Manuscript version: Working paper (or pre-print)
The version presented here is a Working Paper (or ‘pre-print’) that may be later published elsewhere.

Persistent WRAP URL:
http://wrap.warwick.ac.uk/178891

How to cite:
Please refer to the repository item page, detailed above, for the most recent bibliographic citation information. If a published version is known of, the repository item page linked to above, will contain details on accessing it.

Copyright and reuse:
The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions.

Copyright © and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable the material made available in WRAP has been checked for eligibility before being made available.

Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

Publisher’s statement:
Please refer to the repository item page, publisher’s statement section, for further information.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk.
EXCURSION THEORY FOR THE WRIGHT–FISHER DIFFUSION

PAUL A. JENKINS, JERE KOSKELA, JAROMIR SANT, DARIO SPANÒ, AND IVANA VALENTIĆ

ABSTRACT. In this work, we develop excursion theory for the Wright–Fisher diffusion with recurrent mutation. Our construction is intermediate between the classical excursion theory where all excursions begin and end at a single point and the more general approach where excursions of processes from general sets are considered. Since the Wright–Fisher diffusion has two boundaries, it is natural to construct excursions which start from a specified boundary point, and end at one of two boundary points which determine the next starting point. In order to do this we study the killed Wright–Fisher diffusion, which is sent to a cemetery state whenever it hits either boundary. We then construct a marked Poisson process of such killed paths which, when concatenated, produce a pathwise construction of the Wright–Fisher diffusion.

1. INTRODUCTION

The Wright–Fisher diffusion is one of the most prominent forwards-in-time models used within mathematical population genetics to describe the way in which population-level allele frequencies change over time. Various genetic phenomena such as mutation and selection can be straightforwardly incorporated, offering a realistic and rich probabilistic model to explain the genetic variation present around us. This diffusion has been the subject of a number of studies over past decades, leading to a deeper understanding of said process, with many probabilistic aspects now well understood. Despite its prominent status, both in population genetics and as a canonical diffusion on the unit interval, the Wright–Fisher diffusion falls outside the usual scope of general diffusion theory in several ways: its diffusion coefficient vanishes, and fails to be Lipschitz-continuous, at the boundary of its domain.

While a whole theory of Wright–Fisher diffusions (as well as other singular diffusions) has been developed to circumvent these issues (see e.g. [19]), there remain questions to which existing tools based on stochastic differential equations (SDEs) and martingale problems are ill-suited. For example, the fact that the diffusion coefficient vanishes at the boundary implies that the laws of any two Wright–Fisher diffusions whose drift coefficients differ at a boundary point are singular, which complicates their use in path-based inference and simulation applications [7, 9, 16, 17]. All descriptions of the Wright–Fisher diffusion process currently available are either through the diffusion’s generator, or through the lens of an SDE. We thus aim to provide a third way of viewing the process through the machinery of excursion theory.

In this article we present an excursion-based construction of the neutral Wright–Fisher diffusion, whose drift coefficient is an affine function of the state. Excursion theory has been a prominent tool in the study of processes with a single boundary, such as Bessel processes and reflecting Brownian motions [8, 11, 14, 3] and [15, Section VI.8]. It is a classical area of probability whereby diffusion processes are not viewed through the habitual sample path lens as continuous trajectories through time, but rather as a collection of excursions away from a point indexed by the local time at said point. This rich theory has allowed for many properties of the Bessel process to be fully fleshed out, as well as provided an alternative framework within which the Poisson nature of the underlying excursion process allows for easier and more straightforward computations. Diffusions with a single boundary point are natural subjects for excursion

Keywords and phrases. Wright–Fisher diffusion, excursion theory, hypergeometric function, entrance law, local time, Jacobi polynomials.
theory because all excursions begin and end at a single point, facilitating a Markovian construction of a path from concatenated excursions. Excursions of processes from more general sets have also been described, but the results are rarely tractable because it is in general not clear how to concatenate excursions into a path [14, Section 8]. Our construction is intermediate between these two settings, featuring excursions which always start from a specified point, and end at one of two points which determine the next starting point, giving rise to a natural ‘Markov switching process’ alternating between sets of excursions which initiate from one or other of the two boundary points. We are thus able to provide a construction based on a labelled Poisson process of excursions, with the labels determining how excursions connect to form a continuous path.

In order to achieve this, we start by considering a simple one-dimensional Wright–Fisher diffusion with mutation for which a number of results are well known but scattered across the literature. We bring these results together and present them in Section 2 alongside some interesting relations with the hypergeometric \( \text{}_2F_1 \) functions, for which all the necessary known properties are given in Appendix A. Closely related to the Wright–Fisher diffusion and central to the development of an excursion theory for this diffusion is the corresponding killed process, i.e. the Wright–Fisher diffusion which is sent to a cemetery state whenever the process hits either of the boundary points 0 or 1, which we explore in Section 3. In this case, the diffusion behaves as in the usual Wright–Fisher case when inside the boundary of the diffusion, however killing the process at the boundary leads to a distinctly different behaviour in various probabilistic quantities related to the diffusion, such as the resolvent, Laplace transform of the hitting time, and the process’ transition density, all of which we derive for the first time. Section 4 then puts together the results obtained in the previous two sections to develop the full excursion theory for the Wright–Fisher diffusion where sample paths are now viewed as excursions from and terminating at either of the boundary points.

2. Elementary results for Wright–Fisher diffusions

Consider a Wright–Fisher diffusion with mutation parameters \( \theta_1 \) and \( \theta_2 \), that is a diffusion process on \([0, 1]\) with generator whose domain is \( C^2([0, 1]) \) and is given by

\[
\mathcal{G} = \frac{1}{2} x(1 - x) \frac{d^2}{dx^2} + \frac{1}{2} \left( \theta_1 - |\theta| x \right) \frac{d}{dx}.
\]

where we have introduced the notation \( |\theta| = \theta_1 + \theta_2 \). From Feller’s boundary classification (see for example [10, Example 15.6.8 page 240]) we have that both boundaries are regular if the following assumption is satisfied

\[
0 < \theta_1, \theta_2 < 1.
\]

We will only be interested in this case since regularity of the boundary means that process started from the boundary point can enter the interior \((0, 1)\) and again reach both boundaries which will be important in order to build the excursions started from the boundary points. However, some results in this section hold for any \( \theta_1, \theta_2 > 0 \) and we will therefore emphasise when we do not use the additional assumption.

Remark 1. One could also consider excursions from boundary 0 in case \( 0 < \theta_1 < 1 \) and \( \theta_2 \geq 1 \) or from boundary 1 in case \( 0 < \theta_2 < 1 \) and \( \theta_1 \geq 1 \). In those settings, however, the process could never reach the opposite boundary and therefore we would be back to the standard setting of excursions from a single point.

The generator can be rewritten as

\[
\mathcal{G} = \frac{1}{m(x)} \frac{d}{dx} \frac{1}{s'(x)} \frac{d}{dx}
\]
The Green’s function is defined as

\begin{equation}
(7)
m(x) = \frac{\Gamma(|\theta|)}{\Gamma(\theta_1)\Gamma(\theta_2)}x^{\theta_1-1}(1 - x)^{\theta_2-1} = \frac{x^{\theta_1-1}(1 - x)^{\theta_2-1}}{B(\theta_1, \theta_2)}
\end{equation}

is the density on \([0, 1]\) of the speed measure of \(X\) with respect to Lebesgue measure, and

\begin{equation}
(5)
s'(x) = \frac{\Gamma(\theta_1+\theta_2)}{\Gamma(\theta)}x^{-\theta_1}(1 - x)^{-\theta_2} = 2B(\theta_1, \theta_2)x^{-\theta_1}(1 - x)^{-\theta_2}
\end{equation}

is the derivative of the scale function \(s(x)\). We note here that the constants in these definitions can be altered by \(m(dx) \mapsto cm(dx)\), \(s(dx) \mapsto e^{-1}s(dx)\), and (3) still holds. Some authors use other constants such as \(m(x) = 2x^{\theta_1-1}(1 - x)^{\theta_2-1}\). In our normalization \(m\) is also the stationary distribution for \(X\).

According to formula (7.14) in [5] the transition density with respect to the Lebesgue measure for the Wright–Fisher diffusion can, regardless of the assumption \(\theta_1, \theta_2 < 1\), be expressed as

\begin{equation}
(6)
p(t, x, y) = \frac{\Gamma(\theta_1+\theta_2)}{\Gamma(\theta)}x^{\theta_1-1}(1 - y)^{\theta_2-1}\sum_{n=0}^{\infty} \frac{1}{\pi_n} e^{-s_{2}^{(\theta_1, \theta_2)}(x)}R_{n}^{(\theta_1, \theta_2)}(x)R_{n}^{(\theta_1, \theta_2)}(y),
\end{equation}

where

\[R_{n}^{(\theta_1, \theta_2)}(x) = \binom{2F_{1}}{-n, n + |\theta|-1; 1-\theta_1; 1-x}.\]

With \((a)_{n} := a(a + 1)\ldots(a + n - 1)\) defined as the rising factorial, the hypergeometric function \(2F_{1}\) is defined by the following power series

\[2F_{1}(a, b; c; x) = \sum_{n=0}^{\infty} \frac{(a)_{n}(b)_{n}}{(c)_{n}n!}x^{n}\]

under suitable conditions for \(a, b\) and \(c\) (see Appendix A). Since \((-n)_{k} = 0\) for \(k \geq n + 1\), the function \(R_{n}^{(\theta_1, \theta_2)}\) is a polynomial. In fact it is obtained from a Jacobi polynomial by mapping their usual domain \([-1, 1]\) to \([0, 1]\) through \(x \mapsto (x + 1)/2\). The constant \(\pi_n\) is obtained as

\[\pi_n = \int_{0}^{1} x^{\theta_1-1}(1 - x)^{\theta_2-1} \left(R_{n}^{(\theta_1, \theta_2)}(x)\right)^{2} dx = \frac{n!(\theta_1)_{(n)}}{(1\theta_1 + 2n - 1)(\theta_1)_{(n-1)}(\theta_2)_{(n)}}.\]

We denote by \(p_{m}(t, x, y)\) the transition density of the process with respect to the speed measure \(m\), which is such that

\[P_{t}(X, A) = \mathbb{P}_x\left(X_t \in A\right) = \int_A p_{m}(t, x, y)m(dy).\]

Therefore the transition density (with respect to Lebesgue measure) can be expressed as

\[p(t, x, y) = p_{m}(t, x, y)m(y) = \frac{y^{\theta_1-1}(1 - y)^{\theta_2-1}}{B(\theta_1, \theta_2)}p_{m}(t, x, y).\]

The resolvent, defined by

\[R_{\lambda}f(x) := \mathbb{E}_x \left[\int_{0}^{\infty} e^{-\lambda t} f(X_{t}) dt\right] = \int_{0}^{\infty} e^{-\lambda t} P_{t}f(x) dt\]

is therefore given by

\begin{equation}
(7)
R_{\lambda}f(x) = \int_{0}^{\infty} \int_{0}^{1} e^{-\lambda t} p(t, x, y)f(y) dy dt.
\end{equation}

The Green’s function is defined as

\[G_{\lambda}(x, y) := \int_{0}^{\infty} e^{-\lambda t} p_{m}(t, x, y) dt,\]
and it is important to note that we define it with respect to the speed measure \( m \). So we have the following formula which will prove most useful:

\[
\int_A G_A(x, y)m(y)dy = R_A 1_A(x).
\]

From the expansion (6) for the transition density we, regardless of the assumption \( \theta_1, \theta_2 < 1 \), get the following:

\[
G_A(x, y) = \int_0^\infty e^{-\lambda t} \sum_{n=0}^\infty \frac{1}{\pi_n} e^{-\left(k(n+\theta - 1) - \frac{m(n+\theta - 1)}{2}\right)^2 R_n^{\theta_1, \theta_2}(x)R_n^{\theta_1, \theta_2}(y)}dt
\]

\[
= \sum_{n=0}^\infty \frac{2}{2\lambda + n(n + \theta - 1)} \frac{1}{\pi_n} R_n^{\theta_1, \theta_2}(x)R_n^{\theta_1, \theta_2}(y).
\]

In what follows we will obtain an alternative formula for the Green’s function which will in turn, by formula (8), also result in an alternative formula for the resolvent.

First we consider functions \( \Phi_{\lambda, \pm} \) which are respectively the decreasing and increasing solutions to the generator eigenfunction equation

\[
\mathcal{L} \Phi = \lambda \Phi
\]

subject to suitable boundary conditions (which reflect the boundary behaviour of the diffusion). For more details see [4, Section I, Number 10, pg 19]. For a general diffusion with speed measure \( m \), scale function \( s \), and killing measure \( k \), the boundary conditions at zero are given by

\[
\lambda \Phi_{\lambda, -}(0)m(0) = (\Phi_{\lambda, -})^{+} (0) - \Phi_{\lambda, -}(0)k(0)
\]

where \((\Phi_{\lambda, -})^{+} (x)\) denotes the right-hand-side derivative with respect to the scale function, formally

\[
(\Phi_{\lambda, -})^{+} (x) = \lim_{h \searrow 0} \frac{\Phi_{\lambda, -}(x + h) - \Phi_{\lambda, -}(x)}{s(x + h) - s(x)} = \frac{(\Phi_{\lambda, -})^{+} (x)}{s^{+}(x)} = \Phi_{\lambda, -}^{+} (x) 2B(\theta_1, \theta_2)
\]

where \( f^+ \) is a derivative of function \( f \) from the right. So in our case with \( k(0) = 0 = m(0) \), the boundary condition for (10) becomes simply \((\Phi_{\lambda, -})^{+} (0) = 0 \). For \( \Phi_{\lambda, +}(1) \) we have the following similar condition, with \((\Phi_{\lambda, +})^{-} (1)\) defined analogously to \((\Phi_{\lambda, -})^{+} \) as a derivative from the left:

\[
\lambda \Phi_{\lambda, +}(1)m(1) = - (\Phi_{\lambda, +})^{-} (1) - \Phi_{\lambda, +}(1)k(1),
\]

which again becomes simply \((\Phi_{\lambda, +})^{-} (1) = 0 \). Notice that the condition for \( \Phi_{\lambda, -} \) only depends on the behaviour around 0 and the condition for \( \Phi_{\lambda, +} \) only depends on the behaviour around 1. When constructing the excursion theory the killing term will play a crucial role and in that case a good reference for the boundary conditions is [4, pgs. 15–17].

Solving the eigenfunction equation (10) for the Wright-Fisher generator (3) is equivalent to solving the hypergeometric differential equation for a particular parameter configuration. For details see Remark 24. For all \( \lambda \neq \frac{(\theta| - 1)^2}{8} \) the hypergeometric differential equation admits the hypergeometric function as a solution, and thus provides us with the required function \( \Phi_{\lambda} \).

By matching the coefficients in (10) with the generators given by (1) and (A5), we find that

\[
\Phi_{\lambda, +}(x) = 2F_1 \left( \frac{|\theta| - 1 + \sqrt{(|\theta| - 1)^2 - 8\lambda}}{2}, \frac{|\theta| - 1 - \sqrt{(|\theta| - 1)^2 - 8\lambda}}{2}; \theta_2; 1 - x \right)
\]

\[
\Phi_{\lambda, -}(x) = 2F_1 \left( \frac{|\theta| - 1 + \sqrt{(|\theta| - 1)^2 - 8\lambda}}{2}, \frac{|\theta| - 1 - \sqrt{(|\theta| - 1)^2 - 8\lambda}}{2}; \theta_1; x \right),
\]
are respectively the decreasing and increasing solutions to the equation (10).

**Remark 2.** We emphasise that the eigenfunction solutions (14)–(15) hold for almost all \( \lambda \). Contrast this property with the set of eigenfunction solutions to (10) when there is a boundary condition for \( \Phi \) at both zero and one. In that situation, by Sturm–Liouville theory the spectrum of eigenvalues for which a solution exists is countably infinite, namely \( \lambda_n = -n(n + |\theta| - 1)/2, \ n = 0, 1, \ldots \). It is this countable collection of eigen-value/-function pairs which is used to construct the spectral expansion for \( p(t, x, y) \) appearing in (6). Also notice that, in obtaining (14)–(15), the assumption (2) becomes important because if it is not satisfied then functions \( \Phi_{\pm} \) may not be well defined. For details on this point see Remark 21 in Appendix A.

**Remark 3.** For an explanation on why the condition \( \lambda \neq \frac{|\theta| - 1}{\lambda} \) is necessary see Remark 24 in Appendix A. This condition will not, however prove binding because we will mostly be interested in integrals over \( \lambda \in [0, \infty) \) and therefore one point will not cause problems.

We denote by

\[
\begin{align*}
  a(\lambda) &= \frac{|\theta| - 1 + \sqrt{(|\theta| - 1)^2 - 8\lambda}}{2} \\
  b(\lambda) &= \frac{|\theta| - 1 - \sqrt{(|\theta| - 1)^2 - 8\lambda}}{2}.
\end{align*}
\]

Since due to (A4) and (A3) we have

\[
\begin{align*}
  \frac{2F_1(a(\lambda), b(\lambda); \theta_1; \cdot)\uparrow(0)}{s^+(0)} = \lim_{y\searrow 0} \frac{\lambda}{\theta_1} 2F_1(a(\lambda) + 1, b(\lambda) + 1; \theta_1 + 1; y) \frac{y^{\theta_1}(1-y)^{\theta_2}}{2B(\theta_1, \theta_2)} = 0
\end{align*}
\]

(17) and

\[
\begin{align*}
  \frac{2F_1(a(\lambda), b(\lambda); \theta_2; 1 - \cdot)\downarrow(1)}{s^-(1)} = \lim_{y\nearrow 1} \frac{-\lambda}{\theta_2} 2F_1(a(\lambda) + 1, b(\lambda) + 1; \theta_2 + 1; 1 - y) \frac{y^{\theta_1}(1-y)^{\theta_2}}{2B(\theta_1, \theta_2)} = 0,
\end{align*}
\]

we see that \( \Phi_{\pm} \) also satisfy their respective boundary conditions.

---

**Figure 1.** The graphs for \( \Phi_{\pm} \) for \( \lambda = 0.1 \) and \( \theta_1 = \theta_2 = 0.3 \) in blue, \( \theta_1 = 0.3, \theta_2 = 0.7 \) in yellow and \( \theta_1 = 0.7, \theta_2 = 0.3 \) in green.

Due to (A3) we have that

\[
\Phi_{\pm}(0) = 1 = \Phi_{\pm}(1),
\]
from (A2) we have
\[ \Phi_{\lambda,-}(1) = \frac{\Gamma(\theta_1)\Gamma(1 - \theta_2)}{\Gamma\left(\frac{\theta_1 - \theta_2 + 1 - \sqrt{(|\theta_1| - 1)^2 - 8\lambda}}{2}\right)\Gamma\left(\frac{\theta_1 - \theta_2 + 1 + \sqrt{(|\theta_1| - 1)^2 - 8\lambda}}{2}\right)}. \]

and also from (A2) we have
\[ \Phi_{\lambda,+}(0) = \frac{\Gamma(\theta_2)\Gamma(1 - \theta_1)}{\Gamma\left(\frac{\theta_1 - \theta_2 + 1 - \sqrt{(|\theta_1| - 1)^2 - 8\lambda}}{2}\right)\Gamma\left(\frac{\theta_1 - \theta_2 + 1 + \sqrt{(|\theta_1| - 1)^2 - 8\lambda}}{2}\right)}. \]

The Wronskian is defined as
\[ W_{\lambda} = \frac{\Phi'_{\lambda,-}(y)\Phi_{\lambda,+}(y) - \Phi_{\lambda,-}(y)\Phi'_{\lambda,+}(y)}{s'(y)}, \]

which does not depend on \( y \). We will denote by \( W(f, g) = f(x)g'(x) - f'(x)g(x) \) to capture the numerator of the Wronskian in (18) and according to (A11) we can conclude that
\[ W_{\lambda} = \frac{\Gamma(|\theta_1|)}{2\Gamma(\theta_1)\Gamma(\theta_2)} y^{\theta_1 - 1}(1 - y)^{\theta_2 - 1} \]
\[ = \frac{\Gamma(|\theta_1|)\Gamma(1 - \theta_1, \theta_2) (1 - \theta_1)y^{\theta_1 - 1}(1 - y)^{\theta_2 - 1}}{2\Gamma(\theta_1)\Gamma(\theta_2)\Gamma(a(\lambda))\Gamma(b(\lambda))} y^{\theta_1 - 1}(1 - y)^{\theta_2 - 1}. \]

According to equation (37) in [12] we have an alternative to evaluating the Green’s function via the following:
\[ G_{\lambda}(x, y) = \frac{\Phi_{\lambda,-}(x \wedge y)\Phi_{\lambda,+}(x \vee y)}{W_{\lambda}} \]
\[ = \frac{2\Gamma(a(\lambda))\Gamma(1 - a(\lambda))}{\Gamma(|\theta_1|)} \left( \frac{2}{\Gamma(\theta_1)} \right) F_1 \left( a(\lambda) + b(\lambda); \theta_1; x \wedge y \right) \]
\[ \times 2 F_1 \left( a(\lambda), b(\lambda); \theta_2; 1 - (x \vee y) \right). \]

From (8) we therefore obtain an alternative way of evaluating the resolvent.

**Proposition 4.** For \( A \in B([0, 1]) \) and \( x \in [0, 1] \) the following holds:
\[ R_{\lambda}1_A(x) = \int_A 2 \frac{\Gamma(\theta_1)\Gamma(\theta_2)}{\Gamma(\theta_1)\Gamma(\theta_2)} y^{\theta_1 - 1}(1 - y)^{\theta_2 - 1} \]
\[ \times 2 F_1 \left( a(\lambda) + b(\lambda); \theta_1; x \wedge y \right) 2 F_1 \left( a(\lambda), b(\lambda); \theta_2; 1 - (x \vee y) \right) dy. \]

**Remark 5.** If \( Y \) is a beta(\( \theta_1, \theta_2 \)) random variable then
\[ R_{\lambda}1_A(x) \]
\[ = \frac{2\Gamma(a(\lambda))\Gamma(b(\lambda))}{\Gamma(|\theta_1|)} \mathbb{E} \left[ 2 F_1 \left( a(\lambda) + b(\lambda); \theta_1; x \wedge Y \right) 2 F_1 \left( a(\lambda), b(\lambda); \theta_2; 1 - (x \vee Y) \right) 1_A(Y) \right]. \]

One more thing we will be interested in is the first hitting time \( H_y := \inf \{ t > 0 : X_t = y \} \). According to equation (9) in [12] for \( \lambda \geq 0 \)
\[ \mathbb{E}_x \left[ e^{-\lambda H_y} \right] = \begin{cases} \frac{\phi_{\lambda}(x)}{\phi_{\lambda}(y)} & x < y, \\ \frac{\phi_{\lambda}(x)}{\phi_{\lambda}(y)} & x > y, \end{cases} \]
from which it follows [using (14)–(16)] that

$$(23) \quad \mathbb{E}_x [e^{-s H_x}] = \begin{cases} 
\frac{1}{2} \mathbb{F}_1(a, b; \theta_1; x) & x < y, \\
\frac{1}{2} \mathbb{F}_1(a, b; \theta_1; y) & x > y.
\end{cases}$$

3. KILLED WRIGHT–FISHER DIFFUSION

The next question we are interested in is what happens when we introduce killing at one of the boundaries. For this reason we define $X^y$ as a Wright–Fisher diffusion with generator (1) started at $y \in [0, 1]$. For $b \in (0, 1)$ we will denote by $(X^y_{t\wedge H_0})_{t \geq 0} = (X^y_{t\wedge H_b})_{t \geq 0}$ process $X^y$ stopped at time $H_b$ of the first hitting of point $b$. We will do the calculations for killing at the 0 endpoint but also give alternative results for killing at 1. When it comes to the transition density the following formula holds

$$p^0(t, x, y) = p(t, x, y) - \mathbb{E}_x [p(t - H_0, X_{H_0}, y)1_{H_0 < t}].$$

We are interested in how to calculate it in case of a Wright–Fisher diffusion killed at the boundary. We have

$$p^0(t, x, y) = p(t, x, y) - \mathbb{E}_x [p(t - H_0, X_{H_0}, y)]$$
$$= p(t, x, y) - \mathbb{E}_x [p(t - H_0, 0, y)1_{H_0 < t}]$$
$$= \frac{y^{\theta_1 - 1}(1 - y)^{\theta_2 - 1}}{B(\theta_1, \theta_2)} \sum_{n=0}^{\infty} \frac{1}{\pi_n} e^{-n(\theta_1 + \theta_2)H_0} R_n^{(\theta_1, \theta_2)}(y) H_0 < t]$$
$$= \frac{y^{\theta_1 - 1}(1 - y)^{\theta_2 - 1}}{B(\theta_1, \theta_2)} \sum_{n=0}^{\infty} \frac{1}{\pi_n} e^{-n(\theta_1 + \theta_2)H_0} R_n^{(\theta_1, \theta_2)}(y) e^{-n(\theta_1 + \theta_2)H_0} 1_{H_0 < t}$$

where we used (6) and (A2).

An alternative approach to calculating the transition density of the killed process $X^x_{t\wedge 0}$ is to follow the general approach from Section 15.13 in [10] used to determine (6) but change the boundary conditions in order to reflect the killing behaviour at 0.

Denote by $u(t, x) = \mathbb{E}_x [f(X_t)]$ and by $u^{(k)}(t, x) = \mathbb{E}_x [f(X^x_{t\wedge 0})]$. If we consider the spectral decomposition of the Wright–Fisher diffusion, this follows from solving the Kolmogorov backward equation

$$(24) \quad \frac{\partial u}{\partial t}(t, x) = \mathcal{G} u(t, x),$$

where $\mathcal{G}$ is the generator of the Wright–Fisher diffusion with mutation. From comments in Remark 2 we see that there exists a countable set of eigenvalues of form $\lambda_n = -n(n + |\theta| - 1)/2$, $n = 0, 1, \ldots$ and it follows that

$$u(t, x) = \sum_{n=1}^{\infty} c_n(f) e^{-\frac{n(\theta_1 + \theta_2)}{2}H_0} R_n^{(\theta_1, \theta_2)}(x),$$
where \( R_n^{(\theta_1, \theta_2)} \) are Jacobi polynomials linearly transformed to be defined on \([0, 1]\). Constants \( c_n(f) \) need to be chosen such that the initial condition \( f(x) = \sum_{n=1}^{\infty} c_n(f) R_n^{(\theta_1, \theta_2)}(x) \) is satisfied. Since linearly transformed Jacobi polynomials are orthogonal with respect to the scalar product \( \langle f, g \rangle = \int_0^1 f(x)g(x)m(x)dx \) it follows that

\[
c_n(f) = \frac{\langle f, R_n^{(\theta_1, \theta_2)} \rangle}{\langle R_n^{(\theta_1, \theta_2)}, R_n^{(\theta_1, \theta_2)} \rangle},
\]

which implies (6).

In the case of the Wright–Fisher diffusion with killing at the boundary 0 function \( u^0 \) still needs to satisfy (24) with the same initial condition but we also have the following boundary condition

\[
u(t, 0) = 0
\]

and therefore instead of Jacobi polynomials we have to consider the Jacobi functions of the second kind (see Section 4.23 in [18]). We have

\[
u^0(t, x) = \sum_{n=1}^{\infty} c_n^0(f) e^{-\frac{\theta_1(x+y)}{2}} Q_n^{\theta_1, \theta_2}(x),
\]

where \( Q_n^{\theta_1, \theta_2} \) is not a polynomial but instead

\[
Q_n^{\theta_1, \theta_2}(x) = x^{1-\theta_1} F_1 \left( 1 - n - \theta_1, n + \theta_2; 2 - \theta_1; x \right).
\]

Since Jacobi functions of the second kind are not orthogonal with respect to the same scalar product it is not so straightforward to get the constants \( c_n^0(f) \) such that the initial condition is satisfied. Hence, instead of the transition density of the killed diffusion, we compute its Laplace transform, for which this problem is easier to circumvent.

**Lemma 6.** For \( A \in B([0, 1]) \) and \( x \in [0, 1] \) the following holds:

\[
R^0_{\lambda} 1_A(x) = 2 \frac{\Gamma(a) \Gamma(b)}{\Gamma(\theta_1) \Gamma(\theta_2)} \int_A y^{\theta_1-1}(1 - y)^{\theta_2-1}
\]

\[
\times \left[ 2 F_1 \left( a, b; \theta_1; x \wedge y \right) 2 F_1 \left( a, b; \theta_2; 1 - (x \vee y) \right)
\right.
\]

\[
- \frac{\Gamma(\theta_2 - a) \Gamma(\theta_2 - b)}{\Gamma(\theta_2) \Gamma(1 - \theta_1)} 2 F_1 \left( a, b; \theta_2; 1 - x \right) 2 F_1 \left( a, b; \theta_2; 1 - y \right) dx dy,
\]

\[
R^1_{\lambda} 1_A(x) = 2 \frac{\Gamma(a) \Gamma(b)}{\Gamma(\theta_1) \Gamma(\theta_2)} \int_A y^{\theta_1-1}(1 - y)^{\theta_2-1}
\]

\[
\times \left[ 2 F_1 \left( a, b; \theta_1; x \wedge y \right) 2 F_1 \left( a, b; \theta_2; 1 - (x \vee y) \right)
\right.
\]

\[
- \frac{\Gamma(\theta_1 - a) \Gamma(\theta_1 - b)}{\Gamma(\theta_1) \Gamma(1 - \theta_2)} 2 F_1 \left( a, b; \theta_1; x \right) 2 F_1 \left( a, b; \theta_1; y \right) dx dy.
\]

**Proof.** From (1) in [13] we have that

\[
R^0_{\lambda} 1_A(x) = R_{\lambda} 1_A(x) - \mathbb{E}_x \left[ e^{-\lambda H_0} \right] R_{\lambda} 1_A(0).
\]

We proceed by using equation (21) for the resolvent of the process which isn’t killed, (23) for the calculation of \( \mathbb{E}_x \left[ e^{-\lambda H_0} \right] \), where we use the second case since \( x > 0 \) and in the end we use equations (A2) and (A3) for the calculation of the constants. This ends the proof of the first statement and the second statement is proven analogously.

If one was interested in calculating the resolvent of the process killed at time \( H_{0,1} = H_0 \wedge H_1 \), that is a process killed when it reaches either boundary, then you need to apply the same formula
twice,  

\begin{equation}  
R_{x}^{0}1_{A}(x) = R_{x}^{0}1_{A}(x) - \mathbb{E}_{x} \left[ e^{-\lambda H_{0}^{0}} \right] R_{x}^{0}1_{A}(1),  
\end{equation}  

where $H_{0}^{0}$ is the first hitting time of 1 for the Wright–Fisher diffusion killed at 0. To calculate $\mathbb{E}_{x} \left[ e^{-\lambda H_{0}^{0}} \right]$ we can again use (22) but we first need to determine what $\Phi_{1,-}$ and $\Phi_{1,+}$ are for the killed process. Let’s denote by $\Phi^{b}_{\lambda, \pm}$ the decreasing and increasing solutions to the generator eigenfunction equation

$\mathcal{L}\Phi = \lambda \Phi$, 

where now the boundary conditions reflect the fact that the process is killed at time $H_{0}$, that is when it reaches $0$. We will denote the solutions to the generator eigenfunction equation satisfying the boundary conditions with killing at the boundary $b \in \{0, 1\}$ by $\Phi^{b}_{\lambda, \pm}$.

**Proposition 7.** With $a(\lambda)$ and $b(\lambda)$ defined in (16) the following holds

\begin{align*}  
\Phi^{0}_{\lambda,+}(x) &= 2F_{1} \left( a(\lambda), b(\lambda); \theta_{2}; 1 - x \right),  
\Phi^{0}_{\lambda,-}(x) &= x^{-\theta_{1}} 2F_{1} \left( \theta_{2} - b(\lambda), \theta_{2} - a(\lambda); 2 - \theta_{1}; x \right),  
\Phi^{1}_{\lambda,+}(x) &= (1 - x)^{-\theta_{1}} 2F_{1} \left( \theta_{1} - a(\lambda), \theta_{1} - b(\lambda); 2 - \theta_{2}; 1 - x \right),  
\Phi^{1}_{\lambda,-}(x) &= 2F_{1} \left( a(\lambda), b(\lambda); \theta_{1}; x \right).  
\end{align*}  

**Proof.** Since the condition for $\Phi_{\lambda, \pm}$ only depends of the behaviour around $b = 1$ killing the process at boundary $b = 0$ will not affect it. Therefore the boundary condition remains the same as when we had no killing, \( \left( \Phi_{\lambda, \pm}^{0} \right)^{-1} (1) = 0 \), and we have the formula in (26) the same as in (14).

On the other hand since we introduce killing at the boundary $b = 0$ we have $\Phi^{0}_{\lambda,-}(0) = 0$ (see [4, pgs. 15-17]). Since

\( (0)^{1-\theta_{1}} 2F_{1} \left( \theta_{2} - b(\lambda), \theta_{2} - a(\lambda); 2 - \theta_{1}; 0 \right) = 0 \)

holds again by matching the coefficients in (1) and (A5), we observe that (27) holds.

For the case of a process killed once it reaches the boundary $b = 1$, the boundary conditions are

\begin{equation}  
\left( \Phi_{1,-}^{1} \right)^{+} (0) = 0 \text{ and } \Phi_{1,+}^{1}(1) = 0.  
\end{equation}  

Therefore, clearly as before we have (29). For the second condition in (30) we note that

\( (1 - 1)^{1-\theta_{2}} 2F_{1} \left( \theta_{1} - a(\lambda), \theta_{1} - b(\lambda); 2 - \theta_{2}; 1 - 1 \right) = 0 \),

and thus again by matching the coefficients in (1) and (A5), we observe that (28) holds. \qed

**Remark 8.** Equation (A2) implies that

\begin{align*}  
\Phi^{1}_{\lambda,-}(1) &= 2F_{1} \left( a(\lambda), b(\lambda); \theta_{1}; 1 \right) = \frac{\Gamma(\theta_{1}) \Gamma(1 - \theta_{2})}{\Gamma(\theta_{1} - a(\lambda)) \Gamma(\theta_{1} - b(\lambda))},  
\Phi^{1}_{\lambda,+}(0) &= 2F_{1} \left( \theta_{1} - a(\lambda), \theta_{1} - b(\lambda); 2 - \theta_{2}; 1 \right) = \frac{\Gamma(2 - \theta_{2}) \Gamma(1 - \theta_{1})}{\Gamma(1 - b(\lambda)) \Gamma(1 - a(\lambda))},  
\Phi^{0}_{\lambda,-}(1) &= 2F_{1} \left( \theta_{2} - b(\lambda), \theta_{2} - a(\lambda); 2 - \theta_{1}; 1 \right) = \frac{\Gamma(2 - \theta_{1}) \Gamma(1 - \theta_{2})}{\Gamma(1 - a(\lambda)) \Gamma(1 - b(\lambda))},  
\Phi^{0}_{\lambda,+}(0) &= 2F_{1} \left( a(\lambda), b(\lambda); \theta_{2}; 1 \right) = \frac{\Gamma(\theta_{2}) \Gamma(1 - \theta_{1})}{\Gamma(\theta_{2} - b(\lambda)) \Gamma(\theta_{2} - a(\lambda))}.  
\end{align*}
Theorem 9. For $A \in B([0,1])$, $x \in [0,1]$ and $\lambda \in [0, \infty) \setminus \left\{ \frac{(|\theta|-1)^2}{8} \right\}$ the following holds:

$$R^{0,1}_{\lambda} A(x) = 2 \Gamma(a(\lambda)) \Gamma(b(\lambda)) \left[ \int_A y^{\theta_1-1}(1-y)^{\theta_2-1} \times \left( 2F_1 \left( a(\lambda), b(\lambda); \theta_1; x \wedge y \right) \right) \right]$$

Remark 10. We could have also first killed the process at 1 and then at 0 and the formula has to be the same. In that case we have

$$R^{0,1}_{\lambda} A(x) = R^{0,1}_{\lambda} 1_A(x) = R^1 1_A(x) - \mathbb{E}_x \left[ e^{-H^0_1} \right] R^1 1_A(0),$$

where $H^0_1$ is the first hitting time of 0 for the Wright–Fisher diffusion killed at 1. We use Lemma 6 to get $R^1 1_A$ and (22) to conclude that

$$\mathbb{E}_x \left[ e^{-H^0_1} \right] = \frac{\Phi^0_{\lambda,x}(x)}{\Phi^0_{\lambda,0}(0)} = \frac{\Gamma(1-a(\lambda)) \Gamma(1-b(\lambda))}{\Gamma(2-\theta_1) \Gamma(1-\theta_1)} (1-x)^{1-\theta_2} F_1 \left( \theta_1 - a(\lambda), \theta_1 - b(\lambda); 2 - \theta_1; 1 - x \right),$$

which implies that

$$R^{0,1}_{\lambda} A(x) = 2 \frac{\Gamma(a(\lambda)) \Gamma(b(\lambda))}{\Gamma(\theta_1) \Gamma(\theta_2)} \left[ \int_A y^{\theta_1-1}(1-y)^{\theta_2-1} \times \left( 2F_1 \left( a(\lambda), b(\lambda); \theta_1; x \wedge y \right) \right) \right]$$
The excursions away from 
1−y which will be indicated again, by an excursion of infinite length having
been killed at the 0 boundary. At that point we again start the excursion process from 0 and repeat
the procedure. For all of this to be possible, the necessity of the assumption 0 < θ₁, θ₂ < 1 is
clear.

We will construct the excursions of X from the set {0, 1} of boundary points by taking
y ∈ (0, 1), looking at the excursion from y and then letting y → b, for b ∈ {0, 1}.

Definition 11. For y ∈ [0, 1] we define by Uₙ the set of all paths which stay at y once they first
hit it, that is

\[ Uₙ := \{ f \in C(\mathbb{R}_+, [0, 1]) : f^{-1}([0, 1] \setminus \{ y \}) = (0, \eta) \text{ for some } \eta \in (0, \infty) \}, \]

where \( C(\mathbb{R}_+, [0, 1]) \) is the set of all continuous functions mapping \( \mathbb{R}_+ = [0, \infty) \) into [0, 1], and
\( \eta \) can be viewed as the lifetime of the excursion.

We are going to go from looking at diffusions as being entire continuous paths to a collection
of excursions of two types, in either \( U₀ \) or \( U₁ \), interchanging from one type to the other at
the respective stopping times \( H₁ \) and \( H₀ \). The original path can be recovered from an ordered list
of excursions of both types, along with the stopping times which tell us when we shift from one
type to the other.

The local time of \( X' \) at y is a non-decreasing continuous process where the set over which the
process increases is \( \{ t : X'_t = y \} \) and therefore an excursion away from y started at local time
\( l \) comes before another one started at local time \( l' \) if \( l < l' \). The excursions away from y will be

\[
- \frac{Γ(θ₁ - a(λ))Γ(θ₁ - b(λ))}{Γ(θ₁)Γ(1 - θ₂)} F₂ \left( a(λ), b(λ); θ₁; x \right) F₂ \left( a(λ), b(λ); θ₁; y \right)
- \frac{Γ(1 - a(λ))Γ(1 - b(λ))}{Γ(2 - θ₂)Γ(1 - θ₁)} (1 - x)^{θ₂} F₂ \left( θ₁ - b(λ), θ₁ - a(λ); 2 - θ₂; 1 - x \right)
\]

\[
× \left( F₂ \left( a(λ), b(λ); θ₂; 1 - y \right) - \frac{Γ(θ₁ - a(λ))Γ(θ₁ - b(λ))}{Γ(θ₁)Γ(1 - θ₂)} F₂ \left( a(λ), b(λ); θ₁; y \right) \right) \right] dy.
\]

If one compares the two formulas it is clear that (36) follows from (35) by exchanging 0 to 1
(and therefore y to 1 - y) and θ₁ to θ₂. This is the same as taking a Wright–Fisher diffusion \( X \)
with interchanged mutation coefficients and noticing that the distribution of 1 - \( X \) has the same
distribution as \( X \).

4. EXCURSIONS

The standard excursion theory is mostly interested in excursions which start and end at the
same point [3, 14]. Alternatively, one might look at excursions from a general set, which is also
a well-understood theory but much more complicated than the original approach (see Section 8
in [14], Chapter VII in [3]). We will be looking at excursions of a Wright–Fisher diffusion from
the boundary points. Since there are two boundary points this is part of the general theory for
excursions from a set, but since there are only two points in our set and the process cannot jump
from one point to the other the theory will also resemble the the simpler theory of excursions
from a point. We will for example start from point 0, make a lot of excursions that will again
end at 0, and then end in a formally infinite excursion which actually indicates that the process
has ‘switched’ to the other boundary point. Excursions indicating that a boundary switch has
happened are of infinite length because we introduce a killing term at the opposite boundary in
order to identify when such switches take place. Once that happens we will start a new excursion
process from 1 and again have a lot of excursions which all end at 1 until the process switches
again when it reaches 0 which will be indicated, again, by an excursion of infinite length having
been killed at the 0 boundary. At that point we again start the excursion process from 0 and repeat
the procedure. For all of this to be possible, the necessity of the assumption 0 < θ₁, θ₂ < 1 is
clear.
indexed by Markov local time \((L^y_t)_{t \geq 0}\) at \(y\). We use the Itô–McKean normalization
\[
\int_0^t g(X_u)du = \int g(y)L^y_t\,m(dy)
\]
for arbitrary non-negative Borel functions \(g\). We also define the right-continuous inverse of the Markov local time process, that is \(\tau^y_t := \inf \{ u : L^y_u > t \} \).

**Remark 12.** Note here that the definition of the local time depends on the normalization we used in the definition of the speed measure.

**Remark 13.** One might also consider using the semimartingale local time, that is use the Meyer–Tanaka normalization
\[
\int_0^t g(X_u)X_u(1 - X_u)du = \int g(y)L^y_t\,dy
\]
for arbitrary non-negative Borel functions \(g\). This approach, however, would not prove useful since in that case the local time for the Wright–Fisher diffusion at the boundary will be zero, that is \(L^b \equiv 0\) for \(b \in \{0, 1\}\).

This now allows us to formalise the decomposition of \(X\) into a collection of excursions of two types indexed by local time at \(b \in \{0, 1\}\). We define the excursion processes \(\Xi^{y-b}\) on \(\mathbb{R}_+ \times U_y\), such that \((l, f) \in \Xi^{y-b}\) if at local time \(l\) process \(X^{y-b}\) makes an excursion \(f\) away from \(y\). The main reason why phrasing the path of \(X\) in terms of excursions is particularly insightful is the following result by Itô.

**Theorem (Itô [8]).** For \(y \in [0, 1]\) and \(b \in \{0, 1\}\) the excursion process \(\Xi^{y-b}\) is a Poisson point process with excursion measure \(\text{Leb} \otimes n^{y-b}\), where \(n^{y-b}\) is a \(\sigma\)-finite measure on \(U_y\), called the excursion (or characteristic) measure. Thus

1. For any \(A \subset U_y\) for which \(n^{y-b}(A) < \infty\), we have that
\[
\mathbb{P} \left[ \text{no points of } \Xi^{y-b} \text{ fall in } (0, t) \times A \right] = e^{-n^{y-b}(A)}.
\]

2. For disjoint \(A_1, \ldots, A_k \subset U_y\) with \(n^{y-b}(A_i) < \infty\) \(\forall i\), if \(A = \bigcup_{i=1}^k A_i\) and \(\tau := \inf \{ t : \Xi^{y-b}_t \in A \}\), then
\[
\mathbb{P} \left[ \Xi^{y-b}_\tau \in A_i \right] = \frac{n^{y-b}(A_i)}{n^{y-b}(A)}.
\]

**Remark 14.** Due to (38), the distribution of the local time at which the first excursion of process \(X^{y-b}\) in the set \(A\) occurs is \(\text{Exp}(n^{y-b}(A))\).

In view of (37), excursions are segments of the sample paths which avoid point \(y\), and thus we can describe their behaviour if we consider the transition semigroup of the process which is killed upon hitting \(y\). Here we need to be somewhat careful with the notation because we are already considering the excursions of a killed process \(X^{y-b}\).

**Definition 15.** For \(b \in \{0, 1\}\) we denote by \((yP^{x-b}_t)_{t \geq 0}\) the semigroup of the process \(X^{x-b}\) which is killed upon hitting \(y\), that is
\[
yP^{x-b}_tf(x) := \mathbb{E}_x \left[ f(X^{x-b}_t)1_{\{t < H_y\}} \right] = \mathbb{E}_x \left[ f(X_t)1_{\{t < H_y \wedge H_b\}} \right]
\]
and the law of \((X_{t \wedge H_y \wedge H_b})_{t \geq 0}\) by \(yP^{x-b}\) when the process is started at \(x\).
The above describes the behaviour of the process once it is “inside” the excursion, i.e. away from \( y \) and in the case which we are interested in, that is when \( y = 0, b = 1 \) and \( y = 1, b = 0 \), its resolvent has been determined in (9). It remains to specify how the excursions “enters” the state space, which is captured by the follow definition of the entrance law \( n^y_{\cdot} \) associated with the transition semigroup \((yP^y_t)_{t\geq 0}\).

**Definition 16.** For \( t > 0, A \in I \setminus \{y\} \)

\[
n^y_{\cdot}(A) := n^y_{\cdot}(\{ f \in U : f(t) \in A, t < \eta \}) .
\]

Then the following theorem (see Theorem 5 in [14]) illustrates the Markovianity of the excursion process and explains how these excursions behave.

**Theorem.** Given \( 0 < t_1 < \cdots < t_k \) and \( A_1, \ldots, A_k \subset I \setminus \{y\} \),

\[
n^y_{\cdot}\left( \{ f \in U : f_{t_i} \in A_i, \forall i = 1, \ldots, k, t_i < \eta \} \right) = \int_{A_1} \int_{A_2} \cdots \int_{A_k} n^{y_{t_1}}_{\cdot}(d x_1) P^{y_{t_2}}_{t_2-t_1}(x_1, d x_2) \cdots P^{y_{t_k}}_{t_k-t_{k-1}}(x_{k-1}, d x_k).
\]

So the excursion is started off according to the measure \( n^y_{\cdot}(\cdot) \) and propagates inside \( I \setminus \{y\} \) according to \( yP^y_{\cdot} \) until the process hits \( y \). So if we can find both the entrance law and the transition semigroup of the killed process, then we can fully characterise the excursion process.

We will in fact determine the Laplace transform of the entrance law denoted by \( n^y_{\cdot}(\cdot) \) and given by the formula

\[
n^y_{\cdot}(d x) = \int_0^\infty e^{-\lambda t} n^y_{\cdot}(d x) d t
\]

where we point out that \( n^y_{\cdot}(\cdot) \) will always be reserved for the Laplace transform of the entrance law \( n^y_{\cdot}(\cdot) \) such that no confusion can arise. The first step in determining the Laplace transform of the entrance law is the following lemma.

**Lemma 17.** For all \( \lambda \in [0, \infty) \setminus \left\{ \frac{(\theta - 1)^2}{8} \right\} \) the following holds:

\[
\begin{aligned}
\lambda n^0_{\cdot}(1) &= \frac{\Gamma(\theta)\Gamma(1-a(\lambda))\Gamma(1-b(\lambda))}{2\Gamma(1-\theta_1)\Gamma(\theta_2)\Gamma(\theta_1-a(\lambda))\Gamma(\theta_1-b(\lambda))}, \\
\lambda n^0_{\cdot}(1) &= \frac{\Gamma(\theta)\Gamma(1-a(\lambda))\Gamma(1-b(\lambda))}{2\Gamma(1-\theta_1)\Gamma(\theta_2-a(\lambda))\Gamma(\theta_2-b(\lambda))}.
\end{aligned}
\]

**Proof.** We will prove this lemma using [12], from where we get

\[
\lambda n^0_{\cdot}(1) = \frac{\Phi^b_{\lambda,-}'(y)}{s'(y)} \frac{1}{\Phi^b_{\lambda,-}(y)} - \frac{\Phi^b_{\lambda,+}'(y)}{s'(y)} \frac{1}{\Phi^b_{\lambda,+}(y)}
\]

for \( b \in \{0, 1\} \). We have

\[
\lambda n^0_{\cdot}(1) = \lim_{\eta \to 0} \lambda n^y_{\cdot}(1).
\]

According to (40) the following holds:

\[
\begin{aligned}
\lambda n^0_{\cdot}(1) &= \frac{\Phi^1_{\lambda,-}'(y)}{s'(y)} \frac{1}{\Phi^1_{\lambda,-}(y)} - \frac{\Phi^1_{\lambda,+}'(y)}{s'(y)} \frac{1}{\Phi^1_{\lambda,+}(y)} \\
&= \frac{1}{s'(y)} \frac{\Phi^1_{\lambda,-}'(y)\Phi^1_{\lambda,+}(y) - \Phi^1_{\lambda,+}'(y)\Phi^1_{\lambda,-}(y)}{\Phi^1_{\lambda,-}(y)\Phi^1_{\lambda,+}(y)} \\
&= \frac{\Gamma(\theta_1-1)\Gamma(2-\theta_2)\Gamma(\theta)}{\Gamma(\theta_1-a(\lambda))\Gamma(\theta_1-b(\lambda))} \frac{\Gamma(\theta_1)}{2\Gamma(\theta_2)\Gamma(\theta_2)} y^\theta_1(1-y)^{\theta_2}(1-y)^{\theta_1-y_\theta_1(1-y)^{\theta_2}/\Phi^1_{\lambda,-}(y)\Phi^1_{\lambda,+}(y)}.
\end{aligned}
\]
we have
\[
\frac{\Gamma(|\theta|)\Gamma(2 - \theta_2)}{2\Gamma(\theta_2)\Gamma(\theta_1 - a(\lambda))\Gamma(\theta_1 - b(\lambda)) \Phi_{\lambda, -}^1(y)\Phi_{\lambda, +}^1(y)} \cdot \frac{1}{\Phi_{\lambda, -}^0(y)\Phi_{\lambda, +}^0(y)}
\]
where to go from the second line to the third we used (A12) and (5).

Therefore, letting \( y \to 0 \) and using (32) and the fact that \( \Phi_{\lambda, -}^1(0) = 2F_1(a(\lambda), b(\lambda); \theta_1; 0) = 1 \) we get
\[
\lambda n_{\lambda}^{0-1} = \frac{\Gamma(|\theta|)\Gamma(2 - \theta_2)}{2\Gamma(\theta_2)\Gamma(\theta_1 - a(\lambda))\Gamma(\theta_1 - b(\lambda))} \frac{\Gamma(1 - b(\lambda))\Gamma(1 - a(\lambda))}{\Gamma(2 - \theta_2)\Gamma(1 - \theta_1)}.
\]
Similarly we have
\[
\lambda n_{\lambda}^{y-0} = \lim_{y \to 0} \lambda n_{\lambda}^{y-0} = 1.
\]
since from
\[
\lambda n_{\lambda}^{y-0} = \frac{(\Phi_{\lambda, -}^0(y))'}{s'(y)} \cdot \frac{1}{\Phi_{\lambda, +}^0(y)} - \frac{(\Phi_{\lambda, -}^0(y))'}{s'(y)} \cdot \frac{1}{\Phi_{\lambda, +}^0(y)}
\]
\[
= \frac{1}{s'(y)} \cdot \frac{(\Phi_{\lambda, -}^0(y))'}{\Phi_{\lambda, +}^0(y)} - \frac{(\Phi_{\lambda, -}^0(y))'}{s'(y)} \cdot \frac{1}{\Phi_{\lambda, +}^0(y)}
\]
\[
= \frac{\Gamma(1 - \theta_2)\Gamma(\theta_2)}{\Gamma(\theta_1)\Gamma(\theta_2 - a(\lambda))\Gamma(\theta_2 - b(\lambda))} \cdot \frac{\Gamma(|\theta|)}{\Phi_{\lambda, -}^0(y)\Phi_{\lambda, +}^0(y)}
\]
\[
= \frac{\Gamma(|\theta|)\Gamma(2 - \theta_1)}{2\Gamma(\theta_1)\Gamma(\theta_2 - a(\lambda))\Gamma(\theta_2 - b(\lambda))} \cdot \frac{\Gamma(1 - a(\lambda))\Gamma(1 - b(\lambda))}{\Gamma(2 - \theta_2)\Gamma(1 - \theta_1)}.
\]
it follows that
\[
\lambda n_{\lambda}^{y-0} = \frac{\Gamma(|\theta|)\Gamma(2 - \theta_1)}{2\Gamma(\theta_1)\Gamma(\theta_2 - a(\lambda))\Gamma(\theta_2 - b(\lambda))} \cdot \frac{\Gamma(1 - a(\lambda))\Gamma(1 - b(\lambda))}{\Gamma(2 - \theta_2)\Gamma(1 - \theta_1)}.
\]

**Corollary 18.** For all \( \lambda \in (0, \infty) \setminus \left\{ \frac{|\theta| - 1}{8} \right\} \) the following holds:
\[
\lim_{\lambda \to 0} \lambda n_{\lambda}^{0-1} = \frac{\Gamma(|\theta|)\Gamma(2 - |\theta|)}{2\Gamma(\theta_1)\Gamma(2 - \theta_1)} = \frac{1}{2B(\theta_1, \theta_2)B(1 - \theta_1, 1 - \theta_2)},
\]
\[
\lim_{\lambda \to 0} \lambda n_{\lambda}^{y-0} = \frac{\Gamma(|\theta|)\Gamma(2 - |\theta|)}{2\Gamma(\theta_1)\Gamma(2 - \theta_1)} = \frac{1}{2B(\theta_1, \theta_2)B(1 - \theta_1, 1 - \theta_2)}.
\]

**Proof.** Since
\[
\lim_{\lambda \to 0} a(\lambda) = |\theta| - 1 \quad \text{and} \quad \lim_{\lambda \to 0} b(\lambda) = 0
\]
we have
\[
\lim_{\lambda \to 0} \lambda n_{\lambda}^{0-1} = \frac{\Gamma(|\theta|)\Gamma(2 - |\theta|)\Gamma(1)}{2\Gamma(1 - |\theta|)\Gamma(2 - \theta_1)\Gamma(1 - \theta_1)},
\]
\[
\lim_{\lambda \to 0} \lambda n_{\lambda}^{y-0} = \frac{\Gamma(|\theta|)\Gamma(2 - |\theta|)\Gamma(1)}{2\Gamma(1 - |\theta|)\Gamma(2 - \theta_1)\Gamma(1 - \theta_1)}.
\]

According to equation (26) in [12] we have that
\[
n_{\lambda}^{y-0}(\text{excursions with infinite lifetime}) = \lim_{\lambda \to 0} \lambda n_{\lambda}^{y-0},
\]
that is, the previous Corollary gives us the rate of infinite excursions. Since infinite excursions in our setting imply the killing has occurred we have established the rate at which the switching from looking at the excursions from \( U_0 \) to excursions from \( U_1 \) and vice versa will occur.
Now we are ready to obtain the Laplace transform of the entrance law which, together with the resolvent of the killed process obtained in Theorem 9, fully characterises the Wright–Fisher diffusion via its excursions.

**Theorem 19.** For all $\lambda \in [0, \infty) \setminus \left\{ \frac{(\|\theta\| - 1)^2}{8} \right\}$ the following holds:

$$n_\lambda^{0-1}(d\lambda) = \frac{\Gamma(a(\lambda))\Gamma(b(\lambda))\Gamma(1 - a(\lambda))\Gamma(1 - b(\lambda))}{2\Gamma(1 - \theta_1)\Gamma(\theta_2)\Gamma(\theta_1 - a(\lambda))\Gamma(\theta_1 - b(\lambda))\Gamma(\theta_1)\Gamma(\theta_2)} x^{\theta_1 - 1}(1 - x)^{\theta_2 - 1}$$

$$\times \left[ \sum_{\lambda = 0}^{\infty} \binom{-\lambda}{\theta_1, \theta_2} \frac{\Gamma(\theta_1 - a(\lambda))\Gamma(\theta_1 - b(\lambda))}{\Gamma(\theta_1)\Gamma(1 - \theta_1)} \right] + \frac{\Gamma(\theta_2 - a(\lambda))\Gamma(\theta_2 - b(\lambda))}{\Gamma(\theta_2)\Gamma(1 - \theta_1)} \left( \sum_{\lambda = 0}^{\infty} \binom{-\lambda}{\theta_1, \theta_2} \frac{\Gamma(\theta_1 - a(\lambda))\Gamma(\theta_1 - b(\lambda))}{\Gamma(\theta_1)\Gamma(1 - \theta_1)} \right) d\lambda,$$

The entrance law $n_\lambda$ can then be found by applying an inverse Laplace transform.

**Proof.** As in [14, p321] for $y \in (0, 1)$ and $b \in \{0, 1\}$ we have

$$n_\lambda^{x-b}(A) = \lambda \left( n_\lambda^{x-b} \right) R_\lambda^b(A),$$

where for $b \in \{0, 1\}$, as before we denote by $R_\lambda^b$ the resolvent of process killed at $b$, that is of process $X^{x-b}$. From Lemma 6 we know the formula for $R_\lambda^b(A)$. Since all functions under the integral are bounded on the compact space the result will follow from

$$n_\lambda^{0-1}(d\lambda) = \lim_{y \to 0} n_\lambda^{y-1}(d\lambda),$$

using the result from Lemma 17.

**Remark 20.** From a theoretical point of view the Laplace transform characterises the entrance law uniquely and therefore it is equivalent to providing the law itself. In practice, inverting the Laplace transform explicitly is rarely possible and therefore is mostly done numerically. We illustrate that numerical inversion is feasible in Figure 2.

### Appendix A. Hypergeometric Functions

This appendix has all the information about the hypergeometric functions relevant for this paper. The main reference is [1]. In general all the parameters as well as the variable of the hypergeometric function can be complex numbers. In our case the parameters $a$, $b$, and $c$ will sometimes be complex, but the variable $x$ will always be a real number.

The rising factorial is defined by $(a)_n := a(a + 1) \ldots (a + n - 1)$ and we have the following trivial relation to the Gamma function: $(a)_n \Gamma(a) = \Gamma(a + n)$. For $\Re(c - a - b) > -1$, the Gauss hypergeometric function is defined by the power series

\[(A1) \quad _2F_1(a, b; c; x) := \sum_{n=0}^{\infty} \frac{(a)_n(b)_n}{(c)_n n!} x^n,
\]

which converges absolutely on $|x| \leq 1$ when $\Re(c - a - b) > 0$, and converges conditionally on $|x| < 1$ when $-1 < \Re(c - a - b) \leq 0$. The series is not well defined when $c = -m$ for $m = 0, 1, 2 \ldots$ provided $a$ or $b$ is not a negative integer $-n$ with $n < m$. When $a$ or $b$ is equal to $-n$, for $n = 0, 1, 2, \ldots$ the series reduces to a polynomial of degree $n$. This is all due to the fact that $(−n)_{n+1} = −n(−n+1) \ldots (−n+n+1−1) = 0$. 


Remark 21. In all instances of importance in this paper the parameter \( c \) will be one of \( \{ \theta_1, \theta_2, 2-\theta_1, 2-\theta_2 \} \) and therefore the series is well defined. Also in all calculations we have that \( c - a - b > -1 \), mostly \( c - a - b = 1 - \theta_2 > 0 \) or \( c - a - b = 1 - \theta_1 > 0 \) according to (2), therefore the series will not be divergent.

When \( c \neq 0, -1, -2, \ldots \) and \( c - a - b > 0 \) then we have the following important identity:

(A2) \[ _2F_1 (a, b; c; 1) = \frac{\Gamma(c)\Gamma(c - a - b)}{\Gamma(c - a)\Gamma(c - b)}. \]

Remark 22. This formula is used a lot, for example in Remark 8 where to get (31) one needs to check that \( 0 < \theta_1 - a(\lambda) - b(\lambda) = \theta_2 - |(\theta_1 - 1)| = 1 - \theta_2 \). Therefore the identity holds because of the assumption that \( \theta_2 < 1 \). Similarly equation (32) holds because \( 2 - \theta_2 - (2\theta_1 - a(\lambda) - b(\lambda) = 1 - \theta_1 > 0 \), due to symmetry equation (33) holds because \( \theta_2 < 1 \) and equation (34) because \( \theta_1 < 1 \).

Provided that the series (A1) if well defined we have the following identity:

(A3) \[ _2F_1 (a, b; c; 0) = 1. \]

The \( x \)-derivative of a hypergeometric function is again a hypergeometric function given by the following formula:

(A4) \[ \frac{d}{dx} \, _2F_1 (a, b; c; x) = \frac{ab}{c} \, _2F_1 (a + 1, b + 1; c + 1; x). \]

When using this formula one needs to be careful since \( c + 1 - (a + 1 + b + 1) = c - a - b - 1 \) reduces the value of the convergence-determining “\( c - a - b \)” compared to the original function.

Remark 23. We use this formula for example when checking the boundary conditions in equation (17) where as a derivative we get the function \( _2F_1 (a(\lambda) + 1, b(\lambda) + 1; \theta_1 + 1; y) \). In that case we have \( \theta_1 + 1 - a(\lambda) - 1 - b(\lambda) - 1 = \theta_1 - 1 - (\theta_1 + \theta_2 - 1) = -\theta_2 \in (-1, 0) \). Therefore we have the conditional convergence of the series as in (iii).
The hypergeometric differential equation is given by

\begin{equation}
(5) \quad x(1-x)\frac{d^2 f(x)}{dx^2} + (c - (a + b + 1)x) \frac{df(x)}{dx} - abf(x) = 0
\end{equation}

for some parameters \(a, b, c > 0\). Note first that the differential equation (A5) has three singularities: at 0, at 1 and at \(\infty\). We need not worry about this last case as we are only interested in cases when \(x \in [0, 1]\). If none of the numbers \(c, c-a-b, a-b\) is equal to an integer then the two linearly independent solutions around \(x = 0\) are increasing and given by

\begin{equation}
(6) \quad _2F_1(a, b; c; x), \quad x^{1-c}_2F_1(a-c+1, b-c+1; 2-c; x),
\end{equation}

whilst the solutions around \(x = 1\) are decreasing and given by

\begin{equation}
(7) \quad _2F_1(a, b; a+b+1-c; 1-x), \quad (1-x)^{c-a-b}_2F_1(c-a, c-b; c-a-b+1; 1-x).
\end{equation}

Figure 3 displays the behaviour of these functions over the interval \([0, 1]\).

**Figure 3.** The four solutions to (A5). The curve in blue is \(_2F_1(a, b; c; x)\), the one in yellow is \(x^{1-c}_2F_1(a-c+1, b-c+1; 2-c; x)\), the one in green is \(_2F_1(a, b; a+b+1-c; 1-x)\), and the one in red is \((1-x)^{c-a-b}_2F_1(c-a, c-b; c-a-b+1; 1-x)\) all for \(\theta_1 = 0.7, \theta_2 = 0.2,\) and \(\lambda = 0.01\).

**Remark 24.** In our case from (1) and (10) it follows that \(a = a(\lambda), b = b(\lambda),\) and \(c = \theta_1\). Therefore \(c\) is not an integer and also \(c - a - b = \theta_1 - (\theta_1 + \theta_2 - 1) = 1 - \theta_2\) is not an integer. To check that \(a-b = \sqrt{(|\theta| - 1)^2 - 8\lambda}\) is not an integer one needs to first note that \((|\theta| - 1)^2 \in [0, 1)\) and \(\lambda \geq 0\) therefore we only need to check if \(a - b = 0\).

Any three solutions are necessarily linearly related (see Theorem 2.3.2 in [2]); more specifically we have the following relation:

\begin{equation}
(8) \quad \frac{\Gamma(1-c)\Gamma(a+b-c+1)}{\Gamma(a-c+1)\Gamma(b-c+1)}_2F_1(a, b; c; x) + \frac{\Gamma(c-1)\Gamma(a+b-c+1)}{\Gamma(a)\Gamma(b)}x^{1-c}_2F_1(a-c+1, b-c+1; 2-c; x), \quad (1-x)^{c-a-b}_2F_1(c-a, c-b; c-a-b+1; 1-x)
\end{equation}
If we denote by \( W(f, g) = f(x)g'(x) - f'(x)g(x) \) then calculating it for \( f(x) = 2F_1(a, b; c; x) \) and \( g(x) = x^{1-c}2F_1(a + c - 1, b - c + 1; 2 - c; x) \) gives us

\[
W(f, g) = (1 - c)x^{-c}(1 - x)^{c-a-b-1}.
\]

We denote the solutions around \( x = 1 \) by \( h(x) = 2F_1(a, b; a + b + 1 - c; 1 - x) \) and \( k(x) = (1 - x)^{c-a-b}2F_1(a - c, a - c - b; c - a - b + 1; 1 - x) \). Combining (A10) with equations (A8) and (A9) and by observing that if the Wronskian of two linearly independent solutions \( W(f, g) = k \) and a third solution is given by \( r = \alpha_1 f + \alpha_2 g \) for constant \( \alpha_1, \alpha_2 \), then \( W(f, r) = \alpha_2 k \) we have that

\[
W(f, h) = \frac{\Gamma(c - 1)\Gamma(a + b - c + 1)}{\Gamma(a)\Gamma(b)}(1 - c)x^{-c}(1 - x)^{c-a-b-1},
\]

\[
W(f, k) = \frac{\Gamma(c - 1)\Gamma(c - a - b + 1)}{\Gamma(c - a)\Gamma(c - b)}(1 - c)x^{-c}(1 - x)^{c-a-b-1},
\]

\[
W(g, h) = \frac{\Gamma(1 - c)\Gamma(a + b + c + 1)}{\Gamma(a - c + 1)\Gamma(b - c + 1)}(1 - c)x^{-c}(1 - x)^{c-a-b-1},
\]

\[
W(g, k) = \frac{\Gamma(1 - c)\Gamma(c - a - b + 1)}{\Gamma(1 - a)\Gamma(1 - b)}(1 - c)x^{-c}(1 - x)^{c-a-b-1}.
\]

REFERENCES


(Paul A. Jenkins) DEPARTMENT OF STATISTICS, UNIVERSITY OF WARWICK, UNITED KINGDOM, AND DEPARTMENT OF COMPUTER SCIENCE, UNIVERSITY OF WARWICK, UNITED KINGDOM

E-mail address: p.jenkins@warwick.ac.uk

(Jere Koskela) DEPARTMENT OF STATISTICS, UNIVERSITY OF WARWICK, UNITED KINGDOM, AND SCHOOL OF MATHEMATICS, STATISTICS AND PHYSICS, NEWCASTLE UNIVERSITY, UNITED KINGDOM

E-mail address: jere.koskela@ncl.ac.uk

(Jaromir Sant) DEPARTMENT OF STATISTICS, UNIVERSITY OF OXFORD, UNITED KINGDOM

E-mail address: jaromir.sant@stats.ox.ac.uk

(Dario Spanò) DEPARTMENT OF STATISTICS, UNIVERSITY OF WARWICK, UNITED KINGDOM

E-mail address: d.spano@warwick.ac.uk

(Ivana Valentić) DEPARTMENT OF MATHEMATICS, UNIVERSITY OF ZAGREB, CROATIA

E-mail address: ivana.valentic@math.hr