Laderman matrix multiplication algorithm can be constructed using Strassen algorithm and related tensor's isotropies

Alexandre Sedoglavic

Laderman matrix multiplication algorithm can be constructed using Strassen algorithm and related tensor's isotropies

Alexandre.Sedoglavic@univ-lille.fr

March 28, 2017

1 Introduction

In [20], V. Strassen improves the classical 2×2 matrix multiplication algorithm. The current upper bound for 3×3 matrix multiplication was reached by J.B. Laderman in [16]. This note presents a *geometric* relationship between Strassen and Laderman algorithms. By doing so, we retrieve a *geometric* formulation of results very similar to those presented by O. Sýkora in [21].

1.1 Disclaimer: there is no improvement in this note

We do not improve any practical algorithm or prove any theoretical bound in this short note but focus on effective manipulation of tensor associated to matrix multiplication algorithm. To do so, we present only the minimal number of needed definitions and thus leave many facts outside our scope. We refer to [17] for a complete description of the field and to [1] for a state of the art presentation of theoretical complexity issues.

1.2 So, why writing (or reading) it?

We follow the geometric spirit of [12, 8, 4, 6, 5] and related papers: symmetries could be used in practical design of matrix multiplication algorithms. Hence, this note presents another example of this philosophy by giving a precise geometric meaning to the following statement:

> the Laderman matrix multiplication algorithm is composed by four 2×2 optimal matrix multiplication algorithms, an half of the classical 2×2 matrix multiplication algorithm and a correction term.

2 Framework

To do so, we have to present a small part of the classical framework (see [13, 14, 17]) mainly because we do not take it literally and only use a simplified version. Let us start by some basic definitions and notations as the following generic matrices:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}, \; B = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix}, \; C = \begin{pmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{pmatrix}, \quad (1)$$

that will be used in the sequel. Furthermore, as we also consider their 2×2 submatrices, let us introduce some associated notations.

Notations 2.1 *Let n, i, j be positive integers such that $i \leq n$ and $j \leq n$. We denote by $\mathrm{Id}^j_{n\times n}$ the identity $n \times n$ matrix where the jth diagonal term is 0. Given a $n \times n$ matrix A, we denote by $\widetilde{A}^{jk}$ the matrix $\mathrm{Id}^j_{n\times n} \cdot A \cdot \mathrm{Id}^k_{n\times n}$. For example, the matrix $\widetilde{A}^{33}$, $\widetilde{B}^{32}$ and $\widetilde{C}^{23}$ are:*

$$\begin{pmatrix} a_{11} & a_{12} & 0 \\ a_{21} & a_{22} & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} b_{11} & 0 & b_{13} \\ b_{21} & 0 & b_{23} \\ 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} c_{11} & c_{12} & 0 \\ 0 & 0 & 0 \\ c_{31} & c_{32} & 0 \end{pmatrix}. \quad (2)$$

Given a $n \times n$ matrix A, we sometimes consider $\widetilde{A}^{ij}$ as the $(n-1) \times (n-1)$ matrix $\overline{A^{ij}}$ where the line and column composed of 0 are removed.

At the opposite, given any $(n-1) \times (n-1)$ matrix A, we denotes by $\underline{A}_{ij}$ the $n \times n$ matrix where a line and column of 0 were added to A in order to have $\overline{\underline{A}_{ij}{}^{ij}} = A$.

2.1 Strassen multiplication algorithm

Considered as 2×2 matrices, the matrix product $\overline{C^{33}} = \overline{A^{33}} \cdot \overline{B^{33}}$ could be computed using Strassen algorithm (see [20]) by performing the following computations:

$$\begin{aligned} t_1 = (a_{11} + a_{22})(b_{11} + b_{22}), t_2 = (a_{12} - a_{22})(b_{21} + b_{22}), \\ t_3 = (-a_{11} + a_{21})(b_{11} + b_{12}), t_4 = (a_{11} + a_{12})b_{22}, \\ t_5 = a_{11}(b_{12} - b_{22}), t_6 = a_{22}(-b_{11} + b_{21}), t_7 = (a_{21} + a_{22})b_{11}, \\ c_{11} = t_1 + t_2 - t_4 + t_6, c_{12} = t_6 + t_7, \\ c_{21} = t_4 + t_5, c_{22} = t_1 + t_3 + t_5 - t_7. \end{aligned} \quad (3)$$

In order to consider above algorithm under a geometric standpoint, it is usually presented as a tensor.

2.2 Bilinear mappings seen as tensors and associated trilinear forms

Definitions 2.1 *Given a tensor:*

$$\mathcal{T} = \sum_{i=1}^{r} T_{i1} \otimes T_{i2} \otimes T_{i3}, \quad (4)$$

where T_{ij} are $n \times n$ matrices and $\mathcal{T}$ is decomposed as sum of rank-one tensors:

- *the integer r is the* tensor rank *of tensor $\mathcal{T}$;*
- *the unordered list $[(\mathrm{rank}\, M_{ij})_{j=1..3}]_{i=1..r}$ is called the* type *of tensor $\mathcal{T}$ (*$\mathrm{rank}\, A$ *being the classical rank of the matrix A).*

2.3 Tensors' contractions

To explicit the relationship between what is done in the sequel and the bilinear mapping associated to matrix multiplication, let us consider the following tensor's contractions:

Definitions 2.2 *Using the notation of definition 2.1 given a tensor $\mathcal{T}$ and three $n \times n$ matrices A, B and C with coefficients in the algebra $\mathbb{K}$:*

- *the $(1,2)$ contraction of $\mathcal{T} \otimes A \otimes B$ defined by:*

$$\sum_{i=1}^{r} \mathrm{Trace}({}^{t}T_{i1} \cdot A)\, \mathrm{Trace}({}^{t}T_{i2} \cdot B) T_{i3} \tag{5}$$

  corresponds to a bilinear application $\mathbb{K}^{n\times n} \times \mathbb{K}^{n\times n} \mapsto \mathbb{K}^{n\times n}$ with indeterminates A and B.

- *the $(1,2,3)$ (a.k.a full) contraction of $\mathcal{T} \otimes A \otimes B \otimes C$ defined by:*

$$\langle \mathcal{T} | A \otimes B \otimes C \rangle = \sum_{i=1}^{r} \mathrm{Trace}({}^{t}T_{i1} \cdot A)\, \mathrm{Trace}({}^{t}T_{i2} \cdot B)\, \mathrm{Trace}({}^{t}T_{i3} \cdot C) \tag{6}$$

  corresponds to a trilinear form $\mathbb{K}^{n\times n} \times \mathbb{K}^{n\times n} \times \mathbb{K}^{n\times n} \mapsto \mathbb{K}$ with indeterminates A, B and C.

Remarks 2.1 *As the studied object is the tensor, its expressions as full or incomplete contractions are equivalent. Thus, even if matrix multiplication is a bilinear application, we are going to work in the sequel with trilinear forms (see [10] for bibliographic references on this standpoint).*

The definition in 2.2 are taken to express the full contraction as a degenerate inner product between tensors; it is not the usual choice made in the literature and so, we have to explicitly recall some notions used in the sequel.

Strassen multiplication algorithm (3) is equivalent to the tensor $\mathcal{S}$ defined by:

$$\begin{gathered}
\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 0 & -1 \end{pmatrix} \otimes \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \\
\begin{pmatrix} -1 & 0 \\ 1 & 0 \end{pmatrix} \otimes \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \otimes \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \otimes \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} -1 & 0 \\ 1 & 0 \end{pmatrix} + \\
\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \otimes \begin{pmatrix} 0 & 1 \\ 0 & -1 \end{pmatrix} \otimes \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} -1 & 0 \\ 1 & 0 \end{pmatrix} \otimes \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} + \\
\begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \otimes \begin{pmatrix} 0 & 1 \\ 0 & -1 \end{pmatrix}
\end{gathered} \tag{7}$$

and one can check that this tensor defines the matrix multiplication algorithm (3) and that its tensor rank is 7.

2.4 2×2 matrix multiplication tensors induced by a 3×3 matrix multiplication tensor

. Given any 3×3 matrix multiplication tensor, one can define 3^3 induced 2×2 matrix multiplication tensors as follow. First, let us introduce the following operators that generalize to tensor the notations 2.1:

Definitions 2.3 *Using notation introduced in definition 2.1, we define:*

$$\widetilde{A \otimes B \otimes C}^{ijk} = \widetilde{A}^{ij} \otimes \widetilde{B}^{jk} \otimes \widetilde{C}^{ki}, \tag{8a}$$

$$\overline{A \otimes B \otimes C}^{ijk} = \overline{A^{ij}} \otimes \overline{B^{jk}} \otimes \overline{C^{ki}}, \tag{8b}$$

$$\underline{A \otimes B \otimes C}_{ijk} = \underline{A}_{ij} \otimes \underline{B}_{jk} \otimes \underline{C}_{ki} \tag{8c}$$

and we extend the definitions of these operators by additivity in order to be applied on any tensor $\mathcal{T}$ described in definition 2.1.

There is n^3 such projections and given any matrix multiplication tensor $\mathcal{M}$, the full contraction satisfying the following trivial properties:

$$\left\langle \mathcal{M} | \widetilde{A \otimes B \otimes C}^{ijk} \right\rangle = \left\langle \widetilde{\mathcal{M}}^{ijk} \,\middle|\, A \otimes B \otimes C \right\rangle = \left\langle \overline{\mathcal{M}}^{ijk} \,\middle|\, \overline{A \otimes B \otimes C}^{ijk} \right\rangle \tag{9}$$

where the projection operator apply on a $n \times n$ matrix multiplication tensor defines explicitly a $(n-1) \times (n-1)$ matrix multiplication tensor.

The following property holds:

Lemma 2.1

$$(n-1)^3 \langle \mathcal{M} | A \otimes B \otimes C \rangle = \sum_{1 \leq i,j,k \leq n} \left\langle \mathcal{M} \middle| \widetilde{A \otimes B \otimes C}^{ijk} \right\rangle \tag{10}$$

and thus, we have:

$$\langle \mathcal{M} | A \otimes B \otimes C \rangle = \left\langle \frac{1}{(n-1)^3} \sum_{1 \leq i,j,k \leq n} \widetilde{\mathcal{M}}^{ijk} \middle| A \otimes B \otimes C \right\rangle. \tag{11}$$

The obvious facts made in this section underline the relationships between any $n \times n$ matrix multiplication tensor and the n^3 induced $(n-1) \times (n-1)$ algorithms.

Considering the Laderman matrix multiplication tensor, we are going to explore further this kind of relationships. First, let us introduce this tensor.

2.5 Laderman matrix multiplication tensor

The following tensor — described here by giving the full contraction $\langle \mathcal{L} | A \otimes B \otimes C \rangle$:

$$
\begin{array}{ll}
(a_{11} - a_{21} + a_{12} - a_{22} - a_{32} + a_{13} - a_{33})\, b_{22}\, c_{21} & + \\
a_{22}\, (-b_{11} + b_{21} - b_{31} + b_{12} - b_{22} - b_{23} + b_{33})\, c_{12} & + \\
a_{13}\, b_{31}\, (c_{11} + c_{21} + c_{31} + c_{12} + c_{32} + c_{13} + c_{23}) & + \\
(a_{11} - a_{31} + a_{12} - a_{22} - a_{32} + a_{13} - a_{23})\, b_{23}\, c_{31} & + \\
a_{32}\, (-b_{11} + b_{21} - b_{31} - b_{22} + b_{32} + b_{13} - b_{23})\, c_{13} & + \\
a_{11}\, b_{11}\, (c_{11} + c_{21} + c_{31} + c_{12} + c_{22} + c_{13} + c_{33}) & + \\
(-a_{11} + a_{31} + a_{32})\, (b_{11} - b_{13} + b_{23})\, (c_{31} + c_{13} + c_{33}) & + \\
(a_{22} - a_{13} + a_{23})\, (b_{31} + b_{23} - b_{33})\, (c_{31} + c_{12} + c_{32}) & + \\
(-a_{11} + a_{21} + a_{22})\, (b_{11} - b_{12} + b_{22})\, (c_{21} + c_{12} + c_{22}) & + \\
(a_{32} - a_{13} + a_{33})\, (b_{31} + b_{22} - b_{32})\, (c_{21} + c_{13} + c_{23}) & + \\
(a_{21} + a_{22})\, (-b_{11} + b_{12})\, (c_{21} + c_{22}) & + \\
(a_{31} + a_{32})\, (-b_{11} + b_{13})\, (c_{31} + c_{33}) & + \\
(a_{13} - a_{33})\, (b_{22} - b_{32})\, (c_{13} + c_{23}) & + \\
(a_{11} - a_{21})\, (-b_{12} + b_{22})\, (c_{12} + c_{22}) & + \\
(a_{32} + a_{33})\, (-b_{31} + b_{32})\, (c_{21} + c_{23}) & + \\
(-a_{11} + a_{31})\, (b_{13} - b_{23})\, (c_{13} + c_{33}) & + \\
(a_{13} - a_{23})\, (b_{23} - b_{33})\, (c_{12} + c_{32}) & + \\
(a_{22} + a_{23})\, (-b_{31} + b_{33})\, (c_{31} + c_{32}) & + \\
a_{12}\, b_{21}\, c_{11} + a_{23}\, b_{32}\, c_{22} + a_{21}\, b_{13}\, c_{32} + a_{31}\, b_{12}\, c_{23} + a_{33}\, b_{33}\, c_{33} &
\end{array}
\tag{12}
$$

was introduced in [16] (we do not study in this note any other *inequivalent* algorithm of same tensor rank e.g. [15, 9, 19], etc). Considering the projections introduced in definition 2.3, we notice that:

Remark 2.2 *Laderman matrix multiplication tensor defines* 4 *optimal* 2×2 *matrix multiplication tensors and* 19 *other such tensors of tensor rank* 8*.*

Further computations show that:

Remark 2.3 *The type of the Laderman matrix multiplication tensor is*

$$\left[(2,2,2)|4, ((1,3,1),(3,1,1),(1,1,3))|2, (1,1,1)|13\right] \tag{13}$$

where $m|n$ *indicates that* m *is repeated* n *times.*

2.6 Tensors' isotropies

We refer to [13, 14] for a complete presentation of automorphism group operating on varieties defined by algorithms for computation of bilinear mappings and as a reference for the following theorem:

Theorem 2.1 *The isotropy group of the* $n \times n$ *matrix multiplication tensor is*

$$\mathrm{PGL}(\mathbb{C}^n)^{\times 3} \rtimes \mathfrak{S}_6, \tag{14}$$

where PGL *stands for the projective linear group and* $\mathfrak{S}_6$ *for the symmetric group on* 6 *elements.*

Even if we do not completely explicit the concrete action of this isotropy group on matrix multiplication tensor, let us precise some terminologies:

Definitions 2.4 *Given a tensor defining matrix multiplication computations, the orbit of this tensor is called the* multiplication algorithm *and any of the points composing this orbit is a* variant *of this algorithm.*

Remark 2.4 *As shown in [11], matrix multiplication is characterised by its isotropy group.*

Remark 2.5 *We only need the* $\mathrm{PGL}(\mathbb{C}^n)^{\times 3}$ *part of this group (a.k.a. sandwiching) and thus focus on it in the sequel.*

As our framework and notations differ slightly from the framework classically found in the literature, we have to explicitly define several well-known notions for the sake of clarity. Hence, let us recall the *sandwiching* action:

Definition 2.5 *Given* $g = (G_1 \times G_2 \times G_3)$ *an element of* $\mathrm{PGL}(\mathbb{C}^n)^{\times 3}$*, its action on a tensor* $\mathcal{T}$ *is given by:*

$$\begin{aligned} g \diamond \mathcal{T} &= \sum_{i=1}^{r} g \diamond (T_{i1} \otimes T_{i2} \otimes T_{i3}), \\ &= \sum_{i=1}^{r} \left({}^tG_1^{-1} T_{i1} {}^tG_2\right) \otimes \left({}^tG_2^{-1} T_{i2} {}^tG_3\right) \otimes \left({}^tG_3^{-1} T_{i3} {}^tG_1\right). \end{aligned} \tag{15}$$

Example 2.1 *Let us consider the action of the following isotropy*

$$\rho = \begin{pmatrix} 0 & 1/\lambda \\ -1 & 0 \end{pmatrix} \times \begin{pmatrix} 1/\lambda & -1/\lambda \\ 0 & 1 \end{pmatrix} \times \begin{pmatrix} -1/\lambda & 0 \\ 1 & -1 \end{pmatrix} \tag{16}$$

on the Strassen variant of the Strassen algorithm. The resulting tensor $\mathcal{W}$ *is:*

$$\begin{aligned} \sum_{i=1}^{7} w_i = & \begin{pmatrix} -1 & \lambda \\ -\frac{1}{\lambda} & 0 \end{pmatrix} \otimes \begin{pmatrix} 1 & -\lambda \\ \frac{1}{\lambda} & 0 \end{pmatrix} \otimes \begin{pmatrix} 1 & -\lambda \\ \frac{1}{\lambda} & 0 \end{pmatrix} + \begin{pmatrix} -1 & l \\ -\frac{1}{\lambda} & 1 \end{pmatrix} \otimes \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \otimes \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \\ & + \begin{pmatrix} 1 & 0 \\ \frac{1}{\lambda} & 0 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ \frac{1}{\lambda} & 0 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ \frac{1}{\lambda} & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \\ & + \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \otimes \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \otimes \begin{pmatrix} -1 & \lambda \\ -\frac{1}{\lambda} & 1 \end{pmatrix} + \begin{pmatrix} 1 & -\lambda \\ 0 & 0 \end{pmatrix} \otimes \begin{pmatrix} 1 & -\lambda \\ 0 & 0 \end{pmatrix} \otimes \begin{pmatrix} 1 & -\lambda \\ 0 & 0 \end{pmatrix} \\ & + \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \otimes \begin{pmatrix} -1 & \lambda \\ -\frac{1}{\lambda} & 1 \end{pmatrix} \otimes \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \end{aligned} \tag{17}$$

that is the well-known Winograd variant of Strassen algorithm.

Remarks 2.6 *We keep the parameter* λ *useless in our presentation as a tribute to the construction made in [7] that gives an elegant and elementary (i.e. based on matrix eigenvalues) construction of Winograd variant of Strassen matrix multiplication algorithm.*

This variant is remarkable in its own as shown in [3] because it is optimal w.r.t. multiplicative and *additive complexity.*

Remark 2.7 *Tensor's type is an invariant of isotropy's action. Hence, two tensor in the same orbit share the same type. Or equivalently, two tensors with the same type are two variants that represent the same matrix multiplication algorithm.*

This remark will allows us in section 3.4 to recognise the tensor constructed below as a variant of the Laderman matrix multiplication algorithm.

3 A tensor's construction

Let us now present the construction of a variant of Laderman matrix multiplication algorithm based on Winograd variant of Strassen matrix multiplication algorithm.

First, let us give the full contraction of the tensor $\underline{\mathcal{W}}_{111} \otimes A \otimes B \otimes C$:

$$\left(-a_{22} - \frac{a_{32}}{\lambda} + \lambda a_{23}\right)\left(b_{22} + \frac{b_{32}}{\lambda} - \lambda b_{23}\right)\left(c_{22} + \frac{c_{32}}{\lambda} - \lambda c_{23}\right) + \tag{18a}$$

$$(a_{22} - \lambda a_{23})\,(b_{22} - \lambda b_{23})\,(c_{22} - \lambda c_{23}) + \tag{18b}$$

$$\left(a_{22} + \frac{a_{32}}{\lambda}\right)\left(b_{22} + \frac{b_{32}}{\lambda}\right)\left(c_{22} + \frac{c_{32}}{\lambda}\right) + \tag{18c}$$

$$a_{23}\left(-b_{22} - \frac{b_{32}}{\lambda} + \lambda b_{23} + b_{33}\right) c_{32} + \tag{18d}$$

$$\left(-a_{22} - \frac{a_{32}}{\lambda} + \lambda a_{23} + a_{33}\right) b_{32}\, c_{23} + \tag{18e}$$

$$a_{32}\, b_{23}\left(-c_{22} - \frac{c_{32}}{\lambda} + \lambda c_{23} + c_{33}\right) + \tag{18f}$$

$$a_{33}\, b_{33}\, c_{33} \tag{18g}$$

3.1 A Klein four-group of isotropies

Let us introduce now the following notations:

$$\mathrm{Id}_{3\times 3} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad P_{(12)} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{19}$$

used to defined the following group of isotropies:

$$G = \left\{ \begin{array}{cc} g_1 = {\mathrm{Id}_{3\times 3}}^{\times 3}, & g_2 = \left(\mathrm{Id}_{3\times 3} \times P_{(12)} \times P_{(12)}\right), \\ g_3 = \left(P_{(12)} \times P_{(12)} \times \mathrm{Id}_{3\times 3}\right), & g_4 = \left(P_{(12)} \times \mathrm{Id}_{3\times 3} \times P_{(12)}\right) \end{array} \right\} \tag{20}$$

that is isomorphic to the Klein four-group.

3.2 Its action on Winograd variant of Strassen algorithm

In the sequel, we are interested in the action of Klein four-group (20) on our Winograd variant of Strassen algorithm:

$$G \diamond \underline{\mathcal{W}}_{111} = \sum_{g \in G} g \diamond \underline{\mathcal{W}}_{111} = \sum_{g \in G} \sum_{i=1}^{7} g \diamond \underline{w_i}_{111} \tag{21}$$

As we have for any isotropy g:

$$\langle g \diamond \underline{\mathcal{W}}_{111} | A \otimes B \otimes C \rangle = \langle \underline{\mathcal{W}}_{111} | g \diamond (A \otimes B \otimes C) \rangle, \tag{22}$$

the action of isotropies g_i is just a permutation of our generic matrix coefficients. Hence, we have the full contraction of the tensor $(g_2 \diamond \underline{\mathcal{W}}_{111}) \otimes A \otimes B \otimes C$:

$$\left(-a_{21} - \frac{a_{31}}{\lambda} + \lambda a_{23}\right)\left(b_{11} + \frac{b_{31}}{\lambda} - \lambda b_{13}\right)\left(c_{12} + \frac{c_{32}}{\lambda} - \lambda c_{13}\right) + \tag{23a}$$

$$(a_{21} - \lambda a_{23})\,(b_{11} - \lambda b_{13})\,(c_{12} - \lambda c_{13}) + \tag{23b}$$

$$\left(a_{21} + \frac{a_{31}}{\lambda}\right)\left(b_{11} + \frac{b_{31}}{\lambda}\right)\left(c_{12} + \frac{c_{32}}{\lambda}\right) + \tag{23c}$$

$$a_{23}\left(-b_{11} - \frac{b_{31}}{\lambda} + \lambda b_{13} + b_{33}\right) c_{32} + \tag{23d}$$

$$\left(-a_{21} - \frac{a_{31}}{\lambda} + \lambda a_{23} + a_{33}\right) b_{31}\, c_{13} + \tag{23e}$$

$$a_{31}\, b_{13}\left(-c_{12} - \frac{c_{32}}{\lambda} + \lambda c_{13} + c_{33}\right) + \tag{23f}$$

$$a_{33}\, b_{33}\, c_{33}, \tag{23g}$$

the full contraction of the tensor $(g_3 \diamond \underline{\mathcal{W}}_{111}) \otimes A \otimes B \otimes C$:

$$\left(-a_{11} - \frac{a_{31}}{\lambda} + \lambda a_{13}\right)\left(b_{12} + \frac{b_{32}}{\lambda} - \lambda b_{13}\right)\left(c_{21} + \frac{c_{31}}{\lambda} - \lambda c_{23}\right) + \tag{24a}$$

$$(a_{11} - \lambda a_{13})\,(b_{12} - \lambda b_{13})\,(c_{21} - \lambda c_{23}) + \tag{24b}$$

$$\left(a_{11} + \frac{a_{31}}{\lambda}\right)\left(b_{12} + \frac{b_{32}}{\lambda}\right)\left(c_{21} + \frac{c_{31}}{\lambda}\right) + \tag{24c}$$

$$a_{13}\left(-b_{12} - \frac{b_{32}}{\lambda} + \lambda b_{13} + b_{33}\right) c_{31} + \tag{24d}$$

$$\left(-a_{11} - \frac{a_{31}}{\lambda} + \lambda a_{13} + a_{33}\right) b_{32}\, c_{23} + \tag{24e}$$

$$a_{31}\, b_{13}\left(-c_{21} - \frac{c_{31}}{\lambda} + \lambda c_{23} + c_{33}\right) + \tag{24f}$$

$$a_{33}\, b_{33}\, c_{33} \tag{24g}$$

and the full contraction of the tensor $(g_4 \diamond \underline{\mathcal{W}}_{111}) \otimes A \otimes B \otimes C$:

$$\left(-a_{12} - \frac{a_{32}}{\lambda} + \lambda a_{13}\right)\left(b_{21} + \frac{b_{31}}{\lambda} - \lambda b_{23}\right)\left(c_{11} + \frac{c_{31}}{\lambda} - \lambda c_{13}\right) + \tag{25a}$$

$$(a_{12} - \lambda a_{13})\,(b_{21} - \lambda b_{23})\,(c_{11} - \lambda c_{13}) + \tag{25b}$$

$$\left(a_{12} + \frac{a_{32}}{\lambda}\right)\left(b_{21} + \frac{b_{31}}{\lambda}\right)\left(c_{11} + \frac{c_{31}}{\lambda}\right) + \tag{25c}$$

$$a_{13}\left(-b_{21} - \frac{b_{31}}{\lambda} + \lambda b_{23} + b_{33}\right) c_{31} + \tag{25d}$$

$$\left(-a_{12} - \frac{a_{32}}{\lambda} + \lambda a_{13} + a_{33}\right) b_{31}\, c_{13} + \tag{25e}$$

$$a_{32}\, b_{23}\left(-c_{11} - \frac{c_{31}}{\lambda} + \lambda c_{13} + c_{33}\right) + \tag{25f}$$

$$a_{33}\, b_{33}\, c_{33}. \tag{25g}$$

There is several noteworthy points in theses expressions:

Remarks 3.1 • *the term (18g) is a fixed point of G's action;*

- *the term (18d) and (23d), (18e) and (24e), (18f) and (25f), (23e) and (25e), (23f) and (24f), (24d) and (25d) could be* added *in order to obtain new rank-on tensor without changing the tensor rank. For example (18d)+(23d) is equal to:*

$$a_{23}\left(-b_{22}-\frac{b_{32}}{\lambda}+\lambda b_{23}+2b_{33}-b_{11}-\frac{b_{31}}{\lambda}+\lambda b_{13}\right)c_{32}. \tag{26}$$

The tensor rank of the tensor $G \diamond \underline{\mathcal{W}}_{111} = \sum_{i=1} g_i \diamond \underline{\mathcal{W}}_{111}$ is $1+3\cdot 4+6=19$. Unfortunately, this tensor does not define a matrix multiplication algorithm (otherwise according to the lower bound presented in [2], it would be optimal and this note would have another title and impact).

In the next section, after studying the action of isotropy group G on the classical matrix multiplication algorithm, we are going to complete our tensor in order to obtain a matrix multiplication tensor.

3.3 How far are we from a multiplication tensor?

Let us consider the classical 3×3 matrix multiplication algorithm

$$\mathcal{M} = \sum_{1\leq i,j,k\leq 3} e^i_j \otimes e^j_k \otimes e^k_i \tag{27}$$

where e^i_j denotes the matrix with a single non-zero coefficient 1 at the intersection of line i and column j. By considering the trilinear monomial:

$$a_{ij}b_{jk}c_{ki} = \left\langle e^i_j \otimes e^j_k \otimes e^k_i \,\middle|\, A\otimes B\otimes C\right\rangle, \tag{28}$$

we describe below the action of an isotropy g on this tensor by the induced action:

$$\begin{aligned} g\diamond a_{ij}b_{jk}c_{ki} &= \left\langle g\diamond (e^i_j \otimes e^j_k \otimes e^k_i) \,\middle|\, A\otimes B\otimes C\right\rangle, \\ &= \left\langle e^i_j \otimes e^j_k \otimes e^k_i \,\middle|\, g\diamond (A\otimes B\otimes C)\right\rangle. \end{aligned} \tag{29}$$

Remark 3.2 *The isotropies in G act as a permutation on rank-one composant of the tensor $\mathcal{M}$. More precisely, we have the following 9 orbits represented by the sums:*

$$\sum_{i=1}^{4} g_i \diamond a_{11}\, b_{11}\, c_{11} = a_{11}\, b_{11}\, c_{11} + a_{12}\, b_{22}\, c_{21} + a_{22}\, b_{21}\, c_{12} + a_{21}\, b_{12}\, c_{22}, \tag{30a}$$

$$\sum_{i=1}^{4} g_i \diamond a_{22}\, b_{22}\, c_{22} = a_{22}\, b_{22}\, c_{22} + a_{21}\, b_{11}\, c_{12} + a_{11}\, b_{12}\, c_{21} + a_{12}\, b_{21}\, c_{11}, \tag{30b}$$

$$\sum_{i=1}^{4} g_i \diamond a_{23}\, b_{32}\, c_{22} = a_{23}\, b_{32}\, c_{22} + a_{23}\, b_{31}\, c_{12} + a_{13}\, b_{32}\, c_{21} + a_{13}\, b_{31}\, c_{11}, \tag{30c}$$

$$\sum_{i=1}^{4} g_i \diamond a_{22}\, b_{23}\, c_{32} = a_{22}\, b_{23}\, c_{32} + a_{21}\, b_{13}\, c_{32} + a_{11}\, b_{13}\, c_{31} + a_{12}\, b_{23}\, c_{31}, \tag{30d}$$

$$\sum_{i=1}^{4} g_i \diamond a_{32}\, b_{22}\, c_{23} = a_{32}\, b_{22}\, c_{23} + a_{31}\, b_{11}\, c_{13} + a_{31}\, b_{12}\, c_{23} + a_{32}\, b_{21}\, c_{13}, \tag{30e}$$

$$\frac{1}{2}\sum_{i=1}^{4} g_i \diamond a_{23}\, b_{33}\, c_{32} = a_{23}\, b_{33}\, c_{32} + a_{13}\, b_{33}\, c_{31}, \tag{30f}$$

$$\frac{1}{2}\sum_{i=1}^{4} g_i \diamond a_{33}\, b_{32}\, c_{23} = a_{33}\, b_{32}\, c_{23} + a_{33}\, b_{31}\, c_{13}, \tag{30g}$$

$$\frac{1}{2}\sum_{i=1}^{4} g_i \diamond a_{32}\, b_{23}\, c_{33} = a_{32}\, b_{23}\, c_{33} + a_{31}\, b_{13}\, c_{33}, \tag{30h}$$

$$\frac{1}{4}\sum_{i=1}^{4} g_i \diamond a_{33}\, b_{33}\, c_{33} = a_{33}\, b_{33}\, c_{33}. \tag{30i}$$

Hence, the action of G decomposes the classical matrix multiplication tensor $\mathcal{M}$ as a transversal action of G on the implicit projection $\widetilde{\mathcal{M}}^{111}$, its action on the rank-one tensor $e_1^1 \otimes e_1^1 \otimes e_1^1$ and a correction term also related to orbits under G:

$$\begin{aligned} \mathcal{M} &= G \diamond \left(e_1^1 \otimes e_1^1 \otimes e_1^1\right) + G \diamond \widetilde{\mathcal{M}}^{111} - \mathcal{C}, \\ \mathcal{C} &= (1/2)\, G \diamond \left(e_3^2 \otimes e_3^3 \otimes e_2^3\right) + (1/2)\, G \diamond \left(e_3^3 \otimes e_2^3 \otimes e_3^2\right) \\ &\quad + (1/2)\, G \diamond \left(e_2^3 \otimes e_3^2 \otimes e_3^3\right) + 3\, G \diamond \left(e_3^3 \otimes e_3^3 \otimes e_3^3\right). \end{aligned} \tag{31}$$

The term $\widetilde{\mathcal{M}}^{111}$ is a 2×2 matrix multiplication algorithm that could be replaced by any other one. Choosing $\underline{\mathcal{W}}_{111}$, we have the following properties:

- the tensor rank of $G \diamond \underline{\mathcal{W}}_{111}$ is 19;
- its addition with the correction term $\mathcal{C}$ does not change its tensor rank.

3.4 Resulting matrix multiplication algorithm

Hence, we obtain a matrix multiplication tensor with rank $23 = 19 + 4$. Furthermore, the resulting tensor have the same type than the Laderman matrix multiplication tensor, and thus it is a variant of the same algorithm.

We conclude that the Laderman matrix multiplication algorithm can be constructed using the orbit of an optimal 2×2 matrix multiplication algorithm under the action of a given group leaving invariant classical 3×3 matrix multiplication variant/algorithm and with a transversal action on one of its projections.

4 Concluding remarks

All the observations presented in this short note came from an experimental mathematical approach using the computer algebra system Maple [18]. While implementing effectively (if not efficiently) several tools needed to manipulate matrix multiplication tensor—tensors, their isotropies and contractions, etc.—in order to understand the theory, the relationship between the Laderman matrix multiplication algorithm and the Strassen algorithm became clear by simple computations that will be tedious or impossible by hand.

As already shown in [21], this kind of geometric configuration could be found and used with other matrix size.

The main opinion supported by this work is that symmetries play a central role in effective computation for matrix multiplication algorithm and that only a geometrical interpretation may brings further improvement.

Acknowledgment The author would like to thanks Alin Bostan for providing information on the work [21].

References

[1] Ambainis, A., Filmus, Y., and Le Gall, F. Fast matrix multiplication: Limitations of the laser method. Tech. rep., arXiv, Nov. 2014.

[2] Bläser, M. On the complexity of the multiplication of matrices of small formats. *Journal of Complexity 19*, 1 (Feb. 2003), 43–60.

[3] Bshouty, N. H. On the additive complexity of 2×2 matrix multiplication. *Information Processing Letters 56*, 6 (Dec. 1995), 329–335.

[4] Bürgisser, P., and Ikenmeyer, C. Fundamental invariants of orbit closures. Tech. Rep. arXiv:1511.02927v2, arXiv, Dec. 2015.

[5] Burichenko, V. P. On symmetries of the Strassen algorithm. Tech. Rep. arXiv:1408.627v1, arXiv, Aug. 2014.

[6] Burichenko, V. P. Symmetries of matrix multiplications algorithms. I. Tech. rep., arXiv, Aug. 2015.

[7] Chatelin, Ph. On transformations of algorithms to multiply 2×2 matrices. *Information Processing Letters 22*, 1 (Jan. 1986), 1 – 5.

[8] Chiantini, L., Ikenmeyer, C., Landsberg, J. M., and Ottaviani, G. The geometry of rank decompositions of matrix multiplication I: 2×2 matrices. Tech. Rep. 1610.08364, arXiv, Oct. 2016.

[9] Courtois, N. T., Bard, G. V., and Hulme, D. A new general-purpose method to multiply 3×3 matrices using only 23 multiplications. Tech. rep., arXiv:1108.2830v1, 2011.

[10] Dumas, J.-G., and Pan, V. Y. Fast matrix multiplication and symbolic computation. Tech. Rep. 1612.05766, arXiv, Dec. 2016.

[11] Gesmundo, F. Geometric aspect of iterated matrix multiplication. *Journal of Algebra 461* (Sept. 2016), 42–64.

[12] Grochow, J. A., and Moore, C. Matrix multiplication algorithms from group orbits. Tech. Rep. arXiv:1612.01527, arXiv, Dec. 2016.

[13] Groot, de, H. F. On varieties of optimal algorithms for the computation of bilinear mappings I. The isotropy group of a bilinear mapping. *Theoretical Computer Science 7*, 2 (1978), 1–24.

[14] Groot, de, H. F. On varieties of optimal algorithms for the computation of bilinear mappings II. Optimal algorithms for 2×2-matrix multiplication. *Theoretical Computer Science 7*, 2 (1978), 127–148.

[15] Johnson, R. W., and McLoughlin, A. M. Noncommutative bilinear algorithms for 3×3 matrix multiplication. *SIAM Journal on Computing 15*, 2 (May 1986), 595–603.

[16] Laderman, J. B. A noncommutative algorithm for multiplying 3×3 matrices using 23 multiplications. *Bulletin of the American Mathematical Society 82*, 1 (Jan. 1976), 126–128.

[17] Landsberg, J. M. *Tensors: geometry and applications*, vol. 128 of *Graduate Studies in Mathematics*. 2010.

[18] Monagan, M. B., Geddes, K. O., Heal, K. M., Labahn, G., Vorkoetter, S. M., McCarron, J., and DeMarco, P. *Maple* 11 *Introductory programming guide*. Maplesoft, 2007. http://www.maplesoft.com.

[19] Oh, J., Kim, J., and Moon, B.-R. On the inequivalence of bilinear algorithms for 3×3 matrix multiplication. *Information Processing Letters 113*, 17 (Aug. 2013), 640–645.

[20] Strassen, V. Gaussian elimination is not optimal. *Numerische Mathematik 13*, 4 (Aug. 1969), 354–356.

[21] Sykora, O. A fast non-commutative algorithm for matrix multiplication. In *Proceedings of the 6th International Symposium on Mathematical Foundations of Computer Science.* (Tatranská Lomnica, Czechoslovakia, Sept. 5-9 1977), J. Gruska, Ed., vol. 53 of *Lecture Notes in Computer Science*, Springer, pp. 504–512.