ALGEBRAIC THEORY OF DNA RECOMBINATION.

SERGEI R. SVERCHKOV

ABSTRACT. In this paper we investigate the structure and representation of n-ary algebras arising from DNA recombination, where n is a number of DNA segments participating in recombination. Our methods involve a generalization of the Jordan formalization of observables in quantum mechanics in n-ary splicing algebras.

We show that the splicing algebras are an n-ary envelope for algebras of DNA recombination. We have constructed the basis for free algebra of the variety of the n-ary splicing algebras and found the defining identities for n-ary splicing operations. Using the relationship between algebras and its enveloping algebras, we have constructed the basis of the free algebra of the variety of n-ary algebras of DNA recombination. It is proved that every polynomial identity satisfied by n-ary DNA recombination, with no restriction on the degree, is consequence of n-ary commutativity and the special n-ary identity of the degree 3n-2.

We obtain a criterion, analogous to the Specht-Wever Lie criterion, for determining whether an element of n-ary free splicing algebra is an element of the algebra of DNA recombination. Using this criterion, it is proved that all n-ary algebras of DNA recombination are special by module of n-ary splicing algebra (analog of the Poincare-Birkhoff-Witt theorem).

The skew-symmetrization of n-ary splicing operation converts the splicing algebras into n-ary skew-symmetric algebras. In case n = 2, they are Lie algebras. It is proved that every polynomial identity of these algebras, with no restriction on the degree, is consequence of centrally metabelian identity. In case n >2, they are nilpotent of index 3.

The non-homologous recombination is formalized by the algebras of simplified insertions. It is shown that all identities of the algebra of simplified insertion follow from the right-symmetric identity. We construct an infinite series of relations in the algebra of simplified insertion which hold for the words of length n and are not consequences of the right-symmetric identity

1. INTRODUCTION

The DNA recombination provides transfer of the genetic information and realization of the genetic program of development and functioning of all living organisms.

The algebraic formalization of DNA recombination is represented in the form of linear space $F(R)$ over a field F of characteristic 0, where R is an infinite free semigroup generated by the set of DNA nucleotides $\{A, G, C, T\}$. Homologous DNA recombination (exchange of DNA segments) defines the algebraic n-ary operation and transform $F(R)$ in an n-ary algebra J_n, where n is a number of DNA segments, participating in recombination. Algebra J_n is called *Morgan n - algebra.* The algebras from the variety $VAR(J_n)$, generated by J_n, are called *n - algebras of DNA recombination.*

Our purpose is to construct the algebraic theory for DNA recombination and to research structure and representations of the algebras of DNA recombination.

In case $n = 2$, the binary algebras of DNA recombination are Jordan algebras and were introduced by Bremner [4]. They are called *Jordan algebras of intermolecular recombination.* The structural theory of these algebras was developed in author's paper [5]. In Section 2 we will shortly review its main results.

Section 4 provides a detailed exposition of the structure and representation of n-algebras of DNA recombination for $n \geq 3$. We introduce the notion of n-ary splicing algebras and indicate how a generalization of the Jordan formalization of observables in quantum mechanics can be applied for n-ary splicing algebras and algebras of DNA recombination.

We show that the splicing algebras are an n-ary envelope for algebras of DNA recombination. We derive the interesting symmetric formulas in the free splicing algebras. Using these formulas, we construct the basis for free algebra of the variety of the n-ary splicing algebras and find the defining identities for n-ary splicing operations. Using the relationship between algebras and its enveloping algebras, we construct the basis of the free algebra of the variety of n-ary algebras of DNA recombination and its defining identities.

We obtain a criterion, analogous to the Specht-Wever Lie criterion, for determining whether an element of n-ary free splicing algebra is an element of the algebra of DNA recombination. Using this criterion, it is proved that all n-ary algebras of DNA recombination are special by module of n-ary splicing algebra (analog of the Poincare-Birkhoff-Witt theorem).

Finally, we solve the problem for identities of n-ary DNA recombination. It is proved that every polynomial identity satisfied by n-ary DNA recombination, with no restriction on the degree, is consequence of n-ary commutativity and the special n-ary identity of the degree 3n-2.

Section 4 is devoted to the study of the skew-symmetrization of n-ary splicing operation. This transformation converts the splicing algebras into n-ary skew-symmetric algebras. In case n = 2, they are Lie algebras. It is proved that every polynomial identity of these algebras, with no restriction on the degree, is consequence of centrally metabelian identity. In case n >2, they are nilpotent of index 3.

Section 5 deals with the case of non-homologous recombination. The non-homologous recombination is formalized by the algebras of simplified insertions. These algebras were investigated in author's paper [5]. We review its main results. It is shown that all identities of the algebra of simplified insertion follow from the right-symmetric identity. It is constructed an infinite series of relations in the algebra of simplified insertion which hold for the words of length n, $n \in \mathbb{N}$ and are not consequences of the right-symmetric identity.

Let's notice that research of identities of algebras of DNA recombination, construction of universal objects in variety of algebras of DNA recombination, are extremely important for theoretical genetics, mathematical biology and DNA computing. These identities and objects define universal laws of DNA recombination and give possibilities for construction of effective genetic algorithms. Applied value of these identities for gene engineering, is reduction of number of necessary experiments.

The results of Sec.2 and 5 are published with detailed proofs in [5, 6], of Sec. 3 and 4 are announced in [24].

All algebras in this work are considered over a field F of characteristic zero. We use standard definitions and notation from [17], and the theory of n-ary algebraic systems from Malcev's classical work [3].

2. The Jordan algebras arising from DNA recombination

Possibility of research of DNA recombination in theoretical genetic by means of non associative algebras goes back to Thomas Hant Morgan's (1916) classical work, A Critique of the Theory of Evolution [1]. On page 132 of this work the scheme illustrating recombination of (crossing over) two chromosomes $\boxed{a\ \ b}$ and $\boxed{c\ \ d}$ has been presented:

$$\boxed{a\ \ b} \times \boxed{c\ \ d} \to \boxed{a\ \ d} + \boxed{c\ \ b},$$

Where a, b, c, d are segments of homologous chromosomes participating in recombination.

What algebra defines Morgan's binary recombination? Let ab denote the chromosome $\boxed{a\ \ b}$. Then Morgan's recombination defines a commutative algebra J_2 over F with basis $a_i b_j$, $i, j \in \mathbb{N}$ and the multiplication table

$$\forall i,j,k,l \in \mathbb{N} \quad \mathrm{a}_i b_j \bullet \mathrm{a}_k b_l = \mathrm{a}_i b_l + \mathrm{a}_k b_j.$$

It has appeared that J_2 is Jordan algebra. That is, J_2 is a commutative algebra, satisfying to the Jordan identity

$$(x^2 \bullet \mathrm{y}) \bullet x = (x \bullet \mathrm{y}) \bullet x^2.$$

Moreover, J_2 is a special Jordan algebra. Let's remind definition of special Jordan algebra. Let A be an associative algebra over F. We will define new operation of symmetrized multiplication on A by $a \circ b = \dfrac{1}{2}(ab + ba), \ a,b \in A$. Well-known [17] that obtained algebra $A^{(+)}$ is Jordan algebra.

A Jordan algebra J is called *special* if J is isomorphic to a subalgebra of $A^{(+)}$, for some associative algebra A. In this case, the algebra A is called an *associative enveloping algebra* for J. A class $\mathfrak{M}$ of Jordan algebras is called *special* if it consists of special Jordan algebras.

It is easy to check that F-algebra $\tilde{J}_2$ with basis $\mathrm{a}_i b_j$, $i, j \in \mathbb{N}$ and the multiplication table

$$\forall i,j,k,l \in \mathbb{N} \quad \mathrm{a}_i b_j \bullet \mathrm{a}_k b_l = \frac{1}{2}(\mathrm{a}_i b_l + \mathrm{a}_k b_j), \tag{1}$$

is isomorphic to J_2. From now on J_2 stands for $\tilde{J}_2$, i.e. it has multiplication table (1).

Let A be a F-module with basis $\mathrm{a}_i b_j$, $i, j \in \mathbb{N}$. Following Kari [9], we will define on F-module A the *splicing operation* by rule

$$\forall i,j,k,l \in N \quad a_i b_j \triangleright a_k b_l = a_i b_l. \tag{2}$$

We will denote by C_2 the obtained algebra. It is easy to check that the splicing operation is associative:

$$(a_i b_j \triangleright a_k b_l) \triangleright a_s b_t \underset{(2)}{=} a_i b_t \underset{(2)}{=} a_i b_j \triangleright (a_k b_l \triangleright a_s b_t).$$

Therefore C_2 is an associative algebra. From genetic point of view, the splicing operation models only partial fragment of DNA recombination. But, as M. Bremner [4] has noticed, that

$$\mathrm{a}_i b_j \circ \mathrm{a}_k b_l = \frac{1}{2}(\mathrm{a}_i b_j \triangleright \mathrm{a}_k b_l + \mathrm{a}_k b_l \triangleright \mathrm{a}_i b_j) \underset{(2)}{=} \frac{1}{2}(\mathrm{a}_i b_l + \mathrm{a}_k b_j) \underset{(1)}{=} \mathrm{a}_i b_j \bullet \mathrm{a}_k b_l.$$

Hence the symmetrized multiplication $\circ$ on the associative algebra C_2 transforms it into algebra J_2, that is $C_2^{(+)} = J_2$. Hence J_2 is a special Jordan algebra and C_2 is an associative enveloping algebra for J_2.

The structural theory of Jordan algebras of intermolecular recombination has been constructed in author's paper [5]. We will shortly review its main results.

We will denote by $IR = Var(J_2), S = Var(C_2)$ the variety of Jordan algebras is generated by J_2 and the variety of associative algebras is generated by C_2. Algebras from $IR = Var(J_2)$ are called *Jordan algebras of intermolecular recombination*, or IR-algebras for short. Algebras from $S = Var(C_2)$ are called *splicing algebras*, or S-algebras for short.

Let $S[X], IR[X]$ be free algebras in the varieties S and IR, with set of generators $X = \{x_1,...,x_n,...\}$, respectively.

2.1. The basis and multiplication table of algebras $S[X]$ **and** $IR[X]$**.** Let $Ass[x], SJ[X]$ be free associative, free special Jordan algebras (see [17]) with set of generators $X = \{x_1,...,x_n,...\}$. We will define an ordering operator $\langle \ \rangle : Ass[x] \to SJ[X]$ by a rule: If $u = x_{s_1}...x_{s_m}$ is a monomial from $Ass[x]$ then

$$\langle u \rangle = x_{s_1} \circ ... \circ x_{s_m} \in SJ[X],$$

where $i_1 \le ... \le i_m$ and the set $\{s_1,...,s_m\}$ and $\{i_1,...,i_m\}$ coincide with respect to repetitions of all the symbols. Then we will extend the ordering operator on the algebra $Ass[x]$ by linearity: if $f = \sum_i \alpha_i u_i$, where $\alpha_i \in F$, u_i are monomials, then $\langle f \rangle = \sum_i \alpha_i \langle u_i \rangle$.

By definition, the operation of multiplication of the elements of the algebra $Ass[x]$, and consequently $SJ[X]$ within the brackets $\langle \ \rangle$ is associative-commutative. Therefore, for any $v_1,...,v_n \in Ass[x]$ and $\sigma \in S_n$, we have

$$\left\langle v_1 \circ ... \circ v_n \right\rangle = \left\langle v_1...v_n \right\rangle = \left\langle v_{\sigma(1)}...v_{\sigma(n)} \right\rangle,$$

where S_n is the symmetric group on $\{1,...,n\}$.

Let $B = \{x_i, x_i x_j = x_i \langle 1 \rangle x_j, x_i \langle u \rangle x_j\}$, where $x_i, x_i x_j, \langle u \rangle$ runs over all different ordered monomials of $SJ[X]$, 1 is a formal unit. We will denote by $\rhd$ the multiplication in the algebra $S[X]$.

Theorem 1. The algebra $S[X]$ has basis B and the multiplication table: $\forall i,j,k,l \in \mathbb{N}$

$$x_i \rhd x_j = x_i \langle 1 \rangle x_j;$$
$$x_i \rhd x_j \langle u \rangle x_k = x_i \langle u \rangle x_j \rhd x_k = x_i \langle u x_j \rangle x_k;$$
$$x_i \langle u \rangle x_j \rhd x_k \langle v \rangle x_l = x_i \langle u v x_j x_k \rangle x_l.$$

It is easy to check that $IR[X]$ is a special Jordan algebra with the associative enveloping algebra $S[X]$. Hence algebra $IR[X]$ is isomorphic to the subalgebra of $S[X]$ generated by the set X. Consequently we can replace the multiplication $\bullet$ in $IR[X]$ by the symmetrized product $\circ$ in $S[X]$.

Let $\overline{B} = \{x_i, x_i \bullet x_j = \langle 1 \rangle U_{x_i,x_j}, \langle u \rangle U_{x_i,x_j}\}$, where $x_i, x_i \bullet x_j, \langle u \rangle$ runs over all different ordered monomials of $SJ[X]$, 1 is a formal unit, $U_{x,y}$ is the Jordan U–operator of elements x, y, i.e. $zU_{x,y} = (z \bullet x) \bullet y + (z \bullet y) \bullet x - z \bullet (x \bullet y)$.

Theorem 2. The algebra $IR[X]$ has basis $\overline{B}$ and the multiplication table: $\forall i,j,k,l \in \mathbb{N}$

$$x_i \bullet x_j = \langle 1 \rangle U_{x_i,x_j};$$

$$2 \langle u \rangle U_{x_i, x_j} \bullet x_k = \langle ux_i \rangle U_{x_k, x_j} + \langle ux_j \rangle U_{x_i, x_k} \ ;$$

$$4 \langle u \rangle U_{x_i, x_j} \bullet \langle v \rangle U_{x_k, x_l} = \langle uvx_j x_l \rangle U_{x_i, x_k} + \langle uvx_j x_k \rangle U_{x_i, x_l} + \langle uvx_i x_l \rangle U_{x_k, x_j} + \langle uvx_i x_k \rangle U_{x_l, x_j} \ .$$

2.2. The basis of the identities of the algebras C_2 and J_2. Using the computer algebra, M. Bremner [4] has found the identity $Br(2)$ of algebra J_2 of degree 4 and has proved that all identities J_2 of degree ≤ 6 are consequences of Bremner identity $Br(2)$, where

$$Br(2) = (x^2 \cdot y) \cdot z + 2((x \cdot z) \cdot x) \cdot y - 2(x^2 \cdot z) \cdot y - (x^2 \cdot (z \cdot y)).$$

Actually, all identities of the algebra J_2 are consequences of the identity $Br(2)$.

Theorem 3. Basis of the identities of the algebra J_2 consists of single identity $Br(2)$.

The next theorem provides a characterization of the identities of the splicing algebra C_2.

Theorem 4. All identities of the algebra C_2 are consequences of the identity
$$x[y, z]t = 0 \, ,$$
where $[x, y] = xy - yx$ is commutator of x, y.

2.3. Speciality of Jordan algebras of intermolecular recombination. Let $\mathfrak{M}$ be some variety of Jordan algebras. Let a free algebra $\mathfrak{M}[X]$ in $\mathfrak{M}$ be a special Jordan algebra and $A[X]$ be some associative enveloping algebra for.

Let's remind that any antiisomorphism of $A[X]$ of degree 2 is called an *involution*. There exists a natural involution $*$ on $A[X]$ which acts on the monomials by $(x_1...x_n)^* = x_n...x_1$ and linearly extends over the whole algebra $A[X]$. We call it the standard "reversal" involution. We will denote by $HA[X]$ the Jordan algebra of symmetric elements of $A[X]$ in regard to $*$. It is obvious that $\mathfrak{M}[X] \subseteq HA[X]$.

Definition. The variety $\mathfrak{M}$ is called *reflexive*, if $\mathfrak{M}[X] = HA[X]$ for some associative enveloping algebra $A[X]$ of $\mathfrak{M}[X]$.

Theorem 5. Any reflexive variety of Jordan algebras is special.

A straightforward computation shows that $IR = Var(J_2)$ is a reflective variety.

Corollary. All Jordan algebras of intermolecular recombination are special.

2.6. Minimal algebras in the varieties $IR = Var(J_2)$ and $S = Var(C_2)$. Let $\mathfrak{M}$ be a variety of algebras over F.

Definition. The finite dimensional algebra $A \in \mathfrak{M}$ is called *minimal* in $\mathfrak{M}$, if $\mathfrak{M} = Var(A)$ and the algebra A has minimal dimension. The number $k = \dim_F(A)$ is called A-*dimension* of the variety $\mathfrak{M}$ and be denoted by $A \dim_F(\mathfrak{M}) = k$.

Let us denote by $J_2(n,m)$ the subalgebra of J_2 generated by a_ib_j, $1 \le i \le n$, $1 \le j \le m$; by $J_2(n)$ the subalgebra of J_2 generated by a_ib_j, $1 \le i \le n$, $j \in \mathbb{N}$. The algebras $J_2(n,m)$, $J_2(n)$, J_2 are called the *standard algebras* of intermolecular recombination.

For example, algebra $J_2(1,2)$ has basis $a = a_1b_1$, $a = a_1b_2$ and the multiplication table:

$$a^2 = a, \ b^2 = b, \ a{\bullet}b = \frac{1}{2}(a+b).$$

Theorem 6. The algebra $J_2(1,2)$ is minimal in $IR = Var(J_2)$ and $A\dim_F(IR) = 2$.

We will denote by $C_2(n,m)$ the subalgebra of C_2 generated by a_ib_j, $1 \le i \le n$, $1 \le j \le m$; by $C_2(n)$ the subalgebra C_2 of generated by a_ib_j, $1 \le i \le n$, $j \in \mathbb{N}$. The algebras $C_2(n,m)$, $C_2(n)$, C_2 are called the *standard splicing algebras*.

For example, algebra $C_2(1,3)$ has basis $e_1 = a_1b_1$, $e_2 = a_1b_2$, $e_3 = a_1b_3$ and the multiplication table: $e_i \triangleright e_i = e_i$, $1 \le i, j \le 3$.

Theorem 7. The algebra $C_2(2,2)$ is minimal in $S = Var(C_2)$ and $A\dim_F(S) = 4$.

2.5. Jordan Bernstein algebras and Jordan algebras of intermolecular recombination. The multiplication table (1) shows that the standard IR-algebras are algebras with genetic realization (see [17]). Among the algebras with genetic realization the class of Bernstein algebras holds an important position.

Let's remind that the Bernstein algebra over the field F is a commutative algebra J, with a non-zero algebra homomorphism $\omega : J \to F$, satisfying the identity

$$x^2{\cdot}x^2 = \omega(x)^2 x^2.$$

These algebras were introduced by P. Holgate [19], It is known [18] that the algebra J can be represented as $J = Fe \oplus N$, where $N = Ker\omega$, e is an idempotent and $n^2{\cdot}n^2 = 0$ for all $n \in N$.

If $ch(F) \ne 2$, then $N = U \oplus Z$, where $U = \{u \in N \mid e \cdot u = \frac{1}{2}u\}$, $Z = \{z \in N \mid e \cdot z = 0\}$. On the algebra N the following Bernstein graduation is defined:

$$U^2 \subseteq Z, \ U^2 \subseteq Z, \ U \cdot Z \subseteq U.$$

A Bernstein algebra is called Jordan, if it also satisfies the Jordan identity. The Jordan Bernstein algebras were first introduced by P. Holgate [20], who proved that the genetic algebras for the simple Mendel inheritance are special Jordan algebras. Later this result was generalized by A. Wors-Busekros [21]. It was shown that finite-dimensional Bernstein algebras with zero multiplication in N are special Jordan algebras. Also in the paper [21] the necessary and sufficient conditions for a Bernstein algebra to be Jordan were obtained: $Z^2 = 0$ and $N = U \oplus Z$ is nil-index 3 algebra. Jordan Bernstein algebras play an important role in the theory of Bernstein algebras (see [18, 22, 23]).

Definition. Bernstein algebra $B = Fe \oplus U \oplus Z$ is called *annihilator algebra* if Z coincides with annihilator of the algebra B, i.e. $Z = Ann(B)$.

The annihilator algebra $B = Fe \oplus U \oplus Z$ is called of *the type* (V,W), if

$$U = V \bigcup W, V = \{v_1,...,v_n\}, W = \{w_1,...,w_m\}, Z = \{v_i w_j; v_i \in V, w_j \in W\}, 1 \le n,m \le \infty,$$
$$v_i \cdot v_j = w_i \cdot w_j = 0, \mathrm{v}_i \cdot w_j = v_i w_j \text{ for all } v_i \in V, w_j \in W.$$

It is evident that in the annihilator algebras $Z^2 = N^2 = 0$. That is why all annihilator algebras are Jordan Bernstein algebras.

It easily seen that the annihilator algebra the type (V,W) has basis $\{e,\ v_i,\ w_j,\ v_i w_j,\ v_i \in V,\ w_j \in W\}$ and the multiplication table:

$$e^2 = e, e \cdot v_j = \frac{1}{2} v_j, e \cdot w_j = \frac{1}{2} w_j, \mathrm{v}_i \cdot w_j = v_i w_j,$$

all other products of the basis elements are zero.

Theorem 8. The class of standard IR-algebras coincides with the class of annihilator algebras of the type (V,W). An annihilator algebra of the type (V,W) is isomorphic to J_2, if $|V| = |W| = \infty$, to $J_2(n)$, if $|V| = n$, $|W| = \infty$, to $J_2(n,m)$ if $|V| = n$, $|W| = m$.

3. THE N-ALGEBRAS OF DNA RECOMBINATION

Let's consider now the case when n-segments of DNA participate in recombination, where $n \ge 3$. We will denote by $(a_{i_1 1}, a_{i_2 2},...,a_{i_n n})$ the DNA $\boxed{a_{i_1 1}}\boxed{a_{i_2 2}}...\boxed{a_{i_n n}}$, where $a_{i_1 1}, a_{i_2 2},...,a_{i_n n}$ are segments of DNA participating in recombination.

We are going to construct Morgan algebra J_n for n-ary recombination. It is clear that to consider all possible permutations of DNA segments, we have to multiply simultaneously n elements of J_n. Consequently, the operation on J_n will be n-ary. Let $\{x_1, x_2,...,x_n\}$ denote this operation.

By analogy to the algebra J_2 it is received that F-algebra J_n has basis $(a_{i_1 1}, a_{i_2 2},...,a_{i_n n})$, $i_1, i_2,...,i_n \in \mathbb{N}$ and multiplication table

$$\{b_1,...,b_n\} = \sum_{\sigma \in S_n} (b_{\sigma(1)1}, b_{\sigma(2)2},...,b_{\sigma(n)n}), \tag{3}$$

where $b_1 = (b_{11}, b_{12},...,b_{1n}) = (a_{i_1 1}, a_{i_2 2},...,a_{i_n n})$, $b_2 = (b_{21}, b_{22},...,b_{2n}) = (a_{j_1 1}, a_{j_2 2},...,a_{j_n n})$,..., $b_n = (b_{n1}, b_{n2},...,b_{nn}) = (a_{k_1 1}, a_{k_2 2},...,a_{k_n n})$.

Definition. The n-ary algebra J_n is called *Morgan n-algebra.* We call operation $\{x_1, x_2,...,x_n\}$ the *operation of n-ary DNA recombination.* Algebras from the variety $DR = VAR(J_n)$, generated by J_n, are called *n-algebras of DNA recombination* or *DR-algebras* for short.

We will denote by $DR[X]$ the free algebra in the variety DR and by $\{b_1, b_2,...,b_n\}$ the n-ary operation in $DR[X]$.

It follows easily that the operation of n-ary DNA recombination is commutative, e.g.

$$\forall \sigma \in S_n, a_1, ..., a_n \in J_n \quad \{a_1, ..., a_n\} = \{a_{\sigma(1)}, ..., a_{\sigma(n)}\}.$$

For multiplication in algebra J_n it is convenient to use the formula $s\det$ of the symmetric determinant. Set $a_1 = (a_{i1}, a_{i2}, ..., a_{in}) = a_{i1} \otimes a_{i2} \otimes ... \otimes a_{in}, i = 1, ..., n$.

Then

$$\{a_1, ..., a_n\} = s\det \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} = \sum_{i=1}^{n} a_{i1} \otimes s\det(M_{i1}),$$

where $M_{i1} = \begin{pmatrix} \widehat{a_{11}} & ... & a_{1n} \\ \widehat{a_{i1}} & ... & \widehat{a_{in}} \\ \widehat{a_{n1}} & ... & a_{nn} \end{pmatrix}$, the notation $\hat{a}_{ij}$ means that a_{ij} is absent.

For example,

$$n = 2, \quad \{a_i b_j, a_k b_l\} = s\det \begin{pmatrix} a_i & b_j \\ a_k & b_l \end{pmatrix} = a_i b_l + a_k b_j,$$

$$n = 3, \{a_1, a_2, a_3\} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} =$$

$$a_{11} \otimes s\det \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} + a_{21} \otimes s\det \begin{pmatrix} a_{11} & a_{12} \\ a_{32} & a_{33} \end{pmatrix} + a_{31} \otimes s\det \begin{pmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{pmatrix}.$$

Let's notice that operation of intermolecular recombination was first introduced by Kari and Landweber [8] and actively used in DNA computing. Operation of n-ary DNA recombination was first entered by Bremner [7].

3.1. Jordan formalization of DNA recombination. For the algebraic formalization of DNA recombination we will use the basic idea of the formalization of an algebra of observables in quantum mechanics, which was first realized by Pascual Jordan (1933) in classical work [2].

Let $\mathfrak{M}$ be a variety of the n-ary algebras over a field F and an algebra $B \in \mathfrak{M}$ with n-ary operation $f(b_1, ..., b_n)$, where $b_1, ..., b_n \in B$. We will define a new symmetrized n-ary operation on the F-module B by setting

$$j(b_1, ..., b_n) = \sum_{\sigma \in S_n} f(b_{\sigma(1)}, b_{\sigma(2)}, ..., b_{\sigma(n)}), \tag{4}$$

for all $b_1,...,b_n \in B$. The obtained algebra with the n-ary operation $j(b_1,...,b_n)$ will be denoted by $B^{(+)}$. The transformation of the algebra B to $B^{(+)}$ is called *Jordan symmetrization*, the transition from $B^{(+)}$ to B is called *Jordan specialization*.

Definition. A n-ary algebra J is called $\mathfrak{M}$-*special* if J is isomorphic to a subalgebra of $B^{(+)}$ for an algebra $B \in \mathfrak{M}$. The subalgebra of $B^{(+)}$ is generated by image of J under this embedding is called an $\mathfrak{M}$-*enveloping* algebra for J. A n-ary algebra J is called *weak $\mathfrak{M}$-special* if J is homomorphic image of a $\mathfrak{M}$-special algebra. A variety $\mathfrak{N}$ said to be $\mathfrak{M}$-*special (weak $\mathfrak{M}$-special)* if it consists of $\mathfrak{M}$-special *(weak $\mathfrak{M}$-special)* algebras.

We first construct the Jordan specialization for the n-ary algebra J_n. For this purpose, we consider the n-ary F-algebra C_n with the basis $(a_{i_1 1}, a_{i_2 2}, ..., a_{i_n n})$, $i_1, i_2, ..., i_n \in \mathbb{N}$ and multiplication table

$$< b_1,...,b_n > = (b_{11}, b_{22}, ..., b_{nn}) = (a_{i_1 1}, a_{j_2 2}, ..., a_{k_n n}), \tag{5}$$

where $b_1 = (b_{11}, b_{12}, ..., b_{1n}) = (a_{i_1 1}, a_{i_2 2}, ..., a_{i_n n})$, $b_2 = (b_{21}, b_{22}, ..., b_{2n}) = (a_{j_1 1}, a_{j_2 2}, ..., a_{j_n n}), ...,$ $b_n = (b_{n1}, b_{n2}, ..., b_{nn}) = (a_{k_1 1}, a_{k_2 2}, ..., a_{k_n n})$.

Definition. The n-ary algebra C_n is called the *standard splicing n-algebra*. We call operation $< x_1, x_2, ..., x_n >$ the n-*ary splicing operation*. Algebras from variety $S(n) = VAR(C_n)$, generated by C_n, are called the *splicing n-algebras* or $S(n)$-*algebras* for short.

The basis elements in the algebra C_n are easy multiplied by using the main diagonal rule. Let $a_1 = (a_{i1}, a_{i2}, ..., a_{in})$, $i = 1,...,n$. Then

$$< a_1,...,a_n > = < \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} > = (a_{11}, a_{22}, ..., a_{nn}).$$

Note, that the n-ary splicing operation is not associative for $n \geq 3$ and not commutative for $n \geq 2$.

Indeed,

$$n = 2, \ < a_1, a_2 > = (a_{11}, a_{22}) \neq < a_2, a_1 > = (a_{21}, a_{12});$$

$$n = 3, \ << a_1, a_2, a_3 >, a_4, a_5 > \underset{(5)}{=} < (a_{11}, a_{22}, a_{33}), a_4, a_5 > = (a_{11}, a_{42}, a_{53})$$

$$\neq < a_1, a_2, < a_3, a_4, a_5 >> \underset{(5)}{=} < a_1, a_2, (a_{31}, a_{42}, a_{53}) > = (a_{11}, a_{22}, a_{53}).$$

But, how it was noticed in Sec.2, the n-ary splicing operation is associative for $n = 2$.

It is easy to check that algebra J_n is $S(n)$-special.

We have

$$j(<b_1,...,b_n>) \underset{(4)}{=} \sum_{\sigma \in S_n} <b_{\sigma(1)}, b_{\sigma(2)},...,b_{\sigma(n)}> \underset{(5)}{=} \sum_{\sigma \in S_n} (b_{\sigma(1)1}, b_{\sigma(2)2},...,b_{\sigma(n)n}) \underset{(3)}{=} \{b_1,...,b_n\},$$

for any $b_1 = (b_{11}, b_{12},...,b_{1n}) = (a_{i_1 1}, a_{i_2 2},...,a_{i_n n}), b_2 = (b_{21}, b_{22},...,b_{2n}) = (a_{j_1 1}, a_{j_2 2},...,a_{j_n n}),...,$
$b_n = (b_{n1}, b_{n2},...,b_{nn}) = (a_{k_1 1}, a_{k_2 2},...,a_{k_n n})$ from C_n.

From this we conclude that the symmetrized product on C_n coincides with the multiplication on J_n and $J_n \simeq C_n^{(+)}$.

We will denote by $S(n)[X]$ the free algebra in the variety $S(n)$ and by $<b_1, b_2,...,b_n>$ the n-ary operation in $S(n)[X]$.

3.2. Basis of free algebra $S(n)[X]$ and defining identities of C_n.

Definition. The subalgebra of $S(n)^{(+)}[X]$ generated by the set X is called the *free $S(n)$-special algebra* and denoted by $SDR[X]$. The elements of $SDR[X]$ are called J-*polynomials*. The criterion of definition, whether $a \in S(n)[X]$ is J-polynomial, is called J- *criterion*.

Theorem 1. The algebra $DR[X]$ is $S(n)$-special and it is isomorphic to $SDR[X]$. The variety $DR = VAR(J_n)$ is a weak $S(n)$-special variety.

From Theorem 1. it follows that in $S(n)[X]$

$$j(<b_1,...,b_n>) = \sum_{\sigma \in S_n} <b_{\sigma(1)}, b_{\sigma(2)},...,b_{\sigma(n)}> = \sum_{\sigma \in S_n} (b_{\sigma(1)1}, b_{\sigma(2)2},...,b_{\sigma(n)n}) = \{b_1,...,b_n\},$$

for all $b_1,...,b_n \in S(n)[X]$. Hence, without loss of generality we can write $DR[X] = SDR[X]$.

From now on, the notations $u = \{a_1,...,a_m\}, v =< a_1,...,a_m >$ mean that
$m = k(n-1)+1, k \geq 1$ and $u = \{a_1,...,a_m\} = \{\{\{... \{a_1,...,a_n\},...,a_{2(n-1)+1}\},...,a_{k(n-1)+1}\}$,
$v =< a_1,...,a_m >=<<< ... < a_1,...,a_n >,...,a_{2(n-1)+1} >,...,a_{k(n-1)+1} >$.

Let B denote the set

$$\{ x_i, \ < x_{i_1},...,x_{i_n} >, \ < x_{i_1}, \{x_{j_1},...,x_{j_m}\}, x_{j_{m+1}},...,x_{j_{m+n-2}}, x_{i_2},...,x_{i_n} > \},$$

where $m = l(n-1)+1, l \geq 0$, $j_1 \leq ... \leq j_{m+n-2}$; $i, i_1,...,i_n, j_1,..., j_{m+n-2} \in \mathbb{N}$.

Theorem 2. The set B is a basis of the algebra $S(n)[X]$.

Let

$h = < x_1, y_1, ..., y_{n-1}, z_1, ..., z_{n-1}, t_1, ..., t_{n-1} > - < x_1, z_1, y_2, ..., y_{n-1}, y_1, z_2, ..., z_{n-1}, t_1, ..., t_{n-1} >,$

$g_k = << x_1, y_1, ..., y_{n-1} >, z_1, ..., z_{n-1} > - < x_1, z_1, ..., z_{k-1}, < y_1, ..., y_k, z_k, y_{k+1}, ..., y_{n-1} >, z_k, ..., z_{n-1} >,$

$f_k = < x_1, y_1, ..., y_k, ..., y_{n-1}, z_1, ..., z_{n-1} > - < x_1, y_k, ..., y_1, ..., y_{n-1}, z_1, ..., z_{n-1} >,$

where $k = 1, ..., n-1$.

Direct computations show that h and g_k, f_k, $k = 1, ..., n-1$ are identities in the algebra C_n and more delicate that they are the defining identities of C_n.

Theorem 3. All identities in the algebra C_n are consequences of h and g_k, f_k, $k = 1, ..., n-1$.

3.3. The symmetric formulas in free algebra $S(n)[X]$**.** In much the same way as in associative algebras, we can define the standard "reversal" involution of the algebra $S(n)[X]$ by setting:

$\forall x_i \in X \quad x_i^* = x_i$,

$\forall a_1, ..., a_n \in S(n)[X] \quad < a_1, ..., a_n >^* = < a_n^*, ..., a_1^* >.$

We check at once that the set $H(n)[X] = H(S(n)[X], *) = \{ u \in S(n)[X] \mid u^* = u \}$ of the symmetric elements in $S(n)[X]$ under $*$ is a subalgebra of $S(n)^{(+)}[X]$ and $DR[X] \subseteq H[X]$. The algebra $H(n)[X]$ is called *DR-algebra of symmetric elements of* $S(n)[X]$.

It Sec.2 it is shown that $H(2)[X] = DR[X]$, for $n \geq 3$ the situation is completely different, i.e. $H(n)[X] \neq DR[X]$.

Theorem 4. For $n \geq 3$ $H(n)[x_1, ..., x_k] = DR[x_1, ..., x_k]$ if and only if $k = 1$.

In particularly, the element $u = << \underbrace{x_1, ..., x_1}_{n} >, x_2, \underbrace{x_1, ..., x_1}_{n-2} >$ is symmetric but is not J-polynomial.

Let us consider the symmetric element $w = << y, \underbrace{x, ..., x}_{n-1} > \underbrace{y, ..., y}_{n-1} > \in S(n)[X]$. It appears that w is a J-polynomial.

Set

$$F_k = \{ \underbrace{x, ..., x}_{k} \underbrace{y, ..., y}_{n-k} \},$$

$$A_k(x, y) = \{ F_k, \underbrace{x, ..., x}_{n-1-k} \underbrace{y, ..., y}_{k} \},$$

where for k=1,...,n . Note, that F_k, $A_k(x, y) \in DR[x, y]$.

We will denote by $\dfrac{d(z)}{dt}$ the operator of differential substitution $z \to t$.

Theorem 5. For $n \geq 2$,

$$<< y, \underbrace{x,...,x}_{n-1} > \underbrace{y,...,y}_{n-1} >= \frac{1}{(n!)^2} \sum_{k=1}^{n} (-1)^{(k+1)} C_n^k A_{n-k}(x, y)$$

is an identity in $S(n)[X]$; 　　　　　　　　　　　　　　　　　　　　　　　　　　　(6)

$$\{ \underbrace{\{x,...,x\}}_{n}, \underbrace{y,...,y}_{n-1} \} = \frac{(n-1)!}{n!} \frac{d(y)}{dx} (\sum_{k=1}^{n} (-1)^{(k+1)} C_n^k A_{n-k}(x, y))$$

is an identity in any n-ary commutative algebra;

where $C_n^k = \dfrac{n!}{k!(n-k)!}, 0! = 1$.

We call (6) the symmetric formulas in free algebra $S(n)[X]$.

For example, for $n = 3$

$$36 << y, x, x >, y, y >= \{\{y, y, y\}, x, x\} + 3(\{\{y, x, x\}, y, y\} - \{\{y, y, x\}, y, x\}).$$

The symmetric formulas give a main tool for constructing algebraic theory of DNA recombination.

3.4. The basis of the free algebra $DR[X]$ and the defining identities of J_n.

Let

$$S(y_1,..., y_n \mid z_1,..., z_{n-1}) = \frac{1}{n!(n-1)!} \frac{d^n(y)}{dy_1...dy_n} \frac{d^{n-1}(z)}{dz_1...dz_{n-1}} \sum_{k=1}^{n} (-1)^{(k+1)} C_n^k A_{n-k}(z, y),$$

$$S(y_1,..., y_n) = \{y_1,..., y_n\}, \text{ for } y_1,..., y_n, z_1,..., z_{n-1} \in DR[X].$$

By (6), we can easily get identities in the algebra $DR[X]$:

$$b_1 = \{\{\underbrace{x,...,x}_{n}\}\{\underbrace{y,...,y}_{n}\}, \underbrace{z,...z}_{n-2}\} - \sum_{k=1}^{n} (-1)^{k+1} C_n^k \{A_{n-k}(x, y), x, \underbrace{z,...z}_{n-2}\},$$

$$b_2 = \{S(\underbrace{y,...,y}_{n} \mid \underbrace{x,...,x}_{n-1}), \underbrace{z,...z}_{n-1}\} - S(y, \underbrace{z,...z}_{n-1} \mid \{y, \underbrace{x,...,x}_{n-1}\}, \underbrace{y,...,y}_{n-2}),$$

$$b_3 = S(\underbrace{y,...y}_{n} \mid \{x_1,..., x_n\}, z_1,..., z_{n-2}) - S(\underbrace{y,...y}_{n} \mid \{z_1, x_2,..., x_n\}, x_1, z_2,..., z_{n-2}),$$

$$b_4 = S(\underbrace{y,...y}_{n} \mid \{x_1,..., x_n\}, \{z_1,..., z_n\}, \underbrace{t,...,t}_{n-3})$$

$$-S(\underbrace{y,...y}_{n} \mid \{z_1, x_2,..., x_n\}, \{x_1, z_2,..., z_n\}, \underbrace{t,...,t}_{n-3}).$$

Combining the symmetric formulas and the identities $b_1 - b_4$ we can construct a basis of $DR[X]$.

Let B denote the set

$$x_i, \quad S(x_{i_1},...,x_{i_n}), \quad S(x_{i_1},x_{i_2},...,x_{i_n} \mid \{x_{j_1},...,x_{j_m}\},x_{j_{m+1}},...,x_{j_{m+n-2}}),$$

where $m = l(n-1)+1, l \geq 0, i_1 \leq ... \leq i_n, j_1 \leq ... \leq j_{m+n-2}$; $i,i_1,...,i_n, j_1,...,j_{m+n-2} \in \mathbb{N}$.

Theorem 7. The set B is a basis of the algebra $DR[X]$.

The principal point of the proof of the Theorem 7 is in using only identities $b_1 - b_4$ for decomposition of any element of $DR[X]$ in linear combination of the basis ones. Therefore $b_1 - b_4$ are the defining identities for J_n.

Theorem 8. All identities in the algebra J_n are consequences of $b_1 - b_4$.

3.5. The basis of the identities of n-ary DNA recombination.

Let A be an arbitrary n-ary algebra, $n \geq 2$.

Definition. An identity in n-ary algebra A is called r_1-*identity,* if it has form

$$\{\{x_1,...,x_n\},\{x_{n+1},...,x_{2n}\},x_{2n+1},...,x_m\} - \sum_{\sigma \in S_m} \alpha_\sigma \{x_{\sigma(1)},...,x_{\sigma(m)}\},$$

where $\alpha_\sigma \in F$.

An polynomial $a \in A$ is called r_1-*polynomial,* if $a = \sum_{\sigma \in S_m} \alpha_\sigma \{x_{\sigma(1)},...,x_{\sigma(m)}\}$,

where $\alpha_\sigma \in F$, $x_1,...,x_m \in X$.

It is evident that if an arbitrary commutative n-ary algebra A satisfies some r_1-identity, then all elements of A are r_1-polynomials.

By using computer algebra, Bremner [7] has found r_1-identity for algebra J_3 and proved that all identities of degree ≤ 9 in J_3 follow from it. In view of results [4, 5, 7], Bremner conjectured that for $n \geq 3$:

(a). The algebra J_n satisfies some r_1-identity;

(b). The basis of identities of J_n consists of only one identity. We will denote by $Br(n)$ this hypothetical Bremner identity.

The identity b_1 is r_1-identity in J_n and it gives the positive answer to conjecture (a).

We will now design the identity $Br(n)$ and give the positive answer to conjecture (b).

Set

$$\overline{b_2} = \frac{d^n(a)}{dx_1...dx_n}(\{S(\underbrace{a,...,a}_{n} \mid \underbrace{y,...,y}_{n-1}), x, x, \underbrace{z,...,z}_{n-3}\} - S(a,x,x,\underbrace{z,...,z}_{n-3} \mid \{a, \underbrace{y,...,y}_{n-1}\}, \underbrace{a,...,a}_{n-2})),$$

$$\overline{b_3} = S(x_1,...,x_{n-2},x,x \mid \{x_{n-1}, \underbrace{y,...y}_{n-1}\}, x_n, \underbrace{z,...,z}_{n-3}) - S(x_1,...,x_{n-2},x,x \mid \{x_n, \underbrace{y,...y}_{n-1}\}, x_{n-1}, \underbrace{z,...,z}_{n-3}).$$

Let us consider the identity

$$Br(n) = \overline{b_2} + \overline{b_3}.$$

It is clear that $b_2 \Leftrightarrow \overline{b_2}$, $b_3 \Leftrightarrow \overline{b_3}$, $b_2, b_3 \Rightarrow Br(n)$. The identity $\overline{b_2}$ is symmetric in x_{n-1}, x_n, but the identity $\overline{b_3}$ is skew-symmetric in x_{n-1}, x_n. Therefore

$$Br(n)(x_{n-1}, x_n) - Br(n)(x_n, x_{n-1}) = 2\overline{b_3}$$

and $Br(n) \Rightarrow \overline{b_3}$. Consequently, $Br(n) \Rightarrow \overline{b_2} = Br(n) - \overline{b_3}$ and $Br(n) \Rightarrow b_2, b_3$.

More specific and delicate calculations show that $Br(n) \Rightarrow b_1, b_4$. Therefore $Br(n)$ is the defining identity of J_n by Theorem 8.

Theorem 9. All identities in the algebra J_n are consequences of $Br(n)$.

3.6. J - criterion and speciality of n -algebras of DNA recombination.

There are two J - criterion well-known in theory of Lie algebras, that of Friedrichs and that of Specht and Weber [10].

Problem of J - criterion in Jordan algebras is at present far from being solved. Classical Shirshov-Kohn [11, 17] theorem states that $a \in Ass[x_1, x_2, x_3]$ is J-polynomial if and only if a is symmetric element under $*$ - the standard "reversal" involution. But it is not J - criterion for $Ass[x_1, x_2, x_3, x_4]$ (see [17] for more details). Let's note the elegant J - criterion for Jordan r_1-polynomial of Robins [12].

Now we are going to construct J - criterion for $S(n)[X]$, which is an analog of the Specht-Weber criterion for Lie algebras.

For this purpose, we define the partial involutions $\varphi_{ij} : S(n)[X] \rightarrow S(n)[X]$, where $1 \le i, j \le n$, by the following rules:

$\forall x \in X \qquad \varphi_{ij}(x) = x,$

$\forall w_1,..., w_n \in S(n)[X] \quad \varphi_{ij}(< w_1,..., w_n >) = < w_1,..., \varphi_{ij}(w_j)\mid_i,...,\varphi_{ij}(w_i)\mid_j,..., w_n >,$

the notation $< a_1,..., b \mid_i,..., a_n >$ means $< a_1,..., a_{i-1}, b, a_{i+1},..., a_n >$.

It is easily seen that $SH[X] = \{u \in S(n)[X] \mid \varphi_{ij}(u) = u, \ 1 \le i, j \le n\}$ is a subalgebra of $S(n)^{(+)}[X]$ and $DR[X] \subseteq SH[X]$.

Definition. The algebra $SH[X]$ is called the *algebra of supersymmetric elements of $S(n)[X]$.*

Theorem 10. $DR[X] = SH[X]$.

Let us consider an arbitrary permutation $\sigma \in S_n$. Let $\sigma = \tau_1 \cdot \tau_2 \cdot ... \cdot \tau_k$, where $\tau_1, \tau_2, ..., \tau_k$ are transpositions, and $\varphi_\sigma = \varphi_{\tau_1} \circ \varphi_{\tau_2} \circ ... \circ \varphi_{\tau_k}$.

We will denote by

$$T = \sum_{\sigma \in S_n} \varphi_\sigma - (n!) \cdot 1$$

the linear operator $T : S(n)[X] \rightarrow S(n)[X]$, where 1 is the unitary mapping.

The operator T gives us the following J - criterion for $S(n)[X]$.

Theorem 11. A polynomial $a \in S(n)[X]$ is a J - polynomial if and only if $T(a) = 0$.

By J - criterion, it is easy to get speciality of n-algebras of DNA recombination.

Theorem 12. The variety $DR = VAR(J_n)$ is $S(n)$ - special.

4. LIE ALGEBRAS OF DNA RECOMBINATION.

The algebras $L(n,m) = C_2(n,m)^{(-)}$, $L(k) = C_2(k)^{(-)}$, $L = C_2^{(-)}$, $n,m,k \in \mathbb{N}$ are called the *standard Lie algebras* of DNA recombination.

Definition. Algebras from $LD = Var(L)$ are called *Lie algebras of DNA recombination*, or L - algebras for short.

We will denote by $\cdot$ the multiplication in the L-algebras. In particularly,

$$\forall a,b \in L \quad a \cdot b = [a,b] = a \triangleright b - b \triangleright a$$

By definition, the F-algebra L is a Lie algebra with basis $a_i b_j$, $i,j \in \mathbb{N}$ and the multiplication table

$$\forall i,j,k,l \in \mathbb{N} \quad a_i b_j \bullet a_k b_l = a_i b_l - a_k b_j.$$

We will denote by CM the varieties of centrally metabelian algebras, i.e. algebras satisfy the identity

$$((x_1 \cdot x_2) \cdot (x_3 \cdot x_4)) \cdot x_5 = 0, \tag{7}$$

and by M the variety of metabelian algebras, i.e. algebras satisfy the identity

$$(x_1 \cdot x_2) \cdot (x_3 \cdot x_4) = 0. \tag{8}$$

It is easy to check that the algebra L satisfies (7). Actually, all identities of the algebra L are consequences of the identity (7).

Theorem 1. $LD = CM$.

We can describe the subvarieties generated by the standard Lie algebras of DNA recombination

Theorem 2.

$Var(L(1,1))$ is the variety of abelian Lie algebras;
$Var(L(1)) = Var(L(1,k)) = M$;
$Var(L) = Var(L(n)) = Var(L(m,k)) = CM$, for $k,m,n \geq 2$.

Let us consider the Lie algebra $L(2,2)$. It has basis $e_i,$ $1 \leq i \leq 4$ and the multiplication table:
$$e_1 \cdot e_2 = e_2 - e_1,\ e_1 \cdot e_3 = e_1 - e_3,\ e_1 \cdot e_4 = e_2 - e_3,\ e_2 \cdot e_3 = e_1 - e_4,\ e_2 \cdot e_4 = e_2 - e_4,\ e_3 \cdot e_4 = e_4 - e_3.$$

Theorem 3. The algebra $L(2,2)$ is minimal in $LD = Var(L)$ and $A\dim_F(LD) = 4$.

We can define a new skew-symmetric n-ary operation on C_n , $n \geq 3$, by

$$[b_1,...,b_n] = \sum_{\sigma \in S_n} (-1)^{\sigma} < b_{\sigma(1)}, b_{\sigma(2)},...,b_{\sigma(n)} > ,$$

for any $b_1 = (b_{11}, b_{12},...,b_{1n}) = (a_{i_1 1}, a_{i_2 2},...,a_{i_n n}), b_2 = (b_{21}, b_{22},...,b_{2n}) = (a_{j_1 1}, a_{j_2 2},...,a_{j_n n}),...,$
$b_n = (b_{n1}, b_{n2},...,b_{nn}) = (a_{k_1 1}, a_{k_2 2},...,a_{k_n n})$ from C_n .

The algebra with this new operation will be denoted by $C_n^{(-)}$. In case $n \geq 3$, the algebra $C_n^{(-)}$ is nilpotent of index 3.

Theorem 4. , All identities of the algebra $C_n^{(-)}$, $n \geq 3$, are consequences of the identity $[[x_1,...,x_n], y_2,..., y_n] = 0$.

5. NON-HOMOLOGOUS RECOMBINATION AND THE ALGEBRA OF SIMPLIFIED INSERTIONS

The schematic model of non-homologous DNA recombination can be represented in the form:

$$\boxed{a}\boxed{b} \times \boxed{c} \rightarrow \boxed{a}\boxed{c}\boxed{d}, \tag{9}$$

where two chromosomes $\boxed{a}\boxed{b}$ and $\boxed{c}$ are participating in non-homologous recombination.

The algebraic formalization of the recombination (9), with respect of all possible insertions DNA $\boxed{c}$ in DNA $\boxed{a}\boxed{b}$, defines the algebra of simplified insertions.

Given the free associative algebra $Ass[X]$ with set of generators $X = \{x_1,...,x_n,...\}$, we define a new multiplication operation $*$ by the following rule:

$$(x_{i_1}...x_{i_n}) * a = \sum_{k=0}^{n} x_{i_1}...x_{i_k} a x_{i_{k+1}}...x_{i_n},$$

where a is an arbitrary element in $Ass[X]$; and

$$b * a = \sum \alpha_s (u_s * a),$$

where $b = \sum \alpha_s u_s$, $\alpha_s \in F$, and u_s are monomials in $Ass[X]$.

The new operation $*$ is called the *operation of right simplified insertion*. The operation of simplified insertion was firstly introduced by M. Bremner [13], and it is an algebraic formalization of the operation of normal insertion in the theory of DNA computing (see [14, 14]).

The new algebra $\langle Ass[X], +, * \rangle$ is called the *algebra of right simplified insertion* on the generator set X, and it is denoted by $Ass^*[X]$. It is known that $Ass^*[X]$ satisfies the right-symmetric identity

$$(x, y, z) = (x, z, y), \tag{10}$$

where $(x, y, z) = (x * y) * z - x * (y * z)$ is the associator of $x, y, z \in Ass^*[X]$. This fact was first proved by M. Gerstenhaber in [16].

Analogously, we may define the operation of *left simplified insertion*

$$a * (x_{i_1}...x_{i_n}) = \sum_{k=0}^{n} x_{i_1}...x_{i_k} a x_{i_{k+1}}...x_{i_n},$$

and construct the algebra of *left simplified insertion*. This algebra satisfies the left-symmetric identity $(x, y, z) = (y, x, z)$. It is easy to notice that the categories of the algebras of left and right insertion are equivalent; the same happens to the categories of the left-symmetric and right-symmetric algebras. Every algebra of this kind changes its left (right) symmetry to the opposite by the new multiplication $(x, y) = y * x$. This new operation "rewrites" every left (right) relation into its right (left) analog.

From now on, the term *algebra of simplified insertion* stands for an algebra of right simplified insertion. We will shortly review main results on the algebra of simplified insertion from author's paper [6].

5.1. The basis of identities of the algebra of simplified insertion.

By using computer algebra, M. Bremner proved in [13] that all identities of degree 4 in $Ass^*[X]$ follow from (10). For the identities of degree 5 Bremner constructed a relation that is satisfied by words of length 2 of $Ass^*[X]$. Actually, all identities of the algebra $Ass^*[X]$ are consequences of the identity (10).

Theorem 1. All identities of $Ass^*[X]$ follow from the right-symmetric identity.

By analogy with free special Jordan algebras $SJ[X]$, the subalgebra in $Ass^*[X]$ generated by X will be called the *special algebra of simplified insertion* on generators X and denoted by $SI[X]$.

Theorem 2. The linear spaces $SJ[X]$ and $SI[X]$ coincide. All identities of $SI[X]$ follow from the right-symmetric identity.

5.2. The Dirichlet relations on $Ass^*[X]$.

Now, we construct an infinite series of relations that hold for all words in $Ass^*[X]$ of length k, $k \in \mathbb{N}$ and are not consequences of the right-symmetric identity.

The algorithm for construction of these relations is connected with the simple Dirichlet principle: it is impossible to place $n+1$ rabbits in n cages so that every cage contains one rabbit.

Formalize the Dirichlet allocation algorithm. Let $a = x_1...x_n$ be a monomial in $Ass[X]$, and let $r_1,...,r_k$ and $k \leq n+1$, be the variables to be allocated. We need to allocate $r_1,...,r_k$ in $n+1$ cells $\lfloor \ \rfloor x_1 \lfloor \ \rfloor x_2 \lfloor \ \rfloor...\lfloor \ \rfloor x_n \lfloor \ \rfloor$ so that every cell $\lfloor \ \rfloor$ contains no more than one variable. Define the allocation operator $T(r_1, \ldots, r_k): Ass[X] \rightarrow Ass[X \cup \{r_1, \ldots, r_k\}]$, where $k \leq n+1$, by the following rule:

$$aT(r_1, \ldots, r_k) = \sum_{\sigma \in S_k} \sum_{\substack{u_0,...,u_k \\ u_0 u_1...u_k = a \\ \deg u_i \geq 1 \ when \ i=1,...,k-1}} u_0 r_{\sigma(1)} u_1 r_{\sigma(2)}...u_{k-1} r_{\sigma(k)} u_k .$$

For example,
$$x_1 T(a) = ax_1 + x_1 a, \ x_1 x_2 T(a, a) = 2(ax_1 ax_2 + ax_1 x_2 a + x_1 ax_2 a),$$

$$x_1 x_2 x_3 T(a, a, a, a) = 4! ax_1 ax_2 ax_3 a .$$

Denote by R_a, $a \in Ass^*[X]$, the right multiplication operator in $Ass^*[X]$, i.e.,

$$\forall b \in Ass^*[X] \quad bR_a = b * a .$$

Denote by $R(Ass^*[X])$ the algebra of right multiplications $Ass^*[X]$.

By recursion on m, define $D(x_1, \ldots, x_m) \in R(Ass^*[X])$:

$$D(x_1) = Rx_1 ,$$

$$D(x_1, \ldots, x_m) = D(x_1, \ldots, x_{m-1})R_{x_m} - \sum_{i=1}^{m-1} D(x_1,..., x_i * x_m,..., x_{m-1}) .$$

Direct calculations show that the relation

$$aD(y_1, \ldots, y_k) = aT(y_1, \ldots, y_k),$$

holds in $Ass^*[X]$, where $1 \le k \le n+1$; $a, y_1, \ldots, y_k \in Ass^*[X]$ and $\deg(a) = n$. The next theorem gives an infinite series of relations that hold for all words in $Ass^*[X]$ of length n, $n \in \mathbb{N}$ and are not consequences of the right-symmetric identity.

Theorem 3. The following relations

$$aD(y_1, \ldots, y_{n+1}) * y_{n+2} = \sum_{i=1}^{n+1} aD(y_1,...,y_i * y_{n+2},...,x_{n+1}),$$

holds in $Ass^*[X]$ for all $a, y_1, \ldots, y_{n+2} \in Ass^*[X]$ and $\deg(a) = n$.

6. ACKNOWLEDGEMENTS

REFERENCES

1. T. H. Morgan: *A critique of the theory of evolution*. Princeton, NJ: Princeton Univ. Press 1916.
2. P. Jordan : *Uber Verallgemeinerung smoglichkeiten des Formalismus quanten Mechanik*. Nachr. Ges. Wiss., Gottingen, 1933. S. 209–214.
3. A. I. Malcev: *Algebraic systems*. Springer-Verlag, 1973, ISBN 0-387-05792-7.
4. M. R. Bremner: *Jordan algebras arising from intermolecular recombination*. SIGSAM Bulletin (Communications in Computer Algebra). 2005. V. 39, N. 4. P. 106-117.
5. S. R. Sverchkov: *Structure and representation of Jordan algebras arising from intermolecular recombination* . Contemp. Math. AMS. 2009. V. 483. P. 261–285.
6. S.R. Sverchkov: *Basis for identities of the algebra of simplified insertion*, Sib. Math. J. 2009. V.50. N.6.
7. M. R. Bremner: *Polynomial identities for ternary intermolecular recombination*, To appear in Discrete and Continuous Dynamical Systems (accepted 28 September 2009).
8. L. Landweber, L. Kary: *The evolution on cellular computing: natire's solution to a computational problem*. Biosystems. 1999. V. 52. P. 3–13.
9. L. Kari: *DNA computing: Arrival of biological mathematics*. Math. Intelligencer. 1997. V. 19. N 2. P. 9–22.
10. N. Jacobson: *Lie algebras*. New York: Interscience. 1962.
11. P. Cohn: *On homomorphic images of special Jordan algebras*. Canad. J. Math. 1954. V. 6. P. 253–264.
12. D. P. Robbins: *Jordan elements in a free associative algebra. I* .J. of Algebra. 1971. V. 19.P. 354–378.
13. M. R. Bremner: *DNA computing, insertion of words and left-symmetric algebras*. Proc. of the Maple Conf. 2005.Maple Inc. Waterloo. 2005.
14. M. Daley, L. Kari, I. McQuillan: *Families of languages defined by ciliate bio-operations*. Theoret. Comput. Sci. 2004. V. 320. N. 1. P. 51–69.
15. M. Daley, L. Kari: *DNA computing: Models and implementations*. Comm. Theoret. Biol. 2002. V. 7. N. 3. P. 177–198.
16. M. Gerstenhaber: *The cohomology structure of an associative ring*. Ann. Math. 1963. V. 2. N. 78. P. 267–288.

17. K. A. Zhevlakov, A. M. Slinko, I. P. Shestakov., A. I. Shirshov: *Rings That Are Nearly Associative.* AcademicPress. New York. 1982.
18. A. Worz-Busekros: *Algebras in genetics*. Springer-Verl. New York.1980. Lecture notes in Biomath. N 36.
19. P. Holgate: *Genetic algebras satisfying Bernstein's stationary principle*. J. London Math. 1975. Soc. 2. N. 9. P. 613–623.
20. P. Holgate: *Jordan algebras arising in population genetics*. Proc. Edinburgh Math. Soc. 1967.V. 15. N 2, P.291–294.
21. A. Wors-Busekros: *Bernstein algebras*. Arch. Math. 1987. V. 48 P. 388–398.
22. U. Lubich: *Bernstein Algebras*. Uspehi Mat. Nauk 1977.V.32. (in Russian))N. 6. P. 261–262.
23. S. Walcher: *Bernstein algebras which are Jordan algebras*. Arch. Math. 1988.V. 50. P. 218–222.
24. S.R. Sverchkov: *Algebraic theory of DNA recombination*. Inter. Conference Mal'tsev Meeting. Sobolev Institute of Mathematics. Collection of Abstracts (in Russian). 2009. P. 132-133.

SVERCHKOV S.R.

NOVOSIBIRSK STATE UNIVERSITY, NOVOSIBIRSK, RUSSIA

E-mail address: SverchkovSR@yandex.ru