

ASYMPTOTIC BEHAVIOR OF A FELLER EVOLUTION FAMILY INVOLVED IN THE FISHER–WRIGHT MODEL

ADAM BOBROWSKI,* *Polish Academy of Sciences and Lublin University of Technology*

Abstract

We study the evolution in time of the joint distribution of a pair of Feller processes, related by the fact that some random time ago they were identical, evolving as a single Feller process; from that time on, they began to evolve independently, conditional on a state at the time of split, according to the same Feller transition probabilities. Such processes are involved in the Fisher–Wright model: the distribution of the time counted backwards from the present to the time of split in the past is a function of deterministic but time-varying effective size $2N$ of the population from which the two processes are sampled. In terms of a corresponding family of Feller operators, assuming asymptotic stability or ergodicity of the process of mutation, we find the limit form of the distribution of such pairs of processes sampled from decaying, asymptotically constant, and growing populations. In the case where mutation is not asymptotically stable or ergodic, limit distributions are found for the distribution of relative differences.

Keywords: Fisher–Wright–Moran model; Feller process; evolution family; mutation; genetic drift; coalescence

2000 Mathematics Subject Classification: Primary 92D25
Secondary 47N60

1. Introduction

This paper is a continuation of the series [4], [6], [7], [9], and [10] devoted to a study of the interplay between two of the main forces in population genetics, i.e. *mutation* and *genetic drift*. As explained below, from the stochastic point of view, this study is about pairs of random processes that evolved as a single random process, and from a certain time in the past on, conditional on a state at the time of split, began to evolve independently. Joint distributions of such processes are functions of the population size which influences the power of the drift and, hence, the distribution of the time τ counted backwards from the present to the time of the split in the past. In this paper we assume a Fellerian nature to these processes, and we introduce the corresponding evolution family of operators that describes the evolution in time of the distributions of these processes, and study the asymptotic behavior of the family.

Genetic drift is often defined as a random change of the frequency of a particular *allele* (i.e. a variant of the way the chromosome may look at a particular place, this place being called the *locus*) in a finite population; in virtually infinite populations this frequency is constant. Since random events may result in not passing some part of the genetic material to the next generation, if there are no mutations, all the members of a population eventually share one allele, and all the other alleles become extinct: this is the case in the classical model due to Wright and Fisher [20], [46], and this effect of genetic drift is particularly visible in small populations.

Received 1 December 2006; revision received 7 May 2008.

* Postal address: Nadbyszczka 38A, 20-618 Lublin, Poland. Email address: a.bobrowski@pollub.pl

In the presence of mutation these two forces compete, with mutation introducing new variants to the population and drift striving to make the population uniform. Contemporary studies focus on models involving, along with drift, other forces of population genetics: mutation, selection, and recombination [13], [14], [15], [34], [42]. Such Fisher–Wright-type models are of growing interest for both mathematicians and geneticists, but their ‘forward’ mathematical analysis is quite complicated. Fortunately, as shown in the crucial papers by Kingman [31] and Tajima [45], if no selection and recombination is involved, the ‘backward’ structure of these models is quite simple [5], [20], [46]. In particular, in a large population of size $2N$ the time to the most recent common ancestor of two individuals is approximately exponential with parameter $2N$.

Kingman’s [31] approach allows analyzing *selectively neutral loci*. A locus is said to be selectively neutral if it does not play any vital role for the organism, and, hence, it is not under selective pressure. Well-known examples of such loci include some of the *microsatellites* and sequences in the *hypervariable regions 1 and 2* of mitochondrial DNA (see the references in [9] and [10]). At such loci, it is reasonable to assume that the process of mutation is independent of coalescence and, thus, may be superimposed on the ancestral lineages. This allows writing explicit equations for the joint distributions of the attributes, such as the microsatellite length or the number of substitutions in a DNA sequence, of two individuals sampled from a population. These distributions are the main focus of our study. Generalizations of Kingman’s coalescent include models with recombination, but in models involving drift, mutation, and recombination it is difficult, if possible at all, to find an equation for the joint distributions [6], [34], [39], [42].

We are interested in the model where population size varies in time. We assume that, from a time t_0 in the past onwards, the evolution of the size of the population is known and given by a function $2N : [t_0, \infty) \rightarrow \mathbb{R}^+$. We assume that $2N$ is measurable and bounded away from 0 on any finite interval. In applications, $2N$ is usually strictly positive and continuous, except maybe for a finite number of points where it has left- and right-hand limits. ‘The population size’ $2N$ is in fact the ‘relative limit population size’ and does not necessarily take even or integer values, and the factor 2 is only conventional [24]. The time viewed backwards from time $t > t_0$ to the most recent common ancestor of two chromosomes is a generally improper random variable τ_t ,

$$P(\tau_t \in \mathcal{B}) = \int_{\mathcal{B}} g(t, t - s) ds, \tag{1}$$

where \mathcal{B} is a Borel subset of $[0, t - t_0]$ and

$$g(t, s) = \frac{1}{2N(s)} \exp\left(-\int_s^t \frac{1}{2N(u)} du\right).$$

In particular, the probability that there will be no coalescence in the interval $[t_0, t)$ is

$$p(t) = 1 - \int_{t_0}^t g(t, u) du = \exp\left(-\int_{t_0}^t \frac{1}{2N(u)} du\right). \tag{2}$$

Such an approach, which is a variation of [24] and [46], allows treating populations that are not clonal, having several unrelated ancestors (e.g. bacteria or viruses). Even though the time interval $[t_0, t)$ in which such populations are observed may be long, it may happen that there is no coalescence of ancestral lineages in it. Hence, the joint distribution of attributes of interest to us observed at time $t > t_0$ depends on the distribution at t_0 (see (8), below).

2. Definition of the evolution family

2.1. Some functional analytic notions

2.1.1. *The mutation semigroup and its tensor product.* Let \mathcal{S} be a locally compact Hausdorff space; its points will be denoted by p, q , etc. Let $\{P_t, t \geq 0\}$ be a Feller semigroup on $C_0(\mathcal{S})$ with generator A [22], [40]. The same symbol, i.e. $\{P_t, t \geq 0\}$, denotes a transition family of probability kernels on \mathcal{S} associated with this semigroup, so that we have [40]

$$P_t f(p) = \int_{\mathcal{S}} f(q) P_t(p, dq), \quad p \in \mathcal{S}, t \geq 0, f \in C_0(\mathcal{S}). \tag{3}$$

This semigroup models the process of mutation on a single individual (*locus*); a state (an *allele*) of the locus is a point of \mathcal{S} .

The completion $C_0(\mathcal{S}) \otimes_{\varepsilon} C_0(\mathcal{S})$ of the algebraic tensor product $C_0(\mathcal{S}) \otimes C_0(\mathcal{S})$ of two copies of $C_0(\mathcal{S})$ in the injective tensor norm is isometrically isomorphic to $C_0(\mathcal{S} \times \mathcal{S})$; see [12]. Functions $f \in C_0(\mathcal{S} \times \mathcal{S})$ of the form

$$f(p, q) = f_1(p) f_2(q), \tag{4}$$

where $f_i \in C_0(\mathcal{S}), i = 1, 2$, are called elementary tensors and denoted by $f_1 \otimes f_2$. The tensor product semigroup $\{U_t, t \geq 0\}, U_t = P_t \otimes P_t$ in $C_0(\mathcal{S} \times \mathcal{S})$, of two copies of $\{P_t, t \geq 0\}$ is given by

$$U_t f(p, q) = \int_{\mathcal{S} \times \mathcal{S}} f(p', q') P_t(p, dp') P_t(q, dq').$$

This semigroup describes the evolution of two independent Markov processes, both with transition probabilities governed by the semigroup $\{P_t, t \geq 0\}$. The set $\mathcal{D}(A) \otimes \mathcal{D}(A)$, i.e. the set of elementary tensors (4) with both f_i in $\mathcal{D}(A)$, is a core for the generator G of $\{U_t, t \geq 0\}$, and we have $Gf = Af_1 \otimes f_2 + f_1 \otimes Af_2$ for $f_1, f_2 \in \mathcal{D}(A)$; see [2].

2.1.2. *The space of symmetric functions.* The map $Z: \mathcal{S} \times \mathcal{S} \rightarrow \mathcal{S} \times \mathcal{S}$ given by $(p, q) \mapsto Z(p, q) = (q, p)$ is a homeomorphism. Let $C_{0s}(\mathcal{S} \times \mathcal{S})$ be the subspace of $C_0(\mathcal{S} \times \mathcal{S})$ formed by symmetric functions, i.e. functions f satisfying $f \circ Z = f$, and let $\mathcal{M}(\mathcal{S} \times \mathcal{S})$ be the space of finite regular Borel measures on $\mathcal{S} \times \mathcal{S}$. In Appendix A we show that $C_{0s}(\mathcal{S} \times \mathcal{S})$ is isometrically isomorphic to the space of continuous functions $C_0(S_{\Delta})$ on a certain locally compact space S_{Δ} , and its dual is the space of symmetric measures, i.e. measures $\mu \in \mathcal{M}(\mathcal{S} \times \mathcal{S})$ equal to their transports via Z .

2.1.3. *Two operators and a semigroup.* Let $\Theta: C_0(\mathcal{S} \times \mathcal{S}) \rightarrow C_0(\mathcal{S})$ be given by $\Theta f = f \circ \phi, f \in C_0(\mathcal{S} \times \mathcal{S})$, where $\phi(p) = (p, p), p \in \mathcal{S}$. The dual operator Θ^* , mapping the space $\mathcal{M}(\mathcal{S})$ of finite regular Borel measures on \mathcal{S} into $\mathcal{M}(\mathcal{S} \times \mathcal{S})$, assigns to a measure its transport via the map $\phi: \mathcal{S} \rightarrow \mathcal{S} \times \mathcal{S}$.

Let $K^*: \mathcal{M}_s(\mathcal{S} \times \mathcal{S}) \rightarrow \mathcal{M}(\mathcal{S})$ be given by $K^* \mu(\mathcal{B}) = \mu(\mathcal{B} \times \mathcal{S})$, where \mathcal{B} is a Borel subset of \mathcal{S} (if μ is the joint distribution of a pair of exchangeable random variables then $K^* \mu$ is the (marginal) distribution of both of them). The operator K^* is the dual to $K: C_0(\mathcal{S}) \rightarrow C_{0s}(\mathcal{S} \times \mathcal{S})$ given by $(Kf)(p, q) = \frac{1}{2}(f(p) + f(q))$.

We check directly that

$$\Theta Kf = f, \quad f \in C_0(\mathcal{S}), \quad K^* \Theta^* \mu = \mu, \quad \mu \in \mathcal{M}(\mathcal{S}), \tag{5}$$

$$\text{and } U_t Kf = K P_t f, \quad t \geq 0, f \in C_0(\mathcal{S}). \tag{6}$$

We define the operators $P_t^{\boxtimes} := K P_t \Theta$, $t \geq 0$, in $C_0(\mathcal{S} \times \mathcal{S})$. By (5), $P_t^{\boxtimes} P_s^{\boxtimes} = P_{s+t}^{\boxtimes}$ and $[0, \infty) \ni t \mapsto P_t^{\boxtimes}$ is strongly continuous, yet $\{P_t^{\boxtimes}, t \geq 0\}$ is not a strongly continuous semigroup in the usual sense, as $P_0^{\boxtimes} = K \Theta \neq I := I_{C_0(\mathcal{S} \times \mathcal{S})}$ (the identity operator in $C_0(\mathcal{S} \times \mathcal{S})$). In fact, $K \Theta f(p, q) = \frac{1}{2}(f(p, p) + f(q, q))$. By (6),

$$U_t P_s^{\boxtimes} = P_{s+t}^{\boxtimes}, \quad s, t \geq 0. \tag{7}$$

2.2. The definition of the evolution family

Let us consider two exchangeable Feller processes $\{X_1(t), t \geq t_0\}$ and $\{X_2(t), t \geq t_0\}$ with values in \mathcal{S} . We assume that at each $t > t_0$ we know that at time $t - \tau_t$, these two processes were identical and evolved as a single process with transition probabilities governed by a Feller semigroup $\{P_t, t \geq 0\}$ on $C_0(\mathcal{S})$ with generator A ; the distribution of τ_t is given by (1). Furthermore, at time t_0 , the single process referred to above (which is the most recent common ancestor of our processes) was distributed according to the marginal distribution calculated from the joint distribution of the processes. Finally, we assume that from time $\sigma_t = t - \tau_t$ onward, conditional on the state at the time of split, these processes began to evolve independently yet with the same transition probabilities as before.

Suppose that at time t_0 the joint distribution of the two processes was a measure $\mu \in \mathcal{M}_s(\mathcal{S} \times \mathcal{S})$. What is their distribution μ_{t,t_0} at time $t > t_0$? With probability (2), there was no coalescence of ancestral lines in the time interval $[t_0, t]$ and the two processes evolved independently from each other all this time. Conditionally on this event, μ_{t,t_0} equals $U_{t-t_0}^* \mu$ (U_u^* is the dual to U_u). If coalescence occurred at time $s \in [t_0, t]$, i.e. if $\tau_t = t - s$, then at time s the distribution of the processes was concentrated on the diagonal (p, p) , $p \in \mathcal{S}$, and was given by $\Theta^* P_{s-t_0}^* K^* \mu$, where P_u^* is the dual to P_u . Indeed, up to time s , the two processes evolved as a single Feller process with transition probabilities $\{P_t, t \geq 0\}$, and this process at time t_0 had the distribution $K^* \mu$. From the time s onward, the processes evolved independently. Hence, conditional on coalescence at time s , the distribution is $U_{t-s}^* \Theta^* P_{s-t_0}^* K^* \mu$. Therefore,

$$\mu_{t,t_0} = p(t) U_{t-t_0}^* \mu + \int_{t_0}^t g(t, s) U_{t-s}^* \Theta^* P_{s-t_0}^* K^* \mu \, ds. \tag{8}$$

The reasoning presented above is merely formal (it may be made rigorous, for example, if \mathcal{S} is denumerable [9]); in particular, existence of the above integral requires a proof. However, in view of the assumption of the Fellerian nature of the processes involved, we are led to introducing and studying the ‘dual’ operators $S(t, s)$, $t_0 \leq s \leq t < \infty$, in $C_{0s}(\mathcal{S} \times \mathcal{S})$ defined by

$$\begin{aligned} S(t, s) f &= p_s(t) U_{t-s} f + \int_s^t g(t, v) K P_{v-s} \Theta U_{t-v} f \, dv \\ &= p_s(t) U_{t-s} f + \int_s^t g(t, v) P_{v-s}^{\boxtimes} U_{t-v} f \, dv, \end{aligned} \tag{9}$$

where $p_s(t) = \exp(-\int_s^t du/2N(u))$, so that

$$\int_s^t g(t, v) \, dv = 1 - p_s(t). \tag{10}$$

Since the function $v \mapsto K P_{v-s} \Theta U_{t-v} f \in C_{0s}(\mathcal{S} \times \mathcal{S})$ is strongly continuous and $v \mapsto g(t, v)$ is measurable, the integral in (9) is well defined as a strong Bochner integral.

Although S is a family of operators in $C_{0s}(\mathcal{X} \times \mathcal{X})$, the above definition works well for $f \in C_0(\mathcal{X} \times \mathcal{X})$, and $S(t, s)f$ thus defined belongs to $C_0(\mathcal{X} \times \mathcal{X})$. In other words, we may treat S as a family of operators in $C_0(\mathcal{X} \times \mathcal{X})$ leaving $C_{0s}(\mathcal{X} \times \mathcal{X})$ invariant.

Using (7), (9), (10), the semigroup property, and $p_s(t)g(s, w) = g(t, w)$, we check that S is an evolution family of operators in the sense that

$$S(s, u)S(t, s) = S(t, u) \quad \text{for } t_0 \leq u \leq s \leq t < \infty, \tag{11}$$

$S(s, s) = I$, and $(s, t) \mapsto S(t, s)$ is strongly continuous. It is a Feller family since it is strongly continuous, and the operators map $C_0(\mathcal{X} \times \mathcal{X})$ into itself, while leaving the positive cone invariant. We have $\|S(t, s)\| \leq 1$. The reversed order in (11) should be noted: it is the dual operators $S^*(t, s)$, $t_0 \leq s \leq t$, that satisfy the relation $S^*(t, s)S^*(s, u) = S^*(t, u)$ appearing in the definition of an evolution family common in operator theory [17], [33], [37]. The fact that in general the continuity assumption is not satisfied for S^* is precisely the reason why studying Feller evolution families as defined above is important in probability [47, Chapter 8].

By (5), (6), and the semigroup property, $K P_{v-s} \Theta U_{t-v} K f = K P_{t-s} f$, so that, by (9) and (10),

$$S(t, s)Kf = K P_{t-s} f, \quad f \in C_0(\mathcal{X}). \tag{12}$$

By (6), this expresses the fact that in the model with drift, marginal distributions of the processes are the same as in the model without drift (compare [7, Equation (5)]).

2.3. S as a perturbation of U

In order to obtain more information concerning S let us define the operators

$$\begin{aligned} S^\diamond(t, s) &:= \exp\left(\int_s^t \frac{du}{2N(u)}\right) S(t, s) \\ &= U_{t-s} f + \int_s^t \frac{1}{2N(u)} \exp\left(\int_s^u \frac{dv}{2N(v)}\right) P_{u-s}^\boxtimes U_{t-u} f \, du, \quad t_0 \leq s \leq t < \infty, \end{aligned}$$

and note that they also form an evolution family.

Proposition 1. For $t_0 \leq s \leq t$,

$$S^\diamond(t, s) = U_{t-s} + \int_s^t \frac{1}{2N(u)} U_{u-s} K \Theta S^\diamond(t, u) \, du, \tag{13}$$

$$\text{and } S^\diamond(t, s) = U_{t-s} + \int_s^t \frac{1}{2N(u)} S^\diamond(u, s) K \Theta U_{t-u} \, du. \tag{14}$$

Proof. Relation (14) follows immediately from (12) and the definition of S^\diamond . By (5) and (7), $U_{u-s} K \Theta P_{v-u}^\boxtimes U_{t-v} = U_{u-s} P_{v-u}^\boxtimes U_{t-v} = P_{v-s}^\boxtimes U_{t-v}$, and, by (6), $U_{u-s} K \Theta U_{t-u} = P_{u-s}^\boxtimes U_{t-u}$. Hence,

$$U_{u-s} K \Theta S^\diamond(t, u) = P_{u-s}^\boxtimes U_{t-u} + \int_u^t \frac{1}{2N(v)} \exp\left(\int_u^v \frac{dw}{2N(w)}\right) P_{v-s}^\boxtimes U_{t-v} \, dv.$$

Thus, changing the order of integration in the double integral obtained by inserting the definition of $S^\diamond(t, u)$ into (13) shows that the right-hand side of (13) is

$$U_{t-s} + \int_s^t \frac{1}{2N(v)} \left(1 + \int_s^v \frac{1}{2N(u)} \exp\left(\int_u^v \frac{dw}{2N(w)}\right) du\right) P_{v-s}^\boxtimes U_{t-v} \, dv = S^\diamond(t, s),$$

proving (13).

The interpretation of (13) and (14) is that S^\diamond is a perturbation of the evolution family U_{t-s} , $t_0 \leq s \leq t < \infty$, by bounded linear operators $B^\diamond(u) = K\Theta/2N(u)$ (compare [17, p. 487]) or, what is the same, that S is the perturbation of U_{t-s} , $t_0 \leq s \leq t < \infty$, by bounded linear operators $B(u) = (K\Theta - I)/2N(u)$. As the next proposition shows, this remark may be made more specific in the case where $2N$ is continuously differentiable. We omit its straightforward but lengthy proof.

Proposition 2. For $f \in \mathcal{D}(G)$ and t at which $2N$ is continuous,

$$\frac{\partial}{\partial t} S(t, s) = S(t, s) \left(Gf - \frac{1}{2N(t)} f + \frac{1}{2N(t)} K\Theta f \right), \quad t \geq s,$$

with right-hand derivative at $t = s$. Moreover, if $2N$ is continuously differentiable, for $f \in \mathcal{D}(G)$, $S(t, s)f$ belongs to $\mathcal{D}(G)$ and

$$\frac{\partial}{\partial s} S(t, s) = - \left(G - \frac{1}{2N(s)} + \frac{1}{2N(s)} K\Theta \right) S(t, s)f, \quad t_0 \leq s \leq t,$$

with left-hand derivative at $s = t$ and right-hand derivative at $s = t_0$.

By the classical result of Kato [29] (compare [17, p. 479] or [33, Chapter 2, Theorems 3.6 and 3.7]), if $2N$ is continuously differentiable then, for each $t > t_0$, the Cauchy problem in $C_0(\mathcal{X} \times \mathcal{X})$,

$$\frac{dx(s)}{ds} = Gx(s) + \frac{1}{2N(t-s)} K\Theta x(s) - \frac{1}{2N(t-s)} x(s), \tag{15}$$

$0 \leq s \leq t - t_0$ and $x(0) = f \in \mathcal{D}(G)$, is well posed. The function $[0, t - t_0] \ni s \mapsto x(s)$ is a solution to (15) if and only if $[t_0, t] \ni s \mapsto y(s) = x(t - s)$ solves the following (Kolmogorov backward) problem:

$$\frac{dy(s)}{ds} = -Gy(s) - \frac{1}{2N(s)} K\Theta y(s) + \frac{1}{2N(s)} y(s), \tag{16}$$

$t_0 \leq s \leq t$, $y(t) = f \in \mathcal{D}(G)$. Since $x(s) = S(t, t - s)$ solves (15), $y(s) = S(t, s)$ is the solution to (16).

If the population size $2N(t) = 2N$ is constant, (15) and (16) are time homogeneous and $S(t, s) = S_{t-s}^*$, where $\{S_t^*, t \geq 0\}$ is the semigroup generated by $G + K\Theta/2N - I/2N$. Then $g(t, v) = e^{-(t-v)/2N}/2N$ and (9) agrees with the form of $\{S_t^*, t \geq 0\}$ derived by means of the Phillips perturbation theorem (compare [5, p. 325]).

2.4. Examples

Example 1. (Two related standard Brownian motions.) In this case $\mathcal{X} = \mathbb{R}$,

$$P_t f(p) = \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{\infty} \exp\left(-\frac{q^2}{2t}\right) f(p + q) dq, \quad t > 0, p \in \mathbb{R},$$

and $Af = \frac{1}{2}d^2/dp^2$ with maximal domain. Then, see [41], $G = \Delta$ is the closure of $\Delta_0 = \frac{1}{2}(d^2/dp^2 + d^2/dq^2)$ and problem (16) takes the following form. Given $f_0(p, q) \in \mathcal{D}(\Delta)$, we are looking for an f such that $f(p, q, t) = f_0(p, q)$, $f(\cdot, \cdot, s) \in \mathcal{D}(\Delta)$ for each $s \in [t_0, t]$, and, for $t_0 \leq s \leq t$,

$$\frac{\partial f(p, q, s)}{\partial s} = -\Delta f(p, q, s) - \frac{1}{4N(s)} (f(p, p, s) + f(q, q, s)) + \frac{1}{2N(s)} f(p, q, s).$$

Example 2. (Two related simple symmetric random walks.) In this case A is a bounded operator in $C_0(\mathbb{Z})$ given by $Af(p) = \frac{1}{2}f(p + 1) + \frac{1}{2}f(p - 1) - f(p)$. The operator G is bounded as well, and $Gf(p, q) = \frac{1}{2}(f(p + 1, q) + f(p - 1, q) + f(p, q + 1) + f(p, q - 1)) - 2f(p, q)$. Problem (16) takes the form

$$\frac{\partial f(p, q, s)}{\partial s} = -\frac{1}{2}(f(p + 1, q, s) + f(p - 1, q, s) + f(p, q + 1, s) + f(p, q - 1, s)) - \frac{1}{4N(s)}(f(p, p, s) + f(q, q, s)) + \left(\frac{1}{2N(s)} + 2\right)f(p, q, s),$$

$t_0 \leq s \leq t$, with ‘final condition’ $f(p, q, t) = f_0(p, q)$, where $f_0 \in C_0(\mathbb{Z}^2)$.

3. Asymptotic behavior

In this section we study the asymptotic behavior of $S(t, s)$, $t_0 \leq s \leq t < \infty$, as a function of $2N$. Cases of interest are (i) a decaying population, where $\lim_{t \rightarrow \infty} 2N(t) = 0$, (ii) an asymptotically constant population, where $\lim_{t \rightarrow \infty} 2N(t) = 2N_0 > 0$, (iii) slow growth to ∞ , where $\lim_{t \rightarrow \infty} 2N(t) = \infty$ while $\int_{t_0}^{\infty} dt/2N(t) = \infty$, and (iv) fast growth to ∞ , where $\lim_{t \rightarrow \infty} 2N(t) = \infty$ while $\int_{t_0}^{\infty} dt/2N(t) < \infty$. Examples of (iii) and (iv), respectively, are linear growth $2N(t) = t + a$, where $a > -t_0$, and exponential growth $2N(t) = Me^{\omega t}$, where $M, \omega > 0$. For importance of linear growth, see [3].

3.1. Convergence of $S(t, s)$ as $t \rightarrow \infty$

This subsection and the whole section is centered around Theorem 1, below. In the preliminary Proposition 3, below, we show that, for convergence of $S(t, s)$ as $t \rightarrow \infty$, it is necessary for $\{P_t, t \geq 0\}$ to be asymptotically stable in the sense of having a limit P_∞ at ∞ . (Sufficient conditions for asymptotic stability are given in [35].) In Theorem 1 we show that in the main four scenarios of population size behavior, asymptotic stability of $\{P_t, t \geq 0\}$ implies convergence of $S(t, s)$ as $t \rightarrow \infty$, and we provide an explicit form of the limit in terms of population size.

Proposition 3. *Suppose that the strong limit, $\lim_{t \rightarrow \infty} S(t, s)$, exists for some $s \geq t_0$. Then the limit*

$$P_\infty := \lim_{t \rightarrow \infty} P_t \tag{17}$$

exists.

Proof. By (5) and (12), $\Theta S(t, s)K = \Theta K P_{t-s} = P_{t-s}$.

Lemma 1. *Let $\mu_t, t > s$, be the measure on $[0, t - s)$ with density $g(t, t - \cdot)$. Assume that the population decays. Then, as $t \rightarrow \infty$, μ_t converges weakly to the Dirac measure at 0.*

Proof. We have $g(t, t - v) = -(d/dv) \exp(-\int_{t-v}^t dw/2N(w))$. Hence,

$$\int_0^{t-s} g(t, t - v) dv = 1 - \exp\left(-\int_s^t \frac{dw}{2N(w)}\right) \rightarrow 1 \quad \text{as } t \rightarrow \infty$$

and, for any $\delta \in (0, t - s)$,

$$\int_0^\delta g(t, t - v) dv = 1 - \exp\left(-\int_{t-\delta}^t \frac{dw}{2N(w)}\right) \rightarrow 1 \quad \text{as } t \rightarrow \infty.$$

Lemma 2. *Suppose that the population is asymptotically constant, and let $\lambda = 1/2N_0$. Then $\lim_{t \rightarrow \infty} \int_0^{t-s} |\lambda e^{-\lambda v} - g(t, t-v)| dv = 0$.*

Proof. Since $\lim_{t \rightarrow \infty} g(t, t-v) = \lambda e^{-\lambda v}$ for all $v \geq 0$ and $\lim_{t \rightarrow \infty} \int_0^t g(t, t-v) dv = \lim_{t \rightarrow \infty} \lambda \int_0^{t-s} e^{-\lambda v} dv = 1$, the lemma follows by Scheffé’s theorem.

Theorem 1. *Suppose that limit (17) exists. Then the strong limit $U_\infty = \lim_{t \rightarrow \infty} U_t$ also exists. Moreover, in the four main scenarios of population size behavior, the limit $S(\infty, s) := \lim_{t \rightarrow \infty} S(t, s)$ exists for all $s \geq t_0$. Finally, except in the case of fast growth, the limit $S(\infty) = S(\infty, s)$ does not depend on s and*

- (a) *in the case of a decaying population, $S(\infty) = K P_\infty \Theta$;*
- (b) *in the case of an asymptotically constant population, $S(\infty) = \lambda K P_\infty \Theta (\lambda - G)^{-1}$, where $\lambda = (2N_0)^{-1}$;*
- (c) *in the case of slow growth to ∞ , $S(\infty) = K P_\infty \Theta U_\infty$;*
- (d) *in the case of fast growth to ∞ ,*

$$S(\infty, s) = \exp\left(-\int_s^\infty \frac{dv}{2N(v)}\right) U_\infty + \int_s^\infty \frac{1}{2N(u)} \exp\left(-\int_u^\infty \frac{dv}{2N(v)}\right) P_{u-s}^\otimes du U_\infty. \tag{18}$$

Proof. The limit U_∞ exists because on elementary tensors $f \otimes g$, where $f, g \in C_0(\mathcal{S})$, we have $\lim_{t \rightarrow \infty} U_t[f \otimes g] = \lim_{t \rightarrow \infty} P_t f \otimes P_t g = P_\infty f \otimes P_\infty g$, such tensors form a linearly dense set in $C_0(\mathcal{S} \times \mathcal{S})$, and $\|U_t\| \leq 1, t \geq 0$.

(a) We omit this argument as it is a simpler version of the one used to prove Proposition 5, below. The only difference is that here we use

$$S(t, s)f = p_s(t)U_{t-s}f + \int_0^{t-s} g(t, t-v)P_{t-s-v}^\otimes U_v f dv, \tag{19}$$

which is a version of (9), instead of (23), below, used in Proposition 5.

(b) Since $\|p_s(t)U_{t-s}\| \leq p_s(t)$, the first term in (19) vanishes as $t \rightarrow \infty$. By the dominated convergence theorem, $\lambda K P_\infty \Theta (\lambda - G)^{-1} = \lim_{t \rightarrow \infty} \lambda \int_0^{t-s} e^{-\lambda v} P_{t-s-v}^\otimes U_v dv$. However, the distance between this last integral and the integral in (19) does not exceed $\int_0^{t-s} |\lambda e^{-\lambda v} - g(t, t-v)| dv$, which in Lemma 2 was proved to vanish as $t \rightarrow \infty$.

(c) As in the proof of (b), the first term in (19) vanishes as $t \rightarrow \infty$. On the other hand, as in Lemma 1, $\lim_{t \rightarrow \infty} \int_0^{t-s} g(t, t-v) dv = 1$. Moreover, for any $\delta \in (0, t-s)$,

$$\int_0^\delta g(t, t-v) dv = 1 - \exp\left(-\int_{t-\delta}^t \frac{dw}{2N(w)}\right) \rightarrow 0 \quad \text{as } t \rightarrow \infty$$

and

$$\int_{t-s-\delta}^{t-s} g(t, t-v) dv = \exp\left(-\int_{s+\delta}^t \frac{dw}{2N(w)}\right) - \exp\left(-\int_s^t \frac{dw}{2N(w)}\right) \rightarrow 0 \quad \text{as } t \rightarrow \infty.$$

In other words, both the measures μ_t of Lemma 1 and the measures ν_t , where ν_t is the transport of μ_t via the map $v \mapsto t-s-v$, when considered as measures on $[0, \infty]$, converge to the Dirac measure at ∞ (see Remark 3, below).

Given $f \in C_0(\mathcal{X} \times \mathcal{X})$ and $\varepsilon > 0$, we may choose a $\delta > 0$ so that $\|U_v f - U_\infty f\| \leq \varepsilon/3$ and $\|P_v \Theta U_\infty f - P_\infty \Theta U_\infty f\| \leq \varepsilon/3$ for $v > \delta$. Next, we may choose a large $t_1 > 2\delta + s$ so that

$$\int_0^\delta g(t, t - v) \, dv + \int_{t-s-\delta}^{t-s} g(t, t - v) \, dv \leq \frac{\varepsilon}{3\|f\|} \quad \text{for } t \geq t_1.$$

Then, for $v \in [\delta, t - s - \delta]$,

$$\begin{aligned} \|K P_{t-s-v} \Theta U_v f - K P_\infty \Theta U_\infty f\| &\leq \|K P_{t-s-v} \Theta U_v f - K P_{t-s-v} \Theta U_\infty f\| \\ &\quad + \|K P_{t-s-v} \Theta U_\infty f - K P_\infty \Theta U_\infty f\| \\ &\leq \frac{2}{3}\varepsilon. \end{aligned}$$

Writing $K P_\infty \Theta U_\infty$ as $\lim_{t \rightarrow \infty} \int_\delta^{t-s-\delta} g(t, t - v) K P_\infty \Theta U_\infty \, dv$ and splitting the integral in (19) into two integrals over $[0, \delta] \cup [t - s - \delta, t - s]$ and $[\delta, t - s - \delta]$, we see that, for $t > t_1$, the upper limit of the distance between $K P_\infty \Theta U_\infty$ and the integral in (19) does not exceed $\varepsilon/3 + \frac{2}{3}\varepsilon = \varepsilon$.

(d) The first term in (9) converges to $\exp(-\int_s^\infty dv/2N(v))U_\infty$. Moreover, the integral in (9) may be written as

$$\exp\left(\int_t^\infty \frac{dv}{2N(v)}\right) \int_s^t \frac{1}{2N(u)} \exp\left(-\int_u^\infty \frac{dv}{2N(v)}\right) P_{u-s}^\otimes U_{t-u} \, du.$$

This last integral converges, by the Lebesgue dominated convergence theorem, to the second term in (18), while $\lim_{t \rightarrow \infty} \exp(\int_t^\infty dv/2N(v)) = 1$.

Example 3. A probabilistic interpretation of (17) is that distributions of the process of mutation converge weakly to an invariant distribution [21]. A typical example here is the Ornstein–Uhlenbeck process [40], where the invariant distribution is normal. In this context Theorem 1(d) says that in the case of a rapidly growing population the joint distribution of a pair of Ornstein–Uhlenbeck processes related via family ties converges weakly to the distribution of a pair of independent normal random variables. Theorem 1(a) says that if the population decays, the limit distribution is that of a pair of two exact copies of a single normal random variable. In Theorem 1(c) and (d), the limit distribution is a distribution of a pair of two related normal random variables with correlation decreasing as $2N$ increases. It is interesting that even for $2N$ growing to ∞ , but slowly, some correlation of the processes involved is still visible.

3.2. Convergence of averages

In applications, not many mutations may be modeled by asymptotically stable semigroups. Hence, in this subsection we consider a more general case where the semigroup is ergodic [1], [17], [22] to study strong convergence in $C_0(\mathcal{X} \times \mathcal{X})$ of averages

$$A_{s,t} = \frac{1}{t-s} \int_s^t S(t, u) \, du, \quad t_0 \leq s \leq t < \infty, \tag{20}$$

and

$$A_{s,t}^\diamond = \frac{1}{t-s} \int_s^t S(u, s) \, du, \quad t_0 \leq s \leq t < \infty. \tag{21}$$

In Proposition 4, below, we show that, for convergence of any of these averages, it is necessary that $\{P_t, t \geq 0\}$ is ergodic. In Propositions 5 and 6, below, we show that in the case of decaying and asymptotically constant populations ergodicity of the semigroup is sufficient for convergence of $A_{s,t}$. In Proposition 7, below, we show that in the fast growth case ergodicity of the semigroup $\{U_t, t \geq 0\}$ implies convergence of $A_{s,t}^\diamond$. It is still an open problem what happens to $A_{s,t}$ in the case of an expanding population and what happens to $A_{s,t}^\diamond$ in the cases of a decaying, asymptotically constant, and slowly growing population. Our Remark 3, below, explains why it is difficult to treat these cases.

For $t > 0$, let $\bar{P}_t := (1/t) \int_0^t P_u \, du$. We have $\bar{P}_t \in \mathcal{L}(C_0(\mathcal{X}))$ with $\|\bar{P}_t\| \leq 1$. (The converse is also true [16]: if \bar{P}_t are contractions then so are P_t .)

Proposition 4. *Suppose that one of the strong limits $\lim_{t \rightarrow \infty} A_{s,t}$ and $\lim_{t \rightarrow \infty} A_{s,t}^\boxtimes$ exists for some $s \geq t_0$. Then $\{P_t, t \geq 0\}$ is ergodic in that the strong limit*

$$\lim_{t \rightarrow \infty} \bar{P}_t =: \bar{P}_\infty \tag{22}$$

exists.

Proof. By (5) and (12), $\Theta A_{s,t} K = \Theta A_{s,t}^\boxtimes K = \Theta K \bar{P}_{t-s} = \bar{P}_{t-s}$.

Our next two propositions are devoted to averages $A_{s,t}$. In order to obtain a more explicit formula for these operators, we use (19) with $s = u$, integrate from $u = s$ to $u = t$, and change the order of summation in the double integral. This gives

$$A_{s,t} = \frac{1}{t-s} \int_s^t p_u(t) U_{t-u} \, du + \int_0^{t-s} g(t, t-v) K \frac{t-s-v}{t-s} \bar{P}_{t-s-v} \Theta U_v \, dv. \tag{23}$$

If limit (20) exists for some $s \geq t_0$ then it exists for all $s \geq t_0$, and all these limits are equal. The limit operator is then denoted A_∞ .

Proposition 5. *Suppose that limit (22) exists and that the population decays. Then the limit A_∞ exists and*

$$A_\infty = K \bar{P}_\infty \Theta. \tag{24}$$

Proof. The first integral in (23) converges to 0, for $p_u(t)$ converges to 0 and $\|U_{t-u}\| \leq 1$. Given $f \in C_0(\mathcal{X} \times \mathcal{X})$ and $\varepsilon > 0$, we choose $\delta > 0$ so that $\|U_v f - f\| \leq \varepsilon/3$ for $v \in [0, \delta]$. Next, we choose t_1 so large that $\int_\delta^{t-s} g(t, t-v) \, dv \leq \varepsilon/3 \|f\|$ and $\|\bar{P}_t \Theta f - \bar{P}_\infty \Theta f\| \leq \varepsilon/3$ for all $t > t_1 - s - \delta$ (this is possible by Lemma 1). For such t and v ,

$$\left\| K \frac{t-s-v}{t-s} \bar{P}_{t-s-v} \Theta U_v f - K \frac{t-s-v}{t-s} \bar{P}_\infty \Theta f \right\| < \frac{2\varepsilon}{3},$$

and so

$$\left\| K \frac{t-s-v}{t-s} \bar{P}_{t-s-v} \Theta U_v f - K \bar{P}_\infty \Theta f \right\| < \frac{2\varepsilon}{3} + \frac{\delta}{t-s} \|f\|.$$

Writing the right-hand side of (24) as $\lim_{t \rightarrow \infty} \int_0^\delta g(t, t-v) K \bar{P}_\infty \Theta \, dv$ and splitting the second integral in (23) into integrals over $[0, \delta]$ and $[\delta, t-s]$, we see that, as $t \rightarrow \infty$, the upper limit of the distance between this integral and the right-hand side of (24) does not exceed $\limsup_{t \rightarrow \infty} [2\varepsilon/3 + (\delta/(t-s)) \|f\| + \varepsilon/3] = \varepsilon$.

Proposition 6. (Compare [7, Proposition 3].) *Suppose that limit (22) exists and that the population is asymptotically constant. Then the limit A_∞ exists and*

$$A_\infty = \lambda K \bar{P}_\infty \Theta(\lambda - G)^{-1}, \quad \text{where } \lambda = (2N_0)^{-1}. \tag{25}$$

Proof. As in Proposition 5, the first integral in (23) converges to 0 as $t \rightarrow \infty$, for $p_u(t)$ converges to 0 and $\|U_{t-u}\| \leq 1$. In view of $\|((t-s-v)/(t-s))\bar{P}_{t-v-s}\| \leq \|\bar{P}_{t-s-v}\| \leq 1$ for $v \in [0, t-s]$ and the Lebesgue dominated convergence theorem, the right-hand side of (25) is the limit of

$$\lambda \int_0^{t-s} e^{-\lambda v} K \frac{t-s-v}{t-s} \bar{P}_{t-s-v} \Theta U_v \, dv \quad \text{as } t \rightarrow \infty.$$

Consequently, it suffices to show that the distance between this integral and the second integral in (23) converges to 0. This distance, however, does not exceed $\int_0^{t-s} |g(t, t-v) - \lambda e^{-\lambda v}| \, dv$. Lemma 2 completes the proof.

Remark 1. Propositions 5 and 6 agree in the sense that (24) may be obtained from (25) by letting $\lambda \rightarrow \infty$.

We turn to the averages $A_{s,t}^\diamond$. Let $\tilde{s} \geq s \geq t_0$ be given. Using (11), we write

$$A_{s,t}^\diamond = \frac{1}{t-s} \int_s^{\tilde{s}} S(u, s) \, du + \frac{t-\tilde{s}}{t-s} S(\tilde{s}, s) \frac{1}{t-\tilde{s}} \int_s^t S(u, \tilde{s}) \, du,$$

to see that if $A_{\tilde{s},\infty}^\diamond := \lim_{t \rightarrow \infty} A_{\tilde{s},t}^\diamond$ exists (strongly) then so does $A_{s,\infty}^\diamond := \lim_{t \rightarrow \infty} A_{s,t}^\diamond$ and $A_{s,\infty}^\diamond = S(\tilde{s}, s) A_{\tilde{s},\infty}^\diamond$.

For Proposition 7, below, we would like to know that ergodicity of $\{P_t, t \geq 0\}$ implies ergodicity of $\{U_t, t \geq 0\}$, i.e. that the existence of limit (22) implies the existence of the (strong) limit

$$\lim_{t \rightarrow \infty} \bar{U}_t := \bar{U}_\infty, \quad \text{where } \bar{U}_t = \frac{1}{t} \int_0^t U_u \, du. \tag{26}$$

Even though this conjecture seems to be a natural one, its proof has so far eluded us. (Szucs' [44] result concerns the two-parameter semigroup $\{P_t \otimes P_s, s, t \geq 0\}$ and not the one-parameter semigroup $\{U_t = P_t \otimes P_t, t \geq 0\}$, and shows that the strong limit $\lim_{s,t \rightarrow \infty} (1/st) \int_0^s \int_0^t P_u \otimes P_v \, du \, dv$ exists.) Hence, in Proposition 7, instead of assuming the existence of limit (22) we assume the existence of limit (26). (If \mathcal{S} is compact, $\mathbf{1}_\mathcal{S}$ belongs to $C_0(\mathcal{S})$. Then taking a simple tensor $f \otimes \mathbf{1}_\mathcal{S}$, we check that ergodicity of $\{U_t, t \geq 0\}$ implies ergodicity of $\{P_t, t \geq 0\}$.)

Proposition 7. *Suppose that limit (26) exists and that the population grows rapidly to ∞ . Then $A_{s,\infty}^\diamond$ exists for all $s \geq t_0$ and*

$$A_{s,\infty}^\diamond = \exp\left(-\int_s^\infty \frac{dv}{2N(v)}\right) \bar{U}_\infty + \int_s^\infty \frac{1}{2N(u)} \exp\left(-\int_u^\infty \frac{dv}{2N(v)}\right) P_{u-s}^{\boxtimes} \, du \bar{U}_\infty. \tag{27}$$

Proof. Substituting (9) with $t = u$ into (21) and changing the order of integration in the resulting double integral, we obtain

$$A_{s,t}^\diamond = \frac{1}{t-s} \int_0^{t-s} p_s(u+s) U_u \, du + \int_s^t \left(\frac{1}{2N(v)} P_{v-s}^{\boxtimes} \frac{t-v}{t-s} \frac{1}{t-v} \int_0^{t-v} \exp\left(-\int_v^{u+v} \frac{dw}{2N(w)}\right) U_u \, du \right) \, dv. \tag{28}$$

The second integrand does not exceed $\exp(\int_{t_0}^\infty dw/2N(w))\phi(v)$, where $\phi(v) = (1/2N(v)) \times \exp(-\int_v^\infty dw/2N(w))$ is integrable in $[s, \infty]$ with

$$\int_s^\infty \phi(v) dv = 1 - \exp\left(-\int_s^\infty \frac{dw}{2N(w)}\right).$$

Hence, by the dominated convergence theorem, we need to prove that, for any $v \geq s$, the strong limit, as $t \rightarrow \infty$, of

$$V_{t,v} := \frac{1}{t-v} \int_0^{t-v} \exp\left(-\int_v^{u+v} \frac{dw}{2N(w)}\right) U_u du$$

exists and equals $\exp(-\int_v^\infty dw/2N(w))\bar{U}_\infty$. To this end, given $\varepsilon > 0$, we choose k so large that $u + v \geq k$ implies that $\exp(-\int_v^{u+v} dw/2N(w)) - \exp(-\int_v^\infty dw/2N(w)) < \varepsilon$. Then, splitting $V_{t,v}$ and \bar{U}_{t-v} into two integrals over $[0, k]$ and $[k, t - v]$, respectively, we see that $\|V_{t,v} - \exp(-\int_v^\infty dw/2N(w))\bar{U}_{t-v}\| \leq 2k/(t - v) + \varepsilon$, proving convergence of the second integral in (28) to the second term in (27). The rest is done similarly.

Remark 2. As a by-product of the proof, we see that if limit (26) exists in the operator topology (i.e. if the semigroup is uniformly mean ergodic [17]) then so does limit (27).

Remark 3. The proofs presented in the previous and present subsections are given in the spirit of [9]. A somewhat different approach is presented in [4]. To explain this approach in more detail, we note that the limiting behavior of $S(t, t_0)$ is closely related to weak convergence, as $t \rightarrow \infty$, of the pair of random variables (τ_t, σ_t) , where $\sigma_t = t - t_0 - \tau_t$ and τ_t is a generally improper random variable τ_t with distribution given by (1). In fact, by (9), $S(t, t_0)f = p(t)(U_{t-s}f - K \Theta U_{t-s}f) + E K P_{\sigma_t} \Theta U_{\tau_t} f$, where E stands for the expected value. As $t \rightarrow \infty$, the pairs (τ_t, σ_t) converge weakly to a pair (τ, σ) distributed in $[0, \infty] \times [0, \infty]$ according to Table 1 (see [4]). This fact may be used to offer another proof of Theorem 1, and sheds further light on the results of the present section. For example, the key to the proofs of Propositions 5 and 6 is (23); the second integral in this equation is almost the same as $E K ((t - s - \tau_t)/(t - s)) \bar{P}_{\sigma_t} \Theta U_{\tau_t}$. This allows calculating the limit, as $t \rightarrow \infty$, if τ is finite (the first two rows in the table); when τ is infinite, this formula is of no help. The key to the proof of Proposition 7 is (28), which allows calculating the limit only in the case where σ is finite.

TABLE 1.

Behavior of $N(t)$	Random variable τ	Random variable σ
$\lim_{t \rightarrow \infty} N(t) = 0$	0	∞
$\lim_{t \rightarrow \infty} N(t) = N,$ $0 < N < \infty$	Exponential with parameter $2N$	∞
$\lim_{t \rightarrow \infty} N(t) = \infty,$ $\int_{t_0}^\infty du/2N(u) = \infty$	∞	∞
$\lim_{t \rightarrow \infty} N(t) = \infty,$ $\int_{t_0}^\infty du/2N(u) < \infty$	∞	Finite $P(\sigma > w) = 1 - \exp(-\int_{t_0+w}^\infty du/2N(u))$ $P(\sigma = 0) = \exp(-\int_{t_0}^\infty du/2N(u))$

4. Coherence

For the (unrestricted, symmetric) simple random walk or the standard Brownian motion, the semigroup $\{P_t, t \geq 0\}$ is neither asymptotically stable nor ergodic. Although a pair of simple random walks related via family ties has no stationary distribution, these processes are *coherent*. This has been observed by Moran [36]. Inspired by Moran, Kingman [30] considered a discrete-time Fisher–Wright model with constant population size $2N$ and mutation in the form of a random walk, and showed that, as $n \rightarrow \infty$, the vector $X_{n,2} - X_{n,1}, X_{n,3} - X_{n,1}, \dots, X_{n,2N} - X_{n,1}$ converges in distribution; here $X_{n,i}$ is the i th member of the n th generation of the population, $i = 1, \dots, 2N$, and $n \geq 1$. In this section we want to prove a similar result in the case where the population varies in time and its individuals are represented as random processes with values in a locally compact group. We assume that mutation follows a Lévy process in this group. At each instant of time, we draw a pair of processes from such a population and study the asymptotic behavior of the distribution of the difference between these processes, with the difference taken in the sense of the underlying group.

4.1. Mutations in the form of a Lévy process

We assume that \mathcal{S} is a topological group and that the process of mutation is modeled as a Lévy process in \mathcal{S} . Specifically, we assume that we are given a continuous convolution semigroup $\{m_t, t \geq 0\}$ of probability measures on \mathcal{S} . This, by definition, means that (a) $m_t * m_s = m_{t+s}$, $s, t \geq 0$, (b) $m_0 = \delta_e$, where e is the neutral element of \mathcal{S} and δ_e is the corresponding Dirac measure, and (c) $\lim_{t \rightarrow 0+} m_t = m_0$ in the weak* topology. We recall that $m * n$ is the transport of the product measure $m \otimes n$ on $\mathcal{S} \times \mathcal{S}$ via the map $\mathcal{S} \times \mathcal{S} \ni (p, q) \mapsto pq \in \mathcal{S}$. The corresponding semigroup $\{P_t, t \geq 0\}$ on $C_0(\mathcal{S})$, given by

$$P_t f(p) = \int_{\mathcal{S}} f(pq)m_t(dq), \quad t \geq 0, f \in C_0(\mathcal{S}), \tag{29}$$

is strongly continuous [26]. Equation (29) is a special case of (3) with $P_t(p, B) = m_t(p^{-1}B)$, where B is a Borel subset of \mathcal{S} and $p^{-1}B = \{q; pq \in B\}$. If \mathcal{S} is a Lie group then the generator of $\{P_t, t \geq 0\}$ is described by Hunt’s theorem [26]. The corresponding tensor product semigroup $\{U_t, t \geq 0\}$ is given by

$$U_t f(p, q) = \int_{\mathcal{S}} \int_{\mathcal{S}} f(pp', qq')m_t(dp')m_t(dq'), \quad t \geq 0, f \in C_0(\mathcal{S} \times \mathcal{S}). \tag{30}$$

4.2. The difference operator D

Let $\{X_1(t), t \geq t_0\}$ and $\{X_2(t), t \geq t_0\}$ be two \mathcal{S} -valued processes sampled from the Fisher–Wright population. Since \mathcal{S} is not assumed to be abelian, there are four ways in which a difference between these processes may be defined: (a) $(X_1(t))^{-1}X_2(t)$, (b) $X_2(t)(X_1(t))^{-1}$, (c) $(X_2(t))^{-1}X_1(t)$, or (d) $X_1(t)(X_2(t))^{-1}$. Since the processes are exchangeable, though, the distributions of (a) and (c) are the same, and so are the distributions of (b) and (d). Also, definition (b) does not agree well with (29). We could work with (b), but then we would need to replace (29) with $P_t f(p) = \int_{\mathcal{S}} f(qp)m_t(dq)$; the results would be analogous. Hence, in what follows we will consider only (a).

In order to find the distribution of $(X_1(t))^{-1}X_2(t)$, we would need to apply an operator D^* to the joint distribution of $X_1(t)$ and $X_2(t)$; given a symmetric probability measure m on $\mathcal{S} \times \mathcal{S}$, D^* maps it into the transport of m via the map $\mathcal{S} \times \mathcal{S} \ni (p, q) \mapsto p^{-1}q$. Hence, we would like to consider $S(t, s)D$, where $D: C_0(\mathcal{S}) \rightarrow C_{0\mathcal{S}}(\mathcal{S} \times \mathcal{S})$ is a preduel to D^* . For a bounded

measurable f on \mathcal{S} , let

$$Df(p, q) = \frac{1}{2}(f(p^{-1}q) + f(q^{-1}p)), \quad p, q \in \mathcal{S}.$$

Then Df is symmetric and $\int_{\mathcal{S} \times \mathcal{S}} Df \, dm = \int_{\mathcal{S}} f \, dD^*m$. However, D does not map $C_0(\mathcal{S})$ into $C_0(\mathcal{S} \times \mathcal{S})$; for example, if $f \in C_0(\mathbb{R})$ with $f(0) \neq 0$ then $Df(p, p) = f(0) \neq 0$, $p \in \mathbb{R}$. Fortunately, as Lemma 3, below, shows, if the left and right uniform structures on \mathcal{S} are equivalent, D maps the space $C_u(\mathcal{S})$ of bounded, uniformly continuous functions on \mathcal{S} into $C_u(\mathcal{S} \times \mathcal{S})$. Furthermore, see [26], (29) defines a strongly continuous semigroup in $C_u(\mathcal{S})$. Since (30) may be seen as a particular case of (29) (with \mathcal{S} replaced by $\mathcal{S} \times \mathcal{S}$ and m_t replaced by $m_t \otimes m_t$), the tensor product semigroup $\{U_t, t \geq 0\}$ maps $C_u(\mathcal{S} \times \mathcal{S})$ into itself. (Here, $\mathcal{S} \times \mathcal{S}$ is a group with componentwise multiplication and product topology.) Hence, we consider $S(t, s)D$ as operators on $C_u(\mathcal{S})$.

We recall that the left or right uniform structure on a locally compact group is the family of sets $L_{\mathcal{N}}$ or, respectively, $R_{\mathcal{N}}$, where \mathcal{N} is a neighborhood of the neutral element e of \mathcal{S} , $L_{\mathcal{N}}$ is the set of all $(p, q) \in \mathcal{S} \times \mathcal{S}$ such that $p^{-1}q \in \mathcal{N}$, and $R_{\mathcal{N}}$ is the set of all $(p, q) \in \mathcal{S} \times \mathcal{S}$ such that $qp^{-1} \in \mathcal{N}$ (see [25]). A real- or complex-valued function f on \mathcal{S} is said to be uniformly continuous with respect to the left or right uniform structure if, for any $\varepsilon > 0$, there exists a neighborhood \mathcal{N} such that $(p, q) \in L_{\mathcal{N}}$ or, respectively, $(p, q) \in R_{\mathcal{N}}$ implies that $|f(p) - f(q)| < \varepsilon$. The space $C_0(\mathcal{S})$ is a subspace of $C_u(\mathcal{S})$, the space of real bounded functions that are uniformly continuous with respect to left (or right) uniform structures.

Lemma 3. *Suppose that the left and right uniform structures on \mathcal{S} are equivalent. Then D maps $C_u(\mathcal{S})$ into $C_u(\mathcal{S} \times \mathcal{S})$.*

Proof. Let $f \in C_u(\mathcal{S})$. Since f is uniformly continuous with respect to the left uniform structure, given $\varepsilon > 0$, we may find a neighborhood \mathcal{N} of e such that $|f(p) - f(q)| < \varepsilon$ whenever $p^{-1}q \in \mathcal{N}$. Next, the left and right structures being equivalent, we may choose \mathcal{N}' such that $p^{-1}q \in \mathcal{N}'$ implies that $qp^{-1} \in \mathcal{N}$. Since the map $(p, q) \mapsto pq$ is continuous, there is a neighborhood \mathcal{N}'' of e such that $p, q \in \mathcal{N}''$ implies that $pq \in \mathcal{N}'$. Moreover, \mathcal{N}'' may be assumed to have the property that $p \in \mathcal{N}''$ implies that $p^{-1} \in \mathcal{N}''$ (otherwise we may define a new \mathcal{N}'' as $\mathcal{N}'' \cap (\mathcal{N}'')^{-1}$). Finally, there exists a neighborhood \mathcal{N}''' of e such that $qp^{-1} \in \mathcal{N}''$ whenever $p^{-1}q \in \mathcal{N}'''$. For p_1, p_2, q_1 , and q_2 in \mathcal{S} , we have

$$\begin{aligned} p_1^{-1}p_2 \in \mathcal{N}''' \quad \text{and} \quad q_1^{-1}q_2 \in \mathcal{N}''' &\implies p_2p_1^{-1} \in \mathcal{N}'' \quad \text{and} \quad q_2q_1^{-1} \in \mathcal{N}'' \\ &\implies p_1p_2^{-1} \in \mathcal{N}'' \quad \text{and} \quad q_2q_1^{-1} \in \mathcal{N}'' \\ &\implies p_1p_2^{-1}q_2q_1^{-1} \in \mathcal{N}' \\ &\implies q_1^{-1}p_1p_2^{-1}q_2 \in \mathcal{N} \\ &\implies |f(p_1^{-1}q_1) - f(p_2^{-1}q_2)| < \varepsilon. \end{aligned}$$

By symmetry, $p_1^{-1}p_2 \in \mathcal{N}'''$ and $q_1^{-1}q_2 \in \mathcal{N}'''$ imply that $|f(q_1^{-1}p_1) - f(q_2^{-1}p_2)| < \varepsilon$. Hence, $|Df(p_1, q_1) - Df(p_2, q_2)| < \varepsilon$ whenever $p_1^{-1}p_2 \in \mathcal{N}'''$ and $q_1^{-1}q_2 \in \mathcal{N}'''$. This shows that $Df \in C_u(\mathcal{S} \times \mathcal{S})$.

4.3. The evolution family $T(\cdot, \cdot)$

Let $P_t^\sharp f(p) = \int_{\mathcal{S}} f(qp)m_t^\sharp(dq) = \int_{\mathcal{S}} f(q^{-1}p)m_t(dq)$. Then, since P_t commutes with P_t^\sharp , $P_t^\diamond := P_t P_t^\sharp$ defines a strongly continuous semigroup. We have

$$P_t^\diamond f(p) = \int_{\mathcal{S}} \int_{\mathcal{S}} f(r^{-1}pq)m_t(dq)m_t(dr). \tag{31}$$

Moreover, $A_{\diamond}f = Af + A^{\sharp}f$, $f \in \mathcal{D}(A) \cap \mathcal{D}(A^{\sharp})$, where A , A^{\sharp} , and A_{\diamond} are generators of $\{P_t, t \geq 0\}$, $\{P_t^{\sharp}, t \geq 0\}$, and $\{P_t^{\diamond}, t \geq 0\}$, respectively.

Proposition 8. *We have $S(t, s)D = DT(t, s)$, $t_0 \leq s \leq t < \infty$, where*

$$T(t, s)f = p_s(t)P_{t-s}^{\diamond}f + \Theta D \int_s^t g(t, v)P_{t-v}^{\diamond}f \, dv. \tag{32}$$

Moreover, $T(t, s)$, $t_0 \leq s \leq t < \infty$, is an evolution family in $C_0(\mathcal{S})$.

Proof. The following two properties of D are the key to the proof:

$$\begin{aligned} \Theta Df(p) &= f(e), & p \in \mathcal{S}, \\ U_t D &= DP_t^{\diamond}, & t \geq 0. \end{aligned} \tag{33}$$

The first of these is immediate, while the second follows by (31) and

$$U_t Df(p, q) = \frac{1}{2} \int_{\mathcal{S}} \int_{\mathcal{S}} (f((p')^{-1}p^{-1}qq') + f((q')^{-1}q^{-1}pp'))m_t(dp')m_t(dq').$$

Now, by (33), $P_t \Theta D = \Theta D$ and $K P_t \Theta Df(p, q) = f(e) = D \Theta Df(p, q)$ for $f \in C_u(\mathcal{S})$ and $t \geq 0$. This implies the first part of the proposition by (9). The rest is straightforward (use $P_t^{\diamond} \Theta D = \Theta D$, $t \geq 0$).

We see in particular that the distribution of $(X_1(t))^{-1}X_2(t)$ depends on the distribution of $(X_1(t_0), X_2(t_0))$ only via the distribution of $(X_1(t_0))^{-1}X_2(t_0)$. The operators $T(t, s)$ describe the evolution of the distribution of the relative differences in time.

Proposition 9. *Let A_{\diamond} be the generator of the semigroup $\{P_t^{\diamond}, t \geq 0\}$. Suppose that $2N(\cdot)$ is continuously differentiable. Then, for $f \in D(A_{\diamond})$ and $t > t_0$, $x(s) = T(t, s)f$ is the solution to the following (Kolmogorov backward) problem:*

$$\frac{dx(s)}{ds} = -A_{\diamond}x(s) - \frac{1}{2N(s)}\Theta Dx(s) + \frac{1}{2N(s)}x(s), \quad t_0 \leq s \leq t, \quad x(t) = f. \tag{34}$$

Proof. For $f \in \mathcal{D}(A_{\diamond})$, the first term in (32) belongs to $\mathcal{D}(A_{\diamond})$. The second term, say $T_1(t, s)f$, is a constant function and so belongs to $\mathcal{D}(A_{\diamond})$ as well, and $A_{\diamond}T_1(t, s)f = 0$. Moreover, direct differentiation gives

$$\begin{aligned} \frac{dT(t, s)f}{ds} &= \frac{1}{2N(s)}p_s(t)P_{t-s}^{\diamond}f - p_s(t)A_{\diamond}P_{t-s}^{\diamond}f - \frac{1}{2N(s)}\Theta Dp_s(t)P_{t-s}^{\diamond}f \\ &= -A_{\diamond}T(t, s)f + \frac{1}{2N(s)}p_s(t)(I - \Theta D)P_{t-s}^{\diamond}f. \end{aligned}$$

On the other hand, (33) gives $\Theta DT_1(t, s) = T_1(t, s)$. Hence, by (32), $(I - \Theta D)T(t, s) = p_s(t)(I - \Theta D)P_{t-s}^{\diamond}$. This shows that $y(s) = T(t, s)f$ solves (34). Uniqueness of solutions is proved as in Subsection 2.2 (immediately after Proposition 2).

By (33), (34) shows that genetic drift may be interpreted as a tendency to concentrate the distribution of relative differences around the neutral element. We illustrate this fact and Proposition 9 by the following examples.

Example 4. (Two related Brownian motions with constant drift.) Let $b > 0$ and $c \in \mathbb{R}$ be given constants. Also, let $\mathfrak{S} = \mathbb{R}$ and let

$$dm_t := \frac{1}{\sqrt{2\pi bt}} \exp\left(-\frac{(q - ct)^2}{2}\right) dq, \quad t > 0, m_0 = \delta_0.$$

In this case

$$P_t f(p) = \frac{1}{\sqrt{2\pi bt}} \int_{-\infty}^{\infty} \exp\left(-\frac{q^2}{2}\right) f(p + q + ct) dq, \quad t > 0, p \in \mathbb{R},$$

and, since \mathbb{R} is abelian, P_t^\sharp is given by the same formula with c replaced by $-c$. Thus,

$$A = \frac{b}{2} \frac{d^2}{dp^2} + c \frac{d}{dp} \quad \text{and} \quad A^\sharp = \frac{b}{2} \frac{d^2}{dp^2} - c \frac{d}{dp}$$

with common maximal domain in $C_u(\mathbb{R})$, and so $A_\diamond = A + A^\sharp = b d^2 / dp^2$ with the same domain. Therefore, (34) becomes

$$\frac{dx(s, p)}{ds} = -b \frac{d^2 x(s, p)}{dp^2} - \frac{1}{2N(s)} x(s, 0) + \frac{1}{2N(s)} x(s, p), \quad t_0 \leq s \leq t, p \in \mathbb{R},$$

with final condition $x(t, \cdot) = f, f \in \mathcal{D}(A_\diamond)$.

Example 5. (Two related telegraph processes.) Let a and v be given positive constants, and let $\{M(t), t \geq 0\}$ be a Poisson process with $E M(t) = at$. The process

$$p_t = \left(v \int_0^t (-1)^{M(s)} ds, (-1)^{M(t)} \right), \quad t \geq 0,$$

is referred to as the telegraph process since, as shown by Kac [28], who was inspired by Goldstein [23], the solutions to the telegraph equation may be expressed by means of its expected values [5], [19], [38]. Its state space is a noncommutative locally compact group $\mathfrak{S} = \mathbb{R} \times \{-1, 1\}$ with the topology induced from \mathbb{R}^2 and the multiplication rule $(\tau, k)(\xi, l) = (l\tau + \xi, kl)$. The left and right uniform structures on \mathfrak{S} are equivalent. The distributions m_t of $p_t, t \geq 0$, form a continuous convolution semigroup of probability measures on \mathfrak{S} [5], [32], and their explicit form is known [5], [27], [38], [43], but will not be used here. The space $C_u(\mathfrak{S})$ is isometrically isomorphic to the Cartesian product of two copies of $C_u(\mathbb{R})$ with the norm $\|(f_1, f_2)\| = \max_{i=1,2} \|f_i\|_{C_u(\mathbb{R})}$.

Let \mathcal{D} be the set composed of pairs $(f_1, f_2) \in C_u(\mathbb{R}) \times C_u(\mathbb{R})$ such that both f_i are differentiable and $f'_i \in C_u(\mathbb{R}), i = 1, 2$. We claim that \mathcal{D} is the domain of the generator A of $\{P_t, t \geq 0\}$, and $A(f_1, f_2) = v(f'_1, f'_2) + a(\tilde{f}_2, \tilde{f}_1) - a(f_1, f_2)$ for $(f_1, f_2) \in \mathcal{D}$, where $\tilde{f}(\tau) = f(-\tau)$. Since $(\tau, k)(vt, 1) = (\tau + vt, k)$, \mathcal{D} is the domain of the generator A^\blacktriangle of the semigroup $\{P_t^\blacktriangle, t \geq 0\}$ given by $P_t^\blacktriangle f(p) = f(p(vt, 1))$, and we have $A^\blacktriangle(f_1, f_2) = v(f'_1, f'_2)$ for $(f_1, f_2) \in \mathcal{D}$. Moreover, $e^{at} E \mathbf{1}_{\{M(t)=0\}} f(p_{\mathfrak{P}_t}) = P_t^\blacktriangle f(p)$. Hence,

$$t^{-1}(e^{at} P_t f(p) - P_t^\blacktriangle f(p)) = t^{-1} e^{at} E \mathbf{1}_{\{M(t)=1\}} f(p_{\mathfrak{P}_t}) + t^{-1} e^{at} E \mathbf{1}_{\{M(t) \geq 2\}} f(p_{\mathfrak{P}_t}) \quad (35)$$

for $f \in C_u(\mathfrak{S}), p \in \mathfrak{S}$, and $t > 0$. The second term on the right-hand side converges to 0, as $t \rightarrow 0+$, uniformly in $p \in \mathfrak{S}$. On the set where $M(t) = 1$, the distance between p_t and

$(0, -1)$ does not exceed vt . Hence, by continuity of f , the first term converges to $f(p(0, -1))$ uniformly in $p \in \mathcal{S}$. Therefore, (35) shows that f belongs to the domain of A if and only if it belongs to the domain of A^\blacktriangle and $Af(p) = A^\blacktriangle f(p) + af(p(0, -1)) - af(p)$ for $p \in \mathcal{S}$ and f in the domain, as claimed.

The same argument shows that \mathcal{D} is the domain of the generator A^\sharp of $\{P_t^\sharp, t \geq 0\}$ and $A^\sharp(f_1, f_2) = v(-f'_1, f'_2) + a(f_2, f_1) - a(f_1, f_2)$ for $(f_1, f_2) \in \mathcal{D}$.

Since $A_\diamond f = Af + A^\sharp f, f \in \mathcal{D}(A) \cap \mathcal{D}(A^\sharp), \mathcal{D}$ is contained in the domain $\mathcal{D}(A_\diamond)$ of A_\diamond and $A_\diamond(f_1, f_2) = (2a(f_2)_e - 2af_1, 2vf'_2 + 2a(f_1)_e - 2af_2)$, where $f_e(\tau) = \frac{1}{2}(f(\tau) + f(-\tau))$. The operator given by the right-hand side here, with domain \mathcal{D}' composed of pairs $(f_1, f_2) \in C_u(\mathbb{R}) \times C_u(\mathbb{R})$, where f_2 is differentiable with $f'_2 \in C_u(\mathbb{R})$, is a bounded perturbation of the generator of the Feller semigroup $\{P_t^\blacktriangleleft, t \geq 0\}$, where $P_t^\blacktriangleleft f(\tau, 1) = f(\tau, 1)$ and $P_t^\blacktriangleleft f(\tau, -1) = f(2vt + \tau, -1), \tau \in \mathbb{R}$. Hence, by Phillips' perturbation theorem [5], [17], [37], it generates a Feller semigroup. On the other hand, by the Dynkin–Reuter lemma (see, e.g. [41, p. 237]), it cannot be a proper extension of A_\diamond and we must have $\mathcal{D}(A_\diamond) = \mathcal{D}'$. Hence, (34) becomes

$$\begin{aligned} \frac{dx_1(s, p)}{ds} &= -a(x_2(s, p) + x_2(s, -p)) + 2ax_1(s, p) + \frac{1}{2N(s)}(x_1(p) - x_1(0)), \\ \frac{dx_2(s, p)}{ds} &= -2vx'_2(s, p) - a(x_1(s, p) + x_1(s, -p)) + 2ax_2(s, p) \\ &\quad + \frac{1}{2N(s)}(x_2(p) - x_1(0)), \end{aligned}$$

for $t_0 \leq s \leq t$, with $x(t, \cdot) = (f_1, f_2) \in \mathcal{D}'$.

Example 6. (Two related Lévy processes on the Klein group.) Let \mathcal{S} be the (commutative) Klein group with four elements p_1, \dots, p_4 (neutral element p_1 , and multiplication rules $p_2p_3 = p_4, p_2p_4 = p_3$, and $p_3p_4 = p_2$) endowed with the discrete topology. Any finite measure m on \mathcal{S} may be identified with a four-dimensional column vector with nonnegative coordinates, and any continuous convolution semigroup of probability measures on \mathcal{S} is of the form (see [5, pp. 291–292])

$$m_t = \frac{1}{4}G \begin{bmatrix} 1 \\ r_1^t \\ r_2^t \\ r_3^t \end{bmatrix}, \quad \text{where } G = \begin{bmatrix} +1 & +1 & +1 & +1 \\ +1 & -1 & -1 & +1 \\ +1 & -1 & +1 & -1 \\ +1 & +1 & -1 & -1 \end{bmatrix}, \tag{36}$$

and $r_2 = \exp(-\frac{1}{2}(\alpha_2 + \alpha_3)), r_3 = \exp(-\frac{1}{2}(\alpha_2 + \alpha_4)),$ and $r_4 = \exp(-\frac{1}{2}(\alpha_3 + \alpha_4))$ for some nonnegative constants $\alpha_2, \alpha_3,$ and α_4 . The left canonical representation $\mathcal{R}: m \mapsto \mathcal{R}_m$ of the convolution algebra of measures on \mathcal{S} maps a measure represented by numbers $a_i, i = 1, 2, 3, 4,$ into the matrix \mathcal{R}_m with rows $(a_1, a_2, a_3, a_4), (a_2, a_1, a_4, a_3), (a_3, a_4, a_1, a_2),$ and (a_4, a_3, a_2, a_1) so that $m * n$ is the matrix product of \mathcal{R}_m and n . Moreover, if members of $C_u(\mathcal{S}) = C_0(\mathcal{S}) = C(\mathcal{S})$ are identified with row vectors then $P_t f$ is the matrix product of f and \mathcal{R}_{m_t} . Also, A is bounded and may be represented by the matrix of the form of \mathcal{R}_m with a_i replaced by $\alpha_i, i = 1, \dots, 4,$ where $\alpha_1 = -\sum_{i=2}^4 \alpha_i$. Since \mathcal{S} is abelian and $p_i^{-1} = p_i, i = 1, \dots, 4, P_t^\sharp = P_t$ and so $P_t^\diamond = P_{2t}$. Therefore, $A_\diamond f = 2fA$ (the matrix

product) and (34) becomes

$$\begin{aligned} \frac{dx_1(s)}{ds} &= -2\alpha_1x_1(s) - 2\alpha_2x_2(s) - 2\alpha_3x_3(s) - 2\alpha_4x_4(s), \\ \frac{dx_2(s)}{ds} &= -2\alpha_2x_1(s) - 2\alpha_1x_2(s) - 2\alpha_4x_3(s) - 2\alpha_3x_4(s) + \frac{1}{2N(s)}(x_2(s) - x_1(s)), \\ \frac{dx_3(s)}{ds} &= -2\alpha_3x_1(s) - 2\alpha_4x_2(s) - 2\alpha_1x_3(s) - 2\alpha_2x_4(s) + \frac{1}{2N(s)}(x_3(s) - x_1(s)), \\ \frac{dx_4(s)}{ds} &= -2\alpha_4x_1(s) - 2\alpha_3x_2(s) - 2\alpha_2x_3(s) - 2\alpha_1x_4(s) + \frac{1}{2N(s)}(x_4(s) - x_1(s)), \end{aligned}$$

for $t_0 \leq s \leq t$, with $x_i(t) = f_i \in \mathbb{R}$, $i = 1, \dots, 4$.

4.4. Asymptotic behavior of relative differences

4.4.1. *Asymptotically constant population.* Proposition 8 reduces the problem of studying the asymptotic behavior of the distribution of relative differences to that of studying the asymptotic behavior of $T(t, s)$, $t_0 \leq s \leq t$, as $t \rightarrow \infty$. The main result of this subsection says that, as $t \rightarrow \infty$, regardless of the form of the initial distribution, the distribution of the difference of two Lévy processes related via family ties in a population of asymptotically constant size $2N_0$ converges weakly to the measure represented via the functional $f \mapsto \lambda(\lambda - A_\diamond)^{-1}f(e)$, where $\lambda = (2N_0)^{-1}$ and e is the neutral element in \mathcal{S} .

Theorem 2. *Suppose that the population size is asymptotically constant. Then*

$$\lim_{t \rightarrow \infty} T(t, s)f = \lambda \Theta D(\lambda - A_\diamond)^{-1}f = (\lambda(\lambda - A_\diamond)^{-1}f(e)) \mathbf{1}_\mathcal{S}, \quad f \in C_u(\mathcal{S}).$$

Proof. Rewriting (32), $T(t, s)f = p_s(t)P_{t-s}^\diamond f + \Theta D \int_0^\infty \mathbf{1}_{[0,t-s)}(v)g(t, t-v)P_v^\diamond f \, dv$. The first term here is bounded by $p_s(t)\|f\|$ and, hence, converges to 0 as $t \rightarrow \infty$. Since the integrand in the second term is bounded by $(\text{constant})e^{-(\text{constant})v}\|f\|$ and converges, for all v , to $\lambda e^{-\lambda v}P_v^\diamond f$, the claim follows by the dominated convergence theorem.

A similar theorem was obtained in [9] in the case where \mathcal{S} was the group of integers.

Example 7. *(Two related Brownian motions (continued).)* We come back to the situation of Example 4 and recall that [5], [40], [41], [47]

$$R_\lambda f(p) := (\lambda - A_\diamond)^{-1}f(p) = \frac{1}{2\sqrt{b\lambda}} \int_{-\infty}^\infty \exp(-\sqrt{\lambda b^{-1}}|q|)f(p+q) \, dq. \quad (37)$$

Hence,

$$\lambda(\lambda - A_\diamond)^{-1}f(0) = \frac{1}{2}\sqrt{\frac{\lambda}{b}} \int_{-\infty}^\infty \exp(-\sqrt{\lambda b^{-1}}|q|)f(q) \, dq.$$

This means that the distance between two Brownian motions with diffusion coefficient b , drawn from a Fisher–Wright population of approximately constant size $2N_0$, has bilateral exponential distribution (see [21, p. 49]) with parameter $\alpha = (2N_0b)^{-1/2}$.

Example 8. *(Two related telegraph processes (continued).)* In Example 5, $(g_1, g_2) = (\lambda - A_\diamond)^{-1}(f_1, f_2)$, where $\lambda := 1/2N_0$ solves

$$\begin{aligned} \lambda g_1 - 2a(g_2)_e + 2ag_1 &= f_1, \\ \lambda g_2 - 2vg'_2 - 2a(g_1)_e + 2ag_2 &= f_2. \end{aligned}$$

Substituting

$$g_1 = \frac{2a}{\lambda + 2a}(g_2)_e + \frac{1}{\lambda + 2a}f_1$$

into the second equation, we obtain

$$(\lambda + 2a)g_2 - 2vg'_2 - \frac{4a^2}{\lambda + 2a}(g_2)_e = \frac{2a}{\lambda + 2a}(f_1)_e + f_2.$$

By the main result of [8],

$$g_2 = \frac{1}{4v}(R_\mu f_2)' + \frac{a}{2v(\lambda + 2a)}(R_\mu(f_1)_e)' + \frac{\lambda + 2a}{8v^2}R_\mu f_2 + \frac{a}{4v^2}R_\mu(f_1)_e - \frac{4a^2}{8v^2(\lambda + 2a)}R_\mu(f_2)_o,$$

where $\mu = (\lambda^2 + 4a\lambda)/8v^2$, R_μ is given by (37) with $b = \frac{1}{2}$ and $h_o(\tau) := \frac{1}{2}(h(\tau) - h(-\tau))$ for $h \in C_u(\mathbb{R})$. Hence,

$$g_1 = \frac{1}{\lambda + 2a}f_1 + \frac{a}{2v(\lambda + 2a)}(R_\mu(f_2)_o)' + \frac{a}{4v^2}R_\mu(f_2)_e + \frac{a^2}{2v^2(\lambda + 2a)}R_\mu(f_1)_e.$$

Moreover,

$$(R_\mu h_o)'(0) = (R_\mu h)'(0) = \int_0^\infty e^{-\sqrt{2\mu}h(\tau)} d\tau - \int_{-\infty}^0 e^{\sqrt{2\mu}h(\tau)} d\tau$$

and

$$R_\mu h_e(0) = R_\mu h(0) = \frac{1}{\sqrt{2\mu}} \int_{-\infty}^\infty e^{-\sqrt{2\mu}|\tau|} h(\tau) d\tau, \quad h \in C_u(\mathbb{R}).$$

Hence, $\lambda g_1(0) = \lambda(\lambda - A_\diamond)^{-1}(f_1, f_2)((0, 1))$ equals

$$\begin{aligned} &\rho_1 f_1(0) + \rho_2 \sqrt{\frac{\mu}{2}} \int_{-\infty}^\infty e^{-\sqrt{2\mu}|\tau|} f_1(\tau) d\tau \\ &+ \rho_3 \sqrt{2\mu} \int_0^\infty e^{-\sqrt{2\mu}\tau} f_2(\tau) d\tau + \rho_4 \sqrt{2\mu} \int_{-\infty}^0 e^{\sqrt{2\mu}\tau} f_2(\tau) d\tau, \end{aligned} \tag{38}$$

where $\rho_1 = \lambda/(\lambda + 2a)$, $\rho_2 = a^2\lambda/2v^2\mu(\lambda + 2a)$, $\rho_3 = a\lambda/8v^2\mu + a\lambda/2v\sqrt{2\mu}(\lambda + 2a)$, and $\rho_4 = a\lambda/8v^2\mu - a\lambda/2v\sqrt{2\mu}(\lambda + 2a)$ are positive with $\sum_{i=1}^4 \rho_i = 1$.

This means that the limit distribution of the difference of two telegraph processes related via family ties in a population of asymptotically constant size is a convex combination of two probability measures on $\mathbb{R} \times \{1\}$ and two probability measures on $\mathbb{R} \times \{-1\}$. On $\mathbb{R} \times \{1\}$, these are the point mass at $(0, 1)$ and the bilateral exponential distribution with parameter $\alpha = \sqrt{2\mu}$, and on $\mathbb{R} \times \{-1\}$, these are the exponential distribution with parameter $\alpha = \sqrt{2\mu}$ and its mirrored distribution (see [21, p. 49]).

Remark 4. For $v = \sqrt{a}$, letting $a \rightarrow \infty$ in (38), we obtain $(\lambda/4)R_{\lambda/2}f_1(0) + (\lambda/4)R_{\lambda/2}f_2(0)$; if $f_1 = f_2 = f$, this is the quantity obtained in Example 7 (with $b = \frac{1}{2}$). This is a reflection of the fact that, as $a \rightarrow \infty$, $\sqrt{a} \int_0^t (-1)^{M(s)} ds$ converges in distribution to a standard Brownian motion [5], [19], [38].

Example 9. (Two related Lévy processes on the Klein group (continued).) The neutral element of the Klein group is p_1 , and $\lambda(\lambda - A_\diamond)^{-1} f(p_1)$ is the scalar product of the row vector f and the first column of the matrix $\lambda(\lambda - 2A)^{-1} = (\lambda/2)(\lambda/2 - A)^{-1}$. In other words, the distribution of the difference between two exchangeable Lévy processes with values in the Klein group drawn from the Fisher–Wright population of asymptotically constant size $2N_0$ is the first column of $(\lambda/2)(\lambda/2 - A)^{-1}$. On the other hand, by (36), $(\lambda/2)(\lambda/2 - A)^{-1}$ equals

$$\frac{\lambda}{2} \int_0^\infty e^{-(\lambda/2)t} \mathcal{R}_{m_t} dt = \mathcal{R}_{(\lambda/2)} \int_0^\infty \exp(-(\lambda/2)t) m_t dt = \mathcal{R}_{(\lambda/8)} G \bar{v}^\top,$$

where $\bar{v} = \int_0^\infty e^{-(\lambda/2)t} (1, r_2^t, r_3^t, r_4^t) dt = (2/\lambda, 2/(\lambda - 2 \ln r_2), 2/(\lambda - \ln r_3), 2/(\lambda - \ln r_4))$. Since, for any m , the first column of \mathcal{R}_m is simply m , the searched-for distribution is $(\lambda/8)G \bar{v}^\top$.

4.4.2. *Expanding population.* For most of the theory developed in this paper, it has been convenient to focus on spaces of real functions. However, in turning to the Fourier transform, a step we are to take in this subsection, we need to work with complex, uniformly continuous functions on \mathbb{R} . Fortunately, all the results from the previous subsections translate to the case of complex functions, with possible cosmetic changes. Hence, we may and will consider $C_u(\mathbb{R})$ as the space of complex, bounded uniformly continuous functions on \mathbb{R} .

It is the genetic drift that is the reason for the existence of a stationary distribution of relative differences found in Theorem 2, and in the absence of this force, there is no hope for a similar result. In rapidly growing populations, the joint distributions of attributes of pairs of individuals behave essentially as if they were independent. However, if the population size does not grow to infinity sufficiently fast, there may remain subtle dependencies between them. We illustrate this by considering the case where $\mathcal{X} = \mathbb{R}$ and mutation is modeled by a real Lévy process of bounded variance $\sigma^2(t) = bt$ and expected value ct , where $b > 0$ and $c \in \mathbb{R}$ are given constants [21, Chapter IX.4]. In terms of the semigroup $\{P_t, t \geq 0\}$, this means that (compare [21, Chapter XVII])

$$P_t e_\tau = e_\tau e^{t\psi(\tau)}, \tag{39}$$

where $e_\tau(p) = e^{i\tau p}$,

$$\psi(\tau) = b \int_{-\infty}^\infty \frac{e^{i\tau p} - 1 - i\tau p}{p^2} M(dp) + ic\tau,$$

and M is a probability measure. (This is the Lévy–Khintchine formula in the case of finite variances.)

Theorem 3. *Suppose that (39) holds and that the population size grows to infinity rapidly, so that $\int_{t_0}^\infty dt/2N(t) =: \kappa < \infty$. For $t \geq t_0$, let $X(t) = (X_1(t), X_2(t))$ be a random vector such that $S(t, t_0) f(p, q) = E f(p + X_1(t), q + X_2(t))$, $f \in C_u(\mathbb{R}^2)$. Then, as $t \rightarrow \infty$, the vectors $Y(t) = (Y_1(t), Y_2(t)) = (1/\sqrt{t'}) (X_1(t) - ct', X_2(t) - ct')$, where $t' = t - t_0$, tend in distribution to a normal vector with characteristic function $\exp(-b(\tau^2 + \sigma^2)/2)$.*

Proof. Introducing $e_{\tau,\sigma}(p, q) = e_\tau(p)e_\sigma(q)$ for $\tau, \sigma, p, q \in \mathbb{R}$ we obtain $U_t e_{\tau,\sigma} = e^{t(\psi(\tau)+\psi(\sigma))} e_{\tau,\sigma}$, $t \geq 0$. Since $\Theta e_{\tau,\sigma} = e_{\tau+\sigma}$, a straightforward calculation shows that

$$K P_{v-t_0} \Theta U_{t-v} e_{\tau,\sigma} = e^{(v-t_0)\psi(\tau+\sigma)} e^{(t-v)(\psi(\tau)+\psi(\sigma))} K e_{\tau+\sigma}, \quad v \in [t_0, t].$$

Therefore, by (9), $E \exp(i\tau X_1(t)) \exp(i\sigma X_2(t)) = S(t, t_0) e_{\tau, \sigma}(0, 0)$ equals

$$p_{t_0}(t) e^{t'(\psi(\tau) + \psi(\sigma))} + \int_{t_0}^t g(t, v) e^{(v-t_0)\psi(\tau+\sigma)} e^{(t-v)(\psi(\tau) + \psi(\sigma))} dv,$$

and so

$$\begin{aligned} E \exp(i\tau Y_1(t)) \exp(i\sigma Y_2(t)) &= p_{t_0}(t) \exp\left(t' \left(\psi_0\left(\frac{\tau}{\sqrt{t'}}\right) + \psi_0\left(\frac{\sigma}{\sqrt{t'}}\right) \right)\right) \\ &+ \exp\left(\int_t^\infty \frac{ds}{2N(s)}\right) \int_{t_0}^t \frac{1}{2N(v)} \exp\left(-\int_v^\infty \frac{ds}{2N(s)}\right) \\ &\times \exp\left((v-t_0)\psi_0\left(\frac{\tau}{\sqrt{t'}} + \frac{\sigma}{\sqrt{t'}}\right)\right) \\ &\times \exp\left((t-v)\left(\psi_0\left(\frac{\tau}{\sqrt{t'}}\right) + \psi_0\left(\frac{\sigma}{\sqrt{t'}}\right)\right)\right) dv, \end{aligned}$$

where $\psi_0(\tau) = \psi(\tau) - ic\tau$. We have $\lim_{t \rightarrow \infty} t\psi_0(\tau/\sqrt{t}) = \lim_{h \rightarrow 0+} (\psi_0(h\tau) - \psi_0(0))/h^2 = (\tau^2/2)\psi_0''(0) = -(\tau^2/2)b$. Hence, all the moduli of the exponents involving ψ_0 being bounded by 1, by the dominated convergence theorem, we have

$$\begin{aligned} \lim_{t \rightarrow \infty} E \exp(i\tau Y_1(t)) \exp(i\sigma Y_2(t)) &= e^{-\kappa} \exp\left(-b \frac{\tau^2 + \sigma^2}{2}\right) + \int_{t_0}^\infty \frac{1}{2N(v)} \exp\left(-\int_v^\infty \frac{ds}{2N(s)}\right) dv \exp\left(-b \frac{\tau^2 + \sigma^2}{2}\right) \\ &= \exp\left(-b \frac{\tau^2 + \sigma^2}{2}\right). \end{aligned}$$

This completes the proof.

Example 10. This example shows that the assumption of fast growth is crucial for Theorem 3 to hold. We consider simple linear growth: $2N(t) = t' + a$, where $a > 0$. Then $g(t, v) = 1/(t' + a)$, $p_{t_0}(t) = a/(a + t')$, and

$$\begin{aligned} E \exp(i\tau Y_1(t)) \exp(i\sigma Y_2(t)) &= \frac{a}{t' + a} \exp\left(t' \left(\psi_0\left(\frac{\tau}{\sqrt{t'}}\right) + \psi_0\left(\frac{\sigma}{\sqrt{t'}}\right) \right)\right) \\ &+ \frac{1}{t' + a} \int_{t_0}^t \exp\left((v-t_0)\psi_0\left(\frac{\tau}{\sqrt{t'}} + \frac{\sigma}{\sqrt{t'}}\right)\right) \\ &\times \exp\left((t-v)\left(\psi_0\left(\frac{\tau}{\sqrt{t'}}\right) + \psi_0\left(\frac{\sigma}{\sqrt{t'}}\right)\right)\right) dv \\ &= \frac{a}{t' + a} \exp\left(t' \left(\psi_0\left(\frac{\tau}{\sqrt{t'}}\right) + \psi_0\left(\frac{\sigma}{\sqrt{t'}}\right) \right)\right) \\ &+ \frac{1}{t' + a} \\ &\times \frac{\exp((t-t_0)\psi_0(\tau/\sqrt{t'} + \sigma/\sqrt{t'})) - \exp((t-t_0)(\psi_0(\tau/\sqrt{t'}) + \psi_0(\sigma/\sqrt{t'})))}{\psi_0(\tau/\sqrt{t'} + \sigma/\sqrt{t'}) - \psi_0(\tau/\sqrt{t'}) - \psi_0(\sigma/\sqrt{t'})} dv \\ &\rightarrow \exp\left(-b \frac{\tau^2 + \sigma^2}{2}\right) \frac{1 - e^{-\tau\sigma b}}{\tau\sigma b} \quad \text{as } t \rightarrow \infty. \end{aligned}$$

Clearly, the limit vector has normal, but correlated coordinates.

Appendix A

Proposition 10. *The dual to $C_{0s}(\mathcal{S} \times \mathcal{S})$ is (isometrically isomorphic to) the space $\mathcal{M}_s(\mathcal{S} \times \mathcal{S})$ of finite regular symmetric Borel measures on $\mathcal{S} \times \mathcal{S}$.*

Proof. A symmetric Borel measure μ induces a bounded functional $f \mapsto \int_{\mathcal{S} \times \mathcal{S}} f \, d\mu$ on $C_{0s}(\mathcal{S} \times \mathcal{S})$. On the other hand, by the Hahn–Banach and the Riesz representation theorems, a bounded functional F on $C_{0s}(\mathcal{S} \times \mathcal{S})$ may be extended to the functional $f \mapsto \int_{\mathcal{S} \times \mathcal{S}} f \, d\nu$ on $C_0(\mathcal{S} \times \mathcal{S})$ with the same norm, where ν is a finite regular Borel measure on $\mathcal{S} \times \mathcal{S}$. Let ν_Z be the transport of ν via Z . We have

$$\int_{\mathcal{S} \times \mathcal{S}} f \, d\nu_Z = \int_{\mathcal{S} \times \mathcal{S}} f \, d\nu \quad \text{for } f \in C_{0s}(\mathcal{S} \times \mathcal{S}).$$

Hence, $Ff = \int_{\mathcal{S} \times \mathcal{S}} f \, d\mu$, where $\mu = \frac{1}{2}(\nu + \nu_Z)$ is a regular, finite, symmetric Borel measure on $\mathcal{S} \times \mathcal{S}$. Moreover, an $f \in C_0(\mathcal{S} \times \mathcal{S})$ may be written as $f = f_s + f_{as}$, where $f_s = \frac{1}{2}(f + f \circ Z)$ is symmetric and $f_{as} = \frac{1}{2}(f - f \circ Z)$ is asymmetric, so that $\int f \, d\mu = \int f_s \, d\mu$ while $\|f_s\| \leq \|f\|$. Hence,

$$\|F\|_{(C_{0s}(\mathcal{S} \times \mathcal{S}))^*} = \sup_{f \in C_0(\mathcal{S} \times \mathcal{S}), \|f\| \leq 1} \left| \int_{\mathcal{S} \times \mathcal{S}} f \, d\mu \right| = \|\mu\|_{\mathcal{M}_s(\mathcal{S} \times \mathcal{S})}.$$

Let \mathcal{R} be the equivalence relation in $\mathcal{S} \times \mathcal{S}$ given by

$$(p, q)\mathcal{R}(p', q') \iff Z(p, q) = Z(p', q').$$

The quotient space $\mathcal{S}_\Delta = (\mathcal{S} \times \mathcal{S})/\mathcal{R}$ is equipped with the natural topology, termed quotient topology [11], defined as the family of all sets B in \mathcal{S}_Δ such that $M^{-1}B$ is an open set in $\mathcal{S} \times \mathcal{S}$, where M is the canonical map assigning the class $[(p, q)] \in \mathcal{S}_\Delta$ to a point $(p, q) \in \mathcal{S} \times \mathcal{S}$. This topology is the strongest in which M is continuous.

Lemma 4. *The map M is both open and closed. Moreover, for a compact set $\mathcal{K} \subset \mathcal{S}_\Delta$, the counterimage $M^{-1}\mathcal{K}$ of \mathcal{K} is compact.*

Proof. Suppose that \mathcal{O} is an open set in $\mathcal{S} \times \mathcal{S}$. Then $M^{-1}M\mathcal{O} = \mathcal{O} \cup Z\mathcal{O}$ is open, and so is $M\mathcal{O}$. Similarly, suppose that \mathcal{C} is closed in $\mathcal{S} \times \mathcal{S}$. Then $M^{-1}((M\mathcal{C})^c) = (\mathcal{C} \cup Z\mathcal{C})^c$ is open, $\mathcal{C} \cup Z\mathcal{C}$ being closed. Thus, $(M\mathcal{C})^c$ is open in \mathcal{S}_Δ , whence $M\mathcal{C}$ is closed.

Let a nonempty $\mathcal{K} \subset \mathcal{S}_\Delta$ be compact. A subset \mathcal{A} of $\mathcal{S} \times \mathcal{S}$ is said to be antisymmetric if

$$\mathcal{A} \cap Z\mathcal{A} \subset \{(p, p) \in \mathcal{S} \times \mathcal{S}; p \in \mathcal{S}\}.$$

The family of antisymmetric subsets of $M^{-1}\mathcal{K}$ is partially ordered by the relation of inclusion, and nonempty, as it contains at least one singleton. Also, any linearly ordered subfamily of this family has an upper bound, namely the union of all elements of this subfamily. Hence, by the Kuratowski–Zorn lemma, there exists a maximal antisymmetric subset, say \mathcal{A}_m , of $M^{-1}\mathcal{K}$.

The map M establishes a one-to-one correspondence between elements of \mathcal{A}_m and elements of \mathcal{K} . Indeed, M restricted to this set is injective, \mathcal{A}_m being antisymmetric, and ‘onto’, since \mathcal{A}_m is maximal. Hence, a straightforward reasoning based on the fact that M is open shows that \mathcal{A}_m is compact.

On the other hand, $Z\mathcal{A}_m \subset M^{-1}\mathcal{K}$ is also antisymmetric and maximal, and so, by the same reasoning, it is compact. Finally, $M^{-1}\mathcal{K}$ is compact as a union of two compact sets, namely \mathcal{A}_m and $Z\mathcal{A}_m$.

Clearly, \mathcal{S}_Δ is a Hausdorff space. Since, by Lemma 4, it is also the image of the locally compact space $\mathcal{S} \times \mathcal{S}$ via an open map, \mathcal{S}_Δ is locally compact [18].

Proposition 11. *The spaces $C_0(\mathcal{S}_\Delta)$ and $C_{0s}(\mathcal{S} \times \mathcal{S})$ are isometrically isomorphic.*

Proof. For a continuous symmetric function f on $\mathcal{S} \times \mathcal{S}$, define Jf on \mathcal{S}_Δ by $Jf([(p, q)]) = f(p, q)$, where $[\cdot]$ is an equivalence class of \mathcal{R} . By the definition of the topology in \mathcal{S}_Δ , Jf is continuous, and

$$\sup_{\xi \in \mathcal{S}_\Delta} |Jf(\xi)| = \sup_{(p,q) \in \mathcal{S} \times \mathcal{S}} |f(p, q)| \quad (\text{finite or infinite}).$$

Also,

$$\{\xi \in \mathcal{S}_\Delta \mid Jf(\xi) \neq 0\} = M\mathcal{U}, \quad \text{where } \mathcal{U} = \{(p, q) \in \mathcal{S} \times \mathcal{S} \mid f(p, q) \neq 0\}.$$

Hence, M being continuous, the support

$$\text{supp}(Jf) = \overline{M\mathcal{U}} \supset M\overline{\mathcal{U}} = M \text{supp}(f).$$

On the other hand, M being closed, $M \text{supp}(f)$ is closed, and so, since $M \text{supp}(f)$ contains $M\mathcal{U}$,

$$\text{supp}(Jf) = \overline{M\mathcal{U}} \subset M \text{supp}(f).$$

Hence,

$$\text{supp}(Jf) = M \text{supp}(f). \quad (40)$$

If f has compact support then, by (40), $\text{supp}(Jf)$ is compact as the continuous image of the compact set $\text{supp}(f)$. This implies that J maps $C_{0s}(\mathcal{S} \times \mathcal{S})$ into $C_0(\mathcal{S}_\Delta)$ in an isometric manner. On the other hand, if g on \mathcal{S}_Δ is continuous with compact support then $f := g \circ M$ is symmetric (for $M \circ Z = M$) and, by Lemma 4 and (40), has compact support; clearly, $g = Jf$. Thus, J being an isometry and having dense image is onto.

Acknowledgements

This work was supported by the Polish Government research fund (2005–2008) under grant 1 P03A 044 29 (0356/P03/2005/29), and by TODEQ. I would like to thank Y. Tomilov for the reference to Szucs' paper [44].

References

- [1] ARENDT, W., BATTY, C. J. K., HIEBER, M. AND NEUBRANDER, F. (2001) *Vector-Valued Laplace Transforms and Cauchy Problems*. Birkhäuser, Basel.
- [2] ARENDT, W. *et al.* (1986). *One-Parameter Semigroups of Positive Operators* (Lecture Notes Math. **1184**). Springer, Berlin.
- [3] BEAUMONT, M. A. (1999). Detecting population expansion and decline using microsatellites. *Genetics* **153**, 2013–2029.
- [4] BOBROWSKI, A. (2004). Quasi-stationary distributions of a pair of Markov chains related to time evolution of a DNA locus. *Adv. Appl. Prob.* **36**, 57–77.
- [5] BOBROWSKI, A. (2005). *Functional Analysis for Probability and Stochastic Processes*. Cambridge University Press.
- [6] BOBROWSKI, A., AND KIMMEL, M. (2003). A random evolution related to a Fisher–Wright–Moran model with mutation, recombination and drift. *Math. Methods Appl. Sci.* **26**, 1587–1599.
- [7] BOBROWSKI, A. AND KIMMEL, M. (2004). Asymptotic behavior of joint distributions of characteristics of a pair of randomly chosen individuals in discrete-time Fisher–Wright models with mutations and drift. *Theoret. Pop. Biol.* **66**, 355–367.

- [8] BOBROWSKI, A. AND KUBALIŃSKA, M. (2006). On a functional equation with derivative and symmetrization. *Ann. Polon. Math.* **89**, 13–24.
- [9] BOBROWSKI, A., KIMMEL, M., ARINO, O. AND CHAKRABORTY, R. (2001). A semigroup representation and asymptotic behavior of certain statistics of the Fisher–Wright–Moran coalescent. In *Stochastic Processes: Theory and Methods* (Handbook Statist. **19**), eds C. R. Rao and D. N. Shanbhag. Elsevier, Amsterdam, pp. 215–247.
- [10] BOBROWSKI, A., WANG, N., KIMMEL, M. AND CHAKRABORTY, R. (2002). Nonhomogeneous infinite sites model under demographic change: mathematical description and asymptotic behavior of pairwise distributions. *Math. Biosci.* **175**, 83–115.
- [11] BOURBAKI, N. (1940). *Topologie Générale, ch. I et II*. Hermann, Paris.
- [12] DEFANT, A. AND FLORET, K. (1993). *Tensor Norms and Operator Ideals*. North-Holland, Amsterdam.
- [13] DONELLY, P. AND KURTZ, T. G. (1996). A countable representation of the Fleming–Viot measure-valued diffusions. *Ann. Prob.* **24**, 743–760.
- [14] DONELLY, P. AND KURTZ, T. G. (1999). Genealogical processes for Fleming–Viot models with selection and recombination. *Ann. Appl. Prob.* **9**, 1091–1148.
- [15] DONELLY, P. AND KURTZ, T. G. (1999). Particle representations for measure-valued population models. *Ann. Prob.* **27**, 166–205.
- [16] EMLION, R. (1985). Mean-bounded operators and mean ergodic theorems. *J. Funct. Anal.* **61**, 1–14.
- [17] ENGEL, K.-J. AND NAGEL, R. (2000). *One-Parameter Semigroups for Linear Evolution Equations*. Springer, New York.
- [18] ENGELKING, R. (1977). *General Topology*. Polish Scientific Press, New York.
- [19] ETHIER, S. N. AND KURTZ, T. G. (1986) *Markov Processes. Characterization and Convergence*. John Wiley, New York.
- [20] EWENS, W. (2004). *Mathematical Population Genetics. I. Theoretical Introduction*, 2nd edn. Springer, New York.
- [21] FELLER, W. (1971). *An Introduction to Probability Theory and Its Applications*, Vol. 2, 2nd edn. John Wiley, New York.
- [22] GOLDSTEIN, J. A. (1985). *Semigroups of Linear Operators and Applications*. Oxford University Press.
- [23] GOLDSTEIN, S. (1951). On diffusion of discontinuous movements, and on the telegraph equation. *Quart. J. Mech. Appl. Math.* **4**, 129–156.
- [24] GRIFFITHS, R. C. AND TAVARE, S. (1994). Sampling theory for neutral alleles in a varying environment. *Philos. Trans. R. Soc. London B* **344**, 403–410.
- [25] HEWITT, E. AND ROSS, K. A. (1979). *Abstract Harmonic Analysis*, Vol. 1, 2nd edn. Springer, Berlin.
- [26] HEYER, H. (1977). *Probability Measures on Locally Compact Groups*. Springer, Berlin.
- [27] JANSEN, A. AND SIEBERT, E. (1981). Convolution semigroups and generalized telegraph equations. *Mat. Z.* **177**, 519–532.
- [28] KAC, M. (1956). *Some Stochastic Problems in Physics and Mathematics*. Socony Mobil Oil Company.
- [29] KATO, T. (1953). Integration of the equation of evolution in a Banach space. *J. Math. Soc. Japan* **5**, 208–234.
- [30] KINGMAN, J. F. C. (1976). Coherent random walks arising in some genetic models. *Proc. R. Soc. London A* **351**, 19–31.
- [31] KINGMAN, J. F. C. (1982). The coalescent. *Stoch. Process. Appl.* **13**, 235–248.
- [32] KISYŃSKI, J. (1974). On M. Kac’s probabilistic formula for the solution of the telegraphist’s equation. *Ann. Pol. Math.* **29**, 259–272.
- [33] KREĀN, S. G. (1971). *Linear Differential Equations in Banach Spaces* (Transl. Math. Monog. **29**). American Mathematical Society, Providence, RI.
- [34] KRONE, S. M. AND NEUHAUSER, C. (1997) Ancestral process with selection. *Theoret. Pop. Biol.* **51**, 210–237.
- [35] LASOTA, A. AND RUDNICKI, R. (1988). Asymptotic behavior of semigroups of positive operators on $C(X)$. *Bull. Polish Acad. Sci. Math.* **36**, 151–159.
- [36] MORAN, P. A. P. (1975). Wandering distributions and the electrophoretic profile. *Theoret. Pop. Biol.* **9**, 318–330.
- [37] PAZY, A. (1983). *Semigroups of Linear Operators and Applications to Partial Differential Equations* (Appl. Math. Sci. **44**). Springer, New York.
- [38] PINSKY, M. A. (1991). *Lectures on Random Evolution*. World Scientific, River Edge, NJ.
- [39] POLAŃSKA, J. AND KIMMEL, M. (1999). A model of dynamics of mutation, genetic drift and recombination in DNA-repeat genetic loci. *Archives Control Sci.* **9**, 143–157.
- [40] REVUZ, D. AND YOR, M. (1999). *Continuous Martingales and Brownian Motion*, 3rd edn. Springer, Berlin.
- [41] ROGERS, L. C. G. AND WILLIAMS, D. (2000). *Diffusions, Markov Processes and Martingales*, Vol. 1, *Foundations*. Cambridge University Press.
- [42] SLADE, P. F. (2001). Simulation of ‘hitch-hiking’ genealogies. *J. Math. Biol.* **42**, 41–70.
- [43] STEUTEL, F. W. (1984). Poisson processes and a Bessel function integral. *SIAM Rev.* **27**, 73–77.

- [44] SZUCS, J. M. (1985). Ergodic theorems for tensor products. *J. Funct. Anal.* **64**, 125–133.
- [45] TAJIMA, F. (1983). Evolutionary relationship of DNA sequences in finite populations. *Genetics* **105**, 437–460.
- [46] TAVARÉ, S. AND ZEITOUNI, O. (2004). *Lectures on Probability Theory and Statistics* (Lecture Notes Math. **1837**). Springer, Berlin.
- [47] WENTZEL, A. D. (1981). *A Course in the Theory of Stochastic Processes*. McGraw-Hill, New York.