ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Interdepartment Semester Project

Algebraic Approach to Population Genetics

Grégoire Aubry

Tutors :

Professor Kathryn Hess-Bellwald (Director), DMA-EPFL

Dr. Jérôme Goudet (Advisor), IZEA-UNIL

DMA

July 2001

Contents

Chapter 1

Introduction

1.1 Motivation

In this paper we introduce mathematical structures and tools in order to model the genetic development of a given population. We would like to provide answers (at least partial ones) to questions such as the following :

- in which state will be the population considered after r generations ?
- is the given population going to converge to an equilibrium state ? How fast ?

We study here only the inheritance of non sex linked traits. This leads to certain commutative algebras which are in general not associative. *(Note that much of the theory developed here can be made as well for the non commutative case.)* This class of algebras is rather wide and few relevant statements can be made for the behaviour of populations without further insight into the peculiar structure of these algebras.

In the next section, we develop a very basic example of algebra arising in genetics. In chapter 2 we give the algebraic background in order to deal with chapter 3 and 4. In those three chapters, we analyze populations characterized by only one genetic trait. But in chapter 5, we set up constructions able to describe the genetics of populations with several traits.

Throughout this notes we make the following two assumptions :

- the populations considered are infinitely large,
- there is no selection, i.e. all zygotes have the same fertility).

1.2 Introductory example

Consider an infinitely large, randomly mating population of diploid individuals that differ genetically at one or several loci. Let $a_1, \ldots, a_n$ be the genetically distinct gametes produced by this population. The state of the population can be described by a vector of gamete frequencies.

The union of gametes a_i and a_j forms a zygote a_ia_j, which can produce a gamete a_k. In absence of selection, one can define the segregation rates. The **segregation rate** γ_{ijk} is the probability that the zygote a_ia_j produces a gamete a_k. By definition the γ_{ijk}'s satisfy the following relations

$$0 \leq \gamma_{ijk} \leq 1, \ i, j, k = 1, ..., n \tag{1.1}$$

$$\sum_{k=1}^{n} \gamma_{ijk} = 1, \ i, j = 1, ..., n. \tag{1.2}$$

Consider the gametes $a_1, \ldots, a_n$ as abstract elements, free over the field $\mathbb{R}$. They span an n-dimensional vector space

$$\mathcal{W} := \left\{ \sum_{i=1}^{n} \alpha_i a_i \mid \alpha_i \in \mathbb{R} \quad i = 1, \ldots, n \right\}$$

Using the segregation rates γ_{ijk} one can define a multiplication in $\mathcal{W}$ by

$$a_ia_j = \sum_{k=1}^{n} \gamma_{ijk} a_k \quad i, j = 1, ..., n \ ,$$

extended bilinearly onto $\mathcal{W} \times \mathcal{W}$. Like this, $\mathcal{W}$ is equipped with a commutative algebra structure. It is called the **gametic algebra** and denoted by $\mathcal{G}$.

Remark. Each element $a = \sum_{i=1}^{n} \alpha_i a_i$ in $\mathcal{G}$ such that $0 \leq \alpha_i \leq 1$ for all i, and $\sum_{i=1}^{n} \alpha_i = 1$ corresponds to an actual population, in which α_i is the frequency of gametes a_i.

Here is a special case of gametic algebras : the **gametic algebra for simple Mendelian inheritance**. Consider a randomly mating population of diploid individuals that differ in an autosomal locus with alleles $a_0, \ldots, a_n$. Assume that all zygotes have equal fertility and that there is no mutation, i.e. homozygotes a_ia_i produce gametes a_i only and heterozygotes a_ia_j $(i \neq j)$ produce gametes a_i and a_j in equal proportions. So we can write the segregation rates as

$$\gamma_{ijk} = \frac{1}{2}(\delta_{ik} + \delta_{jk}) \quad i, j, k = 1, \ldots, n$$

where δ_{rs} is the Kronecker delta. The γ_{ijk}'s are called the **simple Mendelian segregation rates**.

Assume the zygotes are formed by random union of gametes. Then we can describe the genetics by the gametic algebra $\mathcal{G}_{Mend}$ spanned by the elements $a_0, \ldots, a_n$ whose multiplication is given by

$$a_ia_j = \frac{1}{2}(a_i + a_j)$$

and its bilinear extension.

Chapter 2

Algebraic preliminaries

2.1 Some definitions and results

Let $\mathbb{K}$ be a field, $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$.

Definition 2.1 *Let $\mathcal{A}$ be an algebra over $\mathbb{K}$. Assume that $\mathcal{A}$ admits a basis $\{a_1, \ldots, a_n\}$ such that the* ***multiplication constants*** *γ_{ijk}, with respect to this basis, are given by*

$$a_i a_j = \sum_{k=1}^{n} \gamma_{ijk} a_k \quad i, j = 1, ..., n.$$

We say that $\mathcal{A}$ is an algebra ***with genetic realization*** *if the multiplication constants γ_{ijk} verify :*

$$(i) \qquad \gamma_{ijk} \text{ are real numbers}$$

$$(ii) \qquad 0 \leq \gamma_{ijk} \leq 1 \tag{2.1}$$

$$(iii) \qquad \sum_{k} \gamma_{ijk} = 1 \tag{2.2}$$

In that case, the basis $\{a_1, \ldots, a_n\}$ is called a ***natural basis****.*

Example. Every gametic algebra has a genetic realization.

Remark. An algebra with genetic realization may have many different natural bases. Indeed, let us consider the gametic algebra for simple Mendelian inheritance with $\{a_1, \ldots, a_n\}$ as natural basis and define

$$b_i = \sum_{k=1}^{n} \beta_{ki} a_k \qquad \text{with} \qquad \sum_{k=1}^{n} \beta_{ki} = 1 \qquad \text{for } i = 1, \ldots, n.$$

Then, $\{b_1, \ldots, b_n\}$ is a natural basis.

Proof. Let $\{b_1, \dots, b_n\}$ be a basis of the assumed form. The multiplication constants with respect to this basis follow from

$$\begin{aligned} b_i b_j &= \left(\sum_{k=1}^{n} \beta_{ki} a_k\right) \left(\sum_{l=1}^{n} \beta_{lj} a_l\right) \\ &= \sum_{k,l} \beta_{ki} \beta_{lj} a_k a_l \\ &= \sum_{k,l} \beta_{ki} \beta_{lj} \frac{1}{2}(a_k + a_l) \\ &= \frac{1}{2} \sum_{k} \beta_{ki} a_k \left(\sum_{l} \beta_{lj}\right) + \frac{1}{2} \sum_{l} \beta_{lj} a_l \left(\sum_{k} \beta_{ki}\right) \\ &= \frac{1}{2} b_i + \frac{1}{2} b_j \qquad i, j = 1, \dots, n. \quad \square \end{aligned}$$

Definition 2.2 *Let $\mathcal{A}$ be an algebra over $\mathbb{K}$. $\mathcal{A}$ is called* ***baric****, if it admits a non trivial homomorphism of algebras $\omega : \mathcal{A} \to \mathbb{K}$. In that case ω is called a* ***weight homomorphism****.*

Remark. The weight homomorphism may not be unique.

We have the immediate result :

Proposition 2.3 *Let $(\mathcal{A}, \omega)$ be an n-dimensional baric algebra over $\mathbb{K}$. Then $Ker\omega$ is an $(n-1)$-dimensional ideal of $\mathcal{A}$.*

Definition 2.4 *An algebra $\mathcal{A}$ over a $\mathbb{K}$ equipped with a norm $||.|| : \mathcal{A} \to \mathbb{K}$ is a* ***Banach algebra*** *if*

(i) $\mathcal{A}$ is complete as a vector space

(ii) the norm is submultiplicative, i.e. $||xy|| \leq ||x||\ ||y||$ for all $x, y \in \mathcal{A}$.

Proposition 2.5 *Let $\mathcal{A}$ be an algebra over $\mathbb{K}$. If it has a genetic realization with respect to the natural basis $\{a_1, \dots, a_n\}$, then :*

1) $\mathcal{A}$ is baric

2) $\mathcal{A}$ is a (non associative) Banach algebra with respect to the norm

$$||x|| := \sum_{i=1}^{n} |\xi_i| \quad \textit{for} \quad x = \sum_{i=1}^{n} \xi_i a_i \in \mathcal{A}$$

Proof. The multiplication constants with respect to the natural basis $\{a_1, \dots, a_n\}$ satisfy the axioms (2.1) and (2.2).

1) Define $\omega : \mathcal{A} \to \mathbb{K}$ by $\omega(a_i) = 1, i = 1, \dots, n$, linearly extended onto $\mathcal{A}$. In view of relation (2.2), the mapping ω is compatible with the multiplication in $\mathcal{A}$.

2) First show that $\mathcal{A}$ is complete as a vector space. Let $x_p = \sum_{i=1}^{n} \xi_i^p a_i$ with $p \in \mathbb{N}$ be a Cauchy sequence. Then $\forall \varepsilon > 0$ there exists $p_0 \in \mathbb{N}$ such that, for every $p, q \geq p_0$, we have :

$$||x_p - x_q|| = \sum_i |\xi_i^p - \xi_i^q| \quad < \quad \varepsilon \ ,$$

$$|\xi_i^p - \xi_i^q| \quad < \quad \varepsilon \quad \forall i \ ,$$

thus $\{\xi_i^p\}_{p\in\mathbb{N}}$ is a Cauchy sequence which converges, say to ξ_i, as $\mathbb{K} = \mathbb{R}$ (or $\mathbb{C}$) is complete. Finally, $\{x_p\}_{p\in\mathbb{N}}$ converges to $x = \sum \xi_i a_i$ and $\mathcal{A}$ is complete.

From relations (2.1) and (2.2), it follows that this particular norm is submultiplicative. Indeed, let $x = \sum_{i=1}^{n} \xi_i a_i \ , \ y = \sum_{j=1}^{n} \eta_j a_j \ \in \mathcal{A}$, then

$$\begin{aligned} ||xy|| &= ||\sum_k (\sum_{i,j} \xi_i \gamma_{ijk} \eta_j) a_k|| \\ &= \sum_k |\sum_{i,j} \xi_i \gamma_{ijk} \eta_j| \\ &\leq \sum_k \sum_{i,j} |\xi_i \gamma_{ijk} \eta_j| \\ &= \sum_k \sum_{i,j} |\xi_i| \, |\gamma_{ijk}| \, |\eta_j| \\ &= \sum_i |\xi_i| (\sum_k \gamma_{ijk}) \sum_j |\eta_j| \\ &= ||x|| \, ||y||. \ \square \end{aligned}$$

Remark. Consider an algebra $\mathcal{A}$ with genetic realization and let $\{a_1, \ldots, a_n\}$ be a canonical basis of $\mathcal{A}$. The elements $x = \sum_{i=1}^{n} \xi_i a_i$ of $\mathcal{A}$ with weight 1 (i.e. $\omega(x) = 1$) are of particular interest. Indeed, if the ξ_i's are non negative, x represents the distribution of the genotypes a_i's in a population.

Finally we give some characterizations of baric algebras and of algebras with genetic realization.

Lemma 2.6 *Let $\mathcal{A}$ be an n-dimensional algebra over $\mathbb{R}$. The following conditions are equivalent :*

(i) *$\mathcal{A}$ is baric*

(ii) *$\mathcal{A}$ has a basis $\{a_1, \ldots, a_n\}$, such that the multiplication constants defined by $a_i a_j = \sum_{k=1}^{n} \gamma_{ijk} a_k$ satisfy $\sum_{k=1}^{n} \gamma_{ijk} = 1 \quad i, j = 1, \ldots, n$*

(iii) *$\mathcal{A}$ has an $(n-1)$-dimensional ideal $\mathcal{I}$ and $\mathcal{A}^2 \not\subseteq \mathcal{I}$.*

Proof.

1) ⇒ 2) Let $\omega : \mathcal{A} \to \mathbb{K}$ be a weight homomorphism. By proposition 2.3, $Ker\omega$ is then an $(n-1)$-dimensional ideal of $\mathcal{A}$. Let $\{d_2, \ldots, d_n\}$ be a basis of $\mathcal{I}$. Since ω is non trivial there is an $a \in \mathcal{A}$ such that $\omega(a) \neq 0$. The elements

$$\begin{aligned} a_1 &:= \frac{1}{\omega(a)} \\ a_i &:= a_1 - d_i \text{ for } i = 2, \ldots, n \end{aligned}$$

form a basis of $\mathcal{A}$ with $\omega(a_i) = 1$ for $i = 1, \ldots, n$.
Since $\omega(a_i a_j) = \omega(a_i)\,\omega(a_j)$, the multiplication constants γ_{ijk} of $\mathcal{A}$ with respect to this basis satisfy the relation $\sum_{k=1}^{n} \gamma_{ijk} = 1$ for all i, j.

2) ⇒ 3) Define $d_i := a_1 - a_i$ for $i = 2, \ldots, n$. One can easily check that $\mathcal{I}$ generated by $\{d_2, \ldots, d_n\}$ is an $(n-1)$-dimensional ideal of $\mathcal{A}$.
In addition

$$\begin{aligned} a_1^2 &= \sum_{k=1}^{n} \gamma_{11k} a_k \\ &= \gamma_{111} a_1 + \sum_{k=2}^{n} \gamma_{11k}(a_1 - d_k) \\ &= (\sum_{k=1}^{n}) \gamma_{11k} a_1 - \sum_{k=2}^{n} \gamma_{11k} d_k \\ &= a_1 - \sum_{k=2}^{n} \gamma_{11k} d_k \notin \mathcal{I}. \end{aligned}$$

3) ⇒ 1) By $\varphi : x \longmapsto x + \mathcal{I}$ we define a linear mappping of $\mathcal{A}$ onto $\mathcal{A}/\mathcal{I}$. The vector space $\mathcal{A}/\mathcal{I}$ is isomorphic with $\mathbb{K}$. Since $\mathcal{I}$ is an ideal of $\mathcal{A}$ we have

$$\begin{aligned} \varphi(x) + \varphi(y) &= x + \mathcal{I} + y + \mathcal{I} = x + y + \mathcal{I} = \varphi(x+y) \\ \varphi(x)\,\varphi(y) &= (x + \mathcal{I})(y + \mathcal{I}) = xy + x\mathcal{I} + y\mathcal{I} + \mathcal{I}^2 = xy + \mathcal{I} = \varphi(xy), \end{aligned}$$

thus φ is an algebra homomorphism. And then $\mathcal{A}/\mathcal{I}$ is either isomorphic with $\mathbb{K}$ or with the algebra over $\mathbb{K}$ whose multiplication is nil. The latter possibility can be excluded as $\mathcal{A}^2 \not\subseteq \mathcal{I}$. □

Lemma 2.7 *Let $\mathcal{A}$ be an n-dimensional algebra orver $\mathbb{R}$. The following conditions are equivalent :*

1) $\mathcal{A}$ has a genetic realization

2) $\mathcal{A}$ is baric and the linear manifold of all elements of weight 1 contains an $(n-1)$-dimensional simplex $\mathcal{S}$ with $\mathcal{S}^2 \subseteq \mathcal{S}$

Proof.

1) $\Rightarrow$ 2) Let $\{a_1, \dots, a_n\}$ be a basis of $\mathcal{A}$ satisfying relations (2.1) and (2.2). Define $\omega : \mathcal{A} \to \mathbb{R}$ by $\omega(a_i) = 1$ for $i = 1, \dots, n$. Then $(\mathcal{A}, \omega)$ is baric. Let $\mathcal{S}$ be the convex hull of $a_1, \dots, a_n$. In niew of relations (2.1) and (2.2) we have $a_i a_j \in \mathcal{S}$ for $i = 1, \dots, n$, hence $\mathcal{S}^2 \subset \mathcal{S}$.

2) $\Rightarrow$ 1) Let $\omega : \mathcal{A} \to \mathbb{R}$ be a weight homomorphism. By assumption, the simplex $\mathcal{S}$ is the convex hull of n linearly independent elements $a_1, \dots, a_n$ with $\omega(a_i) = 1$ for all i. These elements form a basis of $\mathcal{A}$. Since $\mathcal{S}^2 \subseteq \mathcal{S}$, the multiplication constants γ_{ijk} of $\mathcal{A}$ with respect to this basis satisfy the relation $0 \le \gamma_{ijk} \le 1$ for all i, j, k.
Using lemma 2.6, we finally have $\sum_k \gamma_{ijk} = 1$. $\square$

We define now two types of products whose interpretations are given in a further remark.

Definition 2.8 *Let $\mathcal{A}$ be an algebra. If x is an element and $\mathcal{U}$ a subalgebra of $\mathcal{A}$,*

*1) the **principal powers** of x and of the subalgebra $\mathcal{U}$ respectively are defined by*

$$x^1 := x, \qquad x^{r+1} := (x^r)x, \qquad r \in \mathbb{N}$$
$$\mathcal{U}^1 := \mathcal{U}, \quad \mathcal{U}^{r+1} := < (\mathcal{U}^r)\mathcal{U} >, \quad r \in \mathbb{N}$$

*2) the **plenary powers** of x and $\mathcal{U}$ are defined by*

$$x^{[1]} := x, \qquad x^{[r+1]} := x^{[r]}x^{[r]}, \qquad r \in \mathbb{N}$$
$$\mathcal{U}^{[1]} := \mathcal{U}, \quad \mathcal{U}^{[r+1]} := < \mathcal{U}^{[r]}\mathcal{U}^{[r]} >, \quad r \in \mathbb{N}$$

where $< B >$ is the algebra generated by B.

By definition, we have the properties stated in the following proposition.

Proposition 2.9 *Let $\mathcal{U}$ be a subalgebra of $\mathcal{A}$. Then :*

1) the principal powers of $\mathcal{U}$ are ideal of $\mathcal{A}$

2) for $r \in \mathbb{N}$, we have $\mathcal{U}^{r+1} \subseteq \mathcal{U}^r$ and $\mathcal{U}^{[r+1]} \subseteq \mathcal{U}^{[r]}$.

Remark. In biological terms, these products represent mating types. If x represents a population, the sequence of principal powers $x^1 = x, x^2, x^3, \dots$ represents the filial generations obtained by random mating with the initial population. Regarding the sequence of plenary powers $x^{[1]} = x, x^{[2]}, x^{[3]}, \dots$, it represents the filial generations obtained by random mating within generations. Proposition 2.5 establishes a norm with respect to which one can define and study the convergence of such sequences.

Remark. Since the product is interpreted as a mating type, it explains why we consider nonassociative algebras : mating the product of populations A and B with population C doesn't produce in general the same population as mating population A with the product of populations B and C, that is : (AB)C $\neq$ A(BC).

We give some definitions about nilpotence in order to state proposition 3.11 which is useful to demonstrate further results.

Definition 2.10 *Let $\mathcal{A}$ be an algebra and $\mathcal{U}$ a subalgebra of $\mathcal{A}$.*

*1) $x \in \mathcal{A}$ is called **nilpotent** of index k if $x^k = 0$ but $x^{k-1} \neq 0$.*

*2) $\mathcal{U}$ is called **nil** if all elements of $\mathcal{U}$ are nilpotent.*

*3) $\mathcal{U}$ is called **nilpotent** of index k if $\mathcal{U}^k = 0$ but $\mathcal{U}^{k-1} \neq 0$.*

2.2 The rank equation

Let $\mathcal{A}$ be an n-dimensional algebra over $\mathbb{K}$. We know that for every element $a \in \mathcal{A}$ there is a unique normalized polynomial that annihilates a, namely the minimal polynomial of a. We would like to find now a polynomial in principal powers that annihilates every element of $\mathcal{A}$.

Let $\mathbb{F} = \mathbb{K}(\xi_1, \dots, \xi_n)$ be the field of rational functions in n indeterminates over $\mathbb{K}$ and let $\mathcal{A}_{\mathbb{F}} := \mathcal{A} \otimes_{\mathbb{K}} \mathbb{F}$ be the algebra over the extension $\mathbb{F}$ of the field $\mathbb{K}$. Let us fix a basis $\{a_1, \dots, a_n\}$ of $\mathcal{A}$. Then every element $x \in \mathcal{A}$ can be represented by

$$x = \sum_{i=1}^{n} \xi_i a_i$$

where the ξ_i's are elements of $\mathbb{K}$. If the ξ_i's are regarded as indeterminates, then $x \in \mathcal{A}_{\mathbb{F}}$ and x is called a **generic element** of $\mathcal{A}$. Therefore, applying to $\mathbb{F}$ instead of $\mathbb{K}$ similar arguments than the ones used to find the minimal polynomial of an element a, we find that there is a normalized polynomial q_x with coefficients in $\mathbb{F}$ which annihilates $x \in \mathcal{A}_{\mathbb{F}}$:

$$q_x(X) = X^r + \beta_1(\xi_1, \dots, \xi_n)X^{r-1} + \dots + \beta_{r-1}(\xi_1, \dots, \xi_n)X. \tag{2.3}$$

In that case, q_x is called the **rank polynomial** of $\mathcal{A}$ and r is the **rank** of $\mathcal{A}$.

Remark. For all $a \in \mathcal{A}$, we have $q_x(a)$.

One can show that the coefficients β_i of q_x are polynomials of degree i, for $i = 1, \dots, s$, in the coordinates $\xi_1, \dots, \xi_n$ of x with respect to the basis $\{a_1, \dots, a_n\}$ of $\mathcal{A}$.

As an abreviation we introduce

$$\beta_i(x) := \beta_i(\xi_1, \dots, \xi_n), \quad i = 1, \dots, r-1$$

Then we can write

$$q_x(X) = X^r + \beta_1(x)X^{r-1} + \dots + \beta_{r-1}(x)X.$$

The **rank equation** of the algebra $\mathcal{A}$ is defined by $q_x(x) = 0$.

Chapter 3

Particular algebras

3.1 Train algebras

In many algebras that arise in connection with genetic problems the coefficients of the rank polynomial have a particular form, depending only on a $\mathbb{K}$-valued function.

Definition 3.1 *Let $(\mathcal{A}, \omega)$ be a baric algebra. $\mathcal{A}$ is called a* ***train algebra*** *if the coefficients $\beta_i(x)$ of the rank polynomial 2.3 are only functions of $\omega(x)$.*

One can show that the coefficients β_i are such that, in a train algebra, the rank equation can be written as follows

$$q_x(x) = x^r + \gamma_1\omega(x)x^{r-1} + \gamma_2\omega^2(x)x^{r-2} + \ldots + \gamma_{r-1}\omega^{r-1}(x)x = 0 \tag{3.1}$$

where the γ_i's are constant belonging to the field $\mathbb{K}$.

Then, in a train algebra $\mathcal{A}$ of rank r, every element $x \in \mathcal{A}$ of weight 1 satisfies

$$x^r + \gamma_1 x^{r-1} + \gamma_2 x^{r-2} + \ldots + \gamma_{r-1}x = 0.$$

Applying the weight homomorphism ω to this equation, we get

$$1 + \gamma_1 + \gamma_2 + \ldots + \gamma_{r-1} = 0.$$

Therefore $\lambda_0 = 1 \in \mathbb{K}$ is a zero of the polynomial

$$p(X) = X^r + \gamma_1 X^{r-1} + \gamma_2 X^{r-2} + \ldots + \gamma_{r-1}X \in \mathbb{K} < X >$$

where $\mathbb{K} < X >$ is the vector space of all "polynomials" in principal powers of a nonassociative indeterminate X over $\mathbb{K}$.
Let $\mathbb{L}$ be any extension field of $\mathbb{K}$ which contains all roots of p. So the polynomial p can be written as

$$p(X) = X(X-1)(X-\lambda_1)\cdots(X-\lambda_{r-2}) \tag{3.2}$$

where $\lambda_i \in \mathbb{L}$ for all i.

Definition 3.2 *Let $\mathcal{A}$ be a train algebra with rank equation (3.1) and let the polynomial p split into linear factors in $\mathbb{L}$ as in (3.2). Then the elements $\lambda_0 := 1, \lambda_1, \ldots, \lambda_{r-2}$ of $\mathbb{L}$ are called the* ***principal train roots*** *of $\mathcal{A}$.*

Then the rank equation (3.1) can be factorized as

$$q_x(x) = x(x - \omega(x))(x - \lambda_1\omega(x))\ldots(x - \lambda_{r-2}\omega(x)) = 0$$

From relation (3.1), we obtain the following result.

Lemma 3.3 *Let $(\mathcal{A}, \omega)$ be a train algebra of rank r. Then every element in $Ker\,\omega$ is nilpotent of an index not greater than r.*

Remark. We recall that a baric algebra may have several weight homomorphisms.

In order to show that a train algebra has exactly one weight homomorphism, we need the following lemma.

Lemma 3.4 *Let $(\mathcal{A}, \omega)$ be a baric algebra. If $Ker\,\omega$ is nil, then ω is uniquely determined.*

Proof. Let $\varphi : \mathcal{A} \to \mathbb{K}$ be a non-trivial homomorphism and let $x \in Ker\,\omega$. Then there is an $l \in \mathbb{N}$ such that $x^l = 0$, hence $\varphi(x) = 0$.
Let $x \in \mathcal{A} \setminus Ker\,\omega$, i.e $\omega(x) \neq 0$. Then we have

$$\omega(\frac{x^2}{\omega(x)} - x) = 0 \qquad \text{that is} \qquad \frac{x^2}{\omega(x)} - x \in Ker\,\omega.$$

Therefore

$$\varphi(\frac{x^2}{\omega(x)} - x) = 0 \qquad \text{i.e.} \qquad \varphi(x)(\frac{\varphi(x)}{\omega(x)} - 1) = 0$$

hence $\varphi(x) = 0$ or $\varphi(x) = \omega(x)$. But φ is non-trivial, thus $\varphi = \omega$. $\square$

Using lemmas 3.3 and 3.4 we obtain

Proposition 3.5 *A train algebra has exactly one weight homomorphism.*

We describe now a method to compute the principal powers. Let $\mathcal{A}$ be a train algebra and assume that the elements $x \in \mathcal{A}$ of weight 1 satisfy (3.1), that is

$$x^r + \gamma_1 x^{r-1} + \gamma_2 x^{r-2} + \ldots + \gamma_{r-1}x = 0.$$

Suppose the principal train roots $\lambda_1, \ldots, \lambda_{r-2}$ of $\mathcal{A}$ are known. If one has computed the first $r - 1$ principal powers $x^1 = x, x^2, \ldots, x^{r-1}$ then the following principal powers $x^{r+k}, k \in \mathbb{N}$ can be computed with the following relation

$$x^{r+k} + \gamma_1 x^{r+k-1} + \ldots + \gamma_{r-1}x^{k+1} = 0.$$

Notice that the multiplication constants of the algebra $\mathcal{A}$ are not needed for this computation.

Introducing in $\mathcal{A}$ the natural topology of an n-dimensional space $\mathbb{R}^n$ or $\mathbb{C}^n$, respectively, we have from matrix theory the following

Proposition 3.6 *Let $\mathcal{A}$ be a train algebra of rank r over $\mathbb{R}$ or $\mathbb{C}$. Assume the principal train roots $\lambda_0 := 1, \lambda_1, \ldots, \lambda_{r-2}$ satisfy $|\lambda_i| < 1, i = 1, \ldots, r-2$. Then the sequence $\{x^j\}_{j\in\mathbb{N}}$ of principal powers converges for every element of weight 1.*

Proof. Omitted. □

Remark. These facts are important in particular with respect to applications in genetics. Assume that the inheritance of a population can be described by a gametic algebra $\mathcal{A}$. Let $x \in \mathcal{A}$ represent the distribution of gametic types in the population. With random mating within the population the filial population is represented by x^2. If the filial population is randomly mated with the initial population we obtain $x^2x = x^3$ and so on. The generation sequence which is created by random mating with the initial population corresponds to the sequence of principal powers of x. Hence, if the gametic algebra $\mathcal{A}$ is a train algebra and the principal train roots of $\mathcal{A}$ are known, we can compute the distribution of gametes in the $(r+k)^{th}$ generation from those in the first $r-1$ generations, $k \in \mathbb{N}$, and we can apply the above convergence result.

As well as sequences of principal powers, sequences of plenary powers $x^{[1]} = x, x^{[2]}, x^{[3]}, \ldots$, $x \in \mathcal{A}$, are of major interest in genetics, because they include all sequences describing the generations of a population with random mating within itself.
One can show there exists a normalized polynomial equation (in plenary powers) of minimal degree which annihilates every element of $\mathcal{A}$. And out of this one can get a method to compute plenary powers (of elements of weight 1) knowing the plenary train roots and the initial data.

3.2 Genetic algebras

We describe now a subclass of train algebras.

Definition 3.7 *Let $\mathcal{A}$ be an (m+1)-dimensional algebra over $\mathbb{K}$. $\mathcal{A}$ is called **genetic** if the extended algebra $\mathcal{A}_{\mathbb{L}} := \mathcal{A} \otimes_{\mathbb{K}} \mathbb{L}$, where $\mathbb{L}$ is a suitable algebraic extension of $\mathbb{K}$, admits a basis $c_0, c_1, \ldots, c_m$ ($c_0 \in \mathcal{A}$) where the multiplication constants λ_{ijk} defined by*

$$c_i c_j = \sum_{k=0}^{m} \lambda_{ijk} c_k$$

for $i, j = 0, 1, \ldots, m$have the following properties

- *(i)* $\lambda_{000} = 1$
- *(ii)* $\lambda_{0jk} = 0$ *for* $k < j$
- *(iii)* $\lambda_{ijk} = 0$ *for* $k \leq max(i, j)$ *if* $i, j > 0$

*Such a basis $c_0, \ldots, c_m$ is called a **canonical basis**. Moreover $\lambda_{000} = 1, \lambda_{011}, \ldots, \lambda_{0mm}$ are called the **train roots** of $\mathcal{A}$.*

If we call $\lambda_0 = 1, \lambda_1, \ldots, \lambda_m$ the train roots of a genetic algebra then we understand that the enumeration is chosen in such a way that there is a basis $c_0, c_1, \ldots, c_m$ which satisfies the axioms i) - iii) above.

Proposition 3.8 *Let $\mathcal{A}$ be a genetic algebra with a canonical basis $\{c_0, c_1, \dots, c_m\}$ over a suitable extension $\mathbb{L}$ of $\mathbb{K}$. Then, for any $a \in \mathcal{A}_{\mathbb{L}}$ with $\omega(a) = 1$, $\{a, c_1, \dots, c_m\}$ form a canonical basis of $a \in \mathcal{A}_{\mathbb{L}}$.*

Proof. Let $a \in \mathcal{A}_{\mathbb{L}}$ with $\omega(a) = 1$. Then we can express a as

$$a = \sum_{i=0}^{m} \eta_i c_i \qquad \text{with } \eta_0 = 1.$$

We can verify that $\{a, c_1, \dots, c_m\}$ is canonical basis of $a \in \mathcal{A}_{\mathbb{L}}$. □

Example : The algebra $\mathcal{G}_{Mend}$ is a genetic algebra. Indeed, the elements

$$c_o := a_0, \qquad c_i := a_0 - a_i \quad \text{for } i = 1, \dots, n$$

form a canonical basis of $\mathcal{G}_{Mend}$ as we have the relations

$$c_0^2 = c_0, \qquad c_0 c_i = \frac{1}{2} c_i, \quad \text{and} \quad c_i c_j = 0 \text{ for } i, j = 1, \dots, n.$$

Thus we deduce the train roots : $\lambda_0 = 1$ and $\lambda_i = \frac{1}{2}$ for $i = 1, \dots, n$. ⊡

Proposition 3.9 *Let $\mathcal{A}$ be a genetic algebra. Then $\mathcal{A}$ is a train algebra.*

Proof. Omitted.□

From proposition 3.5 we obtain

Corollary 3.10 *The weight homomorphism of a genetic algebra is uniquely determined.*

Proposition 3.11 *Let $\mathcal{A}$ be a genetic algebra over a suitable extension $\mathbb{L}$ of $\mathbb{K}$ with weight homomorphism ω. Then $Ker\,\omega$ is nilpotent.*

Proof. Using a characterization of genetic algebras by chains of ideals and doing an induction on the principal powers of $Ker\,\omega$ leads to the conclusion. □

This proposition yields the following corollary that tells us it is worthwhile to define genetic algebras as distinct from train algebras.

Corollary 3.12 *There is a train algebra which is not a genetic algebra.*

Proof. Let $\mathcal{A}$ be the algebra over $\mathbb{R}$ spanned by $c_0, \dots, c_5$ with multiplication table given by

	c_0	c_1	c_2	c_3	c_4	c_5
c_0	c_0	$\frac{1}{2}c_1$	$\frac{1}{2}c_2$	$\frac{1}{2}c_3$	$\frac{1}{2}c_4$	$\frac{1}{2}c_5$
c_1		0	c_3	c_4	0	$-c_3$
c_2			0	c_5	c_3	0
c_3				0	0	0
c_4					0	0
c_5						0

Let $\mathcal{I}$ be spanned by $c_1, \ldots, c_5$. $\mathcal{I}$ is an 5-dimensional ideal of $\mathcal{A}$ and $\mathcal{A}^2 \not\subseteq \mathcal{I}$. By lemma 2.6, $\mathcal{A}$ is then baric with weight homomorphism $\omega : \mathcal{A} \to \mathbb{R}$ defined by $\omega(c_0) = 1$ and $\omega(c_i) = 0$ for $i = 1, \ldots, 5$ and $\mathcal{I} = Ker\,\omega$.

The multiplication table reads $\mathcal{I}^r$, for $r \geq 2$, is spanned by c_3, c_4, c_5 and then the ideal $\mathcal{I}$ is not nilpotent. From proposition 3.11, $\mathcal{A}$ is not a genetic algebra.

Let us show now that $\mathcal{A}$ is a train algebra. Let $y = \xi_0 c_0 + x$, with $x \in \mathcal{I}$ be an arbitrary element of $\mathcal{A}$, thus $\omega(y) = \xi_0$.
First we have

$$\begin{aligned} x &= \sum_{i=1}^{5} \xi_i c_i \\ x^2 &= 2(\xi_1\xi_2 - \xi_1\xi_5 + \xi_2\xi_4)\, c_3 + 2\,\xi_1\xi_3\xi_4 + 2\,\xi_2\xi_3\xi_5 \\ x^3 &= 2\,\xi_1(\xi_1\xi_2 - \xi_1\xi_5 + \xi_2\xi_4)\, c_4 + 2\,\xi_2(\xi_1\xi_2 - \xi_1\xi_5 + \xi_2\xi_4)\, c_5 \\ x^4 &= 0. \end{aligned}$$

Therefore,

$$\begin{aligned} y^2 - \xi_0 &= x^2 \\ (y^2 - \xi_0)\,(y - \tfrac{1}{2}\,\xi_0) &= x^3 \\ (y^2 - \xi_0)\,(y - \tfrac{1}{2}\,\xi_0)^2 &= x^4 = 0. \end{aligned}$$

Thus $\mathcal{A}$ is a train algebra of rank 4 with principal train roots 1 and $\frac{1}{2}$ (twice). □

Remark. For a baric algebra $(\mathcal{A}, \omega)$, there are characterizations of $\mathcal{A}$ as a genetic algebras in terms of particular kinds of powers of $Ker\,\omega$.

3.3 Special train algebras

In general those algebras which arise in connection with genetic problems are more special than train algebras.

Definition 3.13 *Let $\mathcal{A}$ be a baric algebra over $\mathbb{K}$ with weight homomorphism ω. $\mathcal{A}$ is called a* ***special train algebra*** *if $Ker\,\omega$ is nilpotent and the principal powers $(Ker\,\omega)^i$, $i \in \mathbb{N}$, are ideals of $\mathcal{A}$.*

Remark. The latter property is trivial only for associative algebras.

Here is a characterization of special train algebras in terms of genetic algebras with an additional property.

Proposition 3.14 *Let $(\mathcal{A}, \omega)$ be a baric algebra. Then the following statements are equivalent*

1) $\mathcal{A}$ is a special train algebra.

2) $\mathcal{A}$ is genetic and all principal powers of $Ker\,\omega$ are ideals of $\mathcal{A}$.

Proof. Omitted. □

A trivial consequence of this theorem is

Corollary 3.15 *Every special train algebra is genetic.*

The definitions of special train algebras and genetic algebras are not equivalent.

Proposition 3.16 *There is a genetic algebra which is not a special train algebra.*

Proof. Let $\mathcal{A}$ be the algebra spanned by $\{c_0, \ldots, c_5\}$ over $\mathbb{R}$ with multiplication table given by

	c_0	c_1	c_2	c_3	c_4	c_5
c_0	c_0	$\frac{1}{2}c_1$	$\frac{1}{4}c_2$	0	0	0
c_1		$\frac{1}{4}c_2$	$\frac{1}{8}c_4$	0	0	0
c_2			$\frac{1}{16}c_5$	0	0	0
c_3				0	0	0
c_4					0	0
c_5						0

Then $\{c_0, \ldots, c_5\}$ form a canonical basis of $\mathcal{A}$ and thus $\mathcal{A}$ is a genetic algebra. Moreover, $\mathcal{A}$ is baric with weight homomorphism $\omega : \mathcal{A} \to \mathbb{R}$ defined by $\omega(c_0) = 1$ and $\omega(c_i) = 0$ for $i = 1, \ldots, 5$.

The principal powers of $Ker\,\omega$ are

$Ker\,\omega$ is generated by $\{c_1, \dots, c_n\}$ over $\mathbb{R}$

$(Ker\,\omega)^2$ is generated by $\{c_2, c_4, c_5\}$ over $\mathbb{R}$

$(Ker\,\omega)^3$ is generated by $\{c_4, c_5\}$ over $\mathbb{R}$

$(Ker\,\omega)^4$ is generated by $\{0\}$ over $\mathbb{R}$

that is the principal powers form a descending sequence of subalgebras of $\mathcal{A}$. However $c_0c_2 = \frac{1}{2}\,c_3 \notin \mathcal{I}^2$. Thus $\mathcal{I}^2$ is not an ideal of $\mathcal{A}$ and $\mathcal{A}$ is not a special train algebra. $\square$

Remark. In particular cases train algebras are special train algebras. In fact, one can show that a train algebra of rank not greater than 3 over $\mathbb{K}$, char$(\mathbb{K}) \neq 2$, is a special train algebra.

Chapter 4

Idempotence and sequences

In this chapter we give some results about idempotence in train algebras, in genetic algebras and in algebras with genetic realization and results about sequences' convergence. However, we do not develop any proof.

Recall that an element $e \neq 0$ of an algebra is called **idempotent** if $e^2 = e$.

From the biological viewpoint, the existence of idempotent element is of interest because the equilibria of a population that can be described by an algebra correspond to idempotents of this algebra.

4.1 Idempotents in train algebras and genetic algebras

Lemma 4.1 *Let $\mathcal{A}$ be a baric algebra with weight homomorphism ω. For every idempotent e of $\mathcal{A}$ we have either $\omega(e) = 0$ or $\omega(e) = 1$.*

Lemma 4.2 *Let $\mathcal{A}$ be a train algebra with weight homomorphism ω. Then $\omega(e) = 1$ for every idempotent e of $\mathcal{A}$.*

Proposition 4.3 *Let $\mathcal{A}$ be a genetic algebra over $\mathbb{L}$, char$(\mathbb{L}) \neq 2$. $\mathcal{A}_{\mathbb{L}}$ has exactly one idempotent if and only if all train roots are different from $\frac{1}{2}$.*

4.2 Idempotents in algebras with genetic realization

Proposition 4.4 *Let $\mathcal{A}$ be an (m+1)-dimensional algebra over $\mathbb{R}$, $\mathbb{C}$ respectively, which has a genetic realization. Then $\mathcal{A}$ has at least one idempotent and this idempotent is in*

$$\gamma_m := \left\{ \sum_{i=0}^{m} \alpha_i a_i \mid 0 \leq \alpha_i \leq 1,\ i = 0, \ldots, m \ \text{ and } \ \sum_{i=0}^{m} \alpha_i = 1 \right\}$$

i.e. the idempotent has a real biological meaning.

Proposition 4.5 *Let $\mathcal{A}$ be an (m+1)-dimensional genetic algebra over $\mathbb{R}$, $\mathbb{C}$ respectively, which has a genetic realization. If all train roots of $\mathcal{A}$ are different from $\frac{1}{2}$ then $\mathcal{A}_{\mathbb{C}}$ has exactly one idempotent and this idempotent is in γ_m.*

From propositions 4.4 and 4.5, we obtain

Proposition 4.6 *Let $\mathcal{A}$ be a genetic algebra with a genetic realization. Then with the exception of $\lambda_0 = 1$, the absolute value of the train roots does not exceed $\frac{1}{2}$.*

Proposition 4.7 *Let $\mathcal{A}$ be a 2-dimensional algebra over $\mathbb{R}$ with train roots $\lambda_0 = 1$ and λ_1. Then the following assertions are equivalent.*

1) $\mathcal{A}$ has a genetic realization

2) $\mathcal{A}$ has an idempotent and $-\frac{1}{2} \leq \lambda_1 \leq \frac{1}{2}$

4.3 Sequences of principal powers and plenary powers in genetic algebras

If in some population the inheritance can be described by an algebra $\mathcal{A}$, that is an algebra with a genetic realization, 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 principal powers x^r or by the sequence of plenary powers $x^{[r]}$, respectively, where $x \in \mathcal{A}$ represents the initial population. If in addition the algebra $\mathcal{A}$ is genetic one can derive the convergence of sequences of principal and plenary powers directly from the train roots.

Proposition 4.8 *Let $\mathcal{A}$ be a genetic algebra over $\mathbb{R}$ or $\mathbb{C}$. If $\mathcal{A}$ has a genetic interpretation, then the sequence of principal powers of any element $x \in \mathcal{A}$ of weight 1 converges.*

Proposition 4.9 *Let $\mathcal{A}$ be a genetic algebra over $\mathbb{R}$ or $\mathbb{C}$. If all train roots different from λ_0 are less than $\frac{1}{2}$ in absolute value, then the sequence of plenary powers converges towards the idempotent of $\mathcal{A}$ (respectively of $\mathcal{A}_{\mathbb{C}}$) for any element $x \in \mathcal{A}$ of weight 1.*

The following proposition will be useful in chapter 6.

Proposition 4.10 *Let $\mathcal{A}$ be an $(m+1)$-dimensional genetic algebra with a canonical basis $c_0, \ldots, c_m$ over $\mathbb{R}$ or $\mathbb{C}$. Assume that among the train roots*

$$\lambda_0 = 1,\ \lambda_1,\ \ldots, \lambda_m \text{ exactly } \lambda_{t_1} = \cdots = \lambda_{t_r},\ t_1 < \ldots < t_r,$$

and assume that the other train roots except λ_0 satisfy $|\lambda_i| < \frac{1}{2}$. If $c_i c_j$ has no component in the direction of c_{t_ρ} , $i, j = 0, \ldots, t_\rho$, $\rho = 1, \ldots, r$, then the sequence of plenary powers jof any element $x = \sum_{i=0}^{m} \xi_i c_i$ converges towards the idempotent of $\mathcal{A}$ (resp. $\mathcal{A}_{\mathbb{C}}$) determined by the coordinates $\xi_{t_1}, \ldots, \xi_{t_r}$.

Chapter 5

Construction of new algebras

Up to now we have only considered algebras that are able to describe the behaviour of one genetic feature. If a population is characterized by several genetic traits such that each can be described by an algebra then we can try to construct an algebra which models all these features.

Hereafter we state that using linear combinations or tensor products of (basic) algebras, the resulting algebra conserves some properties of the basic algebras. The results stated in this chapter will be useful for the application in the next chapter.

5.1 Linear combination

Let $\mathbb{K}$ be a field. Let $\mathcal{W}$ be a vector space over $\mathbb{K}$ with dimension n and let $\mathcal{A}$ be an algebra over $\mathcal{W}$ which needs not to be associative nor commutative. Denote the multiplication in $\mathcal{A}$ by

$$(x, y) \longmapsto (xy)_{\mathcal{A}}$$

Proposition 5.1 *Let $A(\mathcal{W})$ be the family of all algebras over $\mathcal{W}$. For $\mathcal{A}$, $\mathcal{B} \in A(\mathcal{W})$ and $\alpha \in \mathbb{K}$ the sum $\mathcal{A} + \mathcal{B}$ and the scalar multiple $\alpha\mathcal{A}$ are defined as algebras over $\mathcal{W}$ with multiplication given by*

$$\begin{array}{llllll} (x, y) & \longmapsto & (xy)_{\mathcal{A}+\mathcal{B}} & := & (xy)_{\mathcal{A}} + (xy)_{\mathcal{B}}, & x, y \in \mathcal{W}, \\ (x, y) & \longmapsto & (xy)_{\alpha\mathcal{A}} & := & \alpha(xy)_{\mathcal{A}}, & x, y \in \mathcal{W}. \end{array}$$

1) Then $A(\mathcal{W})$ is a vector space over $\mathcal{W}$ which is called the ***vector space of all algebras over*** *$\mathcal{W}$. The zero element of this vector space is the zero algebra $\mathcal{O}$ over $\mathcal{W}$.*

2) The family $A_c(\mathcal{W})$ of all commutative algebras over $\mathcal{W}$ forms a subspace of $A(\mathcal{W})$.

By this proposition the following is well-defined.

Definition 5.2 *For $\alpha_1, \ldots, \alpha_r \in \mathbb{K}$, the* ***linear combination*** *of the algebras $\mathcal{A}_1, \ldots, \mathcal{A}_r \in A(\mathcal{W})$ is defined by*

$$\sum_{\rho=1}^{r} \alpha_\rho \mathcal{A}_\rho = \alpha_1 \mathcal{A}_1 + \ldots + \alpha_r \mathcal{A}_r$$

with multiplication given by

$$(x,y) \longmapsto (xy)_{\sum \alpha_\rho \mathcal{A}_\rho} := \sum_{\rho=1}^{r} \alpha_\rho \, (xy)_{\mathcal{A}_\rho}$$

We state now some results about linear combinations of baric algebras, of algebras with genetic realization and of genetic algebras.

Proposition 5.3 *Let $\mathcal{A}_1, \ldots, \mathcal{A}_r$ be baric algebras over $\mathcal{W}$ with the same weight homomorphism ω. Then every linear combination $\sum \alpha_\rho \mathcal{A}_\rho$ with $\sum \alpha_\rho = 1$ is a baric algebra with weight homomorphism ω.*

Proof. It suffices to verify that the mapping $\omega : \sum \alpha_\rho \ \mathcal{A}_\rho \to \mathbb{K}$ is compatible with the multiplication :

$$\begin{aligned} \omega((xy)_{\sum \alpha_\rho \, \mathcal{A}_\rho}) &= \omega\left(\sum_\rho \alpha_\rho \, (xy)_{\mathcal{A}_\rho}\right) \\ &= \sum \alpha_\rho \, \omega((xy)_{\mathcal{A}_\rho}) \\ &= \sum \alpha_\rho \, \omega(x) \, \omega(y) \\ &= \omega(x) \, \omega(y) \end{aligned}$$

which is true as the α_ρ's add to 1. □

Proposition 5.4 *Let $\mathcal{W}$ be a vector space over $\mathbb{R}$ or $\mathbb{C}$. Let $\mathcal{A}_1, \ldots, \mathcal{A}_r$ algebras over $\mathcal{W}$ which have a genetic realization with respect to the same natural basis $a_1, \ldots, a_n$ of $\mathcal{W}$. Then every convex combination of these algebras has a genetic realization with respect to $a_1, \ldots, a_n$.*

Proof. It suffices to verify that

$$\begin{aligned} (a_i a_j)_{\sum \alpha_\rho \mathcal{A}_\rho} &= \sum_\rho \alpha_\rho \, (a_i a_j)_{\mathcal{A}_\rho} \\ &= \sum_\rho \alpha_\rho \left(\sum_k \gamma_{ijk}^{(\rho)} a_k\right) \\ &= \tilde{\gamma}_{ijk} \, a_k \end{aligned}$$

where the $\tilde{\gamma}_{ijk} = \sum_{\rho,k} \alpha_\rho \gamma_{ijk}^{(\rho)}$ lie in $[0,1]$ and are such that $\sum_k \tilde{\gamma}_{ijk} = 1$. □

Proposition 5.5 *Let $\mathcal{A}_1, \ldots, \mathcal{A}_r$ be genetic algebras with respect to a common basis of $\mathcal{A}_{\mathbb{L}}$, where $\mathbb{L}$ is a suitable extension field of $\mathbb{K}$. Then every linear combination $\sum \alpha_\rho \mathcal{A}_\rho$ with $\sum \alpha_\rho = 1$ is genetic with respect to this basis and the train roots are the corresponding linear combinations of the train roots of $\mathcal{A}_1, \ldots, \mathcal{A}_r$.*

Proof. Let $\{c_0, \ldots, c_m\}$ be the common basis of $\mathcal{A}_{\mathbb{L}}$ and let $\lambda_{ijk}^{(\rho)}$ denote the train roots of $\mathcal{A}_\rho$. Then we have

$$\begin{aligned}(c_i c_j)_{\sum \alpha_\rho \mathcal{A}_\rho} &= \sum_\rho \alpha_\rho (c_i c_j)_{\mathcal{A}_\rho} \\ &= \sum_\rho \left(\sum_k \lambda_{ijk}^{(\rho)} \, c_k \right) \\ &= \sum_k \tilde{\lambda}_{ijk} \, c_k\end{aligned}$$

where $\tilde{\lambda}_{ijk} = \sum_\rho \alpha_\rho \, \lambda_{ijk}^{(\rho)}$.
Then $\{c_0, \ldots, c_m\}$ is a canonical basis of $\sum \alpha_\rho \mathcal{A}_\rho$ and that its train roots are

$$\begin{aligned}\tilde{\lambda}_0 &= \sum_\rho \alpha_\rho \lambda_0^{(\rho)} \\ \tilde{\lambda}_i &= \sum_\rho \alpha_\rho \lambda_i^{(\rho)} \qquad i = 1, \ldots, m\end{aligned}$$

where $\lambda_0^{(\rho)}, \ldots, \lambda_m^{(\rho)}$ are the principal train roots of $\mathcal{A}_\rho$. □

Remark. We can wonder now if a convex combination of special train algebras is a special train algebra or not.

Proposition 5.6 *A convex combination of special train algebras (with the same weight homomorphism) is not necessarily a special train algebra.*

Proof. Consider the following counter-example. Let $\mathcal{W}$ be the free $\mathbb{R}$-module generated by c_0, c_1, c_2, c_3. In $\mathcal{W}$ define two commutative multiplications by

1	c_0	c_1	c_2	c_3
c_0	c_0	c_1	c_3	c_3
c_1		c_2	c_3	0
c_2			0	0
c_3				0

2	c_0	c_1	c_2	c_3
c_0	c_0	c_1	c_3	c_3
c_1		c_2	$-c_3$	0
c_2			0	0
c_3				0

whereby we obtain the algebras $\mathcal{A}_1$ and $\mathcal{A}_2$. In both cases $\mathcal{I} :=< c_1, c_2, c_3 >$ is an ideal not containing $\mathcal{A}_1^2$ nor $\mathcal{A}_2^2$. By lemma 2.6, $\mathcal{A}_1$ and $\mathcal{A}_2$ are baric with the same weight homomorphism. For $i = 1, 2$ the principal powers of $\mathcal{I}$ in $\mathcal{A}_i$ satisfy

$$\mathcal{I}_{\mathcal{A}_i}^2 =< c_2, c_3 > \qquad \mathcal{I}_{\mathcal{A}_i}^3 =< 0 >$$

and are such ideals in $\mathcal{A}_1$ and $\mathcal{A}_2$, respectively. Thus $\mathcal{A}_1$ and $\mathcal{A}_2$ are special train algebras.
Now let $\mathcal{A} := \frac{1}{2}\mathcal{A}_1 + \frac{1}{2}\mathcal{A}_2$. Then, by proposition 5.3, $\mathcal{A}$ is baric and again $\mathcal{I}$ is the kernel of the weight homomorphism. But $\mathcal{I}_{\mathcal{A}}^2 =< c_2 >$ is not an ideal of $\mathcal{A}$ and so, by lemma 2.6, $\mathcal{A}$ is not a special train algebra. □

5.2 Tensor product

Let $\mathcal{A}$ be an n-dimensional algebra over the field $\mathbb{K}$ which needs not to be associative nor commutative. Let $\mathcal{A} \otimes \ldots \otimes \mathcal{A}$ (r times) be the r-fold tensor product (over $\mathbb{K}$) of the vector space $\mathcal{A}$. Then $\mathcal{A} \otimes \ldots \otimes \mathcal{A}$ is an n^r-dimensional vector space.

Let $\{a_1, \ldots, a_n\}$ be a basis of $\mathcal{A}$, then

$$a_{i_1} \otimes \ldots \otimes a_{i_r} \qquad i_\rho = 1, \ldots, n \quad \rho = 1, \ldots, r$$

is a basis of the r-fold tensor product.

The multiplication defined by

$$(a_{i_1} \otimes \ldots \otimes a_{i_r})\,(a_{j_1} \otimes \ldots \otimes a_{j_r}) := (a_{i_1} a_{j_1}) \otimes \ldots \otimes (a_{i_r} a_{j_r})$$

gives an algebra structure to the r-fold tensor product $\mathcal{A} \otimes \ldots \otimes \mathcal{A}$. This algebra is called the **r-fold tensor product of the algebra** $\mathcal{A}$ and is denoted by $\mathcal{A} \otimes \ldots \otimes \mathcal{A}$.

Let the multiplication table of $\mathcal{A}$ given by

$$a_i a_j = \sum_{k=1}^{n} \lambda_{ijk} a_k \qquad i, j = 1, \ldots, n$$

then the multiplication table of the r-fold tensor product is given by

$$(a_{i_1} \otimes \ldots \otimes a_{i_r})\,(a_{j_1} \otimes \ldots \otimes a_{j_r}) = \sum_{k_1, \ldots, k_r} \lambda_{i_1 j_1 k_1} \cdot \ldots \cdot \lambda_{i_r j_r k_r}\,(a_{k_1} \otimes \ldots \otimes a_{k_r})$$

Let $\mathcal{A}$ and $\mathcal{B}$ be algebras over $\mathbb{K}$ and let $\varphi_\rho : \mathcal{A} \to \mathcal{B}$, $\rho = 1, \ldots, r$ be algebra homomorphisms. Then defining

$$\varphi_1 \otimes \ldots \otimes \varphi_r : \mathcal{A} \otimes \ldots \otimes \mathcal{A} \longrightarrow \mathcal{B} \otimes \ldots \otimes \mathcal{B}$$

by

$$(\varphi_1 \otimes \ldots \otimes \varphi_r)(x_1 \otimes \ldots \otimes x_r) := \varphi_1(x_1) \otimes \ldots \otimes \varphi_r(x_r),$$

we obtain an algebra homomorphism and

$$Im(\varphi_1 \otimes \ldots \otimes \varphi_r) = Im(\varphi_1) \otimes \ldots \otimes Im(\varphi_r).$$

This leads to the following result :

Proposition 5.7 *Let $(\mathcal{A}, \tilde{\omega})$ be a baric algebra. Then the r-fold product $\mathcal{A} \otimes \ldots \otimes \mathcal{A}$ is baric with weight homomorphism*

$$\omega := \tilde{\omega} \otimes \ldots \otimes \tilde{\omega} \qquad (r\ \textit{times})$$

and $Ker\,\omega$ can be represented as

$$Ker\,\omega = (Ker\,\tilde{\omega} \otimes \mathcal{A} \otimes \ldots \otimes \mathcal{A}) + (\mathcal{A} \otimes Ker\,\tilde{\omega} \otimes \ldots \mathcal{A}) + \ldots + (\mathcal{A} \otimes \ldots \otimes \mathcal{A} \otimes Ker\,\tilde{\omega})$$

where "+" denotes the sum of vector spaces.

Proposition 5.8 *Let $\mathcal{A}$ be a commutative special train algebra. Then the r-fold tensor product $\mathcal{A} \otimes \ldots \otimes \mathcal{A}$ is a special train algebra.*

Proof. Omitted. $\square$

Proposition 5.9 *Let $\mathcal{A}$ be a genetic algebra. Then the r-fold tensor product $\mathcal{A} \otimes \ldots \otimes \mathcal{A}$ is a genetic algebra and the train roots are all products of r factors where the r factors run over all train roots of the algebra $\mathcal{A}$.*

Proof. Omitted. $\square$

Chapter 6

Application

Assume the following situation :

1. population of diploid individuals which differ genetically in 3 independent loci each with 2 alleles
2. population infinitely large
3. recombination rate is $\frac{1}{2}$
4. no overlapping
5. no mutation

Remark. The statements given in this chapter can be generalized to r independent loci each with m alleles.

The genetics of this population can be described by the tensor product (over the field $\mathbb{R}$) of 3 gametic algebras for simple Mendelian inheritance, i.e.

$$\mathcal{A} := \mathcal{G} \otimes \mathcal{G} \otimes \mathcal{G}$$

where $\mathcal{G}$ is the gametic algebra over the $\mathbb{R}$-vector space spanned by a_0 and a_1.

Recall that the multiplication in $\mathcal{G}$ is given by $a_i a_j = \frac{1}{2}(a_i + a_j)$ for $i, j = 0, 1$ and linearly extended.

The algebra $\mathcal{A}$ over $\mathbb{R}$ has the following multiplication table :

$$(a_{i_1} \otimes a_{i_2} \otimes a_{i_3})(a_{j_1} \otimes a_{j_2} \otimes a_{j_3}) := \frac{1}{8} \sum_{\substack{k_r \in \{i_r, j_r\} \\ r \in \{1,2,3\}}} (a_{k_1} \otimes a_{k_2} \otimes a_{k_3})$$

Since $\mathcal{G}$ is baric with weight homomorphism $\tilde{\omega} : \mathcal{G} \to \mathbb{R}$ defined by $\tilde{\omega}(a_i) = 1 \quad i = 1, 2$ and linearly extended. By proposition 5.7, $\mathcal{A} = \mathcal{G} \otimes \mathcal{G} \otimes \mathcal{G}$ is baric with weight homomorphism $\omega := \tilde{\omega} \otimes \tilde{\omega} \otimes \tilde{\omega} : \mathcal{A} \to \mathbb{R}$ defined by $\omega(a_i \otimes a_j \otimes a_k) = \omega(a_i) \otimes \omega(a_j) \otimes \omega(a_k) \quad i, j, k = 1, 2$ and linearly extended.
As $\mathbb{R} \otimes \mathbb{R} \otimes \mathbb{R} \cong \mathbb{R}$ (tensor product over $\mathbb{R}$), we can identify $\omega(a_i) \otimes \omega(a_j) \otimes \omega(a_k)$ with

$\omega(a_i) \cdot \omega(a_j) \cdot \omega(a_k)$, where $\cdot$ is the usual product in $\mathbb{R}$.

Since $\mathcal{G}$ is a special train algebra, $\mathcal{A}$ is also a special train algebra (by proposition 5.8). Consider the following basis of $\mathcal{G}$

$$c_0 := a_0 \qquad c_1 := a_0 - a_1 \tag{6.1}$$

Then the multiplication table of $\mathcal{G}$ is given by

	c_0	c_1
c_0	c_0	$\frac{1}{2}c_1$
c_1		0

Thus c_0 and c_1 form a canonical basis of $\mathcal{G}$ and the principal train roots are 1 and $\frac{1}{2}$. Define the following elements :

$$\begin{aligned}
d_0 &:= c_0 \otimes c_0 \otimes c_0 \\
d_1 &:= c_0 \otimes c_0 \otimes c_1 \\
d_2 &:= c_0 \otimes c_1 \otimes c_0 \\
d_3 &:= c_0 \otimes c_1 \otimes c_1 \\
d_4 &:= c_1 \otimes c_0 \otimes c_0 \\
d_5 &:= c_1 \otimes c_0 \otimes c_1 \\
d_6 &:= c_1 \otimes c_1 \otimes c_0 \\
d_7 &:= c_1 \otimes c_1 \otimes c_1
\end{aligned}$$

Then $d_0, \ldots, d_7$ form a basis of $\mathcal{A}$ with multiplication table given by

	d_0	d_1	d_2	d_3	d_4	d_5	d_6	d_7
d_0	d_0	$\frac{1}{2}d_1$	$\frac{1}{2}d_2$	$\frac{1}{4}d_3$	$\frac{1}{2}d_4$	$\frac{1}{4}d_5$	$\frac{1}{4}d_6$	$\frac{1}{8}d_7$
d_1		0	$\frac{1}{4}d_3$	0	$\frac{1}{4}d_5$	0	$\frac{1}{8}d_7$	0
d_2			0	0	$\frac{1}{4}d_6$	$\frac{1}{8}d_7$	0	0
d_3				0	$\frac{1}{8}d_7$	0	0	0
d_4					0	0	0	0
d_5						0	0	0
d_6							0	0
d_7								0

Hence, this basis is a canonical basis of $\mathcal{A}$. In reading this table, we get the principal train roots : 1, $\frac{1}{2}$ (3 times), $(\frac{1}{2})^2$ (3 times) and $(\frac{1}{2})^3$.

Remark. In case of more than 3 independent loci (say r) or more than 2 alleles (say $m+1$) such multiplication tables are not tractable ! But we have a more concise way to express it :

$$(c_{i_1} \otimes \ldots \otimes c_{i_r})(c_{j_1} \otimes \ldots \otimes c_{j_r}) := \begin{cases} \frac{1}{2}^{\alpha(i,j)} c_{k_1} \otimes \ldots \otimes c_{k_r} & \textit{if for every } \rho \in \{1, \ldots, r\} \\ & \textit{either } i_\rho = 0 \textit{ or } j_\rho = 0 \\ 0 & \textit{otherwise} \end{cases}$$

where $c_{k_\rho} := \begin{cases} c_{i_\rho} & \textit{if } j_\rho = 0 \\ c_{j_\rho} & \textit{if } i_\rho = 0 \end{cases}$ and $\alpha(i,j)$ is the number of non vanishing coefficients of $(i_1, \ldots, i_r, j_1, \ldots, j_r) \quad i_\rho, j_\rho = 0, \ldots, m \quad \rho = 1, \ldots, r$.

And then we get the principal train roots of $\mathcal{A}$ in applying theorem 5.9. ⊡

In order to apply proposition 4.10, let us denote the corresponding basis elements to the train root $\frac{1}{2}$ by

$$d_{11} := c_0 \otimes c_0 \otimes c_1 \qquad d_{21} := c_0 \otimes c_1 \otimes c_0 \qquad d_{31} := c_1 \otimes c_0 \otimes c_0 \tag{6.2}$$

Since $d_i d_j$ has no component with respect to $d_{\rho 1}$ for $i, j = 0, \ldots, \rho - 1$ with $\rho \in \{1, 2, 3\}$, we can apply proposition 4.10. Thus $\mathcal{A}$ has a 3-parametric family of idempotents

$$e = \sum_{i,j,k=0}^{1} \eta_{ijk} \, (c_i \otimes c_j \otimes c_k) \qquad i, j, k = 0, 1$$

The parameters, namely the coordinates $\gamma_{\rho 1}$ of $d_{\rho 1}$ for $\rho = 1, 2, 3$, can be arbitralily chosen, then the remaining coordinates have to be computed from

$$\eta_{ijk} = \gamma_{1i} \cdot \gamma_{2j} \cdot \gamma_{3k} \tag{6.3}$$

with $\gamma_{\rho 0} := 1$ for $\rho = 1, 2, 3$.

For every element $x = \sum \xi_{ijk} \, (c_{i_1} \otimes \ldots \otimes c_{i_r})$ of weight 1 (i.e. $\xi_{000} = 1$ as $\omega(c_i \otimes c_j \otimes c_k) = 0$ if $(i, j, k) \neq (0, 0, 0)$), the sequence of plenary powers $x = x^{[1]}, x^{[2]}, \ldots$ (where $x^{[i]} := x^{[i-1]} x^{[i-1]} \quad i = 2, 3, \ldots$) converges towards the following idempotent of $\mathcal{A}$

$$x^e = \sum_{i,j,k} \eta_{ijk} \, (c_i \otimes c_j \otimes c_k) \qquad \text{with } \eta_{ijk} \text{ given by (6.3)}$$

Hereafter we give a method to compute the gametic frequencies of the equilibrium state of a given population.

Let f^0_{ijk} (resp. f^e_{ijk}) be the gametic frequency (resp. the equilibrium frequency) of the gamete $a_i \otimes a_j \otimes a_k$. Then,

the f^0_{ijk}'s and the f^e_{ijk}'s are non-negative and

$\sum_{i,j,k} f^0_{ijk} = 1$ and $\sum_{i,j,k} f^e_{ijk} = 1$

Therefore the initial population can be described by

$$x = \sum_{i,j,k=0}^{1} f_{ijk}^{0} \, (a_i \otimes a_j \otimes a_k) \tag{6.4}$$

and the equililibrium population by

$$x^e = \sum_{i,j,k=0}^{1} f_{ijk}^{e} \, (a_i \otimes a_j \otimes a_k)$$

Proposition 6.1 (Method to compute the f_{ijk}^e's) *The following procedure leads to the equilibrium gametic frequencies f_{ijk}^e :*

Step 1 *From the f_{ijk}^0's compute ξ_{100}, ξ_{010} and ξ_{001} with*

$$\xi_{ijk} = -\sum_{p=i}^{1}\sum_{q=j}^{1}\sum_{r=k}^{1} f_{pqr}^{0} \tag{6.5}$$

Step 2 *For $p, q, r = 0, 1$, compute the η_{pqr} with*

$$\eta_{pqr} = \xi_{100}^{\delta_{p1}} \cdot \xi_{010}^{\delta_{q1}} \cdot \xi_{001}^{\delta_{r1}} \tag{6.6}$$

in setting $0^0 := 1$

Step 3 *We get the f_{ijk}^e's using*

$$f_{ijk}^{e} = (-1)^{i+j+k} \sum_{p=i}^{1}\sum_{q=j}^{1}\sum_{r=k}^{1} \eta_{pqr} \tag{6.7}$$

Proof.

Step 1 Let be x as in (6.4) be the initial population. We have to express x in the canonical basis. Using relation (6.1) and properties of the tensor product, we obtain

$$\begin{aligned}
c_0 \otimes c_0 \otimes c_0 &= a_0 \otimes a_0 \otimes a_0 \\
c_0 \otimes c_0 \otimes c_1 &= a_0 \otimes a_0 \otimes a_0 - a_0 \otimes a_0 \otimes a_1 \\
c_0 \otimes c_1 \otimes c_0 &= a_0 \otimes a_0 \otimes a_0 - a_0 \otimes a_1 \otimes a_0 \\
c_1 \otimes c_0 \otimes c_0 &= a_0 \otimes a_0 \otimes a_0 - a_1 \otimes a_0 \otimes a_0 \\
c_0 \otimes c_1 \otimes c_1 &= a_0 \otimes a_0 \otimes a_0 - a_0 \otimes a_0 \otimes a_1 - a_0 \otimes a_1 \otimes a_0 + a_0 \otimes a_1 \otimes a_1 \\
c_1 \otimes c_0 \otimes c_1 &= a_0 \otimes a_0 \otimes a_0 - a_0 \otimes a_0 \otimes a_1 - a_1 \otimes a_0 \otimes a_0 + a_1 \otimes a_0 \otimes a_1 \\
c_1 \otimes c_1 \otimes c_0 &= a_0 \otimes a_0 \otimes a_0 - a_0 \otimes a_1 \otimes a_0 - a_1 \otimes a_0 \otimes a_0 + a_1 \otimes a_1 \otimes a_0 \\
c_1 \otimes c_1 \otimes c_1 &= a_0 \otimes a_0 \otimes a_0 - a_0 \otimes a_0 \otimes a_1 - a_0 \otimes a_1 \otimes a_0 - a_1 \otimes a_0 \otimes a_0 \\
&+ a_0 \otimes a_1 \otimes a_1 + a_1 \otimes a_0 \otimes a_1 + a_1 \otimes a_1 \otimes a_0 - a_1 \otimes a_1 \otimes a_1
\end{aligned}$$

thus

$$
\begin{array}{rcl}
a_0 \otimes a_0 \otimes a_0 & = & c_1 \otimes c_1 \otimes c_1 \,+\, c_1 \otimes c_1 \otimes c_0 \,+\, c_1 \otimes c_0 \otimes c_1 \,+\, c_0 \otimes c_1 \otimes c_1 \\
 & + & c_1 \otimes c_0 \otimes c_0 \,+\, c_0 \otimes c_1 \otimes c_0 \,+\, c_0 \otimes c_0 \otimes c_1 \,+\, c_0 \otimes c_0 \otimes c_0 \\
a_0 \otimes a_0 \otimes a_1 & = & -\, c_1 \otimes c_1 \otimes c_1 \,-\, c_1 \otimes c_0 \otimes c_1 \,-\, c_0 \otimes c_1 \otimes c_1 \,-\, c_0 \otimes c_0 \otimes c_1 \\
a_0 \otimes a_1 \otimes a_0 & = & -\, c_1 \otimes c_1 \otimes c_1 \,-\, c_1 \otimes c_1 \otimes c_0 \,-\, c_0 \otimes c_1 \otimes c_1 \,-\, c_0 \otimes c_1 \otimes c_0 \\
a_1 \otimes a_0 \otimes a_0 & = & -\, c_1 \otimes c_1 \otimes c_1 \,-\, c_1 \otimes c_1 \otimes c_0 \,-\, c_1 \otimes c_0 \otimes c_1 \,-\, c_1 \otimes c_0 \otimes c_0 \\
a_0 \otimes a_1 \otimes a_1 & = & c_1 \otimes c_1 \otimes c_1 \,+\, c_0 \otimes c_1 \otimes c_1 \\
a_1 \otimes a_0 \otimes a_1 & = & c_1 \otimes c_1 \otimes c_1 \,+\, c_1 \otimes c_0 \otimes c_1 \\
a_1 \otimes a_1 \otimes a_0 & = & c_1 \otimes c_1 \otimes c_1 \,+\, c_1 \otimes c_1 \otimes c_0 \\
a_1 \otimes a_1 \otimes a_1 & = & -\, c_1 \otimes c_1 \otimes c_1
\end{array}
$$

and from which we deduce

$$f^0_{ijk} = (-1)^{i+j+k} \sum_{p=i}^{1} \sum_{q=j}^{1} \sum_{r=k}^{1} \xi_{pqr}$$

Reversing this formula we finally obtain relation (6.5).

Step 2 Relation (6.6) follows directly from (6.2) and (6.3).

Step 3 From Step 1 we have

$$\eta_{pqr} = -\sum_{i=p}^{1} \sum_{j=q}^{1} \sum_{k=r}^{1} f^e_{ijk}$$

and reversing it we finally have relation (6.7). □

Remark. The last thing to study would be to find out how fast the sequence $x = x^{[1]}, x^{[2]}, \ldots$ converges to x^e.

Finally, we give an numerical example computed with the program whose code is given the appendix.
Example

Suppose we have a population with the following gametic frequencies :

$a_0a_0a_0$	0.15
$a_0a_0a_1$	0.18
$a_0a_1a_0$	0.05
$a_1a_0a_0$	0.20
$a_0a_1a_1$	0.27
$a_1a_0a_1$	0.04
$a_1a_1a_0$	0.10
$a_1a_1a_1$	0.01

Then it converges towards the following equilibrium state :

$a_0a_0a_0$	0.211584
$a_0a_0a_1$	0.159616
$a_0a_1a_0$	0.119016
$a_1a_0a_0$	0.089784
$a_0a_1a_1$	0.153216
$a_1a_0a_1$	0.115584
$a_1a_1a_0$	0.086184
$a_1a_1a_1$	0.065016

Remark. We did not investigate numerically how the solution computed depends on the initial frequencies and whether it is stable or not. But, as the problem is quite binary, we could think that the solution computed is numerically stable.

Appendix A

Program code's

Here is the C++ program code's for the equilibrium frequencies computation.

In using this program, we have to be aware of the approximation errors made by the computer.

```
// ***************************************************************************
// *                                                                         *
// * PROJET DE SEMESTRE : ALGEBRES GENETIQUES                                *
// *                                                                         *
// * (dirige par Prof. Kathryn Hess-Bellwald)                                *
// *                                                                         *
// * Soit une population d'individus diploides differant genetiquement      *
// * par 3 loci independants chacun avec 2 alleles. A partir des frequences *
// * gametiques initiales (et sous quelques hypotheses), ce programme       *
// * calcule les frequences gametiques d'equilibre.                         *
// *                                                                         *
// * Gregoire Aubry, le 11 juin 2001.                                        *
// *                                                                         *
// ***************************************************************************


#include <iostream>
#include <vector>
#include <cmath>



// ****************************
// * Definition des variables *
// ****************************

double initial_freq[2][2][2];
double xi[3];
double eta[2][2][2];
```

```
double equ_freq[2][2][2];
double somme;



// ***************************************
// * Definition de la fonction puissance *
// ***************************************

double puissance (double x, double p) {
  if (p == 0) {return 1;}
  if ((p != 0) && (x == 0)) {return 0;}
  if ((p != 0) && (x != 0)) {return pow(x, p);}
}; // puissance



// ***********************************
// * Definition de la fonction signe *
// ***********************************

double signe (double p) {
  if ((p == 0) || (p == 2)) {return 1;}
  else {return -1;}
}; // signe



// ***********************
// * Programme principal *
// ***********************

void main()
{
  cout << endl << endl << endl << endl << endl
       << "*******************************************" << endl
       << "*                                         *" << endl
       << "*  GAMETIC FREQUENCIES COMPUTING PROGRAM  *" << endl
       << "*                                         *" << endl
       << "*******************************************" << endl
       << endl << endl
       << "(Program written by Gregoire Aubry," << endl
       << " Tutored by Prof. Kathryn Hess-Bellwald, June 2001)" << endl
       << endl << endl << endl
       << "Consider a population of diploid individuals which differ " << endl
       << "genetically in 3 independent loci each with 2 alleles "
       << "(denoted " << endl << "by a0 and a1)." << endl
```

```
       << endl
       << "This program compute the gametic frequencies of the " << endl
       << "equilibrium population from the initial gametic frequencies" << endl
       << endl << endl
       << "The main assumptions are the following :" << endl << endl
       << "1) population infinitely large" << endl
       << "2) randomly mating" << endl
       << "3) recombination rate = 0.5" << endl
       << "4) no mutation" << endl
       << "5) no overlapping" << endl;



// ***********************************
// * Initialisation des variables a 0 *
// ***********************************

  for (int i(0); i<=2; ++i) {xi[i] = 0;}

  for (int i(0); i<=1; ++i)
    {
      for(int j(0); j<=1; ++j)
{
  for(int k(0); k<=1; ++k)
    {
      initial_freq[i][j][k] = 0;
      eta[i][j][k] = 0;
      equ_freq[i][j][k] = 0;
    } // for k
} // for j
    } // for i



// ********************************************
// * Entree des frequences gametiques initiales *
// ********************************************

do
  {
    somme = 0;
    cout << endl << endl;
    cout << "Enter the gametic frequencies :" << endl << endl;

for (int i(0); i<=1; ++i)
  {
    for(int j(0); j<=1; ++j)
```

```
      {
for(int k(0); k<=1; ++k)
  {
    do
      {
cout << "a" << i << "a" << j << "a" << k << "    ";
cin >> initial_freq[i][j][k] ;

if (initial_freq[i][j][k] < 0 || initial_freq[i][j][k] > 1 )
  {
    cout << endl <<"You have to enter a number between 0 and "
 << "1 !" << endl;
  } //if
cout << endl;

      } while((initial_freq[i][j][k] < 0) ||
      (initial_freq[i][j][k] > 1));

    somme = somme + initial_freq[i][j][k];

  } // for k
      } // for j
  } // for i


 if ((somme < 0.999999) || (somme > 1.000001))
   {
     cout << endl <<"The frequencies have to add to 1 !" << endl;
   } // if

  } while ((somme < 0.999999) || (somme > 1.000001));



// ******************
// * Calcul des xi's *
// ******************

for (int j(0); j<=1; ++j)
  {
    for (int k(0); k<=1; ++k)
      {
    xi[0] = xi[0] - initial_freq[1][j][k];
      } // for k
  } // for j
```

```
for (int j(0); j<=1; ++j)
  {
    for (int k(0); k<=1; ++k)
      {
    xi[1] = xi[1] - initial_freq[j][1][k];
      } // for k
  } // for j


for (int j(0); j<=1; ++j)
  {
    for (int k(0); k<=1; ++k)
      {
    xi[2] = xi[2] - initial_freq[j][k][1];
      } // for k
  } // for j



// *******************
// * Calcul des eta's *
// *******************

 for (int p(0); p<=1; ++p)
   {
     for(int q(0); q<=1; ++q)
       {
 for(int r(0); r<=1; ++r)
   {
     eta[p][q][r]= puissance(xi[0],p) * puissance(xi[1],q)
                   * puissance(xi[2],r);
   } // for k
       } // for j
   } // for i



// **************************************************
// * Calcul et affichage des frequences d'equilibre *
// **************************************************

 cout << endl << endl
      << "The gametic frequencies of the equilibrium population are the "
      << "following :"
      << endl << endl << endl;

 for (int i(0); i<=1; ++i)
```

```
    {
      for(int j(0); j<=1; ++j)
        {
 for(int k(0); k<=1; ++k)
    {
      for (int p(i); p<=1; ++p)
        {
 for(int q(j); q<=1; ++q)
    {
      for(int r(k); r<=1; ++r)
        {
 equ_freq[i][j][k] = equ_freq[i][j][k] + eta[p][q][r];
        } // for r
    } // for q
        } // for p

      equ_freq[i][j][k] = signe(i+j+k) * equ_freq[i][j][k];
      if (equ_freq[i][j][k] == -0) {equ_freq[i][j][k] = 0;};
      cout << "a" << i << "a" << j << "a" << k << "    "
  << equ_freq[i][j][k] << endl << endl;

    } // for k
        } // for j
    } // for i


 cout << endl << endl;
} // main
```

Bibliography

[1] Atiyah M. F., MacDonald I. G. (1969), *Introduction to Commutative Algebra*, Addison-Wesley Series in Mathematics, Addison-Wesley Publishing Company, London.

[2] Gillespie J. H. (1998), *Population Genetics : A Concise Guide*, The Johns Hopkins University Press, Baltimore.

[3] Wörz-Busekros A. (1980), *Algebras in Genetics*, Lecture Notes in Biomathematics, Springer-Verlag Heidelberg Berlin.