Article

An Application of Total-Colored Graphs to Describe Mutations in Non-Mendelian Genetics

Raúl M. Falcón [1,*] , **Óscar J. Falcón** [2] **and Juan Núñez** [2]

1   Department of Applied Mathematics I, University of Seville, 41012 Sevilla, Spain
2   Department of Geometry and Topology, University of Seville, 41012 Sevilla, Spain;
     oscfalgan@yahoo.es (Ó.J.F.); jnvaldes@us.es (J.N.)
*   Correspondence: rafalgan@us.es

check for updates

Abstract: Any gene mutation during the mitotic cell cycle of a eukaryotic cell can be algebraically represented by an isotopism of the evolution algebra describing the genetic pattern of the inheritance process. We identify any such pattern with a total-colored graph so that any isotopism of the former is uniquely related to an isomorphism of the latter. This enables us to develop some results on graph theory in the context of the molecular processes that occur during the *S*-phase of a mitotic cell cycle. In particular, each monochromatic subset of edges is identified with a mutation or regulatory mechanism that relates any two statuses of the genotypes of a pair of chromatids.

1. Introduction

At present, non-associative algebras are considered an adequate theoretical framework to address important topics in genetics. In fact, there are so many non-associative algebras that have attracted the interest of geneticists that it would be difficult to make an exhaustive list of them. Let us mention, as an example, the following: Mendelian algebras, gametic algebras, zygotic algebras, baric algebras, train algebras, copular algebras, Bernstein algebras, and evolution algebras. These algebras, and generally all the algebras used to model inheritance in genetics, are referred to as "genetic algebras".

In particular, Jin Jung Paul Tian established in his PhD thesis in 2004, later published in 2008, the foundations of the theory of evolution algebras [1]. On page 3 of that thesis, he wrote the following words that in principle could seem like a premonition, but they have been actually confirmed by the numerous publications cited throughout this paper.

> General genetic algebras are the product of interaction between Biology and Mathematics. Mendelian Genetics introduced a new subject to Mathematics: general genetic algebras. The study of these algebras reveals algebraic structures of Mendelian Genetics, which always simplifies and shortens the way to understand genetic and evolutionary phenomena. Indeed, it is the interplay between the purely mathematical structure and the corresponding genetic properties that makes this area so fascinating... Now, non-Mendelian Genetics is a basic language of molecular geneticists. Logically, we can ask what non-Mendelian Genetics offers to Mathematics. The answer is "evolution algebras".

Indeed, evolution algebras were introduced [1–3] to model those probabilistic laws that regulate how the genotypes of a eukaryotic cell undergoing mitosis have influence on the genotypes of the

offspring. Specifically, an n-dimensional algebra over a field $\mathbb{K}$ is an evolution algebra if there exists a natural basis $\{e_1, \ldots, e_n\}$ such that $e_i e_j = 0$, for all $i \neq j$. Each basis vector represents a genotype of a given phenotype so that the role of self-replication of genotypes is played by the products $e_i e_i = \sum_{j=1}^{n} c_{ij} e_j$, for all $1 \leq i \leq n$. In the case of dealing with $\mathbb{K} = \mathbb{R}$ as the base field, if $\sum_{j=1}^{n} c_{ij} = 1$ and $c_{i,j} \geq 0$, for all i, j, then each structure constant c_{ij} represents the probability that the genotype e_i becomes e_j in the offspring. In any case, even if there exist some advances in the algebraic study of such a probabilistic meaning [4,5], evolution algebras are usually studied without probabilistic restrictions on their structure constants [6–16].

The n-tuple $(e_1 e_1, \ldots, e_n e_n)$ is called the genetic pattern of both the evolution algebra and the mitotic cell cycle with respect to the phenotype under study. The distribution into isotopism classes of two- and three-dimensional evolution algebras over any base field is already known [17,18]. This determines, in turn, the spectrum of genetic patterns of mitotic cell cycles with respect to a phenotype that is respectively associated with two and three genotypes. Being introduced by Albert [19] as a generalization of isomorphisms of algebras, the importance of dealing with isotopisms of algebras in genetics was brought to light by Holgate and Campos [20,21]. They realized how isotopisms can be used to formulate gene mutations in any inheritance process algebraically. See [22] for a recent survey on the theory of isotopisms.

The simplicity of describing any evolution algebra by means of its genetic pattern enables one to find interesting connections with graph theory, which was introduced into genetics by Benzer [23] to show the linear internal structure of genes and how this structure is affected by mutations with no detectable tendency to be reverted. More specifically, given a graph G with a finite set of vertices $\{v_1, \ldots, v_n\}$, Tian [1] defined the evolution algebra of natural basis $\{e_1, \ldots, e_n\}$ such that $e_i^2 = \sum_{j \in \Gamma(v_j)} e_j$, for all $1 \leq i \leq n$. Here, $\Gamma(v_j)$ denotes the set of neighbors of the vertex v_j in G. Such a definition induces a faithful functor among graphs and evolution algebras that maps any isomorphism between two finite graphs to an isomorphism between their related evolution algebras. Based on this fact, Tian contemplated as further work the translation of known results on graph theory to the language of evolution algebras. In this regard, for instance, Cadavid, Rodiño, and Rodríguez [5] recently dealt with the connection between the evolution algebra induced by a random walk on a graph and the evolution algebra derived from the same graph. In a subsequent paper [24], they also characterized the space of derivations of evolution algebras associated with finite graphs.

The reverse problem of describing a graph derived from a given evolution algebra has also been dealt with in the literature. In this regard, Núñez et al. [25] and Elduque and Labra [26,27] associated a weighted digraph to any given finite-dimensional evolution algebra and studied the relationship among them. In particular, in the last two contributions, it was proven that the nonexistence of oriented cycles in such a digraph involves the nilpotency of the corresponding algebra. Further, Cabrera et al. [28] described another directed graph whatever the dimension of the evolution algebra is, from which its annihilator and irreducibility may be determined. Unlike Tian's graphs, however, since all of these proposals depend on the basis of the algebra, isomorphic evolution algebras are not related in general to isomorphic graphs.

This paper delves into this last topic in the case of dealing with evolution algebras over finite fields. Specifically, we introduce a total-colored graph that can be associated with any given evolution algebra over a finite field so that any isomorphism of the former is uniquely related to an isotopism of the latter. The underlying idea behind the proposed graph derives from a previous work of the authors [29] in which a pair of colored graphs was introduced in order to describe faithful functors relating the category of finite-dimensional algebras over finite fields with the category of vertex-colored graphs. They were based in turn on a proposal of McKay et al. [30], who identified the isotopisms of Latin squares with isomorphisms of vertex-colored graphs. A first attempt to approach the graphs introduced in [29] to the theory of evolution algebras was carried out in [31], where a step-by-step construction of an edge-colored graph derived from the genetic pattern of an evolution algebra over a finite field was established. The total-colored graph here introduced not only simplifies this last

construction by reducing both the number of vertices and edges under consideration, but also facilitates a relationship among different concepts on evolution algebras, graph theory, and genetics.

The paper is organized as follows. Section 2 shows some preliminary concepts and results on the subjects that are used throughout the paper. Section 3 deals with the description of the total-colored graph that we associate with any given evolution algebra over a finite field. Finally, Section 4 establishes the existence of a faithful functor between both categories of evolution algebras and their total-colored graphs.

2. Preliminaries

This section deals with some basic concepts on genetics, genetic algebras, evolution algebras, graph theory, and isotopisms of algebras that we use throughout the paper. For more details about the last four topics, we refer the reader to [3,19,32,33], respectively. We also refer the reader to [34,35] for different applications of graph theory into genetics and molecular biology.

2.1. Genetics

It is well known that genes are the molecular units of hereditary information within every cell. Disposed in sequential order, they give rise to chromosomes, which are formed by a pair of long twisted strands of deoxyribonucleic acid (DNA) and contain the genetic code to synthesize the proteins that determine each attribute or phenotype of the organism. Genes of a given phenotype always appear in chromosomes at the same position and determine the genotype of the organism with respect to such a phenotype.

In every eukaryotic cell, DNA is mostly contained within its nucleus. The mitotic cell cycle is the series of phases dividing any such cell into two daughter cells with the same DNA. In particular, during the S-phase, both strands of DNA within a nuclear chromosome are separated and serve as templates to synthesize new complementary daughter strands. Up to rare gene mutation, this gives rise to two molecules of identical DNA, called chromatids, which are separated as independent chromosomes into two new nuclei. In turn, the extra-nuclear genetic material of the parent cell is randomly distributed between both daughter cells. Both processes constitute the M-phase of the mitotic cell cycle. During the transition from the S-phase to the M-phase, regulatory mechanisms are produced to check any DNA damage motivated by gene mutations. This stage is known as the G_2-phase.

2.2. Genetic Algebras

General genetic algebras were introduced in the 1920s and 1930s. It was Serebrowski [36] who was the first to interpret algebraically the sign $\times$, which indicated sexual reproduction, and to give a mathematical formulation of Mendel's laws, formulated by G.Mendel in his famous papers of 1865 and 1866 (see [37], for instance). Glivenkov [38] introduced the so-called Mendelian algebras for diploid populations with one locus or two unlinked loci. Independently, Kostitzin [39] also introduced a symbolic multiplication to express Mendel's laws. However, the systematic study of genetic algebras is attributed to Etherington.

Indeed, Mendel's laws were formulated in terms of non-associative algebras by Etherington in several papers in 1939–1941 (and others later), among them [40–42]. In these works, algebraic patterns outlined by Mendel in his famous papers of 1865 and 1866 were developed. Some improvements to the mathematical approach for formulating Mendel's laws were given by Jennings [43] in 1917, by Serebrovski[36] himself in 1934, and by Glivenkov [38] in 1936. Since then, many authors have made relevant contributions to genetics in the context of non-associative algebras: Schafer [44], Gonshör [45], Haldane [46], Holgate [21], Heuch [47], Reiersöl [48], Korol [49], Lyubich [50], Wörz-Busekros [51], and Bertrand [52]. Further information on these algebras can be checked in [51] and the references therein.

2.3. Evolution Algebras

Tian and Vojtechovsky [1,3] showed a complete study of the algebraic properties of an evolution algebra. Let E_1 and E_2 be evolution algebras, and let $B_1 = \{e_1, \ldots, e_n\}$ be a natural basis on E_1. A linear map $\phi : E_1 \mapsto E_2$ is said to be a homomorphism of evolution algebras if $\phi(a\,b) = \phi(a)\,\phi(b)$, for every $a, b \in E_1$. Moreover, if ϕ is bijective, then it is said to be an isomorphism of evolution algebras.

In [15], an evolution algebra was described by identifying the coefficients of inheritance of a bisexual population as the structure constants of the algebra. In this respect, an evolution algebra of a bisexual population is an evolution algebra with a set of females partitioned into finitely many different types and the males having only one type. The idea of treating in a separate way male and female components of the population is due to Etherington [40], and it was later formalized by Holgate [53].

Finally, it is convenient to note that there exist some types of algebras that are not properly evolution algebras in the sense of Tian, but are quite related to them. One of these types is that formed by evolution algebras of a chicken population. Another one is that formed by evolution algebras of a mosquito population (see [54,55], respectively, for further information).

2.4. Graph Theory

A graph is a pair $G = (V, E)$ formed by a set V of vertices and a set E of edges containing two vertices, which are said to be adjacent. Two edges sharing a common vertex are said to be incident. A graph is complete if any pair of its vertices is adjacent. A path between two vertices $v, w \in V$ is any alternating sequence of vertices and edges, $\{v_0, e_0, v_1, \ldots, v_{n-1}, e_{n-1}, v_n\}$, for which:

(a)　$v_0 = v$ and $v_n = w$;
(b)　each edge e_i contains both vertices v_i and v_{i+1}, for all $i \in \{0, \ldots, n-1\}$; and
(c)　all its vertices and all its edges are pairwise distinct.

The number of edges in any path constitutes its length. A graph is said to be connected if there exists a path between any pair of its vertices. The minimum length of a path between two given vertices $v, w \in V$ constitutes their distance $d(v, w)$. The eccentricity $\epsilon(v)$ of a vertex $v \in V$ is the maximum distance from v to any other vertex in V. The radius $\rho(G)$ and the diameter $\delta(G)$ of a graph G are, respectively, the minimum and maximum eccentricity of its vertices.

A graph is total-colored if all its vertices and edges are colored from a given palette of colors. This is proper-edge-colored if no two incident edges are colored with the same color. Two total-colored graphs that are related to the same palette of colors are isomorphic if there exists a bijection between their corresponding sets of vertices that preserves (a) their adjacency, (b) the color of edges, and (c) the fact that two given vertices are colored with the same color. A path between two vertices of a total-colored graph is said to be alternating if it contains at least three vertices and any two adjacent edges have different colors.

2.5. Isotopisms of Algebras

Two n-dimensional algebras A_1 and A_2 are isotopic if there exist three non-singular linear transformations f, g, and h from A_1 to A_2 such that $f(x)g(y) = h(xy)$, for all $x, y \in A_1$. The triple (f, g, h) is called an isotopism from A_1 to A_2. From a strict point of view, the third component h need not be defined for every vector of the algebra A_1, but only for its derived algebra $A_1^2 = \{xy \colon x, y \in A_1\}$. The notion of the isotopism of algebras was introduced by Albert [19] in 1942 as a generalization of the classical concept of the isomorphism of algebras (which arises when $f = g = h$). In particular, isotopisms make it possible to gather together non-isomorphic algebras that, at first sight, seemed to be completely different. Since the original manuscript of Albert, isotopisms have been widely used in the literature to study and classify distinct types of non-associative algebras like division algebras, alternative algebras, Jordan algebras, Lie algebras, Malcev algebras, genetic algebras, or evolution

algebras, with relevance not only in mathematics, but also in natural sciences and engineering. See [22] for a recent survey in this regard.

In the context of evolution algebras, the transformations f and g represent the gene mutation of each one of the two chromatids in which the nuclear genetic material duplicates during the S-phase of a mitotic cell cycle. On the other hand, the transformation h represents the regulatory mechanisms during the G_2-phase that check any possible DNA damage that is produced during the S-phase. More specifically, for each $\alpha \in \{f, g, h\}$, one has $\alpha(e_i) = \sum_{j=1}^{n} \alpha_{ij} e_j$, for all $i \leq n$, where α_{ij} represents the probability that the genotype e_i mutates (if $\alpha \in \{f, g\}$) or is regulated (if $\alpha = h$) to e_j.

Note that a mitotic cell cycle is accomplished only if the genotypes of both chromatids that have been created after the mutations related to the transformations f and g coincide. Otherwise, there is no offspring. To this end, the regulatory mechanism related to the transformation h is crucial.

3. Description of the Total-Colored Graph

In this section, we describe our proposal to associate a total-colored graph $\mathcal{G}(A)$ with any given n-dimensional evolution algebra A over a finite field $\mathbb{K}$. Let $\{e_1, \ldots, e_n\}$ be the natural basis of such an algebra, and suppose $e_i e_i = \sum_{j=1}^{n} c_{ij} e_j$, with $c_{ij} \in \mathbb{K}$, for all i, j. Since we are interested in mutations among different genotypes, we may suppose $n > 1$. Thus, let $\mathcal{G}(A) := (V(A), E(A))$ be the total-colored graph having:

1.   the set of vertices:

$$V(A) := \{v_{x,y} \mid x, y \in A, \ xy \neq 0\}, \tag{1}$$

where $v_{x,y}$ and $v_{x',y'}$ are equally colored in $V(A)$ if and only if $xy = x'y'$;

2.   and whose set of edges $E(A)$ is formed by the disjoint union of both monochromatic subsets:

$$E_1(A) := \{v_{x,y} v_{x,z} \mid x, y, z \in A, \ y \neq z\} \tag{2}$$

$$E_2(A) := \{v_{x,y} v_{z,y} \mid x, y, z \in A, \ x \neq z\} \tag{3}$$

In order to illustrate this definition, Figure 1 shows, for instance, the total-colored graphs of the three isotopism classes of non-trivial two-dimensional evolution algebra over the finite field $\mathbb{F}_2$ [17]. Colors of vertices are represented by the symbols $\bullet$, $\blacktriangle$, and $\blacktriangledown$, while colors of edges are represented by solid and dashed lines.
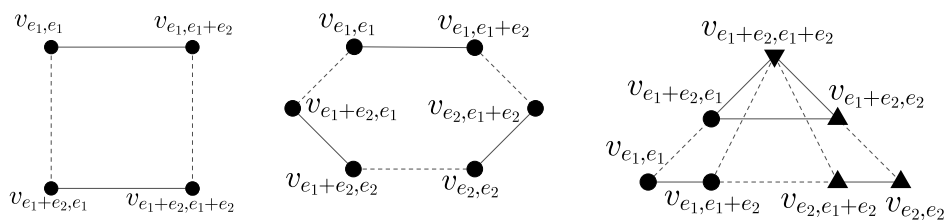
Figure 1. Total-colored graphs of the three two-dimensional evolution algebras over $\mathbb{F}_2$ having respective genetic patterns $(e_1, 0)$, (e_1, e_1), and (e_1, e_2).

The graph $\mathcal{G}(A)$ does not contain multiple edges, because every pair of adjacent vertices is contained in exactly one edge of either $E_1(A)$ or $E_2(A)$. The adjacency within each one of these two sets gives rise to a partition of $V(A)$. More specifically, two vertices $v_{x,y}$ and $v_{x',y'}$ belong to the same class according to the partition derived from $E_1(A)$ (respectively, $E_2(A)$) if and only if $x = x'$ (respectively, $y = y'$). Then, the next result follows readily from the commutativity of the algebra A and the fact that every two vertices within the same equivalence class are adjacent.

Lemma 1. *Both subgraphs $(V(A), E_1(A))$ and $(V(A), E_2(A))$ are isomorphic and constitute a disjoint union of complete graphs within $\mathcal{G}(A)$.*

From a genetic point of view, every vertex in $\mathcal{G}(A)$ represents a probable status of the genotype located in two given chromatids C_1 and C_2 during the S-phase of a mitotic cell cycle, together with the probable status of the genotype of the offspring. More specifically, if $x = \sum_{i=1}^{n} a_i e_i$ and $y = \sum_{i=1}^{n} b_i e_i$ are two vectors in A such that $xy \neq 0$, then the vertex $v_{x,y} \in V(A)$ is related to the following facts.

1. Let $j \in \{1, \ldots, n\}$. If $a_j = 0$ (respectively, $b_j = 0$), then the chromatid C_1 (respectively, C_2) does not contain the genotype e_j. The same occurs for the offspring if $\sum_{i=1}^{n} a_i b_i c_{ij} = 0$, which corresponds to the coefficient of the basis vector e_j in the product xy.

2. Let $j, k \in \{1, \ldots, n\}$ be such that $j \neq k$. If $a_j > a_k$ (respectively, $b_j > b_k$), then it is more probable for the chromatid C_1 (respectively, C_2) to contain the genotype e_j than the genotype e_k. If $a_j = a_k$ (respectively, $b_j = b_k$), then both possibilities are equally probable. The same occurs for the offspring if $\sum_{i=1}^{n} a_i b_i c_{ij} \geq \sum_{i=1}^{n} a_i b_i c_{ik}$.

Furthermore, two different vertices in $\mathcal{G}(A)$ are equally colored if and only if they are associated with the same probable status of the offspring's genotype.

Concerning the edges in $\mathcal{G}(A)$, they represent the regulatory mechanisms produced during the G_2-phase of the mitotic cell cycle. Specifically, every edge $v_{x,y}v_{x,z} \in E_1(A)$ (respectively, $v_{x,y}v_{z,y} \in E_2(A)$) represents a mutation between two probable statuses y and z (respectively, x and z) of the genotype in the chromatid C_2 (respectively, C_1), without changing to this end the probable status of the genotype in C_1 (respectively, C_2). Thus, for instance, the vertex v_{e_1,e_1+e_2} in any of the three graphs in Figure 1 involves the chromatid C_1 containing the genotype e_1, while the chromatid C_2 contains, with the same probability, either the genotype e_1 or e_2. The first case would give rise to an offspring of genotype e_1, because $e_1 e_1 = e_1$, whereas the second one would not give rise to any offspring, because $e_1 e_2 = 0$. In order to avoid the latter, the regulatory mechanism produced during the G_2-phase of the mitotic cell could mutate the possible genotype e_2 in C_2 into the genotype e_1. This is represented in the graph $\mathcal{G}(A)$ by the edge $v_{e_1,e_1+e_2}v_{e_1,e_1} \in E_1(A)$.

Every path within the graph $\mathcal{G}(A)$ is therefore interpreted as a series of mutations from a pair of probable statuses of the genotype of both chromatids C_1 and C_2 into another. The following result shows that these mutations are possible, whatever the probable statuses under consideration are.

Lemma 2. *The graph $\mathcal{G}(A)$ is connected. Moreover, $\delta(\mathcal{G}(A)) \leq 7$.*

Proof. Let $\{v_{x,y}, v_{z,t}\} \subseteq V(A)$, where, for each $\alpha \in \{x, y, z, t\}$, it is $\alpha = \sum_{i=1}^{n} c_i^{\alpha} e_i$. Since $xy \neq 0 \neq zt$, there exist two positive integers $j, j' \leq n$ such that $\sum_{i=1}^{n} c_i^x c_i^y c_{ij} \neq 0 \neq \sum_{i=1}^{n} c_i^z c_i^t c_{ij'}$. Let $i, i' \in \{1, \ldots, n\}$ be such that $c_i^x c_i^y c_{ij} \neq 0 \neq c_{i'}^z c_{i'}^t c_{i'j'}$. Then, the ordered set of vertices:

$$\{v_{x,y},\ v_{x,e_i},\ v_{e_i,e_i},\ v_{e_i,e_i+e_{i'}},\ v_{e_{i'},e_i+e_{i'}},\ v_{e_{i'},e_{i'}},\ v_{x',e_{i'}},\ v_{x',y'}\} \subseteq V(A),$$

determines a path in $\mathcal{G}(A)$ of seven alternating edges in $E_1(A)$ and $E_2(A)$. $\quad\square$

In order to deal with the tightness of the upper bound in Lemma 2, we introduce some new concepts. To this end, let $\{v_{x,y}, v_{z,t}\} \subseteq V(A)$, where, for each $\alpha \in \{x, y, z, t\}$, it is $\alpha = \sum_{i=1}^{n} c_i^{\alpha} e_i$. We say that the vertices $v_{x,y}$ and $v_{z,t}$ are compatible if there exists a positive integer $i \leq n$ such that $c_i^{\alpha} \neq 0$, for all vectors $\alpha \in \{x, y, z, t\}$, and $e_i e_i \neq 0$. Further, a vertex in $V(A)$ is called pure if it has the form $v_{ae_i,be_i} \in V(A)$ for some $a, b \in \mathbb{K} \setminus \{0\}$ and for some i with $e_i e_i \neq 0$. Otherwise, the vertex is called hybrid.

Proposition 1. *The eccentricity of every pure vertex in $\mathcal{G}(A)$ is, at most, three.*

Proof. Let $v_{x,y} \in V(A)$ be a vertex that is not adjacent to a pure vertex $v_{ae_i,be_i} \in V(A)$, where, for each $\alpha \in \{x, y\}$, we have $\alpha = \sum_{i=1}^{n} c_i^{\alpha} e_i$. If $c_i^x \neq 0$ (respectively, $c_i^y \neq 0$), then the ordered set of vertices $\{v_{x,y}, v_{x,be_i}, v_{ae_i,be_i}\} \subseteq V(A)$ (respectively, $\{v_{x,y}, v_{ae_i,y}, v_{ae_i,be_i}\} \subseteq V(A)$) determines a

two-length alternating path in $\mathcal{G}(A)$. Otherwise, if $c_i^x = 0 = c_i^y$, then the ordered set of vertices $\{v_{x,y}, v_{x,be_i+y}, v_{ae_i,be_i+y}, v_{ae_i,be_i}\} \subseteq V(A)$ determines a three-length alternating path. □

Proposition 2. *The eccentricity of every hybrid vertex in $\mathcal{G}(A)$ is, at most, three.*

Proof. Let $v_{x,y}$ and $v_{z,t}$ be two non-adjacent vertices in $\mathcal{G}(A)$, where, for each $\alpha \in \{x, y, z, t\}$, we have $\alpha = \sum_{i=1}^{n} c_i^\alpha e_i$. In addition, suppose that the vertex $v_{x,y}$ is hybrid. From Proposition 1, if $v_{z,t}$ were a pure vertex, then $d(v_{x,y}, v_{z,t}) \leq 3$. Thus, from now on, we focus on the case that $v_{z,t}$ is also hybrid.

If the vertices $v_{x,y}$ and $v_{z,t}$ are compatible, then there exists a positive integer $i \leq n$ such that $c_i^x, c_i^y, c_i^z, c_i^t \in \mathbb{K} \setminus \{0\}$ and $e_i e_i \neq 0$. Thus, the ordered set of vertices $\{v_{x,y}, v_{x,e_i}, v_{z,e_i}, v_{z,t}\}$ determines a three-length alternating path.

Otherwise, if the vertices $v_{x,y}$ and $v_{z,t}$ are not compatible, then there exist two different positive integers $i, j \leq n$ such that $c_i^x, c_i^y, c_j^z, c_j^t \in \mathbb{K} \setminus \{0\}$ and $e_i e_i \neq 0 \neq e_j e_j$. If $c_j^x \neq 0$, then the ordered set of vertices $\{v_{x,y}, v_{x,e_j}, v_{z,e_j}, v_{z,t}\}$ determines a three-length alternating path. Similar paths of the same length can be found if $c_j^y \neq 0$, $c_i^z \neq 0$, or $c_i^t \neq 0$. On the other hand, if $c_j^x = c_j^y = c_i^z = c_i^t = 0$, then the ordered set of vertices $\{v_{x,y}, v_{x,e_i+e_j}, v_{z,e_i+e_j}, v_{z,t}\}$ determines a three-length alternating path. □

Theorem 1. *If the genetic pattern of the evolution algebra A has more than one non-zero component, then:*

$$2 \leq \rho(\mathcal{G}(A)) \leq 3 = \delta(\mathcal{G}(A)).$$

Otherwise, $\rho(\mathcal{G}(A)) = \delta(\mathcal{G}(A)) = 2$.

Proof. The non-triviality of the algebra involves the existence of a positive integer $i \leq n$ such that $e_i e_i \neq 0$. In addition, let $j \leq n$ be a positive integer such that $i \neq j$. This exists because $n > 1$. Then, since the graph $\mathcal{G}(A)$ is connected and contains, for instance, the edge $v_{e_i,e_i} v_{e_i+e_j,e_i}$, we have that $\rho(\mathcal{G}(A)) > 0$.

Now, suppose that $\rho(\mathcal{G}(A)) = 1$. Then, there exists a vertex $v_{x,y} \in V(A)$ that is adjacent to any other vertex in the graph. Since every evolution algebra is commutative, we have that $yx = xy \neq 0$, and hence, $v_{y,x} \in V(A)$. This is adjacent to $v_{x,y}$, and hence, from (2) and (3), we have that $x = y$. Moreover, since the vertex $v_{e_i,e_i} \in V(A)$ is adjacent to $v_{x,x}$, we have that $x = e_i$. This implies that $e_j e_j = 0$. Otherwise, the vertices v_{e_i,e_i} and v_{e_j,e_j} should be adjacent, which is a contradiction with the fact that $e_i \neq e_j$. As a consequence, $(e_i + e_j)(e_i + e_j) = e_i e_i \neq 0$. Thus, the vertices v_{e_i,e_i} and $v_{e_i+e_j,e_i+e_j}$ should be adjacent, and hence, $e_i = e_i + e_j$, which is also a contradiction. Therefore, $\rho(\mathcal{G}(A)) \geq 2$. This fact, together with Propositions 1 and 2, holds the first assertion of the theorem.

Now, in order to prove the second assertion, suppose the existence of a positive integer $i \leq n$ such that $e_i e_i \neq 0$ and $e_j e_j = 0$, for all $j \neq i$. Then, every pair of vertices $v_{x,y}, v_{z,t} \in V(A)$ is compatible so that the coefficient of the basis vector e_i in each one of the four vectors $x, y, z, t \in A$ is not zero. In particular, $xt \neq 0$, and hence, the ordered set of vertices $\{v_{x,y}, v_{x,t}, v_{z,t}\}$ determines a two-length alternating path. Therefore, $\delta(\mathcal{G}(A)) \leq 2$, and the result holds readily. □

Note that all the bounds in the first assertion of Theorem 1 are tight. Thus, for instance, the radius of the second graph in Figure 1 is three, whereas the radii of the first and third graphs in the same figure are two. Furthermore, from a genetic point of view, Theorem 1 can be interpreted as the fact that any pair of probable statuses of the genotype of two chromatids can be transformed into another pair by means of at most two or three mutations.

4. A Faithful Functor among Evolution Algebras and Graphs

Apart from the simplicity of describing the role of genes in mutation problems, the total-colored graph that we have just introduced in the previous section becomes also relevant from a categorical point of view. In this regard, we make use in this section of the mentioned graph in order to introduce

a faithful functor among the category of evolution algebras on a finite field and the category of total-colored graphs. Unlike the graphs associated with evolution algebras that already exist in the literature (see the introductory section), our description enables one to ensure that isomorphic evolution algebras always give rise to isomorphic total-colored graphs. Moreover, we prove in Proposition 3 that the previous assertion holds indeed for isotopic evolution algebras and not only for isomorphic ones.

From here on, the following notations are used:

- Let $\mathcal{E}$ and $\mathcal{M}$ respectively denote the set of finite-dimensional evolution algebras over a finite field and the set formed by their associated total-colored graphs. For each algebra $A \in \mathcal{E}$, the set of vertices of the graph $\mathcal{G}(A)$ is $V(A) := \{v_{x,y}^A \mid x,y \in A, xy \neq 0\}$.
- Let $\mathrm{Isot}(\mathcal{E})$ and $\mathrm{Isom}(\mathcal{M})$ respectively denote the set of isotopisms of algebras in $\mathcal{E}$ and the set of isomorphisms of graphs in $\mathcal{M}$.
- Let $\mathrm{Isot}(A_1, A_2)$ and $\mathrm{Isom}(G_1, G_2)$ respectively denote the set of isotopisms between two given evolution algebras $A_1, A_2 \in \mathcal{E}$ and the set of isomorphisms between two given graphs $G_1, G_2 \in \mathcal{M}$.
- Let $\mathfrak{E}$ denote the category formed by the sets $\mathcal{E}$ and $\mathrm{Isot}(\mathcal{E})$.
- Let $\mathfrak{M}$ denote the category formed by the sets $\mathcal{M}$ and $\mathrm{Isom}(\mathcal{M})$.

Let us finish this paper with the establishment of a faithful functor between the categories $\mathfrak{E}$ and $\mathfrak{M}$. In this regard, let $F : \mathfrak{E} \to \mathfrak{M}$ be such that:

(a) $F(A) := \mathcal{G}(A)$, for all $A \in \mathcal{E}$.
(b) For each isotopism $\Theta = (f, g, h) \in \mathrm{Isot}(\mathcal{E})$ between two isotopic evolution algebras $A_1, A_2 \in \mathcal{E}$, let $F_\Theta : \mathcal{G}(A_1) \to \mathcal{G}(A_2)$ be such that:

$$F_\Theta(v_{x,y}^{A_1}) := v_{f(x),g(y)}^{A_2}, \ \text{ for all } v_{x,y}^{A_1} \in V(A_1).$$

The non-singularity of the three linear transformations f, g, and h involves F_Θ being a well-defined bijection from $V(A_1)$ to $V(A_2)$. Observe to this end that, for each $x,y \in A$, one has that $f(x)g(y) = h(xy) \neq 0$ if and only if $xy \neq 0$.

Proposition 3. *The map $F : \mathfrak{E} \to \mathfrak{M}$ is a faithful functor.*

Proof. Let $A_1, A_2 \in \mathcal{E}$ and $\Theta \in \mathrm{Isot}(A_1, A_2)$. Then, we define the map:

$$\begin{aligned} \mathrm{Isot}(A_1, A_2) & \to & \mathrm{Isom}(\mathcal{G}(A_1), \mathcal{G}(A_2)) \\ \Theta & \to & F_\Theta \end{aligned} \tag{4}$$

Let us prove that this is well defined and injective. Suppose to this end that $\Theta = (f, g, h) \in \mathrm{Isot}(A_1, A_2)$. The non-singularity of both linear transformations f and g implies that $F_\Theta : V(A_1) \to V(A_2)$ maps $E_1(A_1)$ and $E_2(A_1)$ to $E_1(A_2)$ and $E_2(A_2)$, respectively. Furthermore, the non-singularity of the linear transformation h implies that $F_\Theta(V(A_1)) = V(A_2)$. Hence, $F_\Theta \in \mathrm{Isom}(\mathcal{G}(A_1), \mathcal{G}(A_2))$.

Now, suppose the existence of a pair of isotopisms $\Theta_1 = (f_1, g_1, h_1)$ and $\Theta_2 = (f_2, g_2, h_2)$ in the set $\mathrm{Isot}(A_1, A_2)$ such that $F_{\Theta_1} = F_{\Theta_2}$. Then, $v_{f_1(x),g_1(y)}^{A_2} = F_{\Theta_1}(v_{x,y}^{A_1}) = F_{\Theta_2}(v_{x,y}^{A_1}) = v_{f_2(x),g_2(y)}^{A_2}$, for all $x,y \in A_1$ such that $xy \neq 0$. As a consequence, $f_1 = f_2$ and $g_1 = g_2$, and hence, $h_1(xy) = f_1(x)g_1(y) = f_2(x)g_2(y) = h_2(xy)$. Therefore, both isotopisms, Θ_1 and Θ_2, coincide. $\square$

5. Conclusions and Further Work

This paper dealt with the description of a total-colored graph associated with the genetic pattern of any finite-dimensional evolution algebra over a finite field. We proved that both categories formed by such algebraic and combinatorial structures are related by a faithful functor. The faithfulness of

the functor implies that any isotopism of evolution algebras is uniquely related to an isomorphism of total-colored graphs. As a consequence, isotopisms of evolution algebras can be thought from now on as isomorphisms of total-colored graphs satisfying some extra properties. An explicit determination and characterization of the latter were established as further work in order to identify under which circumstances the functor under consideration is indeed fully faithful. In any case, these extra properties derive straightforwardly from the linearity of any isotopism of algebras.

Further, throughout the paper, we interpreted some basic properties and results of the described total-colored graph in the context of the genetic processes that occur during a mitotic cell cycle. Keeping in mind that the spectrum of possible patterns is currently known only for those evolution algebras that are involved with two or three genotypes, we established as further work the study of graph invariants of the proposed total-colored graph in order to deal with a higher number of genotypes.

It is also required to delve into the translation of algebraic concepts and results on evolution algebras to the language of total-colored graphs. A particular focus of interest in this regard is the associator $\mathcal{A}(x, y, z) := (xy)z - x(yz)$, which constitutes, according to Goze and Remms [56], an interesting way of classifying nonassociative algebras over the group algebra $\mathbb{K}[\Sigma_3]$ of the symmetric group Σ_3 over a field $\mathbb{K}$ and also for studying their algebraic properties. In order to illustrate such a classification in the case of dealing with evolution algebras, let A be an n-dimensional evolution algebra over the field $\mathbb{K}$, with natural basis $\{e_1, \ldots, e_n\}$ and such that $e_i e_i = \sum_{j=1}^{n} c_{ij} e_j$, with $c_{ij} \in \mathbb{K}$, for all i, j. Then,

$$\mathcal{A}(e_i, e_i, e_j) + \mathcal{A}(e_j, e_i, e_i) = c_{ij} e_j e_j - c_{ij} e_j e_j = 0, \text{ for all } i \neq j,$$

and hence,

$$\mathcal{A}(e_i, e_k, e_j) + \mathcal{A}(e_k, e_i, e_j) + \mathcal{A}(e_j, e_i, e_k) + \mathcal{A}(e_j, e_k, e_i) = 0, \text{ for all } i, j, k.$$

According to the notation that is used in [56], the previous identity becomes:

$$\mathcal{A} \circ \Phi_v = 0.$$

Here, Φ_v denotes the right action on the group algebra $\mathbb{K}[\Sigma_3]$ by means of the vector:

$$v := \text{Id} + \tau_{12} + c_2 + \tau_{13},$$

where, from here on, τ_{ij} denotes the transposition $(i, j) \in \Sigma_3$ and c_2 is the cycle of order three $(1, 3, 2) \in \Sigma_3$. The orbit of this vector v for the natural action of Σ_3 on the group algebra $\mathbb{K}[\Sigma_3]$ is:

$$\mathcal{O}(v) := \{v, \text{Id} + \tau_{13} + \tau_{23} + c_1, \tau_{23} + c_1 + c_2 + \tau_{12}\},$$

where c_1 denotes the cycle of order three $(1, 2, 3) \in \Sigma_3$. Then, $\mathbb{K}(\mathcal{O}(v))$ is an invariant irreducible space of the group algebra $\mathbb{K}[\Sigma_3]$. In particular,

$$2\text{Id} + 2\tau_{12} + 2\tau_{13} + 2\tau_{23} + 2c_1 + 2c_2 \in \mathbb{K}(\mathcal{O}(v)).$$

According to the terminology in [56], this means that evolution algebras are $\mathbb{K}[\Sigma_3]$-associative. This implies that a development of algebraic properties (like operad, homology, cohomology, amongst others) could be done as further work.

Author Contributions: Investigation, R.M.F., Ó.J.F., and J.N.; methodology, R.M.F.; supervision, R.M.F. and J.N.; validation, R.M.F.; writing, original draft, R.M.F. and Ó.J.F.; writing, review and editing, R.M.F. and J.N.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Tian, J.P. *Evolution Algebras and their Applications*; Lecture Notes in Mathematics; Springer-Verlag: Berlin, Germany, 2008; Volume 1921.
2. Tian, J.P. Evolution Algebra Theory. Ph.D. Thesis, University of California, Riverside, CA, USA, 2004.
3. Tian, J.P.; Vojtechovsky, P. Mathematical concepts of evolution algebras in non-mendelian Genetics. *Quasigroups Relat. Syst.* **2006**, *14*, 111–122.
4. Khudoyberdiyev, A.; Omirov, B.A.; Qaralleh,I. Few remarks on evolution algebras. *J. Algebra Appl.* **2015**, *14*, 1550053. [CrossRef]
5. Cadavid, P.; Rodiño, M.L.; Rodríguez, P.M. The connection between evolution algebras, random walks and graphs. *J. Algebra Appl.* **2019**. [CrossRef]
6. Cabrera Casado, Y.; Siles Molina, M.; Velasco, M.V. Classification of three-dimensional evolution algebras. *Linear Algebra Appl.* **2017**, *524*, 68–108. [CrossRef]
7. Camacho, L.M.; Gómez, J.R.; Omirov, B.A.; Turdibaev, R.M. Some properties of evolution algebras. *Bull. Korean Math. Soc.* **2013**, *50*, 1481–1494. [CrossRef]
8. Camacho, L.M.; Gómez, J.R.; Omirov, B.A.; Turdibaev, R.M. The derivations of some evolution algebras. *Linear Multilinear A* **2013**, 61, 309–322. [CrossRef]
9. Casas, J.M.; Ladra, M.; Rozikov, U.A. A chain of evolution algebras. *Linear Algebra Appl.* **2011**, *435*, 852–870. [CrossRef]
10. Casas, J.M.; Ladra, M.; Omirov, B.A.; Rozikov, U.A. On nilpotent index and dibaricity of evolution algebras. *Linear Algebra Appl.* **2013**, *439*, 90–05. [CrossRef]
11. Casas, J.M.; Ladra, M.; Omirov, B.A.; Rozikov, U.A. On Evolution Algebras. *Algebra Colloq.* **2014**, *21*, 331–342. [CrossRef]
12. Dzhumadildaev, A.; Omirov, B.A.; Rozikov, U.A. Constrained evolution algebras and dynamical systems of a bisexual population. *Linear Algebra Appl.* **2016**, *496*, 351–380. [CrossRef]
13. Labra, A.; Ladra, M.; Rozikov, U.A. An evolution algebra in population Genetics. *Linear Algebra Appl.* **2014**, *457*, 348–362. [CrossRef]
14. Ladra, M.; Omirov, B.A.; Rozikov, U.A. Dibaric and evolution algebras in biology. *Lobachevskii J. Math.* **2014**, *35*, 198–210. [CrossRef]
15. Ladra, M.; Rozikov, U.A. Evolution algebra of a bisexual population. *J. Algebra* **2013**, *378*, 153–172. [CrossRef]
16. Omirov, B. A.; Rozikov, U.A.; Tulenbayev, K.M. On real chains of evolution algebras. *Linear Multilinear Algebra* **2015**, *63*, 586–600. [CrossRef]
17. Falcón, O.J.; Falcón, R.M.; Núñez, J. Classification of asexual diploid organisms by means of strongly isotopic evolution algebras defined over any field. *J. Algebra* **2017**, *472*, 573–593. [CrossRef]
18. Falcón, O.J.; Falcón, R.M.; Núñez, J. Algebraic computation of genetic patterns related to three-dimensional evolution algebras. *Appl. Math. Comput.* **2018**, *319*, 510–517. [CrossRef]
19. Albert, A.A. Non-Associative Algebras: I. Fundamental Concepts and Isotopy. *Ann. Math. Second. Ser.* **1942**, *43*, 685–707. [CrossRef]
20. Campos, T.M.M.; Holgate, P. Algebraic Isotopy in Genetics. *IMA J. Math. Appl. Med. Biol.* **1987**, *4*, 215–222. [CrossRef]
21. Holgate, P. Genetic algebras associated with polyploidy. *Proc. Edinb. Math. Soc.* **1966**, *15*, 1–9. [CrossRef]
22. Falcón, R.M.; Falcón, O.J.; Núñez, J. A Historical Perspective of the Theory of Isotopisms. *Symmetry* **2018**, *10*, 1–21. [CrossRef]
23. Benzer, S. On the topology of genetic fine structure. *Proc. Nat. Acad. Sci. USA* **1959**, *45*, 1607–1620. [CrossRef] [PubMed]
24. Cadavid, P.; Rodiño, M.L.; Rodríguez, P.M. Characterization theorems for the spaces of derivations of evolution algebras associated to graphs. *Linear Multilinear A* **2019**. [CrossRef]
25. Núñez, J.; Silvero, M.; Villar, M.T. Mathematical tools for the future: Graph Theory and graphicable algebras. *Appl. Math. Comput.* **2013**, *219*, 6113–6125. [CrossRef]
26. Elduque, A.; Labra, A. Evolution algebras and graphs. *J. Algebra Appl.* **2015**, *14*, 1550103. [CrossRef]
27. Elduque, A.; Labra, A. On nilpotent evolution algebras. *Linear Algebra Appl.* **2016**, *505*, 11–31. [CrossRef]

28. Cabrera Casado, Y.; Siles Molina, M.; Velasco, M.V. Evolution algebras of arbitrary dimension and their decompositions. *Linear Algebra Appl.* **2016**, *495*, 122–162. [CrossRef]

29. Falcón, O.J.; Falcón, R.M.; Núñez, J.: Pacheco, A.M.; Villar, M.T. Computation of isotopisms of algebras over finite fields by means of graph invariants. *J. Comput. Appl. Math.* **2017**, *318*, 307–315. [CrossRef]

30. McKay, B.D.; Meynert, A.; Myrvold, W. Small Latin Squares, Quasigroups and Loops. *J. Combin. Des.* **2007**, *15*, 98–119. [CrossRef]

31. Falcón, O.J.; Falcón, R.M.; Núñez, J. Mutation graphs of asexual diploid organisms. In *X Encuentro Andaluz de Matemática Discreta*; Universidad de Cádiz: Cádiz, Spain, 2017; pp. 103–106.

32. Brooker, R. *Genetics: Analysis and Principles*; McGraw-Hill Education: New York, NY, USA, 2018.

33. Harary, F. *Graph Theory*; Addison Wesley: Reading, MA, USA, 1969.

34. Jones, N.C.; Pevzner, P.A. *An introduction to Bioinformatics Algoritms*; MIT Press: Cambridge, MA, USA, 2004.

35. Pevzner, P.A. *Computational Molecular Biology*; MIT Press: Cambridge, MA, USA, 2000.

36. Serebrowsky, A. On the properties of the Mendelian equations. *Dokl. Akad. Nauk. SSSR* **1934**, *2*, 33–36.

37. Mendel, G. *Experiments in Plant-Hybridization*; Classic Papers in Genetics; Prentice-Hall Inc.: Upper Saddle River, NJ, USA, 1959.

38. Glivenkov, V. Algebra Mendelienne. *C. R. Acad. Sci. URSS* **1936**, *4*, 385–386.

39. Kostitzin, V.A. Sur les coefficients mendeliens d'heredite. *C. R. Acad. Sci. URSS* **1938**, *206*, 883–885.

40. Etherington, I.M.H. Genetic algebras. *Proc. R. Soc. Edinb.* **1939**, *59*, 242–258. [CrossRef]

41. Etherington, I.M.H. Duplication of linear algebras. *Proc. Edinb. Math. Soc.* **1941**, *6*, 222–230. [CrossRef]

42. Etherington, I.M.H. Non-associative algebra and the symbolism of Genetics. *Proc. R. Soc. Edinb. Sect. B* **1941**, *61*, 24–42. [CrossRef]

43. Jennings, H.S. Observed changes in hereditary characters in relation to evolution. *J. Wash. Acad. Sci.* **1917**, *7*, 281–301.

44. Schafer, R.D. Structure of genetic algebras. *Am. J. Math.* **1949**, *71*, 121–135. [CrossRef]

45. Gonshör, H. Contributions to genetic algebras. *Proc. Edinb. Math. Soc.* **1971**, *17*, 289–298. [CrossRef]

46. Haldane, J.B.S. A mathematical theory of natural and artificial selection. *Bull. Math. Biol.* **1990**, *52*, 209–240.

47. Heuch, I. Sequences in genetic algebras for overlapping generations. *Proc. Edinb.* **1972**, *16*, 291–294. [CrossRef]

48. Reiersöl, O. Genetic algebras studied recursively and by means of differential operators. *Math. Scand.* **1962**, *10*, 25–44. [CrossRef]

49. Korol, A.B.; Preygel, I.A.; Preygel, S.I. *Recombination Variability and Evolution*; Chapman & Hall: London, UK, 1994.

50. Lyubich, Y.I. *Mathematical Structures in Population Genetics*; Springer: Berlin, Germany, 1992.

51. Wörz-Busekros, A. *Algebras in Genetics*; Springer: Berlin/Heidelberg, Germany, 1980.

52. Bertrand, M. *Algèbres non Associatives et Algèbres Gènètiques*; Gauthier-Villars Editeur: Paris, France, 1966.

53. Holgate, P. Characterizations of genetic algebras. *J. Lond. Math. Soc.* **1972**, *2*, 169–174. [CrossRef]

54. Dzhumadil'daev, A.S.; Omirov, B.A.; Rozikov, U.A. On a class of evolution algebras of "chicken" population. *Int. J. Math.* **2014**, *25*, 1450073. [CrossRef]

55. Rozikov, U.A.; Velasco, M.V. A discrete-time dynamical system and an evolution algebra of mosquito population. *J. Math. Biol.* **2019**, *78*, 1225–1244. [CrossRef] [PubMed]

56. Goze, M.; Remm, E. A class of nonassociative algebras. *Algebra Colloq.* **2007**, *14*, 313–326. [CrossRef]