The Annals of Probability
1993, Vol. 21, No. 3, 1571–1590

THE TRANSITION FUNCTION OF A FLEMING–VIOT PROCESS

BY S. N. ETHIER[1] AND R. C. GRIFFITHS

University of Utah and Monash University

Let S be a compact metric space, let $\theta \geq 0$, and let ν_0 be a Borel probability measure on S. An explicit formula is found for the transition function of the Fleming–Viot process with type space S and mutation operator $(Af)(x) = (1/2)\theta\int_S (f(\xi) - f(x))\nu_0(d\xi)$.

1. Introduction and statement of results. The familiar K-type Wright–Fisher diffusion process in population genetics assumes values in the simplex

$$\Delta_K = \{p = (p_1, \dots, p_K)\colon p_1 \geq 0, \dots, p_K \geq 0,\ p_1 + \cdots + p_K = 1\} \tag{1.1}$$

and is characterized in terms of the generator

$$L = \frac{1}{2}\sum_{i,j=1}^{K} p_i(\delta_{ij} - p_j)\frac{\partial^2}{\partial p_i\,\partial p_j} + \sum_{j=1}^{K}\left(\sum_{i=1}^{K} q_{ij}p_i\right)\frac{\partial}{\partial p_j}, \tag{1.2}$$

where the infinitesimal matrix (q_{ij}) describes the mutation structure. Here $\mathscr{D}(L) = \{F|_{\Delta_K}\colon F \in C^2(\mathbf{R}^K)\}$. Fleming and Viot (1979) generalized this process, replacing $\{1, \dots, K\}$ by a compact metric space S, Δ_K by $\mathscr{P}(S)$, the set of Borel probability measures on S with the topology of weak convergence, and L by

$$\begin{aligned}(\mathscr{L}\varphi)(\mu) &= \frac{1}{2}\int_S\int_S \mu(dx)(\delta_x(dy) - \mu(dy))\frac{\delta^2\varphi(\mu)}{\delta\mu(x)\,\delta\mu(y)}\\ &\quad + \int_S \mu(dx) A\left(\frac{\delta\varphi(\mu)}{\delta\mu(\cdot)}\right)(x),\end{aligned} \tag{1.3}$$

where $\delta\varphi(\mu)/\delta\mu(x) = \lim_{\varepsilon\to 0+}\varepsilon^{-1}\{\varphi(\mu + \varepsilon\delta_x) - \varphi(\mu)\}$ and A is the generator of a Feller semigroup on $C(S)$. Here $\mathscr{D}(\mathscr{L}) = \{\varphi\colon \varphi(\mu) \equiv F(\langle f_1, \mu\rangle, \dots, \langle f_k, \mu\rangle),\ F \in C^2(\mathbf{R}^k),\ f_1, \dots, f_k \in \mathscr{D}(A),\ k \geq 1\}$ and $\langle f, \mu\rangle = \int_S f\,d\mu$. We refer to S as the type space and A as the mutation operator. See Ethier and Kurtz (1993a) for a survey article on Fleming–Viot processes.

It was discovered by Wright (1949) that when

$$q_{ij} = \tfrac{1}{2}\theta_j > 0, \qquad i, j \in \{1, \dots, K\},\ i \neq j, \tag{1.4}$$

Received June 1991; revised July 1992.

[1] Research partially supported by NSF Grant DMS-89-02991.

AMS 1991 *subject classifications*. Primary 60G57, 60J35; secondary 60J60, 92D15.

Key words and phrases. Infinite-dimensional diffusion process, measure-valued diffusion, Poisson–Dirichlet distribution, infinitely-many-neutral-alleles diffusion model, population genetics.

the Wright–Fisher diffusion has a unique stationary distribution $\pi \in \mathscr{P}(\Delta_K)$, given by

$$(1.5) \qquad \pi(dp) = \frac{\Gamma(\theta_1 + \cdots + \theta_K)}{\Gamma(\theta_1) \cdots \Gamma(\theta_K)} p_1^{\theta_1 - 1} \cdots p_K^{\theta_K - 1}\, dp_1 \cdots dp_{K-1}.$$

This is the Dirichlet distribution with parameters $\theta_1, \dots, \theta_K$. Shiga (1990) established the analogous result for the Fleming–Viot process with

$$(1.6) \qquad (Af)(x) \equiv \tfrac{1}{2}\theta \int_S (f(\xi) - f(x)) \nu_0(d\xi),$$

where $\theta > 0$ and $\nu_0 \in \mathscr{P}(S)$: There is a unique stationary distribution $\Pi_{\theta, \nu_0} \in \mathscr{P}(\mathscr{P}(S))$, which is the distribution of the $\mathscr{P}(S)$-valued random variable ν characterized by the property that whenever $K \geq 2$ and $\Lambda_1, \dots, \Lambda_K$ is a partition of S into Borel sets, $(\nu(\Lambda_1), \dots, \nu(\Lambda_K))$ has the Dirichlet distribution with parameters $\theta\nu_0(\Lambda_1), \dots, \theta\nu_0(\Lambda_K)$. [It is easy to generalize (1.5), allowing some of the parameters to be 0; see (1.26).] Ethier and Kurtz (1986, 1993b) showed that

$$(1.7) \qquad \Pi_{\theta, \nu_0}(\cdot) = \mathbf{P}\left\{ \sum_{i=1}^{\infty} \rho_i \delta_{\xi_i} \in \cdot \right\},$$

where $(\rho_1, \rho_2, \dots)$ has the Poisson–Dirichlet distribution with parameter θ [Kingman (1975)], and $\xi_1, \xi_2, \dots$ are i.i.d. ν_0, independent of $(\rho_1, \rho_2, \dots)$. Of course, $(\rho_1, \rho_2, \dots)$ assumes values in

$$(1.8) \qquad \nabla_\infty = \left\{ p = (p_1, p_2, \dots) \colon p_1 \geq p_2 \geq \cdots \geq 0,\ \sum_{i=1}^{\infty} p_i = 1 \right\};$$

in particular, Π_{θ, ν_0} is concentrated on $\mathscr{P}_{\mathrm{a}}(S)$, the set of purely atomic Borel probability measures on S.

Shimakura (1977, 1981) and Griffiths (1979) derived an explicit formula for the transition density of the Wright–Fisher diffusion assuming (1.4). This had previously been done in the one-dimensional case ($K = 2$) by Malécot (1948), Goldberg (1950), and Crow and Kimura (1956). Our aim here is to obtain the analogous result for the Fleming–Viot process assuming (1.6), namely, an explicit formula for the transition function of the process; a transition density does not exist in general.

To state the main result, we need some additional notation. For each $n \geq 1$ define $\eta_n \colon S^n \mapsto \mathscr{P}(S)$ by letting $\eta_n(x_1, \dots, x_n)$ be the empirical measure determined by the (not necessarily distinct) points $x_1, \dots, x_n \in S$:

$$(1.9) \qquad \eta_n(x_1, \dots, x_n) = n^{-1}(\delta_{x_1} + \cdots + \delta_{x_n}).$$

Given $\theta \geq 0$, let $\{D_t,\ t \geq 0\}$ be the pure death process in $\mathbf{Z}_+ \cup \{\infty\}$ starting at ∞ with death rates

$$(1.10) \qquad \lambda_n = n(n - 1 + \theta)/2, \qquad n \geq 0,$$

(∞ is an entrance boundary) and define

$$d_n^\theta(t) = \mathbf{P}\{D_t = n\}, \qquad n \geq 0, t > 0. \tag{1.11}$$

It is known [see, e.g., Tavaré (1984)] that

$$d_n^\theta(t) = \begin{cases} 1 - \sum_{m=1}^{\infty} (2m - 1 + \theta)(m!)^{-1}(-1)^{m-1}\theta_{(m-1)} e^{-\lambda_m t}, & \\ & \text{if } n = 0, \\ \sum_{m=n}^{\infty} (2m - 1 + \theta)(m!)^{-1}(-1)^{m-n}\binom{m}{n}(n + \theta)_{(m-1)} e^{-\lambda_m t}, & \\ & \text{if } n \geq 1. \end{cases} \tag{1.12}$$

Here and elsewhere, we use the notation $a_{(0)} = a_{[0]} = 1$ and, for each $n \geq 1$,

$$a_{(n)} = a(a+1)\cdots(a+n-1), \qquad a_{[n]} = a(a-1)\cdots(a-n+1). \tag{1.13}$$

THEOREM 1.1. *Let S be a compact metric space, and let $\theta > 0$ and $\nu_0 \in \mathscr{P}(S)$. Then the Fleming–Viot process with type space S and mutation operator A defined by* (1.6) *has transition function $P(t, \mu, d\nu)$ given for each $t > 0$ and $\mu \in \mathscr{P}(S)$ by*

$$\begin{aligned} P(t, \mu, \cdot) &= d_0^\theta(t)\Pi_{\theta,\nu_0}(\cdot) \\ &\quad + \sum_{n=1}^{\infty} d_n^\theta(t) \int_{S^n} \mu^n(dx_1 \times \cdots \times dx_n) \\ &\qquad\qquad \Pi_{n+\theta,(n+\theta)^{-1}\{n\eta_n(x_1,\ldots,x_n)+\theta\nu_0\}}(\cdot), \end{aligned} \tag{1.14}$$

where $\mu^n \in \mathscr{P}(S^n)$ is the n-fold product measure $\mu \times \cdots \times \mu$.

In particular, for each $t > 0$ and $\mu \in \mathscr{P}(S)$, $P(t, \mu, \cdot)$ is a mixture of probability distributions of the form (1.7). It is therefore concentrated on $\mathscr{P}_a(S)$ [in fact, a stronger conclusion holds; see Ethier and Kurtz (1987) or Shiga (1990)].

When $S = \{1, \ldots, K\}$, the theorem includes the case (1.4) and, more generally, the case in which

$$q_{ij} = \tfrac{1}{2}\theta_j \geq 0, \qquad i, j \in \{1, \ldots, K\}, i \neq j; \qquad \theta_1 + \cdots + \theta_K > 0. \tag{1.15}$$

Here Shimakura (1981) derived the transition function and Griffiths (1979) its absolutely continuous part.

The theorem has a number of corollaries.

Shiga (1990) proved a strong ergodic theorem in this setting; specifically, he showed that, for each $\mu \in \mathscr{P}(S)$,

$$\lim_{t\to\infty} \|P(t, \mu, \cdot) - \Pi_{\theta,\nu_0}(\cdot)\|_{\text{var}} = 0, \tag{1.16}$$

where $\|\cdot\|_{\text{var}}$ denotes the total variation norm. An immediate consequence of Theorem 1.1 is the following estimate of the rate of convergence in (1.16). (We note that the same estimate can be derived by means of a coupling argument [Ethier and Kurtz (1993c)].)

COROLLARY 1.2. *Under the assumptions of Theorem* 1.1 *and for each* $\mu \in \mathscr{P}(S)$,

$$\|P(t,\mu,\cdot) - \Pi_{\theta,\nu_0}(\cdot)\|_{\mathrm{var}} \le 1 - d_0^\theta(t), \qquad t > 0. \tag{1.17}$$

Moreover, equality holds in (1.17) *if* μ *and* ν_0 *are mutually singular.*

Tavaré (1984) has shown that $e^{-\lambda_1 t} \le 1 - d_0^\theta(t) \le (1+\theta)e^{-\lambda_1 t}$ for all $t > 0$. Here $\lambda_1 = \theta/2$.

If ν_0 is nonatomic (and hence S is uncountable), the Fleming–Viot process of Theorem 1.1 is referred to as the *labeled* infinitely-many-neutral-alleles diffusion model. But there is a simpler, albeit less informative, way of describing the model. Topologize ∇_∞ as a subset of the product space $[0,1]^\infty$, let

$$\overline{\nabla}_\infty = \left\{p = (p_1, p_2, \dots) : p_1 \ge p_2 \ge \cdots \ge 0,\ \sum_{i=1}^\infty p_i \le 1\right\} \tag{1.18}$$

be the closure of ∇_∞ in $[0,1]^\infty$, and define $\Phi: \mathscr{P}(S) \mapsto \overline{\nabla}_\infty$ by letting $\Phi(\mu)$ be the sequence of descending order statistics of the sizes (or masses) of the atoms of μ. The image of the Fleming–Viot process of Theorem 1.1 (with ν_0 nonatomic) under the mapping Φ is referred to as the *unlabeled* infinitely-many-neutral-alleles diffusion model, and it was characterized by Ethier and Kurtz (1981). The following result shows that the unlabeled model converges to equilibrium more rapidly than the labeled model.

COROLLARY 1.3. *Suppose, in addition to the assumptions of Theorem* 1.1, *that* ν_0 *is nonatomic. Then, for each* $t > 0$, *the Borel probability measure* $P(t,\mu,\Phi^{-1}(\cdot))$ *on* $\overline{\nabla}_\infty$ *depends on* $\mu \in \mathscr{P}(S)$ *only through* $\Phi(\mu)$. *In addition, for each* $\mu \in \mathscr{P}(S)$,

$$\|P(t,\mu,\Phi^{-1}(\cdot)) - \Pi_{\theta,\nu_0}(\Phi^{-1}(\cdot))\|_{\mathrm{var}} \le 1 - d_0^\theta(t) - d_1^\theta(t), \qquad t > 0. \tag{1.19}$$

Of course, by (1.7), the Borel probability measure $\Pi_{\theta,\nu_0}(\Phi^{-1}(\cdot))$ on $\overline{\nabla}_\infty$ is just the Poisson–Dirichlet distribution with parameter θ. Tavaré (1984) has shown that $e^{-\lambda_2 t} \le 1 - d_0^\theta(t) - d_1^\theta(t) \le (1/2)(2+\theta)(3+\theta)e^{-\lambda_2 t}$ for all $t > 0$. Here $\lambda_2 = 1 + \theta$.

It is not surprising that the rates of convergence in (1.17) and (1.19) differ. Ethier (1992) showed that the complete set of eigenvalues for the generator of the unlabeled model consists of $0, -\lambda_2, -\lambda_3, \dots$. A similar argument shows that, at least when $S = [0,1]$ and ν_0 is Lebesgue measure, the complete set of eigenvalues for the generator of the labeled model consists of $0, -\lambda_1, -\lambda_2, -\lambda_3, \dots$. Thus, ignoring multiplicities, the labeled model has an extra eigenvalue $-\lambda_1$, which Ewens and Kirby (1975) have referred to in the discrete-time context as the labeling eigenvalue.

The two versions of the infinitely-many-neutral-alleles diffusion model have another significant difference: the unlabeled model has a transition density [Griffiths (1979), Ethier (1992)], whereas the labeled model does not. The latter assertion is a consequence of the next corollary. A possible explanation for this behavior is the fact that the nonzero eigenvalues have finite multiplicities in the unlabeled model, thereby permitting an eigenfunction expansion, whereas they have infinite multiplicities in the labeled model.

COROLLARY 1.4. *Suppose, in addition to the assumptions of Theorem* 1.1, *that* S *is uncountable. Then, for each* $t > 0$, *there exists no* σ*-finite positive Borel measure* Π *on* $\mathscr{P}(S)$ *such that* $P(t, \mu, \cdot\,) \ll \Pi(\cdot)$ *for all* $\mu \in \mathscr{P}(S)$.

Next, we consider the case in which $\theta = 0$. By (1.12),

$$d_n^0(t) = \begin{cases} 0, & \text{if } n = 0, \\ 1 - \displaystyle\sum_{m=2}^{\infty} (2m-1)(-1)^m e^{-m(m-1)t/2}, & \text{if } n = 1, \\ \displaystyle\sum_{m=n}^{\infty} (2m-1)(m!)^{-1}(-1)^{m-n}\binom{m}{n} n_{(m-1)} e^{-m(m-1)t/2}, & \\ & \text{if } n \geq 2. \end{cases} \tag{1.20}$$

Note that $d_n^0(t)$ is the probability that there are n equivalence classes at time t in Kingman's (1982) coalescent.

COROLLARY 1.5. *Let* S *be a compact metric space. Then the Fleming–Viot process with type space* S *and mutation operator* $A = 0$ *has transition function* $P(t, \mu, d\nu)$ *given for each* $t > 0$ *and* $\mu \in \mathscr{P}(S)$ *by*

$$P(t, \mu, \cdot) = \sum_{n=1}^{\infty} d_n^0(t) \int_{S^n} \mu^n(dx_1 \times \cdots \times dx_n) \Pi_{n, \eta_n(x_1, \dots, x_n)}(\cdot). \tag{1.21}$$

Here, for each $t > 0$ and $\mu \in \mathscr{P}(S)$, $P(t, \mu, \cdot\,)$ is concentrated on the subset of $\mathscr{P}_{\mathrm{a}}(S)$ consisting of those measures with only finitely many atoms. When $S = \{1, \dots, K\}$, Corollary 1.5 includes the case in which $q_{ij} = 0$ for all $i, j \in \{1, \dots, K\}$; cf. (1.4) and (1.15). In particular, it generalizes results of Kimura (1955, 1956), Littler and Fackerell (1975), Griffiths (1979) and Shimakura (1981).

There are analogues of Corollaries 1.2 and 1.3 when $\theta = 0$.

COROLLARY 1.6. *Under the assumptions of Corollary* 1.5 *and for each* $\mu \in \mathscr{P}(S)$,

$$\left\| P(t, \mu, \cdot) - \int_S \mu(dx)\, \delta_{\delta_x}(\cdot) \right\|_{\mathrm{var}} \leq 1 - d_1^0(t), \qquad t > 0. \tag{1.22}$$

By the inequalities following the statement of Corollary 1.3, we have $e^{-t} \leq 1 - d_1^0(t) \leq 3e^{-t}$ for all $t > 0$.

COROLLARY 1.7. *Under the assumptions of Corollary* 1.5 *and for each* $t > 0$, *the Borel probability measure* $P(t, \mu, \Phi^{-1}(\cdot))$ *on* ∇_∞ *depends on* $\mu \in \mathscr{P}(S)$ *only through* $\Phi(\mu)$. *In addition, for each* $\mu \in \mathscr{P}(S)$,

$$\|P(t,\mu,\Phi^{-1}(\cdot)) - \delta_{(1,0,0,\dots)}(\cdot)\|_{\mathrm{var}} \le 1 - d_1^0(t), \qquad t > 0. \tag{1.23}$$

In particular, when $\theta = 0$, the unlabeled model is ergodic, whereas the labeled model is not.

Corollary 1.5 also yields a description of a probability distribution on ∇_∞ that occurs in a theorem of Cox and Griffeath (1990) on the mean field asymptotics for the planar stepping stone model with infinitely many types. See also Cox (1989).

COROLLARY 1.8. *Under the assumptions of Corollary* 1.5 *and for each* $t > 0$ *and nonatomic* $\mu \in \mathscr{P}(S)$,

$$P(t,\mu,\Phi^{-1}(\cdot)) = \sum_{n=1}^{\infty} d_n^0(t)\mathbf{P}\{(U_{\{1\}}^n,\dots,U_{\{n\}}^n,0,0,\dots) \in \cdot\}, \tag{1.24}$$

where $U_{\{1\}}^n,\dots,U_{\{n\}}^n$ *are the descending order statistics of the coordinates of* $(U_1^n,\dots,U_n^n)$, *which is uniformly distributed over* Δ_n.

We next provide an alternative form for Theorem 1.1 ($\theta > 0$) and Corollary 1.5 ($\theta = 0$) in the special case in which S is finite or countably infinite. In this case it is more conventional to replace the state space $\mathscr{P}(S)$ by Δ_K or

$$\Delta_\infty = \left\{p = (p_1,p_2,\dots): p_1 \ge 0,\, p_2 \ge 0,\dots,\ \sum_{i=1}^{\infty} p_i = 1\right\}, \tag{1.25}$$

the latter topologized as a subset of the product space $[0,1]^\infty$. To state the result, we need to generalize the Dirichlet distribution (1.5) in two directions, allowing $K = \infty$ and allowing some of the parameters to be 0. If $1 \le K < \infty$, let $\theta_1 \ge 0,\dots,\theta_K \ge 0$, assume $\theta_1 + \cdots + \theta_K > 0$, and put $\Theta = (\theta_1,\dots,\theta_K)$. If $K = \infty$, let $\theta_1 \ge 0$, $\theta_2 \ge 0,\dots,$ assume $0 < \theta_1 + \theta_2 + \cdots < \infty$, and put $\Theta = (\theta_1,\theta_2,\dots)$. For $1 \le K \le \infty$, we define $\mathrm{Dirichlet}(\Theta) \in \mathscr{P}(\Delta_K)$ by

$$\mathrm{Dirichlet}(\Theta)(\cdot) = \begin{cases} \mathbf{P}\{(Y_1/Z,\dots,Y_K/Z) \in \cdot\}, & \text{if } K < \infty,\\ \mathbf{P}\{(Y_1/Z,Y_2/Z,\dots) \in \cdot\}, & \text{if } K = \infty, \end{cases} \tag{1.26}$$

where $Y_1, Y_2,\dots$ are independent with Y_i being gamma$(\theta_i, 1)$ distributed (by definition, the gamma(0, 1) distribution is δ_0), and $Z = \sum_{i=1}^K Y_i$. This definition is consistent with (1.5). When all parameters are 0, the Dirichlet distribution can be defined arbitrarily.

COROLLARY 1.9. (i) *Let* $2 \le K < \infty$, *let* $\theta_1 \ge 0,\dots,\theta_K \ge 0$, *and put* $\theta = \theta_1 + \cdots + \theta_K$ *and* $\Theta = (\theta_1,\dots,\theta_K)$. *Then the diffusion process in* Δ_K *with generator* L, *given by* (1.2) *with infinitesimal matrix* (q_{ij}) *satisfying* $q_{ij} = (1/2)\theta_j$ *for all* $i, j \in \{1,\dots,K\}$ *for which* $i \ne j$, *has transition function*

$P(t, p, dq)$ given for each $t > 0$ and $p \in \Delta_K$ by

$$(1.27)\quad P(t,p,\cdot) = \sum_{n=0}^{\infty} d_n^\theta(t) \sum_{\alpha \in (\mathbf{Z}_+)^K:\, |\alpha|=n} \binom{n}{\alpha} \prod_{i=1}^{K} p_i^{\alpha_i}\, \text{Dirichlet}(\alpha + \Theta)(\cdot).$$

(ii) *Let $K = \infty$, let $\theta_1 \geq 0$, $\theta_2 \geq 0, \dots,$ assume that $\theta \equiv \theta_1 + \theta_2 + \cdots < \infty$, and put $\Theta = (\theta_1, \theta_2, \dots)$. Then the diffusion process in Δ_∞ with generator L, given by* (1.2) [*except that $K = \infty$ and $\mathscr{D}(L) = \{F|_{\Delta_\infty}: F \in C^2(\mathbf{R}^\infty)$ depends on only finitely many coordinates$\}$*] *with infinitesimal matrix (q_{ij}) satisfying $q_{ij} = (1/2)\theta_j$ for all $i, j \geq 1$ for which $i \neq j$, has transition function $P(t, p, dq)$ given for each $t > 0$ and $p \in \Delta_\infty$ by* (1.27) *with $K = \infty$.*

If $\theta = 0$, the $n = 0$ term in (1.27) is absent, and the probabilities $d_n^\theta(t)$ are given by (1.20). See Ethier (1981) for the details of the characterization of the infinite-dimensional diffusion process in (ii).

The formula (1.27), which first appeared in this form in Griffiths and Li (1983) (assuming $K < \infty$ and $\theta_1 = \cdots = \theta_K > 0$) and Tavaré (1984) (assuming $K < \infty$ and $\theta_1 > 0, \dots, \theta_K > 0$), has a simple intuitive interpretation based on Griffiths' (1980) work on lines of descent. See Donnelly and Tavaré (1987) for a lucid account, which includes a (nonrigorous) derivation of (1.27) using these ideas. Of course, our formula (1.14) has a similar interpretation.

It follows that, if $\theta > 0$, the diffusion process of Corollary 1.9 has a transition density with respect to its unique stationary distribution [namely, Dirichlet(Θ)] if and only if $\theta_i > 0$ for each i. Part (ii) of the corollary answers a question raised by Shimakura [(1981), Section 6-5].

It is known [Shiga (1990), Ethier (1990)] that the Fleming–Viot process of Theorem 1.1 is reversible with respect to its unique stationary distribution (1.7). In the setting of Corollary 1.9 with $\theta > 0$, the analogous result is clear from (1.27), at least if $K < \infty$ and $\theta_1 > 0, \dots, \theta_K > 0$. It is therefore disappointing that the reversibility of the Fleming–Viot process of Theorem 1.1 does not seem to be an immediate consequence of (1.14). Our last corollary remedies this situation.

COROLLARY 1.10. *Under the assumptions of Theorem* 1.1 *and for each $t > 0$,*

$$\begin{aligned}
&\Pi_{\theta,\nu_0}(d\mu) P(t,\mu,d\nu)\\
&\quad = d_0^\theta(t) \Pi_{\theta,\nu_0}(d\mu) \Pi_{\theta,\nu_0}(d\nu)\\
(1.28)\quad &\qquad + \sum_{n=1}^{\infty} d_n^\theta(t) \int_{\mathscr{P}(S)} \Pi_{\theta,\nu_0}(d\lambda) \int_{S^n} \lambda^n(dx_1 \times \cdots \times dx_n)\\
&\qquad \times \left\{\Pi_{n+\theta,(n+\theta)^{-1}\{n\eta_n(x_1,\dots,x_n)+\theta\nu_0\}}(d\mu)\, \Pi_{n+\theta,(n+\theta)^{-1}\{n\eta_n(x_1,\dots,x_n)+\theta\nu_0\}}(d\nu)\right\}.
\end{aligned}$$

In particular, $\Pi_{\theta,\nu_0}(d\mu)P(t,\mu,d\nu) = \Pi_{\theta,\nu_0}(d\nu)P(t,\nu,d\mu)$ for each $t > 0$.

We now comment briefly on the proof of Theorem 1.1. Perhaps the most efficient proof of the theorem would take Corollary 1.9(i) for granted. This would then give (1.14) for all sets of the form $\Psi^{-1}(B)$, where B is a Borel subset of Δ_K and Ψ: $\mathscr{P}(S) \mapsto \Delta_K$ is given by $\Psi(\nu) = (\nu(\Lambda_1), \dots, \nu(\Lambda_K))$; here $K \geq 2$ and $\Lambda_1, \dots, \Lambda_K$ is a partition of S into Borel sets. The theorem would then follow easily. Alternatively, Corollary 1.9(i) would give (1.14) in the special case in which both μ and ν_0 are purely atomic with only finitely many atoms, and since the set of measures in $\mathscr{P}_{\mathrm{a}}(S)$ with only finitely many atoms is dense in $\mathscr{P}(S)$, the theorem would follow by a continuity argument.

Instead, for the sake of clarity and elegance, we provide a self-contained proof. It does not seem to substantially simplify matters to treat the finite-dimensional case first. (We use the first approach of the preceding paragraph, however, in the proof of Corollary 1.10.) An interesting aspect of the proof is that the formula (1.12) for the pure death probabilities (1.11) is not used; rather, we use the fact that these probabilities satisfy the Kolmogorov forward equation.

As a byproduct of the proof, we obtain an explicit formula for the "moments"

$$\int_{\mathscr{P}(S)} \langle f_1, \nu \rangle \cdots \langle f_m, \nu \rangle P(t, \mu, d\nu), \tag{1.29}$$

where $m \geq 1$, $f_1, \dots, f_m \in C(S)$, $t > 0$, and $\mu \in \mathscr{P}(S)$; see (3.13) below. Dynkin (1989) implicitly used the function-valued dual process introduced by Dawson and Hochberg (1982) to obtain an analytical expression for similar moments in a very general framework. But because of our special choice of A [see (1.6)], a simpler dual process is available and Dynkin's result is not needed here.

Section 2 contains some lemmas, and the proofs of the theorem and the corollaries can be found in Section 3.

Finally, we remark that similarly explicit formulas can be derived for the transition functions of certain measure-valued branching diffusions with immigration. This is not surprising, in view of the relationship [Shiga (1990)] between such diffusions and Fleming–Viot processes. See Ethier and Griffiths (1993).

2. Lemmas. The Poisson–Dirichlet distribution with parameter $\theta > 0$ can be described as follows [Kingman (1975)]: Consider an inhomogeneous Poisson point process on $(0, \infty)$ with intensity function $\theta u^{-1} e^{-u}$, $u > 0$. With probability 1, the points can be arranged in decreasing order $\sigma_1 > \sigma_2 > \cdots$ and have a finite sum $s = \sigma_1 + \sigma_2 + \cdots$. Moreover,

$$(\sigma_1/s, \sigma_2/s, \dots) \text{ is Poisson–Dirichlet}(\theta) \tag{2.1}$$

and is independent of s.

Unless otherwise noted, S (a compact metric space), $\theta > 0$, and $\nu_0 \in \mathscr{P}(S)$ are fixed throughout.

LEMMA 2.1. *Let* $\theta_1, \theta_2 > 0$ *and* $\nu_1, \nu_2 \in \mathscr{P}(S)$. *If the* $\mathscr{P}(S)$-*valued random variables* μ_1 *and* μ_2 *have distributions* Π_{θ_1,ν_1} *and* Π_{θ_2,ν_2}, *if the* $[0,1]$-*valued random variable* ε *is* beta(θ_1,θ_2) *distributed, and if* μ_1, μ_2, *and* ε *are independent, then*

$$(2.2) \qquad \mathbf{P}\{\varepsilon\mu_1 + (1-\varepsilon)\mu_2 \in \cdot\} = \Pi_{\theta_1+\theta_2,(\theta_1+\theta_2)^{-1}(\theta_1\nu_1+\theta_2\nu_2)}(\cdot).$$

PROOF. Let $\sigma_1 > \sigma_2 > \cdots$ and s be as in (2.1) with $\theta = \theta_1 + \theta_2$. Let $\xi_1^{(1)}, \xi_2^{(1)}, \ldots$ be i.i.d. ν_1, let $\xi_1^{(2)}, \xi_2^{(2)}, \ldots$ be i.i.d. ν_2, and let $\chi_1, \chi_2, \ldots$ be i.i.d. with $\mathbf{P}\{\chi_i = 1\} = \theta_1/(\theta_1+\theta_2) = 1 - \mathbf{P}\{\chi_i = 2\}$. Assume that $\{\sigma_i\}$, $\{\xi_j^{(1)}\}$, $\{\xi_j^{(2)}\}$, and $\{\chi_i\}$ are independent. Define the sequences $\sigma_1^{(1)}, \sigma_2^{(1)}, \ldots, \sigma_1^{(2)}, \sigma_2^{(2)}, \ldots,$ and $\xi_1, \xi_2, \ldots$ by

$$(2.3) \qquad \sigma_j^{(\chi_i)} = \sigma_i \quad \text{and} \quad \xi_i = \xi_j^{(\chi_i)} \quad \text{if } j = |\{k\colon 1 \le k \le i,\ \chi_k = \chi_i\}|.$$

Then, with $s_1 = \sigma_1^{(1)} + \sigma_2^{(1)} + \cdots$ and $s_2 = \sigma_1^{(2)} + \sigma_2^{(2)} + \cdots$, we have $s = s_1 + s_2$ and

$$(2.4) \qquad \frac{s_1}{s}\sum_{j=1}^{\infty}\frac{\sigma_j^{(1)}}{s_1}\delta_{\xi_j^{(1)}} + \frac{s_2}{s}\sum_{j=1}^{\infty}\frac{\sigma_j^{(2)}}{s_2}\delta_{\xi_j^{(2)}} = \sum_{i=1}^{\infty}\frac{\sigma_i}{s}\delta_{\xi_i}.$$

It remains to check that $\{\sigma_j^{(1)}\}$ and $\{\sigma_j^{(2)}\}$ are independent Poisson point processes on $(0,\infty)$ with intensity functions $\theta_1 u^{-1}e^{-u}$ and $\theta_2 u^{-1}e^{-u}$ $(u > 0)$; that s_1/s is beta(θ_1,θ_2) [cf. Donnelly and Tavaré (1987)]; that $\xi_1, \xi_2, \ldots$ are i.i.d. $(\theta_1+\theta_2)^{-1}(\theta_1\nu_1 + \theta_2\nu_2)$; and that the required independence holds. The result then follows from (2.4). □

LEMMA 2.2. *Let* $1 \le K < \infty$ *and suppose there exist distinct points* $x_1, \ldots, x_K \in S$ *such that* $\nu_0(\{x_1\}) + \cdots + \nu_0(\{x_K\}) = 1$. *Put* $\Theta = (\theta\nu_0(\{x_1\}), \ldots, \theta\nu_0(\{x_K\}))$, *and let* $(V_1, \ldots, V_K)$ *be* Δ_K-*valued with distribution* Dirichlet(Θ) [*see* (1.26)].

Alternatively, let $K = \infty$ *and suppose there exist distinct points* $x_1, x_2, \ldots \in S$ *such that* $\nu_0(\{x_1\}) + \nu_0(\{x_2\}) + \cdots = 1$. *Put* $\Theta = (\theta\nu_0(\{x_1\}), \theta\nu_0(\{x_2\}), \ldots)$, *and let* $(V_1, V_2, \ldots)$ *be* Δ_∞-*valued with distribution* Dirichlet(Θ).

Then, in either case,

$$(2.5) \qquad \Pi_{\theta,\nu_0}(\cdot) = \mathbf{P}\left\{\sum_{j=1}^{K} V_j\,\delta_{x_j} \in \cdot\right\}.$$

PROOF. Let $\sigma_1 > \sigma_2 > \cdots$ and s be as in (2.1), and let $\xi_1, \xi_2, \ldots$ be i.i.d. ν_0, independent of $\{\sigma_i\}$. Then

$$(2.6) \qquad \sum_{i=1}^{\infty}\frac{\sigma_i}{s}\delta_{\xi_i} = \sum_{j=1}^{K}\left\{\sum_{i\ge 1\colon \xi_i = x_j}\frac{\sigma_i}{s}\right\}\delta_{x_j},$$

and the result follows as before [cf. Donnelly and Tavaré (1987)]. □

Lemma 2.3. *For each $n \geq 1$, $m \geq 1$, $f_1, \dots, f_m \in C(S)$, and $\mu \in \mathscr{P}(S)$,*

$$\begin{aligned}\int_{S^n} \mathbf{E}\left[\prod_{i=1}^{m} \langle f_i, \zeta_n(x_1, \dots, x_n)\rangle\right] \mu^n(dx_1 \times \cdots \times dx_n) \qquad (2.7)\\ = \sum_{k=1}^{m} \frac{n_{[k]}}{n_{(m)}} \sum_{\beta \in \pi(m,k)} |\beta_1|! \cdots |\beta_k|! \prod_{j=1}^{k} \left\langle \prod_{i \in \beta_j} f_i, \mu \right\rangle,\end{aligned}$$

where $\zeta_n(x_1, \dots, x_n)$ is defined in terms of a Δ_n-valued Dirichlet$(1, \dots, 1)$ *(or uniform) random variable $(U_1, \dots, U_n)$ by*

$$\zeta_n(x_1, \dots, x_n) = U_1 \delta_{x_1} + \cdots + U_n \delta_{x_n}, \qquad (2.8)$$

and $\pi(m,k)$ is the set of partitions β of $\{1, \dots, m\}$ into k nonempty subsets $\beta_1, \dots, \beta_k$, labeled so that $\min \beta_1 < \cdots < \min \beta_k$.

Proof. We proceed by induction on n. Let $m \geq 1$, $f_1, \dots, f_m \in C(S)$, and $\mu \in \mathscr{P}(S)$ be arbitrary. If $n = 1$, both sides of (2.7) are equal to $\langle \prod_{i=1}^{m} f_i, \mu\rangle$.

So let us suppose that $n \geq 2$. Let $Y_1, \dots, Y_n$ be independent exponential random variables with parameter 1, and put $Z_n = Y_1 + \cdots + Y_n$. Then we can define [consistently with (2.8)]

$$\zeta_n(x_1, \dots, x_n) = \sum_{j=1}^{n} \frac{Y_j}{Z_n} \delta_{x_j}, \qquad \zeta_{n-1}(x_1, \dots, x_{n-1}) = \sum_{j=1}^{n-1} \frac{Y_j}{Z_{n-1}} \delta_{x_j}, \qquad (2.9)$$

to conclude that

$$\zeta_n(x_1, \dots, x_n) = \frac{Z_{n-1}}{Z_n} \zeta_{n-1}(x_1, \dots, x_{n-1}) + \frac{Y_n}{Z_n} \delta_{x_n}. \qquad (2.10)$$

It follows that

$$\begin{aligned}&\int_{S^n} \mathbf{E}\left[\prod_{i=1}^{m} \langle f_i, \zeta_n(x_1, \dots, x_n)\rangle\right] \mu^n(dx_1 \times \cdots \times dx_n)\\ &= \int_{S^n} \mathbf{E}\left[\prod_{i=1}^{m} \left\{\frac{Z_{n-1}}{Z_n} \langle f_i, \zeta_{n-1}(x_1, \dots, x_{n-1})\rangle \right.\right.\\ &\qquad\qquad \left.\left. + \frac{Y_n}{Z_n} f_i(x_n)\right\}\right] \mu^n(dx_1 \times \cdots \times dx_n)\\ &= \sum_{M \subset \{1, \dots, m\}} \mathbf{E}\left[\left(\frac{Z_{n-1}}{Z_n}\right)^{|M|} \left(\frac{Y_n}{Z_n}\right)^{|M^c|}\right] \qquad (2.11)\\ &\qquad \times \int_{S^n} \mathbf{E}\left[\prod_{i \in M} \langle f_i, \zeta_{n-1}(x_1, \dots, x_{n-1})\rangle \prod_{i \in M^c} f_i(x_n)\right] \mu^n(dx_1 \times \cdots \times dx_n)\\ &= \sum_{M \subset \{1, \dots, m\}} \frac{(n-1)_{(|M|)} 1_{(|M^c|)}}{n_{(m)}}\\ &\qquad \times \int_{S^{n-1}} \mathbf{E}\left[\prod_{i \in M} \langle f_i, \zeta_{n-1}(x_1, \dots, x_{n-1})\rangle\right] \mu^{n-1}(dx_1 \times \cdots \times dx_{n-1})\\ &\qquad\qquad \times \left\langle \prod_{i \in M^c} f_i, \mu \right\rangle,\end{aligned}$$

where the sum over $M \subset \{1, \dots, m\}$ contains 2^m terms and $M^c = \{1, \dots, m\} - M$. Now by the induction hypothesis, the right-hand side of (2.11) is equal to

$$\begin{aligned}&\sum_{M \subset \{1,\dots,m\}:\, M \neq \varnothing} \frac{(n-1)_{(|M|)}\, |M^c|!}{n_{(m)}} \sum_{l=1}^{|M|} \frac{(n-1)_{[l]}}{(n-1)_{(|M|)}} \sum_{\gamma \in \pi(M,l)} |\gamma_1|! \cdots |\gamma_l|! \\ &\qquad \times \prod_{j=1}^{l} \Big\langle \prod_{i \in \gamma_j} f_i, \mu \Big\rangle \Big\langle \prod_{i \in M^c} f_i, \mu \Big\rangle + \frac{m!}{n_{(m)}} \Big\langle \prod_{i \in \{1,\dots,m\}} f_i, \mu \Big\rangle, \end{aligned} \tag{2.12}$$

where $\pi(M, l)$ is the set of partitions γ of M into l nonempty subsets $\gamma_1, \dots, \gamma_l$, labeled so that $\min \gamma_1 < \cdots < \min \gamma_l$.

It remains to check that the right-hand side of (2.7) coincides with (2.12). Fix $k \in \{1, \dots, m\}$ and $\beta \in \pi(m, k)$, and compare coefficients of

$$\frac{1}{n_{(m)}} |\beta_1|! \cdots |\beta_k|! \prod_{j=1}^{k} \Big\langle \prod_{i \in \beta_j} f_i, \mu \Big\rangle. \tag{2.13}$$

On the right-hand side of (2.7) the coefficient of (2.13) is $n_{[k]}$. In (2.12) we get $k + 1$ contributions, depending on whether M^c is $\beta_1, \dots, \beta_k$, or empty. Thus, the coefficient of (2.13) is

$$k(n-1)_{[k-1]} + (n-1)_{[k]} = (k + n - k)(n-1)_{[k-1]} = n_{[k]}, \tag{2.14}$$

as required. □

LEMMA 2.4. *For each* $m \geq 1$ *and* $f_1, \dots, f_m \in C(S)$,

$$\begin{aligned}&\int_{\mathscr{P}(S)} \langle f_1, \nu \rangle \cdots \langle f_m, \nu \rangle \Pi_{\theta, \nu_0}(d\nu) \\ &\quad = \sum_{l=1}^{m} \sum_{\gamma \in \pi(m,l)} (|\gamma_1| - 1)! \cdots (|\gamma_l| - 1)! \frac{\theta^l}{\theta_{(m)}} \prod_{j=1}^{l} \Big\langle \prod_{i \in \gamma_j} f_i, \nu_0 \Big\rangle. \end{aligned} \tag{2.15}$$

PROOF. This is a restatement of Lemma 2.2 of Ethier (1990). □

LEMMA 2.5. *The probabilities* $d_n^\theta(t)$ *defined by* (1.11) *have the following properties*:

(i) $\sum_{n \in \mathbf{Z}_+} d_n^\theta(t) = 1$ *for each* $t > 0$.

(ii) *If* $a_0, a_1, a_2, \dots$ *is a real sequence such that* $\lim_{n \to \infty} a_n = a \in \mathbf{R}$ *exists, then* $\lim_{t \to 0} \sum_{n \in \mathbf{Z}_+} a_n d_n^\theta(t) = a$.

(iii) $\sum_{n \in \mathbf{Z}_+} n^r d_n^\theta(t) < \infty$ *for each* $r \geq 1$ *and* $t > 0$.

(iv) *The Kolmogorov forward equation holds, that is,*

$$\frac{d}{dt} d_n^\theta(t) = -\lambda_n d_n^\theta(t) + \lambda_{n+1} d_{n+1}^\theta(t), \qquad n \in \mathbf{Z}_+, t > 0. \tag{2.16}$$

REMARK. We have implicitly assumed that $\theta > 0$. But the lemma holds also for $\theta = 0$.

PROOF. Define the operator Ω on $C(\mathbf{Z}_+)$ by

$$(\Omega f)(n) = \begin{cases} 0, & \text{if } n = 0, \\ \lambda_n(f(n-1) - f(n)), & \text{if } n \geq 1, \end{cases} \tag{2.17}$$

and let $\mathbf{Z}_+ \cup \{\infty\}$ denote the one-point compactification of $\mathbf{Z}_+$. We begin by showing that the Hille–Yosida theorem applies to the operator Ω_0 on $C(\mathbf{Z}_+ \cup \{\infty\})$ with domain $\mathscr{D}(\Omega_0) = \{f \in C(\mathbf{Z}_+ \cup \{\infty\})\colon \lim_{n\to\infty}(\Omega f)(n)$ exists and is finite$\}$, defined by $(\Omega_0 f)(n) = (\Omega f)(n)$ for each $n \in \mathbf{Z}_+$ and $(\Omega_0 f)(\infty) = \lim_{n\to\infty}(\Omega f)(n)$. Since $\mathscr{D}(\Omega_0)$ contains the functions f for which $(\Omega f)(n) = 0$ for all sufficiently large n, it is dense in $C(\mathbf{Z}_+ \cup \{\infty\})$. Given $g \in C(\mathbf{Z}_+ \cup \{\infty\})$ and $\lambda > 0$, the equation $(\lambda - \Omega)f = g$ can be solved recursively for f, beginning with $f(0) = \lambda^{-1}g(0)$. An inductive argument then shows that $|f(n)| \leq \lambda^{-1}\|g\|$ for all $n \in \mathbf{Z}_+$. Moreover, $\Omega f = \lambda f - g$, so $|f(n-1) - f(n)| \leq 2\lambda_n^{-1}\|g\|$ for each $n \geq 1$. It follows from $\sum \lambda_n^{-1} < \infty$ that $\{f(n)\}$ is a Cauchy sequence, and hence so is $\{(\Omega f)(n)\}$. In other words, $f \in \mathscr{D}(\Omega_0)$ and $(\lambda - \Omega_0)f = g$.

Let $\{S(t)\}$ denote the resulting Feller semigroup on $C(\mathbf{Z}_+ \cup \{\infty\})$. For each $N \geq 1$,

$$\begin{aligned} \int_0^\infty e^{-t} \sum_{n=0}^{N} d_n^\theta(t)\, dt &= \int_0^\infty e^{-t} S(t) I_{\{0,1,\dots,N\}}(\infty)\, dt \\ &= (1 - \Omega_0)^{-1} I_{\{0,1,\dots,N\}}(\infty) = \prod_{n=N+1}^{\infty} \frac{\lambda_n}{1+\lambda_n}, \end{aligned} \tag{2.18}$$

so, letting $N \to \infty$ and using $\sum \lambda_n^{-1} < \infty$ once again, we obtain $\int_0^\infty e^{-t} \sum_{n \in \mathbf{Z}_+} d_n^\theta(t)\, dt = 1$; since the sum in the latter integral is nondecreasing in t and bounded by 1, (i) follows.

As for (ii), define $f \in C(\mathbf{Z}_+ \cup \{\infty\})$ by $f(n) = a_n$ for each $n \in \mathbf{Z}_+$ and $f(\infty) = a$. Then $\lim_{t\to 0} S(t)f(\infty) = f(\infty)$, which by (i) is equivalent to the desired result.

Next, let $T_1, T_2, \dots$ be independent exponential random variables with parameters $\lambda_1, \lambda_2, \dots$. Then

$$\begin{aligned} d_n^\theta(t) &= \mathbf{P}\left\{ \sum_{m=n+1}^{\infty} T_m \leq t < \sum_{m=n}^{\infty} T_m \right\} \\ &\leq \mathbf{P}\left\{ \sum_{m=n}^{\infty} T_m > t \right\} \\ &\leq e^{-\sqrt{n}t}\, \mathbf{E}\left[\exp\left\{ \sqrt{n} \sum_{m=n}^{\infty} T_m \right\} \right] \\ &= e^{-\sqrt{n}t} \prod_{m=n}^{\infty} \left(1 - \frac{\sqrt{n}}{\lambda_n} \right)^{-1} \\ &\leq C e^{-\sqrt{n}t} \end{aligned} \tag{2.19}$$

for all $n \geq 3$, where $C = \prod_{m=3}^{\infty}(1 - \sqrt{n}/\lambda_n)^{-1}$, since $\lambda_n > \sqrt{n}$ for such n. This implies (iii).

Turning to (iv), fix $n \in \mathbf{Z}_+$ and $t > 0$. Then, using (i),

$$\frac{d_n^\theta(t+h) - d_n^\theta(t)}{h} = \sum_{m \in \mathbf{Z}_+:\, m \geq n} d_m^\theta(t) \frac{p_{mn}(h) - \delta_{mn}}{h} \tag{2.20}$$

for all $h > 0$, where $p_{mn}(h) = S(h)I_{\{n\}}(m)$. Now

$$\sup_{m \in \mathbf{Z}_+:\, m \geq n+1} \frac{p_{mn}(h)}{h} \leq \frac{1 - e^{-\lambda_{n+1}h}}{h} \leq \lambda_{n+1} \tag{2.21}$$

for all $h > 0$, so we can apply the dominated convergence theorem to (2.20) to get (2.16). (A continuous function with a continuous right derivative is differentiable.) □

3. Proofs. This section contains the proofs of the results stated in Section 1.

Proof of Theorem 1.1. The formula

$$\mathscr{T}(t)\varphi(\mu) = \int_{\mathscr{P}(S)} \varphi(\nu)P(t, \mu, d\nu) \tag{3.1}$$

defines a Feller semigroup $\{\mathscr{T}(t)\}$ on $C(\mathscr{P}(S))$, which is generated by the closure of $\mathscr{L}$ defined by (1.3) and (1.6) [Ethier and Kurtz (1993a)]. For each $t > 0$ and $\mu \in \mathscr{P}(S)$, let $Q(t, \mu, \cdot\,)$ denote the right-hand side of (1.14), and define the one-parameter family $\{\mathscr{U}(t),\ t > 0\}$ of bounded linear operators on $C(\mathscr{P}(S))$ by

$$\mathscr{U}(t)\varphi(\mu) = \int_{\mathscr{P}(S)} \varphi(\nu)Q(t, \mu, d\nu). \tag{3.2}$$

For each $m \geq 1$ and $f_1, \dots, f_m \in C(S)$, define $\varphi_{f_1, \dots, f_m} \in C(\mathscr{P}(S))$ by

$$\varphi_{f_1, \dots, f_m}(\mu) = \langle f_1, \mu\rangle \cdots \langle f_m, \mu\rangle, \tag{3.3}$$

and note that $\varphi_{f_1, \dots, f_m} \in \mathscr{D}(\mathscr{L})$ and

$$\mathscr{T}(t)\varphi_{f_1, \dots, f_m} = \varphi_{f_1, \dots, f_m} + \int_0^t \mathscr{T}(s)\mathscr{L}\varphi_{f_1, \dots, f_m}\, ds, \qquad t \geq 0. \tag{3.4}$$

Suppose for the moment that we could show that

$$\mathscr{U}(t)\varphi_{f_1, \dots, f_m} = \varphi_{f_1, \dots, f_m} + \int_0^t \mathscr{U}(s)\mathscr{L}\varphi_{f_1, \dots, f_m}\, ds, \qquad t > 0, \tag{3.5}$$

for all $m \geq 1$ and $f_1, \dots, f_m \in C(S)$. Then, in view of the identity

$$\begin{aligned} \mathscr{L}\varphi_{f_1, \dots, f_m} = {} & \sum_{1 \leq i < j \leq m} \varphi_{f_1, \dots, f_{i-1}, f_i f_j, f_{i+1}, \dots, f_{j-1}, f_{j+1}, \dots, f_m} \\ & + \tfrac{1}{2}\theta \sum_{j=1}^{m} \langle f_j, \nu_0\rangle \varphi_{f_1, \dots, f_{j-1}, f_{j+1}, \dots, f_m} \\ & - \lambda_m \varphi_{f_1, \dots, f_m}, \end{aligned} \tag{3.6}$$

(3.4) and (3.5) would imply that, for each $m \geq 1$, the function

$$h_m(t) \equiv \sup_{\substack{f_1,\ldots,f_m \in C(S):\\ \max_{1 \leq i \leq m}\|f_i\| \leq 1}} \sup_{\mu \in \mathscr{P}(S)} \left|\mathscr{T}(t)\varphi_{f_1,\ldots,f_m}(\mu) - \mathscr{U}(t)\varphi_{f_1,\ldots,f_m}(\mu)\right| \tag{3.7}$$

satisfies

$$h_m(t) \leq 2\lambda_m \int_0^t h_m(s)\,ds, \qquad t > 0, \tag{3.8}$$

and hence is identically 0 by Gronwall's inequality. From this we could conclude that $\mathscr{T}(t) = \mathscr{U}(t)$ for all $t > 0$, and consequently that $P(t, \mu, \cdot\,) = Q(t, \mu, \cdot\,)$ for all $t > 0$ and $\mu \in \mathscr{P}(S)$, as required. Thus, to complete the proof, it is enough to verify (3.5).

Fix $m \geq 1$, $f_1, \ldots, f_m \in C(S)$, and $\mu \in \mathscr{P}(S)$. It will suffice to show that

$$\begin{aligned} \frac{d}{dt}\mathscr{U}(t)\varphi_{f_1,\ldots,f_m}(\mu) = {} & \sum_{1 \leq i < j \leq m} \mathscr{U}(t)\varphi_{f_1,\ldots,f_{i-1},f_i f_j,f_{i+1},\ldots,f_{j-1},f_{j+1},\ldots,f_m}(\mu) \\ & + \tfrac{1}{2}\theta \sum_{j=1}^{m} \langle f_j, \nu_0 \rangle \mathscr{U}(t)\varphi_{f_1,\ldots,f_{j-1},f_{j+1},\ldots,f_m}(\mu) \\ & - \lambda_m \mathscr{U}(t)\varphi_{f_1,\ldots,f_m}(\mu) \end{aligned} \tag{3.9}$$

for all $t > 0$, and

$$\lim_{t \to 0} \mathscr{U}(t)\varphi_{f_1,\ldots,f_m}(\mu) = \varphi_{f_1,\ldots,f_m}(\mu). \tag{3.10}$$

For each $n \geq 1$ and $(x_1, \ldots, x_n) \in S^n$, Lemmas 2.1 and 2.2 imply that

$$\begin{aligned} & \Pi_{n+\theta,(n+\theta)^{-1}\{n\eta_n(x_1,\ldots,x_n)+\theta\nu_0\}}(\cdot) \\ & \quad = \mathbf{P}\{\varepsilon(U_1\delta_{x_1} + \cdots + U_n\delta_{x_n}) + (1-\varepsilon)\Lambda \in \cdot\}, \end{aligned} \tag{3.11}$$

where ε is beta(n, θ) distributed, $(U_1, \ldots, U_n)$ is Δ_n-valued Dirichlet$(1, \ldots, 1)$, Λ has distribution Π_{θ,ν_0}, and ε, $(U_1, \ldots, U_n)$, and Λ are independent; therefore, using the notation in (2.8),

$$\begin{aligned} & \int_{\mathscr{P}(S)} \varphi_{f_1,\ldots,f_m}(\nu)\Pi_{n+\theta,(n+\theta)^{-1}\{n\eta_n(x_1,\ldots,x_n)+\theta\nu_0\}}(d\nu) \\ & \quad = \mathbf{E}\left[\prod_{i=1}^{m}\langle f_i, \varepsilon\zeta_n(x_1,\ldots,x_n) + (1-\varepsilon)\Lambda\rangle\right] \\ & \quad = \sum_{M \subset \{1,\ldots,m\}} \mathbf{E}\left[\varepsilon^{|M|}(1-\varepsilon)^{|M^c|}\right]\mathbf{E}\left[\prod_{i \in M}\langle f_i, \zeta_n(x_1,\ldots,x_n)\rangle\right] \\ & \qquad \times \mathbf{E}\left[\prod_{i \in M^c}\langle f_i, \Lambda\rangle\right], \end{aligned} \tag{3.12}$$

where $M^c = \{1, \dots, m\} - M$. It follows from this and Lemmas 2.3 and 2.4 that

$$
\begin{aligned}
\mathscr{U}(t)\varphi_{f_1,\dots,f_m}(\mu) &= \sum_{n=0}^{\infty} d_n^\theta(t) \sum_{M \subset \{1,\dots,m\}} \frac{1}{(n+\theta)_{(m)}} \\
&\quad \times \left\{ \sum_{k=1}^{|M|} n_{[k]} \sum_{\beta \in \pi(M,k)} |\beta_1|! \cdots |\beta_k|! \prod_{j=1}^{k} \left\langle \prod_{i \in \beta_j} f_i, \mu \right\rangle \right\} \\
&\quad \times \left\{ \sum_{l=1}^{|M^c|} \sum_{\gamma \in \pi(M^c,l)} (|\gamma_1|-1)! \cdots (|\gamma_l|-1)! \theta^l \prod_{j=1}^{l} \left\langle \prod_{i \in \gamma_j} f_i, \nu_0 \right\rangle \right\}
\end{aligned}
\tag{3.13}
$$

for all $t > 0$, where $\pi(M,k)$ is as in (2.12) and the first (resp., second) expression within braces is 1 if M (resp., M^c) is empty.

Notice that (3.10) is immediate from (3.13) and Lemma 2.5(ii).

Fix $t > 0$ and $M \subset \{1, \dots, m\}$. If $M \neq \varnothing$, fix $k \in \{1, \dots, |M|\}$ and $\beta \in \pi(M,k)$; if $M = \varnothing$, put $k = 0$. If $M^c \neq \varnothing$, fix $l \in \{1, \dots, |M^c|\}$ and $\gamma \in \pi(M^c, l)$; if $M^c = \varnothing$, put $l = 0$. We verify (3.9) by comparing coefficients of

$$
\begin{aligned}
&\left\{ |\beta_1|! \cdots |\beta_k|! \prod_{j=1}^{k} \left\langle \prod_{i \in \beta_j} f_i, \mu \right\rangle \right\} \\
&\quad \times \left\{ (|\gamma_1|-1)! \cdots (|\gamma_l|-1)! \theta^l \prod_{j=1}^{l} \left\langle \prod_{i \in \gamma_j} f_i, \nu_0 \right\rangle \right\}
\end{aligned}
\tag{3.14}
$$

on both sides of (3.9) [after substituting (3.13)]; the first (resp., second) expression within braces in (3.14) is 1 if M (resp., M^c) is empty. The coefficient of (3.14) on the left-hand side of (3.9) is

$$
\begin{aligned}
\sum_{n=0}^{\infty} \frac{d}{dt} d_n^\theta(t) \frac{n_{[k]}}{(n+\theta)_{(m)}} &= \sum_{n=0}^{\infty} \{-\lambda_n d_n^\theta(t) + \lambda_{n+1} d_{n+1}^\theta(t)\} \frac{n_{[k]}}{(n+\theta)_{(m)}} \\
&= \sum_{n=0}^{\infty} \lambda_{n+1} d_{n+1}^\theta(t) \left\{ \frac{n_{[k]}}{(n+\theta)_{(m)}} - \frac{(n+1)_{[k]}}{(n+1+\theta)_{(m)}} \right\} \\
&= \sum_{n=0}^{\infty} d_n^\theta(t) \frac{n_{[k]}}{(n+\theta)_{(m)}} \\
&\quad \times \left\{ \frac{1}{2}(n+m-1+\theta)(m-k) - \lambda_m \right\},
\end{aligned}
\tag{3.15}
$$

where the interchange of summation and differentiation is justified by Lemma 2.5(iii); the first equality uses Lemma 2.5(iv), and the rest is algebra. The coefficient of (3.14) on the right-hand side of (3.9) is

$$\begin{aligned}
&\sum_{k'=1}^{k} \sum_{1\le i<j\le m:\, i,j\in\beta_{k'}} |\beta_{k'}|^{-1} \sum_{n=0}^{\infty} d_n^{\theta}(t) \frac{n_{[k]}}{(n+\theta)_{(m-1)}}\\
&\quad + \sum_{l'=1}^{l} \sum_{1\le i<j\le m:\, i,j\in\gamma_{l'}} (|\gamma_{l'}|-1)^{-1} \sum_{n=0}^{\infty} d_n^{\theta}(t) \frac{n_{[k]}}{(n+\theta)_{(m-1)}}\\
&\quad + \tfrac{1}{2}\theta \sum_{l'=1}^{l} \sum_{1\le i\le m:\, \gamma_{l'}=\{i\}} \theta^{-1} \sum_{n=0}^{\infty} d_n^{\theta}(t) \frac{n_{[k]}}{(n+\theta)_{(m-1)}}\\
&\quad - \lambda_m \sum_{n=0}^{\infty} d_n^{\theta}(t) \frac{n_{[k]}}{(n+\theta)_{(m)}}
\end{aligned}$$

$$\begin{aligned}
(3.16)\qquad &= \sum_{n=0}^{\infty} d_n^{\theta}(t) \frac{n_{[k]}}{(n+\theta)_{(m)}} \Biggl\{ (n+m-1+\theta) \Biggl[\sum_{k':\, |\beta_{k'}|\ge 2} \binom{|\beta_{k'}|}{2} |\beta_{k'}|^{-1}\\
&\qquad + \sum_{l':\, |\gamma_{l'}|\ge 2} \binom{|\gamma_{l'}|}{2} (|\gamma_{l'}|-1)^{-1} + \sum_{l':\, |\gamma_{l'}|=1} \frac{1}{2} \Biggr] - \lambda_m \Biggr\}\\
&= \sum_{n=0}^{\infty} d_n^{\theta}(t) \frac{n_{[k]}}{(n+\theta)_{(m)}} \Biggl\{ \frac{1}{2}(n+m-1+\theta)\\
&\qquad \times \Biggl[\sum_{k'=1}^{k} (|\beta_{k'}|-1) + \sum_{l'=1}^{l} |\gamma_{l'}| \Biggr] - \lambda_m \Biggr\}\\
&= \sum_{n=0}^{\infty} d_n^{\theta}(t) \frac{n_{[k]}}{(n+\theta)_{(m)}} \Biggl\{ \frac{1}{2}(n+m-1+\theta)(m-k) - \lambda_m \Biggr\}.
\end{aligned}$$

This proves (3.9) and completes the proof. □

PROOF OF COROLLARY 1.2. Observe that if Π_1 and Π_2 are finite positive Borel measures on $\mathscr{P}(S)$, then

$$\begin{aligned}
\|\Pi_1 - \Pi_2\|_{\mathrm{var}} &\equiv \sup_{\Gamma\in\mathscr{B}(\mathscr{P}(S))} |\Pi_1(\Gamma) - \Pi_2(\Gamma)|\\
&\le \max\{\Pi_1(\mathscr{P}(S)), \Pi_2(\mathscr{P}(S))\},
\end{aligned}\qquad (3.17)$$

and equality holds if Π_1 and Π_2 are mutually singular. This implies (1.17).

Now if μ and ν_0 are mutually singular, there exists $\Lambda \in \mathscr{B}(S)$ such that $\mu(\Lambda) = 1$ and $\nu_0(\Lambda) = 0$. Letting $\Gamma = \{\nu \in \mathscr{P}(S): \nu(\Lambda) = 0\}$, we conclude from (1.7) that $\Pi_{\theta,\nu_0}(\Gamma) = 1$ and, if $n \ge 1$ and $x_1, \dots, x_n \in \Lambda$, $\Pi_{n+\theta,(n+\theta)^{-1}\{n\eta_n(x_1,\dots,x_n)+\theta\nu_0\}}(\Gamma) = 0$. Thus, the measures Π_1 and Π_2 to which we apply (3.17) are mutually singular. □

PROOF OF COROLLARY 1.3. For the first assertion, it is enough to show, for fixed $n \geq 1$, that

$$(3.18)\quad \int_{S^n} \mu^n(dx_1 \times \cdots \times dx_n)\Pi_{n+\theta,(n+\theta)^{-1}\{n\eta_n(x_1,\ldots,x_n)+\theta\nu_0\}}(\Phi^{-1}(\cdot))$$

depends on $\mu \in \mathscr{P}(S)$ only through $\Phi(\mu)$. By Lemmas 2.1 and 2.2, (3.18) is equal to

$$(3.19)\quad \int_{S^n} \mu^n(dx_1 \times \cdots \times dx_n)\mathbf{P}\left\{\Phi\left(\varepsilon \sum_{i=1}^{n} U_i\delta_{x_i} + (1-\varepsilon)\sum_{i=1}^{\infty} \rho_i\delta_{\xi_i}\right) \in \cdot\right\},$$

where ε is beta(n, θ) distributed, $(U_1, \ldots, U_n)$ is Δ_n-valued Dirichlet $(1, \ldots, 1)$, $(\rho_1, \rho_2, \ldots)$ is Poisson–Dirichlet(θ), $\xi_1, \xi_2, \ldots$ are i.i.d. ν_0, and ε, $(U_1, \ldots, U_n)$, $(\rho_1, \rho_2, \ldots)$, and $(\xi_1, \xi_2, \ldots)$ are independent. Since ν_0 is nonatomic, the probability in (3.19) depends on $(x_1, \ldots, x_n)$ only through the partition of $\{1, \ldots, n\}$ induced by $(x_1, \ldots, x_n)$ (i.e., the partition of $\{1, \ldots, n\}$ for which i and j belong to the same subset if and only if $x_i = x_j$), and the $\mu^n(dx_1 \times \cdots \times dx_n)$-distribution of the partition of $\{1, \ldots, n\}$ induced by $(x_1, \ldots, x_n)$ depends on μ only through $\Phi(\mu)$.

As for (1.19), it suffices to show that (3.18) with $n = 1$ coincides with $\Pi_{\theta,\nu_0}(\Phi^{-1}(\cdot))$. By (3.19), this is equivalent to the following assertion: If ε_0 is beta$(1, \theta)$ distributed and $(\rho_1, \rho_2, \ldots)$ is Poisson–Dirichlet(θ) and is independent of ε_0, then the descending order statistics of $\varepsilon_0, (1-\varepsilon_0)\rho_1, (1-\varepsilon_0)\rho_2, \ldots$ also have the Poisson–Dirichlet(θ) distribution. But the latter is an immediate consequence of the fact that $\rho_1, \rho_2, \ldots$ are distributed as the descending order statistics of $\varepsilon_1, (1-\varepsilon_1)\varepsilon_2, (1-\varepsilon_1)(1-\varepsilon_2)\varepsilon_3, \ldots,$ where $\varepsilon_1, \varepsilon_2, \ldots$ are i.i.d. beta$(1, \theta)$ [see, e.g., Donnelly and Joyce (1989)]. □

PROOF OF COROLLARY 1.4. For each $x \in S$, let $\Gamma_x = \{\mu \in \mathscr{P}(S): \mu(\{x\}) > 0\}$. By (1.7), $\Pi_{\theta,\nu_0}(\Gamma_x) = 1$ if $x \in S$ and $\nu_0 \in \Gamma_x$. Consequently, Theorem 1.1 implies that $P(t, \delta_x, \Gamma_x) \geq 1 - d_0^\theta(t) > 0$ for all $t > 0$ and $x \in S$. Thus, if for some $t > 0$ there were a σ-finite positive Borel measure Π on $\mathscr{P}(S)$ such that $P(t, \mu, \cdot\,) \ll \Pi(\cdot)$ for all $\mu \in \mathscr{P}(S)$, it would necessarily be the case that $\Pi(\Gamma_x) > 0$ for all $x \in S$.

But we claim that, if S is uncountable, there does not exist a σ-finite positive Borel measure Π on $\mathscr{P}(S)$ such that $\Pi(\Gamma_x) > 0$ for all $x \in S$. Suppose not, that is, suppose that such a Π exists. By the uncountability of S and the σ-finiteness of Π, there exists a *finite* positive Borel measure on $\mathscr{P}(S)$, also denoted by Π, such that $\Pi(\Gamma_x) > 0$ for uncountably many $x \in S$. It follows that there exist $\varepsilon > 0$ and distinct $x_1, x_2, \ldots \in S$ such that $\Pi\{\mu \in \mathscr{P}(S): \mu(\{x_n\}) \geq \varepsilon\} \geq \varepsilon$ for each $n \geq 1$. But this implies that $\Pi\{\mu \in \mathscr{P}(S): \mu(\{x_n\}) \geq \varepsilon$ for infinitely many $n \geq 1\} \geq \varepsilon$, a contradiction. □

PROOF OF COROLLARY 1.5. Temporarily denote $\mathscr{L}$ [defined by (1.3) and (1.6)] by $\mathscr{L}_\theta$, denote $\{\mathscr{T}(t)\}$ [defined by (3.1)] by $\{\mathscr{T}_\theta(t)\}$, and denote $P(t, \mu, \cdot\,)$ [given by (1.14)] by $P_\theta(t, \mu, \cdot\,)$. As $\theta \to 0$, $\mathscr{L}_\theta\varphi \to \mathscr{L}_0\varphi$ for all $\varphi \in \mathscr{D}(\mathscr{L})$, hence $\mathscr{T}_\theta(t)\varphi \to \mathscr{T}_0(t)\varphi$ for all $\varphi \in C(\mathscr{P}(S))$ and $t \geq 0$ by Trotter's semigroup approximation theorem, and therefore $P_\theta(t, \mu, \cdot\,) \Rightarrow P_0(t, \mu, \cdot\,)$ for each $t > 0$ and $\mu \in \mathscr{P}(S)$. Consequently, the corollary will follow from Theorem 1.1, provided the map Ξ: $(0, \infty) \times \mathscr{P}(S) \mapsto \mathscr{P}(\mathscr{P}(S))$ defined by $\Xi(\theta, \nu_0) = \Pi_{\theta, \nu_0}$ is continuous. But this is immediate from Lemma 2.4. □

Proof of Corollary 1.6. Noting that $\Pi_{1, \delta_x} = \delta_{\delta_x}$ for each $x \in S$, the proof is similar to that of Corollary 1.2. □

Proof of Corollary 1.7. The proof of the first assertion is similar to the proof of the corresponding result in Corollary 1.3, and the second assertion is immediate from Corollary 1.6. □

Proof of Corollary 1.8. This follows from Corollary 1.5 and Lemma 2.2. □

Proof of Corollary 1.9. (i) Let $S = \{1, \dots, K\}$ and choose $\nu_0 \in \mathscr{P}(S)$ so that $\theta\nu_0 = \theta_1\delta_1 + \cdots + \theta_K\delta_K$. Given $p \in \Delta_K$, define $\mu \in \mathscr{P}(S)$ by $\mu = p_1\delta_1 + \cdots + p_K\delta_K$, and observe that, for each $n \geq 1$,

$$\begin{aligned} &\int_{S^n} \mu^n(dx_1 \times \cdots \times dx_n)\Pi_{n+\theta, (n+\theta)^{-1}\{n\eta_n(x_1, \dots, x_n) + \theta\nu_0\}}(\cdot\,) \\ &\quad = \sum_{\alpha \in (\mathbf{Z}_+)^K:\, |\alpha| = n} \binom{n}{\alpha} \prod_{i=1}^{K} p_i^{\alpha_i} \Pi_{n+\theta, (n+\theta)^{-1}\{(\alpha_1+\theta_1)\delta_1 + \cdots + (\alpha_K+\theta_K)\delta_K\}}(\cdot\,). \end{aligned} \tag{3.20}$$

Thus, the result follows from Theorem 1.1, Corollary 1.5 and Lemma 2.2.

(ii) Put $S = \mathbf{N} \cup \{\infty\}$ and proceed as above. □

Proof of Corollary 1.10. By Theorem 1.1, it is enough to show, for fixed $n \geq 1$, that

$$\begin{aligned} &\Pi_{\theta, \nu_0}(d\mu) \int_{S^n} \mu^n(dx_1 \times \cdots \times dx_n)\Pi_{n+\theta, (n+\theta)^{-1}\{n\eta_n(x_1, \dots, x_n) + \theta\nu_0\}}(d\nu) \\ &\quad = \int_{\mathscr{P}(S)} \Pi_{\theta, \nu_0}(d\lambda) \int_{S^n} \lambda^n(dx_1 \times \cdots \times dx_n) \\ &\qquad \times \big\{\Pi_{n+\theta, (n+\theta)^{-1}\{n\eta_n(x_1, \dots, x_n) + \theta\nu_0\}}(d\mu) \\ &\qquad\qquad \Pi_{n+\theta, (n+\theta)^{-1}\{n\eta_n(x_1, \dots, x_n) + \theta\nu_0\}}(d\nu)\big\}. \end{aligned} \tag{3.21}$$

For this it suffices to show that the integrals of $\langle f_1, \mu\rangle \cdots \langle f_m, \mu\rangle \cdot \langle g_1, \nu\rangle \cdots \langle g_l, \nu\rangle$ with respect to these measures are equal, whenever $m, l \geq 1$ and $f_1, \dots, f_m, g_1, \dots, g_l$ are simple functions on S. Thus, we need only show that, if $K \geq 2$ and $\Lambda_1, \dots, \Lambda_K$ is a partition of S into Borel sets, and if Ψ: $\mathscr{P}(S) \mapsto \Delta_K$ is defined by $\Psi(\nu) = (\nu(\Lambda_1), \dots, \nu(\Lambda_K))$, then the two measures in (3.21) give the same mass to $\Psi^{-1}(B) \times \Psi^{-1}(C)$, where B and C are arbitrary Borel subsets of Δ_K. But with $\Theta = (\theta\nu_0(\Lambda_1), \dots, \theta\nu_0(\Lambda_K))$, the latter assertion is equivalent to

$$\begin{aligned}
&\sum_{\alpha\in(\mathbf{Z}_+)^K:\ |\alpha|=n} \binom{n}{\alpha}\mathbf{E}\left[\prod_{i=1}^{K} V_i^{\alpha_i} I_B(V_1,\dots,V_K)\right]\mathrm{Dirichlet}(\alpha+\Theta)(C)\\
&\quad=\sum_{\alpha\in(\mathbf{Z}_+)^K:\ |\alpha|=n} \binom{n}{\alpha}\mathbf{E}\left[\prod_{i=1}^{K} V_i^{\alpha_i}\right]\\
&\qquad\times \mathrm{Dirichlet}(\alpha+\Theta)(B)\mathrm{Dirichlet}(\alpha+\Theta)(C),
\end{aligned}\tag{3.22}$$

where $(V_1,\dots,V_K)$ has distribution Dirichlet(Θ), and this is easily seen to hold. □

Acknowledgments. The research for this paper was carried out during the first author's visit to the Department of Mathematics at Monash University. He would like to thank his hosts for their hospitality.

REFERENCES

COX, J. T. (1989). Coalescing random walks and voter model consensus times on the torus in $\mathbf{Z}^d$. *Ann. Probab.* **17** 1333–1366.

COX, J. T. and GRIFFEATH, D. (1990). Mean field asymptotics for the planar stepping stone model. *Proc. London Math. Soc.* **61** 189–208.

CROW, J. F. and KIMURA, M. (1956). Some genetic problems in natural populations. *Proc. Third Berkeley Symp. Math. Statist. Probab.* **4** 1–22. Univ. California Press, Berkley.

DAWSON, D. A. and HOCHBERG, K. J. (1982). Wandering random measures in the Fleming–Viot model. *Ann. Probab.* **10** 554–580.

DONNELLY, P. and JOYCE, P. (1989). Continuity and weak convergence of ranked and size-biased permutations on the infinite simplex. *Stochastic Process. Appl.* **31** 89–103.

DONNELLY, P. and TAVARÉ, S. (1987). The population genealogy of the infinitely-many neutral alleles model. *J. Math. Biol.* **25** 381–391.

DYNKIN, E. B. (1989). Three classes of infinite-dimensional diffusions. *J. Funct. Anal.* **86** 75–110.

ETHIER, S. N. (1981). A class of infinite-dimensional diffusions occurring in population genetics. *Indiana Univ. Math. J.* **30** 925–935.

ETHIER, S. N. (1990). The infinitely-many-neutral-alleles diffusion model with ages. *Adv. in Appl. Probab.* **22** 1–24.

ETHIER, S. N. (1992). Eigenstructure of the infinitely-many-neutral-alleles diffusion model. *J. Appl. Probab.* **29** 487–498.

ETHIER, S. N. and GRIFFITHS, R. C. (1993). The transition function of a measure-valued branching diffusion with immigration. In *Stochastic Processes: A Festschrift for Gopinath Kallianpur* (S. Cambanis, J. K. Ghosh, R. L. Karandikar and P. K. Sen, eds.) 71–79. Springer, New York.

ETHIER, S. N. and KURTZ, T. G. (1981). The infinitely-many-neutral-alleles diffusion model. *Adv. in Appl. Probab.* **13** 429–452.

ETHIER, S. N. and KURTZ, T. G. (1986). *Markov Processes: Characterization and Convergence.* Wiley, New York.

ETHIER, S. and KURTZ, T. G. (1987). The infinitely-many-alleles model with selection as a measure-valued diffusion. In *Stochastic Models in Biology* (M. Kimura, G. Kallianpur, and T. Hida, eds.) **70** 72–86. *Lecture Notes in Biomath.* Springer, Berlin.

ETHIER, S. N. and KURTZ, T. G. (1993a). Fleming–Viot processes in population genetics. *SIAM J. Control Optim.* To appear.

ETHIER, S. N. and KURTZ, T. G. (1993b). Convergence to Fleming–Viot processes in the weak atomic topology. *Stochastic Process. Appl.* To appear.

ETHIER, S. N. and KURTZ, T. G. (1993c). Coupling and ergodic theorems for Fleming–Viot processes. Preprint.

EWENS, W. J. and KIRBY, K. (1975). The eigenvalues of the neutral alleles process. *Theoret. Population Biol.* **7** 212–220.

FLEMING, W. H. and VIOT, M. (1979). Some measure-valued Markov processes in population genetics theory. *Indiana Univ. Math. J.* **28** 817–843.

GOLDBERG, S. (1950). On a singular diffusion equation. Ph.D. dissertation, Cornell Univ.

GRIFFITHS, R. C. (1979). A transition density expansion for a multi-allele diffusion model. *Adv. in Appl. Probab.* **11** 310–325.

GRIFFITHS, R. C. (1980). Lines of descent in the diffusion approximation of neutral Wright–Fisher models. *Theoret. Population Biol.* **17** 37–50.

GRIFFITHS, R. C. and LI, W.-H. (1983). Simulating allele frequencies in a population and the genetic differentiation of populations under mutation pressure. *Theoret. Population Biol.* **32** 19–33.

KIMURA, M. (1955). Solution of a process of random genetic drift with a continuous model. *Proc. Nat. Acad. Sci. U.S.A.* **41** 144–150.

KIMURA, M. (1956). Random genetic drift in a tri-allelic locus; exact solution with a continuous model. *Biometrics* **12** 57–66.

KINGMAN, J. F. C. (1975). Random discrete distributions. *J. Roy. Statist. Soc. Ser. B* **37** 1–22.

KINGMAN, J. F. C. (1982). The coalescent. *Stochastic Process. Appl.* **13** 235–248.

LITTLER, R. A. and FACKERELL, E. D. (1975). Transition densities for neutral multi-allele diffusion models. *Biometrics* **31** 117–123.

MALÉCOT, G. M. (1948). *Les Mathématiques de l'Hérédité.* Masson et Cie, Paris.

SHIGA, T. (1990). A stochastic equation based on a Poisson system for a class of measure-valued diffusion processes. *J. Math. Kyoto Univ.* **30** 245–279.

SHIMAKURA, N. (1977). Equations différentielles provenant de la génétique des populations. *Tôhoku Math. J.* **29** 287–318.

SHIMAKURA, N. (1981). Formulas for diffusion approximations of some gene frequency models. *J. Math. Kyoto Univ.* **21** 19–45.

TAVARÉ, S. (1984). Line-of-descent and genealogical processes, and their applications in population genetics models. *Theoret. Population Biol.* **26** 119–164.

WRIGHT, S. (1949). Adaptation and selection. In *Genetics, Paleontology, and Evolution* (G. L. Jepson, E. Mayr, and G. G. Simpson, eds.) 365–389. Princeton Univ. Press.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF UTAH
SALT LAKE CITY, UTAH 84112

DEPARTMENT OF MATHEMATICS
MONASH UNIVERSITY
CLAYTON, VICTORIA 3168
AUSTRALIA