$. Therefore

$$\begin{aligned}
\|\sum_{J \in P_{\delta(i)}(C_i)} f_{cJ}\|_{L^p(\mathbb{R})} &= \|\sum_{k=1}^{c} \sum_{J \in P_{\delta(i)}(C_i)} (f_{cJ})_{\tau_k(J)}\|_{L^p(\mathbb{R})} \\
&\leq c \sup_k \|\sum_{J \in P_{\delta(i)}(C_i)} (f_{cJ})_{\tau_k(J)}\|_{L^p(\mathbb{R})} \\
&\lesssim_{p,\varepsilon,\dim(C),N(1)} cN(i)^{\kappa_p(C)+\varepsilon} \sup_k (\sum_{J \in P_{\delta(i)}(C_i)} \|(f_{cJ})_{\tau_k(J)}\|^2_{L^p(\mathbb{R})})^{1/2} \\
&\lesssim_{p,\varepsilon,\dim(C),N(1)} cN(i)^{\kappa_p(C)+\varepsilon} (\sum_{J \in P_{\delta(i)}(C_i)} \|f_{cJ}\|^2_{L^p(\mathbb{R})})^{1/2}
\end{aligned}$$

where the third inequality is because decoupling is invariant under translation and (6.2), and the last inequality is by boundedness of the Hilbert transform in $L^p(\mathbb{R})$, $1 < p < \infty$, (see for example [29, p. 59]). This completes the proof of (6.18) and hence the proof of Lemma 6.2.4. □

6.2.3 Iteration

We first have the following lemma which allows us to interchange the last two indices in $M_p(j,k,i_1,i_2)$.

Lemma 6.2.5. *If* $0 \le \delta(j) \le \delta(i_1) \le \delta(i_2) \le \delta(k) \le 1$, *then*

$$M_p(j,k,i_1,i_2) \le M_p(j,k,i_2,i_1)^{1/2} D_{3p}(\delta(j-i_2))^{1/2}.$$

Proof. This lemma follows from $\int F^p G^{2p} \le (\int F^{2p}G^p)^{1/2}(\int G^{3p})^{1/2}$ and applying the definition of $M_p(j,k,i_2,i_1)$ and parabolic rescaling. □

We are now in a good position to conclude the proof of Theorem 6.1.1. After normalization, the iteration is essentially the same as in [51]. The proof follows via a contradiction argument, combining the previous lemmas and using an iteration argument. We start normalizing the main objects that we have been considering in order to simplify our argument. Let

$$D'_{3p}(\delta(i)) := N(i)^{-\kappa_p(C)} D_{3p}(\delta(i))$$

and

$$M'_p(j,k,i_1,i_2) := M_p(j,k,i_1,i_2)(N(j-i_1)N(j-i_2)^2)^{-\kappa_p(C)/3}.$$

With this definition, after multiplying both sides of Lemma 6.2.3 by $N(j-i)^{-\kappa_p(C)}$, we have that if $0 \le \delta(j) \le \delta(i) \le 1$, then

$$D'_{3p}(\delta(j)) \lesssim N(i)^{-\kappa_p(C)} D'_{3p}(\delta(j-i)) + N(i)^{O(1)} M'_p(j,i,i,i). \tag{6.19}$$

The key estimate Lemma 6.2.4 now becomes that if $0 \le \delta(j) \le \delta(i_1), \delta(i'_1), \delta(i_2) \le \delta(k) \le 1$ with $\delta(i_2)^2 \le \delta(i'_1) \le \delta(i_1)$, then for any $\varepsilon > 0$,

$$M'_p(j,k,i_1,i_2) \lesssim_{p,\varepsilon,\dim(C),N(1)} \delta(k)^{-A} N(i'_1 - i_1)^{\varepsilon/3} M'_p(j,k,i'_1,i_2) \tag{6.20}$$

for some absolute constant A. Also, Lemma 6.2.5 above becomes

$$M'_p(j,k,i_1,i_2) \le M'_p(j,k,i_2,i_1)^{1/2} D'_{3p}(\delta(j-i_2))^{1/2}. \tag{6.21}$$

Proof of Theorem 6.1.1. Let λ be the least exponent for which the following statement is true:

$$D'_{3p}(\delta(j)) \lesssim_{p,\varepsilon,\dim(C),N(1)} N(j)^{\lambda+\varepsilon} \qquad \text{for all } j \geq 0 \text{ and } \varepsilon > 0. \tag{6.22}$$

Trivially, $D'_{3p}(\delta(i)) \leq N(i)^{\frac{1}{2}-\kappa_{3p}(C)}$ and so (6.22) is equivalent to the statement that

$$D'_{3p}(\delta(j)) \lesssim_{p,\varepsilon,\dim(C),N(1)} N(j)^{\lambda+\varepsilon} \qquad \text{for all } j \gtrsim 1 \text{ and } 0 < \varepsilon \lesssim 1.$$

If $\lambda = 0$, then we are done, so we assume towards a contradiction that $\lambda > 0$. Fix arbitrary $\varepsilon > 0$, we may assume that $\varepsilon < 1$.

If $1 \leq a \leq \frac{j}{4i}$, then $j \geq 4ai \geq 2ai \geq ai \geq i$ which imply that we can talk about $M'_p(j,i,2ai,i)$ and $M'_p(j,i,4ai,2ai)$. Applying (6.21), (6.20), and (6.22) in that order obtains

$$\begin{aligned}
M'_p(j,i,2ai,ai) &\leq M'_p(j,i,ai,2ai)^{1/2} D'_{3p}(\delta(j-ai))^{1/2} \\
&\lesssim_{p,\varepsilon,\dim(C),N(1)} M'_p(j,i,4ai,2ai)^{1/2}\delta(i)^{-A/2}N(4ai-ai)^{\varepsilon/6}D'_{3p}(\delta(j-ai))^{1/2} \\
&\lesssim_{p,\varepsilon,\dim(C),N(1)} M'_p(j,i,4ai,2ai)^{1/2}\delta(i)^{-A/2}N(4ai-ai)^{\varepsilon/6}N(j-ai)^{\frac{\lambda}{2}+\frac{\varepsilon}{2}} \\
&= M'_p(j,i,4ai,2ai)^{1/2}\delta(i)^{-A/2}N(j)^{\frac{\lambda+\varepsilon}{2}}N(i)^{-a\lambda/2}.
\end{aligned}$$

Hence we have shown that for $1 \leq a \leq \frac{j}{4i}$

$$M'_p(j,i,2ai,ai) \leq C_{p,\varepsilon,\dim(C),N(1)} M'_p(j,i,4ai,2ai)^{1/2}\delta(i)^{-A/2}N(i)^{-a\lambda/2}N(j)^{\frac{\lambda+\varepsilon}{2}}$$

for some constant $C_{p,\varepsilon,\dim(C),N(1)}$ depending only on p, ε, $\dim(C)$ and $N(1)$ and A is an absolute constant.

Then, we multiply both sides of the previous inequality by $N(j)^{-\lambda}$ and raise both sides to the $1/a$ power to obtain that for every integer a such that $1 \leq a \leq \frac{j}{4i}$,

$$\begin{aligned}
&(N(j)^{-\lambda}M'_p(j,i,2ai,ai))^{1/a} \\
&\qquad \leq (C_{p,\varepsilon,\dim(C),N(1)}\delta(i)^{-A/2}N(j)^{\varepsilon/2})^{1/a}N(i)^{-\lambda/2}(N(j)^{-\lambda}M'_p(j,i,4ai,2ai))^{1/(2a)}.
\end{aligned}$$

Therefore, for all $k \in \mathbb{N}$ with $2^{k+1} \le j/i$, the following inequality holds:

$$\begin{aligned}
&N(j)^{-\lambda} M'_p(j,i,2i,i) \\
&\quad \le \left(\prod_{n=0}^{k-1} (C_{p,\varepsilon,\dim(C),N(1)} \delta(i)^{-A/2} N(j)^{\varepsilon/2})^{1/2^n} \right) N(i)^{-k\lambda/2} \left(N(j)^{-\lambda} M'_p(j,i,2^{k+1}i,2^k i) \right)^{1/2^k} \\
&\quad \lesssim_{p,\varepsilon,\dim(C),N(1)} (\delta(i)^{-A/2} N(j)^{\varepsilon/2})^{\sum_{n=0}^{k-1} \frac{1}{2^n}} N(i)^{-k\lambda/2} N(j)^{\varepsilon/2^k} \\
&\quad \lesssim_{p,\varepsilon,\dim(C),N(1)} \delta(i)^{-O(1)} N(i)^{-k\lambda/2} N(j)^{\varepsilon} \qquad (6.23)
\end{aligned}$$

where in the second inequality we have used that

$$\begin{aligned}
M'_p(j,i,2^{k+1}i,2^k i) &\le D'_{3p}(\delta(j-2^{k+1}i))^{1/3} D'_{3p}(\delta(j-2^k i))^{2/3} \\
&\lesssim_{p,\varepsilon,\dim(C),N(1)} N(j-2^{k+1}i)^{(\lambda+\varepsilon)/3} N(j-2^k i)^{2(\lambda+\varepsilon)/3} \le N(j)^{\lambda+\varepsilon}
\end{aligned}$$

which follows from (6.8) and that N is increasing.

Suppose i, j, and k are such that $N(i) = N(j)^{1/2^{k+1}}$ and so by multiplicativity of $N(\cdot)$, $2^{k+1} i = j$. Using (6.1), (6.19), (6.20), (6.22) and (6.23) we conclude that

$$\begin{aligned}
D'_{3p}(\delta(j)) &\lesssim_{p,\varepsilon,\dim(C),N(1)} N(i)^{-\kappa_p(C)} D'_{3p}(\delta(j-i)) + \delta(i)^{-O(1)} N(i)^{\varepsilon} M'_p(j,i,2i,i) \\
&\lesssim_{p,\varepsilon,\dim(C),N(1)} N(i)^{-\kappa_p(C)} N(j-i)^{\lambda+\varepsilon} + \delta(i)^{-O(1)} N(i)^{\varepsilon-k\lambda/2} N(j)^{\lambda+\varepsilon} \\
&\lesssim_{p,\varepsilon,\dim(C),N(1)} N(j)^{\lambda+\varepsilon} N(i)^{-\lambda} + N(i)^{O(\frac{1}{\dim(C)})+\varepsilon-k\lambda/2} N(j)^{\lambda+\varepsilon} \\
&\lesssim_{p,\varepsilon,\dim(C),N(1)} N(j)^{\lambda(1-\frac{1}{2^{k+1}})+\varepsilon} + N(j)^{\lambda[1-\frac{1}{2^{k+1}}(\frac{k}{2}-\frac{O(\frac{1}{\dim(C)})}{\lambda}-\frac{\varepsilon}{\lambda})]} N(j)^{\varepsilon}.
\end{aligned}$$

Choose K so that $\frac{K}{2} - \frac{O(\frac{1}{\dim(C)})}{\lambda} - \frac{\varepsilon}{\lambda} \ge 1$. We have then shown that if $j = 2^{K+1}\mathbb{N}$, then for every $\varepsilon > 0$,

$$D'_{3p}(\delta(j)) \lesssim_{p,\varepsilon,\dim(C),N(1)} N(j)^{\lambda(1-\frac{1}{2^{K+1}})+\varepsilon}.$$

We now upgrade this to be a statement for all $j \ge 0$. We use almost multiplicativity, Corollary 6.2.2. For $n \ge 0$ and j such that $2^{K+1}n \le j \le 2^{K+1}(n+1)$. Note that

$$N(2^{K+1}n) \le N(j) \le N(2^{K+1}(n+1))$$

and

$$\delta(2^{K+1}n) \geq \delta(j) \geq \delta(2^{K+1}(n+1)).$$

From almost multiplicativity and the trivial bound,

$$\begin{aligned} D'_{3p}(\delta(j)) &\leq D'_{3p}(\delta(2^{K+1}n))D'_{3p}(\delta(j-2^{K+1}n)) \\ &\lesssim_{p,\varepsilon,\dim(C),N(1)} N(2^{K+1}n)^{\lambda(1-\frac{1}{2^{K+1}})+\varepsilon} N(j-2^{K+1}n)^{1/2} \\ &\lesssim_{p,\varepsilon,\dim(C),N(1)} N(j)^{\lambda(1-\frac{1}{2^{K+1}})+\varepsilon} (\frac{N(2^{K+1}(n+1))}{N(2^{K+1}n)})^{1/2} \\ &\lesssim_{p,\varepsilon,\dim(C),N(1)} N(j)^{\lambda(1-\frac{1}{2^{K+1}})+\varepsilon} N(1)^{2^K}. \end{aligned}$$

Therefore we have upgraded this estimate to be that for all $j \geq 0$,

$$D'_{3p}(\delta(j)) \lesssim_{p,\varepsilon,\dim(C),N(1),\lambda} N(j)^{\lambda(1-\frac{1}{2^{K+1}})+\varepsilon}.$$

This contradicts the minimality of λ. $\square$

Following the same ideas from the iteration in [51], if there is no dependence on $\dim(C)$ and $N(1)$ in (6.2) (as is the case for our examples in Section 6.3.3), the dependence on $\dim(C)$ and $N(1)$ in $D_{3p}(\delta(i))$ is $\exp(\exp(O(\frac{1}{\varepsilon \dim(C)})) \log N(1))$. If there is some dependence on $\dim(C)$ and $N(1)$ in (6.2), then an examination of the proof above shows that this same exact dependence shows up again in $D_{3p}(\delta(i))$.

6.3 Decoupling for Cantor sets in dimension one

In Theorem 6.1.1, we reduced the study of decoupling for a Cantor set on the parabola to that on the line. We now proceed to carefully study the case of l^2L^{2n} decoupling for a Cantor subset of $[0, 1]$. The use of $2n$ allows us to connect decoupling to number theory.

By rescaling a and f, we have that

$$\mathrm{DE}_{\ell^2 \to L^p}(S) = \sup\{\| \sum_{\ell \in S} a(\ell)e(\ell \cdot x)\|_{L^p([0,1]^m)} \mid a : S \to \mathbb{R}_{\geq 0}, \sum_{\ell \in S} |a(\ell)|^2 = 1\}$$

and

$$K_p(\Omega) = \sup\{\|\sum_I f_I\|_{L^p(\mathbb{R})} \mid f \text{ Schwartz}, \sum_I \|f_I\|^2_{L^p(\mathbb{R})} = 1\}.$$

Making use of that $2n$ is even, we have the following proposition.

Proposition 6.3.1. *Let $S \subset \mathbb{Z}^m$. Then*

$$\mathrm{DE}_{\ell^2\to L^{2n}}(S)^{2n} = \sup\left\{ \sum_{t\in\mathbb{Z}^m} \left(\sum_{\substack{\ell_1,\ldots,\ell_n\in S\\ \ell_1+\cdots+\ell_n=t}} \prod_{i=1}^n a(\ell_i) \right)^2 \,\middle|\, a : S \to \mathbb{R}_{\geq 0} \text{ and } \sum_{\ell\in S} |a(\ell)|^2 = 1 \right\}. \tag{6.24}$$

Proof. This follows immediately from the observation that

$$\|\sum_{\ell\in S} a(\ell)e(\ell\cdot x)\|^{2n}_{L^{2n}([0,1]^m)} = \left\| \sum_{t\in\mathbb{Z}^m} \left(\sum_{\substack{\ell_1,\ldots,\ell_n\in S\\ \ell_1+\cdots+\ell_n=t}} \prod_{i=1}^n a(\ell_i) \right) e^{2\pi i t\cdot x} \right\|^2_{L^2([0,1]^m)}$$

and then applying Plancherel. □

6.3.1 Properties of $\mathrm{DE}_{\ell^2\to L^{2n}}$

For $S \subset \mathbb{Z}^m$ and $S' \subset \mathbb{Z}^{m'}$, we say that $\phi : S \to S'$ is a *Freiman homomorphism of order* n if

$$\text{for all } x_1, \ldots, x_n, y_1, \ldots, y_n \in S, \qquad \sum_{i=1}^n x_i = \sum_{i=1}^n y_i \implies \sum_{i=1}^n \phi(x_i) = \sum_{i=1}^n \phi(y_i)$$

(see, e.g. [72, Section 5.3]). We say that ϕ is a *Freiman isomorphism of order* n if ϕ is a bijection and both ϕ and ϕ^{-1} are Freiman homomorphisms of order n.

It follows immediately from Proposition 6.3.1 that if ϕ is a bijective Freiman homomorphism of order n, then

$$\mathrm{DE}_{\ell^2\to L^{2n}}(S) \leq \mathrm{DE}_{\ell^2\to L^{2n}}(S'), \tag{6.25}$$

and that (6.25) becomes an equality if ϕ is a Freiman isomorphism of order n. We also have the following.

Proposition 6.3.2. *Let $S \subset \mathbb{Z}^m$ and $S' \subset \mathbb{Z}^{m'}$, and let $\phi : S \to S'$ be a bijection. Let*

$$D = \left\{ \sum_{i=1}^n \phi(x_i) - \sum_{i=1}^n \phi(y_i) \;\middle|\; x_1, \ldots, x_n, y_1, \ldots, y_n \in S \text{ and } \sum_{i=1}^n x_i = \sum_{i=1}^n y_i \right\} \tag{6.26}$$

Then

$$\mathrm{DE}_{\ell^2 \to L^{2n}}(S) \le |D|^{\frac{1}{2n}} \, \mathrm{DE}_{\ell^2 \to L^{2n}}(S'). \tag{6.27}$$

Note that if ϕ is a bijective Freiman homomorphism of order n, then $D = \{0\}$, so (6.27) becomes (6.25). Thus, Proposition 6.3.2 is a variant of (6.25) for when the bijection ϕ is not a Freiman homomorphism of order n, but is "close" to being one (in the sense that D is small). This proposition should also be compared to [8, Lemma 2.2].

Proof. Let $a : S \to \mathbb{R}_{\ge 0}$ such that $\sum_{\ell \in S} a(\ell)^2 = 1$. Define $a' : S' \to \mathbb{R}_{\ge 0}$ by $a' = a \circ \phi^{-1}$. Then by the definition of D,

$$\begin{aligned}
\sum_{t \in \mathbb{Z}^m} \left(\sum_{\substack{x_1, \ldots, x_n \in S \\ x_1 + \cdots + x_n = t}} \prod_{i=1}^n a(x_i) \right)^2 &= \sum_{\substack{x_1, \ldots, x_n \in S \\ y_1, \ldots, y_n \in S \\ x_1 + \cdots + x_n = y_1 + \cdots + y_n}} \left(\prod_{i=1}^n a(x_i) \right) \left(\prod_{i=1}^n a(y_i) \right) \\
&\le \sum_{t \in D} \sum_{\substack{x'_1, \ldots, x'_n \in S' \\ y'_1, \ldots, y'_n \in S' \\ \sum_{i=1}^n x'_i - \sum_{i=1}^n y'_i = t}} \left(\prod_{i=1}^n a'(x'_i) \right) \left(\prod_{i=1}^n a'(y'_i) \right)
\end{aligned} \tag{6.28}$$

Define

$$B(t) = \sum_{x'_1, \ldots, x'_n \in S : \sum_{i=1}^n x'_i = t} \prod_{i=1}^n a'(x'_i)$$

so that the right-hand side of (6.28) is

$$\begin{aligned}
= \sum_{s,t \in \mathbb{Z}^{m'} : s - t \in D} B(s)B(t) &\le \sum_{s,t \in \mathbb{Z}^{m'} : s - t \in D} \frac{B(s)^2 + B(t)^2}{2} \\
&= \frac{1}{2} \sum_{\substack{s,t \in \mathbb{Z}^{m'} \\ s - t \in D}} B(s)^2 + \frac{1}{2} \sum_{\substack{s,t \in \mathbb{Z}^{m'} \\ s - t \in D}} B(t)^2 \le |D| \sum_{t \in \mathbb{Z}^{m'}} B(t)^2
\end{aligned}$$

Thus,

$$\sum_{t\in\mathbb{Z}^m}\left(\sum_{\substack{x_1,\dots,x_n\in S\\x_1+\cdots+x_n=t}}\prod_{i=1}^{n}a(x_i)\right)^2\le|D|\sum_{t'\in\mathbb{Z}^{m'}}\left(\sum_{\substack{x'_1,\dots,x'_n\in S'\\x'_1+\cdots+x'_n=t'}}\prod_{i=1}^{n}a'(x'_i)\right)^2$$

which by Proposition 6.3.1 implies (6.27). □

Proposition 6.3.3. *For $S\subset\mathbb{Z}^m$, $S'\subset\mathbb{Z}^{m'}$,*

$$\mathrm{DE}_{\ell^2\to L^{2n}}(S\times S')=\mathrm{DE}_{\ell^2\to L^{2n}}(S)\,\mathrm{DE}_{\ell^2\to L^{2n}}(S')$$

Proof. First, we will show that

$$\mathrm{DE}_{\ell^2\to L^{2n}}(S\times S')\ge\mathrm{DE}_{\ell^2\to L^{2n}}(S)\,\mathrm{DE}_{\ell^2\to L^{2n}}(S'). \tag{6.29}$$

For $a:S\to\mathbb{R}_{\ge0}$ and $a':S'\to\mathbb{R}_{\ge0}$, we define $(a\otimes a'):S\times S'\to\mathbb{R}_{\ge0}$ by

$$(a\otimes a')(l,l')=a(l)a'(l').$$

Observe that

$$\begin{aligned}&\|\sum_{(l,l')\in S\times S'}(a\otimes a')(\ell,\ell')e((\ell,\ell')\cdot(x,x'))\|_{L^{2n}(\mathbb{T}^{m+m'})}\\&\qquad=\|\sum_{\ell\in S}a(\ell)e(\ell\cdot x)\|_{L^{2n}(\mathbb{T}^m)}\|\sum_{\ell'\in S'}a'(\ell')e(\ell'\cdot x')\|_{L^{2n}(\mathbb{T}^{m'})}\end{aligned}$$

and

$$\|a\otimes a'\|_{\ell^2(S\times S')}=\|a\|_{\ell^2(S)}\|a'\|_{\ell^2(S')}.$$

We therefore obtain (6.29).

It now remains to show the reverse inequality

$$\mathrm{DE}_{\ell^2\to L^{2n}}(S\times S')\le\mathrm{DE}_{\ell^2\to L^{2n}}(S)\,\mathrm{DE}_{\ell^2\to L^{2n}}(S'). \tag{6.30}$$

Fix $x' \in \mathbb{T}^{m'}$. Then we view $b_{x'}(\ell) := \sum_{\ell' \in S'} a(\ell, \ell')e(\ell' \cdot x')$ as a function of $\ell \in S$. We have

$$\begin{aligned}\|\sum_{\ell \in S}(\sum_{\ell' \in S'} a(\ell,\ell')e(\ell'\cdot x'))e(\ell \cdot x)\|_{L^{2n}_x(\mathbb{T}^m)}^{2n} &= \|\sum_{\ell\in S} b_{x'}(\ell)e(\ell\cdot x)\|_{L^{2n}_x(\mathbb{T}^m)}^{2n} \\ &\leq \mathrm{DE}_{\ell^2\to L^{2n}}(S)^{2n}(\sum_{\ell\in S}|b_{x'}(\ell)|^2)^{2n/2}.\end{aligned}$$

Next integrating in $\mathbb{T}^{m'}$ gives

$$\|\sum_{\ell\in S,\ell'\in S'} a(\ell,\ell')e(\ell'\cdot x')e(\ell\cdot x)\|_{L^{2n}(\mathbb{T}^{m+m'})} \leq \mathrm{DE}_{\ell^2\to L^{2n}}(S)\|(\sum_{\ell\in S}|\sum_{\ell'\in S'}a(\ell,\ell')e(\ell'\cdot x')|^2)^{1/2}\|_{L^{2n}_{x'}(\mathbb{T}^{m'})}.$$

Since $2n \geq 2$, applying Minkowski's inequality allows us to interchange the $L^{2n}_{x'}$ and the ℓ^2 sum over $\ell \in S$. Thus the above is controlled by

$$\mathrm{DE}_{\ell^2\to L^{2n}}(S)(\sum_{\ell\in S}\|\sum_{\ell'\in S'}a(\ell,\ell')e(\ell'\cdot x')\|^2_{L^{2n}_{x'}(\mathbb{T}^{m'})})^{1/2} \leq \mathrm{DE}_{\ell^2\to L^{2n}}(S)\,\mathrm{DE}_{\ell^2\to L^{2n}}(S')(\sum_{\ell\in S,\ell'\in S'}|a(\ell,\ell')|^2)^{1/2}$$

from which (6.30) follows. □

6.3.2 Arithmetic Cantor sets and ellipsephic sets

Let

$$\alpha_{2n}(\mathcal{E}_q^{\{d_1,\dots,d_k\}}) := \limsup_{j\to\infty}\frac{\log \mathrm{DE}_{\ell^2\to L^{2n}}([\mathcal{E}_q^{\{d_1,\dots,d_k\}}]_j)}{\log k^j} \tag{6.31}$$

and similarly let

$$\kappa_{2n}(C_q^{\{d_1,\dots,d_k\}}) := \limsup_{j\to\infty}\frac{\log K_{2n}([\mathscr{C}_q^{\{d_1,\dots,d_k\}}]_j)}{\log k^j}. \tag{6.32}$$

We call these the decoupling exponents of $\mathrm{DE}_{\ell^2\to L^{2n}}([\mathcal{E}_q^{\{d_1,\dots,d_k\}}]_j)$ and $K_{2n}([\mathscr{C}_q^{\{d_1,\dots,d_k\}}]_j)$, respectively.

In this section we will show that from a decoupling point of view the sets $[C_q^{\{d_1,\dots,d_k\}}]_j$ and $[\mathcal{E}_q^{\{d_1,\dots,d_k\}}]_j$ have similar nature. Namely, we will prove the following proposition. This allows us to upgrade results obtained from discrete restriction of ellipsephic sets to decoupling for arithmetic Cantor sets. In particular, later in Proposition 6.3.5 when the ellipsephic set does not have carryover, the discrete restriction problem has a particularly nice structure.

Proposition 6.3.4. *For an integer $n \geq 1$,*

$$K_{2n}([\mathscr{C}_q^{\{d_1,\dots,d_k\}}]_j) \sim \mathrm{DE}_{\ell^2\to L^{2n}}([\mathcal{E}_q^{\{d_1,\dots,d_k\}}]_j) \tag{6.33}$$

where the implicit constant is an absolute constant. In particular by (6.31) *and* (6.32)*, this implies that*

$$\kappa_{2n}(C_q^{\{d_1,\dots,d_k\}}) = \alpha_{2n}(\mathcal{E}_q^{\{d_1,\dots,d_k\}}).$$

Proof. Let $E_j := [\mathcal{E}_q^{\{d_1,\dots,d_k\}}]_j$ and $C_j := [C_q^{\{d_1,\dots,d_k\}}]_j$. For $\ell \in E_j$, we will denote by I_ℓ the interval $[q^{-j}\ell, q^{-j}(\ell+1)]$, so that $C_j = \bigcup_{\ell\in E_j} I_\ell$.

First we show the $\lesssim$ direction in (6.33). Let $f(x)$ be a Schwartz function Fourier supported on C_j such that $\sum_{\ell\in E_j} \|(f * \check{1}_{I_\ell})\|_{L^{2n}(\mathbb{R})}^2 = 1$. Let $f_\ell = f * \check{1}_{I_\ell}$. Note that for $\ell_1,\dots,\ell_n \in E_j$, the Fourier transform of $\prod_{j=1}^n f_{\ell_i}$ is supported in $[q^{-j}\sum_{i=1}^n \ell_i, q^{-j}(\sum_{i=1}^n \ell_i + n)]$. Therefore, by Plancherel and Hölder,

$$\begin{aligned}
\int_{\mathbb{R}} |\sum_{\ell\in E_j} f_\ell|^{2n} dx = \int_{\mathbb{R}} |\sum_{\ell_1,\dots,\ell_n\in E_j} \prod_{i=1}^n f_{\ell_i}|^2 dx &= \int_{\mathbb{R}} \sum_{\substack{|\sum_{i=1}^n \ell_i - \tilde{\ell}_i| \leq n \\ \ell_1,\dots,\ell_n\in E_j \\ \tilde{\ell}_1,\dots,\tilde{\ell}_n\in E_j}} \prod_{i=1}^n f_{\ell_i}\bar{f}_{\tilde{\ell}_i} dx \\
&\leq \sum_{\substack{|\sum_{i=1}^n \ell_i - \tilde{\ell}_i| \leq n \\ \ell_1,\dots,\ell_n\in E_j \\ \tilde{\ell}_1,\dots,\tilde{\ell}_n\in E_j}} \prod_{i=1}^n \|f_{\ell_i}\|_{L^{2n}(\mathbb{R})} \|f_{\tilde{\ell}_i}\|_{L^{2n}(\mathbb{R})}.
\end{aligned}$$

Then arguing as in the proof of Proposition 6.3.2, we have

$$\begin{aligned}
\sum_{t=-n}^{n} \sum_{\substack{\sum_{i=1}^n \ell_i - \tilde{\ell}_i = t \\ \ell_1,\dots,\ell_n\in E_j \\ \tilde{\ell}_1,\dots,\tilde{\ell}_n\in E_j}} \prod_{i=1}^n \|f_{\ell_i}\|_{2n} \|f_{\tilde{\ell}_i}\|_{2n} &\leq (2n+1) \sum_{t\in\mathbb{Z}} \left(\sum_{\substack{\ell_1,\dots,\ell_n\in E_j \\ \ell_1+\dots+\ell_n = t}} \prod_{i=1}^n \|f_{\ell_i}\|_{2n} \right)^2 \\
&\leq (2n+1)\, \mathrm{DE}_{\ell^2\to L^{2n}}([\mathcal{E}_q^{\{d_1,\dots,d_k\}}]_j)^{2n}
\end{aligned}$$

where the last inequality is by Proposition 6.3.1 and that $\sum_\ell \|f_\ell\|_{L^{2n}(\mathbb{R})}^2 = 1$.

Next we show the $\gtrsim$ direction in (6.33). Let $\phi \in C_c^\infty(\mathbb{R})$ be a smooth nonnegative function which is equal to cn on $[\frac{0.01}{n}, \frac{0.99}{n}]$ and vanishes outside $[0, 1/n]$ and where c is an absolute constant chosen so that $\|\phi\|_1 = 1$. Then observe that $\|\phi\|_2 \sim n^{1/2}$ and $\|\check{\phi}\|_\infty \leq 1$ which imply that $\|\check{\phi}\|_{2n} \lesssim n^{1/2n}$.

Define $\Phi = \phi^{*n}$, the n-fold convolution. Then $\Phi \geq 0$, Φ is supported in $[0, 1]$ and $1 = \|\Phi\|_1 \leq \|\Phi\|_2$. For $\ell \in \mathbb{Z}$, define $\phi_\ell(x) = q^j \phi(q^j x - \ell)$. Also define $\Phi_\ell(x) = q^j \Phi(q^j x - \ell)$, so that $\phi_{\ell_1} * \cdots * \phi_{\ell_n} = \Phi_{\ell_1 + \cdots + \ell_n}$ and Φ_ℓ is supported on I_ℓ.

Since E_j is finite there is a function $a : E_j \to \mathbb{R}$, which attains the supremum in (6.24). Let $a : E_j \to \mathbb{R}$ attain the maximum in (6.24). For $\ell \in E_j$, define f_ℓ by $\widehat{f_\ell} = a(\ell)\phi_\ell$. Observe that

$$\sum_{\ell_1,\ldots,\ell_n \in E_j} \widehat{f}_{\ell_1} * \cdots * \widehat{f}_{\ell_n} = \sum_{\ell_1,\ldots,\ell_n \in E_j} \left(\prod_{i=1}^n a(\ell_i) \right) \Phi_{\ell_1 + \cdots + \ell_n} = \sum_{t \in \mathbb{Z}} \left(\sum_{\sum_{i=1}^n \ell_i = t} \prod_{i=1}^n a(\ell_i) \right) \Phi_t$$

We note that the supports of Φ_t for $t \in \mathbb{Z}$ are disjoint, and that $\|\Phi_t\|_2^2 \geq q^j$, so using Plancherel we obtain

$$\left\| \sum_{\ell \in E_j} f_\ell \right\|_{2n}^{2n} = \left\| \sum_{\ell_1,\ldots,\ell_n \in E_j} \widehat{f}_{\ell_1} * \cdots * \widehat{f}_{\ell_n} \right\|_2^2 \geq \qquad q^j \sum_{t \in \mathbb{Z}} \left(\sum_{\sum_{i=1}^n \ell_i = t} \prod_{i=1}^n a(\ell_i) \right)^2 \tag{6.34}$$

$$= q^j \, \mathrm{DE}_{\ell^2 \to L^{2n}}([\mathcal{E}_q^{\{d_1,\ldots,d_k\}}]_j)^{2n} \tag{6.35}$$

Next, $\|f_\ell\|_{2n} \lesssim n^{1/(2n)} |a(\ell)| q^{j/(2n)}$, so

$$(\sum_{\ell \in E_j} \|f_\ell\|_{2n}^2)^n \lesssim n q^j (\sum_{\ell \in E_j} |a(\ell)|^2)^n = n q^j \tag{6.36}$$

By comparing (6.34) with (6.36), we see that

$$\mathrm{DE}_{\ell^2 \to L^{2n}}([\mathcal{E}_q^{\{d_1,\ldots,d_k\}}]_j) \lesssim n^{1/(2n)} K_{2n}([\mathscr{C}_q^{\{d_1,\ldots,d_k\}}]_j)$$

as desired. $\square$

Recall that given an n we say that $[\mathcal{E}_q^{\{d_1,\ldots,d_k\}}]_j$ has no carryover if $nd_k < q$. In the no carryover case, $\mathrm{DE}_{\ell^2 \to L^{2n}}([\mathcal{E}_q^{\{d_1,\ldots,d_k\}}]_j)$ has a particularly nice structure and we are able to characterize the extremizer of the associated discrete restriction estimate which will allow us the compute the decoupling constant $K_{2n}([\mathscr{C}_q^{\{d_1,\ldots,d_k\}}]_j)$.

Proposition 6.3.5. *Fix $n \geq 1$. Let $\mathcal{E}_q^{\{d_1,\ldots,d_k\}}$ be an ellipsephic set without carryover. Let* $\mathrm{Digits}_q : [\mathcal{E}_q^{\{d_1,\ldots,d_k\}}]_j \to \{0,\ldots,q-1\}^j$ *be the base q expansion of a number. Then*

$$\mathrm{DE}_{\ell^2\to L^{2n}}([\mathcal{E}_q^{\{d_1,\ldots,d_k\}}]_j) = \mathrm{DE}_{\ell^2\to L^{2n}}([\mathcal{E}_q^{\{d_1,\ldots,d_k\}}]_1)^j,$$

and there exists a function $f : \{0,\ldots,q-1\} \to \mathbb{R}_{\geq 0}$ (depending on q and $\{d_1,\ldots,d_k\}$) such that, for all $j \in \mathbb{N}$ the function

$$f_j(x) = \prod_{i=1}^{j} f((\mathrm{Digits}_q(x))_i) \tag{6.37}$$

witnesses the value of $\mathrm{DE}_{\ell^2\to L^{2n}}([\mathcal{E}_q^{\{d_1,\ldots,d_k\}}]_j)$ *where here we use the notation is that given a vector $(x_1,\ldots,x_j)$, $(x_1,\ldots,x_j)_i = x_i$ for $1 \leq i \leq j$.*

Proof. Since there is no carryover, the map $\mathrm{Digits}_q : [\mathcal{E}_q^{\{d_1,\ldots,d_k\}}]_j \to \{d_1,\ldots,d_k\}^j$ defined by $\sum_{s=0}^{j-1} a_s q^s \mapsto (a_0, a_1, \ldots, a_{j-1})$ is a Freiman isomorphism of order n. Hence by (6.25) and Proposition 6.3.3,

$$\mathrm{DE}_{\ell^2\to L^{2n}}([\mathcal{E}_q^{\{d_1,\ldots,d_k\}}]_j) = \mathrm{DE}_{\ell^2\to L^{2n}}(\{d_1,\ldots,d_k\}^j) = \mathrm{DE}_{\ell^2\to L^{2n}}(\{d_1,\ldots,d_k\})^j.$$

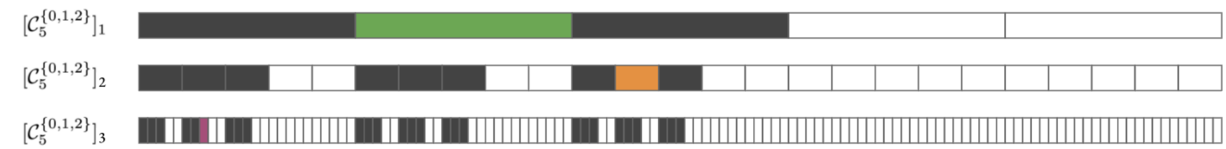
Let f be the function which witnesses the value of

$$\sup_{\substack{a:\{d_1,\ldots,d_k\}\to\mathbb{R}_{\geq 0} \\ \sum_{\ell\in\{d_1,\ldots,d_k\}} a(\ell)^2=1}} \sum_{t\in\mathbb{Z}} \Big(\sum_{\substack{\ell_1,\ldots,\ell_n\in\{d_1,\ldots,d_k\} \\ \ell_1+\cdots+\ell_n=t}} \prod_{i=1}^{n} a(\ell_i)\Big)^2.$$

Such a function exists since $\{d_1,\ldots,d_k\}$ is a finite set. Finally since Digits_q is a Freiman isomorphism of order n, following a proof similar to that of Proposition 6.3.2 shows that f_j as defined in (6.37) witnesses the value of $\mathrm{DE}_{\ell^2\to L^{2n}}([\mathcal{E}_q^{\{d_1,\ldots,d_k\}}]_j)$. □

As an immediate application of having no carryover, we now use Proposition 6.3.4 and Proposition 6.3.5 to show that the decoupling constant for a Cantor subset in $[0,1]$ not only depends on the Hausdorff dimension but also arithmetic properties of the Cantor set.

More precisely we show the following.



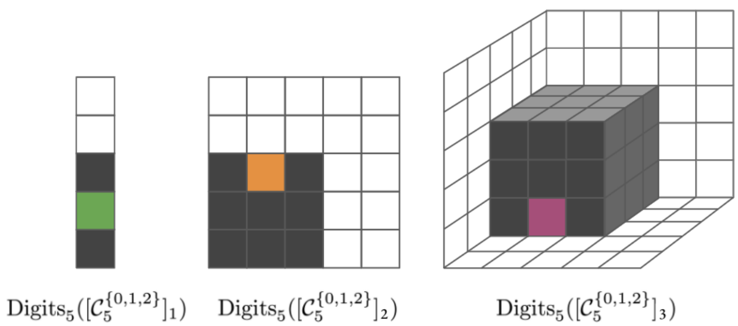


Figure 6.2: Tensor procedure described in Proposition 6.3.5. Each digit in the q-ary expansion of $\left[\mathcal{E}_q^{\{d_1,\ldots,d_k\}}\right]_t$ is mapped to its own axis in $\mathbb{Z}^t$. An element of each $\left[\mathcal{E}_5^{\{0,1,2\}}\right]_t$ in the figure has been highlighted both in the digit expansion and the original ellipsephic/Cantor set.

Proposition 6.3.6. *Fix an integer $n \geq 1$ and fix a Hausdorff dimension $d := \frac{\log r}{\log s}$ with $0 < d < 1$ and $r, s \in \mathbb{N}$. Then there exists an arithmetic Cantor set $C_q^{\{d_1,\ldots,d_k\}}$ of dimension d such that*

$$\kappa_{2n}(C_q^{\{d_1,\ldots,d_k\}}) \geq \frac{1}{2} - \frac{1}{2n}.$$

Proof. Let T be large chosen later. Let $D_T := \{1, \ldots, r^T\}$ and $q_T := s^T$. Then $C_{q_T}^{D_T}$ has Hausdorff dimension equal to $\frac{\log r^T}{\log s^T} = \frac{\log r}{\log s}$. We can also choose T so large so that $nr^T < s^T$ and so the associated ellipsephic set $\mathcal{E}_{q_T}^{D_T}$ has no carryover. Then

$$\begin{aligned}\kappa_{2n}(C_{q_T}^{D_T}) = \alpha_{2n}(\mathcal{E}_{q_T}^{D_T}) &= \limsup_{J\to\infty} \frac{\log \mathrm{DE}_{\ell^2\to L^{2n}}([\mathcal{E}_{q_T}^{D_T}]_J)}{\log(r^T)^J} \\ &= \limsup_{J\to\infty} \frac{\log \mathrm{DE}_{\ell^2\to L^{2n}}([\mathcal{E}_{q_T}^{D_T}]_1)^J}{\log(r^T)^J} = \frac{\log \mathrm{DE}_{\ell^2\to L^{2n}}([\mathcal{E}_{q_T}^{D_T}]_1)}{\log r^T}\end{aligned}$$

where the first equality is an application of Proposition 6.3.4, the second equality is by (6.31), and the third equality is because of Proposition 6.3.5. Since if we choose $a(\ell) = 1$, $\mathrm{DE}_{\ell^2\to L^{2n}}(\{1, \ldots, r^T\}) \geq (r^T)^{\frac{1}{2}-\frac{1}{2n}}$, the claim now follows. □

Note that $\kappa_{2n}(C_q^{\{d_1,\ldots,d_k\}}) \leq \frac{1}{2} - \frac{1}{2n}$. To see this, one can either interpolate the estimates $D_2(\delta(i)) = 1$ and $D_\infty(\delta(i)) \leq N(i)^{1/2}$ (see [71, Exercise 10(iv)] for an interpolation theorem) or alternatively one can follow the same proof as in [37, Proposition 1.12] for a direct proof. Thus Proposition 6.3.6 says that even though our Cantor set has small Hausdorff dimension, it can still have a decoupling constant that is as large as possible.

We had particularly good structure when $\mathcal{E}_q^{\{d_1,\ldots,d_k\}}$ did not have carryover, however the case when one has carryover is much harder. In the general case, from a computational standpoint, the following lemma tells us that we can obtain a good approximation on $\alpha_{2n}(\mathcal{E}_q^{\{d_1,\ldots,d_k\}})$ by estimating $\mathrm{DE}_{\ell^2\to L^{2n}}$ on the finite sets $[\mathcal{E}_q^{\{d_1,\ldots,d_k\}}]_t$.

Proposition 6.3.7. *Let $\mathcal{E}_q^{\{d_1,\ldots,d_k\}}$ be an ellipsephic set potentially with carryover. Let $t > \log_q n$. Then $\alpha_{2n}(\mathcal{E}_q^{\{d_1,\ldots,d_k\}})$ can be approximated by computing $\mathrm{DE}_{\ell^2\to L^{2n}}([\mathcal{E}_q^{\{d_1,\ldots,d_k\}}]_t)$ with the following bound:*

$$|\alpha_{2n}(\mathcal{E}_q^{\{d_1,\ldots,d_k\}}) - \frac{\log \mathrm{DE}_{\ell^2\to L^{2n}}([\mathcal{E}_q^{\{d_1,\ldots,d_k\}}]_t)}{\log k^t}| \leq \frac{\log(2n+1)}{2nt\log k}. \tag{6.38}$$

and therefore

$$\alpha_{2n}(\mathcal{E}_q^{\{d_1,\ldots,d_k\}}) = \lim_{t\to\infty} \frac{\log \mathrm{DE}_{\ell^2\to L^{2n}}([\mathcal{E}_q^{\{d_1,\ldots,d_k\}}]_t)}{\log k^t}.$$

Proof. Choose $t \in \mathbb{N}$ such that $q^t > n$ and note that

$$\left[\mathcal{E}_{q^t}^{\left[\mathcal{E}_q^{\{d_1,\ldots,d_k\}}\right]_t}\right]_j = \left[\mathcal{E}_q^{\{d_1,\ldots,d_k\}}\right]_{jt}.$$

Consider the bijection

$$\begin{aligned} \mathrm{Digit}_{q^t} : \left[\mathcal{E}_q^{\{d_1,\ldots,d_k\}}\right]_{jt} &\longrightarrow \left[\mathcal{E}_q^{\{d_1,\ldots,d_k\}}\right]_t^j, \\ \sum_{s=0}^{j-1} a_s q^{st} &\longmapsto (a_0, a_1, \ldots, a_{j-1}) \end{aligned} \tag{6.39}$$

For this map, the set D in (6.26) satisfies

$$D \subset \{(q^t a_1, q^t a_2 - a_1, \ldots, q^t a_{j-1} - a_{j-2}, -a_{j-1}) : a_1, \ldots, a_{j-1} \in \{-n+1, \ldots, n-1\}\}. \tag{6.40}$$

To see this, note that the inverse of Digit_{q^t} extends to a group homomorphism $\mathbb{Z}^j \to \mathbb{Z}$, so D is contained in the kernel of this group homomorphism. Furthermore, the set $\left[\mathcal{E}_q^{\{d_1,\ldots,d_k\}}\right]_t$ is bounded above by $q^t - 1$. These two observations together imply

$$D \subset \{(b_0, \ldots, b_{j-1}) \in \mathbb{Z}^j : \sum_{s=0}^{j-1} q^{st} b_s = 0\} \cap [-n(q^t-1), n(q^t-1)]^j.$$

To show (6.40), suppose $(b_0, \ldots, b_{j-1}) \in D$. Then $|b_s| \leq n(q^t - 1)$ and

$$\sum_{s=0}^{j-1} q^{st} b_s = 0. \tag{6.41}$$

Taking (6.41) modulo q^t gives $b_0 \equiv 0 \pmod{q^t}$, hence, $b_0 = q^t a_1$ for some $a_1 \in \mathbb{Z}$. Also $|b_0| \leq n(q^t-1)$ implies $|a_1| \leq n-1$. Then taking (6.41) modulo q^{2t} gives $q^t a_1 + q^t b_1 \equiv 0 \pmod{q^{2t}}$, so $b_1 = -a_1 + q^t a_2$ for some $|a_2| \leq n-1$. By repeating this, we get $b_s = -a_s + q^t a_{s+1}$ for $s = 1, \ldots, j-2$. Finally, (6.41) gives us $b_{j-1} = -a_{j-1}$. (We can think of the numbers $(a_1, \ldots, a_{j-1})$ as "carryover digits.")

Equation (6.40) implies $|D| \leq (2n+1)^j$. By Proposition 6.3.2 and Proposition 6.3.3, this tells us that

$$\mathrm{DE}_{\ell^2 \to L^{2n}}([\mathcal{E}_q^{\{d_1,\ldots,d_k\}}]_{jt}) \leq (2n+1)^{\frac{j}{2n}} \, \mathrm{DE}_{\ell^2 \to L^{2n}}([\mathcal{E}_q^{\{d_1,\ldots,d_k\}}]_t)^j.$$

Also, note that the inverse of the map (6.39) is a Freiman homomorphism of order n, so by (6.25)

$$\mathrm{DE}_{\ell^2 \to L^{2n}}([\mathcal{E}_q^{\{d_1,\ldots,d_k\}}]_t)^j \leq \mathrm{DE}_{\ell^2 \to L^{2n}}([\mathcal{E}_q^{\{d_1,\ldots,d_k\}}]_{jt}).$$

Applying (6.31) to the above two inequalities then proves (6.38).

□

Remark 6.3.8. *Note that the right hand side of* (6.38) *is nondecreasing in t (when n and k are kept constant), so increasing t gives strictly better and better approximations to $\alpha_{2n}(\mathcal{E}_q^{\{d_1,\ldots,d_k\}})$.*

6.3.3 Examples

The above results in this section allow for explicit computations (in relatively simple cases) and numerical approximations (in the remaining, more complex cases) of the l^2L^{2n} decoupling constant associated to an arithmetic Cantor set.

To demonstrate some examples, we consider the l^2L^4 decoupling constant for the following arithmetic Cantor sets. To study K_4, we first use Proposition 6.3.4 to reduce to studying $\mathrm{DE}_{\ell^2\to L^4}$. Then we assume q is sufficiently large so that we are in the no carryover case which allows us to use Proposition 6.3.5 and Proposition 6.3.1 which reduces to an optimization problem.

Note that if we take $a(\ell) = 1$ in the definition of $\mathrm{DE}_{\ell^2\to L^4}$, this amounts to studying the additive energy. In the case of an ellipsephic set, one can apply for example, [30, Lemma 3.10]. However this would only give a lower bound on $\mathrm{DE}_{\ell^2\to L^4}$ and the function defined by $a(\ell) = c$ for some c is not always the optimizer of the discrete restriction problem for ellipsephic sets (see for example, Example 6.3.11 below).

Example 6.3.9 (The $(0,1) \pmod q$ arithmetic Cantor set)**.** *Let $k = 2$ and $\{d_1, d_2\} = \{0, 1\}$. At each level j, this Cantor set has 2^j many intervals. By Proposition 6.3.1,*

$$\mathrm{DE}_{\ell^2\to L^4}([\mathcal{E}_q^{\{0,1\}}]_1)^4 = \sup\left\{(a_0^2)^2 + (a_0a_1 + a_1a_0)^2 + (a_1^2)^2 \;\middle|\; a_0^2 + a_1^2 = 1\right\} = \frac{3}{2}$$

It is easy to see that the maximum is attained when $a_0 = a_1 = 2^{-1/2}$. If $q > 2$, then there is no carryover, so Proposition 6.3.5 implies that

$$K_4([\mathscr{C}_q^{\{0,1\}}]_j)^4 \sim \mathrm{DE}_{\ell^2\to L^4}([\mathcal{E}_q^{\{0,1\}}]_j)^4 = (3/2)^j = (2^j)^{\log_2(3/2)}.$$

This should be compared to the trivial bound that $K_4([\mathscr{C}_q^{\{0,1\}}]_j)^4 \leq 2^j$.

Example 6.3.10 (The $(0,2) \pmod 3$ arithmetic Cantor set)**.** *Let $k = 2$ and $\{d_1, d_2\} = \{0, 2\}$. Then $[C_q^{\{0,2\}}]_j$ is the jth level of the middle thirds Cantor set. Since we are studying the l^2L^4 decoupling constant $K_{2\cdot 2}([\mathscr{C}_q^{\{0,2\}}]_j)$, $n = 2$ and so the associated ellipsephic set*

$[\mathcal{E}_3^{\{0,2\}}]_j$ *has carryover. However, note for all levels* j*, the map* $\phi : [\mathcal{E}_3^{\{0,2\}}]_j \to [\mathcal{E}_3^{\{0,1\}}]_j$ *given by* $x \mapsto x/2$ *is a Freiman isomorphism of order* 2 *and the latter set does not have carryover. Therefore from Proposition 6.3.4,*

$$K_4([\mathscr{C}_3^{\{0,2\}}]_j)^4 \sim \mathrm{DE}_{\ell^2\to L^4}([\mathcal{E}_3^{\{0,2\}}]_j)^4 = \mathrm{DE}_{\ell^2\to L^4}([\mathcal{E}_3^{\{0,1\}}]_1)^4 = (3/2)^j$$

where the first equality is because of (6.25) *and the second equality is because of Example 6.3.9. Therefore we have computed precisely the* l^2L^4 *decoupling constant for the middle thirds Cantor set.*

Example 6.3.11 (The $(0,1,2) \pmod q$ arithmetic Cantor set)**.** *Let* $k = 3$ *and* $\{d_1, d_2, d_3\} = \{0, 1, 2\}$*. At each level* j*, this Cantor set has* 3^j *many intervals. By Proposition 6.3.1,*

$$\mathrm{DE}_{\ell^2\to L^4}([\mathcal{E}_q^{\{0,1,2\}}]_1)^4$$
$$= \sup\left\{(a_0^2)^2 + (2a_0a_1)^2 + (2a_0a_2 + a_1^2)^2 + (2a_1a_2)^2 + (a_2^2)^2 \;\middle|\; a_0^2 + a_1^2 + a_2^2 = 1\right\} = \frac{15}{7}$$

One can check that $a_0 = a_2 = (2/7)^{1/2}, a_1 = (3/7)^{1/2}$ *attains the maximum.*

If $q > 4$*, then there is no carryover, so Proposition 6.3.5 implies that*

$$K_4([\mathscr{C}_q^{\{0,1,2\}}]_j)^4 \sim \mathrm{DE}_{\ell^2\to L^4}([\mathcal{E}_q^{\{0,1,2\}}]_j)^4 = (15/7)^j = (3^j)^{\log_3(15/7)}.$$

This once again should be compared to the trivial bound that $K_4([\mathscr{C}_q^{\{0,1,2\}}]_j)^4 \le 3^j$*.*

Example 6.3.12 (The $(0,1,3) \pmod q$ arithmetic Cantor set)**.** *Let* $k = 3$ *and* $\{d_1, d_2, d_3\} = \{0, 1, 3\}$*. At each level* j*, this Cantor set has* 3^j *many intervals. By Proposition 6.3.1,*

$$\mathrm{DE}_{\ell^2\to L^4}([\mathcal{E}_q^{\{0,1,3\}}]_1)^4$$
$$= \sup\left\{(a_0^2)^2 + (2a_0a_1)^2 + (a_1^2)^2 + (2a_0a_3)^2 + (2a_1a_3)^2 + (a_3^2)^2 \;\middle|\; a_0^2 + a_1^2 + a_3^2 = 1\right\} = \frac{5}{3}$$

One can check that $a_0 = a_1 = a_3 = 3^{-1/2}$ *attains the maximum.*

If $q > 6$*, then there is no carryover, so Proposition 6.3.5 implies that*

$$K_4([\mathscr{C}_q^{\{0,1,3\}}]_j)^4 \sim \mathrm{DE}_{\ell^2\to L^4}([\mathcal{E}_q^{\{0,1,3\}}]_j)^4 = (5/3)^j = (3^j)^{\log_3(5/3)}.$$

As in the previous example, we trivially have that $K_4([\mathscr{C}_q^{\{0,1,3\}}]_j)^4 \le 3^j$*.*

Example 6.3.13 (Cantor sets generated by squares). *Let $q > 2$, $S := \{n^2, n \in \mathbb{N}\}$ the set of squares, and $S_q = S \cap [0, q)$ the squares less than q. Then:*

$$\lim_{q \to \infty} \alpha_4(\mathcal{E}_q^{S_q}) = 0 \tag{6.42}$$

By Theorem 6.1.1 and the definition of α in (6.31)*, this implies [8, Corollary 1.4] (note that in [8], q is restricted to be a prime number, while here, this restriction is not needed).*

Equation (6.42) *will follow from Proposition 6.3.7 and a number-theoretic estimate about sums of elements in S. Using* (6.38) *with $t = 1$ (we can do so since $q > 2$) and using that $\#[\mathcal{E}_q^{S_q}]_1 = \lfloor \sqrt{q} \rfloor + 1$, one obtains*

$$|\alpha_4(\mathcal{E}_q^{S_q}) - \frac{\log \mathrm{DE}_{\ell^2 \to L^4}([\mathcal{E}_q^{S_q}]_1)}{\log(\lfloor \sqrt{q} \rfloor + 1)}| \lesssim \frac{1}{\log q}$$

where the implied constant is absolute. Thus (6.42) *will follow from*

$$\lim_{q \to \infty} \frac{\log \mathrm{DE}_{\ell^2 \to L^4}([\mathcal{E}_q^{S_q}]_1)}{\log \sqrt{q}} = 0$$

Since counting diagonal solutions shows that $\mathrm{DE}_{\ell^2 \to L^4} \gtrsim 1$, it suffices to show that

$$\mathrm{DE}_{\ell^2 \to L^4}([\mathcal{E}_q^{S_q}]_1) \lesssim q^{o(1)}. \tag{6.43}$$

We in fact show that the left hand side above is $\lesssim \exp(O(\frac{\log q}{\log \log q}))$ where the implied constant is absolute. Indeed, the divisor bound for $\mathbb{Z}[i]$ implies that

$$\max_{0 \leq j \leq 2q} |\{n_1, n_2 \in S, n_1 + n_2 = j\}| \leq \exp(O(\frac{\log q}{\log \log q}))$$

which leads to

$$\sum_{t \in \mathbb{Z}} | \sum_{\ell_1, \ell_2 \in S_q : \ell_1 + \ell_2 = t} a(\ell_1) a(\ell_2)|^2 \lesssim \exp(O(\frac{\log q}{\log \log q})) \sum_{t \in \mathbb{Z}} \sum_{\ell_1, \ell_2 \in S_q : \ell_1 + \ell_2 = t} |a(\ell_1)|^2 |a(\ell_2)|^2$$
$$= \exp(O(\frac{\log q}{\log \log q}))(\sum_{\ell \in S_q} |a(\ell)|^2)^2$$

which proves (6.43)*. In fact the above proof gives quantitative control on the decoupling exponent and shows*

$$|\alpha_4(\mathcal{E}_q^{S_q})| \lesssim \frac{1}{\log \log q}$$

where the implied constant is absolute.

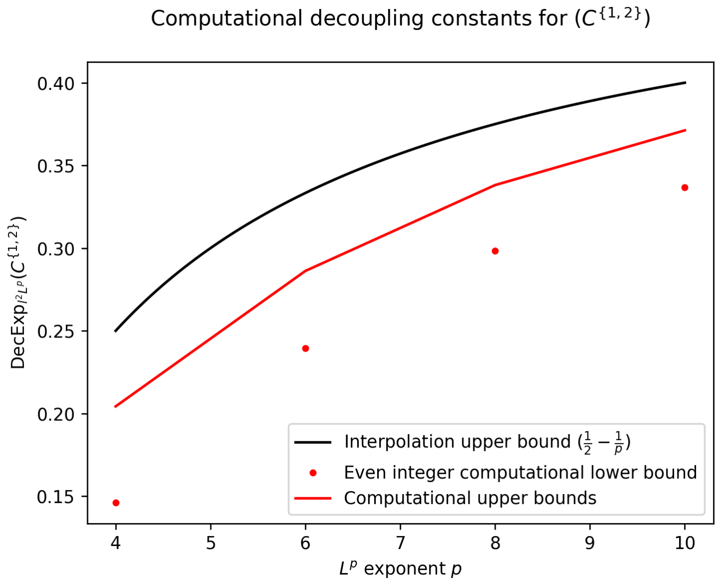


Figure 6.3: Numerical estimation of $\alpha_{2n}(\mathcal{E}_3^{\{1,2\}})$. The optimization has been performed using gradient descent using Torch. At stopping time the l^2 gradients of the optimization where $\leq 10^{-8}$. There is no guarantee, however, that the near-local-optimizers are in fact global optimizers of the problem at hand. The upper bounds on the figure (red line) are the upper bounds from Proposition 6.3.7 assuming the optimization problem resulted in a global optimizer.

6.3.4 Computational results

Proposition 6.3.7 hints of a way of estimating the decoupling exponents of Cantor sets (or at least obtaining an upper bound) by computing the value of $\frac{\log \mathrm{DE}_{\ell^2 \to L^{2n}}([\mathcal{E}_q^{\{d_1,\ldots,d_k\}}]_t)}{\log k^t}$ for finite values of t. Since $[\mathcal{E}_q^{\{d_1,\ldots,d_k\}}]_t$ contains finitely many points, one may attempt to numerically find the extremizers to the decoupling inequality, in other words, to compute:

$$\mathrm{DE}_{\ell^2\to L^{2n}}([\mathcal{E}_q^{\{d_1,\dots,d_k\}}]_t)^{2n} = \underset{\substack{f\in l^2([\mathcal{E}_q^{\{d_1,\dots,d_k\}}]_t)\\ \|f\|_{l^2}=1}}{\arg\max} \sum_{\substack{a_1,\dots a_n\in[\mathcal{E}_q^{\{d_1,\dots,d_k\}}]_t\\ b_1,\dots b_n\in[\mathcal{E}_q^{\{d_1,\dots,d_k\}}]_t\\ a_1+\dots+a_n=b_1+\dots+b_n}} f(a_1)\dots f(a_n)\cdot\bar{f}(b_1)\dots\bar{f}(b_n)$$

$$= \underset{\substack{f\in l^2([\mathcal{E}_q^{\{d_1,\dots,d_k\}}]_t)\\ \|f\|_{l^2}=1}}{\arg\max} \|\underbrace{f*f\dots*f}_{n\text{ times}}\|^2_{l^2(\mathbb{Z})} \tag{6.44}$$

or, as an unconstrained optimization problem,

$$\mathrm{DE}_{\ell^2\to L^{2n}}([\mathcal{E}_q^{\{d_1,\dots,d_k\}}]_t) = \underset{\mathrm{supp}\ f\subseteq[\mathcal{E}_q^{\{d_1,\dots,d_k\}}]_t}{\arg\max} \frac{\|f*f\dots*f\|^{1/n}_{l^2(\mathbb{Z})}}{\|f\|^2_{l^2(\mathbb{Z})}} \tag{6.45}$$

We performed the numerical optimization problem in (6.45) for the $(0,2) \mod 3$ Cantor set and $n = 1,2,3,4$ using gradient descent. The results can be seen in Figure 6.3. While there are no a priori guarantees that the near-local-optimizers obtained from gradient descent are in fact global optimizers of the problem at hand, this method was tested on the previous examples in Section 6.3.3, and converged to the known decoupling exponent.

6.3.4.1 A conjectured fixed point method

Studying equation (6.44), using Lagrange multipliers one may extract information about the solution, more precisely that, at extremizers (which must exist because $l^2([\mathcal{E}_q^{\{d_1,\dots,d_k\}}]_t)$ is a finite-dimensional space) the following equality holds:

$$f = \lambda\cdot\chi_{[\mathcal{E}_q^{\{d_1,\dots,d_k\}}]_t}\cdot\nabla\|\underbrace{f*f\dots*f}_{n\text{ times}}\|^2_{l^2(\mathbb{Z})}$$

where ∇ denotes the gradient with respect to f in $l^2([\mathcal{E}_q^{\{d_1,\dots,d_k\}}]_t))$. Let

$$\Phi(f) := \lambda\cdot\chi_{[\mathcal{E}_q^{\{d_1,\dots,d_k\}}]_t}\cdot\frac{\partial}{\partial f}\|\underbrace{f*f\dots*f}_{n\text{ times}}\|^2_{l^2(\mathbb{Z})}.$$

The functional Φ sends nonnegative functions to nonnegative functions, and by Cauchy-Schwarz we know there exists an extremizer with nonnegative components. This suggests the following numerical method to compute an extremizer:

Require: TOL > 0

Require: $f : \chi_{[\mathcal{E}_q^{\{d_1,\dots,d_k\}}]_t} \to \mathbb{R}^+$

$n \leftarrow 0$

do

$f_{n+1} \leftarrow \frac{\Phi(f_n)}{\|\Phi(f_n)\|_2}$

$n \leftarrow n+1$

while $\|f_n - f_{n-1}\| >$TOL

Convergence of this algorithm to an unique maximum would follow if $f \mapsto \frac{\Phi(f)}{\|\Phi(f)\|}$ was contractive in some norm. Numerical experiments seem to indicate convergence of the algorithm in all situations that were tested at a much faster rate than the gradient descent methods.

6.3.4.2 Code

A commented version of the code can be found at `https://github.com/jaumededios/Decoupling_Cantor`.

CHAPTER 7

Additive energies for product sets

This section reproduces the paper Additive energies on discrete cubes, [25] *by Jaume de Dios Pont, Rachel Greenfeld, Paata Ivanisvili and José Madrid, with minor changes.*

Abstract: *(From the published version)* We prove that for $d \geq 0$ and $k \geq 2$, for any subset A of a discrete cube $\{0,1\}^d$, the $k-$higher energy of A (i.e., the number of $2k-$tuples $(a_1, a_2, \ldots, a_{2k})$ in A^{2k} with $a_1 - a_2 = a_3 - a_4 = \cdots = a_{2k-1} - a_{2k}$) is at most $|A|^{\log_2(2^k+2)}$, and $\log_2(2^k + 2)$ is the best possible exponent. We also show that if $d \geq 0$ and $2 \leq k \leq 10$, for any subset A of a discrete cube $\{0,1\}^d$, the $k-$additive energy of A (i.e., the number of $2k-$tuples $(a_1, a_2, \ldots, a_{2k})$ in A^{2k} with $a_1 + a_2 + \cdots + a_k = a_{k+1} + a_{k+2} + \cdots + a_{2k}$) is at most $|A|^{\log_2 \binom{2k}{k}}$, and $\log_2 \binom{2k}{k}$ is the best possible exponent. We discuss the analogous problems for the sets $\{0, 1, \ldots, n\}^d$ for $n \geq 2$.

7.1 Introduction

The additive energy $E(A)$ of a finite subset A of an additive group G is defined as the number of quadruples $(a_1, a_2, a_3, a_4) \in A \times A \times A \times A$ such that $a_1 + a_2 = a_3 + a_4$ (see [72]). Observe that for any triple (a_1, a_2, a_3) there is at most one a_4 such that $a_1 + a_2 = a_3 + a_4$, so we have the trivial upper bound $E(A) \leq |A|^3$ (here $|A|$ denotes the cardinality of A). This bound is attained, for example, when A is itself a finite group. Considering the diagonal solutions $a_1 = a_3$ and $a_2 = a_4$ we also observe the trivial lower bound $E(A) \geq |A|^2$.

7.1.1 Higher energies

We define the $k-$higher energy of a set $A \subseteq \{0,1\}^d \subset \mathbb{Z}^d$ by

$$\widetilde{E}_k(A) := |\{(a_1, a_2, \ldots, a_{2k-1}, a_{2k}) \in A^{2k} : a_1 - a_2 = a_3 - a_4 = \cdots = a_{2k-1} - a_{2k}\}|.$$

This has been studied by many authors, see [63], [62]. In this case we have the trivial bounds $|A|^k \leq \widetilde{E}_k(A) \leq |A|^{k+1}$.

Theorem 7.1.1. *Let $d \geq 0$, $k \geq 2$, and let $A \subset \{0,1\}^d$. Then $\widetilde{E}_k(A) \leq |A|^{q_k}$, where $q_k := \log_2(2^k + 2)$. Furthermore, the exponent q_k cannot be replaced by any smaller quantity.*

Remark 7.1.2. *This Theorem extends a result obtained by Kane–Tao [45, Theorem 7] for $k = 2$.*

The second claim in our Theorem 7.1.1 follows considering the case $A = \{0,1\}^d$, in this case we have $|A| = |\{0,1\}|^d = 2^d$ and $\widetilde{E}_k(\{0,1\}^d) = (2^k + 2)^d$.

7.1.2 k-additive energies

We discuss another generalization of Kane–Tao result [45, Theorem 7]. We define the k-additive energy $E_k(A)$ of a subset A of an additive group G as the number of $2k-$tuples $(a_1, a_2, \ldots, a_{2k})$ in A^{2k} with $a_1 + a_2 + \cdots + a_k = a_{k+1} + a_{k+2} + \cdots + a_{2k}$. In this case the trivial bounds are $|A|^k \leq E_k(A) \leq |A|^{2k-1}$, and we have the following refinement in the cube $\{0,1\}^d$.

Theorem 7.1.3. *Let $d \geq 0$, $2 \leq k \leq 10$, and let $A \subset \{0,1\}^d$. Then $E_k(A) \leq |A|^{p_k}$, where $p_k := \log_2 \binom{2k}{k}$. Furthermore, the exponent p_k cannot be replaced by any smaller quantity.*

Remark 7.1.4. *Theorem 7.1.3 also extends a result obtained by Kane–Tao ([45, Theorem 7]).*

From the well-known bounds for the central binomial coefficient $\frac{4^k}{2\sqrt{\pi k}} \leq \binom{2k}{k} \leq \frac{4^k}{\sqrt{\pi k}}$, one recovers

$$p_k < 2k - 1. \tag{7.1}$$

As previously, the second claim in our Theorem 7.1.3 follows considering the case $A = \{0,1\}^d$, since in this case we have $|A| = |\{0,1\}|^d = 2^d$ and $E_k(A) = \left[\sum_{i=0}^{k} \binom{k}{i}^2\right]^d = \binom{2k}{k}^d$. We prove this theorem by induction on d together with the following subtle inequality for Legendre polynomials.

Lemma 7.1.5. *Let* $2 \leq k \leq 10$ *and* $p_k = \log_2 \binom{2k}{k}$. *If* $a, b \geq 0$, *then*

$$\sum_{j=0}^{k} \binom{k}{j}^2 a^{p_k \frac{k-j}{k}} b^{p_k \frac{j}{k}} \leq (a+b)^{p_k}. \tag{7.2}$$

The polynomials $Q_k(t)$, $k \geq 0$, defined by

$$Q_k(t) = \frac{1}{2^k k!} \frac{\mathrm{d}^k}{\mathrm{d}t^k} (t^2 - 1)^k = \frac{1}{2^k} \sum_{j=0}^{k} \binom{k}{j}^2 (t-1)^{k-j} (t+1)^j$$

are called Legendre polynomials. They are orthogonal with respect to Lebesgue measure on the interval $[-1, 1]$, each $Q_k(t)$ has degree k, and they satisfy normalization constraint $Q_k(1) = 1$. Dividing both sides of (7.2) by a^{p_k} (without loss of generality assume $a \neq 0$), then (7.2) takes the form $(y-1)^k Q_k\left(\frac{y+1}{y-1}\right) \leq (1 + y^{k/p_k})^{p_k}$ with $y = (b/a)^{p_k/k} \geq 0$. If we let $t := \frac{y+1}{y-1}$ (without loss of generality assume $y \geq 1$), then (7.2) is the same as

$$Q_k(t) \leq \left(\left(\frac{t-1}{2}\right)^{\frac{k}{p_k}} + \left(\frac{t+1}{2}\right)^{\frac{k}{p_k}} \right)^{p_k} \quad \text{for all} \quad t \geq 1.$$

This explains the reason we call Lemma 7.1.5 the inequality for Legendre polynomials.

7.1.3 More general discrete cubes

Let $d \geq 0$. Let us consider additive energies of subsets of general discrete cubes[1] $\{0, 1, \ldots, n\}^d$. Let t_n be the smallest number such that

$$E_2(A) \leq |A|^{t_n}$$

for all $A \subseteq \{0, 1, \ldots, n\}^d$. We have seen that in both Theorem 7.1.1 and Theorem 7.1.3 we have $q_k = \frac{\log \widetilde{E}_k(\{0,1\}^d)}{\log |\{0,1\}|^d}$ and $p_k = \frac{\log E_k(\{0,1\}^d)}{\log |\{0,1\}|^d}$. Thus, one could a-priori expect a similar phenomenon for the additive energy of $\{0, 1, \ldots, n\}^d$. However, it turns out that this is not the case in general, not even for the discrete cube $\{0, 1, 2\}^d$.

Proposition 7.1.6. *The following inequality holds*

$$t_2 > \frac{\log E_2(\{0, 1, 2\}^d)}{\log |\{0, 1, 2\}^d|}.$$

Although finding the precise values of the optimal powers t_n for general discrete cubes $\{0, 1, \ldots, n\}^d$ seems to be a difficult problem, we obtain some bounds describing the asymptotic behavior of t_n as n goes to infinity.

Proposition 7.1.7. *If $n = 2m - 1$, then*

$$3 \geq t_n \geq \log_{2m} \left(\frac{16m^3 + 2m}{3} \right) > 3 - \frac{\log(3/2)}{\log(2m)}.$$

If $n = 2m$, then

$$3 \geq t_n \geq \log_{2m} \left(\frac{16m^3 + 24m^2 + 14m + 3}{3} \right) > 3 - \frac{\log(3/2)}{\log(2m)}.$$

7.2 Proof of Theorem 7.1.1

The proof of Theorem 7.1.1 proceeds via induction on d. Observe that the result is trivial for $d = 0$. Assume now that $d \geq 1$ and that the result has been established for $d - 1$. Any

[1] A related problem about the lower bound for the size of sumsets of subsets of the general discrete cube was studied, e.g., in [13, Theorem 5].

set $A \subseteq \{0,1\}^d$ can be written as

$$A = (A_0 \times \{0\}) \uplus (A_1 \times \{1\})$$

for some $A_0, A_1 \subseteq \{0,1\}^{d-1}$, where $\uplus$ means disjoint union. Then we have

$$\begin{aligned}
\widetilde{E}_k(A) &= |\{(a_1, a_2, \ldots, a_{2k}) \in (A_0 \times A_1)^k : a_1 - a_2 = a_3 - a_4 = \cdots = a_{2k-1} - a_{2k}\}| \\
&\quad + |\{(a_1, a_2, \ldots, a_{2k}) \in (A_1 \times A_0)^k : a_1 - a_2 = a_3 - a_4 = \cdots = a_{2k-1} - a_{2k}\}| \\
&\quad + \sum_{i=0}^{k} \binom{k}{i} |\{(a_1, a_2, \ldots, a_{2k}) \in (A_0^2)^i \times (A_1^2)^{k-i} \\
&\qquad\qquad : a_1 - a_2 = a_3 - a_4 = \cdots = a_{2k-1} - a_{2k}\}| \\
&=: C_1 + C_2 + \widetilde{E}_k(A_0) + \widetilde{E}_k(A_1) + \sum_{i=1}^{k-1} \binom{k}{i} C_{i,k}. \qquad (7.3)
\end{aligned}$$

The next proposition plays a fundamental role in our proof.

Proposition 7.2.1. *For all* $1 \leq i \leq k-1$ *we have that*

$$C_{i,k} \leq |A_0|^{\frac{i}{k} q_k} |A_1|^{\frac{k-i}{k} q_k}.$$

Moreover, we have that

$$C_1 \leq |A_0|^{\frac{q_k}{2}} |A_1|^{\frac{q_k}{2}} \text{ and } C_2 \leq |A_0|^{\frac{q_k}{2}} |A_1|^{\frac{q_k}{2}}.$$

Proof of Proposition 7.2.1. We observe that

$$\widetilde{E}_k(A) := \sum_{x \in \mathbb{Z}^d} (\chi_A \star \chi_A)^k(x),$$

where χ_A denotes the characteristic function of the set A, and $f \star g$ denotes the correlation of the functions f and g defined by $f \star g(x) := \sum_{y \in \mathbb{Z}^d} f(y) g(x+y)$ [62, Equation 7]. Moreover, by Hölder's inequality we have

$$\begin{aligned}
C_{i,k} &= \sum_{x\in\mathbb{Z}^d}(\chi_{A_0}\star\chi_{A_0})^i(x)(\chi_{A_1}\star\chi_{A_1})^{k-i}(x)\\
&\le \left(\sum_{x\in\mathbb{Z}^d}(\chi_{A_0}\star\chi_{A_0})^k(x)\right)^{\frac{i}{k}}\left(\sum_{x\in\mathbb{Z}^d}(\chi_{A_1}\star\chi_{A_1})^k(x)\right)^{\frac{k-i}{k}}\\
&= \widetilde{E}_k^{\frac{i}{k}}(A_0)\widetilde{E}_k^{\frac{k-i}{k}}(A_1)\\
&\le |A_0|^{\frac{q_k i}{k}}|A_1|^{\frac{q_k(k-i)}{k}}.
\end{aligned}$$

The first identity follows from the facts that $\chi_{A_0}\star\chi_{A_0}(x)$ counts the number of pairs $(y,z)\in A_0^2$ such that $z-y=x$, and $\chi_{A_1}\star\chi_{A_1}(x)$ counts the number of pairs $(y,z)\in A_1^2$ such that $z-y=x$. We define

$$f\bullet g := \sum_{\substack{a_1,a_2,\ldots,a_k\in\{0,1\}^d\\ b_1,b_2,\ldots,b_k\in\{0,1\}^d\\ a_1-b_1=a_2-b_2=\cdots=a_k-b_k}} f(a_1)f(a_2)\ldots f(a_k)g(b_1)g(b_2)\ldots g(b_k).$$

Then

$$\begin{aligned}
&f\bullet g =\\
&\sum_{c_2,c_3,\ldots,c_k\in\{-1,0,1\}^d}\left(\sum_{\substack{a_1\in\{0,1\}^d\\ a_1+c_i\in\{0,1\}^d}} f(a_1)f(a_1+c_2)f(a_1+c_3)\ldots f(a_1+c_k)\right)\\
&\quad\times\left(\sum_{\substack{b_1\in\{0,1\}^d\\ b_1+c_i\in\{0,1\}^d}} g(b_1)g(b_1+c_2)g(b_1+c_3)\ldots g(b_1+c_k)\right).
\end{aligned}$$

Therefore, by the Cauchy-Schwarz inequality we obtain

$$\begin{aligned}
C_1 = \chi_{A_0}\bullet\chi_{A_1} &\le (\chi_{A_0}\bullet\chi_{A_0})^{1/2}(\chi_{A_1}\bullet\chi_{A_1})^{1/2}\\
&= \widetilde{E}_k^{1/2}(A_0)\widetilde{E}_k^{1/2}(A_1)\le |A_0|^{\frac{q_k}{2}}|A_1|^{\frac{q_k}{2}}.
\end{aligned}$$

Similarly $C_2\le |A_0|^{\frac{q_k}{2}}|A_1|^{\frac{q_k}{2}}$. □

Then, from (7.3), using Proposition 7.2.1 we obtain

$$\begin{aligned}\widetilde{E}_k(A) &= C_1 + C_2 + \widetilde{E}_k(A_0) + \widetilde{E}_k(A_1) + \sum_{i=1}^{k-1} \binom{k}{i} C_{i,k} \\ &\le 2|A_0|^{\frac{q_k}{2}} |A_1|^{\frac{q_k}{2}} + \sum_{i=0}^{k} \binom{k}{i} |A_0|^{\frac{i}{k} q_k} |A_1|^{\frac{k-i}{k} q_k} \\ &= 2|A_0|^{\frac{q_k}{2}} |A_1|^{\frac{q_k}{2}} + (|A_0|^{\frac{q_k}{k}} + |A_1|^{\frac{q_k}{k}})^k.\end{aligned}$$

Thus, to complete the inductive argument, it is enough to prove that for $x = |A_0|$ and $y = |A_1|$ one has

$$2x^{\frac{q_k}{2}} y^{\frac{q_k}{2}} + (x^{\frac{q_k}{k}} + y^{\frac{q_k}{k}})^k \le (x+y)^{q_k}. \tag{7.4}$$

Lemma 7.2.2. *For all $a \in [0,1]$ we have*

$$(a^{\frac{q_k}{k}} + (1-a)^{\frac{q_k}{k}})^k + 2a^{\frac{q_k}{2}} (1-a)^{\frac{q_k}{2}} \le 1. \tag{7.5}$$

Observe that (7.4) follows from (7.5) by taking $a = \frac{x}{x+y}$. A key ingredient in the proof of Lemma 7.2.2 is the following result established by Carlen, Frank, Ivanisvili and Lieb [14, Proposition 3.1].

Proposition 7.2.3. *For all $a \in [0,1]$ and $p \in (-\infty, 0] \cup [1,2]$*

$$(a^p + (1-a)^p) \left(1 + \left(\frac{2a^{\frac{p}{2}}(1-a)^{\frac{p}{2}}}{a^p + (1-a)^p} \right)^{\frac{2}{p}} \right)^{p-1} \le 1. \tag{7.6}$$

Moreover, the reverse inequality holds if $p \in [0,1] \cup [2,\infty)$.

Proof of Lemma 7.2.2. We observe that (7.5) is equivalent to proving

$$1 + \left(\frac{2^{\frac{1}{k}} a^{\frac{q_k}{2k}} (1-a)^{\frac{q_k}{2k}}}{a^{\frac{q_k}{k}} + (1-a)^{\frac{q_k}{k}}} \right)^k \le \frac{1}{(a^{\frac{q_k}{k}} + (1-a)^{\frac{q_k}{k}})^k}.$$

Since $k < q_k = \log_2(2^k + 2) < k+1$ for all $k \ge 2$, by taking $p = \frac{q_k}{k}$ in Proposition 7.2.3 we obtain

$$\left(1 + \left(\frac{2a^{\frac{q_k}{2k}} (1-a)^{\frac{q_k}{2k}}}{a^{\frac{q_k}{k}} + (1-a)^{\frac{q_k}{k}}} \right)^{\frac{2k}{q_k}} \right)^{\frac{q_k}{k}-1} \le \frac{1}{(a^{\frac{q_k}{k}} + (1-a)^{\frac{q_k}{k}})}. \tag{7.7}$$

Thus, it is enough to prove

$$1+\left(\frac{2^{\frac{1}{k}}a^{\frac{q_k}{2k}}(1-a)^{\frac{q_k}{2k}}}{a^{\frac{q_k}{k}}+(1-a)^{\frac{q_k}{k}}}\right)^k \leq \left(1+\left(\frac{2a^{\frac{q_k}{2k}}(1-a)^{\frac{q_k}{2k}}}{a^{\frac{q_k}{k}}+(1-a)^{\frac{q_k}{k}}}\right)^{\frac{2k}{q_k}}\right)^{q_k-k}.$$

Defining $\mu := \frac{2a^{\frac{q_k}{2k}}(1-a)^{\frac{q_k}{2k}}}{a^{\frac{q_k}{k}}+(1-a)^{\frac{q_k}{k}}}$ (observe that $\mu \in [0,1]$ by AM-GM inequality), it is enough to prove

$$1+\frac{\mu^k}{2^{k-1}} \leq (1+\mu^{\frac{2k}{q_k}})^{q_k-k}$$

for all $\mu \in [0,1]$. By letting $z := \mu^{\frac{2k}{q_k}}$, we reduce the problem to proving

$$1+\frac{z^{\frac{q_k}{2}}}{2^{k-1}} \leq (1+z)^{q_k-k} \tag{7.8}$$

for all $z \in [0,1]$. The equality holds at $z=0$ and $z=1$. Moreover, the left hand side of (7.8) is convex in z (as $2 \leq k < q_k$), and the right hand side is concave (as $k < q_k < k+1$). Therefore (7.8) holds for all $z \in [0,1]$.

□

7.3 Proof of Theorem 7.1.3

In this section we show how to obtain Theorem 7.1.3 from Lemma 7.1.5, and then we prove this lemma. As before, we proceed via induction. Clearly, the result holds for $d=0$. Assume now $d \geq 1$, and the result has been established for $d-1$. Any set $A \subseteq \{0,1\}^d$ can be written as

$$A = (A_0 \times \{0\}) \uplus (A_1 \times \{1\})$$

for some $A_0, A_1 \subseteq \{0,1\}^{d-1}$.

We have

$$
\begin{aligned}
E_k(A) = {} & E_k(A_0) + E_k(A_1) \\
& + \sum_{i=1}^{k-1} \binom{k}{i}^2 |\{(a_1, a_2, \ldots, a_{2k}) \in A_0^i \times A_1^{k-i} \times A_0^i \times A_1^{k-i} \\
& \qquad : a_1 + \cdots + a_k = a_{k+1} + \cdots + a_{2k}\}| \\
= {} & E_k(A_0) + E_k(A_1) + \sum_{i=1}^{k-1} \binom{k}{i}^2 C_{i,k}. \qquad (7.9)
\end{aligned}
$$

Similarly to Proposition 7.2.1, we have

Proposition 7.3.1. *For all $1 \le i \le k-1$ the following inequality holds*

$$C_{i,k} \le |A_0|^{\frac{i}{k}p_k} |A_1|^{\frac{k-i}{k}p_k}. \qquad (7.10)$$

Observe that Theorem 7.1.3 follows from Proposition 7.3.1. Indeed, by (7.9), Proposition 7.3.1 and (7.2) we have

$$
\begin{aligned}
E_k(A) &= E_k(A_0) + E_k(A_1) + \sum_{i=1}^{k-1} \binom{k}{i}^2 C_{i,k} \\
&\le E_k(A_0) + E_k(A_1) + \sum_{i=1}^{k-1} \binom{k}{i}^2 |A_0|^{\frac{i}{k}p_k} |A_1|^{\frac{k-i}{k}p_k} \\
&\le (|A_0| + |A_1|)^{p_k} \\
&= |A|^{p_k}.
\end{aligned}
$$

Proof of Proposition 7.3.1. We observe that

$$C_{i,k} = \sum_{x \in \mathbb{Z}^d} |\chi_{A_0} *_{i-1} \chi_{A_0} * \chi_{A_1} *_{k-i-1} \chi_{A_1}(x)|^2,$$

where, for compactly supported f, g, we define $f * g(x) := \sum_{y \in \mathbb{Z}^d} f(y) g(x-y)$ and $*_k := *(*_{k-1})$. Indeed, this follows from the fact that

$$\chi_{A_0} *_{i-1} \chi_{A_0} * \chi_{A_1} *_{k-i-1} * \chi_{A_1}(x)$$

counts the number of k-tuples $(a_1, a_2, \dots, a_i, a_{i+1}, \dots, a_k) \in A_0^i \times A_1^{k-i}$ such that $a_1 + a_2 + \cdots + a_k = x$. Then, by Plancherel's theorem and Hölder's inequality we obtain

$$\begin{aligned} C_{i,k} &= \sum_{x \in \mathbb{Z}^d} |\chi_{A_0} *_{i-1} \chi_{A_0} * \chi_{A_1} *_{k-i-1} \chi_{A_1}(x)|^2 \\ &= \int_{\mathbb{T}^d} |\hat{\chi}_{A_0}(y)|^{2i} |\hat{\chi}_{A_1}(y)|^{2(k-i)} dm(y) \\ &\leq \left(\int_{\mathbb{T}^d} |\hat{\chi}_{A_0}(y)|^{2k} dm(y) \right)^{\frac{i}{k}} \left(\int_{\mathbb{T}^d} |\hat{\chi}_{A_1}(y)|^{2k} dm(y) \right)^{\frac{k-i}{k}} \\ &= \left(\sum_{x \in \mathbb{Z}^d} |\chi_{A_0} *_{k-1} \chi_{A_0}|^2 \right)^{\frac{i}{k}} \left(\sum_{x \in \mathbb{Z}^d} |\chi_{A_1} *_{k-1} \chi_{A_1}|^2 \right)^{\frac{k-i}{k}} \\ &= E_k^{\frac{i}{k}}(A_0) E_k^{\frac{k-i}{k}}(A_1) \leq |A_0|^{\frac{ip_k}{k}} |A_1|^{\frac{(k-i)p_k}{k}}, \end{aligned}$$

where m is the Haar measure on $\mathbb{T}^d$ with $m(\mathbb{T}^d) = 1$. □

Proof of Lemma 7.1.5. After re-scaling, we observe that to prove (7.2) it is sufficient to show

$$\sum_{i=0}^{k} \binom{k}{i}^2 x^{ip_k/k} \leq (1+x)^{p_k} \tag{7.11}$$

for all $1 \leq x < \infty$. Moreover, after a change of variable, this is equivalent to proving that

$$g_k(y) := \sum_{i=0}^{k} \binom{k}{i}^2 y^i \leq (1 + y^\alpha)^{\frac{k}{\alpha}} =: h_k(y) \tag{7.12}$$

for all $1 \leq y \leq \infty$, where $\alpha := \frac{k}{p_k} \in (1/2, 1)$. Let $f(y) := \log h_k(y) - \log g_k(y)$. We need to show $f(y) \geq 0$ for all $y \geq 1$. Observe that $f(1) = 0$. Moreover

$$\lim_{y \to \infty} f(y) = \lim_{y \to \infty} \log \left(\frac{(\frac{1}{y^\alpha} + 1)^{\frac{k}{\alpha}}}{\sum_{i=0}^{k} \binom{k}{i}^2 y^{-i}} \right) = 0,$$

and, since

$$\left(\frac{1}{y^\alpha} + 1 \right)^{\frac{k}{\alpha}} \geq 1 + \frac{k}{\alpha y^\alpha} \quad \text{and} \quad \sum_{i=0}^{k} \binom{k}{i}^2 y^{-i} = 1 + O(\frac{1}{y}),$$

we have $f(y) > 0$ whenever y is sufficiently large. Thus, it is sufficient to prove that f' changes sign at most once in $(1, \infty)$. Observe that

$$\begin{aligned} yf'(y) &= \frac{ky^\alpha}{1+y^\alpha} - \frac{\sum_{i=0}^k \binom{k}{i}^2 iy^i}{\sum_{i=0}^k \binom{k}{i}^2 y^i} \\ &= \frac{y^\alpha \sum_{i=0}^k \binom{k}{i}^2 y^i(k-i) - \sum_{i=0}^k \binom{k}{i}^2 iy^i}{(1+y^\alpha)\left(\sum_{i=0}^k \binom{k}{i}^2 y^i\right)}. \end{aligned}$$

Thus, we need to prove that $y^\alpha \sum_{i=0}^k \binom{k}{i}^2 y^i(k-i) - \sum_{i=0}^k \binom{k}{i}^2 iy^i$ changes sign in $(1, \infty)$ at most once. We define

$$\phi(y) := \log\left(y^\alpha \sum_{i=0}^k \binom{k}{i}^2 y^i(k-i)\right) - \log\left(\sum_{i=0}^k \binom{k}{i}^2 iy^i\right).$$

We then have $\phi(1) = 0$ and

$$\phi(y) = \alpha\log(y) + \log\left(\frac{n^2y^{n-1} + O(y^{n-2})}{ny^n + O(y^{n-1})}\right) \quad \text{as} \quad y \to \infty.$$

Hence $\lim_{y\to\infty}\phi(y) = -\infty$[2]. It suffices to show that ϕ' changes sign (from + to -) at most once in $(1, \infty)$. Observe that

$$\begin{aligned} \phi'(y) &= \frac{\alpha}{y} + \frac{\sum_{i=0}^k \binom{k}{i}^2 y^{i-1}(k-i)i}{\sum_{i=0}^k \binom{k}{i}^2 y^i(k-i)} - \frac{\sum_{i=0}^k \binom{k}{i}^2 i^2y^{i-1}}{\sum_{i=0}^k \binom{k}{i}^2 iy^i} \\ &= \frac{\sum_{i=0}^{2k} C_i y^i}{y\left(\sum_{i=0}^k \binom{k}{i}^2 y^i(k-i)\right)\left(\sum_{i=0}^k \binom{k}{i}^2 iy^i\right)}, \end{aligned}$$

where

$$\begin{aligned} C_i &:= \sum_{\substack{j+l=i \\ 0\le j,l\le k}} \binom{k}{j}^2\binom{k}{l}^2 [\alpha(k-l)j + (k-l)lj - j^2(k-l)] \\ &= \sum_{\substack{j+l=i \\ 0\le j,l\le k}} \binom{k}{j}^2\binom{k}{l}^2 j(k-l)(\alpha+l-j) \text{ for all } i,\ 0 \le i \le 2k. \end{aligned}$$

[2] Here we use the notation $V(y) = O(U(y))$ at y_0 to denote that an estimate of the form $|V(y)| \le C|U(y)|$, with some constant $C > 0$, holds around y_0.

Let $P(y) := \sum_{i=0}^{2k} C_i y^i$. We would like to show that $P(y)$ changes sign at most once from $+$ to $-$ in $(1,\infty)$. First, we claim $P(y)$ is a palindromic polynomial, i.e., $C_i = C_{2k-i}$ for all $i = 0,\dots,k$. Indeed,

$$C_{2k-i} = \sum_{\substack{j+l=2k-i\\ 0\le j,l\le k}} \binom{k}{j}^2 \binom{k}{l}^2 j(k-l)(\alpha+l-j) =$$

$$\sum_{\substack{(k-j)+(k-l)=i\\ 0\le j,l\le k}} \binom{k}{k-j}^2 \binom{k}{k-l}^2 (k-(k-j))(k-l)(\alpha+(k-j)-(k-l)).$$

If we denote $\tilde{l} = k-j$ and $\tilde{j} = k-l$, then we obtain

$$C_{2k-i} = \sum_{\substack{\tilde{l}+\tilde{j}=i\\ 0\le \tilde{j},\tilde{l}\le k}} \binom{k}{\tilde{l}}^2 \binom{k}{\tilde{j}}^2 \tilde{j}(k-\tilde{l})(\alpha+\tilde{l}-\tilde{j}),$$

which coincides with C_i. Since P is the palindromic polynomial it follows that y_0 is its positive root if and only if $P(1/y_0) = 0$. Therefore, to show that $P(y)$ changes sign from $+$ to $-$ at most once in $(1,\infty)$, it suffices to verify that $P(y)$ has at most two roots in $(0,\infty)$. By Descartes' rule of sign change $P(y)$ has at most two positive roots if there is at most two sign changes between consecutive (nonzero) coefficients C_i, $0 \le i \le 2k$. Since $C_i = C_{2k-i}$ it suffices to show that there is at most one sign change between consecutive (nonzero) coefficients, C_i for $0 \le i \le k$. Since $C_0 = 0$ we should consider coefficients C_i with $1 \le i \le k$. In the table below $C_i^* := \mathrm{sign}(C_i)$, and $2 \le k \le 10$.

□

Remark 7.3.2. *It seems to us that Lemma 7.1.5 holds for all $k \ge 2$. We have verified at most one sign flip of the numbers C_i, $1 \le i \le k$ on a computer for $k \le 100$. It is an interesting question to verify that there is at most one sign flip in the sequence of C_i for all k.*

Table 7.1: Sign flip for the coefficients C_i arising in Lemma 7.1.5.

k	C_1^*	C_2^*	C_3^*	C_4^*	C_5^*	C_6^*	C_7^*	C_8^*	C_9^*	C_{10}^*
2	-1	1								
3	-1	1	1							
4	-1	1	1	1						
5	-1	1	1	1	1					
6	-1	1	1	1	1	1				
7	-1	-1	1	1	1	1	1			
8	-1	-1	1	1	1	1	1	1		
9	-1	-1	-1	1	1	1	1	1	1	
10	-1	-1	-1	1	1	1	1	1	1	1

Note added in proof

Motivated by Remark 7.3.2, Vjekoslav Kovač recently proved inequality (7.2) for all $k \geq 2$, see [46].

Remark 7.3.3. *To prove* (7.11) *it suffices to show*

$$\phi_k(x) := \frac{\sum_{i=0}^{k} \binom{k}{i}^2 x^{p_k(k-i)/k}}{(1+x)^{p_k}} \leq 1 \tag{7.13}$$

for all $x \in [0,1]$*. The inequality (7.13) can be easily verified around* $x = 0$*. One can also verify it around* $x = 1$*. Therefore, to obtain the desired inequality in the whole interval* $[0,1]$ *it would be enough to prove that each* ϕ_k *has only one critical point in* $(0,1)$*. We observe that* x *is a critical point of* ϕ_k *if and only if*

$$(1+x)^{p_k+1}\phi_k'(x) = \sum_{i=0}^{k} \binom{k}{i}^2 \left[\frac{p_k(k-i)}{k} x^{p_k \frac{k-i}{k} - 1}(1+x) - p_k x^{p_k \frac{k-i}{k}} \right] = 0,$$

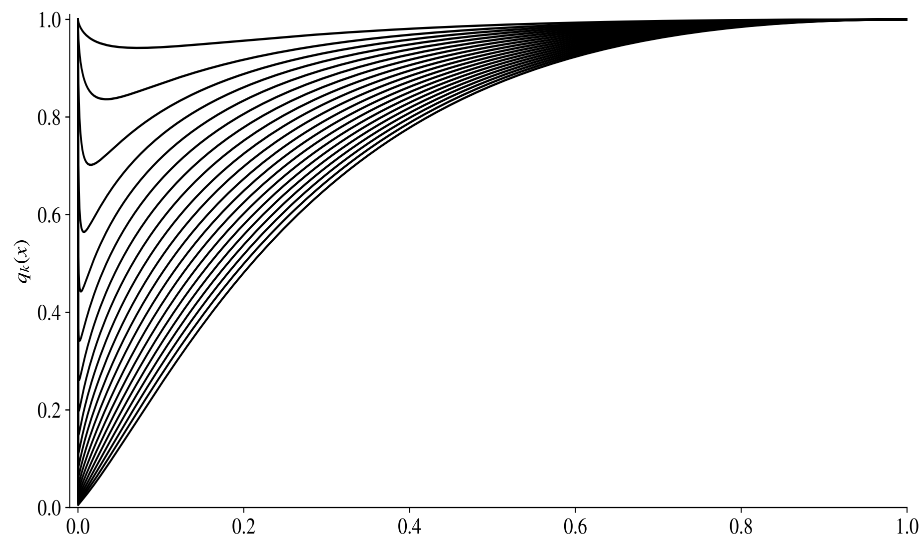


Figure 7.1: Graphs of $q_k(x)$ for $k \in \{2^n; 1 \leq n < 20\}$. The picture suggests that $q_k(x) \leq 1$ for all $x \in [0, 1]$. Lower graphs correspond to larger values of k.

or, equivalently

$$\psi_k(x) := \sum_{i=0}^{k-1} \binom{k}{i}^2 \left[\frac{k-i}{k} x^{p_k \frac{k-i}{k} - 1} - \frac{i}{k} x^{p_k \frac{k-i}{k}} \right] = 1.$$

Therefore, as $\psi_k(0) = 0$ and $\psi_k(1) = 1$, in order to establish the desired inequality, i.e., $\phi_k(x) \leq 1$ for all $x \in (0, 1)$, it would be enough to prove that $\psi_k(x)$ is concave. For small values of k, one can establish the concavity of ψ_k; in particular, this is the approach of Kane–Tao [45] for $k = 2$. Figure 7.2 illustrates that ψ_k is concave for $k = 3$. Unfortunately, this is no longer the case if k is large; e.g., Figure 7.3 illustrates the non-concavity of ψ_k for k as small as 7 already. Another approach to prove Lemma 7.1.5 would be to show $\phi_{k+1}(x) \leq \phi_k(x)$ which numerically seems correct.

7.4 Proofs of Propositions 7.1.6 and 7.1.7

The proof of Kane–Tao [45] of the $\{0, 1\}$-analogue, as well as the proofs of Theorems 7.1.1 and 7.1.3 are based on the following two steps:

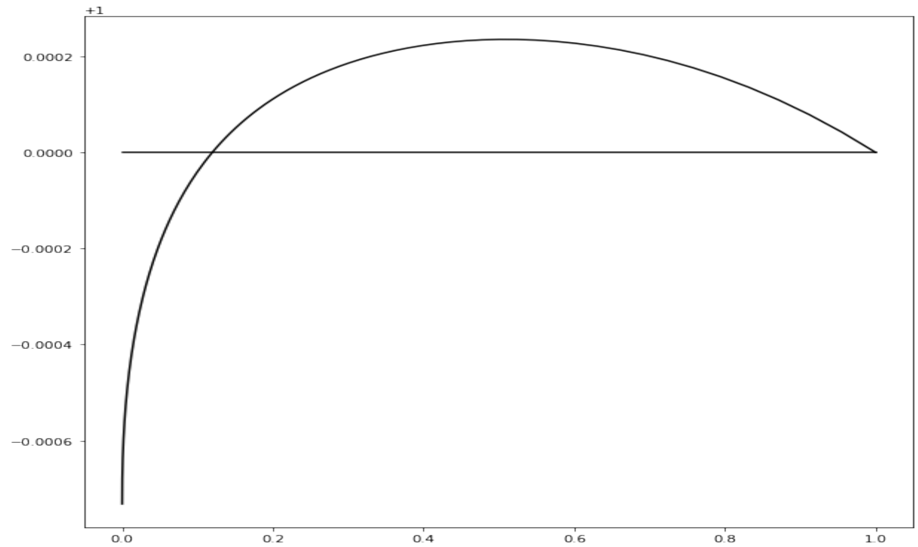


Figure 7.2: Graph of $\psi_3(x)$. We observe that $\psi_3(x)$ is concave, and $\psi_3(x)$ intersects the line $y = 1$ at only one point in $(0, 1)$.

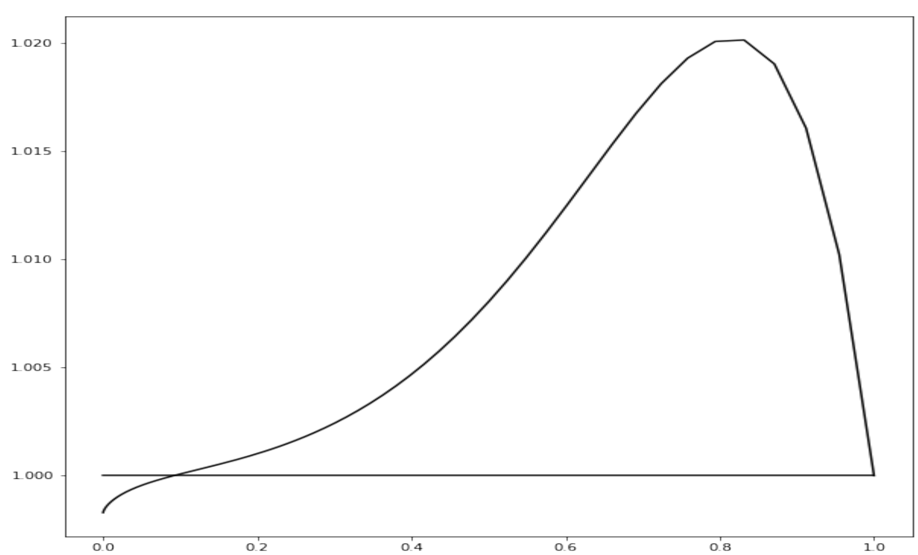


Figure 7.3: Graph of $\psi_7(x)$. We observe that $\psi_7(x)$ is not concave, however $\psi_7(x)$ still intersects the line $y = 1$ at only one point in $(0, 1)$.

- *Guessing* the extremizer to the inequality (which, in those cases, happened to be the entire set).
- Showing an inductive bound that allowed us to see that the extremizer candidate is indeed the extremizer.

In the $\{0, 1, 2\}^n$ or more general cases the entire set is not generally the extremizer, and finding the extremizer becomes a key step of the proof:

- We first construct an auxiliary problem that *inducts*, or, in this case *tensorizes* essentially by construction. Solving this problem is essentially equivalent to *guessing* the extremizers in the previous problems.
- We then show that the solution to this auxiliary problem gives rise to sharp (almost) extremizers of the original problem. This step is new, and necessary due to the fact that the extremizing sets are in general far from being product sets.

7.4.1 The auxiliary (discrete restriction) problem

For each specific instance of interest (in our case $\{0, 1, 2\}$) the auxiliary problem will then reduce to solving a finite-dimensional optimization problem closely related to the inequalities studied in the previous sections. The way to define these problems will be by defining auxiliary quantities frequently appearing in the discrete restriction theory.

Definition 7.4.1 (Discrete extension constants)**.** *Given positive integers k, d, and a finite subset $A \subset \mathbb{Z}^d$, we define:*

- *The discrete extension constant* $\mathrm{DE}_{l^q \to L^{2k}}(A)$ *as the smallest constant such that, for* *any function $f : A \to \mathbb{R}$ it holds that*

$$\left| \sum_{\substack{x_1,\dots,x_k \in A \\ y_1,\dots,y_k \in A \\ \sum x_i = \sum y_i}} f(x_1) \cdot \dots \cdot f(x_k) \cdot f(y_1) \dots f(y_k) \right|^{\frac{1}{2k}} \leq \mathrm{DE}_{l^q \to L^{2k}}(A) \|f\|_{l^q(A)}. \tag{7.14}$$

- *The restricted discrete extension constant* $\mathrm{DE}_{l^{q,1} \to L^{2k}}(A)$*, which is the best possible constant so that* (7.14) *holds for all functions* $f : A \to \{0, 1\}$.

The quantities $\mathrm{DE}_{l^{q,1} \to L^{2k}}$, $\mathrm{DE}_{l^q \to L^{2k}}$ have essentially the same value (Lemma 7.4.6), but DE is much easier to work with (Lemma 7.4.5). Moreover, understanding for which q we have $\mathrm{DE}_{l^{q,1} \to L^{2k}}(\{0,1,2\}^d) \leq 1$ is essentially equivalent to proving Proposition 7.1.6.

Lemma 7.4.2. *Let A be a finite subset of $\mathbb{Z}^d$. Let $1 \leq p = \frac{2k}{q}$, and $C > 0$. The following statements are equivalent:*

1. *For all subsets $B \subset A$, it holds that*

$$E_k(B) \leq C^{2k} |B|^p.$$

2.

$$\mathrm{DE}_{l^{q,1} \to L^{2k}}(A) \leq C.$$

Proof. Set f in the definition of $\mathrm{DE}_{l^{q,1} \to L^{2k}}$ to be equal to χ_B for B as in part (1). □

The constant DE is called the discrete extension constant because it is, indeed, the operator norm of an extension operator.

Lemma 7.4.3 (Fourier transform)**.** *Let A be a finite subset of $\mathbb{Z}^d$. Then $\mathrm{DE}_{l^q \to L^{2k}}(A)$ is the operator norm of the extension operator*[3] $\mathcal{E}(f) = \mathcal{F}\{f\}$ *from $l^q(A) \subseteq l^q(\mathbb{Z}^d)$ to $L^{2k}(\mathbb{T}^d)$.*

[3] *Here we denote by $\mathcal{F}\{f\}$ the Fourier transform of f, i.e., $\mathcal{F}\{f\}(z) = \sum_{k \in \mathbb{Z}^d} f(k) z^k$.*

Proof. By definition, $\mathrm{DE}_{l^q \to L^{2k}}(A)$ is the best constant such that, for any function $f : \mathbb{Z}^d \to \mathbb{R}$ supported on A, it holds that:

$$\| \underbrace{f * f * f \cdots * f}_{k \text{ times}} \|_{l^2(\mathbb{Z}^d)}^{1/k} \leq \mathrm{DE}_{l^q \to L^{2k}}(A) \|f\|_{l^q(\mathbb{Z}^d)}.$$

At the same time, by Plancherel's theorem and the product-convolution rule

$$\|f^{*k}\|_{l^2(\mathbb{Z}^d)} = \|\mathcal{F}\{f^{*k}\}\|_{L^2(\mathbb{T}^d)} = \|\mathcal{F}\{f\}^k\|_{L^2(\mathbb{T}^d)} = \|\mathcal{F}\{f\}\|_{L^{2k}(\mathbb{T}^d)}^k.$$

□

Remark 7.4.4. *Lemma 7.4.3 above shows that the constants* $\mathrm{DE}_{l^{q,1} \to L^{2k}}(A), \mathrm{DE}_{l^q \to L^{2k}}(A)$ *make sense for arbitrary* $2k \in \mathbb{R}$*, and not just even integers.*

The following lemma is essentially [15, Proposition 3.3]. For completeness of the argument we include the proof here.

Lemma 7.4.5 (Tensorization Lemma)**.** *Let* $q \leq 2k$*. Then for* $A \subseteq \mathbb{Z}^{d_1}$*,* $B \subseteq \mathbb{Z}^{d_2}$*,* $A \times B \subseteq \mathbb{Z}^{d_1} \times \mathbb{Z}^{d_2}$ *we have*

$$\mathrm{DE}_{l^q \to L^{2k}}(A \times B) = \mathrm{DE}_{l^q \to L^{2k}}(A)\, \mathrm{DE}_{l^q \to L^{2k}}(B).$$

Proof. The "$\geq$" inequality follows by testing the left hand side operator with the tensor product of (almost) extremizers to the right hand side.

For the opposite direction, let $f : A \times B \to \mathbb{C}$, and $\hat{f} : \mathbb{T}^{d_1} \times \mathbb{T}^{d_2} \to \mathbb{C}$ be its Fourier transform. Let $\mathcal{F}_1, \mathcal{F}_2$ be the Fourier transforms on $\mathbb{Z}^{d_1}$ and $\mathbb{Z}^{d_2}$. The goal is to estimate

$$\| \|\mathcal{F}_2\{\mathcal{F}_1 f\}(x_1, x_2)\|_{L^{2k}(x_2 \in \mathbb{T}^{d_2})} \|_{L^{2k}(x_1 \in \mathbb{T}^{d_1})}.$$

Fixing x_2, we apply the DE inequality

$$\|\mathcal{F}_2\{\mathcal{F}_1 f\}(x_1, x_2)\|_{L^{2k}(x_2 \in \mathbb{T}^{d_2})} \leq \mathrm{DE}_{l^q \to L^{2k}}(B) \|\mathcal{F}_1 f(x_1, b)\|_{l^q(b)}.$$

Now, using the hypothesis that $2k \geq q$, we can reverse the norms

$$\| \| \mathcal{F}_1 f(x_1, b) \|_{l^q(b \in B)} \|_{L^{2k}(x_1 \in \mathbb{T}^{d_1})} \leq \| \| \mathcal{F}_1 f(x_1, b) \|_{L^{2k}(x_1 \in \mathbb{T}^{d_1})} \|_{l^q(b \in B)}.$$

Now the DE inequality can be applied again to $\| \mathcal{F}_1 f(x_1, b) \|_{L^{2k}(x_1 \in \mathbb{T}^{d_1})}$. Joining it all together

$$\| \mathcal{F}_2(\mathcal{F}_1 1)(x_1, x_2) \|_{L^{2k}(x_1 \in \mathbb{T}^{d_1}))L^{2k}(x_2 \in \mathbb{T}^{d_2}))} \leq$$
$$\leq \mathrm{DE}_{l^q \to L^{2k}}(B) \, \mathrm{DE}_{l^q \to L^{2k}}(A) \| f(a, b) \|_{l^q(a \in A) l^q(b \in B)}.$$

□

7.4.2 Relating the Discrete extension problem and the original problem

In this section, we show that the discrete extension constants $\mathrm{DE}_{l^{q,1} \to L^{2k}}(A^d)$ and $\mathrm{DE}_{l^q \to L^{2k}}(A^d)$ grow similarly as d goes to infinity. This will allow us to compute the asymptotic behavior of DE in order to find the (much harder) asymptotics for $\mathrm{DE}_{l^{q,1} \to L^{2k}}$. The next lemma is inspired by Bourgain's logarithmic pigeonhole principle (see [70]).

Lemma 7.4.6 (Comparison Lemma). *For all $q \geq 1$, $k \geq \frac{1}{2}$, $A \subseteq \mathbb{Z}^d$ it holds that*

$$\mathrm{DE}_{l^{q,1} \to L^{2k}}(A) \leq \mathrm{DE}_{l^q \to L^{2k}}(A) \leq (2 + \log |A|) \, \mathrm{DE}_{l^{q,1} \to L^{2k}}(A).$$

Proof. The first inequality follows by the fact that DE is a maximum over a larger class of functions. For the second one, let $f : A \to \mathbb{R}$. Without loss of generality assume $\|f\|_{l^\infty(A)} = 1$, and that f is nonnegative. We can decompose f as a sum

$$f(x) = \sum_{\substack{i \geq 1 \\ 2^i \leq |A|}} 2^{-i} \epsilon_i(x) + f_0(x)$$

with the property that $\epsilon_i : A \to \{0, 1\}$, and $0 \leq f_0(x) \leq |A|^{-1}$. The value of $\epsilon_i(x)$ is the $i-$th digit of the boolean expansion of $f(x)$. Moreover, $\|f_0\|_1 \leq 1$. There are, moreover at most $(\log |A| + 1)$ different ϵ_i. By the triangle inequality, we have

$$\|\hat{f}\|_{L^{2k}(\mathbb{T}^d)} \leq \sum_{\substack{i \geq 1 \\ 2^i \leq |A|}} 2^{-i} \|\hat{\epsilon}_i\|_{L^{2k}(\mathbb{T}^d)} + \|\hat{f}_0\|_{L^{2k}(\mathbb{T}^d)}.$$

We bound the sum by the maximum element in the sum (times the number of elements), and the term $\|\hat{f}_0\|_{L^{2k}(\mathbb{T}^d)}$ by 1, to obtain

$$\|\hat{f}\|_{L^{2k}(\mathbb{T}^d)} \leq (1 + \log(|A|)) \max_{i \geq 1} 2^{-i} \|\hat{\epsilon}_i\|_{L^{2k}(\mathbb{T}^d)} + 1.$$

Now, by applying the $\mathrm{DE}_{l^{q,1} \to L^{2k}}$ bounds on $\hat{\epsilon}_i$ we get

$$\|\hat{f}\|_{L^{2k}(\mathbb{T}^d)} \leq (1 + \log(|A|))\, \mathrm{DE}_{l^{q,1} \to L^{2k}}(A) \max_{i \geq 1} 2^{-i} \|\epsilon_i\|_{L^q(A)} + 1.$$

By construction $2^{-i}\|\epsilon_i\|_{L^q(A)} \leq \|f\|_{L^q(A)}$. By checking against a singleton, $\mathrm{DE}_{l^{q,1} \to L^{2k}}$ is always at least 1, and $\|f\|_{l^q(A)} \geq \|f\|_{l^\infty(A)} = 1$. Combining all this, we obtain

$$\|\hat{f}\|_{L^{2k}(\mathbb{T}^d)} \leq (2 + \log(|A|))\, \mathrm{DE}_{l^{q,1} \to L^{2k}}(A) \|f\|_{l^q(A)}.$$

□

Remark 7.4.7. *The exponent of the* log *in Lemma 7.4.6 is probably not sharp (see, for example, the gains in the log-power in [32, Theorem 1.1] or [55, Lemma 2.4]). Finding the sharp exponent is not necessary for our purposes. We thank A. Mudgal for this remark.*

The results from this section yield the relationship between Proposition 7.1.6 and the discrete extension constant, as follows.

Proposition 7.4.8. *Let A be a finite subset of $\mathbb{Z}$. Let $1 \leq p = \frac{2k}{q}$, and $C > 0$. The following are equivalent:*

1. *An inequality of the form*

$$E_k(X) \leq C|X|^p$$

   holds for all $X \subseteq A^d$, $d \geq 0$.

2. *An inequality of the form*

$$E_k(X) \leq |X|^p.$$

   holds for all $X \subseteq A^d$, $d \geq 0$.

3. $\mathrm{DE}_{l^q\to L^{2k}}(A) \le 1$.

Proof. Clearly, (3) ⇒ (2) (by Lemma 7.4.5 and Lemma 7.4.2), and (2) ⇒ (1). We show that (1) ⇒ (3). By Lemmas 7.4.6 and 7.4.5 we have

$$\mathrm{DE}_{l^{q,1}\to L^{2k}}(A^d) \le \mathrm{DE}_{l^q\to L^{2k}}(A)^d \le (2 + d\log|A|)\,\mathrm{DE}_{l^{q,1}\to L^{2k}}(A^d). \tag{7.15}$$

Observe that by Lemma 7.4.2, (1) is equivalent to

$$\sup_d \mathrm{DE}_{l^{q,1}\to L^{2k}}(A^d) < \infty. \tag{7.16}$$

By equation (7.15), equation (7.16) is equivalent to

$$\mathrm{DE}_{l^q\to L^{2k}}(A)^d \le 1$$

and the result follows. □

Remark 7.4.9. *The proof of Theorem 7.4.8 extends to any finite subset A of an abelian group G without any significant changes, using that the group generated by A inside of G is locally compact and abelian with the discrete topology.*

7.4.3 Concluding the proofs of Propositions 7.1.6 and 7.1.7

Proof of Proposition 7.1.6. Applying Proposition 7.4.8 with $A = \{0, 1, 2\} \subseteq \mathbb{Z}$ and $k = 2$ shows that t_2 is equal to the smallest p such that

$$\frac{x^p + y^p + z^p + 4(x^{p/2}y^{p/2} + x^{p/2}z^{p/2} + y^{p/2}z^{p/2}) + 4x^{p/2}y^{p/4}z^{p/4}}{(x+y+z)^p} \le 1,$$

for all $x, y, z \ge 0$. In particular, taking $x = 1$ and $y = z = 1/2$ we obtain

$$\begin{aligned} t_2 &\ge \inf\{p \in [2,3]\colon\ 4^p - 2^p - 12(2^{p/2}) - 6 \ge 0\} \\ &= \inf\{2\log_2 w\colon w \in [2, 2\sqrt{2}],\ w^4 - w^2 - 12w - 6 \ge 0\} \\ &\ge 2\log_2(2.5664) > \log_3 19 = \frac{\log E(\{0,1,2\}^d)}{\log|\{0,1,2\}^d|}. \end{aligned}$$

□

Proof of Proposition 7.1.7. The upper bound is trivial, so we focus our attention on the lower bounds. Consider the case $n = 2m-1$. We prove that $E(\{0,1,\ldots,2m-1\}) = \frac{16m^3+2m}{3}$. We start observing that for any $a \in \{0,1,\ldots,2m-1\}$ the 4-tuple (a,a,a,a) is a solution. Moreover, for all $a,b \in \{0,1,\ldots,2m\}$ we have that $(a,b,a,b),(a,b,b,a),(b,a,b,a)$ and (b,a,a,b) are also solutions. This gives a total of $2m + 4\binom{2m}{2}$ trivial solutions.

Then, we observe that the m couples $(0,2m-1),(2,2m-3),(3,2m-4),\ldots,(m-1,m)$ add up to $2m-1$, this gives $8\binom{m}{2}$ nontrivial solutions. Similarly, the couples adding up to $2m-2$ and $2m$ give $8\binom{m-1}{2} + 4\binom{m-1}{1}$ solutions. More generally, we have that considering the couples adding k or $4m-2-k$ we obtain $8\binom{\lceil k/2\rceil}{2}$ non-trivial solutions if k is odd and $8\binom{k/2}{2} + 4\binom{k/2}{1}$ if k is even. Therefore

$$\begin{aligned}
&E_2(\{0,1,\ldots,2m-1\}) \\
&= 2m + 4\binom{2m}{2} + 8\binom{m}{2} + 4\left(8\sum_{k=2}^{m-1}\binom{k}{2}\right) + 2\left(4\sum_{k=1}^{m-1}k\right) \\
&= \frac{16m^3+2m}{3}.
\end{aligned}$$

The case $n = 2m$ follows similarly. □

Acknowledgments

We are grateful to the anonymous referees and to Tim Gowers for thoughtful comments and corrections which improved the exposition of the paper. We are also thankful to Terence Tao for helpful discussions.

This work was initiated at the Hausdorff Research Institute for Mathematics, during the trimester program "Harmonic Analysis and Analytic Number Theory"; we are grateful to the institute and the organizers of the program.

APPENDIX A

Supplementary material: Basic analysis and geometry of local fields

The goal of this sections is to provide basic facts of the analysis and geometry of local fields, essentially showing that, for the purposes of our decomposition theorem in Chapter 2. Essentially they show that the geomtery and analysis of local fields of sufficiently high characteristic behaves *no worse* than that of $\mathbb{C}$. In practice, the harmonic analysis on $\mathbb{Q}_p$ is in fact much better behaved than that over $\mathbb{C}, \mathbb{R}$ [52], but we will not use that fact. For an introduction to the analysis of local fields with an eye towards restriction, see [42], and for a general introduction see [69].

A.1 Polynomials with positive coefficients in local fields

The goal of this appendix is to show that, after a suitable decomposition into sectors (as in definition 2.1.1), polynomials with positive coefficients (positive in the sense that the coefficients belong to $\mathbb{Z} \geq 0$) enjoy a certain *reverse triangle inequality.*

On a first read, the reader may wish to focus on the case in which $\mathbb{K} = \mathbb{R}$, or the case $\mathbb{K} = \mathbb{C}$. In this case, the final goal of this section, Lemma A.1.6 can be checked directly (we leave the details to the reader) and the appendix can be omitted.

A.1.1 Sectors

We will let $K = \mathbb{R}$ and K_+ be $\mathbb{R}_{>0}$ if $\mathbb{K} \geq \mathbb{R}$ (in other words, if $\mathbb{K} = \mathbb{R}, \mathbb{C}$), and $K_+ = p^{\mathbb{Z}}$ if $K \geq \mathbb{Q}_p$. The motivation for this definition is the following *reverse triangle inequality* fact:

Lemma A.1.1 (Reverse triangle inequality for K_+)**.** *Let $M \in \mathbb{Z}_{>0}$, let K be either the real numbers or $\mathbb{Q}_p$. Let $a_1, \dots a_M \in K_+$. Then it holds that*

$$|\sum_{i=1}^{M} a_i| \approx \sum_{i=1}^{M} |a_i|, \tag{A.1}$$

where the implicit constant only depends on M.

The proof of Lemma A.1.1 is separated in two cases. The first case, when $K = \mathbb{R}$ is essentially by definition. If $K = \mathbb{Q}_p$ we argue as follows:

Proof of Lemma A.1.1 (when $K = \mathbb{Q}_p$). By scale and permutation invariane of the inequality, assume that the a_i have nonincreasing norm, and that $\|a_1\| = 1$. In that case, the right hand side is approximately one. Now we will see that the value of the left hand side is greater than or equal to $p^{1-\lceil M/(p-1)\rceil}$. We will prove that by induction on $k = \lceil M/(p-1)\rceil$.

If $k = 1$, then $a = \sum_{i=1}^{M} a_i$ cannot be congruent to p,because we are adding less than p numbers which are either 0 or 1 mod p, and at least a_1 is equal to 1 mod p. If $k > 1$ there are two options: Either a is not congruent to p, and $|a| = 1$, or, for some $s \in \mathbb{Z}_{>0}$, the first $s \cdot p$ of the a_i are equal to 1. In the later case, let b_i be a sequence of $M - s(p-1)$ elements defined as:

$$b_k := \begin{cases} 1 & \text{if } k \leq s \\ a_{k-s(p-1)}/p & \text{if } k \geq s \end{cases} \tag{A.2}$$

then $\sum_{i=1}^{M} a_k = p \sum_{i=1}^{M-p(s-1)} b_k$. By induction, $|\sum_{i=1}^{M-p(s-1)} b_k| \geq p^{1+s-\lceil M/(p-1)\rceil}$, and the result closes by multiplicativity of the absolut value.

□

The result above, while having the right intuition, will not be enough for our purposes. The reason is that one cannot generally cover $\mathbb{K}$ by a finite number of sets of the form aK_+, for $a \in \mathbb{K}$. We will need the following generalization of K_+, already defined in the introduction.

Definition A.1.2. *Given a locally compact field $\mathbb{K}$ of characteristic zero we define the sector $\Sigma_\epsilon^{\mathbb{K}}$ of amplitude $\epsilon \in \mathbb{R}$ as the set*

$$\Sigma_\epsilon^{\mathbb{K}} = \{x \in \mathbb{K} : d(x, K_+) < \epsilon|x|\}.$$

Whenever $\mathbb{K}$ is clear by the context we will write Σ_ϵ. For an element $t \in \mathbb{K} \setminus \{0\}$ we will denote by $t\Sigma_\epsilon$ the set $\{tx, x \in \Sigma_\epsilon\} = \{x : d(t^{-1}x, K_+) < |t|^{-1}|x|\}$.

The sets $\Sigma_\epsilon^{\mathbb{K}}$ for arbitrary fields have a very similar behavior to the open complex sectors[1]. In this work we will use the following properties of sectors:

Lemma A.1.3 (Algebraic properties of sectors)**.** *Let $K, \mathbb{K}$ as in definition 2.1.1. Let $M \in \mathbb{Z}_{>0}$. Then the following hold:*

1. *For $t, t' \in \mathbb{K}$, and $\epsilon, \epsilon' \in \mathbb{R}^+$ it holds that:*

$$(t\Sigma_\epsilon^{\mathbb{K}}) \cdot (t'\Sigma_{\epsilon'}^{\mathbb{K}}) \subseteq (tt')(\Sigma_{\epsilon+\epsilon'+\epsilon\epsilon'}^{\mathbb{K}})$$

2. *If ϵ is small enough ($\epsilon = \epsilon_{m,p}$), and $a_1, \ldots a_M \in t\Sigma_\epsilon^{\mathbb{K}}$ then we have a reverse triangle inequality*

$$|\sum_{i=1}^M a_i| \approx \sum_{i=1}^M |a_i|,$$

where the implicit constant may depend on M

[1] In the case $\mathbb{K} = \mathbb{C}$ it holds that $S_\epsilon = \{z \in \mathbb{C}^\times, |\arg z| < \arcsin \epsilon\}$

One thing that is lost in this generalization from $\mathbb{Q}$ to $\mathbb{Q}_p$ is the sub-distributive property for complex sectors: If $\mathbb{K} = \mathbb{R}, \mathbb{C}$ it holds that $(t\Sigma_\epsilon^{\mathbb{K}}) + (t\Sigma_{\epsilon'}^{\mathbb{K}}) \subseteq (t+t')(\Sigma_{\max(\epsilon,\epsilon')}^{\mathbb{K}})$ whenever $\epsilon, \epsilon' < 1$. In the general case this is replaced by the weaker reverse triangle inequality in *(2)* above.

Proof. *(1)* Let $z \in (t\Sigma_\epsilon^{\mathbb{K}})$, and $s \in tK_+$ such that $|z - s| \leq \epsilon|z|$. Define z', s' analogously for $t'\Sigma_{\epsilon'}^{\mathbb{K}}$. Then

$$|zz' - ss'| \leq |zz' - sz'| + |z's - ss'| \leq \epsilon|z||z'| + \epsilon'|z'||s| \leq (\epsilon + \epsilon' + \epsilon\epsilon')|zz'|$$

Therefore $zz' \in tt'\Sigma_{\epsilon+\epsilon'+\epsilon\epsilon'}$

(2) Let C be the implicit constant in (A.1), and $\epsilon < \frac{1}{2C}$. For each a_i let $s_i \in tK_+$ such that $|a_i - s_i| \leq \epsilon|a_i|$. In particular, $\sum_{i=1}^M |s_i| \approx \sum_{i=1}^M |a_i|$. By construction:

$$|\sum_{i=1}^M a_i - \sum_{i=1}^M s_i| \leq \epsilon \sum_{i=1}^M |a_i| \leq \frac{1}{2}|\sum_{i=1}^M a_i| \tag{A.3}$$

and therefore $|\sum_{i=1}^M a_i| \approx |\sum_{i=1}^M s_i|$. The conclusion now follows by applying Lemma A.1.1 to the $(t^{-1}s_i)_{i=1}^M$. □

The main reason to construct sectors is that $\mathbb{K} \setminus \{0\}$ can be covered by finitely many sectors. Using the (local) compactness of $\mathbb{K}$ this can be made explicit as:

Lemma A.1.4 (Finite covers of sectors)**.** *Let $K, \mathbb{K}$ as in definition 2.1.1. Let $\mathbb{K}^\times = \bigcup_{i\in I} t_i S$ denote a family of sectors of $\mathbb{K}$ that cover $\mathbb{K} \setminus \{0\}$. Then there is a finite subcover $I' \subseteq I$ such that $\mathbb{K}^\times = \bigcup_{i\in I'} t_i S$.*

Proof. In all the considered fields, for any $z \in \mathbb{K}^\times$ there exists $t \in K_+$ such that $C^{-1} \leq$

$|tz| \le C$, where C is a constant that depends on $\mathbb{K}$ only[2] . Let $x_1, \dots x_K$ in $\bar{A}_{c^{-1},C}$ so that

$$\bar{A}_{C^{-1},C} \subseteq \bigcup_{i=1}^{K} B_{\epsilon C^{-1}}(x_i).$$

Then the sectors $(x_i \Sigma_\epsilon)_{i=1}^K$ will cover $K^\times$.

Define an annulus $A_{r,R}(0)$ as follows:

Definition A.1.5. *An annulus with center $z_0 \in \mathbb{K}$ and outer and inner radius $0 \le r < R \le \infty$ will be denoted by $A^{\mathbb{K}}_{r,R}(z_0) := \{z \in \mathbb{K}, |z - z_0| \in (r, R)\}$. When $\mathbb{K}$ is clear from the context,the field $\mathbb{K}$ will be dropped and the annulus will be denoted by $A_{r,R}(z_0)$.*

Given a sector $t\Sigma_\epsilon$, we know $z \in t\Sigma_\epsilon$ if an only if $kz \in t\Sigma_\epsilon$ for any $k \in K_+$. Therefore the result will follow if we can show that $A_{(2C)^{-1},2C}(0) \subseteq \bigcup_{i \in I'} t_i S$ for some finite subfamily I'. The closure $\overline{A}_{(2C)^{-1},2C}(0)$ is a compact set, by the local compactness of $\mathbb{K}'$ and $\bigcup_{i \in I} t_i S$ is an open cover of it. Therefore the result follows by compactness. □

A.1.2 Homogeneous polynomials with positive coefficients

We can now state the main result of this section which is a certain reverse triangle inequality for polynomials with positive coefficients. Let q be a homogeneous polynomial in n variables of degree d in $\mathbb{K}^d$ of the form $q(x) = \sum_{i=1}^M x^I$, for $x \in \mathbb{K}^d$ where each I_i is a multi-index with $|I_i| = n$. The arguments in the sequel will need a reverse-triangle inequality for q with respect to its monomials. The proof of the inequality DW in Theorem 2.1.13 for $\mathbb{K} = \mathbb{R}$ given by Dendrinos and Wright (in the case $\mathbb{K} = \mathbb{R}$) strongly uses that $\mathbb{R}$ is an ordered field. The reverse triangle inequality in the following lemma will be the substitute for that fact.

Lemma A.1.6 (Reverse triangle inequality)**.** *Let q be a homogeneous polynomial in n variables of degree n in $\mathbb{K}^d$ of the form $q(x) = \sum_{i=1}^M x^{I_i}$, for $x \in \mathbb{K}^d$ where each I_i is a multi-*

[2] Letting $C = 1$ will suffice if $K = \mathbb{R}$, and $C = p$ will work if $K = \mathbb{Q}_p$ with the usual valuation, because if $|z| > p$, then $|pz| = p^{-1}|z|$ will be closer to 1.

index with $|I_i| = n$. Then there is $\epsilon(M, n, d)$ small enough such that, if all the components of $x = (x_1, \dots x_d)$ belong to the same sector of amplitude ϵ (that is, $x \in (t\Sigma_\epsilon)^d$ for some $t \in K^\times$) it holds that:

$$|q(x)| = |\sum_{i=1}^{M} x^{I_i}| \approx \sum_{i=1}^{M} |x|^{I_i} = q(|x|) \tag{A.4}$$

where $|x| := (|x_1|, \dots, |x_d|) \in \mathbb{R}_+^d$. The implicit constants are may deppend on all te variables but the x_i.

In particular, for any $\epsilon_0 > 0$ there is a $\delta = \delta(\epsilon_0, n, d)$, such that for any polynomial $\tilde{q}(x) = \sum_{i=1}^{M} s_i x^{I_i} \in \mathbb{K}[x]$ of degree n with $|1 - s_i| < \delta$, it holds that

$$|q(x) - \tilde{q}(x)| \leq \epsilon_0 |q(x)| \tag{A.5}$$

whenever $x \in (t\Sigma_\epsilon)^d$.

Proof. First note that (A.5) follows by the trinagle inequality once (A.4) is proven, so we will focus our attention on (A.4).

The first and the third equalities are true by definition. For the approximate equality, a repeated application of *(1)* in Lemma A.1.3 (multiplicativity property for sectors) shows that $x^{I_i} \in t^d \Sigma_{d\epsilon} \subseteq t^d \Sigma_{1/2}$. Now, applying *(2)* in Lemma A.1.3 (reverse triangle for sectors of amplitude $\leq \frac{1}{2}$) to the sum the result follows using that $|x|^{I_i} = |x^{I_i}|$. □

A.2 Calculus over $p-$adic field extensions

This appendix contains the proofs (or sketches) of results that are known over the reals, and whose proof in the complex/p-adic scenario is are essentially the same.

Proof of Lemma 3.1.6. Clearly, the second inequality in Lemma 3.1.6 follows from the first one, so we will follow on the first one.

The usual rules of calculus (such as chain rule) apply on the $p-$adics and complex numbers (with the suitable normalization in the Jacobian) in the same ways they apply over the real numbers (see, for example, [61]). The fact that the Jacobian of Φ doesn't vanish means that Φ is locally smoothly invertible near each preimage (by the usual fixed-point proof of the inverse function theorem), so we may apply the usual chain rule at a neighborhood of each of the pre-images of Φ. □

Proof of Lemma 3.1.7. The proof starts by studying the set of powers $p_i, q_{i,j}$ with $\sum_i p_i^{-1} + \sum_{i,j} q_i^{-1} = l$ for which the estimate

$$\int_{\mathbb{C}^l \sim \mathbb{R}^{2l}} \prod_{1\leq i\leq l} f_i(z_i) \prod_{1\leq i<j\leq l} g_{i,j}(z_i - z_j) dz \lesssim \prod_{i=1}^{l} \|f_i\|_{p_i} \prod_{1\leq i<j\leq l} \|g_{i,j}\|_{q_{i,j}} \tag{A.6}$$

holds. We will denote by capital letters the vectors $(p_1^{-1}, \dots p_l^{-1}, q_{1,2}^{-1}, \dots q_{l-1,l}^{-1})$, which we will think of as elements of the affine subspace $H := \{\sum_i p_i^{-1} + \sum_{i,j} q_i^{-1} = l\}$.

The base cases are $A := (p_i^{-1} = 1, q_{i,j}^{-1} = 0)$ and $B := (p_i^{-1} = \delta_{i,1}, q_{i,j}^{-1} = \delta_{i+1,j})$, which follow from Fubini's theorem. Now, the result is invariant over permutations over all the indices (i, j). This allows us to extend the second base case B to all the cases B_σ permutations obtained from B by permutations.

By Riesz-Thorin, the result is then true for $A' := \frac{1}{l!} \sum_{\sigma \in S_l} B_\sigma = (p_i = \frac{1}{l}, q_i = \frac{1}{2l})$. By Riesz-Thorin again, the result is true for all the points interpolating A and A'. This proves the strong version of the theorem.

To get the weak estimate, it suffices to show that all the points joining A and A' lie on the interior of the interpolation polytope (interior with the affine topology on H). By convexity again, it suffices to show that A' does. The geometric argument can be seen, for example, in [17] (since it is an argument in the *space of exponents* it is exactly the same as in the real case). □

List of Notation

$\mathbb{K}$	A local field of characteristic zero: Either $\mathbb{R}$, $\mathbb{C}$, or a finite extension of $\mathbb{Q}_p$..
.	$A \lesssim_X B$. There exists a constant $C(X)$ so that $A \lesssim C(X)B$. Unless noted otherwise, the constant is allowed to depend on the dimension and the ambient field.
$[N]$	The set of numbers $\{0, 1, \ldots, n-1\}$. .
$\mathbb{K}$	A local field of characteristic zero or sufficiently high (context dependent) characteristic. .
$\\|A\\|_{\mathrm{Op}(X \to Y)}$	The operator norm of A as an operator from X to Y. .
$R_{\gamma,\mu}$	Restriction operator to a curve γ, $f \mapsto \hat{f} \circ \gamma$. See $E_{\gamma,\mu}$ for the need of μ.. See Definition 1.0.1.
$E_{\gamma,\mu}$	Extension operator on a curve, $E_{\gamma,\mu} f(x) = \int_{\mathbb{K}} \exp(i\gamma(t) \cdot x) f(x) d\mu(x)$. Formal adjoint of $R_{\gamma,\mu}$ with respect to the product $(f, g) \mapsto \int_{\mathbb{K}} f(x)g(x)d\mu(x)$. See Definition 1.0.3.
$T_{\gamma,dt}$	The operator $(A_{\gamma,dt} f)(x) := \int_{\mathbb{K}} f(x - \gamma(t))dt$. See Definition 1.0.4.
$\mathrm{Dec}_{l^p L^q}(\mathcal{U})$	The $l^p L^q$ decoupling constant for a family of sets $\mathcal{U}$. See Definition 1.0.6.
$\mathrm{DE}_{l^p \to L^q}(S)$	The $l^p \to L^q$ extension operator norm of a set $S \subseteq \mathbb{Z}^d$. .
$\Sigma_\epsilon^{\mathbb{K}}$	The angular sector $\{x \in \mathbb{K} : d(x, K_+) < \epsilon\|x\|\}$. The field $\mathbb{K}$ is ommited if clear from the context.. See Definition 2.1.1.
$A_{r,R}^{\mathbb{K}}(z_0)$	The annulus $\{z \in \mathbb{K}, \|z - z_0\| \in (r, R)\}$. The field $\mathbb{K}$ is ommited if clear from the context.. See Definition 2.1.2.
$\Lambda[\gamma](z_1, \ldots, z_k)$	The differential form $\gamma'(z_1) \wedge \cdots \wedge \gamma'(z_k)$. .

$\Lambda^{(k)}[\gamma](z)$	The differential form $\gamma'(z) \wedge \gamma''(z) \wedge \cdots \wedge \gamma^{(k)}(z)$. .
$v(z_1, \ldots, z_n)$	The vandermonde determinant $\prod_{i<j}(z_j - z_i)$. .
$\tilde{\Lambda}[\gamma](z_1, \ldots, z_k)$	The ratio $\frac{\Lambda[\gamma](z_1,\ldots,z_k)}{v(z_1,\ldots,z_k)}$, continuously extended to the set $v(z_1, \ldots, z_k) = 0$. .
$M[\gamma]$	For a polynomial curve $\gamma : \mathbb{K} \to \mathbb{K}^d$ of degree N, the $d \times N$ matrix containing the coefficients of degree ≥ 1 of γ. .
Canonical form	Certain form of a polynomial curve where some monomials appear only once. See Definitions 2.1.5, 2.1.6.
ϵ-similar	Quantitative local notion of similarity between a polynomial curve and a monomial curve. See Definitions 2.1.10, 2.1.11.
$[C_q^{\{d_1,\ldots,d_k\}}]_j$	Arithmetic Cantor se, numbers containing only the digits $d_1, \ldots d_k$ in base q. See Section 6.1.2.2.
$[C_q^{\{d_1,\ldots,d_k\}}]_j$	$j-$th level of the construction of $C_q^{\{d_1,\ldots,d_k\}}$. See Section 6.1.2.2.
$[\mathcal{E}_q^{\{d_1,\ldots,d_k\}}]_j$	Set of integers of at most j digits in base q, all of which are in $d_1, \ldots d_k$. See Section 6.1.2.2.
$\mathrm{DE}_{\ell^2 \to L^p}(S)$	Short-hand notation for the discrete extension operator norm $\|\mathrm{DE}(S)\|_{\mathrm{Op}(\ell^2(S) \to L^2([0,1]^m))}$. See Section 6.1.2.1.

Bibliography

[1] Gennady I. Arkhipov, Vladimir N. Chubarikov, and Anatoly A. Karatsuba. *Trigonometric Sums in Number Theory and Analysis.* Walter de Gruyter, Aug. 2008. ISBN: 978-3-11-019798-3.

[2] Jong-Guk Bak and Seheon Ham. "Restriction of the Fourier transform to some complex curves". In: *Journal of Mathematical Analysis and Applications* 409.2 (Jan. 2014), pp. 1107–1127. ISSN: 0022-247X.

[3] Jong-Guk Bak, Daniel M. Oberlin, and Andreas Seeger. "Restriction of Fourier transforms to curves and related oscillatory integrals". In: *American Journal of Mathematics* 131.2 (Mar. 2009), pp. 277–311. ISSN: 1080-6377.

[4] Jong-Guk Bak, Daniel M. Oberlin, and Andreas Seeger. "Restriction of Fourier transforms to curves II: some classes with vanishing torsion". In: *Journal of the Australian Mathematical Society* 85.1 (Aug. 2008), pp. 1–28. ISSN: 1446-8107, 1446-7887.

[5] Jong-Guk Bak, Daniel M. Oberlin, and Andreas Seeger. "Restriction of Fourier transforms to curves: An endpoint estimate with affine arclength measure". In: *Journal für die reine und angewandte Mathematik (Crelles Journal)* 2013.682 (2012), pp. 167–205. ISSN: 1435-5345.

[6] W. Beckner, A. Carbery, S. Semmes, and F. Soria. "A note on Restriction of the Fourier Transform to Spheres". In: *Bulletin of the London Mathematical Society* 21.4 (July 1989), pp. 394–398. ISSN: 0024-6093.

[7] Kirsti D. Biggs. *Efficient congruencing in ellipsephic sets: the general case.* 2019.

[8] Kirsti D. Biggs. "Efficient congruencing in ellipsephic sets: the quadratic case". In: *Acta Arithmetica* 200.4 (2021), pp. 331–348.

[9] J. Bourgain. "Bounded orthogonal systems and the $\Lambda(p)$-set problem". In: *Acta Mathematica* 162.3-4 (1989), pp. 227–245.

[10] Jean Bourgain and Ciprian Demeter. “A study guide for the ℓ^2-decoupling theorem,” in: *Chinese Annals of Mathematics. Series B* 38.1 (2017), pp. 173–200.

[11] Jean Bourgain and Ciprian Demeter. “The proof of the ℓ^2 decoupling conjecture”. In: *Annals of Mathematics* 182.1 (2015), pp. 351–389.

[12] Jean Bourgain, Ciprian Demeter, and Larry Guth. “Proof of the main conjecture in Vinogradov’s mean value theorem for degrees higher than three”. In: *Annals of Mathematics* 184.2 (2016), pp. 633–682.

[13] Jean Bourgain, Stephen Dilworth, Kevin Ford, Sergei Konyagin, and Denka Kutzarova. “Explicit constructions of RIP matrices and related problems”. In: *Duke Mathematical Journal* 159.1 (2011), pp. 145–185.

[14] Eric A. Carlen, Rupert L. Frank, Paata Ivanisvili, and Elliott H. Lieb. “Inequalities for L^p-norms that sharpen the triangle inequality and complement Hanner’s inequality”. In: *The Journal of Geometric Analysis* 31.6 (2021), pp. 4051–4073.

[15] Alan Chang, Jaume de Dios, Rachel Greenfeld, Asgar Jamneshan, Zane Kun Li, and José Madrid. “Decoupling for fractal subsets of the parabola”. In: *Mathematische Zeitschrift* (2020).

[16] Michael Christ. “Convolution, curvature, and combinatorics: a case study”. In: *International Mathematics Research Notices* 1998.19 (1998), pp. 1033–1048.

[17] Michael Christ. “On the restriction of the Fourier transform to curves: endpoint results and the degenerate case”. In: *Transactions of the American Mathematical Society* 287.1 (1985), pp. 223–238. ISSN: 0002-9947, 1088-6850.

[18] Hyunuk Chung and Seheon Ham. “Convolution estimates for measures on some complex curves”. In: *Annali di Matematica Pura ed Applicata (1923 -)* 198.3 (June 2019), pp. 837–867. ISSN: 1618-1891.

[19] Ciprian Demeter. “A decoupling for Cantor-like sets”. In: *Proceedings of the American Mathematical Society* 147.3 (2019), pp. 1037–1050.

[20] Ciprian Demeter. *Fourier restriction, decoupling and applications*. Vol. 184. Cambridge University Press, 2020.

[21] Spyridon Dendrinos, Kevin Hughes, and Marco Vitturi. "Some Subcritical Estimates for the ℓ^p-Improving Problem for Discrete Curves". In: *Journal of Fourier Analysis and Applications* 28.4 (Aug. 5, 2022), p. 69. ISSN: 1531-5851.

[22] Spyridon Dendrinos, Norberto Laghi, and James Wright. "Universal L p improving for averages along polynomial curves in low dimensions". In: *Journal of Functional Analysis* 257.5 (Sept. 2009), pp. 1355–1378. ISSN: 00221236.

[23] Spyridon Dendrinos and Betsy Stovall. "Uniform bounds for convolution and restricted X-ray transforms along degenerate curves". In: *Journal of Functional Analysis* 268.3 (Feb. 2015), pp. 585–633. ISSN: 0022-1236.

[24] Spyridon Dendrinos and James Wright. "Fourier restriction to polynomial curves I: A geometric inequality". In: *American Journal of Mathematics* 132.4 (2010), pp. 1031–1076. ISSN: 0002-9327. JSTOR: 40864468.

[25] Jaume de Dios Pont, Rachel Greenfeld, Paata Ivanisvili, and Jose Madrid. *Additive energies on discrete cubes*. Dec. 17, 2021. arXiv: arXiv:2112.09352. URL: http://arxiv.org/abs/2112.09352 (visited on 04/11/2023). preprint.

[26] S. W. Drury. "Restrictions of Fourier transforms to curves". In: *Annales de l'Institut Fourier* 35.1 (1985), pp. 117–123.

[27] S. W. Drury and B. P. Marshall. "Fourier restriction theorems for curves with affine and Euclidean arclengths". In: *Mathematical Proceedings of the Cambridge Philosophical Society* 97.1 (Jan. 1985), pp. 111–125. ISSN: 1469-8064, 0305-0041.

[28] S. W. Drury and B. P. Marshall. "Fourier restriction theorems for degenerate curves". In: *Mathematical Proceedings of the Cambridge Philosophical Society* 101.3 (May 1987), pp. 541–553. ISSN: 1469-8064, 0305-0041.

[29] Javier Duoandikoetxea. *Fourier analysis*. Vol. 29. Graduate studies in mathematics. American Mathematical Society, Providence, RI, 2001, pp. xviii+222.

[30] Semyon Dyatlov and Long Jin. "Resonances for open quantum maps and a fractal uncertainty principle". In: *Communications in Mathematical Physics* 354.1 (2017), pp. 269–316.

[31] William Fulton. *Young tableaux: With applications to representation theory and geometry*. London mathematical society student texts. Cambridge: Cambridge University Press, 1996.

[32] Philip T. Gressman, Shaoming Guo, Lillian B. Pierce, Jooris Roos, and Po Lam Yung. "Reversing a philosophy: From counting to square functions and decoupling". In: *Journal of Geometric Analysis* 31.12 (2021), pp. 7075–7095.

[33] Shaoming Guo, Zane Kun Li, and Po-Lam Yung. "A bilinear proof of decoupling for the cubic moment curve". In: *Transactions of the American Mathematical Society* 374.8 (2021), pp. 5405–5432.

[34] Shaoming Guo, Zane Kun Li, and Po-Lam Yung. "Improved discrete restriction for the parabola". 2021. arXiv: 2103.09795.

[35] Shaoming Guo, Zane Kun Li, Po-Lam Yung, and Pavel Zorin-Kranich. "A short proof of ℓ^2 decoupling for the moment curve". In: *American Journal of Mathematics* 143.6 (2021), pp. 1983–1998.

[36] Shaoming Guo, Zane Kun Li, Po-Lam Yung, and Pavel Zorin-Kranich. "A short proof of ℓ^2 decoupling for the moment curve". In: *American J. Math.* 143.6 (2021), pp. 1983–1998.

[37] Larry Guth. "18.118 Topics in Analysis: Decoupling, Lecture 2".

[38] Rui Han, Vjekoslav Kovač, Michael T. Lacey, José Madrid, and Fan Yang. "Improving Estimates for Discrete Polynomial Averages". In: *Journal of Fourier Analysis and Applications* 26.3 (May 28, 2020), p. 42. ISSN: 1531-5851.

[39] Kathryn E. Hare and Ivo Klemes. “On permutations of lacunary intervals”. In: *Transactions of the American Mathematical Society* 347.10 (1995), pp. 4105–4127.

[40] D. R. Heath-Brown. *The cubic case of Vinogradov's mean value theorem – a simplified approach to Wooley's “Efficient congruencing”*. 2015.

[41] Jonathan Hickman and James Wright. “The Fourier restriction and Kakeya problems over rings of integers modulo N”. In: *Discrete Analysis* (May 30, 2018).

[42] Jonathan Edward Hickman. “Topics in affine and discrete harmonic analysis”. In: (July 1, 2015).

[43] Luogeng Hua. *Additive Theory of Prime Numbers*. American Mathematical Soc., Dec. 4, 2009. 206 pp. ISBN: 978-0-8218-4942-2. Google Books: `LNeaAwAAQBAJ`.

[44] Kevin Hughes and Trevor D Wooley. “Discrete Restriction for (x, x^3) and Related Topics”. In: *International Mathematics Research Notices* 2022.20 (Oct. 16, 2022), pp. 15612–15631. ISSN: 1073-7928.

[45] Daniel Kane and Terence Tao. “A bound on partitioning clusters”. In: *The Electronic Journal of Combinatorics* 24.2 (2017), Paper P2.31.

[46] Vjekoslav Kovač. *On Binomial Sums, Additive Energies, and Lazy Random Walks*. Jan. 30, 2023. arXiv: `arXiv:2206.01591`. URL: `http://arxiv.org/abs/2206.01591` (visited on 04/12/2023). preprint.

[47] Izabella Łaba and Hong Wang. “Decoupling and near-optimal restriction estimates for Cantor sets”. In: *International Mathematics Research Notices. IMRN* 9 (2018), pp. 2944–2966.

[48] Xudong Lai and Yong Ding. “A note on the discrete Fourier restriction problem”. In: *Proceedings of the American Mathematical Society* 146.9 (May 24, 2018), pp. 3839–3846. ISSN: 0002-9939, 1088-6826.

[49] Jianhui Li and Tongou Yang. “Decoupling for mixed-homogeneous polynomials in $\mathbb{R}^3$”. In: *Mathematische Annalen* 383.3 (Aug. 1, 2022), pp. 1319–1351. ISSN: 1432-1807.

[50] Jianhui Li and Tongou Yang. *Decoupling for smooth surfaces in* $\mathbb{R}^3$. Oct. 15, 2021. arXiv: arXiv:2110.08441. URL: http://arxiv.org/abs/2110.08441 (visited on 04/09/2023). preprint.

[51] Zane Kun Li. "An ℓ^2 decoupling interpretation of efficient congruencing: the parabola". In: *Revista Matematica Iberoamericana* 37.5 (2021), pp. 1761–1802.

[52] Zane Kun Li. "An introduction to decoupling and harmonic analysis over $\mathbb{Q}_p$". 2022. arXiv: 2209.01644.

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

[54] Conor Meade. "Uniform Convolution and Fourier Restriction estimates for complex polynomial curves in $\mathbb{C}^3$". 2020. arXiv: 2012.09651.

[55] Akshat Mudgal. "Additive energies on spheres". In: *Journal of the London Mathematical Society* 106.4 (2022), pp. 2927–2958.

[56] Daniel M. Oberlin. "A Uniform Fourier Restriction Theorem for Surfaces in $\mathbb{R}^3$". In: *Proceedings of the American Mathematical Society* 132.4 (2004), pp. 1195–1199. ISSN: 0002-9939. JSTOR: 4097150.

[57] Daniel M. Oberlin. "Convolution Estimates for Some Measures on Curves". In: *Proceedings of the American Mathematical Society* 99.1 (1987), pp. 56–60. ISSN: 0002-9939.

[58] Daniel M. Oberlin. "Convolution with measures on hypersurfaces". In: *Mathematical Proceedings of the Cambridge Philosophical Society* 129.3 (Nov. 2000), pp. 517–526. ISSN: 1469-8064, 0305-0041.

[59] Lillian B. Pierce. "The Vinogradov mean value theorem [after Wooley, and Bourgain, Demeter and Guth]". In: *Astérisque Exposés Bourbaki* 407 (2019), pp. 479–564.

[60] Elena Prestini. “Restriction theorems for the Fourier transform to some manifolds in $\mathbb{R}^n$”. In: *Proc. Sympos. Pure Math.* Vol. 35. 1979, pp. 101–109.

[61] Yauhen Radyna. “Change of variable in integrals over p-adic and adelic domains”. In: *North-holland mathematics studies.* Vol. 197. Elsevier, 2004, pp. 267–272.

[62] Tomasz Schoen and Ilya Shkredov. “Higher moments of convolutions”. In: *Journal of Number Theory* 133.5 (2013), pp. 1693–1737.

[63] Ilya Shkredov. “Energies and structure of additive sets”. In: *The Electronic Journal of Combinatorics* 21.3 (2014), Paper P3.44.

[64] Peter Sjolin. “Fourier multipliers and estimates of the Fourier transform of measures carried by smooth curves in $\mathbb{R}^2$”. In: *Studia Mathematica* 51 (1974).

[65] John Steinig. “On some rules of Laguerre’s, and systems of equal sums of like powers”. In: *Rendiconti di Matematica. Serie VI.* 4 (1971), 629–644 (1972). ISSN: 0034-4427.

[66] Betsy Stovall. “Endpoint $L^p \to L^q$ bounds for integration along certain polynomial curves”. In: *Journal of Functional Analysis* 259.12 (2010), pp. 3205–3229.

[67] Betsy Stovall. “Endpoint bounds for a generalized Radon transform”. In: *Journal of the London Mathematical Society* 80.2 (2009), pp. 357–374. ISSN: 1469-7750.

[68] Betsy Stovall. “Uniform estimates for Fourier restriction to polynomial curves in $\mathbb{R}^d$”. In: *American Journal of Mathematics* 138.2 (2016), pp. 449–471. ISSN: 0002-9327.

[69] M. H. Taibleson. *Fourier Analysis on Local Fields.* Princeton University Press, Mar. 8, 2015. 308 pp. ISBN: 978-1-4008-7133-9.

[70] T. Tao. “Exploring the toolkit of jean bourgain”. In: *Bulletin of the American Mathematical Society* 58.2 (2021), pp. 155–171.

[71] Terence Tao. “247B, notes 2: Decoupling theory”. In: *What’s new blog, https://terrytao.wordpress.com notes-2-decoupling-theory/* ().

[72] Terence Tao and Van Vu. *Additive combinatorics*. Vol. 105. Cambridge studies in advanced mathematics. Cambridge University Press, Cambridge, 2006.

[73] Trevor D. Wooley. "Nested efficient congruencing and relatives of Vinogradov's mean value theorem". In: *Proceedings of the London Mathematical Society* 118.4 (2019), pp. 942–1016.

[74] James Wright. *A theory of complex oscillatory integrals: A case study*. Dec. 21, 2020. arXiv: `arXiv:2012.11256`. URL: `http://arxiv.org/abs/2012.11256` (visited on 04/09/2023). preprint.

[75] Tongou Yang. "Uniform l^2-Decoupling in $\mathbb{R}^2$ for Polynomials". In: *The Journal of Geometric Analysis* 31.11 (Nov. 1, 2021), pp. 10846–10867. ISSN: 1559-002X.