Uniwersytet Warszawski
Wydział Matematyki, Informatyki i Mechaniki

Filip Rupniewski
Nr albumu: 320330

Tensors of border rank 3 and Strassen's additivity conjecture.

Praca magisterska
na kierunku MATEMATYKA

Praca wykonana pod kierunkiem
dr hab. Jarosława Buczyńskiego
Instytut Matematyki

Styczeń 2017

Oświadczenie kierującego pracą

Potwierdzam, że niniejsza praca została przygotowana pod moim kierunkiem i kwalifikuje się do przedstawienia jej w postępowaniu o nadanie tytułu zawodowego.

Data Podpis kierującego pracą

Oświadczenie autora (autorów) pracy

Świadom odpowiedzialności prawnej oświadczam, że niniejsza praca dyplomowa została napisana przeze mnie samodzielnie i nie zawiera treści uzyskanych w sposób niezgodny z obowiązującymi przepisami.

Oświadczam również, że przedstawiona praca nie była wcześniej przedmiotem procedur związanych z uzyskaniem tytułu zawodowego w wyższej uczelni.

Oświadczam ponadto, że niniejsza wersja pracy jest identyczna z załączoną wersją elektroniczną.

Data Podpis autora (autorów) pracy

Abstract

In this thesis I prove that Strassen's additivity conjecture holds if one of tensors is of border rank 3 and as a corollary I show families of tensors of rank 4 and 5 for which the conjecture holds as well.

Słowa kluczowe

Strassen's additivity theorem, rank of a tensor, border rank of a tensor, The Alexeev-Forbes-Tsimerman method

Dziedzina pracy (kody wg programu Socrates-Erasmus)

11.1 Matematyka

Klasyfikacja tematyczna

14-xx Algebraic geometry
14A10 Varieties and morphisms
15A69 Multilinear algebra, tensor products
15A03 Vector spaces, linear dependence, rank

Tytuł pracy w języku angielskim

Tensors of border rank 3 and Strassen's additivity conjecture.

Contents

Chapter 1

Introduction

Strassen in 1969 showed that it is possible to multiply two 2×2 matrices using seven basic operations rather than eight, see 2.1, and this was proved to have the best possible computational complexity by Winograd in [W]. Using this fact, a better algorithm was produced to multiply matrices of any size.

After Strassen's result, it was clear that even straightforward procedures can require fewer operations than expected. In [S73] Strassen formulated his well known additivity conjecture for bilinear maps: Given bilinear maps φ, ψ and two pairs of matrices M_1, M_2 and M_1', M_2' the computational complexity of simultaneously computing $\varphi(M_1, M_2)$ and $\psi(M_1', M_2')$ is the sum of the complexities of φ and ψ. The conjecture stands open since its formulation in 1973.

Strassen's conjecture can be naturally stated in terms of tensors and the notion of tensor rank, see Conjecture 2.2.2. Note that an analogue of the additivity conjecture for approximate complexity (border rank, in more recent terminology) does not hold (see Example 2.4.1 by Schönhage). For definitions of tensor rank and tensor border rank see Definition 2.1.6 and Definition 2.1.10.

Also the symmetric version of the conjecture stands open. A relevant contribution to its study is the 2012 paper [CCG] where Strassen's additivity conjecture is proved for the sum of (several) monomials.

In this thesis I prove that Strassen's additivity conjecture holds if one of tensors is of border rank 3 (Theorem 3.3.3) and as a corollary I show families of tensors of rank 4 and 5 for which the conjecture holds as well (Corollary 3.3.4).

There is a theorem which summarize cases discussed in the thesis in which the Strassen's additivity conjecture holds.

Theorem 1.0.1. *Let $A_1, A_2, B_1, B_2, C_1, C_2$ be finite dimensional vector spaces, $T_1 \in A_1 \otimes B_1 \otimes C_1$ and $T_2 \in A_2 \otimes B_2 \otimes C_2$ be two tensors. Then the Strassen's additivity conjecture holds for T_1, T_2, i.e. :*

$$R(T_1 \oplus T_2) = R(T_1) + R(T_2)$$

(where R denotes tensor rank and $T_1 \oplus T_2 \in (A_1 \oplus A_2) \otimes (B_1 \oplus B_2) \otimes (C_1 \oplus C_2)$) if one of the following conditions is fulfilled:

1. *dimension of at least one of $A_1, A_2, B_1, B_2, C_1, C_2$ equals 2,*
2. *T_1 is A_1-concise, i.e. T_1 cannot be written as a tensor in a smaller space $A_1' \subsetneq A_1$ (see Definition 3.1.2) and such that $R(T_1) = \dim(A_1)$ (analogous statements hold for permuted situations),*
3. *there exists i such that the border rank $\underline{R}(T_i) = 3$,*

4. at least one of tensors T_1, T_2 is of rank 4 and has form

$$a_1 \otimes (b_1 \otimes c_2 + b_2 \otimes c_1) + a_2 \otimes b_1 \otimes c_1 + a_3 \otimes b_3 \otimes c_3,$$

5. at least one of tensors T_1, T_2 is of rank 5 and has one of the forms:

- $a_1 \otimes (b_1 \otimes c_3 + b_2 \otimes c_2 + b_3 \otimes c_1) + a_2 \otimes (b_1 \otimes c_2 + b_2 \otimes c_1) + a_3 \otimes b_1 \otimes c_1,$
- $a_1 \otimes (b_1 \otimes c_2 + b_2 \otimes c_1) + a_2 \otimes b_1 \otimes c_1 + a_3 \otimes (b_3 \otimes c_1 + b_1 \otimes c_3),$
- $a_1 \otimes (b_1 \otimes c_2 + b_2 \otimes c_1 + b_3 \otimes c_3) + a_2 \otimes b_1 \otimes c_1 + a_3 \otimes b_3 \otimes c_1,$
- $a_1 \otimes (b_1 \otimes c_2 + b_2 \otimes c_1 + b_2 \otimes c_3) + a_2 \otimes b_1 \otimes c_1 + a_3 \otimes b_1 \otimes c_3.$

The point 1 is a result of Ja'Ja' and Takche [JT]. It is also discussed in [LM] Section 4, I give a detailed proof in Proposition 3.2.12. The point 2 is a consequence of results of [L] Theorem 10.3.3.3 and we discuss it in Chapter 3. Points 3-5 are a result of this thesis which use the Alexeev-Forbes-Tsimerman method from article [LM], see Subsection 3.3, Theorem 3.3.3 and Corollary 3.3.4. Note that points 4 and 5 are special cases of point 3 (or point 1 if there is a linear dependence), by Theorem 3.3.1.

Given 1 and 2, the smallest unsolved case of the conjecture is $(A_1 \oplus A_2) \otimes (B_1 \oplus B_2) \otimes (C_1 \oplus C_2) = (\mathbb{C}^3 \oplus \mathbb{C}^3) \otimes (\mathbb{C}^3 \oplus \mathbb{C}^3) \otimes (\mathbb{C}^3 \oplus \mathbb{C}^3)$. Our result 3 (and 4, 5), in particular contributes to this case.

Acknowledgments

I would like to thank my advisor, Jarosław Buczyński and Mateusz Michałek for many helpful discussions that broaden my understanding of the related material.

Chapter 2

Preliminaries

2.1. Ranks and secant varieties

Notation 2.1.1. All *vector spaces* are presumed to be finite-dimensional vector spaces over the complex field $\mathbb{C}$. Letters A, B, C possibly indexed denote vector spaces. The asterisk A^* denotes the dual of a space A, so the space of linear maps $A \to \mathbb{C}$. By a *variety* I always mean an algebraic variety over $\mathbb{C}$.

Notation 2.1.2. For a subset $S \subset A$, by $\langle S \rangle$ we denote the linear space spanned by S.

Definition 2.1.3. For a linear space A, we define a projective space $\mathbb{P}A = (A \setminus \{0\})/\mathbb{C}^*$, where $\mathbb{C}^* = \mathbb{C} \setminus \{0\}$ is the multiplication group of the field $\mathbb{C}$. If $\hat{X} \subset A$ is invariant with respect to rescaling $\forall_{\lambda \in \mathbb{C}} (x \in \hat{X} \Rightarrow \lambda x \in \hat{X})$, we define $X = \mathbb{P}(\hat{X}) \subset \mathbb{P}A$.

Definition 2.1.4. Let us consider an invariant with respect to rescaling subset $\hat{X} \subset A$ spanning A as a linear space. Let $p \in A$. We define $\hat{X}$*-rank* of p as the least integer $r = R_{\hat{X}}(p)$, such that:

$$p = \sum_{i=1}^{r} \lambda_i \hat{x_i} \text{ for certain } \hat{x_i} \in \hat{X} \text{ and } \lambda_i \in \mathbb{C}.$$

Equivalently, r is the minimal integer such that $p \in \langle\{\hat{x_1}, \hat{x_2}, ..., \hat{x_r}\}\rangle$, for certain $\hat{x_i} \in \hat{X}$.

Similarly for $X = \mathbb{P}(\hat{X})$ we define X*-rank* of $p \in \mathbb{P}A$ as the least integer $r = R_X(p)$, such that $p \in \langle\{x_1, x_2, ..., x_r\}\rangle$, for certain $x_i \in X$. Here $\langle\{x_1, x_2, ..., x_r\}\rangle$ denotes the smallest linear subspace $\mathbb{P}^k \subset \mathbb{P}A$ containing $\{x_1, x_2, ..., x_r\}$.

Definition 2.1.5. For $T \in A_1 \otimes A_2 \otimes ... \otimes A_n$, T is a *simple tensor* if there exist elements $a_1, a_2, ..., a_n$ such, that $\forall_i\ a_i \in A_i$ and $T = a_1 \otimes a_2 \otimes ... \otimes a_n$.

Definition 2.1.6. Let $A = A_1 \otimes \cdots \otimes A_k$, $\hat{X}$ be a set of simple tensors, $X = \mathbb{P}(A_1) \times \mathbb{P}(A_2) \times \cdots \times \mathbb{P}(A_k)$. Then for $x = ([a_1], ..., [a_k]) \in X$, we have an inclusion $Seg : X \to \mathbb{P}(A_1 \otimes \cdots \otimes A_k)$, given by $Seg(x) = [a_1 \otimes \cdots \otimes a_k]$. This inclusion is called *the Segre's embedding*. For such X and $\hat{X}$ we define:

- in the affine case, a *tensor rank* of $\hat{p} \in A$ is $R_{\hat{X}}(\hat{p})$,
- in the projective case, a *tensor rank* of $p \in \mathbb{P}(A)$ is $R_{Seg(X)}(p)$.

The tensor rank of a tensor p will be denoted by $R(p)$ if there is no risk of confusion. We will write often just *rank* in place of *tensor rank*.

Similarly we can define the symmetric tensor rank.

Definition 2.1.7. A tensor $T \in A^{\otimes d}$ is a *symmetric tensor* of order d if for every permutation σ of symbols $\{1, 2, ..., d\}$ and every subset of vectors $\{v_1, v_2, ..., v_d\} \in A^d$, $T(v_1, v_2, ..., v_d) = T(v_{\sigma(1)}, v_{\sigma(2)}, ..., v_{\sigma(d)})$. The set of symmetric tensors of order d is denoted by $S^d(A)$.

Definition 2.1.8. For $[l] = x \in X \subset \mathbb{P}(A)$ we have an inclusion $v_d : X \to \mathbb{P}(S^d A)$, given by $v_d(x) = [l \otimes \cdots \otimes l]$. This inclusion is called *the Veronese's embedding of degree d*. We define the *symmetric rank* of $p \in S^d A$ as the $v_d(X)$-rank of this point (see Definition 2.1.4).

Definition 2.1.9. $\langle S \rangle$ denotes space spanned by elements of S like in Notation 2.1.2. We define the *r-th secant variety* of a variety $\hat{X} \subset A$ as the Zariski closure of the sum of all linear subspaces spanned by r points of $\hat{X}$:

$$\sigma_r(\hat{X}) := \overline{\bigcup \{\langle \hat{x_1}, ..., \hat{x_r} \rangle : \hat{x_i} \in \hat{X}\}} \subset A.$$

Equivalently the r-th secant variety of $\hat{X}$ is the Zariski closure of the set of points of $\hat{X}$-rank at most r.

Similarly we define the r-th secant variety of a projective algebraic variety $X \subset \mathbb{P}A$. Here $\langle S \rangle$ denotes the projective space spanned by elements of S, i.e. the smallest linear subspace $\mathbb{P}^k \subset \mathbb{P}A$ which contains S. The *r-th secant variety* of a projective algebraic variety $X \subset \mathbb{P}A$ is the Zariski closure of the sum of all linear subspaces spanned by r points of X:

$$\sigma_r(X) := \overline{\bigcup \{\langle x_1, ..., x_r \rangle : x_i \in X\}} \subset \mathbb{P}A.$$

The secant variety is an Zariski closure of a set of points of X-rank at most r.

Definition 2.1.10. For $\hat{p} \in A, \hat{X} \subset A$ (resp. $p \in \mathbb{P}A, X \subset \mathbb{P}A$) we define $\underline{R}_{\hat{X}}(\hat{p})$, the *$\hat{X}$-border rank* of a point $\hat{p}$ (resp. $\underline{R}_X(p)$, the *X-border rank* of a point p) as the minimal number r, such that $\hat{p} \in \sigma_r(\hat{X})$ (resp. $p \in \sigma_r(X)$).

For $A, \hat{p}, p, \hat{X}, X$ and $Seg(X)$ as in Definition 2.1.6, we define the *tensor border rank* of $\hat{p} \in A$ as $\underline{R}_{\hat{X}}(\hat{p})$ (resp. *tensor border rank* of $p \in \mathbb{P}(A)$ as $\underline{R}_{Seg(X)}(p)$). The tensor border rank of a tensor p will be denoted by $\underline{R}(p)$ if there is no risk of confusion. We will write often just *border rank* in place of *tensor border rank*.

Example 2.1.11. For a tensor $p \in A \otimes B \approx Hom(A^* \to B)$, the tensor rank of p is equal the rank of matrix corresponding to tensor p. The set of tensors in $A \otimes B$ of rank at most r is closed, so for tensors in $A \otimes B$ the tensor rank equals the tensor border rank.

There are fundamental properties of the $\hat{X}$-rank, for $\hat{X} \subset A$ invariant under rescalings:

1. $\underline{R}_{\hat{X}}(p) = 0 \Leftrightarrow R_{\hat{X}}(p) = 0 \Leftrightarrow p = 0$ (convention),
2. $R_{\hat{X}}(p) = 1 \Leftrightarrow p \in \hat{X} \setminus \{0\}$,
3. $\hat{Y} \subset \hat{X}, p \in \langle \hat{Y} \rangle$, then $R_{\hat{Y}}(p) \geq R_{\hat{X}}(p)$,
4. $R_{\hat{X}}(\lambda p) = R_{\hat{X}}(p)$, for any $\lambda \in \mathbb{C}^*$,
5. $R_{\hat{X}}(p_1 + p_2) \leq R_{\hat{X}}(p_1) + R_{\hat{X}}(p_2)$.

2.2. The matrix multiplication

The matrix multiplication is a bilinear map $\mathbb{C}^{fg} \times \mathbb{C}^{gh} \to \mathbb{C}^{fh}$. Hence one can think about it as a tensor:

$$\mathcal{M}_{f,g,h} \in (\mathbb{C}^{fg})^* \otimes (\mathbb{C}^{gh})^* \otimes \mathbb{C}^{fh} = A \otimes B \otimes C.$$

The naive matrix multiplication algorithm needs fgh multiplications of complex numbers and as a tensor, has a form:

$$\mathcal{M}_{f,g,h} = \sum_{i,j,k} a^*_{i,j} \otimes b^*_{j,k} \otimes c_{i,k},$$

where the sum goes by $1 \leq i \leq f, 1 \leq j \leq g, 1 \leq k \leq h$, and elements $a_{i,j} \in \mathbb{C}^{fg}, b_{j,k} \in \mathbb{C}^{gh}, c_{i,k} \in \mathbb{C}^{fh}$ make the standard bases.

This algorithm gives an estimate for the tensor rank $R(\mathcal{M}_{f,g,h}) \leqslant fgh$. Strassen proved that one can multiply two matrices 2×2 using only 7 multiplications of complex numbers (instead of $2 \cdot 2 \cdot 2 = 8$ multiplications) [S69]

$$\begin{aligned}\mathcal{M}_{2,2,2} = {} & (a^*_{1,1} + a^*_{2,2}) \otimes (b^*_{1,1} + b^*_{2,2}) \otimes (c_{1,1} + c_{2,2}) \\ & + (a^*_{2,1} + a^*_{2,2}) \otimes b^*_{1,1} \otimes (c_{2,1} - c_{2,2}) \\ & + a^*_{1,1} \otimes (b^*_{1,2} - b^*_{2,2}) \otimes (c_{2,1} + c_{2,2}) \\ & + a^*_{2,2} \otimes (-b^*_{1,1} + b^*_{2,1}) \otimes (c_{2,1} + c_{2,2}) \\ & + (a^*_{1,1} + a^*_{1,2}) \otimes b^*_{2,2} \otimes (-c_{1,1} + c_{2,1}) \\ & + (-a^*_{1,1} + a^*_{2,1}) \otimes (b^*_{1,1} + b^*_{1,2}) \otimes c_{2,2} \\ & + (a^*_{1,2} - a^*_{2,2}) \otimes (b^*_{2,1} + b^*_{2,2}) \otimes c_{1,1}.\end{aligned} \tag{2.1}$$

Strassen asked if there exists an algorithm that simultaneously computes two different matrix multiplications that costs less than the sum of the best algorithms for the individual matrix multiplications. If not, one says that additivity holds for matrix multiplication.

In [S73] Strassen formulated his well known additivity conjecture for bilinear maps: Given bilinear maps φ, ψ and two pairs of matrices M_1, M_2 and M'_1, M'_2 the computational complexity of simultaneously computing $\varphi(M_1, M_2)$ and $\psi(M'_1, M'_2)$ is the sum of the complexities of φ and ψ. The conjecture stands open since its formulation in 1973.

Strassen's conjecture can be naturally stated in terms of tensors and the notion of tensor rank.

Definition 2.2.1. For two tensors $p_1 = \sum_{i=1}^r a_i \otimes b_i \otimes c_i \in A_1 \otimes B_1 \otimes C_1$ and $p_2 = \sum_{i=1}^{r'} a'_i \otimes b'_i \otimes c'_i \in A_2 \otimes B_2 \otimes C_2$ we define the *direct sum of tensors* as a tensor $p_1 \oplus p_2 = \sum_{i=1}^r a_i \otimes b_i \otimes c_i + \sum_{i=1}^{r'} a'_i \otimes b'_i \otimes c'_i \in (A_1 \oplus A_2) \otimes (B_1 \oplus B_2) \otimes (C_1 \oplus C_2)$.

Conjecture 2.2.2 (Strassen's additivity conjecture (SAC))**.** *Additivity holds for bilinear maps. That is, given $T_1 \in A_1 \otimes B_1 \otimes C_1$ and $T_2 \in A_2 \otimes B_2 \otimes C_2$, then letting $A = A_1 \oplus A_2, B = B_1 \oplus B_2, C = C_1 \oplus C_2$, we have:*

$$R_{Seg(\mathbb{P}A \times \mathbb{P}B \times \mathbb{P}C)}(T_1 \oplus T_2) = R_{Seg(\mathbb{P}A_1 \times \mathbb{P}B_1 \times \mathbb{P}C_1)}(T_1) + R_{Seg(\mathbb{P}A_2 \times \mathbb{P}B_2 \times \mathbb{P}C_2)}(T_2).$$

Note that an analogue of the additivity conjecture for border ranks does not hold (see Example 2.4.1).

In the Chapter 3. we will focus on the 3-way tensors T contained in the space $A \otimes B \otimes C$. In the tensor literature, such tensors are often studied by their images $T(A^*) \subset B \otimes C$ etc. and these images are studied in terms of bases, resulting in a parametrized subspace of a space of matrices.

Definition 2.2.3. For a tensor T contained in the space $A \otimes B \otimes C$, and a given bases of A, B, C, the images $T(A^*) \subset B \otimes C$ etc. are parametrized spaces of matrices called *slices*.

2.3. The join of varieties and the secant varieties

Definition 2.3.1. For a vector space A and $x, y \in \mathbb{P}A$ such that $x \neq y$, let $\mathbb{P}^1_{xy}$ denote the projective line $l \subset \mathbb{P}A$ containing x and y.

Definition 2.3.2. The *join* of two algebraic varieties $Y, Z \subset \mathbb{P}A$ is

$$J(Y,Z) = \overline{\bigcup_{y \in Y, z \in Z, y \neq z} \mathbb{P}^1_{yz}}.$$

The join of k varieties $X_1, ..., X_k \subset \mathbb{P}A$ is defined to be the closure of the union of the corresponding $\mathbb{P}^{k-1}$, or by induction, $J(Y_1, ..., Y_k) = J(Y_1, J(Y_2, ..., Y_k))$.

For $Y = Z$, $J(Y,Y) = \sigma_2(Y)$ denotes *the second secant variety of* Y. If $\forall_{1 \leq k \leq r-1}\ Y_k = Y_{k+1}$ the join $J(Y_1, ..., Y_r) = J(Y, ..., Y) = \sigma_r(Y)$ is the *the r-th secant variety of* Y. [L] Example 5.1.1.3 shows that for $Y = Seg(\mathbb{P}A_1 \times \mathbb{P}A_2 \times \mathbb{P}A_3)$, such defined $\sigma_r(Y)$ agrees with the previous notation (Definition 2.1.9) of the set of tensors in $\mathbb{P}(A_1 \otimes A_2 \otimes A_3)$ of border rank at most r.

Definition 2.3.3. The *expected dimension* of $J(Y,Z) \subset \mathbb{P}A$ is $\min\{\dim Y + \dim Z + 1, \dim \mathbb{P}A\}$ because a point $x \in J(Y,Z)$ is obtained by picking a point of Y, a point of Z, and a point on the line joining the two points. Similarly, the expected dimension of $\sigma_r(Y)$ is $\min\{r(\dim F) + r - 1, \dim \mathbb{P}A\}$.

Definition 2.3.4. If $X^n \subset \mathbb{P}^N$ and $\dim \sigma_r(X) < \min\{rn + r - 1, N\}$, one says $\sigma_r(X)$ is defective, with defect $\delta_r = \delta_r(X) := rn + r - 1 - \dim \sigma_r(X)$. Otherwise one says $\sigma_r(X)$ is nondefective.

Example 2.3.5 ([L] Example 5.1.2.2)**.** Let $a = \dim(A), b = \dim(B)$ and $a, b \geq 3$. Consider $\sigma_2(Seg(\mathbb{P}A \times \mathbb{P}B))$. An open set of this variety may be parametrized as follows: choose bases for A, B and write

$$x_1 = \begin{bmatrix} x_1^1 \\ \vdots \\ x_1^a \end{bmatrix}, x_2 = \begin{bmatrix} x_2^1 \\ \vdots \\ x_2^a \end{bmatrix} \tag{2.2}$$

Choose the column vectors x_1, x_2 arbitrarily and then take the matrix

$$p = (x_1, x_2, c_1^3 x_1 + c_2^3 x_2, ..., c_1^b x_1 + c_2^b x_2)$$

to get a general matrix of rank at most two. Thus the set of matrices of rank at most two in $A \otimes B$, denoted by $\hat{\sigma_2}(Seg(\mathbb{P}A \times \mathbb{P}B))$ is locally parametrized by $2a + 2(b-2) = 2a + 2b - 4$ parameters. Hence $\dim \sigma_2(Seg(\mathbb{P}A \times \mathbb{P}B)) = 2a + 2(b-2) - 1 = 2a + 2b - 5$ compared with the expected $2[(a-1) + (b-1)] + 1 = 2a + 2b - 3$, so $\delta_2(Seg(\mathbb{P}A \times \mathbb{P}B)) = 2$.

Example 2.3.6 ([L] Exercise 5.1.2.4)**.** Let $a = \dim(A), b = \dim(B)$ and $a, b \geq 3$ for $\sigma_r(Seg(\mathbb{P}A \times \mathbb{P}B))$, $r \leq \min\{a, b\}$. Choose the column vectors $x_1, x_2, ..., x_r$ arbitrarily and then take the matrix

$$p = (x_1, x_2, ..., x_r, \sum_{i=1}^{r} c_i^{r+1} x_i, ..., \sum_{i=1}^{r} c_i^{r+1} x_i).$$

Thus $\dim \sigma_r(Seg(\mathbb{P}A \times \mathbb{P}B)) = ra + r(b-r) - 1 = r(a+b-r) - 1$.

One expects $\sigma_7(Seg(\mathbb{P}^3 \times \mathbb{P}^3 \times \mathbb{P}^3)) = \mathbb{P}(\mathbb{C}^4 \otimes \mathbb{C}^4 \otimes \mathbb{C}^4)$, and this is indeed the case, which explains how Strassen's algorithm for multiplying 2×2 matrices could have been anticipated (see [L] Section 5.2.2).

2.4. Schönhage example

Schönhage showed that border rank version of Strassen's conjecture (Conjecture 2.2.2) cannot hold in general. I quote the following example and proposition from [BGL] because I believe it is good to see the counterexample in details.

Example 2.4.1 (Schönhage's example ([BGL] Example 4.5.2))**.** Let

$$\mathcal{M}_{a,b,c} : (\mathbb{C}^a \otimes \mathbb{C}^{b*}) \times (\mathbb{C}^b \otimes \mathbb{C}^{c*}) \times (\mathbb{C}^a \otimes \mathbb{C}^{c*})$$

denote the matrix multiplication operator. Schönhage proved [Sch] that, while $\underline{\mathrm{R}}(\mathcal{M}_{e,1,l}) = el$ and $\underline{\mathrm{R}}(\mathcal{M}_{1,h,1}) = h$, nevertheless, by Proposition 2.4.2, we know that $\underline{\mathrm{R}}(\mathcal{M}_{e,1,l} \oplus \mathcal{M}_{1,h,1}) \leq el + 1$ so for $h > 1$ the border rank Strassen's conjecture does not hold.

Proposition 2.4.2 ([BGL] Proposition 4.5.3)**.** *Let* $h = (e-1)(l-1)$*, then* $\underline{\mathrm{R}}(\mathcal{M}_{e,1,l} \oplus \mathcal{M}_{1,h,1}) \leq el + 1$.

Proof. Write $A = A' \oplus A'', B = B' \oplus B'', C = C' \oplus C''$, where $A', A'', B', B'', C', C''$ are linear spaces. Let $\dim A' = e, \dim B' = l, \dim C' = el, \dim A'' = h, \dim B'' = h$, and $\dim C'' = 1$. Fix index ranges $1 \leq i \leq e, 1 \leq s \leq l, 1 \leq u \leq e-1, 1 \leq v \leq l-1$. Give the vector spaces bases $\{a_i\}, \{b_s\}, \{c_{i,s}\}, \{\alpha_{u,v}\}, \{\beta_{u,v}\}$, and γ.

In the above bases,

$$\mathcal{M}_{e,1,l} = \sum_{1 \leq i \leq e, 1 \leq s \leq l} a_i \otimes b_s \otimes c_{i,s} \text{ and } \mathcal{M}_{1,h,1} = \sum_{1 \leq u \leq e-1, 1 \leq v \leq l-1} \alpha_{u,v} \otimes \beta_{u,v} \otimes \gamma.$$

Set $\alpha_{e,v} = -\sum_u \alpha_{u,v}, \beta_{u,l} = -\sum_v \beta_{u,v}, \alpha_{i,l} = 0, \beta_{e,j} = 0$, which can be illustrated with the following picture:

$$(\alpha_{u,v})_{\substack{u=1,...,e \\ v=1,...,l}} = \begin{bmatrix} \alpha_{1,1} & \alpha_{1,2} & \cdots & \alpha_{1,l-1} & 0 \\ \alpha_{2,1} & \alpha_{2,2} & \cdots & \alpha_{2,l-1} & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \alpha_{e-1,1} & \alpha_{e-1,2} & \cdots & \alpha_{e-1,l-1} & 0 \\ -\sum_u \alpha_{u,1} & -\sum_u \alpha_{u,2} & \cdots & -\sum_u \alpha_{u,l-1} & 0 \end{bmatrix},$$

$$(\beta_{u,v})_{\substack{u=1,...,e \\ v=1,...,l}} = \begin{bmatrix} \beta_{1,1} & \beta_{1,2} & \cdots & \beta_{1,l-1} & -\sum_v \beta_{1,v} \\ \beta_{2,1} & \beta_{2,2} & \cdots & \beta_{2,l-1} & -\sum_v \beta_{2,v} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \beta_{e-1,1} & \beta_{e-1,2} & \cdots & \beta_{e-1,l-1} & -\sum_v \beta_{1,v} \\ 0 & 0 & \cdots & 0 & 0 \end{bmatrix}.$$

Now let

$$\tilde{T}(t) = \sum_{1\le i\le e, 1\le s\le l} (a_i + t\alpha_{i,s}) \otimes (b_s + t\beta_{i,s}) \otimes (\gamma + t^2 c_{i,s}) - (\sum_i a_i) \otimes (\sum_s b_s) \otimes \gamma.$$

Note that for $t \neq 0, R(\tilde{T}(t)) \le el + 1$, and that $\tilde{T}(t) = t^2 T + O(t^3)$ where $T = \mathcal{M}_{e,1,l} \oplus \mathcal{M}_{1,h,1}$. Hence

$$T(t) := \frac{1}{t^2}\tilde{T}(t) = T + O(t)$$

is the desired sequence of tensors, i.e. $\forall_{0\neq t}\ R(T(t)) \le el + 1$, and in the limit we have $T(0) = T$. ■

Chapter 3

Lower bounds for the tensor rank and the main theorem

3.1. Lower bounds for the tensor rank of a sum of tensors $p_1 \oplus p_2$, when p_1 is A_1-concise.

Proposition 3.1.1 ([L] Proposition 3.1.3.1)**.** *Let tensor $T \in A \otimes B \otimes C$ has rank r. Say $T \in A' \otimes B' \otimes C'$, where $A' \subseteq A$, $B' \subseteq B$, $C' \subseteq C$ with at least one inclusion proper, let it be $A' \subsetneq A$. Then any expression $T = \sum_{i=1}^{s} a_i \otimes b_i \otimes c_i$ such that $\exists_i\ a_i \notin A'$ has $s > r$.*

Definition 3.1.2. Tensor $p \in A \otimes B \otimes C$ is *A-concise* if the following implication holds:

$$\exists_{A' \subseteq A}\ p \in A' \otimes B \otimes C \Rightarrow A' = A.$$

If p is A-concise, B-concise and C-concise we say it is *concise.*

Proposition 3.1.3 ([L] Theorem 10.3.3.3.)**.** *If $p_2 \in A_2 \otimes B_2 \otimes C_2$ and $p_1 \in A_1 \otimes B_1 \otimes C_1$ is a A_1-concise tensors, then $R(p_1 \oplus p_2) \geq \dim(A_1) + R(p_2)$.*

Proof. There are presentations: $p_1 = \sum_{i=1}^{r_1} a'_i \otimes b'_i \otimes c'_i$, $p_2 = \sum_{i=1}^{r_2} a''_i \otimes b''_i \otimes c''_i$, $p_1 + p_2 = \sum_{i=1}^{s} a_i \otimes b_i \otimes c_i$, where $r_1 = R(p_1), r_2 = R(p_2), s = R(p_1 \oplus p_2)$. We have the projections $\pi_{A_1} : (A_1 \oplus A_2) \to A_1$, $\pi_{B_2} : (B_1 \oplus B_2) \to B_2$.

From the definition of the A_1-conciseness we get $\dim((p_1 \oplus p_2)(A_1^*)) = \dim(p_1(A_1^*)) = \dim(A_1)$. Thus there exists a subset $S = \{\tilde{a}_1, \ldots, \tilde{a}_{\mathbf{a_1}}\} \subset \{a_1, \ldots, a_n\}$ of cardinality $\mathbf{a_1} = \dim(A_1)$, such that $\pi_{A_1}(\tilde{A_1}) = \langle \pi_{A_1}(\tilde{a}_1), \ldots, \pi_{A_1}(\tilde{a}_{\mathbf{a_1}})\rangle = A_1$, where $\tilde{A_1} = \langle \tilde{a}_1, \ldots, \tilde{a}_{\mathbf{a_1}}\rangle$ is the linear space spanned by elements of S. After a possible permutation of the a_i's, we can forget about tildes, so further we denote $S = \{a_1, \ldots, a_{\mathbf{a_1}}\}$. $A_2 = \ker(\pi_{A_1})$, so $A_2 \cap \tilde{A_1} = 0$. Hence when we define the quotient map $\varphi : (A_1 \oplus A_2) \to (A_1 \oplus A_2)/\tilde{A_1}$, then after restriction to A_2 we have an isomorphism $\varphi|_{A_2} : A_2 \to (A_1 \oplus A_2)/\tilde{A_1}$.

We can define:

$$\varphi^{\#} : (A_1 \oplus A_2) \otimes (B_1 \oplus B_2) \otimes (C_1 \oplus C_2) \to ((A_1 \oplus A_2)/\tilde{A_1}) \otimes B_2 \otimes (C_1 \oplus C_2)$$

$$\varphi^{\#} = \varphi \otimes \pi_{B_2} \otimes Id_{(C_1 \oplus C_2)}.$$

We see that $\varphi^{\#}|_{A_2 \otimes B_2 \otimes C_2}$ is an isomorphism, we define $\tilde{p}_2 := \varphi^{\#}(p_2)$. The kernel of $\varphi^{\#}$ contains $(A_1 \oplus A_2) \otimes B_1 \otimes (C_1 \oplus C_2)$, so $\varphi^{\#}(p_1 \oplus p_2) = \varphi^{\#}(p_2) = \tilde{p}_2$. On the other hand, because of the definition of the quotient map φ, we have:

$$\varphi^{\#}\Big(\sum_{i=1}^{s} a_i \otimes b_i \otimes c_i\Big) = \varphi^{\#}\Big(\sum_{i=\mathbf{a_1}+1}^{s} a_i \otimes b_i \otimes c_i\Big).$$

It implies $\varphi^{\#}(p_1 \oplus p_2) = \tilde{p_2} = \sum_{i=\mathbf{a_1}+1}^{s} a_i \otimes b_i \otimes c_i$. From the definition of rank: $s - \mathbf{a_1} \geq R(\tilde{p_2})$. Because of the isomorphism $\varphi^{\#}|_{A_2 \otimes B_2 \otimes C_2}$, we obtain $s - \mathbf{a_1} \geq R(p_2)$. So

$$R(p_1 \oplus p_2) \geq s - \mathbf{a_1} + \mathbf{a_1} \geq R(p_2) + \mathbf{a_1} = R(p_2) + \dim(A_1).$$

■

As a direct corollary we have:

Corollary 3.1.4. *Let $p_2 \in A_2 \otimes B_2 \otimes C_2$ and $p_1 \in A_1 \otimes B_1 \otimes C_1$ be a A_1-concise tensor, such that $R(p_1) = \dim(A_1)$. Then Strassen's additivity conjecture holds for $p_1 \oplus p_2$, i.e. $R(p_1 \oplus p_2) = R(p_1) + R(p_2)$.*

3.2. The Alexeev-Forbes-Tsimerman method for bounding tensor rank

What follows is the general technique for finding lower bounds for the tensor rank.

Proposition 3.2.1 ([LM] Proposition 2.1)**.** *Let tensor T be an element of $A \otimes B \otimes C$. There exist r rank one elements of $B \otimes C$ such that $T(A^*)$ is contained in their span if and only if $R(T) \leq r$. Similarly, $\underline{R}(T) \leq r$ if and only if there exists a curve E_t in the Grassmannian $G(r, B \otimes C)$, where for $t \neq 0$, E_t is spanned by r rank one elements and $T(A^*) \subseteq E_0$ (which is defined by the compactness of the Grassmannian).*

Proof. For the case of a tensor rank:

For the first implication, T is a tensor of rank r, $T = \sum_{i=1}^{r} a_i \otimes b_i \otimes c_i$. $T(A^*)$ is contained in an r-dimensional linear space spanned by $\langle b_i \otimes c_i \rangle$.

For the reversed implication: We know that $T(A^*)$ is contained in an r-dimensional linear space spanned by $\{M_1, ..., M_r\}$, so we know that $T(\alpha) = \sum_{i=1}^{r} \lambda_i M_i$ for every $\alpha \in A^*$. Thus, for a chosen basis of A $\{a_1, ..., a_{\mathbf{a}}\}$, we have:

$$T = \sum_{j=1}^{\mathbf{a}} a_j \otimes (\sum_{i=1}^{r} \lambda_{i,j} M_i).$$

We can reorder the sum, to write it as a sum of r simple tensors:

$$T = \sum_{i=1}^{r} (\sum_{j=1}^{\mathbf{a}} \lambda_{i,j} a_j) \otimes M_i.$$

For the case of a border tensor rank:

Let us assume that $\underline{R}(T) = r$. Thus we have a curve of $(T_t)_{t \neq 0}$ contained in $A \otimes B \otimes C$ with limit $T = T_0$, such that $\forall_{t \neq 0} R(T_t) = r$. Every single tensor T_t for $t \neq 0$ can be written as $T_t = \sum_{i=1}^{r} a_{i,t} \otimes b_{i,t} \otimes c_{i,t}$. For $t \neq 0$ we associate T_t with the linear space E_t in $B \otimes C$ spanned by $\{b_{i,t} \otimes c_{i,t}\}_{i=1,...,r}$ and denote $E_0 := \lim_{t \to 0} E_t$, which by the compactness of the Grassmannian $G(r, B \otimes C)$ has dimension r. For every $\alpha \in A^*$ $T(\alpha) = \lim_{t \to 0} T_t(\alpha)$, so $T(A^*)$ is contained in E_0.

If $\underline{R}(T) = k < r$ then by the just written method we obtain a curve $\tilde{E_t}$ in $G(k, B \otimes C)$, so we choose additional $r - k$ points $S = \{p_1, ...p_{r-k}\}$ in $B \otimes C$ such that the space spanned by points of $\tilde{E_0}$ and S has dimension r. Our new, desired curve is given by $\forall_t$ $E_t = span(\tilde{E_t}, S)$, the linear space spanned by points from $\tilde{E_t}$ and S.

We need to prove the remaining implication. We have $\forall_{\alpha\in A^*}\ T(\alpha)\in E_0=\langle M_1,\dots,M_r\rangle$, where $\{M_1,...,M_r\}$ is a basis of E_0. We know that $\langle M_1,\dots,M_r\rangle$ is a limit of $\langle \tilde{M}_{1,t},\dots,\tilde{M}_{r,t}\rangle$ when t goes to 0, where $\forall_{k,t}\ \tilde{M}_{k,t}\in B\otimes C$ is of rank 1. Thus every M_i can be approximated by $M_{i,t}=\sum_{k=1}^{r}\zeta_{i,k,t}\tilde{M}_{k,t}$, where $\forall_{i,k,t}\ \zeta_{i,k,t}\in\mathbb{C}$. We define tensors $\tilde{T}_t$ as:

$$\tilde{T}_t=\sum_{j=1}^{\mathbf{a}}a_j\otimes(\sum_{i=1}^{r}\lambda_{i,j}M_{i,t})=\sum_{j=1}^{\mathbf{a}}a_j\otimes(\sum_{i=1}^{r}\lambda_{i,j}(\sum_{k=1}^{r}\zeta_{i,k,t}\tilde{M}_{k,t}))=\sum_{k=1}^{r}(\sum_{j=1}^{\mathbf{a}}\sum_{i=1}^{r}\lambda_{i,j}\zeta_{i,k,t}a_j)\otimes\tilde{M}_{k,t}$$

We received that T can be approximated by the tensors $\tilde{T}_t$, each of rank r. ■

Corollary 3.2.2 ([LM] Corollary 2.2). *Let $T\in A\otimes B\otimes C$ be A-concise and denote $a=\dim(A)$. Then $\underline{R}(T)=a$ implies that $T(A^*)\cap Seg(\mathbb{P}B\times\mathbb{P}C)\neq\emptyset$.*

More generally:

Corollary 3.2.3 ([LM] Corollary 2.3). *Let $T\in\mathbb{C}^m\otimes\mathbb{C}^m\otimes\mathbb{C}^m=A\otimes B\otimes C$. If $\underline{R}(T)=m$, then the flag condition is fulfilled, i.e. there exists a complete flag $A_1\subseteq...\subseteq A_{m-1}\subseteq A_m=A$, with $dim A_j=j$, such that $\mathbb{P}T(A_j^*)\subseteq\sigma_j(Seg(\mathbb{P}B\times\mathbb{P}C))$.*

Proposition 3.2.4 ([LM] Proposition 3.1). *Let $T\in A\otimes B\otimes C$, $R(T)=r$, $\{c_i\}_{i=1,\dots,\mathbf{c}}$ be a fixed basis of C, $T=\sum_{i=1}^{\mathbf{c}}M_i\otimes c_i$, where $M_i\in A\otimes B$ and $M_1\neq 0$. Then:*

1. *there exist constants $\lambda_2,....,\lambda_m$, such that the tensor*

$$\tilde{T}:=\sum_{j=2}^{\mathbf{c}}(M_j-\lambda_jM_1)\otimes c_j\in A\otimes B\otimes c_1^{\perp *} \tag{3.1}$$

   has rank $R(\tilde{T})\leq r-1$.

2. *Moreover, if $R(M_1)=1$ then for any choices of λ_j we have $R(\tilde{T})\geq r-1$.*

Proof. By Proposition 3.2.1 there exist rank one tensors $X_1,...,X_r\in A\otimes B$ and scalars d_j^i such that for every j:

$$M_j=\sum_{i=1}^{r}d_j^iX_i.$$

Since $M_1\neq 0$ we may assume $d_1^1\neq 0$ and define $\lambda_j=\frac{d_j^1}{d_1^1}$. For these λ's the tensor $\tilde{T}$ (see 3.1) defines the subspace $\tilde{T}(c_1^{\perp})$ contained in a span of $X_2,...,X_r$. Proposition 3.2.1 implies $R(\tilde{T})\leq r-1$. The last assertion holds because if $R(M_1)=1$ then we may assume $X_1=M_1$ and apply Proposition 3.2.1. ■

Proposition 3.2.4 is usually implemented by consecutively applying the following steps, which we will call after Michałek and Landsberg *the AFT method*. I have slightly modified the AFT method from [LM], Section 3, adding "break" conditions, to avoid incorrect lower bounds. For examples breaking algorithm see Example 3.2.6 and Example 3.2.7.

The AFT method for bounding rank of a tensor $T\in A\otimes B\otimes C$:

1. Distinguish one of the factors, say A, take its basis $\{a_j\}$ and take bases $\{\beta_i\},\{\gamma_j\}$ of B^*,C^* and represent $T\in A\otimes B\otimes C$ as a matrix M with entries that are linear combinations of the basis vectors a_i. That means in the (i,j) place of the matrix M we have: $m_{i,j}=T(\beta_i^*\otimes\gamma_j^*)$. We will denote by $M(\beta_i)$ the i-th row of M, and by $M(\gamma_j)$ the j-th column of M.

2. Choose a subset $S_B = \{i_1, ...i_{\mathbf{b}'}\} \subset \{1, ..., \mathbf{b}\}$, such that the rows of M corresponding to the numbers belonging to S_B are pairwise different and nonzero. Similarly we choose a subset $S_C = \{j_1, ...j_{\mathbf{c}'}\} \subset \{1, ..., \mathbf{c}\}$, such that the columns of M corresponding to numbers belonging to S_C are pairwise different and nonzero. We are going to change the matrix M in the next steps, so let us denote by $\tilde{M}$ the copy of the beginning matrix M.

3. Inductively, for $1 \le l \le \mathbf{b}'$:

   (a) change the i_l-th row of matrix M to zeros, obtaining new matrix M'.

   If there exists $v \in S_C$ such that the column $M'(\gamma_v)$ is zero, then break the algorithm.

   (b) for every row $M'(\beta_r)$ of number $r \notin \{i_1, ..., i_l\} \subset S_B$ add $M(\beta_{i_l})$ times a variable λ_{r,i_l}, regarding the a_i as formal variables. In a result we obtain the new matrix M''.

   If there exists a choice of values of λ-s in $\mathbb{C}$, such that in a result $\exists_{l<k\le\mathbf{b}'}\ M''(\beta_{i_k}) = 0$ or $\exists_{j\in S_C}\ M''(\gamma_j) = 0$ then break the algorithm.

   (c) From now $M := M''$.

4. Similarly for columns. Inductively, for $1 \le l \le \mathbf{c}'$:

   (a) change the j_l-th column of matrix M to zeros, obtaining new matrix M'.

   (b) for every column $M'(\gamma_v)$ of number $v \notin \{j_1, ..., j_l\} \subset S_C$ add $M(\gamma_{j_l})$ times a variable λ'_{v,j_l}, regarding the a_i as formal variables. In a result we obtain the new matrix M''.

   If there exists a choice of values of λ-s, and λ'-s in $\mathbb{C}$, such that in a result $\exists_{l<k\le\mathbf{c}'}\ M''(\gamma_{j_k}) = 0$ then break the algorithm.

   (c) From now $M := M''$.

5. Recall that the matrix we started with is $\tilde{M}$, as in 2. For every a_j that appeared in any of the selected rows $\tilde{M}(\beta_i)$ or columns $\tilde{M}(\gamma_j)$ for $i \in S_B$ and $j \in S_C$, set a_j equals zero in all entries in whole matrix M obtaining a new matrix M'. Notice, that M' does not depend on the choice of λ-s and λ'-s.

After these 4 steps, if we do not break the algorithm, we obtain:

$$R(T) \ge \mathbf{b}' + \mathbf{c}' + R(T') \tag{3.2}$$

where the tensor T' corresponds to the matrix M', i.e. $T'(A^*) = M'$. The inequality follows by consecutively, $\mathbf{b}' + \mathbf{c}'$ times applying the Proposition 3.2.4 and then, in the 5-th step, a projection to smaller subspace. The above steps can be iterated, interchanging the roles of A, B and C.

I rewrite more precisely the Example 3.3 from [LM] with a mistake corrected.

Example 3.2.5. Let $T = a_1 \otimes (b_1 \otimes c_1 + ... + b_8 \otimes c_8) + a_2 \otimes (b_1 \otimes c_5 + b_2 \otimes c_6 + b_3 \otimes c_7 + b_4 \otimes c_8) + a_3 \otimes (b_1 \otimes c_7 + b_2 \otimes c_8) + a_4 \otimes b_1 \otimes c_8 + a_5 \otimes b_8 \otimes c_1 + a_6 \otimes b_8 \otimes c_2 + a_7 \otimes b_8 \otimes c_3 + a_8 \otimes b_8 \otimes c_4$, so

$$T(A^*) = \begin{bmatrix} a_1 & & & & & & & a_5 \\ & a_1 & & & & & & a_6 \\ & & a_1 & & & & & a_7 \\ & & & a_1 & & & & a_8 \\ a_2 & & & & a_1 & & & \\ & a_2 & & & & a_1 & & \\ a_3 & & a_2 & & & & a_1 & \\ a_4 & a_3 & & a_2 & & & & a_1 \end{bmatrix}.$$

Then $R(T) \geq 18$. Here we start by distinguishing the space B.

$$T(B^*) = \begin{bmatrix} b_1 & b_2 & b_3 & b_4 & b_5 & b_6 & b_7 & b_8 \\ & & & & b_1 & b_2 & b_3 & b_4 \\ & & & & & & b_1 & b_2 \\ & & & & & & & b_1 \\ b_8 & & & & & & & \\ & b_8 & & & & & & \\ & & b_8 & & & & & \\ & & & b_8 & & & & \end{bmatrix}$$

contracting $S_A^1 = \{5, 6, 7, 8\}$. We obtain a tensor T' represented by the matrices

$$T'(B^*) = \begin{bmatrix} b_1 & b_2 & b_3 & b_4 & b_5 & b_6 & b_7 & \\ & & & & b_1 & b_2 & b_3 & b_4 \\ & & & & & & b_1 & b_2 \\ & & & & & & & b_1 \\ & & & & & & & \\ & & & & & & & \\ & & & & & & & \\ & & & & & & & \end{bmatrix},$$

$$T'(A^*) = \begin{bmatrix} a_1 & & & & & & & \\ & a_1 & & & & & & \\ & & a_1 & & & & & \\ & & & a_1 & & & & \\ a_2 & & & & a_1 & & & \\ & a_2 & & & & a_1 & & \\ a_3 & & a_2 & & & & a_1 & \\ a_4 & a_3 & & a_2 & & & & \end{bmatrix}$$

and the inequality $R(T) \geq 4 + R(T')$. The AFT method applied 3 times more, with chosen: $S_C^2 = \{1, 2, 3, 4\}, S_B^2 = \{7, 6, 5\}$, $S_C^3 = \{5, 6\}, S_B^3 = \{4, 3\}, S_C^4 = \{7\}, S_B^4 = \{2\}$ gives $R(T') \geq (3 + 4) + (2 + 2) + (1 + 1) + 1 = 14$ and $R(T) \geq 18$.

In fact, $R(T) = 18$. It is enough to consider 17 matrices with just one nonzero entry corresponding to all nonzero entries of $T(A^*)$, apart from the top left and bottom right corners and 1 matrix with 1 at each corner and all other entries equal to 0.

This generalizes to $T \in \mathbb{C}^{2^k} \otimes \mathbb{C}^{2^k} \otimes \mathbb{C}^{2^k}$ of rank $3 * 2^k - k - 3$.

The examples in which the algorithm breaks:

Example 3.2.6. Tensor $T = a \otimes b \otimes (c_1 + c_2 + c_3) \in A \otimes B \otimes C$ is of rank 1. Trying to apply algorithm AFT for distinguished space A and $S_C = \{c_1, c_2\}$ we obtain $[a, a, a] \mapsto [0, a + \lambda_{2,1} a, a + \lambda_{3,1} a]$ and algorithm breaks because for $\lambda_{2,1} = 1$ the new matrix contains the 0 element for every column belonging to S_C. That agrees with $R(T) \ngeq R(T') + 2$, for any tensor T'.

Example 3.2.7. Tensor $T = a_1 \otimes b_1 \otimes c_1 + a_2 \otimes b_2 \otimes c_2 \in \mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2$ is concise and $R(T) = 2$. Trying to apply algorithm AFT for distinguished space A and $S_B^1 = \{2\}, S_C^1 = \{2\}$ we obtain: $M = \left[\begin{smallmatrix} a_1 & 0 \\ 0 & a_2 \end{smallmatrix}\right] \mapsto M' = \left[\begin{smallmatrix} a_1 & 0 \\ 0 & 0 \end{smallmatrix}\right]$ and algorithm breaks in step 3 of AFT - $M'(\gamma_2) = 0$. That agrees with $R(T) \ngeq R(T') + 2$ for $T' = a_1 \otimes b_1 \otimes c_1$.

Definition 3.2.8. We will say that a rank of a tensor $T \in A \otimes B \otimes C$ *can be determined by the AFT method* if there exist integer $n > 0$ and a way of consecutively applying the AFT method n times, to obtain a sequence of inequalities (see 3.2):

$$R(T) \geq \mathbf{b_1'} + \mathbf{c_1'} + R(T_1') \geq \mathbf{b_1'} + \mathbf{c_1'} + \mathbf{b_2'} + \mathbf{c_2'} + R(T_2') \geq \ldots \geq \sum_{i=1}^{n} (\mathbf{b_i'} + \mathbf{c_i'}) + R(T_n')$$

such that $R(T_n') = 0$, and $\sum_{i=1}^{n} (\mathbf{b_i'} + \mathbf{c_i'}) = R(T)$.

In many cases of low rank the AFT method provides the correct rank. The AFT helps us to prove the following theorem.

Theorem 3.2.9 ([LM] Theorem 4.1)**.** *Let $T_1 \in A_1 \otimes B_1 \otimes C_1$ and $T_2 \in A_2 \otimes B_2 \otimes C_2$ be such that $R(T_1)$ can be determined by the AFT method. Then Strassen's additivity conjecture holds for $T_1 \oplus T_2$, i.e. $R(T_1 \oplus T_2) = R(T_1) + R(T_2)$.*

Proof. In each application of the AFT method for T_1 we choose a subset of columns and subset of rows. Let us assume that in AFT we distinguished A_1. Because of injection $f_{A_1} : A_1 \to A_1 \oplus A_2$ we can interpret S_{B_1} as a rows and S_{C_1} as a columns of $(T_1 \oplus T_2)((A_1 \oplus A_2)^*)$ letting AFT work for $T_1 \oplus T_2$ in a bigger space. We do the same in case of B_1 and C_1. $R(T_1)$ can be determined by the AFT method. Hence we are repeating the AFT method, which give us the proper bound for T_1, in a bigger space, changing A_1 to $A_1 \oplus A_2$ and similarly for B_1 and C_1, and translating the numbers of columns and rows by $f_{A_1}, f_{B_1}, f_{C_1}$.

With each application of the AFT method, T_1 is modified to a tensor of lower rank living in a smaller space and T_2 is unchanged. It is still a direct sum after each application of AFT, because in 5-th step of AFT method we cancel every possible coefficient which could appear in steps 3 and 4 denying $\widetilde{T_1 \oplus T_2}$ is a direct sum.

After all applications, T_1 has been modified to zero and T_2 is still unchanged. Hence $R(T_1 \oplus T_2) = R(T_1) + R(T_2)$ follows from inequality 3.2. ■

To prove Proposition 3.2.11, which we will need in the proof of Proposition 3.2.12, we need the following lemma:

Lemma 3.2.10. *Let $B = \mathbb{C}^{\mathbf{b}}$, $C = \mathbb{C}^{\mathbf{c}}$ and $\mathbf{b} \leq \mathbf{c}$. For a concise tensor $T \in \mathbb{C}^2 \otimes B \otimes C$ there exist $\gamma \in C^*$ such that $T(\gamma)$ is a rank one matrix.*

Proof. Because of conciseness, for $T(C^*) \subseteq \mathbb{C}^2 \otimes B$ $\dim T(C) = \mathbf{c}$. We have $\dim(Seg(\mathbb{P}(\mathbb{C}^2) \times \mathbb{P}(B))) = 1 + \mathbf{b} - 1 + 1 = 1 + \mathbf{b}$, $\dim(\mathbb{C}^2 \otimes B) = 2\mathbf{b}$. Thus $\dim(C) + \dim(Seg(\mathbb{P}(\mathbb{C}^2) \times \mathbb{P}(B))) = c + b + 1 \geq 2b + 1 > 2b$. Hence there is a non-empty intersection $T(C^*) \cap Seg(\mathbb{P}(\mathbb{C}^2) \times \mathbb{P}(B))$. ■

Proposition 3.2.11. *Let $B = \mathbb{C}^{\mathbf{b}}$, $C = \mathbb{C}^{\mathbf{c}}$. For every non-zero tensor $T \in \mathbb{C}^2 \otimes B \otimes C$ there exists γ, an element of B^* or C^* such that $T(\gamma)$ is a rank one matrix.*

Proof. If T is concise then we apply Lemma 3.2.10. In other case we can find $A' = \mathbb{C}^{\mathbf{a}'}, B' = \mathbb{C}^{\mathbf{b}'}, C' = \mathbb{C}^{\mathbf{c}'}$ such that $T \in A' \otimes B' \otimes C' \subseteq \mathbb{C}^2 \otimes B \otimes C$ is concise. If the $\mathbf{a}' = 2$ we apply the lemma to A', B', C' or to A', C', B' if dimension of B' is greater than dimension of C'^*. If not, $\mathbf{a}' = 1$, then for every $\gamma \in C'^*$, $T(\gamma)$ is a rank 1 tensor. ■

In [LM] Section 4, there is noted, with a short proof, that Strassen's additivity conjecture holds if the space A_1 is of dimension 2 and tensor $T_1 \in A_1 \otimes B_1 \otimes C_1$. I give a detailed proof which follows the mentioned one.

Proposition 3.2.12. *Let $T_1 \in A_1 \otimes B_1 \otimes C_1$ and $T_2 \in A_2 \otimes B_2 \otimes C_2$ be such that* $\dim(A_1) = 2$. *Then Strassen's additivity conjecture holds for $T_1 \oplus T_2$, i.e. $R(T_1 \oplus T_2) = R(T_1) + R(T_2)$.*

Proof. We define $T := T_1 \oplus T_2$, $A := A_1 \oplus A_2$ and similarly for B and C. Let us look on $T(A^*) = (T_1 \oplus T_2)(A_1^* \oplus A_2^*)$. We slightly modify the AFT method. We take a basis $\{a_1, ..., a_{\mathbf{a_1}}\}$ of A_1, $\{a'_1, ..., a'_{\mathbf{a_2}}\}$ of A_2, $\{\beta_1, ..., \beta_{\mathbf{b_1}}\}$ of B_1^*, $\{\beta'_1, ..., \beta'_{\mathbf{b_2}}\}$ of B_2^*, $\{\gamma_1, ..., \gamma_{\mathbf{c_1}}\}$ of C_1^*, $\{\gamma'_1, ..., \gamma'_{\mathbf{c_2}}\}$ of C_2^*, and represent $T(A^*)$ as a block matrix M with entries that are linear combinations of the basis vectors of A_1 and A_2.

By Proposition 3.2.11 we can always find either $\beta \in B_1^*$ or $\gamma \in C_1^*$, such that $T_1(\beta)$ or $T_1(\gamma)$ is a rank one matrix. Possibly after changing bases of B_1 and C_1, without loosing the generality, we find β_1. As in point 3a of the AFT algorithm, we change the row corresponding to β_1 to zero, obtaining a new matrix M'. Now as in 3b, for every row v of M', other than $M'(\beta_1)$, we add $M(\beta_1)$ times a coefficients $\lambda_{v,1}$ coming from Proposition 3.2.4 point 1 applied to $T = T_1 \oplus T_2$, regarding the $\{a_1, ..., a_{\mathbf{a_1}}, a'_1, ..., a_{\mathbf{a_2}'}\}$ as formal variables. We obtained new matrix M'' and new tensor $T'_1 \in A_1 \otimes B_1 \otimes C_1$ corresponding to $M''|_{B_1 \otimes C_1}$.

Constantly applying the previous paragraph (each application we start with the tensors $T'_1 \in A_1 \otimes B_1 \otimes C_1$ and $\tilde{T} \in A \otimes B \otimes C$, such that $\tilde{T}(A) = M''$ from the previous application in place of respectively T_1, T, M) we finally obtain the matrix $\tilde{M}$ with submatrix $\tilde{M}|_{B_1 \otimes C_1}$ equal zero. We reach the zero submatrix in not less than $R(T_1)$ steps because of Proposition 3.2.4 point 2. By Proposition 3.2.4 point 1 after each application, we reduce the rank of a tensor by at least 1. We have a lower bound:

$$R(\tilde{T}) \leq R(T) - R(T_1)$$

We change all instances of vectors from A_1 in M to zero, obtaining tensor $\tilde{T}' \in A \otimes B \otimes C$ equivalent to tensor T_2. We have:

$$R(T_2) = R(\tilde{T}') \leq R(\tilde{T}) \leq R(T) - R(T_1),$$

thus:

$$R(T_2) + R(T_1) \leq R(T).$$

The inverse inequality always holds, see property 5, the end of §2.1. ■

3.3. The third secant variety of $A \otimes B \otimes C$ and the main theorem

In this section I prove that the Strassen's additivity conjecture holds if one of the tensor is of border rank 3 which gives a family of tensors of rank 5 for which SAC holds. The crucial facts are the following theorem and Theorem 3.2.9.

Theorem 3.3.1. *([BL] Theorem 1.2.) Let $X := Seg(\mathbb{P}A \times \mathbb{P}B \times \mathbb{P}C)$. Let $p = [v] \in \sigma_3(X) \setminus \sigma_2(X)$. Then v has one of the following normal forms:*

(i) $a_1 \otimes b_1 \otimes c_1 + a_2 \otimes b_2 \otimes c_2 + a_3 \otimes b_3 \otimes c_3$,

(ii) $a_1 \otimes b_1 \otimes c_2 + a_1 \otimes b_2 \otimes c_1 + a_2 \otimes b_1 \otimes c_1 + a_3 \otimes b_3 \otimes c_3$,

(iii) $a_1 \otimes b_2 \otimes c_2 + a_2 \otimes b_1 \otimes c_2 + a_2 \otimes b_2 \otimes c_1 + a_1 \otimes b_1 \otimes c_3 + a_1 \otimes b_3 \otimes c_1 + a_3 \otimes b_1 \otimes c_1$,

(iv) $a_2 \otimes b_1 \otimes c_2 + a_2 \otimes b_2 \otimes c_1 + a_1 \otimes b_1 \otimes c_3 + a_1 \otimes b_3 \otimes c_1 + a_3 \otimes b_1 \otimes c_1$.

For type (iv) there are two other normal forms, where the role of a is switched with that of b and c. We write these types in terms of slices (see Definition 2.2.3) in Table 3.1, if v is not contained in any of $\mathbb{C}^2 \otimes \mathbb{C}^3 \otimes \mathbb{C}^3$, $\mathbb{C}^3 \otimes \mathbb{C}^2 \otimes \mathbb{C}^3$, $\mathbb{C}^3 \otimes \mathbb{C}^3 \otimes \mathbb{C}^2$.

	normal form	slice	$\underline{R}$	R
(i)	$a_1 \otimes b_1 \otimes c_1 + a_2 \otimes b_2 \otimes c_2 + a_3 \otimes b_3 \otimes c_3$	$\begin{bmatrix} t & & \\ & s & \\ & & u \end{bmatrix}$	3	3
(ii)	$a_1 \otimes (b_1 \otimes c_2 + b_2 \otimes c_1) + a_2 \otimes b_1 \otimes c_1 + a_3 \otimes b_3 \otimes c_3$	$\begin{bmatrix} t & s & \\ s & & \\ & & u \end{bmatrix}$	3	4
(iii)	$a_1 \otimes (b_1 \otimes c_3 + b_2 \otimes c_2 + b_3 \otimes c_1) + a_2 \otimes (b_1 \otimes c_2 + b_2 \otimes c_1) + a_3 \otimes b_1 \otimes c_1$	$\begin{bmatrix} t & s & u \\ s & u & \\ u & & \end{bmatrix}$	3	5
(iv)	$a_1 \otimes (b_1 \otimes c_2 + b_2 \otimes c_1) + a_2 \otimes b_1 \otimes c_1 + a_3 \otimes (b_3 \otimes c_1 + b_1 \otimes c_3)$	$\begin{bmatrix} t & s & u \\ s & & \\ u & & \end{bmatrix}$	3	5
(iv)	$a_1 \otimes (b_1 \otimes c_2 + b_2 \otimes c_1 + b_3 \otimes c_3) + a_2 \otimes b_1 \otimes c_1 + a_3 \otimes b_3 \otimes c_1$	$\begin{bmatrix} t & s & \\ s & & \\ u & & s \end{bmatrix}$	3	5
(iv)	$a_1 \otimes (b_1 \otimes c_2 + b_2 \otimes c_1 + b_2 \otimes c_3) + a_2 \otimes b_1 \otimes c_1 + a_3 \otimes b_1 \otimes c_3$	$\begin{bmatrix} t & s & u \\ s & & \\ & & s \end{bmatrix}$	3	5

Table 3.1: Table of the normal forms of tensors in $A \otimes B \otimes C$ of border rank 3. R denotes *rank* and $\underline{R}$ denotes *border rank* of a tensor. [BL]

We can recover the rank of every tensor from the table, equivalently every tensor of border rank 3, using the Alexeev-Forbes-Tsimerman method for bounding tensor rank.

Proposition 3.3.2. *Rank of every tensor in $A \otimes B \otimes C$ of border rank 3, which is not contained in any of $\mathbb{C}^2 \otimes \mathbb{C}^3 \otimes \mathbb{C}^3$, $\mathbb{C}^3 \otimes \mathbb{C}^2 \otimes \mathbb{C}^3$, $\mathbb{C}^3 \otimes \mathbb{C}^3 \otimes \mathbb{C}^2$, can be computed using the Alexeev-Forbes-Tsimerman method for bounding tensor rank.*

Proof. Every tensor in $A \otimes B \otimes C$ of border rank 3, which is not contained in any of $\mathbb{C}^2 \otimes \mathbb{C}^3 \otimes \mathbb{C}^3$, $\mathbb{C}^3 \otimes \mathbb{C}^2 \otimes \mathbb{C}^3$, $\mathbb{C}^3 \otimes \mathbb{C}^3 \otimes \mathbb{C}^2$ can be presented in one of the normal forms from Table 3.1, where $a_1, a_2, a_3, ...$ are linearly independent. Let us show that for every isomorphism class of such a slices the AFT method gives us a proper bound, i.e. is equal to the rank of a corresponding tensor.

(i) For the slice $\begin{bmatrix} t & & \\ & s & \\ & & u \end{bmatrix}$ applying AFT method we obtain: $\begin{bmatrix} t & & \\ & s & \\ & & u \end{bmatrix} \mapsto \begin{bmatrix} t & \\ & s \end{bmatrix} \mapsto \begin{bmatrix} t \end{bmatrix}$. Thus the rank is greater or equal $1 + 1 + 1 = 3$.

(ii) For the slice $\begin{bmatrix} t & s & \\ s & & \\ & & u \end{bmatrix}$ applying AFT method we obtain: $\begin{bmatrix} t & s & \\ s & & \\ & & u \end{bmatrix} \mapsto \begin{bmatrix} t & & \\ & & \\ & & u \end{bmatrix} \mapsto \begin{bmatrix} t & & \\ & & \\ & & \end{bmatrix}$. Thus the rank is greater or equal $2 + 1 + 1 = 4$.

(iii) For a slice $\begin{bmatrix} t & s & u \\ s & u & \\ u & & \end{bmatrix}$ we obtain $\begin{bmatrix} t & s & u \\ s & u & \\ u & & \end{bmatrix} \mapsto \begin{bmatrix} t & s & \\ s & & \\ & & \end{bmatrix} \mapsto \begin{bmatrix} t & & \\ & & \\ & & \end{bmatrix}$. Thus the rank is greater or equal $2 + 2 + 1 = 5$.

(iv) For a tensor of type (iv) normal forms are isomorphic, so we can choose slice $\begin{bmatrix} t & s & u \\ s & & \\ u & & \end{bmatrix}$ which give us the following intermediate matrices: $\begin{bmatrix} t & s & u \\ s & & \\ u & & \end{bmatrix} \mapsto \begin{bmatrix} t & s & \\ s & & \\ & & \end{bmatrix} \mapsto \begin{bmatrix} t & & \\ & & \\ & & \end{bmatrix}$ and rank greater or equal 2+2+1=5.

■

Now as a corollary we obtain:

Theorem 3.3.3. *Let $T_1 \in A_1 \otimes B_1 \otimes C_1$ be a tensor of border rank 3 and $T_2 \in A_2 \otimes B_2 \otimes C_2$ be arbitrary tensor. Then Strassen's additivity conjecture holds for $T_1 \oplus T_2$, i.e. $R(T_1 \oplus T_2) = R(T_1) + R(T_2)$.*

Proof. If one of A_1, B_1, C_1 is of dimension less or equal 2 we have Proposition 3.2.12. If it's not the case, the theorem follows from Theorem 3.2.9 and the Proposition 3.3.2. ■

Corollary 3.3.4. *Then Strassen's additivity conjecture holds for $T_1 \in A_1 \otimes B_1 \otimes C_1$ and $T_2 \in A_2 \otimes B_2 \otimes C_2$, i.e. $R(T_1 \oplus T_2) = R(T_1) + R(T_2)$ if the tensor $T_1 \in A_1 \otimes B_1 \otimes C_1$ fulfills one of the conditions:*

- *T_1 is a tensor of rank 5 and type (iii) or (iv),*
- *T_1 is a tensor of rank 4 and type (ii).*

Bibliography

[BGL] J. Buczyński, A. Ginensky, and J.M. Landsberg, Determinental equations for secant varieties and the Eisenbud-Koh-Stillman conjecture. *Journal of London Mathematical Society* (2013) 88(1), 1-24, doi: 10.1112/jlms/jds073, arXiv:1007.0192

[BL] J. Buczyński and J.M. Landsberg, On the third secant variety, *Journal of Algebraic Combinatorics*, 2014, 40(2), 475-502, DOI: 10.1007/s10801-013-0495-0 (open access)

[CCG] E. Carlini, M.V. Catalisano, and A. V. Geramita. The solution to the Waring problem for monomials and the sum of coprime monomials. *J. Algebra*, 370:5–14, 2012.

[JT] J. Ja'Ja' and J. Takche, On the validity of the direct sum conjecture, *SIAM J. Comput.* 15 (1986), no. 4, 1004–1020. MR MR861366 (88b:68084)

[L] J. M. Landsberg, Tensors: geometry and applications, Graduate Studies in Mathematics, vol. 128

[LM] J.M. Landsberg and Mateusz Michałek, Abelian Tensors, https://arxiv.org/abs/1504.03732, 2016

[S69] V. Strassen. Gaussian elimination is not optimal. *Numer. Math.* , 13:354-356, 1969.

[S73] V. Strassen. Vermeidung von Divisionen. *Journal für die reine und angewandte Mathematik*, 264:184–202, 1973.

[Sch] A. Schönhage, Partial and total matrix multiplication, *SIAM J. Comput.* 10 (1981), no. 3, 434–455. MR MR623057 (82h:68070)

[W] S. Winograd. On multiplication of 2 x 2 matrices. *Linear Algebra and Appl.*, 4:381–388, 1971.