J. Math. Kyoto Univ. (JMKYAZ)
21-1 (1981) 19–45

Formulas for diffusion approximations of some gene frequency models

By

Norio SHIMAKURA

(Received Sept. 12, 1979)

Since the work of W. Feller [4], the diffusion approximations of gene frequency models in population genetics have been treated by many authors as a typical example of stochastic processes.

This problem is also of importance in the theory of partial differential equations. The interest is that the Kolmogorov equations describing the diffusion processes are of degenerated parabolic type. Above all, the domain (state space) is a simplex, hence, it has sides and corners. And the elliptic parts (generators of semi-groups) of the equations are degenerated on the boundary in a very natural way. The present author had been attracted by this particularity of equations, and he published a treatment in L^2-framework ([10]).

But the L^2-theory is not sufficient to describe the diffusion processes, and the movements of the processes on the boundary tell us many things which are characteristic to this type of equations.

Therefore, in this memoir, we treat these equations on the space of continuous functions, and we give the concrete formulas for the semi-groups of diffusion processes. Naturally, for the sake that the formulas be completely written down by means of known functions, we must restrict our considerations to a special class of gene frequency models. That is, we neglect the selection force, and we assume that the mutating pressure is of simple type from the point of view of calculus. If the selection force is taken into account or if the mutating pressure is of general type, the question of giving the formulas is still open (See [1], [3], [7] and [8] for more general models, and see § 3 below for our assumptions on models).

We treat here d types models of the restricted class. The formulas for 2 types models have been given by J. Crow and M. Kimura [1]. And for d types models, we have a work of E. Fackerel and R. Littler [11].

To prepare this memoir, we owe the main ideas to the book of S. Karlin [6] and to the thesis of S. Ethier [3]. The present author is very grateful to Dr. K. Sato and to Dr. T. Maruyama for their suggestions and advices.

§ 1. Multi-type gene frequency model of Wright-Fisher

In this paragraph, we give a brief review of the multi-type gene frequency model of Wright-Fisher. Our reference is the Chapter 13 of [6] (See also [7]).

Let us consider a population of an organism. The quantity controlling (or partially controlling) a character of this organism is called a *gene*. The position at which a gene occurs on a chromosome is called its *locus*. The various alternatives which may occur at a particular locus are called *alleles*.

We are looking at only one locus. Let $A_1,\dots, A_d$ be the alleles occurring there ($2\leqslant d< +\infty$). An individual is called of *type* p if it has the allele A_p ($1\leqslant p\leqslant d$).

Let us assume at first that

(A) A parent of type p produces only the same type of offspring individuals, of which the number is a stochastic variable distributed according to Poisson distribution with parameter ξ_p ($\xi_p>0$, $1\leqslant p\leqslant d$).

Here, $\xi_1,\dots, \xi_d$ express the fitnesses of type 1,..., of type d respectively. Hence, the larger is ξ_p relatively to other ξ_q's, the more advantageous is the type p to others. If in particular $\xi_1=\cdots=\xi_d$, there is neither selective advantage nor disadvantage among $A_1,\dots, A_d$.

The second assumption is that

(B) After the stage (A), each offspring individual is effected by *mutating pressure*. That is to say, an offspring of type p remains of type p with probability m_{pp} ($1\leqslant p\leqslant d$), and it mutates and becomes of type q with probability m_{pq} ($1\leqslant p, q\leqslant d$, $p\neq q$). Naturally, we have

$$m_{pq}\geqslant 0 \quad \text{for} \quad 1\leqslant p, q\leqslant d \quad \text{and} \quad \sum_{q=1}^{d} m_{pq}=1 \quad \text{for} \quad 1\leqslant p\leqslant d. \tag{1.1}$$

If $m_{pp}=1$ in particular, the mutating pressure does not effect to type p.

The third assumption is that

(C) Each individual behaves independently of others.

That is to say, an individual of the k-th generation produces its offsprings of the $(k+1)$-th generation according to (A) independently of others, each of offsprings is effected by mutating pressure according to (B) independently of others and it produces the offsprings of the $(k+2)$-th generation according to (A) independently of others, and so on.

Hence, the probability generating function for progeny distribution, produced by a parent of type p and afterwards effected by mutating pressure, is given by

$$f_p(s_1,\dots, s_d)=\exp\{\xi_p(\sum_{q=1}^{d} m_{pq}s_q-1)\} \quad \text{for} \quad 1\leqslant p\leqslant d. \tag{1.2}$$

The fourth assumption is that

(D) During each period of generation, there are individuals *immigrating* from outside population into our population. And, the number of them is distributed according to the probability generating function

$$h(s_1,\dots, s_d)=\exp\{\sum_{p=1}^{d} c_p(s_p-1)\}, \tag{1.3}$$

where $c_1,\dots, c_d$ are non-negative constants.

There is no immigrant of type p if $c_p=0$.

The fifth and final assumption is that

(E) The *size* N of our population is conditioned to be fixed over generations.

Let us denote by $\alpha_p(k)$ the number of individuals of type p of the k-th generation in our population. The ratio $\alpha_p(k)/N$ is called the *frequency* of the type p at the k-th generation. Let us define the vector $\tilde{\alpha}^{(N)}(k)$ by

$$\tilde{\alpha}^{(N)}(k)=(\alpha_1(k),\dots, \alpha_d(k)). \tag{1.4}$$

Then, $\tilde{\alpha}^{(N)}(k)/N$ is called the *gene frequency* at the k-th generation. The range of $\tilde{\alpha}^{(N)}(k)$ is the set

$$\Omega^{(N)}=\{\tilde{\alpha}=(\alpha_1,\dots, \alpha_d);\ \alpha_p\in \boldsymbol{N} \quad \text{for} \quad 1\leqslant p\leqslant d \quad \text{and} \quad \sum_{p=1}^{d} \alpha_p=N\}, \tag{1.5}$$

where $\boldsymbol{N}$ is the set of all the non-negative integers.

We have obtained therefore a Markov chain $\{\tilde{\alpha}^{(N)}(k)\}_{k=0}^{\infty}$ induced by the direct product branching process. The state space is $\Omega^{(N)}$ and the set of times (generations) is $\boldsymbol{N}$. This Markov chain, constructed under the assumptions (A)$\sim$(E), is called the *multi-type gene frequency model of Wright-Fisher*, or abbreviatedly in this memoir, *discrete model*.

Let $P^{(N)}=(P^{(N)}_{\tilde{\alpha}\tilde{\beta}})$ be the matrix of the one-step transition probability, that is,

$$P^{(N)}_{\tilde{\alpha}\tilde{\beta}}=\text{Prob.}\,[\tilde{\alpha}^{(N)}(k+1)=\tilde{\beta}\mid \tilde{\alpha}^{(N)}(k)=\tilde{\alpha}] \quad \text{for} \quad \tilde{\alpha} \quad \text{and} \quad \tilde{\beta}\in\Omega^{(N)}. \tag{1.6}$$

This $P^{(N)}_{\tilde{\alpha}\tilde{\beta}}$ is independent of k by virtue of the hypothesis (E), and is equal to the coefficient of $s_1^{\beta_1}\cdots s_d^{\beta_d}$ in the power series expansion of $h(s_1,\dots, s_d)\prod_{p=1}^{d} f_p(s_1,\dots, s_d)^{\alpha_p}$ divided by the coefficient of t^N in $h(t,\dots, t)\prod_{p=1}^{d} f_p(t,\dots, t)^{\alpha_p}$. Therefore, by means of (1.2) and (1.3), we have

$$P^{(N)}_{\tilde{\alpha}\tilde{\beta}}=\frac{N!}{\beta_1!\dots\beta_d!}\zeta_1(\tilde{\alpha})^{\beta_1}\cdots\zeta_d(\tilde{\alpha})^{\beta_d} \quad \text{for} \quad \tilde{\alpha} \quad \text{and} \quad \tilde{\beta}\in\Omega^{(N)}, \tag{1.7}$$

where the vector $\zeta(\tilde{\alpha})=(\zeta_1(\tilde{\alpha}),\dots, \zeta_d(\tilde{\alpha}))$ is defined by

$$\zeta_p(\tilde{\alpha})=(c_p+\sum_{q=1}^{d} \xi_q\alpha_q m_{qp})/\sum_{r=1}^{d}(c_r+\xi_r\alpha_r) \quad \text{for} \quad 1\leqslant p\leqslant d. \tag{1.8}$$

$\zeta(\tilde{\alpha})$ (of variable $\tilde{\alpha}$) contains parameters $\xi_1,\dots, \xi_d$; $m_{11}, m_{12},\dots, m_{dd}$; $c_1,\dots, c_d$ and N representing the selection force, the mutating pressure, the migration and the size of population respectively. (1.7) is equivalent to the following identity which will be repeatedly used in the sequel:

$$\sum_{\tilde{\beta}} P^{(N)}_{\tilde{\alpha}\tilde{\beta}} s_1^{\beta_1}\cdots s_d^{\beta_d}=(\sum_{p=1}^{d} \zeta_p(\tilde{\alpha})s_p)^N, \tag{1.9}$$

where the summation with respect to $\tilde{\beta}$ is extended over $\Omega^{(N)}$.

The Markov chain $\{\tilde{\alpha}^{(N)}(k)\}_{k=0}^{\infty}$ is characterized by the matrix $P^{(N)}$. Therefore in the sequel, $P^{(N)}$ itself will be called the multi-type gene frequency model or discrete model.

In this memoir, except in § 2, we replace the hypotheses (A) and (B) by more restrictive ones as follows:

(A′) There is no selection force and

$$\xi_p = 1 \quad \text{for} \quad 1 \leqslant p \leqslant d. \tag{1.10}$$

(B′) The probability of mutation from type p to type q ($p \neq q$) depends only on q for $1 \leqslant p, q \leqslant d$. That is, we assume that

$$\begin{cases} m_{pq} = m_q + \delta_{pq}(1 - \sum_{r=1}^{d} m_r) & \text{for} \quad 1 \leqslant p,\ q \leqslant d, \\ \text{with} \quad m_q \geqslant 0 \quad \text{and} \quad 1 + m_q \geqq \sum_{r=1}^{d} m_r & \text{for} \quad 1 \leqslant q \leqslant d. \end{cases} \tag{1.11}$$

Under (A′), (B′), (C), (D) and (E), $\zeta(\tilde{\alpha})$ defined by (1.8) is written as

$$\zeta_p(\tilde{\alpha}) = \{c_p + Nm_p + \alpha_p(1 - \sum_{r=1}^{d} m_r)\}/(N + \sum_{q=1}^{d} c_q) \quad \text{for} \quad 1 \leqslant p \leqslant d. \tag{1.12}$$

§ 2. Preliminaries for diffusion approximations

Let $\boldsymbol{N}$ be the set of all non-negative integers. For any $\gamma = (\gamma_1, \dots, \gamma_m) \in \boldsymbol{N}^m$, we denote $\sum_{p=1}^{m} \gamma_p = |\gamma|$. We will work mainly with $\boldsymbol{N}^d$ or $\boldsymbol{N}^n$, where the character n is reserved to denote

$$n = d - 1 \tag{2.1}$$

throughout this memoir. The elements of $\boldsymbol{N}^d$ are denoted by $\tilde{\alpha}, \tilde{\beta}, \dots$, while the elements of $\boldsymbol{N}^n$ by $\alpha, \beta, \dots$. The state space $\Omega^{(N)}$ defined by (1.5) is the set $\{\tilde{\alpha} \in \boldsymbol{N}^d; |\tilde{\alpha}| = N\}$. This can be identified with any one of the following three sets

$$\begin{gathered} \{\alpha \in \boldsymbol{N}^n;\ |\alpha| \leqslant N\},\ \{\tilde{\alpha}/N \in \boldsymbol{R}^d;\ \tilde{\alpha} \in \boldsymbol{N}^d \text{ and } |\tilde{\alpha}| = N\}, \\ \text{and} \quad \{\alpha/N \in \boldsymbol{R}^n;\ \alpha \in \boldsymbol{N}^n \text{ and } |\alpha| \leqslant N\}. \end{gathered} \tag{2.2}$$

There are exactly $\binom{n+N}{N}$ points in $\Omega^{(N)}$.

Therefore, the Markov chain $\{\tilde{\alpha}^{(N)}(k)\}_{k=0}^{\infty}$ can also be written as $\{\alpha^{(N)}(k)\}_{k=0}^{\infty}$ deleting the d-th component.

On the other hand, let us consider a closed n-simplex

$$\bar{\Omega} = \{\tilde{x} = (x_1, \dots, x_d) \in \boldsymbol{R}^d;\ x_p \geqslant 0 \text{ for } 1 \leqslant p \leqslant d \text{ and } \sum_{p=1}^{d} x_p = 1\} \tag{2.3}$$

which can be identified with

$$\bar{\Omega}=\{x=(x_1,\dots,x_n)\in \boldsymbol{R}^n;\ x_p\geqslant 0 \text{ for } 1\leqslant p\leqslant n \text{ and } \sum_{p=1}^{n} x_p\leqslant 1\}. \tag{2.3'}$$

Then, $\Omega^{(N)}$ is regarded as a subset of $\bar{\Omega}$ consisting of all the lattice points with mesh $1/N$. The set of interior points of $\bar{\Omega}$ (the open n-simplex) is denoted by Ω.

Let ϕ be a $\binom{n+N}{N}$-dimensional column vector with the α-th component $\phi(\alpha)$, $\alpha\in\Omega^{(N)}$. Then, ϕ can be identified with a function defined only on the lattice points of $\bar{\Omega}$. Reciprocally, given a function $u(x)$ defined on $\bar{\Omega}$, we obtain a vector $[u]$ by restricting it on lattice points:

$$[u](\alpha)=u(\alpha/N) \quad \text{for} \quad \alpha\in\Omega^{(N)}. \tag{2.4}$$

The matrix $P^{(N)}=(P^{(N)}_{\alpha\beta})$ (identifying $(\tilde{\alpha},\tilde{\beta})$ with (α,β)) of the one-step transition probability of the discrete model is then regarded as a linear transformation acting on column vectors:

$$(P^{(N)}\phi)(\alpha)=\sum_{\beta} P^{(N)}_{\alpha\beta}\phi(\beta) \quad \text{for} \quad \alpha\in\Omega^{(N)}, \tag{2.5}$$

where the summation with respect to β is extended over $\Omega^{(N)}$. Therefore, we can define a semi-group $T^{(N)}=\{T^{(N)}(k)\}_{k=0}^{\infty}$ by non-negative powers of $P^{(N)}$:

$$\begin{aligned} T^{(N)}(k)\phi(\alpha) &= (P^{(N)})^k\phi(\alpha)=E_\alpha[\phi(\alpha^{(N)}(k))] \\ &= \sum_{\beta} \text{Prob.}\,[\alpha^{(N)}(k)=\beta \mid \alpha^{(N)}(0)=\alpha]\phi(\beta) \quad \text{for} \quad k\in\boldsymbol{N}. \end{aligned} \tag{2.6}$$

Hence, we can identify the discrete model $\{\alpha^{(N)}(k)\}_{k=0}^{\infty}$ also with the semi-group $T^{(N)}$ of matrices.

On the other hand, let A be a differential operator of the second order in n variables having an expression

$$Au(x)=\frac{1}{2}\sum_{p,q=1}^{n} a_{pq}(x)\frac{\partial^2 u}{\partial x_p\partial x_q}+\sum_{p=1}^{n} b_p(x)\frac{\partial u}{\partial x_p}, \tag{2.7}$$

with coefficients continuously differentiable on $\bar{\Omega}$. Suppose that A generates a semi-group of transformations $\mathscr{T}=\{\mathscr{T}(t)\}_{t\geqslant 0}$ on the function space $C^0(\bar{\Omega})$ under a certain boundary condition. Let us denote $\mathscr{T}^{(N)}=\{\mathscr{T}(t/N)\}_{t\geqslant 0}$. $\mathscr{T}$ will also be called a *diffusion model*, although the corresponding diffusion process $\xi(t,\omega)$ is not yet defined (See § 6).

Definition 2.1. The semi-group $\mathscr{T}^{(N)}$ is called a *diffusion approximation* of the discrete model $T^{(N)}$ as N is large, if $\mathscr{T}$ is generated by some A having an expression (2.7) and if the following condition is satisfied: For any $f\in C^0(\bar{\Omega})$, any positive number ε and for any positive integer K, there exists a number N_0 such that, if $N\geqslant N_0$ and if $0\leqslant k\leqslant KN$, we have

$$|T^{(N)}(k)[f](\alpha)-[\mathscr{T}(k/N)f](\alpha)|<\varepsilon \quad \text{on} \quad \Omega^{(N)}. \tag{2.8}$$

At first, for the existence of a diffusion approximation, we should notice that the parameters contained in the discrete model (See (1.8)) may not be arbitrary but they must have some properties of uniformity as N is large. As a hypothesis of uniformity, we put the following (See [3] and [8]):

$$\sigma_p = N(\xi_p - 1) \text{ are real and independent of } N,\ 1 \leqslant p \leqslant d; \tag{2.9}$$

$$\mu_{pq} = N(m_{pq} - \delta_{pq}) \text{ are independent of } N,\ 1 \leqslant p,\ q \leqslant d; \tag{2.10}$$

$$\mu_{pq} \geqslant 0 \text{ if } p \neq q \text{ and } \sum_{q=1}^{d} \mu_{pq} = 0,\ 1 \leqslant p \leqslant d; \tag{2.11}$$

$$c_p \text{ are non-negative and independent of } N,\ 1 \leqslant p \leqslant d. \tag{2.12}$$

That is, the selective advantages (or disadvantages) are at most of $O(1/N)$, the mutating pressure is also of $O(1/N)$ and the effect of migration is independent of N (If, for example, m_1 is positive and independent of N and if other m_p's are all 0 in the notation of (1.11), the mutation from other types to type 1 is too large that the frequency of the type 1 tends very rapidly to 1. In such a case, no diffusion approximation is correct).

Next, we will calculate the coefficients $a_{pq}(x)$ and $b_p(x)$ to fix the expression (2.7) of A and give a preliminary estimate related to A. We put a priori

$$a_{pq}(x) = \delta_{pq} x_p - x_p x_q \quad \text{for} \quad 1 \leqslant p,\ q \leqslant d; \quad \text{and} \tag{2.13}$$

$$b_p(x) = c_p - x_p \sum_{q=1}^{d} c_q + \sum_{q=1}^{d} a_{pq}(x)\sigma_q + \sum_{q=1}^{d} \mu_{qp} x_q \quad \text{for} \quad 1 \leqslant p \leqslant d. \tag{2.14}$$

A simple computation shows that

$$\zeta_p(\alpha) = \frac{1}{N}\left\{\alpha_p + b_p\left(\frac{\alpha}{N}\right)\right\} + O(N^{-2}) \quad \text{for} \quad 1 \leqslant p \leqslant d \tag{2.15}$$

holds uniformly on $\Omega^{(N)}$ as N is large, where $\zeta(\alpha)$ is defined by (1.8) (α is identified with $\tilde{\alpha}$). We are going to show that the generator A is of the form (2.7) with the coefficients defined by (2.13) and by (2.14). Let us put, for $1 \leqslant p,\ q,\ r \leqslant d$,

$$\begin{cases} b_p^{(N)}(\alpha) = \sum_\beta (\beta_p - \alpha_p) P_{\alpha\beta}^{(N)}, \\ a_{pq}^{(N)}(\alpha) = N^{-1} \sum_\beta (\beta_p - \alpha_p)(\beta_q - \alpha_q) P_{\alpha\beta}^{(N)} \\ c_{pqr}^{(N)}(\alpha) = N^{-2} \sum_\beta (\beta_p - \alpha_p)(\beta_q - \alpha_q)(\beta_r - \alpha_r) P_{\alpha\beta}^{(N)}, \\ e_p^{(N)}(\alpha) = N^{-3} \sum_\beta (\beta_p - \alpha_p)^4 P_{\alpha\beta}^{(N)}, \end{cases} \tag{2.16}$$

where the summations with respect to β are extended over $\Omega^{(N)}$. Then we have the

Lemma 2.2. *As N is large, the following estimates hold uniformly on $\Omega^{(N)}$:*

$$\begin{cases} b_p^{(N)}(\alpha) = b_p\left(\frac{\alpha}{N}\right) + O(1/N), \\ a_{pq}^{(N)}(\alpha) = a_{pq}\left(\frac{\alpha}{N}\right) + O(1/N), \\ c_{pqr}^{(N)}(\alpha) = O(1/N) \text{ and } e_p^{(N)}(\alpha) = O(1/N). \end{cases} \tag{2.17}$$

Proof. We differentiate the both sides of (1.9) successively in s_p's and put $s_1=\cdots=s_d=1$. Then we have

$$\sum_\beta P^{(N)}_{\alpha\beta}=1,\ \sum_\beta \beta_p P^{(N)}_{\alpha\beta}=N\zeta_p,\ \sum_\beta \beta_p\beta_q P^{(N)}_{\alpha\beta}=N^2\zeta_p\zeta_q+Na_{pq}(\zeta),$$

$$\sum_\beta \beta_p\beta_q\beta_r P^{(N)}_{\alpha\beta}=N^3\zeta_p\zeta_q\zeta_r+N^2\{a_{pq}(\zeta)\zeta_r+a_{qr}(\zeta)\zeta_p+a_{rp}(\zeta)\zeta_q\}+O(N),$$

$$\text{and}\quad \sum_\beta \beta_p^4 P^{(N)}_{\alpha\beta}=N^4\zeta_p^4+6N^3a_{pp}(\zeta)\zeta_p^2+O(N^2),$$

where $\zeta=\zeta(\alpha)$. It is quite elementary to derive (2.17) from these equalities with the aid of (2.15). Q.E.D.

Lemma 2.3. (S. Ethier) *Let A be the operator defined by* (2.7), (2.13) *and by* (2.14). *If $u(x)\in C^4(\bar\Omega)$, we have*

$$N(P^{(N)}[u]-[u])(\alpha)=[Au](\alpha)+O(1/N) \tag{2.18}$$

uniformly on $\Omega^{(N)}$ as N is large (*See* (1.13), *p.* 27 *of* [3]).

Proof. Maclaurin expansion of $u(y)$ at x yields

$$u(y)-u(x)=\sum_{p=1}^{n} z_p\frac{\partial u}{\partial x_p}(x)+\frac{1}{2}\sum_{pq,=1}^{n} z_pz_q\frac{\partial^2 u}{\partial x_p\partial x_q}(x)$$
$$+\frac{1}{6}\sum_{p,q,r=1}^{n} z_pz_qz_r\frac{\partial^3 u}{\partial x_p\partial x_q\partial x_r}(x)+R(y,x),$$

where $z=y-x$ and $R(y,x)=O(|z|^4)$. We put $x=\alpha/N$, $y=\beta/N$, multiply $NP^{(N)}_{\alpha\beta}$ to the both sides and sum up with respect to β over $\Omega^{(N)}$. Then, the left hand side is $N(P^{(N)}[u]-[u])(\alpha)$, while the right hand side is $[Au](\alpha)+O(1/N)$ by virtue of Lemma 2.2. This proves (2.18). Q.E.D.

Faithfully to the symmetry with respect to the indices $1,\ldots,d$ in the original problem, we can consider an operator of d variables $\tilde x=(x_1,\ldots,x_d)$ as follows:

$$\tilde A v(\tilde x)=\frac{1}{2}\sum_{p,q=1}^{d} a_{pq}(\tilde x)\frac{\partial^2 v}{\partial x_p\partial x_q}+\sum_{p=1}^{d} b_p(\tilde x)\frac{\partial v}{\partial x_p}, \tag{2.19}$$

with coefficients also given by (2.13) and (2.14). Then we have

Lemma 2.4. (K. Sato) *Let $v(\tilde x)$ be a function of class C^2 in a neighborhood of $\bar\Omega$ in $\boldsymbol{R}^d$. Then we have $(\tilde A v)|_{\bar\Omega}=A(v|_{\bar\Omega})$ on $\bar\Omega$ (See [9]).*

Proof. This will be easily checked, if we use the relations

$$\sum_{q=1}^{d} a_{pq}(\tilde x)=0 \quad\text{for}\quad 1\leqslant p\leqslant d \quad\text{and}\quad \sum_{p=1}^{d} b_p(\tilde x)=0 \quad\text{on}\quad \bar\Omega. \tag{2.21}$$

(A more restrictive but precise version of this lemma will be stated in §4. See Lemma 4.3). Q.E.D.

§ 3. Eigenvectors of $P^{(N)}$ and eigenpolynomials of A

From now on, we discuss and look for the diffusion approximation $\mathscr{T}^{(N)}$ of the discrete model $T^{(N)}$ which satisfies (A′), (B′) in § 1 and (2.9)~(2.12) at the same time. That is, putting $Nm_p=\mu_p$ in (1.12), we assume that $P^{(N)}$ be defined by (1.7) with

$$\begin{cases} \zeta_p(\alpha)=\{b_p+\alpha_p(1-N^{-1}\sum_{q=1}^{d}\mu_q)\}/(N+\sum_{r=1}^{d}c_r) \\ \text{where} \quad c_p\geqslant 0,\ \mu_p\geqslant 0 \quad \text{and} \quad b_p=c_p+\mu_p \quad \text{for} \quad 1\leqslant p\leqslant d, \end{cases} \tag{3.1}$$

c_p and μ_p (hence b_p also) being independent of N.

In this case, $b_p(x)$'s in (2.14) become

$$b_p(x)=b_p-x_p\sum_{q=1}^{d}b_q \quad \text{for} \quad 1\leqslant p\leqslant d. \tag{3.2}$$

Therefore, the operator A has the expression

$$Au(x)=\frac{1}{2}\sum_{p,q=1}^{n}(\delta_{pq}x_p-x_px_q)\frac{\partial^2 u}{\partial x_p\partial x_q}+\sum_{p=1}^{n}(b_p-x_p\sum_{q=1}^{d}b_q)\frac{\partial u}{\partial x_p}, \tag{3.3}$$

or equivalently, the expression in d variables

$$\tilde{A}v(\tilde{x})=\frac{1}{2}\sum_{p,q=1}^{d}(\delta_{pq}x_p-x_px_q)\frac{\partial^2 v}{\partial x_p\partial x_q}+\sum_{p,q=1}^{d}(\delta_{pq}b_p-x_pb_q)\frac{\partial v}{\partial x_p}. \tag{3.4}$$

In this paragraph, we calculate eigenvalues and column eigenvectors of $P^{(N)}$ according to the idea of [6] (See § 7 in Chap. 13). And by passage to the limit as $N\to+\infty$, we derive from them eigenvalues and eigenpolynomials of A.

Take a multi-index $\gamma\in\boldsymbol{N}^n$ with $|\gamma|\leqslant N$. Differentiating the both sides of (1.9) γ times in s and putting $s_1=\cdots=s_d=1$,

$$\sum_{\beta}\frac{\beta!}{(\beta-\gamma)!}P_{\alpha\beta}^{(N)}=\frac{N!}{(N-|\gamma|)!}\zeta(\alpha)^{\gamma}.$$

Therefore, if we put

$$\begin{cases} \psi_\gamma^{(N)}(x)=N^{-|\gamma|}\prod_{p=1}^{n}\dfrac{\Gamma(Nx_p+1)}{\Gamma(Nx_p+1-\gamma_p)} \quad \text{and} \\ \omega_\gamma^{(N)}(x)=\prod_{p=1}^{n}\{x_p+b_p(N-\sum_{q=1}^{d}\mu_q)^{-1}\}^{\gamma_p}, \quad \text{for} \quad \gamma\in\boldsymbol{N}^n, \end{cases} \tag{3.5}$$

the above identity is rewritten as

$$P^{(N)}[\psi_\gamma^{(N)}](\alpha)=\mu_{|\gamma|}^{(N)}[\omega_\gamma^{(N)}](\alpha) \quad \text{for} \quad \gamma\in\boldsymbol{N}^n \quad \text{with} \quad |\gamma|\leqslant N, \tag{3.6}$$

where the numbers $\{\mu_m^{(N)}\}_{m=0}^{N}$ are defined by

$$\mu_m^{(N)}=\frac{N^{-m}N!}{(N-m)!}\{(N-\sum_{q=1}^{d}\mu_q)/(N+\sum_{r=1}^{d}c_r)\}^m \quad \text{for} \quad 0\leqslant m\leqslant N. \tag{3.7}$$

(3.6) suggests that $\{\mu_m^{(N)}\}_{m=0}^N$ are eigenvalues of $P^{(N)}$ because $\psi_\gamma^{(N)}$ and $\omega_\gamma^{(N)}$ are monics in n variables with the common highest degree term x^γ. We summarize the properties of $\mu_m^{(N)}$'s:

Lemma 3.1. (i) $$1=\mu_0^{(N)}\geqslant\mu_1^{(N)}>\mu_2^{(N)}>\cdots>\mu_N^{(N)}>0, \tag{3.8}$$

where $\mu_1^{(N)}=1$ *if and only if* $c_p=\mu_p=b_p=0$ *for* $1\leqslant p\leqslant d$;

(ii) $$\begin{cases} \mu_m^{(N)}=1-(N+\sum_{q=1}^d c_q)^{-1}(\lambda_m-N^{-1}\rho_m^{(N)}), \\ \textit{with}\quad 0\leqslant\rho_m^{(N)}\leqslant m(m-1)(m+\kappa+1)^2/8 \quad \textit{for} \quad 0\leqslant m\leqslant N; \end{cases} \tag{3.9}$$

(iii) λ_m*'s are independent of* N *and defined by*

$$\begin{cases} \lambda_m=m(m+\kappa)/2 \quad \textit{for} \quad m\in\boldsymbol{N}, \quad \textit{where} & (3.10) \\ \kappa+1=2\sum_{p=1}^d b_p. & (3.11) \end{cases}$$

$\kappa+1$ *is non-negative. It is* 0 *if and only if* $c_p=\mu_p=b_p=0$ *for* $1\leqslant p\leqslant d$.

The proof is elementary (We show (3.9) by induction on m).

Proposition 3.2. (S. Karlin) (iv) *For each* $\gamma\in\boldsymbol{N}^n$ *with* $|\gamma|\leqslant N$, *we can find a monic* $\phi_\gamma^{(N)}(x)$ *with the highest degree term* x^γ *such that* $[\phi_\gamma^{(N)}]$ *is a column eigenvector of* $P^{(N)}$ *belonging to the eigenvalue* $\mu_{|\gamma|}^{(N)}$, *that is,*

$$P^{(N)}[\phi_\gamma^{(N)}]=\mu_{|\gamma|}^{(N)}[\phi_\gamma^{(N)}]; \tag{3.12}$$

(v) *If* $0\leqslant m\leqslant N$, $\mu_m^{(N)}$ *is an eigenvalue of* $P^{(N)}$ *of multiplicity* $\binom{m+n-1}{m}$ (1 *is of multiplicity* d *if* $\kappa+1=0$). *And a basis of column eigenvectors belonging to it is* $\{[\phi_\gamma^{(N)}];\ |\gamma|=m\}$ (*See* §7 *in Chap.* 13 *of* [6]).

Proof. Suppose at first that $\kappa+1>0$. Let us expand $\psi_\gamma^{(N)}(x)$ by means of $\omega_\delta^{(N)}(x)$'s and $\phi_\gamma^{(N)}(x)$ by means of $\psi_\delta^{(N)}(x)$'s:

$$\psi_\gamma^{(N)}(x)=\omega_\gamma^{(N)}(x)+\sum_{\delta<\gamma}c_{\gamma,\delta}\omega_\delta^{(N)}(x), \tag{3.13}$$

$$\phi_\gamma^{(N)}(x)=\psi_\gamma^{(N)}(x)+\sum_{\delta<\gamma}a_{\gamma,\delta}\psi_\delta^{(N)}(x), \tag{3.14}$$

where $\delta<\gamma$ for two multi-indices δ and γ means that $\delta_j\leqslant\gamma_j$ for $1\leqslant j\leqslant n$ and that $|\delta|<|\gamma|$. Substituting (3.13) and (3.14) into (3.12) and using (3.6), we obtain a system of equations

$$\sum_{\delta<\gamma}a_{\gamma,\delta}(\mu_{|\delta|}^{(N)}-\mu_m^{(N)})W_\delta=\mu_m^{(N)}\sum_{\delta<\gamma}(c_{\gamma,\delta}W_\delta+a_{\gamma,\delta}\sum_{\eta<\delta}c_{\delta,\eta}W_\eta), \tag{3.15}$$

where $m=|\gamma|$ and $W_\delta=[\omega_\delta^{(N)}]$. Equating the coefficients of W_δ's in the both sides, we obtain a new system of equations. The latter can be solved with respect to $a_{\gamma,\delta}$'s in a unique way, because $\mu_m^{(N)}$'s are distinct. This proves (iv) if $\kappa+1>0$.

Suppose now that $\kappa+1=0$. The above reasoning remains valid if $m\geqslant 2$. And

it is easy to see directly that 1 is an eigenvalue of multiplicity d and that $\{[x_p]\}_{p=1}^{d}$ is a basis of column eigenvectors belonging to 1. (iv) is proved also for $\kappa+1=0$.
(v) is trivial. Q.E.D.

Proposition 3.3. (vi) *For any fixed* $\gamma\in\boldsymbol{N}^n$, $\phi_\gamma^{(N)}(x)$ *converges as* $N\to+\infty$ *to a monic* $\phi_\gamma(x)$ *with the highest degree term* x^γ;
(vii) $\phi_\gamma(x)$ *is an eigenpolynomial of* A *(See* (3.3)) *belonging to the eigenvalue* $-\lambda_{|\gamma|}$, *that is,*

$$A\phi_\gamma(x)=-\lambda_{|\gamma|}\phi_\gamma(x)\quad \textit{for any}\quad \gamma\in\boldsymbol{N}^n; \tag{3.16}$$

(viii) *For each* $m\in\boldsymbol{N}$, $-\lambda_m$ *is an eigenvalue of* A *of multiplicity* $\binom{m+n-1}{m}$ (0 *is of multiplicity* d *if* $\kappa+1=0$). *And,* $\{\phi_\gamma(x);\ |\gamma|=m\}$ *is a basis of eigenpolynomials belonging to it.* $\{\phi_\gamma(x);\ \gamma\in\boldsymbol{N}^n\}$ *spans all the polynomials of* n *variables.*

Proof. We denote by $\boldsymbol{A}_\nu$ the vector space of all convergent power series in $\nu=1/N$, and by $\boldsymbol{A}_\nu^0$ the subspace of $\boldsymbol{A}_\nu$ consisting of all elements vanishing at $\nu=0$. To prove (vi), $\psi_\gamma^{(N)}(x)$ and $\omega_\gamma^{(N)}(x)$ are polynomials with coefficients in $\boldsymbol{A}_\nu$. Moreover, $\psi_\gamma^{(N)}(x)-x^\gamma$ and $\omega_\gamma^{(N)}(x)-x^\gamma$ are with coefficients in $\boldsymbol{A}_\nu^0$. Hence in (3.13), $c_{\gamma,\delta}$'s belong to $\boldsymbol{A}_\nu^0$. In the course of resolution of (3.15) in $a_{\gamma,\delta}$'s, we must do several times of division by $\mu_j^{(N)}-\mu_m^{(N)}$ with $0\leqslant j<m$. By virtue of (3.9), λ_j's are distinct (we assume here that $\kappa+1>0$ or that $m\geqslant 2$). Therefore, we see that an element of $\boldsymbol{A}_\nu^0$ divided by $\mu_j^{(N)}-\mu_m^{(N)}$ belongs to $\boldsymbol{A}_\nu$. Consequently, $a_{\gamma,\delta}$'s in (3.14) belong to $\boldsymbol{A}_\nu$. This proves (vi).

To prove (vii), we put $\phi_\gamma^{(N)}(x)-\phi_\gamma(x)=\nu\varepsilon_\gamma(x)$. Then, $\varepsilon_\gamma(x)$ is a polynomial of degree$\leqslant(m-1)$ with coefficients in $\boldsymbol{A}_\nu$, hence $\phi_\gamma^{(N)}(x)$ and $\varepsilon_\gamma(x)$ remain in a fixed bounded set in $C^4(\bar{\Omega})$ as ν is small (N is large). And we can write

$$\begin{aligned}[A\phi_\gamma+\lambda_m\phi_\gamma]=\{[A\phi_\gamma^{(N)}]-N(P^{(N)}[\phi_\gamma^{(N)}]-[\phi_\gamma^{(N)}])\}-\nu[(A+\lambda_m)\varepsilon_\gamma]\\ +N(\mu_m^{(N)}-1+N^{-1}\lambda_m)[\phi_\gamma^{(N)}].\end{aligned}$$

The first term between { } on the right hand side is of $O(1/N)$ by virtue of Lemma 2.3, so is the second term and the last term is also of $O(1/N)$ by (3.9). Hence $A\phi_\gamma+\lambda_m\phi_\gamma$ is of $O(1/N)$ on $\Omega^{(N)}$. Therefore, it must be identically 0, because $\Omega^{(N)}$ is sufficiently dense in $\bar{\Omega}$ as N is large. (vii) is now proved.

(viii) is trivial. Q.E.D. (See also (B.2), p. 306 of [10]. Our $\phi_\gamma(x)$ coincides with $Q_\gamma(x)$ there within a constant factor).

§4. A characterization of the diffusion approximation

Let $T^{(N)}$ be the discrete model defined by (1.7), (2.6) and by (3.1). Let A and $\{\phi_\gamma(x);\ \gamma\in\boldsymbol{N}^n\}$ be the operator and the basis of its eigenpolynomials studied in §3. We are going to show

Proposition 4.1. *Let* $\mathscr{T}^{(N)}=\{\mathscr{T}(t/N)\}_{t\geqslant 0}$ *be a diffusion approximation of* $T^{(N)}$ *in the sense of Definition* 2.1. *Then,* $\mathscr{T}$ *is characterized by the following condition:*

$$\mathscr{T}(t)\phi_\gamma(x)=e^{-t\lambda_{|\gamma|}}\phi_\gamma(x) \quad \textit{for all} \quad \gamma\in \boldsymbol{N}^n. \tag{4.1}$$

Proof. A semi-group $\mathscr{T}$ satisfying (4.1) is unique (if there exists any), because (4.1) determines the action of $\mathscr{T}$ on the polynomials which are dense in $C^0(\bar{\Omega})$.

For any $\gamma\in\boldsymbol{N}^n$ and any positive integer K, we put

$$\begin{cases} \phi(t, x)=\mathscr{T}(t)\phi_\gamma(x)-e^{-t\lambda_m}\phi_\gamma(x) \quad \text{with} \quad m=|\gamma| \quad \text{and} \\ \tilde{\phi}^{(N)}(k)=[\mathscr{T}(k/N)\phi_\gamma]-T^{(N)}(k)[\phi_\gamma] \quad \text{for} \quad 0\leqslant k\leqslant KN \end{cases}$$

($\phi(t, x)$ is a function and $\tilde{\phi}^{(N)}(k)$ is a vector). Then we have

$$\begin{aligned} [\phi(k/N, \cdot)]=\tilde{\phi}^{(N)}(k)+T^{(N)}(k)[\phi_\gamma-\phi_\gamma^{(N)}]+e^{-k\lambda_m/N}[\phi_\gamma^{(N)}-\phi_\gamma] \\ +\{(\mu_m^{(N)})^k-e^{-k\lambda_m/N}\}[\phi_\gamma^{(N)}]. \end{aligned} \tag{4.2}$$

Let us assume at first that (4.1) holds. The left hand side of (4.2) is 0, two terms involving $[\phi_\gamma-\phi_\gamma^{(N)}]$ are of $O(1/N)$, and $(\mu_m^{(N)})^k-e^{-k\lambda_m/N}$ is also of $O(1/N)$, because $\mu_m^{(N)}-e^{-\lambda_m/N}$ is of $O(1/N^2)$, $0\leqslant k\leqslant KN$ and that $\mu_m^{(N)}$ and $e^{-\lambda_m/N}$ are at most 1. Hence $\tilde{\phi}^{(N)}(k)$ is of $O(1/N)$. Consider generally the vector

$$\tilde{f}^{(N)}(k)=[\mathscr{T}(k/N)f]-T^{(N)}(k)[f] \quad \text{for} \quad 0\leqslant k\leqslant KN.$$

If f is polynomial, f is represented as a linear combination of ϕ_γ's, hence $\tilde{f}^{(N)}(k)$ is of $O(1/N)$. And for general $f\in C^0(\bar{\Omega})$, $\tilde{f}^{(N)}(k)$ is also small because of the uniform boundedness of norm of $\mathscr{T}(t)$ in the interval $0\leqslant t\leqslant K$, of the contraction property of $T^{(N)}$ and of the density of polynomials in $C^0(\bar{\Omega})$. This proves that $\mathscr{T}^{(N)}$ approximates $T^{(N)}$ if (4.1) holds for $\mathscr{T}$.

Conversely, let $\mathscr{T}$ be a semi-group such that $\mathscr{T}^{(N)}$ approximates $T^{(N)}$. We have to derive (4.1). In this case, if N is large, $\tilde{f}^{(N)}(k)$ is small for any $f\in C^0(\bar{\Omega})$, hence in particular, for $f=\phi_\gamma$. Therefore, $\tilde{\phi}^{(N)}(k)$ is small in (4.2), where the remaining terms on the right hand side are of $O(1/N)$. That is, $\phi(t, x)$, which is a continuous function of (t, x) on $[0, +\infty)\times\bar{\Omega}$, is small if $t=0, 1/N,\dots, k/N,\dots, K$ and if $x\in\Omega^{(N)}$. Consequently, $\phi(t, x)$ must be identically zero if $0\leqslant t\leqslant K$ and $x\in\bar{\Omega}$. Thus, we have (4.1) because K is arbitrary. Q. E. D.

Corollary 4.2. *The semi-group $\mathscr{T}$ satisfying* (4.1) *is positivity-preserving and is of contraction.*

Proof. $T^{(N)}$ is positivity-preserving and is of contraction. Therefore, $\mathscr{T}^{(N)}$ approximating $T^{(N)}$ must have the same properties. Hence, the assertion follows from Proposition 4.1. Q. E. D.

From what we have seen in § 3 and by Proposition 4.1, the question of finding the diffusion approximation is reduced to that of determining the semi-group $\mathscr{T}$ acting on $C^0(\bar{\Omega})$, generated by A and satisfying (4.1).

In the remaining part of this paragraph, we will prepare some notions and notations required in the next paragraph.

Let Π be the set of all the non-empty subsets of $\{1,\dots, d\}$. For $K\in\Pi$, the number of elements in K is denoted by $|K|$. Now, we have to distinguish p's according as $b_p>0$ or as $b_p=0$. Therefore we put

$$J_+=\{p;\ 1\leqslant p\leqslant d \text{ and } b_p>0\} \text{ and } J_0=\{p;\ 1\leqslant p\leqslant d \text{ and } b_p=0\}. \tag{4.3}$$

J_+ and J_0 are complementary to each other. There are three cases:
Case 1. *J_+ is empty;*
Case 2. *J_0 is empty;*
Case 3. *Both of J_+ and J_0 are non-empty.* This case consists of two subcases:

Case 3–1. $|J_+|=1$ *and* $|J_0|=n$;

Case 3–2. $|J_+|\geqq 2$ *and* $|J_0|\geqslant 1$.

The formulas obtained in the next paragraph will depend largely on this classification. The probabilistic meaning will be discussed in §6.

In Case 2 and in Case 3, we denote by Π' the set of all the elements of Π containing J_+ (In Case 2, Π' contains only one element $\{1,\dots, d\}$).

Next, for each $K\in\Pi$, let us define the *open K-face* Ω_K and the *closed K-face* $\bar{\Omega}_K$ of $\bar{\Omega}$ in the following way:

Ω_K is the set of points $x\in\bar{\Omega}$ satisfying $x_j>0$ for any $j\in K$ and $x_j=0$ for any $j\notin K$. $\bar{\Omega}_K$ is the closure of Ω_K in $\boldsymbol{R}^n$ (or in $\boldsymbol{R}^d$). Ω_K (or $\bar{\Omega}_K$) is a $(|K|-1)$-dimensional simplex. In particular, if $K=\{j\}$ $(|K|=1)$, $\Omega_K=\bar{\Omega}_K$ is the j-th vertex P_j of $\bar{\Omega}$:

$$P_j=(\delta_{1j},\dots,\delta_{nj}) \quad \text{for} \quad 1\leqslant j\leqslant d. \tag{4.4}$$

These vertices are naturally of special importance.

Lemma 4.3. *Let K be an element of Π in Case* 1 *and of Π' in Case* 2 *or in Case* 3. *Let $v(\tilde{x})$ be a function in d variables and of class C^2 in a neighborhood of $\bar{\Omega}$. Then, $(\tilde{A}v)|_{\bar{\Omega}_K}$ is determined only by $v|_{\bar{\Omega}_K}$. And, there exists a differential operator A_K on $\bar{\Omega}_K$ such that*

$$(\tilde{A}v)|_{\bar{\Omega}_K}=A_K(v|_{\bar{\Omega}_K}). \tag{4.5}$$

In particular, $A_{\{1,\dots,d\}}=A$, and $A_K=0$ if $|K|=1$ (See Lemma 2.4).

Proof. Let us begin with the assertion for $|K|=1$. Such a K occurs only in Case 1 and in Case 3–1. And $\tilde{A}v$ is zero at this vertex (See (3.4)). Next, let us verify the assertion for $K=\{1,\dots, d\}$. Changing the variables, we rewrite $\tilde{A}$ as

$$\begin{cases} 2\tilde{A}v(\tilde{x})=\sum_{p,q=1}^{d}(\delta_{pq}y_p-y_py_q)\dfrac{\partial^2 v}{\partial y_p\partial y_q}+2\sum_{p=1}^{n}b_p(y)\dfrac{\partial v}{\partial y_p}, \\ \text{where } y_j=x_j \text{ for } 1\leqslant j\leqslant n \text{ and } y_d=1-\sum_{p=1}^{d}y_p. \end{cases}$$

Therefore, $(\tilde{A}v)|_{\bar{\Omega}}=A(v|_{\bar{\Omega}})$ because $y_d=0$ on $\bar{\Omega}$.

The proof of (4.5) for general K is reduced to the above argument replacing d by $|K|$. For example, we have

$$A_K w = \frac{1}{2} \sum_{p,q=1}^{k} (\delta_{pq} x_p - x_p x_q) \frac{\partial^2 w}{\partial x_p \partial x_q} + \sum_{p=1}^{k} (b_p - x_p \sum_{q=1}^{k+1} b_q) \frac{\partial w}{\partial x_p} \tag{4.6}$$

$$\text{if} \quad K = \{1, \ldots, k+1\} \quad \text{with} \quad 1 \leqslant k \leqslant n.$$

The lemma is proved. Q.E.D.

Let us define a subset Ω_e of $\bar{\Omega}$ and $(\partial\Omega)_e$ of $\partial\Omega$ as follows:

$$\Omega_e = \{x \in \bar{\Omega};\ x_p > 0 \text{ if } p \in J_+\} \quad \text{and} \quad (\partial\Omega)_e = \partial\Omega \cap \Omega_e. \tag{4.7}$$

Ω_e contains Ω in any case and it can be written as

$$\Omega_e = \bar{\Omega} \text{ in Case 1, and } \Omega_e = \bigcup_{K \in \Pi'} \Omega_K \text{ in Case 2 or in Case 3.} \tag{4.7$'$}$$

We call $(\partial\Omega)_e$ the *effective part of* $\partial\Omega$. The reason for this name is that, if $t > 0$, $\mathscr{T}(t)f$ is determined only by the values of f on Ω_e (This notion is slightly different from that of attainability. For example, $\partial\Omega \cap \{x_j = 0\}$ (or its interior part) is non-attainable if and only if $b_j \geqslant 1/2$ in the sense of Chap. 11 of [5]). By the above definition, we see in particular

$$(\partial\Omega)_e = \partial\Omega \text{ in Case 1, and } (\partial\Omega)_e = \phi \text{ in Case 2.} \tag{4.8}$$

The construction in the next paragraph will give us the following property of $\mathscr{T}$. Let us put $u(t, x) = \mathscr{T}(t)f(x)$, where $f \in C^0(\bar{\Omega})$. Then, u is a solution of the initial boundary value problem

$$\frac{\partial u}{\partial t} = Au \quad \text{on} \quad \Omega_e \quad \text{if} \quad t > 0, \quad \text{and,} \tag{4.9}$$

$$u(0, x) = f(x) \quad \text{on} \quad \bar{\Omega}. \tag{4.10}$$

(The uniqueness of solution follows under the assumptions of regularity: $u(t, x)$ is continuous in (t, x) on $[0, +\infty) \times \bar{\Omega}$, differentiable in t in $C^0(\bar{\Omega})$-topology if $t > 0$ and continuous in t in $C^2(\bar{\Omega})$-topology if $t > 0$).

Here, it should be remarked that (4.9) means not only the differential equation in the interior set Ω but also the boundary condition on $(\partial\Omega)_e$ in the sense of Lemma 4.3. In particular, we have

$$\begin{cases} (\mathscr{T}(t)f)(P_j) = f(P_j) \quad \text{for} \quad t \geqslant 0 \\ \text{if } 1 \leqslant j \leqslant d \text{ in Case 1 and if } j \in J_+ \text{ in Case 3–1} \end{cases} \tag{4.11}$$

(See (5.15) and (5.17)).

§5. Formulas for the diffusion approximation

Let $\{\lambda_m\}_{m=0}^{\infty}$ be defined by (3.10), that is, $\lambda_m = m(m+\kappa)/2$ for $m \in \boldsymbol{N}$, where $\kappa + 1 = 2 \sum_{p=1}^{d} b_p$. This sequence is non-negative and strictly increasing except that $\lambda_0 = \lambda_1 = 0$ in Case 1.

For each $m \in \boldsymbol{N}$, we define $\boldsymbol{E}_m$ as the vector subspace of $C^0(\bar{\Omega})$ spanned by $\{\phi_\gamma(x); \gamma \in \boldsymbol{N}^n$ and $|\gamma| = m\}$, but in Case 1, $\boldsymbol{E}_1$ is the span of $\{x_p\}_{p=1}^d$ and $\boldsymbol{E}_0$ is not defined. That is, $\boldsymbol{E}_m$ is the space of eigenpolynomials of A belonging to the eigenvalue $-\lambda_m$ except that $\boldsymbol{E}_1$ in Case 1 is the eigenspace belonging to 0.

We are going to obtain a kernel representation of the semi-group $\mathscr{T}$ characterized by (4.1). Our first step is to define, for each m, the continuous linear mapping E_m from $C^0(\bar{\Omega})$ into itself having the following two properties:

— E_m is a projection from $C^0(\bar{\Omega})$ onto $\boldsymbol{E}_m$;

— There exists a Radon measure $E_m(x, dy)$ on $\boldsymbol{R}^n$ with respect to variable y with parameter $x \in \bar{\Omega}$ such that $E_m(x, dy)$ is supported by $\bar{\Omega}$ and that

$$E_m f(x) = \int_{\bar{\Omega}} f(y) dE_m(x, dy) \text{ on } \bar{\Omega} \text{ for any } f \in C^0(\bar{\Omega}). \tag{5.1}$$

Let us assume that the mappings E_m were already constructed. Then, we can informally define a semi-group $\mathscr{T} = \{\mathscr{T}(t)\}_{t \geqslant 0}$ by $\mathscr{T}(0)f = f$ and by

$$\begin{cases} \mathscr{T}(t)f = \sum_{m=1}^{\infty} e^{-t\lambda_m} E_m f \text{ in Case 1, and} \\ \mathscr{T}(t)f = \sum_{m=0}^{\infty} e^{-t\lambda_m} E_m f \text{ in Case 2 and in Case 3,} \end{cases} \tag{5.2}$$

for $t > 0$. The series on the right hand side is not yet guaranteed to be convergent for general $f \in C^0(\bar{\Omega})$. But for polynomial f, it is really a finite sum. In particular, (4.1) is immediately checked to be valid.

Therefore, we will define at first the measures $E_m(x, dy)$ in detail. Our second step will be to rearrange the infinite sum in (5.2) and to verify the convergence of the rearranged sum. The results of these two steps will be summarized as Lemma 5.1 and Lemma 5.2 below.

5–1. Definition of $\boldsymbol{E}_m(\boldsymbol{x}, \boldsymbol{dy})$

Let us begin with the definition of $dS_K(x)$, volume element of the open K-face Ω_K, for each $K \in \Pi$ in Case 1, and for each $K \in \Pi'$ in Case 2 and in Case 3. If $K = \{j\}$ ($|K| = 1$), $dS_K(x)$ is the point mass 1 at P_j. If $K = \{j_1, \ldots, j_k\}$ ($|K| = k \geqslant 2$), we put

$$dS_K(x) = m(x) |dx_{j_1} \cdots \widehat{dx_{j_p}} \cdots dx_{j_k}| \quad \text{for} \quad 1 \leqslant p \leqslant k, \tag{5.3}$$

$$\begin{cases} \text{where } m(x) = 1 \text{ in Case 1 and} \\ m(x) = \prod_{j \in J_+} x_j^{2b_j - 1} \text{ in Case 2 and in Case 3.} \end{cases} \tag{5.4}$$

Extending by 0 outside of $\bar{\Omega}_K$, $dS_K(x)$ is regarded as a non-negative Radon measure on $\boldsymbol{R}^n$ supported by $\bar{\Omega}_K$.

Next, we define the functions $F_{p,K}(x, y)$ as follows:

$$F_{p,K}(x, y) = \{ \prod_{j \in K'} x_j \} \sum_{|\tilde{\alpha}| = p} \frac{x^{\tilde{\alpha}} y^{\tilde{\alpha}}}{(\tilde{\alpha} + \tilde{\omega}(K))! \tilde{\alpha}!} \quad \text{for} \quad p \in \boldsymbol{N}, \tag{5.5}$$

$$\begin{cases} \text{where } \tilde{\omega}(K)=(\omega_1(K),\dots,\omega_d(K)) \text{ with } \omega_j(K)=1 \text{ or } 2b_j-1 \\ \text{according as } j\in K' \text{ or otherwise, and } K'=K\cap J_0 \end{cases} \tag{5.6}$$

(The factor $\prod_{j\in K'} x_j$ is 1 by definition if K' is empty). We put conventionally $F_{p,K}(x, y)=0$ if $p<0$. Hereafter (x, y) in $F_{p,K}(x, y)$ are restricted to $\bar{\Omega}\times\bar{\Omega}_K$. Therefore, the summation on the right hand side of (5.5) can be restricted to $\tilde{\alpha}$'s for which $\alpha_p=0$ if $p\notin K$, because $y^{\tilde{\alpha}}$ is 0 on $\bar{\Omega}_K$ if $\alpha_p>0$ for some $p\notin K$.

Thirdly, let us define the functions $E_{m,K}(x, y)$ by

$$E_{m,K}(x, y)=\sum_{q=0}^{m-|K'|}\frac{(2m+\kappa)(2m-q+\kappa-1)!}{(-1)^q q!}F_{m-q-|K'|,K}(x. y) \quad \text{for} \quad m\geqslant|K'|. \tag{5.7}$$

We put conventionally $E_{m,K}(x, y)=0$ if $m<|K'|$.

Now we can define the measure $E_m(x, dy)$ as follows:

$$\begin{cases} E_m(x, dy)=\sum_{K\in\Pi}E_{m,K}(x, y)dS_K(y) \text{ for } m\geqslant 1 \text{ in Case 1, and} \\ E_m(x, dy)=\sum_{K\in\Pi'}E_{m,K}(x, y)dS_K(y) \text{ for } m\geqslant 0 \text{ in Case 2 and in Case 3.} \end{cases} \tag{5.8}$$

Using this, we define the mapping E_m by (5.1).

Lemma 5.1. *For each m, E_m is a continuous projection from $C^0(\bar{\Omega})$ onto $\boldsymbol{E}_m$, that is,*

(i) $AE_mf=-\lambda_mE_mf$ *for any* $f\in C^0(\bar{\Omega})$;

(ii) $E_mAf=-\lambda_mE_mf$ *for any* $f\in C^2(\bar{\Omega})$;

(iii) $E_mf=f$ *for any* $f\in\boldsymbol{E}_m$.

This lemma will be proved in § A. We put now, for $t>0$,

$$\begin{cases} Z(t; x, dy)=\sum_{m=1}^{\infty}e^{-t\lambda_m}E_m(x, dy) \text{ in Case 1, and} \\ Z(t; x, dy)=\sum_{m=0}^{\infty}e^{-t\lambda_m}E_m(x, dy) \text{ in Case 2 and in Case 3.} \end{cases} \tag{5.9}$$

If the series on the right hand side is convergent, $Z(t; x, dy)$ is a Radon measure on $\boldsymbol{R}^n$ supported by $\bar{\Omega}$ with respect to y and with parameters $t>0$ and $x\in\bar{\Omega}$. And the informal definition (5.2) is rewritten as

$$\mathscr{T}(t)f(x)=\int_{\bar{\Omega}}f(y)Z(t; x, dy), \quad \text{for} \quad t>0. \tag{5.10}$$

Therefore, our next work is to justify the sum on the right hand side of (5.9).

5–2. Another expression of $Z(t; x. dy)$

By (5.8), each of $E_m(x, dy)$'s is a linear combination of $dS_K(y)$'s. Hence, we can rearrange $Z(t; x, dy)$ and write it down also as a linear combination of $dS_K(y)$'s:

$$\begin{cases} Z(t;\, x,\, dy) = \sum\limits_{K\in\Pi} z_K(t;\, x,\, y)dS_K(y) \text{ in Case 1, and} \\ Z(t;\, x,\, dy) = \sum\limits_{K\in\Pi'} z_K(t;\, x,\, y)dS_K(y) \text{ in Case 2 and in Case 3,} \end{cases} \tag{5.11}$$

where the number of K's is finite in any case. And the functions $z_K(t;\, x,\, y)$ are defined by

$$z_K(t;\, x,\, y) = \sum_{m=|K'|}^{\infty} e^{-t\lambda_m}E_{m,K}(x,\, y), \quad \text{where} \quad K' = K \cap J_0. \tag{5.12}$$

For any fixed $t>0$, the series on the right hand side converges uniformly on $\bar{\Omega}\times\bar{\Omega}_K$, because we have an estimate of type $|E_{m,K}(x,\, y)| \leqslant C^{m+1}$ with some constant C independent of $(x,\, y,\, m)$ (See (4.13), p. 294 of [10]). Therefore, for any $f\in C^0(\bar{\Omega})$ and for any $t>0$, the integral on the right hand side of (5.10) is convergent. But, we have a more precise assertion which guarantees, as a corollary, the uniform convergence with respect to t. That is,

Lemma 5.2. *For each K, the function $z_K(t;\, x,\, y)$ defined by* (5.12) *is non-negative if $t>0$ and if $(x,\, y)\in\bar{\Omega}\times\bar{\Omega}_K$.*

This will be proved in § B. As a consequence of this lemma, $Z(t;\, x,\, dy)$ is revealed to be a non-negative Radon measure. And moreover we have

$$\int_{\bar{\Omega}} Z(t;\, x,\, dy) = 1 \quad \text{on} \quad \bar{\Omega} \quad \text{if} \quad t>0, \tag{5.13}$$

because this integral is equal to the right hand side of (5.2) (with $f=1$), therefore, to $E_1 1$ (or $E_0 1) = 1$.

Thus, $\mathscr{T}(t)f$, defined by (5.10) for $t>0$ and by $\mathscr{T}(0)f=f$, has the following properties:

$$0 \leqslant \mathscr{T}(t)f(x) \leqslant 1 \quad \text{if} \quad 0 \leqslant f(x) \leqslant 1, \quad \text{and,} \quad \mathscr{T}(t)1 = 1;$$

$$\mathscr{T}(t+t')f = \mathscr{T}(t)\mathscr{T}(t')f \quad \text{if} \quad t \geqslant 0 \quad \text{and} \quad t' \geqslant 0;$$

$\mathscr{T}(t)f$ is strongly continuous with respect to $t\geqslant 0$ in $C^0(\bar{\Omega})$-topology.

The definition of the semi-group $\mathscr{T}$ is now complete, and it coincides with that of (5.2) for polynomials. In particular, $\mathscr{T}$ satisfies (4.1). Let us state the result just proved:

5–3. Main theorem

Let $T^{(N)}$ be the multi-type gene frequency model of Wright-Fisher defined by (1.7), (2.6) *and by* (3.1). *Let $\mathscr{T}$ be the semi-group defined by* (5.10) *for $t>0$ and by $\mathscr{T}(0)f=f$. We put $\mathscr{T}^{(N)} = \{\mathscr{T}(t/N)\}_{t\geqslant 0}$. Then, $\mathscr{T}^{(N)}$ approximates $T^{(N)}$ as N is large in the sense of Definition* 2.1.

5–4. Behavior of $\mathscr{T}(t)$ as t is large

This will be very important in the next paragraph to study the movement of the diffusion process. As is shown in (5.12) above, we see, for any K,

$$z_K(t;\, x,\, y)=O\{\exp(-t\lambda_{|K'|})\} \quad \text{as} \quad t\to+\infty, \quad \text{where} \quad K'=K\cap J_0. \tag{5.14}$$

That is, the contribution of $f|_{\Omega_K}$ to $\mathscr{T}(t)f$ is of this order. In particular, that of $f|_\Omega$ is of $O[\exp\{-t|J_0|(|J_0|+\kappa)/2\}]$.

Let us calculate the limit of $\mathscr{T}(t)f$ as $t\to+\infty$. For this, it suffices to pick up the single term $E_1 f$ in Case 1 or $E_0 f$ in Case 2 and in Case 3 from the right hand side of (5.2). The result is the following:

Case 1 $\quad \mathscr{T}(t)f(x)=\sum_{p=1}^{d} x_p f(P_p)+O(e^{-t});$ (5.15)

Case 2 $\quad \mathscr{T}(t)f(x)=\left\{\int_\Omega f(y)m(y)dy\Big/\int_\Omega m(y)dy\right\}+O(e^{-t(\kappa+1)/2}),$

where $m(y)=\prod_{j=1}^{d} y_j^{2b_j-1}$ and $dy=dy_1\cdots dy_n;$ (5.16)

Case 3–1 $\quad \mathscr{T}(t)f(x)=f(P_k)+O(e^{-tb_k})$ if $J_+=\{k\};$ (5.17)

Case 3–2 $\quad \mathscr{T}(t)f(x)=\left\{\int_{\Omega_{J_+}} f(y)dS_{J_+}(y)\Big/\int_{\Omega_{J_+}} dS_{J_+}(y)\right\}+O(e^{-t(\kappa+1)/2}).$ (5.18)

5–5. Simplifications of formulas

Let us introduce an auxiliary function $V_\nu(t,\zeta)$ with variables $t>0$, $-1\leqslant\zeta\leqslant1$ and a parameter $\nu>-1$:

$$V_\nu(t,\zeta)=\Gamma(\nu)\sum_{m=0}^{\infty}(2m+\nu)C_{2m}^{\nu}(\zeta)e^{-tm(m+\nu)/2}, \tag{5.19}$$

where $\{C_q^\nu(\zeta)\}_{q=0}^{\infty}$ are Gegenbauer polynomials (See [2]):

$$(1-2\xi\zeta+\xi^2)^{-\nu}=\sum_{q=0}^{\infty}\xi^q C_q^\nu(\zeta) \quad \text{for} \quad -1\leqslant\zeta\leqslant1 \quad \text{and} \quad |\xi|<1 \tag{5.20}$$

(The definition of $V_\nu(t,\zeta)$ is not rigorous for $-1<\nu\leqslant0$. See § B).

In Case 1, for each $K\in\Pi$, $z_K(t;\, x,\, y)$ in (5.11) is given by

$$z_K(t;\, x,\, y)=\Big(\prod_{j\in K}x_j\Big)e^{-t\lambda_{|K|}}\times\int_{-1}^{1}\cdots\int_{-1}^{1}V_{2|K|-1}\Big(t,\sum_{p\in K}\phi_p\sqrt{x_p y_p}\Big)M(d\tilde{\phi}), \tag{5.21}$$

the integral being extended with respect to $\tilde{\phi}=(\phi_1,\ldots,\phi_d)$ over the d-dimensional interval $(-1,1)^d$ by means of the measure

$$M(d\tilde{\phi})=\prod_{j=1}^{d}\left(\frac{2}{\pi}\sqrt{1-\phi_j^2}\,d\phi_j\right). \tag{5.22}$$

This z_K with $K=\{1,\ldots,d\}$ is the same as the Green function obtained in [10] (See (7.13), p. 300).

In Case 2, Π' consists only of $K=\{1,\ldots,d\}$. Therefore, (5.11) can be written simply as

$$\begin{cases} Z(t;\, x,\, dy)=z(t;\, x,\, y)m(y)dy \\ \text{with } z(t;\, x,\, y)=\sum_{p=0}^{\infty}F_p(x,y)V_{2p+\kappa}(t,0)e^{-t\lambda_p}, \end{cases} \tag{5.23}$$

where $m(y)dy$ is as in (5.16) and F_p stands for $F_{p,K}$ with $K=\{1,\dots, d\}$. This result is the same as the Green function obtained in [10] (See (6.5), p. 297), that is,

$$\begin{cases} z(t;\, x,\, y)=\int_{-1}^{1}\cdots\int_{-1}^{1} V_\kappa(t,\ \sum_{p=1}^{d}\phi_p\sqrt{x_p y_p})M'\,(d\tilde{\phi}), \\ \text{where } M'(d\tilde{\phi})=\prod_{j=1}^{d}\left[(1-\phi_j^2)^{2b_j-(3/2)}d\phi_j\Big/\left\{\sqrt{\pi}\,\Gamma\left(2b_j-\frac{1}{2}\right)\right\}\right]. \end{cases} \tag{5.23'}$$

In Case 3, for each $K\in\Pi'$, z_K in (5.11) is given by

$$z_K(t;\, x,\, y)=\sum_{p=0}^{\infty}F_{p,K}(x,\, y)V_{2p+2|K'|+\kappa}(t,\, 0)e^{-t\lambda_p}, \tag{5.24}$$

where $K'=K\cap J_0$. An expression analogous to (5.21) or to (5.23′) is also possible.

5–6. A remark

The discrete model $T^{(N)}$ is determined by two vectors of parameters $\boldsymbol{c}=(c_1,\dots, c_d)$ and $\boldsymbol{\mu}=(\mu_1,\dots, \mu_d)$ independently running over the closed set $\bar{Q}=\{\boldsymbol{b}=(b_1,\dots, b_d);\ b_p\geqslant 0$ for $1\leqslant p\leqslant d\}$. And the semi-group $\mathscr{T}$ is determined by $\boldsymbol{b}=\boldsymbol{c}+\boldsymbol{\mu}$ also running over $\bar{Q}$. In Case 2, $\boldsymbol{b}$ lies in the interior of $\bar{Q}$. In Case 1, $\boldsymbol{b}$ is at the vertex $\boldsymbol{O}$. And in Case 3, $\boldsymbol{b}$ is on the remaining part of $\partial\bar{Q}$. Although the formulas of $\mathscr{T}$ are, apparently, written variously case by case, $\mathscr{T}$ depends continuously on $\boldsymbol{b}$. To be more precise, let us denote $\mathscr{T}=\{\mathscr{T}(t,\, \boldsymbol{b})\}_{t\geqslant 0}$. Then, for any $\boldsymbol{b}_0\in\bar{Q}$, $f\in C^0(\bar{\Omega})$, $\varepsilon>0$ and for any $T>0$, there exists a $\delta>0$ such that $|\mathscr{T}(t,\, \boldsymbol{b})f(x)-\mathscr{T}(t,\, \boldsymbol{b}_0)f(x)|<\varepsilon$ on $\bar{\Omega}$ if $0\leqslant t\leqslant T$, $\boldsymbol{b}\in\bar{Q}$ and if $\|\boldsymbol{b}-\boldsymbol{b}_0\|<\delta$ (For the proof, it suffices to verify this essentially for polynomials f). Hence, each model $\mathscr{T}$ in Case 1 or in Case 3 is obtained as the limit of a sequence of models in Case 2.

§ 6. Some probabilistic interpretations

Let $\{\alpha^{(N)}(k)\}_{k=0}^{\infty}$ be the discrete model defined by (1.7) and (3.1). Denote by $T^{(N)}=\{T^{(N)}(k)\}_{k=0}^{\infty}$ be the semi-group of matrices identified with it. Then by (2.6), we can write

$$T^{(N)}(k)\phi(\alpha)=E_\alpha[\phi(\alpha^{(N)}(k))] \quad \text{for} \quad \alpha\in\Omega^{(N)} \quad \text{and} \quad k\in\boldsymbol{N}. \tag{6.1}$$

On the other hand, let us define the diffusion process $\xi(t,\, \omega)$ as the solution of the following stochastic differential equation

$$\begin{cases} d\xi_p(t,\, \omega)=\sum_{q=1}^{n}\sigma_{pq}(\xi(t,\, \omega))dB_q(t,\, \omega)+b_p(\xi(t,\, \omega))dt, \\ \xi_p(t,\, \omega)=x_p \quad \text{for} \quad 1\leqslant p\leqslant n \quad \text{with} \quad x=(x_1,\dots, x_n)\in\bar{\Omega}, \end{cases} \tag{6.2}$$

where $b_p(x)$'s are defined by (3.2), $\{B_p(t,\, \omega)\}_{p=1}^{n}$ is a n-dimensional Brownian motion and $\sigma(x)=(\sigma_{pq}(x))$ is $n\times n$-matrix whose elements are Hölder continuous functions of order 1/2 on $\bar{\Omega}$ such that

$$\sum_{r=1}^{n}\sigma_{pr}(x)\sigma_{qr}(x)=\delta_{pq}x_p-x_px_q \quad \text{on} \quad \bar{\Omega} \quad \text{for} \quad 1\leqslant p,\, q\leqslant n. \tag{6.3}$$

For example, if $\sigma(x)$ is a triangular matrix, this Hölder continuity is guaranteed. The solution $\xi(t, \omega)$ of (6.2) is unique for any $x \in \bar{\Omega}$, remains almost surely on $\bar{\Omega}$ and can be continued up to $t \to +\infty$. Therefore the equation (6.2) defines a diffusion process $\xi(t, \omega)$ (See [8]).

Then, the semi-group $\mathscr{T} = \{\mathscr{T}(t)\}_{t \geqslant 0}$ constructed in §5 can be expressed, by means of $\xi(t, \omega)$, as follows

$$\mathscr{T}(t)f(x) = E_x[f(\xi(t, \omega))] \quad \text{for} \quad t \geqslant 0,\ x \in \bar{\Omega} \quad \text{and} \quad f \in C^0(\bar{\Omega}), \tag{6.4}$$

because each of the both sides is a solution of the initial-boundary value problem (4.9)–(4.10) of which the solution is unique (We do not go into further details of the proof of (6.4). See [5] for example).

We are now going to discuss the Cases 1, 2, 3–1 and 3–2 separately, and give some probabilistic interpretations of our results obtained in the present memoir and in [10].

6–1. Case 1

We have $\zeta_p(\alpha) = \alpha_p/N$ for $1 \leqslant p \leqslant d$ in this case. Therefore, the discrete model represents a population in which effect neither selection force, mutating pressure nor migration.

The vertices $P_1, \dots, P_d$ are states of fixation (homozygosity), that is, the states that the population consists of only one type of individuals. And by (5.15), we have

$$E_x[f(\xi(t/N, \omega))] = \sum_{p=1}^{d} x_p f(P_p) + O(e^{-t/N}) \quad \text{as } t/N \text{ is large.} \tag{6.5}$$

Hence, the diffusion process $\xi(t/N, \omega)$ starting from $\xi(0, \omega) = x \in \bar{\Omega}$ is absorbed almost surely in a finite time to one of the vertices, with probability x_p to P_p $(1 \leqslant p \leqslant d)$, and that the rate of absorption is $1/N$. Genetically, any population governed by this rule becomes ultimately homozygous one of type p with probability x_p (=the frequency of the type p at the 0-th generation).

Let $\tau(\omega)$ be the first exit time from Ω of $\xi(t, \omega)$. Then, its expectation is given by

$$\begin{aligned} E_x[\tau(\omega)] &= 2(d-2)! \int_0^{x_1} \cdots \int_0^{x_d} \Big(\sum_{p=1}^{d} u_p\Big)^{1-d} du_1 \cdots du_d \\ &= 2 \sum_{K \in \Pi} (-1)^{|K|+1} g\Big(\sum_{j \in K} x_j\Big), \quad \text{where} \quad g(u) = u \log \frac{1}{u}. \end{aligned} \tag{6.6}$$

This is calculated as follows. The left hand side $v(x)$ is the solution of the Dirichlet problem

$$Av(x) = -1 \quad \text{in} \quad \Omega \quad \text{and} \quad v(x) = 0 \quad \text{on} \quad \partial\Omega. \tag{6.7}$$

Therefore, $v(x)$ is equal to the right hand side of (6.6) (See (9.9), p. 303 of [10]). Genetically, $NE_x[\tau(\omega)]$ approximates the expectation of generations required for the extinction of at least one type of individuals from the population.

Let us denote by D_p the event that the type p disappears at first from the population: $D_p=\{\omega;\ \xi_p(\tau(\omega),\ \omega)=0\}$. Then, the probability of D_p is given by

$$P_x(D_p)=\mathrm{vol}\,D_p(x)/\mathrm{vol}\,\Omega, \qquad 1\leqslant p\leqslant d, \tag{6.8}$$

where the volume means the n-dimensional Lebesgue measure, and $D_p(x)=\{y\in\Omega;\ y_j/x_j\leqslant y_p/x_p$ for $1\leqslant j\leqslant d\}$. This is proved as follows. $u_p(x)=P_x(D_p)$ is the solution of the Dirichlet problem

$$\begin{cases} Au_p(x)=0 \quad \text{in} \quad \Omega,\ u_p(x)=1 \quad \text{on} \quad \partial\Omega\cap\{x_p=0\} \\ \text{and } u_p(x)=0 \text{ on the remaining part of } \partial\Omega. \end{cases} \tag{6.9}$$

Hence we have (6.8) (See (9.21), p. 304 of [10]). The ratio on the right hand side of (6.8) is nearly equal to the ratio of the numbers of points of $D_p(x)\cap\Omega^{(N)}$ and of $\Omega^{(N)}$.

Reaching $\partial\Omega\cap\{x_p=0\}$, the path starts afresh but remains thereafter almost surely on $\partial\Omega\cap\{x_p=0\}$ and $\bar{\xi}(t,\ \omega)=\xi(t+\tau(\omega),\ \omega)$ is governed by the stochastic differential equation (6.2) (in $(n-1)$-dimensional space with the initial value $\bar{\xi}(0,\ \omega)=\xi(\tau(\omega),\ \omega)$). Genetically, none of types once disappeared from the population does never revive. We could also calculate the expectation of $\tau(\omega)$ under the condition that D_p occurs. For this, we put $w_p(x)=u_p(x)E_x[\tau(\omega)\,|\,D_p]$, and solve the Dirichlet problem

$$Aw_p(x)=-u_p(x) \quad \text{in} \quad \Omega \quad \text{and} \quad w_p(x)=0 \quad \text{on} \quad \partial\Omega \tag{6.10}$$

using the Green function obtained in [10] (See (7.10), p. 300).

Let $\tau_p(\omega)$ the fixation time of $\xi(t,\ \omega)$ at P_p. Then, the expectation of $\xi_p(\omega)$ under the condition that $\xi(t,\ \omega)$ is ultimately fixed at P_p is given by

$$E_x[\tau_p(\omega)\,|\,\xi(+\infty,\ \omega)=P_p]=2g(1-x_p)/x_p \quad \text{for} \quad 1\leqslant p\leqslant d, \tag{6.11}$$

where $g(u)$ is defined in (6.6). In fact, let us put $v_p(x)$ the left hand side multiplied by x_p. It is the solution of the Dirichlet problem

$$Av_p(x)=-x_p \text{ on } \bar{\Omega} \text{ except at } P_p \text{ and } v_p(P_p)=0. \tag{6.12}$$

We can find a particular solution depending only on x_p, which is the unique solution because of the maximum principle. Therefore we have (6.11).

Let $\sigma(\omega)$ be the fixation time of $\xi(t,\ \omega)$ at any one of vertices. Then, its expectation is given by

$$E_x[\sigma(\omega)]=2\sum_{p=1}^{d} g(1-x_p), \tag{6.13}$$

because the left hand side is equal to the sum of $v_p(x)$'s above. Genetically, $NE_x[\sigma(\omega)]$ approximates the expectation of generations required for the given population to be reduced to homozygous one.

6–2. Case 2

This is the case extremely opposite to the Case 1. The discrete model represents a population for which there is no selection force but, for each p, either the mutating pressure or the migration (or both of them) is really effected. That is $\zeta_p(\alpha)>0$ even if $\alpha_p=0$ (See (3.1)). Hence, even if the number of individuals of a type decreases temporarily, it is recovered either by mutation or by migration.

In the diffusion model, no part of the boundary is effective (See (4.8)), because $b_p(x)=b_p>0$ on the hyperplane $\{x_p=0\}$ for each p (See (6.2)). Hence, neither fixation nor extinction of any type can occur. And we have

$$E_x[f(\xi(t/N,\omega))]=\left\{\int_\Omega f(y)m(y)dy\Big/\int_\Omega m(y)dy\right\} + O\left\{\exp\left(-\frac{t}{N}\sum_{p=1}^{d} b_p\right)\right\} \quad \text{as} \quad t/N \text{ is large} \tag{6.14}$$

(See (5.16)). Therefore, independently of the initial state $\xi(0,\omega)=x\in\bar{\Omega}$, $\xi(t/N,\omega)$ is uniformly distributed in Ω with respect to the measure $m(y)dy$ as t/N is large.

6–3. Case 3–1

Assume, for example, that $J_+=\{d\}$ and $J_0=\{1,\dots,n\}$. Then we have, by (3.1) and (3.2),

$$\begin{cases} \zeta_p(\alpha)=\dfrac{\alpha_p(N-\mu_d)}{N(N+c_d)} \text{ for } 1\leqslant p\leqslant n \text{ and } \zeta_d(\alpha)=\dfrac{N(b_d+\alpha_d)-\mu_d\alpha_d}{N(N+c_d)} \\ \text{with } b_d=c_d+\mu_d>0; \end{cases} \tag{6.15}$$

$$b_p(x)=-b_dx_p \quad \text{for} \quad 1\leqslant p\leqslant n \quad \text{and} \quad b_d(x)=(1-x_d)b_d. \tag{6.16}$$

In the discrete model, the fixation occurs only for the type d, because only individuals of type d are recruited by mutation or by migration. In the diffusion model, the closed J_0-face is the non-effective part of the boundary, that is, $(\partial\Omega)_e=\partial\Omega\cap\{x_d>0\}$ (See (4.7)). And, by (5.17), we have

$$E_x[f(\xi(t/N,\omega))]=f(P_d)+O(e^{-tb_d/N}) \quad \text{as} \quad t/N \text{ is large.} \tag{6.17}$$

This means that the diffusion process $\xi(t/N,\omega)$ starting from $\xi(0,\omega)=x\in\bar{\Omega}$ is absorbed almost surely in a finite time to the vertex P_d and that the rate of the absorption is b_d/N. Genetically, any population becomes ultimately homozygous one of type p with probability 1.

Let $\sigma(\omega)$ be the absorption time of $\xi(t,\omega)$ to P_d. Then, its expectation is given by

$$E_x[\sigma(\omega)]=2\int_{x_d}^{1}dz\int_0^z\frac{(y/z)^{2b_d}}{y(1-y)}dy. \tag{6.18}$$

This is calculated as follows. The left hand side $v(x)$ is the solution of the Dirichlet problem

$$Av(x)=-1 \quad \text{on} \quad \Omega_e \quad \text{and} \quad v(P_d)=0. \tag{6.19}$$

This admits a particular solution depending only on x_d, which is the unique solution by virtue of the maximum principle. Genetically, $NE_x[\sigma(\omega)]$ approximates the expectation of generations required for the given population to be reduced to homozygosity of type d.

Analogously to the Case 1, we can calculate explicitly the following expectation and probability (the formulas are omitted here but see §9 of [10] for details of the method):

$E_x[\tau(\omega)]$, where $\tau(\omega)$ is the first exit time of $\xi(t, \omega)$ from Ω;

$P_x(D_p)$ for $1 \leqslant p \leqslant n$, where D_p is as in 6–1.

6–4. Case 3–2

Suppose, for example, that $J_+ = \{1,\dots, m\}$ and $J_0 = \{m+1,\dots, d\}$ with $2 \leqslant m < d$. In this case, fixation does occur to none of types, because individuals of plural number of types $1,\dots, m$ are recruited by mutation or migration.

In the diffusion model, the effective part of the boundary is $(\partial\Omega)_e = \{x \in \partial\Omega;\ x_p > 0 \text{ for } 1 \leqslant p \leqslant m\}$. Above all, the open J_+-face is of special importance, because we have by (5.18)

$$E_x[f(\xi(t/N, \omega))] = \left\{\int_{\Omega_{J_+}} f(y)dS_{J_+}(y) \Big/ \int_{\Omega_{J_+}} dS_{J_+}(y)\right\} + O\left\{\exp\left(-\frac{t}{N}\sum_{p=1}^{m} b_p\right)\right\} \quad \text{as} \quad t/N \text{ is large}, \tag{6.2}$$

where $dS_{J_+}(y)$ is defined by (5.3)–(5.4). This means that the diffusion process starting from $\xi(0, \omega) = x \in \bar{\Omega}$ is absorbed almost surely in a finite time to Ω_{J_+}.

Let $\tau_+(\omega)$ be the absorption time of $\xi(t, \omega)$ to Ω_{J_+}. Then, reaching Ω_{J_+}, the path starts afresh but remains thereafter almost surely in Ω_{J_+} and $\tilde{\xi}(t, \omega) = \xi(t + \tau_+(\omega), \omega)$ is governed by the stochastic differential equation (6.2) in $(m-1)$-dimensional space. And, independently of the starting point, $\xi(t/N, \omega)$ is uniformly distributed in Ω_{J_+} with respect to the measure $dS_{J_+}(y)$ as t/N is large. Genetically, individuals of types $m+1,\dots, d$ disappear sooner or later, and the population behaves thereafter likely as one consisting only of types $1,\dots, m$.

The expectation of $\tau_+(\omega)$ is given by

$$\begin{cases} E_x[\tau_+(\omega)] = 2\displaystyle\int_{x_+}^{1} dz \int_0^z \frac{(y/z)^{2b_+}}{y(1-y)}\, dy, \\ \text{where } x_+ = \displaystyle\sum_{p=1}^{m} x_p \text{ and } b_+ = \sum_{p=1}^{m} b_p. \end{cases} \tag{6.21}$$

This is calculated by solving the Dirichlet problem

$$Av_+(x) = -1 \quad \text{on} \quad \Omega_e \quad \text{and} \quad v_+(x) = 0 \quad \text{on} \quad \Omega_{J_+}. \tag{6.22}$$

Since the solution is unique, we can look for a particular solution depending only on x_+, and we have (6.21) (Compare this with (6.18)).

6–5. An open question

Suppose that the number d of alleles is very large. Let us put $d = \infty$ for example.

We can ask ourselves what happens then to our formulas. As a matter of course, we should reconstruct all of our framework. We should define anew the state space $\bar{\Omega}$, the function space $C^0(\bar{\Omega})$, the notion of the solutions of the differential equation (4.9), and that of the stochastic differential equation (6.2), and so on. We do not go into the details of this point.

But, some of our formulas obtained in § 5 will be available without any modification. At first in Case 1, (5.11) and (5.21) go well whether d is finite or not. Secondly, if the set J_+ is finite, (5.11) and (5.24) in Case 3 are also available. A difficulty arises when J_+ is an infinite set. But in this case also, the majority of the formulas in § 5 may be valid, if we put the hypothesis that $\sum_{p=1}^{\infty} b_p$ be finite.

§ A. Proof of Lemma 5.1

This lemma in Case 2 is exactly the same as the contents in § 3 of [10], hence there is nothing to do in this case. And, if $m=1$ in Case 1 or $m=0$ in Case 3, the assertions of lemma are immediately verified, because we have

$$\begin{cases} E_1 f(x) = \sum_{p=1}^{d} x_p f(P_p) \text{ in Case 1, and,} \\ E_0 f(x) = \int_{\Omega_{J_+}} f(y)\, dS_{J_+}(y) \Big/ \int_{\Omega_{J_+}} dS_{J_+}(y) \text{ in Case 3} \end{cases} \tag{A.1}$$

(See (5.21), (5.23) and (5.24)). Therefore in the sequel, we can assume that $m \geqslant 2$ in Case 1 and $m \geqslant 1$ in Case 3. And by the symbol K, we denote always an element of Π in Case 1 and of Π' in Case 3. We write moreover $K' = K \cap J_0$ ($K' = K$ in Case 1 and K' may be void in Case 3).

We prepare at first some identities required for our proof.

We define the transposed operator ${}^tA_K$ of A_K introduced in Lemma 4.3: If $|K|=1$, we put ${}^tA_K=0$, because $A_K=0$. If $|K| \geqslant 2$, we put as follows (See (4.6)):

$$\begin{aligned} {}^tA_K v(y) = \frac{1}{2} \sum_{p,q=1}^{k} (\delta_{pq} y_p - y_p y_q) \frac{\partial^2 v}{\partial y_p \partial y_q} + \sum_{p=1}^{k} \Big(e_p - y_p \sum_{q=1}^{k+1} e_q\Big) \frac{\partial v}{\partial y_p} \\ - \lambda_{|K'|} v, \quad \text{if} \quad K = \{1, \dots, k+1\}, \end{aligned} \tag{A.2}$$

where e_p is b_p or 1 according as $p \in J_+$ or as $p \in K'$. Then, we have the identity of Green:

$$\begin{aligned} 2\int_{\Omega_K} \{A_K u(y) \cdot v(y) - u(y)\, {}^tA_K v(y)\}\, dS_K(y) \\ = \sum_{j \in K'} \int_{\Omega_{K \setminus \{j\}}} u(y) v(y)\, dS_{K \setminus \{j\}}(y), \end{aligned} \tag{A.3}$$

where the right hand side should be read 0 if $|K|=1$ or if K' is void.

On the other hand, let $F_{p,K}(x, y)$ be the functions defined by (5.5). Then, we have

$$AF_{p,K}(x, y)+\lambda_{p+|K'|}F_{p,K}(x, y)=\frac{1}{2}F_{p-1,K}(x, y), \tag{A.4}$$

$$^tA_KF_{p,K}(x, y)+\lambda_{p+|K'|}F_{p,K}(x, y)=\frac{1}{2}(\sum_{j\in K}x_j)F_{p-1,K}(x, y), \tag{A.5}$$

where A and ^tA_K are operated in x and in y respectively (we put $F_{p,K}=0$ if $p<0$). Therefore if we operate A and ^tA_K to $E_{m,K}(x, y)$, we have by (5.7)

$$AE_{m,K}(x, y)+\lambda_mE_{m,K}(x, y)=0, \quad \text{and} \tag{A.6}$$

$$^tA_KE_{m,K}(x, y)+\lambda_mE_{m,K}(x, y)=\frac{1}{2}(\sum_{j\in K}x_j-1)R_{m,K}(x, y), \tag{A.7}$$

where $R_{m,K}(x, y)$ is defined by

$$\begin{cases} R_{m,K}(x, y)=0 \text{ if } m\leqslant|K'|,\ K'=\phi \text{ or if } |K|=1, \text{ and} \\ \dfrac{R_{m,K}(x, y)}{2m+\kappa}=\sum_{q=0}^{m-1-|K'|}\dfrac{(2m-q+\kappa-1)!}{(-1)^q q!}F_{m-1-q-|K'|,K}(x, y), \quad \text{if otherwise.} \end{cases} \tag{A.8}$$

Now, let us prove (i) and (ii) of lemma. (i) is obvious because of (5.8) and (A.6). To see (ii), we apply (A.3) to $u(y)=f(y)$ and $v(y)=E_{m,K}(x, y)$ for each K and sum up with respect to K. Then we have, by (5.8), (A.3) and (A.7),

$$\begin{aligned} 2(E_mAf+\lambda_mE_mf)(x) &= \sum_K(\sum_{j\in K}x_j-1)R_{m,K}f(x) \\ &\quad +\sum_K(\sum_{j\notin K}x_j)R_{m,K}f(x) \\ &= (\sum_{j=1}^d x_j-1)\sum_K R_{m,K}f(x)=0, \end{aligned}$$

where we used the fact that, if $j\notin K$, $E_{m,K\cup\{j\}}(x, y)=x_jR_{m,K}(x, y)$ for $(x, y)\in\bar{\Omega}\times\bar{\Omega}_K$. This proves (ii).

Next, let us prove (iii). By (i) and (ii) just verified, E_mf belongs to $\boldsymbol{E}_m$ for any $f\in C^0(\bar{\Omega})$ and $E_mf=0$ if f is a polynomial of degree $<m$. Hence it suffices to show that $E_mx^{\tilde{\gamma}}-x^{\tilde{\gamma}}$ is of degree $<m$ for any $\tilde{\gamma}\in\boldsymbol{N}^d$ with $|\tilde{\gamma}|=m$. Let us put

$$\begin{cases} E^0_{m,K}f(x)=(2m+\kappa)!\displaystyle\int_{\Omega_K}f(y)F_{m-|K'|,K}(x, y)dS_K(y), \\ \text{and} \quad E^0_mf(x)=\displaystyle\sum_K E^0_{m,K}f(x). \end{cases} \tag{A.9}$$

The proof of (iii) is now reduced to verify that

$$E^0_mx^{\tilde{\gamma}}-x^{\tilde{\gamma}} \text{ is of degree } <m \text{ if } \tilde{\gamma}\in\boldsymbol{N}^d \text{ with } |\tilde{\gamma}|=m. \tag{A.10}$$

We denote by Supp $\tilde{\alpha}$ the set of p for which $\alpha_p>0$. And put

$$H=\operatorname{Supp}\tilde{\gamma}, \quad H_0=H\cap J_0 \quad \text{and} \quad H_1=H^c\cap J_0 \tag{A.11}$$

If K does not contain H, $E^0_{m,K}x^{\tilde{\gamma}}$ is 0 because $x^{\tilde{\gamma}}=0$ on $\bar{\Omega}_K$.
Therefore, the summation with respect to K in (A.9) can be restricted only to K's

containing H. And, if $K \supset H$, we have

$$E^0_{m,K} x^{\tilde{\gamma}} = \sum_{\tilde{\alpha}} \left\{ \prod_{p \in J_+} \frac{(\alpha_p + \gamma_p + 2b_p - 1)!}{\alpha_p!(\alpha_p + 2b_p - 1)!} x_p^{\alpha_p + 1} \right\} \left\{ \prod_{q \in K'} \frac{(\alpha_q + \gamma_q)!}{\alpha_q!(\alpha_q + 1)!} x_q^{\alpha_q} \right\},$$

and the summation on $\tilde{\alpha}$ is extended over all $\tilde{\alpha}$'s with $\mathrm{Supp}\, \tilde{\alpha} \subset K$ and $|\tilde{\alpha}| = m - |K'|$. Let us introduce d indeterminates $z_1, \ldots, z_d$, and rewrite the sum on the right hand side.

At first, if $K = H_0 \cup J_+$, we see that $E^0_{m,K} x^{\tilde{\gamma}}$ is equal to the value of the following polynomial at $z_1 = x_1, \ldots, z_d = x_d$:

$$\left\{ \prod_{j \in H_0} \left(\frac{\partial}{\partial z_j} \right)^{\gamma_j - 1} \right\} \left\{ \prod_{j \in H_1} \left(\frac{\partial}{\partial z_j} \right)^{\gamma_j} \right\} \left[\left\{ \prod_{j \in J_+} \left(\frac{z_j}{x_j} \right)^{2b_j - 1} \right\} \frac{z^{\tilde{\gamma}} \langle z \rangle^{m - |K'|}}{(m - |K'|)!} \right], \tag{A.12}$$

$$\text{where } \langle z \rangle = \sum_{j=1}^{d} z_j.$$

Therefore, $E^0_{m,K} x^{\tilde{\gamma}} - x^{\tilde{\gamma}}$ is of degree $< m$.

Next, suppose that $K = K_1 \cup H_0 \cup J_+$ with $K_1 = K \cap H_1 \neq \phi$. Then, $E^0_{m,K} x^{\tilde{\gamma}}$ is equal to the value (at $z_1 = x_1, \ldots, z_d = x_d$) of an analogous polynomial to (A.12) but the factor $\langle z \rangle^{m - |K'|}$ must be integrated with respect to z_p from 0 to z_p once for each $p \in K_1$. Multiplying other factors, differentiating as above and putting $z_j = x_j$ $(1 \leqslant j \leqslant d)$, we see that $E^0_{m,K} x^{\tilde{\gamma}}$ is equal to $C x^{\tilde{\gamma}}$ plus terms of degree $< m$. Here C is a constant which is a multiple of the sum of $(-1)^{|L|}$, where the summation is extended over all the subsets L of K_1. Hence, C is a multiple of $(1-1)^{|K_1|}$, that is, $C = 0$. After all, in the sum $E^0_m x^{\tilde{\gamma}} = \sum_K E^0_{m,K} x^{\tilde{\gamma}}$, there is only one term $x^{\tilde{\gamma}}$ of degree m come from $K = H_0 \cup J_+$. This implies (A.10). And the proof of (iii) is complete.

§ B. Proof of Lemma 5.2

Let $V_\nu(t, z)$ be the function defined by (5.19). We are going to prove the following Lemma B. The Lemma 5.2 follows at once from this by virtue of the simplified formulas (5.21), (5.23) and (5.24) ($F_{p,K}(x, y)$'s are non-negative in any case if $(x, y) \in \bar{\Omega} \times \bar{\Omega}_K$).

Lemma B. *$V_\nu(t, z)$ is positive if $t > 0$, $-1 \leqslant z \leqslant 1$ and if $\nu > -1$.*

Proof. (The definition by (5.19) of $V_\nu(t, z)$ is not rigorous for $-1 < \nu \leqslant 0$, because $C^0_q(z)$'s are not well-defined. We will do it over again in what follows).

Suppose at first that $\nu > 0$.

As is known (See § 10.9 of [2]), $V_\nu(t, z)$ satisfies the differential equation

$$8 \frac{\partial V_\nu}{\partial t} = (1 - z^2) \frac{\partial^2 V_\nu}{\partial z^2} - (2\nu + 1) z \frac{\partial V_\nu}{\partial z} \tag{B.1}$$

in the open set $D = \{(t, z);\ t > 0 \text{ and } -1 < z < 1\}$. $V_\nu(t, z)$ is an even function of z. $V_\nu(t, \pm 1)$ is positive, because $C^\nu_{2m}(1)$ is positive. And by an asymptotic expansion as $t \to +0$ (See (F.7), p. 315 of [10]), $V_\nu(+0, z) = 0$ if $-1 < z < 1$. That is, $V_\nu(t, z)$ is positive on the lateral part of ∂D and is 0 on $\partial D \cap \{t=0\}$. Therefore, by the maximum principle applied to (B.1), $V_\nu(t, z)$ is positive if $t>0$ and $-1\leqslant z\leqslant 1$.

Let us derive further properties of $V_\nu(t, z)$ as $\nu>0$. By the recurrence formula $\frac{d}{dz} C_q^\nu(z)=2\nu C_{q-1}^{\nu+1}(z)$ $(C_{-1}^{\nu+1}(z)=0)$,

$$\left(\frac{\partial}{\partial z}\right)^2 V_\nu(t, z)=4e^{-t(\nu+1)/2}V_{\nu+2}(t, z). \tag{B.2}$$

Put $v_\nu(t)=V_\nu(t, 0)$ for simplicity. Then by (B.1) and (B.2),

$$2v_\nu'(t)=e^{-t(\nu+1)/2}v_{\nu+2}(t). \tag{B.3}$$

Integrating the both sides in the interval $(t, +\infty)$ or in $(0, t)$,

$$2v_\nu(t)=2\Gamma(\nu+1)-\int_t^{+\infty} e^{-s(\nu+1)/2}v_{\nu+2}(s)ds \tag{B.4}$$

$$=\int_0^t e^{-s(\nu+1)/2}v_{\nu+2}(s)ds, \tag{B.4'}$$

because $v_\nu(+\infty)=\Gamma(\nu+1)$ and $v_\nu(+0)=0$. More explicitly, we can develop $v_\nu(t)$ in series:

$$v_\nu(t)=\sum_{m=0}^{\infty}\frac{(2m+\nu)\Gamma(m+\nu)}{(-1)^m m!}e^{-tm(m+\nu)/2}. \tag{B.5}$$

And by (B.2), we have Taylor expansion of $V_\nu(t, z)$ at $z=0$:

$$V_\nu(t, z)=\sum_{m=0}^{\infty}\frac{(2z)^{2m}}{(2m)!}v_{\nu+2m}(t)e^{-tm(m+\nu)/2}. \tag{B.6}$$

The right hand side of (B.5) is holomorphic in ν in the half-plane $\operatorname{Re}\nu>-1$, if $t>0$. Hence, $v_\nu(t)$ can be holomorphically extended to this half-plane by means of (B.5). And by (B.4) with $t=+0$, $v_\nu(+0)$ is also extended holomorphically to $\operatorname{Re}\nu>-1$. But this is 0 if ν is real and positive. Thus $v_\nu(+0)=0$ if $\operatorname{Re}\nu>-1$. By this and by (B.3), (B.4') is valid also in this half-plane.

Consequently, defined anew by (B.6), $V_\nu(t, z)$ is revealed to be holomorphic function of ν in the half-plane $\operatorname{Re}\nu>-1$, if $t>0$ and $-1\leqslant z\leqslant 1$.

Suppose now that $\nu>-1$. By (B.4'), $v_\nu(t)$ is positive if $t>0$, because the right hand side is already seen to be positive. Therefore by (B.6), $V_\nu(t, z)$ is also positive if $t>0$ and $-1\leqslant z\leqslant 1$. This completes the proof of Lemma B.

DEPARTMENT OF MATHEMATICS
KYOTO UNIVERSITY

References

[1] J. Crow and M. Kimura, Some genetic problems in natural populations, Proc. 3rd Berkeley Sympos. Math. Statist. and Probab. (1956) Vol. IV, 1–22.

[2] A. Erdelyi et al., Higher Transcendental Functions, Vol. 2 McGraw-Hill.
[3] S. Ethier, An error estimate for the diffusion approximation in population genetics, Ph. D. Thesis, Wisconsin Univ. (1975).
[4] W. Feller, Diffusion processes in genetics, Proc. 2nd Berkeley Sympos. Math. Statist. and Probab. (1951), 227–246.
[5] A. Friedman, Stochastic Differential Equations and Applications, Vol. 1 and 2, Academic Press.
[6] S. Karlin, A First Course in Stochastic Processes, Academic Press.
[7] T. Maruyama, Stochastic Problems in Population Genetics, Lecture Notes in Biomathematics **17** (1977), Springer.
[8] K. Sato, Diffusion processes and a class of Markov chains related to population genetics, Osaka J. Math. **13** (1976), 631–659.
[9] K. Sato, Diffusion operators in population genetics and convergence of Markov chains, Lecture Notes in Mathematics 695 (1978), 127–137, Springer.
[10] N. Shimakura, Equations différentielles provenant de la génétique des populations, Tohoku Math. J. **29** (1977), 287–318.
[11] E. Fackerel and R. Littler, Transition densities for neutral multi-allele diffusion models, Biometrics. **31** (1975), 117–123.