# Mathematical concepts of evolution algebras in non-Mendelian genetics

*Jianjun Paul Tian and Petr Vojtěchovský*

**Abstract**

Evolution algebras are not necessarily associative algebras satisfying $e_ie_j = 0$ whenever $e_i$, $e_j$ are two distinct basis elements. They mimic the self-reproduction of alleles in non-Mendelian genetics. We present elementary mathematical properties of evolution algebras that are of importance from the biological point of view.

Several models of Mendelian [2, 4, 12, 6, 8, 11] and non-Mendelian genetics [1, 5] exist. Based on the self-reproduction rule of non-Mendelian genetics [1, 7], the first author introduced a new type of algebra [10], called *evolution algebra*. In this paper we discuss some basic properties of evolution algebras.

## 1. Evolution algebras and subalgebras

Let $K$ be a field. A vector space $E$ over $K$ equipped with multiplication is an *algebra* (not necessarily associative) if $u(v+w) = uv + uw$, $(u+v)w = uw + vw$, $(\alpha u)v = \alpha(uv) = u(\alpha v)$ for every $u$, $v$, $w \in E$ and $\alpha \in K$.

Let $\{e_i;\ i \in I\}$ be a basis of an algebra $E$. Then $e_ie_j = \sum_{k\in I} a_{ijk}e_k$ for some $a_{ijk} \in K$, where only finitely many *structure constants* $a_{ijk}$ are nonzero for a fixed $i$, $j \in I$. The multiplication in $E$ is fully determined by the structure constants $a_{ijk}$, thanks to the distributive laws.

2000 Mathematics Subject Classification: 17A20, 92C15, 20N05, 37N25
Keywords: Nonassociative algebra, evolution algebra, non-Mendelian genetics, allele, self-reproduction
J. P. Tian would like to acknowledge the support from the National Science Foundation upon agreement No. 0112050

Let $E$ be an algebra. Then $F \subseteq E$ is a *subalgebra* of $E$ if $F$ is a subspace of $E$ closed under multiplication.

It is not difficult to show that the intersection of subalgebras is a subalgebra. Thus, given a subset $S$ of $E$, there is the smallest subalgebra of $E$ containing $S$. We call it the *subalgebra generated by* $S$, and denote it by $\langle S \rangle$. As usual:

**Lemma 1.1.** *Let $S$ be a subset of an algebra $E$. Then $\langle S \rangle$ consists of all elements of the form $\alpha_1(s_{1,1} \cdots s_{1,m_1}) + \cdots + \alpha_k(s_{k,1} \cdots s_{k,m_k})$, where $k \geqslant 1$, $m_i \geqslant 0$, $s_{i,j} \in S$, $\alpha_i \in K$, and where the product $s_{i,1} \cdots s_{i,m_i}$ is parenthesized in some way.*

An *ideal* $I$ of an algebra $E$ is a subalgebra of $E$ satisfying $I \cdot E \subseteq I$, $E \cdot I \subseteq I$. Clearly, 0 and $E$ are ideals of $E$, called *improper* ideals. All other ideals are *proper*. An algebra is *simple* if it has no proper ideals.

An *evolution algebra* is a finite-dimensional algebra $E$ over $K$ with basis $\{e_1, \ldots, e_v\}$ such that $a_{ijk} = 0$ whenever $i \neq j$. Upon renaming the structure constants we can write $e_i e_i = \sum_{j=1}^{v} a_{ij} e_j$. We refer to $\{e_1, \ldots, e_v\}$ as the *natural basis* of an algebra $E$. An evolution algebra is *nondegenerate* if $e_i e_i \neq 0$ for every $i$. *Throughout the paper we will assume that evolution algebras are nondegenerate.*

The multiplication in an evolution algebra is supposed to mimic self-reproduction of non-Mendelian genetics. We think and speak of the generators $e_i$ as *alleles*. The rule $e_i e_j = 0$ for $i \neq j$ is then natural, and the rule $e_i e_i = \sum a_{ij} e_j$ can be interpreted as follows: $a_{ij}$ is the probability that $e_i$ becomes $e_j$ in the next generation, and thus $\sum a_{ij} e_j$ is the superposition of the possible states. Nevertheless, we will develop much of the theory over arbitrary fields and with no (probabilistic) restrictions on the structure constants $a_{ij}$.

Given two elements

$$x = \sum_{i=1}^{v} \alpha_i e_i, \qquad y = \sum_{i=1} \beta_i e_i,$$

of an evolution algebra, we have

$$\begin{aligned} xy &= \sum_{i=1}^{v} \alpha_i e_i \cdot \sum_{j=1}^{v} \beta_j e_j = \sum_{i=1}^{v} \alpha_i \beta_i e_i^2 \\ &= \sum_{i=1}^{v} \left( \alpha_i \beta_i \sum_{j=1}^{v} a_{ij} e_j \right) = \sum_{j=1}^{v} \left( \sum_{i=1}^{v} \alpha_i \beta_i a_{ij} \right) e_j, \end{aligned}$$

a formula we will use without reference.

The natural basis of an evolution algebra plays a privileged role among all other bases, since the generators $e_i$ represent alleles. Importantly, the natural basis is privileged for purely mathematical reasons, too. The following example illustrates this point:

**Example 1.2.** Let $E$ be an evolution algebra with basis $\{e_1, e_2, e_3\}$ and multiplication defined by $e_1e_1 = e_1 + e_2$, $e_2e_2 = -e_1 - e_2$, $e_3e_3 = -e_2 + e_3$. Let $u_1 = e_1 + e_2$, $u_2 = e_1 + e_3$. Then $(\alpha u_1 + \beta u_2)(\gamma u_1 + \delta u_2) = \alpha\gamma u_1^2 + (\alpha\delta + \beta\gamma)u_1u_2 + \beta\delta u_2^2 = (\alpha\delta + \beta\gamma)u_1 + \beta\delta u_2$. Hence $F = Ku_1 + Ku_2$ is a subalgebra of $E$. However, $F$ is not an evolution algebra:

Let $\{v_1, v_2\}$ be a basis of $F$. Then $v_1 = \alpha u_1 + \beta u_2$, $v_2 = \gamma u_1 + \delta u_2$ for some $\alpha$, $\beta$, $\gamma$, $\delta \in K$ such that $D = \alpha\delta - \beta\gamma \neq 0$. By the above calculation, $v_1v_2 = (\alpha\delta + \beta\gamma)u_1 + \beta\delta u_2$. Assume that $v_1v_2 = 0$. Then $\beta\delta = 0$ and $\alpha\delta + \beta\gamma = 0$. If $\beta = 0$, we have $\alpha\delta = 0$. But then $D = 0$, a contradiction. If $\delta = 0$, we reach the same contradiction. Hence $v_1v_2 \neq 0$, and $F$ is not an evolution algebra.

We have just seen that evolution algebras are not closed under subalgebras. We therefore say that a subalgebra $F$ of an evolution algebra $E$ with basis $\{e_1, \ldots, e_v\}$ is an *evolution subalgebra* if, as a vector space, it is spanned by $\{e_i;\ i \in I\}$ for some $I \subseteq \{1, \ldots, v\}$. The subset $I$ determines $F$ uniquely, and we write $F = E(I) = \{\sum_{i=1}^{v} \alpha_i e_i;\ \alpha_i = 0 \text{ when } i \notin I\}$.

Similarly, we define an *evolution ideal* as an ideal $I$ of $E$ that happens to be an evolution subalgebra. This concept is superfluous, however:

**Lemma 1.3.** *Every evolution subalgebra is an evolution ideal.*

*Proof.* Let $F = E(I)$ be an evolution subalgebra . Let $x = \sum_{i \in I} \alpha_i e_i$ be an element of $F$ and $e_j$ an allele. We need to show that $xe_j \in F$. When $j \notin I$ then $xe_j = 0 \in F$. Assume that $j \in I$. Since $F$ is an evolution subalgebra, $e_i \in F$ for every $i \in I$. Then $xe_j = \alpha_j e_j^2 \in F$, since $F$ is a subalgebra. $\square$

Not every ideal of an evolution algebra is an evolution ideal:

**Example 1.4.** *Let $E$ be generated by $e_1$, $e_2$, where $e_1e_1 = e_1 + e_2 = e_2e_2$. Then $K(e_1 + e_2)$ is an ideal of $E$, but not an evolution subalgebra.*

An evolution algebra is *evolutionary simple* if if has no proper evolution ideals (evolution subalgebras).

Clearly, every simple evolution algebra is evolutionary simple. The converse it not true, as is apparent from Example 1.4.

The following theorem gives some basic properties of evolution algebras, all easy to prove (or see [10]). Recall that an algebra is *flexible* if it satisfies $x(yx) = (xy)x$.

**Theorem 1.5.** *Evolution algebras are commutative (and hence flexible), but not necessarily power-associative (hence not necessarily associative). Direct products and direct sums of evolution algebras are evolution algebras. Evolution subalgebra of an evolution algebra is an evolution algebra.*

An algebra is *real* if $K = \mathbb{R}$. An evolution algebra is *nonnegative* if it is real and all structure constants $a_{ij}$ are nonnegative. A *Markov evolution algebra* is a nonnegative evolution algebra such that $\sum_j a_{ij} = 1$ for every $1 \leqslant i \leqslant v$.

When $E$ is a real algebra, let $E^+ = \{\sum \alpha_i e_i;\ \alpha_i \geqslant 0\}$.

**Lemma 1.6.** *Let $E$ be a nonnegative evolution algebra. Then $E^+$ is closed under addition, multiplication, and multiplication by positive scalars.*

*Proof.* Let $x = \sum \alpha_i e_i$, $y = \sum \beta_i e_i \in E^+$. Then $x + y = \sum(\alpha_i + \beta_i)e_i$ clearly belongs to $E^+$. Moreover, $xy = \sum_j \left(\sum_i \alpha_i \beta_i a_{ij}\right) e_j \in E^+$, since $\alpha_i$, $\beta_i$, $a_{ij} \geqslant 0$ for every $i$, $j$. It is clear that $E^+$ is closed under multiplication by nonnegative scalars. □

## 2. The evolution operator

Let $E$ be an evolution algebra with basis $\{e_1, \ldots, e_v\}$. Since we are mainly interested in self-reproduction, we focus on the *evolution operator* $\Lambda : E \to E$, which is the (unique) linear extension of the map $e_i \mapsto e_i^2$.

**Lemma 2.1.** *Let $E$ be an evolution algebra and $x = \sum \alpha_i e_i$. Then $\Lambda(x) = x^2$, i.e., $\sum \alpha_i^2 e_i^2 = (\sum \alpha_i e_i)^2$.*

*Proof.* This is an immediate consequence of the fact that $e_i e_j = 0$ whenever $i \neq j$. □

When $E$ is a real evolution algebra, we can equip it with the usual $L_1$ norm, i.e., $\|\sum \alpha_i e_i\| = \sum |\alpha_i|$. Since $E$ is then isomorphic to $\mathbb{R}^v$ as a vector space, it becomes a complete vector space with respect to the metric $d(x, y) = \|x - y\|$. In other words, $E$ is a Banach space.

Since $v < \infty$, all linear operators defined on $E$ are continuous. In particular, every *left translation by $z$*, defined by $L_z(x) = zx$, is a continuous operator on $E$. However, due to the lack of associativity, the composition of two left translations does not have to be a left translation.

A (not-necessarily associative) *Banach algebra* is an algebra that is also a Banach space with norm $\|\cdot\|$ satisfying $\|xy\| \leqslant \|x\| \cdot \|y\|$. Not every evolution algebra is a Banach algebra. However:

**Lemma 2.2.** *Let $E$ be a real evolution algebra such that $\sum_j |a_{ij}| \leqslant 1$ for every $i$ (eg. a Markov evolution algebra). Then $E$ is a Banach algebra.*

*Proof.* Let $x = \sum_i \alpha_i e_i$, $y = \sum_i \beta_i e_i$. Then $\|x\| \cdot \|y\| = \sum_i |\alpha_i| \cdot \sum_j |\beta_j|$. On the other hand,

$$\begin{aligned}\|xy\| &= \left\| \sum_j \left( \sum_i \alpha_i \beta_i a_{ij} \right) e_j \right\| = \sum_j \left| \sum_i \alpha_i \beta_i a_{ij} \right| \leqslant \sum_j \sum_i (|\alpha_i \beta_i| \cdot |a_{ij}|) \\ &= \sum_i \left( \sum_j |a_{ij}| \right) |\alpha_i \beta_i| \leqslant \sum_i |\alpha_i \beta_i|,\end{aligned}$$

and the needed inequality follows. □

Note that even in the case of a Markov evolution algebra we never have $\|xy\| = \|x\| \cdot \|y\|$ for every $x$, $y$, as long as $v > 1$. For instance, $\|e_i e_j\| = 0 < 1 = \|e_i\| \cdot \|e_j\|$ when $i \neq j$.

Given $x$ in an algebra $E$, we define the *plenary powers* of $x$ by $x^{[0]} = x$, $x^{[n+1]} = x^{[n]} x^{[n]}$. Equivalently, we can set $x^{[n]}$ equal to $\Lambda^n(x)$ for any $n \geqslant 0$.

Recall that composition of maps is an associative binary operation. Thus:

**Lemma 2.3.** *Let $E$ be an algebra, $x \in E$, $\alpha \in K$, and $n$, $m \geqslant 0$. Then:*

(i) $(x^{[n]})^{[m]} = x^{[n+m]}$,

(ii) $(\alpha x)^{[n]} = \alpha^{(2^n)} x^{[n]}$.

*Proof.* It remains to prove (ii), which is easy by an induction on $n$. □

## 3. Occurrence relation

The question we are most interested in is the following: *When does the allele $e_i$ give rise to the allele $e_j$?* The phrase *give rise* can be interpreted in two ways: (i) the self-reproduction of $e_i$ yields $e_j$ with nonzero probability after a given number of generations, or (ii) the self-reproduction of $e_i$ yields $e_j$ with nonzero probability after some number of generations.

The first interpretation is studied below, while the second interpretation is investigated later, starting with Section 5.

Let $E$ be an algebra with basis $\{e_1, \dots, e_v\}$. We say that $e_i$ *occurs* in $x \in E$ if the coefficient $\alpha_i \in K$ is nonzero in $x = \sum_{j=1}^{v} \alpha_j e_j$. When $e_i$ occurs in $x$ we write $e_i \prec x$.

**Lemma 3.1.** *Let $E$ be a nonnegative evolution algebra. Then for every $x$, $y \in E^+$ and $n \geqslant 0$ there is $z \in E^+$ such that $(x+y)^{[n]} = x^{[n]} + z$.*

*Proof.* We proceed by induction on $n$. We have $(x+y)^{[0]} = x+y = x^{[0]}+y$, and it suffices to set $z = y$. Also, $(x+y)^{[1]} = (x+y)(x+y) = x^{[1]}+2xy+y^2$. By Lemma 1.6, $2xy + y^2 = z$ belongs to $E^+$.

Assume the claim is true for some $n \geqslant 1$. In particular, given $x$, $y \in E^+$, let $w \in E^+$ be such that $(x+y)^{[n]} = x^{[n]} + w$. Then $(x+y)^{[n+1]} = ((x+y)^{[n]})^{[1]} = (x^{[n]} + w)^{[1]}$. Since $w \in E^+$ and $x^{[n]} \in E^+$ by Lemma 1.6, we have $(x^{[n]} + w)^{[1]} = (x^{[n]})^{[1]} + z = x^{[n+1]} + z$ for some $z \in E^+$. □

**Proposition 3.2.** *Let $E$ be a nonnegative evolution algebra. When $e_i \prec e_j^{[n]}$ and $e_j \prec e_k^{[m]}$ then $e_i \prec e_k^{[n+m]}$.*

*Proof.* We have $e_k^{[m]} = \alpha_j e_j + y$ for some $\alpha_j \neq 0$ and $y \in E$ such that $e_j \not\prec y$. Moreover, by Lemma 1.6, we have $\alpha_j > 0$ and $y \in E^+$. By Lemma 3.1, $e_k^{[n+m]} = (e_k^{[m]})^{[n]} = (\alpha_j e_j + y)^{[n]} = (\alpha_j e_j)^{[n]} + z = \alpha_j^{(2^n)} e_j^{[n]} + z$ for some $z \in E^+$. Now, $e_j^{[n]} = \beta_i e_i + v$ for some $\beta_i > 0$ and $v \in E$ satisfying $e_i \not\prec v$. We therefore conclude that $e_i \prec e_k^{[n+m]}$ □

The proposition does not generalize to all evolution algebras, as the following example shows:

**Example 3.3.** Let $E$ be an evolution algebra with basis $\{e_i;\ 1 \leqslant i \leqslant 7\}$ such that $e_1e_1 = e_1$, $e_2e_2 = e_4$, $e_3e_3 = e_5 + e_6$, $e_4e_4 = e_1$, $e_5e_5 = e_2$, $e_6e_6 = e_7$, $e_7e_7 = -e_4$. Then $e_2^{[1]} = e_2e_2 = e_4$, $e_2^{[2]} = e_4e_4 = e_1$. Thus $e_1 \prec e_2^{[2]}$. Also, $e_3^{[1]} = e_3e_3 = e_5 + e_6$, $e_3^{[2]} = (e_5 + e_6)^2 = e_5^2 + e_6^2 = e_2 + e_7$. Thus $e_2 \prec e_3^{[2]}$. However, $e_3^{[3]} = (e_2 + e_7)^2 = e_2^2 + e_7^2 = e_4 - e_4 = 0$, and so $e_3^{[n]} = 0$ for every $n \geqslant 3$. This means that $e_1 \not\prec e_3^{[n]}$ for any $n \geqslant 0$.

## 4. Occurrence sets

Let $e_i$, $e_j$ be two alleles of an evolution algebra. Then the *occurrence set of* $e_i$ *with respect to* $e_j$ is the set $O_{i,j} = \{n > 0;\ e_i \prec e_j^{[n]}\}$.

Recall that a *semigroup* is a set with one binary operation that satisfies the associative law. When $E$ is a nonnegative evolution algebra, every occurrence set $O_{i,i}$ is a subsemigroup of $(\{1, 2, \dots\}, +)$ by Proposition 3.2.

The goal of this section is to show that any finite subset of $\{1, 2, \dots\}$ can be realized as an occurrence set of some evolution algebra, and that every subsemigroup of $(\{1, 2, \dots\}, +)$ can be realized as an occurrence set of some nonnegative evolution algebra. Hence the occurrence sets are as rich as one could hope for.

**Example 4.1.** *Let $n > 1$. Consider the evolution algebra $E$ with generators $\{e_1, \dots, e_{n+1}\}$ defined by $e_1e_1 = e_2$, $e_2e_2 = e_3$, $\dots$, $e_{n-1}e_{n-1} = e_n$, $e_ne_n = e_1 + e_{n+1}$, $e_{n+1}e_{n+1} = -e_2$. Then $e_1^{[m]} = e_{m+1}$ for every $1 \leqslant m < n$, $e_1^{[n]} = e_1 + e_{n+1}$, and $e_1^{[m]} = 0$ for every $m > n$. Thus $O_{1,1} = \{n\}$.*

**Lemma 4.2.** *Let $S$ be a finite subset of $\{1, 2, \dots\}$. Then there is an evolution algebra $E$ such that $O_{1,1} = S$.*

*Proof.* Let $S = \{n_1, \dots, n_m\}$. In the following calculations we label basis elements of $E$ also by $e_{i,j}$; these can be relabeled as $e_i$ at the end.

Let $e_1e_1 = e_{2,1} + \cdots + e_{2,m}$. Given $1 \leqslant i \leqslant m$, let $e_{2,i}e_{2,i} = e_{3,i}$, $e_{3,i}e_{3,i} = e_{4,i}$, $\dots$, $e_{n_i,i}e_{n_i,i} = e_1 + e_{n_i+1,i}$, $e_{n_i+1,i}e_{n_i+1,i} = -e_1e_1$. Thus, roughly speaking, we imitate Example 4.1 for every $1 \leqslant i \leqslant m$. It is now not hard to see that $O_{1,1} = S$. □

A semigroup is *finitely generated* if it is generated by a finite subset. Here is a well-known fact:

**Lemma 4.3.** *Every subsemigroup of $(\{1, 2, \dots\}, +)$ is finitely generated.*

*Proof.* Let $S$ be a subsemigroup of $(\{1, 2, \dots\}, +)$. Let $n$ be the smallest element of $S$. For every $1 \leqslant i < n$ let $m_i$ be the smallest element of $S$ such that $m_i$ is congruent to $i$ modulo $n$, if such an element exists, else set $m_i = n$. We claim that $A = \{n, m_1, \dots, m_{n-1}\}$ generates $S$. Suppose that this is not the case and let $s$ be the smallest element of $S$ not generated by $A$. Since $s$ cannot be a multiple of $n$, there is $1 \leqslant i < n$ such that $s$ is congruent to $i$ modulo $n$. Then $m_i \neq n$ and $m_i < s$. But then $s = m_i + kn$ for some $k > 0$, so $s \in A$, a contradiction. □

**Lemma 4.4.** *Let $S$ be a subsemigroup of $(\{1, 2, \dots\}, +)$. Then there is a nonnegative evolution algebra $E$ such that $O_{1,1} = S$.*

*Proof.* Assume that $S$ is 1-generated, i.e., that $S = \{n, 2n, \dots\}$ for some $n \geqslant 1$. Then define $E$ by: $e_1e_1 = e_2$, $e_2e_2 = e_3$, $\dots$, $e_{n-1}e_{n-1} = e_n$, $e_ne_n = e_1$. It is easy to see that $O_{1,1} = S$.

When $S$ is generated by $m$ elements, say $n_1, \dots, n_m$, we can use a similar trick as in the proof of Lemma 4.2.

Every subsemigroup of $(\{1, 2, \dots\}, +)$ is finitely generated by Lemma 4.3. □

**Problem 4.5.** *Can any subset of $\{1, 2, \dots\}$ be realized as an occurrence set of some evolution algebra?*

**Problem 4.6.** *Let $S$ be a subset of $\{1, 2, \dots\}$, $|S| = n$. What is the smallest integer $v$ such that there is an evolution algebra $E$ of dimension $v$ for which $S$ is an occurrence set?*

## 5. Occurrence based on evolution subalgebras

We are now going to look at the second interpretation of "$e_i$ gives rise to $e_j$."

**Lemma 5.1.** *Intersection of evolution subalgebras is an evolution subalgebra.*

*Proof.* Let $F = E(I)$, $G = E(J)$ be two evolution subalgebras of $E$. Then $F \cap G = E(I \cap J)$ as a vector space. Since $F \cap G$ is a subalgebra, we are done. □

Thus for any subset $S$ of $E$ there exists the smallest evolution subalgebra of $E$ containing $S$, and we denote it by $\langle\langle S\rangle\rangle$. The notation is supposed to suggest that the evolution subalgebra generated by $S$ can be larger than the subalgebra generated by $S$.

We now define another occurrence relation as follows: For $x, y \in E$, let $x \ll y$ if $x \in \langle\langle y\rangle\rangle$.

**Lemma 5.2.** *For $x, y, z \in E$ we have:*

(*i*) *if $x \ll y$ and $y \ll x$ then $\langle\langle x\rangle\rangle = \langle\langle y\rangle\rangle$,*

(*ii*) *if $x \ll y$ and $y \ll z$ then $x \ll z$,* (*iii*) *if* $x \ll y^{[n]}$ *for some* $n \geqslant 0$ *then* $x \ll y$.

*Proof.* Easy. □

In view of Lemma 5.2(*iii*), it makes no sense to speak of occurrence sets (analogous to $O_{i,j}$) in the context of $\ll$, since every occurrence set would be either empty or would consists of all nonnegative integers.

**Lemma 5.3.** *Let* $F$, $G$ *be evolutionary simple evolution subalgebras of* $E$. *Then either* $F = G$ *or* $F \cap G = 0$.

*Proof.* Assume that there is $x \in F \cap G$, $x \neq 0$. Then $\langle\langle x \rangle\rangle$ is an evolution subalgebra of both $F$ and $G$. Since both $F$, $G$ are evolutionary simple, it follows that $F = G = \langle\langle x \rangle\rangle$. □

## 6. Algebraically persistent and transient generators

A generator $e_i$ of an evolution algebra $E$ is *algebraically persistent* if $\langle\langle e_i \rangle\rangle$ is evolutionary simple, else it is *algebraically transient.*

**Lemma 6.1.** *If* $E$ *is an evolutionary simple evolution algebra then it has no algebraically transient generators.*

*Proof.* Assume that $e_i$ is an algebraically transient generator, i.e., that $\langle\langle e_i \rangle\rangle$ is not evolutionary simple. If $E = \langle\langle e_i \rangle\rangle$, we see right away that $E$ is not evolutionary simple. If $\langle\langle e_i \rangle\rangle$ is a proper evolution subalgebra of $E$ then it is a proper evolution ideal of $E$ by Lemma 1.3, and $E$ is not evolutionary simple. □

The following example shows that the converse of Lemma 6.1 does not hold (but see Corollary 7.3):

**Example 6.2.** Let $E$ have generators $e_1$, $e_2$ such that $e_1e_1 = e_1$, $e_2e_2 = e_2$. Then $\langle\langle e_1 \rangle\rangle = Ke_1$, $\langle\langle e_2 \rangle\rangle = Ke_2$, which means that both $e_1$, $e_2$ are algebraically persistent. Yet $\langle\langle e_i \rangle\rangle$ is a proper evolution ideal of $E$, and hence $E$ is not evolutionary simple.

**Lemma 6.3.** *Let* $e_i$ *be an algebraically persistent generator of* $E$, *and assume that* $e_j \prec e_ie_i$. *Then* $e_j$ *is algebraically persistent.*

*Proof.* Since $e_j \prec e_ie_i$, we have $\langle\langle e_i \rangle\rangle \supseteq \langle\langle e_j \rangle\rangle$. But $\langle\langle e_i \rangle\rangle$ is evolutionary simple, thus $\langle\langle e_i \rangle\rangle = \langle\langle e_j \rangle\rangle$. Then $\langle\langle e_j \rangle\rangle$ is evolutionary simple, and thus $e_j$ is algebraically persistent. □

# 7. Decomposition of evolution algebras

An evolution algebra $E$ is *indecomposable* if whenever $E = F \oplus G$ for some evolution subalgebras $F$, $G$ of $E$, we have $F = 0$ or $G = 0$. An easy induction proves that every evolution algebra can be written as a direct sum of indecomposable evolution algebras.

Here is an indecomposable evolution algebra that is not evolutionary simple:

**Example 7.1.** Let $E$ be generated by $e_1$, $e_2$, where $e_1e_1 = e_1$, $e_2e_2 = e_1$. Then $\langle\langle e_1\rangle\rangle = Ke_1$, $\langle\langle e_2\rangle\rangle = E$.

An evolution algebra $E$ is *evolutionary semisimple* if it is a direct sum of some of its evolutionary simple evolution subalgebras. Note that every evolutionary simple evolution subalgebra of $E$ can be written as $\langle\langle e_i\rangle\rangle$ for some algebraically persistent generator of $E$.

**Proposition 7.2.** *An evolution algebra $E$ is evolutionary semisimple if and only if all of its alleles $e_i$ are algebraically persistent.*

*Proof.* Assume that $E$ is evolutionary semisimple, and write $E = \langle\langle e_{i_1}\rangle\rangle \oplus \cdots \oplus \langle\langle e_{i_n}\rangle\rangle$, where each $e_{i_j}$ is algebraically persistent. Let $e_j$ be an allele of $E$. Then $e_j$ belongs to some $\langle\langle e_{i_k}\rangle\rangle$. Since $\langle\langle e_j\rangle\rangle$ is an evolution ideal of $\langle\langle e_{i_k}\rangle\rangle$ and $e_{i_k}$ is algebraically persistent, we conclude that $\langle\langle e_j\rangle\rangle = \langle\langle e_{i_k}\rangle\rangle$. Thus $e_j$ is algebraically persistent, too.

Conversely, assume that every allele of $E$ is algebraically persistent. For each $e_i$ let $I_i = \{j;\ e_j \ll e_i\}$. Given $i \neq j$, we have either $I_i = I_j$ or $I_i \cap I_j = \emptyset$, by Lemma 5.3. Thus there exists $\{i_1, \ldots, i_n\} \subseteq \{1, \ldots, v\} = I$ such that $I_{i_1} \cup \cdots \cup I_{i_n} = I$, and the union is disjoint. In other words, $E = \langle\langle e_{i_1}\rangle\rangle \oplus \cdots \oplus \langle\langle e_{i_n}\rangle\rangle$. □

Here is a partial converse of Lemma 6.1:

**Corollary 7.3.** *An indecomposable evolution algebra with no transient generators is evolutionary simple.*

Let $E$ be an evolution algebra. Partition $\{1, \ldots, v\}$ as $I \cup J$, where $e_i \in I$ if and only if $e_i$ is an algebraically persistent generator of $E$. Let $P(E) = \{\sum \alpha_i e_i;\ \alpha_i = 0 \text{ for } i \notin I\}$, and $T(E) = \{\sum \alpha_i e_i;\ \alpha_i = 0 \text{ for } i \notin J\}$.

**Lemma 7.4.** *$P(E)$ is an evolutionary semisimple evolution subalgebra of $E$.*

*Proof.* We first show that $P(E)$ is an evolution subalgebra. Let $x \in P(E)$, $y \in P(E)$, $x = \sum_{i\in I} \alpha_i e_i$, $y = \sum_{i \in I} \beta_i e_i$, where $I$ is as above. Then $xy = \sum_{i\in I} \alpha_i \beta_i e_i^2$. By Lemma 6.3, $e_i^2$ is a linear combination of algebraically persistent generators, and hence $xy \in P(E)$.

Then $P(E)$ is evolutionary semisimple by Proposition 7.2. □

Observe:

**Lemma 7.5.** *Let $E(I)$, $E(J)$ be evolution subalgebras of $E$ such that $E(I)$ is a subalgebra of $E(J)$. Then $I \subseteq J$. If $E(I)$ is a proper subalgebra of $E(J)$, then $I$ is a proper subset of $J$.*

Thus:

**Lemma 7.6.** *Every evolution algebra $E$ has an evolutionary simple evolution subalgebra. In particular, $P(E) \neq 0$.*

*Proof.* We proceed by induction on $v$. If $v = 1$, then $E = \langle\langle e_1 \rangle\rangle$ is evolutionary simple. Assume that the lemma is true for $v - 1$. If $E = E(\{1, \ldots, v\})$ is evolutionary simple, we are done. Else, by Lemma 7.5, there is a proper subset $I$ of $\{1, \ldots, v\}$ such that $E(I)$ is a proper evolution subalgebra. By induction, $E(I)$ contains an evolutionary simple evolution subalgebra. □

Every evolution algebra $E$ decomposes as a vector space into $P(E) \oplus T(E)$, and $P(E) \neq 0$, by the above lemma. Moreover, $P(E)$ is an evolutionary semisimple evolution algebra, and can therefore be written as a direct sum of evolutionary simple evolution algebras $\langle\langle e_{i_j} \rangle\rangle$.

However, the subspace $T(E)$ does not need to be a subalgebra of $E$, hence it does not need to be an evolution algebra. But we can make it into an evolution algebra:

Let $T(E) = \{\sum \alpha_i e_i;\ \alpha_i = 0 \text{ for } i \notin J\}$. Let $J^* = J \setminus \{j;\ e_j^2 \subseteq P(E)\}$. (This will guarantee that the resulting evolution algebra is nondegenerate.) Let $T^*(E)$ be defined on the subspace generated by $\{e_i;\ i \in J^*\}$ by $e_i e_i = \sum_{j \in J^*} a_{ij} e_j$, where the structure constants $a_{ij}$ are inherited from $E$. If $J^* \neq \emptyset$, then $T^*(E)$ is a nondegenerate evolution algebra. If $J^* = \emptyset$ then all algebraically transient generators of $E$ vanish after the first reproduction, and therefore have no impact, biologically speaking.

If $E_1 = T^*(E) \neq 0$, we can iterate the decomposition and form $P(E_1)$, $T(E_1)$ and $T^*(E_1)$, etc. Eventually we reach a point $n$ when $T^*(E_n) = 0$, i.e., every transient generator of $E_n$ disappears after the first generation.

Let us emphasize that the decomposition of $E$ thus obtained results in an evolution algebra not necessarily isomorphic to $E$; some information may be lost in the decomposition $P(E) \oplus T(E)$.

## References


[1] **C. W. Birky, Jr.**: *The inheritance of genes in mitochondria and chloroplasts: laws, mechanisms, and models*, Annu. Rev. Genet. **35** (2001), 125–148.

[2] **I. M. H. Etherington**: *Non-associative algebra and the symbolism of genetics*, Proc. Roy. Soc. Edinburgh B **61** (1941), 24–42.

[3] **R. Costa, A. Grishkov, H. Quzzo and L. Peresi (editors)**: Nonassociative algebra and its applications, the Fourth International Conference, Marcel Dekker, Inc., 2000.

[4] **P. Holgate**: *Selfing in gentic algebras*, J. Math. Biology **6** (1978), 197–206.

[5] **F. Ling and T. Shibata**: *Mhr1p-dependent concatemeric mitochondrial DNA formation for generating yeast mitochondrial homoplasmic cells*, Mol. Biol. Cell **15** (Jan. 2004), 310–322.

[6] **Y. I. Lyubich**: Mathematical structure in population genetics, Springer-Verlag, New York, 1992.

[7] **G. Mendel**: *Experiments in plant-hybridization*, Classic papers in Genetics, 1–20, J. A. Petr (editor), Prentice-Hall Inc. 1959.

[8] **M. L. Reed**: *Algebraic structure of genetic inheritance*, Bull. of Amer. Math. Soc **34** ((2)) (1997), 107–130.

[9] **R. Schafer**: An introduction to nonassociative algebra, Dover Publications, Inc., New York, 1994.

[10] **J. P. Tian**: Introduction to evolution algebras, in preparation.

[11] **J. Tian and B.-L. Li**: *Coalgebraic structure of genetic inheritance*, Mathematical Bioscience and Engineering, vol. **1**, **2** (2004).

[12] **A. Worz-Busekros**: Algebras in Genetics, Lecture Notes in Biomathematics **36**, Springer-Verlag, Berlin, 1980.


Received October 10, 2005


J. P. Tian
Mathematical Biosciences Institute
The Ohio State University
Columbus, OH 43210
U.S.A.
e-mail: tianjj@mbi.ohio-state.edu

P. Vojtěchovský
Department of Mathematics
University of Denver
Denver, CO 80208
U.S.A.
e-mail: petr@math.du.edu