ON VARIETIES OF OPTIMAL ALGORITHMS FOR THE COMPUTATION OF BILINEAR MAPPINGS
I. THE ISOTROPY GROUP OF A BILINEAR MAPPING*

Hans F. de GROOTE

Mathematisches Institut der Universität Tübingen, Tübingen, Federal Republic of Germany

Communicated by A. Schönhage
Received July 1977

Abstract. This paper contains the general frame of a theory of varieties of algorithms for the computation of bilinear mappings. In particular, we shall study linear mappings which operate on algorithm varieties. It will be shown that every bilinear mapping \( \Phi \) defines in a natural way a group of automorphisms operating on the variety of optimal algorithms for \( \Phi \). This group is called the isotropy group of \( \Phi \). For some important classes of bilinear mappings these groups will be determined. Applications of these results will appear in parts II and III of this work.

0. Introduction

This work is devoted to the study of varieties of optimal algorithms for the computation of bilinear mappings. Roughly speaking, our problem is as follows: given a field \( K \) and a bilinear mapping

\[
\Phi : K^l \times K^m \to K^n,
\]

what are the different optimal ways to compute \( \Phi(x, y) \) from \( x \) and \( y \)? For this, first of all we have to furnish a model of computation. Here we adopt the usual non-commutative one: To compute \( (z_1, \ldots, z_n) := \Phi(x, y) \) from \( x = (x_1, \ldots, x_I) \) and \( y = (y_1, \ldots, y_m) \) means to form some products

\[
p_r := \left( \sum_{\lambda=1}^{I} u_{r\lambda} x_\lambda \right) \left( \sum_{\mu=1}^{\ell} v_{r\mu} y_\mu \right) \quad (r = 1, \ldots, R)
\]

such that \( \Phi(x, y) \) is obtained according to

\[
z_\nu = \sum_{r=1}^{R} w_{r\nu} p_r \quad (\nu = 1, \ldots, n), \tag{1}
\]

where \( u_{r\lambda}, v_{r\mu}, w_{r\nu} \in K \).

* This paper coincides with the first chapter of the author's 'Habilitationschrift'.

---

Theoretical Computer Science 7 (1978) 1-27
© North-Holland Publishing Company
The computations with the minimal $R$ possible in (1) are called optimal. As $\Phi$ can be described by

$$z_{\nu} = \sum_{\lambda, \mu} t_{\lambda, \mu} x_{\lambda} y_{\mu},$$

(2)
every computation of $\Phi$ using $R$ products gives rise to a representation of the scheme

$$t = (t_{\lambda, \mu}; \lambda = 1, \ldots, l; \mu = 1, \ldots, m; \nu = 1, \ldots, n) \in K^{lmn}$$
called the tensor corresponding to $\Phi$—as

$$t_{\lambda, \mu} = \sum_{r=1}^{R} u_{\lambda} v_{\mu} w_{rr},$$
i.e.

$$t = \sum_{r=1}^{R} u_{r} \otimes v_{r} \otimes w_{r},$$

(3)
and vice versa [13].

The ingredients $u_{r}, v_{r}, w_{r}$ ($r = 1, \ldots, R$) of such a representation constitute an algorithm of length $R$ for the computation of $\Phi$.

However, the representations (3) of $t$ are in general far from being unique. We will make this more apparent by means of the most prominent bilinear mapping in algebraic complexity theory: the multiplication of $n \times n$-matrices. Let $M_{n}(K)$ be the algebra of all $n \times n$-matrices over the field $K$ and $\Phi: (x, y) \rightarrow xy$ the usual matrix product. We will use the simple fact that $M_{n}(K)$ as a $K$-vector space is isomorphic to $K^{n^2}$ by row-wise numbering of the matrix elements.

Now, for arbitrary non-singular matrices $a, b, c \in M_{n}(K)$

$$xy = a(a^{-1}xbb^{-1}yc)c^{-1}$$
holds for all $x, y \in M_{n}(K)$. Define vector space automorphisms $A, B, C$ of $M_{n}(K)$ and, given any representation (3) of the "matrix-tensor" $t$, we conclude from (4) that for all $x, y \in M_{n}(K)$

$$xy = \sum_{r=1}^{R} (x | u_{r})(y | v_{r})w_{r}$$

$$- \sum_{r=1}^{R} (Ax | u_{r})(By | v_{r})Cw_{r}$$

$$= \sum_{r=1}^{R} (x | A^{T}u_{r})(y | B^{T}v_{r})Cw_{r}$$

(5)
Bilinear mappings

holds, where $(\cdot | \cdot)$ denotes the usual inner product for $K^{n^2}$ and $\tau: X \mapsto X^*$ the transposition of matrices. Hence we have

$$t = (A^\tau \otimes B^\tau \otimes C)t.$$  \hspace{1cm} (6)

According to (4), the mapping $A^\tau \otimes B^\tau \otimes C$ is called sandwiching.

(6) shows that the automorphism $A^\tau \otimes B^\tau \otimes C$ of $K^{n^2} \otimes K^{n^2} \otimes K^{n^2}$ operates on the set of length-$R$-algorithms for $n \times n$-matrix multiplication. Similarly, the relation

$$xy = (y^*x^*)^* \quad (x, y \in M_n(K))$$  \hspace{1cm} (7)

leads to the equivalent statement

$$t = [(\tau \otimes \tau \otimes \tau) \circ \pi_{1,2}]t$$  \hspace{1cm} (8)

where $\pi_{1,2}$ is the automorphism defined by

$$\pi_{1,2}(u \otimes v \otimes w) = v \otimes u \otimes w.$$

Special transformations of type $A \otimes B \otimes C$ were used by Hopcroft, Kerr and Musinski [9, 10] when discussing algorithms for $2 \times 2$-matrix multiplication over the Galois field $GF(2)$.

A related concept is to transform a given tensor into a tensor of the same rank via mappings that are composed of those of the form $A \otimes B \otimes C$ and of permutational mappings, i.e. those which permute the factors of $u \otimes v \otimes w$. These mappings are the core of the concepts of "duality" in [10] and of "equivalence of characteristic functions" in [5].

At present, optimal algorithms for $n \times n$-matrix multiplication are known only for the case $n = 2$. These are Strassen's algorithm [12]

$$\gamma := (u_1 \otimes v_1 \otimes w_1, \ldots; u_7 \otimes v_7 \otimes w_7)$$  \hspace{1cm} (9)

where

$$u_1 \otimes v_1 \otimes w_1 = (1, 0, 0, 1)^T \otimes (1, 0, 0, 1)^T \otimes (1, 0, 0, 1)^T,$$

$$u_2 \otimes v_2 \otimes w_2 = (-1, 0, 0, 1)^T \otimes (1, 1, 0, 0)^T \otimes (0, 0, 0, 1)^T,$$

$$u_3 \otimes v_3 \otimes w_3 = (1, 0, 0, 0)^T \otimes (0, 1, 0, -1)^T \otimes (0, 1, 0, 1)^T,$$

$$u_4 \otimes v_4 \otimes w_4 = (0, 0, 0, 1)^T \otimes (-1, 0, 1, 0)^T \otimes (1, 0, 0, 0)^T,$$

$$u_5 \otimes v_5 \otimes w_5 = (1, 0, 0, 0)^T \otimes (0, 0, 0, 1)^T \otimes (-1, 1, 0, 0)^T,$$

$$u_6 \otimes v_6 \otimes w_6 = (0, 0, 1, 1)^T \otimes (1, 0, 0, 0)^T \otimes (0, 0, 1, 1)^T,$$

$$u_7 \otimes v_7 \otimes w_7 = (0, 1, 0, -1)^T \otimes (0, 0, 1, 1)^T \otimes (1, 0, 0, 0)^T,$$

and algorithms obtained from $\gamma$ by sandwiching. The optimality of $\gamma$ was proved by Winograd [15]. The fact that no other optimal algorithms for the $2 \times 2$-problem
were found gave rise to the conjecture that no others exist. Indeed, if the underlying field $K$ is $\text{GF}(2)$, Hopcroft and Musinski could show that the conjecture is true [9]. Their methods, however, lean heavily on the fact that over $\text{GF}(2)$ there are only finitely many different $2 \times 2$-matrices.

However, our discussion shows that a thorough investigation of the set of optimal algorithms for the computation of a given bilinear mapping $\Phi$ has to enclose the study of mappings of $K^t \otimes K^m \otimes K^n$ which leave the tensor corresponding to $\Phi$ fixed.

Let us now give a short summary of the contents and organization of this work. It is divided into three parts, these parts being divided into paragraphs.

Part I: Here we will develop the general frame of the theory. In the first paragraph we recall some basic notions of algebraic complexity theory and give the projective version of the notion of optimal algorithm. This will show that the set of optimal algorithms for the computation of a bilinear mapping is closely related to a certain projective variety. In Section 2 we study linear mappings which operate on algorithm varieties. It will be shown that every bilinear mapping $\Phi$ defines in a natural way a group of automorphisms which operate on the variety of optimal algorithms for $\Phi$. This group will be called the isotropy group of $\Phi$. The discussion in this paragraph partly follows lines suggested by Strassen. In Section 3 we will determine the isotropy groups for some important classes of bilinear mappings.

Part II contains the core of our work, namely the proof that for an arbitrary ground field $K$, Strassen's algorithm $\gamma$ is "essentially unique". This means that all optimal algorithms for $2 \times 2$-matrix multiplication can be obtained from $\gamma$ by sandwiching. For the proof of this theorem we will develop in Section 1 a general method which, intuitively speaking, enables us to determine explicitly all optimal algorithms for the computation of a given bilinear mapping. Section 2 contains the proof without its purely technical details; these are collected in Section 3. Section 4 contains an application of the uniqueness theorem: we will answer the question to what extent elements of the trivial algorithm for $2 \times 2$-matrix multiplication can be used in an optimal one.

Part III: The methods developed in part II are used to investigate the variety of optimal algorithms for the computation of

$$\Phi : M_2(K) \times M_2(K) \rightarrow M_2(K) \times M_2(K),$$

$\Phi(x, y) := (xy, yx)$. We have obtained the following results:

(i) If the characteristic of $K$ is not two, then optimal algorithms for $\Phi$ have length 9 and the isotropy group of $\Phi$ does not act transitively on the variety of optimal algorithms.

(ii) If $K = \text{GF}(2)$, then every optimal algorithm for the computation of $\Phi$ has length ten.
1. Basic notions

Let $K$ be a field and $U$, $V$, and $W$ finite-dimensional vector spaces over $K$. By $U^*$, $V^*$ we denote the duals of the spaces $U$ and $V$ respectively. Consider a bilinear mapping

$$\Phi: U \times V \to W.$$ (1)

Choosing bases $(x_1, \ldots, x_l), (y_1, \ldots, y_m), (z_1, \ldots, z_n) \subseteq U$, $V$, and $W$ we have for all $\lambda \in \{1, \ldots, l\}$, $\mu \in \{1, \ldots, m\}$

$$\Phi(x_\lambda, y_\mu) = \sum_{\nu=1}^{n} t_{\lambda \mu \nu} z_\nu$$ (2)

and the three-dimensional array $t$ of the coefficients $t_{\lambda \mu \nu}$ determines $\Phi$ uniquely. Of course $t$ depends on the bases of $U$, $V$, $W$ which we have chosen.

For some of our intentions however it seems to be favorable to work coordinate-free. To do this we take advantage of the natural isomorphy $A$ between the spaces $U^* \otimes V^* \otimes W$ and $\mathcal{L}_2(U, V; W)$ of bilinear mappings $U \times V \to W$, determined by

$$A(u^* \otimes v^* \otimes w)(x, y) = (x^\top u^*)(y^\top v^*)w$$ (3)

(see [4]). Representing $t \in U^* \otimes V^* \otimes W$ as

$$t = \sum_{\rho, \sigma, \tau} t_{\rho \sigma \tau} x_{\rho}^* \otimes y_{\sigma}^* \otimes z_{\tau},$$ (4)

where $\{x_1^*, \ldots, x_l^*\}, \{y_1^*, \ldots, y_m^*\}$ are bases dual to $\{x_1, \ldots, x_l\}$ and $\{y_1, \ldots, y_m\}$ respectively, we see that

$$A(t)(x_\lambda, y_\mu) = \sum_{\nu} t_{\lambda \mu \nu} z_{\nu},$$ (5)

holds for all $\lambda, \mu$. Now every computation of $\Phi$ using $R$ products gives rise to a representation

$$t = \sum_{r=1}^{R} u_r^* \otimes v_r^* \otimes w_r,$$ (6)

of the tensor $t$ of $\Phi$ and vice versa. The minimal $R$ possible in such a representation is called the rank of $t$, denoted by $\text{rk}(t)$ (see [13]).

**Definition 1.1.** A $3R$-tuple

$$(u_1^*, v_1^*, w_1, \ldots, u_R^*, v_R^*, w_R)$$

of vectors $u_1^*, \ldots, u_R^* \in U^*$, $v_1^*, \ldots, v_R^* \in V^*$, $w_1, \ldots, w_R \in W$ satisfying $\sum_{r=1}^{R} u_r^* \otimes v_r^* \otimes w_r$, is called an algorithm of length $R$ for the computation of the bilinear mapping $\Phi$ determined by $t$. Algorithms with $R = \text{rk}(t)$ are called optimal.
Having fixed bases in $U$, $V$ and $W$ we represent elements of $U$, $V$, $W$ by column-vectors from $K^l$, $K^m$ and $K^n$ respectively. Then $\Phi$ defines a bilinear mapping

$$\Phi': K^l \times K^m \to K^n,$$

given by

$$Z_\nu = \sum_{\lambda, \mu} t_{\lambda \mu \nu} X_\lambda Y_\mu$$

The bilinear forms $(X, Y) \mapsto Z_\nu$ are determined by the matrices $\theta_\nu$,

$$(\theta_\nu)_\lambda \mu = t_{\lambda \mu \nu} \quad (\nu = 1, \ldots, n),$$

and any representation (6) is equivalent to a representation of $\Theta$:

$$(t_{\lambda \mu \nu}: \lambda = 1, \ldots, l; \mu = 1, \ldots, m; \nu = 1, \ldots, n)$$

of the form

$$\Theta = \sum_{r=1}^{R} \xi_r \otimes \eta_r \otimes \xi_r.$$  

Notice that the set

$$\left\{ (\xi_1, \eta_1, \xi_2, \ldots, \xi_R, \eta_R, \xi_R) \in K^{(l+m+n)}: \Theta = \sum_{r=1}^{R} \xi_r \otimes \eta_r \otimes \xi_r \right\}$$

is an algebraic variety in $K^{(l+m+n)}$ with the defining equations

$$t_{\lambda \mu \nu} = \sum_{r=1}^{R} \xi_{r \lambda} \eta_{r \mu} \xi_{r \nu}$$

$$(\lambda = 1, \ldots, l; \mu = 1, \ldots, m; \nu = 1, \ldots, n).$$

The representation (9) of $\Theta$ is equivalent to the representations

$$\theta_\nu = \sum_{r=1}^{R} (\xi_r \otimes \eta_r) \xi_r \quad (\nu = 1, \ldots, n)$$

of the layers of $\Theta$, and (10) means that the subspace $\operatorname{lin}\{\theta_1, \ldots, \theta_n\}$ of $K^{lm}$ generated by $\theta_1, \ldots, \theta_n$ is contained in $\operatorname{lin}\{\xi_1 \otimes \eta_1, \ldots, \xi_R \otimes \eta_R\}$.

If $R = \operatorname{rk}(t)$, i.e. if $R$ is minimal in (5), then the products $\xi_r \otimes \eta_r$ ($r = 1, \ldots, R$) determine the vectors $\xi_r$ uniquely, hence we may describe optimal algorithms by the tuples $(\xi_1 \otimes \eta_1, \ldots, \xi_R \otimes \eta_R)$ of products $\xi_r \otimes \eta_r$.

In view of the multilinearity of the tensor product we have

$$\alpha \xi \otimes \beta \eta \otimes \gamma \xi = \xi \otimes \eta \otimes \xi$$

where $\alpha, \beta, \gamma$ are non-zero elements of $K$ such that $\alpha \beta \gamma = 1$.

In other words, we may scale two of the factors in $\xi \otimes \eta \otimes \xi$ against the third.
This simple observation leads to the projective viewpoint of algorithms. Scaling as explained above defines an equivalence relation on the set of all such representations. Consider the set of equivalence classes. The equivalence $\alpha \xi \otimes \beta \eta \sim \xi \otimes \eta$ if $\alpha \beta \neq 0$ allows us to look at the factors $\xi$ and $\eta$ as elements of the projective spaces $\mathbb{P}^{l-1}$ and $\mathbb{P}^{m-1}$ respectively. Write $\tilde{\xi} \otimes \tilde{\eta}$ for the equivalence class of $\xi \otimes \eta \cdot \xi \otimes \eta$ lies in the product $\mathcal{P}_{l,m}$ of the projective spaces $\mathbb{P}^{l-1}$ and $\mathbb{P}^{m-1}$. $\mathcal{P}_{l,m}$ is a well-known algebraic subvariety of the projective space $\mathbb{P}^{lm-1}$. It belongs to a class of algebraic varieties usually called Segre varieties [6]. Observe now that every optimal algorithm $(\xi_1, \eta_1, \xi_2, \ldots, \xi_R, \eta_R, \xi_R)$ for $\Phi$ is determined up to scaling-equivalence by the $R$-tuple $(\tilde{\xi}_1 \otimes \tilde{\eta}_1, \ldots, \tilde{\xi}_R \otimes \tilde{\eta}_R)$ contained in $\mathcal{P}_{l,m}$. The scaling-equivalence classes $\tilde{\theta}_i$ of the layers $\theta_i$ of $\theta$ are elements of $\mathbb{P}^{lm-1}$. This leads to the following

**Proposition 1.2.** There is a bijective map from the set of scaling-equivalence classes of optimal algorithms for the computation of the bilinear mapping $\Phi$ onto the set of those $R$-tuples

$$(P_1, \ldots, P_R) \in \mathcal{P}_{l,m}^R \quad (R := \text{rk}(t))$$

which have the property that the projective subspace $[\tilde{\theta}_1, \ldots, \tilde{\theta}_n \in \mathbb{P}^{lm-1}$ generated by $\tilde{\theta}_1, \ldots, \tilde{\theta}_n$ is contained in the projective subspace $[P_1, \ldots, P_R] \subseteq \mathbb{P}^{lm-1}$ generated by $P_1, \ldots, P_R$.

Notice that the condition

$$[\tilde{\theta}_1, \ldots, \tilde{\theta}_n] \subseteq [P_1, \ldots, P_R]$$

can be expressed by a system of homogeneous equations. Hence the set of tuples $(P_1, \ldots, P_R)$ with the above property is a closed algebraic subvariety of $\mathcal{P}_{l,m}^R$.

We end this paragraph with a remark manifesting the geometric nature of the rank of the tensor $t$. Consider the projective space $P_\theta := [\tilde{\theta}_1, \ldots, \tilde{\theta}_n]$ generated by the projective layers $\tilde{\theta}_1, \ldots, \tilde{\theta}_n$ of $\theta$. Let $P_{\theta}$ be the set of all projective subspaces $E_d \subseteq \mathbb{P}^{lm-1}$ of dimension $d$ which

(i) are generated by products, i.e. by their intersection with $\mathcal{P}_{l,m}$; and

(ii) contain the space $P_\theta$.

Then

$$\text{rk}(t) = \min\{d : E_d \subseteq P_{\theta}\} + 1.$$  

It should be mentioned that almost nothing is known about the intersection of $\mathcal{P}_{l,m}$ with a projective subspace $E$ of $\mathbb{P}^{lm-1}$, especially if the codimension of $E$ is greater than $2(l + m) - 1$.

This seems to be one of the major reasons why it is still a difficult task to determine the rank of a given tensor.
It is not difficult to solve this problem completely in the simple case \( l = m = n = 2 \), for in this case we have only to discuss the intersections of projective lines with the Segre-variety \( \mathcal{P}_{2,2} \). This situation however is quite untypical for the general case, so we skip further details.

2. The isotropy group of a bilinear mapping

Let \( U, V, W, \) and \( \Phi \) be as in Section 1 and let \( t \in U^* \otimes V^* \otimes W \) be the tensor corresponding to \( \Phi \). As was pointed out in Section 1, there is a natural bijection from the set of scaling equivalence classes of algorithms of length \( R \) onto the set

\[
\left\{ (u_1 \otimes v_1 \otimes w_1, \ldots, u_R \otimes v_R \otimes w_R): \sum_{r=1}^R u_r \otimes v_r \otimes w_r = t \right\}
\]

of ordered length-\( R \)-decompositions of the tensor \( t \) into tensors of rank one.

From now on, “algorithm” will always be used synonymously to “scaling equivalence class of optimal algorithms for the computation of \( \Phi \)” \( \mathcal{R}_k(U, V, W) \) denotes the set of tensors from \( U \otimes V \otimes W \) whose rank is less or equal to \( k \) (the zero-tensor has rank zero by definition).

Our aim is to exhibit and describe a canonical class of mappings from \( U^* \otimes V^* \otimes W \) into itself which operate on the variety of algorithms.

As we are concerned with additive decompositions of tensors, it is natural to require such mappings to be endomorphisms of the vector space \( U^* \otimes V^* \otimes W \). An endomorphism \( \varphi \) of \( U^* \otimes V^* \otimes W \) operating on the algorithm variety has the following fundamental properties:

(i) \( \varphi \) leaves \( t \) fixed,

(ii) \( \varphi \) maps each tensor of rank one occurring in an algorithm to a tensor of the same kind.

As it was exemplified in the introduction, “sandwiching” of matrices is such a mapping \( \varphi \). In what follows, we will generalize this example.

Let \( U_1, U_2, U_3 \) be finite-dimensional \( K \)-vector spaces \( \neq \{0\} \). As is well-known, every permutation \( \pi \) of \( \{1, 2, 3\} \) induces an isomorphism

\[
U_1 \otimes U_2 \otimes U_3 \rightarrow U_{\pi(1)} \otimes U_{\pi(2)} \otimes U_{\pi(3)},
\]

also denoted by \( \pi \) and defined by

\[
\pi(u_1 \otimes u_2 \otimes u_3) = u_{\pi(1)} \otimes u_{\pi(2)} \otimes u_{\pi(3)}.
\]

We will call \( \pi \) a permutational mapping. Using this notation, the mappings considered in the introduction (cf. (6) and (8)) are of the form

\[
\varphi = \pi^{-1} \circ (A_1 \otimes A_2 \otimes A_3),
\]
where the $A_j: U_j \to U_{\pi(j)}$, ($j \in \{1, 2, 3\}$), are isomorphisms. It is sufficient however to assume only that the $A_j's$ are linear mappings, for we will show in the sequel that, if $t \in U_1 \otimes U_2 \otimes U_3$ is a non-zero tensor and $\varphi$ a linear mapping of the form (1) that keeps $t$ fixed, then there exist isomorphisms $\tilde{A}_j$ such that $A_1 \otimes A_2 \otimes A_3$ and $\tilde{A}_1 \otimes \tilde{A}_2 \otimes \tilde{A}_3$ agree on all elements $u_1 \otimes u_2 \otimes u_3 \in U_1 \otimes U_2 \otimes U_3$ that occur in an optimal decomposition of $t$ into tensors of rank one.

Let $\Phi: U_1^* \times U_2^* \to U_3$ be the bilinear mapping determined by $t$. Since $\Phi$ is bilinear, it possesses a left and a right kernel, namely

$$LK_{\Phi} := \{x \in U_1^*: \Phi(x, y) = 0 \text{ for all } y \in U_2^*\}$$

and

$$RK_{\Phi} := \{y \in U_2^*: \Phi(x, y) = 0 \text{ for all } x \in U_1^*\}.$$  

Then $t$ is an element of the subspace $LK_{\Phi} \otimes RK_{\Phi} \otimes U_3$, where $LK_{\Phi}$, $RK_{\Phi}$ are the orthogonal spaces of $LK_{\Phi}$, $RK_{\Phi}$ respectively. The following result is basic for our discussion:

**Lemma 2.1.** Let $t = \sum_{r=1}^R u_r \otimes v_r \otimes w_r$ be an optimal decomposition of $t \in U_1 \otimes U_2 \otimes U_3$. Then $u_r \in LK_{\Phi}$ and $v_r \in RK_{\Phi}$ holds for all $r \in \{1, \ldots, R\}$.

**Proof.** Let $\rho: U_1 \to U_1 LK_{\Phi}$ be the canonical projection and $\tilde{\rho} := \rho \otimes \text{id}_{V_1} \otimes \text{id}_{U_1}$. Applying $\tilde{\rho}$ to $t$, we obtain

$$0 = \tilde{\rho}(t) = \sum_{r=1}^R \rho(u_r) \otimes v_r \otimes w_r,$$

whence $p(u_r) = 0$ for all $r$ or $\{v_1 \otimes w_1, \ldots, v_R \otimes w_R\}$ is linearly dependent, a contradiction to the optimality of the decomposition.

Hence $u_r \in LK_{\Phi}$ for all $r \in \{1, \ldots, R\}$. A similar reasoning shows $v_r \in RK_{\Phi}$ for all $r$.

**Proposition 2.2.** Let $\pi: U_1 \otimes U_2 \otimes U_3 \to U_{\pi(1)} \otimes U_{\pi(2)} \otimes U_{\pi(3)}$ be a permutational mapping, $A_j: U_j \to U_{\pi(j)}$ ($j \in \{1, 2, 3\}$) linear mappings, $t \in U_1 \otimes U_2 \otimes U_3$, $\Phi: U_1^* \times U_2^* \to U_3$ the bilinear mapping corresponding to $t$, and $\text{lin im} \Phi := \text{lin}\{\Phi(x, y): x \in U_1^*, y \in U_2^*\}$ the linear span of the vectors $\Phi(x, y)$. If $\varphi := \pi^{-1}(A_1 \otimes A_2 \otimes A_3)$ leaves $t$ fixed, then the restrictions

$$A_1_{\text{lin im} \Phi}, A_2_{\text{lin im} \Phi} \quad \text{and} \quad A_3_{\text{lin im} \Phi}$$

are monomorphisms.

**Proof.** $\varphi t = t$ is equivalent to

$$\pi t = (A_1 \otimes A_2 \otimes A_3)t.$$  

(2)
Let $\Phi_\pi: U^*_\pi(1) \times U^*_\pi(2) \to U^*_\pi(3)$ be the bilinear mapping corresponding to the tensor $\pi t \in U^*_\pi(1) \otimes U^*_\pi(2) \otimes U^*_\pi(3)$.

Moreover let $R := r_k(t)$ and

$$t = \sum_{r=1}^R u_r^{(1)} \otimes u_r^{(2)} \otimes u_r^{(3)} \tag{3}$$

be an optimal decomposition of $t$.

Then for any $x \in U^*_\pi(1)$, $y \in U^*_\pi(2)$ we get from (2) and (3):

$$\Phi_\pi(x, y) = \sum_{r=1}^R (A_1 u_r^{(1)} | x)(A_2 u_r^{(2)} | y) A_3 u_r^{(3)}$$

$$= A_3 \sum_{r=1}^R (u_r^{(1)} | A_1^* x)(u_r^{(2)} | A_2^* y) u_r^{(3)},$$

where $(\cdot | \cdot)$ are non-degenerate scalar products. Hence

$$\Phi_\pi(x, y) = A_3(\Phi(A_1^* x, A_2^* y)) \tag{4}$$

which immediately shows that

$$\dim(\ker \Phi) \leq \dim(\ker \Phi_\pi) \tag{5}.$$

Moreover, since $\pi'' t = (A_{\pi''}(1) \otimes A_{\pi''}(2) \otimes A_{\pi''}(3))(\pi'' t)$ and $\pi'' = \text{id}$ for $n = 1, 2$ or 3, the same reasoning as above shows

$$\dim(\ker \Phi) \leq \dim(\ker \Phi_\pi),$$

hence

$$\dim(\ker \Phi_\pi) = \dim(\ker \Phi) \tag{6}.$$

As $\ker \Phi_\pi$ is generated by $\{u_1^{(3)}, \ldots, u_R^{(3)}\}$ and $\ker \Phi$ by $\{u_1^{(3)} , \ldots, u_R^{(3)}\}$, we obtain that $A_3$ maps $\ker \Phi$ isomorphically onto $\ker \Phi_\pi$. (3) is an optimal decomposition of $t$, hence $u_r^{(1)} \in \ker \Phi, u_r^{(2)} \in \ker \Phi, A_1 u_r^{(1)} \in \ker \Phi, A_2 u_r^{(2)} \in \ker \Phi, (r \in \{1, \ldots, R\})$, and it is easy to see that the equalities

$$\ker \Phi_\pi = \ker \Phi, \quad \ker \Phi_\pi = \ker \Phi$$

hold. Now consider the bilinear mappings

$$\Phi^{(1)} : U^*_2 \times U^*_2 \to U_1,$$

$$\Phi^{(2)} : U^*_1 \times U^*_2 \to U_2$$

and their corresponding mappings $\Phi^{(1)}_\pi, \Phi^{(2)}_\pi$ ($\Phi^{(1)}$ and $\Phi^{(2)}$ are defined by $t$ in the obvious way). Like for $\Phi$ and $\Phi_\pi$ we can show that for $j \in \{1, 2\}$

$$\dim(\ker \Phi^{(j)} \pi) = \dim(\ker \Phi^{(j)}).$$
Bilinear mappings

holds. But

\[ \text{lin im } \Phi^{(1)} = L K_{\Phi}, \quad \text{lin im } \Phi^{(2)} = L K_{\Phi}, \]

and

\[ \text{lin im } \Phi^{(2)} = R K_{\Phi}, \quad \text{lin im } \Phi^{(2)} = R K_{\Phi}, \]

whence our assertion.

Proposition 2.2 shows that when discussing endomorphisms of \( U_1 \otimes U_2 \otimes U_3 \) that are of the form \( \pi^{-1} \circ (A_1 \otimes A_2 \otimes A_3) \) and leave a given tensor \( t \) fixed, we may confine ourselves to automorphisms. A decisive property of automorphisms of the form \( \pi^{-1} \circ (A_1 \otimes A_2 \otimes A_3) \) is that they map the set \( \mathcal{R}^0(U_1, U_2, U_3) := \mathcal{R}(U_1, U_2, U_3) \setminus \{0\} \) into itself.

Mappings that operate on a variety are an important tool for the investigation of the variety. Thus it is highly desirable to exhibit a rather general—and feasible—class \( \mathcal{C} \) of mappings \( U_1 \otimes U_2 \otimes U_3 \rightarrow U_1 \otimes U_2 \otimes U_3 \) which operate on the algorithm variety of a tensor \( t \)—provided they keep \( t \) fixed.

Since the definition of \( \mathcal{C} \) should not depend on a special tensor \( t \), the following requirements for a \( \varphi \in \mathcal{C} \) seem to be reasonable:

(i) \( \varphi \) is an endomorphism of \( U_1 \otimes U_2 \otimes U_3 \),

(ii) \( \varphi(\mathcal{R}^0(U_1, U_2, U_3)) \subseteq \mathcal{R}^0(U_1, U_2, U_3) \).

In the following we shall more generally characterize linear mappings \( \varphi \) from \( U_1 \otimes U_2 \otimes U_3 \) to \( V_1 \otimes V_2 \otimes V_3 \) (\( U_i, V_i \) finite-dimensional \( K \)-vector spaces \( \neq \{0\} \)) with the property

\[ \varphi(\mathcal{R}^0(U_1, U_2, U_3)) \subseteq \mathcal{R}^0(V_1, V_2, V_3). \]

Such mappings are called Segre-homomorphisms. The following lemma stated without (the very easy) proof is of technical importance:

**Lemma 2.3.** Let \( \theta := x_1 \otimes x_2 \otimes x_3, \theta' := y_1 \otimes y_2 \otimes y_3 \in \mathcal{R}^0(U_1, U_2, U_3) \). Then \( \theta + \theta' \in \mathcal{R}(U_1, U_2, U_3) \) iff \( [x_j] \neq [y_j] \) for at most one \( j \in \{1, 2, 3\} \).

Our discussion of Segre-homomorphisms \( \varphi \) is based on the investigation of the behavior of \( \varphi \) on subsets of \( \mathcal{R} \) that are linear subspaces of \( U_1 \otimes U_2 \otimes U_3 \). These are characterized by the following lemma whose projective version is well-known [6].

**Lemma 2.4.** \( X_1 \otimes u_2 \otimes u_3, u_1 \otimes X_2 \otimes u_3, u_1 \otimes u_2 \otimes X_3, \), where the \( X_i \)'s are non-zero subspaces and the \( u_i \)'s are non-zero elements of \( U_i \), \( i \in \{1, 2, 3\} \), are the only subsets \( \mathcal{R}(U_1, U_2, U_3) \) that are non-zero linear subspaces of \( U_1 \otimes U_2 \otimes U_3 \).

---

1 If \( X \) is a vector space and \( x \in X \), then \( [x] \) denotes the subspace of \( X \) generated by \( x \).
Proof. Clearly each of $X_1 \otimes U_2 \otimes U_3$, $u_1 \otimes X_2 \otimes U_3$, $u_1 \otimes u_2 \otimes X_3$ is a linear subspace of $U_1 \otimes U_2 \otimes U_3$. For the proof of the converse let $E \neq \{0\}$ be a linear subspace of $U_1 \otimes U_2 \otimes U_3$ contained in $R_1$. Fix $v_1 \otimes v_2 \otimes v_3 \in E \setminus \{0\}$ and let $F_i := \{u_1 \otimes u_2 \otimes u_3 \in E : [u_i] = [v_i] \text{ for } i \neq j, i \in \{1, 2, 3\}\}$. Lemma 2.3 yields

$E = F_1 \cup F_2 \cup F_3$.

Since the $F_i$'s are linear subspaces of $E$, we conclude that there is a $j \in \{1, 2, 3\}$ such that $E = F_j$ and our assertion follows.

Thus the set $L = L(U_1, U_2, U_3)$ of non-zero subspaces of $U_1 \otimes U_2 \otimes U_3$ that are contained in $R_1$ splits into three classes $L_1, L_2, L_3$, where $L_i$ is the set of subspaces $X_1 \otimes X_2 \otimes X_3$ of $U_1 \otimes U_2 \otimes U_3$ with $\dim X_i = 1$ for $i \neq j$. Note that the elements of $L_i \cap L_j (i \neq j)$ are the lines in $U_1 \otimes U_2 \otimes U_3$ that are spanned by rank-one-tensors.

Let $L^{(2)} = \{E \in L^k : \dim E = 2\}$, and

$L^{(3)} := \bigcup_{k=1}^{3} L^{(2)}_k$.

Since $L_i^{(2)} \cap L_j^{(2)} = \emptyset$ for $i \neq j$, $\sigma(E) = k$ iff $E \in L^{(2)}_k$ defines a mapping

$\sigma : L^{(2)} \rightarrow \{1, 2, 3\}$

$\sigma(E)$ is called the positional number of $E \in L^{(2)}$.

Definition 2.5. A Segre-homomorphism

$\varphi : U_1 \otimes U_2 \otimes U_3 \rightarrow V_1 \otimes V_2 \otimes V_3$

is called faithful iff

$\sigma(E_1) \neq \sigma(E_2) \Rightarrow \sigma(\varphi(E_1)) \neq \sigma(\varphi(E_2))$

for all $E_1, E_2 \in L^{(2)}(U_1, U_2, U_3)$. (Observe that $\dim \varphi(E) = \dim E$ for all $E \in L(U_1, U_2, U_3)$.)

Proposition 2.6. For any Segre-homomorphism $\varphi : U_1 \otimes U_2 \otimes U_3 \rightarrow V_1 \otimes V_2 \otimes V_3$ we have

(i) $\varphi$ is semi-faithful, i.e. $E_1, E_2 \in L^{(2)}(U_1, U_2, U_3), \sigma(E_1) = \sigma(E_2)$ implies $\sigma(\varphi(E_1)) = \sigma(\varphi(E_2))$.

(ii) If $\varphi$ is not faithful, then there are $v_i \in V_i$ ($i \in \{1, 2, 3\}$) such that $\varphi(U_1 \otimes U_2 \otimes U_3)$ is contained in one of the spaces $V_1 \otimes V_2 \otimes v_3, V_1 \otimes v_2 \otimes V_3$ or $v_1 \otimes V_2 \otimes V_3$.

Proof. (i) It suffices to prove that $E_1, E_2 \in L^{(2)}_1$ and $\varphi(E_1) \in L^{(2)}_1$ implies $\varphi(E_2) \in L^{(2)}_1$. For the other possible cases can be reduced to this by means of suitable
permutational mappings. Moreover, as \( \varphi \) is injective on linear subspaces contained in \( \mathcal{R} \), we may assume that \( \dim E_1 = \dim E_2 = \dim U_1 \). If for non-zero \( y, z \)

\[ \tilde{\sigma}(y, z) := \sigma(\varphi(U_1 \otimes y \otimes z)) \]

is constant in \( z \) when \( y \) is fixed and constant in \( y \) when \( z \) is fixed, then \( \tilde{\sigma} \) is a constant function. Now assume that the positional number of \( \varphi(E_2) \) is different from 1, i.e., \( \tilde{\sigma} \) is not constant, and assume for instance that there are non-zero \( z_0 \in U_3 \) and linearly independent \( y_0, y_1 \in U_2 \) such that

\[ \sigma(\varphi(U_1 \otimes y_0 \otimes z_0)) = 1, \]
\[ \sigma(\varphi(U_1 \otimes y_1 \otimes z_0)) = 2, \]

i.e. with suitable non-zero vectors \( u_1, v_0, w_0, w_1 \):

\[ \varphi(U_1 \otimes y_0 \otimes z_0) \subseteq V_1 \otimes v_0 \otimes w_0 \]

and

\[ \varphi(U_1 \otimes y_1 \otimes z_0) \subseteq u_1 \otimes V_2 \otimes w_1. \]

Then we obtain for all \( x \in U_1 \)

\[ \varphi(x \otimes (y_0 + y_1) \otimes z_0) = f(x) \otimes v_0 \otimes w_0 + u_1 \otimes g(x) \otimes w_1, \]

where \( f: U_1 \to V_1 \) and \( g: U_1 \to V_2 \) are suitable linear mappings. According to Lemma 2.3, this implies \( f(x) \in [u_1] \) or \( g(x) \in [v_0] \) for all \( x \in U_1 \), i.e.

\[ U_1 = \{x: f(x) \in [u_1]\} \cup \{x: g(x) \in [v_0]\}. \]

Therefore \( f(U_1) \subseteq [u_1] \) or \( g(U_1) \subseteq [v_0] \), both contradicting the injectivity of \( \varphi \) on linear subspaces of \( \mathcal{R}(U_1, U_2, U_3) \).

(ii) As \( \varphi \) is semi-faithful, \( \sigma(\varphi(E_0)) = k \) for a single \( E_0 \in E_1 \) implies \( \sigma(\varphi(E)) = k \) for all \( E \in E_1 \). Thus, if \( \varphi \) is not faithful, an appropriate use of permutational mappings shows that we have to discuss the following case only: For all non-zero vectors \( y_0, z_0 \) there are non-zero vectors \( u_0, v_0 \) such that

\[ \varphi(U_1 \otimes y_0 \otimes z_0) \subseteq u_0 \otimes v_0 \otimes V_3, \]  
(7)

\[ \varphi(x_0 \otimes U_2 \otimes z_0) \subseteq u_0 \otimes v_0 \otimes V_3, \]  
(8)

and \( \sigma(\varphi(x_0 \otimes y_0 \otimes U_3)) \in \{2, 3\} \).

Therefore, using (7) and the fact that \( \varphi \) is semi-faithful, for all \( x \in U_1 \setminus \{0\} \) and all \( y \in U_2 \setminus \{0\} \) there are non-zero vectors \( u_y, v_y, \) and \( w_{xy} \) such that

\[ \varphi(x \otimes y \otimes z_0) = u_y \otimes v_y \otimes w_{xy}, \]

and similarly by (8) there are non-zero vectors \( u_x, v_x, \) and \( w_{xy} \) such that

\[ \varphi(x \otimes y \otimes z_0) = u_x \otimes v_x \otimes w_{xy}. \]
Now fix \( x_0 \in U_1 \backslash \{0\} \). Then \( u_y \in [u_{x_0}] \), \( v_y \in [v_{x_0}] \) for all \( y \in U_2 \backslash \{0\} \). Fixing \( y_0 \in U_2 \backslash \{0\} \), we obtain \( u_x \in [u_{y_0}] = [u_{x_0}] \) and \( v_x \in [v_{y_0}] = [v_{x_0}] \) for all \( x \in U_1 \backslash \{0\} \). Therefore

\[
\phi(U_2 \otimes U_3 \otimes z_0) \subseteq u_0 \otimes v_0 \otimes V_3.
\]

Thus for all \( z \in U_3 \backslash \{0\} \) there are non-zero vectors \( u, v \) such that

\[
\phi(U_1 \otimes U_2 \otimes z) \subseteq u \otimes v \otimes V_3.
\]  

(9)

Assume that \( \sigma(\phi(x_0 \otimes y_0 \otimes U_3)) = 2 \), i.e. that there exist \( u_1 \in V_1 \backslash \{0\} \), \( w_1 \in V_3 \backslash \{0\} \) such that

\[
\phi(x_0 \otimes y_0 \otimes U_3) \subseteq u_1 \otimes V_2 \otimes w_1.
\]

(10)

Then for all \( z \in U_3 \backslash \{0\} \) there exists a \( v'_z \in V_2 \backslash \{0\} \) such that

\[
\phi(x_0 \otimes y_0 \otimes z) = u_1 \otimes v'_z \otimes w_1,
\]

and by (9) this is an element of \( u \otimes v \otimes V_3 \), hence \( u \in [u_1] \) for all \( u \)'s occurring in (9). Thus

\[
\phi(U_1 \otimes U_2 \otimes U_3) \subseteq u_1 \otimes V_2 \otimes V_3.
\]

If \( \sigma(\phi(x_0 \otimes y_0 \otimes U_3)) = 3 \), then a similar argument shows that there are \( u \in V_1 \backslash \{0\} \) and \( v \in V_2 \backslash \{0\} \) such that

\[
\phi(U_1 \otimes U_2 \otimes U_3) \subseteq u \otimes v \otimes V_3.
\]

The following example of a non-faithful Segre-endomorphism shows that the assertion of the theorem cannot be sharpened:

Let \( \mathcal{A} \) be a finite-dimensional division \( K \)-algebra with unit 1. Then

\[
\phi: x \otimes y \otimes z \mapsto 1 \otimes z \otimes xy
\]

determines a Segre-endomorphism of \( \mathcal{A} \otimes \mathcal{A} \otimes \mathcal{A} \) that is not faithful.

Note, however, that this example does not work if \( K \) is algebraically closed for in this case there is no division algebra over \( K \) (except \( K \) itself). This phenomenon is not merely accidental. In fact one can show that, if \( K \) is algebraically closed, every Segre-homomorphism is faithful [14].

However, Proposition 2.6 shows that Segre-endomorphisms that are not faithful can leave fixed only those tensors which essentially are matrices. Since such tensors are of minor interest in algebraic complexity theory we may restrict ourselves to faithful Segre-homomorphisms. The following theorem is the affine version of a result of projective geometry [6]:

**Theorem 2.7.** Let \( \phi: U_1 \otimes U_2 \otimes U_3 \rightarrow V_1 \otimes V_2 \otimes V_3 \) be a faithful Segre-homomorphism. Then there are a permutational mapping \( \pi \) and monomorphisms \( A_i: U_i \rightarrow V_{\pi(i)} \) (\( i \in \{1, 2, 3\} \)) such that \( \phi = \pi^{-1} \circ (A_1 \otimes A_2 \otimes A_3) \).
Proof. There is a unique permutational mapping $\pi$ such that $\varphi := \pi \circ \psi$ satisfies
\[ \sigma(\psi(E)) = \sigma(E) \] (11)
for all $E \in L^{(2)}(U_1, U_2, U_3)$. Now fix non-zero vectors $x_0 \in U_1$, $y_0 \in U_2$, $z_0 \in U_3$ and let
\[ u_0 \otimes v_0 \otimes w_0 = \psi(x_0 \otimes y_0 \otimes z_0). \]
(11) yields that by
\[
\begin{align*}
\psi(x \otimes y_0 \otimes z_0) &= A_1 x \otimes v_0 \otimes w_0 \quad (x \in U_1), \\
\psi(x_0 \otimes y \otimes z_0) &= u_0 \otimes A_2 y \otimes w_0 \quad (y \in U_2), \\
\psi(x_0 \otimes y_0 \otimes z) &= u_0 \otimes v_0 \otimes A_3 z \quad (z \in U_3).
\end{align*}
\] (12)
monomorphisms $A_i: U_i \to V_{n(i)}$ are defined. We have to show that
\[ \psi(x \otimes y \otimes z) = A_1 x \otimes A_2 y \otimes A_3 z \] (13)
holds for all $x \in U_1$, $y \in U_2$, $z \in U_3$.
For this, we first prove
\[ \psi(x_0 \otimes y \otimes z) = u_0 \otimes A_2 y \otimes A_3 z \] (14)
for all $y \in U_2$, $z \in U_3$. (14) is obvious if $\{y, y_0\}$ or $\{z, z_0\}$ are linearly dependent. Otherwise we conclude in the following way: We have
\[
\begin{align*}
\psi(x_0 \otimes y \otimes z) &= u_0 \otimes v_0 \otimes A_3 z \quad \text{by (12),} \\
\psi(x_0 \otimes y \otimes z) &= u_0 \otimes v_0 \otimes A_3 z \quad \text{by (12) and (11),} \\
\psi(x_0 \otimes y \otimes z) &= u_0 \otimes v_0 \otimes A_3 z \quad \text{by (12) and (11).}
\end{align*}
\]
Hence the $\lambda$ is a $\lambda \in K \setminus \{0\}$ such that
\[ v \otimes A_3 z = A_2 y \otimes w = \lambda (A_2 y \otimes A_3 z). \]
Now observe that $\{v_0, A_2 y\}$ and $\{w_0, A_3 z\}$ are linearly independent; hence
\[ \psi(x_0 \otimes (y_0 + y) \otimes (z_0 + z)) = u_0 \otimes (v_0 \otimes w_0 + v_0 \otimes A_3 z + A_2 y \otimes w_0 + \lambda A_2 y \otimes A_3 z) \]
has rank one if and only if $\det \begin{pmatrix} 11 \\ 1\lambda \end{pmatrix} = 0$, i.e., $\lambda = 1$, whence (14) holds.
Similarly one can prove
\[ \psi(x \otimes y_0 \otimes z) = A_1 x \otimes v_0 \otimes A_3 z \] (15)
for all $x \in U_1$, $z \in U_3$ and finally—fixing $A_1 z$—the assertion (13).
Corollary 2.8. Any faithful Segre-endomorphism of $U_1 \otimes U_2 \otimes U_3$ is an automorphism of the vector space $U_1 \otimes U_2 \otimes U_3$. The set $\Gamma(U_1, U_2, U_3)$ of faithful Segre-endomorphisms of $U_1 \otimes U_2 \otimes U_3$ forms a group and the set $\Gamma^0(U_1, U_2, U_3)$ of Segre-endomorphisms of the form $A_1 \otimes A_2 \otimes A_3$, where $A_i$ is a vector space automorphism of $U_i (i \in \{1, 2, 3\})$, is a normal subgroup of $\Gamma(U_1, U_2, U_3)$. Moreover,

$$\#(\Gamma(U_1, U_2, U_3)/\Gamma^0(U_1, U_2, U_3)) = k!,$$

where $k = 4 - \#\{\dim U_i : i \in \{1, 2, 3\}\}$. ($\#M$ denotes the cardinality of the set $M$.)

Proof. If $\pi \circ (A_1 \otimes A_2 \otimes A_3)$ is a Segre-endomorphism of $U_1 \otimes U_2 \otimes U_3$ then

$$A_1 \otimes A_2 \otimes A_3 : U_1 \otimes U_2 \otimes U_3 \rightarrow U_{\pi(1)} \otimes U_{\pi(2)} \otimes U_{\pi(3)}$$

is a Segre-isomorphism, hence

$$\dim U_i = \dim U_{\pi(i)} \quad (i \in \{1, 2, 3\}).$$

Furthermore

$$\pi \circ (A_1 \otimes A_2 \otimes A_3) = (A_{\pi(1)} \otimes A_{\pi(2)} \otimes A_{\pi(3)}) \circ \pi,$$

whence our assertion.

Definition 2.9. Let $\Phi : U \times V \rightarrow W$ be a bilinear mapping, $t \in U^* \otimes V^* \otimes W$ the tensor corresponding to $\Phi$.

$$\Gamma_\Phi := \{ \varphi \in \Gamma(U^*, V^*, W) : \varphi(t) = t \}$$

is called the isotropy group of $\Phi$.

$\Gamma_\Phi$ is a group of automorphisms operating on the variety $\mathcal{A}_{\Phi}$ of algorithms for the computation of $\Phi$. Let

$$t = \sum_{r=1}^{R} u_r^* \otimes v_r^* \otimes w_r,$$

where $R = \text{rk}(t)$, and let $\sigma$ be a permutation of $\{1, \ldots, R\}$. Then we also have

$$t = \sum_{r=1}^{R} u_{\sigma(r)}^* \otimes v_{\sigma(r)}^* \otimes w_{\sigma(r)},$$

hence the group $\Sigma_R$ of permutations of $R$ elements acts on $\mathcal{A}_{\Phi}$. If $\sigma \in \Sigma_R$ and $\varphi \in \Gamma_\Phi$ then the operations of $\sigma$ and $\varphi$ commute on $\mathcal{A}_{\Phi}$, hence

$$G_\Phi := \Gamma_\Phi \cdot \Sigma_R$$

is a group acting on $\mathcal{A}_{\Phi}$. $G_\Phi$ is called the extended isotropy group of $\Phi$.

Definition 2.10. Two (scaling-equivalence classes of) algorithms for $\Phi$ are called equivalent if they belong to the same $G_\Phi$-orbit.
3. The isotropy groups of some important bilinear mappings

In this paragraph we will determine the isotropy groups of an important class of bilinear mappings. We are mainly concerned with bilinear mappings which are defined by the multiplication of a finite-dimensional associative algebra \( \mathcal{A} \) over the field \( K \). We will always assume that \( \mathcal{A} \) has a unit element, usually denoted by 1. L and R denote the left and right regular representations of \( \mathcal{A} \) in the algebra of vector space endomorphisms of \( \mathcal{A} \):

\[
L_a x := ax, \quad R_a x = xa \quad (x \in \mathcal{A}).
\]

Let \( t \) be the tensor of the bilinear mapping \( \Phi : (x, y) \mapsto xy \) defined by the multiplication of \( \mathcal{A} \). Choosing a basis for the vector space \( \mathcal{A} \), the tensor \( \theta \in K^{N^2} \) (\( N := \dim \mathcal{A} \)) corresponding to \( t \) is known as a "scheme of structural constants of \( \mathcal{A} \)." Let \( \Gamma_{\mathcal{A}} \) be the isotropy group of \( \Phi \) (in a more suggestive manner we will call \( \Gamma_{\mathcal{A}} \) the isotropy group of the algebra \( \mathcal{A} \)), \( \Gamma_{\mathcal{A}}^0 \) its normal subgroup consisting of mappings of the form: \( A \otimes B \otimes C \). The following theorem characterizes the group \( \Gamma_{\mathcal{A}}^0 \):

**Theorem 3.1.** Let \( \mathcal{A} \) be a finite-dimensional algebra over the field \( K \) and \( A^* \otimes B^* \otimes C \) an automorphism of the vector space \( \mathcal{A}^* \otimes \mathcal{A}^* \otimes \mathcal{A} \). Then \( A^* \otimes B^* \otimes C \in \Gamma_{\mathcal{A}}^0 \) if and only if there are units \( a, b \in \mathcal{A} \) and an automorphism \( \varphi \) of the algebra \( \mathcal{A} \) such that

\[
A = L_a \circ \varphi, \quad B = R_b \circ \varphi, \quad C = \varphi^{-1} \circ L_a^{-1} \circ R_b^{-1}.
\]

**Proof.** The condition that \( A^* \otimes B^* \otimes C \) leaves the tensor of \( \mathcal{A} \) fixed is equivalent to

\[
\bigwedge_{x,y \in \mathcal{A}} xy = C((Ax)(By)).
\]

Therefore it is evident that, if \( A, B, C \) are as in (1), \( A^* \otimes B^* \otimes C \) is an element of the isotropy group of \( \mathcal{A} \). For the proof of the converse let

\[
a := A1, \quad b := B1.
\]

Fixing \( x = 1 \) in (2) we get

\[
\bigwedge_{y \in \mathcal{A}} y = C(a(By)),
\]

and similarly for \( y = 1 \):

\[
\bigwedge_{x \in \mathcal{A}} x = C((Ax)b).
\]

2 The term "endomorphism of the \( K \)-algebra \( \mathcal{A} \)" includes the rule \( \varphi(xy) = \varphi(x)\varphi(y) \) for \( x, y \in \mathcal{A} \).
(3) and (4) show that \( a \) and \( b \) are units of \( \mathcal{A} \) and that
\[
(Ax)b = a(Bx) \tag{5}
\]
for all \( x \in \mathcal{A} \). We show next that
\[
\psi: x \mapsto (C^{-1}x)b^{-1}a^{-1}
\]
is an automorphism of the algebra \( \mathcal{A} \). Clearly \( \psi \) is an automorphism of the vector space \( \mathcal{A} \). We have to show that \( \psi \) is homomorphic, i.e. \( \psi(xy) = \psi(x)\psi(y) \) for all \( x, y \in \mathcal{A} \). Because of \( C^{-1}1 = ab \) we obtain from (2) and (5):
\[
\psi(xy) = C^{-1}(xy)b^{-1}a^{-1} = (Ax)(By)b^{-1}a^{-1} = (Ax)b^{-1}a^{-1}(By)b^{-1}a^{-1} = \psi(x)\psi(y).
\]
Taking into account (3) and (4) we have for all \( x \in \mathcal{A} \)
\[
Ax = \psi(x)a, \quad Bx = a^{-1}\psi(x)ab, \quad Cx = \psi^{-1}(xb^{-1}a^{-1}).
\]
Hence the theorem follows with \( \varphi \) defined by
\[
\varphi(x) := a^{-1}\psi(x)a \quad (x \in \mathcal{A}).
\]

For the following considerations we must use some concepts and results from the structure theory of finite-dimensional algebras.

All we need can be found in [8], for the convenience of the reader however we recall here the most important definitions and facts.

An algebra \( \mathcal{A} \) over \( K \) is called simple if \( \mathcal{A} \) has no two-sided ideals other than \( \{0\} \) and \( \mathcal{A} \) itself. This means that every non-zero homomorphism of the algebra is an automorphism already. \( \mathcal{A} \) is called central if the center of \( \mathcal{A} \) is \( K \) (via the embedding \( a \mapsto a \cdot 1 \) we look upon \( K \) as a part of \( \mathcal{A} \)).

If \( D \) is a division-algebra over \( K \) then \( D \otimes M_n(K) \) is a simple algebra and it is a fundamental theorem that these examples exhaust the class of simple \( K \)-algebras. For \( K = \mathbb{R} \) this together with a famous theorem of Hurwitz ([8]) gives that each simple \( \mathbb{R} \)-algebra is (with suitable \( n \in \mathbb{N} \)) isomorphic to \( D \otimes M_n(\mathbb{R}) \), where \( D \) equals \( \mathbb{R}, \mathbb{C} \) or \( \mathbb{H} \), the division algebra of quaternions.

The automorphism group of a finite-dimensional central simple algebra \( \mathcal{A} \) coincides with the group of inner automorphisms of \( \mathcal{A} \). This is a corollary of the following important:

**Theorem 3.2.** (Skolem-Noether [8]). Let \( \mathcal{A} \) be a finite-dimensional simple \( K \)-algebra with center \( F \) and let \( \mathcal{B}, \mathcal{C} \) be simple sub-algebras of \( \mathcal{A} \) which contain \( F \). If \( \varphi \) is an isomorphism from \( \mathcal{B} \) onto \( \mathcal{C} \) leaving \( F \) element-wise fixed then there is a unit \( a \in \mathcal{A} \) such that \( \varphi(x) = a^{-1}xa \) for all \( x \in \mathcal{B} \).
Combining this with Theorem 3.1 we get

**Theorem 3.3.** Let \( \mathcal{A} \) be a finite-dimensional central simple \( K \)-algebra and \( A^* \otimes B^* \otimes C \) an automorphism of the vector space \( \mathcal{A}^* \otimes \mathcal{A}^* \otimes \mathcal{A} \). Then \( A^* \otimes B^* \otimes C \in \Gamma_{\mathcal{A}}^0 \) if and only if there are units \( a, b, c \in \mathcal{A} \) such that

\[
A = L_{a^{-1}} R_b, \quad B = L_{b^{-1}} R_c, \quad C = L_{a} R_c^{-1}.
\]

According to this result, the application of a mapping of the above type is called "sandwiching". Theorems 3.1 and 3.3 independently were proved also by Strassen and a result similar to (3.3) for the special case \( \mathcal{A} = M_n(\mathbb{K}) \) appears in [5].

From now on we will confine our discussion to algebras \( \mathcal{A} \) which are central simple and have an anti-automorphism \( \tau \), that is \( \tau(xy) = \tau(y)\tau(x) \) for all \( x, y \in \mathcal{A} \). Of course, if \( \mathcal{A} \equiv D \otimes M_n(\mathbb{K}) \) then \( \mathcal{A} \) possesses an anti-automorphism if and only if the division algebra \( D \) does. There are division algebras which have no anti-automorphism, i.e. which are not isomorphic to their opposite algebra \( D^0 \) (the opposite algebra \( \mathcal{A}^0 \) of an algebra \( \mathcal{A} \) is the vector space \( \mathcal{A} \) with \( (x, y) \mapsto yx \) as its multiplication). An example (over \( \mathbb{Q} \), the rationals) can be found in [3].

We will now consider the effect of permutational mappings on the tensor \( t \) of the algebra \( \mathcal{A} \). For this it seems to be convenient to make an identification of \( \mathcal{A} \) with its dual \( \mathcal{A}^* \). This can be achieved by fixing a basis in \( \mathcal{A} \) and working in \( \mathcal{A}^* \) with the dual basis. Thus we have the usual scalar product

\[
(x \mid y) = \sum_{i=1}^{N} x_i y_i.
\]

We will consider the permutational mappings \( \pi_{ik} \) determined by the transpositions \( i \leftrightarrow k \) of \( \{1, 2, 3\} \) (\( i, k \in \{1, 2, 3\}, i \neq k \)). Let \( \Phi_{ik} \) be the bilinear mapping determined by \( \pi_{ik} t \).

Let us begin with the study of \( \pi_{12} \): for any length-\( R \)-decomposition

\[
t = \sum_{r=1}^{R} u_r \otimes v_r \otimes w_r,
\]

of \( t \) into tensors of rank one we obtain

\[
\pi_{12} t = \sum_{r=1}^{R} v_r \otimes u_r \otimes w_r.
\]

Hence for all \( x, y \in \mathcal{A} \)

\[
\Phi_{12}(x, y) = \sum_{r=1}^{R} (x \mid v_r)(y \mid u_r)w_r = yx.
\]

Because of

\[
yx = \tau^{-1}(\tau(x)\tau(y))
\]
(8) yields

\[ \pi_{12}t = (\tau^* \otimes \tau^* \otimes \tau^{-1})t, \]  

(10)
i.e. \( \pi_{12} \circ (\tau^* \otimes \tau^* \otimes \tau^{-1}) \) leaves \( t \) fixed. (Since we are working with fixed dual bases in \( \mathcal{A} \) and \( \mathcal{A}^* \), the matrix of the dual \( \varphi^* \) of a linear mapping \( \varphi \) is just the transpose of the matrix of \( \varphi \).)

Now from (6) we get for all \( x, y \in \mathcal{A} \) that

\[ \Phi_{13}(x, y) = \sum_{r=1}^{R} (x \mid w_r)(y \mid v_r)u_r, \]

hence for all \( z \in \mathcal{A} \)

\[ (\Phi_{13}(x, y) \mid z) = (zy \mid x). \]  

(11)

Because of

\[ (zy \mid x) = (Ryz \mid x) = (z \mid R^*_y x) \]

we have to study the dual of the right-multiplication operator \( R_y \). Observe that the set \( R_{\mathcal{A}} := \{ R_x : x \in \mathcal{A} \} \) can be viewed as a central simple subalgebra of \( \mathbb{M}_N(K) \) which is isomorphic to the opposite algebra \( \mathcal{A}^0 \) of \( \mathcal{A} \). Similarly, \( \mathcal{A} \) is isomorphic to the central simple sub-algebra \( R_{\mathcal{A}}^0 := \{ R^*_x : x \in \mathcal{A} \} \). Observe further that \( R_1 = R_1^* = I_N \), the \( N \times N \)-unit-matrix. This is a consequence of \( \mathcal{A} \otimes \mathcal{A}^0 = \mathbb{M}_N(K) \) which follows from the simplicity of \( \mathcal{A} \) (see [8]). Hence there is a unique algebra-isomorphism \( \varphi : R_{\mathcal{A}} \rightarrow R_{\mathcal{A}}^* \) which makes the following diagram commutative:

\[ \begin{array}{ccc} 
R_{\mathcal{A}} & \rightarrow & R_{\mathcal{A}}^* \\
\varphi \downarrow & & \uparrow \varphi \\
\mathcal{A}^0 & \xrightarrow{\tau^{-1}} & \mathcal{A} 
\end{array} \]

\( \varphi, R_{\mathcal{A}} \) and \( R_{\mathcal{A}}^* \) satisfy the assumptions of Theorem 3.2, hence there is an automorphism \( S \) of the vector space \( \mathcal{A} \) such that for all \( x \in \mathcal{A} \)

\[ R_x^{\tau^{-1}(x)} = \varphi(R_x) = S^{-1}R_xS \]

(12)
i.e.

\[ \bigwedge_{y \in \mathcal{A}} R_y^{\tau(y)} = S^{-1}R_{\tau(y)}S. \]  

(13)
Consequently

\[ \Phi_{13}(x, y) = S^{-1}((Sx)\tau(y)), \]

i.e.

\[ \pi_{13}t = (S^* \otimes \tau^* \otimes S^{-1})t. \]  

(14)
Because \{\pi_{12}, \pi_{13}\} generates the permutation group \(\mathfrak{S}_3\) of three elements we have proved:

**Theorem 3.4.** Let \(A\) be a finite-dimensional central simple \(K\)-algebra with an anti-automorphism \(\tau\). Then there is an automorphism \(S\) of the vector space \(A\) such that the isotropy group \(\Gamma_{sl}\) of \(A\) is generated by \(\Gamma_{sl}^0\) and the following two mappings:

\[
\pi_{12} \circ (\tau^* \otimes \tau^* \otimes \tau^{-1}),
\]

\[
\pi_{13} \circ (S^* \otimes \tau^* \otimes S^{-1}).
\]

In particular we have \(\# (\Gamma_{sl}/\Gamma_{sl}^0) = 6\).

Let us consider the example \(A = M_n(K)\).

The canonical anti-automorphism of \(M_n(K)\) is defined by the transposition of \(n \times n\)-matrices: \(\tau(x) = x^t\), and we have \(R_x^* = R_{x^t}\) for all \(x \in M_n(K)\). Moreover \(\tau^* = \tau = \tau^{-1}\) and we can choose \(S = \text{id}\) in Theorem 3.4. Hence the elements generated by \(\pi_{12} \circ (\tau \otimes \tau \otimes \tau)\) and \(\pi_{13} \circ (\text{id} \otimes \tau \otimes \text{id})\) are

\[
\pi_{12} \circ (\tau \otimes \tau \otimes \tau), \quad \pi_{13} \circ (\text{id} \otimes \tau \otimes \text{id}),
\]

\[
\pi_{23} \circ (\tau \otimes \text{id} \otimes \text{id}), \quad \pi_{1} \circ (\text{id} \otimes \tau \otimes \tau),
\]

\[
\pi_{2} \circ (\tau \otimes \text{id} \otimes \text{id}), \quad \text{id} \otimes \text{id} \otimes \text{id},
\]

where \(\pi_1 := \pi_{13} \pi_{12}, \pi_2 := \pi_{12} \pi_{13}\), and \(\text{id}\) is the identity of \(M_n(K)\).

Similar results hold for \(A = \mathbb{H}(K)\), the algebra of quaternions over \(K\). Here \(\tau\) is the conjugation of quaternions: \((\xi_1, \xi_2, \xi_3, \xi_4)^T := (\xi_1, -\xi_2, -\xi_3, -\xi_4)\).

We shall now treat a problem closely related to the foregoing one. Consider the bilinear mapping

\[
\Phi : A \times A \to A \times A,
\]

\[
(x, y) \mapsto (xy, yx).
\]

For \(A = M_2(K)\) the algorithm variety for \(\Phi\) will be discussed in part III. An optimal algorithm for \(\Phi\) has length 9 if the characteristic of \(K\) is different from 2. If \(K\) is a real field and \(A = \mathbb{H}(K)\), then optimal algorithms for \(\Phi\) have length 10 ([7]).

First consider elements of the isotropy group of \(\varphi\) that are of the form \(A^* \otimes B^* \otimes C\). Such a map leaves the tensor of \(\Phi\) fixed iff

\[
(xy, yx) = C((Ax)(By), (By)(Ax))
\]

holds for all \(x, y \in A\). Hence \(A^* \otimes B^* \otimes C \in \Gamma_{\varphi}\) if

\[
A = L_{u^{-1}}R_u, \quad B = L_{v^{-1}}R_v, \quad C = \sigma_u \otimes \tau_v \otimes \tau_u
\]

or

\[
A = L_u R_{u^*}, \quad B = L_{v^*} R_v, \quad C = (\tau^{-1} \sigma_u \otimes \tau^{-1} \tau_v \otimes \tau_u) = \delta.
\]
where \( u, v \in \mathcal{A} \) are units, \( \sigma_u : x \mapsto u^{-1} x u \), \( \tau \) an anti-automorphism of \( \mathcal{A} \), and \( \delta : (x, y) \mapsto (y, x) \). For the discussion of the converse we will assume again that the algebra \( \mathcal{A} \) is central, simple and has an anti-automorphism \( \tau \). Moreover we will assume that the characteristic of \( K \) is different from 2. Setting \( a := A1, b := B1 \) we obtain from (15)

\[
\bigwedge_{x \in \mathcal{A}} (x, x) = C(a(Bx), (Bx)a) = C((Ax)b, b(Ax))
\]

which yields

\[
\bigwedge_{x \in \mathcal{A}} a(Bx) = (Ax)b, \quad (Bx)a = b(Ax) \quad (15)
\]

**Lemma 3.5.** There is a non-zero \( \lambda \in K \) such that \( ab = \lambda 1 \).

**Proof.** Put \( x_0 := B^{-1}1 \). From (16) we obtain

\[
(Ax_0)b = a \quad \text{and} \quad b(Ax_0) = a.
\]

Hence \( ba = b(Ax_0)b = ab \). From the first equation of (16) we conclude

\[
(Ax)ba = a(Bx)a \quad \text{for all} \ x \in \mathcal{A}
\]

and from the second

\[
ab(Ax) = a(Bx)a.
\]

In view of \( ab = ba \) we therefore have

\[
\bigwedge_{x \in \mathcal{A}} ab(Ax) = (Ax)ab,
\]

hence \( ab \) is in the center of \( \mathcal{A} \) which is \( K \). By means of (15) we see that \( ab = 0 \) would imply the contradiction

\[
(1, 1) = C(0, 0).
\]

After suitable scaling of \( A \) against \( C \) we may assume that \( ab = 1 \).

Now let \( \varphi(x) := a^{-1}(Ax) \ (x \in \mathcal{A}) \) and let \( p_2 : \mathcal{A} \times \mathcal{A} \to \mathcal{A} \) be the projection onto the second factor: \( p_2(x, y) := y \). Then, according to (16), we obtain for all \( x, y \in \mathcal{A} : (By)(Ax) = (By)a^{-1}(Ax) = a^{-1}(Ay)a^{-1}(Ax) \), thus

\[
(By)(Ax) = \varphi(y)\varphi(x),
\]

hence also

\[
(Bx)(Ay) = \varphi(x)\varphi(y).
\]

(17)

(18)
From (15) we get

\[(By)(Ax) = p_2C^{-1}(y, yx),\]
\[(Bx)(Ay) = p_2C^{-1}(yx, xy)\]  \hspace{1cm} (19)

and therefore, because \(\varphi(x) = b(Ax) = p_2C^{-1}(x, x)\):

\[(By)(Ax) + (Bx)(Ay) = p_2C^{-1}(xy + yx, xy + xy)\]
\[= a^{-1}(A(xy + yx)),\]  \hspace{1cm} (20)

i.e. we have shown that for all \(x, y \in \mathcal{A}\)

\[\varphi(x)\varphi(y) + \varphi(y)\varphi(x) = \varphi(xy) + \varphi(yx)\]  \hspace{1cm} (21)

holds.

This is the functional equation of a so called semi-automorphism.

Semi-automorphisms arose first from the study of certain projectivities over quaternions [1]. Ancochea proved in [2] that for simple algebras \(\mathcal{A}\) there are only two sorts of semi-automorphisms: the automorphisms and the anti-automorphisms. A more concise and at the same time more general treatment of the problem was given by Kaplansky in [11] (It is this characterization of semi-automorphisms where \(\text{Char}(K) \neq 2\) is needed).

If \(\varphi\) is an automorphism, we obtain from (15) for all \(x, y \in \mathcal{A}\):

\[(xy, yx) = C(\varphi(xy)a^{-1}, \varphi(yx)).\]

Thus, according to Proposition 2.2, we may choose

\[C = \varphi^{-1}\sigma_a \oplus \varphi^{-1}\]

where \(\sigma_a\) denotes the inner automorphism \(x \mapsto a^{-1}xa\).

Similarly, if \(\varphi\) is an anti-automorphism, we get

\[C = (\varphi^{-1} \oplus \varphi^{-1}\sigma_a) \circ \delta,\]

where \(\delta(x, y) := (y, x)\). By the Theorem 3.2 there is a unit \(c \in \mathcal{A}\) such that \(\varphi = \sigma_c\) if \(\varphi\) is an automorphism, and \(\varphi = \sigma_c \circ \tau\) if \(\varphi\) is an anti-automorphism.

Concerning the permutational mappings there is only one candidate to study, namely \(\pi_{12}\). Evidently we have

\[\Phi_{12}(x, y) = (yx, xy),\]

hence \(\pi_{12} \circ (id \otimes id \otimes \delta)\) leaves the tensor corresponding to \(\Phi\) fixed. Now put \(u := c \cdot t^{-1}\) and \(v := c\) to obtain

**Theorem 3.6.** Let \(K\) be a field whose characteristic is different from two. \(\mathcal{A}\) a finite-dimensional central simple \(K\)-algebra with an anti-automorphism \(\tau\). Then the
isotropy group $\Gamma_\Phi$ of the bilinear mapping

$$\Phi : A \times A \to A \times A$$

$$(x, y) \mapsto (xy, yx)$$

is generated by $\pi_{12} \circ (\text{id} \otimes \text{id} \otimes \delta)$, where $\delta : (x, y) \mapsto (y, x)$, and the group $\Gamma_\Phi$ of mappings $A^* \otimes B^* \otimes C$ with

$$A = L_{u^{-1}}R_v, \quad B = L_{v^{-1}}R_u, \quad C = \sigma_{u^{-1}} \oplus \sigma_{v^{-1}}$$

or

$$A = L_{u^{-1}}R_v\tau, \quad B = L_{v^{-1}}R_u\tau, \quad C = (\tau^{-1}\sigma_{v^{-1}} \oplus \tau^{-1}\sigma_{u^{-1}}) \circ \delta$$

($u, v$ units of $A$, $\sigma_u : x \mapsto u^{-1}xu$).

Acknowledgement

The author wishes to express his sincere thanks to A. Schönhage who introduced him to this part of mathematics and who sponsored this work by an abundance of fruitful discussions.

References