A NOTE ON THE DEGENERATION ORDER OF BILINEAR MAPS

PETER BÜRGISSER

INSTITUT FÜR MATHEMATIK, UNIVERSITÄT ZÜRICH, WINTERTHURERSTR. 190,
CH-8057 ZÜRICH, SWITZERLAND

JANUARY 1997

Abstract. We investigate whether Strassen's theory of asymptotic spectra [19, 20] could be simplified by extending the degeneration order into a ring. It turns out that this must necessarily fail. This negative result shows once more the intricacy of bilinear complexity. Along the way, we develop a necessary condition for degenerations in terms of dimensions of isotropy groups of tensors.

Key words. Matrix multiplication, bilinear complexity, border rank, degeneration order, asymptotic spectrum.

1. Introduction

We begin by recalling some concepts from the theory of bilinear complexity. For details, the reader may consult the textbooks by Pan [16], de Groote [12], Bini and Pan [3], or Bürgisser et al. [5].

A (coordinate) tensor t of format (m, n, p) is a 3-dimensional array $[t_{ij\ell}]_{i,j,\ell} \in \mathbb{C}^{m\times n\times p}$ of complex numbers. A triple of matrices $\alpha \in \mathbb{C}^{m'\times m}$, $\beta \in \mathbb{C}^{n'\times n}$, $\gamma \in \mathbb{C}^{p'\times p}$ transforms t into the tensor $s = (\alpha \otimes \beta \otimes \gamma)t \in \mathbb{C}^{m'\times n'\times p'}$, which is defined by

$$s_{ij\ell} := \sum_{i_1,j_1,\ell_1} \alpha_{ii_1}\beta_{jj_1}\gamma_{\ell\ell_1}t_{i_1j_1\ell_1}.$$

We abbreviate this reduction by $s \le t$ and call s a restriction of t. If α, β, γ can be chosen as invertible matrices, then s is said to be isomorphic to t: $s \simeq t$. (In this case, s and t have the same format.)

For $r \in \mathbb{N}$ let the tensor $\langle r\rangle \in \mathbb{C}^{r\times r\times r}$ be defined by $\langle r\rangle_{ij\ell} := 1$ if $i = j = \ell$ and $\langle r\rangle_{ij\ell} := 0$ otherwise. The *rank* $R(s)$ of a tensor s is defined as the smallest $r \in \mathbb{N}$ such that $s \le \langle r\rangle$. This is an important quantity, as $R(s)$ equals (up to a factor of two) the minimum number of multiplications or divisions sufficient to compute the bilinear forms

$$\sum_{i,j} s_{ij\ell}X_iY_j, \quad \ell = 1, \dots, p'$$

from the indeterminates X_i, Y_j. (See [5, Sect. 14.1].)

An approximate reduction $\trianglelefteq$ can be defined as follows. Let $s' \in \mathbb{C}^{m'\times n'\times p'}$, $t \in \mathbb{C}^{m\times n\times p}$, and ϵ be an indeterminate over $\mathbb{C}$. We say that s is a *degeneration* of t, $s \trianglelefteq t$, iff there exist matrices α, β, γ over $\mathbb{C}(\epsilon)$ such that

$$(\alpha \otimes \beta \otimes \gamma)t = s + \epsilon s'$$

holds for some tensor $s' \in \mathbb{C}[\epsilon]^{m' \times n' \times p'}$. This reduction is intimately connected with the notion of the *border rank* $\underline{R}(s)$ of s, which is defined as the smallest $r \in \mathbb{N}$ such that s $\trianglelefteq \langle r \rangle$. The border rank was introduced by Bini et al. [2] and Bini [1] as a powerful tool for designing fast matrix multiplication algorithms. (Cf. [5, Chap. 15].)

To show that a specific tensor s is not a degeneration of another tensor t is in general a difficult task; in the special case $t = \langle r \rangle$ this amounts to showing the lower bound $\underline{R}(s) > r$ on the border rank of s. In the next section, we will develop the following necessary condition for a degeneration relation: $s \trianglelefteq t \Rightarrow \gamma(s) \geq \gamma(t)$, where $\gamma(t)$ denotes the dimension of the isotropy group of t. This condition will be applied to prove our main result, and might also be useful in future investigations.

In 1987 Strassen developed a theory of relative bilinear complexity and asymptotic spectra of tensors [19, 20, 21]. His new ideas, extended by Coppersmith and Winograd [6], led to the currently best estimate $\omega < 2.38$ of the exponent ω of matrix multiplication. (See also [5, Chap. 15].) In the sequel, we will discuss an approach to simplify Strassen's theory. For this, let us first recall some definitions and facts from [19, 20].

A tensor $[t_{ij\ell}] \in \mathbb{C}^{m \times n \times p}$ is called *concise* iff the vectors $(t_{1j\ell}, \ldots, t_{mj\ell})$, $1 \leq j \leq n$, $1 \leq \ell \leq p$, generate $\mathbb{C}^m$, and analogous properties hold if the second and third coordinates are distinguished. The direct sum $t^1 \oplus t^2$ and the tensor product $t^1 \otimes t^2$ of two tensors $t^\mu \in \mathbb{C}^{m_\mu \times n_\mu \times p_\mu}$, $\mu = 1, 2$, are defined by

$$t^1 \oplus t^2 \in \mathbb{C}^{(m_1+m_2)\times(n_1+n_2)\times(p_1+p_2)}, \; t^1 \otimes t^2 \in \mathbb{C}^{(m_1 \times m_2)\times(n_1 \times n_2)\times(p_1 \times p_2)},$$

$$\begin{aligned}
(t^1 \oplus t^2)_{ij\ell} &:= \begin{cases} t^1_{ij\ell} & \text{if } i \leq m_1, j \leq n_1, \ell \leq p_1, \\ t^2_{i-m_1, j-n_1, \ell-p_1} & \text{if } i > m_1, j > n_1, \ell > p_1, \\ 0 & \text{otherwise,} \end{cases} \\
(t^1 \otimes t^2)_{(i,i_1),(j,j_1),(\ell,\ell_1)} &:= t^1_{ij\ell} \cdot t^2_{i_1 j_1 \ell_1}.
\end{aligned}$$

(For $\otimes$ we use double indices.) It is straightforward to check that direct sum and tensor product of concise tensors are again concise. (Cf. [5, Lemma 14.40].) Now let $\mathcal{B}^+$ denote the set of all isomorphism classes of concise tensors. Direct sum and tensor product induce operations on $\mathcal{B}^+$, which provide it with the structure of a commutative semiring. Moreover, the degeneration induces a partial order on $\mathcal{B}^+$ which is compatible with these operations. One can prove that additive cancellation holds, which implies that the semiring $\mathcal{B}^+$ can be embedded in an essentially unique commutative ring $\mathcal{B}$ satisfying $\mathcal{B} = \mathcal{B}^+ - \mathcal{B}^+$. (Cf. [19, Prop. 2.1]. This is analogous to the embeddding of the semiring of natural numbers in the ring of integers.) The elements of $\mathcal{B}$ are called *generalized tensor classes.*

Strassen defined an asymptotic version of the degeneration order as follows: $a \in \mathcal{B}^+$ is called an *asymptotic degeneration* of b, $a \lesssim b$, iff there is some sequence $\epsilon_N = o(N)$ of natural numbers such that $a^N \trianglelefteq b^N \cdot 2^{\epsilon_N}$ holds in the ring $\mathcal{B}$ for all N. Note that $a \trianglelefteq b$ implies $a \lesssim b$. It turns out that $\lesssim$ is a preorder on $\mathcal{B}^+$ compatible with addition and multiplication. The surprising and crucial fact is now that $\lesssim$ *allows an extension to the ring* $\mathcal{B}$, *compatible with the ring operations.* In this way, $\mathcal{B}$ becomes a preordered ring, which allows to bring in the structure theory of Stone, Kadison, and Dubois [18, 13, 7, 8]. Using this, Strassen [20] proved the following central result: for any subset $X \subseteq \mathcal{B}^+$ there is a compact space Δ and a ring morphism $\varphi: \mathbb{Z}[X] \to C(\Delta)$ such that $\varphi(X)$ separates the points of Δ, and such that for all $a, b \in \mathbb{Z}[X]$

$$a \lesssim b \Longleftrightarrow (\varphi(a) \le \varphi(b) \text{ pointwise on } \Delta).$$

($\mathbb{Z}[X]$ denotes the subring of $\mathcal{B}$ generated by X, and $C(\Delta)$ stands for the ring of continuous real functions on Δ.) The pair (Δ, φ) is essentially unique and is called the *asymptotic spectrum* of X. For example, let $\langle h, h, h\rangle$ be the structural tensor of the square matrix multiplication $\mathbb{C}^{h\times h} \times \mathbb{C}^{h\times h} \to \mathbb{C}^{h\times h}$ and denote the isomorphism class of $\langle h, h, h\rangle$ by the same symbol. It is known that the asymptotic spectrum of the set $X = \{\langle h, h, h\rangle \mid h \ge 1\} \subseteq \mathcal{B}^+$ is the interval $\Delta_c := [4, 2^\omega]$ together with the map φ that sends $\langle h, h, h\rangle$ to the continuous function $\Delta_c \to \mathbb{R}$, $x \mapsto x^{\log_2 h}$. In particular, $\omega > 2$ iff Δ_c contains more than one point. If we succeed in computing the asymptotic spectrum of a set of tensor classes X, then the question of whether $a \lesssim b$ for $a, b \in \mathbb{Z}[X]$ is reduced to the purely analytical question of deciding whether $\varphi(a) \le \varphi(b)$ holds pointwise on Δ. In particular, this yields also a necessary condition for a degeneration relation $a \trianglelefteq b$.

Now we come to the question treated in this paper. Strassen asked whether it is really necessary to pass from $\trianglelefteq$ to the asymptotic degeneration $\lesssim$ in the above described construction in order to make things work. In fact, if the degeneration order, which is defined on $\mathcal{B}^+$, would allow an extension to the ring $\mathcal{B}$ in a way compatible with the ring operations, then the above theory could be considerably simplified! One can prove that in this case, the asymptotic degeneration would not differ too much from the degeneration: we had $a \lesssim b$ iff $Na \trianglelefteq Nb + \epsilon_N$ for some sequence $\epsilon_N = o(N)$ of natural numbers. The degeneration itself could therefore be rather accurately described in terms of asymptotic spectra, which would represent a major step towards a deeper understanding of the degeneration relation, as well as of border rank.

Unfortunately, this is too nice to be true.

THEOREM 1.1. *The degeneration order cannot be extended to the ring $\mathcal{B}$ in a way compatible with the ring operations.*

The proof relies on constructions due to Pan [15] and Schönhage [17] which show that the border rank is not additive. The second ingredient is the necessary condition for degenerations in our Thm. 2.2.

We remark that the extendability of $\trianglelefteq$ to the ring $\mathcal{B}$ would also imply that a cancellation law holds: $a+c \trianglelefteq b+c$ would imply $a \trianglelefteq b$. In particular, we would have $\underline{R}(a+\langle r\rangle) = \underline{R}(a)+r$. Both of these properties are probably wrong, but we have been unable to prove this.

2. A necessary condition for degeneration

With a tensor $t \in \mathbb{C}^{m\times n\times p}$ we associate the polynomial map

$$\varphi_t: \mathbb{C}^{m\times m} \times \mathbb{C}^{n\times n} \times \mathbb{C}^{p\times p} \to \mathbb{C}^{m\times n\times p},\ (\alpha, \beta, \gamma) \mapsto (\alpha \otimes \beta \otimes \gamma)t.$$

Let $s \in \mathbb{C}^{m\times n\times p}$ be a degeneration of t, say $(\tilde\alpha \otimes \tilde\beta \otimes \tilde\gamma)t = s + \epsilon s'$ for matrices $\tilde\alpha, \tilde\beta, \tilde\gamma$ over $\mathbb{C}(\epsilon)$ and some tensor $s' \in \mathbb{C}[\epsilon]^{m\times n\times p}$. For any complex square matrices α, β, γ of size m, n, p, respectively, we obtain

$$(\alpha\tilde\alpha \otimes \beta\tilde\beta \otimes \gamma\tilde\gamma)t = (\alpha \otimes \beta \otimes \gamma)(\tilde\alpha \otimes \tilde\beta \otimes \tilde\gamma)t = (\alpha \otimes \beta \otimes \gamma)s + \epsilon(\alpha \otimes \beta \otimes \gamma)s'.$$

Now let (θ_n) be a sequence of complex numbers converging to zero and such that $\tilde{\alpha}, \tilde{\beta}, \tilde{\gamma}$ are defined at θ_n for all n. Then $(\alpha\tilde{\alpha}(\theta_n) \otimes \beta\tilde{\beta}(\theta_n) \otimes \gamma\tilde{\gamma}(\theta_n))t$ converges to $(\alpha \otimes \beta \otimes \gamma)s$ as $n \to \infty$, hence $(\alpha \otimes \beta \otimes \gamma)s$ is contained in the closure of the image $\operatorname{im}\varphi_t$ of φ_t. In particular, it is contained in the Zariski closure $\overline{\operatorname{im}\varphi_t}$ of $\operatorname{im}\varphi_t$, which is defined as the smallest zero set of polynomials, which contains $\operatorname{im}\varphi_t$. As α, β, γ were arbitrary, we have thus shown the implication

$$s \trianglelefteq t \Longrightarrow \overline{\operatorname{im}\varphi_s} \subseteq \overline{\operatorname{im}\varphi_t}.$$

In algebraic geometry, one shows that Zariski closed sets have a well defined dimension, which is monotone with respect to inclusion (cf. [14, §1A]). Hence we have $\dim \overline{\operatorname{im}\varphi_s} \leq \dim \overline{\operatorname{im}\varphi_t}$ if $s \trianglelefteq t$.

We define now the following invariant.

DEFINITION 2.1. *For a tensor $t \in \mathbb{C}^{m\times n\times p}$ we put $\gamma(t) := m^2 + n^2 + p^2 - \dim \overline{\operatorname{im}\varphi_t}$.*

We have proved the following necessary condition for degenerations.

THEOREM 2.2. *If $s, t \in \mathbb{C}^{m\times n\times p}$ are such that $s \trianglelefteq t$, then $\gamma(s) \geq \gamma(t)$.*

In order to apply this theorem, we need a method to compute dimensions. For this, the following result from algebraic geometry is useful (cf. [14, §3A, Prop. (3.6)]).

LEMMA 2.3. *Let $\psi = (\psi_1, \ldots, \psi_n)\colon \mathbb{C}^m \to \mathbb{C}^n$ be a polynomial map. Then there is some nonzero polynomial $h \in \mathbb{C}[X_1, \ldots, X_m]$ such that*

$$\dim \overline{\operatorname{im}\varphi_t} = \operatorname{rank}\big[\frac{\partial \psi_i}{\partial X_j}(p)\big]_{i,j}$$

for all $p \in \mathbb{C}^m$ satisfying $h(p) \neq 0$.

Applying this lemma to the map φ_t, we conclude that there exist invertible matrices $\alpha_0, \beta_0, \gamma_0$ such that the dimension of $\overline{\operatorname{im}\varphi_t}$ equals the rank of the Jacobian matrix of φ_t in $p = (\alpha_0, \beta_0, \gamma_0)$. On the other hand, the relation

$$\varphi_t(\alpha, \beta, \gamma) = (\alpha_0^{-1} \otimes \beta_0^{-1} \otimes \gamma_0^{-1})\varphi_t(\alpha_0\alpha, \beta_0\beta, \gamma_0\gamma)$$

shows that the rank of the Jacobian matrix of φ_t in p equals those in $I := (E_m, E_n, E_p)$.

Let ϵ be an indeterminate and X, Y, Z be square matrices of size m, n, p, respectively. A calculation shows that

$$\begin{aligned}
&\varphi_t(E_m + \epsilon X, E_n + \epsilon Y, E_p + \epsilon Z) = \big((E_m + \epsilon X) \otimes (E_n + \epsilon Y) \otimes (E_p + \epsilon Z)\big)t \\
&= \ (E_m \otimes E_n \otimes E_p)t \\
&\quad + \epsilon\big[(X \otimes E_n \otimes E_p)t + (E_m \otimes Y \otimes E_p)t + (E_m \otimes E_n \otimes Z)t\big] + O(\epsilon^2).
\end{aligned}$$

The linear map corresponding to the Jacobian matrix of φ_t in I is called the differential of φ_t in I. It maps the tripel (X, Y, Z) to the coefficient of ϵ (in square brackets). We conclude that the rank of this differential equals $\dim \overline{\operatorname{im}\varphi_t}$, hence $\gamma(t)$ equals the dimension of the kernel of this differential.

Summarizing, we have proved the following.

LEMMA 2.4. *For a tensor $t = [t_{ij\ell}] \in \mathbb{C}^{m\times n\times p}$ the invariant $\gamma(t)$ equals the dimension of the set of solutions of the homogeneous system of linear equations*

$$\forall i \le m, j \le n, \ell \le p : \sum_{\rho=1}^{m} X_{i\rho} t_{\rho j\ell} + \sum_{\rho=1}^{n} Y_{j\rho} t_{i\rho\ell} + \sum_{\rho=1}^{p} Z_{\ell\rho} t_{ij\rho} = 0$$

in the $m^2+n^2+p^2$ variables $X_{ii_1}, Y_{jj_1}, Z_{\ell\ell_1}$.

For instance, $\gamma(\langle 1\rangle)$ equals the dimension of the solution space of $X+Y+Z=0$, hence $\gamma(\langle 1\rangle) = 2$.

REMARK 2.5. One can show that $\gamma(t)$ equals the dimension of the isotropy group $\Gamma(t)$ of t, which is defined as

$$\Gamma(t) := \{(\alpha,\beta,\gamma) \in \mathrm{GL}(m)\times\mathrm{GL}(n)\times\mathrm{GL}(p) \mid (\alpha\otimes\beta\otimes\gamma)t = t\}.$$

The isotropy groups of tensors have been thoroughly investigated by de Groote [9, 10, 11].

As a first application of Lemma 2.4 we show that the invariant γ is additive with respect to direct sums of concise tensors.

PROPOSITION 2.6. *If t^1, t^2 are concise tensors, then $\gamma(t^1\oplus t^2) = \gamma(t^1)+\gamma(t^2)$.*

PROOF. We denote the format of t^μ by (m_μ, n_μ, p_μ). Let X, Y, Z be square matrices of size m_1+m_2, n_1+n_2, p_1+p_2, respectively. We subdivide X into four blocks $X^{\mu\nu} \in k^{m_\mu\times n_\nu}$ $(1 \le \mu,\nu \le 2)$ and proceed similarly for Y and Z. The linear system of equations of Lemma 2.4 for the tensor $t^1\oplus t^2$ then becomes

(a) $\sum_\rho X^{11}_{i\rho} t^1_{\rho j\ell} + \sum_\rho Y^{11}_{j\rho} t^1_{i\rho\ell} + \sum_\rho Z^{11}_{\ell\rho} t^1_{ij\rho} = 0,$

(b) $\sum_\rho X^{21}_{i\rho} t^1_{\rho j\ell} = 0,\ \sum_\rho Y^{21}_{j\rho} t^1_{i\rho\ell} = 0,\ \sum_\rho Z^{21}_{\ell\rho} t^1_{ij\rho} = 0$,

(c) $\sum_\rho X^{12}_{i\rho} t^2_{\rho j\ell} = 0,\ \sum_\rho Y^{12}_{j\rho} t^2_{i\rho\ell} = 0,\ \sum_\rho Z^{12}_{\ell\rho} t^2_{ij\rho} = 0$,

(d) $\sum_\rho X^{22}_{i\rho} t^2_{\rho j\ell} + \sum_\rho Y^{22}_{j\rho} t^2_{i\rho\ell} + \sum_\rho Z^{22}_{\ell\rho} t^2_{ij\rho} = 0,$

where i, j, ℓ vary in the corresponding ranges. The conditions (a) express that (X^{11}, Y^{11}, Z^{11}) is a solution of the system of equations corresponding to t^1 according to Lemma 2.4. Similarly for (d). If t^1 is concise, then conditions (b) are satisfied iff $X^{21} = 0, Y^{21} = 0, Z^{21} = 0$. Similarly for (c). These observations prove the lemma. □

3. Proof of Thm. 1.1

We are going to compute the value of γ for the structural tensor of the matrix multiplication map $\mathbb{C}^{e\times h}\times\mathbb{C}^{h\times\ell}\to\mathbb{C}^{e\times\ell}, (A,B)\mapsto AB$. We denote this tensor as usual by $\langle e,h,\ell\rangle \in k^{e\times h}\otimes k^{h\times\ell}\otimes k^{\ell\times e}$; it is explicitly described by

$$\langle e,h,\ell\rangle_{(pr)(\rho\sigma)(s\pi)} = \delta_{r\rho}\delta_{s\sigma}\delta_{p\pi},$$

where $\delta_{r\rho}$ is the Kronecker delta (cf. [5, (14.20)]). It can be directly checked from the definition that $\langle e,h,\ell\rangle$ is concise (cf. [5, Prop. (14.41)]). Later on, we will need the fundamental relation (cf. [5, Prop. (14.26)])

$$\langle e_1,h_1,\ell_1\rangle \otimes \langle e_2,h_2,\ell_2\rangle \simeq \langle e_1e_2,h_1h_2,\ell_1\ell_2\rangle. \tag{3.1}$$

LEMMA 3.1. *We have* $\gamma(\langle e,h,\ell\rangle) = e^2+h^2+\ell^2-1$ *for all* $e,h,\ell \geq 1$.

PROOF. The system of linear equations in Lemma 2.4 corresponding to the matrix tensor $\langle e,h,\ell\rangle$ reads as

$$\delta_{s\sigma}X_{(pr)(\pi\rho)} + \delta_{p\pi}Y_{(\rho\sigma)(rs)} + \delta_{r\rho}Z_{(s\pi)(\sigma p)} = 0 \tag{3.2}$$

for all $1\leq p,\pi\leq e$, $1\leq r,\rho\leq h$, $1\leq s,\sigma\leq \ell$. This implies for all $p\neq\pi$, $r\neq\rho$, $s\neq\sigma$

$$X_{(pr)(\pi\rho)} = Y_{(\rho\sigma)(rs)} = Z_{(s\pi)(\sigma p)} = 0.$$

Moreover, we obtain from (3.2)

$$\begin{array}{lcl} \forall p\ \forall r\neq\rho\ \forall s & : & X_{(pr)(p\rho)} = X_{(1r)(1\rho)} = -Y_{(\rho 1)(r1)} = -Y_{(\rho s)(rs)},\\ \forall p\ \forall \rho\ \forall s\neq\sigma & : & Y_{(\rho\sigma)(\rho s)} = Y_{(1\sigma)(1s)} = -Z_{(s1)(\sigma 1)} = -Z_{(sp)(\sigma p)},\\ \forall p\neq\pi\ \forall r\ \forall s & : & Z_{(s\pi)(sp)} = Z_{(1\pi)(1p)} = -X_{(p1)(\pi 1)} = -X_{(pr)(\pi r)}. \end{array}$$

Furthermore, (3.2) implies for all p,r,s

$$X_{(pr)(pr)} + Y_{(rs)(rs)} + Z_{(sp)(sp)} = 0. \tag{3.3}$$

From this we see that the solutions (X,Y,Z) of the linear system of equations (3.2) are uniquely determined by the values of $X_{(1r)(1\rho)}$, $Y_{(1\sigma)(1s)}$, $Z_{(1\pi)(1p)}$ for $p\neq\pi$, $r\neq\rho$, $s\neq\sigma$, which may be chosen arbitrarily, and by the values of $X_{(pr)(pr)}$, $Y_{(rs)(rs)}$, $Z_{(sp)(sp)}$, which are to be chosen subject to the constraint (3.3). On the other hand, it is easy to check that the solutions of (3.3) may be uniquely obtained as

$$X_{(pr)(pr)} = -u_p - v_r,\ Y_{(rs)(rs)} = v_r + w_s,\ Z_{(sp)(sp)} = u_p - w_s$$

where $u_1,\ldots,u_e,v_1,\ldots,v_h,w_2,\ldots,w_\ell \in k$ and $w_1 = 0$. Altogether, we conclude that the dimension of the solution space of the linear system of equations (3.2) equals

$$(e^2-e) + (h^2-h) + (\ell^2-\ell) + (e+h+\ell-1) = e^2+h^2+\ell^2-1$$

which shows that $\gamma(\langle e,h,\ell\rangle) = e^2+h^2+\ell^2-1$. □

In [17] Schönhage discovered that the border rank is not additive. His counterexample is

$$\underline{R}(\langle e,1,\ell\rangle \oplus \langle 1,n,1\rangle) = e\ell+1,$$

for $n=(e-1)(\ell-1)$, while $\underline{R}(\langle e,1,\ell\rangle) = e\ell$ and $\underline{R}(\langle 1,n,1\rangle) = n$. Using this, we can now supply the proof of Thm. 1.1. We will denote the r-fold direct sum of a tensor t by rt.

PROOF. (of Thm. 1.1) By Schönhage's example, we have

$$s_1 := \langle e,1,2\rangle \oplus \langle 1,e-1,1\rangle \trianglelefteq \langle 2e+1\rangle =: t_1.$$

Moreover, obviously $s_2 := \langle 1,2,1\rangle \trianglelefteq \langle 2\rangle =: t_2$. Using (3.1) it is easy to check that $s := (s_1 \otimes t_2) \oplus (s_2 \otimes t_1)$ and $t := (s_1 \otimes s_2) \oplus (t_1 \otimes t_2)$ can be expressed as

$$\begin{aligned} s &\simeq 2\langle e,1,2\rangle \oplus 2\langle 1,e-1,1\rangle \oplus (2e+1)\langle 1,2,1\rangle, \\ t &\simeq \langle e,2,2\rangle \oplus \langle 1,2e-2,1\rangle \oplus (4e+2)\langle 1,1,1\rangle. \end{aligned}$$

Let $[t] \in \mathcal{B}^+$ denote the isomorphism class of a concise tensor t. The asymptotic degeneration $\lesssim$ allows an extension to the ring $\mathcal{B}$ of tensor classes, compatible with the ring operations. Thus we may conclude from $[s_i] \lesssim [t_i]$ that $0 \lesssim [t_i] - [s_i]$ for $i = 1, 2$. Hence $0 \lesssim ([t_1] - [s_1])([t_2] - [s_2])$, which implies

$$[s] = [s_1][t_2] + [s_2][t_1] \lesssim [s_1][s_2] + [t_1][t_2] = [t].$$

Thus s is an asymptotic degeneration of t. If the degeneration allowed a compatible extension to the ring $\mathcal{B}$, then we could conclude that $s \trianglelefteq t$ in the same way. It is therefore sufficient to prove that this is not the case.

With Prop. 2.6 and Lemma 3.1 we obtain by a short computation (recall that matrix tensors are concise)

$$\gamma(s) = 4e^2 + 6e + 17, \ \gamma(t) = 5e^2 + 16.$$

This implies that $\gamma(s) < \gamma(t)$ for $e \geq 7$. Now the point is that s and t have the same format $(m,n,p) := (8e, 6e+4, 6e+3)$, so both can be interpreted as (concise) tensors in $\mathbb{C}^{m\times n\times p}$! Therefore, Thm. 2.2 implies that s is not a degeneration of t, which finishes the proof. □

Acknowledgements

The results of this paper form a part of the author's doctoral thesis [4] written under the direction of Professor V. Strassen. I wish to express my gratitude to him for his support and many helpful discussions.

References

[1] D. Bini. Relation between exact and approximate bilinear algorithms. Applications. *Calcolo*, 17:87–97, 1980.

[2] D. Bini, M. Capovani, G. Lotti, and F. Romani. $O(n^{2.7799})$ complexity for matrix multiplication. *Inf. Proc. Letters*, 8:234–235, 1979.

[3] D. Bini and V. Pan. *Polynomial and Matrix Computations*, Volume I of *Progress in Theoretical Computer Science*. Birkhäuser, Boston, 1994.

[4] P. Bürgisser. *Degenerationsordnung und Trägerfunktional bilinearer Abbildungen*. PhD thesis, Konstanz University, 1990.

[5] P. Bürgisser, M. Clausen, and M.A. Shokrollahi. *Algebraic Complexity Theory*, Volume 315 of *Grundlehren der mathematischen Wissenschaften*. Springer Verlag, 1996.

[6] D. Coppersmith and S. Winograd. Matrix multiplication via arithmetic progressions. *J. Symb. Comp.*, 9:251–280, 1990.

[7] D.W. Dubois. A note on David Harrison's theory of preprimes. *Pacific J. Math.*, 21:15–19, 1967.

[8] D.W. Dubois. Second note on David Harrison's theory of preprimes. *Pacific J. Math.*, 24:57–68, 1968.

[9] H.F. de Groote. On varieties of optimal algorithms for the computation of bilinear mappings I: The isotropy group of bilinear mapping. *Theoret. Comp. Sc.*, 7:1–24, 1978.

[10] H.F. de Groote. On varieties of optimal algorithms for the computation of bilinear mappings II: optimal algorithms for 2×2-matrix multiplication. *Theoret. Comp. Sc.*, 7:127–148, 1978.

[11] H.F. de Groote. Characterization of division algebras of minimal rank and the structure of their algorithm varieties. *SIAM J. Comp.*, 12:101–117, 1983.

[12] H.F. de Groote. *Lectures on the Complexity of Bilinear Problems.* Number 245 in LNCS. Springer Verlag, 1986.

[13] R.V. Kadison. A representation theory for commutative topological algebra. *Mem. Amer. Math. Soc.*, 7, 1951.

[14] D. Mumford. *Algebraic Geometry I: Complex Projective Varieties.* Springer Verlag, 1976.

[15] V.Ya. Pan. Field extension and trilinear aggregating, uniting and cancelling for the acceleration of matrix multiplication. In *Proc. of the 20th Ann. IEEE Symp. on Foundations of Comp. Sc.*, pages 28–38, 1979.

[16] V.Ya. Pan. *How to multiply matrices faster.* Number 179 in LNCS. Springer Verlag, 1984.

[17] A. Schönhage. Partial and total matrix multiplication. *SIAM J. Comp.*, 10:434–455, 1981.

[18] M.H. Stone. A general theory of spectra I. *Proc. Nat. Acad. Sc.*, 26:280–283, 1940.

[19] V. Strassen. Relative bilinear complexity and matrix multiplication. *Crelles J. Reine Angew. Math.*, 375/376:406–443, 1987.

[20] V. Strassen. The asymptotic spectrum of tensors. *Crelles J. Reine Angew. Math.*, 384:102–152, 1988.

[21] V. Strassen. Degeneration and complexity of bilinear maps: some asymptotic spectra. *Crelles J. Reine Angew. Math.*, 413:127–180, 1991.