

A General Solution of the Wright–Fisher Model of Random Genetic Drift

Tat Dat Tran¹ · Julian Hofrichter¹ · Jürgen Jost^{1,2,3} 

Published online: 16 March 2016

© Foundation for Scientific Research and Technological Innovation 2016

Abstract We introduce a general solution concept for the Fokker–Planck (Kolmogorov) equation representing the diffusion limit of the Wright–Fisher model of random genetic drift for an arbitrary number of alleles at a single locus. This solution will continue beyond the transitions from the loss of alleles, that is, it will naturally extend to the boundary strata of the probability simplex on which the diffusion is defined. This also takes care of the degeneracy of the diffusion operator at the boundary. We shall then show the existence and uniqueness of a solution. From this solution, we can readily deduce information about the evolution of a Wright–Fisher population.

Keywords Random genetic drift · Fokker–Planck equation · Wright–Fisher model · Several alleles

The research leading to these results has received funding from the European Research Council under the European Union’s Seventh Framework Programme (FP7/2007-2013)/ERC Grant agreement no. 267087. T. D. Tran and J. Hofrichter have also been supported by the IMPRS “Mathematics in the Sciences”, and the material presented in this paper is largely based on the first author’s thesis.

✉ Jürgen Jost
jost@mis.mpg.de

Tat Dat Tran
trandat@mis.mpg.de

Julian Hofrichter
julian.hofrichter@mis.mpg.de

¹ Max Planck Institute for Mathematics in the Sciences, 04103 Leipzig, Germany

² Department of Mathematics, Leipzig University, 04081 Leipzig, Germany

³ Santa Fe Institute for the Sciences of Complexity, Santa Fe, NM 87501, USA

Introduction

The random genetic drift model developed implicitly by Fisher in [11] and explicitly by Wright in [25], and henceforth called the Wright–Fisher model, is one of the most popular stochastic models in population genetics [2, 9]. In its simplest form, it is concerned with the evolution of the probabilities between non-overlapping generations in a population of fixed size of two alleles at a single diploid locus that are obtained from random sampling in the parental generation, without additional biological mechanisms like mutation, selection, or a spatial population structure. Generalizations to multiple alleles, several loci, inclusion of mutations and selection etc. then constituted an important part of mathematical population genetics. It is our aim to develop a general mathematical perspective on the Wright–Fisher model and its generalizations. In the present paper, we treat the case of multiple alleles at a single site. In a companion paper [23], we have discussed the simplest case of two alleles in more detail. Generalizations will be addressed in subsequent papers.

Let us first describe the basic mathematical contributions of Wright and Kimura. In 1945, Wright approximated the discrete process by a diffusion process that is continuous in space and time (continuous process, for short) and that can be described by a Fokker–Planck equation. In 1955, by solving this Fokker–Planck equation derived from the Wright–Fisher model, Kimura obtained an exact solution for the Wright–Fisher model in the case of two alleles (see [15]). Kimura [16] also developed an approximation for the solution of the Wright–Fisher model in the multi-allele case, and in 1956, he obtained ([17]) an exact solution of this model for three alleles and concluded that this can be generalized to arbitrarily many alleles. This yields more information about the Wright–Fisher model as well as the corresponding continuous process. Kimura’s solution, however, is not entirely satisfactory. For one thing, it depends on very clever algebraic manipulations so that the general mathematical structure is not very transparent, and this makes generalizations very difficult. Also, Kimura’s approach is local in the sense that it does not naturally incorporate the transitions resulting from the (irreversible) loss of one or more alleles in the population. Therefore, for instance the integral of his probability density function on its defined domain is not equal to 1.

As mentioned, while the original model of Wright and Fisher works with a finite population in discrete time, many mathematical insights into its behavior are derived from its diffusion approximation. After the original work of Wright and Kimura just described, a more systematic approach was developed within the theory of strongly-continuous semigroups and Markov processes. In this framework, the diffusion approximation for the multi-allele Wright–Fisher model was derived by Ethier and Nagylaki [6, 7], and a proof of convergence of the Markov chain to the diffusion process can be found in [5]. (In this paper, we are not concerned with deriving the diffusion approximation, but actually, this can be derived in a rather direct manner without having to appeal to the general theory, as we shall show elsewhere.) One may then derive existence and uniqueness results for solutions of the Fokker–Planck equation from the theory of strongly continuous semigroups [5, 6, 14]. As the diffusion operator of the diffusion approximation becomes degenerate at the boundary, the analysis at the boundary becomes difficult, and this issue is not addressed by the aforementioned results. Recent work of Epstein and Mazzeo [3, 4], however, treats the boundary regularity with general PDE methods.

The full structure of the Wright–Fisher model and its diffusion approximation, however, is only revealed when one can connect the dynamics before and after the loss of an allele, or in analytic terms, if one can extend the process from the interior of the probability simplex to all its boundary strata. In particular, this is needed to preserve the normalization of the

probability distribution. Therefore, in this paper, we develop the definition of a general solution that naturally includes the transitions resulting from the disappearance of alleles and derive the formalism for its solution. Since this formalism is rather explicit, it will allow us to derive and generalize the known explicit formulas for the quantities associated with the Wright–Fisher diffusion model like expected waiting times for the loss of one or several alleles in a systematic manner. The key for our approach are evolution equations for the moments of the probability density and the duality between the forward and backward Kolmogorov equations. We show that there exists a unique global solution of the Fokker–Planck equation. Since, as explained, our concept of a solution is different from (and, as we believe, better adapted to the structure of the Wright–Fisher model than) those treated in the literature, insofar as it extends to the boundary, these results do not follow from the general results of the literature mentioned above.

In the present paper, we only treat genetic drift in the absence of mutation, selection, and recombination. Extensions that can be obtained on the basis of the formalism presented here, in particular to general recombination schemes, will be presented in subsequent publications.

The Global Solution of the Wright-Fisher Model

In this section, we shall first establish some notation, and then prove some propositions as well as the main theorem of this paper.

Notations

$\Delta_n := \{(x^1, x^2, \dots, x^{n+1}) : x^i \geq 0, \sum_{i=1}^{n+1} x^i = 1\}$ is the standard n -simplex in \mathbb{R}^{n+1} representing the probabilities or relative frequencies of alleles A_1, \dots, A_{n+1} in our population. Often, however, it is advantageous to work in \mathbb{R}^n instead of \mathbb{R}^{n+1} , and with $e_0 := (0, \dots, 0) \in \mathbb{R}^n$, $e_k := (0, \dots, \underbrace{1}_{k\text{th}}, \dots, 0) \in \mathbb{R}^n$, we therefore define

$$\Omega_n := \text{intco}\{e_0, \dots, e_n\} := \left\{ \sum_{k=0}^n x^k e_k, (x, x^0) = \left(x^1, \dots, x^n, 1 - \sum_{k=1}^n x^k \right) \in \text{int}\Delta_n \right\}.$$

Moreover, we shall need the subsimplices corresponding to subsets of alleles, using the following notations

$$\begin{aligned} I_k &:= \{i_0, \dots, i_k\}, 0 \leq i_0 < \dots < i_k \leq n, \quad k \in \{1, \dots, n\}, \\ V_0 &:= \{e_0, \dots, e_n\}, \\ &\quad \text{the domain representing a population of one allele,} \\ V_k^{(i_0, \dots, i_k)} &:= \text{intco}\{e_{i_0}, \dots, e_{i_k}\}, \quad k \in \{1, \dots, n\}, \\ &\quad \text{the domain representing a population of alleles } \{A_{i_0}, \dots, A_{i_k}\}, \\ V_k &:= \{\text{intco}\{e_{i_0}, \dots, e_{i_k}\} \text{ for some } i_0 < \dots < i_k \in \overline{0, n}\}, \quad k \in \{1, \dots, n\}, \\ &= \bigsqcup_{(i_0, \dots, i_k) \in I_k} V_k^{(i_0, \dots, i_k)}, \\ &\quad \text{the domain representing a population of } (k + 1) \text{ alleles,} \end{aligned}$$

$$\begin{aligned} \bar{V}_k &:= \bigcup_{(i_0, \dots, i_k) \in I_k} \bar{V}_k^{(i_0, \dots, i_k)}, \quad k \in \{1, \dots, n\}, \\ &= \bigsqcup_{i=0}^k V_i, \end{aligned}$$

the domain representing a population of at most $(k + 1)$ alleles.

We shall also need some function spaces:

$$\begin{aligned} H_k^{(i_0, \dots, i_k)} &:= C^\infty \left(\bar{V}_k^{(i_0, \dots, i_k)} \right), \\ H_k &:= C^\infty(\bar{V}_k), \quad k \in \{1, \dots, n\}, \\ H &:= \{f : \bar{V}_n \rightarrow [0, \infty] \text{ measurable such that } [f, g]_n < \infty, \forall g \in H_n\}, \end{aligned}$$

$$\begin{aligned} \text{where } [f, g]_n &:= \int_{\bar{V}_n} f(x)g(x)d\mu(x) = \sum_{k=0}^n \int_{V_k} f(x)g(x)d\mu_k(x), \\ &= \sum_{k=0}^n \sum_{(i_0, \dots, i_k) \in I_k} \int_{V_k^{(i_0, \dots, i_k)}} f(x)g(x)d\mu_k^{(i_0, \dots, i_k)}(x), \end{aligned}$$

with $\mu_k^{(i_0, \dots, i_k)}$ a probability measure on $V_k^{(i_0, \dots, i_k)}$.

We can now define the differential operators for our Fokker–Planck equation:

$$\begin{aligned} L_k^{(i_0, \dots, i_k)} : H_k^{(i_0, \dots, i_k)} &\rightarrow H_k^{(i_0, \dots, i_k)}, \quad L_k^{(i_0, \dots, i_k)} f(x) = \frac{1}{2} \sum_{i, j \in \{i_1, \dots, i_k\}} \frac{\partial^2 (a_{ij}(x)f(x))}{\partial x^i \partial x^j}, \\ (L_k^{(i_0, \dots, i_k)})^* : H_k^{(i_0, \dots, i_k)} &\rightarrow H_k^{(i_0, \dots, i_k)}, \quad (L_k^{(i_0, \dots, i_k)})^* g(x) = \frac{1}{2} \sum_{i, j \in \{i_1, \dots, i_k\}} a_{ij}(x) \frac{\partial^2 g(x)}{\partial x^i \partial x^j}, \\ L_k : H_k &\rightarrow H_k, \quad (L_k)|_{H_k^{(i_0, \dots, i_k)}} = L_k^{(i_0, \dots, i_k)}, \\ L_k^* : H_k &\rightarrow H_k, \quad (L_k^*)|_{H_k^{(i_0, \dots, i_k)}} = (L_k^{(i_0, \dots, i_k)})^*, \end{aligned}$$

where the coefficients are defined by

$$a_{ij}(x) := x^i (\delta_{ij} - x^j), \quad i, j \in \{1, \dots, n\}.$$

Finally, we shall need

$$w_k^{(i_0, \dots, i_k)}(x) := \prod_{i \in I_k^{(i_0, \dots, i_k)}} x^i, \quad k \in \{1, \dots, n\}.$$

Proposition 2.1 *For each $1 \leq k \leq n$, $m \geq 0$, $|\alpha| = \alpha^1 + \dots + \alpha^k = m$, the polynomial of degree m in k variables $x = (x^{i_1}, \dots, x^{i_k})$ in $V_k^{(i_0, \dots, i_k)}$*

$$X_{m, \alpha}^{(k)}(x) = x^\alpha + \sum_{|\beta| < m} a_{m, \beta}^{(k)} x^\beta, \tag{1}$$

where the $a_{m,\beta}^{(k)}$ are inductively defined by

$$a_{m,\beta}^{(k)} = -\frac{\sum_{i=1}^k (\beta_i + 2)(\beta_i + 1)a_{m,\beta+e_i}^{(k)}}{(m - |\beta|)(m + \beta + 2k + 1)}, \quad \forall |\beta| < m,$$

is the eigenvector of $L_k^{(i_0, \dots, i_k)}$ corresponding to the eigenvalue $\lambda_m^{(k)} = \frac{(m+k)(m+k+1)}{2}$.

Proof We have

$$\begin{aligned} L_k^{(i_0, \dots, i_k)} X_{m,\alpha}^{(k)}(x) &= \frac{1}{2} \sum_{i \in \{i_1, \dots, i_k\}} \frac{\partial^2}{(\partial x^i)^2} \left[x^i (1 - x^i) \left(x^\alpha + \sum_{|\beta| < m} a_{m,\beta}^{(k)} x^\beta \right) \right] \\ &\quad - \sum_{i \neq j \in \{i_1, \dots, i_k\}} \frac{\partial^2}{\partial x^i \partial x^j} \left[x^i x^j \left(x^\alpha + \sum_{|\beta| < m} a_{m,\beta}^{(k)} x^\beta \right) \right] \\ &= \frac{1}{2} \sum_{i \in \{i_1, \dots, i_k\}} \frac{\partial^2}{(\partial x^i)^2} \left[x^{\alpha+e_i} - x^{\alpha+2e_i} + \sum_{|\beta| < m} a_{m,\beta}^{(k)} x^{\beta+e_i} \right. \\ &\quad \left. - \sum_{|\beta| < m} a_{m,\beta}^{(k)} x^{\beta+2e_i} \right] \\ &\quad - \sum_{i \neq j \in \{i_1, \dots, i_k\}} \frac{\partial^2}{\partial x^i \partial x^j} \left[x^{\alpha+e_i+e_j} + \sum_{|\beta| < m} a_{m,\beta}^{(k)} x^{\beta+e_i+e_j} \right] \\ &= \frac{1}{2} \sum_i \left[(\alpha^i + 1)\alpha^i x^{\alpha-e_i} - (\alpha^i + 2)(\alpha^i + 1)x^\alpha \right. \\ &\quad \left. + \sum_{|\beta| < m} a_{m,\beta}^{(k)} (\beta^i + 1)\beta^i x^{\beta-e_i} - \sum_{|\beta| < m} a_{m,\beta}^{(k)} (\beta^i + 2)(\beta^i + 1)x^\beta \right] \\ &\quad - \sum_{i \neq j} \left[(\alpha^i + 1)(\alpha^j + 1)x^\alpha + \sum_{|\beta| < m} a_{m,\beta}^{(k)} (\beta^i + 1)(\beta^j + 1)x^\beta \right] \\ &= \left[-\frac{1}{2} \sum_i (\alpha^i + 2)(\alpha^i + 1) - \sum_{i \neq j} (\alpha^i + 1)(\alpha^j + 1) \right] x^\alpha \\ &\quad + \text{terms of lower degree} \\ &= \left[-\frac{1}{2} \left(\sum_i \alpha^i + k \right) \left(\sum_i \alpha^i + k + 1 \right) \right] x^\alpha + \text{terms of lower degree} \\ &= -\frac{(m+k)(m+k+1)}{2} x^\alpha + \text{terms of lower degree.} \end{aligned}$$

By equalizing coefficients we obtain

$$\lambda_m^{(k)} = \frac{(m+k)(m+k+1)}{2}$$

and

$$a_{m,\beta}^{(k)} = -\frac{\sum_{i=1}^k (\beta_i + 2)(\beta_i + 1)a_{m,\beta+e_i}^{(k)}}{(m - |\beta|)(m + \beta + 2k + 1)}, \quad \forall |\beta| < m.$$

This completes the proof. □

Remark 2.2 • When $k = 1$, $X_{m,m}^{(1)}(x^1)$ is the m th–Gegenbauer polynomial (up to a constant). Thus, the polynomials $X_{m,\alpha}^{(k)}(x)$ can be understood as a generalization of the Gegenbauer polynomials to higher dimensions.

- Because of this representation of eigenvectors, we can easily see that $X_{m,\alpha}^{(n)}(x)$ is a basis of $C^2(\overline{V_n})$.

Proposition 2.3 *If $X \in C^2(\overline{V_k^{(i_0, \dots, i_k)}})$ is an eigenvector of $L_k^{(i_0, \dots, i_k)}$ corresponding to λ , then $w_k^{(i_0, \dots, i_k)} X$ is an eigenvector of $(L_k^{(i_0, \dots, i_k)})^*$ corresponding to λ .*

Proof If $X \in C^2(\overline{V_k^{(i_0, \dots, i_k)}})$ is an eigenvector of $L_k^{(i_0, \dots, i_k)}$ corresponding to λ , it follows that

$$\begin{aligned} -\lambda(w_k^{(i_0, \dots, i_k)}(x)X) &= \frac{1}{2}w_k^{(i_0, \dots, i_k)}(x) \sum_{i,j \in \{i_1, \dots, i_k\}} \frac{\partial^2}{\partial x^i \partial x^j} \left(x^i (\delta_{ij} - x^j) X \right) \\ &= \frac{1}{2}w_k^{(i_0, \dots, i_k)}(x) \sum_{i,j \in \{i_1, \dots, i_k\}} \left(x^i (\delta_{ij} - x^j) \right) \frac{\partial^2 X}{\partial x^i \partial x^j} \\ &\quad + \frac{1}{2}w_k^{(i_0, \dots, i_k)}(x) \sum_{i,j \in \{i_1, \dots, i_k\}} \frac{\partial (x^i (\delta_{ij} - x^j))}{\partial x^i} \frac{\partial X}{\partial x^j} \\ &\quad + \frac{1}{2}w_k^{(i_0, \dots, i_k)}(x) \sum_{i,j=1}^k \frac{\partial (x^i (\delta_{ij} - x^j))}{\partial x^j} \frac{\partial X}{\partial x^i} \\ &\quad + \frac{1}{2}w_k^{(i_0, \dots, i_k)}(x) \sum_{i,j \in \{i_1, \dots, i_k\}} \frac{\partial^2 (x^i (\delta_{ij} - x^j))}{\partial x^i \partial x^j} X \\ &= \frac{1}{2} \sum_{i,j=1}^k \left(x^i (\delta_{ij} - x^j) \right) \left(w_k^{(i_0, \dots, i_k)}(x) \frac{\partial^2 X}{\partial x^i \partial x^j} \right) \\ &\quad + \frac{1}{2} \sum_{j \in \{i_1, \dots, i_k\}} w_k^{(i_0, \dots, i_k)}(x) \left(1 - (k-1)x^j \right) \frac{\partial X}{\partial x^j} \\ &\quad + \frac{1}{2} \sum_{i \in \{i_1, \dots, i_k\}} w_k^{(i_0, \dots, i_k)}(x) \left(1 - (k-1)x^i \right) \frac{\partial X}{\partial x^i} \\ &\quad - \frac{k(k+1)}{2} w_k^{(i_0, \dots, i_k)}(x) X \\ &= \frac{1}{2} \sum_{i,j \in \{i_1, \dots, i_k\}} \left(x^i (\delta_{ij} - x^j) \right) \left(w_k^{(i_0, \dots, i_k)}(x) \frac{\partial^2 X}{\partial x^i \partial x^j} \right) \\ &\quad + \frac{1}{2} \sum_{i,j \in \{i_1, \dots, i_k\}} \left(x^i (\delta_{ij} - x^j) \right) \frac{\partial w_k^{(i_0, \dots, i_k)}(x)}{\partial x^i} \frac{\partial X}{\partial x^j} \\ &\quad + \frac{1}{2} \sum_{i,j \in \{i_1, \dots, i_k\}} \left(x^i (\delta_{ij} - x^j) \right) \frac{\partial w_k^{(i_0, \dots, i_k)}(x)}{\partial x^j} \frac{\partial X}{\partial x^i} \\ &\quad + \frac{1}{2} \sum_{i,j \in \{i_1, \dots, i_k\}} \left(x^i (\delta_{ij} - x^j) \right) \frac{\partial^2 w_k^{(i_0, \dots, i_k)}(x)}{\partial x^i \partial x^j} X \end{aligned}$$

$$\begin{aligned}
 &= \frac{1}{2} \sum_{i,j \in \{i_1, \dots, i_k\}} (x^i (\delta_{ij} - x^j)) \frac{\partial^2 (w_k^{(i_0, \dots, i_k)} X)(x)}{\partial x^i \partial x^j} \\
 &= \left(L_k^{(i_0, \dots, i_k)} \right)^* \left(w_k^{(i_0, \dots, i_k)}(x) X \right).
 \end{aligned}$$

This completes the proof. □

Proposition 2.4 *Let v be the exterior unit normal vector of the domain $V_k^{(i_0, \dots, i_k)}$. Then we have*

$$\sum_{j \in \{i_1, \dots, i_k\}} a_{ij} v^j = 0 \quad \text{on } \partial V_k^{(i_0, \dots, i_k)}, \quad \forall i \in \{i_1, \dots, i_k\}. \tag{2}$$

Proof In fact, on the surface $(x^s = 0)$, for some $s \in \{i_1, \dots, i_k\}$ we have $v = -e_s$, and hence $\sum_{j \in \{i_1, \dots, i_k\}} a_{ij} v^j = a_{is} = x^s (\delta_{si} - x^i) = 0$. On the surface $(x^{i_0} = 0)$ we have $v = \frac{1}{\sqrt{k}}(e_{i_1} + \dots + e_{i_k})$, hence $\sum_{j \in \{i_1, \dots, i_k\}} a_{ij} v^j = \frac{1}{\sqrt{k}} \sum_{j \in \{i_1, \dots, i_k\}} a_{ij} = \frac{1}{\sqrt{k}} x^i x^{i_0} = 0$. This completes the proof. □

Proposition 2.5 *$L_k^{(i_0, \dots, i_k)}$ and $(L_k^{(i_0, \dots, i_k)})^*$ are weighted adjoints in $H_k^{(i_0, \dots, i_k)}$, i.e.*

$$\left(L_k^{(i_0, \dots, i_k)} X, w_k^{(i_0, \dots, i_k)} Y \right) = \left(X, \left(L_k^{(i_0, \dots, i_k)} \right)^* \left(w_k^{(i_0, \dots, i_k)} Y \right) \right), \quad \forall X, Y \in H_k^{(i_0, \dots, i_k)}.$$

Proof We put $F_i^{(k)}(x) := \sum_{j \in \{i_1, \dots, i_k\}} \frac{\partial (a_{ij}(x) X(x))}{\partial x^j}$. Because of $w_k^{(i_0, \dots, i_k)} Y \in C_0^\infty(\overline{V}_k^{(i_0, \dots, i_k)})$, the second Green formula, and Proposition 2.4, we have

$$\begin{aligned}
 \left(L_k^{(i_0, \dots, i_k)} X, w_k^{(i_0, \dots, i_k)} Y \right) &= \frac{1}{2} \sum_{i,j \in \{i_1, \dots, i_k\}} \int_{\overline{V}_k^{(i_0, \dots, i_k)}} \frac{\partial^2 (a_{ij}(x) X(x))}{\partial x^i \partial x^j} w_k^{(i_0, \dots, i_k)}(x) Y(x) dx \\
 &= \frac{1}{2} \sum_{i \in \{i_1, \dots, i_k\}} \int_{\overline{V}_k^{(i_0, \dots, i_k)}} \frac{\partial F_i^{(k)}(x)}{\partial x^i} w_k^{(i_0, \dots, i_k)}(x) Y(x) dx \\
 &= \frac{1}{2} \sum_{i \in \{i_1, \dots, i_k\}} \int_{\partial V_k^{(i_0, \dots, i_k)}} F_i^{(k)}(x) \nu_i w_k^{(i_0, \dots, i_k)}(x) Y(x) d\sigma(x) \\
 &\quad - \frac{1}{2} \sum_{i \in \{i_1, \dots, i_k\}} \int_{\overline{V}_k^{(i_0, \dots, i_k)}} F_i^{(k)}(x) \frac{\partial (w_k^{(i_0, \dots, i_k)}(x) Y(x))}{\partial x^i} dx \\
 &= -\frac{1}{2} \sum_{i \in \{i_1, \dots, i_k\}} \int_{\overline{V}_k^{(i_0, \dots, i_k)}} F_i^{(k)}(x) \frac{\partial (w_k^{(i_0, \dots, i_k)}(x) Y(x))}{\partial x^i} dx \\
 &= -\frac{1}{2} \sum_{i,j \in \{i_1, \dots, i_k\}} \int_{\overline{V}_k^{(i_0, \dots, i_k)}} \frac{\partial (a_{ij}(x) X(x))}{\partial x^j} \frac{\partial (w_k^{(i_0, \dots, i_k)}(x) Y(x))}{\partial x^i} dx
 \end{aligned}$$

$$\begin{aligned}
 &= -\frac{1}{2} \sum_{i,j \in \{i_1, \dots, i_k\}} \int_{\partial V_k^{(i_0, \dots, i_k)}} a_{ij}(x) v_j X(x) \frac{\partial (w_k^{(i_0, \dots, i_k)}(x) Y(x))}{\partial x^i} d\sigma(x) \\
 &\quad + \left(X, L_k^*(w_k^{(i_0, \dots, i_k)} Y) \right) \\
 &= \left(X, L_k^*(w_k Y) \right).
 \end{aligned}$$

□

Proposition 2.6 In $\overline{V}_k^{(i_0, \dots, i_k)}$, $\{X_{m,\alpha}^{(k)}\}_{m \geq 0, |\alpha|=m}$ is a basis of $H_k^{(i_0, \dots, i_k)}$ which is orthogonal with respect to the weights $w_k^{(i_0, \dots, i_k)}$, i.e.,

$$\left(X_{m,\alpha}^{(k)}, w_k^{(i_0, \dots, i_k)} X_{j,\beta}^{(k)} \right) = 0, \quad \forall j \neq m, |\alpha| = m, |\beta| = j.$$

Proof $\{X_{m,\alpha}^{(k)}\}_{m \geq 0, |\alpha|=m}$ is a basis of $H_k^{(i_0, \dots, i_k)}$ because $\{x^\alpha\}_\alpha$ is a basis of this space. To prove the orthogonality we apply the Propositions 2.1, 2.3, 2.7 as follows

$$\begin{aligned}
 -\lambda_m^{(k)} \left(X_{m,\alpha}^{(k)}, w_k^{(i_0, \dots, i_k)} X_{j,\beta}^{(k)} \right) &= \left(L_k^{(i_0, \dots, i_k)} X_{m,\alpha}^{(k)}, w_k^{(i_0, \dots, i_k)} X_{j,\beta}^{(k)} \right) \\
 &= \left(X_{m,\alpha}^{(k)}, (L_k^{(i_0, \dots, i_k)})^* (w_k^{(i_0, \dots, i_k)} X_{j,\beta}^{(k)}) \right) \\
 &= -\lambda_j^{(k)} \left(X_{m,\alpha}^{(k)}, w_k^{(i_0, \dots, i_k)} X_{j,\beta}^{(k)} \right).
 \end{aligned}$$

Because $\lambda_m^{(k)} \neq \lambda_j^{(k)}$, this finishes the proof. □

Proposition 2.7 (i) The spectrum of the operator $L_k^{(i_0, \dots, i_k)}$ is

$$\text{Spec}(L_k^{(i_0, \dots, i_k)}) = \bigcup_{m \geq 0} \left\{ \lambda_m^{(k)} = \frac{(m+k)(m+k+1)}{2} \right\} =: \Lambda_k$$

and the eigenvectors of $L_k^{(i_0, \dots, i_k)}$ corresponding to $\lambda_m^{(k)}$ are of the form

$$X = \sum_{|\alpha|=m} d_{m,\alpha}^{(k)} X_{m,\alpha}^{(k)},$$

i.e., the eigenspace corresponding to $\lambda_m^{(k)}$ is of dimension $\binom{k+m-1}{k-1}$;

(ii) The spectrum of the operator L_k is the same.

Proof (i) Proposition 2.1 implies that $\Lambda_k \subseteq \text{Spec}(L_k^{(i_0, \dots, i_k)})$. Conversely, for $\lambda \notin \Lambda_k$, we will prove that λ is not an eigenvalue of $L_k^{(i_0, \dots, i_k)}$. In fact, assume that $X \in H_k^{(i_0, \dots, i_k)}$ such that $L_k^{(i_0, \dots, i_k)} X = -\lambda X$ in $H_k^{(i_0, \dots, i_k)}$. Because $\{X_{m,\alpha}^{(k)}\}_{m,\alpha}$ is an orthogonal basis of $H_k^{(i_0, \dots, i_k)}$ with respect to the weights $w_k^{(i_0, \dots, i_k)}$ (Proposition 2.4), we can represent X by

$$X = \sum_{m=0}^{\infty} \sum_{|\alpha|=m} d_{m,\alpha}^{(k)} X_{m,\alpha}^{(k)}.$$

It follows that

$$\begin{aligned}
 \sum_{m=0}^{\infty} \sum_{|\alpha|=m} d_{m,\alpha}^{(k)} (-\lambda_m^{(k)}) X_{m,\alpha}^{(k)} &= \sum_{m=0}^{\infty} \sum_{|\alpha|=m} d_{m,\alpha}^{(k)} L_k^{(i_0, \dots, i_k)} X_{m,\alpha}^{(k)} \\
 &= L_k^{(i_0, \dots, i_k)} X \\
 &= -\lambda \sum_{m=0}^{\infty} \sum_{|\alpha|=m} d_{m,\alpha}^{(k)} X_{m,\alpha}^{(k)}.
 \end{aligned}$$

For any $j \geq 0, |\beta| = j$, multiplying by $w_k X_{j,\beta}^{(k)}$ and then integrating on \bar{V}_n we have

$$\begin{aligned} & \sum_{|\alpha|=j} d_{j,\alpha}^{(k)} \lambda_j^{(k)} \left(X_{j,\alpha}^{(k)}, w_k^{(i_0, \dots, i_k)} X_{j,\beta}^{(k)} \right) \\ &= \sum_{|\alpha|=j} d_{j,\alpha}^{(k)} \lambda \left(X_{j,\alpha}^{(k)}, w_k^{(i_0, \dots, i_k)} X_{j,\beta}^{(k)} \right), \quad \forall j \geq 0, |\beta| = j, \\ &\Rightarrow \left(X_{j,\alpha}^{(k)}, w_k^{(i_0, \dots, i_k)} X_{j,\beta}^{(k)} \right)_{\beta,\alpha} \left(d_{j,\alpha}^{(k)} \lambda_j^{(k)} \right)_\alpha \\ &= \left(X_{j,\alpha}^{(k)}, w_k^{(i_0, \dots, i_k)} X_{j,\beta}^{(k)} \right)_{\beta,\alpha} \left(d_{j,\alpha}^{(k)} \lambda \right)_\alpha, \quad \forall j \geq 0, |\beta| = j, \\ &\Rightarrow d_{j,\alpha}^{(k)} \lambda_j^{(k)} = d_{j,\alpha}^{(k)} \lambda, \quad \forall j \geq 0, |\beta| = j, \text{ because } \det \left(X_{j,\alpha}^{(k)}, w_k^{(i_0, \dots, i_k)} X_{j,\beta}^{(k)} \right)_{\beta,\alpha} \neq 0 \\ &\Rightarrow d_{j,\alpha}^{(k)} = 0, \quad \forall j \geq 0, |\alpha| = j, \text{ because } \lambda \neq \lambda_j^{(k)}. \end{aligned}$$

It follows that $X = 0$ in $H_k^{(i_0, \dots, i_k)}$. Therefore

$$\text{Spec}(L_k^{(i_0, \dots, i_k)}) = \bigcup_{m \geq 0} \left\{ \lambda_m^{(k)} = \frac{(m+k)(m+k+1)}{2} \right\} = \Lambda_k.$$

Moreover, assume that $X \in H_k^{(i_0, \dots, i_k)}$ is an eigenvector of $L_k^{(i_0, \dots, i_k)}$ corresponding to $\lambda_j^{(k)}$, i.e., $L_k^{(i_0, \dots, i_k)} X = -\lambda_j X$. We represent X by

$$X = \sum_{m=0}^{\infty} \sum_{|\alpha|=m} d_{m,\alpha}^{(k)} X_{m,\alpha}^{(k)}.$$

It follows that

$$\begin{aligned} \sum_{m=0}^{\infty} \sum_{|\alpha|=m} d_{m,\alpha}^{(k)} (-\lambda_m^{(k)}) X_{m,\alpha}^{(k)} &= \sum_{m=0}^{\infty} \sum_{|\alpha|=m} d_{m,\alpha}^{(k)} L_k^{(i_0, \dots, i_k)} X_{m,\alpha}^{(k)} \\ &= L_k^{(i_0, \dots, i_k)} X \\ &= -\lambda_j^{(k)} \sum_{m=0}^{\infty} \sum_{|\alpha|=m} d_{m,\alpha}^{(k)} X_{m,\alpha}^{(k)}. \end{aligned}$$

For any $i \neq j, |\beta| = i$, multiplying by $w_k X_{i,\beta}^{(k)}$ and then integrating on \bar{V}_n we have

$$\begin{aligned} & \sum_{|\alpha|=i} d_{i,\alpha}^{(k)} \lambda_i^{(k)} \left(X_{i,\alpha}^{(k)}, w_k^{(i_0, \dots, i_k)} X_{i,\beta}^{(k)} \right) \\ &= \sum_{|\alpha|=i} d_{i,\alpha}^{(k)} \lambda_j^{(k)} \left(X_{i,\alpha}^{(k)}, w_k^{(i_0, \dots, i_k)} X_{i,\beta}^{(k)} \right), \quad \forall i \neq j, |\beta| = i, \\ &\Rightarrow \left(X_{i,\alpha}^{(k)}, w_k^{(i_0, \dots, i_k)} X_{i,\beta}^{(k)} \right)_{\beta,\alpha} \left(d_{i,\alpha}^{(k)} \lambda_i^{(k)} \right)_\alpha \\ &= \left(X_{i,\alpha}^{(k)}, w_k^{(i_0, \dots, i_k)} X_{i,\beta}^{(k)} \right)_{\beta,\alpha} \left(d_{i,\alpha}^{(k)} \lambda_j^{(k)} \right)_\alpha, \quad \forall i \neq j, |\beta| = i, \\ &\Rightarrow d_{i,\alpha}^{(k)} \lambda_i^{(k)} = d_{i,\alpha}^{(k)} \lambda_j^{(k)}, \quad \forall i \neq j, |\beta| = i, \text{ because } \det \left(X_{i,\alpha}^{(k)}, w_k^{(i_0, \dots, i_k)} X_{i,\beta}^{(k)} \right)_{\beta,\alpha} \neq 0, \\ &\Rightarrow d_{i,\alpha}^{(k)} = 0, \quad \forall i \neq j, |\alpha| = i, \text{ because } \lambda_i^{(k)} \neq \lambda_j^{(k)}. \end{aligned}$$

It follows that

$$X = \sum_{|\alpha|=j} d_{j,\alpha}^{(k)} X_{j,\alpha}^{(k)}.$$

This completes the proof.

(ii) is obvious. □

Definition of the Solution

We shall now formally derive the Fokker–Planck equation as the diffusion limit of the Wright–Fisher model and introduce our solution concept for this equation. We consider a diploid population of fixed size N with $n + 1$ possible alleles A_1, \dots, A_{n+1} , at a given locus. Suppose that the individuals in the population are monoecious, that there are no selective differences between these alleles and no mutations. There are $2N$ alleles in the population in any generation, so it is sufficient to focus on the number $Y_m = (Y_m^1, \dots, Y_m^n)$ of alleles A_1, \dots, A_n at generation time m . Assume that $Y_0 = i_0 = (i_0^1, \dots, i_0^n)$ and according to the Wright–Fisher model, the alleles in generation $m + 1$ are derived by sampling with replacement from the alleles of generation m . Thus, the transition probability is

$$\mathbb{P}(Y_{m+1} = j | Y_m = i) = \frac{(2N)!}{(j^0)!(j^1)!\dots(j^n)!} \prod_{k=0}^n \left(\frac{i^k}{2N}\right)^{j^k},$$

where

$$i, j \in S_n^{(2N)} = \left\{ i = (i^1, \dots, i^n) : i^k \in \{0, 1, \dots, 2N\}, \sum_{k=1}^n i^k \leq 2N \right\},$$

and

$$i^0 = 2N - |i| = 2N - i^1 - \dots - i^n; \quad j^0 = 2N - |j| = 2N - j^1 - \dots - j^n.$$

After rescaling

$$t = \frac{m}{2N}, \quad X_t = \frac{Y_t}{2N},$$

we have a discrete Markov chain X_t valued in $\{0, \frac{1}{2N}, \dots, 1\}^n$ with $t = 1$ now corresponding to $2N$ generations. It is easy to see that

$$\begin{aligned} X_0 &= p = \frac{i_0}{2N}, \\ \mathbb{E}(\delta X_t^i) &= 0, \\ \mathbb{E}(\delta X_t^i \cdot \delta X_t^j) &= (X_t^i)(\delta_{ij} - X_t^j), \\ \mathbb{E}(\delta X_t)^\alpha &= (\delta t) \quad \text{for } |\alpha| \geq 3. \end{aligned} \tag{3}$$

We now denote by $m_\alpha(t)$ the α^{th} –moment of the distribution about zero at the t^{th} generation, i.e.,

$$m_\alpha(t) = \mathbb{E}(X_t)^\alpha.$$

Then

$$m_\alpha(t + 1) = \mathbb{E}(X_t + \delta X_t)^\alpha.$$

Expanding the right hand side and noting (3) we obtain the following recursion formula, under the assumption that the population number N is sufficiently large to neglect terms of order $\frac{1}{N^2}$ and higher,

$$m_\alpha(t + 1) = \left\{ 1 - \frac{|\alpha|(|\alpha| - 1)}{2} \right\} m_\alpha(t) + \sum_{i=1}^n \frac{\alpha_i(\alpha_i - 1)}{2} m_{\alpha - e_i}(t). \tag{4}$$

Under this assumption, the moments change very slowly per generation and we can replace this system of difference equations by a system of differential equations:

$$\dot{m}_\alpha(t) = -\frac{|\alpha|(|\alpha| - 1)}{2} m_\alpha(t) + \sum_{i=1}^n \frac{\alpha_i(\alpha_i - 1)}{2} m_{\alpha - e_i}(t). \tag{5}$$

With the aim of finding a continuous process which approximates the above discrete process, we should look for a continuous Markov process $\{X_t\}_{t \geq 0}$ valued in $[0, 1]^n$ with the same conditions as (3) and (5). Denoting by $u(x, t)$ the probability density function of this continuous process, the condition (3) implies (see for example [9, p. 137], or for a more rigorous analysis [6–8]) that u is a solution of the Fokker–Planck (Kolmogorov forward) equation

$$\begin{cases} u_t = L_n u \text{ in } V_n \times (0, \infty), \\ u(x, 0) = \delta_p(x) \text{ in } V_n; \end{cases} \tag{6}$$

and the condition (5) implies

$$[u_t, x^\alpha]_n = \left[u, -\frac{|\alpha|(|\alpha| - 1)}{2} x^\alpha + \sum_{i=1}^n \frac{\alpha_i(\alpha_i - 1)}{2} x^{\alpha - e_i} \right]_n = [u, L^*(x^\alpha)]_n, \quad \forall \alpha,$$

and hence, since the polynomials are dense in H_n w.r.t. the product $[\cdot, \cdot]_n$,

$$[u_t, \phi]_n = [u, L_n^* \phi]_n, \quad \forall \phi \in H_n. \tag{7}$$

This leads us to the following definition of a solution.

Definition 2.8 We call $u \in H$ a solution of the Fokker–Planck equation associated with the Wright–Fisher model if

$$u_t = L_n u \text{ in } V_n \times (0, \infty), \tag{8}$$

$$u(x, 0) = \delta_p(x) \text{ in } V_n, \tag{9}$$

$$[u_t, \phi]_n = [u, L_n^* \phi]_n, \quad \forall \phi \in H_n. \tag{10}$$

We point out that the last of these equations implicitly contains the boundary behavior that we wish to impose upon our solution. This will become clear from our construction in the next section.

The Global Solution

In this subsection, we shall construct the solution and prove the existence as well as the uniqueness of the solution. The process of finding the solution is as follows: We first find the general solution of the Fokker–Planck equation (21) by the separation of variables method. Then we construct a solution depending on certain parameters. We then use the conditions of (9, 10) to determine the parameters. Finally, we verify that we have indeed found the solution.

Step 1 Working on V_n , assume that $u_n(\mathbf{x}, t) = X(\mathbf{x})T(t)$ is a solution of the Fokker–Planck equation (21). Then we have

$$\frac{T_t}{T} = \frac{L_n X}{X} = -\lambda.$$

Clearly λ is a constant which is independent on T, X . From the Proposition 2.7 we obtain the local solution of the Eq. (21) of the form

$$u_n(\mathbf{x}, t) = \sum_{m=0}^{\infty} \sum_{|\alpha|=m} c_{m,\alpha}^{(n)} X_{m,\alpha}^{(n)}(\mathbf{x}) e^{-\lambda_m^{(n)} t},$$

where

$$\lambda_m^{(n)} = \frac{(n+m)(n+m+1)}{2}$$

is the eigenvalue of L_n and

$$X_{m,\alpha}^{(n)}(\mathbf{x}), \quad |\alpha| = m$$

are the corresponding eigenvectors of L_n .

For $m \geq 0, |\beta| = m$, we conclude from Proposition 2.3 that

$$L_n^*(w_n X_{m,\beta}^{(n)}) = -\lambda_m^{(n)} w_n X_{m,\beta}^{(n)}.$$

It follows that

$$\begin{aligned} [u_t, w_n X_{m,\beta}^{(n)}]_n &= [u, L_n^*(w_n X_{m,\beta}^{(n)})]_n \quad (\text{the moment condition}) \\ &= -\lambda_m^{(n)} [u, w_n X_{m,\beta}^{(n)}]_n. \end{aligned}$$

Therefore

$$\begin{aligned} [u, w_n X_{m,\beta}^{(n)}]_n &= [u(\cdot, 0), w_n X_{m,\beta}^{(n)}]_n e^{-\lambda_m^{(n)} t} \\ &= w_n(\mathbf{p}) X_{m,\beta}^{(n)}(\mathbf{p}) e^{-\lambda_m^{(n)} t}. \end{aligned}$$

Thus,

$$\begin{aligned} w_n(\mathbf{p}) X_{m,\beta}^{(n)}(\mathbf{p}) e^{-\lambda_m^{(n)} t} &= [u, w_n X_{m,\beta}^{(n)}]_n \\ &= (u_n, w_n X_{m,\beta}^{(n)})_n \quad (\text{because } w_n \text{ vanishes on boundary}) \\ &= \sum_{|\alpha|=m} c_{m,\alpha}^{(n)} (X_{m,\alpha}^{(n)}, w_n X_{m,\beta}^{(n)})_n e^{-\lambda_m^{(n)} t}. \end{aligned}$$

It follows that

$$(c_{m,\alpha}^{(n)})_{\alpha} = \left[\left((X_{m,\alpha}^{(n)}, w_n X_{m,\beta}^{(n)})_n \right)_{\alpha,\beta} \right]^{-1} (w_n(\mathbf{p}) X_{m,\beta}^{(n)}(\mathbf{p}))_{\beta}.$$

Step 2 The solution $u \in H$ satisfying (21) will be found in the following form

$$u(\mathbf{x}, t) = \sum_{k=1}^n u_k(\mathbf{x}, t) \chi_{V_k}(x) + \sum_{i=0}^n u_0^i(\mathbf{x}, t) \delta_{e^i}(\mathbf{x}). \tag{11}$$

We use the condition (10) to obtain iteratively values of $u_k, k = n - 1, \dots, 0$. In fact, assume that we want to calculate $u_{n-1}^{(0, \dots, n-1)}(x^1, \dots, x^{n-1}, 0, t)$.

We note that, if we choose

$$\phi(\mathbf{x}) = x^1 \cdots x^n X_{k,\beta}^{(n-1)}(x^1, \dots, x^{n-1}), \quad |\beta| = k,$$

then $\phi(\mathbf{x})$ vanishes on faces of dimension at most $n - 1$ except the face $V_{n-1}^{0, \dots, n-1}$. Therefore, the expectation of ϕ will be

$$[u, \phi]_n = (u_n, \phi)_n + \left(u_{n-1}^{(0, \dots, n-1)}, \phi \right)_{n-1}.$$

The left hand side can be calculated easily by the condition (10)

$$[u_t, \phi]_n = [u, L_n^*(\phi)]_n = -\lambda_k^{(n-1)} [u, \phi]_n. \tag{12}$$

It follows that

$$[u, \phi]_n = \phi(\mathbf{p}) e^{-\lambda_k^{(n-1)} t}.$$

The first part of the right hand side is

$$(u_n, \phi)_n = \sum_{m,\alpha} c_{m,\alpha}^{(n)} \left(\int_{V_n} X_{m,\alpha}^{(n)}(\mathbf{x}) \phi(\mathbf{x}) d\mathbf{x} \right) e^{-\lambda_m^{(n)} t}.$$

Therefore we can expand $u_{n-1}^{(0, \dots, n-1)}(x^1, \dots, x^{n-1}, 0, t)$ as follows

$$\begin{aligned} u_{n-1}^{(0, \dots, n-1)}(x^1, \dots, x^{n-1}, 0, t) &= \sum_{m \geq 0} c_m^{(n-1)}(\mathbf{x}) e^{-\lambda_m^{(n-1)} t} \\ &= \sum_{m \geq 0} \sum_{l \geq 0} \sum_{|\alpha|=l} c_{m,l,\alpha}^{(n-1)} X_{l,\alpha}^{(n-1)}(x^1, \dots, x^{n-1}) e^{-\lambda_m^{(n-1)} t}. \end{aligned}$$

Putting this formula into Eq. (12), we shall obtain all the coefficients $c_{m,l,\alpha}^{(n-1)}$. Thus, we shall obtain $u_{n-1}^{(0, \dots, n-1)}(x^1, \dots, x^{n-1}, 0, t)$. Similarly we shall obtain u_{n-1} . And finally we shall obtain all $u_k, k = n - 1, \dots, 0$. Thus, we shall obtain the global solution in the form

$$\begin{aligned} u(\mathbf{x}, t) &= \sum_{k=1}^n u_k \chi_{V_k}(\mathbf{x}) + \sum_{i=0}^n u_0^i(\mathbf{x}, t) \delta_{e_i}(\mathbf{x}). \\ &= \sum_{k=1}^n \sum_{m \geq 0} \sum_{l \geq 0} \sum_{|\alpha|=l} c_{m,l,\alpha}^{(k)} X_{l,\alpha}^{(k)}(\mathbf{x}) e^{-\lambda_m^{(k)} t} \chi_{V_k}(\mathbf{x}) + \sum_{i=0}^n u_0^i(\mathbf{x}, t) \delta_{e_i}(\mathbf{x}). \tag{13} \end{aligned}$$

It is not difficult to show that u is a solution of the Fokker–Planck equation associated with the WF model.

Step 3 We can easily see that this solution is unique. In fact, assume that u_1, u_2 are two solutions of the Fokker–Planck equation associated with the WF model. Then $u = u_1 - u_2$ will satisfy

$$\begin{aligned} u_t &= L_n u \text{ in } V_n \times (0, \infty), \\ u(x, 0) &= 0 \text{ in } \bar{V}_n, \\ [u_t, \phi]_n &= [u, L^* \phi]_n, \quad \forall \phi \in H_n. \end{aligned}$$

It follows that

$$\begin{aligned}
 [u_t, 1]_n &= [u, L_n^*(1)]_n = 0, \\
 [u_t, x^i]_n &= [u, L_n^*(x^i)]_n = 0, \\
 \left[u_t, w_k^{(i_0, \dots, i_k)} X_{j, \alpha}^{(k)} \chi_{V_k^{(i_0, \dots, i_k)}} \right]_n &= \left[u, L_n^*(w_k^{(i_0, \dots, i_k)} X_{j, \alpha}^{(k)} \chi_{V_k^{(i_0, \dots, i_k)}}) \right]_n \\
 &= \left[u, L_k^*(w_k^{(i_0, \dots, i_k)} X_{j, \alpha}^{(k)} \chi_{V_k^{(i_0, \dots, i_k)}}) \right]_n \\
 &= -\lambda_j^{(k)} \left[u, w_k^{(i_0, \dots, i_k)} X_{j, \alpha}^{(k)} \chi_{V_k^{(i_0, \dots, i_k)}} \right]_n.
 \end{aligned}$$

Therefore

$$\begin{aligned}
 [u, 1]_n &= [u(\cdot, 0), 1]_n = 0, \\
 [u, x^i]_n &= [u(\cdot, 0), x^i]_n = 0, \\
 \left[u, w_k^{(i_0, \dots, i_k)} X_{j, \alpha}^{(k)} \chi_{V_k^{(i_0, \dots, i_k)}} \right]_n &= \left[u(\cdot, 0), w_k^{(i_0, \dots, i_k)} X_{j, \alpha}^{(k)} \chi_{V_k^{(i_0, \dots, i_k)}} \right]_n e^{-\lambda_j^{(k)} t} = 0.
 \end{aligned}$$

Since $\{1, \{x^i\}_i, \{w_k^{(i_0, \dots, i_k)} X_{j, \alpha}^{(k)} \chi_{V_k^{(i_0, \dots, i_k)}}\}_{1 \leq k \leq n, (i_0, \dots, i_k) \in I_k, j \geq 0, |\alpha| = j}\}$ is also a basis of H_n it follows that $u = 0 \in H$.

In conclusion, we have established:

Theorem 2.9 *The Fokker Planck equation associated with the Wright–Fisher model with $n + 1$ alleles possesses the unique solution*

$$\begin{aligned}
 u(\mathbf{x}, t) &= \sum_{k=1}^n u_k \chi_{V_k}(\mathbf{x}) + \sum_{i=0}^n u_0^i(\mathbf{x}, t) \delta_{e_i}(\mathbf{x}) \\
 &= \sum_{k=1}^n \sum_{m \geq 0} \sum_{l \geq 0} \sum_{|\alpha|=l} c_{m, l, \alpha}^{(k)} X_{l, \alpha}^{(k)}(\mathbf{x}) e^{-\lambda_m^{(k)} t} \chi_{V_k}(\mathbf{x}) + \sum_{i=0}^n u_0^i(\mathbf{x}, t) \delta_{e_i}(\mathbf{x}). \tag{14}
 \end{aligned}$$

Example 2.10 To illustrate this process, we consider the case of three alleles.

We shall construct the global solution for the problem

$$\begin{cases} \frac{\partial u}{\partial t} = L_2 u, & \text{in } V_2 \times (0, \infty), \\ u(\mathbf{x}, 0) = \delta_{\mathbf{p}}(\mathbf{x}), & \mathbf{x} \in V_2, \\ [u_t, \phi]_2 = [u, L_2^* \phi]_2, & \text{for all } \phi \in H_2, \end{cases}$$

where the global solution of the form

$$u = u_2 \chi_{V_2} + u_1^{0,1} \chi_{V_1^{0,1}} + u_1^{0,2} \chi_{V_1^{0,2}} + u_1^{0,0} \chi_{V_1^{0,0}} + u_0^1 \chi_{V_0^1} + u_0^2 \chi_{V_0^2} + u_0^0 \chi_{V_0^0},$$

and the product is

$$\begin{aligned}
 [u, \phi]_2 &= \int_{V_2} u_2 \phi|_{V_2} d\mathbf{x} + \int_0^1 u_1^{0,1}(x^1, 0, t) \phi(x^1, 0) dx^1 + \int_0^1 u_1^{0,2}(0, x^2, t) \phi(0, x^2) dx^2 \\
 &\quad + \frac{1}{\sqrt{2}} \int_0^1 u_1^{1,2}(x^1, 1 - x^1, t) \phi(x^1, 1 - x^1) dx^1 \\
 &\quad + u_0^1(1, 0, t) \phi(1, 0) + u_0^2(0, 1, t) \phi(0, 1) + u_0^0(0, 0, t) \phi(0, 0).
 \end{aligned}$$

Step 1: We find the local solution u_2

$$u_2(\mathbf{x}, t) = \sum_{m \geq 0} \sum_{\alpha^1 + \alpha^2 = m} c_{m, \alpha^1, \alpha^2}^{(2)} X_{m, \alpha^1, \alpha^2}^{(2)}(\mathbf{x}) e^{-\lambda_m^{(2)} t}.$$

To define the coefficients $c_{m, \alpha^1, \alpha^2}^{(2)}$ we use the initial condition and the orthogonality of the eigenvectors $X_{m, \alpha^1, \alpha^2}^{(2)}$ to get

$$\begin{aligned} w_2(\mathbf{p}) X_{m, \beta^1, \beta^2}^{(2)}(\mathbf{p}) &= \left[u(0), w_2 X_{m, \beta^1, \beta^2}^{(2)} \right]_2 \\ &= \left(u_2(0), w_2 X_{m, \beta^1, \beta^2}^{(2)} \right)_2 \quad \text{because } w_2 \text{ vanishes on the boundary} \\ &= \sum_{\alpha^1 + \alpha^2 = m} c_{m, \alpha^1, \alpha^2}^{(2)} \left(X_{m, \alpha^1, \alpha^2}^{(2)}, w_2 X_{m, \beta^1, \beta^2}^{(2)} \right) \quad \text{for all } \beta^1 + \beta^2 = m. \end{aligned}$$

Since the matrix

$$\left(X_{m, \alpha^1, \alpha^2}^{(2)}, w_2 X_{m, \beta^1, \beta^2}^{(2)} \right)_{(\alpha^1, \alpha^2), (\beta^1, \beta^2)}$$

is positive definite, we therefore have unique values of the $c_{m, \alpha^1, \alpha^2}^{(2)}$. It follows that we have a unique local solution u_2 .

Step 2: We will use the moment condition to define all other coefficients of the global solution.

Firstly, we define the coefficients of $u_1^{1,2}$ as follows

$$u_1^{1,2}(x^1, 1 - x^1, t) = \sum_{m \geq 0} c_m(x^1) e^{-\lambda_m^{(1)} t} \tag{15}$$

$$= \sum_{m, l \geq 0} c_{m, l} X_l^{(1)}(x^1) e^{-\lambda_m^{(1)} t}. \tag{16}$$

We note that

$$L_2^* \left(x^1 x^2 X_k^{(1)}(x^1) \right) = -\lambda_k^{(1)} x^1 x^2 X_k^{(1)}(x^1).$$

Therefore

$$\left[u_t, x^1 x^2 X_k^{(1)}(x^1) \right]_2 = \left[u, L_2^* \left(x^1 x^2 X_k^{(1)}(x^1) \right) \right]_2 = -\lambda_k^{(1)} \left[u, x^1 x^2 X_k^{(1)}(x^1) \right]_2.$$

It follows that

$$\left[u, x^1 x^2 X_k^{(1)}(x^1) \right]_2 = p^1 p^2 X_k^{(1)}(p^1) e^{-\lambda_k^{(1)} t}.$$

Thus we have

$$\begin{aligned}
 p^1 p^2 X_k^{(1)}(p^1) e^{-\lambda_k^{(1)} t} &= \left[u, x^1 x^2 X_k^{(1)}(x^1) \right]_2 \\
 &= \left(u_2, x^1 x^2 X_k^{(1)}(x^1) \right)_2 + \left(u_1^{1,2}, x^1 (1 - x^1) X_k^{(1)}(x^1) \right)_1 \\
 &\quad \text{because } x^1 x^2 \text{ vanish on the other boundaries,} \\
 &= \sum_{m \geq 0} \left(\sum_{|\alpha|=m} c_{m,\alpha}^{(2)} \left(\int_{V_2} x^1 x^2 X_{m,\alpha}^{(2)}(x^1, x^2) X_k^{(1)}(x^1) d\mathbf{x} \right) \right) e^{-\lambda_m^{(2)} t} \\
 &\quad + \sum_{m \geq 0} c_{m,k} \left(X_k^{(1)}, w_1 X_k^{(1)} \right) e^{-\lambda_m^{(1)} t} \\
 &\quad \text{because of the orthogonality of } (\cdot, \cdot)_1 \text{ with respect to } w_1, \\
 &= \sum_{m \geq 0} r_m e^{-\lambda_m^{(2)} t} + \sum_{m \geq 0} c_{m,k} d_k e^{-\lambda_m^{(1)} t}.
 \end{aligned}$$

By equating the coefficients of $e^{\alpha t}$ we obtain $u_1^{1,2}$. Similarly we obtain u_1 . Then, we define the coefficients of u_0^1 from the first moment.

Note that when $\phi = x^i$, $L_2^*(\phi) = 0$, therefore $[u_t, \phi]_2 = 0$ or

$$[u, x^i]_2 = [u(0), x^i] = p^i.$$

It follows that

$$p^1 = [u, x^1] = (u_2, x^1)_2 + (u_1^{0,1}, x^1)_1 + (u_1^{1,2}, x^1)_1 + u_0^1(1, 0, t).$$

Thus we obtain $u_0^1(1, 0, t)$. Similarly we get all u_0 . Therefore we obtain the global solution u .

It is easy to check that u is a global solution. To prove the uniqueness we proceed as follows: Assume that u is the difference of any two global solutions, i.e., u satisfies

$$\begin{cases}
 u_t = L_2 u & \text{in } V_2 \times (0, \infty), \\
 u(\mathbf{x}, 0) = 0 & \text{in } V_2 \\
 [u_t, \phi]_2 = [u, L_2^* \phi]_2 & \text{for all } \phi \in H_2.
 \end{cases}$$

We shall prove that

$$[u, \phi]_2 = 0 \quad \forall \phi \in H_2. \tag{17}$$

In fact,

$$\begin{aligned}
 [u_t, 1]_2 &= [u, L_2^*(1)]_2 = 0 \Rightarrow [u, 1]_2 = [u(0), 1]_2 = 0, \\
 [u_t, x^i]_2 &= [u, L_2^*(x^i)]_2 = 0 \Rightarrow [u, x^i]_2 = [u(0), x^i]_2 = 0, \\
 [u_t, w_1(x^i) X_m^{(1)}(x^i)]_2 &= [u, L_2^*(w_1(x^i) X_m^{(1)}(x^i))]_2 = -\lambda_m^{(1)} [u, w_1(x^i) X_m^{(1)}(x^i)]_2.
 \end{aligned}$$

It implies that

$$\left[u, w_1(x^i) X_m^{(1)}(x^i) \right]_2 = \left[u(0), w_1(x^i) X_m^{(1)}(x^i) \right]_2 e^{-\lambda_m^{(1)} t} = 0.$$

Therefore

$$\begin{aligned} [u_t, w_2(x^1, x^2)X_{m,\alpha}^{(2)}(x^1, x^2)]_2 &= [u, L_2^*(w_2(x^1, x^2)X_{m,\alpha}^{(2)}(x^1, x^2))]_2 \\ &= -\lambda_m^{(2)} [u, w_2(x^1, x^2)X_{m,\alpha}^{(2)}(x^1, x^2)]_2, \\ \Rightarrow [u, w_2(x^1, x^2)X_{m,\alpha}^{(2)}(x^1, x^2)]_2 &= [u(0), w_2(x^1, x^2)X_{m,\alpha}^{(2)}(x^1, x^2)]_2 e^{-\lambda_m^{(2)}t} = 0. \end{aligned}$$

We need only to prove that Eq. (17) holds for all

$$\phi(x^1, x^2) = (x^1)^m(x^2)^n, \quad \forall m, n \geq 0.$$

- (1) If $n = 0, m \geq 0$, we see that ϕ can be generated from $\{1, x^1, w_1(x^1)X_m^{(1)}(x^1)\}$, therefore $[u, \phi]_2 = 0$;
- (2) If $m = 0, n \geq 0$, we see that ϕ can be generated from $\{1, x^2, w_1(x^2)X_n^{(1)}(x^2)\}$, therefore $[u, \phi]_2 = 0$;
- (3) If $n = 1, m \geq 1$, we expand $(x^1)^{m-1}$ by

$$(x^1)^{m-1} = \sum_{k \geq 0} c_k X_k^{(1)}(x^1).$$

Note that

$$L_2^*(x^1 x^2 X_k^{(1)}(x^1)) = -\lambda_k^{(1)} x^1 x^2 X_k^{(1)}(x^1).$$

Therefore

$$[u_t, x^1 x^2 X_k^{(1)}(x^1)]_2 = [u, L_2^*(x^1 x^2 X_k^{(1)}(x^1))]_2 = -\lambda_k^{(1)} [u, x^1 x^2 X_k^{(1)}(x^1)]_2.$$

It follows that

$$[u, x^1 x^2 X_k^{(1)}(x^1)]_2 = [u(0), x^1 x^2 X_k^{(1)}(x^1)]_2 e^{-\lambda_k^{(1)}t} = 0.$$

Therefore

$$[u, \phi]_2 = \sum_{k \geq 0} c_k [u, x^1 x^2 X_k^{(1)}(x^1)]_2 = 0;$$

- (4) If $n \geq 2, m \geq 1$ we proceed by induction w.r.t. n . We have

$$\begin{aligned} (x^1)^m(x^2)^n &= x^1 x^2 (x^1 + x^2 - 1)(x^1)^{m-1}(x^2)^{n-2} + (x^1)^m(1 - x^1)(x^2)^{n-1} \\ &= -w_2(x^1, x^2)(x^1)^{m-1}(x^2)^{n-2} + (x^1)^m(1 - x^1)(x^2)^{n-1}. \end{aligned}$$

By the inductive assumption, we have

$$[u, (x^1)^m(1 - x^1)(x^2)^{n-1}]_2 = 0.$$

Then, we expand $(x^1)^{m-1}(x^2)^{n-2}$ by

$$(x^1)^{m-1}(x^2)^{n-2} = \sum_{m,\alpha} c_{m,\alpha}^{(2)} X_{m,\alpha}^{(2)}(x^1, x^2).$$

Therefore

$$[u, w_2(x^1, x^2)(x^1)^{m-1}(x^2)^{n-2}]_2 = \sum_{m,\alpha} c_{m,\alpha}^{(2)} [u, w_2(x^1, x^2)X_{m,\alpha}^{(2)}(x^1, x^2)]_2 = 0.$$

It follows that $[u, (x^1)^m(x^2)^n]_2 = 0$.

Thus, $u = 0$.

Applications

In this section, we present some applications of our global solution to the evolution of the process $(X_t)_{t \geq 0}$ such as the expectation and the second moment of the absorption time, the probability distribution of the absorption time for having $k + 1$ alleles, the probability of having exactly $k + 1$ alleles, the α th moments, the probability of heterogeneity, and the rate of loss of one allele in a population having $k + 1$ alleles. Several of our formulas are known from other methods, see [9, 15–17, 19, 20], but we emphasize here the general and unifying approach.

The Absorption Time for Having $k + 1$ Alleles

The moments of the sojourn and absorption times were derived by Nagylaki [22] for two alleles, and by Lessard and Lahaie [18] in the multi-allele case. We denote by $T_{n+1}^{k+1}(p) = \inf\{t > 0 : X_t \in \bar{V}_k | X_0 = p\}$ the first time when the population has (at most) $k + 1$ alleles. $T_{n+1}^{k+1}(p)$ is a continuous random variable valued in $[0, \infty)$ and we denote by $\phi(t, p)$ its probability density function. It is easy to see that \bar{V}_k is invariant under the process $(X_t)_{t \geq 0}$, i.e. if $X_s \in \bar{V}_k$ then $X_t \in \bar{V}_k$ for all $t \geq s$ (once an allele is lost from the population, it can never again be recovered). We have the equality

$$\mathbb{P}\left(T_{n+1}^{k+1}(p) \leq t\right) = \mathbb{P}(X_t \in \bar{V}_k | X_0 = p) = \int_{\bar{V}_k} u(x, p, t) d\mu(x).$$

It follows that

$$\phi(t, p) = \int_{\bar{V}_k} \frac{\partial}{\partial t} u(x, p, t) d\mu(x).$$

Therefore the expectation for the absorption time of having $k + 1$ alleles is (see [9, p. 194]):

$$\begin{aligned} \mathbb{E}(T_{n+1}^{k+1}(p)) &= \int_0^\infty t \phi(t, p) dt \\ &= \int_{\bar{V}_k} \int_0^\infty t \frac{\partial}{\partial t} u(x, p, t) dt d\mu(x) \\ &= \sum_{j=1}^k \sum_{(i_0, \dots, i_j) \in I_j} \sum_{m \geq 0} \sum_{|\alpha|=m} c_{m,\alpha}^{(j)} \int_{V_j^{(i_0, \dots, i_j)}} X_{m,\alpha}^{(j)}(x) \left(\int_0^\infty t \frac{\partial}{\partial t} e^{-\lambda_m^{(j)} t} dt \right) d\mu_j^{(i_0, \dots, i_j)}(x) \\ &\quad + \sum_{i=0}^n \sum_{k=1}^n \sum_{m \geq 0} \sum_{|\alpha|=m} c_{m,\alpha}^{(k)} a_{m,\alpha,i}^{(k)} \left(\int_0^\infty t \frac{\partial}{\partial t} e^{-\lambda_m^{(k)} t} dt \right) \\ &= \sum_{j=1}^k \sum_{(i_0, \dots, i_j) \in I_j} \sum_{m \geq 0} \sum_{|\alpha|=m} c_{m,\alpha}^{(j)} \int_{V_j^{(i_0, \dots, i_j)}} X_{m,\alpha}^{(j)}(x) \left(-\frac{1}{\lambda_m^{(j)}} \right) d\mu_j^{(i_0, \dots, i_j)}(x) \\ &\quad + \sum_{i=0}^n \sum_{k=1}^n \sum_{m \geq 0} \sum_{|\alpha|=m} c_{m,\alpha}^{(k)} a_{m,\alpha,i}^{(k)} \left(-\frac{1}{\lambda_m^{(k)}} \right); \end{aligned}$$

and the second moment of this absorption time is (see [1,20]):

$$\begin{aligned} \mathbb{E}(T_{n+1}^{k+1}(p))^2 &= \int_0^\infty t^2 \phi(t, p) dt \\ &= \int_{\bar{V}_k} \int_0^\infty t^2 \frac{\partial}{\partial t} u(x, p, t) dt d\mu(x) \\ &= \sum_{j=1}^k \sum_{(i_0, \dots, i_j) \in I_j} \sum_{m \geq 0} \sum_{|\alpha|=m} c_{m,\alpha}^{(j)} \int_{V_j^{(i_0, \dots, i_j)}} X_{m,\alpha}^{(j)}(x) \left(\int_0^\infty t^2 \frac{\partial}{\partial t} e^{-\lambda_m^{(j)} t} dt \right) d\mu_j^{(i_0, \dots, i_j)}(x) \\ &\quad + \sum_{i=0}^n \sum_{k=1}^n \sum_{m \geq 0} \sum_{|\alpha|=m} c_{m,\alpha}^{(k)} a_{m,\alpha,i}^{(k)} \left(\int_0^\infty t^2 \frac{\partial}{\partial t} e^{-\lambda_m^{(k)} t} dt \right) \\ &= \sum_{j=1}^k \sum_{(i_0, \dots, i_j) \in I_j} \sum_{m \geq 0} \sum_{|\alpha|=m} c_{m,\alpha}^{(j)} \int_{V_j^{(i_0, \dots, i_j)}} X_{m,\alpha}^{(j)}(x) \left(-\frac{2}{(\lambda_m^{(j)})^2} \right) d\mu_j^{(i_0, \dots, i_j)}(x) \\ &\quad + \sum_{i=0}^n \sum_{k=1}^n \sum_{m \geq 0} \sum_{|\alpha|=m} c_{m,\alpha}^{(k)} a_{m,\alpha,i}^{(k)} \left(-\frac{2}{(\lambda_m^{(k)})^2} \right). \end{aligned}$$

In order to see what this means, we consider the case of three alleles:

$$\begin{aligned} u(x^1, x^2; t) &= u_2(x^1, x^2; t) \chi_{V_2} + u_1^{0,1}(x^1, 0; t) \chi_{V_1^{0,1}} \\ &\quad + u_1^{0,2}(0, x^2; t) \chi_{V_1^{0,2}} + u_1^{0,0}(x^1, 1 - x^1; t) \chi_{V_1^{0,0}} \\ &\quad + u_0^1(t) \delta_{e_1} + u_0^2(t) \delta_{e_2} + u_0^0(t) \delta_{e_0}, \end{aligned}$$

and the product is

$$\begin{aligned} [u, \phi]_2 &= (u_2, \phi)_2 + (u_1^{0,1}, \phi(\cdot, 0))_1 + (u_1^{0,2}, \phi(0, \cdot))_1 + (u_1^{0,1}, \phi(\cdot, 1 - \cdot))_1 \\ &\quad + u_0^1(1, 0; t) \phi(1, 0) + u_0^2(0, 1; t) \phi(0, 1) + u_0^0(0, 0; t) \phi(0, 0) \\ &= \int_{V_2} u_2(x^1, x^2; t) \phi(x^1, x^2) dx^1 dx^2 + \int_0^1 u_1^{0,1}(x^1, 0; t) \phi(x^1, 0) dx^1 \\ &\quad + \int_0^1 u_1^{0,2}(0, x^2; t) \phi(0, x^2) dx^2 + \frac{1}{\sqrt{2}} \int_0^1 u_1^{1,2}(x^1, 1 - x^1; t) \phi(x^1, 1 - x^1) dx^1 \\ &\quad + u_0^1(1, 0; t) \phi(1, 0) + u_0^2(0, 1; t) \phi(0, 1) + u_0^0(0, 0; t) \phi(0, 0). \end{aligned}$$

By expansion of eigenvectors, we have

$$u_2(\mathbf{x}; \mathbf{p}; t) = \sum_{m \geq 0} \sum_{|\alpha|=m} c_{m,\alpha}^{(2)}(\mathbf{p}) X_{m,\alpha}^{(2)}(\mathbf{x}) e^{-\lambda_m^{(2)} t},$$

where $c_{m,\alpha}^{(2)}(\mathbf{p})$ is uniquely defined. We represent $u_1(\mathbf{x}; t)$ by

$$u_1^{0,1}(x^1, 0; t) = \sum_{m \geq 0} a_m^{0,1}(x^1) e^{-\lambda_m^{(1)} t}; \tag{18}$$

$$u_1^{0,2}(0, x^2; t) = \sum_{m \geq 0} a_m^{0,2}(x^2) e^{-\lambda_m^{(1)} t}; \tag{19}$$

$$u_1^{1,2}(x^1, 1 - x^1; t) = \sum_{m \geq 0} a_m^{1,2}(x^1) e^{-\lambda_m^{(1)} t}, \tag{20}$$

where the coefficients $a_m^{(1)}(x^1)$ are defined as follows:

Putting

$$\psi_n(x^1) := x^1(1 - x^1)X_n^{(1)}(x^1),$$

we note that $\psi_n(0) = \psi_n(1) = 0$ and

$$L_2^* \psi_n(x^1) = -\lambda_n^{(1)} \psi_n(x^1).$$

It follows that

$$\begin{aligned} [u_t, \psi_n(x^1)]_2 &= [u, L_2^*(\psi_n(x^1))]_2 \\ &= -\lambda_n^{(1)} [u, \psi_n(x^1)]_2. \end{aligned}$$

Therefore

$$\begin{aligned} \psi_n(p^1)e^{-\lambda_n^{(1)}t} &= [u(0), \psi_n(x^1)]_2 e^{-\lambda_n^{(1)}t} = [u, \psi_n(x^1)]_2 \\ &= (u_2, \psi_n(x^1))_2 + (u_1, \psi_n(x^1))_1 + (u_0, \psi_n(x^1))_0 \\ &= \sum_{m \geq 0} \sum_{|\alpha|=m} c_{m,\alpha}^{(2)} (X_{m,\alpha}^{(2)}, \psi_n(x^1))_2 e^{-\lambda_m^{(2)}t} + \sum_{m \geq 0} (a_m(x^1), \psi_n(x^1))_1 e^{-\lambda_m^{(1)}t} \end{aligned}$$

where $a_m(x^1) := a_m^{0,1}(x^1) + a_m^{1,2}(x^1)$ and note that $\psi_n(0) = \psi_n(1) = 0$

$$\begin{aligned} &= (a_0(x^1), \psi_n(x^1))_1 e^{-\lambda_0^{(1)}t} + \sum_{m \geq 1} \left\{ (a_m(x^1), \psi_n(x^1))_1 \right. \\ &\quad \left. + \sum_{|\alpha|=m-1} c_{m,\alpha}^{(2)} (X_{m,\alpha}^{(2)}, \psi_n(x^1))_2 \right\} e^{-\lambda_m^{(1)}t} \\ &\text{(because of } \lambda_m^{(1)} = \lambda_{m-1}^{(2)} \text{).} \end{aligned}$$

We obtain by equating the coefficients in terms of $e^{-\lambda t}$

$$\begin{aligned} (a_0(x^1), \psi_n(x^1))_1 &= \delta_{0,n} \psi_n(p^1) \\ (a_m(x^1), \psi_n(x^1))_1 &= \delta_{m,n} \psi_n(p^1) - \sum_{|\alpha|=m-1} c_{m-1,\alpha}^{(2)} (X_{m-1,\alpha}^{(2)}, \psi_n(x^1))_2, \quad \text{if } m \geq 1. \end{aligned} \tag{21}$$

Remark 3.1 The coefficients of u_2 occur in the representation of the coefficients of u_1 because of the probability flux.

Similarly because of

$$\begin{aligned} L_2^*(x^1) &= 0, \\ [u_t, x^1]_2 &= [u, L_2^*(x^1)]_2 = 0. \end{aligned}$$

We have

$$\begin{aligned} p^1 &= [u(0), x^1]_2 = [u, x^1]_2 \\ &= (u_2, x^1)_2 + (u_1, x^1)_1 + (u_0, x^1)_0. \end{aligned}$$

Thus,

$$\begin{aligned}
 u_0^1(\mathbf{p}; t) &= p^1 - \sum_{m \geq 0} \sum_{|\alpha|=m} c_{m,\alpha}^{(2)}(\mathbf{p}) \left(X_{m,\alpha}^{(2)}, x^1 \right)_2 e^{-\lambda_m^{(2)} t} \\
 &\quad - \sum_{m \geq 0} \left(a_m^{0,1}, x^1 \right)_1 e^{-\lambda_m^{(1)} t} \\
 &\quad - \sum_{m \geq 0} \left(a_m^{1,2}, x^1 \right) e^{-\lambda_m^{(1)} t} \\
 &= p^1 - \left(a_0(x^1), x^1 \right)_1 e^{-\lambda_0^{(1)} t} - \sum_{m \geq 1} \left\{ \left(a_m(x^1), x^1 \right)_1 \right. \\
 &\quad \left. + \sum_{|\alpha|=m-1} c_{m-1,\alpha}^{(2)} \left(X_{m-1,\alpha}^{(2)}, x^1 \right)_2 \right\} e^{-\lambda_m^{(1)} t}.
 \end{aligned}$$

The expectation for the absorption time of having only 1 allele is

$$\begin{aligned}
 \mathbb{E}(T_3^1(\mathbf{p})) &= \int_0^\infty t \phi(t, \mathbf{p}) dt \\
 &= \int_0^\infty t \frac{\partial}{\partial t} \left(u_0^1(\mathbf{p}; t) + u_0^2(\mathbf{p}; t) + u_0^0(\mathbf{p}; t) \right) dt.
 \end{aligned}$$

We first calculate the first term; the other terms will be obtained similarly. To do this, we expand x^1 in terms of the $\psi_n(x^1)$

$$x^1 = \sum_{n \geq 0} d_n \psi_n(x^1).$$

We construct a sequence of entropy functions on $[0, 1]$ as follows.

- $E_0(x) = -x$
- $E_r(x)$ is the unique solution of the boundary value problem

$$\begin{cases} L_1^*(E_r(x)) = -r E_{r-1}(x) \\ E_r(0) = E_r(1) = 0. \end{cases}$$

By some simple calculations, we obtain some first entropy functions

- (1) $E_0(x) = -x$
- (2) $E_1(x) = -2(1-x) \log(1-x)$
- (3) $E_2(x) = -8xz(x) + 8(1-x) \log(1-x)$
- (4) $E_3(x) = 48(1-x)u(x) + 96[xz(x) - (1-x) \log(1-x)]$

where

$$z(x) = \int_x^1 \frac{\ln(1-y)}{y} dy, \quad u(x) = \int_x^1 \frac{z(y)}{1-y} dy.$$

Lemma 3.2 *The entropy functions satisfy*

$$\begin{aligned}
 \frac{\left(X_m^{(1)}, x^1 \right)_1}{\lambda_m^{(1)}} &= \left(E_1(x^1), X_m^{(1)} \right)_1, \\
 \frac{2 \left(X_m^{(1)}, x^1 \right)_1}{\left(\lambda_m^{(1)} \right)^2} &= \left(E_2(x^1), X_m^{(1)} \right)_1,
 \end{aligned}$$

and more generally,

$$\frac{r!(X_m^{(1)}, x^1)_1}{(\lambda_m^{(1)})^r} = (E_r(x^1), X_m^{(1)})_1, \quad r \geq 2.$$

Proof We have

$$\begin{aligned} \lambda_m^{(1)}(E_1(x^1), X_m^{(1)})_1 &= (E_1(x^1), \lambda_m^{(1)} X_m^{(1)})_1 \\ &= (E_1(x^1), -L_1(X_m^{(1)}))_1 \\ &= (-L_1^*(E_1(x^1)), X_m^{(1)})_1, \quad \text{because of } E_1(0) = E_1(1) = 0 \\ &= (-L_1^*(E_1(x^1)), X_m^{(1)})_1 \\ &= (x^1, X_m^{(1)})_1. \end{aligned}$$

Similarly we have

$$\begin{aligned} (\lambda_m^{(1)})^2(E_2(x^1), X_m^{(1)})_1 &= \lambda_m^{(1)}(E_2(x^1), \lambda_m^{(1)} X_{m,\alpha}^{(1)})_1 \\ &= \lambda_m^{(1)}(E_2(x^1), -L_1(X_{m,\alpha}^{(1)}))_1 \\ &= \lambda_m^{(1)}(-L_1^*(E_2(x^1)), X_{m,\alpha}^{(1)})_1, \quad \text{because of } E_2(0) = E_2(1) = 0 \\ &= \lambda_m^{(1)}(-L_1^*(E_2(x^1)), X_m^{(1)})_1 \\ &= \lambda_m^{(1)}(2E_1(x^1), X_m^{(1)})_1 \\ &= (2x^1, X_m^{(1)})_1, \quad \text{because of the above calculation.} \end{aligned}$$

The proof for all r is similar. □

From the Lemma, we can expand $E_1(x^1)$ as

$$E_1(x^1) = \sum_{n \geq 0} \frac{d_n}{\lambda_n^{(1)}} \psi_n(x^1).$$

Therefore we have

$$\begin{aligned} &\int_0^\infty t \frac{\partial u_0^1(\mathbf{p}; t)}{\partial t} dt \\ &= (a_0(x^1), x^1)_1 \int_0^\infty t \lambda_0^{(1)} e^{-\lambda_0^{(1)} t} dt \\ &\quad + \sum_{m \geq 1} \left\{ (a_m(x^1), x^1)_1 + \sum_{|\alpha|=m-1} c_{m-1,\alpha}^{(2)} (X_{m-1,\alpha}^{(2)}, x^1)_2 \right\} \int_0^\infty t \lambda_m^{(1)} e^{-\lambda_m^{(1)} t} dt \\ &= \frac{(a_0(x^1), x^1)_1}{\lambda_0^{(1)}} + \sum_{m \geq 1} \frac{(a_m(x^1), x^1)_1 + \sum_{|\alpha|=m-1} c_{m-1,\alpha}^{(2)} (X_{m-1,\alpha}^{(2)}, x^1)_2}{\lambda_m^{(1)}} \\ &= \sum_{n \geq 0} d_n \left\{ \frac{(a_0(x^1), \psi_n(x^1))_1}{\lambda_0^{(1)}} + \sum_{m \geq 1} \frac{(a_m(x^1), \psi_n(x^1))_1 + \sum_{|\alpha|=m-1} c_{m-1,\alpha}^{(2)} (X_{m-1,\alpha}^{(2)}, \psi_n(x^1))_2}{\lambda_m^{(1)}} \right\} \end{aligned}$$

$$\begin{aligned}
 &= \sum_{n \geq 0} d_n \left\{ \frac{\delta_{0,n} \psi_n(p^1)}{\lambda_0^{(1)}} + \sum_{m \geq 1} \frac{\delta_{m,n} \psi_n(p^1)}{\lambda_m^{(1)}} \right\}, \text{ because of (21)} \\
 &= \sum_{m \geq 0} \frac{d_m}{\lambda_m^{(1)}} \psi_m(p^1) \\
 &= E_1(p^1).
 \end{aligned}$$

Thus, we have

$$\mathbb{E}(T_3^1(\mathbf{p})) = E_1(p^1) + E_1(p^2) + E_1(p^3).$$

Remark 3.3 We can obtain the r th moments of this absorption time by the same method, i.e.

$$\mathbb{E}(T_3^1(\mathbf{p}))^r = E_r(p^1) + E_r(p^2) + E_r(p^3).$$

The Probability Distribution of the Absorption Time for Having $k + 1$ Alleles

We note that $X_{T_{n+1}^{k+1}(p)}$ is a random variable valued in $\overline{V_k}$. We consider the probability that this random variable takes its value in $V_k^{(i_0, \dots, i_k)}$, i.e., the probability of the population at the first time having at most $k + 1$ alleles to consist precisely of the $k + 1$ alleles $\{A_{i_0}, \dots, A_{i_k}\}$. Let g_k be a function of k variables defined inductively by

$$\begin{aligned}
 g_1(p^1) &= p^1; \\
 g_2(p^1, p^2) &= \frac{p^1}{1 - p^2} g_1(p^2) + \frac{p^2}{1 - p^1} g_1(p^1); \\
 g_{k+1}(p^1, \dots, p^{k+1}) &= \sum_{i=1}^{k+1} \frac{p^i}{1 - \sum_{j \neq i} p^j} g_k(p^1, \dots, p^{i-1}, p^{i+1}, \dots, p^{k+1}).
 \end{aligned}$$

Then we have

Theorem 3.4

$$\mathbb{P} \left(X_{T_{n+1}^{k+1}(p)} \in \overline{V_k^{(i_0, \dots, i_k)}} \right) = g_{k+1}(p^{i_0}, \dots, p^{i_k}).$$

Proof Method 1: By proving that

$$\mathbb{P} \left(X_{T_{n+1}^{k+1}(p)} \in \overline{V_k^{(i_0, \dots, i_k)}} \mid X_{T_{n+1}^{k+1}(p)} \in \overline{V_k^{(i_1, \dots, i_k)}} \right) = \frac{p^{i_0}}{1 - p^{i_1} - \dots - p^{i_k}}$$

and elementary combinatorial arguments, we immediately obtain the result (see [19]).

Method 2: By proving that it is the unique solution of the classical Dirichlet problem

$$\begin{cases} (L_k^{(i_0, \dots, i_k)})^* v(p) = 0 & \text{in } V_k \\ \lim_{p \rightarrow q} v(p) = 1, & q \in V_k^{(i_0, \dots, i_k)}, \\ \lim_{p \rightarrow q} v(p) = 0, & q \in \partial V_k \setminus V_k^{(i_0, \dots, i_k)} \setminus V_{k-1}. \end{cases}$$

□

The Probability of Having Exactly $k + 1$ Alleles

The probability of having only the particular allele A_i is (see [12]):

$$\begin{aligned} \mathbb{P}(X_t \in V_0^{(i)} | X_0 = \mathbf{p}) &= \int_{V_0^{(i)}} u_0^{(i)}(\mathbf{x}, t) d\mu_0^{(i)}(\mathbf{x}) \\ &= u_0^{(i)}(e_i, t) \\ &= p^i - \sum_{k=1}^n \sum_{m^{(k)} \geq 0} \sum_{l^{(k)} \geq 0} \sum_{|\alpha^{(k)}|=l^{(k)}} c_{m^{(k)}, l^{(k)}, \alpha^{(k)}}^{(k)} \left(x^i, X_{l^{(k)}, \alpha^{(k)}}^{(k)} \right)_k e^{-\lambda_{m^{(k)}}^{(k)} t}. \end{aligned}$$

The probability of having exactly the $(k + 1)$ alleles $\{A_0, \dots, A_k\}$ (the coexistence probability of alleles $\{A_0, \dots, A_k\}$) is (see [16, 20]):

$$\begin{aligned} \mathbb{P}(X_t \in V_k^{(i_0, \dots, i_k)} | X_0 = \mathbf{p}) &= \int_{V_k^{(i_0, \dots, i_k)}} u_k^{(i_0, \dots, i_k)}(\mathbf{x}, t) d\mu_k^{(i_0, \dots, i_k)}(\mathbf{x}) \\ &= \sum_{m \geq 0} \sum_{l \geq 0} \sum_{|\alpha|=l} c_{m, l, \alpha}^{(k)} \left(\int_{V_k^{(i_0, \dots, i_k)}} X_{m, \alpha}^{(k)}(\mathbf{x}) d\mu_k^{(i_0, \dots, i_k)}(\mathbf{x}) \right) e^{-\lambda_m^{(k)} t}. \end{aligned}$$

The α th Moments

The α th-moments are (see [15–17]):

$$\begin{aligned} m_\alpha(t) &= [u, \mathbf{x}^\alpha]_n \\ &= \int_{\frac{V}{V_n}} x^\alpha u(\mathbf{x}, t) d\mu(\mathbf{x}) \\ &= \sum_{k=0}^n \sum_{(i_0, \dots, i_k) \in I_k} \int_{V_k^{(i_0, \dots, i_k)}} \mathbf{x}^\alpha u_k^{(i_0, \dots, i_k)}(\mathbf{x}, t) d\mu_k^{(i_0, \dots, i_k)}(\mathbf{x}). \end{aligned}$$

The Probability of Heterogeneity

The probability of heterogeneity is (see [16]):

$$\begin{aligned} H_t &= (n + 1)! [u, w_n]_n \\ &= (n + 1)! (u_n, w_n)_n \quad (\text{because } w_n \text{ vanishes on the boundary}) \\ &= (n + 1)! \left(\sum_{m \geq 0} \sum_{|\alpha|=m} c_{m, \alpha}^{(n)} X_{m, \alpha}^{(n)} e^{-\lambda_{m, \alpha}^{(n)} t}, w_n X_{0, \mathbf{0}}^{(n)} \right)_n \\ &= (n + 1)! \left(c_{0, \mathbf{0}}^{(n)} X_{0, \mathbf{0}}^{(n)}, w_n X_{0, \mathbf{0}}^{(n)} \right)_n e^{-\lambda_{0, \mathbf{0}}^{(n)} t} \\ &\quad (\text{because of the orthogonality of the eigenvectors } X_{m, \alpha}^{(n)}) \\ &= H_0 e^{-\frac{(n+1)(n+2)}{2} t}. \end{aligned}$$

The Rate of Loss of One Allele in a Population Having $k + 1$ Alleles

We have the solution of the form

$$u = \sum_{k=0}^n u_k(\mathbf{x}, t) \chi_{V_k}(\mathbf{x})$$

The rate of loss of one allele in a population with $k + 1$ alleles equals the rate of decrease of

$$u_k(\mathbf{x}, t) = \sum_{m \geq 0} \sum_{l \geq 0} \sum_{|\alpha|=l} c_{m,l,\alpha}^{(k)} X_{l,\alpha}^{(k)}(x) \chi_{V_k}(x) e^{-\lambda_m^{(k)} t}$$

which is $\lambda_0^{(k)} = \frac{k(k+1)}{2}$. This means that the rate of loss of alleles in the population decreases as k gets smaller in the course of the process (see [10, 13, 16]).

Population Genetics

The Wright–Fisher model as the basic theoretical model in population genetics then can also be applied to population genetics data, although in most cases, the basic model needs to be suitably extended. For instance, single nucleotide polymorphisms (SNPs) can be modeled in this way. As the name indicates, in a SNP, there is a genetic variant in the population that differs from the rest of the population at a single nucleotide position in the genetic sequence. Thus, we may consider the background sequence and its variant as two different alleles, A_1, A_2 , and ask for the chances of the variant allele A_2 to sweep the population, for the expected when this happens or when the variant will go extinct, etc., and apply the formulae given in Sect. 3. A detailed bioinformatical analysis of SNPs in the human genome was carried out in [21]. In that paper, the effects of recombination and of varying population size were analyzed, again on the basis of the multinomial distribution as in Sect. 2.2. In particular, a population bottleneck in the human history around 40,000 years ago could be identified on the basis of the current SNP distributions in the human population. For our methods to apply, we thus need to extend the basic Wright–Fisher model to include recombination—which we shall present in another paper—as well as varying population size. Similarly, data about mitochondrial DNA [24] indicate a population bottleneck in human history. Since mitochondrial DNA is solely inherited from the mother, we can treat this as a haploid Wright–Fisher model without recombination. Again, however, to be applicable, our model needs to be extended to handle variable population sizes.

Conclusion

We have developed a new global solution concept for the Fokker–Planck equation associated with the Wright–Fisher model, and we have proved the existence and uniqueness of this solution (Theorem 2.9). From this solution, we can easily read off the properties of the considered process, like the absorption time of having $k + 1$ alleles, the probability of having exactly $k + 1$ alleles, the α th moments, the probability of heterogeneity, and the rate of loss of one allele in a population having $k + 1$ alleles.

Acknowledgements We would like to thank the anonymous referee for helpful comments, in particular suggestions concerning SNPs and mitochondrial DNA, as mentioned in Sect. 4. The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC Grant agreement no. 267087.

References

1. Buss, S.R., Clote, P.: Solving the Fisher–Wright and coalescence problems with a discrete Markov chain analysis. *Adv. Appl. Probab.* **36**(4), 1175–1197 (2004)
2. Bürger, R.: *The Mathematical Theory of Selection, Recombination, and Mutation*. John Wiley, New York (2000)
3. Epstein, C.L., Mazzeo, R.: Wright–Fisher diffusion in one dimension. *SIAM J. Math. Anal.* **42**, 568–608 (2010)
4. Epstein, C.L., Mazzeo, R.: *Degenerate Diffusion Operators Arising in Population Biology*. Princeton University Press, Princeton (2013)
5. Ethier, S.N.: A class of degenerate diffusion processes occurring in population genetics. *Commun. Pure Appl. Math.* **29**, 483–493 (1976)
6. Ethier, S.N., Kurtz, T.: *Markov Processes: Characterization and Convergence*. Wiley (1986), New York (2005)
7. Ethier, S.N., Nagylaki, T.: Diffusion approximations of Markov chains with two time scales and applications to population genetics. *Adv. Appl. Probab.* **12**, 14–49 (1980)
8. Ethier, S.N., Nagylaki, T.: Diffusion approximations of the two-locus Wright–Fisher model. *J. Math. Biol.* **27**, 17–28 (1989)
9. Ewens, W.J.: *Mathematical population genetics I. Theoretical introduction*. In: *Interdisciplinary Applied Mathematics*, 2nd edn. Springer, New York (2004)
10. Felsenstein, J.: The rate of loss of multiple alleles in finite haploid populations. *Theor. Popul. Biol.* **2**, 391–403 (1971)
11. Fisher, R.A.: On the dominance ratio. *Proc. R. Soc. Edinb.* **42**, 321–341 (1922)
12. Gill, W.: Modified fixation probability in multiple alleles models in the asymmetric sharply-peaked landscape. *J. Korean Phys. Soc.* **55**(2), 709–717 (2009)
13. Gladstien, K.: Subdivided populations the characteristic values and rate of loss of alleles. *J. Appl. Prob.* **241**, 241–248 (1977)
14. Karlin, S., Taylor, H.M.: *A Second Course in Stochastic Processes*. Academic Press, New York (1981)
15. Kimura, M.: Solution of a process of random genetic drift with a continuous model. *PNAS-USA* **41**(3), 144–150 (1955)
16. Kimura, M.: Random genetic drift in multi-allele locus. *Evolution* **9**, 419–435 (1955)
17. Kimura, M.: Random genetic drift in a tri-allelic locus; exact solution with a continuous model. *Biometrics* **12**, 57–66 (1956)
18. Lessard, S., Lahaie, P.: Fixation probability with multiple alleles and projected average allelic effect on selection. *Theor. Popul. Biol.* **75**, 266–277 (2009)
19. Littler, R.A., Good, A.J.: Ages, extinction times, and first passage probabilities for a multiallele diffusion model with irreversible mutation. *Theor. Popul. Biol.* **1**(3), 214–225 (1978)
20. Littler, R.A.: Loss of variability at one locus in a finite population. *Math. Biosci.* **25**, 151–163 (1975)
21. Marth, G., et al.: Sequence variations in the public human genome data reflect a bottlenecked population history. *Proc. Natl. Acad. Sci.* **100**.1, 376–381 (2003)
22. Nagylaki, T.: The decay of genetic variability in geographically structured populations. *PNAS* **71**, 2932–2936 (1974)
23. Tran, T.D., Hofrichter, J., Jost, J.: An introduction to the mathematical structure of the Wright–Fisher model of population genetics. *Theory Biosci.* 1–10 (2012). doi:[10.1007/s12064-012-0170-3](https://doi.org/10.1007/s12064-012-0170-3)
24. Walker, A.C., Smith, S.D., Smith, S.D.: Mitochondrial DNA and human evolution. *Nature* **325**, 1–5 (1987)
25. Wright, S.: Evolution in Mendelian populations. *Genetics* **16**, 97–159 (1931)