# Max-Planck-Institut für Mathematik in den Naturwissenschaften Leipzig 

An introduction to the mathematical structure of the WrightFisher model of population genetics
by

Tat Dat Tran, Julian Hofrichter, and Jürgen Jost


## An introduction to the mathematical structure of the Wright-Fisher model of population genetics

## Tat Dat Tran, Julian Hofrichter \& Jürgen Jost

Theory in Biosciences
ISSN 1431-7613
Theory Biosci.
DOI 10.1007/s12064-012-0170-3


Springer

Your article is published under the Creative Commons Attribution license which allows users to read, copy, distribute and make derivative works, as long as the author of the original work is cited. You may selfarchive this article on your own website, an institutional repository or funder's repository and make it publicly available immediately.

# An introduction to the mathematical structure of the Wright-Fisher model of population genetics 

Tat Dat Tran • Julian Hofrichter • Jürgen Jost

Received: 18 July 2012 / Accepted: 14 November 2012
© The Author(s) 2012. This article is published with open access at Springerlink.com


#### Abstract

In this paper, we develop the mathematical structure of the Wright-Fisher model for evolution of the relative frequencies of two alleles at a diploid locus under random genetic drift in a population of fixed size in its simplest form, that is, without mutation or selection. We establish a new concept of a global solution for the diffusion approximation (Fokker-Planck equation), prove its existence and uniqueness and then show how one can easily derive all the essential properties of this random genetic drift process from our solution. Thus, our solution turns out to be superior to the local solution constructed by Kimura.


Keywords Random genetic drift • Wright-Fisher model • Fokker-Planck equation

[^0][^1]
## Introduction

In population genetics, one considers the effects of recombination, selection, mutation, and perhaps others like migration on the distribution of alleles in a population, see e.g. (Ewens 2004; Bürger 2000; Rice 2004) as mathematical textbook references. The most basic and at the same time important model is the Wright-Fisher model for random genetic drift [developed implicitly by Fisher (1922) and explicitly by Wright (1931)]. In its simplest version-the one to be treated in the present paper-it is concerned with the evolution of the relative frequencies of two alleles at a single diploid locus in a finite population of fixed size with non-overlapping generations under the sole force of random genetic drift, without any other influences like mutations or selection. The model can be general-ised-and so can our approach-to multiple alleles, several loci, with mutations, selections, spatial population structures, etc, see the above references. To find an exact solution (for the approximating diffusion process for the probability densities of the allele frequencies described by a Fokker-Planck equation) from which the properties of the resulting stochastic process can be deduced, however, is difficult. For the basic two-allele case, this was first achieved in the important work of Kimura (1955), and he then went on to treat the case of several alleles (Kimura 1955, 1956). His solution, however, is local in the sense that it does not naturally incorporate the transitions resulting from the irreversible loss of one or several of the alleles initially present in the population. Consequently, the resulting probability distribution does not integrate to 1 , and it is difficult to read off the quantitative properties of the process from his solution.

In the present paper, we introduce and describe a new global approach. This approach is mathematically more
transparent than Kimura's scheme. We prove the existence of a unique such global solution (see Theorem 3.7), and we can deduce all desired quantities of the underlying stochastic process from our solution. The purpose of the present paper thus is to display the method in the simplest case, that of two alleles at a single locus, so that the structure becomes clear. The case of multiple alleles is presented in our companion paper (Tran et al. 2000) on the basis of the first author's thesis, and further generalisations will be systematically developed elsewhere within the mathematical framework of information geometry (Amari and Nagaoka 2000) and more specifically (Ay and Jost 2000; Jost 2000) on the basis of the second author's thesis.

## The Wright-Fisher model

We consider a diploid population of size $N$. At a given locus, there could be either one of the two alleles $A_{1}, A_{2}$. Thus, an individual can be a homozygote of type $A_{1} A_{1}$ or $A_{2} A_{2}$ or a heterozygote of type $A_{1} A_{2}$ or $A_{2} A_{1}$-but we consider the latter two as the same-at the locus in question. The population reproduces in discrete time steps, and each individual in generation $n+1$ inherits one allele from each of its parents. When a parent is a heterozygote, each allele is chosen with probability $1 / 2$. Here, for each individual in generation $t+1$, randomly two parents in generation $n$ are chosen. Thus, the alleles in generation $n+1$ are chosen by random sampling with replacement from the ones in generation $n$. The quantity of interest is the number $Y_{n}$ of alleles $A_{1}$ in the population at time $n$. This number then varies between 0 and $2 N$. The transition probability then is
$\mathbb{P}\left(Y_{n+1}=j \mid Y_{n}=i\right)=\binom{2 N}{j}\left(\frac{i}{2 N}\right)^{j}\left(1-\frac{i}{2 N}\right)^{2 N-j}$
for $i, j=0, \ldots, 2 N$.
whenever $Y_{n}$ takes the value 0 or $2 N$, that is, if either the allele $A_{1}$ or $A_{2}$ will disappear, it will stay there for all future times. Eventually, this will happen almost surely.

This is the basic model. One can then derive expressions for the expected time for the allele $A_{1}$ to become either fixed, that is, $Y_{n}=2 N$, or become extinct, $Y_{n}=0$, given its initial number $Y_{0}$.

An important idea, first applied in Wright (1945), then is to rescale time and population size via
$t=\frac{n}{2 N}, \quad X_{t}=\frac{Y_{t}}{2 N}$,
and then consider the limit $N \rightarrow \infty$. The rescaling of (2) yields a discrete Markov chain $X_{t}$ valued in $\left\{0, \frac{1}{2 N}, \ldots, 1\right\}$ with $t=1$ now corresponding to $2 N$ generations. One
readily verifies that the expectation values for the variation across generations satisfy

$$
\begin{align*}
X_{0} & =p=\frac{i_{0}}{2 N} \\
\mathbb{E}\left(\delta X_{t}\right) & =0  \tag{3}\\
\mathbb{E}\left(\delta X_{t}\right)^{2} & =X_{t}\left(1-X_{t}\right) \delta t \\
\mathbb{E}\left(\delta X_{t}\right)^{k} & =o(\delta t) \text { for } k \geq 3
\end{align*}
$$

A basic idea of our approach is to consider the $k$ th moment $m_{k}(t)$ of the distribution about zero at the $(2 N t)$ th generation, i.e.
$m_{k}(t)=\mathbb{E}\left(X_{t}\right)^{k}$
We have
$m_{k}(t+1)=\mathbb{E}\left(X_{t}+\delta X_{t}\right)^{k}$
Expanding the right hand side and noting (3) we obtain the following recursion formula
$m_{k}(t+1)=\left\{1-\frac{k(k-1)}{2}\right\} m_{k}(t)+\frac{k(k-1)}{2} m_{k-1}(t)$
when we assume that the population number $N$ is so large that we can neglect all terms of order at least $\frac{1}{N^{2}}$. Under this assumption, the moments change very slowly per generation and we can replace the above system (6) by the system of differential equations
$\dot{m}_{k}(t)=-\frac{k(k-1)}{2} m_{k}(t)+\frac{k(k-1)}{2} m_{k-1}(t)$,
where the dot denotes a derivative w.r.t. the variable $t$.
These formulae now guide us in finding a continuous process that well approximates the above discrete process. We seek a continuous Markov process $\left\{X_{t}\right\}_{t \geq 0}$ valued in [ 0,1 ] with the same conditions as (3) and (7). The conditions (3) imply (see for example Ewens 2004, p. 137, for a derivation) that the probability density function $u(x, t)$ of this continuous process is a solution of the Fokker-Planck (Kolmogorov forward) equation
$\left\{\begin{array}{l}u_{t}(x, t)=\frac{1}{2} \frac{\partial^{2}}{\partial x^{2}}(x(1-x) u(x, t)) \text { in }(0,1) \times(0, \infty), \\ u(x, 0)=\delta_{p}(x) \text { in }(0,1)\end{array}\right.$
where we now use the notation $u_{t}:=\frac{\partial}{\partial t} u(x, t)$ for the partial derivative w.r.t. the time variable $t$. The coefficient $x(1-x)$ in (8) comes from (3) and $\delta_{p}$ denotes the Dirac delta function at $p$. For the definition of this delta function, we use the product
$(f, g):=\int_{0}^{1} f(x) g(x) \mathrm{d} x$
for square integrable functions $f, g:[0,1] \rightarrow \mathbb{R}$ on the unit interval (this will be described in more detail in "Existence and uniqueness of solutions"), and we then put

$$
\left(\delta_{p}, \phi\right):=\phi(p)
$$

whenever $\phi:[0,1] \rightarrow \mathbb{R}$ is a continuous function. ${ }^{1}$
Let us also explain the interpretation of (8) for those not sufficiently versed in this mathematical formalism. The initial condition $u(x, 0)=\delta_{p}(x)$ then simply says that at time 0 , the relative frequency of allele $A_{1}$ is precisely $p$, without any uncertainty (this assumption is not essential, however, and the scheme works also for more general initial condition involving uncertainty about the initial distribution of the alleles). Subsequently, this allele frequence evolves stochastically, according to the equation $u_{t}(x, t)=\frac{1}{2} \frac{\partial^{2}}{\partial x^{2}}(x(1-x) u(x, t))$, and therefore, for $t>0$, we no longer know the precise value of this relative frequency, but only its probability density given by $u(x, t)$. That is, for every $x$, the probability density that the allele frequency at time $t$ has the value $x$ is given by $u(x, t)$.

In the continuum limit, the $k$ th moment becomes $\int_{0}^{1} u(x, t) x^{k} \mathrm{~d} x$, and the condition (7) then implies

$$
\begin{aligned}
\left(u_{t}, x^{k}\right) & =\int_{0}^{1} \frac{\partial}{\partial t} u(x, t) x^{k} \mathrm{~d} x \\
& =\frac{\mathrm{d}}{\mathrm{~d} t} \int_{0}^{1} u(x, t) x^{k} \mathrm{~d} x \\
& =\left(u,-\frac{k(k-1)}{2} x^{k}+\frac{k(k-1)}{2} x^{k-1}\right) \\
& =\left(u, \frac{1}{2} x(1-x) \frac{\partial^{2}}{\partial x^{2}}\left(x^{k}\right)\right), \quad \forall k \geq 0
\end{aligned}
$$

Since the polynomials are dense in the space of (square integrable) functions, this yields
$\left(u_{t}, \phi\right)=\left(u, \frac{1}{2} x(1-x) \frac{\partial^{2}}{\partial x^{2}} \phi\right)$
for all square integrable functions $\phi:[0,1] \rightarrow \mathbb{R}$ that are twice differentiable in the open interval $(0,1)$.

This leads to our concept of a solution of the FokkerPlanck equation in

[^2]Definition 2.1 We call $u \in H$ a solution of the FokkerPlanck equation associated with the Wright-Fisher model if
$u_{t}=L u$ in $(0,1) \times(0, \infty)$,
$u(x, 0)=\delta_{p}(x)$ in $(0,1) ;$
$\left(u_{t}, \phi\right)=\left(u, L^{*} \phi\right), \quad \forall \phi \in H_{0}$,
for all square integrable functions $\phi:[0,1] \rightarrow \mathbb{R}$ that are twice differentiable in the open interval $(0,1)$, with the differential operator
$L u(x):=\frac{1}{2} \frac{\partial^{2}}{\partial x^{2}}(x(1-x) u(x))$
and its formal adjoint
$L^{*} \phi(x)=\frac{1}{2} x(1-x) \frac{\partial^{2}}{\partial x^{2}} \phi(x)$.
This solution concept will allow us to prove the existence of a unique solution from which we can then derive all features of interest of the Wright-Fisher process. We should point out that (12) is not just the integration by parts of (10), but also includes the boundary behaviour (of course, this may not be overt, but the mathematical trick here is to represent this boundary behaviour in an implicit form best suited for formal manipulation). It, thus, reflects transitions from the presence of both alleles to the irreversible loss of one of them. This is the crucial difference to Kimura's (1955) solution concept and the key for the properties of our solution.

## Existence and uniqueness of solutions

We shall now apply a familiar mathematical scheme for the construction of a solution of a differential equation, an expansion in terms of eigenfunctions of the differential operator involved. For our problem, as formalised in Definition 2.1, these eigenfunctions can be constructed from a classical family of polynomials, the Gegenbauer polynomials, which we shall now introduce.

## Preliminaries

For the sequel, we shall need some more notation. We need the function spaces

$$
\begin{aligned}
H_{0}:= & C^{\infty}[0,1] \\
H:= & \{f:[0,1] \rightarrow[0, \infty] \text { measurable with } \\
& \left.\int_{[0,1]} f(x) g(x) \mathrm{d} x<\infty, \quad \forall g \in H_{0}\right\}
\end{aligned}
$$

with the scalar product

$$
(f, g):=\int_{[0,1]} f(x) g(x) \mathrm{d} x, \quad \forall f \in H, g \in H_{0}
$$

To construct solutions in terms of expansions, we shall need a special case of the Gegenbauer polynomials [named after Leopold Gegenbauer (1849-1903)]. ${ }^{2}$ The polynomials $Y_{m}(z)$ we need are defined in terms of their generating function
$\frac{1}{\left(1-2 z t+t^{2}\right)^{\frac{3}{2}}}=\sum_{m \geq 0} Y_{m}(z) t^{m}$.
Lemma 3.1 (Suetin 2001)

- The Gegenbauer polynomials satisfy the recurrence relation

$$
\begin{aligned}
& Y_{0}(z)=1 \\
& Y_{1}(z)=3 z \\
& Y_{n}(z)=\frac{1}{n}\left[2 z\left(n+\frac{1}{2}\right) Y_{n-1}(z)-(n+1) Y_{n-2}(z)\right]
\end{aligned}
$$

- The Gegenbauer polynomials solve the differential equation

$$
\begin{equation*}
\left(1-z^{2}\right) y^{\prime \prime}-4 z y^{\prime}+n(n+3) y=0 \tag{15}
\end{equation*}
$$

Lemma 3.2 [Abramowitz (1965), p. 774] The polynomials $Y_{m}$ are orthogonal polynomials on the interval $[-1,1]$ with respect to the weight function $\left(1-z^{2}\right)$ :

$$
\begin{equation*}
\int_{-1}^{1}\left(1-z^{2}\right) Y_{m}(z) Y_{n}(z) \mathrm{d} z=0 \quad \text { for } m \neq n \tag{16}
\end{equation*}
$$

Auxiliaries

Lemma 3.3 For all $m \geq 0$ we have
$L X_{m}=-\lambda_{m} X_{m}$, in $H_{0}$
with
$\lambda_{m}:=\frac{(m+1)(m+2)}{2}$.
Proof Putting $z=1-2 x$ implies that
$Y_{m}(z):=X_{m}(x)$
is a Gegenbauer polynomial and therefore solves (15),

$$
\left(1-z^{2}\right) \frac{\partial^{2}}{\partial z^{2}} Y_{m}(z)-4 z \frac{\partial}{\partial z} Y_{m}(z)+m(m+3) Y_{m}(z)=0
$$

[^3]This is equivalent to

$$
\begin{aligned}
& x(1-x) \frac{\partial^{2}}{\partial x^{2}} X_{m}(x)-2(1-2 x) \frac{\partial}{\partial x} X_{m}(x) \\
& \quad+m(m+3) X_{m}(x)=0, \Longleftrightarrow \frac{\partial^{2}}{\partial x^{2}}(x(1-x) \\
& \left.\quad \times X_{m}(x)\right)=-(m+1)(m+2) X_{m} \\
& \quad \Longleftrightarrow L X_{m}=-\lambda_{m} X_{m}
\end{aligned}
$$

This completes the proof.
In the sequel, we shall use the abbreviation
$w(x):=x(1-x)$.
Lemma 3.4 If $X$ is an eigenvector of $L$ corresponding to the eigenvalue $\lambda$ then $w X$ is an eigenvector of $L^{*}$ corresponding to the eigenvalue $\lambda$.

Proof Assume that $X$ is an eigenvector of $L$ for the eigenvalue $\lambda$, i.e.
$\frac{1}{2} \frac{\partial^{2}}{\partial x^{2}}(w X)=-\lambda X$.
Multiplying both sides by $w$ yields
$L^{*}(w X)=\frac{1}{2} w \frac{\partial^{2}}{\partial x^{2}}(w X)=-\lambda(w X)$.
This completes the proof.
Lemma 3.5 The spectrum of the operator $L$ is
$\operatorname{Spec}(L)=\bigcup_{m \geq 0}\left\{\lambda_{m}=\frac{(m+1)(m+2)}{2}\right\}:=\Lambda$,
and the eigenvector of $L$ corresponding to $\lambda_{m}$ is the Gegenbauer polynomial $X_{m}(x)$ (up to a constant).

Proof From Lemma 3.3 we have $L\left(X_{m}\right)=-\lambda_{m} X_{m}$ in $H_{0}$. $\operatorname{So}, \Lambda \subseteq \operatorname{Spec}(L)$. Conversely, we shall prove that $\lambda \notin \Lambda$ is not an eigenvalue of $L$. In fact, assume that there is some $X \in H_{0}$ with $L X=-\lambda X \in H_{0}$. Because $\left\{X_{m}\right\}_{m} \geq 0$ is a basis of $H_{0}$, we can represent $X$ by $X=\sum_{m=0}^{\infty} d_{m} X_{m}$. Then

$$
\begin{aligned}
\sum_{m=0}^{\infty} d_{m}\left(-\lambda_{m} X_{m}\right) & =\sum_{m=0}^{\infty} d_{m} L\left(X_{m}\right)=L\left(\sum_{m=0}^{\infty} d_{m} X_{m}\right) \\
& =-\lambda \sum_{m=0}^{\infty} d_{m} X_{m}
\end{aligned}
$$

For any $n \geq 0$, we can multiply this relation by $w X_{n}$ and then integrate on $[0,1]$. From the orthogonality (16) with respect to the weight function $w$, we obtain
$d_{n} \lambda_{n}\left(X_{n}, w X_{n}\right)=d_{n} \lambda\left(X_{n}, w X_{n}\right)$.
Because $\left(X_{n}, w X_{n}\right) \neq 0$ and $\lambda \neq \lambda_{n}$, then $d_{n}=0, \forall n \geq 0$. Therefore, $X=0$, i.e. $\lambda$ is not an eigenvalue of $L$. Thus $\operatorname{Spec}(L)=\Lambda$.

Similarly, if $X$ is an eigenvector of $L$ for the eigenvalue $\lambda_{m}$, we will prove that $X=c X_{m}$. In fact, representing $X=$ $\sum_{n=0}^{\infty} d_{n} X_{n}$, it follows that

$$
\begin{aligned}
\sum_{n=0}^{\infty} d_{n}\left(-\lambda_{n} X_{n}\right) & =\sum_{n=0}^{\infty} d_{n} L\left(X_{n}\right)=L\left(\sum_{n=0}^{\infty} d_{n} X_{n}\right) \\
& =-\lambda_{m} \sum_{n=0}^{\infty} d_{n} X_{n}
\end{aligned}
$$

For any $k \geq 0$, we multiply this relation by $w X_{k}$ and then integrate on $[0,1]$ to obtain
$d_{k} \lambda_{k}\left(X_{k}, w X_{k}\right)=d_{k} \lambda_{m}\left(X_{k}, w X_{k}\right)$.
Because $\left(X_{k}, w X_{k}\right) \neq 0$ and $\lambda_{m} \neq \lambda_{k}$ for all $k \neq m$, then $d_{k}=0, \forall k \neq m$. Hence $X=d_{m} X_{m}$. This completes the proof.

Construction of the solution

In this subsection, we construct the solution and prove its uniqueness. We shall firstly find the general solution of the Fokker-Planck equation (10) by the separation of variables method. Then we shall construct a solution depending on parameters. We shall use $(11,12)$ to determine the parameters. Finally, we shall verify the solution.

Step 1 Assume that $u_{0}(x, t)=X(x) T(t)$ is a solution of the Fokker Planck equation (10). Then we have
$\frac{T_{t}}{T}=\frac{L X}{X}=-\lambda$
which implies that $\lambda$ is a constant which is independent of $t, x$. From Lemma 3.5, we obtain the general solution of the equation (10) as
$u_{0}(x, t)=\sum_{m \geq 0} c_{m} X_{m}(x) \mathrm{e}^{-\lambda_{m} t}$.
Remark $3.6 u_{0}$ is the same as Kimura's solution (see for example Kimura 1955a,b).

Step 2 The general solution $u \in H$ of (10) then is

$$
\begin{align*}
u(x, t)= & \sum_{m \geq 0} c_{m}\left(X_{m}(x)+a_{m, 0} \delta_{0}(x)+a_{m, 1} \delta_{1}(x)\right) \mathrm{e}^{-\lambda_{m} t} \\
& +b_{0} \delta_{0}(x)+b_{1} \delta_{1}(x) \tag{17}
\end{align*}
$$

where $\delta_{0}$ and $\delta_{1}$ are the Dirac delta functionals at 0 and 1 .
Step 3 Checking condition (12) with $\phi=1, \phi=x, \phi$ $=w X_{n}$ yields

$$
\begin{aligned}
\left(u_{t}, 1\right) & =\left(u, L^{*}(1)\right)=0 \\
\left(u_{t}, x\right) & =\left(u, L^{*}(x)\right)=0 \\
\left(u_{t}, w X_{n}\right) & =\left(u, L^{*}\left(w X_{n}\right)\right)=-\lambda_{n}\left(u, w X_{n}\right) .
\end{aligned}
$$

With condition (11), we then obtain

$$
\begin{aligned}
1= & (u(\cdot, 0), 1)=(u(\cdot, \infty), 1)=b_{0}+b_{1}, \\
p= & (u(\cdot, 0), x)=(u(\cdot, \infty), x)=b_{1}, \\
1= & (u(\cdot, 0), 1)=(u, 1)=\sum_{m \geq 0} c_{m} \\
& \times\left(\int_{[0,1]} X_{m}(x) \mathrm{d} x+a_{m, 0}+a_{m, 1}\right) \mathrm{e}^{-\lambda_{m} t}+b_{0}+b_{1}, \\
p= & (u(\cdot, 0), x)=(u, x)=\sum_{m \geq 0} c_{m} \\
& \times\left(\int_{[0,1]} x X_{m}(x) \mathrm{d} x+a_{m, 1}\right) \mathrm{e}^{-\lambda_{m} t}+b_{1},
\end{aligned}
$$

and

$$
\begin{align*}
& \left(u, w X_{n}\right)=\left(u(\cdot, 0), w X_{n}\right) \mathrm{e}^{-\lambda_{n} t}=w(p) X_{n}(p) \mathrm{e}^{-\lambda_{n} t} \\
& \Longleftrightarrow c_{n}\left(X_{n}, w X_{n}\right) \mathrm{e}^{-\lambda_{n} t}=w(p) X_{n}(p) \mathrm{e}^{-\lambda_{n} t} \tag{18}
\end{align*}
$$

Therefore we have all parameters

$$
\begin{align*}
b_{1} & =p ; b_{0}=1-p \\
a_{m, 1} & =-\int_{[0,1]} x X_{m}(x) \mathrm{d} x ; a_{m, 0}=-\int_{[0,1]}(1-x) X_{m}(x) \mathrm{d} x \\
c_{n} & =\frac{w(p) X_{n}(p)}{\left(X_{n}, w X_{n}\right)} \tag{19}
\end{align*}
$$

It follows that the solution should be

$$
\begin{align*}
u(x, t)= & \sum_{m \geq 0} c_{m} X_{m}(x) \mathrm{e}^{-\lambda_{m} t} \\
& +\left\{1-p+\sum_{m \geq 0} c_{m} a_{m, 0} \mathrm{e}^{-\lambda_{m} t}\right\} \delta_{0}(x) \\
& +\left\{p+\sum_{m \geq 0} c_{m} a_{m, 1} \mathrm{e}^{-\lambda_{m} t}\right\} \delta_{1}(x) \tag{20}
\end{align*}
$$

where $X_{m}(x)$ is a Gegenbauer polynomial,

$$
\begin{align*}
\lambda_{m} & =\frac{(m+1)(m+2)}{2} \\
a_{m, 0} & =-\int_{\Omega_{1}}(1-x) X_{m}(x) \mathrm{d} x=-\frac{1}{2} \\
a_{m, 1} & =-\int_{\Omega_{1}} x X_{m}(x) \mathrm{d} x=(-1)^{m+1} \frac{1}{2} \\
c_{m} & =\frac{w(p) X_{m}(p)}{\left(X_{m}, w X_{m}\right)}=\frac{8 w(p) X_{m}(p)(m+3 / 2)}{(m+1)(m+2)} \tag{21}
\end{align*}
$$

Step 4 We will prove the constructed solution $u$ satisfies conditions ( $10,11,12$ ). In fact, because in ( 0,1 ), $u=u_{0}$, it
is clear that $u$ satisfies the Fokker Planck equation (10). Moreover, from the representation (20), we have

$$
\begin{align*}
(u, 1)= & \sum_{m=0}^{\infty} c_{m}\left(\int_{[0,1]} X_{m}(x) \mathrm{d} x+a_{m, 0}+a_{m, 1}\right) \mathrm{e}^{-\lambda_{m} t} \\
& +1-p+p=1, \\
(u, x)= & \sum_{m=0}^{\infty} c_{m}\left(\int_{[0,1]} x X_{m}(x) \mathrm{d} x+a_{m, 1}\right) \mathrm{e}^{-\lambda_{m} t} \\
& +p=p, \\
\left(u, w X_{n}\right)= & c_{n}\left(X_{n}, w X_{n}\right) \mathrm{e}^{-\lambda_{n} t}=w(p) X_{n}(p) \mathrm{e}^{-\lambda_{n} t} . \tag{22}
\end{align*}
$$

Thus,

$$
\begin{align*}
(u(\cdot, 0), 1) & =1=\left(\delta_{p}, 1\right) \\
(u(\cdot, 0), x) & =p=\left(\delta_{p}, x\right)  \tag{23}\\
\left(u(\cdot, 0), w X_{n}\right) & =w(p) X_{n}(p)=\left(\delta_{p}, w X_{n}\right) .
\end{align*}
$$

Because $\left\{1, x,\left\{w X_{n}\right\}_{n} \geq 0\right\}$ is also a basis of $H_{0}$, it follows that
$(u(\cdot, 0), \phi)=\left(\delta_{p}, \phi\right), \quad \forall \phi \in H_{0}$,
i.e. $u(\cdot, 0)=\delta_{p} \in H$, i.e. $u$ satisfies the condition (11).

Finally, from (22) we have

$$
\begin{align*}
\left(u_{t}, 1\right)=0 & =\left(u, L^{*}(1)\right) \\
\left(u_{t}, x\right)=0 & =\left(u, L^{*}(x)\right) \\
\left(u_{t}, w X_{n}\right) & =w(p) X_{n}(p)\left(-\lambda_{n}\right) \mathrm{e}^{-\lambda_{n} t} \\
& =-\lambda_{n}\left(u, w X_{n}\right)=\left(u, L^{*}\left(w X_{n}\right)\right) . \tag{24}
\end{align*}
$$

Because $L^{*}$ is linear and $\left\{1, x,\left\{w X_{n}\right\}_{n \geq 0}\right\}$ is also a basis of $H_{0}$, it follows that
$\left(u_{t}, \phi\right)=\left(u, L^{*}(\phi)\right), \quad \forall \phi \in H_{0}$,
i.e. $u$ satisfies the condition (12).

Therefore, $u$ is a solution of the Fokker-Planck equation associated with the Wright-Fisher model, indeed.

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 Wright-Fisher model. Then $u=u_{1}-u_{2}$ satisfies

$$
\begin{aligned}
u_{t} & =L u \text { in }(0,1) \times(0, \infty), \\
u(x, 0) & =0 \text { in }(0,1) ; \\
\left(u_{t}, \phi\right) & =\left(u, L^{*} \phi\right), \forall \phi \in H_{0} .
\end{aligned}
$$

## Therefore

$$
\begin{aligned}
\left(u_{t}, 1\right) & =\left(u, L^{*}(1)\right)=0, \\
\left(u_{t}, x\right) & =\left(u, L^{*}(x)\right)=0, \\
\left(u_{t}, w X_{n}\right) & =\left(u, L^{*}\left(w X_{n}\right)\right)=-\lambda_{n}\left(u, w X_{n}\right) .
\end{aligned}
$$

Therefore

$$
\begin{aligned}
(u, 1) & =(u(\cdot, 0), 1)=0 \\
(u, x) & =(u(\cdot, 0), x)=0 \\
\left(u, w X_{n}\right) & =\left(u(\cdot, 0), w X_{n}\right) \mathrm{e}^{-\lambda_{n} t}=0
\end{aligned}
$$

Because $\left\{1, x,\left\{w X_{n}\right\}_{n} \geq 0\right\}$ is also a basis of $H_{0}$, it follows that $u=0 \in H$.

Altogether, we obtain our main result.
Theorem 3.7. The Fokker-Planck equation associated with Wright-Fisher model possesses a unique solution.

This new solution continuously deforms the initial state $\delta_{p}(x)$ (the allele $A_{1}$ has relative frequency $p$ ) to $p \delta_{1}(x)+(1-p) \delta_{0}(x)$ (allele $A_{1}$ is fixed with probability $p$ and $A_{2}$ is fixed with probability $1-p$ ) as time proceeds from 0 to $\infty$. In fact, the sequence $\left\{u_{m}(x, t)\right\}_{m} \geq 0$ satisfying

$$
\begin{align*}
u_{m}(x, t)= & \sum_{i=0}^{m} c_{i} X_{i}(x) \mathrm{e}^{-\lambda_{i} t} \\
& +\left\{1-p+\sum_{i=0}^{m} c_{i}\left(-\frac{1}{2}\right) \mathrm{e}^{-\lambda_{i} t}\right\} \frac{m}{\sqrt{2 \pi}} \mathrm{e}^{-x^{2} m^{2} / 2} \\
& +\left\{p+\sum_{i=0}^{m} c_{i}\left(\frac{(-1)^{i+1}}{2}\right) \mathrm{e}^{-\lambda_{i} t}\right\} \frac{m}{\sqrt{2 \pi}} \mathrm{e}^{-(1-x)^{2} m^{2} / 2} \tag{25}
\end{align*}
$$

tends to $u$ for $m \rightarrow \infty$. Therefore, we can visualise the asymptotic behaviour with the help of Mathematica (Fig. 1).

```
w[x_]:= x (1 - x)
X[m_, x_]:= GegenbauerC[m, 3/2, 1-2 x]
c[m_]:= 8w[p] X[m, p] (m+3/2)/{(m+1) (m + 2) }
ld[\mp@subsup{m}{-}{\prime}]:= (m+1)(m+2)/2
u1[m_, x_, t_]:=Sum[c[i] X[i, x] Exp[-(i+1)(i+2) t/2],{i,0,m}]
f[\mp@subsup{n}{-}{}, x_] := n/Sqrt[2 Pi] Exp[-x^2 n^2/2]
g[n_, x_] := n/Sqrt[2 Pi] Exp[-(x-1)^2 n^2/2]
u2[m-, x- , t_] :={1-p+Sum[c[i](-1/2)Exp[-(i+1)(i+2) t/2],{i,0,m}]}f[m, x]
u3[m_ , x - , t_]:={p+Sum[c[i] (-1)^{i+1}1/2 Exp[-(i+1)(i+2)t/2],{i,0,m}]}g[m,x]
u[m_, x_, t_] := u1[m, x, t] + u2[m, x, t] + u3[m, x, t]
Plot3D[u[50, x, t] /. p > .4, {x, 0, 1}, {t, .01, 20}]
```

This behaviour coincides with the discrete one (Figs. 2, 3):

```
a[n_, i_, j_] := If[(i<= 0 && j<= 0) || (i >= n && j >= n), 1,
Binomial[n, j] (i/n)^j (1 - i/n)^(n - j)]
Matran[n_] := Table[a[n-1, j - 1, i - 1], {i, n}, {j, n}]
MP[\mp@subsup{n}{-}{\prime},\mp@subsup{k}{-}{\prime}]:= MatrixPower[Matran[n], k]
XO[\mp@subsup{n}{-}{\prime}, p_] := Table[If[i< p || i>p, 0, 1], {i, n}, {j, 1}]
X[\mp@subsup{n}{-}{\prime},\mp@subsup{k}{-}{\prime},\mp@subsup{p}{-}{\prime}]:= MP[n,k].XO[n, p]
G[\mp@subsup{n}{-}{\prime}, k_, p_] := BarChart[Table[X[n, k, p][[i]], {i, 1, n}]]
Table[G[9, k, 5], {k, 0, 18}], G[9, 32, 5]
```

[^4]

Fig. 1 Behaviour of the new solution from $\delta_{p}$ to $p \delta_{1}+(1-p) \delta_{0}$ in time with $p=0.4$

## Applications

Our global solution readily yields the quantities of interest of the evolution of the process $\left(X_{t}\right)_{t} \geq 0$ such as the expectation and the second moment of the absorption time, $m$ th moments, fixation probabilities, the probability of coexistence, or the probability of heterogeneity.

## Absorption time

Let $V_{0}:=\{0,1\}$ be the domain representing a population of 1 allele. Here, 0 corresponds to the loss of $A_{1}$, that is, the fixation of $A_{2}$, and 1 corresponds to the opposite situation. Either of these irreverible events is called an absorption.

We denote by $T_{2}^{1}(p)=\inf \left\{t>0: X_{t} \in V_{0} \mid X_{0}=p\right\}$ the first time when the population has only 1 allele left, that is, when absorption occurs. $T_{2}^{1}(p)$ is a continuous random variable valued in $[0, \infty)$ with probability density function denoted by $\phi(t, p)$. $V_{0}$ is invariant (absorption set) under the process $X_{t}$, i.e. if $X_{s} \in V_{0}$ then $X_{t} \in V_{0}$ for all $t \geq s$. We have
$\mathbb{P}\left(T_{2}^{1}(p) \leq t\right)=\mathbb{P}\left(X_{t} \in V_{0} \mid X_{0}=p\right)$
It follows that
$\phi(t, p)=\int_{V_{0}} \frac{\partial}{\partial t} u(x, p, t) \mathrm{d} x$.
Therefore the expectation of the absorption time for having only one allele is

$$
\begin{aligned}
\mathbb{E}\left(T_{2}^{1}(p)\right)= & \int_{0}^{\infty} t \phi(t, p) \mathrm{d} t \\
= & \int_{V_{0}}\left(\int_{0}^{\infty} t \frac{\partial}{\partial t} u(x, t) \mathrm{d} t\right) \mathrm{d} x \\
= & \sum_{m=0}^{\infty} \sum_{V_{0}}\left(-\lambda_{m}\right) c_{m}\left(X_{m}(x)+a_{m, 0} \delta_{e_{0}}(x)\right. \\
& \left.+a_{m, 1} \delta_{e_{1}}(x)\right)\left(\int_{0}^{\infty} t \mathrm{e}^{-\lambda_{m} t} \mathrm{~d} t\right) \mathrm{d} x \\
= & -\sum_{m=0}^{\infty} \frac{1}{\lambda_{m}} c_{m}\left(a_{m, 0}+a_{m, 1}\right) \\
= & \sum_{m=0}^{\infty} \frac{1}{\lambda_{2 m}} c_{2 m} \\
= & \sum_{m=0}^{\infty} 16 p(1-p)(2 m+3 / 2) /(2 m+1)^{2} \\
& \times(2 m+2)^{2} X_{2 m}(p) .
\end{aligned}
$$

and its second moment is

$$
\begin{align*}
\mathbb{E}\left(T_{2}^{1}(p)\right)^{2}= & \int_{0}^{\infty} t^{2} \phi(t, p) \mathrm{d} t \\
= & \int_{V_{0}} \int_{0}^{\infty} t^{2} \frac{\partial}{\partial t} u(x, p, t) \mathrm{d} t \mathrm{~d} x \\
= & \sum_{m=0}^{\infty} \int_{V_{0}}\left(-\lambda_{m}\right) c_{m}\left(X_{m}(x)+a_{m, 0} \delta_{e_{0}}(x)\right. \\
& \left.+a_{m, 1} \delta_{e_{1}}(x)\right) \int_{0}^{\infty} t^{2} \mathrm{e}^{-\lambda_{m} t} \mathrm{~d} t \mathrm{~d} x \\
= & -\sum_{m=0}^{\infty} \frac{2}{\lambda_{m}^{2}} c_{m}\left(a_{m, 0}+a_{m, 1}\right) \\
= & \sum_{m=0}^{\infty} \frac{2}{\lambda_{2 m}^{2}} c_{2 m} \\
= & \sum_{m=0}^{\infty} 64 p(1-p)(2 m+3 / 2) /(2 m+1)^{3} \\
& \times(2 m+2)^{3} X_{2 m}(p) . \tag{26}
\end{align*}
$$

Remark $4.1 \mathbb{E}\left(T_{2}^{1}(p)\right)=-2\{p \ln (p)+(1-p) \ln (1-p)\}$ is the unique solution of the one-dimensional boundary value problem

$$
\left\{\begin{array}{c}
L v=-1, \text { in }(0,1) \\
v(0)=v(1)=0
\end{array}\right.
$$



Fig. 2 Behaviour of the discrete solution in time $k=0,1, \ldots, 18$ and $k=32$ with $p=0.5$


Fig. 3 Behaviour of the discrete solution in time $k=0,1, \ldots, 18$ and $k=30$ with $p=0.25$


Fig. 4 Comparison results of expectation of the absorption time

We easily check that this agrees with our formula above by using Mathematica (Fig. 4):

```
X[m_, x_]:=GegenbauerC[m,3/2,1-2 x]
c[m_, p_]:=16 p(1-p)(2m+3/2)/{(2m+1)^2(2m+2)^2}
f[p_] := -2 p Log[p] - 2(1 - p) Log[1 - p]
Plot[{Sum[c[m,p] X[2 m,p],{m,0,200}],f[p]},{p,0.1,0.9}, AxesOrigin->{0,0}]
```

$n$th moments

By induction, it is easy to prove that

$$
\int_{[0,1]} x^{n} X_{m-1}(x) \mathrm{d} x=(-1)^{m} \frac{1}{2}\left\{\frac{(n-1) \ldots(n-m)}{(n+1) \ldots(n+m)}-1\right\} .
$$

Therefore, the $n$th moment is

$$
\begin{aligned}
m_{n}(t)= & \left(u, x^{n}\right) \\
= & \sum_{m=0}^{\infty} c_{m}\left(\int_{[0,1]} x^{n} X_{m}(x) \mathrm{d} x\right) \mathrm{e}^{-\lambda_{m} t} \\
& +\left(p+\sum_{m=0}^{\infty} c_{m} a_{m, 1} \mathrm{e}^{-\lambda_{m} t}\right) \\
= & p+\sum_{i=1}^{\infty} c_{i-1}\left(\int_{[0,1]} x^{n} X_{i-1}(x) \mathrm{d} x+a_{i-1,1}\right) \mathrm{e}^{-\lambda_{i-1} t} \\
= & p+\sum_{i=1}^{\infty} \frac{2(2 i+1)}{i(i+1)} p(1-p)(-1)^{i} \\
& \times X_{i-1}(p) \frac{(n-1) \ldots(n-i)}{(n+1) \ldots(n+i)} \mathrm{e}^{-\frac{i(i+1)}{2} t} .
\end{aligned}
$$

This $n$th moment coincides with Kimura's (1955) one.

Fixation probabilities and probability of coexistence of 2 alleles

The fixation probability for $A_{2}\left(\operatorname{loss}\right.$ of $\left.A_{1}\right)$ is

$$
\begin{aligned}
\mathbb{P}\left(X_{t}=0 \mid X_{0}=p\right) & =\int_{\{0\}} u(x, t) \mathrm{d} x \\
& =1-p+\sum_{m=0}^{\infty} c_{m} a_{m, 0} \mathrm{e}^{-\lambda_{m} t} \\
& =1-p-\frac{1}{2} \sum_{m=0}^{\infty} \frac{8 w(p) X_{m}(p)(m+3 / 2)}{(m+1)(m+2)} \mathrm{e}^{-\lambda_{m} t}
\end{aligned}
$$

Analogously, the fixation probability of $A_{1}$ is

$$
\begin{aligned}
\mathbb{P}\left(X_{t}=1 \mid X_{0}=p\right)= & \int_{\{1\}} u(x, t) \mathrm{d} x \\
= & p+\sum_{m=0}^{\infty} c_{m} a_{m, 1} \mathrm{e}^{-\lambda_{m} t} \\
= & p-\frac{1}{2} \sum_{m=0}^{\infty}(-1)^{m} \\
& \times \frac{8 w(p) X_{m}(p)(m+3 / 2)}{(m+1)(m+2)} \mathrm{e}^{-\lambda_{m} t}
\end{aligned}
$$

The probability of coexistence of the 2 alleles $A_{1}, A_{2}$ therefore is

$$
\begin{aligned}
\mathbb{P}\left(X_{t} \in(0,1) \mid X_{0}=p\right) & =\int_{(0,1)} u(x, t) \mathrm{d} x \\
& =\sum_{m=0}^{\infty} c_{m} \int_{(0,1)} X_{m}(x) \mathrm{d} x \mathrm{e}^{-\lambda_{m} t} \\
& =\sum_{m=0}^{\infty} c_{2 m} \mathrm{e}^{-\lambda_{2 m} t} \\
& =\sum_{m=0}^{\infty} \frac{8 w(p) X_{2 m}(p)(2 m+3 / 2)}{(2 m+1)(2 m+2)} \mathrm{e}^{-\lambda_{2 m} t}
\end{aligned}
$$

These three probabilities sum to 1 , as they should.
We consider their behaviour for $p=0.3$ and $p=0.5$ (Figs. 5, 6):


Fig. $5 p=0.3$


Fig. $6 p=0.5$

Remark 4.2
(i) $\quad \mathbb{P}\left(X_{t} \in[0,1] \mid X_{0}=p\right)=\mathbb{P}\left(X_{t}=0 \mid X_{0}=p\right)$ $+\mathbb{P}\left(X_{t}=1 \mid X_{0}=p\right)+\mathbb{P}\left(X_{t} \in(0,1) \mid X_{0}=p\right)=1 ;$
(ii) $\mathbb{P}\left(X_{t}=0 \mid X_{0}=p\right)$ and $\mathbb{P}\left(X_{t}=1 \mid X_{0}=p\right)$ increase quickly in $t \in(0,5)(10 \mathrm{~N}$ generations) from 0 and then tend slowly to $1-p$ and $p$, respectively;
(iii) When $p=0.5$, the situation is symmetric between the two alleles, that is, $\mathbb{P}\left(X_{t}=0 \mid X_{0}=0.5\right)=$ $\mathbb{P}\left(X_{t}=1 \mid X_{0}=0.5\right)$.

## Heterogeneity

The probability of heterogeneity is

$$
\begin{aligned}
H_{t} & =\int_{[0,1]} 2 x(1-x) u(x, t) \mathrm{d} x \\
& =2\left(u, w X_{0}\right) \\
& =2\left(c_{0} X_{0}, w X_{0}\right) \mathrm{e}^{-\lambda_{0} t} \\
& =2 w(p) X_{0}(p) \mathrm{e}^{-t} \\
& =H_{0} \mathrm{e}^{-t}
\end{aligned}
$$

Of course, this goes to 0 for $t \rightarrow \infty$, as it should.

## Conclusion

We have constructed a unique global solution of the FokkerPlanck equation associated with the Wright-Fisher model. This solution leads to explicit formulae for the absorption time, fixation probabilities, the probability of coexistence, $n$th moments, heterogeneity, and other quantities.

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

## References

Abramowitz M, Stegun I (1965) Handbook of mathematical functions with formulas, graphs, and mathematical tables. Dover, New York
Amari S, Nagaoka H (2000) Methods of information geometry. In: Translations of mathematical monographs, vol 191. American Mathematical Society, Providence
Ay N, Jost J (2012) Information geometry (in preparation)
Bürger R (2000) The mathematical theory of selection, recombination, and mutation. Wiley, New York
Ewens WJ (2004) Mathematical population genetics I. Theoretical introduction. In: Interdisciplinary applied mathematics, 2nd edn. Springer, New York
Fisher RA (1922) On the dominance ratio. Proc. R. Soc. Edinb 42:321-341
Jost J (2012) Mathematical methods in biology and neurobiology (in preparation)
Kimura $M$ (1955) Solution of a process of random genetic drift with a continuous model. PNAS-USA 41(3):144-150
Kimura M (1955) Random genetic drift in multi-allele locus. Evolution 9:419-435
Kimura M (1956) Random genetic drift in a tri-allelic locus; exact solution with a continuous model. Biometrics 12:57-66
Rice S (2004) Evolutionary theory. Sinauer, Sunderland
Suetin PK (2001) Ultraspherical polynomials. In Hazewinkel M (ed) Encyclopaedia of mathematics. Springer, Berlin
Tran TD, Hofrichter J, Jost J (2012) A general solution of the WrightFisher model of random genetic drift. arxiv.org/abs/1207.6623
Wright S (1931) Evolution in Mendelian populations. Genetics 16:97-159
Wright $S$ (1945) The differential equation of the distribution of gene frequencies. Proc. Natl. Acad. Sci. USA 31:382-389


[^0]:    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. The authors T. D. Tran and J. Hofrichter have been supported by the IMPRS "Mathematics in the Sciences".

[^1]:    T. D. Tran • J. Hofrichter • J. Jost ( $\triangle$ )

    Max Planck Institute for Mathematics in the Sciences, 04103 Leipzig, Germany
    e-mail: jost@mis.mpg.de
    T. D. Tran
    e-mail: trandat@mis.mpg.de
    J. Hofrichter
    e-mail: julian.hofrichter@mis.mpg.de
    J. Jost

    Department of Mathematics, Leipzig University, 04081 Leipzig, Germany
    J. Jost

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

[^2]:    ${ }^{1}$ Here is a remark for readers not familiar with this mathematical construction: This is a formal definition, as $\delta_{p}$ defined in this manner is not a function itself, but rather operates on continuous functions by assigning to them their value at the particular point $p$. Thus, while the product $(f, g)$ had been first defined for square integrable functions $f, g$, we now apply it to the pair $\left(\delta_{p}, \phi\right)$ where $\delta_{p}$ is a more general object and in turn $\phi$ is a more restricted function (continuous instead of simply square integrable).

[^3]:    ${ }^{2}$ The Gegenbauer polynomials generalise other important classes of polynomials, like the Legendre and the Chebyshev polynomials, and they constitute in turn special cases of the Jacobi polynomials.

[^4]:    $a\left[n_{-}, i_{-}, j_{-}\right]:=\operatorname{If}[(i<=0 \& \& j<=0)| |(i>=n \& \& j>=n), 1$,
    Binomial[n, $\left.j](i / n)^{\wedge} j(1-i / n)^{\wedge}(n-j)\right]$
    Matran[n_] $:=\operatorname{Table}[a[n-1, j-1, i-1],\{i, n\},\{j, n\}]$ $\operatorname{MP}\left[n_{-}, k_{-}\right]:=$MatrixPower [Matran [n], k]
    $\mathrm{XO}\left[\mathrm{n}_{-}, \mathrm{p}_{-}\right]:=$Table[If[i<p||i>p,0,1],\{i,n\},\{j,1\}]
    $X\left[n_{-}, k_{-}, p_{-}\right]:=M P[n, k] . X O[n, p]$
    $\mathrm{G}\left[\mathrm{n}_{-}, \mathrm{k}_{-}, \mathrm{p}_{-}\right]:=\operatorname{BarChart}[\operatorname{Table}[\mathrm{X}[\mathrm{n}, \mathrm{k}, \mathrm{p}][[\mathrm{i}]],\{i, 1, \mathrm{n}\}]]$
    Table[G[20, k, 5], $\{k, 0,18\}], \mathrm{G}[20,30,5]$

