International Electronic Journal of Algebra
Published Online: September 6, 2026
DOI: 10.24330/ieja.2034027

PHENOTYPE ALGEBRA AND THE DOMINANCE MAP

Manuel Arenas

Received: 14 March 2025; Revised: 10 July 2026; Acceptance: 17 July 2026
Communicated by A. Çiğdem Özcan Çevik

Abstract. In this article, we analyze the algebraic aspects of the phenotype of a specific gene with n different alleles. We study the genetic realization algebra that models the probabilities of an offspring's phenotype given the phenotypes of its parents; we call this algebraic structure the *phenotype algebra*. We introduce a *dominance map* that assigns the probability distribution of an individual's possible phenotypes to a probability distribution of their possible genotypes. We use this dominance map, together with well-known concepts such as gametic algebra, the tensor product of spaces, and the adjoint of a linear map, to compute the phenotype algebra. We provide three simple examples to illustrate how the construction works. We also analyze the case of a gene where the dominance of one allele over another is incomplete. In this case, we obtain a Lotka-Volterra algebra associated with an antisymmetric matrix.

Mathematics Subject Classification (2020): 17D92, 92D10
Keywords: Genetic algebra, allele, phenotype, nonassociative algebra

1. Introduction

The relation between nonassociative algebras and genetics can be traced back to the work of Gregor Mendel. In fact, the symbolism Mendel used to describe his first results is reminiscent of algebraic notation (see his 1866 paper *Experiments in Plant-Hybridization*, in the 1959 edition [12] edited by James Peters). More than seventy years later, in the article [5] of 1942, I.M.H. Etherington introduced the formal language of abstract algebra to the study of genetics. Etherington defined *algebras with genetic realization* and established an entire area within mathematical biology.

Today, there is a very vast literature on the subject. Particularly important are the works [16] of 1949 by R.D. Schafer and [7] of 1971 by H. Gonshor, which independently gave two definitions (which are, in fact, equivalent) of *genetic algebras*

Project supported by the Competition for Research Regular Projects, year 2023, code LPR23-21, Universidad Tecnológica Metropolitana.

(see Figure 1 in the article [15] of 1997 by M.L. Reed). Other classes of algebras have acquired importance, such as *Bernstein algebras* (see the article [3] of 1942 by S.N. Bernstein) or *train algebras* (see the article [6] of 1960 by H. Gonshor).

For us, the most comprehensive references for the mathematical research done in this area are the article [15] of 1997 by M.L. Reed and the book [20] of 1980 by A. Wörz-Busekros. The theory of algebras modeling genetics has experienced significant development in recent years; new classes of algebras have been introduced, such as *evolution algebras of bisexual populations* introduced in the article [11] of 2013 by M. Ladra and U.A. Rozikov, *evolution algebras of "chicken" populations* introduced in the article [10] of 2014 by A. Labra, M. Ladra and U.A. Rozikov, or *gonosomal algebras* introduced in the article [19] of 2016 by R. Varro. We recommend the reader to see also the article [18] of 2004 by J. Tian and B.L. Li, the article [4] of 2015 by J.M. Casas, M. Ladra, B.A. Omirov and R. Turdibaev, the article [13] of 2017 by I. Paniello, or the article [8] of 2019 by J.C. Gutiérrez-Fernández and C. García for recent works on the subject.

An *algebra* over a field $\mathbb{K}$ is a vector space $\mathcal{A}$ over $\mathbb{K}$ endowed with a linear map $\nabla : \mathcal{A} \otimes \mathcal{A} \to \mathcal{A}$ that defines its multiplication, i.e., $x \cdot y = \nabla(x \otimes y)$. We do not require this multiplication to be associative or commutative, nor do we require the existence of a unit element in the algebra.

However, in this article, all algebras are assumed to be commutative. Furthermore, we fix $\mathbb{K} = \mathbb{R}$; that is, the field of scalars is always the set of real numbers.

The dimension of the algebra is its dimension as a vector space; in this article, we work exclusively with finite-dimensional algebras. If the dimension is n and we choose a basis $\mathcal{E} = \{\mathbf{e}_k\}_{k=1}^n$, the real-valued constants $\{\lambda_{i,j,k}\}_{i,j,k=1}^n$ such that

$$\mathbf{e}_i \cdot \mathbf{e}_j = \sum_{k=1}^{n} \lambda_{i,j,k} \mathbf{e}_k \tag{1}$$

are called the *structure constants* of $\mathcal{A}$ relative to $\mathcal{E}$.

We say that a real algebra $\mathcal{A}$ is an *algebra with genetic realization* (or *AGR*) if it possesses a basis $\mathcal{E}$, called its *natural basis*, whose structure constants satisfy the following conditions for every triplet of indices (i, j, k):

$$\sum_{s=1}^{n} \lambda_{i,j,s} = 1, \tag{2}$$

$$\lambda_{i,j,k} \geq 0, \tag{3}$$

$$\lambda_{i,j,k} = \lambda_{j,i,k}. \tag{4}$$

It is straightforward to verify that identity (4) is equivalent to the commutativity of the algebra. Furthermore, identity (2) is equivalent to the fact that the linear map $\mathbf{w} : \mathcal{A} \to \mathbb{R}$, defined by $\mathbf{w}\left(\sum_{i=1}^{n} x_i \mathbf{e}_i\right) = \sum_{i=1}^{n} x_i$, is an $\mathbb{R}$-algebra homomorphism, known as the *weight map*. An algebra equipped with such a homomorphism is called a *baric algebra*.

We say that an algebra is a *Lotka-Volterra algebra* (or *LVA*) if there exists a natural basis $\{\mathbf{e}_k\}_{k=1}^{n}$ and an $n \times n$ skew-symmetric matrix $\mathbf{S}$ such that, for every pair (i, j) we have that $|\mathbf{S}_{i,j}| \leq \frac{1}{2}$ and:

$$\mathbf{e}_i \cdot \mathbf{e}_j = \left(\frac{1}{2} + \mathbf{S}_{i,j}\right) \mathbf{e}_i + \left(\frac{1}{2} + \mathbf{S}_{j,i}\right) \mathbf{e}_j. \tag{5}$$

Lotka-Volterra algebras first appeared in the 1981 work of Y. Itoh [9], where they were associated with a specific type of differential equation (see also [8] by J. C. Gutiérrez-Fernández and C. García for further properties).

A simple calculation shows that an algebra $\mathcal{A}$ is an LVA with natural basis $\{\mathbf{e}_k\}_{k=1}^{n}$ if and only if $\mathcal{A}$ is an AGR relative to the same basis and the structure constants also satisfy:

$$k \notin \{i, j\} \implies \lambda_{i,j,k} = 0 \tag{6}$$

(see, for instance, [1] and [2] by M. Arenas, A. Labra, and I. Paniello).

The relationship between AGRs and the modeling of genetic systems is as follows. Consider a population $\mathcal{U}$ and a trait C that can manifest in n different forms. Let C_i denote the subset of individuals in $\mathcal{U}$ whose trait C takes the i-th form. We define the probability vector of $\mathcal{U}$ as the tuple $x = (x_1, \ldots, x_n)$, where $x_i = \mathcal{P}(C_i)$ is the probability that a randomly chosen individual exhibits the i-th form of trait C. This implies that $x_i \geq 0$ for every index i and $\sum_{i=1}^{n} x_i = 1$.

Consider a population $\mathcal{U}$ with probability vector x and a population $\mathcal{V}$ with probability vector y. If an individual of $\mathcal{U}$ and an individual of $\mathcal{V}$ reproduce, the probability vector of the offspring, denoted by z, is given by $z = x \cdot y = \nabla(x \otimes y)$ for some AGR $(\mathcal{A}, \nabla)$.

Conditions (2) and (3) ensure that z has non-negative coordinates and weight 1, while condition (4)—equivalent to the commutativity of the algebra—represents *panmixia* from a biological perspective. That is, the offspring's probability distribution z is independent of which parent contributes the distribution x or y.

In the study of trait distributions within a population, it is standard to analyze the reproduction of two individuals chosen at random from the same population $\mathcal{U}$, consequently both parents has the same probability vector. Thus, if x is the probability vector of $\mathcal{U}$ in a given generation, the vector $z = x^2 = x \cdot x$ represents

the probability vector of the subsequent generation. For this reason, the map $x \mapsto x^2$ is known as the *evolution operator* of $\mathcal{A}$.

Throughout this article, whenever we describe an event E with n possible outcomes $\{E_1, \ldots, E_n\}$, the boldface symbol $\mathbf{E}_k$ denotes the vector $\mathbf{e}_k$ of the canonical basis of $\mathbb{R}^n$ (with 1 in the k-th position), representing a 100% probability of event E_k occurring. For instance, if the trait C can take the forms $\{C_1, \ldots, C_n\}$, the element $\mathbf{e}_k$ of the canonical basis of $\mathbb{R}^n$ will be denoted by $\mathbf{C}_k$.

Before introducing the most common AGRs appearing in the literature, we provide some essential biological terminology.

A *gene* is the unit of genetic information consisting of a specific segment of DNA, which is related to a particular characteristic or process in an individual. The physical location of a gene on the DNA is called its *locus*. DNA is organized into *chromosomes* present in every cell of an organism. Each gene can take different forms, known as *alleles*. Organisms carrying a double set of chromosomes in each cell are called *diploid*, whereas *haploid* cells (or organisms) carry a single set. In diploid organisms, the process of meiosis produces *gametes*, which are haploid cells. When gametes fuse (e.g., during fertilization), the result is a *zygote*, which is again a diploid cell.

Consider a gene with n different alleles $A_1, \ldots, A_n$. When two gametes—one carrying allele A_i and the other A_j—fuse, the zygote possesses the genotype A_iA_j. In a subsequent meiosis, this zygote will produce a new gamete carrying either allele A_i or A_j. Let $\gamma_{i,j,k}$ denote the probability that, given a zygote formed by alleles A_i and A_j, the resulting gamete will carry allele A_k. The family $\{\gamma_{i,j,k} \mid i, j, k \leq n\}$ defines a product in the space $\mathcal{G}_n = \operatorname{span}\{\mathbf{A}_1, \ldots, \mathbf{A}_n\}$ given by $\mathbf{A}_i \cdot \mathbf{A}_j = \sum_{k=1}^{n} \gamma_{i,j,k}\mathbf{A}_k$. The resulting algebra is called the *gametic algebra*. Note that $\gamma_{i,j,k}$ is a conditional probability; specifically, if A_iA_j denotes the event that the zygote has genotype A_iA_j, then $\gamma_{i,j,k} = \mathcal{P}(A_k \mid A_iA_j)$.

It is straightforward to show that the gametic algebra is an LVA, notice that if $k \notin \{i, j\}$, since a zygote A_iA_j has zero probability to contribute with a gamete with allele A_k, then $\gamma_{i,j,k} = \mathcal{P}(A_k \mid A_iA_j) = 0$. This means that its multiplication follows (5), where its dimension n corresponds to the number of alleles. For instance, under *simple Mendelian inheritance*, i.e., where for every pair of alleles A_i, A_j with $i \neq j$, we have $\mathcal{P}(A_i \mid A_iA_j) = \mathcal{P}(A_j \mid A_iA_j) = 1/2$, the matrix $\mathbf{S}$ in (5) is the null matrix.

As an example, the algebra $\mathcal{G}_2$ with simple Mendelian inheritance is the vector space $\text{span}\{\mathbf{A}_1, \mathbf{A}_2\}$ with commutative multiplication defined by:

$$\mathbf{A}_1^2 = \mathbf{A}_1, \quad \mathbf{A}_2^2 = \mathbf{A}_2, \quad \mathbf{A}_1 \cdot \mathbf{A}_2 = \frac{1}{2}(\mathbf{A}_1 + \mathbf{A}_2). \tag{7}$$

Another example of an AGR is the *zygotic algebra*, which models the probability of all possible genotypes in an offspring based on the genotypes of its parents. This algebra $\mathcal{Z}_n$ is the symmetric duplication of the gametic algebra, i.e.,

$$\mathcal{Z}_n = S^2(\mathcal{G}_n) = (\mathcal{G}_n \otimes \mathcal{G}_n)/\mathcal{I}, \tag{8}$$

where $\mathcal{I} = \text{span}\{\mathbf{A}_i \otimes \mathbf{A}_j - \mathbf{A}_j \otimes \mathbf{A}_i \mid 1 \leq i, j \leq n\}$. This implies that its natural basis is $\{\mathbf{A}_i\mathbf{A}_j\}_{i \leq j}$, where $\mathbf{A}_i\mathbf{A}_j$ denotes the class of $\mathbf{A}_i \otimes \mathbf{A}_j$ in the quotient. For example, the algebra $\mathcal{Z}_2$ is a three-dimensional $\mathbb{R}$-space with basis $\{\mathbf{A}_1\mathbf{A}_1, \mathbf{A}_1\mathbf{A}_2, \mathbf{A}_2\mathbf{A}_2\}$ and commutative product:

$$\begin{aligned}
(\mathbf{A}_i\mathbf{A}_i)^2 &= \mathbf{A}_i\mathbf{A}_i, \quad i = 1, 2, \\
(\mathbf{A}_1\mathbf{A}_1) \cdot (\mathbf{A}_2\mathbf{A}_2) &= \mathbf{A}_1\mathbf{A}_2, \\
(\mathbf{A}_1\mathbf{A}_2)^2 &= \frac{1}{4}(\mathbf{A}_1\mathbf{A}_1 + 2\mathbf{A}_1\mathbf{A}_2 + \mathbf{A}_2\mathbf{A}_2), \\
(\mathbf{A}_1\mathbf{A}_2) \cdot (\mathbf{A}_i\mathbf{A}_i) &= \frac{1}{2}(\mathbf{A}_i\mathbf{A}_i + \mathbf{A}_1\mathbf{A}_2), \quad i = 1, 2.
\end{aligned}$$

For the definition and properties of gametic and zygotic algebras, see Wörz-Busekros [20] (1980), L. A. Peresi [14] (1988), or M. L. Reed [15] (1997).

When a gene has only two alleles, it is customary to denote them as A and a instead of A_1 and A_2. This notation reflects the common case where one allele (A) is dominant and the other (a) is recessive. However, we observe that in both gametic and zygotic algebras, the mechanism of dominance is irrelevant to the algebraic structure.

If a gen with n alleles $A_1, \ldots, A_n$ is associated to a trait C with m posible forms $C_1, \ldots, C_m$, the pair A_iA_j present in the zygote corresponds to the *genotype* of the individual and the form C_k that will be present in the individual is its *phenotype*. For instance, a gen with two alleles A, a as in the above paragraph, we have two possible phenotypes C_1, C_2 where C_1 corresponds to the genotypes AA or Aa and the phenotype C_2 correspond to the genotype aa.

With this in mind it becomes natural to consider a *phenotype algebra*, where each element $\mathbf{C}_i$ of the natural basis represents an individual with phenotype C_i. In this context, the product given by (1) represents the probability of an offspring having phenotype C_k given that the parents have phenotypes C_i and C_j, respectively. It should be clarified that when a gene possesses n alleles, the number of possible

phenotypes is not necessarily n. Various genetic phenomena can occur; for instance, the ABO blood group system in humans involves three alleles (A, B, O) but results in four phenotypes (A, B, AB, O), also referred to as types I, II, III, and IV.

The remainder of this article is organized as follows. In Section 2, we introduce the *dominance map* and describe how to compute the multiplication in the phenotype algebra. In Section 3, we reduce the dimensions of the matrices involved by leveraging the commutativity of our examples and utilizing the symmetric tensor product instead of the standard tensor product. Section 4 demonstrates the application of this method through three specific examples. Finally, in Section 5, we consider the case where both dominance and gamete production during meiosis are stochastic and depend on specific probabilities. From this, we derive a product that depends on two skew-symmetric matrices, $\mathbf{S}$ and $\mathbf{T}$.

2. The phenotype algebra and the dominance map

Let us recall two concepts from linear algebra that will be essential. First, we adopt the convention of representing vectors as column matrices. When evaluating a linear map on a vector, the matrix of the map is placed on the left and the vector on the right, resulting in another column matrix.

Under this convention, a linear map $\mathbf{T} : \mathbb{R}^n \to \mathbb{R}^m$ is called *stochastic* if its matrix relative to the canonical basis has non-negative entries and the sum of the entries in each column is 1. For any linear map $\mathbf{T} : \mathbb{R}^n \to \mathbb{R}^m$, its adjoint is the map $\mathbf{T}^* : \mathbb{R}^m \to \mathbb{R}^n$, whose matrix in the canonical basis is the transpose of the matrix of $\mathbf{T}$. It is immediate that the adjoint of a stochastic map $\mathbf{T}$ is not necessarily stochastic (if a matrix has columns that sum to 1, its transpose will have rows that sum to 1).

Accordingly, for any linear map $\mathbf{T}$ whose matrix has non-negative entries and at least one non-zero entry in every row, we define the linear map $\mathbf{T}^\dagger$ by dividing each entry of $\mathbf{T}^*$ by the sum of the column to which it belongs. The conditions imposed on $\mathbf{T}$ ensure that no division by zero occurs. It is straightforward to verify that if $\mathbf{T}$ is stochastic and has at least one non-zero entry in every row, then $\mathbf{T}^\dagger$ is also stochastic.

Another operation we shall employ is the tensor product of linear maps. Consider two linear maps $\mathbf{T} : \mathbb{R}^n \to \mathbb{R}^m$ and $\mathbf{S} : \mathbb{R}^k \to \mathbb{R}^l$. Their tensor product $\mathbf{T} \otimes \mathbf{S} : \mathbb{R}^n \otimes \mathbb{R}^k \to \mathbb{R}^m \otimes \mathbb{R}^l$ is defined by $(\mathbf{T} \otimes \mathbf{S})(x \otimes y) = \mathbf{T}(x) \otimes \mathbf{S}(y)$. The matrix of this map relative to the natural basis of the tensor product is the *Kronecker product* of the respective matrices.

Remark 2.1. If two linear maps are stochastic and their composition is well-defined, then the composition is also stochastic. Furthermore, the tensor product of two stochastic maps is stochastic. Both facts can be verified by a simple calculation.

Consider a gene with n possible alleles $A_1, \ldots, A_n$ such that the associated phenotype manifests as m possible traits $C_1, \ldots, C_m$. Let $\mathcal{G}_n = \text{span}\{\mathbf{A}_1, \ldots, \mathbf{A}_n\}$ be the *genotype space* (or *allele space*), described previously as the underlying space of the gametic algebra, and let $\mathcal{F}_m = \text{span}\{\mathbf{C}_1, \ldots, \mathbf{C}_m\}$ be the *phenotype space* (or *space of traits*).

Definition 2.2. For a set of n alleles determining m traits, we define the *phenotype algebra* as the pair $(\mathcal{F}_m, \nabla)$, where $\nabla : \mathcal{F}_m \otimes \mathcal{F}_m \to \mathcal{F}_m$ is a linear map such that $\nabla(\mathbf{C}_i \otimes \mathbf{C}_j) = \sum_{k=1}^{m} p_{i,j,k}\mathbf{C}_k$. Here, the coefficient $p_{i,j,k}$ denotes the probability that a pair of parents with traits C_i and C_j produces an offspring with trait C_k.

The definition above may appear somewhat general, as it could describe any instance where a pair of events produces a subsequent event. However, our objective is to provide a systematic method to compute these probabilities when the genotype probabilities are known. To this end, let $\Gamma : \mathcal{G}_n \otimes \mathcal{G}_n \to \mathcal{G}_n$ denote the multiplication in the gametic algebra. We now define the *dominance map*:

Definition 2.3. Let $\mathcal{G}_n$ and $\mathcal{F}_m$ denote the genotype and phenotype spaces, respectively. The *dominance map* $\mathbf{D} : \mathcal{G}_n \otimes \mathcal{G}_n \to \mathcal{F}_m$ is defined by

$$\mathbf{D}(\mathbf{A}_i \otimes \mathbf{A}_j) = \sum_{k=1}^{m} d_{i,j,k}\mathbf{C}_k,$$

where $d_{i,j,k}$ represents the probability of an individual exhibiting trait C_k given the genotype A_iA_j.

For example, if allele A_i is dominant over allele A_j for a characteristic C_i, their image under the dominance map is simply $\mathbf{D}(\mathbf{A}_i \otimes \mathbf{A}_j) = \mathbf{D}(\mathbf{A}_j \otimes \mathbf{A}_i) = \mathbf{C}_i$.

Remark 2.4. The dominance map $\mathbf{D}$ is stochastic. Since each $d_{i,j,k}$ is a probability, it is non-negative and $\sum_{k=1}^{m} d_{i,j,k} = 1$ for every pair (i, j). We further assume that the matrix of the dominance map has at least one positive entry in every row. If a row k consisted entirely of zeros, it would imply that no genotype pair (i, j) could produce trait C_k, meaning C_k is not a possible form of the characteristic and should be removed from the phenotype space. A direct consequence of this is that the stochastic matrix $\mathbf{D}^\dagger$ is always well defined.

Remark 2.5. If C_k and A_iA_j denote the events of having trait C_k and genotype A_iA_j respectively, since we have that $d_{ijk} = P(C_k|A_iA_j)$, then the matrix $\mathbf{D}$ provides the phenotype probability vector from a given genotype probability vector via the law of total probability:

$$\mathcal{P}(C_k) = \sum_{i,j=1}^{n} \mathcal{P}(C_k \mid A_iA_j)\mathcal{P}(A_iA_j). \tag{9}$$

We require a map for the inverse process—one that yields the most plausible genotype probability vector for a given phenotype vector. Note that this is not simply the preimage under $\mathbf{D}$, which is not necessarily unique. For this purpose, we assume that all genotypes in the entire population $\mathcal{U}$ are equally likely. This assumption is expressed by:

$$\mathcal{P}(A_iA_j) = \mathcal{P}(A_rA_s) \quad \forall i,j,r,s \in \{1,\dots,n\} \tag{10}$$

when calculating the entries of the linear map matrix we seek.

Lemma 2.6. *If $\mathbf{D} : \mathcal{G}_n \otimes \mathcal{G}_n \to \mathcal{F}_m$ is the dominance map, then the linear map that assigns to each probability vector of phenotypes its corresponding probability vector of genotypes is $\mathbf{D}^\dagger$. This is equivalent to stating that the probability $\overline{d_{i,j,k}}$ of an individual having genotype A_iA_j given the phenotype C_k is given by the formula:*

$$\overline{d_{i,j,k}} = \frac{d_{i,j,k}}{\sum_{r,s=1}^{n} d_{r,s,k}}. \tag{11}$$

Proof. The definition of $\overline{d_{i,j,k}}$ implies that:

$$\overline{d_{i,j,k}} = \mathcal{P}(A_iA_j \mid C_k).$$

Applying Bayes' theorem, we have:

$$\overline{d_{i,j,k}} = \frac{\mathcal{P}(C_k \mid A_iA_j)\mathcal{P}(A_iA_j)}{\mathcal{P}(C_k)}.$$

Now, substituting (9), we obtain:

$$\overline{d_{i,j,k}} = \frac{\mathcal{P}(C_k \mid A_iA_j)\mathcal{P}(A_iA_j)}{\sum_{r,s=1}^{n} \mathcal{P}(C_k \mid A_rA_s)\mathcal{P}(A_rA_s)}, \tag{12}$$

which is equivalent to:

$$\overline{d_{i,j,k}} = \frac{d_{i,j,k} \cdot \mathcal{P}(A_iA_j)}{\sum_{r,s=1}^{n} d_{r,s,k} \cdot \mathcal{P}(A_rA_s)}. \tag{13}$$

By invoking the assumption in (10), we can cancel the terms $\mathcal{P}(A_iA_j)$ and $\mathcal{P}(A_rA_s)$ in the fraction, thereby arriving at (11). □

We refer the reader to J. Tian and B. L. Li [18] (2004) or I. Paniello [13] (2017) for further applications of this lemma, particularly within the theory of coalgebras. For general information on coalgebras, see the book by M. Sweedler [17] (1969).

Finally, we are ready to prove our main theorem:

Theorem 2.7. *Consider a gene with n alleles $\{A_i\}_{i=1}^n$ and m phenotypes (traits) $\{C_k\}_{k=1}^m$. If $(\mathcal{F}_m, \nabla)$ is the phenotype algebra, then:*

$$\nabla = \mathbf{D} \circ \left((\Gamma \circ \mathbf{D}^\dagger) \otimes (\Gamma \circ \mathbf{D}^\dagger) \right). \tag{14}$$

Note that, according to Remarks 2.1 and 2.4, Theorem 2.7 implies that the map ∇ is stochastic. Since the gametic algebra is commutative, if the dominance map $\mathbf{D}$ satisfies $d_{i,j,k} = d_{j,i,k}$ for every pair (i, j), the phenotype algebra is also commutative. In such cases, the phenotype algebra constitutes an AGR.

Proof. Let i and j be indices in $\{1, \ldots, m\}$, and suppose we have two parents I and J with phenotypes C_i and C_j, respectively. The structure constant $p_{i,j,k}$ represents the probability that the offspring K exhibits phenotype C_k. Thus, we have:

$$p_{i,j,k} = \sum_{p,q=1}^{n} d_{p,q,k} z_{p,q}^{i,j}, \tag{15}$$

where $d_{p,q,k}$ is the probability that an individual with genotype A_pA_q has phenotype C_k, and $z_{p,q}^{i,j}$ is the probability that the offspring K has genotype A_pA_q (knowing that the phenotype of I is C_i and the phenotype J is C_j). Note that the coefficients $d_{p,q,k}$ are the entries of the dominance map $\mathbf{D}$. If we define the element $z^{i,j}$ in the tensor product by:

$$z^{i,j} = \sum_{p,q=1}^{n} z_{p,q}^{i,j} \mathbf{A}_p \otimes \mathbf{A}_q, \tag{16}$$

then identities (15) and (16) imply $\nabla(\mathbf{C}_i \otimes \mathbf{C}_j) = \mathbf{D}(z^{i,j})$. The definition of $z_{p,q}^{i,j}$ further implies that $z_{p,q}^{i,j} = a_p^i b_q^j$, where a_p^i is the probability that parent I contributes allele A_p during reproduction knowing that the phenotype of I is C_i, and b_q^j is the probability that parent J contributes allele A_q during reproduction knowing that the phenotype of J is C_j .

Defining $a^i = \sum_{p=1}^n a_p^i \mathbf{A}_p$ and $b^j = \sum_{q=1}^n b_q^j \mathbf{A}_q$, identity (16) implies that $z^{i,j} = a^i \otimes b^j$. The definition of a_p^i gives:

$$a_p^i = \sum_{r,s=1}^{n} \gamma_{r,s,p} g_{r,s}^i,$$

where $g_{r,s}^i$ is the probability that parent I has genotype A_rA_s with the assumption it has phenotype C_i, and $\gamma_{r,s,p}$ is the probability that a zygote with genotype A_rA_s

contributes allele A_p during meiosis (i.e., the structure constant of the gametic algebra). Defining $g^i = \sum_{r,s} g^i_{r,s} \mathbf{A}_r \otimes \mathbf{A}_s$, we obtain $a^i = \Gamma(g^i)$.

Since we know that parent I has phenotype C_i, it follows from Lemma 2.6 that $g^i_{r,s} = \overline{d_{r,s,i}}$. Hence, identity (11) implies:

$$g^i_{r,s} = \frac{d_{r,s,i}}{\sum_{t,u=1}^{n} d_{t,u,i}},$$

where $\{d_{t,u,i}\}_{t,u=1}^{n}$ are the entries of the dominance map $\mathbf{D}$ corresponding to phenotype C_i. We conclude that g^i is the i^{th}-column of $\mathbf{D}^\dagger$ i.e., $g^i = \mathbf{D}^\dagger(\mathbf{C_i})$, hence $a^i = \Gamma(\mathbf{D}^\dagger(\mathbf{C}_i))$.

Applying the same analysis to parent J, we obtain $b^j = \Gamma(\mathbf{D}^\dagger(\mathbf{C}_j))$. Finally, we have:

$$\nabla(\mathbf{C}_i \otimes \mathbf{C}_j) = \mathbf{D}(z^{i,j}) = \mathbf{D}(a^i \otimes b^j) = \mathbf{D}\Big(\Gamma(\mathbf{D}^\dagger(\mathbf{C}_i)) \otimes \Gamma(\mathbf{D}^\dagger(\mathbf{C}_j))\Big),$$

which completes the proof. □

The composition path can be visualized in the following diagrams:

$$\begin{array}{ccc} \mathcal{F} & \xrightarrow{\mathbf{D}^\dagger} & \mathcal{G} \otimes \mathcal{G} \\ & \searrow^{\Gamma \circ \mathbf{D}^\dagger} & \downarrow \Gamma \\ & & \mathcal{G} \end{array} \qquad \begin{array}{ccc} \mathcal{F} \otimes \mathcal{F} & \xrightarrow{(\Gamma \circ \mathbf{D}^\dagger)^{\otimes 2}} & \mathcal{G} \otimes \mathcal{G} \\ & \searrow^{\nabla} & \downarrow \mathbf{D} \\ & & \mathcal{F} \end{array}$$

3. Reduction to the symmetric tensor space

The construction presented above does not assume the commutativity of the algebra; however, in most examples, this condition is satisfied. In such cases, we can replace the tensor product $\mathcal{V} \otimes \mathcal{V}$ of a space with itself by the *symmetric tensor space* $S^2(\mathcal{V})$, as defined in (8). This replacement is significant because if $\dim(\mathcal{V}) = n$, then $\dim(\mathcal{V} \otimes \mathcal{V}) = n^2$, whereas $\dim(S^2(\mathcal{V})) = \frac{n(n+1)}{2}$. Consequently, the matrices discussed in Section 4 will be smaller.

Although this may not represent a major computational improvement for very large n (since both dimensions grow quadratically), the reduction is useful in low-dimensional cases. This is evident in the three examples provided in this article; for instance, the number of columns in matrix (27) in Remark 4.3 was reduced from 16 to 10. In conclusion, the dimensions are smaller but remain quadratic. This section is devoted to relating the matrices of operators on the tensor product to those of the corresponding operators on the symmetric tensor space.

Consider two vector spaces $\mathcal{V}$ and $\mathcal{W}$. If a map $M : \mathcal{V} \otimes \mathcal{V} \to \mathcal{W}$ satisfies $M(u \otimes v) = M(v \otimes u)$ for every $u, v \in \mathcal{V}$, there exists a unique map $\overline{M} : S^2(\mathcal{V}) \to \mathcal{W}$ such that the following diagram commutes:

$$\begin{array}{ccc} \mathcal{V}\otimes\mathcal{V} & \xrightarrow{\rho} & S^2(\mathcal{V}) \\ & \searrow^{M} & \downarrow{\scriptstyle \overline{M}} \\ & & \mathcal{W} \end{array}$$

where ρ is the canonical projection $\rho : \mathcal{V}\otimes\mathcal{V} \to S^2(\mathcal{V})$, defined by $\rho(t) = t + \mathcal{I}$, where $\mathcal{I}$ is the subspace spanned by $\{v\otimes w - w\otimes v \mid v, w \in \mathcal{V}\}$. Note that if $\{v_i\}_{i=1}^n$ is a basis of $\mathcal{V}$, then $\{v_i \otimes v_j\}_{i,j=1}^n$ is a basis of $\mathcal{V}\otimes\mathcal{V}$, and $\{v_i \cdot v_j\}_{1\le i\le j\le n}$—where $v_i \cdot v_j = \rho(v_i \otimes v_j)$—is a basis of $S^2(\mathcal{V})$.

Conversely, if we have a map $L : \mathcal{W} \to \mathcal{V}\otimes\mathcal{V}$ satisfying $\sigma \circ L = L$ (where σ is the *switch map* $\sigma(u\otimes v) = v\otimes u$), there exists a unique map $\overline{L} : \mathcal{W} \to S^2(\mathcal{V})$ such that the following diagram commutes:

$$\begin{array}{ccc} \mathcal{W} & \xrightarrow{L} & \mathcal{V}\otimes\mathcal{V} \\ & \searrow^{\overline{L}} & \downarrow{\scriptstyle \rho} \\ & & S^2(\mathcal{V}) \end{array}$$

The following two propositions relate the matrices of the previously described maps.

Proposition 3.1. *Consider two vector spaces $\mathcal{V}$ and $\mathcal{W}$ with respective bases $\mathcal{B} = \{v_i\}_{i=1}^n$ and $\mathcal{C} = \{w_j\}_{j=1}^m$. Let $M : \mathcal{V}\otimes\mathcal{V} \to \mathcal{W}$ be a map satisfying $M(u\otimes v) = M(v\otimes u)$ for every $u, v \in \mathcal{V}$. Let $\overline{M}$ be the map associated with M as described above, where*

$$M(v_i \otimes v_j) = \sum_{k=1}^m \mu_{i,j,k} w_k \quad \text{and} \quad \overline{M}(v_i \cdot v_j) = \sum_{k=1}^m \overline{\mu}_{i,j,k} w_k.$$

Then, $\mu_{i,j,k} = \overline{\mu}_{i,j,k}$ for every triplet (i, j, k).

Proof. Since $M = \overline{M}\circ\rho$, for every pair (i, j) we have $M(v_i\otimes v_j) = \overline{M}(\rho(v_i\otimes v_j)) = \overline{M}(v_i \cdot v_j)$. Therefore,

$$\sum_{k=1}^m \mu_{i,j,k} w_k = \sum_{k=1}^m \overline{\mu}_{i,j,k} w_k.$$

Comparing the coefficients of the basis $\mathcal{C}$ yields the desired identity. □

This implies that, in the chosen bases, obtaining the matrix of $\overline{M}$ simply requires removing the redundant columns (j, i) for $i < j$ from the matrix of M.

We now address the case where the tensor product is the codomain of the map.

Proposition 3.2. *Consider two vector spaces $\mathcal{V}$ and $\mathcal{W}$ with respective bases $\mathcal{B} = \{v_i\}_{i=1}^n$ and $\mathcal{C} = \{w_j\}_{j=1}^m$. Let $L : \mathcal{W} \to \mathcal{V}\otimes\mathcal{V}$ be a map satisfying $\sigma \circ L = L$. Let*

$\overline{L}$ be the map defined above such that

$$L(w_k) = \sum_{i,j=1}^{n} \lambda_{i,j,k} v_i \otimes v_j \quad \textit{and} \quad \overline{L}(w_k) = \sum_{1 \le i \le j \le n} \overline{\lambda}_{i,j,k} v_i \cdot v_j.$$

Then, $\lambda_{i,j,k} + \lambda_{j,i,k} = \overline{\lambda}_{i,j,k}$ for $i < j$, and $\lambda_{i,i,k} = \overline{\lambda}_{i,i,k}$ for every pair (i,k).

Proof. Since $\rho \circ L = \overline{L}$, for every index k we have $\overline{L}(w_k) = \rho(L(w_k))$. Thus,

$$\sum_{i \le j} \overline{\lambda}_{i,j,k} v_i \cdot v_j = \rho \left(\sum_{i,j=1}^{n} \lambda_{i,j,k} v_i \otimes v_j \right).$$

Expanding the right side, we obtain:

$$\begin{aligned} \rho(L(w_k)) &= \rho \left(\sum_i \lambda_{i,i,k} v_i \otimes v_i + \sum_{i<j} \lambda_{i,j,k} v_i \otimes v_j + \sum_{j<i} \lambda_{i,j,k} v_i \otimes v_j \right) \\ &= \sum_i \lambda_{i,i,k} v_i \cdot v_i + \sum_{i<j} \lambda_{i,j,k} v_i \cdot v_j + \sum_{i<j} \lambda_{j,i,k} v_j \cdot v_i. \end{aligned}$$

Given that $v_i \cdot v_j = v_j \cdot v_i$, we have:

$$\overline{L}(w_k) = \sum_i \lambda_{i,i,k} v_i \cdot v_i + \sum_{i<j} (\lambda_{i,j,k} + \lambda_{j,i,k}) v_i \cdot v_j.$$

The result follows from comparing the coefficients. □

Consequently, to obtain the matrix of $\overline{L}$ in the given bases, it is sufficient to add row (j,i) to row (i,j) for every $i < j$ in the matrix of L, and then remove the redundant rows.

Remark 3.3. When both the domain and codomain are tensor products—for instance, in the matrix of $(\Gamma \circ \mathbf{D}^\dagger) \otimes (\Gamma \circ \mathbf{D}^\dagger)$ acting from $S^2(\mathcal{F}_m)$ to $S^2(\mathcal{F}_m)$—it is necessary to perform both procedures: eliminate columns (j,i) and add rows (j,i) to rows (i,j) for all $i < j$.

4. Examples

In this first example, we present the maps defined on both the full tensor product and the symmetric tensor space (the former denoted without an overline and the latter with an overline). In all subsequent examples, we shall work exclusively within the symmetric space; that is, using the overlined operators.

Example 4.1. Consider a gene with two alleles: one dominant, A, and one recessive, a. In this case, $\mathcal{G}_2 = \mathcal{F}_2 = \operatorname{span}\{\mathbf{A}, \mathbf{a}\}$. The dominance map is defined by

$\mathbf{D}(\mathbf{A}\otimes\mathbf{A}) = \mathbf{D}(\mathbf{A}\otimes\mathbf{a}) = \mathbf{D}(\mathbf{a}\otimes\mathbf{A}) = \mathbf{A}$ and $\mathbf{D}(\mathbf{a}\otimes\mathbf{a}) = \mathbf{a}$. Passing to the symmetric space, we obtain $\overline{\mathbf{D}}(\mathbf{AA}) = \mathbf{A}$, $\overline{\mathbf{D}}(\mathbf{Aa}) = \mathbf{A}$, and $\overline{\mathbf{D}}(\mathbf{aa}) = \mathbf{a}$. Consequently, the respective matrices in the canonical bases are:

$$\mathbf{D} = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \overline{\mathbf{D}} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{17}$$

Let Γ be the linear map defining the product in the gametic algebra $\mathcal{G}_2$ for two alleles under simple Mendelian inheritance, as defined in (7) with $\mathbf{A}_1 = \mathbf{A}$ and $\mathbf{A}_2 = \mathbf{a}$, i.e. $\Gamma : \mathcal{G}_2 \otimes \mathcal{G}_2 \to \mathcal{G}_2$ is given by:

$$\Gamma(\mathbf{A}\otimes\mathbf{A}) = \mathbf{A}, \quad \Gamma(\mathbf{a}\otimes\mathbf{a}) = \mathbf{a}, \quad \Gamma(\mathbf{A}\otimes\mathbf{a}) = \Gamma(\mathbf{a}\otimes\mathbf{A}) = \frac{1}{2}(\mathbf{A}+\mathbf{a}). \tag{18}$$

So the matrices of Γ and $\overline{\Gamma}$ in the natural basis are:

$$\Gamma = \begin{pmatrix} 1 & 1/2 & 1/2 & 0 \\ 0 & 1/2 & 1/2 & 1 \end{pmatrix}, \quad \overline{\Gamma} = \begin{pmatrix} 1 & 1/2 & 0 \\ 0 & 1/2 & 1 \end{pmatrix}. \tag{19}$$

The other associated matrices are then:

$$\mathbf{D}^* = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \overline{\mathbf{D}^*} = \begin{pmatrix} 1 & 0 \\ 2 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathbf{D}^\dagger = \begin{pmatrix} 1/3 & 0 \\ 1/3 & 0 \\ 1/3 & 0 \\ 0 & 1 \end{pmatrix}, \quad \overline{\mathbf{D}^\dagger} = \begin{pmatrix} 1/3 & 0 \\ 2/3 & 0 \\ 0 & 1 \end{pmatrix}.$$

Computing the composition, we find:

$$\Gamma \circ \mathbf{D}^\dagger = \overline{\Gamma} \circ \overline{\mathbf{D}^\dagger} = \begin{pmatrix} 2/3 & 0 \\ 1/3 & 1 \end{pmatrix}.$$

The matrix $\overline{\Gamma} \circ \overline{\mathbf{D}^\dagger}$ indicates that a parent with the dominant trait has a $2/3$ probability of contributing a dominant gamete and a $1/3$ probability of contributing a recessive one, while a parent with the recessive trait will certainly contribute a recessive gamete. Computing the tensor product, we have:

$$(\Gamma \circ \mathbf{D}^\dagger) \otimes (\Gamma \circ \mathbf{D}^\dagger) = \begin{pmatrix} 4/9 & 0 & 0 & 0 \\ 2/9 & 2/3 & 0 & 0 \\ 2/9 & 0 & 2/3 & 0 \\ 1/9 & 1/3 & 1/3 & 1 \end{pmatrix}.$$

By applying the reduction to the symmetric space, we obtain:

$$\overline{(\Gamma \circ \mathbf{D}^\dagger) \otimes (\Gamma \circ \mathbf{D}^\dagger)} = \begin{pmatrix} 4/9 & 0 & 0 \\ 4/9 & 2/3 & 0 \\ 1/9 & 1/3 & 1 \end{pmatrix}. \tag{20}$$

Finally, composing with $\mathbf{D}$, we arrive at:

$$\overline{\nabla} = \overline{\mathbf{D} \circ ((\Gamma \circ \mathbf{D}^\dagger) \otimes (\Gamma \circ \mathbf{D}^\dagger))} = \begin{pmatrix} 8/9 & 2/3 & 0 \\ 1/9 & 1/3 & 1 \end{pmatrix}. \tag{21}$$

This implies that the multiplication table for the phenotype algebra is:

$$\mathbf{A} \cdot \mathbf{A} = \frac{8}{9}\mathbf{A} + \frac{1}{9}\mathbf{a}, \quad \mathbf{A} \cdot \mathbf{a} = \frac{2}{3}\mathbf{A} + \frac{1}{3}\mathbf{a}, \quad \mathbf{a} \cdot \mathbf{a} = \mathbf{a}. \tag{22}$$

In biological terms, if a gene follows simple Mendelian laws with one dominant and one recessive allele, two parents with the dominant phenotype have an 8/9 probability of producing offspring with the dominant trait and a 1/9 probability of producing offspring with the recessive one. If one parent exhibits the dominant phenotype and the other the recessive one, the offspring has a 2/3 probability of being dominant and a 1/3 probability of being recessive. Finally, if both parents have the recessive phenotype, the offspring will necessarily exhibit the recessive trait.

Example 4.2. Consider the gene responsible for the human ABO blood group system. This gene has three alleles, A, B, and O, alleles A and B are associated to the presence of two antigens called A and B in the blood, while the allele O is associated to the absence of both antigens. The alleles A and B are dominant and the allele O is recessive. The result is four different phenotypes: Individuals with only the antigen A in the blood have blood type A or (I) and their genotype can be AA or AO, individuals with only the antigen B in the blood have blood type B or (II), and their genotype can be BB or BO, individuals with both antigens in the blood have blood type AB or (III) and their genotype is AB, finally, individuals without any of this two antigens have blood type O or (IV) and their genotype is OO. Thus, $\mathcal{G} = \text{span}\{\mathbf{A}, \mathbf{B}, \mathbf{O}\}$, while the phenotype space is $\mathcal{F} = \text{span}\{\mathbf{A}, \mathbf{B}, \mathbf{AB}, \mathbf{O}\}$. We consider the symmetric space

$$S^2(\mathcal{G}) = span\{\mathbf{AA}, \mathbf{AB}, \mathbf{AO}, \mathbf{BB}, \mathbf{BO}, \mathbf{OO}\}.$$

The dominance map is given by:

$$\overline{\mathbf{D}}(\mathbf{AA}) = \overline{\mathbf{D}}(\mathbf{AO}) = \mathbf{A}, \quad \overline{\mathbf{D}}(\mathbf{BB}) = \overline{\mathbf{D}}(\mathbf{BO}) = \mathbf{B},$$
$$\overline{\mathbf{D}}(\mathbf{AB}) = \mathbf{AB}, \quad \text{and} \quad \overline{\mathbf{D}}(\mathbf{OO}) = \mathbf{O}.$$

Its matrix relative to the given bases is:

$$\overline{\mathbf{D}} = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}. \tag{23}$$

To determine the matrix of $\overline{\Gamma}$ we also assume simple Mendelian law, i.e., we have $\overline{\Gamma}(\mathbf{XY}) = \frac{1}{2}(\mathbf{X} + \mathbf{Y})$, for every pair $\mathbf{X}, \mathbf{Y}$ of elements of $\{\mathbf{A}, \mathbf{B}, \mathbf{O}\}$. Therefore:

$$\overline{\Gamma} = \begin{pmatrix} 1 & 1/2 & 1/2 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 1 & 1/2 & 0 \\ 0 & 0 & 1/2 & 0 & 1/2 & 1 \end{pmatrix}. \tag{24}$$

Following the same procedure as in the previous example, the associated matrices are:

$$\overline{\mathbf{D}^\dagger} = \begin{pmatrix} 1/3 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 2/3 & 0 & 0 & 0 \\ 0 & 1/3 & 0 & 0 \\ 0 & 2/3 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \overline{\Gamma \circ \mathbf{D}^\dagger} = \begin{pmatrix} 2/3 & 0 & 1/2 & 0 \\ 0 & 2/3 & 1/2 & 0 \\ 1/3 & 1/3 & 0 & 1 \end{pmatrix}. \tag{25}$$

Finally, by computing the product using Theorem 2.7 and the reduction to the symmetric space, we obtain the matrix

$$\overline{\nabla} = \begin{pmatrix} 8/9 & 2/9 & 1/2 & 2/3 & 0 & 1/6 & 0 & 1/4 & 1/2 & 0 \\ 0 & 2/9 & 1/6 & 0 & 8/9 & 1/2 & 2/3 & 1/4 & 1/2 & 0 \\ 0 & 4/9 & 1/3 & 0 & 0 & 1/3 & 0 & 1/2 & 0 & 0 \\ 1/9 & 1/9 & 0 & 1/3 & 1/9 & 0 & 1/3 & 0 & 0 & 1 \end{pmatrix},$$

and the multiplication table for this algebra:

- $\mathbf{A} \cdot \mathbf{A} = \frac{8}{9}\mathbf{A} + \frac{1}{9}\mathbf{O}$
- $\mathbf{A} \cdot \mathbf{B} = \frac{2}{9}\mathbf{A} + \frac{2}{9}\mathbf{B} + \frac{4}{9}\mathbf{AB} + \frac{1}{9}\mathbf{O}$
- $\mathbf{A} \cdot (\mathbf{AB}) = \frac{1}{2}\mathbf{A} + \frac{1}{6}\mathbf{B} + \frac{1}{3}\mathbf{AB}$
- $\mathbf{A} \cdot \mathbf{O} = \frac{2}{3}\mathbf{A} + \frac{1}{3}\mathbf{O}$
- $\mathbf{B} \cdot \mathbf{B} = \frac{8}{9}\mathbf{B} + \frac{1}{9}\mathbf{O}$
- $\mathbf{B} \cdot (\mathbf{AB}) = \frac{1}{6}\mathbf{A} + \frac{1}{2}\mathbf{B} + \frac{1}{3}\mathbf{AB}$
- $\mathbf{B} \cdot \mathbf{O} = \frac{2}{3}\mathbf{B} + \frac{1}{3}\mathbf{O}$
- $(\mathbf{AB}) \cdot (\mathbf{AB}) = \frac{1}{4}\mathbf{A} + \frac{1}{4}\mathbf{B} + \frac{1}{2}\mathbf{AB}$
- $(\mathbf{AB}) \cdot \mathbf{O} = \frac{1}{2}\mathbf{A} + \frac{1}{2}\mathbf{B}$
- $\mathbf{O} \cdot \mathbf{O} = \mathbf{O}$

From this table, we can deduce the probability distribution of offspring traits based on parental phenotypes. For instance, a child of a couple such that one of the parents has blood type A and the other one has type B, has 2/9 probability of having blood type A, 2/9 blood type B, 4/9 blood type AB and 1/9 blood type

O. Furthermore, the table shows that two parents with blood type AB have a $1/4$ probability of having a child with type A, $1/4$ with type B, and $1/2$ with type AB, but they cannot have a child with type O.

Remark 4.3. In the 2016 article [4] by J.M. Casas, M. Ladra, B.A. Omirov, and R. Turdibaev, the algebra of human blood types is defined by explicitly giving the multiplication in the basis $\{\mathbf{A}, \mathbf{B}, \mathbf{AB}, \mathbf{O}\}$. In their model, the structure constants represent the probability of an offspring having a specific blood type for every parental configuration. These constants depend on a coefficient α, which corresponds to the probability that an individual with type A or B contributes an O allele during meiosis, and a coefficient β, representing the probability that an individual with blood type AB contributes an A allele. In our approach, the interpretation of these coefficients yields:

$$\overline{\Gamma \circ \mathbf{D}^{\dagger}} = \begin{pmatrix} 1-\alpha & 0 & \beta & 0 \\ 0 & 1-\alpha & 1-\beta & 0 \\ \alpha & \alpha & 0 & 1 \end{pmatrix}. \tag{26}$$

By computing its symmetric tensor square and multiplying by the dominance matrix (23), we obtain a matrix $\overline{\nabla}$ whose entries coincide exactly with the structure constants presented in [4] (see the table above Definition 2.1).

$$\begin{pmatrix} 1-\alpha^2 & \alpha(1-\alpha) & \beta & (1-\alpha) & 0 & \alpha\beta & 0 & \beta^2 & \beta & 0 \\ 0 & \alpha(1-\alpha) & \alpha(1-\beta) & 0 & 1-\alpha^2 & 1-\beta & 1-\alpha & (1-\beta)^2 & 1-\beta & 0 \\ 0 & (1-\alpha)^2 & (1-\alpha)(1-\beta) & 0 & 0 & (1-\alpha)\beta & 0 & 2\beta(1-\beta) & 0 & 0 \\ \alpha^2 & \alpha^2 & 0 & \alpha & \alpha^2 & 0 & \alpha & 0 & 0 & 1 \end{pmatrix} \tag{27}$$

Remark 4.4. Given that the dominance map is fixed as in (23), the only way to obtain a matrix such as (26) is to establish arbitrary coefficients for the gametic algebra. Let

$$\overline{\Gamma} = \begin{pmatrix} 1 & r & 1-s & 0 & 0 & 0 \\ 0 & 1-r & 0 & 1 & 1-t & 0 \\ 0 & 0 & s & 0 & t & 1 \end{pmatrix},$$

where r is the probability that an AB zygote contributes allele A, s is the probability that an AO zygote contributes allele O, and t is the probability that a BO zygote contributes allele O. By computing $\overline{\Gamma \circ \mathbf{D}^{\dagger}}$, we obtain (26) with $\alpha = \frac{2t}{3} = \frac{2r}{3}$ and $\beta = s$.

The reason β coincides with s is that the probability of an AB zygote contributing allele A is identical to the probability of an individual with blood type AB

contributing the same allele. However, this differs for the coefficient α because an individual with blood type A can have two possible genotypes: AA or AO.

Here we observe a key distinction between our approach and the one in the cited reference. They state that under Mendelian laws (i.e., when $s = t = r = 1/2$, as in our example), the value of α is $1/4$. However, comparing (25) with (26), we find that $\alpha = 1/3$. The reasoning for $\alpha = 1/4$ assumes that an individual with blood type A is equally likely to have genotype AA as AO. In our approach, we posit that the genotype AO is twice as probable as AA, since we distinguish between the cases where the allele O originates from the father versus the mother.

Example 4.5. In the Peruvian Hairless Dog, the trait determining whether an individual is "coated" (hairy) or "hairless" is governed by the *FOXI3* gene. This follows a pattern similar to Example 4.1, where the dominant allele A corresponds to the hairless phenotype and the recessive allele a to the coated one. However, a key difference is that a zygote with two dominant alleles (AA) is inviable and dies *in utero*. We treat these deceased individuals as a separate phenotype, denoted by b.

Therefore, while $\mathcal{G} = \operatorname{span}\{\mathbf{A}, \mathbf{a}\}$, the phenotype space is $\mathcal{F} = \operatorname{span}\{\mathbf{A}, \mathbf{a}, \mathbf{b}\}$. The dominance map is given by $\overline{\mathbf{D}}(\mathbf{AA}) = \mathbf{b}$, $\overline{\mathbf{D}}(\mathbf{Aa}) = \mathbf{A}$, and $\overline{\mathbf{D}}(\mathbf{aa}) = \mathbf{a}$. Its matrix in the canonical bases is:

$$\overline{\mathbf{D}} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$$

Again we assume simple Mendelian law, hence the matrix $\overline{\Gamma}$ is given by (19) in Example 4.1. It follows that:

$$\overline{\mathbf{D}^\dagger} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \quad \text{and} \quad \overline{\Gamma \circ \mathbf{D}^\dagger} = \begin{pmatrix} 1/2 & 0 & 1 \\ 1/2 & 1 & 0 \end{pmatrix}.$$

Thus, we compute:

$$\overline{(\Gamma \circ \mathbf{D}^\dagger) \otimes (\Gamma \circ \mathbf{D}^\dagger)} = \begin{pmatrix} 1/4 & 0 & 1/2 & 0 & 0 & 1 \\ 1/2 & 1/2 & 1/2 & 0 & 1 & 0 \\ 1/4 & 1/2 & 0 & 1 & 0 & 0 \end{pmatrix},$$

$$\text{implying} \quad \overline{\nabla} = \overline{\mathbf{D} \circ ((\Gamma \circ \mathbf{D}^\dagger) \otimes (\Gamma \circ \mathbf{D}^\dagger))} = \begin{pmatrix} 1/2 & 1/2 & 1/2 & 0 & 1 & 0 \\ 1/4 & 1/2 & 0 & 1 & 0 & 0 \\ 1/4 & 0 & 1/2 & 0 & 0 & 1 \end{pmatrix}.$$

The resulting multiplication table includes terms such as $\mathbf{b}\cdot\mathbf{b}=\mathbf{b}$ and $\mathbf{A}\cdot\mathbf{b}=\frac{1}{2}\mathbf{A}+\frac{1}{2}\mathbf{b}$. Mathematically, the computation is consistent; for instance, "crossing" a hairless dog (Aa) with a "dead" dog (AA) would yield offspring with a $1/2$ probability of being Aa (hairless) and a $1/2$ probability of being AA (dead).

However, since deceased individuals cannot reproduce, these columns (3, 5, and 6) lack biological meaning for a population algebra. To obtain a functional model for the living population, we first remove the columns involving phenotype b:

$$\begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/4 & 1/2 & 1 \\ 1/4 & 0 & 0 \end{pmatrix}.$$

To restore the structure of a baric algebra, we must eliminate the last row (corresponding to the inviable phenotype b) and re-normalize each column so that the entries sum to 1:

$$\overline{\nabla}=\begin{pmatrix} 2/3 & 1/2 & 0 \\ 1/3 & 1/2 & 1 \end{pmatrix}. \tag{28}$$

The final multiplication table for the living population is:

$$\mathbf{A}\cdot\mathbf{A}=\frac{2}{3}\mathbf{A}+\frac{1}{3}\mathbf{a},\quad \mathbf{A}\cdot\mathbf{a}=\frac{1}{2}\mathbf{A}+\frac{1}{2}\mathbf{a},\quad \mathbf{a}\cdot\mathbf{a}=\mathbf{a}.$$

This means that when two hairless dogs (A) mate, the viable offspring has a $2/3$ probability of being hairless and a $1/3$ probability of being coated. A cross between a hairless and a coated dog results in a $1/2$ probability for each phenotype, while two coated dogs will only produce coated offspring.

5. A skew symmetric matrix for incomplete dominance

In the three examples presented in the previous section, the matrix of the dominance map contains only ones and zeros. This occurs because a specific genotype unambiguously produces a specific phenotype. For instance, in all three cases, if both parents exhibit a trait related to a recessive allele (such as a in Examples 4.1 and 4.5, or O in Example 4.2), it is certain that the offspring will also exhibit this characteristic. Furthermore, to compute the map Γ, we assumed that both alleles in a zygote have an equal probability of appearing in the resulting gamete. This corresponds to a gametic algebra associated with the null matrix.

In the general case, a gametic algebra has a multiplication as defined in (5) for a skew-symmetric matrix $\mathbf{S}$. We denote by $\overline{\Gamma}_{\mathbf{S}}$ the factorization—via the symmetric tensor product—of the multiplication in the gametic algebra defined by $\mathbf{S}$. We now analyze the case of *incomplete dominance*. Suppose that for a specific gene,

the phenotype of an individual corresponds to one of the alleles of its genotype according to some probability distribution. We introduce a skew-symmetric matrix $\mathbf{T}$ such that the dominance map $\mathbf{D}$ satisfies:

$$\mathbf{D}(\mathbf{A}_i \otimes \mathbf{A}_j) = \mathbf{D}(\mathbf{A}_j \otimes \mathbf{A}_i) = \left(\frac{1}{2} + \mathbf{T}_{i,j}\right)\mathbf{C}_i + \left(\frac{1}{2} + \mathbf{T}_{j,i}\right)\mathbf{C}_j.$$

Under this assumption, the genotype space $\mathcal{G}$ coincides with the phenotype space $\mathcal{F}$. Thus, the pair $(\mathcal{G}, \mathbf{D})$ becomes the *dominance algebra*, which is itself an LVA. We denote such a dominance map as $\mathbf{D}_\mathbf{T}$ to indicate its dependence on the matrix $\mathbf{T}$; consequently, $\mathbf{D}^\dagger_\mathbf{T}$ also depends on $\mathbf{T}$.

Therefore, for a gene with n alleles where the gametic algebra is a Lotka-Volterra algebra (LVA) defined by the matrix $\mathbf{S}$, and the dominance algebra is an LVA defined by the matrix $\mathbf{T}$, the phenotype algebra is given by $(\mathcal{F}, \nabla_{\mathbf{S},\mathbf{T}})$, where:

$$\nabla_{\mathbf{S},\mathbf{T}} = \mathbf{D}_\mathbf{T} \circ \left((\Gamma_\mathbf{S} \circ \mathbf{D}^\dagger_\mathbf{T}) \otimes (\Gamma_\mathbf{S} \circ \mathbf{D}^\dagger_\mathbf{T})\right). \tag{29}$$

Suppose we have a gene with two alleles such that a zygote with genotype A_1A_2 has a probability s of contributing the allele A_2. This implies that the probability of contributing allele A_1 is $1 - s$. Consequently, the skew-symmetric matrix $\mathbf{S}$ and the matrix $\overline{\Gamma_\mathbf{S}}$ of the gametic algebra are:

$$\mathbf{S} = \begin{pmatrix} 0 & \frac{1}{2} - s \\ s - \frac{1}{2} & 0 \end{pmatrix}, \overline{\Gamma_\mathbf{S}} = \begin{pmatrix} 1 & 1-s & 0 \\ 0 & s & 1 \end{pmatrix}.$$

Now, assume that every individual with genotype A_1A_2 has a small probability t of exhibiting phenotype A_2. Then the skew-symmetric matrix $\mathbf{T}$ and the matrix $\overline{\mathbf{D}_\mathbf{T}}$ of the dominance algebra are:

$$\mathbf{T} = \begin{pmatrix} 0 & \frac{1}{2} - t \\ t - \frac{1}{2} & 0 \end{pmatrix}, \quad \overline{\mathbf{D}_\mathbf{T}} = \begin{pmatrix} 1 & 1-t & 0 \\ 0 & t & 1 \end{pmatrix}, \quad \text{and} \quad \overline{\mathbf{D}^\dagger_\mathbf{T}} = \begin{pmatrix} \frac{1}{3-2t} & 0 \\ \frac{2(1-t)}{3-2t} & \frac{2t}{1+2t} \\ 0 & \frac{1}{1+2t} \end{pmatrix}.$$

By defining the following polynomials in the variables s and t:

$$\begin{aligned} a(s,t) &= \frac{1 + 2(1-s)(1-t)}{3-2t}, & b(s,t) &= \frac{2t(1-s)}{1+2t}, \\ c(s,t) &= \frac{2s(1-t)}{3-2t} & d(s,t) &= \frac{1+2st}{1+2t} \end{aligned} \tag{30}$$

we obtain:

$$\overline{\Gamma_\mathbf{S} \circ \mathbf{D}^\dagger_\mathbf{T}} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Therefore, the matrix $\overline{(\Gamma_{\mathbf{S}} \circ \mathbf{D}_{\mathbf{T}}^{\dagger}) \otimes (\Gamma_{\mathbf{S}} \circ \mathbf{D}_{\mathbf{T}}^{\dagger})}$ is:

$$\begin{pmatrix} a^2 & ab & b^2 \\ 2ac & ad+bc & 2bd \\ c^2 & cd & d^2 \end{pmatrix}.$$

Finally, the multiplication in the phenotype algebra is given by:

$$\overline{\nabla_{\mathbf{S},\mathbf{T}}} = \begin{pmatrix} a^2 + 2(1-t)ac & ab + (1-t)(ad+bc) & b^2 + 2(1-t)bd \\ 2tac + c^2 & t(ad+bc) + cd & 2tbd + d^2 \end{pmatrix}. \tag{31}$$

To provide a concrete example, one could choose specific values for s and t and substitute them into (30) to find a, b, c, and d, subsequently using these results in (31) to obtain $\overline{\nabla_{\mathbf{S},\mathbf{T}}}$. However, it is computationally simpler to substitute s and t into the initial matrices and derive the subsequent operators. We illustrate this process as follows:

Example 5.1. Suppose a gene has two alleles, A_1 and A_2, with a corresponding pair of traits. Assume the probability that an A_1A_2 zygote produces a gamete with allele A_2 is $s = 1/5$. Furthermore, assume that an individual with genotype A_1A_2 has a probability $t = 1/10$ of exhibiting phenotype A_2.

The skew-symmetric matrices $\mathbf{S}$ and $\mathbf{T}$, along with the respective matrices for the gametic algebra and the dominance map, are:

$$\mathbf{S} = \begin{pmatrix} 0 & 3/10 \\ -3/10 & 0 \end{pmatrix}, \quad \mathbf{T} = \begin{pmatrix} 0 & 4/10 \\ -4/10 & 0 \end{pmatrix},$$

$$\overline{\Gamma_{\mathbf{S}}} = \begin{pmatrix} 1 & 4/5 & 0 \\ 0 & 1/5 & 1 \end{pmatrix}, \quad \overline{\mathbf{D}_{\mathbf{T}}} = \begin{pmatrix} 1 & 9/10 & 0 \\ 0 & 1/10 & 1 \end{pmatrix}.$$

Consequently, the derived matrices are:

$$\overline{\mathbf{D}_{\mathbf{T}}^{\dagger}} = \begin{pmatrix} 5/14 & 0 \\ 9/14 & 1/6 \\ 0 & 5/6 \end{pmatrix}, \quad \overline{\Gamma_{\mathbf{S}} \circ \mathbf{D}_{\mathbf{T}}^{\dagger}} = \begin{pmatrix} 61/70 & 2/15 \\ 9/70 & 13/15 \end{pmatrix},$$

$$\overline{(\Gamma_{\mathbf{S}} \circ \mathbf{D}_{\mathbf{T}}^{\dagger}) \otimes (\Gamma_{\mathbf{S}} \circ \mathbf{D}_{\mathbf{T}}^{\dagger})} = \begin{pmatrix} \frac{3721}{4900} & \frac{122}{1050} & \frac{4}{225} \\ \frac{1098}{4900} & \frac{811}{1050} & \frac{52}{225} \\ \frac{81}{4900} & \frac{117}{1050} & \frac{169}{225} \end{pmatrix}.$$

Finally, we obtain:

$$\overline{\nabla_{\mathbf{S},\mathbf{T}}} = \begin{pmatrix} \frac{47092}{49000} & \frac{8519}{10500} & \frac{508}{2250} \\ \frac{1908}{49000} & \frac{1981}{10500} & \frac{1742}{2250} \end{pmatrix}.$$

To clarify the biological implications of this matrix, we present the probabilities as percentages:

$$\overline{\nabla_{\mathbf{S},\mathbf{T}}} \approx \begin{pmatrix} 96.11\% & 81.13\% & 22.58\% \\ 3.89\% & 18.87\% & 77.42\% \end{pmatrix}.$$

In conclusion, if a gene has two alleles A_1, A_2, where an A_1A_2 zygote has an 80% probability of producing an A_1 gamete (since $s = 0.2$ for A_2) and an A_1A_2 individual has a 10% probability of exhibiting phenotype A_2, the trait distribution in the offspring is as follows:

- If both parents have phenotype A_1, the offspring has a 96.11% probability of having phenotype A_1 and a 3.89% probability of having phenotype A_2.
- If one parent has phenotype A_1 and the other has A_2, the offspring has an 81.13% probability of having phenotype A_1 and an 18.87% probability of having phenotype A_2.
- If both parents have phenotype A_2, the offspring has a 22.58% probability of having phenotype A_1 and a 77.42% probability of having phenotype A_2.

Acknowledgement. The author thanks the reviewers for their suggestions and improvements.

Disclosure statement. The author declares no conflicts of interest. Furthermore, the author confirms that there are no datasets associated with this article.

References

[1] M. Arenas, A. Labra and I. Paniello, *Lotka-Volterra coalgebras*, Linear Multilinear Algebra, 70(19) (2022), 4483-4497.

[2] M. Arenas, A. Labra and I. Paniello, *On the structure of Lotka-Volterra coalgebras*, Comm. Algebra, 52(6) (2024), 2386-2403.

[3] S. N. Bernstein, *Solution of a mathematical problem connected with the theory of heredity*, Ann. Math. Statistics, 13 (1942), 53-61.

[4] J. M. Casas, M. Ladra, B. A. Omirov and R. Turdibaev, *On the algebraic properties of human ABO-blood group inheritance pattern*, ANZIAM J., 58(1) (2016), 78-95.

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

[6] H. Gonshor, *Special train algebras arising in genetics*, Proc. Edinburgh Math. Soc., 12(1) (1960), 41-53.

[7] H. Gonshor, *Contributions to genetic algebras*, Proc. Edinburgh Math. Soc., 17(4) (1971), 289-298.

[8] J. C. Gutiérrez-Fernández and C. I. García, *Derivations of Lotka-Volterra algebras*, Sao Paulo J. Math. Sci., 13(1) (2019), 292-304.

[9] Y. Itoh, *Nonassociative algebra and Lotka-Volterra equation with ternary interaction*, Nonlinear Anal., 5(1) (1981), 53-56.

[10] A. Labra, M. Ladra and U. A. Rozikov, *An evolution algebra in population genetics*, Linear Algebra Appl., 457 (2014), 348-362.

[11] M. Ladra and U. A. Rozikov, *Evolution algebra of a bisexual population*, J. Algebra, 378 (2013), 153-172.

[12] G. Mendel, *Experiments in plant-hybridization*, in Classic Papers in Genetics, J. Peters, ed., Prentice-Hall, Inc. (1959), 1-20.

[13] I. Paniello, *Backwards genetic inheritance through coalgebra-graphs*, Linear Multilinear Algebra, 65(5) (2017), 943-961.

[14] L. A. Peresi, *The derivation algebra of gametic algebra for linked loci*, Math. Biosci., 91(2) (1988), 151-156.

[15] M. L. Reed, *Algebraic structure of genetic inheritance*, Bull. Amer. Math. Soc. (N.S.), 34(2) (1997), 107-130.

[16] R. D. Schafer, *Structure of genetic algebras*, Amer. J. Math., 71 (1949), 121-135.

[17] M. E. Sweedler, Hopf Algebras, Mathematics Lecture Note Series, W. A. Benjamin Inc., New York, 1969.

[18] J. Tian and B. L. Li, *Coalgebraic structure of genetic inheritance*, Math. Biosci. Eng., 1(2) (2004), 243-266.

[19] R. Varro, *Gonosomal algebra*, J. Algebra, 447 (2016), 1-30.

[20] A. Wörz-Busekros, Algebras in Genetics, Lecture Notes in Biomathematics, 36, Springer-Verlag, Berlin-Heidelberg-New York, 1980.

Manuel Arenas
Departamento De Matemática
Facultad De Ciencias Naturales Matemática y Del Medio Ambiente
Universidad Tecnológica Metropolitana
Santiago, Chile
e-mail: m.arenasc@utem.cl