Around Smyth's Conjecture

Which Coefficients Arise in Linear Relations Among Galois Conjugates?

by

Will Hardt

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Mathematics

at the
University of Wisconsin-Madison
2023

Date of Final Oral Exam: August 18, 2023

The dissertation is approved by the following members of the Final Oral Committee:

Jordan Ellenberg, Professor, Mathematics
Betsy Stovall, Professor, Mathematics
Simon Marshall, Associate Professor, Mathematics
Jose Israel Rodriguez, Associate Professor, Mathematics

Around Smyth's Conjecture

A Study of Smyth's Conjecture over $\mathbb{Q}$ and a Proof of a Function Field Analogue

Will Hardt

Abstract In 1986, Smyth conjectured an elegant classification of the tuples of coefficients $(a_1, \dots a_k) \in \mathbb{Z}^k$ that appear in a linear relation $\sum_{i=1}^{k} a_i \gamma_i = 0$ among Galois conjugates $\gamma_1, \dots \gamma_k$ over $\mathbb{Q}$. Thirty-seven years later, the conjecture remains open with little direct progress made on it. This thesis compiles evidence in favor of the conjecture, including a proof of a function field analogue and a proposed number field generalization. Additionally, we establish a surprising connection between the conjecture and the recent notion of slice rank from additive combinatorics, via Strassen's asymptotic spectrum. We show that Smyth's Conjecture would be implied by certain "Smyth tensors" having full asymptotic slice rank, and we prove that the Smyth tensors do have full slice rank. We discuss the obstacle to extending this argument to asymptotic slice rank. Finally, we provide a counterexample to an old conjecture of Brualdi and Csima regarding the support patterns of stochastic tensors.

Dedication

To Anna, whose companionship and support I appreciate every day.

Acknowledgements

I am very grateful to my advisor, Jordan Ellenberg for his guidance throughout graduate school and the role he played in mentoring this project.

Chapter 2 of this thesis is joint work with John Yin, who was an excellent collaborator. I also had productive conversations with Philip Harris about Smyth's Conjecture.

I may never have pursued graduate studies in math without the influence of my brother Andy. I am deeply grateful for his encouragement over many years.

Thanks to Bobby Grizzard for introducing Smyth's Conjecture to the Wisconsin math department.

Finally, I'd like to thank my committee for their time and flexibility.

Contents

List of Notation

(a, b, c) (potential) Smyth triple

$(a_1, \dots a_k)$ (potential) Smyth tuple, where $k \geq 3$ is not necessarily fixed

$[r, s]$ $\{r, r+1, \dots, s\}$

$[s]$ $\{1, 2, \dots, s\}$

Φ_{abc} The set $\{(x, y, z) \in S^3 : ax + by + ca = 0\}$

$\mathbb{S}^{\mathbb{F}}_{a,b,c,N}$ nonnegative tensors indexed by $[-n, n]$, with entries in the field $\mathbb{F}$, and whose support is contained in $\{(x, y, z) : ax + by + cz = 0\}$

supp(T) {(i,j,k): t$_{ijk} \neq 0$}

Ten(m, m, m) $\mathbb{R}^m \otimes \mathbb{R}^m \otimes \mathbb{R}^m$

d^B_A The density of A as a subset of B

T_{abc} The 0-1 tensor with supp$(T) = \Phi_{abc}$

[m] The set $\{1, 2, \dots, m\}$; equivalent to the notation $[1, m]$

I$_1 \times I_2 \times I_3$ The index set of a 3-tensor

N Equal to 2n+1, the size of S

n Equal to the product abc

S The integer interval [-n,n]

SR(T) The slice-rank of T

Chapter 1

An Introduction to Smyth's Conjecture

1.1 The Conjecture and an Equivalence

In 1986, Smyth asked which coefficients $(a_1, ..., a_k) \in \mathbb{Q}^k$ arise in linear relations $\sum_{i=1}^{k} a_i\gamma_i = 0$ among Galois conjugates $\gamma_1, ..., \gamma_k$ algebraic over $\mathbb{Q}$ [Smy86]. We will call such an $(a_1, \ldots a_k)$ a *Smyth tuple.* Notice that we may normalize this problem by assuming that $(a_1, \ldots, a_k)$ is a coprime integer tuple, because for any nonzero $\alpha \in \mathbb{Q}$, we have $\sum_{i=1}^{k} a_i\gamma_i = 0 \iff \sum_{i=1}^{k}(\alpha a_i)\gamma_i = 0$; thus $(a_1, \ldots a_k)$ is a Smyth tuple if and only if $(\alpha a_1, \ldots \alpha a_k)$ is.

Smyth proved that the following conditions on $(a_1, \ldots, a_k)$ are necessary for being a Smyth tuple, and conjectured that they are jointly sufficient.

Definition 1.1.1. *A coprime integer tuple $(a_1, \ldots, a_k) \in \mathbb{Z}^k$ is said to satisfy the absolute value criteria if*

1. *$|a_i| \leq \sum_{j \neq i} |a_j|$ for all i, where $|\cdot|$ denotes the usual archimedean absolute value, and*
2. *Every prime p divides at most $k-2$ of the a_i.*

We will call $(a_1, \ldots, a_k)$ satisfying the absolute value criteria *potential Smyth tuples.*

Theorem 1.1.2. *[Smy86] If a coprime integer tuple $(a_1, ..., a_k) \in \mathbb{Z}^k$ is a Smyth tuple, then it satisfies the absolute value criteria.*

Conjecture 1.1.3 (Smyth's Conjecture)**.** *[Smy86] If an integer tuple $(a_1, ..., a_k) \in \mathbb{Z}^n$ satisfies the absolute value criteria, then it is a Smyth tuple.*

Smyth also proved another characterization of Smyth tuples involving multisets of solutions to the integral equations $\sum_{i=1}^k a_i x_i = 0$.

Definition 1.1.4. *Let $(a_1, \ldots a_k) \in \mathbb{Z}^k$ be a coprime integer tuple. A balanced multiset with respect to $(a_1, \ldots, a_k)$ is a finite multiset $\{(x_{i1}, ..., x_{ik}) \in \mathbb{Z}^k\}_{i=1}^m$ of integral solutions to the equation $\sum_{i=1}^k a_i x_i = 0$ such that the multiset $\{x_{ij} : i = 1, \ldots, m\}$ is independent of j.*

Theorem 1.1.5. *[Smy86] A coprime integer tuple $(a_1, ..., a_k) \in \mathbb{Z}^k$ is a Smyth tuple if and only if there exists a balanced multiset with respect to $(a_1, ..., a_k)$.*

We illustrate the notion of balanced multisets with an example.

Example 1.1.6. *Consider the coprime integer triple $(a, b, c) = (3, 4, 5)$. Consider the 8-element (multi)set of solutions to the equation $3x + 4y + 5z = 0$*

$$S = \left\{ \begin{array}{ccc} x & y & z \\ (3, & 4, & -5), \\ (-3, & -4, & 5), \\ (4, & -3, & 0), \\ (-4, & 3, & 0), \\ (5, & 0, & -3), \\ (-5, & 0, & 3), \\ (0, & 5, & -4), \\ (0, & -5, & 4) \end{array} \right\}$$

Notice that the multiset of entries in each of the $x, y,$ and z columns is the same: $\{\pm 3, \pm 4, \pm 5, 0, 0\}$.

Thus S is a balanced set with respect to $(3,4,5)$, and it follows from Theorem 1.1.5 that $(3,4,5)$ is a Smyth triple.

Thus one can approach Smyth's Conjecture as follows; for each (a,b,c) satisfying the absolute value criteria, fix an integer $n := n(a,b,c)$, and then for each triple $(x,y,z) \in [-n,n] \cap \mathbb{Z}^3$ such that $ax+by+cz=0$, prescribe how many times to include (x,y,z) in a multiset. This approach is successful if the resulting multisets are balanced in the sense of Definition 1.1.4.

1.2 A Heuristic

In this section, we give a heuristic that predicts the existence of balanced sets (in fact, 1-factors) with high probability, regardless of the potential Smyth tuple. The heuristic is sharp enough to not falsely predict abundant Smyth pairs (i.e. Smyth tuples when $k=2$) but is blunt enough to not see the necessity of the local conditions for $k \geq 3$.

To set up the heuristic, we need a slight change of perspective, which we will describe after the following proposition.

Proposition 1.2.1. *Let K be any global field. The following are equivalent for coprime $(a_1,\ldots,a_k) \in \mathbb{Z}^k$.*

1. *$(a_1,\ldots,a_k)$ is a Smyth tuple.*
2. *There exists a balanced multiset of tuples with respect to $(a_1,\ldots,a_k)$.*
3. *There exist permutation matrices $X_1,\ldots,X_k$ such that $\det(\sum_{i=1}^k a_iX_i)=0$.*

Proof. Smyth [Smy86, Thm. 2] proved (1) $\iff$ (2); this was our Theorem 1.1.5.

(2) $\implies$ (3): Let $T=\{(x_{i1},x_{i2},\ldots,x_{ik})\}_{i=1}^m$ be a balanced multiset of tuples of size m. For $j: 1 \leq j \leq k$, let $v_j=(x_{ij})_{i=1}^m$ be the vector in $\mathbb{Q}^m$ obtained by taking the j^{th} entry from each tuple in T. By definition of balanced multiset, there exist (not necessarily unique) $m \times m$ permutation matrices $X_1,\ldots,X_k$ such that $X_iv_k=v_i$ for all i. Thus $(\sum_{i=1}^k a_iX_i)v_k = \sum_{i=1}^k a_iv_i = 0$, so $\sum_{i=1}^k a_iX_i$ has nontrivial kernel.

(3) $\implies$ (2): Reverse the previous argument as follows. Let v_k be any nonzero vector in the kernel of $\sum_{i=1}^k a_i X_i$ and let $v_i := X_i v_k$ for $1 \leq i \leq k$. Then the coordinates of the vectors $v_1, \ldots, v_k$ give a balanced multiset of tuples as above. $\square$

The setting of Smyth's Conjecture involves fixing a tuple $(a_1, \ldots a_k)$ and asking for a balanced multiset $T = \{(x_{i1}, \ldots x_{ik}) \in \mathbb{Z}^k\}_{i=1}^m$ with respect to the tuple. The heuristic involves fixing a particular vector $v_n \in \mathbb{R}^{2n+1}$ for each positive integer n, and asking for the probability, under a certain assumption of randomness, that there exist $X_1, \ldots X_k \in S_{2n+1}$ such that $(\sum_{i=1}^k a_i X_i) v_n = 0$.

Heuristic 1.2.2. *Fix a positive integer n and the vector $v_n = (-n, -n+1, \ldots 0, \ldots, n-1, n)^T$. Let $N = 2n+1$ and let $X_k = I$ be the $N \times N$ identity matrix. Then choose random permutations $X_1, \ldots, X_{k-1} \in G \subset S_N$ and assume that for each $j : 1 \leq j \leq q^N$, the sum $\sum_{i=1}^n a_i v_{X_i^{-1}(j)}$ takes values in [-An, An] uniformly and independently at random, where $A := \sum_{i=1}^k a_i$.*

The result of this heuristic is as follows. For a fixed coprime tuple $(a_1, \ldots a_k) \in \mathbb{Z}^k$ and a fixed n, the probability of there *not* existing permutation matrices $X_1, \ldots X_{k-1}$ such that $(\sum_{i=1}^k a_i X_i) v_n = 0$ is

$$\left(1 - \frac{1}{2An+1}\right)^{(k-1)N!}$$

This probabillity goes to zero as $n \to \infty$, hence this heuristic says that with high probability, $(a_1, \ldots, a_k)$ has a 1-factor.

As mentioned, a weakness of this heuristic is that it does not notice the necessary local conditions on the a_i.

1.3 The Tensor Perspective

Before proceeding, we need some additional notation. Given integers $r < s$, we will write $[r, s]$ for the set of integers $\{r, r+1, \ldots, s\}$, and when $r = 1$, we will just write $[s]$.

This thesis will be concerned almost exclusively with Smyth's Conjecture for triples, i.e. the case where $k = 3$. Suppose that we've fixed a potential Smyth triple (a, b, c) and a positive integer n, and we're looking for a balanced multiset $\{(x_i, y_i, z_i) \in [-n, n]^3\}_{i=1}^m$ with respect to (a, b, c).

A multiset of solutions (x, y, z) to $ax + by + cz = 0$ can be conveniently represented as a 3d-array $T = [t_{ijk}]_{i,j,k=-n}^n$, where t_{ijk} represents the number of times that the triple (i, j, k) occurs in our multiset. This means in particular that the support of T (the set of (i, j, k) for which $t_{ijk} \neq 0$) must be contained in the zero locus of $ai + by + ck = 0$.

It will eventually be useful to view T not just as a 3d array, but as a tensor sitting in the vector space $V = \mathbb{R}^{2n+1} \otimes \mathbb{R}^{2n+1} \otimes \mathbb{R}^{2n+1}$. Specifically we identify T with the tensor that shares the coefficients of T in the standard basis for V, i.e. $T = \sum_{i,j,k=-n}^n t_{ijk} e_i \otimes e_j \otimes e_k$. Viewing T as a tensor will allow us to leverage algebraic tools, specifically Strassen's asymptotic spectrum and slice rank, to understand our combinatorial questions about T.

The "balanced" property of a balanced multiset manifests in the corresponding tensor as follows. A tensor T is *balanced* if its vector of co-dimension one slices is the same in each direction, that is if for all $\alpha \in [-n, n]$, we have

$$\sum_{i,j=-n}^{n} t_{i,j,\alpha} = \sum_{i,k=-n}^{n} t_{i,\alpha,k} = \sum_{j,k=-n}^{n} t_{\alpha,j,k} \tag{1.3.1}$$

If all of these sums are equal to 1, then we will say that T is *stochastic.*

Thus Smyth's Conjecture (for triples) is equivalent to the following.

Conjecture 1.3.1. *Given any potential Smyth triple* $(a, b, c) \in \mathbb{Z}^3$*, there exists a positive integer* n *and a nonnegative integer tensor* T *indexed by* $[-n, n]^3$ *whose support* $\operatorname{supp}(T)$ *is contained in* $\{(x, y, z) : ax + by + cz = 0\}$*, and whose entries satisfy 1.3.1.*

Notice that the existence of such tensors would be an immediate consequence of elementary linear algebra without the nonnegativity requirement on the entries of T. We can, however, relax the requirement that the entries are integers, as we now show. We will write $\mathbb{S}_{a,b,c,N}^{\mathbb{F}}$ to denote the set of nonnegative tensors indexed by $[-n, n]^3$, with entries in the field $\mathbb{F}$, and whose support is contained in $\{(x, y, z) : ax + by + cz = 0\}$.

Lemma 1.3.2. *Let m, n be positive integers and $V = \ker(A) \subset \mathbb{R}^n$ where $A = (a_{ij}) \in \mathbb{Q}^{m\times n}$. If $V \cap (\mathbb{R}^{\geq 0})^n \neq 0$, then $V \cap (\mathbb{Q}^{\geq 0})^n \neq 0$.*

Proof. Fix a basis $B = \{b_i\}_{i\in I}$ for $\mathbb{R}$ as a $\mathbb{Q}$-vector space. Let $0 \neq (r_1, \dots, r_n) \in V \cap (\mathbb{R}^{\geq 0})^n$. Then we can write

$$r_j = \sum_{i\in I_j} c_{ij} b_{ij}$$

for finite subsets $I_j \subset I$, where $b_{ij} \in B$ and $c_{ij} \in \mathbb{Q}$.

Now re-label the b_{ij} as $b_1, b_2, \dots, b_N$; then there exist $k_{ij} \in \mathbb{Q}$ (each one equal to some $c_{i'j'}$ or to 0) such that

$$a_{j1}(\sum_{i=1}^{N} k_{ij} b_i) + \ldots + a_{jn}(\sum_{i=1}^{N} k_{ij} b_i) = 0, \text{ for } j = 1, 2, \dots, m$$

Now fix an index i so that $k_{ij} \neq 0$ for some j. By linear independence of the b's, and the fact that $b_i \neq 0$, we have

$$a_{j1}k_{i1} + a_{j2}k_{i2} + \ldots + a_{jn}k_{in} = 0 \text{ for } j = 1, 2, \dots, m$$

Thus, $0 \neq (k_{i1} \dots k_{in})^T \in \ker(A) \cap (\mathbb{Q}^{\geq 0})^n$, as desired. $\square$

We are now ready to show that Smyth's Conjecture can be formulated in terms of the existence of real balanced tensors.

Proposition 1.3.3. *Let (a, b, c) be a potential Smyth triple and $N \in \mathbb{Z}^+$. If $\mathbb{S}^{\mathbb{R}}_{a,b,c;N} \neq \emptyset$ then $\mathbb{S}^{\mathbb{Q}}_{a,b,c;N} \neq \emptyset$.*

Proof. Note that a real (resp. rational) nonnegative $N \times N \times N$ tensor T with support contained in $\{(x, y, z) : ax + by + cz = 0\}$ lies in $\mathbb{S}^{\mathbb{R}}_{a,b,c;N}$ (resp. $\mathbb{S}^{\mathbb{Q}}_{a,b,c;N}$) if and only if its entries satisfy a system of linear equations dictating that it is balanced. We will call the corresponding coefficient matrix $A^{\mathbb{R}}_{a,b,c;N}$ (resp. $A^{\mathbb{Q}}_{a,b,c;N}$). So if we vectorize T as $\vec{t}$ in an appropriate way, the balanced condition becomes $A^{\mathbb{R}}_{a,b,c;N}\vec{t} = 0$ (resp. $A^{\mathbb{Q}}_{a,b,c;N}\vec{t} = 0$). Note that the entries of $A^{\mathbb{R}}_{a,b,c;N}$ and $A^{\mathbb{Q}}_{a,b,c;N}$ are exactly the same; we write superscripts here only to indicate the field where the entries of $\vec{t}$ must lie.

It now follows from Lemma 1.3.2 that there exists a tensor in $\mathbb{S}^{\mathbb{R}}_{a,b,c;N}$ if and only if there exists one in $\mathbb{S}^{\mathbb{Q}}_{a,b,c;N}$. □

The shift to constructing real balanced tensors opens up some new avenues. In particular, it would be enough to find a sequence of nonnegative tensors with rational entries which, in the limit, is balanced.

1.4 Related Work

The question of how prevalent linear relations among Galois conjugates are has been studied from multiple angles. In [Ber+04, Thm. 14(ii)] it is shown that for any global field K of characteristic not equal to 2 and all but finitely many nonnegative integers n, there exists $\alpha \in \overline{K}$ of degree $2^n n!$ whose conjugates span a vector space of dimension n. In these cases, the dimension of relations between conjugates is $2^n n! - n$, and so in this sense, linear relations among conjugates are plentiful.

On the other hand, there are results constraining the supply of linear relations among Galois conjugates. Dixon [Dix97, Thm. 1'] showed that if K is any subfield of $\mathbb{C}$ (e.g. a number field) and $f(x) \in K[x]$ is an irreducible polynomial whose Galois group acts 2-transitively on its set of roots, then there are no nontrivial K-linear relations among the roots of $f(x)$.

1.5 The Structure of this Thesis

The structure of this thesis is as follows. In Chapter 1, we've introduced the statement of Smyth's Conjecture and seen equivalent formulations in terms of balanced multisets and balanced tensors. We also introduced a heuristic and explained how it predicts a positive resolution to Smyth's Conjecture.

Chapter 2, which has been published jointly with John Yin [**HY21**], examines analogues of

Smyth's Conjecture. In particular we prove a function field analogue and formulate a number field analogue, which is not a straightforward generalization due to a subtlety occurring at the Archimedean places.

Chapter 3 explores the surprising connection between Smyth's Conjecture and the slice rank of tensors. The chapter begins with an introduction to stochastic patterns, Strassen's asymptotic spectrum, slice rank, and related concepts. Next we prove that Smyth's Conjecture would be implied by certain "Smyth tensors" having full asymptotic slice rank. We go on to prove that these tensors have full (non-asymptotic) slice rank. In the penultimate section of the chapter, we discuss the difficulties of extending this approach to show full asymptotic slice rank, which leads us into a high-dimensional sumset problem. Finally, we provide a computer-verified counterexample to an old conjecture of Brualdi and Csima regarding stochastic patterns.

Chapter 2

Analogues of Smyth's Conjecture

2.1 A Proof of a Function Field Analogue

Global fields are central objects in number theory, and come in two varieties – number fields and function fields. There is a strong parallel between number fields and function fields, and in particular between $\mathbb{Q}$ and $\mathbb{F}_q(t)$. In each field, the ring of integers is a Principal Ideal Domain, and many important number theoretic theorems over $\mathbb{Q}$ have analogies that are also true in $\mathbb{F}_q(t)$. For details of this analogy, see [Ell14] and [Poo06].

Despite these similarities, the relationship between $\mathbb{Q}$ and $\mathbb{F}_q(t)$ is somewhat asymmetrical; while $\mathbb{F}_q(t)$ carries a similar structure to $\mathbb{Q}$, it often exhibits less complexity. Most critically for our purposes, the "scales" in $\mathbb{F}_q(t)$ are well preserved by addition (the sum of two polynomials of degree $\leq d$ also has degree $\leq d$), but there is no such separation of scales in $\mathbb{Q}$.

With this context in mind, we now turn toward the proof of the main result of this section, Theorem 2.1.2, which will follow essentially as a corollary from Proposition 2.1.3.

First, we lay out some more conventions. The *height* of a coprime tuple $(a_1, \ldots, a_k) \in \mathbb{F}_q[t]^k$ is $\max_i \deg(a_i)$. Let $V_D \subset \mathbb{F}_q[t]$ be the set of polynomials of degree $< D$ for any positive integer m.

Throughout this thesis, we will use the standard normalizations of absolute values over global fields. Namely, the absolute values over $\mathbb{Q}$ are given by the usual archimedean one $|\cdot|$ and

$|a|_p = p^{-\mathrm{ord}_p(a)}$ for each positive prime p. For the function field $\mathbb{F}_q(t)$ and $g \in \mathbb{F}_q(t)$, we take the normalizations $|g|_f = q^{-\deg(f)\mathrm{ord}_f(g)}$ for irreducible polynomials $f \in \mathbb{F}_q(t)$ and $|g|_\infty = q^{\deg(g)}$. These normalizations extend uniquely to any global field.

To set up the function field analogue we will prove, we recall the absolute value criteria,

A natural way to generalize (1) and (2) to any global field K is as follows.

Definition 2.1.1. *A tuple $(a_1, \ldots, a_k) \in K^k$ satisfies the absolute value criteria over K if*

(1') For any archimedean absolute value $|\cdot|$ of K, we have $|a_i| \leq \sum_{j \neq i} |a_j|$ for all i.

(2') For any nonarchimedean absolute value $|\cdot|$ of K, we have $|a_i| \leq \max_{j \neq i} |a_j|$ for all i.

When the field K is clear from context we may omit it from our terminology.

Thus, the most natural function field analogue of Smyth's Conjecture is:

Theorem 2.1.2 (Smyth's Conjecture Over $\mathbb{F}_q(t)$)**.** *Let $k \geq 3$ be an integer. A coprime tuple $(a_1, \ldots, a_k) \in \mathbb{F}_q[t]^k$ is a Smyth tuple if and only if $(a_1, \ldots, a_k)$ satisfies the absolute value criteria over $\mathbb{F}_q(t)$.*

We will show that, surprisingly, for any coprime $(a_1, \ldots, a_k) \in \mathbb{F}_q[t]^k$ satisfying the absolute value criteria and any $D \geq d := \mathrm{height}(a_1, \ldots, a_k)$, the set of *all* solutions $(x_1, \ldots, x_k) \in V_D^k$ to the equation $\sum_{i=1}^k a_i x_i = 0$ is a balanced set.

Proposition 2.1.3. *Let $k \geq 3$ be an integer. Let $(a_1, \ldots, a_k) \in \mathbb{F}_q[t]^k$ of height d be a coprime tuple satisfying the absolute value criteria. Let $D \geq d$ be an integer and let j be an integer so that $1 \leq j \leq k$. Fix $x_j \in V_D$. Then the number of tuples $(x_1, \ldots, x_{j-1}, x_{j+1}, \ldots x_k) \in V_D^{k-1}$ satisfying $\sum_{i=1}^k a_i x_i = 0$ is $q^{D(k-2)-d}$. In particular, this count does not depend on j.*

Proof. Without loss of generality, we let $j = 1$. By the absolute value criteria, the maximum degree of $a_1, \ldots, a_k$ is achieved at least twice. Hence, some a_i with $i \neq 1$ has degree d; without loss of generality, assume that a_k does. Let $c = a_1 x_1$. Define

$$S = \{(x_2, \ldots, x_k) \in V_D^{k-1} : c + \sum_{i=2}^k a_i x_i = 0\}.$$

Our goal is to compute $\#S$. To do so, we will project onto $\mathbb{F}_q[t]/a_k$, so we define

$$\overline{S} = \{(\overline{x_2}, \ldots, \overline{x_{k-1}}) \in (\mathbb{F}_q[t]/a_k)^{k-2} : \overline{c} + \sum_{i=2}^{k-1} \overline{a_i x_i} = 0\}.$$

Reducing modulo a_k in each coordinate and throwing out the last coordinate gives a surjective $q^{(D-d)(k-2)}$-to-1 map $S \to \overline{S}$; the pre-image of any $(\overline{x}_2, \ldots, \overline{x}_{k-1}) \in \overline{S}$ is

$$\left\{ \left(x_2 + h_2 a_k, \ldots, x_{k-1} + h_{k-1} a_k, -\left(\frac{c + \sum_{i=2}^{k-1} a_i x_i}{a_k} + \sum_{i=2}^{k-1} a_i h_i \right) \right) : h_i \in V_{D-d} \right\},$$

where x_i is the unique polynomial of degree $< d$ equal to $\overline{x}_i \bmod a_k$. Thus, we have $\#S = q^{(D-d)(k-2)} \#\overline{S}$.

So we want to count the number of solutions $(\overline{x}_2, \ldots \overline{x}_{k-1}) \in (\mathbb{F}_q[t]/(a_k))^{k-2}$ to $\overline{c} + \sum_{i=2}^{k-1} \overline{a_i x_i} = 0$. Let $a_k = \prod p_j^{e_j}$ be the prime factorization of a_k. Let $R_j := \mathbb{F}_q[t]/(p_j^{e_j})$; by the Chinese Remainder Theorem, it will suffice to count the number of solutions in R_j for each j.

Specifically, let $\overline{S}_j = \{(\overline{x_2}, \ldots, \overline{x}_{k-1}) \in R_j^{k-2} : \overline{c} + \sum_{i=2}^{k-1} \overline{a_i x_i} = 0\}$; then the Chinese Remainder Theorem implies that $\#\overline{S} = \prod_j \#\overline{S}_j$.

We now compute $\#\overline{S}_j$. Recall that by the absolute value criteria, for all j, there are at least two a_i that are not divisible by p_j. Of course a_k is divisible by all p_j; hence, for all j, there is at least one a_i, with $1 < i < k$, such that $p_j \nmid a_i$, in which case $\overline{a_i}$ is a unit in R_j. Thus, we can write $\overline{x_i} = \frac{\overline{c} + \sum_{\ell \neq i, \ell \neq 1} \overline{a_\ell x_\ell}}{\overline{a_i}}$, and so any collection of choices of $\overline{x_\ell}$ for $\ell \in \{2, \ldots, k-1\} \setminus \{i\}$, will give a unique choice of $\overline{x_i}$. There are $\#R_j = q^{\deg(p_j^{e_j})}$ choices for each $\overline{x}_\ell$, so $\#\overline{S}_j = q^{\deg(p_j^{e_j})(k-3)}$. Thus, since $\sum_j \deg(p_j^{e_j}) = \deg(a_k) = d$, we have

$$\#S = q^{(D-d)(k-2)} \prod_j q^{\deg(p_j^{e_j})(k-3)} = q^{(D-d)(k-2)} q^{d(k-3)} = q^{D(k-2)-d}, \text{ as desired.}$$

$\square$

Now Theorem 2.1.2 follows easily.

Proof of Theorem 2.1.2. ($\Rightarrow$): Smyth proved this statement over $\mathbb{Q}$ [Smy86, Cor. 2] (and we recorded this as Theorem 1.1.2), and his proof is valid over any global field K.

($\Leftarrow$): Without loss of generality, we assume $(a_1, \ldots, a_k)$ is a coprime tuple in $\mathbb{F}_q[t]^k$. Let T_D be the set of all tuples $(x_1, \ldots, x_k) \in V_D^k$ satisfying $\sum_{i=1}^k a_i x_i = 0$ and enumerate $T_D = \{(x_{i1}, \ldots, x_{ik})\}_{i=1}^t$ where $t = |T_D|$. In Proposition 2.1.3 we showed that for every $x \in V_D$ and all $i : 1 \le i \le n$, the number of tuples $(x_1, \ldots, x_k)$ in T_D with $x_i = x$ is $q^{D(k-2)-d}$. This means that for each $j : 1 \le j \le k$, the multiset $\{(x_{ij})\}_{i=1}^t$ is precisely $q^{D(k-2)-d}$ copies of V_D. Thus, T_D is a balanced (multi)set of tuples. So by Proposition 1.2.1, $(a_1, \ldots, a_k)$ is a Smyth tuple. $\square$

Remark 2.1.4. *By setting $D = d$ and $k = 3$ in Proposition 2.1.3, we see that if (a, b, c) is a Smyth triple, then T_d is a "1-factor," to borrow a term from (hyper)graph theory. That is, if one considers the hypergraph $H = (V, E)$ with $V = [-n, n]^3$ and edges $E = \{(x, y, z) : ax + by + cz = 0\}$, then T_d is a 1-regular subgraph of H, as each integer in $[-n, n]$ appears in each of the x,y, and z positions in T_d exactly once.*

2.2 The General Number Field Case

Recall that in any global field, the absolute value criteria are necessary conditions for being a Smyth tuple [Smy86, Cor. 2]. We showed in Theorem 2.1.2 that these criteria are sufficient for being a Smyth tuple over $\mathbb{F}_q(t)$, and Smyth conjectured the same over $\mathbb{Q}$ (Conjecture 1.1.3).

However, an example presented by David Speyer [Spe] in a MathOverflow post shows that the absolute value criteria are not sufficient for being a Smyth tuple in a general number field. In particular, the triple $(1, 1, \frac{1+\sqrt{-15}}{2})$ satisfies the absolute value criteria, but is not a Smyth triple. Note that this triple achieves equality in the archimedean absolute value inequalities.

Speyer showed in the same post that for triples of the form $(1, 1, a_3)$, if one amends the absolute value criteria to be strict inequalities for the archimedean absolute values, then they become a sufficient condition for being a Smyth triple. On the other hand, examples such as $(2, 3, -5)$ show that we cannot simply amend the archimedean absolute value criteria to be strict inequalities, as

$(2,3,-5)$ trivially *is* a Smyth triple. Instead, if some analogue of Smyth's Conjecture is true in number fields, it must be a little more sensitive to the cases in which there is equality in one of the archimedean absolute value criteria.

In order to formulate what we think the right conjecture is, we define the *strong absolute value criteria over a number field K* as follows.

(1") For any archimedean absolute value $|\cdot|$ of K, we have $|a_i| < \sum_{j\neq i} |a_j|$ for all i.

(2") For any nonarchimedean absolute value $|\cdot|$ of K, we have $|a_i| \leq \max_{j\neq i} |a_j|$ for all i.

The strong absolute value criteria are obtained from the absolute value criteria by making the archimedean inequalities strict.

We are now ready to formulate our generalization of Conjecture 1.1.3.

Conjecture 2.2.1 (Smyth's Conjecture over Number Fields)**.** *Let K be a number field and $\mathcal{O}_K$ its ring of integers. Then $(a_1,\ldots,a_k) \in \mathcal{O}_K^k$ is a Smyth tuple if and only if $(a_1,\ldots,a_k)$ satisfy the strong absolute value criteria over K or there exist roots of unity $\omega_1,\ldots,\omega_k$ in some extension of K such that $\sum_{i=1}^k a_i\omega_i = 0$.*

Remark 2.2.2. *The $K=\mathbb{Q}$ case of Conjecture 2.2.1 is equivalent to Conjecture 1.1.3.*

We will show in Proposition 2.2.6 that Conjecture 2.2.1 correctly deals with the tuples in which equality is achieved in one of the archimedean absolute value criteria. In particular, if $(a_1,\ldots,a_k)$ is a tuple such that equality holds in one of the archimedean absolute value criteria, then any tuple in a balanced multiset with respect to $(a_1,\ldots,a_k)$, if one exists, is a scalar multiple of a tuple of roots of unity.

But first we need two lemmas, the first of which shows that the property of being a Smyth tuple is preserved by multiplying the coordinates by (possibly different) roots of unity.

Lemma 2.2.3. *Let K be a number field and $\mathcal{O}_K$ its ring of integers. Let $(a_1,\ldots,a_k) \in \mathcal{O}_K^k$. If $\omega_1,\ldots,\omega_k$ are roots of unity in some extension of K and $(a_1,\ldots,a_k)$ is a Smyth tuple, then $(\omega_1 a_1,\ldots,\omega_k a_k)$ is a Smyth tuple in $\mathcal{O}_{K(\omega_1,\ldots,\omega_k)}^k$.*

Proof. Without loss of generality we may assume that $\omega_2 = \cdots = \omega_k = 1$, as we can make the following argument about each coordinate in turn. Denote $\omega := \omega_1$ and $L := K(\omega)$. Suppose that $\omega^r = 1$.

Let $\{(x_{i_1}, \ldots, x_{i_k}) \in K^k\}_{i=1}^r$ be a balanced multiset with respect to $(a_1, \ldots, a_k)$. Then $\bigcup_{j=0}^{m-1} \{(\omega^{j-1} x_{i_1}, \omega^j x_{i_2}, \ldots, \omega^j x_{i_k}) \in L^k\}_{i=1}^m$ is a balanced multiset with respect to $(\omega a_1, a_2, \ldots, a_k)$. □

Remark 2.2.4. In particular, Lemma 2.2.3 shows that if there are roots of unity $\omega_1, \ldots, \omega_k$ such that $\sum_{i=1}^k a_i \omega_i = 0$, then $(a_1, \ldots, a_k)$ is a Smyth tuple. Linear relations among roots of unity are a well-studied topic, going back at least to the 1960s. There are several results constraining the prevalence of such relations, indicating that such coefficients represent quite a small subset of Smyth tuples. A survey of some of these results is given in [Zan95]. For instance, when $a_1, \ldots, a_k$ are rational, a result of Mann [Man65] gives an explicit upper bound depending only on n for the order of the roots of unity ω_i occurring in a minimal relation $\sum_{i=1}^k a_i \omega_i = 0$. (Here minimality means that no nonempty proper sub-sum vanishes, and that the equation is normalized so that $\omega_1 = 1$.)

Lemma 2.2.5. *Let K be a number field and $\mathcal{O}_K$ its ring of integers. Let $(a_1, \ldots, a_k) \in \mathcal{O}_K^k$. Suppose that there exists an archimedean absolute value $|\cdot|_\nu$ of K and some i for which $|a_i|_\nu = \sum_{j \neq i} |a_j|_\nu$. If there exists a balanced multiset with respect to $(a_1, \ldots, a_k)$, then there exists a balanced multiset $\{(y_{i1}, \ldots, y_{ik}) \in K^k\}_{i=1}^m$ with respect to $(a_1, \ldots, a_k)$ whose coordinates y_{ij} all satisfy $|y_{ij}|_\nu = 1$.*

Proof. Without loss of generality assume that $|a_1|_\nu = \sum_{j>1} |a_j|_\nu$. Let $S = \{(x_{i1}, \ldots, x_{ik}) \in K^k\}_{i=1}^m$ be a balanced multiset with respect to $(a_1, \ldots, a_k)$. Let $X = \{x_{ij} : 1 \leq i \leq m, 1 \leq j \leq k\}$ be the set of all coordinates appearing in S. Write $M = \max_{x \in X} |x|_\nu$. Any reference to "absolute value" in this proof refers to $|\cdot|_\nu$.

We claim that if a tuple in S has a coordinate of absolute value M, then all coordinates of that tuple have absolute value M. To see this, first suppose that $|x_{i_0 1}| = M$ for some i_0. Along with the assumptions that $\sum_{j=1}^k a_j x_{i_0 j} = 0$ and $|a_1|_\nu = \sum_{j>2} |a_j|_\nu$, this implies that $|x_{i_0 j}|_\nu = M$ for all

$j = 1, \dots, n$. What we've shown so far is that if the first coordinate in a tuple in S has absolute value M, then all coordinates in that tuple do.

But S is balanced, which means that the multiset of first coordinates is the same as the multiset of j^{th} coordinates for every $j = 1, 2, \dots, k$. In particular, each of these multisets has the same number of elements of absolute value M, with the same multiplicities. Therefore coordinates of absolute value M can only occur in tuples whose first coordinate has absolute value M, and the claim is proved.

Thus the tuples whose coordinates have absolute value M form a balanced sub-multiset of S, and dividing all of these coordinates by an element of K of absolute value M, we obtain the desired balanced multiset. $\square$

Proposition 2.2.6. *Let K be a number field and $\mathcal{O}_K$ its ring of integers. Let $(a_1, \dots, a_k) \in \mathcal{O}_K^k$. Suppose that there exists an archimedean absolute value $|\cdot|_\nu$ of K and some i for which $|a_i|_\nu = \sum_{j \neq i} |a_j|_\nu$. Then $(a_1, \dots, a_k)$ is a Smyth tuple if and only if there exist roots of unity $\omega_1, \dots, \omega_k$ (not necessarily in K) such that $\sum_{i=1}^k a_i \omega_i = 0$.*

Proof. ($\Leftarrow$): By assumption, $(\omega_1 a_1, \dots, \omega_k a_k)$ is a Smyth tuple. The result now follows from Lemma 2.2.3.

($\Rightarrow$): Let $\phi : K \hookrightarrow \mathbb{C}$ be an embedding corresponding to the archimedean absolute value $|\cdot|_\nu$ and let $\psi : \overline{\mathbb{Q}} \hookrightarrow \mathbb{C}$ be an embedding of the algebraic closure of $\mathbb{Q}$ which extends ϕ. We will write $|\cdot|$ for the standard absolute value of complex numbers. Without loss of generality assume that $|\phi(a_1)| = \sum_{j>1} |\phi(a_j)|$.

By Lemma 2.2.5, there exists a balanced multiset $S = \{(x_{i1}, \dots, x_{ik}) \in K^k\}_{i=1}^k$ with respect to $(a_1, \dots, a_k)$ such that all $|x_{ij}|_\nu = 1$. By definition of balanced multiset, we have

$$\sum_{j=1}^k a_j x_{ij} = 0. \tag{2.2.1}$$

Now (2.2.1) along with $|\phi(a_1)| = \sum_{j>1} |\phi(a_j)|$ and the assumption that $|\phi(x_{ij})| = 1$ implies that

$$\arg\phi(a_j x_{ij}) = \pi + \arg\phi(a_1 x_{i1}) (\text{mod } 2\pi), \text{ for all } i, j \text{ with } j > 1. \tag{2.2.2}$$

In words, (2.2.2) is saying that given a fixed i, the $\phi(a_j x_{ij})$ all “point in the same direction” for $j > 1$, and $\phi(a_1 x_{i1})$ “points in the opposite direction.”

The rest of the argument is most easily expressed in polar coordinates. For all j, let $\phi(a_j) = r_j\theta_j$ where $r_j \in \mathbb{R}^{\geq 0}$ and $|\theta_j| = 1$. Fix any $i_0 \in \{1, 2, .., m\}$ and any $j \in \{2, \ldots, k\}$. Then by (2.2.2) and the fact that all $|\phi(x_{ij})| = 1$, we have $\phi(x_{i_0 j}) = -\frac{\theta_1}{\theta_j}\phi(x_{i_0 1})$.

By balancedness, there is some i_1 so that $x_{i_1 1} = x_{i_0 j}$, so repeating the above argument, we get $\phi(x_{i_1 j}) = -\frac{\theta_1}{\theta_j}\phi(x_{i_1 1}) = -\frac{\theta_1}{\theta_j}\phi(x_{i_0 j}) = (-\frac{\theta_1}{\theta_j})^2\phi(x_{i_0 1})$. Iterating, this argument shows that $(-\frac{\theta_1}{\theta_j})^m\phi(x_{i_0 1}) \in \{\phi(x) : x \in X\}$ for all $m \in \mathbb{Z}$, implying that $-\frac{\theta_1}{\theta_j}$ is a root of unity.

Now let $\omega_1 = 1$ and $\omega_j = -\frac{\theta_1}{\theta_j}$ for $j > 1$. Dividing the equation (2.2.1) with $i = i_0$ by $x_{i_0 1}$ and applying ϕ to both sides, we have $\sum_{i=1}^k \phi(a_i)\omega_i = 0$. Finally, letting $\rho_i = \psi^{-1}(\omega_i)$, we see that $\psi(\sum_{i=1}^k a_i\rho_i) = \sum_{i=1}^k \phi(a_i)\omega_i = 0$, and hence $\sum_{i=1}^k a_i\rho_i = 0$. □

The above work, along with the known necessity of the absolute value criteria, reduces Conjecture 2.2.1 to the following.

Conjecture 2.2.7. *Let K be a number field and $\mathcal{O}_K$ its ring of integers. If $(a_1, \ldots, a_k) \in \mathcal{O}_K^k$ satisfies the strong absolute value criteria, then $(a_1, \ldots, a_k)$ is a Smyth tuple.*

Speyer [Spe] gives a proof of Conjecture 2.2.7 in the case where $k = 3$ and $a_1 = a_2$.

Speyer's argument works for general k and $a_1 = \cdots = a_{k-1}$ with minimal modification; this result is our final proposition of the section.

Proposition 2.2.8. *Let $k \geq 3$ be an integer. Let K be a number field and $\mathcal{O}_K$ its ring of integers. Let $\alpha \in \mathcal{O}_K$ so that every archimedean absolute value of α is less than $k - 1$. Then $(1, 1, \ldots, 1, \alpha) \in \mathcal{O}_K^k$ is a Smyth tuple.*

Proof. By Lemma 2.2.3, it suffices to show $(1, 1, \ldots, 1, -\alpha)$ is a Smyth tuple. By Proposition 1.2.1, it suffices for us to show that there are permutation matrices X_i so that $\sum_{i=1}^{k-1} X_i$ has α as an eigenvalue.

We will follow the argument from [Spe], starting with a slight generalization of Speyer's Step 1, which we write out in full for the sake of clarity.

Step 1: There is a nonnegative integer matrix C, with eigenvalue α, all of whose row sums are $k-1$.

Consider the lattice $A = \mathbb{Z}[\alpha]$ and the vector space $V = A \otimes_{\mathbb{Z}} \mathbb{R}$. Since α is an algebraic integer, A is a discrete full sublattice of V. We take the norm $\sum_\nu |x|_\nu^2$, where the sum runs over all archimedean places. Let $c = \max_\nu |\alpha|_\nu$. By hypothesis, $c < k-1$. Denote by B_R the closed ball of radius R around 0.

Let M be large enough so that any ball of radius M around any point in V contains a point in A. Take R large enough so that $\frac{c}{k-1}R + (k-2)M < R$. Now, for any $z \in A \cap B_R$, let $z_1 \in A \cap B_R$ be the closest point to $\frac{\alpha z}{k-1}$. Let $z_2 = \alpha z - (k-2)z_1$. Now,

$$|z_1| \leq |z_1 - \frac{\alpha z}{k-1}| + |\frac{\alpha z}{k-1}| \leq M + \frac{c}{k-1}R < R.$$

Similarly,

$$|z_2| = |\alpha z - (k-2)z_1| \leq |\alpha z - \frac{k-2}{k-1}\alpha z| + (k-2)|\frac{\alpha z}{k-1} - z_1| \leq \frac{c}{k-1}R + (k-2)M < R.$$

Thus, for any $z \in A \cap B_R$, we can find $z_1, z_2 \in A \cap B_R$ so that $(k-2)z_1 + z_2 = z$. Enumerate the elements of $A \cap B_R$ as $z_1, z_2, \ldots, z_l$. Then, we can form an $l \times l$ matrix C with the following entries. For the t-th row, consider z_t. As before, we may write $(k-2)z_r + z_s = z_t$ for some $1 \leq r, s \leq l$. In the t-th row, put $k-2$ in the r-th column and 1 in the s-th column if $r \neq s$; if $r = s$, put an $k-1$ in the r^{th} column. Every row sum of the matrix C is equal to $k-1$. By construction, it has α as an eigenvalue with right eigenvector $(z_1, z_2, \ldots, z_l)^T$.

The rest of Speyer's argument can now be applied with virtually no modification; using the Perron-Frobenius theorem, one obtains a matrix D from C which is the sum of $k-1$ permutation matrices and still has α as an eigenvalue. $\square$

Chapter 3

The Slice Rank Approach to Smyth's Conjecture: Stochastic Patterns and Strassen's Asymptotic Spectrum

This chapter connects Smyth's Conjecture for triples to recent work in additive combinatorics around slice rank and less recent work from multi-linear algebra concerning asymptotic spectra. The first section of this chapter introduces preliminaries about tensors, the main objects of focus in this chapter. Next we introduce stochastic tensors/patterns and describe how they relate to balanced tensors. We then introduce the notions of asymptotic spectra and Strassen's support functionals, and explain how work in these areas provide a path towards proving Smyth's Conjecture via asymptotic slice rank. In particular, Smyth's Conjecture would follow from certain "Smyth tensors" having full asymptotic slice rank. (Along the way, we discuss matrix multiplication algorithms, which were the original motivation for Strassen's work.) Once this context is set, we are ready to prove the main result of this chapter, Theorem 3.4.4, which says that Smyth tensors have full slice rank. We go on to discuss the obstacle to extending our argument about slice rank to asymptotic slice rank. Finally, we conclude the section by presenting a computer-verified counterexample to an old conjecture of Brualdi and Csima regarding stochastic patterns.

3.1 Tensor Preliminaries

This chapter will focus on real 3-tensors $T \in V_1 \otimes V_2 \otimes V_3$, where $V_1 = V_2 = V_3 = \mathbb{R}^m$ for some m. We will write $\mathrm{Ten}(m,m,m)$ for the set of such tensors. Due to the combinatorial nature of this chapter, much of the section will treat these tensors as 3-d arrays of real numbers, with the basis for each tensor leg being the standard basis. That is, when we refer to the "entries" of a tensor, or otherwise are presuming certain bases, we are (implicitly or explicitly) expanding $T = \sum_{i,j,k=1}^{m} t_{ijk} e_i \otimes e_j \otimes e_k$, and identifying T with the 3-d array of coefficients $[t_{ijk}]_{i,j,k=1}^{m}$. The support of T is $\mathrm{supp}(T) = \{(i,j,k) : t_{ijk} \neq 0\}$.

This chapter will at times be concerned with *tensor powers*, by which we mean the following. A tensor $T \in V_1 \otimes V_2 \otimes V_3$, can be thought of as a multilinear map $T : V_1 \times V_2 \times V_3 \to \mathbb{R}$. The d^{th} tensor power $T^{\otimes d}$ of T corresponds to the multilinear map $T^{\otimes d} : V_1^{\otimes d} \times V_2^{\otimes d} \times V_3^{\otimes d} \to \mathbb{R}$ given by $(x_1 \otimes \cdots \otimes x_d, y_1 \otimes \cdots \otimes y_d, z_1 \otimes \cdots \otimes z_d) \mapsto \prod_{i=1}^{d} T(x_i, y_i, z_i)$. Thus the coefficients of $T^{\otimes d}$ with respect to the standard bases are all d-fold products of the corresponding coefficients of T, the support $\mathrm{supp}(T^{\otimes d})$ is the d-fold Cartesian product $\mathrm{supp}(T) \times \cdots \times \mathrm{supp}(T)$.

We will use the notation $\langle n,n,n \rangle$ to refer to the 0-1 diagonal $n \times n \times n$ tensor, $\sum_{i=1}^{d} e_i \otimes e_i \otimes e_i$.

3.2 Stochastic Patterns

Recall that by Proposition 1.3.3, Smyth's Conjecture (for triples) is equivalent to: for each potential Smyth triple (a,b,c), there exists a nonnegative real balanced tensor with support contained in $\{(x,y,z) \in \mathbb{Z}^3 : ax + by + cz = 0\}$. Let us fix a potential Smyth triple (a,b,c) and as before, set $n = abc$, and $N = 2n+1$. Let $S = [-n,n]$, $\Phi_{abc} = \{(x,y,z) \in S^3 : ax+by+cz = 0\}$, and write $T_{abc} \in (\mathbb{R}^N)^{\otimes 3}$ for the 0-1 tensor with support $\mathrm{supp}(T_{abc}) = \Phi_{abc}$; that is, $T_{abc} = \sum_{(i,j,k) \in \Phi_{abc}} (e_i \otimes e_j \otimes e_k)$.

A nonnegative 3-tensor $T = \sum t_{ijk} e_i \otimes e_j \otimes e_k$ is said to be *stochastic* if its plane sums are all 1, i.e. if for every fixed k, $\sum_{i,j} t_{ijk} = 1$, and similarly $\sum_{j,k} t_{ijk} = 1$ and $\sum_{i,k} t_{ijk} = 1$ for any fixed i and j respectively. A *pattern* in our context is any subset of S^3, and a pattern is stochastic if there exists a tensor which is stochastic and has support contained in the pattern. Similarly, we will say a pattern is *balanced* if there exists a balanced tensor with support contained in the pattern.

Let $\Delta = \{(s, s, s) : s \in S\}$ denote the main diagonal of S^3; we will say a pattern P is *non-diagonally stochastic* if there exists a stochastic tensor T whose support is contained in P, but is not equal to Δ.

To prove Smyth's Conjecture, it would certainly be sufficient to show that Φ_{abc} is a stochastic pattern; the following proposition explains more precisely the relationship between balanced patterns and stochastic patterns.

Proposition 3.2.1. *Let $P \subset S^3$ be a pattern. P is balanced if and only if $P \cup \Delta$ is non-diagonally stochastic.*

Proof. ($\Rightarrow$): Suppose that P is balanced; let T be a balanced tensor with support contained in P. Then the plane sums of T form the same vector $v = \{v_i\}_{i\in S}$ in each direction. Let $m = \max_{i\in S} v_i$; then the scaled tensor $\frac{1}{m}T$ is balanced and has plane sums between 0 and 1. For each $i \in S$, let $d_i = 1 - v_i$. Adding d_i to the i^{th} diagonal entry of $\frac{1}{m}T$ yields a stochastic tensor, which is not diagonal since balanced tensors are by definition nonzero.

($\Leftarrow$): If T is a non-diagonal stochastic tensor with support contained in P, then setting the entries of T along Δ equal to 0 yields a tensor which is still balanced, and now has support contained in P. $\square$

This proposition has utility for algorithmic approaches to Smyth's Conjecture. There is some work on algorithms [Bur+18] for finding stochastic tensors with a given support (or support contained in a certain set) and this proposition means one can extend these algorithms to look for balanced tensors by first adding the diagonal to the input tensor's support, running the algorithm, and then zeroing out the diagonal. If the algorithm found a non-diagonal stochastic tensor, the end result will be a nonzero balanced tensor with appropriate support.

At first glance, Proposition 3.2.1 gives the impression that we can dispose of balanced tensors and instead pursue an equivalent problem involving the better-studied stochastic tensors. However, this is not quite the approach we will take. We will turn our attention to stochastic tensors, but our specific approach – which will eventually be studying the asymptotic slice rank of Smyth tensors – will discourage enlarging the support of Smyth tensors. So, over the course of the rest of this chapter, we will end up trying, and failing, to show that the Smyth pattern itself is stochastic.

3.3 Strassen's Asymptotic Spectrum

This section builds towards Corollary 3.3.4, which gives a criterion for certain patterns – including Smyth patterns – to be stochastic. We arrive at this result via the theory of asymptotic spectra developed principally by Volker Strassen starting in the 1980s. To set the scene, we begin by discussing his initial motivation.

3.3.1 Matrix Multiplication Algorithms

The problem of finding the fastest algorithms for matrix multiplication has been of great interest for a long time, particularly recently with the increasingly widespread deployment of machine learning algorithms built out of matrix multiplications. The definitional algorithm requires n^3 arithmetic operations to multiply two $n \times n$ matrices, but Volker Strassen showed in 1969 that one can multiply 2×2 matrices using only 7 arithmetic operations [Str69]. (The algorithm appears quite unilluminating and has largely resisted attempts to extract meaning, although recently Grochow and Moore gave a convincing conceptual justification for the existence of a 7-operation algorithm [GM17]. I am not aware, however, of any illumination of why Strassen's 7 operations in particular work.) This improvement in the 2×2 case automatically extends to general $n \times n$ matrices, by treating an $n \times n$ matrix as a 2×2 block matrix.

The exponent of matrix multiplication ω is defined to be the minimum real number such that the multiplication of $n \times n$ matrices can be done in $O(n^{\omega+o(1)})$ arithmetic operations. The definitional algorithm means that $\omega \leq 3$. Strassen's algorithm proved that $\omega \leq \log_2 7 = 2.807....$ The tightest bounds currently known are $2 \leq \omega \leq 2.37286$. (The lower bound is simply due to the algorithm outputting n^2 entries.)

The multiplication of $m \times n$ and $n \times p$ matrices is a bilinear map between vector spaces $\mathbb{R}^{m\times n} \times \mathbb{R}^{n\times p} \to \mathbb{R}^{m\times p}$, and hence is represented by a single tensor $\langle m, n, p\rangle$ with respect to the standard bases. It has been shown that ω can alternatively be defined as $\omega = \{\inf \beta : R(\langle n, n, n\rangle) = O(n^{\beta})\}$. Thus the asymptotic complexity of matrix multiplication is controlled by the tensor ranks of the tensors $\langle n, n, n\rangle$.

In his Ph.D. thesis, Garterberg generalized the notion of the exponent of matrix multiplication as follows. [Gar85] Noting that the d^{th} tensor power $\langle n, n, n\rangle^{\otimes d} = \langle n^d, n^d, n^d\rangle$, one can define the asymptotic (tensor) rank of a tensor T as $\underset{\sim}{R}(T) = \lim_{d\to\infty}(R(T^{\otimes d})^{1/d}$; this limit is guaranteed to exist. Thus $\omega = \log_n(\underset{\sim}{R}(\langle n, n, n\rangle))$, for any integer $n > 1$, and the complexity of matrix multiplication is determined by the asymptotic rank of a single tensor $\langle 2, 2, 2\rangle$.

3.3.2 The Asymptotic Restriction Problem

Strassen, motivated to understand the complexity of matrix multiplication, began to study the asymptotic restriction problem, which is as follows.

Given multilinear maps $f, g : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \mathbb{R}^{n_3} \to \mathbb{R}$, one says that f *restricts to* g (written $f \geq g$) if there are linear maps (r_1, r_2, r_3) such that $g = f \circ (r_1, r_2, r_3)$. Moreover, f *asymptotically restricts to* g (written $f \gtrsim g$) if there exists a sequence of natural numbers $a_d \in o(d)$ such that $f^{\otimes d + a_d} \geq g^{\otimes d}$. The asymptotic restriction problem asks, given f and g, whether $f \gtrsim g$.

One way to prove a negative answer to this question would be to identify a family of tensors $\mathcal{X}$ containing f and g and produce a map $\phi : \mathcal{X} \to \mathbb{R}^{\geq 0}$ such that $\phi(f) < \phi(g)$, and such that ϕ also has the following general properties.

1. monotone under restriction $\geq$
2. multiplicative under tensor product $\otimes$
3. additive under direct sum $\oplus$

Note that by Property 2, $\phi(f^{\otimes d}) < \phi(g^{\otimes d})$ for any d; and for large enough d depending on a_d, $\phi(f^{\otimes d + a_d}) < \phi(g^{\otimes d})$. By Property 1, it follows that f does not asymptotically restrict to g.

Strassen called such maps ϕ *spectral points.* He normalized them by assuming additionally that $\phi(\langle n \rangle) = n$. He then showed that, remarkably, spectral points are the *only* obstacles to $f \gtrsim g$. [Str88] In particular, one can study the asymptotic restriction problem for all pairs of maps $f, g \in \mathcal{X}$ by determining the spectral points ϕ of $\mathcal{X}$. Then, given particular f and g, one compares each of the spectral points evaluated at f and g. If $\phi(f) \geq \phi(g)$ for all spectral point ϕ, then $f \gtrsim g$; otherwise f does not asymptotically restrict to g.

It has since been of great interest to construct spectral points for interesting families of tensors. Strassen himself constructed "support functionals" which are spectral points for the family of oblique tensors. (As the name suggests, support functionals depend only on the support of the tensors.) We'll define oblique tensors a little later; for now, suffice it to say that Smyth tensors are oblique.

More recently, Cristandl, Vrana, and Zuiddam constrcuted the first nontrivial spectral points for the family of all complex tensors ("universal spectral points") [CVZ18]. They also used Strassen's support functionals to prove a formula for the asymptotic slice rank of tight tensors, a refinement of oblique tensors. It is this result that directly applies to our situation, so we will now introduce the relevant notions – slice rank, oblique tensors, and tight tensors.

3.3.3 Introduction to Slice Rank and Asymptotic Slice Rank

Our main attempt to prove Smyth's Conjecture interrogates the *slice rank* of the Smyth tensors and their tensor powers. The notion of slice rank arose from Tao's reformulation [Tao16] of the work of Croot, Lev, Pach, Ellenberg, and Gijswijt that resolved the cap set conjecture. [EG17] [CLP17]

A tensor $T : V_1 \otimes \ldots \otimes V_k \to \mathbb{R}$ has slice rank one if it can be written as $T(x_1, \ldots, x_k) = f(x_i)g(x_1, \ldots x_{i-1}, x_{i+1}, \ldots x_k)$ for some index i and some functions f and g.

More generally, the *slice rank* of T is the minimal r such that T can be written as the sum of r slice-rank-one tensors (note that the index i in the above form does not have to be the same for each slice-rank-one tensor in the decomposition; if it did, the definition would reduce to the minimum matrix rank of the "flattenings" of T). Slice rank has the following notable properties.

- Slice rank is in invariant under the action of $\prod_{i=1}^k \mathrm{GL}(V_i)$.
- Slice rank is a generalization of the ordinary notion of rank of matrices. That is, the slice rank of a 2-tensor is its rank as a matrix.
- Given $T \in \mathrm{Ten}(n_1, \ldots, n_k)$, we have $SR(T) \leq \min_i n_i$, as can be seen for each i by flattening T along its i^{th} coordinate.
- The slice rank of a "diagonal tensor" equals its number of nonzero entries. That is, given $T = \sum_{i=1}^n c_i e_i \otimes \ldots e_i \in \mathrm{Ten}(n, \ldots, n)$, we have $SR(T) = \#\{i : c_i \neq 0\}$. [Tao16]

This last bullet point was key in Tao's reformulated proof of the cap set conjecture. We are going to make use of a generalization of this result, which was jointly proved by Tao and Sawin (we will only state it in the generality we need).

Proposition 3.3.1. *[TS16, Prop. 4] For each $1 \leq j \leq 3$, let $(v_{j,s})_{s \in S}$ be a linearly independent subset of $V_j = \mathbb{R}^d$ indexed by some finite subset S_j. Let Γ be a nonempty subset of $S_1 \times S_2 \times S_3$.*

Suppose further that there are total orderings $\leq_i$ on S_i such that Γ is an anti-chain (i.e. every element of Γ is maximal).

Let $v = \sum_{(s_1,s_2,s_3) \in \Gamma} c_{s_1,s_2,s_3} v_{1,s_1} \otimes v_{2,s_2} \otimes v_{3,s_3}$, where the coefficients c_{s_1,s_2,s_3} are nonzero and lie in $\mathbb{R}$.

Then we have

$$SR(v) = \min_{\Gamma = \Gamma_1 \sqcup \Gamma_2 \sqcup \Gamma_3} |\pi_1(\Gamma)| + |\pi_2(\Gamma)| + |\pi_3(\Gamma)|,$$

where the minimum ranges over all coverings of Γ by sets $\Gamma_{1'}, \Gamma_{2'}, \Gamma_{3'}$ and $\pi_i : S_1 \times S_2 \times S_3 \to S_i$ are the projection maps.

A tensor T is said to be *oblique* if its support is an anti-chain in the sense described in the statement of Proposition 3.3.1. Thus Proposition 3.3.1 is saying that given a tensor $T \in$

$\mathrm{Ten}(n,n,n)$ which is *oblique*, computing the slice-rank of T is equivalent to the combinatorial problem of determining the size of the minimum covering of the support of T by axis-parallel "slices." Note that in particular, for oblique tensors, slice-rank depends only on the tensor's support. Hence we can refer to the slice rank of a tensor's support, when it is oblique, without ambiguity.

Moreover, a 3-tensor $T \in \mathrm{Ten}(I_1, I_2, I_3)$ is *tight* if there are injective functions $u_i : I_i \to \mathbb{R}$ such that $\sum_{i=1}^{3} u_i(\alpha_i) = 0$ for all $\alpha = (\alpha_1, \alpha_2, \alpha_3) \in \mathrm{supp}(T)$. By definition, Smyth tensors T_{abc}, and their tensor powers, are tight.

We can see that tight tensors are oblique as follows.

Proposition 3.3.2. *Tight tensors are oblique.*

Proof. Let T be a tight k-tensor indexed by $I_1 \times \cdots \times I_k$. Let $u_1, \ldots, u_k$ be the corresponding injective maps. For each i, let $\leq_i$ be the ordering determined by pulling back the real ordering along u_i, i.e. for $x, y \in I_i$ $x \leq_i y \iff u_i(x) \leq u_i(y)$.

Now if $\alpha = (\alpha_1, \ldots \alpha_k)$ and $\beta = (\beta_1, \ldots, \beta_k) \in \mathrm{supp}(T)$ with $\alpha_i \leq_i \beta_i$ for all i, then also $\sum_{i=1}^{k} u_i(\alpha_i) = 0 = \sum_{i=1}^{k} u_i(\beta_i)$. Therefore each inequality $u_i(\alpha_i) \leq u_i(\beta_i)$ must be an equality and since the maps are injective, it follows that $\alpha = \beta$. $\square$

The previous two propositions show that computing the slice rank of tensor powers $T_{abc}^{\otimes d}$ of Smyth tensors is equivalent to the combinatorial problem of determining the size of the minimum covering of the support of $T^{\otimes d}$ by axis-parallel slices. We will now see how this relates to the question of whether or not the Smyth support Φ_{abc} is stochastic.

First, we need a little more terminology/notation. We will write I_1, I_2, I_3 for the index sets of our tensor legs. Given a subset $\Phi \subset I_1 \times I_2 \times I_3$, we'll write $\mathscr{P}(\Phi)$ denote the set of all probability distributions supported on Φ. We will also need the notion of entropy of a probability distribution P. Entropy can be thought of as a measure of the expected information gain of observing a sample from P. When P is a discrete random variable on a set X, the *entropy* of P is given by the formula

$$H(P) = -\sum_{x \in X} P(x) \log_2(x)$$

In this context, $\log_2 0$ is understood to be 0.

The following theorem gives a formula relating entropy of probability distributions supported on Φ to asymptotic slice rank, written $\underset{\sim}{SR}$. Recall that in the case of tight/oblique tensors, slice rank depend only on the support Φ.

Theorem 3.3.3. *[CVZ18, Thm 4.4, Cor 5.10] Let $\Phi \subset I_1 \times I_2 \times I_3$ be tight. We will write P_i to denote the i^{th} marginal of a probability distribution P.*

Then

$$\underset{\sim}{SR}(\Phi) = \max_{P \in \mathscr{P}(\Phi)} \min\{2^{H(P_1)}, 2^{H(P_2)}, 2^{H(P_3)}\}$$

Corollary 3.3.4. *A tight pattern $\Phi \subset S^3$ is stochastic if and only if $\underset{\sim}{SR}(\Phi) = N$.*

Proof. This follows from some basic properties of entropy, namely that $H : \mathscr{P}(N) \to \mathbb{R}$ achieves its unique maximum, which is $\log_2(N)$, on the uniform distribution. Therefore, given a pattern $\Phi \subset S^3$,

$$\begin{aligned} \Phi \text{ is stochastic } &\iff \exists P \in \mathscr{P}(\Phi) : H(P_i) = \log_2(N) \text{ for all } i \\ &\iff N = \max_{P \in \mathscr{P}(\Phi)} \min\{2^{H(P_1)}, 2^{H(P_2)}, 2^{H(P_3)}\}. \end{aligned}$$

By Theorem 3.3.3, this latter quantity equals $\underset{\sim}{SR}(\Phi)$ if Φ is additionally tight. $\square$

To recap: since Smyth patterns Φ_{abc} are tight, proving that they are stochastic is equivalent to proving that T_{abc} has full asymptotic slice rank, which is equivalent to proving that one needs to use 100% of the slices $\{x = s : s \in S^d\} \cup \{y = s : s \in S^d\} \cup \{z = s : s \in S^d\}$ to cover the support of $T_{abc}^{\otimes d}$ as $d \to \infty$.

3.4 Smyth Tensors Have Full Slice Rank

The purpose of this section is to prove Theorem 3.4.4, but first, we need a couple of lemmas. In this section we will make use of the following notation. Given two finite sets A, B, we will write d_A^B to mean $\frac{|A \cap B|}{|B|}$, i.e. the density of A in B. Once B has been made clear in a certain context we may omit it from subsequent occurrences of the notation.

Lemma 3.4.1. *[TV06] Let $A, B \subset \mathbb{Z}$. Then $|A+B| \geq |A| + |B| - 1$*

Proof. This is a fundamental result in additive combinatorics with a short proof:

Order the elements of each set $A = \{a_1, \dots, a_r\}$ and $B = \{b_1, \dots b_s\}$ under the standard real ordering $\leq$. Then $a_1 + b_1 < a_2 + b_1 < \dots < a_r + b_1 < a_r + b_2 < a_r + b_3 < \dots < a_r + b_s$ are $r + s - 1$ distinct elements of $A + B$. □

Lemma 3.4.2. *Let P, Q be arithmetic progressions in $\mathbb{Z}$ with the same step size, and let $R \subset (P+Q)$ such that $|R| > \max(|P|, |Q|)$. Then for any subsets $A \subset P, B \subset Q$, we have $d_{A+B}^R \geq d_A^P + d_B^Q - 1$. (Furthermore, equality holds if and only if $A = P$ and $B = Q$.)*

Proof. We have that

$$\begin{aligned}
|(A+B) \cap R| &\geq |A+B| - |(P+Q) \setminus R|, \text{ since } (A+B) \subset (P+Q) \\
&= |A+B| - (|P+Q| - |R|), \text{ since } R \subset (P+Q) \\
&\geq |A| + |B| - 1 - (|P+Q|) + |R|, \text{ by Lemma 3.4.1} \\
&= |A| + |B| - 1 - (|P| + |Q| - 1) + |R|, \text{ since } P, Q \text{ are APs with the same step size} \\
&= |A| + |B| - |P| - |Q| + |R|
\end{aligned}$$

Therefore,

$$\begin{aligned}
d_{A+B}^R = \frac{|(A+B) \cap R|}{|R|} &\geq \frac{|A| + |B| + |R| - |P| - |Q|}{|R|} \\
&= 1 + \frac{|A| + |B| - |P| - |Q|}{|R|}
\end{aligned}$$

So it will suffice to show that

$$\frac{|A|}{|P|} + \frac{|B|}{|Q|} - \frac{|A| + |B| - |P| - |Q|}{|R|} \leq 2$$

Indeed, viewing P, Q, and R as fixed, the quantity on the LHS is increasing in both $|A|$ and $|B|$ over their whole mutual domain, and hence is maximized when $A = P$ and $B = Q$, at which point we have equality.

Thus we have shown the desired inequality. We have also shown that equality can only hold if $A = P$ and $B = Q$. It is straightforward to see that equality does in fact hold in this case. $\square$

Lemma 3.4.3. *For any subsets $X, Y \subset S$, the quantity $d^{cS}_{aX+bY} - d^S_X - d^S_Y$ is a weighted average of the quantities $d^{(cS)_{bj;ai;}}_{aX_i+bY_j} - d^{S_{;i;}}_{X_i} - d^{S_{j;;}}_{Y_j}$ as i, j range over $[b]$ and $[a]$ respectively.*

Proof. For each $i \in [b], j \in [a]$, let $w_{ji} := d^S_{S_{j;i;}}$ (these will be the weights in the weighted average). Thus $w_{ji} = \begin{cases} \frac{N+ab-1}{Nab}, & i = j = 0 \\ \frac{N-1}{Nab}, & \text{otherwise.} \end{cases}$

Note that the w_{ji} satisfy the following properties.

- $w_{ji} = w_{bj\ ai}$ for all $i \in [b], j \in [a]$.
- $w_{ji} = d^{cS}_{(cS)_{j;i;}}$ for all $i \in [b], j \in [a]$.
- For any fixed $j \in [a]$, $\sum_{i \in [b]} w_{ji} = d^S_{S_{j;;}}$, and for any fixed $i \in [b]$, $\sum_{j \in [a]} w_{ji} = d^S_{S_{;i;}}$.

Putting these properties together with definitions, we see that

$$\begin{aligned}
d^{cS}_{aX+bY} - d^S_X - d^S_Y &= \frac{|(aX+bY)\cap cS|}{|cS|} - \frac{|X|}{|S|} - \frac{|Y|}{|S|} \\
&= \sum_{i\in[b],j\in[a]} \Big(\frac{|(aX_i+bY_j)\cap cS|}{|cS|} - \frac{|X_i|}{a|S|} - \frac{|Y_j|}{b|S|} \Big) \\
&= \sum_{i\in[b],j\in[a]} \frac{|S_{j;i;}|}{|S|} \frac{|(aX_i+bY_j)\cap (cS)_{bj;ai;}| - \frac{1}{a}|X_i| - \frac{1}{b}|Y_j|}{|S_{j;i;}|} \\
&= \sum_{i\in[b],j\in[a]} w_{ji} \Big(\frac{|(aX_i+bY_j)\cap (cS)_{bj;ai;}|}{|(cS)_{bj;ai;}|} - \frac{|X_i|}{a|S_{j;i}|} - \frac{|Y_j|}{b|S_{j;i}|} \Big) \\
&= \sum_{i\in[b],j\in[a]} w_{ji} \Big(\frac{|(aX_i+bY_j)\cap (cS)_{bj;ai;}|}{|(cS)_{bj;ai;}|} - \frac{|X_i|}{|S_{;i;}|} - \frac{|Y_j|}{|S_{j;;}|} \Big) \\
&= \sum_{i\in[b],j\in[a]} w_{ji} \Big(d^{(cS)_{bj;ai;}}_{aX_i+bY_j} - d^{S_{;i;}}_{X_i} - d^{S_{j;;}}_{Y_j} \Big)
\end{aligned}$$

To verify the second-last line, we compute the coefficient of $|X_i|$ in the second-last line. This coefficient is

$$-\sum_{j\in[a]} \frac{w_{ji}}{|S_{;i;}|} = -\sum_{j\in[a]} \frac{-|S_{j;i;}|}{|S||S_{;i;}|} = -\frac{1}{|S|}.$$

This matches the coefficient of $|X_i|$ in the previous line. A similar argument shows that the coefficient of $|Y_j|$ in the second- and third-last lines match too. This completes the proof of the lemma.

□

We are now ready to prove the main theorem of this section. The reader is invited to refer to the example following the proof for concreteness.

Theorem 3.4.4. *If (a, b, c) be a potential Smyth triple then $SR(T_{abc}) = N$.*

Proof. First recall from Proposition 3.3.2 that Smyth tensors are oblique, and so Proposition 3.3.1 applies.

By Proposition 3.3.1, $SR(T) = N$ is equivalent to the following: given any subsets $X, Y, Z \subset S$ such that $0 \notin aX + bY + cZ$, we have

$$|X| + |Y| + |Z| \leq 2N, \text{ or equivalently, that}$$
$$d_X^S + d_Y^S + d_Z^S \leq 2.$$

And this in turn is equivalent to showing that for any subsets $X, Y \subset S$ we have

$$d_{aX+bY}^{cS} \geq d_X^S + d_Y^S - 1.$$

For each $i \in [b]$, let $X_i = \{x \in X : x \equiv i \pmod{b}\}$, and similarly for each $j \in [a]$, let $Y_j = \{y \in Y : y \equiv j \pmod{a}\}$. We will decompose $aX + bY$ into mod-ab residue classes; this decomposition can be written as the disjoint union

$$aX + bY = \sqcup_{i \in [b], j \in [a]} (aX_i + bY_j).$$

We now fix $i \in [b]$ and $j \in [a]$. In what follows we will write $S_{i;j;k}$ to denote elements of S congruent to i mod a, j mod b, and k mod c. When we only wish to specify one or two residue values, we will leave empty spaces for i, j or k, but will still write the semi-colons (e.g. $S_{;j;}$ denotes the subset of elements congruent to j mod b). Notice that the elements of $aX_i + bY_j$ are congruent to ai mod b and bj mod a. Thus $(aX_i + bY_j) \cap (cS)_{i';j';}$ can be nonempty only if $i' \equiv bj \pmod{a}$ and $j' \equiv ai \pmod{b}$.

By Lemma 3.4.3, it will suffice to show that

$$d_{aX_i+bY_j}^{(cS)_{bj;ai;}} - d_{X_i}^{S_{;i;}} - d_{Y_j}^{S_{j;;}} \geq -1.$$

Towards showing this, we will now consider each set modulo c. First, let $f : [c] \to [c]$ denote the bijection such that $ak + bf(k) \equiv 0 \pmod{c}$ for all $k \in [c]$.

For each $r \in [c]$, let $X_{ik} = \{x \in X_i : x \equiv k \pmod{c}\}$, and similarly $Y_{jk} = \{y \in Y_j : y \equiv k \pmod{c}\}$. Now fix a k for which $d^{S_{;i;k}}_{X_{ik}} + d^{S_{j;;f(k)}}_{Y_{jf(k)}}$ is maximal among all $k \in [c]$. Notice that it follows that $d^{S_{;i;k}}_{X_{ik}} + d^{S_{j;;f(k)}}_{Y_{jf(k)}} \geq d^{S_{;i;}}_{X_i} + d^{S_{j;;}}_{Y_j}$, since $d^{S_{;i;}}_{X_i} + d^{S_{j;;}}_{Y_j}$ is a weighted average of $d^{S_{;i;r}}_{X_{ir}} + d^{S_{j;;f(r)}}_{Y_{jf(r)}}$ over all $r \in [c]$. (The weights are nearly uniform, but the 0 residue class is slightly over-represented.)

Our proof is to show the following.

$$\begin{aligned} d^{(cS)_{bj;ai;}}_{aX_i+bY_j} &\geq d^{(cS)_{bj;ai;}}_{aX_{ik}+bY_{jk}} \\ &\geq d^{S_{;i;k}}_{X_{ik}} + d^{S_{j;;f(k)}}_{Y_{jf(k)}} - 1 \\ &\geq d^{S_{;i;}}_{X_i} + d^{S_{j;;}}_{Y_j} - 1 \end{aligned}$$

The last line holds by choice of k, as just described; and the first line is clear since $X_{ik} \subset X_i$ and $Y_{jf(k)} \subset Y_j$.

To complete the proof, we will explain how the middle line follows from Lemma 3.4.2, with $A = aX_{ik}$, $B = bY_{jf(k)}$, $P = aS_{;i;k}$, $Q = bS_{j;;f(k)}$, and $R = (cS)_{bj;ai;}$.

We just need to confirm that the hypotheses of the lemma are satisfied. The sets $aS_{;i;k}$ and $bS_{j;;k}$ are each arithmetic progressions with step size abc. The set containments $aX_{ik} \subset aS_{;i;k}$ and $bY_{jf(k)} \subset bS_{j;;f(k)}$ are by definition.

Additionally,

$$|(cS)_{bj;ai;}| = \begin{cases} \frac{2n}{ab}, & \text{if } bj \neq 0 \text{ (mod a) or } ai \neq 0 \text{ (mod b)} \\ \frac{2n}{ab} + 1, & \text{else} \end{cases}$$

The cardinalities $|aS_{;i;k}|$ and $|bS_{j;;f(k)}|$ are computed similarly. Now notice that $\frac{2n}{ab} = 2c > 2b+1 = \frac{2n}{ac} + 1$, thus showing that $|(cS)_{bj;ai;}| > \max(|aS_{;i;k}|, |bS_{j;;f(k)}|)$.

It remains to establish that $(cS)_{bj;ai} \subset aS_{;i;k} + bS_{j;;f(k)}$. This is where the Archimedean absolute value criterion (triangle inequality) is required. First, observe that both sets $(cS)_{bj;ai}$ and $aS_{;i;k} + bS_{j;;f(k)}$ are arithmetic progressions with step size abc. Additionally, every element of both sets is congruent to bj mod a; ai mod b; and 0 mod c. So we just need to show that the maximum and minimum of $aS_{;i;k}+bS_{j;;f(k)}$ are more extreme than those of $(cS)_{bj;ai;}$. First notice that we have $\max S_{;i;k} \geq n-bc+1$, and similarly, $\max S_{j;;k} \geq n-ac+1$. Next, observe that the maximum element of $(cS)_{bj;ai;}$ lies in the range $[(c-1)n+1, cn]$ (recalling that $n = abc$). So in order to show that $\max(aS_{;i;k} + bS_{j;;f(k)}) \geq \max(cS)_{bj;ai;}$, it will suffice to show that $\max aS_{;i;k} + bS_{j;;f(k)} > (c-1)n$, since the two progressions occupy the same mod-n residue class. Indeed,

$$\begin{aligned}
\max(aS_{;i;k} + bS_{j;;f(k)}) &\geq a(n-bc+1) + b(n-ac+1) \\
&= (a+b)n - 2abc + a + b \\
&= (a+b-2)n + a + b \\
&\geq (c-1)n + a + b \\
&> (c-1)n
\end{aligned}$$

An analogous argument shows that $\min aS_{;i;k} + bS_{j;;f(k)} \leq \min(cS)_{bj;ai;}$. □

Before proceeding to discuss the obstacle to extending this argument to higher dimensions, we walk through an example to make things more concrete.

Example 3.4.5. *Consider the Smyth triple* $(a,b,c) = (3,4,5)$. *Then* $n = 3 \cdot 4 \cdot 5 = 60$ *and* $S = [-60, 60]$. *Let* $i = j = 1$. *The sets* $X_1 = \{-59, -55, \ldots, -3, 1, 5, \ldots, 57\}$ *and* $Y_1 = \{-59, -56, \ldots, -2, 1, 4, \ldots, 58\}$. *All elements of* $3X_1 + 4Y_1$ *are congruent to* 3 *mod 4, and* 1 *mod 3. Let* $k = 2$. *Then* $f(k) = 1$, *as* $3 \cdot 2 + 1 \cdot 4 = 0$ *(mod 5). We now have* $X_{1\,2} = \{-43, -23, -3, 17, 37, 57\}$ *and* $Y_{1\,1} = \{-59, -44, -29, -14, 1, 16, 31, 46\}$. *The sumset* $3X_{1\,2} + 4Y_{1\,1} = \{-365, -305, \ldots, 295, 355\}$. *Meanwhile,* $(cS)_{bj;ai;} = (5S)_{1;3;} = \{-245, -185, \ldots, 235, 295\}$.

3.5 The Obstacle to Full Asymptotic Slice Rank

The problem of showing that a Smyth tensor power $T^{\otimes d}$ has full slice-rank for $d \geq 2$ (which may not be necessary, as we only need to show T has full asymptotic slice rank) is in many ways similar to showing full slice rank for T itself. The tensor $T^{\otimes d}$ is indexed by S^{3d}, with support

$\{(x, y, z) : x, y, z \in S^d \text{ and } ax + by + cz = 0\}$.

Nearly everything from Section 3.4 holds in an analogous way. The support of $T^{\otimes d}$ is still tight and oblique, and so the slice rank problem connects to sumsets in the same way. And we can similarly decompose subsets $X, Y \subset S^d$ into mod-ab residue classes (in this case, there are $(ab)^d$ of them), and then into mod-c residue classes. It again would be sufficient to show that $d^{(cS)_{bj;ai;}}_{aX_{ik}+bY_{jk}} \geq d^{S_{;i;k}}_{X_{ik}} + d^{S_{j;;f(k)}}_{Y_{jf(k)}} - 1$, for a k chosen in the same way as in the one-dimensional case. However, this bound is not true in general, as we'll now see.

The obstacle lies in establishing a higher-dimensional analogue of Lemma 3.4.2. This lemma was written in a general fashion, but we can be a little more specific about our situation. Recall that $P = aS_{;i;k}$, $Q = bS_{j;;f(k)}$, and $R = (cS)_{bj;ai;}$ were each arithmetic progression is of step size $n = abc$. Notice that P has exactly one element p^* in the interval $[-\frac{n}{2}, \frac{n}{2})$; and similarly Q has one element $q^* \in (-\frac{n}{2}, \frac{n}{2}]$. Let $r^* := p^* + q^* \in (-n, n)$; by construction, $r^* \in R$. Now translating and dilating the sets P, Q, R by $\tilde{P} = \frac{1}{n}(P - p^*)$, $\tilde{Q} = \frac{1}{n}(Q - q^*)$, $\tilde{R} = \frac{1}{n}(R - r^*)$, we get that

$$\tilde{P} = \begin{cases} [-a, a-1], & \text{if } p^* > 0 \\ [-(a-1), a] & \text{if } p^* < 0 \\ [-a, a], & \text{if } p^* = 0 \end{cases} \qquad \tilde{Q} = \begin{cases} [-b, b-1], & \text{if } q^* > 0 \\ [-(b-1), b] & \text{if } q^* < 0 \\ [-b, b], & \text{if } q^* = 0 \end{cases}$$

$$\tilde{R} = \begin{cases} [-c, c-1], & \text{if } r^* > 0 \\ [-(c-1), c] & \text{if } r^* < 0 \cdot \\ [-c, c], & \text{if } r^* = 0 \end{cases}$$

These transformations are sum-preserving in the sense that $p + q = r \iff \tilde{p} + \tilde{q} = \tilde{r}$, where $\tilde{\cdot}$ represents the image of each element under their respective translation/dilation.

Example 3.5.1. *We continue from Example 3.4.5. Given (a, b, c) and X_{12}, Y_{11}, and $(5S)_{1;3;}$ as before, we have $P = 3X_{12} = \{-129, -69, -9, 51, 111, 171\}$, $Q = 4Y_{11} = \{-236, -176, -116, -56, 4, 64, 124, 184\}$, and $R = (5S)_{1;3;} = \{-245, -185, -125, -65, -5, 55, 115, 175, 225, 285\}$. In this case $p^* = -9$, $q^* = 4$, and so $r^* = -9 + 4 = -5$. Then $\tilde{P} = [-2, 3]$, $\tilde{Q} = [-4, 3]$, and $\tilde{R} = [-4, 5]$.*

Whenever $p^*, q^*, r^* \neq 0$, the intervals $\tilde{P}, \tilde{Q}$, and $\tilde{R}$ are asymmetric – their averages are $\pm\frac{1}{2}$ –

and we have $\operatorname{avg}(\tilde{P}) + \operatorname{avg}\tilde{Q} \neq \operatorname{avg}(\tilde{R})$. This observation gives rise to a simple proof that there exist mod-ab residue classes where the inequality fails.

Proposition 3.5.2. *For large enough $d > 0$, the following holds.*

Let $\bar{i} = (i, i, \ldots, i) \in (\mathbb{Z}/b\mathbb{Z})^d$, $\bar{j} = (j, j, \ldots, j) \in (\mathbb{Z}/a\mathbb{Z})^d$. Then for any $k = (k_1, k_2, \ldots, k_d) \in (\mathbb{Z}/c\mathbb{Z})^d$ with all $k_i \neq 0$, there exist subsets $aX_{ik} \subset P = aS_{;i;k}$, $bY_{jk} \subset Q = bS_{j;;f(k)}$, such that $d^{(cS)_{bj;ai;}}_{aX_{ik}+bY_{jk}} < d^{S_{;i;k}}_{X_{ik}} + d^{S_{j;;f(k)}}_{Y_{jf(k)}} - 1$. (Thus, we cannot have a high dimensional analogue of Lemma 3.4.3.)

Proof. We describe the sets aX_{ik} and bY_{jk} by their images under the maps ϕ_P and ϕ_Q. We will let $\sigma : \mathbb{R}^d \to \mathbb{R}$ denote the homomorphism which averages coordinates, i.e. $(x_1, \ldots, x_d) \mapsto \frac{x_1 + \cdots + x_d}{d}$. Given a finite subset $A \subset \mathbb{R}^d$, we will write $\sigma(A)$ to denote $\frac{\sum_{a \in A} \sigma(a)}{|A|}$, the average value of σ on A.

By choice of i, j, k, the sets $\tilde{P}, \tilde{Q}, \tilde{R}$ are each centered at $\pm\frac{1}{2}$, as discussed above, and therefore $|\sigma(\tilde{P}) + \sigma(\tilde{Q}) - \sigma(\tilde{R})| \geq \frac{1}{2}$. Let $\epsilon = \frac{1}{6}$ and let $\widetilde{aX_{ik}} = \{x \in \tilde{P} : |\sigma(x) - \sigma(\tilde{P})| < \epsilon\}$. Similarly define $\widetilde{bY_{jk}} = \{y \in \tilde{Q} : |\sigma(y) - \sigma(\tilde{Q})| < \epsilon\}$. By the Central Limit Theorem, $d^{\tilde{P}}_{a\tilde{X}_{ik}} \to 1$ and $d^{\tilde{Q}}_{b\tilde{Y}_{jk}} \to 1$ as $d \to \infty$. Also, since σ is a homomorphism, we have that $|\sigma(x + y) - \sigma(R)| > \epsilon$ for all $x \in \widetilde{aX_{ik}}$, $y \in \widetilde{bY_{jk}}$. So the Central Limit Theorem also says that $d^{\tilde{R}}_{a\tilde{X}_{ik}+b\tilde{Y}_{jk}} \to 0$ as $d \to \infty$. From this, it easily follows that for large enough d,

$$d^{\tilde{R}}_{a\tilde{X}_{ik}+b\tilde{Y}_{jk}} < d^{\tilde{P}}_{a\tilde{X}_{ik}} + d^{\tilde{Q}}_{b\tilde{Y}_{jk}} - 1.$$

Finally, pulling these sets back by the bijections ϕ_R, ϕ_P, and ϕ_Q, we get the desired inequality. $\square$

What we have shown in this section is that, for general d, unlike for $d = 1$, one cannot use a single mod-c residue class of the sumset $aX_i + bY_j$ to show that $d^{(cS)_{bj;ai;}}_{aX_i+bY_j} \geq d^{S_{;i;}}_{X_i} + d^{S_{j;;}}_{Y_j} - 1$.

3.6 A Counterexample to a Conjecture of Brualdi and Csima

In this final section, we give a counterexample to a conjecture of Brualdi and Csima regarding the support patterns of stochastic 3-tensors [BC75].

We begin by introducing some additional terminology. A *plane* of a 3-tensor T is a 2-dimensional pattern obtained by fixing one of the coordinates of $[n]^3$. A *plane section* of a pattern P is the intersection of P with a plane.

Given a plane section Γ of a pattern P, the associated *characteristic function* is $h : P \to \mathbb{R}$ given by $h(p) = \begin{cases} 1 & \text{if } p \in \Gamma \\ 0 & \text{otherwise.} \end{cases}$

Csima [Csi69] gave a characterization of stochastic patterns involving these characteristic functions.

Theorem 3.6.1. *[Csi69] Let $P \subset [n]^3$ be a pattern. Let $h_1, ..., h_{3n}$ be the characteristic functions of the plane sections of P. Then P is stochastic if and only if for all $(c_1, ..., c_{3n}) \in \mathbb{Z}^{3n}$ such that $\sum_{i=1}^{3n} c_i = 0$, the function $\sum_{i=1}^{3n} c_i h_i$ is either identically zero or assumes both positive and negative values.*

Brualdi and Csima [BC75] then conjectured that this statement still holds if the c_i are restricted to the set $\{-1, 0, 1\}$.

Conjecture 3.6.2. *[BC75] Let $S \subset [n]^3$ be a pattern. Let $h_1, ..., h_{3n}$ be the characteristic functions of the plane sections of S. Then S is stochastic if for all $(\epsilon_1, ..., \epsilon_{3n}) \in \{-1, 0, 1\}^{3n}$ such that $\sum_{i=1}^{3n} \epsilon_i = 0$, the function $\sum_{i=1}^{3n} \epsilon_i h_i$ is either identically zero or assumes both positive and negative values.*

In the same paper, Brualdi and Csima proved [BC75, Theorem 2.7] the matrix analogue of Conjecture 3.6.2.

Our counterexample to Conjecture 3.6.2 is a $5 \times 5 \times 5$ pattern, which we present as a subset of $[-2, 2]^3$ instead of $[5]^3$ for notational convenience. We will write x_i, y_i, z_i $(i = -2, -1, 0, 1, 2)$ to denote the characteristic functions of the plane sections in the x, y, and z directions respectively. Define the function $L = \sum_{i=-2}^{2} i(x_i + y_i + z_i)$, and let $T = \text{NSupp}(L) := \{(x, y, z) \in [-2, 2]^3 : L(x, y, z) \geq 0\}$ be the nonnegative support of L. Thus T is the set of integral points in the cube $[-2, 2]^3$ whose coordinates have nonnegative sum.

By Theorem 3.6.1 and construction of T, T is not a stochastic pattern. However, we have verified by exhaustive computer search that every $\{-1, 0, 1\}$-linear combination of the characteristic functions x_i, y_i, z_i with coefficients summing to zero is either identically zero on T or assumes both positive and negative values, in violation of Conjecture 3.6.2. Our code can be found at [HY23].

Bibliography

[BC75] R.A. Brualdi and J. Csima. "Stochastic Patterns". In: *Journal of Combinatorial Theory, Series A* 19 (1 1975), pp. 1–12. DOI: 10.1016/0097-3165(75)90088-6.

[Ber+04] N. Berry et al. "The conjugate dimension of algebraic numbers". In: *The Quarterly Journal of Mathematics* 55.3 (2004), pp. 237–252. DOI: 10.1093/qmath/hah003.

[Bur+18] P. Burgisser et al. "Alternating minimization, scaling algorithms, and the null-cone problem from invariant theory". In: *Proceedings of the Innovations in Theoretical Computer Science (ITCS).* 2018, 24:1–24:20. DOI: https://doi.org/10.4230/LIPIcs.ITCS.2018.24.

[CLP17] E. Croot, V. F Lev, and P. P. Pach. "Progression-free sets in $\mathbf{Z}_4^n$ are exponentially small". In: *Ann. of Math.* 185 (2017), pp. 000–000. DOI: 10.4007/annals.2017.185.1.

[Csi69] J. Csima. "Stochastic Functions on Hypergraphs". In: *Proceedings of Colloquia Mathematics Societatis János Bolyai* 4: Combinatorial Theory and Its Applications (1969), pp. 247–255.

[CVZ18] Matthias Christandl, Péter Vrana, and Jeroen Zuiddam. "Universal points in the asymptotic spectrum of tensors". In: June 2018, pp. 289–296. DOI: 10.1145/3188745.3188766.

[Dix97] J. Dixon. "Polynomials with nontrivial relations between their roots". In: *Acta Arithmetica* 82 (1997), pp. 293–302. DOI: 10.4064/aa-82-3-293-302.

[EG17] J. Ellenberg and D. Gijswijt. "On large subsets of $\mathbb{F}_q^n$ with no three-term arithmetic progression". In: *Annals of Mathematics* 185 (1 2017), pp. 339–343.

[Ell14] J. Ellenberg. *Arizona Winter School 2014 Notes: Geometric Analytic Number Theory.* Online; accessed 21-July-2023. 2014. URL: https://swc-math.github.io/aws/2014/2014EllenbergNotes.pdf.

[Gar85] P. A. Gartenberg. "Fast Rectangular Matrix Multiplication (Algebraic Computational, Asymptotic Complexity)". PhD thesis. 1985, p. 121. ISBN: 979-8-204-48818-2. URL: https://ezproxy.library.wisc.edu/login?url=https://www.proquest.com/dissertations-theses/fast-rectangular-matrix-multiplication-algebraic/docview/303332114/se-2.

[GM17] Joshua A. Grochow and Cristopher Moore. *Designing Strassen's algorithm.* arXiv:1708.09398 [cs.DS]. 2017.

[HY23] W. Hardt and J. Yin. *Stochastic Patterns Counterexample.* https://github.com/whardt/Stochastic-Patterns-Counterexample. GitHub repository. 2023.

[Man65] H.B. Mann. "On linear relations between roots of unity". In: *Mathematika* 12 (1965), pp. 107–117. DOI: 10.1112/s0025579300005210.

[Poo06] B. Poonen. *Lectures on rational points on curves*. Online; accessed 21-July-2023. 2006. URL: http://www-math.mit.edu/~poonen/papers/curves.pdf.

[Smy86] C.J. Smyth. "Additive and multiplicative relations connecting conjugate algebraic numbers". In: *Journal of Number Theory* 23.2 (1986), pp. 243–254. DOI: 10.1016/0022-314x(86)90094-6.

[Spe] David E Speyer. *Tell me an algebraic integer that isn't an eigenvalue of the sum of two permutations*. MathOverflow. URL:https://mathoverflow.net/q/264035 (version: 2017-03-09). eprint: https://mathoverflow.net/q/264035. URL: https://mathoverflow.net/q/264035.

[Str69] V. Strassen. "Gaussian elimination is not optimal". In: *Numerische Mathematik* 13 (1969), pp. 354–356. DOI: doi:10.1007/BF02165411.

[Str88] V. Strassen. "The asymptotic spectrum of tensors". In: *J. reine angew.* 384 (1988), pp. 102–152. DOI: doi:10.1515/crll.1988.384.102.

[Tao16] T. Tao. *A symmetric formulation of the Croot-Lev-Pach-Ellenberg-Gijswijt capset bound*. Online; accessed 1-July-2023. 2016. URL: https://terrytao.wordpress.com/2016/05/18/a-symmetric-formulation-of-the-croot-lev-pach-ellenberg-gijswijt-capset-bound/.

[TS16] T. Tao and W. Sawin. *Notes on the "slice rank" of tensors*. Online; accessed 1-July-2023. 2016. URL: https://terrytao.wordpress.com/2016/08/24/notes-on-the-slice-rank-of-tensors/#stom.

[TV06] Terence Tao and Van H. Vu. *Additive Combinatorics*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2006. DOI: 10.1017/CBO9780511755149.

[Zan95] U. Zannier. "Vanishing sums of roots of unity". In: *Rend. Sem. Mat. Univ. Pol. Torino* 53, 4 (1995), pp. 487–495.