

COALGEBRAIC STRUCTURE OF GENETIC INHERITANCE

JIANJUN TIAN

Department of Mathematics, University of California
Riverside, CA 92521-0135, USA

BAI-LIAN LI

Department of Botany and Plant Sciences, University of California
Riverside, CA 92521-0124, USA

(Communicated by Yang Kuang)

ABSTRACT. Although in the broadly defined genetic algebra, multiplication suggests a forward direction of from parents to progeny, when looking from the reverse direction, it also suggests to us a new algebraic structure — coalgebraic structure, which we call genetic coalgebras. It is not the dual coalgebraic structure and can be used in the construction of phylogenetic trees. Mathematically, to construct phylogenetic trees means we need to solve equations $x^{[n]} = a$, or $x^{(n)} = b$. It is generally impossible to solve these equations in algebras. However, we can solve them in coalgebras in the sense of tracing back for their ancestors. A thorough exploration of coalgebraic structure in genetics is apparently necessary. Here, we develop a theoretical framework of the coalgebraic structure of genetics. From biological viewpoint, we defined various fundamental concepts and examined their elementary properties that contain genetic significance. Mathematically, by genetic coalgebra, we mean any coalgebra that occurs in genetics. They are generally noncoassociative and without counit; and in the case of non-sex-linked inheritance, they are cocommutative. Each coalgebra with genetic realization has a baric property. We have also discussed the methods to construct new genetic coalgebras, including cocommutative duplication, the tensor product, linear combinations and the skew linear map, which allow us to describe complex genetic traits. We also put forward certain theorems that state the relationship between gametic coalgebra and gametic algebra. By Brower's theorem in topology, we prove the existence of equilibrium state for the in-evolution operator.

1. Genetic motivation. While modern genetic inheritance initiated with the theory of Charles Darwin, it was the Augustinian Monk Gregor Mendel, who first discovered the mathematical character of heredity. In his first paper [1], Mendel exploited some symbolism, which is quite algebraically suggestive, to express his law. In fact, it was later termed “Mendelian algebras” by several authors. In the 1920s and 1930s, general genetic algebras were introduced. Apparently Serebrowsky [2] was the first to give an algebraic interpretation of the sign “ \times ”, which indicated sexual reproduction, and to give a mathematical formulation of the Mendelian laws. Glivenkov [3] continued to work at this direction and introduced the so-called Mendelian algebras for diploid populations with one locus or with two unlinked loci.

2000 *Mathematics Subject Classification.* 16W99.

Key words and phrases. general genetic algebras, general genetic coalgebras, baric coalgebras, conilpotent coalgebras, in-evolution operators.

Independently, Kostitzin [4] also introduced a “symbolic multiplication” to express the Medelian laws. The systematic study of algebras occurring in genetics was due to I. M. H. Etherington. In his series of seminal papers [5], he succeeded in giving a precise mathematical formulation of Mendel’s laws in terms of non-associative algebras. He pointed out that the nilpotent property is essential to these genetic algebras and formulized it in his definition of train algebra and baric algebra. He also introduced the concept of commutative duplication by which the gametic algebra of a randomly mating population is associated with a zygotic algebra. Besides Etherington, fundamental contributions have been made by Gonshor [6], Schafer [7] Holgate [8, 9], Hench [10], Reiser [11], Abraham [12], Lyubich, and Worz-Busekros [13]. During the early days in this area, it appeared that general genetic algebras or broadly defined genetic algebra, (by these terms we mean all algebras or any algebra having been used in genetics,) can be developed into a field of independent mathematical interest, because these algebras are in general not associative and do not belong to any of the well-known classes of non-associative algebras such as Lie algebra, alternative algebra, or Jordan algebra. They possess some distinguished properties that lead to many interesting mathematical results. For example, baric algebra, which has nontrivial representation over the underlying field, and train algebra, whose coefficients of rank equation are only functions of the image under this representation, are new objectives for mathematicians. From the viewpoint of mathematics, Gonshor’s and Schafer’s papers are particularly important. They introduced the (narrowly defined) concept of “genetic algebra” mathematically and proved two fundamental theorems on the existence and uniqueness of idempotents and on the convergence behavior of sequences of plenary powers in special train algebras. Until 1980, the most comprehensive reference in this area was Worz-Busekros. More recent results and direction, such as evolution in genetic algebras, can be found in the book of Lyubich’s work [13]. A good survey article is Reed’s [14].

General genetic algebras are the product of interaction between biology and mathematics. Mendel’s genetics offers a new object to mathematics: general genetic algebras. The study of these algebras reveals the algebraic structure of genetics, which always simplifys 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. However, we found that a very important aspect of Mendelian genetics is lost concerning the existing various genetic algebras, because Mendelian genetic processes in itself implies two directions: forward from parents to progeny, and backward from progeny to their ancestors. Multiplication (referred to sexual reproduction), while it suggests an evolutionary dynamics over generations, was only considered, in general genetic algebras, in a direction from parents to their progeny. By looking at the genetic processes in the backward direction, we put forward a new algebraic structure—general genetic coalgebras. From mathematical viewpoint, they are interesting objectives, with new concepts and new elements put forward, such as a character of coalgebra, baric coalgebras, co-powers, co-nilpotent, and so on. It will also become a new foundation and perspective to study genetic inheritance.

This new mathematical framework could help evolutionary biologists to trace back through generations over time and space in search of certain common ancestors or ancestral distributions. For example, in general genetic algebras, we study

the sequences of plenary powers or principle powers, $x^{[n]}$ or $x^{(n)}$, to look for the genetic information after n generations from the initial generation x . However, when we know the present generation y , and want to trace back for the genetic information of n generations before, we need to solve equations $x^{[n]} = y$ or $x^{(n)} = y$. Generally, it is impossible to solve these equations in algebras. This is the main reason one gives restriction of low-degree polynomials (rank) when people study general nonassociative algebras. In our proposed coalgebraic structure, we can solve these equations in the sense of dynamic viewpoint. That is, we once view comultiplication as the backward dynamic process whenever we think of multiplication as forward dynamic process, then we just need to take plenary co-powers or principle co-powers, $\Delta^{[n]}(y) = x$ or $\Delta^{(n)}(y) = x$. This way, we provide a method for evolutionary biologists to trace back through generations over time and space in search of certain common ancestors, and to construct the phylogenetic trees. Since any genetic algebra is not sufficient to study evolutionary processes, a thorough exploration of coalgebraic structure in genetics is necessary.

The article is organized as follows: In section 2, we recall the basic concepts in broadly defined genetic algebras, particularly algebras with genetic realization and their dual coalgebras. In this section we will note that these dual coalgebras can not give us the genetic information when we study backward evolution, construct phylogenetic trees. In section 3, we introduce the coalgebras with genetic realization, which is the basic genetic coalgebras. As examples of this, we will study the gametic coalgebras, the zygotic coalgebras and the co-commutative duplications. In section 4, we will discuss elementary properties of the general genetic coalgebras, in particular, the baric coalgebras and their genetic aspect. In section 5, we will discuss methods to construct new genetic coalgebras from simple ones which corresponds to complex genetic traits. In section 6, we will introduce various co-powers and co-nilpotent coalgebras that are essentially important in constructing phylogenetic trees. In section 7, we introduce the in-evolution operators and discuss their properties that can be used to specify the tracing path. Finally, we post some interesting open questions in section 8.

2. The basic genetic algebras and their dual coalgebras. In this section, we recall some basic definitions in general genetic algebras and give their dual coalgebraic structure, which may offer a different perspective about genetic algebras. These dual coalgebras themselves do not provide any new information about the underlying genetic processes, but they serve as a starting point for us to develop a new coalgebraic structure that will provide new genetic information.

2.1. The Basic Genetic Algebras. By a population space, we mean that a vector

space, Ω , spanned by a set $\{e_i, | i \in \Lambda\}$, which are free over the real number field R , where the generator set $\{e_i, | i \in \Lambda\}$ is a certain genotype set or a hereditary type set that is related to a trait that we are interested in, and Λ is an index set which is usually finite, although it may be infinite. Now, we can define a linear map m from the tensor product space $\Omega \otimes \Omega$ to Ω as follows:

$$m : \Omega \otimes \Omega \longrightarrow \Omega$$

$$m(e_i \otimes e_j) = \sum_{k=1}^n \gamma_{ij}^k e_k, \quad i, j = 1, 2, \dots, n$$

and linearly extend $m(e_i \otimes e_j)$ onto $\Omega \otimes \Omega$. These structural coefficients satisfy $0 \leq \gamma_{ij}^k \leq 1$ and $\sum_{k=1}^n \gamma_{ij}^k = 1$. Then, the vector space Ω becomes an algebra. We call the pair (Ω, m) , in a general sense, a basic genetic algebra. Generally, a basic genetic algebra is not associative. Its biological significance is that when three genotypic parents cross to beget genotypic grandsons, there are two different combinations of crossing. As a result, the genotype of grandsons arising from two different combinations of crossing are generally different. Some authors call an algebra defined by this way an algebra with genetic realization. Lyubich call it stochastic algebra when it is commutative. Two examples follow: the gametic algebras and zygotic algebras.

2.1.1. *The gametic algebras.* We consider an infinitely large randomly mating population of diploid individuals which differ genetically at several autosomal loci and can be represented by genetically finitely distinct gametes as our population space Θ . We denote these genetical distinct gametes by e_i , $i = 1, 2, \dots, n$, then Θ is spanned by these gametes:

$$\Theta = \left\{ \sum_{i=1}^n \alpha_i e_i \mid \alpha_i \in R, \quad i = 1, 2, \dots, n \right\}.$$

If we take the standard n -dimensional complex in the linear space R^n and denote it by S_0 , there is a bijection between the state set of the population space and S_0 . For convenience, we denote a subspace of Θ by Θ_0 , which is linearly isomorphic to S_0 . Then, each vector, $x = \sum_{i=1}^n \alpha_i e_i$, in Θ_0 represents a state of the population, and the coefficient α_i of e_i is the frequency of the gamete e_i . For the multiplication $m(e_i \otimes e_j) = \sum_{k=1}^n \gamma_{ij}^k e_k$, $i, j = 1, 2, \dots, n$, the structural coefficient γ_{ij}^k can be interpreted as the probability that the zygote $e_i \otimes e_j$ produces the gamete e_k . It is clear that, when we consider non-sex-linked inheritance, the gametic algebra is commutative and not associative. Specifically, the structural coefficients satisfy an additional condition

$$\gamma_{ij}^k = \gamma_{ji}^k, \quad i, j, k = 1, 2, \dots, n.$$

2.1.2. *The zygotic algebras.* In a way similar to constructing a gametic algebra, we can construct a zygotic algebra. Let e_{ij} be a kind of zygote that is begotten from two kinds of genetically different gametes. Then, we can represent the zygote as

$$e_{ij} = e_i \otimes e_j, \quad i \leq j, \quad i, j = 1, 2, \dots, n,$$

where e_i and e_j are gametes. We also denote the zygote space by Z

$$Z = \left\{ \sum_{i,j=1, i \leq j}^n \alpha_{ij} e_{ij} \mid \alpha_{ij} \in R, \quad i \leq j, \quad i, j = 1, 2, \dots, n \right\}.$$

Random mating of zygotes e_{ij} and e_{pq} in the population yields zygote e_{ks} with probability $\gamma_{ij,pq,ks}$. And the multiplication is given by

$$m(e_{ij} \otimes e_{pq}) = \sum_{k \leq s} \gamma_{ij,pq,ks} e_{ks}, \quad i \leq j, p \leq q, \quad i, j, p, q = 1, 2, \dots, n.$$

These coefficients satisfy:

$$0 \leq \gamma_{ij,pq,ks} \leq 1, \quad \sum_{k,s=1, k \leq s}^n \gamma_{ij,pq,ks} = 1, \quad \text{and} \quad \gamma_{ij,pq,ks} = \gamma_{pq,ij,ks},$$

where $i \leq j, p \leq q, i, j, p, q = 1, 2, \dots, n$.

Thus, we get the basic zygotic algebras.

2.2. The dual coalgebras of basic genetic algebras. Now, we define the co-population space to be the linear dual space of the population space,

$$\Omega^* = \{f \mid f : \Omega \longrightarrow R, \text{ is a linear map}\}.$$

The dual basis $\{\eta_i \mid i \in \Lambda\}$ is called a co-genotype or co-hereditary type set, which satisfies $\eta_i(e_j) = \delta_{ij}$, where δ_{ij} is Kronecker delta. There is a dual coalgebraic structure over Ω^* . Its comultiplication is given by the composition

$$\Delta : \quad \Omega^* \xrightarrow{m^*} (\Omega \otimes \Omega)^* \xrightarrow{\rho^{-1}} \Omega^* \otimes \Omega^*,$$

where map ρ is the bijection from the tensor space $\Omega^* \otimes \Omega^*$ to $(\Omega \otimes \Omega)^*$, since Ω is finite dimensional. Explicitly, the composition says $\langle \Delta(f), a \otimes b \rangle = \langle f, m(a \otimes b) \rangle = \langle f, ab \rangle$, for each $f \in \Omega^*$ and $a \otimes b \in \Omega \otimes \Omega$.

If we write the comultiplication as $\Delta(\eta_k) = \sum_{i,j} \alpha_{ijk} \eta_i \otimes \eta_j$, then

$$\begin{aligned} \langle \Delta(\eta_k), e_i \otimes e_j \rangle &= \langle \eta_k, e_i \cdot e_j \rangle = \left\langle e_k^*, \sum_{l=1}^n \gamma_{ij}^l e_l \right\rangle \\ &= \sum_{l=1}^n \gamma_{ij}^l \langle e_k^*, e_l \rangle = \gamma_{ij}^k \end{aligned}$$

so we have

$$\Delta(\eta_k) = \sum_{i,j=1}^n \gamma_{ij}^k \eta_i \otimes \eta_j.$$

Thus, we get the dual coalgebra (Ω^*, Δ) . When the coefficients satisfy $0 \leq \gamma_{ij}^k \leq 1$, and $\sum_{k=1}^n \gamma_{ij}^k = 1$, it is the dual basic genetic coalgebra, and, in general it is not coassociative.

REMARK 1. We give an interpretation of the definition as follows. If we take the gametic space Θ as the population space Ω here, we will get the co-gametic space Θ^ , and we call η_k to be a co-gamete, and $\eta_i \otimes \eta_j$ to be a co-zygote. The coefficient γ_{ij}^k is still the probability that a co-zygote $\eta_i \otimes \eta_j$ produces a co-gamete η_k . Thus, we will get the dual coalgebra of the gametic algebra (Θ^*, Δ) . Similarly, we can get the dual coalgebra of the zygotic algebra. Although the summation in the comultiplication operation is taken in a different way, this coalgebra is “equivalent” to the original algebra to describe various gametes’ or zygotes’ genetic situations. This also means that we will not obtain new information about genetic processes from dual coalgebras, which just gives us a different viewpoint. Pursuing different coalgebras or coalgebraic structures in order to study genetic evolutionary problems is needed, which is the focus of this paper.*

3. The basic genetic coalgebras. In this section, we will define the basic genetic coalgebras. We will use the term “general genetic coalgebra(s)” or “broadly defined genetic algebras” for any coalgebra or all coalgebras used in genetics, and to avoid any confusion, we will not use the term “genetic coalgebra.” As examples, we will give the gametic coalgebras, and the zygotic coalgebras. We will also put forward an approach—cocommutative duplication, which can be used to construct a new genetic structure from the old ones to reveal some more genetic information. For

example, we can construct zygotic coalgebraic structures from gametic coalgebraic structures. We will also give the relation between gametic algebras and gametic coalgebras under certain specific conditions.

3.1. The gametic coalgebras. Let Θ be the gametic space, e_1, e_2, \dots, e_n be the genetically distinct gametes in this space, every element, $a_i \otimes a_j$, in tensor product space $\Theta \otimes \Theta$, can be viewed as a zygote. Now, we trace from a gamete back to its “parental generation” to see various possibilities that those parents can “beget” this gamete. Then, if a gamete e_k comes from a zygote of type $e_i \otimes e_j$ with probability β_{ij}^k , we define a comultiplication as follows:

$$\begin{aligned} \Delta &: \Theta \longrightarrow \Theta \otimes \Theta \\ \Delta(e_k) &= \sum_{i,j=1}^n \beta_{ij}^k e_i \otimes e_j, \quad k = 1, 2, \dots, n \end{aligned}$$

and linearly extend it onto Θ . The coefficients satisfy

$$0 \leq \beta_{ij}^k \leq 1 \text{ and } \sum_{i,j=1}^n \beta_{ij}^k = 1, \quad i, j, k = 1, 2, \dots, n.$$

Then we call pair (Θ, Δ) gametic coalgebra.

REMARK 2. *In general, this gametic coalgebra is not coassociative. That is, the coassociativity $(id \otimes \Delta)\Delta = (\Delta \otimes id)\Delta$ is not satisfied. When we trace from a gamete back to “its grandparental generation” along distinct individuals in “its parental generation,” we will get different distributions of probabilities of the genotype of two generations ago that gave rise to the gamete.*

REMARK 3. *In the study of the inheritance of non-sex-linked traits, this coalgebra is cocommutative; that is, $\tau\Delta = \Delta$, where τ is the permutation of $\Theta \otimes \Theta$, $\tau(a \otimes b) = b \otimes a$. Or, structural coefficients satisfy $\beta_{ij}^k = \beta_{ji}^k$, $i, j, k = 1, 2, \dots, n$. In some special cases, it is not cocommutative.*

REMARK 4. *In general, there is no counit; that is, there is no coalgebraic map ε from Θ to R , such that $(id \otimes \varepsilon)\Delta = id = (\varepsilon \otimes id)\Delta$. However, we will give another concept, character, to describe a similar property later.*

As an example, let’s consider a randomly mating population of diploid individuals that differ in a locus alleles e_1, e_2, \dots, e_n ; then, the population space is spanned by $\{e_1, e_2, \dots, e_n\}$. We can still use them to indicate the genetically distinct gametes. Since comultiplication is determined by structural coefficients, we give these coefficients as follows:

$$\beta_{ij}^k = \frac{1}{2n} (\delta_{ik} + \delta_{jk}), \quad i, j, k = 1, 2, \dots, n.$$

Let’s look at a special case: when $e_1 = A$, $e_2 = a$, we have

$$\Delta(A) = \frac{1}{2}A \otimes A + \frac{1}{4}A \otimes a + \frac{1}{4}a \otimes A,$$

and

$$\Delta(a) = \frac{1}{4}A \otimes a + \frac{1}{4}a \otimes A + \frac{1}{2}a \otimes a.$$

3.2. The zygotic coalgebras. For simplicity, we denote the zygote $e_i \otimes e_j$ by e_{ij} . Then, our population space will be the zygotic space spanned by

$$\{e_{ij}, \quad i \leq j, \quad i, j = 1, 2, \dots, n\}$$

where we consider each e_{ij} to be genetically distinct zygotes in a population. We also think e_{ij} and e_{ji} are the same type of zygote here. The zygotic space Z is the same as that in Section 2. For each zygote e_{ij} , we trace back to its parental generation to see the probability distribution of zygotes in its parental generation involved. Or, we consider the zygote e_{ij} comes from two zygotes e_{pq} and e_{st} in its parental generation with probability $\beta_{pq,st}^{ij}$. Then, we have:

$$\begin{aligned} 0 &\leq \beta_{pq,st}^{ij} \leq 1, & \beta_{pq,st}^{ij} &= \beta_{st,pq}^{ij} \\ i &\leq j, \quad p \leq q, & i, j, p, q, s, t &= 1, 2, \dots, n, \end{aligned}$$

and

$$\sum_{p \leq q, s \leq t} \beta_{pq,st}^{ij} = 1, \quad i, j = 1, 2, \dots, n.$$

We define the comultiplication for basis elements:

$$\Delta(e_{ij}) = \sum_{p \leq q, s \leq t} \beta_{pq,st}^{ij} e_{pq} \otimes e_{st}$$

and linearly extend it onto Z . Thus there is a cocommutative coalgebraic structure in Z , which is not coassociative and without counit. We call the pair (Z, Δ) the zygotic coalgebra.

3.3. The cocommutative duplication. First, let us look at an example to see how to construct a zygotic coalgebra from a gametic coalgebra. Here, we still consider non-sex linked traits. Both the gametic coalgebra and the zygotic coalgebra are cocommutative. We assume that zygotes are formed by random mating of gametes; and gametes are e_1, e_2, \dots, e_n , and zygotes are $e_{ij} = e_i \otimes e_j \quad i \leq j, i, j = 1, 2, \dots, n$. It is clear that the probability of zygote e_{ij} coming from zygotes e_{pq} and e_{st} is the product of the probabilities that the gamete e_i comes from zygote e_{pq} and the probabilities that the gamete e_j comes from zygote e_{st} . Since $e_{ij} = e_{ji}$, we should add these probabilities appropriately. Then the comultiplication coefficients, $\beta_{pq,st}^{ij}$, of zygotic coalgebra Z are obtained by the comultiplication coefficient β_{pq}^i of gametic coalgebra Θ as follows:

$$\beta_{pq,st}^{ij} = \begin{cases} \beta_{pq}^i \beta_{st}^j + \beta_{st}^i \beta_{pq}^j & i \leq j \\ \beta_{pq}^i \beta_{st}^i & i = j \end{cases}$$

Thus, we have derived the zygotic coalgebra from the gametic coalgebra. We call this process cocommutative duplication of the gametic coalgebra. We give the definition as follows:

DEFINITION 3.1. Let (C, Δ) be a cocommutative coalgebra, Σ be a subspace of the tensor product $C \otimes C$, which is given by

$$\Sigma = \left\{ \sum_{i \in I} (x_i \otimes y_i - y_i \otimes x_i) \mid x_i, y_i \in C, \quad i \in I, \quad |I| < \infty \right\}.$$

The equivalent class of $x \otimes y$ in the quotient, the symmetric tensor product, $C \otimes C / \Sigma = C \vee C$ is denoted by $x \vee y$. We can define comultiplication over this

quotient to be

$$\nabla(x \vee y) = \sum_{(x),(y)} x_{(1)} \vee y_{(1)} \otimes x_{(2)} \vee y_{(2)}$$

where $\Delta(x) = \sum_{(x)} x_{(1)} \otimes x_{(2)}$ and $\Delta(y) = \sum_{(y)} y_{(1)} \otimes y_{(2)}$ are Sweedler notations.

Then, the symmetric tensor product with comultiplication ∇ forms a cocommutative coalgebra. We call $(C \vee C, \nabla)$ the cocommutative duplication of (C, Δ) .

3.4. Coalgebras with genetic realization. Let (C, Δ) be a coalgebra over a field K . K may be taken as real number field R , if C admits a basis e_1, e_2, \dots, e_n such that the comultiplication constants β_{ij}^k with respect to this basis satisfy

$$0 \leq \beta_{ij}^k \leq 1 \quad i, j, k = 1, 2, \dots, n$$

and

$$\sum_{i,j=1}^n \beta_{ij}^k = 1, \quad k = 1, 2, \dots, n$$

where

$$\Delta(e_k) = \sum_{i,j=1}^n \beta_{ij}^k e_i \otimes e_j, \quad k = 1, 2, \dots, n.$$

We say (C, Δ) is a coalgebra with genetic realization and its basis is called a natural basis.

REMARK 5. *Gametic coalgebras and zygotic coalgebra all have a genetic realization. If a coalgebra has a genetic realization, then element e_1, e_2, \dots, e_n of the natural basis can be interpreted as genotypes or hereditary types of a population and the real non-negative number β_{ij}^k can be considered as the probability that e_k comes from e_i and e_j by mating, $i, j, k = 1, 2, \dots, n$. If a coalgebra has a natural basis, it may have many (finite or even infinite many) natural bases.*

THEOREM 3.1. *Suppose that the population space is a randomly mating diploid population without selection, and non-sex linked and all zygotes have the same fertility. If the genetic distinct gametes e_1, e_2, \dots, e_n span the gametic space Θ , then the gametic algebra (Θ, m) and the gametic coalgebra (Θ, Δ) have the relation*

$$\beta_{ij}^k = \frac{\gamma_{ij}^k}{\sum_{i,j=1}^n \gamma_{ij}^k}, \quad i, j, k = 1, 2, \dots, n$$

where these coefficients are structural constants, that $m(e_i \otimes e_j) = \sum_{k=1}^n \gamma_{ij}^k e_k$, $i, j = 1, 2, \dots, n$ and $\Delta(e_k) = \sum_{i,j=1}^n \beta_{ij}^k e_i \otimes e_j$, $k = 1, 2, \dots, n$.

Proof. Since it is non-sex linked, the gametic algebra (Θ, m) is commutative and the gametic coalgebra (Θ, Δ) is cocommutative. By the definition of the gametic algebra, γ_{ij}^k is the probability that zygote $e_i \otimes e_j$ produces gamete e_k , and since there is no selection, we can think of γ_{ij}^k as the number that zygote $e_i \otimes e_j$ produces gametes e_k . Then the total number of gamete e_k would be $\sum_{i,j=1}^n \gamma_{ij}^k$, but γ_{ij}^k gametes e_k come from zygote $e_i \otimes e_j$, by the definition of the gametic coalgebra, $\beta_{ij}^k = \frac{\gamma_{ij}^k}{\sum_{i,j=1}^n \gamma_{ij}^k}$. We complete the proof. \square

4. The baric coalgebras. To characterize various genetic situations by coalgebras, we need some new specific concepts. In this section, we will put forward the baric coalgebra which contains certain genetic information.

4.1. The character of a coalgebra.

DEFINITION 4.1. *Let (C, Δ) be a coalgebra over a field K . A character is defined to be a nonzero coalgebraic map from C to the underlying field K .*

That is, if ϕ is a character, ϕ is a linear map which preserves the coalgebraic structure. We take field K as a coalgebra while the comultiplication is taken as $\Delta_k(1) = 1 \otimes 1$. Since $K \otimes K \simeq K$, character ϕ satisfies $(\phi \otimes \phi)\Delta = \Delta_k\phi = \phi$. In fact, the character can be viewed as a representation of C with dimension one.

REMARK 6. *Since a coalgebra need not to have counit in genetic case, no coalgebra admits a nonzero character. For instance, coalgebra A with comultiplication $\Delta(x) = 0$ has no character. For any character ϕ , the kernel of ϕ , $\ker\phi$, is a coideal with codimension one as a vector subspace of C , and $C \otimes C$ is not contained in $\Delta(\ker\phi)$; that is, $C \otimes C$ is not contained in $\ker\phi \otimes C + C \otimes \ker\phi$.*

PROPOSITION 4.1. *Suppose that C is a coalgebra over a field K and C has a character ϕ , then $\ker\phi$ is a $(n - 1)$ -dimensional coideal of C and the factor coalgebra $C/\ker\phi$ is isomorphic to the field K .*

Proof. Since $(\phi \otimes \phi)\Delta = \phi$, so $(\phi \otimes \phi)\Delta(\ker\phi) = \phi(\ker\phi) = 0$, then

$$\Delta(\ker\phi) \subseteq \ker(\phi \otimes \phi) = C \otimes \ker\phi + \ker\phi \otimes C$$

so $\ker\phi$ is a coideal. By coalgebraic fundamental isomorphism theorem, the factor $C/\ker\phi$ has a unique coalgebra structure such that the induced map is a coalgebra isomorphism $\phi_* : C/\ker\phi \rightarrow K$. It is clear that the dimension of the $\ker\phi$ is $(n - 1)$. We complete the proof. \square

THEOREM 4.1. *Let C be a coalgebra over K with characters, then the correspondence from character ϕ to its $\ker\phi$ is a bijection between the set of characters of C and the set of coideals Π of codimension one such that $C \otimes C$ is not contained in $\Pi \otimes C + C \otimes \Pi$.*

Proof. First, we prove the onto-ness:

Let Π be a coideal with codimensional one, then we have $C = C_0 \oplus \Pi$ as a decomposition of vector space, where $\dim(C_0) = 1$. We define a map $\phi : C \rightarrow K$ which satisfies $\phi(\Pi) = 0$ and $\phi(C_0) \neq 0$ as follows. Take any element $e_0 \in C_0$ as its basis, for any $x \in C_0$, there is $x = k_x e_0$, for some $k_x \in K$. Then, let $\phi(x) = k_x \phi(e_0)$. Now, linearly extend ϕ onto C . Thus, it is clear that ϕ is unique up to a scalar coefficient. Let us explain ϕ is a character for C .

Since $C \otimes C$ is not contained in $\Pi \otimes C + C \otimes \Pi$, we surely have $\Delta(e_0) = \beta e_0 \otimes e_0$ for some nonzero constants $\beta \in K$. If we set $(\phi \otimes \phi)\Delta(e_0) = \phi(e_0)$, then, $\beta\phi(e_0)\phi(e_0) = \phi(e_0)$. We have $\phi(e_0) = \frac{1}{\beta}$. Since this β is determined by multiplication Δ , so ϕ is determined uniquely.

Now,

$$\forall y \in \Pi, \phi(y) = 0$$

since

$$\Delta(y) \in \Delta(\Pi) \subseteq \Pi \otimes C + C \otimes \Pi, (\phi \otimes \phi)\Delta(y) = 0,$$

so,

$$(\phi \otimes \phi) \Delta(y) = \phi(y).$$

And $\forall x \in C_0$, we have

$$\begin{aligned} (\phi \otimes \phi) \Delta(x) &= (\phi \otimes \phi) \Delta(k_x e_0) \\ &= k_x (\phi \otimes \phi) \Delta(e_0) = k_x \phi(e_0) \\ &= \phi(k_x e_0) = \phi(x). \end{aligned}$$

Thus, for any $z \in C = C_0 \oplus \Pi$, write $z = z_0 + z_1$, $z_0 \in C_0$, and $z_1 \in \Pi$, then

$$\begin{aligned} (\phi \otimes \phi) \Delta(z) &= (\phi \otimes \phi) (\Delta(z_0) + \Delta(z_1)) \\ &= (\phi \otimes \phi) \Delta(z_0) + (\phi \otimes \phi) \Delta(z_1) \\ &= \phi(z_0) = \phi(z). \end{aligned}$$

So, $\phi: C \rightarrow K$ is a nonzero coalgebraic map. It is a character.

Now, we prove the into-ness:

If characters $\phi_1 \neq \phi_2$, we need show that $\ker \phi_1 \neq \ker \phi_2$. But, if $\ker \phi_1 = \ker \phi_2$, according to what have been done above, $C = C_0^1 \oplus \ker \phi_1 = C_0^2 \oplus \ker \phi_2$, and we get that C_0^1 and C_0^2 are the same. Then,

$$\phi_1(e_0) = \frac{1}{\beta} = \phi_2(e_0), \quad \beta = \frac{\Delta(e_0)}{e_0 \otimes e_0},$$

Thus, $\phi_1 = \phi_2$. We complete the proof. \square

4.2. The baric coalgebras.

DEFINITION 4.2. Let (C, Δ) be a coalgebra over a field K , if C has a nontrivial character ϕ , we say that C is a baric coalgebra. Sometimes, we call the character ϕ a weight function. We use the notation (C, Δ, ϕ) to denote a baric coalgebra.

DEFINITION 4.3. Let (C, Δ, ϕ) be a baric coalgebra over a field K and C_1 be a subset of C , then C_1 is called a baric subcoalgebra if C_1 is a subcoalgebra of C and C_1 is not contained in $\ker \phi$, or, equivalently, $\phi|_{C_1}$ is not zero.

DEFINITION 4.4. Let (C, Δ, ϕ) be a baric coalgebra over a field K and Π be a subset of C , then Π is called a baric coideal if Π is a coideal of C and $\Pi \subseteq \ker \phi$. In a natural way, the quotient coalgebra C/Π will be a baric quotient coalgebra.

THEOREM 4.2. If (C, Δ) is a coalgebra over R , which has a genetic realization with respect to a natural basis e_1, e_2, \dots, e_n , then (C, Δ) is a baric coalgebra.

Proof. We define a character $\phi: C \rightarrow K$ by $\phi(e_k) = 1$ and linearly extend it onto C . Therefore,

$$\begin{aligned} (\phi \otimes \phi) \Delta(e_k) &= (\phi \otimes \phi) \left(\sum_{i,j=1}^n \beta_{ij}^k e_i \otimes e_j \right) \\ &= \sum_{i,j=1}^n \beta_{ij}^k \phi(e_i) \otimes \phi(e_j) \\ &= \sum_{i,j=1}^n \beta_{ij}^k = 1 = \phi(e_k) \end{aligned}$$

and for every $x \in C$, $x = \sum_{k=1}^n \alpha_k e_k$, we have

$$\begin{aligned} (\phi \otimes \phi) \Delta(x) &= (\phi \otimes \phi) \left(\sum_{k=1}^n \alpha_k \Delta(e_k) \right) \\ &= \sum_{k=1}^n \alpha_k (\phi \otimes \phi) \Delta(e_k) \\ &= \sum_{k=1}^n \alpha_k = \phi(x) \end{aligned}$$

so ϕ is indeed a character. We finish the proof. □

THEOREM 4.3. *Let (C, Δ) be a n -dimensional coalgebra over R ; the following conditions are equivalent:*

- ⟨1⟩ (C, Δ) is baric.
- ⟨2⟩ (C, Δ) has a basis e_1, e_2, \dots, e_n , such that the comultiplication constant can be defined by

$$\Delta(e_k) = \sum_{i,j=1}^n \beta_{ij}^k e_i \otimes e_j$$

and satisfies

$$\sum_{i,j=1}^n \beta_{ij}^k = 1, \quad k = 1, 2, \dots, n.$$

- ⟨3⟩ (C, Δ) has a $(n - 1)$ -dimensional coideal Π and $C \otimes C$ is not contained in $\Delta(\Pi)$.

Proof. ⟨1⟩ \implies ⟨2⟩ Let $\phi : C \longrightarrow K$ be a character, then $\ker \phi = \Pi$ is a $(n - 1)$ -dimensional coideal of C . Let d_2, d_3, \dots, d_n be a basis of Π . Since ϕ is nontrivial, there is an element $a \in C$ such that $\phi(a) \neq 0$. Then, set $e_1 = \frac{1}{\phi(a)}a$, $e_2 = e_1 - d_2$, $e_3 = e_1 - d_3, \dots, e_n = e_1 - d_n$. $e_1, e_2, e_3, \dots, e_n$ forms a basis for C , and $\phi(e_k) = 1, k = 1, 2, \dots, n$. Denote $\Delta(e_k) = \sum_{i,j=1}^n \beta_{ij}^k e_i \otimes e_j$. Since $(\phi \otimes \phi) \Delta(e_k) = \phi(e_k)$, we have

$$\begin{aligned} (\phi \otimes \phi) \Delta(e_k) &= (\phi \otimes \phi) \left(\sum_{i,j=1}^n \beta_{ij}^k e_i \otimes e_j \right) \\ &= \sum_{i,j=1}^n \beta_{ij}^k (\phi \otimes \phi)(e_i \otimes e_j) = \sum_{i,j=1}^n \beta_{ij}^k \\ &= \phi(e_k) = 1 \end{aligned}$$

Thus, we get $\sum_{i,j=1}^n \beta_{ij}^k = 1, k = 1, 2, \dots, n$.

⟨2⟩ \implies ⟨3⟩

We define $d_k = e_1 - e_k, k = 2, 3, \dots, n$. $C_0 = \langle d_2, d_3, \dots, d_n \rangle$ is a subspace spanned by those elements. We will check that C_0 is a $(n - 1)$ -dimensional coideal

of C as follows:

$$\begin{aligned}
\Delta(d_k) &= \Delta(e_1 - e_k) = \sum_{i,j=1}^n (\beta_{ij}^1 - \beta_{ij}^k) (e_i \otimes e_j) \\
&= - \sum_{i,j=1}^n (\beta_{ij}^1 - \beta_{ij}^k) [(e_1 - e_i) \otimes e_j - e_1 \otimes e_j] \\
&= \sum_{i,j=1}^n (\beta_{ij}^k - \beta_{ij}^1) [(e_1 - e_i) \otimes e_j + e_1 \otimes (e_1 - e_j) - e_1 \otimes e_1] \\
&= \sum_{i,j=1}^n (\beta_{ij}^k - \beta_{ij}^1) (d_i \otimes e_j) + \sum_{i,j=1}^n (\beta_{ij}^k - \beta_{ij}^1) e_1 \otimes d_j \\
&\in C_0 \otimes C + C \otimes C_0
\end{aligned}$$

and

$$\begin{aligned}
\Delta(e_1) &= - \sum_{ij}^n \beta_{ij}^1 d_i \otimes e_j - \sum_{ij}^n e_1 \otimes d_j + e_1 \otimes e_1 \\
&\notin C_0 \otimes C + C \otimes C_0.
\end{aligned}$$

It is done.

$\langle 3 \rangle \Rightarrow \langle 1 \rangle$

We define $\phi : C \rightarrow C/\Pi$ by $x \mapsto x + \Pi$ and linearly extend it onto C . The vector space C/Π is isomorphic to the underlying field K . Since Π is a coideal of C , the map ϕ is coalgebraic, so C/Π is either isomorphic to K or to zero. The latter possibility can be excluded by the dimension one of C/Π . So ϕ is a character function. Thus, we complete the proof. \square

THEOREM 4.4. *Let (C, Δ) be a n -dimensional coalgebra over R , then the following conditions are equivalent:*

- $\langle 1 \rangle$ C has a genetic realization;
- $\langle 2 \rangle$ C is baric and the linear manifold of all elements of weight (the value of the character function) 1 contain a $(n - 1)$ -dimensional simplex L with $\Delta(L) \subseteq L \otimes L$.

Proof. $\langle 1 \rangle \Rightarrow \langle 2 \rangle$ Let e_1, e_2, \dots, e_n be a natural basis of C . We define a map $\phi : C \rightarrow R$ by $\phi(e_k) = 1$ and linearly extend it. Then it is easy to see ϕ is a coalgebraic map. So C is baric. Let L be the convex hull of e_1, e_2, \dots, e_n , then $\Delta(e_k) = \sum_{ij}^n \beta_{ij}^k e_i \otimes e_j \in L \otimes L$, so $\Delta(L) \subseteq L \otimes L$.

$\langle 2 \rangle \Rightarrow \langle 1 \rangle$ Let $\phi : C \rightarrow R$ be any character function. By assumption, this simplex L is the convex hull of n linearly independent elements e_1, e_2, \dots, e_n with $\phi(e_k) = 1, k = 1, 2, \dots, n$. These elements form a basis of C . Since $\Delta(L) \subseteq L \otimes L$, if we write $\Delta(e_k) = \sum_{i,j=1}^n \beta_{ij}^k e_i \otimes e_j$, these coefficients satisfy $0 \leq \beta_{ij}^k \leq 1, i, j, k = 1, 2, \dots, n$. Moreover, since

$$(\phi \otimes \phi) \Delta(e_k) = \phi(e_k) = 1,$$

$$\begin{aligned}
 (\phi \otimes \phi) \Delta(e_k) &= (\phi \otimes \phi) \left(\sum_{i,j=1}^n \beta_{ij}^k e_i \otimes e_j \right) \\
 &= \sum_{i,j=1}^n \beta_{ij}^k \phi(e_i) \phi(e_j) = \phi(e_k),
 \end{aligned}$$

it is $\sum_{i,j=1}^n \beta_{ij}^k = 1$. We complete the proof. □

5. Construction of new genetic coalgebras. To understand the complicated genetic situation in term of coalgebras, we need some approaches to construct a new coalgebra from old ones. For example, there are several coalgebras with each one describing a particular trait in a biological population, so we can combine these coalgebras to make a new one in order to describe and model the population. We will deal with polyploid individual populations as an example. In this section, we will discuss the linear combinations of coalgebras, the tensor product of coalgebras and skew coalgebra by linear maps. In the algebra case, these constructions are well known. Here we just give some of the counterparts in coalgebra motivated by applications in genetics and skip their detailed proofs.

5.1. Linear combinations. Let V be an n -dimensional vector space over a field K , if C is a coalgebra over V , which need not be cocommutative or coassociative, we denote the comultiplication by $\Delta(x) = \Delta_C(x)$.

PROPOSITION 5.1. *Let $C_o(V)$ be the family of all coalgebras over V , for $C_1, C_2 \in C_o(V)$; and $\alpha \in K$, the sum $C_1 + C_2$, and the scalar product αC_1 are defined as coalgebras over V with comultiplications given by*

$$\begin{aligned}
 \Delta_{C_1+C_2}(x) &= \Delta_{C_1}(x) + \Delta_{C_2}(x) \\
 \Delta_{\alpha C_1}(x) &= \alpha \Delta_{C_1}(x).
 \end{aligned}$$

Then, we have

(1) $C_o(V)$ is a vector space over K , which is called the vector space of all coalgebras over V . The zero element of this vector space is the zero coalgebra 0 over V .

(2) The family $C_{oc}(V)$ of all cocommutative coalgebras over V forms a subspace of $C_o(V)$.

Proof. The proof is skipped here. □

REMARK 7. (1) *This proposition is obvious; we do not need to give a proof here.*

(2) *It is clear that a linear combination of the coalgebras C_1, C_2, \dots, C_l is $\sum_{i=1}^l \alpha_i C_i$, its comultiplication is given by*

$$\Delta_{\sum_{i=1}^l \alpha_i C_i}(x) = \sum_{i=1}^l \alpha_i \Delta_{C_i}(x).$$

(3) *If the combination coefficients above are non-negative and sum to one, we call it a convex combination of coalgebras.*

(4) *Coalgebras $C_1, C_2, \dots, C_r \in C_o(V)$ are called linearly independent if and only if the trivial linear combination of these coalgebras is the zero coalgebra.*

PROPOSITION 5.2. *Let e_1, e_2, \dots, e_n be a basis of V . Then the coalgebras C_1, C_2, \dots, C_r are linearly independent if and only if $\Delta_{\sum_{i=1}^r \alpha_i C_i}(e_k) = \sum_{i=1}^r \alpha_i \Delta_{C_i}(e_k) = 0$, $k = 1, 2, \dots, n$, which implies that $\alpha_1 = \alpha_2 = \dots = \alpha_r = 0$.*

Proof. We skip it here. \square

THEOREM 5.1. *Let C_1, C_2, \dots, C_r be baric coalgebras with the same character function ϕ , then every linear combination $\sum_{i=1}^r \alpha_i C_i$ with $\sum_{i=1}^r \alpha_i = 1$ is a baric coalgebra with the character ϕ .*

Proof. For every $x \in V$,

$$\begin{aligned} (\phi \otimes \phi) \Delta_{\sum_{i=1}^r \alpha_i C_i}(x) &= (\phi \otimes \phi) \left(\sum_{i=1}^r \alpha_i \Delta_{C_i}(x) \right) \\ &= \sum_{i=1}^r \alpha_i (\phi \otimes \phi) \Delta_{C_i}(x) \\ &= \sum_{i=1}^r \alpha_i \phi(x) = \phi(x) \end{aligned}$$

Note $K \otimes K \otimes \dots \otimes K \cong K$. So, ϕ is a character for the coalgebra linear combination $\sum_{i=1}^r \alpha_i C_i$. We prove it. \square

THEOREM 5.2. *Let V be a vector space over a field K and C_1, C_2, \dots, C_r be coalgebras over V , which have a genetic realization with respect to the same basis e_1, e_2, \dots, e_n of V . Then every convex combination of these coalgebras has a genetic realization with respect to e_1, e_2, \dots, e_n .*

Proof. We denote the comultiplication constants of coalgebra C_i by $\beta_{pq}^{i,k}$, that is to say

$$\Delta_{C_i}(e_k) = \sum_{p,q=1}^n \beta_{pq}^{i,k} e_p \otimes e_q,$$

then

$$\sum_{p,q=1}^n \beta_{pq}^{i,k} = 1 \text{ for } i = 1, 2, \dots, r, k = 1, 2, \dots, n,$$

and

$$0 \leq \beta_{pq}^{i,k} \leq 1, \text{ for } i = 1, 2, \dots, r, k, p, q = 1, 2, \dots, n.$$

Now, for every convex combination $\sum_{i=1}^r \alpha_i C_i$, where $\sum_{i=1}^r \alpha_i = 1$ and each α_i is non-negative number, if we write $\Delta_{\sum_{i=1}^r \alpha_i C_i}(e_k) = \sum_{p,q=1}^n \lambda_{pq}^k e_p \otimes e_q$, we need check that $\sum_{p,q=1}^n \lambda_{pq}^k = 1$ for each k and $0 \leq \lambda_{pq}^k \leq 1$ for every k, p, q . Let us compute

$$\begin{aligned} \Delta_{\sum_{i=1}^r \alpha_i C_i}(e_k) &= \sum_{i=1}^r \alpha_i \Delta_{C_i}(e_k) \\ &= \sum_{i=1}^r \alpha_i \sum_{p,q=1}^n \beta_{pq}^{i,k} e_p \otimes e_q \\ &= \sum_{p,q=1}^n \left(\sum_{i=1}^r \alpha_i \beta_{pq}^{i,k} \right) e_p \otimes e_q \end{aligned}$$

so

$$\lambda_{pq}^k = \sum_{i=1}^r \alpha_i \beta_{pq}^{i,k}.$$

By Cauchy-Schwarz inequality,

$$\begin{aligned} 0 &\leq \lambda_{pq}^k = \sum_{i=1}^r \alpha_i \beta_{pq}^{i,k} \\ &\leq \left(\sum_{i=1}^r \alpha_i^2 \right)^{\frac{1}{2}} \left(\sum_{i=1}^r (\beta_{pq}^{i,k})^2 \right)^{\frac{1}{2}} \leq 1, \end{aligned}$$

and

$$\begin{aligned} \sum_{p,q=1}^n \lambda_{pq}^k &= \sum_{p,q=1}^n \sum_{i=1}^r \alpha_i \beta_{pq}^{i,k} \\ &= \sum_{i=1}^r \sum_{p,q=1}^n \alpha_i \beta_{pq}^{i,k} = 1 \end{aligned}$$

We complete the proof. □

THEOREM 5.3. *Let V be an n -dimensional vector space over a field K , $n \geq 2$, and $C_{oP}(V)$ be the family of coalgebras over V with property P . Let M_P be the smallest linear manifold $C_o(V)$, which contains $C_{oP}(V)$. Then the dimension of M_P is given as follows:*

property	$\dim(M_P)$
-	n^3
cocommutative	$\frac{1}{2}n^2(n+1)$
baric with the same character	$n(n^2-1)$
cocommutative, baric with the same character	$\frac{1}{2}(n+2)n(n-1)$
genetic realization with the same natural basis	$n(n^2-1)$
cocommutative, with realization under the same natural basis	$\frac{1}{2}(n+2)n(n-1)$

TABLE 1. Dimensions of manifolds.

Proof. We skip the proof. □

5.2. Tensor products. Let C_1, C_2, \dots, C_r be coalgebras over a field K with dimensions n_1, n_2, \dots, n_r , which need not be cocommutative or coassociative, and $C_1 \otimes C_2 \otimes \dots \otimes C_r$ be the tensor product of the underlying vector spaces. Then it is an $(n_1 n_2 \dots n_r)$ -dimensional vector space over K . We can define a linear map as follows:

$$\begin{aligned} \Delta &= P(\Delta_{C_1} \otimes \Delta_{C_2} \otimes \dots \otimes \Delta_{C_r}) : \\ C_1 \otimes C_2 \otimes \dots \otimes C_r &\longrightarrow (C_1 \otimes C_2 \otimes \dots \otimes C_r) \otimes (C_1 \otimes C_2 \otimes \dots \otimes C_r) \end{aligned}$$

where P is an appropriate permutation. Then we have two propositions as follows.

PROPOSITION 5.3. *The pair $(C_1 \otimes C_2 \otimes \dots \otimes C_r, \Delta)$ is a coalgebra.*

PROPOSITION 5.4. *If the comultiplication table of factor C_k is given by*

$$\Delta_{C_k}(e_{kj_k}) = \sum_{p_k, q_k=1}^{n_k} \beta_{p_k q_k}^{kj_k} e_{kp_k} \otimes e_{kq_k}$$

where $\{e_{k1}, e_{k2}, \dots, e_{kn_k}\}$ is a basis for C_k , $k = 1, 2, \dots, r$. Then the comultiplication table of $C_1 \otimes C_2 \otimes \dots \otimes C_r$ is given by

$$\sum_{p_1, q_1, \dots, p_r, q_r}^{n_1, n_2, \dots, n_r} \beta_{p_1 q_1}^{1j_1} \beta_{p_2 q_2}^{2j_2} \dots \beta_{p_r q_r}^{rj_r} (e_{1p_1} \otimes e_{2p_2} \otimes \dots \otimes e_{rp_r}) \otimes (e_{1q_1} \otimes e_{2q_2} \otimes \dots \otimes e_{rq_r}),$$

which is equal to $\Delta(e_{1j_1} \otimes e_{2j_2} \otimes \dots \otimes e_{rj_r})$

Proof. We skip the proof. □

Now we give two statements that have evident significance for genetics. Because complex traits are affected by multiple factors, we can “tensor” these multiple genetic factors up as one coalgebra; such a tensor product of coalgebras will appear very useful to describe polygenetic traits.

THEOREM 5.4. *Let C_k be a baric coalgebra with character ϕ_k , $k = 1, 2, \dots, r$, then the tensor product $C_1 \otimes C_2 \otimes \dots \otimes C_r$ is baric with character $\phi = \phi_1 \otimes \phi_2 \otimes \dots \otimes \phi_r$.*

Proof. For any $x = x^1 \otimes x^2 \otimes \dots \otimes x^r \in C_1 \otimes C_2 \otimes \dots \otimes C_r$, let us verify

$$(\phi \otimes \phi) \Delta(x) = \phi(x).$$

By $(\phi_k \otimes \phi_k) \Delta_{C_k}(x^k) = \phi_k(x^k)$ (that is, $\sum_{(x^k)} \phi_k(x^k_{(1)}) \phi_k(x^k_{(2)}) = \phi_k(x^k)$), we have

$$\begin{aligned} & [(\phi_1 \otimes \phi_2 \otimes \dots \otimes \phi_r) \otimes (\phi_1 \otimes \phi_2 \otimes \dots \otimes \phi_r)] \cdot \\ & [P(\Delta_{C_1} \otimes \Delta_{C_2} \otimes \dots \otimes \Delta_{C_r})](x^1 \otimes x^2 \otimes \dots \otimes x^r) \\ &= \sum_{k=1}^r \sum_{(x^k)} \phi_k(x^k_{(1)}) \phi_k(x^k_{(2)}) = \sum_{k=1}^r \phi_k(x^k). \end{aligned}$$

We get the proof. □

THEOREM 5.5. *Suppose that each coalgebra C_k has a genetic realization with respect to the natural basis $\{e_{k1}, e_{k2}, \dots, e_{kn_k}\}$, $k = 1, 2, \dots, r$, then the tensor product $C_1 \otimes C_2 \otimes \dots \otimes C_r$ also has a genetic realization with respect to the basis $\{e_{1j_1} \otimes e_{2j_2} \otimes \dots \otimes e_{rj_r} \mid 1 \leq j_1 \leq n_1, 1 \leq j_2 \leq n_2, \dots, 1 \leq j_r \leq n_r\}$.*

Proof. Note $\Delta_{C_k}(e_{kj_k}) = \sum_{p_k, q_k=1}^{n_k} \beta_{p_k q_k}^{kj_k} e_{kp_k} \otimes e_{kq_k}$ and the comultiplication table in above proposition, it is easy to see

$$0 \leq \beta_{p_1 q_1}^{1j_1} \beta_{p_2 q_2}^{2j_2} \dots \beta_{p_r q_r}^{rj_r} \leq 1,$$

and

$$\sum_{p_1, q_1, \dots, p_r, q_r}^{n_1, n_2, \dots, n_r} \beta_{p_1 q_1}^{1j_1} \beta_{p_2 q_2}^{2j_2} \dots \beta_{p_r q_r}^{rj_r} = 1,$$

since $\sum_{p_k, q_k=1}^{n_k} \beta_{p_k q_k}^{kj_k} e_{kp_k} = 1$, for each k and j_k . Thus, we complete the proof. □

5.3. Construction of new genetic coalgebras by linear maps. When we consider genetic situation of sex-linked inheritance and mutation in a population of autopolyploid individuals in term of coalgebras, it seems reasonable to introduce a skew or new comultiplication by applying a linear map to one factor of the given comultiplication. We will give some genetic applications later on.

DEFINITION 5.1. *Let V be an n -dimensional vector space over a field K , L_0, L_1, L_2 be linear maps from V to V , and C be a coalgebra over V with comultiplication Δ . We define a map $\overline{\Delta}$ to be a composite*

$$\overline{\Delta} = (L_1 \otimes L_2) \Delta L_0 : C \longrightarrow C \otimes C.$$

Then, the pair $(C, \overline{\Delta})$ is a coalgebra over V .

THEOREM 5.6. *Let V be a vector space over K and C a baric coalgebra over V with character ϕ . If the linear maps $L_0, L_1,$ and $L_2 : V \longrightarrow V$ preserve character; that is, $\phi L_i = \phi, i = 1, 2, 3$, then the coalgebra $(C, \overline{\Delta})$ is baric with character ϕ .*

Proof. By definitions,

$$\begin{aligned} (\phi \otimes \phi) \overline{\Delta} &= (\phi \otimes \phi) (L_1 \otimes L_2) \Delta L_0 \\ &= (\phi L_1 \otimes \phi L_2) \Delta L_0 = (\phi \otimes \phi) \Delta L_0 \\ &= \phi L_0 = \phi. \end{aligned}$$

we get the proof. □

THEOREM 5.7. *Let (C, Δ) be a coalgebra over a vector space V and C have a genetic realization with respect to a natural basis e_1, e_2, \dots, e_n . If the linear map $L_0, L_1, L_2 : V \longrightarrow V$ leaves the simplex*

$$M = \left\{ \sum_{i=1}^n \alpha_i e_i \mid 0 \leq \alpha_i \leq 1, \quad i = 1, 2, \dots, n, \quad \sum_{i=1}^n \alpha_i = 1 \right\}$$

invariant; then the coalgebra $(C, \overline{\Delta})$ also has a genetic realization.

Proof. Write $L_t(e_k) = \sum_{i=1}^n \alpha_i^{t,k} e_i$, $t = 1, 2, 3$, $k = 1, 2, \dots, n$, then $\sum_{i=1}^n \alpha_i^{t,k} = 1$, since L_t leaves M invariant. We see

$$\begin{aligned}
\overline{\Delta}(e_k) &= (L_1 \otimes L_2) \Delta L_0(e_k) \\
&= (L_1 \otimes L_2) \Delta \left(\sum_{i=1}^n \alpha_i^{0,k} e_i \right) \\
&= \sum_{i=1}^n \alpha_i^{0,k} (L_1 \otimes L_2) \Delta(e_i) \\
&= \sum_{i=1}^n \alpha_i^{0,k} (L_1 \otimes L_2) \left(\sum_{p,q=1}^n \beta_{pq}^i e_p \otimes e_q \right) \\
&= \sum_{i=1}^n \alpha_i^{0,k} \sum_{p,q=1}^n \beta_{pq}^i L_1(e_p) L_2(e_q) \\
&= \sum_{i=1}^n \alpha_i^{0,k} \sum_{p,q=1}^n \beta_{pq}^i \sum_{j=1}^n \alpha_j^{1,p} \sum_{s=1}^n \alpha_s^{2,q} e_j \otimes e_s \\
&= \sum_{j,s=1}^n \left(\sum_{i,p,q=1}^n \alpha_i^{0,k} \beta_{pq}^i \alpha_j^{1,p} \alpha_s^{2,q} \right) e_j \otimes e_s
\end{aligned}$$

By Cauchy-Schwarz inequality, we have

$$0 \leq \sum_{i,p,q=1}^n \alpha_i^{0,k} \beta_{pq}^i \alpha_j^{1,p} \alpha_s^{2,q} \leq 1$$

and

$$\sum_{j,s=1}^n \left(\sum_{i,p,q=1}^n \alpha_i^{0,k} \beta_{pq}^i \alpha_j^{1,p} \alpha_s^{2,q} \right) = 1$$

Thus, we complete the proof. \square

6. Conilpotent coalgebras. In this section, we will define several concepts that capture certain interesting genetic features. Since we are considering non-coassociative coalgebras (which may be or may be not be cocommutative) without counit, we need to give a kind of order to take comultiplication. The order may give us a way to specify the path that traces back over the past generations (phylogenetic genealogical trees). We first define various copowers, then define conilpotent coalgebras, which is an analogy of nilpotentness in coalgebraic structure and their possible genetic implications. We also give simple propositions and applications as examples.

DEFINITION 6.1. Let Δ be the comultiplication of a coalgebra (C, Δ) . We define a copower to be a special order that the comultiplication can be performed consecutively. The left principal copower of an element or a subcoalgebra is defined as

$$\begin{aligned} \overset{1}{\Delta} &= \Delta \\ \overset{2}{\Delta} &= (\Delta \otimes id) \Delta \\ \overset{3}{\Delta} &= (\Delta \otimes id \otimes id) (\Delta \otimes id) \Delta \\ &\dots\dots\dots \\ \overset{m}{\Delta} &= \left(\Delta \otimes id^{\otimes(m-1)} \right)^{m-1} \Delta \quad . \end{aligned}$$

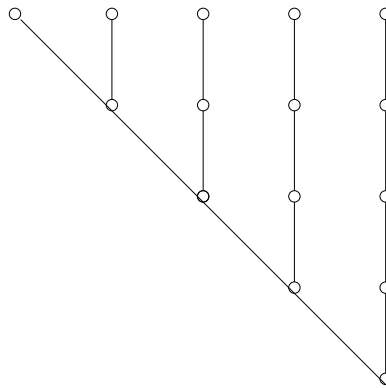


FIGURE 1. Fourth left principle copower of x .

The right principal copower of an element or a subcoalgebra is defined as

$$\begin{aligned} \Delta^1 &= \Delta \\ \Delta^2 &= (id \otimes \Delta) \Delta \\ \Delta^3 &= (id \otimes id \otimes \Delta) (id \otimes \Delta) \Delta \\ &\dots\dots\dots \\ \Delta^m &= \left(id^{\otimes(m-1)} \otimes \Delta \right) \Delta^{(m-1)}. \end{aligned}$$

The principle copower of an element or a coalgebra is defined as

$$\begin{aligned} \overset{(1)}{\Delta} &= \Delta \\ \overset{(2)}{\Delta} &= (\Delta \otimes \Delta) \Delta \\ \overset{(3)}{\Delta} &= (\Delta \otimes id \otimes id \otimes \Delta) \overset{(2)}{\Delta} \\ &\dots\dots\dots \\ \overset{(m)}{\Delta} &= (\Delta \otimes id^{\otimes 2(m-2)} \otimes \Delta) \overset{(m-1)}{\Delta} \quad . \end{aligned}$$

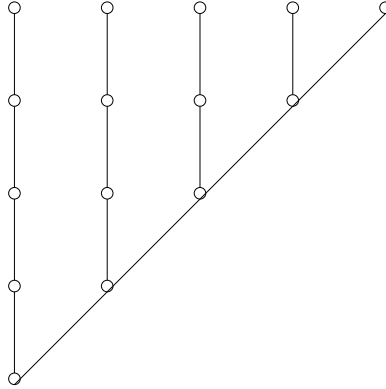


FIGURE 2. Fourth right principle copower of x .

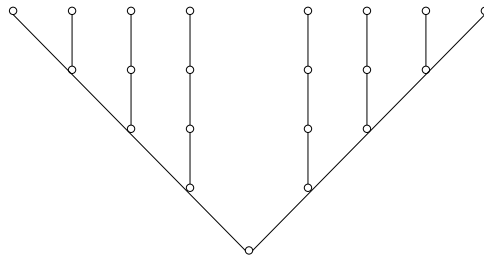


FIGURE 3. Fourth principle copower of x .

The plenary copower of an element or a coalgebra is defined as

$$\begin{aligned}
 \overset{[1]}{\Delta} &= \Delta \\
 \overset{[2]}{\Delta} &= (\Delta \otimes \Delta) \Delta \\
 \overset{[3]}{\Delta} &= (\Delta \otimes \Delta \otimes \Delta \otimes \Delta) \overset{[2]}{\Delta} \\
 &\dots\dots\dots \\
 \overset{[m]}{\Delta} &= \left(\Delta^{\otimes 2^{m-1}} \right) \overset{[m-1]}{\Delta}.
 \end{aligned}$$

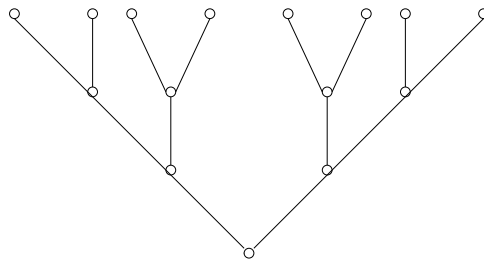


FIGURE 4. Fourth right principle copower of x .

DEFINITION 6.2. An element $x \in (C, \Delta)$ is called copower coassociative if the relation holds for any integer $k \geq 2$ and any non-negative partition $k = l + p + 1$,

$$\Delta^k(x) = (id^{\otimes l} \otimes \Delta \otimes id^{\otimes p}) \Delta^{k-1}(x)$$

DEFINITION 6.3. 1. An element $x \in C$ is called left (right) conilpotent of index k , if

$${}^k\Delta(x) = 0, \text{ but } {}^{k-1}\Delta(x) \neq 0,$$

$$(\Delta^k(x) = 0, \text{ but } \Delta^{k-1}(x) \neq 0).$$

2. A subcoalgebra C_1 of C is called left conil (right conil) if all elements of C_1 are left conilpotent (right conilpotent).

3. A subcoalgebra C_0 of C is called left conilpotent (right conilpotent, conilpotent) of index k , if

$${}^k\Delta(C_0) = 0, \text{ but } {}^{k-1}\Delta(x) \neq 0.$$

$$\left(\Delta^k(C_0) = 0, \text{ but } \Delta^{k-1}(C_0) \neq 0; \binom{(k)}{\Delta}(C_0) = 0, \text{ but } \binom{(k-1)}{\Delta}(C_0) \neq 0 \right)$$

4. A subcoalgebra C_2 of C is called cosolvable if there is a positive integer k , such that

$$\binom{[k]}{\Delta}(C_2) = 0.$$

REMARK 8. The biological significance of left conilpotent of element x is that if we trace back by the left path, the information of ancestors will be lost. This is a kind of evolution path. The other conilpotent concepts have the similar significance.

LEMMA 6.1. Let (C, Δ) be a baric coalgebra with character ϕ . If $\ker\phi$ is right conilpotent, then ϕ is uniquely determined.

Proof. Assume that $\omega : C \rightarrow K$ is any nontrivial coalgebraic map. Let $y \in \ker\phi$. Since $\ker\phi$ is right conilpotent, there is a positive integer l , such that $\Delta^l(y) = 0$. Then $\omega^{\otimes(k+1)}\Delta^k(y) = \omega(y)$, so $\omega(y) = 0$. Thus, for any $y \in \ker\phi$, we have $\omega(y) = \phi(y)$. Let $x \in C - \ker\phi$, $\phi(x) \neq 0$, but

$$(\phi \otimes \phi) \left(\frac{x \otimes x}{\phi(x)} - \Delta(x) \right) = \frac{\phi(x)\phi(x)}{\phi(x)} - \phi(x) = 0$$

This means $\frac{x \otimes x}{\phi(x)} - \Delta(x) \in \ker(\phi \otimes \phi)$. Now we claim that

$$\forall p \in \ker(\phi \otimes \phi), (\omega \otimes \omega)(p) = 0.$$

Then,

$$(\omega \otimes \omega) \left(\frac{x \otimes x}{\phi(x)} - \Delta(x) \right) = \frac{\omega(x)\omega(x)}{\phi(x)} - \omega(x) = 0$$

So we have that $\omega(x) = \phi(x)$, or, $\omega(x) = 0$. But, since ω is nontrivial, we get $\omega = \phi$.

The proof of our claim: since $\ker(\phi \otimes \phi) = \ker\phi \otimes C + C \otimes \ker\phi$, $p \in \ker(\phi \otimes \phi)$ can be written as $p = \sum_i a_i \otimes b_i$. For simplicity, we take two terms as $p = a_1 \otimes b_1 + a_2 \otimes b_2$ to check.

Now suppose $a_1 \in \ker\phi$ and $b_2 \in \ker\phi$, then $\Delta^{k_1}(a_1) = 0$ and $\Delta^{k_2}(b_2) = 0$ for some integers k_1 and k_2 . Thus

$$\begin{aligned} & \left(id^{\otimes(k_1+1)} \otimes \Delta^{k_2} \right) (\Delta^{k_1} \otimes id)(p) \\ &= \left(id^{\otimes(k_1+1)} \otimes \Delta^{k_2} \right) (\Delta^{k_1} \otimes id)(a_1 \otimes b_1 + a_2 \otimes b_2) \\ &= \Delta^{k_1}(a_1) \otimes \Delta^{k_2}(b_1) + \Delta^{k_1}(a_2) \otimes \Delta^{k_2}(b_2) \\ &= 0 \end{aligned}$$

The proof is done. \square

For example, if in some population the genetic inheritance can be described by an algebra A , this is a basic genetic algebra. Then the generation sequence arising from random mating with the initial population or from random mating within a generation is represented by the sequence of principle powers $x^{(r)}$ or by the sequence of plenary powers $x^{[r]}$, respectively, where $x \in A$ is the initial population. Now, supposing that the present population or gene y is known, what about the ancestral or ancestral gene distribution of n generation before? To answer this question, we have to give a coalgebraic structure of A , as we have defined basic genetic coalgebra by tracing back in backward dynamic process. Then $\overset{(n)}{\Delta}(y)$ or $\overset{[n]}{\Delta}(y)$ or other order of comultiplications can give us some ancestral or ancestral gene distribution information. In this sense, we can solve the equations $x^{(n)} = y$ and $x^{[n]} = y$. For example, in simple Mendelian inheritance, let $y = \alpha A + \beta a$. Then

$$\Delta(y) = \frac{\alpha}{2}A \otimes A + \frac{\alpha + \beta}{4}A \otimes a + \frac{\alpha + \beta}{4}a \otimes A + \frac{\beta}{2}a \otimes a.$$

This tells us that y comes from $A \otimes A$ with probability $\frac{\alpha}{2}$, from $A \otimes a$ with probability $\frac{\alpha + \beta}{2}$, from $a \otimes A$ with probability $\frac{\alpha + \beta}{2}$ and from $a \otimes a$ with probability $\frac{\beta}{2}$.

$$\begin{aligned} \overset{[2]}{\Delta}(y) &= \frac{\alpha}{8}A \otimes A \otimes A \otimes A + \left(\frac{3\alpha + \beta}{32} \right) \left(\begin{array}{l} a \otimes A \otimes A \otimes A + A \otimes a \otimes A \otimes A + \\ + A \otimes A \otimes a \otimes A + A \otimes A \otimes A \otimes a \end{array} \right) \\ &+ \frac{\alpha + \beta}{16} (A \otimes A \otimes a \otimes a + a \otimes a \otimes A \otimes A) + \\ &+ \left(\frac{3\alpha + 3\beta}{32} \right) \left(\begin{array}{l} A \otimes a \otimes A \otimes a + a \otimes A \otimes a \otimes A + \\ + A \otimes a \otimes a \otimes A + a \otimes A \otimes A \otimes a \end{array} \right) \\ &+ \left(\frac{\alpha + 3\beta}{32} \right) \left(\begin{array}{l} A \otimes a \otimes a \otimes a + a \otimes A \otimes a \otimes a + \\ + a \otimes a \otimes A \otimes a + a \otimes a \otimes a \otimes A \end{array} \right) + \frac{\beta}{8}a \otimes a \otimes a \otimes a \end{aligned}$$

This tells us two generation before the possible distribution of all ancestral genes. Generally, we can obtain n generations before ancestral distributions by taking plenary copowers. Of course, we also can take other types of copowers. Then after from that distribution by random mating, we will have the present generation y . In this sense, we can solve the power equations like $x^{(n)} = y$ or $x^{[n]} = y$ to get satisfying biological information.

7. The in-evolution operator. In this section, we will discuss a kind of special map in a population space. We call it in-evolution operator. We also define the equilibrium state which captures the biological essence in the coalgebraic structure of inheritance. We apply the Brower's fixed point theorem in topology to prove that there exists a equilibrium state.

Let Ω be a population space and V be a subspace of Ω , which is isomorphic to S_0 in section 2. Suppose there is a genetic coalgebra (V, Δ) , which has a character ϕ . We give the following definitions:

- The linear map $S_l: V \rightarrow V$, as defined by $S_l = (\phi \otimes id) \Delta$, is called left in-evolution operator. Similarly, we call $S_r = (id \otimes \phi) \Delta$ the right in-evolution operator. If S is both left and right in-evolution operator, we call it the in-evolutionary operator.
- For any state $x \in V$, if it satisfies $S_l(x) = x$, we say x is a left equilibrium state. Similarly, we can define a right equilibrium state. If a state is both left and right equilibrium state, we call it an equilibrium.

THEOREM 7.1. *Let (C, Δ) be a cocommutative coalgebra which has a genetic realization, then there is an equilibrium state.*

Proof. Since C has a genetic realization, we can choose a natural basis for it, for instance, $\{e_1, e_2, \dots, e_n\}$. Let C_0 be the convex combination of e_1, e_2, \dots, e_n , then we define a character ϕ by $\phi(e_k) = 1$ and linearly extend it onto C_0 . Since a character exists by our theorem above, it is easy to check this ϕ is a character. Since C is cocommutative, we have

$$S_l = (\phi \otimes id) \Delta = (\phi \otimes id) \tau \Delta = (id \otimes \phi) \Delta = S_r.$$

Then, $S (= S_l = S_r)$ is the in-evolution operator.

It is clear that S is a linear map from C_0 to C_0 . By the Brower's fixed point theorem, there is a point $x \in C_0$ such that $S(x) = x$. Up to now, we complete the proof. \square

REMARK 9. *The significance of left in-evolution operators is that they enable us to trace from the present generation back to their parental generation by the left path (for example, a father's genealogical trees). Similarly, we have the significance of right in-evolution operators and in-evolution operators. The interpretation of equilibrium states is that if we trace back one generation before, the population is still in equilibrium.*

8. Conclusion. As we have demonstrated, coalgebraic structures of genetic inheritance come into existence naturally, and they are new mathematical structures for genetics. In algebraic structure of genetic inheritance, the multiplication represents a forward dynamic system over generations. That is, once we take a product of two populations or genes of the reproduction process, we moves to the next generation. So it is logical to look at the reproduction process backward. Once we trace back from the present generation to previous generations, we get the coalgebraic structure of the reproduction process. These coalgebras are not the dual coalgebras of general genetic algebras. They are new coalgebraic structures involved in genetics. The evolution of a population in terms of algebras requires the study of all kinds of powers of the population x . Although it is tough to study power $x^{(n)}$ or $x^{[n]}$, it is generally impossible to solve power equations $x^{(n)} = a$ or $x^{[n]} = b$ in an algebraic system. Resolving these power equations $x^{(n)} = a$ or $x^{[n]} = b$ means to get a population x that is a population of n generations before of population a or b . We call this backward evolution. In coalgebraic structures of genetics, we can solve them in a sense that we can get genetic information of previous generations. To construct a phylogenetic tree from the present gene or population, all kinds of

copowers provide tools; and in-evolution operators can specify a path. Mathematically, to establish coalgebraic structures for genetic inheritance, we have to define many basic and fundamental concepts as we have already done in this paper. As a new theoretical framework, it is obvious that there are lots of theoretical issues and applied questions that need to be addressed properly and resolved. For example:

1. How to establish general in-evolution equation in genetic coalgebra setting?
2. How to characterize all stationary in-evolution operators?
3. How to develop a theory about coalgebraic dynamics under selection and migration?
4. How to apply genetic coalgebra to DNA sequence evolution?

It is also obvious that we need detailed coalgebraic models for many different and specific situations. We post them as our open problems.

Acknowledgements. We would like to express our thanks to Professor Xiao-Song Lin who gave us very valuable suggestions and help in writing the paper. We are also grateful to Professor Keh-Shin Lii who gave us useful suggestions when we discussed some ideas last summer. We also would like to thank Professor Shi-Zhong Xu for encouraging the first author to study quantitative genetics.

REFERENCES

- [1] Mendel, G., Experiments in Plant-Hybridization. In James A. Peters, editor, *Classic Papers in Genetics*, pages 1-20. Prentice-Hall, Inc., 1959.
- [2] Serebrowsky, A. On the properties of the Mendelian equations. *Doklady A.N.SSSR.* 2, 1934, 33-36. (in Russian)
- [3] Glivenkov, V. Algebra mendelienne *Comptes rendus (Doklady) de l'Acad. des Sci. de l'URSS* 4, (13), 1936, 385-386. (in Russian).
- [4] Kostitzin, V. A. Sur les coefficients mendeliens d'heredite. *Comptes rendus de l'Acad. des Sci.* 206, 1938, 883-885. (in French)
- [5] Etherington, I. M. H. Non-associative algebra and the symbolism of genetics. *Proc. Roy. Soc. Edinburgh B* 61, 1941, 24-42.
- [6] Gonshor, H. Contributions to genetic algebras. *Proc. Edinburgh Math. Soc* (2) 1973, 273-279.
- [7] Schafer, R. D. *An Introduction to Non-associative Algebras.* Acad. Press, New York, 1966.
- [8] Holgate, P. Sequences of powers in genetic algebras. *J. London Math.*, 42, 489-496, 1967.
- [9] Holgate, P. Selfing in genetic algebras. *J. Math. Biology*, 6, 197-206, 1978.
- [10] Hench, I. Sequences in genetic algebras for overlapping generations. *Proc. Edinburgh Math. Soc.* (2) 18, 19-29, 1972.
- [11] Reiser, O. Genetic algebras studied recursively and by means of differential operators. *Math. Scand.* 10, 25-44, 1962.
- [12] Abraham, V. M. *Linearising quadratic transformations in genetic algebras.* Thesis, Univ. of London 1976.
- [13] Worz-Busekros, A. *Algebras in Genetics, Lecture Notes in Biomathematics*, 36, Springer-Verlag, Berlin, 1980.
- [13] Lyubich, Y. I. *Mathematical Structures in Population Genetics.* Springer-Verlag, New York, 1992.
- [14] Reed, M. L. Algebraic Structure of Genetic Inheritance. *Bulletin of the American Mathematical Society.* 34, 2, 107-130, 1997.

Received on Dec. 6, 2003. Revised on May. 16, 2004.

E-mail address: tian@math.ucr.edu

E-mail address: bai-lian.li@ucr.edu